



PRODUCT DOCUMENTATION

Pivotal Cloud Foundry[®]

Version v2.6

Published: 21 June 2019

© 2019 Pivotal Software, Inc. All Rights Reserved.

Table of Contents

Table of Contents	2
PCF v2.6 Feature Highlights	11
Pivotal Cloud Foundry Release Notes	13
PCF v2.6 Breaking Changes	14
PCF Ops Manager v2.6 Release Notes	16
Pivotal Application Service v2.6 Release Notes	21
PAS for Windows v2.6 Release Notes	26
PKS Release Notes	29
Enterprise PKS Release Notes	45
PCF Isolation Segment v2.6 Release Notes	50
Stemcell Release Notes	53
Stemcell (Linux) Release Notes	55
Stemcell v2019.x (Windows Server version 2019) Release Notes	95
Tiles Using Xenial Stemcells in PCF	96
Platform Architecture and Planning	97
AWS Reference Architecture	103
Azure Reference Architecture	107
GCP Reference Architecture	114
OpenStack Reference Architecture	121
vSphere Reference Architecture	124
Using Edge Services Gateway on VMware NSX	135
Upgrading vSphere without PCF Downtime	148
Migrating PCF to a New Datastore in vSphere	150
Control Plane Reference Architectures	152
Implementing a Multi-Foundation PKS Deployment	160
Global DNS Load Balancers for Multi-Foundation Environments	162
Architecture and Installation Overview	164
Preparing Your Firewall	167
IaaS Permissions Guidelines	169
Installing PCF in Airgapped Environments	170
Installing Pivotal Cloud Foundry on AWS	174
AWS Permissions Guidelines	178
Installing PCF on AWS Manually	179
Configuring BOSH Director on AWS	202
Installing PCF on AWS using Terraform	221
Deploying Ops Manager on AWS Using Terraform	222
Configuring BOSH Director on AWS Using Terraform	226
Deleting PCF from AWS	245
Creating a Proxy ELB for Diego SSH	249
Upgrading BOSH Director on AWS	252
Installing Pivotal Cloud Foundry on Azure	256
Installing PCF on Azure Manually	259
Preparing to Deploy Ops Manager on Azure Manually	260
Deploying Ops Manager on Azure Manually	265
Configuring BOSH Director on Azure Manually	274
Installing PCF on Azure Using Terraform	294
Deploying Ops Manager on Azure Using Terraform	295
Configuring BOSH Director on Azure Using Terraform	298

Troubleshooting PCF on Azure	318
Deleting a PCF on Azure Installation	321
Upgrading BOSH Director on Azure	322
Installing Pivotal Cloud Foundry on GCP	327
Recommended GCP Quotas	330
Installing PCF on GCP Manually	331
Preparing to Deploy Ops Manager on GCP Manually	332
Deploying Ops Manager on GCP Manually	333
Configuring BOSH Director on GCP Manually	336
Installing PCF on GCP Using Terraform	372
Deploying Ops Manager on GCP Using Terraform	373
Configuring BOSH Director on GCP Using Terraform	378
Configuring a Shared VPC on GCP	392
Deleting PCF from GCP	396
Troubleshooting PCF on GCP	398
Upgrading BOSH Director on GCP	400
Installing Pivotal Cloud Foundry on OpenStack	404
Deploying Ops Manager to OpenStack	407
Configuring BOSH Director on OpenStack	417
Installing Pivotal Cloud Foundry on vSphere	437
vSphere Service Account Requirements	439
PAS on vSphere Requirements	443
Deploying Ops Manager on vSphere	445
Configuring BOSH Director on vSphere	451
Deploying PAS with NSX-T Networking	470
vSphere Virtual Disk Types	488
Using the Cisco Nexus 1000v Switch with Ops Manager	490
Using BOSH Resurrector and vSphere HA	493
Configuring Pivotal Cloud Foundry SSL Termination for vSphere Deployments	496
Availability Zones in vSphere	498
Updating NSX-V Security Group and Load Balancer Information	500
Updating NSX-T Load Balancer Server Pool Membership	502
Configuring PAS	505
Quick Start PAS Configuration	537
Configuring File Storage for PAS	539
Configuring Load Balancing for PAS	552
Installing PCF Isolation Segment	559
Getting Started with Small Footprint PAS	572
Upgrading Pivotal Cloud Foundry	579
What Happens During PAS Upgrades	584
Upgrade Preparation Checklist for PCF v2.6	587
Configuring PAS for Upgrades	595
Upgrade Load Example: Pivotal Web Services	599
Upgrading PAS and Other Pivotal Cloud Foundry Products	601
cf push Availability During Pivotal Application Service Upgrades	604
Backing Up and Restoring Pivotal Cloud Foundry	605
Incident Response Guides	606
Restoring Lost BOSH Director Persistent Disk	607
Restoring Lost PAS Data	609
Recovering from an Ops Manager and PAS Upgrade Failure	612

Disaster Recovery in Pivotal Cloud Foundry	614
Backing Up Pivotal Cloud Foundry with BBR	619
PAS Component Availability During Backup	634
Restoring PCF from Backup with BBR	637
Enabling External Blobstore Backups	651
Troubleshooting BBR	652
Using Ops Manager	655
Using the Ops Manager UI	657
Using the Ops Manager Interface	658
Reviewing Pending Product Changes	666
Adding and Deleting Products	669
Importing and Managing Stemcells	673
Logging In to Apps Manager	675
Managing Errands in Ops Manager	676
Using Ops Manager Programmatically and from the Command Line	679
Using the Ops Manager API	681
Creating UAA Clients for BOSH Director	688
Using Your Own Load Balancer	690
Creating and Managing Ops Manager User and Client Accounts	692
Configuring Role-Based Access Control (RBAC) in Ops Manager	697
Adding Existing SAML or LDAP Users to a PCF Deployment	702
Modifying Your Ops Manager Installation and Product Template Files	704
Monitoring PCF VMs from Ops Manager	707
Advanced Features in Ops Manager	709
Managing Ops Manager Verifiers	710
Managing Custom VM Extensions	714
Ops Manager Fields That Lock On Deploy	717
Property and Template References	719
PAS Concepts	779
Cloud Foundry Overview	780
Security	783
Cloud Foundry Security	784
Container Security	789
Orgs, Spaces, Roles, and Permissions	793
Application Security Groups	796
Application SSH Components and Processes	803
High Availability	804
High Availability in Cloud Foundry	805
How Cloud Foundry Maintains High Availability	808
Running Apps in the Cloud	809
How Apps Are Staged	810
App Container Lifecycle	813
How Diego Balances App Processes	815
Runtime Components	818
Diego Components and Architecture	821
Cloud Foundry Routing Architecture	827
Service Mesh Architecture	830
Cloud Controller	832
Cloud Controller Blobstore	834
Messaging (NATS)	836

Gorouter	838
User Account and Authentication (UAA) Server	841
Garden	851
GrootFS Disk Usage	853
HTTP Routing	855
Container-to-Container Networking	861
PCF Operator Guide	865
Deploying Service Mesh (Beta)	866
Identifying the API Endpoint for Your PAS Instance	868
Creating New PAS User Accounts	869
Configuring SSL/TLS Termination at HAProxy	870
Configuring Frontend Idle Timeout for Gorouter and HAProxy	872
Configuring Route Service Lookup	873
Configuring Proxy Settings for All Applications	877
Restricting App Access to Internal PCF Components	879
Configuring Application Security Groups for Email Notifications	883
Configuring SSH Access for PCF	885
Securing Services Instance Credentials with Runtime CredHub	887
Identifying PAS Jobs Using vCenter	890
Configuring Logging in PAS	892
Configuring UAA Password Policy	898
Managing Internal MySQL for PAS	900
Scaling Internal MySQL for PAS	901
Running mysql-diag	904
Recovering From MySQL Cluster Downtime	907
Configuring Authentication and Enterprise SSO for PAS	914
Configuring ADFS as an Identity Provider	920
Configuring CA as an Identity Provider	922
Configuring PingFederate as an Identity Provider	925
Switching Application Domains	926
Scaling PAS	928
Using Docker Registries	932
Configuring Cell Disk Cleanup Scheduling	936
Custom Branding Apps Manager	938
Reporting App, Task, and Service Instance Usage	943
Reporting Instance Usage with Apps Manager	950
Providing a Certificate for Your TLS Termination Point	952
Enabling Volume Services	954
Rotating Runtime CredHub Encryption Keys	957
Administrating PAS	959
Managing the Runtime	960
Stopping and Starting Virtual Machines	961
Creating and Modifying Quota Plans	964
Using Feature Flags	967
Examining GrootFS Disk Usage	969
Using Metadata	974
Managing Custom Buildpacks	980
Using Docker in Cloud Foundry	982
User Accounts and Communications	984
Creating and Managing Users with the cf CLI	985

Creating and Managing Users with the UAA CLI (UAAC)	987
Getting Started with the Notifications Service	993
Routing	996
Enabling IPv6 for Hosted Applications	997
Supporting WebSockets	998
Configuring Load Balancer Healthchecks for Cloud Foundry Routers	1000
Securing Traffic into Cloud Foundry	1002
Enabling TCP Routing	1011
Service Mesh (Beta)	1014
Isolation Segments	1016
Managing Isolation Segments	1017
Routing for Isolation Segments	1021
Monitoring PAS	1027
Key Performance Indicators	1028
Key Capacity Scaling Indicators	1052
Selecting and Configuring a Monitoring System	1060
Pivotal Application Service for Windows	1062
Product Architecture	1064
Downloading or Creating Windows Stemcells	1066
Creating a Windows Stemcell for vSphere Manually	1067
Creating a Windows Stemcell for vSphere Using stembuild (Beta)	1076
Installing and Configuring PASW	1083
Windows Cells in Isolation Segments	1088
Migrating Apps to PAS for Windows	1090
Upgrading PAS for Windows and Windows Cells	1093
Troubleshooting Windows Cells	1094
Deploying .NET Apps	1100
Using SMB Volumes in .NET Apps	1101
Changing Windows Stacks	1105
Using Apps Manager	1107
Getting Started with Apps Manager	1108
Managing Orgs and Spaces Using Apps Manager	1110
Managing User Roles with Apps Manager	1115
Managing Apps and Service Instances Using Apps Manager	1118
Scaling an App Using App Autoscaler	1139
Using the App Autoscaler CLI	1145
Viewing ASGs in Apps Manager	1150
Configuring Spring Boot Actuator Endpoints for Apps Manager	1152
Using Spring Boot Actuators with Apps Manager	1155
Configuring Multi-Foundation Support in Apps Manager	1163
Using the Cloud Foundry Command Line Interface (cf CLI)	1167
Installing the cf CLI	1168
Getting Started with the cf CLI	1172
Using the cf CLI with a Proxy Server	1177
Using the cf CLI with a Self-Signed Certificate	1180
Using cf CLI Plugins	1181
Developing cf CLI Plugins	1184
Cloud Foundry CLI Reference Guide	1185
Using Experimental cf CLI Commands	1193
Developer Guide	1195

cf push	1196
Pushing an App	1197
Deploying with App Manifests	1202
App Manifest Attribute Reference	1204
Deploying an App with Docker	1213
Deploying a Large App	1217
Starting, Restarting, and Restaging Apps	1219
Pushing an App with Multiple Processes (Beta)	1221
Running cf push Sub-Step Commands (Beta)	1223
Rolling App Deployments (Beta)	1225
Pushing Apps with Sidecar Processes (Beta)	1228
Using Blue-Green Deployment to Reduce Downtime and Risk	1233
Troubleshooting app Deployment and Health	1236
SSH for Apps and Services	1241
App SSH Overview	1242
Accessing Apps with SSH	1244
Accessing Services with SSH	1249
Routes and Domains	1251
Configuring Routes and Domains	1252
Detailed documentation to help you install, understand, and succeed with Pivotal's enterprise-grade software.	1263
owner: Routing	1263
Overview	1263
Configure Routing Weights	1263
Confirm Routing Weights	1266
Configuring CF to Route Traffic to Apps on Custom Ports	1267
Managing Services	1269
Managing Service Instances with the cf CLI	1270
Sharing Service Instances	1275
Delivering Service Credentials to an App	1277
Managing Service Keys	1280
Configuring Play Framework Service Connections	1282
Using an External File System (Volume Services)	1283
User-Provided Service Instances	1289
Streaming App Logs	1291
Streaming App Logs to Log Management Services	1292
Service-Specific Instructions for Streaming App Logs	1295
Streaming App Logs to Splunk	1302
Streaming App Logs with Fluentd	1304
Streaming App Logs to Azure OMS Log Analytics (Beta)	1305
Managing Apps with the cf CLI	1309
Running Tasks	1310
Scaling an App Using cf scale	1313
Using Application Health Checks	1314
Configuring Container-to-Container Networking	1317
Administering Dynamic Egress Policies (Beta)	1320
Cloud Foundry Environment Variables	1326
Cloud Controller API Client Libraries	1334
Considerations for Designing and Running an App in the Cloud	1335
App Revisions	1338

PCF Security and Compliance Guide	1341
Security Processes	1342
Pivotal Cloud Foundry Security Overview and Policy	1343
PCF Testing, Release, and Security Lifecycle	1345
Security Concepts	1348
PCF Infrastructure Security	1349
Stemcell Security	1350
Linux Stemcell Hardening	1351
Certificates and TLS in PCF	1355
TLS Connections in PCF	1356
Custom Certificate Authorities	1363
Certificate Types	1364
Managing Certificates with the Ops Manager API	1365
Rotating Certificates	1367
Trusted System Certificates	1375
Disk Encryption	1376
Security Event Logging	1379
Network Security	1385
BOSH DNS Network Communications	1386
Cloud Controller Network Communications	1389
Container-to-Container Networking Communications	1391
CredHub Network Communications	1392
Diego Network Communications	1393
Loggregator Network Communications	1395
MySQL Network Communications	1397
NATS Network Communications	1399
Routing Network Communications	1400
UAA Network Communications	1402
Credential and Identity Management	1403
Pivotal Cloud Foundry User Types	1405
Retrieving Credentials from Your Deployment	1408
CredHub	1414
CredHub Credential Types	1417
Security for Apps and Services	1418
Compliance and Other Security-Related Topics	1419
General Data Protection Regulation	1420
Security-Related PCF Tiles and Add-ons	1425
Security Guidelines for Your IaaS Provider	1426
How to Use This Topic	1426
Buildpacks	1427
Using Buildpacks	1428
How Buildpacks Work in Cloud Foundry	1429
Stack Association	1434
Pushing an Application with Multiple Buildpacks	1436
Using a Proxy Server	1437
Supported Binary Dependencies	1438
Production Server Configuration	1439
System Buildpacks	1441
Binary Buildpack	1442
Go Buildpack	1444

HWC Buildpack	1448
Creating an Extension Buildpack for .NET Apps	1452
Tips for .NET Framework Developers	1455
Java Buildpack	1459
Tips for Java Developers	1461
Getting Started Deploying Java Apps	1468
Getting Started Deploying Grails Apps	1469
Getting Started Deploying Ratpack Apps	1476
Getting Started Deploying Spring Apps	1482
Configuring Service Connections	1488
Configuring Service Connections for Grails	1489
Configuring Service Connections for Play Framework	1492
Configuring Service Connections for Spring	1493
Cloud Foundry Java Client Library	1500
.NET Core Buildpack	1503
NGINX Buildpack	1509
Node.js Buildpack	1511
Tips for Node.js Applications	1514
Environment Variables Defined by the Node Buildpack	1517
Configuring Service Connections for Node.js	1518
PHP Buildpack	1520
Tips for PHP Developers	1523
Getting Started Deploying PHP Apps	1524
PHP Buildpack Configuration	1527
Composer	1531
Sessions	1533
New Relic	1534
Python Buildpack	1535
R Buildpack	1539
Ruby Buildpack	1541
Tips for Ruby Developers	1544
Getting Started Deploying Ruby Apps	1550
Getting Started Deploying Ruby Apps	1551
Getting Started Deploying Ruby on Rails Apps	1556
Configure Rake Tasks for Deployed Apps	1559
Environment Variables Defined by the Ruby Buildpack	1560
Configure Service Connections for Ruby	1561
Support for Windows Gemfiles	1564
Staticfile Buildpack	1565
Customizing and Developing Buildpacks	1570
Creating Custom Buildpacks	1571
Packaging Dependencies for Offline Buildpacks	1574
Merging from Upstream Buildpacks	1577
Upgrading Dependency Versions	1578
Using CI for Buildpacks	1583
Releasing a New Buildpack Version	1584
Updating Buildpack-Related Gems	1586
Services	1587
Overview	1588
Managing Service Brokers	1590

Managing Access to Service Plans	1594
Dashboard Single Sign-On	1597
Example Service Brokers	1601
Binding Credentials	1602
Enabling Service Instance Sharing	1604
Application Log Streaming	1606
Route Services	1607
Supporting Multiple Cloud Foundry Instances	1612
Logging and Metrics	1613
Overview of Logging and Metrics	1614
Loggregator Architecture	1616
Loggregator Guide for Cloud Foundry Operators	1618
Deploying a Nozzle to the Loggregator Firehose	1620
Installing the Loggregator Firehose Plugin for cf CLI	1626
App Logging in Cloud Foundry	1627
Customizing Platform Log Forwarding	1631
Configuring the Metric Registrar	1633
Registering and Emitting Custom App Metrics	1635
Identifying the Source Deployment of Metrics	1639
Troubleshooting and Diagnostics	1640
Diagnosing Problems in PCF	1641
Troubleshooting Problems in PCF	1646
Advanced Troubleshooting with the BOSH CLI	1650
Troubleshooting Slow Requests in Cloud Foundry	1657
Troubleshooting TCP Routes	1660
Troubleshooting Router Error Responses	1663
Troubleshooting Ops Manager for VMware vSphere	1665

PCF v2.6 Feature Highlights

This topic highlights important new features included in Pivotal Cloud Foundry (PCF) v2.6.

PCF Ops Manager Highlights

Ops Manager v2.6 includes the important major features. For additional information about these and other features included in Ops Manager v2.6, see [PCF Ops Manager v2.6 Release Notes](#).

Ops Manager Supports Multiple Stemcells for Products

Ops Manager supports products that require multiple stemcells. This means it is no longer necessary to deploy separate tiles for products that require different stemcells.

Product Deployment Times from API and Change Log Page

Ops Manager publishes deployment times for individual products and whole deployments through the `/api/v0/installations` endpoint and [Change Log](#) page.

Compiled Release Assets Included in BOSH Deployment Manifest

Ops Manager now automatically includes the `exported_from` field in the `releases` section of the BOSH deployment manifest file. The `exported_from` field contains the compiled release assets for the deployment and prevents BOSH from updating every VM with later versions of packages during a BOSH deployment. This saves time and reduces unexpected VM updates.

API Endpoint Returns Deployment Configuration Details

The `/api/v0/staged/products/:product_guid/pre_deploy_check` Ops Manager API endpoint returns configuration details about a deployment. This allows you to verify the details of your deployment before clicking [Apply Changes](#), and to troubleshoot failed deploys.

AWS Deployments Use 5th Generation Instances

Ops Manager v2.6 uses the latest generation of instances for AWS deployments. When you upgrade to Ops Manager v2.6, your VM types convert to the corresponding 5th generation instance types, such as `t2` to `t3` and `c4` to `c5`.

Tile Developers Can Implement a Form Verifier to Confirm vSphere Properties

An optional form verifier is available that notifies tile operators when they have configured vSphere properties incorrectly.

Pivotal Application Service (PAS) Highlights

PAS v2.6 includes the following important major features. For additional information about these and other features included in PAS v2.6, see [Pivotal Application Service v2.6 Release Notes](#).

Push Apps with Sidecar Processes (Beta)

You can run additional processes, or sidecars, in the same container as your app. Sidecars are useful for processes that depend on each other or must run in the same container. This includes processes that must communicate through localhost or share the same filesystem, such as an Application Performance Monitoring (APM) tool.

Enable, View, and Rollback Revisions for Apps

PAS supports revisions for apps. A revision is an object that represents code and configuration used by an app at a specific time. Some uses cases for revisions include rolling back your app to a previous version and viewing changes in your app over time. You can also add metadata to revisions.

Monitor System Metrics with System Metrics Agent

System Metrics Agent allows you to more easily correlate VM metrics with BOSH jobs. These VM metrics help with troubleshooting and diagnosing issues for potential infrastructure problems.

Loggregator Syslog Agent Increases Scale For Syslog Drains

Syslog Agents run on PCF component VMs and host VMs to manage connections with and write to syslog drains for app logs. The addition of Syslog Agents increases the number of syslog drain service bindings supported by the Loggregator system and reduces the workload for Loggregator VMs.

Pivotal Cloud Foundry Release Notes

Pivotal Cloud Foundry is certified by the Cloud Foundry Foundation for 2019.

Read more about the [certified provider program](#) and the [requirements of providers](#).

This topic provides links to the release notes for Pivotal Cloud Foundry (PCF). Release notes include new features, breaking changes, bug fixes, and known issues.

To find release notes for PCF services, navigate to the documentation for the desired service from [Pivotal Documentation](#).

PCF Release Notes

- [PCF v2.6 Breaking Changes](#)
- [PCF Ops Manager v2.6 Release Notes](#)

PCF Runtime Release Notes

- [Pivotal Application Service v2.6 Release Notes](#)
- [PAS for Windows v2.6 Release Notes](#)
- [PCF Isolation Segment v2.6 Release Notes](#)
- [Pivotal Container Service \(PKS\) v1.2 Release Notes](#)

Stemcell Release Notes

- [Stemcell Release Notes](#)

PCF v2.6 Breaking Changes

This topic describes the breaking changes you need to be aware of when upgrading to Pivotal Cloud Foundry (PCF) v2.6. For more information about important preparation steps you must follow before beginning an upgrade, see [Upgrading Pivotal Cloud Foundry](#).

PCF Ops Manager

See the following PCF Ops Manager breaking changes:

Passwords Not Supported for Ops Manager VM on vSphere

You can only SSH onto the Ops Manager VM in a vSphere deployment with a private SSH key. You cannot SSH onto the Ops Manager VM with a password.

To avoid errors when authenticating, add a public key to the OVF template for the Ops Manager VM. Then, use the private key to SSH onto the Ops Manager VM.

For more information about adding a public key to the OVF template, see [Deploy Ops Manager](#) in *Deploying Ops Manager on vSphere*.

Stemcell Property for Diagnostic Report API Endpoint is Removed

The `stemcell` property for the `/api/v0/diagnostic_report` endpoint is removed. Update any automation that uses the stemcell information in the `/api/v0/diagnostic_report` endpoint to avoid errors.

Information about all stemcells for both staged and deployed products is listed under a new `added_products` property for the `/api/v0/diagnostic_report` endpoint. Stemcell information under the `added_products` property includes the stemcell filename, OS, and version.

For more information about the `/api/v0/diagnostic_report` endpoint, see [Diagnostic Report](#) in the Ops Manager API documentation.

Stemcell Version Key for Pending Changes is Moved and Renamed

For the `/api/v0/staged/pending_changes` endpoint, the `stemcell_version` key from previous versions of the Ops Manager API is now the `version` sub-key under the `stemcells` key.

Update any automation that uses the stemcell information in the `/api/v0/pending_changes` endpoint to avoid errors.

For more information about the `/api/v0/staged/pending_changes` endpoint, see [View a list of pending changes](#) in the Ops Manager API documentation.

AWS Deployments Use 5th Generation Instances

Ops Manager v2.6 uses the latest generation of instances for AWS deployments. When you upgrade to Ops Manager v2.6, your VM types convert to the corresponding 5th generation instance types, such as `t2` to `t3` and `c4` to `c5`.

If you are using a custom instance type catalog, your VM types that are used and available do not change.

These instance types are supported by Xenial stemcells and Trusty stemcell 3586 and later. Stemcells earlier than Trusty stemcell 3586 fail to boot Ops Manager. To avoid this, ensure that you are using Trusty stemcell 3586 or later.

For more information about AWS instance types, see [Amazon EC2 Instance Types](#) in the AWS documentation.

Stemcell Assignments Endpoint Deprecated

The `stemcells_assignments` endpoint in the Ops Manager API has been deprecated. The deprecated endpoint only returns a single stemcell for a product.

Instead, operators should use the `stemcell_associations` endpoint, which returns information about all stemcells associated with a product.

For more information, see [Stemcell Associations for Products](#) in the Ops Manager API documentation.

Pivotal Application Service (PAS)

See the following PAS breaking changes:

HTTP Healthcheck Server Disabled for SSH Proxy

If your SSH load balancer is configured with an HTTP healthcheck, remove the healthcheck before upgrading. This is the load balancer specified in the **Diego Brain** row of the **Resource Config** pane.

 **Note:** Neither the Terraform templates for installing PAS nor the PAS installation instructions specify a HTTP healthcheck for the SSH Load balancer.

The `ssh_proxy` job running on the Diego Brain VM no longer starts a HTTP healthcheck server. This feature was previously added to allow GCP load balancers to healthcheck the service. However, GCP load balancers no longer require an HTTP endpoint for healthchecks.

Additionally, making this healthcheck server inaccessible improves the security of the platform.

Default VM Type Change for Router VMs

PAS v2.6 increases the default and minimum CPU core count from one to two for Router VMs.

Verify that you have sufficient infrastructure resources allocated to the deployment before upgrading to PAS v2.6.

PAS for Windows

See the following PAS for Windows breaking changes:

- Option to override deployment name to CF is removed.
- Windows stemcell v2019.2 is not compatible with Ops Manager v2.6. Before upgrading to Ops Manager v2.6, upgrade your Windows stemcell to v2019.3 or later.

PCF Ops Manager v2.6 Release Notes

Pivotal Cloud Foundry is certified by the Cloud Foundry Foundation for 2019.

Read more about the [certified provider program](#) and the [requirements of providers](#).

How to Upgrade

The [Upgrading Pivotal Cloud Foundry](#) topic contains instructions for upgrading to Pivotal Cloud Foundry (PCF) Ops Manager v2.6.

Releases

Ops Manager v2.6 includes the following minor releases:

2.6.0

Ops Manager v2.6.0 uses the following component versions:

Component	Version
Ops Manager	2.6.0-build.138*
Stemcell	315.34*
BBR SDK	1.15.1
BOSH Director	269.0.2*
BOSH DNS	1.11.0
Metrics Server	0.0.22
CredHub	2.4.0
Syslog	11.4.0
Windows Syslog	1.0.3
UAA	71.0
BPM	1.0.4
Networking	9
OS Conf	20.0.0
AWS CPI	75
Azure CPI	36.0.0*
Google CPI	29.0.1
OpenStack CPI	42
vSphere CPI	52.1.1
BOSH CLI	5.5.1*
Credhub CLI	2.4.0
BBR CLI	1.5.1

** Components marked with an asterisk have been updated.*

New Features in Ops Manager v2.6

Ops Manager v2.6 includes the following major features:

Ops Manager Supports Multiple Stemcells for Products

 **Breaking Changes:** There are breaking changes for API users related to this feature. For more information, see [PCF Ops Manager](#) in PCF 2.6 [Breaking Changes](#).

Ops Manager supports products that require multiple stemcells. This means it is no longer necessary to deploy separate tiles for products that require different stemcells.

The following features and API endpoints enable support for products with multiple stemcells:

- The **Review Pending Changes** page includes the OS and version number of all stemcells associated with each product. The **Review Pending Changes** page also displays a message that indicates when a new stemcell is added to a product.
- The **Stemcell Library** includes a dropdown for each additional required stemcell. You can use this dropdown to add multiple stemcells to each product. The **Stemcell Library** also displays the OS and version number of the stemcells associated with each deployed and staged product.
- You can assign one or more stemcells to a product with the multi-stemcell API, `PATCH /api/v0/stemcells_associations`. `GET /api/v0/stemcells_associations` returns information on all stemcells assigned to products in your deployment.
- The `added_products` property of the `/api/v0/diagnostic_report` endpoint returns the filename, OS, and version number of all stemcells associated with each product in your deployment.
- For products with multiple stemcells, you can view available stemcell updates using the Pivotal Network integration sidebar or the `pivotal_network/stemcell_updates` API endpoint.

For more information about the API endpoints that enable support for products with multiple stemcells, see the [Ops Manager API documentation](#).

For more information about using multiple stemcells as a tile developer, see [Support for Multiple Stemcells](#) in the [PCF v2.6 Partners Release Notes](#).

vSphere Operators Can Use Whitespace in Folder Names

vSphere's vCenter allows you to designate folder names with multiple words separated by spaces. That ability is reflected in Ops Manager. You can specify folder names in vCenter using spaces between words and Ops Manager will accept them as valid.

For an example of folder names with spaces, see [Using the Cisco Nexus 1000v Switch with Ops Manager](#).

Tile Developers Can Implement a Form Verifier to Confirm vSphere Properties

An operator configuring a product tile on a vSphere deployment must configure vSphere properties. Incorrect configuration can cause deployment failures, and can be difficult to troubleshoot effectively. An optional form verifier is available that notifies tile operators when they have configured vSphere properties incorrectly.

For more information about this form verifier, see the `install_time_verifiers` section in *Property and Template References* in the tile developer documentation.

Added Strategies for Backing Up BOSH Blobstore to S3

The BOSH Director tile > **Director Config** pane supports multiple strategies for backing up the BOSH Director blobstore to external S3 buckets:

- **S3 with versioning:** Back up to prior version of blobstore bucket.
- **S3 without versioning:** Back up to S3 bucket dedicated to blobstore backup.

See the *Director Config Page* section of the *Configuring BOSH Director* topic for your IaaS, which you can navigate to from the [IaaS-Specific Deployment Guidelines](#) section of the *Deploying BOSH and Ops Manager* topic.

API Endpoint Returns Deployment Configuration Details

The `/api/v0/staged/products/:product_guid/pre_deploy_check` Ops Manager API endpoint now validates for missing configurations. The endpoint performs a dry run of Apply Changes verifiers and reports any failures that might occur when triggering Apply Changes. This endpoint allows operators to discover which configurations need attention without having to manually search the Ops Manager UI or wait for Apply Changes to fail.

For example, you can check whether a product has not been assigned a stemcell, network, or availability zone. You can also check exactly which properties are invalid and which `install_time` verifiers might fail.

For more information, see [Running a pre-deploy check for the staged BOSH Director](#) in the Ops Manager API documentation.

Operators Can Specify a Monitor Port When Defining an NSX Load Balancer

The Ops Manager API allows you to specify a monitor port when defining an NSX load balancer with the `monitor_port` parameter. This means the monitor port value does not default to the same value as the traffic port.

For more information, see [Configuring resources for a job \(Experimental\)](#) in the Ops Manager API documentation.

Operators Not Required to Specify a Port for NSX Load Balancers

The Ops Manager API no longer requires you to specify a port for NSX load balancers. When a port is not specified, the TCP router relies on a default port that is specified by other components in the system.

This feature helps to support NSX load balancer compatibility with Gorouter, which requires a monitor port of `8080` and a traffic port of `80`.

For more information, see [Configuring resources for a job \(Experimental\)](#) in the Ops Manager API documentation.

Fresh Installations of Ops Manager Create New Default Certificate Authority in CredHub

When installing Ops Manager v2.6, Ops Manager creates a new default certificate authority (CA) in the BOSH Director CredHub. This CA can generate TLS certificates.

When upgrading to Ops Manager v2.6 from Ops Manager v2.5, any CA that already exists does not get overwritten.

For more information about certificates and certificate authorities, see [Certificates and TLS in PCF](#).

Improved Change Log Page

The Ops Manager Change Log page has an improved UI, with deployment status and build times listed individually for each product, as well as for the deployment as a whole.

For more information, see the [Change Log Page](#) section in the *Using the Ops Manager Interface* topic.

Product Deployment Times from API and Change Log Page

Ops Manager publishes deployment times for individual products and whole deployments.

The Ops Manager API returns `started_at` and `finished_at` timestamps from its `/api/v0/installations` endpoint.

The Ops Manager UI [Change Log](#) page publishes these start and end timestamps, and durations calculated from them. For more information about viewing start and end timestamps on the [Change Log](#) page, see [Improved Change Log Page](#).

Ops Manager Syslog Template Available for Service Broker Tiles

Tile developers can use the Ops Manager Syslog template form for their service broker tiles. The template makes tile users' experience more consistent and secure. Tile developers can apply the Syslog template using the `opsmanager_syslog` key in their tile's `metadata.yml` file.

For more information about the Syslog template, see [Syslog Form Template Available for Tile Authors](#).

Ops Manager Adds Support for All AWS Regions

The Ops Manager [AWS Config](#) pane includes an open text field for AWS region. This field allows you to input any AWS region when configuring BOSH Director on AWS.

For more information about adding an AWS region in Ops Manager, see [Step 2: AWS Config Page](#).

AWS Deployments Use 5th Generation Instances

Ops Manager v2.6 uses the latest generation of instances for AWS deployments. When you upgrade to Ops Manager v2.6, your VM types convert to the corresponding 5th generation instance types, such as `i2` to `i3` and `c4` to `c5`.

If you are using a custom instance type catalog, your VM types that are used and available do not change.

These instance types are supported by Xenial stemcells and Trusty stemcell 3586 and later. Stemcells earlier than Trusty stemcell 3586 fail to boot Ops Manager.

For more information about AWS instance types, see [Amazon EC2 Instance Types](#) in the AWS documentation.

Compiled Release Assets Included in BOSH Deployment Manifest

Ops Manager now automatically includes the `exported_from` field in the `releases` section of the BOSH deployment manifest file. The `exported_from` field contains the compiled release assets for the deployment and prevents BOSH from updating every VM with later versions of packages during a BOSH deployment. This saves time and reduces unexpected VM updates.

For example, the `exported_from` field in the following manifest file specifies that the deployment uses the packages from the release compiled against version `250.4` of `ubuntu-xenial`:

```
releases:  
- name: bpm  
  url: https://s3.amazonaws.com/bosh-compiled-release-tarballs/bpm-0.12.3-ubuntu-xenial-250.4.tgz  
  version: 0.12.3  
  exported_from:  
    - os: ubuntu-xenial  
      version: 250.4
```

For more information about using the `exported_from` field, see [Locking Compiled Releases](#) in the BOSH documentation.

ADFS Authentication Available for Microsoft Azure

You can deploy PCF on Microsoft Azure using Active Directory Federation Services (ADFS) for secured identity federation and single sign-on (SSO). To implement ADFS for Azure, you must use a certificate to authenticate on Azure, rather than a client and secret.

For more information about using ADFS for Azure, see [Use Service Principal with certificate](#) on GitHub.

Other Ops Manager API Improvements

In addition to features described above, the Ops Manager v2.6 API also includes the following changes:

- Operators can override the NATS maximum payload. Deployments with a large number of custom domains may use certificates that are larger than the default 1 MB NATS payload size. As a workaround, operators can override the default maximum. For more information, see [Updating director and IaaS properties \(Experimental\)](#) in the Ops Manager API documentation.
- When a request to `POST /api/v0/installations` fails to trigger Apply Changes, the response contains more details about which products are not fully configured. For more information, see [Triggering an installation](#) in the Ops Manager API documentation.

Known Issues

This release contains the following known issues:

Ops Manager Syslog Template Is Not Supported for Service Broker Tiles

Ops Manager syslog template does not propagate to service broker tiles. Tile developers with service broker tiles should continue to use their own syslog configurations. Service broker tiles with the syslog template configured do not display the syslog template.

For more information about the Syslog template, see [Syslog Form Template Available for Tile Authors](#).

Incorrect Link for Failed Installations in Ops Manager UI

When an installation fails, the Ops Manager UI presents a hyperlink to the installation log in the pop-up window. This link is currently non-functional.

To work around this issue, click on **Change Log** in the Ops Manager Installation Dashboard to view product installation logs.

Pivotal Application Service v2.6 Release Notes

Pivotal Cloud Foundry is certified by the Cloud Foundry Foundation for 2019.

Read more about the [certified provider program](#) and the [requirements of providers](#).

Releases

2.6.0

- See [New Features in PAS v2.6](#)
- See [Breaking Changes](#)
- [Breaking Change] Disables ssh proxy HTTP healthcheck server used for some GCP load balancers
- [Feature] Apps Manager sidebar is collapsible
- [Feature] Google Cloud Storage (GCS) blobstores are now backed up with BOSH backup and restore (BBR)
- [Feature] Improve scalability of application syslog drain system with new syslog agent architecture
- [Feature] Apps Manager announces page title changes to screen readers
- [Feature] Space Developers can now terminate a specific instance of an app in Apps Manager
- [Feature] Update the side bar, header and footer visuals in Apps Manager
- [Feature Improvement] Garden will now delegate container creates and destroys to containerd, an industry standard container runtime.
- [Feature] Enable Metric Registrar by default to enable App Developers to output custom application metrics that can be monitored by platform-provided tooling
- [Feature] Add optional System Metric Agent to allow monitoring a more complete set of metrics for all VMs in the deployment
- [Feature] `push-apps-manager` job can read configured UAA SAML providers
- [Feature] New color palette and other improvements for better accessibility when navigating Apps Manager
- [Feature Improvement] Improved performance in loading Apps Manager foundation home page
- [Feature Improvement] Apps Manager styling and accessibility improvements
- [Feature Improvement] Increase default number of CPUs for router from 1 to 2
- [Feature Improvement] Move ‘Enable SMB Volume Services’ checkbox from Advanced pane to Application Containers pane
- [Feature Improvement] Removes `nfs-legacy` NFS mounter. Existing `nfs-legacy` service instances will continue to work, but will use the newer `nfs` mounter
- [Security Fix] Usage service BBR script now uses ssl validation by default when logging into CAPI
- [Security Fix] Introduce and trust new Diego “root CA” in advance of existing CA expiration

Component	Version
ubuntu-xenial stemcell	315.36
backup-and-restore-sdk	1.15.1
binary-offline-buildpack	1.0.32
bosh-dns-aliases	0.0.3
bosh-system-metrics-forwarder	0.0.18
bpm	1.0.4
capi	1.80.0
cf-autoscaling	218
cf-backup-and-restore	0.0.11
cf-cli	1.13.0
cf-networking	2.22.2
cf-smoke-tests	40.0.108
cf-syslog-drain	10.2
cflinuxfs3	0.101.0

Component	Version
consul	198
credhub	2.4.0
diego	2.30.0
dotnet-core-offline-buildpack	2.2.11
garden-runc	1.19.1
go-offline-buildpack	1.8.39
haproxy	9.5.2
istio	1.0.1
java-offline-buildpack	4.18
leadership-election	1.4
log-cache	2.1.4
loggregator-agent	3.14
loggregator	105.5
mapfs	1.1.5
metric-registrar	1.0.4
mysql-monitoring	9.3.0
nats	26
nfs-volume	2.1.0
nginx-offline-buildpack	1.0.11
nodejs-offline-buildpack	1.6.49
notifications-ui	36
notifications	57
php-offline-buildpack	4.3.76
push-apps-manager-release	669.0.4
push-usage-service-release	670.0.6
pxc	0.16.0
python-offline-buildpack	1.6.32
r-offline-buildpack	1.0.9
routing	0.188.1
ruby-offline-buildpack	1.7.38
silk	2.22.2
smb-volume	1.1.0
staticfile-offline-buildpack	1.4.42
statsd-injector	1.10.0
syslog	11.4.0
uaa	71.0

How to Upgrade

The procedure for upgrading to Pivotal Application Service (PAS) v2.6 is documented in the [Upgrading Pivotal Cloud Foundry](#) topic.

When upgrading to PAS v2.6, be aware of the following upgrade considerations:

- If you previously used an earlier version of PAS, you must first upgrade to PAS v2.5 to successfully upgrade to PAS v2.6.
- Some partner service tiles may be incompatible with PCF v2.6. Pivotal is working with partners to ensure their tiles are updated to work with the latest versions of PCF.

For information about which partner service releases are currently compatible with PCF v2.6, review the appropriate partners services release

documentation at <https://docs.pivotal.io>, or contact the partner organization that produces the tile.

New Features in PAS v2.6

See the following new features for PAS v2.6:

Monitor System Metrics with System Metrics Agent

System Metrics Agent provides more visibility into VM compute, network, and storage metrics. These VM metrics help with troubleshooting and diagnosing issues for potential infrastructure problems.

When enabled, the metrics are emitted through Loggregator. For more information about viewing logs and metrics, see the *Platform components* row of [Viewing Logs and Metrics](#).

For a list of the VM metrics that the System Metric Agent emits, see [VM Metrics](#) in GitHub.

To enable the System Metric Agent, go to the **System Logging** pane in the PAS tile and select **Enable System Metrics**.

Metric Registrar Enabled by Default

In the **Metric Registrar** pane, Metric Registrar is enabled by default. This allows you to output custom application metrics that can be monitored by platform-provided tooling.

For information about configuring Metric Registrar, see the [\(Optional\) Configure Metric Registrar](#) section of the *Configuring PAS* topic.

Enable, View, and Rollback Revisions for Apps

PAS supports revisions for apps. A revision is an object that represents code and configuration used by an app at a specific time. Some uses cases for revisions include rolling back your app to a previous version and viewing changes in your app over time. You can also add metadata to revisions.

For more information, see [App Revisions](#).

Push Apps with Sidecar Processes (Beta)

You can run additional processes, or sidecars, in the same container as your app. Sidecars are useful for processes that depend on each other or must run in the same container. This includes processes that must communicate through localhost or share the same filesystem, such as an Application Performance Monitoring (APM) tool.

For more information, see [Pushing Apps with Sidecar Processes \(Beta\)](#).

Garden Delegates Container Creation and Destruction to containerd by Default

Traditionally, Garden uses runc directly to create, delete, and run container processes. Garden v1.15.0 and later support delegating some of the container lifecycle through containerd, which is the industry standard wrapper around runc.

Containerd mode is enabled by default to bring PAS in line with other container providers. To disable containerd mode, deselect the **Enable Containerd Delegation** checkbox in the **Application Containers** pane of the PAS tile.

For more information, see [opsguide-containerd.md](#) on GitHub.

NFS Legacy Mounter Removed

PAS v2.6 removes the `nfs-legacy` NFS mounter. Existing `nfs-legacy` service instances will continue to work, but will use the newer `nfs` mounter.

PAS v2.3 introduced the `nfs-experimental` service. PAS v2.4 made this the default `nfs` service, and the original service became `nfs-legacy`. For more

information, see [Experimental NFS Volume Service Supports NFSv4](#) and [NFS-Experimental Service Graduation](#).

Extra Diego Root CA Added to Prevent Future Expiration Issues

The current Diego root CA included with PAS v2.3 and later expires in mid-2020. To prevent potential expiration issues with intermediate CAs when the root CA expires, PAS v2.6 includes another Diego root CA that extends the root CA expiration date.

The Diego root CA is used to sign intermediate CAs, which are used to sign app identity certificates. In turn, app identity certificates are used to establish trust between applications and various components.

PAS v2.6 includes both the original root CA and the new root CA for Diego. This release note is informational only. No operator intervention is required.

Performance Improvements for Read-Write File Systems

PAS v2.6 improves the speed of read-write mounted file systems that use UID mapping. This improvement addresses performance issues in mapfs, which is enabled when you specify a UID that maps to an NFS volume.

For more information about using external file systems and NFS volume services, see [Using an External File System \(Volume Services\)](#).

Increased CPUs for Router VMs

To improve reliability, PAS v2.6 increases the default and minimum CPU core count from one to two for Router VMs.

PAS migrates Router VMs with a single core to a VM type with two CPU cores during the upgrade.

Loggregator Syslog Agent Increases Scale For Syslog Drains

 **warning:** See the following known issue related to this feature: [App Syslog Drains Fail After Enabling Agent-Based Syslog](#).

The Loggregator architecture includes optional Syslog Agents. Syslog Agents run on PCF component VMs and host VMs to manage connections with and write to syslog drains for app logs. The addition of Syslog Agents increases the number of syslog drain service bindings supported by the Loggregator system and reduces the workload for Loggregator VMs.

To enable Syslog Agents, select **Enable agent-based syslog egress for app logs** in the **System Logging** PAS configuration pane.

 **Note:** Enabling this feature disables the Syslog Adapter and Syslog Scheduler to avoid log duplication.

This update resolves a known issue where app log loss occurs at 10,000 syslog drain service bindings. For more information about the known issue, see the [Known Loggregator Scaling Issues](#) knowledge base article.

For more information about Syslog Agents, see [Loggregator Architecture](#).

Terminate Specific Instances of an App in Apps Manager UI

Apps Manager allows you to terminate a specific instance of an app through the UI. On the list of apps in the Apps Manager UI, a menu option called **Terminate Instance** appears.

For more information about terminating an instance, see [Terminate a process instance](#) in the Cloud Foundry API documentation.

View and Edit Labels and Annotations Associated with an Organization

In the **Settings** pane of the Apps Manager UI, the **Metadata** section contains lists of labels and annotations associated with an organization. You can edit these labels and annotations.

For more information, see the [Add Metadata](#) section of the *Managing Orgs and Spaces Using Apps Manager* topic.

Known Issues

App Syslog Drains Fail After Enabling Agent-Based Syslog

If you select **Enable agent-based syslog egress for app logs** in the **System Logging** pane of the PAS tile, external syslog drains that are bound to Windows Apps cannot collect logs. For more information, see [“Enable agent-based syslog egress for app logs” interrupts external log collection for PAS Windows apps](#).

PAS for Windows v2.6 Release Notes

This topic contains release notes for Pivotal Application Service (PAS) for Windows.

How to Upgrade

The PAS for Windows v2.6 tile is available with the release of Pivotal Cloud Foundry (PCF) version v2.6. To use the PAS for Windows v2.6 tile, you must install Ops Manager v2.6 or later and PAS v2.6 or later.

If you are running PAS for Windows v2.4.4 or later on PCF v2.5 and want to upgrade to PCF v2.6, you must upgrade to PAS for Windows v2.5 before upgrading to PCF v2.6.

Releases

2.6.0

 **Breaking Change:** Option to override deployment name to CF is removed.

- [Feature] App developers can gracefully shutdown their .NET applications in PASW
- [Feature] Add optional System Metric Agent to allow monitoring a more complete set of metrics for all VMs in the deployment
- [Feature] Operators can provide trusted certs in OpsMgr for .NET apps to use automatically
- [Feature Improvement] Allow vcap WriteKey access to ODBC registry entry
- [Feature Improvement] Improve scalability of application syslog drain system with new syslog agent architecture
- [Security Fix] Introduce and trust new Diego “root CA” in advance of existing CA expiration
- Bump windows2019 stemcell to version `2019.5`
- Bump hwc-offline-buildpack to version `3.1.9`
- Bump loggregator-agent to version `3.14`
- Bump loggregator to version `105.5`
- Bump winc to version `1.13.0`
- Bump windowsfs-release to version `1.5.0`

Component	Version
windows2019 stemcell	2019.5
cf-windows-smoke-tests	40.0.51
diego	2.30.0
event-log	0.8.0
garden-runc	1.19.1
hwc-offline-buildpack	3.1.9
loggregator-agent	3.14
loggregator	105.5
winc	1.13.0
windows-utilities	0.11.0
windowsfs-release	1.5.0

New Features in PAS for Windows v2.6

PAS for Windows v2.6.0 includes the following major features:

App Developers Can Gracefully Shutdown Their .NET Applications in PASW

PAS for Windows v2.6 allows ten seconds for a .NET app to shutdown gracefully after the app is stopped using `cf stop`. Graceful shutdown is supported on deployments using Windows Server 2019 or later for apps that provide a control handler for the `CTRL_SHUTDOWN_EVENT` event.

For more information, see [Graceful Shutdown](#) in the .NET on Pivotal Cloud Foundry Cookbook.

Introduction of System Metrics Agent

PAS for Windows v2.6 includes the introduction of the System Metrics Agent. When the System Metrics Agent is enabled, Windows cells emit system-level metrics about each VM in your deployment.

You can enable this feature in the PAS for Windows v2.6 tile by selecting **Enable System Metrics** within the **System Logging** pane.

For a list of the VM metrics that the System Metric Agent emits for Windows VMs, see [VM Metrics](#) in GitHub.

For additional information about system-level metrics in PCF, see [BOSH VM and PCF Component Health Metrics](#) and [Key Performance Indicators](#).

Deprecation of the windows2016 Stack

You should migrate any apps that run on the `windows2016` stack to the `windows` stack.

You can migrate your apps from `windows2016` to `windows` using Stack Auditor, a Cloud Foundry CLI plugin. For more information, see [Using the Stack Auditor Plugin](#).

ODBC Driver Installation Support in Windows Containers

PAS for Windows v2.6 supports ODBC driver installation by giving developers the ability to add ODBC driver entries to the Windows Registry. Previously, ODBC drivers were installed to Windows cells using BOSH add-ons. Installing ODBC drivers by using BOSH add-ons is no longer possible in true Windows containerized runtimes.

BOSH SSH Enabled by Default

In PAS for Windows v2.6, use of the `bosh ssh` command is enabled by default for deployed VMs. Developers can use the `bosh ssh` command to troubleshoot Windows-based Diego cells.

Deploy Windows Stemcells on BOSH Director with GCS

PAS for Windows v2.6 supports deploying Windows stemcells with a BOSH Director that uses GCS as the BOSH blobstore.

For more information on configuring a GCS blobstore, see *Director Config Page* in [Configuring BOSH Director on GCP Manually](#) or *Director Config Page* in [Configuring BOSH Director on GCP Using Terraform](#).

Improved Application Syslog Drain Scalability

PAS for Windows v2.6 has improved application syslog drain scalability and incorporates a new syslog agent architecture.

Known Issues

This section includes known issues with PAS for Windows v2.6.

App Syslog Drains Fail After Enabling Agent-Based Syslog

If you select [Enable agent-based syslog egress for app logs](#) in the **System Logging** pane of the PAS tile, external syslog drains that are bound to Windows Apps cannot collect logs. For more information, see [“Enable agent-based syslog egress for app logs” interrupts external log collection for PAS Windows apps](#).

Graceful Shutdown Is Not Supported on Deployments with One Windows Cell

Graceful shutdown is supported only on Windows Server 2019 deployments with two or more Windows cells. For more information, see [Graceful shutdown works only for CF deployments with multiple Windows 2019 cells](#).

PKS Release Notes

Page last updated:

This topic contains release notes for Pivotal Container Service (PKS) v1.3.x.

v1.3.6

Release Date: April 8, 2019

Product Snapshot

Element	Details
Version	v1.3.6
Release date	April 8, 2019
Compatible Ops Manager versions	v2.3.1+, v2.4.0+
Stemcell version	v170.15
Kubernetes version	v1.12.7
On-Demand Broker version	v0.24
CFCR	v0.25.11
NSX-T versions *	v2.3.1, v2.4.0.1
NCP version	v2.4.0
Docker version	v18.06.3-ce CFCR v0.25.11

Note: Ops Manager v2.3.10 and later in the v2.3 version line and Ops Manager v2.4.4 and later in the v2.4 version line do not support PKS v1.3 on Azure. Before deploying PKS v1.3 on Azure, you must install Ops Manager v2.3.9 or earlier in the 2.3 version line or Ops Manager v2.4.3 or earlier in the 2.4 version line.

Note: NSX-T v2.4 implements a new Policy API that PKS v1.3.6 does not support. If you are using NSX-T v2.4 with PKS 1.3.6, you must use the “Advanced Networking” tab in NSX Manager to create, read, update, and delete network object required for PKS.

vSphere Version Requirements

If you are installing PKS on vSphere or vSphere with NSX-T, note that Ops Manager and PKS support the following vSphere component versions:

Versions	Editions
<ul style="list-style-type: none"> VMware vSphere 6.7 U1 EP06 (ESXi670-201901001) – for NSX-T 2.4 VMware vSphere 6.7 U1 VMware vSphere 6.7.0 VMware vSphere 6.5 U2 P03 (ESXi650-201811002) – for NSX-T 2.4 VMware vSphere 6.5 U2 VMware vSphere 6.5 U1 	<ul style="list-style-type: none"> vSphere Enterprise Plus vSphere with Operations Management Enterprise Plus

Note: VMware vSphere 6.7 is only supported with Ops Manager v2.3.1 or later and NSX-T v2.3.

For more information, see [Upgrading vSphere in an NSX Environment](#) in the VMware documentation.

Feature Support by IaaS

	AWS	Azure	GCP	vSphere	vSphere with NSX-T
Automatic Kubernetes Cluster API load balancer					✓
HTTP proxy				✓	✓
Multi-AZ storage				✓	✓
Per-namespace subnets					✓
Service <code>type:LoadBalancer</code>	✓*	✓	✓		✓

Upgrade Path

The supported upgrade paths to PKS v1.3.6 are as follows:

- PKS v1.3.4 or later

When upgrading to NSX-T 2.4:

- Use the official VMware NSX-T Data Center 2.4 build.
- Apply the NSX-T v2.4.0.1 hot-patch. For more information, see VMware Knowledge Base [KB article 67499](#).
- To obtain the NSX-T v2.4.0.1 hot-patch, open a support ticket with VMware Global Support Services (GSS) for NSX-T Engineering.

What's New

PKS v1.3.6 adds the following:

- Telemetry property `environment_provider`.
- Support for `nsx-cf-cni` with 2.4.0.12511604.
- Remaining plans to the osb-proxy configuration.

Breaking Changes and Known Issues

 **Breaking Change:** Heapster is deprecated in PKS v1.3.x, and Kubernetes has retired Heapster. For more information, see the [kubernetes-retired/heapster](#) repository on GitHub.

PKS v1.3.6 has the following known issues:

Azure Resource Group Field in the Kubernetes Cloud Provider Is Ignored

The PKS tile's **Resource Group** configuration is ignored on Azure IaaS platform deployments. On Azure, the PKS VM is always deployed to the same **Resource Group** as the Ops Manager and BOSH VMs. The **Resource Group** field is in the PKS tile's **Kubernetes Cloud Provider** section.

NSX-T Upgrades from v2.3.X to v2.4.0.1 Fail for Bare Metal Edge Node

Upgrading NSX-T v2.3.X to v2.4.0.1 fails for Bare Metal Edge Nodes.

If you are using a Bare Metal Edge Nodes, please refrain from upgrading NSX-T v2.3.x to NSX-T v2.4.0.1.

Worker Nodes with Small Ephemeral Disks Can Cause Upgrade Failure

PKS deploys packages to the ephemeral disk, `/var/vcap/data`, during installations and upgrades. If master or worker node VMs have ephemeral disks smaller than 8 GB, the disk can fill during an upgrade and cause the upgrade to fail. Cluster upgrades can present error messages such as the following:

```
{"time":999999999,"error":{"code":450001,"message":"Response exceeded maximum allowed length"}}
```

Workaround: In the plans you use to deploy clusters, ensure that worker and master node ephemeral disks are set to greater than 8 GB. For plan configuration instructions, see the [Plans](#) section of the *Installing PKS* topic for your IaaS.

This issue should not affect new installations of PKS v1.3.x as the default ephemeral disk size in plans is larger than 8 GB.

PKS Flannel Network Gets Out of Sync with Docker Bridge Network (cni0)

When VMs have been powered down for multiple days, turning them back on and issuing a `bosh recreate` to re-create the VMs causes the pods to get stuck in a `ContainerCreating` state.

Workaround: See [PKS Flannel network gets out of sync with docker bridge network \(cni0\)](#) in the Pivotal Knowledge Base.

Cluster Upgrades from PKS v1.3.0 on Azure Fail If Services Are Exposed

If you install PKS v1.3.0 on Azure, clusters might fail with the following error when you upgrade to PKS v1.3.1 or later:

```
result: 1 of 2 post-start scripts failed. Failed Jobs: kubelet. Successful Jobs: bosh-dns
```

This issue is caused by a timeout condition. The issue affects nodes hosting Kubernetes pods that are exposed externally by a Kubernetes service.

New cluster creations and cluster scaling operations are not affected by this issue.

Workaround: If you install PKS on Azure and experience this issue, contact Support for assistance.

The kubelet customization feature is only enabled for Plan 1

PKS 1.3.4 introduces the ability to configure kubelet startup parameters `system-reserved` and `eviction-hard` within a [plan](#). This capability is only functional in Plan 1 for PKS 1.3.4 and will be enabled in additional plans in the next release.

v1.3.5

Release Date: March 28, 2019

Product Snapshot

Element	Details
Version	v1.3.5
Release date	March 28, 2019
Compatible Ops Manager versions	v2.3.1+, v2.4.0+
Stemcell version	v170.15
Kubernetes version	v1.12.7
On-Demand Broker version	v0.24
CFCR	v0.25.11
NSX-T versions *	v2.2, v2.3.0.2, v2.3.1
NCP version	v2.3.2
Docker version	v18.06.3-ce CFCR v0.25.11

 **Note:** Ops Manager v2.3.10 and later in the v2.3 version line and Ops Manager v2.4.4 and later in the v2.4 version line do not support PKS v1.3 on Azure. Before deploying PKS v1.3 on Azure, you must install Ops Manager v2.3.9 or earlier in the 2.3 version line or Ops Manager v2.4.3 or earlier in the 2.4 version line.

vSphere Version Requirements

If you are installing PKS on vSphere or vSphere with NSX-T, note that Ops Manager and PKS support the following vSphere component versions:

Versions	Editions
<ul style="list-style-type: none"> VMware vSphere 6.7 U1 EP06 (ESXi670-201901001) – for NSX-T 2.4 VMware vSphere 6.7 U1 VMware vSphere 6.7.0 VMware vSphere 6.5 U2 P03 (ESXi650-201811002) – for NSX-T 2.4 VMware vSphere 6.5 U2 VMware vSphere 6.5 U1 	<ul style="list-style-type: none"> vSphere Enterprise Plus vSphere with Operations Management Enterprise Plus

 **Note:** VMware vSphere 6.7 is only supported with Ops Manager v2.3.1 or later and NSX-T v2.3.

For more information, see [Upgrading vSphere in an NSX Environment](#) in the VMware documentation.

Feature Support by IaaS

	AWS	Azure	GCP	vSphere	vSphere with NSX-T
Automatic Kubernetes Cluster API load balancer					✓
HTTP proxy				✓	✓
Multi-AZ storage				✓	✓
Per-namespace subnets					✓
Service <code>type:LoadBalancer</code>	✓*	✓	✓		✓

Upgrade Path

The supported upgrade paths to PKS v1.3.5 are as follows:

- PKS v1.3.4 or later

What's New

PKS v1.3.5 adds the following:

- Support for Kubernetes v1.12.7.
- Fix: [CVE-2019-1002101](#). Kubernetes v1.12.7 address this CVE.
- Fix: [CVE-2019-9946](#). Kubernetes v1.12.7 address this CVE.

Breaking Changes and Known Issues

 **Breaking Change:** Heapster is deprecated in PKS v1.3.x, and Kubernetes has retired Heapster. For more information, see the [kubernetes-retired/heapster](#) repository on GitHub.

PKS v1.3.5 has the following known issues:

Worker Nodes with Small Ephemeral Disks Can Cause Upgrade Failure

PKS deploys packages to the ephemeral disk, `/var/vcap/data`, during installations and upgrades. If master or worker node VMs have ephemeral disks smaller than 8 GB, the disk can fill during an upgrade and cause the upgrade to fail. Cluster upgrades can present error messages such as the following:

```
{"time":99999999,"error":{"code":450001,"message":"Response exceeded maximum allowed length"}}
```

Workaround: In the plans you use to deploy clusters, ensure that worker and master node ephemeral disks are set to greater than 8 GB. For plan configuration instructions, see the [Plans](#) section of the *Installing PKS* topic for your IaaS.

This issue should not affect new installations of PKS v1.3.x as the default ephemeral disk size in plans is larger than 8 GB.

PKS Flannel Network Gets Out of Sync with Docker Bridge Network (cni0)

When VMs have been powered down for multiple days, turning them back on and issuing a `bosh recreate` to re-create the VMs causes the pods to get stuck in a `ContainerCreating` state.

Workaround: See [PKS Flannel network gets out of sync with docker bridge network \(cni0\)](#) in the Pivotal Knowledge Base.

Cluster Upgrades from PKS v1.3.0 on Azure Fail If Services Are Exposed

If you install PKS v1.3.0 on Azure, clusters might fail with the following error when you upgrade to PKS v1.3.1 or later:

```
result: 1 of 2 post-start scripts failed. Failed Jobs: kubelet. Successful Jobs: bosh-dns
```

This issue is caused by a timeout condition. The issue affects nodes hosting Kubernetes pods that are exposed externally by a Kubernetes service.

New cluster creations and cluster scaling operations are not affected by this issue.

Workaround: If you install PKS on Azure and experience this issue, contact Support for assistance.

The kubelet customization feature is only enabled for Plan 1

PKS 1.3.4 introduces the ability to configure kubelet startup parameters `system-reserved` and `eviction-hard` within a [plan](#). This capability is only functional in Plan 1 for PKS 1.3.4 and will be enabled in additional plans in the next release.

v1.3.4

Release Date: March 26, 2019

Product Snapshot

Element	Details
Version	v1.3.4
Release date	March 26, 2019
Compatible Ops Manager versions	v2.3.1+, v2.4.0+
Stemcell version	v170.15
Kubernetes version	v1.12.6
On-Demand Broker version	v0.24
CFCR	v0.25.11
NSX-T versions *	v2.2, v2.3.0.2, v2.3.1
NCP version	v2.3.2
Docker version	v18.06.3-ce CFCR v0.25.11

Note: Ops Manager v2.3.10 and later in the v2.3 version line and Ops Manager v2.4.4 and later in the v2.4 version line do not support PKS v1.3 on Azure. Before deploying PKS v1.3 on Azure, you must install Ops Manager v2.3.9 or earlier in the 2.3 version line or Ops Manager v2.4.3 or earlier in the 2.4 version line.

vSphere Version Requirements

If you are installing PKS on vSphere or vSphere with NSX-T, note that Ops Manager and PKS support the following vSphere component versions:

Versions	Editions
<ul style="list-style-type: none"> VMware vSphere 6.7 U1 EP06 (ESXi670-201901001) – for NSX-T 2.4 VMware vSphere 6.7 U1 VMware vSphere 6.7.0 VMware vSphere 6.5 U2 P03 (ESXi650-201811002) – for NSX-T 2.4 VMware vSphere 6.5 U2 VMware vSphere 6.5 U1 	<ul style="list-style-type: none"> vSphere Enterprise Plus vSphere with Operations Management Enterprise Plus

Note: VMware vSphere 6.7 is only supported with Ops Manager v2.3.1 or later and NSX-T v2.3.

For more information, see [Upgrading vSphere in an NSX Environment](#) in the VMware documentation.

Feature Support by IaaS

	AWS	Azure	GCP	vSphere	vSphere with NSX-T
Automatic Kubernetes Cluster API load balancer					✓
HTTP proxy				✓	✓
Multi-AZ storage				✓	✓
Per-namespace subnets					✓
Service <code>type:LoadBalancer</code>	✓*	✓	✓		✓

Upgrade Path

The supported upgrade paths to PKS v1.3.3 are as follows:

- When upgrading from PKS v1.3.x: PKS v1.3.1 or later
- When upgrading from PKS v1.2.x: PKS v1.2.8 or later

What's New

PKS v1.3.4 adds the following:

- Custom DNS configuration for Kubernetes clusters using NSX-T and Network Profiles. For more information, see [DNS Configuration for Kubernetes Clusters](#) in [Defining Network Profiles](#).
- Support for NSX-T NCP v2.3.2. For more information, see the [VMware NSX Container Plug-in 2.3.2 Release Notes](#).
- Support for additional plans. Operators can configure up to ten sets of resource types, or **Plans**, in the PKS tile. All plans except the first can made available or unavailable to developers deploying clusters. **Plan 1** must be configured and made available as a default for developers.
- Kubelet customization. You can enable Kubelet to reserve compute resources for system daemons by configuring the startup parameters `system-reserved` and `eviction-hard` in the **Plans** pane of the PKS tile. For more information, see the *Plans* section of the *Installing PKS* topic for your IaaS, such as [Installing PKS on vSphere](#).
- Fix: [CVE-2019-1002100](#). Kubernetes v1.12.6 address this CVE.
- Fix: Updated the Telemetry URL.

- **Fix:** Resolved an issue where vSphere Cloud Provider configuration could fail if credentials contained non-alphanumeric characters. For example, `#`, `\`, and `"`.

Breaking Changes and Known Issues

 **Breaking Change:** Heapster is deprecated in PKS v1.3.x, and Kubernetes has retired Heapster. For more information, see the [kubernetes-retired/heapster](#) repository on GitHub.

PKS v1.3.4 has the following known issues:

Master and Worker Nodes with Small Ephemeral Disks Can Cause Upgrade Failure

PKS deploys packages to the ephemeral disk, `/var/vcap/data`, during installations and upgrades. If master and worker node VMs have ephemeral disks smaller than 8 GB, the disk can fill during an upgrade and cause the upgrade to fail. Cluster upgrades can present error messages such as the following:

```
{"time":99999999,"error":{"code":450001,"message":"Response exceeded maximum allowed length"}}
```

Workaround: In the plans you use to deploy clusters, ensure that the master and worker node ephemeral disks are set to greater than 8 GB. For plan configuration instructions, see the *Plans* section of the *Installing PKS* topic for your IaaS, such as [Installing PKS on vSphere](#).

This issue should not affect new installations of PKS v1.3.x as the default ephemeral disk size in plans is larger than 8 GB.

PKS Flannel Network Gets Out of Sync with Docker Bridge Network (cni0)

When VMs have been powered down for multiple days, turning them back on and issuing a `bosh recreate` to re-create the VMs causes the pods to get stuck in a `ContainerCreating` state.

Workaround: See [PKS Flannel network gets out of sync with docker bridge network \(cni0\)](#) in the Pivotal Knowledge Base.

Cluster Upgrades from PKS v1.3.0 on Azure Fail If Services Are Exposed

If you install PKS v1.3.0 on Azure, clusters might fail with the following error when you upgrade to PKS v1.3.1 or later:

```
result: 1 of 2 post-start scripts failed. Failed Jobs: kubelet. Successful Jobs: bosh-dns
```

This issue is caused by a timeout condition. The issue affects nodes hosting Kubernetes pods that are exposed externally by a Kubernetes service.

New cluster creations and cluster scaling operations are not affected by this issue.

Workaround: If you install PKS on Azure and experience this issue, contact Support for assistance.

Kubelet Customization Feature Only Enabled for Plan 1

PKS v1.3.4 introduces the ability to configure Kubelet startup parameters `system-reserved` and `eviction-hard` within a plan. For more information, see the *Plans* section of the *Installing PKS* topic for your IaaS, such as [Installing PKS on vSphere](#).

This feature is only functional in Plan 1 for PKS v1.3.4 and will be enabled in additional plans in the next release.

v1.3.3

Release Date: February 22, 2019

Product Snapshot

Element	Details
Version	v1.3.3
Release date	February 22, 2019
Compatible Ops Manager versions	v2.3.1+, v2.4.0+
Stemcell version	v170.15
Kubernetes version	v1.12.5
On-Demand Broker version	v0.24
CFCR	v0.25.9
NSX-T versions *	v2.2, v2.3.0.2, v2.3.1
NCP version	v2.3.1
Docker version	v18.06.3-ce CFCR v0.25.9

Note: Ops Manager v2.3.10 and later in the v2.3 version line and Ops Manager v2.4.4 and later in the v2.4 version line do not support PKS v1.3 on Azure. Before deploying PKS v1.3 on Azure, you must install Ops Manager v2.3.9 or earlier in the 2.3 version line or Ops Manager v2.4.3 or earlier in the 2.4 version line.

Feature Support by IaaS

	AWS	Azure	GCP	vSphere	vSphere with NSX-T
Automatic Kubernetes Cluster API load balancer					✓
HTTP proxy				✓	✓
Multi-AZ storage				✓	✓
Per-namespace subnets					✓
Service <code>type:LoadBalancer</code>	✓*	✓	✓		✓

Upgrade Path

The supported upgrade paths to PKS v1.3.3 are as follows:

- When upgrading from PKS v1.3.x: PKS v1.3.1 or v1.3.2
- When upgrading from PKS v1.2.x: PKS v1.2.8 through v1.2.11

What's New

PKS v1.3.3 adds the following:

- Fix: [CVE-2019-5736](#). This release updates the version of Docker deployed by PKS to v18.06.3-ce. This Docker version addresses a runc vulnerability whereby a malicious image could run in privileged mode and elevate to root access on worker nodes. Docker v18.06.2-ce, deployed by PKS v1.3.2, did not contain the correct compiled binary. This Docker version includes the correct runc binary to address the CVE.

Breaking Changes and Known Issues

Breaking Change: Heapster is deprecated in PKS v1.3.x, and Kubernetes has retired Heapster. For more information, see the [kubernetes-retired/heapster](#) repository on GitHub.

PKS v1.3.3 has the following known issues:

PKS Flannel Network Gets Out of Sync with Docker Bridge Network (cni0)

When VMs have been powered down for multiple days, turning them back on and issuing a `bosh recreate` to re-create the VMs causes the pods to get stuck in a `ContainerCreating` state.

Workaround: See [PKS Flannel network gets out of sync with docker bridge network \(cni0\)](#) in the Pivotal Knowledge Base.

Deploy Fails if vSphere Master Credentials Field Has Special Characters Without Quotes

If you install PKS on vSphere and you enter credentials in the **vCenter Master Credentials** field of the **Kubernetes Cloud Provider** pane of the PKS tile that contain special characters, such as `#`, `$`, `,`, `!`, or `-`, your deployment might fail with the following error:

```
ServerFaultCode: Cannot complete login due to an incorrect user name or password.
```

Workaround: If you install PKS on vSphere **without** NSX-T integration, place quotes around the credentials in the cloud provider configuration. For example, `"SomeP4$$w0rd#!"`. Then redeploy the PKS tile by clicking **Apply Changes**.

If you install PKS on vSphere **with** NSX-T integration, avoid using special characters in this field until this issue is resolved.

Cluster Upgrades from PKS v1.3.0 on Azure Fail If Services Are Exposed

If you install PKS v1.3.0 on Azure, clusters might fail with the following error when you upgrade to PKS v1.3.1 or later:

```
result: 1 of 2 post-start scripts failed. Failed Jobs: kubelet. Successful Jobs: bosh-dns
```

This issue is caused by a timeout condition. The issue affects nodes hosting Kubernetes pods that are exposed externally by a Kubernetes service.

New cluster creations and cluster scaling operations are not affected by this issue.

Workaround: If you install PKS on Azure and experience this issue, contact Support for assistance.

v1.3.2

Release Date: February 13, 2019

Product Snapshot

Element	Details
Version	v1.3.2
Release date	February 13, 2019
Compatible Ops Manager versions	v2.3.1+, v2.4.0+
Stemcell version	v170.15
Kubernetes version	v1.12.4
On-Demand Broker version	v0.24
CFCR	v0.25.8
NSX-T versions *	v2.2, v2.3.0.2, v2.3.1
NCP version	v2.3.1
Docker version	v18.06.2-ce CFCR v0.25.8

 **Note:** Ops Manager v2.3.10 and later in the v2.3 version line and Ops Manager v2.4.4 and later in the v2.4 version line do not support PKS v1.3 on Azure. Before deploying PKS v1.3 on Azure, you must install Ops Manager v2.3.9 or earlier in the 2.3 version line or Ops Manager v2.4.3 or earlier in the 2.4 version line.

Feature Support by IaaS

	AWS	Azure	GCP	vSphere	vSphere with NSX-T
Automatic Kubernetes Cluster API load balancer					✓
HTTP proxy				✓	✓
Multi-AZ storage				✓	✓
Per-namespace subnets					✓
Service <code>type:LoadBalancer</code>	✓*	✓	✓		✓

Upgrade Path

The supported upgrade paths to PKS v1.3.2 are as follows:

- When upgrading from PKS v1.3.x: PKS v1.3.1
- When upgrading from PKS v1.2.x: PKS v1.2.8 or v1.2.9

What's New

PKS v1.3.2 adds the following:

- Fix: [CVE-2019-3779](#). This fix addresses a vulnerability where certs signed by the Kubernetes API could be used to gain access to a PKS-deployed cluster's etcd service.
- Fix: [CVE-2019-3780](#). This fixes a regression bug in PKS where vCenter IaaS credentials intended for the vSphere Cloud Provider were written on worker node VM disks.
- Fix: Clusters can now be successfully created if there are pre-existing Kubernetes clusters using the same hostname.

Breaking Changes and Known Issues

 **Breaking Change:** Heapster is deprecated in PKS v1.3.x, and Kubernetes has retired Heapster. For more information, see the [kubernetes-retired/heapster](#) repository on GitHub.

PKS v1.3.2 has the following known issues:

PKS Flannel Network Gets Out of Sync with Docker Bridge Network (cni0)

When VMs have been powered down for multiple days, turning them back on and issuing a `bosh recreate` to re-create the VMs causes the pods to get stuck in a `ContainerCreating` state.

Workaround: See [PKS Flannel network gets out of sync with docker bridge network \(cni0\)](#) in the Pivotal Knowledge Base.

Deploy Fails if vSphere Master Credentials Field Has Special Characters Without Quotes

If you install PKS on vSphere and you enter credentials in the **vCenter Master Credentials** field of the **Kubernetes Cloud Provider** pane of the PKS tile that contain special characters, such as `#`, `$`, `,`, `,`, `!`, or `-`, your deployment might fail with the following error:

ServerFaultCode: Cannot complete login due to an incorrect user name or password.

Workaround: If you install PKS on vSphere **without** NSX-T integration, place quotes around the credentials in the cloud provider configuration. For example, `"SomeP4$$w0rd#!"`. Then redeploy the PKS tile by clicking **Apply Changes**.

If you install PKS on vSphere **with** NSX-T integration, avoid using special characters in this field until this issue is resolved.

Cluster Upgrades from PKS v1.3.0 on Azure Fail If Services Are Exposed

If you install PKS v1.3.0 on Azure, clusters might fail with the following error when you upgrade to either PKS v1.3.1 or later:

```
result: 1 of 2 post-start scripts failed. Failed Jobs: kubelet. Successful Jobs: bosh-dns
```

This issue is caused by a timeout condition. The issue affects nodes hosting Kubernetes pods that are exposed externally by a Kubernetes service.

New cluster creations and cluster scaling operations are not affected by this issue.

Workaround: If you install PKS on Azure and experience this issue, contact Support for assistance.

v1.3.1

Release Date: February 8, 2019

⚠ warning: PKS v1.3.1 and earlier includes a critical CVE. Follow the procedures in the [PKS upgrade approach for CRITICAL CVE](#) article in the Pivotal Support Knowledge Base to perform an upgrade to PKS v1.3.2.

Product Snapshot

Element	Details
Version	v1.3.1
Release date	February 8, 2019
Compatible Ops Manager versions	v2.3.1+, v2.4.0+
Stemcell version	v170.15
Kubernetes version	v1.12.4
On-Demand Broker version	v0.24
CFCR	v0.25.8
NSX-T versions *	v2.2, v2.3.0.2, v2.3.1
NCP version	v2.3.1
Docker version	v18.06.1-ce CFCR v0.25.8

💡 Note: Ops Manager v2.3.10 and later in the v2.3 version line and Ops Manager v2.4.4 and later in the v2.4 version line do not support PKS v1.3 on Azure. Before deploying PKS v1.3 on Azure, you must install Ops Manager v2.3.9 or earlier in the 2.3 version line or Ops Manager v2.4.3 or earlier in the 2.4 version line.

vSphere Version Requirements

If installing PKS on vSphere or vSphere with NSX-T, please note Ops Manager and PKS support the following vSphere component versions:

Versions	Editions
<ul style="list-style-type: none"> VMware vSphere 6.7 U1 EP06 (ESXi670-201901001) – for NSX-T 2.4 VMware vSphere 6.7 U1 VMware vSphere 6.7.0 VMware vSphere 6.5 U2 P03 (ESXi650-201811002) – for NSX-T 2.4 VMware vSphere 6.5 U2 VMware vSphere 6.5 U1 	<ul style="list-style-type: none"> vSphere Enterprise Plus vSphere with Operations Management Enterprise Plus

💡 Note: VMware vSphere 6.7 is only supported with Ops Manager v2.3.1 or later and NSX-T v2.3.

For more information, see [Upgrading vSphere in an NSX Environment](#) in the VMware documentation.

Feature Support by IaaS

	AWS	Azure	GCP	vSphere	vSphere with NSX-T
Automatic Kubernetes Cluster API load balancer					✓
HTTP proxy				✓	✓
Multi-AZ storage				✓	✓
Per-namespace subnets					✓
Service <code>type:LoadBalancer</code>	✓*	✓	✓		✓

Upgrade Path

The supported upgrade paths to PKS v1.3.1 are as follows:

- When upgrading from PKS v1.3.x: PKS v1.3.0
- When upgrading from PKS v1.2.x: PKS v1.2.7 or v1.2.8

Follow the procedures in the [PKS upgrade approach for CRITICAL CVE](#) article in the Pivotal Support Knowledge Base to perform an upgrade to PKS v1.3.2.

What's New

PKS v1.3.1 adds support for the following:

- Certificates for the Etcd instance for each Kubernetes cluster provisioned by PKS are generated with a four-year lifetime and signed by a new Etcd Certificate Authority (CA).
- Fix: Upgrading PKS no longer fails during upgrades if there are Kubernetes clusters with duplicate hostnames.
- Fix: Deploying PKS no longer fails if an entry in the **No Proxy** field contains special characters such as (-) character.
- Fix: The Kubernetes API now responds with the CA certificate that signed the Kubernetes cluster's certificate so that customer scripts such as the [get-pks-k8s-config.sh](#) tool will function again.

Breaking Changes and Known Issues

 **Breaking Change:** Heapster is deprecated in PKS v1.3.x, and Kubernetes has retired Heapster. For more information, see the [kubernetes-retired/heapster](#) repository on GitHub.

PKS v1.3.1 has the following known issues:

PKS Flannel Network Gets out of Sync with Docker Bridge Network (cni0)

When VMs have been powered down for multiple days, turning them back on and issuing a `bosh recreate` to re-create the VMs causes the pods to get stuck in a `ContainerCreating` state.

Workaround: See [PKS Flannel network gets out of sync with docker bridge network \(cni0\)](#) in the Pivotal Knowledge Base.

Deploy Fails if vSphere Master Credentials Field Has Special Characters Without Quotes

If you install PKS on vSphere and you enter credentials in the **vCenter Master Credentials** field of the **Kubernetes Cloud Provider** pane of the PKS tile that contain special characters, such as #, \$, , , !, or -, your deployment might fail with the following error:

ServerFaultCode: Cannot complete login due to an incorrect user name or password.

Workaround: If you install PKS on vSphere **without** NSX-T integration, place quotes around the credentials in the cloud provider configuration. For example, `"SomeP4$$w0rd#!"`. Then redeploy the PKS tile by clicking **Apply Changes**.

If you install PKS on vSphere **with** NSX-T integration, avoid using special characters in this field until this issue is resolved.

Cluster Upgrades from PKS v1.3.0 on Azure Fail If Services Are Exposed

If you install PKS v1.3.0 on Azure, clusters might fail with the following error when you upgrade to either PKS v1.3.1 or later:

```
result: 1 of 2 post-start scripts failed. Failed Jobs: kubelet. Successful Jobs: bosh-dns
```

This issue is caused by a timeout condition. The issue affects nodes hosting Kubernetes pods that are exposed externally by a Kubernetes service.

New cluster creations and cluster scaling operations are not affected by this issue.

v1.3.0

Release Date: January 16, 2019

A warning: PKS v1.3.0 has a known vulnerability and is no longer available. Install or upgrade to [PKS v1.3.1](#).

Product Snapshot

Element	Details
Version	v1.3.0
Release date	January 16, 2019
Compatible Ops Manager versions	v2.3.1+, v2.4.0+
Stemcell version	v170.15
Kubernetes version	v1.12.4
On-Demand Broker version	v0.24
NSX-T versions *	v2.2, v2.3.0.2, v2.3.1
NCP version	v2.3.1
Docker version	v18.06.1-ce CFCR

* PKS v1.3 supports NSX-T v2.2 and v2.3 with the following caveats:

- To use the network profile features available in PKS v1.3, you must use NSX-T v2.3.
- NSX-T 2.3.0 has a known critical issue: [ESX hosts lose network connectivity rendering the host inaccessible from network \(60293\)](#). This issue is patched in NSX-T v2.3.0.2 and fixed in the NSX-T v2.3.1 release.

Note: Ops Manager v2.3.10 and later in the v2.3 version line and Ops Manager v2.4.4 and later in the v2.4 version line do not support PKS v1.3 on Azure. Before deploying PKS v1.3 on Azure, you must install Ops Manager v2.3.9 or earlier in the 2.3 version line or Ops Manager v2.4.3 or earlier in the 2.4 version line.

vSphere Version Requirements

If installing PKS on vSphere or vSphere with NSX-T, please note Ops Manager and PKS support the following vSphere component versions:

Versions	Editions
• VMware vSphere 6.7 U1 EP06 (ESXi670-201901001) – for NSX-T 2.4	

- | | |
|--|---|
| <ul style="list-style-type: none"> • VMware vSphere 6.7 U1 • VMware vSphere 6.7.0 • VMware vSphere 6.5 U2 P03 (ESXi650-201811002) – for NSX-T 2.4 • VMware vSphere 6.5 U2 • VMware vSphere 6.5 U1 | <ul style="list-style-type: none"> • vSphere Enterprise Plus • vSphere with Operations Management Enterprise Plus |
|--|---|

 **Note:** VMware vSphere 6.7 is only supported with Ops Manager v2.3.1 or later and NSX-T v2.3.

For more information, see [Upgrading vSphere in an NSX Environment](#) in the VMware documentation.

Feature Support by IaaS

	AWS	Azure	GCP	vSphere	vSphere with NSX-T
Automatic Kubernetes Cluster API load balancer					✓
HTTP proxy				✓	✓
Multi-AZ storage				✓	✓
Per-namespace subnets					✓
Service <code>type:LoadBalancer</code>	✓*	✓	✓		✓

* For more information about configuring Service `type:LoadBalancer` on AWS, see the [Access Workloads Using an Internal AWS Load Balancer](#) section of *Deploying and Exposing Basic Workloads*.

Upgrade Path

The supported upgrade paths to PKS v1.3.0 are from PKS v1.2.5 and later.

For more information, see [Upgrading PKS](#) and [Upgrading PKS with NSX-T](#).

 **Note:** Upgrading from PKS v1.2.5+ to PKS v1.3.x causes all certificates to be automatically regenerated. The old certificate authority is still trusted, and has a validity of one year. But the new certificates are signed with a new certificate authority, which is valid for four years.

What's New

PKS v1.3.0 adds the following:

- Support for PKS on Azure. For more information, see [Azure](#).
- BOSH Backup and Restore (BBR) for single-master clusters. For more information, see [Backing up the Single Master Cluster](#) and [Restoring the Single Master Cluster](#).
- Routable pods on NSX-T. For more information, see [Routable Pod Networks](#) in *Defining Network Profiles*.
- Large size NSX-T load balancers with Bare Metal NSX-T edge nodes. For more information, see [Hardware Requirements for PKS on vSphere with NSX-T](#).
- HTTP proxy for NSX-T components. For more information, see [Using Proxies with PKS on NSX-T](#).
- Ability to specify the size of the Pods IP Block subnet using a network profile. For more information, see [Pod Subnet Prefix](#) in *Defining Network Profiles*.
- Support for bootstrap security groups, custom floating IPs, and edge router selection using network profiles. For more information, see [Bootstrap Security Group](#), [Custom Floating IP Pool](#), and [Edge Router Selection](#) in *Defining Network Profiles*.
- Support for sink resources in air-gapped environments.
- Support for creating sink resources with the PKS Command Line Interface (PKS CLI). For more information, see [Creating Sink Resources](#).
- Sink resources include both pod logs as well as events from the Kubernetes API. These events are combined in a shared format that provides operators with a robust set of filtering and monitoring options. For more information, see [Monitoring PKS with Sinks](#).
- Support for multiple NSX-T Tier-0 (T0) logical routers for use with PKS multi-tenant environments. For more information, see [Configuring Multiple Tier-0 Routers for Tenant Isolation](#).

- Support for multiple PKS foundations on the same NSX-T. For more information, see [Implementing a Multi-Foundation PKS Deployment](#).
- Smoke tests errand that uses the PKS CLI to create a Kubernetes cluster and then delete it. If the creation or deletion fails, the errand fails and the installation of the PKS tile is aborted. For more information, see the *Errands* section of the *Installing PKS* topic for your IaaS, such as [Installing PKS on vSphere](#).
- Support for scaling down the number of worker nodes. For more information, see [Scaling Existing Clusters](#).
- Support for defining the CIDR range for Kubernetes pods and services on Flannel networks. For more information, see the *Networking* section of the *Installing PKS* topic for your IaaS, such as [Installing PKS on vSphere](#).
- Kubernetes v1.12.4.
- **Bug Fix:** The `No Proxy` property for vSphere now accepts wildcard domains like `*.example.com` and `example.com`. See [Networking](#) in *Installing PKS on vSphere* for more information.
- **Bug Fix:** The issue with NSX-T where special characters in username and password doesn't work is resolved.
- **Security Fix:** [CVE 2018-18264](#): This CVE allows unauthenticated secret access to the Kubernetes Dashboard.
- **Security Fix:** [CVE-2018-15759](#): This CVE contains an insecure method of verifying credentials. A remote unauthenticated malicious user may make many requests to the service broker with a series of different credentials, allowing them to infer valid credentials and gain access to perform broker operations.

Breaking Changes and Known Issues

 **Breaking Change:** Heapster is deprecated in PKS v1.3, and Kubernetes has retired Heapster. For more information, see the [kubernetes-retired/heapster](#) repository on GitHub.

PKS v1.3.0 has the following known issues:

Upgrades Fail When Clusters Share an External Hostname

If you use the same external hostname across more than one PKS-deployed Kubernetes cluster, upgrades from PKS v1.2.x to PKS v1.3.0 might fail. The external hostname is the value you set with either the `-e` or `--external-hostname` argument when you created the cluster. For more information, see [Create a Kubernetes Cluster](#).

PKS v1.3.0 introduces restrictions that prevent you from deploying clusters with duplicate hostnames, so this issue does not affect upgrades from PKS v1.3.0 and later.

If you have existing clusters that use the same external hostname, do not upgrade to PKS v1.3.x. Contact your Support representative for more information.

Upgrades Fail with a Hyphen in the No Proxy Field on vSphere

If you install PKS on vSphere and you enable the **HTTP/HTTPS Proxy** setting, you cannot use the `-` character in the **No Proxy** field. Entering `-` in the **No Proxy** field can cause validation errors when trying to upgrade to PKS v1.3.0. For more information, see the [Networking](#) section of *Installing PKS on vSphere*.

If you experience this issue during an upgrade, contact Support for a hotfix that will be applied in a future PKS v1.3.x release.

PKS Flannel Network Gets Out of Sync with Docker Bridge Network (cni0)

When VMs have been powered down for multiple days, turning them back on and issuing a `bosh recreate` to re-create the VMs causes the pods to get stuck in a `ContainerCreating` state.

Workaround: See [PKS Flannel network gets out of sync with docker bridge network \(cni0\)](#) in the Pivotal Knowledge Base.

Deploy Fails if vSphere Master Credentials Field Has Special Characters Without Quotes

If you install PKS on vSphere and you enter credentials in the **vCenter Master Credentials** field of the **Kubernetes Cloud Provider** pane of the PKS tile that contain special characters, such as `#`, `$`, `,`, `!`, or `-`, your deployment might fail with the following error:

ServerFaultCode: Cannot complete login due to an incorrect user name or password.

Workaround: If you install PKS on vSphere **without** NSX-T integration, place quotes around the credentials in the cloud provider configuration. For example, `"SomeP4$$w0rd#!"`. Then redeploy the PKS tile by clicking **Apply Changes**.

If you install PKS on vSphere **with** NSX-T integration, avoid using special characters in this field until this issue is resolved.

PKS Selects the First AZ Only During Cluster Creation

If the first availability zone (AZ) used by a plan with multiple AZs runs out of resources, cluster creation fails with an error like the following:

```
L Error: CPI error 'Bosh::Clouds::CloudError' with message 'No valid placement found for requested memory: 4096'
```

Explanation: BOSH creates VMs for your PKS deployment using a round-robin algorithm. It tries to create the first VM in the first AZ that your plan uses. If the first AZ runs out of resources, cluster creation fails and BOSH does not try to create more VMs.

Please send any feedback you have to pkss-feedback@pivotal.io.

Enterprise PKS Release Notes

Page last updated:

This topic contains release notes for Enterprise Pivotal Container Service (Enterprise PKS) v1.4.x.

v1.4.1

Release Date: June 20, 2019

Product Snapshot

Element	Details
Version	v1.4.1
Release date	June 20, 2019
Compatible Ops Manager versions	v2.4.2+, v2.5.x, v2.6.x
Stemcell version	v250.25 Note: To address the Zombieload CVE, upload stemcell v250.48.
Kubernetes version	v1.13.5
On-Demand Broker version	v0.26.0
NSX-T versions	v2.3.1, v2.4.0.1, (recommended) v2.4.1
NCP version	v2.4.1
Docker version	v18.06.3-ce docker-boshrelease ↗
BBR version	v1.8.0

A **warning:** NSX-T v2.4 implements a new password expiration policy. By default, the administrator password expires after 90 days. If the password expires, you cannot deploy new Kubernetes clusters using PKS with NSX-T. For more information, see the following [VMware KB article ↗](#) and [Updating the NSX Manager Password for NSX-T v2.4 ↗](#) in the Enterprise PKS documentation.

💡 Note: NSX-T v2.4 implements a new user interface (UI) based on the NSX Policy API. PKS v1.4.1 does not support the NSX Policy API. Any objects created through the new UI cannot be used with PKS v1.4.1. If you are installing PKS v1.4.1 with NSX-T v2.4.x or upgrading to PKS v1.4.1 and NSX-T 2.4.x, you must use the **Advanced Networking** tab in NSX Manager to create, read, update, and delete all networking objects required for Enterprise PKS. For more information, see [NSX-T v2.4 Management Interfaces ↗](#) in the Enterprise PKS documentation.

vSphere Version Requirements

If you are installing Enterprise PKS on vSphere or on vSphere with NSX-T Data Center, see the [VMware Product Interoperability Matrices ↗](#) for compatibility information.

Feature Support by IaaS

	AWS	Azure	GCP	vSphere	vSphere with NSX-T
Automatic Kubernetes Cluster API load balancer					✓
HTTP proxy				✓	✓
Multi-AZ storage				✓	✓
Per-namespace subnets					✓
Service <code>type:LoadBalancer</code>	✓*	✓	✓		✓

* For more information about configuring Service `type:LoadBalancer` on AWS, see the [Access Workloads Using an Internal AWS Load Balancer ↗](#) section of [Deploying and Exposing Basic Workloads](#).

Upgrade Path

The supported upgrade paths to PKS v1.4.1 are from PKS v1.4.0 and PKS v1.3.2 and later.

To upgrade, see [Upgrading Enterprise PKS](#) and [Upgrading Enterprise PKS with NSX-T](#).

What's New

Enterprise PKS v1.4.1 adds the following:

- Support for VMware NSX-T v2.4.1. For more information, see [VMware NSX-T Data Center 2.4.1 Release Notes](#).
- Support for VMware NCP v2.4.1. For more information, see [NSX Container Plugin 2.4 Release Notes](#) in the VMware documentation.
- Support for provisioning an NSX-T load balancer in front of the NSX management cluster. For more information, see [Provisioning a Load Balancer for the NSX-T v2.4 Management Cluster](#).
- Ability to tune NCP rate limits.
- Configurable Kubernetes service network CIDR range. For more information, see [Networking](#) in *Installing Enterprise PKS on vSphere with NSX-T*.
- Configurable node IP block. For more information, see [Defining Network Profiles](#).
- Improved functionality for the `/v1beta1/cluster` API endpoint to allow users to list and resize clusters that have compute profiles associated with them. For more information, see [List Clusters with Compute Profile](#) and [Resize a Cluster with a Compute Profile](#) in *Using Compute Profile (vSphere Only)*.

Breaking Changes and Known Issues

 **Breaking change:** Log sink entries in PKS v1.4 are tagged with the human-readable cluster name instead of the host ID of the BOSH-defined VM.

Enterprise PKS v1.4.1 has the following known issues:

Upgrade All Service Instances Errand Fails

Symptom

After clicking **Apply Changes** in Ops Manager, the `upgrade-all-service-instances` errand fails.

Explanation

The `upgrade-all-service-instances` errand fails if one or more of your existing clusters are in a **FAILED** state.

To check cluster status, run `pks clusters`.

Workaround

If you experience this issue, delete and re-create the failed cluster. Follow the procedure in [Cannot Re-Create a Cluster that Failed to Deploy](#).

v1.4.0

Release Date: April 25, 2019

Product Snapshot

Element	Details
Version	v1.4.0
Release date	April 25, 2019

Compatible Ops Manager versions	v2.4.2+, v2.5.x
Stemcell version	v250.25
Kubernetes version	v1.13.5
On-Demand Broker version	v0.26.0
NSX-T versions *	v2.3.1, v2.4.0.1
NCP version	v2.4.0
Docker version	v18.06.3-ce CFCR
BBR version	v1.8.0

WARNING: NSX-T v2.4 implements a new *password expiration policy*. By default the administrator password expires after 90 days. If the password expires, you cannot deploy new Kubernetes clusters using PKS with NSX-T. For more information, refer to the following VMware KB article:
<https://kb.vmware.com/s/article/70691>. See also [Updating the NSX Manager Password for NSX-T v2.4](#) in the Enterprise PKS documentation.

Note: NSX-T v2.4 implements a new user interface (UI) based on the NSX Policy API. PKS v1.4.0 does not support the NSX Policy API. Any objects created via the new Policy-based UI cannot be used with PKS v1.4.0. If you are installing PKS v1.4.0 with NSX-T v2.4.x, or upgrading to PKS v1.4.0 and NSX-T 2.4.x, you must use the “Advanced Networking” tab in NSX Manager to create, read, update, and delete all networking objects required for PKS. See also [NSX-T v2.4 Management Interfaces](#) in the Enterprise PKS documentation.

Note: PKS v1.4.0 on Azure is compatible only with Ops Manager v2.4.2 and v2.4.3. Before deploying PKS v1.4.0 on Azure, you must install Ops Manager v2.4.2 or v2.4.3.

vSphere Version Requirements

For Enterprise PKS installations on vSphere or on vSphere with NSX-T Data Center, refer to the [VMware Product Interoperability Matrices](#).

Feature Support by IaaS

	AWS	Azure	GCP	vSphere	vSphere with NSX-T
Automatic Kubernetes Cluster API load balancer					✓
HTTP proxy				✓	✓
Multi-AZ storage				✓	✓
Per-namespace subnets					✓
Service <code>type:LoadBalancer</code>	✓*	✓	✓		✓

* For more information about configuring Service `type:LoadBalancer` on AWS, see the [Access Workloads Using an Internal AWS Load Balancer](#) section of [Deploying and Exposing Basic Workloads](#).

Upgrade Path

The supported upgrade paths to PKS v1.4.0 are from PKS v1.3.2 and later.

To upgrade, see [Upgrading Enterprise PKS](#) and [Upgrading Enterprise PKS with NSX-T](#).

When upgrading to NSX-T 2.4:

- Use the official VMware NSX-T Data Center 2.4 build.
- Apply the NSX-T v2.4.0.1 hot-patch. For more information, see [KB article 67499](#) in the VMware Knowledge Base.
- To obtain the NSX-T v2.4.0.1 hot-patch, open a support ticket with VMware Global Support Services (GSS) for NSX-T Engineering.

What's New

Enterprise PKS v1.4.0 adds the following:

- Operators can configure up to ten sets of resource types, *orplans*, in the Enterprise PKS tile. All plans except the first can be made available or unavailable to developers deploying clusters. Plan 1 must be configured and made available as a default for developers. For more information, see the *Plans* section of the installation topic for your IaaS, such as [Installing Enterprise PKS on vSphere with NSX-T](#).
- Operators can deploy up to 5 master nodes per plan. For more information, see the *Plans* section of the installation topic for your IaaS, such as [Installing Enterprise PKS on vSphere with NSX-T](#).
- Operators can install PKS and Pivotal Application Service (PAS) on the same instance of Ops Manager.
- Improved workflow for managing cluster access. For more information, see [Grant Cluster Access](#) in *Managing Users in Enterprise PKS with UAA*.
- Operators can create webhook ClusterSink resources. A webhook ClusterSink resource batches logs into 1 second units, wraps the resulting payload in JSON, and uses the POST method to deliver the logs to the address of your log management service. For more information see, [Create a Webhook ClusterSink Resource with YAML and kubectl](#) in *Creating Sinks*.
- Operators can set quotas for maximum memory and CPU utilization in a PKS deployment. For more information, see [Managing Resource Usage](#). This is a beta feature.

⚠ warning: This feature is a beta component and is intended for evaluation and test purposes only. Do not use this feature in a production environment. Product support and future availability are not guaranteed for beta components.

- Operators can enable the PodSecurityPolicy admission plugin on a per-plan basis requiring cluster users to have policy, role, and role binding permissions to deploy pod workloads. See [Pod Security Policies](#) for more information.
- Operators can enable the SecurityContextDeny admission plugin on a per-plan basis to prohibit the use of security context configurations on pods and containers. See [Security Context Deny](#) for more information.
- Operators can enable the DenyEscalatingExec admission plugin on a per-plan basis to prohibit the use of certain commands for containers that allow host access. See [Deny Escalating Execution](#) for more information.
- Operators using vSphere can use HostGroups to define Availability Zones (AZs) for clusters in BOSH. See [Using vSphere Host Groups](#).
- Operators using vSphere can configure compute profiles to specify which vSphere resources are used when deploying Kubernetes clusters in a PKS deployment. For more information, see [Using Compute Profiles \(vSphere Only\)](#).

⚠ warning: This feature is a beta component and is intended for evaluation and test purposes only. Do not use this feature in a production environment. Product support and future availability are not guaranteed for beta components.

- Operators using vSphere with NSX-T can update a Network Profile and add to or reorder the Pods IP Block IDs. For more information, see the [Change the Network Profile for a Cluster](#) section of *Using Network Profiles (NSX-T Only)*.

Breaking Changes and Known Issues

⚠ Breaking change: Log sink entries in PKS v1.4 are tagged with the human-readable cluster name instead of the host ID of the BOSH-defined VM.

Enterprise PKS v1.4.0 has the following known issues:

Azure Default Security Group Is Not Automatically Assigned to Cluster VMs

Symptom

You experience issues when configuring a load balancer for a multi-master Kubernetes cluster or creating a service of type `LoadBalancer`. Additionally, in the Azure portal, the **VM > Networking** page does not display any inbound and outbound traffic rules for your cluster VMs.

Explanation

As part of configuring the Enterprise PKS tile for Azure, you enter **Default Security Group** in the **Kubernetes Cloud Provider** pane. When you create a Kubernetes cluster, Enterprise PKS automatically assigns this security group to each VM in the cluster. However, in Enterprise PKS v1.4, the automatic assignment may not occur.

As a result, your inbound and outbound traffic rules defined in the security group are not applied to the cluster VMs.

Workaround

If you experience this issue, manually assign the default security group to each VM NIC in your cluster.

Cluster Creation Fails When First AZ Runs out of Resources

Symptom

If the first availability zone (AZ) used by a plan with multiple AZs runs out of resources, cluster creation fails with an error like the following:

```
L Error: CPI error 'Bosh::Clouds::CloudError' with message 'No valid placement found for requested memory: 4096'
```

Explanation

BOSH creates VMs for your Enterprise PKS deployment using a round-robin algorithm, creating the first VM in the first AZ that your plan uses. If the AZ runs out of resources, cluster creation fails because BOSH cannot create the cluster VM.

For example, if your three AZs each have enough resources for ten VMs, and you create two clusters with four worker VMs each, BOSH creates VMs in the following AZs:

	AZ1	AZ2	AZ3
Cluster 1	Worker VM 1	Worker VM 2	Worker VM 3
	Worker VM 4		
Cluster 2	Worker VM 1	Worker VM 2	Worker VM 3
	Worker VM 4		

In this scenario, AZ1 has twice as many VMs as AZ2 or AZ3.

Azure Worker Node Communication Fails after Upgrade

Symptom

Outbound communication from a worker node VM fails after an upgrade to Enterprise PKS v1.4.0.

Explanation

Enterprise PKS 1.4.0 uses Azure Availability Sets to improve the uptime of workloads and worker nodes in the event of Azure platform failures. Worker node VMs are distributed evenly across Availability Sets.

Azure Standard SKU Load Balancers are recommended for the Kubernetes control plane and Kubernetes ingress and egress. This load balancer type provides an IP address for outbound communication using SNAT.

During an upgrade, when BOSH rebuilds a given worker instance in an Availability Set, Azure can time out while re-attaching the worker node network interface to the back-end pool of the Standard SKU Load Balancer.

For more information, see [Outbound connections in Azure](#) in the Azure documentation.

Workaround

You can manually re-attach the worker instance to the back-end pool of the Azure Standard SKU Load Balancer in your Azure console.

Please send any feedback you have to pkss-feedback@pivotal.io.

PCF Isolation Segment v2.6 Release Notes

Known Issues

Releases

2.6.0

- See [New Features in PCF Isolation Segment v2.6](#)
- [Feature] Add optional System Metric Agent to allow monitoring a more complete set of metrics for all VMs in the deployment
- [Feature] Improve scalability of application syslog drain system with new syslog agent architecture
- [Feature Improvement] Increase default number of CPUs for router from 1 to 2
- [Feature Improvement] Garden will now delegate container creates and destroys to containerd, an industry standard container runtime.
- [Feature Improvement] Update default polling interval and idle connection limits for networking components to reduce resource contention on PAS database
- [Security Fix] Introduce and trust new Diego “root CA” in advance of existing CA expiration
- Removed loggregator release

PCF Isolation Segment v2.6 includes the following component versions:

Component	Version
ubuntu-xenial stemcell	315.36
bpm	1.0.4
cf-networking	2.22.2
cflinuxfs3	0.101.0
diego	2.30.0
garden-runc	1.19.1
haproxy	9.5.2
loggregator-agent	3.14
mapfs	1.1.5
nfs-volume	2.1.0
routing	0.188.1
silk	2.22.2
smb-volume	1.1.0
syslog	11.4.0

About PCF Isolation Segment

The PCF Isolation Segment v2.6 tile is available for installation with PCF v2.6.

[Isolation segments](#) provide dedicated pools of resources where you can deploy apps and isolate workloads. Using isolation segments separates app resources as completely as if they were in different CF deployments but avoids redundant management and network complexity.

For more information about using isolation segments in your deployment, see the [Managing Isolation Segments](#) topic.

How to Install

The procedure for installing PCF Isolation Segment v2.6 is documented in the [Installing PCF Isolation Segment](#) topic.

To install a PCF Isolation Segment, you must first install PCF v2.6.

New Features in PCF Isolation Segment v2.6

See the following new features for PCF Isolation Segment v2.6:

Monitor System Metrics with System Metrics Agent

System Metrics Agent provides more visibility into VM compute, network, and storage metrics. These VM metrics help with troubleshooting and diagnosing issues for potential infrastructure problems.

When enabled, the metrics are emitted through Loggregator. For more information about viewing logs and metrics, see the *Platform components* row of [Viewing Logs and Metrics](#).

For a list of the VM metrics that the System Metric Agent emits, see [VM Metrics](#) in GitHub.

To enable the System Metric Agent, go to the **System Logging** pane in the PCF Isolation Segment tile and select **Enable System Metrics**.

Garden Delegates Container Creation and Destruction to containerd by Default

Traditionally, Garden uses runc to create, delete, and run container processes. Garden v1.15.0 and later support delegating the creation and destruction of some container processes through containerd instead of runc.

To disable containerd mode, deselect the **Enable Containerd Delegation** checkbox in the **Application Containers** pane of the PCF Isolation Segment tile.

Increased CPUs for Router VMs

To improve reliability, PAS v2.6 increases the default and minimum CPU core count from one to two for Router VMs.

PAS migrates Router VMs with a single core to a VM type with two CPU cores during the upgrade.

Loggregator Syslog Agent Increases Scale For Syslog Drains

The Loggregator architecture in PAS v2.6 includes optional Syslog Agents through the **Enable agent-based syslog egress for app logs** field in the **System Logging** configuration pane.

 **Note:** Enabling this feature disables the Syslog Adapter and Syslog Scheduler to avoid log duplication.

 **warning:** See the following known issue related to this feature: [App Syslog Drains Fail After Enabling Agent-Based Syslog](#).

Syslog Agents run on PCF component VMs and host VMs to manage connections with and write to syslog drains for app logs. The addition of Syslog Agents increases the number of syslog drain service bindings supported by the Loggregator system and reduces the workload for Loggregator VMs.

This update resolves a known issue where app log loss occurs at 10,000 syslog drain service bindings. For more information about the known issue, see the [Known Loggregator Scaling Issues](#) knowledge base article.

For more information about Syslog Agents, see [Loggregator Architecture](#).

About Advanced Features

The Advanced Features section of the PCF Isolation Segment v2.6 tile includes new functionality that may have certain constraints.

Although these features are fully supported, Pivotal recommends caution when using them in production.

Known Issues

There are no known issues for PCF Isolation Segment v2.6 at this time.

Stemcell Release Notes

This topic provides an overview of the Stemcell versions used by Pivotal Cloud Foundry (PCF).

Stemcells are available for download on Pivotal Network. Once downloaded, you can upload them to your deployment using the **Stemcell** configuration pane of a given tile or by navigating to the [Stemcell Library](#) page in Ops Manager v2.1 or later.

Before you upgrade PCF, always verify that any tiles that you have installed, either from Pivotal or a partner, are compatible with the version of PCF you are deploying and its supported stemcells. For more information on checking PCF and stemcell compatibility, see [Tile Compatibility](#) in the Upgrade Checklist.

To find out which stemcell version is used by which tile version, check the release notes or the Pivotal Network download page for the tile.

 **Note:** A stemcell is a versioned OS image that BOSH uses to create VMs for the BOSH Director, Pivotal Application Service (PAS), and other PCF Service tiles, such as MySQL for PCF. For more information, see [What is a Stemcell?](#) in the BOSH documentation.

Ubuntu Linux Stemcell Lines

This section lists Linux-based stemcells used in Pivotal Cloud Foundry.

Trusty Stemcells

Each stemcell line has a corresponding release notes section. See the links below:

- [Stemcell 3586.x Release Notes](#)
- [Stemcell 3541.x Release Notes](#)
- [Stemcell 3468.x Release Notes](#)
- [Stemcell 3445.x Release Notes](#)
- [Stemcell 3421.x Release Notes](#)
- [Stemcell 3363.x Release Notes](#)

Trusty End-of-Support

Ubuntu 14.04 LTS (Trusty) stemcells are reaching end-of-support in April 2019. Starting in September 2018, Pivotal and its partners will begin releasing product tiles for PCF that support Ubuntu 16.04 LTS (Xenial) stemcells instead. Using supported stemcells is necessary to avoid exposure to security vulnerabilities.

Xenial Stemcells

Each stemcell line has a corresponding release notes section. See the links below:

- [Stemcell 97.x Release Notes](#)
- [Stemcell 87.x Release Notes](#)
- [Stemcell 81.x Release Notes](#)
- [Stemcell 60.x Release Notes](#)
- [Stemcell 50.x Release Notes](#)
- [Stemcell 40.x Release Notes](#)

Add-on Support for Xenial Stemcells

If you are using any of the following PCF add-ons, you must update the add-on and its configuration to be compatible with the Ubuntu Xenial 16.04 stemcell before deploying any tiles that use Xenial. The following add-on versions support the Xenial stemcell:

- ClamAV Add-on for PCF v1.2.22 or later. For information on how to update this add-on, see [Updating ClamAV Add-on for PCF to Run with Xenial Stemcells ↗](#).
- File Integrity Monitoring (FIM) Add-on for PCF v1.4.28 or later. For information on how to update this add-on, see [Updating FIM Add-on for PCF to Run with Xenial Stemcells ↗](#).
- IPsec Add-on for PCF v1.9.9 or later. For information on how to update this add-on, see [Updating IPsec Add-on for PCF to Run with Xenial Stemcells ↗](#).

For information on which PCF tile releases now use Xenial, see [Tiles Using Xenial Stemcells in PCF](#).

Windows Stemcell Lines

This section lists Windows-based stemcells used in Pivotal Cloud Foundry. Each stemcell line has a corresponding release notes section. See the links below:

- [Stemcell 2019.x \(Windows Server version 2019\) Release Notes](#)
- [Stemcell 1803.x \(Windows Server version 1803\) Release Notes ↗](#)
- [Stemcell 1709.x \(Windows Server version 1709\) Release Notes ↗](#)
- [Stemcell 1200.x \(Windows2012R2\) Release Notes ↗](#)

Stemcell (Linux) Release Notes

This topic includes release notes for Linux stemcells used with Pivotal Cloud Foundry (PCF).

Xenial Stemcells

The following sections describe each Xenial stemcell release.

315.x

This section includes release notes for the 315 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

315.41

 Available in Pivotal Network

Release Date: June 17, 2019

CVE fixes for <https://usn.ubuntu.com/4017-1/> ↗

250.x

This section includes release notes for the 250 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

250.63

 Available in Pivotal Network

Release Date: June 17, 2019

CVE fixes for <https://usn.ubuntu.com/4017-1/> ↗

250.56

 Available in Pivotal Network

Release Date: May 22, 2019

USN 3977-2

250.29

 Available in Pivotal Network

Release Date: April 08, 2019

Periodic stemcell bump (Apr 9, 2019)

250.25

 Available in Pivotal Network

Release Date: March 25, 2019

Periodic stemcell bump (Mar 26, 2019)

250.23

 Available in Pivotal Network

Release Date: March 21, 2019

Periodic stemcell bump (Mar 22, 2019)

250.21

 Available in Pivotal Network

Release Date: March 12, 2019

Periodic stemcell bump (Mar 15, 2019)

250.17

 Available in Pivotal Network

Release Date: February 25, 2019

Periodic stemcell bump (Mar 06, 2019)

250.9

Release Date: February 12, 2019

Periodic Ubuntu Xenial stemcell bump (Feb 13, 2019)

250.4

Release Date: January 29, 2019

First published xenial 250. stemcell.

170.x

This section includes release notes for the 170 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

170.82

 Available in Pivotal Network

Release Date: June 17, 2019

CVE fixes for <https://usn.ubuntu.com/4017-1/> ↗

170.48

 Available in Pivotal Network

Release Date: April 08, 2019

Periodic stemcell bump (Apr 8, 2019)

170.45

 Available in Pivotal Network

Release Date: March 25, 2019

Periodic stemcell bump (Mar 26, 2019)

170.43

Release Date: March 21, 2019

Periodic stemcell bump (Mar 22, 2019)

170.39

 Available in Pivotal Network

Release Date: March 11, 2019

Periodic stemcell bump (Mar 15, 2019)

170.38

 Available in Pivotal Network

Release Date: February 25, 2019

Periodic stemcell bump (Mar 06, 2019)

170.30

 Available in Pivotal Network

Release Date: February 12, 2019

Periodic Ubuntu Xenial stemcell bump (Feb 12, 2019)

170.25

 Available in Pivotal Network

Release Date: January 28, 2019

Periodic Ubuntu Xenial stemcell bump (Jan 28, 2019)

170.24

 Available in Pivotal Network

Release Date: January 23, 2019

Addresses “USN-3866-1: Ghostscript vulnerability”

170.23

 Available in Pivotal Network

Release Date: January 22, 2019

Addresses “USN-3863-1: APT vulnerability”

170.21

Release Date: January 15, 2019

Periodic Ubuntu Xenial stemcell bump (Jan 15, 2019)

170.19

 Available in Pivotal Network

Release Date: January 11, 2019

Addresses “USN-3855-1: systemd vulnerabilities”

170.15

 Available in Pivotal Network

Release Date: December 20, 2018

Periodic Ubuntu Xenial stemcell bump (Dec 26, 2018)

170.14

 Available in Pivotal Network

Release Date: December 17, 2018

Periodic Ubuntu Xenial stemcell bump (Dec 17, 2018)

170.13

Release Date: December 11, 2018

Fixes

- *Google:* hostname should always be BOSH Agent ID ([#57](#), [#162225262](#))

170.12

Release Date: December 04, 2018

Periodic Ubuntu Xenial stemcell bump (Dec 05, 2018)

170.9

Release Date: November 19, 2018

Periodic Ubuntu Xenial stemcell bump (Nov 19, 2018)

170.6

Release Date: November 15, 2018

Includes updates to address:

- [USN-3820-2](#): Linux kernel (HWE) vulnerabilities

170.5

Release Date: November 05, 2018

Periodic Ubuntu Xenial stemcell bump (Nov 05, 2018)

97.x

This section includes release notes for the 97 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

97.113

Available in Pivotal Network

Release Date: June 17, 2019

CVE fixes for <https://usn.ubuntu.com/4017-1/>

97.74

Available in Pivotal Network

Release Date: April 08, 2019

Periodic stemcell bump (Apr 8, 2019)

97.71

Available in Pivotal Network

Release Date: March 25, 2019

Periodic stemcell bump (Mar 26, 2019)

97.67

Available in Pivotal Network

Release Date: March 15, 2019

Periodic stemcell bump (Mar 22, 2019)

97.66

Available in Pivotal Network

Release Date: March 11, 2019

Periodic stemcell bump (Mar 15, 2019)

97.65

 Available in Pivotal Network

Release Date: February 25, 2019

Periodic stemcell bump (Mar 06, 2019)

97.57

 Available in Pivotal Network

Release Date: February 12, 2019

Periodic Ubuntu Xenial stemcell bump (Feb 13, 2019)

97.53

 Available in Pivotal Network

Release Date: January 28, 2019

Periodic Ubuntu Xenial stemcell bump (Jan 28, 2019)

97.52

 Available in Pivotal Network

Release Date: January 23, 2019

Addresses “USN-3866-1: Ghostscript vulnerability”

97.51

 Available in Pivotal Network

Release Date: January 22, 2019

Addresses “USN-3863-1: APT vulnerability”

97.49

Release Date: January 15, 2019

Periodic Ubuntu Xenial stemcell bump (Jan 15, 2019)

97.47

 Available in Pivotal Network

Release Date: January 11, 2019

Addresses “USN-3855-1: systemd vulnerabilities”

97.43

 Available in Pivotal Network

Release Date: December 20, 2018

Periodic Ubuntu Xenial stemcell bump (Dec 26, 2018)

97.42

 Available in Pivotal Network

Release Date: December 17, 2018

Periodic Ubuntu Xenial stemcell bump (Dec 17, 2018)

97.41

 Available in Pivotal Network

Release Date: December 11, 2018

- *Google*: hostname should always be BOSH Agent ID ([#57](#), [#162225262](#))
- Unpin rsyslog (was v8.22; backported from 170.x; [#162514665](#))
- Periodic Ubuntu Xenial updates

97.39

 Available in Pivotal Network

Release Date: December 04, 2018

Periodic Ubuntu Xenial stemcell bump (Dec 05, 2018)

97.34

 Available in Pivotal Network

Release Date: November 19, 2018

Periodic Ubuntu Xenial stemcell bump (Nov 19, 2018)

97.33

 Available in Pivotal Network

Release Date: November 15, 2018

Includes updates to address:

- [USN-3820-2](#): Linux kernel (HWE) vulnerabilities

97.32

 Available in Pivotal Network

Release Date: November 05, 2018

Periodic Ubuntu Xenial stemcell bump (Nov 08, 2018)

This stemcell addresses a bug introduced in the 97.31 release. The bug impacts AWS light stemcell users only.

97.31

 Available in Pivotal Network

Release Date: November 05, 2018

Do not use - this release has a bug that impacts the light stemcells Periodic Ubuntu Xenial stemcell bump (Nov 05, 2018)

97.22

Release Date: October 04, 2018

Addresses “USN-3777-2: Linux kernel (HWE) vulnerabilities” (Oct 04, 2018)

97.19

 Available in Pivotal Network

Release Date: October 02, 2018

(Oct 02, 2018)

97.18

 Available in Pivotal Network

Release Date: September 24, 2018

Periodic Ubuntu Xenial stemcell bump (Sep 25, 2018)

97.17

 Available in Pivotal Network

Release Date: September 18, 2018

Fixes mounting persistent disk issue with bosh-agent. (Sep 19, 2018)

97.16

 Available in Pivotal Network

Release Date: September 10, 2018

Periodic Ubuntu Xenial stemcell bump (Sep 11, 2018)

97.15

 Available in Pivotal Network

Release Date: August 27, 2018

Bump Ubuntu Xenial stemcells for “USN-3756-1: Intel Microcode vulnerabilities”

97.12

 Available in Pivotal Network

Release Date: August 14, 2018

Bump Ubuntu Xenial stemcells for “USN-3740-2: Linux kernel (HWE) vulnerabilities”

97.10

Release Date: August 13, 2018

Periodic Ubuntu Xenial stemcell bump (Aug 14, 2018)

97.5

Release Date: August 08, 2018

Bump Ubuntu Xenial stemcells for “USN-3732-2: Linux kernel (HWE) vulnerability”

97.3

Release Date: July 30, 2018

Periodic Ubuntu Xenial stemcell bump (July 31, 2018)

87.x

This section includes release notes for the 87 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

87.4

Release Date: July 16, 2018

Periodic Ubuntu Xenial stemcell bump (July 16, 2018)

87.3

Release Date: July 11, 2018

Periodic Ubuntu Xenial stemcell bump (July 12, 2018)

87

Release Date: July 02, 2018

Periodic Ubuntu Xenial stemcell bump (July 2, 2018)

81.x

This section includes release notes for the 81 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

81

Release Date: June 19, 2018

- Periodic Ubuntu Xenial stemcell bump (June 18, 2018)

60.x

This section includes release notes for the 60 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

60

Release Date: June 04, 2018

- Periodic Ubuntu Xenial stemcell bump (June 4, 2018)

50.x

This section includes release notes for the 50 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

50

Release Date: May 23, 2018

- Light stemcells are available

40.x

This section includes release notes for the 40 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

40

Release Date: May 22, 2018

- First release of Ubuntu Xenial stemcells
- Notable differences from Ubuntu Trusty
 - Includes systemd instead of upstart
 - Includes 4.15 Linux Kernel instead of 4.4
 - Uses chronyd to sync time (runs as a daemon) instead of ntpdate
 - Does not include NFS utils by default

Warning: Do not *downgrade* instances from Ubuntu Xenial to Ubuntu Trusty stemcells as it may corrupt persistent disk content since Trusty stemcells may decide to use sfdisk partitioner instead of parted partitioner selected by Xenial stemcells.

7.x

This section includes release notes for the 7 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

7

Release Date: June 03, 2019

Periodic stemcell bump (Jun 4, 2019)

7

Release Date: April 08, 2019

Periodic stemcell bump (Apr 8, 2019)

7

Release Date: February 14, 2019

Periodic CentOS 7 stemcell bump (Feb 14, 2019)

7

Release Date: January 28, 2019

Periodic CentOS 7 stemcell bump (Jan 28, 2019)

7

Release Date: January 16, 2019

Periodic CentOS 7 stemcell bump (Jan 17, 2019)

Trusty Stemcells

The following sections describe each Trusty stemcell release.

3763.x

This section includes release notes for the 3763 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3763.14

Release Date: March 11, 2019

Periodic stemcell bump (Mar 14, 2019)

3763.13

Release Date: February 25, 2019

Periodic stemcell bump (Mar 08, 2019)

3586.x

This section includes release notes for the 3586 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3586.93

Release Date: March 25, 2019

Periodic stemcell bump (Mar 26, 2019)

3586.91

Release Date: March 12, 2019

Periodic stemcell bump (Mar 15, 2019)

3586.86

Release Date: February 25, 2019

Periodic stemcell bump (Mar 06, 2019)

3586.79

Release Date: February 12, 2019

Periodic Ubuntu Trusty stemcell bump (Feb 14, 2019)

3586.71

Release Date: January 28, 2019

Periodic Ubuntu Trusty stemcell bump (Jan 28, 2019)

3586.70

Release Date: January 23, 2019

Addresses “USN-3866-1: Ghostscript vulnerability” and “USN-3863-1: APT vulnerability”

3586.67

Release Date: January 14, 2019

Periodic Ubuntu Trusty stemcell bump (Jan 15, 2019)

3586.65

Release Date: December 21, 2018

Periodic Ubuntu Trusty stemcell bump (Dec 21, 2018)

3586.63

Release Date: December 19, 2018

Periodic Ubuntu Trusty stemcell bump (Dec 19, 2018)

3586.60

Release Date: December 03, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Dec 05, 2018)

3586.57

Release Date: November 19, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Nov 19, 2018)

3586.56

Release Date: November 15, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Nov 15, 2018)

3586.54

Release Date: November 05, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Nov 05, 2018)

3586.52

Release Date: October 22, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Oct 23, 2018)

3586.48

Release Date: October 08, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Oct 11, 2018)

3586.46

Release Date: October 02, 2018

Addresses “USN-3776-2: Linux kernel (Xenial HWE) vulnerabilities” (Oct 02, 2018)

3586.43

Release Date: September 24, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Sep 25, 2018)

3586.42

Release Date: September 10, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Sep 11, 2018)

3586.40

Release Date: August 27, 2018

Bump Ubuntu Trusty stemcells for “USN-3756-1: Intel Microcode vulnerabilities”

Known Issue

- On GCP, writing moderate amounts of data to a persistent disk and then migrating the disk will fail with:

Error: Timed out sending 'update_settings'. See #159511884 ↗

3586.36

Release Date: August 14, 2018

Bump Ubuntu Trusty stemcells for “USN-3741-2: Linux kernel (Xenial HWE) vulnerabilities”

3586.35

Release Date: August 13, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Aug 14, 2018)

3586.27

Release Date: July 30, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (July 31, 2018)

3586.26

Release Date: July 16, 2018

Periodic Ubuntu Trusty stemcell bump (July 16, 2018)

3586.25

Release Date: July 02, 2018

- Periodic Ubuntu Trusty stemcell bump (July 2, 2018)

3586.24

Release Date: June 18, 2018

- Periodic Ubuntu Trusty stemcell bump (June 18, 2018)

3586.23

Release Date: June 13, 2018

- We are continuing to investigate GCP stemcell compatibility issue from earlier version, but we did roll back BOSH Agent to an earlier version that seems to not trigger this problem
 - Note: This build does not include fixes to recently published CVE-2018-3665

3586.18

Release Date: June 04, 2018

WARNING: We are currently investigating [unresponsive agent](#) issues when using the Google Cloud Platform version of this stemcell. In the meantime, please use [3586.16](#) when deploying to GCP.

- Periodic Ubuntu Trusty stemcell bump (June 4, 2018)

3586.16

Release Date: May 24, 2018

- Bump Ubuntu Trusty stemcells for “USN-3654-2: Linux kernel (Xenial HWE) vulnerabilities”

TLS for Internal Blobstore Supported

For Ops Manager v2.2 and later, you can enable TLS for your internal blobstore. Make sure you configured all tiles with a stemcell v3586 or later before enabling TLS for your internal blobstore.

For more information, see [TLS for Internal Blobstore Supported](#) in the Ops Manager release notes.

3586.8

Release Date: May 21, 2018

- Periodic Ubuntu Trusty stemcell bump (May 21, 2018)

3586.7

Release Date: May 09, 2018

- Bump Ubuntu Trusty stemcells for “USN-3641-1: Linux kernel vulnerabilities”

3586.5

Release Date: May 08, 2018

- Bump s3cli to include AliCloud support
- Bump bosh-agent
 - Support network aliases (used by Softlayer CPI)
 - Support static routes for networks (used by Softlayer CPI)
 - Support iSCSI for persistent disks (used by Softlayer CPI)
 - Use parted when GPT partitions are detected
 - Refactor retryable strategy usages

3541.x

This section includes release notes for the 3541 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3541.65

Release Date: December 07, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Dec 07, 2018)

3541.64

Release Date: December 03, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Dec 05, 2018)

3541.61

Release Date: November 19, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Nov 19, 2018)

3541.60

Release Date: November 15, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Nov 15, 2018)

3541.57

Release Date: October 22, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Oct 23, 2018)

3541.54

Release Date: October 08, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Oct 11, 2018)

3541.52

Release Date: October 02, 2018

Addresses “USN-3776-2: Linux kernel (Xenial HWE) vulnerabilities” (Oct 02, 2018)

3541.49

Release Date: September 24, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Sep 25, 2018)

3541.48

Release Date: September 10, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Sep 11, 2018)

3541.46

Release Date: August 27, 2018

Bump Ubuntu Trusty stemcells for “USN-3756-1: Intel Microcode vulnerabilities”

3541.44

Release Date: August 14, 2018

Bump Ubuntu Trusty stemcells for “USN-3741-2: Linux kernel (Xenial HWE) vulnerabilities”

3541.43

Release Date: August 13, 2018

Periodic Ubuntu Trusty/CentOS stemcell bump (Aug 14, 2018)

3541.37

Release Date: July 30, 2018

Periodic Ubuntu Trusty stemcell bump (July 31, 2018)

3541.36

Release Date: July 16, 2018

Periodic Ubuntu Trusty stemcell bump (July 16, 2018)

3541.35

Release Date: July 02, 2018

- Periodic Ubuntu Trusty stemcell bump (July 2, 2018)

3541.34

Release Date: June 18, 2018

- Periodic Ubuntu Trusty stemcell bump (June 18, 2018)

3541.31

Release Date: June 04, 2018

- Periodic Ubuntu Trusty stemcell bump (June 4, 2018)

3541.25

Release Date: May 09, 2018

- Bump Ubuntu Trusty stemcells for “USN-3641-1: Linux kernel vulnerabilities”

3541.30

Release Date: May 23, 2018

- Bump Ubuntu Trusty stemcells for “USN-3654-2: Linux kernel (Xenial HWE) vulnerabilities”

3541.26

Release Date: May 21, 2018

- Periodic Ubuntu Trusty stemcell bump (May 21, 2018)

3541.24

Release Date: May 07, 2018

- Ubuntu Trusty stemcells periodic update (May 7, 2018)

3541.12

Release Date: April 06, 2018

- Bump Ubuntu Trusty stemcells for USN-3619-2: Linux kernel (Xenial HWE) vulnerabilities

3541.10

Release Date: March 26, 2018

- Periodic Ubuntu and CentOS stemcell bump (March 26/27, 2018)

3541.9

Release Date: March 12, 2018

- Periodic Ubuntu and CentOS stemcell bump (March 12, 2018)

3541.8

Release Date: March 08, 2018

- Bump bosh-agent to 2.67.1
 - Agent will now respect previously set permissions and owner on sys/run, sys/log and data job directories
 - This should fix stemcell compatibility with Diego/Garden if Agent restarts
 - If you were using 3541.x stemcell for any of your deployments, it's recommended to update your deployments to this version before updating Director since that would cause Agent restart

3541.5

Release Date: February 22, 2018

- Bump Ubuntu Trusty stemcells for USN-3582-2: Linux kernel (Xenial HWE) vulnerabilities

3541.4

Release Date: February 14, 2018

- Rolled back custom umask configuration as we found out it was different in some cases (depends on how processes were started)
 - Hardening of /var/vcap/jobs/* is still applied by the agent

3541.2

Release Date: February 08, 2018

- [breaking] Set default umask to 077 and further harden several /var/vcap/* directories
 - Note that you may have to change your release to adapt to this change
- [breaking] Renamed /var/vcap/bosh/bin/ntpdate to /var/vcap/bosh/bin/sync-time
- [breaking] Stop forwarding SSH events to bosh-agent
 - Agent no longer receives and forwards such events to HM. This should remove a lot of noisy generated by releases that expect a lot of SSH sessions (eg Gitlab). This information will continue to be available in logs forwarded to remote destinations (and locally /var/log/auth.log).
- Fixes `env.bosh.swap_size: 0` to work on more clouds (including GCP)

Misc

- Order stemcell tarballs so that upload-stemcell command can execute faster
- Generate `packages.txt` within stemcell tarball that includes list of installed packages (previously known under different name)

3469.x

This section includes release notes for the 3469 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3469.1

Release Date: February 27, 2018

- Stemcell produced for testing rsyslog bump to the latest version
 - Unless testing rsyslog, use 3468.x or 3541.x stemcell lines

3468.x

This section includes release notes for the 3468 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3468.78

Release Date: October 22, 2018

Periodic Ubuntu Trusty stemcell bump (Oct 23, 2018)

3468.75

Release Date: October 08, 2018

Periodic Ubuntu Trusty stemcell bump (Oct 11, 2018)

3468.73

Release Date: October 02, 2018

Addresses “USN-3776-2: Linux kernel (Xenial HWE) vulnerabilities” (Oct 02, 2018)

3468.71

Release Date: September 24, 2018

Periodic Ubuntu Trusty stemcell bump (Sep 25, 2018)

3468.69

Release Date: September 10, 2018

Periodic Ubuntu Trusty stemcell bump (Sep 11, 2018)

3468.67

Release Date: August 27, 2018

Bump Ubuntu Trusty stemcells for “USN-3756-1: Intel Microcode vulnerabilities”

3468.64

Release Date: August 14, 2018

Bump Ubuntu Trusty stemcells for “USN-3741-2: Linux kernel (Xenial HWE) vulnerabilities”

3468.63

Release Date: August 13, 2018

Periodic Ubuntu Trusty stemcell bump (Aug 14, 2018)

3468.55

Release Date: July 30, 2018

Periodic Ubuntu Trusty stemcell bump (July 31, 2018)

3468.54

Release Date: July 16, 2018

Periodic Ubuntu Trusty stemcell bump (July 16, 2018)

3468.53

Release Date: July 02, 2018

- Periodic Ubuntu Trusty stemcell bump (July 2, 2018)

3468.51

Release Date: June 18, 2018

- Periodic Ubuntu Trusty stemcell bump (June 18, 2018)

3468.47

Release Date: June 04, 2018

- Periodic Ubuntu Trusty stemcell bump (June 4, 2018)

3468.42

Release Date: May 09, 2018

- Bump Ubuntu Trusty stemcells for “USN-3641-1: Linux kernel vulnerabilities”

3468.46

Release Date: May 23, 2018

- Bump Ubuntu Trusty stemcells for “USN-3654-2: Linux kernel (Xenial HWE) vulnerabilities”

3468.44

Release Date: May 21, 2018

- Periodic Ubuntu Trusty stemcell bump (May 21, 2018)

3468.41

Release Date: May 07, 2018

- Ubuntu Trusty stemcells periodic update (May 7, 2018)

3468.30

Release Date: April 06, 2018

- Bump Ubuntu Trusty stemcells for USN-3619-2: Linux kernel (Xenial HWE) vulnerabilities

3468.28

Release Date: March 26, 2018

- Periodic Ubuntu and CentOS stemcell bump (March 26/27, 2018)

3468.27

Release Date: March 12, 2018

- Periodic Ubuntu and CentOS stemcell bump (March 12, 2018)

3468.26

Release Date: March 01, 2018

- Includes updated `ixgbevf 4.3.4`

3468.25

Release Date: February 22, 2018

- Bump Ubuntu Trusty stemcells for USN-3582-2: Linux kernel (Xenial HWE) vulnerabilities

3468.22

Release Date: February 05, 2018

- [Feb 5] Periodic stemcell bump

3468.21

Release Date: January 23, 2018

- No functional change from 3468.20, except version number

3468.20

Release Date: January 23, 2018

- Bump Ubuntu Trusty stemcells for [USN-3540-2: Linux kernel \(Xenial HWE\) vulnerabilities](#) (This flaw is known as Spectre.)

3468.19

Release Date: January 17, 2018

- Bump Ubuntu Trusty stemcells for USN-3534-1: GNU C Library vulnerabilities

3468.17

Release Date: January 10, 2018

- Bump Ubuntu Trusty stemcells for [USN-3522-4: Linux kernel \(Xenial HWE\) regression](#)
- Configure vSphere stemcells to use VM hardware version 9 (requires ESXi 5.1) to be compatible with <https://www.vmware.com/us/security/advisories/VMSA-2018-0004.html>

3468.16

Release Date: January 10, 2018

- Bump Ubuntu Trusty stemcells for [USN-3522-2: Linux \(Xenial HWE\) vulnerability](#) (This flaw is known as Meltdown.)

3468.15

Release Date: December 15, 2017

- Bump Ubuntu Trusty stemcells for USN-3509-4: Linux kernel (Xenial HWE) regression

3468.13

Release Date: December 08, 2017

- Bump Ubuntu Trusty stemcell USN-3509-2: Linux kernel (Xenial HWE) vulnerabilities

3468.12

Release Date: December 06, 2017

- Bump Ubuntu Trusty stemcells for USN-3505-1: Linux firmware vulnerabilities

3468.11

Release Date: November 21, 2017

- Periodic Ubuntu stemcells update
- Includes Agent changes to support IPv6 on vSphere (manual networking)

3468.5

Release Date: October 26, 2017

- Configure `/tmp` to have sticky bit set

3468.1

Release Date: October 23, 2017

- Periodic stemcell bump

3468

Release Date: October 05, 2017

- Removed password authentication for Warden stemcells
- Various minor tweaks that were already backported to older lines

Upcoming features on this stemcell line:

- IPv6 support for vSphere

3445.x

This section includes release notes for the 3445 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3445.76

Release Date: October 22, 2018

Periodic Ubuntu Trusty stemcell bump (Oct 23, 2018)

3445.73

Release Date: October 08, 2018

Periodic Ubuntu Trusty stemcell bump (Oct 11, 2018)

3445.71

Release Date: October 02, 2018

Addresses “USN-3776-2: Linux kernel (Xenial HWE) vulnerabilities” (Oct 02, 2018)

3445.68

Release Date: September 24, 2018

Periodic Ubuntu Trusty stemcell bump (Sep 25, 2018)

3445.67

Release Date: September 10, 2018

Periodic Ubuntu Trusty stemcell bump (Sep 10, 2018)

3445.66

Release Date: August 27, 2018

Bump Ubuntu Trusty stemcells for “USN-3756-1: Intel Microcode vulnerabilities”

3445.64

Release Date: August 14, 2018

Bump Ubuntu Trusty stemcells for “USN-3741-2: Linux kernel (Xenial HWE) vulnerabilities”

3445.63

Release Date: August 13, 2018

Periodic Ubuntu Trusty stemcell bump (Aug 14, 2018)

3445.55

Release Date: July 30, 2018

Periodic Ubuntu Trusty stemcell bump (July 31, 2018)

3445.54

Release Date: July 16, 2018

Periodic Ubuntu Trusty stemcell bump (July 16, 2018)

3445.53

Release Date: July 02, 2018

- Periodic Ubuntu Trusty stemcell bump (July 2, 2018)

3445.51

Release Date: June 18, 2018

- Periodic Ubuntu Trusty stemcell bump (June 18, 2018)

3445.49

Release Date: June 04, 2018

- Periodic Ubuntu Trusty stemcell bump (June 4, 2018)

3445.48

Release Date: May 23, 2018

- Bump Ubuntu Trusty stemcells for “USN-3654-2: Linux kernel (Xenial HWE) vulnerabilities”

3445.42

Release Date: May 01, 2018

- Ubuntu Trusty stemcells periodic update (Apr 30, 2018)

3445.46

Release Date: May 21, 2018

- Periodic Ubuntu Trusty stemcell bump (May 21, 2018)

3445.45

Release Date: May 09, 2018

- Bump Ubuntu Trusty stemcells for “USN-3641-1: Linux kernel vulnerabilities”

3445.44

Release Date: May 07, 2018

- Ubuntu Trusty stemcells periodic update (May 7, 2018)

3445.32

Release Date: April 06, 2018

- Bump Ubuntu Trusty stemcells for USN-3619-2: Linux kernel (Xenial HWE) vulnerabilities

3445.30

Release Date: March 26, 2018

- Periodic Ubuntu Trusty stemcell bump (March 26/27, 2018)

3445.29

Release Date: March 12, 2018

- Periodic Ubuntu Trusty stemcell bump (March 12, 2018)

3445.28

Release Date: February 22, 2018

- Bump Ubuntu Trusty stemcells for USN-3582-2: Linux kernel (Xenial HWE) vulnerabilities

3445.25

Release Date: February 05, 2018

- [Feb 5] Periodic stemcell bump

3445.24

Release Date: January 23, 2018

- Bump Ubuntu Trusty stemcells for [USN-3540-2: Linux kernel \(Xenial HWE\) vulnerabilities](#) (This flaw is known as Spectre.)

3445.23

Release Date: January 17, 2018

- Bump Ubuntu Trusty stemcells for USN-3534-1: GNU C Library vulnerabilities

3445.22

Release Date: January 10, 2018

- Bump Ubuntu Trusty stemcells for [USN-3522-4: Linux kernel \(Xenial HWE\) regression ↗](#)
- Configure vSphere stemcells to use VM hardware version 9 (requires ESXi 5.1) to be compatible with [https://www.vmware.com/us/security/advisories/VMSA-2018-0004.html ↗](https://www.vmware.com/us/security/advisories/VMSA-2018-0004.html)

3445.21

Release Date: January 10, 2018

- Bump Ubuntu Trusty stemcells for [USN-3522-2: Linux \(Xenial HWE\) vulnerability ↗](#) (This flaw is known as Meltdown.)

3445.19

Release Date: December 08, 2017

- Bump Ubuntu Trusty stemcell USN-3509-2: Linux kernel (Xenial HWE) vulnerabilities

3445.18

Release Date: December 06, 2017

- Bump Ubuntu Trusty stemcells for [USN-3505-1: Linux firmware vulnerabilities](#)

3445.17

Release Date: November 21, 2017

- Periodic Ubuntu stemcells update

3445.11

Release Date: September 19, 2017

- Bump Ubuntu stemcells for [USN-3420-2: Linux kernel \(Xenial HWE\) vulnerabilities](#)

3445.7

Release Date: August 30, 2017

- Logrotate /var/log/wtmp and utmp more aggressively
- Updated BOSH agent to include aggressive 5 minute timeout on NATS connection failure
- Set auditd rules to be mutable by default
 - Please use [auditd](#) job from os-conf-release to make rules immutable

3445.2

Release Date: August 16, 2017

- Bump Ubuntu stemcells for [USN-3392-2: Linux kernel \(Xenial HWE\) regression ↗](#)

3445

Release Date: August 11, 2017

- Bump Ubuntu stemcells for [USN-3385-2: Linux kernel \(Xenial HWE\) vulnerabilities](#)
- Include gcscli for native GCS support
- Bump bosh-agent to include support for colocated errand jobs
- Fix occasional rsyslog hang on startup (as a workaround for <https://github.com/rsyslog/rsyslog/issues/1188>)

3431.x

This section includes release notes for the 3431 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3431.13

Release Date: August 03, 2017

- Bump version (no change)

3431.11

Release Date: August 03, 2017

- Bump Ubuntu Trusty stemcells for [USN-3378-2: Linux kernel \(Xenial HWE\) vulnerabilities](#)
- Fix occasional rsyslog hang on startup
 - Workaround for <https://github.com/rsyslog/rsyslog/issues/1188>

3431.10

Release Date: July 31, 2017

- Periodic Ubuntu stemcells update

3422.x

This section includes release notes for the 3422 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3422.7

Release Date: November 22, 2017

- Test stemcell for umask changes (based on 3468.x stemcell line)

3421.x

This section includes release notes for the 3421 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3421.88

Release Date: October 08, 2018

Periodic Ubuntu Trusty stemcell bump (Oct 11, 2018)

3421.86

Release Date: October 02, 2018

Addresses “USN-3776-2: Linux kernel (Xenial HWE) vulnerabilities” (Oct 02, 2018)

3421.83

Release Date: September 24, 2018

Periodic Ubuntu Trusty stemcell bump (Sep 25, 2018)

3421.82

Release Date: September 10, 2018

Periodic Ubuntu Trusty stemcell bump (Sep 10, 2018)

3421.81

Release Date: August 27, 2018

Bump Ubuntu Trusty stemcells for “USN-3756-1: Intel Microcode vulnerabilities”

3421.79

Release Date: August 14, 2018

Bump Ubuntu Trusty stemcells for “USN-3741-2: Linux kernel (Xenial HWE) vulnerabilities”

3421.78

Release Date: August 13, 2018

Periodic Ubuntu Trusty stemcell bump (Aug 14, 2018)

3421.70

Release Date: July 30, 2018

Periodic Ubuntu Trusty stemcell bump (July 31, 2018)

3421.69

Release Date: July 16, 2018

Periodic Ubuntu Trusty stemcell bump (July 16, 2018)

3421.68

Release Date: July 02, 2018

- Periodic Ubuntu Trusty stemcell bump (July 2, 2018)

3421.66

Release Date: June 18, 2018

- Periodic Ubuntu Trusty stemcell bump (June 18, 2018)

3421.64

Release Date: June 04, 2018

- Periodic Ubuntu Trusty stemcell bump (June 4, 2018)

3421.63

Release Date: May 23, 2018

- Bump Ubuntu Trusty stemcells for “USN-3654-2: Linux kernel (Xenial HWE) vulnerabilities”

3421.56

Release Date: May 01, 2018

- Ubuntu Trusty stemcells periodic update (Apr 30, 2018)

3421.60

Release Date: May 21, 2018

- Periodic Ubuntu Trusty stemcell bump (May 21, 2018)

3421.59

Release Date: May 09, 2018

- Bump Ubuntu Trusty stemcells for “USN-3641-1: Linux kernel vulnerabilities”

3421.58

Release Date: May 07, 2018

- Ubuntu Trusty stemcells periodic update (May 7, 2018)

3421.46

Release Date: April 06, 2018

- Bump Ubuntu Trusty stemcells for USN-3619-2: Linux kernel (Xenial HWE) vulnerabilities

3421.44

Release Date: March 26, 2018

- Periodic Ubuntu Trusty stemcell bump (March 26/27, 2018)

3421.43

Release Date: March 12, 2018

- Periodic Ubuntu Trusty stemcell bump (March 12, 2018)

3421.42

Release Date: February 22, 2018

- Bump Ubuntu Trusty stemcells for [USN-3582-2: Linux kernel \(Xenial HWE\) vulnerabilities](#) (This flaw is known as Meltdown.)

3421.39

Release Date: February 05, 2018

- [Feb 5] Periodic stemcell bump

3421.38

Release Date: January 23, 2018

- Bump Ubuntu Trusty stemcells for [USN-3540-2: Linux kernel \(Xenial HWE\) vulnerabilities](#) (This flaw is known as Spectre.)

3421.37

Release Date: January 17, 2018

- Bump Ubuntu Trusty stemcells for [USN-3534-1: GNU C Library vulnerabilities](#)

3421.36

Release Date: January 10, 2018

- Bump Ubuntu Trusty stemcells for [USN-3522-4: Linux kernel \(Xenial HWE\) regression](#)
- Configure vSphere stemcells to use VM hardware version 9 (requires ESXi 5.1) to be compatible with <https://www.vmware.com/us/security/advisories/MSA-2018-0004.html>

3421.35

Release Date: January 10, 2018

- Bump Ubuntu Trusty stemcells for [USN-3522-2: Linux \(Xenial HWE\) vulnerability](#) (This flaw is known as Meltdown.)

3421.34

Release Date: December 08, 2017

- Bump Ubuntu Trusty stemcell USN-3509-2: Linux kernel (Xenial HWE) vulnerabilities

3421.33

Release Date: December 06, 2017

- Bump Ubuntu Trusty stemcells for [USN-3505-1: Linux firmware vulnerabilities](#)

3421.32

Release Date: November 21, 2017

- Periodic Ubuntu stemcells update

3421.20

Release Date: August 16, 2017

- Bump Ubuntu stemcells for [USN-3392-2: Linux kernel \(Xenial HWE\) regression ↗](#)

3421.19

Release Date: August 11, 2017

- Bump Ubuntu stemcells for [USN-3385-2: Linux kernel \(Xenial HWE\) vulnerabilities ↗](#)

3421.18

Release Date: August 03, 2017

- Bump Ubuntu Trusty stemcells for [USN-3378-2: Linux kernel \(Xenial HWE\) vulnerabilities ↗](#)
- Fix occasional rsyslog hang on startup
 - Workaround for [https://github.com/rsyslog/rsyslog/issues/1188 ↗](https://github.com/rsyslog/rsyslog/issues/1188)

3421.11

Release Date: June 29, 2017

- Bump Ubuntu stemcells for USN-3344-2: Linux kernel (Xenial HWE) vulnerabilities

3421.9

Release Date: June 21, 2017

- Bump Ubuntu stemcells for USN-3334-1: Linux kernel (Xenial HWE) vulnerabilities

3421.6

Release Date: June 09, 2017

- Bump Ubuntu stemcells for USN-3312-2 - Linux kernel vulnerabilities

3421.4

Release Date: June 01, 2017

- Bump CentOS stemcells for CESA-2017:1382 - sudo vulnerability

3421.3

Release Date: May 30, 2017

- Bump Ubuntu stemcells for USN-3304-1: Sudo vulnerability

3421

Release Date: May 22, 2017

New:

- Added `env.bosh.remove_static_libraries` (bool) to remove static libraries
 - Useful to enable this option when exporting compiled releases
- Added `env.bosh.ipv6.enable` (bool) to remove ipv6.disable kernel functionality at bootup time

Fixes:

- Fixed sysstat logging
- Fixed anacron's RANDOM_DELAY configuration

Bumps:

- Bumped s3cli v0.0.60
 - Updated aws-sdk-go to solve network timeout edge case
- Bumped davcli v0.0.19
 - Use TCP keep alive to solve network timeout edge case
- Bumped bosh-agent v0.0.35
 - Add `-v` to the Agent binary
 - Prepared `sync_dns` action to work with future Director's DNS integration

3363.x

This section includes release notes for the 3363 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3363.65

Release Date: June 18, 2018

- Periodic Ubuntu Trusty stemcell bump (July 18, 2018)

3363.64

Release Date: June 04, 2018

- Periodic Ubuntu Trusty stemcell bump (June 4, 2018)

3363.63

Release Date: May 23, 2018

- Bump Ubuntu Trusty stemcells for "USN-3654-2: Linux kernel (Xenial HWE) vulnerabilities"

3363.62

Release Date: May 21, 2018

- Periodic Ubuntu Trusty stemcell bump (May 21, 2018)

3363.61

Release Date: May 09, 2018

- Bump Ubuntu Trusty stemcells for “USN-3641-1: Linux kernel vulnerabilities”

3363.60

Release Date: May 07, 2018

- Ubuntu Trusty stemcells periodic update (May 7, 2018)

3363.53

Release Date: April 06, 2018

- Bump Ubuntu Trusty stemcells for USN-3619-2: Linux kernel (Xenial HWE) vulnerabilities

3363.52

Release Date: March 26, 2018

- Periodic Ubuntu Trusty stemcell bump (March 26/27, 2018)

3363.51

Release Date: March 12, 2018

- Periodic Ubuntu Trusty stemcell bump (March 12, 2018)

3363.50

Release Date: February 22, 2018

- Bump Ubuntu Trusty stemcells for USN-3582-2: Linux kernel (Xenial HWE) vulnerabilities

3363.49

Release Date: February 05, 2018

- [Feb 5] Periodic stemcell bump

3363.48

Release Date: January 23, 2018

- Bump Ubuntu Trusty stemcells for [USN-3540-2: Linux kernel \(Xenial HWE\) vulnerabilities](#) (This flaw is known as Spectre.)

3363.47

Release Date: January 17, 2018

- Bump Ubuntu Trusty stemcells for USN-3534-1: GNU C Library vulnerabilities

3363.46

Release Date: January 10, 2018

- Bump Ubuntu Trusty stemcells for [USN-3522-4: Linux kernel \(Xenial HWE\) regression](#)
- Configure vSphere stemcells to use VM hardware version 9 (requires ESXi 5.1) to be compatible with <https://www.vmware.com/us/security/advisories/VMSA-2018-0004.html>

3363.45

Release Date: January 10, 2018

- Bump Ubuntu Trusty stemcells for [USN-3522-2: Linux \(Xenial HWE\) vulnerability](#) (This flaw is known as Meltdown.)

3363.44

Release Date: December 08, 2017

- Bump Ubuntu Trusty stemcell USN-3509-2: Linux kernel (Xenial HWE) vulnerabilities

3363.43

Release Date: December 06, 2017

- Bump Ubuntu Trusty stemcells for [USN-3505-1: Linux firmware vulnerabilities](#)

3363.42

Release Date: November 21, 2017

- Periodic Ubuntu stemcells update

3363.24

Release Date: May 17, 2017

- Periodic Ubuntu stemcells update

3363.22

Release Date: May 11, 2017

- Periodic Ubuntu stemcells update
- Run `cron` in BOSH Lite stemcells so that logrotation is performed

3363.20

Release Date: April 25, 2017

- Bump Ubuntu stemcells for [USN-3265-2: Linux kernel \(Xenial HWE\) vulnerabilities](#)

3363.19

Release Date: April 17, 2017

- Periodic bump for CentOS stemcells to include CESA-2017:0933
- Disable IPv6 through /proc/cmdline to eliminate possibility of listening on tcp6/udp6

3363.15

Release Date: April 05, 2017

- Bump Ubuntu stemcells for USN-3256-2: Linux kernel (HWE) vulnerability

Misc:

- Made AWS AMI backing snapshot public to support encryption of boot disks

3363.14

Release Date: March 30, 2017

- Bump Ubuntu stemcells for USN-3249-2: Linux kernel (Xenial HWE) vulnerability

3363.10

Release Date: March 08, 2017

- Bumps Ubuntu stemcells for USN-3220-2: Linux kernel (Xenial HWE) vulnerability

3363.9

Release Date: February 22, 2017

Changes: - Bumps Ubuntu stemcells for USN-3208-2: Linux kernel (Xenial HWE) vulnerabilities - Fixes excessive “out of memory” errors in kernel - <https://bugs.launchpad.net/ubuntu/+source/linux/+bug/1655842> - Fixes regression to rsyslog by locking it down again to rsyslog 8.22.0

Agent: - Fixes Azure stemcell persistent disk formatting - Fixes Warden stemcells SSH access

3363.1

Release Date: February 15, 2017

Reported Problems: - DO NOT USE azure stemcell as it may cause data loss. - rsyslog version updated to 8.24.0, regressing on issue #1537 - [Out of memory errors still exists in Kernel 4.4.0.62](#) - will be fixed around Feb 20.

Changes: - Fixes double -hvm- suffix problem for AWS Light stemcells

3363

Release Date: February 15, 2017

Reported Problems: - DO NOT USE azure stemcell as it may cause data loss. - [Out of memory errors still exists in Kernel 4.4.0.62](#) - will be fixed around Feb 20. - rsyslog version updated to 8.24.0, regressing on issue #1537 - AWS Light stemcell has incorrect name once imported - BOSH SSH does not work on BOSH Lite

Changes: - Add more auditd rules - Fix CentOS initramfs to load necessary kernel modules - Disable boot loader login - Increasing tcp_max_sync_backlog - Disabling any DSA host keys - Add `bosh_ssers` group and assign it to vcap user - Only allow users in `bosh_ssers` group to SSH

Agent: - Log Agent API access events in CEF format to syslog (vcap.agent topic) - Allow configuring swap size through `env.bosh.swap_size` (example: `env.bosh.swap_size: 0`) - Prepare for SHA2 releases - Allow setting fetching to work with base64 encoded user data - Do not delaycompress in logrotate

3312.x

This section includes release notes for the 3312 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3312.51

Release Date: January 23, 2018

- Bump Ubuntu Trusty stemcells for [USN-3540-2: Linux kernel \(Xenial HWE\) vulnerabilities](#) (This flaw is known as Spectre.)

3312.50

Release Date: January 10, 2018

- Bump Ubuntu Trusty stemcells for [USN-3522-4: Linux kernel \(Xenial HWE\) regression](#)
- Configure vSphere stemcells to use VM hardware version 9 (requires ESXi 5.1) to be compatible with <https://www.vmware.com/us/security/advisories/VMSA-2018-0004.html>

3312.49

Release Date: January 09, 2018

- Bump Ubuntu Trusty stemcells for [USN-3522-2: Linux \(Xenial HWE\) vulnerability](#) (This flaw is known as Meltdown.)

3312.48

Release Date: December 08, 2017

- Bump Ubuntu Trusty stemcell USN-3509-2: Linux kernel (Xenial HWE) vulnerabilities

3312.47

Release Date: December 06, 2017

- Bump Ubuntu Trusty stemcells for USN-3505-1: Linux firmware vulnerabilities

3312.46

Release Date: November 21, 2017

- Periodic Ubuntu stemcells update

3312.17

Release Date: January 31, 2017

Reported Problems: - [Memory leak bug in Kernel 4.4.0-59](#)

Changes: - Periodic stemcell update

3312.15

Release Date: January 12, 2017

Reported Problems: - [Memory leak bug in Kernel 4.4.0-59](#)

Changes: - Periodic stemcell update

3312.12

Release Date: December 20, 2016

- Lock down rsyslog to 8.22.0 to avoid high memory usage issue [https://github.com/cloudfoundry/bosh/issues/1537 ↗](https://github.com/cloudfoundry/bosh/issues/1537)

3312.9

Release Date: December 15, 2016

- Updates Azure Ubuntu stemcells to use parted partitioner to fix slow disk partitioning
 - Related issue: [https://github.com/cloudfoundry-incubator/bosh-azure-cpi-release/issues/227 ↗](https://github.com/cloudfoundry-incubator/bosh-azure-cpi-release/issues/227)

3312.8

Release Date: December 14, 2016

- Bumps Ubuntu stemcells for USN-3156-1: APT vulnerability

3312.7

Release Date: December 06, 2016

- Bump Ubuntu stemcells for [USN-3151-2: Linux kernel \(Xenial HWE\) vulnerability ↗](#)

3312.6

Release Date: December 02, 2016

- Periodic stemcell update

3312.3

Release Date: November 30, 2016

- Periodic stemcell update
 - Includes [USN-3134-1 ↗](#) as requested by a community member

3312

Release Date: November 16, 2016

- Properly includes `libpam_cracklib.so` to avoid errors in `/var/log/auth.log`

3309.x

This section includes release notes for the 3309 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3309

Release Date: November 10, 2016

- Fixes persistent disk mounting on OpenStack described in [Stemcell 3308 ↗](#)

3308.x

This section includes release notes for the 3308 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3308

Release Date: November 09, 2016

Reported Problems: - On OpenStack: Mounting persistent disks not working when using `config-drive: disk` while nova is configured to use a `cdrom` config-drive due to <https://github.com/cloudfoundry/bosh/issues/1503>

Fixes: - Fixes SSH key installation issue introduced in [Stemcell 3306](#)

3306.x

This section includes release notes for the 3306 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3306

Release Date: November 08, 2016

Reported Problems - `bosh-init` doesn't work with this stemcell on OpenStack and AWS due to <https://github.com/cloudfoundry/bosh/issues/1500> - Booting the stemcell image directly in your IaaS (without using BOSH/bosh-init) does no longer provision the ssh key for user `vcap`, so you need to login differently

Changes - Agent will now wait for monit to complete stop all processes before carrying on - Added google stemcells - Default `dmesg_restrict` to 1 - Disable all IPv6 configurations - Reenabled UDF kernel module for Azure - Increase `root_maxkeys` and `maxkeys` kernel configurations - Changed default hostname to `bosh-stemcell` instead of `localhost` to avoid boot problems on GCP - Lower TCP keepalive configuration by default - Mount `/var/log` directory to `/var/vcap/data/root_log` - Restrict Access to the `su` command - Add `pam_cracklib` requirements to `common-password` and `password-auth` - Enable auditing for processes that start prior to `auditd` - Set log rotation interval to 15 min in stemcell - Made ownership & permissions for `/etc/cron*` files more restrictive - Customize shell prompt to show instance name and ID - Removed floppy drives from vSphere stemcells - Removed `bosh` assets hence

making `bosh` unsupported
making `micro`

Misc: - Stemcells are now built through Concourse via <https://main.bosh-ci.cf-app.com/teams/main/pipelines/bosh:stemcells>

3263.x

This section includes release notes for the 3263 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3263.10

Release Date: November 03, 2016

- Updates CentOS kernel to the latest version for “Dirty COW”
 - Ubuntu stemcells were updated in previous versions at the time of Ubuntu USN updates
- Includes fix to the bosh-agent to better support 1TB+ disk partitioning

3263.8

Release Date: October 21, 2016

- Bump Ubuntu stemcells for USN-3106-2: Linux kernel (Xenial HWE) vulnerability
- Includes a fix to the bosh-agent to work more reliably with 2TB+ persistent disks

3263.7

Release Date: October 12, 2016

- Bump Ubuntu stemcells for USN-3099-2: Linux kernel (Xenial HWE) vulnerabilities

3263.5

Release Date: September 30, 2016

- Periodic bump
- Delay start of rsyslogd using systemd on CentOS

3263.4

Release Date: September 28, 2016

- google-kvm: improve the `google-*` daemon configurations
 - fixes `ssh: handshake failed` errors on boot

3263.3

Release Date: September 26, 2016

- Bumps Ubuntu stemcells for [USN-3087-2](#) (OpenSSL regression)

3263

Release Date: September 19, 2016

- Bumps Ubuntu to Linux kernel to 4.4

Based on 3262 stemcells. Note: OpenStack stemcells series 3263 is broken due to <https://github.com/cloudfoundry/bosh-agent/issues/98> and should not be used

3262.x

This section includes release notes for the 3262 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3262.21

Release Date: October 13, 2016

- Bump Ubuntu stemcells for USN-3099-2: Linux kernel (Xenial HWE) vulnerabilities

3262.19

Release Date: September 28, 2016

- google-kvm: improve the `google-*` daemon configurations
 - fixes `ssh: handshake failed` errors on boot

3262.16

Release Date: September 26, 2016

- Bumps Ubuntu stemcells for [USN-3087-2](#) (OpenSSL regression)

3262.15

Release Date: September 23, 2016

- Bumps Ubuntu stemcells for USN-3087-1: OpenSSL vulnerabilities

3233.x

This section includes release notes for the 3233 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3233.1

Release Date: September 27, 2016

- Bumps Ubuntu stemcells for [USN-3087-2](#) (OpenSSL regression)

USN-3522-2 Addresses Meltdown Vulnerabilities

Meltdown exploits critical vulnerabilities in modern processors. For more information about Meltdown, see the [Meltdown and Spectre Attacks](#) blog post. [USN-3522-2](#) addresses the critical vulnerability in Ubuntu associated with Meltdown.

This update may include degradations to performance if your VM's CPU and memory usage are currently at near-capacity levels. Prior to upgrading to this stemcell, monitor your PCF VM's current CPU and memory usage and scale those components if necessary. If any of your VMs are currently operating at 60% or above, Pivotal recommends scaling that VM. For more information about the performance impact of Meltdown-related stemcell patches on PCF components and guidance on scaling, see this [KB article](#).

For more information about monitoring and scaling PCF, see the [Monitoring PCF VMs from Ops Manager](#), [Key Capacity Scaling Indicators](#), and [Scaling PAS](#) topics. Performance degradation is likely to vary by workload type, IaaS, and other factors. Pivotal recommends testing your deployment thoroughly after upgrading to this stemcell.

3232.x

This section includes release notes for the 3232 line of Linux stemcells used with Pivotal Cloud Foundry (PCF).

3232.21

Release Date: September 26, 2016

- Bumps Ubuntu stemcells for [USN-3087-2](#) (OpenSSL regression)

Stemcell v2019.x (Windows Server version 2019) Release Notes

This topic includes release notes for Windows stemcells used with Pivotal Application Service for Windows (PASW) v2.5 and v2.6.

The stemcell is based on Windows Server, version 2019.

To download a stemcell, see [Stemcells for PCF \(Windows\)](#) on Pivotal Network.

2019.6

Release Date: June 19, 2019

Security Fix

- Includes Microsoft Security Updates [June 11, 2019—KB4503327](#)
- Introduces 2.3.1.2 (L1) and 1.1.1 (L1) CIS L1 policy hardenings based on the CIS Security Benchmark.

2019.5

Release Date: May 30, 2019

Security Fix

- Based on [Microsoft's guidance](#), additional fixes to protect against speculative execution side-channel vulnerabilities

2019.4

Release Date: May 22, 2019

Features

- Platform Engineers can deploy Windows Stemcells on a BOSH Director with Google Cloud Storage as their external Blobstore.
- Improved Troubleshooting of Windows VMs, with ssh enabled by default for all Windows VMs. You can still disable SSH in the PASW tile.
- Includes Microsoft Security Updates to protect against Microarchitectural Data Sampling side-channel vulnerabilities. For more information, see [May 14, 2019—KB4494441 \(OS Build 17763.503\)](#) in the Windows support documentation.

2019.3

Release Date: April 25, 2019

Features

- This is the first 2019 stemcell.
- Includes [Microsoft Security Updates Patch Tuesday April 2019](#)

Tiles Using Xenial Stemcells in PCF

This table provides a quick reference of PCF product versions that are now using Xenial stemcells.

Products Using Xenial Stemcells

If a product is not yet listed, the product still uses Trusty stemcells.

Product/Component	Starting Version	Release Notes Link
PCF Operations Manager	v2.3	Release Notes ↗
Pivotal Application Service (PAS)	v2.3	Release Notes ↗
Pivotal Container Service (PKS)	v1.2	Release Notes ↗
PCF Event Alerts	v1.2.3	Release Notes ↗
Metrics Forwarder for PCF	v1.11.3	Release Notes ↗
MySQL for PCF	v2.4	Release Notes ↗
RabbitMQ for PCF	v1.14.1	Release Notes ↗
Redis for PCF	v1.14.0	Release Notes ↗
Scheduler for PCF	v1.2.3	Release Notes ↗
Single Sign-On (SSO) for PCF	v1.7.1	Release Notes ↗
Spring Cloud Data Flow	v1.2.0	Release Notes ↗
Spring Cloud Services for PCF	v2.0.2	Release Notes ↗

Xenial Stemcell Support in Ops Manager

You can import and use Xenial stemcell-based tiles in the following versions of Ops Manager:

- v2.1.15 and later
- v2.2.2 and later
- v2.3.x and later

For more information about the introduction of Xenial stemcells to your PCF deployment, see [Updates for Xenial Stemcell Support ↗](#) in the *Pivotal Cloud Foundry v2.3 Breaking Changes Release Notes*.

Platform Architecture and Planning

This documentation describes reference architectures and other plans for installing Pivotal Cloud Foundry (PCF) on any infrastructure to support the runtime environments Pivotal Application Service (PAS) and Pivotal Container Service (PKS).

Overview

A Pivotal Cloud Foundry (PCF) reference architecture describes a proven approach for deploying PCF on a specific IaaS, such as AWS, Azure, GCP, OpenStack, or vSphere. PCF reference architectures meet the following requirements:

- Secure
- Publicly-accessible
- Includes common PCF-managed services such as MySQL, RabbitMQ, and Spring Cloud Services
- Can host at least 100 app instances, or far more
- Have been deployed and validated by Pivotal to support Ops Manager, PAS, and PKS

You can use PCF reference architectures to help plan the best configuration for your PCF deployment on your IaaS.

Reference Architecture and Planning Topics

All PCF reference architectures start with the [Base PAS Architecture](#) and [Base PKS Architecture](#), below.

The following infrastructure-specific topics build on these two common base architectures:

- [AWS Reference Architecture](#)
- [Azure Reference Architecture](#)
- [GCP Reference Architecture](#)
- [OpenStack Reference Architecture](#)
- [vSphere Reference Architecture](#)

The following topic describes a broader architecture that automates the installation and updating of multiple PCF foundations, multiple instances of each architecture above:

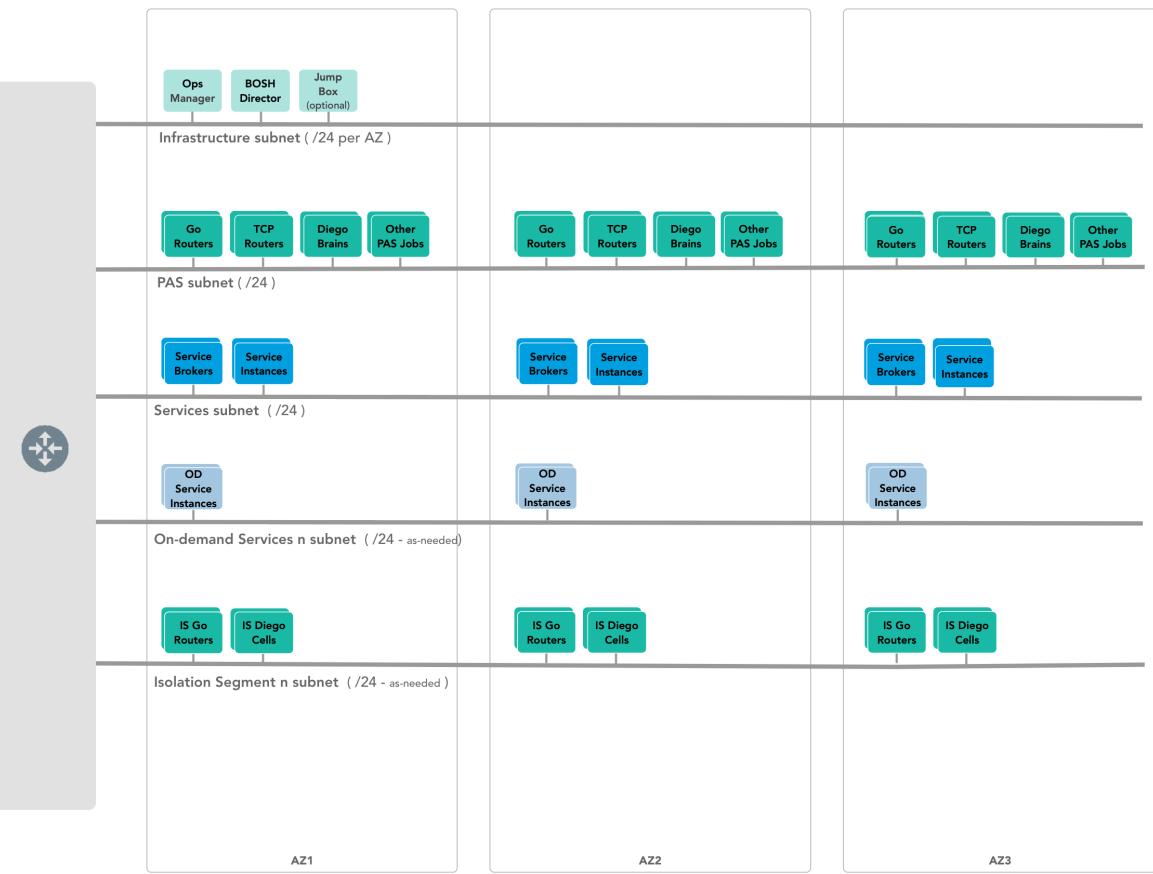
- [Control Plane Reference Architectures](#)

The following topics address aspects of platform architecture and planning that the PCF reference architectures do not cover:

- [Implementing a Multi-Foundation PKS Deployment](#)
- [Using Global DNS Load Balancers for Multi-Foundation](#)

PAS Architecture

The following diagram shows a base architecture for PAS, and how its network topology places and replicates PCF and PAS components across subnets and Availability Zones.



[View a larger version of this diagram](#)

Internal Components

The following table describes the internal component placements shown in the diagram above:

Component	Placement and Access Notes
Ops Manager	Deployed on one of the three public subnets. Accessible by FQDN or through an optional jumpbox.
BOSH Director	Deployed on the infrastructure subnet.
Jumpbox	Optional. Deployed on the infrastructure subnet for accessing PAS management components such as Ops Manager and the Cloud Foundry command-line interface (cf CLI).
Gorouters (HTTP routers in CF)	Deployed on all three PAS subnets, one per Availability Zone (AZ). Accessed through the HTTP, HTTPS, and SSL load balancers.
Diego Brain	Deployed on all three PAS subnets, one per AZ. The Diego Brain component is required, but SSH container access support through the Diego Brain is optional, and enabled through the SSH load balancers.
TCP Routers	Optional. Deployed on all three PAS subnets, one per AZ, to support TCP routing.
Service Tiles	Service brokers and shared services instances are deployed to the Services subnet. Dedicated on-demand service instances are deployed to an On-demand services subnet.
Isolation Segments	Deployed on an Isolation Segment subnet. Includes Diego Cells and Gorouters for running and accessing apps hosted within isolation segments.

Networks

Pivotal recommends defining your networks and load-balancing their incoming requests as follows:

Required Subnets

PAS requires the following statically-defined networks to host its main component systems.

- Infrastructure subnet - `/24` segment This subnet contains VMs that require access only for Platform Administrators, for example Ops Manager, the BOSH Director, and an optional jumpbox.
- PAS subnet - `/24` segment This subnet contains PAS runtime VMs, such as Gorouters, Diego Cells and Cloud Controllers.
- Services subnet - `/24` segment The Services and On-Demand Services networks support Ops Manager tiles that you might add in addition to PAS. You can think of them as the “everything not PAS” networks. Some services tiles can call for additional network capacity to grow into on-demand. If you use services with this capability, Pivotal recommends that you add an On-Demand Services network for each on-demand service, as described below.
- On-Demand Services subnet(s) - `/24` segments This is for services that can allocate network capacity on-demand from BOSH for their worker VMs; see the Services subnet above. Pivotal recommends allocating a dedicated subnet to each on-demand service. For example, you can configure the Redis tile as follows:
 - **Network:** enter the existing `Services` network, to host the service broker
 - **Services Network:** deploy a new network `OD-Services1`, to host the Redis worker VMs

Another on-demand service tile can then also use `Services` for its broker and a new `OD-Services2` network for its workers, and so on.

- Isolation Segments subnet(s) - `/24` segments You can add one or more Isolation Segment tiles to an PAS installation to compartmentalize hosting and routing resources. For each Isolation Segment you deploy, you should designate a `/24` network for its range of address space.

Load Balancing

Any PAS installation needs a suitable load balancer to send incoming HTTP, HTTPS, SSH, and SSL traffic to its Gorouters and application containers. All installations approaching production-level use rely on external load balancing from hardware appliance vendors or other network-layer solutions.

The load balancer can also perform Layer 4 or Layer 7 load balancing functions. SSL can be terminated at the load balancer or used as a pass-through to the Gorouter.

Common deployments of load balancing in PAS are:

- HTTP/HTTPS traffic to and from Gorouters
- TCP traffic to and from TCP routers
- Traffic from the Diego Brain, when developers access app containers via SSH

To balance load across multiple PAS foundations, use an IaaS- or vendor-specific Global Traffic Manager or Global DNS load balancer.

For additional information, see [Global DNS Load Balancers for Multi-Foundation Environments](#).

High Availability

PAS is not considered High Availability (HA) until it runs across at least two Availability Zones (AZs). Pivotal recommends defining three AZs.

On IaaSes with their own HA capabilities, using the IaaS HA in conjunction with a PCF HA topology provides the best of both worlds. Multiple AZs give PCF redundancy, so that losing an AZ is not catastrophic. The BOSH Resurrector can then replace lost VMs as needed to repair a foundation.

To back up and restore a foundation externally, use BOSH Backup and Restore (BBR). For more information, see the [BOSH Backup and Restore](#) documentation.

Storage

PAS requires disk storage for each component, for both persistent data and to allocate to ephemeral data. You size these disks in the [Resource Config](#) pane of the PAS tile. For details on storage configuration and capacity planning, see the corresponding chapter for the specific IaaS.

The platform also requires you to configure file storage for large shared objects. These blobstores can be external or internal. For details, see [Configuring](#)

[File Storage for PAS](#).

Security

See the topics below for how PAS implements security:

- [PCF Infrastructure Security](#)
- [Security Concepts](#)
- [Certificates and TLS in PCF](#)
- [Network Communication Paths in PCF](#) - includes firewall ports

Domain Names

PAS requires following domain names to be registered:

- System domain, for PAS and other tiles: `sys.domain.name`
- App domain, for your applications: `app.domain.name`

You must also define the following wildcard domain names and include them when creating certificates that access PAS and its hosted apps:

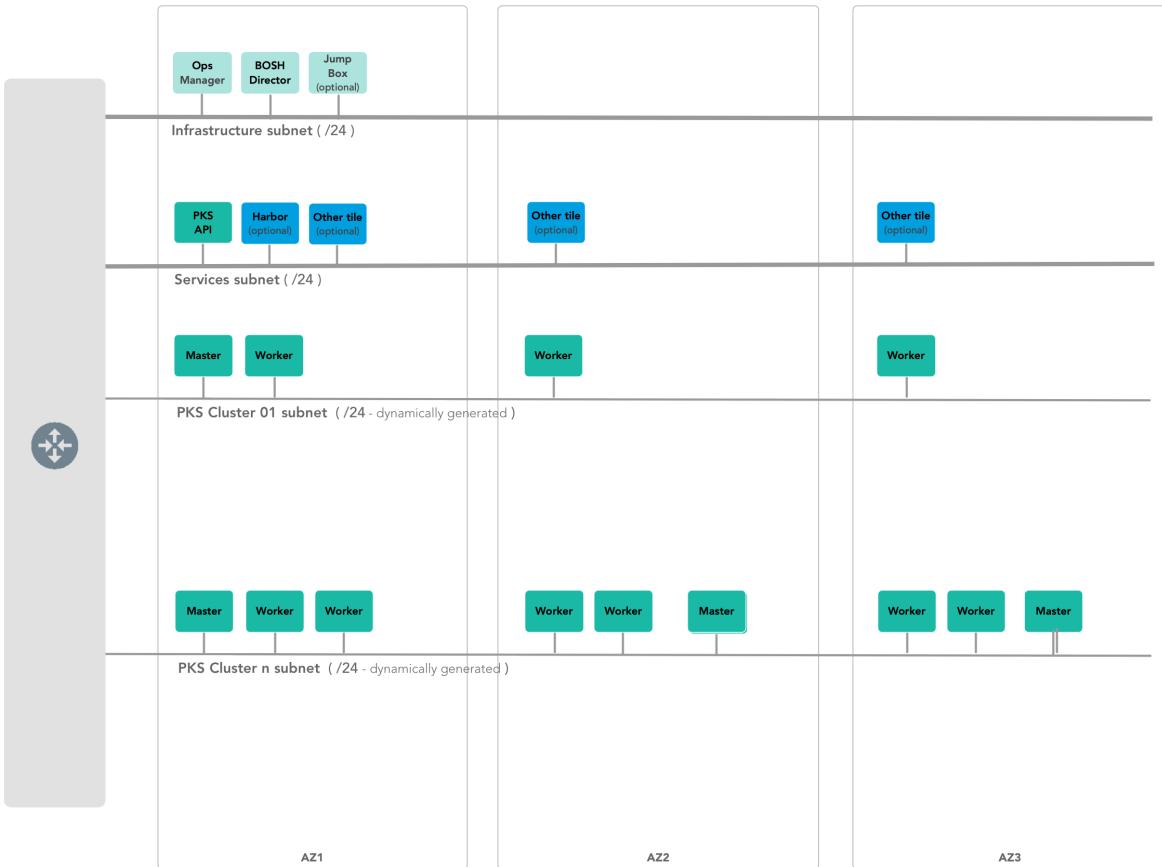
- *.SYSTEM-DOMAIN
- *.APPS-DOMAIN
- *.login.SYSTEM-DOMAIN
- *.uaa.SYSTEM-DOMAIN

Component Scaling

See [Scaling PAS](#) for recommendations on how to scale PAS for different deployment scenarios.

PKS Architecture

The following diagram shows a base architecture for PKS, and how its network topology places and replicates PCF and PKS components across subnets and Availability Zones.



[View a larger version of this diagram](#)

Internal Components

The following table describes the internal component placements shown in the diagram above:

Component	Placement and Access Notes
Ops Manager	Deployed on one of the subnets. Accessible by fully-qualified domain name (FQDN) or through an optional jumpbox.
BOSH Director	Deployed on the infrastructure subnet.
Jumpbox	Optional. Deployed on the infrastructure subnet for accessing PKS management components such as Ops Manager and the PKS + Kubectl command-line interface.
PKS API	Deployed as a service broker VM on the PKS Services subnet. Handles PKS API and service adapter requests, and manages PKS clusters. For more information, see Pivotal Container Service (PKS)
Harbor Tile	Optional container images registry, typically deployed to the Services subnet.
PKS Cluster	Deployed to a dynamically-created, dedicated PKS cluster subnet. Each cluster consists of worker nodes that run the workloads (applications) and one or more master nodes.

Networks

Pivotal recommends defining your networks and load-balancing their incoming requests as follows:

Subnets Requirements

PKS requires two defined networks to host the main elements that compose it.

- Infrastructure subnet - `/24` Contains VMs that require access only for Platform Administrators, for example Ops Manager, BOSH Director and an optional jumpbox.
- PKS Services subnet - `/24` Hosts PKS API VM and other optional service tiles such as Harbor.
- PKS Clusters subnets - each one a `/24` from a pool of pre-allocated IPs Hosts PKS clusters.

Load Balancing

Load balancers can be used to manage traffic across master nodes of a PKS cluster or for deployed workloads. For more information on how to configure load balancers for PKS, refer to the corresponding documentation for the specific IaaS.

High Availability

PKS has no inherent HA capabilities to design for. Make the best efforts to have HA design at the IaaS, storage, power and access layers to support PKS.

Storage

PKS requires shared storage across all Availability Zones for the deployed workloads to appropriately allocate their required storage.

Security

See the topics below for how PKS implements security:

- [Enterprise PKS Security](#)
- [Firewall Ports](#)

Domain Names

PKS requires the following domain names to be registered when creating a wildcard certificate and PKS tile configurations: `*.pks.domain.name`

The wildcard certificate covers both the PKS API domain e.g. `api.pks.domain.name` and the PKS Clusters domains e.g. `mycluster.pks.domain.name`.

Cluster Management

See [Managing Clusters](#) for recommendations on managing Enterprise PKS clusters.

AWS Reference Architecture

This guide presents a reference architecture for Pivotal Cloud Foundry (PCF) on Amazon Web Services (AWS). This architecture is valid for most production-grade PCF deployments using three availability zones (AZs).

See [PCF on AWS Requirements](#) for general requirements for running PCF and specific requirements for running PCF on AWS.

PCF Reference Architectures

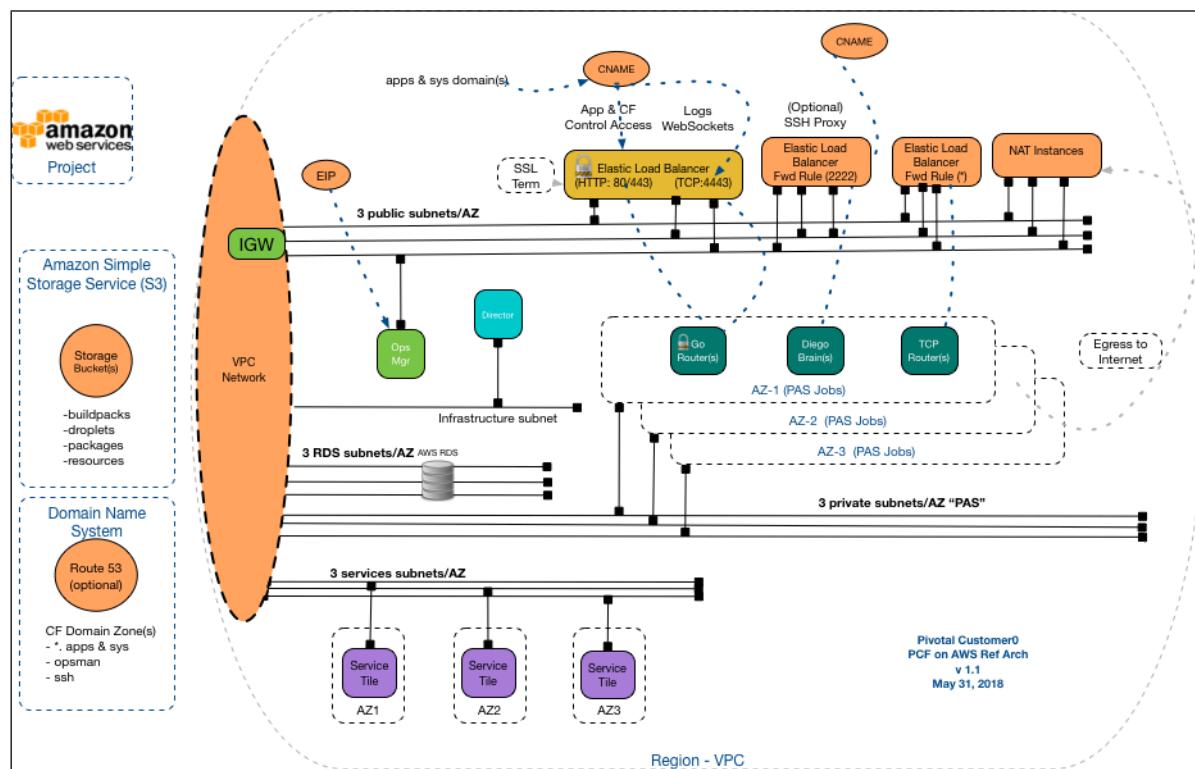
A PCF reference architecture describes a proven approach for deploying Pivotal Cloud Foundry on a specific IaaS, such as AWS, that meets the following requirements:

- Secure
- Publicly-accessible
- Includes common PCF-managed services such as MySQL, RabbitMQ, and Spring Cloud Services
- Can host at least 100 app instances, or far more

Pivotal provides reference architectures to help you determine the best configuration for your PCF deployment.

Base AWS Reference Architecture

The following diagram provides an overview of a reference architecture deployment of PCF on AWS using three AZs.



[View a larger version of this diagram ↗](#)

Note: Each AWS subnet must reside entirely within one AZ. As a result, a multi-AZ deployment topology requires a subnet for each AZ.

Base Reference Architecture Components

The following table lists the components that are part of a base reference architecture deployment on AWS with three AZs.

Component	Reference Architecture Notes
Domains & DNS	<p>CF Domain Zones and routes in use by the reference architecture include:</p> <ul style="list-style-type: none"> • domains for *.apps and *.sys (required) • a route for Ops Manager (required) • a route for ssh access to app containers (optional) <p>Using Route 53 to manage domains is optional.</p>
Ops Manager	Deployed on one of the three public subnets and accessible by FQDN or through an optional jumpbox.
BOSH Director	Deployed on the infrastructure subnet.
Elastic Load Balancers - HTTP, HTTPS, and SSL	Required. Load balancer that handles incoming HTTP, HTTPS, and SSL traffic and forwards them to the Gorouters. Deployed on all three public subnets.
Elastic Load Balancers - SSH	Optional. Load balancer that provides SSH access to app containers. Deployed on all three public subnets, one per AZ.
Gorouters	Accessed through the HTTP, HTTPS, and SSL Elastic Load Balancers. Deployed on all three Pivotal Application Service (PAS) subnets, one per AZ.
Diego Brains	Required. However, the SSH container access functionality is optional and enabled through the SSH Elastic Load Balancers. Deployed on all three PAS subnets, one per AZ.
TCP Routers	Optional feature for TCP routing. Deployed on all three PAS subnets, one per AZ.
CF Database	Reference architecture uses AWS RDS. Deployed on all three RDS subnets, one per AZ.
Storage Buckets	Reference architecture uses 4 S3 buckets: buildpacks, droplets, packages, and resources.
Service Tiles	Deployed on all three service subnets, one per AZ.
Service User & Roles	<p>One IAM role and one IAM user are recommended: the IAM role for Terraform, and the IAM user for Ops Manager and BOSH. Consult the following list:</p> <ul style="list-style-type: none"> • Admin Role: Terraform uses this IAM role to provision required AWS resources as well as an IAM user. • IAM User: This IAM user with IAM security credentials (access key ID and secret access key) is automatically provisioned with restrict access only to resources needed by PCF.
EC2 Instance Quota	The default EC2 instance quota on a new AWS subscription only has around 20 EC2 instances, which is not enough to host a multi-AZ deployment. The recommended quota for EC2 instances is 100. AWS requires the instances quota tickets to include Primary Instance Types, which should be <code>t3.micro</code> .

Network Objects

The following table lists the network objects in this reference architecture.

Network Object	Notes	Estimated Number
External Public IPs	One per deployment, assigned to Ops Manager.	1
Virtual Private Network (VPC)	One per deployment. A PCF deployment exists within a single VPC and a single AWS region, but should distribute PCF jobs and instances across 3 AWS AZs to ensure a high degree of availability.	1
Subnets	<p>The reference architecture requires the following subnets:</p> <ul style="list-style-type: none"> • 1 x (/24) infrastructure (BOSH Director) subnet • 3 x (/24) public subnets (Ops Manager, Elastic Load Balancers, NAT instances), one per AZ • 3 x (/20) PAS subnets (Gorouters, Diego Cells, Cloud Controllers, etc.), one per AZ • 3 x (/20) services subnets (RabbitMQ, MySQL, Spring Cloud Services, etc.), one per AZ • 3 x (/24) RDS subnets (Cloud Controller DB, UAA DB, etc.), one per AZ. 	13
	This reference architecture requires 4 route tables: one for the public subnet, and one each for all 3 private subnets across 3 AZs. Consult the following list:	

Route Tables	<ul style="list-style-type: none"> PublicSubnetRouteTable: This routing table enables the ingress/egress routes from/to Internet through the Internet gateway for Ops Manager and the NAT Gateway. PrivateSubnetRouteTable: This routing table enables the egress routing to the Internet through the NAT Gateway for the BOSH Director and PAS. <p>For more information, see the Terraform script that creates the route tables and the script that performs the route table association.</p> <p>Note: If an EC2 instance sits on a subnet with an Internet gateway attached as well as a public IP address, it is accessible from the Internet through the public IP address; for example, Ops Manager. PAS needs Internet access due to the access needs of using an S3 bucket as a blobstore.</p>	4																																													
Security Groups	<p>The reference architecture requires 5 Security Groups. For more information, see the Terraform Security Group rules script. The following table describes the Security Group ingress rules:</p> <p>Note: The extra port of 4443 with the Elastic Load Balancer is due to the limitation that the Elastic Load Balancer does not support WebSocket connections on HTTP/HTTPS.</p> <table border="1" data-bbox="187 747 1269 1129"> <thead> <tr> <th>Security Group</th><th>Port</th><th>From CIDR</th><th>Protocol</th><th>Description</th></tr> </thead> <tbody> <tr> <td>OpsMgrSG</td><td>22</td><td>0.0.0.0/0</td><td>TCP</td><td>Ops Manager SSH access</td></tr> <tr> <td>OpsMgrSG</td><td>443</td><td>0.0.0.0/0</td><td>TCP</td><td>Ops Manager HTTP access</td></tr> <tr> <td>VmsSG</td><td>ALL</td><td>VPC_CIDR</td><td>ALL</td><td>Open up connections among BOSH-deployed VMs</td></tr> <tr> <td>MysqlSG</td><td>3306</td><td>VPC_CIDR</td><td>TCP</td><td>Enable network access to RDS</td></tr> <tr> <td>ElbSG</td><td>80</td><td>0.0.0.0/0</td><td>TCP</td><td>HTTP to PAS</td></tr> <tr> <td>ElbSG</td><td>443</td><td>0.0.0.0/0</td><td>TCP</td><td>HTTPS to PAS</td></tr> <tr> <td>ElbSG</td><td>4443</td><td>0.0.0.0/0</td><td>TCP</td><td>WebSocket connection to Loggregator endpoint</td></tr> <tr> <td>SshElbSG</td><td>2222</td><td>0.0.0.0/0</td><td>TCP</td><td>SSH connection to containers</td></tr> </tbody> </table>	Security Group	Port	From CIDR	Protocol	Description	OpsMgrSG	22	0.0.0.0/0	TCP	Ops Manager SSH access	OpsMgrSG	443	0.0.0.0/0	TCP	Ops Manager HTTP access	VmsSG	ALL	VPC_CIDR	ALL	Open up connections among BOSH-deployed VMs	MysqlSG	3306	VPC_CIDR	TCP	Enable network access to RDS	ElbSG	80	0.0.0.0/0	TCP	HTTP to PAS	ElbSG	443	0.0.0.0/0	TCP	HTTPS to PAS	ElbSG	4443	0.0.0.0/0	TCP	WebSocket connection to Loggregator endpoint	SshElbSG	2222	0.0.0.0/0	TCP	SSH connection to containers	5
Security Group	Port	From CIDR	Protocol	Description																																											
OpsMgrSG	22	0.0.0.0/0	TCP	Ops Manager SSH access																																											
OpsMgrSG	443	0.0.0.0/0	TCP	Ops Manager HTTP access																																											
VmsSG	ALL	VPC_CIDR	ALL	Open up connections among BOSH-deployed VMs																																											
MysqlSG	3306	VPC_CIDR	TCP	Enable network access to RDS																																											
ElbSG	80	0.0.0.0/0	TCP	HTTP to PAS																																											
ElbSG	443	0.0.0.0/0	TCP	HTTPS to PAS																																											
ElbSG	4443	0.0.0.0/0	TCP	WebSocket connection to Loggregator endpoint																																											
SshElbSG	2222	0.0.0.0/0	TCP	SSH connection to containers																																											
Load Balancers	<p>PCF on AWS requires the Elastic Load Balancer, which can be configured with multiple listeners to forward HTTP/HTTPS/TCP traffic. Two Elastic Load Balancers are recommended: one to forward the traffic to the Gorouters, PcfElb, the other to forward the traffic to the Diego Brain SSH proxy, PcfSshElb. For more information, see the Terraform load balancers script.</p> <p>The following table describes the required listeners for each load balancer:</p> <table border="1" data-bbox="187 1343 1269 1534"> <thead> <tr> <th>ELB</th><th>Instance/Port</th><th>LB Port</th><th>Protocol</th><th>Description</th></tr> </thead> <tbody> <tr> <td>PcfElb</td><td>gorouter/80</td><td>80</td><td>HTTP</td><td>Forward traffic to Gorouters</td></tr> <tr> <td>PcfElb</td><td>gorouter/80</td><td>443</td><td>HTTPS</td><td>SSL termination and forward traffic to Gorouters</td></tr> <tr> <td>PcfElb</td><td>gorouter/80</td><td>4443</td><td>SSL</td><td>SSL termination and forward traffic to Gorouters</td></tr> <tr> <td>PcfSshElb</td><td>diego-brain/2222</td><td>2222</td><td>TCP</td><td>Forward traffic to Diego Brain for container SSH connections</td></tr> </tbody> </table> <p>Each ELB binds with a health check to check the health of the back-end instances:</p> <ul style="list-style-type: none"> PcfElb checks the health on Gorouter port 80 with TCP PcfSshElb checks the health on Diego Brain port 2222 with TCP 	ELB	Instance/Port	LB Port	Protocol	Description	PcfElb	gorouter/80	80	HTTP	Forward traffic to Gorouters	PcfElb	gorouter/80	443	HTTPS	SSL termination and forward traffic to Gorouters	PcfElb	gorouter/80	4443	SSL	SSL termination and forward traffic to Gorouters	PcfSshElb	diego-brain/2222	2222	TCP	Forward traffic to Diego Brain for container SSH connections	2																				
ELB	Instance/Port	LB Port	Protocol	Description																																											
PcfElb	gorouter/80	80	HTTP	Forward traffic to Gorouters																																											
PcfElb	gorouter/80	443	HTTPS	SSL termination and forward traffic to Gorouters																																											
PcfElb	gorouter/80	4443	SSL	SSL termination and forward traffic to Gorouters																																											
PcfSshElb	diego-brain/2222	2222	TCP	Forward traffic to Diego Brain for container SSH connections																																											
Jumpbox	<p>Optional. Provides a way of accessing different network components. For example, you can configure it with your own permissions and then set it up to access to Pivotal Network to download tiles. Using a jumpbox is particularly useful in IaaSes where Ops Manager does not have a public IP address. In these cases, you can SSH into Ops Manager or any other component through the jumpbox.</p>	1																																													

Integrate PCF with Customer Data Center through VPN

At times, applications on PCF need to access on-premise data. The connection between an AWS VPC and an on-premise datacenter is made through [VPN peering](#). When employing non-VPN peering, there are several points to consider:

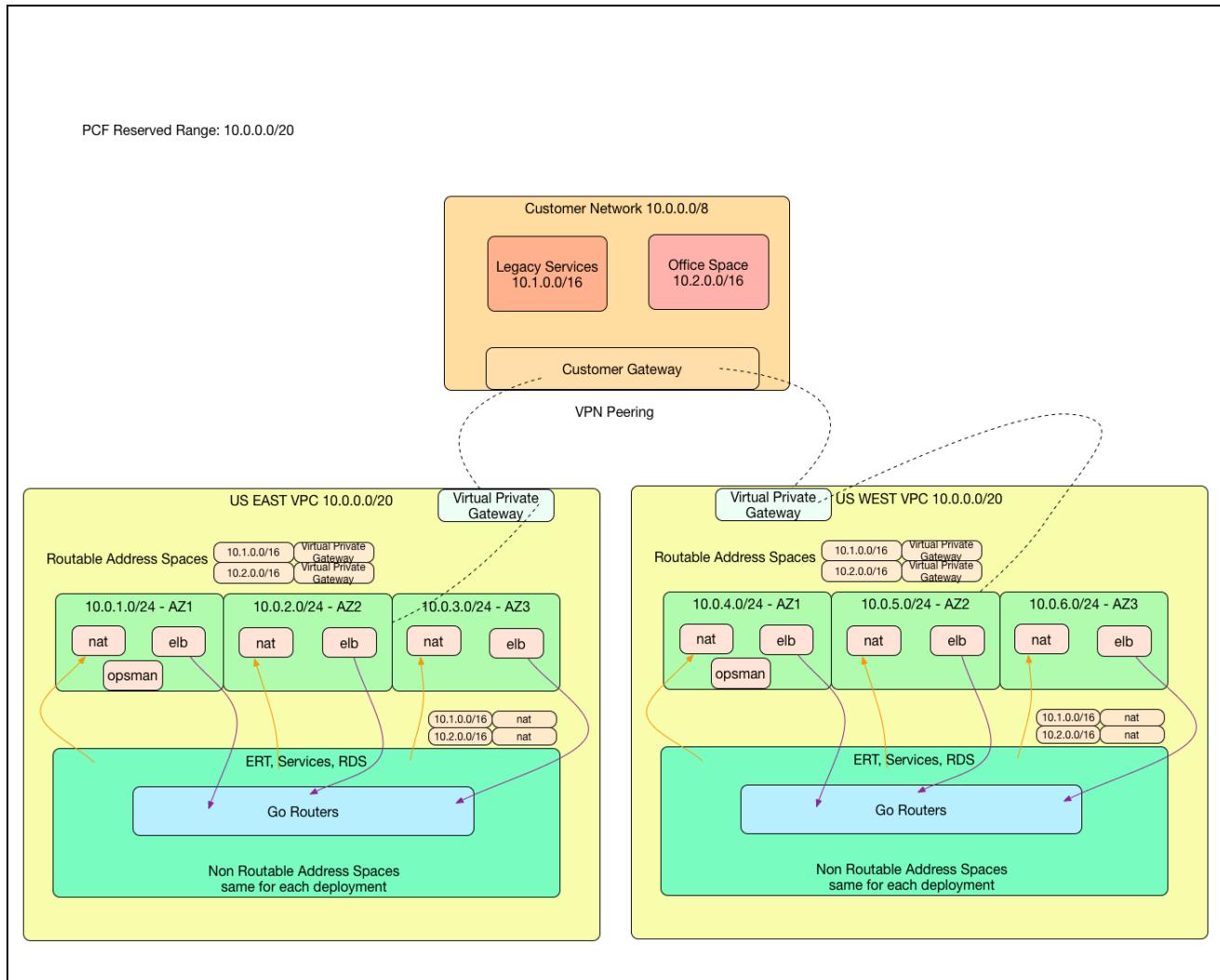
1. Assign routable IP addresses with the following in mind:

- It may not be realistic to request multiple routable /22 address spaces, due to IP exhaustion.
- Using different VPC address spaces can cause snowflakes deployments and present difficulties in automation.

- Only make the load balancer, NAT devices, and Ops Manager routable.
- PCF components can route egress through a NAT instance. As a result, operators do not need to assign routable IP addresses to PCF components.

2. Inbound traffic from the datacenter should come through an [internal load balancer](#).

3. Outbound traffic to the datacenter should go through AWS NAT instances.



[View a larger version of this diagram](#).

Azure Reference Architecture

This guide presents a reference architecture for Pivotal Cloud Foundry (PCF) on Azure.

Azure does not provide resources in a way that translates directly to PCF availability zones. Instead, Azure provides high availability through fault domains and [availability sets](#).

All reference architectures described in this topic are validated for production-grade PCF deployments using fault domains and availability sets that include multiple job instances.

See [Azure on PCF Requirements](#) for general requirements for running PCF and specific requirements for running PCF on Azure.

PCF Reference Architectures

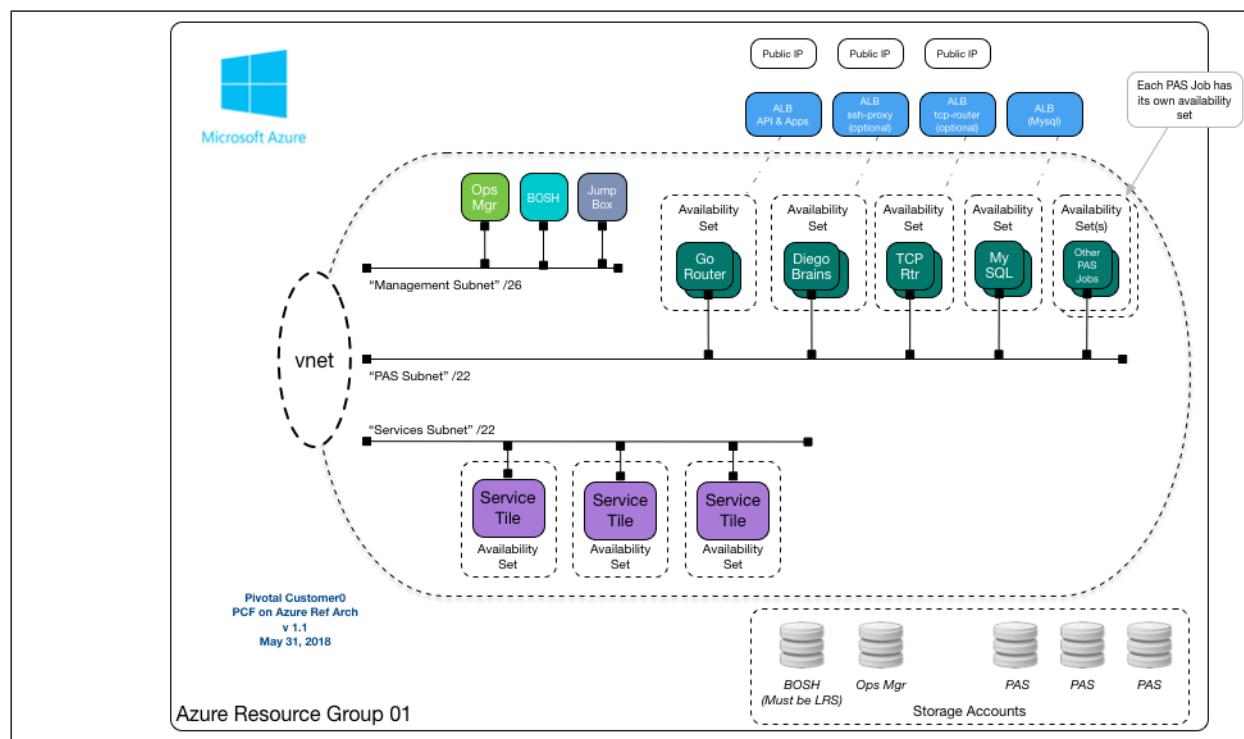
A PCF reference architecture describes a proven approach for deploying PCF on a specific IaaS, such as Azure, that meets the following requirements:

- Secure
- Publicly-accessible
- Includes common PCF-managed services such as MySQL, RabbitMQ, and Spring Cloud Services
- Can host at least 100 app instances, or far more

Pivotal provides reference architectures to help you determine the best configuration for your PCF deployment.

Base Azure Reference Architecture

The following diagram provides an overview of a reference architecture deployment of PCF on Azure.



[View a larger version of this diagram](#).

Base Reference Architecture Components

The following table lists the components that are part of a base reference architecture deployment on Azure using a single resource group.

Component	Reference Architecture Notes
Domains and DNS	CF Domain Zones and routes in use by the reference architecture include: <ul style="list-style-type: none">• domains for *.apps and *.system (required),• a route for Ops Manager (required),• a route for Doppler (required),• a route for Loggregator (required),• a route for SSH access to app containers (optional),• and a route for TCP routing to apps (optional).
Ops Manager	Deployed on the management subnet and accessible by FQDN or through an optional jumpbox.
BOSH	Deployed on the management subnet.
Azure Load Balancer - API and Apps	Required. Load balancer that handles incoming API and apps requests and forwards them to the Gorouters.
Azure Load Balancer - ssh-proxy	Optional. Load balancer that provides SSH access to app containers.
Azure Load Balancer - tcp-router	Optional. Load balancer that handles TCP routing requests for apps.
Azure Load Balancer - MySQL	Required to provide high availability for MySQL backend to Pivotal Application Service (PAS).
Gorouters	Accessed through the API and Apps load balancer. Deployed on the PAS subnet, one job per Azure availability set.
Diego Brains	Required. However, the SSH container access functionality is optional and enabled through the SSH Proxy load balancer. Deployed on the PAS subnet, one job per Azure availability set.
TCP Routers	Optional feature for TCP routing. Deployed on the PAS subnet, one job per availability zone.
MySQL	Reference architecture uses internal MySQL provided with PCF. Deployed on the PAS subnet, one job per Azure availability set.
PAS	Required. Deployed on the PAS subnet, one job per Azure availability set.
Storage Accounts	PCF on Azure requires 5 standard storage accounts: BOSH, Ops Manager, and three PAS storage accounts. Each account comes with a set amount of disk. Reference architecture recommends using 5 storage accounts because Azure Storage Accounts have an IOPs limit of approximately 20k per account, which generally relates to a BOSH JOB/VM limit of approximately 20 VMs each.
Service Tiles	Deployed on the Services subnet. Each service tile is deployed to an availability set.

Alternative Network Layouts for Azure

This section describes the possible network layouts for PCF deployments as covered by the reference architecture of PCF on Azure.

At a high level, there are currently two possible ways of deploying PCF as described by the reference architecture:

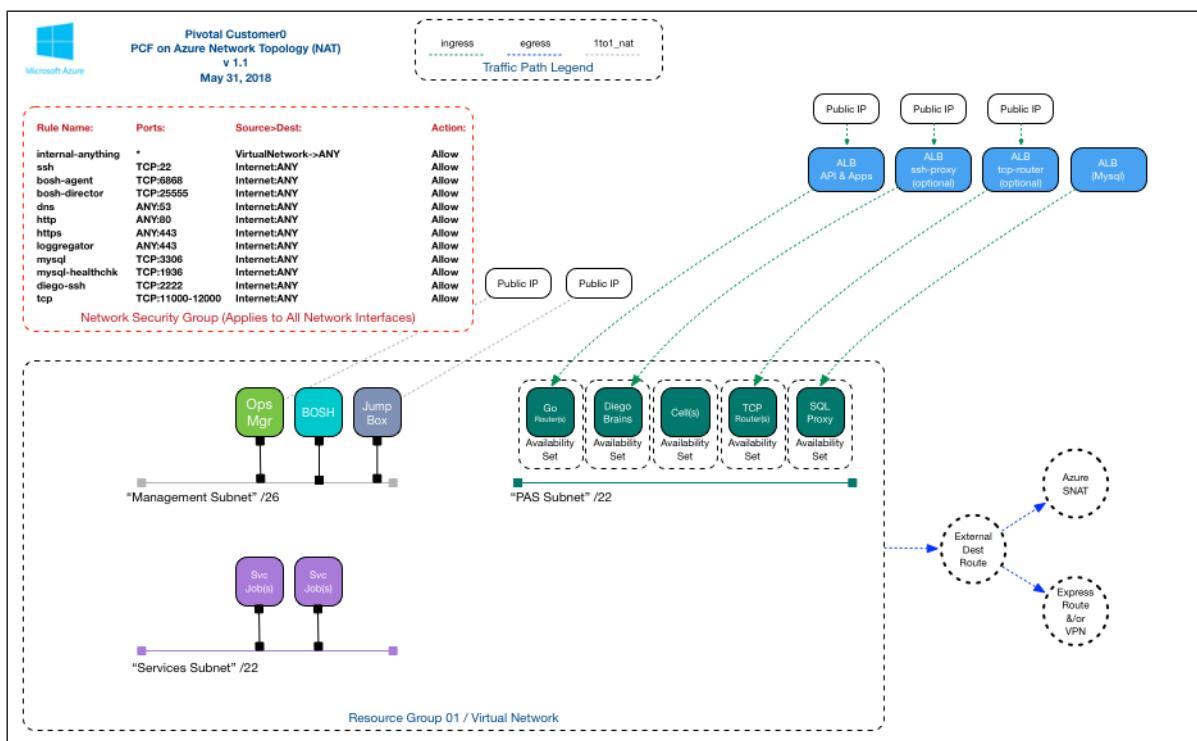
1. Single resource group, or
2. Multiple resource groups.

The first scenario is outlined in [Installing PCF on Azure](#). It models a single PCF deployment in a single Azure Resource Group.

If you require multiple resource groups, refer to the [Multiple Resource Group deployment](#) section of this topic.

Network Layout

This diagram illustrates the network topology of the base reference architecture for PCF on Azure. In this deployment, you expose only a minimal number of public IP addresses and deploy only one resource group.



[View a larger version of this diagram ↗](#)

Network Objects

The following table lists the network objects in PCF on Azure reference architecture.

Network Object	Notes	Estimated Number
External Public IP addresses	Use <ol style="list-style-type: none"> global IP address for apps and system access Ops Manager or optional jumpbox. Optionally, you can use a public IP address for the ssh-proxy and tcp-router load balancers.	1-4
Virtual Network	One per deployment. Azure virtual network objects allow multiple subnets with multiple CIDRs, so a typical deployment of PCF will likely only ever require one Azure Virtual Network object.	1
Subnets	Separate subnets for <ol style="list-style-type: none"> management (Ops Manager, BOSH Director, Jumpbox), PAS, and services. Using separate subnets allows you to configure different firewall rules due to your needs.	4
Routes	Routes are typically created by Azure dynamically when subnets are created, but you may need to create additional routes to force outbound communication to dedicated SNAT nodes. These objects are required to deploy PCF without public IP addresses.	3+
Firewall Rules	Azure firewall rules are collected into a Network Security Group (NSG) and bound to a Virtual Network object and can be created to use IP ranges, subnets, or instance tags to match for source and destination fields in a rule. One NSG can be used for all firewall rules.	12
Load Balancers	Used to handle requests to Gorouters and infrastructure components. Azure uses 1 or more load balancers. The API and Apps load balancer is required. The TCP Router load balancer used for TCP routing feature and the SSH load balancer that allows SSH access to Diego apps are both optional. In addition, you can use a MySQL load balancer to provide high availability to MySQL. This is also optional.	1-4
	Optional. Provides a way of accessing different network components. For example, you can configure it with your own	

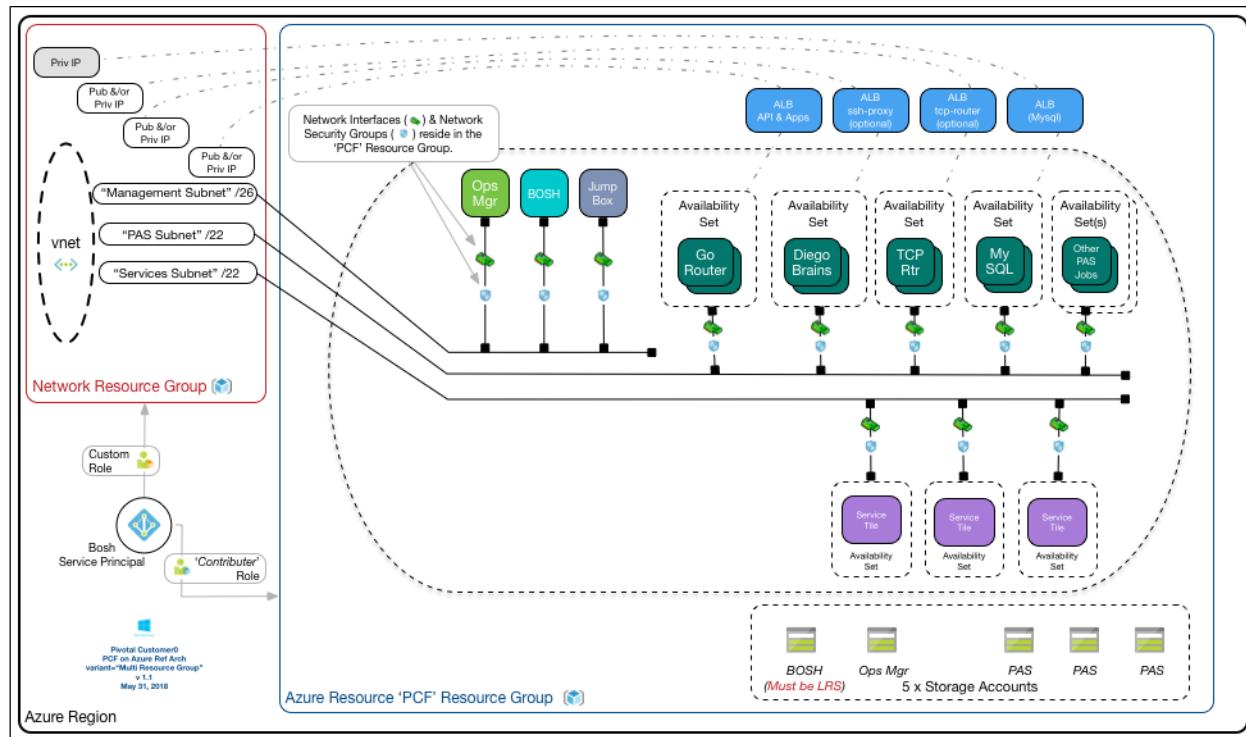
Jumpbox	permissions and then set it up to access to Pivotal Network to download tiles. Using a jumpbox is particularly useful in IaaSes where Ops Manager does not have a public IP address. In these cases, you can SSH into Ops Manager or any other component through the jumpbox.	1
---------	---	---

Multiple Resource Group Deployment

This diagram illustrates the case where you want to use additional resource groups in your PCF deployment on Azure.

Shared network resources may already exist in an Azure subscription. In this type of deployment, using multiple resource groups allows you to reuse existing resources instead of provisioning new ones.

To use multiple resource groups, you need to provide the BOSH Service Principal with access to the existing network resources.



[View a larger version of this diagram ↗](#)

Multiple Resource Groups Deployment Notes

To deploy PCF on Azure with multiple resource groups, you can define custom roles to grant resource group access to your BOSH Service Principal. For example, you might develop the following:

- Dedicated **Network** Resource Group, limits BOSH Service Principal so that it does not have admin access to network objects.
- Custom Role for BOSH Service Principal, applied to **Network** Resource Group, limits the BOSH Service Principal to minimum read-only access.

```
az role definition create --role-definition
{ \
  "Name": "PCF Network Read Only", \
  "IsCustom": true, \
  "Description": "MVP PCF Read Network Resgroup", \
  "Actions": [ \
    "Microsoft.Network/virtualNetworks/read", \
    "Microsoft.Network/virtualNetworks/subnets/read", \
    "Microsoft.Network/virtualNetworks/subnets/join/action", \
    "Microsoft.Network/networkSecurityGroups/read", \
    "Microsoft.Network/networkSecurityGroups/join/action", \
    "Microsoft.Network/loadBalancers/read", \
    "Microsoft.Network/publicIPAddresses/read", \
    "Microsoft.Network/publicIPAddresses/join/action" \
  ], \
  "NotActions": [], \
  "AssignableScopes": ["/subscriptions/[YOUR_SUBSCRIPTION_ID]"] \
}
```

The actions prefixed with `Microsoft.Network/publicIPAddresses` are only required if using IP addresses.

- Custom Role for BOSH Service Principal, applied to Subscription, allowing the Operator to deploy PCF components

```
az role definition create --role-definition
{ \
  "Name": "PCF Deploy Min Perms", \
  "IsCustom": true, \
  "Description": "MVP PCF Terraform Perms", \
  "Actions": [ \
    "Microsoft.Compute/register/action" \
  ], \
  "NotActions": [], \
  "AssignableScopes": ["/subscriptions/[YOUR_SUBSCRIPTION_ID]"] \
}
```

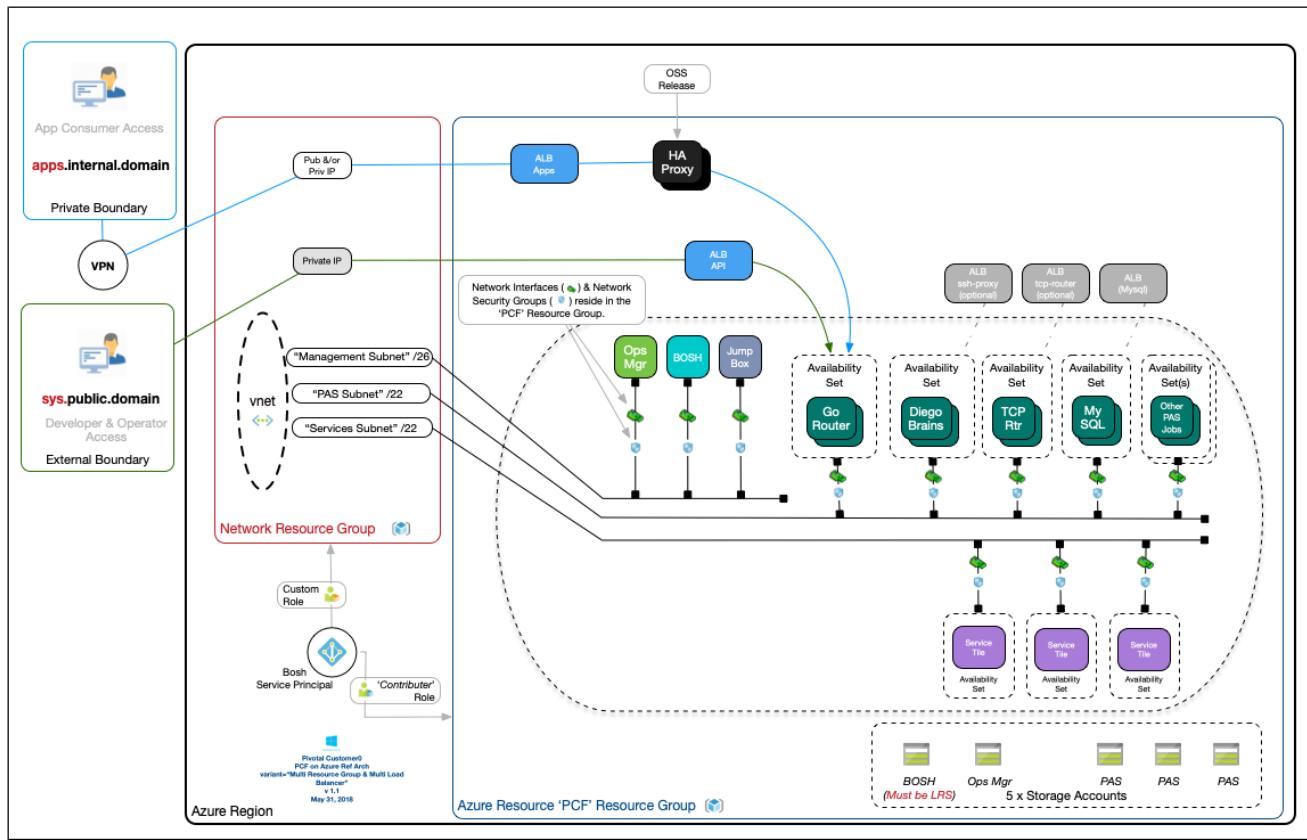
- Custom roles can be assigned to service principals with `az role assignment create --role [ROLE] --assignee [SERVICE_PRINCIPAL_ID]`.

Multiple Resource Group and Multiple Load Balancer Deployment

This diagram illustrates the case where you want to deploy multiple load balancers with your multiple resource group deployment.

The key design points of this deployment are the following:

- Two Azure load balancer (ALBs) to the Gorouter exist. The first ALB is for API access, which utilizes a private IP address. The system domain should resolve to this ALB. The second ALB is for application access, which can either use a public or private IP address. The apps domain should resolve to this ALB.
- Azure does not allow the Gorouters to be members of more than one ALB member pool for the same ports, for example `80` and `443`. This restriction requires an additional reverse proxy to front the Gorouters to allow them to expose traffic on these ports for the system domain.



[View a larger version of this diagram ↗](#)

Log Analytics Integration

Azure Log Analytics [↗](#) is a service that helps you collect and analyze data generated by resources in your cloud and on-premises environments. The Microsoft Azure Log Analytics Nozzle for PCF [↗](#) receives logs and metrics from the Loggregator Firehose, filters and resolves the events, and then sends them to Log Analytics, where they appear on unified dashboards and can be correlated with and alerted on other Azure resources.

DNS Delegation

Azure DNS supports DNS delegation, allowing for sub-level domains to be hosted within Azure. This functionality is fully supported within PCF.

Pivotal recommends that use a sub-zone for your PCF deployment. For example, if your company's domain is `example.com`, your PCF zone in Azure DNS would be `pcf.example.com`.

As Azure DNS does not support recursion, in order to properly configure Azure DNS, create an `NS` record with your registrar which points to the four name servers supplied by your Azure DNS Zone configuration. Once your `NS` records have been created, you can then create the required wildcard `A` records for the PCF application and system domains, as well as any other records desired for your PCF deployments.

You do not need to make any configuration changes in PCF to support Azure DNS.

Azure Blob Storage

Due to limitations in PCF, it is not possible to support a high-availability deployment of the backing NFS store needed for droplets, buildpacks, and other resources. Azure Blob Storage provides fully-redundant hot, cold, or archival storage in either local, regional, or global offerings. It is recommended to use Azure Blob Storage as the external File Storage to provide unlimited scaling and redundancy for high-availability deployments of PCF.

To enable Azure Blob Storage, do the following:

1. Create a Storage Account with the level of redundancy you require. The storage account does not need to be in the same resource Group as PCF.
2. Create Containers for the buildpacks, droplets, packages, and resources required by PCF.

3. Open Ops Manager.
4. Select **Pivotal Application Services**.
5. Click the **Settings Tab**.
6. Select **File Storage**.
7. Click **External Azure Account**.
8. Enter the Storage Account details.
9. In the Ops Manager main page, apply the changes when you are ready to reconfigure.

 **Note:** There is no direct path to migrate previous objects to Azure Blob Storage. Contact Pivotal Support if you need assistance with this migration.

Load Balancer Migrations

The Azure Load Balancer has two SKUs: Standard and Basic. The Standard SKU Load Balancer is a new load balancer for all TCP and UDP applications with an expanded and more granular feature set over the Basic Load Balancer. Azure's Standard SKU Load Balancer lets you scale your applications and create high availability in environments ranging from small scale deployments to large and complex multi-zone architectures. While the Basic SKU Load Balancer works within the scope of an availability set, a Standard Load Balancer covers an entire virtual network. Pivotal recommends using the Standard SKU Load Balancer instead of the Basic SKU Load Balancer. If you currently use the Basic SKU Load Balancer and wish to migrate to the the Standard SKU, please see the [Migrate Basic SKU Load Balancer to Standard SKU Load Balancer](#) documentation on GitHub.

GCP Reference Architecture

This topic presents two reference architectures for installing Pivotal Cloud Foundry (PCF) on Google Cloud Platform (GCP): on a [shared virtual private cloud](#) (VPC) and on a [single-project VPC](#). This topic also outlines multiple networking variants for VPC deployment. The architectures are validated for production-grade PCF deployments using multiple AZs.

See [PCF on GCP Requirements](#) for general requirements for running PCF and specific requirements for running PCF on GCP.

PCF Reference Architectures

A PCF reference architecture describes a proven approach for deploying PCF on a specific IaaS, such as GCP.

A PCF reference architecture must meet the following requirements:

- Be secure
- Be publicly-accessible
- Include common PCF-managed services such as MySQL, RabbitMQ, and Spring Cloud Services
- Be able to host at least 100 app instances

Pivotal provides reference architectures to help you determine the best configuration for your PCF deployment.

Shared vs Single-Project VPCs

A [shared VPC](#) installation is harder to configure than a PCF deployment on a single-project VPC, because the required account privileges and resource allocations are more granular and complex. But the shared VPC architecture allows network assets to be centrally located, which simplifies auditing and security. Pivotal recommends the shared VPC model for:

- Deployments with deep auditing and security requirements
- When networks hosting the foundation need to connect back to an internal network via VPN or interconnect

A [single-project VPC](#) lets the platform architect give PCF full access to the VPC and its resources, which makes configuration easier. Pivotal recommends single-project VPC architecture for:

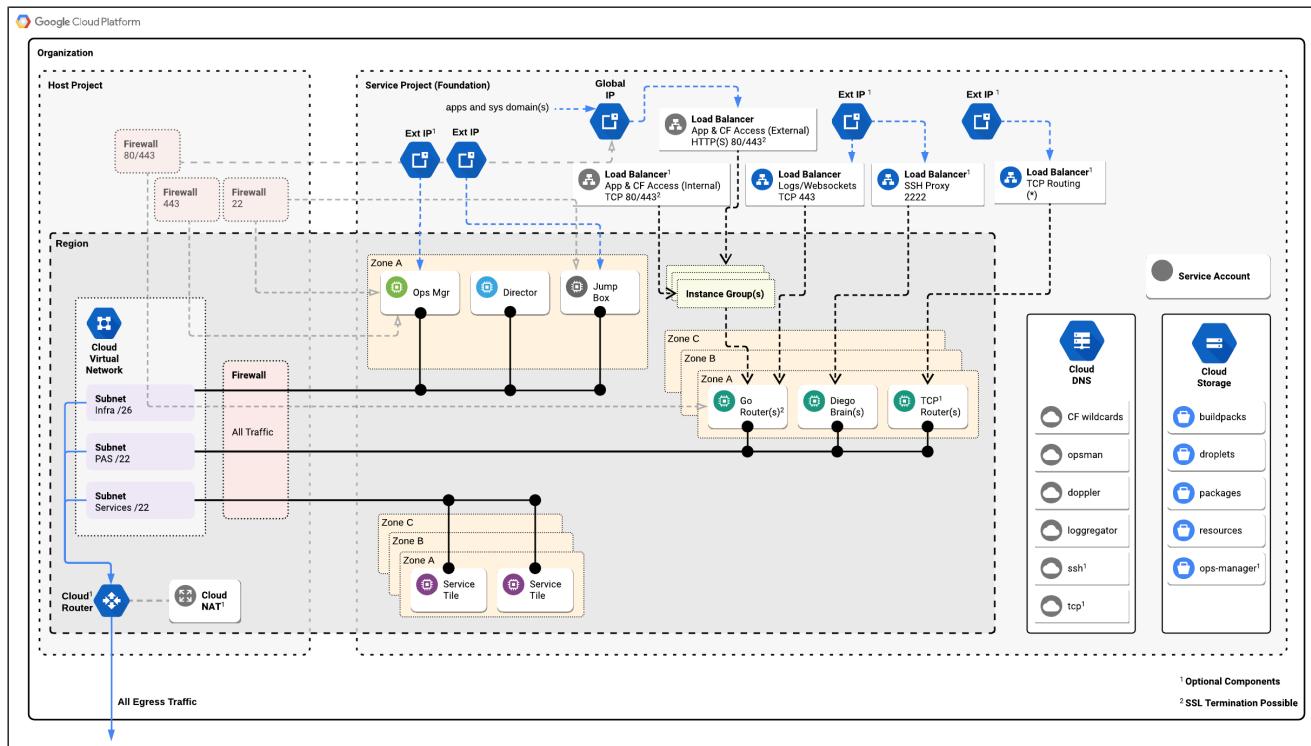
- Standalone deployments that do not connect to an internal network
- Test and experimental deployments, and for projects which do not belong to an organization

Shared VPC GCP Reference Architecture

The following diagram provides an overview of a reference architecture deployment of PCF on a shared VPC on GCP. This architecture requires an organization on the VPC that contains a host project and a service project.

For more information about shared VPCs on GCP, see [Shared VPC Overview](#) in the Google Cloud documentation.

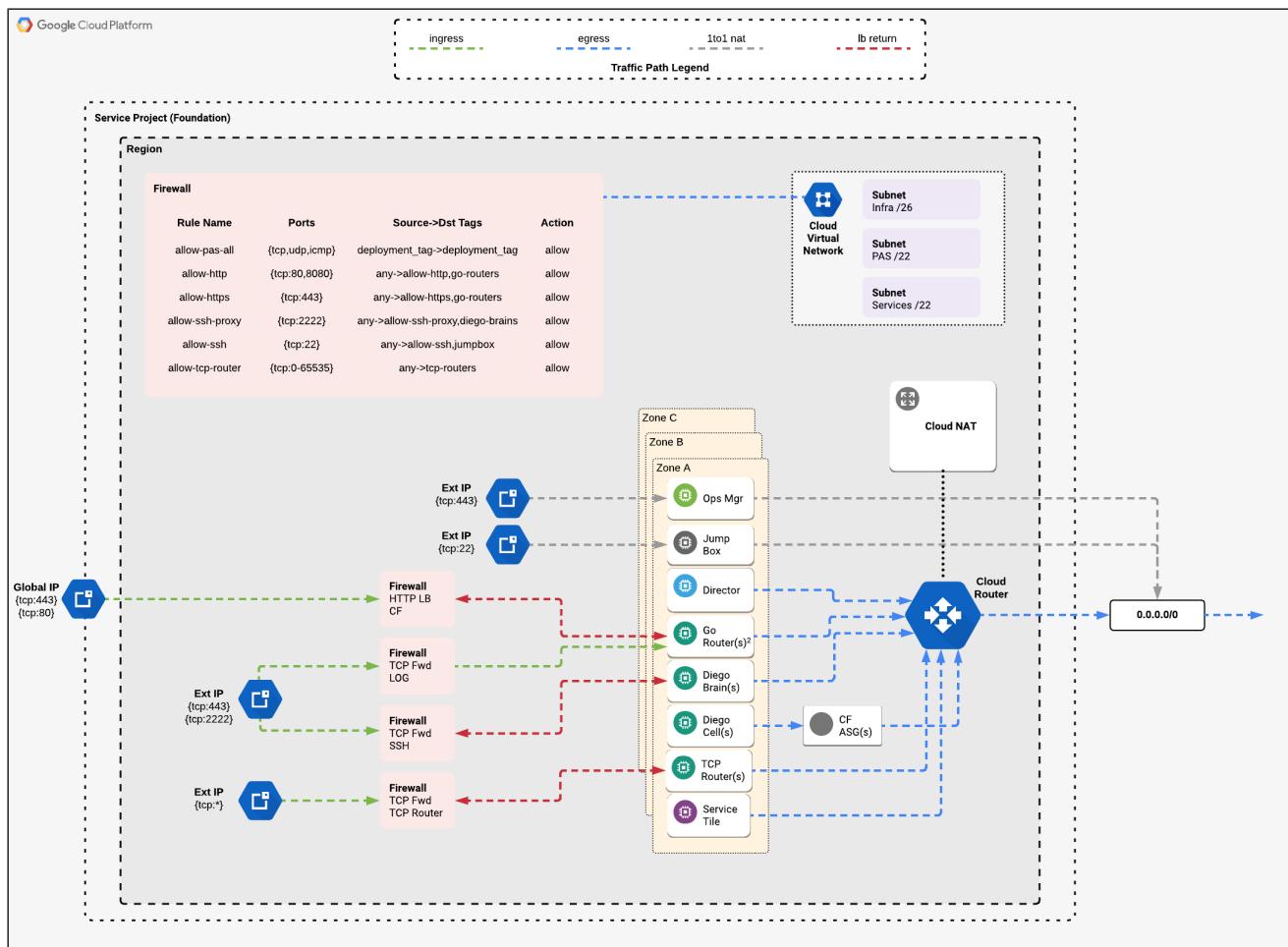
For more information about how this architecture divides resources between projects, see [Host / Service Architecture](#).



[View a larger version of this diagram.](#)

NAT Network Topology

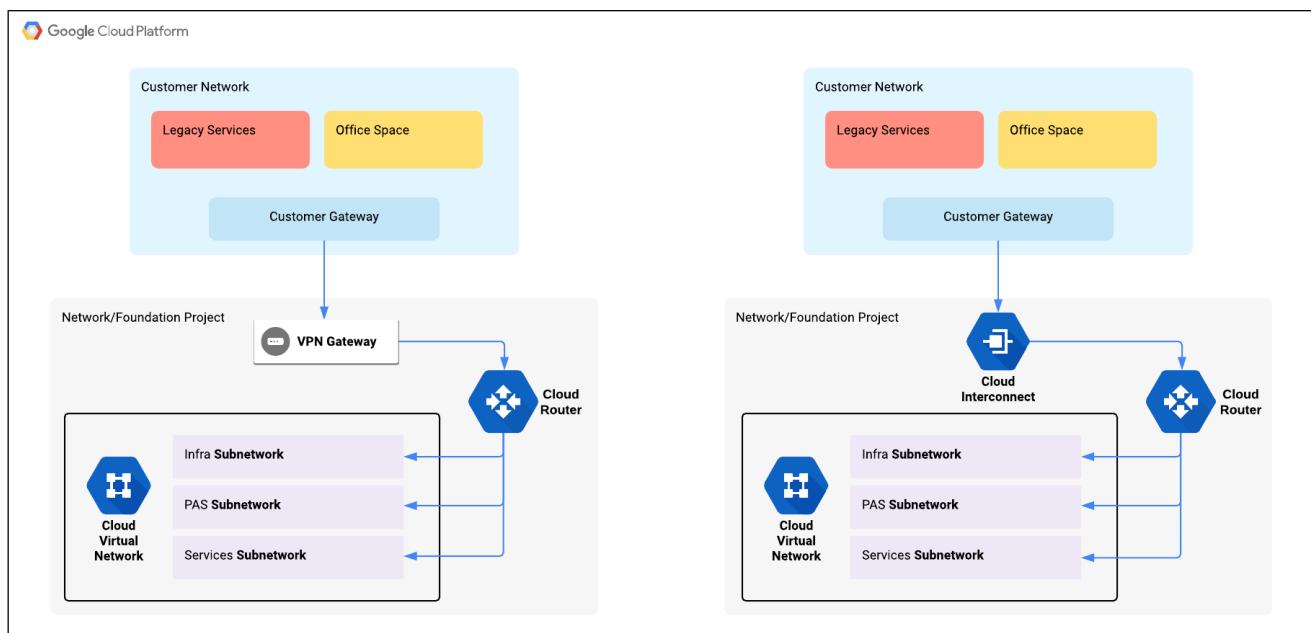
To expose a minimal number of public IP addresses, set up your NAT as shown in this diagram.



[View a larger version of this diagram.](#)

Cloud Interconnect

To speed communication between data centers, use Google Cloud Interconnect as shown in this diagram.



[View a larger version of this diagram.](#)

Host / Service Architecture

GCP allocates resources using a hierarchy that centers around projects. To create a VPC, architects define a *host project* that allocates network resources for the VPC, such as address space and firewall rules. Then they can define one or more *service projects* to run within the VPC, which share the network resources allocated by the host project and include their own non-network resources, such as VMs and storage buckets.

To install PCF in a shared VPC on GCP, you create a host project for the VPC and a service project dedicated to running PCF.

Host Project Resources

The host project centrally manages the following shared VPC network resources for PCF:

- Infra subnet (Ops Manager and BOSH Director)
- PAS subnet
- Services subnet
- Isolation Segments
- Firewall rules
- NAT instances and gateway
- VPN/interconnect
- Routes, e.g. egress internet through NAT or egress on-premises through VPN

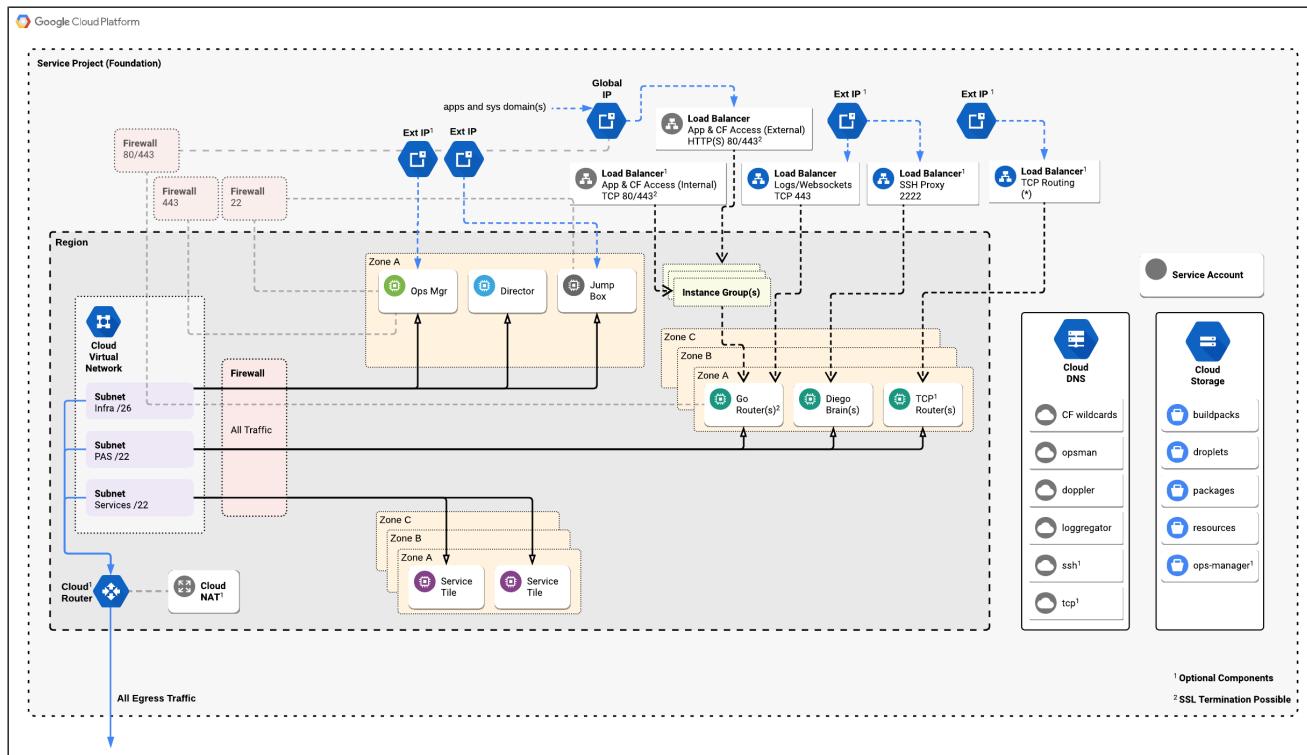
Service Project Resources

The PCF service project manages the following resources:

- Google Cloud Compute instances (VMs)
 - BOSH
 - Ops Manager
 - VMs deployed by BOSH, such as PCF and service components
- Google Cloud Storage buckets, for blobstore
 - BOSH Director
 - Resources
 - Buildpacks
 - Droplets
 - Packages
- Service account and a service account key for PCF to access the storage buckets
- A service account for PCF
- Load Balancers
- Google Cloud SQL instances, if using external databases

Single-Project VPC Base GCP Reference Architecture

The following diagram provides an overview of a reference architecture deployment of PCF on a single-project [VPC](#) on GCP.



[View a larger version of this diagram.](#)

Base Reference Architecture Components

The following table lists the components that are part of a reference architecture deployment with three availability zones.

Component	Reference Architecture Notes
Domains & DNS	CF Domain Zones and routes in use by the reference architecture include: domains for *.apps and *.system (required), a route for Ops Manager (required), a route for doppler (required), a route for Loggregator (required), a route for ssh access to app containers (optional) and a route for TCP routing to apps (optional). Reference architecture uses GCP Cloud DNS as the DNS provider.
Ops Manager	Deployed on the infrastructure subnet and accessible by FQDN or through an optional jumpbox.
BOSH Director	Deployed on the infrastructure subnet.
Gorouters	Accessed through the HTTP and TCP WebSockets load balancers. Deployed on the Pivotal Application Service (PAS) subnet, one job per availability zone.
Diego Brains	Required. However, the SSH container access functionality is optional and enabled through the SSH Proxy load balancer. Deployed on the PAS subnet, one job per availability zone.
TCP Routers	Optional feature for TCP routing. Deployed on the PAS subnet, one job per availability zone.
CF Database	Reference architecture uses GCP Cloud SQL rather than internal databases. Configure your database with a strong password and limit access only to components that require database access.
CF Blob Storage and Buckets	For buildpacks, droplets, packages and resources. Reference architecture uses Google Cloud Storage rather than internal file storage.
Services	Deployed on the PCF managed services subnet. Each service is deployed to each availability zone.

Alternative GCP Network Layouts for PCF

This section describes the possible network layouts for PCF deployments as covered by the reference architecture of PCF on GCP.

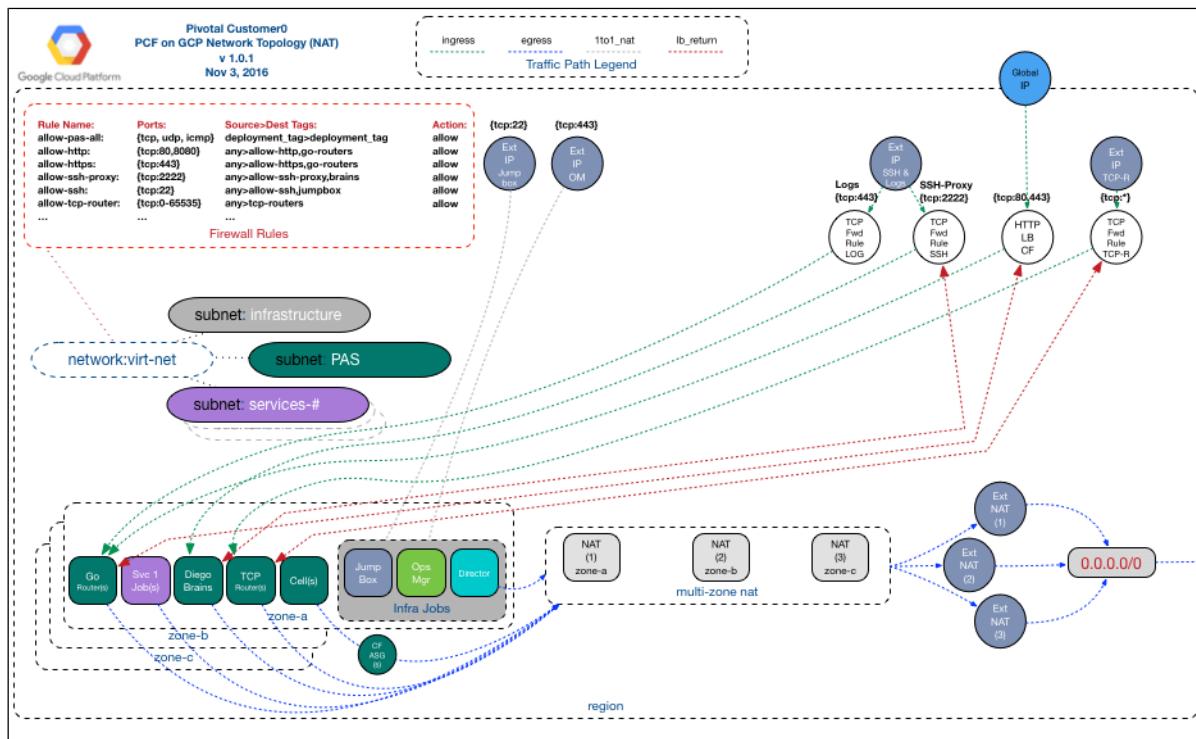
At a high level, there are currently two possible ways of granting public Internet access to PCF as described by the reference architecture:

- NAT provides connectivity from PCF internals to the public Internet.
 - The instructions for [Installing PCF on GCP Manually](#) use this method.
- Every PCF VM receives its own public IP address (no NAT).
 - The instructions for [Installing PCF on GCP using Terraform](#) use this method.

However, if you require NAT, you may refer to the following section.

NAT-based Solution

This diagram illustrates the case where you want to expose only a minimal number of public IP addresses.



[View a larger version of this diagram ↗](#)

Public IP addresses Solution

If you prefer not to use a NAT solution, you can configure PCF on GCP to assign public IP addresses for all components. This type of deployment may be more performant since most of the network traffic between Cloud Foundry components are routed through the front end load balancer and the Gorouter.

Network Objects

The following table lists the network objects expected for each type of reference architecture deployment with three availability zones (assumes you are using NAT).

Network Object	Notes	Minimum Number: NAT-based	Minimum Number: Public IP Addresses
External IPs	For a NAT solution, use global IP address for apps and system access, and Ops Manager or optional jumpbox	2	30+
NAT	One NAT per availability zone.	3	0
Network	One per deployment. GCP Network objects allow multiple subnets with multiple CIDRs, so a typical deployment of PCF will likely only ever require one GCP Network object	1	1
	Separate subnets for infrastructure (Ops Manager, BOSH Director, Jumpbox), PAS, and services. Using separate		

Subnets	subnets allows you to configure different firewall rules due to your needs.	Minimum	Minimum
Network Object Routes	Notes Routes are typically created by GCP dynamically when subnets are created, but you may need to create additional routes to force outbound communication to dedicated SNAT nodes. These objects are required to deploy PCF without public IP addresses.	Number: NAT-Based	Number: Public IP Addresses
Firewall Rules	GCP firewall rules are bound to a Network object and can be created to use IP ranges, subnets, or instance tags to match for source & destination fields in a rule. The preferred method use in the reference architecture deployment is instance tags.	6+	6+
Load balancers	Used to handle requests to Gorouters and infrastructure components. GCP uses two or more load balancers. The HTTP load balancer and TCP WebSockets load balancer are both required. The TCP Router load balancer used for TCP routing feature and the SSH load balancer that allows SSH access to Diego apps are both optional. The HTTP load balancer provides SSL termination.	2+	2+
Jumpbox	Optional. Provides a way of accessing different network components. For example, you can configure it with your own permissions and then set it up to access to Pivotal Network to download tiles. Using a jumpbox is particularly useful in IaaS where Ops Manager does not have a public IP address. In these cases, you can SSH into Ops Manager or any other component through the jumpbox.	(1)	(1)

Network Communication in GCP Deployments

This section provides more background on the reasons behind certain network configuration decisions, specifically for the Gorouter.

Load Balancer to Gorouter Communications and TLS Termination

In a PCF on GCP deployment, the Gorouter receives two types of traffic:

1. Unencrypted HTTP traffic on port 80 that is decrypted by the HTTP(S) load balancer.
2. Encrypted secure web socket traffic on port 443 that is passed through the TCP WebSockets load balancer.

TLS is terminated for HTTPS on the HTTP load balancer and is terminated for WebSockets (WSS) traffic on the Gorouter.

PCF deployments on GCP use two load balancers to handle Gorouter traffic because HTTP load balancers currently do not support WebSockets.

ICMP

GCP routers do not respond to ICMP; therefore, Pivotal recommends disabling ICMP checks in [BOSH Director network configuration](#).

OpenStack Reference Architecture

This guide presents a reference architecture for Pivotal Cloud Foundry (PCF) on OpenStack. This architecture is valid for most production-grade PCF deployments in a single project using three availability zones (AZs).

See [OpenStack on PCF Requirements](#) for general requirements for running PCF and specific requirements for running PCF on OpenStack.

PCF Reference Architectures

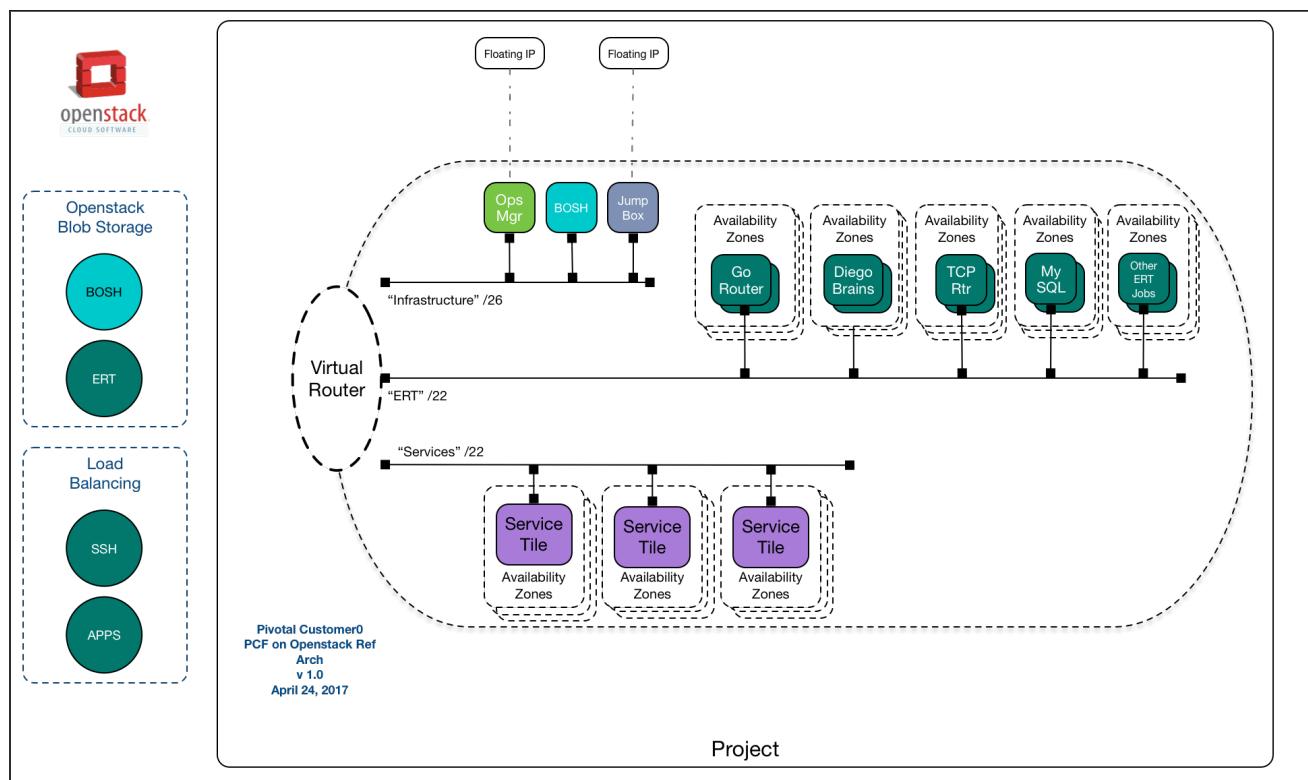
A PCF reference architecture describes a proven approach for deploying Pivotal Cloud Foundry on a specific IaaS, such as OpenStack, that meets the following requirements:

- Secure
- Publicly-accessible
- Includes common PCF-managed services such as MySQL, RabbitMQ, and Spring Cloud Services
- Can host at least 100 app instances, or far more

Pivotal provides reference architectures to help you determine the best configuration for your PCF deployment.

Base OpenStack Reference Architecture

The following diagram provides an overview of a reference architecture deployment of PCF on OpenStack using three AZs.



[View a larger version of this diagram.](#)

Base Reference Architecture Components

The following table lists the components that are part of a base reference architecture deployment on OpenStack with three AZs.

Component	Reference Architecture Notes
	CF Domain zones and routes in use by the reference architecture include:

Domains & DNS	<ul style="list-style-type: none"> • zones for *.apps and *.sys (required) • a route for Ops Manager (required) • a route for Doppler (required) • a route for Loggregator (required) • a route for ssh access to app containers (optional) • a route for tcp access to tcp routers (optional)
Ops Manager	Deployed on the infrastructure network and accessible by FQDN or through an optional jumpbox.
BOSH Director	Deployed on the infrastructure network.
Application Load Balancer	Required. Load balancer that handles incoming HTTP, HTTPS, TCP, and SSL traffic and forwards them to the Gorouters. Load balancers are outside the scope of this document.
SSH Load Balancer	Optional. Load balancer that provides SSH access to application containers for developers. Load balancers are outside the scope of this document.
Gorouters	Accessed through the Application Load Balancer. Deployed on the Pivotal Application Service (PAS) network, one per AZ.
Diego Brains	This component is required. However, the SSH container access functionality is optional and enabled through the SSH Load Balancers. Deployed on the PAS network, one per AZ.
TCP Routers	Optional feature for TCP routing. Deployed on the PAS network, one per AZ.
CF Database	Reference architecture uses internal MySQL.
Storage Buckets	Reference architecture uses customer provided blobstore. Buckets are needed for BOSH and PAS.
Service Tiles	Deployed on the services network.
Service Accounts	<p>Two service accounts are recommended: one for OpenStack “paving,” and the other for Ops Manager and BOSH. Consult the following list:</p> <ul style="list-style-type: none"> • Admin Account: Concourse will use this account to provision required OpenStack resources as well as a Keystone service account. • Keystone Service Account: This service account will be automatically provisioned with restricted access only to resources needed by PCF.
OpenStack Quota	The default compute quota on a new OpenStack subscription is typically not enough to host a multi-AZ deployment. The recommended quota for instances is 100. Your OpenStack network quotas may also need to be increased.

OpenStack Objects

The following table lists the network objects in this reference architecture.

Network Object	Notes	Estimated Number
Floating IP addresses	Two per deployment. One assigned to Ops Manager, the other to your jumpbox.	2
Project	One per deployment. A PCF deployment exists within a single project and a single OpenStack region, but should distribute PCF jobs and instances across three OpenStack AZs to ensure a high degree of availability.	1
Networks	<p>The reference architecture requires the following Tenant Networks:</p> <ul style="list-style-type: none"> • 1 x (/24) Infrastructure (Ops Manager, BOSH Director, Jumpbox). • 1 x (/20) PAS (Gorouters, Diego Cells, Cloud Controllers, etc.). • 1 x (/20) Services (RabbitMQ, MySQL, Spring Cloud Services, etc.) • 1 x (/24) On-demand services (Various.) <p>An Internet-facing network is also required:</p> <ul style="list-style-type: none"> • 1 x Public network. <p>Note: In many cases, the public network is an “under the cloud” network that is shared across projects.</p>	5
	This reference architecture requires one router attached to all networks:	

Routers	<ul style="list-style-type: none"> VirtualRouter: This router table enables the ingress/egress routes from/to Internet to the project networks and provides sNAT services. 	1																				
Security Groups	<p>The reference architecture requires one Security Groups. The following table describes the Security Group ingress rules:</p> <table border="1"> <thead> <tr> <th>Security Group</th><th>Port</th><th>From CIDR</th><th>Protocol</th><th>Description</th></tr> </thead> <tbody> <tr> <td>OpsMgrSG</td><td>22</td><td>0.0.0.0/0</td><td>TCP</td><td>Ops Manager SSH access</td></tr> <tr> <td>OpsMgrSG</td><td>443</td><td>0.0.0.0/0</td><td>TCP</td><td>Ops Manager HTTP access</td></tr> <tr> <td>VmsSG</td><td>ALL</td><td>VPC_CIDR</td><td>ALL</td><td>Open up connections among BOSH-deployed VMs</td></tr> </tbody> </table> <p>Additional security groups may be needed which are specific to your chosen load balancing solution.</p>	Security Group	Port	From CIDR	Protocol	Description	OpsMgrSG	22	0.0.0.0/0	TCP	Ops Manager SSH access	OpsMgrSG	443	0.0.0.0/0	TCP	Ops Manager HTTP access	VmsSG	ALL	VPC_CIDR	ALL	Open up connections among BOSH-deployed VMs	5
Security Group	Port	From CIDR	Protocol	Description																		
OpsMgrSG	22	0.0.0.0/0	TCP	Ops Manager SSH access																		
OpsMgrSG	443	0.0.0.0/0	TCP	Ops Manager HTTP access																		
VmsSG	ALL	VPC_CIDR	ALL	Open up connections among BOSH-deployed VMs																		
Load Balancers	<p>PCF on OpenStack requires a load balancer, which can be configured with multiple listeners to forward HTTP/HTTPS/TCP traffic. Two load balancers are recommended: one to forward the traffic to the Gorouters, AppsLB, the other to forward the traffic to the Diego Brain SSH proxy, SSHLB.</p> <p>The following table describes the required listeners for each load balancer:</p> <table border="1"> <thead> <tr> <th>Name</th><th>Instance/Port</th><th>LB Port</th><th>Protocol</th><th>Description</th></tr> </thead> <tbody> <tr> <td>AppsLB</td><td>gorouter/80</td><td>80</td><td>HTTP</td><td>Forward traffic to Gorouters</td></tr> <tr> <td>AppsLB</td><td>gorouter/80</td><td>443</td><td>HTTPS</td><td>SSL termination and forward traffic to Gorouters</td></tr> <tr> <td>SSHLB</td><td>diego-brain/2222</td><td>2222</td><td>TCP</td><td>Forward traffic to Diego Brain for container SSH connections</td></tr> </tbody> </table> <p>Each load balancer needs a check to validate the health of the back-end instances:</p> <ul style="list-style-type: none"> AppsLB checks the health on Gorouter port 80 with TCP SSHLB checks the health on Diego Brain port 2222 with TCP <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> 💡 Note: In many cases, the load balancers are provided as an “under the cloud” service that is shared across projects. </div>	Name	Instance/Port	LB Port	Protocol	Description	AppsLB	gorouter/80	80	HTTP	Forward traffic to Gorouters	AppsLB	gorouter/80	443	HTTPS	SSL termination and forward traffic to Gorouters	SSHLB	diego-brain/2222	2222	TCP	Forward traffic to Diego Brain for container SSH connections	2
Name	Instance/Port	LB Port	Protocol	Description																		
AppsLB	gorouter/80	80	HTTP	Forward traffic to Gorouters																		
AppsLB	gorouter/80	443	HTTPS	SSL termination and forward traffic to Gorouters																		
SSHLB	diego-brain/2222	2222	TCP	Forward traffic to Diego Brain for container SSH connections																		
Jumpbox	<p>Optional. Provides a way of accessing different network components. For example, you can configure it with your own permissions and then set it up to access to Pivotal Network to download tiles. Using a jumpbox is particularly useful in IaaSes where Ops Manager does not have a public IP address. In these cases, you can SSH into Ops Manager or any other component through the jumpbox.</p>	1																				

vSphere Reference Architecture

This guide provides reference architectures for Pivotal Cloud Foundry (PCF), including Pivotal Application Service (PAS) and Pivotal Container Service (PKS), on vSphere.

Overview

Pivotal validates the reference architectures described in this topic against multiple production-grade usage scenarios.

This document does not replace the installation instructions provided in the [PCF on vSphere Requirements](#) topic. Instead, it gives examples of how to apply those instructions to real-world production environments.

PCF Products Validated	Version
vSphere	6.5
VMware NSX-T Data Center	2.2 or later
Pivotal Cloud Foundry Operations Manager	2.2 or later
PAS	2.2 or later
PKS	1.1.5 or later

Base vSphere Reference Architecture

This reference architecture includes VMware vSphere and NSX-T Data Center, a software-defined network virtualization platform that runs on VMware ESXi virtual hosts.

The reference architecture supports capacity growth at the vSphere and PCF levels as well as installation security through the NSX-T firewall. It allocates a minimum of three servers to each vSphere cluster and spreads PCF components across three clusters for high availability (HA). Each vSphere cluster corresponds to an Availability Zone (AZ). For information about HA in PCF, see [High Availability in Cloud Foundry](#).

 **Note:** Small and non-critical deployments can maintain high availability with two vSphere clusters instead of three, provided that the PAS system databases are configured as external. For more information, see [Configuring PAS](#).

To use all features listed in this topic for your PAS installation, you must have **Advanced** or above licensing from VMware for NSX-T. For PKS, the required licensing is included by default.

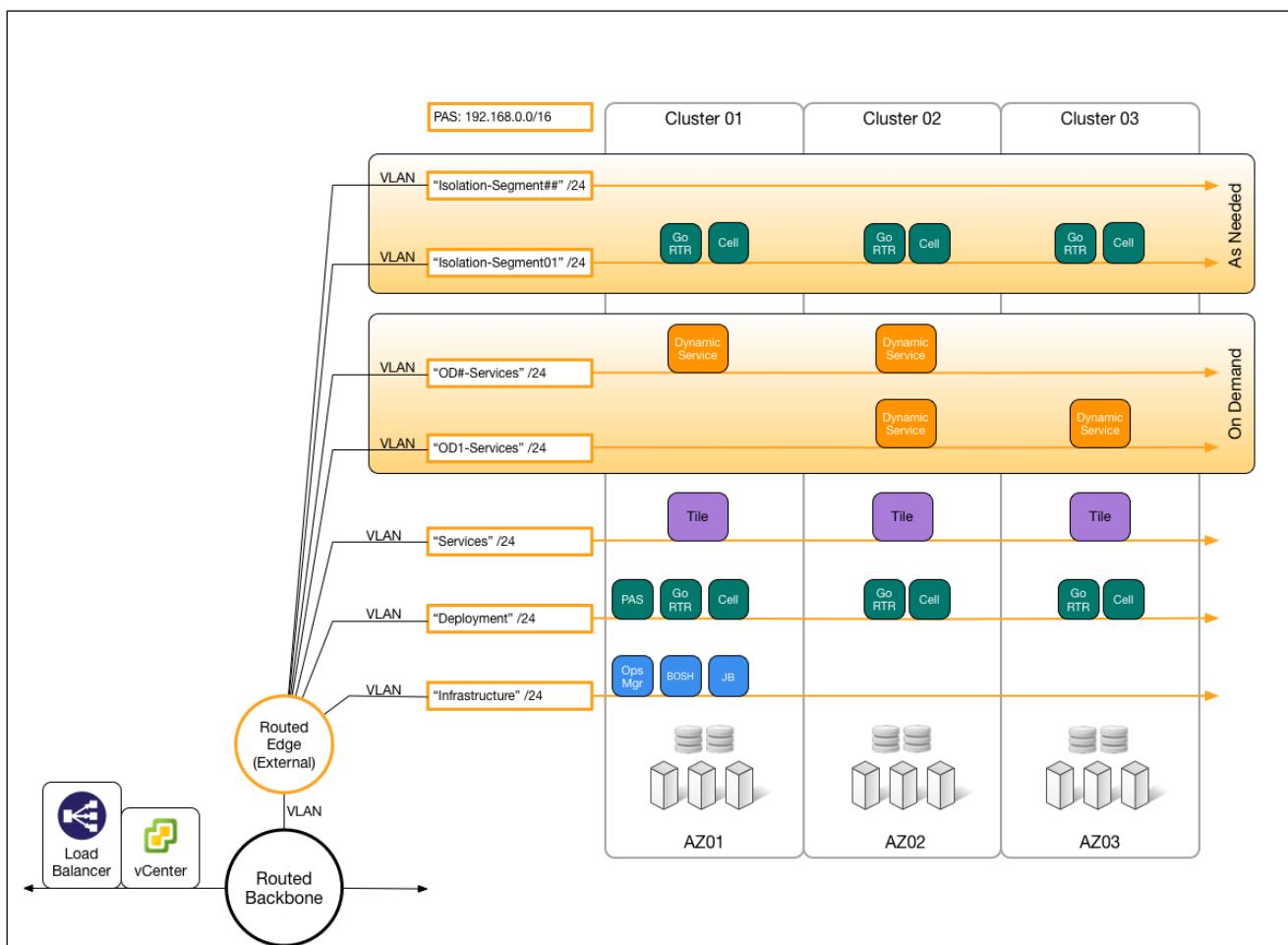
If you want to deploy PCF without NSX-T, see [vSphere Reference Architecture](#), which includes NSX-V design considerations.

Non-SDN Deployments

This section provides information about VLAN and container-to-container networking considerations for PCF deployments that do not use software-defined networking (SDN).

VLAN Considerations

Without an SDN environment, you draw network address space and broadcast domains from the greater data center pool. For this approach, you need to deploy PAS and PKS in a particular alignment.



[View a larger version of this diagram](#)

When designing a non-SDN environment, ensure that VLANs and address space for PAS and PKS do not overlap. In addition, note that non-SDN PCF environments have the following characteristics:

- Load balancing is performed external to the PCF installation.
- Client SSL termination must happen either at the load balancer or Gorouters, or both. Ensure your certificates can accomplish this.
- Firewall layer is accomplished external to the PCF installation.

Container-to-Container Networking and CNI Considerations

Changing your container-to-container networking strategy after deployment is possible. However, with NSX-T, you need both the VMware NSX-T Container Plug-In for PCF tile and NSX-T Data Center. For information about deploying PAS on vSphere with NSX-T internal networking using the VMware NSX-T Container Plug-In for PCF tile, see [Deploying PAS with NSX-T Networking](#).

Migrating from a non-SDN environment to an SDN-enabled solution is possible but best considered as a greenfield deployment. Inserting an SDN layer under an active PCF installation is disruptive.

PAS Without SDN

You can use PAS without an SDN overlay. For more information, see [PCF on vSphere Requirements](#).

PKS Without SDN

NSX-T SDN is included in PKS by default, along with the associated licensing. Alternatively, you can use a built-in network stack, Flannel, which handles container networking.

If you want to deploy PKS without NSX-T, select **Flannel** as your container networking interface in the **Networking** pane of the PKS tile. For information

about configuring your container networking interface, see the [Networking](#) section in *Installing PKS on vSphere*.

In addition, you must define networks for deploying PKS and PKS-provisioned Kubernetes clusters. For information about defining these networks in Ops Manager, see the [Create Networks Page](#) section in *Configuring Ops Manager on vSphere*.

SDN-Enabled Deployments With NSX-T

This section provides information about network requirements, routing, and load balancing for PAS and PKS deployments that use SDN.

Network Requirements for PAS

PAS requires a number of statically defined networks to host the main components it is composed of. For the Tenant side of an NSX-T deployment, a series of non-routable address banks that the NSX-T routers will manage is defined as follows:

 **Note:** These static networks have a smaller host address space assigned than reference designs for PCF v1.xx.

- Infrastructure: `192.168.1.0/24`
- Deployment: `192.168.2.0/24`
- Services: `192.168.3.0/24`

The Services network can be used with other Ops Manager tiles that you install in addition to PAS. Some of these tiles may require on-demand network capacity. Pivotal recommends that you consider adding a network per tile that needs on-demand resources and pair them up in the tile's configuration. For more information, see *OD-Services#* below.

- OD-Services#: `192.168.4.0 - 192.168.9.0` in /24 segments

For example, the Redis for PCF tile asks for Network and Services Network. The first one is for placing the broker, and the second one is for deploying worker VMs to support the service. In this scenario, you can deploy a new network OD-Services1 and instruct Redis for PCF to use the Services network for the broker and the OD-Services1 network for the workers. The next tile can use the Services network for the broker and a new OD-Services2 network for workers and so on.

- Isolation Segments: `192.168.10.0 - 192.168.63.0` in /24 segments

Isolation segments can be added to an existing PAS installation. This range of address space is used when you add one or more segments. A /24-network in this range should be deployed for each new isolation segment. The capacity of this address bank is sufficient for more than 50 isolation segments, each having more than 250 VMs.

- PAS Dynamic Orgs: `192.168.128.0/17` = 128 orgs/foundation

Dynamically assigned Org networks are attached to automatically generated NSX-T Tier-1 routers. Instead of defining these networks in Ops Manager, the operator provides a non-overlapping block of address space. This is configurable in the NCP pane of the VMware NSX-T Container Plug-In for PCF tile in Ops Manager. Every Org receives a new /24 network.

This reference uses a pattern that follows previous references. However, all networks now break on the /24 boundary, and the network octet is numerically sequential (1-2-3).

Network Requirements for PKS

When deploying PKS through Ops Manager, you must allocate a block of address space for dynamic networks that PKS will deploy per namespace.

Network requirements for PKS are as follows:

- PKS Clusters: `172.24.0.0/14`
- PKS Pods: `172.28.0.0/14`

Summary

The complete addressing strategy for both PAS and PKS is as follows:

- Infrastructure: `192.168.1.0/24`
- Deployment: `192.168.2.0/24`
- Services: `192.168.3.0/24`
- OD-Services#: `192.168.4.0 - 192.168.9.0` in /24 segments

- Isolation Segments: `192.168.10.0` - `192.168.63.0` in /24 segments
- Undefined: `192.168.64.0` - `192.168.127.0`
- PAS Dynamic Orgs: `192.168.128.0/17`
- PKS Clusters: `172.24.0.0/14`
- PKS Pods: `172.28.0.0/14`

Network Requirements for External Routing

Routable external IPs on the provider side, for example, for NATs, PAS orgs, and load balancers, are assigned to the T0 router, which is located in front of the PCF installation. There are two approaches to assigning address space blocks to this job:

- Without PKS or with PKS Ingress: The T0 router needs some routable address space to advertise on the BGP network with its peers. Select a network range with ample address space that can be split into two logical jobs: one job is advertised as a route for traffic, and the other job is for aligning T0 DNATs and SNATs, load balancers, and other jobs. Unlike with NSX-V, SNATs consume much more address space than before.
- With PKS No Ingress: Compared to the approach above, this approach has much higher address space consumption for load balancer VIPs. Therefore, allow for 4x the address space because Kubernetes service types allocate addresses frequently.

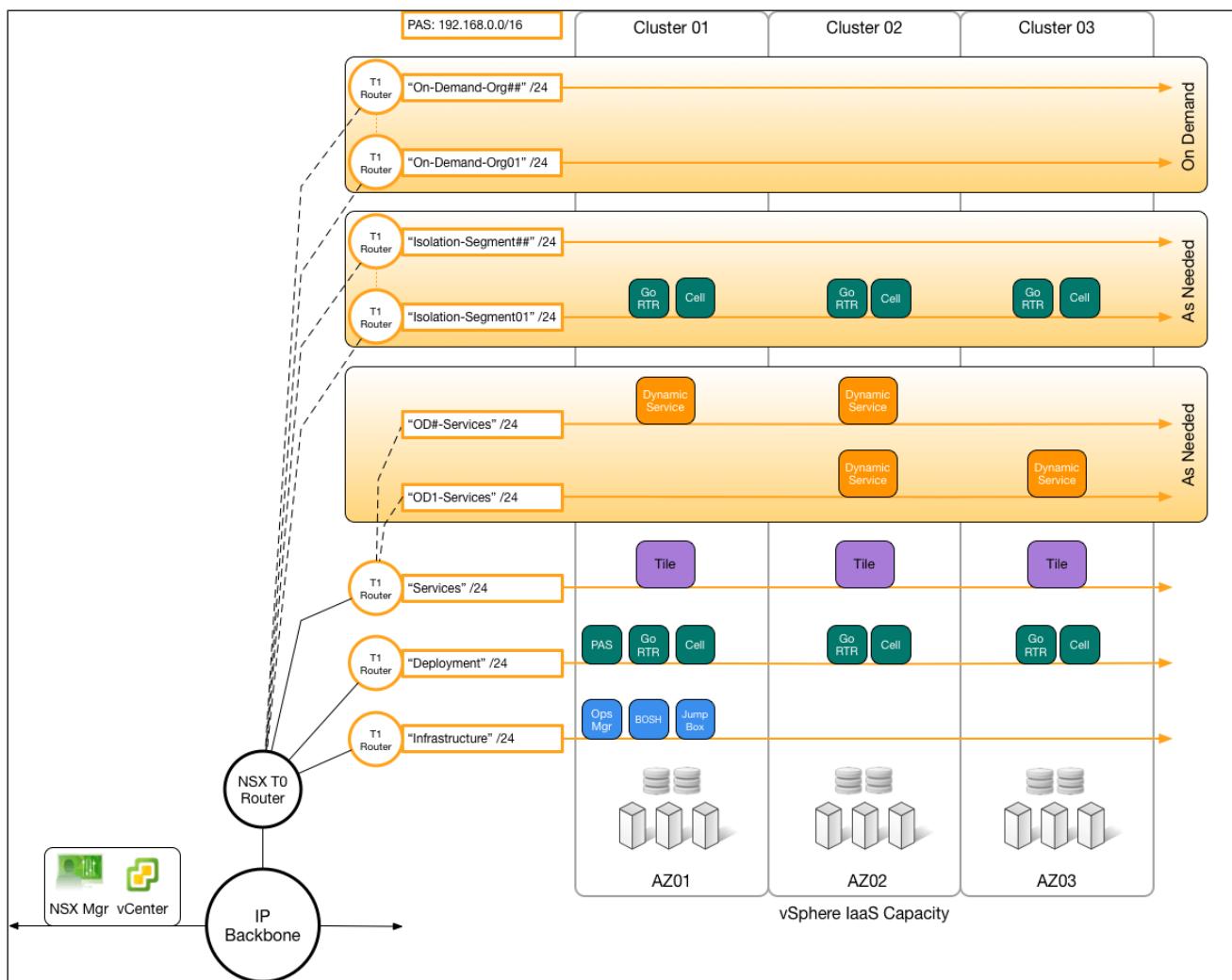
Provider routable address space is /25 or /23 for PKS No Ingress.

NSX-T handles the routing between a T0 router and any T1 routers associated with it.

PAS with NSX-T

Expanding PAS with SDN features is best considered as a greenfield effort. Inserting an SDN layer under a working PAS installation is non-trivial and likely triggers a rebuild. NSX-T constructs that can be used by PAS include the following:

- Logical networks, encapsulated broadcast domains
- VLAN exhaustion avoidance through the use of logical networks
- Routing services and NAT/SNAT to network fence the PCF installation
- Load balancing services to pool systems such as Gorouters
- SSL termination at the load balancer
- Distributed routing and firewall services at the hypervisor



[View a larger version of this diagram](#)

The VMware NSX-T Container Plug-In for PCF tile, which provides a container networking stack instead of the built-in Silk solution, interlocks with NSX-T already deployed on the IaaS. This tile cannot be used unless NSX-T has already been established on the IaaS. These technologies cannot be retrofitted into an established PKS or PAS installation.

To deploy PAS with NSX-T, you need to provide the following in the VMware NSX-T Container Plug-In for PCF tile:

- NSX Manager host
- Username and password
- NSX Manager CA certificate
- PAS foundation name that must match the tag added to T0 router, external NAT pool, and IP block
- Subnet prefix, which controls the size of each org and defaults to /24, that is, 254 addresses

In addition, you must select **Enable SNAT for Container Networks** in the tile. For more information about configuring PAS with NSX-T networking, see [Deploying PAS with NSX-T Networking](#).

Each new job, such as an isolation segment, falls to a broadcast domain or logical switch connected to a T1 router acting as the gateway to that network. This approach provides a DNAT and SNAT control point and a firewall boundary.

The Services network is an exception. It shares a T1 router with any associated OD-Services# networks. Because these networks are considered part of the same system, inserting any NATs or firewall rules between them is not needed.

Load Balancing for PAS

Without NSX-T, you need to choose a suitable load balancer to send traffic to the Gorouters and other systems. Installations approaching production level typically use external load balancing from hardware appliance vendors or other network-layer solutions.

With NSX-T, load balancing is available in the SDN layer. These load balancers are a logical entity tied to the resources in the Edge Cluster and align to the network or networks represented by a T1 router. They function as a Layer 4 load balancer. SSL termination is available on the NSX-T load balancer. However, Pivotal recommends passing the SSL through to your Gorouters.

Common deployments of load balancing in PAS are as follows:

- HTTP/HTTPS traffic to and from Gorouters
- TCP traffic to and from TCP routers
- Diego Brain

NSX-T load balancers can support many VIPs. Consider deploying one load balancer per network (T1) and one-to-many VIPs on that load balancer per job. Edge Cluster resources are consumed for load balancing. When designing your deployment, consider how many load balancers are needed and how much capacity is available for them.

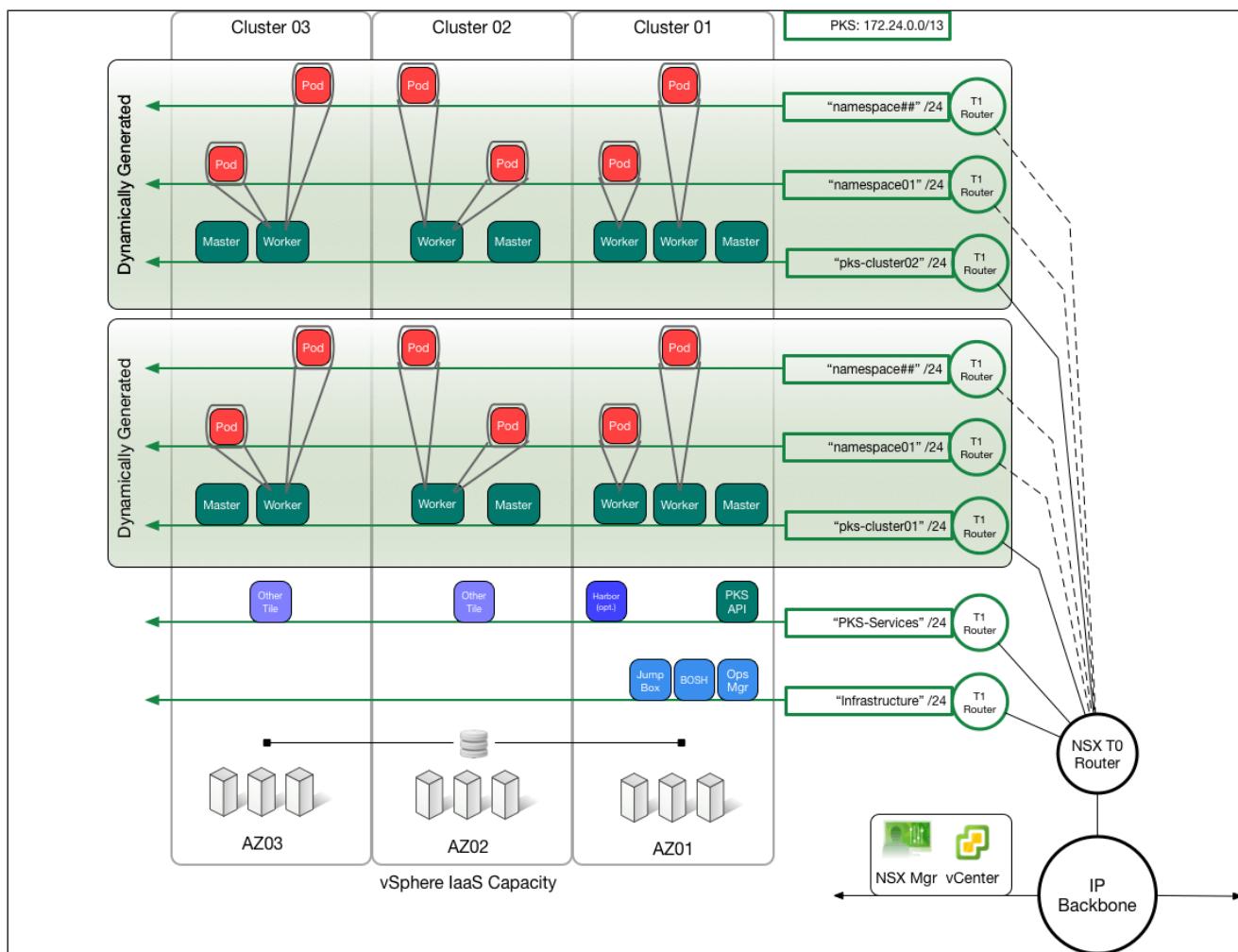
BOSH can manage the members of the server pools for the NSX-T load balancers using NS Groups.

PKS with NSX-T

NSX-T SDN is included in PKS by default, along with the associated licensing. To use NSX-T SDN and its dynamic constructs, you need to provide the following information when configuring PKS:

- NSX Manager host
- Username and password
- NSX Manager CA certificate
- T0 router to connect dynamically created namespace networks to
- NSX IP block from which to pull networks for pods per namespace
- NSX IP block from which to pull networks for each new cluster
- Floating IP pool ID created for externally facing IPs (NATs and VIPs)

For more information about configuring PKS with NSX-T networking, see [Installing and Configuring PKS with NSX-T Integration](#).



[View a larger version of this diagram](#)

New T1 routers are deployed on-demand as new namespaces are added to PKS. Because these can grow rapidly, allocate a large Pod IP block such as /14 in NSX-T, and then NSX-T provisions /24 address blocks for new namespaces. Reference the Pod IP block from your PKS tile configuration.

PKS v1.1.5 and later supports multiple master nodes. The number of master nodes is defined per plan in the PKS tile. You must use an odd number of master nodes to allow etcd to form a quorum. Pivotal recommends using at least 1 master node per AZ for HA and disaster recovery.

With NSX-T SDN, networks for both PKS clusters and pods are created dynamically. Pivotal recommends using multiple clusters instead of a single cluster with multiple namespaces. Multiple clusters provide security, customization per cluster, privileged containers, failure domains, and version choice.

Ingress Routing and Load Balancing for PKS

This section provides information about ingress routing (Layer 7) and load balancing (Layer 4) for your PKS deployment.

Ingress Routing

NSX-T provides a native ingress router. Third-party options include Istio or Nginx that are running as containers in the cluster.

Wildcard DNS entries are needed for pointing at the ingress service in the style of Gorouters in PAS. Domain information for ingress is defined in the manifest of your Kubernetes deployment. See the example below.

```
apiVersion: extensions/v1beta1
kind: Ingress
metadata:
  name: music-ingress
  namespace: music1
spec:
  rules:
    - host: music1.pks.domain.com
      http:
        paths:
          - path: /*
            backend:
              serviceName: music-service
              servicePort: 8080
```

Load Balancing

When pushing a Kubernetes deployment with `type` set to `LoadBalancer`, NSX-T automatically creates a new VIP for the deployment on the existing load balancer for that namespace.

You need to specify a listening and translation port in the service, along with a name for tagging. You also specify a protocol to use. See the example below.

```
apiVersion: v1
kind: Service
metadata:
  ...
spec:
  type: LoadBalancer
  ports:
    - port: 80
      targetPort: 8080
      protocol: TCP
      name: web
```

Storage Design

Shared storage is a requirement for PCF. You can allocate networked storage to the host clusters following one of two common approaches: horizontal or vertical. The approach you follow should reflect how your data center arranges its storage and host blocks in its physical layout.

The following describes the vertical and horizontal approaches to configuring shared storage:

- Vertical: You grant each cluster its own dedicated datastores, creating a cluster-aligned storage strategy. vSphere VSAN is an example of this architecture. The vertical alignment is the preferred choice for PCF as it matches the AZ model of the PaaS. Loss of storage in a cluster constitutes loss on an AZ, which is a recoverable failure when at least two more AZs are operational.
For example, with six datastores `ds01` through `ds06`, you assign datastores `ds01` and `ds02` to your first cluster, `ds03` and `ds04` to your second cluster, and `ds05` and `ds06` to your third cluster. You then instruct your first PCF installation to use `ds01`, `ds03`, and `ds05` and your second PCF installation to use `ds02`, `ds04`, and `ds06`. In this arrangement, all VMs in the same installation and cluster share a dedicated datastore.
- Horizontal: You grant all hosts access to all datastores and assign a subset to each installation.
For example, with six datastores `ds01` through `ds06`, you grant all nine hosts access to all six datastores. You then instruct your first PCF installation to use stores `ds01` through `ds03` and your second PCF installation to use `ds04` through `ds06`.

⚠ Warning: If you use the horizontal storage design approach, any storage loss could affect datastores in all AZs. This could result in downtime and unrecoverable loss of data. To avoid downtime and data loss, Pivotal recommends implementing the vertical storage design approach or implementing strong redundancy design at the storage array.

Highly redundant storage systems, such as hyperconverged systems, are the optimal choice for a fully HA PaaS. In addition, hyperconverged systems scale with capacity growth and are aligned to the AZ strategy of PCF.

To improve the resiliency of your deployment, you can use separate storage for the management plane of PCF, Ops Manager and BOSH, as well as separate storage for the blobstore.

Note: If a datastore is part of a vSphere storage cluster using sDRS (Storage DRS), you must disable the s-vMotion feature on any datastores used by PCF. Otherwise, s-vMotion activity can rename independent disks and cause BOSH to malfunction. For more information, see [How to Migrate PCF to a New Datastore in vSphere](#).

Storage Capacity

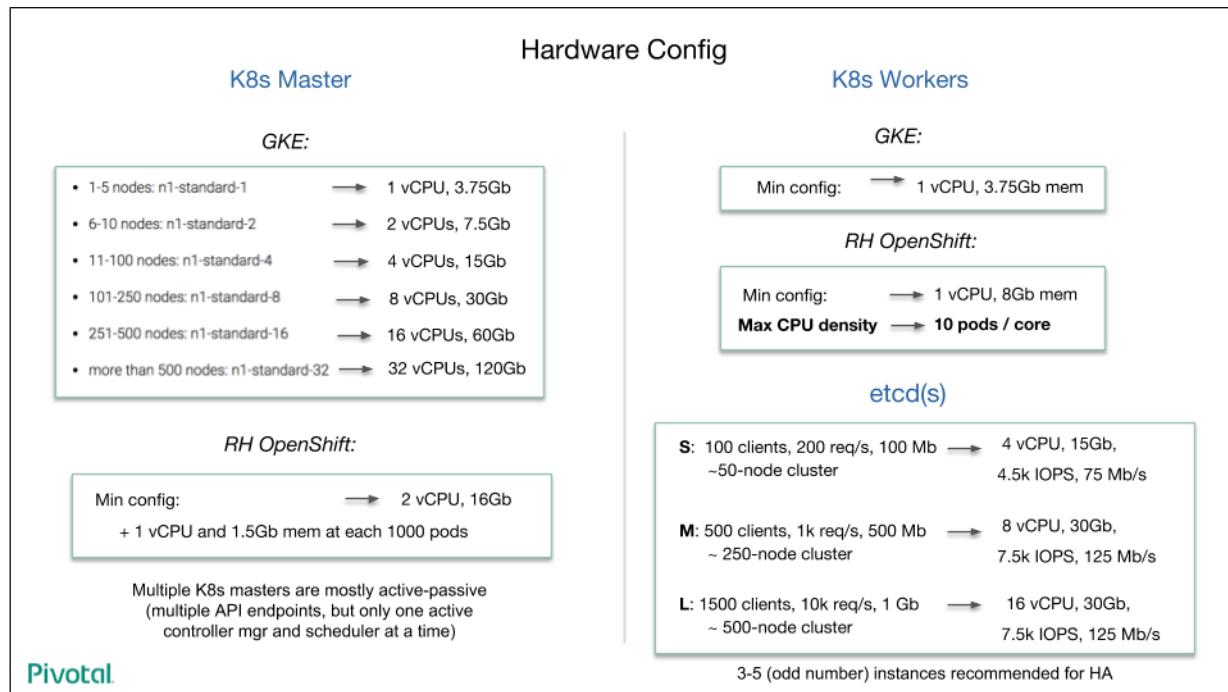
Pivotal recommends the following capacity allocation for PAS installations:

- For production use, at least 8 TB of data storage, either as one 8-TB store or a number of smaller volumes adding up to 8 TB. Frequent development may require significantly more storage to accommodate new code and buildpacks.
- For small installations, 4-6 TB of data storage is recommended.

The primary consumer of storage is NFS or WebDAV.

Note: PCF does not currently support using vSphere storage clusters with the versions of PCF validated for this reference architecture. Datastores should be listed in the vSphere tile by their native name, not the cluster name created by vCenter for the storage cluster.

When planning storage allocation for PKS installations based on Pod and Node needs, consider the following chart.



Compute and HA Considerations

In PAS, a consolidation ratio for containers should follow a conservative 4:1 ratio of vCPUs to pCPUs. You can use a more conservative ratio, which is 2:1.

A standard Diego cell is 4x16 (XL) by default. At a 4:1 ratio, you have a 16-vCPU budget, or about 12 containers.

If you want to run more containers in a cell, scale the vCPUs accordingly. However, high core count VMs become increasingly hard to schedule on the pCPU without high physical core counts in the socket.

HA considerations include the following:

- For non-production environments, the number of AZs that PAS requires for HA depends on the location of its system databases.
 - With external databases, PAS requires at least two AZs for HA.
 - With internal system databases, PAS requires at least three AZs for HA.
- For any production environment, Pivotal recommends using at least three AZs for HA.
- PKS has no inherent HA capabilities to design for. To support PKS, design HA at the IaaS, storage, power, and access layers.

PCF achieves redundancy through the AZ construct, and the loss of an AZ is not considered catastrophic. BOSH Resurrector replaces lost VMs as needed to repair a foundation.

Foundation backup and restoration is accomplished externally by BOSH Backup and Restore (BBR). For information about BBR, see [Backing Up and Restoring Pivotal Cloud Foundry](#).

Scaling and Capacity Management

This section provides information about scaling and capacity management for PCF.

Small PCF Installations

A small-sized PCF foundation looks as follows:

- 1 vSphere cluster/1 AZ
- 2 resource pools: 1 for NSX components and 1 for PAS or PKS
- 3 hosts minimum for vSphere HA (4 hosts for vSphere VSAN)
- Shared storage/VSAN
- 1 NSX Manager
- 3 NSX controllers
- 2 large edge VMs in a cluster
- PCF: Ops Manager, the BOSH Director, and Small Footprint Runtime

This approach is intended for designing a proof of concept or development-only system. Small Footprint Runtime has no upgrade path to the standard PAS tile.

Medium PCF Installations

A medium-sized PCF foundation looks as follows:

- 2 vSphere clusters/2 AZs
- 3 resource pools: 1 for NSX components and 2 for PAS/1 for PKS
- 3 hosts minimum for vSphere HA (4 hosts for vSphere VSAN) per cluster
- Shared storage/VSAN
- 1 NSX Manager
- 3 NSX controllers
- 2 large edge VMs in a cluster
- PCF: Ops Manager, the BOSH Director, and PAS/PKS

For this design, Pivotal recommends replacing Small Footprint Runtime with PAS. The second AZ doubles compute capacity and expands the NSX-T footprint.

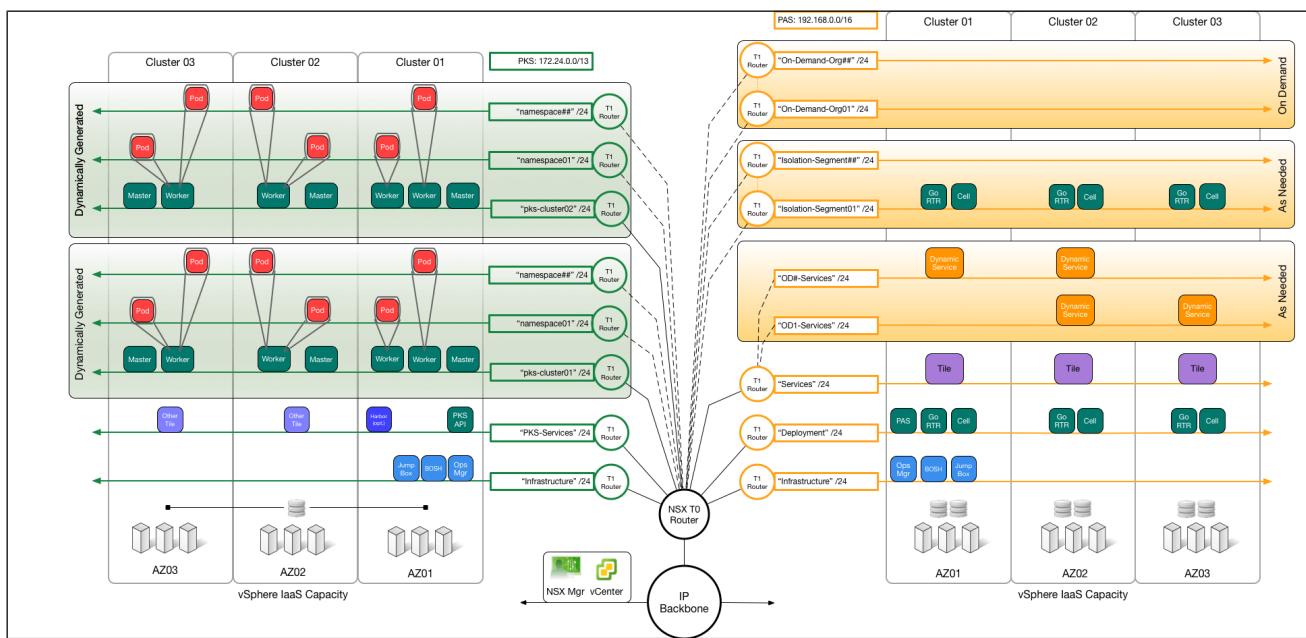
Production-Ready PCF Installations

A production-ready PCF foundation looks as follows:

- 3 vSphere clusters/3 AZs
- 4 resource pools: 1 for NSX components and 3 for PAS/1 for PKS
- 3 hosts minimum for vSphere HA (4 hosts for vSphere VSAN)
- Shared Storage/VSAN
- 1 NSX Manager
- 3 NSX Controllers
- 3 large edge VMs in a cluster
- PCF: Ops Manager, the BOSH Director, and PAS/PKS

PAS and PKS with NSX-T

A fully meshed installation of PCF v2.2 or later designed using the recommendations provided in this topic looks as follows:



[View a larger version of this diagram](#)

Common components are the NSX T0 router and the associated T1 routers. This approach ensures that any cross-traffic between PKS and PAS apps stays within the bounds of the T0. This also provides a one-stop access point to the whole installation, which simplifies deployment automation for multiple, identical installations.

AZs are aligned to vSphere clusters, with resource pools as an optional organizational tool to place multiple foundations into the same capacity. You can align PKS to any AZ or cluster. Keeping PKS and PAS in completely separate AZs is not required.

You can use resource pools in vSphere clusters as AZ constructs to stack different installations of PCF. As server capacity continues to increase, the efficiency of deploying independent server clusters only for one installation is low. For example, customers are commonly deploying servers with 768-GB RAM and greater.

To allow for max capacity growth, consider using an NSX-T installation per foundation.

If you want to split the PAS and PKS installations into separate network trees, behind separate T0 routers, ensure that approach meets your needs by reviewing VMware's recommendations for T0 to T0 routing. For more information, see [Reference Design Guide for PAS and PKS with VMware NSX-T Data Center](#) in the VMware documentation.

Using Edge Services Gateway on VMware NSX

This cookbook provides guidance on how to configure the NSX firewall, load balancing, and NAT/SNAT services for Pivotal Cloud Foundry (PCF) on vSphere installations. These NSX-provided services take the place of an external device or the bundled HAProxy VM in PCF.

This document presents the reader with fundamental configuration options of an Edge Services Gateway (ESG) with PCF and vSphere NSX. Its purpose is not to dictate the settings required on every deployment, but instead to empower the NSX Administrator with the ability to establish a known good “base” configuration and apply specific security configurations as required.

If you are using NSX, the specific configurations described here supersede any general recommendations in the [Preparing Your Firewall](#) topic.

Assumptions

This document assumes that the reader has the level of skill required to install and configure the following products:

- VMware vSphere v5.5 or greater
- NSX v6.1.x or greater
- PCF v1.6 or greater

For detailed installation and configuration information about these products, see the following:

- [vSphere Documentation ↗](#)
- [NSX Installation and Upgrade Guide ↗](#)
- [Reference Design: VMware NSX for vSphere \(NSX\) Network Virtualization Design Guide ↗](#)
- [Pivotal Cloud Foundry Documentation ↗](#)

General Overview

This cookbook follows a three-step recipe to deploy PCF behind an ESG:

1. Configure Firewall
2. Configure Load Balancer
3. Configure NAT/SNAT

The ESG can scale to accommodate very large PCF deployments as needed.

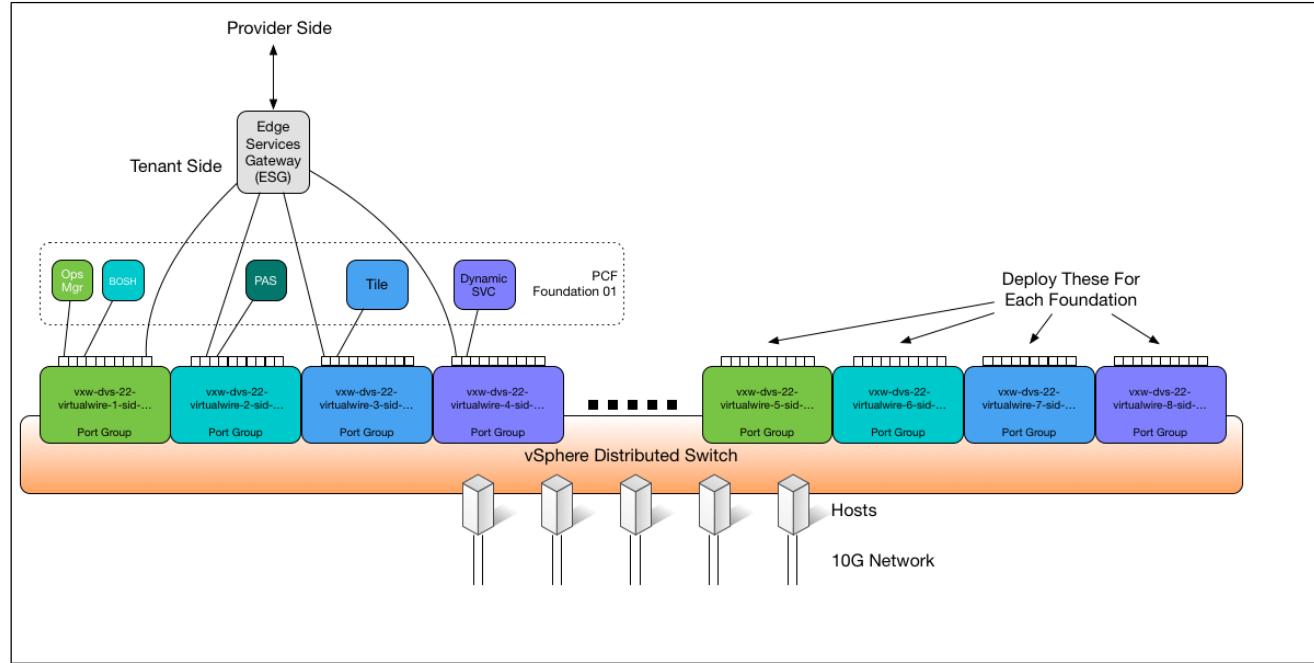
This cookbook focuses on a single-site deployment and makes the following design assumptions:

- There are five non-routable networks on the tenant (inside) side of the ESG.
 - The **Infra** network is used to deploy Ops Manager and BOSH Director.
 - The **Deployment** network is used exclusively by Pivotal Application Service (PAS) to deploy Cells that host apps and related elements.
 - The **CF Tiles** network is used for all other deployed tiles in a PCF installation.
 - The **Services** network is used by BOSH Director for service tiles.
 - The **Container-to-Container** network is used for container to container communication in the Cells.
- There is a single service provider (outside) interface on the ESG that provides Firewall, Load Balancing and NAT/SNAT services.
- The service provider (outside) interface is connected appropriately to the network backbone of the environment, as either routed or non-routed depending on the design. This cookbook does not cover provisioning of the uplink interface.
- Routable IP addresses should be applied to the service provider (outside) interface of the ESG. Pivotal recommends that you apply 10 consecutive routable IP addresses to each ESG.
 - One reserved for NSX use (Controller to Edge I/F)
 - One for NSX Load Balancer to Gorouters
 - One for NSX Load Balancer to Diego Brains for SSH to apps
 - One routable IP address, used to access the Ops Manager frontend
 - One routable IP address, used with SNAT egress
 - Five for future use

Pivotal recommends that operators deploy the ESGs as high availability (HA) pairs in vSphere. Also, Pivotal recommends that they be sized “large” or greater for any pre-production or production use. The deployed size of the ESG impacts its overall performance, including how many SSL tunnels it can terminate.

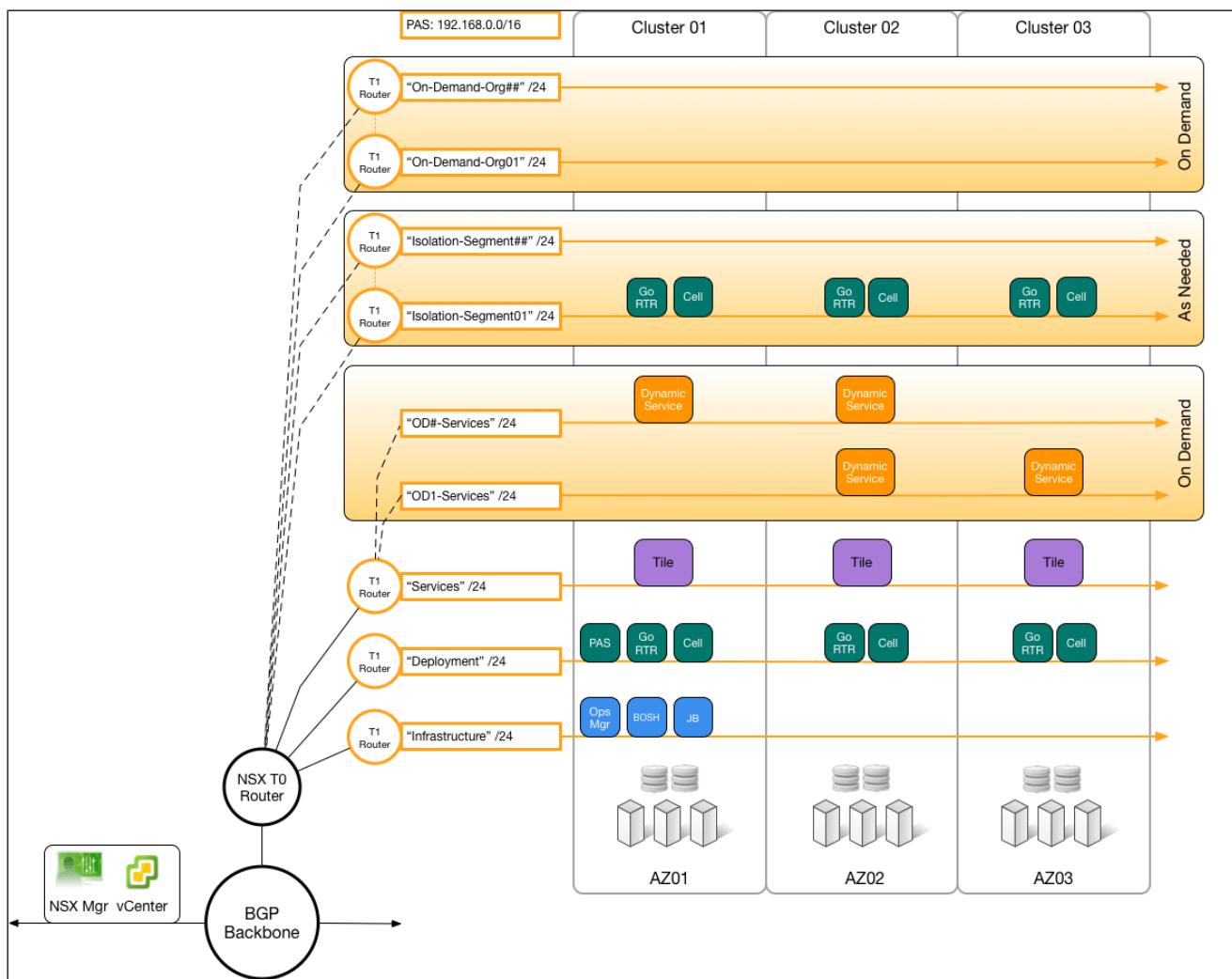
The ESGs have an interface in each port group used by PCF as well as a port group on the service provider (outside), often called the “transit network.” Each PCF installation has a set of port groups in a vSphere DVS to support connectivity, so that the ESG arrangement is repeated for every PCF install. It is not necessary to build a DVS for each ESG/PCF install. You do not re-use an ESG amongst PCF deployments. NSX Logical Switches (VXLAN vWires) are ideal candidates for use with this architecture.

The following diagram provides an example of port groups used with an ESG:



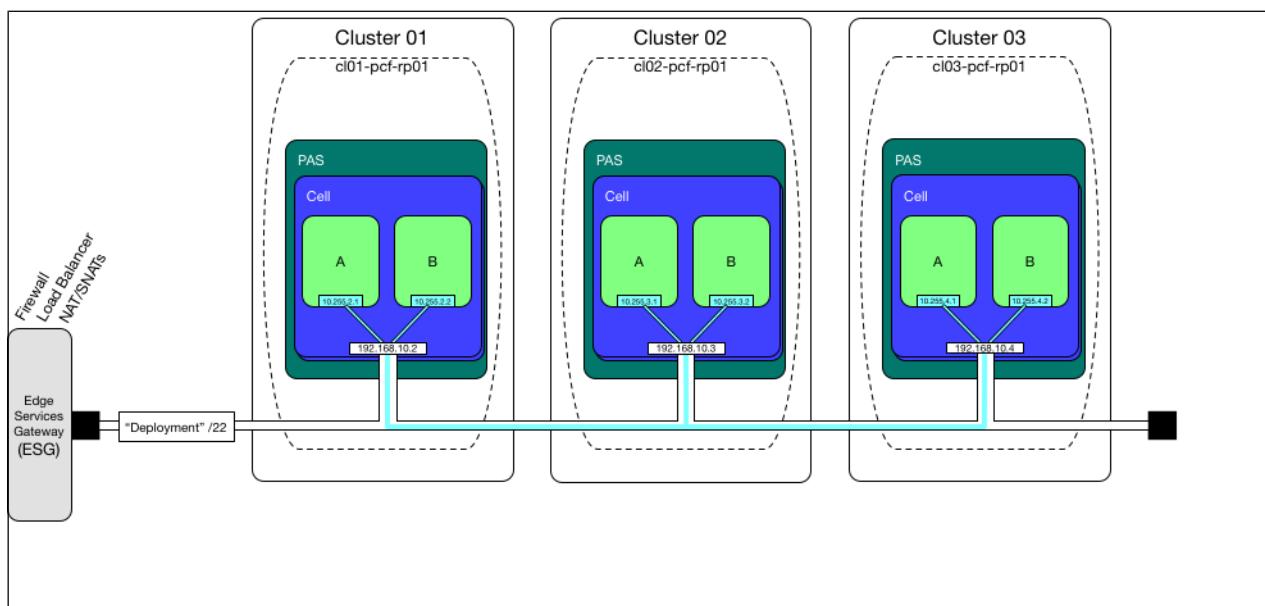
[View a larger version of this diagram.](#)

The following is an example of a network architecture deployment.



[View a larger version of this diagram.](#)

The following diagram illustrates container-to-container networking. The overlay addresses are wrapped and transported using the underlay deployment subnet.



[View a larger version of this diagram.](#)

Prep Step: Configure DNS and Network Prerequisites

As a prerequisite, create wildcard DNS entries for system and apps domains in PCF. Map these domains to the selected IP address on the uplink (outside) interface of the ESG in your DNS server.

The wildcard DNS **A** record must resolve to an IP address associated with the outside interface of the ESG for it to function as a load balancer. You can either use a single IP address to resolve both the system and apps domain, or one IP address for each.

In addition, assign the following IP addresses and address ranges within your network:

1. Assign IP Addresses to the “Uplink” (outside) interface
 - Typically you have one SNAT and three DNATs per ESG.
 - IP associated for SNAT use: All PCF internal IP addresses appear to be coming from this IP address at the ESG.
 - IP associated with Ops Manager DNAT: This IP address is the publicly routable interface for Ops Manager UI and SSH access.
2. Assign “Internal” Interface IP Address Space to the Edge Gateway.
 - 192.168.10.0/26 = PCF Deployment Network (Logical Switch or Port Group)
 - 192.168.20.0/22 = Deployment Network for PAS tile
 - 192.168.24.0/22 = CF Tiles Network for all Tiles besides PAS
 - 192.168.28.0/22 = Dynamic Services network for BOSH Director-managed service tiles.
 - 10.255.0.0/16 = Container-to-Container network for intercontainer communication.

You must update the security group and load balancer information for your PCF deployments using NSX-V on vSphere through the Ops Manager API. See [Updating NSX Security Group and Load Balancer Information](#) for more information.

Step 1: Configure Firewall

This procedure populates the ESG internal firewall with rules to protect a PCF installation.

These rules provide granular control on what can be accessed within a PCF installation. For example, rules can be used to allow or deny another PCF installation behind a different ESG access to apps published within the installation you are protecting.

This step is not required for the installation to function properly when the firewall feature is disabled or set to “Allow All.”

To configure the ESG firewall, navigate to **Edge, Manage, Firewall** and set the following:

Name	Source	Destination	Service	Action
Allow Ingress -> Ops Manager	any	IP_of_OpsMgr	SSH, HTTP, HTTPS	Accept
Allow Ingress -> PAS	any	IP_of_NSX-LB	HTTP, HTTPS	Accept
Allow Ingress -> SSH for Apps	any	tcp:IP_of_DiegoBrain:2222	any	Accept
Allow Ingress -> TCPRouter	any	tcp:IP_of_NSX-TCP-LB:5000	any	Accept
Allow Inside <-> Inside	192.168.10.0/26 192.168.20.0/22 192.168.24.0/22 192.168.28.0/22	192.168.10.0/26 192.168.20.0/22 192.168.24.0/22 192.168.28.0/22	any	Accept
Allow Egress -> IaaS	192.168.10.0/26	IP_of_vCenter IPs_of_ESXi-Svrs	HTTP, HTTPS	Accept
Allow Egress -> DNS	192.168.0.0/16	IPs_of_DNS	DNS, DNS-UDP	Accept
Allow Egress -> NTP	192.168.0.0/16	IPs_of_NTP	NTP	Accept
Allow Egress -> SYSLOG	192.168.0.0/16	IPs_of_Syslog:514	SYSLOG	Accept
Allow ICMP	192.168.10.0/26	*	ICMP	Accept
Allow Egress -> LDAP	192.168.10.0/26 192.168.20.0/22	IPs_of_LDAP:389	LDAP, LDAP-over-ssl	Accept
Allow Egress -> All Outbound	192.168.0.0/16	any	any	Accept
Default Rule	any	any	any	Deny

Step 2: Configure Load Balancer

The ESG provides software load balancing functionality, equivalent to the bundled HAProxy that is included with PCF, or hardware appliances such as an F5 or A10 load balancer.

This step is required for the installation to function properly.

There are seven high level steps to this procedure:

1. Import SSL certificates to the Edge for SSL termination.
2. Enable the load balancer.
3. Create Application Profiles in the Load Balancing tab of NSX.
4. Create Application Rules in the load balancer.
5. Create Service Monitors for each pool type.
6. Create Application Pools for the multiple groups needing load balancing.
7. Create a virtual server (also known as a VIP) to pool balanced IP addresses.

What you will need:

- PEM files of SSL certificates provided by the certificate supplier for only this installation of PCF, or the self-signed SSL certificates generated during PCF installation.

In this procedure you marry the ESG's IP address used for load balancing with a series of internal IP addresses provisioned for Gorouters in PCF. It is important to know the IP addresses used for the Gorouters beforehand.

These IP addresses can be pre-selected or reserved prior to deployment (recommended) or discovered after deployment by looking them up in BOSH Director, which lists them in the release information of the PAS installation.

Step 2.1: Import SSL Certificate

PCF requires SSL termination at the load balancer.

 **Note:** If you intend to pass SSL termination through the load balancer directly to the Gorouters, you can skip the step below and select **Enable SSL Passthru** on the **HTTPS Application Profile**.

To enable SSL termination at the load balancer in ESG, access the ESG UI and do the following:

1. Select **Edge**, **Manage**, **Settings**, and then **Certificates**.
2. Click Green Plus button to Add Certificate.
3. Insert PEM file contents from the Networking configuration screen of PAS.
4. Save the results.

Step 2.2: Enable the Load Balancer

To enable the load balancer, access the ESG UI and do the following:

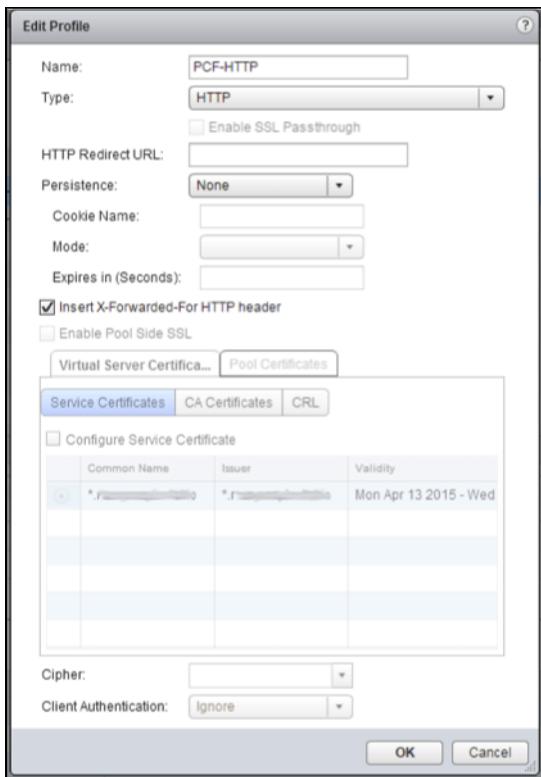
1. Select **Edge**, **Manage**, **Load Balancer**, and then **Global Configuration**.
2. Edit load balancer global configuration.
3. Enable load balancer.
4. Enable acceleration.
5. Set logging to desired level (**Info** or greater).

Step 2.3: Create Application Profiles

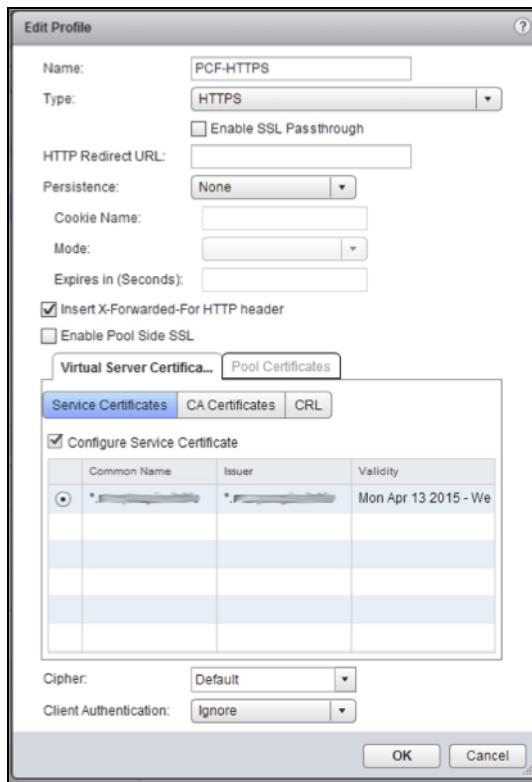
The Application Profiles allow advanced `x-forwarded` options as well as linking to the SSL Certificate. You must create three Profiles: **PCF-HTTP**, **PCF-HTTPS** and **PCF-TCP**.

To create the application profiles, access the ESG UI and do the following:

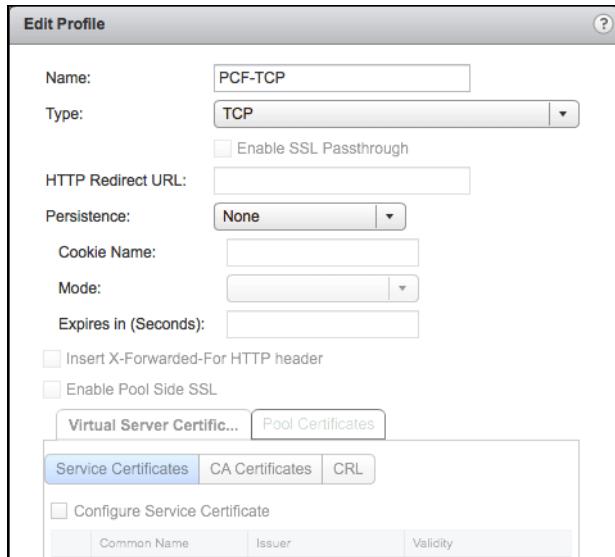
1. Select **Edge**, **Manage**, **Load Balancer**, and then **Global Application Profiles**.
2. Create/Edit Profile and make the **PCF-HTTP** rule, turning on **Insert X-Forwarded-For HTTP header**.



3. Create/Edit Profile and make the **PCF-HTTPS** rule, same as before, but add the service certificate inserted before. If encrypting TLS traffic to the Gorouters, turn on **Enable Pool Side SSL**. Otherwise, leave it unchecked.



4. Create/Edit Profile and make PCF-TCP rule, with the Type set to TCP.



Step 2.4: Create Application Rules

In order for the ESG to perform proper `x-forwarded` requests, you need to add a few HAProxy directives to the ESG Application Rules. NSX supports most directives that HAProxy supports.

To create the application rules, access the ESG UI and do the following:

1. Select **Edge**, **Manage**, **Load Balancer**, and then **Application Rules**.
2. Copy and paste the table entries below into each field, one per rule.

Rule Name	Script
option httplog	option httplog
reqadd X-Forwarded-Proto:\ https	reqadd X-Forwarded-Proto:\ https

Rule Id	Name	Script
applicationRule-1	option httplog	option httplog
applicationRule-2	option forwardfor	option forwardfor
applicationRule-3	reqadd X-Forwarded-Proto:\ https	reqadd X-Forwarded-Proto:\ https

Step 2.5: Create Monitors For Pools

NSX ships with several load balancing monitoring types pre-defined. These types are for HTTP, HTTPS and TCP. For this installation, operators build new monitors matching the needs of each pool to ensure correct 1:1 monitoring for each pool type.

To create monitors for pools, access the ESG UI and do the following:

1. Select **Edge**, **Manage**, **Load Balancer**, and then **Service Monitoring**.
2. Create a new monitor for `http-routers`, and keep the defaults.
3. Set the Type to `HTTP`.
4. Set the Method to `GET`.
5. Set the URL to `/health`.
6. Create a new monitor for `tcp-routers`, and keep the defaults.
7. Set the type to `HTTP`.
8. Set the Method to `GET`.
9. Set the URL to `/health`.
10. Create a new monitor for `diego-brains`, and keep the defaults.
11. Set the type to `TCP`.

These monitors are selected during the next step when pools are created. A pool and a monitor are matched 1:1.

Step 2.6: Create Pools of Multi-Element PCF Targets

The following steps creates the pools of resources that ESG load-balances to: the Gorouter, TCP Router, and Diego Brain jobs deployed by BOSH Director. If the IP addresses specified in the configuration do not exactly match the IP addresses reserved or used for the resources, then the pool will not effectively load balance.

Step 2.6a: Create Pool for `http-routers`

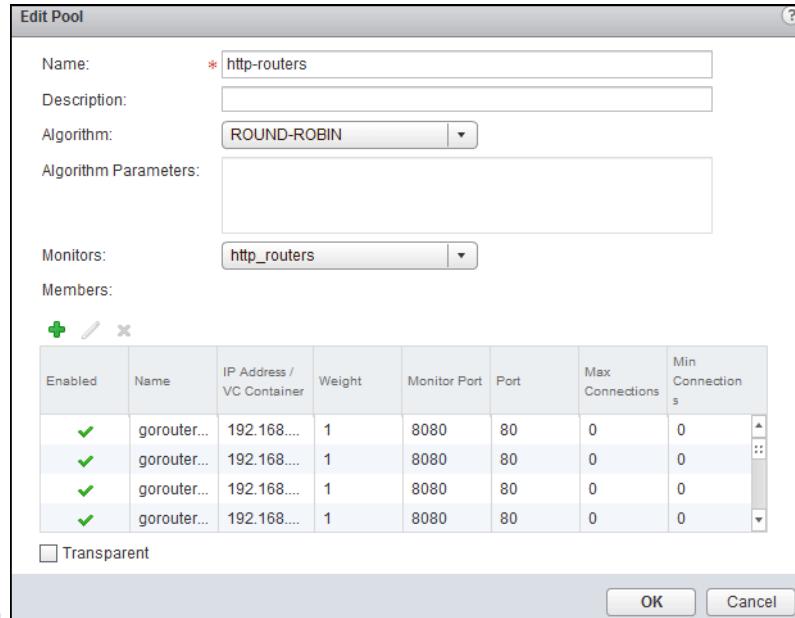
To create pool for `http-routers`, access the ESG UI and do the following:

1. Select **Edge**, **Manage**, **Load Balancer**, and then **Pools**.

- Enter ALL the IP addresses reserved for the Gorouters into this pool. If you reserved more addresses than you have Gorouters, enter the addresses anyway and the load balancer ignores the missing resources as “down”.

Note: If your deployment matches the [Reference Architecture for PCF on vSphere](#), these IP addresses are in the 192.168.20.0/22 address space.

- Set the **Port** to 80 and **Monitor Port** to 8080. The assumption is that internal traffic from the ESG load balancer to the Gorouters is trusted because it is on a VXLAN secured within NSX. If using encrypted TLS traffic to the Gorouter inside the VXLAN, set the **Port** to 443.
- Set the **Algorithm** to `ROUND-ROBIN`.



- Set **Monitors** to `http-routers`.

Step 2.6b: Create Pool for `tcp-routers`

- Select **Edge**, **Manage**, **Load Balancer**, and then **Pools**.
- Enter ALL the IP addresses reserved for TCP Routers into this pool. If you reserved more addresses than you have VMs, enter the addresses anyway and the load balancer ignores the missing resources as “down”.

Note: If your deployment matches the [Reference Architecture for PCF on vSphere](#), these IP addresses are in the 192.168.20.0/22 address space.

- Set the **Port** to empty (these numbers vary) and the **Monitor Port** to 80.
- Set the **Algorithm** to `ROUND-ROBIN`.
- Set the **Monitors** to `tcp-routers`.

Step 2.6c: Create Pool for `diego-brains`

- Select **Edge**, **Manage**, **Load Balancer**, and then **Pools**.
- Enter ALL the IP addresses reserved for Diego Brains into this pool. If you reserved more addresses than you have VMs, enter the addresses anyway and the load balancer will just ignore the missing resources as “down”.

Note: If your deployment matches the [Reference Architecture for PCF on vSphere](#), these IP addresses are in the 192.168.20.0/22 address space.

- Set the **Port** to 2222 and the **Monitor Port** to 2222.
- Set the Algorithm to `ROUND-ROBIN`.

- Set the Monitors to `diego-brains`.

Step 2.7: Create Virtual Servers

This is the Virtual IP (VIP) that the load balancer uses to represent the pool of Gorouters to the outside world. This also links the Application Policy, Application Rules, and backend pools to provide PCF load balancing services. This is the interface that the load balancer balances **from**. You create three Virtual Servers.

- Select **Edge**, **Manage**, **Load Balancer**, and then **Virtual Servers**.
- Select an IP address from the available routable address space allocated to the ESG. For information about reserved IP addresses, see [General Overview](#).
- Create a new Virtual Server named `GoRtr-HTTP` and select Application Profile `PCF-HTTP`.
 - Use **Select IP Address** to select the IP address to use as a VIP on the uplink interface.
 - Set **Protocol** to match the **Application Profile** protocol (HTTP) and set **Port** to match the protocol (80).
 - Set **Default Pool** to the pool name set in [Step 2.6a: Create Pool for http-routers](#) (`http-routers`). This connects this VIP to the pool of resources being balanced to.
 - Ignore **Connection Limit** and **Connection Rate Limit** unless these limits are desired.
 - Switch to **Advanced Tab** on this Virtual Server.
 - Use the green plus to add/attach three Application Rules to this Virtual Server:
 - option `httplog`
 - `reqadd X-Forwarded-Proto:\ http`

 **Note:** Be careful to match protocol rules to the protocol `VIP- HTTP` to `HTTP` and `HTTPS` to `HTTPS`.

- Create a new Virtual Server named `GoRtr-HTTPS` and select Application Profile `PCF-HTTPS`.
 - Use **Select IP Address** to select the **same IP address** to use as a VIP on the uplink interface.
 - Set **Protocol** to match the **Application Profile** protocol (`HTTPS`) and set **Port** to match the protocol (443).
 - Set **Default Pool** to the pool name set in [Step 2.6a: Create Pool for http-routers](#) (`http-routers`). This connects this VIP to that pool of resources being balanced to.
 - Ignore **Connection Limit** and **Connection Rate Limit** unless these limits are desired.
 - Switch to **Advanced Tab** on this Virtual Server.
 - Use the green plus to add/attach three Application Rules to this Virtual Server:
 - option `httplog`
 - `reqadd X-Forwarded-Proto:\ https`

 **Note:** Be careful to match protocol rules to the protocol `VIP- HTTP` to `HTTP` and `HTTPS` to `HTTPS`.

- Create a new Virtual Server named `TCPRtrs` and select Application Profile `PCF-TCP`.
 - Use **Select IP Address** to select the IP address to use as a VIP on the uplink interface.
 - Set **Protocol** to match the **Application Profile** protocol (TCP) and set **Port** to match the protocol (5000).
 - Set **Default Pool** to the pool name set in [Step 2.6b: Create Pool for tcp-routers](#) (`tcp-routers`). This connects this VIP to the pool of resources being balanced to.
 - Ignore **Connection Limit** and **Connection Rate Limit** unless these limits are desired.

The screenshot shows the Pivotal Load Balancer's configuration interface. The top navigation bar includes links for Settings, Firewall, DHCP, NAT, Routing, Load Balancer (which is selected and highlighted in blue), VPN, SSL VPN-Plus, and Grouping Objects. On the left, a sidebar lists Global Configuration, Application Profiles, Service Monitoring, Pools, Virtual Servers (selected and highlighted in blue), and Application Rules. The main content area displays a table of existing virtual servers and an 'Edit Virtual Server' dialog box.

Virtual Server ID	Name	Description	Default Pool	IP Address
virtualServer-1	GoRtr-HTTP		pool-1	10.193.251.16
virtualServer-2	GoRtr-HTTPS		pool-1	10.193.251.16
virtualServer-4	TCPRtrs		pool-2	10.193.251.16
virtualServer-5	SSH-DiegoBrain		pool-2	10.193.251.16

Edit Virtual Server

General

Enable Virtual Server
 Enable Acceleration

Application Profile: * PCF-TCP

Name: * TCPRtrs

Description:

IP Address: * 10.193.251.16

Protocol: TCP

Port: * 5000

Default Pool: tcp-routers

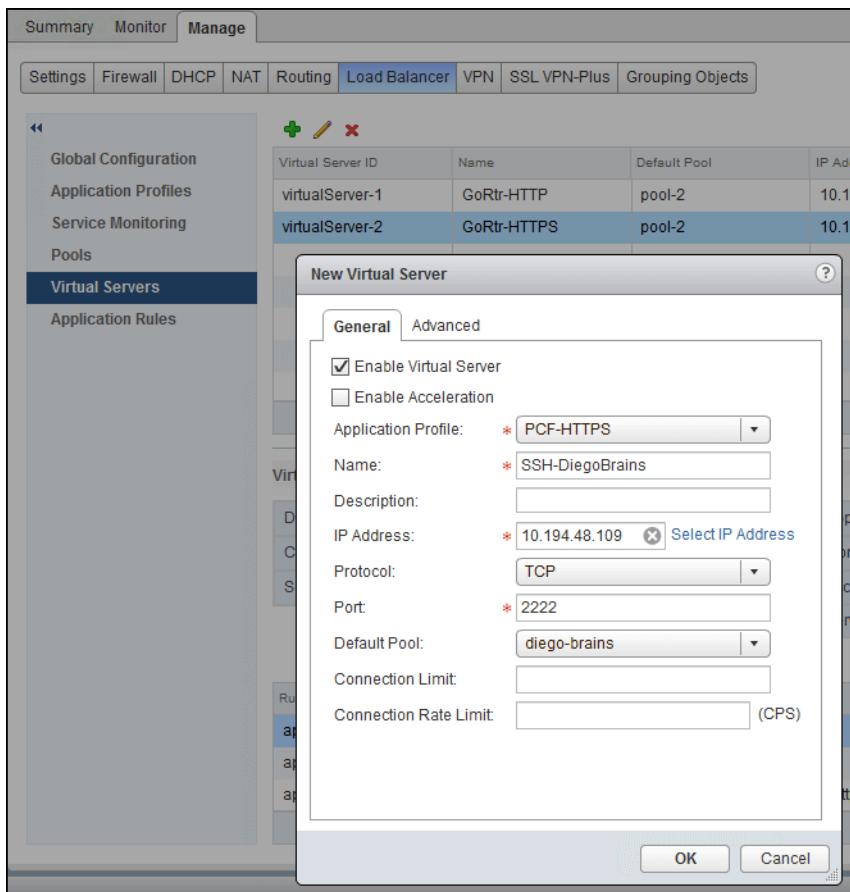
Connection Limit: 0

Connection Rate Limit: 0 (CPS)

OK Cancel

6. Create a new Virtual Server named `SSH-DiegoBrains` and select Application Profile `PCF-HTTPS`.

- o Use **Select IP Address** to select the same IP address to use as a VIP on the uplink interface if you want to use this address for SSH access to apps. If not, select a different IP address to use as the VIP.
- o Set **Protocol** to TCP and set **Port** to 2222.
- o Set **Default Pool** to the pool name set in [Step 2.6c: Create Pool for diego-brains](#) (`diego-brains`). This connects this VIP to that pool of resources being balanced to.
- o Ignore **Connection Limit** and **Connection Rate Limit** unless these limits are desired.



Step 3: Configure NAT/SNAT

The ESG obfuscates the PCF installation through network translation. The PCF installation is placed entirely on non-routable RFC-1918 network address space, so you must translate routable IP addresses to non-routable IP addresses to make connections.

Note: Correct NAT/SNAT configuration is required for the PCF installation to function as expected.

Action	Applied on Interface	Original IP	Original Port	Translated IP	Translated Port	Protocol	Description
SNAT	uplink	192.168.0.0/16	any	IP_of_PCF	any	any	All Nets Egress
DNAT	uplink	IP_of_OpsMgr	any	192.168.10.OpsMgr	any	tcp	OpsMgr Mask for external use
SNAT	infra	192.168.10.OpsMgr	any	IP_of_OpsMgr	any	tcp	OpsMgr Mask for internal use
DNAT	infra	IP_of_OpsMgr	any	192.168.10.OpsMgr	any	tcp	OpsMgr Mask for internal use

Note: The NAT/SNAT on the `infra` network in this table is an example of an optional **Hairpin NAT** rule to allow VMs within the **Infrastructure** network to access the Ops Manager API. This is because the Ops Manager hostname and the API HTTPS endpoint are registered to the Ops Manager external IP address. A pair of Hairpin NAT rules are necessary on **each** internal network interface that requires API access to Ops Manager. You should create these rules only if the network must access the Ops Manager API.

NAT/SNAT functionality is not required if routable IP address space is used on the Tenant Side of the ESG. At that point, the ESG simply performs routing between the address segments.

Note: NSX generates a number of DNAT rules based on load balancing configs. You can safely ignore these.

Additional Notes

The ESG also supports scenarios where Private RFC subnets and NAT are not utilized for **Deployment** or **Infrastructure** networks, and the guidance in

this document can be modified to meet those scenarios.

Additionally, the ESG supports up to 10 Interfaces allowing for more Uplink options if necessary.

The use of Private RFC-1918 subnets for PCF Deployment networks was chosen due to its popularity with customers. ESG devices are capable of leveraging ECMP, OSPF, BGP, and IS-IS to handle dynamic routing of customer and/or public L3 IP space. That design is out of scope for this document, but is supported by VMware NSX and Pivotal PCF.

Upgrading vSphere without PCF Downtime

This topic describes how to upgrade the vSphere components that host your Pivotal Cloud Foundry (PCF) installation without service disruption.

Minimum Requirements

At a bare minimum, vSphere contains the following components:

- vCenter Server
- one or more ESXi hosts

You cannot perform an in-place upgrade of vSphere without at least two ESXi hosts in your cluster.

If you do not meet this requirement (in other words, you have insufficient resources to evacuate an entire host), then you may experience PCF downtime during the upgrade.

To upgrade vSphere with only one ESXi host or without sufficient headroom capacity, you must reduce your PCF installation size. In other words, you can either reduce the number of Diego cells in your deployment or pause PCF VMs to make more capacity available. These actions can result in PCF downtime.

Recommended Starting Configuration

If you are running a PCF deployment as recommended by the base reference architecture for PCF on vSphere (recommended), then your vSphere installation should have the following components:

- One vCenter Server
- Three ESXi hosts per cluster
- Three or more clusters
- One (or HA pair) NSX Edge appliances

 **Note:** Pivotal recommends having at least three ESXi hosts in your cluster to maintain PCF high availability during your upgrade.

For more information, see the [Reference Architecture for Pivotal Cloud Foundry on vSphere](#).

Procedure to Upgrade vSphere

To upgrade the vSphere management layer underneath PCF, perform the following steps:

Step 1. Upgrade vCenter

For example, you might be upgrading vCenter 6.0 to vCenter 6.5.

For more information about how to upgrade vCenter, see [Overview of the vCenter Server Upgrade Process](#) in VMware documentation.

Step 2. Upgrade ESXi Hosts

After a successful vCenter upgrade, upgrade your ESXi hosts one at a time.

Starting with the first ESXi host, perform the following steps:

1. Verify that your ESXi hosts have sufficient resources and headroom to evacuate the VM workload of a single ESXi host to the two remaining hosts.

 **Note:** If you have enabled vSphere HA on your ESXi host, then each ESXi host should have sufficient headroom capacity since HA reserves 66% of available memory.

2. Use vMotion to move all the PCF VMs on the host you want to upgrade to the other ESXi hosts. vMotion places the VMs on the other hosts based on available capacity. For more information, see [Migration with vMotion](#) in VMware documentation.
3. Upgrade the evacuated ESXi host. For example, you may be upgrading from ESXi v6.0 to ESX v6.5. For instructions, see [Upgrading ESXi Hosts](#) in VMware documentation.

After successfully upgrading the first ESXi host, repeat the above steps for each remaining host one at a time. vSphere automatically rebalances all PCF VMs back onto the upgraded hosts via DRS after all the hosts are done.

Step 3. Upgrade ESG on VMware NSX

If your PCF deployment lives on a network behind an Edge Services Gateway (ESG) as recommended by the reference architecture, then upgrade each ESG only after completing the upgrade of vCenter and your ESXi hosts.

When you upgrade an ESG on VMware NSX, you upgrade the NSX Manager software. This upgrade can cause some slight downtime, the amount of which depends on the number of ESGs you are using.

- If your deployment only has one ESG, you can expect a downtime of 5 minutes for network reconvergence.
- If your ESGs are deployed in HA, upgrade the first ESG. Then upgrade the second ESG. This upgrade results only in 15-20 seconds of downtime.

For more information, see the [NSX Upgrade Guide](#) in VMware documentation.

Migrating PCF to a New Datastore in vSphere

This topic describes how to migrate your Pivotal Cloud Foundry (PCF) installation to a new vSphere datastore.

Prerequisites

Both the new and existing vSphere datastores must reside in the same datacenter.

To avoid service disruption, Pivotal recommends that you configure your overall PCF deployment for [high availability](#). In addition, check for configurations necessary to achieve high availability in each of your installed product tiles.

If your environment has any single points of failure, service may be disrupted as a result of the migration.

Before You Begin

This section describes the steps you should perform prior to the migration.

Step 1: Back Up Your Environment

Ensure that your PCF environment is fully backed up.

For more information about how to backup PCF, see [Backing up Pivotal Cloud Foundry](#).

Step 2: Document Current Environment Settings

Document your current environment settings before proceeding with the datastore migration. Record which VMs are running and in which datastore they reside. If you experience any issues during or after the migration, you must have this information to restore your environment.

To obtain this information, perform the following steps:

1. Run the `bosh instances` command.
 - o If you use [BOSH CLI v2+](#), run the following command, replacing `MY-ENV` with the alias you assigned to your BOSH Director:

```
$ bosh -e MY-ENV instances --details > instances.txt
```

- o If you use the original version of the BOSH CLI, run the following command:

```
$ bosh-old instances --details > instances.txt
```

2. Save the resulting file `instances.txt` to a safe location.
3. Note the datastore where each VM resides in vSphere.

Step 3: Check System Health

In Pivotal Application Service (PAS), check the **Status** tab and make sure there are no errors or reported issues.

Step 4: Check Installed Products Health

In each tile installed in your PCF deployment, check the **Status** tab and make sure there are no errors or reported issues.

Step 5: Check BOSH Director Status

1. Check that there are no pending changes and that the status of all tiles is green.
2. Make sure the last Installation Log does not contain any errors.
3. Before proceeding with the migration, click **Review Pending Changes** and **Apply Changes** to make sure there are no errors in the Installation Log.

Procedure: Migrate PCF to a New Datastore

1. In BOSH Director, navigate to the **vCenter Config** page.
2. Update the **Ephemeral Datastore Names** and **Persistent Datastore Names** field to reflect the new datastore names, then click **Save**.

Note: If you use the Datastore Clustering feature in vSphere, provide only the individual names of the datastores in the cluster. Do not provide the name of the cluster that contains them.

3. Click **Review Pending Changes**, then **Apply Changes**.
4. Confirm that the BOSH Director VM has persistent disk on the new datastore.
 - a. Navigate to **vCenter Resource Pools** and select the **Resource Pool** that contains your PCF deployment VMs and new datastore.
 - b. Click the **Related Objects** and **Virtual Machines**.
 - c. Locate the Ops Manager VM and verify that the VM has an expected value in the **Provisioned Space** column.

Name	State	Status	Provisioned Space	Used Space
linuxvm2	Powered On	Normal	22.09 GB	4.04 GB

5. In BOSH Director, navigate to the **Director Config** page, and select the **Recreate all VMs** option.
6. Click **Review Pending Changes**, then **Apply Changes**.

After the Migration

When BOSH moves disks, it waits for up to 60 minutes for the operation to complete. If the operation does not complete in time, BOSH can enter a state where it claims that the disks are **out of sync**.

Fix Failed BOSH Deployment with Out-of-Sync Error

If your PCF deployment gets into this state, you can resolve the issue by performing the steps in the KB article [How to recover from a failed bosh deployment when VMs are out of sync on vSphere](#).

Prevent Out-of-Sync Error

You can also prevent the **out of sync** BOSH error by increasing the CPI timeout to a larger value before performing the migration. Follow the instructions in the KB article [How to Increase the Timeout on Bosh CPI Command Calls](#).

Control Plane Reference Architectures

Introduction

Concourse is the main continuous integration and continuous delivery (CI/CD) tool that the Pivotal and open-source Cloud Foundry communities use to develop, test, deploy and manage Cloud Foundry foundations.

This topic describes topologies and best practices for deploying Concourse on [BOSH](#), and using Concourse to manage Pivotal Cloud Foundry (PCF) foundations. For production environments, Pivotal recommends deploying Concourse with BOSH.

The three topologies described in this document govern the network placement and relationship between two systems:

- The **control plane** runs Concourse to gather sources for, integrate, update, and otherwise manage PCF foundations. This layer may also host internal Docker registries, S3 buckets, git repositories, and other tools.
- Each **PCF foundation** runs PCF on an instance of BOSH.

Each topology described below has been developed and validated in multiple PCF customer and Pivotal Labs environments.

Deployment Topologies

Security policies are usually the main factor that determines which CI/CD deployment topology best suits a site's needs. Specifically, the decision depends on what network connections are allowed between the control plane and the PCF foundations that it manages.

The following three topologies answer a range of security needs, ordered by increasing level of security around the PCF foundations:

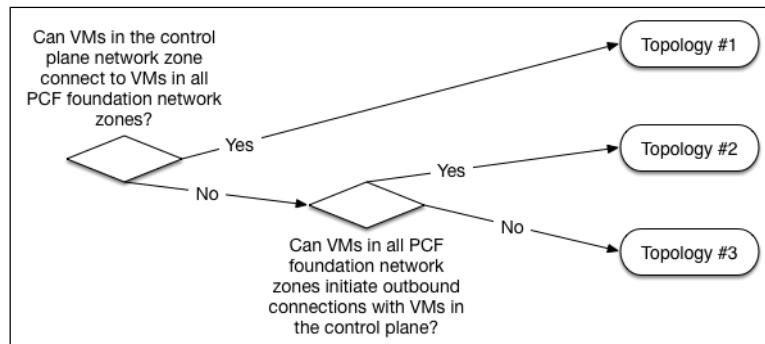
- [Topology 1](#): Concourse server and worker VMs all colocated on control plane
- [Topology 2](#): Concourse server on control plane, and remote workers colocated with PCF foundations
- [Topology 3](#): Multiple Concourse servers and workers colocated with PCF foundations

Across these three topologies, the increasing level of PCF foundation security correlates with:

- Increasing network complexity
- Increasing effort required for initial deployment and ongoing maintenance

Security Decision Factors

The graph below captures the decision factors dictated by network security policy, and recommends Concourse deployment topologies that adhere to those policies.



Additional Decision Factors

In addition to security policy, deciding on a CI/CD deployment topology may also depend on factors such as:

- Network latency across network zones

- Air-gapped vs. Internet-connected environments
- Resource limitations

The [Credentials Management](#) and [Storage Services](#) sections below includes notes and recommendations regarding credentials management, Docker registries, S3 buckets, and git repositories, but does not cover these tools extensively.

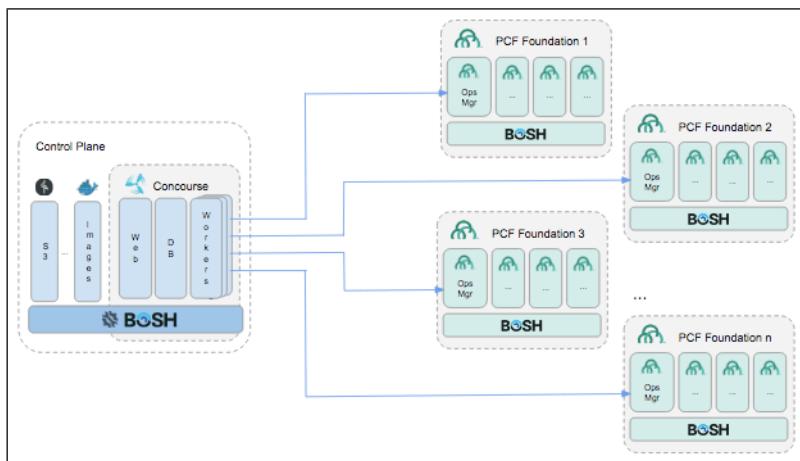
Topology Objects

Complete deployment topologies dictate the internal placement or remote use of the following:

- Concourse server and worker VMs, as discussed [above](#)
- Credentials managers such as CredHub or Vault
- Docker registries, public or private
- S3 or other storage buckets, public or private
- git or other code repositories, public or private

Topology 1: Concourse Server and Worker VMs All Colocated on Control Plane

This simple topology follows network security policies that allow a single control plane to connect to all PCF foundations deployed across multiple network zones or data centers.



In this topology, the Concourse server and worker VMs, along with other tools (e.g. Docker registry, S3 buckets, Vault server), are all deployed to the same subnet.

Connectivity Requirements

Concourse worker VMs must be allowed to connect to:

- The Ops Manager VM or a jumpbox VM in all of the PCF foundations networks
- (on vSphere) The vCenter API for each PCF foundation

Performance Notes

Network data transfers between the control plane and each PCF foundation network zone may carry large files such as PCF tiles, release files, and PCF foundation backup pipelines output.

Pivotal recommends that you test network throughput and latency between those network zones to make sure that data transfer rates do not make pipeline execution times unacceptably long.

Firewall Requirements

All Concourse worker VMs are required to connect to certain VMs on PCF foundation subnets across networks zones or data centers.

See [PCF CI/CD Pipelines](#) section for required VMs, ports, and external websites required for PCF CI/CD pipelines.

Pros

- Simplified deployment and maintenance of centralized control plane, which requires only one BOSH deployment and runs one BOSH Director.
- Simplified setup and maintenance of PCF CI/CD pipelines. All pipelines and Concourse teams use a single, centralized server.

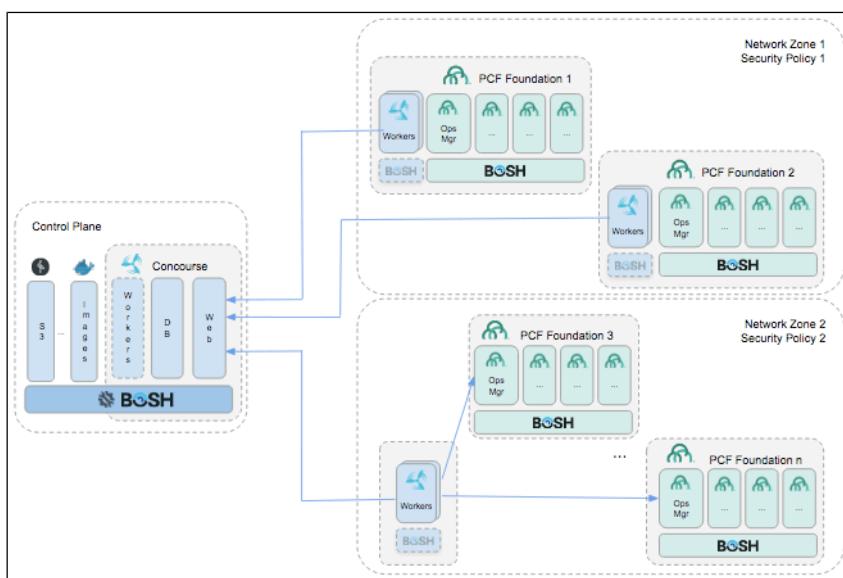
Cons

- You may have to configure firewall rules in each PCF network to allow connectivity from workers in CI/CD zone, as mentioned above.

Topology 2: Concourse Server on Control Plane, and Remote Workers Colocated with PCF Foundations

This topology supports environments where PCF foundation VMs cannot receive incoming traffic from IP addresses outside their network zone or data center, but they can initiate outbound connections to outside zones.

As in [Topology 1](#), the Concourse server, and potentially other VMs for tools such as Docker registries or S3 buckets, are deployed to a dedicated subnet. The difference here is that Concourse worker VM pools are deployed inside each PCF foundation subnet, network zone, or data center.



Connectivity Requirements

Concourse worker VMs in each PCF foundation network zone or data center must:

- Connect to the Ops Manager VM or a jumpbox VM in each PCF foundations network
- (on vSphere) Connect to the vCenter API for each PCF foundation
- Have outbound connection access to the Concourse web/ATC server on port 2222. This lets them handshake with the Concourse server to open a reverse SSH tunnel for ongoing communication between them and the Concourse ATC.

For more details, see [Concourse Architecture](#).

Performance Notes

Remote workers have to download large installation files from either the Internet or from a configured S3 artifacts repository. For PCF backup pipelines, workers may also have to upload large backup files to the S3 repository.

Pivotal recommends that you test network throughput and latency between those network zones to make sure that data transfer rates do not make pipeline execution times unacceptably long.

Firewall Requirements

Remote worker VMs require outbound access to the Concourse web/ATC server on port 2222.

Remote worker VMs inside each PCF foundation network zone or data center are required to connect to VMs in their colocated foundation. They may also require access to external web sites or S3 repositories for downloading installation files.

See the [PCF CI/CD Pipelines](#) section for required VMs and ports required for PCF CI/CD pipelines.

Pros

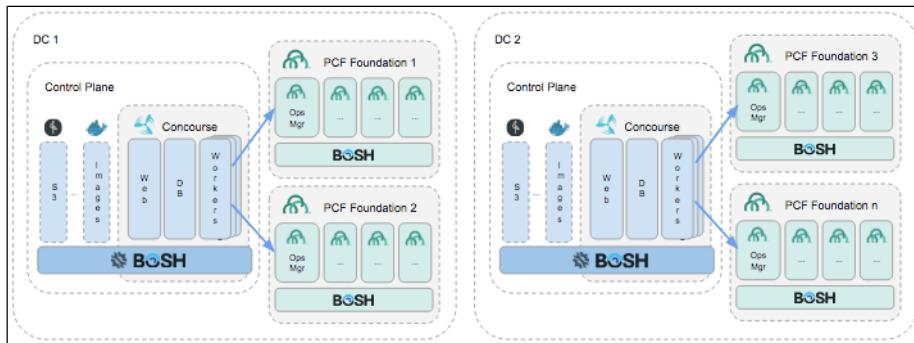
- Relatively simple maintenance of centralized control plane, which contains single Concourse server and other tools.
- Simplified setup and maintenance of PCF CI/CD pipelines. All pipelines and Concourse teams use a single, centralized server.

Cons

- You have to reconfigure firewalls to grant outbound access to remote worker VMs.
- In addition to deploying the control plane, you need an additional BOSH deployment and running BOSH Director for each PCF Foundation network zone or data center.
- You need to manage multiple Concourse worker pools in multiple locations.

Topology 3: Multiple Concourse Servers and Workers Colocated with PCF Foundations

This topology supports environments where PCF foundation VMs can only be accessed from within the same network zone or data center. This scenario requires deploying complete and dedicated control planes within each deployment zone.



Performance Notes

Since workers run in the same network zone or data center as the PCF foundation, data transfer throughput should not limit pipeline performance.

Firewall Requirements

- **Air-gapped environments** require some way to bootstrap S3 repository with Docker images and PCF releases files for pipelines from external sites.
- **Non-air-gapped environments**, in which workers can download required files from external websites, need those websites to be whitelisted in the proxy or firewall setup.

Pros

- Requires little or no firewall rules configuration for control plane VMs. In non-air-gapped environments, worker VMs download PCF releases and Docker images for pipelines from external websites.

Cons

- Requires deploying and maintaining multiple Concourse and other tools.
- Requires deploying and maintaining multiple PCF pipelines for each Concourse server.
- For air-gapped environments, requires setting up an S3 repository for each control plane.

Deploying CI/CD to the Control Plane

There are a few alternatives to deploy BOSH Directors, Concourse servers, and other tool releases to the control plane. Details on those alternatives are outside of the scope of this document, but refer to the links below for the most common options:

Manual deployments

- BOSH Director: [BOSH create-env ↗](#)
- Concourse
 - [Concourse ↗](#)
 - [Concourse with CredHub ↗](#)
 - [Concourse with Vault ↗](#)

- Docker registries
 - [Private Docker Registry](#)
 - [VMware Harbor Registry](#)
- S3 buckets
 - [Minio S3](#)
 - [EMC Cloud Storage](#)

Automated deployments

- [Bosh Bootloader \(BBL\)](#) deploys BOSH Director and Concourse on multiple IaaSes.

Control Plane Deployment Best Practices

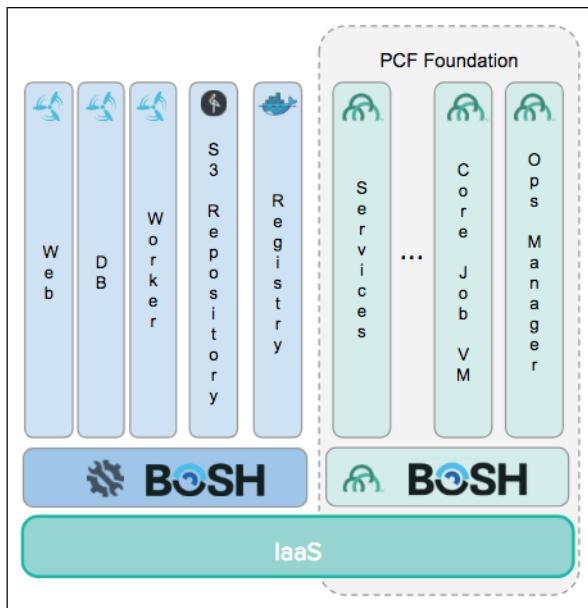
Here are some best practices for control plane components that apply across all deployment topologies:

Dedicated BOSH Director

Pivotal recommends deploying Concourse and other control plane tools on their own, dedicated BOSH layer, with their own BOSH Director that runs separately from the PCF foundations that the control plane manages.

Pivotal does **not** recommend using an existing PCF BOSH Director instance to deploy Concourse and other software (e.g. Minio S3, private Docker registry, CredHub). Sharing the same BOSH Director with PCF deployments increases the risk of accidental or undesired updates or deletion of those deployments.

Dedicating a BOSH Director to Concourse other control plane tools also provides higher flexibility for PCF foundation upgrades and updates, such as stemcell patches.



Credentials Management

All credentials and other sensitive information that feeds into a Concourse pipeline should be encrypted and stored using credentials management software such as CredHub or Vault. Never store credentials as plain text in parameter files in file repositories.

Credentials Management with CredHub

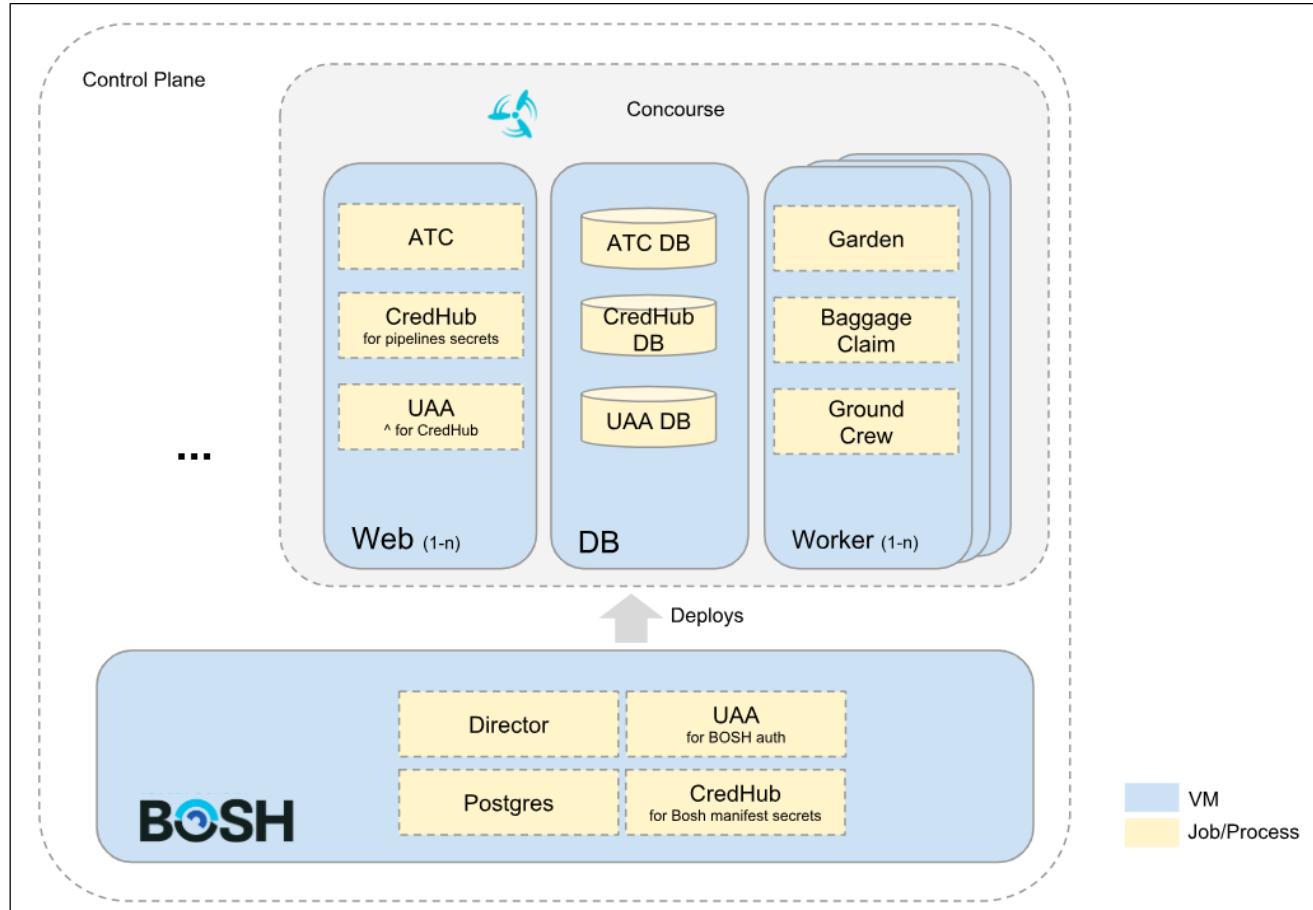
Concourse integrates with CredHub to manage credentials in its pipelines. The pipelines reference encrypted secrets stored in a CredHub server and retrieve them automatically during execution of tasks.

To integrate Concourse with a CredHub server, you configure its ATC job's deployment properties with information about the CredHub server and corresponding UAA authentication credentials.

You can deploy CredHub in multiple ways: as a dedicated VM, or integrated with other VMs, such as colocating the CredHub server with the BOSH Director VM or Concourse's ATC/web VM.

Colocating a CredHub server with Concourse's ATC VM dedicates it to the Concourse pipelines and lets Concourse administrators manage the credentials server. This configuration also means that during Concourse upgrades, the CredHub server only goes down when the Concourse ATC job is also down, which minimizes potential credential server outages for running pipelines.

The diagram below illustrates the jobs of Concourse VMs, along with the ones for the BOSH Director VM, when a dedicated CredHub server is deployed with Concourse.



The [Concourse Pipelines Integration with CredHub](#) documentation in the PCF Pipelines repository describes how to deploy a CredHub server integrated with Concourse.

Credentials Management with Vault

For how to configure Vault to manage credentials for Concourse pipelines, see [Secure credential automation with Vault and Concourse](#) in the PCF Pipelines repository.

Storage Services

Git Server

BOSH and Concourse implement the concepts of infrastructure-as-code and pipelines-as-code. As such, it is important to store all source code for deployments and pipelines in a version-controlled repository.

PCF CI/CD pipelines assume that their source code is kept in git-compatible repositories. [GitHub](#) is the most popular git-compatible repository for Internet-connected environments.

GitLab, BitBucket and [GOGS](#) are examples of git servers that can be used for both connected and air-gapped environments.

A git server that contains configuration and pipeline code for a PCF foundation needs to be accessible by the corresponding worker VMs that run CI/CD pipelines for that foundation.

S3 Repository

In all environments Concourse requires an S3 repository to store PCF backups and possibly other files. If your deployment requires a private, internal S3 repository but your IaaS lacks built-in options, you can use BOSH to deploy your own S3 releases, such as [Minio S3](#) and [EMC Cloud Storage](#) to your control plane.

For air-gapped environments, an S3 repository is also the preferred method to store release files for PCF tiles, stemcells and buildpacks. Docker images can also be stored to an S3 repository as an alternative to a private Docker registry. See [Offline Pipelines for Airgapped Environments](#) for details.

Private Docker Registry

For air-gapped environments, Docker Images for Concourse pipelines need to be stored either on a private Docker registry or in an S3 repository. For BOSH-deployed private registry alternatives, check [Docker Registry](#) or [VMWare Harbor](#).

High Availability

For details on how to set up a load balancer to handle traffic across multiple instances of the ATC/web VM, and how to deploy multiple worker instances, see the [Concourse Architecture](#) topic.

PCF CI/CD Pipelines

PCF Pipelines

PCF Platform Automation with Concourse (PCF Pipelines) is a collection of Concourse pipelines for installing and upgrading Pivotal Cloud Foundry. See the [source on Github](#) or download from [Pivotal Network](#) (sign-in required).

 **warning:** At time of publication, the PCF Pipelines repository is undergoing planned deprecation.

To run PCF Pipelines, you need:

- Ops Manager web UI, API and VM installed
- (vSphere environments) vCenter API installed
- For proxy- or Internet-connected environments, whitelist the following sites:
 - [Docker Hub](#): hub.docker.io
 - [Pivotal Network](#): network.pivotal.io
 - bosh.io

BOSH Backup and Restore (BBR) PCF Pipelines

BBR PCF Pipelines automate PCF foundation backups. See and download the [source on Github](#).

To run BBR PCF Pipelines, you need an S3 repository for storing backup artifacts.

 **Note:** BBR PCF Pipelines is a PCF community project not officially supported by Pivotal.

Pipelines Orchestration Frameworks

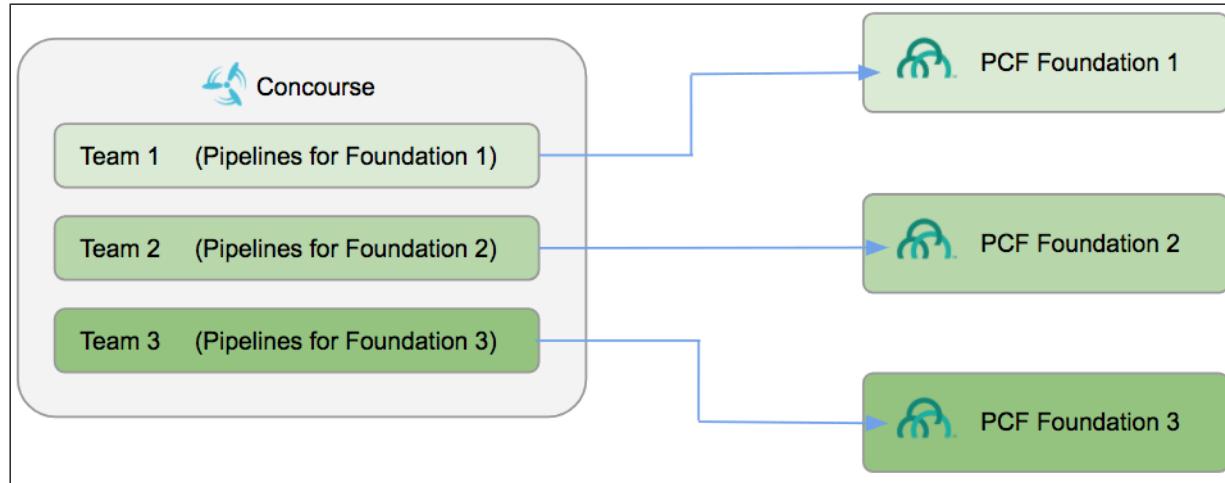
[PCF Pipelines Maestro](#) uses a Maestro framework to:

- Automate pipeline creation and management for multiple PCF foundations
- Promote and audit configuration changes and version upgrades across all foundations

 **Note:** PCF Pipelines Maestro is a PCF community project not officially supported by Pivotal.

Concourse Team Management Best Practices

When a single Concourse server hosts CI/CD pipelines for more than one PCF foundation, Pivotal recommends creating one Concourse team (not `main`) specific to each foundation, and associating that team with all pipelines for that foundation, e.g. install, upgrade, backup, and metering.



Dedicating a Concourse team to each PCF foundation has the following benefits:

- **It avoids the clutter of pipelines in a single team.** The list of pipelines for each foundation may be long, depending how many tiles are deployed to it.
- **It avoids the risk of operators running a pipeline for the wrong foundation.** When a single team hosts maintenance pipelines for multiple foundations, the clutter of dozens of pipelines may lead operators to accidentally run a pipeline (e.g. upgrade or delete tile) targeted at the wrong PCF foundation.
- **It allows for more granular access control settings per team.** PCF pipelines for higher environments (e.g. Production) may require a more restricted access control than ones from lower environments (e.g. Sandbox). Authentication settings for Concourse teams enable that level of control.
- **It allows workers to be assigned to pipelines of a specific foundation.** Concourse deployment configuration allows for the assignment of workers to a single team. If that team contains pipelines of only one foundation, then the corresponding group of workers run pipelines only for that foundation. This is useful when security policy requires tooling and automation for a foundation (e.g. Production) to run on specific VMs.

Concourse Teams for App Development Orgs and Spaces

When a Concourse server hosts pipelines for an app development team, Pivotal recommends creating a Concourse team associated with that development team, and associating Concourse team membership with org and space membership defined in the Pivotal Application Service (PAS) user authentication and authorization (UAU) server.

Associating Concourse teams with PAS org and space member lists synchronizes access between PAS and Concourse, letting developers see and operate the build pipelines for the apps they develop.

Implementing a Multi-Foundation PKS Deployment

Page last updated:

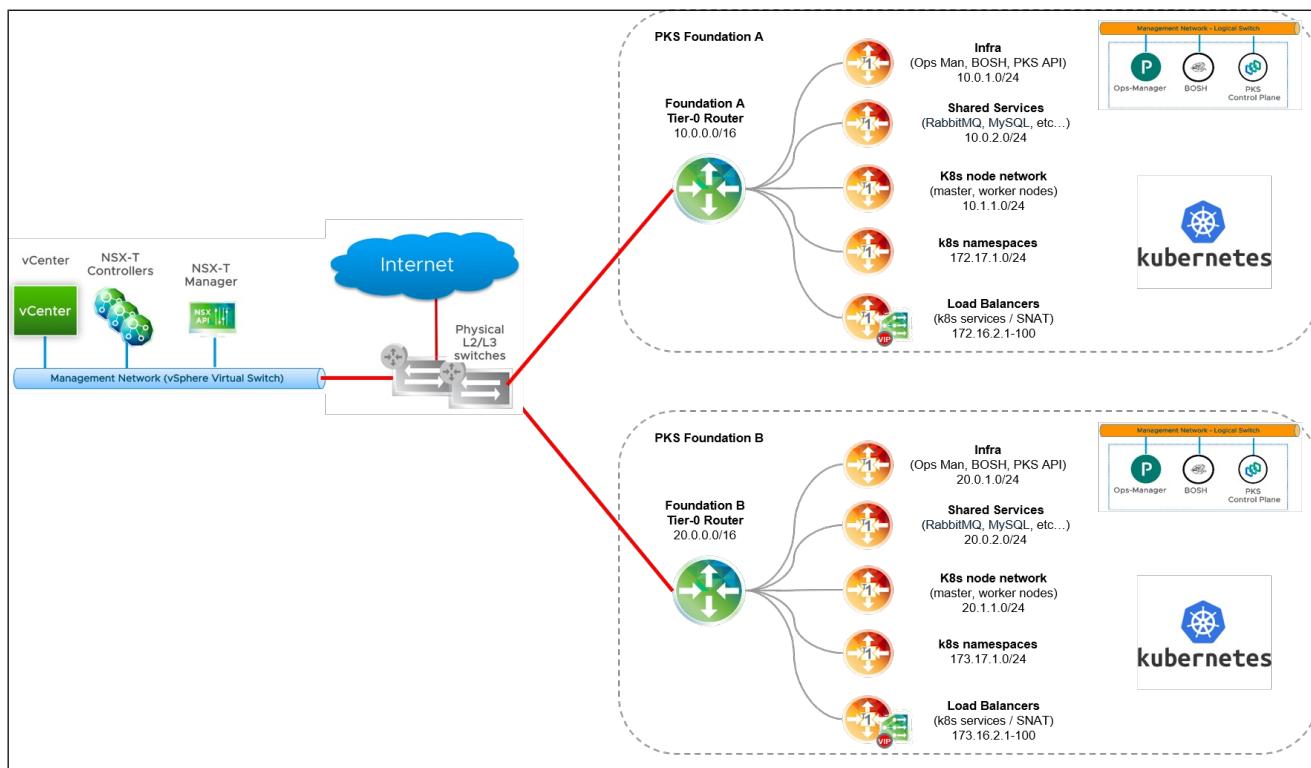
This topic describes how to deploy multiple instances of PKS on vSphere with NSX-T infrastructure.

About Multi-Foundation PKS

A multi-foundation deployment of PKS lets you install and run multiple instances of PKS. The purpose of a multi-foundation deployment of PKS is to share a common vSphere and NSX-T infrastructure across multiple foundations, while providing complete networking isolation across foundations.

As shown in the diagram, with a multi-foundation PKS topology, each PKS instance is deployed to a dedicated NSX-T Tier-0 router. Foundation A T0 router with Management CIDR 10.0.0.0/16 connects to the vSphere and NSX-T infrastructure. Similarly, Foundation B T0 router with Management CIDR 20.0.0.0/16 connects to the same vSphere and NSX-T components.

As with a single instance deployment, PKS management components are deployed to a dedicated network, for example, 10.0.0.0/24 for PKS Foundation A; 20.0.0.0/24 for PKS Foundation B. When PKS is deployed, networks are defined for nodes, pods, and load balancers. Because of the dedicated Tier-0 router, there is complete networking isolation between each PKS instance.



Requirements

To implement a multi-foundation PKS topology, adhere to the following requirements:

- One Tier-0 router for each PKS instance. For more information, see [Configuring Multiple Tier-0 Routers for Tenant Isolation](#).
- The Floating IP pool must not overlap. The CIDR range for each Floating IP Pool must be unique and not overlapping across foundations. For more information, see [Create Floating IP Pool](#).
- PKS instances can be deployed in NAT and no-NAT mode. If more than one PKS instance is deployed in no-NAT mode, the Nodes IP Block networks cannot overlap.
- For any Pods IP Block used to deploy Kubernetes clusters in no-NAT (routable) mode, the Pods IP Block cannot overlap across foundations.

The image below shows three PKS installations across three Tier-0 foundations. Key considerations to keep in mind with a multi-foundation PKS topology include the following:

- Each foundation must rely on a dedicated Tier-0 router

- You can mix-and-match NAT and no-NAT mode across foundations for Node and Pod networks
- If you are using non-routable Pods IP Block networks, the Pods IP Block addresses can overlap across foundations
- Because Kubernetes nodes are behind a dedicated Tier-0 router, if clusters are deployed in NAT mode the Nodes IP Block addresses can also overlap across foundations
- For each foundation you must define a unique Floating IP Pool with non-overlapping IPs

PKS Foundation A	PKS Foundation B	PKS Foundation C
<input checked="" type="checkbox"/> NAT mode	<input type="checkbox"/> NAT mode	<input type="checkbox"/> NAT mode
Pods IP Block ID *	Pods IP Block ID *	Pods IP Block ID *
927d2aff-fa86-4df8-bb21-c45b5314f547	927d2aff-fa86-4df8-bb21-c45b5314f547	1b81b967-e269-4a62-9ff5-e2e39a5feaae
Nodes IP Block ID *	Nodes IP Block ID *	Nodes IP Block ID *
3d577e5c-dcaf-4921-9458-a12b0e1318e6	3d577e5c-dcaf-4921-9458-a12b0e1318e6	3d577e5c-dcaf-4921-9458-a12b0e1318e6
T0 Router ID *	T0 Router ID *	T0 Router ID *
40445803-8c3c-417e-bb24-e84cf9a330b5	5c579a37-5318-4255-9658-1a2a99a1d1e9	791f220b-155b-4fe9-bf3b-50199e4a911d
Floating IP Pool ID *	Floating IP Pool ID *	Floating IP Pool ID *
86213c33-9b7a-4a91-b470-7145941bccb3	31e0fd4e-19e7-4122-b300-438d465a406f	803c7aa5-16c9-4bdf-a404-e9bbdf0eb7ce
Nodes DNS *	Nodes DNS *	Nodes DNS *
10.40.53.1	10.40.53.1	10.40.53.1
vSphere Cluster Names *	vSphere Cluster Names *	vSphere Cluster Names *
Cluster-A	Cluster-B	Cluster-C

Please send any feedback you have to pkcs-feedback@pivotal.io.

Global DNS Load Balancers for Multi-Foundation Environments

This topic describes global DNS load balancers (GLBs) for multi-foundation environments. This topic also describes concepts such as foundation affinity and healthchecks.

 **Note:** If you want to configure a load balancer dedicated to one PCF foundation and you are using an F5 LTM, see [Configuring an F5 Load Balancer for PAS](#).

About Multi-Foundation Environments

Multi-foundation environments are multiple instances of BOSH and Ops Manager that can communicate with each other. Each foundation can use different infrastructures to fit your preferences.

Multi-foundation environments are commonly deployed in an active-active or active-passive pattern. See below for descriptions of each term:

Term	Description
Active-active pattern	Instances of apps run on two foundations and they are both in use.
Active-passive pattern	Instances of apps run on two foundations, but the instances may only be active on one foundation. The other foundation becomes active only in the event of a failover.

Configure Your GLB

The typical setup for foundation failover requires the GLB be authoritative for a wildcard app domain. The wildcard app domain is not the same domain as the foundation-default app domain.

To configure your GLB, do the following:

- Find and record your wildcard app domain. For Pivotal Application Service (PAS), your wildcard app domain is typically the **Apps Domain** you configure in the **Domains** pane of the PAS tile.
- Add the domain you recorded to both PCF foundations using a shared domain. You can create a shared domain using the Cloud Foundry Command Line Interface (cf CLI). For more information about the cf CLI command to create shared domains, see [create-shared-domain](#) in the cf CLI reference guide.
- To support failover, set the time-to-live (TTL) in the wildcard DNS record. Set the TTL to about 30 to 180 seconds. When determining your TTL, consider the tradeoff between the app performance impact and the resulting time for failover to occur.

About Foundation Affinity

Foundation affinity occurs when the GLB favors one foundation over another during a route request. For example, users experience less latency if they are routed to a foundation that is geographically closer, so the GLB may favor that foundation.

Different GLBs have their own mechanisms to achieve this. See the following table for common concepts:

Foundation Affinity Concepts	
Term	Description
Topology/geographically-based affinity	The GLB attempts to direct traffic to the geographically nearest foundation based on IP geolocation or provided topology for private networks of the LDNS server performing the lookup.
Static-persist/member-hashing affinity	The GLB attempts to direct traffic to the geographically nearest foundation based on IP geolocation or provided topology for private networks of the LDNS server performing the lookup.
Active/passive foundations	If one foundation is usually idle, you can always pick the active foundation IP as long as it remains available.
Microservice-to-microservice affinity	For microservice apps, you typically use a Services Registry to manage traffic between microservices. There are two ways to do this: <ul style="list-style-type: none"> If the microservices are using the domain which maps to the GLB, then their traffic is routed through the GLB. If the microservices are communicating using IP address or internal domains, such as when using Envoy, then the

traffic does not pass through the GLB.

About Healthchecks

Healthchecks determine whether a foundation for an app is healthy or not.

See the following table for the levels at which you can check the health of your foundation:

Level	Description
Foundation	GLB relies on a local load balancer in front of the PCF Gorouters to determine the overall health of the foundation. GLB can perform healthchecks on TCP port 443 or 80 or on the local load balancer. The local load balancer healthchecks the backend pool of Gorouters on port 8080.
App	Healthchecks are set only for apps which have instances on both foundations. Each app instance has canary DNS records. The canary DNS records are the same for the app instances on each foundation. You would also need to add more VIPs dedicated to these canary apps that would be used to check their health.

⚠ warning Pivotal does not recommend setting healthchecks at the app level. Configuring healthchecks in this way can cause more frequent failover and delays while pushing apps. Complete failover of a foundation affects all apps on the platform. Healthchecks on a per-app basis can require additional overhead beyond the control of an app developer.

Architecture and Installation Overview

For PCF Architects and Operators

This guide shows you how to design a Pivotal Cloud Foundry (PCF) platform and install it on an IaaS. If you are doing this, you have one or both of the following roles:

- **Architects** design a PCF platform. They know the IaaS that they will deploy it to, and what other relevant resources they have. In their design, they consider needs for the platform's capacity, availability, security, geography, budget, and other factors. If they do not install PCF themselves, they provide the architectural specifications to whoever does.
- **Operators** run a PCF platform, keep it up-to-date, monitor its health and performance, and fix any problems. They may also install the platform, or perform "Day 2" configurations that expand its functionality and integrate it with external systems.

This guide helps people in both roles create a PCF platform that does what they want it to. The contents of this guide follow the phases of a typical PCF planning and installation effort.

 **Note:** Elastic Runtime has been renamed Pivotal Application Service.

Planning and Installation Overview

PCF is a suite of products that runs on multiple IaaSes. Planning and installing PCF means building layers from the bottom up, starting with the details of your IaaS and ending with "Day 2" configurations that you perform on a installed and running PCF deployment.

Here's the typical PCF planning and installation process:

1. Plan Your PCF Architecture

- Review the Requirements for your IaaS ([AWS](#), [Azure](#), [GCP](#), [OpenStack](#), [vSphere](#)).
- Under [Reference Architectures](#), refer to the reference architecture for your IaaS.
- Assess your platform needs, including capacity, availability, container support, host OS, resource isolation, and geographical distribution. Discuss with your Pivotal contact.

2. Deploy BOSH and Ops Manager

- [BOSH](#) is an open-source tool that lets you run software systems in the cloud.
 - BOSH and its IaaS-specific Cloud Provider Interfaces (CPIs) are what enable PCF to run on multiple IaaSes.
 - See [Deploying with BOSH](#) for a brief description of the BOSH deployment process.
- [Ops Manager](#) is a GUI application deployed by BOSH that streamlines deploying subsequent software to the cloud via BOSH.
 - Ops Manager represents PCF products as tiles with multiple configuration panes that let you input or select configuration values needed for the product.
 - Ops Manager generates BOSH manifests containing the user-supplied configuration values, and sends them to the Director.
 - After you install Ops Manager and BOSH, you use Ops Manager to deploy almost all PCF products.
- Deploying Ops Manager deploys both BOSH and Ops Manager with a single procedure.
 - On AWS, you can deploy Ops Manager manually, or automatically with a CloudFormation template.
 - On Azure, you can deploy Ops Manager manually, or automatically with an Azure Resource Manager (ARM) template. On Azure Government Cloud and Azure Germany you can only deploy Ops Manager manually.

3. Deploy BOSH Add-ons (Optional)

- BOSH add-ons include the [IPsec](#), [ClamAV](#), and [File Integrity Monitoring](#), which enhance PCF platform security and security logging.
- You deploy these add-ons via BOSH rather than installing them with Ops Manager tiles.

4. Install Runtimes

- [PAS](#) (Pivotal Application Service) lets developers develop and manage cloud-native apps and software services.
 - PAS is based on the Cloud Foundry Foundation's open-source Application Runtime (formerly Elastic Runtime) project.
- [PKS](#) (Pivotal Container Service) uses BOSH to run and manage Kubernetes container clusters.

- PKS is based on the Cloud Foundry Foundation's open-source Container Runtime (formerly Kubo) project.
- [PCF Isolation Segment](#) lets a single PAS deployment run apps from separate, isolated pools of computing, routing, and logging resources.
 - Operators replicate and configure an Isolation Segment tile for each new resource pool they want to create.
 - You must install PAS before you can install Isolation Segment.
- [PAS for Windows 2012R2](#) enables PAS to manage Windows Server 2012 cells hosting .NET apps, and can also be replicated to create multiple isolated resource pools.
 - As with Isolation Segment, Operators replicate and configure a PAS for Windows 2012R2 tile for each new resource pool they want to create.
 - You must install PAS before you can install PAS for Windows 2012R2.
- [Small Footprint PAS](#) is an alternative to PAS that uses far fewer VMs than PAS but has limitations.

5. Install Services

- Install software services for PCF developers to use in their apps.
 - Services include the databases, caches, and message brokers that stateless cloud apps rely on to save information.
 - Installing and managing software services on PCF is an ongoing process, and is covered in the [PCF Operator Guide](#).

6. Day 2 Configurations

- Day 2 configurations set up internal operations and external integrations on a running PCF platform.
 - Examples include front-end configuration, user accounts, logging and monitoring, internal security, and container and stemcell images.

Guide Contents

This guide has two parts. The first part explains the PCF planning and installation process, and the second describes the main tools that operators use when installing PCF.

Planning and Installation

PCF is a suite of products that runs on multiple IaaSes. Planning and installing PCF means building layers from the bottom up, starting with the details of your IaaS and ending with “Day 2” configurations that you perform on a installed and running PCF deployment.

This guide follows this bottom-up progression:

- [Planning PCF](#) - Design a PCF platform that runs on your IaaS and fits your needs.
- [Deploying BOSH and Ops Manager](#) - Build the foundation of PCF.
- [Deploying BOSH Add-Ons](#) - Use BOSH to extend foundation-level PCF capabilities.
- [Using Ops Manager](#) - Ops Manager provides a dashboard for installing and configuring PCF products and services.
- [Using the Cloud Foundry Command Line Interface \(cf CLI\)](#) - The cf CLI runs PCF operations in the cloud from a shell on your local workstation.
- [Installing Runtimes](#) - Install the application and container runtime environments that PCF exists for.
- [Day 2 Configurations](#) - Set up internal operations and external integrations for PCF.
- [Troubleshooting and Diagnostics](#)

After installing PCF, Operators install the software services that PCF developers use in their apps. These PCF services include the databases, caches, and message brokers that stateless cloud apps rely on to save information.

Installing and managing software services on PCF is an ongoing process, and is covered in the [PCF Operator Guide](#).

Related Documentation

The [PCF Operator Guide](#) explains how to maintain a running PCF platform, including monitoring, tuning, troubleshooting, and upgrading.

The [PCF Security Guide](#) explains how PCF security systems work and how to keep your PCF platform secure.

[Getting Started with PCF](#) gives a high-level overview of how PCF works and explains how you can try a simple deployment on your own local

machine.

See [Pivotal Cloud Foundry Documentation](#) for all PCF documentation.

Preparing Your Firewall

This topic describes how to configure your firewall for [Pivotal Cloud Foundry](#) (PCF) and how to verify that PCF resolves DNS entries behind your firewall.

Configure Your Firewall for PCF

Ops Manager and Pivotal Application Service (PAS) require the following open TCP ports:

- **25555**: Routes from Ops Manager to the BOSH Director.
- **443**: Routes to HAProxy or, if configured, your own load balancer.
- **80**: Routes to HAProxy or, if configured, your own load balancer.
- **8844**: Routes from Ops Manager to BOSH CredHub.
- **8443**: Routes from Ops Manager to BOSH Director UAA.
- **6868**: Routes to the BOSH Agent.
- **2222**: Necessary for using Application SSH. For details, see the [Diego Architecture Diagram](#).
- **25595**: Routes from the Traffic Controller to the BOSH Director, to enable sending BOSH health metrics to the Firehose.

UDP port **123** must be open if you want to use an external NTP server.

For more information about required ports for additional installed products, see [Network Communication Paths](#).

Example: Configure Firewall with iptables

The following example procedure uses `iptables` commands to configure a firewall.

Note: `GATEWAY_EXTERNAL_IP` is a placeholder. Replace this value with your `PUBLIC_IP`.

1. Open `/etc/sysctl.conf`, a file that contains configurations for Linux kernel settings, with the command below:

```
$ sudo vi /etc/sysctl.conf
```

2. Add the line `net.ipv4.ip_forward=1` to `/etc/sysctl.conf` and save the file.

3. If you want to remove all existing filtering or Network Address Translation (NAT) rules, run the following commands:

```
$ iptables --flush  
$ iptables --flush -t nat
```

4. Add environment variables to use when creating the IP rules:

```
$ export INTERNAL_NETWORK_RANGE=10.0.0.0/8  
$ export GATEWAY_INTERNAL_IP=10.0.0.1  
$ export GATEWAY_EXTERNAL_IP=203.0.113.242  
$ export PIVOTALCF_IP=10.0.0.2  
$ export HA_PROXY_IP=10.0.0.254
```

5. Run the following commands to configure IP rules for the specified chains:

- FORWARD:

```
$ iptables -A FORWARD -i eth1 -j ACCEPT  
$ iptables -A FORWARD -o eth1 -j ACCEPT
```

- POSTROUTING:

```
$ iptables -t nat -A POSTROUTING -o eth0 -j MASQUERADE
$ iptables -t nat -A POSTROUTING -d SHA_PROXY_IP -s $INTERNAL_NETWORK_RANGE \
    -p tcp --dport 80 -j SNAT --to $GATEWAY_INTERNAL_IP
$ iptables -t nat -A POSTROUTING -d SHA_PROXY_IP -s $INTERNAL_NETWORK_RANGE \
    -p tcp --dport 443 -j SNAT --to $GATEWAY_INTERNAL_IP
```

- PREROUTING:

```
$ iptables -t nat -A PREROUTING -d $GATEWAY_EXTERNAL_IP -p tcp --dport \
    25555 -j DNAT --to $PIVOTALCF_IP
$ iptables -t nat -A PREROUTING -d $GATEWAY_EXTERNAL_IP -p tcp --dport \
    443 -j DNAT --to $HA_PROXY_IP
$ iptables -t nat -A PREROUTING -d $GATEWAY_EXTERNAL_IP -p tcp --dport \
    80 -j DNAT --to $HA_PROXY_IP
$ iptables -t nat -A PREROUTING -i eth0 -p tcp --dport 8443 -j DNAT \
    --to $PIVOTALCF_IP:443
$ iptables -t nat -A PREROUTING -i eth0 -p tcp --dport 80 -j DNAT \
    --to $HA_PROXY_IP:80
$ iptables -t nat -A PREROUTING -i eth0 -p tcp --dport 8022 -j DNAT \
    --to $PIVOTALCF_IP:22
$ iptables -t nat -A PREROUTING -i eth0 -p tcp --dport 8080 -j DNAT \
    --to $PIVOTALCF_IP:80
```

6. Run the following command to save the iptables:

```
$ service iptables save
```

For more information about administering IP tables with `iptables`, refer to the [iptables documentation](#).

Verify PCF Resolves DNS Entries Behind a Firewall

When you install PCF in an environment that uses a strong firewall, the firewall might block DNS resolution. For example, if you use [xip.io](#) to test your DNS configuration, the tests will fail without warning if the firewall prevents PAS from accessing `*.xip.io`.

To verify that PAS can correctly resolve DNS entries:

1. SSH into the Pivotal Ops Manager VM. For more information, refer to the [SSH into Ops Manager](#) section of the Advanced Troubleshooting with the BOSH CLI topic.
2. Run any of the following network administration commands with the IP address of the VM:
 - `nslookup`
 - `dig`
 - `host`
 - The appropriate `traceroute` command for your OS
3. Review the output of the command and fix any blocked routes. If the output displays an error message, review the firewall logs to determine which blocked route or routes you need to clear.
4. Repeat steps 1-3 with the BOSH Director VM and the HAProxy VM.

IaaS Permissions Guidelines

This topic describes practices recommended by Pivotal for creating secure IaaS user roles.

Pivotal Cloud Foundry (PCF) is an automated platform that connects to IaaS providers such as AWS and OpenStack. This connectivity typically requires accounts with appropriate permissions to act on behalf of the operator to access IaaS functionality such as creating virtual machines (VMs), managing networks and storage, and other related services.

Ops Manager and Pivotal Application Service (PAS) can be configured with IaaS users in different ways depending on your IaaS. Other product tiles and services might also use their own IaaS credentials. Refer to the documentation for those product tiles or services to configure them securely.

Least Privileged Users (LPUs)

Pivotal recommends following the principle of least privilege by scoping privileges to the most restrictive permissions possible for a given role. In the event that someone gains access to credentials by mistake or through malicious intent, LPUs limit the scope of the breach. Pivotal recommends following best practices for the particular IaaS you are deploying.

AWS Guidelines

See the recommendations detailed in the [AWS Permissions Guidelines](#) topic.

Azure Guidelines

See the permissions recommendations in [Preparing to Deploy Ops Manager on Azure Manually](#), and use the minimum permissions necessary when creating your service principal.

GCP Guidelines

For GCP, Pivotal recommends using two different accounts with the least privilege.

Use one account with the minimum permissions required to create desired GCP resources in your GCP project, then create a separate service account with the minimum permissions required to deploy PCF components such as Pivotal Ops Manager and PAS. For more information about creating the service account, see *Step 1: Set up IAM Service Accounts* in [Preparing to Deploy Ops Manager on GCP Manually](#).

OpenStack Guidelines

Pivotal recommends following the principle of least privilege by scoping privileges to the most restrictive permissions possible for a given role.

vSphere Guidelines

See the vCenter permissions recommendations in the [Installing Pivotal Cloud Foundry on vSphere](#) topic.

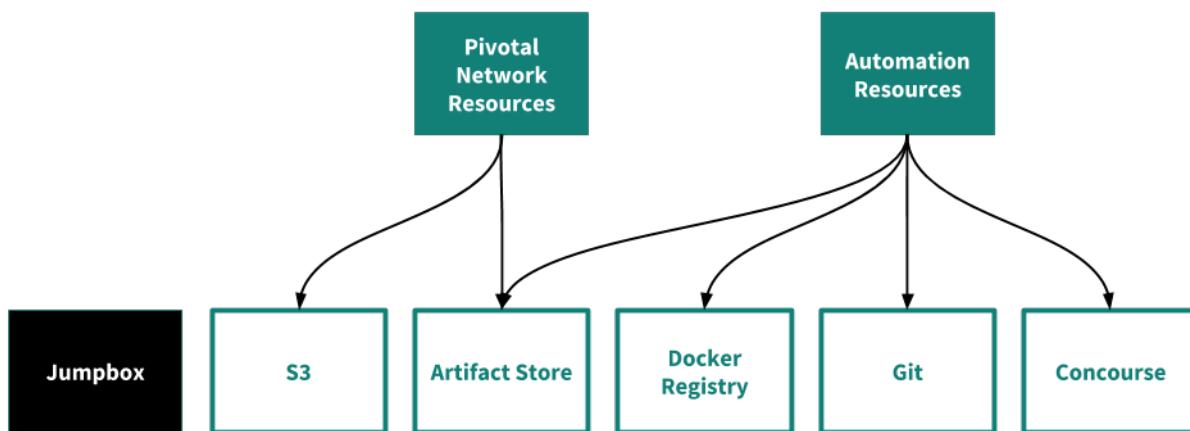
Installing PCF in Airgapped Environments

This topic provides an overview of the components and resources needed to install PCF in airgapped environments, including the typical corresponding automation resources.

Offline Components

To run PCF offline, you must obtain resources from the internet and move them into offline components that store and use them. The method you use to move a resource from Pivotal Network or GitHub into an offline environment can vary from setting up a designated proxy to burning a DVD.

The following image displays types of resources and the components you must move them to in your offline environment. It also includes a jumpbox. The jumpbox is a Linux host for running commands such as `bosh`, `uaac`, and `fly`. You could use the Ops Manager VM for this purpose.



The following table provides more detail about resources and the component you must move them to:

Resource	Component
Pivotal Network products such as tiles, stemcells, BOSH releases, and Ops Manager	S3 and artifact store
Pipelines and scripts	Git
Configuration	Git
Container images	Docker Registry, S3, or artifact store
Third Party Resources such as NSX-T and OSS BOSH releases	Artifact store
Backup artifacts	S3

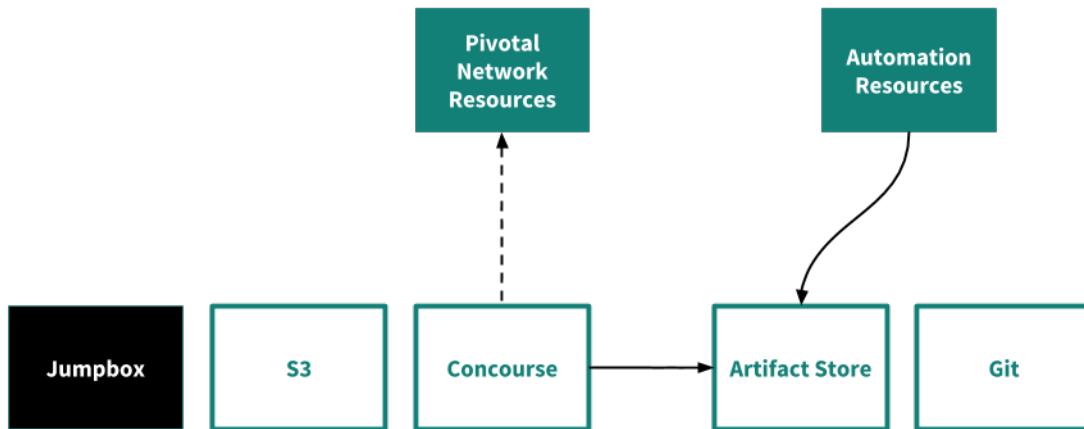
Architectural Patterns

The following sections describe three architectural patterns for running PCF in airgapped environments. The pattern you use depends on how your environment is set up.

Pattern One: Artifact Store with Internet Access

It is common for organizations to deploy an artifact store such as Artifactory or Nexus as a gateway to the Internet. In this configuration, the artifact store is typically configured as a Docker mirror.

This pattern includes an S3 component that is used only for backups of the PCF installation.

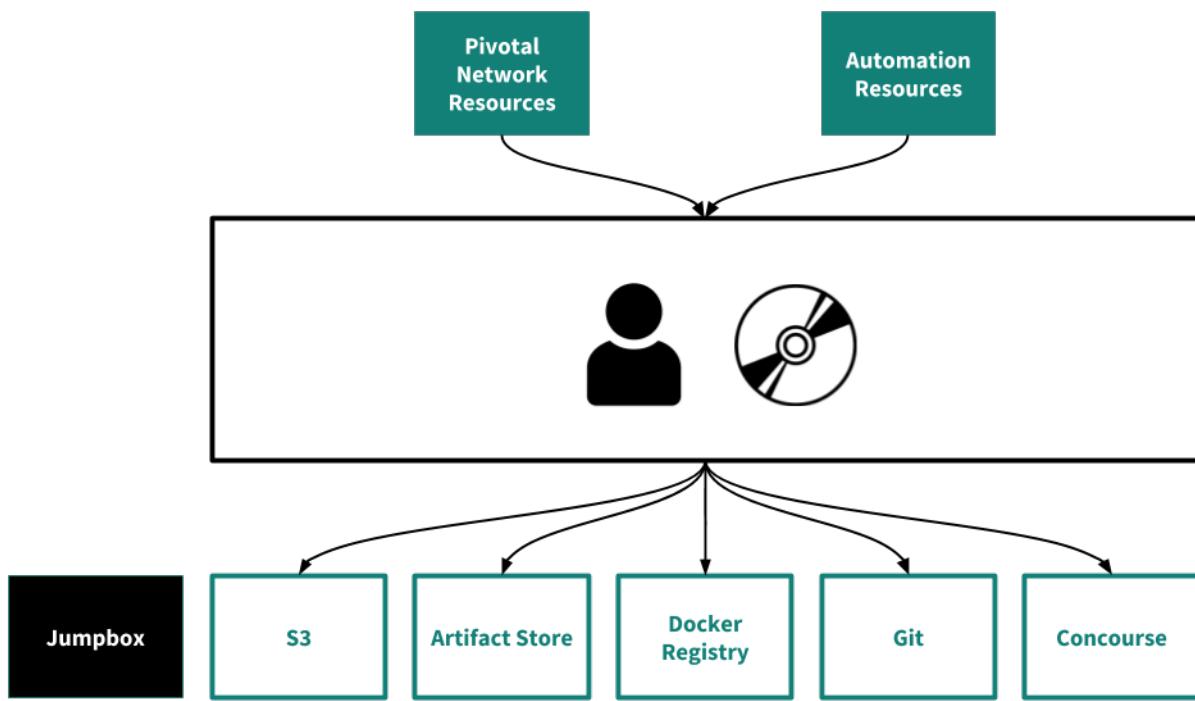


The following table provides an example of how you might handle resources in this scenario:

Resource	Component
Pivotal Network products such as tiles, stemcells, BOSH releases, and Ops Manager	This architectural pattern presents the challenge of getting Pivotal Network resources from the Internet into the artifact store. For example, Artifactory and Nexus cannot interact directly with Pivotal Network. You could place Pivotal Network resources in an artifact store either through a remote Concourse worker with access to the Internet, by downloading them from the Internet and placing them in the artifact store manually, or some other method.
Pipelines and scripts	Store in Git. Use manual clone and push. Modify pipelines to use <code>registry_mirror</code> .
Configuration	Store in Git. Use manual clone and push.
Container images	Store in the artifact store that you configured as a mirror.
Third Party Resources such as NSX-T and OSS BOSH releases	Store in a generic artifact store repository.
Backup artifacts	Store in a S3 blobstore.

Pattern Two: Completely Offline

In completely offline environments, you must bring in all resources manually and store them in the available local components.

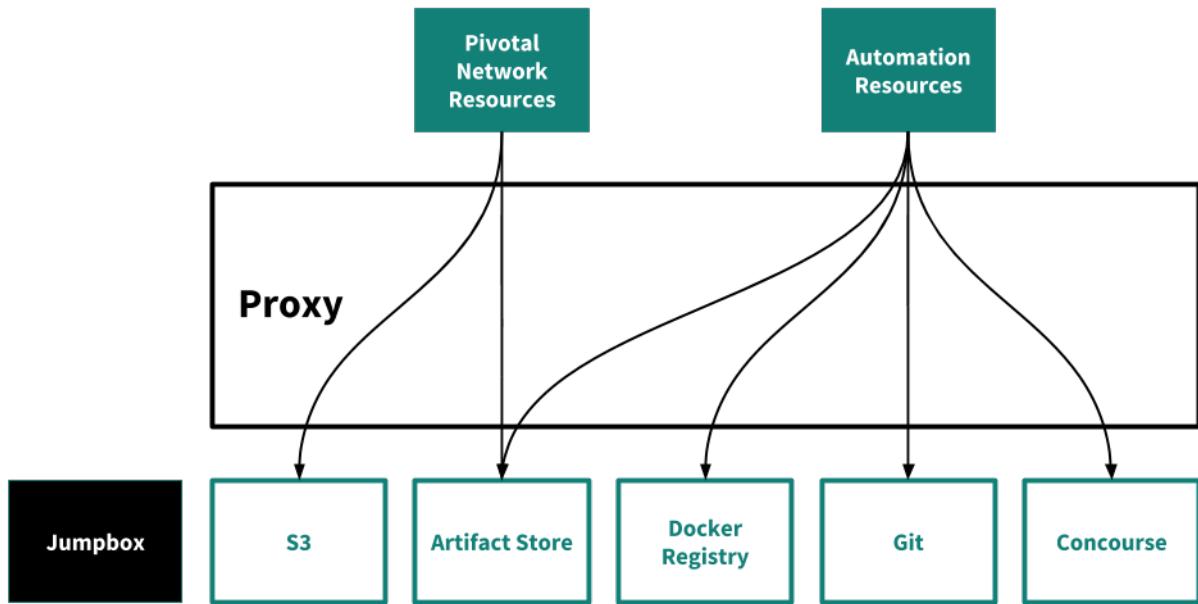


In this architectural pattern, you must configure pipelines to watch components for changes and apply updates when available. The following table provides an example of how you might handle resources in this scenario:

Resource	Component
Pivotal Network products such as tiles, stemcells, BOSH releases, and Ops Manager	Manually upload to Nexus.
Pipelines and scripts	Modify pipelines to use a local Harbor Docker registry. Manually clone an online environment, bring it to the offline environment on DVD, and push it to offline Git.
Configuration	Store in your offline Git.
Container Images	Get from the Internet, transfer to USB, and push to offline Harbor.
Third Party Resources such as NSX-T and OSS BOSH releases	Store in a generic artifact store repository.
Backup artifacts	Store in a S3 blobstore.

Pattern Three: TLS Intercepting Proxy

This pattern allows Internet access, but only through a proxy that decrypts and rewrites TLS certificates. In general, this resembles an online install, however it requires that you add your corporate certificate to all BOSH-deployed VMs.



Some of the challenges in this pattern include the following:

- Any automation that calls to the Internet will fail if it does not have the corporate certificate in its trust store.
- Concourse tasks that do not use resources do not receive updated BOSH root certificates. This means you must configure tasks to ignore TLS errors or update the root certificates as part of the tasks.

Installing Pivotal Cloud Foundry on AWS

This guide describes how to install [Pivotal Cloud Foundry](#) (PCF) on Amazon Web Services (AWS).

Overview

You can install PCF on AWS with either the Pivotal Application Service (PAS) or Pivotal Container Service (PKS) runtime. There are resource requirements specific to each runtime. Ensure you meet the requirements for your runtime and the requirements specific to AWS before installing PCF on AWS.

Requirements

This section lists the following resource requirements for installing PCF on AWS:

- General PCF resource requirements. See [PCF Resource Requirements](#).
- AWS-specific resource requirements. See [AWS Resource Requirements](#).

PCF Resource Requirements

This section lists PCF resource requirements for installing PCF on AWS. It includes general PCF resource requirements for both the PAS and PKS runtimes.

View one of the following, depending on your PCF runtime:

- PAS-specific PCF resource requirements. See [PAS Resource Requirements](#).
- PKS-specific PCF resource requirements. See [PKS Resource Requirements](#).

PAS Resource Requirements

The following are general resource requirements for deploying and managing a PCF deployment with Ops Manager and PAS:

- PAS requires sufficient IP allocation. The following lists the minimum required IP allocations:
 - One static IP address for either HAProxy or one of your Gorouters
 - One static IP address for each job in the Ops Manager tile. See the Ops Manager **Resource Config** pane for each tile for a full list.
 - One static IP address for each job listed below:
 - Consul
 - NATS
 - File Storage
 - MySQL Proxy
 - MySQL Server
 - Backup Restore Node
 - HAProxy
 - Router
 - MySQL Monitor
 - Diego Brain
 - TCP Router
 - One IP for each VM instance created by the service.
 - An additional IP address for each compilation worker. Use the following formula to determine the total IPs required:
$$\text{IPs needed} = \text{static IPs} + \text{VM instances} + \text{compilation workers}$$
 - Pivotal recommends that you allocate at least 36 dynamic IP addresses when deploying Ops Manager and PAS. BOSH requires additional dynamic IP addresses during installation to compile and deploy VMs, install PAS, and connect to services.
- Pivotal recommends using a network without DHCP for deploying PAS VMs.

 **Note:** If you have DHCP, refer to the [Troubleshooting Guide](#) to avoid issues with your installation.

PKS Resource Requirements

For PKS-specific resource requirements, see [AWS Prerequisites and Resource Requirements](#).

AWS Resource Requirements

The following are AWS-specific resource requirements for installing PCF on AWS with an external database and external file storage:

- Installing PCF on AWS requires a minimum of the following VM instance limits in your AWS account. The number of VMs required depends on the number of tiles and availability zones you plan to deploy. The following VM guidelines apply to the PAS, Small Footprint PAS, and PKS runtimes:
 - **PAS:** At a minimum, a new AWS deployment requires the following VMs for PAS:

AWS Requirements	VM Name	VM Type	Default VM Count	Required or Optional VM
PAS	NATS	t3.micro	2	Required
	File Storage	m5.large	1	Optional
	MySQL Proxy	t3.micro	2	Optional
	MySQL Server	r5.large	3	Optional
	Backup Restore Node	t3.micro	1	Optional
	Diego BBS	t3.micro	3	Required
	UAA	m5.large	2	Required
	Cloud Controller	m5.large	2	Required
	HAProxy	t3.micro	0	Optional
	Router	t3.micro	3	Required
	MySQL Monitor	t3.micro	1	Optional
	Clock Global	t3.medium	2	Required
	Cloud Controller Worker	t3.micro	2	Required
	Diego Brain	t3.small	3	Required
	Diego Cell	r5.xlarge	3	Required
	Loggregator Traffic Controller	t3.micro	2	Required
	Syslog Adapter	t3.micro	3	Required
	Syslog Scheduler	t3.micro	2	Required
	Doppler Server	m5.large	3	Required
	TCP Router	t3.micro	0	Optional
	CredHub	r5.large	2	Optional
	Istio Router	r5.large	0	Optional
	Istio Control	r5.large	0	Optional
	Route Syncer	r5.large	0	Optional
Ops Manager	BOSH Director	m5.large	1	Required

 **Note:** If you are deploying a test or sandbox PCF that does not require high availability, then you can scale down the number of VM instances in your deployment. For more information, see [Scaling PAS](#).

- **Small Footprint PAS:** To run Small Footprint PAS, a new AWS deployment requires:

AWS Requirements	VM Name	VM Type	Default VM Count	Minimum HA VM Count	Required or Optional VM
Small Footprint PAS	Compute	r5.xlarge	1	3	Required
	Control	r5.xlarge	1	2	Required
	Database	r5.large	1	3	Required
	Router	t3.micro	1	3	Required
	File Storage	m5.large	1	N/A	Optional
	Backup Restore Node	t3.micro	1	1	Optional

	MySQL Monitor	t3.micro	1	1	Optional
	HAProxy	t3.micro	0	2	Optional
	TCP Router	t3.micro	0	1	Optional
	Istio Router	r5.large	0	1	Optional
	Istio Control	r5.large	0	2	Optional
	Route Syncer	r5.large	0	1	Optional
Ops Manager	BOSH Director	m5.large	1	N/A	Required

- **PKS:** See [AWS Prerequisites and Resource Requirements](#).
- The following AWS resources are required for installing PCF on AWS with PAS:
 - 3 Elastic Load Balancers (ELBs)
 - 1 Relational Database Service. As a minimum, Pivotal recommends using a db.m5.xlarge instance with at least 100 GB of allocated storage.
 - 5 S3 Buckets

Prerequisites

To install PCF on AWS, you must do the following:

- Increase or remove the VM instance limits in your AWS account. Installing PCF requires more than the default 20 concurrent instances. For more information about VM resource requirements, see [Requirements](#).
- Configure your AWS account with the appropriate AWS region. For more information about selecting the correct region for your deployment, see [Region and Availability Zone Concepts](#) in the AWS documentation.
- Install the AWS CLI. Configure the AWS CLI with the user credentials that have admin access to your AWS account. To download the AWS CLI, see [AWS CLI](#).
- Configure an AWS EC2 key pair to use with your PCF deployment. For more information, see [Creating an EC2 Key Pair](#) in the AWS documentation.
- Register a wildcard domain for your PCF installation. For more information, see [SSL/TLS Certificates for Classic Load Balancers](#) in the AWS documentation.
- Create an SSL certificate for your PCF domain. For more information, see the [AWS documentation about SSL certificates](#).

 **Note:** To deploy PCF to a production environment, you must obtain a certificate from a certificate authority. Pivotal recommends using a self-signed certificate generated by Ops Manager for development and testing purposes only.

- **(PAS-only)** Configure sufficient IP allocation. For more information about IP allocation requirements, see [PAS Resource Requirements](#).
- **(Optional) (PAS-only)** Configure external storage. Pivotal recommends using external storage if possible. For more information about how file storage location affects platform performance and stability during upgrades, see [Configure File Storage](#).
- **(Optional) (PAS and Ops Manager-only)** Configure external databases. Pivotal recommends using external databases in production deployments for BOSH Director and PAS. An external database must be configured to use the UTC timezone.
- **(Optional) (PAS and Ops Manager-only)** Configure external user stores. When you deploy PCF, you can select a SAML user store for Ops Manager or a SAML or LDAP user store for PAS, to integrate existing user accounts.

Install PCF on AWS

You can install PCF on AWS either manually or using Terraform.

To install PCF on AWS, do one of the following:

- Install PCF on AWS manually. See [Installing PCF on AWS Manually](#).
- Install PCF on AWS using Terraform. See [Install PCF on AWS Using Terraform](#).

Additional Resources

The following are additional resources related to installing PCF on AWS:

- For information about AWS identity and access management, see [What is IAM?](#) in the AWS documentation.

- For information about users, groups, and roles in AWS, see [Identities \(Users, Groups, and Roles\)](#) in the AWS documentation.
- For best practices for managing IaaS users and permissions, see [Temporary Security Credentials](#) in the AWS documentation.
- For recommendations on how to create and scope AWS accounts for PCF, see [AWS Permissions Guidelines](#).
- For more information about certificate requirements for installing PCF, see [Certificate Requirements](#).

AWS Permissions Guidelines

Pivotal recommends that you minimize the use of master account credentials by creating an IAM role and instance profile with the minimum required EC2, VPC, and EBS credentials.

In addition, Pivotal recommends that you follow AWS account security best practices such as disabling root keys, using multi-factor authentication on the root account, and using CloudTrail for auditing API actions.

For more Amazon-specific best practices, refer to the following Amazon documentation:

- [IAM Roles Best Practices ↗](#)
- [AWS Security Best Practices Whitepaper ↗](#)
- [AWS Well-Architected Framework ↗](#)

Installing PCF on AWS Manually

This topic describes how to manually configure the Amazon Web Services (AWS) components that you need to run [Pivotal Cloud Foundry](#) (PCF) on AWS.

To deploy PCF on AWS, you must perform the procedures in this topic to create objects in the AWS Management Console that PCF requires.

To view the list of AWS objects created by the procedures in this topic, see the [Required AWS Objects](#) section.

After completing the procedures in this topic, proceed to [Configuring BOSH Director on AWS](#) to continue deploying PCF.

Step 1: File a Ticket

Log in to the AWS Management Console, and file a ticket with Amazon to ensure that your account can launch more than the default 20 instances. In the ticket, ask for a limit of 50 `t3.micro` instances and 20 `c5.large` instances in the region you are using.

Note: To deploy PCF to AWS GovCloud (US), log in to the [AWS GovCloud \(US\) Console](#) instead of the standard AWS Management Console and select the `us-gov-west-1` region.

Note: To deploy PCF to AWS China, set up an [AWS China](#) account and contact the Platform Architect assigned for your Pivotal account.

You can check the limits on your account by visiting the EC2 Dashboard on the AWS Management Console and clicking **Limits** on the left navigation.

Step 2: Create S3 Buckets

1. Navigate to the S3 Dashboard.

Note: S3 bucket names must be globally unique. When naming buckets, Pivotal recommends that you prefix the generic names below with an unique and helpfully identifiable string (i.e. ID-STRING-pcf-ops-manager-bucket, MY-IDENTIFIER-pcf-buildpacks-bucket, and so on). Then you should use the same prefix when naming other associated resources, such as IAM policies.

2. Perform the following steps to create five S3 buckets:

- o Click **Create Bucket**
- o For **Bucket name**, enter `ID-STRING-pcf-ops-manager-bucket`.
- o For **Region**, select your region.
- o Click **Next** three times.
- o Click **Create bucket**.
- o Repeat the above steps to create four more S3 buckets: `ID-STRING-pcf-buildpacks-bucket`, `ID-STRING-pcf-packages-bucket`, `ID-STRING-pcf-resources-bucket`, and `ID-STRING-pcf-droplets-bucket`.

Step 3: Create an IAM User for PCF

Perform the following steps to create an Amazon Identity and Access Management (IAM) user with the minimal permissions necessary to run and install PCF:

1. Click **IAM** to access the IAM Dashboard.
2. Click **Users** and then click **Add user**.

Add user
1
2
3
4

Set user details

You can add multiple users at once with the same access type and permissions. [Learn more](#)

User name*	<input type="text" value="pcf-user"/>
+ Add another user	

Select AWS access type

Select how these users will access AWS. Access keys and autogenerated passwords are provided in the last step. [Learn more](#)

Access type*	<input checked="" type="checkbox"/> Programmatic access Enables an access key ID and secret access key for the AWS API, CLI, SDK, and other development tools.
	<input type="checkbox"/> AWS Management Console access Enables a password that allows users to sign-in to the AWS Management Console.

* Required

[Cancel](#)
[Next: Permissions](#)

3. Enter a user name, such as `pcf-user`.

4. For AWS access type, select **Programmatic access**.

💡 Note: If you prefer to create your keys locally and import them into AWS, see the [Amazon documentation](#).

5. Click **Next: Permissions**.

6. Click **Next: Review** and review your choices.

💡 Note: On the **Review** page you may see a warning that the user has no permissions. You can disregard this message. You do not need to set user permissions.

7. Click **Create user**.

8. Click **Download .csv** to download the user security credentials.

⚠️ warning: The `credentials.csv` contains the IDs for your user security access key and secret access key. Keep the `credentials.csv` file for your currently active key pairs in a secure directory. You cannot recover a lost key pair.

9. Click **Close**.

10. On the **Users** page, click the user name to access the user details page.

💡 Note: On the **Users** page you may see a warning that the user has no permissions. You can disregard this message. You do not need to set user permissions.

11. Click **Add inline policy**. You can review your existing inline policies by clicking the down arrow.

The screenshot shows the 'Create policy' page in the AWS IAM console. At the top, there are two tabs: 'Visual editor' (selected) and 'JSON'. A button 'Import managed policy' is located at the top right. Below the tabs, a code editor displays a JSON policy document:

```

1  {
2   "Version": "2012-10-17",
3   "Statement": [
4     {
5       "Effect": "Deny",
6       "Action": [
7         "iam:*"
8       ],
9       "Resource": [
10      "*"
11     ]
12   },
13   {
14     "Sid": "OpsMgrInfrastructureIaaSConfiguration",
15     "Effect": "Allow"
16   }
17 ]
18 }

```

A validation error message is displayed in a yellow box: "This policy validation failed and might have errors converting to JSON : The policy must have at least one statement For more information about the IAM policy grammar, see [AWS IAM Policies](#)".

At the bottom right of the page are 'Cancel' and 'Review policy' buttons.

12. On the **Create policy** page, define a policy:

- Copy the policy document included in the [Pivotal Cloud Foundry for AWS Policy Document](#) topic. You must edit the policy document so the names of the S3 buckets match the ones you created in [Step 2: Create S3 Buckets](#).
- Paste the policy document into the **JSON** tab on the **Create policy** page.

13. Click **Review policy**.

14. In the **Name** field, enter `pcf-iam-policy`.

15. Click **Create policy**. The **Summary** page displays a list of available policies and actions.

Step 4: Create a VPC

1. Navigate to the VPC Dashboard.

2. Click **Start VPC Wizard**.

The screenshot shows the 'VPC Dashboard' page. At the top, there is a 'Resources' section with a note: 'Note: Your Instances will launch in the US East (N. Virginia) region.' Below this, there is a summary of resources in the 'US East (N. Virginia) region':

	2 VPCs	2 Internet Gateways
Subnets	4 Subnets	4 Route Tables
Route Tables	2 Network ACLs	3 Elastic IPs
Internet Gateways	6 Security Groups	18 Running Instances
DHCP Options Sets	0 VPC Peering Connections	0 Customer Gateways
Elastic IPs	0 VPN Connections	0 Virtual Private Gateways

On the left, there is a sidebar with links: 'Virtual Private Cloud', 'Your VPCs', 'Subnets', 'Route Tables', 'Internet Gateways', 'DHCP Options Sets', and 'Elastic IPs'. At the top center, there are buttons: 'Start VPC Wizard' (highlighted in blue) and 'Launch EC2 Instances'.

3. Select **VPC with Public and Private Subnets** and click **Select**.

Step 1: Select a VPC Configuration

VPC with a Single Public Subnet	In addition to containing a public subnet, this configuration adds a private subnet whose instances are not addressable from the Internet. Instances in the private subnet can establish outbound connections to the Internet via the public subnet using Network Address Translation (NAT).
VPC with Public and Private Subnets	A /16 network with two /24 subnets. Public subnet instances use Elastic IPs to access the Internet. Private subnet instances access the Internet via a Network Address Translation (NAT) instance in the public subnet. (Hourly charges for NAT instances apply.)
VPC with Public and Private Subnets and Hardware VPN Access	
VPC with a Private Subnet Only and Hardware VPN Access	

Select

4. Specify the following details for your VPC:

- IPv4 CIDR block: Enter `10.0.0.0/16`.
- IPv6 CIDR block: Select **No IPv6 CIDR Block**.
- VPC name: `pcf-vpc`.
- Public subnet's IPv4 CIDR: Enter `10.0.0.0/24`.
- Set the Availability Zone fields for both subnets to `REGION-#a`. For example, `us-west-2a`.
- Public subnet name: Enter `pcf-public-subnet-az0`.
- Private subnet's IPv4 CIDR: Enter `10.0.16.0/28`.
- Private subnet name: Enter `pcf-management-subnet-az0`.
- Click **Use a NAT instance instead** and do the following:
 - Under **Specify the details of your NAT instance**, set the **Instance type** to `t3.medium`
 - Create a key pair titled `pcf-ops-manager-key`. For more information about creating the key pair, see [Amazon EC2 Key Pairs](#) in the AWS documentation.
 - Select your newly-created `pcf-ops-manager-key` for the **Key Pair name**.
- Enable DNS hostnames: Click **Yes**.
- Hardware tenancy: Select **Default**.
- Click **Create VPC**.

5. After the VPC is successfully created, click **Subnets** in the left navigation.

6. Click **Create Subnet**.

7. Add the following subnets to the `pcf-vpc` VPC:



Note: You created the first two subnets in the previous step: `pcf-public-subnet-az0` and `pcf-management-subnet-az0`.

Name	AZ	IPv4 CIDR block
<code>pcf-public-subnet-az1</code>	<code>REGION-#b</code> (for example, <code>us-west-2b</code>)	<code>10.0.1.0/24</code>
<code>pcf-public-subnet-az2</code>	<code>REGION-#c</code> (for example, <code>us-west-2c</code>)	<code>10.0.2.0/24</code>
<code>pcf-management-subnet-az1</code>	<code>REGION-#b</code> (for example, <code>us-west-2b</code>)	<code>10.0.16.16/28</code>
<code>pcf-management-subnet-az2</code>	<code>REGION-#c</code> (for example, <code>us-west-2c</code>)	<code>10.0.16.32/28</code>
<code>pcf-ert-subnet-az0</code>	<code>REGION-#a</code> (for example, <code>us-west-2a</code>)	<code>10.0.4.0/24</code>
<code>pcf-ert-subnet-az1</code>	<code>REGION-#b</code> (for example, <code>us-west-2b</code>)	<code>10.0.5.0/24</code>
<code>pcf-ert-subnet-az2</code>	<code>REGION-#c</code> (for example, <code>us-west-2c</code>)	<code>10.0.6.0/24</code>
<code>pcf-services-subnet-az0</code>	<code>REGION-#a</code> (for example, <code>us-west-2a</code>)	<code>10.0.8.0/24</code>
<code>pcf-services-subnet-az1</code>	<code>REGION-#b</code> (for example, <code>us-west-2b</code>)	<code>10.0.9.0/24</code>
<code>pcf-services-subnet-az2</code>	<code>REGION-#c</code> (for example, <code>us-west-2c</code>)	<code>10.0.10.0/24</code>
<code>pcf-rds-subnet-az0</code>	<code>REGION-#a</code> (for example, <code>us-west-2a</code>)	<code>10.0.12.0/24</code>
<code>pcf-rds-subnet-az1</code>	<code>REGION-#b</code> (for example, <code>us-west-2b</code>)	<code>10.0.13.0/24</code>
<code>pcf-rds-subnet-az2</code>	<code>REGION-#c</code> (for example, <code>us-west-2c</code>)	<code>10.0.14.0/24</code>

Step 5: Configure a Security Group for Ops Manager

1. Return to the EC2 Dashboard.
2. Select **Security Groups>Create Security Group**.
3. For **Security group name**, enter `pcf-ops-manager-security-group`.
4. For **Description**, enter a description to identify this security group.
5. For **VPC**, select the VPC where you want to deploy Ops Manager.
6. Click the **Inbound** tab and add rules according to the table below.

Note: Pivotal recommends limiting access to Ops Manager to IP ranges within your organization, but you may relax the IP restrictions after configuring authentication for Ops Manager.

Type	Protocol	Port Range	Source
HTTP	TCP	80	My IP
HTTPS	TCP	443	My IP
SSH	TCP	22	My IP
BOSH Agent	TCP	6868	10.0.0.0/16
BOSH Director	TCP	25555	10.0.0.0/16

7. Click **Create**.

Step 6: Configure a Security Group for PCF VMs

1. From the **Security Groups** page, click **Create Security Group** to create another security group.
2. For **Security group name**, enter `pcf-vms-security-group`.
3. For **Description**, enter a description to identify this security group.
4. For **VPC**, select the VPC where you want to deploy the PCF VMs.
5. Click the **Inbound** tab and add rules for all traffic from your public and private subnets to your private subnet, as the table and image show. This rule configuration does the following:
 - Enables BOSH to deploy PAS and other services.
 - Enables application VMs to communicate through the router.
 - Allows the load balancer to send traffic to Pivotal Application Service (PAS).

Type	Protocol	Port Range	Source
All traffic	All	0 - 65535	Custom IP 10.0.0.0/16

6. Click **Create**.

The screenshot shows the AWS EC2 Security Group creation interface. The 'Inbound' tab is selected. The configuration is as follows:

- Security group name:** pcf-vms
- Description:** pcf-vms
- VPC:** vpc-1ed91c/b (10.0.0.0/16) |pcf-vpc|
- Security group rules:**
 - Inbound:** Type: All traffic, Protocol: All, Port Range: 0 - 65535, Source: Custom IP 10.0.0.0/16

Step 7: Configure a Security Group for the Web ELB

1. From the **Security Groups** page, click **Create Security Group** to create another security group.

2. For **Security group name**, enter `pcf-web-elb-security-group`.
3. For **Description**, enter a description to identify this security group.
4. For **VPC**, select the VPC where you want to deploy this Elastic Load Balancer (ELB).
5. Click the **Inbound** tab and add rules to allow traffic to ports `80`, `443`, and `4443` from `0.0.0.0/0`, as the table and image show.

Note: Allow traffic to port `4443` only if you are in an AWS cloud region that does not support AWS ALBs. For example, the GovCloud region. For more information about AWS regions and availability zones, see [AWS Global Infrastructure](#).

Note: For finer control over what can reach PAS, change `0.0.0.0/0` to be more restrictive. This security group governs external access to PAS from apps such as the cf CLI and app URLs.

Type	Protocol	Port Range	Source
Custom TCP rule	TCP	4443	Anywhere
HTTP	TCP	80	Anywhere
HTTPS	TCP	443	Anywhere

6. Click **Create**.

Create Security Group

Security group name: `pcf-web-elb-security-group`

Description: `pcf-web-elb-security-group`

VPC: `vpc-00bd5179 | jj-vpc`

Security group rules:

Inbound **Outbound**

Type	Protocol	Port Range	Source	Description
Custom TCP	TCP	4443	Anywhere	0.0.0.0/0, ::/0 e.g. SSH for Admin Desktop
HTTP	TCP	80	Anywhere	0.0.0.0/0, ::/0 e.g. SSH for Admin Desktop
HTTPS	TCP	443	Anywhere	0.0.0.0/0, ::/0 e.g. SSH for Admin Desktop

Add Rule

Cancel **Create**

Step 8: Configure a Security Group for the SSH ELB

1. From the **Security Groups** page, click **Create Security Group** to create another security group.
2. For **Security group name**, enter `pcf-ssh-elb-security-group`.
3. For **Description**, enter a description to identify this security group.
4. For **VPC**, select the VPC where you want to deploy this ELB.
5. Click the **Inbound** tab and add the following rule:

Type	Protocol	Port Range	Source
Custom TCP rule	TCP	2222	Anywhere

6. Click **Create**.

Step 9: Configure a Security Group for the TCP ELB

1. From the **Security Groups** page, click **Create Security Group** to create another security group.
2. For **Security group name**, enter `pcf-tcp-elb-security-group`.
3. For **Description**, enter a description to identify this security group.
4. For **VPC**, select the VPC where you want to deploy this ELB.
5. Click the **Inbound** tab and add the following rule:

Type	Protocol	Port Range	Source
Custom TCP rule	TCP	1024 - 1123	Anywhere 0.0.0.0/0

6. Click **Create**.

Step 10: Configure a Security Group for the Outbound NAT

1. From the **Security Groups** page, click **Create Security Group** to create another security group.
2. For **Security group name**, enter `pcf-nat-security-group`.
3. For **Description**, enter a description to identify this security group.
4. For **VPC**, select the VPC where you want to deploy the Outbound NAT.
5. Click the **Inbound** tab and add a rule to allow all traffic from your VPCs, as the table and image show.

Type	Protocol	Port Range	Source
All traffic	All	All	Custom IP 10.0.0.0/16

6. Click **Create**.

The screenshot shows the 'Inbound' tab of the AWS Security Groups rules configuration. It displays a single rule with the following details:

- Security group name:** OutboundNAT
- Description:** OutboundNAT
- VPC:** vpc-8ec593eb (10.0.0.0/16) | pcf-vpc
- Security group rules:** (Inbound selected)
- Type:** All traffic
- Protocol:** All
- Port Range:** 0 - 65535
- Source:** Custom IP 10.0.0.0/16

Step 11: Configure a Security Group for MySQL

Note: If you plan to use an internal database, skip this step. If you are using RDS, you must configure a security group that enables the Ops Manager VM and BOSH Director VM to access the database.

1. From the **Security Groups** page, click **Create Security Group** to create another security group.
2. For **Security group name**, enter `pcf-mysql-security-group`.
3. For **Description**, enter a description to identify this security group.
4. For **VPC**, select the VPC where you want to deploy MySQL.
5. Click the **Inbound** tab. Add a rule of type `MySQL` and specify the subnet of your VPC in **Source**, as the table and image show.

Type	Protocol	Port Range	Source
MySQL	TCP	3306	Custom IP 10.0.0.0/16

6. Click the **Outbound** tab. Add a rule of type **All traffic** and specify the subnet of your VPC in **Destination**, as the table and image show.

Type	Protocol	Port Range	Destination	
All traffic	All	All	Custom IP	10.0.0.0/16

7. Click **Create**.

Step 12: Launch an Ops Manager AMI

To launch an Amazon Machine Image (AMI) for Ops Manager, do the following:

1. Navigate to the **Pivotal Cloud Foundry Operations Manager** section of [Pivotal Network](#).
2. Select the version of PCF you want to install from the **Releases** dropdown.
3. In the **Release Download Files**, click the file named **Pivotal Cloud Foundry Ops Manager for AWS** to download a PDF.
4. Open the PDF and identify the AMI ID for your region.
5. Return to the EC2 Dashboard.
6. Click **AMIs** from the **Images** menu.
7. Select **Public images** from the drop-down filter that says **Owned by me**.
8. Paste the AMI ID for your region into the search bar and press enter.

Note: There is a different AMI for each region. If you cannot locate the AMI for your region, verify that you have set your AWS Management Console to your desired region. If you still cannot locate the AMI, log in to the [Pivotal Network](#) and file a support ticket.

9. (Optional) If you want to encrypt the VM that runs Ops Manager with AWS Key Management Service (KMS), perform the following additional steps:
 - a. Right click the row that lists your AMI and click **Copy AMI**.
 - b. Select your **Destination region**.
 - c. Enable **Encryption**. For more information about AMI encryption, see [Encryption and AMI Copy](#) from the *Copying an AMI* topic in the AWS documentation.
 - d. Select your **Master Key**. To create a new custom key, see [Creating Keys](#) in the AWS documentation.
 - e. Click **Copy AMI**. You can use the new AMI you copied for the following steps.
10. Select the row that lists your Ops Manager AMI and click **Launch**.
11. Choose **m3.large** for your instance type and click **Next: Configure Instance Details**.

	Family	Type	vCPUs	Memory (GiB)	Instance Storage (GB)
<input type="checkbox"/>	Micro instances	t1.micro <small>Free tier eligible</small>	1	0.613	EBS only
<input checked="" type="checkbox"/>	General purpose	t2.micro <small>Free tier eligible</small>	1	1	EBS only
<input checked="" type="checkbox"/>	General purpose	t2.small	1	2	EBS only
<input checked="" type="checkbox"/>	General purpose	t2.medium	2	4	EBS only
<input type="checkbox"/>	General purpose	m3.medium	1	3.75	1 x 4 (SSD)
<input checked="" type="checkbox"/>	General purpose	m3.large	2	7.5	1 x 32 (SSD)

12. Configure the following for your instance:

- Network: Select the VPC that you created.
- Subnet: Select `pcf-public-subnet-az0` to allow traffic from public IP addresses, or `pcf-management-subnet-az0` to allow traffic only from private IP addresses.
- Auto-assign for Public IP: Select **Enable** to allow traffic from public IP addresses, or **Disable** to allow traffic only from private IP addresses.
- IAM role: Select the IAM role associated with your pcf-user profile. If you have not created one, click **Create new IAM role** and follow the [Guidelines for Creating User Roles on AWS](#).
- For all other fields, accept the default values.

Step 3: Configure Instance Details

Configure the instance to suit your requirements. You can launch multiple instances from the same AMI, request Spot Instances to take advantage of the lower pricing, assign an access management role to the instance, and more.

Number of instances	<input type="text" value="1"/>
Purchasing option	<input type="checkbox"/> Request Spot Instances
Network	<input type="text" value="vpc-8ec593eb (10.0.0.0/16) pcf-vpc"/> <small>Create new VPC</small>
Subnet	<input type="text" value="subnet-fe8129d5(10.0.0.0/24) public-az1 l"/> <small>Create new subnet 250 IP Addresses available</small>
Auto-assign Public IP	<input type="text" value="Enable"/>
IAM role	<input type="text" value="None"/> <small>Create new IAM role</small>
Shutdown behavior	<input type="text" value="Stop"/>
Enable termination protection	<input type="checkbox"/> Protect against accidental termination
Monitoring	<input type="checkbox"/> Enable CloudWatch detailed monitoring <small>Additional charges apply.</small>
Tenancy	<input type="text" value="Shared tenancy (multi-tenant hardware)"/> <small>Additional charges will apply for dedicated tenancy.</small>

13. Click **Next: Add Storage** and adjust the **Size (GiB)** value. The default persistent disk value is 50 GB. Pivotal recommends increasing this value to a minimum of 100 GB.

Step 4: Add Storage

Your instance will be launched with the following storage device settings. You can attach additional EBS volumes and instance store volumes to your instance, or edit the settings of the root volume. You can also attach additional EBS volumes after launching an instance, but not instance store volumes. [Learn more](#) about storage options in Amazon EC2.

Type	Device	Snapshot	Size (GiB)	Volume Type	IOPS	Delete on Termination	Encrypted
Root	/dev/sda1	snap-f1ce6d7e	100	General Purpose (SSD)	150 / 3000	<input checked="" type="checkbox"/>	Not Encrypted
Instance Store 0	/dev/sdb	N/A	N/A	N/A	N/A	N/A	Not Encrypted
Add New Volume							

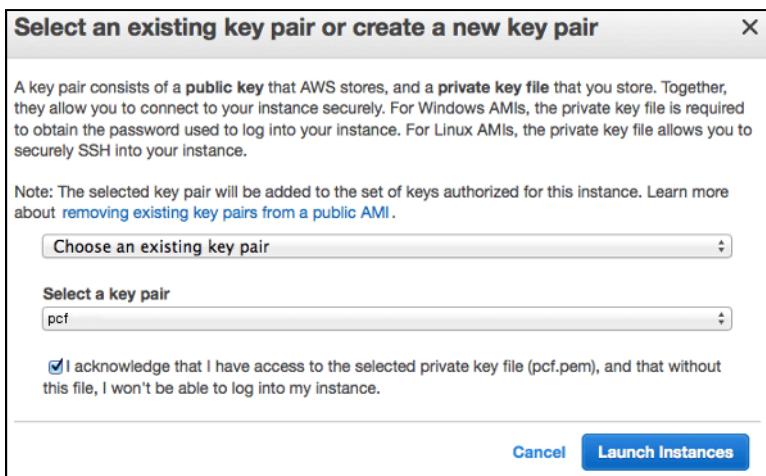
14. Click **Next: Tag Instance**.

15. On the **Add Tags** page, add a tag with the key `Name` and value `pcf-ops-manager`.

16. Click **Next: Configure Security Group**.

17. Select the `pcf-ops-manager-security-group` that you created in [Step 5: Configure a Security Group for Ops Manager](#).

18. Click **Review and Launch** and confirm the instance launch details.
19. Click **Launch**.
20. Select the `pcf-ops-manager-key` key pair, confirm that you have access to the private key file, and click **Launch Instances**. You use this key pair to access the Ops Manager VM.



21. Click **View Instances** to access the **Instances** page on the EC2 Dashboard.

Step 13: Create Web Load Balancer

1. On the EC2 Dashboard, click **Load Balancers**.
2. Click **Create Load Balancer**.
3. Under **Application Load Balancer**, click **Create**.
4. For **Step 1: Configure Load Balancer**, do the following:
 - a. Under **Basic Configuration**, do the following:
 - For **Name**, enter `pcf-web-elb`.
 - For **Scheme**, select **internet-facing** to allow traffic from public IP addresses, or **internal** to allow traffic only from private IP addresses.
 - For **IP address type**, select the type of IP addresses you want to allow.
 - b. Under **Listeners**, click **Add listener**. For **Load Balancer Protocol**, select **HTTPS**.
 - c. Under **Availability Zones**, select all availability zones.
 - d. Click **Next: Configure Security Settings**.
5. For **Step 2: Configure Security Settings**, do the following:
 - a. Under **Select default certificate**, do one of the following:
 - If you already have a certificate from AWS Certificate Manager (ACM), select **Choose a certificate from ACM**.
 - If you do not have a certificate from ACM, select **Upload a certificate to ACM**. For more information, see [Importing Certificates into AWS Certificate Manager](#) in the AWS documentation.

Note: For a production or production-like environment, use a certificate from a Certificate Authority (CA). This can be an internal certificate or a purchased certificate. For a sandbox environment, you can use a self-signed certificate.
 - b. For **Certificate Name**, select the desired certificate.
 - c. For **Security Policy**, select the policy you created in [Step 3: Create an IAM User for PCF](#).
 - d. Click **Next: Configure Security Groups**.
6. For **Step 3: Configure Security Groups**, do the following:
 - a. Under **Assign a security group**, select **Select an existing security group**.
 - b. From the list of security groups, select the `pcf-web-elb-security-group` security group that you configured in [Step 7: Configure a Security Group for the Web ELB](#).

c. Click **Next: Configure Routing**.

7. For **Step 4: Configure Routing**, do the following:

a. Under **Target Group**, enter the following values:

- **Name:** Enter `pcf-web-elb-target-group`.
- **Protocol:** Select **HTTP**.

b. Under **Health checks**, set **Path** to `/health`.

c. Under **Advanced health check settings**, enter the following values:

- **Port:** Set to `8080`.
- **Interval:** Set to `5` seconds.
- **Timeout:** Set to `3` seconds.
- **Unhealthy threshold:** Set to `3`.
- **Health threshold:** Set to `6`.

d. Click **Next: Register Targets**.

8. For **Step 5: Register Targets**, accept the default values and click **Next: Review**.

9. For **Step 6: Review**, review the load balancer details and then click **Create**. A message appears to confirm AWS successfully created the load balancer.

Step 14: Create SSH Load Balancer

1. From the **Load Balancers** page, click **Create Load Balancer**.

2. Select **Classic Load Balancer**.

3. Configure the load balancer with the following information:

- **Load Balancer name:** Enter `pcf-ssh-elb`.
- **Create LB Inside:** Select the `pcf-vpc` VPC that you created in [Step 4: Create a VPC](#).
- If you want to allow traffic from public IP addresses, ensure that the **Create an internal load balancer** checkbox is not selected. If you want to allow traffic only from private IP addresses, select this checkbox.

1. Define Load Balancer 2. Assign Security Groups 3. Configure Security Settings 4. Configure Health Check 5. Add EC2 Instances 6. Add Tags 7. Review

Step 1: Define Load Balancer

Basic Configuration

This wizard will walk you through setting up a new load balancer. Begin by giving your new load balancer a unique name so that you can identify it from other load balancers you might create. You will also need to configure ports and protocols for your load balancer. Traffic from your clients can be routed from any load balancer port to any port on your EC2 instances. By default, we've configured your load balancer with a standard web server on port 80.

Load Balancer name:	pcf-aws-lb		
Create LB Inside:	vpc-e2210d87 (10.0.0.0/16) pcf-vpc		
Create an internal load balancer:	<input type="checkbox"/> (what's this?)		
Enable advanced VPC configuration:	<input checked="" type="checkbox"/>		
Listener Configuration:			
Load Balancer Protocol	Load Balancer Port	Instance Protocol	Instance Port
HTTP	80	HTTP	80
Add			

Select Subnets

You will need to select a Subnet for each Availability Zone where you wish traffic to be routed by your load balancer. If you have instances in only one Availability Zone, please select at least two Subnets in different Availability Zones to provide higher availability for your load balancer.

VPC vpc-e2210d87 (10.0.0.0/16) | pcf-vpc

A Please select at least two Subnets in different Availability Zones to provide higher availability for your load balancer.

Available Subnets				
Actions	Availability Zone	Subnet ID	Subnet CIDR	Name
+	us-east-1c	subnet-ac5fc2f5	10.0.1.0/24	Private subnet

Selected Subnets				
Actions	Availability Zone	Subnet ID	Subnet CIDR	Name
-	us-east-1c	subnet-af5fc2f6	10.0.0.0/24	Public subnet

4. Under Listener Configuration, add the following rules:

Load Balancer Protocol	Load Balancer Port	Instance Protocol	Instance Port
TCP	2222	TCP	2222

5. Under Select Subnets, select either the public or private subnets you configured in Step 4: Create a VPC, and click Next: Assign Security Groups.

6. On the Assign Security Groups page, select the security group pcf-ssh-elb-security-group you configured in Step 8: Configure a Security Group for the SSH ELB, and click Next: Configure Security Settings.

1. Define Load Balancer 2. Assign Security Groups 3. Configure Security Settings 4. Configure Health Check 5. Add EC2 Instances 6. Add Tags 7. Review

Step 2: Assign Security Groups

You have selected the option of having your Elastic Load Balancer inside of a VPC, which allows you to assign security groups to your load balancer. Please select the security groups to assign to this load balancer. This can be changed at any time.

Assign a security group:	<input type="radio"/> Create a new security group	<input checked="" type="radio"/> Select an existing security group	
Filter VPC security groups			
Security Group ID	Name	Description	Actions
sg-41693b25	default	default VPC security group	Copy to new
sg-f6c29092	MySQL	MySQL	Copy to new
sg-10dd8f74	OpsManager	OpsManager	Copy to new
sg-3ec2905a	OutboundNAT	OutboundNAT	Copy to new
sg-88dd8fec	PCF VMs	PCF VMs	Copy to new
sg-48c2902c	PCF_ELB_SecurityGroup	PCF_ELB_SecurityGroup	Copy to new

7. On the Configure Security Settings page, ignore the Improve your load balancer's security error message and click Next: Configure Health Check.

8. On the Configure Health Check page, enter the following values:

- Ping Protocol: Select [TCP](#).
- Ping Port: Set to [2222](#).

- **Interval:** Set to seconds.
- **Response Timeout:** Set to seconds.
- **Unhealthy threshold:** Set to .
- **Health threshold:** Set to .

9. Click **Next: Add EC2 Instances**.

Step 4: Configure Health Check

Your load balancer will automatically perform health checks on your EC2 instances and only route traffic to instances that pass the health check. If an instance fails the health check, it is automatically removed from the load balancer. Customize the health check to meet your specific needs.

Ping Protocol	<input type="text" value="TCP"/>
Ping Port	<input type="text" value="80"/>
Advanced Details	
Response Timeout	<input type="text" value="5"/> seconds
Health Check Interval	<input type="text" value="10"/> seconds
Unhealthy Threshold	<input type="text" value="2"/>
Healthy Threshold	<input type="text" value="10"/>

10. Accept the defaults on the **Add EC2 Instances** page and click **Next: Add Tags**.

11. Accept the defaults on the **Add Tags** page and click **Review and Create**.

12. Review and confirm the load balancer details, and click **Create**.

Step 15: Create TCP Load Balancer

1. From the **Load Balancers** page, click **Create Load Balancer**.

2. Select **Classic Load Balancer**.

3. Configure the load balancer with the following information:

- **Load Balancer name:** Enter .
- **Create LB Inside:** Select the VPC that you created in [Step 4: Create a VPC](#).
- If you want to allow traffic from public IP addresses, ensure that the **Create an internal load balancer** checkbox is not selected. If you want to allow traffic only from private IP addresses, select this checkbox.

1. Define Load Balancer 2. Assign Security Groups 3. Configure Security Settings 4. Configure Health Check 5. Add EC2 Instances 6. Add Tags 7. Review

Step 1: Define Load Balancer

Basic Configuration

This wizard will walk you through setting up a new load balancer. Begin by giving your new load balancer a unique name so that you can identify it from other load balancers you might create. You will also need to configure ports and protocols for your load balancer. Traffic from your clients can be routed from any load balancer port to any port on your EC2 instances. By default, we've configured your load balancer with a standard web server on port 80.

Load Balancer name:	pcf-aws-lb		
Create LB Inside:	vpc-e2210d87 (10.0.0.0/16) pcf-vpc		
Create an internal load balancer:	<input type="checkbox"/> (what's this?)		
Enable advanced VPC configuration:	<input checked="" type="checkbox"/>		
Listener Configuration:			
Load Balancer Protocol	Load Balancer Port	Instance Protocol	Instance Port
HTTP	80	HTTP	80
<input type="button" value="Add"/>			

Select Subnets

You will need to select a Subnet for each Availability Zone where you wish traffic to be routed by your load balancer. If you have instances in only one Availability Zone, please select at least two Subnets in different Availability Zones to provide higher availability for your load balancer.

VPC vpc-e2210d87 (10.0.0.0/16) | pcf-vpc

⚠ Please select at least two Subnets in different Availability Zones to provide higher availability for your load balancer.

Available Subnets				
Actions	Availability Zone	Subnet ID	Subnet CIDR	Name
+	us-east-1c	subnet-ac5fc2f5	10.0.1.0/24	Private subnet

Selected Subnets				
Actions	Availability Zone	Subnet ID	Subnet CIDR	Name
-	us-east-1c	subnet-af5fc2f6	10.0.0.0/24	Public subnet

4. Under **Listener Configuration**, add the following rules:

Load Balancer Protocol	Load Balancer Port	Instance Protocol	Instance Port
TCP	1024	TCP	1024
TCP	1025	TCP	1025
TCP	1026	TCP	1026
...
TCP	1123	TCP	1123

The entry above indicates that you must add listening rules for each port between 1026 and 1123.

5. Under **Select Subnets**, select either the public or private subnets you configured in [Step 4: Create a VPC](#), and click [Next: Assign Security Groups](#).
6. On the **Assign Security Groups** page, select the security group you configured in [Step 9: Configure a Security Group for the TCP ELB](#), and click [Next: Configure Security Settings](#).

[1. Define Load Balancer](#) [2. Assign Security Groups](#) [3. Configure Security Settings](#) [4. Configure Health Check](#) [5. Add EC2 Instances](#) [6. Add Tags](#) [7. Review](#)

Step 2: Assign Security Groups

You have selected the option of having your Elastic Load Balancer inside of a VPC, which allows you to assign security groups to your load balancer. Please select the security groups to assign to this load balancer. This can be changed at any time.

Assign a security group:

- Create a new security group
- Select an existing security group

Filter

Security Group ID	Name	Description	Actions
sg-41693b25	default	default VPC security group	Copy to new
sg-f6c29092	MySQL	MySQL	Copy to new
sg-10dd8f74	OpsManager	OpsManager	Copy to new
sg-3ec2905a	OutboundNAT	OutboundNAT	Copy to new
sg-88dd8fec	PCF VMs	PCF VMs	Copy to new
sg-48c2902c	PCF_ELB_SecurityGroup	PCF_ELB_SecurityGroup	Copy to new

7. On the **Configure Security Settings** page, ignore the **Improve your load balancer's security** error message and click **Next: Configure Health Check**.
8. On the **Configure Health Check** page, enter the following values:
 - **Ping Protocol:** Select **TCP**.
 - **Ping Port:** Set to **80**.
 - **Interval:** Set to **5** seconds.
 - **Response Timeout:** Set to **3** seconds.
 - **Unhealthy threshold:** Set to **3**.
 - **Health threshold:** Set to **6**.

9. Click **Next: Add EC2 Instances**.

[1. Define Load Balancer](#) [2. Assign Security Groups](#) [3. Configure Security Settings](#) [4. Configure Health Check](#) [5. Add EC2 Instances](#) [6. Add Tags](#) [7. Review](#)

Step 4: Configure Health Check

Your load balancer will automatically perform health checks on your EC2 instances and only route traffic to instances that pass the health check. If an instance fails the health check, it is automatically removed from the load balancer. Customize the health check to meet your specific needs.

Ping Protocol	<input type="text" value="TCP"/>
Ping Port	<input type="text" value="80"/>
Advanced Details	
Response Timeout	<input type="text" value="5"/> seconds
Health Check Interval	<input type="text" value="10"/> seconds
Unhealthy Threshold	<input type="text" value="2"/>
Healthy Threshold	<input type="text" value="10"/>

10. Accept the defaults on the **Add EC2 Instances** page and click **Next: Add Tags**.
11. Accept the defaults on the **Add Tags** page and click **Review and Create**.
12. Review and confirm the load balancer details, and click **Create**.

Step 16: Configure DNS Records

1. Perform the following steps for all three of the load balancers you created in previous steps, named `pcf-web-elb`, `pcf-ssh-elb`, and `pcf-tcp-elb`:
 - a. From the **Load Balancers** page, select the load balancer.
 - b. On the **Description** tab, locate the **Basic Configuration** section and record the **DNS name** of the load balancer.
2. Click **Instances** on the left navigation to view your EC2 instances.
3. Select the `PcfOpsManInstance` instance created by Cloudformation.
4. On the **Description** tab, record the value for **IPv4 Public IP**.

5. Navigate to your DNS provider and create the following CNAME and A records:

- o CNAME: `*.apps.YOUR-SYSTEM-DOMAIN.com` and `*.system.YOUR-SYSTEM-DOMAIN.com` points to the DNS name of the `pcf-web-elb` load balancer.
- o CNAME: `ssh.YOUR-SYSTEM-DOMAIN.com` points to the DNS name of the `pcf-ssh-elb` load balancer.
- o CNAME: `tcp.YOUR-SYSTEM-DOMAIN.com` points to the DNS name of the `pcf-tcp-elb` load balancer.
- o A: `pcf.YOUR-SYSTEM-DOMAIN.com` points to the public IP address of the `pcf-ops-manager` EC2 instance.

Step 17: Secure the NAT Instance

1. On the EC2 Dashboard, click **Instances**.
2. Select the NAT instance, which has an instance type of **t3.medium**.
3. From the **Actions** menu, select **Networking>Change Security Groups**.
4. Change the NAT security group from the default group to the `pcf-nat-security-group` NAT security group that you created in [Step 10: Configure a NAT Instance](#)

Security Group ID	Name	Description
<input type="checkbox"/> sg-04bdda60	default	default VPC security group
<input type="checkbox"/> sg-6d85e209	ElasticRuntime	ElasticRuntime
<input type="checkbox"/> sg-7e80e71a	MySQL	MySQL
<input type="checkbox"/> sg-37b8df53	OpsManager	OpsManager
<input checked="" type="checkbox"/> sg-1482e570	OutboundNAT	OutboundNAT
<input type="checkbox"/> sg-0283e466	PCF_ELB_SecurityGroup	PCF_ELB_SecurityGroup

[Security Group for the Outbound NAT](#).

5. Click **Assign Security Groups**.

Step 18: Create RDS Subnet Group

1. Navigate to the RDS Dashboard.
2. Perform the following steps to create a RDS Subnet Group for the two RDS subnets:
 - a. Click **Subnet Groups>Create DB Subnet Group**.
 - b. Enter the following values:
 - **Name:** Enter `pcf-rds-subnet-group`.
 - **Description:** Enter a description to identify this subnet group.
 - **VPC ID:** Select `pcf-vpc`.
 - **Availability Zone and Subnet ID:** Choose the AZ and subnet for `pcf-rds-subnet-az0` and click **Add**.
 - c. Repeat the steps above to add `pcf-rds-subnet-az1` and `pcf-rds-subnet-az2` to the group.
 - d. Click **Create**.

The following screenshot shows a completed subnet group.

Create DB Subnet Group

To create a new Subnet Group give it a name, description, and select an existing VPC below. Once you select an existing VPC, you will be able to add subnets related to that VPC.

Name	<input type="text" value="PCF_RDSGroup"/>	
Description	<input type="text" value="PCF_RDSGroup"/>	
VPC ID	<input type="text" value="pcf-vpc (vpc-8ec593eb)"/>	

Add Subnet(s) to this Subnet Group. You may add subnets one at a time below or [add all the subnets](#) related to this VPC. You may make additions/edits after this group is created. A minimum of 2 subnets is required.

Availability Zone	<input type="text" value="us-east-1b"/>
Subnet ID	<input type="text" value="subnet-970765e0 (10.0.3.0/24)"/>
Add	

Availability Zone	Subnet ID	CIDR Block	Action
us-east-1b	subnet-970765e0	10.0.3.0/24	Remove
us-east-1a	subnet-119c343a	10.0.2.0/24	Remove

[Cancel](#) [Create](#)

Note: On the Subnet Group page, you may need to refresh the page to view the new group.

Step 19: Create a MySQL Database Using AWS RDS

Note: You must have an empty MySQL database when you install or reinstall PCF on AWS.

1. Navigate to the RDS Dashboard.
2. Click **Instances>Launch DB Instance** to launch the wizard.
3. Select **MySQL**.
4. Select the **MySQL** radio button under **Production** to create a database for production environments.
5. Click **Next Step**.
6. Specify the following database details:
 - o **DB Instance Class:** Select **db.m3.large - 2 vCPU, 7.5 GiB RAM**.
 - o **Multi-AZ Deployment:** Select **Yes**.
 - o **Storage Type:** Select **Provisioned IOPS (SSD)**.
 - o **Allocated Storage:** Enter **100 GB**.
 - o **DB Instance Identifier:** Enter **pcf-ops-manager-director**.
 - o Enter a secure **Master Username** and **Master Password**.

Note: Record the username and password. You need these credentials later when configuring the **Director Config** page in the BOSH Director tile.

Specify DB Details

Instance Specifications

DB Engine: mysql
 License Model: general-public-license
 DB Engine Version: 5.6.22

Review the Known Issues/Limitations to learn about potential compatibility issues with specific database versions.

DB Instance Class: db.m3.large — 2 vCPU, 7.5 GiB RA
 Multi-AZ Deployment: Yes
 Storage Type: General Purpose (SSD)
 Allocated Storage*: 100 GB

Settings

DB Instance Identifier*: pcf-bosh
 Master Username*: admin
 Master Password*:
 Confirm Password*:

Retype the value you specified for Master Password.

* Required [Cancel](#) [Previous](#) **Next Step**

7. Click **Next Step**.

8. On the **Configure Advanced Settings** page, enter the following values:

- o **VPC:** Select `pcf-vpc`.
- o **Subnet Group:** Select the `pcf-rds-subnet-group` you created in [Step 18: Create RDS Subnet Group](#).
- o **Publicly Accessible:** Select **No**.
- o **VPC Security Groups:** Select the `pcf-rds-security-group` that you created in [Step 11: Configure a Security Group for MySQL](#).
- o **Database Name:** Enter `bosh`.
- o Accept the default values for the remaining fields.

Configure Advanced Settings

Network & Security

This instance will be created with the new Certificate Authority rds-ca-2015. If you are using SSL to connect to this instance, you should use the [new certificate bundle](#). Learn more [here](#)

VPC*	<input type="text" value="pcf-vpc (vpc-e2210d87)"/>
Subnet Group	<input type="text" value="pcf_rdsgroup"/>
Publicly Accessible	<input type="text" value="No"/>
Availability Zone	<input type="text" value="No Preference"/>
VPC Security Group(s)	<input type="text" value="Create new Security Group"/> MySQL (VPC)
	OpsManager (VPC)
	OutboundNAT (VPC)

Database Options

Database Name	<input type="text" value="bosh"/>
---------------	-----------------------------------

Note: if no database name is specified then no initial MySQL database will be created on the DB Instance.

Database Port	<input type="text" value="3306"/>
DB Parameter Group	<input type="text" value="default.mysql5.6"/>
Option Group	<input type="text" value="default:mysql-5-6"/>
Enable Encryption	<input type="text" value="No"/>

Select to encrypt the given instance. Master key ids and aliases appear in the list after they have been created using the Key Management Service(KMS) console. [Learn More](#).

Backup

Please note that automated backups are currently supported for InnoDB storage engine only. If you are using MyISAM, refer to detail [here](#).

Backup Retention Period	<input type="text" value="7"/> days
Backup Window	<input type="text" value="No Preference"/>

Maintenance

Auto Minor Version Upgrade	<input type="text" value="Yes"/>
Maintenance Window	<input type="text" value="No Preference"/>

* Required
[Cancel](#)
[Previous](#)
Launch DB Instance

9. Click **Launch DB Instance**. Launching the instance may take several minutes.
10. When the instance has launched, proceed to [Configuring BOSH Director on AWS](#) to continue deploying PCF.

Required AWS Objects

This section describes the AWS objects you create in the procedures above in order to deploy PCF.

Use this section to determine the resource requirements of PCF on AWS, or to verify that you created the correct resources after completing the procedures above.

S3 Buckets for Ops Manager and PAS

You must create the following S3 buckets from the S3 Dashboard:

- pcf-ops-manager-bucket
- pcf-buildpacks-bucket
- pcf-packages-bucket
- pcf-resources-bucket
- pcf-droplets-bucket

These buckets must be empty when you install or reinstall PCF.

See [Step 2: Create S3 Buckets](#).

IAM User for PCF

You must create an IAM user for PCF named `pcf-user` from the Identity and Access Management Dashboard, using the policy document included in the [Pivotal Cloud Foundry for AWS Policy Document](#) topic.

See [Step 3: Create an IAM User for PCF](#).

Key Pair

You must generate a key pair named `pcf-ops-manager-key`. For more information about setting up a key pair, see [Amazon EC2 Key Pairs](#) in the AWS documentation.

VPC (Public and Private Subnets)

You must create a VPC with public and private subnets from the VPC Dashboard.

The following table lists the subnets in CIDR block `10.0.0.0/16`.

Name	AZ	IPv4 CIDR block
pcf-public-subnet-az0	REGION-#a (for example, us-west-2a)	10.0.0.0/24
pcf-public-subnet-az1	REGION-#b (for example, us-west-2b)	10.0.1.0/24
pcf-public-subnet-az2	REGION-#c (for example, us-west-2c)	10.0.2.0/24
pcf-management-subnet-az0	REGION-#a (for example, us-west-2a)	10.0.16.0/28
pcf-management-subnet-az1	REGION-#b (for example, us-west-2b)	10.0.16.16/28
pcf-management-subnet-az2	REGION-#c (for example, us-west-2c)	10.0.16.32/28
pcf-ert-subnet-az0	REGION-#a (for example, us-west-2a)	10.0.4.0/24
pcf-ert-subnet-az1	REGION-#b (for example, us-west-2b)	10.0.5.0/24
pcf-ert-subnet-az2	REGION-#c (for example, us-west-2c)	10.0.6.0/24
pcf-services-subnet-az0	REGION-#a (for example, us-west-2a)	10.0.8.0/24
pcf-services-subnet-az1	REGION-#b (for example, us-west-2b)	10.0.9.0/24
pcf-services-subnet-az2	REGION-#c (for example, us-west-2c)	10.0.10.0/24
pcf-rds-subnet-az0	REGION-#a (for example, us-west-2a)	10.0.12.0/24
pcf-rds-subnet-az1	REGION-#b (for example, us-west-2b)	10.0.13.0/24
pcf-rds-subnet-az2	REGION-#c (for example, us-west-2c)	10.0.14.0/24

See [Step 4: Create a VPC](#).

NAT Instance

You must create a NAT instance when creating a VPC. The NAT instance must have the following configuration:

- **Instance type:** t3.medium
- **Key pair name:** pcf-ops-manager-key
- **Enable DNS hostnames:** Yes
- **Hardware tenancy:** Default

See [Step 4: Create a VPC](#).

You must also assign the NAT instance to the pcf-nat-security-group. See [Step 17: Secure the NAT Instance](#).

Security Groups

The following sections describe the security groups you must create from the EC2 Dashboard.

Ops Manager

The Ops Manager Security Group must be named pcf-ops-manager-security-group and have the following inbound rules:

Type	Protocol	Port Range	Source
HTTP	TCP	80	My IP
HTTPS	TCP	443	My IP
SSH	TCP	22	My IP
BOSH Agent	TCP	6868	10.0.0.0/16
BOSH Director	TCP	25555	10.0.0.0/16

See [Step 5: Configure a Security Group for Ops Manager](#).

PCF VMs

The PCF VMs Security Group must be named pcf-vms-security-group and have the following inbound rule:

Type	Protocol	Port Range	Source
All traffic	All	0 - 65535	Custom IP 10.0.0.0/16

See [Step 6: Configure a Security Group for PCF VMs](#).

Web ELB

The Web ELB Security Group must be named pcf-web-elb-security-group and have the following inbound rules:

Type	Protocol	Port Range	Source
Custom TCP rule	TCP	4443	Anywhere 0.0.0.0/0
HTTP	TCP	80	Anywhere 0.0.0.0/0
HTTPS	TCP	443	Anywhere 0.0.0.0/0

See [Step 7: Configure a Security Group for the Web ELB](#).

SSH ELB

The SSH ELB Security Group must be named pcf-ssh-elb-security-group and have the following inbound rule:

Type	Protocol	Port Range	Source

Custom TCP rule	TCP	2222	Anywhere	0.0.0.0/0
-----------------	-----	------	----------	-----------

The SSH ELB Security Group must have the following outbound rule:

Type	Protocol	Port Range	Source	
All traffic	All	All	Anywhere	0.0.0.0/0

See [Step 8: Configure a Security Group for the SSH ELB.](#)

TCP ELB

The TCP ELB Security Group must be named `pcf-tcp-elb-security-group` and have the following inbound rule:

Type	Protocol	Port Range	Source	
Custom TCP rule	TCP	1024 - 1123	Anywhere	0.0.0.0/0

The TCP ELB Security Group must have the following outbound rule:

Type	Protocol	Port Range	Source	
All traffic	All	All	Anywhere	0.0.0.0/0

See [Step 9: Configure a Security Group for the TCP ELB.](#)

Outbound NAT

The Outbound NAT Security Group must be named `pcf-nat-security-group` and have the following inbound rule:

Type	Protocol	Port Range	Source	
All traffic	All	All	Custom IP	10.0.0.0/16

See [Step 10: Configure a Security Group for the Outbound NAT.](#)

MySQL

The MySQL Security Group must be named `pcf-mysql-security-group` and have the following inbound rules:

Type	Protocol	Port Range	Source	
MySQL	TCP	3306	Custom IP	10.0.0.0/16

The MySQL Security Group must have the following outbound rules:

Type	Protocol	Port Range	Destination	
All traffic	All	All	Custom IP	10.0.0.0/16

See [Step 11: Configure a Security Group for MySQL.](#)

Ops Manager AMI

You must locate the public Ops Manager AMI using the AMI ID provided by the PDF downloaded when clicking **Pivotal Cloud Foundry Ops Manager for AWS** on Pivotal Network.

See [Step 12: Launch a Pivotal Ops Manager AMI.](#)

ELBs

The following sections describe the ELBs you must create from the EC2 Dashboard.

Web ELB

You must create a web ELB with the following configuration:

- **Name:** `pcf-web-elb`
- **LB Inside:** `pcf-vpc`
- **Selected Subnet:** `pcf-public-subnet-az0`, `pcf-public-subnet-az1`, `pcf-public-subnet-az2`
- **Security Group:** `pcf-elb-security-group`
- **Health Check:** TCP Port 8080, Path: `/health`

See [Step 13: Create Web Load Balancer.](#)

SSH ELB

- **Name:** `pcf-ssh-elb`
- **LB Inside:** `pcf-vpc`
- **Selected Subnet:** `pcf-public-subnet-az0`, `pcf-public-subnet-az1`, `pcf-public-subnet-az2`
- **Security Group:** `pcf-ssh-security-group`
- **Health Check:** TCP Port 2222

See [Step 14: Create SSH Load Balancer.](#)

TCP ELB

- **Name:** `pcf-tcp-elb`
- **LB Inside:** `pcf-vpc`
- **Selected Subnet:** `pcf-public-subnet-az0`, `pcf-public-subnet-az1`, `pcf-public-subnet-az2`
- **Security Group:** `pcf-tcp-security-group`
- **Health Check:** TCP Port 80

See [Step 15: Create TCP Load Balancer.](#)

DNS Configuration

You must navigate to your DNS provider and create CNAME and A records for all three of your load balancers.

See [Step 16: Configure DNS Records.](#)

RDS Subnet Group

You must create a subnet group for RDS named `pcf-rds-subnet-group` from the RDS Dashboard.

See [Step 18: Create RDS Subnet Group.](#)

MySQL Database

You must create a MySQL database from the RDS Dashboard.

See [Step 19: Create a MySQL Database using AWS RDS.](#)

Configuring BOSH Director on AWS

This topic describes how to configure Ops Manager to deploy Pivotal Cloud Foundry (PCF) on Amazon Web Services (AWS).

Before beginning this procedure, ensure that you have successfully completed all of the steps in the [Installing PCF on AWS Manually](#) topic.

Step 1: Access Ops Manager

1. In a web browser, navigate to the fully qualified domain you created in the [Configure DNS Records](#) step of *Installing PCF on AWS Manually*.
2. When Ops Manager starts for the first time, you must choose one of the following:
 - [Internal Authentication](#): If you use Internal Authentication, Ops Manager maintains your user database.
 - [SAML Identity Provider](#): If you use a SAML Identity Provider (IdP), an external identity server maintains your user database.
 - [LDAP Server](#): If you use a LDAP Server, an external identity server maintains your user database.

Internal Authentication

1. When redirected to the [Internal Authentication](#) page, do the following:
 - Enter a **Username**, **Password**, and **Password confirmation** to create an Admin user.
 - Enter a **Decryption passphrase** and the **Decryption passphrase confirmation**. This passphrase encrypts the Ops Manager datastore, and is not recoverable.
 - If you are using an **HTTP proxy** or **HTTPS proxy**, follow the instructions in the [Configuring Proxy Settings for the BOSH CPI](#) topic.
 - Read the [End User License Agreement](#), and select the checkbox to accept the terms.
 - Click **Setup Authentication**.
2. Log in to Ops Manager with the Admin username and password you created in the previous step.

SAML Identity Provider

1. Log in to your IdP console and download the IdP metadata XML. Optionally, if your IdP supports metadata URL, you can copy the metadata URL instead of the XML.
2. Copy the IdP metadata XML or URL to the Ops Manager [SAML Identity Provider](#) login page.

 **Note:** The same IdP metadata URL or XML is applied for the BOSH Director. If you use a separate IdP for BOSH, copy the metadata XML or URL from that IdP and enter it into the BOSH IdP Metadata text box in the Ops Manager login page.

3. Enter values for the fields listed below. Failure to provide values in these fields results in a `500` error.
 - **SAML admin group**: Enter the name of the SAML group that contains all Ops Manager administrators. This field is case-sensitive.
 - **SAML groups attribute**: Enter the groups attribute tag name with which you configured the SAML server. This field is case-sensitive.
4. Enter your **Decryption passphrase**. Read the [End User License Agreement](#), and select the checkbox to accept the terms.
5. Your Ops Manager login page appears. Enter your username and password. Click **Login**.
6. Download your SAML Service Provider metadata (SAML Relying Party metadata) by navigating to the following URLs:
 - **6a. Ops Manager SAML service provider metadata**: <https://OPS-MAN-FQDN:443/uaa/saml/metadata>
 - **6b. BOSH Director SAML service provider metadata**: <https://BOSH-IP-ADDRESS:8443/saml/metadata>

 **Note:** To retrieve your `BOSH-IP-ADDRESS`, navigate to the **Status** tab in the **BOSH Director** tile. Record the **BOSH Director** IP address.

7. Configure your IdP with your SAML Service Provider metadata. Import the Ops Manager SAML provider metadata from Step 6a above to your IdP. If your IdP does not support importing, provide the values below.
 - **Single sign on URL**: <https://OPS-MAN-FQDN:443/uaa/saml/SSO/alias/OPS-MAN-FQDN>
 - **Audience URI (SP Entity ID)**: <https://OP-MAN-FQDN:443/uaa>
 - **Name ID**: Email Address

- SAML authentication requests are always signed
8. Import the BOSH Director SAML provider metadata from Step 6b to your IdP. If the IdP does not support an import, provide the values below.
- Single sign on URL: `https://BOSH-IP:8443/saml/SSO/alias/BOSH-IP`
 - Audience URI (SP Entity ID): `https://BOSH-IP:8443`
 - Name ID: Email Address
 - SAML authentication requests are always signed
9. Return to the **BOSH Director** tile, and continue with the configuration steps below.

LDAP Server

- For **Server URL**, enter the URL that points to your LDAP server. With multiple LDAP servers, separate their URLs with spaces. Each URL must include one of the following protocols:
 - `ldap://`: This specifies that the LDAP server uses an unencrypted connection.
 - `ldaps://`: This specifies that the LDAP server uses SSL for an encrypted connection and requires that the LDAP server holds a trusted certificate or that you import a trusted certificate to the JVM truststore.
- For **LDAP Username** and **LDAP Password**, enter the LDAP Distinguished Name (DN) and the password for binding to the LDAP Server. Example DN: `cn=administrator,ou=Users,dc=example,dc=com`

 **Note:** Pivotal recommends that you provide LDAP credentials that grant read-only permissions on the LDAP Search Base and the LDAP Group Search Base. In addition to this, if the bind user belongs to a different search base, you must use the full DN.

 **warning:** Pivotal recommends against reusing LDAP service accounts across environments. LDAP service accounts should not be subject to manual lockouts, such as lockouts that result from users utilizing the same account. Also, LDAP service accounts should not be subject to automated deletions, since disruption to these service accounts could prevent user logins.

- For **User Search Base**, enter the location in the LDAP directory tree from which any LDAP User search begins. The typical LDAP Search Base matches yourdomain name.

For example, a domain named “cloud.example.com” typically uses the following LDAP User Search Base: `ou=Users,dc=example,dc=com`

- For **User Search Filter**, enter a string that defines LDAP User search criteria. These search criteria allow LDAP to perform more effective and efficient searches. For example, the standard LDAP search filter `cn=Smith` returns all objects with a common name equal to `Smith`.

In the LDAP search filter string that you use to configure your runtime, use `{0}` instead of the username. For example, use `cn={0}` to return all LDAP objects with the same common name as the username.

In addition to `cn`, other attributes commonly searched for and returned are `mail`, `uid` and, in the case of Active Directory, `sAMAccountName`.

 **Note:** For instructions for testing and troubleshooting your LDAP search filters, see [Configuring LDAP Integration with Pivotal Cloud Foundry](#) in the Pivotal Support Knowledge Base.

- For **Group Search Base**, enter the location in the LDAP directory tree from which the LDAP Group search begins.

For example, a domain named “cloud.example.com” typically uses the following LDAP Group Search Base: `ou=Groups,dc=example,dc=com`

- For **Group Search Filter**, enter a string that defines LDAP Group search criteria. The standard value is `member={0}`.
- For **Email Attribute**, enter the attribute name in your LDAP directory that corresponds to the email address in each user record, for example `mail`.
- For **LDAP RBAC Admin Group Name**, enter the DN of the LDAP group you want to have admin permissions in Ops Manager.

- From the dropdown, select how the UAA handles LDAP server referrals out to other external user stores. The UAA can:

- Automatically follow any referrals.
 - Ignore referrals and return partial result.
 - Throw exception for each referral and abort.
- For **Server SSL Cert**, paste in the root certificate from your CA certificate or your self-signed certificate.
 - Enter a **Decryption passphrase** and the **Decryption passphrase confirmation**. This passphrase encrypts the Ops Manager datastore, and is not recoverable.
 - If you are using an **HTTP proxy** or **HTTPS proxy**, follow the instructions in [Configuring Proxy Settings for the BOSH CPI](#).
 - Read the **End User License Agreement**, and select the checkbox to accept the terms.

- Select **Provision an admin client in the BOSH UAA**. You can use this to enable BOSH automation with scripts and tooling. For more information, see [Provision Admin Client](#) in *Creating UAA Clients for BOSH Director*.
- Click **Setup Authentication**.
- Return to the **BOSH Director** tile, and continue with the configuration steps below.

Step 2: AWS Config Page

1. Click the **BOSH Director** tile.

2. Select **AWS Config**.

AWS Management Console Config

Use AWS Keys

Access Key ID*

(i)

AWS Secret Key*

(i) The AWS Secret Key

Use AWS Instance Profile

AWS IAM Instance Profile*

Security Group ID*

Key Pair Name*

SSH Private Key*

Region*

Encrypt Linux EBS Volumes

Custom Encryption Key

Save

3. Select **Use AWS Keys** or **Use AWS Instance Profile**.

- **Access Key ID** and **AWS Secret Key**: To retrieve your AWS key information, use the AWS Identity and Access Management (IAM) credentials that you generated in the [Step 3: Create an IAM User for PCF](#) section of the *Installing PCF on AWS Manually* topic.

- **AWS IAM Instance Profile:** Enter the name of the IAM profile you created in the [Step 3: Create an IAM User for PCF](#) section of the *Installing PCF on AWS Manually* topic.

4. Complete the rest of the **AWS Management Console Config** page:

- **Security Group ID:** Enter the **Group ID** of the `pcf-vms-security-group` you created for your PCF VMs in the [Step 6: Configure a Security Group for PCF VMs](#) section of the *Installing PCF on AWS Manually* topic. Locate the Group ID in the **Security Groups** tab of your EC2 Dashboard.
- **Key Pair Name:** Enter `pcf-ops-manager-key`.
- **SSH Private Key:** Open the AWS key pair `pcf-ops-manager-keys.pem` file you generated in the [Step 3: Create an IAM User for PCF](#) section of the *Installing PCF on AWS Manually* topic. Copy the contents of the `.pem` file and paste it into the **SSH Private Key** field.
- **Region:** Select the region where you deployed Ops Manager.
- **Encrypt Linux EBS Volumes:** Select this checkbox to enable full encryption on persistent disks of all BOSH-deployed VMs except the Ops Manager VM and Director VM. See the [Configuring Amazon EBS Encryption for PCF on AWS](#) topic for details about using EBS encryption.

Note: Enabling EBS encryption only encrypts Linux VMs. The Windows VMs deployed with Pivotal Application Service (PAS) for Windows or PAS for Windows 2012R2 are not encrypted.

5. Click **Save**.

Step 3: Director Config Page

1. In Ops Manager, select **Director Config**.

2. Enter at least two of the following NTP servers in the **NTP Servers (comma delimited)** field, separated by a comma:

`0.amazon.pool.ntp.org,1.amazon.pool.ntp.org,2.amazon.pool.ntp.org,3.amazon.pool.ntp.org`

Note: The NTP server configuration only updates after VM recreation. Ensure that you select the **Recreate all VMs** checkbox if you modify the value of this field.

3. Leave the **Bosh HM Forwarder IP Address** field blank.

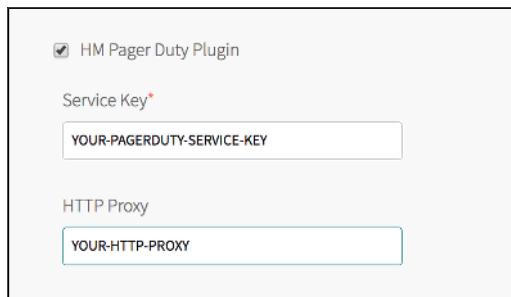
Note: Starting in PAS v2.0, BOSH-reported component metrics are available in the Loggregator Firehose by default. If you continue to use the BOSH HM Forwarder to consume these component metrics, you may receive duplicate data. To prevent this, leave the **Bosh HM Forwarder IP Address** field blank.

4. Select the **Enable VM Resurrector Plugin** checkbox to enable the BOSH Resurrector functionality and increase your runtime availability.

5. Select **Enable Post Deploy Scripts** to run a post-deploy script after deployment. This script allows the job to execute additional commands against a deployment.

 **Note:** If you intend to install PKS, you must enable post-deploy scripts.

6. Select **Recreate all VMs** to force BOSH to recreate all VMs on the next deploy. This process does not destroy any persistent disk data.
7. Select **Recreate All Persistent Disks** to force BOSH to migrate and recreate persistent disks for the BOSH Director and all tiles. This process does not destroy any persistent disk data.
8. Select **Enable bash deploy retries** to instruct Ops Manager to retry failed BOSH operations up to five times.
9. Select **Skip Director Drain Lifecycle** to prevent drain scripts from running when the BOSH Director is recreated.
10. Select **Store BOSH Job Credentials on tmpfs (beta)** to store credentials for BOSH jobs on temporary file storage (tmpfs) memory, rather than on disk. You must recreate all VMs for this setting to take effect.
11. Select **Keep Unreachable Director VMs** if you want to preserve BOSH Director VMs after a failed deployment for troubleshooting purposes.
12. Select **HM Pager Duty Plugin** to enable Health Monitor integration with PagerDuty.



HM Pager Duty Plugin

Service Key*

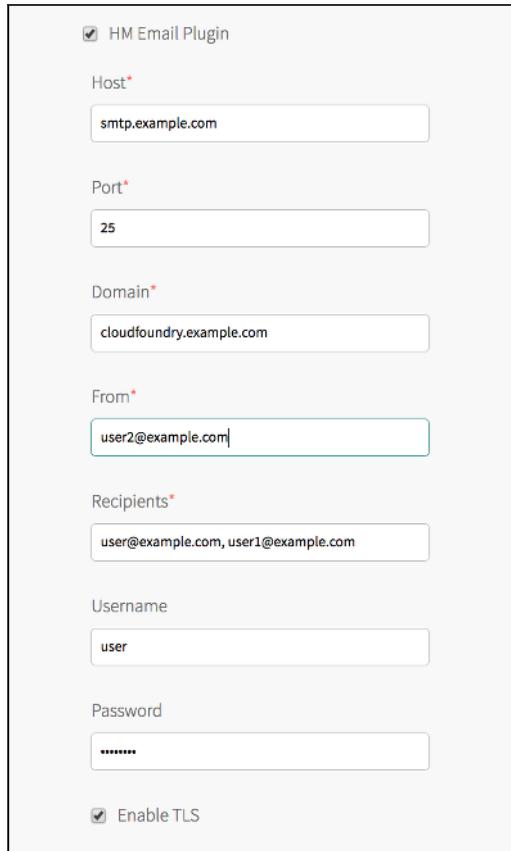
YOUR-PAGERDUTY-SERVICE-KEY

HTTP Proxy

YOUR-HTTP-PROXY

- **Service Key:** Enter your API service key from PagerDuty.
- **HTTP Proxy:** Enter an HTTP proxy for use with PagerDuty.

13. Select **HM Email Plugin** to enable Health Monitor integration with email.



HM Email Plugin

Host*

smtp.example.com

Port*

25

Domain*

cloudfoundry.example.com

From*

user2@example.com

Recipients*

user@example.com, user1@example.com

Username

user

Password

.....

Enable TLS

- **Host:** Enter your email hostname.
- **Port:** Enter your email port number.
- **Domain:** Enter your domain.
- **From:** Enter the address for the sender.
- **Recipients:** Enter comma-separated addresses of intended recipients.
- **Username:** Enter the username for your email server.
- **Password:** Enter the password for your email server.
- **Enable TLS:** Select this checkbox to enable Transport Layer Security to the email host.

14. For **CredHub Encryption Provider**, you can choose whether BOSH CredHub stores its encryption key internally on the BOSH Director and CredHub VM, or in an external hardware security module (HSM). The HSM option is more secure.

Before configuring an HSM encryption provider in the **Director Config** pane, you must follow the procedures and collect information described in [Preparing CredHub HSMs for Configuration](#).

 **Note:** After you deploy Ops Manager with an HSM encryption provider, you cannot change BOSH CredHub to store encryption keys internally.

CredHub Encryption Provider

Internal
 Luna HSM

Encryption Key Name*

Provider Partition*

Provider Partition Password*

Provider Client Certificate*

Provider Client Certificate Private Key*

HSM Host Address*

HSM Port Address*

Partition Serial Number*

HSM Certificate*

- **Internal:** Select this option for internal CredHub key storage. This option is selected by default and requires no additional configuration.
 - **Luna HSM:** Select this option to use a SafeNet Luna HSM as your permanent CredHub encryption provider, and fill in the following fields:
 1. **Encryption Key Name:** Any name to identify the key that the HSM uses to encrypt and decrypt the CredHub data. Changing this key name after you deploy Ops Manager can cause service downtime.
 2. **Provider Partition:** The partition that stores your encryption key. Changing this partition after you deploy Ops Manager could cause service downtime. For this value and the ones below, use values gathered in [Preparing CredHub HSMs for Configuration](#).
 3. **Provider Partition Password**
 4. **Provider Client Certificate:** The certificate that validates the identity of the HSM when CredHub connects as a client.
 5. **Provider Client Certificate Private Key**
 6. **HSM Host Address**
 7. **HSM Port Address:** If you do not know your port address, enter `1792`.
 8. **Partition Serial Number**
 9. **HSM Certificate:** The certificate that the HSM presents to CredHub to establish a two-way mTLS connection.
15. Select a **Blobstore Location** to either configure the blobstore as an internal server or an external endpoint. Because the internal server is unscalable and less secure, Pivotal recommends that you configure an external blobstore.

Note: After you deploy Ops Manager, you cannot change the blobstore location.

- **Internal:** Select this option to use an internal blobstore. Ops Manager creates a new VM for blob storage. No additional configuration is required.
 - **Enable TLS:** Select this checkbox to enable TLS to the blobstore.

Note: If you are using PASW 2016, make sure you have downloaded [Windows stemcell v1709.10 or higher](#) before enabling TLS.

- **S3 Compatible Blobstore:** Select this option to use an external S3-compatible endpoint. When you have created an S3 bucket, complete the

The form is a configuration panel for an S3 Compatible Blobstore. It includes the following sections and fields:

- Blobstore Location:**
 - Internal
 - S3 Compatible Blobstore
- S3 Endpoint***: An input field.
- Bucket Name***: An input field.
- Access Key***: An input field.
- Secret Key***: An input field.
- S3 Signature Version:**
 - V2 Signature
 - V4 Signature
- Region***: An input field.
- S3 Backup Strategy:**
 - No backups
 - Use a versioned bucket
 - Copy into an additional bucket
- Backup Bucket Region***: An input field.
- Backup Bucket Name***: An input field.

following steps:

Follow the procedures in [Sign up for Amazon S3](#) and [Creating a Bucket](#) in the AWS documentation. When you have created an S3 bucket, complete the following steps:

1. **S3 Endpoint:** Navigate to the [Regions and Endpoints](#) topic in the AWS documentation.
 - a. Locate the endpoint for your region in the **Amazon Simple Storage Service (S3)** table and construct a URL using your region's endpoint. For example, if you are using the `us-west-2` region, the URL you create would be `https://s3-us-west-2.amazonaws.com`. Enter this URL into the **S3 Endpoint** field.
 - b. On a command line, run `ssh ubuntu@OPS-MANAGER-FQDN` to SSH into the Ops Manager VM. Replace `OPS-MANAGER-FQDN` with the fully qualified domain name of Ops Manager.
 - c. Copy the custom public CA certificate you used to sign the S3 endpoint into `/etc/ssl/certs` on the Ops Manager VM.

- d. On the Ops Manager VM, run `sudo update-ca-certificates -f -v` to import the custom CA certificate into the Ops Manager VM truststore.

Note: You must also add this custom CA certificate into the **Trusted Certificates** field in the **Security** page. See [Security Page](#) for instructions.

2. **Bucket Name:** Enter the name of the S3 bucket.
3. **Access Key** and **Secret Key:** Enter the keys you generated when creating your S3 bucket.
4. Select **V2 Signature** or **V4 Signature**. If you select **V4 Signature**, enter your **Region**.

Note: AWS recommends using Signature Version 4. For more information about AWS S3 Signatures, see [Authenticating Requests](#) in the AWS documentation.

5. **S3 Backup Strategy:** Select one of the following to configure whether and how to back up the BOSH Director's S3 blobstore:

- **No backups**
- **Use a versioned bucket:** This option uses blobstore bucket versioning to save backups. It requires an S3 blobstore that supports versioning.
- **Copy into an additional bucket:** This option saves blobstore backups to a separate bucket, which is useful with blobstores that do not support versioning. The option requires you to fill in **Backup Bucket Region** and **Backup Bucket Name**.

- **GCS Blobstore:** Select this option to use an external Google Cloud Storage (GCS) endpoint.

The screenshot shows a configuration form for a GCS Blobstore. At the top, there is a radio button labeled "GCS Blobstore". Below it, there are three input fields: "Bucket Name*" with a placeholder box, "Storage Class*" with a dropdown menu set to "Regional", and "Service Account Key*" with a large text area for pasting JSON content.

Follow the procedures in [Creating Storage Buckets](#) in the GCS documentation to create a GCS bucket. To create a GCS bucket, you must have a GCS account. When you have created a GCS bucket, complete the following steps:

1. **Bucket Name:** Enter the name of your GCS bucket.
2. **Storage Class:** Select the storage class for your GCS bucket. See [Storage Classes](#) in the GCP documentation for more information.
3. **Service Account Key:** Follow the steps in the [Set up IAM Service Accounts](#) section of [Preparing to Deploy Ops Manager on GCP Manually](#) to download a JSON file with a private key. Enter the contents of the JSON file into the field.

16. For **Database Location**, select **External MySQL Database** and complete the following steps:

Database Location

Internal

External MySQL Database

Host*

Port*

Username*

Password*

Database*

- From the AWS Console, navigate to the RDS Dashboard.
- Select **Instances**, then click the arrow to the left of your instance and select the second icon to display the **Details** information.
- Refer to the following table to retrieve the values for the **Director Config** page:

RDS Instance Field	BOSH Director Field
Endpoint	Host
Port	Port, which is <code>3306</code> .
DB Name	Database, which is <code>bosh</code> .
Username	Username

- For **Password**, enter the password that you defined for your MySQL database when you created in the [Step 19: Create a MySQL Database using AWS RDS](#) section of the *Installing PCF on AWS Manually* topic. In addition, if you selected the **Enable TLS for Director Database** checkbox, you can fill out the following optional fields:
- Enable TLS:** Selecting this checkbox enables TLS communication between the BOSH Director and the database.
- TLS CA:** Enter the Certificate Authority for the TLS Certificate.
- TLS Certificate:** Enter the client certificate for mutual TLS connections to the database.
- TLS Private Key:** Enter the client private key for mutual TLS connections to the database.
- Advanced DB Connection Options:** If you would like to provide additional options for the database, use this field to provide a JSON-formatted options string.

- (Optional) Modify the **Director Workers** value, which sets the number of workers available to execute Director tasks. This field defaults to `5`.
- (Optional) **Max Threads** sets the maximum number of threads that the BOSH Director can run simultaneously. Pivotal recommends that you leave the field blank to use the default value, unless doing so results in rate limiting or errors on your IaaS.
- (Optional) To add a custom URL for your BOSH Director, enter a valid hostname in **Director Hostname**. You can also use this field to configure a load balancer in front of your BOSH Director. For more information, see [How to Set Up a Load Balancer in Front of Operations Manager Director](#) in the Pivotal Support Knowledge Base.

Director Workers

Max Threads

Director Hostname

20. (Optional) To set a custom banner that users see when logging in to the Director using SSH, enter text in the **Custom SSH Banner** field.
21. (Optional) Enter your comma-separated custom **Identification Tags**. For example, `iaas:foundation1, hello:world`. You can use the tags to identify your foundation when viewing VMs or disks from your IaaS.
22. Click **Save**.

Step 4: Create Availability Zones Page

 **Note:** Pivotal recommends at least three availability zones (AZs) for a highly available installation of PAS. The procedures in [Installing PCF on AWS Manually](#) use 3 AZs.

1. Select **Create Availability Zones**.

Create Availability Zones

Availability Zones

▼ us-west-1c

Amazon Availability Zone*

The Amazon Availability Zone name (ex: 'us-east-1b')

Add 

Save

2. To add the three AZs you specified in the [Step 4: Create a VPC](#) section of the [Installing PCF on AWS Manually](#) topic, do the following:
 - a. Click **Add**.
 - b. For **Amazon Availability Zone**, enter the name of the AZ.
 - c. Repeat until you have entered all three AZs, in the format `REGION-#a`, `REGION-#b`, and `REGION-#c`. For example, `us-west-2a`, `us-west-2b`, and `us-west-2c`.
3. Click **Save**.

Step 5: Create Networks Page

Create Networks

Warning: Pivotal recommends keeping the IP settings throughout the life of your installation. Ops Manager may prevent you from changing them in the future.
Contact Pivotal support for help completing such a change.

Verification Settings

Enable ICMP checks

Networks

One or many IP ranges upon which your products will be deployed

Subnets

Add Network

Name*
infrastructure

VPC Subnet ID*
pcf-infrastructure-subnet-az0

CIDR*
10.0.16.0/28

Reserved IP Ranges
10.0.16.0-10.0.16.4

DNS*
10.0.16.2

Gateway*
10.0.16.1

Availability Zones*

us-west-2a
 us-west-2b
 us-west-2c

1. Select **Create Networks**.
2. (Optional) Select **Enable ICMP checks** to enable ICMP on your networks. Ops Manager uses ICMP checks to confirm that components within your network are reachable.
3. Perform the following steps to add the network configuration that you created for your VPC in the [Create a VPC](#) section of *Installing PCF on AWS Manually*. Record your VPC CIDR if you set a CIDR other than the recommendation!
 - a. Click **Add Network**.
 - b. For **Name**, enter `infrastructure`.
 - c. Create a subnet for each availability zone by clicking **Add Subnet**. Refer to the table below for the information required to create all three subnets:

First Subnet	VPC Subnet ID	pcf-infrastructure-subnet-az0
	CIDR	10.0.16.0/28
	Reserved IP Ranges	10.0.16.0-10.0.16.4
	DNS	10.0.0.2 *
	Gateway	10.0.16.1
	Availability Zones	REGION-#a . Example: us-west-2a

Second Subnet	VPC Subnet ID	pcf-infrastructure-subnet-az1
	CIDR	10.0.16.16/28
	Reserved IP Ranges	10.0.16.16-10.0.16.20
	DNS	10.0.0.2 *
	Gateway	10.0.16.17
Third Subnet	Availability Zones	REGION-#b . Example: us-west-2b
	VPC Subnet ID	pcf-infrastructure-subnet-az2
	CIDR	10.0.16.32/28
	Reserved IP Ranges	10.0.16.32-10.0.16.36
	DNS	10.0.0.2 *
	Gateway	10.0.16.33
Third Subnet	Availability Zones	REGION-#c . Example: us-west-2c

* If you set a VPC CIDR other than recommended, enter the second IP in your VPC CIDR. For example, for a 10.0.0/24 VPC CIDR, enter 10.0.0.2 in each subnet.

- d. Click Add Network.
- e. For Name, enter the name of your runtime. For example, pas or pks .
- f. Create a subnet for each availability zone by clicking Add Subnet. See the table below for the information required to create all three subnets:

First Subnet	VPC Subnet ID	pcf-pas-subnet-az0 or pcf-pks-subnet-az0
	CIDR	10.0.4.0/24
	Reserved IP Ranges	10.0.4.0-10.0.4.4
	DNS	10.0.0.2 *
	Gateway	10.0.4.1
	Availability Zones	REGION-#a . Example: us-west-2a
Second Subnet	VPC Subnet ID	pcf-pas-subnet-az1 or pcf-pks-subnet-az1
	CIDR	10.0.5.0/24
	Reserved IP Ranges	10.0.5.0-10.0.5.4
	DNS	10.0.0.2 *
	Gateway	10.0.5.1
	Availability Zones	REGION-#b . Example: us-west-2b
Third Subnet	VPC Subnet ID	pcf-pas-subnet-az2 or pcf-pks-subnet-az2
	CIDR	10.0.6.0/24
	Reserved IP Ranges	10.0.6.0-10.0.6.4
	DNS	10.0.0.2 *
	Gateway	10.0.6.1
	Availability Zones	REGION-#c . Example: us-west-2c

* If you set a VPC CIDR other than recommended, enter the second IP in your VPC CIDR. For example, for a 10.0.0/24 VPC CIDR, enter 10.0.0.2 in each subnet.

- g. Click Add Network.
- h. For Name, enter services .
- i. Create a subnet for each availability zone by clicking Add Subnet. Refer to the table below for the information required to create all three subnets:

First Subnet	VPC Subnet ID	pcf-services-subnet-az0
	CIDR	10.0.8.0/24
	Reserved IP Ranges	10.0.8.0-10.0.8.3
	DNS	10.0.0.2 *
	Gateway	10.0.8.1
	Availability Zones	REGION-#a . Example: us-west-2a

Second Subnet	VPC Subnet ID	pcf-services-subnet-az1
	CIDR	10.0.9.0/24
	Reserved IP Ranges	10.0.9.0-10.0.9.3
	DNS	10.0.0.2 *
	Gateway	10.0.9.1
	Availability Zones	REGION-#b . Example: us-west-2b
Third Subnet	VPC Subnet ID	pcf-services-subnet-az2
	CIDR	10.0.10.0/24
	Reserved IP Ranges	10.0.10.0-10.0.10.3
	DNS	10.0.0.2 *
	Gateway	10.0.10.1
	Availability Zones	REGION-#c . Example: us-west-2c

* If you set a VPC CIDR other than recommended, enter the second IP in your VPC CIDR. For example, for a 10.0.0.0/24 VPC CIDR, enter 10.0.0.2 in each subnet.

- Click **Save**.

Note: If you are deploying PKS with a PKS workload load balancer, you must tag each AWS subnet with your PKS Kubernetes cluster unique identifier before you create the load balancer. For more information about tagging subnets with a PKS cluster unique identifier, see [AWS Prerequisites](#).

Note: After you deploy Ops Manager, you add subnets with overlapping Availability Zones to expand your network. For more information about configuring additional subnets, see [Expanding Your Network with Additional Subnets](#).

Step 6: Assign AZs and Networks

- Select **Assign AZs and Networks**.

Assign AZs and Networks

The BOSH Director is a single instance.

Choose the availability zone in which to place that instance. It is highly recommended that you backup this VM on a regular basis to preserve settings.

Singleton Availability Zone

AZ-MGMT

Network

infrastructure

Save

- Use the dropdown to select a **Singleton Availability Zone**. The BOSH Director is deployed into this AZ.

- Use the dropdown to select `infrastructure` under **Network**. The BOSH Director is deployed into this network.

- Click **Save**.

Step 7: Security Page

- Select **Security**.

Security

Trusted Certificates

These certificates enable BOSH-deployed components to trust a custom root certificate.

Include OpsManager Root CA in Trusted Certs

Checking this box will place the OpsManager Root CA into the trusted certificate field. The Bosh Director will then place this CA into the trust store of every VM that it deploys.

Generate VM passwords or use single password for all VMs

Generate passwords

Use default BOSH password

Save

2. In **Trusted Certificates**, enter your custom certificate authority (CA) certificates to insert into your organization's certificate trust chain. This feature enables all BOSH-deployed components in your deployment to trust custom root certificates.

To enter multiple certificates, paste your certificates one after the other. For example, format your certificates like the following:

```
-----BEGIN CERTIFICATE-----
ABCDEFGH12345678ABCDEFGH12345678ABCDEFGH12345678AB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
BCDEFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
CDEFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABBB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
```

Note: If you want to use Docker registries to run PAS app instances in Docker containers, enter the certificate for your private Docker registry in this field. See [Using Docker Registries](#) for more information about running app instances in PAS using Docker registries.

3. (Optional) Select the **Include OpsManager Root CA in Trusted Certs** checkbox to include the Ops Manager root CA in the Trusted Certificates field. BOSH Director includes this CA in the trust store of every VM that it deploys.
4. Choose **Generate passwords** or **Use default BOSH password**. Pivotal recommends that you use the **Generate passwords** option for greater security.
5. Click **Save**. To view your saved Director password, click the **Credentials** tab.

Step 8: BOSH DNS Config Page

BOSH DNS Config

Excluded Recursors

Recurser Timeout

Handlers

[]

Enter a JSON-formatted list of optional custom domain handlers

Save

1. Select **BOSH DNS Config**.
2. (Optional) In **Excluded Recursors**, enter a list of prohibited recursor addresses.
3. (Optional) In **Recurser Timeout**, enter a time limit for contacting the connected recursors. This includes dialing, writing, and reading from the recursor. If any of these actions exceeds the time limit you set, the action fails.

Note: This time limit must include one of the Go parse duration time units. For example, entering `5s` sets the timeout limit to five seconds. For more information about supported time units, see [func ParseDuration](#) in the Go Programming Language documentation.

4. (Optional) In **Handlers**, enter a list of custom domain handlers in JSON format.
5. Click **Save**.

Step 9: Syslog Page

Settings

Syslog

Syslog settings make take up to 1 min to take effect.

Do you want to configure Syslog for OpsManager?

No
 Yes

Address*
 The address or host for the syslog server

Port*

Transport Protocol*

Enable TLS

Permitted Peer*

SSL Certificate*

Queue Size

Forward Debug Logs

Custom rsyslog Configuration

SAVE SYSLOG SETTINGS

1. Select **Syslog**.
2. (Optional) Select **Yes** to send BOSH Director system logs to a remote server.
3. Enter the IP address or DNS name for the remote server in **Address**.
4. Enter the port number that the remote server listens on in **Port**.
5. Select **TCP** or **UDP** from the **Transport Protocol** dropdown. This selection determines which transport protocol is used to send the logs to the remote server.
6. (Optional) Select the **Enable TLS** checkbox to send encrypted logs to remote server with TLS. After you select the checkbox, do the following:
 - a. Enter either the name or SHA1 fingerprint of the remote peer in **Permitted Peer**.
 - b. Enter the SSL certificate for the remote server in **SSL Certificate**.



Note: Pivotal strongly recommends that you enable TLS encryption when you are forwarding logs. Logs can contain sensitive

information, such as cloud provider credentials.

7. (Optional) Enter an integer in **Queue Size**. This value specifies the number of log messages held in the buffer. The default value is 100,000.
8. (Optional) Select the checkbox to **Forward Debug Logs** to an external source. This option is deselected by default. If you select it, you may generate a large amount of log data.
9. (Optional) Enter configuration details for rsyslog in the **Custom rsyslog Configuration** field. This field requires the rainerscript syntax.
10. Click **Save Syslog Settings**.

Step 10: Resource Config Page

1. Select **Resource Config**.

The screenshot shows the 'Resource Config' page with two jobs listed:

- BOSH Director**: Instances: Automatic: 1, Persistent Disk Type: Automatic: 50 GB, VM Type: Automatic: large.disk (cpu: 2, ram: 8 GB, disk: 50 GB), Load Balancers: None, Internet Connected: checked.
- Master Compilation Job**: Instances: 1, Persistent Disk Type: None, VM Type: small (cpu: 1, ram: 2 GB, disk: 8 GB), Load Balancers: None, Internet Connected: checked.

A blue 'Save' button is at the bottom left.

2. Adjust any values as necessary for your deployment. Under the **Instances**, **Persistent Disk Type**, and **VM Type** fields, choose **Automatic** from the dropdown to allocate the recommended resources for the job. If the **Persistent Disk Type** field reads **None**, the job does not require persistent disk space.

Note: Pivotal recommends provisioning a BOSH Director VM with at least 8 GB memory.

Note: If you set a field to **Automatic** and the recommended resource allocation changes in a future version, Ops Manager automatically uses the new recommended allocation.

Note: If you install PAS for Windows, provision your **Master Compilation Job** with at least 100 GB of disk space.

3. (Optional) Enter your AWS target group name in the **Load Balancers** column for each job. Prepend the name with `alb:`. For example, enter `alb:target-group-name`. To create an Application Load Balancer (ALB) and target group, follow the procedures in [Getting Started with Application Load Balancers](#) in the AWS documentation. Then navigate to **Target Groups** in the **EC2 Dashboard** menu to find your target group Name.

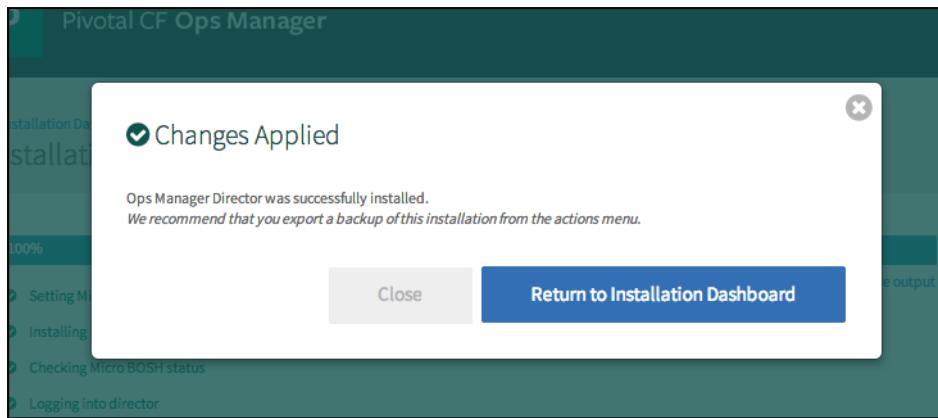
Note: To configure an ALB, you must have the following AWS IAM permissions.

```
"elasticloadbalancing:DescribeLoadBalancers",
"elasticloadbalancing:DeregisterInstancesFromLoadBalancer",
"elasticloadbalancing:RegisterInstancesWithLoadBalancer",
"elasticloadbalancing:DescribeTargetGroups",
"elasticloadbalancing:RegisterTargets"
```

4. Click **Save**.

Step 11: Complete the BOSH Director Installation

1. Return to the **Installation Dashboard**.
2. Click **Review Pending Changes**, then **Apply Changes**.
3. BOSH Director begins to install. The **Changes Applied** window displays when the installation process successfully completes.



Next Steps

After completing the procedures in this topic, you must configure a runtime for PCF. You can install PCF on AWS with the Pivotal Application Service (PAS) or Pivotal Container Service (PKS) runtime.

To configure a runtime, do one of the following:

- Configure PAS. See [Configuring PAS](#).
- Configure PKS. See [Installing PKS on AWS](#).

Installing PCF on AWS using Terraform

This topic explains how to install PCF on AWS using Terraform.

Overview

You can install PCF on AWS with either the Pivotal Application Service (PAS) or Pivotal Container Service (PKS) runtime.

To install PCF on AWS using Terraform, do one of the following:

- Install with PAS. See [Install with PAS Using Terraform](#).
- Install with PKS. See [Install with PKS Using Terraform](#).

Install with PAS Using Terraform

To install with PAS using Terraform, do the following:

1. Deploy Ops Manager. See [Deploying Ops Manager on AWS Using Terraform](#).
2. Configure BOSH Director. See [Configuring BOSH Director on AWS Using Terraform](#).
3. Configure PAS. See [Configuring PAS](#).

Install with PKS Using Terraform

To install with PKS using Terraform, do the following:

1. Deploy Ops Manager. See [Deploying Ops Manager on AWS Using Terraform](#).
2. Configure BOSH Director. See [Configuring BOSH Director on AWS Using Terraform](#).
3. Configure PKS. See [Installing PKS on AWS](#).

Deploying Ops Manager on AWS Using Terraform

This guide describes the preparation steps required to deploy Ops Manager on Amazon Web Services (AWS) using Terraform templates.

The Terraform template for Ops Manager on AWS describes a set of AWS resources and properties. For more information about how Terraform creates resources in AWS, see the [AWS Provider ↗](#) topic on the Terraform site.

If you are deploying Pivotal Application Service (PAS), you may also find it helpful to review different deployment options in the [Reference Architecture for Pivotal Cloud Foundry on AWS ↗](#).

Prerequisites

Before you deploy Ops Manager on AWS, review the following:

- If you intend to install PAS, see [PCF on AWS Requirements](#).
- If you intend to install Pivotal Container Service (PKS), see [AWS Prerequisites and Resource Requirements ↗](#).

In addition to reviewing the prerequisites for your runtime, ensure you have the following:

- The [Terraform CLI ↗](#)
- In your AWS account, ensure you have an IAM user with the following permissions:
 - AmazonEC2FullAccess
 - AmazonRDSFullAccess
 - AmazonRoute53FullAccess
 - AmazonS3FullAccess
 - AmazonVPCFullAccess
 - IAMFullAccess
 - AWSKeyManagementServicePowerUser

Step 1: Download Templates and Edit Variables File

Before you can run Terraform commands to provision infrastructure resources, you must download the AWS Terraform templates and create a Terraform template variables file as described below:

1. On [Pivotal Network ↗](#), navigate to the [Pivotal Application Service \(formerly Elastic Runtime\)](#) release.
2. Download the AWS Terraform templates ZIP file.
3. Extract the contents of the ZIP file.
4. Move the extracted folder to the `workspace` directory on your local machine.
5. On the command line, navigate to the directory. For example:

```
$ cd ~/workspace/pivotal-cf-terraforming-aws
```

6. Navigate to the `terraforming-pas` or `terraforming-pks` directory that contains the Terraform files for your runtime.
7. In the runtime directory, create a text file named `terraform.tfvars`.
8. Open the `terraform.tfvars` file and add the following:

```

env_name      = "YOUR-ENVIRONMENT-NAME"
access_key     = "YOUR-ACCESS-KEY"
secret_key     = "YOUR-SECRET-KEY"
region         = "YOUR-AWS-REGION"
availability_zones = ["YOUR-AZ-1", "YOUR-AZ-2", "YOUR-AZ-3"]
ops_manager_ami  = "YOUR-OPS-MAN-IMAGE-AMI"
dns_suffix     = "YOUR-DNS-SUFFIX"

ssl_cert = <<SSL_CERT
-----BEGIN CERTIFICATE-----
YOUR-CERTIFICATE
-----END CERTIFICATE-----
SSL_CERT

ssl_private_key = <<SSL_KEY
-----BEGIN EXAMPLE RSA PRIVATE KEY-----
YOUR-PRIVATE-KEY
-----END EXAMPLE RSA PRIVATE KEY-----
SSL_KEY

```

9. Edit the values in the file according to the table below.

Value to replace	Guidance
YOUR-ENVIRONMENT-NAME	<p>Enter a name to use to identify resources in AWS. Terraform prepends the names of the resources it creates with this environment name.</p> <p>Note: You can only enter lowercase alphanumeric characters and hyphens. Examples include <code>pcf</code>, <code>pas</code>, and <code>pks</code>.</p>
YOUR-ACCESS-KEY	Enter your AWS Access Key ID of the AWS account in which you want Terraform to create resources.
YOUR-SECRET-KEY	Enter your AWS Secret Access Key of the AWS account in which you want Terraform to create resources.
YOUR-AWS-REGION	Enter the name of the AWS region in which you want Terraform to create resources. Example: <code>us-central1</code> .
YOUR-AZ-1 YOUR-AZ-2 YOUR-AZ-3	Enter three availability zones from your region. Example: <code>us-central-1a</code> , <code>us-central-1b</code> , <code>us-central-1c</code> .
YOUR-OPS-MAN-IMAGE-AMI	<p>Enter the source code for the Ops Manager Amazon Machine Image (AMI) you want to boot. You can find this code in the PDF included with the Ops Manager release on Pivotal Network.</p> <p>If you want to encrypt your Ops Manager VM, create an encrypted AMI copy from the AWS EC2 dashboard and enter the source code for the copied Ops Manager image instead. For more information about copying an AMI, see step 9 of Launch an Ops Manager AMI in the manual AWS configuration topic.</p> <p>To prevent the creation of an Ops Manager VM, set this value to an empty string (<code>" "</code>). When using Platform Automation, you must disable the creation of the Ops Manager VM from Terraform. For more information, see Platform Automation.</p>
YOUR-DNS-SUFFIX	Enter a domain name to use as part of the system domain for your deployment. Terraform creates DNS records in AWS using <code>YOUR-ENVIRONMENT-NAME</code> and <code>YOUR-DNS-SUFFIX</code> . For example, if you enter <code>example.com</code> for your DNS suffix and have <code>pcf</code> as your environment name, Terraform will create DNS records at <code>pcf.example.com</code> .
YOUR-CERTIFICATE	<p>Enter a certificate to use for HTTP load balancing. For production environments, use a certificate from a Certificate Authority (CA). For test environments, you can use a self-signed certificate.</p> <p>Your certificate must specify your system domain as the common name. Your system domain is <code>YOUR-ENVIRONMENT-NAME.YOUR-DNS-SUFFIX</code>.</p> <p>It also must include the following subdomains: <code>*.sys.YOUR-SYSTEM-DOMAIN</code>, <code>*.login.sys.YOUR-SYSTEM-DOMAIN</code>, <code>*.uaa.sys.YOUR-SYSTEM-DOMAIN</code>, <code>*.apps.YOUR-SYSTEM-DOMAIN</code>.</p>
YOUR-PRIVATE-KEY	Enter a private key for the certificate you entered.

Step 2: Add Optional Variables

Complete this step if you want to do any of the following in PAS:

- Use an AWS Relational Database Service (RDS) for your deployment. For more information, see [Getting started with Amazon RDS](#) in Amazon RDS resources.
- Deploy the Isolation Segment tile.

In your `terraform.tfvars` file, specify the appropriate variables from the sections below.

 **Note:** You can see the configurable options by opening the `variables.tf` file and looking for variables with default values.

Isolation Segments

If you plan to deploy the Isolation Segment tile, add the following variables to your `terraform.tfvars` file, replacing `YOUR-CERTIFICATE` and `YOUR-PRIVATE-KEY` with a certificate and private key. This causes terraform to create an additional HTTP load balancer across three availability zones to use for the Isolation Segment tile.

```
create_isoseg_resources = 1

iso_seg_ssl_cert = <<ISO_SEG_SSL_CERT
-----BEGIN CERTIFICATE-----
YOUR-CERTIFICATE
-----END CERTIFICATE-----
ISO_SEG_SSL_CERT

iso_seg_ssl_cert_private_key = <<ISO_SEG_SSL_KEY
-----BEGIN EXAMPLE RSA PRIVATE KEY-----
YOUR-PRIVATE-KEY
-----END EXAMPLE RSA PRIVATE KEY-----
ISO_SEG_SSL_KEY
```

RDS

1. If you want to use an RDS for Ops Manager and PAS, add the following to your `terraform.tfvars` file:

```
rds_instance_count = 1
```

2. If you want to specify a username for RDS authentication, add the following variable to your `terraform.tfvars` file.

```
rds_db_username = "your-database-name"
```

Step 3: Create AWS Resources with Terraform

Follow these steps to use the Terraform CLI to create resources on AWS:

1. From the directory that contains the Terraform files, run `terraform init` to initialize the directory based on the information you specified in the `terraform.tfvars` file.

```
terraform init
```

2. Run `terraform plan -out=plan` to create the execution plan for Terraform.

```
terraform plan -out=plan
```

3. Run `terraform apply plan` to execute the plan from the previous step. It may take several minutes for Terraform to create all the resources in AWS.

```
terraform apply plan
```

Step 4: Create DNS Record

1. In a browser, navigate to the DNS provider for the DNS suffix you entered in your `terraform.tfvars` file.
2. Create a new NS record for your system domain. Your system domain is `YOUR-ENVIRONMENT-NAME.YOUR-DNS-SUFFIX`.
3. In this record, enter the name servers included in `env_dns_zone_name_servers` from your Terraform output.

What to Do Next

Proceed to the next step in the deployment, [Configuring BOSH Director on AWS Using Terraform](#).

Configuring BOSH Director on AWS Using Terraform

This topic describes how to configure the BOSH Director tile in Ops Manager on Amazon Web Services (AWS) after [Deploying Ops Manager on AWS Using Terraform](#).

 **Note:** You can also perform the procedures in this topic using the Ops Manager API. For more information, see the [Using the Ops Manager API](#) topic.

Prerequisites

To complete the procedures in this topic, you must have access to the output from when you ran `terraform apply` to create resources for this deployment. You can view this output at any time by running `terraform output`. You use the values in your Terraform output to configure the BOSH Director tile.

Step 1: Access Ops Manager

1. In a web browser, navigate to the fully qualified domain name (FQDN) of Ops Manager. Use the `ops_manager_dns` value from running `terraform output`.
2. When Ops Manager starts for the first time, you must choose one of the following:
 - o **Internal Authentication:** If you use Internal Authentication, Ops Manager maintains your user database.
 - o **SAML Identity Provider:** If you use a SAML Identity Provider (IdP), an external identity server maintains your user database.
 - o **LDAP Server:** If you use a LDAP Server, an external identity server maintains your user database.

Internal Authentication

1. When redirected to the **Internal Authentication** page, do the following:
 - o Enter a **Username**, **Password**, and **Password confirmation** to create an Admin user.
 - o Enter a **Decryption passphrase** and the **Decryption passphrase confirmation**. This passphrase encrypts the Ops Manager datastore, and is not recoverable.
 - o If you are using an **HTTP proxy** or **HTTPS proxy**, follow the instructions in the [Configuring Proxy Settings for the BOSH CPI](#) topic.
 - o Read the **End User License Agreement**, and select the checkbox to accept the terms.
 - o Click **Setup Authentication**.
2. Log in to Ops Manager with the Admin username and password you created in the previous step.

SAML Identity Provider

1. Log in to your IdP console and download the IdP metadata XML. Optionally, if your IdP supports metadata URL, you can copy the metadata URL instead of the XML.
2. Copy the IdP metadata XML or URL to the Ops Manager **SAML Identity Provider** login page.

 **Note:** The same IdP metadata URL or XML is applied for the BOSH Director. If you use a separate IdP for BOSH, copy the metadata XML or URL from that IdP and enter it into the BOSH IdP Metadata text box in the Ops Manager login page.

3. Enter values for the fields listed below. Failure to provide values in these fields results in a `500` error.
 - o **SAML admin group:** Enter the name of the SAML group that contains all Ops Manager administrators. This field is case-sensitive.
 - o **SAML groups attribute:** Enter the groups attribute tag name with which you configured the SAML server. This field is case-sensitive.
4. Enter your **Decryption passphrase**. Read the **End User License Agreement**, and select the checkbox to accept the terms.
5. Your Ops Manager login page appears. Enter your username and password. Click **Login**.
6. Download your SAML Service Provider metadata (SAML Relying Party metadata) by navigating to the following URLs:
 - o **6a. Ops Manager SAML service provider metadata:** `https://OPS-MAN-FQDN:443/uaa/saml/metadata`

- o 6b. BOSH Director SAML service provider metadata: `https://BOSH-IP-ADDRESS:8443/saml/metadata`

Note: To retrieve your `BOSH-IP-ADDRESS`, navigate to the **Status** tab in the **BOSH Director** tile. Record the **BOSH Director** IP address.

7. Configure your IdP with your SAML Service Provider metadata. Import the Ops Manager SAML provider metadata from Step 6a above to your IdP. If your IdP does not support importing, provide the values below.

- o Single sign on URL: `https://OPS-MAN-FQDN:443/uaa/saml/SSO/alias/OPS-MAN-FQDN`
- o Audience URI (SP Entity ID): `https://OP-MAN-FQDN:443/uaa`
- o Name ID: Email Address
- o SAML authentication requests are always signed

8. Import the BOSH Director SAML provider metadata from Step 6b to your IdP. If the IdP does not support an import, provide the values below.

- o Single sign on URL: `https://BOSH-IP:8443/saml/SSO/alias/BOSH-IP`
- o Audience URI (SP Entity ID): `https://BOSH-IP:8443`
- o Name ID: Email Address
- o SAML authentication requests are always signed

9. Return to the **BOSH Director** tile, and continue with the configuration steps below.

LDAP Server

- For **Server URL**, enter the URL that points to your LDAP server. With multiple LDAP servers, separate their URLs with spaces. Each URL must include one of the following protocols:
 - o `ldap://`: This specifies that the LDAP server uses an unencrypted connection.
 - o `ldaps://`: This specifies that the LDAP server uses SSL for an encrypted connection and requires that the LDAP server holds a trusted certificate or that you import a trusted certificate to the JVM truststore.
- For **LDAP Username** and **LDAP Password**, enter the LDAP Distinguished Name (DN) and the password for binding to the LDAP Server. Example DN: `cn=administrator,ou=Users,dc=example,dc=com`

Note: Pivotal recommends that you provide LDAP credentials that grant read-only permissions on the LDAP Search Base and the LDAP Group Search Base. In addition to this, if the bind user belongs to a different search base, you must use the full DN.

warning: Pivotal recommends against reusing LDAP service accounts across environments. LDAP service accounts should not be subject to manual lockouts, such as lockouts that result from users utilizing the same account. Also, LDAP service accounts should not be subject to automated deletions, since disruption to these service accounts could prevent user logins.

- For **User Search Base**, enter the location in the LDAP directory tree from which any LDAP User search begins. The typical LDAP Search Base matches your domain name.

For example, a domain named “cloud.example.com” typically uses the following LDAP User Search Base: `ou=Users,dc=example,dc=com`

- For **User Search Filter**, enter a string that defines LDAP User search criteria. These search criteria allow LDAP to perform more effective and efficient searches. For example, the standard LDAP search filter `cn=Smith` returns all objects with a common name equal to `Smith`.

In the LDAP search filter string that you use to configure your runtime, use `{0}` instead of the username. For example, use `cn={0}` to return all LDAP objects with the same common name as the username.

In addition to `cn`, other attributes commonly searched for and returned are `mail`, `uid` and, in the case of Active Directory, `sAMAccountName`.

Note: For instructions for testing and troubleshooting your LDAP search filters, see [Configuring LDAP Integration with Pivotal Cloud Foundry](#) in the Pivotal Support Knowledge Base.

- For **Group Search Base**, enter the location in the LDAP directory tree from which the LDAP Group search begins.

For example, a domain named “cloud.example.com” typically uses the following LDAP Group Search Base: `ou=Groups,dc=example,dc=com`

- For **Group Search Filter**, enter a string that defines LDAP Group search criteria. The standard value is `member={0}`.
- For **Email Attribute**, enter the attribute name in your LDAP directory that corresponds to the email address in each user record, for example `mail`.
- For **LDAP RBAC Admin Group Name**, enter the DN of the LDAP group you want to have admin permissions in Ops Manager.
- From the dropdown, select how the UAA handles LDAP server referrals out to other external user stores. The UAA can:

- Automatically follow any referrals.
- Ignore referrals and return partial result.
- Throw exception for each referral and abort.
- For **Server SSL Cert**, paste in the root certificate from your CA certificate or your self-signed certificate.
- Enter a **Decryption passphrase** and the **Decryption passphrase confirmation**. This passphrase encrypts the Ops Manager datastore, and is not recoverable.
- If you are using an **HTTP proxy** or **HTTPS proxy**, follow the instructions in [Configuring Proxy Settings for the BOSH CPI](#).
- Read the **End User License Agreement**, and select the checkbox to accept the terms.
- Select **Provision an admin client in the BOSH UAA**. You can use this to enable BOSH automation with scripts and tooling. For more information, see [Provision Admin Client](#) in *Creating UAA Clients for BOSH Director*.
- Click **Setup Authentication**.
- Return to the **BOSH Director** tile, and continue with the configuration steps below.

Step 2: AWS Config Page

1. Click the **BOSH Director** tile.
2. Select **AWS Config** to open the **AWS Management Console Config** page.

AWS Management Console Config

Use AWS Keys

Access Key ID*

AWS Secret Key*
 The AWS Secret Key

Use AWS Instance Profile

AWS IAM Instance Profile*

Security Group ID*

Key Pair Name*

SSH Private Key*

Region*

Encrypt Linux EBS Volumes

Custom Encryption Key

Save

3. Select **Use AWS Keys** or **Use AWS Instance Profile**.

- o If you choose to use AWS keys, complete the following fields:
 - **Access Key ID:** Enter the value of `ops_manager_iam_user_access_key` from the Terraform output.
 - **AWS Secret Key:** Enter the value of `ops_manager_iam_user_secret_key` from the Terraform output.
- o If you choose to use an AWS instance profile, enter the name of your AWS Identity and Access Management (IAM) profile or enter the value of `ops_manager_iam_instance_profile_name` from the Terraform output.

4. Complete the remainder of the **AWS Management Console Config** page with the following information.

- o **Security Group ID:** Enter the value of `vms_security_group_id` from the Terraform output.
- o **Key Pair Name:** Enter the value of `ops_manager_ssh_public_key_name` from the Terraform output.
- o **SSH Private Key:** Run `terraform output` to view the value of `ops_manager_ssh_private_key` and enter it into this field. `ops_manager_ssh_private_key` is a `sensitive` value and does not display when you run `terraform apply`.
- o **Region:** Input the AWS region where you deployed Ops Manager.
- o **Encrypt Linux EBS Volumes:** Select this checkbox to enable full encryption on persistent disks of all BOSH-deployed VMs, except for the Ops

Manager VM and BOSH Director VM. See the [Configuring Amazon EBS Encryption](#) topic for details about using Elastic Block Store (EBS) encryption.

Note: Enabling EBS encryption only encrypts Linux VMs. The Windows VMs deployed with Pivotal Application Service for Windows (PASW) are not encrypted.

- **Custom Encryption Key (Optional)** Once you enable EBS encryption, you may want to specify a custom Key Management Service (KMS) encryption key. If you don't enter a value, your custom encryption key will default to the account key. For more information, see [Configuring Amazon EBS Encryption](#).

5. Click **Save**.

Step 3: Director Config Page

1. In Ops Manager, select **Director Config**.

NTP Servers (comma delimited)*

Bosh HM Forwarder IP Address

Enable VM Resurrector Plugin

Enable Post Deploy Scripts

Recreate All VMs
This will force BOSH to recreate all VMs on the next deploy. Persistent disk will be preserved.

Recreate All Persistent Disks
Check this box to recreate all Persistent Disks for the Director and all other Tiles.

Enable bosh deploy retries
This will attempt to re-deploy a failed deployment up to 5 times.

2. Enter at least two of the following NTP servers in the **NTP Servers (comma delimited)** field, separated by a comma:

0.amazon.pool.ntp.org,1.amazon.pool.ntp.org,2.amazon.pool.ntp.org,3.amazon.pool.ntp.org

Note: The NTP server configuration only updates after VM recreation. Ensure that you select the **Recreate all VMs** checkbox if you modify the value of this field.

3. Leave the **Bosh HM Forwarder IP Address** field blank.

Note: Starting in PAS v2.0, BOSH-reported component metrics are available in the Loggregator Firehose by default. If you continue to use the BOSH HM Forwarder to consume these component metrics, you may receive duplicate data. To prevent this, leave the **Bosh HM Forwarder IP Address** field blank.

4. Select the **Enable VM Resurrector Plugin** checkbox to enable the BOSH Resurrector functionality and increase your runtime availability.

5. Select **Enable Post Deploy Scripts** to run a post-deploy script after deployment. This script allows the job to execute additional commands against a deployment.

Note: If you intend to install PKS, you must enable post-deploy scripts.

6. Select **Recreate all VMs** to force BOSH to recreate all VMs on the next deploy. This process does not destroy any persistent disk data.

7. Select **Recreate All Persistent Disks** to force BOSH to migrate and recreate persistent disks for the BOSH Director and all tiles. This process does not destroy any persistent disk data.
8. Select **Enable bash deploy retries** to instruct Ops Manager to retry failed BOSH operations up to five times.
9. Select **Skip Director Drain Lifecycle** to prevent drain scripts from running when the BOSH Director is recreated.
10. Select **Store BOSH Job Credentials on tmpfs (beta)** to store credentials for BOSH jobs on temporary file storage (tmpfs) memory, rather than on disk. You must recreate all VMs for this setting to take effect.
11. Select **Keep Unreachable Director VMs** if you want to preserve BOSH Director VMs after a failed deployment for troubleshooting purposes.
12. Select **HM Pager Duty Plugin** to enable Health Monitor integration with PagerDuty.

HM Pager Duty Plugin

Service Key*

YOUR-PAGERDUTY-SERVICE-KEY

HTTP Proxy

YOUR-HTTP-PROXY

- **Service Key:** Enter your API service key from PagerDuty.
- **HTTP Proxy:** Enter an HTTP proxy for use with PagerDuty.

13. Select **HM Email Plugin** to enable Health Monitor integration with email.

HM Email Plugin

Host*

smtp.example.com

Port*

25

Domain*

cloudfoundry.example.com

From*

user2@example.com

Recipients*

user@example.com, user1@example.com

Username

user

Password

Enable TLS

- **Host:** Enter your email hostname.
- **Port:** Enter your email port number.
- **Domain:** Enter your domain.
- **From:** Enter the address for the sender.
- **Recipients:** Enter comma-separated addresses of intended recipients.
- **Username:** Enter the username for your email server.
- **Password:** Enter the password for your email server.

- **Enable TLS:** Select this checkbox to enable Transport Layer Security to the email host.

14. For **CredHub Encryption Provider**, you can choose whether BOSH CredHub stores its encryption key internally on the BOSH Director and CredHub VM, or in an external hardware security module (HSM). The HSM option is more secure.

Before configuring an HSM encryption provider in the **Director Config** pane, you must follow the procedures and collect information described in [Preparing CredHub HSMs for Configuration](#).

 **Note:** After you deploy Ops Manager with an HSM encryption provider, you cannot change BOSH CredHub to store encryption keys internally.

CredHub Encryption Provider

Internal
 Luna HSM

Encryption Key Name*

Provider Partition*

Provider Partition Password*

Provider Client Certificate*

Provider Client Certificate Private Key*

HSM Host Address*

HSM Port Address*

Partition Serial Number*

HSM Certificate*

- **Internal:** Select this option for internal CredHub key storage. This option is selected by default and requires no additional configuration.
- **Luna HSM:** Select this option to use a SafeNet Luna HSM as your permanent CredHub encryption provider, and fill in the following fields:

1. **Encryption Key Name:** Any name to identify the key that the HSM uses to encrypt and decrypt the CredHub data. Changing this key name after you deploy Ops Manager can cause service downtime.
2. **Provider Partition:** The partition that stores your encryption key. Changing this partition after you deploy Ops Manager could cause

service downtime. For this value and the ones below, use values gathered in [Preparing CredHub HSMs for Configuration](#).

3. **Provider Partition Password**
4. **Provider Client Certificate:** The certificate that validates the identity of the HSM when CredHub connects as a client.
5. **Provider Client Certificate Private Key**
6. **HSM Host Address**
7. **HSM Port Address:** If you do not know your port address, enter **1792**.
8. **Partition Serial Number**
9. **HSM Certificate:** The certificate that the HSM presents to CredHub to establish a two-way mTLS connection.

15. Select a **Blobstore Location** to either configure the blobstore as an internal server or an external endpoint. Because the internal server is unscalable and less secure, Pivotal recommends that you configure an external blobstore.

 **Note:** After you deploy Ops Manager, you cannot change the blobstore location.

- **Internal:** Select this option to use an internal blobstore. Ops Manager creates a new VM for blob storage. No additional configuration is required.
 - **Enable TLS:** Select this checkbox to enable TLS to the blobstore.

 **Note:** If you are using PASW 2016, make sure you have downloaded [Windows stemcell v1709.10 or higher](#) before enabling TLS.

- **S3 Compatible Blobstore:** Select this option to use an external S3-compatible endpoint. When you have created an S3 bucket, complete the

Blobstore Location

Internal

Enable TLS

S3 Compatible Blobstore

S3 Endpoint*

Bucket Name*

Access Key*

Secret Key*

S3 Signature Version

V2 Signature

V4 Signature

Region*

S3 Backup Strategy

No backups

Use a versioned bucket

Copy into an additional bucket

Backup Bucket Region*

Backup Bucket Name*

following steps:

Follow the procedures in [Sign up for Amazon S3](#) and [Creating a Bucket](#) in the AWS documentation. When you have created an S3 bucket, complete the following steps:

1. **S3 Endpoint:** Navigate to the [Regions and Endpoints](#) topic in the AWS documentation.
 - a. Locate the endpoint for your region in the **Amazon Simple Storage Service (S3)** table and construct a URL using your region's endpoint. For example, if you are using the `us-west-2` region, the URL you create would be `https://s3-us-west-2.amazonaws.com`. Enter this URL into the **S3 Endpoint** field.
 - b. On a command line, run `ssh ubuntu@OPS-MANAGER-FQDN` to SSH into the Ops Manager VM. Replace `OPS-MANAGER-FQDN` with the fully qualified domain name of Ops Manager.
 - c. Copy the custom public CA certificate you used to sign the S3 endpoint into `/etc/ssl/certs` on the Ops Manager VM.
 - d. On the Ops Manager VM, run `sudo update-ca-certificates -f -v` to import the custom CA certificate into the Ops Manager VM truststore.



Note: You must also add this custom CA certificate into the **Trusted Certificates** field in the **Security** page. See [Security Page](#) for instructions.

2. **Bucket Name:** Enter the name of the S3 bucket.
3. **Access Key** and **Secret Key:** Enter the keys you generated when creating your S3 bucket.
4. Select **V2 Signature** or **V4 Signature**. If you select **V4 Signature**, enter your **Region**.



Note: AWS recommends using Signature Version 4. For more information about AWS S3 Signatures, see [Authenticating Requests](#)

[🔗](#) in the AWS documentation.

5. **S3 Backup Strategy:** Select one of the following to configure whether and how to back up the BOSH Director's S3 blobstore:

- **No backups**
- **Use a versioned bucket:** This option uses blobstore bucket versioning to save backups. It requires an S3 blobstore that supports versioning.
- **Copy into an additional bucket:** This option saves blobstore backups to a separate bucket, which is useful with blobstores that do not support versioning. The option requires you to fill in **Backup Bucket Region** and **Backup Bucket Name**.

- **GCS Blobstore:** Select this option to use an external Google Cloud Storage (GCS) endpoint.

The screenshot shows a configuration form for a GCS Blobstore. At the top left is a blue circular icon with a white dot. To its right, the text "GCS Blobstore" is displayed. Below this are three input fields: "Bucket Name*" with a red asterisk, "Storage Class*" with a dropdown menu set to "Regional", and "Service Account Key*" with a large text input area.

Follow the procedures in [Creating Storage Buckets](#) in the GCS documentation to create a GCS bucket. To create a GCS bucket, you must have a GCS account. When you have created a GCS bucket, complete the following steps:

1. **Bucket Name:** Enter the name of your GCS bucket.
 2. **Storage Class:** Select the storage class for your GCS bucket. See [Storage Classes](#) in the GCP documentation for more information.
 3. **Service Account Key:** Follow the steps in the [Set up IAM Service Accounts](#) section of *Preparing to Deploy Ops Manager on GCP Manually* to download a JSON file with a private key. Enter the contents of the JSON file into the field.
16. For **Database Location**, if you choose to configure an external MySQL database with Amazon Relational Database Service (RDS) or another service, select **External MySQL Database** and complete the fields below. Otherwise, select **Internal**. For more information about creating a RDS MySQL instance, see [Creating a MySQL DB Instance](#) in the AWS documentation.
- **Host:** Enter the value of your host.
 - **Port:** Enter your port number. For example, `3306`.
 - **Username:** Enter your username.
 - **Password:** Enter your password.
 - **Database:** Enter your database name.
- In addition, if you selected the **Enable TLS for Director Database** checkbox, you can fill out the following optional fields:
- **Enable TLS:** Selecting this checkbox enables TLS communication between the BOSH Director and the database.
 - **TLS CA:** Enter the Certificate Authority for the TLS Certificate.
 - **TLS Certificate:** Enter the client certificate for mutual TLS connections to the database.
 - **TLS Private Key:** Enter the client private key for mutual TLS connections to the database.
 - **Advanced DB Connection Options:** If you would like to provide additional options for the database, use this field to provide a JSON-formatted options string.
17. (Optional) Modify the **Director Workers** value, which sets the number of workers available to execute Director tasks. This field defaults to `5`.
18. (Optional) **Max Threads** sets the maximum number of threads that the BOSH Director can run simultaneously. Pivotal recommends that you leave the field blank to use the default value, unless doing so results in rate limiting or errors on your IaaS.
19. (Optional) To add a custom URL for your BOSH Director, enter a valid hostname in **Director Hostname**. You can also use this field to configure a load balancer in front of your BOSH Director. For more information, see [How to Set Up a Load Balancer in Front of Operations Manager Director](#) in the Pivotal Support Knowledge Base.

Director Workers

Max Threads

Director Hostname

20. (Optional) To set a custom banner that users see when logging in to the Director using SSH, enter text in the **Custom SSH Banner** field.
21. (Optional) Enter your comma-separated custom **Identification Tags**. For example, `iaas:foundation1, hello:world`. You can use the tags to identify your foundation when viewing VMs or disks from your IaaS.
22. Click **Save**.

 **Note:** For more information about AWS S3 Signatures, see the AWS [Authenticating Requests](#) documentation.

Step 4: Create Availability Zones Page

1. Select **Create Availability Zones**.

Create Availability Zones

Availability Zones	
▼ us-west-1c Amazon Availability Zone* <input type="text" value="us-west-1c"/> <small>The Amazon Availability Zone name (ex: 'us-east-1b')</small>	Add Remove
Save	

2. To create three Availability Zones for your apps to use, do the following:

- a. Click **Add** three times.
- b. For **Amazon Availability Zone**, enter the values corresponding to the key `infrastructure_subnet_availability_zones` from the Terraform output.
- c. Click **Save**.

Step 5: Create Networks Page

Create Networks

Warning: Pivotal recommends keeping the IP settings throughout the life of your installation. Ops Manager may prevent you from changing them in the future.
Contact Pivotal support for help completing such a change.

Verification Settings

Enable ICMP checks

Networks

One or many IP ranges upon which your products will be deployed

Subnets

Name*

VPC Subnet ID*

CIDR*

Reserved IP Ranges

DNS*

Gateway*

Availability Zones*

us-west-2a
 us-west-2b
 us-west-2c

1. Select **Create Networks**.

2. (Optional) Select **Enable ICMP checks** to enable ICMP on your networks. Ops Manager uses ICMP checks to confirm that components within your network are reachable.

3. To add the network configuration you created for your VPC, do the following:

- Click **Add Network**.
- For **Name**, enter .
- Create a subnet for each availability zone by clicking **Add Subnet**. Refer to the table below for the information required to create all three subnets:

First Subnet	VPC Subnet ID	The first value of <code>infrastructure_subnet_ids</code> from the Terraform output.
	CIDR	<input type="text" value="10.0.16.0/28"/>
	Reserved IP Ranges	<input type="text" value="10.0.16.0-10.0.16.4"/>
	DNS	<input type="text" value="10.0.0.2"/> *
	Gateway	<input type="text" value="10.0.16.1"/>
	Availability Zones	The first value of <code>infrastructure_subnet_availability_zones</code> from the Terraform output.
	VPC Subnet ID	The second value of <code>infrastructure_subnet_ids</code> from the Terraform output.

Second Subnet	CIDR	10.0.16.16/28
	Reserved IP Ranges	10.0.16.16-10.0.16.20
	DNS	10.0.0.2 *
	Gateway	10.0.16.17
	Availability Zones	The second value of <code>infrastructure_subnet_availability_zones</code> from the Terraform output.
Third Subnet	VPC Subnet ID	The third value of <code>infrastructure_subnet_ids</code> from the Terraform output.
	CIDR	10.0.16.32/28
	Reserved IP Ranges	10.0.16.32-10.0.16.36
	DNS	10.0.0.2 *
	Gateway	10.0.16.33
	Availability Zones	The third value of <code>infrastructure_subnet_availability_zones</code> from the Terraform output.

* If you set a VPC CIDR other than recommended, enter the second IP in your VPC CIDR. For example, for a `10.0.0.0/24` VPC CIDR, enter `10.0.0.2` in each subnet.

d. Click **Add Network**.

e. For **Name**, enter the name of your runtime. For example, `pas` or `pks`.

f. Create a subnet for each availability zone by clicking **Add Subnet**. Refer to the table below for the information required to create all three subnets:

First Subnet	VPC Subnet ID	The first value of <code>RUNTIME_subnet_ids</code> from the Terraform output, where <code>RUNTIME</code> is either <code>pas</code> or <code>pks</code> .
	CIDR	10.0.4.0/24
	Reserved IP Ranges	10.0.4.0-10.0.4.4
	DNS	10.0.0.2 *
	Gateway	10.0.4.1
	Availability Zones	The first value of <code>RUNTIME_subnet_availability_zones</code> from the Terraform output.
Second Subnet	VPC Subnet ID	The second value of <code>RUNTIME_subnet_ids</code> from the Terraform output.
	CIDR	10.0.5.0/24
	Reserved IP Ranges	10.0.5.0-10.0.5.4
	DNS	10.0.0.2 *
	Gateway	10.0.5.1
	Availability Zones	The second value of <code>RUNTIME_subnet_availability_zones</code> from the Terraform output.
Third Subnet	VPC Subnet ID	The third value of <code>RUNTIME_subnet_ids</code> from the Terraform output.
	CIDR	10.0.6.0/24
	Reserved IP Ranges	10.0.6.0-10.0.6.4
	DNS	10.0.0.2 *
	Gateway	10.0.6.1
	Availability Zones	The third value of <code>RUNTIME_subnet_availability_zones</code> from the Terraform output.

* If you set a VPC CIDR other than recommended, enter the second IP in your VPC CIDR. For example, for a `10.0.0.0/24` VPC CIDR, enter `10.0.0.2` in each subnet.

g. Click **Add Network**.

h. For **Name**, enter `services`.

i. Create a subnet for each availability zone by clicking **Add Subnet**. Refer to the table below for the information required to create all three subnets:

First Subnet	VPC Subnet ID	The first value of <code>services_subnet_ids</code> from the Terraform output.
	CIDR	10.0.8.0/24
	Reserved IP Ranges	10.0.8.0-10.0.8.3
	DNS	10.0.0.2 *

	Gateway	10.0.8.1	
	Availability Zones	The first value of <code>services_subnet_availability_zones</code> from the Terraform output.	
Second Subnet	VPC Subnet ID	The second value of <code>services_subnet_ids</code> from the Terraform output.	
	CIDR	10.0.9.0/24	
	Reserved IP Ranges	10.0.9.0-10.0.9.3	
	DNS	10.0.0.2*	
	Gateway	10.0.9.1	
	Availability Zones	The second value of <code>services_subnet_availability_zones</code> from the Terraform output.	
Third Subnet	VPC Subnet ID	The third value of <code>services_subnet_ids</code> from the Terraform output.	
	CIDR	10.0.10.0/24	
	Reserved IP Ranges	10.0.10.0-10.0.10.3	
	DNS	10.0.0.2*	
	Gateway	10.0.10.1	
	Availability Zones	The third value of <code>services_subnet_availability_zones</code> from the Terraform output.	

* If you set a VPC CIDR other than recommended, enter the second IP in your VPC CIDR. For example, for a 10.0.0.0/24 VPC CIDR, enter 10.0.0.2 in each subnet.

- Click **Save**.

Note: If you are deploying PKS with a PKS workload load balancer, you must tag each AWS subnet with your PKS Kubernetes cluster unique identifier before you create the load balancer. For more information about tagging subnets with a PKS cluster unique identifier, see [AWS Prerequisites](#).

Note: After you deploy Ops Manager, you add subnets with overlapping Availability Zones to expand your network. For more information about configuring additional subnets, see [Expanding Your Network with Additional Subnets](#).

Step 6: Assign AZs and Networks Page

- Select **Assign AZs and Networks**.

Assign AZs and Networks

The BOSH Director is a single instance.

Choose the availability zone in which to place that instance. It is highly recommended that you backup this VM on a regular basis to preserve settings.

Singleton Availability Zone

AZ-MGMT

Network

infrastructure

Save

- Use the dropdown to select a **Singleton Availability Zone**. The BOSH Director installs in this availability zone (AZ).

- Use the dropdown to select the `infrastructure` network for your BOSH Director.

- Click **Save**.

Step 7: Security Page

1. Select **Security**.

Security

Trusted Certificates

These certificates enable BOSH-deployed components to trust a custom root certificate.

Include OpsManager Root CA in Trusted Certs

Checking this box will place the OpsManager Root CA into the trusted certificate field. The Bosh Director will then place this CA into the trust store of every VM that it deploys.

Generate VM passwords or use single password for all VMs

Generate passwords
 Use default BOSH password

Save

2. In **Trusted Certificates**, enter your custom certificate authority (CA) certificates to insert into your organization's certificate trust chain. This feature enables all BOSH-deployed components in your deployment to trust custom root certificates.

To enter multiple certificates, paste your certificates one after the other. For example, format your certificates like the following:

```
-----BEGIN CERTIFICATE-----
ABCDEFH12345678ABCDEFGH12345678ABCDEFGH12345678AB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
BCDEFH12345678ABCDEFGH12345678ABCDEFGH12345678ABB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
CDEFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABBB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
```

Note: If you want to use Docker registries to run PAS app instances in Docker containers, enter the certificate for your private Docker registry in this field. See [Using Docker Registries](#) for more information about running app instances in PAS using Docker registries.

- (Optional) Select the **Include OpsManager Root CA in Trusted Certs** checkbox to include the Ops Manager root CA in the Trusted Certificates field. BOSH Director includes this CA in the trust store of every VM that it deploys.
- Choose **Generate passwords** or **Use default BOSH password**. Pivotal recommends that you use the **Generate passwords** option for greater security.
- Click **Save**. To view your saved Director password, click the **Credentials** tab.

Step 8: BOSH DNS Config Page

BOSH DNS Config

Excluded Recursors

Recurser Timeout

Handlers

[]

Enter a JSON-formatted list of optional custom domain handlers

Save

1. Select **BOSH DNS Config**.
2. (Optional) In **Excluded Recursors**, enter a list of prohibited recursor addresses.
3. (Optional) In **Recurser Timeout**, enter a time limit for contacting the connected recursors. This includes dialing, writing, and reading from the recursor. If any of these actions exceeds the time limit you set, the action fails.

Note: This time limit must include one of the Go parse duration time units. For example, entering `5s` sets the timeout limit to five seconds. For more information about supported time units, see [func ParseDuration](#) in the Go Programming Language documentation.

4. (Optional) In **Handlers**, enter a list of custom domain handlers in JSON format.
5. Click **Save**.

Step 9: Syslog Page

Settings

Change Decryption Passphrase

SAML Settings

LDAP Settings

SSL Certificate

Pivotal Network Settings

Proxy Settings

Custom Banner

Export Installation Settings

Syslog

Syslog settings make take up to 1 min to take effect.

Do you want to configure Syslog for OpsManager?

No

Yes

Address*

The address or host for the syslog server

Port*

Transport Protocol*

Enable TLS

Permitted Peer*

SSL Certificate*

Queue Size

Forward Debug Logs

Custom rsyslog Configuration

1. Select **Syslog**.
2. (Optional) Select **Yes** to send BOSH Director system logs to a remote server.
3. Enter the IP address or DNS name for the remote server in **Address**.
4. Enter the port number that the remote server listens on in **Port**.
5. Select **TCP** or **UDP** from the **Transport Protocol** dropdown. This selection determines which transport protocol is used to send the logs to the remote server.
6. (Optional) Select the **Enable TLS** checkbox to send encrypted logs to remote server with TLS. After you select the checkbox, do the following:
 - a. Enter either the name or SHA1 fingerprint of the remote peer in **Permitted Peer**.
 - b. Enter the SSL certificate for the remote server in **SSL Certificate**.



Note: Pivotal strongly recommends that you enable TLS encryption when you are forwarding logs. Logs can contain sensitive

information, such as cloud provider credentials.

7. (Optional) Enter an integer in **Queue Size**. This value specifies the number of log messages held in the buffer. The default value is 100,000.
8. (Optional) Select the checkbox to **Forward Debug Logs** to an external source. This option is deselected by default. If you select it, you may generate a large amount of log data.
9. (Optional) Enter configuration details for rsyslog in the **Custom rsyslog Configuration** field. This field requires the rainerscript syntax.
10. Click **Save Syslog Settings**.

Step 10: Resource Config Page

1. Select **Resource Config**.

The screenshot shows the 'Resource Config' page with the following settings:

JOB	INSTANCES	PERSISTENT DISK TYPE	VM TYPE	LOAD BALANCERS	INTERNET CONNECTED
BOSH Director	Automatic: 1	Automatic: 50 GB	Automatic: large.disk (cpu: 2, ram: 8 GB, disk: 50 GB)		<input checked="" type="checkbox"/>
Master Compilation Job	1	None	small (cpu: 1, ram: 2 GB, disk: 8 GB)		<input checked="" type="checkbox"/>

A blue 'Save' button is at the bottom left.

2. Adjust any values as necessary for your deployment. Under the **Instances**, **Persistent Disk Type**, and **VM Type** fields, choose **Automatic** from the dropdown to allocate the recommended resources for the job. If the **Persistent Disk Type** field reads **None**, the job does not require persistent disk space.

Note: Pivotal recommends provisioning a BOSH Director VM with at least 8 GB memory.

Note: If you set a field to **Automatic** and the recommended resource allocation changes in a future version, Ops Manager automatically uses the new recommended allocation.

Note: If you install PASW, provision your **Master Compilation Job** with at least 100 GB of disk space.

Note: Ignore the Load Balancers field. See [Step 11](#) for details.

3. Click **Save**.

Step 11: Add VM Extensions for Load Balancers

You must assign each of your load balancer VMs to the correct AWS target groups and security groups. The following steps use the `om` CLI to configure both target groups and security groups. For more information, see [pivotalcf/om](#) in GitHub.

To assign your load balancer VMs to the correct target groups and security groups, do the following:

1. Run the following command to create a VM extension for the web VM:

```
om -k create-vm-extension \
--name "web-lb-security-groups" \
--cloud-properties '{ "security_groups": [ "web_lb_security_group", "vms_security_group"], "target_groups": [ "web-lb-target-groups" ] }'
```

2. Run the following command to create a VM extension for the SSH VM:

```
om -k create-vm-extension \
--name "ssh-lb-security-groups" \
--cloud-properties '{ "security_groups": [ "ssh_lb_security_group", "vms_security_group"], "target_groups": [ "ssh-lb-target-groups" ] }'
```

- Run the following command to create a VM extension for the TCP VM:

```
om -k create-vm-extension \
--name "tcp-lb-security-groups" \
--cloud-properties '{ "security_groups": [ "tcp_lb_security_group", "vms_security_group" ], "target_groups": [ "tcp-lb-target-groups" ] }'
```

Step 12: (Optional) Add Custom VM Extensions

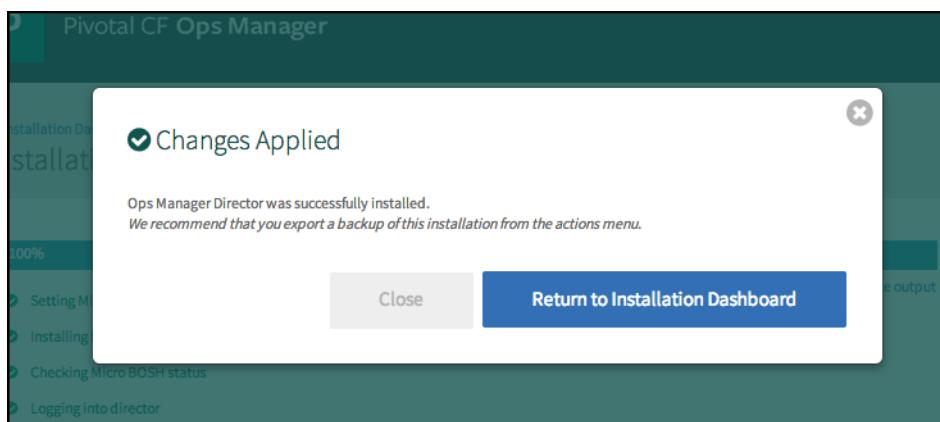
Use the Ops Manager API to add custom properties to your VMs such as associated security groups and load balancers. For more information, see [Managing Custom VM Extensions](#).

Step 13: Complete the BOSH Director Installation

- Click the **Installation Dashboard** link to return to the Installation Dashboard.
- Click **Review Pending Changes**, then **Apply Changes**. If the following ICMP error message appears, click **Ignore errors** and start the install.



- BOSH Director installs. This may take a few moments. When the installation process successfully completes, the **Changes Applied** window appears.

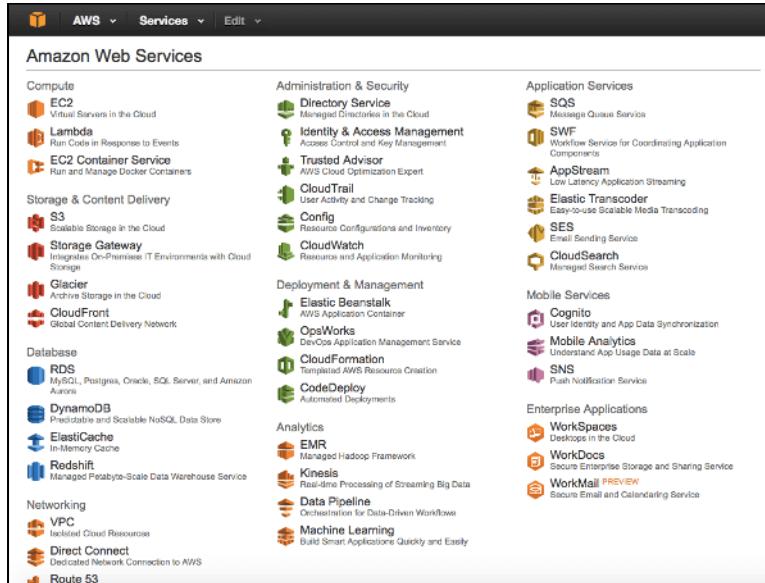


- After you complete this procedure, continue to the topic for the runtime you are installing:

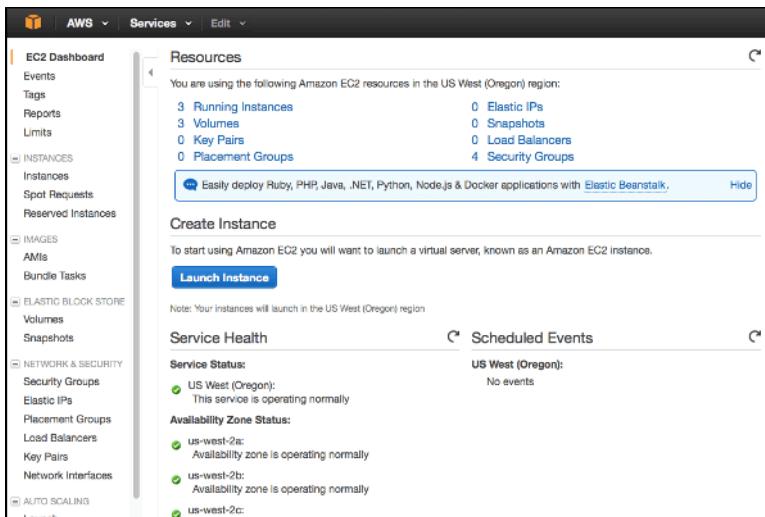
- PAS: [Configuring PAS](#)
- PKS v1.2: [Installing PKS on AWS](#)
- PKS v1.3: [Installing PKS on AWS](#)

Deleting PCF from AWS

When you deploy [Pivotal Cloud Foundry](#) (PCF) to Amazon Web Services (AWS), you provision a set of resources. This topic describes how to delete the AWS resources associated with a PCF deployment. You can use the AWS console to remove an installation of all components, but retain the objects in your bucket for a future deployment.



1. Log into your AWS Console.
2. Navigate to your EC2 dashboard. Select **Instances** from the menu on the left side.

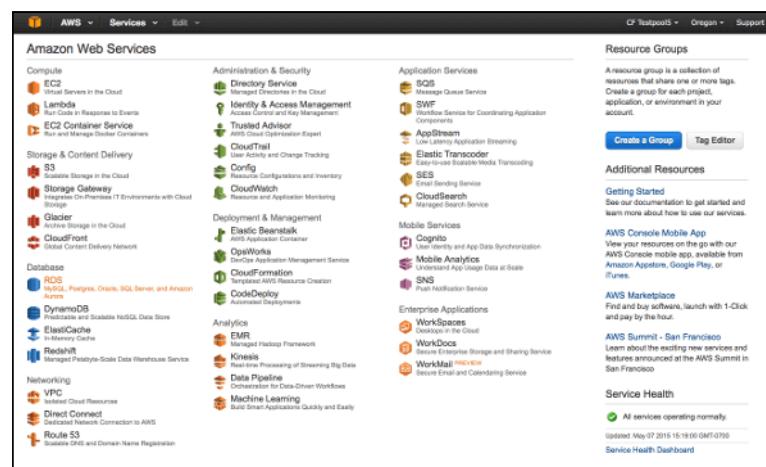


3. Terminate all your instances.

The screenshot shows the AWS EC2 Instances page. On the left, there's a sidebar with links like EC2 Dashboard, Events, Tags, Reports, Limits, Instances, Spot Requests, Reserved Instances, AMIs, Bundle Tasks, Volumes, Snapshots, Security Groups, Elastic IPs, Placement Groups, Load Balancers, Key Pairs, and Network Interfaces. The 'Instances' link is currently selected. In the main pane, three instances are listed: i-060963f0, i-0c0e64fa, and i-ba0e644c. Each instance has a green 'running' status indicator. An 'Actions' dropdown menu is open over the first instance, showing options: Connect, Get Windows Password, Launch More Like This, Instance State (with Start, Stop, Reboot, Terminate), Image, Networking, and CloudWatch Monitoring. The 'Terminate' option is highlighted with a red box. Below the instances, a list of their public DNS names is shown.

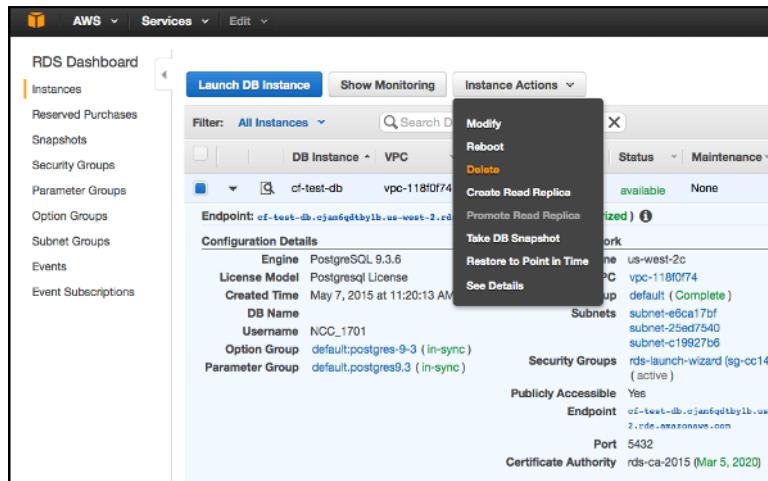
The screenshot shows the AWS Load Balancers page. The sidebar on the left is identical to the one in the previous screenshot. In the main pane, there is a single load balancer named 'CF-Test'. A context menu is open over this load balancer, listing options: Delete (highlighted with a red box), Edit health check, Edit subnets, Edit instances, Edit listeners, and Edit security groups. The 'Delete' option is also highlighted with a red box.

4. Select **Load Balancers**. Delete all load balancers.

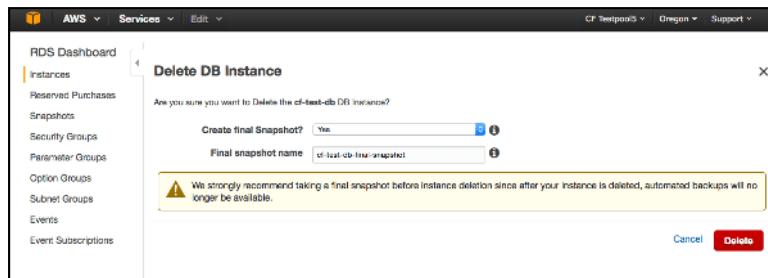


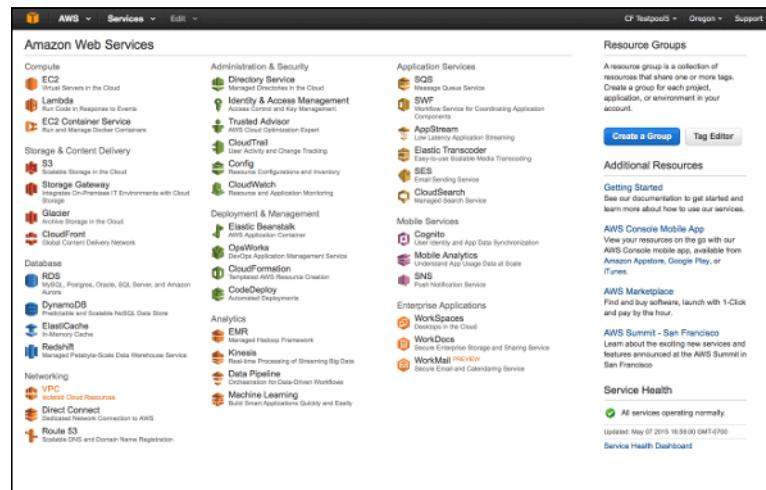
5. From the AWS Console, select RDS.

6. Select Instances from the menu on the left side. Delete the RDS instances.



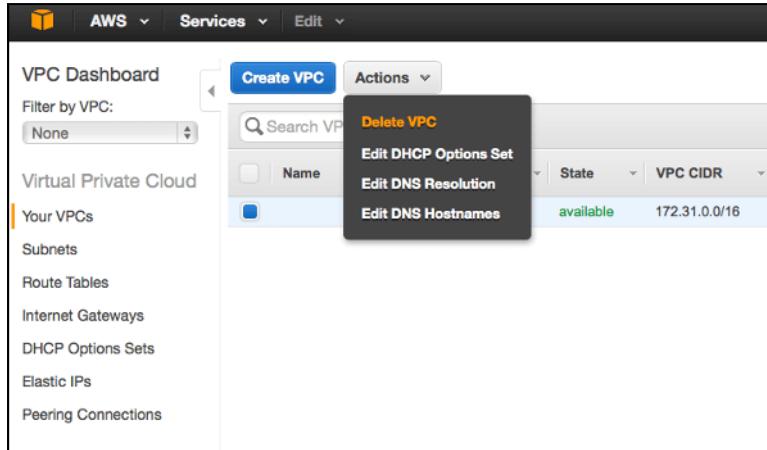
7. Select Create final Snapshot from the dropdown. Click Delete.



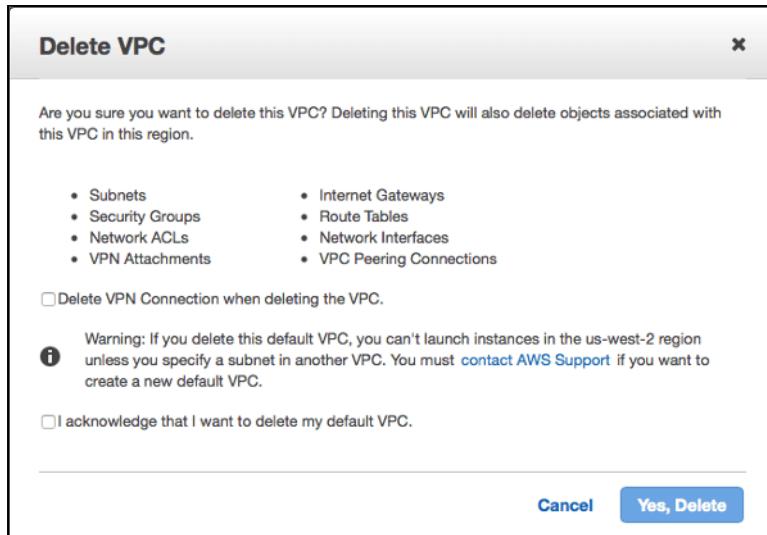


8. From the AWS Console, select VPC.

9. Select Your VPCs from the menu on the left. Delete the VPCs.



10. Check the box to acknowledge that you want to delete your default VPC. Click Yes, Delete.



Creating a Proxy ELB for Diego SSH

If you want to allow SSH connections to application containers, you may want to use an Elastic Load Balancer (ELB) as the SSH proxy.

Perform the steps below to create this ELB:

1. On the EC2 Dashboard, click **Load Balancers**.
2. Click **Create Load Balancer** and configure a classic load balancer with the following information:

Step 1: Define Load Balancer

Basic Configuration

This wizard will walk you through setting up a new load balancer. Begin by giving your new load balancer a unique name so that you can identify it from other load balancers you might create. You will also need to configure ports and protocols for your load balancer. Traffic from your clients can be routed from any load balancer port to any port on your EC2 instances. By default, we've configured your load balancer with a standard web server on port 80.

Load Balancer name:	my-SSH-ELB		
Create LB Inside:	vpc-4b38852e (10.0.0.0/16) pcf-vpc		
Create an internal load balancer:	<input type="checkbox"/> (what's this?)		
Enable advanced VPC configuration:	<input checked="" type="checkbox"/>		
Listener Configuration:			
Load Balancer Protocol	Load Balancer Port	Instance Protocol	Instance Port
TCP	2222	TCP	2222
<input type="button" value="Add"/>			

Select Subnets

You will need to select a Subnet for each Availability Zone where you wish traffic to be routed by your load balancer. If you have instances in only one Availability Zone, please select at least two Subnets in different Availability Zones to provide higher availability for your load balancer.

VPC vpc-4b38852e (10.0.0.0/16) | pcf-vpc

Available Subnets

Actions	Availability Zone	Subnet ID	Subnet CIDR	Name
+	us-west-1b	subnet-b63be1ef	10.0.16.0/20	pcf-private-subnet
+	us-west-1b	subnet-b73be1ee	10.0.2.0/24	pcf-rds-subnet-1
+	us-west-1c	subnet-c8f095ad	10.0.3.0/24	pcf-rds-subnet-2

Selected Subnets

Actions	Availability Zone	Subnet ID	Subnet CIDR	Name
-	us-west-1b	subnet-b43be1ed	10.0.0.0/24	pcf-public-subnet

- Enter a load balancer name.
 - **Create LB Inside:** Select the `pcf-vpc` VPC where your PCF installation lives.
 - Ensure that the **Create an internal load balancer** checkbox is not selected.
3. Under **Load Balancer Protocol**, ensure that this ELB is listening on TCP port `2222` and forwarding to TCP port `2222`.
 4. Under **Select Subnets**, select the public subnet.
 5. On the **Assign Security Groups** page, create a new Security Group. This Security Group should allow inbound traffic on TCP port `2222`.

Create Security Group

Security group name: PCF_SSH_ELB_SecurityGroup

Description:

VPC: vpc-4b38852e (10.0.0.0/16) |pcf-vpc| * denotes default VPC

Security group rules:

Inbound Outbound

Type	Protocol	Port Range	Source
Custom TCP Rule	TCP	2222	Anywhere 0.0.0.0/0

Add Rule

Cancel **Create**

6. The **Configure Security Settings** page displays a security warning because your load balancer is not using a secure listener. You can ignore this warning.

1. Define Load Balancer 2. Assign Security Groups 3. Configure Security Settings 4. Configure Health Check 5. Add EC2 Instances 6. Add Tags 7. Review

Step 3: Configure Security Settings

⚠ Improve your load balancer's security. Your load balancer is not using any secure listener.

If your traffic to the load balancer needs to be secure, use either the HTTPS or the SSL protocol for your front-end connection. You can go back to the first step to add/configure secure listeners under [Basic Configuration](#) section. You can also continue with current settings.

7. Click **Next: Configure Health Check**.

1. Define Load Balancer 2. Assign Security Groups 3. Configure Security Settings 4. Configure Health Check 5. Add EC2 Instances 6. Add Tags 7. Review

Step 4: Configure Health Check

Your load balancer will automatically perform health checks on your EC2 instances. Instances removed from the load balancer will be removed from the load balancer. Customize the health check to meet your needs.

Ping Protocol: TCP

Ping Port: 2222

Advanced Details

Response Timeout	5	seconds
Health Check Interval	30	seconds
Unhealthy Threshold	2	
Healthy Threshold	10	

8. Select **TCP** in **Ping Protocol** on the **Configure Health Check** page. Ensure that the **Ping Port** value is **2222** and set the **Health Check Interval** to **30** seconds.
9. Click **Next: Add EC2 Instances**.
10. Accept the defaults on the **Add EC2 Instances** page and click **Next: Add Tags**.
11. Accept the defaults on the **Add Tags** page and click **Review and Create**.
12. Review and confirm the load balancer details and click **Create**.
13. With your DNS service (for example, Amazon Route 53), create an **ssh.system.YOUR-SYSTEM-DOMAIN** DNS record that points to this ELB that you just created.

Create Record Set

Name: ssh.your-system-domain.com.

Type: CNAME – Canonical name

Alias: Yes No

TTL (Seconds): 300 | 1m | 5m | 1h | 1d

Value: your-elb-domain

The domain name that you want to resolve to instead of the value in the Name field.

Example:
www.example.com

14. You can now use this ELB to the SSH Proxy of your Pivotal Application Service (PAS) installation.

15. In PAS, select **Resource Config** and enter the ELB that you just created in the **Diego Brain** row, under the Load Balancers column.

UAA	1	None	Automatic: t2.small (cpu: 1, ram: 2 GB, disk: 2 GB)		
Diego Brain	1	Automatic: 1 GB	Automatic: m3.medium (cpu: 1, ram: 3.75 GB, disk: 4 GE)		passion-elb
Diego Cell	3	None	m3.2xlarge (cpu: 8, ram: 30 GB, disk: 64 GB)		

Upgrading BOSH Director on AWS

This topic describes how to upgrade BOSH Director for Pivotal Cloud Foundry (PCF) on Amazon Web Services (AWS).

Complete the tasks in this topic as part of the Ops Manager upgrade process. For more information, see [Upgrading Pivotal Cloud Foundry](#).

Overview

In this procedure, you create an Ops Manager VM instance to host the upgraded version of Ops Manager. Then, to complete the Ops Manager upgrade, you export your existing Ops Manager installation onto this new VM.

For PCF installations on AWS, the VM that hosts the new version of Ops Manager uses an Amazon Machine Image (AMI) specific to your region.

To create an Ops Manager VM instance, do the following:

1. Retrieve the Ops Manager AMI ID. See [Retrieve Ops Manager AMI ID](#).
2. Launch the AMI. See [Launch Ops Manager AMI](#).
3. Edit the Ops Manager DNS A Record. See [Edit Ops Manager DNS A Record](#).

Prerequisites

To complete the Ops Manager upgrade, you must have your Ops Manager decryption passphrase. You defined this decryption passphrase during the initial installation of Ops Manager.

Retrieve Ops Manager AMI ID

Ops Manager has a different AMI ID for each region. Retrieve and record the AMI ID for your region. You use this AMI ID when launching the Ops Manager AMI instance.

To retrieve the AMI ID for Ops Manager, perform the following steps:

1. Navigate to the **Pivotal Cloud Foundry Operations Manager** section of [Pivotal Network](#).
2. Select the version of PCF you want to install from the **Releases** dropdown.
3. In the **Release Download Files**, click the file named **Pivotal Cloud Foundry Ops Manager for AWS** to download a PDF.
4. Open the PDF and record the AMI ID for your region.

Launch Ops Manager AMI

Use the AMI ID for your region to launch an AMI instance that hosts the new version of Ops Manager.

To launch the AMI for Ops Manager, perform the following steps:

1. Navigate to your EC2 Dashboard.
2. Click **AMIs** from the **Images** menu.
3. Select **Public images** from the drop-down filter that says **Owned by me**.
4. Paste the AMI ID for your region into the search bar and press enter.

 **Note:** There is a different AMI for each region. If you cannot locate the AMI for your region, verify that you have set your AWS Management Console to your desired region. If you still cannot locate the AMI, log in to the [Pivotal Network](#) and file a support ticket.

5. (Optional) If you want to encrypt the VM that runs Ops Manager with AWS Key Management Service (KMS), perform the following additional steps:

- Right click the row that lists your AMI and click **Copy AMI**.
- Select your **Destination region**.
- Enable **Encryption**. For more information about AMI encryption, see [Encryption and AMI Copy](#) from the *Copying an AMI* topic in the AWS documentation.
- Select your **Master Key**. To create a new custom key, see [Creating Keys](#) in the AWS documentation.
- Click **Copy AMI**. You can use the new AMI you copied for the following steps.

6. Select the row that lists your Ops Manager AMI and click **Launch**.

7. Choose **m5.large** for your instance type and click **Next: Configure Instance Details**.

<input checked="" type="checkbox"/>	General purpose	m5d.12xlarge	48	192	2 x 900 (SSD)
<input type="checkbox"/>	General purpose	m5d.24xlarge	96	384	4 x 900 (SSD)
<input checked="" type="checkbox"/>	General purpose	m5.large	2	8	EBS only
<input type="checkbox"/>	General purpose	m5.xlarge	4	16	EBS only
<input type="checkbox"/>	General purpose	m5.2xlarge	8	32	EBS only

8. Configure the following for your instance:

- **Network:** Select the VPC that you created.
- **Subnet:** Select `pcf-public-subnet-az0`.
- **Auto-assign for Public IP:** Select **Enable**.
- **IAM role:** Select the IAM role associated with your pcf-user profile. If you have not created one, click **Create new IAM role** and follow the [Guidelines for Creating User Roles on AWS](#).
- For all other fields, accept the default values.

9. Click **Next: Add Storage** and adjust the **Size (GiB)** value. The default persistent disk value is 50 GB. Pivotal recommends increasing this value to a minimum of 100 GB.

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Tag Instance 6. Configure Security Group 7. Review

Step 4: Add Storage

Your instance will be launched with the following storage device settings. You can attach additional EBS volumes and instance store volumes to your instance, or edit the settings of the root volume. You can also attach additional EBS volumes after launching an instance, but not instance store volumes. [Learn more](#) about storage options in Amazon EC2.

Type	Device	Snapshot	Size (GiB)	Volume Type	IOPS	Delete on Termination	Encrypted
Root	/dev/sda1	snap-f1ce6d7e	100	General Purpose (SSD)	150 / 3000	<input checked="" type="checkbox"/>	Not Encrypted
Instance Store 0	/dev/sdb	N/A	N/A	N/A	N/A	N/A	Not Encrypted
Add New Volume							

10. Click **Next: Tag Instance**
11. On the **Add Tags** page, add a tag with the key `Name` and value `pcf-ops-manager`.
12. Click **Next: Configure Security Group**.
13. Select the `pcf-ops-manager-security-group` that you created in [Step 5: Configure a Security Group for Ops Manager](#).
14. Click **Review and Launch** and confirm the instance launch details.
15. Click **Launch**.

16. Select the `pcf-ops-manager-key` key pair, confirm that you have access to the private key file, and click **Launch Instances**. You use this key pair to access the Ops Manager VM.

Select an existing key pair or create a new key pair

A key pair consists of a **public key** that AWS stores, and a **private key file** that you store. Together, they allow you to connect to your instance securely. For Windows AMIs, the private key file is required to obtain the password used to log into your instance. For Linux AMIs, the private key file allows you to securely SSH into your instance.

Note: The selected key pair will be added to the set of keys authorized for this instance. Learn more about [removing existing key pairs from a public AMI](#).

Choose an existing key pair

Select a key pair
pcf

I acknowledge that I have access to the selected private key file (pcf.pem), and that without this file, I won't be able to log into my instance.

[Cancel](#) [Launch Instances](#)

Edit Ops Manager DNS A Record

After you deploy the new Ops Manager VM, edit the Ops Manager DNS record in the EC2 dashboard to point to the IP address for the new VM.

To edit the Ops Manager DNS A record, do the following:

1. Click **View Instances** to access the **Instances** page on the EC2 Dashboard.
2. Select the VM that you created in the previous section.
3. Locate the **IPv4 Public IP** value in the instance **Description** tab, and record this value for use in the next step.
4. In your DNS provider, edit the A record for `pcf.YOUR-SYSTEM-DOMAIN` to point to the IP address recorded in the previous step.

Next Steps

After you complete this procedure, continue the upgrade instructions in Upgrading Pivotal Cloud Foundry topic.

Installing Pivotal Cloud Foundry on Azure

This topic describes how to install Pivotal Cloud Foundry (PCF) on Microsoft Azure.

It includes resource requirements, prerequisites, instructions for installing PCF on Azure, and additional resources.

Overview

You can install PCF on Azure, Azure Government Cloud, or Azure Germany. For more information about Azure regions, see [Azure regions](#) in the Microsoft Azure documentation.

You can install PCF on Azure with the Pivotal Application Service (PAS) runtime. There are resource requirements specific to both PAS and Azure. Ensure you meet the requirements for PAS and the requirements specific to Azure before installing PCF on Azure. For more information, see [Requirements](#).

Requirements

This section lists the following resource requirements for installing PCF on Azure:

- PAS resource requirements. See [PAS Resource Requirements](#).
- Azure-specific resource requirements. See [Azure Resource Requirements](#).

PAS Resource Requirements

The following are general resource requirements for deploying and managing a PCF deployment with Ops Manager and PAS:

- PAS requires sufficient IP allocation. The following lists the minimum required IP allocations:
 - One static IP address for either HAProxy or one of your Gorouters
 - One static IP address for each job in the Ops Manager tile. See the Ops Manager **Resource Config** pane for each tile for a full list.
 - One static IP address for each job listed below:
 - Consul
 - NATS
 - File Storage
 - MySQL Proxy
 - MySQL Server
 - Backup Restore Node
 - HAProxy
 - Router
 - MySQL Monitor
 - Diego Brain
 - TCP Router
 - One IP for each VM instance created by the service.
 - An additional IP address for each compilation worker. Use the following formula to determine the total IPs required:

$$\text{IPs needed} = \text{static IPs} + \text{VM instances} + \text{compilation workers}$$
 - Pivotal recommends that you allocate at least 36 dynamic IP addresses when deploying Ops Manager and PAS. BOSH requires additional dynamic IP addresses during installation to compile and deploy VMs, install PAS, and connect to services.
- Pivotal recommends using a network without DHCP for deploying PAS VMs.

 **Note:** If you have DHCP, refer to the [Troubleshooting Guide](#) to avoid issues with your installation.

Azure Resource Requirements

The following are the minimum resource requirements for deploying a PCF deployment with Ops Manager and PAS on Azure:

- An OS disk of 120 GB for the Ops Manager VM.

 **Note:** Ops Manager v1.11 requires a Director VM with at least 8 GB memory.

- A minimum of the following VM instance limits in your Azure account. The number of VMs required depends on the number of tiles and availability zones you plan to deploy. The following VM guidelines apply to the PAS and Small Footprint PAS runtimes:

- **PAS:** A new Azure deployment requires the following VMs for PAS and Ops Manager:

VM Type	VM Count
F1s	22
F2s	22
DS1 v2	2
DS11 v2	10
DS12 v2	24

 **Note:** If you are deploying a test or sandbox PCF that does not require high availability, then you can scale down the number of VM instances in your deployment. For more information, see [Scaling PAS](#).

- **Small Footprint PAS:** A new Azure deployment requires the following VMs for Small Footprint PAS and Ops Manager:

	VM Type	VM Count	Notes
Small Footprint PAS	DS11 v2	1	
	DS12 v2	2	
	F2s	0	Add 1 to count if using HAProxy
	F1s	5	
Ops Manager	DS2 v2	1	
	F4s	4	

 **Note:** Some VM instance types are only supported in certain Azure regions. For more information, see [Products available by region](#) in the Azure documentation. To deploy PCF on an Azure region that does not support the above VM instance types, you must override the default VM sizes. For more information, see [Overriding defaults with custom disk types](#) in the Ops Manager API documentation. Changing the default VM sizes may increase the cost of your deployment.

Prerequisites

Before installing PCF on Azure, you must do the following:

- Install the Azure CLI v2.0. For instructions on how to install the Azure CLI for your operating system, see [Preparing to Deploy Ops Manager on Azure Manually](#).
- Configure sufficient IP allocation for PAS. For more information about IP allocation requirements, see [PAS Resource Requirements](#).
- Create an SSL certificate for your PCF domain.

 **Note:** To deploy PCF to a production environment, you must obtain a certificate from a certificate authority. Pivotal recommends using a self-signed certificate generated by Ops Manager for development and testing purposes only.

- Assign administrative rights to a domain for PCF. You must be able to add wildcard records to this domain.
- Create a wildcard DNS record that points to your router or load balancer. Alternatively, you can use a service such as xip.io. For example, 203.0.113.0.xip.io. Then, create at least one wildcard TLS certificate that matches the DNS record you configured.

 **Note:** With a wildcard DNS record, every hostname in your domain resolves to the IP address of your router or load balancer. For example, if you create a DNS record *.example.com pointing to your router, every app deployed to the example.com domain resolves to the IP address of your router.

- Create one or more NTP servers, if the NTP servers are not already provided by your Azure project.
- Install the most recent version of the Cloud Foundry Command Line Interface (cf CLI). See [Cloud Foundry Command Line Interface](#) in GitHub.
- **(Optional)** Configure external storage. Pivotal recommends using external storage if possible. For more information about how file storage location affects platform performance and stability during upgrades, see [Configure File Storage](#).

- **(Optional)** Configure external databases. Pivotal recommends using external databases in production deployments for BOSH Director and PAS. An external database must be configured to use the UTC timezone.
- **(Optional)** Configure external user stores. When you deploy PCF, you can select a SAML user store for Ops Manager or a SAML or LDAP user store for PAS, to integrate existing user accounts.

Install PCF on Azure

You can install PCF on Azure either manually or using Terraform.

To install PCF on Azure, do one of the following:

- Install PCF on Azure manually. See [Installing PCF on Azure Manually](#).
- Install PCF on Azure using Terraform. See [Installing PCF on Azure Using Terraform](#).

Install PCF on Azure Government Cloud

To install PCF on Azure Government Cloud, see [Configuring BOSH Director on Azure Government Cloud](#).

Install PCF on Azure Germany

To install PCF in Azure Germany, see [Configuring BOSH Director on Azure Germany](#).

Additional Resources

The following are additional resources related to installing PCF on GCP:

- For information about troubleshooting the PCF on Azure install process, see [Troubleshooting PCF on Azure](#).
- For production-level deployment options for PCF on Azure, see [Reference Architecture for Pivotal Cloud Foundry on Azure](#).
- For information about how to manage IaaS users and credentials in Azure, see [Azure security documentation](#).
- For recommendations on how to create and scope Azure accounts for PCF, see [IaaS Permissions Guidelines](#).
- For information about deleting a PCF on Azure installation, see [Deleting a PCF on Azure Installation](#).
- For information about upgrading PCF on Azure, see [Upgrading BOSH Director on Azure](#).

Installing PCF on Azure Manually

This topic explains how to install Pivotal Cloud Foundry (PCF) on Microsoft Azure manually.

To install PCF on Azure manually, do the following:

1. Deploy Ops Manager. See [Deploying Ops Manager on Azure Manually](#).
2. Configure BOSH Director. See [Configuring BOSH Director on Azure Manually](#).
3. Configure PAS. See [Configuring PAS](#).

Preparing to Deploy Ops Manager on Azure Manually

This topic describes how to prepare Azure to deploy Ops Manager. You must deploy Ops Manager to deploy Pivotal Application Service (PAS) or Pivotal Container Service (PKS).

Step 1: Install and Configure the Azure CLI

1. Install the Azure CLI 2.0 by following the instructions for your operating system in [Install the Azure CLI](#) in the Microsoft Azure documentation.
2. On the command line, set your cloud with a `name` value by running the following command:

```
az cloud set --name CLOUD-NAME
```

Where `CLOUD-NAME` is one of the following values. Use the value that corresponds to the Azure environment on which you are installing Ops Manager:

- o **Azure:** `AzureCloud`.
- o **Azure China:** `AzureChinaCloud`. If logging in to `AzureChinaCloud` fails with a `CERT_UNTRUSTED`, use the latest version of node, 4.x or later. For more information about this error, see [Failed to login AzureChinaCloud](#) in the Azure/azure-xplat-cli GitHub repository.
- o **Azure Government Cloud:** `AzureUSGovernment`.
- o **Azure Germany:** `AzureGermanCloud`.

For example:

```
$ az cloud set --name AzureCloud
```

 **Note:** For more information about installing Ops Manager in the China Region, see [Install in the China Region](#) in *Installing PCF in Airgapped Environments* in the Pivotal documentation.

3. To log in, run the following command:

```
az login
```

Authenticate by navigating to the URL in the output from the command, entering the provided code, and clicking your account.

Step 2: Set Your Default Subscription

1. To list your Azure subscriptions, run the following command:

```
az account list
```

For example:

```
$ az account list
[
{
  "id": "12345678-1234-5678-1234-567891234567",
  "name": "Sample Subscription",
  "user": {
    "name": "Sample Account",
    "type": "user"
  },
  "tenantId": "11111111-1234-5678-1234-678912345678",
  "state": "Enabled",
  "isDefault": true,
  "registeredProviders": [],
  "environmentName": "AzureCloud"
},
{
  "id": "87654321-1234-5678-1234-678912345678",
  "name": "Sample Subscription1",
  "user": {
    "name": "Sample Account1",
    "type": "user"
  },
  "tenantId": "22222222-1234-5678-1234-678912345678",
  "state": "Enabled",
  "isDefault": false,
  "registeredProviders": [],
  "environmentName": "AzureCloud"
}
]
```

- Identify your default subscription in the output of this command by locating the subscription listed that has `isDefault` set to `true`. To deploy Ops Manager to a different subscription, you must set the other subscription as the default. To set a new default subscription, run the following command:

```
az account set --subscription SUBSCRIPTION-ID
```

Where `SUBSCRIPTION-ID` is the value of the `id` field of the subscription that you want to set as the default.

The following example sets the default to a subscription that has an `id` with the value of `"87654321-1234-5678-1234-567891234567"`.

```
$ az account set --subscription "87654321-1234-5678-1234-567891234567"
```

- Record the value of the `id` of the default subscription. You use this value in future configuration steps.
- Record the value of `tenantID` of the default subscription. This is your `TENANT_ID` for creating a service principal. If your `tenantID` value is not defined, you may be using a personal account to log in to your Azure subscription.

Step 3: Create an AAD Application

- To create an Azure Active Directory (AAD) application, run the following command:

```
az ad app create --display-name "Service Principal for BOSH" \
--password "PASSWORD" --homepage "http://BOSHAzureCPI" \
--identifier-uris "http://BOSHAzureCPI"
```

Where `PASSWORD` is a password of your choice. This is your `CLIENT_SECRET` for creating a service principal.

For example:

```
$ az ad app create --display-name "Service Principal for BOSH" \
--password "Swordfish" --homepage "http://BOSHAzureCPI" \
--identifier-uris "http://BOSHAzureCPI"
```

Note: You can provide any string for the `homepage` and `identifier-uris` flags, but the value of `identifier-uris` must be unique within the organization associated with your Azure subscription. For the `homepage`, Pivotal recommends using `http://BOSHAzureCPI` as shown in the example above.

- From the output of the `az ad app create` that you ran in the previous step, locate and record the value of `appId`. This is your `APPLICATION_ID` for creating a service principal.

The following example output shows `"5c552e8f-b977-45f5-a50b-981cf17cb9d"` as the value of `appId`.

```
{
  "appId": "5c552e8f-b977-45f5-a50b-981cf17cb9d",
  "appPermissions": null,
  "availableToOtherTenants": false,
  "displayName": "Service Principal for BOSH",
  "homepage": "http://BOSHAzureCIP",
  "identifierUris": [
    "http://BOSHAzureCPI"
  ],
  "objectId": "f3884df4-7d1d-4894-a78c-c1fe75750436",
  "objectType": "Application",
  "replyUrls": []
}
```

Step 4: Create and Configure a Service Principal

1. To create a service principal, run the following command:

```
az ad sp create --id YOUR-APPLICATION-ID
```

Where `YOUR-APPLICATION-ID` is the `APPLICATION_ID` that you recorded in [Create an AAD Application](#).

For example:

```
$ az ad sp create --id 5c552e8f-b977-45f5-a50b-981cf17cb9d
{
  "appId": "5c552e8f-b977-45f5-a50b-981cf17cb9d",
  "displayName": "Service Principal for BOSH",
  "objectId": "cc13c685-4c3b-461e-ae96-7a0563960b83",
  "objectType": "ServicePrincipal",
  "servicePrincipalNames": [
    "5c552e8f-b977-45f5-a50b-981cf17cb9d",
    "http://BOSHAzureCPI"
  ]
}
```

2. You must have the Contributor role on your service principal to deploy Ops Manager to Azure. To assign the Contributor role on your service principal, run the following command:

```
az role assignment create --assignee "SERVICE-PRINCIPAL-NAME" \
--role "Contributor" --scope /subscriptions/SUBSCRIPTION-ID
```

Where:

- `SERVICE-PRINCIPAL-NAME` is any value of `servicePrincipalNames` from the output above, such as `YOUR-APPLICATION-ID`.
- `SUBSCRIPTION-ID` is the value of the `id` of the default subscription that you recorded in [Set Your Default Subscription](#).

For example:

```
$ az role assignment create \
--assignee "5c552e8f-b977-45f5-a50b-981cf17cb9d" \
--role "Contributor" \
--scope /subscriptions/87654321-1234-5678-1234-567891234567
```

 **Note:** If you need to use multiple resource groups for your deployment on Azure, you can define custom roles for your Service Principal. These roles allow BOSH to deploy to pre-existing network resources outside of the resource group. For more information, see [Multiple Resource Group Deployment](#) in *Reference Architecture for Pivotal Cloud Foundry on Azure* in the Pivotal documentation.

For more information about Azure Role-Based Access Control, refer to [Built-in roles for Azure resource](#) in the Microsoft Azure documentation.

3. To verify the role assignment, run the following command:

```
az role assignment list --assignee "SERVICE-PRINCIPAL-NAME"
```

Where `"SERVICE-PRINCIPAL-NAME"` is any value of `servicePrincipalNames` from the output above, such as `YOUR-APPLICATION-ID`.

For example:

```
$ az role assignment list --assignee "5c552e8f-b977-45f5-a50b-981cf17cb9d"
[
{
  "id": "/subscriptions/995b7eed-77ef-45ff-a5c9-1a405ffb8243/providers/Microsoft.Authorization/roleAssignments/32e644cf-ba1a-4f43-bf7c-68bf4583e463",
  "name": "32e644cf-ba1a-4f43-bf7c-68bf4583e463",
  "properties": {
    "principalId": "cc13c685-4c3b-461e-ae96-7a0563960b83",
    "principalName": "http://BOSHAzureCPI",
    "roleDefinitionId": "/subscriptions/995b7eed-77ef-45ff-a5c9-1a405ffb8243/providers/Microsoft.Authorization/roleDefinitions/b24988ac-6180-42a0-ab88-20f7382dd24c",
    "roleDefinitionName": "Contributor",
    "scope": "/subscriptions/995b7eed-77ef-45ff-a5c9-1a405ffb8243"
  },
  "type": "Microsoft.Authorization/roleAssignments"
}
]
```

Step 5: Verify Your Service Principal

To verify your service principal, run the following command to log in to your service principal:

```
az login --username APPLICATION_ID --password CLIENT_SECRET \
--service-principal --tenant TENANT_ID
```

Where:

- `APPLICATION_ID` is the `APPLICATION_ID` that you recorded in [Create an AAD Application](#).
- `CLIENT_SECRET` is the password that you provided in [Create an AAD Application](#).
- `TENANT_ID` is the value of `tenantID` of the default subscription that you recorded in [Set Your Default Subscription](#).

For example:

```
$ az login --username 5c552e8f-b977-45f5-a50b-981cf17cb9d \
--password "Swordfish" \
--service-principal --tenant 22222222-1234-5678-1234-678912345678
[
{
  "cloudName": "AzureCloud",
  "id": "995b7eed-77ef-45ff-a5c9-1a405ffb8243",
  "isDefault": true,
  "name": "my-user-name",
  "state": "Enabled",
  "tenantId": "22222222-1234-5678-1234-678912345678",
  "user": {
    "name": "5c552e8f-b977-45f5-a50b-981cf17cb9d",
    "type": "servicePrincipal"
  }
}
```

If you cannot log in, the service principal is invalid. Create a new service principal and try again.

Step 6: Perform Registrations

1. To register your subscription with Microsoft.Storage, run the following command:

```
az provider register --namespace Microsoft.Storage
```

2. To register your subscription with Microsoft.Network, run the following command:

```
az provider register --namespace Microsoft.Network
```

3. To register your subscription with Microsoft.Compute, run the following command:

```
az provider register --namespace Microsoft.Compute
```

Next Steps

After you complete this procedure, follow the instructions in [Deploying Ops Manager on Azure Manually](#).

Deploying Ops Manager on Azure Manually

This topic describes how to deploy BOSH and Ops Manager for Pivotal Cloud Foundry (PCF) on Azure by using individual commands to create resources. Pivotal recommends this manual procedure for deploying to Azure China, Azure Germany, and Azure Government Cloud.

Before you perform the procedures in this topic, you must have completed the procedures in [Preparing to Deploy Ops Manager on Azure Manually](#). After you complete the procedures in this topic, follow the instructions in [Configuring BOSH Director on Azure](#).

 **Note:** If you are deploying BOSH and Ops Manager on Azure Stack, complete the procedures in [Install and configure CLI for use with Azure Stack](#) in the Microsoft documentation before following the procedures in this topic.

 **Note:** The Azure portal sometimes displays the names of resources with incorrect capitalization. Always use the Azure CLI to retrieve the correctly capitalized name of a resource.

Step 1: Create Network Resources

To create network resources for your deployment, do the following:

1. Navigate to the Azure portal, click **Resource groups**, and click **Add** to create a new resource group for your deployment.
2. Enter a **Resource group name**, select your **Subscription**, and select a **Resource group location**. Click **Create**.
3. Export the name of your resource group as the environment variable `$RESOURCE_GROUP`.

```
$ export RESOURCE_GROUP="YOUR-RESOURCE-GROUP-NAME"
```

 **Note:** If you are on a Windows machine, you can use `set` instead of `export`.

4. Export your location. For example, `westus`.

```
$ export LOCATION=westus
```

For a list of available locations, run `az account list-locations`.

5. Create a network security group named `pcf-nsg`.

```
$ az network nsg create --name pcf-nsg \
--resource-group $RESOURCE_GROUP \
--location $LOCATION
```

6. Add network security group rules to the `pcf-nsg` group to allow traffic to known ports from the public Internet.

```
$ az network nsg rule create --name ssh \
--nsg-name pcf-nsg --resource-group $RESOURCE_GROUP \
--protocol Tcp --priority 100 \
--destination-port-range '22'
```

```
$ az network nsg rule create --name http \
--nsg-name pcf-nsg --resource-group $RESOURCE_GROUP \
--protocol Tcp --priority 200 \
--destination-port-range '80'
```

```
$ az network nsg rule create --name https \
--nsg-name pcf-nsg --resource-group $RESOURCE_GROUP \
--protocol Tcp --priority 300 \
--destination-port-range '443'
```

```
$ az network nsg rule create --name diego-ssh \
--nsg-name pcf-nsg --resource-group $RESOURCE_GROUP \
--protocol Tcp --priority 400 \
--destination-port-range '2222'
```

- To block traffic from the public Internet, append `--source-address-prefixes AzureLoadBalancer` to allow traffic from only the Azure load balancer or `--source-address-prefixes VirtualNetwork` to only allow traffic from the virtual network.
- To allow traffic from both the Azure load balancer and the virtual network, create duplicates of each rule, one specifying `--source-address-prefixes AzureLoadBalancer` and one specifying `--source-address-prefixes VirtualNetwork`.

7. Create a network security group named `opsmgr-nsg`.

```
$ az network nsg create --name opsmgr-nsg \
--resource-group $RESOURCE_GROUP \
--location $LOCATION
```

8. Add a network security group rule to the `opsmgr-nsg` group to allow HTTP traffic to the Ops Manager VM.

```
$ az network nsg rule create --name http \
--nsg-name opsmgr-nsg --resource-group $RESOURCE_GROUP \
--protocol Tcp --priority 100 \
--destination-port-range 80
```

9. Add a network security group rule to the `opsmgr-nsg` group to allow HTTPS traffic to the Ops Manager VM.

```
$ az network nsg rule create --name https \
--nsg-name opsmgr-nsg --resource-group $RESOURCE_GROUP \
--protocol Tcp --priority 200 \
--destination-port-range 443
```

10. Add a network security group rule to the `opsmgr-nsg` group to allow SSH traffic to the Ops Manager VM.

```
$ az network nsg rule create --name ssh \
--nsg-name opsmgr-nsg --resource-group $RESOURCE_GROUP \
--protocol Tcp --priority 300 \
--destination-port-range 22
```

- To block traffic from the public Internet, append `--source-address-prefixes AzureLoadBalancer` to allow traffic from only the Azure load balancer or `--source-address-prefixes VirtualNetwork` to only allow traffic from the virtual network.
- To allow traffic from both the Azure load balancer and the virtual network, create duplicates of each rule, one specifying `--source-address-prefixes AzureLoadBalancer` and one specifying `--source-address-prefixes VirtualNetwork`.
- Optionally, if you want to use private IP ranges with Ops Manager and allow all internal traffic, you can create the Network Security Groups to allow all internal traffic.

```
$ az network nsg rule create --name internal-virtual-network \
--nsg-name internal-traffic --resource-group $RESOURCE_GROUP \
--protocol Tcp --priority 100 \
--destination-port-range * \
--source-address-prefixes VirtualNetwork
$ az network nsg rule create --name internal-from-lb \
--nsg-name internal-traffic --resource-group $RESOURCE_GROUP \
--protocol Tcp --priority 110 \
--destination-port-range * \
--source-address-prefixes AzureLoadBalancer
```

11. Create a virtual network named `pcf-virtual-network`.

```
$ az network vnet create --name pcf-virtual-network \
--resource-group $RESOURCE_GROUP --location $LOCATION \
--address-prefixes 10.0.0.0/16
```

12. Add subnets to the network for Ops Manager, BOSH director, and the VMs for your runtime, and attach the Network Security Group.

```
$ az network vnet subnet create --name pcf-infrastructure-subnet \
--vnet-name pcf-virtual-network \
--resource-group $RESOURCE_GROUP \
--address-prefix 10.0.4.0/26 \
--network-security-group pcf-nsg
$ az network vnet subnet create --name pcf-pas-subnet \
--vnet-name pcf-virtual-network \
--resource-group $RESOURCE_GROUP \
--address-prefix 10.0.12.0/22 \
--network-security-group pcf-nsg
$ az network vnet subnet create --name pcf-services-subnet \
--vnet-name pcf-virtual-network \
--resource-group $RESOURCE_GROUP \
--address-prefix 10.0.8.0/22 \
--network-security-group pcf-nsg
```

Step 2: Create BOSH and Deployment Storage Accounts

PCF on Azure uses multiple general-purpose Azure storage accounts. The BOSH and Ops Manager VMs use one main BOSH account, and the other components share five or more deployment storage accounts.

To create storage accounts for BOSH and your deployment, do the following:

1. Choose a name for your BOSH storage account, and export it as the environment variable `$STORAGE_NAME`. Storage account names must be globally unique across Azure, between 3 and 24 characters in length, and contain only lowercase letters and numbers.

```
$ export STORAGE_NAME="YOUR-BOSH-STORAGE-ACCOUNT-NAME"
```

2. Create a Standard storage account for BOSH with the following command. This account will be used for BOSH bookkeeping and running the Ops Manager VM itself, but does not have to be used for running any other VMs.

```
$ az storage account create --name $STORAGE_NAME \
--resource-group $RESOURCE_GROUP \
--sku Standard_LRS \
--location $LOCATION
```

Note: `Standard_LRS` refers to a Standard Azure storage account. The BOSH Director requires table storage to store stemcell information. Azure Premium storage does not support table storage and cannot be used for the BOSH storage account.

If the command fails, ensure you have followed the rules for naming your storage account. Export another new storage account name if necessary.

3. Configure the Azure CLI to use the BOSH storage account as its default.

- a. Retrieve the connection string for the account.

```
$ az storage account show-connection-string \
--name $STORAGE_NAME --resource-group $RESOURCE_GROUP
```

The command returns output similar to the following:

```
{
  "connectionString": "DefaultEndpointsProtocol=https;EndpointSuffix=core.windows.net;AccountName=cfdocsboshsstorage;AccountKey=EXAMPLEaaaaabbbrcn5igFxYWsgq016"
}
```

- b. Record the full value of `connectionString` from the output above, starting with and including `DefaultEndpointsProtocol=`.
- c. Export the value of `connectionString` as the environment variable `$CONNECTION_STRING`.

```
$ export CONNECTION_STRING="YOUR-ACCOUNT-KEY-STRING"
```

4. Create three blob containers in the BOSH storage account, named `opsmanager`, `bosh`, and `stemcell`.

```
$ az storage container create --name opsmanager \
--connection-string $CONNECTION_STRING
$ az storage container create --name bosh \
--connection-string $CONNECTION_STRING
$ az storage container create --name stemcell --public-access blob \
--connection-string $CONNECTION_STRING
```

5. Create a table named `stemcells`.

```
$ az storage table create --name stemcells \
--connection-string $CONNECTION_STRING
```

6. Choose a set of unique names for five or more deployment storage accounts. As with the BOSH storage account above, the names must be unique, alphanumeric, lowercase, and 3-24 characters long. The account names must also be sequential or otherwise identical except for the last character. For example: `xyzdeploystorage1`, `xyzdeploystorage2`, `xyzdeploystorage3`, `xyzdeploystorage4`, and `xyzdeploystorage5`.

7. Decide which type of storage to use and run the corresponding command below.

 **Note:** Pivotal recommends five Premium storage accounts, which provides a reasonable amount of initial storage capacity. You can use either Premium or Standard storage accounts, but they have very different scalability metrics. Pivotal recommends creating 1 Standard storage account for every 30 VMs, or 1 Premium storage account for every 150 VMs. You can increase the number of storage accounts later by provisioning more and following the naming sequence.

- o To use Premium storage (recommended):

```
$ export STORAGE_TYPE="Premium_LRS"
```

- o To use Standard storage:

```
$ export STORAGE_TYPE="Standard_LRS"
```

8. For each deployment storage account you create, do the following:

- a. Create the storage account with the following command, replacing `MY_DEPLOYMENT_STORAGE_X` with one of your deployment storage account names.

```
$ az storage account create --name MY_DEPLOYMENT_STORAGE_X \
--resource-group $RESOURCE_GROUP --sku $STORAGE_TYPE \
--kind Storage --location $LOCATION
```

If the command fails, try a different set of account names.

- b. Retrieve the connection string for the account.

```
$ az storage account show-connection-string \
--name MY_DEPLOYMENT_STORAGE_X --resource-group $RESOURCE_GROUP
```

The command returns output similar to the following:

```
{
  "connectionString": "DefaultEndpointsProtocol=https;EndpointSuffix=core.windows.net;AccountName=cfdocsdeploystorage1;AccountKey=EXAMPLEaaaaaaaaQjSAmqj1OocsGh
}
```

- c. Record the full value of `connectionString` from the output above, starting with and including `DefaultEndpointsProtocol=`.
- d. Create two blob containers named `bosh` and `stemcell` in the account.

```
$ az storage container create --name bosh \
--connection-string $CONNECTION_STRING
```

```
$ az storage container create --name stemcell \
--connection-string $CONNECTION_STRING
```

Step 3: Create Load Balancers

Your load balancer configuration depends on whether you want apps to be available from public IP addresses, private IP addresses, or both.

To create load balancers, do the following:

1. **Required:** PAS Load Balancer

- a. Create a load balancer named `pcf-lb`. A static IP address will be automatically created for Standard SKU load balancers unless specified otherwise with `--public-ip-address-allocation Dynamic`.

```
$ az network lb create --name pcf-lb \
--resource-group $RESOURCE_GROUP --location $LOCATION \
--backend-pool-name pcf-lb-be-pool --frontend-ip-name pcf-lb-fe-ip \
--public-ip-address pcf-lb-ip --public-ip-address-allocation Static \
--sku Standard
```

 **Note:** If the Standard SKU is not available in Azure China, you can use the Basic SKU.

This back end pool is empty when you create it.

- Add a probe to the load balancer.

```
$ az network lb probe create --lb-name pcf-lb \
--name http8080 --resource-group $RESOURCE_GROUP \
--protocol Http --port 8080 --path health
```

- Add a load balancing rule for HTTP.

```
$ az network lb rule create --lb-name pcf-lb \
--name http --resource-group $RESOURCE_GROUP \
--protocol Tcp --frontend-port 80 \
--backend-port 80 --frontend-ip-name pcf-lb-fe-ip \
--backend-pool-name pcf-lb-be-pool \
--probe-name http8080
```

- Add a load balancing rule for HTTPS.

```
$ az network lb rule create --lb-name pcf-lb \
--name https --resource-group $RESOURCE_GROUP \
--protocol Tcp --frontend-port 443 \
--backend-port 443 --frontend-ip-name pcf-lb-fe-ip \
--backend-pool-name pcf-lb-be-pool \
--probe-name http8080
```

- (Optional) For private IPs, do the following:

- Create a load balancer named `pcf-lb`.

```
$ az network lb create --name pcf-lb \
--resource-group $RESOURCE_GROUP --location $LOCATION \
--backend-pool-name pcf-lb-be-pool --frontend-ip-name pcf-lb-fe-ip \
--private-ip-address 10.0.0.6 --sku Standard
```

- Add a probe to the load balancer.

```
$ az network lb probe create --lb-name pcf-lb \
--name http8080 --resource-group $RESOURCE_GROUP \
--protocol Http --port 8080
```

- Add a load balancing rule for HTTP.

```
$ az network lb rule create --lb-name pcf-lb \
--name http --resource-group $RESOURCE_GROUP \
--protocol Tcp --frontend-port 80 \
--backend-port 80 --frontend-ip-name pcf-lb-fe-ip \
--backend-pool-name pcf-lb-be-pool \
--probe-name http8080
```

- Add a load balancing rule for HTTPS.

```
$ az network lb rule create --lb-name pcf-lb \
--name https --resource-group $RESOURCE_GROUP \
--protocol Tcp --frontend-port 443 \
--backend-port 443 --frontend-ip-name pcf-lb-fe-ip \
--backend-pool-name pcf-lb-be-pool \
--probe-name http8080
```

 **Note:** If the Standard SKU is not available in Azure China, you can use the Basic SKU.

This back end pool is empty when you create it.

f. Navigate to your DNS provider and create an entry that points `*.YOUR-SUBDOMAIN` to the public IP address of your load balancer. For example, create an entry that points `azure.example.com` to `198.51.100.1`. You can retrieve it by running
`az network public-ip show --name pcf-lb-ip --resource-group $RESOURCE_GROUP`.

2. Optional: Diego SSH Load Balancer

- a. Create a load balancer named `pcf-ssh-lb`.

```
$ az network lb create --name pcf-ssh-lb \
--resource-group $RESOURCE_GROUP --location $LOCATION \
--backend-pool-name pcf-ssh-lb-be-pool --frontend-ip-name pcf-ssh-lb-fe-ip \
--public-ip-address pcf-ssh-lb-ip --public-ip-address-allocation Static \
--sku Standard
```

 **Note:** If the Standard SKU is not available in Azure China, you can use the Basic SKU.

This back end pool is empty when you create it.

- b. (Optional) For private IPs, create a load balancer named `pcf-ssh-lb`.

```
$ az network lb create --name pcf-ssh-lb \
--resource-group $RESOURCE_GROUP --location $LOCATION \
--backend-pool-name pcf-ssh-lb-be-pool --frontend-ip-name pcf-ssh-lb-fe-ip \
--private-ip-address 10.0.0.7 \
--sku Standard
```

 **Note:** If the Standard SKU is not available in Azure China, you can change to use the Basic SKU.

This back end pool is empty when you create it.

- c. Add a probe to the load balancer.

```
$ az network lb probe create --lb-name pcf-ssh-lb \
--name tcp2222 --resource-group $RESOURCE_GROUP \
--protocol Tcp --port 2222
```

- d. Add a load balancing rule for SSH.

```
$ az network lb rule create --lb-name pcf-ssh-lb \
--name diego-ssh --resource-group $RESOURCE_GROUP \
--protocol Tcp --frontend-port 2222 \
--backend-port 2222 --frontend-ip-name pcf-ssh-lb-fe-ip \
--backend-pool-name pcf-ssh-lb-be-pool \
--probe-name tcp2222
```

- e. Navigate to your DNS provider, and create an entry that points `ssh.sys.YOUR-SUBDOMAIN` to the public IP address of your load balancer. For example, create an entry that points `azure.example.com` to `198.51.100.1`. You can retrieve it by running

```
az network public-ip show --name pcf-ssh-lb-ip --resource-group $RESOURCE_GROUP .
```

Step 4: Boot Ops Manager

To boot Ops Manager, do the following:

1. Navigate to [Pivotal Network](#) and download the latest release of Pivotal Cloud Foundry Ops Manager for Azure.
2. View the downloaded PDF and locate the Ops Manager image URL appropriate for your region.
3. Export the Ops Manager image URL as an environment variable.

```
$ export OPS_MAN_IMAGE_URL="YOUR-OPS-MAN-IMAGE-URL"
```

4. Download the Ops Manager image. For compatibility when upgrading to future versions of Ops Manager, choose a unique name for the image that includes the Ops Manager version number. For example, replace `opsman-image-2.6.x` in the following examples with `opsman-image-2.6.1`.

- o If you use unmanaged disks, perform the following steps:

 **Note:** Azure Stack requires unmanaged disks.

- Download the Ops Manager image to your local machine. The image size is 10 GB.

```
$ wget $OPS_MAN_IMAGE_URL -O opsman-image-2.6.x.vhd
```

- Upload the image to your storage account using the Azure CLI.

```
$ az storage blob upload --name opsman-image-2.6.x.vhd \
--connection-string $CONNECTION_STRING \
--container-name opsmanager \
--type page \
--file opsman-image-2.6.x.vhd
```

- If you use managed disks, do the following:

- Copy the Ops Manager image into your storage account using the Azure CLI.

```
$ az storage blob copy start --source-uri $OPS_MAN_IMAGE_URL \
--connection-string $CONNECTION_STRING \
--destination-container opsmanager \
--destination-blob opsman-image-2.6.x.vhd
```

- Copying the image may take several minutes. Run the following command and examine the output under "copy":

```
$ az storage blob show --name opsman-image-2.6.x.vhd \
--container-name opsmanager \
--connection-string $CONNECTION_STRING
...
"copy": {
"completionTime": "2017-06-26T22:24:11+00:00",
"id": "b9c8b272-a562-4574-baa6-f1a04afcefd",
"progress": "53687091712/53687091712",
"source": "https://opsmanagerwestus.blob.core.windows.net/images/opsman-image-2.6.x.vhd",
"status": "success",
"statusDescription": null
},
```

- Wait a few moments and re-run the command above if `status` is `pending`. When `status` reads `success`, continue to the next step.

- Create a public IP address named `ops-manager-ip`.

```
$ az network public-ip create --name ops-manager-ip \
--resource-group $RESOURCE_GROUP --location $LOCATION \
--allocation-method Static
{
  "publicIp": {
    "dnsSettings": null,
    "etag": "W/\"4450ebe2-9e97-4b17-9cf2-44838339e661\"",
    "id": "/subscriptions/995b7eed-77ef-45ff-a5c9-1a405ffb8243/resourceGroups/cf-docs/providers/Microsoft.Network/publicIPAddresses/ops-manager-ip",
    "idleTimeoutInMinutes": 4,
    "ipAddress": "40.83.148.183",
    "ipConfiguration": null,
    "location": "westus",
    "name": "ops-manager-ip",
    "provisioningState": "Succeeded",
    "publicIpAddressVersion": "IPv4",
    "publicIpAddressAllocationMethod": "Static",
    "resourceGroup": "cf-docs",
    "resourceGuid": "950d4831-1bec-42da-8a79-959bcddea9dd",
    "tags": null,
    "type": "Microsoft.Network/publicIPAddresses"
  }
}
```

If you do not want to use a public IP for Ops Manager, skip this step.

- Record the `ipAddress` from the output above. This is the public IP address of Ops Manager.

- Create a network interface for Ops Manager.

```
$ az network nic create --vnet-name pcf-virtual-network \
--subnet pcf-infrastructure-subnet --network-security-group opsmgr-nsg \
--private-ip-address 10.0.4.4 \
--public-ip-address ops-manager-ip \
--resource-group $RESOURCE_GROUP \
--name opsman-nic --location $LOCATION
```

If you do not want use a public IP address for Ops Manager, remove the `--public-ip-address ops-manager-ip` flag and value.

- Create a keypair on your local machine with the username `ubuntu`. For example, enter the following command:

```
$ ssh-keygen -t rsa -f opsman -C ubuntu
```

When prompted for a passphrase, press the `enter` key to provide an empty passphrase.

- Create the Ops Manager VM.

- If you are using unmanaged disks, run the following command to create your Ops Manager VM, replacing `PATH-TO-PUBLIC-KEY` with the path to your public key `.pub` file:

```
$ az vm create --name opsman-2.6.x --resource-group $RESOURCE_GROUP \
--location $LOCATION \
--nics opsman-nic \
--image https://$STORAGE_NAME.my-azure-instance.com/opsmanager/opsman-image-2.6.x.vhd \
--os-disk-name opsman-2.6.x-osdisk \
--os-disk-size-gb 128 \
--os-type Linux \
--use-unmanaged-disk \
--storage-account $STORAGE_NAME \
--storage-container-name opsmanager \
--admin-username ubuntu \
--ssh-key-value PATH-TO-PUBLIC-KEY
```

Replace `my-azure-instance.com` with the URL of your Azure instance. Find the complete source URL in the Azure UI by viewing the **Blob properties** of the Ops Manager image you created earlier in this procedure.

- If you are using Azure managed disks, do the following:

- Create a managed image from the Ops Manager VHD file:

```
$ az image create --resource-group $RESOURCE_GROUP \
--name opsman-image-2.6.x \
--source https://$STORAGE_NAME.blob.core.windows.net/opsmanager/image-2.6.x.vhd \
--location $LOCATION \
--os-type Linux
```

If you are using Azure China, Azure Government Cloud, or Azure Germany, replace `blob.core.windows.net` with the following:

- For Azure China, use `blob.core.chinacloudapi.cn`. See the [Azure documentation](#) for more information.
- For Azure Government Cloud, use `blob.core.usgovcloudapi.net`. See the [Azure documentation](#) for more information.
- For Azure Germany, use `blob.core.cloudapi.de`. See the [Azure documentation](#) for more information.

- Create your Ops Manager VM, replacing `PATH-TO-PUBLIC-KEY` with the path to your public key `.pub` file.

```
$ az vm create --name opsman-2.6.x --resource-group $RESOURCE_GROUP \
--location $LOCATION \
--nics opsman-nic \
--image opsman-image-2.6.x \
--os-disk-size-gb 128 \
--os-disk-name opsman-2.6.x-osdisk \
--admin-username ubuntu \
--size Standard_DS2_v2 \
--storage-sku Standard_LRS \
--ssh-key-value PATH-TO-PUBLIC-KEY
```

- If you plan to install more than one tile in this Ops Manager installation, do the following to increase the size of the Ops Manager VM disk. You can repeat this process and increase the disk again at a later time if necessary.

 **Note:** If you use Azure Stack, you must increase the Ops Manager VM disk size using the Azure Stack UI.

- Run the following command to stop the VM and detach the disk:

```
$ az vm deallocate --name opsman-2.6.x \
--resource-group $RESOURCE_GROUP
```

- Run the following command to resize the disk to 128 GB:

```
$ az disk update --size-gb 128 --name opsman-2.6.x-osdisk \
--resource-group $RESOURCE_GROUP
```

- c. Run the following command to start the VM:

```
$ az vm start --name opsman-2.6.x --resource-group $RESOURCE_GROUP
```

Step 5: Complete BOSH Director Configuration

To finish configuring BOSH Director, do the following:

1. Navigate to your DNS provider, and create an entry that points a fully qualified domain name (FQDN) to the public IP address of Ops Manager. As a best practice, always use the FQDN to access Ops Manager.
2. Continue to the [Configuring BOSH Director on Azure Manually](#) topic.

Configuring BOSH Director on Azure Manually

This topic describes how to configure the BOSH Director tile within Ops Manager on Azure after [Deploying Ops Manager on Azure Manually](#).

Prerequisites

See the following sections to prepare for configuring BOSH Director on Azure.

General Prerequisites

Before you perform the procedures in this topic, you must have completed the procedures in the following topics:

- [Preparing to Deploy Ops Manager on Azure Manually](#)
- [Deploying Ops Manager on Azure Manually](#)

 **Note:** You can also perform the procedures in this topic using the Ops Manager API. For more information, see the [Using the Ops Manager API](#) topic.

Prerequisites for Azure Stack or Azure Government Cloud

If you use Azure Government Cloud, you must set custom VM types for a successful deployment. You can set custom VM types with the Ops Manager API. For more information, see [Using the Ops Manager API](#).

Example `curl` command to set a VM types:

```
$ curl -k https://YOUR-OPS-MAN-FQDN/api/v0/vm_types -X  
PUT -H "Authorization: bearer UAA-ACCESS-TOKEN"  
"Content-Type: application/json" -d '{  
  "vm_types": [  
    {"name": "Standard_D1_v2", "ram": 3584, "cpu": 1, "ephemeral_disk": 51200},  
    {"name": "Standard_D2_v2", "ram": 7168, "cpu": 2, "ephemeral_disk": 102400},  
    {"name": "Standard_D3_v2", "ram": 14336, "cpu": 4, "ephemeral_disk": 204800},  
    {"name": "Standard_D4_v2", "ram": 28672, "cpu": 8, "ephemeral_disk": 409600},  
    {"name": "Standard_D5_v2", "ram": 57344, "cpu": 8, "ephemeral_disk": 819200},  
    {"name": "Standard_D11_v2", "ram": 14336, "cpu": 2, "ephemeral_disk": 102400},  
    {"name": "Standard_D12_v2", "ram": 28672, "cpu": 4, "ephemeral_disk": 204800},  
    {"name": "Standard_D13_v2", "ram": 57344, "cpu": 8, "ephemeral_disk": 409600},  
    {"name": "Standard_D14_v2", "ram": 114688, "cpu": 16, "ephemeral_disk": 819200},  
    {"name": "Standard_F1", "ram": 2048, "cpu": 1, "ephemeral_disk": 16384},  
    {"name": "Standard_F2", "ram": 4096, "cpu": 2, "ephemeral_disk": 32768},  
    {"name": "Standard_F4", "ram": 8192, "cpu": 4, "ephemeral_disk": 65536},  
    {"name": "Standard_F8", "ram": 16384, "cpu": 8, "ephemeral_disk": 131072}]}
```

To verify that you have successfully set your custom VM types, use the `GET /api/v0/vm_types` Ops Manager API endpoint.

Step 1: Access Ops Manager

1. In a web browser, navigate to the fully qualified domain name (FQDN) of Ops Manager that you created in [Deploying Ops Manager on Azure Manually](#).
2. When Ops Manager starts for the first time, you must choose one of the following:
 - [Internal Authentication](#): If you use Internal Authentication, Ops Manager maintains your user database.
 - [SAML Identity Provider](#): If you use a SAML Identity Provider (IdP), an external identity server maintains your user database.
 - [LDAP Server](#): If you use a LDAP Server, an external identity server maintains your user database.

Internal Authentication

1. When redirected to the [Internal Authentication](#) page, do the following:

- Enter a **Username**, **Password**, and **Password confirmation** to create an Admin user.
- Enter a **Decryption passphrase** and the **Decryption passphrase confirmation**. This passphrase encrypts the Ops Manager datastore, and is not recoverable.
- If you are using an **HTTP proxy** or **HTTPS proxy**, follow the instructions in [Configuring Proxy Settings for the BOSH CPI](#).
- Read the **End User License Agreement**, and select the checkbox to accept the terms.
- Click **Setup Authentication**.

2. Log in to Ops Manager with the Admin username and password you created in the previous step.

SAML Identity Provider

1. Log in to your IdP console and download the IdP metadata XML. Optionally, if your IdP supports metadata URL, you can copy the metadata URL instead of the XML.

2. Copy the IdP metadata XML or URL to the Ops Manager **SAML Identity Provider** login page.

Note: The same IdP metadata URL or XML is applied for the BOSH Director. If you use a separate IdP for BOSH, copy the metadata XML or URL from that IdP and enter it into the BOSH IdP Metadata text box in the Ops Manager login page.

3. Enter values for the fields listed below. Failure to provide values in these fields results in a **500** error.

- **SAML admin group:** Enter the name of the SAML group that contains all Ops Manager administrators. This field is case-sensitive.
- **SAML groups attribute:** Enter the groups attribute tag name with which you configured the SAML server. This field is case-sensitive.

4. Enter your **Decryption passphrase**. Read the **End User License Agreement**, and select the checkbox to accept the terms.

5. Your Ops Manager login page appears. Enter your username and password. Click **Login**.

6. Download your SAML Service Provider metadata (SAML Relying Party metadata) by navigating to the following URLs:

- **6a.** Ops Manager SAML service provider metadata: <https://OPS-MAN-FQDN:443/uaa/saml/metadata>
- **6b.** BOSH Director SAML service provider metadata: <https://BOSH-IP-ADDRESS:8443/saml/metadata>

Note: To retrieve your **BOSH-IP-ADDRESS**, navigate to the **Status** tab in the **BOSH Director** tile. Record the **BOSH Director** IP address.

7. Configure your IdP with your SAML Service Provider metadata. Import the Ops Manager SAML provider metadata from Step 6a above to your IdP. If your IdP does not support importing, provide the values below.

- **Single sign on URL:** <https://OPS-MAN-FQDN:443/uaa/saml/SSO/alias/OPS-MAN-FQDN>
- **Audience URI (SP Entity ID):** <https://OP-MAN-FQDN:443/uaa>
- **Name ID:** Email Address
- SAML authentication requests are always signed

8. Import the BOSH Director SAML provider metadata from Step 6b to your IdP. If the IdP does not support an import, provide the values below.

- **Single sign on URL:** <https://BOSH-IP:8443/saml/SSO/alias/BOSH-IP>
- **Audience URI (SP Entity ID):** <https://BOSH-IP:8443>
- **Name ID:** Email Address
- SAML authentication requests are always signed

9. Return to the **BOSH Director** tile, and continue with the configuration steps below.

LDAP Server

- For **Server URL**, enter the URL that points to your LDAP server. With multiple LDAP servers, separate their URLs with spaces. Each URL must include one of the following protocols:
 - **ldap://**: This specifies that the LDAP server uses an unencrypted connection.
 - **ldaps://**: This specifies that the LDAP server uses SSL for an encrypted connection and requires that the LDAP server holds a trusted certificate or that you import a trusted certificate to the JVM truststore.
- For **LDAP Username** and **LDAP Password**, enter the LDAP Distinguished Name (DN) and the password for binding to the LDAP Server. Example DN: <cn=administrator,ou=Users,dc=example,dc=com>

Note: Pivotal recommends that you provide LDAP credentials that grant read-only permissions on the LDAP Search Base and the LDAP Group Search Base. In addition to this, if the bind user belongs to a different search base, you must use the full DN.

warning: Pivotal recommends against reusing LDAP service accounts across environments. LDAP service accounts should not be subject to manual lockouts, such as lockouts that result from users utilizing the same account. Also, LDAP service accounts should not be subject to automated deletions, since disruption to these service accounts could prevent user logins.

- For **User Search Base**, enter the location in the LDAP directory tree from which any LDAP User search begins. The typical LDAP Search Base matches your domain name.

For example, a domain named “cloud.example.com” typically uses the following LDAP User Search Base: `ou=Users,dc=example,dc=com`

- For **User Search Filter**, enter a string that defines LDAP User search criteria. These search criteria allow LDAP to perform more effective and efficient searches. For example, the standard LDAP search filter `cn=Smith` returns all objects with a common name equal to `Smith`.

In the LDAP search filter string that you use to configure your runtime, use `{0}` instead of the username. For example, use `cn={0}` to return all LDAP objects with the same common name as the username.

In addition to `cn`, other attributes commonly searched for and returned are `mail`, `uid` and, in the case of Active Directory, `sAMAccountName`.

Note: For instructions for testing and troubleshooting your LDAP search filters, see [Configuring LDAP Integration with Pivotal Cloud Foundry](#) in the Pivotal Support Knowledge Base.

- For **Group Search Base**, enter the location in the LDAP directory tree from which the LDAP Group search begins.

For example, a domain named “cloud.example.com” typically uses the following LDAP Group Search Base: `ou=Groups,dc=example,dc=com`

- For **Group Search Filter**, enter a string that defines LDAP Group search criteria. The standard value is `member={0}`.
- For **Email Attribute**, enter the attribute name in your LDAP directory that corresponds to the email address in each user record, for example `mail`.
- For **LDAP RBAC Admin Group Name**, enter the DN of the LDAP group you want to have admin permissions in Ops Manager.
- From the dropdown, select how the UAA handles LDAP server referrals out to other external user stores. The UAA can:
 - Automatically follow any referrals.
 - Ignore referrals and return partial result.
 - Throw exception for each referral and abort.
- For **Server SSL Cert**, paste in the root certificate from your CA certificate or your self-signed certificate.
- Enter a **Decryption passphrase** and the **Decryption passphrase confirmation**. This passphrase encrypts the Ops Manager datastore, and is not recoverable.
- If you are using an **HTTP proxy** or **HTTPS proxy**, follow the instructions in [Configuring Proxy Settings for the BOSH CPI](#).
- Read the **End User License Agreement**, and select the checkbox to accept the terms.
- Select **Provision an admin client in the BOSH UAA**. You can use this to enable BOSH automation with scripts and tooling. For more information, see [Provision Admin Client](#) in *Creating UAA Clients for BOSH Director*.
- Click **Setup Authentication**.
- Return to the **BOSH Director** tile, and continue with the configuration steps below.

Step 2: Azure Config Page

- Click the **BOSH Director** tile.
- Select **Azure Config**.

BOSH Director for Azure

Settings Status Credentials

Azure Config

Director Config

Create Networks

Assign AZs and Networks

Security

BOSH DNS Config

Syslog

Resource Config

Azure Config

Subscription ID*

Tenant ID*

Application ID*

Client Secret* [Change](#)

Resource Group Name*

BOSH Storage Account Name*

Cloud Storage Type
 Use Managed Disks
 Use Storage Accounts

Storage Account Type

Deployments Storage Account Name

Default Security Group

SSH Public Key*

SSH Private Key*

Availability Mode
 Availability Sets
 Availability Zones

Azure Environment
 Azure Commercial Cloud
 Azure US Gov Cloud
 Azure Germany Cloud
 Azure China Cloud
 Azure Stack

Tenant Management Resource Endpoint

Domain*

Authentication*

Azure Stack Endpoint Prefix*

Azure Stack CA Certificate*

[Save](#)

3. Complete the following fields with information you obtained in [Preparing to Deploy Ops Manager on Azure Manually](#).

- **Subscription ID:** Enter the ID of your Azure subscription.
- **Tenant ID:** Enter your `TENANT_ID`.
- **Application ID:** Enter the `APPLICATION_ID` that you created in the [Create an AAD Application](#) step of *Preparing to Deploy Ops Manager on Azure Manually*.
- **Client Secret:** Enter your `CLIENT_SECRET`.

4. Complete the following fields:

- **Resource Group Name:** Enter the name of your resource group, which you exported as the `$RESOURCE_GROUP` environment variable.
- **BOSH Storage Account Name:** Enter the name of your storage account, which you exported as the `$STORAGE_NAME` environment variable.

5. For **Cloud Storage Type**, select one of the following options based on your Azure VM storage settings.

- **Use Managed Disks:** Select this option if you use Azure Managed Disks. See [Azure Managed Disks Overview](#) in the Microsoft documentation for more information. For **Storage Account Type**, select the storage option that corresponds with your Azure subscription. Select **Standard_LRS** for HDD-based storage or **Premium_LRS** for SSD-based storage.

Cloud Storage Type

Use Managed Disks

Storage Account Type

Premium_LRS Storage Account Type to use with managed disks

Use Storage Accounts

- **Use Storage Accounts:** Select this option if you use storage accounts to store your Azure VMs. Enter the base storage name that you used to create your deployment storage accounts, prepended and appended with the wildcard character `*`. For example, if you created accounts named `xyzdeploymentstorage1`, `xyzdeploymentstorage2`, and `xyzdeploymentstorage3`, enter `*deploymentstorage*`. Ops Manager requires that you specify an asterisk at both the beginning and the end of the base storage account name.

Use Storage Accounts

Deployments Storage Account Name

A Storage Account into which Ops Manager will deploy other VMs. Name must contain only lower-case letters and numbers, may start AND end with an asterisk. Required if Cloud Storage Type is set to "Storage Accounts".

⚠ warning: You can update your deployment from using storage accounts to using managed disks. However, after you select **Use Managed Disks** and deploy Ops Manager, you cannot change your deployment back to use storage accounts.

6. (Optional) Enter your **Default Security Group**. For more information about Network Security Groups, see [Filter network traffic with network security groups](#) in Microsoft Azure's documentation.

Note: The Azure portal sometimes displays the names of resources with incorrect capitalization. Always use the Azure CLI to retrieve the correctly capitalized name of a resource.

7. For **SSH Public Key**, copy and paste the contents of your public key in the `opsman.pub` file. You created this file in [Deploying Ops Manager on Azure Manually](#).

8. For **SSH Private Key**, copy and paste the contents of your private key in the `opsman` file.

9. For **Availability Mode**, choose whether to use **Availability Sets** or **Availability Zones**. You can further configure this option in the **Assign AZs and Networks** pane. For more information, see [Assign AZs and Networks](#).

- **For new installations of Ops Manager v2.5:** Availability Zones are selected by default, but you can choose Availability Sets.
- **If you are upgrading from v2.4:** You must use Availability Sets.

Note: You can only choose between Availability Zones and Availability Sets in a staged BOSH Director tile that has not yet deployed. After you click **Apply Changes** and the BOSH Director deploys successfully, you cannot switch between Availability Zones and Availability Sets.

10. For **Azure Environment**, select the Azure cloud you want to use. Pivotal recommends **Azure Commercial Cloud** for most Azure environments.

Note: If you selected Azure Stack or Azure Government Cloud as your Azure environment, you must set custom VM types. For more

information, see [Prerequisites for Azure Stack or Azure Government Cloud](#) in *Configuring BOSH Director on Azure Manually*.

Note: To deploy PCF on Azure using Active Directory Federation Services (ADFS), select **Azure Stack**. For more information about enabling ADFS for Azure, see [ADFS Authentication Available for Microsoft Azure](#).

11. (Optional) If you selected **Azure Stack** for your **Azure Environment**, complete the following Azure Stack-only fields:
 - a. For **Domain**, enter the domain for your Azure Stack deployment. For example, `local.azurestack.external`.
 - b. Enter the **Tenant Management Resource Endpoint** from your AzureRM Environment Context. This URL changes with every ASDK installation.
 - c. Enter `AzureAD` for **Authentication**. If you want to use Azure China Cloud Active Directory for your authentication type, enter `AzureChinaCloudAD`.
 - d. Enter your **Azure Stack Endpoint Prefix**. For example `management`.
 - e. Enter your **Azure Stack CA Certificate**. Azure Stack requires a custom CA certificate. Copy the certificate from your Azure Stack environment.
12. Click **Save**.

Step 3: Director Config Page

1. In Ops Manager, select **Director Config**.

NTP Servers (comma delimited)*

Bosh HM Forwarder IP Address

Enable VM Resurrector Plugin

Enable Post Deploy Scripts

Recreate All VMs
This will force BOSH to recreate all VMs on the next deploy. Persistent disk will be preserved.

Recreate All Persistent Disks
Check this box to recreate all Persistent Disks for the Director and all other Tiles.

Enable bosh deploy retries
This will attempt to re-deploy a failed deployment up to 5 times.

2. In the **NTP Servers (comma delimited)** field, enter a comma-separated list of valid NTP servers.

Note: The NTP server configuration only updates after VM recreation. Ensure that you select the **Recreate all VMs** checkbox if you modify the value of this field.

3. Leave the **Bosh HM Forwarder IP Address** field blank.

Note: Starting in PAS v2.0, BOSH-reported component metrics are available in the Loggregator Firehose by default. If you continue to use the BOSH HM Forwarder to consume these component metrics, you may receive duplicate data. To prevent this, leave the **Bosh HM Forwarder IP Address** field blank.

4. Select the **Enable VM Resurrector Plugin** checkbox to enable the BOSH Resurrector functionality and increase your runtime availability.
5. Select **Enable Post Deploy Scripts** to run a post-deploy script after deployment. This script allows the job to execute additional commands against a deployment.

Note: If you intend to install PKS, you must enable post-deploy scripts.

6. Select **Recreate all VMs** to force BOSH to recreate all VMs on the next deploy. This process does not destroy any persistent disk data.
7. Select **Recreate All Persistent Disks** to force BOSH to migrate and recreate persistent disks for the BOSH Director and all tiles. This process does not destroy any persistent disk data.
8. Select **Enable bash deploy retries** to instruct Ops Manager to retry failed BOSH operations up to five times.
9. Select **Skip Director Drain Lifecycle** to prevent drain scripts from running when the BOSH Director is recreated.
10. Select **Store BOSH Job Credentials on tmpfs (beta)** to store credentials for BOSH jobs on temporary file storage (tmpfs) memory, rather than on disk. You must recreate all VMs for this setting to take effect.
11. Select **Keep Unreachable Director VMs** if you want to preserve BOSH Director VMs after a failed deployment for troubleshooting purposes.
12. Select **HM Pager Duty Plugin** to enable Health Monitor integration with PagerDuty.

HM Pager Duty Plugin

Service Key*

YOUR-PAGERDUTY-SERVICE-KEY

HTTP Proxy

YOUR-HTTP-PROXY

- **Service Key:** Enter your API service key from PagerDuty.
- **HTTP Proxy:** Enter an HTTP proxy for use with PagerDuty.

13. Select **HM Email Plugin** to enable Health Monitor integration with email.

HM Email Plugin

Host*

smtp.example.com

Port*

25

Domain*

cloudfoundry.example.com

From*

user2@example.com

Recipients*

user@example.com, user1@example.com

Username

user

Password

Enable TLS

- **Host:** Enter your email hostname.
- **Port:** Enter your email port number.
- **Domain:** Enter your domain.
- **From:** Enter the address for the sender.

- **Recipients:** Enter comma-separated addresses of intended recipients.
- **Username:** Enter the username for your email server.
- **Password:** Enter the password for your email server.
- **Enable TLS:** Select this checkbox to enable Transport Layer Security to the email host.

14. For **CredHub Encryption Provider**, you can choose whether BOSH CredHub stores its encryption key internally on the BOSH Director and CredHub VM, or in an external hardware security module (HSM). The HSM option is more secure.

Before configuring an HSM encryption provider in the **Director Config** pane, you must follow the procedures and collect information described in [Preparing CredHub HSMs for Configuration](#).

 **Note:** After you deploy Ops Manager with an HSM encryption provider, you cannot change BOSH CredHub to store encryption keys internally.

CredHub Encryption Provider

Internal
 Luna HSM

Encryption Key Name*

Provider Partition*

Provider Partition Password*

Provider Client Certificate*

Provider Client Certificate Private Key*

HSM Host Address*

HSM Port Address*

Partition Serial Number*

HSM Certificate*

- **Internal:** Select this option for internal CredHub key storage. This option is selected by default and requires no additional configuration.
- **Luna HSM:** Select this option to use a SafeNet Luna HSM as your permanent CredHub encryption provider, and fill in the following fields:

1. **Encryption Key Name:** Any name to identify the key that the HSM uses to encrypt and decrypt the CredHub data. Changing this key name after you deploy Ops Manager can cause service downtime.
 2. **Provider Partition:** The partition that stores your encryption key. Changing this partition after you deploy Ops Manager could cause service downtime. For this value and the ones below, use values gathered in [Preparing CredHub HSMs for Configuration](#).
 3. **Provider Partition Password**
 4. **Provider Client Certificate:** The certificate that validates the identity of the HSM when CredHub connects as a client.
 5. **Provider Client Certificate Private Key**
 6. **HSM Host Address**
 7. **HSM Port Address:** If you do not know your port address, enter **1792**.
 8. **Partition Serial Number**
 9. **HSM Certificate:** The certificate that the HSM presents to CredHub to establish a two-way mTLS connection.
15. Select a **Blobstore Location** to either configure the blobstore as an internal server or an external endpoint. Because the internal server is unscalable and less secure, Pivotal recommends that you configure an external blobstore.

 **Note:** After you deploy Ops Manager, you cannot change the blobstore location.

- **Internal:** Select this option to use an internal blobstore. Ops Manager creates a new VM for blob storage. No additional configuration is required.
 - **Enable TLS:** Select this checkbox to enable TLS to the blobstore.

 **Note:** If you are using PASW 2016, make sure you have downloaded [Windows stemcell v1709.10 or higher](#) before enabling TLS.

- **S3 Compatible Blobstore:** Select this option to use an external S3-compatible endpoint.

Blobstore Location

Internal

Enable TLS

S3 Compatible Blobstore

S3 Endpoint*

Bucket Name*

Access Key*

Secret Key*

S3 Signature Version

V2 Signature

V4 Signature

Region*

S3 Backup Strategy

No backups

Use a versioned bucket

Copy into an additional bucket

Backup Bucket Region*

Backup Bucket Name*

Follow the procedures in [Sign up for Amazon S3](#) and [Creating a Bucket](#) in the AWS documentation. When you have created an S3 bucket, complete the following steps:

1. **S3 Endpoint:** Navigate to the [Regions and Endpoints](#) topic in the AWS documentation.
 - a. Locate the endpoint for your region in the **Amazon Simple Storage Service (S3)** table and construct a URL using your region's endpoint. For example, if you are using the `us-west-2` region, the URL you create would be `https://s3-us-west-2.amazonaws.com`. Enter this URL into the **S3 Endpoint** field.
 - b. On a command line, run `ssh ubuntu@OPS-MANAGER-FQDN` to SSH into the Ops Manager VM. Replace `OPS-MANAGER-FQDN` with the fully qualified domain name of Ops Manager.
 - c. Copy the custom public CA certificate you used to sign the S3 endpoint into `/etc/ssl/certs` on the Ops Manager VM.
 - d. On the Ops Manager VM, run `sudo update-ca-certificates -f -v` to import the custom CA certificate into the Ops Manager VM truststore.

Note: You must also add this custom CA certificate into the **Trusted Certificates** field in the [Security](#) page. See [Security Page](#) for instructions.

2. **Bucket Name:** Enter the name of the S3 bucket.
3. **Access Key** and **Secret Key:** Enter the keys you generated when creating your S3 bucket.
4. Select **V2 Signature** or **V4 Signature**. If you select **V4 Signature**, enter your **Region**.

Note: AWS recommends using Signature Version 4. For more information about AWS S3 Signatures, see [Authenticating Requests](#)

[🔗](#) in the AWS documentation.

5. **S3 Backup Strategy:** Select one of the following to configure whether and how to back up the BOSH Director's S3 blobstore:

- **No backups**
- **Use a versioned bucket:** This option uses blobstore bucket versioning to save backups. It requires an S3 blobstore that supports versioning.
- **Copy into an additional bucket:** This option saves blobstore backups to a separate bucket, which is useful with blobstores that do not support versioning. The option requires you to fill in **Backup Bucket Region** and **Backup Bucket Name**.

- **GCS Blobstore:** Select this option to use an external Google Cloud Storage (GCS) endpoint.

GCS Blobstore

Bucket Name*

Storage Class*

Regional

Service Account Key*

Follow the procedures in [Creating Storage Buckets](#) in the GCS documentation to create a GCS bucket. To create a GCS bucket, you must have a GCS account. When you have created a GCS bucket, complete the following steps:

1. **Bucket Name:** Enter the name of your GCS bucket.
2. **Storage Class:** Select the storage class for your GCS bucket. See [Storage Classes](#) in the GCP documentation for more information.
3. **Service Account Key:** Follow the steps in the [Set up IAM Service Accounts](#) section of *Preparing to Deploy Ops Manager on GCP Manually* to download a JSON file with a private key. Enter the contents of the JSON file into the field.

16. Select a **Database Location**. By default, Ops Manager deploys and manages an **Internal** database for you. If you choose to use an **External MySQL Database**, complete the associated fields with information obtained from your external MySQL Database provider: **Host**, **Port**, **Username**,

Database Location

Internal

External MySQL Database

Host*

Port*

Username*

Password*

Database*

Password, and Database

selected the **Enable TLS for Director Database** checkbox, you can complete the following optional fields:

In addition, if you

- **Enable TLS:** Select this checkbox enables TLS communication between the BOSH Director and the database.
- **TLS CA:** Enter the Certificate Authority for the TLS Certificate.

- **TLS Certificate:** Enter the client certificate for mutual TLS connections to the database.
- **TLS Private Key:** Enter the client private key for mutual TLS connections to the database.
- **Advanced DB Connection Options:** If you would like to provide additional options for the database, use this field to provide a JSON-formatted options string.

17. (Optional) Modify the **Director Workers** value, which sets the number of workers available to execute Director tasks. This field defaults to .
18. (Optional) **Max Threads** sets the maximum number of threads that the BOSH Director can run simultaneously. Pivotal recommends that you leave the field blank to use the default value, unless doing so results in rate limiting or errors on your IaaS.
19. (Optional) To add a custom URL for your BOSH Director, enter a valid hostname in **Director Hostname**. You can also use this field to configure a load balancer in front of your BOSH Director. For more information, see [How to Set Up a Load Balancer in Front of Operations Manager Director](#) in the Pivotal Support Knowledge Base.

The screenshot shows a configuration form for a BOSH Director. It includes three input fields: 'Director Workers' with the value '5', 'Max Threads' (empty), and 'Director Hostname' with the value 'opsdirector.example.com'.

20. (Optional) To set a custom banner that users see when logging in to the Director using SSH, enter text in the **Custom SSH Banner** field.
21. (Optional) Enter your comma-separated custom **Identification Tags**. For example, `iaas:foundation1, hello:world`. You can use the tags to identify your foundation when viewing VMs or disks from your IaaS.

Note: Azure has limited space for identification tags. Pivotal recommends entering no more than five tags.

22. Click **Save**.

Step 4: Create Networks Page

Create the following three networks:

- [Infrastructure Network](#)
- [Runtime Network](#)
- [Services Network](#)

Note: The Azure portal sometimes displays the names of resources with incorrect capitalization. Always use the Azure CLI to retrieve the correctly capitalized name of a resource.

Infrastructure Network

To create the infrastructure network, do the following:

1. Select **Create Networks**.
2. (Optional) Select **Enable ICMP checks** if you want to enable ICMP on your networks. Ops Manager uses ICMP checks to confirm that components within your network are reachable.
3. Click **Add Network**. For name, enter `infrastructure`.
4. Under **Subnets**, create a subnet using the information in the following table:

Network Name	infrastructure
Azure Network Name	Enter <code>pcf-virtual-network/pcf-infrastructure-subnet</code> . You can use either the <code>NETWORK-NAME/SUBNET-NAME</code> format or the <code>RESOURCE-GROUP/NETWORK-NAME/SUBNET-NAME</code> format. If you specify a resource group, it must exist under the same subscription ID that you provided in the Azure Config Page .
CIDR	Enter <code>10.0.4.0/26</code> .
Reserved IP Ranges	Enter the first nine IP addresses of the subnet. For example, <code>10.0.4.1-10.0.4.9</code> .
DNS	Enter <code>168.63.129.16</code> .
Gateway	Enter the first IP address of the subnet. For example, <code>10.0.4.1</code> .

Create Networks

Warning: Pivotal recommends keeping the IP settings throughout the life of your installation. Ops Manager may prevent you from changing them in the future.
Contact Pivotal support for help completing such a change.

Verification Settings

Enable ICMP checks

Networks

Add Network

One or many IP ranges upon which your products will be deployed

▼ infrastructure

Name*	<input type="text" value="infrastructure"/>	
Subnets	Add Subnet	
Azure Network Name*	<input type="text" value="pcf-virtual-network/pcf-infrastructure-subnet"/>	
CIDR*	<input type="text" value="10.0.4.0/26"/>	
Reserved IP Ranges	<input type="text" value="10.0.4.1-10.0.4.9"/>	
DNS*	<input type="text" value="168.63.129.16"/>	
Gateway*	<input type="text" value="10.0.4.1"/>	

5. Click **Save**.

Note: After you deploy Ops Manager, you add subnets with overlapping Availability Zones to expand your network. For more information about configuring additional subnets, see [Expanding Your Network with Additional Subnets](#).

Runtime Network

To create the runtime network, do the following:

1. Click **Add Network**. For **Name**, enter the name of the runtime that you intend to deploy in this environment. For example, `pas` or `pks`.

2. Under **Subnets**, create a subnet using the information in the following table:

Network Name	<code>pas</code> or <code>pks</code>
Azure Network Name	Enter <code>pcf-virtual-network/pcf-pas-subnet</code> or <code>pcf-virtual-network/pcf-pks-subnet</code> , depending on your runtime. You can use either the <code>NETWORK-NAME/SUBNET-NAME</code> format or the <code>RESOURCE-GROUP/NETWORK-NAME/SUBNET-NAME</code> format. If you specify a resource group, it must exist under the same subscription ID that you provided in the Azure Config Page .
CIDR	Enter <code>10.0.12.0/22</code> .
Reserved IP Ranges	Enter the first nine IP addresses of the subnet. For example, <code>10.0.12.1-10.0.12.9</code> .
DNS	Enter <code>168.63.129.16</code> .
Gateway	Enter the first IP address of the subnet. For example, <code>10.0.12.1</code> .

The screenshot shows the 'Add Network' configuration dialog for the 'pas' runtime. The 'Name' field is set to 'pas'. Under the 'Subnets' section, the 'Azure Network Name' is set to 'pcf-virtual-network/pcf-pas-subnet'. The 'CIDR' field contains '10.0.12.0/22'. The 'Reserved IP Ranges' field contains '10.0.12.1-10.0.12.9'. The 'DNS' field contains '168.63.129.16'. The 'Gateway' field contains '10.0.12.1'. An 'Add Subnet' button is visible in the top right corner of the subnet configuration area.

3. Click **Save**.

Services Network

To create the services network, do the following:

1. Click **Add Network**. For **Name**, enter `services`.

2. Under **Subnets**, create a subnet using the information in the following table:

Network Name	<code>services</code>
Azure Network Name	Enter <code>pcf-virtual-network/pcf-services-subnet</code> . You can use either the <code>NETWORK-NAME/SUBNET-NAME</code> format or the <code>RESOURCE-GROUP/NETWORK-NAME/SUBNET-NAME</code> format. If you specify a resource group, it must exist under the same subscription ID that you provided in the Azure Config Page .
CIDR	Enter <code>10.0.8.0/22</code> .
Reserved IP Ranges	Enter the first nine IP addresses of the subnet. For example, <code>10.0.8.1-10.0.8.9</code> .
DNS	Enter <code>168.63.129.16</code> .

Gateway | Enter the first IP address of the subnet. For example, `10.0.8.1`.

Name*
services

Subnets

Azure Network Name*
pcf-virtual-network/pcf-services-subnet

CIDR*
10.0.8.0/22

Reserved IP Ranges
10.0.8.1-10.0.8.9

DNS*
168.63.129.16

Gateway*
10.0.8.1

Add Subnet

- Click **Save**. If **Enable ICMP checks** is not selected, you may see red warnings. You can ignore these warnings.

Step 5: Assign AZs and Networks

- Select **Assign AZs and Networks**.

Note: The appearance of this page changes based on whether you are using Availability Zones or Availability Sets. For more information, see [Azure Config Page](#).

- If you use Availability Zones, select one of the three available zones from the **Singleton Availability Zone** dropdown, then select a network from the **Networks** dropdown. The networks that appear in this list are networks that you created earlier in this procedure.
- If you use Availability Sets, select a network from the **Networks** dropdown. The networks that appear in this list are networks that you created earlier in this procedure.

- Click **Save**.

Step 6: Security Page

- Select **Security**.

Security

Trusted Certificates

These certificates enable BOSH-deployed components to trust a custom root certificate.

Include OpsManager Root CA in Trusted Certs

Checking this box will place the OpsManager Root CA into the trusted certificate field. The Bosh Director will then place this CA into the trust store of every VM that it deploys.

Generate VM passwords or use single password for all VMs

Generate passwords

Use default BOSH password

Save

2. In **Trusted Certificates**, enter your custom certificate authority (CA) certificates to insert into your organization's certificate trust chain. This feature enables all BOSH-deployed components in your deployment to trust custom root certificates.

To enter multiple certificates, paste your certificates one after the other. For example, format your certificates like the following:

```
-----BEGIN CERTIFICATE-----
ABCDEFGH12345678ABCDEFGH12345678ABCDEFGH12345678AB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
BCDEFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
CDEFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABBB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
```

Note: If you want to use Docker registries to run PAS app instances in Docker containers, enter the certificate for your private Docker registry in this field. See [Using Docker Registries](#) for more information about running app instances in PAS using Docker registries.

3. (Optional) Select the **Include OpsManager Root CA in Trusted Certs** checkbox to include the Ops Manager root CA in the Trusted Certificates field. BOSH Director includes this CA in the trust store of every VM that it deploys.
4. Choose **Generate passwords** or **Use default BOSH password**. Pivotal recommends that you use the **Generate passwords** option for greater security.
5. Click **Save**. To view your saved Director password, click the **Credentials** tab.

Step 7: BOSH DNS Config Page

BOSH DNS Config

Excluded Recursors

Recurser Timeout

Handlers

[]

Enter a JSON-formatted list of optional custom domain handlers

Save

1. Select **BOSH DNS Config**.
2. (Optional) In **Excluded Recursors**, enter a list of prohibited recursor addresses.
3. (Optional) In **Recurser Timeout**, enter a time limit for contacting the connected recursors. This includes dialing, writing, and reading from the recursor. If any of these actions exceeds the time limit you set, the action fails.

Note: This time limit must include one of the Go parse duration time units. For example, entering `5s` sets the timeout limit to five seconds. For more information about supported time units, see [func ParseDuration](#) in the Go Programming Language documentation.

4. (Optional) In **Handlers**, enter a list of custom domain handlers in JSON format.
5. Click **Save**.

Step 8: Syslog Page

Settings

Syslog

Syslog settings make take up to 1 min to take effect.

Do you want to configure Syslog for OpsManager?

No
 Yes

Address*
 The address or host for the syslog server

Port*

Transport Protocol*

Enable TLS

Permitted Peer*

SSL Certificate*

Queue Size

Forward Debug Logs

Custom rsyslog Configuration

SAVE SYSLOG SETTINGS

1. Select **Syslog**.
2. (Optional) Select **Yes** to send BOSH Director system logs to a remote server.
3. Enter the IP address or DNS name for the remote server in **Address**.
4. Enter the port number that the remote server listens on in **Port**.
5. Select **TCP** or **UDP** from the **Transport Protocol** dropdown. This selection determines which transport protocol is used to send the logs to the remote server.
6. (Optional) Select the **Enable TLS** checkbox to send encrypted logs to remote server with TLS. After you select the checkbox, do the following:
 - a. Enter either the name or SHA1 fingerprint of the remote peer in **Permitted Peer**.
 - b. Enter the SSL certificate for the remote server in **SSL Certificate**.



Note: Pivotal strongly recommends that you enable TLS encryption when you are forwarding logs. Logs can contain sensitive

information, such as cloud provider credentials.

7. (Optional) Enter an integer in **Queue Size**. This value specifies the number of log messages held in the buffer. The default value is 100,000.
8. (Optional) Select the checkbox to **Forward Debug Logs** to an external source. This option is deselected by default. If you select it, you may generate a large amount of log data.
9. (Optional) Enter configuration details for rsyslog in the **Custom rsyslog Configuration** field. This field requires the rainerscript syntax.
10. Click **Save Syslog Settings**.

Step 9: Resource Config Page

1. Select **Resource Config**.

JOB	INSTANCES	PERSISTENT DISK TYPE	VM TYPE	LOAD BALANCERS	INTERNET CONNECTED
Ops Manager Director	Automatic: 1	Automatic: 50 GB	Automatic: Standard_DS2_v2 (cpu: 2, ram: 8 G)	<input type="checkbox"/>	<input type="checkbox"/>
Master Compilation Job	Automatic: 4	None	Automatic: Standard_F4s (cpu: 4, ram: 8 G)	<input type="checkbox"/>	<input type="checkbox"/>

Save

2. Ensure that the **Internet Connected** checkboxes are deselected for all jobs.
3. Adjust any values as necessary for your deployment. Under the **Instances**, **Persistent Disk Type**, and **VM Type** fields, choose **Automatic** from the dropdown to allocate the recommended resources for the job. If the **Persistent Disk Type** field reads **None**, the job does not require persistent disk space.

Note: Ops Manager requires a Director VM with at least 8 GB memory.

Note: If you set a field to **Automatic** and the recommended resource allocation changes in a future version, Ops Manager automatically uses the updated recommended allocation.

Note: If you install PAS for Windows, provision your **Master Compilation Job** with at least 100 GB of disk space.

4. (Optional) For **Load Balancers**, enter your Azure application gateway name for each associated job. To create an application gateway, follow the procedures in [Configure an application gateway for SSL offload by using Azure Resource Manager](#) from the Azure documentation. When you create the application gateway, associate the gateway's IP address with your system domain.

warning: This feature is not recommended for production use. The Azure load balancer does not support an override port in the healthcheck configuration.

5. Click **Save**.

Step 10: (Optional) Add Custom VM Extensions

Use the Ops Manager API to add custom properties to your VMs such as associated security groups and load balancers. For more information, see [Managing Custom VM Extensions](#).

Step 11: Complete the BOSH Director Installation

1. Click **Review Pending Changes**, then **Apply Changes**. If a red ICMP error message appears and you have disabled ICMP, click **Ignore errors and start the install**.

2. BOSH Director installs. This may take a few moments. When the installation process successfully completes, the **Changes Applied** window appears.
3. Click the **Installation Dashboard** link to return to the Installation Dashboard.

Next Steps

- (Optional) When Ops Manager finishes deploying, deploy BOSH Add-ons to your system. For more information, see [Deploying BOSH Add-Ons](#).
- Install one or more PCF runtime environments to complete your installation. For more information, see [Installing Runtimes](#).

Installing PCF on Azure Using Terraform

This topic explains how to install Pivotal Cloud Foundry (PCF) on Microsoft Azure using Terraform.

To install PCF on Azure using Terraform, do the following:

1. Deploy Ops Manager. See [Deploying Ops Manager on Azure Using Terraform](#).
2. Configure BOSH Director. See [Configuring BOSH Director on Azure Using Terraform](#).
3. Configure PAS. See [Configuring PAS](#).

Deploying Ops Manager on Azure Using Terraform

This guide describes the preparation steps required to deploy Ops Manager to Azure using Terraform templates.

The Azure Terraform templates describe a set of Azure resources and properties. For more information about how Terraform creates resources in Azure, see the [Azure Provider](#) topic on the Terraform site.

You may also find it helpful to review different deployment options in the [Reference Architecture for Pivotal Cloud Foundry on Azure](#).

Prerequisites

In addition to fulfilling the prerequisites listed in the [Installing Pivotal Cloud Foundry on Azure](#) topic, ensure you have the following:

- The [Terraform CLI](#)
- In your Azure project, ensure you have completed the steps in [Preparing to Deploy Ops Manager on Azure Using Terraform](#) to create a service principal.

Step 1: Download Templates and Edit Variables File

Before you can run Terraform commands to provision infrastructure resources, you must download the Azure Terraform Templates and create a Terraform template variables file as described below:

1. On [Pivotal Network](#), navigate to the Pivotal Application Service (PAS) release.
2. Download the Azure Terraform Templates ZIP file. You can use the templates to deploy PAS or Pivotal Container Service (PKS).
3. Extract the contents of the ZIP file.
4. Move the extracted folder to the `workspace` directory on your local machine.
5. On the command line, navigate to the directory. For example:

```
$ cd ~/workspace/pivotal-cf-terraforming-azure
```

6. Navigate to the `terraforming-pas` or `terraforming-pks` directory that contains the Terraform files for your runtime.
7. In the runtime directory, create a text file named `terraform.tfvars`.
8. Open the `terraform.tfvars` file and add the following:

```
subscription_id      = "YOUR-SUBSCRIPTION-ID"
tenant_id          = "YOUR-TENANT-ID"
client_id          = "YOUR-CLIENT-ID"
client_secret       = "YOUR-CLIENT-SECRET"

env_name           = "YOUR-ENVIRONMENT-NAME"
env_short_name     = "YOUR-ENVIRONMENT-SHORTNAME"
location           = "YOUR-AZURE-LOCATION"
ops_manager_image_uri = "YOUR-OPS-MAN-IMAGE-URI"
dns_suffix         = "YOUR-DNS-SUFFIX"
vm_admin_username   = "YOUR-ADMIN-USERNAME"
```

9. Edit the values in the file according to the table below. For information about locating your Azure service principal values, see [Preparing to Deploy Ops Manager on Azure Using Terraform](#).

Value to replace	Guidance
<code>YOUR-SUBSCRIPTION-ID</code>	Enter the subscription ID of your Azure service principal. Terraform uses this ID when creating resources.
<code>YOUR-TENANT-ID</code>	Enter the tenant ID of your Azure service principal. Terraform uses this ID when creating resources.
<code>YOUR-CLIENT-ID</code>	Enter the client ID of your Azure service principal. Terraform uses this ID when creating resources.

<code>YOUR-CLIENT-SECRET</code>	Enter your Azure service client secret. Terraform requires this secret to create resources.
<code>YOUR-ENVIRONMENT-NAME</code>	Enter a name to use to identify resources in Azure. Terraform prepends the names of the resources it creates with this environment name. This environment variable is also used to name the Azure resource group created for the deployment. Example: <code>mypcf</code> .
<code>YOUR-ENVIRONMENT-SHORTNAME</code>	Enter a name to use when creating storage accounts in Azure. Must be a-z only and no longer than 10 characters. Example: <code>myazure</code> .
<code>YOUR-AZURE-LOCATION</code>	Enter the name of the Azure location in which you want Terraform to create resources. Example: <code>Central US</code> .
<code>YOUR-OPS-MAN-IMAGE-URI</code>	Enter the URL for the Ops Manager Azure image you want to boot. You can find this code in the PDF included with the Ops Manager release on Pivotal Network .
<code>YOUR-DNS-SUFFIX</code>	To prevent the creation of an Ops Manager VM, set this value to an empty string (<code>" "</code>). When using Platform Automation, you must disable the creation of the Ops Manager VM from Terraform. For more information, see Platform Automation .
<code>YOUR-ADMIN-USERNAME</code>	Enter the admin username you want to use for your Ops Manager deployment.

Step 2: (Optional) Add Variables for Isolation Segment

Complete this section if you plan to deploy the Isolation Segment tile.

 **Note:** You can see the configurable options by opening the `variables.tf` file and looking for variables with default values.

Add the following variable to your `terraform.tfvars` file. This causes Terraform to create an additional HTTP load balancer and DNS record to use for the Isolation Segment tile.

```
isolation_segment = "true"
```

Step 3: Create Azure Resources with Terraform

Follow these steps to use the Terraform CLI to create resources on Azure:

- From the directory that contains the Terraform files, run `terraform init` to initialize the directory based on the information you specified in the `terraform.tfvars` file.

```
$ terraform init
```

- Run the following command to create the execution plan for Terraform.

```
$ terraform plan -out=plan
```

- Run the following command to execute the plan from the previous step. It may take several minutes for Terraform to create all the resources in Azure.

```
$ terraform apply plan
```

Step 4: Create DNS Record

- In a browser, navigate to the DNS provider for the DNS suffix you entered in your `terraform.tfvars` file.
- Create a new name server (NS) record for your system domain. Your system domain is `YOUR-ENVIRONMENT-NAME.YOUR-DNS-SUFFIX`.

3. In this record, enter the name servers included in `env_dns_zone_name_servers` from your Terraform output.

Next Steps

Proceed to the next step in the deployment, [Configuring BOSH Director on Azure Using Terraform](#).

Configuring BOSH Director on Azure Using Terraform

This topic describes how to configure the BOSH Director tile within Ops Manager on Azure after [Deploying Ops Manager to Azure](#).

 **Note:** You can also perform the procedures in this topic using the Ops Manager API. For more information, see the [Using the Ops Manager API](#) topic.

Prerequisite

To complete the procedures in this topic, you must have access to the output generated when you ran `terraform apply` to create resources for this deployment. To view this output, run `terraform output` in the directory where you ran `terraform apply`. You use these values in this output to configure the BOSH Director tile.

Step 1: Access Ops Manager

1. In a web browser, navigate to the fully qualified domain name (FQDN) of the BOSH Director. Use the `ops_manager_dns` value from running `terraform output`.
2. When Ops Manager starts for the first time, you must choose one of the following:
 - o [Internal Authentication](#): If you use Internal Authentication, Ops Manager maintains your user database.
 - o [SAML Identity Provider](#): If you use a SAML Identity Provider (IdP), an external identity server maintains your user database.
 - o [LDAP Server](#): If you use a LDAP Server, an external identity server maintains your user database.

Internal Authentication

1. When redirected to the [Internal Authentication](#) page, you must complete the following steps:
 - o Enter a **Username**, **Password**, and **Password confirmation** to create an Admin user.
 - o Enter a **Decryption passphrase** and the **Decryption passphrase confirmation**. This passphrase encrypts the Ops Manager datastore, and is not recoverable.
 - o If you are using an **HTTP proxy** or **HTTPS proxy**, follow the instructions in the [Configuring Proxy Settings for the BOSH CPI](#) topic.
 - o Read the [End User License Agreement](#), and select the checkbox to accept the terms.
 - o Click **Setup Authentication**.
2. Log in to Ops Manager with the Admin username and password you created in the previous step.

SAML Identity Provider

1. Log in to your IdP console and download the IdP metadata XML. Optionally, if your IdP supports metadata URL, you can copy the metadata URL instead of the XML.
2. Copy the IdP metadata XML or URL to the Ops Manager [SAML Identity Provider](#) login page.

 **Note:** The same IdP metadata URL or XML is applied for the BOSH Director. If you use a separate IdP for BOSH, copy the metadata XML or URL from that IdP and enter it into the BOSH IdP Metadata text box in the Ops Manager login page.

3. Enter values for the fields listed below. Failure to provide values in these fields results in a `500` error.
 - o **SAML admin group**: Enter the name of the SAML group that contains all Ops Manager administrators. This field is case-sensitive.
 - o **SAML groups attribute**: Enter the groups attribute tag name with which you configured the SAML server. This field is case-sensitive.
4. Enter your **Decryption passphrase**. Read the [End User License Agreement](#), and select the checkbox to accept the terms.
5. Your Ops Manager login page appears. Enter your username and password. Click **Login**.
6. Download your SAML Service Provider metadata (SAML Relying Party metadata) by navigating to the following URLs:

- 6a. Ops Manager SAML service provider metadata: <https://OPS-MAN-FQDN:443/uaa/saml/metadata>
- 6b. BOSH Director SAML service provider metadata: <https://BOSH-IP-ADDRESS:8443/saml/metadata>

 **Note:** To retrieve your `BOSH-IP-ADDRESS`, navigate to the **Status** tab in the **BOSH Director** tile. Record the **BOSH Director** IP address.

7. Configure your IdP with your SAML Service Provider metadata. Import the Ops Manager SAML provider metadata from Step 6a above to your IdP. If your IdP does not support importing, provide the values below.

- Single sign on URL: <https://OPS-MAN-FQDN:443/uaa/saml/SSO/alias/OPS-MAN-FQDN>
- Audience URI (SP Entity ID): <https://OP-MAN-FQDN:443/uaa>
- Name ID: Email Address
- SAML authentication requests are always signed

8. Import the BOSH Director SAML provider metadata from Step 6b to your IdP. If the IdP does not support an import, provide the values below.

- Single sign on URL: <https://BOSH-IP:8443/saml/SSO/alias/BOSH-IP>
- Audience URI (SP Entity ID): <https://BOSH-IP:8443>
- Name ID: Email Address
- SAML authentication requests are always signed

9. Return to the **BOSH Director** tile, and continue with the configuration steps below.

LDAP Server

- For **Server URL**, enter the URL that points to your LDAP server. With multiple LDAP servers, separate their URLs with spaces. Each URL must include one of the following protocols:
 - `ldap://`: This specifies that the LDAP server uses an unencrypted connection.
 - `ldaps://`: This specifies that the LDAP server uses SSL for an encrypted connection and requires that the LDAP server holds a trusted certificate or that you import a trusted certificate to the JVM truststore.
- For **LDAP Username** and **LDAP Password**, enter the LDAP Distinguished Name (DN) and the password for binding to the LDAP Server. Example DN: `cn=administrator,ou=Users,dc=example,dc=com`

 **Note:** Pivotal recommends that you provide LDAP credentials that grant read-only permissions on the LDAP Search Base and the LDAP Group Search Base. In addition to this, if the bind user belongs to a different search base, you must use the full DN.

 **warning:** Pivotal recommends against reusing LDAP service accounts across environments. LDAP service accounts should not be subject to manual lockouts, such as lockouts that result from users utilizing the same account. Also, LDAP service accounts should not be subject to automated deletions, since disruption to these service accounts could prevent user logins.

- For **User Search Base**, enter the location in the LDAP directory tree from which any LDAP User search begins. The typical LDAP Search Base matches your domain name.

For example, a domain named “cloud.example.com” typically uses the following LDAP User Search Base: `ou=Users,dc=example,dc=com`

- For **User Search Filter**, enter a string that defines LDAP User search criteria. These search criteria allow LDAP to perform more effective and efficient searches. For example, the standard LDAP search filter `cn=Smith` returns all objects with a common name equal to `Smith`.

In the LDAP search filter string that you use to configure your runtime, use `{0}` instead of the username. For example, use `cn={0}` to return all LDAP objects with the same common name as the username.

In addition to `cn`, other attributes commonly searched for and returned are `mail`, `uid` and, in the case of Active Directory, `sAMAccountName`.

 **Note:** For instructions for testing and troubleshooting your LDAP search filters, see [Configuring LDAP Integration with Pivotal Cloud Foundry](#) in the Pivotal Support Knowledge Base.

- For **Group Search Base**, enter the location in the LDAP directory tree from which the LDAP Group search begins.

For example, a domain named “cloud.example.com” typically uses the following LDAP Group Search Base: `ou=Groups,dc=example,dc=com`

- For **Group Search Filter**, enter a string that defines LDAP Group search criteria. The standard value is `member={0}`.
- For **Email Attribute**, enter the attribute name in your LDAP directory that corresponds to the email address in each user record, for example `mail`.
- For **LDAP RBAC Admin Group Name**, enter the DN of the LDAP group you want to have admin permissions in Ops Manager.

- From the dropdown, select how the UAA handles LDAP server referrals out to other external user stores. The UAA can:
 - Automatically follow any referrals.
 - Ignore referrals and return partial result.
 - Throw exception for each referral and abort.
- For **Server SSL Cert**, paste in the root certificate from your CA certificate or your self-signed certificate.
- Enter a **Decryption passphrase** and the **Decryption passphrase confirmation**. This passphrase encrypts the Ops Manager datastore, and is not recoverable.
- If you are using an **HTTP proxy** or **HTTPS proxy**, follow the instructions in [Configuring Proxy Settings for the BOSH CPI](#).
- Read the **End User License Agreement**, and select the checkbox to accept the terms.
- Select **Provision an admin client in the BOSH UAA**. You can use this to enable BOSH automation with scripts and tooling. For more information, see [Provision Admin Client](#) in *Creating UAA Clients for BOSH Director*.
- Click **Setup Authentication**.
- Return to the **BOSH Director** tile, and continue with the configuration steps below.

Step 2: Azure Config Page

- Click the **BOSH Director** tile.
- Select **Azure Config**.

BOSH Director for Azure

Settings Status Credentials

Azure Config

Director Config

Create Networks

Assign AZs and Networks

Security

BOSH DNS Config

Syslog

Resource Config

Azure Config

Subscription ID*

Tenant ID*

Application ID*

Client Secret* [Change](#)

Resource Group Name*

BOSH Storage Account Name*

Cloud Storage Type
 Use Managed Disks
 Use Storage Accounts

Storage Account Type

Deployments Storage Account Name

Default Security Group

SSH Public Key*

SSH Private Key*

Availability Mode
 Availability Sets
 Availability Zones

Azure Environment
 Azure Commercial Cloud
 Azure US Gov Cloud
 Azure Germany Cloud
 Azure China Cloud
 Azure Stack

Tenant Management Resource Endpoint

Domain*

Authentication*

Azure Stack Endpoint Prefix*

Azure Stack CA Certificate*

[Save](#)

3. Complete the following fields with information you obtained in the [Preparing to Deploy Ops Manager on Azure Using Terraform](#) topic.

- **Subscription ID:** Enter the ID of your Azure subscription.
- **Tenant ID:** Enter your `TENANT_ID`.
- **Application ID:** Enter the `APPLICATION_ID` that you created in the [Step 3: Create an AAD Application](#) step of the *Preparing to Deploy Ops Manager on Azure Using Terraform* topic.
- **Client Secret:** Enter your `CLIENT_SECRET`.

4. Complete the following fields:

- **Resource Group Name:** Enter the name of your resource group, which is exported from Terraform as the output `pcf_resource_group_name`.
- **BOSH Storage Account Name:** Enter the name of your storage account, which is exported from Terraform as the output `bosh_root_storage_account`.

5. For **Cloud Storage Type**, select **Use Managed Disks**. For **Storage Account Type**, select **Premium_LRS** which corresponds to SSD-based storage. See [Azure Managed Disks Overview](#) in the Microsoft documentation for more information.

Cloud Storage Type

Use Managed Disks

Storage Account Type

Premium_LRS Storage Account Type to use with managed disks

Use Storage Accounts

⚠ **warning:** You can update your deployment from using storage accounts to using managed disks. However, after you select **Use Managed Disks** and deploy Ops Manager, you cannot change your deployment back to use storage accounts.

6. For **Default Security Group**, enter the `bosh_deployed_vms_security_group_name` output from Terraform.

7. For **SSH Public Key**, enter the `ops_manager_ssh_public_key` output from Terraform.

8. For **SSH Private Key**, enter the `ops_manager_ssh_private_key` output from Terraform.

9. For **Availability Mode**, choose whether to use **Availability Sets** or **Availability Zones**. You can further configure this option in the [Assign AZs and Networks Page](#).

- **For new installations of Ops Manager v2.5:** Availability Zones are selected by default, but you can choose Availability Sets.
- **If you are upgrading from v2.4:** You must use Availability Sets.

💡 **Note:** You can only choose between Availability Zones and Availability Sets in a staged BOSH Director tile that has not yet deployed. After you click **Apply Changes** and the BOSH Director deploys successfully, you cannot switch between Availability Zones and Availability Sets.

10. For **Azure Environment**, select **Azure Commercial Cloud**.

💡 **Note:** To deploy PCF on Azure using Active Directory Federation Services (ADFS), select **Azure Stack**. For more information about enabling ADFS for Azure, see [ADFS Authentication Available for Microsoft Azure](#).

11. Click **Save**.

Step 3: Director Config Page

1. In Ops Manager, select **Director Config**.

Director Config

NTP Servers (comma delimited)*

Bosh HM Forwarder IP Address

Enable VM Resurrector Plugin

Enable Post Deploy Scripts

Recreate All VMs
 This will force BOSH to recreate all VMs on the next deploy. Persistent disk will be preserved.

Recreate All Persistent Disks
 Check this box to recreate all Persistent Disks for the Director and all other Tiles.

Enable bosh deploy retries
 This will attempt to re-deploy a failed deployment up to 5 times.

2. In the **NTP Servers (comma delimited)** field, enter a comma-separated list of valid NTP servers.

Note: The NTP server configuration only updates after VM recreation. Ensure that you select the **Recreate all VMs** checkbox if you modify the value of this field.

3. Leave the **Bosh HM Forwarder IP Address** field blank.

Note: Starting in PAS v2.0, BOSH-reported component metrics are available in the Loggregator Firehose by default. If you continue to use the BOSH HM Forwarder to consume these component metrics, you may receive duplicate data. To prevent this, leave the **Bosh HM Forwarder IP Address** field blank.

4. Select the **Enable VM Resurrector Plugin** checkbox to enable the BOSH Resurrector functionality and increase your runtime availability.
5. Select **Enable Post Deploy Scripts** to run a post-deploy script after deployment. This script allows the job to execute additional commands against a deployment.
- Note:** If you intend to install PKS, you must enable post-deploy scripts.
6. Select **Recreate all VMs** to force BOSH to recreate all VMs on the next deploy. This process does not destroy any persistent disk data.
7. Select **Recreate All Persistent Disks** to force BOSH to migrate and recreate persistent disks for the BOSH Director and all tiles. This process does not destroy any persistent disk data.
8. Select **Enable bosh deploy retries** to instruct Ops Manager to retry failed BOSH operations up to five times.
9. Select **Skip Director Drain Lifecycle** to prevent drain scripts from running when the BOSH Director is recreated.
10. Select **Store BOSH Job Credentials on tmpfs (beta)** to store credentials for BOSH jobs on temporary file storage (tmpfs) memory, rather than on disk. You must recreate all VMs for this setting to take effect.
11. Select **Keep Unreachable Director VMs** if you want to preserve BOSH Director VMs after a failed deployment for troubleshooting purposes.
12. Select **HM Pager Duty Plugin** to enable Health Monitor integration with PagerDuty.

HM Pager Duty Plugin

Service Key*

YOUR-PAGERDUTY-SERVICE-KEY

HTTP Proxy

YOUR-HTTP-PROXY

- **Service Key:** Enter your API service key from PagerDuty.
- **HTTP Proxy:** Enter an HTTP proxy for use with PagerDuty.

13. Select **HM Email Plugin** to enable Health Monitor integration with email.

HM Email Plugin

Host*

smtp.example.com

Port*

25

Domain*

cloudfoundry.example.com

From*

user2@example.com

Recipients*

user@example.com, user1@example.com

Username

user

Password

.....

Enable TLS

- **Host:** Enter your email hostname.
- **Port:** Enter your email port number.
- **Domain:** Enter your domain.
- **From:** Enter the address for the sender.
- **Recipients:** Enter comma-separated addresses of intended recipients.
- **Username:** Enter the username for your email server.
- **Password:** Enter the password for your email server.
- **Enable TLS:** Select this checkbox to enable Transport Layer Security to the email host.

14. For **CredHub Encryption Provider**, you can choose whether BOSH CredHub stores its encryption key internally on the BOSH Director and CredHub VM, or in an external hardware security module (HSM). The HSM option is more secure.

Before configuring an HSM encryption provider in the **Director Config** pane, you must follow the procedures and collect information described in [Preparing CredHub HSMs for Configuration](#).

 **Note:** After you deploy Ops Manager with an HSM encryption provider, you cannot change BOSH CredHub to store encryption keys internally.

CredHub Encryption Provider

Internal
 Luna HSM

Encryption Key Name*

Provider Partition*

Provider Partition Password*

Provider Client Certificate*

Provider Client Certificate Private Key*

HSM Host Address*

HSM Port Address*

Partition Serial Number*

HSM Certificate*

- **Internal:** Select this option for internal CredHub key storage. This option is selected by default and requires no additional configuration.
 - **Luna HSM:** Select this option to use a SafeNet Luna HSM as your permanent CredHub encryption provider, and fill in the following fields:
 1. **Encryption Key Name:** Any name to identify the key that the HSM uses to encrypt and decrypt the CredHub data. Changing this key name after you deploy Ops Manager can cause service downtime.
 2. **Provider Partition:** The partition that stores your encryption key. Changing this partition after you deploy Ops Manager could cause service downtime. For this value and the ones below, use values gathered in [Preparing CredHub HSMs for Configuration](#).
 3. **Provider Partition Password**
 4. **Provider Client Certificate:** The certificate that validates the identity of the HSM when CredHub connects as a client.
 5. **Provider Client Certificate Private Key**
 6. **HSM Host Address**
 7. **HSM Port Address:** If you do not know your port address, enter `1792`.
 8. **Partition Serial Number**
 9. **HSM Certificate:** The certificate that the HSM presents to CredHub to establish a two-way mTLS connection.
15. Select a **Blobstore Location** to either configure the blobstore as an internal server or an external endpoint. Because the internal server is unscalable and less secure, Pivotal recommends that you configure an external blobstore.

Note: After you deploy Ops Manager, you cannot change the blobstore location.

- **Internal:** Select this option to use an internal blobstore. Ops Manager creates a new VM for blob storage. No additional configuration is required.
 - **Enable TLS:** Select this checkbox to enable TLS to the blobstore.

Note: If you are using PASW 2016, make sure you have downloaded [Windows stemcell v1709.10 or higher](#) before enabling TLS.

- **S3 Compatible Blobstore:** Select this option to use an external S3-compatible endpoint.

Blobstore Location

Internal

Enable TLS

S3 Compatible Blobstore

S3 Endpoint*

Bucket Name*

Access Key*

Secret Key*

S3 Signature Version

V2 Signature

V4 Signature

Region*

S3 Backup Strategy

No backups

Use a versioned bucket

Copy into an additional bucket

Backup Bucket Region*

Backup Bucket Name*

Follow the procedures in [Sign up for Amazon S3](#) and [Creating a Bucket](#) in the AWS documentation. When you have created an S3 bucket, complete the following steps:

1. **S3 Endpoint:** Navigate to the [Regions and Endpoints](#) topic in the AWS documentation.
 - a. Locate the endpoint for your region in the **Amazon Simple Storage Service (S3)** table and construct a URL using your region's endpoint. For example, if you are using the `us-west-2` region, the URL you create would be `https://s3-us-west-2.amazonaws.com`. Enter this URL into the **S3 Endpoint** field.
 - b. On a command line, run `ssh ubuntu@OPS-MANAGER-FQDN` to SSH into the Ops Manager VM. Replace `OPS-MANAGER-FQDN` with the fully qualified domain name of Ops Manager.
 - c. Copy the custom public CA certificate you used to sign the S3 endpoint into `/etc/ssl/certs` on the Ops Manager VM.

- d. On the Ops Manager VM, run `sudo update-ca-certificates -f -v` to import the custom CA certificate into the Ops Manager VM truststore.

 **Note:** You must also add this custom CA certificate into the **Trusted Certificates** field in the **Security** page. See [Security Page](#) for instructions.

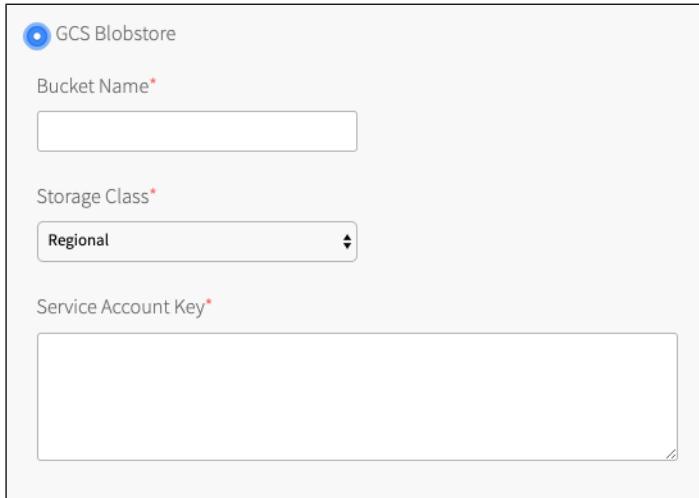
2. **Bucket Name:** Enter the name of the S3 bucket.
3. **Access Key** and **Secret Key:** Enter the keys you generated when creating your S3 bucket.
4. Select **V2 Signature** or **V4 Signature**. If you select **V4 Signature**, enter your **Region**.

 **Note:** AWS recommends using Signature Version 4. For more information about AWS S3 Signatures, see [Authenticating Requests](#) in the AWS documentation.

5. **S3 Backup Strategy:** Select one of the following to configure whether and how to back up the BOSH Director's S3 blobstore:

- **No backups**
- **Use a versioned bucket:** This option uses blobstore bucket versioning to save backups. It requires an S3 blobstore that supports versioning.
- **Copy into an additional bucket:** This option saves blobstore backups to a separate bucket, which is useful with blobstores that do not support versioning. The option requires you to fill in **Backup Bucket Region** and **Backup Bucket Name**.

- **GCS Blobstore:** Select this option to use an external Google Cloud Storage (GCS) endpoint.



The screenshot shows a configuration form for a GCS Blobstore. At the top, there is a radio button labeled "GCS Blobstore". Below it, there are three input fields: "Bucket Name*" with a placeholder box, "Storage Class*" with a dropdown menu set to "Regional", and "Service Account Key*" with a large text area for pasting JSON content.

Follow the procedures in [Creating Storage Buckets](#) in the GCS documentation to create a GCS bucket. To create a GCS bucket, you must have a GCS account. When you have created a GCS bucket, complete the following steps:

1. **Bucket Name:** Enter the name of your GCS bucket.
2. **Storage Class:** Select the storage class for your GCS bucket. See [Storage Classes](#) in the GCP documentation for more information.
3. **Service Account Key:** Follow the steps in the [Set up IAM Service Accounts](#) section of [Preparing to Deploy Ops Manager on GCP Manually](#) to download a JSON file with a private key. Enter the contents of the JSON file into the field.
16. For **Database Location**, select **Internal**.
17. (Optional) Modify the **Director Workers** value, which sets the number of workers available to execute Director tasks. This field defaults to `5`.
18. (Optional) **Max Threads** sets the maximum number of threads that the BOSH Director can run simultaneously. Pivotal recommends that you leave the field blank to use the default value, unless doing so results in rate limiting or errors on your IaaS.
19. (Optional) To add a custom URL for your BOSH Director, enter a valid hostname in **Director Hostname**. You can also use this field to configure a load balancer in front of your BOSH Director. For more information, see [How to Set Up a Load Balancer in Front of Operations Manager Director](#) in the Pivotal Support Knowledge Base.

Director Workers	<input type="text" value="5"/>
Max Threads	<input type="text"/>
Director Hostname	<input type="text" value="opsdirector.example.com"/>

20. (Optional) To set a custom banner that users see when logging in to the Director using SSH, enter text in the **Custom SSH Banner** field.
21. (Optional) Enter your comma-separated custom **Identification Tags**. For example, `iaas:foundation1, hello:world`. You can use the tags to identify your foundation when viewing VMs or disks from your IaaS.

 **Note:** Azure has limited space for identification tags. Pivotal recommends entering no more than five tags.

22. Click **Save**.

Step 4: Create Networks Page

Select **Create Networks** and follow the procedures in this section to add the network configuration that you created for your VPC.

- [Create Network for PAS](#)
- [Create Network for PKS](#)

Create Network for PAS

Create the following three networks:

- [PAS Infrastructure Network](#)
- [PAS Runtime Network](#)
- [PAS Services Network](#)

 **Note:** The Azure portal sometimes displays the names of resources with incorrect capitalization. Always use the Azure CLI to retrieve the correctly capitalized name of a resource.

PAS Infrastructure Network

To create the infrastructure network, do the following:

1. (Optional) Select **Enable ICMP checks** to enable ICMP on your networks. Ops Manager uses ICMP checks to confirm that components within your network are reachable.
2. Click **Add Network**.
3. For **Name**, enter `infrastructure`.
4. Under **Subnets**, create a subnet using the information in the following table:

Network Name	<input type="text" value="infrastructure"/>
	Enter <code>NETWORK-NAME/SUBNET-NAME</code> , where:

Azure Network Name	<ul style="list-style-type: none"> <code>NETWORK-NAME</code> is the <code>network_name</code> output from Terraform <code>SUBNET-NAME</code> is the <code>infrastructure_subnet_name</code> output from Terraform
CIDR	Enter the CIDR listed under <code>infrastructure_subnet_cidrs</code> output from Terraform. For example, <code>10.0.8.0/26</code> .
Reserved IP Ranges	Enter the first nine IP addresses of the subnet. For example, <code>10.0.8.1-10.0.8.9</code> .
DNS	Enter <code>168.63.129.16</code> .
Gateway	Enter the first IP address of the subnet. For example, <code>10.0.8.1</code> .

Create Networks

Warning: Pivotal recommends keeping the IP settings throughout the life of your installation. Ops Manager may prevent you from changing them in the future.
Contact Pivotal support for help completing such a change.

Verification Settings

Enable ICMP checks

Networks

One or many IP ranges upon which your products will be deployed

Add Network

infrastructure Delete

Name*

Subnets

Add Subnet

Azure Network Name*

CIDR*

Reserved IP Ranges

DNS*

Gateway*

5. Click **Save**.

Note: After you deploy Ops Manager, you add subnets with overlapping Availability Zones to expand your network. For more information about configuring additional subnets, see [Expanding Your Network with Additional Subnets](#).

PAS Runtime Network

- Click **Add Network**.
- For **Name**, enter `pas`.
- Under **Subnets**, create a subnet using the information in the following table:

Network Name	<code>pas</code>
--------------	------------------

Azure Network Name	Enter <code>NETWORK-NAME/SUBNET-NAME</code> , where: <ul style="list-style-type: none"> o <code>NETWORK-NAME</code> is the <code>network_name</code> output from Terraform o <code>SUBNET-NAME</code> is the <code>pas_subnet_name</code> output from Terraform
CIDR	Enter the CIDR listed under <code>pas_subnet_cidrs</code> output from Terraform. For example, <code>10.0.0.0/22</code> .
Reserved IP Ranges	Enter the first nine IP addresses of the subnet. For example, <code>10.0.0.1-10.0.0.9</code> .
DNS	Enter <code>168.63.129.16</code> .
Gateway	Enter the first IP address of the subnet. For example, <code>10.0.0.1</code> .

Name*
pas

Subnets

Azure Network Name*
pcf-virtual-network/pcf-pas-subnet

CIDR*
10.0.12.0/22

Reserved IP Ranges
10.0.12.1-10.0.12.9

DNS*
168.63.129.16

Gateway*
10.0.12.1

Add Subnet

4. Click **Save**.

PAS Services Network

1. Click **Add Network**.
2. For **Name**, enter `services`.
3. Under **Subnets**, create a subnet using the information in the following table:

Network Name	<code>services</code>
Azure Network Name	Enter <code>NETWORK-NAME/SUBNET-NAME</code> , where: <ul style="list-style-type: none"> o <code>NETWORK-NAME</code> is the <code>network_name</code> output from Terraform o <code>SUBNET-NAME</code> is the <code>services_subnet_name</code> output from Terraform
CIDR	Enter the CIDR listed under <code>services_subnet_cidrs</code> output from Terraform. For example, <code>10.0.4.0/22</code> .
Reserved IP Ranges	Enter the first nine IP addresses of the subnet. For example, <code>10.0.4.1-10.0.4.9</code> .
DNS	Enter <code>168.63.129.16</code> .
Gateway	Enter the first IP address of the subnet. For example, <code>10.0.4.1</code> .

services

Name*
services

Subnets

Azure Network Name*
pcf-virtual-network/pcf-services-subnet

CIDR*
10.0.8.0/22

Reserved IP Ranges
10.0.8.1-10.0.8.9

DNS*
168.63.129.16

Gateway*
10.0.8.1

Add Subnet

4. Click **Save**. If **Enable ICMP checks** is not selected, you may see red warnings. You can ignore these warnings.

If you complete this step for PAS, proceed to [Assign AZs and Networks Page](#).

Create Network for PKS

Create the following three networks:

- [PKS Infrastructure Network](#)
- [PKS Runtime Network](#)
- [PKS Services Network](#)

Note: The Azure portal sometimes displays the names of resources with incorrect capitalization. Always use the Azure CLI to retrieve the correctly capitalized name of a resource.

PKS Infrastructure Network

To create the infrastructure network, do the following:

1. (Optional) Select **Enable ICMP checks** to enable ICMP on your networks. Ops Manager uses ICMP checks to confirm that components within your network are reachable.
2. Click **Add Network**.
3. For **Name**, enter `infrastructure`.
4. Under **Subnets**, create a subnet using the information in the following table:

Network Name	<code>infrastructure</code>
Azure Network Name	<p>Enter <code>NETWORK-NAME/SUBNET-NAME</code>, where:</p> <ul style="list-style-type: none"> ◦ <code>NETWORK-NAME</code> is the <code>network_name</code> output from Terraform ◦ <code>SUBNET-NAME</code> is the <code>infrastructure_subnet_name</code> output from Terraform

CIDR	Enter the CIDR listed under <code>infrastructure_subnet_cidrs</code> output from Terraform. For example, <code>10.0.8.0/26</code> .
Reserved IP Ranges	Enter the first nine IP addresses of the subnet. For example, <code>10.0.8.1-10.0.8.9</code> .
DNS	Enter <code>168.63.129.16</code> .
Gateway	Enter the first IP address of the subnet. For example, <code>10.0.8.1</code> .

5. Click **Save**.

PKS Runtime Network

1. Click **Add Network**.

2. For **Name**, enter `pks`.

3. Under **Subnets**, create a subnet using the information in the following table:

Network Name	<code>pks</code>
Azure Network Name	Enter <code>NETWORK-NAME/SUBNET-NAME</code> , where: <ul style="list-style-type: none"> o <code>NETWORK-NAME</code> is the <code>network_name</code> output from Terraform o <code>SUBNET-NAME</code> is the <code>pks_subnet_name</code> output from Terraform
CIDR	Enter the CIDR listed under <code>pks_subnet_cidrs</code> output from Terraform. For example, <code>10.0.12.0/22</code> .
Reserved IP Ranges	Enter the first nine IP addresses of the subnet. For example, <code>10.0.12.1-10.0.12.9</code> .
DNS	Enter <code>168.63.129.16</code> .
Gateway	Enter the IP address listed under <code>pks_subnet_gateway</code> output from Terraform. This should be the first IP address of the subnet. For example, <code>10.0.12.1</code> .

4. Click **Save**.

PKS Services Network

1. Click **Add Network**.

2. For **Name**, enter `services`.

3. Under **Subnets**, create a subnet using the information in the following table:

Network Name	<code>services</code>
Azure Network Name	Enter <code>NETWORK-NAME/SUBNET-NAME</code> , where: <ul style="list-style-type: none"> o <code>NETWORK-NAME</code> is the <code>network_name</code> output from Terraform o <code>SUBNET-NAME</code> is the <code>services_subnet_name</code> output from Terraform
CIDR	Enter the CIDR listed under <code>services_subnet_cidrs</code> output from Terraform. For example, <code>10.0.16.0/22</code> .
Reserved IP Ranges	Enter the first nine IP addresses of the subnet. For example, <code>10.0.16.1-10.0.16.9</code> .
DNS	Enter <code>168.63.129.16</code> .
Gateway	Enter the IP address listed under <code>services_subnet_gateway</code> output from Terraform. This should be the first IP address of the subnet. For example, <code>10.0.16.1</code> .

4. Click **Save**. If **Enable ICMP checks** is not selected, you may see red warnings. You can ignore these warnings.

Step 5: Assign AZs and Networks Page

1. Select **Assign AZs and Networks**.

 **Note:** The appearance of this page changes based on whether you are using Availability Zones or Availability Sets. For more information,

see [Step 2: Azure Config Page](#).

- If you use Availability Zones, select one of the three available zones from the **Singleton Availability Zone** dropdown, then select a network from the **Networks** dropdown. The networks that appear in this list are networks that you created earlier in this procedure.
- If you use Availability Sets, choose a network from the **Networks** dropdown. The networks that appear in this list are networks you created earlier in this procedure.

2. Click **Save**.

Step 6: Security Page

1. Select **Security**.

Security

Trusted Certificates

These certificates enable BOSH-deployed components to trust a custom root certificate.

Include OpsManager Root CA in Trusted Certs

Checking this box will place the OpsManager Root CA into the trusted certificate field. The Bosh Director will then place this CA into the trust store of every VM that it deploys.

Generate VM passwords or use single password for all VMs

Generate passwords

Use default BOSH password

Save

2. In **Trusted Certificates**, enter your custom certificate authority (CA) certificates to insert into your organization's certificate trust chain. This feature enables all BOSH-deployed components in your deployment to trust custom root certificates.

To enter multiple certificates, paste your certificates one after the other. For example, format your certificates like the following:

```
-----BEGIN CERTIFICATE-----
ABCDEF12345678ABCDEFGH12345678ABCDEFGH12345678AB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
BCDEFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
CDEFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABBB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
```

 **Note:** If you want to use Docker registries to run PAS app instances in Docker containers, enter the certificate for your private Docker registry in this field. See [Using Docker Registries](#) for more information about running app instances in PAS using Docker registries.

3. (Optional) Select the **Include OpsManager Root CA in Trusted Certs** checkbox to include the Ops Manager root CA in the Trusted Certificates field.

BOSH Director includes this CA in the trust store of every VM that it deploys.

4. Choose **Generate passwords** or **Use default BOSH password**. Pivotal recommends that you use the **Generate passwords** option for greater security.
5. Click **Save**. To view your saved Director password, click the **Credentials** tab.

Step 7: BOSH DNS Config Page

The screenshot shows the 'BOSH DNS Config' page. It has three main sections: 'Excluded Recursors' (with an empty input field), 'Recursor Timeout' (with an empty input field), and 'Handlers' (a text area containing a single placeholder icon). A note on the right says 'Enter a JSON-formatted list of optional custom domain handlers'. At the bottom is a blue 'Save' button.

1. Select **BOSH DNS Config**.
2. (Optional) In **Excluded Recursors**, enter a list of prohibited recursor addresses.
3. (Optional) In **Recursor Timeout**, enter a time limit for contacting the connected recursors. This includes dialing, writing, and reading from the recursor. If any of these actions exceeds the time limit you set, the action fails.

Note: This time limit must include one of the Go parse duration time units. For example, entering `5s` sets the timeout limit to five seconds. For more information about supported time units, see [func ParseDuration](#) in the Go Programming Language documentation.

4. (Optional) In **Handlers**, enter a list of custom domain handlers in JSON format.
5. Click **Save**.

Step 8: Syslog Page

Settings

Syslog

Syslog settings make take up to 1 min to take effect.

Do you want to configure Syslog for OpsManager?

No
 Yes

Address*
 The address or host for the syslog server

Port*

Transport Protocol*

Enable TLS

Permitted Peer*

SSL Certificate*

Queue Size

Forward Debug Logs

Custom rsyslog Configuration

SAVE SYSLOG SETTINGS

1. Select **Syslog**.
2. (Optional) Select **Yes** to send BOSH Director system logs to a remote server.
3. Enter the IP address or DNS name for the remote server in **Address**.
4. Enter the port number that the remote server listens on in **Port**.
5. Select **TCP** or **UDP** from the **Transport Protocol** dropdown. This selection determines which transport protocol is used to send the logs to the remote server.
6. (Optional) Select the **Enable TLS** checkbox to send encrypted logs to remote server with TLS. After you select the checkbox, do the following:
 - a. Enter either the name or SHA1 fingerprint of the remote peer in **Permitted Peer**.
 - b. Enter the SSL certificate for the remote server in **SSL Certificate**.



Note: Pivotal strongly recommends that you enable TLS encryption when you are forwarding logs. Logs can contain sensitive

information, such as cloud provider credentials.

7. (Optional) Enter an integer in **Queue Size**. This value specifies the number of log messages held in the buffer. The default value is 100,000.
8. (Optional) Select the checkbox to **Forward Debug Logs** to an external source. This option is deselected by default. If you select it, you may generate a large amount of log data.
9. (Optional) Enter configuration details for rsyslog in the **Custom rsyslog Configuration** field. This field requires the rainerscript syntax.
10. Click **Save Syslog Settings**.

Step 9: Resource Config Page

1. Select **Resource Config**.

JOB	INSTANCES	PERSISTENT DISK TYPE	VM TYPE	LOAD BALANCERS	INTERNET CONNECTED
Ops Manager Director	Automatic: 1	Automatic: 50 GB	Automatic: Standard_DS2_v2 (cpu: 2, ram: 8 GB)	<input type="checkbox"/>	<input type="checkbox"/>
Master Compilation Job	Automatic: 4	None	Automatic: Standard_F4s (cpu: 4, ram: 8 G)	<input type="checkbox"/>	<input type="checkbox"/>

2. Adjust any values as necessary for your deployment. Under the **Instances**, **Persistent Disk Type**, and **VM Type** fields, choose **Automatic** from the dropdown to allocate the recommended resources for the job. If the **Persistent Disk Type** field reads **None**, the job does not require persistent disk space.

Note: Pivotal recommends provisioning a BOSH Director VM with at least 8 GB memory.

Note: If you set a field to **Automatic** and the recommended resource allocation changes in a future version, Ops Manager automatically uses the new recommended allocation.

Note: If you install PAS for Windows, provision your **Master Compilation Job** with at least 100 GB of disk space.

3. (Optional) For **Load Balancers**, enter your Azure application gateway name for each associated job. To create an application gateway, follow the procedures in [Create an application gateway with SSL termination using Azure PowerShell](#) from the Azure documentation. When you create the application gateway, associate the gateway's IP with your system domain.

4. Click **Save**.

Step 10: (Optional) Add Custom VM Extensions

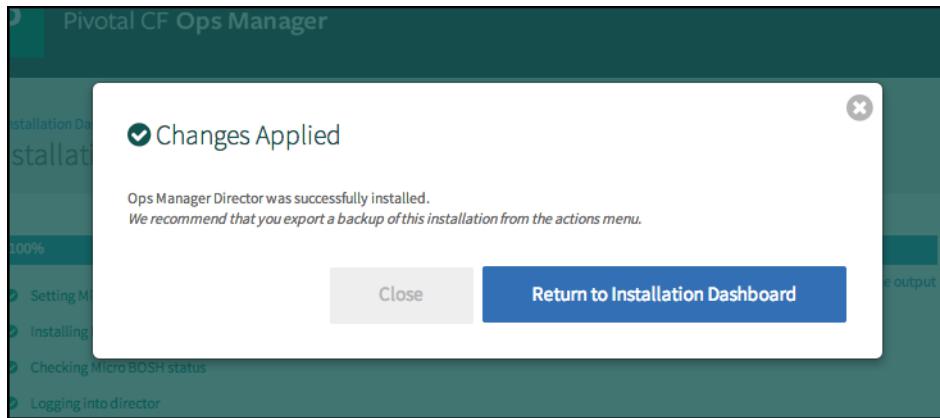
Use the Ops Manager API to add custom properties to your VMs such as associated security groups and load balancers. For more information, see [Managing Custom VM Extensions](#).

Step 11: Complete the BOSH Director Installation

1. Click the **Installation Dashboard** link to return to the Installation Dashboard.
2. Click **Review Pending Changes**, then **Apply Changes**. If the following ICMP error message appears, click **Ignore errors and start the install**.



3. BOSH Director installs. This may take a few moments. When the installation process successfully completes, the **Changes Applied** window appears.



Next Steps

After you complete this procedure, do one of the following:

- If you are deploying PAS, follow the procedures in [Configuring PAS](#).
- If you are deploying PKS, follow the procedures in [Creating Managed Identities in Azure for PKS](#).

Troubleshooting PCF on Azure

This topic describes how to troubleshoot known issues when deploying or running Pivotal Cloud Foundry (PCF) on Azure.

Installation Issues

Cannot Copy the Ops Manager Image

Symptom

Cannot copy the Ops Manager image into your storage account as part of Step 4: Boot Ops Manager in [Deploying Ops Manager on Azure Manually](#).

Explanation

You have an outdated version of the Azure CLI. You need the Azure CLI version 2.0.0 or greater. Run `az --version` from the command line to display your current Azure CLI version.

Solution

Install the Azure CLI 2.0 by following the instructions for your operating system in the [Azure documentation](#).

Deployment Fails at “create-env”

Symptom

After clicking **Review Pending Changes**, then **Apply Changes** to install Ops Manager and PAS, the deployment fails at `create-env` with an error message similar to the following:

```
Command 'deploy' failed:  
Deploying:  
Creating instance 'bosh/0':  
Waiting until instance is ready:  
Starting SSH tunnel:  
  Parsing private key file '/tmp/bosh_ec2_private_key.pem':  
    asn1: structure error: tags don't match (16 vs {class:3 tag:28 length:127  
isCompound:false}) {optional:false explicit:false application:false  
defaultValue:<nil> tag:<nil> stringType:0 timeType:0 set:false omitEmpty:false} pkcs1PrivateKey @2  
===== 2016-09-29 16:28:22 UTC Finished "bosh create-env"  
/var/tempst/workspaces/default/deployments/bosh.yml";  
Duration: 328s; Exit Status: 1  
Exited with 1.
```

Explanation

You provided a passphrase when creating your key pair in Step 4: Boot Ops Manager of [Deploying Ops Manager on Azure Manually](#).

Solution

Create a new key pair with no passphrase and redo the installation. See Step 4: Boot Ops Manager in [Deploying Ops Manager on Azure Manually](#).

Insufficient External Database Permissions

Upgrade issues can be caused when the external database user used for the network policy DB is given insufficient permissions. To avoid this upgrade issue, ensure that the networkpolicyserver database user has the **ALL PRIVILEGES** permission.

Operation Issues

Slow Performance or Timeouts

Symptom

Developers suffer from slow performance or timeouts when pushing or managing apps, and end users suffer from slow performance or timeouts when accessing apps

Explanation

The Azure Load Balancer (ALB) disconnects active TCP connections lying idle for over four minutes.

Solution

To mitigate slow performance or timeouts, the default value of the **Router Timeout to Backends (in seconds)** field is set to 900 seconds. This default value is set high to mitigate performance issues but operators should tune this parameter to fit their infrastructure.

To edit the **Router Timeout to Backends (in seconds)** field:

1. Select the Pivotal Application Service (PAS) tile that is located within your **Installation Dashboard**.
2. Select the **Networking** tab.
3. Enter your desired time, in seconds, within the **Router Timeout to Backends (in seconds)** field.

Router Timeout to Backends (in seconds) (min: 1) *	<input type="text"/>
Timeout for connections from Router (and HAProxy, if you use it) to applications and system components. Increase this to accommodate larger uploads over connections with high latency.	

4. Click **Save**.

Service Instance Creation Times Out

Symptom

You are unable to provision a service instance of a Java or Go service. `cf create-service` fails with error

Failure provisioning service instance... Timed out after 8 minutes... or similar.

Explanation

HTTP libraries for Java and Go, running with default settings, prune idle (for 240 seconds) connections from their connection pool without sending a TCP reset message back to the client service broker. This removes the ability of the broker to provision new service instances.

Solution

In the PAS tile **Configure Networking** pane, set **Frontend Idle Timeout** to 240 seconds or less, so that Cloud Foundry regenerates the front-end connection before it times out.

See the Knowledge Base article [Azure Networking Connection Idle for more than Four minutes](#) for details.

Deleting a PCF on Azure Installation

When you deploy [Pivotal Cloud Foundry](#) (PCF) to Azure, you provision a set of resources. This topic describes how to delete the resources associated with a PCF deployment.

The fastest way to remove resources is to delete the resource group, or resource groups, associated with your PCF on Azure installation.

Delete the Resource Group

Perform the following steps to delete a resource group:

1. Navigate to the [Azure Portal](#).
2. Within your subscription, select **Resource Groups**.
3. Click on the resource group you wish to delete.
4. In the details pane for the resource group, click on the trash can icon. Review the information in the confirmation screen before proceeding.
5. To confirm deletion, type in the resource group name and click **Delete**.

For more information about managing resource groups in Azure, see the [Azure documentation](#).

Upgrading BOSH Director on Azure

This topic describes how to upgrade BOSH Director for Pivotal Cloud Foundry (PCF) on Azure.

Complete the tasks in this topic as part of the Ops Manager upgrade process. For more information, see [Upgrading Pivotal Cloud Foundry](#).

Overview

In this procedure, you create an Ops Manager VM instance to host the upgraded version of Ops Manager. Then, to complete the Ops Manager upgrade, you export your existing Ops Manager installation onto this new VM.

Note: The Azure portal sometimes displays the names of resources with incorrect capitalization. Always use the Azure CLI to retrieve the correctly capitalized name of a resource.

To create an Ops Manager VM instance, do the following:

1. Export environment variables. See [Export Environment Variables](#).
2. Copy the Ops Manager image into your storage account. See [Copy the Ops Manager Image](#).
3. Configure the Ops Manager IP Address. See [Configure Ops Manager IP Address](#).
4. Create Ops Manager VM instance. See [Create Ops Manager VM Instance](#).

Prerequisites

To complete the Ops Manager upgrade, you must have your Ops Manager decryption passphrase. You defined this decryption passphrase during the initial installation of Ops Manager.

Export Environment Variables

Export the `connectionString` variable for your Azure account and assign it to a new environment variable. You use the `connectionString` environment variable when you copy the Ops Manager image in the following procedure.

To export environment variables, do the following:

1. Install the Azure CLI v2.0 by following the instructions for your operating system in the [Azure documentation](#).
2. Set your cloud:

```
$ az cloud set --name AzureCloud
```

If you deployed PCF in an environment other than Azure Cloud, consult the following list:

- For Azure China, replace `AzureCloud` with `AzureChinaCloud`. If logging in to `AzureChinaCloud` fails with a `CERT_UNTRUSTED` error, use the latest version of node, 4.x or later.
- For Azure Government Cloud, replace `AzureCloud` with `AzureUSGovernment`.
- For Azure Germany, replace `AzureCloud` with `AzureGermanCloud`.

3. Log in:

```
$ az login
```

Authenticate by navigating to the URL in the output, entering the provided code, and clicking your account.

4. Ensure that the following environment variables are set to the names of the resources you created when deploying Ops Manager as part of the procedures in [Deploying Ops Manager on Azure Manually](#).
 - `$RESOURCE_GROUP`: This should be set to the name of your resource group. Run `az group list` to list the resource groups for your subscription.

- **\$LOCATION** : This should be set to your location, such as `westus`. For a list of available locations, run `az account list-locations`.
 - **\$STORAGE_NAME** : This should be set to your BOSH storage account name. Run `az storage account list` to list your storage accounts.

5. Retrieve the connection string for the account.

```
$ az storage account show-connection-string \
--name $STORAGE_NAME --resource-group $RESOURCE_GROUP
```

The command returns output similar to the following:

6. From the `data:` field in the output above, record the full value of `connectionString` from the output above, starting with and including `DefaultEndpointsProtocol=`

7. Export the value of `connectionString` as the environment variable `$AZURE_STORAGE_CONNECTION_STRING`.

```
$ export AZURE_STORAGE_CONNECTION_STRING="YOUR-ACCOUNT-KEY-STRING"
```

Copy the Ops Manager Image

Copy the Ops Manager image into your storage account. You use the Ops Manager image to create the VM that hosts that the new version of Ops Manager.

To copy the Ops Manager image into your storage account, do the following:

1. Navigate to [Pivotal Network](#) and download the release of **Pivotal Cloud Foundry Ops Manager for Azure** you want to upgrade to.
 2. View the downloaded PDF and locate the Ops Manager image URL appropriate for your region.
 3. Export the Ops Manager image URL as an environment variable.

```
$ export OPS_MAN_IMAGE_URL="YOUR-OPS-MAN-IMAGE-URL"
```

- This command overrides the old Ops Manager image URL with the new Ops Manager image URL.

4. Copy the Ops Manager image into your storage account. For compatibility when upgrading to future versions of Ops Manager, choose a unique name for the image that includes the Ops Manager version number. For example, replace `opsman-image-version` in the following examples with `opsman-image-2018-01-01`.

```
$ az storage blob copy start --source-uri $OPS_MAN_IMAGE_URL \
--connection-string $AZURE_STORAGE_CONNECTION_STRING \
--destination-container opsmanager \
--destination-blob canman-image-version-1hd
```

5. Copying the image may take several minutes. Run the following command and examine the output under "copy":

```
$ az storage blob show --name opsman-image-version.vhd \
--container-name opsmanager \
--account-name $STORAGE_NAME
...
"copy": {
  "completionTime": "2017-06-26T22:24:11+00:00",
  "id": "b9c8b272-a562-4574-baa6-f1a04acfefdf",
  "progress": "53687091712/53687091712",
  "source": "https://opsmanagerverwestus.blob.core.windows.net/images/ops-manager-2.0.x.vhd",
  "status": "success",
  "statusDescription": null
},
},
```

When `status` reads `success`, continue to the next step.

Configure Ops Manager IP Address

Remove the old Ops Manager VM and record the IP address for the Ops Manager network interface. You use this IP address when creating the new Ops Manager VM that hosts the upgraded version of Ops Manager.

To configure the Ops Manager IP address, do one of the following:

- [Reuse the existing dynamic public IP address.](#)
- [Use a new dynamic public IP address.](#)

Reuse Existing Dynamic Public IP Address

To reuse an existing dynamic public IP address, do the following:

1. List your VMs and record the name of your Ops Manager VM:

```
$ az vm list
```

2. Delete your old Ops Manager VM:

```
$ az vm delete --name YOUR-OPS-MAN-VM --resource-group $RESOURCE_GROUP
```

3. List your network interfaces and record the name of the Ops Manager network interface:

```
$ az resource list --resource-group $RESOURCE_GROUP \
--resource-type Microsoft.Network/networkInterfaces
```

Use a New Dynamic Public IP Address

To use a new dynamic public IP address, do the following:

1. Create a new public IP address named `ops-manager-ip-new`.

```
$ az network public-ip create --name ops-manager-ip-new \
--resource-group $RESOURCE_GROUP --location $LOCATION \
--allocation-method Static
{
  "publicIp": {
    "dnsSettings": null,
    "etag": "W/"4450cbe2-9e97-4b17-9cf2-44838339c661"",
    "id": "/subscriptions/995b7eed-77ef-45ff-a5c9-1a405ffb8243/resourceGroups/cf-docs/providers/Microsoft.Network/publicIPAddresses/ops-manager-ip-new",
    "idleTimeoutInMinutes": 4,
    "ipAddress": "40.83.148.183",
    "ipConfiguration": null,
    "location": "westus",
    "name": "ops-manager-ip-new",
    "provisioningState": "Succeeded",
    "publicIpAddressVersion": "IPv4",
    "publicIpAllocationMethod": "Static",
    "resourceGroup": "cf-docs",
    "resourceGuid": "950d4831-1bec-42da-8a79-959bcddea9dd",
    "tags": null,
    "type": "Microsoft.Network/publicIPAddresses"
  }
}
```

2. Record the value for `ipAddress` from the output above. This is the public IP address of Ops Manager.

3. Create a network interface for Ops Manager.

```
$ az network nic create --vnet-name pcf-net \
--subnet pcf --network-security-group opsmgr-nsg \
--private-ip-address 10.0.0.5 \
--public-ip-address ops-manager-ip-new \
--resource-group $RESOURCE_GROUP \
--name ops-manager-nic-new --location $LOCATION
```

4. Shut down your old Ops Manager VM, if it still exists:

```
$ az vm deallocate --name ops-manager --resource-group $RESOURCE_GROUP
```

If your Ops Manager VM is not named `ops-manager`, provide the correct name. To list all VMs in your account, use `az vm list`.

5. Update your DNS record to point to your new public IP address of Ops Manager.

Create Ops Manager VM Instance

Create an Ops Manager VM instance to host the new version of Ops Manager.

To create an Ops Manager VM instance, do the following:

1. To use the key pair from your previous Ops Manager, locate the path to the file on your local machine. To create a new key pair, enter the following command:

```
$ ssh-keygen -t rsa -f opsman -C ubuntu
```

When prompted for a passphrase, press the `enter` key to provide an empty passphrase.

2. Create the Ops Manager VM.

- o If you are using unmanaged disks, run the following command to create your Ops Manager VM, replacing `PATH-TO-PUBLIC-KEY` with the path to your public key `.pub` file, and replacing `opsman-image-version` with the version of Ops Manager you are deploying, for example `opsman-image-2.0.1`:

```
$ az vm create --name opsman-version --resource-group $RESOURCE_GROUP \
--location $LOCATION \
--nics opsman-nic \
--image https://$STORAGE_NAME.my-azure-instance.com/opsmanager/opsman-image-version.vhd \
--os-disk-name opsman-version-osdisk \
--os-disk-size-gb 128 \
--os-type Linux \
--use-unmanaged-disk \
--storage-account $STORAGENAME \
--storage-container-name opsmanager \
--admin-username ubuntu \
--ssh-key-value PATH-TO-PUBLIC-KEY
```

Replace `my-azure-instance.com` with the URL of your Azure instance. Find the complete source URL in the Azure UI by viewing the **Blob properties** of the Ops Manager image you created earlier in this procedure.

- o If you are using Azure managed disks, perform the following steps, replacing `opsman-image-version` with the version of Ops Manager you are deploying, for example `opsman-image-2.0.1`:

1. Create a managed image from the Ops Manager VHD file:

```
$ az image create --resource-group $RESOURCE_GROUP \
--name opsman-image-version \
--source https://$STORAGE_NAME.blob.core.windows.net/opsmanager/opsman-image-version.vhd \
--location $LOCATION \
--os-type Linux
```

If you are using Azure China, Azure Government Cloud, or Azure Germany, replace `blob.core.windows.net` with the following:

- For Azure China, use `blob.core.chinacloudapi.cn`. See the [Azure documentation](#) for more information.
- For Azure Government Cloud, use `blob.core.usgovcloudapi.net`. See the [Azure documentation](#) for more information.
- For Azure Germany, use `blob.core.cloudapi.de`. See the [Azure documentation](#) for more information.

2. Create your Ops Manager VM, replacing `PATH-TO-PUBLIC-KEY` with the path to your public key `.pub` file, and replacing `opsman-version` and `opsman-image-version` with the version of Ops Manager you are deploying, for example `opsman-2.0.1`, `opsman-2.0.1-osdisk`, and `opsman-image-2.0.1`.

```
$ az vm create --name opsman-version --resource-group $RESOURCE_GROUP \
--location $LOCATION \
--nics opsman-nic \
--image opsman-image-version \
--os-disk-name opsman-version-osdisk \
--admin-username ubuntu \
--size Standard_DS2_v2 \
--storage-sku Standard_LRS \
--ssh-key-value PATH-TO-PUBLIC-KEY
```

3. If you have deployed more than one tile in this Ops Manager installation, perform the following steps to increase the size of the Ops Manager VM disk. You can repeat this process and increase the disk again at a later time if necessary.

 **Note:** If you use Azure Stack, you must increase the Ops Manager VM disk size using the Azure Stack UI.

- a. Run the following command to stop the VM and detach the disk, replacing `opsman-version` with the version of Ops Manager you are deploying, for example `opsman-2.0.1`:

```
$ az vm deallocate --name opsman-version \
--resource-group $RESOURCE_GROUP
```

- b. Run the following command to resize the disk to 128 GB, replacing `opsman-version` with the version of Ops Manager you are deploying, for example `opsman-2.0.1`:

```
$ az disk update --size-gb 128 --name opsman-version-osdisk \
--resource-group $RESOURCE_GROUP
```

- c. Run the following command to start the VM, replacing `opsman-version` with the version of Ops Manager you are deploying, for example `opsman-2.0.1`:

```
$ az vm start --name opsman-version --resource-group $RESOURCE_GROUP
```

Next Steps

After you complete this procedure, continue the upgrade instructions in Upgrading Pivotal Cloud Foundry topic.

Installing Pivotal Cloud Foundry on GCP

This topic describes how to install [Pivotal Cloud Foundry](#) (PCF) on Google Cloud Platform (GCP).

It includes resource requirements, prerequisites, instructions for installing PCF on GCP, and additional resources.

Overview

You can install PCF on GCP with either the Pivotal Application Service (PAS) or Pivotal Container Service (PKS) runtime. There are resource requirements specific to each runtime. Ensure you meet the requirements for your runtime and the requirements specific to GCP before installing PCF on GCP.

Requirements

This section lists the following resource requirements for installing PCF on GCP:

- General PCF resource requirements. See [PCF Resource Requirements](#).
- GCP-specific resource requirements. See [GCP Resource Requirements](#).

PCF Resource Requirements

This section lists PCF resource requirements for installing PCF on GCP. It includes general PCF resource requirements for both the PAS and PKS runtimes.

View one of the following, depending on your PCF runtime:

- PAS-specific PCF resource requirements. See [PAS Resource Requirements](#).
- PKS-specific PCF resource requirements. See [PKS Resource Requirements](#).

PAS Resource Requirements

The following are general resource requirements for deploying and managing a PCF deployment with Ops Manager and PAS:

- PAS requires sufficient IP allocation. The following lists the minimum required IP allocations:
 - One static IP address for either HAProxy or one of your Gorouters
 - One static IP address for each job in the Ops Manager tile. See the Ops Manager **Resource Config** pane for each tile for a full list.
 - One static IP address for each job listed below:
 - Consul
 - NATS
 - File Storage
 - MySQL Proxy
 - MySQL Server
 - Backup Restore Node
 - HAProxy
 - Router
 - MySQL Monitor
 - Diego Brain
 - TCP Router
 - One IP for each VM instance created by the service.
 - An additional IP address for each compilation worker. Use the following formula to determine the total IPs required:
$$\text{IPs needed} = \text{static IPs} + \text{VM instances} + \text{compilation workers}$$
 - Pivotal recommends that you allocate at least 36 dynamic IP addresses when deploying Ops Manager and PAS. BOSH requires additional dynamic IP addresses during installation to compile and deploy VMs, install PAS, and connect to services.
- Pivotal recommends using a network without DHCP for deploying PAS VMs.

 **Note:** If you have DHCP, refer to the [Troubleshooting Guide](#) to avoid issues with your installation.

PKS Resource Requirements

For PKS-specific resource requirements, see [GCP Prerequisites and Resource Requirements](#).

GCP Resource Requirements

The following are GCP-specific resource requirements for installing PCF on GCP:

- Installing PCF on GCP requires a minimum of the following VM instance limits in your GCP account. The number of VMs required depends on the number of tiles and availability zones you plan to deploy. The following VM guidelines apply to the PAS and Small Footprint PAS runtimes:
 - **PAS:** At a minimum, a new GCP deployment requires the following custom VMs for PAS and Ops Manager:

VM Count	vCPU Count per VM	RAM (GB)
28	1	1
2	2	1
3	1	2
4	2	4
3	2	8
3	4	16

 **Note:** If you are deploying a test or sandbox PCF that does not require high availability, then you can scale down the number of VM instances in your deployment. For more information, see [Scaling PAS](#).

- **Small Footprint PAS:** At a minimum, a new GCP deployment requires the following VMs to run Small Footprint PAS:

	VM Type	VM Count	vCPU Count per VM	RAM (GB)	Notes
Small Footprint PAS	micro	11	1	1	Add 1 to VM count if using HAProxy
	small	3	1	2	
	highcpu	1	2	2	
	xlarge.disk	1	4	16	
	xlarge	1	4	16	
	medium.mem	1	1	6	
large.disk	1	2	8		
Ops Manager	large.disk	1	2	8	
	large.cpu	4	4	4	

- **PKS:** See [GCP Prerequisites and Resource Requirements](#).
- (PAS-only) Your GCP project must have sufficient quota to deploy all the VMs needed to install PCF with PAS. For a list of suggested quotas, see [Recommended GCP Quotas](#).

Prerequisites

To install PCF on GCP, you must do the following:

- Install the Google Cloud SDK on your machine and authenticate it to your GCP account. To download the Google Cloud SDK, see [Google Cloud SDK](#).
- Increase or remove the VM instance limits in your GCP account. For VM instance requirements, see [GCP Resource Requirements](#).
- Update your GCP account with the following required permissions:
 - The following permissions are required to create firewalls, networks, load balancers, and other resources:
 - Compute Engine > Compute Instances Admin (beta)
 - Compute Engine > Compute Network Admin
 - Compute Engine > Compute Security Admin
 - To use Google Cloud Storage (GCS) for Cloud Controller file storage, the following permission is required to create buckets:

- Storage > Storage Admin
 - To use Cloud DNS, the following permission is required to add and modify DNS entries:
 - Project > Editor
- Create an SSL certificate for your PCF domain.

 **Note:** To deploy PCF to a production environment, you must obtain a certificate from a certificate authority. Pivotal recommends using a self-signed certificate generated by Ops Manager for development and testing purposes only.

- Assign administrative rights to a domain for PCF. You need to be able to add wildcard records to this domain.
- Create a wildcard DNS record that points to your router or load balancer. Alternatively, you can use a service such as xip.io. For example, `203.0.113.0.xip.io`. Then, create at least one wildcard TLS certificate that matches the DNS record you configured.

 **Note:** With a wildcard DNS record, every hostname in your domain resolves to the IP address of your router or load balancer. For example, if you create a DNS record `*.example.com` pointing to your router, every app deployed to the `example.com` domain resolves to the IP address of your router.

- Create one or more NTP servers, if the NTP servers are not already provided by your GCP project.
- Install the most recent version of one of the following CLIs, depending on your PCF runtime:
 - **PAS:** The Cloud Foundry Command Line Interface (cf CLI). See [Cloud Foundry Command Line Interface](#) in GitHub.
 - **PKS:** The PKS CLI. See [Installing the PKS CLI](#).
- **(PAS-only)** Request a quota increase for your GCP project. For GCP quota requirements, see [GCP Resource Requirements](#). To request an increase, see [Quotas](#) in the GCP console.
- **(PAS-only)** Configure sufficient IP allocation. For more information about IP allocation requirements, see [PAS Resource Requirements](#).
- **(Optional) (PAS-only)** Configure external storage. Pivotal recommends using external storage if possible. For more information about how file storage location affects platform performance and stability during upgrades, see [Configure File Storage](#).
- **(Optional) (PAS and Ops Manager-only)** Configure external databases. Pivotal recommends using external databases in production deployments for BOSH Director and PAS. An external database must be configured to use the UTC timezone.
- **(Optional) (PAS and Ops Manager-only)** Configure external user stores. When you deploy PCF, you can select a SAML user store for Ops Manager or a SAML or LDAP user store for PAS, to integrate existing user accounts.

Install PCF on GCP

You can install PCF on GCP either manually or using Terraform.

To install PCF on GCP, do one of the following:

- Install PCF on GCP manually. See [Installing PCF on GCP Manually](#).
- Install PCF on GCP using Terraform. See [Install PCF on GCP Using Terraform](#).

Additional Resources

The following are additional resources related to installing PCF on GCP:

- For general authentication guidelines for GCP, see [GCP authentication documentation](#).
- For more information about troubleshooting the PCF on GCP infrastructure, see [Troubleshooting PCF on GCP](#).
- For production-level deployment options for PCF on GCP, see [Reference Architecture for Pivotal Cloud Foundry on GCP](#).
- For general certificate requirements for deploying PCF, see [Certificate Requirements](#).

Recommended GCP Quotas

Default quotas on a new GCP subscription do not have enough quota for a typical production-level PCF deployment.

The following table lists recommended resource quotas for a single PCF deployment.

Metric	Resource	Suggested Minimum Quota
CPUs	Regional	150*
Firewall rules	Global	15
Forwarding rules	Global	15
Backend services	Global	5
Health checks	Global	5
Images	Global	10
Static IP addresses**	Regional	5
In-use IP addresses	Global	5
In-use IP addresses**	Regional	5
Networks	Global	5
Subnetworks	Global	5
Routes	Global	20
Target pools	Global	5
Target HTTP proxies	Global	5
Target HTTPS proxies	Global	5
Persistent Disk Standard (GB)	Regional	15,000

* Assuming a deployment with 100 app instances.

** Assuming a SNAT topology.

To view production-level deployment options for PCF on GCP, see the [Reference Architecture for Pivotal Cloud Foundry on GCP](#).

For instructions on how to set up GCP resources required to deploy PCF, see [Preparing to Deploy Ops Manager on GCP Manually](#).

Installing PCF on GCP Manually

This topic explains how to install PCF on GCP manually.

Overview

You can install PCF on GCP with either the Pivotal Application Service (PAS) or Pivotal Container Service (PKS) runtime.

To install PCF on GCP manually, do one of the following:

- Install with PAS. See [Install Manually with PAS](#).
- Install with PKS. See [Install Manually with PKS](#).

Install Manually with PAS

To install manually with PAS, do the following:

1. Prepare to deploy Ops Manager. See [Preparing to Deploy Ops Manager on GCP Manually](#).
2. Deploy Ops Manager. See [Deploying Ops Manager on GCP Manually](#).
3. Configure BOSH Director. See [Configuring BOSH Director on GCP Manually](#).
4. (Optional). Configure a shared VPC. See [\(Optional\) Configuring a Shared VPC on GCP](#).
5. Configure PAS. See [Configuring PAS](#).

Install Manually with PKS

To install manually with PKS, do the following:

1. Prepare to deploy Ops Manager. See [Preparing to Deploy Ops Manager on GCP Manually](#).
2. Deploy Ops Manager. See [Deploying Ops Manager on GCP Manually](#).
3. Configure BOSH Director. See [Configuring BOSH Director on GCP Manually](#).
4. (Optional). Configure a shared VPC. See [\(Optional\) Configuring a Shared VPC on GCP](#).
5. Configure PKS. See [Installing PKS on GCP](#).

Preparing to Deploy Ops Manager on GCP Manually

This topic describes the preparation steps required to install Ops Manager for Pivotal Cloud Foundry (PCF) on Google Cloud Platform (GCP).

Prerequisites

Before you prepare your Ops Manager installation, do the following depending on the runtime you intend to deploy:

- If you are deploying Pivotal Application Service (PAS), see [PCF on GCP Requirements](#).
- If you are deploying Pivotal Container Service (PKS), see [GCP Prerequisites and Resource Requirements](#).

Configuration and Components

This section outlines high-level infrastructure options for PCF on GCP. A PCF deployment includes Ops Manager and your chosen runtime. For example, both Ops Manager with PAS and Ops Manager with PKS are PCF deployments. For more information, review the deployment options and recommendations in [Reference Architecture for Pivotal Cloud Foundry on GCP](#).

You can deploy PCF using one of two main configurations on a GCP virtual private cloud (VPC):

- A *single-project* configuration that gives Ops Manager full access to VPC resources
- A *shared VPC* configuration in which Ops Manager shares VPC resources

See [Shared vs Single-Project VPCs](#) in the *Reference Architecture for Pivotal Cloud Foundry on GCP* topic for a full discussion and recommendations.

When deploying PCF on GCP, Pivotal recommends using the following GCP components:

- [Google Cloud SQL](#) for external database services
- [NAT Gateway Instances](#) to limit the number of VMs with public IP addresses
- [Google Cloud Storage](#) for external file storage

Step 1: Set up IAM Service Accounts

Ops Manager uses IAM service accounts to access GCP resources.

For a single-project installation: Complete the following steps to create a service account for Ops Manager.

For a shared-VPC installation: Complete the following steps twice, to create a host account and service account for Ops Manager.

1. From the GCP console, select **IAM & Admin**, then **Service accounts**.

2. Click **Create Service Account**

- **Service account name:** Enter a name. For example, `bosh`.
- **Role:** Select the following roles:
 - **Service Accounts > Service Account User**
 - **Service Accounts > Service Account Token Creator**
 - **Compute Engine > Compute Instance Admin (v1)**
 - **Compute Engine > Compute Network Admin**
 - **Compute Engine > Compute Storage Admin**
 - **Storage > Storage Admin**

 **Note:** You must scroll down in the pop-up windows to select all required roles.

The **Service Account User** role is only required if you plan to use **The Ops Manager VM Service Account** to deploy Ops Manager. For more information about **The Ops Manager VM Service Account**, see [Step 2: Google Cloud Platform Config](#) in *Configuring BOSH Director on GCP*.

- **Service account ID:** The field automatically generates a unique ID based on the username.

- **Furnish a new private key:** Select this checkbox and JSON as the **Key type**.

Create service account

Service account name ?
bosh

Service account ID
bosh-48 @cf-sf-onboarding-env-3.iam.gserviceaccount.com

Furnish a new private key
Downloads a file that contains the private key. Store the file can't be recovered if lost.

Key type
 JSON
Recommended
 P12
For backward compatibility with code using the P12 format.

Enable G Suite Domain-wide Delegation
Allows this service account to be authorized to access all domain without manual authorization on their part. [Learn more](#)

Role ?
Compute Instance Admin

Selected

- Compute Instance Admin (v1)
- Compute Network Admin
- Compute Storage Admin
- Service Account Token Creator
- Service Account User
- Storage Admin

Project

App Engine

BigQuery

Billing

Cloud IAP

Cloud KMS

Cloud SQL

Cloud Security Scanner

3. Click **Create**. Your browser automatically downloads a JSON file with a private key for this account. Save this file in a secure location.

Step 2: Enable Google Cloud APIs

Ops Manager manages GCP resources using the Google Compute Engine and Cloud Resource Manager APIs. To enable these APIs, do the following:

1. Log in to the Google Developers Console at <https://console.developers.google.com>.
2. In the console, navigate to the GCP projects where you want to install Ops Manager.
 - For a single-project installation, complete the following steps for the Ops Manager project.
 - For a shared-VPC installation, complete the following steps for both host and service projects, to enable them to access the Google Cloud API.
3. Select **API Manager > Library**.
4. Under **Google Cloud APIs**, select **Compute Engine API**.
5. On the **Google Compute Engine API** page, click **Enable**.
6. In the search field, enter **Google Cloud Resource Manager API**.
7. On the **Google Cloud Resource Manager API** page, click **Enable**.
8. To verify that the APIs have been enabled, perform the following steps:
 - a. Log in to GCP using the IAM service account you created in [Set up IAM Service Accounts](#):

```
$ gcloud auth activate-service-account --key-file JSON_KEY_FILENAME
```

- b. List your projects:

```
$ gcloud projects list
PROJECT_ID          NAME          PROJECT_NUMBER
my-host-project-id my-host-project-name #####
my-service-project-id my-service-project-name #####
```

This command lists the projects where you enabled Google Cloud APIs.

Step 3: Create a GCP Network with Subnets

1. Log in to the [GCP console](#).
2. Navigate to the GCP project where you want to install Ops Manager. For a shared-VPC installation, navigate to the host project.
3. Select **VPC network**, then **CREATE VPC NETWORK**.

	Name	Region	Subnets	Mode
<input checked="" type="checkbox"/> VPC networks	pcf-virt-net		3	Custom
<input type="checkbox"/> External IP addresses				

4. In the **Name** field, enter a name of your choice for the VPC network. This name helps you identify resources for this deployment in the GCP console. Network names must be lowercase. For example, `pcf-virt-net`.

← Create a VPC network

Name ?

- a. Under **Subnets**, complete the form as follows to create an infrastructure subnet for Ops Manager and NAT instances:

Name	<code>pcf-infrastructure-subnet-GCP-REGION</code> Example: <code>pcf-infrastructure-subnet-us-west1</code>
Region	A region that supports three availability zones. For help selecting the correct region for your deployment, see the Google documentation about regions and zones .
IP address range	A CIDR ending in <code>/26</code> Example: <code>192.168.101.0/26</code>

See the following image for an example:

X ^

New subnet

Name ?

Add a description

Region ?

IP address range ?

Create secondary IP range

Private Google access ?
 On Off

Flow logs
 On Off

Done **Cancel**

! **Note:** For deployments that do not use external IP addresses, enable **Private Google access** to allow your runtime to make API calls to Google services.

- b. Click **Add subnet** to add a second subnet for the BOSH Director and components specific to your runtime. Complete the form as follows:

Name	<input type="text" value="pcf-RUNTIME-subnet-GCP-REGION"/> Example: <input type="text" value="pcf-pas-subnet-us-west1"/>
Region	The same region you selected for the infrastructure subnet
IP address range	A CIDR ending in <input type="text" value="22"/> Example: <input type="text" value="192.168.16.0/22"/>

- c. Click **Add subnet** to add a third **Subnet** with the following details:

Name	<input type="text" value="pcf-services-subnet-GCP-REGION"/> Example: <input type="text" value="pcf-services-subnet-us-west1"/>
Region	The same region you selected for the previous subnets

IP address range	A CIDR in /22 Example: 192.168.20.0/22
------------------	---

See the following image for an example:

VPC networks		CREATE VPC NETWORK	REFRESH				
Name	Region	Subnets	Mode	IP addresses ranges	Gateways	Firewall Rules	Global dynamic routing
pcf-virt-network	3		Custom			0	Off
	us-west1	pcf-subnet-ert-us-west1		192.168.16.0/22	192.168.16.1		
	us-west1	pcf-subnet-infrastructure-us-west1		192.168.101.0/26	192.168.101.1		
	us-west1	pcf-subnet-services-us-west1		192.168.20.0/22	192.168.20.1		

5. Under **Dynamic routing mode**, leave **Regional** selected.

6. Click **Create**.

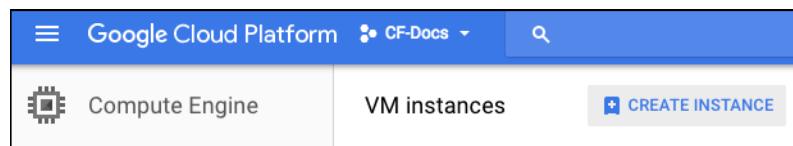
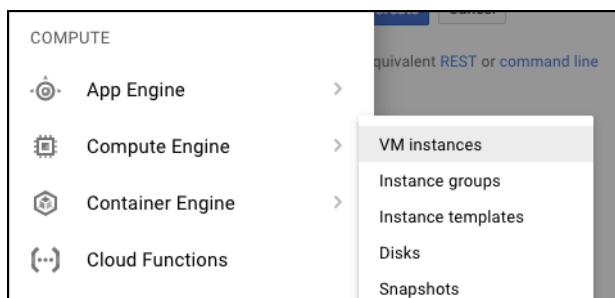
Step 4: Create NAT Instances

Use NAT instances when you want to expose only a minimal number of public IP addresses.

Creating NAT instances permits internet access from cluster VMs. You might, for example, need this internet access for pulling Docker images or enabling internet access for your workloads.

For more information, see [Reference Architecture for Pivotal Cloud Foundry on GCP](#) and [GCP documentation](#).

1. In the GCP console, with your single project or shared-VPC host project selected, navigate to **Compute Engine > VM instances**.



2. Click **CREATE INSTANCE**.

3. Complete the following fields:

- o **Name:** Enter `pcf-nat-gateway-pri`.
This is the first, or primary, of three NAT instances you need. If you use a single AZ, you need only one NAT instance.
- o **Zone:** Select the first zone from your region.
Example: For region `us-west1`, select zone `us-west1-a`.
- o **Machine type:** Select `n1-standard-4`.
- o **Boot disk:** Click **Change** and select `Ubuntu 14.04 LTS`.

[← Create an instance](#)

Name [?](#)
pcf-nat-gateway-pri

Zone [?](#)
us-west1-a

Machine type
4 vCPUs 15 GB memory [Customize](#)

Boot disk [?](#)
 New 10 GB standard persistent disk
 Image Ubuntu 14.04 LTS [Change](#)

4. Expand the additional configuration fields by clicking **Management, disks, networking, SSH keys**.

Firewall [?](#)
Add tags and firewall rules to allow specific network traffic from the Internet

Allow HTTP traffic
 Allow HTTPS traffic

[▼ Management, disks, networking, SSH keys](#)

- a. In the **Startup script** field under **Automation**, enter the following text:

```
#!/bin/bash
sudo sysctl -w net.ipv4.ip_forward=1
sudo sh -c 'echo net.ipv4.ip_forward=1 >> /etc/sysctl.conf'
sudo iptables -t nat -A POSTROUTING -o eth0 -j MASQUERADE
```

Automation

Startup script (Optional)
You can choose to specify a startup script that will run when your instance boots up or restarts. Startup scripts can be used to install software and updates, and to ensure that services are running within the virtual machine. [Learn more](#)

```
#!/bin/bash
sudo sh -c 'echo 1 > /proc/sys/net/ipv4/ip_forward'
```

5. Click **Networking** to open additional network configuration fields:

[Management](#) [Disks](#) **Networking** [SSH Keys](#)

Network tags [?](#) (Optional)
nat-traverse  pcf-nat-instance 

- a. In the **Network tags** field, add the following: `nat-traverse` and `pcf-nat-instance`.
- b. Click on the **Networking** tab and the pencil icon to edit the **Network interface**.
- c. For **Network**, select `pcf-virt-net`. You created this network in [Step 1: Create a GCP Network with Subnets](#).
- d. For **Subnetwork**, select `pcf-infrastructure-subnet-GCP-REGION`.
- e. For **Primary internal IP**, select `Ephemeral (Custom)`. Enter an IP address, for example, `192.168.101.2`, in the **Custom ephemeral IP** address field. The IP address must meet the following requirements:

- The IP address must exist in the CIDR range you set for the `pcf-infrastructure-subnet-GCP-REGION` subnet.
- The IP address must exist in a reserved IP range set later in BOSH Director. The reserved range is typically the first `.1` through `.9` addresses in the CIDR range you set for the `pcf-infrastructure-subnet-GCP-REGION` subnet.
- The IP address cannot be the same as the Gateway IP address set later in Ops Manager. The Gateway IP address is typically the first `.1` address in the CIDR range you set for the `pcf-infrastructure-subnet-GCP-REGION` subnet.

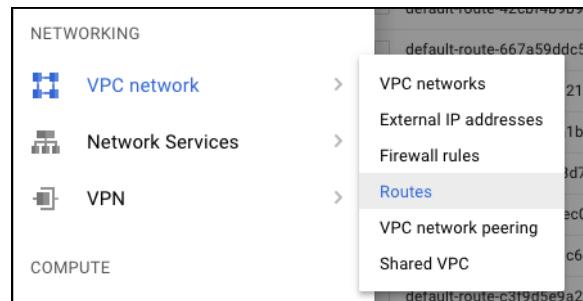
- f. For **External IP**, select `Ephemeral`.

 **Note:** If you select a static external IP address for the NAT instance, then you can use the static IP to further secure access to your CloudSQL instances.

- g. Set IP forwarding to **On**.
- h. Click **Done**.
6. Click **Create** to finish creating the NAT instance.
7. Repeat steps 2-6 to create two additional NAT instances with the names and zones specified in the table below. The rest of the configuration remains the same.

	Name	<input type="text"/> pcf-nat-gateway-sec
	Zone	Select the second zone from your region. Example: For region <input type="text"/> us-west1, select zone <input type="text"/> us-west1-b.
Instance 2	Internal IP	Select <input type="checkbox"/> Custom and enter an IP address in the Internal IP address field. Example: <input type="text"/> 192.168.101.3.
		As described above, this address must in the CIDR range you set for the <input type="text"/> pcf-infrastructure-subnet-GCP-REGION subnet, must exist in a reserved IP range set later in BOSH Director, and cannot be the same as the Gateway IP address set later in Ops Manager.
	Name	<input type="text"/> pcf-nat-gateway-ter
	Zone	Select the third zone from your region. Example: For region <input type="text"/> us-west1, select zone <input type="text"/> us-west1-c.
Instance 3	Internal IP	Select <input type="checkbox"/> Custom and enter an IP address in the Internal IP address field. Example: <input type="text"/> 192.168.101.4.
		As described above, this address must in the CIDR range you set for the <input type="text"/> pcf-infrastructure-subnet-GCP-REGION subnet, must exist in a reserved IP range set later in BOSH Director, and cannot be the same as the Gateway IP address set later in Ops Manager.

Create Routes for NAT Instances



1. Navigate to **VPC Networks > Routes**.

2. Click **CREATE ROUTE**.

3. Complete the form as follows:

- o Name: pcf-nat-pri
- o Network: pcf-virt-net
- o Destination IP range: 0.0.0.0/0
- o Priority: 800
- o Instance tags: pcf
- o Next hop: Specify an instance
- o Next hop instance: pcf-nat-gateway-pri

4. Click **Create** to finish creating the route.

5. Repeat steps 2-4 to create two additional routes with the names and next hop instances specified in the table below. The rest of the configuration remains the same.

Route 2	Name: <input type="text"/> pcf-nat-sec Next hop instance: <input type="text"/> pcf-nat-gateway-sec
Route 3	Name: <input type="text"/> pcf-nat-ter Next hop instance: <input type="text"/> pcf-nat-gateway-ter

Step 5: Create Firewall Rules for the Network

GCP lets you assign tags to VM instances and create firewall rules that apply to VMs based on their tags. For more information about tags, see [Labeling Resources](#) in the Google Cloud documentation. This step assigns tags and firewall rules to Ops Manager components and VMs that handle incoming traffic.

1. With your single project or shared-VPC host project selected, navigate to the **Networking > VPC network** pane and select **Firewall rules**.
2. Apply the firewall rules in the following table:

Firewall Rules	
Rule 1	<p>This rule allows SSH from public networks.</p> <p>Name: <code>pcf-allow-ssh</code> Network: <code>pcf-virt-net</code> Allowed protocols and ports: <code>tcp:22</code> Source filter: IP ranges Source IP ranges: <code>0.0.0.0/0</code> Target tags: <code>allow-ssh</code></p>
Rule 2	<p>This rule allows HTTP from public networks.</p> <p>Name: <code>pcf-allow-http</code> Network: <code>pcf-virt-net</code> Allowed protocols and ports: <code>tcp:80</code> Source filter: IP ranges Source IP ranges: <code>0.0.0.0/0</code> Target tags: <code>allow-http</code>, <code>router</code></p>
Rule 3	<p>This rule allows HTTPS from public networks.</p> <p>Name: <code>pcf-allow-https</code> Network: <code>pcf-virt-net</code> Allowed protocols and ports: <code>tcp:443</code> Source filter: IP ranges Source IP ranges: <code>0.0.0.0/0</code> Target tags: <code>allow-https</code>, <code>router</code></p>
Rule 4	<p>This rule allows Gorouter health checks.</p> <p>Name: <code>pcf-allow-http-8080</code> Network: <code>pcf-virt-net</code> Allowed protocols and ports: <code>tcp:8080</code> Source filter: IP ranges Source IP Ranges: <code>0.0.0.0/0</code> Target tags: <code>router</code></p>
Rule 5	<p>This rule allows communication between BOSH-deployed jobs.</p> <p>Name: <code>pcf-allow-pas-all</code> Network: <code>pcf-virt-net</code> Allowed protocols and ports: <code>tcp;udp;icmp</code> Source filter: Source tags Target tags: <code>pcf</code>, <code>pcf-opsman</code>, <code>nat-traverse</code> Source tags: <code>pcf</code>, <code>pcf-opsman</code>, <code>nat-traverse</code></p>
Rule 6 (Optional)	<p>This rule allows access to the TCP router.</p> <p>Name: <code>pcf-allow-cf-tcp</code> Network: <code>pcf-virt-net</code> Source filter: IP ranges Source IP ranges: <code>0.0.0.0/0</code> Allowed protocols and ports: <code>tcp:1024-65535</code> Target tags: <code>pcf-cf-tcp</code></p>
	<p>This rule allows access to the SSH proxy.</p> <p>Name: <code>pcf-allow-ssh-proxy</code></p>

Rule 7 (Optional)	Network: <code>pcf-virt-net</code> Source filter: IP ranges Source IP ranges: <code>0.0.0.0/0</code> Allowed protocols and ports: <code>tcp:2222</code> Target tags: <code>pcf-ssh-proxy</code> , <code>diego-brain</code>
-------------------	--

 **Note:** If you want your firewall rules to only permit traffic within your private network, modify the **Source IP Ranges** from the table accordingly.

- If you are only using your GCP project to deploy Ops Manager, then you can delete the following default firewall rules:

- o `default-allow-http`
- o `default-allow-https`
- o `default-allow-icmp`
- o `default-allow-internal`
- o `default-allow-rdp`
- o `default-allow-ssh`

If you are deploying **PKS only**, continue to [Next Steps](#).

If you are deploying **PAS or other runtimes**, proceed to the following step.

Step 6: Create Database Instance and Databases

Create Database Instance

- For a shared-VPC installation, select the service project in the GCP console. This step and the following steps allocate resources to the service project, not the host project.
- From the GCP console, select **SQL** and click **CREATE INSTANCE**.
- Ensure **MySQL** is selected and click **Next**.
- Under **MySQL**, select instance type **Second Generation**.
- Click **Configure MySQL** under your choice for instance type: Development, Staging, or Production.
- Configure the instance as follows:
 - o **Instance ID:** `pcf-pas-sql`
 - o **Root password:** Set a password for the root user.
 - o **Region:** Select the region you specified when creating networks.
 - o **Zone:** Any.
 - o **Configure machine type and storage:**
 - Click **Change** and then select **db-n1-standard-2**.
 - Ensure that **Enable automatic storage increases** is selected. This allows DB storage to grow automatically when space is required.
 - o **Enable auto backups and high availability:** Make the following selections:
 - Leave **Automate backups** and **Enable binary logging** selected.
 - Under **High availability**, select the **Create failover replica** checkbox.
 - o **Authorize Networks:** Click **Add network** and create a network named `all` that allows traffic from `0.0.0.0/0`.
-  **Note:** If you assigned static IP addresses to your NAT instances, you can instead limit access to the database instances by specifying the NAT IP addresses.
- Click **Create**.

Create Databases

1. Navigate to the **Instances** page and select the database instance you just created.
2. Select the **Databases** tab.
3. Click **Create database** to create the following databases:

- o account
- o app_usage_service
- o autoscale
- o ccdb
- o console
- o diego
- o locket
- o networkpolicyserver
- o nfsvolume
- o notifications
- o routing
- o silk
- o uaa
- o credhub

4. Select the **USERS** tab.
5. Click **Create user account** to create a unique username and password for each database you created above. For **Host name**, select **Allow any host**. You must create a total of fourteen user accounts.

 **Note:** Ensure that the networkpolicyserver database user has the **ALL PRIVILEGES** permission.

Step 7: Create Storage Buckets

1. With your single project or shared-VPC service project selected in the GCP console, select **Storage > Browser**.
2. Using **CREATE BUCKET**, create buckets with the following names. For **Default storage class**, select **Multi-Regional**:

- o pcf-buildpacks
- o pcf-droplets
- o pcf-packages
- o pcf-resources

Step 8: Create HTTP Load Balancer

For load balancing, you can use a global HTTP load balancer or an internal, regional load balancer with a private IP address.

Single-project, standalone installations typically use a global HTTP load balancer. See [Create HTTP Load Balancer](#) for how to set this up.

Shared-VPC installation typically use an internal TCP/UDP load balancer to minimize public IP addresses. See [Create Internal Load Balancer](#) for how to set this up.

Create Internal Load Balancer

To create an internal load balancer for Ops Manager on GCP, do the following.

1. Create an internal-facing TCP/UDP load balancer for each region of your PCF deployment.

 **Note:** GCP Internal Load Balancer (iLB) is a regional product. Within the same VPC/network, client VMs in a different region from the iLB cannot access the iLB. See [Global Routing Issue](#) in the Google Cloud Setting Up Internal Load Balancing documentation.

2. Assign private IP addresses to the load balancers.

3. After you have deployed Ops Manager, follow instructions in [Create or Update a VM Extension](#) to add a custom VM extension that applies internal load balancing to all VMs deployed by BOSH.

- For example, the following manifest code adds a VM extension `backend-pool` to PCF VMs:

```
vm_extensions:  
- name: backend-pool  
cloud_properties:  
ephemeral_external_ip: true  
backend_service:  
name: name-of-backend-service  
scheme: INTERNAL
```

Create HTTP Load Balancer

To create a global HTTP load balancer for PCF on GCP, do the following:

- [Create Instance Group](#)
- [Create Health Check](#)
- [Configure Back End](#)
- [Configure Front End](#)

Create Instance Group

1. Navigate to **Compute Engine > Instance groups**.
2. Click **CREATE INSTANCE GROUP**.
3. Complete the form as follows:
 - For **Name**, enter `pcf-http-lb`
 - For **Location**, select **Single-zone**.
 - For **Zone**, select the first zone from your region.
Example: For region `us-west1`, select zone `us-west1-a`.
 - Under **Group type**, select **Unmanaged instance group**.
 - For **Network**, select `pcf-virt-net`.
 - For **Subnetwork**, select the `pcf-pas-subnet-my-gcp-region` subnet that you created previously.
 - Click **Create**.
4. Create a second instance group with the following details:
 - Name:** `pcf-http-lb`
 - Location:** **Single-zone**
 - Zone:** Select the second zone from your region.
Example: For region `us-west1`, select zone `us-west1-b`.
 - Group type:** Select **Unmanaged instance group**.
 - Network:** Select `pcf-virt-net`.
 - Subnetwork:** Select the `pcf-pas-subnet-my-gcp-region` subnet that you created previously.
5. Create a third instance group with the following details:
 - Name:** `pcf-http-lb`
 - Location:** **Single-zone**
 - Zone:** Select the third zone from your region.
Example: For region `us-west1`, select zone `us-west1-c`.
 - Group type:** Select **Unmanaged instance group**.
 - Network:** Select `pcf-virt-net`.
 - Subnetwork:** Select the `pcf-pas-subnet-my-gcp-region` subnet that you created previously.

Create Health Check

1. Navigate to **Compute Engine > Health checks**.

2. Click **CREATE HEALTH CHECK**.

3. Complete the form as follows:

- o **Name:** pcf-cf-public
- o **Port:** 8080
- o **Request path:** /health
- o **Check interval:** 30
- o **Timeout:** 5
- o **Healthy threshold:** 10
- o **Unhealthy threshold:** 2

4. Click **Create**.

Configure Back End

1. Navigate to **Network services > Load balancing**.

2. Click **CREATE LOAD BALANCER**.

3. Under **HTTP(S) Load Balancing**, click **Start configuration**.

4. For the **Name**, enter `pcf-global-pcf`.

5. Select **Backend configuration**

a. From the dropdown, select **Backend services > Create a backend service**.

b. Complete the form as follows:

c. **Name:** pcf-http-lb-backend.

d. **Protocol:** HTTP.

e. **Named port:** http.

f. **Timeout:** 10 seconds.

g. Under **Backends > New backend**, select the **Instance group** that corresponds to the first zone of the multi-zone instance group you created.

For example: `pcf-http-lb (us-west1-a)`. Click **Done**.

h. Click **Add backend**, select the **Instance group** that corresponds to the second zone of the multi-zone instance group you created. For example: `pcf-http-lb (us-west1-b)`. Click **Done**.

i. Click **Add backend**, select the **Instance group** that corresponds to the third zone of the multi-zone instance group you created. For example: `pcf-http-lb (us-west1-c)`. Click **Done**.

j. **Health check:** Select the `pcf-cf-public` health check that you created.

k. **Cloud CDN:** Ensure Cloud CDN is disabled.

l. Click **Create**.

Configure Front End

1. Click **Host and path rules** to populate the default fields and a green check mark.

2. Select **Frontend configuration**, and add the following:

o **Name:** pcf-cf-lb-http

o **Protocol:** HTTP

o **IP:** Perform the following steps:

1. Select **Create IP address**.

2. Enter a **Name** for the new static IP address and an optional description. For example, `pcf-global-pcf`.

3. Click **Reserve**.

o **Port:** 80

3. Click **Add Frontend IP and port** and add the following:

Note: Skip this step if you do not have either a self-signed or trusted SSL certificate. When you configure the tile for your chosen runtime, you are given the opportunity to create a new self-signed certificate. Upon creating a certificate, you can complete the **Add Frontend IP and port** section.

- **Name:** `pcf-cf-lb-https`
- **Protocol:** `HTTPS`
- **IP address:** Select the `pcf-global-pcf` address you create for the previous **Frontend IP and Port**.
- **Port:** `443`
- Select **Create a new certificate**. The **Create a New Certificate** dialog is displayed.

Create a new certificate

Name [?](#)

Add a description

Public key certificate [Upload](#)

```
--BEGIN CERTIFICATE--
(paste or upload a public key certificate in .pem format)
--END CERTIFICATE--
```

Certificate chain [Upload](#)

```
--BEGIN CERTIFICATE--
(paste or upload a certificate chain in .pem format)
--END CERTIFICATE--
```

Private key [Upload](#)

```
--BEGIN PRIVATE KEY--
(paste or upload a private key in .pem format)
--END PRIVATE KEY--
```

- In the **Name** field, enter a name for the certificate.
- In the **Public key certificate** field, copy in the contents of your public certificate, or upload your certificate as a .pem file. If the certificate is runtime-generated, copy and paste the generated contents from the runtime's Certificate field into the Ops Manager **Public key certificate** field.
- In the **Certificate chain** field, enter or upload your certificate chain in the .pem format. If you are using a self-signed certificate, such as a PAS or PKS-generated certificate, do not enter a value in the **Certificate Chain** field.
- In the **Private key** field, copy in the contents or upload the .pem file of the private key for the certificate. If the certificate is runtime-generated, copy and paste the generated contents from the runtime's Private Key field into the Ops Manager **Private key** field.

4. Review the completed frontend configuration.

5. Click **Review and finalize** to verify your configuration.

6. Click **Create**.

Step 9: Create TCP WebSockets Load Balancer

The load balancer for tailing logs with WebSockets for PCF on GCP operates on TCP port `443`.

1. From the GCP console, select **Network services > Load balancing > Create load balancer**

2. Under **TCP Load Balancing**, click **Start configuration**.

[←](#) **Create a load balancer**

HTTP(S) Load Balancing Layer 7 load balancing for HTTP and HTTPS applications Learn more Configure HTTP LB HTTPS LB Options Internet-facing only Single or multi-region Start configuration	TCP Load Balancing Layer 4 load balancing or proxy for applications that rely on TCP/SSL protocol Learn more Configure TCP LB SSL Proxy TCP Proxy Options Internet-facing or internal Single or multi-region Start configuration	UDP Load Balancing Layer 4 load balancing for applications that rely on UDP protocol Learn more Configure UDP LB Options Internet-facing or internal Single-region Start configuration
--	--	--

3. In the **Create a load balancer** configuration screen, make the following selections:

- Under **Internet facing or internal only**, select **From Internet to my VMs**.
- Under **Multiple regions or single region**, select **Single region only**.

Please answer a few questions to help us select the right load balancing type for your application

Internet facing or internal only
Do you want to load balance traffic from the Internet to your VMs or only between VMs in your network?

From Internet to my VMs
 Only between my VMs

Multiple regions or single region
Do you want to place the backends for your load balancer in a single region or across multiple regions?

Multiple regions (or not sure yet)
 Single region only

Connection termination
Do you want to offload TCP or SSL processing to the Load Balancer?

Yes (TCP Proxy or SSL Proxy - recommended)
 No (TCP)

Continue

- Under **Connection termination**, select **No (TCP)**.

4. Click **Continue**.

5. In the **New TCP load balancer** window, enter `pcf-wss-logs` in the **Name** field.

6. Click **Backend configuration** to configure the **Backend service**:

Backend configuration

Name ?

Region ?

Backends ?

Select existing instance groups Select existing instances

No instance groups in this region

Backup pool ? (Optional)

Failover ratio ?

%

Health check ?

port: 8080, timeout: 5s, check interval: 5s, unhealthy threshold: 2 attempts

Session affinity ?

- Region: Select the region you used to create the network in [Create a GCP Network with Subnets](#).
- From the Health check dropdown, create a health check with the following details:

- Name:
- Port:
- Request path:
- Check interval:
- Timeout:
- Healthy threshold:
- Unhealthy threshold: The Backend configuration section shows a green check mark.

7. Click **Frontend configuration** to open its configuration window and complete the fields:

- Protocol:
- IP: Perform the following steps:
 1. Select **Create IP address**.
 2. For name **Name** for the new static IP address and an optional description. For example, .
 3. Click **Reserve**.

- Port:

Review and finalize

Backend
Name: **pcf-wss-logs** Region: **us-central1** Session affinity: **None** Health check: **pcf-gorouter**

Frontend

Protocol	IP:Port
TCP	<input type="text" value="████████:443"/>

8. Click **Review and finalize** to verify your configuration.

9. Click **Create**.

Step 10: Create SSH Proxy Load Balancer

1. From the GCP console, select **Network services > Load balancing > Create load balancer**

2. Under **TCP Load Balancing**, click **Start configuration**.

3. Under **Internet facing or internal only**, select **From Internet to my VMs**.

[←](#) **Create a load balancer**

Please answer a few questions to help us select the right load balancing type for your application

Internet facing or internal only
Do you want to load balance traffic from the Internet to your VMs or only between VMs in your network?

From Internet to my VMs
 Only between my VMs

Multiple regions or single region
Do you want to place the backends for your load balancer in a single region or across multiple regions?

Multiple regions (or not sure yet)
 Single region only

Connection termination
Do you want to offload TCP or SSL processing to the Load Balancer?

Yes (TCP Proxy or SSL Proxy - recommended)
 No (TCP)

Continue

4. Under **Connection termination**, select **No (TCP)**.

5. Click **Continue**.

6. In the **New TCP load balancer** window, enter in the **Name** field.

7. Select **Backend configuration**, and enter the following field values:

- **Region:** Select the region you used to create the network in [Create a GCP Network with Subnet](#).
- **Backup pool:**
- **Failover ratio:**
- **Health check:**

Backend configuration

Name ?

Region ?

Backends ?

Select existing instance groups Select existing instances

No instance groups in this region

Backup pool ? (Optional)

Failover ratio ?

%

Health check ?

Session affinity ?

8. Select **Frontend configuration**, and add the following:

- **Protocol:**
- **IP:** Perform the following steps:
 1. Select **Create IP address**.
 2. Enter a **Name** for the new static IP address and an optional description. For example, .
 3. Click **Reserve**.
- **Port:**

9. (Optional) Review and finalize your load balancer.

10. Click **Create**.

Step 11: Create Load Balancer for TCP Router

Note: This step is optional and only required if you enable TCP routing in your deployment.

To create a load balancer for TCP routing in GCP, do the following:

1. From the GCP console, select **Network services > Load balancing > Create load balancer**
2. Under **TCP Load Balancing**, click **Start configuration**.
3. Under **Connection termination**, select **No (TCP)**. Click **Continue**.

Please answer a few questions to help us select the right load balancing type for your application

Internet facing or internal only

Do you want to load balance traffic from the Internet to your VMs or only between VMs in your network?

From Internet to my VMs
 Only between my VMs

Multiple regions or single region

Do you want to place the backends for your load balancer in a single region or across multiple regions?

Multiple regions (or not sure yet)
 Single region only

Connection termination

Do you want to offload TCP or SSL processing to the Load Balancer?

Yes (TCP Proxy or SSL Proxy - recommended)
 No (TCP)

Continue

4. On the **New TCP load balancer** screen, enter a unique name for the load balancer in the **Name** field. For example, `pcf-cf-tcp-lb`.

5. Select **Backend configuration**, and enter the following field values:

- o **Region:** Select the region you used to create the network in [Create a GCP Network with Subnet](#).
- o From the **Health check** dropdown, create a health check with the following details:

- **Name:** `pcf-tcp-lb`
- **Port:** `80`
- **Request path:** `/health`
- **Check interval:** `30`
- **Timeout:** `5`
- **Healthy threshold:** `10`
- **Unhealthy threshold:** `2`
- Click **Save and continue**.

Backend configuration

Name ?
pcf-cf-tcp-lb

Region ?
us-central1

Backends ?

Select existing instance groups Select existing instances

No instance groups in this region

Backup pool ? (Optional)
None

Failover ratio ?
10 %

Health check ?
pcf-tcp-lb (HTTP)
port: 80, timeout: 5s, check interval: 30s, unhealthy threshold: 2 attempts

Session affinity ?
None

6. Select **Frontend configuration**, and add the front end IP and port entry as follows:

- o **Protocol:** `TCP`
- o **IP:** Perform the following steps:
 1. Select **Create IP address**.
 2. Enter a **Name** for the new static IP address and an optional description. For example, `pcf-cf-tcp-lb`.
 3. Click **Reserve**.
- o **Port:** `1024-65535`

[New TCP load balancer](#)

Name [?](#)
pcf-cf-tcp-lb

- Backend configuration**
Your backend is configured
- Frontend configuration**
Your frontend is configured →
- Review and finalize**
Optional

[Create](#) [Cancel](#)

Frontend configuration

Specify an IP address, port and protocol. This IP address is the frontend IP for your clients requests.

Protocol:TCP, IP:35.197.101.100, Port:1024-65535 Not saved [Edit](#)

[+ Add Frontend IP and port](#)

7. Click **Review and finalize** to verify your configuration.

8. Click **Create**.

Step 12: Add DNS Records for Your Load Balancers

In this step, you redirect queries for your domain to the IP addresses of your load balancers.

1. Locate the static IP addresses of the load balancers you created in [Preparing to Deploy Ops Manager on GCP Manually](#):

- An HTTP(S) load balancer named `pcf-global-pcf`
- A TCP load balancer for WebSockets named `pcf-wss-logs`
- A TCP load balancer named `pcf-ssh-proxy`
- A TCP load balancer named `pcf-cf-tcp-lb`

Note: You can locate the static IP address of each load balancer by clicking its name under Network services > Load balancing in the GCP console.

2. Log in to the DNS registrar that hosts your domain. Examples of DNS registrars include Network Solutions, GoDaddy, and Register.com.

3. Create **A records** with your DNS registrar that map domain names to the public static IP addresses of the load balancers located above:

Create and map this record...	To the IP of this load balancer	Required
<code>*.sys.MY-DOMAIN</code> Example: <code>*.sys.example.com</code>	<code>pcf-global-pcf</code>	Yes
<code>*.apps.MY-DOMAIN</code> Example: <code>*.apps.example.com</code>	<code>pcf-global-pcf</code>	Yes
<code>doppler.sys.MY-DOMAIN</code> Example: <code>doppler.sys.example.com</code>	<code>pcf-wss-logs</code>	Yes
<code>loggregator.sys.MY-DOMAIN</code> Example: <code>loggregator.sys.example.com</code>	<code>pcf-wss-logs</code>	Yes
<code>ssh.sys.MY-DOMAIN</code> Example: <code>ssh.sys.example.com</code>	<code>pcf-ssh-proxy</code>	Yes, to allow SSH access to apps
<code>tcp.MY-DOMAIN</code> Example: <code>tcp.example.com</code>	<code>pcf-cf-tcp-lb</code>	No, only set up if you have enabled the TCP routing feature

4. Save changes within the web interface of your DNS registrar.

5. In a terminal window, run the following `dig` command to confirm that you created your A record successfully:

```
dig SUBDOMAIN.EXAMPLE-URL.com
```

Where `SUBDOMAIN.EXAMPLE-URL` is the subdomain for your load balancer.

You should see the A record that you just created:

```
; ANSWER SECTION:  
xyz.EXAMPLE.COM. 1767 IN A 203.0.113.1
```

Next Steps

- (Optional) To prepare for deploying either a PAS or PKS tile on GCP, you can download the needed runtime tile in advance:
 - To download PAS, log in to [Pivotal Network](#), select your desired release version, and download **Pivotal Application Service**.
 - To download PKS, log in to [Pivotal Network](#), select your desired release version, and download **Pivotal Container Service**.
- After initiating the tile download, proceed to the next step, [Deploying Ops Manager on GCP Manually](#).

Deploying Ops Manager on GCP Manually

This topic describes how to deploy Ops Manager for Pivotal Cloud Foundry (PCF) on Google Cloud Platform (GCP).

Before you deploy Ops Manager, see the preparation steps in [Preparing to Deploy Ops Manager on GCP Manually](#).

After you complete this procedure, follow the instructions in [Configuring BOSH Director on GCP Manually](#).

Step 1: Locate the Ops Manager Installation File

1. Log in to the [Pivotal Network](#), and click **Pivotal Cloud Foundry Operations Manager**.
2. From the **Releases** dropdown, select the release to install.
3. Select one of the following download files:
 - o **Pivotal Cloud Foundry Ops Manager for GCP**
 - o **Pivotal Cloud Foundry Ops Manager YAML for GCP**

When you click the download link, your browser downloads or opens the `OpsManager_VERSION_onGCP.pdf` or `OpsManager_VERSION_onGCP.yml` file.

These documents provide the GCP location of the Ops Manager `.tar.gz` installation file based on the geographic location of your installation.

4. Copy the filepath string of the Ops Manager image based on your deployment location.

Step 2: Create a Private VM Image

1. Log in to the [GCP console](#).
2. In the left navigation panel, click **Compute Engine**, and select **Images**.
3. Click **Create Image**.
4. Complete the following fields:
 - o **Name:** Enter a name. For example, `om-pcf`.
 - o **Encryption:** Leave Automatic (recommended) selected.
 - o **Source:** Choose **Cloud Storage file**.
 - o **Cloud Storage file:** Paste in the Google Cloud Storage filepath you copied from the PDF file in the [previous step](#).

Create an image

Name om-pcf

Family (Optional)

Description (Optional)

Encryption Automatic (recommended)

Source Cloud Storage file

Cloud Storage file r-images/pivotal-ops-manager-201608161101000-514500.tar.gz

Create **Cancel**

Equivalent REST or command line

- Click **Create**. The file might take a few minutes to import.

Step 3: Create the Ops Manager VM Instance

- Select the checkbox for the image that you created above.

Images				
CREATE IMAGE REFRESH CREATE INSTANCE DEPRECATE DELETE				
<input type="text" value="Filter by label or name"/> Columns Labels				
Name	Size	Created by	Family	Creation time
<input checked="" type="checkbox"/> om-pcf	50 GB	CF-Docs		Nov 22, 2016, 10:07:54 AM

- Click **Create Instance**.

- In the **Create an instance form**, complete the following fields:

- Name:** Enter a name that matches the naming conventions of your deployment.
- Zone:** Choose a zone from the region in which you created your network.
- Machine type:** Choose `n1-standard-2`.
- Click **Customize** to manually configure the vCPU and memory. An Ops Manager VM instance requires the following minimum specifications:

Machine Spec	Minimum Value
CPU	2 vCPUs
Memory	8 GB

- Boot disk:** Click **Change**, then perform the following steps:

- Click **Custom images** if it is not already selected.
- Select the **Boot disk type**. If your environment has high performance needs, select **SSD**. As an example, environments used to [develop PCF tiles](#) might benefit from a higher performing Ops Manager VM boot disk. For most environments, you can select **Standard**.
- Set the **Size (GB)** of the boot disk to the minimum or higher.

Machine Spec	Minimum Value
Boot disk	100 GB

- Select the Ops Manager image you created in the previous step if it is not already selected.

Boot disk

Select an image or snapshot to create a boot disk; or attach an existing disk.

OS images Application images **Custom images** Snapshots

Existing disks

om-pcf
Created from CF-Docs on Oct 18, 2016, 11:38:02 AM

- Click **Select** to save.

- Under **Identity and API access**, select the **Service account** you created when preparing your environment during the step [Set up IAM Service Accounts](#).
- **Allow HTTP traffic:** Leave this checkbox unselected.
- **Allow HTTPS traffic:** Leave this checkbox unselected.
- **Networking:** Select the **Networking** tab, and perform the following steps:
 - Under **Network interfaces**, perform the following steps:
 - Remove the `default` network interface if this interface still exists.
 - Select the network you created when preparing your environment in the [Create a GCP Network with Subnet](#) section of the *Preparing to Deploy Ops Manager on GCP Manually* topic. For example, `pcf-virt-net`.
 - Under **Subnetwork**, select the `pcf-infrastructure-subnet-GCP-REGION` subnet that you created when preparing your environment in the [Create a GCP Network with Subnet](#) section of the *Preparing to Deploy Ops Manager on GCP Manually* topic.
 - For **Primary internal IP**, select **Ephemeral (Custom)**. Enter an IP address (for example, `192.168.101.5`) in the **Custom ephemeral IP address** field. Specify the next available internal IP address located within the reserved IP range that you [configure later in BOSH Director](#). Do not use the **Gateway IP**, for example `192.168.101.1`.
 - For **External IP**, select **Create IP address**. In the next form, enter a name for the static IP. For example, `om-public-ip`. Click **Reserve**. In the **External IP** dropdown, select the static IP address you just reserved.
 - For **Network tags**, enter `pcf-opsman` and `allow-https`. These tags apply the firewall rules you created in [Create Firewall Rules for the Network](#) to the Ops Manager VM.

4. Click **Create** to deploy the new Ops Manager VM. This might take a few moments.
5. Navigate to your DNS provider, and create an entry that points a fully qualified domain name (FQDN) `opsman.MY-DOMAIN` to the `pcf-opsman` static IP address of Ops Manager that you created in a previous step.

Note: In order to set up Ops Manager authentication correctly, Pivotal recommends using an FQDN to access Ops Manager. Using an ephemeral IP address to access Ops Manager can cause authentication errors upon subsequent access.

Next Steps

After you complete this procedure, follow the instructions in the [Configuring BOSH Director on GCP Manually](#) topic.

Later on, if you need to SSH into the Ops Manager VM to perform diagnostic troubleshooting, see [SSH into Ops Manager](#).

Configuring BOSH Director on GCP Manually

This topic describes how to configure the BOSH Director in Ops Manager for Pivotal Cloud Foundry (PCF) on Google Cloud Platform (GCP).

 **Note:** You can also perform the procedures in this topic using the Ops Manager API. For more information, see the [Using the Ops Manager API](#) topic.

Step 1: Access Ops Manager

1. In a web browser, navigate to the fully qualified domain name (FQDN) of Ops Manager that you set up in [Deploying Ops Manager on GCP Manually](#).

 **Note:** In order to set up Ops Manager authentication correctly, Pivotal recommends using an FQDN to access Ops Manager. Using an ephemeral IP address to access Ops Manager can cause authentication errors upon subsequent access.

2. When Ops Manager starts for the first time, you must choose one of the following:

- [Internal Authentication](#): If you use internal authentication, Ops Manager maintains your user database.
- [SAML Identity Provider](#): If you use a SAML Identity Provider (IdP), an external identity server maintains your user database.
- [LDAP Server](#): If you use an LDAP server, an external identity server maintains your user database.

Internal Authentication

1. When redirected to the [Internal Authentication](#) page, do the following:

- Enter a **Username**, **Password**, and **Password confirmation** to create an Admin user.
- Enter a **Decryption passphrase** and the **Decryption passphrase confirmation**. This passphrase encrypts the Ops Manager datastore, and is not recoverable.
- If you are using an **HTTP proxy** or **HTTPS proxy**, follow the instructions in the [Configuring Proxy Settings for the BOSH CPI](#) topic.
- Read the **End User License Agreement**, and select the checkbox to accept the terms.
- Click **Setup Authentication**.

2. Log in to Ops Manager with the Admin username and password that you created in the previous step.

SAML Identity Provider

1. Log in to your IdP console and download the IdP metadata XML. Optionally, if your IdP supports metadata URL, you can copy the metadata URL instead of the XML.
2. Copy the IdP metadata XML or URL to the Ops Manager [SAML Identity Provider](#) login page.

 **Note:** The same IdP metadata URL or XML is applied for the BOSH Director. If you use a separate IdP for BOSH, copy the metadata XML or URL from that IdP and enter it into the BOSH IdP Metadata text box in the Ops Manager login page.

3. Enter values for the fields listed below. Failure to provide values in these fields results in a **500** error.
 - **SAML admin group**: Enter the name of the SAML group that contains all Ops Manager administrators. This field is case-sensitive.
 - **SAML groups attribute**: Enter the groups attribute tag name with which you configured the SAML server. This field is case-sensitive.
4. Enter your **Decryption passphrase**. Read the **End User License Agreement**, and select the checkbox to accept the terms.
5. Your Ops Manager login page appears. Enter your username and password. Click **Login**.
6. Download your SAML Service Provider metadata (SAML Relying Party metadata) by navigating to the following URLs:
 - **6a. Ops Manager SAML service provider metadata**: <https://OPS-MAN-FQDN:443/uaa/saml/metadata>
 - **6b. BOSH Director SAML service provider metadata**: <https://BOSH-IP-ADDRESS:8443/saml/metadata>

 **Note:** To retrieve your **BOSH-IP-ADDRESS**, navigate to the **Status** tab in the **BOSH Director** tile. Record the **BOSH Director** IP address.

7. Configure your IdP with your SAML Service Provider metadata. Import the Ops Manager SAML provider metadata from Step 6a above to your IdP. If your IdP does not support importing, provide the values below.

- Single sign on URL: `https://OPS-MAN-FQDN:443/uaa/saml/SSO/alias/OPS-MAN-FQDN`
- Audience URI (SP Entity ID): `https://OP-MAN-FQDN:443/uaa`
- Name ID: Email Address
- SAML authentication requests are always signed

8. Import the BOSH Director SAML provider metadata from Step 6b to your IdP. If the IdP does not support an import, provide the values below.

- Single sign on URL: `https://BOSH-IP:8443/saml/SSO/alias/BOSH-IP`
- Audience URI (SP Entity ID): `https://BOSH-IP:8443`
- Name ID: Email Address
- SAML authentication requests are always signed

9. Return to the **BOSH Director** tile, and continue with the configuration steps below.

LDAP Server

- For **Server URL**, enter the URL that points to your LDAP server. With multiple LDAP servers, separate their URLs with spaces. Each URL must include one of the following protocols:
 - `ldap://`: This specifies that the LDAP server uses an unencrypted connection.
 - `ldaps://`: This specifies that the LDAP server uses SSL for an encrypted connection and requires that the LDAP server holds a trusted certificate or that you import a trusted certificate to the JVM truststore.
- For **LDAP Username** and **LDAP Password**, enter the LDAP Distinguished Name (DN) and the password for binding to the LDAP Server. Example DN: `cn=administrator,ou=Users,dc=example,dc=com`

 **Note:** Pivotal recommends that you provide LDAP credentials that grant read-only permissions on the LDAP Search Base and the LDAP Group Search Base. In addition to this, if the bind user belongs to a different search base, you must use the full DN.

 **warning:** Pivotal recommends against reusing LDAP service accounts across environments. LDAP service accounts should not be subject to manual lockouts, such as lockouts that result from users utilizing the same account. Also, LDAP service accounts should not be subject to automated deletions, since disruption to these service accounts could prevent user logins.

- For **User Search Base**, enter the location in the LDAP directory tree from which any LDAP User search begins. The typical LDAP Search Base matches your domain name.

For example, a domain named “cloud.example.com” typically uses the following LDAP User Search Base: `ou=Users,dc=example,dc=com`

- For **User Search Filter**, enter a string that defines LDAP User search criteria. These search criteria allow LDAP to perform more effective and efficient searches. For example, the standard LDAP search filter `cn=Smith` returns all objects with a common name equal to `Smith`.

In the LDAP search filter string that you use to configure your runtime, use `{0}` instead of the username. For example, use `cn={0}` to return all LDAP objects with the same common name as the username.

In addition to `cn`, other attributes commonly searched for and returned are `mail`, `uid` and, in the case of Active Directory, `sAMAccountName`.

 **Note:** For instructions for testing and troubleshooting your LDAP search filters, see [Configuring LDAP Integration with Pivotal Cloud Foundry](#) in the Pivotal Support Knowledge Base.

- For **Group Search Base**, enter the location in the LDAP directory tree from which the LDAP Group search begins.

For example, a domain named “cloud.example.com” typically uses the following LDAP Group Search Base: `ou=Groups,dc=example,dc=com`

- For **Group Search Filter**, enter a string that defines LDAP Group search criteria. The standard value is `member={0}`.
- For **Email Attribute**, enter the attribute name in your LDAP directory that corresponds to the email address in each user record, for example `mail`.
- For **LDAP RBAC Admin Group Name**, enter the DN of the LDAP group you want to have admin permissions in Ops Manager.
- From the dropdown, select how the UAA handles LDAP server referrals out to other external user stores. The UAA can:
 - Automatically follow any referrals.
 - Ignore referrals and return partial result.
 - Throw exception for each referral and abort.

- For **Server SSL Cert**, paste in the root certificate from your CA certificate or your self-signed certificate.
- Enter a **Decryption passphrase** and the **Decryption passphrase confirmation**. This passphrase encrypts the Ops Manager datastore, and is not recoverable.
- If you are using an **HTTP proxy** or **HTTPS proxy**, follow the instructions in [Configuring Proxy Settings for the BOSH CPI](#).
- Read the **End User License Agreement**, and select the checkbox to accept the terms.
- Select **Provision an admin client in the BOSH UAA**. You can use this to enable BOSH automation with scripts and tooling. For more information, see [Provision Admin Client](#) in *Creating UAA Clients for BOSH Director*.
- Click **Setup Authentication**.
- Return to the **BOSH Director** tile, and continue with the configuration steps below.

Step 2: Google Cloud Platform Config

1. Click the **Google Cloud Platform** tile within the **Installation Dashboard**.
2. Select **Google Config**. Complete the following fields:
 - **Project ID**: Enter your GCP project ID in lowercase, such as: `your-gcp-project-id`.
 - **Default Deployment Tag**: Enter the prefix that you used when creating the GCP resources for this PCF installation. For example, `pcf`.
 - Select **AuthJSON** and in the field below enter the contents of the JSON file that you downloaded in the [Set up IAM Service Accounts](#) section of *Preparing to Deploy Ops Manager on GCP Manually*.

Note: As an alternative, you can select **The Ops Manager VM Service Account** option to use the service account automatically created by GCP for the Ops Manager VM. To use this option, the project-wide service account that you set up in [Set up IAM Service Accounts](#) must be assigned the **Service Account Actor** role.

Google Cloud Platform Config

Project ID*

Default Deployment Tag

The Ops Manager VM Service Account
 AuthJSON

`[Empty text area for AuthJSON]`

Save

3. Click **Save**.

Step 3: Director Config Page

1. In Ops Manager, select **Director Config**.

Director Config

NTP Servers (comma delimited)*

Bosh HM Forwarder IP Address

Enable VM Resurrector Plugin

Enable Post Deploy Scripts

Recreate All VMs
 This will force BOSH to recreate all VMs on the next deploy. Persistent disk will be preserved.

Recreate All Persistent Disks
 Check this box to recreate all Persistent Disks for the Director and all other Tiles.

Enable bosh deploy retries
 This will attempt to re-deploy a failed deployment up to 5 times.

2. In the **NTP Servers (comma delimited)** field, enter `169.254.169.254` to designate the `metadata.google.internal` host as NTP server.

Note: The NTP server configuration only updates after VM recreation. Ensure that you select the **Recreate all VMs** checkbox if you modify the value of this field.

3. Leave the **Bosh HM Forwarder IP Address** field blank.

Note: Starting in PAS v2.0, BOSH-reported component metrics are available in the Loggregator Firehose by default. If you continue to use the BOSH HM Forwarder to consume these component metrics, you may receive duplicate data. To prevent this, leave the **Bosh HM Forwarder IP Address** field blank.

4. Select the **Enable VM Resurrector Plugin** checkbox to enable the BOSH Resurrector functionality and increase your runtime availability.
5. Select **Enable Post Deploy Scripts** to run a post-deploy script after deployment. This script allows the job to execute additional commands against a deployment.
- Note:** If you intend to install PKS, you must enable post-deploy scripts.
6. Select **Recreate all VMs** to force BOSH to recreate all VMs on the next deploy. This process does not destroy any persistent disk data.
7. Select **Recreate All Persistent Disks** to force BOSH to migrate and recreate persistent disks for the BOSH Director and all tiles. This process does not destroy any persistent disk data.
8. Select **Enable bosh deploy retries** to instruct Ops Manager to retry failed BOSH operations up to five times.
9. Select **Skip Director Drain Lifecycle** to prevent drain scripts from running when the BOSH Director is recreated.
10. Select **Store BOSH Job Credentials on tmpfs (beta)** to store credentials for BOSH jobs on temporary file storage (tmpfs) memory, rather than on disk. You must recreate all VMs for this setting to take effect.
11. Select **Keep Unreachable Director VMs** if you want to preserve BOSH Director VMs after a failed deployment for troubleshooting purposes.
12. Select **HM Pager Duty Plugin** to enable Health Monitor integration with PagerDuty.

HM Pager Duty Plugin

Service Key*

YOUR-PAGERDUTY-SERVICE-KEY

HTTP Proxy

YOUR-HTTP-PROXY

- **Service Key:** Enter your API service key from PagerDuty.
- **HTTP Proxy:** Enter an HTTP proxy for use with PagerDuty.

13. Select **HM Email Plugin** to enable Health Monitor integration with email.

HM Email Plugin

Host*

smtp.example.com

Port*

25

Domain*

cloudfoundry.example.com

From*

user2@example.com

Recipients*

user@example.com, user1@example.com

Username

user

Password

.....

Enable TLS

- **Host:** Enter your email hostname.
- **Port:** Enter your email port number.
- **Domain:** Enter your domain.
- **From:** Enter the address for the sender.
- **Recipients:** Enter comma-separated addresses of intended recipients.
- **Username:** Enter the username for your email server.
- **Password:** Enter the password for your email server.
- **Enable TLS:** Select this checkbox to enable Transport Layer Security to the email host.

14. For **CredHub Encryption Provider**, you can choose whether BOSH CredHub stores its encryption key internally on the BOSH Director and CredHub VM, or in an external hardware security module (HSM). The HSM option is more secure.

Before configuring an HSM encryption provider in the **Director Config** pane, you must follow the procedures and collect information described in [Preparing CredHub HSMs for Configuration](#).

 **Note:** After you deploy Ops Manager with an HSM encryption provider, you cannot change BOSH CredHub to store encryption keys internally.

CredHub Encryption Provider

Internal
 Luna HSM

Encryption Key Name*

Provider Partition*

Provider Partition Password*

Provider Client Certificate*

Provider Client Certificate Private Key*

HSM Host Address*

HSM Port Address*

Partition Serial Number*

HSM Certificate*

- **Internal:** Select this option for internal CredHub key storage. This option is selected by default and requires no additional configuration.
 - **Luna HSM:** Select this option to use a SafeNet Luna HSM as your permanent CredHub encryption provider, and fill in the following fields:
 1. **Encryption Key Name:** Any name to identify the key that the HSM uses to encrypt and decrypt the CredHub data. Changing this key name after you deploy Ops Manager can cause service downtime.
 2. **Provider Partition:** The partition that stores your encryption key. Changing this partition after you deploy Ops Manager could cause service downtime. For this value and the ones below, use values gathered in [Preparing CredHub HSMs for Configuration](#).
 3. **Provider Partition Password**
 4. **Provider Client Certificate:** The certificate that validates the identity of the HSM when CredHub connects as a client.
 5. **Provider Client Certificate Private Key**
 6. **HSM Host Address**
 7. **HSM Port Address:** If you do not know your port address, enter `1792`.
 8. **Partition Serial Number**
 9. **HSM Certificate:** The certificate that the HSM presents to CredHub to establish a two-way mTLS connection.
15. Select a **Blobstore Location** to either configure the blobstore as an internal server or an external endpoint. Because the internal server is unscalable and less secure, Pivotal recommends that you configure an external blobstore.

Note: After you deploy Ops Manager, you cannot change the blobstore location.

- **Internal:** Select this option to use an internal blobstore. Ops Manager creates a new VM for blob storage. No additional configuration is required.
 - **Enable TLS:** Select this checkbox to enable TLS to the blobstore.

Note: If you are using PASW 2016, make sure you have downloaded [Windows stemcell v1709.10 or higher](#) before enabling TLS.

- **S3 Compatible Blobstore:** Select this option to use an external S3-compatible endpoint.

Blobstore Location

Internal

Enable TLS

S3 Compatible Blobstore

S3 Endpoint*

Bucket Name*

Access Key*

Secret Key*

S3 Signature Version

V2 Signature

V4 Signature

Region*

S3 Backup Strategy

No backups

Use a versioned bucket

Copy into an additional bucket

Backup Bucket Region*

Backup Bucket Name*

Follow the procedures in [Sign up for Amazon S3](#) and [Creating a Bucket](#) in the AWS documentation. When you have created an S3 bucket, complete the following steps:

1. **S3 Endpoint:** Navigate to the [Regions and Endpoints](#) topic in the AWS documentation.
 - a. Locate the endpoint for your region in the **Amazon Simple Storage Service (S3)** table and construct a URL using your region's endpoint. For example, if you are using the `us-west-2` region, the URL you create would be `https://s3-us-west-2.amazonaws.com`. Enter this URL into the **S3 Endpoint** field.
 - b. On a command line, run `ssh ubuntu@OPS-MANAGER-FQDN` to SSH into the Ops Manager VM. Replace `OPS-MANAGER-FQDN` with the fully qualified domain name of Ops Manager.
 - c. Copy the custom public CA certificate you used to sign the S3 endpoint into `/etc/ssl/certs` on the Ops Manager VM.

- d. On the Ops Manager VM, run `sudo update-ca-certificates -f -v` to import the custom CA certificate into the Ops Manager VM truststore.

 **Note:** You must also add this custom CA certificate into the **Trusted Certificates** field in the **Security** page. See [Security Page](#) for instructions.

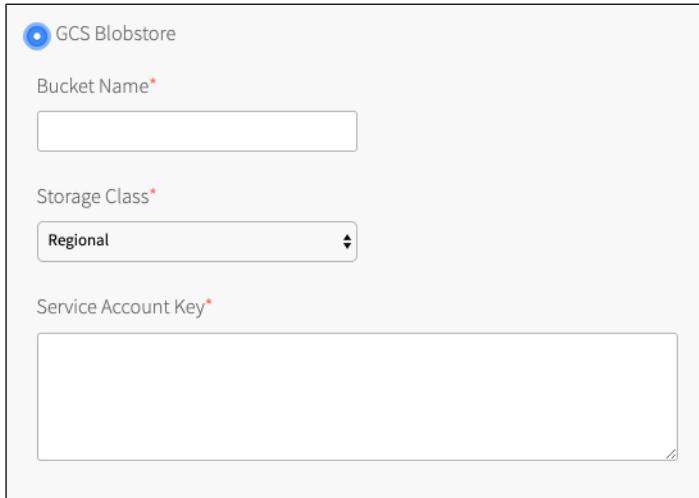
2. **Bucket Name:** Enter the name of the S3 bucket.
3. **Access Key** and **Secret Key:** Enter the keys you generated when creating your S3 bucket.
4. Select **V2 Signature** or **V4 Signature**. If you select **V4 Signature**, enter your **Region**.

 **Note:** AWS recommends using Signature Version 4. For more information about AWS S3 Signatures, see [Authenticating Requests](#) in the AWS documentation.

5. **S3 Backup Strategy:** Select one of the following to configure whether and how to back up the BOSH Director's S3 blobstore:

- **No backups**
- **Use a versioned bucket:** This option uses blobstore bucket versioning to save backups. It requires an S3 blobstore that supports versioning.
- **Copy into an additional bucket:** This option saves blobstore backups to a separate bucket, which is useful with blobstores that do not support versioning. The option requires you to fill in **Backup Bucket Region** and **Backup Bucket Name**.

- o **GCS Blobstore:** Select this option to use an external Google Cloud Storage (GCS) endpoint.



The screenshot shows a configuration form for a GCS Blobstore. It includes fields for Bucket Name, Storage Class (set to Regional), and Service Account Key. The Bucket Name field is empty, and the Storage Class dropdown is set to "Regional". The Service Account Key field is also empty.

Follow the procedures in [Creating Storage Buckets](#) in the GCS documentation to create a GCS bucket. When you have created a GCS bucket, complete the following steps:

1. **Bucket Name:** Enter the name of your GCS bucket.
2. **Storage Class:** Select the storage class for your GCS bucket. See [Storage Classes](#) in the GCP documentation for more information.
3. **Service Account Key:** Enter the contents of the JSON file that you downloaded in the [Set up IAM Service Accounts](#) section of *Preparing to Deploy Ops Manager on GCP Manually*.
16. For **Database Location**, if you configured an external MySQL database such as Cloud SQL, select **External MySQL Database** and complete the fields below. Otherwise, select **Internal**. For more information about creating a Cloud SQL instance, see [Quickstart for Cloud SQL for MySQL](#) in the Google Cloud documentation.
 - o **Host:** Enter the value of your host.
 - o **Port:** Enter your port number. For example, `3306`.
 - o **Username:** Enter your username.
 - o **Password:** Enter your password.
 - o **Database:** Enter your database name.
 In addition, if you selected the **Enable TLS for Director Database** checkbox, you can fill out the following optional fields:
 - o **Enable TLS:** Selecting this checkbox enables TLS communication between the BOSH Director and the database.
 - o **TLS CA:** Enter the Certificate Authority for the TLS Certificate.
 - o **TLS Certificate:** Enter the client certificate for mutual TLS connections to the database.
 - o **TLS Private Key:** Enter the client private key for mutual TLS connections to the database.
 - o **Advanced DB Connection Options:** If you would like to provide additional options for the database, use this field to provide a JSON-formatted options string.
17. (Optional) Modify the **Director Workers** value, which sets the number of workers available to execute Director tasks. This field defaults to `5`.

18. (Optional) **Max Threads** sets the maximum number of threads that the BOSH Director can run simultaneously. Pivotal recommends that you leave the field blank to use the default value, unless doing so results in rate limiting or errors on your IaaS.

19. (Optional) To add a custom URL for your BOSH Director, enter a valid hostname in **Director Hostname**. You can also use this field to configure a load balancer in front of your BOSH Director. For more information, see [How to Set Up a Load Balancer in Front of Operations Manager Director](#) in the Pivotal Support Knowledge Base.

The screenshot shows a configuration interface for a BOSH Director. It has three main sections: 'Director Workers' with a value of '5'; 'Max Threads' which is empty; and 'Director Hostname' with the value 'opsdirector.example.com'.

20. (Optional) To set a custom banner that users see when logging in to the Director using SSH, enter text in the **Custom SSH Banner** field.

21. (Optional) Enter your comma-separated custom **Identification Tags**. For example, `iaas:foundation1, hello:world`. You can use the tags to identify your foundation when viewing VMs or disks from your IaaS.

22. Click **Save**.

Step 4: Create Availability Zones Page

Note: Pivotal recommends at least three availability zones for a highly available installation.

- For configuring a PAS deployment with multiple availability zones, see [Reference Architecture for Pivotal Cloud Foundry on GCP](#).
- For PKS best practices for distributing workloads across multiple availability zones and clusters, see [Maintaining Workload Uptime](#).

1. Select **Create Availability Zones**.

2. Click **Add**.

3. For **Google Availability Zone**:

- Enter one of the zones that you associated to the backend service instance groups of the HTTP(S) Load Balancer. For example, if you are using the `us-central1` region and selected `us-central1-a` as one of the zones for your HTTP(S) Load Balancer instance groups, enter `us-central1-a`.
- Click **Add**.
- Repeat the above step for all the availability zones that you associated to instance groups in [Preparing to Deploy Ops Manager on GCP Manually](#).

Create Availability Zones

Availability Zones

- ▶ us-central1-b
- ▼ us-central1-a

Google Availability Zone*

The Google Availability Zone name
- ▶ us-central1-c

Save

- Click **Save**.
4. Repeat the above step for all the availability zones you are using in your deployment. When you are done, click **Save**.

Step 5: Create Networks Page

1. Select **Create Networks**.
2. Make sure **Enable ICMP checks** is not selected. GCP routers do not respond to ICMP pings.
3. Click **Add Network**. Create the three networks described in the tables below.

Note: To use a shared VPC network, enter the shared VPC host project name before the network name in the format `VPC-PROJECT-NAME/NETWORK-NAME/SUBNET-NAME/REGION-NAME`. For example, `vpc-project/opsmgr/central/us-central1`. For more information, see [Configuring a Shared VPC on GCP](#).

Infrastructure Network

Network Name	<input type="text" value="infrastructure"/>
Google Network Name	Enter the name of the infrastructure network that you created in the Create a GCP Network with Subnets step of <i>Preparing to Deploy Ops Manager on GCP Manually</i> . The format is: <code>pcf-virt-net/pcf-infrastructure-subnet-MY-REGION/MY-GCP-REGION</code>
CIDR	Enter the name of the CIDR ending in <code>/26</code> that you used when you created the infrastructure subnet in GCP. Example: <code>192.168.101.0/26</code>
Reserved IP Ranges	Enter the first <code>.1</code> through <code>.9</code> addresses from the CIDR. For example, if the CIDR is <code>192.168.101.0/26</code> , enter the range <code>192.168.101.1-192.168.101.9</code> .
DNS	<input type="text" value="169.254.169.254"/>
Gateway	Enter the first <code>.1</code> address from the CIDR. For example, if the CIDR is <code>192.168.101.0/26</code> , enter <code>192.168.101.1</code> .
Availability Zones	Select all three availability zones.

Runtime Network

Name	Enter the name of the runtime that you intend to deploy in this environment. For example, <code>pas</code> or <code>pks</code> .
Google Network Name	Enter the name of the runtime network that you created when you prepared your GCP environment. The format is: <code>pcf-virt-net/pcf-RUNTIME-subnet-MY-REGION/MY-GCP-REGION</code>
	Enter the name of the CIDR ending in <code>/22</code> that you used when you created the runtime subnet in GCP.

CIDR	Example: <code>192.168.16.0/22</code>
Reserved IP Ranges	Enter the first <code>.1</code> through <code>.9</code> addresses from the CIDR. For example, if the CIDR is <code>192.168.16.0/22</code> , enter the range <code>192.168.16.1-192.168.16.9</code> .
DNS	<code>169.254.169.254</code>
Gateway	Enter the first <code>.1</code> address from the CIDR. For example, if the CIDR is <code>192.168.16.0/22</code> , enter <code>192.168.16.1</code> .
Availability Zones	Select all three availability zones.

Services Network

Network Name	<code>services</code>
Google Network Name	Enter the name of the services network that you created when you prepared your GCP environment. The format is: <code>pcf-virt-net/pcf-services-subnet-MY-REGION/MY-GCP-REGION</code>
CIDR	Enter the name of the CIDR ending in <code>/22</code> that you used when you created the services subnet in GCP. Example: <code>192.168.20.0/22</code>
Reserved IP Ranges	Enter the first <code>.1</code> through <code>.9</code> addresses from the CIDR. For example, if the CIDR is <code>192.168.20.0/22</code> , enter the range <code>192.168.20.1-192.168.20.9</code> .
DNS	<code>169.254.169.254</code>
Gateway	Enter the first <code>.1</code> address from the CIDR. For example, if the CIDR is <code>192.168.16.0/22</code> , enter <code>192.168.16.1</code> .
Availability Zones	Select all three availability zones.

 **Note:** After you deploy Ops Manager, you add subnets with overlapping Availability Zones to expand your network. For more information about configuring additional subnets, see [Expanding Your Network with Additional Subnets](#).

 **Note:** Ensure `169.254.169.254`, the IP address of the `metadata.google.internal` metadata server, is always specified at the beginning of this step's DNS resolver lists.

Step 6: Assign AZs and Networks Page

1. Select **Assign AZs and Networks**.

Assign AZs and Networks

The BOSH Director is a single instance.

Choose the availability zone in which to place that instance. It is highly recommended that you backup this VM on a regular basis to preserve settings.

Singleton Availability Zone

AZ-MGMT

Network

infrastructure

Save

2. Use the dropdown to select a **Singleton Availability Zone**. The BOSH Director installs in this Availability Zone.
3. Under **Network**, select the `infrastructure` network for your BOSH Director.
4. Click **Save**.

Step 7: Security Page

1. Select **Security**.

Security

Trusted Certificates

These certificates enable BOSH-deployed components to trust a custom root certificate.

Include OpsManager Root CA in Trusted Certs

Checking this box will place the OpsManager Root CA into the trusted certificate field. The Bosh Director will then place this CA into the trust store of every VM that it deploys.

Generate VM passwords or use single password for all VMs

Generate passwords

Use default BOSH password

Save

2. In **Trusted Certificates**, enter your custom certificate authority (CA) certificates to insert into your organization's certificate trust chain. This feature enables all BOSH-deployed components in your deployment to trust custom root certificates.

To enter multiple certificates, paste your certificates one after the other. For example, format your certificates like the following:

```
-----BEGIN CERTIFICATE-----
ABCDEFH12345678ABCDEFGH12345678ABCDEFGH12345678AB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
BCDEFH12345678ABCDEFGH12345678ABCDEFGH12345678ABB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
CDEFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABBB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
```

Note: If you want to use Docker registries to run PAS app instances in Docker containers, enter the certificate for your private Docker registry in this field. See [Using Docker Registries](#) for more information about running app instances in PAS using Docker registries.

- (Optional) Select the **Include OpsManager Root CA in Trusted Certs** checkbox to include the Ops Manager root CA in the Trusted Certificates field. BOSH Director includes this CA in the trust store of every VM that it deploys.
- Choose **Generate passwords** or **Use default BOSH password**. Pivotal recommends that you use the **Generate passwords** option for greater security.
- Click **Save**. To view your saved Director password, click the **Credentials** tab.

Step 8: BOSH DNS Config Page

BOSH DNS Config

Excluded Recursors

Recursor Timeout

Handlers

[]

Enter a JSON-formatted list of optional custom domain handlers

Save

1. Select **BOSH DNS Config**.
2. (Optional) In **Excluded Recursors**, enter a list of prohibited recursor addresses.
3. (Optional) In **Recursor Timeout**, enter a time limit for contacting the connected recursors. This includes dialing, writing, and reading from the recursor. If any of these actions exceeds the time limit you set, the action fails.

Note: This time limit must include one of the Go parse duration time units. For example, entering `5s` sets the timeout limit to five seconds. For more information about supported time units, see [func ParseDuration](#) in the Go Programming Language documentation.

4. (Optional) In **Handlers**, enter a list of custom domain handlers in JSON format.
5. Click **Save**.

Step 9: Syslog Page

Settings

Syslog

Syslog settings make take up to 1 min to take effect.

Do you want to configure Syslog for OpsManager?

No
 Yes

Address*
 The address or host for the syslog server

Port*

Transport Protocol*

Enable TLS

Permitted Peer*

SSL Certificate*

Queue Size

Forward Debug Logs

Custom rsyslog Configuration

SAVE SYSLOG SETTINGS

1. Select **Syslog**.
2. (Optional) Select **Yes** to send BOSH Director system logs to a remote server.
3. Enter the IP address or DNS name for the remote server in **Address**.
4. Enter the port number that the remote server listens on in **Port**.
5. Select **TCP** or **UDP** from the **Transport Protocol** dropdown. This selection determines which transport protocol is used to send the logs to the remote server.
6. (Optional) Select the **Enable TLS** checkbox to send encrypted logs to remote server with TLS. After you select the checkbox, do the following:
 - a. Enter either the name or SHA1 fingerprint of the remote peer in **Permitted Peer**.
 - b. Enter the SSL certificate for the remote server in **SSL Certificate**.



Note: Pivotal strongly recommends that you enable TLS encryption when you are forwarding logs. Logs can contain sensitive

information, such as cloud provider credentials.

7. (Optional) Enter an integer in **Queue Size**. This value specifies the number of log messages held in the buffer. The default value is 100,000.
8. (Optional) Select the checkbox to **Forward Debug Logs** to an external source. This option is deselected by default. If you select it, you may generate a large amount of log data.
9. (Optional) Enter configuration details for rsyslog in the **Custom rsyslog Configuration** field. This field requires the rainerscript syntax.
10. Click **Save Syslog Settings**.

Step 10: Resource Config Page

1. Select **Resource Config**.
2. Ensure that the **Internet Connected** checkboxes are not selected for any jobs. The checkbox gives VMs a public IP address that enables outbound internet access. In [Preparing to Deploy Ops Manager on GCP Manually](#), you provisioned a Network Address Translation (NAT) box to provide internet connectivity to your VMs. For more information about using NAT in GCP, see the [GCP documentation](#).

Note: If you install Pivotal Application Service for Windows (PASW), provision your **Master Compilation Job** with at least 100 GB of disk space.

Step 11: (Optional) Add Custom VM Extensions

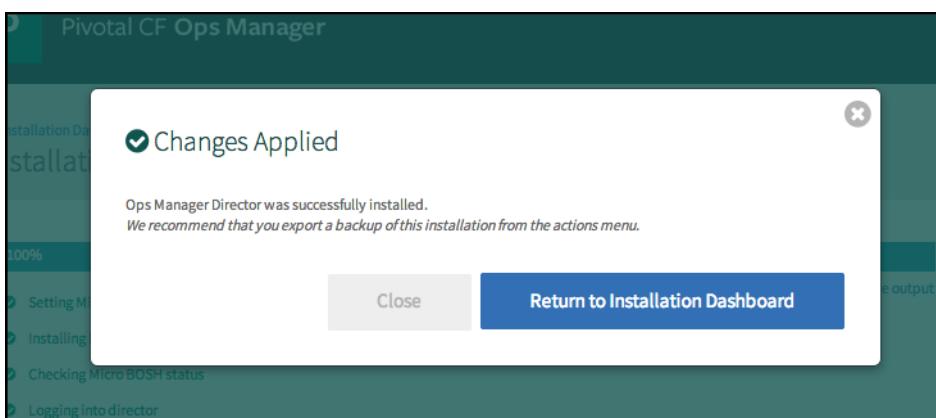
Use the Ops Manager API to add custom properties to your VMs such as associated security groups and load balancers. For more information, see [Managing Custom VM Extensions](#).

Step 12: Complete the BOSH Director Installation

1. Click the **Installation Dashboard** link to return to the Installation Dashboard.
2. Click **Review Pending Changes**, then **Apply Changes**. If the following ICMP error message appears, return to the **Network Config** screen, and ensure you have deselected the **Enable ICMP Checks**. Click **Review Pending Changes**, then **Apply Changes** again.



3. BOSH Director installs. This might take a few moments. When the installation process successfully completes, the **Changes Applied** window appears.



Next Steps

After you complete this procedure, follow the instructions for the runtime you intend to install.

- To deploy PAS, see [Configuring PAS](#).
- To prepare to deploy PKS, see [Creating Service Accounts in GCP for PKS](#).

Installing PCF on GCP Using Terraform

This topic explains how to install PCF on GCP using Terraform.

Overview

You can install PCF on GCP with either the Pivotal Application Service (PAS) or Pivotal Container Service (PKS) runtime.

To install PCF on GCP using Terraform, do one of the following:

- Install with PAS. See [Install on PAS Using Terraform](#).
- Install with PKS. See [Install on PKS Using Terraform](#).

Install with PAS Using Terraform

To install with PAS using Terraform, do the following:

1. Deploy Ops Manager. See [Deploying Ops Manager on GCP Using Terraform](#).
2. Configure BOSH Director. See [Configuring BOSH Director on GCP Using Terraform](#).
3. (Optional) Configure a shared VPC. See [\(Optional\) Configuring a Shared VPC on GCP](#).
4. Configure PAS. See [Configuring PAS](#).

Install with PKS Using Terraform

To install with PKS using Terraform, do the following:

1. Deploy Ops Manager. See [Deploying Ops Manager on GCP Using Terraform](#).
2. Configure BOSH Director. See [Configuring BOSH Director on GCP Using Terraform](#).
3. (Optional). Configure a shared VPC. See [\(Optional\) Configuring a Shared VPC on GCP](#).
4. Configure PKS. See [Installing PKS on GCP](#).

Deploying Ops Manager on GCP Using Terraform

This guide describes the preparation steps required to install Pivotal Cloud Foundry (PCF) on Google Cloud Platform (GCP) using Terraform templates.

The Terraform template for PCF on GCP describes a set of GCP resources and properties. For more information about how Terraform creates resources in GCP, see the [Google Cloud Provider](#) topic on the Terraform site.

You may also find it helpful to review different deployment options in the [Reference Architecture for Pivotal Cloud Foundry on GCP](#).

Prerequisites

In addition to fulfilling the prerequisites listed in the [Installing Pivotal Cloud Foundry on GCP](#) topic, ensure you have the following:

- The [Terraform CLI](#)
- The [Google Cloud SDK](#)
- In your GCP project, enable the following APIs:
 - [Identity and Access Management](#)
 - [Cloud Resource Manager](#)
 - [Cloud DNS](#)
 - [Cloud SQL API](#)
 - [Compute Engine API](#)

Step 1: Obtain a GCP Service Account Key File

To use the Terraform templates to create the necessary infrastructure resources for PCF, you need a service account key file.

To create an account key file, follow the procedure below corresponding to your own use case.

- I already have a service account I want to use:

1. Navigate to the GCP console.
2. Select IAM and locate your service account.
3. From the Options column, open the dropdown and click Create Key.

- I want to create a new service account:

1. Open a terminal window.
2. To create a service account using the gcloud CLI, run the following command:

```
gcloud iam service-accounts create ACCOUNT-NAME
```

3. To create a key file for your service account, run the following command:

```
gcloud iam service-accounts keys create "terraform.key.json" --iam-account "ACCOUNT-NAME@PROJECT-ID.iam.gserviceaccount.com"
```

4. To bind the service account to your project and give it the owner role, run the following command:

```
gcloud projects add-iam-policy-binding PROJECT-ID --member 'serviceAccount:ACCOUNT-NAME@PROJECT-ID.iam.gserviceaccount.com' --role 'roles
```

Where:

- `ACCOUNT-NAME` is the name you want to apply to the new account.
- `PROJECT-ID` is your Google Cloud Platform Project ID.

Step 2: Download Templates and Edit Variables File

Before you can run Terraform commands to provision infrastructure resources, you must download the GCP Terraform Templates and create a Terraform

template variables file as described below:

1. On [Pivotal Network](#), navigate to the Pivotal Application Service (PAS) release.
2. Download the GCP Terraform ZIP file.
3. Extract the contents of the ZIP file.
4. Move the extracted folder to the `workspace` directory on your local machine.
5. On the command line, navigate to the directory. For example:

```
$ cd ~/workspace/pivotal-cf-terraforming-gcp
```

6. Navigate to the `terraforming-pas` or `terraforming-pks` directory that contains the Terraform files for your runtime.
7. In the runtime directory, create a text file named `terraform.tfvars`.
8. Open the `terraform.tfvars` file and add the following:

 **Note:** Ensure that you insert a new line at the end of the file.

```
env_name      = "YOUR-ENVIRONMENT-NAME"
opsman_image_url = "YOUR-OPS-MAN-IMAGE-URL"
region        = "YOUR-GCP-REGION"
zones         = ["YOUR-AZ-1", "YOUR-AZ-2", "YOUR-AZ-3"]
project       = "YOUR-GCP-PROJECT"
dns_suffix    = "YOUR-DNS-SUFFIX"

ssl_cert = <<SSL_CERT
-----BEGIN CERTIFICATE-----
YOUR-CERTIFICATE
-----END CERTIFICATE-----
SSL_CERT

ssl_private_key = <<SSL_KEY
-----BEGIN EXAMPLE RSA PRIVATE KEY-----
YOUR-PRIVATE-KEY
-----END EXAMPLE RSA PRIVATE KEY-----
SSL_KEY

service_account_key = <<SERVICE_ACCOUNT_KEY
YOUR-KEY-JSON
SERVICE_ACCOUNT_KEY
```

9. Edit the values in the file according to the table below.

Value to replace	Guidance
<code>YOUR-ENVIRONMENT-NAME</code>	Enter a name to use to identify resources in GCP. Terraform prepends the names of the resources it creates with this environment name. Example: <code>pcf</code> .
<code>YOUR-OPS-MAN-IMAGE-URL</code>	Enter the source URL of the Ops Manager image you want to boot. You can find this URL in the PDF included with the Ops Manager release on Pivotal Network .
<code>YOUR-GCP-REGION</code>	To prevent the creation of an Ops Manager VM, set this value to an empty string (<code>" "</code>). When using Platform Automation, you must disable the creation of the Ops Manager VM from Terraform. For more information, see Platform Automation .
<code>YOUR-AZ-1</code> <code>YOUR-AZ-2</code> <code>YOUR-AZ-3</code>	Enter the name of the GCP region in which you want Terraform to create resources. Example: <code>us-central1</code> .
<code>YOUR-GCP-PROJECT</code>	Enter the name of the GCP project in which you want Terraform to create resources.
<code>YOUR-DNS-SUFFIX</code>	Enter a domain name to use as part of the system domain for your PCF deployment. Terraform creates DNS records in GCP using <code>YOUR-ENVIRONMENT-NAME</code> and

	<p><code>YOUR-DNS-SUFFIX</code>. For example, if you enter <code>example.com</code> for your DNS suffix and have <code>pcf</code> as your environment name, Terraform creates DNS records at <code>pcf.example.com</code>. Enter a certificate to use for HTTP load balancing. For production environments, use a certificate from a Certificate Authority (CA). For test environments, you can use a self-signed certificate.</p> <p>Your certificate must specify your system domain as the common name. Your system domain is <code>YOUR-ENVIRONMENT-NAME.YOUR-DNS-SUFFIX</code>. It also must include the following subdomains: <code>*.sys.YOUR-SYSTEM-DOMAIN</code>, <code>*.login.sys.YOUR-SYSTEM-DOMAIN</code>, <code>*.uaa.sys.YOUR-SYSTEM-DOMAIN</code>, <code>*.apps.YOUR-SYSTEM-DOMAIN</code>.</p>
<code>YOUR-CERTIFICATE</code>	
<code>YOUR-PRIVATE-KEY</code>	Enter a private key for the certificate you entered.
<code>YOUR-KEY-JSON</code>	Enter the contents of your service account key file. This file is in JSON format.

Step 3: Add Optional Variables

Complete this step if you want to do any of the following:

- Change the default CIDR ranges
- Deploy the Isolation Segment tile
- Use an external Google Cloud SQL database
- Use external Google Storage buckets
- Disable generated GCP service account key for blobstore

In your `terraform.tfvars` file, specify the appropriate variables from the sections below.

 **Note:** You can see the configurable options by opening the `variables.tf` file and looking for variables with default values.

CIDR Ranges for Subnets

If you want to change the CIDR ranges for the infrastructure, your runtime, or services networks that Terraform creates, add the following variables to your `terraform.tfvars` file, replacing `YOUR-INFRASTRUCTURE-CIDR`, `YOUR-RUNTIME-CIDR` and `YOUR-SERVICES-CIDR` with your desired values.

```
infrastructure_cidr = YOUR-INFRASTRUCTURE-CIDR
pas_cidr = YOUR-RUNTIME-CIDR
services_cidr = YOUR-SERVICES-CIDR
```

Isolation Segments

If you plan to deploy the Isolation Segment tile, add the following variables to your `terraform.tfvars` file, replacing `YOUR-CERTIFICATE` and `YOUR-PRIVATE-KEY` with a certificate and private key. This causes Terraform to create an additional HTTP load balancer across three availability zones to use for the Isolation Segment tile.

```
isolation_segment = true
iso_seg_ssl_cert = <<ISO_SEG_SSL_CERT
-----BEGIN CERTIFICATE-----
YOUR-CERTIFICATE
-----END CERTIFICATE-----
ISO_SEG_SSL_CERT
iso_seg_ssl_private_key = <<ISO_SEG_SSL_KEY
-----BEGIN EXAMPLE RSA PRIVATE KEY-----
YOUR-PRIVATE-KEY
-----END EXAMPLE RSA PRIVATE KEY-----
ISO_SEG_SSL_KEY
```

External Database

1. If you want to use an external Google Cloud SQL database for Ops Manager and Pivotal Application Service (PAS), add the following to your `terraform.tfvars` file:

```
external_database = true
```

2. If you want to specify a single host from which users can connect to the Ops Manager and runtime databases, add the following to your `terraform.tfvars` file.

```
opsman_sql_db_host = HOST-IP-ADDRESS  
pas_sql_db_host = HOST-IP-ADDRESS
```

Where `HOST-IP-ADDRESS` is your desired IP address(es).

External Storage Buckets

If you want to use Google Cloud Storage buckets for the PAS Cloud Controller, add the following to your `terraform.tfvars` file:

```
create_gcs_buckets = true
```

GCP Service Account Key for Blobstore

If you want to provide your own service account for blob storage instead of using a generated service account, add the following to your `terraform.tfvars` file:

```
create_blobstore_service_account_key = false
```

Step 4: Create GCP Resources with Terraform

Follow these steps to use the Terraform CLI to create resources on GCP:

1. From the directory that contains the Terraform files, run the following command to initialize the directory based on the information you specified in the `terraform.tfvars` file.

```
terraform init
```

2. Run the following command to create the execution plan for Terraform.

```
terraform plan -out=plan
```

3. To execute the plan from the previous step, run the following command:

```
terraform apply plan
```

 **Note:** It may take several minutes for Terraform to create all the resources in GCP.

Step 5: Create DNS Record

1. In a browser, navigate to the DNS provider for the DNS suffix you entered in your `terraform.tfvars` file.
2. Create a new NS (Name server) record for your PCF system domain. Your system domain is `YOUR-ENVIRONMENT-NAME.YOUR-DNS-SUFFIX`.
 - a. In this record, enter the name servers included in `env_dns_zone_name_servers` from your Terraform output.

Next Steps

After you complete this procedure, follow the instructions in the [Configuring BOSH Director on GCP Using Terraform](#) topic.

Configuring BOSH Director on GCP Using Terraform

This topic describes how to configure the BOSH Director for Pivotal Cloud Foundry (PCF) on Google Cloud Platform (GCP) after [Preparing to Deploy Ops Manager on GCP Using Terraform](#).

 **Note:** You can also perform the procedures in this topic using the Ops Manager API. For more information, see the [Using the Ops Manager API](#) topic.

Prerequisite

To complete the procedures in this topic, you must have access to the output from when you ran `terraform apply` to create resources for this deployment. You can view this output at any time by running `terraform output`. You use the values in your Terraform output to configure the BOSH Director tile.

Step 1: Access Ops Manager

1. In a web browser, navigate to the fully qualified domain name (FQDN) of the BOSH Director. Use the `ops_manager_dns` value from running `terraform output`.

 **Note:** In order to set up Ops Manager authentication correctly, Pivotal recommends using a Fully Qualified Domain Name (FQDN) to access Ops Manager. Using an ephemeral IP address to access Ops Manager can cause authentication errors upon subsequent access.

1. When Ops Manager starts for the first time, you must choose one of the following:

- o [Internal Authentication](#): If you use internal authentication, Ops Manager maintains your user database.
- o [SAML Identity Provider](#): If you use a SAML Identity Provider (IdP), an external identity server maintains your user database.
- o [LDAP Server](#): If you use an LDAP Server, an external identity server maintains your user database.

Internal Authentication

1. When redirected to the [Internal Authentication](#) page, do the following:
 - o Enter a **Username**, **Password**, and **Password confirmation** to create an Admin user.
 - o Enter a **Decryption passphrase** and the **Decryption passphrase confirmation**. This passphrase encrypts the Ops Manager datastore, and is not recoverable.
 - o If you are using an **HTTP proxy** or **HTTPS proxy**, follow the instructions in the [Configuring Proxy Settings for the BOSH CPI](#) topic.
 - o Read the [End User License Agreement](#), and select the checkbox to accept the terms.
 - o Click **Setup Authentication**.
2. Log in to Ops Manager with the Admin username and password you created in the previous step.

SAML Identity Provider

1. Log in to your IdP console and download the IdP metadata XML. Optionally, if your IdP supports metadata URL, you can copy the metadata URL instead of the XML.
2. Copy the IdP metadata XML or URL to the Ops Manager [SAML Identity Provider](#) login page.

 **Note:** The same IdP metadata URL or XML is applied for the BOSH Director. If you use a separate IdP for BOSH, copy the metadata XML or URL from that IdP and enter it into the BOSH IdP Metadata text box in the Ops Manager login page.

3. Enter values for the fields listed below. Failure to provide values in these fields results in a `500` error.
 - o **SAML admin group**: Enter the name of the SAML group that contains all Ops Manager administrators. This field is case-sensitive.
 - o **SAML groups attribute**: Enter the groups attribute tag name with which you configured the SAML server. This field is case-sensitive.
4. Enter your **Decryption passphrase**. Read the [End User License Agreement](#), and select the checkbox to accept the terms.

5. Your Ops Manager login page appears. Enter your username and password. Click **Login**.
6. Download your SAML Service Provider metadata (SAML Relying Party metadata) by navigating to the following URLs:
 - o 6a. Ops Manager SAML service provider metadata: <https://OPS-MAN-FQDN:443/uaa/saml/metadata>
 - o 6b. BOSH Director SAML service provider metadata: <https://BOSH-IP-ADDRESS:8443/saml/metadata>

Note: To retrieve your `BOSH-IP-ADDRESS`, navigate to the **Status** tab in the **BOSH Director** tile. Record the **BOSH Director** IP address.
7. Configure your IdP with your SAML Service Provider metadata. Import the Ops Manager SAML provider metadata from Step 6a above to your IdP. If your IdP does not support importing, provide the values below.
 - o **Single sign on URL:** <https://OPS-MAN-FQDN:443/uaa/saml/SSO/alias/OPS-MAN-FQDN>
 - o **Audience URI (SP Entity ID):** <https://OP-MAN-FQDN:443/uaa>
 - o **Name ID:** Email Address
 - o SAML authentication requests are always signed
8. Import the BOSH Director SAML provider metadata from Step 6b to your IdP. If the IdP does not support an import, provide the values below.
 - o **Single sign on URL:** <https://BOSH-IP:8443/saml/SSO/alias/BOSH-IP>
 - o **Audience URI (SP Entity ID):** <https://BOSH-IP:8443>
 - o **Name ID:** Email Address
 - o SAML authentication requests are always signed
9. Return to the **BOSH Director** tile, and continue with the configuration steps below.

LDAP Server

- For **Server URL**, enter the URL that points to your LDAP server. With multiple LDAP servers, separate their URLs with spaces. Each URL must include one of the following protocols:
 - o `ldap://`: This specifies that the LDAP server uses an unencrypted connection.
 - o `ldaps://`: This specifies that the LDAP server uses SSL for an encrypted connection and requires that the LDAP server holds a trusted certificate or that you import a trusted certificate to the JVM truststore.
 - For **LDAP Username** and **LDAP Password**, enter the LDAP Distinguished Name (DN) and the password for binding to the LDAP Server. Example DN: `cn=administrator,ou=Users,dc=example,dc=com`
- Note:** Pivotal recommends that you provide LDAP credentials that grant read-only permissions on the LDAP Search Base and the LDAP Group Search Base. In addition to this, if the bind user belongs to a different search base, you must use the full DN.
- warning:** Pivotal recommends against reusing LDAP service accounts across environments. LDAP service accounts should not be subject to manual lockouts, such as lockouts that result from users utilizing the same account. Also, LDAP service accounts should not be subject to automated deletions, since disruption to these service accounts could prevent user logins.

- For **User Search Base**, enter the location in the LDAP directory tree from which any LDAP User search begins. The typical LDAP Search Base matches yourdomain name.

For example, a domain named “cloud.example.com” typically uses the following LDAP User Search Base: `ou=Users,dc=example,dc=com`

- For **User Search Filter**, enter a string that defines LDAP User search criteria. These search criteria allow LDAP to perform more effective and efficient searches. For example, the standard LDAP search filter `cn=Smith` returns all objects with a common name equal to `Smith`.

In the LDAP search filter string that you use to configure your runtime, use `{0}` instead of the username. For example, use `cn={0}` to return all LDAP objects with the same common name as the username.

In addition to `cn`, other attributes commonly searched for and returned are `mail`, `uid` and, in the case of Active Directory, `sAMAccountName`.

Note: For instructions for testing and troubleshooting your LDAP search filters, see [Configuring LDAP Integration with Pivotal Cloud Foundry](#) in the Pivotal Support Knowledge Base.

- For **Group Search Base**, enter the location in the LDAP directory tree from which the LDAP Group search begins.

For example, a domain named “cloud.example.com” typically uses the following LDAP Group Search Base: `ou=Groups,dc=example,dc=com`

- For **Group Search Filter**, enter a string that defines LDAP Group search criteria. The standard value is `member={0}`.
- For **Email Attribute**, enter the attribute name in your LDAP directory that corresponds to the email address in each user record, for example `mail`.
- For **LDAP RBAC Admin Group Name**, enter the DN of the LDAP group you want to have admin permissions in Ops Manager.
- From the dropdown, select how the UAA handles LDAP server referrals out to other external user stores. The UAA can:
 - Automatically follow any referrals.
 - Ignore referrals and return partial result.
 - Throw exception for each referral and abort.
- For **Server SSL Cert**, paste in the root certificate from your CA certificate or your self-signed certificate.
- Enter a **Decryption passphrase** and the **Decryption passphrase confirmation**. This passphrase encrypts the Ops Manager datastore, and is not recoverable.
- If you are using an **HTTP proxy** or **HTTPS proxy**, follow the instructions in [Configuring Proxy Settings for the BOSH CPI](#).
- Read the **End User License Agreement**, and select the checkbox to accept the terms.
- Select **Provision an admin client in the BOSH UAA**. You can use this to enable BOSH automation with scripts and tooling. For more information, see [Provision Admin Client](#) in *Creating UAA Clients for BOSH Director*.
- Click **Setup Authentication**.
- Return to the **BOSH Director** tile, and continue with the configuration steps below.

Step 2: Google Cloud Platform Config

1. Click the **Google Cloud Platform** tile within the **Installation Dashboard**.

2. Select **Google Config**. Complete the following fields:

- Project ID:** Enter the value of `project` from your `terraform.tfvars` file.
- Default Deployment Tag:** Enter the value of `env_name` from your `terraform.tfvars` file.
- Select **AuthJSON** and in the field below enter the contents of the JSON file for your service account key.

Note: As an alternative, you can select **The Ops Manager VM Service Account** option to use the service account automatically created by GCP for the Ops Manager VM.

Google Cloud Platform Config

Project ID*
your-gcp-project-id

Default Deployment Tag
pcf

The Ops Manager VM Service Account
 AuthJSON

Save

3. Click **Save**.

Step 3: Director Config Page

1. In Ops Manager, select **Director Config**.

Director Config

NTP Servers (comma delimited)*

Bosh HM Forwarder IP Address

Enable VM Resurrector Plugin

Enable Post Deploy Scripts

Recreate All VMs
 This will force BOSH to recreate all VMs on the next deploy. Persistent disk will be preserved.

Recreate All Persistent Disks
 Check this box to recreate all Persistent Disks for the Director and all other Tiles.

Enable bosh deploy retries
 This will attempt to re-deploy a failed deployment up to 5 times.

2. In the **NTP Servers (comma delimited)** field, enter `169.254.169.254` to designate the `metadata.google.internal` host as NTP server.

Note: The NTP server configuration only updates after VM recreation. Ensure that you select the **Recreate all VMs** checkbox if you modify the value of this field.

3. Leave the **Bosh HM Forwarder IP Address** field blank.

Note: Starting in PAS v2.0, BOSH-reported component metrics are available in the Loggregator Firehose by default. If you continue to use the BOSH HM Forwarder to consume these component metrics, you may receive duplicate data. To prevent this, leave the **Bosh HM Forwarder IP Address** field blank.

4. Select the **Enable VM Resurrector Plugin** checkbox to enable the BOSH Resurrector functionality and increase your runtime availability.
5. Select **Enable Post Deploy Scripts** to run a post-deploy script after deployment. This script allows the job to execute additional commands against a deployment.
- Note:** If you intend to install PKS, you must enable post-deploy scripts.
6. Select **Recreate all VMs** to force BOSH to recreate all VMs on the next deploy. This process does not destroy any persistent disk data.
7. Select **Recreate All Persistent Disks** to force BOSH to migrate and recreate persistent disks for the BOSH Director and all tiles. This process does not destroy any persistent disk data.
8. Select **Enable bosh deploy retries** to instruct Ops Manager to retry failed BOSH operations up to five times.
9. Select **Skip Director Drain Lifecycle** to prevent drain scripts from running when the BOSH Director is recreated.
10. Select **Store BOSH Job Credentials on tmpfs (beta)** to store credentials for BOSH jobs on temporary file storage (tmpfs) memory, rather than on disk. You must recreate all VMs for this setting to take effect.
11. Select **Keep Unreachable Director VMs** if you want to preserve BOSH Director VMs after a failed deployment for troubleshooting purposes.
12. Select **HM Pager Duty Plugin** to enable Health Monitor integration with PagerDuty.

HM Pager Duty Plugin

Service Key*

YOUR-PAGERDUTY-SERVICE-KEY

HTTP Proxy

YOUR-HTTP-PROXY

- **Service Key:** Enter your API service key from PagerDuty.
- **HTTP Proxy:** Enter an HTTP proxy for use with PagerDuty.

13. Select **HM Email Plugin** to enable Health Monitor integration with email.

HM Email Plugin

Host*

smtp.example.com

Port*

25

Domain*

cloudfoundry.example.com

From*

user2@example.com

Recipients*

user@example.com, user1@example.com

Username

user

Password

.....

Enable TLS

- **Host:** Enter your email hostname.
- **Port:** Enter your email port number.
- **Domain:** Enter your domain.
- **From:** Enter the address for the sender.
- **Recipients:** Enter comma-separated addresses of intended recipients.
- **Username:** Enter the username for your email server.
- **Password:** Enter the password for your email server.
- **Enable TLS:** Select this checkbox to enable Transport Layer Security to the email host.

14. For **CredHub Encryption Provider**, you can choose whether BOSH CredHub stores its encryption key internally on the BOSH Director and CredHub VM, or in an external hardware security module (HSM). The HSM option is more secure.

Before configuring an HSM encryption provider in the **Director Config** pane, you must follow the procedures and collect information described in [Preparing CredHub HSMs for Configuration](#).

 **Note:** After you deploy Ops Manager with an HSM encryption provider, you cannot change BOSH CredHub to store encryption keys internally.

CredHub Encryption Provider

Internal
 Luna HSM

Encryption Key Name*

Provider Partition*

Provider Partition Password*

Provider Client Certificate*

Provider Client Certificate Private Key*

HSM Host Address*

HSM Port Address*

Partition Serial Number*

HSM Certificate*

- **Internal:** Select this option for internal CredHub key storage. This option is selected by default and requires no additional configuration.
 - **Luna HSM:** Select this option to use a SafeNet Luna HSM as your permanent CredHub encryption provider, and fill in the following fields:
 1. **Encryption Key Name:** Any name to identify the key that the HSM uses to encrypt and decrypt the CredHub data. Changing this key name after you deploy Ops Manager can cause service downtime.
 2. **Provider Partition:** The partition that stores your encryption key. Changing this partition after you deploy Ops Manager could cause service downtime. For this value and the ones below, use values gathered in [Preparing CredHub HSMs for Configuration](#).
 3. **Provider Partition Password**
 4. **Provider Client Certificate:** The certificate that validates the identity of the HSM when CredHub connects as a client.
 5. **Provider Client Certificate Private Key**
 6. **HSM Host Address**
 7. **HSM Port Address:** If you do not know your port address, enter `1792`.
 8. **Partition Serial Number**
 9. **HSM Certificate:** The certificate that the HSM presents to CredHub to establish a two-way mTLS connection.
15. Select a **Blobstore Location** to either configure the blobstore as an internal server or an external endpoint. Because the internal server is unscalable and less secure, Pivotal recommends that you configure an external blobstore.

Note: After you deploy Ops Manager, you cannot change the blobstore location.

- **Internal:** Select this option to use an internal blobstore. Ops Manager creates a new VM for blob storage. No additional configuration is required.
 - **Enable TLS:** Select this checkbox to enable TLS to the blobstore.

Note: If you are using PASW 2016, make sure you have downloaded [Windows stemcell v1709.10 or higher](#) before enabling TLS.

- **S3 Compatible Blobstore:** Select this option to use an external S3-compatible endpoint.

Blobstore Location

Internal

Enable TLS

S3 Compatible Blobstore

S3 Endpoint*

Bucket Name*

Access Key*

Secret Key*

S3 Signature Version

V2 Signature

V4 Signature

Region*

S3 Backup Strategy

No backups

Use a versioned bucket

Copy into an additional bucket

Backup Bucket Region*

Backup Bucket Name*

Follow the procedures in [Sign up for Amazon S3](#) and [Creating a Bucket](#) in the AWS documentation. When you have created an S3 bucket, complete the following steps:

1. **S3 Endpoint:** Navigate to the [Regions and Endpoints](#) topic in the AWS documentation.
 - a. Locate the endpoint for your region in the **Amazon Simple Storage Service (S3)** table and construct a URL using your region's endpoint. For example, if you are using the `us-west-2` region, the URL you create would be `https://s3-us-west-2.amazonaws.com`. Enter this URL into the **S3 Endpoint** field.
 - b. On a command line, run `ssh ubuntu@OPS-MANAGER-FQDN` to SSH into the Ops Manager VM. Replace `OPS-MANAGER-FQDN` with the fully qualified domain name of Ops Manager.
 - c. Copy the custom public CA certificate you used to sign the S3 endpoint into `/etc/ssl/certs` on the Ops Manager VM.

- d. On the Ops Manager VM, run `sudo update-ca-certificates -f -v` to import the custom CA certificate into the Ops Manager VM truststore.

Note: You must also add this custom CA certificate into the **Trusted Certificates** field in the **Security** page. See [Security Page](#) for instructions.

2. **Bucket Name:** Enter the name of the S3 bucket.
3. **Access Key** and **Secret Key:** Enter the keys you generated when creating your S3 bucket.
4. Select **V2 Signature** or **V4 Signature**. If you select **V4 Signature**, enter your **Region**.

Note: AWS recommends using Signature Version 4. For more information about AWS S3 Signatures, see [Authenticating Requests](#) in the AWS documentation.

5. **S3 Backup Strategy:** Select one of the following to configure whether and how to back up the BOSH Director's S3 blobstore:

- **No backups**
- **Use a versioned bucket:** This option uses blobstore bucket versioning to save backups. It requires an S3 blobstore that supports versioning.
- **Copy into an additional bucket:** This option saves blobstore backups to a separate bucket, which is useful with blobstores that do not support versioning. The option requires you to fill in **Backup Bucket Region** and **Backup Bucket Name**.

- **GCS Blobstore:** Select this option to use an external Google Cloud Storage (GCS) endpoint.

The screenshot shows a configuration form for a GCS Blobstore. At the top, there is a radio button labeled "GCS Blobstore". Below it, there are three input fields: "Bucket Name*" with a placeholder box, "Storage Class*" with a dropdown menu set to "Regional", and "Service Account Key*" with a large text area for pasting JSON content.

Follow the procedures in [Creating Storage Buckets](#) in the GCS documentation to create a GCS bucket. When you have created a GCS bucket, complete the following steps:

1. **Bucket Name:** Enter the name of your GCS bucket.
 2. **Storage Class:** Select the storage class for your GCS bucket. See [Storage Classes](#) in the GCP documentation for more information.
 3. **Service Account Key:** Enter the contents of the JSON file that you downloaded in the [Set up IAM Service Accounts](#) section of *Preparing to Deploy Ops Manager on GCP Manually*.
16. For **Database Location**, if you configured your `terraform.tfvars` file to create an external database for Ops Manager, select **External MySQL Database** and complete the fields below. Otherwise, select **Internal**.

- **Host:** Enter the value of `sql_db_ip` from your Terraform output.
- **Port:** Enter `3306`.
- **Username:** Enter the value of `opsman_sql_username` from your Terraform output.
- **Password:** Enter the value of `opsman_sql_password` from your Terraform output.
- **Database:** Enter the value of `opsman_sql_db_name` from your Terraform output.

In addition, if you selected **External MySQL Database**, you can fill out the following optional fields:

- **Enable TLS:** Selecting this checkbox enables TLS communication between the BOSH Director and the database.
- **TLS CA:** Enter the Certificate Authority for the TLS Certificate.
- **TLS Certificate:** Enter the client certificate for mutual TLS connections to the database.
- **TLS Private Key:** Enter the client private key for mutual TLS connections to the database.

Note: You must select **Enable TLS** for **Director Database** to configure the TLS-related fields.

- **Advanced DB Connection Options:** If you would like to provide additional options for the database, use this field to provide a JSON-formatted options string.

17. (Optional) Modify the **Director Workers** value, which sets the number of workers available to execute Director tasks. This field defaults to `5`.
18. (Optional) **Max Threads** sets the maximum number of threads that the BOSH Director can run simultaneously. Pivotal recommends that you leave the field blank to use the default value, unless doing so results in rate limiting or errors on your IaaS.
19. (Optional) To add a custom URL for your BOSH Director, enter a valid hostname in **Director Hostname**. You can also use this field to configure a load balancer in front of your BOSH Director. For more information, see [How to Set Up a Load Balancer in Front of Operations Manager Director](#) in the Pivotal Support Knowledge Base.

The screenshot shows a configuration form for a BOSH Director. It includes three input fields: 'Director Workers' with the value '5', 'Max Threads' (empty), and 'Director Hostname' with the value 'opsdirector.example.com'.

20. (Optional) To set a custom banner that users see when logging in to the Director using SSH, enter text in the **Custom SSH Banner** field.
21. (Optional) Enter your comma-separated custom **Identification Tags**. For example, `iaas:foundation1, hello:world`. You can use the tags to identify your foundation when viewing VMs or disks from your IaaS.

22. Click **Save**.

Step 4: Create Availability Zones Page

- Note:** Pivotal recommends at least three availability zones for a highly available installation.
- For configuring a PAS deployment with multiple availability zones, see [Reference Architecture for Pivotal Cloud Foundry on GCP](#).
 - For PKS best practices for distributing workloads across multiple availability zones and clusters, see [Maintaining Workload Uptime](#).

1. Select **Create Availability Zones**.
2. Use the **Add** button to add three availability zones corresponding to those listed in the `azs` field in your Terraform output.
3. Click **Save**.

Step 5: Create Networks Page

1. Select **Create Networks**.
2. Make sure **Enable ICMP checks** is not selected. GCP routers do not respond to ICMP pings.
3. Use the **Add Network** button to create the following three networks:

Note: To use a shared VPC network, enter the shared VPC host project name before the network name in the format `VPC-PROJECT-NAME/NETWORK-NAME/SUBNET-NAME/REGION-NAME`. For example, `vpc-project/opsmgr/central/us-central1`. For more information, see [Configuring a Shared VPC on GCP](#).

Infrastructure Network

Network Name	infrastructure
--------------	----------------

Google Network Name	Use the <code>network_name</code> , <code>infrastructure_subnet_name</code> , and <code>region</code> fields from your Terraform output to enter the name of the infrastructure network created by Terraform. The format is: <code>network_name/infrastructure_subnet_name/region</code>
CIDR	Enter the value of <code>infrastructure_subnet_cidrs</code> from your Terraform output.
Reserved IP Ranges	Enter the first <code>.1</code> through <code>.9</code> addresses from the CIDR. For example, if the CIDR is <code>192.168.101.0/26</code> , enter the range <code>192.168.101.1-192.168.101.9</code> .
DNS	<code>169.254.169.254</code>
Gateway	Enter the value of <code>infrastructure_subnet_gateway</code> from your Terraform output.
Availability Zones	Select all three availability zones.

Runtime Network

Network Name	Enter the name of the runtime that you intend to deploy in this environment. For example, <code>pas</code> or <code>pks</code> .
Google Network Name	Use the <code>network_name</code> , <code>pas_subnet_name</code> , and <code>region</code> fields from your Terraform output to enter the name of the runtime network created by Terraform. The format is: <code>network_name/pas_subnet_name/region</code>
CIDR	Enter the value of <code>pas_subnet_cidrs</code> from your Terraform output.
Reserved IP Ranges	Enter the first <code>.1</code> through <code>.9</code> addresses from the CIDR. For example, if the CIDR is <code>192.168.16.0/22</code> , enter the range <code>192.168.16.1-192.168.16.9</code> .
DNS	<code>169.254.169.254</code>
Gateway	Enter the value of <code>pas_subnet_gateway</code> from your Terraform output.
Availability Zones	Select all three availability zones.

Services Network

Network Name	<code>services</code>
Google Network Name	Use the <code>network_name</code> , <code>services_subnet_name</code> , and <code>region</code> fields from your Terraform output to enter the name of the services network created by Terraform. The format is: <code>network_name/services_subnet_name/region</code>
CIDR	Enter the value of <code>services_subnet_cidrs</code> from your Terraform output.
Reserved IP Ranges	Enter the first <code>.1</code> through <code>.9</code> addresses from the CIDR. For example, if the CIDR is <code>192.168.16.0/22</code> , enter the range <code>192.168.20.1-192.168.20.9</code> .
DNS	<code>169.254.169.254</code>
Gateway	Enter the value of <code>services_subnet_gateway</code> from your Terraform output.
Availability Zones	Select all three availability zones.

 **Note:** After you deploy Ops Manager, you add subnets with overlapping Availability Zones to expand your network. For more information about configuring additional subnets, see [Expanding Your Network with Additional Subnets](#).

Step 6: Assign AZs and Networks Page

1. Select **Assign AZs and Networks**.
2. Use the dropdown to select a **Singleton Availability Zone**. The BOSH Director installs in this Availability Zone.
3. Under **Network**, select the `infrastructure` network for your BOSH Director.
4. Click **Save**.

Step 7: Security Page

1. Select **Security**.

Security

Trusted Certificates

These certificates enable BOSH-deployed components to trust a custom root certificate.

Include OpsManager Root CA in Trusted Certs

Checking this box will place the OpsManager Root CA into the trusted certificate field. The Bosh Director will then place this CA into the trust store of every VM that it deploys.

Generate VM passwords or use single password for all VMs

Generate passwords

Use default BOSH password

Save

2. In **Trusted Certificates**, enter your custom certificate authority (CA) certificates to insert into your organization's certificate trust chain. This feature enables all BOSH-deployed components in your deployment to trust custom root certificates.

To enter multiple certificates, paste your certificates one after the other. For example, format your certificates like the following:

```
-----BEGIN CERTIFICATE-----
ABCDEF12345678ABCDEFGH12345678ABCDEFGH12345678AB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
BCDEF12345678ABCDEFGH12345678ABCDEFGH12345678ABB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
CDEFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABBB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
```

Note: If you want to use Docker registries to run PAS app instances in Docker containers, enter the certificate for your private Docker registry in this field. See [Using Docker Registries](#) for more information about running app instances in PAS using Docker registries.

- (Optional) Select the **Include OpsManager Root CA in Trusted Certs** checkbox to include the Ops Manager root CA in the Trusted Certificates field. BOSH Director includes this CA in the trust store of every VM that it deploys.
- Choose **Generate passwords** or **Use default BOSH password**. Pivotal recommends that you use the **Generate passwords** option for greater security.
- Click **Save**. To view your saved Director password, click the **Credentials** tab.

Step 8: Syslog Page

Settings

Syslog

Syslog settings make take up to 1 min to take effect.

Do you want to configure Syslog for OpsManager?

No
 Yes

Address*
 The address or host for the syslog server

Port*

Transport Protocol*

Enable TLS

Permitted Peer*

SSL Certificate*

Queue Size

Forward Debug Logs

Custom rsyslog Configuration

SAVE SYSLOG SETTINGS

1. Select **Syslog**.
2. (Optional) Select **Yes** to send BOSH Director system logs to a remote server.
3. Enter the IP address or DNS name for the remote server in **Address**.
4. Enter the port number that the remote server listens on in **Port**.
5. Select **TCP** or **UDP** from the **Transport Protocol** dropdown. This selection determines which transport protocol is used to send the logs to the remote server.
6. (Optional) Select the **Enable TLS** checkbox to send encrypted logs to remote server with TLS. After you select the checkbox, do the following:
 - a. Enter either the name or SHA1 fingerprint of the remote peer in **Permitted Peer**.
 - b. Enter the SSL certificate for the remote server in **SSL Certificate**.



Note: Pivotal strongly recommends that you enable TLS encryption when you are forwarding logs. Logs can contain sensitive

information, such as cloud provider credentials.

7. (Optional) Enter an integer in **Queue Size**. This value specifies the number of log messages held in the buffer. The default value is 100,000.
8. (Optional) Select the checkbox to **Forward Debug Logs** to an external source. This option is deselected by default. If you select it, you may generate a large amount of log data.
9. (Optional) Enter configuration details for rsyslog in the **Custom rsyslog Configuration** field. This field requires the rainerscript syntax.
10. Click **Save Syslog Settings**.

Step 9: Resource Config Page

1. Select **Resource Config**.
2. Verify that the **Internet Connected** checkbox for every job is checked. The Terraform templates do not provision a Network Address Translation (NAT) box for internet connectivity to your VMs, so instead VMs are provided with ephemeral public IP addresses to allow the jobs to reach the internet.

Note: If you want to provision a Network Address Translation (NAT) box to provide internet connectivity to your VMs instead of providing them with public IP addresses, deselect the **Internet Connected** checkboxes. For more information about using NAT in GCP, see the [GCP documentation](#).

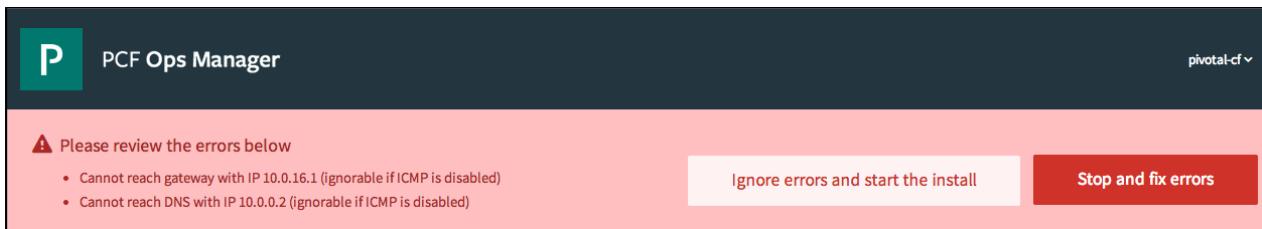
Note: If you install Pivotal Application Service for Windows (PASW), provision your **Master Compilation Job** with at least 100 GB of disk space.

Step 10: (Optional) Add Custom VM Extensions

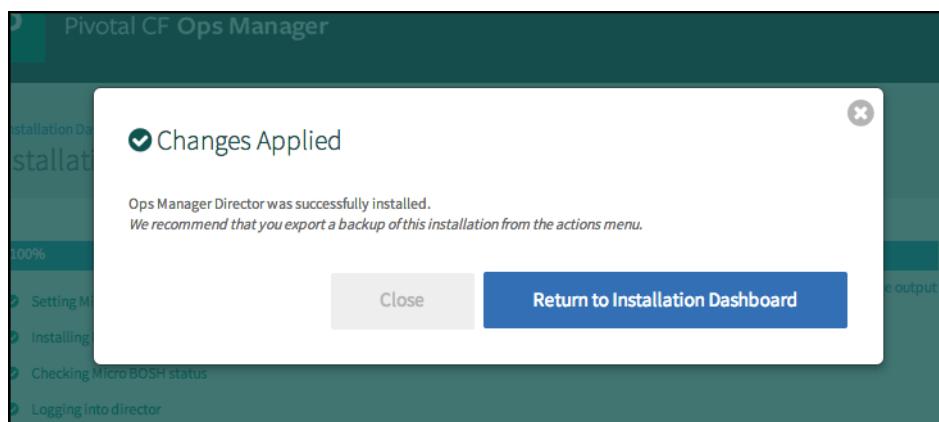
Use the Ops Manager API to add custom properties to your VMs such as associated security groups and load balancers. For more information, see [Managing Custom VM Extensions](#).

Step 11: Complete the BOSH Director Installation

1. Click the **Installation Dashboard** link to return to the Installation Dashboard.
2. Click **Review Pending Changes**, then **Apply Changes**. If the following ICMP error message appears, return to the **Network Config** screen, and make sure you have deselected the **Enable ICMP Checks** box. Click **Review Pending Changes**, then **Apply Changes** again.



3. BOSH Director installs. This may take a few moments. When the installation process successfully completes, the **Changes Applied** window appears.



Next Steps

After you complete this procedure, follow the instructions for the runtime you intend to install.

- To deploy PAS, see [Configuring PAS](#).
- To prepare to deploy PKS, see [Creating Service Accounts in GCP for PKS](#).

Configuring a Shared VPC on GCP

This guide describes the preparation steps required to configure and integrate a shared Virtual Private Cloud (VPC) on Google Cloud Platform (GCP) with Pivotal Cloud Foundry (PCF).

GCP Shared VPC, formerly known as Google Cross-Project Networking (XPN), enables you to assign GCP resources to individual projects within an organization but allows communication and shared services between projects. For more information about shared VPCs, see [Shared VPC Overview](#) in the GCP documentation.

Prerequisites

To configure a shared VPC, you must assign your project to a Cloud Organization. Confirm that you have a Cloud Organization associated with your GCP account using one of the following methods:

- **GCP Console:** From <https://console.cloud.google.com>, click the **Organization** dropdown at the top of the page to display all organizations you belong to.
- **gcloud Command Line Interface (CLI):** From the command line, run `gcloud organizations list` to display all organizations you belong to. See [gcloud Overview](#) in the Google documentation to install the gcloud CLI.

For more information, see [Creating and Managing Organizations](#) in the GCP documentation. If you do not have a Cloud Organization, contact GCP support.

Step 1: Provision the Shared VPC

Follow the [Enabling a shared VPC host project](#) procedure in the GCP documentation. This procedure requires shared VPC admin permissions.

Step 2: Create a Shared VPC Network

Use the procedures in the [Preparing to Deploy Ops Manager on GCP Manually](#) topic to create a new network with firewall rules. Do the following:

- [Step 3: Create a GCP Network with Subnet](#)
- [Step 5: Create Firewall Rules for the Network](#)

Step 3: Connect the Shared VPC to Ops Manager

You can use the [GCP console](#) or the gcloud CLI to connect the shared VPC host project with Ops Manager.

For more information, see [VPC Network Peering](#) in the GCP documentation.

 **warning:** VPC Network Peering is currently in beta and intended for evaluation and test purposes only.

Set Up VPC Network Peering with GCP Console

To set up VPC network peering with the GCP console, do the following:

1. From the GCP console, click **Networking**, then **VPC network peering**.

VPC network

VPC network peering

- VPC networks
- External IP addresses
- Firewall rules
- Routes
- VPC network peering**
- Shared VPC

VPC Network
VPC Network Peering

Cloud VPC Network Peering lets you privately connect two VPC networks, which can reduce latency, cost, and increase security. To get started click "Create connection". [Learn more](#)

[Create connection](#) [Learn more](#)

2. Click **Create Connection**.
3. Enter a name for the network connection from the Ops Manager project to the new shared network, such as `opsmanager-to-xpn`.

Networking

VPC Network Peering

Name	Your VPC network	Peered VPC network	Peered project ID	Status
<code>opsmanager-to-xpn</code>	opsmanager-pcf-network	xpn-project-network	cf-opsmanager-xpn-integration	Connected.

[CREATE PEERING CONNECTION](#) [REFRESH](#) [DELETE](#)

- VPC networks
- External IP addresses
- Firewall rules
- Routes
- Load balancing
- Cloud DNS
- VPN
- Cloud Routers
- VPC network peering**
- Cloud CDN
- Shared VPC

4. Click **Save**.
5. Click **Create Connection**.
6. Enter a name for the network connection from the new shared network to the Ops Manager project, such as `xpn-to-opsmanager`.

Name	Your VPC network	Peered VPC network	Peered project ID	Status
xpn-to-opsmanager	xpn-project-network	opsmanager-pcf-network	ops-manager-project	Connected.

- Click Save.

Set Up VPC Network Peering with gcloud CLI

To set up VPC network peering with the gcloud CLI, do the following:

- Enter the following command, replacing `OPSMANAGER-PROJECT` with the name of the project that contains your Ops Manager installation:

```
$ gcloud config set project OPSMANAGER-PROJECT
```

- Enter the following command to create a connection from the Ops Manager project to the new shared VPC project:

```
$ gcloud beta compute networks peerings create OPSMANAGER-TO-VPN \
--network OPSMANAGER-NETWORK \
--peer-project VPC-HOST-PROJECT \
--peer-network VPC-NETWORK \
--auto-create-routes
```

Replace the following text in the command above:

- `OPSMANAGER-TO-VPN`: Choose a name for the connection, such as `om-to-vpc`.
- `OPSMANAGER-NETWORK`: Enter the name of the network assigned to the Ops Manager project in GCP, such as `my-om-project`.
- `VPC-HOST-PROJECT`: Enter the name you gave the shared VPC project in [Step 1: Provision the Shared VPC](#).
- `VPC-NETWORK`: Enter the name of the network you gave the shared VPC project in [Step 2: Create Shared VPC Networks](#).

- Enter the following command, replacing `VPC-HOST-PROJECT` with the new shared VPC project you created in [Step 1: Provision the Shared VPC](#):

```
$ gcloud config set project VPC-HOST-PROJECT
```

- Enter the following command to create a connection from the new shared VPC project to the Ops Manager project:

```
$ gcloud beta compute networks peerings create VPC-TO-OPSMANAGER \
--network VPC-NETWORK \
--peer-project OPSMANAGER-PROJECT \
--peer-network OPSMANAGER-NETWORK \
--auto-create-routes
```

Replace the following text and run the following command:

- `VPC-TO-OPSMANAGER`: Choose a name for the connection, such as `vpc-to-om`.
- `VPC-NETWORK`: Enter the name of the network you gave the shared VPC project in [Step 2: Create Shared VPC Networks](#).
- `OPSMANAGER-PROJECT`: Enter the name of the project that contains your Ops Manager installation.

- OPSMANAGER-NETWORK : Enter the name of the network assigned to the Ops Manager project in GCP.

Step 4: Verify the Shared VPC Configuration

After configuring a shared VPC, use the following procedure to verify that the shared VPC host project VM appears in the Ops Manager project.

1. From <https://console.cloud.google.com>, select the Ops Manager project from the drop-down menu at the top of the page.
2. Click **Networking**, then **VPC networks**.

The screenshot shows the Google Cloud Platform interface for managing VPC networks. The left sidebar lists several options: VPC networks, External IP addresses, Firewall rules, Routes, VPC network peering (which is selected and highlighted in blue), and Shared VPC. The main pane is titled "VPC Network Peering" and contains a sub-section titled "VPC Network Peering". It provides a brief description: "Cloud VPC Network Peering lets you privately connect two VPC networks, which can reduce latency, cost, and increase security. To get started click "Create connection". [Learn more](#)". Below this are two buttons: "Create connection" and "Learn more".

3. Confirm that the shared VPC network name appears in the **Subnets** list.
4. Confirm that the shared VPC network **IP address ranges** match what you set for the new VPC project in [Step 2: Create a Shared VPC Network](#).

Deleting PCF from GCP

When you deploy [Pivotal Cloud Foundry](#) (PCF) to Google Cloud Platform (GCP), you provision a set of resources. This topic describes how to delete the resources associated with a PCF deployment.

You can delete the resources in one of two ways:

- If you created a separate project for your PCF deployment, perform the procedure in the [Delete the Project](#) delete the project.
- If the project that contains your PCF deployment also contains other resources that you want to preserve, perform the procedure in the [Delete PCF Resources](#) section.

Delete the Project

Perform the following steps to delete the project for your PCF deployment:

1. Navigate to the GCP Console Dashboard.
2. Under your **Project**, click **Manage project settings**.
3. Click **DELETE PROJECT**.
4. Enter your project ID and click **SHUT DOWN** to confirm.

Delete PCF Resources

Perform the following steps to delete the resources associated with your PCF deployment:

1. Navigate to the GCP Console Dashboard.
2. Click the upper left icon and select **Networking**.
3. Click **Load balancing**.
4. Perform the following steps for all load balancers associated with your PCF deployment:
 - a. Click the trashcan icon next to the load balancer.
 - b. In the next dialog, select any **health checks** and **backend services** associated with the load balancer.
 - c. Click **DELETE LOAD BALANCER AND THE SELECTED RESOURCES**.
5. Click the upper left icon and select **Compute Engine**.
6. Perform the following steps for **VM instances**, **Instance groups**, and **Disks**:
 - a. Select the checkbox next to the PCF resource.
 - b. When all PCF resources are selected, click **DELETE** in the upper right.
 - c. Click **DELETE** to confirm.
7. Click the upper left icon and select **Networking**.
8. Click **External IP addresses**.
9. Select all external IP addresses associated with your PCF deployment, and click **RELEASE STATIC ADDRESS**.
10. Click on **Networks**, and perform the following steps for any networks you created for PCF:
 - a. Click the name of the network.
 - b. Click **DELETE NETWORK**.
 - c. Click **DELETE** to confirm.
11. Click the upper left icon and select **IAM & Admin**.
12. Click the trashcan icon next to the **bosh** service account you created for PCF and click **REMOVE**.
13. Navigate to **Compute Engine > Metadata > SSH Keys** Delete the **vcap** SSH key that you created for the project.

Troubleshooting PCF on GCP

This topic describes how to troubleshoot known issues when deploying Pivotal Cloud Foundry (PCF) on Google Cloud Platform (GCP).

Problems Connecting with Single Sign-On (SSO)

Users may be unable to connect to applications running on PCF using SSO.

Explanation

SSO does not support multi-subnets.

Solution

Ensure that you have configured only one subnet. See [Preparing to Deploy Ops Manager on GCP Manually](#).

Uploading PAS Tile Causes Ops Manager Rails Application Crash

Uploading the Pivotal Application Service (PAS) tile causes the Ops Manager Rails application to crash.

Explanation

In compressed format, the PAS tile is 5 GB in size. However, when uncompressed during installation, the PAS tile requires additional disk space that can exhaust the space allocated to the boot disk.

Solution

Ensure that the boot disk is allocated at least 50 GB of space. See *Step 3: Create the Ops Manager VM Instance* in [Deploying Ops Manager on GCP Manually](#).

PAS Deployment Fails - MySQL Monitor replication-canary Job

During installation of the PAS tile, the replication-canary job fails to start. The error reported in the installation log resembles the following:

```
Started updating job mysql_monitor > mysql_monitor/0
(48e7ec82-3cdf-41af-9d0f-90d1f12683c8) (canary). Failed: 'mysql_monitor/0
(48e7ec82-3cdf-41af-9d0f-90d1f12683c8)' is not running after update.
Review logs for failed jobs: replication-canary (00:05:13)

Error 400007: 'mysql_monitor/0 (48e7ec82-3cdf-41af-9d0f-90d1f12683c8)'
is not running after update.
Review logs for failed jobs: replication-canary
```

Explanation

This error can appear as a result of incorrect configuration of network traffic and missed communication between the Gorouter and a load balancer.

Solution

1. Make sure you have selected the **Forward SSL to PAS Router** option in your [PAS Network Configuration](#).

2. Verify that you have configured the firewall rules properly and that TCP ports 80, 443, 2222, and 8080 are accessible on your GCP load balancers. See Step 5: *Create Firewall Rules for the Network* in [Preparing to Deploy Ops Manager on GCP Manually](#).
3. Verify that you have configured the proper SSL certificates on your HTTP(S) load balancer in GCP. See Step 8: *Create HTTP Load Balancer* in [Preparing to Deploy Ops Manager on GCP Manually](#).
4. If necessary, re-upload a new certificate and update any required SSL Certificate and SSH Key fields in your [PAS Network Configuration](#).

Insufficient External Database Permissions

Upgrade issues can be caused when the external database user used for the network policy DB is given insufficient permissions. To avoid this upgrade issue, ensure that the networkpolicyserver database user has the **ALL PRIVILEGES** permission.

Upgrading BOSH Director on GCP

This topic describes how to upgrade BOSH Director for Pivotal Cloud Foundry (PCF) on Google Cloud Platform (GCP).

Complete the tasks in this topic as part of the Ops Manager upgrade process. For more information, see [Upgrading Pivotal Cloud Foundry](#).

Overview

In this procedure, you create an Ops Manager VM instance to host the upgraded version of Ops Manager. Then, to complete the Ops Manager upgrade, you export your existing Ops Manager installation onto this new VM.

To create an Ops Manager VM instance, do the following tasks:

1. Locate the Ops Manager installation file. See [Locate the Pivotal Ops Manager Installation File](#).
2. Create a private VM image. See [Create a Private VM Image](#).
3. Create the Ops Manager VM. See [Create the Ops Manager VM Instance](#).

Prerequisites

To complete the Ops Manager upgrade, you must have your Ops Manager decryption passphrase. You defined this decryption passphrase during the initial installation of Ops Manager.

Locate the Pivotal Ops Manager Installation File

The Ops Manager installation file includes a filepath to the Ops Manager image. You use this image file to create a private VM image for the new Ops Manager VM.

To locate the Ops Manager installation file, do the following:

1. Log in to the [Pivotal Network](#), and click on **Pivotal Cloud Foundry Operations Manager**.
2. From the **Releases** drop-down, select the release for your upgrade.
3. Select one of the following download files:
 - o **Pivotal Cloud Foundry Ops Manager for GCP**
 - o **Pivotal Cloud Foundry Ops Manager YAML for GCP** When you click on the download link, your browser downloads or opens the `OpsManager_version_onGCP.pdf` or `OpsManager_version_onGCP.yml` file.

These documents provide the GCP location of the Ops Manager `.tar.gz` installation file based on the geographic location of your installation.

4. Copy the filepath string of the Ops Manager image based on your existing deployment location.

Create a Private VM Image

Create a private VM image to use when you create the new Ops Manager VM.

To create a private VM image, do the following:

1. Log in to the [GCP Console](#).
2. In the left navigation panel, click **Compute Engine**, and select **Images**.
3. Click **Create Image**.
4. Complete the following fields:

- **Name:** Enter a name that matches the naming convention of your existing Ops Manager image files.
- **Encryption:** Leave Automatic (recommended) selected.
- **Source:** Choose Cloud Storage file.
- **Cloud Storage file:** Paste in the Google Cloud Storage filepath you copied from the PDF or YAML file in the previous task.

The screenshot shows the 'Create an image' dialog box. On the left is a sidebar with various icons. The main area has fields for 'Name' (containing 'om-pcf'), 'Family (Optional)', 'Description (Optional)', 'Encryption' (set to 'Automatic (recommended)'), 'Source' (set to 'Cloud Storage file'), and 'Cloud Storage file' (a dropdown menu with a selected item). At the bottom are 'Create' and 'Cancel' buttons, and links for 'Equivalent REST or command line'.

5. Click **Create**. The file may take a few minutes to import.

Create the Ops Manager VM Instance

Create an Ops Manager VM instance to host the new version of Ops Manager.

To create an Ops Manager VM instance, do the following:

1. Select the checkbox for the image that you created above.

Images					CREATE IMAGE	REFRESH	CREATE INSTANCE	DEPRECATE	DELETE
<input type="text" value="Filter by label or name"/>					Columns	Labels			
Name	Size	Created by	Family	Creation time					
<input checked="" type="checkbox"/> om-pcf	50 GB	CF-Docs		Nov 22, 2016, 10:07:54 AM					

2. Click **Create Instance**.

3. In the **Create an instance form**, complete the following fields:

- **Name:** Enter a name that matches the naming conventions of your existing deployment.
- **Zone:** Choose a zone from the region of your existing deployment.
- **Machine type:** Click **Customize** to manually configure the vCPU and memory. An Ops Manager VM instance requires the following minimum specifications:

Machine Spec	Minimum Value
CPU	2 vCPUs
Memory	8 GB

- **Boot disk:** Click **Change**, then perform the following steps:

- Click **Custom images** if it is not already selected.
- Select the **Boot disk type**. If you have an Ops Manager environment with high performance needs, select **SSD**. As an example,

environments used to [develop PCF tiles](#) may benefit from a higher performing Ops Manager VM boot disk. For most environments, however, you can select **Standard**.

- Set the **Size (GB)** of the boot disk to the minimum or higher.

Machine Spec	Minimum Value
Boot disk	100 GB

- Select the Ops Manager image you created in the previous step if it is not already selected.

Boot disk

Select an image or snapshot to create a boot disk; or attach an existing disk.

OS images Application images **Custom images** Snapshots

Existing disks

om-pcf
Created from CF-Docs on Oct 18, 2016, 11:38:02 AM

- Click **Select** to save.
- Under **Identity and API access**, choose the **Service account** you created when you initially installed Pivotal Cloud Foundry. See *Step 1: Set up IAM Service Accounts* in [Preparing to Deploy Ops Manager on GCP Manually](#).
- Allow HTTP traffic:** Only select this checkbox if you selected it in your original Ops Manager VM configuration.
- Allow HTTPS traffic:** Only select this checkbox if you selected it in your original Ops Manager VM configuration.

Name [?](#)
om-pcf-1a

Zone [?](#)
us-central1-b

Machine type

Cores [Basic view](#)
2 vCPU 1 - 64

Memory
8 GB 1.8 - 13

Extend memory [?](#)

CPU platform [?](#)
Automatic

[▼ GPUs](#)

Choosing a machine type [?](#)

Boot disk [?](#)

New 100 GB standard persistent disk
Image om-pcf [Change](#)

Identity and API access [?](#)

Service account [?](#)
bosh

Access scopes [?](#)
Use IAM roles with service accounts to control VM access [Learn more](#)

Firewall [?](#)
Add tags and firewall rules to allow specific network traffic from the Internet

Allow HTTP traffic

Allow HTTPS traffic

[▼ Management, disks, networking, SSH keys](#)

You will be billed for this instance. [Learn more](#)

Create **Cancel**

- Networking: Select the **Networking** tab, and perform the following steps:

- For **Network** and **Subnetwork**, select the network and subnetwork you created when you initially deployed Pivotal Cloud Foundry. See [Step 3: Create a GCP Network with Subnets](#) in [Preparing to Deploy Ops Manager on GCP Manually](#).
- For **Network tags**, enter any tags that you applied to your original Ops Manager. For example, if you used the `pcf-opsmanager` tag to apply the firewall rule you created in [Step 5: Create Firewall Rules for the Network](#), then apply the same tag to this Ops Manager VM.
- For **Internal IP**, select `Custom`. In the **Internal IP address** field, enter a spare address located within the reserved IP range you configured in your existing BOSH Director as part of [Step 5: Create Networks Page](#). Do not use `10.0.0.1`, which is configured for the Gateway.
- For **External IP**, select **New static IP address....** In the next form, enter a name for the static IP. For example, `om-public-ip`. Click **Reserve**. In the **External IP** drop-down, select the static IP address you just reserved.

Firewall Add tags and firewall rules to allow specific network traffic from the Internet
 Allow HTTP traffic
 Allow HTTPS traffic

Management Disks **Networking** SSH Keys

Network opsmgr

Subnetwork opsmgr-subnet (10.0.0.0/20)

Network tags (Optional) pcf-opsmanager

Internal IP Custom
10.0.0.4

External IP Ephemeral
None
pbj-websockets-ip (10.0.0.1)
New static IP address...

[Less](#)

- Click **Create** to deploy the new Ops Manager VM. This may take a few moments.
- Navigate to your DNS provider, and modify the entry that points a fully qualified domain name (FQDN) to the Ops Manager VM. Replace the original Ops Manager static IP address with the public IP address of the new Ops Manager VM you created in a previous step.

warning: In order to set up Ops Manager authentication correctly, Pivotal recommends using a Fully Qualified Domain Name (FQDN) to access Ops Manager. Using an ephemeral IP address to access Ops Manager can cause authentication errors upon subsequent access. If you are importing older Ops Manager settings, you must use the FQDN.

Next Steps

After you complete this procedure, continue the upgrade instructions in [Upgrading Pivotal Cloud Foundry](#) topic.

Later on, if you need to SSH into the Ops Manager VM to perform diagnostic troubleshooting, see [SSH into Ops Manager](#).

Installing Pivotal Cloud Foundry on OpenStack

This guide describes how to install Pivotal Cloud Foundry (PCF) on OpenStack with Ops Manager and Pivotal Application Service (PAS).

Overview

OpenStack is a cloud operating system that controls large pools of compute, storage, and networking resources throughout a datacenter. For guidance on OpenStack service credential management, see [Open Stack Security Documents](#) below.

OpenStack Security Documents

These documents provide a general reference for OpenStack service credential management.

- [OpenStack credential configuration ↗](#)
- [OpenStack credential creation ↗](#)
- [OpenStack deployment configuration ↗](#)

Requirements

This section describes the requirements for installing PCF on OpenStack, including general requirements for installing PCF with Ops Manager and PAS as well as OpenStack requirements.

 **Note:** You can install PCF on OpenStack with the Pivotal Application Service (PAS) runtime. The Pivotal Container Service (PKS) runtime is not supported for OpenStack. For more information about PAS, see [PAS Concepts](#). For more information about PKS, see [Pivotal Container Service \(PKS\)](#).

General Resource Requirements

The following are general requirements for deploying and managing a PCF deployment with Ops Manager and Pivotal Application Service (PAS):

- A wildcard DNS record that points to your router or load balancer. Alternatively, you can use a service such as xip.io. For example, `203.0.113.0.xip.io`.
 - PAS gives each application its own hostname in your app domain.
 - With a wildcard DNS record, every hostname in your domain resolves to the IP address of your router or load balancer, and you do not need to configure an A record for each app hostname. For example, if you create a DNS record `*.example.com` pointing to your load balancer or router, every application deployed to the `example.com` domain resolves to the IP address of your router.
- At least one wildcard TLS certificate that matches the DNS record you set up above, `*.example.com`.
- Sufficient IP allocation:
 - One static IP address for either HAProxy or one of your goRouters
 - One static IP address for each job in the Ops Manager tile. See the Resource Config pane for each tile for a full list.
 - One static IP address for each job listed below:
 - Consul
 - NATS
 - File Storage
 - MySQL Proxy
 - MySQL Server
 - Backup Restore Node
 - HAProxy
 - Router
 - MySQL Monitor
 - Diego Brain
 - TCP Router
 - One IP for each VM instance created by the service.

- An additional IP address for each compilation worker. So the formula for total IPs needed is

`IPs needed = static IPs + VM instances + compilation workers`

 **Note:** Pivotal recommends that you allocate at least 36 dynamic IP addresses when deploying Ops Manager and PAS. BOSH requires additional dynamic IP addresses during installation to compile and deploy VMs, install PAS, and connect to services.

- One or more NTP servers if not already provided by your IaaS.
- (Recommended) A network without DHCP available for deploying the PAS VMs.

 **Note:** If you have DHCP, refer to the [Troubleshooting Guide](#) to avoid issues with your installation.

- (Optional) External storage. When you deploy PCF, you can select internal file storage or external file storage, either network-accessible or IaaS-provided, as an option in the PAS tile. Pivotal recommends using external storage whenever possible. See [Configure File Storage](#) for a discussion of how file storage location affects platform performance and stability during upgrades.
- (Optional) External databases. When you deploy PCF, you can select internal or external databases for the BOSH Director and for PAS. Pivotal recommends using external databases in production deployments. An external database must be configured to use the UTC timezone.
- (Optional) External user stores. When you deploy PCF, you can select a SAML user store for Ops Manager or a SAML or LDAP user store for PAS, to integrate existing user accounts.
- The most recent version of the [Cloud Foundry Command Line Interface](#) (cf CLI).

OpenStack Requirements

The following are OpenStack requirements for deploying PCF:

- PCF is supported on the OpenStack Ocata, Pike, Queens, Rocky, and Stein releases. OpenStack is a collection of inter-operable components and requires general OpenStack expertise to troubleshoot issues that may occur when installing Pivotal Cloud Foundry on particular releases and distributions. To verify that your OpenStack platform is compatible with PCF, use the OpenStack Validator tool. To access the OpenStack Validator tool, see [CF OpenStack Validator](#) on GitHub.
- Pivotal recommends granting complete access to the OpenStack logs to the operator managing the PCF installation process.
- For OpenStack accounts for PCF, Pivotal recommends following the principle of least privilege by scoping privileges to the most restrictive permissions possible for a given role.
- You must have a dedicated OpenStack project, formerly known as an OpenStack tenant.
- You must have Keystone access to the dedicated OpenStack project, including the following:
 - Auth URL
 - Username and password. The `PrimaryProject` for the user must be the project you want to use to deploy PCF. For more information, see [Manage projects and users](#) in the OpenStack documentation.
 - Project name
 - Region (with multiple availability zones if you require high availability)
 - SSL certificate for your wildcard domain (see below)
- You must have the ability to do the following in OpenStack:
 - Create and modify VM flavors
 - Enable DHCP if required
 - Create a network and then connect that network with a router to an external network
 - Create an external network with a pool of floating IP addresses
 - Boot VMs directly from image
 - Create two wildcard domains for separate system and app domains
- The following are resource requirements for the dedicated OpenStack project:
 - 118 GB of RAM
 - 22 available instances
 - 14 small VMs (1 vCPU, 1024 MB of RAM, 10 GB of root disk)
 - 2 high-CPU VMs (2 vCPU, 1024 MB of RAM, 10 GB of root disk)
 - 3 large VMs (4 vCPU, 16384 MB of RAM, 10 GB of root disk)
 - 3 extra-large VMs (8 vCPU, 16 GB of RAM, 160 GB of ephemeral disk)
 - 58 vCPUs
 - 1 TB of storage

- Nova or Neutron networking with floating IP support

 **Note:** By default, PAS deploys the number of VM instances required to run a highly available configuration of PCF. If you are deploying a test or sandbox PCF that does not require HA, then you can scale down the number of instances in your deployment. For information about the number of instances required to run a minimal, non-HA PCF deployment, see [Scaling PAS](#).

- The following are requirements for the OpenStack Cinder back end:
 - PCF requires RAW root disk images. The Cinder back end for your OpenStack project must support RAW.
 - Pivotal recommends that you use a Cinder back end that supports snapshots. This is required for some BOSH functionalities.
 - Pivotal recommends enabling your Cinder back end to delete block storage asynchronously. If this is not possible, it must be able to delete multiple 20 GB volumes within 300 seconds.
- The following are requirements for using an Overlay Network with VXLAN or GRE Protocols:
 - If an overlay network is being used with VXLAN or GRE protocols, the MTU of the created VMs must be adjusted to the best practices recommended by the plugin vendor (if any).
 - DHCP must be enabled in the internal network for the MTU to be assigned to the VMs automatically.
 - Review the [Configuring PAS](#) topic to adjust your MTU values.
 - Failure to configure your overlay network correctly could cause Apps Manager to fail since applications will not be able to connect to the UAA.

 **Note:** If you are using IPsec, your resource usage will increase by approximately 36 bytes. View the [Installing IPsec topic](#) for information, including setting correct MTU values.

Install PCF on OpenStack with PAS

To install PCF on OpenStack with the PAS runtime, do the following:

1. Deploy Ops Manager. See [Deploying Ops Manager on OpenStack](#).
2. Configure BOSH Director on OpenStack. See [Configuring BOSH Director on OpenStack](#).
3. Configure PAS. See [Configuring PAS](#).

Deploying Ops Manager to OpenStack

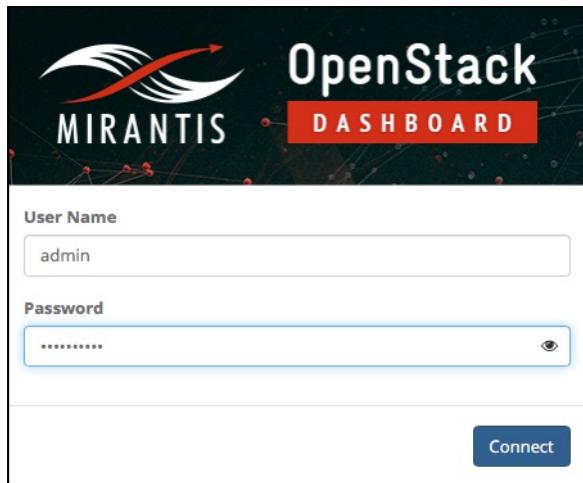
This guide describes how to provision the OpenStack infrastructure where you need to install [Pivotal Cloud Foundry](#). Use this topic when [Installing Pivotal Cloud Foundry on OpenStack](#).

After completing this procedure, complete all of the steps in the [Configuring BOSH Director on OpenStack](#) and [Configuring PAS](#) topics.

Note: This document uses Mirantis OpenStack for screenshots and examples. The screens of your OpenStack vendor configuration interface may differ.

Step 1: Log in to the OpenStack Horizon Dashboard

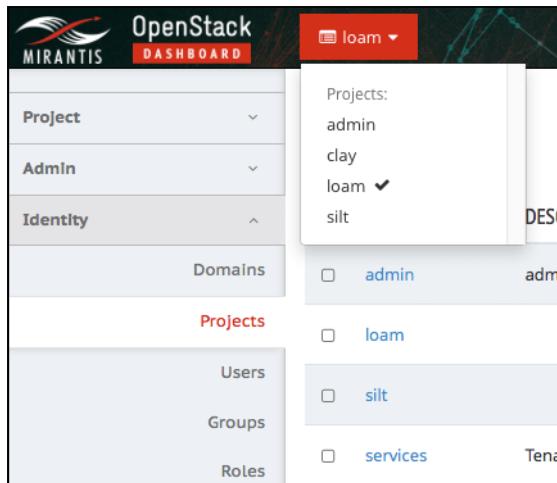
1. Log in to the OpenStack Horizon dashboard.



The screenshot shows the Mirantis OpenStack Horizon login interface. At the top, there's a logo for 'MIRANTIS' with a stylized red swoosh, followed by 'OpenStack DASHBOARD'. Below the header, there are two input fields: 'User Name' containing 'admin' and 'Password' with masked input. At the bottom right of the form is a blue 'Connect' button.

2. Click **Connect**.

3. From the OpenStack project list dropdown, set the active project by selecting the project where you will deploy PCF.



Step 2: Configure Security

warning: If you are using OpenStack Liberty or Mitaka, do not create the key pair with the OpenStack Horizon dashboard. Instead make sure that you generate the SSH key pair manually. For example, use the `ssh-keygen` command. Then follow the procedure below to import that key pair into OpenStack. This is due to an [OpenStack bug](#).

1. In the left navigation of your OpenStack Horizon dashboard, click Project > Compute > Access & Security.

2. Select the **Key Pairs** tab on the **Access & Security** page.
3. Click **Import Key Pair**.
4. Enter a **Key Pair Name** and the contents of your public key in the **Public Key** field.

Import Key Pair

Key Pair Name *
pcf

Public Key *
ssh-rsa
AA[REDACTED]

Description:
Key Pairs are how you login to your instance after it is launched.
Choose a key pair name you will recognise and paste your SSH public key into the space provided.
SSH key pairs can be generated with the ssh-keygen command:
`ssh-keygen -t rsa -f cloud.key`
This generates a pair of keys: a key you keep private (cloud.key) and a public key (cloud.key.pub). Paste the contents of the public key file here.
After launching an instance, you login using the private key (the username might be different depending on the image you launched):
`ssh -i cloud.key <username>@<instance_ip>`

Cancel **Import Key Pair**

5. Click **Import Key Pair**.
6. In the left navigation, click **Access & Security** to refresh the page. The new key pair appears in the list.
7. Select the **Security Groups** tab. Click **Create Security Group** and create a group with the following properties:

- o **Name:** opsmanager
- o **Description:** Ops Manager

Create Security Group

Name *
opsmanager

Description *
Ops Manager

Description:
Security groups are sets of IP filter rules that are applied to the network settings for the VM. After the security group is created, you can add rules to the security group.

Cancel **Create Security Group**

8. Select the checkbox for the `opsmanager` Security Group and click **Manage Rules**.

Security Groups Key Pairs Floating IPs API Access

Security Groups

[+ Create Security Group](#) [X Delete Security Groups](#)

<input type="checkbox"/>	Name	Description	Actions
<input type="checkbox"/>	default	default	Manage Rules
<input checked="" type="checkbox"/>	opsmanager	Ops Manager	Manage Rules ▼

Displaying 2 items

9. Add the following ingress access rules for HTTP, HTTPS, and SSH as shown in the table below. The rules with `opsmanager` in the Remote column have restricted access to that particular Security Group.

Note: Adjust the remote sources as necessary for your own security compliance. Pivotal recommends limiting remote access to Ops Manager to IP ranges within your organization.

Direction	Ether Type	IP Protocol	Port/Port Range	Remote
Ingress	IPv4	TCP	22 (SSH)	0.0.0.0/0 (CIDR)
Ingress	IPv4	TCP	80 (HTTP)	0.0.0.0/0 (CIDR)
Ingress	IPv4	TCP	443 (HTTPS)	0.0.0.0/0 (CIDR)
Ingress	IPv4	TCP	4222 (NATS)	opsmanager
Ingress	IPv4	TCP	6868 (BOSH Agent)	opsmanager
Ingress	IPv4	TCP	8844 (CredHub)	opsmanager
Ingress	IPv4	TCP	8853 (BOSH Health Monitor)	opsmanager
Ingress	IPv4	TCP	25250 (BOSH Blobstore)	opsmanager
Ingress	IPv4	TCP	25555 (BOSH Director)	opsmanager
Ingress	IPv4	TCP	25777 (BOSH Registry)	opsmanager
Egress	IPv4	TCP	1-65535	0.0.0.0/0 (CIDR)

10. Leave the existing default egress access rules as shown in the screenshot below.

Access & Security

/ Manage Security Group Rules: opsmanager (af1 [REDACTED])

[+ Add Rule](#) [Delete Rules](#)

<input type="checkbox"/>	DIRECTION	ETHER TYPE	IP PROTOCOL	PORT RANGE	REMOTE IP PREFIX	REMOTE SECURITY GROUP	ACTIONS
<input type="checkbox"/>	Egress	IPv4	Any	Any	0.0.0.0/0	-	Delete Rule
<input type="checkbox"/>	Egress	IPv6	Any	Any	::/0	-	Delete Rule
<input type="checkbox"/>	Ingress	IPv4	TCP	1 - 65535	-	opsmanager	Delete Rule
<input type="checkbox"/>	Ingress	IPv4	TCP	22 (SSH)	0.0.0.0/0	-	Delete Rule
<input type="checkbox"/>	Ingress	IPv4	TCP	80 (HTTP)	0.0.0.0/0	-	Delete Rule
<input type="checkbox"/>	Ingress	IPv4	TCP	443 (HTTPS)	0.0.0.0/0	-	Delete Rule
<input type="checkbox"/>	Ingress	IPv4	TCP	25555	0.0.0.0/0	-	Delete Rule
<input type="checkbox"/>	Ingress	IPv4	UDP	1 - 65535	-	opsmanager	Delete Rule

Displaying 8 items

Step 3: (Optional) Run the CF OpenStack Validator Tool

As an optional but recommended step, you can now run the CF OpenStack Validator tool against your OpenStack tenant to verify support for PCF.

1. Follow the directions for running the [CF OpenStack Validator Tool](#).
2. When configuring the CPI version used by the Validator, specify the OpenStack CPI version indicated in the [PCF Ops Manager Release Notes](#) for the PCF release that you are planning to deploy.

Troubleshooting the output of the CF OpenStack Validator tool is beyond the scope of this document.

Step 4: Create Ops Manager Image

You can create the Ops Manager image in OpenStack using the OpenStack Horizon dashboard.

 **Note:** If your Horizon Dashboard does not support file uploads, you must use the [Glance CLI](#) client.

To create an Ops Manager image in OpenStack, perform the following steps:

1. Download the Pivotal Cloud Foundry Ops Manager for OpenStack image file from [Pivotal Network](#).
2. In the left navigation of your OpenStack dashboard, click Project > Compute > Images.
3. Click **Create Image**. Complete the **Create An Image** page with the following information:
 - **Name:** Enter `Ops Manager`.
 - **Image Source:** Select **Image File**.
 - **Image File:** Click **Choose File**. Browse to and select the image file that you downloaded from [Pivotal Network](#).
 - **Format:** Select **Raw**.
 - **Minimum Disk (GB):** Enter `80`.
 - **Minimum RAM (MB):** Enter `8192`.
 - Deselect the **Public** checkbox.
 - Select the **Protected** checkbox.

Create An Image

Name *	Ops Manager	Description:
		Currently only images available via an HTTP/HTTPS URL are supported. The image location must be accessible to the Image Service.
Please note: The Image Location field MUST be a valid and direct URL to the image binary. URLs that redirect or serve error pages will result in unusable images.		
Image Source		
Image File		
Image File ? <input type="button" value="Choose File"/> pcf-openstack-1.12.2.raw		
Format * Raw		
Architecture		
Minimum Disk (GB) ? 80		
Minimum RAM (MB) ? 8192		
<input type="checkbox"/> Public <input checked="" type="checkbox"/> Protected		
<input type="button" value="Cancel"/>		<input type="button" value="Create Image"/>

4. Click **Create Image**.

Step 5: Launch Ops Manager VM

1. In the left navigation of your OpenStack dashboard, click **Project > Compute > Images**.

2. Click **Launch**.

Images								+ Create Image	x Delete Images
	Image Name	Type	Status	Public	Protected	Format	Size	Actions	
<input checked="" type="checkbox"/>	Ops Manager	Image	Active	No	No	RAW	3.0 GB	<input type="button" value="Launch"/>	<input type="button" value="Delete"/>
<input type="checkbox"/>	pcf-openstack-1.12.2	Image	Active	No	No	RAW	3.0 GB	<input type="button" value="Launch"/>	<input type="button" value="Delete"/>

3. In the **Details** tab, specify the following values:

- o **Instance Name:** Enter `Ops Manager`.
- o **Availability Zone:** Use the dropdown to select an availability zone. You specify this availability zone in the [Complete the Availability Zones Pages](#) step of [Configuring Ops Manager Director](#).
- o **Count:** Do not change from the default value of **1**.

Launch Instance

Details

Please provide the initial hostname for the instance, the availability zone where it will be deployed, and the instance count. Increase the Count to create multiple instances with the same settings.

Source

Instance Name * Ops Manager

Availability Zone nova

Count * 1

Total Instances (No Limit) 0 Current Usage
1 Added

Source

Flavor *

Networks

Network Ports

Security Groups

Key Pair

Configuration

Metadata

Cancel < Back Next > **Launch Instance**

4. In the **Source** tab, specify the following values:

- **Select Boot Source:** Select **Image**.
- **Create New Volume:** Leave **No** selected.
- **Allocated:** Make sure **Ops Manager** is selected.

Launch Instance

Source

Instance source is the template used to create an instance. You can use a snapshot of an existing instance, an image, or a volume (if enabled). You can also choose to use persistent storage by creating a new volume.

Select Boot Source Image **Create New Volume** Yes No

Allocated

NAME	UPDATED	SIZE	TYPE	VISIBILITY
Ops Manager	10/10/17 6:49 PM	4.39 GB	RAW	Private

Available 2

Select one

Available 2

Click here for filters.

Metadata

NAME	UPDATED	SIZE	TYPE	VISIBILITY
TestVM	9/28/17 1:50 AM	21.54 MB	QCOW2	Public
ubuntu-xenial	9/29/17 10:36 AM	277.13 MB	QCOW2	Public

Cancel < Back Next > **Launch Instance**

5. In the **Flavor** tab, configure the OpenStack VM flavors as follows:

Note: Do not change the names of the VM flavors.

ID	Name	Memory_MB	Disk	Ephemeral	VCPUs
1	m1.small	2048	20	0	1

2	m1.medium	4096	40	0	2
3	m1.large	8192	80	0	4
4	m1.xlarge	16384	160	0	8

6. In the **Networks** tab, select a private subnet. You add a Floating IP to this network in a later step.

Launch Instance

Details

Source

Flavor

Networks

Network Ports

Security Groups

Key Pair

Configuration

Metadata

Allocated 1

NETWORK	SUBNETS ASSOCIATED	SHARED	ADMIN STATE	STATUS
loam_net	No	Up	Active	-

Available 0

Select at least one network

Click here for filters.

NETWORK	SUBNETS ASSOCIATED	SHARED	ADMIN STATE	STATUS
No available items				

< Back | Next > | **Launch Instance**

7. Skip the **Network Ports** tab.

8. In the **Security Groups** tab, select the **opsmanager** security group that you created in [Step 2: Configure Security](#). Deselect all other Security Groups.

Launch Instance

Details *

Source

Flavor *

Networks

Network Ports

Security Groups

Key Pair

Configuration

Metadata

Allocated 1

NAME	DESCRIPTION
opsmanager	Ops Manager

Available 2

Select one or more

NAME	DESCRIPTION
default	Default security group
loam	loam Security Group

Filter

< Back | Next > | **Launch Instance**

9. In the **Key Pair** tab, select the key pair that you imported in [Step 2: Configure Security](#).

Launch Instance

Details	A key pair allows you to SSH into your newly created instance. You may select an existing key pair, import a key pair, or generate a new key pair. ?		
Source	+ Create Key Pair Import Key Pair		
Flavor	Allocated		
Networks	NAME	FINGERPRINT	
Network Ports	> pcf	9c [REDACTED]	-
Security Groups	Available 5 Select one		
Key Pair	<input type="text"/> Filter		
Configuration	NAME ▲	FINGERPRINT	
Metadata	> clay	9e [REDACTED]	+
	> clay-bosh	53 [REDACTED]	+
	> id_rsa_bosh	d8 [REDACTED]	+
	> loam	42 [REDACTED]	+
	> silt	39 [REDACTED]	+

10. Skip the Configuration and Metadata tabs.
11. Click **Launch Instance**. This step starts your new Ops Manager instance.

Step 6: Associate a Floating IP Address

1. In the left navigation of your OpenStack dashboard, click Project > Compute > Instances.
2. Wait until the **Power State** of the Ops Manager instance shows as *Running*.
3. Record the private **IP Address** of the Ops Manager instance.

Instances

<input type="checkbox"/>	INSTANCE NAME	IMAGE NAME	IP ADDRESS	SIZE	KEY PAIR	STATUS	AVAILABILITY ZONE	TASK	POWER STATE	TIME SINCE CREATED	ACTIONS
<input type="checkbox"/>	Ops Manager	Ops Manager	192.168.125.3	m1.large	pcf	Active	nova	None	Running	0 minutes	Create Snapshot

Displaying 1 item

You must provide this IP Address when you perform [Step 6: Complete the Create Networks Page](#) in Ops Manager.

4. Select the **Ops Manager** checkbox. Click the **Actions** dropdown and select **Associate Floating IP**. The **Manage Floating IP Associations** screen

Manage Floating IP Associations

IP Address *

IP Address *
No floating IP addresses allocated +

Select the IP address you wish to associate with the selected instance or port.

Port to be associated *
Ops Manager: 192.168.125.3

Cancel Associate

appears.

5. Under **IP Address**, click the plus button (+). The **Allocate Floating IP** screen appears.

6. Under **Pool**, select an IP Pool and click **Allocate IP**.

Allocate Floating IP

Pool *
admin_floating_net

Description:
Allocate a floating IP from a given floating IP pool.

Project Quotas

Floating IP (1)	49 Available
-----------------	--------------

Cancel Allocate IP

7. Under **Port to be associated**, select your Ops Manager instance.

Manage Floating IP Associations

IP Address *

IP Address *
10. [REDACTED] +

Select the IP address you wish to associate with the selected instance or port.

Port to be associated *
Ops Manager: 192.168.125.3

Cancel Associate

8. Click **Associate**.

Step 7: Add Blob Storage

1. In the left navigation of your OpenStack dashboard, click **Project > Object Store > Containers**.

2. Click **Create Container**. Create a container with the following properties:

- **Container Name:** Enter `pcf`.
- **Container Access:** Leave `public` unselected.

Create Container

Container Name * pcf	A container is a storage compartment for your data and provides a way for you to organize your data. You can think of a container as a folder in Windows® or a directory in UNIX®. The primary difference between a container and these other file system concepts is that containers cannot be nested. You can, however, create an unlimited number of containers within your account. Data must be stored in a container so you must have at least one container defined in your account prior to uploading data. Note: A Public Container will allow anyone with the Public URL to gain access to your objects in the container.
<input type="checkbox"/> Public	
<input type="button" value="Cancel"/> <input type="button" value="Create"/>	

3. Click **Create**.

Step 8: Download Credentials for S3 Blob Storage

- In the left navigation of your OpenStack dashboard, click **Project > Compute > Access & Security**. Select the **API Access** tab.

Access & Security

Security Groups	Key Pairs	Floating IPs	API Access						
API Endpoints <div style="display: flex; justify-content: space-around;"> Download OpenStack RC File Download EC2 Credentials View Credentials </div> <table border="1"> <thead> <tr> <th>Service</th> <th>Service Endpoint</th> </tr> </thead> <tbody> <tr> <td>Compute</td> <td>http://216.58.199.102</td> </tr> <tr> <td>Network</td> <td>http://216.58.199.102</td> </tr> </tbody> </table>				Service	Service Endpoint	Compute	http://216.58.199.102	Network	http://216.58.199.102
Service	Service Endpoint								
Compute	http://216.58.199.102								
Network	http://216.58.199.102								

2. Click **Download EC2 Credentials**.

3. Unzip the downloaded credentials.

4. If you select **S3 Compatible Blobstore** in your [BOSH Director Config](#), you need the contents of this file to complete the configuration.

Step 9: Create a DNS Entry

Note: For security, Ops Manager v1.7 and later require you to create a fully qualified domain name in order to access Ops Manager during the [initial configuration](#).

Create a DNS entry for the floating IP address that you assigned to Ops Manager in [Step 6: Associate a Floating IP Address](#).

You must use this fully qualified domain name when you log into Ops Manager for the first time.

Step 10: Configure BOSH Director for OpenStack

After completing this procedure, complete all of the steps in the [Configuring BOSH Director on OpenStack](#) and [Configuring PAS](#) topics.

Return to [Installing Pivotal Cloud Foundry on OpenStack](#).

Configuring BOSH Director on OpenStack

This topic describes how to configure BOSH Director after deploying [Pivotal Cloud Foundry](#) (PCF) on OpenStack. Use this topic when [Installing Pivotal Cloud Foundry on OpenStack](#).

- Before beginning this procedure, ensure that you have successfully completed all steps in [Provisioning the OpenStack Infrastructure](#).
- After you complete this procedure, follow the instructions in [Configuring PAS](#).

 **Note:** You can also perform the procedures in this topic using the Ops Manager API. For more information, see [Using the Ops Manager API](#).

Step 1: Access Ops Manager

1. In a web browser, navigate to the fully qualified domain you created in the [Create a DNS Entry](#) step of *Provisioning the OpenStack Infrastructure*.
2. When Ops Manager starts for the first time, you must choose one of the following:
 - **Internal Authentication:** If you use Internal Authentication, Ops Manager maintains your user database.
 - **SAML Identity Provider:** If you use a SAML Identity Provider (IdP), an external identity server maintains your user database.
 - **LDAP Server:** If you use a LDAP Server, an external identity server maintains your user database.

Internal Authentication

1. When redirected to the **Internal Authentication** page, do the following:
 - Enter a **Username**, **Password**, and **Password confirmation** to create an Admin user.
 - Enter a **Decryption passphrase** and the **Decryption passphrase confirmation**. This passphrase encrypts the Ops Manager datastore, and is not recoverable.
 - If you are using an **HTTP proxy** or **HTTPS proxy**, follow the instructions in the [Configuring Proxy Settings for the BOSH CPI](#) topic.
 - Read the **End User License Agreement**, and select the checkbox to accept the terms.
 - Click **Setup Authentication**.
2. Log in to Ops Manager with the Admin username and password you created in the previous step.

SAML Identity Provider

1. Log in to your IdP console and download the IdP metadata XML. Optionally, if your IdP supports metadata URL, you can copy the metadata URL instead of the XML.
2. Copy the IdP metadata XML or URL to the Ops Manager **SAML Identity Provider** login page.

 **Note:** The same IdP metadata URL or XML is applied for the BOSH Director. If you use a separate IdP for BOSH, copy the metadata XML or URL from that IdP and enter it into the BOSH IdP Metadata text box in the Ops Manager login page.

3. Enter values for the fields listed below. Failure to provide values in these fields results in a **500** error.
 - **SAML admin group:** Enter the name of the SAML group that contains all Ops Manager administrators. This field is case-sensitive.
 - **SAML groups attribute:** Enter the groups attribute tag name with which you configured the SAML server. This field is case-sensitive.
4. Enter your **Decryption passphrase**. Read the **End User License Agreement**, and select the checkbox to accept the terms.
5. Your Ops Manager login page appears. Enter your username and password. Click **Login**.
6. Download your SAML Service Provider metadata (SAML Relying Party metadata) by navigating to the following URLs:
 - **6a.** Ops Manager SAML service provider metadata: <https://OPS-MAN-FQDN:443/uaa/saml/metadata>
 - **6b.** BOSH Director SAML service provider metadata: <https://BOSH-IP-ADDRESS:8443/saml/metadata>

 **Note:** To retrieve your **BOSH-IP-ADDRESS**, navigate to the **Status** tab in the **BOSH Director** tile. Record the **BOSH Director** IP address.

7. Configure your IdP with your SAML Service Provider metadata. Import the Ops Manager SAML provider metadata from Step 6a above to your IdP. If

your IdP does not support importing, provide the values below.

- Single sign on URL: `https://OPS-MAN-FQDN:443/uaa/saml/SSO/alias/OPS-MAN-FQDN`
- Audience URI (SP Entity ID): `https://OP-MAN-FQDN:443/uaa`
- Name ID: Email Address
- SAML authentication requests are always signed

8. Import the BOSH Director SAML provider metadata from Step 6b to your IdP. If the IdP does not support an import, provide the values below.

- Single sign on URL: `https://BOSH-IP:8443/saml/SSO/alias/BOSH-IP`
- Audience URI (SP Entity ID): `https://BOSH-IP:8443`
- Name ID: Email Address
- SAML authentication requests are always signed

9. Return to the **BOSH Director** tile, and continue with the configuration steps below.

LDAP Server

- For **Server URL**, enter the URL that points to your LDAP server. With multiple LDAP servers, separate their URLs with spaces. Each URL must include one of the following protocols:
 - `ldap://`: This specifies that the LDAP server uses an unencrypted connection.
 - `ldaps://`: This specifies that the LDAP server uses SSL for an encrypted connection and requires that the LDAP server holds a trusted certificate or that you import a trusted certificate to the JVM truststore.
- For **LDAP Username** and **LDAP Password**, enter the LDAP Distinguished Name (DN) and the password for binding to the LDAP Server. Example DN: `cn=administrator,ou=Users,dc=example,dc=com`

Note: Pivotal recommends that you provide LDAP credentials that grant read-only permissions on the LDAP Search Base and the LDAP Group Search Base. In addition to this, if the bind user belongs to a different search base, you must use the full DN.

warning: Pivotal recommends against reusing LDAP service accounts across environments. LDAP service accounts should not be subject to manual lockouts, such as lockouts that result from users utilizing the same account. Also, LDAP service accounts should not be subject to automated deletions, since disruption to these service accounts could prevent user logins.

- For **User Search Base**, enter the location in the LDAP directory tree from which any LDAP User search begins. The typical LDAP Search Base matches your domain name.

For example, a domain named “cloud.example.com” typically uses the following LDAP User Search Base: `ou=Users,dc=example,dc=com`

- For **User Search Filter**, enter a string that defines LDAP User search criteria. These search criteria allow LDAP to perform more effective and efficient searches. For example, the standard LDAP search filter `cn=Smith` returns all objects with a common name equal to `Smith`.

In the LDAP search filter string that you use to configure your runtime, use `{0}` instead of the username. For example, use `cn={0}` to return all LDAP objects with the same common name as the username.

In addition to `cn`, other attributes commonly searched for and returned are `mail`, `uid` and, in the case of Active Directory, `sAMAccountName`.

Note: For instructions for testing and troubleshooting your LDAP search filters, see [Configuring LDAP Integration with Pivotal Cloud Foundry](#) in the Pivotal Support Knowledge Base.

- For **Group Search Base**, enter the location in the LDAP directory tree from which the LDAP Group search begins.

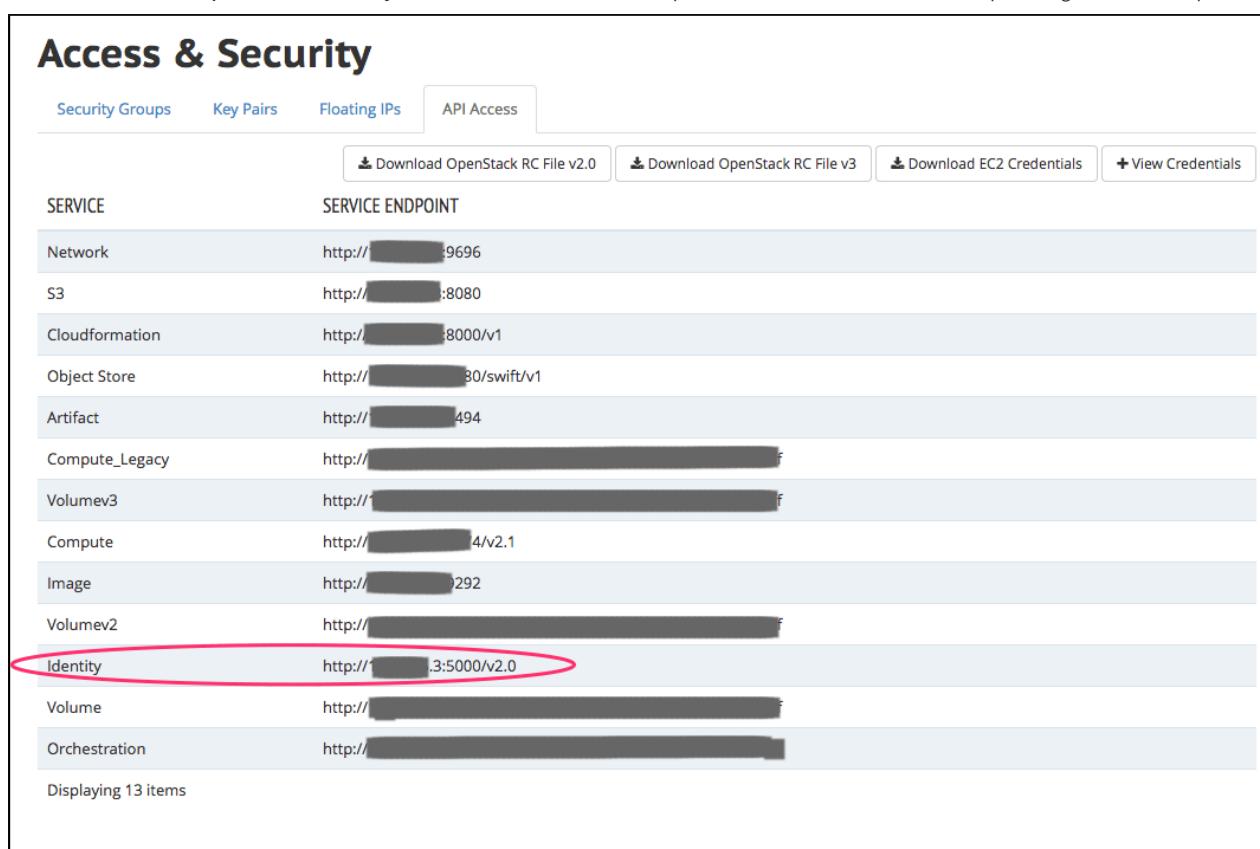
For example, a domain named “cloud.example.com” typically uses the following LDAP Group Search Base: `ou=Groups,dc=example,dc=com`

- For **Group Search Filter**, enter a string that defines LDAP Group search criteria. The standard value is `member={0}`.
- For **Email Attribute**, enter the attribute name in your LDAP directory that corresponds to the email address in each user record, for example `mail`.
- For **LDAP RBAC Admin Group Name**, enter the DN of the LDAP group you want to have admin permissions in Ops Manager.
- From the dropdown, select how the UAA handles LDAP server referrals out to other external user stores. The UAA can:
 - Automatically follow any referrals.
 - Ignore referrals and return partial result.
 - Throw exception for each referral and abort.
- For **Server SSL Cert**, paste in the root certificate from your CA certificate or your self-signed certificate.

- Enter a **Decryption passphrase** and the **Decryption passphrase confirmation**. This passphrase encrypts the Ops Manager datastore, and is not recoverable.
- If you are using an **HTTP proxy** or **HTTPS proxy**, follow the instructions in [Configuring Proxy Settings for the BOSH CPI](#).
- Read the **End User License Agreement**, and select the checkbox to accept the terms.
- Select **Provision an admin client in the BOSH UAA**. You can use this to enable BOSH automation with scripts and tooling. For more information, see [Provision Admin Client](#) in *Creating UAA Clients for BOSH Director*.
- Click **Setup Authentication**.
- Return to the **BOSH Director** tile, and continue with the configuration steps below.

Step 2: OpenStack Configs Page

- In the left navigation of your OpenStack dashboard, click **Project > Compute > Access & Security**. Select the **API Access** tab.
- Record the **Service Endpoint** for the **Identity** service. You use this Service Endpoint as the Authentication URL for Ops Manager in a later step.



The screenshot shows the 'Access & Security' interface with the 'API Access' tab selected. It lists various services and their corresponding service endpoints. The 'Identity' service endpoint is circled in red at the bottom of the list.

SERVICE	SERVICE ENDPOINT
Network	http://[REDACTED]:9696
S3	http://[REDACTED]:8080
Cloudformation	http://[REDACTED]:8000/v1
Object Store	http://[REDACTED]:80/swift/v1
Artifact	http://[REDACTED]:494
Compute_Legacy	http://[REDACTED]
Volumev3	http://[REDACTED]
Compute	http://[REDACTED]:4/v2.1
Image	http://[REDACTED]:292
Volumev2	http://[REDACTED]
Identity	http://[REDACTED]:5000/v2.0
Volume	http://[REDACTED]
Orchestration	http://[REDACTED]

Displaying 13 items

- In the PCF Ops Manager Installation Dashboard, click the **BOSH Director** tile.
- Select **OpenStack Configs**.
- Complete the **OpenStack Management Console Config** page with the following information:
 - Name:** Enter a unique name for the OpenStack config.
 - Authentication URL:** Enter the Service Endpoint for the Identity service that you recorded in a previous step.
 - Keystone Version:** Choose a Keystone version, either **v2** or **v3**.
 - If you choose **v3**, enter the OpenStack Keystone domain to authenticate against in the **Domain** field. For more information about

OpenStack Config

Name*

default

Authentication URL*

http://10.87.16.4:5000/v2.0

Keystone Version

v2

v3

Domain

Keystone domains in OpenStack, see [Domains](#) in the *OpenStack documentation*.

- **Username:** Enter your OpenStack Horizon username. The [PrimaryProject](#) for the user must be the project you are using to deploy PCF. For more information, see [Manage projects and users](#) in the OpenStack documentation.
- **Password:** Enter your OpenStack Horizon password.
- **Tenant:** Enter your OpenStack tenant name.
- **Region:** Enter [RegionOne](#), or another region if recommended by your OpenStack administrator.

Username*	<input type="text" value="admin"/>
Password*	<input type="password" value="*****"/> Change
Tenant*	<input type="text" value="loam"/>
Region*	<input type="text" value="RegionOne"/>

- **Select OpenStack Network Type:** Select either **Nova**, the legacy OpenStack networking model, or **Neutron**, the newer networking model.
- **Ignore Server Availability Zone:** Do not select the checkbox.
- **Security Group Name:** Enter [opsmanager](#). You created this Security Group in the [Configure Security](#) step of *Provisioning the OpenStack Infrastructure*.
- **Key Pair Name:** Enter the name of the key pair that you created in the [Configure Security](#) step of *Provisioning the OpenStack Infrastructure*.

Select Openstack Network Type*	<input type="text" value="Neutron"/>
<input type="checkbox"/> Ignore Server Availability Zone	
Security Group Name	<input type="text" value="opsmanager"/>
Key Pair Name*	<input type="text" value="pcf"/>

- **SSH Private Key:** In a text editor, open the key pair file that you downloaded in the [Configure Security](#) step of *Provisioning the OpenStack Infrastructure*. Copy and paste the contents of the key pair file into the field.
- **(Optional) API SSL Certificate:** If you configured API SSL termination in your OpenStack Dashboard, enter your [API SSL Certificate](#).
- **Disable DHCP:** Do not select the checkbox unless your configuration requires it.
- **Boot From Volume:** Enable to boot VMs from a Cinder volume.

The screenshot shows a configuration form for an OpenStack environment. It includes fields for 'SSH Private Key' (with a placeholder '*****' and a 'Change' link), 'API SSL Certificate' (with a large empty text area), and checkboxes for 'Disable DHCP' (unchecked) and 'Boot From Volume' (checked). A blue 'Save' button is at the bottom.

6. Click **Save**.

7. (Optional) Click **Add OpenStack Config** to configure additional data centers. Click **Save** for each additional OpenStack config to add it successfully. For more information, see [Managing Multiple Data Centers](#).

Step 3: (Optional) Advanced Config Page

Note: This is an advanced option. Most users leave this field blank.

1. In Ops Manager, select **Advanced Infrastructure Config**.

The screenshot shows the 'Advanced Infrastructure Configuration' page. It features a 'Connection Options' field (an empty text area) and a blue 'Save' button at the bottom.

2. If your OpenStack environment requires specific connection options, enter them in the **Connection Options** field in JSON format. For example:
`'connection_options' => { 'read_timeout' => 200 }`

 **Note:** Your connection options apply to all of your OpenStack configs.

3. Click **Save**.

Step 4: Director Config Page

1. In Ops Manager, select **Director Config**.

Director Config

NTP Servers (comma delimited)*

Bosh HM Forwarder IP Address

Enable VM Resurrector Plugin

Enable Post Deploy Scripts

Recreate All VMs
This will force BOSH to recreate all VMs on the next deploy. Persistent disk will be preserved.

Recreate All Persistent Disks
Check this box to recreate all Persistent Disks for the Director and all other Tiles.

Enable bosh deploy retries
This will attempt to re-deploy a failed deployment up to 5 times.

2. Enter one or more NTP servers in the **NTP Servers (comma delimited)** field. For example, `[us.pool.ntp.org]`.

 **Note:** The NTP server configuration only updates after VM recreation. Ensure that you select the **Recreate all VMs** checkbox if you modify the value of this field.

3. Leave the **Bosh HM Forwarder IP Address** field blank.

 **Note:** Starting in PAS v2.0, BOSH-reported component metrics are available in the Loggregator Firehose by default. If you continue to use the BOSH HM Forwarder to consume these component metrics, you may receive duplicate data. To prevent this, leave the **Bosh HM Forwarder IP Address** field blank.

4. Select the **Enable VM Resurrector Plugin** checkbox to enable the BOSH Resurrector functionality and increase your runtime availability.
 5. Select **Enable Post Deploy Scripts** to run a post-deploy script after deployment. This script allows the job to execute additional commands against a deployment.
-  **Note:** If you intend to install PKS, you must enable post-deploy scripts.
6. Select **Recreate all VMs** to force BOSH to recreate all VMs on the next deploy. This process does not destroy any persistent disk data.
 7. Select **Recreate All Persistent Disks** to force BOSH to migrate and recreate persistent disks for the BOSH Director and all tiles. This process does not destroy any persistent disk data.
 8. Select **Enable bosh deploy retries** to instruct Ops Manager to retry failed BOSH operations up to five times.
 9. Select **Skip Director Drain Lifecycle** to prevent drain scripts from running when the BOSH Director is recreated.
 10. Select **Store BOSH Job Credentials on tmpfs (beta)** to store credentials for BOSH jobs on temporary file storage (tmpfs) memory, rather than on disk. You must recreate all VMs for this setting to take effect.

11. Select **Keep Unreachable Director VMs** if you want to preserve BOSH Director VMs after a failed deployment for troubleshooting purposes.

12. Select **HM Pager Duty Plugin** to enable Health Monitor integration with PagerDuty.

HM Pager Duty Plugin

Service Key*

YOUR-PAGERDUTY-SERVICE-KEY

HTTP Proxy

YOUR-HTTP-PROXY

- **Service Key:** Enter your API service key from PagerDuty.
- **HTTP Proxy:** Enter an HTTP proxy for use with PagerDuty.

13. Select **HM Email Plugin** to enable Health Monitor integration with email.

HM Email Plugin

Host*

smtp.example.com

Port*

25

Domain*

cloudfoundry.example.com

From*

user2@example.com

Recipients*

user@example.com, user1@example.com

Username

user

Password

.....

Enable TLS

- **Host:** Enter your email hostname.
- **Port:** Enter your email port number.
- **Domain:** Enter your domain.
- **From:** Enter the address for the sender.
- **Recipients:** Enter comma-separated addresses of intended recipients.
- **Username:** Enter the username for your email server.
- **Password:** Enter the password for your email server.
- **Enable TLS:** Select this checkbox to enable Transport Layer Security to the email host.

14. For **CredHub Encryption Provider**, you can choose whether BOSH CredHub stores its encryption key internally on the BOSH Director and CredHub VM, or in an external hardware security module (HSM). The HSM option is more secure.

Before configuring an HSM encryption provider in the **Director Config** pane, you must follow the procedures and collect information described in [Preparing CredHub HSMs for Configuration](#).

Note: After you deploy Ops Manager with an HSM encryption provider, you cannot change BOSH CredHub to store encryption keys internally.

CredHub Encryption Provider

- Internal
- Luna HSM

Encryption Key Name*

Provider Partition*

Provider Partition Password*

Provider Client Certificate*

Provider Client Certificate Private Key*

HSM Host Address*

HSM Port Address*

Partition Serial Number*

HSM Certificate*

- **Internal:** Select this option for internal CredHub key storage. This option is selected by default and requires no additional configuration.
 - **Luna HSM:** Select this option to use a SafeNet Luna HSM as your permanent CredHub encryption provider, and fill in the following fields:
 1. **Encryption Key Name:** Any name to identify the key that the HSM uses to encrypt and decrypt the CredHub data. Changing this key name after you deploy Ops Manager can cause service downtime.
 2. **Provider Partition:** The partition that stores your encryption key. Changing this partition after you deploy Ops Manager could cause service downtime. For this value and the ones below, use values gathered in [Preparing CredHub HSMs for Configuration](#).
 3. **Provider Partition Password**
 4. **Provider Client Certificate:** The certificate that validates the identity of the HSM when CredHub connects as a client.
 5. **Provider Client Certificate Private Key**
 6. **HSM Host Address**
 7. **HSM Port Address:** If you do not know your port address, enter `1792`.
 8. **Partition Serial Number**
 9. **HSM Certificate:** The certificate that the HSM presents to CredHub to establish a two-way mTLS connection.
15. Select a **Blobstore Location** to either configure the blobstore as an internal server or an external endpoint. Because the internal server is unscalable

and less secure, Pivotal recommends that you configure an external blobstore.

 **Note:** After you deploy Ops Manager, you cannot change the blobstore location.

- **Internal:** Select this option to use an internal blobstore. Ops Manager creates a new VM for blob storage. No additional configuration is required.
 - **Enable TLS:** Select this checkbox to enable TLS to the blobstore.

 **Note:** If you are using PASW 2016, make sure you have downloaded [Windows stemcell v1709.10 or higher](#) before enabling TLS.

- **S3 Compatible Blobstore:** Select this option to use an external S3-compatible endpoint.

Blobstore Location

Internal

Enable TLS

S3 Compatible Blobstore

S3 Endpoint*

Bucket Name*

Access Key*

Secret Key*

S3 Signature Version

V2 Signature

V4 Signature

Region*

S3 Backup Strategy

No backups

Use a versioned bucket

Copy into an additional bucket

Backup Bucket Region*

Backup Bucket Name*

Follow the procedures in [Sign up for Amazon S3](#) and [Creating a Bucket](#) in the AWS documentation. When you have created an S3 bucket, complete the following steps:

1. **S3 Endpoint:** Navigate to the [Regions and Endpoints](#) topic in the AWS documentation.
 - a. Locate the endpoint for your region in the **Amazon Simple Storage Service (S3)** table and construct a URL using your region's endpoint. For example, if you are using the `us-west-2` region, the URL you create would be `https://s3-us-west-2.amazonaws.com`. Enter this URL into the **S3 Endpoint** field.
 - b. On a command line, run `ssh ubuntu@OPS-MANAGER-FQDN` to SSH into the Ops Manager VM. Replace `OPS-MANAGER-FQDN` with the fully qualified domain name of Ops Manager.

- c. Copy the custom public CA certificate you used to sign the S3 endpoint into `/etc/ssl/certs` on the Ops Manager VM.
- d. On the Ops Manager VM, run `sudo update-ca-certificates -f -v` to import the custom CA certificate into the Ops Manager VM truststore.

Note: You must also add this custom CA certificate into the **Trusted Certificates** field in the **Security** page. See [Security Page](#) for instructions.

2. **Bucket Name:** Enter the name of the S3 bucket.
3. **Access Key** and **Secret Key:** Enter the keys you generated when creating your S3 bucket.
4. Select **V2 Signature** or **V4 Signature**. If you select **V4 Signature**, enter your **Region**.

Note: AWS recommends using Signature Version 4. For more information about AWS S3 Signatures, see [Authenticating Requests](#) in the AWS documentation.

5. **S3 Backup Strategy:** Select one of the following to configure whether and how to back up the BOSH Director's S3 blobstore:

- **No backups**
- **Use a versioned bucket:** This option uses blobstore bucket versioning to save backups. It requires an S3 blobstore that supports versioning.
- **Copy into an additional bucket:** This option saves blobstore backups to a separate bucket, which is useful with blobstores that do not support versioning. The option requires you to fill in **Backup Bucket Region** and **Backup Bucket Name**.

- o **GCS Blobstore:** Select this option to use an external Google Cloud Storage (GCS) endpoint.

The screenshot shows a configuration form for a GCS Blobstore. At the top, there is a radio button labeled "GCS Blobstore" which is selected. Below it are three input fields: "Bucket Name*" with a placeholder box, "Storage Class*" with a dropdown menu set to "Regional", and "Service Account Key*" with a large text area for pasting JSON content.

Follow the procedures in [Creating Storage Buckets](#) in the GCS documentation to create a GCS bucket. To create a GCS bucket, you must have a GCS account. When you have created a GCS bucket, complete the following steps:

1. **Bucket Name:** Enter the name of your GCS bucket.
 2. **Storage Class:** Select the storage class for your GCS bucket. See [Storage Classes](#) in the GCP documentation for more information.
 3. **Service Account Key:** Follow the steps in the [Set up IAM Service Accounts](#) section of [Preparing to Deploy Ops Manager on GCP Manually](#) to download a JSON file with a private key. Enter the contents of the JSON file into the field.
16. Select a **Database Location**. By default, Ops Manager deploys and manages an **Internal** database for you. If you choose to use an **External MySQL Database**, complete the associated fields with information obtained from your external MySQL Database provider: **Host**, **Port**, **Username**,

Database Location

Internal
 External MySQL Database

Host*

Port*

Username*

Password*

Database*

Password, and Database.

In addition, if you selected the **Enable TLS for Director Database** checkbox, you can complete the following optional fields:

- **Enable TLS:** Select this checkbox enables TLS communication between the BOSH Director and the database.
- **TLS CA:** Enter the Certificate Authority for the TLS Certificate.
- **TLS Certificate:** Enter the client certificate for mutual TLS connections to the database.
- **TLS Private Key:** Enter the client private key for mutual TLS connections to the database.
- **Advanced DB Connection Options:** If you would like to provide additional options for the database, use this field to provide a JSON-formatted options string.

17. (Optional) Modify the **Director Workers** value, which sets the number of workers available to execute Director tasks. This field defaults to .
18. (Optional) **Max Threads** sets the maximum number of threads that the BOSH Director can run simultaneously. Pivotal recommends that you leave the field blank to use the default value, unless doing so results in rate limiting or errors on your IaaS.
19. (Optional) To add a custom URL for your BOSH Director, enter a valid hostname in **Director Hostname**. You can also use this field to configure a load balancer in front of your BOSH Director. For more information, see [How to Set Up a Load Balancer in Front of Operations Manager Director](#) in the Pivotal Support Knowledge Base.

Director Workers

Max Threads

Director Hostname

20. (Optional) To set a custom banner that users see when logging in to the Director using SSH, enter text in the **Custom SSH Banner** field.
21. (Optional) Enter your comma-separated custom **Identification Tags**. For example, `iaas:foundation1, hello:world`. You can use the tags to identify your foundation when viewing VMs or disks from your IaaS.
22. Click **Save**.

Step 5: Create Availability Zones Page

1. In Ops Manager, select **Create Availability Zones**.

The screenshot shows the 'Availability Zones' configuration screen. At the top right is an 'Add' button. Below it, under the 'nova' section, there is a 'Name*' field containing 'nova' with a copy icon, an 'IaaS Configuration*' dropdown set to 'second_iaas', and an 'Openstack Availability Zone*' field also set to 'nova'. A tooltip for the Openstack Availability Zone field explains that it is the name of the Availability Zone on Openstack. Below this is a row for '2nd az' with a delete icon. At the bottom is a large blue 'Save' button.

2. Enter the name of the availability zone that you selected in the [Launch Ops Manager VM](#) step of *Provisioning the OpenStack Infrastructure*.
3. (Optional) Select an OpenStack config name from the **IaaS Configuration** dropdown. The default is set to your first OpenStack config.
4. Enter the **OpenStack Availability Zone** of your OpenStack environment. Many OpenStack environments default to `nova`.
5. Click **Add** for each additional OpenStack config you created in [Step 2: OpenStack Configs Page](#). Give each AZ a unique **Name** and an **IaaS Configuration** with a different OpenStack config.
6. Click **Save**.

Step 6: Create Networks Page

1. In the left navigation of your OpenStack dashboard, click **Project > Network > Networks**.
2. Click the name of the network that contains the private subnet where you deployed the Ops Manager VM. The OpenStack Network Detail page displays your network settings.

[Networks / loam_net](#) Edit Network ▾

Network Overview

Name	loam_net
ID	f92facc6-b7c7-4e00-8a4d-1234567890
Project ID	8a4d1234-5678-4e00-8a4d-1234567890
Status	Active
Admin State	UP
Shared	No
External Network	No
MTU	1450
Provider Network	Network Type: vxlan Physical Network: - Segmentation ID: 93

Subnets

+ Create Subnet >Delete Subnets

<input type="checkbox"/> NAME	NETWORK ADDRESS	IP VERSION	GATEWAY IP	ACTIONS
<input type="checkbox"/> (9c622786-2276)	192.168.125.0/24	IPv4	192.168.125.1	Edit Subnet ▾

Displaying 1 item

Ports

NAME	FIXED IPS	ATTACHED DEVICE	STATUS	ADMIN STATE	ACTIONS
(592facc6-b7c7)	192.168.125.2	network:dhcp	Active	UP	Edit Port
(be836ab3-e9ed)	192.168.125.3	compute:nova	Active	UP	Edit Port
(75b1acc9-48ea)	192.168.125.1	network:router_interface	Active	UP	Edit Port
(1e7026a4-83ce)	192.168.125.4	compute:nova	Active	UP	Edit Port

Displaying 4 items

3. In Ops Manager, select **Create Networks**.

Create Networks

Warning: Pivotal recommends keeping the IP settings throughout the life of your installation. Ops Manager may prevent you from changing them in the future. Contact Pivotal support for help completing such a change.

Verification Settings

Enable ICMP checks

Networks

One or many IP ranges upon which your products will be deployed

loam

Name*	<input type="text" value="loam"/> A unique name for this network
-------	---

Subnets

Network ID*	<input type="text" value="f1..."/>
CIDR*	<input type="text" value="192.168.125.0/24"/>
Reserved IP Ranges	<input type="text" value="192.168.125.0-192.168.125.10"/>
DNS*	<input type="text" value="8.8.8.8"/>
Gateway*	<input type="text" value="192.168.125.1"/>
Availability Zones*	<input checked="" type="checkbox"/> nova

4. Select **Enable ICMP checks** to enable ICMP on your networks. Ops Manager uses ICMP checks to confirm that components within your network are reachable. Review the [Configure Security](#) step of *Deploying BOSH and Ops Manager to OpenStack* to ensure you have configured ICMP in your Security Group.

5. Use the following steps to create one or more Ops Manager networks using information from your OpenStack network:
 - Click **Add Network**.
 - Enter a unique **Name** for the network.
 - Click **Add Subnet** to create one or more subnets for the network.
 - For **Network ID**, use the ID from the OpenStack page.
 - For **CIDR**, use the **Network Address** from the OpenStack page.
 - For **Reserved IP Ranges**, use the first 10 IP addresses of the **Network Address** range, and the private IP address of the Ops Manager instance that you recorded in the [Associate a Floating IP Address](#) step of *Provisioning the OpenStack Infrastructure*.
 - For **DNS**, enter one or more Domain Name Servers.
 - For **Gateway**, use the **Gateway IP** from the OpenStack page.
 - For **Availability Zones**, select which Availability Zones to use with the network.

6. Click **Save**.

 **Note:** After you deploy Ops Manager, you add subnets with overlapping Availability Zones to expand your network. For more information about configuring additional subnets, see [Expanding Your Network with Additional Subnets](#).

Step 7: Assign AZs and Networks Page

1. Select **Assign Availability Zones**.

Assign AZs and Networks

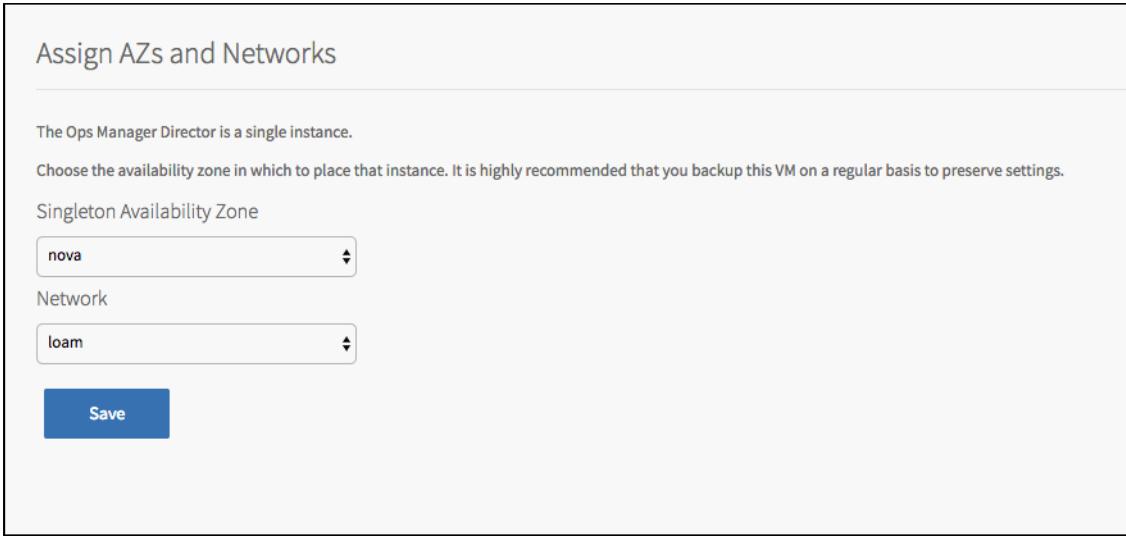
The Ops Manager Director is a single instance.

Choose the availability zone in which to place that instance. It is highly recommended that you backup this VM on a regular basis to preserve settings.

Singleton Availability Zone

Network

Save



2. From the **Singleton Availability Zone** dropdown, select the availability zone that you created in a previous step. The BOSH Director installs in this Availability Zone.
3. Use the dropdown to select the **Network** that you created in a previous step. BOSH Director installs in this network.
4. Click **Save**.

Step 8: Security Page

1. Select **Security**.

Security

Trusted Certificates

These certificates enable BOSH-deployed components to trust a custom root certificate.

Include OpsManager Root CA in Trusted Certs

Checking this box will place the OpsManager Root CA into the trusted certificate field. The Bosh Director will then place this CA into the trust store of every VM that it deploys.

Generate VM passwords or use single password for all VMs

Generate passwords

Use default BOSH password

Save

2. In **Trusted Certificates**, enter your custom certificate authority (CA) certificates to insert into your organization's certificate trust chain. This feature enables all BOSH-deployed components in your deployment to trust custom root certificates.

To enter multiple certificates, paste your certificates one after the other. For example, format your certificates like the following:

```
-----BEGIN CERTIFICATE-----
ABCDEFGH12345678ABCDEFGH12345678ABCDEFGH12345678AB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
BCDEFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
CDEFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABBB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
```

Note: If you want to use Docker registries to run PAS app instances in Docker containers, enter the certificate for your private Docker registry in this field. See [Using Docker Registries](#) for more information about running app instances in PAS using Docker registries.

3. (Optional) Select the **Include OpsManager Root CA in Trusted Certs** checkbox to include the Ops Manager root CA in the Trusted Certificates field. BOSH Director includes this CA in the trust store of every VM that it deploys.
4. Choose **Generate passwords** or **Use default BOSH password**. Pivotal recommends that you use the **Generate passwords** option for greater security.
5. Click **Save**. To view your saved Director password, click the **Credentials** tab.

Step 9: BOSH DNS Config Page

BOSH DNS Config

Excluded Recursors

Recurser Timeout

Handlers

[]

Enter a JSON-formatted list of optional custom domain handlers

Save

1. Select **BOSH DNS Config**.
2. (Optional) In **Excluded Recursors**, enter a list of prohibited recursor addresses.
3. (Optional) In **Recurser Timeout**, enter a time limit for contacting the connected recursors. This includes dialing, writing, and reading from the recursor. If any of these actions exceeds the time limit you set, the action fails.

Note: This time limit must include one of the Go parse duration time units. For example, entering `5s` sets the timeout limit to five seconds. For more information about supported time units, see [func ParseDuration](#) in the Go Programming Language documentation.

4. (Optional) In **Handlers**, enter a list of custom domain handlers in JSON format.
5. Click **Save**.

Step 10: Syslog Page

Settings

Syslog

Syslog settings make take up to 1 min to take effect.

Do you want to configure Syslog for OpsManager?

No
 Yes

Address*
 The address or host for the syslog server

Port*

Transport Protocol*

Enable TLS

Permitted Peer*

SSL Certificate*

Queue Size

Forward Debug Logs

Custom rsyslog Configuration

SAVE SYSLOG SETTINGS

1. Select **Syslog**.
2. (Optional) Select **Yes** to send BOSH Director system logs to a remote server.
3. Enter the IP address or DNS name for the remote server in **Address**.
4. Enter the port number that the remote server listens on in **Port**.
5. Select **TCP** or **UDP** from the **Transport Protocol** dropdown. This selection determines which transport protocol is used to send the logs to the remote server.
6. (Optional) Select the **Enable TLS** checkbox to send encrypted logs to remote server with TLS. After you select the checkbox, do the following:
 - a. Enter either the name or SHA1 fingerprint of the remote peer in **Permitted Peer**.
 - b. Enter the SSL certificate for the remote server in **SSL Certificate**.



Note: Pivotal strongly recommends that you enable TLS encryption when you are forwarding logs. Logs can contain sensitive

information, such as cloud provider credentials.

7. (Optional) Enter an integer in **Queue Size**. This value specifies the number of log messages held in the buffer. The default value is 100,000.
8. (Optional) Select the checkbox to **Forward Debug Logs** to an external source. This option is deselected by default. If you select it, you may generate a large amount of log data.
9. (Optional) Enter configuration details for rsyslog in the **Custom rsyslog Configuration** field. This field requires the rainerscript syntax.
10. Click **Save Syslog Settings**.

Step 11: Resource Config Page

1. Select **Resource Config**.

JOB	INSTANCES	PERSISTENT DISK TYPE	VM TYPE
Ops Manager Director	Automatic: 1	Automatic: 50 GB	Automatic: medium.disk (cpu: 2, ram: 4 GB)
Master Compilation Job	Automatic: 4	None	Automatic: large.cpu (cpu: 4, ram: 4 GB, disk: 100 GB)

2. Adjust any values as necessary for your deployment, such as increasing the persistent disk size. Select **Automatic** from the dropdown to provision the amount of persistent disk predefined by the job. If the persistent disk field reads **None**, the job does not require persistent disk space.

Note: Ops Manager requires a Director VM with at least 8 GB memory.

Note: If you set a field to **Automatic** and the recommended resource allocation changes in a future version, Ops Manager automatically uses the updated recommended allocation.

3. Click **Save**.

Step 12: (Optional) Add Custom VM Extensions

Use the Ops Manager API to add custom properties to your VMs such as associated security groups and load balancers. For more information, see [Managing Custom VM Extensions](#).

Step 13: Complete BOSH Director Installation

1. Click the **Installation Dashboard** link to return to the Installation Dashboard.
2. Click **Review Pending Changes**, then **Apply Changes**. If the following ICMP error message appears, click **Ignore errors and start the install**.

PCF Ops Manager

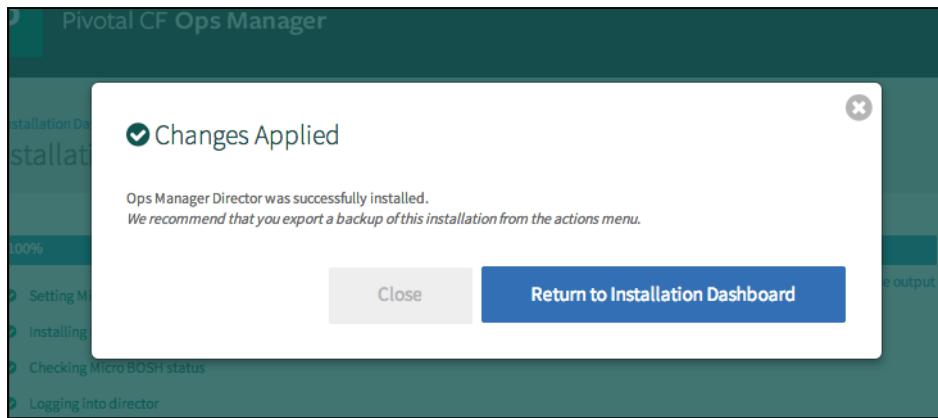
pivotal-cf

⚠ Please review the errors below

- Cannot reach gateway with IP 10.0.16.1 (ignorable if ICMP is disabled)
- Cannot reach DNS with IP 10.0.0.2 (ignorable if ICMP is disabled)

Ignore errors and start the install **Stop and fix errors**

3. BOSH Director installs. The image shows the **Changes Applied** message that Ops Manager displays when the installation process successfully completes.



4. After you complete this procedure, follow the instructions in [Configuring PAS](#).

Return to [Installing Pivotal Cloud Foundry on OpenStack](#)

Installing Pivotal Cloud Foundry on vSphere

This guide describes how to install [Pivotal Cloud Foundry](#) (PCF) on vSphere.

If you experience a problem while following the steps below, refer to the [Known Issues](#) topics or to [Diagnosing Problems in PCF](#).

Requirements

This section describes the resource requirements for installing PCF.

 **Note:** If you are using the Cisco Nexus 1000v Switch, refer to the [Using the Cisco Nexus 1000v Switch with Ops Manager](#) topic for more information.

 **Note:** When installing Ops Manager on a vSphere environment with multiple ESXi hosts, you must use network-attached or shared storage devices. Local storage devices do not support sharing across multiple ESXi hosts.

Resource Requirements

You can install PCF with one of two runtimes. For more information about the requirements for each runtime, see the following:

- **Pivotal Application Service (PAS):** [PAS on vSphere Requirements](#). For information requirements when deploying with NSX-T, see the *Requirements* section of [Deploying PAS with NSX-T Networking](#).
- **Pivotal Container Service (PKS):**
 - **Without NSX-T:** [vSphere Prerequisites and Resource Requirements](#)
 - **With NSX-T:** See the *Preparing to Install PKS on vSphere with NSX-T* section of [Installing PKS on vSphere with NSX-T Data Center](#).

vSphere Service Account Requirements

Ops Manager requires read/write permissions to the data center level of the [vSphere Inventory Hierarchy](#) to successfully install. Pivotal recommends defining a custom role for the service account that has all privileges for all objects in the data center, including propagating privileges to children.

Since Ops Manager passes all required credentials through to BOSH, you only need one service account with the required vSphere privileges to complete the installation. Setting up separate service accounts for Ops Manager and BOSH is not necessary or recommended.

For more information about service account requirements, see the following topics:

- **PAS:** [vSphere Service Account Requirements](#)
- **PKS:** Step 3: Create the BOSH/Ops Manager Service Account of [Preparing vSphere Before Deploying PKS](#)

 **Note:** You can also apply the default [VMware Administrator System Role](#) to the service account to achieve the appropriate permission level.

vSphere Security Documents

- [vSphere Security guide \(PDF\)](#)

This guide contains best practices for securing and managing a vSphere installation.

Procedure

This section describes the steps for installing PCF.

Step 1: Install Ops Manager

Complete the following procedures to install Ops Manager on vSphere:

1. [Deploying Ops Manager to vSphere](#)
2. [Configuring BOSH Director on vSphere](#)

Step 2: Install a Runtime

Follow the procedure that corresponds to your desired runtime and networking configuration:

- PAS:
 - Without NSX-T: [Configuring PAS](#)
 - With NSX-T: [Deploying PAS with NSX-T Networking](#)
- PKS:
 - Without NSX-T: [Installing PKS](#)
 - With NSX-T: [Deploying NSX-T for PKS](#)

 Note: You must install the NSX-T tile before you install or Upgrade the PAS tile.

Additional Configuration

See the following topics for additional configuration options:

- [Provisioning a Virtual Disk in vSphere](#)
- [Using BOSH Resurrector and vSphere HA](#)
- [Availability Zones in vSphere](#)
- [Updating NSX Security Group and Load Balancer Information](#)

vSphere Service Account Requirements

This topic describes the minimum privileges required by the vSphere BOSH CPI. A vSphere admin must grant the following privileges to the vSphere service account that Pivotal Cloud Foundry (PCF) uses to manage vSphere resources.

The PCF account needs privileges at both the vCenter server level and the Datacenter level. See [Hierarchical Inheritance of Permissions](#) in the VMware documentation for how permission levels and inheritance work in vSphere.

vCenter-Level Privileges

Ops Manager assigns custom attributes to the virtual machines (VMs) it deploys to identify BOSH releases and job index information about each VM. vCenter APIs require vCenter server level access to manage these custom attributes.

The following table summarizes the privileges that a PCF account requires at the vCenter Server instance level. Some of these privileges are inherited, and others must be granted by a vCenter admin:

Object	Privilege (UI)	Privilege (API)
Role	Read-only	System.Anonymous
		System.Read
		System.View
Global	Manage custom attributes	Global.ManageCustomFields
	Register Extensions	Extension.Register
Profile-Driven Storage	Profile-driven Storage	StorageProfile.Update
		StorageProfile.View

Datacenter-Level Privileges

The following privileges must be set at the data center level:

Object	Privilege (UI)	Privilege (API)
Datastore	Low level file operations	Datastore.FileManagement
Network	Assign network	Network.Assign

Folder and Datastore-Level Privileges

You must grant the following privileges on any entities in a datacenter where you will deploy PCF:

Datastore Object

Privilege (UI)	Privilege (API)
Allocate space	Datastore.AllocateSpace
Browse datastore	Datastore.Browse
Remove file	Datastore.DeleteFile
Update virtual machine files	Datastore.UpdateVirtualMachineFiles

Folder Object

Ops Manager creates a folder for VMs, stemcells, and persistent disks during installation. The folder contents change frequently as Ops Manager applies changes.

Privilege (UI)	Privilege (API)

Delete folder	Folder.Delete
Create folder	Folder.Create
Move folder	Folder.Move
Rename folder	Folder.Rename

Inventory Service Object

Privilege (UI)	Privilege (API)
vSphere Tagging > Create vSphere Tag	InventoryService.Tagging.CreateTag
vSphere Tagging > Delete vSphere Tag	InventoryService.Tagging.EditTag
vSphere Tagging > Edit vSphere Tag	InventoryService.Tagging.DeleteTag

Resource Object

When using `vAppImport` to clone a VM, BOSH requires the resource migration privileges to create a new, powered-off VM based on a given stemcell. BOSH migrates the VM to the destination datastore, where Ops Manager deploys the VM and powers it on.

Privilege (UI)	Privilege (API)
Assign virtual machine to resource pool	Resource.AssignVMToPool
Migrate powered off virtual machine	Resource.ColdMigrate
Migrate powered on virtual machine	Resource.HotMigrate

Virtual Machine Object

Configuration

Privilege (UI)	Privilege (API)
Add existing disk	VirtualMachine.Config.AddExistingDisk
Add new disk	VirtualMachine.Config.AddNewDisk
Add or remove device	VirtualMachine.Config.AddRemoveDevice
Advanced	VirtualMachine.Config.AdvancedConfig
Change CPU count	VirtualMachine.Config.CPUCount
Change resource	VirtualMachine.Config.Resource
Configure managedBy	VirtualMachine.Config.ManagedBy
Disk change tracking	VirtualMachine.Config.ChangeTracking
Disk lease	VirtualMachine.Config.DiskLease
Display connection settings	VirtualMachine.Config.MksControl
Extend virtual disk	VirtualMachine.Config.DiskExtend
Memory	VirtualMachine.Config.Memory
Modify device settings	VirtualMachine.Config.EditDevice
Raw device	VirtualMachine.Config.RawDevice
Reload from path	VirtualMachine.Config.ReloadFromPath
Remove disk	VirtualMachine.Config.RemoveDisk
Rename	VirtualMachine.Config.Rename
Reset guest information	VirtualMachine.Config.ResetGuestInfo
Set annotation	VirtualMachine.Config.Annotation
Settings	VirtualMachine.Config.Settings

Swapfile placement	VirtualMachine.Config.SwapPlacement
Unlock virtual machine	VirtualMachine.Config.Unlock
Upgrade virtual machine hardware	VirtualMachine.Config.UpgradeVirtualHardware

Guest Operations

Privilege (UI)	Privilege (API)
Guest Operation Program Execution	VirtualMachine.GuestOperations.Execute
Guest Operation Modifications	VirtualMachine.GuestOperations.Modify
Guest Operation Queries	VirtualMachine.GuestOperations.Query

Interaction

Privilege (UI)	Privilege (API)
Answer question	VirtualMachine.Interact.AnswerQuestion
Configure CD media	VirtualMachine.Interact.SetCDMedia
Device connection	VirtualMachine.Interact.DeviceConnection
Guest operating system management by VIX API	VirtualMachine.Interact.GuestControl
Power off	VirtualMachine.Interact.PowerOff
Power on	VirtualMachine.Interact.PowerOn
Reset	VirtualMachine.Interact.Reset
Suspend	VirtualMachine.Interact.Suspend
VMware Tools install	VirtualMachine.Interact.ToolsInstall

Inventory

Privilege (UI)	Privilege (API)
Create from existing	VirtualMachine.Inventory.CreateFromExisting
Create new	VirtualMachine.Inventory.Create
Move	VirtualMachine.Inventory.Move
Remove	VirtualMachine.Inventory.Delete

Provisioning

When cloning a stemcell, BOSH sets custom specifications, such as hostnames and network configurations, based on the stemcell operating system.

The VM download privilege allows BOSH to modify files within a VM, including links between VMs and persistent disks. When vMotion migrates disks in vSphere, BOSH uses these links to maintain the connections between VMs and their persistent disks.

Privilege (UI)	Privilege (API)
Allow disk access	VirtualMachine.Provisioning.DiskRandomAccess
Allow read-only disk access	VirtualMachine.Provisioning.DiskRandomRead
Allow virtual machine download	VirtualMachine.Provisioning.GetVmFiles
Allow virtual machine files upload	VirtualMachine.Provisioning.PutVmFiles
Clone template	VirtualMachine.Provisioning.CloneTemplate
Clone virtual machine	VirtualMachine.Provisioning.Clone
Customize	VirtualMachine.Provisioning.Customize
Deploy template	VirtualMachine.Provisioning.DeployTemplate
Mark as template	VirtualMachine.Provisioning.MarkAsTemplate

Mark as virtual machine	VirtualMachine.Provisioning.MarkAsVM
Modify customization specification	VirtualMachine.Provisioning.ModifyCustSpecs
Promote disks	VirtualMachine.Provisioning.PromoteDisks
Read customization specifications	VirtualMachine.Provisioning.ReadCustSpecs

Snapshot Management

Before Ops Manager deploys a new VM, it uses a snapshot to clone the stemcell image to the destination.

Privilege (UI)	Privilege (API)
Create snapshot	VirtualMachine.State.CreateSnapshot
Remove snapshot	VirtualMachine.State.RemoveSnapshot
Rename snapshot	VirtualMachine.State.RenameSnapshot
Revert snapshot	VirtualMachine.State.RevertToSnapshot

vApp Object

These privileges must be set at the resource pool level. `VApp.ApplicationConfig` is required when attaching or detaching persistent disks.

Privilege (UI)	Privilege (API)
Import	VApp.Import
vApp application configuration	VApp.ApplicationConfig

PAS on vSphere Requirements

This topic describes requirements for installing Pivotal Application Service (PAS) on vSphere.

General Requirements

The following are general requirements for deploying and managing a PCF deployment with Ops Manager and Pivotal Application Service (PAS):

- A wildcard DNS record that points to your router or load balancer. Alternatively, you can use a service such as xip.io. For example, `203.0.113.0.xip.io`.
 - PAS gives each application its own hostname in your app domain.
 - With a wildcard DNS record, every hostname in your domain resolves to the IP address of your router or load balancer, and you do not need to configure an A record for each app hostname. For example, if you create a DNS record `*.example.com` pointing to your load balancer or router, every application deployed to the `example.com` domain resolves to the IP address of your router.
- At least one wildcard TLS certificate that matches the DNS record you set up above, `*.example.com`.
- Sufficient IP allocation:
 - One static IP address for either HAProxy or one of your goRouters
 - One static IP address for each job in the Ops Manager tile. See the Resource Config pane for each tile for a full list.
 - One static IP address for each job listed below:
 - Consul
 - NATS
 - File Storage
 - MySQL Proxy
 - MySQL Server
 - Backup Restore Node
 - HAProxy
 - Router
 - MySQL Monitor
 - Diego Brain
 - TCP Router
 - One IP for each VM instance created by the service.
 - An additional IP address for each compilation worker. So the formula for total IPs needed is
`IPs needed = static IPs + VM instances + compilation workers`.

 **Note:** Pivotal recommends that you allocate at least 36 dynamic IP addresses when deploying Ops Manager and PAS. BOSH requires additional dynamic IP addresses during installation to compile and deploy VMs, install PAS, and connect to services.

- One or more NTP servers if not already provided by your IaaS.
- **(Recommended)** A network without DHCP available for deploying the PAS VMs.

 **Note:** If you have DHCP, refer to the [Troubleshooting Guide](#) to avoid issues with your installation.

- **(Optional)** External storage. When you deploy PCF, you can select internal file storage or external file storage, either network-accessible or IaaS-provided, as an option in the PAS tile. Pivotal recommends using external storage whenever possible. See [Configure File Storage](#) for a discussion of how file storage location affects platform performance and stability during upgrades.
- **(Optional)** External databases. When you deploy PCF, you can select internal or external databases for the BOSH Director and for PAS. Pivotal recommends using external databases in production deployments. An external database must be configured to use the UTC timezone.
- **(Optional)** External user stores. When you deploy PCF, you can select a SAML user store for Ops Manager or a SAML or LDAP user store for PAS, to integrate existing user accounts.
- The most recent version of the [Cloud Foundry Command Line Interface](#) (cf CLI).

vSphere Requirements

The following are requirements for installing PAS that are specific to vSphere.

Minimum Resource Requirements for PCF Deployment with Pivotal Application Service

The following are the minimum resource requirements for maintaining a [Pivotal Cloud Foundry](#) (PCF) deployment with Ops Manager and Pivotal Application Service (PAS) on vSphere:

- vSphere v6.7, v6.5, v6.0, or v5.5
- Disk space: 2 TB recommended
- Memory: 120 GB
- Two public IP addresses: One for PAS and one for Ops Manager
- vCPU cores: 82
- Overall CPU: 28 GHz
- vSphere editions: standard and above
- Ops Manager must have HTTPS access to vCenter and ESX hosts on TCP port 443.
- A configured vSphere cluster:
 - If you enable vSphere DRS (Distributed Resource Scheduler) for the cluster, you must set the Automation level to **Partially automated** or **Fully automated**. If you set the Automation level to **Manual**, the BOSH automated installation will fail with a `power_on_vm` error when BOSH attempts to create virtual machines (VMs).
 - Disable hardware virtualization if your vSphere hosts do not support VT-x/EPT. If you are unsure whether the VM hosts support VT-x/EPT, you should disable this setting. If you leave this setting enabled and the VM hosts do not support VT-x/EPT, then each VM requires manual intervention in vCenter to continue powering on without the Intel virtualized VT-x/EPT. Refer to the vCenter help topic at [Configuring Virtual Machines > Setting Virtual Processors and Memory > Set Advanced Processor Options](#) for more information.
- If you configure an external load balancer, an HTTP keep-alive connection timeout greater than five seconds

 **Note:** If you are deploying PCF behind a firewall, see the [Preparing Your Firewall for Deploying Pivotal Cloud Foundry](#) topic.

Minimum Resource Requirements for PCF Deployment with Small Footprint PAS

The following are the minimum resource requirements for maintaining a PCF deployment with Ops Manager and Small Footprint PAS on vSphere:

- 1 vSphere cluster/1 AZ
- 2 resource pools: 1 for NSX-T components and 1 for PAS or PKS
- 3 hosts minimum for vSphere HA (4 hosts for vSphere VSAN)
- Shared storage/VSAN
- 1 NSX Manager
- 3 NSX controllers
- 4 large edge VMs in a cluster
- PCF: Ops Manager, the BOSH Director, and Small Footprint PAS

Instance Number and Scaling Requirements

By default, PAS deploys the number of VM instances required to run a highly available configuration of PCF. If you are deploying a test or sandbox PCF that does not require HA, then you can scale down the number of instances in your deployment.

For information about the number of instances required to run a minimal, non-HA PCF deployment, see [Scaling PAS](#).

Deploying Ops Manager on vSphere

This topic provides instructions for deploying Pivotal Cloud Foundry (PCF) Ops Manager to VMware vSphere.

The instructions in this topic are based on vSphere 6.5 using the vSphere Client (HTML5) interface.

Runtime Considerations

For security purposes, VMware requires a dedicated instance of Ops Manager for use with Pivotal Container Service (PKS).

⚠ warning: If you are installing PKS on vSphere with NSX-T integration, follow the instructions in [Deploying Ops Manager with NSX-T for PKS](#) instead of performing the steps described in this topic.

Prepare vSphere

1. Log into vCenter using the vSphere Client (HTML5).
2. Select the vSphere Datacenter where you are deploying Ops Manager.
3. Create one or more vSphere clusters. A vSphere cluster is a collection of ESXi hosts and associated virtual machines (VMs) with shared resources and a shared management interface. Follow the steps below to create a vSphere cluster:

Note: For highly available, production-ready environments, you must create more than one cluster. For more information, see [Compute and HA Considerations](#) and [Scaling and Capacity Management](#) in *vSphere Reference Architecture*.

- a. Right-click the vSphere Datacenter object.
 - b. Select **New Cluster**.
 - c. Name the Cluster. For example, `Cluster1`.
 - d. Select **Turn ON vSphere DRS** to enable DRS. Set DRS Automation to **Partially Automated** or **Full Automated**.
 - e. Select **Turn ON vSphere HA** to enable vSphere HA.
4. If you created more than one vSphere cluster, configure shared storage for the clusters. You can allocate shared storage to the clusters using one of two common approaches: horizontal or vertical. For more information about configuring shared storage for clusters, see [Storage Design](#) in *vSphere Reference Architecture*.

⚠ Warning: Failure to properly configure shared storage for vSphere clusters could result in downtime and unrecoverable loss of data.

5. Create a resource pool for the management components in your deployment. Resource pools are vSphere objects used to organize resources.

Note: The management resource pool separates non-HA components such as Ops Manager and BOSH Director. For example, if you deploy Ops Manager for use with PKS, the PKS management plane includes Ops Manager, BOSH Director, the PKS control plane, and Harbor(optional).

Follow the steps below to create a vSphere resource pool:

- a. Right-click the vSphere cluster object that you created.
- b. Select **New Resource Pool**.
- c. Name the resource pool. For example, `RP-MGMT`.
- d. Review the default CPU and Memory settings and adjust if necessary. This is not common.
- e. Click **OK**.

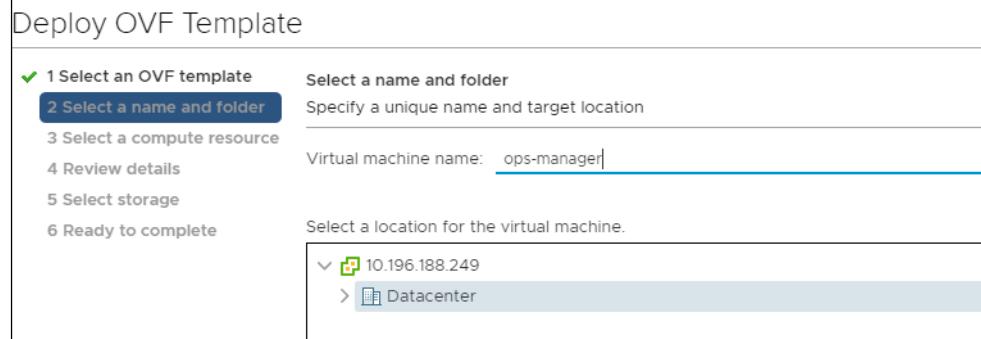
Deploy Ops Manager

Ops Manager for vSphere is provided as an OVA file. An OVA file is a template for a VM. To deploy Ops Manager on vSphere, you import the OVA file into your vSphere environment.

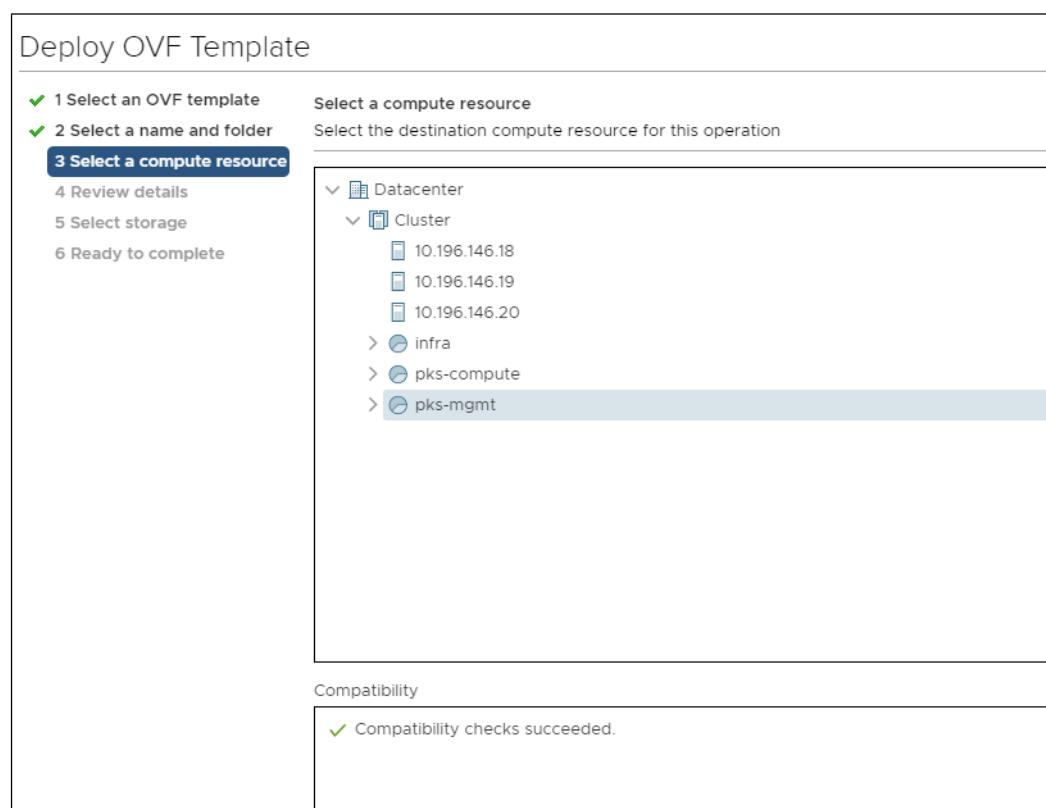
1. Download the [Pivotal Cloud Foundry Operations Manager for vSphere](#) OVA file from [Pivotal Network](#). Use the **Releases** dropdown menu to

select an Ops Manager release.

2. Right-click the management resource pool that you created and select **Deploy OVF Template**. Alternatively, from the vCenter menu, select **Actions > Deploy OVF Template**.
3. At the **Select an OVF template** screen, click **Browse**.
4. Select the Ops Manager OVA file that you downloaded and click **Open**.
5. Review template selection and click **Next**.
6. At the **Select a name and folder** screen:
 - o Enter a name for the Ops Manager virtual machine, or use the default name.
 - o Select the **Datacenter** object.
 - o Click **Next**.

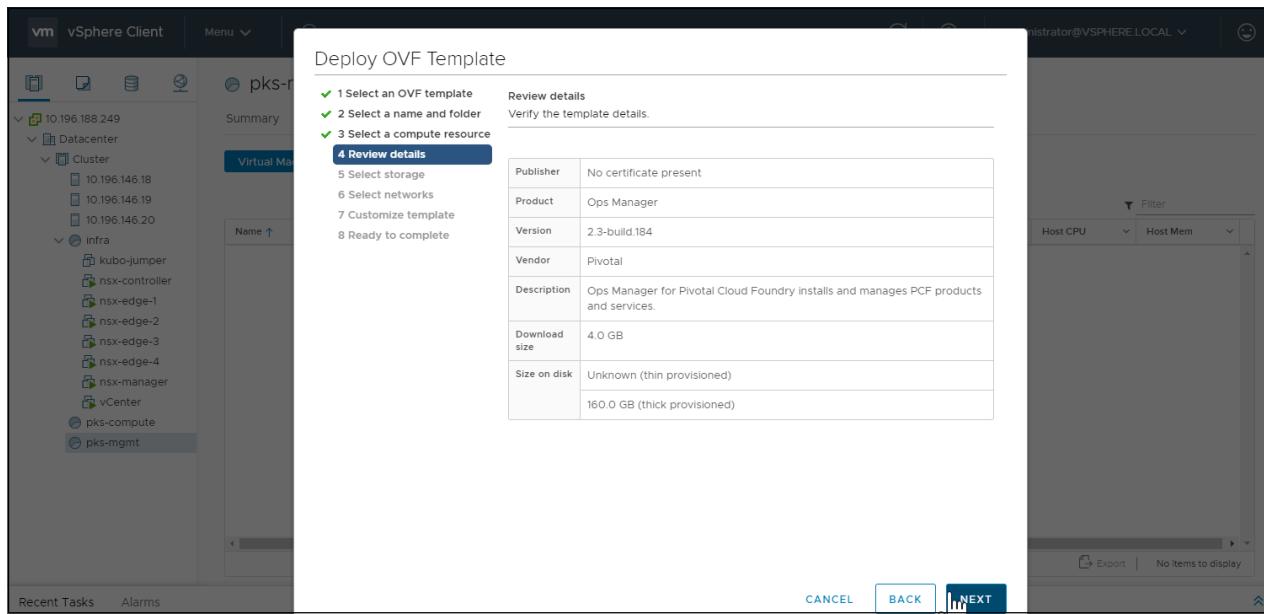


7. At the **Select a compute resource** screen, select the compute resource where the Ops Manager VM will run and click **Next**. Select the Resource Pool for the management components.



Note: Hardware virtualization must be turned off if your vSphere host does not support VT-X/EPT. Refer to [vSphere Requirements](#) in *PCF on vSphere Requirements* for more information.

8. At the **Review Details** screen, confirm the configuration up to this point and click **Next**.



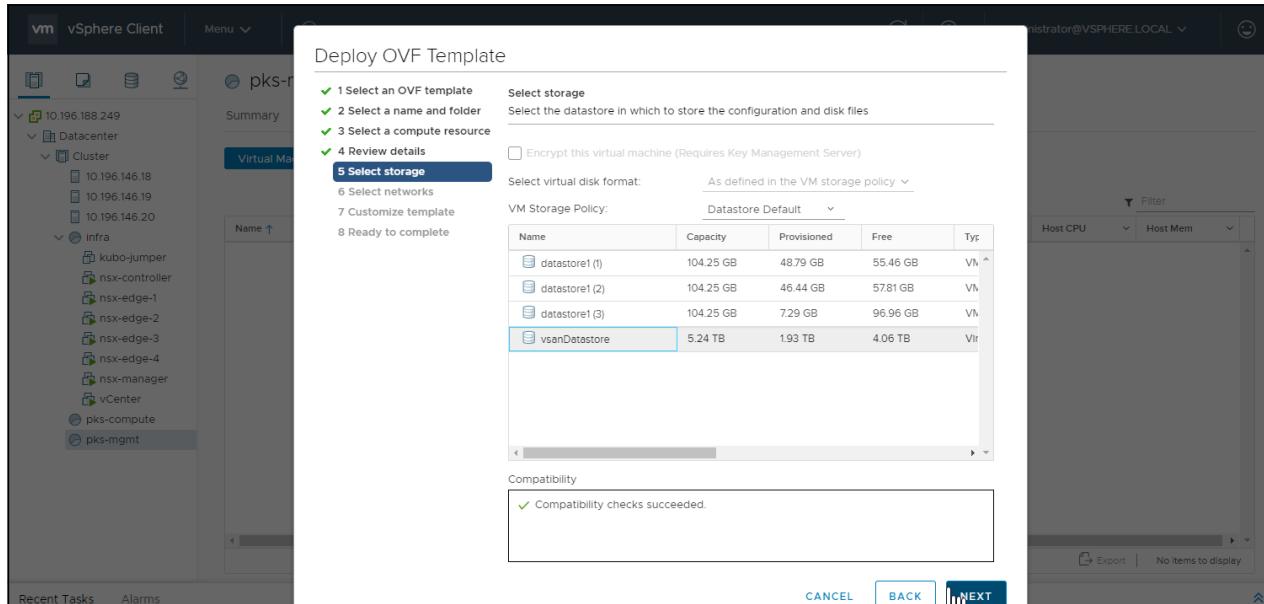
9. At the **Select Storage** screen, select the virtual disk format.

a. Choose a Datastore.

- If you use vSAN, select the vSAN datastore and apply the default storage policy.
- If you use a dedicated disk, select the appropriate disk format for your deployment needs. If in doubt, choose **Thin Provision**. For more information about disk formats, see [vSphere Virtual Disk Types](#).

b. Click **Next**.

⚠ warning: Ops Manager requires a VM with at least 8 GB memory.



10. At the **Select Networks** screen, select the standard **VM Network**, and click **Next**.

Deploy OVF Template

✓ 1 Select an OVF template **Select networks**
 ✓ 2 Select a name and folder Select a destination network for each source network.
 ✓ 3 Select a compute resource
 ✓ 4 Review details
 ✓ 5 Select storage
6 Select networks
 7 Customize template
 8 Ready to complete

Source Network	Destination Network
Network 1	VM Network

IP Allocation Settings

IP allocation: Static - Manual
 IP protocol: IPv4

CANCEL BACK **NEXT**

11. At the **Customize template** screen, enter the following information, and click **Next**.

- **IP Address:** The IP address of the Ops Manager network interface. For example, `10.0.0.2`.
- **Netmask:** The network mask for Ops Manager. For example, `255.255.255.0`.
- **Default Gateway:** The default gateway for Ops Manager to use. For example, `10.0.0.1`.
- **DNS:** One or more DNS servers for the Ops Manager VM to use. For example, `10.20.20.1`.
- **NTP Servers:** The IP address of one or more NTP servers for Ops Manager. For example, `10.113.60.176`.
- **Public SSH Key:** Enter pre-existing public SSH key value to allow SSH access to the Ops Manager VM.
- **Custom hostname:** The hostname for the Ops Manager VM. For example, `ops-manager`.

Deploy OVF Template

Customize template
Customize the deployment properties of this software solution.

All properties have valid values X

Uncategorized		8 settings
IP Address	The IP address for the Ops Manager. Leave blank if DHCP is desired. <input type="text" value="10.196.188.27"/>	
Netmask	The netmask for the Ops Manager's network. Leave blank if DHCP is desired. <input type="text" value="255.255.255.0"/>	
Default gateway	The default gateway address for the Ops Manager's network. Leave blank if DHCP is desired. <input type="text" value="10.196.188.253"/>	
DNS	The domain name servers for the Ops Manager (comma separated). Leave blank if DHCP is desired. <input type="text" value="10.40.53.1"/>	

CANCEL **BACK** **NEXT**

vm vSphere Client Menu vSphere Client Deploy OVF Template

Administrator@VSPHERE.LOCAL

Deploy OVF Template

1 Select an OVF template
2 Select a name and folder
3 Select a compute resource
4 Review details
5 Select storage
6 Select networks
7 Customize template
8 Ready to complete

DNS	The domain name servers for the Ops Manager (comma separated). Leave blank if DHCP is desired. <input type="text" value="10.40.53.1"/>
NTP Servers	Comma-delimited list of NTP servers <input type="text" value="10.142.7.21"/>
Admin Password	This password is used to SSH into the Ops Manager. The username is 'ubuntu'. One or both of Admin Password and SSH Key is required. Password <input type="password"/> Confirm Password <input type="password"/>
Public SSH Key	The Public SSH Key is used to allow SSHing into the Ops Manager with your private ssh key. The username is 'ubuntu'. One or both of Admin Password and SSH Key is required. <input type="text" value="AAAAAB3tNzaC1yc2EAAAQAB"/>
Custom Hostname	This will be set as the hostname on the VM. Default: 'pivotal-ops-manager'. <input type="text" value="ops-manager"/>

CANCEL **BACK** **NEXT**

12. At the **Ready to complete** screen, review the configuration settings and click **Finish**.

This action begins the OVA import and deployment process.

Deploy OVF Template

Click Finish to start creation.

Provisioning type	Deploy from template
Name	ops-manager
Template name	pivotal-ops-manager
Download size	4.0 GB
Size on disk	160.0 GB
Folder	Datacenter
Resource	pks-mgmt
Storage mapping	1
All disks	Datastore: vsanDatastore; Format: As defined in the VM storage policy
Network mapping	1
Network 1	VM Network
IP allocation settings	
IP protocol	IPV4
IP allocation	Static - Manual

8 Ready to complete

CANCEL **BACK** **FINISH**

13. Use the **Recent Tasks** panel at the bottom of the vCenter dashboard to check the progress of the OVA import and deployment.

If the import or deployment is unsuccessful, check the configuration for errors. Fix any errors and restart the deployment process.

Start Ops Manager

1. Create a DNS entry for the IP address that you used for Ops Manager. You must use this fully qualified domain name when you log into Ops Manager.

Note: Ops Manager security features require you to create a fully qualified domain name to access Ops Manager during the initial configuration. For more information, see [PCF on vSphere Requirements](#).

2. Select the Ops Manager VM and click **Actions > Power > Power On**.

Once the VM boots, the Ops Manager web interface is available at the IP address that you specified.

Note: It is normal to experience a brief delay before the interface is accessible while the web server and VM start.

3. Navigate to the fully qualified domain of your Ops Manager in a web browser.

Note: It is normal to experience a brief delay before the interface is accessible while the web server and VM start.

Next Steps

After you complete this procedure, follow the instructions in [Configuring BOSH Director on vSphere](#).

Configuring BOSH Director on vSphere

This topic describes how to configure the BOSH Director tile for VMware vSphere.

You can also perform the procedures in this topic using the Ops Manager API. For more information, see [Using the Ops Manager API](#).

⚠ warning: If you are installing Pivotal Container Service (PKS) on vSphere with NSX-T integration, follow the instructions in [Configuring BOSH Director with NSX-T for PKS](#) instead of performing the procedure described in this topic.

After you complete this procedure, follow the configuration instructions for the runtime that you choose to install.

For example:

- To install Pivotal Application Service (PAS), see [Configuring PAS](#).
- To install PKS, see [Installing PKS on vSphere](#).

See [Installing Runtimes](#) for more information about Pivotal Cloud Foundry (PCF) runtimes.

Prerequisite

Before you begin this procedure, you must complete all steps in [Deploying Ops Manager to vSphere](#).

Step 1: Set Up Ops Manager Authentication

1. In a web browser, navigate to the fully qualified domain of your Ops Manager.
2. When Ops Manager starts for the first time, you must choose one of the following:
 - [Internal Authentication](#): If you use Internal Authentication, Ops Manager maintains your user database.
 - [SAML Identity Provider](#): If you use a SAML Identity Provider (IdP), an external identity server maintains your user database.
 - [LDAP Server](#): If you use a LDAP Server, an external identity server maintains your user database.

Internal Authentication

1. When redirected to the [Internal Authentication](#) page, you must complete the following steps:
 - Enter a **Username**, **Password**, and **Password confirmation** to create a user with administrative privileges.
 - Enter a **Decryption passphrase** and the **Decryption passphrase confirmation**. This passphrase encrypts the Ops Manager datastore, and is not recoverable.
 - If you are using an **HTTP proxy** or **HTTPS proxy**, follow the instructions in the [Configuring Proxy Settings for the BOSH CPI](#) topic.
 - Read the **End User License Agreement**, and select the checkbox to accept the terms.
 - Click **Setup Authentication**.
2. Log in to Ops Manager with the username and password you created in the previous step.
3. Verify success.

You should be able to log in, and you should see the BOSH Director tile is present and ready for configuration, indicated by the orange color.

SAML Identity Provider

1. Log in to your IdP console and download the IdP metadata XML. Optionally, if your IdP supports metadata URL, you can copy the metadata URL instead of the XML.
2. Copy the IdP metadata XML or URL to the Ops Manager **SAML Identity Provider** login page.

Note: The same IdP metadata URL or XML is applied for the BOSH Director. If you use a separate IdP for BOSH, copy the metadata XML or URL from that IdP and enter it into the BOSH IdP Metadata text box in the Ops Manager login page.

3. Enter values for the fields listed below. Failure to provide values in these fields results in a `500` error.
 - **SAML admin group:** Enter the name of the SAML group that contains all Ops Manager administrators. This field is case-sensitive.
 - **SAML groups attribute:** Enter the groups attribute tag name with which you configured the SAML server. This field is case-sensitive.
4. Enter your **Decryption passphrase**. Read the **End User License Agreement**, and select the checkbox to accept the terms.
5. Your Ops Manager login page appears. Enter your username and password. Click **Login**.
6. Download your SAML Service Provider metadata (SAML Relying Party metadata) by navigating to the following URLs:
 - **6a. Ops Manager SAML service provider metadata:** `https://OPS-MAN-FQDN:443/uaa/saml/metadata`
 - **6b. BOSH Director SAML service provider metadata:** `https://BOSH-IP-ADDRESS:8443/saml/metadata`

Note: To retrieve your `BOSH-IP-ADDRESS`, navigate to the **Status** tab in the **BOSH Director** tile. Record the **BOSH Director** IP address.
7. Configure your IdP with your SAML Service Provider metadata. Import the Ops Manager SAML provider metadata from Step 6a above to your IdP. If your IdP does not support importing, provide the values below.
 - **Single sign on URL:** `https://OPS-MAN-FQDN:443/uaa/saml/SSO/alias/OPS-MAN-FQDN`
 - **Audience URI (SP Entity ID):** `https://OP-MAN-FQDN:443/uaa`
 - **Name ID:** Email Address
 - SAML authentication requests are always signed
8. Import the BOSH Director SAML provider metadata from Step 6b to your IdP. If the IdP does not support an import, provide the values below.
 - **Single sign on URL:** `https://BOSH-IP:8443/saml/SSO/alias/BOSH-IP`
 - **Audience URI (SP Entity ID):** `https://BOSH-IP:8443`
 - **Name ID:** Email Address
 - SAML authentication requests are always signed
9. Return to the **BOSH Director** tile, and continue with the configuration steps below.

LDAP Server

- For **Server URL**, enter the URL that points to your LDAP server. With multiple LDAP servers, separate their URLs with spaces. Each URL must include one of the following protocols:
 - `ldap://`: This specifies that the LDAP server uses an unencrypted connection.
 - `ldaps://`: This specifies that the LDAP server uses SSL for an encrypted connection and requires that the LDAP server holds a trusted certificate or that you import a trusted certificate to the JVM truststore.
 - For **LDAP Username** and **LDAP Password**, enter the LDAP Distinguished Name (DN) and the password for binding to the LDAP Server. Example DN: `cn=administrator,ou=Users,dc=example,dc=com`
- Note:** Pivotal recommends that you provide LDAP credentials that grant read-only permissions on the LDAP Search Base and the LDAP Group Search Base. In addition to this, if the bind user belongs to a different search base, you must use the full DN.
- warning:** Pivotal recommends against reusing LDAP service accounts across environments. LDAP service accounts should not be subject to manual lockouts, such as lockouts that result from users utilizing the same account. Also, LDAP service accounts should not be subject to automated deletions, since disruption to these service accounts could prevent user logins.

- For **User Search Base**, enter the location in the LDAP directory tree fromwhich any LDAP User search begins. The typical LDAP Search Base matches yourdomain name.

For example, a domain named “cloud.example.com” typically uses the following LDAP User Search Base: `ou=Users,dc=example,dc=com`

- For **User Search Filter**, enter a string that defines LDAP User search criteria. These search criteria allow LDAP to perform more effective and efficient searches. For example, the standard LDAP search filter `cn=Smith` returns all objects with a common name equal to `Smith`.

In the LDAP search filter string that you use to configure your runtime, use `{0}` instead of the username. For example, use `cn={0}` to return all LDAP objects with the same common name as the username.

In addition to `cn`, other attributes commonly searched for and returned are `mail`, `uid` and, in the case of Active Directory, `sAMAccountName`.

Note: For instructions for testing and troubleshooting your LDAP search filters, see [Configuring LDAP Integration with Pivotal Cloud Foundry](#) in the Pivotal Support Knowledge Base.

- For **Group Search Base**, enter the location in the LDAP directory tree from which the LDAP Group search begins.

For example, a domain named “cloud.example.com” typically uses the following LDAP Group Search Base: `ou=Groups,dc=example,dc=com`

- For **Group Search Filter**, enter a string that defines LDAP Group search criteria. The standard value is `member={0}`.
- For **Email Attribute**, enter the attribute name in your LDAP directory that corresponds to the email address in each user record, for example `mail`.
- For **LDAP RBAC Admin Group Name**, enter the DN of the LDAP group you want to have admin permissions in Ops Manager.
- From the dropdown, select how the UAA handles LDAP server referrals out to other external user stores. The UAA can:
 - Automatically follow any referrals.
 - Ignore referrals and return partial result.
 - Throw exception for each referral and abort.
- For **Server SSL Cert**, paste in the root certificate from your CA certificate or your self-signed certificate.
- Enter a **Decryption passphrase** and the **Decryption passphrase confirmation**. This passphrase encrypts the Ops Manager datastore, and is not recoverable.
- If you are using an **HTTP proxy** or **HTTPS proxy**, follow the instructions in [Configuring Proxy Settings for the BOSH CPI](#).
- Read the **End User License Agreement**, and select the checkbox to accept the terms.
- Select **Provision an admin client in the BOSH UAA**. You can use this to enable BOSH automation with scripts and tooling. For more information, see [Provision Admin Client](#) in *Creating UAA Clients for BOSH Director*.
- Click **Setup Authentication**.
- Return to the **BOSH Director** tile, and continue with the configuration steps below.

Step 2: Configure vCenter

- Click the **BOSH Director for vSphere** tile.
- Select **vCenter Config**. Configure the one or more vCenters that host your PCF foundation. The following image shows the **vCenter Config** pane.

vCenter Config
[Add vCenter Config](#)

Name*
 [Edit](#)

vCenter Host*

vCenter Username*

vCenter Password*
 [Change](#)

Datacenter Name*

Virtual Disk Type*
 [Change](#)

Ephemeral Datastore Names (comma delimited)*

NOTE: Removing an Ephemeral Datastore after an initial deploy can result in a system outage and/or data loss.

Persistent Datastore Names (comma delimited)*

NOTE: Removing a Persistent Datastore after an initial deploy can result in a system outage and/or data loss.

Standard vCenter
 NSX Networking

NSX Mode*
 NSX-V
 NSX-T

NSX Address*

NSX Username*

NSX Password*

NSX CA Cert

3. Complete the fields on the **vCenter Config** pane as follows:

- **Name:** A name that you provide for your vCenter configuration. This field is used to identify the datacenter configuration in Ops Manager if you are configuring multiple datacenters.
- **vCenter Host:** The hostname of the vCenter that manages ESXi/vSphere.
- **vCenter Username:** A vCenter username with create and delete privileges for virtual machines (VMs) and folders.
- **vCenter Password:** The password for the vCenter user specified above.
- **Datacenter Name:** The name of the datacenter as it appears in vCenter.
- **Virtual Disk Type:** The Virtual Disk Type to provision for all VMs. For guidance on selecting a virtual disk type, see [vSphere Virtual Disk Types](#).

 **Note:** Datastores do not support whitespace characters in their names. Including whitespace characters in datastore names causes an error.

- **Ephemeral Datastore Names (comma delimited):** The names of the datastores that store ephemeral VM disks deployed by Ops Manager.
- **Persistent Datastore Names (comma delimited):** The names of the datastores that store persistent VM disks deployed by Ops Manager.

4. Select a network configuration from one of the following:

- **Standard vCenter Networking:** This is the default option when upgrading Ops Manager.

 **warning:** If you intend to deploy PKS using Flannel as the CNI (as opposed to NSX-T), then select this option to configure the BOSH Director. Do not choose **NSX Networking**. When you configure the PKS tile later on, you will configure Flannel as the container network interface (CNI).

If you intend to deploy PKS with NSX-T, see [Installing PKS on vSphere with NSX-T Data Center](#).

- **NSX Networking:** Select this option to enable VMware NSX Network Virtualization for non-PKS runtimes such as Pivotal Application Service (PAS).

Configure NSX networking by entering the following information:

- **NSX Mode:** Select either **NSX-V** or **NSX-T**.
- **NSX Address:** The address of the NSX manager.
- **NSX Username:** The username to connect to the NSX manager.
- **NSX Password:** The password for the username specified above.
- **NSX CA Cert:** A CA certificate in PEM format that authenticates to the NSX server. If the NSX Manager generated a self-signed certificate, use the following command to retrieve the CA certificate using OpenSSL:

```
openssl s_client -showcerts -connect NSX-MANAGER-ADDRESS:443 </dev/null 2> /dev/null | openssl x509
```

Where `NSX-MANAGER-ADDRESS` is the address of the NSX manager.

5. Configure the following folder names:

- **VM Folder:** The vSphere datacenter folder where Ops Manager places VMs. Enter `[YOUR-DEPLOYMENT]_vms` where `YOUR-DEPLOYMENT` corresponds to a descriptive name for your deployment. For example, `my_pcf_vms`.
- **Template Folder:** The vSphere datacenter folder where Ops Manager places VMs. Enter `[YOUR-DEPLOYMENT]_templates`. For example, `my_pcf_templates`.
- **Disk path Folder:** The vSphere datastore folder where Ops Manager creates attached disk images. You must not nest this folder. Enter `[YOUR-DEPLOYMENT]_disk`. For example, `my_pcf_disk`.

 **Note:** After initial deployment, you cannot edit folder names.

6. Click **Save**.

7. (Optional) Click **Add vCenter Config** toward the top of the form to configure additional vCenters. Once you click **Save**, your multiple vCenter Configs are listed in the **vCenter Configs** pane. For more information about multiple vCenter configs, see [Managing Multiple Data Centers](#).

Step 3: Director Config Page

1. In Ops Manager, select **Director Config**.

Director Config

NTP Servers (comma delimited)*

Bosh HM Forwarder IP Address

Enable VM Resurrector Plugin

Enable Post Deploy Scripts

Recreate All VMs
 This will force BOSH to recreate all VMs on the next deploy. Persistent disk will be preserved.

Recreate All Persistent Disks
 Check this box to recreate all Persistent Disks for the Director and all other Tiles.

Enable bosh deploy retries
 This will attempt to re-deploy a failed deployment up to 5 times.

2. In the **NTP Servers (comma delimited)** field, enter a comma-separated list of valid NTP servers.

Note: The NTP server configuration only updates after VM recreation. Ensure that you select the **Recreate all VMs** checkbox if you modify the value of this field.

3. Leave the **Bosh HM Forwarder IP Address** field blank.

Note: Starting in PAS v2.0, BOSH-reported component metrics are available in the Loggregator Firehose by default. If you continue to use the BOSH HM Forwarder to consume these component metrics, you may receive duplicate data. To prevent this, leave the **Bosh HM Forwarder IP Address** field blank.

4. Select the **Enable VM Resurrector Plugin** checkbox to enable the BOSH Resurrector functionality and increase your runtime availability. If you intend to install PAS, see [Using BOSH Resurrector and vSphere HA](#) for more information about the Resurrector plugin.
5. Select **Enable Post Deploy Scripts** to run a post-deploy script after deployment. This script allows the job to execute additional commands against a deployment.
- Note:** If you intend to install PKS, you must enable post-deploy scripts.
6. Select **Recreate all VMs** to force BOSH to recreate all VMs on the next deploy. This process does not destroy any persistent disk data.
7. Select **Recreate All Persistent Disks** to force BOSH to migrate and recreate persistent disks for the BOSH Director and all tiles. This process does not destroy any persistent disk data.
8. Select **Enable bosh deploy retries** to instruct Ops Manager to retry failed BOSH operations up to five times.
9. Select **Skip Director Drain Lifecycle** to prevent drain scripts from running when the BOSH Director is recreated.
10. Select **Store BOSH Job Credentials on tmpfs (beta)** to store credentials for BOSH jobs on temporary file storage (tmpfs) memory, rather than on disk. You must recreate all VMs for this setting to take effect.
11. Select **Keep Unreachable Director VMs** if you want to preserve BOSH Director VMs after a failed deployment for troubleshooting purposes.
12. Select **HM Pager Duty Plugin** to enable Health Monitor integration with PagerDuty.

HM Pager Duty Plugin

Service Key*

YOUR-PAGERDUTY-SERVICE-KEY

HTTP Proxy

YOUR-HTTP-PROXY

- **Service Key:** Enter your API service key from PagerDuty.
- **HTTP Proxy:** Enter an HTTP proxy for use with PagerDuty.

13. Select **HM Email Plugin** to enable Health Monitor integration with email.

HM Email Plugin

Host*

smtp.example.com

Port*

25

Domain*

cloudfoundry.example.com

From*

user2@example.com

Recipients*

user@example.com, user1@example.com

Username

user

Password

.....

Enable TLS

- **Host:** Enter your email hostname.
- **Port:** Enter your email port number.
- **Domain:** Enter your domain.
- **From:** Enter the address for the sender.
- **Recipients:** Enter comma-separated addresses of intended recipients.
- **Username:** Enter the username for your email server.
- **Password:** Enter the password for your email server.
- **Enable TLS:** Select this checkbox to enable Transport Layer Security to the email host.

14. For **CredHub Encryption Provider**, you can choose whether BOSH CredHub stores its encryption key internally on the BOSH Director and CredHub VM, or in an external hardware security module (HSM). The HSM option is more secure.

Before configuring an HSM encryption provider in the **Director Config** pane, you must follow the procedures and collect information described in [Preparing CredHub HSMs for Configuration](#).

 **Note:** After you deploy Ops Manager with an HSM encryption provider, you cannot change BOSH CredHub to store encryption keys internally.

CredHub Encryption Provider

Internal
 Luna HSM

Encryption Key Name*

Provider Partition*

Provider Partition Password*

Provider Client Certificate*

Provider Client Certificate Private Key*

HSM Host Address*

HSM Port Address*

Partition Serial Number*

HSM Certificate*

- **Internal:** Select this option for internal CredHub key storage. This option is selected by default and requires no additional configuration.
 - **Luna HSM:** Select this option to use a SafeNet Luna HSM as your permanent CredHub encryption provider, and fill in the following fields:
 1. **Encryption Key Name:** Any name to identify the key that the HSM uses to encrypt and decrypt the CredHub data. Changing this key name after you deploy Ops Manager can cause service downtime.
 2. **Provider Partition:** The partition that stores your encryption key. Changing this partition after you deploy Ops Manager could cause service downtime. For this value and the ones below, use values gathered in [Preparing CredHub HSMs for Configuration](#).
 3. **Provider Partition Password**
 4. **Provider Client Certificate:** The certificate that validates the identity of the HSM when CredHub connects as a client.
 5. **Provider Client Certificate Private Key**
 6. **HSM Host Address**
 7. **HSM Port Address:** If you do not know your port address, enter `1792`.
 8. **Partition Serial Number**
 9. **HSM Certificate:** The certificate that the HSM presents to CredHub to establish a two-way mTLS connection.
15. Select a **Blobstore Location** to either configure the blobstore as an internal server or an external endpoint. Because the internal server is unscalable and less secure, Pivotal recommends that you configure an external blobstore.

Note: After you deploy Ops Manager, you cannot change the blobstore location.

- **Internal:** Select this option to use an internal blobstore. Ops Manager creates a new VM for blob storage. No additional configuration is required.
 - **Enable TLS:** Select this checkbox to enable TLS to the blobstore.

Note: If you are using PASW 2016, make sure you have downloaded [Windows stemcell v1709.10 or higher](#) before enabling TLS.

- **S3 Compatible Blobstore:** Select this option to use an external S3-compatible endpoint.

Blobstore Location

Internal

Enable TLS

S3 Compatible Blobstore

S3 Endpoint*

Bucket Name*

Access Key*

Secret Key*

S3 Signature Version

V2 Signature

V4 Signature

Region*

S3 Backup Strategy

No backups

Use a versioned bucket

Copy into an additional bucket

Backup Bucket Region*

Backup Bucket Name*

Follow the procedures in [Sign up for Amazon S3](#) and [Creating a Bucket](#) in the AWS documentation. When you have created an S3 bucket, complete the following steps:

1. **S3 Endpoint:** Navigate to the [Regions and Endpoints](#) topic in the AWS documentation.
 - a. Locate the endpoint for your region in the **Amazon Simple Storage Service (S3)** table and construct a URL using your region's endpoint. For example, if you are using the `us-west-2` region, the URL you create would be `https://s3-us-west-2.amazonaws.com`. Enter this URL into the **S3 Endpoint** field.
 - b. On a command line, run `ssh ubuntu@OPS-MANAGER-FQDN` to SSH into the Ops Manager VM. Replace `OPS-MANAGER-FQDN` with the fully qualified domain name of Ops Manager.
 - c. Copy the custom public CA certificate you used to sign the S3 endpoint into `/etc/ssl/certs` on the Ops Manager VM.

- d. On the Ops Manager VM, run `sudo update-ca-certificates -f -v` to import the custom CA certificate into the Ops Manager VM truststore.

Note: You must also add this custom CA certificate into the **Trusted Certificates** field in the **Security** page. See [Security Page](#) for instructions.

2. **Bucket Name:** Enter the name of the S3 bucket.
3. **Access Key** and **Secret Key:** Enter the keys you generated when creating your S3 bucket.
4. Select **V2 Signature** or **V4 Signature**. If you select **V4 Signature**, enter your **Region**.

Note: AWS recommends using Signature Version 4. For more information about AWS S3 Signatures, see [Authenticating Requests](#) in the AWS documentation.

5. **S3 Backup Strategy:** Select one of the following to configure whether and how to back up the BOSH Director's S3 blobstore:
 - **No backups**
 - **Use a versioned bucket:** This option uses blobstore bucket versioning to save backups. It requires an S3 blobstore that supports versioning.
 - **Copy into an additional bucket:** This option saves blobstore backups to a separate bucket, which is useful with blobstores that do not support versioning. The option requires you to fill in **Backup Bucket Region** and **Backup Bucket Name**.

- **GCS Blobstore:** Select this option to use an external Google Cloud Storage (GCS) endpoint.

The screenshot shows a configuration form for a GCS Blobstore. At the top, there is a radio button labeled "GCS Blobstore". Below it, there are three input fields: "Bucket Name*" with a placeholder box, "Storage Class*" with a dropdown menu set to "Regional", and "Service Account Key*" with a large text area for pasting JSON key contents.

Follow the procedures in [Creating Storage Buckets](#) in the GCS documentation to create a GCS bucket. To create a GCS bucket, you must have a GCS account. When you have created a GCS bucket, complete the following steps:

1. **Bucket Name:** Enter the name of your GCS bucket.
 2. **Storage Class:** Select the storage class for your GCS bucket. See [Storage Classes](#) in the GCP documentation for more information.
 3. **Service Account Key:** Follow the steps in the [Set up IAM Service Accounts](#) section of *Preparing to Deploy Ops Manager on GCP Manually* to download a JSON file with a private key. Enter the contents of the JSON file into the field.
16. Select a **Database Location**. By default, Ops Manager deploys and manages an **Internal** database for you. If you choose to use an **External MySQL Database**, complete the associated fields with information obtained from your external MySQL Database provider: **Host**, **Port**, **Username**,

Database Location

Internal
 External MySQL Database

Host*

Port*

Username*

Password*

Database*

Password, and Database.

In addition, if you selected the **Enable TLS for Director Database** checkbox, you can complete the following optional fields:

- **Enable TLS:** Select this checkbox enables TLS communication between the BOSH Director and the database.
- **TLS CA:** Enter the Certificate Authority for the TLS Certificate.
- **TLS Certificate:** Enter the client certificate for mutual TLS connections to the database.
- **TLS Private Key:** Enter the client private key for mutual TLS connections to the database.
- **Advanced DB Connection Options:** If you would like to provide additional options for the database, use this field to provide a JSON-formatted options string.

17. (Optional) Modify the **Director Workers** value, which sets the number of workers available to execute Director tasks. This field defaults to .
18. (Optional) **Max Threads** sets the maximum number of threads that the BOSH Director can run simultaneously. Pivotal recommends that you leave the field blank to use the default value, unless doing so results in rate limiting or errors on your IaaS.
19. (Optional) To add a custom URL for your BOSH Director, enter a valid hostname in **Director Hostname**. You can also use this field to configure a load balancer in front of your BOSH Director. For more information, see [How to Set Up a Load Balancer in Front of Operations Manager Director](#) in the Pivotal Support Knowledge Base.

Director Workers

Max Threads

Director Hostname

20. (Optional) To set a custom banner that users see when logging in to the Director using SSH, enter text in the **Custom SSH Banner** field.
21. (Optional) Enter your comma-separated custom **Identification Tags**. For example, . You can use the tags to identify your foundation when viewing VMs or disks from your IaaS.
22. Click **Save**.

Step 4: Create Availability Zones Page

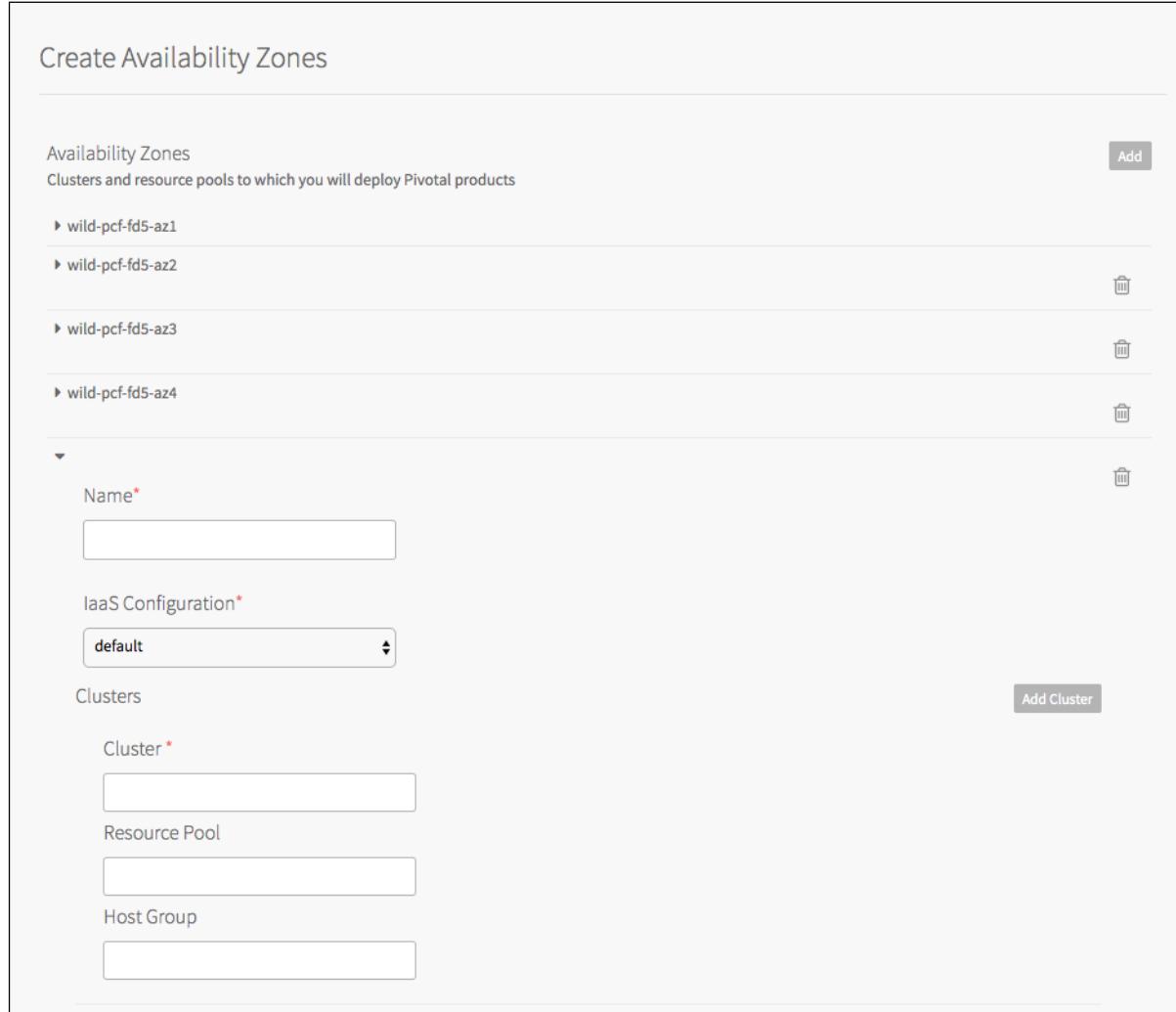
Ops Manager Availability Zones (AZs) correspond to your vCenter clusters and resource pools.

Multiple AZs allow you to provide high availability and load balancing to your applications. When you run more than one instance of an application, Ops Manager balances those instances across all of the AZs assigned to the application.

Pivotal recommends that you use at least three AZs for a highly available installation of your chosen runtime.

 **Note:** For more information about AZs and high availability in vSphere, see [Compute and HA Considerations](#) in *vSphere Reference Architecture*.

1. Select **Create Availability Zones** to navigate to the pane.



Create Availability Zones

Availability Zones
Clusters and resource pools to which you will deploy Pivotal products

Add

- ▶ wild-pcf-fd5-az1
- ▶ wild-pcf-fd5-az2
- ▶ wild-pcf-fd5-az3
- ▶ wild-pcf-fd5-az4

Name*

IaaS Configuration*

default

Clusters

Cluster*

Resource Pool

Host Group

Add Cluster

2. Use the following steps to create the first AZ.

- Click **Add**.
- Enter a unique **Name** for the AZ.
- Select a vCenter config name from the **IaaS Configuration** dropdown to associate your AZ with a vCenter. If you have only one vCenter config, or if you are creating your first AZ, IaaS Configuration defaults to the first vCenter and cannot be configured.
- Enter the name of the vCenter **Cluster** where you deployed Ops Manager. You created this cluster in the [Prepare vSphere](#) step of *Deploying Ops Manager to vSphere*. For example, `Cluster1`.
- Enter the name of the **Resource Pool** that you created in the [Prepare vSphere](#) step of *Deploying Ops Manager to vSphere* to organize management components like Ops Manager and BOSH Director. For example, `RP-MGMT`. The jobs running in this AZ share the CPU and memory resources defined by the pool. You only need to specify this resource group in one AZ.
- (Optional) Enter the name of a **Host Group** for the AZ to use. The host group must be within the cluster you designated in the **Clusters** field.
- (Optional) Click **Add Cluster** to add more clusters to the AZ as required by your installation size and high availability requirements.

3. (Optional) Use the following steps to create additional AZs.

- Click **Add**.

- Specify Cluster and optionally, a Resource Pool.
- Click Add Cluster to assign more clusters to the AZ as necessary for high availability. Click the trash icon to delete a cluster. The first cluster cannot be deleted.

4. Click Save.

Step 5: Create Networks Page

- Select Create Networks.
- Select Enable ICMP checks to enable ICMP on your networks. Ops Manager uses ICMP checks to confirm that components within your network are reachable.
- Depending on the runtime that you are deploying, do one of the following:

- For PAS, click Add Network and create the following networks:

- infrastructure** : This network is for Ops Manager and the BOSH Director.
- pas** : This network is for all the PAS VMs including Gorouter, Diego cells, and Cloud Controller.
- services** : This network is for any service tiles to be deployed alongside PAS.

Use the values from the following table as a guide when you create each network, replacing the IP addresses with ranges that are available in your vSphere environment:

	Field	Configuration Example
Infrastructure Network	Name	infrastructure
	vSphere Network Name *	pcf-virt-net/infrastructure
	CIDR	192.168.101.0/24
	Reserved IP Ranges †	192.168.101.1-192.168.101.9
	DNS	192.168.101.2
	Gateway	192.168.101.1
Deployment Network	Field	Configuration Example
	Name	pas
	vSphere Network Name *	pcf-virt-net/pcf-pas-subnet
	CIDR	192.168.16.0/24
	Reserved IP Ranges †	192.168.16.1-192.168.16.9
	DNS	192.168.16.2
	Gateway	192.168.16.1
Services Network	Field	Configuration Example
	Name	services
	vSphere Network Name *	pcf-virt-net/pcf-services-subnet
	CIDR	192.168.20.0/24
	Reserved IP Ranges †	192.168.20.1-192.168.20.9
	DNS	192.168.20.2
	Gateway	192.168.20.1

* In vSphere Network Name, enter the full path of the network as it displays in vCenter. For example, enter `YOUR-DIRECTORY-NAME/YOUR-NETWORK-NAME`. If your vSphere Network Name contains a forward slash character, replace the forward slash with the URL-encoded forward slash character `%2f`.

† For Reserved IP Ranges, enter any IP addresses from the CIDR that you want to blacklist from the installation. Ops Manager will not deploy VMs to any address in this range.

- For PKS, click Add Network and create the following networks:

- infrastructure** : This network is for Ops Manager, the BOSH Director, the PKS broker, and the PKS API.
- pks** : If you have a large deployment with multiple tiles, you can choose to deploy the PKS broker and PKS API to a separate network

named `pks`. See the table below for more information.

- `services` : Network for creating the master and worker VMs for Kubernetes clusters. The CIDR should not conflict with the pod overlay network `10.200.0.0/16` or the reserved Kubernetes services CIDR of `10.100.200.0/24`.

 **Note:** Multiple networks allow you to place vCenter on a private network and the rest of your deployment on a public network. Isolating vCenter in this manner denies access to it from outside sources and reduces possible security vulnerabilities.

Use the values from the following table as a guide when you create each network, replacing the IP addresses with ranges that are available in your vSphere environment:

	Field	Configuration Example
Infrastructure Network	Name	<code>infrastructure</code>
	vSphere Network Name *	<code>pcf-virt-net/pks-infrastructure-subnet</code>
	CIDR	<code>192.168.101.0/26</code>
	Reserved IP Ranges †	<code>192.168.101.1-192.168.101.9</code>
	DNS	<code>192.168.101.2</code>
	Gateway	<code>192.168.101.1</code>
Main Network (Optional)	Field	Configuration Example
	Name	<code>pks</code>
	vSphere Network Name *	<code>pcf-virt-net/pks-subnet</code>
	CIDR	<code>192.168.16.0/26</code>
	Reserved IP Ranges †	<code>192.168.16.1-192.168.16.9</code>
	DNS	<code>192.168.16.2</code>
Service Network	Field	Configuration Example
	Name	<code>services</code>
	vSphere Network Name *	<code>pcf-virt-net/pks-service-subnet</code>
	CIDR	<code>192.168.20.0/22</code>
	Reserved IP Ranges †	<code>192.168.20.1-192.168.20.9</code>
	DNS	<code>192.168.20.2</code>
	Gateway	<code>192.168.20.1</code>

* In **vSphere Network Name**, enter the full path of the network as it displays in vCenter. For example, enter `YOUR-DIRECTORY-NAME/YOUR-NETWORK-NAME`. If your vSphere Network Name contains a forward slash character, replace the forward slash with the URL-encoded forward slash character `%2f`.

† For **Reserved IP Ranges**, enter any IP addresses from the **CIDR** that you want to blacklist from the installation. Ops Manager will not deploy VMs to any address in this range.

- For each network that you create, select the **Availability Zones** to use with the network. Assign as many AZs to your network as needed by your deployment. For more information, see [Compute and HA Considerations](#) and [Scaling and Capacity Management](#) sections of the vSphere Reference Architecture.

- Click **Save**.

Step 6: Assign AZs and Networks

- Select **Assign AZs and Networks**.

Assign AZs and Networks

The BOSH Director is a single instance.

Choose the availability zone in which to place that instance. It is highly recommended that you backup this VM on a regular basis to preserve settings.

Singleton Availability Zone

AZ-MGMT

Network

infrastructure

Save

2. Use the dropdown to select a **Singleton Availability Zone**. The BOSH Director installs as a single instance in this Availability Zone.

3. Use the dropdown to select a **Network** for your BOSH Director.

4. Click **Save**.

Step 7: Security Page

1. Select **Security**.

Security

Trusted Certificates

These certificates enable BOSH-deployed components to trust a custom root certificate.

Include OpsManager Root CA in Trusted Certs

Checking this box will place the OpsManager Root CA into the trusted certificate field. The Bosh Director will then place this CA into the trust store of every VM that it deploys.

Generate VM passwords or use single password for all VMs

Generate passwords

Use default BOSH password

Save

2. In **Trusted Certificates**, enter your custom certificate authority (CA) certificates to insert into your organization's certificate trust chain. This feature enables all BOSH-deployed components in your deployment to trust custom root certificates.

To enter multiple certificates, paste your certificates one after the other. For example, format your certificates like the following:

```
-----BEGIN CERTIFICATE-----  
ABCDEF12345678ABCDEF12345678ABCDEF12345678ABC  
EFGH12345678ABCDEF12345678ABCDEF12345678ABCDEF
```

```
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
BCDEFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
CDEFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABBB
EFGH12345678ABCDEFGH12345678ABCDEFGH12345678ABCDEF
GH12345678ABCDEFGH12345678ABCDEFGH12345678...
-----END CERTIFICATE-----
```

Note: If you want to use Docker registries to run PAS app instances in Docker containers, enter the certificate for your private Docker registry in this field. See [Using Docker Registries](#) for more information about running app instances in PAS using Docker registries.

3. (Optional) Select the **Include OpsManager Root CA in Trusted Certs** checkbox to include the Ops Manager root CA in the Trusted Certificates field. BOSH Director includes this CA in the trust store of every VM that it deploys.
4. Choose **Generate passwords** or **Use default BOSH password**. Pivotal recommends that you use the **Generate passwords** option for greater security.
5. Click **Save**. To view your saved Director password, click the **Credentials** tab.

Step 8: BOSH DNS Config Page

BOSH DNS Config

Excluded Recursors

Recursor Timeout

Handlers

Enter a JSON-formatted list of optional custom domain handlers

Save

1. Select **BOSH DNS Config**.
2. (Optional) In **Excluded Recursors**, enter a list of prohibited recursor addresses.
3. (Optional) In **Recursor Timeout**, enter a time limit for contacting the connected recursors. This includes dialing, writing, and reading from the recursor. If any of these actions exceeds the time limit you set, the action fails.

Note: This time limit must include one of the Go parse duration time units. For example, entering `5s` sets the timeout limit to five seconds. For more information about supported time units, see [func ParseDuration](#) in the Go Programming Language documentation.

4. (Optional) In **Handlers**, enter a list of custom domain handlers in JSON format.
5. Click **Save**.

Step 9: Syslog Page

Note: BOSH Director logs contain sensitive information that should be considered privileged. For example, these logs may contain cloud provider credentials. If you choose to forward logs to an external syslog endpoint, use TLS encryption to prevent information from being intercepted by a third party.

Settings

- [Change Decryption Passphrase](#)
- [SAML Settings](#)
- [LDAP Settings](#)
- [SSL Certificate](#)
- [Pivotal Network Settings](#)
- [Proxy Settings](#)
- [Custom Banner](#)
- [Export Installation Settings](#)
- [Syslog](#)
- [Advanced Options](#)

Syslog

Syslog settings make take up to 1 min to take effect.

Do you want to configure Syslog for OpsManager?

No

Yes

Address*

The address or host for the syslog server

Port*

Transport Protocol*

Enable TLS

Permitted Peer*

SSL Certificate*

Queue Size

Forward Debug Logs

Custom rsyslog Configuration

SAVE SYSLOG SETTINGS

1. Select **Syslog**.
2. (Optional) Select **Yes** to send BOSH Director system logs to a remote server.
3. Enter the IP address or DNS name for the remote server in **Address**.
4. Enter the port number that the remote server listens on in **Port**.
5. Select **TCP** or **UDP** from the **Transport Protocol** dropdown. This selection determines which transport protocol is used to send the logs to the remote server.
6. (Optional) Select the **Enable TLS** checkbox to send encrypted logs to remote server with TLS. After you select the checkbox, do the following:

- a. Enter either the name or SHA1 fingerprint of the remote peer in **Permitted Peer**.
- b. Enter the SSL certificate for the remote server in **SSL Certificate**.

Note: Pivotal strongly recommends that you enable TLS encryption when you are forwarding logs. Logs can contain sensitive information, such as cloud provider credentials.

7. (Optional) Enter an integer in **Queue Size**. This value specifies the number of log messages held in the buffer. The default value is 100,000.
8. (Optional) Select the checkbox to **Forward Debug Logs** to an external source. This option is deselected by default. If you select it, you may generate a large amount of log data.
9. (Optional) Enter configuration details for rsyslog in the **Custom rsyslog Configuration** field. This field requires the rainerscript syntax.
10. Click **Save Syslog Settings**.

Step 10: Resource Config Page

1. Select **Resource Config**.

JOB	INSTANCES	PERSISTENT DISK TYPE	VM TYPE
Ops Manager Director	Automatic: 1	Automatic: 50 GB	Automatic: medium.disk (cpu: 2, ram: 4 GB)
Master Compilation Job	Automatic: 4	None	Automatic: large.cpu (cpu: 4, ram: 4 GB, di)

Save

2. Adjust any values as necessary for your deployment. Under the **Instances**, **Persistent Disk Type**, and **VM Type** fields, choose **Automatic** from the dropdown to allocate the recommended resources for the job. If the **Persistent Disk Type** field reads **None**, the job does not require persistent disk space.

Note: Ops Manager requires a Director VM with at least 8 GB memory.

Note: If you set a field to **Automatic** and the recommended resource allocation changes in a future version, Ops Manager automatically uses the updated recommended allocation.

Note: If you install PAS for Windows, provision your **Master Compilation Job** with at least 100 GB of disk space.

3. Click **Save**.

Step 11: (Optional) Add Custom VM Extensions

Use the Ops Manager API to add custom properties to your VMs such as associated security groups and load balancers. For more information, see [Managing Custom VM Extensions](#).

Step 12: Complete the BOSH Director Installation

1. Click the **Installation Dashboard** link to return to the Installation Dashboard.
2. Click **Review Pending Changes**. Select the product that you intend to deploy and review the changes. For more information, see [Reviewing Pending Product Changes](#).

Next Steps

After you complete this procedure, follow the instructions for the runtime that you intend to install.

- To install PAS, see [Configuring PAS](#).
- To install PKS, see [Installing PKS on vSphere](#).

For more information about PCF runtimes, see [Installing Runtimes](#).

Deploying PAS with NSX-T Networking

This topic describes how to install Pivotal Application Service (PAS) on vSphere with NSX-T internal networking, using the VMware NSX-T Container Plug-in for PCF.

Overview

PAS uses a Container Network Interface (CNI) plugin to support secure and direct internal communication between containers. This plugin can be:

- The internal Silk plugin that comes packaged with PAS, or
- On vSphere, [VMware NSX-T Container Plug-in for PCF](#), which installs as an Ops Manager tile.

Prerequisites

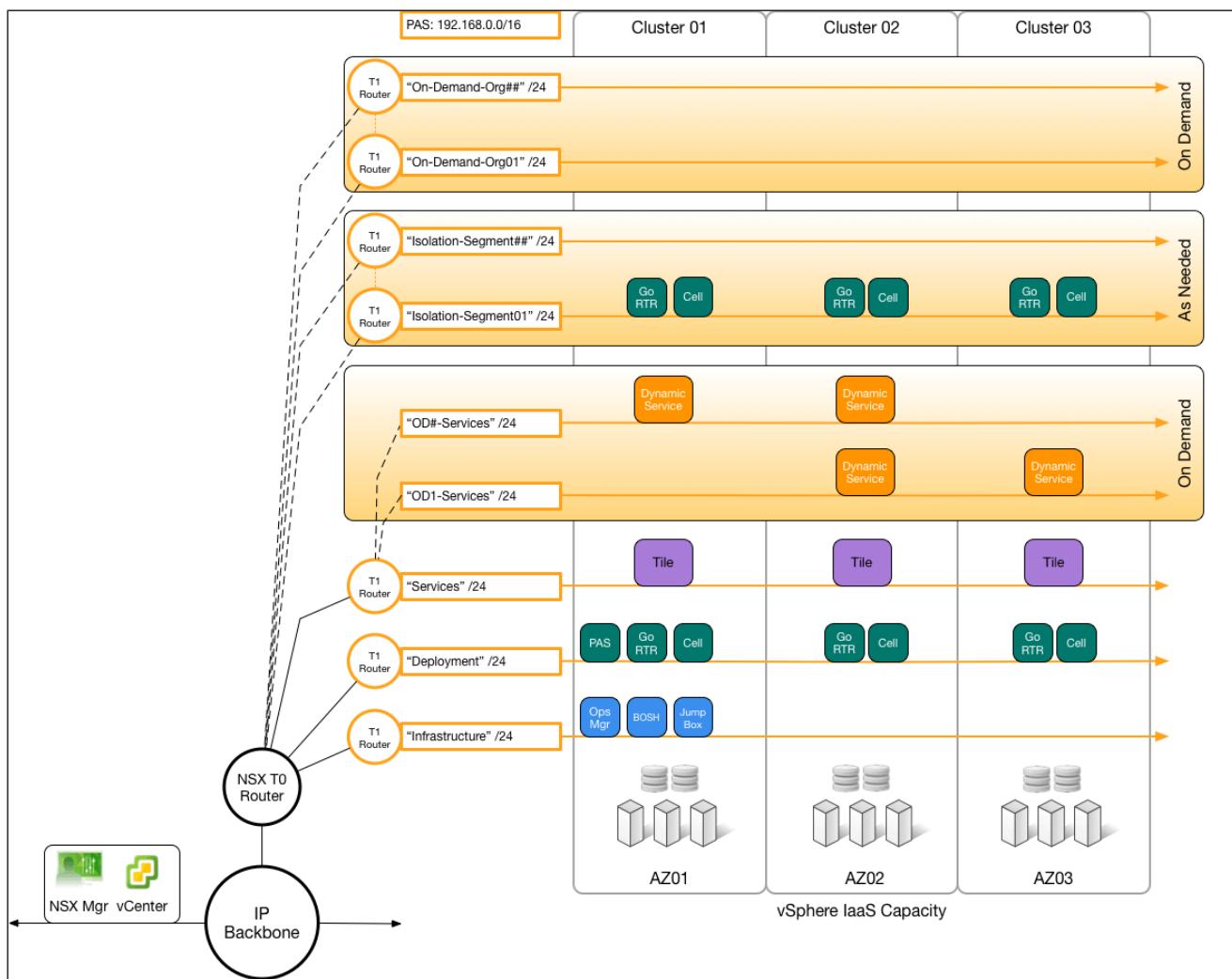
Before deploying PAS with NSX-T networking, you must have the following:

- An NSX-T environment with NSX-T components installed and configured. The NSX-T version must support the versions of NCP and PAS you intend to use. See the [Support Matrix](#) for relevant supported version combinations and see the [VMware Documentation](#) for NSX-T installation instructions.
- BOSH and Ops Manager installed and configured on vSphere. See [Deploying Ops Manager on vSphere](#) and [Configuring BOSH Director on vSphere](#).
- The [VMware NSX-T Container Plug-in for PCF tile](#) downloaded from Pivotal Network and imported to the Ops Manager [Installation Dashboard](#). See [Adding and Importing Products](#) for how to download and import Pivotal products to the [Installation Dashboard](#).
- The [PAS tile](#) downloaded from Pivotal Network and imported to the Ops Manager [Installation Dashboard](#). The PAS tile must be in one of the following two states:
 - Not deployed yet; you have not yet clicked **Review Pending Changes**, then **Apply Changes** on this version of PAS.
 - Deployed previously, with the **Networking** pane > **Container Network Interface Plugin** set to **External**.

 **Note:** Deploying PAS with its Container Network Interface (CNI) set to **Silk** configures Diego cells to use an internally-managed container network. Subsequently switching the CNI interface to **External** NSX-T leads to errors.

Architecture

The following diagram shows how to deploy an NSX-T machine to run PAS across multiple vSphere hardware clusters. NSX-T runs a Tier-0 (T0) router and multiple Tier-1 (T1) routers, each connecting to a network within Pivotal Cloud Foundry. Each vSphere hardware column cluster corresponds to an Availability Zone in PCF:



When a developer pushes an app to a new Org for the first time, the NSX-T plugin triggers NSX-T to create a new T1 router and allocate an address range for the Org, on demand.

Install and Configure PAS and NSX-T

Installing NSX-T to run with PAS requires the following actions, which are described below:

1. [Set up NSX-T to Integrate with PAS](#)
2. [Enable NSX-T Mode in the BOSH Director](#)
3. [Configure PAS for External Container Networking](#)
4. [Install and Configure the NSX-T Tile](#)

Set up NSX-T to Integrate with PAS

To set up NSX-T to integrate with PAS, complete the procedures in the following sections:

- [Set up Logical Switches](#)
- [Set up Routers](#)
- [Set up Load Balancer](#)

Set up Logical Switches

To set up logical switches, do the following:

1. In vSphere, create logical network switches to correspond to the networks that PCF uses.
 - a. Log into the **NSX-T Manager Dashboard**.
 - b. Open the **Advanced Networking & Security** tab at the top.
 - c. Open the **Switching** tab on the left.
 - d. For each of these networks...
 - **Infrastructure** (BOSH and Ops Manager, defined in BOSH Director tile > **Assign AZs and Networks** pane)
 - **Deployment** (PAS, defined in PAS tile > **Assign AZs and Networks** pane)
 - **Services and Dynamic Services** (marketplace services and on-demand services, also defined in the PAS tile)
 - **Isolation Segment** (optional, defined in the Isolation Segment tile > **Assign AZs and Networks** pane) ...do the following:
 1. Click **+ADD**.
 2. Enter a name for the logical switch (e.g. `PAS-Infrastructure`, `PAS-Deployment`).
 3. Click **ADD**.

Add New Logical Switch

(?)
X

General
Switching Profiles

Name *
PAS-Infrastructure

Description

Transport Zone *
overlay-tz

Uplink Teaming Policy Name *
[Use Default]

Admin Status
 Up

Replication Mode

 Hierarchical Two-Tier replication
 Head replication

VLAN

Only VLAN Trunk Spec is allowed (eg: 1, 5, 10-12, 31-35).

CANCEL
ADD

Set up Routers

To set up routers, do the following:

1. Create T0 network address translation (NAT) rules to communicate with Ops Manager

- a. Open the **Advanced Networking & Security** tab at the top
- b. Open the **Routers** tab on the left
- c. Select your T0 router
- d. Select the **Services** drop-down menu and choose **NAT**.
- e. Add a rule for destination NAT (DNAT) with:
 - The externally-specified destination IP address of incoming requests. If your Ops Manager has a DNS entry (e.g. `opsmgr.example.com`), this would be its IP address.
 - The internal network address of the Ops Manager

The screenshot shows a modal dialog titled "New NAT Rule". The form contains the following fields:

- Priority:** 1024 (with a dropdown arrow)
- Action ***: DNAT (with a dropdown arrow)
- Protocol:** Any Protocol
 Specific Protocol
- Source IP:** (empty field)
- Destination IP ***: 10.195.68.16
- Translated IP ***: 192.168.1.10
- Translated Ports:** (empty field)
- Applied To:** (empty field with close and dropdown arrows)
- Status:** Enabled
- Logging:** Disabled
- Firewall Bypass:** Enabled

At the bottom right are two buttons: **CANCEL** and **ADD**.

- f. Add the corresponding source NAT (SNAT) rule with:
 - The externally-specified destination IP address
 - The internal network address of the Ops Manager

New NAT Rule

X

Priority	1024	^ ▼
Action *	SNAT	▼
Protocol	<input checked="" type="radio"/> Any Protocol <input type="radio"/> Specific Protocol	
Source IP	192.168.1.10	
Destination IP		
Translated IP *	10.195.68.16	
Applied To	X ▼	
Status	<input checked="" type="checkbox"/> Enabled	
Logging	<input type="checkbox"/> Disabled	
Firewall Bypass	<input checked="" type="checkbox"/> Enabled	

CANCEL
ADD

g. Add a rule for source NAT (SNAT) for the infrastructure and deployment networks:

- The externally-specified destination IP address
- The internal network address in CIDR notation

New NAT Rule

Priority: 1024

Action*: SNAT

Protocol: Any Protocol Specific Protocol

Source IP: 192.168.0.0/16

Destination IP:

Translated IP*: 10.195.68.251

Applied To:

Status: Enabled

Logging: Disabled

Firewall Bypass: Enabled

CANCEL **ADD**

2. Create T1 Routers for PAS, to connect from the T0 router. For each PCF network, Infrastructure, Deployment, and so on, create a T1 router as follows:
3. In the NSX-T Manager UI, navigate to **Advanced Networking & Security > Routing > Routers**
 - a. Click **+ADD > Tier-1 Router**.
 - b. Configure the router. Include the Edge Cluster and Edge Cluster Members; they are required to enable the Load Balancer. The Infrastructure network router configuration might look like:

New Tier-1 Router

[?](#) [X](#)

Tier-1 Router Advanced

Name *	T1-Router-PAS-Infrastructure	
Description	<input type="text"/>	
Tier-0 Router	TO-Router X ▼	
Edge Cluster	edge-cluster-1 X ▼	
StandBy Relocation	<input checked="" type="checkbox"/> Disable	
Failover Mode	<input type="radio"/> Preemptive	<input checked="" type="radio"/> Non-Preemptive
Edge Cluster Members i	<input type="text" value="edge-transp-node-1"/> X <input type="text" value="edge-transp-node-2"/> X X ▼	

[CANCEL](#) [ADD](#)

4. Create T1 router downlink ports for PAS. For each T1 router you created, add a **New Router Port** as follows, to to allow traffic in and out:
 - a. In the NSX-T Manager UI, select the T1 router.
 - b. In **Configuration > Router Ports**, click **+ADD** to add a new router port.
 - c. For **Logical Switch**, enter the name of the logical switch you defined for the network in **Add New Logical Switch**, above.
 - d. For **IP Address**, use the first IP of the appropriate subnet. In this example, we have set aside 192.168.1.0/24 for Infrastructure (Operations Manager and BOSH Director), and 192.168.2.0/24 for the Deployment, so we used 192.168.1.1 and 192.168.2.1 respectively.

New Router Port

Name*

Description

Type

URPF Mode Strict None

Logical Switch x v OR Create a New Switch

Logical Switch Port Attach to new switch port Attach to existing switch port
 Switch Port Name

Subnets

[+ ADD](#) [DELETE](#)

IP Address*	Prefix Length*
<input checked="" type="checkbox"/> 192.168.1.1	24

Relay Service x v

[CANCEL](#) [ADD](#)

5. Advertise the routes of the T1 routers to the T0 router, so the T0 router can correctly route incoming requests based on their destination IP address:

- a. Select your T1 Router and navigate to **Routing > Route Advertisement**.
- b. Under **Edit Route Advertisement Configuration**, enable route advertisement by setting **Status** to **Enabled**.
- c. Set **Advertise All Connected Routes** to **Yes**.
- d. Set **Advertise All LB VIP Routes** to **Yes** (necessary if Load Balancing service is configured).

Edit Route Advertisement Configuration

X

Status	<input checked="" type="checkbox"/> Enabled
Advertise All Connected Routes	<input checked="" type="checkbox"/> Yes
Advertise All NAT Routes	<input type="checkbox"/> No
Advertise All Static Routes	<input type="checkbox"/> No
Advertise All LB VIP Routes	<input checked="" type="checkbox"/> Yes
Advertise All LB SNAT IP Routes	<input type="checkbox"/> No
Advertise All DNS Forwarder Routes	<input type="checkbox"/> No

CANCEL
SAVE

6. Allocate an IP Block for PAS Orgs.

- a. From the NSX-T Manager, navigate to **Advanced Networking & Security > Networking > IPAM** and click **+ADD**.
- b. Enter a name (e.g. `PAS-container-ip-block`). Later you will enter this IP Block name on the **VMware NSX-T** tile in the **NCP** section under **IP Blocks of Container Networks**.
- c. Enter a description, such as `Subnets are allocated from this pool to each newly-created Org`
- d. Enter a CIDR to allocate an address block large enough to accommodate all PAS apps. A `/14` CIDR is large enough for ~1,000 Orgs with ~250 apps each. If you're planning such a large foundation, pay heed to the [VMware NSX-T PAS limits](#).

New IP Block

[?](#) [X](#)

Name *	PAS-container-ip-block
Description	Subnets are allocated from this pool to each newly-created Org
CIDR *	10.12.0.0/14

[CANCEL](#) [ADD](#)

7. Create an external SNAT IP pool.

- Navigate to **Advanced Networking & Security > Inventory > Groups > IP Pools** and click **+ADD**.
- Enter a name (e.g. `external-ip-pool`) and a description (e.g. `IP pool that provides 1 public IP for each Cloud Foundry Org`). Later you will enter this pool name on the VMware NSX-T tile in the NCP section under **IP Pools used to provide External (NAT) IP Addresses to Org Networks**.

Add New IP Pool

[?](#) [X](#)

Name *	external-ip-pool
Description	IP pool that provides 1 public IP for each Cloud Foundry Org

Subnets

[+ ADD](#) [DELETE](#)

<input checked="" type="checkbox"/> IP Ranges *	Gateway	CIDR *	DNS Servers	DNS Suffix
<input checked="" type="checkbox"/> 10.195.68.128 - 10.195.68.250	10.195.68.1	10.195.68.0/24		

[CANCEL](#) [ADD](#)

Set up Load Balancer

To set up a load balancer, do the following:

1. Create Active Health Monitors (health checks) for use by the Virtual Server later on:

- a. In the NSX-T Manager UI, navigate to **Advanced Networking & Security > Networking > Load Balancing > Monitors > Active Health Monitors**
- b. Create the health monitor for web load balancing:

c. Click **+ADD**

d. Fill out **Monitor Properties**:

- **Name:** pas-web-monitor
- **Health Check Protocol:** LbHttpMonitor
- **Monitoring Port:** 8080

e. Click **Next**

f. Fill out **Health Check Parameters**:

- **HTTP Method:** GET
- **HTTP Request URL:** /health
- **HTTP Response Code:** 200

g. Click **Finish**

h. Create the health monitor for TCP load balancing:

i. Click **+ADD**

j. Fill out **Monitor Properties**:

- **Name:** pas-tcp-monitor
- **Health Check Protocol:** LbHttpMonitor
- **Monitoring Port:** 80

k. Click **Next**

l. Fill out **Health Check Parameters**:

- **HTTP Method:** GET
- **HTTP Request URL:** /health
- **HTTP Response Code:** 200

m. Create the health monitor for SSH load balancing:

n. Click **+ADD**

o. Fill out **Monitor Properties**:

- **Name:** pas-ssh-monitor
- **Health Check Protocol:** LbTcpMonitor
- **Monitoring Port:** 2222

p. Click **Next > Finish**

2. Create Server Pools (collections of VMs which handle traffic) for use by the Virtual Server later on:

a. In the NSX-T Manager UI, navigate to **Advanced Networking & Security > Networking > Load Balancing > Server Pools**

b. Create the server pool for web load balancing:

c. Click **+ADD** to add a new pool

d. Fill out **General Properties**:

- **Name:** pas-web-pool

e. Click **Next**

f. Fill out **SNAT Translation**:

- **Translation Mode:** Auto Map

g. Click **Next**

h. Fill out **Pool Members**:

- **Membership Type:** Static

i. Click **Next**

j. Fill out **Health Monitors**:

- **Active Health Monitor:** pas-web-monitor

- k. Click **Finish**
- l. Create the server pool for TCP load balancing:
- m. Click **+ADD** to add new pool
- n. Fill out **General Properties**:

- **Name:** `pas-tcp-pool`

- o. Click **Next**
- p. Fill out **SNAT Translation**:

- **Translation Mode:** `Transparent`

- q. Click **Next**
- r. Fill out **Pool Members**:

- **Membership Type:** `Static`

- s. Click **Next**
- t. Fill out **Health Monitors**:

- **Active Health Monitor:** `pas-tcp-monitor`

- u. Click **Finish**
- v. Create the server pool for ssh load balancing:
- w. Click **+ADD** to add new pool
- x. Fill out **General Properties**:

- **Name:** `pas-ssh-pool`

- y. Click **Next**
- z. Fill out **SNAT Translation**:

- **Translation Mode:** `Transparent`

- aa. Click **Next**
- ab. Fill out **Pool Members**:

- **Membership Type:** `Static`

- ac. Click **Next**
- ad. Fill out **Health Monitors**:

- **Active Health Monitor:** `pas-ssh-monitor`

- ae. Click **Finish**

3. Create Virtual Servers:

- a. In the NSX-T Manager UI, navigate to **Advanced Networking & Security > Networking > Load Balancing > Virtual Servers**
- b. Create the Virtual Server which forwards unencrypted web (HTTP) traffic to the Foundation:

Note: Foundations requiring end-to-end encryption should not enable the virtual server on port 80, or, if enabled, should configure it to redirect traffic to the encrypted port (443)

- c. Click **+ADD**
- d. Fill out **General Properties**:

- **Name:** `pas-web-vs`
 - **Application Type:** `Layer 4 (TCP)`
 - **Application Profile:** `nsx-default-lb-fast-tcp-profile`

- e. Click **Next**
- f. Fill out **Virtual Server Identifiers**:

- **IP Address:** use the address of the DNS record of `*.system.YOUR-SYSTEM-DOMAIN.com`
 - **Port:** `80,443`

- g. Fill out **Server Pool and Rules**:

- **Default Server Pool:** `pas-web-pool`

- h. Click **Next** several times, then **Finish**

- i. Create the Virtual Server which forwards traffic to apps with custom ports to the Foundation:
- j. Click **+ADD** to add a new Virtual Server
- k. Fill out **General Properties**:

- **Name:** `pas-tcp-vs`
- **Application Type:** `Layer 4 (TCP)`
- **Application Profile:** `nsx-default-lb-fast-tcp-profile`

- l. Click Next
- m. Fill out **Virtual Server Identifiers**:

- **IP Address:** use the address of the DNS record of `tcp.apps.YOUR-SYSTEM-DOMAIN.com`
- **Port:** use the same ports as configured in the **PAS Tile > Networking > TCP Routing Ports**, e.g. `1024-1123,5900`

- n. Click Next
- o. Fill out **Server Pool and Rules**:

- **Default Server Pool:** `pas-tcp-pool`

- p. Click Next > Finish
- q. Create the Virtual Server which forwards ssh traffic to the Foundation:
- r. Click **+ADD** to add a new Virtual Server
- s. Fill out **General Properties**:

- **Name:** `pas-ssh-vs`
- **Application Type:** `Layer 4 (TCP)`
- **Application Profile:** `nsx-default-lb-fast-tcp-profile`

- t. Click Next
- u. Fill out **Virtual Server Identifiers**:

- **IP Address:** use the address of the DNS record of `ssh.system.YOUR-SYSTEM-DOMAIN.com`
- **Port:** `2222`

- v. Click Next
- w. Fill out **Server Pool and Rules**:

- **Default Server Pool:** `pas-ssh-pool`

- x. Click Next > Finish

4. Create the Load Balancer:

- a. In the NSX-T Manager UI, navigate to **Advanced Networking & Security > Networking > Load Balancing > Load Balancers**
 - b. Click **+ADD**
 - c. Fill out the fields:
- **Name:** `pas-lb`
 - **Load Balancer Size:** Choose `Small` unless you have a larger Foundation
- d. Click OK
 - e. Select `pas-lb`
 - f. Click **Actions > Attach to a Virtual Server**, and then select `pas-web-vs`. Repeat this procedure for the Virtual Servers `pas-tcp-vs` and `pas-ssh-vs`.
 - g. Click **Action > Attach to a Logical Router**, and then select `T1-Router-PAS-Deployment`

Enable NSX-T Mode in the BOSH Director

To enable NSX-T mode in the BOSH Director, do the following:

1. From the Ops Manager **Installation Dashboard**, open the **BOSH Director** tile.
2. In the **vCenter Configs** pane, click the pencil icon for the vCenter Config you want to edit. Then select **NSX Networking**.

Standard vCenter Networking

NSX Networking

NSX Mode*
 NSX-V
 NSX-T

NSX Address*

NSX Username*
 User to connect to the NSX manager

NSX Password*

[Change](#)

NSX CA Cert

```
-----BEGIN CERTIFICATE-----
MIIDmTCCAoGgAwIBAgIEShW2VDANBgkqhkiG9w0BAQsFADB9MQswCQYDVQQGEwJ
V
UzELMAkGA1UECBMCQ0ExEjAQBgNVBAcTCVBhbG8gQWx0bzEUMBIGA1UEChMLVk1
3
VVJIElVavDnAKRanVRActAOCTWDE5MCcGA1UEAxMgdnbNahWfUVMdciOQDw
-----
```




3. Configure NSX networking by entering the following:

- **NSX Mode:** Select NSX-T.
- **NSX Address:** Enter the address of the NSX-T Manager.
- **NSX Username:** Enter the username to connect to the NSX-T Manager.
- **NSX Password:** The password for the username specified above.
- **NSX CA Cert:** A CA certificate in PEM format that authenticates to the NSX-T server. If the NSX-T Manager generated a self-signed certificate, you can retrieve the CA certificate using OpenSSL with the command `openssl s_client -host NSX-ADDRESS -port 443 -prexit -showcerts`.

Configure PAS for External Container Networking

To configure PAS for external container networking, do the following:

1. If you have not already done so, download the PAS tile from [Pivotal Network](#) and import it to the [Installation Dashboard](#). See [Adding and Importing Products](#) for directions.

The screenshot shows the PCF Ops Manager interface with the title "PCF Ops Manager". On the left, there's a sidebar with a "Import a Product" button and a section for "VMware Harbor Registry" containing an item "1.4.1-build.1". Below this is a link to "Download PCF compatible products at Pivotal Network" and a "Delete All Unused Products" button. The main area is titled "Installation Dashboard" and displays four tiles representing installed products:

- BOSH Director for vSphere**: Version v2.0.6
- Pivotal Application Service**: Version v2.0.6
- Pivotal Container Service**: Version v1.0.0-build.3
- VMware NSX-T**: Version v2.1.0.0.0.7371408

- Configure PAS, following the directions in [Configuring PAS](#). When you configure **Networking**, select **External** under **Container Networking Interface Plugin**.

Container Network Interface Plugin*

Silk

External

- Update the server pool membership for the NSX-T Load Balancers by following [Updating NSX-T Load Balancer Server Pool Membership](#). Make sure to make a VM extension for each of the three server pools: `pas-web-pool`, `pas-tcp-pool`, and `pas-ssh-pool`.

Install and Configure the NSX-T tile

To install and configure the NSX-T tile, do the following:

- If you have not already done so, download the VMware NSX-T tile version 2.3 or later from [Pivotal Network](#) and import it to the **Installation Dashboard**. See [Adding and Importing Products](#) for directions.

The screenshot shows the PCF Ops Manager interface with the title "PCF Ops Manager". On the left, there's a sidebar with a "Import a Product" button and a section for "VMware Harbor Registry" containing an item "1.4.1-build.1". Below this is a link to "Download PCF compatible products at Pivotal Network" and a "Delete All Unused Products" button. The main area is titled "Installation Dashboard" and displays four tiles representing installed products:

- BOSH Director for vSphere**: Version v2.0.6
- Pivotal Application Service**: Version v2.0.6
- Pivotal Container Service**: Version v1.0.0-build.3
- VMware NSX-T**: Version v2.1.0.0.0.7371408

2. Click the **VMware NSX-T** tile to open its **Settings** tab, and configure the **NSX Manager** pane as follows:

- **NSX Manager Address:** NSX-T Manager host address or IP address
- **Use Client Certificates or Username/Password:** Select **Basic Authentication with Username and Password** and enter **NSX Manager Admin Username** and **Admin Password** credentials in the fields underneath.
- **NSX Manager CA Cert:** Obtain this certificate from NSX-T Manager as follows:
 1. `ssh` into NSX-T Manager using the admin account that you created when you deployed NSX-T Manager.
 2. From the NSX-T Manager command line, run `get certificate api` to retrieve the certificate.

The screenshot shows the VMware NSX-T Installation Dashboard. The top navigation bar has a back arrow labeled "Installation Dashboard" and the title "VMware NSX-T". Below the title is a horizontal menu with four tabs: "Settings" (which is selected and highlighted in blue), "Status", "Credentials", and "Logs". On the left side, there is a sidebar with three items: "NSX Manager" (selected and highlighted in blue), "NCP", and "NSX Node Agent", each preceded by a checkmark icon. The main content area is titled "NSX Manager Configuration". It contains the following fields:

- "NSX Manager Address *": A text input field containing the value "nsx-manager.haas-108.pez.pivotal.io".
- "Use Client Certificates or Username/Password for NSX Manager Authentication *": A radio button group where the "Basic Authentication with Username and Password" option is selected (indicated by a blue dot).
- "NSX Manager Admin Username *": A text input field containing the value "admin". To the right of the field is a tooltip: "Admin username of NSX Manager".
- "NSX Manager Admin Password *": A password input field showing masked text "*****" and a "Change" link below it.
- "NSX Manager CA Cert": A text input field containing the text "-----BEGIN CERTIFICATE-----".

3. Open and configure the **NCP** (NSX-T Container Plugin) pane as follows:

- **PAS foundation Name:** If unsure, use `PAS`. If multiple foundations co-exist on the same NSX-T Manager, choose a unique string, e.g. `PAS-beta`. NCP creates artifacts (e.g. T1 routers) and prefixes their names with this string for easy identification.
- **Overlay Transport Zone:** A uniquely identifying string for the **Transport Zone** that you chose when you [created logical switches](#) for each network. This can be the name of the transport zone if no other zones in NSX-T share the same name, or else the UUID for the transport zone.
- **Tier-0 Router:** A uniquely identifying string for the T0 router. This can be the tag string that you gave the router in NSX-T Manager if no other T0 routers in NSX-T share the same name, or else the UUID for the router.
- **Subnet Prefix of Container Networks:** Subnet mask to set the address range size for apps in a single org. Defaults to `24`. This number must be higher than the mask for all PAS orgs in the NSX-T Manager [New IP Block](#) pane, to define a each Org's fraction of the total PAS address space.
- **Enable SNAT for Container Network:** Enable this checkbox.

NCP Configuration

PAS Foundation Name *

Overlay Transport Zone *

This string uniquely identifies the NSX overlay transport zone for logical switches for container networking. You can specify the name or the UUID of the transport zone to be used.

Tier-0 Router *

IP Blocks of Container Networks

Subnet Prefix of Container Networks *

Enable SNAT for Container Networks

IP Pools used to provide External (NAT) IP Addresses to Org Networks

Top Firewall Section Marker

Bottom Firewall Section Marker

Log Dropped Application Traffic

Enable Debug Level for NCP Logging

Save

- In the **NSX Node Agent** pane, enable the **Enable Debug Level of Logging for NSX Node Agent** checkbox.

NSX Node Agent Configuration

Enable Debug Level of Logging for NSX Node Agent

Save

- Click **Save** and return to the **Installation Dashboard**.
- After you have configured both the **PAS** tile and the **VMware NSX-T** tile, click **Review Pending Changes**, then **Apply Changes** to deploy PAS with NSX-T networking.

Automation

For more information about automation, see the following resources:

- [Concourse Pipelines: Configure NSX-T for PAS](#): This sample Concourse pipeline provides jobs setup switches, routers, an IP block, and an IP pool to be used by PAS.

- [PyNSXT](#): This is a Python library that can be used as a CLI to run commands against a VMware NSX-T installation. It exposes operations for working with logical switches, logical routers, and pools. It uses version 2.1 of NSX-T for the swagger client.

vSphere Virtual Disk Types

When you create a virtual machine in VMware vSphere, vSphere creates a new virtual hard drive for that virtual machine. The virtual hard drive is contained in a virtual machine disk (VMDK). The disk format you choose for the new virtual hard drive can have a significant impact on performance.

You can choose one of three formats when creating a virtual hard drive:

- Thin Provisioned
- Thick Provisioned Lazy Zeroed
- Thick Provisioned Eager Zeroed

Thin Provisioned

Advantages:

- Fastest to provision
- Allows disk space to be over-committed to VMs

Disadvantages:

- Slowest performance due to metadata allocation overhead and additional overhead during initial write operations
- Over-commitment of storage can lead to application disruption or downtime if resources are actually used
- Does not support clustering features

When vSphere creates a thin provisioned disk, it only writes a small amount of metadata to the datastore. It does not allocate or zero out any disk space. At write time, vSphere first updates the allocation metadata for the VMDK, then zeros out the block or blocks, then finally writes the data. Because of this overhead, thin provisioned VMDKs have the lowest performance of the three disk formats.

Thin provisioning allows you to overcommit disk spaces to VMs on a datastore. For example, you could put 10 VMs, each with a 50 GB VMDK attached to it, on a single 100 GB datastore, as long as the sum total of all data written by the VMs never exceeded 100 GB. Thin provisioning allows administrators to use space on datastores that would otherwise be unavailable if using thick provisioning, possibly reducing costs and administrative overhead.

Thick Provisioned Lazy Zeroed

Advantages:

- Faster to provision than Thick Provisioned Eager Zeroed
- Better performance than Thin Provisioned

Disadvantages:

- Slightly slower to provision than Thin Provisioned
- Slower performance than Thick Provisioned Eager Zero
- Does not support clustering features

When vSphere creates a thick provisioned lazy zeroed disk, it allocates the maximum size of the disk to the VMDK, but does nothing else. At the initial access to each block, vSphere first zeros out the block, then writes the data. Performance of a thick provisioned lazy zeroed disk is not as good as a thick provisioned eager zero disk because of this added overhead.

Thick Provisioned Eager Zeroed

Advantages:

- Best performance
- Overwriting allocated disk space with zeros reduces possible security risks
- Supports clustering features such as Microsoft Cluster Server (MSCS) and VMware Fault Tolerance

Disadvantages:

- Longest time to provision

When vSphere creates a thick provisioned eager zeroed disk, it allocates the maximum size of the disk to the VMDK, then zeros out all of that space.

Example: If you create an 80 GB thick provisioned eager zeroed VMDK, vSphere allocates 80 GB and writes 80 GB of zeros.

By overwriting all data in the allocated space with zeros, thick provisioned eager zeroed eliminates the possibility of reading any residual data from the disk, thereby reducing possible security risks.

Thick provisioned eager zeroed VMDKs have the best performance. When a write operation occurs to a thick provisioned eager zeroed disk, vSphere writes to the disk, with none of the additional overhead required by thin provisioned or thick provisioned lazy zeroed formats.

Using the Cisco Nexus 1000v Switch with Ops Manager

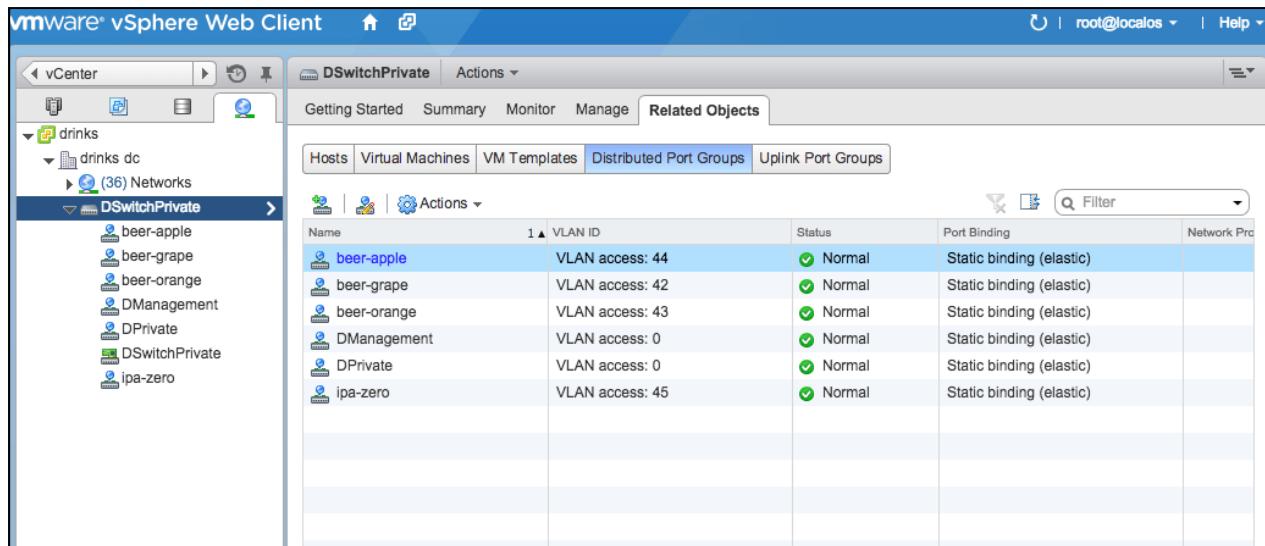
Refer to the procedure in this topic to use Ops Manager with the Cisco Nexus 1000v Switch. First, configure Ops Manager through Step 4 in [Configuring BOSH Director on vSphere](#). Then configure your network according to the following steps.

1. From your [Pivotal Cloud Foundry](#) (PCF) Ops Manager **Installation Dashboard**, click the **BOSH Director** tile.
2. Select **Create Networks**.
3. Click the network name to configure the network settings. This is **default** if you have not changed the name.

The screenshot shows the 'Create Networks' configuration page for the 'default' network. On the left, a sidebar lists configuration sections: vCenter Config, Director Config, Create Availability Zones, **Create Networks** (selected), Assign AZs and Networks, Security, and Resource Config. The main area is titled 'Create Networks' and contains a warning about keeping IP settings. Under 'Verification Settings', 'Enable ICMP checks' is checked. The 'Networks' section shows the 'default' network selected. It includes fields for 'Name*' (set to 'default'), 'Subnets' (with 'vSphere Network Name*' and 'CIDR*' fields), 'Reserved IP Ranges', 'DNS*', 'Gateway*', and 'Availability Zones*' (with a checkbox for 'first-az'). A 'Save' button is at the bottom.

4. Find the folder name and port group name for the switch, as you configured them in vCenter. For the example vSphere environment pictured below, a user might want to use the switch configured on the `beer-apple` port group, which is in the `drinks dc` folder.

Note: Folder names can be strings of multiple words separated by spaces, or strings of words separated by other punctuation, including dashes.



- In the **vSphere Network Name** field, instead of entering your network name, enter the folder name and port group name for the switch, as you configured them in vCenter. For the example vSphere environment pictured above, you would enter `drinks dc/beer-apple` to use the switch configured on the `beer-apple` port group.

[Installation Dashboard](#)

Ops Manager Director

- Settings
- Status
- Credentials

vCenter Config

Director Config

Create Availability Zones

Create Networks

Assign AZs and Networks

Security

Resource Config

Create Networks

Warning: Pivotal recommends keeping the IP settings throughout the life of your installation. Ops Manager may prevent you from changing them in the future. Contact Pivotal support for help completing such a change.

Verification Settings

Enable ICMP checks

Networks

One or many IP ranges upon which your products will be deployed

Add Network

YOUR_NETWORK_NAME

Name*

Subnets

Add Subnet

vSphere Network Name*

The name of the network as it appears in vCenter

CIDR*

Reserved IP Ranges

DNS*

Gateway*

Availability Zones*

Save

6. Click **Save**.

7. Return to [Configuring BOSH Director on vSphere](#) to complete the Ops Manager installation.

Using BOSH Resurrector and vSphere HA

This topic describes the BOSH Resurrector and vSphere High Availability (HA), as well as how to enable them for your Pivotal Cloud Foundry (PCF) deployment.

Pivotal recommends BOSH Resurrector, vSphere HA, and a highly-available storage solution to protect the VMs in your deployment.

BOSH Resurrector

The following sections describe the BOSH Resurrector and how to enable it.

About BOSH Resurrector

The BOSH Resurrector increases Pivotal Application Service (PAS) availability in the following ways:

- Reacts to hardware failure and network disruptions by recreating virtual machines on active, stable hosts
- Detects operating system failures by continuously monitoring virtual machines and recreating them as required
- Continuously monitors the BOSH Agent running on each virtual machine and recreates the VMs as required

The BOSH Resurrector continuously monitors the status of all virtual machines in an PAS deployment. The Resurrector also monitors the BOSH Agent on each VM. If either the VM or the BOSH Agent fail, the Resurrector recreates the virtual machine on another active host.

BOSH Resurrector Limitations

The following limitations apply to using the BOSH Resurrector:

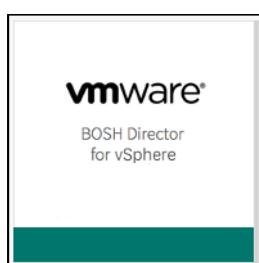
- The Resurrector does not monitor or protect the Ops Manager VM or the BOSH Director VM.
- The Resurrector might not be able to resolve issues caused by the loss of an entire host.
- The Resurrector does not monitor or protect data storage.

To mitigate these limitations, Pivotal recommends vSphere HA. For more information, see [vSphere HA](#) below.

Enable BOSH Resurrector

To enable the BOSH Resurrector:

1. Log into the Ops Manager web interface.
2. On the Product Dashboard, select **BOSH Director**.



3. In the left navigation menu, select **Director Config**.

The screenshot shows the 'Ops Manager Director' interface under the 'Settings' tab. On the left, a sidebar lists several configuration steps: vCenter Config, Director Config (which is selected and highlighted in grey), Create Availability Zones, Create Networks, Assign AZs and Networks, Security, and Resource Config. The main panel is titled 'Director Config' and contains fields for 'NTP Servers (comma delimited)*' (with 'time1.sf.cf-app.com' entered) and 'Metrics IP Address' (an empty input field). Below these are two checkboxes: 'Enable VM Resurrector Plugin' (checked) and 'Recreate all VMs' (unchecked). A note at the bottom states: 'This will force BOSH to recreate all VMs on the next deploy. Persistent disk will be preserved.'

4. Select **Enable VM Resurrector Plugin** and click **Save**.

vSphere HA

The following sections describe vSphere HA and how to enable it.

About vSphere HA

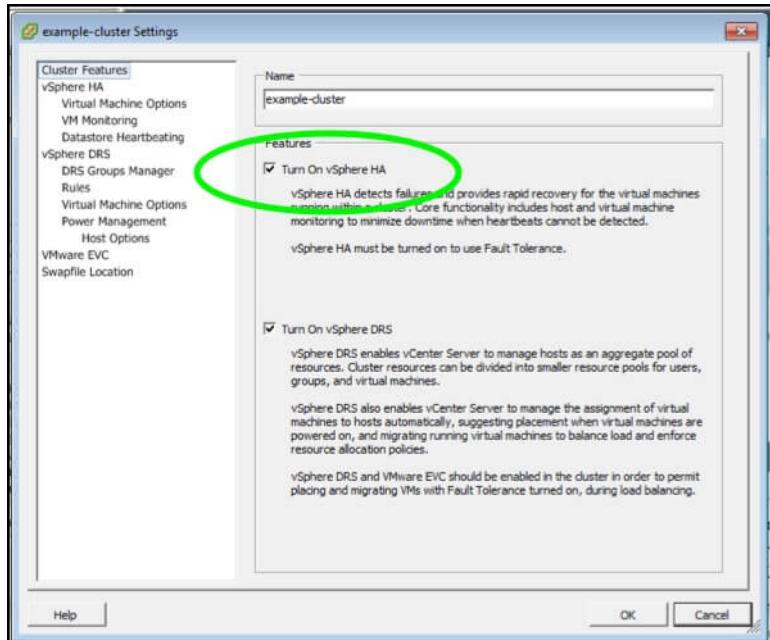
In the event of a host failure, vSphere HA restarts but does not recreate VMs on alternate ESXi hosts in the same cluster. A host failure could result from a sudden power off, crash due to hardware fault, or the vSphere host becoming disconnected from the network. BOSH Resurrector cannot recover from a host failure.

For more information, see [vSphere High Availability](#) in the BOSH documentation.

Enable vSphere HA

Follow the steps below to enable vSphere High Availability:

1. Launch the vSphere Management Console.
2. Right-click the cluster that contains the [Pivotal Cloud Foundry](#) deployment and select **Edit Settings**.
3. Select the **Turn on vSphere High Availability** checkbox.



4. Click **OK** to enable vSphere High Availability on the cluster.

Configuring Pivotal Cloud Foundry SSL Termination for vSphere Deployments

To use SSL termination in [Pivotal Cloud Foundry](#) (PCF), you must configure the Pivotal-deployed HAProxy load balancer or your own load balancer.

Pivotal recommends that you use HAProxy in lab and test environments only. Production environments should instead use a highly-available customer-provided load balancing solution.

Select an SSL termination method to determine the steps you must take to configure Pivotal Application Service (PAS).

Using the Pivotal HAProxy Load Balancer

PCF deploys with a single instance of HAProxy for use in lab and test environments. You can use this HAProxy instance for SSL termination and load balancing to the PCF Routers. HAProxy can generate a self-signed certificate if you do not want to obtain a signed certificate from a well-known certificate authority.

Note: Certificates generated in PAS are signed by the Operations Manager Certificate Authority. They are not technically self-signed, but they are referred to as “Self-Signed Certificates” in the Ops Manager GUI and throughout this documentation.

To use the HAProxy load balancer, you must create a wildcard A record in your DNS and configure three fields in the PAS product tile.

1. Create an A record in your DNS that points to the HAProxy IP address. The A record associates the **System Domain** and **Apps Domain** that you configure in the **Domains** section of the PAS tile with the HAProxy IP address.

For example, with `cf.example.com` as the main subdomain for your CF install and an HAProxy IP address `203.0.113.1`, you must create an A record in your DNS that serves `example.com` and points `*.cf` to `203.0.113.1`.

Name	Type	Data	Domain
<code>*.cf</code>	A	<code>203.0.113.1</code>	<code>example.com</code>

2. Use the Linux `host` command to test your DNS entry. The `host` command should return your HAProxy IP address.

Example:

```
$ host cf.example.com
cf.example.com has address 203.0.113.1
$ host anything.example.com
anything.cf.example.com has address 203.0.113.1
```

3. From the PCF Ops Manager Dashboard, click on the PAS tile.
4. Select **Networking**.
5. Leave the **Router IPs** field blank. HAProxy assigns the router IPs internally.
6. Enter the IP address for HAProxy in the **HAProxy IPs** field.
7. Provide your SSL certificate in the **SSL Termination Certificate and Private Key** field. See [Providing a Certificate for your SSL Termination Point](#) for details.

[Return to the Getting Started Guide](#)

Using Another Load Balancer

Production environments should use a highly-available customer-provided load balancing solution that does the following:

- Provides SSL termination with wildcard DNS location
- Provides load balancing to each of the PCF Router IPs
- Adds appropriate `x-forwarded-for` and `x-forwarded-proto` HTTP headers

You must register static IP addresses for PCF with your load balancer and configure three fields in the PAS product tile.

1. Register one or more static IP address for PCF with your load balancer.
2. Create an A record in your DNS that points to your load balancer IP address. The A record associates the **System Domain** and **Apps Domain** that you configure in the **Domains** section of the PAS tile with the IP address of your load balancer.

For example, with `cf.example.com` as the main subdomain for your CF install and a load balancer IP address `198.51.100.1`, you must create an A record in your DNS that serves `example.com` and points `*.cf` to `198.51.100.1`.

Name	Type	Data	Domain
<code>*.cf</code>	A	<code>198.51.100.1</code>	<code>example.com</code>

3. From the PCF Ops Manager Dashboard, click on the PAS tile.
4. Select **Networking**.
5. In the **Router IPs** field, enter the static IP address for PCF that you have registered with your load balancer.
6. Leave the **HAProxy IPs** field blank.
7. Provide your SSL certificate in the **SSL Termination Certificate and Private Key** field. See [Providing a Certificate for your SSL Termination Point](#) for details.

 **Note:** When adding or removing PCF routers, you must update your load balancing solution configuration with the appropriate IP addresses.

[Return to the Installing Pivotal Cloud Foundry Guide](#) 

Availability Zones in vSphere

This topic describes availability zones (AZs) in vSphere and an example of how an operator might scale an app across AZs.

Overview

Pivotal defines an AZ as an operator-assigned, functionally-independent segment of network infrastructure. In cases of partial infrastructure failure, a tile distributes and balances all instances of running apps across remaining AZs. Strategic use of Availability Zones contributes to the fault tolerance and high availability of a tile deployment.

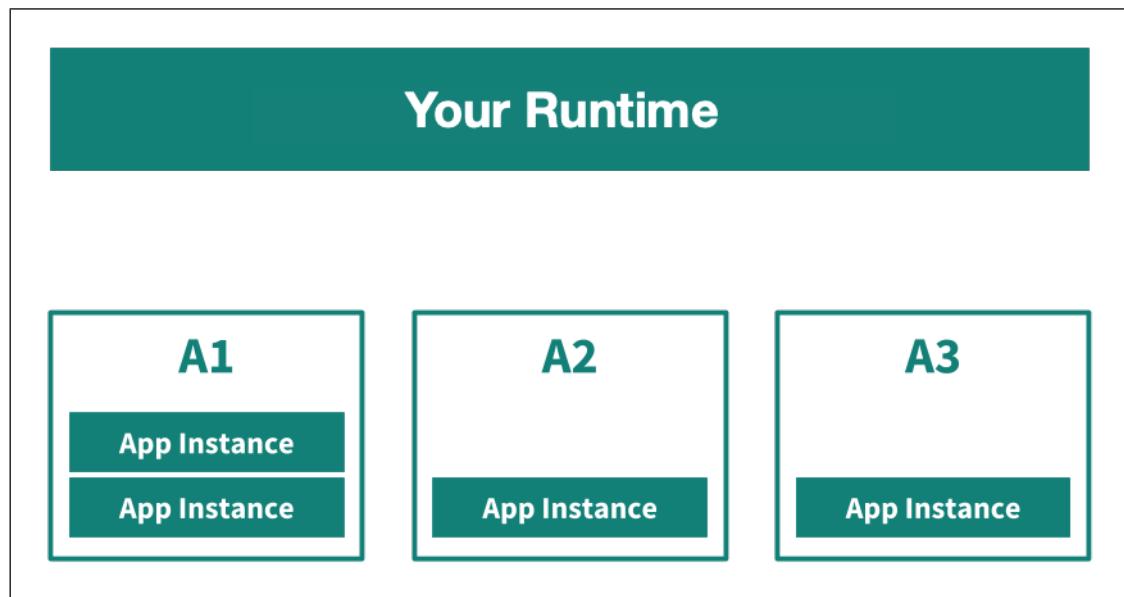
vSphere supports distributing deployments across multiple AZs. For more information, see [Step 4: Create Availability Zones in Configuring BOSH Director on vSphere](#).

It is recommended that customers use three Availability Zones to operate a highly available tile deployment.

Balancing Across AZs During Failure: Example Scenario

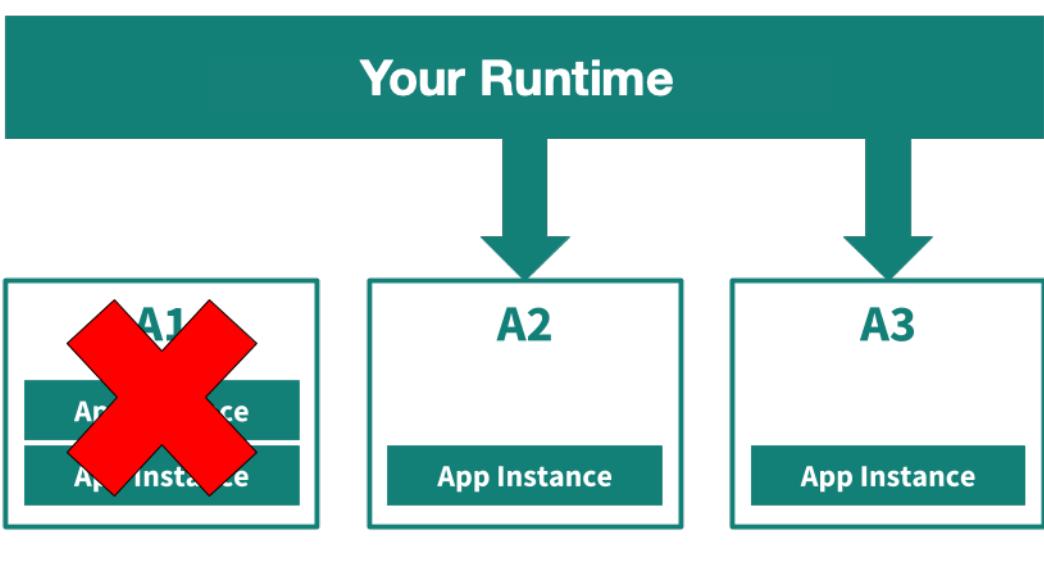
An operator scales an app to four instances in a tile environment distributed across three availability zones: A1, A2, and A3. The environment allocates the instances according to the [Diego Auction](#).

For a visual representation of this scenario, see the following diagram:



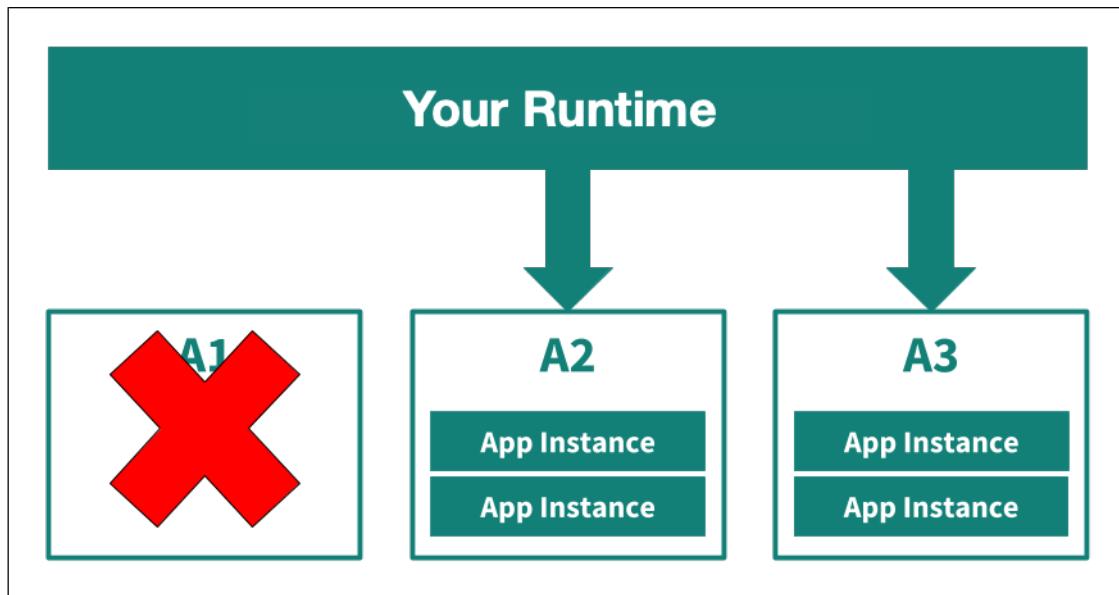
If A1 experiences a power outage or hardware failure, the two app instances running in A1 terminate while the app instances in zones A2 and A3 continue to run.

For a visual representation of this scenario, see the following diagram:



If A1 remains unavailable, the tile balances new instances of the app across the remaining availability zones.

For a visual representation of this scenario, see the following diagram:



Updating NSX-V Security Group and Load Balancer Information

This topic describes how to update security group and load balancer information for Pivotal Cloud Foundry (PCF) deployments using NSX-V on vSphere. To update this information, you must use the Ops Manager API.

See the [Ops Manager API documentation](#) for more information about the API.

See [Using Edge Services Gateway on VMware NSX](#) for guidance on how to configure the NSX firewall, load balancing, and NAT/SNAT services for PCF on vSphere installations.

 **Note:** Ops Manager v1.11 supports NSX-V v6.2 and later.

Authenticate

To use the Ops Manager API, you must authenticate and retrieve a token from the Ops Manager User Account and Authentication (UAA) server. For instructions, see the [Using Ops Manager API](#) topic.

Update Security Group or Load Balancer Information

To update either NSX security group or load balancer information, you use `curl` to make a PUT request against the `api/v0/staged/products/product_guid/jobs/job_guid/resource_config` endpoint.

You must first retrieve the GUID of your PCF deployment, and the GUID of the job whose information you want to update.

Do the following:

1. Retrieve a list of staged products:

```
$ curl 'https://OPS-MAN-FQDN/api/v0/staged/products' \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
[ {
  "product_version" : "1.10.6.0",
  "guid" : "p-bosh-dee11e111e111ee1e1a",
  "installation_name" : "p-bosh",
  "type" : "p-bosh"
},
{
  "type" : "cf",
  "product_version" : "1.10.8-build.7",
  "installation_name" : "cf-01222ab111111aa1a",
  "guid" : "cf-01222ab111111aa1a"
}
]
```

Record the GUID of the `cf` product. In the above example, the GUID is `cf-01222ab111111aa1a`.

2. Retrieve a list of jobs for your product:

```
$ curl 'https://OPS-MAN-FQDN/api/v0/staged/products/PRODUCT-GUID/jobs' \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
{
  "jobs" : [
    {
      "guid" : "consul_server-9c37cf48ae7412f2af1",
      "name" : "consul_server"
    },
    {
      "name" : "nats",
      "guid" : "nats-6af18efdd18d198edee9"
    },
    {
      "name" : "nfs_server",
      "guid" : "nfs_server-b49b0b2aed247302c0e1"
    },
    ...
  ]
}
```

Record the GUID of the job whose security groups you want to update.

3. You can update either your security group information, load balancer information, or both.

- o **Security groups:** To update the security groups for your job, use the following command:

```
$ curl "https://OPS-MAN-FQDN/api/v0/staged/products/PRODUCT-GUID/jobs/JOB-GUID/resource_config" \
-X PUT \
-H "Content-Type: application/json" \
-H "Authorization: Bearer UAA-ACCESS-TOKEN" \
-d '{"instance_type": {"id": "INSTANCE-TYPE"}, \
"persistent_disk": {"size_mb": "DISK-SIZE"}, \
"nsx_security_groups": ["SECURITY-GROUP1", "SECURITY-GROUP2"]}'
```

Replace the placeholder values as follows:

- **INSTANCE-TYPE** : The instance type for the job. For the default instance type, use `"automatic"`.
- **DISK-SIZE** : The disk size for the job. For the default persistent disk size, use `"automatic"`.

Note: The `persistent_disk` parameter is required to make this API request. For jobs that do not have persistent disks, you must set the value of the parameter to `"automatic"`.

- **SECURITY-GROUP1** , **SECURITY-GROUP2** : The value of the `nsx_security_groups` parameter is a list of the security groups that you want to set for the job. To clear all security groups for a job, pass an empty list with the `[]` value.

- o **Load balancers:** To update the load balancers for your job, use the following command:

```
$ curl "https://OPS-MAN-FQDN/api/v0/staged/products/PRODUCT-GUID/jobs/JOB-GUID/resource_config" \
-X PUT \
-H "Content-Type: application/json" \
-H "Authorization: Bearer UAA-ACCESS-TOKEN" \
-d '{"instance_type": {"id": "INSTANCE-TYPE"}, \
"persistent_disk": {"size_mb": "DISK-SIZE"}, \
"nsx_lbs": [ { \
"edge_name": "EDGE-NAME", \
"pool_name": "POOL-NAME", \
"security_group": "SECURITY-GROUP", \
"port": "PORT-NUMBER" \
} ]}'
```

Replace the placeholder values as follows:

- **INSTANCE-TYPE** : The instance type for the job. For the default instance type, use `"automatic"`.
- **DISK-SIZE** : The disk size for the job. For the default persistent disk size, use `"automatic"`.

Note: The `persistent_disk` parameter is required to make this API request. For jobs that do not have persistent disks, you must set the value of the parameter to `"automatic"`.

- **EDGE-NAME** : The name of the NSX Edge.
- **POOL-NAME** : The name of the NSX Edge's server pool.
- **SECURITY-GROUP** : The name of the NSX server pool's target security group.
- **PORT** : The name of the port that the VM service is listening on, such as `5000`. You can configure more than one load balancer for a job by using additional hashes in the `nsx_lbs` array.

4. Navigate to `OPS-MAN-FQDN` in a browser and log in to the Ops Manager Installation Dashboard.

5. Click **Review Pending Changes**, then **Apply Changes** to redeploy.

Updating NSX-T Load Balancer Server Pool Membership

This topic describes how to update load balancer server pool membership for Pivotal Cloud Foundry (PCF) deployments using NSX-T on vSphere. To update this information, you must use the Ops Manager API.

See the [Ops Manager API documentation](#) for more information about the API.

 **Note:** This procedure requires NSX-T v2.2 or later.

Authenticate

To use the Ops Manager API, you must authenticate and retrieve a token from the Ops Manager User Account and Authentication (UAA) server. For instructions, see the [Using Ops Manager API](#) topic.

Update Load Balancer Server Pool Membership

To update load balancer server pool membership, do the following:

1. [Create VM Extensions](#)
2. [Stage and Associate VM Extensions](#)
3. [Apply Changes](#)

Create VM Extensions

To create VM extensions, do the following:

1. Create a file named `web.json` and paste the following. This is the VM extension for the VMs that serve as the HTTP or HTTPS endpoint of the foundation. The BOSH instance group `router` serves as the HTTP or HTTPS endpoint.

```
{  
  "name": "pas_web_vm_extension",  
  "cloud_properties": {  
    "nsxt": {  
      "lb": {  
        "server_pools": [  
          {  
            "name": "pas-web-pool"  
          }  
        ]  
      }  
    }  
  }  
}
```

2. Create a file named `ssh.json` and paste the following. This is the VM extension for the VMs that serve as the SSH endpoint of the foundation. The BOSH instance group `diego_brain` serves as the SSH endpoint. On Small Footprint PAS the instance group is called `control`.

```
{  
  "name": "pas_ssh_vm_extension",  
  "cloud_properties": {  
    "nsxt": {  
      "lb": {  
        "server_pools": [  
          {  
            "name": "pas-ssh-pool"  
          }  
        ]  
      }  
    }  
  }  
}
```

3. Create a file named `tcp.json`. This is the VM extension for the VMs that serve as the endpoint of apps that require custom TCP ports. The BOSH instance group `tcp_router` serves as the custom TCP port endpoint On Small Footprint PAS this instance group is not deployed by default.

```
{
  "name": "pas_tcp_vm_extension",
  "cloud_properties": {
    "nsxt": {
      "lb": {
        "server_pools": [
          {
            "name": "pas-tcp-pool"
          }
        ]
      }
    }
  }
}
```

Stage and Associate VM Extensions

To stage and associate VM extensions with the load balancer, do the following:

1. Run the following command to stage the `web.json` VM extension:

```
curl "https://OPS-MAN-FQDN/api/v0/staged/vm_extensions" \
-X POST \
-H "Authorization: Bearer UAA-ACCESS-TOKEN" \
-H "Content-Type: application/json" \
-d "$(cat $VM_EXTENSION_FILE_PATH)"
```

Where:

- o `OPS-MAN-FQDN` is the fully qualified domain name of your Ops Manager installation.
- o `UAA-ACCESS-TOKEN` is the access token you retrieved when authenticating with the Ops Manager API.
- o `VM_EXTENSION_FILE_PATH` is the path to the VM extension file.

2. Repeat the previous step for the `ssh.json` and `tcp.json` VM extensions.

3. Run the following command to retrieve a list of staged products:

```
curl 'https://OPS-MAN-FQDN/api/v0/staged/products' \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
```

4. Record the GUID of the `cf` product. In the following example output, the GUID is `cf-01222ab111111aa1a`.

```
[
  {
    "product_version": "1.10.6.0",
    "guid": "p-bosh-dee11e11e111ee1e1a",
    "installation_name": "p-bosh",
    "type": "p-bosh"
  },
  {
    "type": "cf",
    "product_version": "1.10.8-build.7",
    "installation_name": "cf-01222ab111111aa1a",
    "guid": "cf-01222ab111111aa1a"
  }
]
```

5. Retrieve the list of jobs for `cf`:

```
curl https://OPS-MAN-FQDN/api/v0/staged/products/PRODUCT-GUID/jobs' \
-H 'Authorization: Bearer UAA-ACCESS-TOKEN'
```

6. Record the value of `guid` for the `router`, `tcp_router`, and `diego_brain` jobs. See the following example output:

```
{
"jobs": [
{
  "name": "router",
  "guid": "router-9c37cf48ae7412f2afdf1"
},
{
  "name": "tcp\_router",
  "guid": "tcp\_router-6af18cfdd18d198edee9"
},
{
  "name": "diego\_brain",
  "guid": "diego\_\_brain-b49b0b2aed247302c0e1"
},
...
...
```

- To associate your job with its appropriate VM extension, run the following command:

```
curl " https://OPS-MAN-FQDN/api/v0/staged/products/PRODUCT-GUID/jobs/JOB-GUID/resource\_config" \
-X PUT \
-H "Content-Type: application/json" \
-H "Authorization: Bearer UAA-ACCESS-TOKEN" \
-d '{
  "instance_type": {"id": "INSTANCE-TYPE"},
  "instances": INSTANCE-COUNT,
  "additional_vm_extensions": ["VM-EXTENSION-NAME"]
}'
```

⚠ warning: Do not use the resource config fields for `nsx_lbs` and `nsx_security_groups`. These fields are only for NSX-V.

Where:

- `PRODUCT-GUID` is the GUID you recorded for the `cf` product.
- `JOB-GUID` is the GUID you recorded for the `router` job.
- `INSTANCE-TYPE` is the instance type for the job. For the default instance type, use `"automatic"`.
- `INSTANCE-COUNT` is the number of instances for the job. The default number for each job is visible in the Resource Config section of the Ops Manager UI.
- `VM-EXTENSION-NAME` is `pas_web_vm_extension`.

- Run the previous command again.

Where:

- `JOB-GUID` is the GUID you recorded for the `tcp_router` job.
- `VM-EXTENSION-NAME` is `pas_tcp_vm_extension`.

- Run the previous command again.

Where:

- `JOB-GUID` is the GUID you recorded for the `diego_brain` job.
- `VM-EXTENSION-NAME` is `pas_ssh_vm_extension`.

Apply Changes

To apply changes, do the following:

- Log in to the Ops Manager Installation Dashboard.
- Click **Review Pending Changes**, then **Apply Changes** to redeploy.

Configuring PAS

This topic describes how to configure Pivotal Application Service (PAS) as part of deploying Pivotal Cloud Foundry (PCF).

Before beginning this procedure, ensure that you have successfully completed the steps to prepare your environment for PCF and install and configure the BOSH Director.

Note: If you plan to [install the PCF IPsec add-on](#), you must do so before installing any other tiles. Pivotal recommends installing IPsec immediately after Ops Manager, and before installing the PAS Runtime tile.

Add PAS to Ops Manager

1. If you have not already downloaded PAS, log in to [Pivotal Network](#), and click PAS.
2. From the **Releases** drop-down, select the release to install and choose one of the following:
 - a. Click PAS to download the PAS `.pivotal` file.
 - b. Click **PCF Small Footprint Runtime** to download the Small Footprint Runtime `.pivotal` file. For more information, see [Getting Started with Small Footprint Runtime](#).
3. Navigate to the Pivotal Cloud Foundry Operations Manager Installation Dashboard.
4. Click **Import a Product** to add your tile to Ops Manager. For more information, refer to the [Adding and Deleting Products](#) topic.
5. Click the **Pivotal Application Service** tile in the Installation Dashboard.

Assign AZs and Networks

Note: For Azure environments, this configuration pane is **Assign Networks** and does not include AZ configuration.

1. Select **Assign AZ and Networks**.
2. Select the first Availability Zone under **Place singleton jobs**. Ops Manager runs any job with a single instance in this Availability Zone.
3. Select all Availability Zones under **Balance other jobs**. Ops Manager balances instances of jobs with more than one instance across the Availability Zones that you specify.
4. From the **Network** dropdown, choose the runtime network that you created when configuring the BOSH Director tile.
5. Click **Save**.

Note: For production deployments, Pivotal recommends at least three Availability Zones for a highly available installation.

Configure Domains

1. Select **Domains**. The following image shows the **Domains** page.

Application Service hosts applications at subdomains under its apps domain and assigns system components to subdomains under its system domain. You need to configure a wildcard DNS for both the apps domain and system domain. The two domains can be the same, although this is not recommended.

System Domain *	<input type="text" value="sys.metallicseaweed.cf-app.com"/>
Apps Domain *	<input type="text" value="apps.metallicseaweed.cf-app.com"/>
<input type="button" value="Save"/>	

2. Enter the system and application domains. The **System Domain** defines your target when you push apps to PAS. For example, `system.example.com`. The **Apps Domain** defines where PAS should serve your apps. For example, `apps.example.com`. For additional guidance based on your installation method, see the following table:

Installation Method	Guidance
Manual	Enter the domains you created when preparing your environment for PCF.
Terraform	Enter the values for <code>sys_domain</code> and <code>apps_domain</code> from the Terraform output.

 **Note:** Pivotal recommends that you use the same domain name but different subdomain names for your system and app domains. Doing so allows you to use a single wildcard certificate for the domain while preventing apps from creating routes that overlap with system routes.

3. Click **Save**.

Configure Networking

1. For **Router IPs** **HAProxy IPs**, see the following guidance:

- For **AWS**, **Azure**, and **GCP**, leave these fields blank. You do not need to complete these fields when deploying PCF on these infrastructures.
- For **OpenStack** and **vSphere** the values you enter in the **Router IPs** and **HAProxy IPs** fields depend on whether you are using HAProxy in your deployment. Use the table below to determine how to complete these fields.

 **Note:** If you choose to assign specific IP addresses in either the **Router IPs** or **HAProxy IPs** field, ensure that these IP addresses are in the subnet that you configured for PAS in Ops Manager.

Using HAProxy?	Router IPs Field	HAProxy IPs Field
No	<ol style="list-style-type: none"> 1. Choose IP addresses from the subnet you configured in Ops Manager. 2. Enter these IP addresses in the Router IPs field. You should specify more than one IP address for high availability. 3. Configure your load balancer to forward requests for the domains that you have configured for your deployment to these IP addresses. 	Leave this field blank.
Yes	Leave this field blank.	<ol style="list-style-type: none"> 1. Choose IP addresses from the subnet you configured in Ops Manager. 2. Enter these IP addresses in the HAProxy IPs field. You should specify more than one IP address for high availability. 3. Configure your load balancer to forward requests for the domains you have configured for your deployment to these IP addresses.

2. For **SSH Proxy IPs** and **TCP Router IPs**, see the following guidance:

- For **AWS**, **Azure**, and **GCP**, leave these fields blank. You do not need to complete these fields when deploying PCF on these infrastructures.
- For **OpenStack** and **vSphere**:
 - (Optional) In **SSH Proxy IPs**, add the IP address for your Diego Brain, which will accept requests to SSH into application containers on port `2222`.
 - (Optional) In **TCP Router IPs**, add the IP address(es) you would like assigned to the TCP Routers. You enable this feature at the bottom of this screen.

 **Note:** If you have mutual TLS app identity verification enabled, app containers accept incoming communication only from the Gorouter. This disables TCP routing.

3. Under **Certificates and Private Key for HAProxy and Router**, you must provide at least one **Certificate and Private Key** name and certificate key pair for HAProxy and Gorouter. HAProxy and Gorouter are enabled to receive TLS communication by default. You can configure multiple certificates for HAProxy and Gorouter.

Note: When providing custom certificates, enter them in the following order: `wildcard`, `Intermediate`, `CA`. For more information, see [Creating a .pem File for SSL Certificate Installations](#) in the DigiCert documentation.

- Click **Add** to add a name for the certificate chain and its private key pair. This certificate is the default used by Gorouter and HAProxy. You can either provide a certificate signed by a Certificate Authority (CA) or click on the **Generate RSA Certificate** link to generate a certificate generated by the Ops Manager CA.

Note: If you configured Ops Manager Front End without a certificate, you can use this new certificate to complete Ops Manager configuration. To configure your Ops Manager Front End certificate, see *Configure Front End* in [Preparing to Deploy Ops Manager on GCP Manually](#).

- If you want to configure multiple certificates for HAProxy and Gorouter, click **Add** and fill in the appropriate fields for each additional certificate key pair. For details about generating certificates in Ops Manager for your wildcard system domains, see the [Providing a Certificate for Your SSL/TLS Termination Point](#) topic.

Note: Ensure that you add any certificates that you generate in this pane to your infrastructure load balancer.

- (Optional) When validating client requests using mutual TLS, the Gorouter trusts multiple certificate authorities (CAs) by default. If you want to configure the Gorouter and HAProxy to trust additional CAs, enter your CA certificates under **Certificate Authorities Trusted by Router and HAProxy**. All CA certificates should be appended together into a single collection of PEM-encoded entries.
- In the **Minimum version of TLS supported by HAProxy and Router** field, select the minimum version of TLS to use in HAProxy and Gorouter communications. HAProxy and Gorouter use TLS v1.2 by default. If you need to accommodate clients that use an older version of TLS, select a lower minimum version. For a list of TLS ciphers supported by the Gorouter, see [Securing Traffic into Cloud Foundry](#).
- Configure **Logging of Client IPs in CF Router**. The **Log client IPs** option is set by default. To comply with the General Data Protection Regulation (GDPR), select one of the following options to disable logging of client IP addresses:
 - If your load balancer exposes its own source IP address, disable logging of the `X-Forwarded-For` HTTP header only.
 - If your load balancer exposes the source IP of the originating client, disable logging of both the source IP address and the `X-Forwarded-For` HTTP header.
- Under **Configure support for the X-Forwarded-Client-Cert header**, configure PCF handles `x-forwarded-client-cert` (XFCC) HTTP headers based on where TLS is terminated for the first time in your deployment.

The following table indicates which option to choose based on your deployment configuration.

Deployment Configuration	TLS Option	Additional Notes
<ul style="list-style-type: none"> ◦ The Load Balancer is terminating TLS, and ◦ Load balancer is configured to put the client certificate from a mutual authentication TLS handshake into the X-Forwarded-Client-Cert HTTP header 	TLS terminated for the first time at infrastructure load balancer (default).	Both HAProxy and the Gorouter forward the XFCC header when included in the request.
<ul style="list-style-type: none"> ◦ The Load Balancer is configured to pass through the TLS handshake via TCP to the instances of HAProxy, and ◦ HAProxy instance count is > 0 	TLS terminated for the first time at HAProxy.	<p>HAProxy sets the XFCC header with the client certificate received in the TLS handshake. The Gorouter forwards the header.</p> <p>Breaking Change: In the Router behavior for Client Certificates field, you cannot select the Router does not request client certificates option.</p>
<ul style="list-style-type: none"> ◦ The Load Balancer is configured to pass through the TLS handshake via TCP to instances of the Gorouter 	TLS terminated for the first time at the Gorouter.	<p>The Gorouter strips the XFCC header if it is included in the request and forwards the client certificate received in the TLS handshake in a new XFCC header.</p> <p>If you have deployed instances of HAProxy, app traffic bypasses those instances in this configuration. If you have also configured your load balancer to route requests for ssh directly to the Diego Brain, consider reducing HAProxy instances to 0.</p> <p>Breaking Change: In the Router behavior for Client Certificates field, you cannot select the Router does not request client certificates option.</p>

certificates option.

For a description of the behavior of each configuration option, see [Forward Client Certificate to Applications](#).

8. To configure HAProxy to handle client certificates, select one of the following options in the **HAProxy behavior for Client Certificate Validation** field:

- **HAProxy does not request client certificates.** This option requires mutual authentication, which makes it incompatible with XFCC option **TLS terminated for the first time at HAProxy**. HAProxy does not request client certificates, so the client does not provide them and no validation occurs. This is the default configuration.
- **HAProxy requests but does not require client certificates.** The HAProxy requests client certificates in TLS handshakes, validates them when presented, but does not require them.

A warning: Upon upgrade, PAS will fail to receive requests if your load balancer is configured to present a client certificate in the TLS handshake with HAProxy but HAProxy has not been configured with the certificate authority used to sign it. To mitigate this issue, select **HAProxy does not request client certificates** in the [Networking](#) pane or configure the HAProxy with the appropriate CA.

9. To configure Gorouter behavior for handling client certificates, select one of the following options in the **Router behavior for Client Certificate Validation** field:

- **Router does not request client certificates.** This option is incompatible with the XFCC configuration options **TLS terminated for the first time at HAProxy** and **TLS terminated for the first time at the Router** in PAS because these options require mutual authentication. As client certificates are not requested, client will not provide them, and thus validation of client certificates will not occur.
- **Router requests but does not require client certificates.** The Gorouter requests client certificates in TLS handshakes, validates them when presented, but does not require them. This is the default configuration.
- **Router requires client certificates.** The Gorouter validates that the client certificate is signed by a Certificate Authority that the Gorouter trusts. If the Gorouter cannot validate the client certificate, the TLS handshake fails.

A warning: Requests to the platform will fail upon upgrade if your load balancer is configured with client certificates and the Gorouter does not have the certificate authority. To mitigate this issue, select **Router does not request client certificates** for **Router behavior for Client Certificate Validation** in the [Networking](#) pane.

10. In the **TLS Cipher Suites for Router** field, review the TLS cipher suites for TLS handshakes between Gorouter and front-end clients such as load balancers or HAProxy. The default value for this field is `ECDHE-RSA-AES128-GCM-SHA256:TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384`.

To modify the default configuration, use an ordered, colon-delimited list of Golang-supported TLS cipher suites in the OpenSSL format.

Operators should verify that the ciphers are supported by any clients or front-end components that will initiate TLS handshakes with Gorouter. For a list of TLS ciphers supported by Gorouter, see [Securing Traffic into Cloud Foundry](#).

Verify that every client participating in TLS handshakes with Gorouter has at least one cipher suite in common with Gorouter.

Note: Specify cipher suites that are supported by the versions configured in the **Minimum version of TLS supported by HAProxy and Router** field.

Note: AWS Classic Load Balancers do not support PCF's default cipher suites. See [TLS Cipher Suite Support by AWS Load Balancers](#) for information about configuring your AWS load balancers and Gorouter.

11. In the **TLS Cipher Suites for HAProxy** field, review the TLS cipher suites for TLS handshakes between HAProxy and its clients such as load balancers and Gorouter. The default value for this field is the following:

`DHE-RSA-AES128-GCM-SHA256:DHE-RSA-AES256-GCM-SHA384:ECDHE-RSA-AES128-GCM-SHA256:ECDHE-RSA-AES256-GCM-SHA384`

To modify the default configuration, use an ordered, colon-delimited list of TLS cipher suites in the OpenSSL format.

Operators should verify that the ciphers are supported by any clients or front-end components that will initiate TLS handshakes with HAProxy.

Verify that every client participating in TLS handshakes with HAProxy has at least one cipher suite in common with HAProxy.

Note: Specify cipher suites that are supported by the versions configured in the **Minimum version of TLS supported by HAProxy and Router** field.

12. Under **HAProxy forwards requests to Router over TLS**, select **Enable** or **Disable** based on your deployment layout.

- **Enable HAProxy forwarding of requests to Router over TLS.** To enable communication between HAProxy and the Gorouter, do the following:
 1. Leave **Enable** selected.
 2. In the **Certificate Authority for HAProxy Backend** field, specify the Certificate Authority (CA) that signed the certificate you configured in the **Certificate and Private Key for HAProxy and Router** field.

Note: If you used the [Generate RSA Certificate](#) link to generate a certificate, then the CA to specify is the Ops Manager CA, which you can locate at the </api/v0/certificateAuthorities> endpoint in the Ops Manager API.

3. Make sure that Gorouter and HAProxy have TLS cipher suites in common in the **TLS Cipher Suites for Router** and **TLS Cipher Suites for HAProxy** fields.

For more information, see [Terminating SSL/TLS at the Load Balancer and Gorouter](#), [Providing a Certificate for Your SSL/TLS Termination Point](#), and [Using the Ops Manager API](#).

- **Disable HAProxy forwarding of requests to Router over TLS.** To use non-encrypted communication between HAProxy and Gorouter, do the following:
 1. Select **Disable**.
 2. If you are not using HAProxy, set the number of HAProxy job instances to on the **Resource Config** page. See [Disable Unused Resources](#).

For more information, see [Terminating SSL/TLS at the Gorouter Only](#) and [Terminating SSL/TLS at the Load Balancer Only](#).

13. If you want to force browsers to use HTTPS when making requests to HAProxy, select **Enable** in the **HAProxy support for HSTS** field, and complete the following optional configuration steps:
 - a. Enter a **Max Age in Seconds** for the HSTS request. By default, the age is set to one year. HAProxy will force HTTPS requests from browsers for the duration of this setting.
 - b. Select the **Include Subdomains** checkbox to force browsers to use HTTPS requests for all component subdomains.
 - c. Select the **Enable Preload** checkbox to force instances of Google Chrome, Firefox, and Safari that access your HAProxy to refer to their built-in lists of known hosts that require HTTPS, of which HAProxy is one. This ensures that the first contact a browser has with your HAProxy is an HTTPS request, even if the browser has not yet received an HSTS header from HAProxy.
14. If you are not using SSL encryption or if you are using self-signed certificates, select **Disable SSL certificate verification for this environment**. Selecting this checkbox also disables SSL verification for route services and disables mutual TLS app identity verification.

Note: For production deployments, Pivotal does not recommend disabling SSL certificate verification.

15. (Optional) If you want HAProxy or the Gorouter to reject any HTTP (non-encrypted) traffic, select the **Disable HTTP on HAProxy and Gorouter** checkbox. When selected, HAProxy and Gorouter will not listen on port 80.
16. (Optional) Select the **Disable insecure cookies on the Router** checkbox to set the secure flag for cookies generated by the router.
17. (Optional) To disable the addition of Zipkin tracing headers on the Gorouter, deselect the **Enable Zipkin tracing headers on the router** checkbox. Zipkin tracing headers are enabled by default. For more information about using Zipkin trace logging headers, see [Zipkin Tracing in HTTP Headers](#).
18. (Optional) To stop the Router from writing access logs to local disk, deselect the **Enable Router to write access logs locally** checkbox. You should consider disabling this checkbox for high traffic deployments since logs may not be rotated fast enough and can fill up the disk.
19. By default, the PAS routers handle traffic for applications deployed to an isolation segment created by the PCF Isolation Segment tile. To configure the PAS routers to reject requests for applications within isolation segments, select the **Routers reject requests for Isolation Segments** checkbox. Do not enable this option without deploying routers for each isolation segment. See the following topics for more information:
 - [Installing PCF Isolation Segment](#)
 - [Sharding Routers for Isolation Segments](#).
20. (Optional) By default, Gorouter support for the PROXY protocol is disabled. To enable the PROXY protocol, select **Enable support for PROXY protocol in CF Router**. When enabled, client-side load balancers that terminate TLS but do not support HTTP can pass along information from the originating client. Enabling this option may impact Gorouter performance. For more information about enabling the PROXY protocol in Gorouter, see the [HTTP Header Forwarding](#) sections in the [Securing Traffic in Cloud Foundry](#) topic.
21. In the **Choose whether to enable route services** section, choose either **Enable route services** or **Disable route services**. Route services are a class of [marketplace services](#) that perform filtering or content transformation on application requests and responses. See the [Route Services](#) topic for details.
 - a. If you enabled route services, you can also configure the **Bypass security checks for route service lookup** field. Pivotal recommends that you do not enable this field because it has potential security concerns. However, you may need to enable it if your load balancer requires mutual TLS from clients. For more information, see [Configuring Route Service Lookup](#).

22. (Optional) If you want to limit the number of app connections to the backend, enter a value in the **Max Connections Per Backend** field. You can use this field to prevent a poorly behaving app from all the connections and impacting other apps.

To choose a value for this field, review the peak concurrent connections received by instances of the most popular apps in your deployment. You can determine the number of concurrent connections for an app from the `httpStartStop` event metrics emitted for each app request.

If your deployment uses PCF Metrics, you can also obtain this peak concurrent connection information from [Network Metrics](#). The default value is `500`.

23. Under **Enable Keepalive Connections for Router**, select **Enable** or **Disable**. Keep-alive connections are enabled by default. For more information, see [Keepalive Connections in HTTP Routing](#).

24. (Optional) To accommodate larger uploads over connections with high latency, increase the number of seconds in the **Router Timeout to Backends** field.

25. (Optional) Use the **Frontend Idle Timeout for Gorouter and HAProxy** field to help prevent connections from your load balancer to Gorouter or HAProxy from being closed prematurely. The value you enter sets the duration, in seconds, that Gorouter or HAProxy maintains an idle open connection from a load balancer that supports keep-alive.

In general, set the value higher than your load balancer's backend idle timeout to avoid the race condition where the load balancer sends a request before it discovers that Gorouter or HAProxy has closed the connection.

See the following table for specific guidance and exceptions to this rule:

IaaS	Guidance
AWS	AWS ELB has a default timeout of 60 seconds, so Pivotal recommends a value greater than <code>60</code> .
Azure	By default, Azure load balancer times out at 240 seconds without sending a TCP RST to clients, so as an exception, Pivotal recommends a value lower than <code>240</code> to force the load balancer to send the TCP RST.
GCP	GCP has a default timeout of 600 seconds. For GCP HTTP load balancers, Pivotal recommends a value greater than <code>600</code> . For GCP TCP load balancers, pivotal recommends a value less than <code>600</code> to force the load balancer to send a TCP RST.
Other	Set the timeout value to be greater than that of the load balancer's backend idle timeout.

 **Note:** Do not set a frontend idle timeout lower than six seconds.

26. (Optional) Increase the value of **Load Balancer Unhealthy Threshold** to specify the amount of time, in seconds, that the router continues to accept connections before shutting down. During this period, healthchecks may report the router as unhealthy, which causes load balancers to failover to other routers. Set this value to an amount greater than or equal to the maximum time it takes your load balancer to consider a router instance unhealthy, given contiguous failed healthchecks.

27. (Optional) Modify the value of **Load Balancer Healthy Threshold**. This field specifies the amount of time, in seconds, to wait until declaring the Router instance started. This allows an external load balancer time to register the Router instance as healthy.

28. (Optional) If app developers in your organization want certain HTTP headers to appear in their app logs with information from the Gorouter, specify them in the **HTTP Headers to Log** field. For example, to support app developers that deploy Spring apps to PCF, you can enter [Spring-specific HTTP headers](#).

29. If you expect requests larger than the default maximum of 16 Kbytes, enter a new value (in bytes) for **HAProxy Request Max Buffer Size**. You may need to do this, for example, to support apps that embed a large cookie or query string values in headers.

30. If your PCF deployment uses HAProxy and you want it to receive traffic only from specific sources, use the following fields:

- **HAProxy Protected Domains:** Enter a comma-separated list of domains to protect from unknown source requests.
- **HAProxy Trusted CIDRs:** Optionally, enter a space-separated list of CIDRs to limit which IP addresses from the **Protected Domains** can send traffic to PCF.

31. The **Loggregator Port** defaults to `443` if left blank. For AWS environments that are not using an Application Load Balancer, enter `4443`.

32. For **Container Network Interface Plugin**, select one of the following:

- **Silk:** This option is the default Container Network Interface (CNI) for PAS.
- **External:** Select this if you are deploying the [VMware NSX-T Container Plug-in for PCF](#).
 - If you select **External**, follow the instructions in [Deploying PAS with NSX-T Networking](#) in addition to the PAS configuration instructions in this topic.

warning: The NSX-T integration only works for fresh installs of PCF. If your PAS is already deployed and running with Silk as its CNI, you cannot change the CNI plugin to NSX-T.

33. If you selected **Silk** in the previous step, review the following fields:

- (Optional) You can change the value in the **Applications Network Maximum Transmission Unit (MTU)** field. Pivotal recommends setting the MTU value for your application network to `1454`. Some configurations, such as networks that use GRE tunnels, may require a smaller MTU value.
- (Optional) Enter an IP range for the overlay network in the **Overlay Subnet** box. If you do not set a custom range, Ops Manager uses `10.255.0.0/16`.

warning: The overlay network IP range must not conflict with any other IP addresses in your network.

- Enter a UDP port number in the **VXLAN Tunnel Endpoint Port** box. If you do not set a custom port, Ops Manager uses `4789`.
- For **Denied logging interval**, set the per-second rate limit for packets blocked by either a container-specific [networking policy](#) or by [Application Security Group](#) rules applied across the space, org, or deployment. This field defaults to `1`.
- For **UDP logging interval**, set the per-second rate limit for UDP packets sent and received. This field defaults to `100`.
- To enable logging for app traffic, select **Log traffic for all accepted/denied application packets**. See [Manage Logging for Container-to-Container Networking](#) for more information.
- The **Enable Silk Policy Enforcement** checkbox is enabled by default. To disable Silk network policy enforcement between apps, disable the checkbox. Disabling network policy enforcement allows all apps to send network traffic to all other apps in the foundation despite no policy specifically allowing it.

34. For **DNS Search Domains**, enter DNS search domains for your containers as a comma-separated list. DNS on your containers appends these names to its host names, to resolve them into full domain names.

35. For **Database Connection Timeout**, set the connection timeout for clients of the policy server and silk databases. The default value is `120`. You may need to increase this value if your deployment experiences timeout issues related to Container-to-Container Networking.

36. (Optional) TCP Routing is disabled by default. You should enable this feature if your DNS sends TCP traffic through a load balancer rather than directly to a TCP router. To enable TCP routing:

- Select **Enable TCP Routing**.
- For **TCP Routing Ports**, enter a single port or a range of ports for the load balancer to forward to. These are the same ports that you configured in the [Pre-Deployment Steps](#) of the [Enabling TCP Routing](#) topic.
 - To support multiple TCP routes, Pivotal recommends allocating multiple ports.
 - To allocate a list of ports rather than a range:
 1. Enter a single port in the **TCP Routing Ports** field.
 2. After deploying PAS, follow the directions in [Configuring a List of TCP Routing Ports](#) to add TCP routing ports using the cf CLI.

Note: If you have mutual TLS app identity verification enabled, app containers accept incoming communication only from the Gorouter. This disables TCP routing.

- For GCP, you also need to specify the name of a GCP TCP load balancer in the **LOAD BALANCER** column of TCP Router job of the [Resource Config](#) screen. You configure this later on in PAS. See the [Configure Resources](#) section of this topic.
- For AWS, you also need to specify the name of a TCP ELB in the **LOAD BALANCER** column of TCP Router job of the [Resource Config](#) screen. You configure this later on in PAS. For more information, see the [Configure Resources](#) section of this topic.
- For Azure, you also need to specify the name of Azure load balancer in the **LOAD BALANCER** column of TCP Router job of the [Resource Config](#) screen. You configure this later on in PAS. See the [Configure Resources](#) section of this topic.
- For OpenStack and vSphere: Return to the top of the **Networking** screen. In **TCP Router IPs** field, make sure you have entered IP addresses within your subnet CIDR block. These will be the same IP addresses you configured your load balancer with in [Pre-Deployment Steps](#), unless you configured DNS to resolve the TCP domain name directly to an IP you've chosen for the TCP router.

37. (Optional) To disable TCP routing, click **Select this option if you prefer to enable TCP Routing at a later time**. For more information, see the [Configuring TCP Routing in PAS](#) topic.

38. (Optional) To enable the creation of dynamic egress policies that allow apps to reach external services, select **Enable Beta Dynamic Egress Enforcement**. For more information about this feature, see [Administering Dynamic Egress Policies \(Beta\)](#).

39. (Optional) For additional security, enter headers that you want the router to remove from app responses in **Remove Specified HTTP Response Headers**.

40. Click **Save**.

Configure Networking - Service Mesh (Beta)

PAS includes an optional, beta routing plane that uses a service mesh. A service mesh provides traffic management, security, and observability for microservices. It enables new features such as weighted routing. For more information, see [Service Mesh \(Beta\)](#).

To add service mesh to your deployment, do the following:

1. Under **Service Mesh (Beta)**, select **Enable**.
2. For **IP Addresses for Ingress Router**, do the following depending on your IaaS:
 - **vSphere**: Enter static IPs for the Istio Routers. You must configure your load balancer with these IPs as well.
 - **Other**: Leave this field blank.
3. For **Ingress Router TLS Keypairs**, complete the following fields. You can add more than one keypair if desired using the **Add** button.
 - **Name**: Enter a name for the keypair.
 - **Certificate and Private Key for Istio Router**: Enter the Private key and certificate for TLS handshakes with clients. These must be in PEM block format.
4. Click **Save**.
5. Create a load balancer and add it to the **Resource Config** pane as described in [Deploying Service Mesh \(Beta\)](#).

Configure Application Containers

1. Select **Application Containers**. The following image shows the **Application Containers** page.

Enable microservice frameworks, private Docker registries, and other services that support your applications at a container level.

Enable Custom Buildpacks

Allow SSH access to app containers

Enable SSH when an app is created

Router application identity verification*

Router uses TLS to verify application identity (Default)

Router and applications use mutual TLS to verify each other's identity (incompatible with TCP Routing)

Private Docker Insecure Registry Whitelist

10.10.10.10:8888,example.com:8888

Docker Images Disk-Cleanup Scheduling on Cell VMs*

Never clean up Cell disk-space

Routinely clean up Cell disk-space

Clean up disk-space once usage fills disk

Reserved disk space for other jobs(MB) (min: 15360) *

15360

Enable Containerd Delegation

Max Inflight Container Starts *

200

Enabling NFSv3 volume services will allow application developers to bind existing NFS volumes to their applications for shared file access. *

Enable

LDAP Service Account User

LDAP Service Account Password

LDAP Server Host

LDAP Server Port

LDAP User Search Base

LDAP Server CA Cert

Disable

Enable SMB volume services

Format of timestamps in Diego logs*
 RFC3339 timestamps (e.g. 2018-02-09T00:54:13.479724884Z)
 Seconds since the Unix epoch (e.g. 1518137653.479724884)

Default health check timeout *

Save

2. The **Enable Custom Buildpacks** checkbox governs the ability to pass a custom buildpack URL to the `-b` option of the `cf push` command. By default, this ability is enabled, letting developers use custom buildpacks when deploying apps. Disable this option by disabling the checkbox. For more information about custom buildpacks, refer to the [buildpacks](#) section of the PCF documentation.
3. The **Allow SSH access to app containers** checkbox controls SSH access to application instances. Enable the checkbox to permit SSH access across your deployment, and disable it to prevent all SSH access. See the [Application SSH Overview](#) topic for information about SSH access permissions at the space and app scope.
4. If you want to enable SSH access for new apps by default in spaces that allow SSH, select **Enable SSH when an app is created**. If you deselect the checkbox, developers can still enable SSH after pushing their apps by running `cf enable-ssh APP-NAME`.
5. To enable the Gorouter to verify app identity using TLS, click **Router uses TLS to verify application identity**. To enable the Gorouter and your apps to verify each other's identity using TLS, click **Router and applications use mutual TLS to verify each other's identity**.

Note: Mutual TLS app identity verification does not work for Windows containers.

Verifying app identity using TLS or mutual TLS enables encryption between router and app containers and guards against misrouting during control plane failures. For more information about Gorouter route consistency modes, see [Preventing Misrouting in HTTP Routing](#).

Note: To support mutual TLS app identity verification, you need v2.3 or later of both PAS and PCF Isolation Segment (IST). The Gorouter and cell components in PCF v2.2 and earlier do not support mTLS handshakes.

Note: This feature does not work if the **Disable SSL certificate verification for this environment** checkbox is selected.

6. You can configure Pivotal Application Service (PAS) to run app instances in Docker containers by supplying their IP address ranges in the **Private Docker Insecure Registry Whitelist** textbox. See the [Using Docker Registries](#) topic for more information.
7. Select your preference for **Docker Images Disk-Cleanup Scheduling on Cell VMs**. If you choose **Clean up disk-space once usage fills disk**, enter a **Reserved amount of disk space for other jobs** in megabytes. For more information about the configuration options and how to configure a reserved amount, see [Configuring Docker Images Disk-Cleanup Scheduling](#).

8. The **Enable Containerd Delegation** checkbox governs whether or not Garden delegates container create and destroy operations to the containerd tool. By default, this option is enabled and Garden uses containerd. Disable this option by disabling the checkbox. For more information about the containerd tool, see [containerd](#).
9. Enter a number in the **Max Inflight Container Starts** textbox. This number configures the maximum number of started instances across the Diego cells in your deployment. For more information about this feature, see [Setting a Maximum Number of Started Containers](#).
10. Under **Enabling NFSv3 volume services**, select **Enable** or **Disable**. NFS volume services allow application developers to bind existing NFS volumes to their applications for shared file access. For more information, see the [Enabling NFS Volume Services](#) topic.

Note: In a fresh install, NFSv3 volume services is enabled by default. In an upgrade, NFSv3 volume services is set to the same setting as it was in the previous deployment.

11. (Optional) To configure LDAP for NFSv3 volume services, do the following:
 - For **LDAP Service Account User**, enter the username of the service account in LDAP that will manage volume services.
 - For **LDAP Service Account Password**, enter the password for the service account.
 - For **LDAP Server Host**, enter the hostname or IP address of the LDAP server.
 - For **LDAP Server Port**, enter the LDAP server port number. If you do not specify a port number, Ops Manager uses 389.
 - For **LDAP User Search Base**, enter the location in the LDAP directory tree from which any LDAP User search begins. The typical LDAP Search Base matches your domain name.
For example, a domain named `cloud.example.com` typically uses the following LDAP User Search Base: `ou=Users,dc=example,dc=com`.
 - For **LDAP Server CA Cert**, you can optionally enter a certificate if your LDAP server supports TLS and you want to enable TLS connections from the NFS driver to your LDAP server. Paste in the root certificate from your CA certificate or your self-signed certificate.

Note: UAA can only parse one certificate entered into this field. If you enter multiple certificates, UAA only uses the first one you entered and ignores the rest. You only need to include one root certificate or self-signed certificate.

12. Select the **Format of timestamps in Diego logs**, either **RFC3339 timestamps** or **Seconds since the Unix epoch**. Fresh PAS v2.2 installations default to **RFC3339 timestamps**, while upgrades to PAS v2.2 from previous versions default to **Seconds since the Unix epoch**.
13. (Optional) Select the **Enable SMB volume services** checkbox to enable SMB volume services. Enabling SMB volume services allows developers to bind existing SMB shares to their apps. For more information, see [Enabling Volume Services](#).

Note: If you enable SMB volume services, in the **Errands** pane, set the **SMB Broker Errand** to **On**.

14. (Optional) Modify the **Default health check timeout**. The value configured for this field is the amount of time allowed to elapse between starting up an app and the first healthy response from the app. If the health check does not receive a healthy response within the configured timeout, then the app is declared unhealthy. The default timeout is `60` seconds and the maximum configurable timeout is `600` seconds.

15. Click **Save**.

Configure Application Developer Controls

1. Select **Application Developer Controls**. The following image shows the **Application Developer Controls** page.

Configure restrictions and default settings for applications pushed to Application Service.

Maximum File Upload Size (MB) (min: 1024, max: 2048) *

1024



Default App Memory (MB) (min: 64, max: 2048) *

1024

Memory allocated to a newly-pushed app when no value is specified at `cf push`.

Default App Memory Quota per Org (MB) (min: 10240, max: 102400) *

10240

Maximum Disk Quota per App (MB) (min: 512, max: 20480) *

2048

Default Disk Quota per App (MB) (min: 512, max: 20480) *

1024

Default Service Instances Quota per Org (min: 0, max: 1000) *

100

Staging Timeout (Seconds) *

900

Internal Domain *

apps.internal

Allow Space Developers to manage network policies

Save

2. Enter the **Maximum File Upload Size (MB)**. This is the maximum size of an application upload.
3. Enter the **Default App Memory (MB)**. This is the amount of RAM allocated by default to a newly pushed application if no value is specified with the cf CLI.
4. Enter the **Default App Memory Quota per Org**. This is the default memory limit for all applications in an org. The specified limit only applies to the first installation of PAS. After the initial installation, operators can use the cf CLI to change the default value.
5. Enter the **Maximum Disk Quota per App (MB)**. This is the maximum amount of disk allowed per application.

Note: If you allow developers to push large applications, PAS may have trouble placing them on Cells. Additionally, in the event of a system upgrade or an outage that causes a rolling deploy, larger applications may not successfully re-deploy if there is insufficient disk capacity. Monitor your deployment to ensure your Cells have sufficient disk to run your applications.

6. Enter the **Default Disk Quota per App (MB)**. This is the amount of disk allocated by default to a newly pushed application if no value is specified with the cf CLI.
7. Enter the **Default Service Instances Quota per Org**. The specified limit only applies to the first installation of PAS. After the initial installation, operators can use the cf CLI to change the default value.
8. Enter the **Staging Timeout (Seconds)**. When you stage an application droplet with the Cloud Controller, the server times out after the number of seconds you specify in this field.
9. For **Internal Domain**, enter a domain that apps use for internal DNS service discovery. If you specify this domain using `cf push -d`, other PAS apps can reach the pushed app at `APP-NAME.INTERNAL-DOMAIN`. This value defaults to `apps.internal`.

10. Select the **Allow Space Developers to manage network policies** checkbox to permit developers to manage their own network policies for their applications.

11. Click **Save**.

Configure Application Security Groups

Setting appropriate [Application Security Groups](#) is critical for a secure deployment. Type in the box to acknowledge that once the Pivotal Application Service (PAS) deployment completes, you will review and set the appropriate application security groups. See [Restricting App Access to Internal PCF Components](#) for instructions.

Configure Authentication and Enterprise SSO

1. Select **Authentication and Enterprise SSO**. The following image shows the **Authentication and Enterprise SSO** page.

Configure your user store access, which can be an internal user store (managed by Cloud Foundry's UAA) or an external user store (LDAP or SAML). You can also adjust the lifetimes of authentication tokens.

Configure your UAA user account store with either internal or external authentication mechanisms*

Internal UAA (provided by Elastic Runtime; configure your password policy below)

Minimum Password Length *

Minimum Uppercase Characters Required for Password *

Minimum Lowercase Characters Required for Password *

Minimum Numerical Digits Required for Password *

Minimum Special Characters Required for Password *

Maximum Password Entry Attempts Allowed *

2. To authenticate user sign-ons, your deployment can use one of three types of user database: the UAA server's internal user store, an external SAML identity provider, or an external LDAP server.
 - To use the internal UAA, select the **Internal** option and follow the instructions in the [Configuring UAA Password Policy](#) topic to configure your password policy.
 - To connect to an external identity provider through SAML, scroll down to select the **SAML Identity Provider** option and follow the instructions in the [Configuring PCF for SAML](#) section of the *Configuring Authentication and Enterprise SSO for Pivotal Application Service (PAS)* topic.
 - To connect to an external LDAP server, scroll down to select the **LDAP Server** option and follow the instructions in the [Configuring LDAP](#) section of the *Configuring Authentication and Enterprise SSO for PAS* topic.

3. Click **Save**.

Configure UAA

- Select **UAU**. The following image shows the **UAU** page.

Choose the location of your UAA database *

PAS database (configured on the Databases pane)

Other external database

Hostname *

TCP Port *

Username *

Password *

CA Certificate

- Under **Choose the location of your UAA database** select one of the following:

- PAS database (configured on the Databases pane)**: Use the same database server that other PAS components use. This system database is configured in the **Databases** pane, and it can be either internal or external.
- Other external database** - Use a separate, dedicated database server for UAA.

⚠ **warning:** Protect whichever database you use in your deployment with a password.

💡 **Note:** For GCP installations, Pivotal recommends using an external database on Google Cloud SQL.

- (Optional) If you selected **Other external database**, complete the fields as follows. Each field includes additional guidance for specific IaaSes and installation methods.

- For **Hostname**, enter the hostname of the database server.
 - AWS Terraform**: Enter the value of `rds_address` in your Terraform output.
 - GCP Terraform**: Enter the value of `sql_db_ip` from your Terraform output.
- For **TCP Port**, enter the port of the database server.
 - AWS Terraform**: Enter the value of `rds_port` in the Terraform output.
 - GCP and GCP Terraform**: Enter `3306`.
- For **Username**, specify a unique username that can access this specific database on the database server.
 - AWS Terraform**: Enter the value of `rds_username` from your Terraform output.
 - GCP Terraform**: Enter the value of `pas_sql_username` from your Terraform output.
- For **Password**, specify a password for the provided username.
 - AWS Terraform**: Enter the value of `rds_password` from your Terraform output.
 - GCP Terraform**: Enter the value of `pas_sql_password` from your Terraform output.
- For **CA Certificate**, enter a certificate to use for encrypting traffic to and from the database.

Note: The CA Certificate field only works if your external database hostname matches a name specified in the certificate. This is not true with GCP CloudSQL.

3. (Optional) Under **JWT Issuer URI**, enter the URI that UAA uses as the issuer when generating tokens.
4. Under **SAML Service Provider Credentials**, enter a certificate and private key to be used by UAA as a SAML Service Provider for signing outgoing SAML authentication requests. You can provide an existing certificate and private key from your trusted Certificate Authority or generate a certificate. The following domain must be associated with the certificate: `*.login.YOUR-SYSTEM-DOMAIN`.

Note: The Pivotal Single Sign-On Service and Pivotal Spring Cloud Services tiles require the `*.login.YOUR-SYSTEM-DOMAIN`.
5. If the private key specified under **Service Provider Credentials** is password-protected, enter the password under **SAML Service Provider Key Password**.
6. (Optional) To override the default value, enter a custom SAML Entity ID in the **SAML Entity ID Override** field. By default, the SAML Entity ID is `http://login.YOUR-SYSTEM-DOMAIN` where `YOUR-SYSTEM-DOMAIN` is set in the **Domains > System Domain** field.
7. For **Signature Algorithm**, choose an algorithm from the dropdown to use for signed requests and assertions. The default value is `SHA256`.
8. (Optional) In the **Apps Manager Access Token Lifetime**, **Apps Manager Refresh Token Lifetime**, **Cloud Foundry CLI Access Token Lifetime**, and **Cloud Foundry CLI Refresh Token Lifetime** fields, change the lifetimes of tokens granted for Apps Manager and Cloud Foundry Command Line Interface (cf CLI) login access and refresh. Most deployments use the defaults.
9. (Optional) In the **Global Login Session Max Timeout** and **Global Login Session Idle Timeout** fields, change the maximum number of seconds before a global login times out. These fields apply to the following:
 - **Default zone sessions:** Sessions in Apps Manager, PCF Metrics, and other web UIs that use the UAA default zones
 - **Identity zone sessions:** Sessions in apps that use a UAA identity zone, such as a Single Sign-On service plan
10. (Optional) Customize the text prompts used for username and password from the cf CLI and Apps Manager login popup by entering values for **Customize Username Label (on login page)** and **Customize Password Label (on login page)**.
11. (Optional) The **Proxy IPs Regular Expression** field contains a pipe-delimited set of regular expressions that UAA considers to be reverse proxy IP addresses. UAA respects the `x-forwarded-for` and `x-forwarded-proto` headers coming from IP addresses that match these regular expressions. To configure UAA to respond properly to Gorouter or HAProxy requests coming from a public IP address, append a regular expression or regular expressions to match the public IP address.
12. Click **Save**.

Configure CredHub

Note: Enabling CredHub is not required. However, you cannot leave the fields under **Encryption Keys** blank. If you do not intend to use CredHub, enter any text in the **Name** and **Key** fields as placeholder values.

1. Select **CredHub**.
2. Select the location of your CredHub database. PAS includes this CredHub database for services to store their service instance credentials.
 - **PAS database (configured on the Databases pane):** If you select this option, CredHub uses the same database as other PAS components, whether the database is internal or external.

Note: If you deploy PAS on GCP, you must select **PAS database (configured on the Databases pane)**. For more information, see [CredHub Database Cannot be External on GCP](#).

 - **Other external database:** If you select this option, enter the following:

Field	Instructions
Hostname	Enter the IP address of your database server.
TCP Port	Enter the port of your database server, such as <code>3306</code> .
Username	Enter a unique username that can access your CredHub database on the database server.
Password	Enter the password for the provided username.

Database CA Certificate

This certificate is used when encrypting traffic to and from the database.



Note: If you deploy PAS on AWS using Terraform, you can locate the following values in your Terraform output:

- **Hostname:** `rds_address`
- **TCP Port:** `rds_port`
- **Username:** `rds_username`
- **Password:** `rds_password`

3. Under **Encryption Keys**, specify one or more keys to use for encrypting and decrypting the values stored in the CredHub database.

- If you plan to use internal encryption, configure the following values:

Field	Instructions
Name	Enter any key name.
Provider	Select Internal .
Key	Enter a randomly generated value that is at least 20 characters long. This key is used for encrypting all data.
Primary	This checkbox is used for marking the key you specified above as the primary encryption key. You must mark one key as Primary . Do not mark more than one key as Primary .

- If you plan to use a hardware security module (HSM) as your encryption provider, configure the following values:

Field	Instructions
Name	Enter a key name that already exists on your HSM or a new key name. For each new key name, CredHub generates a key in HSM Provider Partition that you configure below.
Provider	Select HSM .
Key	Enter a placeholder value under Key . CredHub does not use this key for encryption. However, you cannot leave the Key field blank.
Primary	This checkbox is used for marking the key you specified above as the primary encryption key. You must mark one key as Primary . Do not mark more than one key as Primary .



Note: For information about using additional keys for key rotation, see the [Rotating Runtime CredHub Encryption Keystopic](#).

4. (Optional) To configure CredHub to use an HSM, complete the following fields:

Field	Instructions
HSM Provider Partition	Enter the name of the HSM provider partition.
HSM Provider Partition Password	Enter a password to use to access the HSM provider partition.
HSM Provider Client Certificate	Enter the client certificate for the HSM. For more information, see Create and Register HSM Clients in the <i>Preparing CredHub HSMs for Configuration</i> topic.

- In the **HSM Provider Servers** section, click **Add** to add an HSM server. You can add multiple HSM servers. For each HSM server, complete the following fields:

Field	Instructions
Host Address	Enter the hostname or IP address of the HSM server.
Port	Enter the port of the HSM server. If you do not know your port address, enter <code>1792</code> .
Partition Serial Number	Enter the serial number of the HSM partition.
HSM Certificate	Enter the certificate for the HSM server. The HSM presents this certificate to CredHub to establish a two-way TLS connection.

5. If your deployment uses any PCF services that support storing service instance credentials in CredHub and you want to enable this feature, select the **Secure Service Instance Credentials** checkbox.

6. Click **Save**.

7. Select the **Resource Config** pane.
8. Under the **Job** column of the **CredHub** row, ensure that the number of instances is set to **2**. This is the minimum instance count required for high availability.
9. Click **Save**.

For more information about using CredHub for securing service instance credentials, see [Securing Service Instance Credentials with Runtime CredHub](#).

Configure System Databases

You can configure PAS to use an internal MySQL database provided with PCF, or you can configure an external database provider for the databases required by PAS.

Note: If you are performing an upgrade, do not modify your existing internal database configuration or you may lose data. You must migrate your existing data first before changing the configuration. See [Upgrading Pivotal Cloud Foundry](#) for additional upgrade information.

Internal Database Configuration

Note: For Runtime CredHub to work on GCP installations, you must use internal MySQL. See [CredHub Database Cannot be External on GCP](#).

Note: For GCP installations, Pivotal recommends selecting **External** and using Google Cloud SQL. Only use internal MySQL for non-production or test installations on GCP.

If you want to use internal databases for your deployment, do the following:

1. Select **Databases**.
2. Select **Internal Databases - MySQL - Percona XtraDB Cluster**.
3. Click **Save**.

Then proceed to [Configure Internal MySQL](#) to configure high availability for your internal MySQL databases.

External System Database Configuration

Note: If you use external MySQL, you cannot use Runtime CredHub. See [CredHub Database Cannot be External on GCP](#).

1. Ensure that you have a database instance with the following databases created. The steps vary depending on your database type. For an example procedure, see [Creating Databases for PAS](#).

- o account
- o app_usage_service
- o autoscale
- o credhub
- o diego
- o locket
- o networkpolicyserver
- o nfsvolume
- o notifications
- o routing
- o silk
- o uaa

2. In PAS, select **Databases**.
3. Select the **External Databases** option.

 **Note:** If you configure external databases, you cannot configure an internal database in the **UAA** pane.

- For **Hostname**, enter the hostname of the database server. If you are installing PCF using Terraform, this value corresponds to the following variable:

- AWS Terraform:** `rds_address`
- GCP Terraform:** `sql_db_ip`

- For **TCP Port**, enter the port of the database server.

- If you are using GCP CloudSQL, enter `3306`.
- If you are installing PCF on AWS using Terraform, enter the value for `rds_port`.

- Each component that requires a relational database has two corresponding fields: one for the database username and one for the database password. For each set of fields, specify a unique username that can access this specific database on the database server and a password for the provided username. If you are installing PCF using Terraform, these values correspond to the following variables:

- AWS Terraform:** `rds_username` and `rds_password`
- GCP Terraform:** `pas_sql_username` and `pas_sql_password`

 **Note:** Ensure that the `networkpolicyserver` database user has the `ALL PRIVILEGES` permission.

- (Optional) If you want to enable TLS for your external databases, paste your Certificate Authority (CA) certificate in the **Database CA Certificate** field.

 **Note:** TLS is not currently supported for databases that do not include a matching hostname in their server certificate, e.g. GCP. For more information, see [Connection Options for External Applications](#) in the GCP documentation.

- Click **Save**.

(Optional) Configure Internal MySQL

 **Note:** You only need to configure this section if you have selected one of the **Internal Databases - MySQL** options in the **Databases** section.

- Select **Internal MySQL**.
- In the **Replication canary time period** field, leave the default of 30 seconds or modify the value based on the needs of your deployment. Lower numbers cause the canary to run more frequently, which means that the canary reacts more quickly to replication failure but adds load to the database.
- In the **Replication canary read delay** field, leave the default of 20 seconds or modify the value based on the needs of your deployment. This field configures how long the canary waits, in seconds, before verifying that data is replicating across each MySQL node. Clusters under heavy load can experience a small replication lag as write-sets are committed across the nodes.
- (Required): In the **E-mail address** field, enter the email address where the MySQL service sends alerts when the cluster experiences a replication issue or when a node is not allowed to auto-rejoin the cluster.
- To prohibit the creation of command line history files on the MySQL nodes, disable the **Allow Command History** checkbox.
- To allow the admin and roadmin to connect from any remote host, enable the **Allow Remote Admin Access** checkbox. When the checkbox is disabled, admins must `bosh ssh` into each MySQL VM to connect as the MySQL super user.

 **Note:** Network configuration and Application Security Groups restrictions may still limit a client's ability to establish a connection with the databases.

- For **Cluster Probe Timeout**, enter the maximum amount of time, in seconds, that a new node will search for existing cluster nodes. If left blank, the default value is 10 seconds.
- For **Max Connections**, enter the maximum number of connections allowed to the database. If left blank, the default value is 1500.
- If you want to log audit events for internal MySQL, select **Enable server activity logging** under **Server Activity Logging**.

- a. For the **Event types** field, you can enter the events you want the MySQL service to log. By default, this field includes `connect` and `query`, which tracks who connects to the system and what queries are processed.

10. Enter values for the following fields:

- o **Load Balancer Healthy Threshold:** Specifies the amount of time, in seconds, to wait until declaring the MySQL Proxy instance started. This allows an external load balancer time to register the instance as healthy.
- o **Load Balancer Unhealthy Threshold:** Specifies the amount of time, in seconds, that the MySQL Proxy continues to accept connections before shutting down. During this period, the Healthcheck reports as unhealthy to cause load balancers to fail over to other proxies. You must enter a value greater than or equal to the maximum time it takes your load balancer to consider a proxy instance unhealthy, given repeated failed healthchecks.

11. If you want to enable the MySQL interruptor feature, select the checkbox to **Prevent node auto re-join**. This feature stops all writes to the MySQL database if it notices an inconsistency in the dataset between the nodes. For more information, see the [Interruptor](#) section in the MySQL for PCF documentation.

12. Click **Save**.

For more information on how to monitor the node health of your MySQL Proxy instances, see [Using the MySQL Proxy](#).

Configure File Storage

1. In the PAS tile, select **File Storage**.
2. For **Max Valid Packages per App**, Pivotal recommends leaving the default value of `5`. However, you may want to lower the value if you have strict storage requirements and want to use less disk space.
3. For **Max Staged Droplets per App**, Pivotal recommends leaving the default value of `5`. However, you may want to lower the value if you have strict storage requirements and want to use less disk space.
4. For **Configure your Cloud Controller's filesystem**, see [Configuring File Storage for PAS](#).

(Optional) Configure System Logging

You can configure system logging in PAS to forward log messages from PAS component VMs to an external service. Pivotal recommends forwarding logs to an external service for use in troubleshooting.

Note: The following instructions explain how to configure system logging for PAS component VMs. To forward logs from PCF tiles to an external service, you must also configure system logging in each tile. See the documentation for the given tiles for information about configuring system logging.

To configure system logging in PAS, do the following:

1. In the PAS **Settings** tab, select the **System Logging** pane. The following image shows the **System Logging** pane.

Optional configuration for rsyslog to forward platform component logs to an external service. If you do not fill these fields, platform logs will not be forwarded but will remain available on the component VMs and for download via Ops Manager.

Address

Port

Transport Protocol

Encrypt syslog using TLS?*

- No
- Yes

Syslog Drain Buffer Size (# of messages) *

Include container metrics in SysLog Drains

Enable Cloud Controller security event logging

Use TCP for file forwarding local transport

Don't Forward Debug Logs

Custom rsyslog Configuration

Save

2. For **Address**, enter the IP address of the syslog server.

3. For **Port**, enter the port of the syslog server. The default port for a syslog server is **514**.

Note: The host must be reachable from the PAS network and accept UDP or TCP connections. Ensure the syslog server listens on external interfaces.

4. For **Transport Protocol**, select a transport protocol for log forwarding.

5. For **Encrypt syslog using TLS?**, select **Yes** to use TLS encryption when forwarding logs to a remote server.

- a. For **Permitted Peer**, enter either the name or SHA1 fingerprint of the remote peer.
- b. For **TLS CA Certificate**, enter the TLS CA certificate for the remote server.

6. For **Syslog Drain Buffer Size**, enter the number of messages from the Loggregator Agent that the Doppler server can store before it begins to drop messages. See the [Loggregator Guide for Cloud Foundry Operators](#) topic for more details.

7. Disable the **Include container metrics in Syslog Drains** checkbox to prevent the [CF Drain CLI plugin](#) from including app container metrics in syslog drains. This feature is enabled by default.

8. Enable **Enable agent-based syslog egress for app logs** to use agent-based syslog. Agent-based syslog provides an alternative to scalable-syslog architecture. For more information about agent-based syslog, see [Loggregator Architecture](#).

 **Note:** Selecting this checkbox disables the Syslog Adapter and Syslog Scheduler VMs to prevent log duplication.

9. Enable the **Enable Cloud Controller security event logging** checkbox to include security events in the log stream. This feature logs all API requests, including the endpoint, user, source IP address, and request result, in the Common Event Format (CEF).
10. Enable the **Use TCP for file forwarding local transport** checkbox to transmit logs over TCP. This prevents log truncation, but may cause performance issues.
11. Disable the **Don't Forward Debug Logs** checkbox to forward DEBUG syslog messages to an external service. This checkbox is enabled by default.

 **Note:** Some PAS components generate a high volume of DEBUG syslog messages. Enabling the **Don't Forward Debug Logs** checkbox prevents PAS components from forwarding the DEBUG syslog messages to external services. However, PAS still writes the messages to the local disk.

12. For **Custom syslog Configuration**, enter a custom syslog rule. For more information about adding custom syslog rules, see [Customizing Syslog Rules](#).
13. Select **Enable System Metrics** to emit system-level metrics about all VMs in the deployment. For a list of the VM metrics that the System Metric Agent emits, see [VM Metrics](#) in GitHub.
14. Click **Save**.

To configure Ops Manager for system logging, see the [Settings](#) section in the *Using the Ops Manager Interface* topic.

(Optional) Configure Custom Branding and Apps Manager

This section describes how to configure **Custom Branding** and **Apps Manager** to customize the appearance and functionality of Apps Manager. For more information about the **Custom Branding** configuration settings, see [Custom Branding Apps Manager](#).

1. Select **Custom Branding**. The following image shows the **Custom Branding** page.

Customize colors, images, and text for Apps Manager and the Cloud Foundry login portal.

Company Name



Accent Color

Main Logo (PNGs only)



Square Logo/Favicon (PNGs only)



Footer Text

Defaults to 'Pivotal Software Inc. All rights reserved.'

Add

Footer Links

You may configure up to three links in the Apps Manager footer

Classification Header/Footer Background Color

Classification Header/Footer Text Color

Classification Header Content



Classification Footer Content



Save

2. Click **Save** to save your settings in this section.
3. Select **Apps Manager**. The following image shows the **Configure Apps Manager** page.

Configure Apps Manager

Enable Invitations *

Display Marketplace Service Plan Prices

Supported currencies as JSON *

{ "usd": "\$", "eur": "€" }



Product Name

Marketplace Name

Customize Sidebar Links *

Add

You may configure up to 10 links in the Apps Manager sidebar.

▶ Marketplace



▶ Docs



▶ Tools



Apps Manager Memory Usage (MB) (min: 128)

Invitations Memory Usage (MB) (min: 256)

Memory limit with which to deploy the Invitations app. Leave blank to use the default value.

Apps Manager Polling Interval *

30

App Details Polling Interval *

10

Multi-Foundation Configuration (BETA)

Save

4. Select **Enable Invitations** to enable invitations in Apps Manager. Space Managers can invite new users for a given space, Org Managers can invite new users for a given org, and Admins can invite new users across all orgs and spaces. See the [Inviting New Users](#) section of the *Managing User Roles with Apps Manager* topic for more information.

5. Select **Display Marketplace Service Plan Prices** to display the prices for your services plans in the Marketplace.

6. Enter the **Supported currencies as JSON** to appear in the Marketplace. Use the format `{ "CURRENCY-CODE": "SYMBOL" }`. This defaults to `{ "usd": "$", "eur": "€" }`.
7. Use **Product Name**, **Marketplace Name**, and **Customize Sidebar Links** to configure page names and sidebar links in the **Apps Manager** and **Marketplace** pages.
8. The **Apps Manager Memory Usage (MB)** field sets the memory limit with which to deploy the Apps Manager app. Use this field to increase the memory limit if the app fails to start with an `out of memory` error.
9. The **Invitations Memory Usage (MB)** field sets the memory limit with which to deploy the Invitations app. Use this field to increase the memory limit if the app fails to start with an `out of memory` error.
10. The **Apps Manager Polling Interval** field provides a temporary fix if Apps Manager usage degrades Cloud Controller response times. In this case, you can use this field to reduce the load on the Cloud Controller and ensure Apps Manager remains available while you troubleshoot the Cloud Controller. Pivotal recommends that you do not keep this field modified as a long term fix because it can degrade Apps Manager performance. You can optionally do the following:
 - Increase the polling interval above the default of `30` seconds.

Note: If you enter a value between `0` and `30`, the field is automatically set to `30`.
 - Disable polling by entering `0`. This stops Apps Manager from refreshing data automatically, but users can update displayed data by reloading Apps Manager manually.
11. The **App Details Polling Interval** field provides an additional way to reduce the load on the Cloud Controller when the **Apps Manager Polling Interval** field is not sufficient. This field controls the rate at which Apps Manager polls for data when a user views the **Overview** page of an app. Pivotal recommends that you do not keep this field modified as a long term fix because it can degrade Apps Manager performance. You can optionally do the following:
 - Increase the polling interval above the default of `10` seconds.

Note: If you enter a value between `0` and `30`, the field is automatically set to `10`.
 - Disable polling by entering `0`. This stops Apps Manager from refreshing data automatically, but users can update displayed data by reloading Apps Manager manually.
12. For **Multi-Foundation Configuration (BETA)**, enter a JSON object to enable multi-foundation support for Apps Manager. Enabling multi-foundation support allows you to manage orgs, spaces, apps, and service instances from multiple PCF foundations from a single Apps Manager interface. For more information, see [Configuring Multi-Foundation Support in Apps Manager](#).
13. Click **Save** to save your settings in this section.

(Optional) Configure Email Notifications

PAS uses SMTP to send invitations and confirmations to Apps Manager users. You must complete the **Email Notifications** page if you want to enable end-user self-registration.

1. Select **Email Notifications**. The following image shows the **Email Notifications** page.

Configure Simple Mail Transfer Protocol for the Notifications application to send email notifications about your deployment. This application is deployed as an errand in Elastic Runtime. If you do not need this service, leave this section blank and disable the Notifications and Notifications UI errands.

From Email

Address of SMTP Server

Port of SMTP Server

SMTP Server Credentials

[Change](#)

SMTP Enable Automatic STARTTLS

SMTP Authentication Mechanism*

SMTP CRAMMD5 secret

Save

2. Enter your reply-to and SMTP email information.

Note: For GCP, you must use port `2525`. Ports `25` and `587` are not allowed on GCP Compute Engine.

3. Verify your authentication requirements with your email administrator and use the **SMTP Authentication Mechanism** dropdown to select `None`, `Plain`, or `CRAMMD5`. If you have no SMTP authentication requirements, select `None`.
4. If you selected `CRAMMD5` as your authentication mechanism, enter a secret in the **SMTP CRAMMD5 secret** field.
5. Click **Save**.

Note: If you do not configure the SMTP settings using this form, the administrator must create orgs and users using the cf CLI. See [Creating and Managing Users with the cf CLI](#) for more information.

(Optional) Configure App Autoscaler

To use App Autoscaler, you must create an instance of the service and bind it to an app. To create an instance of App Autoscaler and bind it to an app, see [Set Up App Autoscaler](#) in the *Scaling an Application Using App Autoscaler* topic.

1. Click **App Autoscaler**. The following image shows the **Configure App Autoscaler** page.

Configure the App Autoscaler

Autoscaler Instance Count (min: 1) *

Autoscaler API Instance Count (min: 1) *

Metric Collection Interval (min: 60, max: 3600) (min: 60, max: 3600) *

Scaling Interval (min: 15, max: 120) (min: 15, max: 120) *

Verbose Logging

Disable API Connection Pooling

Enable Notifications

Save

2. Review the following settings:

- **Autoscaler Instance Count:** How many instances of the App Autoscaler service you want to deploy. The default value is . For high availability, set this number to or higher. You should set the instance count to an odd number to avoid split-brain scenarios during leadership elections. Larger environments may require more instances than the default number.
- **Autoscaler API Instance Count:** How many instances of the App Autoscaler API you want to deploy. The default value is . Larger environments may require more instances than the default number.
- **Metric Collection Interval:** How many seconds of data collection you want App Autoscaler to evaluate when making scaling decisions. The minimum interval is 60 seconds, and the maximum interval is 3600 seconds. The default value is . Increase this number if the metrics you use in your scaling rules are emitted less frequently than the existing Metric Collection Interval.
- **Scaling Interval:** How frequently App Autoscaler evaluates an app for scaling. The minimum interval is 15 seconds, and the maximum interval is 120 seconds. The default value is .
- **Verbose Logging (checkbox):** Enables verbose logging for App Autoscaler. Verbose logging is disabled by default. Select this checkbox to see more detailed logs. Verbose logs show specific reasons why App Autoscaler scaled the app, including information about minimum and maximum instance limits, App Autoscaler's status. For more information about App Autoscaler logs, see [App Autoscaler Events and Notifications](#).
- **Disable API Connection Pooling:** If you select this option, the Autoscaler API does not reuse HTTP connections. This may be necessary if your [Frontend Idle Timeout](#) for Gorouter is set to a low value, such as 1 second.

3. Click **Save**.

Configure Cloud Controller

1. Click **Cloud Controller**. The following image shows the **Configure the Cloud Controller** page.

1

Configure the Cloud Controller

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

The Smoke Tests errand runs basic functionality tests against your Pivotal Application Service (PAS) deployment after an installation or update. In this section, choose where to run smoke tests. In the **Errands** section, you can choose whether or not to run the Smoke Tests errand.

1. Select **Smoke Tests**. The following image shows the **Smokes Tests** page.

Specify a Cloud Foundry organization and space where smoke tests can run if in the future you delete your Elastic Runtime deployment domains.

Choose where to deploy applications when running the smoke tests *

Temporary space within the system organization (This is deleted after smoke tests finish.)
 Specified org and space (The org and space must have a domain available for routing.)

Organization *

Space *

Domain *

Save

2. If you have a shared apps domain, select **Temporary space within the system organization**, which creates a temporary space within the **system** organization for running smoke tests and deletes the space afterwards. Otherwise, select **Specified org and space** and complete the fields to specify where you want to run smoke tests.
3. Click **Save**.

(Optional) Configure Advanced Features

The **Advanced Features** section of Pivotal Application Service (PAS) includes new functionality that may have certain constraints. Although these features are fully supported, Pivotal recommends caution when using them in production environments.

Diego Cell Memory and Disk Overcommit

If your apps do not use the full allocation of disk space and memory set in the **Resource Config** tab, you might want use this feature. These fields control the amount to overcommit disk and memory resources to each Diego Cell VM.

For example, you might want to use the overcommit if your apps use a small amount of disk and memory capacity compared to the amounts set in the **Resource Config** settings for **Diego Cell**.

Note: Due to the risk of app failure and the deployment-specific nature of disk and memory use, Pivotal has no recommendation about how much, if any, memory or disk space to overcommit.

To enable overcommit, do the following:

1. Select **Advanced Features**.
2. Enter the total desired amount of Diego cell memory value in the **Cell Memory Capacity (MB)** field. Refer to the **Diego Cell** row in the **Resource**

Config tab for the current Cell memory capacity settings that this field overrides.

3. Enter the total desired amount of Diego cell disk capacity value in the **Cell Disk Capacity (MB)** field. Refer to the **Diego Cell** row in the **Resource Config** tab for the current Cell disk capacity settings that this field overrides.
4. Click **Save**.

 **Note:** Entries made to each of these two fields set the total amount of resources allocated, not the overage.

Whitelist for Non-RFC-1918 Private Networks

Some private networks require extra configuration so that internal file storage (WebDAV) can communicate with other PCF processes.

The **Whitelist for non-RFC-1918 Private Networks** field is provided for deployments that use a non-RFC 1918 private network. This is typically a private network other than `10.0.0.0/8`, `172.16.0.0/12`, or `192.168.0.0/16`.

Most PCF deployments do not require any modifications to this field.

To add your private network to the whitelist, do the following:

1. Select **Advanced Features**.
2. Append a new `allow` rule to the existing contents of the **Whitelist for non-RFC-1918 Private Networks** field. Include the word `allow`, the network CIDR range to allow, and a semi-colon (`:`) at the end. For example: `allow 172.99.0.0/24;`
3. Click **Save**.

CF CLI Connection Timeout

The **CF CLI Connection Timeout** field allows you to override the default five second timeout of the Cloud Foundry Command Line Interface (cf CLI) used within your PCF deployment. This timeout affects the cf CLI command used to push PAS errand apps such as Notifications, Autoscaler, and Apps Manager.

Set the value of this field to a higher value, in seconds, if you are experiencing domain name resolution timeouts when pushing errands in PAS.

To modify the value of the **CF CLI Connection Timeout**, perform the following steps:

1. Select **Advanced Features**.
2. Add a value, in seconds, to the **CF CLI Connection Timeout** field.
3. Click **Save**.

Enable TLS for Internal System Database

You can optionally enable TLS for clients of the internal system database. This feature is in beta, and the checkbox is disabled by default. For more information about the internal system database, see [Managing Internal Databases](#).

To enable TLS for clients of the internal system database, do the following:

1. Select **BETA: Enable TLS for internal system database**.
2. Click **Save**.

Database Connection Limits

You can configure the maximum number of concurrent database connections that diego and container networking components can have. Use the field beginning with **Maximum number of open connections...** for a given component. The placeholder values for each field are the default values.

When there are not enough connections available, such as during a time of heavy load, components may experience degraded performance and sometimes failure. To resolve or prevent this, you can increase and fine-tune database connection limits for the component.

⚠ warning: Decreasing the value of this field for a component may affect the amount of time it takes for it to respond to requests.

Rolling App Deployments

You can optionally disable rolling app deployments. For more information, see the [Rolling App Deployments \(Beta\)](#) topic.

To disable zero downtime app deployments, do the following:

1. Select **Disable Zero Downtime App Deployments**.
2. Click **Save**.

(Optional) Configure Metric Registrar

The PAS tile includes a **Metric Registrar** configuration pane. You can configure the Metric Registrar from this pane in the following ways:

- Enable / Disable
- Edit default scraping interval
- Add blacklisted tags

Enable and Disable the Metric Registrar

The PAS tile deploys the Metric Registrar by default.

To disable the Metric Registrar, do the following:

1. In the PAS tile, click **Metric Registrar**.
2. Clear the **Enable Metric Registrar** checkbox.

Edit Default Scraping Interval

The scraping interval defines how often the Metric Registrar polls custom metric endpoints. The default is 35 seconds.

To edit the Metric Registrar scraping interval, do the following:

1. In the PAS tile, click **Metric Registrar**.
2. Edit the **Endpoint Scraping Interval** field.

Add Blacklisted Tags

To prevent the Metric Registrar from consuming the value of a metric or event tag, you can add the tag to the **Blacklisted Tags** field. For example, if you tag your metrics with a `customer_id`, you may want to add `customer_id` to the list of blacklisted tags. The field accepts a comma-separated list.

By default, the following tags are blacklisted to prevent interference with other products like PCF Metrics that use and rely on such tags.

- `deployment`
- `job`
- `index`
- `id`

To prevent the Metric Registrar from consuming the value of a metric or event tag, do the following:

1. In the PAS tile, click **Metric Registrar**.
2. Add the desired tag to the **Blacklisted Tags** field.

Configure Errands

Errands are scripts that Ops Manager runs automatically when it installs or uninstalls a product, such as a new version of Pivotal Application Service (PAS). There are two types of errands: *post-deploy errands* run after the product is installed, and *pre-delete errands* run before the product is uninstalled.

By default, Ops Manager always runs all errands.

The PAS tile **Errands** pane lets you change these run rules. For each errand, you can select **On** to run it always or **Off** to never run it.

For more information about how Ops Manager manages errands, see the [Managing Errands in Ops Manager](#) topic.

Note: Several errands, such as App Autoscaler and Notifications, deploy apps that provide services for your deployment. When one of these apps is running, selecting **Off** for the corresponding errand on a subsequent installation does not stop the app.

- **Smoke Test Errand** verifies that your deployment can do the following:
 - Push, scale, and delete apps
 - Create and delete orgs and spaces
- **Usage Service Errand** deploys the Pivotal Usage Service application, which Apps Manager depends on.
- **Apps Manager Errand** deploys Apps Manager, a dashboard for managing apps, services, orgs, users, and spaces. Until you deploy Apps Manager, you must perform these functions through the cf CLI. After Apps Manager has been deployed, Pivotal recommends setting this errand to **Off** for subsequent PAS deployments. For more information about Apps Manager, see the [Getting Started with the Apps Manager](#) topic.
- **Notifications Errand** deploys an API for sending email notifications to your PCF platform users.

Note: The Notifications app requires that you [configure SMTP](#) with a username and password, even if you set the value of **SMTP Authentication Mechanism** to `none`.

- **Notifications UI Errand** deploys a dashboard for users to manage notification subscriptions.
- **App Autoscaler Errand** pushes the App Autoscaler app, which enables you to configure your apps to automatically scale in response to changes in their usage load. See the [Scaling an Application Using Autoscaler](#) topic for more information.
- **NFS Broker Errand** pushes the NFS Broker app, which supports NFS Volume Services for PAS. For more information, see [Enabling NFS Volume Services](#).
- **Metric Registrar Smoke Test Errand** verifies the Metric Registrar can access custom metrics emitted by an app and convert them into Loggregator metrics.

Note: The Metric Registrar Smoke Test errand runs only if the Metric Registrar is enabled. For more information about configuring the Metric Registrar, see [Configuring the Metric Registrar](#).

- **SMB Broker Application Errand** pushes the SMB Broker app, which supports SMB Volume Services for PAS. For more information, see [Enabling SMB Volume Services](#).

Configure Resources

Configure Load Balancing

In the Resource Config pane, you must associate load balancers with the VMs in your deployment to enable traffic. See [Configure Load Balancing for PAS](#).

Note: This step does not apply to vSphere deployments.

(Optional) Scale Down and Disable Resources

Note: The Resource Config pane has fewer VMs if you are installing the [Small Footprint Runtime](#).

Note: The Small Footprint Runtime does not default to a highly available configuration. It defaults to the minimum configuration. If you want to make the Small Footprint Runtime highly available, scale the **Compute**, **Router**, and **Database** VMs to `3` instances and scale the **Control** VM to

2 instances.

Pivotal Application Service (PAS) defaults to a highly available resource configuration. However, you may need to perform additional procedures to make your deployment highly available. See the [Zero Downtime Deployment and Scaling in CF](#) and the [Scaling Instances in PAS](#) topics for more information.

If you do not want a highly available resource configuration, you must scale down your instances manually by navigating to the **Resource Config** section and using the dropdowns under **Instances** for each job.

By default, PAS also uses an internal filestore and internal databases. If you configure PAS to use external resources, you can disable the corresponding system-provided resources in Ops Manager to reduce costs and administrative overhead.

To disable specific VMs in Ops Manager, do the following:

1. Click **Resource Config**.
2. If you configured PAS to use an external S3-compatible filestore, enter 0 in **Instances** in the **File Storage** field.
3. If you selected **External** when configuring the UAA, System, and CredHub databases, edit the following fields:
 - **MySQL Proxy**: Enter 0 in **Instances**.
 - **MySQL Server**: Enter 0 in **Instances**.
 - **MySQL Monitor**: Enter 0 in **Instances**.
4. If you disabled TCP routing, enter 0 **Instances** in the **TCP Router** field.
5. If you are not using HAProxy, enter 0 **Instances** in the **HAProxy** field.
6. Click **Save**.

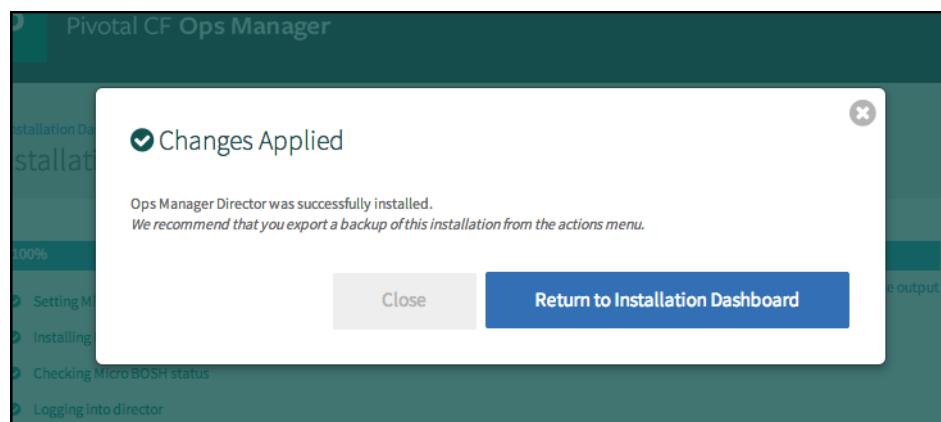
Download Stemcell

This step is only required if your Ops Manager does not already have the stemcell version required by PAS. For more information about importing stemcells, see [Importing and Managing Stemcells](#).

1. Open the [Stemcell product page](#) in the Pivotal Network. You may have to log in.
2. Download the appropriate stemcell version targeted for your IaaS.
3. Navigate to **Stemcell Library** in the **Installation Dashboard**.
4. Click **Import Stemcell** to import the downloaded stemcell **.tgz** file.
5. When prompted, enable the Ops Manager product checkbox to stage your stemcell.
6. Click **Apply Stemcell to Products**.

Complete the PAS Installation

1. Click the **Installation Dashboard** link to return to the Installation Dashboard.
2. Click **Review Pending Changes**, then **Apply Changes**. The install process generally requires a minimum of 90 minutes to complete. The following image shows the Changes Applied window that displays when the installation process successfully completes.



Quick Start PAS Configuration

This topic describes how to minimally configure Pivotal Application Service (PAS) for evaluation or testing purposes. It does not include optional configurations such as external databases or external file storage.

For production deployments, Pivotal recommends following the instructions in the [Configuring PAS](#) topic.

Prerequisites

Before beginning this procedure, ensure that you have successfully completed the steps to prepare your environment for PCF and install and configure the BOSH Director.

Add PAS to Ops Manager

To add PAS to Ops Manager, do the following:

1. If you have not already downloaded PAS, log in to [Pivotal Network](#), and click the PAS tile.
2. From the **Releases** drop-down, select the release to install and choose one of the following:
 - a. Click PAS to download the PAS `.pivotal` file.
 - b. Click PCF Small Footprint Runtime to download the Small Footprint Runtime `.pivotal` file. For more information, see [Getting Started with Small Footprint Runtime](#).
3. Navigate to the Pivotal Cloud Foundry Operations Manager Installation Dashboard.
4. Click **Import a Product** to add your tile to Ops Manager. For more information, refer to the [Adding and Deleting Products](#) topic.
5. Click the PAS tile in the Installation Dashboard.

Configure PAS

To install PAS with minimal configuration, do the following:

1. Do the procedure in the [Assign AZs and Networks](#) section of *Configuring PAS*.
2. Do the procedure in the [Configure Domains](#) section of *Configuring PAS*.
3. Click the **Networking** pane.
4. Under **Certificates and Private Key for HAProxy and Router**, you must provide at least one **Certificate and Private Key** name and certificate key pair for HAProxy and Gorouter. HAProxy and Gorouter are enabled to receive TLS communication by default. You can configure multiple certificates for HAProxy and Gorouter.

 **Note:** When providing custom certificates, enter them in the following order: `wildcard`, `Intermediate`, `CA`. For more information, see [Creating a .pem File for SSL Certificate Installations](#) in the DigiCert documentation.

- a. Click the **Add** button to add a name for the certificate chain and its private key pair. This certificate is the default used by Gorouter and HAProxy. You can either provide a certificate signed by a Certificate Authority (CA) or click on the **Generate RSA Certificate** link to generate a self-signed certificate in Ops Manager.

 **Note:** If you configured Ops Manager Front End without a certificate, you can use this new certificate to complete Ops Manager configuration. To configure your Ops Manager Front End certificate, see *Configure Front End* in [Preparing to Deploy Ops Manager on GCP Manually](#).

 **Note:** Ensure that you add any certificates that you generate in this pane to your infrastructure load balancer.

5. If you are not using SSL encryption or if you are using self-signed certificates, select **Disable SSL certificate verification for this environment**. Selecting this checkbox also disables SSL verification for route services and disables mutual TLS app identity verification.

 **Note:** For production deployments, Pivotal does not recommend disabling SSL certificate verification.

6. Disable **HAProxy forwards requests to Router over TLS**. By default, PAS does not deploy HAProxy.
 7. Click the **Application Security Groups** pane.
Setting appropriate [Application Security Groups](#) is critical for a secure deployment. Type in the box to acknowledge that once the Pivotal Application Service (PAS) deployment completes, you will review and set the appropriate application security groups. See [Restricting App Access to Internal PCF Components](#) for instructions.
 8. Under **SAML Service Provider Credentials**, enter a certificate and private key to be used by UAA as a SAML Service Provider for signing outgoing SAML authentication requests. You can provide an existing certificate and private key from your trusted Certificate Authority or generate a self-signed certificate. The following domain must be associated with the certificate: `*.login.YOUR-SYSTEM-DOMAIN`.
-  **Note:** The Pivotal Single Sign-On Service and Pivotal Spring Cloud Services tiles require the `*.login.YOUR-SYSTEM-DOMAIN`.
9. If the private key specified under **Service Provider Credentials** is password-protected, enter the password under **SAML Service Provider Key Password**.
 10. In the **CredHub** pane, Under **Encryption Keys**, specify one or more keys to use for encrypting and decrypting the values stored in the CredHub database.
 - **Name.** This is the name of the encryption key.
 - **Provider.** This is the provider of the encryption key. If you have configured an HSM provider and HSM servers above, select **HSM**. Otherwise, select **Internal**.
 - **Key.** This key is used for encrypting all data. The key must be at least 20 characters long.
 - **Primary.** This checkbox is used for marking the key you specified above as the primary encryption key. You must mark one key as **Primary**. Do not mark more than one key as **Primary**.
 11. In the **E-mail address** field, enter the email address where the MySQL service sends alerts when the cluster experiences a replication issue or when a node is not allowed to auto-rejoin the cluster.
 12. In the **Resource Config** pane, you must associate load balancers with the VMs in your deployment to enable traffic. See [Configure Load Balancing for PAS](#).

Complete the PAS Installation

To complete the PAS installation, do the following:

1. Click the **Installation Dashboard** link to return to the Installation Dashboard.
2. Click **Review Pending Changes**, then **Apply Changes**.

Configuring File Storage for PAS

This topic provides instructions for configuring file storage for PAS based on your IaaS and installation method. See the section that applies to your use case.

To minimize system downtime, Pivotal recommends using highly resilient and redundant [external filestores](#) for your Pivotal Application Service (PAS) file storage. For more factors to consider when selecting file storage, see the [Considerations for Selecting File Storage in Pivotal Cloud Foundry](#) topic.

Internal File Storage

Internal file storage is only appropriate for small, non-production deployments.

To use the PCF internal filestore, select **Internal WebDAV**, and click **Save**.

AWS

This section describes how to configure file storage for AWS.

 **Note:** If you followed the instructions in [Installing PCF on AWS Manually](#), you created the necessary resources for [external S3-compatible file storage](#).

For production-level PCF deployments on AWS, Pivotal recommends selecting the **External S3-Compatible File Store**. For more information about production-level PCF deployments on AWS, see the [Reference Architecture for Pivotal Cloud Foundry on AWS](#).

To use an external S3-compatible filestore for PAS file storage, perform the following steps:

1. Select the **External S3-Compatible Filestore** option and complete the following fields:

- Enter the **https://** URL Endpoint for your region.
For example, in the **us-west-2** region, enter **https://s3-us-west-2.amazonaws.com/**.
- If you use an AWS instance profile to manage role information for your filestore, select the **S3 AWS with Instance Profile** checkbox. For more information, see [AWS Identity and Access Management](#) in the AWS documentation.



The dialog box contains the following fields:
 S3 AWS with Instance Profile
 Access Key (required if not using S3 AWS with Instance Profile)

 Secret Key (required if not using S3 AWS with Instance Profile)

 Secret
 Cancel

- Alternatively, enter the **Access Key** and **Secret Key** of the **pcf-user** you created when configuring AWS for PCF. If you select the **S3 AWS with Instance Profile** checkbox and also enter an **Access Key** and **Secret Key**, the instance profile will overrule the Access Key and Secret Key.
- From the **S3 Signature Version** dropdown, select **V4 Signature**. For more information about S4 signatures, see [Signing AWS API Requests](#) in the AWS documentation.
- For **Region**, enter the region in which your S3 buckets are located. **us-west-2** is an example of an acceptable value for this field.
- Select **Server-side Encryption** to encrypt the contents of your S3 filestore. This option is only available for AWS S3.
- (Optional) If you selected **Server-side Encryption**, you can also specify a **KMS Key ID**. PAS uses the KMS key to encrypt files uploaded to the blobstore. If you do not provide a KMS Key ID, PAS uses the default AWS key. For more information, see [Protecting Data Using Server-Side Encryption with AWS KMS-Managed Keys \(SSE-KMS\)](#).
- Enter names for your S3 buckets:

Ops Manager Field	Value	Description
Buildpacks Bucket Name	pcf-buildpacks-bucket	This S3 bucket stores app buildpacks.
Droplets Bucket Name	pcf-droplets-bucket	This S3 bucket stores app droplets. Pivotal recommends that you use a unique bucket name for droplets, but you can also use the same name as above.

Packages Bucket Name	pcf-packages-bucket	This S3 bucket stores app packages. Pivotal recommends that you use a unique bucket name for packages, but you can also use the same name as above.
Resources Bucket Name	pcf-resources-bucket	This S3 bucket stores app resources. Pivotal recommends that you use a unique bucket name for app resources, but you can also use the same name as above.

- Configure the following depending on whether your S3 buckets have versioning enabled:

- Versioned S3 buckets:** Enable the **Use versioning for backup and restore** checkbox to archive each bucket to a version.
- Unversioned S3 buckets:** Disable the **Use versioning for backup and restore** checkbox, and enter a backup bucket name for each active bucket. The backup bucket name must be different from the name of the active bucket it backs up.

Use versioning for backup and restore.
(All of the above buckets must have versioning enabled and use the V4 S3 Signature Version).

Backup Region

Backup Buildpacks Bucket Name

Backup Droplets Bucket Name

Backup Packages Bucket Name

If versioning is not enabled or supported, you can configure separate buckets below for backups. More information at <https://docs.pivotal.io/pivotalcf/2-1/customizing/backup-restore/backup-pcf-bbr.html>.

For more information

about setting up external S3 blobstores, see the [Backup and Restore with External Blobstores](#) topic in the Cloud Foundry documentation.

- Click **Save**.

 **Note:** For more information regarding AWS S3 Signatures, see the [Authenticating Requests](#) topic in the AWS documentation.

Azure

This section describes how to configure file storage for Azure.

For production-level PCF deployments on Azure, the recommended selection is Azure Storage. For more information about production-level PCF deployments on Azure, see the [Reference Architecture for Pivotal Cloud Foundry on Azure](#).

For more factors to consider when selecting file storage, see [Considerations for Selecting File Storage in Pivotal Cloud Foundry](#).

To use external Azure file storage for your Pivotal Application Service (PAS) filestore, perform the following steps:

- Select the **External AzureStorage** option.

Configure your Cloud Controller's filesystem*

<input checked="" type="checkbox"/> Authentication and Enterprise SSO <input type="radio"/> UAA <input type="radio"/> CredHub <input checked="" type="checkbox"/> Databases <input type="radio"/> Internal MySQL <input checked="" type="checkbox"/> File Storage <input checked="" type="checkbox"/> System Logging <input checked="" type="checkbox"/> Custom Branding <input checked="" type="checkbox"/> Apps Manager <input checked="" type="checkbox"/> Email Notifications <input checked="" type="checkbox"/> App Autoscaler <input checked="" type="checkbox"/> Cloud Controller <input checked="" type="checkbox"/> Smoke Tests <input checked="" type="checkbox"/> Advanced Features <input checked="" type="checkbox"/> Errands <input checked="" type="checkbox"/> Resource Config	<input type="radio"/> Internal WebDAV (provided by Application Service) <input type="radio"/> External S3-Compatible File Store (if you want to use a service like S3 or Ceph) <input type="radio"/> External Google Cloud Storage with Access Key and Secret Key <input type="radio"/> External Google Cloud Storage with Service Account <input checked="" type="radio"/> External Azure Storage
	Account Name *
	<input type="text"/>
	Access Key *
	<input type="text"/> Secret
	Environment *
	<input type="text"/> AzureCloud
	Buildpacks Container Name *
	<input type="text"/>
	Droplets Container Name *
	<input type="text"/>
	Packages Container Name *
	<input type="text"/>
	Resources Container Name *
	<input type="text"/>
	<input type="checkbox"/> Enable backup and restore (Soft deletes must be enabled for all storage containers listed above).
	Restore from Storage Account
	<input type="text"/>
	Restore using Access Key
	<input type="text"/> Secret

Save

2. To create a new storage account and storage containers for the PAS filestore, perform the following steps.

- o In the Azure Portal, navigate to the **Storage accounts** tab.
- o Click on the plus icon to add a new storage account.
- o In the **Name** field, enter a unique name (all lowercase, 3 to 24 alphanumeric characters) for the storage account.
- o For the **Deployment model**, select **Resource manager**.
- o For **Account kind**, select **General purpose**.
- o For **Performance**, select **Standard**.
- o From the **Replication** dropdown, select **Locally-redundant storage (LRS)**.
- o For **Storage service encryption**, select **Disabled**.
- o From the **Subscription** dropdown, select the subscription where you want to deploy PCF resources.
- o For **Resource group**, select **Use existing** and enter the name of the resource group where you deployed PAS.
- o From **Location** the dropdown, select the **Location** where you are deploying PCF.
- o Click **Create**.
- o After the storage account is created, select the new storage account from the dashboard.
- o Navigate to the **Blob Service** section of the storage account, and then click on **Containers** to create one or more containers in this storage account for buildpacks, droplets, resources, and packages.
- o In **Blob Service**, select **Soft Delete**. In **Soft Delete**, click **Enabled** to enable soft delete in your Azure storage account.



Note: BBR requires that you enable soft delete in your Azure storage account before you enable backup and restore for your Azure blobstores in Ops Manager. You should set a reasonable retention policy to minimize storage costs. For more information on enabling soft delete in your Azure storage account, see the [Azure documentation](#).

- o For each container that you create, set the **Access type** to **Private**.

3. In PAS, enter the name of the storage account you created for **Account Name**.
4. In the **Access Key** field, enter one of the access keys provided for the storage account. To obtain a value for this fields, visit the Azure Portal, navigate to the **Storage accounts** tab and click on **Access keys**.
5. For **Environment**, enter the name of the Azure Cloud environment that contains your storage. This value defaults to AzureCloud.
6. For the **Buildpacks Container Name**, enter the container name for storing your app buildpacks.
7. For **Droplets Container Name**, enter the container name for your app droplet storage. Pivotal recommends that you use a unique container name, but you can use the same container name as the previous step.
8. For **Packages Container Name**, enter the container name for packages. Pivotal recommends that you use a unique container name, but you can use the same container name as the previous step.
9. For **Resources Container Name**, enter the container name for resources. Pivotal recommends that you use a unique container name, but you can use the same container name as the previous step.
10. (Optional) To enable backup and restore for your Azure blobstores in PAS, select the **Enable backup and restore (Soft deletes must be enabled for all storage containers listed above)** checkbox.
11. (Optional) To enable PAS to restore your containers to a different Azure storage account than the account where you take backups, do the following:
 - a. Under **Restore from Storage Account**, enter the name of the Azure storage account you want to restore your containers from. Leave this field blank if restoring to the same storage account where you take backups.
 - b. Under **Restore using Access Key**, enter the access key for the Azure storage account you specified in **Restore from Storage Account**. Leave this field blank if restoring to the same storage account where you take backups.
12. Click **Save**.

 **Note:** To enable backup and restore of your PAS tile that uses S3 compatible blobstore, see [Enable External Blobstore Backups](#).

GCP

This section describes how to configure file storage for GCP. Follow the instructions that correspond to your installation method: [Manual](#) or [Terraform](#).

For production-level PCF deployments on GCP, Pivotal recommends selecting **External Google Cloud Storage**. For more information about production-level PCF deployments on GCP, see the [Reference Architecture for Pivotal Cloud Foundry on GCP](#).

Manual

This section describes how to configure file storage for GCP if you installed PCF manually.

PAS can use Google Cloud Storage (GCS) as its external filestore by using either a GCP interoperable storage access key or your GCS Service Account. Choose the procedure that corresponds to your use case.

PAS can use Google Cloud Storage (GCS) as its external filestore by using either a GCP inter-operable storage access key or your GCS Service Account. See below for how to configure each option.

 **Note:** The Pivotal Application Service for Windows (PASW) tile is incompatible with GCP configured with a GCS file store. If you are deploying PASW in your GCP environment, Pivotal recommends that you select the S3-compatible file store for your environment.

External Google Cloud Storage with Access Key and Secret Key

1. Select the **External Google Cloud Storage with Access Key and Secret Key** option

This section determines where you would like to place your Application Service Cloud Controller's file storage.

Configure your Cloud Controller's filesystem*

- Internal WebDAV (provided by Application Service)
- External S3-Compatible File Store (if you want to use a service like S3 or Ceph)
- External Google Cloud Storage with Access Key and Secret Key

Access Key *

Secret Key *

Secret

Buildpacks Bucket Name *

Bucket for storing app buildpacks.

Droplets Bucket Name *

Packages Bucket Name *

Resources Bucket Name *

- External Google Cloud Storage with Service Account

- External Azure Storage

2. Enter values for **Access Key** and **Secret Key**. To obtain the values for these fields:

- a. In the GCP Console, navigate to the **Storage** tab, then click **Settings**.
- b. Click **Interoperability**.
- c. If necessary, click **Enable interoperability access**. If interoperability access is already enabled, confirm that the default project matches the project where you are installing PCF.

The screenshot shows the Google Cloud Platform Settings page with the Interoperability tab selected. It includes sections for Default project for interoperable access (with 'cf-docs' as the default), Interoperable storage access keys (with a key labeled 'GOO...' and '4Kk...'), and a 'Create a new key' button.

- d. Click **Create a new key**.
- e. Copy and paste the generated values into the corresponding PAS fields. PCF uses these values for authentication when connecting to Google Cloud Storage.
- f. Enter the names of the storage buckets you created in [Preparing to Deploy Ops Manager on GCP Manually](#).
 - **Buildpacks Bucket Name:** MY-PCF-buildpacks
 - **Droplets Bucket Name:** MY-PCF-droplets
 - **Resources Bucket Name:** MY-PCF-resources
 - **Packages Bucket Name:** MY-PCF-packages
- g. Click **Save**.

External Google Cloud Storage with Service Account

1. Select the **External Google Cloud Storage with Service Account** option

Configure your Cloud Controller's filesystem*

Internal WebDAV (provided by Application Service)
 External S3-Compatible File Store (if you want to use a service like S3 or Ceph)
 External Google Cloud Storage with Access Key and Secret Key
 External Google Cloud Storage with Service Account

GCP Project ID *

GCP Service Account Email *

 Your Google Cloud Platform Service Account Email.

GCP Service Account Key JSON *

Buildpacks Bucket Name *

- a. For **GCP Project ID** enter the Project ID on your GCP Console that you want to use for your PAS file storage.
- b. For **GCP Service Account Email** enter the email address associated with your GCP account.
- c. For **GCP Service Account JSON** enter the account key that you use to access the specified GCP project, in JSON format.
- d. Enter the names of the storage buckets you created in [Preparing to Deploy Ops Manager on GCP Manually](#).

- Buildpacks Bucket Name: MY-PCF-buildpacks
- Droplets Bucket Name: MY-PCF-droplets
- Resources Bucket Name: MY-PCF-resources
- Packages Bucket Name: MY-PCF-packages

e. Click Save.

Terraform

This section describes how to configure file storage for GCP if you installed PCF with Terraform.

PAS can use Google Cloud Storage (GCS) as its external filestore by using either a GCP interoperable storage access key or your GCS Service Account. Choose the procedure that corresponds to your use case.

PAS can use Google Cloud Storage (GCS) as its external filestore by using either a GCP inter-operable storage access key or your GCS Service Account. See below for how to configure each option.

Note: The Pivotal Application Service for Windows (PASW) tile is incompatible with GCP configured with a GCS file store. If you are deploying PASW in your GCP environment, Pivotal recommends that you select the S3-compatible file store for your environment.

External Google Cloud Storage with Access Key and Secret Key

1. Select the **External Google Cloud Storage with Access Key and Secret Key** option

This section determines where you would like to place your Application Service Cloud Controller's file storage.

Configure your Cloud Controller's filesystem*

- Internal WebDAV (provided by Application Service)
- External S3-Compatible File Store (if you want to use a service like S3 or Ceph)
- External Google Cloud Storage with Access Key and Secret Key

Access Key *

Secret Key *

Secret

Buildpacks Bucket Name *

Bucket for storing app buildpacks.

Droplets Bucket Name *

Packages Bucket Name *

Resources Bucket Name *

- External Google Cloud Storage with Service Account

- External Azure Storage

2. Enter values for **Access Key** and **Secret Key**. To obtain the values for these fields:

- a. In the GCP Console, navigate to the **Storage** tab, then click **Settings**.
- b. Click **Interoperability**.
- c. If necessary, click **Enable interoperability access**. If interoperability access is already enabled, confirm that the default project matches the project where you are installing PCF.

The screenshot shows the Google Cloud Platform Settings page with the Interoperability tab selected. The page includes sections for Default project for interoperable access (with 'cf-docs' as the default project), Interoperable storage access keys (with an access key and secret displayed), and a 'Create a new key' button.

- d. Click **Create a new key**.
- e. Copy and paste the generated values into the corresponding PAS fields. PCF uses these values for authentication when connecting to Google Cloud Storage.
- f. Enter the names of the storage buckets you created in [Deploying Ops Manager on GCP Using Terraform](#):
 - **Buildpacks Bucket Name:** Enter the value of `buildpacks_bucket` from your Terraform output.
 - **Droplets Bucket Name:** Enter the value of `droplets_bucket` from your Terraform output.
 - **Resources Bucket Name:** Enter the value of `packages_bucket` from your Terraform output.
 - **Packages Bucket Name:** Enter the value of `resources_bucket` from your Terraform output.
- g. Click **Save**.

External Google Cloud Storage with Service Account

1. Select the **External Google Cloud Storage with Service Account** option

Configure your Cloud Controller's filesystem*

- Internal WebDAV (provided by Application Service)
- External S3-Compatible File Store (if you want to use a service like S3 or Ceph)
- External Google Cloud Storage with Access Key and Secret Key
- External Google Cloud Storage with Service Account

GCP Project ID *

GCP Service Account Email *

Your Google Cloud Platform Service Account Email.

GCP Service Account Key JSON *

Buildpacks Bucket Name *

- a. For **GCP Project ID** enter the Project ID on your GCP Console that you want to use for your PAS file storage.
- b. For **GCP Service Account Email** enter the email address associated with your GCP account.
- c. For **GCP Service Account JSON** enter the account key that you use to access the specified GCP project, in JSON format.
- d. Enter the names of the storage buckets you created in [Deploying Ops Manager on GCP Using Terraform](#):

- **Buildpacks Bucket Name:** Enter the value of `buildpacks_bucket` from your Terraform output.
- **Droplets Bucket Name:** Enter the value of `droplets_bucket` from your Terraform output.
- **Resources Bucket Name:** Enter the value of `packages_bucket` from your Terraform output.
- **Packages Bucket Name:** Enter the value of `resources_bucket` from your Terraform output.

- e. Click **Save**.

Openstack

For production-level PCF deployments on OpenStack, the recommended selection is **External S3-Compatible**.

For more factors to consider when selecting file storage, see [Considerations for Selecting File Storage in Pivotal Cloud Foundry](#).

To use an external S3-compatible filestore for PAS file storage, perform the following steps:

1. Select the **External S3-Compatible Filestore** option and complete the following fields:

- o Enter the `https://` URL Endpoint for your region.
For example, in the `us-west-2` region, enter `https://s3-us-west-2.amazonaws.com/`.
- o If you use an AWS instance profile to manage role information for your filestore, select the **S3 AWS with Instance Profile** checkbox. For more information, see [AWS Identity and Access Management](#) in the AWS documentation.

S3 AWS with Instance Profile

Access Key (required if not using S3 AWS with Instance Profile)

Secret Key (required if not using S3 AWS with Instance Profile)

- o Alternatively, enter the **Access Key** and **Secret Key** of the `pcf-user` you created when configuring AWS for PCF. If you select the **S3 AWS with Instance Profile** checkbox and also enter an **Access Key** and **Secret Key**, the instance profile will overrule the Access Key and Secret Key.

- From the **S3 Signature Version** dropdown, select **V4 Signature**. For more information about S4 signatures, see [Signing AWS API Requests](#) in the AWS documentation.
- For **Region**, enter the region in which your S3 buckets are located. `us-west-2` is an example of an acceptable value for this field.
- Select **Server-side Encryption** to encrypt the contents of your S3 filestore. This option is only available for AWS S3.
- (Optional) If you selected **Server-side Encryption**, you can also specify a **KMS Key ID**. PAS uses the KMS key to encrypt files uploaded to the blobstore. If you do not provide a KMS Key ID, PAS uses the default AWS key. For more information, see [Protecting Data Using Server-Side Encryption with AWS KMS-Managed Keys \(SSE-KMS\)](#).
- Enter names for your S3 buckets:

Ops Manager Field	Value	Description
Buildpacks Bucket Name	<code>pcf-buildpacks-bucket</code>	This S3 bucket stores app buildpacks.
Droplets Bucket Name	<code>pcf-droplets-bucket</code>	This S3 bucket stores app droplets. Pivotal recommends that you use a unique bucket name for droplets, but you can also use the same name as above.
Packages Bucket Name	<code>pcf-packages-bucket</code>	This S3 bucket stores app packages. Pivotal recommends that you use a unique bucket name for packages, but you can also use the same name as above.
Resources Bucket Name	<code>pcf-resources-bucket</code>	This S3 bucket stores app resources. Pivotal recommends that you use a unique bucket name for app resources, but you can also use the same name as above.

- Configure the following depending on whether your S3 buckets have versioning enabled:
 - Versioned S3 buckets:** Enable the **Use versioning for backup and restore** checkbox to archive each bucket to a version.
 - Unversioned S3 buckets:** Disable the **Use versioning for backup and restore** checkbox, and enter a backup bucket name for each active bucket. The backup bucket name must be different from the name of the active bucket it backs up.

Use versioning for backup and restore.
(All of the above buckets must have versioning enabled and use the V4 S3 Signature Version).

Backup Region
`us-east-1`

Backup Buildpacks Bucket Name

Backup Droplets Bucket Name

Backup Packages Bucket Name

For more information

about setting up external S3 blobstores, see the [Backup and Restore with External Blobstores](#) topic in the Cloud Foundry documentation.

- Click **Save**.

 **Note:** For more information regarding AWS S3 Signatures, see the [Authenticating Requests](#) topic in the AWS documentation.

vSphere

For production-level PCF deployments on vSphere, the recommended selection is **External S3-Compatible** and the use of an external filestore. For more information about production-level PCF deployments on vSphere, see the [Reference Architecture for Pivotal Cloud Foundry on vSphere](#).

For more factors to consider when selecting file storage, see [Considerations for Selecting File Storage in Pivotal Cloud Foundry](#).

To use an external S3-compatible filestore for PAS file storage, perform the following steps:

1. Select the **External S3-Compatible Filestore** option and complete the following fields:

- o Enter the **https:// URL Endpoint** for your region.
For example, in the **us-west-2** region, enter **https://s3-us-west-2.amazonaws.com/**.
- o If you use an AWS instance profile to manage role information for your filestore, select the **S3 AWS with Instance Profile** checkbox. For more information, see [AWS Identity and Access Management](#) in the AWS documentation.

S3 AWS with Instance Profile

Access Key (required if not using S3 AWS with Instance Profile)

Secret Key (required if not using S3 AWS with Instance Profile)

Secret

Cancel

- o Alternatively, enter the **Access Key** and **Secret Key** of the **pcf-user** you created when configuring AWS for PCF. If you select the **S3 AWS with Instance Profile** checkbox and also enter an **Access Key** and **Secret Key**, the instance profile will overrule the Access Key and Secret Key.
- o From the **S3 Signature Version** dropdown, select **V4 Signature**. For more information about S4 signatures, see [Signing AWS API Requests](#) in the AWS documentation.
- o For **Region**, enter the region in which your S3 buckets are located. **us-west-2** is an example of an acceptable value for this field.
- o Select **Server-side Encryption** to encrypt the contents of your S3 filestore. This option is only available for AWS S3.
- o (Optional) If you selected **Server-side Encryption**, you can also specify a **KMS Key ID**. PAS uses the KMS key to encrypt files uploaded to the blobstore. If you do not provide a KMS Key ID, PAS uses the default AWS key. For more information, see [Protecting Data Using Server-Side Encryption with AWS KMS-Managed Keys \(SSE-KMS\)](#).
- o Enter names for your S3 buckets:

Ops Manager Field	Value	Description
Buildpacks Bucket Name	pcf-buildpacks-bucket	This S3 bucket stores app buildpacks.
Droplets Bucket Name	pcf-droplets-bucket	This S3 bucket stores app droplets. Pivotal recommends that you use a unique bucket name for droplets, but you can also use the same name as above.
Packages Bucket Name	pcf-packages-bucket	This S3 bucket stores app packages. Pivotal recommends that you use a unique bucket name for packages, but you can also use the same name as above.
Resources Bucket Name	pcf-resources-bucket	This S3 bucket stores app resources. Pivotal recommends that you use a unique bucket name for app resources, but you can also use the same name as above.

- o Configure the following depending on whether your S3 buckets have versioning enabled:
 - **Versioned S3 buckets:** Enable the **Use versioning for backup and restore** checkbox to archive each bucket to a version.
 - **Unversioned S3 buckets:** Disable the **Use versioning for backup and restore** checkbox, and enter a backup bucket name for each active bucket. The backup bucket name must be different from the name of the active bucket it backs up.

Use versioning for backup and restore.
(All of the above buckets must have
versioning enabled and use the V4 S3
Signature Version).

If versioning is not enabled or supported, you can configure separate buckets below for
backups. More information at <https://docs.pivotal.io/pivotalcf/2-1/customizing/backup-restore/backup-pcf-bbr.html>.

Backup Region

Backup Buildpacks Bucket Name

Backup Droplets Bucket Name

Backup Packages Bucket Name

For more information

about setting up external S3 blobstores, see the [Backup and Restore with External Blobstores](#) topic in the Cloud Foundry documentation.

2. Click **Save**.

 **Note:** For more information regarding AWS S3 Signatures, see the [Authenticating Requests](#) topic in the AWS documentation.

Configuring Load Balancing for PAS

This topic describes how to configure load balancing for Pivotal Application Service (PAS) by entering the names of your load balancers in the **Resource Config** pane of the PAS tile. This procedure varies by IaaS an installation method. See the section below that corresponds to your use case.

AWS

To configure the Gorouter or HAProxy to AWS Elastic Load Balancers, do the following:

1. Record the names of your ELBs. If you followed the procedures in the [Installing PCF on AWS Manually](#) topic, you created the following:

- o `pcf-ssh-elb` : A SSH load balancer. This is a Classic Load Balancer.
- o `pcf-tcp-elb` : A TCP load balancer. This is a Classic Load Balancer.
- o `pcf-web-elb` : A web load balancer. This is an Application Load Balancer.
- o `pcf-web-elb-target-group` : a target group for the web load balancer

2. In the PAS tile, click **Resource Config**.

Resource Config					
JOB	INSTANCES	PERSISTENT DISK TYPE	VM TYPE	LOAD BALANCERS	INTERNET CONNECTED
Consul	1	Automatic: 1 GB	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
NATS	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
File Storage	Automatic: 1	Automatic: 100 GB	Automatic: medium.mem (cpu: 1, ram: 6 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
MySQL Proxy	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
MySQL Server	1	Automatic: 100 GB	Automatic: large.disk (cpu: 2, ram: 8 GB, disk: 64 GB)		<input checked="" type="checkbox"/>
Backup Prepare Node	Automatic: 1	Automatic: 200 GB	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Diego BBS	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
UAA	1	None	Automatic: medium.disk (cpu: 2, ram: 4 GB, disk: 32 GB)		<input checked="" type="checkbox"/>
Cloud Controller	1	None	Automatic: medium.disk (cpu: 2, ram: 4 GB, disk: 32 GB)		<input checked="" type="checkbox"/>
HAProxy	Automatic: 1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Router	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)	<code>pcf-web-elb</code>	<input checked="" type="checkbox"/>
MySQL Monitor	Automatic: 1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Clock Global	Automatic: 1	None	Automatic: medium.disk (cpu: 2, ram: 4 GB, disk: 32 GB)		<input checked="" type="checkbox"/>
Cloud Controller Worker	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Diego Brain	1	Automatic: 1 GB	Automatic: small (cpu: 1, ram: 2 GB, disk: 8 GB)	<code>pcf-ssh-elb</code>	<input checked="" type="checkbox"/>
Diego Cell	1	None	Automatic: xlarge.disk (cpu: 4, ram: 16 GB, disk: 128 GB)		<input checked="" type="checkbox"/>
Loggregator Trafficcontroller	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Syslog Adapter	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Syslog Scheduler	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Doppler Server	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
TCP Router	0	Automatic: 1 GB	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)	<code>pcf-tcp-elb</code>	<input checked="" type="checkbox"/>
CredHub	Automatic: 0	None	Automatic: large (cpu: 2, ram: 8 GB, disk: 16 GB)		<input checked="" type="checkbox"/>

Save

3. Enter the name of your SSH load balancer depending on which release you are using.

- o **PAS:** In the **Load Balancers** field of the **Diego Brain** row, enter the name of your SSH load balancer: `pcf-ssh-elb`.
- o **Small Footprint Runtime:** In the **Load Balancers** field of the **Control** row, enter the name of your SSH load balancer: `pcf-ssh-elb`.

4. In the **Load Balancers** field of the **Router** row, enter the value determined by the type of load balancer you are using:

- o **Application Load Balancer:** Enter the name of the target group of your web load balancer, prefixed with `alb:`: `alb:pcf-web-elb-target-group`. The prefix indicates to Ops Manager that you entered the name of a target group, and is required for AWS Application Load Balancers or Network Load Balancers.
- o **Clasic Load Balancer:** Enter the name of the load balancer: `pcf-web-elb`.



Note: If you are using HAProxy in your deployment, then put the name of the load balancers in the **Load Balancers** field of the **HAProxy** row instead of the **Router** row. For a high availability configuration, scale up the HAProxy job to more than one instance.

5. In the **Load Balancers** field of the **TCP Router** row, enter the name of your TCP load balancer if you enabled TCP routing: `pcf-tcp-elb`.

AWS Terraform

If you are installing PAS on AWS using Terraform, do the following to set up load balancing:

1. Create a file named `vm_extensions_config.yml` with the following content, depending on which release you are using:

- o **Pivotal Application Service (PAS):**

```
---
product-name: cf
resource-config:
  diego_brain:
    elb_names:
      - alb:SSH_TARGET_GROUP_1
      - alb:SSH_TARGET_GROUP_2
    additional_vm_extensions:
      - ssh-lb-security-groups
  router:
    elb_names:
      - alb:WEB_TARGET_GROUPS_1
      - alb:WEB_TARGET_GROUPS_2
    additional_vm_extensions:
      - web-lb-security-groups
  tcp_router:
    elb_names:
      - alb:TCP_TARGET_GROUP_1
      - alb:TCP_TARGET_GROUP_2
    additional_vm_extensions:
      - tcp-lb-security-groups
```

- o **Small Footprint Runtime:**

```
---
product-name: cf
resource-config:
  control:
    elb_names:
      - alb:SSH_TARGET_GROUP_1
      - alb:SSH_TARGET_GROUP_2
    additional_vm_extensions:
      - ssh-lb-security-groups
  router:
    elb_names:
      - alb:WEB_TARGET_GROUPS_1
      - alb:WEB_TARGET_GROUPS_2
    additional_vm_extensions:
      - web-lb-security-groups
  tcp_router:
    elb_names:
      - alb:TCP_TARGET_GROUP_1
      - alb:TCP_TARGET_GROUP_2
    additional_vm_extensions:
      - tcp-lb-security-groups
```

2. Replace values in the file as follows;

- o `SSH_TARGET_GROUP_X` : Enter your SSH target groups. You can find these values by running `terraform output ssh_target_groups`.
- o `WEB_TARGET_GROUPS_X` : Enter your web target groups. You can find these values by running `terraform output web_target_groups`.
- o `TCP_TARGET_GROUP_X` : Enter your TCP target groups. You can find these values by running `terraform output tcp_target_groups`.

3. Apply the VM extension configuration using the `om` CLI. For more information about `om`, see [pivotalscf/om](#) in GitHub.

```
om -k \
  -t "YOUR-OPS-MANAGER-FQDN" \
  -u "USERNAME" \
  -p "PASSWORD" \
  configure-product \
  -c vm_extensions_config.yml
```

Where:

- `YOUR-OPS-MANAGER-FQDN` is the URL at which you access your Ops Manager instance. This corresponds to `ops_manager_dns` in the Terraform output.
- `USERNAME` is the user name you entered when configuring internal authentication.
- `PASSWORD` is the password you entered when configuring internal authentication.

Note: If you did not configure internal authentication, you must modify this command to use a client ID and secret instead of user name and password. See [Authentication](#) in the `om` repository.

Azure

To configure the Gorouter to Azure Load Balancers, do the following:

1. Select **Resource Config**.

Resource Config				
JOB	INSTANCES	PERSISTENT DISK TYPE	VM TYPE	LOAD BALANCER
Consul	Automatic: 3	Automatic: 1 GB	Automatic: Standard_F1s (cpu: 1, ram: 2 GB, disk: 1	
NATS	Automatic:	None	Automatic: Standard_F1s (cpu: 1, ram: 2 GB, disk: 1	
etcd	Automatic: 3	Automatic: 1 GB	Automatic: Standard_F1s (cpu: 1, ram: 2 GB, disk: 1	
File Storage	Automatic: 1	Automatic: 100 GB	Automatic: Standard_F2s (cpu: 2, ram: 4 GB, disk: 3	
MySQL Proxy	Automatic:	None	Automatic: Standard_F1s (cpu: 1, ram: 2 GB, disk: 1	
MySQL Server	Automatic: 3	Automatic: 100 GB	Automatic: Standard_DS11_v2 (cpu: 2, ram: 14 GB,	
Backup Prepare Node	0	Automatic: 200 GB	Automatic: Standard_F1s (cpu: 1, ram: 2 GB, disk: 1	
UAA	Automatic: 2	None	Automatic: Standard_F2s (cpu: 2, ram: 4 GB, disk: 3	
Cloud Controller	Automatic: 2	Automatic: 1 GB	Automatic: Standard_F2s (cpu: 2, ram: 4 GB, disk: 3	
HAProxy	Automatic: 1	None	Automatic: Standard_F1s (cpu: 1, ram: 2 GB, disk: 1	
Router	Automatic: 3	None	Automatic: Standard_F1s (cpu: 1, ram: 2 GB, disk: 1	

2. Ensure a `Standard` VM type is selected for the **Router** VM. The PAS deployment fails if you select a `Basic` VM type.
3. Retrieve the name(s) of your external ALB by navigating to the Azure portal, clicking **All resources**, and locating your **Load balancer** resource.

Note: The Azure portal sometimes displays the names of resources with incorrect capitalization. Always use the Azure CLI to retrieve the correctly capitalized name of a resource. `az network lb list`

4. Locate the **Router** job in the **Resource Config** pane and enter the name of your external ALB in the field under **Load Balancers**.
5. Retrieve the name of your Diego SSH Load Balancer by navigating to the Azure portal, clicking **All resources**, and locating your **Load balancer** resource.
6. Locate the **Diego Brain** job in the **Resource Config** pane and enter the name of the Diego SSH Load Balancer in the field under **Load Balancers**.
7. Ensure that the **Internet Connected** checkboxes are deselected for all jobs.
8. Scale the number of instances as appropriate for your deployment.

Note: For a high availability deployment of PCF on Azure, Pivotal recommends scaling the number of each PAS job to a minimum of three (3) instances. Using three or more instances for each job creates a sufficient number of availability sets and fault domains for your deployment. For more information, see [Reference Architecture for Pivotal Cloud Foundry on Azure](#).

Azure Terraform

To configure the Gorouter to Azure Load Balancers, do the following:

1. Select **Resource Config**.

Resource Config

JOB	INSTANCES	PERSISTENT DISKTYPE	VM TYPE	LOAD BALANCER
Consul	Automatic: 3	Automatic: 1 GB	Automatic: Standard_F1s (cpu: 1, ram: 2 GB, disk: 1)	
NATS	Automatic:	None	Automatic: Standard_F1s (cpu: 1, ram: 2 GB, disk: 1)	
etcd	Automatic: 3	Automatic: 1 GB	Automatic: Standard_F1s (cpu: 1, ram: 2 GB, disk: 1)	
File Storage	Automatic: 1	Automatic: 100 GB	Automatic: Standard_F2s (cpu: 2, ram: 4 GB, disk: 3)	
MySQL Proxy	Automatic:	None	Automatic: Standard_F1s (cpu: 1, ram: 2 GB, disk: 1)	
MySQL Server	Automatic: 3	Automatic: 100 GB	Automatic: Standard_DS11_v2 (cpu: 2, ram: 14 GB, disk: 1)	
Backup Prepare Node	0	Automatic: 200 GB	Automatic: Standard_F1s (cpu: 1, ram: 2 GB, disk: 1)	
UAA	Automatic: 2	None	Automatic: Standard_F2s (cpu: 2, ram: 4 GB, disk: 3)	
Cloud Controller	Automatic: 2	Automatic: 1 GB	Automatic: Standard_F2s (cpu: 2, ram: 4 GB, disk: 3)	
HAProxy	Automatic: 1	None	Automatic: Standard_F1s (cpu: 1, ram: 2 GB, disk: 1)	
Router	Automatic: 3	None	Automatic: Standard_F1s (cpu: 1, ram: 2 GB, disk: 1)	

- a. Ensure a **Standard** VM type is selected for the **Router** VM. The PAS deployment fails if you select a **Basic** VM type.
2. Enter the value of `web_lb_name` from your Terraform output in the **Resource Config** pane under **Load Balancers** for the **Router** job.
3. Enter the value of `diego_ssh_lb_name` from your Terraform output in the **Resource Config** pane under **Load Balancers** for the **Diego Brain** job.
4. Ensure that the **Internet Connected** checkboxes are deselected for all jobs.
5. Scale the number of instances as appropriate for your deployment.

Note: For a high availability deployment of PCF on Azure, Pivotal recommends scaling the number of each PAS job to a minimum of three (3) instances. Using three or more instances for each job creates a sufficient number of availability sets and fault domains for your deployment. For more information, see [Reference Architecture for Pivotal Cloud Foundry on Azure](#).

GCP

To Configure Gorouter to GCP Load Balancers, do the following:

1. Navigate to the GCP Console and click **Load balancing**.

Load balancer	Protocol	Action
pcf-router	HTTP(S)	trash Edit
pcf-ssh	TCP	trash Edit
pcf-websockets	TCP	trash Edit

To edit load balancing resources like forwarding rules and target proxies, go to the [advanced menu](#).

You should see the SSH load balancer, the HTTP(S) load balancer, the TCP WebSockets load balancer, and the TCP router that you created in the [Preparing to Deploy PCF on GCP](#) topic.

2. Record the name of your SSH load balancer and your TCP WebSockets load balancer, `MY-PCF-wss-logs` and `MY-PCF-ssh-proxy`.
3. Click your HTTP(S) load balancer, `MY-PCF-global-pcf`.

4. Under **Backend services**, record the name of the backend service of the HTTP(S) load balancer, `MY-PCF-http-lb-backend`.

5. In the PAS tile, click **Resource Config**.

JOB	INSTANCES	PERSISTENT DISK TYPE	VM TYPE	LOAD BALANCERS	INTERNET CONNECTED
Consul	1	Automatic: 1 GB	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
NATS	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
File Storage	Automatic: 1	Automatic: 100 GB	Automatic: medium.mem (cpu: 1, ram: 6 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
MySQL Proxy	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
MySQL Server	1	Automatic: 100 GB	Automatic: large.disk (cpu: 2, ram: 8 GB, disk: 64 GB)		<input checked="" type="checkbox"/>
Backup Prepare Node	Automatic: 1	Automatic: 200 GB	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Diego BBS	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
UAA	1	None	Automatic: medium.disk (cpu: 2, ram: 4 GB, disk: 32 GB)		<input checked="" type="checkbox"/>
Cloud Controller	1	None	Automatic: medium.disk (cpu: 2, ram: 4 GB, disk: 32 GB)		<input checked="" type="checkbox"/>
HAProxy	Automatic: 1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Router	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)	pcf-web-elb	<input checked="" type="checkbox"/>
MySQL Monitor	Automatic: 1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Clock Global	Automatic: 1	None	Automatic: medium.disk (cpu: 2, ram: 4 GB, disk: 32 GB)		<input checked="" type="checkbox"/>
Cloud Controller Worker	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Diego Brain	1	Automatic: 1 GB	Automatic: small (cpu: 1, ram: 2 GB, disk: 8 GB)	pcf-ssh-elb	<input checked="" type="checkbox"/>
Diego Cell	1	None	Automatic: xlarge.disk (cpu: 4, ram: 16 GB, disk: 128 GB)		<input checked="" type="checkbox"/>
Loggregator Trafficcontroller	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Syslog Adapter	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Syslog Scheduler	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Doppler Server	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
TCP Router	0	Automatic: 1 GB	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)	pcf-tcp-elb	<input checked="" type="checkbox"/>
CredHub	Automatic: 0	None	Automatic: large (cpu: 2, ram: 8 GB, disk: 16 GB)		<input checked="" type="checkbox"/>

6. Under the **LOAD BALANCERS** column of the **Router** row, enter a comma-delimited list consisting of the name of your TCP WebSockets load balancer and the name of your HTTP(S) load balancer backend with the protocol prepended. For example, `tcp:MY-PCF-wss-logs,http:MY-PCF-http-lb-backend`.

Note: Do not add a space between key/value pairs in the `LOAD BALANCER` field or it will fail.

Note: If you are using HAProxy in your deployment, then enter the above load balancer values in the `LOAD BALANCERS` field of the **HAProxy** row instead of the **Router** row. For a high availability configuration, scale up the HAProxy job to more than one instance.

7. If you have enabled TCP routing in the **Networking** pane and set up the [TCP Load Balancer in GCP](#), add the name of your TCP load balancer, prepended with `tcp:`, to the **LOAD BALANCERS** column of the **TCP Router** row. For example, `tcp:pcf-tcp-router`.
8. Enter the name of your SSH load balancer depending on which release you are using.
 - **PAS:** Under the **LOAD BALANCERS** column of the **Diego Brain** row, enter the name of your SSH load balancer prepended with `tcp:`. For example, `tcp:MY-PCF-ssh-proxy`.
 - **Small Footprint Runtime:** Under the **LOAD BALANCERS** column of the **Control** row, enter the name of your SSH load balancer prepended with `tcp:`.
9. Verify that the **Internet Connected** checkbox for every job is unchecked. When preparing your GCP environment, you provisioned a Network Address Translation (NAT) box to provide Internet connectivity to your VMs instead of providing them with public IP addresses to allow the jobs to reach the Internet.

Note: If you want to provision a Network Address Translation (NAT) box to provide Internet connectivity to your VMs instead of providing them with public IP addresses, deselect the **Internet Connected** checkboxes. For more information about using NAT in GCP, see the [GCP documentation](#).

10. Click **Save**.

GCP Terraform

To Configure Gorouter to GCP Load Balancers, do the following:

1. Click **Resource Config**.
2. Under the **LOAD BALANCERS** column of the **Router** row, enter a comma-delimited list consisting of the values of `ws_router_pool` and `http_lb_backend_name` from your Terraform output. For example, `tcp:pcf-cf-ws,http:pcf-httpslb`. These are the names of the TCP WebSockets and HTTP(S) load balancers for your deployment.

Note: Do not add a space between key/value pairs in the **LOAD BALANCER** field or it will fail.

Note: If you are using HAProxy in your deployment, then enter the above load balancer values in the **LOAD BALANCERS** field of the **HAPRoxy** row instead of the **Router** row. For a high availability configuration, scale up the HAProxy job to more than one instance.
3. If you have enabled TCP routing in the **Networking** pane, add the value of `tcp_router_pool` from your Terraform output, prepended with `tcp:`, to the **LOAD BALANCERS** column of the **TCP Router** row. For example, `tcp:pcf-cf-tcp`.
4. Enter the name of your SSH load balancer depending on which release you are using.
 - **PAS:** Under the **LOAD BALANCERS** column of the **Diego Brain** row, enter the value of `ssh_router_pool` from your Terraform output, prepended with `tcp:`. For example, `tcp:MY-PCF-ssh-proxy`.
 - **Small Footprint Runtime:** Under the **LOAD BALANCERS** column of the **Control** row, enter the value of `ssh_router_pool` from your Terraform output, prepended with `tcp:`.
5. Verify that the **Internet Connected** checkbox for every job is checked. The terraform templates do not provision a Network Address Translation (NAT) box for Internet connectivity to your VMs so instead they will be provided with ephemeral public IP addresses to allow the jobs to reach the Internet.

Note: If you want to provision a Network Address Translation (NAT) box to provide Internet connectivity to your VMs instead of providing them with public IP addresses, deselect the **Internet Connected** checkboxes. For more information about using NAT in GCP, see the [GCP documentation](#).

6. Click **Save**.

Openstack

Unless you are using your own load balancer, you must enable traffic flow to the OpenStack private subnet as follows. Give each HAProxy a way of routing traffic into the private subnet by providing public IP addresses as floating IP addresses.

1. Click **Resource Config**.

Resource Config

JOB	INSTANCES	PERSISTENT DISK TYPE	VM TYPE	LOAD BALANCERS	INTERNET CONNECTED
Consul	1	Automatic: 1 GB	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
NATS	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
File Storage	Automatic: 1	Automatic: 100 GB	Automatic: medium.mem (cpu: 1, ram: 6 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
MySQL Proxy	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
MySQL Server	1	Automatic: 100 GB	Automatic: large.disk (cpu: 2, ram: 8 GB, disk: 64 GB)		<input checked="" type="checkbox"/>
Backup Prepare Node	Automatic: 1	Automatic: 200 GB	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Diego BBS	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
UAA	1	None	Automatic: medium.disk (cpu: 2, ram: 4 GB, disk: 32 GB)		<input checked="" type="checkbox"/>
Cloud Controller	1	None	Automatic: medium.disk (cpu: 2, ram: 4 GB, disk: 32 GB)		<input checked="" type="checkbox"/>
HAProxy	Automatic: 1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Router	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)	pcf-web-elb	<input checked="" type="checkbox"/>
MySQL Monitor	Automatic: 1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Clock Global	Automatic: 1	None	Automatic: medium.disk (cpu: 2, ram: 4 GB, disk: 32 GB)		<input checked="" type="checkbox"/>
Cloud Controller Worker	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Diego Brain	1	Automatic: 1 GB	Automatic: small (cpu: 1, ram: 2 GB, disk: 8 GB)	pcf-ssh-elb	<input checked="" type="checkbox"/>
Diego Cell	1	None	Automatic: xlarge.disk (cpu: 4, ram: 16 GB, disk: 128 GB)		<input checked="" type="checkbox"/>
Loggregator Trafficcontroller	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Syslog Adapter	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Syslog Scheduler	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Doppler Server	1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
TCP Router	0	Automatic: 1 GB	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)	pcf-tcp-elb	<input checked="" type="checkbox"/>
CredHub	Automatic: 0	None	Automatic: large (cpu: 2, ram: 8 GB, disk: 16 GB)		<input checked="" type="checkbox"/>

Save

2. Enter one or more IP addresses in **Floating IPs** for each HAProxy.
3. (Optional) If you have enabled the TCP Routing feature, enter one or more IP addresses in **Floating IPs** column for each TCP Router.
4. Click **Save**.

Installing PCF Isolation Segment

This topic describes how to install the PCF Isolation Segment tile, which allows operators to isolate deployment workloads into dedicated resource pools called *isolation segments*.

Installing the tile installs a single isolation segment. However, you can install multiple isolation segments using the Replicator tool documented in [Step 4](#).

After installing the tile, you must perform the steps in the [Create an Isolation Segment](#) section of the *Managing Isolation Segments* topic to create the isolation segment in the Cloud Controller Database (CCDB). The topic also includes information about managing an isolation segment.

For more information about how isolation segments work, see [Isolation Segments](#).

Step 1: (Optional) Configure Routing

By default, the Pivotal Application Service (PAS) Router handles traffic for your isolation segment. However, you can deploy a dedicated router for your isolation segment instead. For information about configuring and managing routing for isolation segments, see the [Routing for Isolation Segments](#) topic.

To deploy a dedicated router, perform the following steps:

1. Add a load balancer in front of the PAS Router. The steps to do this depend on your IaaS, but the setup of the load balancer should mirror the setup of the load balancer for the PAS Router that you configured in the PAS tile.
2. Create a wildcard DNS entry for traffic routed to any app in the isolation segment. For example, `*.iso.example.com`.
3. Attach the wildcard DNS entry to the load balancer you created.

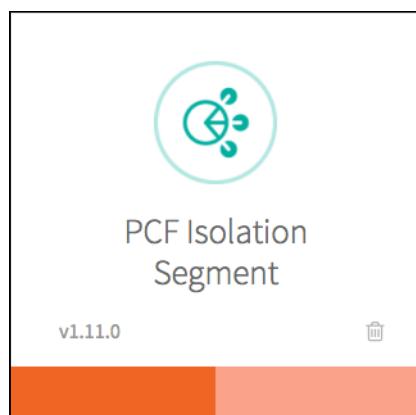
Step 2: Install the Tile

Perform the following steps to install the PCF Isolation Segment tile:

1. Download the product file from [Pivotal Network](#).
2. Navigate to `YOUR-OPSMAN-FQDN` in a browser to log in to the Ops Manager Installation Dashboard.
3. Click **Import a Product** and select the downloaded product file.
4. Under **PCF Isolation Segment** in the left column, click the plus sign.

Step 3: Configure the Tile

Click the orange **PCF Isolation Segment** tile to start the configuration process.



Assign AZs and Networks

Perform the following steps to configure the PCF Isolation Segment tile:

1. Click **Assign AZs and Networks**.

AZ and Network Assignments

Place singleton jobs in

us-central1-a
 us-central1-b
 us-central1-c

Balance other jobs in

us-central1-a
 us-central1-b
 us-central1-c

Network

pcf-network

Save

2. Select an availability zone for your singleton jobs, and one or more availability zones to balance other jobs in.
3. Select a network. This network does not need to be the same network where you deployed PAS. For most deployments, operators should create unique networks in which to deploy the Isolation Segment tile. These networks should maintain network reachability with the Diego components so that the cells can reach the Diego Brain and Diego Database VMs.
4. Click **Save**.

Networking

Perform the following steps to configure the PCF Isolation Segment tile:

1. Click **Networking**.

Configure security and routing services for your isolation segment.

Router IPs

HAProxy IPs

2. (Optional): Under **Router IPs**, enter one or more static IP addresses for the routers that handle this isolation segment. These IP addresses must be within the subnet CIDR block that you defined in the Ops Manager network configuration for your Isolation Segment. If you have a load balancer, configure it to point to these IP addresses.

Note: Entering the static IP addresses is not necessary for deployments running on a public IaaS such as AWS, GCP, or Azure because PCF users specify the IaaS load balancer in the **Resource Config** section of the **PCF Isolation Segment** tile.

3. If you want to use HAProxy for this isolation segment, enter at least one address in the **HAProxy IPs** field. You should specify more than one IP

address for high availability. Then configure your load balancer to forward requests for the domains you have set up for your deployment to these IP addresses. For more information, see [Configuring SSL/TLS Termination at HAProxy](#).

Note: If you rely on HAProxy for a feature in Pivotal Application Service (PAS) and you want isolated networking for this isolation segment, you may want to deploy the HAProxy provided by the Isolation Segment tile.

- Under **Certificates and Private Key for HAProxy and Router**, you must provide at least one **Certificate and Private Key** name and certificate key pair for HAProxy and Gorouter. HAProxy and Gorouter are enabled to receive TLS communication by default. You can configure multiple certificates for HAProxy and Gorouter.

Note: When providing custom certificates, enter them in the following order: `wildcard`, `Intermediate`, `CA`. For more information, see [Creating a .pem File for SSL Certificate Installations](#) in the DigiCert documentation.

- Click **Add** to add a name for the certificate chain and its private key pair. This certificate is the default used by Gorouter and HAProxy. You can either provide a certificate signed by a Certificate Authority (CA) or click on the **Generate RSA Certificate** link to generate a certificate generated by the Ops Manager CA.

Note: If you configured Ops Manager Front End without a certificate, you can use this new certificate to complete Ops Manager configuration. To configure your Ops Manager Front End certificate, see [Configure Front End](#) in [Preparing to Deploy Ops Manager on GCP Manually](#).

- If you want to configure multiple certificates for HAProxy and Gorouter, click **Add** and fill in the appropriate fields for each additional certificate key pair. For details about generating certificates in Ops Manager for your wildcard system domains, see the [Providing a Certificate for Your SSL/TLS Termination Point](#) topic.

Note: Ensure that you add any certificates that you generate in this pane to your infrastructure load balancer.

- (Optional) When validating client requests using mutual TLS, the Gorouter trusts multiple certificate authorities (CAs) by default. If you want to configure the Gorouter and HAProxy to trust additional CAs, enter your CA certificates under **Certificate Authorities Trusted by Router and HAProxy**. All CA certificates should be appended together into a single collection of PEM-encoded entries.
- In the **Minimum version of TLS supported by HAProxy and Router** field, select the minimum version of TLS to use in HAProxy and Gorouter communications. HAProxy and Gorouter use TLS v1.2 by default. If you need to accommodate clients that use an older version of TLS, select a lower minimum version. For a list of TLS ciphers supported by the Gorouter, see [Securing Traffic into Cloud Foundry](#).
- Configure **Logging of Client IPs in CF Router**. The **Log client IPs** option is set by default. To comply with the General Data Protection Regulation (GDPR), select one of the following options to disable logging of client IP addresses:
 - If your load balancer exposes its own source IP address, disable logging of the `X-Forwarded-For` HTTP header only.
 - If your load balancer exposes the source IP of the originating client, disable logging of both the source IP address and the `X-Forwarded-For` HTTP header.
- Under **Configure support for the X-Forwarded-Client-Cert header**, configure PCF handles `x-forwarded-client-cert` (XFCC) HTTP headers based on where TLS is terminated for the first time in your deployment.

The following table indicates which option to choose based on your deployment configuration.

Deployment Configuration	TLS Option	Additional Notes
<ul style="list-style-type: none"> The Load Balancer is terminating TLS, and Load balancer is configured to put the client certificate from a mutual authentication TLS handshake into the X-Forwarded-Client-Cert HTTP header 	TLS terminated for the first time at infrastructure load balancer (default).	Both HAProxy and the Gorouter forward the XFCC header when included in the request.
<ul style="list-style-type: none"> The Load Balancer is configured to pass through the TLS handshake via TCP to the instances of HAProxy, and HAProxy instance count is > 0 	TLS terminated for the first time at HAProxy.	<p>HAProxy sets the XFCC header with the client certificate received in the TLS handshake. The Gorouter forwards the header.</p> <p>Breaking Change: In the Router behavior for Client Certificates field, you cannot select the Router does not request client certificates option.</p>
		The Gorouter strips the XFCC header if it is included in the request and

<ul style="list-style-type: none"> The Load Balancer is configured to pass through the TLS handshake via TCP to instances of the Gorouter 	<p>TLS terminated for the first time at the Gorouter.</p>	<p>forwards the client certificate received in the TLS handshake in a new XFCC header.</p> <p>If you have deployed instances of HAProxy, app traffic bypasses those instances in this configuration. If you have also configured your load balancer to route requests for ssh directly to the Diego Brain, consider reducing HAProxy instances to 0.</p> <p>Breaking Change: In the Router behavior for Client Certificates field, you cannot select the Router does not request client certificates option.</p>
--	---	---

For a description of the behavior of each configuration option, see [Forward Client Certificate to Applications](#).

- To configure HAProxy to handle client certificates, select one of the following options in the **HAProxy behavior for Client Certificate Validation** field:
 - HAProxy does not request client certificates.** This option requires mutual authentication, which makes it incompatible with XFCC option **TLS terminated for the first time at HAProxy**. HAProxy does not request client certificates, so the client does not provide them and no validation occurs. This is the default configuration.
 - HAProxy requests but does not require client certificates.** The HAProxy requests client certificates in TLS handshakes, validates them when presented, but does not require them.

⚠ warning: Upon upgrade, PAS will fail to receive requests if your load balancer is configured to present a client certificate in the TLS handshake with HAProxy but HAProxy has not been configured with the certificate authority used to sign it. To mitigate this issue, select **HAProxy does not request client certificates** in the **Networking** pane or configure the HAProxy with the appropriate CA.

- To configure Gorouter behavior for handling client certificates, select one of the following options in the **Router behavior for Client Certificate Validation** field:

- Router does not request client certificates.** This option is incompatible with the XFCC configuration options **TLS terminated for the first time at HAProxy** and **TLS terminated for the first time at the Router** in PAS because these options require mutual authentication. As client certificates are not requested, client will not provide them, and thus validation of client certificates will not occur.
- Router requests but does not require client certificates.** The Gorouter requests client certificates in TLS handshakes, validates them when presented, but does not require them. This is the default configuration.
- Router requires client certificates.** The Gorouter validates that the client certificate is signed by a Certificate Authority that the Gorouter trusts. If the Gorouter cannot validate the client certificate, the TLS handshake fails.

⚠ warning: Requests to the platform will fail upon upgrade if your load balancer is configured with client certificates and the Gorouter does not have the certificate authority. To mitigate this issue, select **Router does not request client certificates** for **Router behavior for Client Certificate Validation** in the **Networking** pane.

- In the **TLS Cipher Suites for Router** field, review the TLS cipher suites for TLS handshakes between Gorouter and front-end clients such as load balancers or HAProxy. The default value for this field is `ECDHE-RSA-AES128-GCM-SHA256:TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384`.

To modify the default configuration, use an ordered, colon-delimited list of Golang-supported TLS cipher suites in the OpenSSL format.

Operators should verify that the ciphers are supported by any clients or front-end components that will initiate TLS handshakes with Gorouter. For a list of TLS ciphers supported by Gorouter, see [Securing Traffic into Cloud Foundry](#).

Verify that every client participating in TLS handshakes with Gorouter has at least one cipher suite in common with Gorouter.

Note: Specify cipher suites that are supported by the versions configured in the **Minimum version of TLS supported by HAProxy and Router** field.

- In the **TLS Cipher Suites for HAProxy** field, review the TLS cipher suites for TLS handshakes between HAProxy and its clients such as load balancers and Gorouter. The default value for this field is the following:

`DHE-RSA-AES128-GCM-SHA256:DHE-RSA-AES256-GCM-SHA384:ECDHE-RSA-AES128-GCM-SHA256:ECDHE-RSA-AES256-GCM-SHA384`

To modify the default configuration, use an ordered, colon-delimited list of TLS cipher suites in the OpenSSL format.

Operators should verify that the ciphers are supported by any clients or front-end components that will initiate TLS handshakes with HAProxy.

Verify that every client participating in TLS handshakes with HAProxy has at least one cipher suite in common with HAProxy.

Note: Specify cipher suites that are supported by the versions configured in the **Minimum version of TLS supported by HAProxy and Router** field.

13. Under **HAProxy forwards requests to Router over TLS**, select **Enable** or **Disable** based on your deployment layout.
 - o **Enable HAProxy forwarding of requests to Router over TLS.** To enable communication between HAProxy and the Gorouter, do the following:
 1. Leave **Enable** selected.
 2. In the **Certificate Authority for HAProxy Backend** field, specify the Certificate Authority (CA) that signed the certificate you configured in the **Certificate and Private Key for HAProxy and Router** field.

Note: If you used the **Generate RSA Certificate** link to generate a certificate, then the CA to specify is the Ops Manager CA, which you can locate at the `/api/v0/certificateAuthorities` endpoint in the Ops Manager API.

3. Make sure that Gorouter and HAProxy have TLS cipher suites in common in the **TLS Cipher Suites for Router** and **TLS Cipher Suites for HAProxy** fields.

For more information, see [Terminating SSL/TLS at the Load Balancer and Gorouter](#), [Providing a Certificate for Your SSL/TLS Termination Point](#), and [Using the Ops Manager API](#).

- o **Disable HAProxy forwarding of requests to Router over TLS.** To use non-encrypted communication between HAProxy and Gorouter, do the following:
 1. Select **Disable**.
 2. If you are not using HAProxy, set the number of HAProxy job instances to `0` on the **Resource Config** page. See [Disable Unused Resources](#).

For more information, see [Terminating SSL/TLS at the Gorouter Only](#) and [Terminating SSL/TLS at the Load Balancer Only](#).

14. If you are not using SSL encryption or if you are using self-signed certificates, select **Disable SSL certificate verification for this environment**. Selecting this checkbox also disables SSL verification for route services and disables mutual TLS app identity verification.

Note: For production deployments, Pivotal does not recommend disabling SSL certificate verification.

15. (Optional) If you want HAProxy or the Gorouter to reject any HTTP (non-encrypted) traffic, select the **Disable HTTP on HAProxy and Gorouter** checkbox. When selected, HAProxy and Gorouter will not listen on port 80.
16. (Optional) Select the **Disable insecure cookies on the Router** checkbox to set the secure flag for cookies generated by the router.
17. (Optional) To disable the addition of Zipkin tracing headers on the Gorouter, deselect the **Enable Zipkin tracing headers on the router** checkbox. Zipkin tracing headers are enabled by default. For more information about using Zipkin trace logging headers, see [Zipkin Tracing in HTTP Headers](#).
18. (Optional) To stop the Router from writing access logs to local disk, deselect the **Enable Router to write access logs locally** checkbox. You should consider disabling this checkbox for high traffic deployments since logs may not be rotated fast enough and can fill up the disk.
19. (Optional) By default, Gorouter support for the PROXY protocol is disabled. To enable the PROXY protocol, select **Enable support for PROXY protocol in CF Router**. When enabled, client-side load balancers that terminate TLS but do not support HTTP can pass along information from the originating client. Enabling this option may impact Gorouter performance. For more information about enabling the PROXY protocol in Gorouter, see the [HTTP Header Forwarding](#) sections in the [Securing Traffic in Cloud Foundry](#) topic.
20. **Bypass security checks for route service lookup** has potential security concerns, but may be needed for backwards compatibility. See the [Route Service Internal Lookup Considerations](#) section of [Route Services](#) for details.

21. (Optional) If you want to limit the number of app connections to the backend, enter a value in the **Max Connections Per Backend** field. You can use this field to prevent a poorly behaving app from all the connections and impacting other apps.

To choose a value for this field, review the peak concurrent connections received by instances of the most popular apps in your deployment. You can determine the number of concurrent connections for an app from the `httpStartStop` event metrics emitted for each app request.

If your deployment uses PCF Metrics, you can also obtain this peak concurrent connection information from [Network Metrics](#). The default value is `500`.

22. Under **Enable Keepalive Connections for Router**, select **Enable** or **Disable**. Keep-alive connections are enabled by default. For more information, see [Keepalive Connections](#) in [HTTP Routing](#).
23. (Optional) To accommodate larger uploads over connections with high latency, increase the number of seconds in the **Router Timeout to Backends** field.

24. (Optional) Increase the value of **Load Balancer Unhealthy Threshold** to specify the amount of time, in seconds, that the router continues to accept connections before shutting down. During this period, healthchecks may report the router as unhealthy, which causes load balancers to failover to other routers. Set this value to an amount greater than or equal to the maximum time it takes your load balancer to consider a router instance unhealthy, given contiguous failed healthchecks.

25. (Optional) Modify the value of **Load Balancer Healthy Threshold**. This field specifies the amount of time, in seconds, to wait until declaring the Router instance started. This allows an external load balancer time to register the Router instance as healthy.

Load Balancer Unhealthy Threshold *

Load Balancer Healthy Threshold *

26. (Optional) If app developers in your organization want certain HTTP headers to appear in their app logs with information from the Gorouter, specify them in the **HTTP Headers to Log** field. For example, to support app developers that deploy Spring apps to PCF, you can enter [Spring-specific HTTP headers ↗](#).

HTTP Headers to Log

27. If you expect requests larger than the default maximum of 16 Kbytes, enter a new value (in bytes) for **HAProxy Request Max Buffer Size**. You may need to do this, for example, to support apps that embed a large cookie or query string values in headers.

28. If your PCF deployment uses HAProxy and you want it to receive traffic only from specific sources, use the following fields:

- **HAProxy Protected Domains:** Enter a comma-separated list of domains to protect from unknown source requests.
- **HAProxy Trusted CIDRs:** Optionally, enter a space-separated list of CIDRs to limit which IP addresses from the **Protected Domains** can send traffic to PCF.

29. For **DNS Search Domains**, enter DNS search domains as a comma-separated list. Your containers append these search domains to hostnames to resolve them into full domain names.

DNS Search Domains

DNS search domains to be used in containers. A comma-separated list can be specified.

30. The **Enable Silk Policy Enforcement** checkbox is enabled by default. To disable Silk network policy enforcement between apps, disable the checkbox. Disabling network policy enforcement allows all apps to send network traffic to all other apps in the foundation despite no policy specifically allowing it.

Enable Silk Policy Enforcement *

WARNING!!! Disabling this feature will allow all application containers to access any other application container with no restrictions.

31. Select a sharding mode in the **Router Sharding Mode** field. The options are explained below. For more information, see [Sharding Routers for Isolation Segments](#).

Router Sharding Mode: When 'Isolation Segment Only' is selected, the routers of this tile will only have knowledge of applications deployed to the Cells of this tile; all other requests will receive a 404 response. When 'No Isolation Segments' is selected, the routers of this tile will reject requests for any isolation segment. Choose 'No Isolation Segments' to add a group of routers for the Elastic Runtime tile, as when a private point of entry for the system domain is desired.*

- Isolation Segment Only
 No Isolation Segment

Save

- **Isolation Segment Only:** The routers for the tile acknowledge requests only from apps deployed within the cells of the tile. All other requests

fail.

- **No Isolation Segment** The routers for the tile reject requests for any isolation segment. Choose this option to add a group of routers for the PAS tile, such as when you want a private point of entry for the system domain.

32. Click **Save**.

Application Containers

Perform the following steps to configure the PCF Isolation Segment tile:

1. Click **Application Containers**.

Enable custom configuration that supports your applications at a container level.

Private Docker Insecure Registry Whitelist

10.10.10.10:8888,example.com:8888 [\[x\]](#)

Segment Name *

DNS Servers

Docker Images Disk-Cleanup Scheduling on Cell VMs*

Never clean up Cell disk-space
 Routinely clean up Cell disk-space
 Clean up disk-space once usage fills disk

Reserved disk space for other jobs(MB) (min: 1) *

10240

2. (Optional): Under **Private Docker Insecure Registry Whitelist**, enter one or more private Docker image registries that are secured with self-signed certificates. Use a comma-delimited list in the format `IP:Port` or `Hostname:Port`.
3. Under **Segment Name**, enter the name of your isolation segment. This name must be unique across your PCF deployment. You use this name when performing the steps in the [Create an Isolation Segment](#) section of the *Managing Isolation Segments* topic to create the isolation segment in the Cloud Controller Database (CCDB).
4. Select your preference for **Docker Images Disk-Cleanup Scheduling on Cell VMs**. If you choose **Clean up disk-space once usage fills disk**, enter a **Reserved amount of disk space for other jobs** in megabytes. For more information about the configuration options and how to configure a reserved amount, see [Configuring Docker Images Disk-Cleanup Scheduling](#).
5. The **Enable Containerd Delegation** checkbox governs whether or not Garden delegates container create and destroy operations to the containerd tool. By default, this option is enabled and Garden uses containerd. Disable this option by disabling the checkbox. For more information about the containerd tool, see [containerd](#).
6. Under **Enabling NFSv3 volume services**, select **Enable** or **Disable**. NFS volume services allow application developers to bind existing NFS volumes to their applications for shared file access. For more information, see the [Enabling NFS Volume Services](#) topic.

Note: In a clean install, NFSv3 volume services are enabled by default. In an upgrade, NFSv3 volume services match the setting of the previous deployment.

Enabling NFSv3 volume services will allow application developers to bind existing NFS volumes to their applications for shared file access. *

Enable

LDAP Service Account User

LDAP Service Account Password
 Secret 

LDAP Server Host

LDAP Server Port

LDAP User Search Base
 cn=Users,dc=corp,dc=test,dc=com

LDAP Server CA Cert

Disable

7. (Optional) To configure LDAP for NFSv3 volume services, do the following:

Enabling NFSv3 volume services will allow application developers to bind existing NFS volumes to their applications for shared file access. *

Enable

LDAP Service Account User

LDAP Service Account Password

?

LDAP Server Host

LDAP Server Port

LDAP User Search Base

LDAP Server CA Cert

Disable

- For **LDAP Service Account User**, enter the username of the service account in LDAP that will manage volume services.
- For **LDAP Service Account Password**, enter the password for the service account.
- For **LDAP Server Host**, enter the hostname or IP address of the LDAP server.
- For **LDAP Server Port**, enter the LDAP server port number. If you do not specify a port number, Ops Manager uses 389.
- For **LDAP User Search Base**, enter the location in the LDAP directory tree from which any LDAP User search begins. The typical LDAP Search Base matches your domain name.
For example, a domain named `cloud.example.com` typically uses the following LDAP User Search Base: `ou=Users,dc=example,dc=com`.
- For **LDAP Server CA Cert**, you can optionally enter a certificate if your LDAP server supports TLS and you want to enable TLS connections from the NFS driver to your LDAP server. Paste in the root certificate from your CA certificate or your self-signed certificate.

8. To enable Gorouter to verify app identity using TLS, select the **Router uses TLS to verify application identity** checkbox. Verifying app identity using TLS improves resiliency and consistency for app routes. For more information about Gorouter route consistency modes, see [Preventing Misrouting in HTTP Routing](#).

 **Note:** This feature does not work if the **Disable SSL certificate verification for this environment** checkbox is selected.

9. Click **Save**.

System Logging

1. In the **System Logging** menu, select an option underneath **Do you want to configure syslog for system components?**. **No** is selected by default. This setting only affects Diego cell, router, and HAProxy components within the isolation segment, not shared PAS system components.

Configure system logging. Complete the External Syslog fields only if using an external syslogd server.

Do you want to configure syslog for system components?*

- No
 Yes

Save

2. If you want to use syslog, select Yes.

Do you want to configure syslog for system components?*

- No
 Yes

Address *

Port *

Transport Protocol*

 TCP protocol

Enable TLS

Permitted Peer

TLS CA Certificate

Use TCP for file forwarding local transport

Workaround to avoid truncation of very long (> 1024 bytes) log lines. May negatively impact performance.

Don't Forward Debug Logs

Custom rsyslog Configuration

Save

3. Enter the address of your external syslog aggregation service in the **Address** field. The address can be a hostname or IP address.
 4. Enter a port number in the **Port** field.
 5. Select a protocol from the **Transport Protocol** menu. This is the protocol the system uses to transmit logs to syslog.
 6. (Optional) Select the **Enable TLS** option if you want to transmit logs over TLS.
 7. Enter a **Permitted Peer**.
 8. Paste the certificate for your TLS certificate authority (CA) in the **TLS CA Certificate** field.
 9. (Optional) Select the **Use TCP for file forwarding local transport** to transmit logs over TCP, rather than UDP. This prevents log truncation, but may cause performance issues.
 10. (Optional) To forward DEBUG syslog messages to an external service, disable the **Don't Forward Debug Logs** checkbox. This checkbox is enabled in PCF Isolation Segment by default.
- Note:** Some PCF Isolation Segment components generate a high volume of DEBUG syslog messages. Using the **Don't Forward Debug Logs** checkbox prevents them from being forwarded to external services. PCF Isolation Segment still writes the messages to the local disk.
11. (Optional) To specify a custom syslog rule, enter it in the **Custom syslog Configuration** field in [RainerScript](#) syntax. For more information about customizing syslog rules, see [Customizing Syslog Rules](#).
 12. Select **Enable System Metrics** to emit system-level metrics about all VMs in the deployment. For a list of the VM metrics that the System Metric Agent emits, see [VM Metrics](#) in GitHub. When you enable this, ensure that you open port `9100` for the isolation segment. For more information, see [Configure Firewall Rules](#).
 13. Click **Save**.

Advanced Features

The **Advanced Features** section of Pivotal Application Service (PAS) includes new functionality that may have certain constraints. Although these features are fully supported, Pivotal recommends caution when using them in production environments.

Diego Cell Memory and Disk Overcommit

If your apps do not use the full allocation of disk space and memory set in the **Resource Config** tab, you might want use this feature. These fields control the amount to overcommit disk and memory resources to each Diego Cell VM.

For example, you might want to use the overcommit if your apps use a small amount of disk and memory capacity compared to the amounts set in the **Resource Config** settings for **Diego Cell**.

Note: Due to the risk of app failure and the deployment-specific nature of disk and memory use, Pivotal has no recommendation about how much, if any, memory or disk space to overcommit.

To enable overcommit, do the following:

1. Select **Advanced Features**.
2. Enter the total desired amount of Diego cell memory value in the **Cell Memory Capacity (MB)** field. Refer to the **Diego Cell** row in the **Resource Config** tab for the current Cell memory capacity settings that this field overrides.
3. Enter the total desired amount of Diego cell disk capacity value in the **Cell Disk Capacity (MB)** field. Refer to the **Diego Cell** row in the **Resource Config** tab for the current Cell disk capacity settings that this field overrides.
4. Click **Save**.

Note: Entries made to each of these two fields set the total amount of resources allocated, not the overage.

Configure Router Resources

1. Select Resource Config.

Resource Config					
JOB	INSTANCES	PERSISTENT DISK TYPE	VM TYPE	LOAD BALANCERS	INTERNET CONNECTED
Router	Automatic: 3	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>
Diego Cell	Automatic: 3	None	Automatic: xlarge.disk (cpu: 4, ram: 16 GB, disk: 128 GB)		<input checked="" type="checkbox"/>
HAProxy	Automatic: 3	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)		<input checked="" type="checkbox"/>

Save

2. If you are using a dedicated router for your isolation segment, follow the instructions below.

Note: The configuration settings available in **Resource Config** vary depending on your IaaS.

- If you have the Load Balancers column in your Resource Config:
 - Enter the wildcard DNS entry attached to your load balancer into the Router row under Load Balancers.
- If you do not have the Load Balancers column in your Resource Config:
 - Navigate to the Networking section of the PCF Isolation Segment tile.
 - Specify Router IPs.
 - Attach the IP addresses to your load balancer manually.

3. If you are not using HAProxy for this isolation segment, set the number of Instances to **0**.

After Tile Configuration

After you configure the PCF Isolation Segment tile, perform the following steps:

1. Create the isolation segment in the Cloud Controller Database (CCDB) by following the procedure in the [Create an Isolation Segment](#) section of the [Managing Isolation Segments](#) topic.
2. Return to the Ops Manager Installation Dashboard.
3. Click **Review Pending Changes**, then **Apply Changes** to deploy the tile.

After the tile finishes deploying, see the [Managing Isolation Segments](#) topic for more information about managing an isolation segment.

Step 4: (Optional) Copy the Tile

If you want to create multiple isolation segments, perform the following steps to copy the PCF Isolation Segment tile with the Replicator tool:

1. Download the Replicator tool from the **Pivotal Cloud Foundry Isolation Segment** section of [Pivotal Network](#).
2. Navigate to the directory where you downloaded the Replicator tool.
3. Replicate the tile:

```
./replicator \
--name "YOUR-NAME" \
--path /PATH/TO/ORIGINAL.pivotl \
--output /PATH/TO/COPY.pivotl
```

Replace the values above with the following:

- `YOUR-NAME` : Provide a unique name for the new PCF Isolation Segment tile. The name must be ten characters or less and only contain alphanumeric characters, dashes, underscores, and spaces.
- `/PATH/TO/ORIGINAL` : Provide the absolute path to the original PCF Isolation Segment tile you downloaded from Pivotal Network.
- `/PATH/TO/COPY` : Provide the absolute path for the copy that the Replicator tool produces.

4. Follow the procedures in this topic using the new `.pivotal` file, starting with [Step 1](#).

Upgrading Replicated Tiles

For information about upgrading replicated Isolation Segment tiles, see [Upgrading Replicated Tiles](#).

Getting Started with Small Footprint PAS

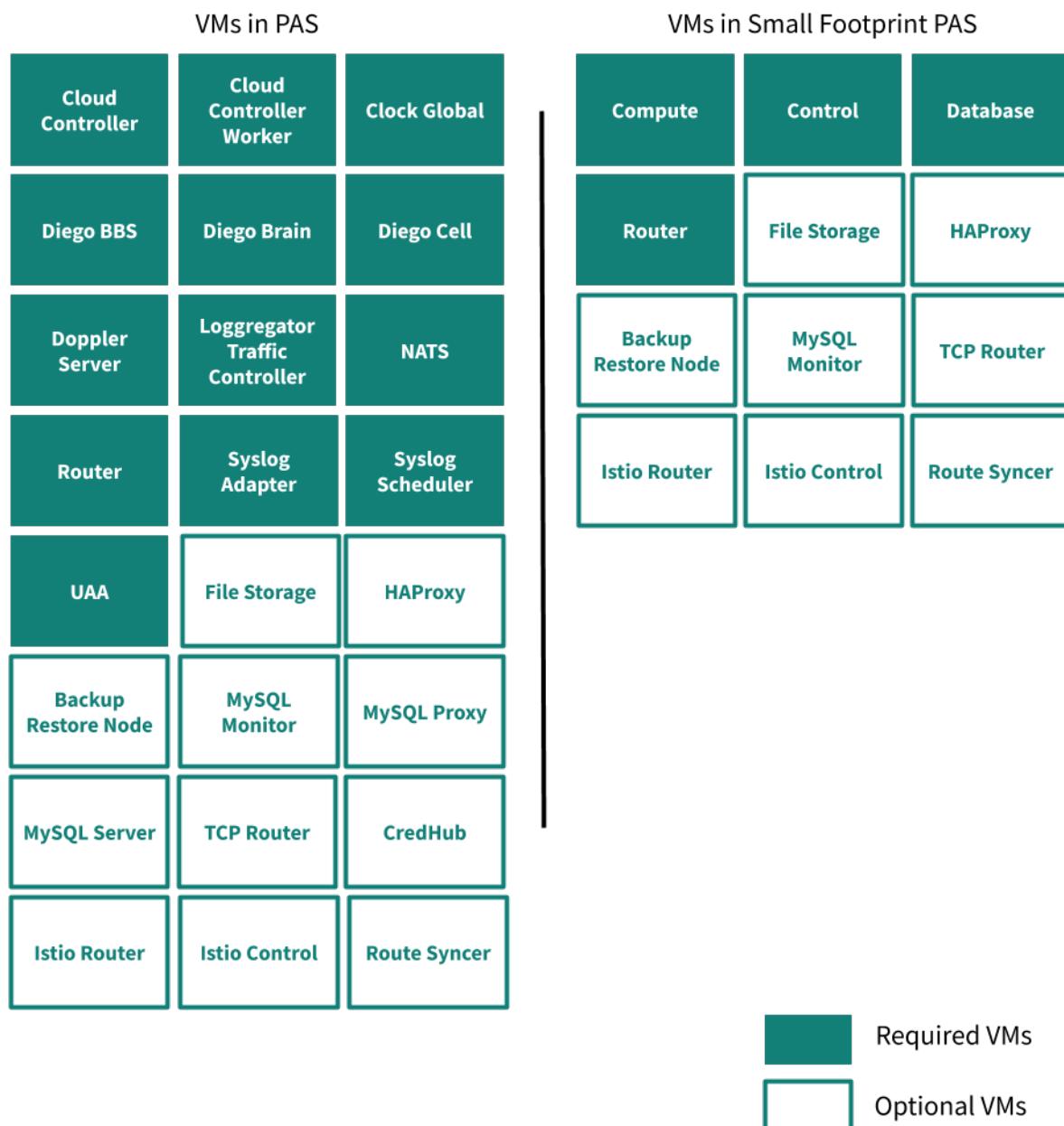
This topic describes the Small Footprint Pivotal Application Services (PAS) tile for Pivotal Cloud Foundry (PCF).

The Small Footprint PAS is a repackaging of the PAS components into a smaller deployment with fewer virtual machines (VMs). The [Limitations](#) section describes the limitations that come with a smaller deployment.

Differentiate Small Footprint PAS and PAS

A standard PAS deployment must have at least 13 VMs, but the Small Footprint PAS requires only four.

The following image displays a comparison of the number of VMs deployed by PAS and Small Footprint PAS.



Use Cases

Use the Small Footprint PAS tile for smaller PCF deployments on which you intend to host 2500 or fewer apps, as described in the [Limitations](#) section. If

you want to use Small Footprint PAS in a production environment, ensure the *Limitations* described below are not an issue in your use case.

Note: The Small Footprint PAS is compatible with PCF service tiles.

The Small Footprint PAS tile is also ideal for the following use cases:

- **Proof-of-concept installations:**
 - Deploy PCF quickly and with a small footprint for evaluation or testing purposes.
- **Sandbox installations:**
 - Use Small Footprint PAS as a PCF operator sandbox for tasks such as testing compatibility.
- **Service tile R&D:**
 - Test a service tile against Small Footprint PAS instead of a standard PAS deployment to increase efficiency and reduce cost.

Limitations

The Small Footprint PAS has the following limitations:

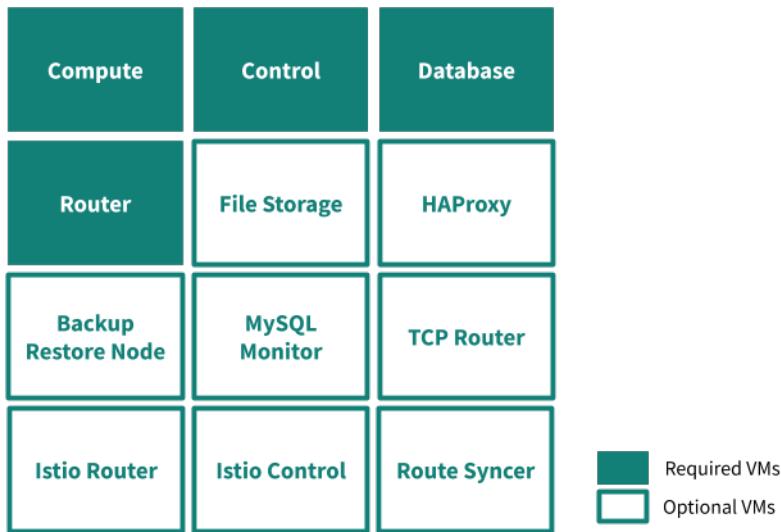
- **Number of app instances:**
 - The tile is not designed to support large numbers of app instances. You cannot scale the number of [Compute VMs](#) beyond 10 instances in the [Resource Config](#) pane. The Small Footprint PAS is designed to support 2500 or fewer apps.
- **Increasing platform capacity:**
 - You cannot upgrade the Small Footprint PAS to the standard PAS tile. If you expect platform usage to increase beyond the capacity of the Small Footprint PAS, Pivotal recommends using the standard PAS tile.
- **Management plane availability during tile upgrades:**
 - You may not be able to perform management plane operations like deploying new apps and accessing APIs for brief periods during tile upgrades. The management plane is colocated on the [Control VM](#).
- **App availability during tile upgrades:**
 - If you require availability during your upgrades, you must scale your [Compute VMs](#) to a highly available configuration. Ensure sufficient capacity exists to move app instances between Compute VM instances during the upgrade.

Architecture

You can deploy the Small Footprint PAS tile with a minimum of four VMs, as shown in the image below.

Note: The following image assumes that you are using an external blobstore.

VMs in Small Footprint PAS



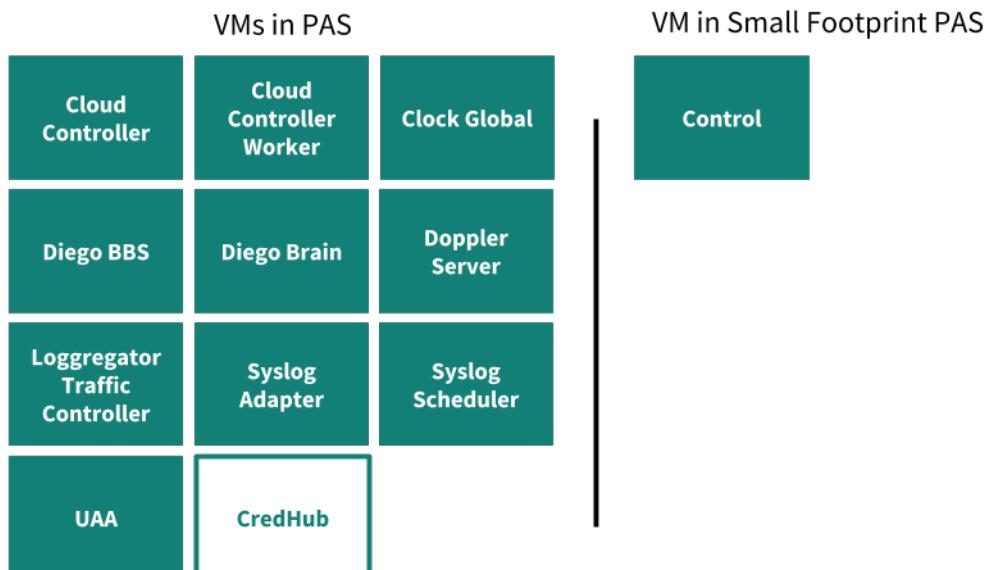
To reduce the number of VMs required for Small Footprint PAS, the *Control* and *Database* VMs include colocated jobs that run on a single VM in PAS. See the next sections for details.

For more information about the components mentioned on this page, see the *Architecture* section of the [Cloud Foundry Concepts](#) guide.

Control VM

The Control VM includes the PAS jobs that handle management plane operations, app lifecycles, logging, and user authorization and authentication. Additionally, all errands run on the Control VM, eliminating the need for a VM for each errand and significantly reducing the time it takes to run errands.

The following image shows all the jobs from PAS that are colocated on the Control VM in Small Footprint PAS.



Database VM

The database VM includes the PAS jobs that handle internal storage and messaging.

The following image shows all the jobs from PAS that are colocated on the Database VM in Small Footprint PAS.



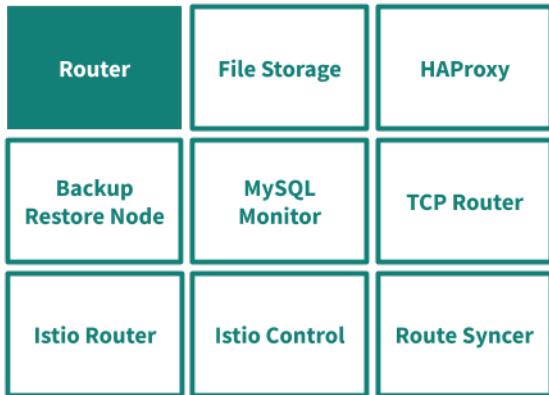
Compute VM

The *Compute* VM is the same as the Diego Cell VM in PAS.



Other VMs (Unchanged)

The following image shows the VMs performs the same functions in both versions of the PAS tile.



Requirements

The following topics list the minimum resources needed to run Small Footprint PAS and Ops Manager on the public IaaSes that PCF supports:

- [Installing PCF on AWS](#)
- [Installing PCF on Azure](#)
- [Installing PCF on GCP](#)
- [Installing PCF on vSphere](#)

Installing Small Footprint PAS

To install the Small Footprint PAS tile, follow the instructions for [Installing Pivotal Cloud Foundry](#) on your IaaS.

Follow the same installation and configuration steps as for PAS, with the following differences:

- Selecting a product in Pivotal Network:

- When you navigate to the PAS tile on [Pivotal Network](#), select the **Small Footprint** release.

- Configuring resources:

- The **Resource Config** pane in the Small Footprint PAS tile reflects the differences in VMs discussed in the *Architecture* section of this topic.
- Small Footprint PAS does not default to a highly available configuration like PAS does. It defaults to a minimum configuration. To make Small Footprint PAS highly available, scale the VMs to the following instance counts:

▪ Compute:	<input type="text" value="3"/>
▪ Control:	<input type="text" value="2"/>
▪ Database:	<input type="text" value="3"/>
▪ Router:	<input type="text" value="3"/>

- Configuring load balancers:

- If you are using an SSH load balancer, you must enter its name in the **Control VM** row of the **Resource Config** pane. There is no **Diego Brain** row in Small Footprint Runtime because the Diego Brain is colocated on the Control VM. You can still enter the appropriate load balancers in the **Router** and **TCP Router** rows as normal.

Troubleshooting Colocated Jobs Using Logs

If you need to troubleshoot a job that runs on the Control or Database VMs, do the following:

1. Follow the procedures in *Advanced Troubleshooting with the BOSH CLI* to log in to the BOSH Director for your deployment:

- a. [Gather Credential and IP Address Information](#)
- b. [SSH into Ops Manager](#)
- c. [Log in to the BOSH Director](#)

2. Use BOSH to list the VMs in your Small Footprint PAS deployment:

```
bosh -e MY-ENV -d MY-DEPLOYMENT vms
```

 **Note:** If you do not know the name of your deployment, you can run `bosh -e MY-ENV deployments` to list the deployments for your BOSH Director.

See the following example output:

```
$ bosh -e example-env -d example-deployment vms
Using environment 'example-env' as client 'ops_manager'

Task 182. Done

Deployment 'example-deployment'

Instance          Process State AZ     IPs      VM CID           VM Type
backup-restore/8fd07242-cf7c-4a4d-ba69-85fe078114f9 running us-central1-a 10.0.4.10 vm-6ec72a47-55b0-4767-78af-759f1f295183 micro
compute/01c6947d-477e-4605-9e6b-5d130a58c70c running us-central1-b 10.0.4.8 vm-ce14173c-d93e-414c-6830-afbe0c713fc5 xlarge.disk
compute/28045395-5048-4c8d-8363-e22f7b66847 running us-central1-c 10.0.4.9 vm-e3d8f696-5802-4552-4006-a1260563ed49 xlarge.disk
compute/2e3ed7dc-baa4-42ef-814d-980c6ab1c36b running us-central1-a 10.0.4.7 vm-6d34e53-71f3-4741-674a-c42c4df9e559 xlarge.disk
control/12b1b027-7ffd-43ca-9dc9-7f4ff204d86a running us-central1-a 10.0.4.6 vm-9760b74e-e13e-4483-79b6-78ab3818b628 xlarge
ha_proxy/b0587c68-45a8-40e2-94d3-5d2ffcdaf858 running us-central1-a 10.0.4.11 vm-27d62bfc-af6d-4c8b-6e2a-cbba09eddd1e micro
mysql_monitor/5185d04e-e038-4664-a26a-d160d295a7f running us-central1-a 10.0.4.15 vm-6d215888-913b-44a3-4db3-52329c5ada53 micro
router/2043b22d-0c3b-4a02-873f-80a724c3ed08 running us-central1-a 10.0.4.12 vm-2b7cf5f4-5926-4f70-6e47-d994a6eff93b micro
router/72b54793-e0d0-4301-8932-76da5375e654 running us-central1-c 10.0.4.14 vm-e77bcdff1-0e26-46cd-7783-6d766f4c5098 micro
router/e3d2ab7b-6191-46bb-ab62-c1db7268a942 running us-central1-b 10.0.4.13 vm-3c84523b-1988-475e-49e8-de80fd76c656 micro
database/681bcad5-fa8b-4cf1-912f-45140d96123f running us-central1-a 10.0.4.5 vm-e3cded4f-cf47-499f4c96-992b3c6eb9c large.disk
tcp_router/61a06e83-a62b-4afb-b452-441dc2dc1e4c running us-central1-a 10.0.4.17 vm-cc1f0a62-409f-47f9-58b0-0b8f46cf9ac0 micro

13 vms
```

Succeeded

3. Use BOSH to SSH into one of the Small Footprint PAS VMs.

```
bosh -e MY-ENV -d MY-DEPLOYMENT ssh VM-NAME/GUID
```

For example, to SSH into the Control VM, run the following:

```
$ bosh -e example-env -d example-deployment ssh control/12b1b027-7ffd-43ca-9dc9-7f4ff204d86a
```

4. Run `sudo su` to act as super user.

5. Use `monit` to list the processes running on the VM.

```
monit summary
```

See the following example output that lists the processes running on the Control VM. The processes listed reflect the colocation of jobs as outlined in the *Architecture* section of this topic.

```
control/12b1b027-7ffd-43ca-9dc9-7f4ff204d86a:/var/vcap/bosh_ssh/bosh_f9d2446b18b445e# monit summary
The Monit daemon 5.2.5 uptime: 5d 21h 10m

Process 'bbs'          running
Process 'metron_agent' running
Process 'locket'        running
Process 'route_registrar' running
Process 'policy-server' running
Process 'silk-controller' running
Process 'uaa'           running
Process 'statsd_injector' running
Process 'cloud_controller_ng' running
Process 'cloud_controller_worker_local_1' running
Process 'cloud_controller_worker_local_2' running
Process 'nginx_cc'      running
Process 'routing-api'   running
Process 'cloud_controller_clock' running
Process 'cloud_controller_worker_1' running
Process 'auctioneer'    running
Process 'cc_uploader'   running
Process 'file_server'   running
Process 'nsync_listener' running
Process 'ssh_proxy'     running
Process 'tps_watcher'   running
Process 'stager'        running
Process 'loggregator_trafficcontroller' running
Process 'reverse_log_proxy' running
Process 'adapter'        running
Process 'doppler'        running
Process 'syslog_drain_binder' running
System 'system_localhost' running
```

6. To access logs, navigate to `/var/vcap/sys/log`:

```
cd /var/vcap/sys/log
```

7. Run `ls` to list the log directories for each process. See the following example output from the Control VM:

```
control/12b1b027-7ffd-43ca-9dc9-7f4ff204d86a:/var/vcap/sys/log# ls
adapter  cloud_controller_clock  file_server      nginx_cc      route_registrar statsd_injector  uaa_ctl.err.log
auctioneer  cloud_controller_ng  locket         nginx_newrelic_plugin routing-api  syslog_drain_binder uaa_ctl.log
bbs       cloud_controller_worker  loggregator_trafficcontroller nsync      silk-controller syslog_forwarder
cc_uploader metron_agent        policy-server   ssh_proxy    tps
cfdot     doppler               monit          reverse_log_proxy stager      uaa
```

8. Navigate to the directory of the process that you want to view logs for. For example, for the Cloud Controller process, run `cd cloud_controller_ng/`. From the directory of the process, you can list and view its logs. See the following example output:

```
control/12b1b027-7ffd-43ca-9dc9-7f4ff204d86a:/var/vcap/sys/log/cloud_controller_ng# ls
cloud_controller_ng_ctl.err.log  cloud_controller_ng.log.2.gz  cloud_controller_ng.log.6.gz  drain      pre-start.stdout.log
cloud_controller_ng_ctl.log      cloud_controller_ng.log.3.gz  cloud_controller_ng.log.7.gz  post-start.stderr.log
cloud_controller_ng.log          cloud_controller_ng.log.4.gz  cloud_controller_worker_ctl.err.log  post-start.stdout.log
cloud_controller_ng.log.1.gz      cloud_controller_ng.log.5.gz  cloud_controller_worker_ctl.log  pre-start.stderr.log
```

Release Notes

The Small Footprint PAS tile releases alongside the PAS tile. For more information, see the [PAS Release Notes](#).

Upgrading Pivotal Cloud Foundry

This topic describes upgrading Pivotal Cloud Foundry (PCF) to v2.6. The upgrade procedure below describes upgrading Ops Manager, Pivotal Application Service (PAS), and product tiles.

 **Breaking Changes:** Read the [Release Notes](#) and [Breaking Changes](#) for this release, including the Known Issues sections, before starting the upgrade process.

 **Warning:** Pivotal does not recommend that you skip minor versions when upgrading PCF. Skipping minor versions when upgrading PCF may result in breaking changes. To avoid additional breaking changes, upgrade PCF to the minor version that directly follows your current version of PCF.

The apps in your deployment continue to run during the upgrade. However, you cannot write to your deployment or make changes to apps during the upgrade.

For details about how upgrading PCF impacts individual PAS components, see [What Happens During PAS Upgrades](#).

Prepare to Upgrade

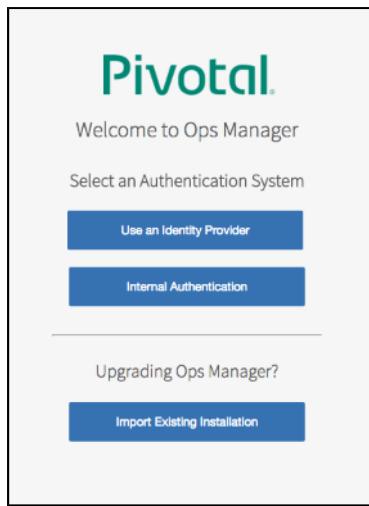
If you have not already, complete the steps in the [Upgrade Preparation Checklist for PCF v2.6](#).

Upgrade Ops Manager and Installed Products to v2.6

Follow these steps to upgrade Ops Manager and Installed Products to PCF v2.6.

Import Installation to Ops Manager v2.6 VM

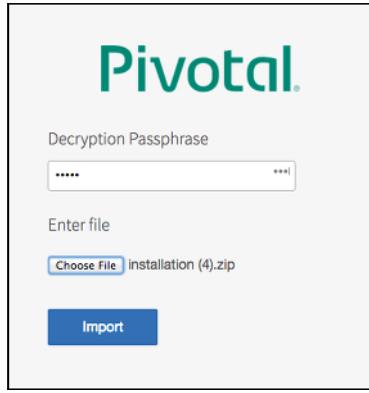
1. Download the Ops Manager VM Template v2.6 from the [Pivotal Network](#) site.
2. Record the FQDN address of the existing Ops Manager VM.
3. To avoid conflicts, power off the existing Ops Manager VM.
4. Deploy the new Ops Manager VM by following the steps in one of these topics:
 - [AWS: Upgrading BOSH Director on AWS](#)
 - [Azure: Upgrading BOSH Director on Azure](#)
 - [GCP: Upgrading BOSH Director on GCP](#)
 - [OpenStack: Deploying Ops Manager to OpenStack](#)
 - [vSphere: Deploying Ops Manager on vSphere](#)
5. When redirected to the [Welcome to Ops Manager](#) page, select **Import Existing Installation**.



- When prompted, enter the **Decryption Passphrase** for this Ops Manager installation. You set this passphrase during your initial installation of Ops Manager.

Note: If lost, the **Decryption Passphrase** cannot be recovered.

- Click **Choose File** and browse to the installation ZIP file exported in the [Export Your Installation](#) step of your upgrade preparation.



- Click **Import**.

Note: Some browsers do not provide feedback on the status of the import process, and might appear to hang.

- A **Successfully imported installation** message appears upon completion.



Import New PAS and Product Tiles

After upgrading to Ops Manager v2.6, upgrade your product versions:

- Import the product file to your Ops Manager [Installation Dashboard](#).
- Hover over the product name in [Available Products](#) and click [Add](#).
- Click the newly-added tile to review any configurable options.
- (Optional) If you use other service tiles, you can upgrade them following the same procedure. See the [Upgrading PAS and Other Pivotal Cloud](#)

[Foundry Products](#) topic for more information.

Perform Your Upgrade

⚠ warning: If the installation fails or returns errors, contact [Support](#). Do not attempt to roll back the upgrade by restarting the previous (v2.3) Ops Manager VM.

1. Navigate to the Ops Manager [Installation Dashboard](#).
2. Click **Review Pending Changes**, then **Apply Changes**. This immediately imports and applies upgrades to all tiles in a single transaction.
3. Click each service tile, select the **Status** tab, and confirm that all VMs appear and are in good health.
4. After confirming that the new installation functions correctly, remove the Ops Manager v2.3 VM.

Monitor Upgrade

During the upgrade, you can do the following to monitor your PCF foundation and help troubleshoot any issues.

Check Status and Performance

Note/Reason:

Monitor the progress of the upgrade, checking the status of the foundation at various locations.

Product/Component:

PCF

Pivotal recommends live-monitoring your upgrade with PCF Healthwatch, which monitors and alerts on the current health, performance, and capacity of PCF. Healthwatch captures, calculates, stores, visualizes, and alerts on PCF platform metrics, including:

- BOSH-reported VM metrics
- Runtime component metrics
- [Key Performance Indicators](#) and [Key Scaling Indicators](#)
- Healthwatch-generated [super metrics](#)

For more information, see the [PCF Healthwatch documentation](#).

If you are not using PCF Healthwatch, you can do some or all of the following to monitor upgrade progress:

- Run BOSH CLI status checks:
 - `bosh -e ALIAS task TASK_NUMBER`
 - `bosh -e ALIAS vms --vitals, bosh -e ALIAS instances --ps`

- Check app availability
- Run cf CLI Commands
- Check the availability of the Ops Manager UI
- Check NAS performance (if using NAS)
- Check vSphere performance (if on vSphere)

Check Diego State

Note/Reason:

See the [cf-dot documentation](#) on GitHub for details.

Product/Component:

PAS

Use the CF Diego Operator Toolkit (cf-dot) to check Diego component instance count by current state.

(Optional) Take Snapshots of Storage Metrics

Reason:

Pivotal recommends this if you have a large foundation and have experienced storage issues in the past.

Product/Component:

PCF

Periodically take snapshots of storage metrics.

(Optional) Collect Logs

Reason:

This information helps determine the cause of upgrade issues.

Product/Component:

PCF

If you encounter problems during upgrade, collect the following information:

- All job logs
- Task debug logs for VM upgrade tasks
- Installation log from Ops Manager

After Upgrade

After your upgrade, do the following to prepare for use of your new environment, check its health status, and clean up.

Re-Create BOSH Alias

Re-create your alias using BOSH: `bosh alias-env ALIAS -e DIRECTOR_IP`

- **Reason:** To log in to BOSH after upgrading PCF, you need to re-create your alias.
- **Product/Component:** BOSH

Install New cf CLI

Download the version of the Cloud Foundry Command Line Interface (cf CLI) packaged with the [PAS](#) tile on Pivotal Network.

Check the Health of Your Deployment

1. Run BOSH CLI commands to check system status:

- a. `bosh -e ALIAS -d DEPLOYMENT_NAME instances --ps`
- b. `bosh -e ALIAS vms --vitals`
- c. `bosh -e ALIAS -d DEPLOYMENT_NAME cck --report`

Note/Reason: To ensure that all jobs and process are running as expected

Product/Component: PCF and all tiles

2. Push and horizontally scale a test application.

- **Reason:** This is a performance test for PAS.
- **Product/Component:** PCF

3. If you are running PAS MySQL as a cluster, run the `mysql-diag` tool to validate health of the cluster.

- **Reason:** See the BOSH CLI v2 instructions in the [Running mysql-diag](#) topic.
- **Product/Component:** PAS

Check Resource Settings

If you added custom **VM Type** or **Persistent Disk Type** options, ensure that these values are correctly set and were not overwritten.

- **Note/Reason:** Verify that the Ops Manager **Resource Config** pane still lists your custom options.
- **Product/Component:** PCF

Run BOSH Clean-Up

Run `bosh -e ALIAS clean-up --all` to clean up old stemcells, releases, orphaned disks, and other unused resources.

- **Product/Component:** Tiles

What Happens During PAS Upgrades

This topic explains what happens to Pivotal Application Service (PAS) components and apps during a PAS upgrade.

BOSH Drains Diego Cell VMs

During a PAS upgrade, BOSH drains all Diego cell VMs that host app instances. BOSH manages this process by upgrading a batch of cells one at a time.

When BOSH triggers an upgrade, each upgrading Diego cell enters evacuation mode. In evacuation mode, BOSH stops Diego cells and then schedules replacements for its app instances.

For more information, see the [Specific Guidance for Diego Cells](#) section of the *Configuring PAS for Upgrades* topic.

cf push Can Become Unavailable

`cf push` is mostly available for the duration of a PAS upgrade. However, `cf push` can become unavailable when a single VM is in use or during BOSH Backup and Restore (BBR).

For more information, see [cf push Availability During Pivotal Application Service Upgrades](#).

PAS Components Upgrade

This section describes the order in which Ops Manager upgrades components and runs tasks during a full platform upgrade. It also explains how the scale of different Pivotal Application Service (PAS) components affects uptime during upgrades, and which components are scalable.

When performing an upgrade, Ops Manager first upgrades individual components, and then runs one-time tasks.

1. The [Components](#) section describes how Ops Manager upgrades PAS components and explains how individual component upgrades affect broader PAS capabilities.
2. The [One-Time Tasks](#) section lists the tasks that Ops Manager runs after it upgrades the PAS components.

Components

Ops Manager upgrades PAS components in a fixed order that honors component dependencies and minimizes downtime and other system limitations during the upgrade process.

The type and duration of downtime and other limitations that you can expect during a PAS upgrade reflect the following:

- Component instance scaling. See [How Single-Component Scaling Affects Upgrades](#)
- Component upgrade order. See [Component Upgrade Order and Behavior](#)

How Single-Component Scaling Affects Upgrades

In [Pivotal Cloud Foundry](#) (PCF) Ops Manager, the Pivotal Application Service (PAS) tile > **Resource Config** pane shows the components that the BOSH Director installs:

- **Scalable** component fields let you select the instance count from a range of settings or enter a custom value.
- **Unscalable** component fields allow a maximum of one instance.

When a component is scaled at a single instance, it can experience downtime and other limitations while the single VM restarts. This behavior might be acceptable for a test environment. To avoid downtime in a production environment, you should scale any scalable components, such as HAProxy, Router, and Diego cells, to more than one instance.

For more information about how the scale of each component affects upgrade behavior, see the [Component Upgrade Order and Behavior](#) table below.

Note: A full Ops Manager upgrade may take close to two hours, and you will have limited ability to deploy an application during this time.

Component Upgrade Order and Behavior

The table below lists components in the order that Ops Manager upgrades each. It also lists which components are scalable and explains how component downtime affects PAS app and control availability. The table includes the following columns:

- **Scalable:** Indicates whether the component is scalable above a single instance.

Note: For components marked with a checkmark in this column, we recommend that you change the preconfigured instance value of `1` to a value that best supports your production environment. For more information about scaling a deployment, refer to the [Scaling Cloud Foundry topic](#).

- **Extended Downtime:** Indicates that if there is only one instance of the component, that component is unavailable for up to five minutes during an Ops Manager upgrade.
- **Downtime Affects...:** Indicates the plane of the PAS platform that component downtime affects, if the component is scaled at single instance:
 - **Apps:** Downtime can affect app availability.
 - **Platform:** Apps remain available during component downtime, but you cannot push, stage, or restart apps, or run other Cloud Foundry command-line interface (`cf` CLI) commands.
- **Other Limitations and Information:** Provides the following information:
 - Component availability, behavior, and usage during an upgrade
 - Guidance on disabling the component before an upgrade

Component	Scalable	Extended Downtime	Downtime Affects...		Other Limitations and Information
			Apps	Platform	
NATS	✓	✓			
File Storage		✓		✓	
MySQL Proxy	✓	✓		✓	The MySQL Proxy is responsible for managing failover of the MySQL Servers. If the Proxy becomes unavailable, then access to the MySQL Server could be broken.
MySQL Server	✓	✓		✓	The MySQL Server is responsible for persisting internal databases for the platform. If the MySQL Server becomes unavailable, then platform services that rely upon a database (Cloud Controller, UAA) will also become unavailable. See Effects of MySQL Downtime for details.
Backup Restore Node					
Diego BBS	✓	✓		✓	
UAA	✓			✓	If a user has an active authorization token prior to performing an upgrade, they can still log in using either a UI or the CLI.
Cloud Controller	✓	✓		✓	
HAProxy	✓	✓	✓		HAProxy is used to load-balance incoming requests to the Router. If HAProxy is unavailable, you may lose the ability to make requests to applications unless there is another routing path from your load balancer to the Router.
Router	✓	✓	✓		The Router is responsible for routing requests to their application containers. If the Router is not available, then applications cannot receive requests.
MySQL Monitor					
Clock Global	✓				
Cloud Controller Worker	✓	✓			
Diego Brain	✓	✓		✓	
Diego Cell	✓	✓	✓	✓	If you only have one Diego Cell, upgrading causes downtime for all apps that run on it. These include apps pushed with <code>cf push</code> and apps automatically installed by PAS, like Apps Manager and the App Usage Service.

Loggregator Trafficcontroller	✓				Ops Manager operators experience 2-5 minute gaps in logging.
Syslog Adapter	✓				
Syslog Scheduler	✓				
Doppler Server	✓				Ops Manager operators experience 2-5 minute gaps in logging.
TCP Router (if enabled)	✓				
CredHub	✓	✓	✓	✓	App downtime for apps that use secure credentials. Platform downtime for cf CLI commands such as <code>bind-service</code> and <code>unbind-service</code> applied to services configured with CredHub.
Istio Control		✓		✓	
Istio Router	✓	✓	✓		Downtime for this component only affects routes on mesh domains.
Route Syncer	✓	✓		✓	Downtime for this component only affects routes on mesh domains.

One-Time Tasks

After Ops Manager upgrades components, it performs system checks and launches UI apps and other PAS components as Cloud Foundry apps. These tasks run in the following order:

1	Apps Manager Errand - Push Apps Manager
2	Smoke Test Errand - Run smoke tests
3	Usage Service Errand - Push Usage Service app
4	Notifications Errand - Push Notifications app
5	Notifications UI Errand - Push Notifications UI
6	App Autoscaler Errand - Push App Autoscaler
7	App Autoscaler Smoke Test Errand - Run smoke tests against App Autoscaler
8	Register Autoscaling Service Broker
9	Destroy Autoscaling Service Broker
10	Bootstrap Errand - Recover MySQL cluster
11	MySQL Rejoin Unsafe Errand

Upgrading Installation Example

For sample performance measurements of an upgrading Cloud Foundry installation, see [Upgrade Load Example: Pivotal Web Services ↗](#).

Upgrade Preparation Checklist for PCF v2.6

This topic serves as a checklist for preparing to upgrade Pivotal Cloud Foundry (PCF) from v2.5 to v2.6.

This topic contains important preparation steps that you must follow before beginning your upgrade. Failure to follow these instructions may jeopardize your existing deployment data and cause the upgrade to fail.

After completing the steps in this topic, you can continue to [Upgrading Pivotal Cloud Foundry](#).

⚠ Warning: Pivotal does not recommend that you skip minor versions when upgrading PCF. Skipping minor versions when upgrading PCF may result in breaking changes. To avoid additional breaking changes, upgrade PCF to the minor version that directly follows your current version of PCF.

Back Up Your PCF Deployment

Pivotal recommends backing up your PCF deployment before upgrading, to restore in the case of failure. To do this, follow the instructions in the [Backing Up Pivotal Cloud Foundry with BBR](#) topic.

Find Your Decryption Passphrase for Ops Manager

To complete the Ops Manager upgrade, you must have your Ops Manager decryption passphrase. You defined this decryption passphrase during the initial installation of Ops Manager.

Review Changes in PCF v2.6

Review each of the following links to understand the changes in the new release, such as new features, known issues, and breaking changes.

- **Release Notes**
 - [Ops Manager Release v2.6 Release Notes ↗](#)
 - [PAS v2.6 Release Notes ↗](#)
- **Known Issues**
 - [Ops Manager v2.6 Known Issues ↗](#)
 - [PAS v2.6 Known Issues ↗](#)
- **Breaking Changes**
 - [PCF v2.6 Breaking Changes ↗](#)
- **KPI Changes**
 - [KPI Changes ↗](#)
- **Diego Network Communications:**
 - [Diego Network Communications ↗](#)

Restage Windows Apps

The `windows2016` stack is deprecated in PAS for Windows v2.6.

Before upgrading to PAS for Windows v2.6 from PAS for Windows v2.5, all Windows apps that were pushed with the `windows2016` stack must be restaged to use the `windows` stack.

For more information, see [Deprecation of the windows2016 Stack](#) in the PAS for Windows v2.6 Release Notes.

If you are running PAS for Windows v2.4.4 or later on PCF v2.5, you can also restage your apps before upgrading to PAS for Windows v2.5. For more

information about upgrading to PAS for Windows v2.5, see [Upgrade to PAS for Windows v2.5](#) below.

Update Tiles and Add-Ons

The following section describes changes you must make to your product tiles and add-ons before upgrading PCF.

Upgrade to PAS for Windows v2.5

If you are running PAS for Windows v2.4.4 or later on PCF v2.5 and want to upgrade to PCF v2.6, you must upgrade to PAS for Windows v2.5 before upgrading to PCF v2.6. To download PAS for Windows, navigate to [Pivotal Network](#).

Upgrade Stemcell to v2019.3 or Later for PAS for Windows v2.5

If you are running PAS for Windows v2.5 using Windows stemcell v2019.2, you must upgrade your stemcell to v2019.3 or later before you upgrade to Ops Manager v2.6.

 **Warning:** Windows stemcell v2019.2 is not compatible with Ops Manager v2.6.

Review Service Tile Compatibility

Before you upgrade, check whether the service tiles that you currently have are compatible with the new version of PCF.

To check PCF versions supported by a service tile, either from Pivotal or a Pivotal partner:

- Navigate to the tile's download page on [Pivotal Network](#).
- Select the tile version in the **Releases** dropdown.
- See the **Depends On** section under **Release Details**. For more information, refer to the tile's release notes.

If the currently-deployed version of a tile is not compatible with PCF v2.6, you must upgrade the tile to a compatible version before you upgrade PCF. You do not need to upgrade tiles that are compatible with both PCF v2.5 and v2.6.

Some partner service tiles may be incompatible with PCF v2.6. Pivotal works with partners to ensure their tiles are updated to work with the latest versions of PCF. For more information about which partner service release compatibility, review the **Depends On** section of the partners tile download page, the partners services release documentation in [Pivotal Documentation](#), or contact the partner organization that produces the service tile.

The [Product Compatibility Matrix](#) provides an overview of which PCF versions support which versions of the most popular service tiles from Pivotal.

Environment Details

Pivotal provides the empty table below as a model to print out or adapt for recording and tracking the tile versions that you have deployed in all of your environments.

		Sandbox	Non-Prod	Prod	Other...
Pivotal Cloud Foundry	Ops Manager				
	Pivotal Application Service (PAS)				
Pivotal Cloud Foundry Services	MySQL v2				
	Redis				
	RabbitMQ				
	Single Sign On (SSO)				
	Spring Cloud Services				
	Concourse				
	...				
Pivotal Cloud Foundry Partner Services	New Relic				

	...	Sandbox	Non-Prod	Prod	Other...
--	-----	---------	----------	------	----------

Upgrade Services Tiles

Upgrade all service tiles to versions that are compatible with PCF v2.6. Service tiles are add-on products you install alongside your runtime. For example, MySQL for PCF, PCF Healthwatch, and RabbitMQ are service tiles.

Do not upgrade runtime tiles, such as PAS, PAS for Windows, or Pivotal Container Service (PKS), at this time.

Review the [Compatibility Matrix](#) and tile documentation to check version compatibility.

Configure BOSH Director

With each release of a new PCF version, BOSH Director may require specific updates before upgrading to the new version. See the following for what action to take before upgrading to PCF v2.6:

Check Required Machine Specifications

Check the required machine specifications for Ops Manager v2.6. These specifications are specific to your IaaS. If these specifications do not match your existing Ops Manager, modify the values of your Ops Manager VM instance. For example, if the boot disk of your existing Ops Manager is 50 GB and the new Ops Manager requires 100 GB, then increase the size of your Ops Manager boot disk to 100 GB.

Configure PAS

With each release of a new PCF version, PAS may require specific updates before upgrading to the new version. See the following sections for what action to take before upgrading to PCF v2.6:

Remove Healthcheck from SSH Load Balancer

If your SSH load balancer is configured with an HTTP healthcheck, remove the healthcheck before upgrading. This is the load balancer specified in the **Diego Brain** row of the **Resource Config** pane.

For more information, see [HTTP Healthcheck Server Disabled for SSH Proxy](#).

(Optional) Disable Unused Errands

To save upgrade time, you can disable unused PAS post-deploy errands. See the [Post-Deploy Errands](#) section of the Errands topic for details. Only disable these errands if your environment does not need them.

In some cases, if you have previously disabled lifecycle errands for any installed product to reduce deployment time, you may want to re-enable these errands before upgrading. For more information, see the [Adding and Deleting Products](#) topic.

Configure Diego Cell Garbage collection

In the PAS tile **Application Container** pane, if **Docker Images Disk-Cleanup Scheduling on Cell VMs** is set to **Clean up disk-space once threshold is reached** and the value of **Threshold of Disk-Used (MB)** is different than the default of **10240**, then you must specify a new threshold.

If the scheduling is set to **Never clean up...** or **Routinely clean up...**, or the threshold value is set to the default of **10240**, then no action is necessary. All thresholds will migrate to a sensible value when you upgrade PAS. For more information see [Options for Disk Cleanup](#).

Check OS Compatibility of BOSH-Managed Add-Ons and Tiles

Before upgrading to PCF v2.6, operators who have deployed any PCF add-ons such as IPsec for PCF, ClamAV for PCF, or File Integrity Monitoring for PCF and who have deployed or are planning to deploy PAS for Windows must modify the add-on manifest to specify a compatible OS stemcell. For more

information, see [Pivotal Application Service for Windows](#).

For example, **File Integrity Monitor for PCF (FIM)** is not supported on Windows. Therefore, the manifest must use an `include` directive to specify the target OS stemcell of `ubuntu-trusty` and `ubuntu-xenial`.

 **Note:** To upgrade to a Xenial stemcell, see the documentation for each add-on and follow the instructions.

To update an add-on manifest, do the following:

1. Locate your existing add-on manifest file. For example, for FIM, locate the `fim.yml` you uploaded to the Ops Manager VM.
2. Modify the manifest to include following `include` directive to your manifest:

```
include:  
stemcell:  
- os: ubuntu-xenial
```

3. Upload the modified manifest file to your PCF deployment. For example instructions, see [Create the FIM Manifest](#).

If you use any other BOSH-managed add-ons in your deployment, you should verify OS compatibility for those component as well. For more information about configuring BOSH add-on manifests, see the [BOSH documentation](#).

Check Backup and Restore External Blobstore Add-On

If you have enabled external blobstore backups for an Azure Blobstore using the [Blobstore Add-On](#), you must update your runtime configuration to remove the `sdk-preview` add-on before upgrading. If you do not remove this job, upgrading PAS fails with the error:

```
Preparing deployment: Preparing deployment (00:00:01)  
L Error: Colocated job 'azure-blobstore-backup-restorer' is already added to the instance group 'backup-restore'.
```

After removing this job from your runtime configuration, ensure that the **Enable backup and restore** checkbox is enabled in the PAS tile > **File Storage** pane.

Check Certificate Authority Expiration Dates

Depending on the requirements of your deployment, you may need to rotate your Certificate Authority (CA) certificates. The non-configurable certificates in your deployment expire every two years. You must regenerate and rotate them so that critical components do not face a complete outage.

 **Note:** PCF uses SHA-2 certificates and hashes by default. You can convert existing SHA-1 hashes into SHA-2 hashes by rotating your Ops Manager certificates using the procedure described in the [Regenerating and Rotating Non-Configurable TLS/SSL Certificates](#) section of *Managing TLS Certificates*.

To retrieve information about all the RSA and CA certificates for the BOSH Director and other products in your deployment, you can use the `GET /api/v0/deployed/certificates?expires_within=TIME` request of the Ops Manager API.

In this request, the `expires_within` parameter is optional. Valid values for the parameter are `d` for days, `w` for weeks, `m` for months, and `y` for years. For example, to search for certificates expiring within one month, replace `TIME` with `1m`:

```
$ curl "https://OPS-MAN-FQDN/api/v0/deployed/certificates?expires_within=1m" \  
-X GET \  
-H "Authorization: Bearer UAA_ACCESS_TOKEN"
```

For information about regenerating and rotating CA certificates, see [Managing TLS Certificates](#).

Check the Capacity of Your Deployment

The following sections describe steps for ensuring your deployment has adequate capacity to perform the upgrade.

Confirm Adequate Disk Space

Confirm that you have adequate disk space for your upgrades. You need at least 20 GB of free disk space to upgrade PCF Ops Manager and Pivotal Application Service. If you plan to upgrade other products, the amount of disk space required depends on how many tiles you plan to deploy to your upgraded PCF deployment.

To check current persistent disk usage, select the **BOSH Director** tile from the **Installation Dashboard**. Select **Status** and review the value of the **PERS. DISK** column. If persistent disk usage is higher than 50%, select **Settings > Resource Config**, and increase your persistent disk space to handle the size of the resources. If you do not know how much disk space to allocate, set the value to at least **100 GB**.

Check Diego Cell RAM and Disk

Check that Diego cells have sufficient available RAM and disk capacity to support app containers.

The KPIs that monitor these resources are:

- `rep.CapacityRemainingMemory`
- `rep.CapacityRemainingDisk`

Adjust Diego Cell Limits

If needed, adjust the maximum number of Diego cells that the platform can upgrade simultaneously, to avoid overloading the other cells. See [Limit PCF Component Instances During Restart](#).

For PCF v1.10 and later, the maximum number of cells that can update at once, `max_in_flight` is 4%. This setting is configured in the BOSH manifest's Diego cell definition. For more information, see the [Prevent Overload](#) section.

Review the [Diego Cell Metrics](#) section of the KPI topic for more information about these KPIs.

Review File Storage IOPS and Other Upgrade Limiting Factors

During the PCF upgrade process, a large quantity of data is moved around on disk.

To ensure a successful upgrade of PCF, verify that your underlying PAS file storage is performant enough to handle the upgrade. For more information about the configurations to evaluate, see [Upgrade Considerations for Selecting Pivotal Cloud Foundry Storage](#).

In addition to file storage IOPS, consider additional existing deployment factors that can impact overall upgrade duration and performance:

Factor	Impact
Network latency	Network latency can contribute to how long it takes to move app instance data to new containers.
Number of ASGs	A large number of Application Security Groups in your deployment can contribute to an increase in app instance container startup time.
Number of app instances and application growth	A large increase in the number of app instances and average droplet size since the initial deployment can increase the upgrade impact on your system.

To review example upgrade-related performance measurements of an existing production Cloud Foundry deployment, see the [Pivotal Web Services Performance During Upgrade](#) topic.

Run BOSH Clean-Up

Run `bosh -e ALIAS clean-up --all` to clean up old stemcells, releases, orphaned disks, and other resources before upgrade. This cleanup helps prevent the product and stemcell upload process from exceeding the BOSH Director's available persistent disk space.

Check the Health of Your Deployment

The following sections describe steps for ensuring your deployment is healthy before you perform the upgrade.

Collect Foundation Health Status

For collecting foundation health status, Pivotal recommends PCF Healthwatch, which monitors and alerts on the current health, performance, and capacity of PCF. For more information, see the [PCF Healthwatch documentation](#).

If you are not using PCF Healthwatch, you can do some or all of the following to collect foundation health status:

- If your PCF deployment has external metrics monitoring set up, verify that VM CPU, RAM, and disk use levels are within reasonable levels.
- Run BOSH CLI commands to check system status:
 - `bosh -e ALIAS -d DEPLOYMENT_NAME instances --ps`.
bosh instances with the flags `--ps`, `--vitals`, or `--failing` highlights individual job failure.
 - `bosh -e ALIAS vms --vitals`. This reveals VMs with high CPU, high memory, high disk utilization, and with `state != running`.
 - `bosh -e ALIAS -d DEPLOYMENT_NAME cck --report`
- Check Ops Manager GUI each PAS/Tiles the status page for CPU/RAM/DISK utilization
- Validate Ops Manager persistent disk usage is below 50%. If not, follow the procedure in [Confirm Adequate Disk Space](#).

(Optional) Check the logs for errors before proceeding with the upgrade. For more information, see [Viewing Logs in the Command Line Interface](#).

Push and Scale a Test App

Check that a test app can be pushed and scaled horizontally, manually or through automated testing. This check ensures that the platform supports apps as expected before the upgrade.

Validate MySQL Cluster Health

If you are running PAS MySQL as a cluster, run the `mysql-diag` tool to validate health of the cluster.

See the BOSH CLI v2 instructions in the [Running mysql-diag](#) topic.

Review Pending and Recent Changes

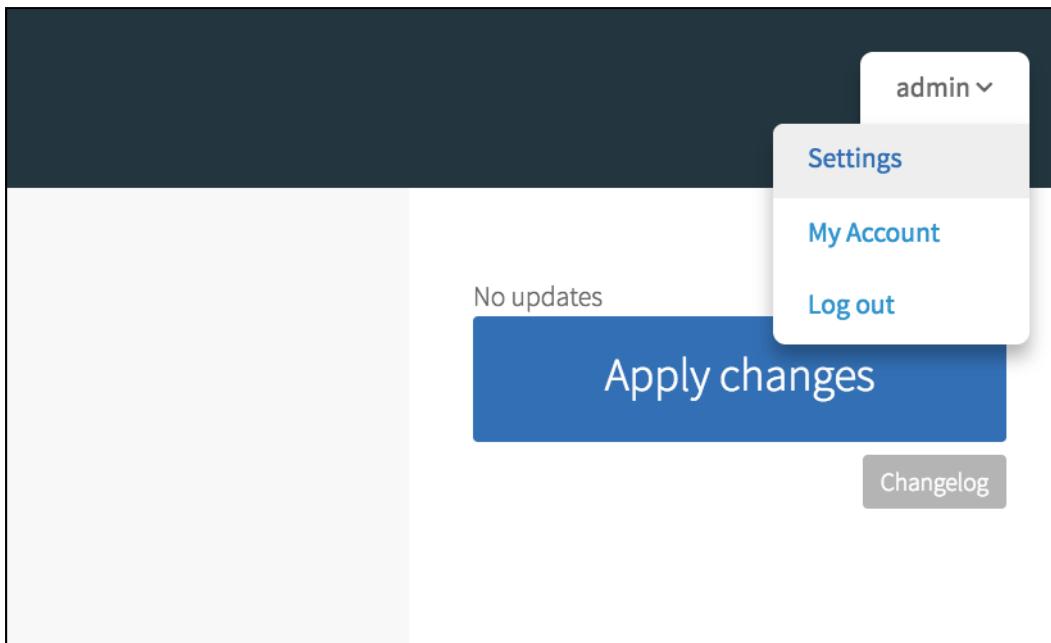
1. Confirm there are no outstanding changes in Ops Manager or any other tile. All tiles should be green. Click **Review Pending Changes**, then **Apply Changes** if necessary.
2. After applying changes, click **Recent Install Logs** to confirm that the changes completed cleanly:

```
Cleanup complete
{"type": "step_finished", "id": "clean_up_bosh.cleaning_up"}
Exited with 0.
```

Export Your Installation

To export your installation, do the following:

1. In your Ops Manager **Installation Dashboard**, click the account dropdown and select **Settings**.



2. On the **Settings** screen, select **Export Installation Settings** from the left menu, then click **Export Installation Settings**.

A screenshot of the PCF Ops Manager Settings screen. The title bar says "PCF Ops Manager". Below it, there is a breadcrumb navigation "Installation Dashboard" and a lock icon followed by the text "Settings". On the left side, there is a vertical navigation menu with the following items: "Decryption Passphrase", "Authentication Method", "External API Access", "Proxy Settings", "Export Installation Settings" (which is highlighted with a grey background), and "Advanced". The main content area has a heading "Export Installation". Below the heading, there is a warning message: "⚠️ Upgrading Ops Manager may block the upgrade path for certain products. Please read [Upgrading Ops Manager](#) and verify an upgrade path exists for your installed products." At the bottom of the main content area, there is a blue button labeled "Export Installation Settings".

Decryption Passphrase

Authentication Method

External API Access

Proxy Settings

Export Installation Settings

Advanced

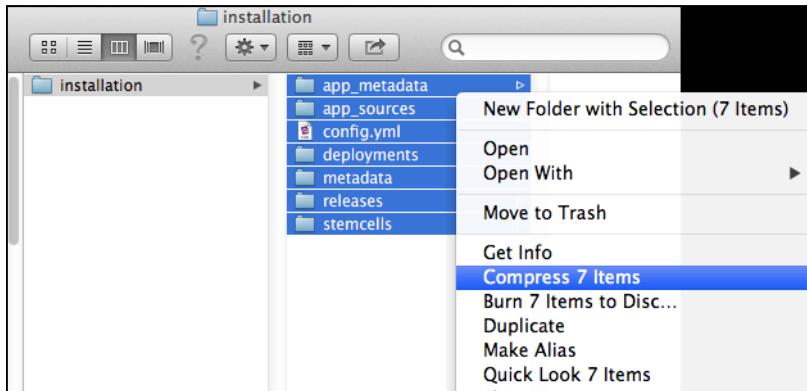
This exports the current PCF installation with all of its assets.

When you export an installation, the export contains the base VM images, necessary packages, and configuration settings, but does not include releases between upgrades if Ops Manager has already uploaded them to BOSH. When backing up PCF, you must take this into account by backing up the BOSH blobstore that contains the uploaded releases. BOSH Backup and Restore (BBR) backs up the BOSH blobstore. For more information, see [Backing Up Pivotal Cloud Foundry with BBR](#).

- The export time depends on the size of the exported file.

- Some browsers do not provide feedback on the status of the export process and might appear to hang.

Note: Some operating systems automatically unzip the exported installation. If this occurs, create a ZIP file of the unzipped export. Do not start compressing at the “installation” folder level. Instead, start compressing at the level containing the `config.yml` file:



warning: If you fail to perform the remedial steps for this issue, this upgrade process may corrupt your existing usage data.

Next Steps

Now that you have completed the Upgrade Preparation Checklist, continue to [Upgrading Pivotal Cloud Foundry](#).

Complete Survey

Please take some time to help us improve this document by completing the [Upgrade Checklist Survey ↗](#).

Configuring PAS for Upgrades

This topic describes several configuration options for Pivotal Application Service (PAS) that can help ensure successful upgrades. In addition to following the [Upgrade Preparation Checklist](#), review the sections in this document to better understand how to prepare for PAS Upgrades.

Limit PCF Component Instance Restarts

The `max_in_flight` variable limits how many instances of a component can restart simultaneously during updates or upgrades. Increasing the value of `max_in_flight` can make updates run faster, but setting it too high risks overloading VMs and causing failure. See [Best Practices](#) for guidance on setting `max_in_flight` values.

Values for `max_in_flight` can be any integer between 1 and 100, or a percentage of the total number of instances. For example, a `max_in_flight` value of `20%` in a deployment with 10 Diego cell instances would make no more than two cell instances restart at once.

Set max_in_flight

The `max_in_flight` variable is a system-wide value with optional component-specific overrides. You can override the default value for individual jobs using an API endpoint.

Use the max_in_flight API Endpoint

Use the `max_in_flight` API endpoint to configure the maximum value for component instances that can start at a given time. This endpoint overrides product defaults. You can specify values as a percentage or an integer.

Use the string “default” as the `max_in_flight` value to force the component to use the deployment’s default value.

 **Note:** The example below lists three JOB_GUIDs. These three GUIDs are examples of the three different types of values you can use to configure `max_in_flight`. The endpoint only requires one GUID.

```
curl "https://EXAMPLE.com/api/v0/staged/products/PRODUCT-TYPE1-GUID/max_in_flight" \
-X PUT \
-H "Authorization: Bearer UAA_ACCESS_TOKEN" \
-H "Content-Type: application/json" \
-d'{
  "max_in_flight": {
    "JOB_1_GUID": 1,
    "JOB_2_GUID": "20%",
    "JOB_3_GUID": "default"
  }
}'
```

Specific Guidance for Diego Cells

To upgrade Cloud Foundry, BOSH must drain all Diego cell VMs that host app instances. BOSH manages this process by upgrading a batch of cells at a time.

The number of cells that undergo upgrade simultaneously (either in a state of shutting down or coming back online) is controlled by the `max_in_flight` value of the Diego cell job. For example, if `max_in_flight` is set to `10%` and your deployment has 20 Diego cell job instances, then the maximum number of cells that BOSH can upgrade at a single time is `2`.

When BOSH triggers an upgrade, each Diego cell undergoing upgrade enters “evacuation” mode. Evacuation mode means that the cell stops accepting new work and signals the rest of the Diego system to schedule replacements for its app instances. This scheduling is managed by the [Diego auctioneer process](#).

The evacuating cells continue to interact with the Diego system as replacements come online. The cell undergoing upgrade waits until either its app instance replacements run successfully before shutting down the original local instances, or for the evacuation process to time out. This “evacuation timeout” defaults to 10 minutes.

If cell evacuation exceeds this timeout, then the cell stops its app instances and shuts down. The Diego system continues to re-emit start requests for the

app replacements.

Prevent Overload

A potential issue arises if too many app instance replacements are slow to start or do not start successfully at all.

If too many app instances are starting concurrently, then the load of these starts on the rest of the system can cause other applications that are already running to crash and be rescheduled. These events can result in a cascading failure.

To prevent this issue, PCF provides two throttle configurations: the maximum number of in-flight diego cell instances and the maximum number of starting containers.

The values of the above throttle configurations depend on the version of PCF that you have deployed and whether you have overridden the default values.

Refer to the following table for existing defaults and, if necessary, determine the override values in your deployment.

PCF Version	Starting Container Count Maximum	Starting Container Count Overridable?	Maximum In Flight Diego Cell Instances	Maximum In Flight Diego Cell Instances Overridable?
PCF 1.7.43 and earlier	No limit set	No	1 instance	No
PCF 1.7.44 to 1.7.49	200	No	1 instance	No
PCF 1.7.50 +	200	No	1 instance	No
PCF 1.8.0 to 1.8.29	No limit set	No	10% of total instances	No
PCF 1.8.30 +	200	Yes	10% of total instances	No
PCF 1.9.0 to 1.9.7	No limit set	No	4% of total instances	Yes
PCF 1.9.8 +	200	Yes	4% of total instances	Yes
PCF 1.10.0 and later	200	Yes	4% of total instances	Yes
PCF 1.12.0 and later	200	Yes	4% of total instances	Yes

Best Practices

Set the `max_in_flight` variable high enough that the remaining component instances are not overloaded by typical use. If component instances are overloaded during updates, upgrades, or typical use, users may experience downtime.

Some more precise guidelines include:

- For jobs with high resource usage, set `max_in_flight` low. For example, for Diego cells, `max_in_flight` allows non-migrating cells to pick up the work of cells stopping and restarting during migration. If resource usage is already close to 100%, scale up your jobs before making any updates.
- For quorum-based components (these are components with odd-numbered settings in the manifest), such as etcd, consul, and Diego BBS, set `max_in_flight` to 1. This preserves quorum and prevents a [split-brain scenario](#) from occurring as jobs restart.
- For other components, set `max_in_flight` to the number of instances that you can afford to have down at any one time. The best values for your deployment vary based on your capacity planning. In a highly redundant deployment, you can make the number high so that updates run faster. If your components are at high utilization, however, you should keep the number low to prevent downtime.
- Never set `max_in_flight` to a value greater than or equal to the number of instances you have running for a component.

Set a Maximum Number of Starting Containers

This section describes how to use the auctioneer job to configure the maximum number of app instances starting at a given time. This prevents Diego from scheduling too much new work for your platform to handle concurrently. A lower default can prevent server overload during cold start, which may be important if your infrastructure is not sized for a large number of concurrent cold starts.

The auctioneer only schedules a fixed number of app instances to start concurrently. This limit applies to both single and multiple Diego Cells. For

example, if you set the limit to five starting instances, it does not matter if you have one Diego Cell with ten instances or five Diego Cells with two instances each. The auctioneer will not allow more than five instances to start at the same time.

If you are using a cloud-based IaaS, rather than a smaller on-premise solution, Pivotal recommends setting a larger default. By default, the maximum number of started instances is 200.

You can configure the maximum number of started instances in the Settings tab of the Pivotal Application Service (PAS) tile.

1. Log in to Operations Manager.
2. Click the PAS tile.
3. Click **Application instances** in the sidebar.
4. In the Max Inflight Container Starts field, type the maximum number of started instances.

Application Containers

Application Developer Controls

Application Security Groups

Authentication and Enterprise SSO

Databases

Internal MySQL

File Storage

System Logging

Custom Branding

Apps Manager

Email Notifications

Allow SSH access to app containers

Enable SSH when an app is created

Private Docker Insecure Registry Whitelist
10.10.10.10:8888,example.com:8888

Docker Images Disk-Cleanup Scheduling on Cell VMs*

Never clean up Cell disk-space

Routinely clean up Cell disk-space

Clean up disk-space once threshold is reached

Threshold of Disk-Used (MB) (min: 1)*
10240

Max Inflight Container Starts *

200

Save

5. Click **Save**.

Configure File Storage

This section describes critical factors to consider when evaluating the type of file storage to use in your Pivotal Cloud Foundry (PCF) deployment. The [Pivotal Application Service \(PAS\) blobstore](#) relies on the file storage system to read and write resources, app packages, and droplets.

During an upgrade of PCF, file storage with insufficient IOPS numbers can negatively impact the performance and stability of your PCF deployment.

If disk processing time takes longer than the evacuation timeout for Diego cells, then Diego cells and app instances may take too long to start up, resulting in a cascading failure.

However, the minimum required IOPS depends upon a number of deployment-specific factors and configuration choices. Use this section as a guide when deciding on the file storage configuration for your deployment.

To see an example of system performance and IOPS load during an upgrade, refer to [Upgrade Load Example: Pivotal Web Services](#).

Select Internal or External File Storage

When you deploy PCF, you can select internal file storage or external file storage, either network-accessible or IaaS-provided, as an option in the PAS

tile.

Selecting internal storage causes PCF to deploy a dedicated virtual machine (VM) that uses either NFS or WebDAV for file storage. Selecting external storage allows you to configure file storage provided in network-accessible location or by an IaaS, such as Amazon S3, Google Cloud Storage, or Azure Storage.

Whenever possible, Pivotal recommends using external file storage.

Calculate Potential Disk Load Requirements

As a best-effort calculation, estimate the total number of bits needed to move during a system upgrade to determine how IOPS-performant your file storage needs to be.

- **Number of Diego Cells**
 - As a first calculation, determine the number of Diego cells that your deployment currently uses. To view the number of Diego cell instances currently running in your deployment, see the **Resource Config** section of your PAS tile. If you expect to scale up the number of instances, use the anticipated scaled number.

 **Note:** If your deployment uses more than 20 Diego cells, you should avoid using internal file storage. Instead, you should always select external or IaaS-provided file storage.
- **Maximum In-Flight Load and Container Starts for Diego Cells**
 - Operators can limit the number of containers and Diego cell instances that Diego starts concurrently. If operators impose no limits, your file storage may experience exceptionally heavy load during an upgrade.

Upgrade Load Example: Pivotal Web Services

This topic provides sample performance measurements of a Cloud Foundry installation undergoing the workload associated with an upgrade.

To obtain these measurements, Pivotal repaved its production Pivotal Web Services (PWS) deployment. The repave process simulates system load that would be incurred when performing a rolling upgrade of Diego cells.

Use the measurements and configuration values published in this document as guidance when ensuring you have adequate file storage hardware prior to a platform upgrade.

For more information about the impact of upgrade on file storage performance, see [Configure File Storage](#).

Platform Configuration

The following table details the starting parameters and configuration of PWS.

Configuration	Value	How to Locate
IaaS	Amazon Web Services	Refer to your BOSH Director configuration or BOSH deployment manifest.
File Storage	AWS EBS (External with some elastic capacity)	Refer to your Pivotal Application Service (PAS) configuration or BOSH deployment manifest.
Version of CF	v252	Refer to your BOSH Director and PAS configuration or BOSH deployment manifest.
Number of Diego Cells	218	To view the number of Diego cell instances currently running in your deployment, see the <code>Resource Config</code> section of your PAS tile or consult your Diego deployment manifest.
Maximum Number of Started Containers	250	See Configuring PAS for Upgrades .
max_in_flight Configuration for Diego Cells	6	To retrieve the existing <code>max_in_flight</code> value for the Diego Cell job in BOSH Director, use the Ops Manager API. See the Ops Manager API documentation. If you are running open source CF, consult your BOSH deployment manifest.
Number of Availability Zones (AZ)	2	Consult your PAS or BOSH deployment AZ configuration.
Number of App Instances	16231	<code>datadog.nozzle.bbs.LRPsRunning</code>
Number of Application Security Groups (ASGs)	43	As admin user, run the <code>cf security-groups</code> command. For more information, see Application Security Groups .

System Performance Measurements During Cell Repave

This table presents performance measurements taken during the Diego cell repave.

 **Note:** These measurements indicate the peak cumulative values of the entire system (250 Diego cells, ~15,000 application instances, and 2 AZs.)

Use these measurements as a baseline for expected system load during Diego cell upgrade.

Measurement	Value	Metric Used
Cell CPU Consumption	36%	<code>bosh.healthmonitor.system.cpu.user</code>
Cell Memory Consumption	~50%	<code>bosh.healthmonitor.system.mem.percentage</code>
Cell I/O Consumption (Read) During Normal Operations	43 Read I/O Operations per second	<code>aws.ebs.volume_read_ops</code>
Cell I/O Consumption (Read) During Upgrade	1 943 Read I/O Operations per second	<code>aws.ebs.volume_read_ops</code>

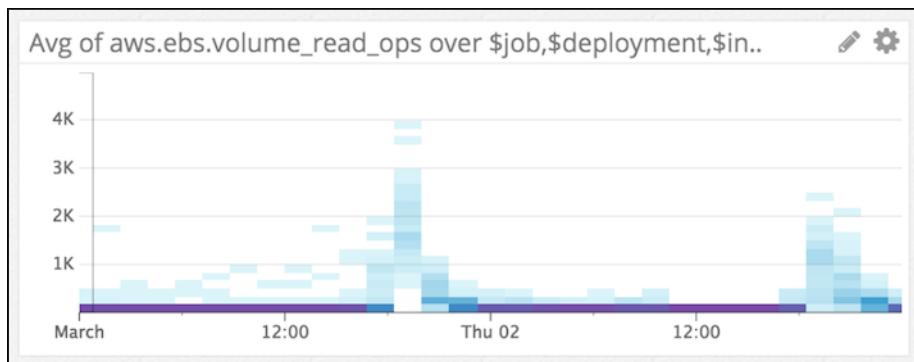
Cell IO Consumption (Read) During Operations	~2,166 Read I/O Operations per second	aws.ebs.volume_read_ops
Cell IO Consumption (Write) During Normal Operations	2,166 Write I/O Operations per second	aws.ebs.volume_write_ops
Cell IO Consumption (Write) During Upgrade	21,000 Write I/O Operations per second	aws.ebs.volume_write_ops
Cell Network Consumption (Network Out) During Normal Operations	~1.25 GB per minute	aws.ec2.network_out
Cell Network Consumption (Network Out) During Upgrade	~1.25 GB per minute (no significant change)	aws.ec2.network_out
Cell Network Consumption (Network In) During Normal Operations	2.11 GB per minute	aws.ec2.network_in
Cell Network Consumption (Network In) During Upgrade	16.75GB per minute	aws.ec2.network_in

Sample Performance Graphs

These DataDog graphs represent a timeline visualization of read and write operations during the repave event.

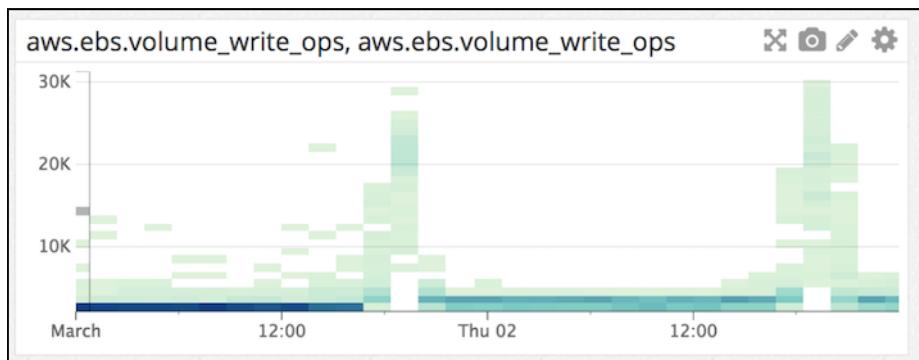
Read I/O Operations

The read I/O operations sample was taken over 115 VMs and represent the number of read operations over 300 seconds for a single Diego cell.



Write I/O Operations

The write I/O operations sample was taken over 115 VMs and represent the number of read I/O operations over 300 seconds for a single Diego cell.



Summary

During the repave process, 250 Diego cells were updated. The repave process took 6 hours overall or about 3 hours for each Availability Zone.

Upgrading PAS and Other Pivotal Cloud Foundry Products

This topic describes how to upgrade to a point release of Pivotal Application Service (PAS) and other product tiles without upgrading Ops Manager. For example, use this topic to upgrade from PAS v1.9.0 to v1.9.1. You might need to perform this upgrade if a security update for PAS is released, or if new features are introduced in a point release of a product tile.

For PAS component and version information, see the [PAS release notes](#).

 **Note:** If you cannot download products from Pivotal Network due to restricted network connectivity, see [Installing PCF in Airgapped Environments](#).

Before You Upgrade to Point Releases

- You must have completed the [Upgrading Pivotal Cloud Foundry](#) procedure.
- Refer to the [Product Compatibility Matrix](#) before upgrading PAS.
- **Important:** Read the Known Issues sections of the products you plan on installing before starting. See [Pivotal Cloud Foundry Release Notes](#) for all available product release notes.

Upgrading PAS

 **Note:** If you are using the [Pivotal Network API](#), the latest product versions will automatically appear in your [Installation Dashboard](#).

To upgrade PAS without upgrading Ops Manager, do the following:

1. Download the product file from [Pivotal Network](#).
2. Import the product file to your Ops Manager [Installation Dashboard](#).
3. Click the plus icon next to the uploaded product description to add this product to your staging area.
4. Click the newly added tile to review any configurable options.
5. Click **Review Pending Changes**, then **Apply Changes** to install the service.

Upgrading PCF Products

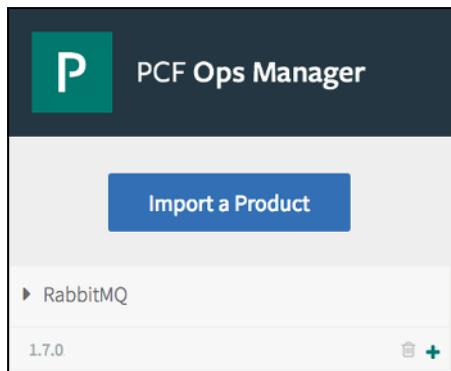
 **Note:** If you are using the [Pivotal Network API](#), the latest product versions will automatically appear in your [Installation Dashboard](#).

This section describes how to upgrade individual products like [Single Sign-On for PCF](#), [MySQL for PCF](#), [RabbitMQ® for PCF](#), and [Metrics for PCF](#) in your Pivotal Cloud Foundry (PCF) deployment. Ensure you review the individual product upgrade procedure for each tile.

1. Browse to [Pivotal Network](#) and sign in.
2. Download the latest PCF release for the product or products you want to upgrade. Every product is tied to exactly one stemcell. Download the stemcell that matches your product and version.
3. Confirm that you have adequate disk space for your upgrades. You need at least 20 GB of free disk space to upgrade PCF Ops Manager and Pivotal Application Service. If you plan to upgrade other products, the amount of disk space required depends on how many tiles you plan to deploy to your upgraded PCF deployment.

To check current persistent disk usage, select the **BOSH Director** tile from the [Installation Dashboard](#). Select **Status** and review the value of the **PERS. DISK** column. If persistent disk usage is higher than 50%, select **Settings > Resource Config**, and increase your persistent disk space to handle the size of the resources. If you do not know how much disk space to allocate, set the value to at least **100 GB**.

4. Browse to the Pivotal Cloud Foundry Ops Manager web interface and click **Import a Product**.



5. Select the `.pivotal` file that you downloaded from Pivotal Network or received from your software distributor, then click **Open**. If the product is successfully added, it appears in your product list. If the product that you selected is not the latest version, the most up-to-date version will appear on your product list.
6. Click the plus icon next to the product description to add the product tile to the **Installation Dashboard**.
7. Repeat the import, upload, and upgrade steps for each product you downloaded.
8. If you are upgrading a product that uses a self-signed certificate from v1.1 to v1.2, you must configure the product to trust the self-signed certificate. To configure a product to trust the self-signed certificate, do the following:
 - a. Click the product tile.
 - b. In the left-hand column, select the setting page containing the SSL certificate configuration. For example, for PAS, select the **HAProxy** page.
 - c. Check the **Trust Self-Signed Certificates** box.
 - d. Click **Save**.
9. Click **Review Pending Changes**, then **Apply Changes**.

Upgrading Replicated Tiles

To upgrade a replicated tile, you must do the following:

- Generate a replica of the newer version of the tile using the replicator
- Give the new replica the same name as the existing replica.

See the following example workflow.

Example Workflow

This example assumes the following:

- You used the replicator to generate a replica of v1 of the Isolation Segments tile.
- You used the `--name seg-one`.
- You installed the tile in Ops Manager.

Here is the sample replicator command you used for the initial installation:

```
./replicator-darwin
--name seg-one\
--path /download/p-isolation-segment-v1.pivotal\
--output /output/p-isolation-segment-v1-seg-one.pivotal
```

To upgrade to v2, follow these steps:

1. Download and unzip the new Isolation Segment Tile Replicator from [Pivotal Network](#). You must download the version of Tile Replicator that corresponds with the version of the Isolation Segment Tile you want to replicate.
2. Run the replicator command below to create the replica:

```
./replicator-darwin\  
--name NAME-OF-EXISTING-REPLICA \  
--path PATH-TO-NEW-TILE \  
--output /output/p-isolation-segment-v2-seg-one.pivot
```

Where:

- **NAME-OF-EXISTING-REPLICA** must be the same as the name used for the existing replica. This is **seg-one** in this example.
- **PATH-TO-NEW-TILE** is the path to the new Isolation Segment Tile.

3. After you have the replica tile **p-isolation-segment-v2-seg-one.pivot**, upload it to Ops Manager. This upgrades the v1 replica tile in place.

cf push Availability During Pivotal Application Service Upgrades

This topic describes what you can expect regarding the availability of `cf push` during upgrades of Pivotal Application Service (PAS).

Overview

Availability of `cf push` during PAS upgrades varies from release to release. There are various considerations, such as Cloud Controller database (CCDB) migrations or the number of VMs in use, that can impact `cf push` availability. However, `cf push` is mostly available for the duration of an upgrade.

Impact on Single VMs vs. High-Availability Infrastructure

Having a single VM in use, such as a WebDAV, a HAProxy, a Gorouter, a UAA, or a Cloud Controller, impacts whether `cf push` is unavailable during an upgrade.

If you have scaled out your application to achieve high availability and are not using WebDAV, `cf push` should be available for the entire duration of the upgrade. However, upgrades to certain versions of PAS sometimes require a CCDB schema or data migration, which may cause `cf push` to be unavailable while Cloud Controllers are rolling during the upgrade.

Availability During a BBR Backup

The Cloud Controller API is taken down during the `pre-backup-lock` stage of a BBR backup and put back up again during the `post-backup-unlock` stage. As a result, `cf push` becomes unavailable during that time. However, the backup that takes place between those stages is very short, no longer than a few minutes. The bulk of the BBR backup operation happens after the `post-backup-unlock` stage, so the Cloud Controller API and `cf push` are available for most of the duration of a BBR backup.

The [Uptimer](#) tool can help you measure `cf push` availability during an upgrade.

Backing Up and Restoring Pivotal Cloud Foundry

Consider the following when backing up data in your Pivotal Cloud Foundry (PCF) deployment:

- If your deployment uses external databases, check that it is a [supported database configuration](#).
- If your PCF deployment uses internal databases, follow the backup and restore instructions for the MySQL server included in the [BBR](#) documentation.

General Data Protection Regulation

The General Data Protection Regulation (GDPR) came into effect on May 25, 2018 and impacts any company processing the data of EU citizens or residents, even if the company is not EU-based. The GDPR sets forth how companies should handle privacy issues, securely store data, and respond to security breaches.

Backup artifacts may contain personal data covered by GDPR. For example, a backup of a PAS could contain a user email. For further information regarding personal data that may be stored in PCF, see [here](#).

Backup and Restore with BBR

BOSH Backup and Restore (BBR) is a command-line tool for backing up and restoring BOSH deployments.

To perform a backup of your PCF deployment with BBR, see [Backing Up Pivotal Cloud Foundry with BBR](#).

To restore your PCF deployment with BBR, see [Restoring Pivotal Cloud Foundry from Backup with BBR](#).

To troubleshoot problems with BBR, see [Troubleshooting BBR](#).

BBR Restore Scenarios

There are many different restore cases that can arise with BOSH Directors and PCF deployments. We have provided some guidance on how to use BBR in common restore scenarios below:

- [Restoring Lost BOSH Director Persistent Disk](#)
- [Restoring Lost PAS Data](#)
- [Recovering from an Ops Manager and PAS Upgrade Failure](#)

These guides do not cover all possible restore scenarios. If you encounter another restore case, then some information in these guides may be applicable. Contact [Pivotal Support](#) for guidance.

Incident Response Guides

This guide describes how Pivotal Cloud Foundry (PCF) operators can use BOSH Backup and Restore (BBR) to respond to upgrade failures and data loss within their deployment.

Overview

This guide includes the following topics:

- [Restoring Lost BOSH Director Persistent Disk](#)
- [Restoring Lost PAS Data](#)
- [Recovering from an Ops Manager and PAS Upgrade Failure](#)

Restoring Lost BOSH Director Persistent Disk

This guide explains how to use BOSH Backup and Restore (BBR) to restore a BOSH Director that has lost its persistent disk.

Prerequisite

Restoring a BOSH Director persistent disk with BBR is only possible if you have previously taken a backup using the following command:

```
bbr director backup
```

This command produces a backup artifact in the format `DIRECTOR_IP_TIMESTAMP`.

A **warning:** BOSH Director backups require identical references to all BOSH-deployed VMs. If BOSH recreates your deployment VMs for any reason, such as changes to stemcells, networks, or availability zones (AZs), your BBR backup artifact is no longer compatible for performing a restore.

Prepare Your Environment

Your BOSH Director must be in a healthy state before you can perform a BOSH Director restore. The BOSH Director is in a healthy state when all jobs are running.

If your BOSH Director has lost its persistent disk, create a new disk to return your BOSH Director to a healthy state.

Create a New BOSH Director Persistent Disk

To create a new persistent disk for the BOSH Director, do the following:

1. Create a new persistent disk using your IaaS console.
2. Log in to the Ops Manager VM. See the [Log in to the Ops Manager VM with SSH](#) section for your IaaS in *Advanced Troubleshooting with the BOSH CLI*.
3. Open the `bosh-state.json` file in a text editor. Edit the `disks.cid` value to match the new persistent disk you created in the first step.
4. Update the persistent disk size in Ops Manager and then click **Apply Changes**. This creates a new BOSH Director VM and attaches the new persistent disk to it.

Perform the Restore

A **warning:** BBR restore is a destructive process which removes any current data in your deployment. Performing a restore overwrites all new data since you created your most recent backup artifact.

Perform a BOSH Director restore by following the BBR instructions in [Step 8: Restore the BOSH Director](#) of *Restoring PCF From Backup with BBR*.

Consolidate Your Environment

After you successfully restore your BOSH Director, the BOSH Director uses the VM references that were stored in the backup artifact. If you clicked **Apply Changes** in Ops Manager between when the backup and the restore took place, check if any BOSH-deployed VMs were deleted or added.

If changes have occurred, ensure that the BOSH Director database is consistent with the current state of your IaaS after the restore. If VMs were deleted, the BOSH Director still has references to these deleted VMs after the restore. If VMs were added, the BOSH Director has no knowledge of the new VMs.

To ensure that the BOSH Director database matches your IaaS console, do the following:

1. Run the following command to find and delete outdated VM references for an environment:

```
bosh --deployment DEPLOYMENT-NAME cloud-check
```

Where `[DEPLOYMENT-NAME]` is the name of your BOSH deployment.

2. Log in to your IaaS console and delete any orphaned VMs that your BOSH Director does not reference.

Your BOSH Director should now be in a healthy state with the new persistent disk attached. Your BOSH deployment VMs should be aligned with the current state of your IaaS.

Restoring Lost PAS Data

This guide explains how to use BOSH Backup and Restore (BBR) to restore a piece of Pivotal Application Service (PAS) data, such as a user or space, which has been lost.

Prerequisite

Restoring PAS data with BBR is only possible if you have previously taken a backup using the following command:

```
bbr director backup
```

This command produces a backup artifact in the format `[DIRECTOR_IP_TIMESTAMP]`.

Validate Your Backup Artifact

Before restoring, verify that the backup artifact you have is suitable for your restore need. You can `untar` and explore the PAS backup artifact to verify that it contains the data that you need to restore.

Use any the checks in this section depending on the data you need to restore.

Verify an Org Is Present

To verify that your backup artifact contains the org you need to restore, do the following:

```
tar -xvf backup_restore-0-bbr-cloudcontrollerdb.tar  
cat cloudcontroller-artifact-file | grep "ORG-TO-RESTORE"
```

Where `ORG-TO-RESTORE` is the org you need to restore.

If you see the following line in the output of the above command, then you can restore this org:

```
INSERT INTO `organizations` VALUES (...,'ORG-TO-RESTORE', ...);
```

Verify a Space Is Present

To verify that your backup artifact contains the space you need to restore, do the following:

```
tar -xvf backup_restore-0-bbr-cloudcontrollerdb.tar  
cat cloudcontroller-artifact-file | grep "SPACE-TO-RESTORE"
```

Where `SPACE-TO-RESTORE` is the space you need to restore.

If you see the following line in the output of the above command, then you can restore this space:

```
INSERT INTO `spaces` VALUES (...,'SPACE-TO-RESTORE', ...);
```

Verify a User Is Present

To verify that your backup artifact contains the user you need to restore, do the following:

```
tar -xvf backup_restore-0-bbr-uaadb.tar  
cat uaadb-artifact-file | grep "USER-TO-RESTORE"
```

Where `USER-TO-RESTORE` is the user you need to restore.

If you see the following line in the output of the above command, then you can restore this user:

```
INSERT INTO `users` VALUES (...,'USER-TO-RESTORE', ...);
```

Verify a Route Is Present

To verify that your backup artifact contains the route you need to restore, do the following:

```
tar -xvf backup_restore-0-bbr-cloudcontrollerdb.tar  
cat cloudcontroller-artifact-file | grep "ROUTE-TO-RESTORE"
```

Where `ROUTE-TO-RESTORE` is the route you need to restore.

If you see the following line in the output of the above command, then you can restore this user:

```
INSERT INTO `routes` VALUES (...,'ROUTE-TO-RESTORE', ...);
```

Verify a Domain Is Present

To verify that your backup artifact contains the domain you need to restore, do the following:

```
tar -xvf backup_restore-0-bbr-cloudcontrollerdb.tar  
cat cloudcontroller-artifact-file | grep "DOMAIN-TO-RESTORE"
```

Where `DOMAIN-TO-RESTORE` is the domain you need to restore.

If you see the following line in the output of the above command, then you can restore this domain:

```
INSERT INTO `domains` VALUES (...,'DOMAIN-TO-RESTORE', ...);
```

Verify an App Is Present

To verify that your backup artifact contains the app you need to restore, do the following:

```
tar -xvf backup_restore-0-bbr-cloudcontrollerdb.tar  
cat cloudcontroller-artifact-file | grep "APP-TO-RESTORE"
```

Where `APP-TO-RESTORE` is the app you need to restore.

To verify that your blobstore backup artifact contains the droplet for the app you need to restore, do the following:

```
tar -xvf blobstore-backup.tar  
tree blobstore-backup | grep "APP-DROPLET"
```

Where `APP-DROPLET` is the droplet for the app you need to restore.

If you see the following line in the output of the above command, then you can restore this app:

```
INSERT INTO `apps` VALUES (...,'APP-TO-RESTORE', 'APP-DROPLET', ...);
```

 **Note:** If you use an external blobstore, go to your IaaS console and check the backup bucket to verify that the droplet is present.

Perform the Restore

⚠ warning: BBR restore is a destructive process which removes any current data in your deployment. Performing a restore overwrites all new data since you created your most recent backup artifact.

Pivotal recommends that you verify the compatibility of your backup artifact by attempting the restore on an environment that is identical to the target environment. If this is not possible, refer to the [Compatibility of Restore](#) section of *Restoring PCF from Backup with BBR* to confirm that your backup is compatible with your restore environment.

Run a PAS restore by following the BBR instructions in [Step 13: Restore PAS](#) of *Restoring PCF from Backup with BBR*.

After a restore, PAS should roll back to its state at the time when the backup was taken, and contain the data that you had previously lost.

Consolidate Your Environment

If you have service instances that were not present when you performed the backup, then the BOSH service deployments become orphaned. PAS has no reference to these service deployments, but they continue to exist. You can locate orphaned service deployments using the BOSH command-line interface (CLI).

You can locate orphaned on-demand service instances by running the `orphan-deployments` errand. For more information, see [Orphan Deployments](#) in *Broker and Service Management*.

Recovering from an Ops Manager and PAS Upgrade Failure

Page last updated:

This topic explains how to use BOSH Backup and Restore (BBR) to restore Ops Manager and Pivotal Application Service (PAS) back to a healthy state after a failed upgrade between PAS minor versions.

 **Note:** BBR can only restore across the same PAS minor versions. See `metadata/cf-VERSION.yml` in the backup artifact to verify which version of PAS was used to create the backup artifact.

Evaluate Your Deployment

You can restore Ops Manager and PAS to a healthy state after a failed upgrade between PAS minor versions if your backups, BOSH Director, and PAS deployment meet certain conditions. Follow the procedures in this section to confirm that your deployment can be restored by BBR.

Review Your Backup

Before you perform a restore, you must do the following:

1. Run the following `bbr` commands from your Ops Manager VM:

- o Back up your deployments by running the following command:

```
bbr deployment --deployment DEPLOYMENT_NAME backup
```

- o Back up your BOSH Director by running the following command:

```
bbr director backup
```

The above `bbr` commands produce the following backup artifacts:

- `DEPLOYMENT_NAME_TIMESTAMP/`
- `DIRECTOR_IP_TIMESTAMP/`
- `installation.zip`

2. Export your Ops Manager settings. For more information, see the [Settings Page](#) section of *Using the Ops Manager Interface*.

Validate Your Environment for Restore

Before you can restore after a failed upgrade between PAS minor versions, you must confirm that your BOSH Director and PAS deployment are healthy.

Confirm That Your BOSH Director Is Healthy

To check that the BOSH Director is in a healthy state for a `bbr restore`, SSH to the Ops Manager VM and ensure that all jobs are running.

To make sure that all jobs are running, do the following:

1. SSH into your BOSH Director. For more information, see [Log in to the BOSH Director VM with SSH](#) in *Advanced Troubleshooting with the BOSH CLI*.
2. View all deployment job instances by running the following command:

```
monit summary
```

Monit should report that all jobs are `running`.

3. Ensure that the BOSH API is responsive by running the following command:

```
bosh configs
```

If this command runs successfully, then the BOSH Director API is in a healthy and responsive state and can be restored by BBR.

If your BOSH Director is unhealthy, see [Restoring an Unhealthy BOSH Director](#) below.

Confirm That Your PAS Deployment Is Healthy

To confirm that PAS is in a healthy state, run the following command:

```
bosh -d PAS-DEPLOYMENT-NAME instances
```

Where `PAS-DEPLOYMENT-NAME` is the name of your PAS BOSH deployment.

If your PAS deployment is healthy, then the output states that all of your instances are in the `running` state. If all instances are running, PAS can be restored by BBR.

If your PAS deployment is unhealthy, see [Restoring an Unhealthy PAS Deployment](#).

Restore Your Deployment

After you confirm that your BOSH Director is healthy with its data intact, and all instances in your PAS deployment are responsive, you can restore PAS with BBR.

Restoring PAS

If your BOSH Director is healthy, you can restore across the same PAS minor version by following the procedures in [Step 13: Restore PAS in Restoring PCF from Backup with BBR](#).

Recover to a Healthy State

If either the BOSH Director or PAS are not in a healthy state, they must be recovered before you can run `bbr restore`. If you need guidance on the following subsections, contact [Pivotal Support](#).

Restoring an Unhealthy BOSH Director

If the BOSH Director is not in a healthy state, Pivotal recommends following our full restore workflow to regain your Ops Manager, BOSH Director, and BOSH deployments.

See [Restore Workflow](#) in *Restoring PCF from Backup with BBR*.

Restoring an Unhealthy PAS Deployment

If your BOSH Director is healthy, you can restore your previous BOSH Director state by following [Step 8: Restore the BOSH Director](#) in *Restoring PCF from Backup with BBR*.

After your BOSH Director is restored, you can then restore your deployments by following [Step 13: Restore PAS](#) in *Restoring PCF from Backup with BBR*.

Disaster Recovery in Pivotal Cloud Foundry

This document provides an overview of the options and considerations for disaster recovery in Pivotal Cloud Foundry (PCF).

Operators have a range of approaches for ensuring they can recover Pivotal Cloud Foundry, apps, and data in case of a disaster. The approaches fall into the following two categories:

- Using data from a backup to restore the data in the PCF Deployment. See [Back up and Restore Using BOSH Backup and Restore \(BBR\)](#) for more information.
- Recreating the data in PCF by automating the creation of state in PCF. See [Disaster Recovery by Recreating the Deployment](#) for more information.

Back up and Restore using BOSH Backup and Restore (BBR)

What is BBR?

BOSH Backup and Restore (BBR) is a CLI for orchestrating backing up and restoring BOSH deployments and BOSH Directors. BBR triggers the backup or restore process on the deployment or Director, and transfers the backup artifact to and from the deployment or Director.

Use BOSH Backup and Restore to reliably create backups of core PCF components and their data. These core components include CredHub, UAA, BOSH Director, and PAS.

Each component includes its own backup scripts. This decentralized structure helps keep scripts synchronized with the components. At the same time, *locking* features ensure data integrity and consistent, distributed backups across your deployment.

For more information about the BBR framework, see [BOSH Backup and Restore](#) in the open source Cloud Foundry documentation.

Backing up PCF

Backing up PCF requires backing up the following components:

- Ops Manager settings
- BOSH Director, including CredHub and UAA
- Pivotal Application Service (PAS)
- Data services

For more information, see [Backing up Pivotal Cloud Foundry with BBR](#). With these backup artifacts, operators can recreate PCF exactly as it was when the backup was taken.

Restoring PCF

The restore process involves creating a new PCF deployment starting with the Ops Manager VM. For more information, see [Restoring Pivotal Cloud Foundry from Backup with BBR](#).

The time required to restore the data is proportionate to the size of the data because the restore process includes copying data. For example, restoring a 1 TB blobstore takes one thousand times as long as restoring a 1 GB blobstore.

Benefits

Unlike other backup solutions, using BBR to back up PCF enables the following:

- **Completeness:** BBR supports backing up BOSH, including releases, CredHub, UAA, and service instances created with an on-demand service broker. With PCF v1.12, Ops Manager export no longer includes releases.
- **Consistency:** BBR provides referential integrity between the database and the blobstore because a lock is held while both the database and blobstore are backed up.
- **Correctness:** Using the BBR restore flow addresses C2C and routing issues that can occur during restore.

API Downtime During Backups

Apps are not affected during backups, but certain APIs are unavailable. The downtime occurs only while the backup is being taken, not while the backup is being copied to the jumpbox.

In a consistent backup, the blobs in the blobstore match the blobs in the Cloud Controller Database. To take a consistent backup, changes to the data are prevented during the backup. This means that the CF API, Routing API, Usage Service, Autoscaler, Notification Service, Network Policy Server, and CredHub are unavailable while the backup is being taken. UAA is in read-only mode during the backup.

Backup Timings

The first three phases of the backup are lock, backup, and unlock. During this time, the API is unavailable. The drain and checksum phase starts once the backup scripts finish. BBR downloads the backup artifacts from the instances to the BBR VM, and performs a checksum to ensure the artifacts are not corrupt. The size of the blobstore significantly influences backup time.

The table below gives an indication of the downtime that you can expect. Actual downtime varies based on hardware and PCF configuration. These example timings were recorded with Pivotal Application Service (PAS) deployed on Google Cloud Platform (GCP) with all components scaled to one and only one app pushed.

Backup Timings				
API State	Backup phase	Duration for External Versioned S3-Compatible Blobstore	Duration for External Unversioned S3-Compatible Blobstore	Duration for Internal Blobstore
API unavailable	lock	15 seconds	15 seconds	15 seconds
	backup	<30 seconds	Proportional to blobstore size	10 seconds
	unlock	3 minutes	3 minutes	3 minutes
API available	drain and checksum	<10 seconds	<10 seconds	Proportional to blobstore size

Blobstore backup and restore

Blobstores can be very large. To minimize downtime, BBR only takes blob metadata during the backup. For example, in the case of internal blobstores (Webdav/NFS), BBR takes a list of hard links that point to the blobs. Once the API becomes available, BBR makes copies of the blobs.

Unsupported Products

- **Data services.** The Pivotal data services listed below do not support BBR. Operators of these services should use the automatic backups feature of each tile, available within Ops Manager.
 - MySQL for PCF
 - Pivotal Cloud Cache for PCF
 - RabbitMQ for PCF
 - Redis for PCF
- **External blobstores and databases.** BBR support for backing up and restoring external databases and blobstores varies across PCF versions. For more information, see [Supported Components](#) and [External Storage Support Across PCF Versions in Backing up Pivotal Cloud Foundry with BBR](#).

Best Practices

Frequency of Backups

Pivotal recommends that you take backups in proportion to the rate of change of the data in PCF to minimize the number of changes lost if a restore is required. We suggest starting with backing up every 24 hours. If app developers make frequent changes, you should increase the frequency of backups.

Retention of Backup Artifacts

Operators should retain backup artifacts based on the timeframe they need to be able to restore to. For example, if backups are taken every 24 hours and

PCF must be able to be restored to three days prior, three sets of backup artifacts should be retained.

Artifacts should be stored in two data centers other than the PCF data center. When deciding the restore timeframe, you should take other factors such as compliance and audit-ability into account.

Security

Pivotal strongly recommends that you encrypt artifacts and stored them securely.

Disaster Recovery by Recreating the Deployment

An alternative strategy for recovering PCF after a disaster is to have automation in place so that all the data can be recreated. This requires that every modification to PCF settings and state be automated, typically through use of a pipeline.

Recovery steps include creating a new PCF, recreating orgs, spaces, users, services, service bindings and other state, and re-pushing apps.

For more information about this approach, see the following Cloud Foundry Summit presentation: [Multi-DC Cloud Foundry: What, Why and How? ↗](#).

Disaster Recovery for Different Topologies

Active-Active

To prevent app downtime, some Pivotal customers run active-active, where they run two or more identical PCF deployments in different data centers. If one PCF deployment becomes unavailable, traffic is seamlessly routed to the other deployment. To achieve identical deployments, all operations to PCF are automated so they can be applied to both PCF deployments in parallel.

Because all operations have been automated, the automation approach to disaster recovery is a viable option for active-active. Disaster recovery requires recreating PCF, then running all the automation to recreate state.

This option requires discipline to automate all changes to PCF. Some of the operations that need to be automated are the following:

- App push, re-stage, scale
- Org, space, and user create, read, update, and delete (CRUD)
- Service instance CRUD
- Service bindings CRUD
- Routes CRUD
- Security groups CRUD
- Quota CRUD

Human-initiated changes always make their way into the system. These changes can include quotas being raised, new settings being enabled, and incident responses. For this reason, Pivotal recommends taking backups even when using an automated disaster recovery strategy.

Using BBR Backup and Restore versus Recreating a Failed PCF Deployment in Active-Active

Disaster Recovery		
	Restore the PCF Data	Recreate the PCF Data
Preconditions	IaaS prepared for PCF install	IaaS prepared for PCF install
Steps	1. Recreate PCF 2. Restore 3. Apply changes to make restored PCF match the other active PCF	1. Recreate PCF 2. Trigger automation to recreate orgs, spaces, etc. 3. Notify app developers to re-push apps, recreate service instances and bindings
RTO (Recovery Time Objective)		

Platform	Time to recreate PCF	Time to recreate PCF
Apps	Time to restore	Time until orgs/spaces/etc have been recreated + apps have been re-pushed
RPO (Recovery Point Objective)		
Platform	Time of the last backup	Current time
Apps	Time of the last backup	Current time

Active-Passive

Instead of having a true active-active deployment across all layers, some Pivotal customers prefer to install a PCF or PAS deployment on a backup site. The backup site resides on-premises, in a co-location facility, or the public cloud. The backup site includes an operational deployment, with only the most critical apps ready to accept traffic should a failure occur in the primary data center. Disaster recovery in this scenario involves the following:

1. Switching traffic to the passive PCF, making it active.
2. Recovering the formerly-active PCF. Operators can choose to do this through automation, if that option is available, or by using BBR and the restore process.

The RTO and RPO for recreating the active PCF are the same as outlined in the table above.

Reducing RTO

Both the restore and recreate data disaster recovery options require standing up a new PCF, which can take hours. If you require shorter RTO, several options involving a pre-created standby hardware and PCF are available:

Active-cold	Public cloud environment ready for PCF installation, no PCF installed. This saves both IaaS costs and PCF instance costs. For on-on-premise installations, this requires hardware on standby, ready to install on, which may not be a realistic option.
Active-warm	PCF installed on standby hardware and kept up to date, VMs scaled down to zero (spin them up each time there is a platform update), no apps installed, no orgs or spaces defined.
Active-inflate platform	Bare minimum PCF install, either with no applications, or a small number of each app in a stopped state. On recovery, push a small number of apps or start current apps, while simultaneously triggering automation to scale the platform to the primary node size, or a smaller version if large percentages of loss are acceptable. This mode allows you to start sending some traffic immediately, while not paying for a full non-primary platform. This method requires data seeded, but it is usually acceptable to complete data sync while platform is scaling up.
Active-inflate apps	Non-primary deployment scaled to the primary node size, or smaller version if large percentages of loss are acceptable, with a small number of Diego cells (VMs). On fail-over, scale Diego cells up to primary node counts. This mode allows you to start sending most traffic immediately, while not paying for all the AIs of a fully fledged node. This method requires data to be there very quickly after failure. It does not require real-time sync, but near-real time.

There is a trade-off between cost and RTO: the less the replacement PCF needs to be deployed and scaled, the faster the restore.

Automating Backups

BBR generates the backup artifacts required for PCF, but does not handle scheduling, artifact management, or encryption. The BBR team has created a [starter Concourse pipeline](#) to automate backups with BBR.

Also, Stark & Wayne's Shield can be used as a front end management tool using the BBR plugin.

Validating Backups

To ensure that backup artifacts are valid, the BBR tool creates checksums of the generated backup artifacts, and ensures that the checksums match the

artifacts on the jumpbox.

However, the only way to be sure that the backup artifact can be used to successfully recreate PCF is to test it in the restore process. This is a cumbersome, dangerous process so should be done with care. For instructions, see [Step 11: \(Optional\) Validate Your Backup](#) of the *Backing Up Pivotal Cloud Foundry with BBR*.

Backing Up Pivotal Cloud Foundry with BBR

This topic describes the procedure for backing up your critical backend Pivotal Cloud Foundry (PCF) components with BOSH Backup and Restore (BBR), a command-line tool for backing up and restoring BOSH deployments. To restore your backup, see the [Restoring Pivotal Cloud Foundry from Backup with BBR](#) topic.

To view the BBR release notes, see the [Cloud Foundry documentation](#).

During the backup, BBR stops the Cloud Controller API and the Cloud Controller workers to create a consistent backup. Only the API functionality, like pushing applications or using the Cloud Foundry Command Line Interface (cf CLI) are affected. The deployed applications do not experience downtime.

warning

- **Security:** Backup artifacts can contain secrets. Secure backup artifacts by using encryption or other means.
- **Compatibility:** A BBR backup can only be restored to environments matching its source environment. To ensure that a restore environment is compatible with a backup see [Compatibility of Restore](#).
- **Failed Restore:** Restore is a destructive operation. BBR is designed to restore PCF after a disaster. If a restore fails, the environment might be left in an unusable state and require re-provisioning. For more information, see [Restoring Pivotal Cloud Foundry from Backup with BBR](#).
- **Service Data:** BBR does not back up any service data.
- **API Outage:** The Cloud Controller API does not send or receive calls during PCF restoration.

Recommendations

Pivotal recommends the following:

- Follow the full procedure documented in this topic when creating a backup. This ensures that you always have a consistent backup of Ops Manager and PAS to restore from.
- Back up frequently, especially before making any changes to your PCF deployment, such as the configuration of any tiles in Ops Manager.

Supported Components

BBR is a binary that can back up and restore BOSH deployments and BOSH Directors. BBR requires that the backup targets supply scripts that implement the backup and restore functions. BBR can communicate securely with external blobstores and databases, using TLS, if these are configured accordingly.

BBR can back up the following components:

- [Pivotal Application Service \(PAS\)](#)
- [BOSH Director](#)

PAS

BBR can back up and restore PAS configured with the following:

- An internal MySQL database or a supported external database. For a list of supported external databases, see the [Supported External Databases](#) section of *Configuring Cloud Foundry for BOSH Backup and Restore* in the Cloud Foundry documentation.
- An internal WebDAV/NFS blobstore, an external Amazon S3 or S3-compatible blobstore, or an external Azure blobstore.

For guidance about backing up unsupported databases and blobstores, see [Unsupported External Blobstores and Databases in PAS](#) below.

BOSH Director

BBR can back up and restore the BOSH Director configured with the following:

- An internal PostgreSQL database or a supported external database. As part of backing up the BOSH Director, BBR backs up the BOSH UAA database and the CredHub database. For a list of supported external databases, see the [Supported External Databases](#) section of *Configuring Cloud Foundry for BOSH Backup and Restore* in the Cloud Foundry documentation.

- An internal blobstore or an external versioned S3 blobstore.

Note: BBR support for backing up and restoring external databases and blobstores varies across PCF versions. For more information, see [External Storage Support Across PCF Versions](#) below.

External Storage Support Across PCF Versions

The following table shows what types of external databases and blobstores you can back up and restore with BBR.

Deployment Type		Ops Manager Version				
		1.0	1.1	1.2	1.3	1.4
	PAS with External Database	Yes	Yes	Yes	Yes	Yes
	BOSH Director with External Database	No	Yes	Yes	Yes	Yes
PAS	PAS with Amazon S3 Blobstore*	Yes	Yes	Yes	Yes	Yes
	BOSH Director with Amazon S3 Blobstore*	No	No	No	No	Yes
	PAS with S3-Compatible Blobstore	Yes	Yes	Yes	Yes	Yes
	BOSH Director with S3-Compatible Blobstore	No	No	No	No	Yes
	PAS with Azure Blobstore	Add-on	Yes	Yes	Yes	Yes
	PAS with GCS Blobstore	No	No	No	No	No
	BOSH Director with GCS Blobstore	No	No	No	No	No

* Any S3 clone that supports the versioning API.

Unsupported External Blobstores and Databases in PAS

If you configured an unsupported external blobstore or an unsupported external database in PAS, see the following guidelines to perform a backup successfully:

Scenario	Action
You configured a supported database and an unsupported external blobstore in PAS.	Follow the PCF backup process and copy the external blobstore by using the IaaS-specific tool.
You configured an unsupported external database and a supported blobstore in PAS.	Follow the PCF backup process and copy the external database by using the IaaS-specific tool.
You configured both an unsupported external database and an unsupported external blobstore in PAS.	Follow the PCF backup process but skip the Back Up Your PAS Deployment section below. Copy the external database and blobstore by using the IaaS-specific tool.

warning: You might encounter inconsistencies between the blobstore and database. If apps do not appear during a restore, re-push those apps.

Backing Up Services

warning: BBR does not currently back up any service data.

Keep in mind the following when backing up services:

- BBR backs up and restores the service bindings, but not the service data.
- You can redeploy on-demand service instances manually during restore, but the data in the instance is not backed up.
- BBR does not back up managed services or their data.
- You can back up and restore brokered services with the procedures documented in this topic and in the [Restoring Pivotal Cloud Foundry from Backup with BBR](#) topic.

Backup Workflow

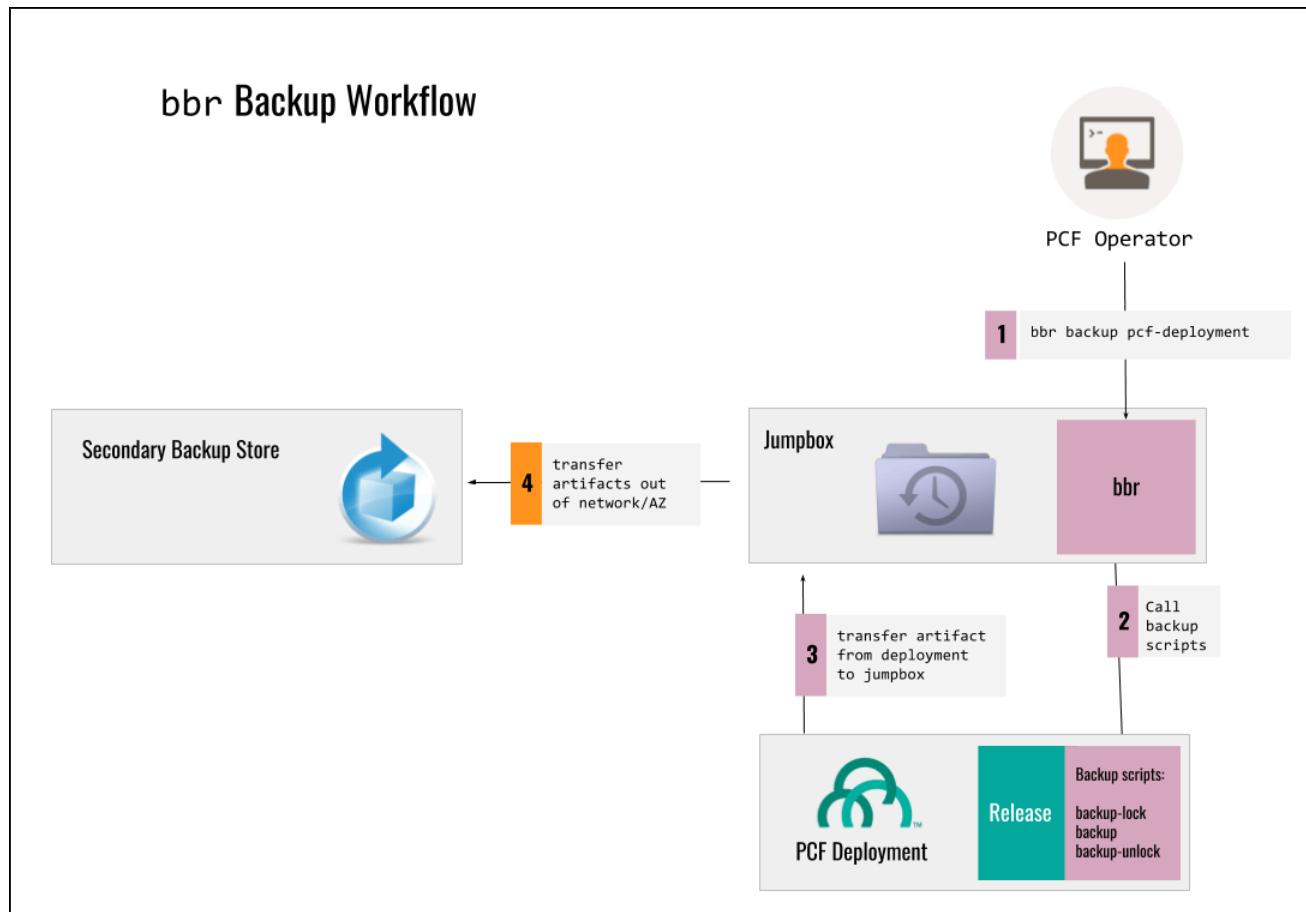
BBR examines the jobs in the BOSH deployment, and triggers the scripts in the following stages:

1. **Pre-backup lock:** The pre-backup lock scripts locks the job so backups are consistent across the cluster.
2. **Backup:** The backup script backs up the release.
3. **Post-backup unlock:** The post-backup unlock script unlocks the job after the backup is complete.

Scripts in the same stage are all triggered together. For instance, BBR triggers all pre-backup lock scripts before any backup scripts. Scripts within a stage can be triggered in any order.

The backup artifacts are drained to the jumpbox, where the operator can transfer them to storage and use them to restore PCF.

The following diagram shows a sample backup flow.



Preparing to Create Your Backup

Step 1: Set Up Your Jumpbox

You must have a jumpbox before you can install BBR to the jumpbox. A jumpbox is a separate, hardened server on your network that provides a controlled means of access to the VMs other computers on your network.

Pivotal recommends using the Ops Manager VM as your jumpbox. If you use the Ops Manager VM as your jumpbox, the path to the root Certificate Authority (CA) certificate is:

```
/var/tempest/workspaces/default/root_ca_certificate
```

You must have this path to run BBR commands.

As an alternative to using the Ops Manager VM as your jumpbox, you can do the following:

- Use Concourse to back up a PCF installation using BBR. For more information, see the [BBR PCF Pipeline Tasks](#) GitHub repository.
- Configure a custom jumpbox. For more information, see [Configuring a custom jumpbox for BBR](#) in the Pivotal Support knowledge base.

Step 2: Record the Cloud Controller Database Encryption Credentials

Perform the following steps to retrieve the Cloud Controller database encryption credentials from the PAS tile:

1. Navigate to Ops Manager in a browser and log in to the Ops Manager Installation Dashboard.
2. Select PAS Credentials and locate the Cloud Controller section.
3. Record the Db Encryption Credentials. You must provide these credentials when you contact [Pivotal Support](#) for help restoring your installation.

Cloud Controller	VM Credentials	Link to Credential
	Db Encryption Credentials	Link to Credential
	Encrypt Key	Link to Credential
	Locket Client Cert	Link to Credential

Step 3: Retrieve BOSH Director Credentials

To use BBR, you must retrieve and record the following credentials:

- Bosh Director Credentials
- Bbr Ssh Credentials
- Uaa Bbr Client Credentials

There are two ways to retrieve BOSH Director credentials:

- Ops Manager Installation Dashboard
- Ops Manager API

Option 1: Ops Manager Installation Dashboard

To retrieve your BOSH Director credentials using the Ops Manager Installation Dashboard, perform the following steps:

1. Navigate to the Ops Manager Installation Dashboard.
2. Click the BOSH Director tile.
3. Click the **Credentials** tab.

BOSH Director

Settings Status Credentials

JOB	NAME	CREDENTIALS
BOSH Director	Vm Credentials	Link to Credential
	Agent Credentials	Link to Credential
	Registry Credentials	Link to Credential
	Director Credentials	Link to Credential
	Postgres Credentials	Link to Credential
	Blobstore Credentials	Link to Credential
	Health Monitor Credentials	Link to Credential
	Uaa Admin User Credentials	Link to Credential
	Uaa Login Client Credentials	Link to Credential
	Uaa Jwt Key	Link to Credential
	Bbr Ssh Credentials	Link to Credential
	Uaa Bbr Client Credentials	Link to Credential
Bosh Commandline Credentials	Link to Credential	
Blobstore Certificate	Link to Credential	

4. Locate Director Credentials.

- Click [Link to Credentials](#) next to it.
- Verify the value of the `identity` field. It should be `director`.
- Copy and record the value of the `password` field.

5. Locate Bbr Ssh Credentials.

- Click [Link to Credentials](#) next to it.
- Copy the value of the `private_key_pem` field.
- Run the following command to reformat the copied value, and save it in the current directory to a file named `PRIVATE-KEY-FILE`:

```
printf -- "YOUR-PRIVATE-KEY" > PRIVATE-KEY-FILE
```

Where:

- `YOUR-PRIVATE-KEY` is the text of your private key.
- `PRIVATE-KEY-FILE` is the path to the private key file you are creating.

For example:

```
$ printf -- "-----BEGIN RSA PRIVATE KEY----- MIIIEkeycontents -----END RSA PRIVATE KEY-----" > bbr_key.pem
```

6. Locate the Uaa Bbr Client Credentials

- Click [Link to Credentials](#) next to it.

- b. Verify the value of the `identity` field. It should be `bbr_client`.
- c. Record the value of the `password` field.

Option 2: Ops Manager API

To retrieve BOSH Director credentials using the Ops Manager API, perform the following steps:

1. Obtain your UAA access token. For more information, see [Access the API](#).

2. Retrieve the **Director Credentials** by performing the following steps:

- a. Run the following command:

```
curl "https://OPS-MAN-FQDN/api/v0/deployed/director/credentials/director_credentials" \
-X GET \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
```

Where:

- `OPS-MAN-FQDN` is the fully-qualified domain name (FQDN) for your Ops Manager deployment.
- `UAA-ACCESS-TOKEN` is your UAA access token.

- b. Verify the value of the `identity` field. It should be `director`.

- c. Record the value of the `password` field.

3. Retrieve the **Bbr Ssh Credentials** by performing the following steps:

- a. Run the following command:

```
curl "https://OPS-MAN-FQDN/api/v0/deployed/director/credentials/bbr_ssh_credentials" \
-X GET \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
```

Where:

- `OPS-MAN-FQDN` is the fully-qualified domain name (FQDN) for your Ops Manager deployment.
- `UAA-ACCESS-TOKEN` is your UAA access token.

- b. Copy the value of the `private_key_pem` field.

c. Run the following command to reformat the copied value, and save it in the current directory to a file named `PRIVATE-KEY-FILE`:

```
printf -- "YOUR-PRIVATE-KEY" > PRIVATE-KEY-FILE
```

Where:

- `YOUR-PRIVATE-KEY` is the text of your private key.
- `PRIVATE-KEY-FILE` is the path to the private key file you are creating.

For example:

```
$ printf -- "-----BEGIN RSA PRIVATE KEY----- MIIIEkeycontents -----END RSA PRIVATE KEY-----" > bbr_key.pem
```

4. Retrieve the **Uaa Bbr Client Credentials** by performing the following steps:

- a. Run the following command:

```
curl "https://OPS-MAN-FQDN/api/v0/deployed/director/credentials/uaa_bbr_client_credentials" \
-X GET \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
```

Where:

- `OPS-MAN-FQDN` is the fully-qualified domain name (FQDN) for your Ops Manager deployment.
- `UAA-ACCESS-TOKEN` is your UAA access token.

- b. Verify the value of the `identity` field. It should be `bbr_client`.

- c. Record the value of the `password` field.

For more information about using the Ops Manager API, see the [Using the Ops Manager API](#) topic.

Step 4: Retrieve BOSH Director Address

Perform the following steps to retrieve the IP address of your BOSH Director from the BOSH Director tile:

1. If you are not using the Ops Manager VM as your jumpbox, install the latest [BOSH CLI](#) on your jumpbox.
2. From the Installation Dashboard in Ops Manager, select **BOSH Director > Status** and record the IP address listed for the Director. You access the

JOB	INDEX	IPS
BOSH Director	0	10.0.0.5

BOSH Director using this IP address.

3. From the command line, log into the BOSH Director, using the IP address that you recorded above, by running the following command:

```
bosh -e DIRECTOR-IP \
--ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE log-in
```

Where:

- o `DIRECTOR-IP` is the BOSH Director IP address recorded above.
- o `PATH-TO-BOSH-SERVER-CERTIFICATE` is the path to the root Certificate Authority (CA) certificate as outlined in [Step 1: Set Up your Jumpbox](#), above.

4. When prompted for **Email ()**, specify `director`.

5. When prompted for **Password ()**, enter the **Director Credentials** that you obtained in [Retrieve BOSH Director Credentials](#).

For example:

```
$ bosh -e 10.0.0.3 \
--ca-cert /var/tempest/workspaces/default/root_ca_certificate log-in
Email (): director
Password (): ****
Successfully authenticated with UAA
Succeeded
```

Step 5: Check Your BOSH Director

Perform the following steps to confirm that your BOSH Director is reachable and can be backed up.

Connect to Your Jumpbox

You can establish a connection to your jumpbox in one of the following ways.

- [Connect with SSH](#)
- [Connect with BOSH_ALL_PROXY](#)

Note: These procedures require you have a jumpbox configured for SSH access. If you do not have a jumpbox, see [Enabling SSH Access](#) and [Tunneling](#) in the BOSH documentation.

Connect with SSH

There are two options available to you to connect to your jumpbox with SSH:

- Ops Manager VM.
- SSH through the command line.

1. To connect to your jumpbox through an Ops Manager VM, perform the following steps:

- a. Log in to the Ops Manager VM:
- b. If you are using the Ops Manager VM as your jumpbox, see the [Log in to the Ops Manager VM with SSH](#) section of *Advanced Troubleshooting with the BOSH CLI* for information about how to use SSH to connect to the Ops Manager VM.

2. To connect to your jumpbox through the CLI, perform the following steps:

- a. Run the following command to SSH into your jumpbox:

```
ssh -i LOCAL-PATH-TO-JUMPBOX-PRIVATE-KEY JUMPBOX-USER@JUMPBOX-ADDRESS
```

Where:

- `LOCAL-PATH-TO-JUMPBOX-PRIVATE-KEY` is the local path to your private key file for the jumpbox host.
- `JUMPBOX-USER` is your jumpbox username.
- `JUMPBOX-ADDRESS` is the IP address of your jumpbox.

 **Note:** If you connect to your jumpbox with SSH, you must run the BBR commands in the following sections from within your jumpbox.

Connect with BOSH_ALL_PROXY

Set and use `BOSH_ALL_PROXY`. Using `BOSH_ALL_PROXY` opens an SSH tunnel with SOCKS5 to the jumpbox. This tunnel enables you to forward requests to the BOSH Director through the jumpbox from your local machine.

Use one of the following methods to create the tunnel:

- **Tunnel created by BOSH CLI:** To provide the BOSH CLI with the SSH credentials it needs to create the tunnel, run the following command:

```
export BOSH_ALL_PROXY=ssh+socks5://JUMPBOX@JUMPBOX-IP:SOCKS-PORT?private_key=JUMPBOX-KEY-FILE
```

Where:

- `JUMPBOX` is the name of your jumpbox.
- `JUMPBOX-IP` is the IP address of the jumpbox.
- `SOCKS-PORT` is the local SOCKS port.
- `JUMPBOX-KEY-FILE` is the local SSH private key for accessing the jumpbox.

- **Tunnel established separately:**

1. To establish the tunnel and make it available on a local port, run the following command:

```
ssh -4 -D SOCKS-PORT -fNC JUMPBOX@JUMPBOX-IP -i JUMPBOX-KEY-FILE
```

Where:

- `SOCKS-PORT` is the local SOCKS port.
- `JUMPBOX` is the name of your jumpbox.
- `JUMPBOX-IP` is the IP address of the jumpbox.
- `JUMPBOX-KEY-FILE` is the local SSH private key for accessing the jumpbox.

2. To provide the BOSH CLI with access to the tunnel through use of the `BOSH_ALL_PROXY` environment variable, run the following command:

```
export BOSH_ALL_PROXY=socks5://localhost:SOCKS-PORT
```

 **Note:** Ensure the SOCKS port is not already in use by a different tunnel/process.

 **Note:** Using `BOSH_ALL_PROXY` can result in longer backup and restore times due to network performance degradation. Because all operations must pass through the proxy, moving backup artifacts can be significantly slower.

Check Your BOSH Director

- Run the BBR `pre-backup-check` command:

```
bbr director \
--private-key-path PRIVATE-KEY-FILE \
--username bbr \
--host HOST \
pre-backup-check
```

Where:

- `PRIVATE-KEY-FILE` is the path to the private key file that you created from **Bbr Ssh Credentials** in [Step 3: Retrieve BOSH Director Credentials](#).
- `HOST` is the address of the BOSH Director. If the BOSH Director is public, `HOST` is a URL, such as `https://my-bosh.xxx.cf-app.com`. Otherwise, `HOST` is the `BOSH-DIRECTOR-IP` which you retrieved in the [Step 4: Retrieve BOSH Director Address](#) section of this topic above.



Note: Use the optional `--debug` flag to enable debug logs. See the [Logging](#) section below for more information.

- After the pre-backup check succeeds, continue to [Step 5: Confirm Backup Restore Node is Deployed](#) below.

If the pre-backup check fails, the BOSH Director might not have the correct backup scripts, or the connection to the BOSH Director might have failed.

Step 6: Confirm Backup Restore Node is Deployed

When BBR backs up PAS, it needs a Backup Restore node. This procedure confirms that the Backup Restore node is deployed.

- In Ops Manager, open the **Pivotal Application Service (PAS)** tile.
- In the **Resource Config** pane, find the **Backup Restore Node** job and check if the **Instances** dropdown is set to `1`.



- If the **Instances** dropdown is not set to `1`:
 - Set the **Instances** dropdown to `1`.
 - Click **Save**.
 - Navigate to the Ops Manager Installation Dashboard.
 - Click **Review Pending Changes**.
 - Review your changes. For more information, see [Reviewing Pending Product Changes](#).
 - Click **Apply Changes**.

Step 7: Identify Your Deployment

- Log in to your BOSH Director.
- To identify the name of the BOSH deployment that contains PCF, run the following command:

```
bosh -e BOSH-DIRECTOR-IP --ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE deployments
```

Where:

- `BOSH-DIRECTOR-IP` is the BOSH Director IP retrieved in the [Step 4: Retrieve BOSH Director Address](#) section above.
- `PATH-TO-BOSH-SERVER-CERTIFICATE` is the path to the Certificate Authority (CA) certificate for the BOSH Director.

```
$ bosh -e BOSH-DIRECTOR-IP --ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE deployments
```

Name	Release(s)
cf-example	push-apps-manager-release/661.1.24 cf-backup-and-restore/0.0.1 binary-buildpack/1.0.11 capi/1.28.0 cf-autoscaling/91 cf-mysql/35 ...

In the above example, the name of the BOSH deployment that contains PCF is `cf-example`.

Step 8: Check Your Deployment

Perform the following steps to check that your BOSH Director is reachable and has a deployment that can be backed up:

1. Run the BBR pre-backup check:

```
bbbr deployment \
--target BOSH-DIRECTOR-IP \
--username BOSH-CLIENT \
--password BOSH-PASSWORD \
--deployment DEPLOYMENT-NAME \
--ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE \
pre-backup-check
```

Where:

- `BOSH-DIRECTOR-IP` is the BOSH Director IP retrieved in [Step 4: Retrieve BOSH Director Address](#).
- `BOSH-CLIENT`, `BOSH-PASSWORD` are the [Uaa Bbr Client Credentials](#), identity and password, that you retrieved in [Step 3: Retrieve BOSH Director Credentials](#).
- `DEPLOYMENT-NAME` is the deployment name retrieved in [Step 6: Identify Your Deployment](#).
- `PATH-TO-BOSH-SERVER-CERTIFICATE` is the path to the root CA certificate for the BOSH Director. If you are using the Ops Manager VM as your jumpbox, the path is `/var/tempest/workspaces/default/root_ca_certificate`.

2. After the pre-backup check succeeds, continue to [Create Your Backup](#) below.

If the pre-backup check fails, the deployment you selected might not have the correct backup scripts, or the connection to the BOSH Director might have failed.

Creating Your Backup

Step 9: Export Installation Settings

Pivotal recommends that you back up your Ops Manager installation settings by exporting frequently.

When exporting your installation settings, keep in mind the following:

- This option is only available after you have deployed at least once.
- Always export your installation settings before following the steps in the [Deploy Ops Manager and Import Installation Settings](#) section of the *Restoring Pivotal Cloud Foundry from Backup with BBR* topic.
- Exporting your installation settings only backs up the settings you configured in Ops Manager. It does not back up your VMs or any external MySQL databases.
- Your Ops Manager settings are encrypted. Make sure you keep track of your Decryption Passphrase because this is needed to restore the Ops Manager settings.

 **Note:** When exporting installation settings for Ops Manager v1.12 and later, releases are not included in the output file. Releases are now included in the BOSH Director backup, therefore, the output file is much smaller than in previous Ops Manager versions.

There are two ways to export Ops Manager installation settings:

- [Use the Ops Manager UI](#)
- [Use the Ops Manager API](#)

Option 1: Use the Ops Manager UI

To export your installation settings using the Ops Manager UI, do the following:

1. From the **Installation Dashboard** in the Ops Manager interface, click your username at the top right navigation.
2. Select **Settings**.
3. Select **Export Installation Settings**.

4. Click **Export Installation Settings**.

Option 2: Use the Ops Manager API

If you want to automate the back up process, you can use the Ops Manager API to export your installation settings.

1. To export your installation settings using the Ops Manager API, run the following command:

```
curl https://OPS-MAN-FQDN/api/v0/installation_asset_collection \
-H "Authorization: Bearer UAA-ACCESS-TOKEN" > installation.zip
```

Where:

- o `OPS-MAN-FQDN` is the fully-qualified domain name (FQDN) for your Ops Manager deployment.
- o `UAA-ACCESS-TOKEN` is your UAA access token. For more information, see [Access the API](#).

Step 10: Back Up Your BOSH Director

Notes:

- The BOSH Director backup can take more than 10 minutes.
- Because the BBR `backup` command can take a long time to complete, Pivotal recommends you run it independently of the SSH session so that the process can continue running even if your connection to the jumpbox fails.
Use `nohup`, a `screen`, or a `tmux` session.

1. Run the BBR `backup` command to back up your BOSH Director:

```
bbr director \
--private-key-path PRIVATE-KEY-FILE \
--username bbr \
--host HOST \
backup
```

Where:

- o `PRIVATE-KEY-FILE` is the path to the private key file that you created from **Bbr Ssh Credentials** in [Step 3: Retrieve BOSH Director Credentials](#).
- o `HOST` is the address of the BOSH Director:
 - If the BOSH Director is public, `HOST` is a URL, such as `https://my-bosh.xxx.cf-app.com`.
 - If the BOSH Director is not public, `HOST` is the `BOSH-DIRECTOR-IP` retrieved in the [Step 4: Retrieve BOSH Director Address](#) section above.

 Note: Use the optional `--debug` flag to enable debug logs. See [Logging](#) below for more information.

If your backup terminates early or fails you need to perform a cleanup.

- To clean up, run the following `backup-cleanup` command:

```
bbr director \
--private-key-path PRIVATE-KEY-FILE \
--username bbr \
--host HOST \
backup-cleanup
```

Where:

- `PRIVATE-KEY-FILE` is the path to the private key file that you created from **Bbr Ssh Credentials** in [Step 3: Retrieve BOSH Director Credentials](#).
- `HOST` is the address of the BOSH Director:
 - If the BOSH Director is public, `HOST` is a URL, such as `https://my-bosh.xxx.cf-app.com`.
 - If the BOSH Director is not public, `HOST` is the `BOSH-DIRECTOR-IP` retrieved in the [Step 4: Retrieve BOSH Director Address](#) section above.

Step 11: Back Up Your PAS Deployment

 Notes:

- Backing up PAS can take more than 5 minutes, and can take considerably longer with larger blobstores or slow network connections. The backup also incurs Cloud Controller downtime during which users are unable to push, scale, or delete apps. Your apps are not affected.
- Because the BBR `backup` command can take a long time to complete, Pivotal recommends you run it independently of the SSH session, so that the process can continue running even if your connection to the jumpbox fails. The command above uses `nohup` but you could also run the command in a `screen` or `tmux` session.

- (Optional) If you are not using an internal blobstore or a versioned S3-compatible external blobstore, create a copy of the blobstore with your IaaS specific tool. Your blobstore backup might be slightly inconsistent with your PAS backup depending on the duration of time between performing the backups.
- Run the BBR `backup` command to back up your PAS deployment:

```
bbr deployment \
--target BOSH-DIRECTOR-IP \
--username BOSH-CLIENT \
--password BOSH-PASSWORD \
--deployment DEPLOYMENT-NAME \
--ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE \
backup
```

Where:

- `BOSH-DIRECTOR-IP` is the BOSH Director IP retrieved in the [Step 4: Retrieve BOSH Director Address](#) section above.
- `BOSH-CLIENT`, `BOSH-PASSWORD` are the **Uaa Bbr Client Credentials**, identity and password, that you retrieved in the [Step 3: Retrieve BOSH Director Credentials](#) section above.
- `DEPLOYMENT-NAME` is the deployment name retrieved in the [Step 6: Identify Your Deployment](#) section above.
- `PATH-TO-BOSH-SERVER-CERTIFICATE` is the path to the Certificate Authority (CA) certificate for the BOSH Director.
- Use the optional `--debug` flag to enable debug logs. For more information, see [Logging](#) below.
- Use the optional `--with-manifest` flag to ensure that BBR downloads your current deployment manifest when backing up. These manifests are included in the Ops Manager export, but are useful for reference.

- After the command successfully completes, continue to [Step 11: Transfer Backup Artifacts](#).

If the command fails, perform the following steps:

- Run the BBR `backup-cleanup` command:

```
bbr deployment \
--target BOSH-DIRECTOR-IP \
--username BOSH-CLIENT \
--password BOSH-PASSWORD \
--deployment DEPLOYMENT-NAME \
--ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE \
backup-cleanup
```

Where:

- `BOSH-DIRECTOR-IP` is the BOSH Director IP retrieved in the [Step 4: Retrieve BOSH Director Address](#) section above.
- `BOSH-CLIENT`, `BOSH-PASSWORD` are the [Uaa Bbr Client Credentials](#), identity and password, that you retrieved in the [Step 3: Retrieve BOSH Director Credentials](#) section above.
- `DEPLOYMENT-NAME` is the deployment name retrieved in the [Step 6: Identify Your Deployment](#) section above.
- `PATH-TO-BOSH-SERVER-CERTIFICATE` is the path to the Certificate Authority (CA) certificate for the BOSH Director.
- Use the optional `--debug` flag to enable debug logs. For more information, see [Logging](#) below.
- Use the optional `--with-manifest` flag to ensure that BBR downloads your current deployment manifest when backing up. These manifests are included in the Ops Manager export, but are useful for reference.

b. Run the BBR `backup` command again after checking the following:

- All the parameters in the command are set.
- The BOSH Director credentials are valid.
- The specified deployment exists.

After Taking Your Backup

Step 12: Transfer Backup Artifacts

After creating your backup successfully, perform the following steps:

1. Move the backup artifacts off the jumpbox to your preferred storage space. The backup created by BBR consists of a directory with the backup artifacts and metadata files. Pivotal recommends compressing and encrypting the files.
2. Make redundant copies of your backup artifacts and store them in multiple locations to minimize the risk of losing backups in the event of a disaster.
3. Attempt to restore every backup to validate the artifacts. Perform the procedures in the next step, [Step 12: Validate Your Backup](#).

Note: Backup artifacts might contain data covered by European Union regulations. For more information, see [European Union's General Data Protection Regulation \(GDPR\)](#) below.

Step 13: (Optional) Validate Your Backup

After backing up PCF, consider validating your backup by restoring it and checking the applications. BBR is designed for disaster recovery and BBR backups are only compatible with environments matching their source environment's configuration.

A **warning:**

- The VMs and disks from the backed-up BOSH Director must not be visible to the new BOSH Director when validating your backup. As a result, Pivotal recommends that you deploy the new BOSH Director to a different IaaS network and account than the VMs and disks of the backed-up BOSH Director.
- **Data services outside of PCF might produce unexpected side effects during restoration.** Restored apps and services might attempt to connect to data services when you restore to a new environment. For example, consider an app that processes mail queues and connects to an external database. When you validate your backup in a test environment, the app might start processing the queue, and this work might be lost.

Validate Your Backup in a Second Environment

To spin up a second environment and test the backup, perform the following steps:

1. Export your Ops Manager installation by performing the steps in the [Step 8: Export Installation Settings](#) section above.
2. Create a new Ops Manager VM in a different network to the original. Ensure that the Ops Manager VM has enough persistent disk to accommodate the files exported in the previous step. Consult the topic for [deploying PCF on your IaaS](#).
3. Ensure that the restore environment is compatible with the backup. For more information, see [Compatibility of Restore](#).
4. Follow the instructions in the [Restoring Pivotal Cloud Foundry from Backup with BBR](#) topic.

Validate Your PAS Backup Only

For a sandbox or other non-production environment, you can optionally perform an *in-place* restore of PAS only. In this case, you restore the PAS backup to the same PCF environment that the backup was created from. Follow the procedures in [Restoring PCF from Backup with BBR](#).

Logging

BBR outputs logs to stdout. By default, BBR logs:

- The backup and restore scripts that it finds
- When it starts or finishes a stage, such as pre-backup scripts or backup scripts
- When the process is complete
- When any error occurs

If you require additional logging, use the optional `--debug` flag to print the following information:

- Logs about the API requests made to the BOSH server
- All commands executed on remote instances
- All commands executed on local environment
- Standard in and standard out streams for the backup and restore scripts when they are executed

Canceling a Backup

If you need to cancel a backup, perform the following steps:

1. Enter `CTRL-C` to terminate the BBR process, then enter `yes` to confirm.
2. Run the BBR `backup-cleanup` command.
 - For a canceled director backup, run the following:

```
bbr director \
--private-key-path PRIVATE-KEY-FILE \
--username bbr \
--host HOST \
backup-cleanup
```

Where:

- `PRIVATE-KEY-FILE` is the path to the private key file that you created from [Bbr Ssh Credentials](#) in [Step 3: Retrieve BOSH Director Credentials](#).

- `HOST` is the address of the BOSH Director:

- If the BOSH Director is public, `HOST` is a URL, such as `https://my-bosh.xxx.cf-app.com`.
- If the BOSH Director is not public, `HOST` is the `BOSH-DIRECTOR-IP` retrieved in the [Step 4: Retrieve BOSH Director Address](#) section above.

- For a canceled deployment backup, run the following:

```
bbr deployment \
--target BOSH-DIRECTOR-IP \
--username BOSH-CLIENT \
--password BOSH-PASSWORD \
--deployment DEPLOYMENT-NAME \
--ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE \
backup-cleanup
```

Where:

- `BOSH-DIRECTOR-IP` is the BOSH Director IP retrieved in the [Step 4: Retrieve BOSH Director Address](#) section above.
- `BOSH-CLIENT`, `BOSH-PASSWORD` are the [Uaa Bbr Client Credentials](#), identity and password, that you retrieved in the [Step 3: Retrieve BOSH Director Credentials](#) section above.
- `DEPLOYMENT-NAME` is the deployment name retrieved in the [Step 6: Identify Your Deployment](#) section above.
- `PATH-TO-BOSH-SERVER-CERTIFICATE` is the path to the Certificate Authority (CA) certificate for the BOSH Director.
- Use the optional `--debug` flag to enable debug logs. For more information, see [Logging](#) above.
- Use the optional `--with-manifest` flag to ensure that BBR downloads your current deployment manifest when backing up. These manifests are included in the Ops Manager export, but are useful for reference.

European Union's General Data Protection Regulation (GDPR)

Backup artifacts might contain personal and other data covered by the European Union's *General Data Protection Regulation (GDPR)*. For example, a PAS backup could contain user email addresses. As such, handle backup artifacts in accordance with organizational policies and applicable laws as they pertain to security, confidentiality, and privacy.

PAS Component Availability During Backup

This topic describes the operational impact of backing up Pivotal Application Service (PAS) components with BBR. To ensure correctness of backups, each component that requires backup has its own set of scripts. The sections in this topic describe the availability of each component during backup of its database.

Cloud Controller

The Cloud Controller is unavailable during backup. Apps and Services continue to run as normal, but you cannot perform operations that require the [Cloud Controller API](#). This includes the following:

- Pushing new apps or creating new services
- Modifying existing apps or services
- Using the clients of the Cloud Controller API, such as the following:
 - The Cloud Foundry Command Line Interface (cf CLI)
 - Apps Manager and its integrations
 - The [Java client](#) used by Spring apps

The backup process for the Cloud Controller is as follows:

Stage	Description
1: Pre-backup lock	The processes running on the Cloud Controller Worker, Cloud Controller, and Clock Global VMs are stopped.
2: Backup	The BBR SDK backup script runs to backup the Cloud Controller database (CCDB), which contains state information for apps on your deployment.
3: Post-backup unlock	The processes start again on the Cloud Controller Worker, Cloud Controller, and Clock Global VMs.

UAA

UAA remains available in read-only mode during backup. This means that you cannot perform write operations for clients, users, groups, identity providers, or zone configuration. However, you can continue performing read operations, such as generating, validating, and revoking tokens. Additionally, UAA continues to authenticate users and authorize requests for users and clients.

The read-only behavior during backup applies to all of the following ways of accessing UAA: the UAA API, the UAA CLI, cf CLI, login screens, and services such as the Single Sign-On Service tile.

The backup process for UAA is as follows:

Stage	Description
1: Pre-backup lock	UAA enters read-only mode.
2: Backup	The BBR SDK backup script runs to backup the UAA database, which contains Cloud Foundry user credentials.
3: Post-backup unlock	UAA exits read-only mode.

Routing API

The [Routing API](#) remains available during backup. However, you cannot perform write operations using the Routing API because the routing database is locked. All read operations offered by the Routing API remain available.

The BBR SDK backup script for the Routing API backs up its database, which contains router groups, routes, and internal implementation information.

Usage Service

The Usage Service is unavailable during backup. You cannot access the API as described in [Monitoring App, Task, and Service Instance Usage](#). Additionally, you cannot view usage and accounting reports as described in [Monitoring Instance Usage with Apps Manager](#).

The backup process for Usage Service is as follows:

 **Note:** The Usage Service runs as a set of Cloud Foundry apps in the `system` org.

Stage	Description
1: Pre-backup lock	The Usage Service apps in the <code>system</code> org stop. This lock occurs before the Cloud Controller and UAA components lock.
2: Backup	The BBR SDK backup script runs to backup the Usage Service database.
3: Post-backup unlock	The Usage Service apps in the <code>system</code> org start again. This unlock occurs after the Cloud Controller and UAA components unlock.

App Autoscaler

The [App Autoscaler service](#) is unavailable during backup. You cannot access the UI or API. For any apps configured to use the App Autoscaler, the service does not scale these apps during backup.

The backup process for App Autoscaler is as follows:

 **Note:** The App Autoscaler service runs as a set of Cloud Foundry apps in the `system` org.

Stage	Description
1: Pre-backup lock	The Autoscaler apps in the <code>system</code> org stop. This lock occurs before the Cloud Controller and UAA components lock.
2: Backup	The BBR SDK backup script runs to backup the App Autoscaler database.
3: Post-backup unlock	The Autoscaler apps in the <code>system</code> org start again. This unlock occurs after the Cloud Controller and UAA components unlock.

NFS Volume Service

The NFS service broker backup scripts rely on the locking of the Cloud Controller to stop traffic to its service. This is because the Cloud Controller is responsible for invoking the NFS service broker.

When the Cloud Controller locks during backup, you cannot create or delete new instances or bindings of a volume service. However, apps already bound to a volume service continue to operate normally during backups.

The NFS service broker backup script performs a backup of the database used to store service instances and service bindings for the NFS service broker.

Notification Service

The Notification Service is not available during backup with BBR due to its dependency on the Cloud Controller. Notifications cannot be sent while the Cloud Controller is unavailable.

Network Policy Server

The Network Policy Server is unavailable during backup. While existing policies are still enforced, you cannot use the cf CLI to add or remove policies for Container Networking as documented in [Administering Container-to-Container Networking](#).

CredHub

Runtime CredHub is unavailable during backup. If the service instance credentials for an app are stored in CredHub, the app cannot fetch those credentials during backup. In some cases, apps may not start if they cannot fetch credentials for a service instance binding.

Restoring PCF from Backup with BBR

This topic describes the procedure for restoring your critical backend Pivotal Cloud Foundry (PCF) components with BOSH Backup and Restore (BBR), a command-line tool for backing up and restoring BOSH deployments. To perform the procedures in this topic, you must have backed up PCF by following the steps in the [Backing Up Pivotal Cloud Foundry with BBR](#) topic.

To view the BBR release notes, see [BOSH Backup and Restore Release Notes](#).

The procedures described in this topic prepare your environment for PCF, deploy Ops Manager, import your installation settings, and use BBR to restore your PCF components.

A warning: Restoring PCF with BBR is a destructive operation. If the restore fails, the new environment may be left in an unusable state and require re-provisioning. Only perform the procedures in this topic for the purpose of disaster recovery, such as recreating PCF after a storage-area network (SAN) corruption.

A warning: When validating your backup, the VMs and disks from the backed-up BOSH Director should not be visible to the new BOSH Director. As a result, Pivotal recommends that you deploy the new BOSH Director to a different IaaS network and account than the VMs and disks of the backed up BOSH Director.

Note: If the BOSH Director you are restoring had any deployments that were deployed manually rather than through an Ops Manager tile, you must restore them manually at the end of the process. For more information, see [\(Optional\) Step 16: Restore Non-Tile Deployments](#).

Compatibility of Restore

This section describes the restrictions for a backup artifact to be restorable to another environment. This section is for guidance only, and Pivotal highly recommends that operators validate their backups by using the backup artifacts in a restore.

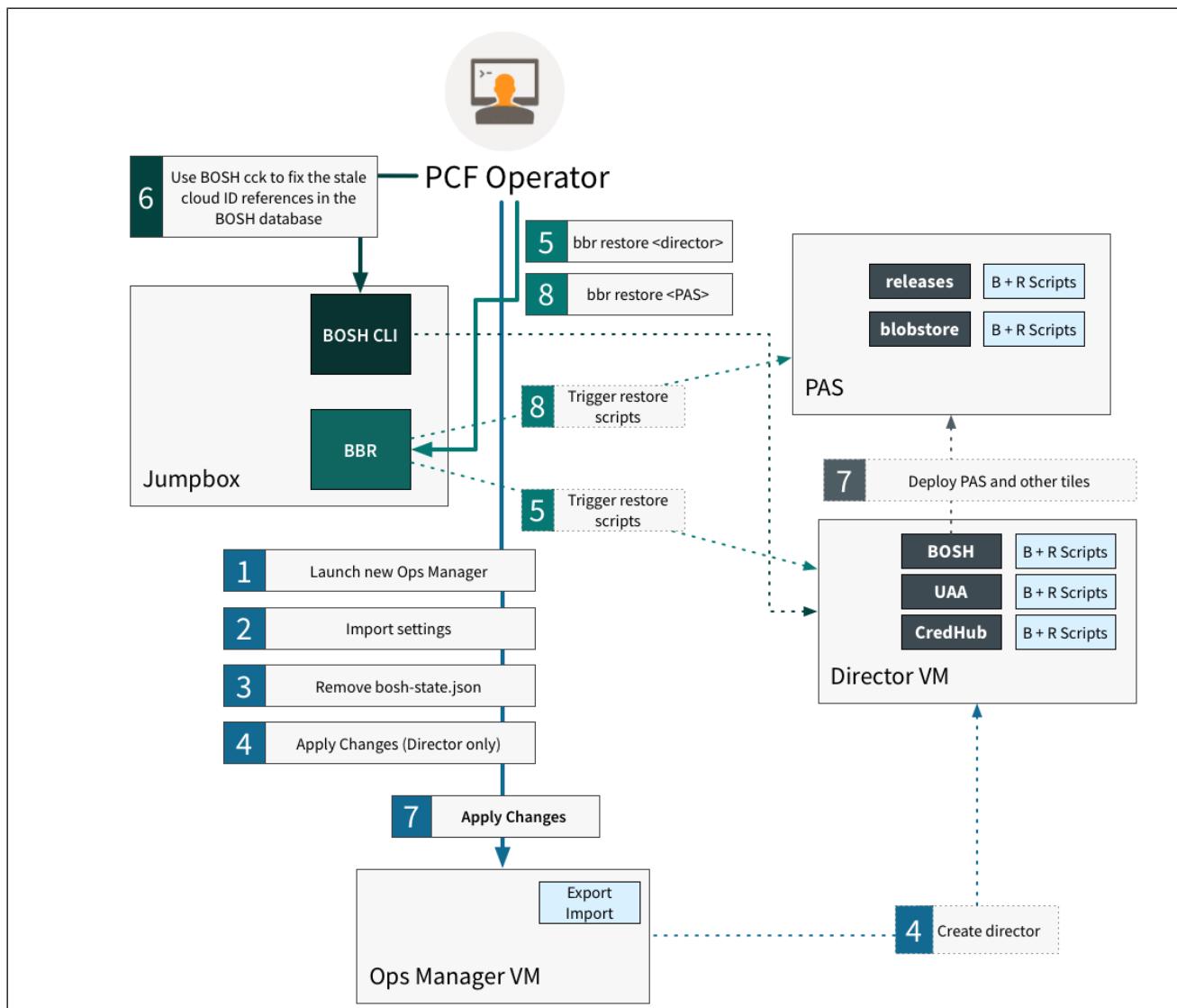
Consult the following restrictions for a backup artifact to be restorable:

- **CIDR ranges:** BBR requires the IP address ranges to be the same in the restore environment as in the backup environment.
- **Topology:** BBR requires the BOSH topology of a deployment to be the same in the restore environment as it was in the backup environment.
- **Naming of instance groups and jobs:** For any deployment that implements the backup and restore scripts, the instance groups and jobs must have the same names.
- **Number of instance groups and jobs:** For instance groups and jobs that have backup and restore scripts, there must be the same number of instances.
- **Limited validation:** BBR puts the backed up data into the corresponding instance groups and jobs in the restored environment, but can't validate the restore beyond that. For example, if the MySQL encryption key is different in the restore environment, the BBR restore might succeed although the restored MySQL database is unusable.
- **PCF version:** BBR can restore to the same version of PCF that was backed up. BBR does not support restoring to other major, minor, or patch releases.

Note: A change in VM size or underlying hardware will not affect BBR's ability to restore data, as long as there is adequate storage space to restore the data.

Restore Workflow

Click the diagram below to see the full-size image.



The diagram above shows the flow of the PCF restore process in a series of steps performed by the PCF operator. The following steps will be covered in more detail throughout this topic.

1. **Launch new Ops Manager:** Perform the procedures for your IaaS to deploy Ops Manager. See part one of the [Deploy Ops Manager and Import Installation Settings](#) step below for more information.
2. **Import settings:** You can import settings either with the Ops Manager UI or API. See part two of the [Deploy Ops Manager and Import Installation Settings](#) step below for more information.
3. **Remove `bosh-state.json`:** SSH into your Ops Manager VM and delete the `bosh-state.json` file. See the [Remove BOSH State File](#) step below for more information.
4. **Apply Changes (Director only):** Use the Ops Manager API to only deploy the BOSH Director. See the [Deploy the BOSH Director](#) step below for more information.
5. **bbr restore <director>:** Run the BBR restore command from your jumpbox to restore the BOSH Director. See the [Restore the BOSH Director](#) step below for more information.
6. **Use BOSH cck to fix the stale cloud ID references in the BOSH database:** For each deployment in the BOSH Director, you will need to run a `bosh cloud-check` command. See the [Remove Stale Cloud IDs for All Deployments](#) step for more information.
7. **Apply Changes:** On the Ops Manager Installation Dashboard, click [Review Pending Changes](#), review your changes, and then click [Apply Changes](#). For more information, see [Reviewing Pending Product Changes](#).
8. **bbr restore <PAS>:** Run the BBR restore command from your jumpbox to restore PAS. See the [Restore PAS](#) step below for more information.

Prepare to Restore

This section provides the steps you need to perform before restoring your PCF backup with BBR.

Step 1: (Optional) Prepare Your Environment

In an event of a disaster, you may lose not only your VMs and disks, but your IaaS resources as well, such as networks and load balancers.

If you need to recreate your IaaS resources, prepare your environment for PCF by following the instructions specific to your IaaS in [Installing Pivotal Cloud Foundry](#).

Note: The instructions for installing PCF on Amazon Web Services (AWS) and OpenStack combine the procedures for preparing your environment and deploying Ops Manager into a single topic. The instructions for the other supported IaaSes split these procedures into two separate topics.

If you recreate your IaaS resources, you must also add those resources to Ops Manager by performing the procedures in the [Step 3: \(Optional\) Configure Ops Manager for New Resources](#) section.

Step 2: Deploy Ops Manager and Import Installation Settings

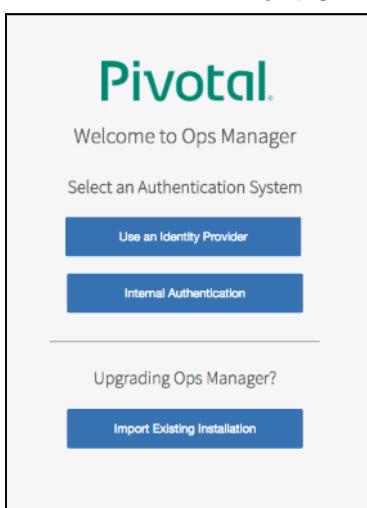
1. Perform the procedures for your IaaS to deploy Ops Manager:

- AWS:
 - If you configured AWS manually, see [Step 12: Launch an Ops Manager API](#) through [Step 19: Create a MySQL Database Using AWS RDS](#) of [Installing PCF on AWS Manually](#).
 - If you used Terraform to install PCF on AWS, see [Installing PCF on AWS Using Terraform](#).
- GCP: See [Deploying Ops Manager on GCP Manually](#).
- OpenStack: See [Step 4: Create Ops Manager Image](#) through [Step 9: Create a DNS Entry](#) in [Deploying Ops Manager to OpenStack](#).
- vSphere: See [Deploying Ops Manager on vSphere](#).

2. Import your installation settings. This can be done in two ways:

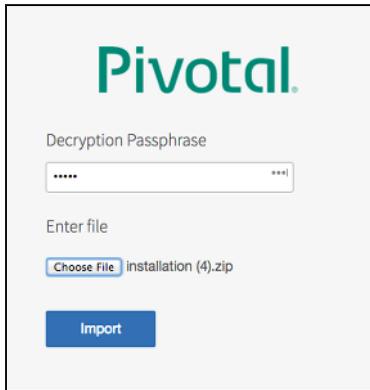
a. Using the Ops Manager UI

- i. Access your new Ops Manager by navigating to `YOUR-OPS-MAN-FQDN` in a browser.
- ii. On the **Welcome to Ops Manager** page, click **Import Existing Installation**.



iii. In the import panel, perform the following tasks:

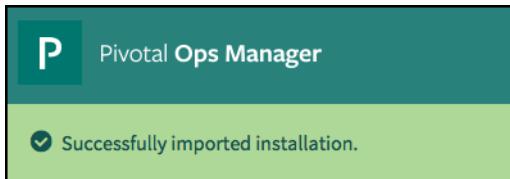
- Enter the **Decryption Passphrase** in use when you exported the installation settings from Ops Manager.
- Click **Choose File** and browse to the installation zip file that you exported in the [Step 9: Export Installation Settings](#) section of the [Backing Up Pivotal Cloud Foundry with BBR](#) topic.



iv. Click **Import**.

Note: Some browsers do not provide feedback on the status of the import process, and may appear to hang. The import process takes at least 10 minutes, and takes longer the more tiles that were present on the backed-up Ops Manager.

v. A **Successfully imported installation** message appears upon completion.



b. Using the Ops Manager API:

```
curl "https://OPS-MAN-FQDN/api/v1/installation_asset_collection" \
-X POST \
-H "Authorization: Bearer UAA-ACCESS-TOKEN" \
-F 'installation[file]=@installation.zip' \
-F 'passphrase=DECRYPTION-PASSPHRASE'
```

Where:

- **OPS-MAN-FQDN** is the fully-qualified domain name (FQDN) for your Ops Manager deployment.
- **UAA-ACCESS-TOKEN** is the UAA access token. For more information about how to retrieve this token, see [Using the Ops Manager API](#).
- **DECRYPTION-PASSPHRASE** is the decryption passphrase in use when you exported the installation settings from Ops Manager.

warning: Do not click **Apply Changes** in Ops Manager until the instruction in Step 14: Redeploy PAS.

Step 3: (Optional) Configure Ops Manager for New Resources

If you recreated IaaS resources such as networks and load balancers by following the steps in the [Step 1: \(Optional\) Prepare Your Environment](#) section above, perform the following steps to update Ops Manager with your new resources:

1. Enable Ops Manager advanced mode. For more information, see [How to Enable Advanced Mode in the Ops Manager](#) in the Pivotal Knowledge Base.

Note: In advanced mode Ops Manager will allow you to make changes that are normally disabled. You may see warning messages when you save changes.

2. Navigate to the Ops Manager Installation Dashboard and click the BOSH Director tile.

3. If you are using Google Cloud Platform (GCP), click **Google Config** and update:

- a. **Project ID** to reflect the GCP project ID.
- b. **Default Deployment Tag** to reflect the environment name.
- c. **AuthJSON** to reflect the service account.

4. Click **Create Networks** and update the network names to reflect the network names for the new environment.

5. If your BOSH Director had an external hostname, you must change it in **Director Config > Director Hostname** to ensure it does not conflict with the hostname of the backed up Director.
6. Return to the Ops Manager Installation Dashboard and click the Pivotal Application Service (PAS) tile.
7. Click **Resource Config**. If necessary for your IaaS, enter the name of your new load balancers in the **Load Balancers** column.
8. If necessary, click **Networking** and update the load balancer SSL certificate and private key under **Certificates and Private Keys for HAProxy and Router**.
9. If your environment has a new DNS address, update the old environment DNS entries to point to the new load balancer addresses. For more information, see the [Step 4: Configure Networking](#) section of the *Using Your Own Load Balancer* topic and follow the link to the instructions for your IaaS.
10. If your PAS uses an external blobstore, ensure that the PAS tile is configured to use a different blobstore, otherwise it will attempt to connect to the blobstore that the backed up PAS is using.
11. Ensure your **System Domain** and **Apps Domain** under **PAS Domains** are updated to refer to the new environment's domains.
12. Ensure that there are no outstanding warning messages in the BOSH Director tile, then disable Ops Manager advanced mode. For more information, see [How to Enable Advanced Mode in the Ops Manager](#) in the Pivotal Knowledge Base.

Step 4: Remove BOSH State File

1. SSH into your Ops Manager VM. For more information, see the [SSH into Ops Manager](#) section of the *Advanced Troubleshooting with the BOSH CLI* topic.
2. To delete the `/var/tempest/workspaces/default/deployments/bosh-state.json` file, run the following on the Ops Manager VM:

```
sudo rm /var/tempest/workspaces/default/deployments/bosh-state.json
```
3. In a browser, navigate to your Ops manager's fully-qualified domain name.
4. Log in to Ops Manager.

Step 5: Deploy the BOSH Director

Use the Ops Manager API or the checkbox on the Review Pending Changes page to deploy the BOSH Director by itself.

Step 6: Transfer Artifacts to Jumpbox

In the [Step 9: Back Up Your PAS Deployment](#) section of the *Backing Up Pivotal Cloud Foundry with BBR* topic, in the *After Taking the Backups* section you moved the TAR and metadata files of the backup artifacts off your jumpbox to your preferred storage space. Now you must transfer those files back to your jumpbox.

Restore Your Backup

This section provides the steps you need to perform to restore your PCF backup with BBR.

Step 7: Retrieve BOSH Director Credentials

To use BBR, you must retrieve and record the following credentials:

- **Bosh Director Credentials**
- **Bbr Ssh Credentials**
- **Uaa Bbr Client Credentials**

There are two ways to retrieve BOSH Director credentials:

- Ops Manager Installation Dashboard
- Ops Manager API

Option 1: Ops Manager Installation Dashboard

To retrieve your BOSH Director credentials using the Ops Manager Installation Dashboard, perform the following steps:

1. Navigate to the Ops Manager Installation Dashboard.
2. Click the BOSH Director tile.
3. Click the **Credentials** tab.

JOB	NAME	CREDENTIALS
BOSH Director	Vm Credentials	Link to Credential
	Agent Credentials	Link to Credential
	Registry Credentials	Link to Credential
	Director Credentials	Link to Credential
	Postgres Credentials	Link to Credential
	Blobstore Credentials	Link to Credential
	Health Monitor Credentials	Link to Credential
	Uaa Admin User Credentials	Link to Credential
	Uaa Login Client Credentials	Link to Credential
	Uaa Jwt Key	Link to Credential
	Bbr Ssh Credentials	Link to Credential
	Uaa Bbr Client Credentials	Link to Credential
	Bosh Commandline Credentials	Link to Credential
	Blobstore Certificate	Link to Credential

4. Locate Director Credentials.

- a. Click **Link to Credentials** next to it.
- b. Verify the value of the `identity` field. It should be `director`.
- c. Copy and record the value of the `password` field.

5. Locate Bbr Ssh Credentials.

- a. Click **Link to Credentials** next to it.
- b. Copy the value of the `private_key_pem` field.
- c. Run the following command to reformat the copied value, and save it in the current directory to a file named `PRIVATE-KEY-FILE`:

```
printf -- "YOUR-PRIVATE-KEY" > PRIVATE-KEY-FILE
```

Where:

- `YOUR-PRIVATE-KEY` is the text of your private key.
- `PRIVATE-KEY-FILE` is the path to the private key file you are creating.

For example:

```
$ printf -- "-----BEGIN RSA PRIVATE KEY----- MIIkeycontents -----END RSA PRIVATE KEY-----" > bbr_key.pem
```

6. Locate the Uaa Bbr Client Credentials

- a. Click **Link to Credentials** next to it.
- b. Verify the value of the `identity` field. It should be `bbr_client`.
- c. Record the value of the `password` field.

Option 2: Ops Manager API

To retrieve BOSH Director credentials using the Ops Manager API, perform the following steps:

1. Obtain your UAA access token. For more information, see [Access the API](#).

2. Retrieve the **Director Credentials** by performing the following steps:

- a. Run the following command:

```
curl "https://OPS-MAN-FQDN/api/v0/deployed/director/credentials/director_credentials" \
-X GET \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
```

Where:

- `OPS-MAN-FQDN` is the fully-qualified domain name (FQDN) for your Ops Manager deployment.
 - `UAA-ACCESS-TOKEN` is your UAA access token.
- b. Verify the value of the `identity` field. It should be `director`.
 - c. Record the value of the `password` field.

3. Retrieve the **Bbr Ssh Credentials** by performing the following steps:

- a. Run the following command:

```
curl "https://OPS-MAN-FQDN/api/v0/deployed/director/credentials/bbr_ssh_credentials" \
-X GET \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
```

Where:

- `OPS-MAN-FQDN` is the fully-qualified domain name (FQDN) for your Ops Manager deployment.
 - `UAA-ACCESS-TOKEN` is your UAA access token.
- b. Copy the value of the `private_key_pem` field.
 - c. Run the following command to reformat the copied value, and save it in the current directory to a file named `PRIVATE-KEY-FILE`:

```
printf -- "YOUR-PRIVATE-KEY" > PRIVATE-KEY-FILE
```

Where:

- `YOUR-PRIVATE-KEY` is the text of your private key.
- `PRIVATE-KEY-FILE` is the path to the private key file you are creating.

For example:

```
$ printf -- "-----BEGIN RSA PRIVATE KEY----- MIIkeycontents -----END RSA PRIVATE KEY-----" > bbr_key.pem
```

4. Retrieve the **Uaa Bbr Client Credentials** by performing the following steps:

- a. Run the following command:

```
curl "https://OPS-MAN-FQDN/api/v0/deployed/director/credentials/uaa_bbr_client_credentials" \
-X GET \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
```

Where:

- `OPS-MAN-FQDN` is the fully-qualified domain name (FQDN) for your Ops Manager deployment.
- `UAA-ACCESS-TOKEN` is your UAA access token.

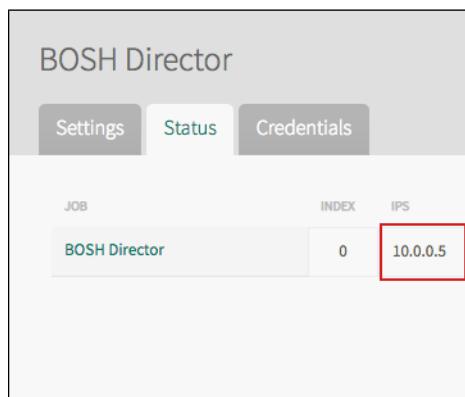
- Verify the value of the `identity` field. It should be `bbr_client`.
- Record the value of the `password` field.

For more information about using the Ops Manager API, see the [Using the Ops Manager API](#) topic.

Step 8: Retrieve BOSH Director Address

Perform the following steps to retrieve the IP address of your BOSH Director from the BOSH Director tile:

1. If you are not using the Ops Manager VM as your jumpbox, install the latest [BOSH CLI](#) on your jumpbox.
2. From the Installation Dashboard in Ops Manager, select **BOSH Director > Status** and record the IP address listed for the Director. You access the



BOSH Director using this IP address.

3. From the command line, log into the BOSH Director, using the IP address that you recorded above, by running the following command:

```
bosh -e DIRECTOR-IP \
--ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE log-in
```

Where:

- `DIRECTOR-IP` is the BOSH Director IP address recorded above.
- `PATH-TO-BOSH-SERVER-CERTIFICATE` is the path to the root Certificate Authority (CA) certificate as outlined in [Step 1: Set Up your Jumpbox](#), above.

4. When prompted for **Email ()**, specify `director`.

5. When prompted for **Password ()**, enter the **Director Credentials** that you obtained in [Retrieve BOSH Director Credentials](#).

For example:

```
$ bosh -e 10.0.0.3 \
--ca-cert /var/tempest/workspaces/default/root_ca_certificate log-in
Email (): director
Password (): ****
Successfully authenticated with UAA
Succeeded
```

Step 9: Restore the BOSH Director

Notes:

- The BBR BOSH Director restore command can take at least 15 minutes to complete.

- Pivotal recommends that you run it independently of the SSH session, so that the process can continue running even if your connection to the jumpbox fails. The command above uses `nohup` but you could also run the command in a `screen` or `tmux` session.

- SSH into your jumpbox. If you are using the Ops Manager VM as your jumpbox, see the [Log in to the Ops Manager VM with SSH](#) section of *Advanced Troubleshooting with the BOSH CLI* for procedures on how to use SSH to connect to the Ops Manager VM.
- Ensure the BOSH Director backup artifact is in the folder you from which you will run BBR.
- Run the BBR restore command from your jumpbox to restore the BOSH Director:

```
bbr director \
--private-key-path PRIVATE-KEY-FILE \
--username bbr \
--host HOST \
restore \
--artifact-path PATH-TO-DIRECTOR-BACKUP
```

Where:

- `PATH-TO-DIRECTOR-BACKUP` is the path to the Director backup you want to restore.
- `PRIVATE-KEY-FILE` is the path to the private key file you created in [Step 7: Retrieve BOSH Director Credentials](#).
- `HOST` is the address of the BOSH Director.
 - If the BOSH Director is public, `HOST` is a URL, such as `https://my-bosh.xxx.cf-app.com`.
 - If the BOSH Director is not public, `HOST` is the `BOSH-DIRECTOR-IP` retrieved in [Step 8: Retrieve BOSH Director Address](#).
- Use the optional `--debug` flag if you want to enable debug logs. See the [Logging](#) section of the *Backing Up Pivotal Cloud Foundry with BBR* topic for more information.

- After the command succeeds, continue to [Step 10: Identify Your Deployment](#).

If the command fails, complete the following procedure:

- Run the BBR restore-cleanup command:

```
bbr director \
--private-key-path PRIVATE-KEY-FILE \
--username bbr \
--host HOST \
restore-cleanup
```

Where:

- `PRIVATE-KEY-FILE` is the path to the private key file you created in [Step 7: Retrieve BOSH Director Credentials](#).
- `HOST` is the address of the BOSH Director.
 - If the BOSH Director is public, `HOST` is a URL, such as `https://my-bosh.xxx.cf-app.com`.
 - If the BOSH Director is not public, `HOST` is the `BOSH-DIRECTOR-IP` retrieved in [Step 8: Retrieve BOSH Director Address](#).

- Run the BBR restore command again after checking the following:

- All the parameters in the command are set.
- The BOSH Director credentials are valid.
- The specified deployment exists.
- The source deployment is compatible with the target deployment.
- That the jumpbox can reach the BOSH Director.

Step 10: Identify Your Deployment

- Log in to your BOSH Director.
- To identify the name of the BOSH deployment that contains PCF, run the following command:

```
bosh -e BOSH-DIRECTOR-IP --ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE deployments
```

Where:

- `BOSH-DIRECTOR-IP` is the BOSH Director IP retrieved in [Step 8: Retrieve BOSH Director Address](#).

- `PATH-TO-BOSH-SERVER-CERTIFICATE` is the path to the Certificate Authority (CA) certificate for the BOSH Director, if the certificate is not verifiable by the local machine's certificate chain.

For example:

```
$ bosh -e BOSH-DIRECTOR-IP --ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE deployments
```

Name	Release(s)
cf-example	push-apps-manager-release/661.1.24 cf-backup-and-restore/0.0.1 binary-buildpack/1.0.11 capi/1.28.0 cf-autoscaling/91 cf-mysql/35 ...

In the above example, the name of the BOSH deployment that contains PCF is `cf-example`. `PATH-TO-BOSH-SERVER-CERTIFICATE` is the path to the root Certificate Authority (CA) certificate for the BOSH Director. If you are using the Ops Manager VM as your jumpbox, the path is `/var/tempest/workspaces/default/root_ca_certificate`.

Step 11: Remove Stale Cloud IDs for All Deployments

1. Review the deployments listed when performing [Step 10: Identify Your Deployment](#).
2. To reconcile the BOSH Director's internal state with the state in the IaaS, run the following command for each deployment:

```
bosh -e BOSH-DIRECTOR-IP \
--ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE \
-d DEPLOYMENT-NAME -n cck \
--resolution delete_disk_reference \
--resolution delete_vm_reference
```

Where:

- `BOSH-DIRECTOR-IP` is the BOSH Director IP retrieved in [Step 8: Retrieve BOSH Director Address](#).
- `PATH-TO-BOSH-SERVER-CERTIFICATE` is the path to the Certificate Authority (CA) certificate for the BOSH Director, if the certificate is not verifiable by the local machine's certificate chain.
- `DEPLOYMENT-NAME` is the deployment name retrieved in [Step 10: Identify Your Deployment](#).

3. To delete disk references, run the following command:

```
bosh cloud-check
```

If the `bosh cloud-check` command does not successfully delete disk references, and you see a message similar to the following, perform the additional procedures in the [Remove Unused Disks](#) section below.

```
Scanning 19 persistent disks: 19 OK, 0 missing ...
```

Step 12: Redeploy PAS

Determine the Required Stemcells

Perform either the following procedures to determine which stemcell is used by PAS:

- Review the Stemcell Library:
 1. Go to [Stemcell Library](#).
 2. Record the PAS stemcell release number from the **Staged** column.
- Review a Stemcell List Using BOSH CLI:
 1. To retrieve the stemcell release using the BOSH CLI, run the following command:

```
bosh -e BOSH-DIRECTOR-IP deployments
```

Where `BOSH-DIRECTOR-IP` is the BOSH Director IP retrieved in [Step 8: Retrieve BOSH Director Address](#).

For example:

```
$ bosh -e BOSH-DIRECTOR-IP deployments
Using environment '10.0.0.5' as user 'director' (bosh.*.read, openid, bosh.*.admin, bosh.read, bosh.admin)

Name          Release(s)      Stemcell(s)      Team(s) Cloud Config
cf-9cb6995b7d746cd77438 push-apps-manager-release/661.1.24 bosh-google-kvm-ubuntu-trusty-go_agent/3421.9 - latest
...
...
```

For more information about stemcells in Ops Manager, see [Importing and Managing Stemcells](#).

Upload Stemcells

1. Download the stemcell from [Pivotal Network](#).
2. Run the following command to upload the stemcell used by PAS:

```
bosh -e BOSH-DIRECTOR-IP \
-d DEPLOYMENT-NAME \
--ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE \
upload-stemcell \
--fix PATH-TO-STEMCELL
```

Where:

- `BOSH-DIRECTOR-IP` is the BOSH Director IP retrieved in [Step 8: Retrieve BOSH Director Address](#).
- `DEPLOYMENT-NAME` is the deployment name retrieved in [Step 10: Identify Your Deployment](#).
- `PATH-TO-BOSH-SERVER-CERTIFICATE` is the path to the Certificate Authority (CA) certificate for the BOSH Director, if the certificate is not verifiable by the local machine's certificate chain.
- `PATH-TO-STEMCELL` is the path to your tile's stemcell.

3. To ensure the stemcells for all of your other installed tiles have been uploaded, repeat the last step, running the `bosh upload-stemcell --fix PATH-TO-STEMCELL` command, for each stemcell that is different from the already uploaded PAS stemcell.

R redeploy PAS

1. From the Ops Manager Installation Dashboard, navigate to [PAS Resource Config](#).
 2. Ensure the number of instances for MySQL Server is set to `1`.
- A warning:** Restore will fail if there is not exactly one MySQL Server instance deployed.
3. Ensure that all errands needed by your system are set to run.
 4. Return to the Ops Manager Installation Dashboard.
 5. Click [Review Pending Changes](#).
 6. Review your changes. For more information, see [Reviewing Pending Product Changes](#).
 7. Click [Apply Changes](#) to redeploy.

Step 13: (Optional) Restore Service Data

A warning: BBR does not back up or restore any service data.

For this step, restore data to pre-provisioned service tiles.

The procedures for restoring service data vary. Consult the documentation for your service tile for more information.

For example, if you are using Redis for PCF v1.14, see [Using BOSH Backup and Restore \(BBR\)](#).

Step 14: Restore PAS

Notes:

- The BBR PAS restore command can take at least 15 minutes to complete.
- Pivotal recommends that you run it independently of the SSH session, so that the process can continue running even if your connection to the jumpbox fails. The command above uses `nohup` but you could also run the command in a `screen` or `tmux` session.

1. Refer to the table in [Enabling External Blobstore Backups](#). If external blobstore support is not included in the version of Ops Manager and PAS you are using, restore the external blobstore with your IaaS-specific tools before restoring PAS.
2. Run the BBR restore command from your jumpbox to restore PAS:

```
bbr deployment \
--target BOSH-DIRECTOR-IP \
--username BOSH-CLIENT \
--password BOSH-PASSWORD \
--deployment DEPLOYMENT-NAME \
--ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE \
restore \
--artifact-path PATH-TO-PAS-BACKUP
```

Where:

- `BOSH-DIRECTOR-IP` is the BOSH Director IP retrieved in [Step 8: Retrieve BOSH Director Address](#).
- `BOSH-CLIENT`, `BOSH-PASSWORD` are the [Uaa Bbr Client Credentials](#), identity and password, that you retrieved in [Step 7: Retrieve BOSH Director Credentials](#).
- `DEPLOYMENT-NAME` is the deployment name retrieved in [Step 10: Identify Your Deployment](#).
- `PATH-TO-BOSH-SERVER-CERTIFICATE` is the path to the Certificate Authority (CA) certificate for the BOSH Director, if the certificate is not verifiable by the local machine's certificate chain.
- `PATH-TO-PAS-BACKUP` is the path to the PAS backup you want to restore.

3. If desired, scale the MySQL Server job back up to its previous number of instances by navigating to the **Resource Config** section of the PAS tile. After scaling the job, return to the Ops Manager Installation Dashboard.
4. Click **Review Pending Changes**.
5. Review your changes. For more information, see [Reviewing Pending Product Changes](#).
6. Click **Apply Changes** to deploy.

Step 15: (Optional) Restore On-Demand Service Instances

If you have on-demand service instances provisioned by an on-demand service broker, perform the following steps to restore them after successfully restoring PAS:

1. Navigate to an on-demand service tile in the **Installation Dashboard**.
2. Select the **Errands** tab.
3. Ensure the **Upgrade All Service Instances** errand is **On**.
4. Repeat for all on-demand service tiles.
5. Return to the **Installation Dashboard**.
6. Click **Review Pending Changes**, review your changes, and then click **Apply Changes**. For more information, see [Reviewing Pending Product Changes](#). This will include running the Upgrade All Service Instances errand for the on-demand service, which will redeploy the on-demand service instances.
7. (Optional) Restore service data to every on-demand service instance.

8. Any app on PAS bound to an on-demand service instance may need to be restarted to start consuming the restored on-demand service instances.

Step 16: (Optional) Restore Non-Tile Deployments

If you have any deployments that were deployed manually with the BOSH Director rather than through an Ops Manager tile, perform the following steps to restore the VMs.

1. To obtain a list of all deployments on your BOSH Director, run the following command:

```
bosh -e BOSH-DIRECTOR-IP \
--ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE \
deployments
```

Where:

- o `BOSH-DIRECTOR-IP` is the BOSH Director IP retrieved in [Step 8: Retrieve BOSH Director Address](#).
- o `PATH-TO-BOSH-SERVER-CERTIFICATE` is the path to the Certificate Authority (CA) certificate for the BOSH Director, if the certificate is not verifiable by the local machine's certificate chain.

2. Identify the names of the deployments that you need to restore. Do not include the deployments from Ops Manager tiles.

3. Run the following command for each deployment you need to restore:

```
bosh -n -e BOSH-DIRECTOR-IP \
--ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE \
-d DEPLOYMENT-NAME \
cck --resolution=recreate_vm
```

Where:

- o `BOSH-DIRECTOR-IP` is the BOSH Director IP retrieved in [Step 8: Retrieve BOSH Director Address](#).
- o `PATH-TO-BOSH-SERVER-CERTIFICATE` is the path to the Certificate Authority (CA) certificate for the BOSH Director, if the certificate is not verifiable by the local machine's certificate chain.
- o `DEPLOYMENT-NAME` is the deployment name retrieved in [Step 10: Identify Your Deployment](#).

4. Run the following command to verify the status of the VMs in each deployment:

```
bosh -e BOSH-DIRECTOR-IP \
--ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE \
-d DEPLOYMENT-NAME \
vms
```

Where:

- o `BOSH-DIRECTOR-IP` is the BOSH Director IP retrieved in [Step 8: Retrieve BOSH Director Address](#).
- o `PATH-TO-BOSH-SERVER-CERTIFICATE` is the path to the Certificate Authority (CA) certificate for the BOSH Director, if the certificate is not verifiable by the local machine's certificate chain.
- o `DEPLOYMENT-NAME` is the deployment name retrieved in [Step 10: Identify Your Deployment](#).

The process state for all VMs should show as `running`.

After Restoring Your Backup

This section provides the steps you need to perform after restoring your PCF backup with BBR.

Step 17: Remove Unused Disks

 **warning:** This is a very destructive operation.

Disks from a previous deployment will prevent recreated deployments from working.

Use BOSH To Clean Up Disks

1. To clean up disk references, run the following command:

```
bosh cloud-check
```

Manually Clean Up Disks

If `bosh cloud-check` does not clean up all disk references, you must manually delete the remaining disks.

1. To delete the remaining disks, perform one of the following procedures:

- o Use the BOSH CLI to delete the disks by performing the following steps:

1. Target the redeployed BOSH Director using the BOSH CLI by performing the procedures in [Step 8: Retrieve BOSH Director Address](#).
2. List the deployments by running the following command:

```
bosh -e BOSH-DIRECTOR-IP \
--ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE deployments
```

Where:

- `BOSH-DIRECTOR-IP` is the BOSH Director IP retrieved in [Step 8: Retrieve BOSH Director Address](#).
- `PATH-TO-BOSH-SERVER-CERTIFICATE` is the path to the Certificate Authority (CA) certificate for the BOSH Director, if the certificate is not verifiable by the local machine's certificate chain.

3. Delete each deployment with the following command:

```
bosh -d DEPLOYMENT-NAME delete-deployment
```

Where:

- `DEPLOYMENT-NAME` is the deployment name retrieved in [Step 10: Identify Your Deployment](#).

- o Log in to your IaaS account and delete the disks manually.

1. To retrieve a list of disk IDs, run the following command:

```
bosh -e BOSH-DIRECTOR-IP \
--ca-cert PATH-TO-BOSH-SERVER-CERTIFICATE instances -i
```

Where:

- `BOSH-DIRECTOR-IP` is the BOSH Director IP retrieved in [Step 8: Retrieve BOSH Director Address](#).
- `PATH-TO-BOSH-SERVER-CERTIFICATE` is the path to the Certificate Authority (CA) certificate for the BOSH Director, if the certificate is not verifiable by the local machine's certificate chain.

Once the disks are deleted, continue with [Step 11: Remove Stale Cloud IDs for All Deployments](#).

Enabling External Blobstore Backups

This topic provides instructions for enabling external blobstore backups in your Pivotal Application Service (PAS) tile.

BOSH Backup and Restore (BBR) supports the following:

- Versioned S3 or S3-Compatible Blobstores
- Unversioned S3 or S3-Compatible Blobstores
- Azure Blobstores

For more information, see [Backup and Restore with External Blobstores](#).

 **Note:** To enable external blobstore backups for PAS, the **Backup Restore Node** must be enabled. See [Enable Backup Restore Node](#) in *Backing Up Pivotal Cloud Foundry with BBR*.

External Blobstore Support

External blobstore backup support varies based on which version of Ops Manager you are running and what type of blobstore you are backing up.

In some cases, external blobstore support is included in the version of Ops Manager you are using. In other cases, installing Blobstore Add-On is required.

Refer to the table below to determine if external blobstore support is included in the version of Ops Manager you are using.

Blobstore Type	Ops Manager Version						
	Versioned	Add-On	Add-On	Add-On	Included	Included	Included
Unversioned	Add-On	Add-On	Add-On	Add-On	Included	Included	Included
Azure	Add-On	Add-On	Add-On	Add-On	Add-On	Add-On	Included

S3 or S3-compatible Blobstores

In PAS v2.2 and later, BBR supports backing up and restoring versioned and unversioned S3 or S3-compatible blobstores by default. For more information about configuring an S3 or S3-compatible blobstore, see [Configuring Load Balancing for PAS](#).

 **Note:**

- Backup artifacts of external, S3-compatible, versioned blobstores do not contain the physical blobs. BBR requires that the original buckets still exist to be restored.
- To protect yourself from losing a bucket, see [Enable Replication on Your External Blobstore](#) in *Backup and Restore with External Blobstores* in the Cloud Foundry documentation.

Azure Blobstores

To configure your Azure blobstore for backups, you must enable Soft Delete in your Azure Storage account. You can also enable Soft Delete for existing Azure storage accounts that were not initially created with Soft Delete enabled. For more information, see [Soft delete for Azure Storage blobs](#) in the Microsoft documentation.

After enabling Soft Delete, you can enable Azure blobstore backups in Ops Manager for new or existing blobstores. For more information, [Configuring Load Balancing for PAS](#).

To save storage space and cost, Pivotal recommends that you configure a retention policy to permanently delete objects after a period of time.

Troubleshooting BBR

This topic lists common troubleshooting scenarios and their solutions when using BOSH Backup and Restore (BBR) to back up and restore Pivotal Cloud Foundry (PCF).

Troubleshooting During a Restore

This section provides solutions for errors that occur during a restore.

Restore Fails with a MySQL Monit Start Timeout

Symptom

While running the BBR restore command, restoring the job `mysql-restore` fails with the following error:

```
1 error occurred:  
* restore script for job mysql-restore failed on mysql/0.  
...  
Monit start failed: Timed out waiting for monit: 2m0s
```

Explanation

This happens when the MySQL job fails to start within the timeout period. It ends up in an **Execution Failed** state and `monit start` never tries to start it again.

Solution

Ensure that your MySQL Server cluster has only one instance. If there is more than one instances of MySQL Server, the restore fails with a `monit start` timeout. Scale down to one instance and retry.

If your MySQL Server cluster is already scaled down to one node, it may have taken longer than normal to restart the cluster. Follow the procedure below to manually verify and retry.

1. To list the VMs in your deployment, run the following command:

```
bosh -e BOSH-DIRECTOR-IP --ca-cert /var/tempest/workspaces/default/root_ca_certificate \  
-d DEPLOYMENT-NAME \  
ssh
```

Where:

- o `BOSH-DIRECTOR-IP` is the BOSH Director IP. For how to retrieve this IP address, see [Retrieve BOSH Director Address](#).
- o `DEPLOYMENT-NAME` is the deployment name. For how to retrieve the deployment name, see [Identify Your Deployment](#).

2. Locate the `mysql` VM.
3. SSH into the `mysql` VM.
4. To check that the MySQL job process is running, perform one of the following from the `mysql` VM:
 - o If you selected **Internal Databases - MySQL - Percona XtraDB Cluster** when you configured the PAS tile, run the following command:

```
ps aux | grep galera-init
```

- o If you selected **Internal Databases - MySQL - MariaDB Galera Cluster** when you configured the PAS tile, run the following command:

```
ps aux | grep mariadb_ctrl
```

For example:

```
$ ps aux | grep galera-init
```

For more information, see the [Configuring PAS](#).

- Run the following command to check that `monit` reports that the MySQL job process is in an **Execution Failed** state:

```
sudo monit summary
```

- If so, run the following command from the `mysql` VM to disable monitoring:

```
monit unmonitor
```

- Run the following command to enable monitoring:

```
monit monitor
```

- After a few minutes, run the following command:

```
monit summary
```

The command should report that all the processes are running.

- Re-attempt the restore with BBR.

Deployment Does Not Match the Structure of the Backup

Symptom

The following error displays:

```
Deployment 'deployment-name' does not match the structure of the provided backup
```

Explanation

The instance groups with the restore scripts in the destination environment don't match the backup metadata. For example, they may have the wrong number of instances of a particular instance group, or the metadata names an instance group that does not exist in the destination environment.

Solution

BBR only supports restoring to an environment that matches your original environment. Pivotal recommends altering the destination environment to match the structure of the backup.

General Troubleshooting

This section provides solutions for general errors.

Connection Error

Symptom

BBR displays an error message containing "SSH Dial Error" or another connection error.

Explanation

The jumpbox and the VMs in the deployment are experiencing connection problems.

Solution

Perform the following steps:

1. To ensure your deployment is healthy, run the following command:

```
bosh vms
```

2. To clean up the data from the failed backup on the instances, run the following command:

```
bbr deployment backup-cleanup
```

 **Note:** This step must be performed, otherwise, further BBR commands fail.

3. Repeat the BBR operation.

Error Running Metadata Script

Symptom

BBR backup or restore fails with a metadata error:

```
1 error occurred:  
error 1:  
An error occurred while running metadata script for job redis-server on redis/0ce9f81f-1756-480b-8e3e-a4609b14b6a6: error from metadata
```

Explanation

There is a problem with your PCF install.

Solution

Contact [Pivotal Support](#) ↗

Using Ops Manager

Ops Manager is a web application that you use to deploy and manage a [Pivotal Cloud Foundry](#) (PCF) PaaS. This is a guide to deploying and using Ops Manager.

Ops Manager UI

Use the Ops Manager UI to install PCF products and configure their settings.

For more information on the Ops Manager UI, see [Using the Ops Manager UI](#).

Using Ops Manager Programmatically and from the Command Line

Use the Ops Manager API, `om` CLI, and Platform Automation (Beta) tasks to manage PCF installations programmatically, from the command line, or from an automation pipeline.

See the [Using Ops Manager Programmatically and from the Command Line](#) topic to learn how to get started using these Ops Manager tools.

Related Links

Using the Ops Manager UI and Installed Products

- [Using the Ops Manager Interface](#)
- [Adding and Deleting Products](#)
- [Retrieving Credentials from Your Deployment](#)
- [Floating Stemcells](#)
- [Configuring BOSH Director on vSphere](#)
- [vSphere Service Account Requirements](#)
- [Creating UAA Clients for BOSH Director](#)
- [Configuring BOSH Director on OpenStack](#)
- [Using Your Own Load Balancer](#)
- [Pivotal Cloud Foundry User Types](#)
- [Creating and Managing Ops Manager User Accounts](#)
- [Creating New PAS User Accounts](#)
- [Logging In to Apps Manager](#)
- [Adding Existing SAML or LDAP Users to a Pivotal Cloud Foundry Deployment](#)
- [Deleting an AWS Installation from the Console](#)
- [Modifying Your Ops Manager Installation and Product Template Files](#)
- [Managing Errands in Ops Manager](#)

Using the Ops Manager API, CLI, and Platform Automation

- [Using Ops Manager Programmatically and from the Command Line](#)
- [Using the Ops Manager API](#)

Backing Up

- [Backing Up and Restoring Pivotal Cloud Foundry](#)

Monitoring, Logging, and Troubleshooting

- [Monitoring Virtual Machines in Pivotal Cloud Foundry](#)
- [Diagnosing Problems in PCF](#)
- [Troubleshooting Problems in PCF](#)
- [Troubleshooting Ops Manager for VMware vSphere](#)
- [Recovering MySQL from PAS Downtime](#)
- [Advanced Troubleshooting with the BOSH CLI](#)

Using the Ops Manager UI

Overview

This topic explains how to start using the Ops Manager UI, and lists other content about the Ops Manager UI.

Ops Manager operators can configure Ops Manager using a dashboard-style interface. The Ops Manager UI lets you make installation-wide modifications to Ops Manager, including updates to authentication methods, security and networking configuration, and more.

Advanced users can configure Ops Manager by accessing the Ops Manager API. For information about configuring Ops Manager with the API, see [Using the Ops Manager API](#).

Access the Ops Manager UI

Ops Manager is accessible at its unique URL, which was assigned during the initial PCF setup process. For more information about setting up PCF, see [Architecture and Installation Overview](#).

Browser Support

The Ops Manager UI is accessible from any internet-enabled host. It is compatible with current and recent versions of all major browsers.

Pivotal recommends using the current version of Chrome, Firefox, or Safari for the best Ops Manager experience.

Related Links

- [Using the Ops Manager Interface](#)
- [Reviewing Pending Product Changes](#)
- [Adding and Deleting Products](#)
- [Importing and Managing Stemcells](#)
- [Logging in to Apps Manager](#)
- [Managing Errands in Ops Manager](#)

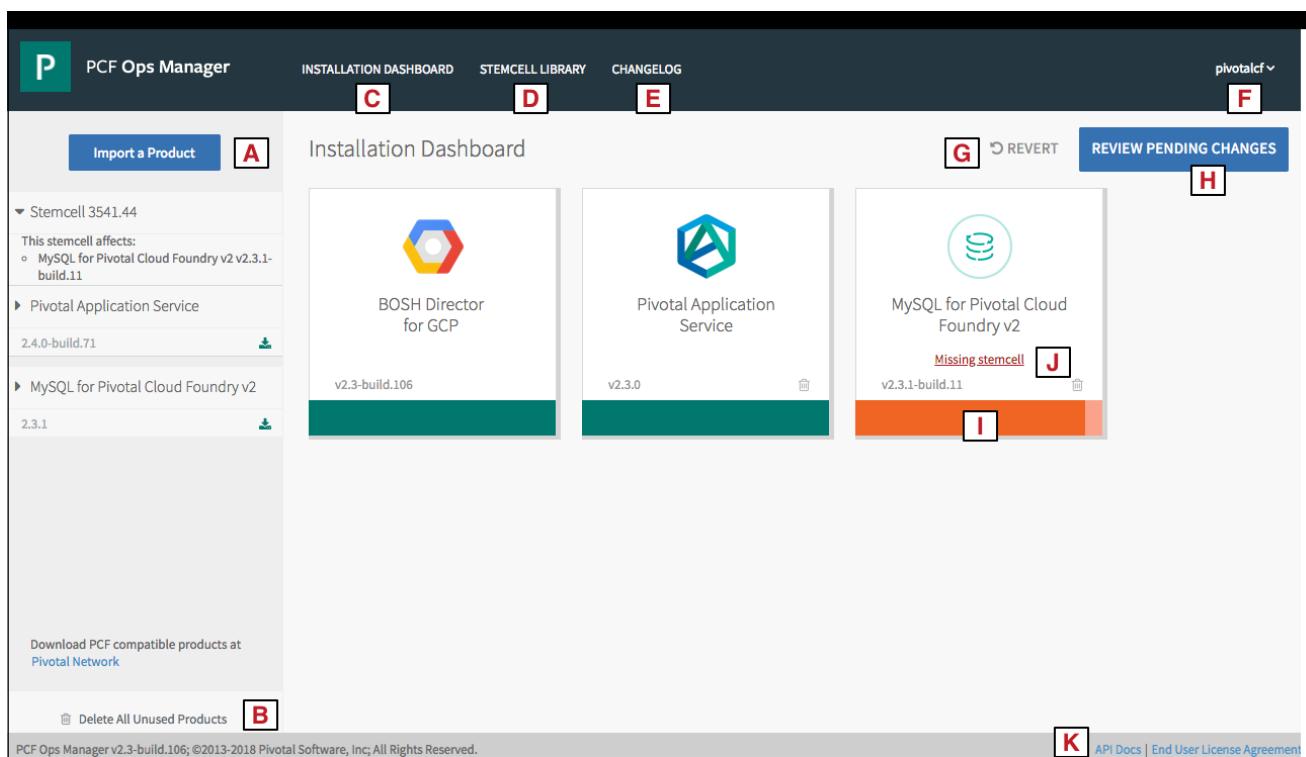
Using the Ops Manager Interface

This topic describes key features of the [Pivotal Cloud Foundry](#) (PCF) Operations Manager interface.

Installation Dashboard Page

The Installation Dashboard is the central Ops Manager interface for managing your PCF deployment. It displays the BOSH Director configuration for your IaaS and product tiles for your PCF deployment.

The following screenshot shows the Installation Dashboard. Each section is labeled with a red letter. Click the image to see it at full size.



The following list describes each labeled section of the Installation Dashboard:

- **A— Import a Product:** Click this button to add a new product to Ops Manager. You can download PCF-compatible product files from [Pivotal Network](#).
- If you have configured a [Pivotal Network API](#) token in the **Settings** section of Ops Manager, a list of the latest versions of already imported products appears automatically.
- **B— Delete All Unused Products:** Click this link to delete any unused products.
- **C— Installation Dashboard:** Click this link to return to the **Installation Dashboard** page from other Ops Manager pages.
- **D— Stemcell Library:** Click this link to open the **Stemcell Library**. In the **Stemcell Library** you can import stemcells, stage stemcells, and review your stemcell version numbers. For more information, see [Importing and Managing Stemcells](#).
- **E— Change Log:** Click this link to view and search a log of your previous installations. See the [Change Log Page](#) section of this topic for details.
- **F— User Account Menu:** Click the arrow next to your username to view the menu. Use this menu to navigate to your **Settings** page, view **My Account** to change your email and password, or log out of the **Installation Dashboard**. See [My Account](#) of this topic for details on account information. See the [Settings Page](#) section of this topic for details on available Ops Manager settings.
- **G— Revert:** Click the revert button to return to the previous successful installation. Any pending changes you have made to the current installation settings will be lost.
- **H— Review Pending Changes:** Click the button to go to the **Review Pending Changes** page, which organizes pending changes by tile. You are able to enable or disable each tile to selectively deploy individual tiles and their dependencies. For more information, see [Reviewing Pending Product Changes](#).
- **I— Orange Bar:** Indicates that additional configuration for the product tile is required before deployment. Click the product tile to complete its configuration.

- J—Missing Stemcell Link:** If an imported product is missing a required stemcell, a **Missing Stemcell** link appears on the tile. Click this link to open the Stemcell Library. For more information, see [Importing and Managing Stemcells](#).
- K—API Docs:** Click this link to go to the Ops Manager API documentation, which details how you can manage Ops Manager through the API rather than with the user interface. For more information about the Ops Manager API, see [Using the Ops Manager API](#).

Change Log Page

Navigate to the **Change Log** page by clicking the corresponding link in the dashboard header. This page shows Ops Manager's deployment history, a record of all of its **Apply Changes** actions.

The table lists the following top-level attributes for each deployment:

Label	Type	Description
STATUS	Icon	Whether the deployment succeeded or failed
DEPLOYMENT ID	Text	A sequentially-numbered identifier for the deployment
USER	Username	The user who initiated the deployment

A **LOGS** button for each deployment opens the deployment's full Installation Log.

Within each deployment listing, a table details for the individual products deployed in or deleted from each installation.

The details listed for each product are as follows:

Label	Type	Description
PRODUCT	Text	The name of the product
ACTION	One of the following labels: ADDED, UPDATED, DELETED, NO CHANGES, FAILED	The action, if any, that the deployment took with the product
STARTED	UTC timestamp	When Ops Manager began deploying the product
FINISHED	UTC timestamp	When Ops Manager stopped deploying or attempting to deploy the product
DURATION	Time, in minutes	How long Ops Manager took to deploy or attempt to deploy the product

A **Total** row, below the individual product rows, lists all-product totals for the **STARTED**, **FINISHED**, and **DURATION** fields.

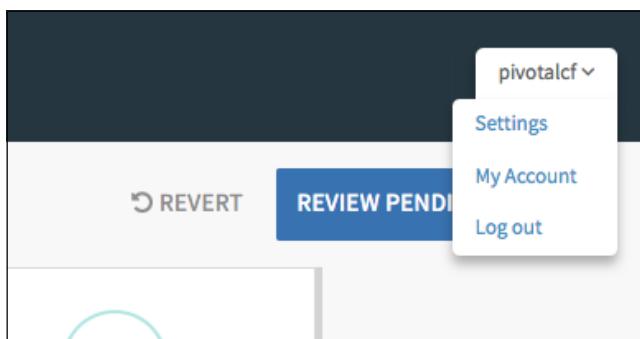
Using the Change Log Page

Configure the **Change Log** page by modifying the following fields:

- Installation Dashboard**—Click **Installation Dashboard** to return to Ops Manager's Installation Dashboard. Alternatively, click the **Back** button in your web browser.
- Show X entries**—Click the number displayed in the **Show X entries** dropdown to choose between 10, 25, 50, and 100 entries.
- Search**—Type in the search box to sort the Change Log page by text or integer matches. As you type, matching entries appear on the screen.
- Previous / Next**—Click **Previous**, **Next**, or the number between them to load older or newer entries.

Settings Page

Navigate to the **Settings** page by clicking on your user name located at the upper right corner of the screen and selecting **Settings**.



The Settings configuration screen displays.

The screenshot shows the 'Settings' configuration screen. On the left, a sidebar lists various settings sections: Change Decryption Passphrase, SAML Settings, LDAP Settings, SSL Certificate, Pivotal Network Settings, Proxy Settings, Custom Banner, Export Installation Settings, Syslog, and Advanced Options. The 'Change Decryption Passphrase' section is currently active, indicated by a grey background. It contains fields for 'Current Decryption Passphrase*', 'New Decryption Passphrase*', and 'Confirm New Decryption Passphrase*'. A large blue button labeled 'CHANGE DECRYPTION PASSPHRASE' is at the bottom. The top of the page has a header with a 'REVERT' button, a 'REVIEW PENDING' button, and a user dropdown menu labeled 'pivotalcf'.

The following section describes each configuration pane:

Note: Modifying these settings does not require you to return to the Installation Dashboard and click **Apply Changes**. These settings apply to the Ops Manager VM. The BOSH Director does not apply them to your PCF deployment.

Change Decryption Passphrase

Use this pane to reset your **decryption passphrase**.

SAML Settings

To change your Identity Provider (IdP) to SAML, configure the following fields:

- **Current Decryption Passphrase:** Enter your decryption passphrase.
- **SAML IDP Metadata:** Enter the full URL or XML SAML IdP metadata.
- **BOSH IDP Metadata:** (Optional) Enter the full URL or XML BOSH IdP metadata. If left blank, the default is the same metadata as the field above.
- **SAML Admin Group:** Enter the name of the SAML group that contains all of the Ops Manager administrators. This field is case-sensitive.
- **Groups Attribute:** Enter the groups attribute tag name with which you configured the SAML server. This field is case-sensitive.
- **Provision an admin client in the Bosh UAA:** Enable if you want to provision an admin client in the BOSH UAA. For more information, see [Provision Admin Client](#) in the *Creating UAA Clients for BOSH Director* topic.

LDAP Settings

Use this pane to change your IdP to LDAP.

For more information about changing your IdP to SAML or LDAP, view the following instructions for your IaaS configuration:

- Amazon Web Services. See [Configuring BOSH Director on AWS](#).
- Google Cloud Platform. See [Configuring BOSH Director on GCP Manually](#).
- Microsoft Azure. See [Configuring BOSH Director on Azure Manually](#).
- OpenStack. See [Configuring BOSH Director on OpenStack](#).
- vSphere. See [Configuring BOSH Director on vSphere](#).

The screenshot shows the 'Settings' page with the 'SSL Certificate' tab selected. On the left, a sidebar lists other settings: Change Decryption Passphrase, SAML Settings, LDAP Settings, SSL Certificate (which is highlighted in grey), Pivotal Network Settings, Proxy Settings, Custom Banner, Export Installation Settings, Syslog, and Advanced Options. The main content area is titled 'SSL Certificate' and contains a sub-section 'Revert to self-signed certificate'. Below this is a note: 'This page lets you configure Ops Manager to use a custom SSL certificate. This certificate serves all traffic from the UI and API. Allow up to 1 minute for the new certificate to take effect.' There are two large input fields: one for 'Certificate*' and one for 'Private Key*'. Both fields have their content masked by asterisks. A 'Change' link is located below the private key field. At the bottom right of the main content area is a blue 'ADD CERTIFICATE' button.

SSL Certificate

Configure Ops Manager to use a custom SSL certificate for all Ops Manager traffic both through the UI and API.

If you leave the fields blank, Ops Manager uses an auto-generated self-signed certificate rather than your own custom certificate and private key. If you wish to replace a custom certificate with a self-signed certificate, select **Revert to self-signed certificate**.

Settings

- Change Decryption Passphrase
- SAML Settings
- LDAP Settings
- SSL Certificate
- Pivotal Network Settings**
- Proxy Settings
- Custom Banner
- Export Installation Settings
- Syslog
- Advanced Options

Pivotal Network Settings

Configure your [Pivotal Network API](#) token to allow checks for updates.

Set API Token

ADD TOKEN

Pivotal Network Settings

Enter your [Pivotal Network API](#) token and click **Add Token** to connect your **Installation Dashboard** to Pivotal Network.

Proxy Settings

If you are using a proxy to connect to Ops Manager, update your **Proxy Settings** by providing a **HTTP proxy**, **HTTPS proxy**, or **No proxy**.

The screenshot shows the 'Settings' page in the Pivotal Ops Manager interface. On the left, a sidebar lists various configuration options: Decryption Passphrase, Change Authentication Method to SAML, Change Authentication Method to LDAP, SSL Certificate, External API Access, Proxy Settings, Custom Banner (which is selected and highlighted in grey), Export Installation Settings, Syslog, and Advanced. The main content area is titled 'Update Custom Banner'. It contains two text input fields: 'UI Banner' and 'SSH Banner', each with a placeholder text area. A blue 'Save' button is located at the bottom right of the form.

Custom Banner

Create a custom text banner to communicate important messages to operators. For **UI Banner**, enter the text you want to be shown on each page of the Ops Manager UI. For **SSH Banner**, enter the text that appears when an operator shells into Operations Manager.

Export Installation Settings

Exports the current installation with all of its assets. When you export an installation, the exported file contains references to the base VM images, necessary packages, and configuration settings.

Syslog

Viewable by administrators only. Configure a custom Syslog server for Ops Manager. When you select **Yes** and fill the following fields, Ops Manager produces and sends all Ops Manager logs to the configured syslog endpoint.

1. Select **Syslog**.

Syslog

Do you want to configure Syslog for Bosh Director?

- No
 Yes

Address*

The address or host for the syslog server

Port*

Transport Protocol*

TCP



Enable TLS

Permitted Peer*

SSL Certificate*

Queue Size

100000

Save

2. (Optional) Select Yes to send BOSH Director system logs to a remote server.

3. Enter the IP address or DNS name for the remote server in **Address**.
4. Enter the port number that the remote server listens on in **Port**.
5. Select **TCP** or **UDP** from the **Transport Protocol** dropdown. This selection determines which transport protocol is used to send the logs to the remote server.
6. (Optional) Select the **Enable TLS** checkbox to send encrypted logs to remote server with TLS. After you select the checkbox, do the following:
 - a. Enter either the name or SHA1 fingerprint of the remote peer in **Permitted Peer**.
 - b. Enter the SSL certificate for the remote server in **SSL Certificate**.



Note: Pivotal strongly recommends that you enable TLS encryption when you are forwarding logs. Logs can contain sensitive information, such as cloud provider credentials.

7. (Optional) Enter an integer in **Queue Size**. This value specifies the number of log messages held in the buffer. The default value is 100,000.
8. (Optional) Select the checkbox to **Forward Debug Logs** to an external source. This option is deselected by default. If you select it, you may generate a large amount of log data.
9. (Optional) Enter configuration details for rsyslog in the **Custom rsyslog Configuration** field. This field requires the rainerscript syntax.
10. Click **Save Syslog Settings**.

Advanced Options

- **Download Activity Data:** Downloads a directory containing the config file for the installation, the deployment history, and version information.
- **Download Root CA Cert:** Use this to download the root CA certificate of your deployment as an alternative to curling the Ops Manager API.
- **View Diagnostic Report:** Displays various types of information about the configuration of your deployment.
- **Delete This Installation:** Deletes your Ops Manager installation permanently.

My Account Page

To change your email and password, navigate to the **My Account** page by clicking on your user name located at the upper right corner of the screen and selecting **My Account**.

Account Settings

Profile

admin@test.org

Change Email
Change Password

Third Party Access

You have not yet authorized any third party applications.

Reviewing Pending Product Changes

This topic describes the **Review Pending Changes** page and how to selectively deploy individual or multiple tiles in Ops Manager. Deploying tiles selectively can drastically reduce deployment time.

Review Pending Changes Page

The following screenshot shows the **Review Pending Changes** page. Features called out with red letters are described below.

Color	Description
Green	Select All Products
Orange	Select Product
Red	Product Listing

Each labeled section of the **Review Pending Changes** page works as follows:

- A—Select All Products:** When enabled, the checkbox selects all available products. When disabled, the checkbox deselects all available products.
- B—Select Product:** The checkboxes next to each product listing allow you to select the tiles you want to deploy. The BOSH Director is always selected.
- C—Product Listing:** Each product listing displays the version of the product staged for deployment. The top of each listing shows one of the following three colors:
 - Green**—Tile is configured and ready for deployment.
 - Orange**—Tile is not fully configured and cannot be deployed yet.
 - Red**—Tile is pending deletion.
- D—Depends on:** This section lists tile dependencies and versions. If you choose to deploy a tile that has dependencies, you must also deploy all of that tile's dependencies in the same deployment.

- **E— Warnings:**—Red text in orange product listings indicates changes you need to make before you deploy the tile. You might see some of the following warnings:
 - **Missing stemcell(s)**—You must import and stage one or more stemcells with Stemcell Library. The warning links to the Stemcell Library.
 - **Stemcell(s) out of date**—You must update one or more stemcells for your tile in Stemcell Library. For more information, see [Importing and Managing Stemcells](#).
 - **Configuration is incomplete**—You must complete the configuration of your tile.
 - **Configuration is invalid**—You must resolve your tile’s invalid configurations.
- **F— Changes:** This section lists high-level pending changes to the product that Ops Manager has not yet attempted to deploy. When Ops Manager tries to deploy a product but the deployment fails, it clears this list and the pending changes are retained for the next deployment. To see if your product deployed successfully, see [Change Log Page](#).

Listed changes include:

- Tile deletion is staged.
 - Tile version is updated.
 - Stemcell is added.
 - Stemcell is updated.
- **G— Errands:** If you expand this section, you can enable or disable errands that trigger when that tile deploys. For example, you may enable an errand to log properties. For more information, see [Managing Errands in Ops Manager](#).
 - **H— Staged:** This section indicates the status of the tile.
 - **I— See Changes:** Clicking this button takes you to the [Pending Changes](#) page for that tile.
 - **J— Apply Changes:** Clicking this button deploys all enabled tiles.

About Pending Changes for a Specific Tile

If you have changed the configuration of a tile in your deployment, use the [Pending Changes](#) page to review the changes to the existing manifest or configs line-by-line.

This page summarizes staged changes to the existing tile’s configuration. The changes displayed in this page have not yet been applied to the deployment.

This page allows operators to verify changes to the manifest and configs so they can approve any staged changes before applying changes to their deployment.

Use the Pending Changes Page

The change summary fields display features that have been enabled or disabled using the Ops Manager UI or API. These changes are staged, but have not been implemented yet. You must click **Apply Changes** to put these changes into effect.

If you want to see changes for a specific tile, do the following:

1. On the Review Pending Changes page, click **See Changes** on the tile you wish to review. A new tab opens.
2. Review the changes displayed. Settings that have been **disabled** are displayed in red with a minus sign (-). Settings that have been **enabled** are displayed in green with a plus sign (+).
3. Close the tab and return to the Review Pending Changes page to apply the changes or review other tiles.

The screenshot below shows changes to the BOSH Director’s manifest.

PCF Ops Manager pivotalcf ▾

Pending Changes for BOSH Director

Manifest

Runtime Configs

- ops_manager_dns_runtime
- director_runtime

CPI Configs

- default

Cloud Config

Manifest

```

instance_groups:
- name: bosh
  properties:
    director:
+   enable_post_deploy: "<redacted>"
-   enable_post_deploy: "<redacted>"
  agent:
    env:
      bosh:
+     agent:
+       settings:
+         tmpfs: "<redacted>"
+       job_dir:
+         tmpfs: "<redacted>"
```

Runtime Configs

ops_manager_dns_runtime

No changes

director_runtime

No changes

CPI Configs

default

No changes

Cloud Config

No changes

When you make changes to a tile, change summaries display for the following tile areas:

- Manifest
- Runtime configs
- CPI configs
- Cloud configs

If an aspect of the tile has not changed, the change summary displays `No changes`.

 **Note:** Changes to errands do not display in the page UI.

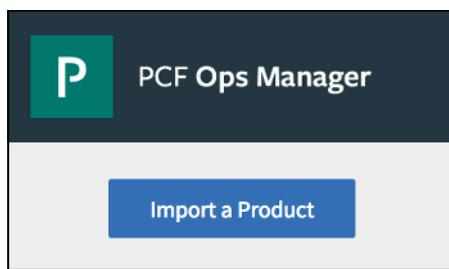
Adding and Deleting Products

Refer to this topic for help adding and deleting additional products from your [Pivotal Cloud Foundry](#) (PCF) installation, such as [Pivotal RabbitMQ® for PCF](#).

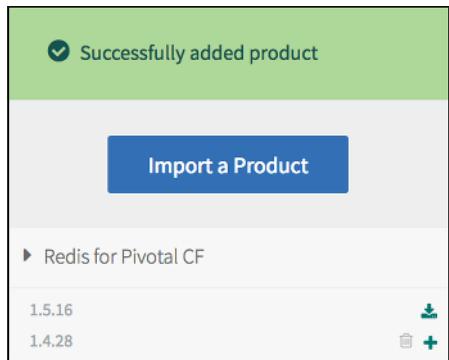
 **Note:** In Ops Manager 2.6, all product tiles use floating stemcells by default. This increases the security of your deployment by enabling tiles to automatically use the latest patched version of a stemcell, but it may significantly increase the amount of time required by a tile upgrade. Review the [Floating Stemcells](#) topic for more information.

Adding and Importing Products

1. Download PCF-compatible products at [Pivotal Network](#). If you cannot download products from Pivotal Network due to restricted network connectivity, see [Installing PCF in Airgapped Environments](#).
2. Navigate to your Ops Manager **Installation Dashboard** and log in.
3. Click **Import a Product**.



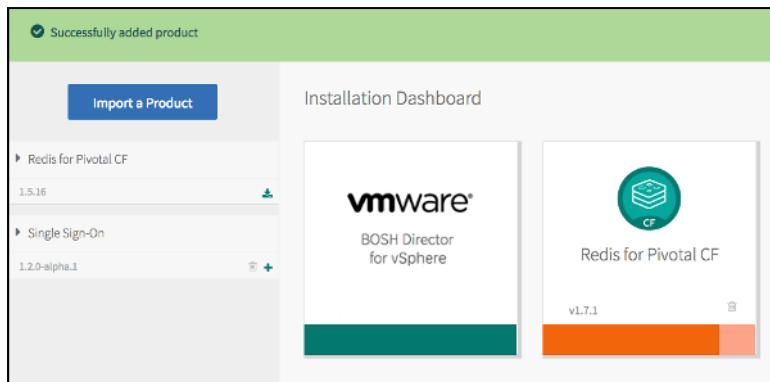
4. Select the .pivotal file that you downloaded from Pivotal Network or received from your software distributor, then click **Open**. If the product is successfully added, it appears in your product list. If the product that you selected is not the latest version, the most up-to-date version will appear on your product list.



5. Add the product tile to the **Installation Dashboard** by clicking the green plus sign icon.



6. The product tile appears in the Installation Dashboard. If the product requires configuration, the tile appears orange. If necessary, configure the product.



- (Optional) In the product configuration view, select the **Errands** pane to configure post-install errands or review the default settings. Post-install errands are scripts that automatically run after a product installs, before Ops Manager makes the product available for use. For more information about post-install errands, see the [Errands](#) topic.

Note: By default, Ops Manager reruns errands even if they are not necessary due to settings left from a previous install. Leaving errands checked at all times can cause updates and other processes to take longer. To prevent an errand from running, deselect the checkbox for the errand in the **Settings** tab on the product tile before installing the product.

The screenshot shows the Redis for Pivotal CF product configuration view. At the top, there are tabs for "Settings", "Status", "Credentials", and "Logs". The "Settings" tab is selected. On the left, a sidebar lists several configuration items with checkboxes: "Assign AZs and Networks" (checked), "Shared-VM Plan" (checked), "Syslog" (checked), "Errands" (checked, highlighted in grey), "Resource Config" (checked), and "Stemcell" (unchecked). On the right, the "Errands" section is expanded, showing two categories: "Post-Deploy Errands" and "Pre-Delete Errands". Under "Post-Deploy Errands", the "Broker Registrar" checkbox is checked. Under "Pre-Delete Errands", the "Broker Deregistrar" checkbox is checked. A large blue "Save" button is located at the bottom of the section.

The **Broker Registrar** checkbox is an example of an errand available for a product. When you select this checkbox, this errand registers service brokers with the Cloud Controller and also updates any broker URL and credential values that have changed since the previous registration.

- Click **Review Pending Changes**, then **Apply Changes** to start installation and run post-install lifecycle errands for the product.

Using Pivotal Network API to Upgrade Products

Ops Manager provides a way to upgrade products by connecting your **Installation Dashboard** with Pivotal Network using a API token. Once you have uploaded a [product](#), all subsequent product upgrades appear automatically in your **Installation Dashboard**.

Note: Using the Pivotal Network API is only available if you have access to the Internet since communication between Ops Manager and the Pivotal Network is necessary to import your products. If you are on an isolated network, do not save your API token.

1. Navigate to [Pivotal Network](#) and log in.
2. Click your user name, located in the upper top right side of the page.
3. Select **Edit Profile**.



4. In the **Edit Profile** tab, copy your **API Token**.
5. Navigate to your Ops Manager **Installation Dashboard** and log in.
6. Click your user name, located in the upper top right side of the page.
7. Select **Settings**.
8. In the **External API Access** tab, paste your **API Token**.

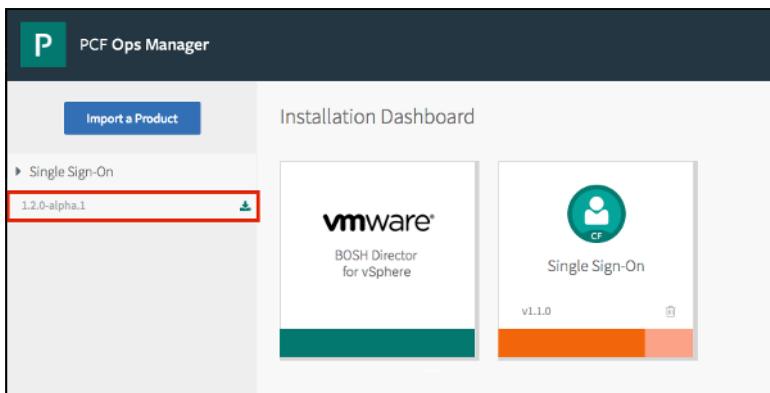
The screenshot shows the 'Settings' page of the 'Installation Dashboard'. On the left, there's a sidebar with links: 'Decryption Passphrase', 'Authentication Method', 'External API Access' (which is highlighted in blue), 'Proxy Settings', 'Export Installation Settings', and 'Advanced'. The main content area has a title 'Pivotal Network Settings' and a sub-instruction 'Configure your [Pivotal Network](#) API token to enable update checking.' Below this is a 'Set API token' section with an input field containing 'JWT [REDACTED]' and a blue 'Save' button.

9. Click **Save**.

Update Existing Products

Once you save the Pivotal Network **API Token** to the Ops Manager **Installation Dashboard**, the latest versions of your existing products will appear in your **Installation Dashboard**. Upgrade your product to the latest version by following these instructions.

1. Locate and download the product version you want to upgrade to by clicking on the green download icon.



2. When the download is complete, refresh the page to use the product.
3. If necessary, configure the product.
4. Click **Review Pending Changes**, then **Apply Changes**.

Applying Changes to BOSH Director

You can use the Ops Manager API or the Review Pending Changes page to apply pending changes only to BOSH Director when you stage multiple products in a new installation or as part of an upgrade.

Deleting a Product

1. From the Installation Dashboard, click the trash icon on a product tile to remove that product. In the **Delete Product** dialog box that appears, click **Confirm**.

 **Note:** You cannot delete the BOSH Director product.

2. Click **Review Pending Changes**, then **Apply Changes**.

After you delete a product, the product tile is removed from the installation and the Installation Dashboard. However, the product appears in the Available Products view.

Importing and Managing Stemcells

This topic explains how to use the Stemcell Library to import and stage stemcells to products.

The screenshot shows the 'Stemcell Library' page. At the top right are two buttons: 'IMPORT STEMCELL' and 'SAVE'. Below is a table with four columns: 'Product', 'Required', 'Deployed', and 'Staged'. A single row is present for 'Ops Manager: Example Product'. The 'Required' column shows 'ubuntu-trusty 3541.4'. The 'Deployed' column shows '3541.4'. The 'Staged' column contains a dropdown menu set to '3541.4' and a checked checkbox labeled 'Latest stemcell.'

For more conceptual information about floating stemcells and stemcell upgrades, see [Floating Stemcells](#).

⚠ warning: Some product releases require Xenial stemcells, as listed in [Tiles Using Xenial Stemcells in PCF](#). If you are importing and using Xenial stemcells for the first time, see [Updates for Xenial Stemcell Support](#).

Import and Stage a Stemcell

The screenshot shows the 'Stemcell Library' page. The 'IMPORT STEMCELL' button at the top right is circled in red. Below is a table with columns 'Product', 'Required', 'Deployed', and 'Staged'. A single row is present for 'Ops Manager: Example Product'. The 'Required' column shows 'ubuntu-trusty 3541.4'. The 'Deployed' column shows '3541.4'. The 'Staged' column contains a dropdown menu set to '3541.4' and a checked checkbox labeled 'Latest stemcell.'

Download the appropriate `.tgz` stemcell file from [Pivotal Network](#), then click **Import Stemcell** to permanently import a stemcell into Ops Manager.

The screenshot shows the 'IMPORT STEMCELL' dialog box. It lists the products to be staged: 'light-bosh-stemcell-3541.5-google-kvm-ubuntu-trusty'. Two checkboxes are checked: 'Product' and 'Ops Manager: Example Product v2.1-committed.20180209.125552-stemcell.3541.4'. At the bottom are 'DISMISS' and 'APPLY STEMCELL TO PRODUCTS' buttons.

To stage your stemcell to compatible products, enable the products in the **Import Stemcell** dialog. Click **Apply Stemcell to Products** to save your selections. Click **Dismiss** to dismiss the **Import Stemcell** dialog.

Choose a Stemcell Version

If you have uploaded multiple versions of a stemcell, you can use the dropdown in the **Staged** column to choose which version to use.

You can choose a different version until you deploy. After deployment, older stemcell versions are no longer available.

	Pivotal Cloud Foundry Isolation Segments v1.12.8	3345.17	3345.17	3345.17
	Pivotal Cloud Foundry Metrics v1.4.2	3363.*	3363.0	3345.17

Available Stemcells

- 3345.30
- 3345.22
- 3345.17

If the stemcell displays with a green checkmark and the words **Latest stemcell** below the **Staged** dropdown, the stemcell is the latest available version on your host. An outdated stemcell displays **Stemcell out-of-date**.

	Spring Cloud Services for PCF v1.4.7	3445.*	3445.0	3445.5
--	--------------------------------------	--------	--------	--------

Latest stemcell.

Logging In to Apps Manager

Log In as Admin User

Complete the following steps to log in to Apps Manager as the Admin user:

1. If you do not know the system domain for the deployment, then select Pivotal Application Service (PAS) **Settings > Domains** to locate the configured system domain.
2. Open a browser and navigate to `apps.YOUR-SYSTEM-DOMAIN`. For example, if the system domain is `system.example.com`, then point your browser to `apps.system.example.com`.
3. Log in using UAA credentials for the Admin user. To obtain these credentials, refer to PAS **Credentials > UAA > Admin Credentials**.

UAA	VM Credentials	Link to Credential
	Admin Credentials	Link to Credential

Managing Errands in Ops Manager

This topic describes product errands and how to configure them in Pivotal Cloud Foundry (PCF) Ops Manager.

Errands are scripts that can run at the beginning and at the end of an installed product's availability time. You can use Ops Manager to adjust whether and when these errands run.

Product tiles include two types of errands:

- **Post-deploy errands** run after the product installs but before Ops Manager makes the product available for use. One example is an errand that publishes a newly-installed service to the Services Marketplace.
- **Pre-delete errands** run after an operator chooses to delete the product but before Ops Manager actually deletes it. One example is a clean-up task that removes all data objects used by the errand.

When you click **Review Pending Changes**, then **Apply Changes** in Ops Manager, BOSH either creates a VM for each errand that runs or co-locates errands on existing VMs. Tile developers determine where BOSH deploys the errands for their product.

Pivotal Application Service (PAS) provides several post-deploy errands including smoke tests, Apps Manager, notification, Pivotal Account, and autoscaling errands. For more information about PAS errands, see the [Configuring PAS](#) topic.

For information about the errands associated with any other PCF product, see the documentation provided with the product tile.

Errand Run Rules

Operators can configure two different run rules for errands: **On** and **Off**. These rules control when Ops Manager executes the errand.

When the errand is configured to be **On**, then it always runs, even if there are no changes to the product manifest. When the errand is configured to be **Off**, it never runs.

Ops Manager Defaults and Tile Defaults

By default, Ops Manager applies the **On** rule to all errands.

For any errand, the tile developer can override these Ops Manager defaults with their own tile-specific defaults defined in the tile [property blueprints](#).

Configure Run Rules in Ops Manager

You can configure the run rules for errands in two places in Ops Manager. The [Errands](#) pane saves your configuration and applies the configuration to future installations. The [Pending Changes](#) view applies the rules only to the next time you run an Ops Manager install, without saving them.

Errands Pane: Persistent Rules

Product tiles for PAS and other PCF products have an **Errands** pane that configures the run rules for the product's errands and saves the settings for later.

The **Errands** pane lists all errands for the product and lets you select **On** or **Off**. The [default](#) option differs depending on the errand, and reflects the default setting used by Ops Manager for the errand or any tile-specific default that overrides it.

Errands

Errands are scripts that run at designated points during an installation.

Post-Deploy Errands

Smoke Test Errand	Runs Smoke Tests against your Application Service installation
Default (On)	▼
Usage Service Errand	Pushes the Pivotal Usage Service application to your Application Service installation. Pivotal Apps Manager depends on this application.
Default (On)	▼
Apps Manager Errand	Pushes the Pivotal Apps Manager application to your Application Service installation
Default (On)	▼
Notifications Errand	Pushes the Pivotal Notifications application to your Application Service installation
Default (On)	▼
Notifications UI Errand	Pushes the Notifications UI component to your Application Service installation
Default (On)	▼
App Autoscaler Errand	Pushes the Pivotal App Autoscaler application to your Application Service installation
Default (On)	▼
App Autoscaler Smoke Test Errand	Runs Smoke Tests against the App Autoscaling Service.
Default (On)	▼
NFS Broker Errand	Pushes the NFS Broker application to your Application Service installation
Default (On)	▼
Delete Pivotal Account Application	Deletes the Pivotal Account Application from your Application Service installation
Default (On)	▼

There are no pre-delete errands for this product.

Save

To configure the run rules for a tile, do the following:

1. Navigate to Ops Manager and click the tile to open it.
2. Under the **Settings** tab, open the **Errands** pane.
3. Use the dropdowns to configure the run rule choice for each errand: **On** or **Off**.
4. Click **Save** to save the configuration values and return to the Installation Dashboard.
5. Click **Review Pending Changes**, then **Apply Changes** to redeploy the tile with the new settings.

Pending Changes: One-Time Rules

Ops Manager lets you quickly configure one-time errand run rules for any product queued up for installation:

1. Navigate to Ops Manager. The **Pending Changes** section at top right shows products that Ops Manager has yet to install or update.

2. Under **Pending Changes**, click the product you wish to configure. A list of errands associated with the product appears.

The screenshot shows the 'Pending Changes' interface for the 'INSTALL Pivotal Elastic Runtime' product. The interface includes a 'Revert' button at the top right. Below the product name, there's a list of errands with dropdown menus for configuration. The errands listed are:

- Smoke Test Errand
- Usage Service Errand
- Apps Manager Errand
- Notifications Errand
- Notifications UI Errand
- Pivotal Account Errand
- Autoscaling Errand
- Autoscaling Registration Errand
- NFS Broker Errand

Each errand has a dropdown menu with 'Default' selected. At the bottom of the interface are buttons for 'Apply changes' and 'Changelog'.

3. Use the dropdowns to configure the [run rule](#) choice for each errand: **On** or **Off**. Ops Manager applies these settings once you click **Review Pending Changes**, then **Apply Changes** to install the product, but does not save the settings for future installations.
4. Click **Review Pending Changes**, then **Apply Changes** to redeploy the tile.

Related Links

If you are a product developer and want to learn more about adding errands to your product tile for PCF, see the [Errands](#) topic in the *PCF Tile Developer Guide*.

Using Ops Manager Programmatically and from the Command Line

This topic introduces three tools for running Ops Manager operations programmatically and from the command line, rather than from the Ops Manager GUI: The Ops Manager API, the `om` command-line interface (CLI), and Platform Automation (Beta) tasks.

Overview

Pivotal Cloud Foundry (PCF) includes three related tools that let you run Ops Manager operations programmatically and from the command line, rather than from the Ops Manager GUI:

- The Ops Manager API. For more information, see the [Ops Manager API](#) documentation.
- The `om` command-line interface (CLI). For more information about `om`, see the [om repository](#).
- Platform Automation (Beta) tasks. For more information, see the [Platform Automation](#) documentation.

These three tools control the Ops Manager VM at different levels of abstraction. Platform operators and tile developers use them as described below.

For more information about tile development, see the [PCF Tile Developer Guide](#).

Ops Manager API

The Ops Manager API offers the most granular control of the Ops Manager VM. The API enables users to do the following:

- Automate component-level testing of PCF product tiles
- Manage Ops Manager user accounts
- Develop and debug product tiles
- Access product property values that are not exposed in a product tile GUI
- Otherwise control the Ops Manager VM directly, programmatically or manually, bypassing the Ops Manager UI. For more information about the Ops Manager interface, see [Using the Ops Manager interface](#)

For information about getting started with the Ops Manager API, see [Access the Ops Manager API](#) in the *Using the Ops Manager API* topic.

Your Ops Manager serves a local copy of the complete Ops Manager API documentation at <https://YOUR-OPS-MANAGER-FQDN/docs>.

Ops Manager CLI (`om`)

The Ops Manager CLI, `om`, calls the Ops Manager API to do the following:

- Perform Ops Manager operations from the command line or scripts
- Manage credentials from the command line or scripts
- Generate configuration files from, and apply saved configurations to, staged product tiles.

For information, including how to install `om` and a usage example, see the [om repository](#).

Platform Automation (Beta)

Platform Automation tasks call the `om` CLI to do the following:

- Serve as building blocks for Continuous Integration (CI) pipelines that automate installation and updating of PCF foundations
- Within pipelines, store and retrieve product configuration values from secrets vaults like CredHub and code repositories like GitHub.

For a complete list of Platform Automation tasks, see the [Concourse Tasks](#) topic in the Platform Automation documentation.

Examples of Equivalent Operations

The following table lists some common operations using the Ops Manager UI, Ops Manager API, `om` CLI, and Platform Automation tasks, as examples of how different Ops Manager interfaces relate to one another.

For complete lists of operations, consult the references in above.

UI Action	API Command	<code>om</code> Command	Platform Automation (Beta) Task
Manually configure the BOSH Director tile	Multiple commands; see Configure the BOSH Director in the <i>Using the Ops Manager API</i> topic	<code>om configure-director</code>	<code>configure-director</code>
Click IMPORT A PRODUCT	<code>POST</code> product file to <code>/available_products</code>	<code>om upload-product</code>	<code>upload-product</code>
Click the + symbol next to a product listing left of Installation Dashboard	<code>POST</code> product name and version to <code>/staged/products</code>	<code>om stage-product</code>	<code>stage-product</code>
Manually configure a product tile	Multiple commands; see Configure a Product in the <i>Using the Ops Manager API</i> topic	<code>om configure-product</code>	<code>configure-product</code>
Click STEMCELL LIBRARY > IMPORT STEMCELL	<code>POST</code> stemcell file to <code>/stemcells</code>	<code>om upload-stemcell</code>	<code>upload-stemcell</code>
Click Apply Changes	<code>POST</code> specified products (or <code>all</code>) and errand lists to <code>/installations</code>	<code>om apply-changes</code>	<code>apply-changes</code>

Using the Ops Manager API

This topic explains how to access the Ops Manager API, and gives examples of how to use the API to perform some common Ops Manager operations.

Overview

The Ops Manager API controls the Ops Manager VM directly, bypassing the Ops Manager UI.

Platform operators use the Ops Manager API to automate deployments, retrieve and manage credentials, and otherwise work with Ops Manager. Tile Developers use the Ops Manager API to test and debug Pivotal Cloud Foundry (PCF) product tiles.

For more information about the Ops Manager API, see the [Ops Manager API](#) documentation. Your Ops Manager serves a local copy of this documentation at <https://YOUR-OPS-MANAGER-FQDN/docs>.

Related Tools

For running Ops Manager operations from the command line or within shell scripts, the Ops Manager command-line interface (CLI) `om` is usually a better tool than the Ops Manager API.

For more information about `om`, see the [om repository](#).

For an overview of the Ops Manager API and the tools based on it, see [Using Ops Manager Programmatically and from the Command Line](#).

Access the Ops Manager API

To access the Ops Manager API, you must authenticate to the Ops Manager User Account and Authentication (UAA) server and log in as described below.

For more information about UAA, see [User Account and Authentication \(UAA\) Server](#).

Install the UAAC

If you haven't already, install the UAA Command Line Interface (UAAC) by running the following command from a terminal window:

```
gem install cf-uaac
```

Retrieve an Authorization Token

Every call to the Ops Manager API must include an authorization token that is acceptable to the Ops Manager UAA.

How you retrieve this token depends on whether your Ops Manager user store is an internal IaaS component or an external server. With an internal UAA, the procedure depends on which IaaS you are on.

To retrieve your authorization token, perform the procedure below that corresponds to your Ops Man UAA location and IaaS.

Internal UAA on vSphere

You need the credentials used to import the PCF .ova or .ovf file into your virtualization system.

1. SSH into the Ops Manager VM:

```
ssh ubuntu@OPS-MANAGER-FQDN
```

Where `OPS-MANAGER-FQDN` is the fully qualified domain name of Ops Manager.

For example:

```
$ ssh ubuntu@my-opsmanager.example.com  
Password: *****
```

- When prompted, enter the password that you set during the .ova deployment into vCenter. For example:

```
$ ssh ubuntu@my-opsmanager.example.com  
Password: *****
```

- Proceed to [Log in to Ops Manager](#).

Internal UAA on AWS, Azure, or OpenStack

- Locate the Ops Manager FQDN on the AWS **EC2 instances** page, Azure **Virtual machines** page, or the OpenStack **Access & Security** page.

- Run `chmod 600` to change the permissions on the `.pem` file to be more restrictive:

```
chmod 600 OPS-MGR-PEM
```

Where `OPS-MGR-PEM` is the public key from the keypair used when creating the Ops Manager VM.

For example:

```
$ chmod 600 ops_mgr.pem
```

- Run `ssh -i OPS-MGR-PEM ubuntu@OPS-MANAGER-FQDN` to SSH into the Ops Manager VM:

```
ssh -i ops_mgr.pem ubuntu@OPS-MANAGER-FQDN
```

Where:

- `OPS-MGR-PEM` is the public key from the keypair used when creating the Ops Manager VM.
- `OPS-MANAGER-FQDN` is the fully qualified domain name of your Ops Manager.

For example:

```
$ ssh -i ops_mgr.pem ubuntu@my-opsmanager.example.com
```

- Proceed to [Log in to Ops Manager](#).

Internal UAA on GCP

- Confirm that you have installed the gcloud CLI. If you do not have the gcloud CLI, see [the Google Cloud Platform documentation](#).

- To configure your Google Cloud Platform project, run the following command:

```
$ gcloud config set project MY-PROJECT
```

Where `PROJECT` is the name of your GCP project.

For example:

```
$ gcloud config set project gcp
```

- Run the following command:

```
gcloud auth login MY-GCP-ACCOUNT
```

Where `MY-GCP-ACCOUNT` is your GCP account ID.

For example:

```
$ gcloud auth login user@example.com
```

- Run the following command:

```
gcloud compute ssh MY-INSTANCE --zone MY-ZONE'
```

Where:

- `MY-INSTANCE` is your Ops Manager VM instance.
- `MY-ZONE` is the zone where your Ops Manager VM is running.

For example:

```
$ gcloud compute ssh om-pcf-1a --zone us-central1-b
```

5. Switch to the `ubuntu` user by running the following command:

```
sudo su - ubuntu
```

6. Proceed to [Log in to Ops Manager](#).

External UAA

If you configured your Ops Manager for an external Identity Provider with SAML or LDAP, you do not need any IaaS-specific setup. You can proceed to [Log in to Ops Manager](#).

Log in to Ops Manager

To log in to the Ops Manager VM and retrieve your token, do the following:

1. Use the UAAC to target your Ops Manager UAA server:

```
target https://OPS-MAN-FQDN/uaa
```

Where `OPS-MANAGER-FQDN` is the fully qualified domain name of Ops Manager.

For example:

```
$ target https://my-opsmanager.example.com/uaa
```

2. Depending on whether your Ops Manager UAA is internal or external, run a command to retrieve your UAA token and respond to the authentication prompts as follows:

- **Internal**

```
uaac token owner get
Client ID: opsman
Client secret: [Leave Blank]
User name: OPS-MAN-USERNAME
Password: OPS-MAN-PASSWORD
```

Where `OPS-MAN-USERNAME` and `OPS-MAN-PASSWORD` are the credentials that you use to log in to the Ops Manager web interface. For example:

```
$ uaac token owner get
Client ID: opsman
Client secret:
User name: my-opsman-username
Password: my-opsman-password
```

- **External**

```
uaac token sso get
Client ID: opsman
Client secret: [Leave Blank]
Passcode: OPS-MAN-PASSCODE
```

Where `OPS-MAN-PASSCODE` is the value you retrieve from <https://OPS-MAN-FQDN/uaa/passcode>. For example:

```
$ uaac token sso get  
Client ID: opsman  
Client secret:  
Password: my-opsman-passcode
```

If authentication is successful, the UAAC displays the following message:

Successfully fetched token via owner password grant.
--

Run API Commands

To run any Ops Manager API command, you pass your authorization token to the API command endpoint in a header that follows the format `Authorization: Bearer YOUR-ACCESS-TOKEN`.

If you are calling Ops Manager API commands from the command line, the machine that you can be logged into depends on the location of your Ops Manager UAA, as follows:

- **Internal:** You must run Ops Manager API commands from the Ops Manager VM.
 - **External:** You may run commands from your local machine.

API Access Test

The following procedure tests whether you can access the Ops Manager API from the command line by retrieving a list of deployed products:

1. List your tokens by running the following command:

uaac contexts

In the command output, locate the entry for your Ops Manager FQDN. Under `client_id: opsman`, record the value for `access_token`.

2. Use the `GET /api/v0/deployed/products` endpoint to retrieve a list of deployed products:

```
curl "https://OPS-MAN-FQDN/api/v0/deployed/products"  
-X GET \  
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
```

Where `UAA-ACCESS-TOKEN` is the access token recorded in the previous step.

The command output should look like the following:

Example Ops Manager API Procedures

The following procedures illustrate how to test API calls from the command line. If you are a developer writing deployment or test routines in programming languages, these procedures show how you might test the routines line-by-line while building them.

The Ops Manager command-line interface (CLI) `om` is usually a better tool than the Ops Manager API for running Ops Manager operations directly from the command-line or within shell scripts.

These methods are intended for advanced PCF operators and administrators.

Configure the BOSH Director

The following procedure configures the BOSH Director for your IaaS using the Ops Manager API.

Pivotal recommends that beginning users configure the BOSH Director through the Ops Manager UI, which provides context and explanations for each option. For information about configuring the BOSH Director tile in the Ops Manager UI, see one of the following:

- AWS ([Terraform](#) | [manual](#))
- Azure ([Terraform](#) | [manual](#))
- GCP ([Terraform](#) | [manual](#))
- [OpenStack](#)
- [vSphere](#)

To configure the BOSH Director with the Ops Manager API, do the following:

1. Access the Ops Manager API by following the procedure in [Access the Ops Manager API](#) procedure above.
2. To perform initial setup, enter a `POST` request to the Ops Manager API `setup` endpoint, passing metadata to the `-d` flag:

```
curl "https://YOUR-OPS-MANAGER-URL.com/api/v0/setup" \
-X POST \
-H "Content-Type: application/json" \
-d'{
  "setup": {
    "decryption_passphrase": "PASSPHRASE",
    "decryption_passphrase_confirmation": "PASSPHRASE",
    "eula_accepted": "EULA-STATUS",
    "identity_provider": "IDP-LOCATION",
    "admin_user_name": "ADMIN-USERNAME",
    "admin_password": "ADMIN-PASSWORD",
    "admin_password_confirmation": "ADMIN-PASSWORD",
    "http_proxy": HTTP-PROXY,
    "https_proxy": HTTPS-PROXY,
    "no_proxy": "127.0.0.1"
  }
}'
```

Where:

- `PASSPHRASE` is your deployment's decryption passphrase
- `EULA-STATUS` is your EULA status
- `IDP-LOCATION` is your identity provider location, `internal` or `external`
- `ADMIN-USERNAME` is your Ops Manager admin username
- `ADMIN-PASSWORD` is your Ops Manager admin password
- `HTTP-PROXY` is your HTTP proxy, if applicable
- `HTTPS-PROXY` is your HTTPS proxy, if applicable

For example:

```
$ curl "https://my-opsman.example.com/api/v0/setup" \
-X POST \
-H "Content-Type: application/json" \
-d'{
  "setup": {
    "decryption_passphrase": "my-passphrase",
    "decryption_passphrase_confirmation": "my-passphrase",
    "eula_accepted": "true",
    "identity_provider": "internal",
    "admin_user_name": "my-username",
    "admin_password": "my-password",
    "admin_password_confirmation": "my-password",
    "http_proxy": "http://myproxy.example.com",
    "https_proxy": "https://myproxy.example.com",
    "no_proxy": "127.0.0.1"
  }
}'
```

 **Note:** You do not need to specify UAA or other authentication details on a first-time deploy. Making the `setup` request automatically creates a UAA client when it completes successfully. If you specify a UAA client for a first-time deployment, the deployment will fail.

3. A `200 OK` response appears.
4. After the successful response, a UAA client with the metadata you specified is created and launched. Ops Manager users can authenticate with UAAC, and if they have a pre-created client they can target UAAC with the client name and password. For more information on pre-created clients, see [Creating and Managing Ops Manager User and Client Accounts](#).
5. Send an HTTP `PUT` request to the `/api/v0/staged/director/properties` endpoint to configure your IaaS and BOSH Director. There are many configuration parameters available to customize your BOSH Director. All the commands are IaaS-agnostic except the IaaS configuration key. Missing required fields should cause an error. For more information about configuring the BOSH Director, see [Fetching Director, IaaS, and Security Properties](#).

- Send an HTTP `POST` request to the `api/v0/staged/director/availability_zones` endpoint to create Availability Zones (AZs) for your product.

 Note: The `create-azs` endpoint is optional for Azure-based deployments, because you cannot manually configure Azure AZs.

- Send an HTTP `PUT` request to the `api/v0/staged/director/networks` endpoint to create networking rules for the deployment. Specify whether or not to use ICMP checks by setting the `icmp_checks_enabled` parameter to `true` or `false`.
- Send an HTTP `PUT` request to the `api/v0/staged/director/network_and_az` endpoint to assign a singleton AZ and a network where your BOSH Director will be located.
- You must update the BOSH Director's resource configuration settings before deploying BOSH. To update the resource config, do the following:
 - Send a `GET` request to `api/v0/staged/products` to find your Director's GUID.
 - Use the following HTTP request to list all jobs on the Director: `GET api/v0/staged/products/:BOSH-DIRECTOR-GUID/jobs`.
 - Use the following HTTP request to update the resource config of each job on the BOSH Director:
`PUT api/v0/staged/products/:BOSH-DIRECTOR-GUID/jobs/:JOB-GUID/resource_config`.

Configure a Product

The following procedure uploads, stages, and configures a product tile using the Ops Manager API:

- Upload or import the product into Ops Manager.
 - Send an HTTP `POST` request to the `api/v0/available_products` endpoint to upload the product.
 - Send an HTTP `GET` request to the `api/v0/available_products` endpoint to list the names of available products.
 - In the output list from `available_products`, find the product you uploaded by referencing its name and version.
 - Send an HTTP `POST` request to the `/api/v0/staged/products` endpoint to add the uploaded product to Ops Manager.
- Send an HTTP `GET` request to the `api/v0/staged_products` endpoint to confirm that the product is staged for deployment in Ops Manager and display the product GUID.
- Send an HTTP `PUT` request to the `PUT /api/v0/staged/products/:PRODUCT-GUID/networks_and_azs` endpoint to assign availability zones (AZs) and networks to the product. Where `PRODUCT-GUID` is the product GUID.
- Send an HTTP `PUT` request to the `/api/v0/staged/products/:PRODUCT-GUID/properties` endpoint to update the product's properties. Where `PRODUCT-GUID` is the product GUID.
- Send an HTTP `PUT` request to the `/api/v0/staged/products/:PRODUCT-GUID/syslog_configuration` endpoint to configure syslog for the product. Where `PRODUCT-GUID` is the product GUID.
- Update the product's resource configuration by doing the following:
 - Send an HTTP `GET` request to the `api/v0/staged/products/:PRODUCT-GUID/jobs` endpoint to list all jobs on a product and display the GUID for each job. Where `PRODUCT-GUID` is the product GUID.
 - Send an HTTP `PUT` request to the `api/v0/staged/products/:PRODUCT-GUID/jobs/:JOB-GUID/resource_config` endpoint to update the resource config for a particular job. Where:
 - `PRODUCT-GUID` is the product GUID.
 - `JOB-GUID` is the job GUID.

Upgrade Ops Manager

The following procedure upgrades Ops Manager using the Ops Manager API:

Prerequisites

Before you upgrade Ops Manager with the Ops Manager API, you must have the following:

- Two versions of Ops Manager: an existing deployment and a new deployment of a later version. The new deployment cannot be configured in any way.
- DNS information for the existing deployment of Ops Manager

Procedure

Follow this procedure to upgrade your Ops Manager:

1. Send an HTTP `GET` request to the `/api/v0/installation_asset_collection` endpoint of your old Ops Manager to export your existing installation data.
2. Send an HTTP `POST` request to the `/api/v0/installation_asset_collection` endpoint of your new Ops Manager to import your installation data to the new deployment.
3. Send an HTTP `POST` request to the `/api/v0/installations` endpoint of your new Ops Manager to trigger a new installation process.
4. Test the new deployment.
5. (Optional) Delete the old deployment.

Creating UAA Clients for BOSH Director

This topic assumes you are using [BOSH CLI v2+ ↗](#).

This topic describes the process of creating a UAA client for the BOSH Director. You must create an automation client to run BOSH from a script or set up a continuous integration pipeline.

To create an automation client, do one of the following:

- **Local Authentication:** When using Internal Authentication, use the existing admin client to manually create an automation client with the correct privileges.
- **Provision Admin Client:** When setting up SAML or LDAP authentication, ensure Ops Manager provisions an admin client.

Local Authentication

To perform this procedure, the UAAC client must be installed on the Ops Manager virtual machine (VM).

1. Open a terminal and SSH into the Ops Manager VM by following the instructions for your IaaS in the [SSH into Ops Manager](#) topic.
2. Navigate to the Ops Manager **Installation Dashboard** and select the **BOSH Director** tile. In BOSH Director, click the **Status** tab, and record the IP address.
3. Using the `uaac target` command, target BOSH Director UAA on port `8443` using the IP address you copied, and specify the location of the root certificate. The default location is `/var/tempest/workspaces/default/root_ca_certificate`.

Target: `https://10.85.16.4:8443`

Where:

- `BOSH-DIRECTOR-IP` is the IP you recorded in the **Status** tab of the BOSH Director.

Note: You can also curl or point your browser to the following endpoint to obtain the root certificate: `https://OPS-MANAGER-FQDN/api/v0/security/root_ca_certificate`

4. Log in to the BOSH Director UAA and retrieve the owner token. Perform the following step to obtain the values for `UAA-LOGIN-CLIENT-PASSWORD` and `UAA-ADMIN-CLIENT-PASSWORD`:
 - a. Select the **BOSH Director** tile from the Ops Manager **Installation Dashboard**.
 - b. Click the **Credentials** tab, and record the entries for **Uaa Login Client Credentials** and **Uaa Admin User Credentials**.
 - c. For each entry, click **Link to Credential** to obtain the password.

```
$ uaac token owner get login -s UAA-LOGIN-CLIENT-PASSWORD  
User name: admin  
Password: UAA-ADMIN-CLIENT-PASSWORD  
  
Successfully fetched token via owner password grant.  
Target: https://10.85.16.4:8443  
Context: admin, from client login
```

Note: To obtain the password for the UAA login and admin clients, you can also curl or point your browser to the following endpoints:
`https://OPS-MANAGER-FQDN/api/v0/deployed/director/credentials/uaa_login_client_credentials` and
`https://OPS-MANAGER-FQDN/api/v0/deployed/director/credentials/uaa_admin_user_credentials`

5. Create a new UAA Client with `bosh.admin` privileges.

```
$ uaa client add ci --authorized_grant_types client_credentials \
--authorities bosh.admin --secret CI-SECRET
```

```
scope: uaa.none
client_id: ci
resource_ids: none
authorized_grant_types: client_credentials
autoapprove:
action: none
authorities: bosh.admin
name: ci
lastmodified: 1469727130702
id: ci
```

- Set the client and secret as environment variables on the VM.

```
$ ubuntu@ip-10-0-0-12:~$ export BOSH_CLIENT=ci
$ ubuntu@ip-10-0-0-12:~$ export BOSH_CLIENT_SECRET=CI-SECRET
```

- Set an alias for the BOSH Director environment.

```
$ bosh alias-env MY-ENVIRONMENT-NAME -c BOSH-DIRECTOR-IP \
--ca-cert /var/tempest/workspaces/default/root_ca_certificate
```

You can now use the UAA client you created to run BOSH in automated or scripted environments, such as continuous integration pipelines.

Provision Admin Client

Pivotal does not support LDAP or SAML authentication to the BOSH Director. Ops Manager provides an option to create UAA clients during SAML or LDAP configuration so that BOSH can be automated using scripts and tooling.

- Select **Provision an admin client in the Bosh UAA** when configuring Ops Manager for SAML or LDAP.
- Click the **Status** tab, and record the IP address, after deploying BOSH Director (BOSH).
- Click the **Credentials** tab in the BOSH Director tile.
- Click the link for the **Uaa Bosh Client Credentials** to record the client name and secret.
- Open a terminal and SSH into the Ops Manager VM. Follow the instructions for your SSH in the [SSH into Ops Manager](#) topic.
- Set the client and secret as environment variables on the Ops Manager VM. `export BOSH_CLIENT=bosh_admin_client export BOSH_CLIENT_SECRET=UAA-BOSH-CLIENT-SECRET` Where:

- `UAA-BOSH-CLIENT-SECRET` is the client secret you recorded in Step 4. For example:

```
$ ubuntu@ip-10-0-0-12:~$ export BOSH_CLIENT=bosh_admin_client
$ ubuntu@ip-10-0-0-12:~$ export BOSH_CLIENT_SECRET=aBcDeFGhijKabdsadfdfdf
```

- Set an alias for the BOSH Director environment.

```
$ bosh alias-env MY-ENVIRONMENT-NAME -c BOSH-DIRECTOR-IP \
--ca-cert /var/tempest/workspaces/default/root_ca_certificate
```

Where:

- `BOSH-DIRECTOR-IP` is the IP you recorded in the **Status** tab of the BOSH Director.

Using Your Own Load Balancer

This guide describes how to use your own load balancer and forward traffic to your Pivotal Application Service (PAS) router IP address.

[Pivotal Cloud Foundry](#) (PCF) includes a tier of reverse proxies that dynamically track the location of application containers and system components, enabling routing of requests to those endpoints even as IPs and ports change.

In order for the PCF Routers to be horizontally scalable and highly available, a load balancer must be deployed in front of them. The simplest solution is to use a layer-4 TCP load balancer, provided by your IaaS or IT team, which passes all HTTP and TLS handling to the PCF Routers. For details on TLS termination, see [Securing Traffic into Cloud Foundry](#). For a description of features supported by the PCF routing tier, see [HTTP Routing](#).

If you have requirements that are not fulfilled by the PCF Routers alone, you can choose to use your own layer-7 load balancer (provided by your IaaS or IT team), or the HAProxy load balancer included with PCF. If you choose to use HAProxy, you must use a layer-4 TCP load balancer in front of it, in order for HAProxy itself to be highly available. Singleton instances of HAProxy are only for use in lab and test environments.

If you choose to use your own layer-7 load balancer, it must fulfill the following requirements:

- Provides load balancing to each of the PCF Router IP addresses
- Supports TLS termination for wildcard hostnames
- Adds appropriate `x-forwarded-for` and `x-forwarded-proto` HTTP headers to incoming requests
- Sets an HTTP keep-alive connection timeout greater than five seconds
- (Optional) Supports WebSockets

The choice to use HAProxy or your own load balancer depends on what features you need out of a load balancer, and whether you want the ability to configure it yourself.

 **Note:** Application logging with [Loggregator](#) requires WebSockets. To use another logging service, see the [Using Third-Party Log Management Services](#) topic.

For how to install an [F5 Local Traffic Manager \(LTM\)](#) as a load balancer for PCF and Pivotal Application Service (PAS), see [Configuring an F5 Load Balancer for PAS](#).

Prerequisites

To integrate your own load balancer with PCF, you must ensure the following:

- WebSocket connections are not blocked for Loggregator functionality.
- The load balancer must be able to reach the Gorouter IP addresses.

Follow the instructions below to use your own load balancer.

Step 1: Deploy PCF Installation VM

Deploy a PCF Installation virtual machine. See [Deploying Ops Manager on vSphere](#) for more information.

Step 2: Register PCF IP Address

In your load balancer, register the IP addresses that you assigned to PCF.

Step 3: Configure Pivotal Ops Manager and BOSH Director

Configure your Pivotal Operations Manager and BOSH Director as described in [Installing Pivotal Cloud Foundry](#), then add PAS.

Do not click **Install** after adding PAS.

Step 4: Configure Networking

Complete the **Networking** configuration page in PAS. Load balancer configuration in PAS varies depending on which IaaS you are using for PCF. See [Configuring PAS](#).

Step 5: Finalize Changes

1. Return to the **Ops Manager Installation Dashboard**
2. Click **Install**.

Creating and Managing Ops Manager User and Client Accounts

This topic describes how to add and remove Ops Manager users. It also describes how you can use an admin Ops Manager account to create client accounts that you can use for Ops Manager automation.

Overview

Pivotal Cloud Foundry supports multiple user accounts in Ops Manager. A User Account and Authentication (UAA) module co-located on the Ops Manager VM manages access permissions to Ops Manager.

When Ops Manager boots for the first time, you create an admin user. However, you do not create additional users through the Ops Manager web interface. If you want to create additional users who can log into Ops Manager, you must use the UAA API, either through `curl` or the UAA Command Line Client (UAAC).

Users are not the only type of account you can create for Ops Manager. You can also create client accounts, which connect automation tools and scripts to Ops Manager. Pivotal recommends using clients to handle automated tasks.

Client accounts are not bound to the same authentication protocols as user accounts. A user account that controls automated components can cause those components to fail if the account experiences inconsistent availability due to permission or authentication issues.

You can create client accounts after deploying Ops Manager, or during configuration for an initial deployment. For more information about adding clients during initial configuration or after deployment, see [Add Pre-Created Client](#).

Add Ops Manager Users

This section describes how to add or remove users with UAAC. If you do not already have the UAAC installed, run `gem install cf-uaac` on the command line.

 **Note:** You can only manage users on the Ops Manager UAA module if you chose to use Internal Authentication instead of an external Identity Provider when configuring Ops Manager.

To add Ops Manager users, do the following:

1. Target your Ops Manager UAA:

```
uaac target https://YOUR-OPSMANAGER-FQDN/uaa/
```

Where:

- `YOUR-OPSMANAGER-FQDN` is the fully qualified domain name of your Ops Manager installation.

2. Get your token:

```
uaac token owner get  
Client ID: opsman  
Client Secret:  
Username: OPSMANAGER-ADMIN-USERNAME  
Password: OPSMANAGER-ADMIN-PASSWORD
```

```
Successfully fetched token via client credentials grant.  
Target https://YOUR-OPSMANAGER-FQDN/uaa/
```

Where:

- `YOUR-OPSMANAGER-FQDN` is the fully qualified domain name of your Ops Manager installation.
- `OPSMANAGER-ADMIN-USERNAME` and `OPSMANAGER-ADMIN-PASSWORD` are the username and password for the Ops Manager admin user.

 **Note:** The `Client Secret` field does not require a value.

3. Add a user.

```
uaac user add USER-NAME -p USER-PASSWORD --emails USER-EMAIL@EXAMPLE.COM
```

Where:

- `USER-NAME` is the username of the user you are adding.
- `USER-PASSWORD` is the password with which this user authenticates.
- `USER-EMAIL` is the email address associated with this user.

4. (Optional) Set the Role-Based Access Control (RBAC) permissions for your user. For more information, see [Configuring Role-Based Access Control \(RBAC\) in Ops Manager](#).

Remove Ops Manager Users

To remove Ops Manager users, do the following:

1. Target your Ops Manager UAA:

```
uaac target https://YOUR-OPSMANAGER-FQDN/uaa/
```

2. Get your token:

```
uaac token owner get  
Client ID: opsman  
Client Secret:  
Username: OPSMANAGER-ADMIN-USERNAME  
Password: OPSMANAGER-ADMIN-PASSWORD
```

```
Successfully fetched token via client credentials grant.  
Target https://YOUR-OPSMAN-FQDN/uaa/
```

Where:

- `YOUR-OPSMANAGER-FQDN` is the fully qualified domain name of your Ops Manager installation.
- `OPSMANAGER-ADMIN-USERNAME` and `OPSMANAGER-ADMIN-PASSWORD` are the username and password for the Ops Manager admin user.

 **Note:** The `Client Secret` field does not require a value.

3. Delete a user:

```
uaac user delete USER-NAME
```

Where:

- `USER-NAME` is the username of the user you wish to delete.

Add Ops Manager Client Accounts

The following sections describe how to create client accounts for Ops Manager automation using an admin account.

Log in to UAAC as an Admin

In order to configure a client, you must first log in to UAAC as an admin.

Use one of the following two methods to authenticate to UAAC:

- [Authenticate Using SAML or SSO](#)
- [Authenticate Using LDAP](#)

Authenticate Using SAML or SSO

If you're using SAML or SSO, authenticate to UAAC as an admin before creating a client.

To authenticate to UAAC, do the following:

1. Target your UAA server.

```
uaac target https://YOUR-OPSMANAGER-FQDN/uaa
```

Where:

- `YOUR-OPSMANAGER-FQDN` is the fully qualified domain name of your Ops Manager installation.

2. Log in as an admin.

```
uaac token sso get
```

3. When prompted, type the Client ID and passcode. Leave the client secret blank.

```
Client ID: opsman  
Client secret:  
Passcode (from http://YOUR-OPSMANAGER-FQDN/uaa/passcode): YOUR-UAA-PASSCODE
```

Where:

- `YOUR-OPSMANAGER-FQDN` is the fully qualified domain name of your Ops Manager installation.
- `YOUR-UAA-PASSCODE` is the passcode with which you authenticate to UAA.

Authenticate Using LDAP

If you're not using SAML or SSO, authenticate to UAAC as an admin before creating a client.

To authenticate to UAAC, do the following:

1. Target your UAA server.

```
uaac target https://YOUR-OPSMANAGER-FQDN/uaa
```

Where:

- `YOUR-OPSMANAGER-FQDN` is the fully qualified domain name of your Ops Manager installation.

2. Log in as an admin.

```
uaac token owner get
```

3. When prompted, type the Client ID, your username, and your password. Leave the client secret blank.

```
Client ID: opsman  
Client secret:  
User name: admin  
Password: *****
```

4. A message appears confirming that UAAC has `Successfully fetched token`.

Create a Client

After you authenticate to UAAC, create a client to manage automated components and tasks.

To create a client, do the following:

1. Create a client with role-based permissions, an ID, and an authentication secret using UAAC:

```
uaac client add CLIENT-ID --authorized_grant_types client_credentials --authorities opsman.admin --secret CLIENT-SECRET
```

Where:

- o `CLIENT-ID` is the name of your client.
- o `CLIENT-SECRET` is the secret you use to authenticate to your client.

The `opsman.admin` authority referenced in the example above grant specific permissions to the client. You can choose which permissions you wish the client to have by assigning a different role.

For more information about Ops Manager roles and what they do, see [Understanding Roles in Ops Manager](#).

Add Pre-Created Client

You can add a pre-created client to Ops Manager using the Ops Manager API in either of the following ways:

- [Add a Pre-Created Client Before Initial Deployment](#), or
- [Add or Modify a Pre-Created Client for an Existing Deployment](#)

Add a Pre-Created Client Before Initial Deployment

This section describes how to add a pre-created client before you deploy Ops Manager for the first time.

If you do not already have the UAAC installed, run `gem install cf-uaac` on the command line.

To add a pre-created client before you deploy Ops Manager, do the following:

1. Authenticate to and access the Ops Manager API by following the steps in [Using the Ops Manager API](#).
2. Pass the following command:

```
POST /api/v0/setup
```

Include any configuration details your deployment requires. For more information about configuring Ops Manager with the API, see the [Setting Up with SAML](#) in the Ops Manager API documentation.

3. In the configuration details, specify this parameter:

```
"precreated_client_secret": "YOUR-SECRET"
```

Where:

- o `YOUR-SECRET` is the authentication secret you use to access the pre-created client.

 **Note:** The authentication secret for the pre-created client must be between 1 and 255 ASCII characters.

When `precreated_client_secret` is passed, a UAA client is created called `precreated-client`. To use this client, see [Authenticate as the Pre-Created Client](#).

Add or Modify a Pre-Created Client for an Existing Deployment

This section describes how to add a pre-created client after you have already deployed Ops Manager. You may only have one pre-created client at a time. If you have an existing pre-created client, you may also use this procedure to change its authentication secret.

To add a pre-created client to an Ops Manager that has already been deployed, or change the secret of an existing pre-created client, do the following:

1. Authenticate to and access the Ops Manager API by following the steps in [Using the Ops Manager API](#).
2. Pass the following command:

```
PUT /api/v0/uaa/precreated_client
```

```
"precreated_client_secret": "NEW-SECRET"
```

Where:

- `NEW-SECRET` is the authentication secret you use to access the pre-created client. This secret must differ from the previous secret you used. The secret cannot be blank.

 **Note:** The authentication secret for the pre-created client must be between one and 255 ASCII characters.

To use this client, see [Authenticate as the Pre-Created Client](#).

Authenticate as the Pre-Created Client

To authenticate to UAAC as the pre-created client, do the following:

1. Target your UAA server.

```
uaac target https://YOUR-OPSMANAGER-FQDN/uaa
```

Where:

- `YOUR-OPSMANAGER-FQDN` is the fully qualified domain name of your Ops Manager installation.

2. Log in as a client.

```
uaac token client get precreated-client -s "PRECREATED-CLIENT-SECRET"
```

3. A message appears confirming that UAAC has `Successfully fetched token`.

You have authenticated to UAAC as the pre-created client. Use this client to perform any automation tasks you wish.

For more information about the Ops Manager API, see [Using the Ops Manager API](#).

Configuring Role-Based Access Control (RBAC) in Ops Manager

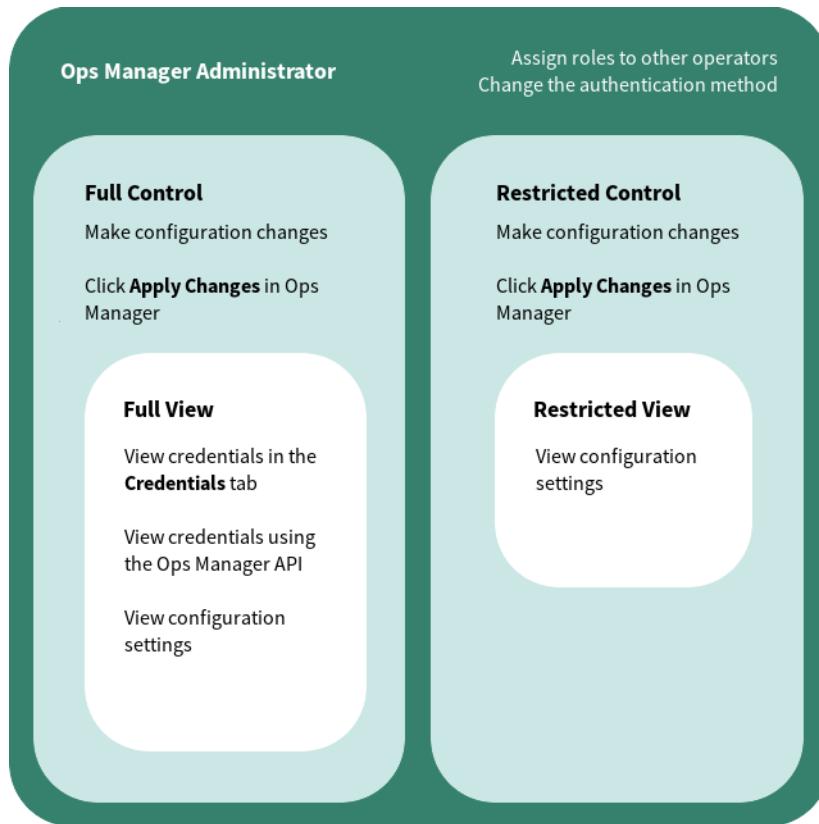
This topic describes how to customize role-based access control (RBAC) in Ops Manager. Use RBAC to manage which operators in your organization can make deployment changes, view credentials, and manage user roles in Ops Manager.

For information about configuring Ops Manager to use internal authentication or SAML authentication, see the Ops Manager configuration topic for your IaaS:

- [Configuring BOSH Director on AWS](#)
- [Configuring BOSH Director on Azure Manually](#)
- [Configuring BOSH Director on GCP](#)
- [Configuring BOSH Director on OpenStack](#)
- [Configuring BOSH Director on vSphere](#)

Roles in Ops Manager

You can assign the following roles to determine which operators in your organization make deployment changes, view credentials, and manage user roles in Ops Manager:



Ops Manager administrators can use the roles defined in the diagram above to meet the security needs of their organization. The roles provide a range of privileges that are appropriate for different types of users. For example, assign either **Restricted Control** or **Restricted View** to an operator to prevent access to all Ops Manager credentials.

See the following table for more information about each role:

Ops Manager Role	Role Definition	UAA Scope
Ops Manager Administrator	Administrators can make configuration changes and click Review Pending Changes and Apply Changes in Ops Manager, view credentials in the Credentials tab and Ops Manager API endpoints, change the authentication method, and assign roles to other operators.	opsman.admin
Full Control	Operators can make configuration changes and click Review Pending Changes and Apply Changes in Ops Manager, and view credentials in the Credentials tab and Ops Manager API endpoints.	opsman.full_control

Ops Manager Role	Role Definition	UAA Scope
Restricted Control	Operators can make configuration changes and click Review Pending Changes and Apply Changes in Ops Manager. They cannot view credentials in the Credentials tab or Ops Manager API endpoints.	opsman.restricted_control
Full View	Operators can view Ops Manager configuration settings and view credentials in the Credentials tab and Ops Manager API endpoints. They cannot make configuration changes or click Apply Changes .	opsman.full_view
Restricted View	Operators can view Ops Manager configuration settings. They cannot make configuration changes or view credentials in the Credentials tab or Ops Manager API endpoints.	opsman.restricted_view

To assign one of the above roles to an operator, follow the procedure for granting access using either [internal authentication](#) or [SAML authentication](#).

When you install a new Ops Manager instance, all existing users have the Ops Manager Administrator role by default.

Simultaneous Ops Manager Administrators

Ops Manager allows multiple administrators to log in to Ops Manager simultaneously and make changes.

The interface does not provide visibility to other administrators that are logged in. Pivotal recommends that administrators communicate with each other and coordinate their changes.

Precedence for Apply Changes

Only one deployment takes precedence when two administrators try to deploy around the same time.

If two administrators are working at the same time, the administrator who first clicks **Apply Changes** takes precedence. Ops Manager overwrites all configurations made by other administrators during deployment.

Pivotal recommends coordinating changes between administrators to avoid overwriting configurations.

Note: If you are having deployment issues or changes to your Ops Manager are not persisting correctly, confirm that your work is not conflicting with an automated administrator.

Enable RBAC in Ops Manager After Upgrade

When you install a new instance of Ops Manager, RBAC is permanently enabled by default.

If your organization has operators who are devoted to managing certain services like MySQL for PCF, you can use RBAC to assign those services operators a more restricted role.

If you upgrade from an older Ops Manager instance, you must enable RBAC and assign roles to users before they can access Ops Manager. If you do not assign any roles to a user, they cannot log in to Ops Manager.

warning: Do not assign roles before you enable RBAC.

Enable RBAC with Internal Authentication

If you are upgrading from an older version of Ops Manager and use internal authentication, do the following to enable RBAC:

1. Log in to the Ops Manager dashboard.
2. Click **Settings** from the user account menu.
3. Click **Advanced**.
4. Click **Enable RBAC**. When the confirmation dialog box appears, click **Confirm and Logout**.

Notes:

- Enabling RBAC is permanent. You cannot undo this action. When you upgrade Ops Manager, your RBAC settings remain configured.
- You will not see this dialog box if RBAC is already configured. With new instances of Ops Manager, RBAC is permanently configured by default.

Enable RBAC with SAML Authentication

If you are upgrading from an older version of Ops Manager and use SAML authentication, perform the steps in this section to enable RBAC. To enable RBAC in Ops Manager when using SAML authentication, you must configure groups in SAML for admins and non-admins and then map the admin group to Ops Manager.

Step 1: Configure SAML Groups

To gather information from your SAML dashboard, do the following:

1. Log in to your SAML provider dashboard.
2. Create or identify the name of the SAML group that contains Ops Manager admin users.
3. Identify the groups attribute tag you configured for your SAML server.

Step 2: Enable RBAC in Ops Manager

Perform the steps above in [Enable RBAC with Internal Authentication](#) to configure Ops Manager to recognize your SAML admin user group.

 **Note:** When RBAC is enabled, only users with the Ops Manager Administrator role can edit SAML configuration.

Create User Accounts in Ops Manager

To assign RBAC roles to operators, you must first create user accounts for them. For more information about creating user accounts in Ops Manager with the User Account and Authentication (UAA) module, see [Creating and Managing Ops Manager User Accounts](#).

In addition to user accounts, you can create a client account to add to Ops Manager. Client accounts manage automation tasks, such as upgrade scripts, log management, and other behaviors that might be negatively impacted if managed by a user account. You can add a client account either before initial deployment or to an existing deployment.

For more information about client accounts, see [Creating and Managing Ops Manager User and Client Accounts](#).

Manage RBAC Roles in Ops Manager

You can assign the roles defined in [Roles in Ops Manager](#) to determine which operators in your organization make deployment changes, view credentials, and manage user roles in Ops Manager.

Manage Roles with Internal Authentication

If you configured Ops Manager to use internal authentication, do the following to configure roles using the [UAA Command Line Interface \(UAAC\)](#):

1. Target your UAA server and log in as an admin:

```
uaac target https://YOUR-OPSMAN-DOMAIN/uaa  
uaac token owner get
```

2. When prompted, enter the following credentials. Enter `opsman` for **Client ID** and leave **Client secret** blank, then enter your username and password:

```
Client ID: opsman  
Client secret:  
User name: USERNAME  
Password: YOUR-PASSWORD
```

3. Assign one of the following roles to a user, replacing `USERNAME` with their username.

- o **Ops Manager Administrator:**

```
uaac member add opsman.admin USERNAME
```

- o **Full Control:**

```
uaac member add opsman.full_control USERNAME
```

- o **Restricted Control:**

```
uaac member add opsman.restricted_control USERNAME
```

- o **Full View:**

```
uaac member add opsman.full_view USERNAME
```

- o **Restricted View:**

```
uaac member add opsman.restricted_view USERNAME
```

Manage Roles with SAML Authentication

If you configured Ops Manager with SAML authentication, do the following to assign non-admin user roles using UAAC:

1. Target your UAA server and log in as an admin:

```
uaac target https://YOUR-OPSMAN-DOMAIN/uaa  
uaac token sso get
```

2. When prompted, enter **Client ID** and **Passcode**, leaving **Client secret** blank:

```
Client ID: opsman  
Client secret:  
Passcode (from http://YOUR-OPSMAN-DOMAIN/uaa/passcode): YOUR-UAA-PASSCODE
```

3. Run the following command:

```
uaac group map SAML-GROUP --name 'OPSMAN-SCOPE' --origin 'saml'
```

Replace the placeholder text as follows:

- o `SAML-GROUP` : Replace with name of the SAML group the user belongs to.
- o `OPSMAN-SCOPE` : Replace with an Ops Manager UAA scope. See the table in [Understand Roles in Ops Manager](#) to determine which UAA scope to use.

4. Add new and existing users to the appropriate SAML groups in the SAML provider dashboard. Users must log out of both Ops Manager and the SAML provider for role changes to take effect.

Manage Roles with LDAP Authentication

If you configured Ops Manager with LDAP authentication, do the following to assign non-admin user roles using UAAC:

1. Target your UAA server and log in as an admin:

```
uaac target https://YOUR-OPSMAN-DOMAIN/uaa  
uaac token sso get
```

- When prompted, enter **Client ID** and **Passcode**, leaving **Client secret** blank:

```
Client ID: opsman  
Client secret:  
Passcode (from http://YOUR-OPSMAN-DOMAIN/uaa/passcode): YOUR-UAA-PASSCODE
```

- Run the following command:

```
uaac group map LDAP-GROUP --name 'OPSMAN-SCOPE'
```

Replace the placeholder text as follows:

- LDAP-GROUP : Replace with name of the LDAP group the user belongs to.
- OPSMAN-SCOPE : Replace with an Ops Manager UAA scope. See the table in [Understand Roles in Ops Manager](#) to determine which UAA scope to use.

- Add new and existing users to the appropriate LDAP groups in the LDAP provider dashboard. Users must log out of both Ops Manager and the LDAP provider for role changes to take effect.

Adding Existing SAML or LDAP Users to a PCF Deployment

This topic describes the procedure for adding existing SAML or LDAP users to a [Pivotal Cloud Foundry](#) (PCF) deployment enabled with SAML or LDAP.

The following two ways exist to add existing SAML or LDAP users to your PCF deployment:

- [Option 1: Import Users in Bulk](#)
- [Option 2: Add Users Manually](#)

Prerequisites

You must have the following to perform the procedures in this topic:

- Admin access to the Ops Manager Installation Dashboard for your PCF deployment
- The [Cloud Foundry Command Line Interface](#) (cf CLI) v6.23.0 or later

Option 1: Import Users in Bulk

You can import SAML or LDAP users in bulk by using the UAA Bulk Import Tool. See the [UAA Users Import README](#) for instructions about installing and using the tool.

Option 2: Add Users Manually

Perform the procedures below to add existing SAML or LDAP users to your PCF deployment manually.

Step 1: Create User

Perform the following steps to add a SAML or LDAP user:

1. Run `cf target https://api.YOUR-SYSTEM-DOMAIN` to target the API endpoint for your PCF deployment. Replace `YOUR-SYSTEM-DOMAIN` with your system domain. For example:

```
$ cf target https://api.example.com
```

2. Run `cf login` and provide credentials for an account with the [Admin user role](#):

```
$ cf login
```

3. Run `cf create-user EXAMPLE-USERNAME --origin YOUR-PROVIDER-NAME` to create the user in UAA. Replace `EXAMPLE-USERNAME` with the username of the SAML or LDAP user you wish to add, and select one of the options below:

- For LDAP, replace `YOUR-PROVIDER-NAME` with `ldap`. For example:

```
$ cf create-user j.smith@example.com --origin ldap
```

- For SAML, replace `YOUR-PROVIDER-NAME` with the name of the SAML provider you provided when configuring Ops Manager. For example:

```
$ cf create-user j.smith@example.com --origin example-saml-provider
```

Step 2: Associate User with Org or Space Role

After creating the SAML or LDAP user, you must associate the user with either an Org or Space role.

For more information about roles, see the [Roles and Permissions](#) section of the *Orgs, Spaces, Roles, and Permissions* topic.

Associate User with Org Role

Run `cf set-org-role USERNAME YOUR-ORG ROLE` to associate the SAML or LDAP user with an Org role. Replace `USERNAME` with the name of the SAML or LDAP user, and replace `YOUR-ORG` with the name of your Org.

For `ROLE`, enter one of the following:

- `OrgManager` : Org Managers can invite and manage users, select and change plans, and set spending limits.
- `BillingManager` : Billing Managers can create and manage the billing account and payment information.
- `OrgAuditor` : Org Auditors have read-only access to Org information and reports.

Example:

```
$ cf set-org-role j.smith@example.com my-org OrgManager
```

Associate User with Space Role

Run `cf set-space-role USERNAME YOUR-ORG YOUR-SPACE ROLE` to associate the SAML or LDAP user with a Space role. Replace `USERNAME` with the name of the SAML or LDAP user, replace `YOUR-ORG` with the name of your Org, and `YOUR-SPACE` with the name of a Space in your Org.

For `ROLE`, enter one of the following:

- `SpaceManager` : Space Managers can invite and manage users, and enable features for a given Space.
- `SpaceDeveloper` : Space Developers can create and manage apps and services, and see logs and reports.
- `SpaceAuditor` : Space Auditors can view logs, reports, and settings on this Space.

Example:

```
$ cf set-space-role j.smith@example.com my-org my-space SpaceDeveloper
```

Modifying Your Ops Manager Installation and Product Template Files

This topic describes how to modify your Ops Manager installation by decrypting and editing the YAML files that Ops Manager uses to store configuration data. Operators can use these procedures to view and change values that they cannot access through the Ops Manager web interface. They can also modify the product templates that Ops Manager uses to create forms and obtain user input.

Operators may want to modify the Ops Manager installation and product template files for a number of reasons, including the following:

- To change the User Account and Authentication (UAA) admin password of their deployment
- To retrieve key values
- To migrate content across different Pivotal Cloud Foundry (PCF) releases

⚠ warning: Be careful when making changes to your Ops Manager installation and product template files. Use spaces instead of tabs, and remember that YAML files use whitespace as a delimiter. Finally, Pivotal does not officially support these procedures, so use them at your own risk.

Understand Installation and Product Template Files

During the installation process, Ops Manager combines information from the installation and product template files to generate the manifests that define your deployment.

- **Installation files:** PCF stores user-entered data and automatically generated values for Ops Manager in installation YAML files on the Ops Manager virtual machine (VM). PCF encrypts and stores these files in the directory `/var/tempest/workspaces/default`. You must decrypt the files to view their contents, edit them as necessary, then re-encrypt them.
- **Product templates:** Ops Manager uses product templates to create forms and obtain user input. The `job_types` and `property_blueprint` key-value pairs in a product template determine how the `jobs` and `properties` sections display in the installation file. Ops Manager stores product templates as YAML files in the directory `/var/tempest/workspaces/default/metadata` on the Ops Manager VM. These files are not encrypted, so you can edit them without decrypting. User input does not alter these files.

Note: Upgrading Ops Manager may eliminate your changes to the installation and product template files.

Modify the Installation Files

Perform the following steps to locate, decrypt, and edit the Ops Manager installation files `installation.yml` and `actual-installation.yml`.

1. SSH into the Ops Manager VM by following the steps in the [SSH into Ops Manager](#) section of the [Advanced Troubleshooting with the BOSH CLI](#) topic.
2. On the command line, navigate to the scripts directory:

```
$ cd /home/tempest-web/tempest/web/scripts/
```

3. Run the following command to decrypt the installation YAML file and make a temporary copy of the decrypted file. When prompted for a passphrase, enter the decryption passphrase you created when you launched Ops Manager for the first time:

```
$ sudo -u tempest-web ./decrypt /var/tempest/workspaces/default/installation.yml /tmp/installation.yml
```

4. Open `/tmp/installation.yml` to view or edit values.
5. If you plan to make changes, make a backup of the original installation YAML file:

```
$ cp /var/tempest/workspaces/default/installation.yml ~/installation-orig.yml
```

6. If you have made changes to your copy of the installation YAML file, you must encrypt it and overwrite the original with it:

```
$ sudo -u tempest-web RAILS_ENV=production /home/tempest-web/tempest/web/scripts/encrypt /tmp/installation.yml /var/tempest/workspaces/default/installation.yml
```

When prompted, enter a passphrase.

7. Delete the temporary copy of the decrypted file:

```
$ rm /tmp/installation.yml
```

8. Repeat steps 2 through 7 for `/tmp/actual-installation.yml`. Each step you see `installation.yml`, replace with `actual-installation.yml`. For example, for step 3 you run:

```
$ sudo -u tempest-web ./decrypt /var/tempest/workspaces/default/actual-installation.yml /tmp/actual-installation.yml
```

9. Restart the Ops Manager web interface:

```
$ sudo service tempest-web stop && sudo service tempest-web start
```

10. Navigate to Ops Manager in a browser and enter your decryption passphrase.

11. Log in to Ops Manager.

12. Click **Review Pending Changes**, then **Apply Changes**.

13. If Ops Manager cannot load your changes, see the [Revert To Your Backup](#) section of this topic to restore your previous settings.

Modify Product Template Files

Perform the following steps to locate and edit your Ops Manager product template files:

1. SSH into the Ops Manager VM by following the steps in the [SSH into Ops Manager](#) section of the *Advanced Troubleshooting with the BOSH CLI* topic.

2. On the Ops Manager VM, navigate to the `/var/tempest/workspaces/default/metadata` directory.

```
$ cd /var/tempest/workspaces/default/metadata
```

3. The `/var/tempest/workspaces/default/metadata` directory contains the product templates as YAML files. If you plan to make changes, make a backup of the original product template YAML file:

```
$ cp /var/tempest/workspace/default/metadata/YOUR-PRODUCT-TEMPLATE.yml ~YOUR-PRODUCT-TEMPLATE-orig.yml
```

4. Open and edit the product template YAML file as necessary. For more information about product templates, see the [Product Template Reference](#) topic.

5. Navigate to Ops Manager to see your changes.

6. If Ops Manager cannot load your changes, see the [Revert To Your Backup](#) section of this to restore your previous settings.

Revert to Your Backup

Perform the following steps to revert to your backup of an installation or product template file:

1. SSH into the Ops Manager VM by following the steps in the [SSH into Ops Manager](#) section of the *Advanced Troubleshooting with the BOSH CLI* topic.

2. Overwrite the modified file with the backup:

- For the installation file, run the following command:

```
$ cp ~/installation-orig.yml /var/tempest/workspaces/default/installation.yml
```

- For a product template file, run the following command:

```
$ cp ~/YOUR-PRODUCT-TEMPLATE-orig.yml /var/tempest/workspaces/default/metadata/YOUR-PRODUCT-TEMPLATE.yml
```

3. Restart the Ops Manager web interface:

```
$ sudo service tempest-web stop && sudo service tempest-web start
```

4. Navigate to Ops Manager in a browser and enter your decryption passphrase.
5. Log in to Ops Manager.
6. Click **Review Pending Changes**, then **Apply Changes**.

Monitoring PCF VMs from Ops Manager

This topic describes how to check current VM status in Pivotal Cloud Foundry (PCF) Ops Manager.

For a complete guide to monitoring PCF, see [Monitoring Pivotal Cloud Foundry](#).

Monitoring VMs Using the Ops Manager Interface

Click any product tile and select the **Status** tab to view monitoring information.

Pivotal Elastic Runtime												
JOB	INDEX	IPS	CID	LOAD AVG15	CPU	MEMORY	SWAP	SYSTEM DISK	EPHEM. DISK	PERS. DISK	LOGS	
HAProxy	0	10.0.0.254	vm-9985a13c-106a-48d1-a3de-d0e0e816c857	0.06%	0.1%	9.6%	0.0%	41%	5%	N/A		
NATS	0	10.0.0.5	vm-dee49615-aea8-4f4f-bf0f-b1060083ddef	0.12%	0.1%	9.7%	0.0%	41%	21%	N/A		
			vm-6d43e59e-									

The columns display the following information:

VM Data Point	Details
Job	Each job represents a component running on one or more VMs that Ops Manager deployed.
Index	For jobs that run across multiple VMs, the index value indicates the order in which the job VMs were deployed. For jobs that run on only one VM, the VM has an index value of 0.
IPs	IP address of the job VM.
CID	Uniquely identifies the VM.
Load Avg15	CPU load average over 15 minutes.
CPU	Current CPU usage.
Memory	Current memory usage.
Swap	Swap file percentage.
System Disk	System disk space usage.
Ephem. Disk	Ephemeral disk space usage.
Pers. Disk	Persistent disk space usage.
Logs	Download link for the most recent log files.

Operations Manager VM Disk Space

The Ops Manager stores its logs on the Ops Manager VM in the /tmp directory.

Note: The logs collect over time and do not self-delete. To prevent the VM from running out of disk space, restart the VM to clear the log entries from /tmp.

Advanced Features in Ops Manager

This guide describes advanced features for Ops Manager operators. Use these features to perform functions in Ops Manager that are outside the scope of typical Ops Manager behavior.

 **warning:** Misuse of any advanced feature could disable or destroy deployments. These features are for advanced operators only. Pivotal recommends contacting [Pivotal Support](#) before using these features.

In This Guide

This guide includes the following topics:

- [Managing Ops Manager Verifiers](#)
- [Managing Custom VM Extensions](#)
- [Ops Manager Fields That Lock On Deploy](#)
- [Property and Template References](#)

Managing Ops Manager Verifiers

This topic describes how to disable and re-enable verifiers for fields in the BOSH Director and other product tiles.

Ops Manager provides this option for troubleshooting purposes. For example, your deployment may have a unique configuration that the verifier cannot detect. In this case, you can unblock your deployment by disabling the verifier.

 **Note:** Pivotal recommends contacting support before you disable Ops Manager verifiers.

 **warning:** Pivotal recommends recording and routinely revisiting which verifiers you disable. The Ops Manager UI does not indicate which verifiers you disable. Disabled verifiers persist through upgrades.

About Verifiers

Ops Manager verifiers narrow the feedback loop for operators by checking that fields are configured correctly. Ops Manager verifiers reach out and find objects in the world. For example, given an IP, a verifier can ping the IP to see that it responds.

Verifiers run early in a deployment. With verifiers turned off, deployments may fail much later.

Verifiers run on two occasions:

Verifier Label	Triggers when you...	Blocks deployment?
form	Click Save in a pane.	Does not block deployment. Error message is a warning.
install_time	Click Apply Changes.	Blocks you from deploying. Disabling verifiers unblocks deployment.

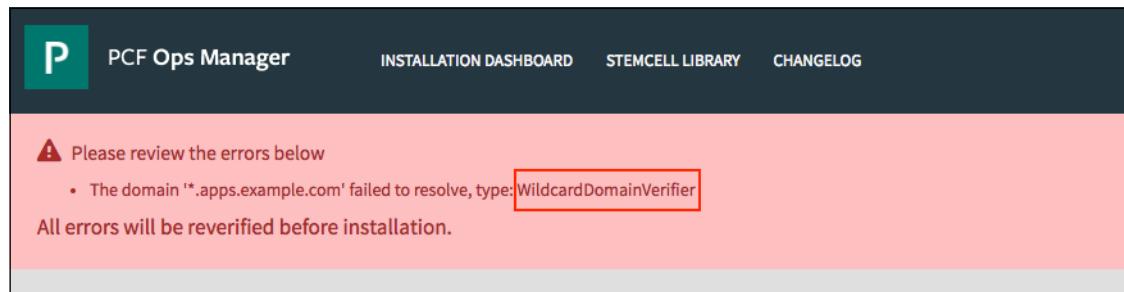
Disable Verifiers by Type

You can disable verifiers by verifier type. For example, you can disable all MySQL verifiers for a tile. To disable a verifier type for a tile, you must find and record the type, then use the Ops Manager API.

Step 1: Record the Verifier Type

To record the verifier type, do the following:

1. Log in to Ops Manager and go to **Review Pending Changes**.
2. Click **Apply Changes** and see the verifier error message.
3. Record the `type` in the verifier error message. For example, `WildcardDomainVerifier`.



The screenshot shows the PCF Ops Manager Installation Dashboard. At the top, there are navigation links: PCF Ops Manager, INSTALLATION DASHBOARD, STEMCELL LIBRARY, and CHANGELOG. Below the header, a red warning box contains the following text:

 Please review the errors below

- The domain '*.apps.example.com' failed to resolve, type: `WildcardDomainVerifier`

All errors will be reverified before installation.

Step 2: Use the Ops Manager API

To disable verifiers, you must use the Ops Manager API. To get started with the Ops Manager API, follow the procedures in [Using the Ops Manager API](#).

The API endpoint varies by tile. Go to one of the procedures below:

- [Disable Verifiers for the BOSH Director Tile](#)
- [Disable Verifiers for Other Tiles](#)

Disable Verifiers for the BOSH Director Tile

To disable a verifier type for the BOSH Director, do the following:

1. Run the following API command:

```
curl "https://OPS-MAN-FQDN/api/v0/staged/director/verifiers/install_time/VERIFIER-TYPE" \
-X PUT \
-H "Authorization: Bearer UAA-ACCESS-TOKEN" \
-H "Content-Type: application/json" \
-d '{ "enabled": false }'
```

Where:

- `OPS-MAN-FQDN` is the fully-qualified domain name (FQDN) for your Ops Manager deployment.
- `VERIFIER-TYPE` is the verifier type you recorded in [Step 1: Record the Verifier Type](#).
- `UAA_ACCESS_TOKEN` is your UAA access token. For more information, see [Access the API](#).

For more information about this endpoint, see [Turning off a verifier for the director](#) in the Ops Manager API documentation.

2. To confirm your verifier type is disabled, run the following command:

```
curl "https://OPS-MAN-FQDN/api/v0/staged/director/verifiers/install_time" \
-X GET \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
```

See that the Ops Manager API shows `false` for the verifier type you disabled. In the following example, `DirectorConfigurationVerifier` is set to `false`:

```
{
  "verifiers": [
    {
      "type": "AllocatedIpAddressVerifier",
      "enabled": true
    },
    {
      "type": "AvailabilityZonesVerifier",
      "enabled": true
    },
    {
      "type": "DirectorConfigurationVerifier",
      "enabled": false
    },
    {
      "type": "IaaSConfigurationVerifier",
      "enabled": true
    }
  ]
}
```

For more information about this endpoint, see [Retrieving a list of install time verifiers for a specific product](#) in the Ops Manager API documentation.

3. Record the product and the verifier type that you disabled in a secure location. Pivotal recommends recording and routinely revisiting which verifiers you disable.

Disable Verifiers for Other Tiles

To disable a verifier type for tiles other than the BOSH Director, do the following:

1. Run the following command and record the `guid` corresponding to your product.

```
curl "https://OPS-MAN-FQDN/api/v0/staged/products" \
-X GET \
-H "Authorization: Bearer UAA_ACCESS_TOKEN"
```

Where:

- `OPS-MAN-FQDN` is the fully-qualified domain name (FQDN) for your Ops Manager deployment.
- `UAA_ACCESS_TOKEN` is your UAA access token. For more information, see [Access the API](#).

See the following example output:

```
[[{"installation_name": "component-type1-installation-name",
 "guid": "component-type1-guid",
 "type": "component-type1"}, {"installation_name": "p-bosh-installation-name",
 "guid": "p-bosh-guid",
 "type": "p-bosh"}]]
```

For more information about this endpoint, see [Listing all staged products](#) in the Ops Manager API documentation.

- To disable your verifier by type, run the following command.

```
curl "https://OPS-MAN-FQDN/api/v0/staged/products/PRODUCT-GUID/verifiers/install_time/VERIFIER-TYPE" \
-X PUT \
-H "Authorization: Bearer UAA_ACCESS_TOKEN" \
-H "Content-Type: application/json" \
-d '{ "enabled": false }'
```

Where:

- `OPS-MAN-FQDN` is the fully-qualified domain name (FQDN) for your Ops Manager deployment.
- `PRODUCT-GUID` is the GUID you recorded in the previous step.
- `VERIFIER-TYPE` is the verifier type you recorded in [Step 1: Record the Verifier Type](#).

For more information about this endpoint, see [Turning off a verifier for a specific product](#) in the Ops Manager API documentation.

- To confirm your verifier type is disabled, run the following command.

```
curl "https://OPS-MAN-FQDN/api/v0/staged/products/PRODUCT-GUID/verifiers/install_time" \
-X GET \
-H "Authorization: Bearer UAA_ACCESS_TOKEN"
```

See that the Ops Manager API shows `false` for the verifier type you disabled. In the following example, `MysqlDatabaseVerifier` is set to `false`:

```
{
  "verifiers": [
    {
      "type": "BlobstoreVerifier",
      "enabled": true
    },
    {
      "type": "MysqlDatabaseVerifier",
      "enabled": false
    },
    {
      "type": "SsoUrlVerifier",
      "enabled": true
    },
    {
      "type": "WildcardDomainVerifier",
      "enabled": true
    }
  ]
}
```

For more information about this endpoint, see [Retrieving a list of install time verifiers for a specific product](#) in the Ops Manager API documentation.

- Record the product and the verifier type that you disabled in a secure location. Pivotal recommends recording and routinely revisiting which verifiers you disable.

Re-enable Verifiers by Type

Pivotal recommends recording and routinely revisiting which verifiers you disable. The Ops Manager UI does not indicate which verifiers you disable. Disabled verifiers persist through upgrades.

To re-enable verifiers, follow the steps detailed in [Use the Ops Manager API](#) above. Instead of setting `enabled` to `false`, set it to `true`.

If you do not know what verifiers are disabled, `curl` the `verifiers` endpoint for each installed product. For more information, see [Retrieving a list of install time verifiers for a specific product](#) in the Ops Manager API documentation.

Managing Custom VM Extensions

This topic describes how to add, delete, apply, and list custom VM extensions using the Ops Manager API.

VM extensions allow you to specify IaaS-specific configurations for your VMs such as associated security groups and load balancers. You can use the Ops Manager API to create custom VM extensions, give them names, and associate those names with BOSH jobs.

⚠ warning: If you use custom VM extensions, you may override other settings. For example, if you use a custom VM extension to add tags, the default tags are removed from all instance groups. To use default tags with custom VM extensions, apply the tags manually.

Prerequisites

To manage custom VM extensions, you need the following:

- **Access to the Ops Manager API:** For more information about the Ops Manager API, see [Using the Ops Manager API](#).
- **Your UAA Access Token Value:** Follow the procedure in [Access the API](#) in the [Using the Ops Manager API](#) topic to retrieve your access token. In each example, replace `UAA_ACCESS_TOKEN` with the value of your access token.
- **Knowledge of your IaaS-specific configurations:** To find configurations you can set for custom VM extensions, see the `VM Types` section of your IaaS CPI in the [BOSH documentation](#). For example, see [VM Types](#) in the AWS CPI docs for a list of AWS VM extension options.

Create or Update a VM Extension

From the `/api/v0/staged/vm_extensions/NAME` endpoint, use `PUT` to enter your custom VM extension. Replace `NAME` in the endpoint with the name of the VM extension.

Enter the following properties:

- `name` : Enter a memorable string value as the name of your new VM extension. You later associate this name to a job.
- `cloud_properties` : Enter your IaaS-specific VM extension properties.

Note: If you are updating a custom VM extension rather than creating a new one, enter the name of the VM extension you want to override.

See the following AWS example:

```
curl "https://example.com/api/v0/staged/vm_extensions/example_vm_extension_name" \
-X PUT \
-H "Authorization: Bearer UAA_ACCESS_TOKEN" \
-H "Content-Type: application/json" \
-d '{"name": "example_vm_extension_name", "cloud_properties": { "source_dest_check": false }}'
```

For more information about this command, see [Creating a new vm extension](#) in the Ops Manager API documentation.

Delete a VM Extension

From the `/api/v0/staged/vm_extensions/NAME` endpoint, use `DELETE` to delete a custom VM extension by `name`. Replace `NAME` in the API endpoint with the name of the VM extension.

See the following example:

```
curl "https://example.com/api/v0/staged/vm_extensions/vm_ext1" \
-X DELETE \
-H "Authorization: Bearer UAA_ACCESS_TOKEN" \
-H "Content-Type: application/json" \
-d '{"name": "vm_ext1"}'
```

For more information about this command, see [Deleting a vm extension](#) in the Ops Manager API docs.

Apply VM Extensions to a Job

From the `/api/v0/staged/products/PRODUCT_GUID/jobs/JOB_GUID/resource_config` endpoint use `PUT` to configure a job with the names of your VM extensions.

For the API endpoint, replace `PRODUCT_GUID` with the Globally Unique Identifier (GUID) of your job. You can find this value by requesting `GET /api/v0/deployed/products`. Also replace `JOB_GUID` with the GUID of your job. You can find this value by requesting `GET /api/v0/staged/products/PRODUCT_GUID/jobs`.

For the `additional_vm_extensions` property, list the names of each of the custom VM extensions you want to associate with that job.

See the following GCP example:

```
curl "https://example.com/api/v0/staged/products/product-type1-guid/jobs/example-job-guid/resource_config" \
-X PUT \
-H "Authorization: Bearer UAA_ACCESS_TOKEN" \
-H "Content-Type: application/json" \
-d '{
  "instances": 1,
  "instance_type": {
    "id": "automatic"
  },
  "persistent_disk": {
    "size_mb": "20480"
  },
  "internet_connected": true,
  "elb_names": ["tcp:load_balancer_1", "http:load_balancer_2"],
  "additional_vm_extensions": ["vm_ext_setting_automatic_restart", "vm_ext_setting_on_host_maintenance"]
}'
```

 **Note:** You must apply changes in order for the new custom VM extension to take effect.

For more information about this command, see [Configuring resources for a job](#) in the Ops Manager API docs.

See Deployed VM Extensions

From the `/api/v0/deployed/vm_extensions` endpoint, use `GET` to see your applied custom VM extensions.

See the following example command:

```
curl "https://example.com/api/v0/deployed/vm_extensions" \
-X GET \
-H "Authorization: Bearer UAA_ACCESS_TOKEN"
```

The output lists all your VM extension names and its associated custom properties defined in `cloud_properties`.

See the following example output:

```
{
  "vm_extensions": [
    {
      "name": "vm_ext1",
      "cloud_properties": {
        "source_dest_check": false
      }
    },
    {
      "name": "vm_ext2",
      "cloud_properties": {
        "key_name": "operations_keypair"
      }
    }
  ]
}
```

For more information about this command, see [Creating a new VM extension](#) in the Ops Manager API documentation.

Ops Manager Fields That Lock On Deploy

This topic describes how locked fields work in Ops Manager, and lists the unlockable and permanently-locked fields in installation tiles.

Overview

Some fields in Ops Manager lock during deployment. After deployment, a locked field is no longer configurable via the Ops Manager UI or API endpoints. Product tile authors configure fields as locked to prevent irrevocable VM malfunction, deployment corruption, and data loss. Most users will not need to edit locked fields after initial deployment.

Because deployment requirements may evolve in unanticipated ways over time, some locked fields are *unlockable*, which means that an advanced PCF operator can unlock them to allow reconfiguration. Other fields are *permanently locked*, and their values cannot be changed. Unlocking the unlockable fields requires you to use Ops Manager's Advanced Mode. For more information about Advanced Mode, see [Unlockable Fields](#).

In summary, there are three types of configurable fields in Ops Manager:

- **Standard fields**, which can be reconfigured after deployment as many times as you need.
- **Unlockable fields**, which lock on deployment but can be unlocked and configured with [Advanced Mode](#).
- **Permanently-locked fields**, which lock on deployment and cannot be unlocked.

⚠ warning: Pivotal recommends that only skilled operators use Ops Manager's Advanced Mode. If you use Advanced Mode incorrectly, you could disable or destroy your deployment.

Unlockable Fields

This section describes the fields that lock during deployment and can be unlocked and configured after deployment. To unlock these fields, you must use Ops Manager's [Advanced Mode ↗](#).

After you have enabled Advanced Mode, you can use the API or Ops Manager UI to reconfigure some fields that were locked previously. Those fields are listed in the following section.

⚠ warning: Pivotal recommends that only skilled operators use Ops Manager's Advanced Mode. If you use Advanced Mode incorrectly, you could disable or destroy your deployment.

Unlockable Fields in the Director Tile

In the **Director Config** pane, you can unlock the following fields after deployment:

In the **CredHub Encryption Provider** fields, when you are using a Luna HSM, you can unlock the following fields after deployment:

- **Encryption Key Name**
- **Provider Partition**
- **HSM Host Address**
- **HSM Port Address**
- **Partition Serial Number**

In the **Assign AZs and Networks** pane, you can unlock the following fields after deployment:

- The **Singleton Availability Zone** dropdown

Unlockable Fields in Product Tiles

Product tiles also have unlockable fields.

In the **Assign AZs and Networks** pane, you can unlock the following fields after deployment:

- Place singleton jobs in
- Balance other jobs

Permanently-Locked Fields

This section describes the fields that lock during deployment and cannot be unlocked. These fields cannot be configured after deployment, either via the Ops Manager UI or via API endpoints.

Permanently-Locked Fields in the Director tile

In the **Director Config** pane, the following fields lock permanently after deployment:

In the **Blobstore Location** section:

- **Enable TLS** in the **Internal** section
- **S3 Endpoint** in the **S3 Compatible Blobstore** section
- **Bucket Name** in the **S3 Compatible Blobstore** section
- **V2 Signature** in the **S3 Compatible Blobstore** section
- **V4 Signature** in the **S3 Compatible Blobstore** section
 - **Region** in the **S3 Compatible Blobstore > V4 Signature** section

 **Note:** Changes to the *Internal* or *S3 Compatible Blobstore* radio buttons in the *Blobstore Location* section will not affect a deployed Ops Manager.

In the **Database Location** section:

- **Host** in the **External MySQL Database** section
- **Port** in the **External MySQL Database** section
- **Username** in the **External MySQL Database** section
- **Password** in the **External MySQL Database** section
- **Database** in the **External MySQL Database** section

 **Note:** Changes to the *Internal* or *External MySQL Database* radio buttons in the *Database Location* section will not affect a deployed Ops Manager.

In the **Create Availability Zone** pane, the following fields lock permanently after deployment:

- You cannot delete an in-use Availability Zone (AZ)

In the **Assign AZs and Networks** pane, all fields lock permanently after deployment.

Configurable Fields in Product Tiles

No fields other than the ones listed above lock in third-party product tiles. However, tile authors can specify which additional fields should lock on deploy by passing the `freeze-on-deploy` flag.

For more information about this flag, see [Common Property Blueprint Attributes](#).

Property and Template References

Page last updated:

Overview

A tile's `metadata` subdirectory contains a product template file in `.yml` format. The product template file specifies how the tile interface collects configurable properties from the user, and how Ops Manager incorporates these properties into its deployment manifest.

You can modify your tile's product template in many different ways. Two important concepts to understand are what the product template contains and how to modify it to include different properties.

This topic explains these two concepts and gives examples of different configuration options for tiles.

Product Template Contents

The product template, a `.yml` file in the tile's `metadata` subdirectory, includes or points to the following:

- **Metadata:** High level information about your tile
- **Dependencies:** Your opportunity to specify product dependencies
- **Property Blueprints:** The templates to represent values
- **Form Types:** The way to expose property blueprints to generated forms
- **Job Types:** configuration properties for jobs defined in the BOSH release
- **Runtime Configs:** specifies the manifest for any number of global deployment configurations

For the purpose of explanation, examples that refer to the product template use the [PCF example tile](#), a functional tile provided by the Ops Manager engineering team that deploys the NGINX web server.

For more information about product template configuration, see [Top-Level Properties](#).

To learn how to configure your tile's forms, see [Form Properties](#).

Modifying the Product Template

You can modify the product template to designate property values.

Within the product template, there are two sections with `manifest` snippets. These sections are:

- `form_types` : This defines the **tile interface**, or how users of your tile view and interact with different tile features.
- `job_types` : This defines the **jobs that the manifest deploys**, or what the tile does in response to user input, developer input, or at set intervals.

You can use special expressions in these snippets to include property values that the tile would otherwise not know about, such as user provided configurable properties or dynamically generated system properties.

These special expressions take two forms:

Double Parentheses Expressions

Double parentheses expressions, such as `((PROPERTY-VALUE))` designate property values that Ops Manager fills in when it generates the deployment manifest. This occurs after the user clicks **Apply Changes**.

These values include configurable properties and properties supplied by Ops Manager.

To include these properties in a manifest snippet, use “double-parens” syntax, which consists of a variable name surrounded by two sets of parentheses, as shown in the following example:

```
manifest: |
pizza_toppings:
  peppers: (( .properties.example_selector.pizza_option.peppers.value ))
```

Triple Parentheses Expressions

Triple-parentheses expressions, such as `((PROPERTY-VALUE))` designate property values that BOSH supplies when it deploys instances of the tile service. For example, CredHub credentials are designated in triple parentheses.

To include these BOSH deploy-time properties in a manifest snippet, use “triple-parens” notation, as shown in the following example:

```
manifest: |
credhub:
  concatenated_password: prefix-(( credhub-password ))-suffix
  password: ((( credhub-password )))
```

When Ops Manager parses a product template and BOSH parses a manifest, they both fill in properties designated by double-parens syntax.

Some property values in a product template, such as CredHub credentials, must be filled in by BOSH on the BOSH Director VM, rather than by Ops Manager.

Ops Manager strips the outer parentheses from these expressions and includes the resulting double-parens expressions in the manifest it generates, for BOSH to evaluate at deploy time.

For more information about product properties, continue reading [Referencing Properties](#).

Referencing Properties

Each property is represented by two segments:

- The **location** of the property.
- The information from the property you want to access, commonly known as **accessors**.

Using these two segments, you can write an expression as:

```
(( LOCATION_OF_PROPERTY.ACCESSOR ))
```

Referencing a Location

The method of referencing the location of the property varies. The following table describes the different methods of referencing a property:

Method name	Description
<code>.properties.top_level_property</code>	Refers to the property blueprint called “top_level_property” and found in the global list of properties of the same product
<code>.job_one.job_level_property</code>	Refers to the property blueprint called “job_level_property” and found in the list of properties of the job “job_one” of the same product
<code>job_level_property</code>	Refers to the property blueprint called “top_level_property” and found in the same product and job whose manifest is currently being evaluated
<code>..other_product.properties.top_level_property</code>	Refers to the property blueprint called “top_level_property” and found in the global list of properties of the product “other_product”
<code>..other_product.job_two.job_level_property</code>	Refers to the property blueprint called “job_level_property” and found in the list of properties of the job “job_one” of the product “other_product”

Referencing an Accessor

Accessors vary between property blueprint types. The following example uses the property blueprint type `string` with its one accessor, `value`. A valid

double-parentheses expression to access the value of this property (assuming it is top-level, and has the name `example-string`) would look like:

```
(( .properties.example-string.value ))
```

Ops Manager allows empty arrays in double-parentheses expressions. For example:

```
(( .properties.example-string.value || [] ))
```

For more information about the available properties and their accessors, see the [Property Blueprint Reference](#).

Product Context

Product context is useful for accessing information about a product. It can be used in a manifest using this syntax:

```
(( ..PRODUCT-NAME.accessor ))
```

These are the allowed accessors:

Accessor	Description
deployment_name	Provides the deployment name passed to BOSH for the product
network_name	Provides the name of the network assigned for the product

Dollar Contexts

Outside of properties, you can also retrieve information about various configuration details of your product and Ops Manager.

- **\$ops_manager:** Any product can use this to obtain information about specific OpsManagers.
- **\$director:** Any product can use this to obtain information about the Director.
- **\$runtime:** Any product can use this to obtain information about the PAS tile.
- **\$self:** Your own product uses this to obtain information about its configuration.

\$ops_manager

ca_certificate	Provides the root CA cert that is used to sign the Director VM
trusted_certificates	Provides a list of certificates that the Director applies to all VMs
http_proxy	Provides the comma separated values (CSV) that are entered when Ops Manager traffic is directed to an HTTP Proxy
https_proxy	Provides the CSVs that are entered when Ops Manager traffic is directed to an HTTPS Proxy
no_proxy	Provides the CSVs that should not go through a proxy
instance_types_catalog	Provides a list of all available VM types. Includes custom VM types.
restricted_view_api_access_credentials	Provides a payload for a UAA client to make read-only non-credential requests against Ops Manager. The payload includes <code>uaa_client_name</code> , <code>uaa_client_secret</code> , <code>hostname</code> , and <code>port</code>

\$director

deployment_ip	Provides the IP address where the BOSH Director is deployed
username	Provides the username for the Director VM
password	Provides the password for the Director VM
ntp_servers	Provides a list of Network Time Protocol (NTP) servers that the Director deploys
ca_public_key	Provides the public key that is used to sign the Director VM
hostname	Provides the hostname for the Director VM

tld	Returns the string <code>bosh</code> as the top-level domain (TLD) of the BOSH Director
bosh_metrics_forwarder_client_name	Provides the BOSH Metrics Forwarder client name
bosh_metrics_forwarder_client_secret	Provides the BOSH Metrics Forwarder client secret
dns_release_present	Exposes the Director configuration for <code>disable_dns_release</code>

\$runtime

system_domain	Provides domain for system-level PCF components, such as the API, UAA, or Apps Manager.
apps_domain	Provides the default domain used by applications
system_api_url	Provides the url to the PCF API component
login_url	Provides the url to authenticate with the PCF components
uaa_url	Provides the url to the PCF UAA component
console_url	Provides the url to the PCF Apps Manager component
default_reply_to	Provides the default reply-to email address for the system domain
default_from	Provides the default from email address of the system domain

\$self

uaa_client_name	Provides the UAA client name created for your product to communicate with the BOSH Director
uaa_client_secret	Provides the UAA client secret created for your product to communicate with the BOSH Director
service_network	Provides the name of the service network assigned to your product
stemcell_version	Provides the version of the stemcell your product uses

Complex Properties: Selectors and Collections

Most properties are simple values such as strings, integers, URL addresses, or IP addresses. Others are complex, such as selectors or collections.

- **Selectors:** Give the user a choice of a set of inputs.
- **Collections:** Give the user the ability to enter an array of values to create a hash.

Selectors appear as follows:

A selector example form

Food Choices*

Pizza

Add Pepperoni

Add Pineapple

Other toppings

Filet Mignon

How rare?*

Save

Collections appear as follows:

A collection example form

Albums collection

The albums

Add

▶ Total Eclipse of the Heart

Remove

Name of the Album *

Remove

Name of the Artist *

Remove

Explicit?

Save

The selector and collections inputs are referenced by their selector and collection property blueprints.

Most properties are simple values such as strings, integers, URL addresses, or IP addresses. Selectors and collections are more complicated than simple properties because they contain manifest snippets, which are further referenced in other manifest snippets. Selectors and collections can exist alone, or nest within each other or within different parts of the manifest.

Top-Level Properties

The following is an example of the properties that appear at the top of a product template. Definitions of each property follow this example.

```
---  
name: example-product
```

```

product_version: "1.0-build.0"
minimum_version_for_upgrade: "1.7.0"
pivnet_filename_regex: "product-*.pivotal"
metadata_version: "1.11"
opsmanager_syslog: true
label: 'Ops Manager: Example Product'
description: An example product to demonstrate Ops Manager product-author features
rank: 1
service_broker: false # Default value
stemcell_criteria:
  os: ubuntu-xenial
  version: "97.0"

enable_patch_security_updates: true

requires_product_versions:
  - name: p-mysql
    version: '~> 2.4'

releases:
  - name: example-release
    file: "example-release.tgz"
    version: "15"

variables:
  - name: credhub-password
    type: password

post_deploy_errands:
  - name: example-errand

pre_delete_errands:
  - name: example-errand

runtime_configs:
  - name: example-runtime-config
    runtime_config: |
      releases:
        - name: os-conf
          version: 15
      addons:
        - name: login
          jobs:
            - name: login-banner
              release: os-conf
      properties:
        login_banner:
          text: |
            (( .properties.example_string.value )).


```

name

- **Format:** String
- **Type:** Required

The internal name of the product. You must keep the name of your product consistent for migrations to function properly. Changing the name indicates the installation of a completely different product.

product_version

- **Format:** String
- **Type:** Required

The version of the product.

You can only import this version into Ops Manager once. If you want to import the same product or version, you must delete the existing one from the `/metadata` folder and delete the installation files from Ops Manager's disk.

The version number is important for [migrations ↗](#).

minimum_version_for_upgrade

- **Format:** String
- **Type:** Required

You must set a minimum version for upgrading to your current product version. This example shows a current product version of v1.7 that only upgrades from a v1.6.x version of the same product:

```
- product_version: 1.7.0.0
minimum_version_for_upgrade: 1.6.0.0
```

metadata_version

- **Format:** String
- **Type:** Required

The version of the schema of the product template, the file that this documentation describes.

Bumping the version number enables new properties that are not present in an older `metadata_version` but may also require restructuring of the product template to conform to the new `metadata_version`.

The `metadata_version` corresponds to a major and/or minor release of Ops Manager. Ops Manager can accept a product with a `metadata_version` that is older than the Ops Manager version, but Ops Manager cannot accept a product with a `metadata_version` that is newer than the Ops Manager version. It is therefore best practice to set the `metadata_version` to be the same as the oldest version of Ops Manager that the tile version supports.

label

- **Format:** String
- **Type:** Optional

The label that appears in the product tile when it displays in the Ops Manager Dashboard.

description

- **Format:** String
- **Type:** Optional

A description of the product.

rank

- **Format:** Integer
- **Type:** Optional

The order in which a product tile appears on the dashboard. The BOSH Director always appears at rank 100.

For your product to appear to the right of BOSH Director, you must set this value to an integer less than 100.

Pivotal recommends that you set it to 1.

If all tiles have the same rank, Ops Manager sorts them alphabetically.

pivnet_filename_regex

- **Format:** String
- **Type:** Optional

This regular expression allows Ops Manager's Pivotal Network integration to pull a specific product file. You must do this when there are multiple products with the same product slug.

service_broker

- **Format:** Boolean
- **Type:** Optional, default `false`

Set `service_broker` to `true` for on-demand service brokers. Setting `service_broker` to `true` does the following:

- Enables the service network selector property type
- Requires the operator to select a service network during tile configuration. Tile authors can reference the selected service network with `(($self.service_network))`.
- Includes a UAA client for the service to use. Tile authors can reference the UAA client credentials with `(($self.uaa_client_name))` and `(($self.uaa_client_secret))`.

stemcell_criteria

- **Format:** Hash
- **Type:** Required

`enable_patch_security_updates` enables you to automatically use the latest patched version of a stemcell. This key is set to `true` by default. For products using static compilations, you can disable this feature. This feature increases security by automatically using the latest patched version of a stemcell. However, operators might experience longer than expected upgrade times.

Example:

```
stemcell_criteria
  os: ubuntu-xenial
  version: "97.0"
  enable_patch_security_updates: false
```

Pivotal recommends setting the `enable_patch_security_updates` property to `true`. If you set the property to `false`, your product does not receive security patches through automatic stemcell updates.

You do not specify which IaaS the stemcell targets. This keeps your product template IaaS agnostic so that one product template can be deployed on any IaaS.

For a list of stemcells, including OS and version, see [the BOSH hub](#).

For more information, see [Floating Stemcells](#).

additional_stemcells_criteria

- **Format:** Array of hashes
- **Type:** Optional

A list of stemcell criteria that can be selected using the `stemcell_selector` property type. These stemcells are listed in addition to the stemcell configured with `stemcell_criteria`. After an operator has selected a stemcell, it becomes required.

For more information about the `stemcell_selector` property type, see [stemcell_selector](#) below.

Example:

```
additional_stemcells_criteria:
  - os: ubuntu-xenial
    version: 97.0
    enable_patch_security_updates: false
  - os: windows2016
    version: 1709.1
```

requires_product_versions

- **Format:** Array of hashes
- **Type:** Required

A list of product dependencies. If the required product is not present in the PCF installation, Ops Manager lists the needed dependencies and does not install your tile until the dependencies are satisfied.

releases

- **Format:** Array of hashes
- **Type:** Required

The list of releases contained in your product's releases directory.

The version of the release must be exactly the same as the version contained in the release (BOSH releases are versioned and signed by BOSH).

Each release requires the following keys:

- `name`
- `file`
- `version`

variables

- **Format:** Array of hashes
- **Type:** Optional

A list of variables, that are generated after a deploy succeeds.

You can reference variables in a manifest snippet using triple-parentheses expressions.

Each release requires the following keys:

- `name`
- `type`

post_deploy_errands

- **Format:** Array of hashes
- **Type:** Optional

A list of errands that run after a deploy succeeds.

Set the `run_post_deploy_errand_default:` property to `on` or `off` to set the default for the errand's run rule selector in Ops Manager.

See [Lifecycle Errands](#). If this property is not supplied, the selector defaults to `On`.

pre_delete_errands

- **Format:** Array of hashes
- **Type:** Optional

A list of errands that run before a deployment is deleted.

Set the `run_pre_delete_errand_default:` property to `on` or `off` to set the default for the errand's run rule selector in Ops Manager. If you do not specify this property, the selector defaults to `On`.

For more information, see [Lifecycle Errands](#).

runtime_configs

- **Format:** Array of hashes
- **Type:** Optional

These are keys that define global deployment configurations. For more information, see [Managing Runtime Configs ↗](#).

install_time_verifiers

- **Format:** Array of hashes
- **Type:** Optional

Install-time verifiers confirm that properties for external services are set correctly. Install-time verifiers run in Ops Manager before changes are applied. For example, `Verifiers::MysqlDatabaseVerifier` checks that MySQL database properties are valid by using the properties to connect to the MySQL database.

Each install-time verifier requires the following keys:

- `name` : The type of verifier. The value must be one of the verifiers listed in the table below.
- `properties` : The properties to verify. The required properties for each verifier are listed in the table below.
- `ignorable` : If this key is set to `false`, an operator cannot **Apply Changes** until the verification errors are resolved. If this key is set to `true`, an operator can ignore the warning message and **Apply Changes**. This key is set to `false` by default.

The following table lists the available install-time verifiers:

Verifier Type	Verifies that...	Required Properties
<code>Verifiers::BlobstoreVerifier</code>	the specified AWS s3 bucket exists and is accessible.	<ul style="list-style-type: none"> <code>access_key_id</code> <code>bucket_name</code> <code>secret_access_key</code> <code>signature_version</code>
<code>Verifiers::LDAPBindVerifier</code>	the specified LDAP server exists and that the provided credentials are valid.	<ul style="list-style-type: none"> <code>url</code> <code>credentials</code>
<code>Verifiers::MysqlDatabaseVerifier</code>	the specified MySQL server exists and that the provided credentials and database are valid.	<ul style="list-style-type: none"> <code>database</code> <code>host</code> <code>password</code> <code>port</code> <code>username</code>
<code>Verifiers::SmtpAuthenticationVerifier</code>	the specified SMTP server exists and that the provided credentials are valid.	<ul style="list-style-type: none"> <code>credentials</code> <code>address</code> <code>port</code> <code>hello_domain</code> <code>authentication</code> <code>enable_starttls_auto</code>
<code>Verifiers::SsoUrlVerifier</code>	the specified SSO server exists.	<ul style="list-style-type: none"> <code>url</code>
<code>Verifiers::StaticIpsVerifier</code>	all requested static IPs fall within the subnet assigned to the job. See <code>static_ips</code> in Job Types below.	(none)
<code>Verifiers::WildcardDomainVerifier</code>	the specified domain resolves using DNS.	<ul style="list-style-type: none"> <code>domain</code>
<code>Verifiers::VsphereConfigurationVerifier</code>	the vSphere configuration is valid.	<ul style="list-style-type: none"> <code>vcenter_hosturl</code> <code>vcenter_username</code> <code>vcenter_password</code> <code>datacenter</code>

The following is an example of a `install_time_verifiers` section:

```
install_time_verifiers:
- ignorable: true
  name: Verifiers::WildcardDomainVerifier
  properties:
    domain: .cloud_controller.system_domain
- name: Verifiers::SsoUrlVerifier
  properties:
    url: .properties.uaa.saml.sso_url
- ignorable: true
  name: Verifiers::BlobstoreVerifier
  properties:
    access_key_id: .properties.system_blobstore.external.access_key
    bucket_name: .properties.system_blobstore.external.buildpacks_bucket
    endpoint: .properties.system_blobstore.external.endpoint # Optional. Defaults to AWS
    region: .properties.system_blobstore.external.region # Optional. Defaults to default IaaS setting
    secret_access_key: .properties.system_blobstore.external.secret_key
    signature_version: .properties.system_blobstore.external.signature_version
- ignorable: true
  name: Verifiers::MysqlDatabaseVerifier
  properties:
    database: .properties.account_database_name
    host: .properties.system_database.external.host
    password: .properties.system_database.external.account_password
    port: .properties.system_database.external.port
    username: .properties.system_database.external.account_username
```

icon_image

- **Format:** Base64 Image
- **Type:** Required

This is the icon that displays on the tile in the Ops Manager Installation Dashboard.

base_releases_url

- **Format:** String
- **Type:** Optional

A publicly available URL that is passed to BOSH in order to download your releases during installation. If you use this field, BOSH Director must have internet connectivity. This allows you to package your tile without releases, decreasing the tile size.

opsmanager_syslog

- **Format:** Boolean
- **Type:** Optional, default `false`

Set `opsmanager_syslog` to `true` to opt in to the Ops Manager provided syslog form. If you opt in, you do not need provide the Syslog BOSH release for your tile.

For more information about the Syslog BOSH release, see the [syslog release](#) GitHub repository.

For more information about migrating an existing form to the Ops Manager provided syslog form, see [Migrating Existing Syslog Configuration to Ops Manager](#).

Form Properties

Each form type you write is composed of form properties. Form properties dictate the structure of the form fields that appear in the Ops Manager UI. The `name` of each form appears on the left-hand side as navigational tabs.

Form properties reference [property_blueprints](#). Property blueprints define each field's data type. For a corresponding example to the [form_types](#) example below, see [property_blueprints](#).

The following is an example of the properties that appear in the [form_types](#) section of a product template:

```

form_types:
- name: example-form
  label: Configurable Properties
  description: All the properties that you can configure!
  markdown: |
    ## I am markdown, hear me _roar_.

  ![[Alt text](http://placekitten.com/g/400/200)

  Things to do:
  1. Learn [markdown](https://daringfireball.net/projects/markdown/).
  1. ...
  1. Profit!

  property_inputs:
  - reference: .web_server.example_string
    label: Example string
    description: 'Configure a property of type string'
  - reference: .web_server.example_string_with_placeholder
    label: Example string containing Placeholder text
    description: 'Optional field. Configuration not necessary'
    placeholder: 'Ghost text. Spooky!'
  - reference: .web_server.example_migrated_integer
    label: Example integer
    description: 'Configure a property of type integer'
  - reference: .web_server.example_boolean
    label: Example boolean
    description: 'Configure a property of type boolean'
  - reference: .web_server.example_dropdown
    label: Example dropdown
    description: 'Configure a property of type dropdown'
  - reference: .web_server.example_domain
    label: Example domain
    description: 'Configure a property of type domain'
  - reference: .web_server.example_wildcard_domain
    label: Example wildcard_domain
    description: 'Configure a property of type wildcard_domain'
  - reference: .web_server.example_string_list
    label: Example string_list
    description: 'Configure a property of type string_list'
  - reference: .web_server.example_text
    label: Example text
    description: 'Configure a property of type text (setting to "magic value" causes the web server job instance count to go to 0)'
  - reference: .web_server.example_ldap_url
    label: Example ldap_url
    description: 'Configure a property of type ldap_url'
  - reference: .web_server.example_email
    label: Example email
    description: 'Configure a property of type email'
  - reference: .web_server.example_http_url
    label: Example http_url
    description: 'Configure a property of type http_url'
  - reference: .web_server.example_ip_address
    label: Example ip_address
    description: 'Configure a property of type ip_address'
  - reference: .web_server.example_ip_ranges
    label: Example ip_ranges
    description: 'Configure a property of type ip_ranges'
  - reference: .web_server.example_multi_select_options
    label: Example multi_select_options
    description: 'Configure a property of type multi_select_options'
  - reference: .web_server.example_network_address_list
    label: Example network_address_list
    description: 'Configure a property of type network_address_list (this property was marked with freeze_on_deploy, and so will not be editable after changes are first applied)'
  - reference: .web_server.example_network_address
    label: Example network_address
    description: 'Configure a property of type network_address'
  - reference: .web_server.example_port
    label: Example port
    description: 'Configure a property of type port'
  - reference: .web_server.example_smtp_authentication
    label: Example smtp_authentication
    description: 'Configure a property of type smtp_authentication'
  - reference: .web_server.client_certificate
    label: Example certificate
    description: 'Configure a certificate'

  verifiers:
  - name: Verifiers::WildcardDomainVerifier
    properties:
      domain: .web_server.example_wildcard_domain
  - name: Verifiers::StaticIpsVerifier
    properties:
      domain: .web_server.example_ip_address

```

name

- **Format:** String
- **Type:** Required

The internal name of the form.

label

- **Format:** String
- **Type:** Required

The label of the form as it appears as a link on the left hand side of each form.

description

- **Format:** String
- **Type:** Optional

The description of the form. Appears at the top of the form as a header.

markdown

- **Format:** Markdown
- **Type:** Optional

Provide a block of markdown to display at the top of the form, including image support. Use this property to document the tile and provide explanations or references.

property_inputs

- **Format:** Array of hashes
- **Type:** Required

References to properties defined in the [property_blueprints](#) section of the product template.

verifiers

Verifiers confirm that properties are set correctly in forms. If a verifier displays a warning, the form can still be saved and the warning can be ignored.

Verifiers are similar to install-time verifiers except that they run after a form is saved and not before **Apply Changes**. If you want to have a verifier run when a form is saved and before **Apply Changes**, add the verifier to both the `verifiers` and `install_time_verifiers` sections.

For information about how verifiers are configured, see [install_time_verifiers](#).

The following is an example of a `verifiers` section:

```

form_types:
- name: example-form
  label: Configurable Properties
  description: All the properties that you can configure!
  property_inputs:
    - reference: .web_server.example_wildcard_domain
      label: Example wildcard_domain
      description: 'Configure a property of type wildcard_domain'
    - reference: .web_server.static_ips
      label: Example ip_address
      description: 'Configure a property of type ip_address'
  verifiers:
    - name: Verifiers::WildcardDomainVerifier
      properties:
        domain: .web_server.example_wildcard_domain
    - name: Verifiers::StaticIpsVerifier

```

placeholder

- **Format:** String
- **Type:** Optional

Specify placeholder text for a field. The text appears in light gray to show an example value for the user. The text disappears when the user types in the field and reappears if the user leaves the field empty.

The `placeholder` attribute displays for the following form types:

- string
- integer
- domain
- wildcard_domain
- string_list
- text
- ldap_url
- email
- http_url
- ip_address
- ip_ranges
- network_address_list
- network_address
- port

Job Types

The following is an example of the `job_types` section that appears in a product template.

This section defines the jobs that end up in a BOSH manifest. Those jobs are defined in your BOSH release.

Jobs require many different settings in order to function properly. That is the core of what Ops Manager does for you: Ops Manager asks a user to designate values for those settings and generates a manifest based on what the user specifies.

Ops Manager does not require tile authors to provide `vm_credentials` in the `property_blueprints` for each `job_type`. This is because `vm_credentials` are generated automatically. You can find them in the release manifest.

 **Note:** Ops Manager ignores `static_ip` and `dynamic_ip` keys. To configure the static IPs of a `job_type`, create a property blueprint with `name: static_ips` and `type: ip_ranges`.
`static_ips` is a special keyword that Ops Manager uses to find static IPs and pass them to the BOSH manifest. For more information about `ip_ranges`, see [ip_ranges](#).

```
job_types:
  - name: web_server
    resource_label: Web Server
    templates:
      - name: web_server
        release: example-release
      - name: time_logger
        release: example-release
    release: example-release
    static_ip: 1
    dynamic_ip: 0
    max_in_flight: 1
    single_az_only: true:
    instance_definition:
      name: instances
      type: integer
      configurable: true
      default: 1
      constraints:
        max: 1
      zero_if:
        property_reference: '.web_server.example_text'
        property_value: 'magic value'
    resource_definitions:
      - name: ram
        type: integer
```

name

- **Format:** String
- **Type:** Required

The name of the job as it will be created in the Ops Manager generated BOSH manifest.

resource_label

- **Format:** String
- **Type:** Required

The label of the job as it will appear in the resources page of the tile.

templates

- **Format:** Array of hashes
- **Type:** Required

Each element has the following fields:

name

- **Format:** Name of the job template to use
- **Type:** Required

release

- **Format:** Name of the release the template is from
- **Type:** Required

consumes

- **Format:** YAML string defining [BOSH links](#) that this job consumes

- **Type:** Optional

provides

- **Format:** YAML string defining [BOSH links](#) that this job provides
- **Type:** Optional

This is a BOSH feature (creating jobs from different releases). See the [BOSH documentation](#) for more information.

release

- **Format:** String
- **Type:** Required

The name of the BOSH release contained in your product archive, which is the `.pivotal` file.

single_az_only

- **Format:** Boolean
- **Type:** Required

You can give users control of balancing jobs across availability zones (AZs) by setting `single_az_only` to `false`.

To limit a job to a single AZ, set `single_az_only` to `true`.

A **warning:** If you change the `single_az_only` setting, your VMs may switch AZs. This change can cause an orphaned disk.

max_in_flight

- **Format:** Integer
- **Type:** Required

A BOSH setting that controls the number of instances of this job that BOSH will deploy in parallel.

resource_definitions

- **Format:** Array of hashes
- **Type:** Required

A set of resource settings for the job along with maximum and minimum constraints, defaults, and enabled or disabled configurability.

The resources you can set are:

- ram
- ephemeral_disk
- persistent_disk
- cpu

For examples of supported constraint types, see [integer](#) below.

💡 Note: If you set the `default` property for `persistent_disk` to `0`, users cannot edit this value and the **Resource Config** page in **Ops Manager** displays `None` under the persistent disk field.

instance_definition

- **Format:** Hash
- **Type:** Required

The number of default instances for a job, including maximum, minimum, odd, and the ability to decrease sizing after deploy.

If your product uses an external service that performs the same job as a service in Pivotal Application Service (PAS), you can reduce resource usage by setting the instance count of a job to `0` with the `zero_if` property.

For example, your product uses Amazon Relational Database Service (RDS) instead of MySQL, which is the default system database for PAS. Remove all instance counts of MySQL by setting `property reference` to `.properties.system.database` and `property value` to `magic value`.

For examples of supported constraint types, see [integer](#) below.

 **Note:** If you are using the `zero_if` property, you should not set the `may_only_increase` constraint to true.

manifest

- **Format:** Text snippet, prefaced by pipe symbol: `|`
- **Type:** Optional

Ops Manager generates a BOSH manifest that defines properties for each job the manifest deploys. Some of these properties are not set until the user clicks **Apply Changes**, because the user configures them in the tile or because Ops Manager has to generate them.

For more information about referencing these properties, see [Referencing Properties](#).

Job Accessors

Ops Manager provides double parentheses expressions to access information about the jobs that the manifest deploys.

<code>((name))</code>	Name of the job
<code>((cpu))</code>	CPU of the VM type configured for the job
<code>((ram))</code>	RAM of the VM type configured for the job
<code>((ephemeral_disk))</code>	Ephemeral disk of the VM type configured for the job
<code>((persistent_disk))</code>	Persistent disk of the VM type configured for the job
<code>((instances))</code>	Number of instances configured for the job
<code>((any_instances))</code>	Returns true if there is a non-zero number of instances for the job
<code>((subnet_cidrs))</code>	List of CIDRs associated with the job

Property Blueprint Reference

Common Property Blueprint Attributes

All property blueprints can have the following attributes:

<code>name</code>	Required. The name of the property. The name is used to reference a property in <code>form_types</code> and in <code>(())</code> accessors.
<code>type</code>	Required. The type of property. Must be one of the property types listed in this documentation.
<code>optional</code>	Optional. Default: False When set to true, Ops Manager will not require this property to be set in order to deploy, nor will it validate that this field has a value when saving a form with this property. <code>optional</code> cannot be set to true if a <code>default</code> value is set (see documentation below for which types support <code>default</code>).

	<code>optional</code> cannot be set to true if the property also has <code>configurable: false</code> .
configurable	<p>Optional. Default: false.</p> <p>When set to true for property types which support operator configuration, the operator will be allowed to configure this value.</p> <p>Do not set <code>configurable: true</code> for property types which do not support operator configuration.</p> <p>When set to false, Ops Manager will not render this property in any form, even if it's specified in a <code>form_type</code>, nor will it allow the property to be updated via the API.</p> <p>For property types which support auto-generation of values, when <code>configurable</code> is false, Ops Manager will generate and save a value for this property when the product to which this property belongs is deployed for the first time.</p>
freeze_on_deploy	<p>Optional. Default: false.</p> <p>When set to true, Ops Manager will not allow this property to be changed after the product this property belongs to is successfully deployed.</p>

Configurable and Auto-Generated Properties

When a property has `configurable: true` and is added to a `form_type`, the operator can configure the value of that property.

Alternatively, a property with `configurable: false` will have a value auto-generated by Ops Manager. See below for which property types support each of these usages.

Properties with `configurable` set to `false`:

- cannot be edited by the operator.
- will not show up in forms, even if added under `form_types`.
- will have auto-generated values filled in by OpsManager if that specific type of property supports auto-generation of values.

named_manifest for selector and collection type properties

Specify a property for collection within the `named_manifest` section of the metadata.

```
- name: certificate_collection
  type: collection
  configurable: true
  property_blueprints:
    - name: some_cert_name
      type: string
    - name: some_cert
      type: rsa_cert_credentials
  named_manifests:
    - name: for_routing
      manifest: |
        name: (( current_record.some_cert_name.value ))
        private_key: (( current_record.some_cert.private_key_pem ))
        public_key: (( current_record.some_cert.public_key_pem ))
        certificate: (( current_record.some_cert.cert_pem ))
```

Use the `current_record` property within a collection record to refer to other properties in the same record.

For example, the properties in the `for_routing` named manifest refer to the values for `name`, `private_key`, `public_key`, and `certificate` within this record only.

 Note: The `current_record` property is reserved. You cannot create a new property named `current_record`.

After defining a named manifest, you can reference it using a manifest snippet in the following format:

```
routing_certificates: (( .properties.certificate_collection.parsed_manifest(for_routing) ))
```

Ops Manager renders the following manifest from this example:

```
routing_certificates:
  - name: foo_cert
    private_key: PRIVATE-KEY
```

```

public_key: PUBLIC-KEY
certificate: CERTIFICATE
- name: bar_cert
private_key: PRIVATE-KEY
public_key: PUBLIC-KEY
certificate: CERTIFICATE

```

For more information, see the sections on [Selector and Collection types](#).

All Property Blueprint Types

Below is a reference to every property blueprint type.

boolean

This holds a single boolean value.

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> .
---------	--

Accessors

value	Returns the boolean value of the property.
value_present?	Returns true only if the value of the property is set.

Example

```

property_blueprints:
- name: example_boolean
  type: boolean
  configurable: true
  default: false

form_types:
- name: example_form
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_boolean
      label: 'Example checkbox'
      description: 'This is an example checkbox'

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: user_add
      release: os-conf
      manifest: |
        persistent_homes: (( .properties.example_boolean.value ))

```

An example form

Example checkbox This is an example checkbox

Save

ca_certificate

This holds a string value

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> .
---------	--

Accessors

value	Returns a string or null.
value_present?	Returns true only if the value of the property is set.

Example

```

property_blueprints:
- name: ca_cert
  type: ca_certificate
  configurable: true

form_types:
- name: example_form
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.ca_cert
      label: 'Trusted CA certs'
      description: 'These certificates will be trusted by the deployed VM'

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: ca_certs
      release: os-conf
      manifest: |
        certs: (( .properties.ca_cert.value ))
  
```

An example form

Trusted CA certs *

```
-----BEGIN CERTIFICATE-----
MIIEkjCCAxqAwIBAgIBATANBgkqhkiG9w0BAQUFADB7MQswCQYDVQQGEwJHQjEb
MBkGA1UECAwSR3JlYXRlcibNYW5jaGVzdGVyMRAwDgYDVQQHDAdTlWxmb3JkMRow
GAYDVQQKDBFDb21vZG8gQ0EgTltaXRIZDehMB8GA1UEAwvYQUFBIENlcnRpZmlj
YXRIIFNlcnZpY2VzMB4XDTA0MDEwMTAwMDAwMFoXDTI4MTIzMThNTk1OVowezEL
-----END CERTIFICATE-----
```

These certificates will be trusted by the deployed VM

Save

collection

This holds multiple records of a group of custom defined properties.

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

default	<p>Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code>.</p> <p>If specified, the default value must be an array where each item of the array has the keys and values which are appropriate for the property blueprints provided. For the example below, a default value could be:</p> <pre>default: - name: sam public_key: ssh-rsa AAAAB3NzaC1yc2EAAAQ... sudo: true</pre>
property_blueprints	A group of property blueprints that each record will have. Any type and combination of property types can be used except collections and selectors which cannot be nested inside of collections.
named_manifests	An array of objects each containing a <code>name</code> and a <code>manifest</code> . The manifest specifies the YAML which can be used with <code>parsed_manifest(name)</code> where the manifest will be interpolated for each record in the collection. The manifest has a special property level, <code>current_record</code> , which has all the properties in a record of the collection. The provided example below shows how <code>named_manifest</code> , <code>current_record</code> , and <code>parsed_manifest</code> together to render a collection in a BOSH manifest.

Accessors

<code>parsed_manifest(named)</code>	An array of interpolated YAML elements (often objects, but could be strings or numbers). Where each element of the array is an interpolation of the <code>named_manifest</code> for each record.
<code>value</code>	An array of objects whose keys are the property names specified under <code>property_blueprints</code> , and whose values are the <code>.value</code> for each property. For example, if a collection "albums" had properties: title (string) and explicit (boolean), <code>((albums.value))</code> would render:

Example using `.value` accessor

```
property_blueprints:  
- name: users_to_add  
  type: collection  
  configurable: true  
property_blueprints:  
- name: name  
  type: string  
  configurable: true  
- name: public_key  
  type: string  
  configurable: true  
- name: sudo  
  type: boolean  
  configurable: true  
  
form_types:  
- name: example_form  
  label: 'Example form'  
  description: 'An example form'  
property_inputs:  
- reference: .properties.users_to_add  
  label: 'Users to add'  
  description: 'will added to deployed VM'  
property_inputs:  
- reference: name  
  label: Username  
  description: username of the added user  
- reference: public_key  
  label: Public key  
  description: RSA public key used to SSH onto the VM  
- reference: sudo  
  label: make sudoer  
  description: if checked, user will be able to sudo  
  
job_types:  
- name: example-job  
  resource_label: An Example Job  
  <<: *job_type_boilerplate  
templates:  
- name: user_add  
  release: os-conf  
  manifest: |  
    users: ((.properties.users_to_add.value))
```

An example form

Users to add
will added to deployed VM

Add

▼ tom

Username *

Public key *

make sudoer

▼ sameer

Username *

Public key *

make sudoer

Save

Resulting manifest snippet

```
...
instance_groups:
- name: example-job
  jobs:
  - name: user_add
    release: os-conf
  ...
  properties:
    users:
    - name: tom
      public_key: ssh-rsa MFwwDQYJKoZIhvcNAQEBBQ...
      sudo: false
    - name: sameer
      public_key: ssh-rsa MFwwDQYJKoZIhvcNAQEBBQ...
      sudo: true
...

```

Example with `.parsed_manifest(name)`

```

property_blueprints:
- name: sysctls
  type: collection
  configurable: true
  property_blueprints:
    - name: setting_name
      type: string
      configurable: true
    - name: setting_value
      type: string
      configurable: true
  named_manifests:
    - name: name_value_format
      manifest: (( current_record.setting_name.value ))=(( current_record.setting_value.value ))

```

```

form_types:
- name: example_form
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.sysctls
      label: 'Custom sysctls'
      description: 'sysctl configuration settings to be placed in /etc/sysctl.d'
  property_inputs:
    - reference: setting_name
      label: Setting name
      description: the sysctl setting name, e.g. net.core.somaxconn
    - reference: setting_value
      label: Setting value
      description: value for this sysctl setting

```

```

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: sysctl
      release: os-conf
      manifest: |
        sysctl: (( .properties.sysctls.parsed_manifest(name_value_format) ))

```

An example form

Custom sysctls
sysctl configuration settings to be placed in /etc/sysctl.d

Add

▼ foo

Setting name *



Setting value *

▼ fooz

Setting name *



Setting value *

Save

Resulting manifest snippet

```
...
instance_groups:
- name: example-job
  jobs:
  - name: sysctl
    release: os-conf
  ...
  properties:
    sysctl:
      - foo=bar
      - fooz=baz
...
...
```

disk_type_dropdown

This holds a single string value selected from allowed disk types.

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

resource_definitions	Optional. An array of objects with the same schema as <code>resource_definitions</code> for <code>job_types</code> . For <code>disk_type_dropdown</code> , <code>resource_definitions</code> should only have one element with <code>name</code> “persistent_disk”. This element can also specify a <code>default</code> which OpsManager uses as a minimum bound when it selects a default value. It can also specify <code>constraints</code> , which can be any set of constraints that can be specified for an integer property. For examples of supported constraint types, see integer below.
----------------------	---

Accessors

value	Returns a string containing the size of the selected disk in megabytes, e.g. “1024”. If nothing or “Automatic” was specified by the operator, then this returns the smallest disk that is greater than the <code>default</code> value specified in <code>resource_definitions</code> . If no default was specified, then this returns the smallest disk in the catalog.
value_present?	Returns true only if a disk type is selected. If the user has selected “Automatic”, <code>value_present?</code> will be false.

Example

```

property_blueprints:
- name: example_disk_type_dropdown
  type: disk_type_dropdown
  configurable: true
  resource_definitions:
    - name: persistent_disk
      default: 4096
      constraints:
        min: 2048

form_types:
- name: example_form
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_disk_type_dropdown
      label: 'Example disk_type_dropdown'
      description: 'This is an example'

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: login_banner
      release: os-conf
      manifest: |
        login_banner:
          text: "Hello user, I see that your favorite disk size greater than 2 Gb is: (( .properties.example_disk_type_dropdown.value ))."

```

An example form

Example disk_type_dropdown*

Automatic: 5 GB
 ▼

This is an example

Save

domain

This holds a string value that is a valid domain (e.g. `example.com` or `x.example.com:80/a/b`).

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> .
---------	--

Accessors

value	Returns the domain as a string or null.
value_present?	Returns true only if the value is a non-empty string.

Example

```

property_blueprints:
- name: example_domain
  type: domain
  configurable: true

form_types:
- name: example_form
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_domain
      label: 'Example domain'
      description: 'This is an example'

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: login_banner
      release: os-conf
      manifest: |
        login_banner:
          text: Hello, user. I see your favorite domain is (( .properties.example_domain.value )).

```

An example form

Example domain *

Save

dropdown_select

This holds an array of strings selected string values.

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> . The default value should be one of the options' names. For example, in the example shown below, "tomato" could be used as the default.
options	An array of hashes, each item specifying a: <code>name</code> , which becomes the value of the property if this option is selected, and <code>label</code> shows up as the text in the dropdown for this option.

Accessors

value	Returns the <code>name</code> of the selected option.
value_present?	Returns true if an option is selected, and that option's name is a non-empty string.

Example

```

property_blueprints:
- name: example_dropdown_select
  type: dropdown_select
  default: kiwi
  configurable: true
  options:
    - name: kiwi
      label: 'Kiwi'
    - name: lime
      label: 'Lime'
    - name: tomato
      label: 'Tomato'

form_types:
- name: example_dropdown
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_dropdown_select
      label: Favorite fruit
      description: 'What is your favorite fruit?'

job_types:
- name: example-job
  resource_label: An Example Job
  <<< *job_type_boilerplate
  templates:
    - name: login_banner
      release: os-conf
      manifest: |
        login_banner:
          text: Hello, user. I see your favorite fruit is (( .properties.example_dropdown_select.value )).

```

An example form

Favorite fruit*

What is your favorite fruit?

Save

email

This ensures the string value is formatted as an email address.

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> .
---------	--

Accessors

value	Returns a string or null.
value_present?	Returns true if value is a non-empty string.

Example

```

property_blueprints:
- name: example_email
  type: email
  configurable: true

form_types:
- name: example_dropdown
  label: 'Example form'
  description: 'An example form'

property_inputs:
- reference: .properties.example_email
  label: User email
  description: 'What is your email?'

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
- name: login_banner
  release: os-conf
  manifest: |
    login_banner:
    text: Hello, user. Your email is (( .properties.example_email.value )).


```

An example form

User email *

What is your email?

Save

http_url

This holds an HTTP URL.

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> .
---------	--

Accessors

value	Returns a string or null.
value_present?	Returns true if value is a non-empty string.

Example

```

property_blueprints:
- name: example_http_url
  type: http_url
  configurable: true
  default: 'http://example.com'

form_types:
- name: example_form
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_http_url
      label: Example http_url
      description: 'Configure a property of type http_url'

job_types:
- name: example-job
  resource_label: An Example Job
<<< *job_type_boilerplate
templates:
- name: login_banner
  release: os-conf
  manifest: |
    login_banner:
    text: Hello, user. For fun, please visit (( .properties.example_http_url.value )).
  
```

An example form

Example http_url *

Save

integer

This holds a single integer value.

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> .
constraints	Constrains which integer values are considered valid. See the below table for all supported constraint types.

Supported Constraints

max	Sets the maximum value of the integer.
may_only_be_odd_or_zero	Sets the integer to only allow odd or zero values.
may_only_increase	Sets the integer to only increase in value.
min	Sets the minimum value of the integer.
modulo	Sets the integer to only be divisible by the given value.
power_of_two	Sets the integer to only be a power of two.
zero_or_min	Sets the minimum value of the integer but allows zero.

Accessors

value	Returns an integer or null.
value_present?	Returns true if value is an integer.

Example

```

property_blueprints:
- name: example_integer1
  type: integer
  configurable: true
  default: 9
  constraints: # the following constraints allow the values 9, 15, and 21
    min: 8
    max: 21
    may_only_increase: true
    may_only_be_odd_or_zero: true
    modulo: 3
- name: example_integer2 # the following constraints allow the values 8 and 16
  type: integer
  configurable: true
  optional: true
  constraints:
    zero_or_min: 5
    max: 20
    power_of_two: true

form_types:
- name: example_integer
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_integer1
      label: Example Integer 1
      description: 'Configure a property of type integer'
    - reference: .properties.example_integer2
      label: Example Integer 2
      description: 'Configure another property of type integer'

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: tcp_keepalive
      release: os-conf
      manifest: |
        tcp_keepalive:
          time: (( .properties.example_integer1.value ))
          interval: (( .properties.example_integer2.value ))

```

An example form

Example Integer 1 (min: 8, max: 21) *

Example Integer 2 (max: 20)

Configure another property of type integer

Save

ip_address

Ensures the string value is an IP address.

credential		no
supports auto-generation		no
supports operator configuration		yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> .
---------	--

Accessors

value	Returns a string or null.
value_present?	Returns true if value is a non-empty string.

Example

```

property_blueprints:
- name: example_ip_address
  type: ip_address
  configurable: true
  default: 192.168.0.1

form_types:
- name: example_form
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_ip_address
      label: Example IP Address
      description: 'Configure a property of type ip_address'

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: login_banner
      release: os-conf
      manifest: |
        login_banner:
          text: Hello, user. I see your favorite IP Address is (( .properties.example_ip_address.value )).


```

An example form

Example IP Address *

Configure a property of type ip_address

ip_ranges

This holds a string that is a comma-separated list of IP addresses and IP address ranges.

credential		no
supports auto-generation		no
supports operator configuration		yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> .
---------	--

Accessors

<code>parsed_ip_ranges</code>	Returns an array of strings containing each IP range
<code>value</code>	Returns a string containing a comma-separated list of IP ranges and IP addresses, or null.
<code>value_present?</code>	Returns true if value is a non-empty string.

Example

```

property_blueprints:
- name: example_ip_ranges
  type: ip_ranges
  configurable: true
  default: 1.1.1.1-1.1.1.4,2.2.2.1-2.2.2.4,3.3.3.3

form_types:
- name: example_form
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_ip_ranges
      label: Example IP Ranges
      description: 'Configure a property of type ip_ranges'

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: login_banner
      release: os-conf
      manifest: |
        login_banner:
          text: Hello, user. I see your favorite IP ranges are (( .properties.example_ip_ranges.value )).

```

An example form

Example IP Ranges *

Configure a property of type ip_ranges

ldap_url

Ensures the inputted string matches a URL of the LDAP protocol.

credential		no
supports auto-generation		no
supports operator configuration		yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> .
---------	--

Accessors

value	Returns a string or null.
value_present?	Returns true if value is a non-empty string.

Example

```

property_blueprints:
- name: example_ldap_url
  type: ldap_url
  configurable: true
  default: 'ldap://example.com'

form_types:
- name: example_form
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_ldap_url
      label: Example ldap_url
      description: 'Configure a property of type ldap_url'

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: login_banner
      release: os-conf
      manifest: |
        login_banner:
          text: Hello, user. I see your favorite LDAP URL is (( .properties.example_ldap_url.value )).

```

An example form

Example ldap_url *

Save

multi_select_options

This holds an array of selected string values. Must be an array of non-blank strings.

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> .
---------	--

Accessors

value	Returns an array of strings for the selected options or null.
value_present?	Returns true if any of the checkboxes are checked.

Example

```

property_blueprints:
- name: example_multi_select_options
  type: multi_select_options
  configurable: true
  default: ['earth', 'mercury']
  options:
    - name: mercury
      label: 'label for mercury'
    - name: venus
      label: 'label for venus'
    - name: earth
      label: 'label for earth'

form_types:
- name: example_multi_select_options
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_multi_select_options
      label: Example Multi-Select Options
      description: 'Configure a property of type multi_select_options'

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: modprobe
      release: os-conf
      manifest: |
        # the value at the 'modules' key will render as an array of strings
        modules: (( .properties.example_multi_select_options.value ))

```

An example form

Example Multi-Select Options *

label for mercury Configure a property of type multi_select_options

label for venus

label for earth

Save

network_address

This holds a string which is a hostname, domain name, or IP addresses.

credential		no
supports auto-generation		no
supports operator configuration		yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> .
---------	--

Accessors

value	Returns a string or null.
value_present?	Returns true if value is a non-empty string.

Example

```

property_blueprints:
- name: example_network_address
  type: network_address
  configurable: true
  default: 'localhost'

form_types:
- name: example_network_address
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_network_address
      label: Example Network Address
      description: 'Configure a property of type network_address'

job_types:
- name: example-job
  resource_label: An Example Job
<<: *job_type_boilerplate
templates:
- name: login_banner
  release: os-conf
  manifest: |
    login_banner:
    text: Hello, user. I see your favorite network address is (( .properties.example_network_address.value )).
  
```

An example form

Example Network Address *

 Configure a property of type network_address

Save

network_address_list

This holds a string which is a comma-separated list of hostnames and IP addresses.

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> .
---------	--

Accessors

parsed_network_addresses	Returns the list of hostnames and IP addresses as an array of strings
--------------------------	---

value	Returns a string which is a comma-separated list of hostnames and IP addresses or null.
value_present?	Returns true if value is a non-empty string.

Example

```

property_blueprints:
- name: example_network_address_list
  type: network_address_list
  configurable: true
  default: 'localhost,8.8.8.8'

form_types:
- name: example_network_address_list
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_network_address_list
      label: Example Network Address List
      description: 'Configure a property of type network_address_list'

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: login_banner
      release: os-conf
      manifest: |
        login_banner:
          text: Hello, user. I see your favorite network addresses are (( .properties.example_network_address_list.value )).
          an_ignored_array_of_network_addresses: (( .properties.example_network_address_list.parsed_network_addresses ))

```

An example form

Example Network Address List *

localhost,8.8.8.8

Save

port

This holds a single integer value. Allowed values are 0 through 65535 and null.

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> .
---------	--

Accessors

value	Returns an integer or null.
value_present?	Returns true if value is not null.

Example

```

property_blueprints:
- name: example_port
  type: port
  configurable: true
  default: 3000

form_types:
- name: example_port
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_port
      label: Example Port
      description: 'Configure a property of type port'

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: login_banner
      release: os-conf
      manifest: |
        login_banner:
          text: Hello, user. I see your favorite network port is (( .properties.example_port.value )).

```

An example form

Example Port *

Configure a property of type port

Save

rsa_cert_credentials

This holds SSL certificate generated from root CA

credential	yes
supports auto-generation	yes
supports operator configuration	yes

Property Blueprint Attributes

default	For properties which use <code>configurable: false</code> , the default can specify a list of domain names to use when auto-generating the RSA certificate. The first domain listed will be the value of the Common Name field of the cert. The full list of domains will be the value of the Alternative Names field of the cert. See below for an example.
---------	--

Accessors

cert_pem	Returns a string
private_key_pem	Returns a string
cert_and_private_key_pems	Returns a string
public_key_pem	Returns a string
value_present?	Returns true if value is an RSA certificate. Returns false if the value is null.

Example

```

property_blueprints:
# This is configurable by the operator and appears in the form_types section
- name: example_configurable_rsa_cert_credentials
  type: rsa_cert_credentials
  configurable: true
  optional: true
# This is not configurable by the operator and is auto-generated
- name: example_non_configurable_rsa_cert_credentials
  type: rsa_cert_credentials
  configurable: false
  default:
    domains:
      - 'cell.service.cf.internal'
      - '*cell.service.cf.internal'
form_types:
- name: example_form
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_configurable_rsa_cert_credentials
      label: Example configurable rsa_cert_credentials
      description: 'Configure a property of type rsa_cert_credentials'

job_types:
- name: example-job
  resource_label: An Example Job
<<: *job_type_boilerplate
templates:
- name: login_banner
  release: os-conf
  manifest: |
    login_banner:
    text: |
      Hello, user.
      Here is your public_key_pem
      (( .properties.example_non_configurable_rsa_cert_credentials.public_key_pem ))
      Here is your private_key_pem
      (( .properties.example_non_configurable_rsa_cert_credentials.private_key_pem ))
      Here is your cert_pem
      (( .properties.example_non_configurable_rsa_cert_credentials.cert_pem ))
      Here is your cert_and_private_key_pems
      (( .properties.example_non_configurable_rsa_cert_credentials.cert_and_private_key_pems ))

```

An example form

Example configurable rsa_cert_credentials

Certificate PEM

Private Key PEM

[Generate RSA Certificate](#)

Configure a property of type rsa_cert_credentials

[Save](#)

rsa_pkey_credentials

This holds RSA public and private keys. This is a non-configurable property.

credential	yes
supports auto-generation	yes
supports operator configuration	no

Accessors

public_key_pem	Returns a string
private_key_pem	Returns a string
public_key_fingerprint	Returns a string
public_key_openssh	Returns a string
value_present?	Returns true if value of the pkey already exists.

Example

```
property_blueprints:
- name: example_rsa_pkey_credentials
  type: rsa_pkey_credentials
  configurable: false

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: user_add
      release: os-conf
      manifest: |
        users:
          - name: nick
            public_key: ((.properties.example_rsa_pkey_credentials.public_key_pem))
```

salted_credentials

This holds credentials with salted and hashed passwords. This is a non-configurable property.

credential	yes
supports auto-generation	yes
supports operator configuration	no

Property Blueprint Attributes

default	<p>Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code>.</p> <p>If specified, the default value must be an object which is allowed to have the key <code>identity</code>. For example, a default value could be:</p> <pre>default: - identity: ryan</pre>
---------	---

Accessors

identity	Returns a string
password	Returns a string
salt	Returns a string
sha512_hashed_password	Returns a string
value_present?	Returns true if value of the credential has been generated already

Example

```
property_blueprints:
# Example without defaults
- name: example_salt_credentials
  type: salted_credentials
# Example with default identity
- name: example_salt_credentials_with_default
  type: salted_credentials
  default:
    identity: vcap

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: user_add
      release: os-conf
      manifest: |
        users:
          - name: vcap
            encrypted_password: ((.properties.example_salt_credentials.sha512_hashed_password))
```

secret

This holds a single string value for a secret. Adding `display_type: text_area` to the property inputs creates a multi-line credential.

credential	yes
supports auto-generation	yes
supports operator configuration	yes

Property Blueprint Attributes

constraints.must_match_regex	A regular expression that the user input must match. Create a validator that runs on the form save event. If the user input does not match the <code>must_match_regex</code> constraint, the form displays the specified <code>error_message</code> . Multiple <code>must_match_regex</code> constraints for a single property blueprint are evaluated in the order listed. See below for an example.
------------------------------	---

Accessors

value	Returns the secret as a string or null.
value_present?	Returns true if value is a non-empty string.

Example

```

property_blueprints:
- name: example_secret
  type: secret
  configurable: true
  constraints:
    - must_match_regex: '^[a-zA-Z]{20,}$'
      error_message: 'This password must be at least 20 alphabetic characters'

form_types:
- name: example_secret
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_secret
      label: Example Secret
      description: 'Configure a property of type secret'

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: login_banner
      release: os-conf
      manifest: |
        login_banner:
          text: Shh, today's secret is ((.properties.example_secret.value)).

```

An example form

Example Secret*

Configure a property of type secret

Save

Example of a multi-line credential

```
form_types:
- name: example_secret
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_secret
      label: Example Secret
      description: 'Configure a property of type secret'
      display_type: text_area
```

An example form

Example Secret*

`Secret`

Configure a property of type secret

SAVE

selector

Provides the ability to switch between groups of properties. Shows in the UI as a radio button group, potentially with additional options nested under each radio button.

Selectors are unique in the way that property information is accessed. Ops Manager provides accessors available at the top-level selector property, accessors for retrieving a specific property in an option group, and the ability to provide manifest snippets for a selector option group.

Each selector group may provide manifest snippets. This is because Ops Manager does not support conditionally adding manifest snippets. A manifest snippet should be present within all option groups, and can be used to create a simple kind of branching logic in manifest generation. Only one of these sets is evaluated and inserted into the job's manifest.

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> . If specified, the default value should be one of the values of <code>select_value</code> .
option_templates	An array of objects which defines the options available. See example below.

Accessors

<code>selected_option.parsed_manifest(manifest_snippet_name)</code>	Returns a parsed named manifest snippet for the currently selected option. Example: <code>.properties.example_selector.selected_option.parsed_manifest(m y_snippet)</code>
<code>SPECIFIC_SELECTOR_OPTION_GROUP.OPTION_GROUP_PROPERTY.value</code>	Scopes the accessor to a specific selector option group. Does not return meaningful information alone. Must be followed with the name and accessor of a specific property in the option group.

	<p>Example:</p> <pre>.properties.example_selector.filet_mignon_option.review.value</pre>
value	Returns the <code>select_value</code> of the selected option. For example, the selector shown in the example below might return "Filet Mignon" as its value.

Example

```

property_blueprints:
- name: example_selector
  type: selector
  configurable: true
  default: Pizza
  freeze_on_deploy: true
  option_templates:
    - name: pizza_option
      select_value: Pizza
      named_manifests:
        - name: my_snippet
          manifest: |
            pizza_toppings:
              pepperoni: (( .properties.example_selector.pizza_option.pepperoni.value ))
              other: (( .properties.example_selector.pizza_option.other_toppings.value ))
  property_blueprints:
    - name: pepperoni
      type: boolean
      configurable: true
      freeze_on_deploy: true
    - name: other_toppings
      type: string
      configurable: true
      optional: true
    - name: filet_mignon_option
      select_value: Filet Mignon
      named_manifests:
        - name: my_snippet
          manifest: |
            rarity: (( .properties.example_selector.filet_mignon_option.rarity_dropdown.value ))
            review: (( .properties.example_selector.filet_mignon_option.review.value ))
  property_blueprints:
    - name: rarity_dropdown
      type: dropdown_select
      configurable: true
      default: rare
      options:
        - name: rare
          label: 'Rare'
        - name: medium
          label: 'Medium'
        - name: well-done
          label: 'Well done'
    - name: review
      type: string
      configurable: true
      default: A+ power seller of mail order steak
      optional: false

form_types:
- name: example_selector_form
  label: 'Selector Example'
  description: 'A selector example form'
  property_inputs:
    - reference: .properties.example_selector
      label: 'Food Choices'
      selector_property_inputs:
        - reference: .properties.example_selector.pizza_option
          label: 'Pizza'
          property_inputs:
            - reference: .properties.example_selector.pizza_option.pepperoni
              label: 'Add Pepperoni'
              description: 'Yum!'
            - reference: .properties.example_selector.pizza_option.other_toppings
              label: 'Other toppings'
    - reference: .properties.example_selector.filet_mignon_option
      label: 'Filet Mignon'
      property_inputs:
        - reference: .properties.example_selector.filet_mignon_option.rarity_dropdown
          label: 'How rare?'
          description: 'Extinct.'

```

```

- reference: .properties.example_selector.filet_mignon_option.review
  label: 'Write your review'
  description: "E.g. Would not buy again"

job_types:
- name: example-job
  resource_label: An Example Job
<<: *job_type_boilerplate
templates:
- name: login_banner
  release: os-conf
  manifest: |
    login_banner:
    text: |
      Would you like some (( .properties.example_selector.value ))?
      I'm asking because you mentioned (( .properties.example_selector.pizza_option.other_toppings.value ))
      more_manifest_ignored_by_os_conf: (( .properties.example_selector.selected_option.parsed_manifest(my_snippet) ))

```

A selector example form

Food Choices*

Pizza

Add Pepperoni

Other toppings

Filet Mignon

Save

service_network_az_multi_select

This holds an arrays of string value selected from allowed availability zones (AZs).

credential	no
supports auto-generation	no
supports operator configuration	yes

Accessors

value	Returns an array of strings for the selected options.
value_present?	Returns true if value is set.

Example

```

service_broker: true

property_blueprints:
- name: service_instance_azs
  type: service_network_az_multi_select
  configurable: true
  optional: false

form_types:
- name: example_service_network_az_multi_select
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.service_instance_azs
      label: Example Service Network AZ Multi-Select
      description: 'Configure a property of type service_network_az_multi_select'

job_types:
- name: example-job
  resource_label: An Example Job
<<: *job_type_boilerplate
templates:
- name: login_banner
  release: os-conf
  manifest: |
  login_banner:
    text: "Hello, user. Here is an array of AZ names: (( .properties.service_instance_azs.value ))"

```

An example form

Example Service Network AZ Multi-Select *

us-central1-a Configure a property of type service_network_az_multi_select

Save

service_network_az_single_select

This holds a single string value selected from allowed azs. Appears as a radio button group in the UI.

credential		no
supports auto-generation		no
supports operator configuration		yes

Accessors

value	Returns a string or null.
value_present?	Returns true if value is a non-empty string

Example

```

service_broker: true

property_blueprints:
- name: example_service_network_az_single_select
  type: service_network_az_single_select
  configurable: true
  optional: false

form_types:
- name: example_service_network_az_single_select
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_service_network_az_single_select
      label: Example Service Network AZ Single-Select
      description: 'Configure a property of type service_network_az_single_select'

job_types:
- name: example-job
  resource_label: An Example Job
<<: *job_type_boilerplate
templates:
- name: login_banner
  release: os-conf
  manifest: |
    login_banner:
      text: "Hello, user. Here is a single AZ name: (( .properties.example_service_network_az_single_select.value ))"

```

An example form

Example Service Network AZ Single-Select *

us-central1-a Configure a property of type service_network_az_single_select

Save

simple_credentials

This holds an identity and password.

credential	yes
supports auto-generation	yes
supports operator configuration	yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> .
	If specified, the default value must be an object which is allowed to have the key <code>identity</code> . For example, a default value could be:
	<pre> default: - identity: ryan </pre>

Accessors

--	--

identity	Returns a string
password	Returns a string
value_present?	Returns true if value is a non-empty string.

Example

```

property_blueprints:
- name: example_configurable_simple_credentials
  type: simple_credentials
  configurable: true
- name: example_non_configurable_simple_credentials
  type: simple_credentials
  configurable: false
  default:
    identity: vcap
form_types:
- name: example_form
  label: 'Example form'
  description: 'An example form'
property_inputs:
- reference: .properties.example_configurable_simple_credentials
  label: Example configurable simple_credentials
  description: 'Configure a property of type simple_credentials'

job_types:
- name: example-job
  resource_label: An Example Job
<<< *job_type_boilerplate
templates:
- name: login_banner
  release: os-conf
  manifest: |
    login_banner:
    text: |
      Shh, today's secret password is (( .properties.example_non_configurable_simple_credentials.password ))
      And the username is (( .properties.example_non_configurable_simple_credentials.identity ))

```

An example form

Example configurable simple_credentials *

ryan

Change

[Save](#)

smtp_authentication

This holds string with a possible value of plain, login, or cram_md5.

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

`default` | Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has `optional: true`.

Accessors

value	Returns a string with possible value of <code>plain</code> , <code>login</code> , <code>cram_md5</code> .
value_present?	Returns true if value is a non-empty string.

Example

```
property_blueprints:
- name: example_smtp_authentication
  type: smtp_authentication
  configurable: true
  default: plain

form_types:
- name: example_text
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_smtp_authentication
      label: Example SMTP Authentication
      description: 'Configure a property of type smtp_authentication'

job_types:
- name: example-job
  resource_label: An Example Job
<<: *job_type_boilerplate
templates:
- name: login_banner
  release: os-conf
  manifest: |
  login_banner:
    text: Hello. The SMTP auth type is (( .properties.example_smtp_authentication.value ))
```

An example form

Example SMTP Authentication *

Save

stemcell_selector

Provides the ability to select between all available stemcell criteria. The stemcell selector is automatically populated with a list of all stemcell operating systems that are listed in the `stemcell_criteria` and `additional_stemcells_criteria` top-level properties.

After an operator has selected a stemcell listed in `stemcell_selector`, that stemcell is required to deploy the product.

For more information about the `stemcell_criteria` top-level property, see [stemcell_criteria](#) above.

For more information about the `additional_stemcells_criteria` top-level property, see [additional_stemcells_criteria](#) above.

credential	no
supports auto-generation	no

`supports operator configuration`

yes

Property Blueprint Attributes

default	<p>Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code>.</p> <p>If specified, the default value should be the operating system of a stemcell criteria.</p>
---------	--

Accessors

os	The operating system of the stemcell that is assigned to the product. This is not always the same as <code>value</code> . If a stemcell is selected but not assigned, this is set to <code>nil</code> .
version	The version of the stemcell that is assigned to the product. If a stemcell is selected but not assigned, this is set to <code>nil</code> .
assigned?	Boolean. Returns true if there is an assigned stemcell that matches the selected stemcell.
value	Returns the selected operating system.
value_present?	Returns true if the operator has selected an operating system.

Example

```

stemcell_criteria:
  os: ubuntu-xenial
  version: 1.0.0
additional_stemcells_criteria:
  os: windows2016
  version: 1.0.0

property_blueprints:
  - name: example_stemcell_selector
    type: stemcell_selector
    configurable: true
    optional: true
    form_types:
      - name: example_stemcell_selector_form
        label: 'Stemcell Selector Example'
        description: 'A stemcell selector example form'
    property_inputs:
      - reference: .properties.example_stemcell_selector
        label: 'Example stemcell selector'
        description: 'Select which stemcell you want to use'

job_types:
  - name: example-job
    resource_label: An Example Job
    <<: *job_type_boilerplate
    templates:
      - name: login_banner
        release: os-conf
        manifest: |
          login_banner:
            text: |
              Stemcell Assigned?: (( .properties.example_stemcell_selector.assigned? ))
              Assigned stemcell OS: (( .properties.example_stemcell_selector.os ))
              Assigned stemcell version: (( .properties.example_stemcell_selector.version ))
              Selected stemcell?: (( .properties.example_stemcell_selector.value_present? ))
              Selected stemcell OS: (( .properties.example_stemcell_selector.value ))
```

A stemcell selector example form

Example stemcell selector

- ubuntu-xenial
- windows2016

Select which stemcell you want to use

Save

string_list

This holds an array of strings.

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> .
---------	--

Accessors

parsed_strings	Returns an array of strings for each string entry
parsed_regex	Returns a string containing a regex of the format <code>^(string1 string2 string3)\$</code> where the value of this property is <code>string1,string2,string3</code>
value	Returns a string or null.
value_present?	Returns true if value is a non-empty string.

Example

```

property_blueprints:
- name: example_string_list
  type: string_list
  configurable: true
  default: foo,bar,baz

form_types:
- name: example_string_list
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_string_list
      label: Example String List
      description: 'Configure a property of type string_list'

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: login_banner
      release: os-conf
      manifest: |
        login_banner:
          text: Hello. Your array of strings is (( .properties.example_string_list.value ))

```

An example form

Example String List *

Configure a property of type string_list

string

This holds a single string value.

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

constraints.must_match_regex	A regular expression that the user input must match. Create a validator that runs on the form save event. If the user input does not match the <code>must_match_regex</code> constraint, the form displays the specified <code>error_message</code> . Multiple <code>must_match_regex</code> constraints for a single property blueprint are evaluated in the order listed. See below for an example.
------------------------------	---

Accessors

value	Returns the string value or null.
value_present?	Returns true only if the value is a non-empty string.

Example

```

property_blueprints:
- name: example_string
  type: string
  configurable: true
  default: 'Karl'
  constraints:
    - must_match_regex: 'A[^!@#$%^&*()]*z'
      error_message: 'This name cannot contain special characters.'
    - must_match_regex: 'A[^0-9]*z'
      error_message: 'This name cannot contain digits.'

form_types:
- name: example_form
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_string
      label: 'Example string'
      description: 'This is an example'

job_types:
- name: example-job
  resource_label: An Example Job
<<< *job_type_boilerplate
templates:
- name: login_banner
  release: os-conf
  manifest: |
    login_banner:
      text: Hello, {{ .properties.example_string.value }}! Welcome.

```

An example form

Example string *

Save

text

This holds a string value. For configurable properties of this type, the UI allows multi-line strings to be entered.

credential		no
supports auto-generation		no
supports operator configuration		yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> .
---------	--

Accessors

value	Returns a string or null.
-------	---------------------------

<code>value_present?</code>	Returns true if value is a non-empty string.
-----------------------------	--

Example

```

property_blueprints:
- name: example_text
  type: text
  configurable: true
  default: |
    Example
    Text

form_types:
- name: example_text
  label: 'Example form'
  description: 'An example form'
  property_inputs:
  - reference: .properties.example_text
    label: Example Text
    description: 'Configure a property of type text'

job_types:
- name: example-job
  resource_label: An Example Job
<<: *job_type_boilerplate
templates:
- name: login_banner
  release: os-conf
  manifest: |
    login_banner:
    text: |
      Hello. Check out all this text!
      (( .properties.example_text.value ))
  
```

An example form

Example Text *

Example
Text

Save

uuid

This holds a string uuid value.

credential	no
supports auto-generation	yes
supports operator configuration	no

Accessors

<code>value</code>	Returns a string or null.
<code>value_present?</code>	Returns true if value is a non-empty string.

Example

```

property_blueprints:
- name: example_uuid
  type: uuid
  configurable: false

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: login_banner
      release: os-conf
      manifest: |
        login-banner:
          text: Hello. Your UUID is (( .properties.example_uuid.value ))

```

vm_type_dropdown

This holds single string value selected from allowed vm_types.

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

resource_definitions	Optional. An array of objects with the same schema as <code>resource_definitions</code> for <code>job_types</code> . For <code>vm_type_dropdown</code> , <code>resource_definitions</code> can have elements with <code>name</code> “ram”, “cpu”, or “ephemeral_disk”. This element can also specify a <code>default</code> which OpsManager uses as a minimum bound when it selects a default value. It can also specify <code>constraints</code> , which can be any set of constraints that can be specified for an integer property. For examples of supported constraint types, see integer above.
----------------------	--

Accessors

value	Returns a string which is the name of the selected VM type.
-------	---

Example

```

property_blueprints:
- name: example_vm_type_dropdown
  type: vm_type_dropdown
  configurable: true
  resource_definitions:
    - name: ram
      default: 4096
    - name: cpu
      default: 1
    - name: ephemeral_disk
      default: 4096
  constraints:
    min: 2048

form_types:
- name: example_vm_type_dropdown
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_vm_type_dropdown
      label: Example VM Type Dropdown
      description: 'Configure a property of type vm_type_dropdown'

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: login_banner
      release: os-conf
      manifest: |
        login_banner:
          text: Hello. Your chosen VM Type is (( .properties.example_vm_type_dropdown.value )).

```

An example form

Example VM Type Dropdown*

Configure a property of type vm_type_dropdown

Save

wildcard_domain

Ensures the string value is a domain prefixed with `*`.

credential	no
supports auto-generation	no
supports operator configuration	yes

Property Blueprint Attributes

default	Optional. Specifies a default value. When provided, must be set to a value which is valid for this type. Cannot be used when the property also has <code>optional: true</code> .
---------	--

Accessors

to_wildcard	Returns a string of the value prefixed with <code>*</code> if not present
-------------	---

value	Returns a string or null.
value_present?	Returns true if value is a non-empty string.

Example

```

property_blueprints:
- name: example_wildcard_domain
  type: wildcard_domain
  configurable: true
  default: 'example.com'

form_types:
- name: example_form
  label: 'Example form'
  description: 'An example form'
  property_inputs:
    - reference: .properties.example_wildcard_domain
      label: Example Wildcard Domain
      description: 'Configure a property of type wildcard_domain'

job_types:
- name: example-job
  resource_label: An Example Job
  <<: *job_type_boilerplate
  templates:
    - name: login_banner
      release: os-conf
      manifest: |
        login_banner:
          text: Hello. Your wildcard domain is (( .properties.example_wildcard_domain.value )).

```

An example form

Example Wildcard Domain *

Configure a property of type wildcard_domain

Save

Using the Examples Provided

You can copy each example into the `metadata.yml` template below and package it together along with the `os-conf` release as a `.pivotal` file to make a working tile.

```
---
name: p-example
label: Example Tile
description: An example tile
product_version: '1.1'
minimum_version_for_upgrade: '1.0'
metadata_version: '2.2'
rank: 1
stemcell_criteria:
  os: ubuntu-xenial
  version: '97.3'
releases:
- name: os-conf
  file: os-conf
  version: '15'
# a base 64 encoded PNG with a teal square
icon_image: iVBORw0KGgoAAAANSUhEUgAAAAEAAAABCIAAACQdIPeAAADEIEQVQI2NgaC8BAAGGAPwUz8ygAAAAAEIFTkSuQmCC

includes: # This top-level key will be ignored by OpsManager.
# This anchor is only to make the examples below more terse, and is not required.
job_type_boilerplate: &job_type_boilerplate
max_in_flight: 1
single_az_only: true
resource_definitions:
- name: ram
  configurable: true
  default: 1024
- name: ephemeral_disk
  configurable: true
  default: 1024
- name: persistent_disk
  configurable: true
  default: 1024
  constraints:
    min: 1024
- name: cpu
  configurable: true
  default: 1
instance_definition:
  name: instances
  configurable: true
  default: 1
```

PAS Concepts

Pivotal Application Service (PAS) is based on Cloud Foundry, which is an open source cloud application platform, providing a choice of clouds, developer frameworks, and application services. Cloud Foundry makes it faster and easier to build, test, deploy, and scale applications. It is an [open source project](#) and is available through a variety of private cloud distributions and public cloud instances.

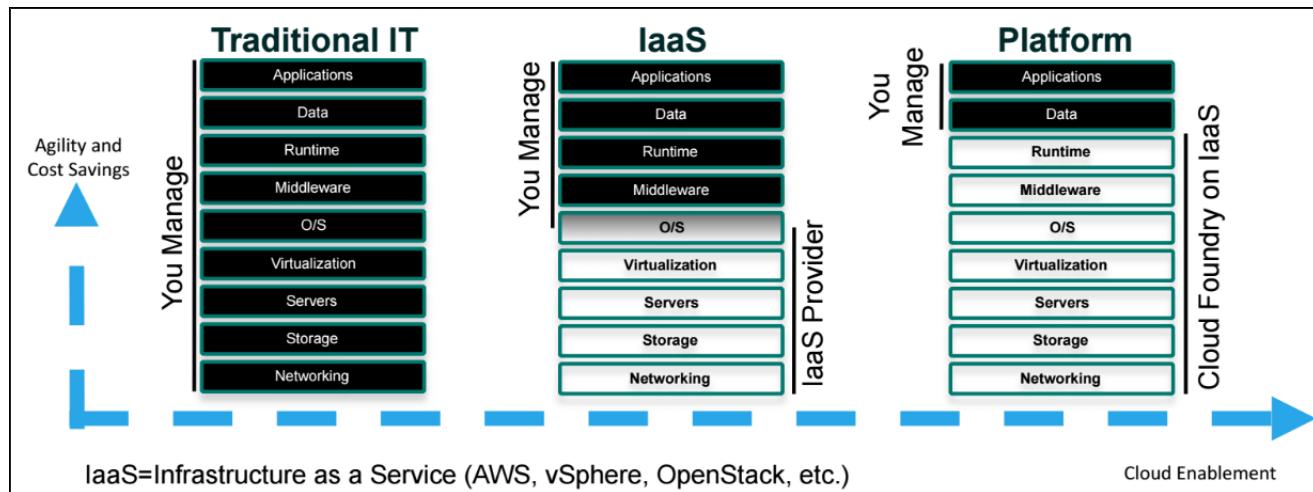
This documentation presents an overview of how Cloud Foundry works and a discussion of key concepts. See the following topics to learn more about Cloud Foundry fundamentals.

- [Cloud Foundry Overview](#)
- [Security](#)
- [High Availability](#)
- [Running Apps in the Cloud](#)
- [Runtime Components](#)
- [Container-to-Container Networking](#)

Cloud Foundry Overview

The Industry-Standard Cloud Platform

Cloud platforms let anyone deploy network apps or services and make them available to the world in a few minutes. When an app becomes popular, the cloud easily scales it to handle more traffic, replacing with a few keystrokes the build-out and migration efforts that once took months. Cloud platforms represent the next step in the evolution of IT, enabling you to focus exclusively on your applications and data without worrying about underlying infrastructure.



Not all cloud platforms are created equal. Some have limited language and framework support, lack key app services, or restrict deployment to a single cloud. Cloud Foundry (CF) has become the industry standard. It is an [open source](#) platform that you can deploy to run your apps on your own computing infrastructure, or deploy on an IaaS like AWS, vSphere, or OpenStack. You can also use a PaaS deployed by a commercial [CF cloud provider](#). A broad [community](#) contributes to and supports Cloud Foundry. The platform's openness and extensibility prevent its users from being locked into a single framework, set of app services, or cloud.

Cloud Foundry is ideal for anyone interested in removing the cost and complexity of configuring infrastructure for their apps. Developers can deploy their apps to Cloud Foundry using their existing tools and with zero modification to their code.

How Cloud Foundry Works

To flexibly serve and scale apps online, Cloud Foundry has subsystems that perform specialized functions. Here's how some of these main subsystems work.

How the Cloud Balances Its Load

Clouds balance their processing loads over multiple machines, optimizing for efficiency and resilience against point failure. A Cloud Foundry installation accomplishes this at three levels:

1. [BOSH](#) creates and deploys virtual machines (VMs) on top of a physical computing infrastructure, and deploys and runs Cloud Foundry on top of this cloud. To configure the deployment, BOSH follows a manifest document.
2. The CF [Cloud Controller](#) runs the apps and other processes on the cloud's VMs, balancing demand and managing app lifecycles.
3. The [router](#) routes incoming traffic from the world to the VMs that are running the apps that the traffic demands, usually working with a customer-provided load balancer.

How Apps Run Anywhere

Cloud Foundry designates two types of VMs: the component VMs that constitute the platform's infrastructure, and the host VMs that host apps for the outside world. Within CF, the Diego system distributes the hosted app load over all of the host VMs, and keeps it running and balanced through demand

surges, outages, or other changes. Diego accomplishes this through an auction algorithm.

To meet demand, multiple host VMs run duplicate instances of the same app. This means that apps must be portable. Cloud Foundry distributes app source code to VMs with everything the VMs need to compile and run the apps locally. This includes the OS stack that the app runs on and a buildpack containing all languages, libraries, and services that the app uses. Before sending an app to a VM, the Cloud Controller stages it for delivery by combining stack, buildpack, and source code into a droplet that the VM can unpack, compile, and run. For simple, standalone apps with no dynamic pointers, the droplet can contain a pre-compiled executable instead of source code, language, and libraries.

How CF Organizes Users and Workspaces

CF manages user accounts through two [User Authentication and Authorization](#) (UAA) servers, which support access control as [OAuth2](#) services and can store user information internally, or connect to external user stores through LDAP or SAML.

One UAA server grants access to BOSH, and holds accounts for the CF operators who deploy runtimes, services, and other software onto the BOSH layer directly. The other UAA server controls access to the Cloud Controller, and determines who can tell it to do what. The Cloud Controller UAA defines different user roles, such as admin, developer, or auditor, and grants them different sets of privileges to run CF commands. The Cloud Controller UAA also scopes the roles to separate, compartmentalized [Orgs and Spaces](#) within an installation, to manage and track use.

Where CF Stores Resources

Cloud Foundry uses the git system on [GitHub](#) to version-control source code, buildpacks, documentation, and other resources. Developers on the platform also use GitHub for their own apps, custom configurations, and other resources. To store large binary files, such as droplets, CF maintains an internal or external blobstore. CF uses MySQL to store and share temporary information, such as internal component states.

How CF Components Communicate

Cloud Foundry components communicate in two of the following ways:

- By sending messages internally using HTTP and HTTPS protocols
- By sending [NATS](#) messages to each other directly

BOSH Director collocates a BOSH DNS server on every deployed VM. All VMs keep up-to-date DNS records for all the other VMs in the same foundation. This enables service discovery between VMs.

BOSH DNS allows deployments to continue communicating with VMs even when the VMs' IP addresses change. It also provides client-side load-balancing by randomly selecting a healthy VM when multiple VMs are available.

For more information about BOSH DNS, see [Native DNS Support](#) in the BOSH documentation.

How to Monitor and Analyze a CF Deployment

Cloud Foundry generates system logs from Cloud Foundry components and app logs from hosted apps

As Cloud Foundry runs, its component and host VMs generate logs and metrics. Cloud Foundry apps also typically generate logs. The [Loggregator](#) system aggregates the component metrics and app logs into a structured, usable form, the *Firehose*. You can use all of the output of the Firehose, or direct the output to specific uses, such as monitoring system internals, triggering alerts, or analyzing user behavior, by applying *nozzles*.

The component logs follow a different path. They stream from rsyslog agents, and the cloud operator can configure them to stream out to a syslog drain.

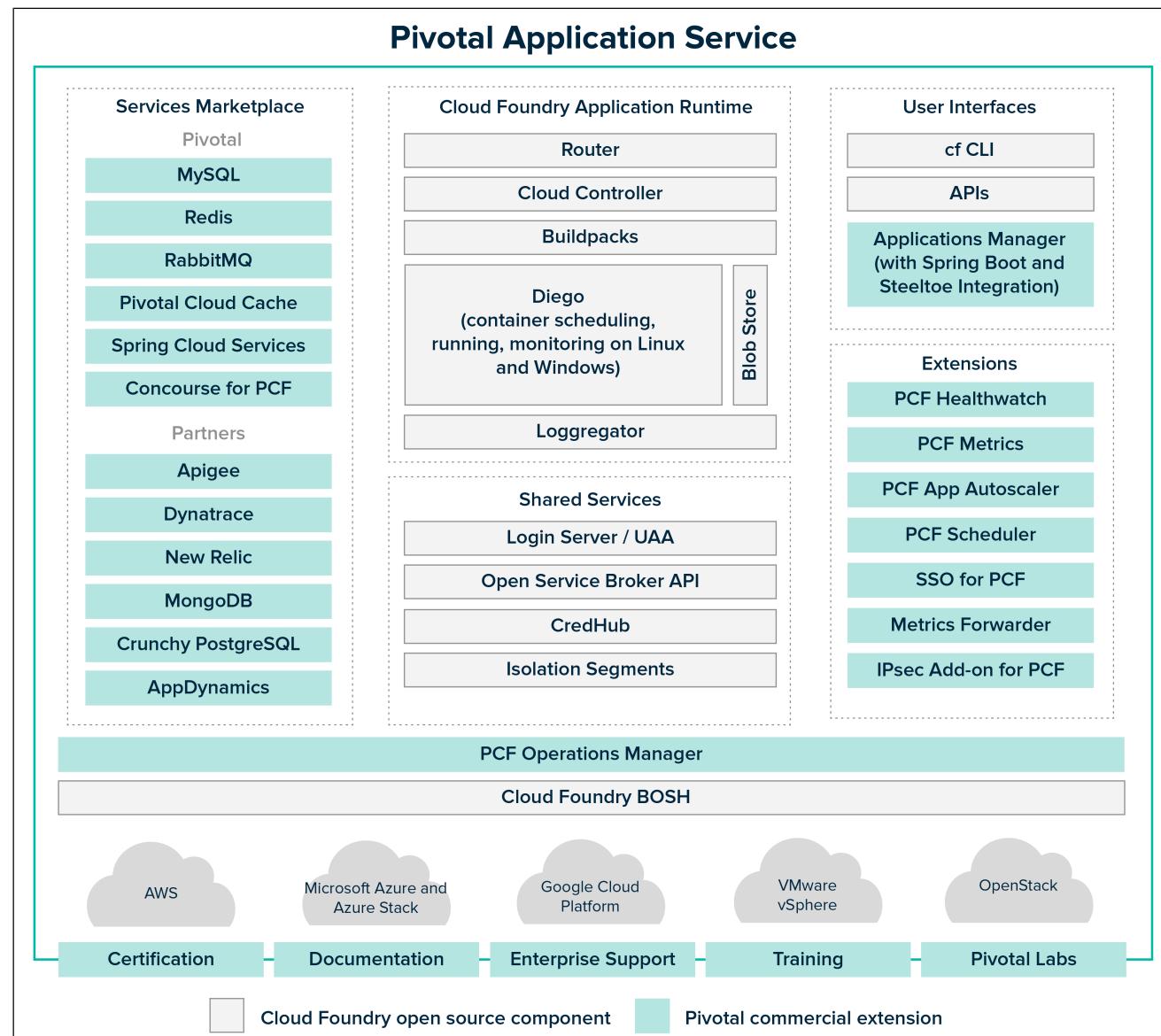
Using Services with CF

Typical apps depend on free or metered [services](#) such as databases or third-party APIs. To incorporate these into an app, a developer writes a Service Broker, an API that publishes to the Cloud Controller the ability to list service offerings, provision the service, and enable apps to make calls out to it.

How Pivotal Cloud Foundry Differs from Open Source Cloud Foundry

Open source software provides the basis for the Pivotal Cloud Foundry platform. Pivotal Application Service (PAS) is the Pivotal distribution of Cloud

Foundry software for hosting apps. Pivotal offers additional commercial features, enterprise services, support, documentation, certificates, and others value-adds.



Security

The following topics provide information about security:

- [Cloud Foundry Security](#)
- [Container Security](#)
- [Orgs, Spaces, Roles, and Permissions](#)
- [Application Security Groups](#)
- [Application SSH Components and Processes](#)

Cloud Foundry Security

This topic provides an overview of Cloud Foundry (CF) security. For an overview of container security, see the [Container Security](#) topic.

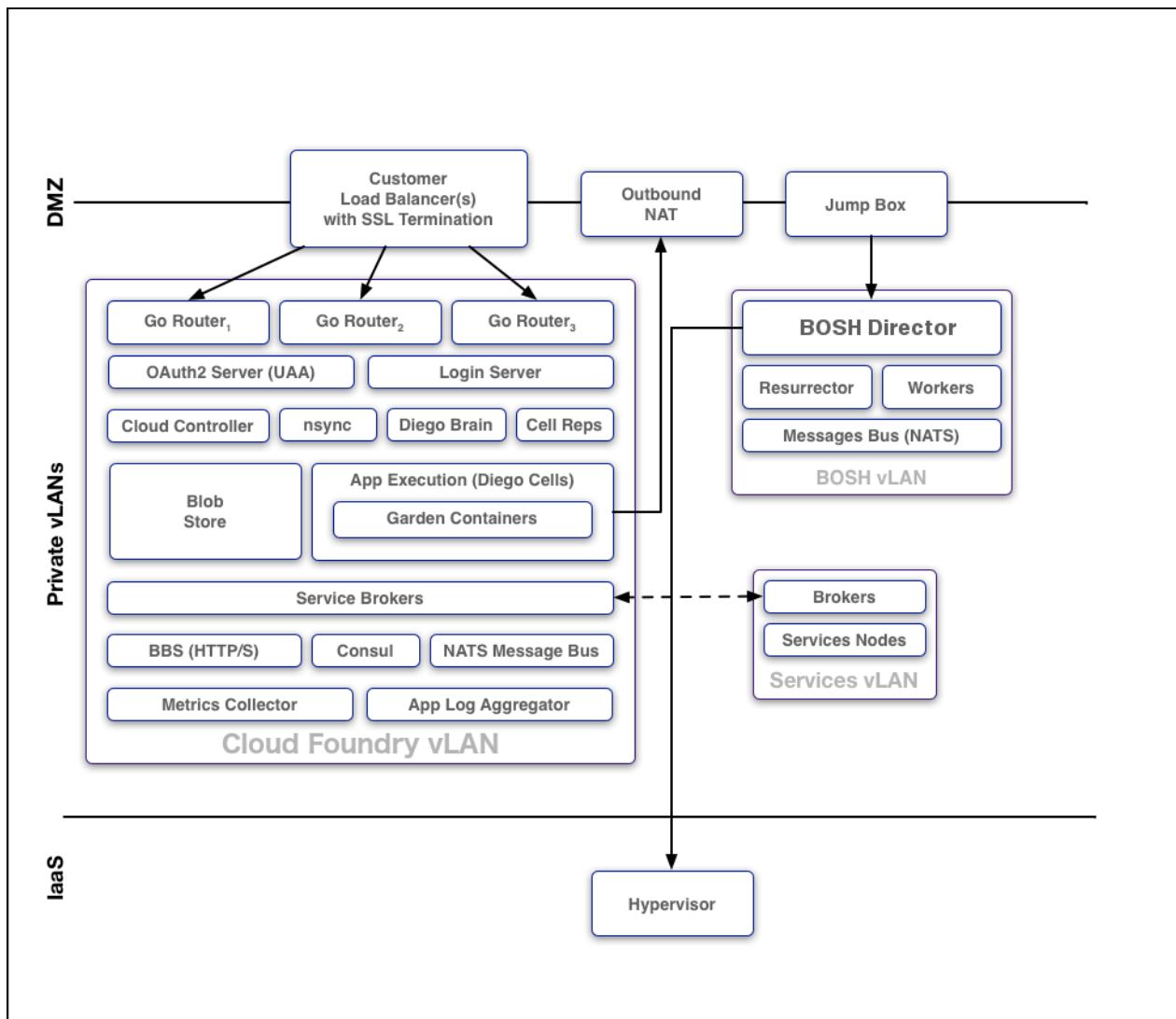
Cloud Foundry implements the following measures to mitigate against security threats:

- Minimizes network surface area
- Isolates customer applications and data in containers
- Encrypts connections
- Uses role-based access controls, applying and enforcing roles and permissions to ensure that users can only view and affect the spaces for which they have been granted access
- Ensures security of application bits in a multi-tenant environment
- Prevents possible denial of service attacks through resource starvation

System Boundaries and Access

As the image below shows, in a typical deployment of Cloud Foundry, the components run on virtual machines (VMs) that exist within a VLAN. In this configuration, the only access points visible on a public network are a load balancer that maps to one or more Cloud Foundry routers and, optionally, a NAT VM and a jumpbox. Because of the limited number of contact points with the public internet, the surface area for possible security vulnerabilities is minimized.

 **Note:** Pivotal recommends that you also install a NAT VM for outbound requests and a Jumpbox to access the BOSH Director, though these access points are optional depending on your network configuration.



Protocols

All traffic from the public internet to the Cloud Controller and UAA happens over HTTPS. Inside the boundary of the system, components communicate over a publish-subscribe (pub-sub) message bus NATS, HTTP, and SSL/TLS.

BOSH

Operators deploy Cloud Foundry with BOSH. The BOSH Director is the core orchestrating component in BOSH: it controls VM creation and deployment, as well as other software and service lifecycle events. You use HTTPS to ensure secure communication to the BOSH Director.

Note: Pivotal recommends that you deploy the BOSH Director on a subnet that is not publicly accessible, and access the BOSH Director from a Jumpbox on the subnet or through VPN.

BOSH includes the following functionality for security:

- Communicates with the VMs it launches over NATS. Because NATS cannot be accessed from outside Cloud Foundry, this ensures that published messages can only originate from a component within your deployment.
- Provides an audit trail through the `bosh tasks --all` and `bosh tasks --recent=VALUE` commands. `bosh tasks --all` returns a table that shows all BOSH actions taken by an operator or other running processes. `bosh tasks --recent=VALUE` returns a table of recent tasks, with `VALUE` being the number of recent tasks you want to view.
- Allows you to set up individual login accounts for each operator. BOSH operators have root access.

Note: BOSH does not encrypt data stored on BOSH VMs. Your IaaS might encrypt this data.

Isolation Segments

Isolation segments provide dedicated pools of resources to which apps can be deployed to isolate workloads. Using isolation segments separates app resources as completely as if they were in different CF deployments but avoids redundant management components and unneeded network complexity.

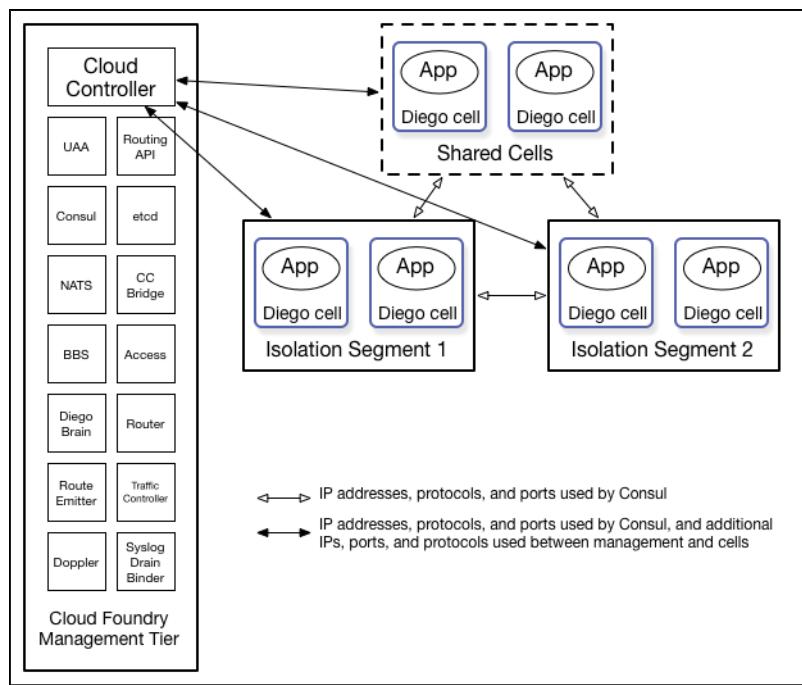
You can designate isolation segments for exclusive use by [orgs and spaces](#) within CF. This guarantees that apps within the org or space use resources that are not also used by other orgs or spaces.

Customers can use isolation segments for different reasons, including the following:

- To follow regulatory restrictions that require separation between different types of applications. For example, a health care company may not be able to host medical records and billing systems on the same machines.
- To dedicate specific hardware to different isolation segments. For example, to guarantee that high-priority apps run on a cluster of high-performance hosts.
- To separate data on multiple clients, to strengthen a security story, or offer different hosting tiers.

In CF, the Cloud Controller Database (CCDB) identifies isolation segments by name and GUID, for example `30dd879c-ce2f-11db-8314-0800200c9a66`. The isolation segment object has no internal structure beyond these two properties at the Cloud Foundry level, but BOSH associates the name of the isolation segment with Diego cells, through their `placement_tag` property.

This diagram shows how isolation segments keep apps running on different pools of cells, and how the cells communicate with each other and with the management components:



See the [Installing PCF Isolation Segment](#) and [Managing Isolation Segments](#) topics for more information about how to create and manage isolation segments in a PCF deployment.

See the [Isolation Segments](#) section of the *Cloud Controller API (CAPI) Reference* for API commands related to isolation segments.

Authentication and Authorization

[User Account and Authentication](#) (UAA) is the central identity management service for Cloud Foundry and its various components.

UAA acts as an [OAuth2](#) Authorization Server and issues access tokens for applications that request platform resources. The tokens are based on the [JSON Web Token](#) and are digitally signed by UAA.

Operators can configure the identity store in UAA. If users register an account with the Cloud Foundry platform, UAA acts as the user store and stores user passwords in the UAA database using [bcrypt](#). UAA also supports connecting to external user stores through LDAP and SAML. Once an operator has

configured the external user store, such as a corporate Microsoft Active Directory, users can use their LDAP credentials to gain access to the Cloud Foundry platform instead of registering a separate account. Alternatively, operators can use SAML to connect to an external user store and enable single sign-on for users into the Cloud Foundry platform.

Managing User Access with Role-Based Access Control

Applications that users deploy to Cloud Foundry exist within a space. Spaces exist within orgs. To view and access an org or a space, a user must be a member of it. Cloud Foundry uses role-based access control (RBAC), with each role granted permissions to either an org or a specified space. For more information about roles and permissions, refer to the [Orgs, Spaces, Roles, and Permissions](#) topic.

For more information, see [Getting Started with the Apps Manager](#) and [Managing User Accounts and Permissions Using the Apps Manager](#).

Security for Service Broker Integration

The Cloud Controller authenticates every request with the Service Broker API using HTTP or HTTPS, depending on which protocol that you specify during broker registration. The Cloud Controller rejects any broker registration that does not contain a username and password.

Service instances bound to an app contain credential data. Users specify the binding credentials for [user-provided service instances](#), while third-party brokers specify the binding credentials for managed service instances. The VCAP_SERVICES environment variable contains credential information for any service bound to an app. Cloud Foundry constructs this value from encrypted data that it stores in the Cloud Controller Database (CCDB).

 **Note:** The selected third-party broker controls how securely to communicate managed service credentials.

A third-party broker might offer a dashboard client in its catalog. Dashboard clients require a text string defined as `client_secret`. Cloud Foundry does not store this secret in the CCDB. Instead, Cloud Foundry passes the secret to the UAA component for verification using HTTP or HTTPS.

Software Vulnerability Management

Cloud Foundry manages software vulnerability using releases and BOSH stemcells. New Cloud Foundry releases are created with updates to address code issues, while new stemcells are created with patches for the latest security fixes to address any underlying operating system issues.

Ensuring Security for Application Artifacts

Cloud Foundry secures both the code and the configuration of an application using the following functionality:

- Application developers push their code using the [Cloud Foundry API](#). Cloud Foundry secures each call to the CF API using the [UAA](#) and SSL.
- The Cloud Controller uses [RBAC](#) to ensure that only authorized users can access a particular application.
- The Cloud Controller stores the configuration for an application in an encrypted database table. This configuration data includes user-specified environment variables and service credentials for any services bound to the app.
- Cloud Foundry runs the app inside a secure container. For more information, see the [Container Security](#) topic.
- Cloud Foundry operators can configure network traffic rules to control inbound communication to and outbound communication from an app. For more information, see the [Network Traffic Rules](#) section of the [Container Security](#) topic.

Security Event Logging and Auditing

For operators, Cloud Foundry provides an audit trail through the `bosh tasks` command. This command shows all actions that an operator has taken with the platform. Additionally, operators can redirect Cloud Foundry component logs to a standard syslog server using the `syslog_daemon_config` property in the `metron_agent` job of `cf-release`.

For users, Cloud Foundry records an audit trail of all relevant API invocations of an app. The Cloud Foundry Command Line Interface (cf CLI) command `cf events` returns this information.

Recommendations for Running a Secure Deployment

To help run a secure deployment, Pivotal recommends the following:

- Configure UAA clients and users using a BOSH manifest. Limit and manage these clients and users as you would any other kind of privileged account.
- Deploy within a VLAN that limits network traffic to individual VMs. This reduce the possibility of unauthorized access to the VMs within your BOSH-managed cloud.
- Enable HTTPS for applications and SSL database connections to protect sensitive data transmitted to and from applications.
- Ensure that the Jumpbox is secure, along with the load balancer and NAT VM.
- Encrypt stored files and data within databases to meet your data security requirements. Deploy using industry standard encryption and the best practices for your language or framework.
- Prohibit promiscuous network interfaces on the trusted network.
- Review and monitor data sharing and security practices with third-party services that you use to provide additional functionality to your application.
- Store SSH keys securely to prevent disclosure, and promptly replace lost or compromised keys.
- Use Cloud Foundry's RBAC model to restrict your users' access to only what is necessary to complete their tasks.
- Use a strong passphrase for both your Cloud Foundry user account and SSH keys.
- Use the [IPsec add-on](#) to encrypt IP data traffic within your deployment.

Container Security

This topic describes how Cloud Foundry (CF) secures the containers that host application instances on Linux. For an overview of other CF security features, see the [Cloud Foundry Security](#) topic.

- [Container Mechanics](#) provides an overview of container isolation.
- [Inbound and Outbound Traffic from CF](#) provides an [overview](#) of container networking and describes how CF administrators [customize](#) container network traffic rules for their deployment.
- [Container Security](#) describes how CF secures containers by running application instances in [unprivileged](#) containers and by [hardening](#) them.

Container Mechanics

Each instance of an app deployed to CF runs within its own self-contained environment, a [Garden container](#). This container isolates processes, memory, and the filesystem using operating system features and the characteristics of the virtual and physical infrastructure where CF is deployed.

CF achieves container isolation by namespacing kernel resources that would otherwise be shared. The intended level of isolation is set to prevent multiple containers that are present on the same host from detecting each other. Every container includes a private root filesystem, which includes a Process ID (PID), namespace, network namespace, and mount namespace.

CF creates container filesystems using the [Garden Roots](#) (GrootFS) tool. It stacks the following using OverlayFS:

- A **read-only base filesystem**: This filesystem has the minimal set of operating system packages and Garden-specific modifications common to all containers. Containers can share the same read-only base filesystem because all writes are applied to the read-write layer.
- A **container-specific read-write layer**: This layer is unique to each container and its size is limited by XFS project quotas. The quotas prevent the read-write layer from overflowing into unallocated space.

Resource control is managed using Linux control groups. Associating each container with its own cgroup or job object limits the amount of memory that the container may use. Linux cgroups also require the container to use a fair share of CPU compared to the relative CPU share of other containers.

 **Note:** CF does not support a RedHat Enterprise Linux OS stemcell. This is due to an inherent security issue with the way RedHat handles user namespaces and container isolation.

CPU

Each container is placed in its own cgroup. Cgroups make each container use a fair share of CPU relative to the other containers. This prevents oversubscription on the host VM where one or more containers hog the CPU and leave no computing resources to the others.

The way cgroups allocate CPU time is based on shares. CPU shares do not work as direct percentages of total CPU usage. Instead, a share is relative in a given time window to the shares held by the other containers. The total amount of CPU that can be overall divided among the cgroups is what is left by other processes that may run in the host VM.

Generally, cgroups offers two possibilities for limiting the CPU usage:*CPU affinity* and *CPU bandwidth*, the latter of which is used in Cloud Foundry.

- **CPU affinity**: It consists of binding a cgroup to certain CPU cores. The actual amount of CPU cycles that can be consumed by the cgroup is thus limited to what is available on the bound CPU cores.
- **CPU bandwidth**: Sets the weight of a cgroup with the process scheduler. The process scheduler divides the available CPU cycles among cgroups depending on the shares held by each cgroup, relative to the shares held by the others.

For example, consider two cgroups, one holding two shares and one holding four. Assuming the process scheduler gets to administer 60 CPU cycles, the first cgroup with two shares will get one third of those available CPU cycles, as it holds one third of the overall shares. Similarly, the second cgroup will get 40 cycles, as it holds two thirds of the collective shares.

The calculation of the CPU usage based on the percentage of the total CPU power available is quite sophisticated and is performed regularly as the CPU demands of the various containers fluctuates. Specifically, the percentage of CPU cycles a cgroup gets can be calculated by dividing the `cpu.shares` it holds by the sum of the `cpu.shares` of all the cgroups that are currently doing CPU activity:

```
process_cpu_share_percent = cpu.shares / sum_cpu_shares *  
100
```

In Cloud Foundry, cgroup shares range from 10 to 1024, with 1024 being the default. The actual amount of shares a cgroup gets can be read from the `cpu.shares` file of the cgroup configurations pseudo-file-system available in the container at `/sys/fs/cgroup/cpu`.

The amount of shares given to an application's cgroup depends on the amount of memory the application declares to need in the manifest. Cloud Foundry scales the number of allocated shares linearly with the amount of memory, with an app instance requesting 8G of memory getting the upper limit of 1024 shares.

```
process.cpu.shares = min(1024 * (application_memory / 8 GB), 1024)
```

The next example helps to illustrate this better. Consider three processes P1, P2 and P3, which are assigned `cpu.shares` of 5, 20 and 30, respectively.

P1 is active, while P2 and P3 require no CPU. Hence, P1 may use the whole CPU. When P2 joins in and is doing some actual work (e.g. a request comes in), the CPU share between P1 and P2 will be calculated as follows:

- P1 $\rightarrow 5 / (5+20) = 0.2 = 20\%$
- P2 $\rightarrow 20 / (5+20) = 0.8 = 80\%$
- P3 (idle)

At some point process P3 joins in as well. Then the distribution will be recalculated again:

- P1 $\rightarrow 5 / (5+20+30) = 0.0909 \approx 9\%$
- P2 $\rightarrow 20 / (5+20+30) = 0.363 \approx 36\%$
- P3 $\rightarrow 30 / (5+20+30) = 0.545 \approx 55\%$

Should P1 become idle, the following recalculation between P2 and P3 takes place:

- P1 (idle)
- P2 $\rightarrow 20 / (20+30) = 0.4 = 40\%$
- P3 $\rightarrow 30 / (20+30) = 0.6 = 60\%$

If P3 finishes or becomes idle then P2 can consume all the CPU as another recalculation will be performed.

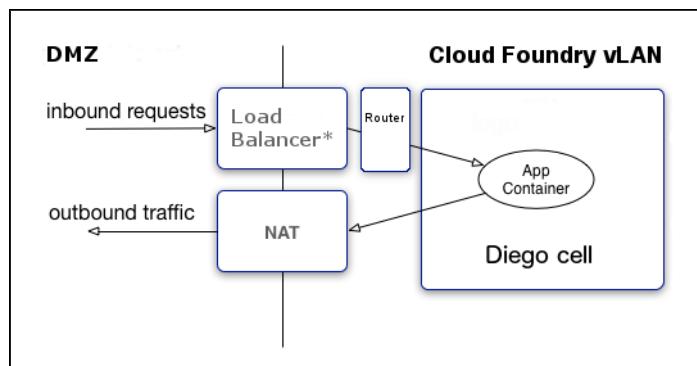
Summary: The cgroup shares are the minimum guaranteed CPU share that the process can get. This limitation becomes effective only when processes on the same host compete for resources.

Inbound and Outbound Traffic from CF

Networking Overview

A host VM has a single IP address. If you configure the deployment with the cluster on a VLAN, as recommended, then all traffic goes through the following levels of network address translation, as shown in the diagram below.

- **Inbound** requests flow from the load balancer through the router to the host cell, then into the application container. The router determines which application instance receives each request.
- **Outbound** traffic flows from the application container to the cell, then to the gateway on the cell's virtual network interface. Depending on your IaaS, this gateway may be a NAT to external networks.



Network Traffic Rules

Administrators configure rules to govern container network traffic. This is how containers send traffic outside of CF and receive traffic from outside, the Internet. These rules can prevent system access from external networks and between internal components and determine if apps can establish connections over the virtual network interface.

Administrators configure these rules at two levels:

- Application Security Groups (ASGs) apply network traffic rules at the container level. For information about creating and configuring ASGs, see [Application Security Groups](#).
- Container-to-Container networking policies determine app-to-app communication. Within CF, apps can communicate directly with each other, but the containers are isolated from outside CF. For information about administering container-to-container network policies, see [Administering Container-to-Container Networking](#).

Container Security

CF secures containers through the following measures:

- Running application instances in `unprivileged` containers by default
- [Hardening](#) containers by limiting functionality and access rights
- Allowing administrators to configure ASGs to block outbound connections from application containers. For information about creating and configuring ASGs, see [Application Security Groups](#).

Types

Garden has two container types: unprivileged and privileged. Currently, CF runs all application instances and staging tasks in unprivileged containers by default. This measure increases security by eliminating the threat of root escalation inside the container.

Hardening

CF mitigates against container breakout and denial of service attacks in the following ways:

- CF uses the full set of [Linux namespaces](#) (IPC, Network, Mount, PID, User, UTS) to provide isolation between containers running on the same host. The User namespace is not used for privileged containers.
- In unprivileged containers, CF maps UID/GID 0 (root) inside the container user namespace to a different UID/GID on the host to prevent an app from inheriting UID/GID 0 on the host if it breaks out of the container.
 - CF uses the same UID/GID for all containers.
 - CF maps all UIDs except UID 0 to themselves. CF maps UID 0 inside the container namespace to `MAX_UID-1` outside of the container namespace.
 - Container Root does not grant Host Root permissions.
- CF mounts `/proc` and `/sys` as read-only inside containers.
- CF disallows `dmesg` access for unprivileged users and all users in unprivileged containers.
- CF uses `chroot` when importing docker images from docker registries.
- CF establishes a container-specific overlay filesystem mount. CF uses `pivot_root` to move the root filesystem into this overlay, in order to isolate the container from the host system's filesystem.
- CF does not call any binary or script inside the container filesystem, in order to eliminate any dependencies on scripts and binaries inside the root filesystem.
- CF avoids side-loading binaries in the container through bind mounts or other methods. Instead, it re-executes the same binary by reading it from `/proc/self/exe` whenever it needs to run a binary in a container.
- CF establishes a virtual Ethernet pair for each container for network traffic. See the [Container Network Traffic](#) section above for more information. The virtual Ethernet pair has the following features:
 - One interface in the pair is inside the container's network namespace, and is the only non-loopback interface accessible inside the container.
 - The other interface remains in the host network namespace and is bridged to the container-side interface.
 - Egress whitelist rules are applied to these interfaces according to Application Security Groups (ASGs) configured by the administrator.
 - First-packet logging rules may also be enabled on TCP whitelist rules.
 - DNAT rules are established on the host to enable traffic ingress from the host interface to whitelisted ports on the container-side interface.

- CF applies disk quotas using container-specific XFS quotas with the specified disk-quota capacity.
- CF applies a total memory usage quota through the memory cgroup and destroys the container if the memory usage exceeds the quota.
- CF applies a fair-use limit to CPU usage for processes inside the container through the `cpu.shares` control group.
- CF allows administrators to rate limit the maximum bandwidth consumed by single application containers, configuring `rate` and `burst` properties on the `silk-cni` job.
- CF limits access to devices using cgroups but explicitly whitelists the following safe device nodes:
 - /dev/full
 - /dev/fuse
 - /dev/null
 - /dev/ptmx
 - /dev/pts/*
 - /dev/random
 - /dev/tty
 - /dev/tty0
 - /dev/tty1
 - /dev/urandom
 - /dev/zero
 - /dev/tap
 - /dev/tun
- CF drops the following [Linux capabilities](#) for all container processes. Every dropped capability limits the actions the root user can perform.
 - CAP_DAC_READ_SEARCH
 - CAP_LINUX_IMMUTABLE
 - CAP_NET_BROADCAST
 - CAP_NET_ADMIN
 - CAP_IPC_LOCK
 - CAP_IPC_OWNER
 - CAP_SYS_MODULE
 - CAP_SYS_RAWIO
 - CAP_SYS_PTRACE
 - CAP_SYS_PACCT
 - CAP_SYS_BOOT
 - CAP_SYS_NICE
 - CAP_SYS_RESOURCE
 - CAP_SYS_TIME
 - CAP_SYS_TTY_CONFIG
 - CAPLEASE
 - CAP_AUDIT_CONTROL
 - CAP_MAC_OVERRIDE
 - CAP_MAC_ADMIN
 - CAP_SYSLOG
 - CAP_WAKE_ALARM
 - CAP_BLOCK_SUSPEND
 - CAP_SYS_ADMIN (for unprivileged containers)

Orgs, Spaces, Roles, and Permissions

This topic describes orgs and spaces in PCF foundations. It also describes the default permissions for user roles in PCF.

Overview

PCF uses a role-based access control (RBAC) system to grant appropriate permissions to Pivotal Application Service (PAS) users.

Admins, Org Managers, and Space Managers can assign user roles using the Cloud Foundry Command Line Interface (cf CLI). For more information, see [Users and Roles in Getting Started with the cf CLI](#) or [Apps Manager](#).

Orgs

An org is a development account that an individual or multiple collaborators can own and use. All collaborators access an org with user accounts. Collaborators in an org share a resource quota plan, applications, services availability, and custom domains.

By default, an org has the status of *active*. An admin can set the status of an org to *suspended* for various reasons such as failure to provide payment or misuse. When an org is suspended, users cannot perform certain activities within the org, such as push apps, modify spaces, or bind services. For details on what activities are allowed for suspended orgs, see [Roles and Permissions for Suspended Orgs](#).

Spaces

Every application and service is scoped to a space. An org can contain multiple spaces. A space provides users with access to a shared location for application development, deployment, and maintenance. Each space role applies only to a particular space.

User Roles

A user account represents an individual person within the context of a PCF foundation. A user can have one or more roles. These roles define the user's permissions in orgs and spaces.

Roles can be assigned different scopes of User Account and Authentication (UAA) privileges. For more information about UAA scopes, see [Scopes](#) in *Component: User Account and Authentication (UAA) Server*.

The following describes each type of user role in PCF:

- **Admin:** Perform operational actions on all orgs and spaces using the Cloud Controller API. Assigned the `cloud_controller.admin` scope in UAA.
- **Admin Read-Only:** Read-only access to all Cloud Controller API resources. Assigned the `cloud_controller.admin_read_only` scope in UAA.
- **Global Auditor:** Read-only access to all Cloud Controller API resources except for secrets, such as environment variables. The Global Auditor role cannot access those values. Assigned the `cloud_controller.global_auditor` scope in UAA.
- **Org Managers:** Administer the org.
- **Org Auditors:** Read-only access to user information and org quota usage information.
- **Org Users:** Read-only access to the list of other org users and their roles. When an Org Manager gives a person an Org or Space role, that person automatically receives Org User status in that org.
- **Space Managers:** Administer a space within an org.
- **Space Developers:** Manage applications and services in a space.
- **Space Auditors:** Read-only access to a space.

For non-admin users, the `cloud_controller.read` scope is required to view resources, and the `cloud_controller.write` scope is required to create, update, and delete resources.

Before you assign a space role to a user, you must assign an org role to the user. The error message

Server error, error code: 1002, message: cannot set space role because user is not part of the `org` occurs when you try to set a space role before setting an org role for the user.

User Role Permissions

Each user role includes different permissions in a PCF foundation. The following sections describe the permissions associated with each user role in both active and suspended orgs in PCF.

Roles and Permissions for Active Orgs

The following table describes the default permissions for various PCF roles in active orgs.

Note: You can use feature flags to edit some of the default permissions in the following table. For more information, see [Using Feature Flags](#).

Activity	Admin	Admin Read-Only	Global Auditor	Org Manager	Org Auditor	Space Manager	Space Developer	Space Auditor
Scope of operation	Org	Org	Org	Org	Org	Space	Space	Space
Add and edit users and roles	Yes			1		1		
View users and roles	Yes							
Create and assign org quota plans	Yes							
View org quota plans	Yes							
Create orgs	Yes			2	2	2	2	2
View all orgs	Yes	Yes	Yes					
View orgs where user is a member	Yes ³	Yes ³	Yes ³	Yes	Yes	Yes	Yes	Yes
Edit, rename, and delete orgs	Yes			Yes ⁴				
Suspend or activate an org	Yes							
Create and assign space quota plans	Yes			Yes				
Create spaces	Yes			Yes				
View spaces	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Edit spaces	Yes			Yes		Yes		
Delete spaces	Yes			Yes				
Rename spaces	Yes			Yes		Yes		
View the status, number of instances, service bindings, and resource use of apps	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Add private domains ⁵	Yes			Yes				
Deploy, run, and manage apps	Yes						Yes	
Use app SSH ⁶	Yes						Yes	
Instantiate and bind services to apps	Yes						Yes	
Associate routes ⁵ , instance counts, memory allocation, and disk limit of apps	Yes						Yes	
Rename apps	Yes						Yes	
Create and manage Application Security Groups	Yes							
Create, update, and delete an Isolation Segment	Yes							
List all Isolation Segments for an org	Yes	Yes	Yes ⁷					
Entitle or revoke an Isolation Segment	Yes							
List all orgs entitled to an Isolation Segment	Yes	Yes	Yes ⁷					
Assign a default Isolation Segment to an org	Yes			Yes				
List and manage Isolation Segments for spaces	Yes			Yes				
List entitled Isolation Segment for a space	Yes	Yes	Yes	Yes		Yes	Yes	Yes
List the Isolation Segment on which an app runs	Yes	Yes	Yes	Yes		Yes	Yes	Yes

¹Not by default, unless feature flag `set_roles_by_username` is set to `true`.

²Not by default, unless feature flag `user_org_creation` is set to `true`.

³Admin, admin read-only, and global auditor roles do not need to be added as members of orgs or spaces to view resources.

⁴Org Managers can rename their orgs and edit some fields; they cannot delete orgs.

⁵Unless disabled by feature flags.

⁶This assumes that SSH is enabled for the platform, space, and app. For more information, see [SSH Access Control Hierarchy](#).

⁷Applies only to orgs to which the user account belongs.

Roles and Permissions for Suspended Orgs

The following table describes roles and permissions applied after an operator sets the status of an org to *suspended*.

Activity	Admin	Admin Read-Only	Global Auditor	Org Manager	Org Auditor	Space Manager	Space Developer	Space Auditor
Scope of operation	Org	Org	Org	Org	Org	Space	Space	Space
Add and edit users and roles	Yes							
View users and roles	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Create and assign org quota plans	Yes							
View org quota plans	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Create orgs	Yes							
View all orgs	Yes	Yes	Yes					
View orgs where user is a member	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Edit, rename, and delete orgs	Yes							
Suspend or activate an org	Yes							
Create and assign space quota plans	Yes							
Create spaces	Yes							
View spaces	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Edit spaces	Yes							
Delete spaces	Yes							
Rename spaces	Yes							
View the status, number of instances, service bindings, and resource use of applications	Yes	Yes	Yes	Yes		Yes	Yes	Yes
Add private domains [†]	Yes							
Deploy, run, and manage applications	Yes							
Instantiate and bind services to applications	Yes							
Associate routes [†] , instance counts, memory allocation, and disk limit of applications	Yes							
Rename applications	Yes							
Create and manage Application Security Groups	Yes							

Application Security Groups

This topic provides an overview of Application Security Groups (ASGs), and describes how to manage and administer them. Many of the steps below require the [Cloud Foundry Command Line Interface](#) (cf CLI) tool.

 **Note:** If you are creating ASGs for the first time, see [Restricting App Access to Internal PCF Components](#).

About Application Security Groups

Application Security Groups (ASGs) are a collection of egress rules that specify the protocols, ports, and IP address ranges where app or task instances send traffic.

ASGs define **allow** rules, and their order of evaluation is unimportant when multiple ASGs apply to the same space or deployment. The platform sets up rules to filter and log outbound network traffic from app and task instances. ASGs apply to both buildpack-based and Docker-based apps and tasks.

Staging and Running ASGs

Administrators can define a `staging` ASG for app and task staging, and a `running` ASG for app and task runtime.

When apps or tasks begin staging, they require traffic rules permissive enough to allow them to pull resources from the network. A running app or task no longer needs to pull resources, so traffic rules can be more restrictive and secure. To distinguish between these two security requirements, administrators can define a `staging` ASG for app and task staging with more permissive rules, and a `running` ASG for app and task runtime with less permissive rules.

Platform-Wide and Space-Scooped ASGs

To provide granular control when securing a deployment, administrators can assign platform-wide ASGs that apply to all app and task instances for the entire deployment, or space-scooped ASGs that apply only to apps and tasks in a particular space.

In environments with both platform-wide and space-specific ASGs, the ASGs for a particular space combine with the platform ASGs to determine the rules for that space.

Simplifying ASGs with a Services Subnet

ASGs can be complicated to configure correctly, especially when the specific IP addresses listed in a group change.

To simplify securing a deployment while still allowing apps reach external services, operators can deploy the services into a subnet that is separate from their Cloud Foundry deployment, then create ASGs for the apps that whitelist those service subnets, while denying access to any virtual machine (VM) hosting other apps.

For examples of typical ASGs, see the [Typical Application Security Groups](#) section of this topic.

Default ASGs

Pivotal Application Service (PAS) defines one default ASG, `default_security_group`. This group allows all outbound traffic from application containers on public and private networks except for the link-local range, `169.254.0.0/16`, which is blocked.

 **warning:** For security, PAS administrators must [modify the default ASGs](#) so that outbound network traffic cannot access internal components.

The ASG is defined in the Cloud Controller configuration as follows:

```
security_group_definitions:
- name: default_security_group
  rules:
  - protocol: all
    destination: 0.0.0.0-169.253.255.255
  - protocol: all
```

destination: 169.255.0.0-255.255.255.255

ASG Sets

ASGs are applied by configuring ASG sets differentiated by *scope*, platform-wide or space specific, and *lifecycle*, staging or running.

Currently, four ASG sets exist in Cloud Foundry:

- Platform-wide staging ASG set, also called “default-staging”
- Platform-wide running ASG set, also called “default-running”
- Space-scoped staging ASG set
- Space-scoped running ASG set

In environments with both platform-wide and space-specific ASG sets, combine the ASG sets for a particular space with the platform ASG sets to determine the rules for that space.

The following table indicates the differences between the four sets.

When an ASG is bound to the...	the ASG rules are applied to...
Platform-wide staging ASG set	the staging lifecycle for all apps and tasks.
Platform-wide running ASG set	the running lifecycle for all app and task instances.
Space-scoped staging ASG set	the staging lifecycle for apps and tasks in a particular space.
Space-scoped running ASG set	the running lifecycle for app and task instances in a particular space.

Typically, ASGs applied during the staging lifecycle are more permissive than the ASGs applied during the running lifecycle. This is because staging often requires access to different resources, such as dependencies.

You use different commands to apply an ASG to each of the four sets. For more information, see the [Procedures](#) section of this topic.

 **Note:** To apply a staging ASG to apps within a space, you must use cf CLI v6.28.0 or later.

The Structure and Attributes of ASGs

ASG rules are specified as a JSON array of ASG objects. An ASG object has the following attributes:

Attribute	Description	Notes
<code>protocol</code>	<code>tcp</code> , <code>udp</code> , <code>icmp</code> , or <code>all</code>	Required
<code>destination</code>	A single IP address, an IP address range like <code>192.0.2.0-192.0.2.50</code> , or a CIDR block that can receive traffic	
<code>ports</code>	A single port, multiple comma-separated ports, or a single range of ports that can receive traffic. Examples: <code>443</code> , <code>80,8080,8081</code> , <code>8080-8081</code>	Only possible if <code>protocol</code> is <code>tcp</code> or <code>udp</code> .
<code>code</code>	ICMP code	Required when <code>protocol</code> is <code>icmp</code> . A value of <code>-1</code> allows all codes.
<code>type</code>	ICMP type	Required when <code>protocol</code> is <code>icmp</code> . A value of <code>-1</code> allows all types.
<code>log</code>	Set to <code>true</code> to enable logging. For more information about how to configure system logs to be sent to a syslog drain, see the Using Log Management Services topic.	Logging is only supported with protocol type <code>tcp</code>
<code>description</code>	An optional text field for operators managing security group rules	

Process for Administering ASGs

The following table outlines the flow of tasks that the administrator carries out over the lifecycle of ASGs. Procedures for each of these tasks are given in

Managing ASGs with the cf CLI below.

Note: If you are creating ASGs for the first time, see [Restricting App Access to Internal PCF Components](#).

Task	For more information, see
1. Review the existing ASGs. If this is a new deployment, these consist of only the Default ASGs .	View ASGs
2. Create new ASGs.	Create ASGs
3. Update the existing ASGs.	Update ASGs
4. Bind ASGs to an ASG set.	Bind ASGs
5. If you need to delete an ASG, first unbind it, then delete it.	Unbind ASGs Delete ASGs

Managing ASGs with the cf CLI

This section provides the commands you need to create and manage ASGs.

View ASGs

Run the following cf CLI commands to view information about existing ASGs:

Command	Output
<code>cf security-groups</code>	All ASGs
<code>cf staging-security-groups</code>	All ASGs applied to the platform-wide staging ASG set
<code>cf running-security-groups</code>	All ASGs applied to the platform-wide running ASG set
<code>cf security-group SECURITY-GROUP</code>	All rules in the ASG named <code>SECURITY-GROUP</code> , for example, <code>cf security-group dns</code>

Note: You can also view ASGs in Apps Manager under the **Settings** tab of a space or an app.

Create ASGs

To create an ASG, perform the following steps:

1. Create a rules file: a JSON-formatted single array containing objects that describe the rules. See the following example, which allows ICMP traffic of code `1` and type `0` to all destinations, and TCP traffic to `10.0.11.0/24` on ports `80` and `443`. Also see [The Structure and Attributes of ASGs](#).

```
[  
  {  
    "protocol": "icmp",  
    "destination": "0.0.0.0/0",  
    "type": 0,  
    "code": 0  
  },  
  {  
    "protocol": "tcp",  
    "destination": "10.0.11.0/24",  
    "ports": "80,443",  
    "log": true,  
    "description": "Allow http and https traffic from ZoneA"  
  }  
]
```

2. Run `cf create-security-group SECURITY-GROUP PATH-TO-RULES-FILE`. Replace `SECURITY-GROUP` with the name of your security group, and `PATH-TO-RULES-FILE` with the absolute or relative path to a rules file.

In the following example, `my-asg` is the name of a security group, and `~/workspace/my-asg.json` is the path to a rules file.

```
$ cf create-security-group my-asg ~/workspace/my-asg.json
```

After the ASG is created, you must bind it to an ASG set before it takes effect. See [Bind ASGs](#) below.

Bind ASGs

 **Note:** Binding an ASG does not affect started apps until you restart them. To restart all of the apps in an org or a space, use the [app-restarter](#) cf CLI plugin.

To apply an ASG, you must first bind it to an ASG set.

To bind an ASG to the platform-wide staging ASG set, run `cf bind-staging-security-group SECURITY-GROUP`. Replace `SECURITY-GROUP` with the name of your security group.

Example:

```
$ cf bind-staging-security-group my-asg
```

To bind an ASG to the platform-wide running ASG set, run `cf bind-running-security-group SECURITY-GROUP` command. Replace `SECURITY-GROUP` with the name of your security group.

Example:

```
$ cf bind-running-security-group my-asg
```

To bind an ASG to a space-scoped running ASG set, run `cf bind-security-group SECURITY-GROUP ORG SPACE`. Replace `SECURITY-GROUP` with the name of your security group. Replace `ORG` and `SPACE` with the org and space where you want to bind the ASG set.

Example:

```
$ cf bind-security-group my-asg my-org my-space
```

To bind an ASG to a space-scoped staging ASG set, run `cf bind-security-group SECURITY-GROUP ORG SPACE --lifecycle staging`. Replace `SECURITY-GROUP` with the name of your security group. Replace `ORG` and `SPACE` with the org and space where you want to bind the ASG set.

Update ASGs

To update an existing ASG, perform the following steps.

1. Edit the ASG rules in the JSON file.
2. Run `cf update-security-group SECURITY-GROUP PATH-TO-RULES-FILE`. Replace `SECURITY-GROUP` with the name of the existing ASG you want to change, and `PATH-TO-RULES-FILE` with the absolute or relative path to a rules file.

In the following example, `my-asg` is the name of a security group, and `~/workspace/my-asg-v2.json` is the path to a rules file.

```
$ cf update-security-group my-asg ~/workspace/my-asg-v2.json
```

 **Note:** Updating an ASG does not affect started apps until you restart them. To restart all of the apps in an org or a space, use the [app-restarter](#) cf CLI plugin.

Unbind ASGs

 **Note:** Unbinding an ASG does not affect started apps until you restart them. To restart all of the apps in an org or a space, use the [app-restarter](#) cf CLI plugin.

To unbind an ASG from the platform-wide staging ASG set, run `cf unbind-staging-security-group SECURITY-GROUP`. Replace `SECURITY-GROUP` with the name of your security group.

Example:

```
$ cf unbind-staging-security-group my-asg
```

To unbind an ASG from the platform-wide running ASG set, run `cf unbind-running-security-group SECURITY-GROUP`. Replace `SECURITY-GROUP` with the name of your security group.

Example:

```
$ cf unbind-running-security-group my-asg
```

To unbind an ASG from a specific space, run `cf unbind-security-group SECURITY-GROUP ORG SPACE --lifecycle running`. Replace `SECURITY-GROUP` with the name of your security group. Replace `ORG` and `SPACE` with the org and space where you want to unbind the ASG set, and replace `running` with `staging` if you want to unbind from the staging ASG set.

Example:

```
$ cf unbind-security-group my-asg my-org my-space --lifecycle staging
```

Delete ASGs

 Note: You can only delete unbound ASGs. To unbind ASGs, see [Unbind ASGs above](#).

To delete an ASG, run `cf delete-security-group SECURITY-GROUP`. Replace `SECURITY-GROUP` with the name of your security group.

Example:

```
$ cf delete-security-group my-asg
```

Typical ASGs

Below are examples of typical ASGs. Configure your ASGs in accordance with your organization's network access policy for untrusted apps.

ASG	For access to
<code>dns</code>	DNS, either public or private
<code>public-networks</code>	Public networks, excluding IaaS metadata endpoints
<code>private-networks</code>	Private networks in accordance with RFC-1918
<code>load-balancers</code>	The internal Pivotal Application Service (PAS) load balancer and others
<code>internal-proxies</code>	Internal proxies
<code>internal-databases</code>	Internal databases

DNS

To resolve hostnames to IP addresses, apps require DNS server connectivity, which typically use port 53. Administrators should create or update a `dns` ASG with appropriate rules. Administrators may further restrict the DNS servers to specific IP addresses or ranges of IP addresses.

Example `dns` ASG:

```
[  
 {  
 "protocol": "tcp",  
 "destination": "0.0.0.0/0",  
 "ports": "53"  
 },  
 {  
 "protocol": "udp",  
 "destination": "0.0.0.0/0",  
 "ports": "53"  
 }  
]
```

Public Networks

Apps often require public network connectivity to retrieve app dependencies, or to integrate with services available on public networks. Example app dependencies include public Maven repositories, NPM, RubyGems, and Docker registries.

Note: You should exclude IaaS metadata endpoints, such as `169.254.169.254`, because the metadata endpoint can expose sensitive environment information to untrusted apps. The `public_networks` example below accounts for this recommendation.

Example `public_networks` ASG:

```
[  
 {  
 "destination": "0.0.0.0-9.255.255.255",  
 "protocol": "all"  
 },  
 {  
 "destination": "11.0.0.0-169.253.255.255",  
 "protocol": "all"  
 },  
 {  
 "destination": "169.255.0.0-172.15.255.255",  
 "protocol": "all"  
 },  
 {  
 "destination": "172.32.0.0-192.167.255.255",  
 "protocol": "all"  
 },  
 {  
 "destination": "192.169.0.0-255.255.255.255",  
 "protocol": "all"  
 }  
]
```

Private Networks

Network connections that are commonly allowable in private networks include endpoints such as proxy servers, Docker registries, load balancers, databases, messaging servers, directory servers, and file servers. Configure appropriate private network ASGs as appropriate. You may find it helpful to use a naming convention with `private_networks` as part of the ASG name, such as `private_networks_databases`.

Note: You should exclude any private networks and IP addresses that app and task instances should not have access to.

Example `private_networks` ASG:

```
[  
 {  
 "protocol": "tcp",  
 "destination": "10.0.0.0-10.255.255.255",  
 "ports": "443"  
 },  
 {  
 "protocol": "tcp",  
 "destination": "172.16.0.0-172.31.255.255",  
 "ports": "443"  
 },  
 {  
 "protocol": "tcp",  
 "destination": "192.168.0.0-192.168.255.255",  
 "ports": "443"  
 }  
 ]
```

Marketplace Services

Each installed Marketplace Service requires its own set of ASG rules to function properly. See the installation instructions for each installed Marketplace Service to determine which ASG rules it requires. For more information about how to provision and integrate services, see the [Services Overview](#) topics.

About the ASG Creator Tool

The ASG Creator is a command line tool that you can use to create JSON rules files. The ASG Creator lets you specify IP addresses, CIDRs, and IP address ranges that you want to disallow traffic to, as well as the addresses that you want to allow traffic to. Based on these disallow/allow (exclude/include) lists that you provide as input, the ASG Creator formulates a JSON file of allow rules.

In turn, the JSON file is the input for the `cf create-security-group` command that creates an ASG.

You can download the latest release of the ASG Creator from the Cloud Foundry incubator repository on Github: <https://github.com/cloudfoundry-incubator/asg-creator/releases/latest>

ASG Logging

The KB article [How to use Application Security Group \(ASG\) logging](#) describes how you can use ASGs to correlate emitted logs back to an app.

Application SSH Components and Processes

This document describes details about the Pivotal Application Service (PAS) SSH components for access to deployed application instances. Pivotal Application Service (PAS) supports native SSH access to applications and load balancing of SSH sessions with the load balancer for your PAS deployment.

The [SSH Overview](#) document describes procedural and configuration information about application SSH access.

SSH Components

The PAS SSH includes the following central components, which are described in more detail below:

- An implementation of an SSH [proxy server](#).
- A lightweight SSH [daemon](#).

If these components are deployed and configured correctly, they provide a simple and scalable way to access containers apps and other long running processes (LRPs).

SSH Daemon

The SSH daemon is a lightweight implementation that is built around the Go SSH library. It supports command execution, interactive shells, local port forwarding, and secure copy. The daemon is self-contained and has no dependencies on the container root file system.

The daemon is focused on delivering basic access to application instances in PAS. It is intended to run as an unprivileged process, and interactive shells and commands will run as the daemon user. The daemon only supports one authorized key, and it is not intended to support multiple users.

The daemon can be made available on a file server and Diego LRPs that want to use it can include a download action to acquire the binary and a run action to start it. PAS applications will download the daemon as part of the lifecycle bundle.

SSH Proxy Authentication

The SSH proxy hosts the user-accessible SSH endpoint and is responsible for authentication, policy enforcement, and access controls in the context of PAS. After a user has successfully authenticated with the proxy, the proxy will attempt to locate the target container and create an SSH session to a daemon running inside the container. After both sessions have been established, the proxy will manage the communication between the user's SSH client and the container's SSH Daemon.

High Availability

The following topics provide information about high availability:

- [High Availability in Cloud Foundry](#)
- [How Cloud Foundry Maintains High Availability](#)

High Availability in Cloud Foundry

This topic describes the components used to ensure high availability in Cloud Foundry, vertical and horizontal scaling, and the infrastructure required to support scaling component VMs for high availability.

Components of a High Availability Deployment

This section describes the system components needed to ensure high availability.

Availability Zones

During product updates and platform upgrades, the VMs in a deployment restart in succession, rendering them temporarily unavailable. During outages, VMs go down in a less orderly way. Spreading components across Availability Zones (AZs) and scaling them to a sufficient level of redundancy maintains high availability during both upgrades and outages and can ensure zero downtime.

Deploying Cloud Foundry across three or more AZs and assigning multiple component instances to different AZ locations lets a deployment operate uninterrupted when entire AZs become unavailable. Cloud Foundry maintains its availability as long as a majority of the AZs remain accessible. For example, a three-AZ deployment stays up when one entire AZ goes down, and a five-AZ deployment can withstand an outage of up to two AZs with no impact on uptime.

External Load Balancers

Production environments should use a highly-available customer-provided load balancing solution that does the following:

- Provides load balancing to each of the Cloud Foundry Router IP addresses
- Supports SSL termination with wildcard DNS location
- Adds appropriate x-forwarded-for and x-forwarded-proto HTTP headers to incoming requests
- (Optional) Supports WebSockets

If you are deploying in lab and test environments, the `use-haproxy.yml` ops file enables HAProxy for your foundation.

For more information, see [Using Your Own Load Balancer](#).

Blob Storage

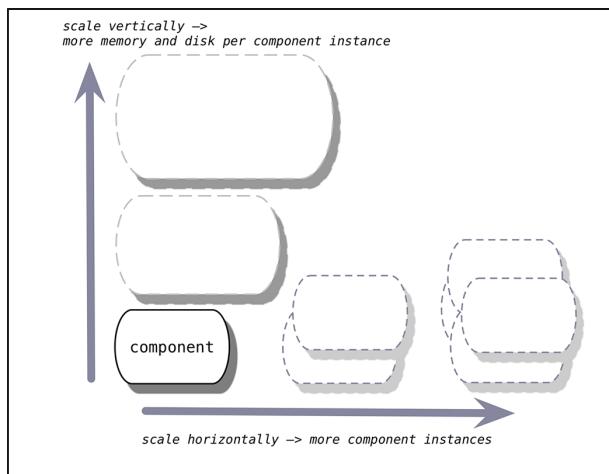
For storing blobs, large binary files, the best approach for high availability is to use external storage such as Amazon S3 or an S3-compatible service.

If you store blobs internally using WebDAV or NFS, these components run as single instances and you cannot scale them. For these deployments, use the high availability features of your IaaS to immediately recover your WebDAV or NFS server VM if it fails. Contact [Pivotal Support](#) if you need assistance.

The singleton compilation components do not affect platform availability.

Vertical and Horizontal Scaling for High Availability

You can scale platform capacity vertically by adding memory and disk, or horizontally by adding more VMs running instances of Cloud Foundry components. The nature of the applications you host on Cloud Foundry should determine whether you should scale vertically or horizontally.



For more information about scaling applications and maintaining app uptime, see [Scaling an Application Using cf scale](#) and [Using Blue-Green Deployment to Reduce Downtime and Risk](#).

Scale Vertically

Scaling vertically means adding memory and disk to your component VMs.

To scale vertically, ensure that you allocate and maintain enough of the following:

- Free space on host Diego cell VMs so that apps expected to deploy can successfully be staged and run.
- Disk space and memory in your deployment such that if one host VM is down, all instances of apps can be placed on the remaining Host VMs.
- Free space to handle one AZ going down if deploying in multiple AZs.

Scale Horizontally

Scaling horizontally means increasing the number of VM instances dedicated to running a functional component of the system.

You can horizontally scale most Cloud Foundry components to multiple instances to achieve the redundancy required for high availability.

You should also distribute the instances of multiply-scaled components across different AZs to minimize downtime during ongoing operation, product updates, and platform upgrades. If you use more than three AZs, ensure that you use an odd number of AZs.

For more information regarding zero downtime deployment, see [Scaling Instances in PAS](#).

The following table provides recommended instance counts for a high-availability deployment and the minimum instances for a functional deployment:

Pivotal Application Service (PAS) Job	Recommended Instance Number for HA	Minimum Instance Number	Notes
Diego Cell	≥ 3	1	The optimal balance between CPU/memory sizing and instance count depends on the performance characteristics of the apps that run on Diego cells. Scaling vertically with larger Diego cells makes for larger points of failure, and more apps go down when a cell fails. On the other hand, scaling horizontally decreases the speed at which the system re-balances apps. Re-balancing 100 cells takes longer and demands more processing overhead than re-balancing 20 cells.
Diego Brain	≥ 2	1	For high availability, use at least one per AZ, or at least two if only one AZ.
Diego BBS	≥ 2	1	For high availability in a multi-AZ deployment, use at least one instance per AZ. Scale Diego BBS to at least two instances for high availability in a single-AZ deployment.
MySQL Server	3	1	If you use an external database in your deployment, then you can set the MySQL Server instance count to <input type="text" value="0"/> . For instructions about scaling down an internal MySQL cluster, see Scaling Down Your MySQL Cluster .
MySQL Proxy	2	1	If you use an external database in your deployment, then you can set the MySQL Proxy instance count to <input type="text" value="0"/> .
			In a high availability deployment, you might run a single NATS instance if your deployment lacks the

NATS Server	≥ 2	1	resources to deploy two stable NATS servers. Components using NATS are resilient to message failures and the BOSH Resurrector recovers the NATS VM quickly if it becomes non-responsive.
Cloud Controller	≥ 2	1	Scale the Cloud Controller to accommodate the number of requests to the API and the number of apps in the system.
Clock Global	≥ 2	1	For a high availability deployment, scale the Clock Global job to a value greater than 1 or to the number of AZs you have.
Router	≥ 2	1	Scale the router to accommodate the number of incoming requests. Additional instances increase available bandwidth. In general, this load is much less than the load on Diego cells.
UAA	≥ 2	1	
Doppler Server	≥ 2	1	Deploying additional Doppler servers splits traffic across them. For a high availability deployment, Pivotal recommends at least two per Availability Zone.
Loggregator Traffic Controller	≥ 2	1	Deploying additional Loggregator Traffic Controllers allows you to direct traffic to them in a round-robin manner. For a high availability deployment, Pivotal recommends at least two per Availability Zone.
Syslog Scheduler	≥ 2	1	The Syslog Scheduler is a scalable component. For high availability, use at least one instance per AZ, or at least two instances if only one AZ is present.

Configure Support for High Availability Components

This section describes the surrounding infrastructure required to support scaling component VMs for high availability.

BOSH Resurrector

The BOSH Resurrector increases Pivotal Application Service (PAS) availability in the following ways:

- Reacts to hardware failure and network disruptions by recreating virtual machines on active, stable hosts
- Detects operating system failures by continuously monitoring virtual machines and recreating them as required
- Continuously monitors the BOSH Agent running on each virtual machine and recreates the VMs as required

The BOSH Resurrector continuously monitors the status of all virtual machines in an PAS deployment. The Resurrector also monitors the BOSH Agent on each VM. If either the VM or the BOSH Agent fail, the Resurrector recreates the virtual machine on another active host. To enable the BOSH Resurrector, see [Enable BOSH Resurrector](#).

Resource Pools

To configure your resource pools according to the requirements of your deployment, see the [Ops Manager configuration topic for your IaaS](#).

Each IaaS has different ways of limiting resource consumption for scaling VMs. Consult with your IaaS administrator to ensure additional VMs and related resources, like IPs and storage, will be available when scaling.

For information about configuring resource pools for Amazon Web Services, see [Amazon EC2 FAQs](#) in the Amazon documentation. For information about configuring resource pools for OpenStack, see [Manage projects and users](#) in the OpenStack documentation. For information about configuring resource pools for vSphere, see [Configuring BOSH Director on vSphere](#).

Databases

For database services deployed outside Cloud Foundry, plan to leverage your infrastructure's high availability features and to configure backup and restore where possible. For more information about scaling internal database components, see the [Scaling Instances in PAS](#) topic.

 **Note:** Data services may have single points of failure depending on their configuration.

Contact [Pivotal Support](#) if you need assistance.

How Cloud Foundry Maintains High Availability

This topic explains how Pivotal Cloud Foundry (PCF) deployments include several layers of high availability to keep applications running during system failure. These layers include AZs, application health management, process monitoring, and VM resurrection.

Availability Zones

PCF supports deploying applications instances across multiple AZs. This level of high availability requires that you define AZs in your IaaS. PCF balances the applications you deploy across the AZs you defined. If an AZ goes down, you still have application instances running in another.

You can configure your deployment so that Diego cells are created across these AZs. See the [Assign AZs and Networks](#) pane of the PAS tile.

Health Management for App Instances

If you lose application instances for any reason, such as a bug in the app or an AZ going down, PCF restarts new instances to maintain capacity. Under Diego architecture, the nsync, BBS, and Cell Rep components track the number of instances of each application that are running across all of the Diego cells. When these components detect a discrepancy between the actual state of the app instances in the cloud and the desired state as known by the Cloud Controller, they advise the Cloud Controller of the difference and the Cloud Controller initiates the deployment of new application instances.

For more information about the nsync, BBS, and Cell Rep components, see the [nsync, BBS, and Cell Rep](#) section of the *Cloud Foundry Components* topic.

Process Monitoring

PCF uses a BOSH agent, monit, to monitor the processes on the component VMs that work together to keep your applications running, such as nsync, BBS, and Cell Rep. If monit detects a failure, it restarts the process and notifies the BOSH agent on the VM. The BOSH agent notifies the BOSH Health Monitor, which triggers responders through plugins such as email notifications or paging.

Resurrection for VMs

BOSH detects if a VM is present by listening for heartbeat messages that are sent from the BOSH agent every 60 seconds. The BOSH Health Monitor listens for those heartbeats. When the Health Monitor finds that a VM is not responding, it passes an alert to the Resurrector component. If the Resurrector is enabled, it sends the IaaS a request to create a new VM instance to replace the one that failed.

To enable the Resurrector, see the following pages for your particular IaaS: [AWS](#), [Azure](#), [GCP](#), [OpenStack](#), or [vSphere](#).

Running Apps in the Cloud

The following topics provide information about how Pivotal Application Service (PAS) runs apps:

- [How Apps are Staged](#)
- [Cloud Controller Blobstore](#)
- [App Container Lifecycle](#)
- [How the Diego Auction Allocates Jobs](#)

How Apps Are Staged

This topic describes how Diego stages buildpack apps and Docker images.

Overview

Cloud Foundry uses Diego to manage app containers. It is a self-healing system that attempts to keep the correct number of instances running in Diego Cells to avoid network failures and crashes. For more information about Diego, see [Diego Components and Architecture](#).

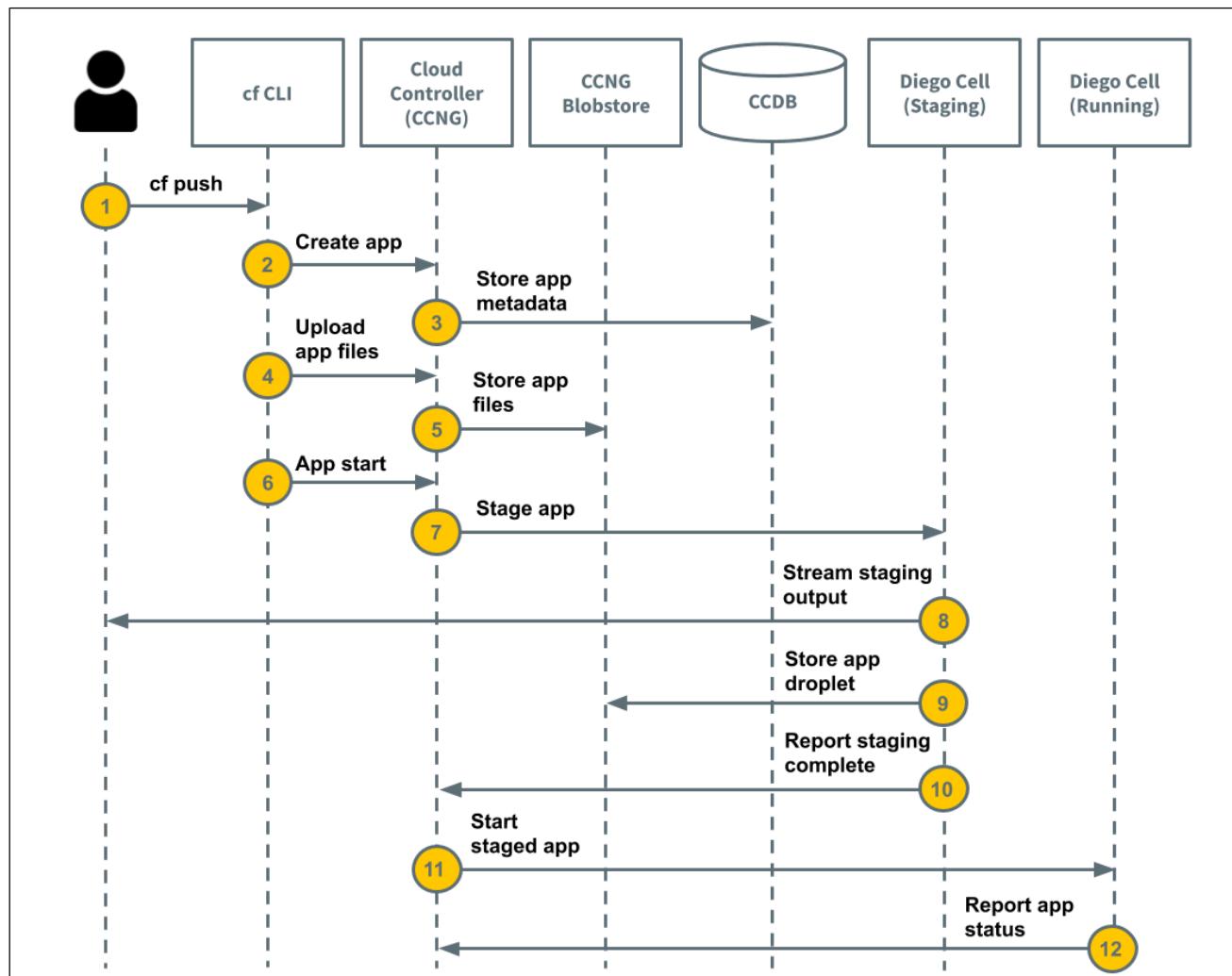
This topic references tasks and long-running processes (LRPs). For more information about these, see [Tasks and Long-Running Processes](#).

To better understand the flow and caching of source bits during staging, see [Cloud Controller Blobstore](#).

How Diego Stages Buildpack Apps

This section describes how Diego stages buildpack apps.

The following diagram illustrates the steps and components involved in the process of staging a buildpack app.



For a description of each step in the process of staging a buildpack app, see the following:

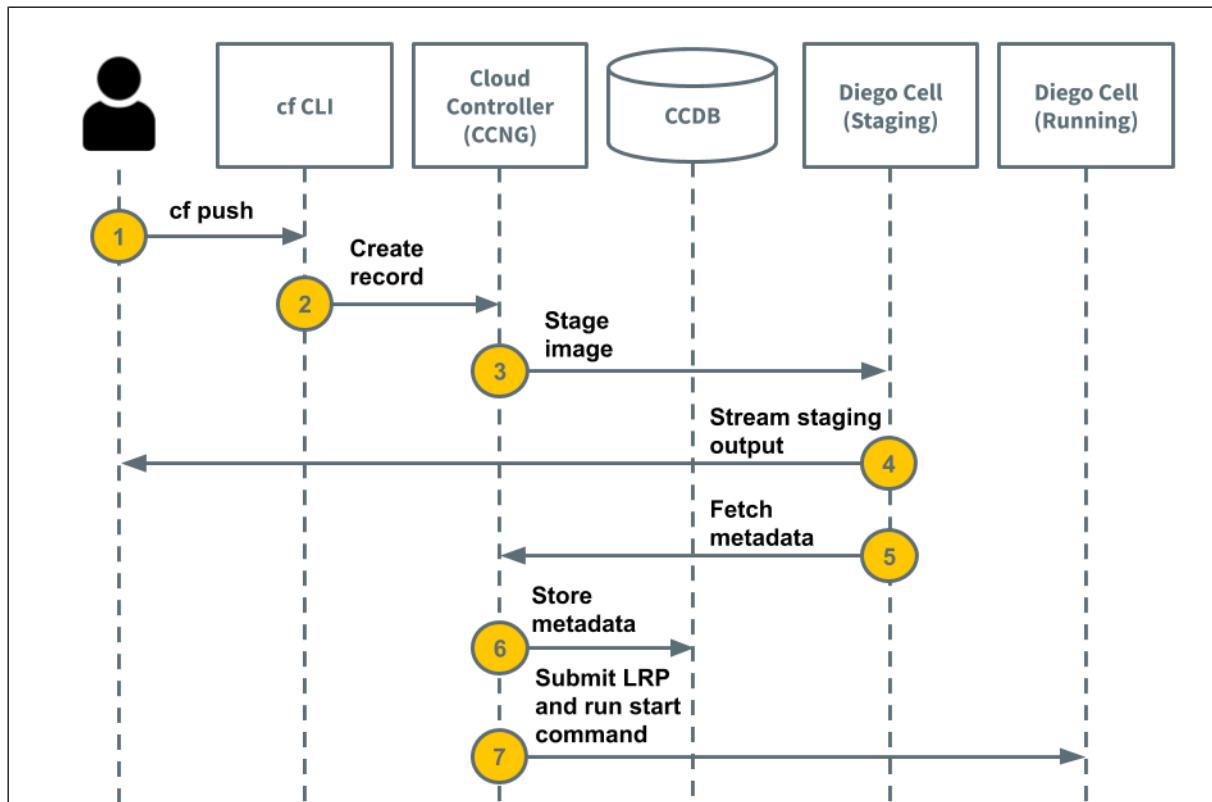
1. A developer runs `cf push`.
2. The Cloud Foundry Command Line Interface (cf CLI) tells the Cloud Controller to create a record for the app. See the [Cloud Controller](#) topic for more information about the Cloud Controller.

3. The Cloud Controller stores the app metadata. App metadata can include the app name, number of instances, buildpack, and other information about the app.
4. This step includes the following:
 - a. The cf CLI requests a resource match from the Cloud Controller.
 - b. The cf CLI uploads the app source files, omitting any app files that already exist in the resource cache.
 - c. The Cloud Controller combines the uploaded app files with files from the resource cache to create the app package.
5. The Cloud Controller stores the app package in the blobstore. For more information, see the [Blobstore](#) section in the *Cloud Foundry Components* topic.
6. The cf CLI issues a request to start the app.
7. This step includes the following:
 - a. The Cloud Controller issues a staging request to Diego.
 - b. Diego schedules a Diego Cell to run the staging task.
 - c. The task downloads buildpacks and the app buildpack cache, if present.
 - d. The task uses the buildpack to compile and stage the app.
8. The Diego Cell streams the output of the staging process. A developer may need to view the output to troubleshoot staging problems.
9. This step includes the following:
 - a. The task creates a tarball, or droplet, with the compiled and staged app.
 - b. The Diego Cell stores the droplet in the blobstore.
 - c. The task uploads the buildpack cache to the blobstore for use the next time the app is staged.
10. The Diego Bulletin Board System (BBS) reports to the Cloud Controller that staging is complete. If staging does not complete within 15 minutes, it fails.
11. Diego schedules the app as a LRP on one or more Diego Cells.
12. The Diego Cells report the status of the app to the Cloud Controller.

How Diego Stages Docker Images

This section describes how Diego stages Docker images.

The following diagram illustrates the steps and components involved in the process of staging a Docker image.



For a description of each step in the process of staging a Docker image, see the following:

1. A developer runs `cf push` and includes the name of a Docker image in an accessible Docker Registry.
2. The cf CLI tells the Cloud Controller to create a record for the Docker image.
3. This step includes the following:
 - a. The Cloud Controller issues a staging request to Diego.
 - b. Diego schedules a Diego Cell to run the task.
4. The Diego Cell streams the output of the staging process. A developer may need to view the output to troubleshoot staging problems.
5. The task fetches the metadata associated with the Docker image and returns a portion of it to the Cloud Controller.
6. The Cloud Controller stores the metadata in the Cloud Controller database (CCDB).
7. This step includes the following:
 - a. The Cloud Controller uses the Docker image metadata to construct a LRP that runs the start command specified in the Dockerfile.
 - b. The Cloud Controller submits the LRP to Diego.
 - c. Diego schedules the LRP on one or more Diego Cells.
 - d. The Cloud Controller instructs Diego and the Gorouter to route traffic to the Docker image.

Note: The Cloud Controller takes into account any user-specified overrides specified in the Dockerfile, such as environment variables.

App Container Lifecycle

This topic describes the lifecycle of an app container for Cloud Foundry (CF) deployments running on the Diego architecture.

Deployment

The app deployment process involves uploading, staging, and starting the app in a container. Your app must successfully complete each of these phases within certain time limits. The default time limits for the phases are as follows:

- Upload: 15 minutes
- Stage: 15 minutes
- Start: 60 seconds

 **Note:** Your administrator can change these defaults. Check with your administrator for the actual time limits set for app deployment.

Developers can change the time limit for starting apps through an app manifest or on the command line. For more information, see [The timeout attribute](#) section of the Deploying with App Manifests topic and [Using Application Health Checks](#).

Crash Events

If an app instance crashes, CF automatically restarts it by rescheduling the instance on another container three times. After three failed restarts, CF waits thirty seconds before attempting another restart. The wait time doubles each restart until the ninth restart, and remains at that duration until the 200th restart. After the 200th restart, CF stops trying to restart the app instance.

Evacuation

Certain operator actions require restarting VMs with containers hosting app instances. For example, an operator who updates stemcells or installs a new version of CF must restart all the VMs in a deployment.

CF automatically relocates the instances on VMs that are shutting down through a process called evacuation. CF recreates the app instances on another VM, waits until they are healthy, and then shuts down the old instances. During an evacuation, developers may see their app instances in a duplicated state for a brief period.

During this app duplication process, singleton app instances may become temporarily unavailable if the replacement instance does not become healthy within the cell's evacuation timeout, which defaults to 10 minutes. Because of this, app developers with a low tolerance for brief downtime may prefer to run several instances of their app. See [Run Multiple Instances to Increase Availability](#).

Shutdown

PCF requests a shutdown of your app instance in the following scenarios:

- When a user runs `cf scale`, `cf stop`, `cf push`, `cf delete`, or `cf restart-app-instance`
- As a result of a system event, such as the replacement procedure during Diego cell evacuation or when an app instance stops because of a failed health-check probe

To shut down the app, CF sends the app process in the container a SIGTERM. The process has ten seconds to shut down gracefully. If the process has not exited after ten seconds, CF sends a SIGKILL.

Apps must finish their in-flight jobs within ten seconds of receiving the SIGTERM before CF terminates the app with a SIGKILL. For instance, a web app must finish processing existing requests and stop accepting new requests.

 **Note:** One exception to the cases mentioned above is when monit restarts a crashed Diego cell rep or Garden server. In this case, CF immediately stops the apps that are still running using SIGKILL.

How Diego Balances App Processes

Diego balances app processes over the virtual machines (VMs) in a Cloud Foundry (CF) installation using the Diego Auction. When new processes need to be allocated to VMs, the Diego Auction determines which ones should run on which machines. The auction algorithm balances the load on VMs and optimizes app availability and resilience. This topic explains how the Diego Auction works.

Refer to the [Auction repository](#) on GitHub for source code and more information.

Tasks and Long-Running Processes

The Diego Auction distinguishes between two types of jobs: **Tasks** and **Long-Running Processes** (LRPs).

- **Tasks** run once, for a finite amount of time. A common example is a staging task that compiles an app’s dependencies, to form a self-contained droplet that makes the app portable and runnable on multiple VMs. Other examples of tasks include making a database schema change, bulk importing data to initialize a database, and setting up a connected service.
- **Long-Running Processes** run continuously, for an indefinite amount of time. LRPCs terminate only if stopped or killed, or if they crash. Examples include web servers, asynchronous background workers, and other applications and services that continuously accept and process input. To make high-demand LRPCs more available, Diego may allocate multiple instances of the same application to run simultaneously on different VMs, often spread across Availability Zones that serve users in different geographic regions.

The Diego Auction process repeats whenever new jobs need to be allocated to VMs. Each auction distributes a current **batch** of work, Tasks and LRPCs, that can include newly-created jobs, jobs left unallocated in the previous auction, and jobs left orphaned by failed VMs. Diego does not redistribute jobs that are already running on VMs. Only one auction can take place at a time, which prevents placement collisions.

Ordering the Auction Batch

The Diego Auction algorithm allocates jobs to VMs to fulfill the following outcomes, in decreasing **priority** order:

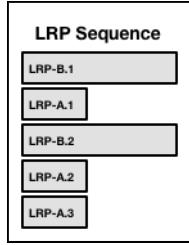
1. Keep at least one instance of each LRP running.
2. Run all of the Tasks in the current batch.
3. Distribute as much of the total desired LRP load as possible over the remaining available VMs, by spreading multiple LRP instances broadly across VMs and their Availability Zones.

To achieve these outcomes, each auction begins with the [Diego Auctioneer](#) component arranging the batch’s jobs into a priority order. Some of these jobs may be duplicate instances of the same process that Diego needs to allocate for high-traffic LRPCs, to meet demand. So the Auctioneer creates a list of multiple LRP instances based on the desired instance count configured for each process.

For example, if the process LRP-A has a desired instance count of 3 and a memory load of 2, and process LRP-B has 2 desired instances and a load of 5, the Auctioneer creates a list of jobs for each process as follows:

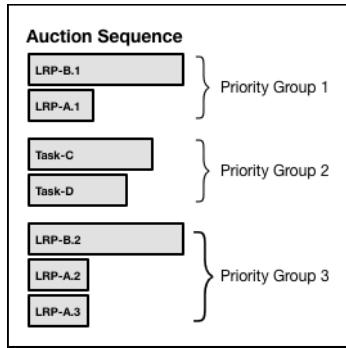
Process	Desired Instances	Load	Jobs
LRP-A	3	2	LRP-A.1 LRP-A.2 LRP-A.3
LRP-B	2	5	LRP-B.1 LRP-B.2

The Auctioneer then builds an ordered sequence of LRP instances by cycling through the list of LRPCs in decreasing order of load. With each cycle, it adds another instance of each LRP to the sequence, until all desired instances of the LRP have been added. With the example above, the Auctioneer would order the LRPCs like this:



The Auctioneer then builds an ordered sequence for all jobs, both LRPs and Tasks. Reflecting the auction batch [priority order](#), the first instances of LRPs are first priority. Tasks are next, in decreasing order of load. Duplicate LRP jobs come last.

Adding one-time Task-C (load = 4) and Task-D (load = 3) to the above example, the priority order becomes:



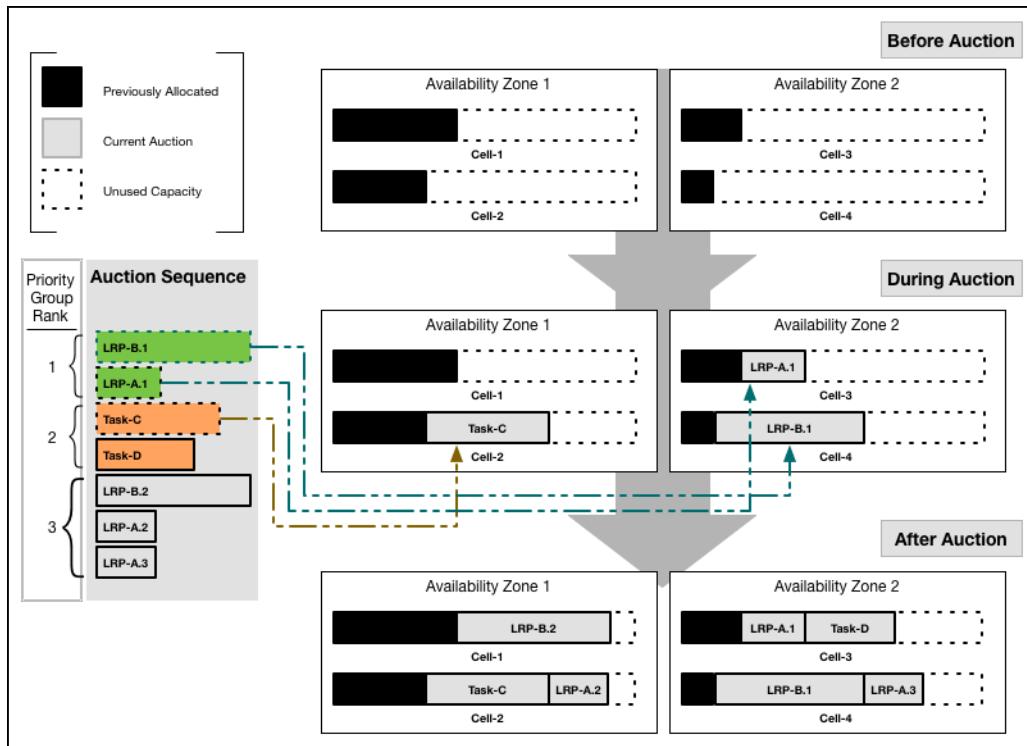
Auctioning the Batch to the Cells

With all jobs sorted in priority order, the Auctioneer allocates each in turn to one of the VMs. The process resembles an auction, where VMs “bid” with their suitability to run each job. Facilitating this process, each app VM has a resident [Cell](#) that monitors and allocates the machine’s operation. The Cell participates in the auction on behalf of the virtual machine that it runs on.

Starting with the highest-priority job in the ordered sequence, the Auctioneer polls all the Cells on their fitness to run the currently-auctioned job. Cells “bid” to host each job according to the following priorities, in decreasing order:

1. Allocate all jobs only to Cells that have the correct software stack to host them, and sufficient resources given their allocation so far during this auction.
2. Allocate LRP instances into Availability Zones that are not already hosting other instances of the same LRP.
3. Within each Availability Zone, allocate LRP instances to run on Cells that are not already hosting other instances of the same LRP.
4. Allocate any job to the Cell that has lightest load, from both the current auction and jobs it has been running already. In other words, distribute the total load evenly across all Cells.

Our example auction sequence has seven jobs: five LRP instances and two Tasks. The following diagram shows how the Auctioneer might distribute this work across four Cells running in two Availability Zones:



If the Auctioneer reaches the end of its sequence of jobs, having distributed all jobs to the Cells, it submits requests to the Cells to execute their allotted work. If the Cells ran out of capacity to handle all jobs in the sequence, the Auctioneer carries the unallocated jobs over and merges them into the next auction batch, to be allocated in the next auction.

Triggering Another Auction

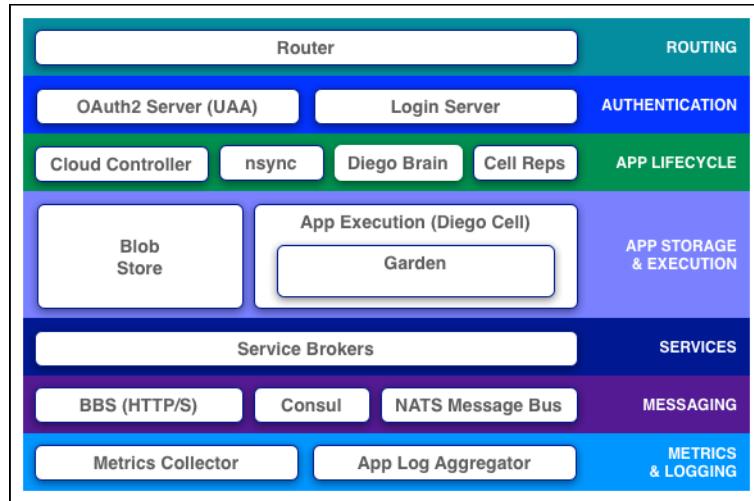
The Diego Auction process repeats to adapt a Cloud Foundry deployment to its changing workload. For example, the BBS initiates a new auction when it detects that the actual number of running instances of LRPs does not match the number desired. Diego's BBS component monitors the number of instances of each LRP that are currently running. The BBS component periodically compares this number with the desired number of LRP instances, as configured by the user. If the actual number falls short of what is desired, the BBS triggers a new auction. In the case of a surplus of application instances, the BBS kills the extra instances and initiates another auction.

The Cloud Controller also triggers an auction whenever a Cell fails. After any auction, if a Cell responds to its work request with a message that it cannot perform the work after all, the Auctioneer carries the unallocated work over into the next batch. But if the Cell fails to respond entirely, for example if its connection times out, the unresponsive Cell may still be running its work. In this case, the Auctioneer does not automatically carry the Cell's work over to the next batch. Instead, the Auctioneer defers to the BBS to continue monitoring the states of the Cells, and to re-assign unassigned work later if needed.

Runtime Components

Cloud Foundry components include a self-service application execution engine, an automation engine for application deployment and lifecycle management, and a scriptable command line interface (CLI), as well as integration with development tools to ease deployment processes. Cloud Foundry has an open architecture that includes a buildpack mechanism for adding frameworks, an application services interface, and a cloud provider interface.

See the descriptions below for more information about Cloud Foundry components. Some descriptions include links to more detailed documentation.



Routing

Router

The [router](#) routes incoming traffic to the appropriate component, either a Cloud Controller component or a hosted application running on a Diego Cell.

The router periodically queries the Diego Bulletin Board System (BBS) to determine which cells and containers each application currently runs on. Using this information, the router recomputes new routing tables based on the IP addresses of each cell virtual machine (VM) and the host-side port numbers for the cell's containers.

Authentication

OAuth2 Server (UAA) and Login Server

The OAuth2 server (the [UAA](#)) and Login Server work together to provide identity management.

App Lifecycle

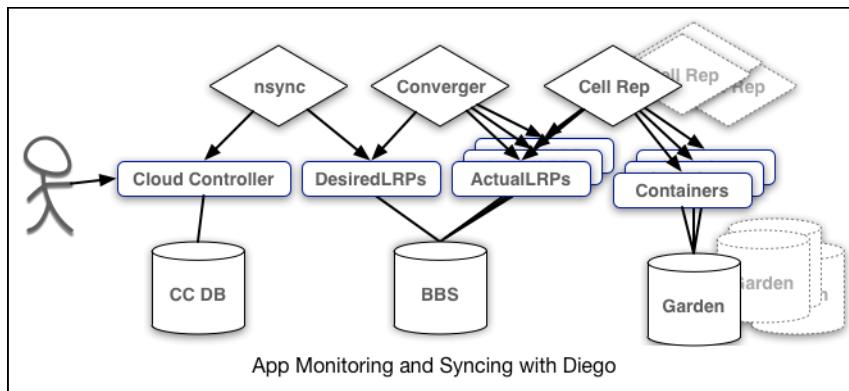
Cloud Controller and Diego Brain

The [Cloud Controller](#) (CC) directs the deployment of applications. To push an app to Cloud Foundry, you target the Cloud Controller. The Cloud Controller then directs the Diego Brain through the CC-Bridge components to coordinate individual [Diego cells](#) to stage and run applications.

The Cloud Controller also maintains records of [orgs](#), [spaces](#), [user roles](#), [services](#), and more.

nsync, BBS, and Cell Reps

To keep applications available, cloud deployments must constantly monitor their states and reconcile them with their expected states, starting and stopping processes as required.



The nsync, BBS, and Cell Rep components work together along a chain to keep apps running. At one end is the user. At the other end are the instances of applications running on widely-distributed VMs, which may crash or become unavailable.

Here is how the components work together:

- **nsync** receives a message from the Cloud Controller when the user scales an app. It writes the number of instances into a `DesiredLRP` structure in the Diego BBS database.
- **BBS** uses its convergence process to monitor the `DesiredLRP` and `ActualLRP` values. It launches or kills application instances as appropriate to ensure the `ActualLRP` count matches the `DesiredLRP` count.
- **Cell Rep** monitors the containers and provides the `ActualLRP` value.

App Storage and Execution

Blobstore

The blobstore is a repository for large binary files, which GitHub cannot easily manage because GitHub is designed for code. The blobstore contains the following:

- Application code packages
- Buildpacks
- Droplets

You can configure the blobstore as either an internal server or an external S3 or S3-compatible endpoint. See this [Knowledge Base article](#) for more information about the blobstore.

Diego Cell

Application instances, application tasks, and staging tasks all run as `Garden` containers on the Diego Cell VMs. The Diego cell rep component manages the lifecycle of those containers and the processes running in them, reports their status to the Diego BBS, and emits their logs and metrics to `Loggregator`.

Services

Service Brokers

Applications typically depend on [services](#) such as databases or third-party SaaS providers. When a developer provisions and binds a service to an application, the service broker for that service is responsible for providing the service instance.

Messaging

Internal HTTPS and BBS

Cloud Foundry component VMs communicate with each other internally through HTTP and HTTPS protocols, sharing temporary messages and data stored in Diego's [Bulletin Board System \(BBS\)](#).

- BOSH Director collocates a [BOSH DNS](#) server on every deployed VM. All VMs keep up-to-date DNS records for all the other VMs in the same foundation, enabling service discovery between VMs. BOSH DNS also provides client-side load-balancing by randomly selecting a healthy VM when multiple VMs are available.
- Diego's [Bulletin Board System](#) (BBS) stores more frequently updated and disposable data such as cell and app status, unallocated work, and heartbeat messages, as well as longer-lived distributed locks. The BBS stores data in MySQL, using the [Go MySQL Driver](#).

The route-emitter component uses the NATS protocol to broadcast the latest routing tables to the routers.

Metrics and Logging

Loggregator

The [Loggregator](#) (log aggregator) system streams application logs to developers.

Diego Components and Architecture

This topic describes the components that form and interact with the Diego system.

Overview

Cloud Foundry uses the Diego system to manage app containers. Diego components assume app scheduling and management responsibility from the Cloud Controller.

Diego is a self-healing container management system that attempts to keep the correct number of instances running in Diego cells to avoid network failures and crashes. Diego schedules and runs Tasks and Long-Running Processes (LRP). For more about Tasks and LRPCs, see [How the Diego Auction Allocates Jobs](#).

You can submit, update, and retrieve the desired number of Tasks and LRPCs using the [Bulletin Board System \(BBS\) API](#).

How Diego Runs an App

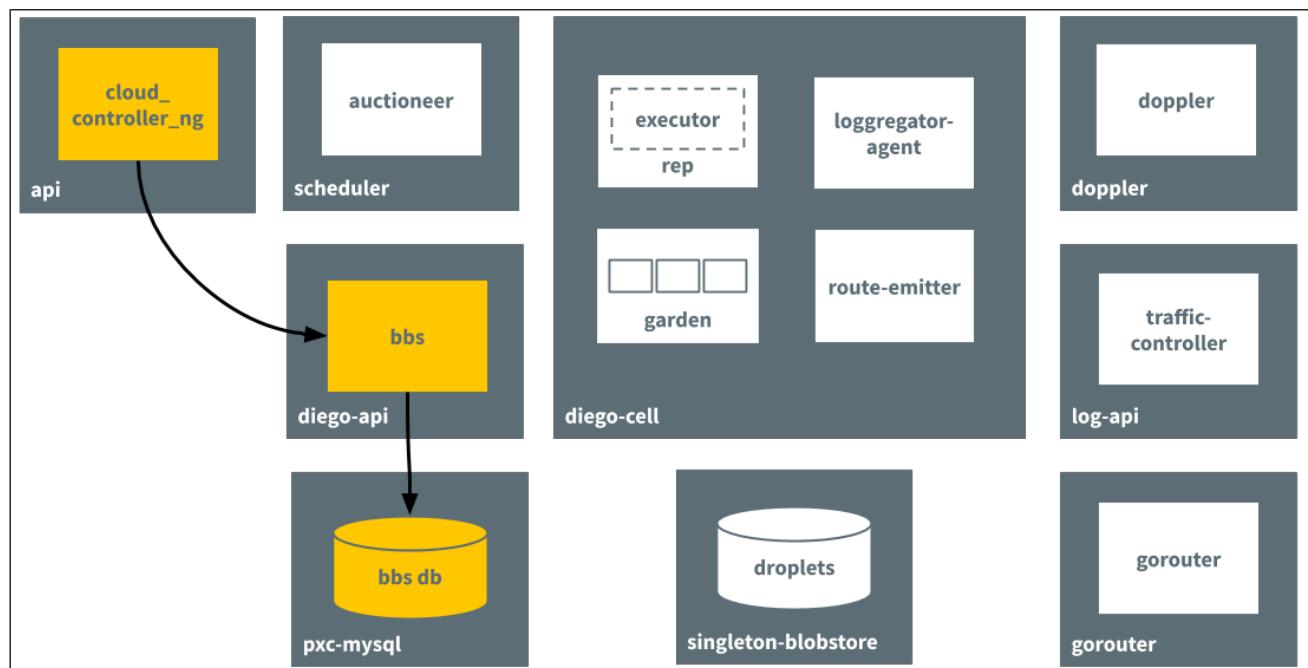
The following sections describe how Diego handles a request to run an app. This is only one of the processes that happen in Diego. For example, running an app assumes the app has already been staged. For more information about the staging process, see [How Applications are Staged](#).

The diagrams below do not include all of the components Diego. For information about each Diego component, see [Diego Components](#).

Note: The images below are based on the VM names in an open-source deployment of Cloud Foundry Application Runtime. In Pivotal Application Service (PAS), the processes interact in the same way, but are on different VMs. Correct VM names for each process are in the components sections of this topic.

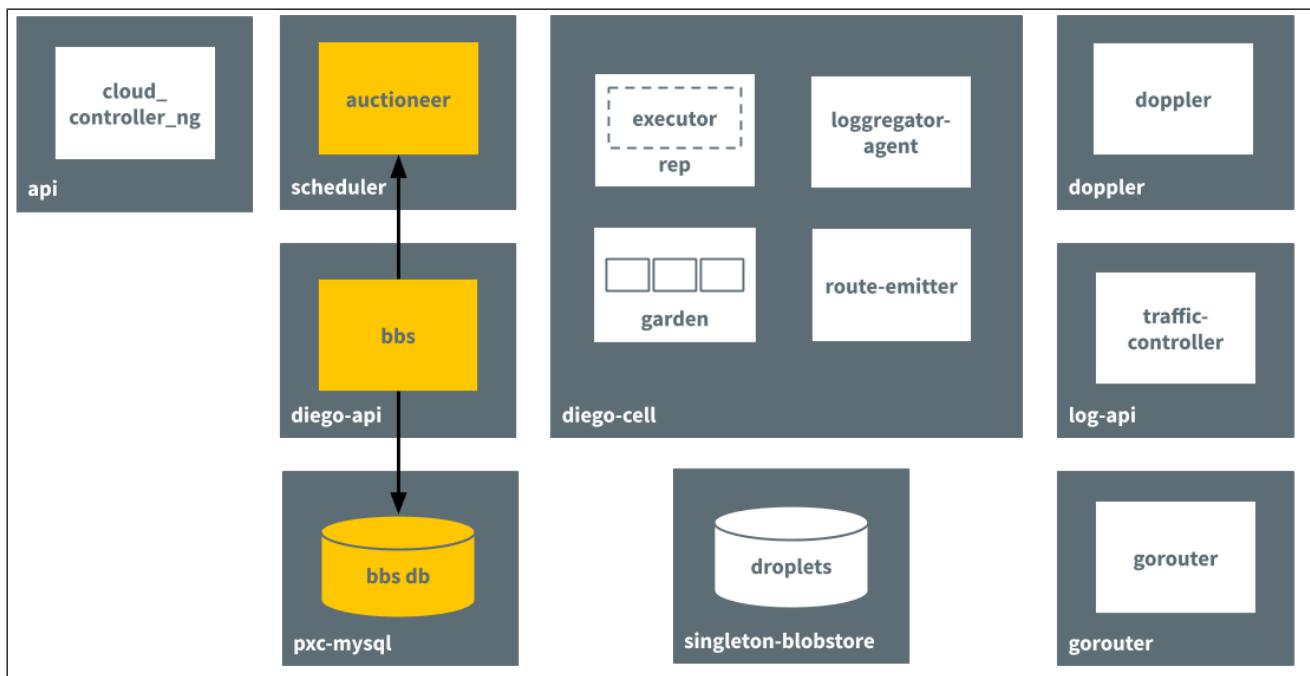
Receives the Request to Run an App

The Cloud Controller passes requests to run apps to the Diego BBS, which stores information about the request in its database.



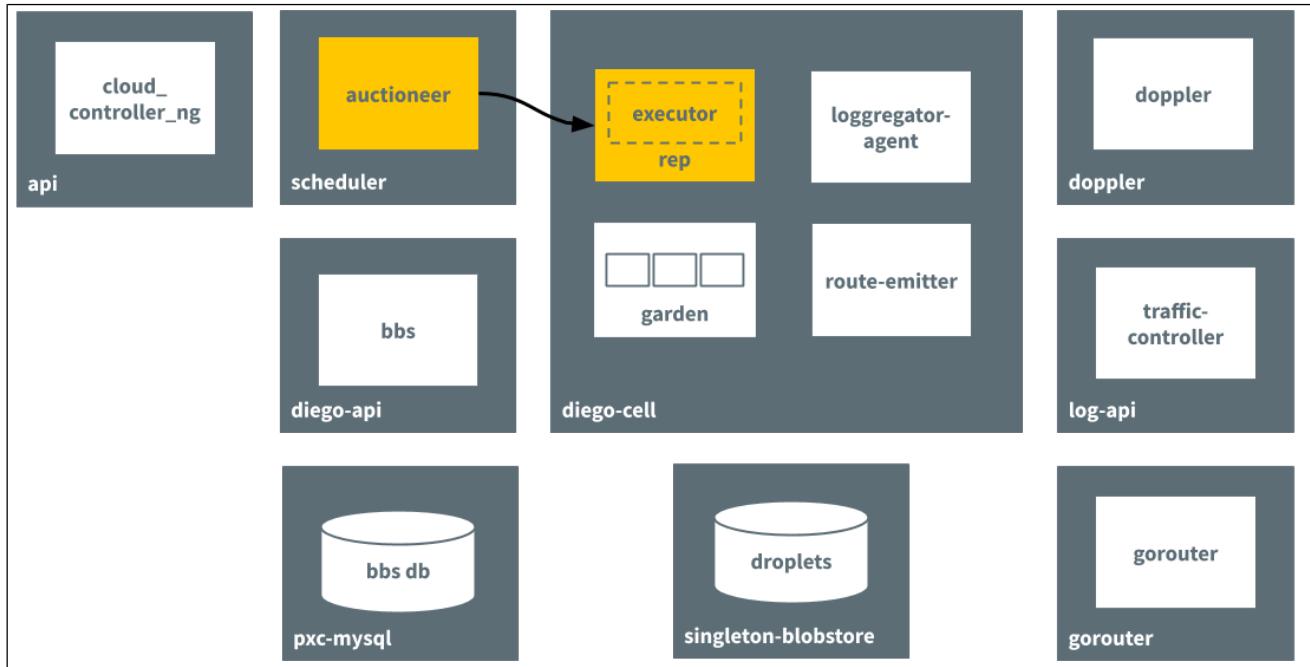
Passes Request to the Auctioneer Process

The BBS contacts the Auctioneer to create an auction based on the desired resources for the app. It references the information stored in its database.



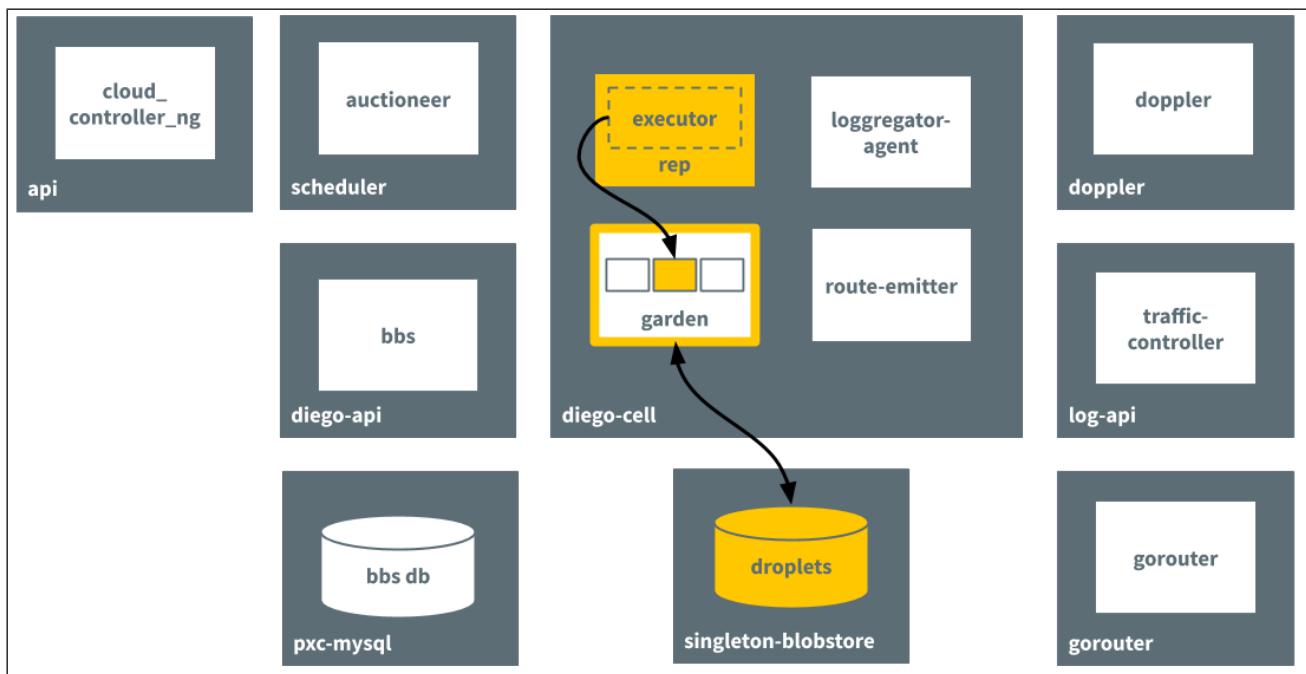
Performs Auction

Through an auction, the Auctioneer finds a Diego cell to run the app on. The Rep job on the Diego cell accepts the auction request.



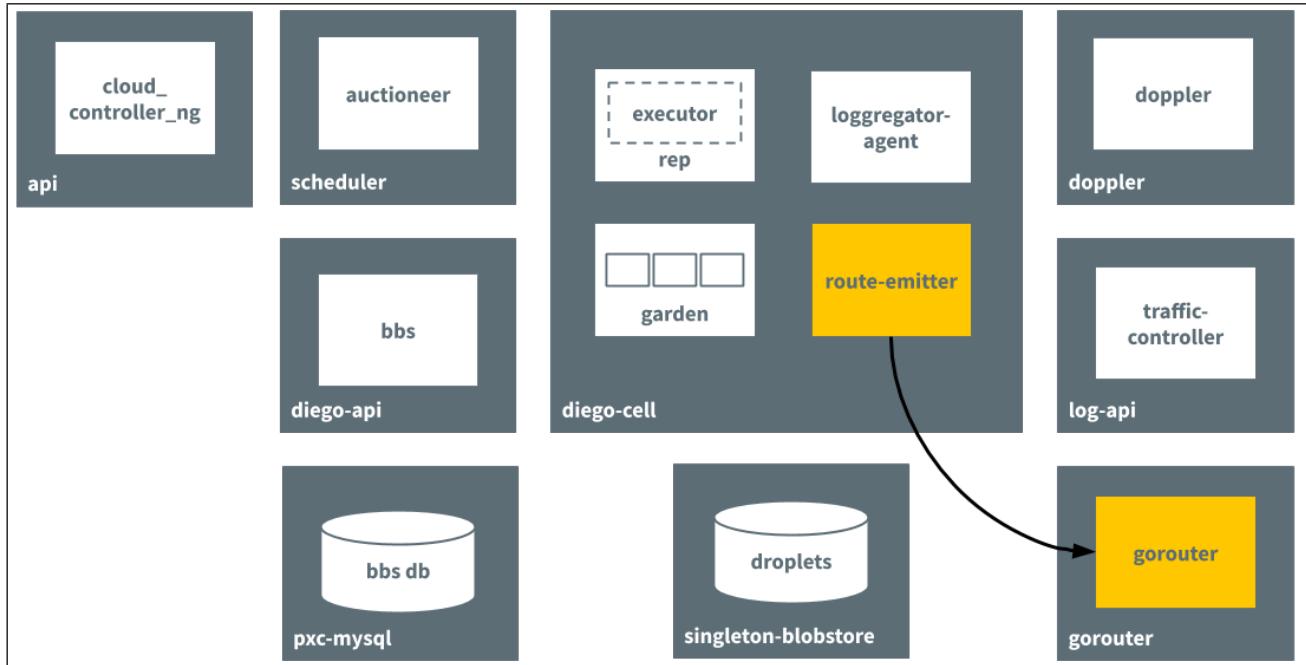
Creates Container and Runs App

The in-process Executor creates a Garden container in the Diego cell. Garden downloads the droplet that resulted from the staging process and runs the app in the container.



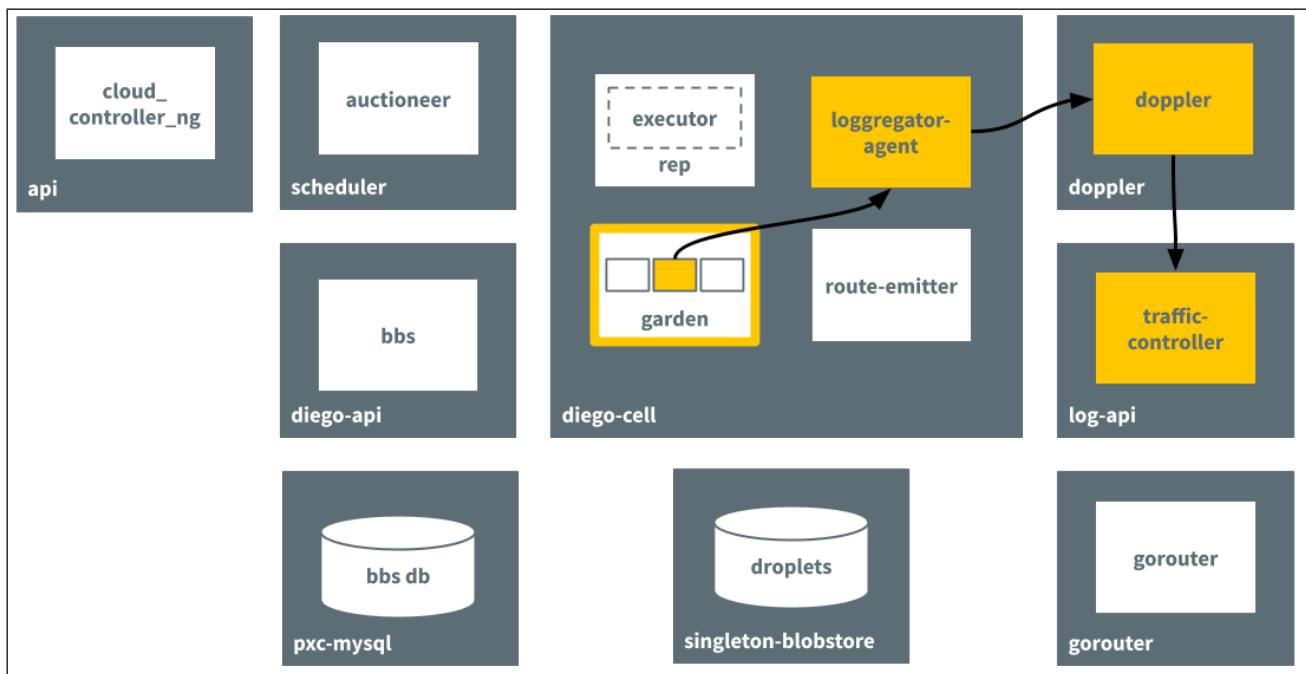
Emits Route for App

The `route-emitter` process emits a route registration message to Gorouter for the new app running on the the Diego cell.



Sends Logs to Loggregator

The Loggregator agent forwards app logs, errors, and metrics to the Cloud Foundry Loggregator. For more information, see the [Application Logging in Cloud Foundry](#) topic.



Diego Components

The following table describes the jobs that are part of the Cloud Foundry Diego BOSH release.

Component	Function
Job: auctioneer VM: diego_brain	<ul style="list-style-type: none"> Distributes work through auction to Cell Reps over SSL/TLS. See How the Diego Auction Allocates Jobs. Maintains a lock in Locket to ensure only one auctioneer handles auctions at a time.
Job: bbs VM: diego_database	<ul style="list-style-type: none"> Maintains a real-time representation of the state of the Diego cluster, including desired LRPCs, running LRPCs, and in-flight Tasks. Provides an RPC-style API over HTTP to Diego Core components for external clients as well as internal clients, including the SSH Proxy and Route Emitter. Ensures consistency and fault tolerance for Tasks and LRPCs by comparing desired state with actual state. Keeps <code>DesiredLRP</code> and <code>ActualLRP</code> counts synchronized. If the <code>DesiredLRP</code> count exceeds the <code>ActualLRP</code> count, requests a start auction from the Auctioneer. If the <code>ActualLRP</code> count exceeds the <code>DesiredLRP</code> count, sends a stop message to the Rep on the Cell hosting an instance
Job: file_server VM: diego_brain	<ul style="list-style-type: none"> Serves static assets that can include general-purpose App Lifecycle binaries
Job: locket VM: diego_database	<ul style="list-style-type: none"> Provides a consistent key-value store for maintenance of distributed locks and component presence
Job: rep VM: diego_cell	<ul style="list-style-type: none"> Represents a Cell in Diego Auctions for Tasks and LRPCs Runs Tasks and LRPCs by creating a container and then running actions in it Periodically ensures its set of Tasks and ActualLRPs in the BBS is in sync with the containers actually present on the Cell Manages container allocations against resource constraints on the Cell, such as memory and disk space Streams stdout and stderr from container processes to the metron-agent running on the Cell, which in turn forwards to the Loggregator system Periodically collects container metrics and emits them to Loggregator Mediates all communication between the Cell and the BBS Maintains a presence record for the Cell in Locket

Job: route_emitter VM: diego_cell	<ul style="list-style-type: none"> Monitors <code>DesiredLRP</code> and <code>ActualLRP</code> states, emitting route registration and unregistration messages to Gorouter when it detects changes. Periodically emits the entire routing table to the Cloud Foundry Gorouter.
Job: ssh_proxy VM: diego_brain	<ul style="list-style-type: none"> Brokers connections between SSH clients and SSH servers Runs inside instance containers and authorizes access to app instances based on Cloud Controller roles

Additional Information

The following resources provide more information about Diego components:

- The [Diego Release repository](#) on GitHub.
- The [Auctioneer repository](#) on GitHub.
- The [Bulletin Board System repository](#) on GitHub.
- The [File Server repository](#) on GitHub.
- The [Rep repository](#) on GitHub.
- The [Executor repository](#) on GitHub.
- The [Route-Emitter repository](#) on GitHub.
- [Application SSH](#), [Application SSH Overview](#), and the [Diego SSH repository](#) on GitHub.

Components from Other Releases

The following table describes jobs that interact closely with Diego but are not part of the Diego Cloud Foundry BOSH release.

Component	Function
Job: boshs-dns-aliases VM: all	<ul style="list-style-type: none"> Provides service discovery through colocated DNS servers on all BOSH-deployed VMs Provides client-side load-balancing by randomly selecting a healthy VM when multiple VMs are available
Job: cc_uploader VM: diego_brain	<ul style="list-style-type: none"> Mediates uploads from the Executor to the Cloud Controller Translates simple HTTP POST requests from the Executor into complex multipart-form uploads for the Cloud Controller
Job: database VM: mysql	<ul style="list-style-type: none"> Provides a consistent key-value data store to Diego
Job: loggregator-agent VM: all	<ul style="list-style-type: none"> Forwards app logs, errors, and app and Diego metrics to the Loggregator Doppler component
Job: cloud_controller_clock VM: clock_global	<ul style="list-style-type: none"> Runs a Diego sync process to ensure desired app data in Diego is in sync with the Cloud Controller.

App Lifecycle Binaries

The following three platform-specific binaries deploy apps and govern their lifecycle:

- The **Builder**, which stages a CF app. The Builder runs as a Task on every staging request. It performs static analysis on the app code and does any necessary pre-processing before the app is first run.
- The **Launcher**, which runs a CF app. The Launcher is set as the Action on the `DesiredLRP` for the app. It executes the start command with the correct

system context, including working directory and environment variables.

- The [Healthcheck](#), which performs a status check on running CF app from inside the container. The Healthcheck is set as the Monitor action on the [DesiredLRP](#) for the app.

Current Implementations

- [Buildpack App Lifecycle](#) implements the Cloud Foundry buildpack-based deployment strategy.
- [Docker App Lifecycle](#) implements a Docker deployment strategy.

Additional Information

The following resources provide more information about components from other releases that interact closely with Diego:

- The [CC-Uploader repository](#) on GitHub.
- The [Garden](#) topic or the [Garden repository](#) on GitHub.
- The [Loggregator Release repository](#) on GitHub.
- The [BOSH DNS documentation](#).
- The [TPS repository](#) on GitHub.

Cloud Foundry Routing Architecture

This topic explains routing flow and architecture in Cloud Foundry. Cloud Foundry's routing tables must be updated before it can route external client requests to an app.

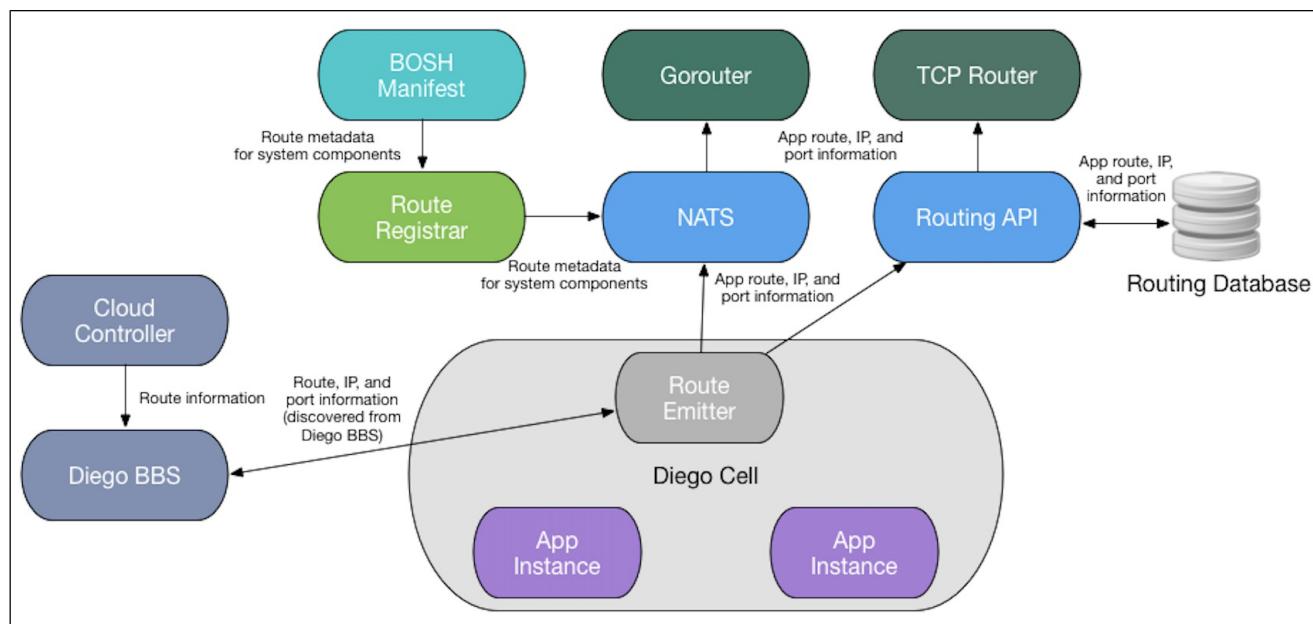
How Cloud Foundry Maintains Updated Routing Tables

Because each app can have many instances, one app route can go to multiple containers. As Diego moves and scales app instances, the route emitter on each cell sends messages to NATS or the Routing API to register and de-register routes to the cell's app instances.

The route emitter periodically emits the routes it discovers from Diego BBS to NATS and the Routing API as registration and unregistration messages every twenty seconds. By default, the Gorouter uses TLS to verify app identity and confirm that its routes are up-to-date. With TLS disabled, the Gorouter prunes any routes that the route emitters have not re-registered within the past two minutes. For more information about how CF maintains route consistency when TLS is or is not enabled, see the [Preventing Misrouting](#) section of the [HTTP Routing](#) topic.

The following process describes how a router obtains information about routes for an app running on Cloud Foundry:

1. The Cloud Controller component sends app route information to Diego BBS. For HTTP routing, route information includes the host and path of an external URL, as shown in the format of the [router.register message](#) in the Gorouter documentation on GitHub. For TCP routing, route information includes the route port on which the TCP connection was received; for more information, see the [Routing API documentation on GitHub](#).
2. Diego BBS coordinates the back end IP address and port where each instance of the app runs. When queried by the route emitter, the BBS sends this information along with Cloud Controller's app route information to the route emitter on the Diego cell where instances of the app are located.
3. If a route is HTTP, the route emitter on the Diego cells sends app route, IP, and port information to NATS, which then sends it to the Gorouter. If a route is TCP, the route emitter sends that information to the Routing API, which then sends it to the TCP router.
4. The Gorouter and TCP router use the route, IP, and port information from the route emitter to map incoming app requests to back end app instance locations.



Routing Table Recovery After Component Failure

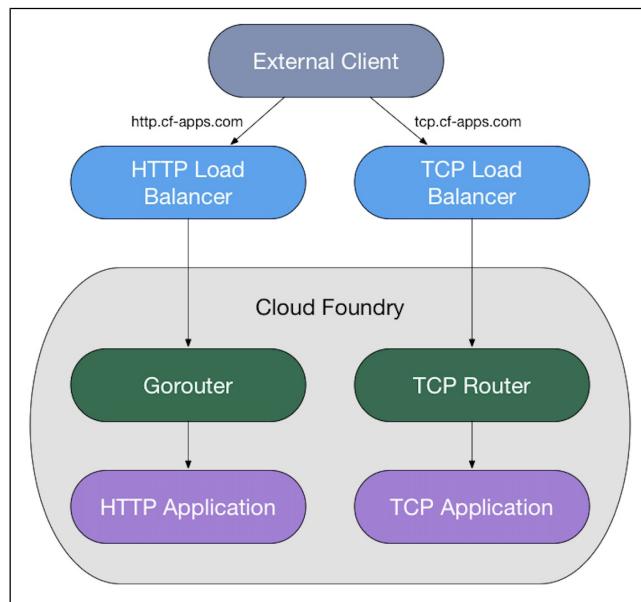
Cloud Controller and Diego BBS have their own databases, while NATS and the Gorouter only store their data in memory. If NATS or the Gorouter are restarted, they lose all of their data and must wait for the route emitter to send data to them again. If Diego BBS is restarted, it can retrieve its data from Cloud Controller.

If Cloud Controller is restarted, you must retrieve its data from a backup.

External Client Request Flow

When Cloud Foundry's routing tables are updated, Cloud Foundry can process requests from external clients to an app. The following process describes how an external client makes a request to an app running on Cloud Foundry:

1. The external client sends its request.
2. Your DNS service sends the request to the HTTP or TCP load balancer based on the prefix of the DNS name in the client request, such as `http` in `http.example.com`.
3. The load balancer sends the request to the load balancer's corresponding router.
4. The router sends the request to the app.



Routing Architecture Components

The following summarizes the roles and responsibilities of various components depicted in the Cloud Foundry routing architecture diagrams above. These summaries are limited to the roles and responsibilities these components have pertaining to routing. For more complete descriptions of these components, see [Cloud Foundry Concepts](#), [Cloud Foundry Components](#), and [Diego Components and Architecture](#).

Component Name	Summary
BOSH manifest	Used to configure route registrar with route(s) for system components such as UAA and Loggregator.
Cloud Controller	Contains route metadata, including whether they are HTTP or TCP.
Diego BBS	Contains IP and port metadata as well as route metadata from Cloud Controller, which route emitter discovers.
Diego cell	Manages app instances and tasks and long-running processes related to them. A route emitter runs on each cell.
Gorouter	Routes HTTP traffic coming into Cloud Foundry to the appropriate component. Receives route updates through NATS. Routes that have not been updated in two minutes are pruned from the Gorouter's database.
NATS	Receives routing configuration from route emitter and provides same to Gorouter.
Route registrar	Sends routing metadata described in BOSH manifest for system components such as UAA and Loggregator to NATS. This is because the Diego cell does not have information about system components, only about user spaces.
Route emitter	Periodically emits route, IP, and port metadata to NATS or Routing API as registration and unregistration messages. Does not know about app instances on Diego cell, but knows what cell it belongs to and learns about what app instances are running on its cell by asking Diego BBS for information about app instances on the same cell.
Routing API	Receives routing configuration from route emitter and other internal clients, and provides routing configuration for TCP router.
Routing database	Saves some routing data from Routing API. If the Gorouter misses a message about an unmapped route from NATS, it will not get it again, so TCP router and Routing API can consult routing database for current state of routes.

Component	Summary
Name	Routes TCP traffic coming into Cloud Foundry to the appropriate component. Receives route updates through the routing API.

Service Mesh Architecture

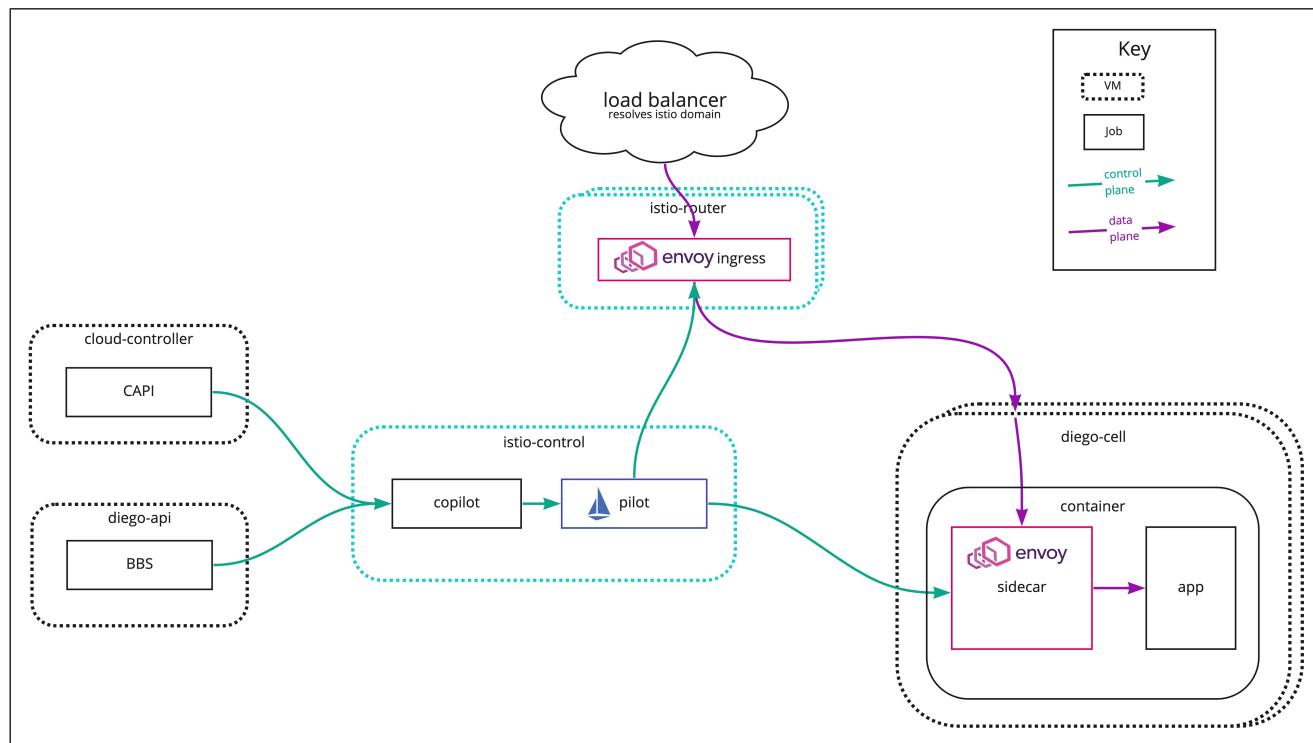
This topic describes the routing flow and architecture of the service mesh data and control plane in Pivotal Application Service (PAS).

Overview

The service mesh data plane is a parallel routing path for ingress traffic for apps on Pivotal Application Service. It is deployed alongside the existing Cloud Foundry routing tier and manages istio routes for applications.

We use [Istio's Pilot component](#) to configure ingress [Envoy Proxies](#), and these proxies are the routers. We use a custom component called Copilot to push Cloud Foundry configuration to into Pilot.

A route is managed by istio if it is associated with an istio managed domain. These are specified in the manifest.



Control Plane

1. A new route is added to CAPI and mapped to one or more applications
2. The route and mapping are sent to copilot
3. Copilot then exposes that configuration in a way Pilot can understand, Pilot polls for it
4. Pilot distributes the configuration to the ingress envys

Data Plane

1. The request hits your load balancer.
2. The load balancer directs the request to one of your ingress envys (on the istio-router vm)
3. The ingress envoy then chooses which app container to send the request to
4. The app container has an iptables rule which DNATs the request to its local envoy sidecar

5. The envoy sidecar passes the request along to the application

Service Mesh Components

The following table lists each component in the service mesh architecture and describes its function.

Component Name	Summary
CAPI	Cloud Controller receives API requests from the cf CLI and stores information about routes. It distributes this route information to Copilot.
BBS	BBS sends information about apps across all Diego cells to Copilot.
Copilot ↗	Copilot acts as an interface between Cloud Foundry routes and Istio configuration types. It sends configuration to Pilot through Mesh Configuration Protocol (MCP).
Pilot ↗	Pilot is an Istio component that can accept configuration from multiple sources simultaneously and distribute configuration intelligently across ingress and sidecar envoys.
Envoy ↗	Envoy Proxy is a lightweight edge proxy designed for microservices. It routes traffic based on configuration it receives from Pilot and emits in-depth metrics based on that traffic.
Load Balancer	The load balancer is a reverse proxy provided by the IaaS, or a physical machine, that distributes network traffic across the ingress envoys while presenting a single public endpoint. This is not the same load balancer used by Gorouter.
istio-release ↗	A BOSH release that deploys Istio-related components and configures any existing components to use them.

Cloud Controller

The Cloud Controller provides REST API endpoints for clients to access the system. The Cloud Controller maintains a database with tables for orgs, spaces, services, user roles, and more.

Diego Auction

The Cloud Controller uses the [Diego Auction](#) to balance application processes over the [cells](#) in a Cloud Foundry installation.

Database (CC_DB)

The Cloud Controller database has been tested with MySQL.

Blobstore

To stage and run apps, Cloud Foundry manages and stores the following types of binary large object (blob) files:

Blob Type	Description	Location in Blobstore
App Packages	Full contents of app directories, including source code and resource files, zipped into single blob files.	/cc-packages
Buildpacks	Buildpack directories, which Diego cells download to compile and stage apps with.	/cc-buildpacks
Resource Cache	Large files from app packages that the Cloud Controller stores with a SHA for later re-use. To save bandwidth, the Cloud Foundry Command Line Interface (cf CLI) only uploads large application files that the Cloud Controller has not already stored in the resource cache.	/cc-resources
Buildpack Cache	Large files that buildpacks generate during staging, stored for later re-use. This cache lets buildpacks run more quickly when staging apps that have been staged previously.	cc-droplets/buildpack_cache
Droplets	Staged apps packaged with everything needed to run in a container.	/cc-droplets

Cloud Foundry blobstores use the [Fog](#) Ruby gem to store blobs in services like Amazon S3, WebDAV, or the NFS filesystem. The file system location of an internal blobstore is `/var/vcap/store/shared`.

A single blobstore typically stores all five types of blobs, but you can configure the Cloud Controller to use separate blobstores for each type.

Automatic Blob Cleanup

After a blob deletion fails silently or something else goes wrong, the blobstore may contain blobs that the Cloud Controller no longer needs or lists in its database. These are called orphan blobs, and they waste blobstore capacity.

The Cloud Controller detects and removes orphan blobs by scanning part of the blobstore daily and checking for any blobs that its database does not account for. The process scans through the entire blobstore every week, and only removes blobs that show as orphans for three consecutive days.

The Cloud Controller performs this automatic cleanup when the `cloud_controller_worker` job property `cc.perform_blob_cleanup` is set to `true`.

Manual Blob Cleanup

The Cloud Controller does not track resource cache and buildpack cache [blob types](#) in its database, so it does not [clean them up automatically](#) as it does with app package, buildpack, and droplet type blobs.

To clean up the buildpack cache, admin users can run `cf curl -X DELETE /v2/blobstores/buildpack_cache`. This empties the buildpack cache completely, which is a safe operation.

To clean up the resource cache, delete it as follows:

- **Internal blobstore:** Run `bosh ssh` to connect to the blobstore VM (NFS or WebDav) and `rm *` the contents of the `/var/vcap/store/shared/cc-resources` directory.
- **External blobstore:** Use the file store's API to delete the contents of the `resources` bucket.

Do not manually delete app package, buildpack, or droplet blobs from the blobstore. To free up resources from those locations, run `cf delete-buildpack` for buildpacks or `cf delete` for app packages and droplets.

Testing

By default `rspec` runs a test suite with the SQLite in-memory database. Specify a connection string using the `DB_CONNECTION` environment variable to test against MySQL. For example:

```
DB_CONNECTION="mysql2://root:password@localhost:3306/ccng" rspec
```

Cloud Controller Blobstore

This topic describes how Cloud Controllers and Diego interact with the CC's blobstore.

Overview

Cloud Foundry uses a blobstore to store the source code that developers push, stage, and run.

This topic references staging and treats all blobstores as generic (S3-style) object stores.

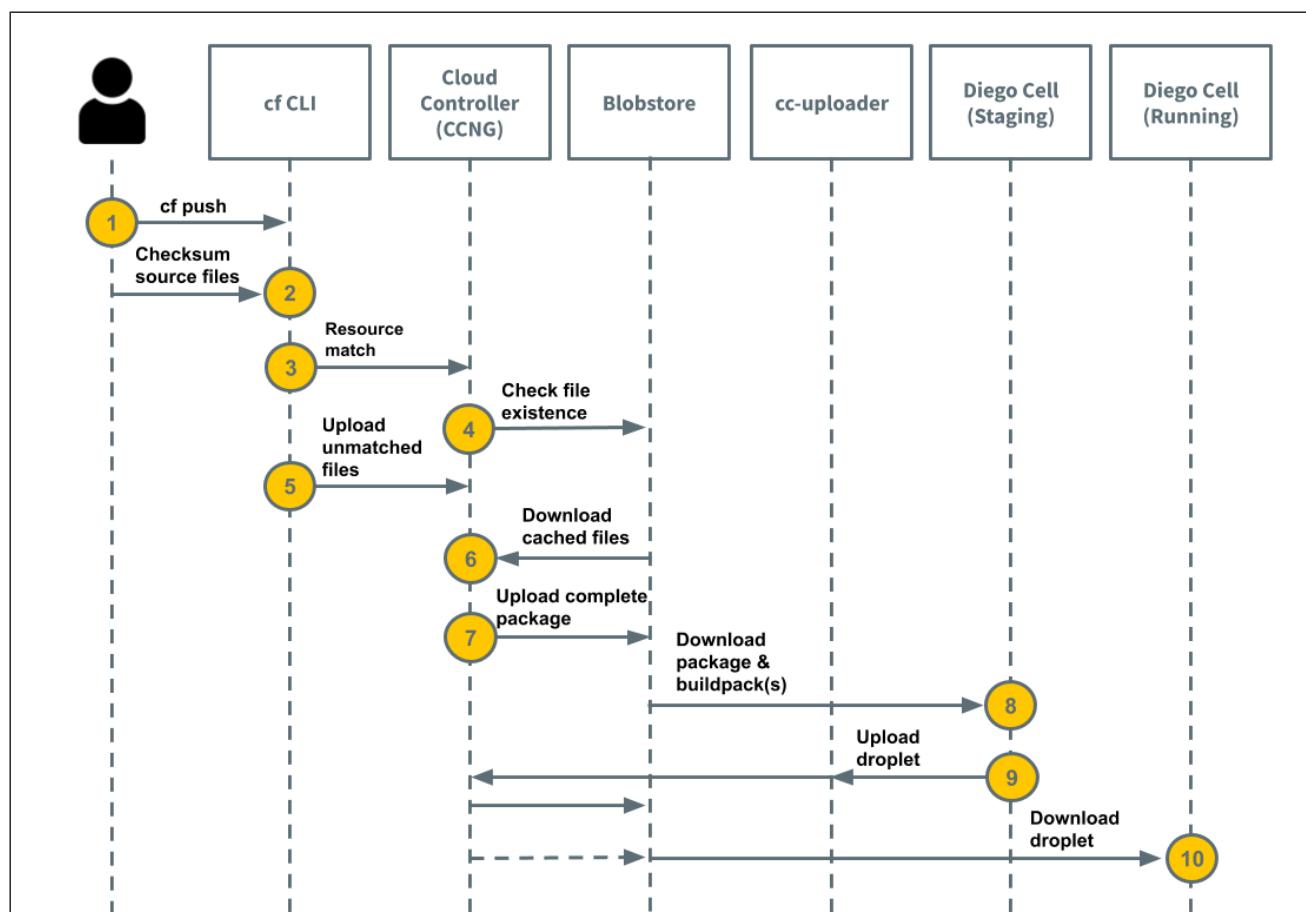
For more general information about staging, see [How Apps Are Staged](#).

For more information about how specific 3rd party blobstores can be configured, see [Configuring File Storage for PAS](#).

How staging uses the blobstore

This section describes how staging buildpack apps utilizes the blobstore.

The following diagram illustrates, in-depth, how the staging process uses the blobstore:



For a description of each step where staging a buildpack app interacts with the blobstore, see the following:

1. A developer runs `cf push`.
2. The cf CLI gets all the developer's local source code files and computes a checksum of each.
3. The cf CLI makes a "Resource Match" request to the cloud controller, listing those files and their checksums.
4. This step, "Check file existence", includes the following:
 - The cf CLI sends a "Check file existence" request to the Blobstore.
 - The Blobstore checks if the files exist and returns the results to the cf CLI.

- a. The Cloud Controller makes a series of “HEAD”-type requests to the blobstore to find out which files it has cached.
 - b. Cloud Controller content-addresses its cached files so that changes to a file result in it being stored as a different object.
 - c. Eventually Cloud Controller is able to compute which files it has and which it needs the CLI to upload.
 - d. In response to the resource match request, Cloud Controller lists the files the CLI needs to upload .
5. The CLI compresses and uploads the unmatched files to the Cloud Controller.
 6. Asynchronously, The Cloud Controller downloads the matched files from the blobstore to its local disk.
 7. This step includes the following:
 - a. The Cloud Controller zips together the freshly uploaded files with the downloaded cached files.
 - b. The Cloud Controller uploads the complete package to the blobstore.
 8. The Diego Cell downloads the package and its buildpack(s) into a container and stages the app.
 9. This step includes the following:
 - a. After the app has been staged, the Diego Cell uploads the complete droplet to the cc-uploader.
 - b. The cc-uploader makes a multipart upload request to upload the droplet to the Cloud Controller.
 - c. The Cloud Controller enqueues an asynchronous job to upload the blobstore.
 10. This step includes the following:
 - a. The Diego Cell attempts to download the droplet from the Cloud Controller into the app’s container so it can run it.
 - b. The Cloud Controller redirects the Cell’s droplet download request to the blobstore.
 - c. The Diego Cell downloads the app’s droplet from the blobstore and runs it.

Blobstore load

The shape of the load CC generates on its blobstore is a bit unusual because of resource matching. Many blobstores that perform very well under normal read, write, and delete load are overwhelmed by CC’s heavy usage of HEAD requests during resource matching.

Pushing an app with large number of files will cause the CC to check the blobstore for the existence of each and every one of those files.

Parallel BOSH deploys of Diego Cells can also generate significant read load on CC’s Blobstore as the cells perform [evacuation](#).

How CC reaps expired packages, droplets, and buildpacks

As new droplets and packages are created, the oldest ones associated with an app are marked as “EXPIRED” if they exceed the configured limits for packages and droplets stored per app.

Nightly, starting at midnight, the Cloud Controller runs a series of jobs to delete the data associated with expired packages, droplets, and buildpacks.

Enabling the native [versioning](#) feature on your blobstore will increase the number of resources stored and increase costs.

Blobstore Interaction timeouts

The Cloud Controller inherits its default blobstore operation timeouts from [excon](#). Excon defaults to a 60 second read, write, and connect timeouts.

Messaging (NATS)

This information was adapted from the [NATS](#) README. NATS is a lightweight publish-subscribe and distributed queueing messaging system written in Ruby. Ops Manager sends all NATS traffic using Transport Layer Security (TLS) encryption by default.

Install NATS

```
$ gem install nats
# or
$ rake geminstall

$ nats-sub foo &
$ nats-pub foo 'Hello World!'
```

Basic Usage

```
require "nats/client"

NATS.start do

  # Simple Subscriber
  NATS.subscribe('foo') { |msg| puts "Msg received : '#{msg}'" }

  # Simple Publisher
  NATS.publish('foo.bar.baz', 'Hello World!')

  # Unsubscribing
  sid = NATS.subscribe('bar') { |msg| puts "Msg received : '#{msg}'" }
  NATS.unsubscribe(sid)

  # Requests
  NATS.request('help') { |response| puts "Got a response: '#{response}'" }

  # Replies
  NATS.subscribe('help') { |msg, reply| NATS.publish(reply, "I'll help!") }

  # Stop using NATS.stop, exits EM loop if NATS.start started the loop
  NATS.stop

end
```

Wildcard Subscriptions

```
# "*" matches any token, at any level of the subject.
NATS.subscribe('foo.*.baz') { |msg, reply, sub| puts "Msg received on [#{sub}] : '#{msg}'" }
NATS.subscribe('foo.bar.*') { |msg, reply, sub| puts "Msg received on [#{sub}] : '#{msg}'" }
NATS.subscribe('*.*.bar.*') { |msg, reply, sub| puts "Msg received on [#{sub}] : '#{msg}'" }

# ">" matches any length of the tail of a subject and can only be the last token
# E.g. 'foo.>' will match 'foo.bar', 'foo.bar.baz', 'foo.foo.bar.bax.22'
NATS.subscribe('foo.>') { |msg, reply, sub| puts "Msg received on [#{sub}] : '#{msg}'" }
```

Queues Groups

```
# All subscriptions with the same queue name will form a queue group
# Each message will be delivered to only one subscriber per queue group, queuing semantics
# You can have as many queue groups as you want
# Normal subscribers will continue to work as expected.
NATS.subscribe(subject, :queue => 'job.workers') { |msg| puts "Received '#{msg}'" }
```

Advanced Usage

```
# Publish with closure, callback fires when server has processed the message
NATS.publish('foo', 'You done?') { puts 'msg processed!' }

# Timeouts for subscriptions
sid = NATS.subscribe('foo') { received += 1 }
NATS.timeout(sid, TIMEOUT_IN_SECS) { timeout_recd = true }

# Timeout unless a certain number of messages have been received
NATS.timeout(sid, TIMEOUT_IN_SECS, :expected => 2) { timeout_recd = true }

# Auto-unsubscribe after MAX_WANTED messages received
NATS.unsubscribe(sid, MAX_WANTED)

# Multiple connections
NATS.subscribe('test') do |msg|
  puts "received msg"
  NATS.stop
end

# Form second connection to send message on
NATS.connect { NATS.publish('test', 'Hello World!') }
```

Gorouter

Gorouter routes traffic coming into Cloud Foundry to the appropriate component, whether the request comes from an operator addressing the [Cloud Controller](#) or from an application user accessing an app running on a Diego Cell. Handling both platform and app requests with the same process centralizes routing logic and simplifies support for WebSockets and other types of traffic (for example, through HTTP CONNECT).

Refer to the following instructions for help getting started with Gorouter in a standalone environment.

Setup

```
$ git clone https://github.com/cloudfoundry/gorouter.git
$ cd gorouter
$ git submodule update --init
$ ./bin/go install gorouter/gorouter
$ gem install nats
```

Start

```
# Start NATS server in daemon mode
$ nats-server -d

# Start gorouter
$ ./bin/gorouter
```

Usage

Gorouter receives route updates through [NATS](#). By default, routes that have not been updated in two minutes are pruned. Therefore, to maintain an active route, you must ensure that the route is updated at least every two minutes. The format of these route updates is as follows:

```
{
  "host": "127.0.0.1",
  "port": 4567,
  "uris": [
    "my_first_url.vcap.me",
    "my_second_url.vcap.me"
  ],
  "tags": {
    "another_key": "another_value",
    "some_key": "some_value"
  }
}
```

Such a message can be sent to both the `gorouter.register` subject to register URIs, and to the `gorouter.unregister` subject to unregister URIs, respectively.

```
$ nohup ruby -rsinatra -e 'get("/") { "Hello!" }' &
$ nats-pub 'gorouter.register' '{"host":"127.0.0.1","port":4567,
  "uris":["my_first_url.vcap.me","my_second_url.vcap.me"],
  "tags":{"another_key":"another_value","some_key":"some_value"}}'
Published [gorouter.register] : {"host":"127.0.0.1","port":4567,
  "uris":["my_first_url.vcap.me","my_second_url.vcap.me"],
  "tags":{"another_key":"another_value","some_key":"some_value"}}
$ curl my_first_url.vcap.me:8080
Hello!
```

Instrumentation

Gorouter provides `/varz` and `/healthz` http endpoints for monitoring.

The `/routes` endpoint returns the entire routing table as JSON. Each route has an associated array of `host:port` entries.

All of the endpoints require http basic authentication, credentials for which you can acquire through NATS. You can explicitly set the `port`, `user` and `password` (`pass` is the config attribute) in the `gorouter.yml` config file `status` section.

```
status:
  port: 8080
  user: some_user
  pass: some_password
```

Example interaction with `curl`:

```
$ curl -vvv "http://someuser:somepass@127.0.0.1:8080/routes"
* About to connect() to 127.0.0.1 port 8080 (#0)
*   Trying 127.0.0.1...
* Connected
* Connected to 127.0.0.1 (127.0.0.1) port 8080 (#0)
* Server auth using Basic with user 'someuser'
> GET /routes HTTP/1.1
> Authorization: Basic c29tZXVzZXI6c29tZXBlc3M=
> User-Agent: curl/7.24.0 (x86_64-apple-darwin12.0) libcurl/7.24.0 OpenSSL/0.9.8r zlib/1.2.5
> Host: 127.0.0.1:8080
> Accept: */*
>
< HTTP/1.1 200 OK
< Content-Type: application/json
< Date: Mon, 25 Mar 2013 20:31:27 GMT
< Transfer-Encoding: chunked
<
{"0295dd314aaf582f201e655cbd74ade5.cloudfoundry.me":["127.0.0.1:34567"],
 "03e316d6aa375d1dc1153700da5f1798.cloudfoundry.me":["127.0.0.1:34568"]}
```

Logs

This section provides details about Gorouter logging.

Levels

The following table describes the log levels supported by Gorouter. The log level is set to `debug` and is not configurable.

Message	Description	Examples
<code>fatal</code>	Gorouter is unable to handle any requests due to a fatal error.	Gorouter cannot bind to its TCP port, a CF component has published invalid data to Gorouter.
<code>error</code>	An unexpected error has occurred.	Gorouter failed to fetch token from UAA service.
<code>info</code>	An expected event has occurred.	Gorouter started or exited, Gorouter has begun to prune routes for stale droplets.
<code>debug</code>	A lower-level event has occurred.	Route registration, route unregistration.

Message Contents

This section provides a sample Gorouter log message and an explanation of the contents.

```
[2017-02-01 22:54:08+0000] {"log_level":0,"timestamp":1485989648.0895808,"message":"endpoint-registered","source":"vcap.Gorouter.registry","data":{"uri":"0-* .login.bosh-lite.com","backend":"10.123.0.134:8080","modification_tag":{"guid":"","index":0}}}
```

Property	Description
<code>log_level</code>	Logging level of the message
<code>timestamp</code>	Epoch time of the log
<code>message</code>	Content of the log line
<code>source</code>	Gorouter function that initiated the log message
<code>data</code>	Additional information that varies based on the message

About Access Logs

Access logs provide information for the following fields when receiving a request:

```
<Request Host> - [<Start Date>] "<Request Method> <Request URL> <Request Protocol>" <Status Code> <Bytes Received> <Bytes Sent> "<Referer>" "<User-Agent>" <Remote Address>
<Backend Address> x_forwarded_for:<X-Forwarded-For> x_forwarded_proto:<X-Forwarded-Proto> vcap_request_id:<X-Vcap-Request-ID> response_time:<Response Time> app_id:
<Application ID> app_index:<Application Index> <Extra Headers>
```

The following are optional fields: Status Code, Response Time, Application ID, Application Index, and Extra Headers.

If the access log lacks a Status Code, Response Time, Application ID, or Application Index, the corresponding field shows -.

 **Note:** Access logs are also redirected to syslog.

User Account and Authentication (UAA) Server

This topic provides an overview of the User Account and Authentication (UAA) Server, the identity management service for Cloud Foundry.

The primary role of the UAA is as an OAuth2 provider, issuing tokens for client apps to use when they act on behalf of Cloud Foundry users. In collaboration with the login server, the UAA can authenticate users with their Cloud Foundry credentials, and can act as an SSO service using those, or other, credentials.

The UAA has endpoints for managing user accounts and for registering OAuth2 clients, as well as various other management functions.

Quick Start

You can deploy the UAA locally or to Cloud Foundry.

- [Deploy UAA Locally](#)
- [Deploy UAA to Cloud Foundry](#)

Deploy UAA Locally

Follow the instructions below to deploy and run the UAA locally.

1. In a terminal window, clone the UAA GitHub repository.

```
$ git clone git://github.com/cloudfoundry/uaa.git
```

2. Navigate to the directory where you cloned the UAA GitHub repository, then run the `./gradlew run` command to build and run all the components that comprise the UAA and the example programs, `uaa`, `samples/api`, and `samples/app`.

```
$ cd uaa  
$ ./gradlew run
```

3. If successful, the three apps run together on a single instance of Tomcat listening on port 8080, with endpoints `/uaa`, `/app`, and `/api`.

Use Local UAA

Follow the steps below to access and use a locally-deployed UAA server.

1. Run the UAA server as described in the [Deploy Locally](#) section, above.
2. Open another terminal window. From the project base directory, run `curl localhost:8080/uaa/info -H "Accept: application/json"` to confirm the UAA is running. You should see basic information about the system. For example:

```
$ curl localhost:8080/uaa/info -H "Accept: application/json"
{
  "app": {
    "version": "4.19.0"
  },
  "links": {
    "uaa": "http://localhost:8080/uaa",
    "passwd": "forgot_password",
    "login": "http://localhost:8080/uaa",
    "register": "create_account"
  },
  "zone_name": "uaa",
  "entityID": "cloudfoundry-saml-login",
  "commit_id": "af93628",
  "idpDefinitions": {},
  "prompts": {
    "username": [
      "text",
      "Email"
    ],
    "password": [
      "password",
      "Password"
    ]
  },
  "timestamp": "2018-05-25T15:34:31-0700"
}
```

3. Run `gem install cf-uaac` to install the Cloud Foundry UAA Command Line Client (UAAC) Ruby gem.

```
$ gem install cf-uaac
```

4. Run `uaac target http://localhost:8080/uaa` to target the local UAA server endpoint.

```
$ uaac target http://localhost:8080/uaa
```

5. Run `uaac token client get CLIENT_NAME -s CLIENT_SECRET` to obtain an access token. Replace `CLIENT_NAME` and `CLIENT_SECRET` with actual values. For example, when starting up the UAA locally for development there should be a predefined `admin client` you can use:

```
$ uaac token client get admin -s adminsecret
```

If you do not specify the client secret, UAAC will show an interactive prompt where you must enter the client secret value.

The `uaac token client get` command requests an access token from the server using the OAuth2 `client credentials` grant type.

6. View your UAAC token context. When UAAC obtains a token, the token and other metadata is stored in the `~/.uaac.yml` file on your local machine. To view the token you have obtained, use the command `uaac context`. For example:

```
$ uaac context
[0]*[http://localhost:8080/uaa]

[0]*[admin]
client_id: admin
access_token: eyJhbGciOiJIUzI1NiIsImtpZCI6ImxJZ2FjeS10b2tbi1rZXkiLCJ0eXAiOiJKV1QiQifQ.eyJqdGkiOiIxOWYyNWU2Y2E5Y2M0ZWlyYTdmNTAxNmU0NDFjZThkNCIsIn
token_type: bearer
expires_in: 43199
scope: clients.read clients.secret clients.write uaa.admin clients.admin scim.write scim.read
jti: 19f256eca9cc4eb2a7f5016e441ce8d4
```

Copy the access token from this output for the following step.

7. Run `uaac token decode ACCESS-TOKEN-VALUE` to view information in the token, which is encoded using the [JSON Web Token](#) format. Replace `ACCESS-TOKEN-VALUE` with your access token, copied from the `uaac context` output. The UAAC should display all the claims inside the token body. For example:

```
$ uaac token decode eyJhbGciOiJIUzI1NiIsImtpZCI6ImxlZ2FjeS10b2tlbi1rZXkiLCJ0eXAiOiJKV1QiAQ.eyJqdGkiOiIxOWYyNWU2Y2E5Y2M0ZWlyYTdmNTAxNmU0NDEjZThkNC
Note: no key given to validate token signature

jti: 19f25e6ca9cc4eb2a7f5016e441ce8d4
sub: admin
authorities: clients.read clients.secret clients.write uaa.admin clients.admin scim.write scim.read
scope: clients.read clients.secret clients.write uaa.admin clients.admin scim.write scim.read
client_id: admin
cid: admin
azp: admin
grant_type: client_credentials
rev_sig: 267199d1
iat: 1528232017
exp: 1528275217
iss: http://localhost:8080/uaa/oauth/token
zid: uaa
aud: scim clients uaa admin
```

Deploy UAA to Cloud Foundry

Follow the instructions below to build the UAA as an app and push it to Cloud Foundry using the Cloud Foundry Command Line Interface (cf CLI).

1. In a terminal window, clone the UAA GitHub repository.

```
$ git clone git://github.com/cloudfoundry/uaa.git
```

2. Navigate to the directory where you cloned the UAA GitHub repository, then run the `./gradlew :cloudfoundry-identity-uaa:war` command build the UAA as a WAR file.

```
$ cd uaa
$ ./gradlew :cloudfoundry-identity-uaa:war
```

3. Run the cf CLI `cf push APP-NAME -m 512M -p PATH-TO-WAR-FILE --no-start` command to push the app to Cloud Foundry. Replace `APP-NAME` with a name for your UAA app, and `PATH-TO-WAR-FILE` with the path to the WAR file you created in the previous step. For example:

```
$ cf push MYUAA -m 512M -p uaa/build/libs/cloudfoundry-identity-uaa-1.8.0.war --no-start
```

4. Run `cf set-env APP-NAME SPRING_PROFILES_ACTIVE default` to set the `SPRING_PROFILES_ACTIVE` environment variable with the value `default`. Replace `APP-NAME` with the name of your app that you used in the previous step. For example:

```
$ cf set-env MYUAA SPRING_PROFILES_ACTIVE default
```

5. Run `cf start APP-NAME` to start your app. Replace `APP-NAME` with the name of your app. For example:

```
$ cf start MYUAA
```

Use Remote UAA

You use a UAA server that you pushed as an app to Cloud Foundry in a similar way to one you [run locally](#). You do not need app token encoding because you do not have the client secret.

Follow the steps below to access and use a UAA server that you pushed as an app to Cloud Foundry.

1. [Deploy UAA to Cloud Foundry](#) as described above.
2. From the project base directory, run `curl -H "Accept: application/json" APP-FQDN/login` to query the external login endpoint about the system. Replace `APP-FQDN` with the FQDN of your app. For example:

```
$ curl -H "Accept: application/json" uaa.example.org/login
{
  "timestamp": "2014-09-15T18:25:04+0000",
  "app": {"version": "1.8.3"},
  "commit_id": "git-metadata-not-found",
  "prompts": {"username": ["text", "Email"],
              "password": ["password", "Password"]}
}
```

3. Run `gem install cf-uaac` to install the Cloud Foundry UAA Command Line Client (UAAC) Ruby gem.

```
$ gem install cf-uaac
```

4. Run `uaac target APP-FQDN` to target the remote UAA Server endpoint. Replace `APP-FQDN` with the FQDN of your app.

```
$ uaac target uaa.example.org
```

5. Run `uaac token get USERNAME PASSWORD` to log in. Replace `USERNAME` with your user name, and `PASSWORD` with your password. For example:

```
$ uaac token get marissa koala
```

If you do not specify a username and password, the UAAC prompts you to supply them.

The `uaac token client get` command authenticates and obtains an access token from the server using the OAuth2 implicit grant, similar to the approach intended for a standalone client like the Cloud Foundry Command Line Interface (cf CLI).

Integration Tests

Run the integration tests with the following command:

```
$ ./gradlew integrationTest
```

This command starts a UAA server running in a local Apache Tomcat instance. By default, the service URL is set to `http://localhost:8080/uaa`.

You can set the environment variable `CLOUD_FOUNDRY_CONFIG_PATH` to a directory containing a `uaa.yml` file where you change the URLs used in the tests, and where you can set the UAA server context root.

Custom YAML Configuration

Follow the steps below to modify the runtime parameters.

1. Create a `uaa.yml` file in the following format:

```
uaa:
  host: UAA-HOSTNAME
  test:
    username: USERNAME
    password: PASSWORD
    email: EMAIL-ADDRESS
```

Replace the values in the above format as follows:

- **UAA-HOSTNAME:** The FQDN of UAA app. Example: `uaa.example.org`
- **USERNAME:** A valid username. Example: `dev@example.org`
- **PASSWORD:** Password for the above username.
- **EMAIL-ADDRESS:** Email address for the above user. Example: `dev@example.org`

2. From the `uaa/uaa` directory, run `CLOUD_FOUNDRY_CONFIG_PATH=/tmp ./gradlew test`.

3. The web app looks for a YAML file in the following locations when it starts, with later entries overriding earlier ones:

```
classpath:uaa.yml
file:${CLOUD_FOUNDRY_CONFIG_PATH}/uaa.yml
file:${UAA_CONFIG_FILE}
${UAA_CONFIG_URL}
```

Test with PostgreSQL or MySQL

The default UAA unit tests, `./gradlew test`, use HyperSQL (hsqldb).

To use a different database management system, create a `uaa.yml` file containing `spring_profiles: default,OTHER-DBMS` in the `src/main/resources/` directory. Replace `OTHER-DBMS` with the name of the other database management system to use.

For example, run the following command to run the unit tests using PostgreSQL instead of hsqldb:

```
$ echo "spring_profiles: default,postgresql" > src/main/resources/uaa.yml
$ ./gradlew test integrationTest
```

Run the following command to run the unit tests using MySQL instead of hsqldb:

```
$ echo "spring_profiles: default,mysql" > src/main/resources/uaa.yml
$ ./gradlew test integrationTest
```

You can find the database configuration for the common and scim modules at `common/src/test/resources/(mysql|postgresql).properties` and `scim/src/test/resources/(mysql|postgresql).properties`.

UAA Projects

The following UAA projects exist:

- `common`: A module containing a JAR with all the business logic. `common` is used in the web apps listed below.
- `uaa`: The UAA server. `uaa` provides an authentication service and authorized delegation for back-end services and apps by issuing OAuth2 access tokens.
- `api`: A sample OAuth2 resource service that returns a mock list of deployed apps. `api` provides resources that other apps might want to access on behalf of the resource owner.
- `app`: A sample user app that uses both `api` and `uaa`. `app` is a web app that requires single sign-on and access to the `api` service on behalf of users.
- `scim`: The [SCIM](#) user management module used by UAA.

UAA Server

The authentication service, `uaa`, is a Spring MVC web app. You can deploy it in Tomcat or your container of choice, or execute `./gradlew run` to run it directly from the `uaa` directory in the source tree. When run with Gradle, `uaa` listens on port `8080` and has URL `http://localhost:8080/uaa`.

The UAA Server supports the APIs defined in the UAA-APIs document which include the following:

- The OAuth2 `/authorize` and `/token` endpoints
- A `/login_info` endpoint to allow querying for required login prompts
- A `/check_token` endpoint to allow resource servers to obtain information about an access token submitted by an OAuth2 client
- SCIM user provisioning endpoint
- OpenID connect endpoints to support authentication `/userinfo` and `/check_id`

Command line clients can perform authentication by submitting credentials directly to the `/authorize` endpoint.

An `ImplicitAccessTokenProvider` exists in Spring Security OAuth to use if your client is Java.

By default, `uaa` will launch with a context root `/uaa`.

Configuration

A `uaa.yml` file exists in the app. This file provides defaults to the placeholders in the Spring XML.

You can override any occurrences of `${placeholder.name}` in the XML by adding it to the `uaa.yml` file, or by providing a System property, `-D` to JVM, with the same name.

All passwords and client secrets in the config files are in plain text, but are inserted into the UAA database encrypted with BCrypt.

User Account Data

The default uses an in-memory RDBMS user store, pre-populated with a single test user `marissa` with password `koala`.

To use PostgreSQL for user data, activate the Spring profile `postgresql`.

You can configure the active profiles in `uaa.yml` using the following:

```
spring_profiles: postgresql,default
```

To use PostgreSQL instead of HyperSQL (hsqldb):

```
$ echo "spring_profiles: postgresql,default" > src/main/resources/uaa.yml  
$ ./gradlew run
```

To bootstrap a microcloud-type environment, you need an admin client. A database initializer component that inserts an admin client exists. If the default profile is active, for example not `postgresql`, a cf CLI client exists so that the Gem login works with no additional configuration required. You can override the default settings and add additional clients in the `uaa.yml` file:

```
oauth:  
clients:  
admin:  
authorized-grant-types: client_credentials  
scope: read,write,password  
authorities: ROLE_CLIENT,ROLE_ADMIN  
id: admin  
secret: adminclientsecret  
resource-ids: clients
```

You can use the admin client to create additional clients. You must have a client with read/write access to the `scim` resource to create user accounts. The integration tests handle this automatically by inserting client and user accounts as necessary for the tests.

Sample Apps

Two sample apps are included with the UAA: `/api` and `/app`.

Run them with `./gradlew run` from the `uaa` root directory. All three apps, `uaa`, `/api`, and `/app`, are simultaneously deployed.

API Sample App

The `api` sample app is an example resource server. It hosts a service that returns a list of mock apps under `/apps`.

APP Sample App

The `app` sample app is a user interface app, primarily aimed at browsers, that uses OpenID Connect for authentication and OAuth2 for access grants. `app` authenticates with the Auth service, then accesses resources in the API service. Run it with `./gradlew run` from the `uaa` root directory.

The app can operate in multiple different profiles according to the location and presence of the UAA server and the login app. By default, the app looks for a UAA on `localhost:8080/uaa`, but you can change this by setting the `UAA_PROFILE` environment variable or System Property.

The app source code, `samples/app/src/main/resources`, contains multiple properties files pre-configured with different likely locations for those servers. The names of these properties files follow the form `app-UAA_PROFILE.properties`.

The naming convention for the `UAA_PROFILE` is as follows:

- `local`: a localhost deployment
- `vcap`: a `vcap.me` deployment
- `staging`: a staging deployment

Profile names can be hyphenated to indicate multiple contexts. For example, `local-vcap` can be used when the login server is in a different location than the UAA server.

Use Cases

1. See all apps:

```
GET /app/apps
```

Browser is redirected through a series of authentication and access grant steps, and then the photos are shown.

2. See the currently logged in user details, a selection of attributes from the OpenID provider:

```
GET /app
```

LOGIN App

The `login` app is a user interface for authentication. The UAA can also authenticate user accounts, but only if it manages them itself, and it only provides a basic UI. You can brand and customize the login app for non-native authentication and for more complicated UI flows, like user registration and password reset.

The login app is itself an OAuth2 endpoint provider, but delegates those features to the UAA server. Therefore, configuration for the login app consists of locating the UAA through its OAuth2 endpoint URLs and registering the login app itself as a client of the UAA. A `login.yml` for the UAA locations exists, such as for a local `vcap` instance:

```
uaa:  
  url: http://uaa.vcap.example.net  
  token:  
    url: http://uaa.vcap.example.net/oauth/token  
  login:  
    url: http://uaa.vcap.example.net/login.do
```

An environment variable or Java System property also exists, `LOGIN_SECRET`, for the client secret that the app uses when it authenticates itself with the UAA. The login app is registered by default in the UAA only if there are no active Spring profiles. In the UAA, the registration is located in the `oauth-clients.xml` config file:

```
id: login  
secret: loginsecret  
authorized-grant-types: client_credentials  
authorities: ROLE_LOGIN  
resource-ids: oauth
```

Use Cases

1. Authenticate:

```
GET /login
```

The sample app presents a form login interface for the backend UAA, and an OpenID widget where a user can authenticate using Google or other credentials.

2. Approve OAuth2 token grant:

GET /oauth/authorize?client_id=app&response_type=code...

Standard OAuth2 Authorization Endpoint. The UAA handles client credentials and all other features in the back end, and the login app is used to render the UI.

3. Obtain access token:

POST /oauth/token

Standard OAuth2 Authorization Endpoint passed through to the UAA.

Scopes

UAA covers multiple scopes of privilege, including access to UAA, access to [Cloud Controller](#), and access to the [router](#).

See the tables below for a description of the scopes covered by UAA:

- [UAA Scopes](#)
- [Cloud Controller Scopes](#)
- [Router Scopes](#)
- [Other Scopes](#)

UAA Scopes

Scope	Description
uaa.user	This scope indicates that this is a user. It is required in the token if submitting a GET request to the OAuth 2 /authorize endpoint.
uaa.none	This scope indicates that this client will not be performing actions on behalf of a user.
uaa.admin	This scope indicates that this is the superuser.
scim.write	This scope gives admin write access to all SCIM endpoints, /Users, and /Groups.
scim.read	This scope gives admin read access to all SCIM endpoints, /Users, and /Groups.
scim.create	This scope gives the ability to create a user with a POST request to the /Users endpoint, but not to modify, read, or delete users.
scim.userids	This scope is required to convert a username and origin into a user ID and vice versa.
scim.invite	This scope is required to participate in invitations using the /invite_users endpoint.
groups.update	This scope gives the ability to update a group. This ability can also be provided by the broader scim.write scope.
password.write	This admin scope gives the ability to change a user's password.
openid	This scope is required to access the /userinfo endpoint. It is intended for OpenID clients.
idps.read	This scope gives read access to retrieve identity providers from the /identity-providers endpoint.
idps.write	This scope gives the ability to create and update identity providers from the /identity-providers endpoint.
clients.admin	This scope gives the ability to create, modify, and delete clients.
clients.write	This scope is required to create and modify clients. The scopes are prefixed with the scope holder's client ID. For example, id:testclient authorities:client.write gives the ability to create a client that has scopes with the testclient. prefix. Authorities are limited to uaa.resource.
clients.read	This scope gives the ability to read information about clients.
clients.secret	This admin scope is required to change the password of a client.
zones.read	This scope is required to invoke the /identity-zones endpoint to read identity zones.
zones.write	This scope is required to invoke the /identity-zones endpoint to create and update identity zones.
scim.zones	This is a limited scope that only allows adding a user to, or removing a user from, zone management groups under the path /Groups/zones.

<code>oauth.approval</code>	<code>/approvals endpoint</code> . This scope is required to approve or reject clients to act on a user's behalf. This is a default scope defined in the <code>uaa.yml</code> file.
<code>oauth.login</code>	This scope is used to indicate a login app, such as external login servers, can perform trusted operations, such as creating users not authenticated in the UAA.
<code>approvals.me</code>	This scope is not currently used.
<code>uaa.resource</code>	This scope indicates that this is a resource server, used for the <code>/check_token</code> endpoint.
<code>zones.ZONE-ID.admin</code>	This scope permits operations in a designated zone, such as creating identity providers or clients in another zone, by authenticating against the default zone. This scope is used with the <code>X-Identity-Zone-Id header</code> .
<code>zones.ZONE-ID.read</code>	This scope permits reading the given identity zone. This scope is used with the <code>X-Identity-Zone-Id header</code> .
<code>zones.ZONE-ID.clients.admin</code>	This scope translates into <code>clients.admin</code> after zone switch completes. This scope is used with the <code>X-Identity-Zone-Id header</code> .
<code>zones.ZONE-ID.clients.read</code>	This scope translates into <code>clients.read</code> after zone switch completes. This scope is used with the <code>X-Identity-Zone-Id header</code> .
<code>zones.ZONE-ID.clients.write</code>	This scope translates into <code>clients.write</code> after zone switch completes. This scope is used with the <code>X-Identity-Zone-Id header</code> .
<code>zones.ZONE-ID.clients.scim.read</code>	This scope translates into <code>scim.read</code> after zone switch completes. This scope is used with the <code>X-Identity-Zone-Id header</code> .
<code>zones.ZONE-ID.clients.scim.create</code>	This scope translates into <code>scim.create</code> after zone switch completes. This scope is used with the <code>X-Identity-Zone-Id header</code> .
<code>zones.ZONE-ID.clients.scim.write</code>	This scope translates into <code>scim.write</code> after zone switch completes. This scope is used with the <code>X-Identity-Zone-Id header</code> .
<code>zones.ZONE-ID.idps.read</code>	This scope translates into <code>idps.read</code> after zone switch completes. This scope is used with the <code>X-Identity-Zone-Id header</code> .

Cloud Controller Scopes

Scope	Description
<code>cloud_controller.read</code>	This scope gives the ability to read from any Cloud Controller route the token has access to.
<code>cloud_controller.write</code>	This scope gives the ability to post to Cloud Controller routes the token has access to.
<code>cloud_controller.admin</code>	This admin scope gives full permissions to Cloud Controller.
<code>cloud_controller.admin_read_only</code>	This admin scope gives read permissions to Cloud Controller.
<code>cloud_controller.global_audit</code>	This scope gives read-only access to all Cloud Controller API resources except for secrets such as environment variables.

Routing Scopes

Scope	Description
<code>routing.routes.read</code>	This scope gives the ability to read the full routing table from the router.
<code>routing.routes.write</code>	This scope gives the ability to write the full routing table from the router.
<code>routing.router_groups.read</code>	This scope gives the ability to read the full list of routing groups.
<code>routing.router_groups.write</code>	This scopes gives the ability to write the full list of routing groups.

Other Scopes

Scope	Description
<code>doppler.firehose</code>	This scope gives the ability to read logs from the Loggregator Firehose endpoint.
<code>notifications.write</code>	This scope gives the ability to send notifications through the Notification Service .

Garden

This topic describes Garden, the component that Cloud Foundry uses to create and manage isolated environments called containers. Each instance of an application deployed to Cloud Foundry runs within a container. For more information about how containers work, see the [Container Mechanics](#) section of the *Container Security* topic.

Backends

Garden has pluggable backends for different platforms and runtimes, and specifies a set of interfaces that each platform-specific backend must implement. These interfaces contain methods to perform the following actions:

- Create and delete containers
- Apply resource limits to containers
- Open and attach network ports to containers
- Copy files into and out of containers
- Run processes within containers
- Stream `STDOUT` and `STDERR` data out of containers
- Annotate containers with arbitrary metadata
- Snapshot containers for redeployments without downtime

For more information, see the [Garden repository](#) on GitHub.

Garden-runC

Cloud Foundry currently uses the [Garden-runC](#) backend, a Linux-specific implementation of the Garden interface using the [Open Container Interface \(OCI\)](#) standard. Previous versions of Cloud Foundry used the [Garden-Linux](#) backend.

 **Note:** PAS versions v1.8.8 and above use Garden-runC instead of Garden-Linux.

Garden-runC has the following features:

- Uses the same OCI low-level container execution code as Docker and Kubernetes, so container images run identically across all three platforms
- [AppArmor](#) is configured and enforced by default for all unprivileged containers
- Seccomp whitelisting restricts the set of system calls a container can access, reducing the risk of container breakout
- Allows pluggable networking and rootfs management

For more information, see the [Garden-runC repository](#) on GitHub.

Garden RootFS (GrootFS)

Garden manages container filesystems through a plugin interface. Cloud Foundry uses the [GrootFS](#) plugin for this task. GrootFS is a Linux-specific implementation of the Garden volume plugin interface.

GrootFS performs the following actions:

- Creates container filesystems based on buildpacks and droplets
- Creates container filesystems based on remote docker images
- Authenticates with remote registries when using remote images
- Properly maps UID/GID for all files inside an image
- Executes garbage collection to remove unused volumes
- Applies per container disk quotas
- Provides per container disk usage stats

For more information, see [GrootFS Disk Usage](#) and the [GrootFS repository](#) on GitHub.

GrootFS Disk Usage

This topic explains the concepts related to GrootFS disk space management in Pivotal Application Service (PAS).

GrootFS Stores

GrootFS is the container root filesystem management component for Garden. A container root filesystem `orootfs` is often referred to as an **image**.

A **GrootFS store** is the directory in which `orootfs` layers and container images are cached. This directory is configured by GrootFS and mounted on an XFS-formatted volume by the Garden job during BOSH VM creation.

Individual container root filesystems are provided via OverlayFS mounts.

Supplying GrootFS with an already formatted XFS volume for its store is not yet supported for BOSH-controlled deployments.

General Garbage Collection Behavior in GrootFS Stores

GrootFS stores are initialized to use the entirety of `/var/vcap/data`. If the `reserved_space_for_other_jobs_in_mb` is not set high enough, or if there are many images with few shared volumes, the store can use up all available space.

The threshold component calculates and sets a value so that GrootFS's garbage collector can attempt to ensure that a small reserved space is kept free for other jobs. GrootFS only tries to garbage collect when that threshold is reached. However, if all the `orootfs` layers are actively in use by images, then garbage collection cannot occur and that space is used up.

Volumes

Underlying layers in `orootfs` images are known as `volumes` in GrootFS. They are read-only and their changesets are layered together through an **OverlayFS** mount to create the root filesystems for containers.

When GrootFS writes each filesystem volume to disk, it also stores the number of bytes written to a file in a `meta` directory. The size of an individual volume is available in its corresponding metadata file. GrootFS also stores the SHA of each underlying volume used by an image in the `meta` folder.

For each container, GrootFS mounts the underlying `volumes` using overlay to a point in the `images` directory. This mount point is the `orootfs` for the container and is read-write.

On disk, the read-write layer for each container can be found at `/var/vcap/data/grootfs/store/unprivileged/images/CONTAINER-ID/diff` (or `/var/vcap/data/grootfs/store/privileged/images/CONTAINER-ID/diff` for privileged containers.)

When GrootFS calls on the built-in XFS quota tooling to get disk usage for a container, it takes into account data written to those `diff` directories and not the data in the read-only volumes.

Volume Cleanup Example

When `clean` is called in GrootFS, any layers that are not being used by an existing `orootfs` are deleted from the store. The cleanup only takes into account the `volumes` folders in the store.

For example, imagine that there are two `orootfs` images from different base images, Image A and Image B:

```
- Image A
  Layers:
    - layer-1
    - layer-2
    - layer-3
```

```
- Image B
  Layers:
    - layer-1
    - layer-4
    - layer-5
```

They have a layer in common, layer-1. And after deleting Image B, layer-4 and layer-5 can be collected by clean, but not layer-1 because Image A still uses that layer.

For more information on how to calculate GrootFS disk usage in your deployment, see [Examining GrootFS Disk Usage](#).

Additional Information

For more information, see the following sections of [garden-runc-release](#):

- [overlay-xfs-setup ↗](#)
- [grootfs-utils ↗](#)
- [thresholder ↗](#)

HTTP Routing

This topic describes how the Gorouter, the main component in Cloud Foundry's [routing](#) tier, routes HTTP traffic within Cloud Foundry (CF).

HTTP Headers

Header Size Limit

Gorouter has a limit of 1MB for HTTP Headers.

X-Forwarded-Proto

The `X-Forwarded-Proto` header gives the scheme of the HTTP request from the client.

If an incoming request includes the `X-Forwarded-Proto` header, Gorouter does the following:

- Appends it to the existing header
- Sets the scheme to HTTP if the client made an insecure request, meaning a request on port 80
- Sets the scheme to HTTPS if the client made a secure request, meaning a request on port 443

Developers can configure their apps to reject insecure requests by inspecting the `X-Forwarded-Proto` HTTP header on incoming traffic. The header may have multiple values represented as a comma-separated list, so developers must ensure the app rejects traffic that includes any `X-Forwarded-Proto` values that are not HTTPS.

X-Forwarded-For

If `X-Forwarded-For` is present, the Gorouter appends the load balancer's IP address to it and forwards the list. If `X-Forwarded-For` is not present, then the Gorouter sets it to the IP address of the load balancer in the forwarded request (some load balancers masquerade the client IP). If a load balancer sends the client IP using the PROXY protocol, then the Gorouter uses the client IP address to set `X-Forwarded-For`.

If your load balancer terminates TLS on the client side of the Gorouter, it must append these headers to requests forwarded to the Gorouter. For more information, see the [Securing Traffic into Cloud Foundry](#) topic.

HTTP Headers for Zipkin Tracing

Zipkin is a tracing system that enables app developers to troubleshoot failures or latency issues. Zipkin provides the ability to trace requests and responses across distributed systems. See [Zipkin.io](#) for more information.

When the [Zipkin feature is enabled in Cloud Foundry](#), the Gorouter examines the HTTP request headers and performs the following:

- If the `X-B3-TraceId` and `X-B3-SpanId` HTTP headers are not present in the request, the Gorouter generates values for these and inserts the headers into the request forwarded to an application. These values are also found in the Gorouter access log message for the request: `x_b3_traceid` and `x_b3_spanid`.
- If the `X-B3-TraceId` and `X-B3-SpanId` HTTP headers are present in the request, the Gorouter forwards them unmodified. In addition to these trace and span ids, the Gorouter access log message for the request includes `x_b3_parentspanid`.

Developers can then add Zipkin trace IDs to their application logging in order to trace app requests and responses in Cloud Foundry.

After adding Zipkin HTTP headers to app logs, developers can use `cf logs myapp` to correlate the trace and span ids logged by the Gorouter with the trace ids logged by their app. To correlate trace IDs for a request through multiple apps, each app must forward appropriate values for the headers with requests to other applications.

HTTP Headers for App Instance Routing

Developers who want to obtain debug data for a specific instance of an app can use the HTTP header `X-CF-APP-INSTANCE` to make a request to an app instance.

Perform the following steps to make an HTTP request to a specific app instance:

1. Obtain the GUID of your app:

```
$ cf app YOUR-APP --guid
```

2. List your app instances and retrieve the index number of the instance you want to debug:

```
$ cf app YOUR-APP
```

3. Make a request to the app route using the HTTP header `X-CF-APP-INSTANCE` set to the concatenated values of the app GUID and the instance index:

```
$ curl app.example.com -H "X-CF-APP-INSTANCE":"YOUR-APP-GUID:YOUR-INSTANCE-INDEX"
```

Forward Client Certificate to Applications

Applications that require mutual TLS (mTLS) need metadata from client certificates to authorize requests. Cloud Foundry supports this use case without bypassing layer-7 load balancers and the Gorouter.

The HTTP header `X-Forwarded-Client-Cert` (XFCC) may be used to pass the originating client certificate along the data path to the application. Each component in the data path must trust that the back end component has not allowed the header to be tampered with.

If you configure the load balancer to terminate TLS and set the XFCC header from the received client certificate, then you must also configure the load balancer to strip this header if it is present in client requests. This configuration is required to prevent spoofing of the client certificate.

The following sections describe supported deployment configurations.

Terminating TLS at Load Balancer

By default, Gorouter forwards arbitrary headers that are not otherwise mentioned in the docs; this includes the XFCC header.

For applications to receive the XFCC header, configure your load balancer to set the XFCC header with the contents of the client certificate received in the TLS handshake.

This mode is enabled when the **TLS terminated for the first time at infrastructure load balancer** option is selected in the **Networking** configuration screen of the PAS tile.

Terminating TLS at HAProxy

This option allows you to configure support for the XFCC header while leveraging HAProxy. When selected, HAProxy sets the XFCC header to the contents of the client certificate received in the TLS handshake.

Selecting this configuration requires that the load balancer in front of HAProxy is configured to pass through the TLS handshake to HAProxy via TCP.

This mode is enabled when the **TLS terminated for the first time at HAProxy** option is selected in the **Networking** configuration screen of the PAS tile.

HAProxy trusts the Diego intermediate certificate authority. This trust is enabled automatically and permits mutual authentication between applications that are running on Pivotal Cloud Foundry.

Terminating TLS at Gorouter

If the Gorouter is the first component to terminate TLS, such that it receives the certificate of the originating client in the mutual TLS handshake, the operator should select this option. When selected, Gorouter sets the XFCC header to the contents of the client certificate received in the TLS handshake and strips the XFCC header when present in a request.

Selecting this configuration requires that the load balancer in front of Gorouter is configured to pass through TLS handshake to Gorouter via TCP.

This mode is enabled when the **TLS terminated for the first time at the Router** option is selected in the **Networking** configuration screen of the PAS tile.

Gorouter trusts the Diego intermediate certificate authority. This trust is enabled automatically and permits mutual authentication between applications that are running on Pivotal Cloud Foundry.

Client-Side TLS

Depending on your needs, you can configure your deployment to terminate TLS at the Gorouter, at the Gorouter and the load balancer, or at the load balancer only. For more information, see the [Securing Traffic into Cloud Foundry](#) topic.

TLS to Apps and Other Back End Services

The Gorouter supports TLS and mutual authentication to back end destinations, including app instances, platform services, and any other routable endpoints.

This is enabled by following the steps in [Configure Validation of App Instance Identity with TLS](#).

Preventing Misrouting

As CF manages and balances apps, the internal IP address and ports for app instances change. To keep the Gorouter's routing tables current, a Route-Emitter on each Diego cell sends a periodic update to all Gorouters through NATS, reminding them of the location of all app instances; the default frequency of these updates is 20 seconds. The Gorouter tracks a time-to-live (TTL) for each route to back end mapping; this TTL defaults to 120 seconds and is reset when the Gorouter receives an updated registration message.

Network partitions or NATS failures can cause the Gorouter's routing table to fall out of sync, as CF continues to re-create containers across hosts to keep apps running. This can lead to routing of requests to incorrect destinations.

You can configure the Gorouter to handle this problem in two ways:

Consistency mode	Security of traffic between Gorouter and Containers	Gorouter Route Pruning
With TLS Enabled	Encrypted via TLS	Routes are pruned on failure of TLS handshake only. See Gorouter TLS pruning behavior .
Without TLS Enabled	Plain text	Routes are pruned on TTL expiry

Both of these consistency modes are described below.

With TLS Enabled

This consistency mode is newer and has the following benefits:

- Improved availability for applications by keeping routes in the Gorouter's routing table when TTL expires
- Increased guarantees against misrouting by validating the identity of backends before forwarding requests
- Increased security by encrypting data in flight from the Gorouter to backends

A warning: TLS routing requires an additional 32 MB of RAM capacity on your Diego cells per app instance, as well as additional CPU capacity. Check the total amount of Diego cell memory available to allocate in your environment, and if it is less than 32 MB times the number of running app instances, scale out your Diego cells.

A warning: You may see an increase of memory and CPU usage for your Gorouters after enabling TLS routing. Check the total amount of memory and CPU usage of the Gorouters in your environment, and if they are close to the size limit, consider scaling out your Gorouters before enabling TLS routing.

In this mode, the Diego Route-Emitters send a modified route registration message to NATS that includes a unique identifier for the app instance, as well as instructions to use TLS when communicating with the instance. See [TLS to Apps and Other Back End Services](#) for details.

Before forwarding traffic to an app instance, the Gorouter initiates a TLS handshake with an [Envoy proxy](#) running in each app container. In the TLS handshake, the Envoy proxy presents a certificate generated by Diego for each container which uniquely identifies the container using the same app instance identifier sent by the Route-Emitter, configured in the certificate as a domain Subject Alternative Name (SAN).

If the Gorouter confirms that the app instance identifier in the certificate matches the one received in the route registration message, the Gorouter forwards the HTTP request over the TLS session, and the Envoy proxy then forwards it to the app process. If the instance identifiers do not match, the Gorouter removes the app instance from its routing table and transparently retries another instance of the app.

Configure Validation of App Instance Identity with TLS

Verifying app identity using TLS improves resiliency and consistency for app routes. The **Application Containers** configuration pane of the PAS tile includes the following options:

 **Breaking Change:** If you have mutual TLS app identity verification enabled, app containers accept incoming communication only from the Gorouter. This disables TCP routing.

 **Note:** This feature does not work if the **Disable SSL certificate verification for this environment** checkbox is selected.

- **Router uses TLS to verify application identity**
 - Enables the Gorouter to verify app identity using TLS. This is the default option.
- **Router and applications use mutual TLS to verify each other's identity**
 - Enables the Gorouter and your apps to verify each other's identity using TLS.

Without TLS Enabled

In this consistency mode, the Diego Route-Emitters on each cell send route registration messages that include instructions for the Gorouter to send unencrypted requests to the app instance. If the Gorouter does not receive an update for the route within the time-to-live (TTL) interval, the route is pruned from the Gorouter's routing table. See [TLS to Apps and Other Back End Services](#) for details.

This pruning method was developed before the alternative was available.

Consistency Mode Can Differ by Instance Group

The Gorouter can [validate app instance identity using TLS](#) only when Diego cells are configured appropriately. Because cells are configured for TLS through the instance group that they belong to, the Gorouter can run in different consistency modes with cells in different instance groups. For example, the Gorouter can communicate over TLS and validate the cells in one Isolation Segment, while communicating with cells in another Isolation Segment via plain text and [without validating instance identity](#).

Currently, only Linux cells support the Gorouter validating app instance identities [using TLS](#). With Windows cells, the Gorouter runs [without TLS enabled](#), forwarding requests to Windows apps over plain text and pruning based on route TTL.

Round-Robin Load Balancing

The Gorouter uses the round-robin algorithm for load balancing incoming requests to application instances. The Gorouter maintains a dynamically updated list of application instances for each route, and forwards each request for a given route to the next application instance in the list.

WebSockets

WebSockets is a protocol providing bi-directional communication over a single, long-lived TCP connection, commonly implemented by web clients and servers. WebSockets are initiated through HTTP as an upgrade request. The Gorouter supports this upgrade handshake, and holds the TCP connection open with the selected application instance. To support WebSockets, the operator must configure the load balancer correctly. Depending on the configuration, clients may have to use a different port for WebSocket connections, such as port 4443, or a different domain name. For more information, see the [Supporting WebSockets](#) topic.

Session Affinity

The Gorouter supports session affinity, or *sticky sessions*, for incoming HTTP requests to compatible apps.

With sticky sessions, when multiple instances of an app are running on CF, requests from a particular client always reach the same app instance. This allows apps to store session data specific to a user session.

- To support sticky sessions, configure your app to return a `JSESSIONID` cookie in responses. The app generates a `JSESSIONID` as a long hash in the following format:

```
1A530637289A03B07199A44E8D531427
```

- If an app returns a `JSESSIONID` cookie to a client request, the CF routing tier generates a unique `VCAP_ID` for the app instance based on its GUID in the following format:

```
323f211e-fca3-4161-9bd1-615392327913
```

- On subsequent requests, the client must provide both the `JSESSIONID` and `VCAP_ID` cookies.

The CF routing tier uses the `VCAP_ID` cookie to forward client requests to the same app instance every time. The `JSESSIONID` cookie is forwarded to the app instance to enable session continuity. If the app instance identified by the `VCAP_ID` crashes, the Gorouter attempts to route the request to a different instance of the app. If the Gorouter finds a healthy instance of the app, it initiates a new sticky session.

 **Note:** CF does not persist or replicate HTTP session data across app instances. If an app instance crashes or is stopped, session data for that instance is lost. If you require session data to persist across crashed or stopped instances, or to be shared by all instances of an app, store session data in a CF marketplace service that offers data persistence.

Keepalive Connections

From Front End Clients

The Gorouter supports keepalive connections from clients and does not close the TCP connection with clients immediately after returning an HTTP response. Clients are responsible for closing these connections.

To Back End Servers

If keepalive connections are disabled, the Gorouter closes the TCP connection with an app instance or system component after receiving an HTTP response.

If keepalive connections are enabled, the Gorouter maintains established TCP connections to back ends. the Gorouter supports up to 100 idle connections to each back end:

- If an idle connection exists for a given back end, the Gorouter reuses it to route subsequent requests.
- If no idle connection exists for this back end, the Gorouter creates a new connection.

Transparent Retries

If the Gorouter cannot establish a TCP connection with a selected application instance, the Gorouter considers the instance ineligible for requests for 30 seconds and transparently attempts to connect to another application instance. Once the Gorouter has established a TCP connection with an application instance, the Gorouter forwards the HTTP request.

When you deploy an app that requires Diego cells to restart or recreate, the app may not respond to a Gorouter request before the keepalive connection breaks. The following table describes how the Gorouter behaves if it cannot establish a TCP connection to an app:

If the Gorouter...	and the backend...	then the Gorouter...
cannot establish a TCP connection to a backend	n/a	retries another backend, no more than 3 times
establishes a TCP connection to a backend and	does not respond	waits 15 minutes for a response, and if it errors, does not retry another

forwards the request	backend
establishes a TCP connection to a backend and forwards the request	returns a TCP connection error

In all cases, if the app still does not respond to the request, the Gorouter returns a [502 error ↗](#).

Container-to-Container Networking

This topic provides an overview of how Container-to-Container Networking works.

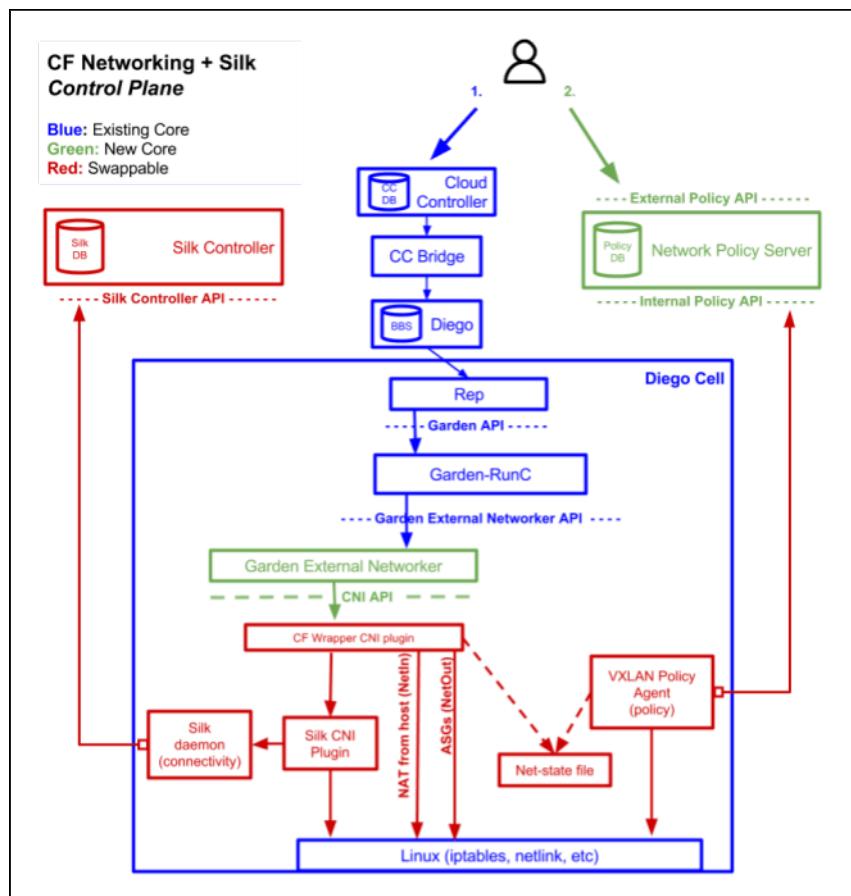
The Container-to-Container Networking feature enables app instances to communicate with each other directly. Container-to-Container Networking is always enabled in PAS. For more information about how to configure Container-to-Container Networking, see the [Administering Container-to-Container Networking](#) topic.

Architecture

Overview

Container-to-Container Networking integrates with [Garden-runC](#) in a [Diego](#) deployment. The [Container-to-Container Networking BOSH release](#) includes several core components, as well as swappable components.

To understand the components and how they work, see the diagram and tables below. The diagram highlights Pivotal Application Service (PAS) components in blue and green. The diagram also highlights swappable components in red.



Core Components

The Container-to-Container Networking BOSH release includes the following core components:

Part	Function
Cloud Foundry Command Line Interface (CF CLI) plugin	A plugin that you download to control network access policies between apps.
	A central management node that does the following: <ul style="list-style-type: none"> Maintains a database of policies for traffic between apps. The CF CLI plugin calls an API to create or update a

Policy Server	<p>record in the policy database whenever you create or remove a policy.</p> <ul style="list-style-type: none"> Exposes a JSON REST API used by the CF CLI plugin
Garden External Networker	<p>A Garden-runC add-on deployed to every Diego cell that does the following:</p> <ul style="list-style-type: none"> Invokes the CNI plugin component to set up the network for each app Forwards ports to support incoming connections from the Gorouter, TCP Router, and Diego SSH Proxy. This keeps apps externally reachable. Installs outbound whitelist rules to support Application Security Groups (ASGs).

Swappable Components

The Container-to-Container Networking BOSH release includes the following swappable components:

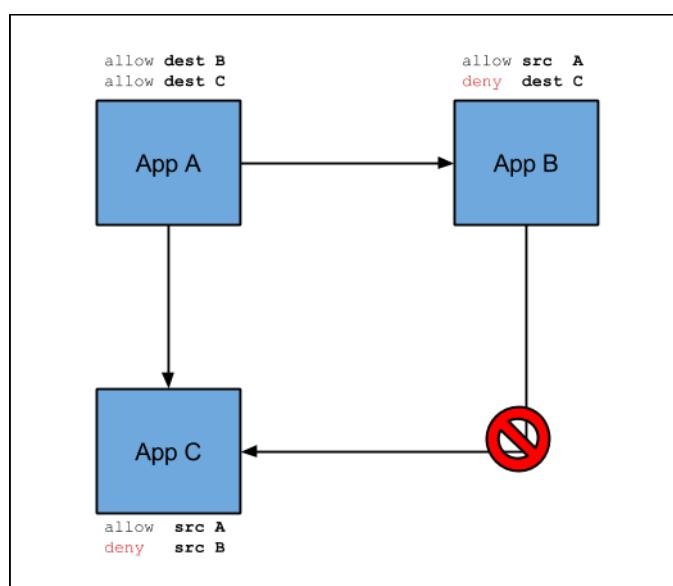
Part	Function
Silk CNI plugin	<p>A plugin that provides IP address management and network connectivity to app instances as follows:</p> <ul style="list-style-type: none"> Uses a shared VXLAN overlay network to assign each container a unique IP address Installs network interface in container using the Silk VXLAN backend. This is a shared, flat L3 network.
VXLAN Policy Agent	<p>Enforces network policy for traffic between apps as follows:</p> <ul style="list-style-type: none"> Discovers desired network policies from the Policy Server Internal API Updates iptables rules on the Diego cell to allow whitelisted inbound traffic Tags outbound traffic with the unique identifier of the source app using the VXLAN Group-Based Policy (GBP) header

App Instance Communication

The diagram below illustrates how app instances communicate in a deployment with Container-to-Container Networking enabled. In this example, the operator creates two policies to regulate the flow of traffic between **App A**, **App B**, and **App C**.

- Allow traffic from **App A** to **App B**
- Allow traffic from **App A** to **App C**

If traffic and its direction is not explicitly allowed, it is denied. For example, **App B** cannot send traffic to **App C**.



Overlay Network

Container-to-Container Networking uses an overlay network to manage communication between app instances.

Overlay networks are not externally routable, and traffic sent between containers does not exit the overlay. You can use the same overlay network range for different Cloud Foundry deployments in your environment.

The overlay network range defaults to `10.255.0.0/16`. You can modify the default to any [RFC 1918](#) range that meets the following requirements:

- The range is not used by services that app containers access.
- The range is not used by the underlying Cloud Foundry infrastructure.

All Diego cells in your Cloud Foundry deployment share this overlay network. By default, each cell is allocated a /24 range that supports 254 containers per cell, one container for each of the usable IP addresses, `.1` through `.254`. To modify the number of Diego cells your overlay network supports, see [Overlay Network](#) in [Configuring Container-to-Container Networking](#).

Cloud Foundry container networking is currently supported only on Linux.

⚠ warning: The overlay network IP address range must not conflict with any other IP addresses in the network. If a conflict exists, Diego cells cannot reach any endpoint that has a conflicting IP address.

💡 Note: Traffic to app containers from the Gorouter or from app containers to external services uses cell IP addresses and NAT, not the overlay network.

Policies

Enabling Container-to-Container Networking for your deployment allows you to create policies for communication between app instances. The Container-to-Container Networking feature also provides a unique IP address to each app container and provides direct IP reachability between app instances.

The policies you create specify a source app, destination app, protocol, and port so that app instances can communicate directly without going through the Gorouter, a load balancer, or a firewall. Container-to-Container Networking supports UDP and TCP, and you can configure policies for multiple ports. These policies apply immediately without having to restart the app.

Additionally, policies use and track the GUIDs of the apps. The policies continue to work when apps redeploy, or if they crash and Diego places them in a new container. Pushing a brand new app requires a new policy, but not updates to an existing app because an app always retains its GUID.

App Service Discovery

The Pivotal Application Service (PAS) platform supports DNS-based service discovery that lets apps find each others' internal addresses. For example, a front end app instance can use the service discovery mechanism to establish communications with a back end app instance. See the [Developer Guide](#) for how to set up and use app service discovery.

Container-to-Container app service discovery does not provide client-side load balancing or circuit-breaking, and it does not apply to [cf marketplace](#) services or require [application binding](#). It just lets apps publish service endpoints to each other, unbrokered and unmediated.

Alternative Network Stacks

The BOSH release that contains the Container-to-Container Networking feature is composed of a pluggable network stack. Advanced users or third-party vendors can integrate a different network stack. For more information about third-party plugins, see the [Container-to-Container Networking BOSH release](#) documentation.

Container-to-Container Networking versus ASGs

Both application security groups (ASGs) and Container-to-Container Networking policies affect traffic from app instances. The following table highlights differences between ASGs and Container-to-Container Networking policies.

	ASGs	Container-to-Container Networking Policies
Policy granularity	From a space to an IP address range	From a source app to a destination app
Scope	For a space, org, or deployment	For app to app only
Traffic direction	Outbound control	Policies apply for incoming packets from other app instances

Source app	Is not known	Is identified because of direct addressability
Policies take affect	After app restart	Immediately

PCF Operator Guide

For PCF Operators

This guide shows you how to run a PCF platform. This ongoing responsibility may include but is not limited to:

- Configuring PCF capabilities
- Integrating PCF with external systems
- Updating PCF and installed products
- Monitoring PCF health and performance
- Adjusting PCF resources and options to fix health or performance issues
- Diagnosing and troubleshooting PCF problems
- Managing Pivotal Application Service (PAS) users, resources, and infrastructure
- Installing software services and otherwise enabling PCF developers
- Maintaining PCF
- Keeping PCF secure

If you do these things, you are a PCF **operator**, and the contents of this guide are for you.

Guide Contents

- [Day 2 Configurations](#) - Setting up internal operations and external integrations for PCF.
- [Ongoing Operations](#) - Routine procedures for running and growing the platform, including:
 - PCF Upgrades
 - IaaS Changes
 - Monitoring, Logging, and Reporting
 - Platform Tuning
 - Enabling Developers
 - Backing Up
 - Security
- [Managing PAS Runtimes](#) - Procedures performed by people with administrator or manager roles in Pivotal Application Service (PAS), such as managing users, orgs, spaces, and service instances. Operators can perform these actions by [logging in with Admin credentials](#), which grants them the role of Org Manager across all PAS orgs.
- [Using Ops Manager](#) - Ops Manager's dashboard interface streamlines the installation, configuration, and upgrading of PCF platform services and add-ons.
- [Using the Cloud Foundry Command Line Interface \(cf CLI\)](#) - Using the cf CLI to send commands to the Cloud Controller, the executive component of PAS.
- [Troubleshooting and Diagnostics](#) - Tools and procedures for troubleshooting PCF.

Deploying Service Mesh (Beta)

This topic describes how to deploy service mesh for Pivotal Application Service (PAS).

For more information about service mesh, see [Service Mesh \(Beta\)](#).

Enable in PAS

To deploy service mesh, configure PAS as follows:

1. Navigate to the **Networking - Service Mesh** pane of the PAS tile.

Service Mesh currently only supports ingress routing to applications; support for management of egress communications (including app-to-app) will be offered in future releases.

Service Mesh (Beta)*

Disable When enabled, a new kind of platform ingress router will be deployed, in parallel to the existing HTTP and TCP routers, offering new capabilities. Enable this only for deployments with fewer than 20,000 routes; core platform functions will be impacted at greater scale.
 Enable

IP Addresses for Ingress Router

Ingress Router TLS Keypairs

Add

Name *

Certificate and Private Key for Istio Router *

Certificate PEM

Private Key PEM

Generate RSA Certificate

Save

2. Under **Service Mesh (Beta)**, select **Enable**.
3. For **IP Addresses for Ingress Router**, do the following depending on your IaaS:
 - **vSphere**: Enter static IPs for the Istio Routers. You must configure your load balancer with these IPs as well.
 - **Other**: Leave this field blank.
4. For **Ingress Router TLS Keypairs**, complete the following fields. You can add more than one keypair if desired using the **Add** button.

- **Name:** Enter a name for the keypair.
- **Certificate and Private Key for Istio Router:** Enter the Private key and certificate for TLS handshakes with clients. These must be in PEM block format.

5. Click **Save**.

Create a Load Balancer

To configure a load balancer for service mesh, do the following. The exact procedure varies by IaaS.

1. Create a load balancer with the following:
 - A static IP
 - Health check port `8002` and path `/healthcheck`
 - Firewall rules to allow the following:
 - HTTP on port `80`
 - HTTP on port `8002`
 - TLS on port `443`
2. Navigate to your DNS provider and create a DNS name, `*.mesh.YOUR-CF-APPS-DOMAIN`, that resolves to the IP of the load balancer. This is the default domain for service mesh and it is not configurable.

Add Load Balancer to Resource Config

If your deployment is on an IaaS other than vSphere, do the following after you create your load balancer:

1. Navigate to the **Resource Config** pane of the PAS tile.
2. In the **Load Balancer** column of the **istio-router** row, enter the name of the load balancer you created.
3. Click **Apply Changes**.

Identifying the API Endpoint for Your PAS Instance

The API endpoint for your Pivotal Application Service (PAS) deployment, its target URL, is the API endpoint of the deployment's Cloud Controller. Find your Cloud Controller API endpoint by consulting your cloud operator, from the Apps Manager, or from the command line.

From the Apps Manager

Log in to the Apps Manager for your PAS instance, then click **Tools** in the left navigation panel. The **Getting Started** section of the Tools page shows your API endpoint.

```
GETTING STARTED

$ cf help
$ cf login -a https://api.your_endpoint.com
API endpoint: https://api.your_endpoint.com
Username> your_username
Password> your_password
Org> your_org
Space> your_space
$ cf push your_app
```

From the Command Line

From a command line, use the `cf api` command to view your API endpoint.

Example:

```
$ cf api
API endpoint: https://api.example.com (API version: 2.2.0)
```

Creating New PAS User Accounts

When you first deploy your [Pivotal Application Service \(PAS\)](#), there is only one user: an administrator. At this point, you can add accounts for new users who can then push applications using the Cloud Foundry Command Line Interface (cf CLI).

How to add users depends on whether or not you have SMTP enabled, as described in the options below.

Option 1: Adding New Users when SMTP is Enabled

If you have enabled SMTP, your users can sign up for accounts and create their own orgs. They do this using the [Pivotal Cloud Foundry](#) (PCF) Apps Manager, a self-service tool for managing organizations, users, applications, and application spaces.

Instruct users to complete the following steps to log in and get started using the Apps Manager.

1. Browse to `apps.YOUR-SYSTEM-DOMAIN`. Refer to [PAS Domains](#) to locate your system domain.
2. Select [Create an Account](#).
3. Enter your email address and click [Create an Account](#). You will receive an email from the Apps Manager when your account is ready.
4. When you receive the new account email, follow the link in the email to complete your registration.
5. You will be asked to choose your organization name.

You now have access to the Apps Manager. Refer to the Apps Manager documentation at [docs.pivotal.io](#) for more information about using the Apps Manager.

Option 2: Adding New Users when SMTP is Not Enabled

If you have not enabled SMTP, only an administrator can create new users, and there is no self-service facility for users to sign up for accounts or create orgs.

The administrator creates users with the cf CLI. See [Creating and Managing Users with the cf CLI](#).

[Return to the Installing Pivotal Cloud Foundry Guide](#)

Configuring SSL/TLS Termination at HAProxy

Both Pivotal Application Service (PAS) and Isolation Segments for Pivotal Cloud Foundry include an HAProxy instance.

HAProxy is appropriate to use in a deployment when features are needed that are offered by HAProxy but are not offered by the CF Routers or IaaS-provided load balancers such as with Azure load balancers. These include filtering of protected domains from trusted networks.

While HAProxy instances provide load balancing for the Gorouters, HAProxy is not itself highly available. For production environments, use a highly-available load balancer to scale HAProxy horizontally. The load balancer does not need to terminate TLS or even operate at layer 7 (HTTP); it can simply provide layer 4 load balancing of TCP connections. Use of HAProxy does not remove the need for CF Routers; the Gorouter must always be deployed for HTTP applications, and TCP Router for non-HTTP applications.

You can generate a self-signed certificate for HAProxy if you do not want to obtain a signed certificate from a well-known certificate authority.

Procedure: Terminate SSL/TLS at HAProxy

In PCF, perform the following steps to configure SSL termination on HAProxy:

1. Navigate to the Ops Manager Installation Dashboard.
2. Click the **Pivotal Application Service** tile in the Installation Dashboard.
3. Click **Networking**.
4. Configure the following based on the IaaS of your PCF deployment.

If your PCF deployment is on:	Then configure the following:	See also:
OpenStack or vSphere	<p>Decide whether you want your HAProxy to be highly available.</p> <ul style="list-style-type: none"> o If you need highly available HAProxy, then perform the following steps: <ol style="list-style-type: none"> 1. Choose an IP address for each HAProxy instance on the subnet where you deployed PCF. 2. In the HAProxy IPs field of the Networking page, enter the IP addresses you have selected for your HAProxy instances. 3. Configure your load balancer (for example, F5 or NSX) to forward domain names to the HAProxy IP addresses. o If you do not require high availability (for example, you are setting up a development environment), then perform the following steps: <ol style="list-style-type: none"> 1. Skip setting up the load balancer. 2. Choose one IP address for the single HAProxy instance. 3. Configure DNS to point at the IP address. See How to Set Up DNS for HAProxy. 	For more information, see Configuring PAS .
AWS, GCP or Azure	<ol style="list-style-type: none"> 1. Leave the HAProxy IP address blank. 2. In the Resource Config page of PAS tile, locate the HAProxy job. 3. In the Load Balancer column for the HAProxy job, specify the appropriate IaaS load balancer resource. 	For more information, see Configuring PAS .

5. In the **Certificates and Private Keys for HAProxy and Router** field, click the **Add** button to define at least one certificate keypair for HAProxy and Router. For each certificate keypair that you add, assign a name, enter the PEM-encoded certificate chain and PEM-encoded private key. You can either upload your own certificate or generate an RSA certificate in PAS. For options and instructions on creating a certificate for your wildcard domains, see [Creating a Wildcard Certificate for PCF Deployments](#).
6. In the **Minimum version of TLS supported by HAProxy and Router**, select the minimum version of TLS to use in HAProxy communications. HAProxy use TLS v1.2 by default. If you need to accommodate clients that use an older version of TLS, select a lower minimum version. For a list of TLS ciphers supported by the HAProxy, see [TLS Cipher Suites](#).
7. Under **HAProxy forwards requests to Router over TLS**, leave **Enabled** selected and provide the backend certificate authority.
8. If you want to use a specific set of TLS ciphers for HAProxy, configure **TLS Cipher Suites for HAProxy**. Enter an ordered, colon-separated list of TLS cipher suites in the OpenSSL format. For example, if you have selected support for an earlier version of TLS, you can enter cipher suites supported by

this version. For a list of TLS ciphers supported by the HAProxy, see [TLS Cipher Suites](#).

9. If you expect requests larger than the default maximum of 16 Kbytes, enter a new value (in bytes) for **HAProxy Request Max Buffer Size**. You may need to do this, for example, to support apps that embed a large cookie or query string values in headers.
10. If you want to force browsers to use HTTPS when making requests to HAProxy, select **Enable** in the **HAProxy support for HSTS** field and complete the following optional configuration steps:
 - a. (Optional) Enter a **Max Age in Seconds** for the HSTS request. By default, the age is set to one year. HAProxy will force HTTPS requests from browsers for the duration of this setting.
 - b. (Optional) Select the **Include Subdomains** checkbox to force browsers to use HTTPS requests for all component subdomains.
 - c. (Optional) Select the **Enable Preload** checkbox to force instances of Google Chrome, Firefox, and Safari that access your HAProxy to refer to their built-in lists of known hosts that require HTTPS, of which HAProxy is one. This ensures that the first contact a browser has with your HAProxy is an HTTPS request, even if the browser has not yet received an HSTS header from your HAProxy.
11. (Optional) If you are not using SSL encryption or if you are using self-signed certificates, you can select **Disable SSL certificate verification for this environment**. Selecting this checkbox also disables SSL verification for route services.

 Use this checkbox only for development and testing environments. Do not select it for production environments.
12. (Optional) If you do not want HAProxy or the Gorouter to accept any non-encrypted HTTP traffic, select the **Disable HTTP on HAProxy and Router** checkbox.
13. In the **Configure the CF Router support for the X-Forwarded-Client-Cert header** field, select **Always forward the XFCC header in the request, regardless of the whether the client connection is mTLS**.
14. (Optional) If your PCF deployment uses HAProxy and you want it to receive traffic only from specific sources, use the following fields:
 - o **Protected Domains:** Enter a comma-separated list of domains from which PCF can receive traffic.
 - o **Trusted CIDRs:** Optionally, enter a space-separated list of CIDRs to limit which IP addresses from the Protected Domains can send traffic to PCF.
15. Click **Save**.

How to Set Up DNS for HAProxy

You only need to perform this procedure if you are using one instance of HAProxy such as in a development environment. If you would like HAProxy to be highly available, you must have a load balancer in front of it. In this case, you would point DNS at the load balancer.

To use a single instance HAProxy load balancer in a vSphere or OpenStack deployment, create a wildcard A record in your DNS and configure some fields in the PAS product tile.

1. Create an A record in your DNS that points to the HAProxy IP address. The A record associates the **System Domain** and **Apps Domain** that you configure in the **Domains** section of the PAS tile with the HAProxy IP address.

For example, with `cf.example.com` as the main subdomain for your Cloud Foundry (CF) deployment and an HAProxy IP address `203.0.113.1`, you must create an A record in your DNS that serves `example.com` and points `*.cf` to `203.0.113.1`.

Name	Type	Data	Domain
<code>*.cf</code>	A	<code>203.0.113.1</code>	<code>example.com</code>

2. Use the Linux `host` command to test your DNS entry. The `host` command should return your HAProxy IP address.

Example:

```
$ host cf.example.com
cf.example.com has address 203.0.113.1
$ host anything.example.com
anything.cf.example.com has address 203.0.113.1
```

Configuring Frontend Idle Timeout for Gorouter and HAProxy

This topic describes how to configure the **Frontend Idle Timeout for Gorouter and HAProxy** field in the Pivotal Application Service (PAS) **Networking** pane.

You can optionally use the **Frontend Idle Timeout for Gorouter and HAProxy** field to help prevent connections from your load balancer to Gorouter or HAProxy from being closed prematurely. The value you enter sets the duration, in seconds, that Gorouter or HAProxy maintains an idle open connection from a load balancer that supports keep-alive.

In general, set the value higher than your load balancer's backend idle timeout to avoid the race condition where the load balancer sends a request before it discovers that Gorouter or HAProxy has closed the connection.

See the following table for specific guidance and exceptions to this rule:

IaaS	Guidance
AWS	AWS ELB has a default timeout of 60 seconds, so Pivotal recommends a value greater than <code>60</code> .
Azure	By default, Azure load balancer times out at 240 seconds without sending a TCP RST to clients, so as an exception, Pivotal recommends a value lower than <code>240</code> to force the load balancer to send the TCP RST.
GCP	GCP has a default timeout of 600 seconds. For GCP HTTP load balancers, Pivotal recommends a value greater than <code>600</code> . For GCP TCP load balancers, pivotal recommends a value less than <code>600</code> to force the load balancer to send a TCP RST.
Other	Set the timeout value to be greater than that of the load balancer's backend idle timeout.

Configuring Route Service Lookup

This topic describes configuring route service lookup in Pivotal Application Service (PAS).

Overview

Developers can bind their app to a route service to preprocess requests before they reach an app. Example use cases include authentication, rate limiting, and caching services. For more information, see [Route Services](#).

The **Bypass security checks for route service lookup** field in the PAS tile allows you to configure how the router handles traffic for apps that are bound to route services. The configuration options are as follows:

- **Default Lookup:**
 - Default lookup is configured when you do not select **Bypass security checks for route service lookup**. In this case, the router does not check for an existing route. It sends traffic back through the load balancer when the traffic is for an internal route service.
- **Bypass Mode:**
 - Bypass mode is configured when you select **Bypass security checks for route service lookup**. The router checks for an existing route. If the router finds the route and the route service is internal, it sends the traffic directly to the route service and skips the load balancer. This improves performance, but introduces the security risk described in [Bypass Mode and External Route Service \(Security Risk\)](#).

For more details, see [Summary of Behavior in Different Configurations](#) below.

For configuration guidance and procedures, see [Configure Route Service Lookup](#) below.

Configure Route Service Lookup

This following sections provide guidance and configuration steps for route service lookup.

Guidance

Pivotal recommends that you do not configure PAS for bypass mode because of the security risk described in [Bypass Mode and External Route Service \(Security Risk\)](#). However, you may need to do so if your load balancer requires mutual TLS from clients.

If your load balancer requires mutual TLS from clients and PAS is configured for default lookup, the router cannot handle traffic successfully for internal route services. This is because the router does not have the necessary certificates from the client to communicate back with the load balancer for DNS lookup. Therefore you must configure bypass mode so the router can send the traffic directly to the route service.

Configuring PAS for Bypass Mode

To configure bypass mode, do the following:

1. Select **Bypass security checks for route service lookup** in the **Networking** pane of the PAS tile.
2. Follow the steps in [Mitigate Security Risk](#).

Mitigate Security Risk

To prevent users from intercepting traffic for externally hosted route services, do the following.

1. Create an org for use by the PAS administrator.
2. Register all external route service domains as private domains in the org you created.
3. Monitor PAS for the addition of new external route services and ensure you follow the same process for those. One way to do this is by using `curl` to regularly view a list of user-provided service instances. You can do this by running `curl /v2/user_provided_service_instances`.

Note: Since route services can be added by any space developer, this may be difficult to manage.

Configuring PAS for Default Lookup

To configure PAS for default lookup behavior, with bypass mode disabled, do the following:

1. Ensure that **Bypass security checks for route service lookup** in the **Networking** pane of the PAS tile is not selected.
2. Communicate to developers of route services that the domain name for their internally hosted route services must resolve to the load balancer.
3. If your load balancer or router terminates TLS, do the following:
 - a. Work with developers of route services to ensure the load balancer or router have TLS certificates that are valid for the route service URL.
 - b. Ensure that the TLS certificate from your load balancer is either signed by a well known CA, or the CA has been added to the **Certificate Authorities Trusted by Router and HAProxy** field in the **Networking** pane of the PAS and Isolation Segment tiles. The CA for the TLS certificate provided by the load balancer must be trusted by the Router.
4. Work with route service developers to verify that their internal route service apps are reachable. You can do this by visiting the HTTPS URL of the route service directly and confirming that the app received the request with the `| cf logs` output for the route service app.

Summary of Behavior in Different Configurations

The following sections describe how the router behaves when bypass mode is enabled or disabled and when a route service is internal or external.

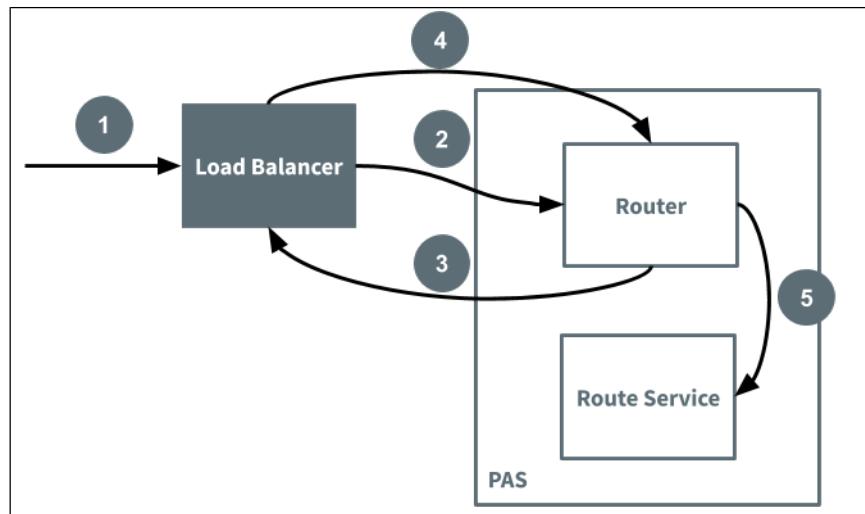
Default Lookup and Internal Route Service

This section describes how the router handles app requests when the following is true:

- The **Bypass security checks for route service lookup** field in PAS is not selected.
- The app is bound to a route service that is hosted on PAS.

In this case, when the router receives the request, it sends the traffic back to the load balancer to resolve DNS. The load balancer then sends the traffic back to the router.

The following diagram illustrates the flow of the request and numbers the steps to indicate order of occurrence.



Bypass Mode and Internal Route Service

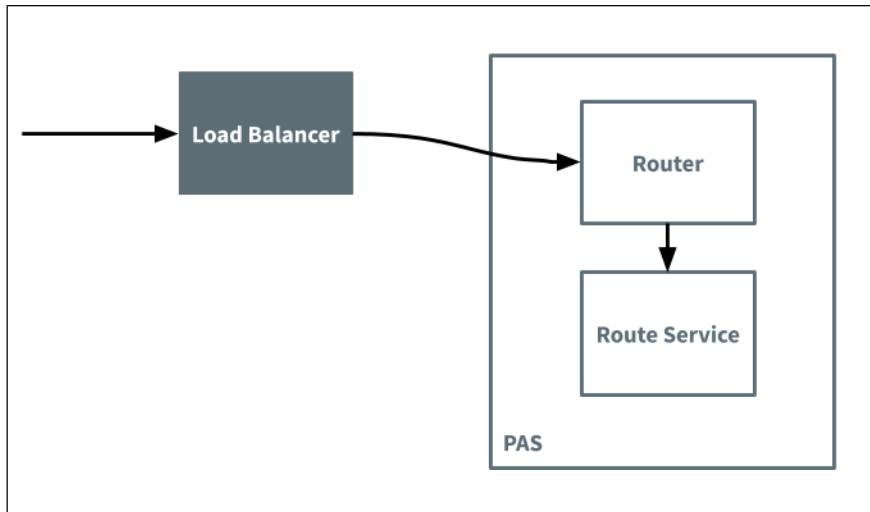
This section describes how the router handles app requests when the following is true:

- The **Bypass security checks for route service lookup** field in PAS is selected.

- The app is bound to a route service that is hosted on PAS

In this case, when the router receives the request, it sends it directly to the route service. This assumes the router finds an existing route for the route service.

The following diagram illustrates the flow of the request.



Bypass Mode and External Route Service (Security Risk)

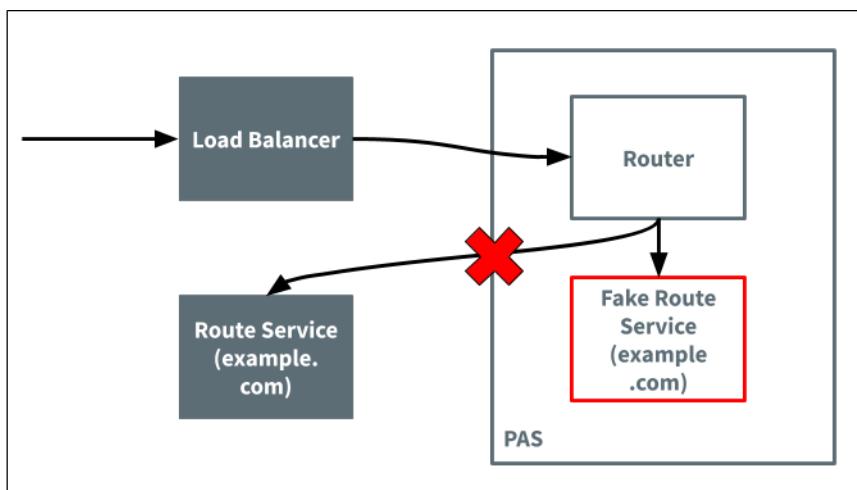
This section describes how the router handles app requests when the following is true:

- The **Bypass security checks for route service lookup** field in PAS is selected.
- The app is bound to a route service that is hosted outside of PAS.

In this case, when the router receives the request, it checks for an existing route and then sends the request directly to the route service. This introduces the ability for route service traffic to be intercepted. A developer can register the external route service domain as a private domain in PAS and map it to their own, malicious app. When the router receives a request for the original app bound to the external route service, it will find the domain internally and send the request to the malicious app.

Note: This vulnerability exists for both externally hosted route services and route services hosted on a separate foundation. If all of your route services are hosted internally on the same foundation, you are not at risk for this vulnerability. However, you would be at risk if externally hosted route services are later configured.

The following diagram illustrates the flow of the request in the case that it is intercepted:



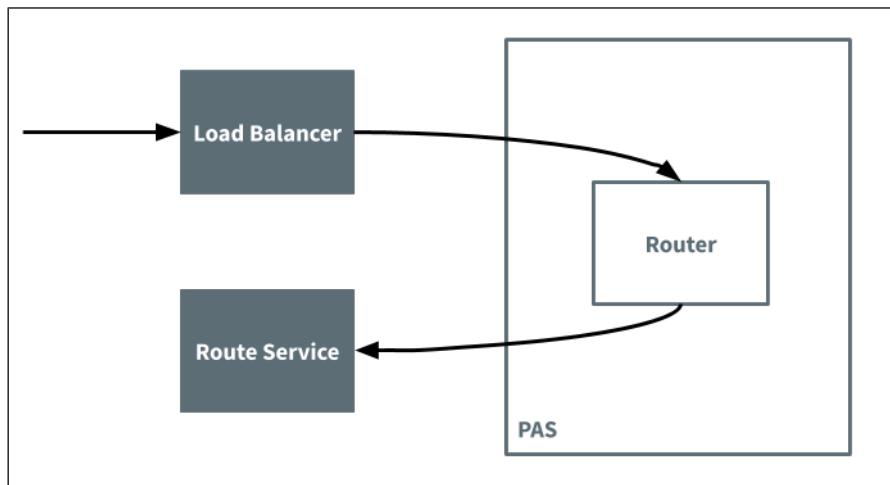
Default Lookup and External Route Service

This section describes how the router handles app requests when the following is true:

- The **Bypass security checks for route service lookup** field in PAS is not selected.
- The app is bound to a route service that is hosted outside of PAS.

In this case, the router sends traffic directly to the external route service without checking for an existing route.

The following diagram illustrates the flow of the request.



Configuring Proxy Settings for All Applications

This topic describes how to globally configure proxy settings for all applications in your Pivotal Cloud Foundry (PCF) deployment. Some environments restrict access to the Internet by requiring traffic to pass through an HTTP or HTTPS proxy. PCF operators can use the Cloud Foundry Command Line Interface (cf CLI) to provide the proxy settings to all applications, including system applications and service brokers.

Note: Incorrectly configuring proxy settings can prevent applications from connecting to the Internet or accessing required resources. They can also cause errands to fail and break system applications and service brokers. Although errands, system applications, and service brokers do not need to connect to the Internet, they often need to access other resources on PCF. Incorrect proxy settings can break these connections.

Set Environment Variables

To globally configure proxy settings for PCF applications, perform the following steps to set three environment variables for both the staging environment variable group and the running environment variable group.

For more information about variable groups, see the [Environment Variable Groups](#) section in the *Cloud Foundry Environment Variables* topic.

This procedure explains how to set proxy information for both staging and running applications. However, you can set proxy settings for only staging or only running applications.

1. Target your Cloud Controller with the cf CLI. If you have not installed the cf CLI, see the [Installing the cf CLI](#) topic.

```
$ cf api api.YOUR-SYSTEM-DOMAIN
Setting api endpoint to api.YOUR-SYSTEM-DOMAIN...
OK
API endpoint: https://api.YOUR-SYSTEM-DOMAIN (API version: 2.54.0)
Not logged in. Use 'cf login' to log in.
```

2. Log in with your UAA administrator credentials. To retrieve these credentials, navigate to the Pivotal Application Service (PAS) tile in the Ops Manager Installation Dashboard and click **Credentials**. Under **UAA**, click **Link to Credential** next to **Admin Credentials** and record the password.

```
$ cf login
API endpoint: https://api.YOUR-SYSTEM-DOMAIN

Email> admin
Password>
Authenticating...
OK
```

3. To configure proxy access for applications that are staging, run the following command, replacing the placeholder values:

```
$ cf set-staging-environment-variable-group '{"http_proxy": "http://YOUR-PROXY:8080/", "https_proxy": "http://YOUR-PROXY:8080/", "no_proxy": "NO-PROXY.EXAMPLE.COM"}'
```

- o `http_proxy` : Set this value to the proxy to use for HTTP requests.
- o `https_proxy` : Set this value to the proxy to use for HTTPS requests. In most cases, this will be the same as `http_proxy`.
- o `no_proxy` : Set this value to a comma-separated list of DNS names or IP addresses that can be accessed without passing through the proxy. This value may not be needed, because it depends on your proxy configuration. From now on, the proxy settings are applied to staging applications.

4. To configure proxy access for applications that are running, run the following command, replacing the placeholder values as above:

```
$ cf set-running-environment-variable-group '{"http_proxy": "http://YOUR-PROXY:8080/", "https_proxy": "http://YOUR-PROXY:8080/", "no_proxy": "NO-PROXY.EXAMPLE.COM"}'
```

To configure proxy settings for Java-based applications, use the following command instead, replacing the placeholder values. For `http.nonProxyHosts`, use a pipe-delimited list rather than a comma-separated list.

```
$ cf set-running-environment-variable-group '{"JAVA_OPTS": "-Dhttp.proxyHost=YOUR-PROXY -Dhttp.proxyPort=8080 -Dhttp.nonProxyHosts=NO-PROXY.EXAMPLE.COM"}'
```

For more information about these Java proxy settings, see [Java Networking and Proxies](#).

5. To apply the proxy configuration for the running environment variable group, you must restart each application that you want to use the new configuration.

Troubleshooting

If an application fails after you apply the global proxy settings, try the following solutions.

Exclude an App From Global Proxy Settings

If your application fails, try instructing the application to ignore the global proxy settings. Perform the following commands to manually unset the proxy environment variables for the failing application:

1. Set the proxy environment variables for `http_proxy` to an empty value:

```
$ cf set-env YOUR-APP http_proxy "
```

2. Set the proxy environment variables for `https_proxy` to an empty value:

```
$ cf set-env YOUR-APP https_proxy "
```

3. Set the proxy environment variables for `no_proxy` to an empty value:

```
$ cf set-env YOUR-APP no_proxy "
```

Change Case of HTTP

Your application and language runtime may be case-sensitive. Try performing the steps in the [Set Environment Variables](#) section using uppercase for `HTTP_PROXY`, `HTTPS_PROXY`, and `NO_PROXY` instead of lowercase. Refer to the following example.

```
$ cf set-staging-environment-variable-group '{"HTTP_PROXY": "http://YOUR-PROXY:8080", "HTTPS_PROXY": "http://YOUR-PROXY:8080"}';
```

Check Proxy Settings

If you have set up your proxy so that it can only send traffic to the Internet, then a request to an internal resource like PCF fails. You must set `no_proxy` so that traffic destined for PCF and other internal resources is sent directly and does not go through the proxy. For instance, setting `no_proxy` to include your system and application domains will ensure that requests destined for those domains are sent directly.

Verify Interpretation

The interpretation of `no_proxy` depends on the application and the language runtime. Most support `no_proxy`, but the specific implementation may vary. For example, some match DNS names that end with the value set in `no_proxy`: `example.com` would match `test.example.com`. Others support the use of the asterisk as a wildcard to provide basic pattern matching in DNS names: `*.example.com` would match `test.example.com`. Most applications and language runtimes do not support pattern matching and wildcards for IP addresses.

Restricting App Access to Internal PCF Components

This topic describes how to secure the component virtual machines (VMs) of your Pivotal Cloud Foundry (PCF) deployment from being accessed by apps.

Introduction

See the following list to understand the concepts for this topic:

- **How PCF determines where apps can send traffic:**
 - PCF uses *Application Security Groups (ASGs)*, which are network policy rules specifying protocols, ports, and IP ranges that apply to outbound network connections initiated from apps. See [ASGs](#).
- **Why you must create new rules for outbound app traffic:**
 - PCF installs with a [default ASG](#) that allows apps running on your deployment to send traffic to almost any IP address. This means apps are not blocked from initiating connections to most network destinations unless an administrator takes action to update the ASGs with a more restrictive policy.
- **How you can set up new rules:**
 - To help secure your component VMs against apps while ensuring your apps can access the services they need, follow the [procedure](#) below, which includes these steps:

Step	Description
1	Determine Your Network Layout: The procedure for securing your deployment with ASGs varies depending on your network layout, which you can determine using Ops Manager.
2	Ensure Access for PCF System Apps: Bind the default ASG to the <code>system</code> org so that PCF system apps can continue accessing the system components they need after you remove the deployment-wide default ASG in Step 4.
3	Create New ASGs: Block apps from sending traffic to system components, but allow them to send traffic to the services they need.
4	Remove the Default ASG: After you create and bind new ASGs, you no longer need the deployment-wide default ASG bindings that allow apps to send traffic to any IP.
5	Restart your Apps: To apply the ASG changes, you must restart all of the apps in your deployment.

- **When to set up new rules:**
 - Pivotal recommends that you complete this procedure directly after installing PCF, prior to developers pushing apps to the platform. If you complete the procedure after apps have been pushed to the platform, you must restart all the apps in your deployment.

Prerequisites

The procedure below requires that you have the latest release of [ASG Creator](#) from the Cloud Foundry incubator repository on Github. See [About the ASG Creator Tool](#).

Procedure

Follow these steps to apply ASGs that prevent apps running on your deployment from accessing internal PCF components.

Step 1: Determine Your Network Layout

The procedure for securing your deployment with ASGs varies depending on your network layout, which you can determine by following these steps:

1. Log in to Ops Manager.

2. For each tile, click **Assign AZs and Networks** and record the selected **Network** that the tile is installed on.

3. Based on the information you gathered, determine which of the following network layouts you have:

Layout Name	Layout Description
One Network	<ul style="list-style-type: none"> o One network for Ops Manager and the BOSH Director, Pivotal Application Service (PAS), and services. <div style="background-color: #ffffcc; padding: 5px;"> 💡 Note: You cannot secure your deployment with ASGs if you have this network layout. Because PCF dynamically allocates IPs, they cannot be easily excluded in the case of a single network. </div>
Two Networks	<ul style="list-style-type: none"> o One network for Ops Manager and the BOSH Director. o One network for PAS and Services.
Three Networks	<ul style="list-style-type: none"> o One network for Ops Manager and the BOSH Director. o One network for PAS. o One network for all services.
Three or More Networks	<ul style="list-style-type: none"> o One network for Ops Manager and the BOSH Director. o One network for PAS. o One network for each service.

4. If your network layout includes two or more networks, continue [Step 2: Ensure Access for PCF System Apps](#).

Step 2: Ensure Access for PCF System Apps

Follow these steps to apply the default ASG to the `system` org. This provides network access to PCF system apps without restrictions, which enables them to continue functioning properly after you perform [Step 4: Remove the Deployment-wide Default ASG Binding](#).

1. Bind the default ASG to the `staging set` in the `system` org:

```
$ cf bind-staging-security-group default_security_group
```

2. Bind the default ASG to the `running set` in the `system` org:

```
$ cf bind-running-security-group default_security_group
```

Step 3: Create New ASGs

Follow these steps to create ASGs that block apps from accessing PCF components and create any additional ASGs that allow apps to access the services they require.

Part A: Record CIDRs

Gather the CIDRs for each network in your deployment:

1. From the BOSH Director tile, click **Create Networks** within the **Settings** tab.
2. In the **Networks** section, expand each network in your deployment by clicking its name.
3. Record the **CIDR** for each network.

Part B: Create and Bind ASGs that Block Network Access

Create ASGs that block apps from sending traffic to the networks that host Ops Manager, PAS, and, optionally, any installed services.

1. Create a `config.yml` containing the appropriate content for your network layout and replace the indicated values with the CIDRs you gathered:

- o **Two Network Layout:**

```
exclude:  
- YOUR-OPS-MANAGER-CIDR  
- YOUR-PAS-AND-SERVICES-CIDR
```

- o **Three Network Layout:**

 **Note:** If you want to secure only the Ops Manager and PAS components, you can remove the services CIDR from the `exclude` section.

```
exclude:  
- YOUR-OPS-MANAGER-CIDR  
- YOUR-PAS-CIDR  
- YOUR-SERVICES-CIDR
```

- o **Three or More Network Layout:**

 **Note:** If you want to secure only the Ops Manager and PAS components, you can remove the services CIDRs from the `exclude` section.

```
exclude:  
- YOUR-OPS-MANAGER-CIDR  
- YOUR-PAS-CIDR  
- YOUR-SERVICE-CIDR-1  
- YOUR-SERVICE-CIDR-2  
etc...
```

2. Run the following command to generate the default `public-networks.json` and `private-networks.json` files that contain your ASG rules, specifying the location of the `config.yml` file as input:

```
$ asg-creator create --config config.yml
```

3. Create the `public-networks` ASG by running the following command:

```
$ cf create-security-group public-networks public-networks.json
```

4. Bind the ASG to the default staging set:

```
$ cf bind-staging-security-group public-networks
```

5. Bind the ASG to the default running set:

```
$ cf bind-running-security-group public-networks
```

6. Create the `private-networks` ASG by running the following command:

```
$ cf create-security-group private-networks private-networks.json
```

7. Bind the ASG to the default staging set:

```
$ cf bind-staging-security-group private-networks
```

8. Bind the ASG to the default running set:

```
$ cf bind-running-security-group private-networks
```

 **Note:** You can create and bind additional ASGs by following the procedures in [Create ASGs](#) and [Bind ASGs](#).

Part C: Create and Bind ASGs for Service Access

Note: This part is only necessary if you blocked apps from a network that hosts services in the previous part. If you did not block apps from a network that hosts services, proceed to [Step 4: Remove the Default ASG](#).

warning: In the two network layout, PAS and services share the same network. This means that each time you create an ASG that allows apps to access a new port and protocol within the network, you further expose the PAS component VMs. This is a limitation of a two network layout. For guidance on network topology, see [Reference Architectures](#).

Now that you have created ASGs to secure the Ops Man, PAS, and service components, work with developers to create additional ASGs that give apps access to the services they need.

For more information about creating and binding ASGs, see the following:

- [Managing ASGs with the cf CLI](#)
- [Typical ASGs](#)

Step 4: Remove the Default ASG

Now that you have bound new ASGs to determine outbound traffic rules, you no longer need the default ASG bindings that allow apps to send traffic to any IP address.

1. Unbind the default ASG from the staging set:

```
$ cf unbind-staging-security-group default_security_group
```

2. Unbind the default ASG from the running set:

```
$ cf unbind-running-security-group default_security_group
```

Step 5: Restart your Apps

To apply the ASG changes, you must restart all of the apps in your deployment. To mitigate app downtime during the restart, Pivotal recommends a [blue-green](#) deployment strategy.

Notes: You do not need to restart the apps in the `system` org.

1. Work with developers to restart a few of their apps individually and test that they still work correctly with the new ASGs in place. If an app does not work as expected, you likely must create another ASG that allows the app to send traffic to a service it requires.

Note: To quickly roll back to the original overly-permissive state, you can re-bind the `default_security_group` ASG to the `default-staging` and `default-running` sets. You must then restart your apps to re-apply the original ASGs.

2. Restart the rest of the apps running on your deployment. Optionally, you can use the [app-restarter cf CLI plugin](#) to restart all apps in a particular space, org, or deployment.

Configuring Application Security Groups for Email Notifications

To allow the Notifications Service to have network access you need to create [Application Security Groups](#) (ASGs).

 **Note:** Without Application Security Groups the service is not usable.

Prerequisite

Review the [Getting Started with the Notifications Service](#) topic to ensure you have setup the service.

Configure Network Connections

The Notifications Service is deployed as a suite of applications to the `notifications-with-ui` space in the `system` org, and requires the following outbound network connections:

Destination	Ports	Protocol	Reason
<code>SMTP_SERVER</code>	587 (default)	tcp (default)	This service is used to send out email notifications
<code>LOAD_BALANCER_IP</code>	80, 443	tcp	This service will access the load balancer
<code>ASSIGNED_NETWORK</code>	3306	tcp	This service requires access to internal services. <code>ASSIGNED_NETWORK</code> is the CIDR of the network assigned to this service.

 **Note:** The SMTP Server port and protocol are dependent on how you configure your server.

Create a SMTP Server ASG

1. Navigate to the Ops Manager **Installation Dashboard** and click the **Pivotal Application Service** tile > **Settings** tab.
2. Record the information in the **Address of SMTP Server** and **Port of SMTP Server** fields.
3. Using the **Address of SMTP Server** information you obtained in the previous step, find the IP addresses and protocol of your SMTP Server from the service you are using. You might need to contact your service provider for this information.
4. Create a `smtp-server.json` file. For `destination`, you must enter the IP address of your SMTP Server.

```
[  
  {  
    "protocol": "tcp",  
    "destination": "SMTP_SERVER_IPS",  
    "ports": "587"  
  }  
]
```

5. Create a security group called `smtp-server`:

```
cf create-security-group smtp-server smtp-server.json
```

Create a Load Balancer ASG

 **Note:** If you already have a ASG setup for a Load Balancer, you do not need to perform this step. Review your [ASGs](#) to check which groups you have setup.

If you are using the built-in HAProxy as your load balancer, follow this procedure. If you are using an external load balancer, you must obtain your HAProxy IPs from the service you are using.

1. Record the HAProxy IPs in the **Pivotal Application Service** tile > **Settings** > **Networking** tab.
2. Create a `load-balancer-https.json` file. For `destination`, use the **HAProxy IPs** you recorded above.

```
[  
  {  
    "protocol": "tcp",  
    "destination": "10.68.196.250",  
    "ports": "80,443"  
  }  
]
```

3. Create a security group called `load-balancer-https`:

```
$ cf create-security-group load-balancer-https load-balancer-https.json
```

Create an Assigned Network ASG

 **Note:** If you use external services, the IP addresses, ports, and protocols depend on the service.

1. Navigate to the Ops Manager **Installation Dashboard** > **Pivotal Application Service** tile > **Settings** > **Assign AZs and Networks** section.
2. Navigate to the network selected in the dropdown.
3. Record the **BOSH Director** tile > **Settings** tab > **Create Networks** > **CIDR** for the network identified in the previous step. Ensure the subnet mask allows the space to access `p-mysql`, `p-rabbitmq`, and `p-redis`.
4. Create a file `assigned-network.json`. For the `destination`, enter the **CIDR** you recorded above.

```
[  
  {  
    "protocol": "tcp",  
    "destination": "10.68.0.0/20",  
    "ports": "3306,5672,6379"  
  }  
]
```

5. Create a security group called `assigned-network`:

```
$ cf create-security-group assigned-network assigned-network.json
```

Bind the ASGs

1. Target the `system` org:

```
$ cf target -o system
```

2. Create a `notifications-with-ui` space:

```
$ cf create-space notifications-with-ui
```

3. Bind the ASGs you created in this topic to the `notifications-with-ui` space:

```
$ cf bind-security-group smtp-server system notifications-with-ui  
$ cf bind-security-group load-balancer-https system notifications-with-ui  
$ cf bind-security-group assigned-network system notifications-with-ui
```

Configuring SSH Access for PCF

To help troubleshoot applications hosted by a deployment, [Pivotal Cloud Foundry \(PCF\)](#) supports SSH access into running applications. This document describes how to configure a PCF deployment to allow SSH access to application instances, and how to configure load balancing for those application SSH sessions.

Pivotal Application Service Configuration

This section describes how to configure Pivotal Application Service (PAS) to enable or disable deployment-wide SSH access to application instances. In addition to this deployment-wide configuration, Space Managers have SSH access control over their Space, and Space Developers have SSH access control over their Applications. For details about SSH access permissions, see the [Application SSH Overview](#) topic.

To configure PAS SSH access for app instances:

1. Open the PAS tile in Ops Manager.
2. Under the **Settings** tab, select the **Application Containers** section.
3. Enable or disable the **Allow SSH access to app containers** checkbox.
4. Optionally, select **Enable SSH when an app is created** to enable SSH access for new apps by default in spaces that allow SSH. If you deselect this checkbox, developers can still enable SSH after pushing their apps by running `cf enable-ssh APP-NAME`.

Enable microservice frameworks, private Docker registries, and other services that support your applications at a container level.

Enable Custom Buildpacks

Allow SSH access to app containers

Enable SSH when an app is created

When enabled and the space has SSH enabled, a new app will have SSH enabled. Turning this feature off will disable SSH for new apps but allow developers to enable it when necessary.

Router application identity verification*

Router uses TLS to verify application identity (Default)

Router and applications use mutual TLS to verify each other's identity (incompatible with TCP Routing and SSH access to app containers)

Private Docker Insecure Registry Whitelist

10.10.10.10:8888,example.com:8888



Docker Images Disk-Cleanup Scheduling on Cell VMs*

Never clean up Cell disk-space

Routinely clean up Cell disk-space

Clean up disk-space once usage fills disk

Reserved amount of Disk (MB) (min: 1) *

15360

SSH Load Balancer Configuration

For IaaSes where load-balancing is available as a service, you should provision a load balancer to balance load across SSH proxy instances. Configure this load balancer to forward incoming TCP traffic on port `2222` to a target pool where you deploy `diego_brain` instances.

For AWS, Azure, and GCP IaaSes, you configure SSH load balancers in the **Resource Config** pane. To register SSH proxies with a load balancer, enter your load balancer name in the **Load Balancers** field in the **Diego Brain** row.

Ops Manager supports an API-only `nsx_lbs` field. You can configure load balancers in vSphere using this field.

Securing Services Instance Credentials with Runtime CredHub

This topic describes how Pivotal Cloud Foundry (PCF) operators can ensure service instance credentials are securely stored in CredHub.

- **What is runtime CredHub?**
 - The Pivotal Application Service (PAS) tile includes its own CredHub component, separate from the CredHub component included with the BOSH Director tile. For more information about this centralized credential management component, see the [CredHub](#) documentation.
- **What is runtime CredHub used for?**
 - Runtime CredHub exists to securely store service instance credentials. Previously, PCF could only use the Cloud Controller database for storing these credentials.
- **What are service instance credentials?**
 - When developers want their app to use a service, such as those provided by the Spring Cloud Services tile for PCF, they must bind their app to an instance of that service. Service bindings include credentials that developers can use to access the service from their app. For more information, see [Binding Credentials](#).
- **How can I ensure that service instance credentials are stored in runtime CredHub?**
 - You must configure PAS to enable this functionality by following the procedure below. Not all services support the use of runtime CredHub.
- **Can I use runtime CredHub to store service instance credentials if some of my services do not support the use of runtime CredHub?**
 - PAS supports both services that do and do not use runtime CredHub. Services that do not use runtime CredHub continue to pass their credentials to the Cloud Controller database.
- **Can I rotate service instance credentials in runtime CredHub?**
 - Runtime CredHub supports credential rotation. For more information, see [Rotating Runtime CredHub Encryption Keys](#).

PCF Services that Use Secure Binding Credentials

The procedures in this document are only effective for services that support storing their instance credentials in runtime CredHub. To learn whether a service supports this feature, see the table below or the documentation for that service.

Service	Versions with Secure Binding Credentials	Pivotal Network Listing
CredHub Service Broker for PCF	All	CredHub Service Broker for PCF
MySQL for PCF v2	v2.3 and later (available to user groups only)	MySQL for PCF v2
Spring Cloud Services for PCF	v1.5 and later	Spring Cloud Services for PCF
RabbitMQ for PCF	v1.12 and later	RabbitMQ for PCF
Redis for PCF	v1.13 and later	Redis for PCF

Prerequisites

 **Breaking Change:** If you opt out of the [BOSH DNS feature](#), your PCF deployment cannot support Secure Service Instance Credentials.

Runtime CredHub allows you to use one or more Hardware Security Modules (HSMs) to store encryption keys. If you wish to use an HSM with CredHub, you must configure the HSM before completing the procedures below. For more information, see [Preparing CredHub HSMs for Configuration](#).

Step 1: Configure the PAS Tile

To configure the PAS tile to support securing service instance credentials in CredHub, you need to:

- In the **CredHub** pane, provide at least one encryption key. CredHub supports multiple encryption key providers.
- In the **Resource Config** pane, set the number of **CredHub** instances to at least one.

CredHub configuration options include:

- Internal or external databases
- Encryption keys stored internally, externally in a Hardware Security Module (HSM), or both

To configure the PAS tile, follow the instructions in [Configuring PAS](#).

Step 2: Create Application Security Groups

Application Security Groups (ASGs) are network policy rules specifying protocols, ports, and IP ranges that apply to outbound network connections initiated from apps. You must follow the steps below to ensure the ASGs for your deployment allow apps to communicate with the runtime CredHub API.

Note: The default ASGs in PCF allow apps running on your deployment to send traffic to almost any IP address. This step is only required if your ASGs restrict network access from apps to the network that PAS runs on, as documented in [Restricting App Access to Internal PCF Components](#).

1. From the PAS tile, click **Assign AZs and Networks** and record the selected Network where the tile is installed.
2. From the BOSH Director tile, within the **Settings** tab, click **Create Networks**.
3. In the **Networks** section, click the name of the PAS network to expand it.
4. Record the CIDR for the PAS network.
5. Create a file named `runtime-credhub.json` for specifying your ASG rules. Copy the content below into the file. Replace `YOUR-PAS-CIDR` with the CIDR you recorded in the previous step.

```
[  
 {  
   "protocol": "tcp",  
   "destination": "YOUR-PAS-CIDR",  
   "ports": "8844"  
 }  
]
```

6. Run the following command to create an ASG that allows apps to access the CredHub API:

```
$ cf create-security-group runtime-credhub ~/workspace/runtime-credhub runtime-credhub.json
```

7. Bind this ASG to your deployment or the specific space in which you want apps to access CredHub. For more information about binding ASGs, see [Bind ASGs](#). Ensure that apps deployed as part of the service tile installation process have access to CredHub in addition to the apps pushed to the platform by developers. For example, the Spring Cloud Services tile deploys the `spring-cloud-broker` app to the `p-spring-cloud-services` space of the `system` org.
8. Restart apps for the ASGs to take effect. Optionally, you can use the [app-restarter cf CLI plugin](#) to restart all apps in a particular space, org, or deployment.

Step 3: Unbind and Rebind Service Instances

For any service instance bindings that existed before runtime CredHub was supported for that service, you must work with your developers to unbind and rebind the service instances to their apps. If you do not unbind and rebind the service, apps continue functioning as normal and fetching credentials from the Cloud Controller database.

Note: This step is not required for bindings created after you installed the new version of the service tile that supports CredHub and you completed the procedures in steps 1 and 2 of this topic.

1. Unbind the service instance from the app:

```
$ cf unbind-service YOUR-APP YOUR-SERVICE-INSTANCE
```

2. Rebind the service instance to the app:

```
$ cf bind-service YOUR-APP YOUR-SERVICE-INSTANCE
```

3. Review the `VCAP_SERVICES` environment variable to verify that the new service instance binding includes CredHub pointers:

```
$ cf env YOUR-APP
```

See the [VCAP_SERVICES](#) section of *Cloud Foundry Environment Variables* for help parsing the output of the `cf-env` command.

4. Restart the app to apply the service instance binding:

```
$ cf restart YOUR-APP
```

If you run `cf-env` again, you can see the `VCAP-SERVICES` environment variable now contains the credentials for the service instance binding.

Identifying PAS Jobs Using vCenter

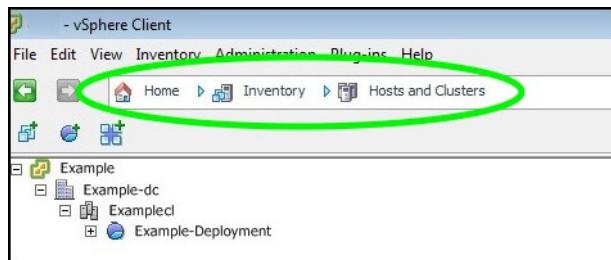
To effectively monitor, control, and manage the virtual machines making up your Pivotal Application Service (PAS) deployment, you may need to identify which VM corresponds to a particular job in PAS. You can find the CID of a particular VM from [Pivotal Cloud Foundry](#) (PCF) Operations Manager by navigating to **PAS Status**.

If you have deployed PAS to VMware vSphere, you can also identify which PAS job corresponds to which VM using the vCenter vSphere client.

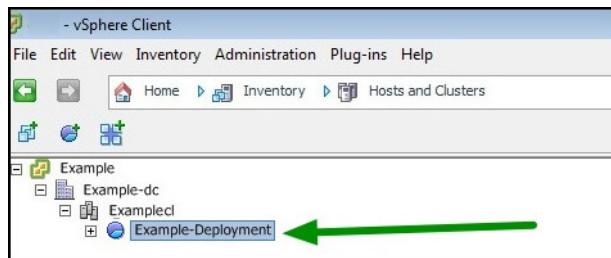
 **Note:** The CID shown in Ops Manager is the name of the machine in vCenter.

Identifying PAS Jobs Using vCenter

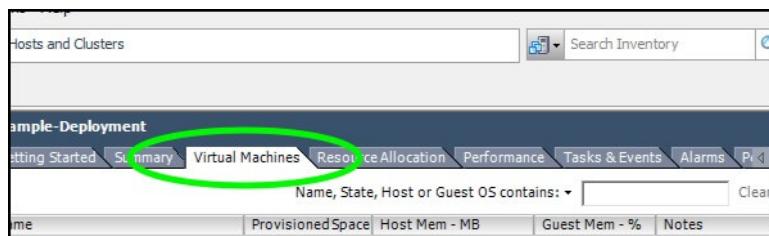
1. Launch the vSphere client and log in to the vCenter Server system.
2. Select the **Inventory > Hosts and Clusters** view.



3. Select the Resource Pool containing your PAS deployment.



4. Select the **Virtual Machines** tab.



5. Right-click the column label heading and check **job**.

Example-Deployment

Getting Started Summary Virtual Machines Resource Allocation Performance Tasks & Events Alarms Permissions Maps

Name, State, Host or Guest OS contains:

	Provisioned Space	Host Mem - MB	Guest
Host	9.09 GB	1040	
Provisioned Space	11.09 GB	597	
Used Space	151.09 GB	1036	45
Host CPU - MHz	10.09 GB	598	
Host Mem - MB	19.09 GB	1023	88
Guest Mem - %	6.69 GB	0	
Guest OS	9.09 GB	1045	
VM Version	9.09 GB	470	
Memory Size	9.09 GB	468	
Reservation - MB	20.09 GB	3130	
CPU Count	6.69 GB	0	
NIC Count	9.09 GB	551	
Uptime	6.69 GB	0	
IP Address	6.69 GB	0	
VMware Tools Running Status			
VMware Tools Version Status			
DNS Name			
EVC Mode			
UUID			
Notes			
Alarm Action			
vSphere HA Protection			
Needs Consolidation			
Name			
compiling			
deployment			
director			
index			
job			

Name, Target or Status contains:

	Requested Start Time	Start Time
	3/6/2014 12:21:36 PM	3/6/2014 12:21:36 PM
	3/6/2014 12:21:32 PM	3/6/2014 12:21:32 PM

6. The job column displays the PAS job associated with each virtual machine.

Example-Deployment

Getting Started Summary Virtual Machines Resource Allocation Performance Tasks & Events Alarms Permissions Maps

Name, State, Host or Guest OS contains:

Name	job
vm-347a89c6-115f-433a-83d1-38cd2c31451b	ccdb
sc-d3520bac-11c2-460f-aa02-60fc11cba375	cloud_controller
vm-d537f816-fc51-450c-8ada-cb536a9e765	consoledb
vm-85d43147-7cd2-4cc5-ac6c-47cce18cccd9	dea
vm-5c04a27a-ca70-4f67-a221-767efb76922a	ha_proxy
vm-e9b3e319-690f-44be-87fb-ce97f90c4b2	health_manager
vm-a8bd3d1-7e31-47f3-800f-3f81ff23ed81	loggregator
vm-ee8b5374-bbc9-436b-bfe5-1a7e2b5fee9a	loggregator_trafficcontroller
vm-5896908f-15c5-4989-8c90-36a36e0eb4fd	nats
vm-61fc0d19-01f5-4434-82e0-a11e15b38888	nfs_server
vm-e88725cd-36fe-4a79-b8ea-f35b2d2cd85b	riak-cs
vm-9ab56b339-4cae-443a-842c-3651c9bf228c	router
vm-2726dcda-7f92-440c-906e-3e5e544fe628	saml_login
vm-273df76f-7141-432d-9e53-9cb6549d870a	uaa
vm-307a34f4-ed7f-4c48-94d1-bb7a61a695bc	uaadb
vm-e6c2dfcd-ef2e-40d8-af0b-06e5f3e73fa9	
vm-0e9e9384-0057-4efc-9cea-234b5956c510	
vm-5e33e793-0118-44bb-948d-3ea6526b4b2f	

Configuring Logging in PAS

This topic describes the types of logs that Pivotal Cloud Foundry (PCF) generates. It also explains how to forward system logs to an external aggregator service, how to scale [Loggregator](#) component VMs to keep up with app log volume, and how to manage app traffic logging.

System Logs, App Logs, App Traffic Logs

Pivotal Cloud Foundry (PCF) generates two types of logs, *system logs* from PCF components and *app logs* from hosted apps, as differentiated in the table below:

Log Type	Originate from	Follow format	Stream from	Can to stream out to (configurable)	Visible to	
System Logs	Platform components	Syslog standard	rsyslog agent	Component syslog drain	Operators	
App Logs	Hosted apps	Format is up to the developer	Firehose ¹	External data platform, optionally via nozzles	Developers and Operators	
		(Optional) With Syslog Adapter				
		Converted to syslog standard	Syslog Adapter	External syslog drain		

¹The Loggregator Firehose also streams component metrics.

App traffic logs are system logs. When app containers communicate, or attempt to communicate, their host cells generate *app traffic logs*. App traffic logs are system logs, not app logs. These logs come from host cells, not apps, and they carry no information from within the app. App traffic logs only show app communication behavior, as detected from outside by the host cell.

Log Cache

Log Cache is a Loggregator feature that lets you filter and query app logs through a CLI plugin or API endpoints. Cached app logs are available on demand; you do not need to stream them continuously.

Example Uses of the Log Cache CLI Plugin

To access cached logs with the Log Cache CLI plugin, you must first download and install the plugin.

[Download the Log Cache CLI plugin ↗](#).

Once you have installed the plugin, the basic command to access cached app logs is:

```
$ cf tail [OPTIONS] [SOURCE ID/APP]
```

Here are some ways you can use Log Cache to filter app logs:

- `--start-time` : Displays the start of the cache or the start of a time period you specify. Results display as a UNIX timestamp, in nanoseconds. Pair with `--end-time` to view logs within a time period.
- `--end-time` : Displays the end of the cache or the end of a time period you specify. Results display as a UNIX timestamp, in nanoseconds. Pair with `--start-time` to view logs within a time period.
- `--json` : Output logs in JSON.
- `--follow` : Append exported logs to `stdout`.

For more information on using the Log Cache CLI, see [Log Cache CLI: Usage ↗](#).

Example Uses of the Log Cache API

The Log Cache API is hosted on PCF, and references your system domain to return responses. The root URL for API calls is [https://log-cache.\[YOUR-SYSTEM-DOMAIN\]](https://log-cache.[YOUR-SYSTEM-DOMAIN]).

The basic call to access and filter cached app logs is:

```
GET https://log-cache.[YOUR-SYSTEM-DOMAIN]/v1/read/[SOURCE-ID]
```

Append the following parameters to the `GET` to customize app logs:

- `start_time` : Displays the start of the cache or the start of a time period you specify. Results display as a UNIX timestamp, in nanoseconds. Pair with `end_time` to view logs within a time period.
- `end_time` : Displays the end of the cache or the end of a time period you specify. Results display as a UNIX timestamp, in nanoseconds. Pair with `start_time` to view logs within a time period.
- `envelope_types` : Filters by Envelope Type. The available filters are: `LOG`, `COUNTER`, `GAUGE`, `TIMER`, and `EVENT`. Set an envelope type filter to emit logs of only that type. Specify this parameter multiple times to include more types.
- `limit` : Sets a maximum number of envelopes to request. The max limit is 1000. This value defaults to 100.

More API parameters are available to customize retrieved app logs. For more information, see [Log Cache: RESTful API Gateway](#).

Enable Syslog Forwarding

You can configure system logging in PAS to forward log messages from PAS component VMs to an external service. Pivotal recommends forwarding logs to an external service for use in troubleshooting.

 **Note:** The following instructions explain how to configure system logging for PAS component VMs. To forward logs from PCF tiles to an external service, you must also configure system logging in each tile. See the documentation for the given tiles for information about configuring system logging.

To configure system logging in PAS, do the following:

1. In the PAS **Settings** tab, select the **System Logging** pane. The following image shows the **System Logging** pane.

Optional configuration for rsyslog to forward platform component logs to an external service. If you do not fill these fields, platform logs will not be forwarded but will remain available on the component VMs and for download via Ops Manager.

Address

Port

Transport Protocol

Encrypt syslog using TLS?*

- No
- Yes

Syslog Drain Buffer Size (# of messages) *

Include container metrics in SysLog Drains

Enable Cloud Controller security event logging

Use TCP for file forwarding local transport

Don't Forward Debug Logs

Custom rsyslog Configuration

Save

2. For **Address**, enter the IP address of the syslog server.

3. For **Port**, enter the port of the syslog server. The default port for a syslog server is **514**.

Note: The host must be reachable from the PAS network and accept UDP or TCP connections. Ensure the syslog server listens on external interfaces.

4. For **Transport Protocol**, select a transport protocol for log forwarding.

5. For **Encrypt syslog using TLS?**, select **Yes** to use TLS encryption when forwarding logs to a remote server.

- a. For **Permitted Peer**, enter either the name or SHA1 fingerprint of the remote peer.
- b. For **TLS CA Certificate**, enter the TLS CA certificate for the remote server.

6. For **Syslog Drain Buffer Size**, enter the number of messages from the Loggregator Agent that the Doppler server can store before it begins to drop messages. See the [Loggregator Guide for Cloud Foundry Operators](#) topic for more details.

7. Disable the **Include container metrics in Syslog Drains** checkbox to prevent the [CF Drain CLI plugin](#) from including app container metrics in syslog drains. This feature is enabled by default.

8. Enable **Enable agent-based syslog egress for app logs** to use agent-based syslog. Agent-based syslog provides an alternative to scalable-syslog architecture. For more information about agent-based syslog, see [Loggregator Architecture](#).

 **Note:** Selecting this checkbox disables the Syslog Adapter and Syslog Scheduler VMs to prevent log duplication.

9. Enable the **Enable Cloud Controller security event logging** checkbox to include security events in the log stream. This feature logs all API requests, including the endpoint, user, source IP address, and request result, in the Common Event Format (CEF).
10. Enable the **Use TCP for file forwarding local transport** checkbox to transmit logs over TCP. This prevents log truncation, but may cause performance issues.
11. Disable the **Don't Forward Debug Logs** checkbox to forward DEBUG syslog messages to an external service. This checkbox is enabled by default.

 **Note:** Some PAS components generate a high volume of DEBUG syslog messages. Enabling the **Don't Forward Debug Logs** checkbox prevents PAS components from forwarding the DEBUG syslog messages to external services. However, PAS still writes the messages to the local disk.

12. For **Custom syslog Configuration**, enter a custom syslog rule. For more information about adding custom syslog rules, see [Customizing Syslog Rules](#).
13. Select **Enable System Metrics** to emit system-level metrics about all VMs in the deployment. For a list of the VM metrics that the System Metric Agent emits, see [VM Metrics](#) in GitHub.
14. Click **Save**.

To configure Ops Manager for system logging, see the [Settings](#) section in the *Using the Ops Manager Interface* topic.

Include Container Metrics in Syslog Drains

Developers can monitor container metrics over the syslog protocol using the [CF Drain CLI plugin](#). With the CF Drain CLI plugin, you can use the Cloud Foundry Command Line Interface (cf CLI) tool to set the app container to deliver container metrics to a syslog drain. Developers can then monitor the app container based on those metrics.

In PAS, the **Include container metrics in Syslog Drains** checkbox is enabled by default in the **System Logging** pane. If you have security concerns about streaming container metrics from your app, you can disable this checkbox.

For more information, see [Including Container Metrics in Syslog Drains](#) in *Application Logging in Cloud Foundry*.

Scale Loggregator

Apps constantly generate app logs and PCF platform components constantly generate component metrics. The Loggregator system combines these data streams and handles them as follows. See [Loggregator Architecture](#) for more information.

- The [Loggregator](#) agent running on each component or application VM collects and sends this data out to Doppler components.
- [Doppler](#) components temporarily buffer the data before periodically forwarding it to the Traffic Controller. When the log and metrics data input to a Doppler exceeds its buffer size for a given interval, data can be lost.
- The [Traffic Controller](#) serves the aggregated data stream through the Firehose WebSocket endpoint.

Follow the instructions below to scale the Loggregator system. For guidance on monitoring and capacity planning, see [Monitoring Pivotal Cloud Foundry](#).

Add Component VM Instances

1. From the PAS tile, select **Resource Config**.
2. Increase the number in the **Instances** column of the component you want to scale. You can add instances for the following Loggregator components:
 - [Loggregator Traffic Controller](#)

Note: The [Reverse Log Proxy \(RLP\)](#) BOSH job is colocated on the Traffic Controller VM. If you want to scale Loggregator to handle more logs for syslog drains, you can add instances of the Traffic Controller.

Note: The [BOSH System Metrics Forwarder](#) job is colocated on the Traffic Controller VM. If you want to scale Loggregator to handle more BOSH component metrics, you can add instances of the Traffic Controller.

- [Syslog Adapter](#)
- [Doppler Server](#)

Loggregator	Automatic: 3	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: !)	None	<input checked="" type="checkbox"/>
Trafficcontroller	Automatic: 3	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: !)	None	<input checked="" type="checkbox"/>
Syslog Adapter	Automatic: 3	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: !)	None	<input checked="" type="checkbox"/>
Syslog Scheduler	Automatic: 1	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: !)	None	<input checked="" type="checkbox"/>
Doppler Server	Automatic: 3	None	Automatic: micro (cpu: 1, ram: 1 GB, disk: !)	None	<input checked="" type="checkbox"/>

3. Click **Save**.

4. Click **Review Pending Changes**, then **Apply Changes**.

App Traffic Logging

App traffic logging generates logs when app containers communicate with each other directly, or attempt to communicate, as allowed by [container-to-container networking \(C2C\) policies](#) and [Application Security Groups \(ASGs\)](#).

App traffic logging lets network security teams audit C2C traffic, by seeing allowed and denied packets, without needing access to the Cloud Controller or the apps themselves.

Enable App Traffic Logging

To enable app traffic logging:

1. From Ops Manager, navigate to the **Pivotal Application Service** tile > **Networking** pane.
2. Under **Log traffic for all accepted/denied application packets**, select **Enable (will increase log volume)** or **Disable** to enable or disable app traffic logging.

Log traffic for all accepted/denied application packets.*

Enable (will increase log volume)

Denied logging interval (packets/second) *

UDP logging interval (packets/second) *

Disable

App Logging Behavior

App traffic logging generates log messages as follows:

- **TCP traffic** - Logs the first packet of every new TCP connection.
- **UDP traffic** - Logs UDP packets sent and received, up to a maximum per-second rate for each container. Set this rate limit in the **UDP logging interval** field (default: 100).
- **Packets denied** - Logs packets blocked by either a container-specific [networking policy](#) or by [Application Security Group \(ASG\)](#) rules applied across

the space, org, or deployment. Logs packet denials up to a maximum per-second rate for each container, set in the **Denied logging interval** field (default: 1).

App Traffic Log Format and Contents

App traffic logs are formatted as described in the [cf-networking-release Traffic logging](#) documentation, following the `iptables-logger` format but without line breaks. For example, the first part of an app traffic log line looks like:

```
{"timestamp": "1500924070.182554722", "source": "cfnetworking.iptables", "message": "cfnetworking.iptables.ingress-allowed", "log_level": 1, "data": { "destination": { "container_id": "d5978989-1401-49ff-46cd-33e5", "app_guid": "bc6f229d-5e4a-4c41-a63f-e8795496c283",
```

Each log message includes the following:

- Timestamp
- The GUID for the source or destination app that sent or was designated to receive the packet
- The protocol of the communication, `TCP` or `UDP`
- GUIDs for the container, space, and org running the source or destination app
- IP addresses and ports for both source and destination apps
- A `message` field recording whether the packet was allowed or denied, with one of the following four possibilities:
 - ASG allowed packet to exit source app container
 - C2C policy allowed packet to enter destination app container
 - ASG prevented packet from exiting source app container
 - C2C policy prevented packet from entering destination app container
- Additional information described in the [cf-networking-release](#).

Denied Packet Causes

You can determine whether a denied-packet log resulted from a container networking policy or an ASG rule as follows:

- **Container networking policy:** Log `message` string includes `ingress-denied` and `packet direction` is `ingress`.
- **ASG rule:** Log `message` string includes `egress-denied` and `packet direction` is `egress`.

Global vs. ASG-Level App Traffic Logging

PCF supports two mechanisms for enabling app traffic logging. [Setting Log traffic to Enable](#) in Ops Manager enables app traffic logging globally for all ASGs and container policies. Setting the `log` property of an ASG to `true` enables app traffic logging at the individual ASG level.

Because these two mechanisms operate independently, PCF generates duplicate logs when app traffic logging is enabled globally and an ASG's `log` property is set to `true`. To avoid duplicate logs, Pivotal recommends setting the `log` property to `false` for all ASGs, or leaving it out entirely, when app traffic logging is [enabled globally](#).

To focus on specific ASGs for log analysis, Pivotal recommends enabling app traffic logs globally and using a logging platform to filter traffic logs by ASG, rather than setting `log` at the individual ASG level.

Configuring UAA Password Policy

If your Pivotal Cloud Foundry (PCF) deployment uses the internal user store for authentication, you can configure its password policy within the Pivotal Application Service (PAS) tile.

Open the Internal UAA Configuration

1. In a browser, navigate to the fully qualified domain name (FQDN) of your Ops Manager and log in.
2. Click the **Pivotal Application Service** tile.
3. Select **Authentication and Enterprise SSO** on the **Settings** tab.

Configure your user store access, which can be an internal user store (managed by Cloud Foundry's UAA) or an external user store (LDAP or SAML). You can also adjust the lifetimes of authentication tokens.

Configure your UAA user account store with either internal or external authentication mechanisms*

- Internal UAA (provided by Elastic Runtime; configure your password policy below)

Minimum Password Length *

Minimum Uppercase Characters Required for Password *

Minimum Lowercase Characters Required for Password *

Minimum Numerical Digits Required for Password *

Minimum Special Characters Required for Password *

Maximum Password Entry Attempts Allowed *

4. Confirm that the **Internal UAA** option is selected.

Set Password Requirements and Entry Attempts

1. For **Minimum Password Length**, enter the minimum number of characters for a valid password.
2. For **Minimum Uppercase Characters Required for Password**, enter the minimum number of uppercase characters required for a valid password.
3. For **Minimum Lowercase Characters Required for Password**, enter the minimum number of lowercase characters required for a valid password.
4. For **Minimum Numerical Digits Required for Password**, enter the minimum number of digits required for a valid password.
5. For **Minimum Special Characters Required for Password**, enter the minimum number of special characters required for a valid password.
6. For **Maximum Password Entry Attempts Allowed**, enter the maximum number of failures allowed to enter a password within a five-minute

timespan before the account is locked.

Managing Internal MySQL for PAS

This topic is a reference for the concepts and procedures related to internal MySQL for Pivotal Application Service (PAS).

See the following topics:

- [Scaling Internal MySQL for PAS](#)
- [Running mysql-diag](#)
- [Recovering From MySQL Cluster Downtime](#)
- [Using the MySQL Proxy ↗](#)
- [MySQL Network Communications](#)

Scaling Internal MySQL for PAS

This topic describes scaling down your Pivotal Application Service (PAS) MySQL cluster and provides example sizing data from two environments that have significant load on their MySQL clusters.

For additional resources about scaling Internal MySQL, see the following documentation:

- *Deployments Using Internal MySQL* section of [Scaling PAS](#)
- *Scalable Components* section of [High Availability in Cloud Foundry](#)

 **Note:** The procedure does not apply to databases configured as external in the PAS tile **Databases** pane.

PAS components that use system databases include the Cloud Controller, Diego brain, Gorouter, and the User Authorization and Authentication (UAA) server. See [Cloud Foundry Components](#).

Scale Down Your MySQL Cluster

This procedure explains how to safely scale your MySQL cluster down to a single node. If you are already running the MySQL cluster with a single node, you do not need to perform these steps.

By default, internal MySQL deploys as a single node. To take advantage of the high availability features of MySQL, you may have scaled the configuration up to three or more server nodes.

Check the Health of Your Cluster

Before scaling down your MySQL cluster, perform the following actions to ensure the cluster is healthy.

1. Use the Cloud Foundry Command Line Interface (cf CLI) to target the API endpoint of your Pivotal Cloud Foundry (PCF) deployment:

```
$ cf api api.YOUR-SYSTEM-DOMAIN
Setting api endpoint to api.YOUR-SYSTEM-DOMAIN...
OK

API endpoint: https://api.YOUR-SYSTEM-DOMAIN... (API version: 2.54.0)
Not logged in. Use 'cf login' to log in.
```

2. Log in with your User Account and Authentication (UAA) Administrator user credentials. Obtain these credentials by clicking the **Credentials** tab of the Pivotal Application Service (PAS) tile, locating the **Admin Credentials** entry in the **UAA** section, and clicking **Link to Credential**.

```
$ cf login -u admin
API endpoint: https://api.YOUR-SYSTEM-DOMAIN

Password>
Authenticating...
OK
```

3. Create a test organization to verify the database across all nodes:

```
$ cf create-org data-integrity-test-organization
Creating org data-integrity-test-organization as admin...
OK

Assigning role OrgManager to user admin in org data-integrity-test-organization ...
OK

TIP: Use 'cf target -o data-integrity-test-organization' to target new org
```

4. Obtain the IP addresses of your MySQL server by performing the following steps:

- a. From the PCF **Installation Dashboard**, click the **Pivotal Application Service** tile.
- b. Click the **Status** tab.

- c. Record the IP addresses for all instances of the MySQL Server job.
5. Retrieve Cloud Controller database credentials from CredHub using the Ops Manager API:
- Perform the procedures in the [Using the Ops Manager API](#) topic to authenticate and access the Ops Manager API.
 - Use the `GET /api/v0/deployed/products` endpoint to retrieve a list of deployed products, replacing `UAA-ACCESS-TOKEN` with the access token recorded in the previous step:
- ```
$ curl "https://OPS-MAN-FQDN/api/v0/deployed/products" \
-X GET \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
```

- c. In the response to the above request, locate the product with an `installation_name` starting with `cf-` and copy its `guid`.
- d. Run the following `curl` command, replacing `PRODUCT-GUID` with the value of `guid` from the previous step:
- ```
$ curl "https://OPS-MAN-FQDN/api/v0/deployed/products/PRODUCT-GUID/variables?name=cc-db-credentials" \
-X GET \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
```

- e. Record the Cloud Controller database `username` and `password` from the response to the above request.
6. SSH into the Ops Manager VM. Because the procedures vary by IaaS, review the [SSH into Ops Manager](#) section of the Advanced Troubleshooting with the BOSH CLI topic for specific instructions.
7. For each of the MySQL server IP addresses recorded above, perform the following steps from the Ops Manager VM:
- Query the new organization with the following command, replacing `YOUR-IP` with the IP address of the MySQL server and `YOUR-IDENTITY` with the `identity` value of the CCDB credentials obtained above:
- ```
$ mysql -h YOUR-IP -u YOUR-IDENTITY -D ccdb -p -e "select created_at, name from organizations where name = 'data-integrity-test-organization'"
```
- When prompted, provide the `password` value of the CCDB credentials obtained above.
  - Examine the output of the `mysql` command and verify the `created_at` date is recent.
- | created_at          | name                             |
|---------------------|----------------------------------|
| 2016-05-28 01:11:42 | data-integrity-test-organization |

8. If each MySQL server instance does not return the same `created_at` result, contact [Pivotal Support](#) before proceeding further or making any changes to your deployment. If each MySQL server instance does return the same result, then you can safely proceed to scaling down your cluster to a single node by performing the steps in the following section.

## Set Server Instance Count to 1

- From the PCF Installation Dashboard, click the **Pivotal Application Service** tile.
- Select **Resource Config**.
- Use the dropdown to change the **Instances** count for **MySQL Server** to `1`.
- Click **Save**, then **Review Pending Changes** and **Apply Changes** to apply the changes.
- Delete your test organization with the following cf CLI command:

```
$ cf delete-org data-integrity-test-organization
```

## MySQL Cluster Sizing Examples

This topic describes two sizing examples for internal MySQL in PAS.

Use this data as guidance to ensure your MySQL Clusters are scaled to handle the number of app instances running on your deployment.

## Example 1: Pivotal Web Services Production Environment

The information in this section comes from the Pivotal-managed Cloud Foundry deployment, Pivotal Web Services (PWS).

 **Note:** This deployment differs from most PCF deployments in that the MySQL database is used for Diego, but not the Cloud Controller. This means that while there are a large number of queries per second, most of them are reading data, and the number of writes is not realistically reflected.

- **IAAS:** AWS
- **App Instances:** ~23,000
- **Average SQL Queries Per Minute:** ~308,000
- **Average IOPS:** 220
- **Storage Volume Usage:** 80%

### VM Sizing

The following table displays MySQL VM settings for this environment.

| Setting             | Value      |
|---------------------|------------|
| VM Type             | c4.2xlarge |
| Storage Volume Type | io1        |
| Storage Volume Size | 98GB       |
| Storage Volume IOPS | 2000       |

## Example 2: Cloud Foundry Diego Test Environment

The information in this section comes from an environment used by the Diego to test the MySQL cluster with a high load similar to one generated large deployment.

- **IAAS:** GCP
- **App Instances:** 250,000
- **Average SQL Queries Per Minute:** ~5,100,000
- **Average IOPS:**
  - **Reads:** 3.63
  - **Writes:** 363.98

### VM Sizing

The following table displays MySQL VM settings for this environment.

| Setting             | Value          |
|---------------------|----------------|
| VM Type             | n1-standard-16 |
| Storage Volume Type | pd-ssd         |
| Storage Volume Size | 1TB            |

## Running mysql-diag

This topic discusses how to use the `mysql-diag` tool in MySQL for Pivotal Cloud Foundry (PCF). `mysql-diag` prints the state of your MySQL highly available (HA) cluster and suggests solutions if your node fails. Pivotal recommends running this tool against your HA cluster before all deployments.

`mysql-diag` checks the following information about the status of your HA cluster:

- Membership status of all nodes
- Size as it appears to all nodes
- If it needs to be bootstrapped
- If replication is working
- Used disk space and inodes per server

## Run mysql-diag Using the BOSH Command Line Interface (CLI)

To use the BOSH CLI to run `mysql-diag`, do the following:

1. Obtain the information needed to use the BOSH CLI by doing the procedure in [Gather Credential and IP Address Information](#).
2. SSH into your Ops Manager VM by doing the procedure in [Log in to the Ops Manager VM with SSH](#) for your IaaS.
3. Log in to your BOSH Director by doing the procedure in [Log in to the BOSH Director](#).
4. Identify the VM to log in to with SSH by running the following command:

```
bosh -e MY-ENV -d MY-DEPLOYMENT vms
```

Where:

- `MY-ENV` is the name of your environment.
- `MY-DEPLOYMENT` is the name of your deployment.

5. Record the GUID associated with the `mysql-monitor` VM, also known as the jumpbox VM.

6. SSH into your `mysql-monitor` VM by running the following command:

```
bosh -e MY-ENV -d MY-DEP ssh mysql-monitor/GUID
```

Where:

- `MY-ENV` is the name of your environment.
- `MY-DEPLOYMENT` is the name of your deployment.
- `GUID` is the GUID you recorded in the previous step.

7. View the status of your HA cluster by running the following command:

```
mysql-diag
```

## Example Healthy Output

The `mysql-diag` command returns the following message if your canary status is healthy:

```
Checking canary status...healthy
```

Here is a sample `mysql-diag` output after the tool identified a healthy HA cluster:

```
mysql> monitor/5faae86a-a325-44fa-a8ab-24e736e42390:~$ mysql-diag
Wed Jul 19 19:03:54 UTC 2017
Checking canary status...
Checking canary status... healthy
Checking cluster status of mysql/fefced1e-83ce-4f7d-96f1-fb658b69cc81 at 10.244.9.2...
Checking cluster status of mysql/32665833-6a66-45dc-bad1-7a6476787d59 at 10.244.7.2...
Checking cluster status of mysql/026d5b79-6844-4227-9e6a-923dd24de9a6 at 10.244.8.2...
Checking cluster status of mysql/32665833-6a66-45dc-bad1-7a6476787d59 at 10.244.7.2... done
Checking cluster status of mysql/026d5b79-6844-4227-9e6a-923dd24de9a6 at 10.244.8.2... done
Checking cluster status of mysql/fefced1e-83ce-4f7d-96f1-fb658b69cc81 at 10.244.9.2... done
+-----+-----+-----+-----+
| HOST | NAME/UUID | WSREP LOCAL STATE | WSREP CLUSTER STATUS | WSREP CLUSTER SIZE |
+-----+-----+-----+-----+
10.244.7.2	mysql/32665833	Synced	Primary	3
10.244.8.2	mysql/026d5b79	Synced	Primary	3
10.244.9.2	mysql/fefced1e	Synced	Primary	3
+-----+-----+-----+-----+				
I don't think bootstrap is necessary				
Checking disk status of mysql/fefced1e-83ce-4f7d-96f1-fb658b69cc81 at 10.244.9.2...				
Checking disk status of mysql/32665833-6a66-45dc-bad1-7a6476787d59 at 10.244.7.2...				
Checking disk status of mysql/026d5b79-6844-4227-9e6a-923dd24de9a6 at 10.244.8.2...				
Checking disk status of mysql/026d5b79-6844-4227-9e6a-923dd24de9a6 at 10.244.8.2... done				
Checking disk status of mysql/32665833-6a66-45dc-bad1-7a6476787d59 at 10.244.7.2... done				
Checking disk status of mysql/fefced1e-83ce-4f7d-96f1-fb658b69cc81 at 10.244.9.2... done				
+-----+-----+-----+-----+				
HOST	NAME/UUID	PERSISTENT DISK USED	EPHEMERAL DISK USED	
+-----+-----+-----+-----+				
10.244.8.2	mysql/026d5b79	28.9% of 9G (0.0% of 0.61M inodes)	28.3% of 157.4G (6.4% of 10.49M inodes)	
10.244.7.2	mysql/32665833	28.7% of 9G (0.0% of 0.61M inodes)	28.3% of 157.4G (6.4% of 10.49M inodes)	
10.244.9.2	mysql/fefced1e	28.9% of 9G (0.0% of 0.61M inodes)	28.3% of 157.4G (6.4% of 10.49M inodes)	
+-----+-----+-----+-----+
```

[View a larger version of this image.](#)

## Example Unhealthy Output

The `mysql-diag` command returns the following message if your canary status is unhealthy:

```
Checking canary status...unhealthy
```

In the event of a broken HA cluster, running `mysql-diag` outputs actionable steps meant to expedite the recovery of that HA cluster. Below is a sample `mysql-diag` output after the tool identified an unhealthy HA cluster:

```

mysql-monitor/5faae86a-a325-44fa-a8ab-24e736e42390:~$ mysql-diag
Wed Jul 19 17:59:43 UTC 2017
Checking canary status...
Checking canary status... healthy
Checking cluster status of mysql/fefced1e-83ce-4f7d-96f1-fb658b69cc81 at 10.244.9.2...
Checking cluster status of mysql/32665833-6a66-45dc-bad1-7a6476787d59 at 10.244.7.2...
Checking cluster status of mysql/026d5b79-6844-4227-9e6a-923dd24de9a6 at 10.244.8.2...
Checking cluster status of mysql/fefced1e-83ce-4f7d-96f1-fb658b69cc81 at 10.244.9.2... dial tcp 10.244.9.2:3306: getsockopt: connection refused
Checking cluster status of mysql/32665833-6a66-45dc-bad1-7a6476787d59 at 10.244.7.2... dial tcp 10.244.7.2:3306: getsockopt: connection refused
Checking cluster status of mysql/026d5b79-6844-4227-9e6a-923dd24de9a6 at 10.244.8.2... dial tcp 10.244.8.2:3306: getsockopt: connection refused
+-----+-----+-----+-----+
| HOST | NAME/UUID | WSREP LOCAL STATE | WSREP CLUSTER STATUS | WSREP CLUSTER SIZE |
+-----+-----+-----+-----+
10.244.9.2	mysql/fefced1e	N/A - ERROR	N/A - ERROR	N/A - ERROR
10.244.7.2	mysql/32665833	N/A - ERROR	N/A - ERROR	N/A - ERROR
10.244.8.2	mysql/026d5b79	N/A - ERROR	N/A - ERROR	N/A - ERROR
+-----+-----+-----+-----+				
Checking disk status of mysql/fefced1e-83ce-4f7d-96f1-fb658b69cc81 at 10.244.9.2...				
Checking disk status of mysql/026d5b79-6844-4227-9e6a-923dd24de9a6 at 10.244.8.2...				
Checking disk status of mysql/32665833-6a66-45dc-bad1-7a6476787d59 at 10.244.7.2...				
Checking disk status of mysql/32665833-6a66-45dc-bad1-7a6476787d59 at 10.244.7.2... done				
Checking disk status of mysql/026d5b79-6844-4227-9e6a-923dd24de9a6 at 10.244.8.2... done				
Checking disk status of mysql/fefced1e-83ce-4f7d-96f1-fb658b69cc81 at 10.244.9.2... done				
+-----+-----+-----+				
HOST	NAME/UUID	PERSISTENT DISK USED	Ephemeral Disk Used	
+-----+-----+-----+				
10.244.7.2	mysql/32665833	28.7% of 9G (0.0% of 0.61M inodes)	28.2% of 157.4G (6.3% of 10.49M inodes)	
10.244.8.2	mysql/026d5b79	28.9% of 9G (0.0% of 0.61M inodes)	28.2% of 157.4G (6.3% of 10.49M inodes)	
10.244.9.2	mysql/fefced1e	28.9% of 9G (0.0% of 0.61M inodes)	28.2% of 157.4G (6.3% of 10.49M inodes)	
+-----+-----+-----+
[CRITICAL] You must bootstrap the cluster. Follow these instructions: http://docs.pivotal.io/p-mysql/bootstrapping.html
[CRITICAL] Run the download-logs command:
$ download-logs -d /tmp/output -n 10.244.7.2 -n 10.244.8.2 -n 10.244.9.2
For full information about how to download and use the download-logs command see https://discuss.pivotal.io/hc/en-us/articles/221504408
[WARNING] NOT RECOMMENDED
Do not perform the following unless instructed by Pivotal Support:
- Do not scale down the cluster to one node then scale back. This puts user data at risk.
- Avoid "bosh recreate" and "bosh cck". These options remove logs on the VMs making it harder to diagnose cluster issues.
```

[View a larger version of this image.](#)

## Recovering From MySQL Cluster Downtime

Page last updated:

This topic describes how to bootstrap your MySQL cluster in the event of a cluster failure.

You can bootstrap your cluster by using one of two methods:

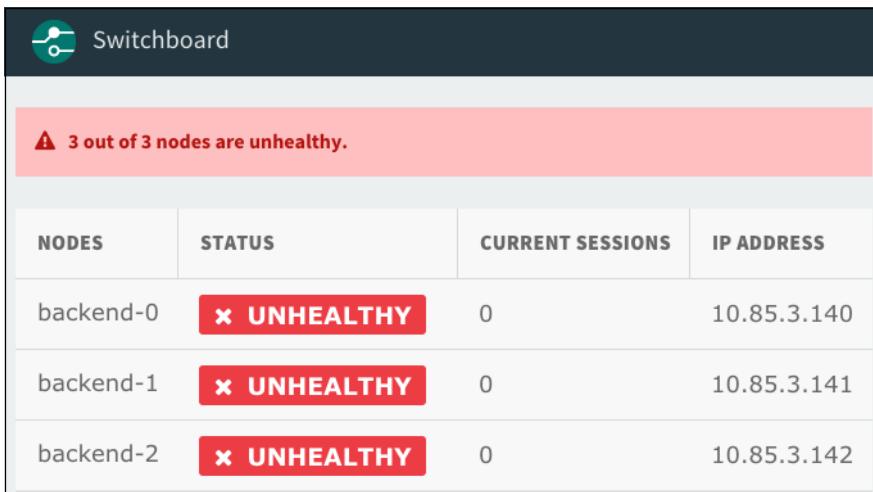
- Run the bootstrap errand. See [Run the Bootstrap Errand](#) below.
- Bootstrap manually. See [Bootstrap Manually](#) below.

### When to Bootstrap

You must bootstrap a cluster that loses quorum. A cluster loses quorum when less than half of the nodes can communicate with each other for longer than the configured grace period. If a cluster does not lose quorum, individual unhealthy nodes automatically rejoin the cluster after resolving the error, restarting the node, or restoring connectivity.

To check whether your cluster has lost quorum, look for the following symptoms:

- All nodes appear “Unhealthy” on the proxy dashboard, viewable at `https://BOSH-JOB-INDEX-proxy-p-mysql-ert.YOUR-SYSTEM-DOMAIN`:



The screenshot shows the Switchboard proxy dashboard interface. At the top, there is a dark header bar with the "Switchboard" logo. Below it, a red banner displays the message "⚠ 3 out of 3 nodes are unhealthy." The main area is a table with four columns: "NODES", "STATUS", "CURRENT SESSIONS", and "IP ADDRESS". There are three rows corresponding to the nodes "backend-0", "backend-1", and "backend-2". Each row has a "STATUS" cell containing a red button with the text "✗ UNHEALTHY". The "IP ADDRESS" column lists the IP addresses 10.85.3.140, 10.85.3.141, and 10.85.3.142 respectively.

| NODES     | STATUS      | CURRENT SESSIONS | IP ADDRESS  |
|-----------|-------------|------------------|-------------|
| backend-0 | ✗ UNHEALTHY | 0                | 10.85.3.140 |
| backend-1 | ✗ UNHEALTHY | 0                | 10.85.3.141 |
| backend-2 | ✗ UNHEALTHY | 0                | 10.85.3.142 |

- All responsive nodes report the value of `wsrep_cluster_status` as `non-Primary`:

```
mysql> SHOW STATUS LIKE 'wsrep_cluster_status';
+-----+-----+
| Variable_name | Value |
+-----+-----+
| wsrep_cluster_status | non-Primary |
+-----+-----+
```

- All unresponsive nodes respond with `ERROR 1047` when using most statement types in the MySQL client:

```
mysql> select * from mysql.user;
ERROR 1047 (08S01) at line 1: WSREP has not yet prepared node for application use
```

For more information about checking the state of your cluster, see [Check Cluster State](#).

### Run the Bootstrap Errand

MySQL for PCF includes a [BOSH errand](#) that automates the manual bootstrapping procedure in the [Bootstrap Manually](#) section below.

It finds the node with the highest transaction sequence number and asks it to start up by itself (i.e. in bootstrap mode) and then asks the remaining nodes to join the cluster.

In most cases, running the errand recovers your cluster. But, certain scenarios require additional steps.

## Determine Type of Cluster Failure

To determine which set of instructions to follow, do the following:

- Run the following command.

```
bosh -e YOUR-ENV -d YOUR-DEPLOYMENT instances
```

Where:

- `YOUR-ENV` is the environment where you deployed the cluster.
- `YOUR-DEPLOYMENT` is the deployment cluster name.

For example:

```
$ bosh -e prod -d mysql instances
```

- Find and record the `Process State` for your MySQL instances. In the following example output, the MySQL instances are in the `failing` process state.

| Instance                                                           | Process State | AZ            | IPs       |
|--------------------------------------------------------------------|---------------|---------------|-----------|
| backup-restore/c635410e-917d-46aa-b054-86d222b6d1c0                | running       | us-central1-b | 10.0.4.47 |
| bootstrap/a31af4ff-e1df-4ff1-a781-abc3c6320ed4                     | -             | us-central1-b | -         |
| broker-registrar/1a93e53d-a7fc-4308-85d4-3b2b80d504e4              | -             | us-central1-b | 10.0.4.58 |
| cf-mysql-broker/137d52b8-a1b0-41f3-847f-c44f51f87728               | running       | us-central1-c | 10.0.4.57 |
| cf-mysql-broker/28b463b1-cc12-42bf-b34b-82ca7c417c41               | running       | us-central1-b | 10.0.4.56 |
| deregister-and-purge-instances/4cb93432-4d90-4f1d-8152-d0c238fa5ab | -             | us-central1-b | -         |
| monitoring/f7117db-1c22-495e-a99e-cf2add90dea9                     | running       | us-central1-b | 10.0.4.48 |
| mysql/220fe72a-9026-4e2e-9fc3-1f5c0b6b09b                          | failing       | us-central1-b | 10.0.4.44 |
| mysql/28a210ac-cb98-4ab4-9672-9f4c661c57b8                         | failing       | us-central1-f | 10.0.4.46 |
| mysql/c1639373-26a2-44ce-85db-c9fe5a42964b                         | failing       | us-central1-c | 10.0.4.45 |
| proxy/87c5683d-12f5-426c-b925-62521529f64a                         | running       | us-central1-b | 10.0.4.60 |
| proxy/b0115cccd-7973-42d3-b6de-edb5ae53c63e                        | running       | us-central1-c | 10.0.4.61 |
| rejoin-unsafe/8cc9370a-e86b-4638-bf76-e103f858413f                 | -             | us-central1-b | -         |
| smoke-tests/e026aaef-efd9-4644-8d14-0811cb1ba733                   | -             | us-central1-b | 10.0.4.59 |

- Choose your scenario:

- If your MySQL instances are in the `failing` state, continue to [Scenario 1](#).
- If your MySQL instances are in the `-` state, continue to [Scenario 2](#).

## Scenario 1: VMs Running, Cluster Disrupted

In this scenario, the VMs are running, but the cluster has been disrupted.

To bootstrap in this scenario, follow these steps:

- To run the bootstrap errand, run the following command:

```
bosh -e YOUR-ENV -d YOUR-DEPLOYMENT run-errand bootstrap
```

Where:

- `YOUR-ENV` is the name of your environment.
- `YOUR-DEPLOYMENT` is the name of your deployment.

**Note:** The errand runs for a long time, during which no output is returned.

The command returns many lines of output, eventually followed by:

```
Bootstrap errand completed
[stderr]
+ echo 'Started bootstrap errand ...'
+ JOB_DIR=/var/vcap/jobs/bootstrap
+ CONFIG_PATH=/var/vcap/jobs/bootstrap/config/config.yml
+ /var/vcap/packages/bootstrap/bin/cf-mysql-bootstrap -configPath=/var/vcap/jobs/bootstrap/config/config.yml
+ echo 'Bootstrap errand completed'
+ exit 0
Errand 'bootstrap' completed successfully (exit code 0)
```

2. If the errand fails, run the bootstrap errand command again after a few minutes. The bootstrap errand might not work the first time.
3. If the errand fails after several tries, bootstrap your cluster manually. See [Bootstrap Manually](#) below.

## Scenario 2: VMs Terminated or Lost

In severe circumstances, such as a power failure, it is possible to lose all your VMs. You must recreate them before you can begin recovering the cluster.

When MySQL instances are in the `-` state, the VMs are lost. The procedures in this scenario bring the instances from a `-` state to a `failing` state. Then you run the bootstrap errand similar to [Scenario 1](#) above and restore configuration.

To recover terminated or lost VMs, do the procedures in the sections below:

1. [Recreate the Missing VMs](#): Bring MySQL instances from a `-` state to a `failing` state.
2. [Run the Bootstrap Errand](#): Since your instances are now in the `failing` state, you continue similarly to [Scenario 1](#) above.
3. [Restore the BOSH Configuration](#): Go back to unignoring all instances and redeploy. This is a critical and mandatory step.

**⚠ warning:** If you do not set each of your ignored instances to `unignore`, your instances are not updated in future deploys. You must perform the procedure in the final section of *Scenario 2*, [Restore the BOSH Configuration](#).

### Recreate the Missing VMs

The procedure in this section uses BOSH to recreate the VMs, install software on them, and try to start the jobs.

The procedure below allows you to do the following:

- Redeploy your cluster while expecting the jobs to fail.
- Instruct BOSH to ignore the state of each instance in your cluster. This allows BOSH to deploy the software to each instance even if the instance is failing.

To recreate your missing VMs, do the following:

1. To SSH into the BOSH VMs, follow the procedure in [BOSH SSH](#).
2. If BOSH resurrection is enabled, disable it by running the following command:

```
bosh -e YOUR-ENV update-resurrection off
```

Where `YOUR-ENV` is the name of your environment.

3. To download the current manifest, run the following command:

```
bosh -e YOUR-ENV -d YOUR-DEPLOYMENT manifest > /tmp/manifest.yml
```

Where:

- `YOUR-ENV` is the name of your environment.
- `YOUR-DEPLOYMENT` is the name of your deployment.

4. To redeploy deployment, run the following command:

```
bosh -e YOUR-ENV -d YOUR-DEPLOYMENT deploy /tmp/manifest.yml
```

Where:

- `YOUR-ENV` is the name of your environment.
- `YOUR-DEPLOYMENT` is the name of your deployment.

**💡 Note:** Expect one of the MySQL VMs to fail. Deploying causes BOSH to create new VMs and install the software. Forming a cluster is in a subsequent step.

- To view the instance GUID of the VM that attempted to start, run the following command:

```
bosh -e YOUR-ENV -d YOUR-DEPLOYMENT instances
```

Where:

- YOUR-ENV is the name of your environment.
- YOUR-DEPLOYMENT is the name of your deployment.

Record the instance GUID, which is the string after mysql in your BOSH instances output.

- To instruct BOSH to ignore each MySQL VM, run the following command:

```
bosh -e YOUR-ENV -d YOUR-DEPLOYMENT ignore mysql/INSTANCE-GUID
```

Where:

- YOUR-ENV is the name of your environment.
- YOUR-DEPLOYMENT is the name of your deployment.
- INSTANCE-GUID is the GUID of your instance you recorded in the above step.

- Repeat steps 4 through 6 until all instances have attempted to start.

- If you disabled BOSH resurrection in step 2, to re-enable it, run the following command:

```
bosh -e YOUR-ENV update-resurrection on
```

Where YOUR-ENV is the name of your environment.

- To confirm that your MySQL instances have gone from the - state to the failing state, run the following command

```
bosh -e YOUR-ENV -d YOUR-DEPLOYMENT instances
```

Where:

- YOUR-ENV is the name of your environment.
- YOUR-DEPLOYMENT is the name of your deployment.

## Run the Bootstrap Errand

After you recreate the VMs, all instances now have a failing process state and have the MySQL code. You must run the bootstrap errand to recover the cluster.

To bootstrap, do the following:

- To run the bootstrap errand, run the following command:

```
bosh -e YOUR-ENV -d YOUR-DEPLOYMENT run-errand bootstrap
```

Where:

- YOUR-ENV is the name of your environment.
- YOUR-DEPLOYMENT is the name of your deployment.

 **Note:** The errand runs for a long time, during which no output is returned.

The command returns many lines of output, eventually with the following successful output:

```
Bootstrap errand completed
[stderr]
echo 'Started bootstrap errand ...'
JOB_DIR=/var/vcap/jobs/bootstrap
CONFIG_PATH=/var/vcap/jobs/bootstrap/config/config.yml
/var/vcap/packages/bootstrap/bin/cf-mysql-bootstrap -configPath=/var/vcap/jobs/bootstrap/config/config.yml
echo 'Bootstrap errand completed'
exit 0
Errand 'bootstrap' completed successfully (exit code 0)
```

2. If the errand fails, run the bootstrap errand command again after a few minutes. The bootstrap errand might not work immediately.
3. See that the errand completes successfully in the shell output and continue to [Restore the BOSH Configuration](#) below.

 **Note:** After you complete the bootstrap errand, you might still see instances in the `failing` state. Continue to the next section anyway.

## Restore the BOSH Configuration

 **warning:** If you do not set each of your ignored instances to `unignore`, your instances are never updated in future deploys.

To restore your BOSH configuration to its previous state, this procedure unignores each instance that was previously ignored:

1. For each ignored instance, run the following command:

```
bosh -e YOUR-ENV -d YOUR-DEPLOYMENT unignore mysql/INSTANCE-GUID
```

Where:

- o `YOUR-ENV` is the name of your environment.
- o `YOUR-DEPLOYMENT` is the name of your deployment.
- o `INSTANCE-GUID` is the GUID of your instance.

2. To redeploy your deployment, run the following command:

```
bosh -e YOUR-ENV -d YOUR-DEPLOYMENT deploy
```

Where:

- o `YOUR-ENV` is the name of your environment.
- o `YOUR-DEPLOYMENT` is the name of your deployment.

3. To validate that all `mysql` instances are in a `running` state, run the following command:

```
bosh -e YOUR-ENV -d YOUR-DEPLOYMENT instances
```

Where:

- o `YOUR-ENV` is the name of your environment.
- o `YOUR-DEPLOYMENT` is the name of your deployment.

## Bootstrap Manually

If the bootstrap errand is not able to automatically recover the cluster, you might need to do the steps manually.

Do the procedures in the sections below to manually bootstrap your cluster.

 **warning:** The following procedures are prone to user-error and can result in lost data if followed incorrectly. Follow the procedure in [Bootstrap with the BOSH Errand](#) above first, and only resort to the manual process if the errand fails to repair the cluster.

## Shut Down MySQL

Do the following for each node in the cluster:

1. To SSH into the node, follow the procedure in [BOSH SSH ↗](#).
2. To shut down the `mysqld` process on the node, run the following command:

```
monit stop galera-init
```

Re-bootstrapping the cluster is not successful unless you shut down the `mysqld` process on all nodes in the cluster.

## Choose Node to Bootstrap

To avoid losing data, you must bootstrap from a node in the cluster that has the highest transaction sequence number (`seqno`).

For each node in the cluster, do the following to find its `seqno`:

1. SSH into the node, using the procedure in [BOSH SSH ↗](#).
2. Run the following command to find `seqno` values in the node's Galera state file, `grastate.dat`:
 

```
cat /var/vcap/store/pxc-mysql/grastate.dat | grep 'seqno:'
```
3. Do one of the following, based on the last and highest `seqno` value in the `grep` output above:
  - o If `seqno` is a positive integer, then the node shut down gracefully. Record this number for comparison with the latest `seqno` of other nodes in the cluster.
  - o If `seqno` is `-1`, then node crashed or was killed. Proceed as follows to possibly recover the `seqno` from the database:
    1. Run the following command to temporarily start the database and append the last sequence number to its error log:
 

```
/var/vcap/packages/pxc/bin/mysqld --defaults-file=/var/vcap/jobs/pxc-mysql/config/my.cnf --wsrep-recover
```
    2. Scan the end of the error log `mysql.err.log` for the recovered `seqno`. It appears after the group ID (a UUID) and a colon. For example, it is `15026` in the following output:
 

```
$ grep "Recovered position" /var/vcap/sys/log/pxc-mysql/mysql.err.log | tail -1
150225 18:09:42 mysqld_safe WSREP: Recovered position e93955c7-b797-11e4-9faa-9a6f0b73cb46:15026
```
    3. If the node never connected to the cluster before crashing, it may never have been assigned a group ID. In this case, there is nothing to recover. Do not choose this node for bootstrapping unless all the other nodes also crashed.
4. After you determine the `seqno` for all nodes in your cluster, identify the one with the highest `seqno`. If multiple nodes share the same highest `seqno`, and it is not `-1`, you can bootstrap from any of them.

## Bootstrap the First Node

After determining the node with the highest `sequence_number`, do the following to bootstrap the node:

**Note:** Only run these bootstrap commands on the node with the highest `sequence_number`. Otherwise the node with the highest `sequence_number` is unable to join the new cluster unless its data is abandoned. Its `mysqld` process exits with an error. For more information about intentionally abandoning data, see [Architecture ↗](#).

1. To update the state file and restart the `mysqld` process on the new bootstrap node, run the following commands:
 

```
echo -n "NEEDS_BOOTSTRAP" > /var/vcap/store/pxc-mysql/state.txt
monit start galera-init
```
2. It can take up to ten minutes for `monit` to start the `mysqld` process. To check if the `mysqld` process has started successfully, run the following command:
 

```
watch monit summary
```

## Restart Remaining Nodes

After the bootstrapped node is running, to restart the nodes do the following:

1. To start the `mysqld` process with `monit`, run the following command:

```
monit start galera-init
```

If the node is prevented from starting by the [Interruptor](#), do the manual procedure to force the node to rejoin the cluster, documented in [Pivotal Knowledge Base](#).

**⚠ warning:** Forcing a node to rejoin the cluster is a destructive procedure. Only do the procedure with the assistance of [Pivotal Support](#).

2. If the `monit start` command fails, it might be because the node with the highest `sequence_number` is `mysql/0`.

In this case, do the following:

- a. To make BOSH ignore updating `mysql/0`, run the following command:

```
bosh -e YOUR-ENV -d YOUR-DEPLOYMENT ignore mysql/0
```

Where:

- `YOUR-ENV` is the name of your environment.
- `YOUR-DEPLOYMENT` is the name of your deployment.

- b. Navigate to Ops Manager in a browser, log in, and click **Apply Changes**.
- c. When the deploy finishes, run the following command from the Ops Manager VM:

```
bosh -e YOUR-ENV -d YOUR-DEPLOYMENT unignore mysql/0
```

Where:

- `YOUR-ENV` is the name of your environment.
- `YOUR-DEPLOYMENT` is the name of your deployment.

3. To verify that the new nodes have successfully joined the cluster, SSH into the bootstrap node and run the following command to output the total number of nodes in the cluster:

```
mysql> SHOW STATUS LIKE 'wsrep_cluster_size';
```

## Configuring Authentication and Enterprise SSO for PAS

This topic describes [Pivotal Cloud Foundry](#) (PCF) Pivotal Application Service (PAS) authentication and single sign-on configuration with Lightweight Directory Access Protocol (LDAP) and Security Assertion Markup Language (SAML).

Refer to the instructions below to configure your deployment with SAML or LDAP.

Connecting Pivotal Application Service (PAS) to either the LDAP or SAML external user store allows the User Account and Authentication (UAA) server to delegate authentication to existing enterprise user stores.

If your enterprise user store is exposed as a SAML or LDAP Identity Provider for single sign-on (SSO), you can configure SSO to allow users to access the Apps Manager and Cloud Foundry Command Line Interface (cf CLI) without creating a new account or, if using SAML, without re-entering credentials.

See the [Adding Existing SAML or LDAP Users to a PCF Deployment](#) topic for information about managing user identity and pre-provisioning user roles with SAML or LDAP in PCF.

For an explanation of the process used by the UAA Server when it attempts to authenticate a user through LDAP, see [Configuring LDAP Integration with Pivotal Cloud Foundry](#) in the Pivotal Knowledge Base.

**Note:** When integrating with an external identity provider, such as LDAP, authentication within the UAA becomes chained. An authentication attempt with a user's credentials is first attempted against the UAA user store before the external provider, LDAP. For more information, see [Chained Authentication](#) in the *User Account and Authentication LDAP Integration GitHub documentation*.

## Configure PAS to Use a SAML Identity Provider

In SAML terminology, the SAML protocol communicates user data between an identity provider (IdP) and a service provider (SP).

To connect PAS with SAML, do the following:

- [Configure PAS as a Service Provider for SAML](#)
- [Configure SAML as an Identity Provider for PAS](#)

### Configure PAS as a Service Provider for SAML

To configure PAS to use a SAML IdP, do the following:

1. From the **Installation Dashboard**, click the PAS tile.
2. Select the **Domains** tab and record your system domain.

The screenshot shows the Pivotal Elastic Runtime Installation Dashboard. The left sidebar has tabs for Settings, Status, Credentials, and Logs. The Domains tab is selected. On the right, there is a note: "Elastic Runtime hosts applications at subdomains under its apps domain and assigns system components to subdomains under its system domain. You need to configure a wildcard DNS for both the apps domain and system domain. The two domains can be the same, although this is not recommended." Below this note are fields for "System Domain" (containing "sy") and "Apps Domain" (containing "ap"). A note next to the Apps Domain field says: "Default parent domain that pushed apps use for their hostnames. This domain also requires a wildcard DNS record. Use the Cloud Foundry command line interface (cf CLI) to add or delete subdomains assigned to individual apps." At the bottom is a "Save" button.

3. Select **Authentication and Enterprise SSO**.

**SAML Identity Provider**

**Provider Name \***  
 Name of Identity Provider. Allowed characters are alphanumeric, plus `\_, and `-`.

**Display Name \***

**Provider Metadata** (if you would rather provide an SSO endpoint URL, skip to the next field)

**(OR) Provider Metadata URL**

**Name ID Format\***  
 Email Address

**Email Domain(s)**

**First Name Attribute**

**Last Name Attribute**

**Email Attribute**

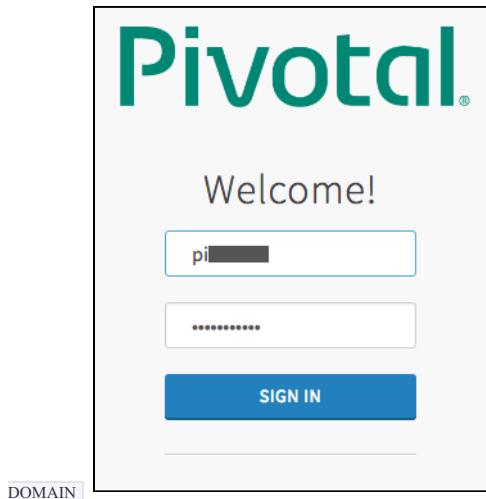
**External Groups Attribute**

Sign Authentication Requests

Require Signed Assertions

4. Select **SAML Identity Provider**.

5. Set the **Provider Name**. This is a unique name you create for the Identity Provider. This name can include only alphanumeric characters, **[+]**, **[-]**, and **[.]**. You should not change this name after deployment because all external users use it to link to the provider.
6. Enter a **Display Name**. Your provider display name appears as a link on your Pivotal login page, which you can access at <https://login.YOUR-SYSTEM>



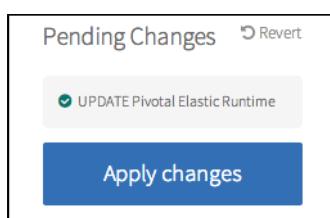
7. Retrieve the metadata from your Identity Provider and copy it into either the **Provider Metadata** or the **Provider Metadata URL** fields, depending on whether your Identity Provider exposes a Metadata URL or not. Refer to the [Configure SAML as an Identity Provider for PAS](#) section of this topic for more information. Pivotal recommends that you use the Provider Metadata URL rather than Provider Metadata because the metadata can change. You can do this in either of the following ways:

- If your Identity Provider exposes a Metadata URL, provide the Metadata URL.
- Download your Identity Provider metadata and paste this XML into the **Provider Metadata** area.

**Note:** You only need to select one of the above configurations. If you configure both, your Identity Provider defaults to the (OR) **Provider Metadata URL**.

**Note:** See [Adding Existing SAML or LDAP Users to a PCF Deployment](#) for information about on-boarding SAML users and mapping them to PAS user roles.

8. Select the **Name ID Format** for your SAML Identity Provider. This translates to `username` on PAS. The default is `Email Address`.
9. For **Email Domain(s)**, enter a comma-separated list of the email domains for external users who will receive invitations to Apps Manager.
10. For **First Name Attribute** and **Last Name Attribute**, enter the attribute names in your SAML database that correspond to the first and last names in each user record, for example `first_name` and `last_name`. This field is case sensitive.
11. For **Email Attribute**, enter the attribute name in your SAML assertion that corresponds to the email address in each user record, for example `EmailID`. This field is case sensitive.
12. For **External Groups Attribute**, enter the attribute name in your SAML database that defines the groups that a user belongs to, for example `group_memberships`. To map the groups from the SAML assertion to admin roles in PAS, follow the instructions in the [Grant Admin Permissions to an External Group \(SAML or LDAP\)](#) section of the *Creating and Managing Users with the UAA CLI (UAAC)* topic. This field is case sensitive.
13. By default, all SAML Authentication Request from PAS are signed. To change this, disable the **Sign Authentication Requests** checkbox and configure your Identity Provider to verify SAML authentication requests.
14. To validate the signature for the incoming SAML assertions, enable the **Required Signed Assertions** checkbox and configure your Identity Provider to send signed SAML assertions.
15. Click **Save**.
16. Return to the **Installation Dashboard** by clicking the link.
17. On the Installation Dashboard, click **Review Pending Changes**, then **Apply Changes**.



## Configure SAML as an Identity Provider for PAS

The following procedure configures a SAML IdP to designate PAS as an SP.

Download the Service Provider Metadata from <https://login.YOUR-SYSTEM-DOMAIN/saml/metadata>. Consult the documentation from your Identity Provider for configuration instructions.

Refer to the table below for information about certain industry-standard Identity Providers and how to integrate them with PAS:

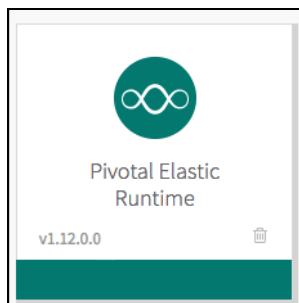
| Solution Name                                          | Integration Guide    |
|--------------------------------------------------------|----------------------|
| <a href="#">CA Single Sign-On aka CA SiteMinder ↗</a>  | <a href="#">Link</a> |
| <a href="#">Ping Federate ↗</a>                        | <a href="#">Link</a> |
| <a href="#">Active Directory Federation Services ↗</a> | <a href="#">Link</a> |

 **Note:** Some Identity Providers allow uploads of Service Provider Metadata. Other providers require you to manually enter the Service Provider Metadata into a form. If your Identity Provider requires manual entry but is not listed above, see the [CA SiteMinder SSO Integration Guide](#).

## Configure LDAP as an Identity Provider for PAS

To integrate the UAA with one or more LDAP servers, configure PAS with your LDAP endpoint information as follows:

1. Log into the Operations Manager web interface.
2. On the Product Dashboard, select the PAS tile.



3. In the left navigation menu, select **Authentication and Enterprise SSO**.

LDAP Server

Server URL \*

LDAP Credentials \*

Username

Password

User Search Base \*

User Search Filter \*

Group Search Base

Group Search Filter \*

Server SSL Cert

Server SSL Cert AltName

First Name Attribute

4. Under **Configure your UAA**, select **LDAP Server**.

- For **Server URL**, enter the URL that point your LDAP server. With multiple LDAP servers, separate their URLs with spaces. Each URL must include one of the following protocols:
  - `ldap://` : This specifies that the LDAP server uses an unencrypted connection.
  - `ldaps://` : This specifies that the LDAP server uses SSL for an encrypted connection and requires that the LDAP server holds a trusted certificate or that you import a trusted certificate to the JVM truststore.
- For **LDAP Credentials**, enter the LDAP Distinguished Name (DN) and password for binding to the LDAP Server. Example DN:  
`cn=administrator,ou=Users,dc=example,dc=com`

**Note:** Pivotal recommends that you provide LDAP credentials that grant read-only permissions on the LDAP Search Base and the LDAP Group Search Base. In addition to this, if the bind user belongs to a different search base, you must use the full DN.

**warning:** Pivotal recommends against reusing LDAP service accounts across environments. LDAP service accounts should not be subject to manual lockouts, such as lockouts that result from users utilizing the same account. Also, LDAP service accounts should not be subject to automated deletions, since disruption to these service accounts could prevent user logins.

- For **User Search Base**, enter the location in the LDAP directory tree from which any LDAP User search begins. The typical LDAP Search Base matches your domain name.

For example, a domain named “cloud.example.com” typically uses the following LDAP User Search Base: `ou=Users,dc=example,dc=com`

- For **User Search Filter**, enter a string that defines LDAP User search criteria. These search criteria allow LDAP to perform more effective and efficient searches. For example, the standard LDAP search filter `cn=Smith` returns all objects with a common name equal to `Smith`.

In the LDAP search filter string that you use to configure PAS, use `{0}` instead of the username. For example, use `cn={0}` to return all LDAP objects with the same common name as the username.

In addition to `cn`, other attributes commonly searched for and returned are `mail`, `uid` and, in the case of Active Directory, `sAMAccountName`.

**Note:** For instructions for testing and troubleshooting your LDAP search filters, see the [Configuring LDAP Integration with Pivotal Cloud Foundry](#) Knowledge Base article.

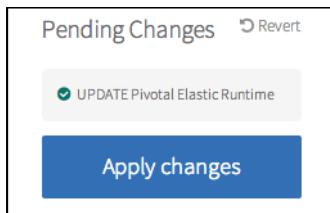
- For **Group Search Base**, enter the location in the LDAP directory tree from which the LDAP Group search begins.

For example, a domain named “cloud.example.com” typically uses the following LDAP Group Search Base: `ou=Groups,dc=example,dc=com`

Follow the instructions in the [Grant Admin Permissions to an External Group \(SAML or LDAP\)](#) section of *Creating and Managing Users with the UAA CLI (UAAC)* to map the groups under this search base to admin roles in PAS.

**Note:** See [Adding Existing SAML or LDAP Users to a PCF Deployment](#) to on-board individual LDAP users and map them to PAS Roles.

- For **Group Search Filter**, enter a string that defines LDAP Group search criteria. The standard value is `member={0}`.
  - For **Server SSL Cert**, paste in the root certificate from your CA certificate or your self-signed certificate.
  - For **First Name Attribute** and **Last Name Attribute**, enter the attribute names in your LDAP directory that correspond to the first and last names in each user record, for example `cn` and `sn`.
  - For **Email Attribute**, enter the attribute name in your LDAP directory that corresponds to the email address in each user record, for example `mail`.
  - For **Email Domain(s)**, enter a comma-separated list of the email domains for external users who will receive invitations to Apps Manager.
    - For **LDAP Referrals**, select how the UAA handles LDAP server referrals out to other external user stores. The UAA can:
      - Automatically follow any referrals.
      - Ignore referrals and return partial result.
      - Throw exception for each referral and abort.
- Click **Save**.
  - Return to the **Installation Dashboard** by clicking the link.
  - On the Installation Dashboard, click **Review Pending Changes**, then **Apply Changes**.



## Configuring ADFS as an Identity Provider

This topic describes the process of configuring Active Directory Federation Services (ADFS) as your identity provider (IdP) in Pivotal Cloud Foundry (PCF) and ADFS.

### Configure SAML Integration in PCF

You can use ADFS as your SAML IdP for Ops Manager and Pivotal Application Service (PAS):

- If you want to use ADFS as your SAML IdP for Ops Manager, do the following:
  - [Configure SAML Integration in Ops Manager](#)
  - [Configure SAML Integration in ADFS](#)
- If you want to use ADFS as your SAML IdP for PAS, do the following:
  - [Configure SAML Integration in PAS](#)
  - [Configure SAML Integration in ADFS](#)

#### Configure SAML Integration in Ops Manager

To configure Ops Manager to use ADFS as your SAML IdP, do the following:

1. Download your IdP metadata from <https://YOUR-ADFS-HOSTNAME/federationmetadata/2007-06/federationmetadata.xml>.
2. Perform the steps in the *Use an Identity Provider* section of the BOSH Director configuration topic for you IaaS:
  - [Configuring BOSH Director on AWS](#)
  - [Configuring BOSH Director on Azure Manually](#)
  - [Configuring BOSH Director on GCP](#)
  - [Configuring BOSH Director on OpenStack](#)
  - [Configuring BOSH Director on vSphere](#)

 **Note:** You can set up SAML access for Ops Manager during the initial PCF installation or later by navigating to **Settings** in the user menu on the Ops Manager Installation Dashboard, configuring the **Authentication Method** pane, and then clicking **Review Pending Changes** and **Apply Changes**.

#### Configure SAML Integration in PAS

To configure PAS to use ADFS as your SAML IdP, do the following:

1. Download your IdP metadata from <https://YOUR-ADFS-HOSTNAME/federationmetadata/2007-06/federationmetadata.xml>.
2. Perform the steps in the [Configure PCF as a Service Provider for SAML](#) section of the *Configuring Authentication and Enterprise SSO for PAS* topic.

### Configure SAML Integration in ADFS

To designate PCF as your SAML service provider (SP) in ADFS, do the following:

1. Download your SP metadata from <https://login.YOUR-SYSTEM-DOMAIN/saml/metadata>.
2. Open your ADFS Management console and add a relying party trust as follows:
  - a. Click **Add Relying Party Trust...** in the **Actions** pane.
  - b. On the Welcome step, click **Start**.
  - c. Select **Import data about the relying party from a file**, import the downloaded SP metadata file, and click **Next**.
  - d. Enter a **Display name** for the new relying party trust and click **Next**.
  - e. Leave the default multi-factor authentication selection and click **Next**.
  - f. Select **Permit all users to access this relying party** and click **Next**.

- g. Review your settings and click **Next**.
- h. Click **Close** to finish the wizard.

3. Modify your relying party trust as follows:

- a. Double-click the new relying party trust.
- b. Select the **Encryption** tab and click **Remove** to remove the encryption certificate you imported.
- c. In the **Advanced** tab, select **SHA256** for the **Secure hash algorithm**.

4. (Optional) If you are using a self-signed certificate and want to disable CRL checks, do the following:

- a. Open **Windows Powershell** as an Administrator.
- b. Execute the following command:

```
set-ADFSRelyingPartyTrust -TargetName "RELYING-PARTY-TRUST" -SigningCertificateRevocationCheck None
```

5. To add claim rules for your relying party trust, select your relying party trust and click **Edit Claim Rules....**

6. In the **Issuance Transform Rules** tab, create two claim rules as follows:

- a. Click **Add Rule**.
- b. Select **Send LDAP Attributes as Claims** for **Claim rule template** and click **Next**.
- c. Enter a **Claim rule name**.
- d. Select **Active Directory** for **Attribute store**.
- e. Select **E-Mail-Addresses** for **LDAP Attribute** and **E-Mail Address** for **Outgoing Claim Type**. Alternatively, if you do not have the email attribute configured for users, you can select **User-Principal-Name** under **LDAP Attribute**.
- f. Click **Finish**.
  
- g. Click **Add Rule**.
- h. Select **Transform an Incoming Claim** for **Claim rule template** and click **Next**.
- i. Enter a **Claim rule name**.
- j. Select **E-Mail Address** for **Incoming claim type**.
- k. Select **Name ID** for **Outgoing claim type**.
- l. Select **Email** for **Outgoing name ID format**.
- m. Click **Finish**.

7. To permit access to users based on a security group, navigate to the **Issuance Authorization Rules** tab and create an authorization claim rule as follows:

- a. Click **Add Rule**.
- b. Select **Permit or Deny Users Based on an Incoming Claim** for **Claim rule template** and click **Next**.
- c. Enter a **Claim rule name**.
- d. Select **Group SID** for **Incoming claim type**.
- e. Click **Browse** and locate the security group in your domain that PCF developers are a part of and click **OK**.
- f. Ensure **Permit access to users with this incoming claim** is selected.
- g. Click **Finish**.

## Configuring CA as an Identity Provider

This topic explains how to configure single sign-on (SSO) between CA and Pivotal Cloud Foundry (PCF).

Partnership creation between CA and PCF involves the following steps:

1. Installing and configuring the prerequisites
2. Configuring CA Single Sign-On as an Identity Provider
3. Configuring the Service Provider

### Prerequisites

- CA Single Sign-On v12.52 installation
- User store and session store configuration
- Creation of Signed Certificate by a Certificate Authority
- Protect Identity Provider URL with CA SSO by creating the following objects:
  - Authentication scheme
  - Domain
  - Realm
  - Rules and policy
- PCF environment at <https://console.YOUR-SYSTEM-DOMAIN>

 **Note:** Replace YOUR-SYSTEM-DOMAIN with the name of your PCF installation.

## Configuring CA as the SAML 2.0 Identity Provider on PCF

1. Download the service provider metadata.
  - a. Navigate <https://login.YOUR-SYSTEM-DOMAIN/saml/metadata> and log in to CA SSO.
  - b. Navigate to **Federation**.
  - c. Select **Partnership Federation**
  - d. In the Actions menu, select **Export Metadata**.
  - e. Save the exported metadata in an XML file.
2. Follow the steps in [Configuring Authentication and Enterprise SSO for Elastic Runtime](#) to set the identity provider metadata on PCF.
3. Paste the contents of the XML file into the **Identity Provider Metadata** field.
4. Click **Save**.
5. Click **Review Pending Changes**, then **Apply Changes**.

## Configuring PCF as the SAML 2.0 Service Provider on CA Single Sign-On

### Configure Identity Provider and Service Provider Entities

1. Navigate to <https://login.YOUR-SYSTEM-DOMAIN/> and log in to CA SSO.
2. Navigate to **Federation**.
3. Click **Partnership Federation**.
4. Click **Entity**.

5. Click **Create Entity**.

6. To create a local entity, use the values below:

- **Entity Location:** Local
- **Entity Type:** SAML2 IDP
- **Entity ID:** Enter an ID for your local identity provider. For example, <https://ca-technologies.xxx.com>.
- **Entity Name:** Create a name for your local identity provider.
- **Base URL:** Enter the fully-qualified domain name for the host service CA SSO Federation Web Services.
- **Signing Private Key Alias:** Select the private key alias or import a private key.
- **Signed Authentication Requests Required:** Select **No**.
- **Supported NameID format:** Enter `urn:oasis:names:tc:SAML:1.1:nameid-format:emailAddress` and `urn:oasis:names:tc:SAML:1.1:nameid-format:unspecified` to select both email address and unspecified as supported NameID formats.

7. To create a remote entity, click **Import Metadata Button** and do the following:

- a. Download the service provider metadata from <https://login.{systemdomain}/saml/metadata> and save to an XML file.
- b. Browse and select the saved XML Metadata you downloaded in the previous step.
- c. Provide a name for the Remote Service Provider Entity.
- d. Provide an alias for the Signing Certificate imported from the metadata.

 **Note:** PCF signs the outgoing SAML authentication requests.

e. Click **Save**.

## Configure Partnership Between CA SSO and PCF

1. Navigate to <https://login.YOUR-SYSTEM-DOMAIN/> and log in to CA SSO.

2. Navigate to **Federation**.

3. Click **Partnership Federation**.

4. Click **Create Partnership**.

5. To configure the partnership, use the values below to fill in the fields:

- **Add Partnership Name:** Enter a name for your partnership.
- (optional) **Description:** Enter a relevant description for your partnership.
- **Local IPD ID:** Enter the Local Service Provider ID you created in the [Configure Identity Provider and Service Provider Entities](#) section.
- **Remote SP ID:** Enter the Remote SP ID you created in the [Configure Identity Provider and Service Provider Entities](#) section.
- **Base URL:** This field will be pre-populated.
- **Skew Time:** Enter any skew time required by your environment.
- **User Directories and Search Order:** Select the required directories in the required search order.

6. Click **Next**.

7. On the Federation Users page, accept the default values.

8. Click **Next**.

9. To complete the Name ID Format section:

- a. Select **Email Address** from the Name ID Format dropdown.
- b. Select **User Attribute** from the Name ID Type dropdown.

 **Note:** PCF does not support processing SAML Assertion Attributes at this time. You can skip filling out the Assertion Attributes fields.

10. Click **Next**.

11. To complete the SSO and SLO section:

- a. Enter the Authentication URL that is protected by CA SSO under prerequisites.
- b. For SSO Binding, click **HTTP-POST**.
- c. In the Audience field, enter <http://login.YOUR-SYSTEM-DOMAIN>.

 **Note:** The Audience field requires <http://> instead of <https://>. This is only a naming convention within the schema and does not determine connection security.

- d. Select **Both IDP and DP Initiated** from the Transactions Allowed dropdown.
- e. The Assertion Consumer Service URL field will be pre-populated using information from the service provider entity.

12. Click **Next**.

13. To complete the Configure Signature and Encryption section:

- a. In the **Signing Private Key Alias** dropdown, verify that the correct Private Key Alias is selected.
- b. Verify that the correct Verification Certificate Alias is selected in the **Verification Certificate Analysis** dropdown. This alias should be the same certificate created when you import the [Remote Service Provider Entity ID](#).
- c. Select **Sign Both** from the Post Signature Options dropdown.

 **Note:** PCF does not support encryption options at this time.

- d. Click **Finish**.

14. To activate the partnership, expand the Action dropdown for your partnership and click **Activate**.

## Configuring PingFederate as an Identity Provider

This topic explains how to configure single sign-on (SSO) between PingFederate and Pivotal Application Service (PAS).

### Configuring PingFederate as the SAML 2.0 Identity Provider on PAS

1. Download your Identity Provider Metadata from PingFederate Server. Click **Metadata Export** under **Administrative Functions** on the Main Menu of the PingFederate Administrative Console. If your PingFederate server is configured to act as both an Identity Provider (IdP) and a service provider (SP), indicate which type of configuration you want to export and click **Next**. The Signing key can be exported. You can skip the options related to Encryption Keys and Metadata Attribute Contract because they are not supported at this time.
2. Follow the steps in [Configuring Authentication and Enterprise SSO for PAS](#) to set your IdP metadata on PAS.

### Configuring PAS as the SAML 2.0 Service Provider on PingFederate

1. Download your Service Provider Metadata from <https://login.YOUR-SYSTEM-DOMAIN/saml/metadata>.
2. Import the Service Provider Metadata to PingFederate. Navigate to **Main Menu** → **IdP Configuration** → **SP Connection** and click **Import**. In the **Import Connection** screen, browse and select the **.xml** file downloaded in the previous step. Click **Import** and **Done**.
3. PAS expects the NameID format to be an email address (for example, `urn:oasis:names:tc:SAML:1.1:nameid-format:emailAddress`) and the value to be the email address of the currently logged in user. The SSO does not function without this setting.
  - a. Click the connection name on **Main Menu**. To see a full list of connections, click **Manage All SP**.
  - b. Click **Browser SSO** under the **SP Connection** tab.
  - c. Click **Configure Browser SSO**.
  - d. Click **Assertion Creation** under the **Browser SSO** tab.
  - e. Click **Configure Assertion Creation**.
  - f. Click **Identity Mapping** on the **Summary** screen.
  - g. Select **Standard** as the option and map the NameID format to be an email address and the value to be the email address of the user.
4. Select the Authentication Source.
  - a. Click **Browser SSO** under the **SP Connection** tab.
  - b. Click **Configure Browser SSO**.
  - c. Click **Assertion Creation** under the **Browser SSO** tab.
  - d. Click **Configure Assertion Creation**.
  - e. Click **IdP Adapter Mapping** on the **Summary** screen.
  - f. Click **Adapter Instance Name**.
  - g. Click **Adapter Instance** on the **Summary** screen.
5. Enable the SSO Browser Profiles.
  - a. Click **Browser SSO** under the **SP Connection** tab.
  - b. Click **Configure Browser SSO**.
  - c. Click **SAML Profiles** on the **Summary** screen.
  - d. Ensure that IdP & SP initiated SSO are selected.
6. Activate the SP Connection.

 Note: PAS does not support SLO profiles at this time, and you can leave them unchecked.

## Switching Application Domains

This topic describes how to change the domain of an existing [Pivotal Cloud Foundry](#) (PCF) installation, using an example domain change from `myapps.mydomain.com` to `newapps.mydomain.com`.

1. In PCF Ops Manager, select the **Pivotal Application Service** tile.
2. Select **Domains** from the menu to see the current **Apps Domain** for your Pivotal Application Service (PAS) deployment. In the following example it is `myapps.mydomain.com`.

Elastic Runtime hosts applications at subdomains under its apps domain and assigns system components to subdomains under its system domain. You need to configure a wildcard DNS for both the apps domain and system domain. The two domains can be the same, although this is not recommended.

|                                                                                                                                                                                                                                                |                                    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------|
| System Domain *                                                                                                                                                                                                                                | <code>mysystem.mydomain.com</code> |
| Apps Domain *                                                                                                                                                                                                                                  | <code>myapps.mydomain.com</code>   |
| <small>Default parent domain that pushed apps use for their hostnames. This domain also requires a wildcard DNS record. Use the Cloud Foundry command line interface (cf CLI) to add or delete subdomains assigned to individual apps.</small> |                                    |
| <input type="button" value="Save"/>                                                                                                                                                                                                            |                                    |

3. In the terminal, run `cf login -a YOUR_API_ENDPOINT`. The cf CLI prompts you for your PCF username and password, as well as the org and space you want to access. See [Identifying the API Endpoint for your PAS Instance](#) if you don't know your API endpoint.
4. Run `cf domains` to view the domains in the space. If you have more than one shared domain, ensure that the domain you want to change is at the top of the list before you apply the new domain to your PAS tile configuration. You can delete and re-create the other shared domains as necessary to push the domain you want to change to the top of the list. If you do this, make sure to [re-map the routes for each domain](#).

```
$ cf domains
Getting domains in org my-org as admin...
name status
myapps.mydomain.com shared
```

5. Run `cf routes` to confirm that your apps are assigned to the domain you plan to change.

```
$ cf routes
Getting routes as admin ...
space host domain apps
my-space myapp myapps.mydomain.com myapp
```

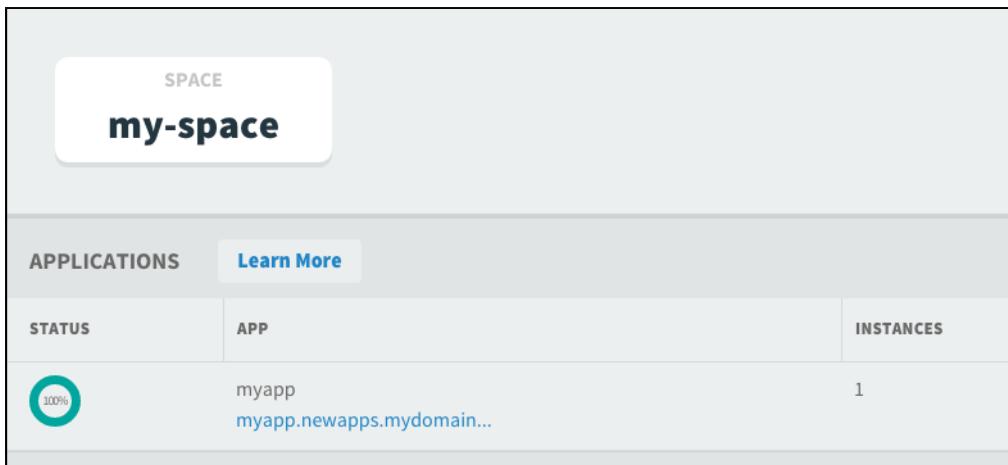
6. Run `cf create-shared-domain YOUR_DESIRED_NEW_DOMAIN` to create the new domain you want to use:

```
$ cf create-shared-domain newapps.mydomain.com
Creating shared domain newapps.mydomain.com as admin...
OK
```

7. Run `cf map-route APP_NAME NEW_DOMAIN -n HOST_NAME` to map the new domain to your app. In this example both the `NEW_DOMAIN` and `HOST_NAME` arguments are `myapp`, since this is both the name of the app to which we are mapping a route, and the intended hostname for the URL.

```
$ cf map-route myapp newapps.mydomain.com -n myapp
Creating route myapp.newapps.mydomain.com for org my-org / space my-space as admin...
OK
Adding route myapp.newapps.mydomain.com to app myapp in org my-org / space my-space as admin...
OK
```

8. Repeat the previous step for each app in this space. Afterwards, check Apps Manager to confirm that the route URL has updated correctly for each app:



| STATUS | APP                                | INSTANCES |
|--------|------------------------------------|-----------|
|        | myapp<br>myapp.newapps.mydomain... | 1         |

9. Repeat the above steps for each space in your PCF installation except for the System org, beginning with logging into the org and space and ending with confirming the URL update.

**Note:** Ordinarily the System org contains only PCF apps that perform utility functions for your installation. Pivotal does not recommend pushing apps to this org. However, if you have pushed apps to System, you must also repeat the above steps for these apps.

10. Once you have confirmed that every app in every space has been mapped to the new domain, delete the old domain by running `cf delete-shared-domain OLD_DOMAIN_TO_DELETE`:

```
$ cf delete-shared-domain myapps.mydomain.com
Deleting domain myapps.mydomain.com as admin...
```

This domain is shared across all orgs.  
 Deleting it will remove all associated routes, and will make any app with this domain unreachable.  
 Are you sure you want to delete the domain myapps.mydomain.com?  
 > yes  
 OK

11. Configure your PAS tile to use the new domain, and apply changes. Apps that you push after your update finishes use this new domain.

Elastic Runtime hosts applications at subdomains under its apps domain and assigns system components to subdomains under its system domain. You need to configure a wildcard DNS for both the apps domain and system domain. The two domains can be the same, although this is not recommended.

---

System Domain \*

This domain is for system-level PCF components, such as Apps Manager, service brokers, etc. You must set up a wildcard DNS record for this domain that points to your entry point load balancer or HAProxy.

Apps Domain \*

**Save**

## Scaling PAS

This topic discusses how to scale Pivotal Application Service (PAS) for different deployment scenarios. To increase the capacity and availability of the Pivotal Cloud Foundry (PCF) platform, and to decrease the chances of downtime, you can scale a deployment up using the instructions below.

If you want to make a PCF configuration highly available, see the [High Availability in Cloud Foundry](#) topic.

 **Note:** In PCF v1.11 and later, PAS defaults to a highly available resource configuration.

## Scaling Recommendations

The following table provides recommended instance counts for a high-availability deployment and the minimum instances for a functional deployment:

| Pivotal Application Service (PAS) Job | Recommended Instance Number for HA | Minimum Instance Number | Notes                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|---------------------------------------|------------------------------------|-------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Diego Cell                            | $\geq 3$                           | 1                       | The optimal balance between CPU/memory sizing and instance count depends on the performance characteristics of the apps that run on Diego cells. Scaling vertically with larger Diego cells makes for larger points of failure, and more apps go down when a cell fails. On the other hand, scaling horizontally decreases the speed at which the system re-balances apps. Re-balancing 100 cells takes longer and demands more processing overhead than re-balancing 20 cells. |
| Diego Brain                           | $\geq 2$                           | 1                       | For high availability, use at least one per AZ, or at least two if only one AZ.                                                                                                                                                                                                                                                                                                                                                                                                 |
| Diego BBS                             | $\geq 2$                           | 1                       | For high availability in a multi-AZ deployment, use at least one instance per AZ. Scale Diego BBS to at least two instances for high availability in a single-AZ deployment.                                                                                                                                                                                                                                                                                                    |
| MySQL Server                          | 3                                  | 1                       | If you use an external database in your deployment, then you can set the MySQL Server instance count to <a href="#">0</a> . For instructions about scaling down an internal MySQL cluster, see <a href="#">Scaling Down Your MySQL Cluster</a> .                                                                                                                                                                                                                                |
| MySQL Proxy                           | 2                                  | 1                       | If you use an external database in your deployment, then you can set the MySQL Proxy instance count to <a href="#">0</a> .                                                                                                                                                                                                                                                                                                                                                      |
| NATS Server                           | $\geq 2$                           | 1                       | In a high availability deployment, you might run a single NATS instance if your deployment lacks the resources to deploy two stable NATS servers. Components using NATS are resilient to message failures and the BOSH Resurrector recovers the NATS VM quickly if it becomes non-responsive.                                                                                                                                                                                   |
| Cloud Controller                      | $\geq 2$                           | 1                       | Scale the Cloud Controller to accommodate the number of requests to the API and the number of apps in the system.                                                                                                                                                                                                                                                                                                                                                               |
| Clock Global                          | $\geq 2$                           | 1                       | For a high availability deployment, scale the Clock Global job to a value greater than 1 or to the number of AZs you have.                                                                                                                                                                                                                                                                                                                                                      |
| Router                                | $\geq 2$                           | 1                       | Scale the router to accommodate the number of incoming requests. Additional instances increase available bandwidth. In general, this load is much less than the load on Diego cells.                                                                                                                                                                                                                                                                                            |
| UAA                                   | $\geq 2$                           | 1                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| Doppler Server                        | $\geq 2$                           | 1                       | Deploying additional Doppler servers splits traffic across them. For a high availability deployment, Pivotal recommends at least two per Availability Zone.                                                                                                                                                                                                                                                                                                                     |
| Loggregator Traffic Controller        | $\geq 2$                           | 1                       | Deploying additional Loggregator Traffic Controllers allows you to direct traffic to them in a round-robin manner. For a high availability deployment, Pivotal recommends at least two per Availability Zone.                                                                                                                                                                                                                                                                   |
| Syslog Scheduler                      | $\geq 2$                           | 1                       | The Syslog Scheduler is a scalable component. For high availability, use at least one instance per AZ, or at least two instances if only one AZ is present.                                                                                                                                                                                                                                                                                                                     |

## Scaling Up PAS

To scale up PAS instances, do the following:

1. Navigate to the Pivotal Cloud Foundry Operations Manager Installation Dashboard.
2. Click the PAS tile in the Installation Dashboard.

3. Select **Resource Config**.

| JOB                           | INSTANCES    | PERSISTENT DISK TYPE | VM TYPE                                                   |
|-------------------------------|--------------|----------------------|-----------------------------------------------------------|
| Consul                        | Automatic: 3 | Automatic: 1 GB      | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| NATS                          | Automatic: 2 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| File Storage                  | Automatic: 1 | Automatic: 100 GB    | Automatic: medium.mem (cpu: 1, ram: 6 GB, disk: 8 GB)     |
| MySQL Proxy                   | Automatic: 2 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| MySQL Server                  | Automatic: 3 | Automatic: 100 GB    | Automatic: large.disk (cpu: 2, ram: 8 GB, disk: 64 GB)    |
| Backup Prepare Node           | Automatic: 1 | Automatic: 200 GB    | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| Diego BBS                     | Automatic: 3 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| UAA                           | Automatic: 2 | None                 | Automatic: medium.disk (cpu: 2, ram: 4 GB, disk: 32 GB)   |
| Cloud Controller              | Automatic: 2 | None                 | Automatic: medium.disk (cpu: 2, ram: 4 GB, disk: 32 GB)   |
| HAProxy                       | Automatic: 1 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| Router                        | Automatic: 3 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| MySQL Monitor                 | Automatic: 1 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| Clock Global                  | Automatic: 1 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| Cloud Controller Worker       | Automatic: 2 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| Diego Brain                   | Automatic: 3 | Automatic: 1 GB      | Automatic: small (cpu: 1, ram: 2 GB, disk: 8 GB)          |
| Diego Cell                    | Automatic: 3 | None                 | Automatic: xlarge.disk (cpu: 4, ram: 16 GB, disk: 128 GB) |
| Loggregator Trafficcontroller | Automatic: 3 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| Syslog Adapter                | Automatic: 3 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| Syslog Scheduler              | Automatic: 2 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| Doppler Server                | Automatic: 3 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| TCP Router                    | Automatic: 1 | Automatic: 1 GB      | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| CredHub                       | Automatic: 2 | None                 | Automatic: large (cpu: 2, ram: 8 GB, disk: 16 GB)         |

4. To scale your deployment horizontally, increase the number of **Instances** of a job. See [Scaling Recommendations](#) for guidance about the number of job instances required to ensure high availability.

 **Note:** In PCF v1.12, you cannot scale the Autoscaler job to greater than one instance.

5. To scale your deployment vertically, adjust the **Persistent Disk Type** and **VM Type** of a job to allocate more disk space and memory. If you choose **Automatic** from the dropdown, PAS uses the recommended amount of resources for the job.

6. Click **Save**.

7. Return to the **Installation Dashboard**, click **Review Pending Changes**, and click **Apply Changes**.

## Scaling Down PAS

If you are deploying a PCF configuration that does not need to be highly available, Pivotal recommends scaling down job instances to the minimum number required for a functional deployment.

To scale down your deployment, do the following:

1. Navigate to the Pivotal Cloud Foundry Operations Manager Installation Dashboard.
2. Click the PAS tile in the Installation Dashboard.
3. Select **Resource Config**.

## Resource Config

| JOB                           | INSTANCES    | PERSISTENT DISK TYPE | VM TYPE                                                   |
|-------------------------------|--------------|----------------------|-----------------------------------------------------------|
| Consul                        | Automatic: 3 | Automatic: 1 GB      | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| NATS                          | Automatic: 2 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| File Storage                  | Automatic: 1 | Automatic: 100 GB    | Automatic: medium.mem (cpu: 1, ram: 6 GB, disk: 8 GB)     |
| MySQL Proxy                   | Automatic: 2 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| MySQL Server                  | Automatic: 3 | Automatic: 100 GB    | Automatic: large.disk (cpu: 2, ram: 8 GB, disk: 64 GB)    |
| Backup Prepare Node           | Automatic: 1 | Automatic: 200 GB    | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| Diego BBS                     | Automatic: 3 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| UAA                           | Automatic: 2 | None                 | Automatic: medium.disk (cpu: 2, ram: 4 GB, disk: 32 GB)   |
| Cloud Controller              | Automatic: 2 | None                 | Automatic: medium.disk (cpu: 2, ram: 4 GB, disk: 32 GB)   |
| HAProxy                       | Automatic: 1 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| Router                        | Automatic: 3 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| MySQL Monitor                 | Automatic: 1 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| Clock Global                  | Automatic: 1 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| Cloud Controller Worker       | Automatic: 2 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| Diego Brain                   | Automatic: 3 | Automatic: 1 GB      | Automatic: small (cpu: 1, ram: 2 GB, disk: 8 GB)          |
| Diego Cell                    | Automatic: 3 | None                 | Automatic: xlarge.disk (cpu: 4, ram: 16 GB, disk: 128 GB) |
| Loggregator Trafficcontroller | Automatic: 3 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| Syslog Adapter                | Automatic: 3 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| Syslog Scheduler              | Automatic: 2 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| Doppler Server                | Automatic: 3 | None                 | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| TCP Router                    | Automatic: 1 | Automatic: 1 GB      | Automatic: micro (cpu: 1, ram: 1 GB, disk: 8 GB)          |
| CredHub                       | Automatic: 2 | None                 | Automatic: large (cpu: 2, ram: 8 GB, disk: 16 GB)         |

- In the **Resource Config** screen, decrease the number of **Instances** for each job. Choose the suggested values outlined in [Scaling Recommendations](#), or in the [Scaling Recommendations for Specific Deployment Configurations](#) and click **Save**.
- Return to the [Installation Dashboard](#), click **Review Pending Changes**, and click **Apply Changes**.

## Scaling Down Jobs with Persistent Disk

If you scale down or delete a job that uses persistent disk, PAS marks the disk as **orphaned**. Orphaned disks are not attached to any job, and PAS deletes them after five days.

Use the BOSH CLI to list and recover orphaned disks. Follow the instructions in the [Prepare to Use the BOSH CLI](#) section of the [Advanced Troubleshooting with the BOSH CLI](#) topic to log in to the BOSH Director, and then follow the procedures in [Orphaned Disks](#) in the BOSH documentation.

## Scaling Recommendations for Specific Deployment Configurations

If you use one of the following configurations, choose the values in the corresponding table to scale instances for your particular deployment:

- [Deployments Using External Databases](#)
- [Deployments Using Internal MySQL](#)
- [Deployments Using an External Blobstore](#)
- [Deployments Using External Load Balancers](#)

### Deployments Using External Databases

If you use an external database, you can scale down the instance counts for internal MySQL jobs.

Select the following values in the Resource Config:

| Job | Instance Count |
|-----|----------------|
|-----|----------------|

|              |   |
|--------------|---|
| MySQL Server | 0 |
| MySQL Proxy  | 0 |

## Deployments Using Internal MySQL

If you use the internal MySQL database on a clean install, or on an upgrade from a configuration that previously used internal MySQL databases, you do not need to change the default values shown in the table below.

If you need to change back to this configuration, choose the values shown below in the Resource Config.

 **Note:** Changing back to this configuration deletes any data written to your other database option.

| Job          | Instance Count |
|--------------|----------------|
| MySQL Server | 3              |
| MySQL Proxy  | 2              |

 **Note:** Apps that do not use MySQL for PCF are not affected by the scaling process when you redeploy PAS. In addition, redeploying PAS with the MySQL cluster means that the PCF API will not be available for a brief period of time. For example, you are not able to push apps or query their state during this time.

 **Note:** For MySQL high availability, you need to configure an external load balancer. For more information about configuring a load balancer for MariaDB-based deployments, see the [Configure a Load Balancer](#) section of the *Installing MySQL for PCF* topic.

## Deployments Using an External Blobstore

If you use an external blobstore, choose the following value in the Resource Config:

| Job          | Instance Count |
|--------------|----------------|
| File Storage | 0              |

## Deployments Using External Load Balancers

If you use an external load balancer, choose the following values in the Resource Config:

| Job         | Instance Count |
|-------------|----------------|
| HAProxy     | 0              |
| Router      | $\geq 1$       |
| Diego Brain | $\geq 1$       |

For more information about configuring load balancers in the Resource Config section of PAS, see [Configuring Load Balancing for PAS](#).

## Using Docker Registries

This topic describes how to configure Pivotal Cloud Foundry (PCF) to access Docker registries such as [Docker Hub](#), by using either a root certificate authority (CA) certificate or by adding its IP address to a whitelist. It also explains how to configure PCF to access Docker registries through a proxy.

Docker registries store Docker container images. PCF uses these images to create the Docker containers that it runs apps in.

### Prerequisite: Enable Docker Support

PCF can only access Docker registries if an operator has enabled Docker support with the `cf enable-feature-flag diego_docker` command, as described in the [Using Docker in Cloud Foundry](#) topic.

With Docker enabled, developers can push an app with a Docker image using the Cloud Foundry Command Line Interface (cf CLI). For more information, see the [Deploying an App with Docker](#) topic.

### Use a CA Certificate

If you provide your root CA certificate in the Ops Manager configuration, follow this procedure:

1. In the Ops Manager Installation Dashboard, click the **BOSH Director** tile.
2. Click **Security**.

3. In the **Trusted Certificates** field, paste one or more root CA certificates. The Docker registry does not use the CA certificate itself but uses a certificate that is signed by the CA certificate.
4. Click **Save**.
5. Choose one of the following:
  - o If you are configuring Ops Manager for the first time, return to your specific IaaS installation instructions ([AWS](#), [Azure](#), [GCP](#), [OpenStack](#), [vSphere](#)) to continue the installation process.
  - o If you are modifying an existing Ops Manager installation, return to the Ops Manager Installation Dashboard, click **Review Pending Changes**, and click **Apply Changes**.

After configuration, BOSH propagates your CA certificate to all application containers in your deployment. You can then push and pull images from your Docker registries.

## Use an IP Address Whitelist

If you choose not to provide a CA certificate, you must provide the IP address of your Docker registry.

**Note:** Using a whitelist skips SSL validation. If you want to enforce SSL validation, enter the IP address of the Docker registry in the **No proxy** field described [below](#).

1. Navigate to the Ops Manager Installation Dashboard.
2. Click the **Pivotal Application Service** tile, and navigate to the **Application Containers** tab.

Enable microservice frameworks, private Docker registries, and other services that support your applications at a container level.

Enable Custom Buildpacks

Allow SSH access to app containers

Enable SSH when an app is created

When enabled and the space has SSH enabled, a new app will have SSH enabled. Turning this feature off will disable SSH for new apps but allow developers to enable it when necessary.

Router application identity verification\*

Router uses TLS to verify application identity (Default)

Router and applications use mutual TLS to verify each other's identity (incompatible with TCP Routing and SSH access to app containers)

Private Docker Insecure Registry Whitelist

10.10.10.10:8888,example.com:8888



Docker Images Disk-Cleanup Scheduling on Cell VMs\*

Never clean up Cell disk-space

Routinely clean up Cell disk-space

Clean up disk-space once usage fills disk

Reserved amount of Disk (MB) ( min: 1 ) \*

15360

3. Select **Allow SSH access to app containers** to enable app containers to accept SSH connections. If you use a load balancer instead of HAProxy, you must open port `2222` on your load balancer to enable SSH traffic. To open an SSH connection to an app, a user must have Space Developer privileges for the space where the app is deployed. Operators can grant those privileges in Apps Manager or using the cf CLI.
4. For **Private Docker Insecure Registry Whitelist**, provide the hostname or IP address and port that point to your private Docker registry. For example, enter `198.51.100.1:80` or `mydockerregistry.com:80`. Enter multiple entries in a comma-delimited sequence. SSL validation is ignored for private Docker image registries secured with self-signed certificates at these locations.

5. Under **Docker Images Disk-Cleanup Scheduling** on Cell VMs, choose one of the options listed below. For more information about these options, see [Configuring Docker Images Disk-Cleanup Scheduling](#).

- **Never clean up Cell disk-space**
- **Routinely clean up Cell disk-space**
- **Clean up disk-space once threshold is reached.** If you choose this option, enter the amount of disk space limit the Cell must reach before disk cleanup initiates under **Threshold of Disk-Used (MB)**.

6. Click **Save**.

7. Choose one of the following:

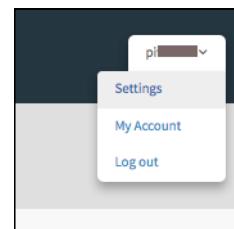
- If you are configuring Pivotal Application Service (PAS) for the first time, return to your specific IaaS installation instructions ([AWS](#), [Azure](#), [GCP](#), [OpenStack](#), [vSphere](#)) to continue the installation process.
- If you are modifying an existing PAS installation, return to the Ops Manager Installation Dashboard, click **Review Pending Changes**, and click **Apply Changes**.

After configuration, PAS allows Docker images to pass through the specified IP address without checking certificates.

## Configure PCF to Access Proxies for Docker Registries

If you have proxies already set up for Docker registries, you should configure PCF to access your Docker registries through a proxy.

To configure PCF to access a Docker registry through a proxy, do the following:



1. On the Installation Dashboard, navigate to **USERNAME > Settings > Proxy Settings**.

2. On the **Update Proxy Settings** pane, complete one of the following fields:

- **HTTP proxy:** If you have an HTTP proxy server for your Docker registry, enter its IP address.
- **HTTPS proxy:** If you have an HTTPS proxy server for your Docker registry, enter its IP address.
- **No proxy:** If you do not use a proxy server, enter the IP address for the Docker registry. This field may already contain proxy settings for the BOSH Director.

Enter multiple IP addresses as a comma-separated list.

| Update Proxy Settings        |                                                     |
|------------------------------|-----------------------------------------------------|
| Decryption Passphrase        | Http proxy                                          |
| Authentication Method        | <input type="text" value="Http proxy"/>             |
| External API Access          | Https proxy                                         |
| Proxy Settings               | <input type="text" value="Https proxy"/>            |
| Export Installation Settings | No proxy                                            |
| Advanced                     | <input type="text" value="10.10.10.4, 10.10.10.5"/> |
| <b>Update</b>                |                                                     |

3. Click **Update**.
4. Return to the Ops Manager dashboard, click **Review Pending Changes**, and click **Apply Changes**.

## Configuring Cell Disk Cleanup Scheduling

This topic describes how to configure disk cleanup scheduling on Diego cells in Pivotal Cloud Foundry (PCF).

### What is Disk Cleanup

PCF isolates application instances (AIs) from each other using containers that run inside Diego cells. Containers enforce a set of isolation layers including file system isolation. A PCF container file system can either be a PCF stack or the result of pulling a Docker image.

For performance reasons, the cells cache the Docker image layers and the PCF stacks used by running AIs. When PCF destroys an AI or reschedules an AI to a different cell, a chance exist that certain Docker image layers or an old PCF stack becomes unused. If PCF does not clean these unused layers, the cell ephemeral disk will slowly fill.

Disk cleanup is the process of removing unused layers from the cell disk. The disk cleanup process removes all unused Docker image layers and old PCF stacks, regardless of their size or age.

To perform a detailed analysis of disk usage in your PAS deployment, see [Examining GrootFS Disk Usage](#).

### Options for Disk Cleanup

PCF provides the following options for scheduling the disk cleanup process on Diego cells:

- **Never clean up the Cell disk-space:** Pivotal does not recommend selecting this option for production environments.
- **Routinely clean up the Cell disk-space:** This option makes the cell schedule a disk cleanup whenever a container is created. Running the disk cleanup process this frequently can result in a negative impact on the cell performance.
- **Clean up disk-space once usage fills disk:** This option makes the cell schedule the disk cleanup process only when disk usage encroaches on the space reserved for other jobs.

See the [Configure Disk Cleanup Scheduling](#) section of this topic to select one of these options.

### Recommendations

To select the best option for disk cleanup, you must consider the workload that the Diego cells run.

For PCF installations that primarily run buildpack-based apps, Pivotal recommends selecting the **Routinely clean up Cell disk-space** option. The **Routinely clean up Cell disk-space** option ensures that when a new stack becomes available on a cell, the old stack is dropped immediately from the cache.

For PCF installations that primarily run Docker images, or both Docker images and buildpack-based apps, Pivotal recommends selecting the **Clean up disk-space once usage fills disk** option.

### Calculating a Reserve

By default, the space reserved for other jobs is set to a reasonable value for the majority of PCF deployments. The reserve default value accounts for the most typical deployment case where the only other component that writes to `/var/vcap/data` is the Diego executor cache.

However, if you are deploying any BOSH add-ons that require space in the ephemeral disk, you may need to increase the reserve.

For more information on disk usage and garbage collection in GrootFS, see [GrootFS Disk Usage](#) and [Examining GrootFS Disk Usage](#).

**⚠ warning:** Setting the reserve to an excessive value can cause garbage collection to occur more frequently. An increase in frequency can result in regular slowdowns in performance of apps on particular Diego Cells. Pivotal recommends leaving the default value for the reserve unless you are advised to modify it.

## Configure Disk Cleanup Scheduling

1. Navigate to the PCF Operations Manager **Installation Dashboard**.
2. Click the Pivotal Application Service (PAS) tile, and navigate to the **Application Containers** tab.

Enable microservice frameworks, private Docker registries, and other services that support your applications at a container level.

Enable Custom Buildpacks

Allow SSH access to app containers

Enable SSH when an app is created

When enabled and the space has SSH enabled, a new app will have SSH enabled. Turning this feature off will disable SSH for new apps but allow developers to enable it when necessary.

Router application identity verification\*

Router uses TLS to verify application identity (Default)

Router and applications use mutual TLS to verify each other's identity (incompatible with TCP Routing and SSH access to app containers)

Private Docker Insecure Registry Whitelist

10.10.10.10:8888,example.com:8888



Docker Images Disk-Cleanup Scheduling on Cell VMs\*

Never clean up Cell disk-space

Routinely clean up Cell disk-space

Clean up disk-space once usage fills disk

Reserved amount of Disk (MB) ( min: 1 ) \*

15360

3. Under **Docker Images Disk-Cleanup Scheduling on Cell VMs**, select an option.

4. If you select **Clean up disk-space once usage fills disk**, you can modify the **Reserved amount of Disk (MB)** field. Pivotal recommends that you leave the default value of **15360** MB, or 15GB, in **Reserved amount of Disk (MB)** unless you are instructed to change it. For more information, see [Calculating a Reserve](#).

5. Click **Save**.

## Next Steps

If you are configuring PAS for the first time, return to your specific IaaS configuration to continue the installation process.

If you are modifying an existing PAS installation, return to the **PCF Ops Manager Installation Dashboard**, click **Review Pending Changes**, and click **Apply Changes**.

## Custom Branding Apps Manager

This topic describes how Pivotal Cloud Foundry operators can visually brand Apps Manager by changing certain text, colors, and images of the interface. Developers view the customized interface when logging in, creating an account, resetting a password, or using Apps Manager.

Operators customize Apps Manager by configuring the **Custom Branding** and **Apps Manager Config** pages of the Pivotal Application Service (PAS) tile.

### Custom Branding Page

1. In a browser, navigate to the fully qualified domain name (FQDN) of your Ops Manager and log in.
2. Click **Pivotal Application Service**.
3. Click the **Custom Branding** tab.

Company Name  
 Defaults to 'Pivotal'

Accent Color  
 Enter a hexadecimal color code like '#71ffda'

Main Logo (PNGs only)  
  
Enter a base64-encoded PNG image string, but leave out the mime-type (data:image/png;base64,) string. Only enter the base64 encoded data.

Square Logo (PNGs only)  
  
Enter a base64-encoded PNG image string, but leave out the mime-type (data:image/png;base64,) string. Only enter the base64 encoded data.

Favicon (PNGs only)  
  
Enter a base64-encoded PNG image string, but leave out the mime-type (data:image/png;base64,) string. Only enter the base64 encoded data.

Footer Text  
 Defaults to 'Pivotal Software Inc. All rights reserved.'

Footer Links  
You may configure up to three links in the Apps Manager footer.

Classification Header/Footer Background Color  
 Enter a hexadecimal color code like '#71ffda'

Classification Header/Footer Text Color  
 Enter a hexadecimal color code like '#71ffda'

Classification Header Content  
  
Plain text or HTML Markup

Classification Footer Content  
  
Plain text or HTML Markup

**Save**

4. For **Company Name**, enter the name of your organization. If left blank, the name defaults to **Pivotal**.
5. For **Accent Color**, enter the hexadecimal code for the color used to accent various visual elements, such as the currently selected space in the sidebar. For example, `#71ffda`.
6. For **Main Logo**, enter a Base64-encoded URL string for a PNG image to use as your main logo. The image can be square or wide. For example, `[data:image/png;base64,iVBORw0....]`. If left blank, the image defaults to the Pivotal Logo.
7. For **Square Logo**, enter a Base64-encoded URL string for a PNG image to use in the Apps Manager header and in places that require a smaller logo. For example, `[data:image/png;base64,iVBORw0....]`. If left blank, the image defaults to the Pivotal Logo.
8. For **Favicon**, enter a Base64-encoded URL string for a PNG image to use as your favicon. For example, `[data:image/png;base64,iVBORw0....]`. If left blank, the image defaults to the Pivotal Logo.
9. For **Footer Text**, enter a string to be displayed as the footer. If left blank, the footer text defaults to **Pivotal Software Inc. All rights reserved..**
10. To add up to three footer links that appear to the right of the footer text, complete the following steps:
  - o Click **Add**.
  - o For **Link text**, enter a label for the link.
  - o For **Url**, enter an external or relative URL. For example, `http://docs.pivotal.io` or `/tools.html`.
11. For special notification purposes such as governmental or restricted usage, use the Classification fields to create a special Header and Footer. Enter values in the following fields:
  - o For **Classification Header/Footer Background Color**, enter the hexadecimal code for the desired background color of the header and footer.
  - o For **Classification Header/Footer Text Color**, enter the hexadecimal code for the desired color of header and footer text.
  - o For **Classification Header Content**, enter content for the header in either plain text or HTML. If you enter HTML content, eliminate white spaces and new lines. If you do not provide any content, the custom header will not appear.
  - o For **Classification Footer Content**, enter content for the footer in either plain text or HTML. If you enter HTML content, eliminate white spaces and new lines. If you do not provide any content, the custom footer will not appear. The Classification footer appears below the normal footer, which you can customize in the **Footer Text** and **Footer Links** fields.

 Note: The Header and Footer do not appear on the [User Account and Authentication \(UAA\)](#) login page.

## Apps Manager Config Page

1. In a browser, navigate to the fully qualified domain name (FQDN) of your Ops Manager and log in.
2. Click **Pivotal Application Service**.
3. Click the **Apps Manager Config** tab.

### Configure Apps Manager

Enable Invitations

Display Marketplace Service Plan Prices

Supported currencies as JSON \*

{ "usd": "\$", "eur": "€" }

Product Name

Replace "Apps Manager" in the header with a custom product

Marketplace Name

Customize Sidebar Links \*

You may configure up to 10 links in the Apps Manager sidebar.

▶ Marketplace

▶ Docs

▶ Tools

Apps Manager Memory Usage (MB) ( min: 128 )

Invitations Memory Usage (MB) ( min: 256 )

Apps Manager Polling Interval \*

App Details Polling Interval \*

**Save**

4. For **Product Name**, enter text to replace **Apps Manager** in the header and the title of Apps Manager. This text defaults to **Apps Manager** if left blank.
5. For **Marketplace Name**, enter text to replace the header in the Marketplace pages. This text defaults to **Marketplace** if left blank.
6. By default, Apps Manager includes three sidebar links: **Marketplace**, **Docs**, and **Tools**. You can edit existing sidebar links by clicking the name of the link and editing the **Link text** and **Url** fields. Or, you can remove the link by clicking the trash icon next to its name. If you want to add a new sidebar link, click **Add** and complete the **Link text** and **Url** fields.

**Note:** Removing any of the default links will remove them from the sidebar for all users.

7. (Optional) Enter your desired **Apps Manager Memory Usage** in MBs. The minimum number you can enter is **128**.

8. (Optional) Enter your desired **Invitations Memory Usage** in MBs. This is the memory limit used to deploy the Invitations app. The minimum number you can enter is `256`.

9. The **Apps Manager Polling Interval** field provides a temporary fix if Apps Manager usage degrades Cloud Controller response times. In this case, you can use this field to reduce the load on the Cloud Controller and ensure Apps Manager remains available while you troubleshoot the Cloud Controller. Pivotal recommends that you do not keep this field modified as a long term fix because it can degrade Apps Manager performance. You can optionally do the following:

- Increase the polling interval above the default of `30` seconds.

 **Note:** If you enter a value between `0` and `30`, the field is automatically set to `30`.

- Disable polling by entering `0`. This stops Apps Manager from refreshing data automatically, but users can update displayed data by reloading Apps Manager manually.

10. The **App Details Polling Interval** field provides an additional way to reduce the load on the Cloud Controller when the **Apps Manager Polling Interval** field is not sufficient. This field controls the rate at which Apps Manager polls for data when a user views the **Overview** page of an app. Pivotal recommends that you do not keep this field modified as a long term fix because it can degrade Apps Manager performance. You can optionally do the following:

- Increase the polling interval above the default of `10` seconds.

 **Note:** If you enter a value between `0` and `30`, the field is automatically set to `10`.

- Disable polling by entering `0`. This stops Apps Manager from refreshing data automatically, but users can update displayed data by reloading Apps Manager manually.

## Reporting App, Task, and Service Instance Usage

This topic describes how to use the Cloud Foundry Command Line Interface (cf CLI) to retrieve historical system- and org-level usage information about your apps, tasks, and service instances through the Cloud Controller and Usage Service APIs.

Usage reports are compiled from the `/v2/app_usage_events` endpoint. For more information, see [List all App Usage Events](#) in the Cloud Foundry API documentation.

You can also access usage information by using Apps Manager. For more information, see the [Reporting Instance Usage with Apps Manager](#) topic.

To retrieve current app and event information from the Cloud Controller, see [Retrieving App and Event Information](#) in the open-source Cloud Foundry documentation.

### Prerequisite

Before performing the procedures in this topic, ensure that your user is a member of the `cloud_controller.admin` or `usage_service.audit` group. For example, a UAAC admin can add a user to the `usage_service.audit` group with the following command:

```
uaac member add usage_service.audit USERNAME
```

## Obtain System Usage Information

You can obtain the following system-wide usage information:

- [App Usage](#)
- [Task Usage](#)
- [Service Usage](#)
- [All Orgs App Usage](#)

### App Usage

Use `curl` to make a request to `/system_report/app_usages` on the Usage Service endpoint to show system-wide app usage data:

```
$ curl "https://app-usage.YOUR-SYSTEM-DOMAIN/system_report/app_usages" -k -v \
-H "authorization: cf oauth-token"
{
 "report_time": "2017-04-11 22:28:24 UTC",
 "monthly_reports": [
 {
 "month": 4,
 "year": 2017,
 "average_app_instances": 17855.256153713308,
 "maximum_app_instances": 18145,
 "app_instance_hours": 4686533.080277303
 }
],
 "yearly_reports": [
 {
 "year": 2017,
 "average_app_instances": 16184.9,
 "maximum_app_instances": 18145,
 "app_instance_hours": 39207433.0802773
 }
]
}
```

The following table describes the data generated by this report:

| Value                           | Calculation Method                                                                                 |
|---------------------------------|----------------------------------------------------------------------------------------------------|
| <code>app_instance_hours</code> | The total number of hours app instances have been running on the foundation for the month or year. |

|                                           |                                                                                                                                                                                                                                                                                                                                                                                                                         |
|-------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><code>average_app_instances</code></p> | <p>The total <code>app_instance_hours</code> on the foundation for the month or year divided by the total hours for that time period.</p> <p><b>Example:</b> For the month of January, 100 apps run for 300 hours each, generating a value of 30000 for <code>app_instance_hours</code>. When divided by 744, the total number of hours in January, you get a value of 40.3 for <code>average_app_instances</code>.</p> |
| <p><code>maximum_app_instances</code></p> | <p>The highest concurrent number of app instances running on the foundation for the month or year.</p>                                                                                                                                                                                                                                                                                                                  |

## Task Usage

Use `curl` to make a request to `/system_report/task_usages` on the Usage Service endpoint to show system-wide task usage data:

```
$ curl "https://app-usage.YOUR-SYSTEM-DOMAIN/system_report/task_usages" -k -v \
-H "authorization: cf oauth-token"
{
 "report_time": "2017-04-11T22:33:48.971Z",
 "monthly_reports": [
 {
 "month": 4,
 "year": 2017,
 "total_task_runs": 235,
 "maximum_concurrent_tasks": 7,
 "task_hours": 43045.20194444445
 }
],
 "yearly_reports": [
 {
 "year": 2017,
 "total_task_runs": 2894,
 "maximum_concurrent_tasks": 184,
 "task_hours": 45361.26694444445
 }
]
}
```

## Service Usage

Use `curl` to make a request to `/system_report/service_usages` on the Usage Service endpoint to show system-wide service usage data:

```
$ curl "https://app-usage.YOUR-SYSTEM-DOMAIN/system_report/service_usages" -k -v \
-H "authorization: cf oauth-token"
{
 "report_time": "2017-05-11 18:29:14 UTC",
 "monthly_service_reports": [
 {
 "service_name": "fake-service-0507f1fd-2340-49a6-9d43-a347a5f5f6be",
 "service_guid": "177dcfde-cd51-4058-bd86-b98015c295f5",
 "usages": [
 {
 "month": 1,
 "year": 2017,
 "duration_in_hours": 0,
 "average_instances": 0,
 "maximum_instances": 0
 },
 {
 "month": 2,
 "year": 2017,
 "duration_in_hours": 0,
 "average_instances": 0,
 "maximum_instances": 0
 },
 {
 "month": 3,
 "year": 2017,
 "duration_in_hours": 4.182222222222227,
 "average_instances": 0,
 "maximum_instances": 2
 },
 {
 "month": 4,
 "year": 2017,
 "duration_in_hours": 0,
 "average_instances": 0,
 "maximum_instances": 0
 }
]
 }
]
}
```

```

"duration_in_hours": 2176.962222222186,
"average_instances": 3,
"maximum_instances": 7
},
{
"month": 5,
"year": 2017,
"duration_in_hours": 385.6138888888854,
"average_instances": 1.5,
"maximum_instances": 3
}
],
"plans": [
{
"usages": [
{
"month": 1,
"year": 2017,
"duration_in_hours": 0,
"average_instances": 0,
"maximum_instances": 0
},
{
"month": 2,
"year": 2017,
"duration_in_hours": 0,
"average_instances": 0,
"maximum_instances": 0
},
{
"month": 3,
"year": 2017,
"duration_in_hours": 4.18222222222227,
"average_instances": 0,
"maximum_instances": 2
},
{
"month": 4,
"year": 2017,
"duration_in_hours": 1465.638888888941,
"average_instances": 2,
"maximum_instances": 5
},
{
"month": 5,
"year": 2017,
"duration_in_hours": 385.6138888888854,
"average_instances": 1.5,
"maximum_instances": 3
}
],
"service_plan_name": "fake-plan",
"service_plan_guid": "ac09f607-f4e5-4807-af16-e95856061bd7"
}
]
}
]
}
}

```

## All Orgs Usage (Beta)

The `app_usages` endpoint directly under your system domain returns system-wide app usage for all orgs that you have access to. The command works like the org-specific [App Usage](#) command, except you do not supply an org GUID. You can set `order_by` to either `total_instance_hours` or `megabyte_hours`.

The `order_by` ordering defaults to ascending, which lists orgs with the most usage at the end of the output. To change the sort ordering to descending, prepend the `order_by` argument with a `-`, as in `order_by=-megabyte_hours`.

Use `curl` to make a request to `/app_usages` on the Usage Service endpoint to show system-wide organization usage data:

```
$ curl "https://app-usage.YOUR-SYSTEM-DOMAIN/app_usages?start=2018-01-01&end=2018-11-01&order_by=total_instance_hours" -k -v \
-H "authorization: cf oauth-token"
{
 "period_start": "2018-01-10T00:00:00Z",
 "period_end": "2018-11-01T23:59:59Z",
 "Organizations": [
 {
 "guid": "d5c9f947-df85-45af-abab-00a2174598f9",
 "total_instance_hours": "91535.6",
 "megabyte_hours": "1830712.2"
 },
 {
 "guid": "a9605fa4-99da-4932-84f5-3db8698c1fc1",
 "total_instance_hours": "99440.8",
 "megabyte_hours": "96698582.2"
 }
]
}
```

This beta API endpoint may change; please try it out and help Pivotal make it better.

## Obtain Org Usage Information

You can obtain the following org-specific usage information:

- [App Usage](#)
- [Task Usage](#)
- [Service Usage](#)

You must have the GUID of the org you want to obtain information about in order to perform the procedures in this section. To retrieve your org GUID, run the `cf org` command:

```
$ cf org YOUR-ORG --guid
```

## App Usage

Use `curl` to make a request to `/app_usages` on the Usage Service endpoint to show app usage in an org. You must complete the placeholders in `start=YYYY-MM-DD&end=YYYY-MM-DD` to define a date range.

```
$ curl "https://app-usage.YOUR-SYSTEM-DOMAIN/organizations/YOUR-ORG-GUID/app_usages?start=YYYY-MM-DD&end=YYYY-MM-DD" \
-k -v \
-H "authorization: `cf oauth-token`"

{
 "organization_guid": "88e485fa-2222-4993-8ef5-bac3711adfee",
 "period_start": "2019-01-16T00:00:00Z",
 "period_end": "2019-01-17T23:59:59Z",
 "app_usages": [
 {
 "space_guid": "d6988d4d-6175-4dee-b219-a6eabbc05514",
 "space_name": "development",
 "app_name": "static2",
 "app_guid": "1afc2b20-dd41-4ba4-8c6e-c894b5a2cb23",
 "parent_application_name": "static2",
 "parent_application_guid": "1afc2b20-dd41-4ba4-8e6e-c894b5a2cb23",
 "instance_count": 1,
 "memory_in_mb_per_instance": 1024,
 "duration_in_seconds": 172800
 },
 {
 "space_guid": "d6988d4d-6175-4dee-b219-a6eabbc05514",
 "space_name": "development",
 "app_name": "static",
 "app_guid": "54e767ff-4f83-45b1-bf06-46c501217845",
 "parent_application_name": "static",
 "parent_application_guid": "54e767ff-4f83-45b1-bf06-46c501217845",
 "instance_count": 1,
 "memory_in_mb_per_instance": 1024,
 "duration_in_seconds": 172800
 }
]
}
```

You may see results with “instance\_count”: 0. This happens when a v3 app is pushed but has processes that aren’t scaled up. The memory configuration and duration are still stored and can be useful for resource management. Apps with ‘0’ instances are not counted in the App Usages System Report above.

## Task Usage

Use `curl` to make a request to `/task_usages` on the Usage Service endpoint to show task usage in an org. You must complete the placeholders in `start=YYYY-MM-DD&end=YYYY-MM-DD` to define a date range.

```
$ curl "https://app-usage.YOUR-SYSTEM-DOMAIN/organizations/YOUR-ORG-GUID/task_usages?start=YYYY-MM-DD&end=YYYY-MM-DD" \
-k -v \
-H "authorization: `cf oauth-token`"
{
 "organization_guid": "8f88362f-547c-4158-808b-4605468387d5",
 "period_start": "2014-01-01",
 "period_end": "2017-04-04",
 "spaces": [
 {
 "e6445eb3-fdac-4049-bafc-94d1703d5e78": {
 "space_name": "smoketest",
 "task_summaries": [
 {
 "parent_application_guid": "04cd29d5-1f9e-4900-ac13-2e903f6582a9",
 "parent_application_name": "task-dummy-app",
 "memory_in_mb_per_instance": 256,
 "task_count_for_range": 54084,
 "concurrent_task_count_for_range": 5,
 "total_duration_in_seconds_for_range": 37651415
 }
]
 },
 "b66665e4-873f-4482-acf1-89307ba9c6e4": {
 "space_name": "smoketest-experimental",
 "task_summaries": [
 {
 "parent_application_guid": "d941b689-4a27-44ec-91d3-1f97434dbed9",
 "parent_application_name": "console-blue",
 "memory_in_mb_per_instance": 256,
 "task_count_for_range": 14,
 "concurrent_task_count_for_range": 2,
 "total_duration_in_seconds_for_range": 20583
 }
]
 }
 }
 }
}
```

 Note: In the `/task_usages` endpoint, `memory_in_mb_per_instance` is the memory of the task.

## Service Usage

Use `curl` to make a request to `/service_usages` on the Usage Service endpoint to show service usage in an org. You must complete the placeholders in `start=YYYY-MM-DD&end=YYYY-MM-DD` to define a date range.

```
$ curl "https://app-usage.YOUR-SYSTEM-DOMAIN/organizations/`cf org YOUR-ORG --guid`/service_usages?start=YYYY-MM-DD&end=YYYY-MM-DD" -k -v -H "authorization: `cf oauth-token`"
{
 "organization_guid": "8d83362f-587a-4148-806b-4407428887b5",
 "period_start": "2016-06-01T00:00:00Z",
 "period_end": "2016-06-13T23:59:59Z",
 "service_usages": [
 {
 "deleted": false,
 "duration_in_seconds": 1377982.52,
 "service_guid": "02802293-b769-44cc-807f-eee331ba9b2b",
 "service_instance_creation": "2016-01-20T18:48:16.000Z",
 "service_instance_deletion": null,
 "service_instance_guid": "b25b4481-19aa-4504-82c9-f303e7e9ed6e",
 "service_instance_name": "something-usage-db",
 "service_instance_type": "managed_service_instance",
 "service_name": "my-mysql-service",
 "service_plan_guid": "70915a70-7311-4f3e-ab0d-4a7dfd3ef2d9",
 "service_plan_name": "20gb",
 "space_guid": "e6445eb3-fdac-4049-bafc-94d1703d5e78",
 "space_name": "outer-space"
 }
]
}
```

## Example: Autogenerated Database Usage Reports

For security compliance and recordkeeping, some PCF customers create Concourse CI/CD pipelines that regularly run JavaScript scripts to do the following:

1. Call the `app-usage` APIs in the above sections of this topic, as well as the [Security Event Logging](#) topic, to collect data about service instances from a PCF foundation. This data includes database instance information such as tile or database type, database creator, and date of creation or deletion.
2. Merge and format the data about services instances into a JSON output file and save it to an S3 bucket.
3. Input the file to another process that does the following:
  - a. Parses the data
  - b. Eliminates duplicate entries from previous reports
  - c. Sends the data to a database compliance system through APIs

## Reporting Instance Usage with Apps Manager

This topic describes how to retrieve app, task, and service instance usage information using Apps Manager.

You can also retrieve app, task, and service instance usage information using the Usage service API, or the [Cloud Foundry API](#) from the Cloud Foundry Command Line Interface (cf CLI). For more information, see the [Monitoring App, Task, and Service Instance Usage](#) topic.

There are two ways to monitor app, task, and service instance usage from Apps Manager. The Accounting Report provides a summarized report, and the Usage Report provides a more detailed view of the data.

### View the Accounting Report

The Accounting Report displays instance usage information for all orgs in your [Pivotal Cloud Foundry](#) (PCF) deployment except the **system** org. The Accounting Report provides a high-level overview of your usage by displaying your monthly count of app, task, and service instances.

Follow the steps below to access the Accounting Report.

1. [Log into the Apps Manager](#) as an admin.
2. From the left navigation of Apps Manager, select **Accounting Report**.
3. Under **App Statistics** and **Service Usage**, view the average and maximum instances in use per month.

**Max Concurrent** displays the largest number of instances in use at a time over the indicated time period. The Accounting Report calculates these values from the `start`, `stop`, and `scale` app usage events in Cloud Controller.

| Accounting Report                                                                 |  |      |         |          |       |
|-----------------------------------------------------------------------------------|--|------|---------|----------|-------|
| Monthly count of App Instances in use for all orgs (not including the system org) |  |      |         |          |       |
| APP STATISTICS                                                                    |  | 2017 | JANUARY | FEBRUARY | MARCH |
| Avg Instances                                                                     |  | 0.0  | 0       | 0        | 0.0   |
| Max Concurrent                                                                    |  | 2.0  | 0       | 0        | 2.0   |
| Apps                                                                              |  | 2.0  | 0       | 0        | 2.0   |
| Tasks                                                                             |  | 0.0  | 0       | 0        | 0.0   |
| Instance Hours                                                                    |  | 0.0  | 0       | 0        | 0.0   |
| Apps                                                                              |  | 0.0  | 0       | 0        | 0.0   |
| Tasks                                                                             |  | 0.0  | 0       | 0        | 0.0   |
| SERVICE USAGE                                                                     |  | 2017 | JANUARY | FEBRUARY | MARCH |
|                                                                                   |  |      |         |          | APRIL |

### View the Usage Report

The Usage Report provides a more granular view of your usage by detailing app, task, and service instance usage information for all spaces in a particular org, excluding the **system** org.

Follow the steps below to access the Usage Report.

1. [Log into the Apps Manager](#) as an admin, or as an account with the **Org Manager** or **Org Auditor** role. For more information about managing roles, see the [Managing User Accounts and Permissions Using the Apps Manager](#) topic.
2. From the dropdown on the left, select the org for which you want to view a usage report.
3. Click **Usage Report** in the upper right.

ORG QUOTA  
My-Org 1.37 GB / 10 GB 14%

[Usage Report](#)

[Spaces \(3\)](#) [Domains \(2\)](#) [Members \(8\)](#) [Settings](#)

The top of the Usage Report displays total App + Task Instance Hours, App + Task Memory, and Service Instance Hours by all spaces in the org.

| Usage Report |                           | Period            | May 1st - Current      | <a href="#">DOWNLOAD ZIP</a> |
|--------------|---------------------------|-------------------|------------------------|------------------------------|
| Org          | App + Task Instance Hours | App + Task Memory | Service Instance Hours |                              |
| My-Org       | 576.7 hrs                 | 111.4 GB hrs      | 486.7 hrs              |                              |

The report provides total usage information for each of your spaces.

| Space       | App + Task Instance Hours | App + Task Memory | Service Instance Hours |
|-------------|---------------------------|-------------------|------------------------|
| development | 74.1 hrs                  | 74.1 GB hrs       | 120.1 hrs              |
| production  | 0.0 hrs                   | 0.0 GB hrs        | 120.0 hrs              |
| staging     | 0.0 hrs                   | 0.0 GB hrs        | 0.0 hrs                |

To display more detailed information about app, task, and service instance usage in a space, click the name of the space for an expanded view.

| Space                              | App + Task Instance Hours | App + Task Memory | Service Instance Hours |
|------------------------------------|---------------------------|-------------------|------------------------|
| development                        | 74.1 hrs                  | 74.1 GB hrs       | 120.1 hrs              |
| <a href="#">Apps(1)</a>            |                           |                   |                        |
| spring-music-sample-app            |                           |                   |                        |
| <a href="#">Services (1)</a>       |                           |                   |                        |
| <a href="#">app-autoscaler (1)</a> |                           |                   |                        |
| my-autoscaler                      | standard                  | 120.1             |                        |

## Providing a Certificate for Your TLS Termination Point

This topic describes how to configure Transport Layer Security (TLS) termination for HTTP traffic in Pivotal Application Service (PAS) with a TLS certificate, as part of the process of configuring PAS for deployment.

### Configure TLS Termination

When you deploy PCF, you must configure the TLS termination for HTTP traffic in your PAS configuration. You can terminate TLS at all of the following points:

- Load Balancer
- Load Balancer and Gorouter
- Gorouter

Follow the guidance in [Securing Traffic into Cloud Foundry](#) to choose and configure the TLS termination option for your deployment.

 **Note:** If you are using HAProxy in a PCF deployment, you can choose to terminate SSL/TLS at HAProxy in addition to any of the SSL/TLS termination options above. For more information, see [Configuring SSL/TLS Termination at HAProxy](#).

### Obtain TLS Certificates

To secure traffic into PCF, you must obtain at least one TLS certificate. See [Certificate Requirements](#) for general certificate requirements for deploying PCF.

See the following sections for additional IaaS-specific certificate requirements:

- AWS: [Certificate Requirements on AWS](#)
- GCP: [Certificate Requirements on GCP](#)

### Creating a Wildcard Certificate for PCF Deployments

This section describes how to create or generate a certificate for your PAS environment. If you are deploying to a production environment, you should obtain a certificate from a trusted authority (CA).

For internal development or testing environments, you have two options for creating a required TLS certificates.

- You can create a self-signed certificate, or
- You can have [PAS generate the certificate](#) for you.

To create a certificate, you can use a wide variety of tools including OpenSSL, Java's keytool, Adobe Reader, and Apple's Keychain to generate a Certificate Signing Request (CSR).

In either case for either self-signed or trusted single certificates, apply the following rules when creating the CSR:

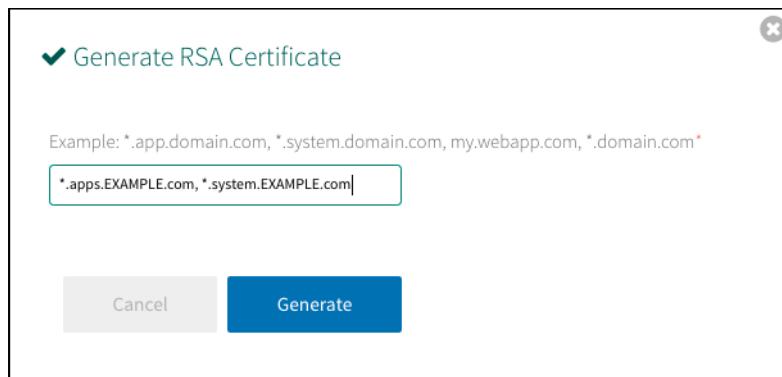
- Specify your registered wildcard domain as the `Common Name`. For example, `*.yourdomain.com`.
- If you are using a split domain setup that separates the domains for `apps` and `system` components (recommended), then enter the following values in the `Subject Alternative Name` of the certificate:
  - `*.apps.yourdomain.com`
  - `*.system.yourdomain.com`
  - `*.login.system.yourdomain.com`
  - `*.uaa.system.yourdomain.com`
- If you are using a single domain setup, then use the following values as the `Subject Alternative Name` of the certificate:
  - `*.login.system.yourdomain.com`
  - `*.uaa.system.yourdomain.com`

**Note:** TLS certificates generated for wildcard DNS records only work for a single domain name component or component fragment. For example, a certificate generated for \*.EXAMPLE.com does not work for \*.apps.EXAMPLE.com and \*.system.EXAMPLE.com. The certificate must have both \*.apps.EXAMPLE.com and \*.system.EXAMPLE.com attributed to it.

## Generating a RSA Certificate in PAS

1. Navigate to the Ops Manager Installation Dashboard.
2. Click the **Pivotal Application Service** tile in the Installation Dashboard.
3. Click **Networking**.
4. Click **Generate RSA Certificate** to populate the **Certificate and Private Key for HAProxy and Router** fields with RSA certificate and private key information.
5. If you are using a split domain setup that separates the domains for **apps** and **system** components (recommended), then enter the following domains for the certificate:
  - o \*.yourdomain.com
  - o \*.apps.yourdomain.com
  - o \*.system.yourdomain.com
  - o \*.login.system.yourdomain.com
  - o \*.uaa.system.yourdomain.com

For example, \*.example.com , \*.apps.example.com , \*.system.example.com , \*.login.system.example.com , \*.uaa.system.example.com



## Enabling Volume Services

This topic describes how Pivotal Cloud Foundry (PCF) operators can deploy NFS or SMB volume services.

### Overview

A volume service gives apps access to a persistent filesystem, such as NFS or SMB. By performing the procedures in this topic, operators can add a volume service to the [Marketplace](#) that provides an NFS filesystem or an SMB share.

Developers can then use the [Cloud Foundry Command Line Interface](#) (cf CLI) to create service instances of the volume service and bind them to their apps. For more information, see the [Using an External File System \(Volume Services\)](#) topic.

 **Note:** You must have an NFS or SMB server running to enable NFS or SMB volume services. If you want to use NFS or SMB and do not currently have a server available, you can do one of the following:

- Deploy the test NFS or SMB server bundled with the volume release or you can enable NFS or SMB volume services with the NFS or SMB Broker Errand for Pivotal Application Service (PAS). For more information, see [Adding Volume Services to Your Deployment](#) in the Cloud Foundry documentation.
- Enable this errand during PAS configuration.

### Enable NFS Volume Services

To enable NFS volume services in PCF, do the following:

1. Navigate to the Ops Manager Installation Dashboard.
2. Click the **Pivotal Application Service** tile.
3. Click **Application Containers**.
4. Under **Enabling NFS volume services**, select **Enable**.

 **Note:** In a clean install, NFS volume services are enabled by default. In an upgrade, NFS volume services match the setting of the previous deployment.

Enabling NFSv3 volume services will allow application developers to bind existing NFS volumes to their applications for shared file access. \*

Enable

LDAP Service Account User

LDAP Service Account Password

?

LDAP Server Host

LDAP Server Port

LDAP User Search Base

LDAP Server CA Cert

Disable

5. (Optional) To configure LDAP for NFS volume services, perform the following steps:

**Note:** If you already use an LDAP server with your network-attached storage (NAS) file server, enter its information below. This ensures that the identities known to the file server match those checked by the PCF NFS driver.

- For **LDAP Service Account User**, enter either the full domain name (DN) for the service account or the username of the service account that manages volume services, depending on how your LDAP server is configured. This value needs to be exactly what you would enter when binding the account to your LDAP server.
- For **LDAP Service Account Password**, enter the password for the service account.
- For **LDAP Server Host**, enter the hostname or IP address of the LDAP server.
- For **LDAP Server Port**, enter the LDAP server port number. If you do not specify a port number, Ops Manager uses 389.
- For **LDAP User Search Base**, enter the location in the LDAP directory tree from which any LDAP User search begins. The typical LDAP Search Base matches your domain name.  
For example, a domain named `cloud.example.com` typically uses the following LDAP User Search Base: `ou=Users,dc=example,dc=com`.
- For **LDAP Server CA Cert**, you can optionally enter a certificate if your LDAP server supports TLS and you want to enable TLS connections from the NFS driver to your LDAP server. Paste in the root certificate from your CA certificate or your self-signed certificate.

6. Click **Save**.

7. Return to the Ops Manager Installation Dashboard, click **Review Pending Changes**, and click **Apply Changes** to redeploy.

8. Using the cf CLI, enable access to the service:

```
$ cf enable-service-access nfs
```

To limit access to a specific org, use the `-o` flag, followed by the name of the org where you want to enable access. For more information, see the [Access Control](#) topic.

9. (Optional) Enable access to the `nfs-legacy` service. See [NFS Volume Service](#) for details about the differences between the two `nfs` services.

```
$ cf enable-service-access nfs-legacy
```

After completing these steps, developers can use the cf CLI to create service instances of the `nfs` service and bind them to their apps.

## NFS Security

You can use ASGs and LDAP to secure your NFS server against traffic apps running on PCF:

- **Application Security Groups (ASGs)**—Prevent apps from sending traffic directly to your NFS ports. Apps should never need to use NFS ports directly. Pivotal recommends defining an ASG that blocks direct access to your NFS server IP, especially ports 111 and 2049. For more information on setting up ASGs, see [Application Security Groups](#).
- **LDAP**—In addition to ASGs, secures the NFS volume service so that app developers cannot bind to the service using an arbitrary UID. App developers also cannot gain access to sensitive data. With LDAP support enabled, app developers must provide credentials for any user they wish to bind as.

The Diego cells running on PCF must be able to reach your LDAP server on the port you use for connections, which are typically `389` or `636`. You cannot limit which Diego cells have access to your NFS or LDAP servers.

## Enable SMB Volume Services

To enable SMB volume services in PCF, do the following:

1. Navigate to the Ops Manager Installation Dashboard.
  2. Click the **Pivotal Application Service** tile.
  3. Select the **Application Containers** pane.
  4. Check the box next to **Enable SMB volume services**.
- Enable SMB volume services** Enabling SMB volume services will allow application developers to bind existing SMB shares to their applications for shared file access.
5. Click **Save**.
  6. Select the **Errands** pane.
  7. Set the **SMB Broker Errand** to **On**.
  8. Click **Save**.
  9. Return to the Ops Manager Installation Dashboard, click **Review Pending Changes**, and click **Apply Changes** to redeploy.
  10. Using the cf CLI, enable access to the service:

```
$ cf enable-service-access smb
```

To limit access to a specific org, use the `-o` flag, followed by the name of the org where you want to enable access. For more information, see the [Access Control](#) topic.

After completing these steps, developers can use the cf CLI to create service instances of the `smb` service and bind them to their apps.

## Rotating Runtime CredHub Encryption Keys

This topic discusses how to rotate runtime CredHub encryption keys. Encryption keys are values that CredHub uses to obscure stored secrets. When an operator marks an additional key as primary, CredHub can rotate in that additional key as the encryption key.

During this credential rotation process, the initial encryption key is used to access the hidden value, That value is then stored again by the additional encryption key.

**A** **warning:** If you remove an encryption key and click **Apply Changes** before the rotation completes, the deployment breaks. If this happens, you can no longer access data stored with the deleted key.

### Rotate PAS Encryption Keys

To rotate PAS encryption keys, do the following:

1. Navigate to the **Ops Manager Installation Dashboard**.
2. Click the **Pivotal Application Service** tile.
3. Select the **CredHub** tab.

The screenshot shows a form titled "Encryption Keys". At the top right is a grey "Add" button. Below it is a "Name \*" field with a placeholder "Name of the encryption key." and a "Key \*" field with a placeholder "Secret". To the right of the "Key" field is a small icon. Below these fields is a "Primary" checkbox. The entire form is contained within a light grey box with a thin border.

4. In the **Encryption Keys** section, click **Add**.
5. For **Name**, enter the name of your new encryption key.
6. For **Key**, enter your new encryption key.
7. Select the **Primary** check box.
8. Click **Save**.
9. Navigate to the **Ops Manager Installation Dashboard**.
10. Click **Review Pending Changes**, then **Apply Changes**.

### Verify PAS Encryption Key Rotation

Follow the steps below to verify that the rotation completes.

1. Click the **Pivotal Application Service** tile.
2. Select the **Status** tab.
3. Within the **CredHub** job, locate **Index 0**.

| JOB                           | INDEX | IPS       | AZ            | CID                                     | LOAD AVG15 | CPU  | MEMORY | SWAP | SYSTEM DISK | Ephem. Disk | Pers. Disk | Logs |
|-------------------------------|-------|-----------|---------------|-----------------------------------------|------------|------|--------|------|-------------|-------------|------------|------|
| Consul                        | 0     | 10.0.4.5  | us-central1-a | vm-dca0a175-3a56-4cb0-6b64-2d3f0a39ddc7 | 0.00       | 1.2% | 14%    | 0%   | 40%         | 3%          | 3%         |      |
| NATS                          | 0     | 10.0.4.6  | us-central1-a | vm-a5b7a6e3-d80d-44d2-54cc-a160d3f66916 | 0.00       | 1.1% | 14%    | 0%   | 40%         | 3%          | N/A        |      |
| File Storage                  | 0     | 10.0.4.7  | us-central1-a | vm-c4aa5642-54f9-436e-5c31-03b2607de077 | 0.00       | 1.0% | 4%     | 0%   | 40%         | 7%          | 3%         |      |
| MySQL Proxy                   | 0     | 10.0.4.8  | us-central1-a | vm-c6df54cc-128c-440b-4835-e2826b9502be | 0.00       | 0.3% | 15%    | 0%   | 40%         | 4%          | N/A        |      |
| MySQL Server                  | 0     | 10.0.4.9  | us-central1-a | vm-5ec3aa62-f066-42e0-60ec-ce4464cf628  | 0.08       | 1.1% | 24%    | 0%   | 40%         | 3%          | 5%         |      |
| Backup Prepare Node           | 0     | 10.0.4.10 | us-central1-a | vm-55e93a88-ce00-4519-4ccb-39b12bd82495 | 0.00       | 0.3% | 19%    | 0%   | 40%         | 20%         | 0%         |      |
| Diego BBS                     | 0     | 10.0.4.11 | us-central1-a | vm-c95a4858-b9fb-460e-4693-1fdf13be88d  | 0.03       | 1.1% | 20%    | 0%   | 40%         | 5%          | N/A        |      |
| UAA                           | 0     | 10.0.4.12 | us-central1-a | vm-9f5c1fdc-704e-484c-7e89-58c95a19a319 | 0.02       | 0.6% | 20%    | 0%   | 40%         | 3%          | N/A        |      |
| Cloud Controller              | 0     | 10.0.4.13 | us-central1-a | vm-6ecf788-a390-45ed-4990-9cbd3f6aeee6  | 0.07       | 1.0% | 27%    | 0%   | 40%         | 16%         | N/A        |      |
| HAProxy                       | 0     | 10.0.4.14 | us-central1-a | vm-9466dc03-d2bc-48f0-401b-d9082e0f8778 | 0.00       | 0.5% | 14%    | 0%   | 40%         | 3%          | 40%        |      |
| Router                        | 0     | 10.0.4.15 | us-central1-a | vm-32ba827f-564b-4462-7b20-46a658a8e8df | 0.00       | 1.2% | 16%    | 0%   | 40%         | 4%          | N/A        |      |
| MySQL Monitor                 | 0     | 10.0.4.16 | us-central1-a | vm-98c9c06c-e381-455f-517c-62e9ad9d70d1 | 0.00       | 0.2% | 11%    | 0%   | 40%         | 2%          | N/A        |      |
| Clock Global                  | 0     | 10.0.4.17 | us-central1-a | vm-65a93076-941e-4b4c-50a2-e067bfe2d0ed | 0.00       | 0.3% | 41%    | 1%   | 40%         | 86%         | N/A        |      |
| Cloud Controller Worker       | 0     | 10.0.4.18 | us-central1-a | vm-de45f832-e290-4e38-7692-236fa5b417d5 | 0.00       | 1.1% | 48%    | 1%   | 40%         | 9%          | N/A        |      |
| Diego Brain                   | 0     | 10.0.4.19 | us-central1-a | vm-247bcf3-9241-49df-57fd-7c3d8bb9fef   | 0.00       | 0.4% | 99%    | 0%   | 41%         | 6%          | 0%         |      |
| Diego Cell                    | 0     | 10.0.4.20 | us-central1-a | vm-99761423-b711-4871-7aa9-848537f5096e | 0.07       | 0.6% | 12%    | 0%   | 40%         | 8%          | N/A        |      |
| Loggregator Trafficcontroller | 0     | 10.0.4.21 | us-central1-a | vm-37c47f42-f226-4728-6acc-d4351a7d06cc | 0.00       | 1.4% | 18%    | 0%   | 40%         | 4%          | N/A        |      |
| Syslog Adapter                | 0     | 10.0.4.22 | us-central1-a | vm-9c577460-1a29-4741-7a0b-e70d6d93b286 | 0.00       | 1.1% | 14%    | 0%   | 40%         | 3%          | N/A        |      |
| Syslog Scheduler              | 0     | 10.0.4.23 | us-central1-a | vm-43a14a07-aaae-466d-6839-23767c3ce141 | 0.00       | 0.4% | 14%    | 0%   | 40%         | 3%          | N/A        |      |
| Doppler Server                | 0     | 10.0.4.24 | us-central1-a | vm-66359c07-0ef4-4a42-542c-2e7ce9e4be5  | 0.05       | 1.5% | 19%    | 0%   | 40%         | 4%          | N/A        |      |
| <b>CredHub</b>                | 0     | 10.0.4.25 | us-central1-a | vm-84b31970-374a-4e0c-411c-3977b6a20ff1 | 0.00       | 0.2% | 12%    | 0%   | 40%         | 7%          | N/A        |      |
|                               | 1     | 10.0.4.26 | us-central1-b | vm-b67b89d6-0415-4021-7950-48c6c3aa2b14 | 0.00       | 0.2% | 12%    | 0%   | 40%         | 7%          | N/A        |      |

4. Within the **Logs** column, click the correlating download icon.
5. Select the **Logs** tab.
6. Click the corresponding link to the retrieve the downloaded log file.
7. Unzip the log file.
8. Unzip the larger of the two nested directories.
9. Ops Manager generates a compressed file for each CredHub VM that exists on your deployment. Unzip each of these compressed files.
10. Open the `credhub` directory.
11. Open the `credhub.log` file. If the PAS credential rotation completed successfully, the CredHub log contains the following string: `Successfully rotated NUMBER-OF-CREDENTIALS items`
12. Remove the old encryption key.
13. Click the trashcan icon that corresponds to the old encryption key.
14. Click **Save**.
15. Navigate to the **Ops Manager Installation Dashboard**.
16. Click **Review Pending Changes**, then **Apply Changes**.

## Administrating PAS

This documentation describes administrating Pivotal Application Service (PAS). This includes procedures specific to managing the runtime, such as creating user accounts, modifying quota plans, and managing isolation segments.

See the following topics:

- [Managing the Runtime](#)
- [User Accounts and Communications](#)
- [Routing](#)
- [Isolation Segments](#)

## Managing the Runtime

The following topics provide information about managing Pivotal Application Service (PAS):

- [Stopping and Starting Virtual Machines](#)
- [Creating and Modifying Quota Plans](#)
- [Using Feature Flags](#)
- [Examining GrootFS Disk Usage](#)
- [Using Metadata](#)
- [Managing Custom Buildpacks](#)
- [Using Docker in Cloud Foundry](#)

## Stopping and Starting Virtual Machines

This topic assumes you are using [BOSH CLI version 2](#).

This topic describes stopping and starting the component virtual machines (VMs) that make up a Pivotal Cloud Foundry (PCF) deployment.

In some cases you may want to stop all your VMs (for example, power down your deployment) or start all of your PAS VMs (for example, recover from a power outage.) You can stop or start all PAS VMs with a single `bosh` command.

If you want to shut down or start up a single VM in your deployment, you can use the [manual process](#) described on this page.

This procedure uses the BOSH Command Line Interface (BOSH CLI). See [Advanced Troubleshooting with the BOSH CLI](#) for more information about using this tool.

### Stopping and Starting All PAS VMs

This section describes how to stop and start all the VMs in your deployment.

#### Stopping All PAS VMs

To shut down all the VMs in your deployment, perform the following steps:

1. Scale down the following jobs to one instance:
  - o `consul_server`
  - o `mysql`
2. Run the following command for each of the deployments listed in the previous step:

```
bosh -e MY-ENV -d MY-DEPLOYMENT stop --hard
```

Replace the text above with the following:

- o `MY-ENV` : the alias you set for the BOSH Director.
- o `MY-DEPLOYMENT` : the name of your deployment.

For example:

```
$ bosh -e prod -d mysql stop --hard
```

This command stops all VMs in the specified deployment. The `--hard` flag instructs BOSH to delete the VMs but retain any persistent disks.

#### Starting All PAS VMs

Perform the following steps to start all the VMs in your deployment:

1. Select the product deployment for the VMs you want to shut down. You can run the following command to locate CF deployment manifests:

```
$ find /var/tempest/workspaces/default/deployments -name cf-*.*.yml
```

2. Run the following command:

```
bosh -e MY-ENV -d MY-DEPLOYMENT start
```

Replace the text above with the following:

- o `MY-ENV` : the alias you set for the BOSH Director.
- o `MY-DEPLOYMENT` : the name of your deployment.

For example:

```
$ bosh -e prod -d mysql start
```

This command starts all VMs in the specified deployment.

3. If you require [high availability](#) in your deployment, scale up all instance groups to the original or desired counts.

## Stopping and Starting Individual PAS VMs

This section describes how to stop and start individual VMs in your deployment.

### Find the Names of Your PAS Virtual Machines

You need the full names of the VMs to stop and start them using the BOSH CLI. To find full names for the VMs running each component, run

```
bosh -e MY-ENV instances , replacing [MY-ENV] with the alias you set for your BOSH Director. To filter the list of instances by deployment, run
```

```
bosh -e MY-ENV -d MY-DEPLOYMENT instances .
```

For example:

```
$ bosh -e prod -d mysql instances
...
Deployment 'mysql'

Instance Process State AZ IPs
mysql/0123-abcd-4567ef89 running - 10.244.0.6
mysql/abcd-0123-ef4567ab running - 10.244.0.2

2 instances
...
```

You can see the full name of each VM in the `Instance` column of the terminal output. Each full name includes:

- A prefix indicating the component function of the VM.
- An identifier string specific to the VM.

For any component, you can look for its prefix in the `bosh instances` output to find the full name of the VM or VMs that run it.

### Stopping an Individual PAS VM

To stop a job, run the following command for the component in your PAS deployment, replacing `[MY-ENV]` with the alias you set for your BOSH Director and `[MY-DEPLOYMENT]` with the name of the deployment:

```
bosh -e MY-ENV -d MY-DEPLOYMENT stop VM-NAME
```

To delete the instance that contains the job, run the following command for the component in your PAS deployment:

```
bosh -e MY-ENV -d MY-DEPLOYMENT stop VM-NAME --hard
```

Use the full name of the component VM as listed in your `bosh instances` [terminal output](#) without the unique identifier string.

For example, the following command stops the Loggregator Traffic Controller job:

```
$ bosh -e prod -d loggregator stop loggregator_trafficcontroller
```

To stop a specific instance of a job, include the identifier string at the end of its full name.

For example, the following command stops the Loggregator Traffic Controller job on only one Diego cell instance:

```
$ bosh -e prod -d loggregator stop loggregator_trafficcontroller/0123-abcd-4567ef89
```

To delete the VM, include `--hard` at the end of the command. This command does not delete persistent disks.

For example, the following command deletes a specific Loggregator Traffic Controller instance:

```
$ bosh -e prod -d loggregator stop loggregator_trafficcontroller/0123-abcd-4567ef89 --hard
```

## Starting an Individual PAS VM

Run the following command for the component in your PAS deployment you wish to start, replacing `MY-ENV` with the alias you set for your BOSH Director and `MY-DEPLOYMENT` with the name of the deployment. Use the full name of the component VM as listed in your `bosh vms` [terminal output](#) without the unique identifier string.

```
bosh -e MY-ENV -d MY-DEPLOYMENT start VM-NAME
```

The following example command starts the Loggregator Traffic Controller VM:

```
$ bosh -e prod -d loggregator start loggregator_trafficcontroller
```

To start a specific instance of a VM, include the identifier string at the end of its full name.

For example, the following command starts the Loggregator Traffic Controller job on one Diego cell instance:

```
$ bosh -e prod -d loggregator start loggregator_trafficcontroller/0123-abcd-4567ef89
```

## Creating and Modifying Quota Plans

Quota plans are named sets of memory, service, and instance usage quotas. For example, one quota plan might allow up to 10 services, 10 routes, and 2 GB of RAM, while another might offer 100 services, 100 routes, and 10 GB of RAM. Quota plans have user-friendly names, but are referenced in Cloud Foundry (CF) internal systems by unique GUIDs.

Quota plans are not directly associated with user accounts. Instead, every org has a list of available quota plans, and the account admin assigns a specific quota plan from the list to the org. Everyone in the org shares the quotas described by the plan. There is no limit to the number of defined quota plans an account can have, but only one plan can be assigned at a time.

You must set a quota plan for an [org](#), but you can choose whether to set a [space](#) quota.

## Org Quota Plan Attributes

| Name                       | Description                                                                                                                                                                                                              | Valid Values                                                                                                 | Example Value |
|----------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------|---------------|
| name                       | The name you use to identify the plan                                                                                                                                                                                    | A sequence of letters, digits, and underscore characters. Quota plan names within an account must be unique. | silver_quota  |
| memory_limit               | Maximum memory usage allowed                                                                                                                                                                                             | An integer and a unit of measurement like M, MB, G, or GB                                                    | 2048M         |
| app_instance_limit         | Maximum app instances allowed.<br><br>Stopped apps do not count toward this instance limit. Crashed apps count toward the limit because their desired state is <a href="#">starting</a> .                                | An integer                                                                                                   | 25            |
| non_basic_services_allowed | Determines whether users can provision instances of non-free service plans. Does not control plan visibility. When false, non-free service plans may be visible in the marketplace but instances can not be provisioned. | <code>true</code> or <code>false</code>                                                                      | true          |
| total_routes               | Maximum routes allowed                                                                                                                                                                                                   | An integer                                                                                                   | 500           |
| total_reserved_route_ports | Maximum routes with reserved ports                                                                                                                                                                                       | An integer not greater than total_routes                                                                     | 60            |
| total_services             | Maximum services allowed                                                                                                                                                                                                 | An integer                                                                                                   | 25            |
| trial_db_allowed           | Legacy Field. Value can be ignored.                                                                                                                                                                                      | <code>true</code> or <code>false</code>                                                                      | true          |

## Default Quota Plan for an Org

Cloud Foundry installs with a quota plan named `default` with the following values:

- Memory Limit: `10240 MB`
- Total Routes: `1000`
- Total Services: `100`
- Non-basic Services Allowed: `True`
- Trial DB Allowed: `True`

## Create a New Quota Plan for an Org

 **Note:** The org manager sets and manages quotas. See the [Orgs, Spaces, Roles, and Permissions topic](#) for more information.

You must set an org quota. You can create a new quota plan for org with `cf create-quota`.

## Use cf create-quota

In a terminal window, run the following command. Replace the placeholder attributes with the values for this quota plan:

```
cf create-quota QUOTA [-m TOTAL-MEMORY] [-i INSTANCE-MEMORY] [-r ROUTES] [-s SERVICE-INSTANCES] [--allow-paid-service-plans]
```

This command accepts the following flags:

- `-m` : Total amount of memory
- `-i` : Maximum amount of memory an application instance can have (`-1` represents an unlimited amount)
- `-r` : Total number of routes
- `-s` : Total number of service instances
- `--allow-paid-service-plans` : Can provision instances of paid service plans

Example:

```
$ cf create-quota small -m 2048M -i 1024M -r 10 -s 10 --allow-paid-service-plans
```

## Modify an Existing Quota Plan for an Org

### Use cf update-quota

1. Run `cf quotas` to find the names of all quota definitions available to your org. Record the name of the quota plan to be modified.

```
$ cf quotas
Getting quotas as admin@example.com...
OK

name total memory limit instance memory limit routes service instances paid service plans
free 0 0 1000 0 disallowed
paid 10G 0 1000 -1 allowed
small 2G 0 10 10 allowed
trial 2G 0 1000 10 disallowed
```

2. Run the following command, replacing QUOTA with the name of your quota.

```
cf update-quota QUOTA [-i INSTANCE-MEMORY] [-m MEMORY] [-n NEW-NAME] [-r ROUTES] [-s SERVICE-INSTANCES] [--allow-paid-service-plans] [--disallow-paid-service-plans]
```

This command accepts the following flags:

- `-i` : Maximum amount of memory an application instance can have (`-1` represents an unlimited amount)
- `-m` : Total amount of memory a space can have
- `-n` : New name
- `-r` : Total number of routes
- `-s` : Total number of service instances
- `--allow-paid-service-plans` : Can provision instances of paid service plans
- `--disallow-paid-service-plans` : Can not provision instances of paid service plans

Example:

```
$ cf update-quota small -i 2048M -m 4096M -n medium -r 20 -s 20 --allow-paid-service-plans
```

## Create and Modify Quota Plans for a Space

For each org, Org Managers create and modify quota plans for spaces in the org. If an Org Manager allocates a space quota, CF verifies that resources do not exceed the allocated space limit. For example, when a Space Developer deploys an application, CF first checks the memory allocation at the space level, then at the org level.

Perform the following procedures to create and modify quota plans for individual spaces within an org.

## Create a New Quota Plan for a Space

In a terminal window, run the following command to create a quota for a space. Replace the placeholder attributes with the values for this quota plan:

```
cf create-space-quota QUOTA [-i INSTANCE-MEMORY] [-m MEMORY] [-r ROUTES] [-s SERVICE-INSTANCES] [--allow-paid-service-plans]
```

Example:

```
$ cf create-space-quota big -i 1024M -m 4096M -r 20 -s 20 --allow-paid-service-plans
```

## Modify a Quota Plan for a Space

Run `cf space-quotas` to find the names of all space quota available to your org. Record the name of the quota plan to be modified.

```
$ cf space-quotas
Getting quotas as admin@example.com...
OK

name total memory limit instance memory limit routes service instances paid service plans
big 2G unlimited 0 10 allowed
trial 2G 0 0 10 allowed
```

To modify that quota, use the `update-space-quota` command. Replace the placeholder attributes with the values for this quota plan.

```
cf update-space-quota SPACE-QUOTA-NAME [-i MAX-INSTANCE-MEMORY] [-m MEMORY] [-n NEW-NAME] [-r ROUTES] [-s SERVICES] [--allow-paid-service-plans] [--disallow-paid-service-plans]
```

Example:

```
$ cf update-space-quota big -i 20 -m 4096M -n bigger -r 20 -s 20 --allow-paid-service-plans
```

## Run cf help

For more information regarding quotas, run `cf help` to view a list and brief description of all cf CLI commands. Scroll to view org and space quotas usage and information.

```
$ cf help
...
ORG ADMIN:
quotas List available usage quotas
quota Show quota info
set-quota Assign a quota to an org

create-quota Define a new resource quota
delete-quota Delete a quota
update-quota Update an existing resource quota

share-private-domain Share a private domain with an org
unshare-private-domain Unshare a private domain with an org

SPACE ADMIN:
space-quotas List available space resource quotas
space-quota Show space quota info
create-space-quota Define a new space resource quota
update-space-quota update an existing space quota
delete-space-quota Delete a space quota definition and unassign the space quota from all spaces
set-space-quota Assign a space quota definition to a space
unset-space-quota Unassign a quota from a space
```

## Using Feature Flags

This topic describes how Cloud Foundry (CF) administrators can set feature flags using the Cloud Foundry Command Line Interface (cf CLI) to enable or disable the features available to users.

## View and Edit Feature Flags

To perform the following procedures, you must be logged in to your deployment as an administrator using the cf CLI.

1. Use the `cf feature-flags` command to list the feature flags:

```
$ cf feature-flags

Features State
user_org_creation disabled
private_domain_creation enabled
app_bits_upload enabled
app_scaling enabled
route_creation enabled
service_instance_creation enabled
diego_docker disabled
set_roles_by_username enabled
unset_roles_by_username enabled
task_creation enabled
env_var_visibility enabled
spaceScoped_private_broker_creation enabled
space_developer_env_var_visibility enabled
service_instance_sharing disabled
hide_marketplace_from_unauthenticated_users disabled
```

For descriptions of the features enabled by each feature flag, see the [Feature Flags](#) section below.

2. To view the status of a feature flag, use the `cf feature-flag FEATURE-FLAG-NAME` command:

```
$ cf feature-flag user_org_creation

Retrieving status of user_org_creation as admin...
OK

Features State
user_org_creation disabled
```

3. To enable a feature flag, use the `cf enable-feature-flag FEATURE-FLAG-NAME` command:

```
$ cf enable-feature-flag user_org_creation
```

4. To disable a feature flag, use the `cf disable-feature-flag FEATURE-FLAG-NAME` command:

```
$ cf disable-feature-flag user_org_creation
```

## Feature Flags

Only administrators can set feature flags. All flags are enabled by default except `user_org_creation` and `diego_docker`. When disabled, these features are only available to administrators.

The following list provides descriptions of the features enabled or disabled by each flag, and the minimum Cloud Controller API (CC API) version necessary to use the feature. To determine your CC API version, follow the instructions in [Identifying the API Endpoint for your PAS Instance](#).

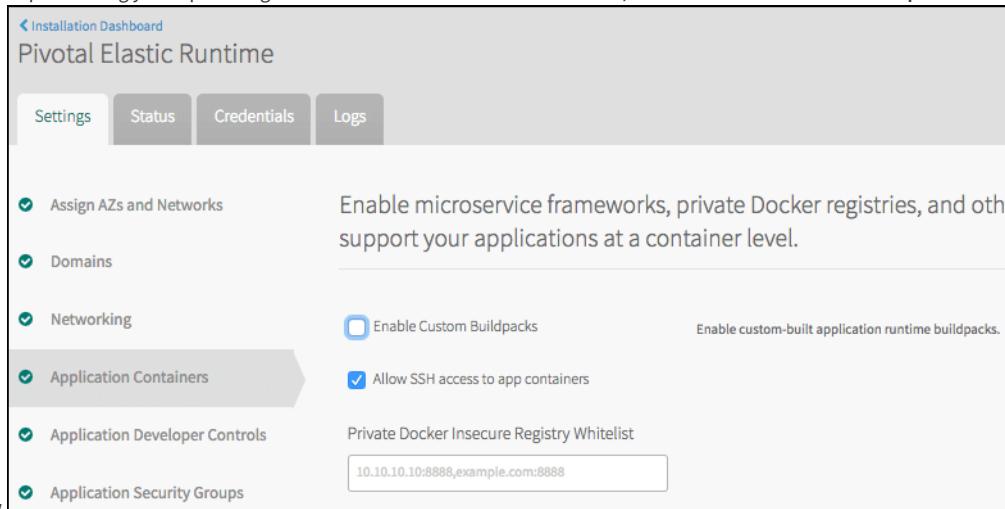
- `user_org_creation`: Any user can create an organization. If enabled, this flag activates the [Create a New Org](#) link in the dropdown menu of the left navigation menu in Apps Manager. Minimum CC API version: 2.12.
- `private_domain_creation`: An Org Manager can create private domains for that organization. Minimum CC API version: 2.12.
- `app_bits_upload`: Space Developers can upload app bits. Minimum CC API version: 2.12.
- `app_scaling`: Space Developers can perform scaling operations (i.e. change memory, disk, or instances). Minimum CC API version: 2.12.

- `route_creation` : Space Developers can create routes in a space. Minimum CC API version: 2.12.
- `service_instance_creation` : Space Developers can create service instances in a space. Minimum CC API version: 2.12.
- `diego_docker` : Space Developers can push Docker apps. Minimum CC API version 2.33.
- `set_roles_by_username` : Org Managers and Space Managers can add roles by username. Minimum CC API version: 2.37.
- `unset_roles_by_username` : Org Managers and Space Managers can remove roles by username. Minimum CC API version: 2.37.
- `task_creation` : Space Developers can create tasks on their application. Minimum CC API version: 2.47.
- `env_var_visibility` : All users can view environment variables. Minimum CC API version: 2.58.
- `spaceScopedPrivateBrokerCreation` : Space Developers can create space-scoped private service brokers. Minimum CC API version: 2.58.
- `spaceDeveloperEnvVarVisibility` : Space Developers can view their v2 environment variables. Org Managers and Space Managers can view their v3 environment variables. Minimum CC API version: 2.58.
- `serviceInstanceSharing` : Space Developers can share service instances between two spaces (across orgs) in which they have the Space Developer role.
- `hideMarketplaceFromUnauthenticatedUsers` : Do not allow unauthenticated users to see the service offerings available in the marketplace.

For more information about feature flag commands, see the [Feature Flags](#) section of the [Cloud Foundry API documentation](#).

## Disabling Custom Buildpacks

You can disable custom buildpacks using your Ops Manager PAS tile. From the Cloud Controller tab, check the **Disable Custom Buildpacks** checkbox, as



shown in the image below.

By default, the cf CLI gives developers the option of using a custom buildpack when they deploy apps to PAS. To do so, they use the `-b` option to provide a custom buildpack URL with the `cf push` command. The **Disable Custom Buildpacks** checkbox prevents the `-b` option from being used with external buildpack URLs.

For more information about custom buildpacks, refer to the [buildpacks](#) section of the PCF documentation.

## Examining GrootFS Disk Usage

This topic describes how to analyze GrootFS disk space usage in Pivotal Application Service (PAS).

You run the commands in this topic as root on any BOSH-deployed VM that hosts the Garden job in a PAS deployment.

 **Note:** This topic provides different commands depending on whether you are using privileged or unprivileged containers in your deployment. By default, all deployments use unprivileged containers. For more information about these container types, see [Container Security](#).

For more information about the GrootFS concepts, see [GrootFS Disk Usage](#).

## Reconcile Container Disk Usage and Host Disk Usage

To reconcile disk allocations for containers with the actual disk usage on the host VM, you need to understand how GrootFS uses its disk.

Command line tools such as `du` and `df` can provide misleading information because of the way container file systems work.

For example, the following situations can occur:

| If...                                                                                                                                                                               | Then...                                                                  |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------|
| the Diego cell rep appears to be out of disk capacity, but the actual disk usage on the Garden host is low                                                                          | Diego does not schedule containers on the cell.                          |
| the Diego cell rep appears to have available disk capacity, but the combined space used by containers and system components prevents Diego from allocating the remaining disk space | Diego continues to place containers on the cell, but they fail to start. |

## About Container Disk Usage

On disk, the read-write layer for each container can be found at `/var/vcap/data/grootfs/store/unprivileged/images/CONTAINER-ID/diff`.

When GrootFS calls on the built-in XFS quota tooling to get disk usage for a container, it takes into account data written to that directory and not the data in the read-only volumes.

Running `grootfs stats` returns the following values:

- `total_bytes_used` : This is the disk usage of the container **including** the rootfs image volumes.
- `exclusive_bytes_used` : This is the disk usage of the container **not including** the rootfs image volumes.

## Retrieve Disk Usage Stats for a Single Container

To obtain the disk usage stats of a single container, perform the following steps:

1. In your PAS deployment, use `bosh ssh` to connect to the BOSH-deployed VM running the Garden job.
2. On the VM, run the following command to look up the container ID:

```
ls /var/vcap/data/garden/depot/
```

The command above returns a container ID in the following format:

```
55afbfb5-5cbf-49c6-4461-f803
```

3. Based on the container type used in your deployment, run one of the following commands on the VM:

- For unprivileged containers, run the following command:

```
/var/vcap/packages/grootfs/bin/grootfs --config \
/var/vcap/jobs/garden/config/grootfs_config.yml stats CONTAINER-ID
```

Where `CONTAINER-ID` corresponds to the container ID you obtained in step 2.

For example:

```
$ /var/vcap/packages/grootfs/bin/grootfs --config \
/var/vcap/jobs/garden/config/grootfs_config.yml stats 55afb65-5cbf-49c6-4461-f803
```

- For privileged containers, run the following command:

```
/var/vcap/packages/grootfs/bin/grootfs --config \
/var/vcap/jobs/garden/config/privileged_grootfs_config.yml stats CONTAINER-ID
```

Where `CONTAINER-ID` corresponds to the container ID you obtained in step 2.

For example:

```
$ /var/vcap/packages/grootfs/bin/grootfs --config \
/var/vcap/jobs/garden/config/privileged_grootfs_config.yml stats 55afb65-5cbf-49c6-4461-f803
```

The commands above return output in the following format:

```
{"disk_usage": {"total_bytes_used": 23448093, "exclusive_bytes_used": 8192}}
```

## Retrieve Exclusive Disk Usage Stats for All Running Containers

To check the disk usage of all running containers, perform the following steps:

- In your PAS deployment, use `bosh ssh` to connect to the BOSH-deployed VM running the Garden job.
- Based on the container type used in your deployment, run one of the following commands on the VM:

- For unprivileged containers, run the following command:

```
ls /var/vcap/data/garden/depot/ \
|xargs -I{} /var/vcap/packages/grootfs/bin/grootfs --config \
/var/vcap/jobs/garden/config/grootfs_config.yml stats {} \
| cut -d: -f4 | cut -d- -f1 | awk '{sum += $1} END {print sum}'
```

- For privileged containers, run the following command:

```
ls /var/vcap/data/garden/depot/ \
|xargs -I{} /var/vcap/packages/grootfs/bin/grootfs --config \
/var/vcap/jobs/garden/config/privileged_grootfs_config.yml stats {} \
| cut -d: -f4 | cut -d- -f1 | awk '{sum += $1} END {print sum}'
```

The commands above return the total disk usage in bytes for all running containers.

## About Volumes in GrootFS

Underlying layers are known as `volumes` in GrootFS.

They are read-only and their changesets are layered together through an `OverlayFS` mount to create the rootfs for containers. When GrootFS writes each file system volume to disk, it also stores the number of bytes written to a file in the `meta` directory.

## Check Volume Disk Size

To find out the size of an individual volume, you can read the corresponding metadata file or run `du` on the volume itself. Perform the following steps:

- In your PAS deployment, use `bosh ssh` to connect to the BOSH-deployed VM running the Garden job.
- Based on the container type used in your deployment, run one of the following commands on the VM:
  - For unprivileged containers, run the following command on the VM:

```
cat /var/vcap/data/grootfs/store/unprivileged/meta/volume-VOLUME-SHA
```

Where `VOLUME-SHA` corresponds to the SHA value of the volume.

- For privileged containers, run the following command:

```
cat /var/vcap/data/grootfs/store/privileged/meta/volume-VOLUME-SHA
```

Where `VOLUME-SHA` corresponds to the SHA value of the volume.

The `cat` commands above return the volume size in bytes in the following format:

```
{"Size":5607885}
```

- Alternatively, use `du` and pass the absolute path to the volume. Run one of the following commands on the VM:

- For unprivileged containers, run the following command:

```
du -sch /var/vcap/data/grootfs/store/unprivileged/volumes/VOLUME-SHA/
```

Where `VOLUME-SHA` corresponds to the SHA value of the volume.

- For privileged containers, run the following command:

```
du -sch /var/vcap/data/grootfs/store/privileged/volumes/VOLUME-SHA/
```

Where `VOLUME-SHA` corresponds to the SHA value of the volume.

The `du` commands above return the volume size in the following format:

```
5.4M /var/vcap/data/grootfs/store/unprivileged/volumes/VOLUME-SHA/
```

## Determine Disk Usage and Reclaimable Disk Space

This section describes how to calculate the amount of disk space is in use and estimate how much space is reclaimable.

### Calculate Disk Use by All Active Volumes

For each container, GrootFS mounts the underlying volumes using overlay to a point in the `images` directory. This point is the rootfs for the container and is read-write.

GrootFS also stores the SHA of each underlying volume used by an image in the `meta` folder.

You can determine the bytes of all active volumes on disk by running one of the following commands:

 If you do not have `python3` is not installed, replace `python3` with `python` in the commands below.

- For unprivileged containers, run the following command:

```
for image in $(ls /var/vcap/data/grootfs/store/unprivileged/meta/dependencies/image\.*.json); \
do cat $image | python3 -c 'import json,sys;obj=json.load(sys.stdin); \
print("\n".join(obj))'; done | sort -u \
|xargs -I{} cat /var/vcap/data/grootfs/store/unprivileged/meta/volume-{} \
| cut -d : -f2 | cut -d\} -f1 \
| awk '{sum += $1} END {print sum}'
```

- For privileged containers, run the following command:

```
for image in $(ls /var/vcap/data/grootfs/store/privileged/meta/dependencies/image\.*.json); \
do cat $image | python3 -c 'import json,sys;obj=json.load(sys.stdin); \
print("\n".join(obj))'; done | sort -u \
|xargs -I{} cat /var/vcap/data/grootfs/store/privileged/meta/volume-{} \
| cut -d : -f2 | cut -d\} -f1 \
| awk '{sum += $1} END {print sum}'
```

The commands above return the total number of bytes used by all active volumes on disk.

## Calculate GrootFS Store Disk Usage

To determine how much total disk space the store is using, run the following command:

```
df | grep -E "/var/vcap/data/grootfs/store/(privileged|unprivileged)$" \
| awk '{sum += $3} END {print sum}'
```

The command above returns the total number of bytes used by the store.

## Calculate Reclaimable Disk Space

You can use values gathered from the commands above to calculate how much space can be cleared in GrootFS. Garbage collection reclaims disk space by pruning unused volumes.

The overall formula to calculate reclaimable disk space is to subtract the total disk in use by the store from the total disk used by active volumes.

For example, perform the following steps:

1. Calculate how much disk space the store is using by following the instructions in [Calculate GrootFS Store Disk Usage](#). For example, your result might be:

```
Total disk store = 5607885 bytes
```

2. Calculate how much disk space active volumes are using by following the instructions in [Calculate Disk Use by All Active Volumes](#). For example, your result might be:

```
Active volumes = 3212435 bytes
```

3. Subtract the amount of space used by active volumes from the space used by the store. For example, your result might be:

```
5607885 - 3212435 = 2395450 bytes
```

In this example, you can reclaim 2395450 bytes through garbage collection.

## How GrootFS Reclaims Disk Space

The thresholder component calculates and sets a value so that GrootFS's garbage collector can attempt to ensure that a small reserved space is kept free for other jobs. GrootFS only tries to garbage collect or reclaim space when that threshold is reached. However, if all the rootfs layers are actively in use by images, then garbage collection cannot occur and that space is used up.

If you determine that there is not enough reclaimable disk and more space is needed on disk, you should scale up your VMs to a larger size or add more VMs to provide more disk space.

Alternatively, you can configure a lower threshold for cell disk cleanup in Pivotal Application Service (PAS). For more information, see [Configuring Disk Cell Cleanup](#).

## Other Categories of GrootFS Disk Usage

There may be categories of GrootFS disk usage other than those listed in the above sections. However, the bulk of disk usage is stored in the `images/CONTAINER-ID/diff` and `volumes` directories, so these are rarely taken into consideration when calculating store usage.

You can find these directories under `/var/vcap/data/grootfs/store/unprivileged` for unprivileged containers and `/var/vcap/data/grootfs/store/privileged` for privileged containers.

GrootFS also stores information in the following directories:

- `l`: link directories. Shorter directory names are symlinked to volume directories to allow Groot to union mount more file paths.
- `locks`: file system lock directory to ensure safety during concurrent cleans and creates.
- `meta`: per image and volume metadata.
- `projectids`: empty numbered directories used to track image quotas.
- `tmp`: normal temporary directory contents.

These directories typically use less than 2 MB disk in total.

## Using Metadata

This topic describes metadata in Pivotal Application Service (PAS) and provides instructions for adding, updating, removing, and viewing metadata.

For additional information about adding metadata with the Cloud Foundry API (CAPI), see the *Metadata* section in the [CAPI documentation](#).

## About Metadata

PAS allows you to add metadata to objects such as spaces and apps. You can use metadata to provide additional information about the objects in your PAS deployment. This can help with operating, monitoring, and auditing.

For example, you can tag objects with metadata that describes the type of environment they belong to. You could also use metadata to describe app characteristics, such as front end or back end. Other examples include billing codes, points of contact, resource consumption, and information about security or risk.

Here is the full list of objects to which you can add metadata using the Cloud Foundry API:

- Apps
- Builds
- Buildpacks
- Deployments
- Droplets
- Isolation Segments
- Orgs
- Packages
- Processes
- Spaces
- Stacks
- Tasks
- Revisions

You can add metadata to apps and spaces using Apps Manager. For more information about adding metadata to apps using Apps Manager, see [Managing Apps and Service Instances Using Apps Manager](#). For more information about adding metadata to spaces using Apps Manager, see [Managing Orgs and Spaces Using Apps Manager](#).

## Types of Metadata

You can add two types of metadata to objects in PAS:

- **Labels:** Labels allow you to identify and select PAS objects. For example, if you have labeled all apps running in production, or all spaces that contain Internet-facing apps, you can then search for them.
- **Annotations:** Annotations allow you to add non-identifying metadata to PAS objects. You cannot query based on annotations. Also, there are fewer restrictions for key-value pairs of annotations than there are for labels. For example, you can include contact information of persons responsible for the object, or tool information for debugging purposes.

## Metadata Requirements

The following tables describe requirements for creating metadata.

### Requirements for Labels

The following table describes the requirements for creating labels.

| Label Requirements |  |  |  |
|--------------------|--|--|--|
|                    |  |  |  |

| Part of Label         | Length in characters | Allowed characters                                                                                                                                                                      | Other Requirements                                                                                                                                                                              |
|-----------------------|----------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| (Optional) Key Prefix | 0-253                | <ul style="list-style-type: none"> <li>• Alphanumeric</li> <li>• <input type="checkbox"/></li> <li>• <input checked="" type="checkbox"/></li> <li>• <input type="checkbox"/></li> </ul> | <ul style="list-style-type: none"> <li>• DNS subdomain format, with at least one <input checked="" type="checkbox"/>.</li> <li>• Must end with <input checked="" type="checkbox"/> /</li> </ul> |
| Key Name              | 1-63                 | <ul style="list-style-type: none"> <li>• Alphanumeric</li> <li>• <input type="checkbox"/></li> <li>• <input checked="" type="checkbox"/></li> <li>• <input type="checkbox"/></li> </ul> | Must begin and end with an alphanumeric character                                                                                                                                               |
| Value                 | 0-63                 | <ul style="list-style-type: none"> <li>• Alphanumeric</li> <li>• <input type="checkbox"/></li> <li>• <input checked="" type="checkbox"/></li> <li>• <input type="checkbox"/></li> </ul> | <ul style="list-style-type: none"> <li>• Must begin and end with an alphanumeric character</li> <li>• Empty values allowed</li> </ul>                                                           |

## Requirements for Annotations

The following table describes the requirements for creating annotations.

| Annotation Requirements |                      |
|-------------------------|----------------------|
| Part of Annotation      | Length in characters |
| Key                     | 1-1000               |
| Value                   | 0-5000               |

## Label Prefixes

You can ensure a label key is easily differentiated from other keys by using a prefix. A prefix is a namespacing pattern that helps you more clearly identify objects. Prefixes are in DNS subdomain format: `prefix.example.com`.

Consider an example in which you have two scanner tools: one for security and one for compliance. Both tools use a `scanned` label. You can disambiguate between the two tools using a prefix. The security tool can prefix a label with `security.example.com/scanned` and the compliance tool can prefix a label with `compliance.example.com/scanned`.

## Add Metadata to an Object

The sections below describe how to add labels and annotations to objects.

### Add a Label

The following procedure describes how to add a label:

1. To add a label to an object, run the following command:

```
cf curl v3/OBJECT-ENDPOINT/GUID \
-X PATCH \
-d'{
 "metadata": {
 "labels": {
 "LABEL-KEY": "LABEL-VALUE"
 }
 }
}'
```

Where:

- `OBJECT-ENDPOINT` is the CAPI endpoint for the type of object you want to label, such as `apps` or `organizations`
- `GUID` is the GUID of the object you want to label

- `LABEL-KEY` is the key for the label
- `LABEL-VALUE` is the corresponding value for the label key

For example, the following command labels an app as `"environment": "production"`.

```
$ cf curl v3/apps/1cb006ee-fb05-47e1-b541-c34179ddc446 \
-X PATCH \
-d'{
 "metadata": {
 "labels": {
 "environment": "production"
 }
 }
}'
```

## Add an Annotation

The following procedure describes how to add an annotation:

1. To add an annotation to an object, run the following command:

```
cf curl v3/OBJECT-ENDPOINT/GUID \
-X PATCH \
-d'{
 "metadata": {
 "annotations": {
 "ANNOTATION-KEY": "ANNOTATION-VALUE"
 }
 }
}'
```

Where:

- `OBJECT-ENDPOINT` is the CAPI endpoint for the type of object you want to label, such as `apps` or `organizations`
- `GUID` is the GUID of the object you want to label
- `ANNOTATION-KEY` is the key for the label
- `ANNOTATION-VALUE` is the corresponding value for the annotation key

For example, the following command provides a `contacts` annotation for an app.

```
$ cf curl v3/apps/1cb006ee-fb05-47e1-b541-c34179ddc446 \
-X PATCH \
-d'{
 "metadata": {
 "annotations": {
 "contacts": "Bill tel(1111111) email(bill@fixme), Bob tel(2222222) pager(3333333#555) email(bob@fixme)"
 }
 }
}'
```

## Update Metadata for an Object

To update metadata for an object, follow the procedure for adding metadata and provide a new value for an existing key. See [Add Metadata to an Object](#) above.

## Remove Metadata from an Object

To remove metadata from an object, follow the procedure for adding metadata and provide a `null` value for an existing key. See [Add Metadata to an Object](#) above.

## View Metadata for an Object

The following procedure describes how to view metadata:

- To view metadata using the list endpoint of an object, run the following command:

```
cf curl /v3/OBJECT-ENDPOINT/GUID
```

Where:

- OBJECT-ENDPOINT** is the CAPI endpoint for the type of object you want to view, such as `apps` or `organizations`
- GUID** is the GUID of the object you want to view

For example:

```
$ cf curl /v3/apps/1cb006ee-fb05-47e1-b541-c34179ddc446

{
 "guid": "1cb006ee-fb05-47e1-b541-c34179ddc446",
 "name": "my_app",
 "state": "STOPPED",
 "created_at": "2016-03-17T21:41:30Z",
 "updated_at": "2016-06-08T16:41:26Z",
 "lifecycle": {
 "type": "buildpack",
 "data": {
 "buildpacks": ["java_buildpack"],
 "stack": "cflinuxfs3"
 }
 },
 "relationships": {
 "space": {
 "data": {
 "guid": "2f35885d-0c9d-4423-83ad-fd05066f8576"
 }
 }
 },
 "links": {
 ...
 },
 "metadata": {
 "labels": {
 "environment": "production"
 },
 "annotations": {
 "contacts": "Bill tel(1111111) email(bill@fixme), Bob tel(2222222) pager(3333333#555) email(bob@fixme)"
 }
 }
}
```

## List Objects by Querying Labels

The following procedure describes how to list objects by querying label metadata:

- To query an object by using the `label_selector` parameter on its list endpoint, run the following command:

```
cf curl /v3/OBJECT-ENDPOINT/?label_selector=SELECTOR-REQUIREMENTS
```

Where:

- OBJECT-ENDPOINT** is the CAPI endpoint for the type of object you want to view, such as `apps` or `organizations`
- SELECTOR-REQUIREMENTS** is one of requirement types specified in [Selector Requirement Reference](#) below. You can add multiple selector requirements using a comma-separated list.

 **Note:** Ensure that this part of the URL is appropriately escaped.

For example, the following command selects all apps tagged with a label that has a key `environment` and value `production`.

```
$ cf curl /v3/apps/?label_selector=environment=production
```

## Selector Requirement Reference

The following table describes how to form selector requirements:

| Requirement   | Format                                            | Description                                                                          |
|---------------|---------------------------------------------------|--------------------------------------------------------------------------------------|
| existence     | <code>KEY</code>                                  | Returns all resources labeled with the given key                                     |
| existence     | <code>!KEY</code>                                 | Returns all resources not labeled with the given key                                 |
| equality      | <code>KEY==VALUE</code> or <code>KEY=VALUE</code> | Returns all resources labeled with the given key and value                           |
| inequality    | <code>KEY!=VALUE</code>                           | Returns all resources not labeled with the given key and value                       |
| set inclusion | <code>KEY in (VALUE1,VALUE2...)</code>            | Returns all resources labeled with the given key and one of the specified values     |
| set inclusion | <code>KEY notin (VALUE1,VALUE2...)</code>         | Returns all resources not labeled with the given key and one of the specified values |

## Example: Label Objects with a Git Commit

This section provides the following:

- A procedure for labeling an app, package, and droplet with a git commit SHA. See [Manually Label Objects](#).
- A script that automates the procedure. See [Automate Labeling Objects](#).

Labeling your app and related resources with a git commit SHA allows you to track which version of your code is running on PAS.

For more information about app packages and droplets, see the [API documentation](#).

### Manually Label Objects

To label an app, droplet, and package with a git commit SHA, do the following:

1. Log in to the Cloud Foundry Command Line Interface (cf CLI).
2. Run the following command and record the app GUID:

```
cf app APP_NAME --guid
```

Where `APP_NAME` is the name of the app.

3. Run the following command to return the GUID of the droplet and package associated with the app:

```
cf curl /v3/apps/APP_GUID/droplets/current
```

Where `APP_GUID` is the GUID of the app.

4. Record the GUID of the droplet and package:
  - o The droplet GUID is the value for the `"guid"` key.
  - o The package GUID is the end of the `"href"` URL for the `"package"` key.

For example, the droplet and package GUIDs are highlighted in blue in the following output:

```
{
 "guid": "fd35633f-5c5e-4e4e-a5a9-0722c970a9d2",
 ...
 "links": {
 "package": {
 "href": "https://api.run.pivotal.io/v3/packages/fd35633f-5c5e-4e4e-a5a9-0722c970a9d2"
 }
 }
}
```

5. Run the following command to label the app with a git commit SHA:

```
cf curl /v3/apps/APP_GUID -X PATCH -d '{"metadata": { "labels": { "commit": COMMIT_SHA } }}'
```

Where:

- o `APP_GUID` is the GUID of the app.

- o `COMMIT_SHA` is the SHA of the git commit.

6. Run the following command to label the app droplet with the same git commit SHA:

```
cf curl /v3/droplets/DROPLET_GUID -X PATCH -d '{"metadata": { "labels": { "commit": COMMIT_SHA } } }'
```

Where:

- o `DROPLET_GUID` is the GUID of the droplet.
- o `COMMIT_SHA` is the SHA of the git commit.

7. Run the following command to label the app package with the same git commit SHA:

```
cf curl /v3/packages/PACKAGE_GUID -X PATCH -d '{"metadata": { "labels": { "commit": COMMIT_SHA } } }'
```

Where:

- o `PACKAGE_GUID` is the GUID of the package.
- o `COMMIT_SHA` is the SHA of the git commit.

## Automate Labeling Objects

You can automate labeling objects by running a script either programatically or manually in the app repository.

### Prerequisite

To run the following example script, you must install `jq`. To download `jq`, see [jq](#).

### Example Script

The following script retrieves the GUID of the app, droplet, and package. It then uses the CAPI to label each resource with the current git commit SHA.

Replace `APP-NAME` with the name of your app.

```
#!/usr/bin/env bash

set -ex

APP_GUID=$(cf app APP-NAME --guid)
APP_URI="/v3/apps/${APP_GUID}"

DROPLET_GUID=$(cf curl "/v3/apps/${APP_GUID}/droplets/current" | jq -r .guid)
DROPLET_URI="/v3/droplets/${DROPLET_GUID}"

PACKAGE_GUID=$(cf curl "/v3/droplets/${DROPLET_GUID}" | jq -r .links.package.href | xargs basename)
PACKAGE_URI="/v3/packages/${PACKAGE_GUID}"

COMMIT_SHA=$(git rev-parse --short HEAD)
REQUEST_BODY=$(jq -nc --arg commit "$COMMIT_SHA" '{"metadata": { "labels": { "commit": $commit } } }'

cf curl "${APP_URI}" -X PATCH -d "${REQUEST_BODY}"
cf curl "${PACKAGE_URI}" -X PATCH -d "${REQUEST_BODY}"
cf curl "${DROPLET_URI}" -X PATCH -d "${REQUEST_BODY}"
```

## Managing Custom Buildpacks

This topic describes how an admin can manage additional buildpacks in Cloud Foundry using the Cloud Foundry Command Line Interface tool (cf CLI). If your application uses a language or framework that the Cloud Foundry system buildpacks do not support, you can take one of the following actions:

- [Write your own buildpack](#)
- Customize an existing buildpack
- Use a [Cloud Foundry Community Buildpack ↗](#)
- Use a [Heroku Third-Party Buildpack ↗](#)

### Add a Buildpack

 Note: You must be a Cloud Foundry admin user to run the commands discussed in this section.

To add a buildpack, run the `cf create-buildpack BUILDPACK PATH POSITION [--enable|--disable]` command. The arguments to [cf create-buildpack ↗](#) specify the following:

- **BUILDPACK** specifies the buildpack name.
- **PATH** specifies the location of the buildpack. PATH can point to a zip file, the URL of a zip file, or a local directory.
- **POSITION** specifies where to place the buildpack in the detection priority list. For more information, see the [Buildpack Detection ↗](#) topic.
- **enable** or **disable** specifies whether to allow apps to be pushed with the buildpack. This argument is optional, and defaults to enable. When a buildpack is disabled, app developers cannot push apps using that buildpack.

To confirm that you have successfully added a buildpack, run `cf buildpacks`.

The following example shows the output from running the `cf buildpacks` command after an administrator adds a Python buildpack:

```
$ cf buildpacks
Getting buildpacks...

buildpack position enabled locked filename
ruby_buildpack_1 true false buildpack_ruby_v46-245-g2fc4ad8.zip
nodejs_buildpack_2 true false buildpack_nodejs_v8-177-g2b0a5cf.zip
java_buildpack_3 true false buildpack_java_v2.1.zip
python_buildpack_4 true false buildpack_python_v2.7.6.zip
```

### Rename a Buildpack

To rename a buildpack, run `cf rename-buildpack BUILDPACK-NAME NEW-BUILDPACK-NAME`. Replace **BUILDPACK-NAME** with the original buildpack name, and **NEW-BUILDPACK-NAME** with the new buildpack name.

For more information about renaming buildpack, see the [cf CLI documentation ↗](#).

### Update a Buildpack

```
$ cf update-buildpack BUILDPACK [-p PATH] [-i POSITION] [-s STACK] [--enable|--disable] [-lock|--unlock]
```

For more information about updating buildpacks, see the [cf CLI documentation ↗](#).

### Delete a Buildpack

```
$ cf delete-buildpack BUILDPACK [-s STACK] [-f]
```

For more information about deleting buildpacks, see the [cf CLI documentation](#).

## Lock and Unlock a Buildpack

Every new version of Cloud Foundry includes an updated buildpack. By default, your deployment applies the most recent buildpack when you upgrade. In some cases, however, you may want to preserve an existing buildpack, rather than upgrade to the latest version. For example, if an app you deploy depends on a specific component in Buildpack A that is not available in Buildpack B, you may want to continue using Buildpack A.

The `--lock` flag lets you continue to use your existing buildpack even after you upgrade. Locked buildpacks are not updated when PCF updates. You must manually unlock them to update them.

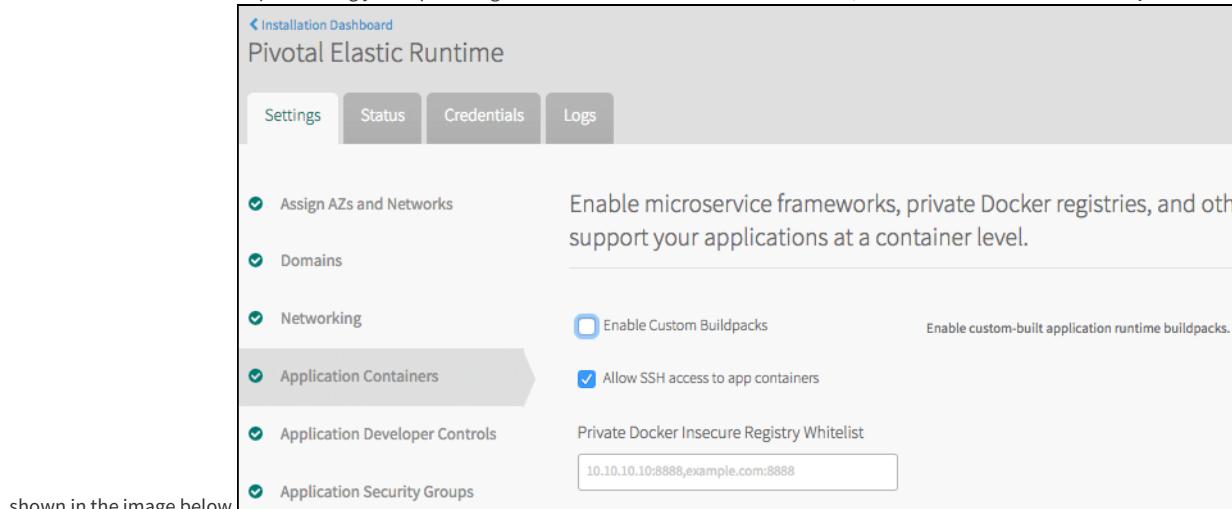
If you elect to use the `--unlock` flag, your deployment will apply the most recent buildpack when you upgrade PCF.

```
cf update-buildpack BUILDPACK [-p PATH] [-i POSITION] [-s STACK] [--enable|--disable] [--lock]--unlock]
```

This feature is also available through the API. For more information, see [Lock or unlock a Buildpack](#) in the [Cloud Foundry API](#) documentation.

## Disabling Custom Buildpacks

You can disable custom buildpacks using your Ops Manager PAS tile. From the [Cloud Controller](#) tab, check the **Disable Custom Buildpacks** checkbox, as



shown in the image below.

By default, the cf CLI gives developers the option of using a custom buildpack when they deploy apps to PAS. To do so, they use the `-b` option to provide a custom buildpack URL with the `cf push` command. The **Disable Custom Buildpacks** checkbox prevents the `-b` option from being used with external buildpack URLs.

For more information about custom buildpacks, refer to the [buildpacks](#) section of the PCF documentation.

## Using Docker in Cloud Foundry

This topic describes how Cloud Foundry (CF) operators can enable CF developers to run their apps in Docker containers, and explains how Docker works in Cloud Foundry.

For information about Diego, the Cloud Foundry component that manages application containers, see the [Diego Architecture](#) topic. For information about how CF developers push apps with Docker images, see [Deploying an App with Docker](#).

## Enable Docker

By default, apps deployed with the `cf push` command run in standard Cloud Foundry Linux containers. With Docker support enabled, Cloud Foundry can also deploy and manage apps running in Docker containers.

To deploy apps to Docker, developers run `cf push` with the `--docker-image` option and the location of a Docker image to create the containers from. See the [Push a Docker Image](#) topic for information about how CF developers push apps with Docker images.

To enable Docker support on a CF deployment, an operator must do the following:

- [Enable](#) the `diego_docker` feature flag.
- [Configure](#) access to any Docker registries that developers want to use images from.

### Enable and Disable the `diego_docker` Feature Flag

The `diego_docker` feature flag governs whether a CF deployment supports Docker containers.

To enable Docker support, run:

```
$ cf enable-feature-flag diego_docker
```

To disable Docker support, run:

```
$ cf disable-feature-flag diego_docker
```

 **Note:** Disabling the `diego_docker` feature flag stops all Docker-based apps in your deployment within a few convergence cycles, on the order of a minute.

## Configure Docker Registry Access

To support Docker, Pivotal Cloud Foundry needs the ability to access Docker registries using either a Certificate Authority or an IP address whitelist. The [Using Docker Registries](#) topic explains how to configure this access.

## Docker Image Contents

A Docker image consists of a collection of layers. Each layer consists of one or both of the following:

- Raw bits to download and mount. These bits form the file system.
- Metadata that describes commands, users, and environment for the layer. This metadata includes the `ENTRYPOINT` and `CMD` directives, and is specified in the Dockerfile.

## How Garden-runC Creates Containers

Diego currently uses Garden-runC to construct Linux containers.

Both Docker and Garden-runC use libraries from the [Open Container Initiative \(OCI\)](#) to build Linux containers. After creation, these containers use name space isolation, or *namespaces*, and control groups, or *cgroups*, to isolate processes in containers and limit resource usage. These are common kernel resource isolation features used by all Linux containers.

 **Note:** PAS versions v1.8.8 and above use Garden-runC instead of Garden-Linux.

Before Garden-runC creates a Linux container, it creates a file system that is mounted as the root file system of the container. Garden-runC supports mounting Docker images as the root file systems for the containers it creates.

When creating a container, both Docker and Garden-runC perform the following actions:

- Fetch and cache the individual layers associated with a Docker image
- Combine and mount the layers as the root file system

These actions produce a container whose contents exactly match the contents of the associated Docker image.

Earlier versions of Diego used Garden-Linux. For more information, see the [Garden](#) topic.

## How Diego Runs and Monitors Processes

After Garden-runC creates a container, Diego runs and monitors the processes inside of it.

To determine which processes to run, the [Cloud Controller](#) fetches and stores the metadata associated with the Docker image. The Cloud Controller uses this metadata to perform the following actions:

- Runs the start command as the user specified in the Docker image
- Instructs Diego and the [Gorouter](#) to route traffic to the lowest-numbered port [exposed](#) in the Docker image, or port 8080 if the Dockerfile does not explicitly expose a listen port.

 **Note:** When launching an application on Diego, the Cloud Controller honors any user-specified overrides such as a custom start command or custom environment variables.

## Docker Security Concerns in a Multi-Tenant Environment

The attack surface area for a Docker-based container running on Diego remains somewhat higher than that of a buildpack application because Docker allows users to fully specify the contents of their root file systems. A buildpack application runs on a trusted root filesystem.

Garden-runC provides features that allow the platform to run Docker images more securely in a multi-tenant context. In particular, Cloud Foundry uses the [user-namespacing](#) feature found on modern Linux kernels to ensure that users cannot gain escalated privileges on the host even if they escalate privileges within a container.

The Cloud Controller always runs Docker containers on Diego with user namespaces enabled. This security restriction prevents certain features, such as the ability to mount FuseFS devices, from working in Docker containers. Docker applications can use fuse mounts through [volume services](#), but they cannot directly mount fuse devices from within the container.

To mitigate security concerns, Cloud Foundry recommends that you run only trusted Docker containers on the platform. By default, the Cloud Controller does not allow Docker-based applications to run on the platform.

## User Accounts and Communications

The following topics provide information about managing user accounts and user communications in Pivotal Application Service (PAS):

- [Creating and Managing Users with the cf CLI](#)
- [Creating and Managing Users with the UAA CLI \(UAAC\)](#)
- [Creating New PAS User Accounts](#)
- [Adding Existing SAML or LDAP Users to a Pivotal Cloud Foundry Deployment](#)
- [Getting Started with the Notifications Service](#)

## Creating and Managing Users with the cf CLI

Using the Cloud Foundry Command Line Interface (cf CLI), administrators, Org Managers, and Space Managers can manage users. Cloud Foundry uses role-based access control, with each role granting permissions in either an organization or an application space.

For more information, see [Organizations, Spaces, Roles, and Permissions](#).

### About Roles

To manage all users, organizations, and roles with the cf CLI, log in with your admin credentials. In Pivotal Operations Manager, refer to [PAS > Credentials](#) for the admin name and password.

If the feature flag `set_roles_by_username` is enabled, Org Managers can [assign org roles](#) to existing users in their org and Space Managers can [assign space roles](#) to existing users in their space. For more information about using feature flags, see the [Feature Flags](#) topic.

## Creating and Deleting Users

| FUNCTION                                                                              | COMMAND                                              | EXAMPLE                                         |
|---------------------------------------------------------------------------------------|------------------------------------------------------|-------------------------------------------------|
| Create a new user                                                                     | <code>cf create-user USERNAME PASSWORD</code>        | <code>cf create-user Alice pa55w0rd</code>      |
| Create a new user, specifying LDAP as an external identity provider                   | <code>cf create-user USERNAME --origin ORIGIN</code> | <code>cf create-user Aayah ldap</code>          |
| Create a new user, specifying SAML or OpenID Connect as an external identity provider | <code>cf create-user USERNAME --origin ORIGIN</code> | <code>cf create-user Aiko provider-alias</code> |
| Delete a user                                                                         | <code>cf delete-user USERNAME</code>                 | <code>cf delete-user Alice</code>               |

### Creating Administrator Accounts

To create a new administrator account, use the [UAA CLI](#).

 **Note:** The cf CLI cannot create new administrator accounts.

## Org and App Space Roles

A user can have one or more roles. The combination of these roles defines the user's overall permissions in the org and within specific app spaces in that org.

### Org Roles

Valid [org roles](#) are OrgManager, BillingManager, and OrgAuditor.

| FUNCTION                                       | COMMAND                                                        | EXAMPLE                                                        |
|------------------------------------------------|----------------------------------------------------------------|----------------------------------------------------------------|
| View the organizations belonging to an account | <code>cf orgs</code>                                           | <code>cf orgs</code>                                           |
| View all users in an organization by role      | <code>cf org-users ORGANIZATION-NAME</code>                    | <code>cf org-users my-example-org</code>                       |
| Assign an org role to a user                   | <code>cf set-org-role USERNAME ORGANIZATION-NAME ROLE</code>   | <code>cf set-org-role Alice my-example-org OrgManager</code>   |
| Remove an org role from a user                 | <code>cf unset-org-role USERNAME ORGANIZATION-NAME ROLE</code> | <code>cf unset-org-role Alice my-example-org OrgManager</code> |

If multiple accounts share a username, `set-org-role` and `unset-org-role` return an error. See [Identical Usernames in Multiple Origins](#) for details.

## App Space Roles

Each app space role applies to a specific app space.

Valid [app space roles](#) are SpaceManager, SpaceDeveloper, and SpaceAuditor.

| FUNCTION                          | COMMAND                                                        | EXAMPLE                                                                        |
|-----------------------------------|----------------------------------------------------------------|--------------------------------------------------------------------------------|
| View the spaces in an org         | cf spaces                                                      | <code>cf spaces</code>                                                         |
| View all users in a space by role | cf space-users ORGANIZATION-NAME SPACE-NAME                    | <code>cf space-users my-example-org development</code>                         |
| Assign a space role to a user     | cf set-space-role USERNAME ORGANIZATION-NAME SPACE-NAME ROLE   | <code>cf set-space-role Alice my-example-org development SpaceAuditor</code>   |
| Remove a space role from a user   | cf unset-space-role USERNAME ORGANIZATION-NAME SPACE-NAME ROLE | <code>cf unset-space-role Alice my-example-org development SpaceAuditor</code> |

If multiple accounts share a username, `set-space-role` and `unset-space-role` return an error. See [Identical Usernames in Multiple Origins](#) for details.

## Creating and Managing Users with the UAA CLI (UAAC)

Using the UAA Command Line Interface (UAAC), an administrator can create users in the User Account and Authentication (UAA) server.

 **Note:** The UAAC only creates users in UAA, and does not assign roles in the Cloud Controller database (CCDB). In general, administrators create users using the Cloud Foundry Command Line Interface (cf CLI). The cf CLI both creates user records in the UAA and associates them with org and space roles in the CCDB. Before administrators can assign roles to the user, the user must log in through Apps Manager or the cf CLI for the user record to populate the CCDB. Review the [Creating and Managing Users with the cf CLI](#) topic for more information.

For additional details and information, refer to the following topics:

- [UAAC Overview](#)
- [UAAC Sysadmin Guide](#)
- [Other UAAC Documentation](#)

 **Note:** UAAC requires Ruby v2.3.1 or later. If you have an earlier version of Ruby installed, install v2.3.1 or later before using the UAAC.

For more information about which roles can perform various operations in Cloud Foundry, see the [Roles and Permissions](#) topic.

### Create an Admin User

1. Install the UAAC CLI, `uaac`.

```
$ gem install cf-uaac
```

2. Use the `uaac target uaa.YOUR-DOMAIN` command to target your UAA server.

```
$ uaac target uaa.example.com
```

3. Record the `uaa:admin:client_secret` from your deployment manifest.

4. Run `uaac token client get admin -s ADMIN-CLIENT-SECRET` to authenticate and obtain an access token for the admin client from the UAA server. Replace `ADMIN-CLIENT-SECRET` with the admin secret you have retrieved in previous step. UAAC stores the token in `~/.uaac.yml`.

```
$ uaac token client get admin -s MyAdminSecret
```

5. Use the `uaac contexts` command to display the users and applications authorized by the UAA server, and the permissions granted to each user and application.

```
$ uaac contexts

[1]*[admin]
client_id: admin
access_token: aBcdEfg0hIJKlm123.e
token_type: bearer
expires_in: 43200
scope: uaa.admin clients.secret scim.read
jti: 91b3-abed1233
```

6. In the output from `uaac contexts`, search in the `scope` section of the `client_id: admin` user for `scim.write`. The value `scim.write` represents sufficient permissions to create accounts.

7. If the admin user lacks permissions to create accounts, add the permissions by following these steps:

- o Run `uaac client update admin --authorities "EXISTING-PERMISSIONS scim.write"` to add the necessary permissions to the admin user account on the UAA server. Replace EXISTING-PERMISSIONS with the current contents of the `scope` section from `uaac contexts`.
- o Run `uaac token delete` to delete the local token.
- o Run `uaac token client get admin` to obtain an updated access token from the UAA server.

```
$ uaac contexts

[1]*[admin]
client_id: admin
...
scope: uaa.admin clients.secret scim.read
...

$ uaac client update admin --authorities "uaac client get admin | \
awk '/:{e=0}/authorities:{e=1;if(e==1){$1="";print}}'" scim.write"

$ uaac token delete
$ uaac token client get admin
```

8. Run the following command to create an admin user: `uaac user add NEW-ADMIN-USERNAME -p NEW-ADMIN-PASSWORD --emails NEW-ADMIN-EMAIL`. Replace `NEW-ADMIN-USERNAME`, `NEW-ADMIN-PASSWORD`, and `NEW-ADMIN-EMAIL` with appropriate information.

```
$ uaac user add Adam -p newAdminSecretPassword --emails newadmin@example.com
```

9. Run `uaac member add GROUP NEW-ADMIN-USERNAME` to add the new admin to the groups `cloud_controller.admin`, `uaa.admin`, `scim.read`, and `scim.write`.

```
$ uaac member add cloud_controller.admin Adam
$ uaac member add uaa.admin Adam
$ uaac member add scim.read Adam
$ uaac member add scim.write Adam
```

## Create an Admin Read-Only User

The admin read-only account can view but not modify almost all Cloud Controller API resources. The admin read-only account cannot view process `stats` or `logs`.

If you want to create an admin read-only user account, then perform the following steps:

1. Obtain the credentials of an admin client created using UAAC as above, or refer to the `uaa: scim` section of your deployment manifest for the user name and password of an admin user.
2. Run `uaac token client get admin -s ADMIN-CLIENT-SECRET` to authenticate and obtain an access token for the admin client from the UAA server. Replace `ADMIN-CLIENT-SECRET` with your admin secret. UAAC stores the token in `~/.uaac.yml`.

```
$ uaac token client get admin -s MyAdminSecret
```

3. Run the following command to create an admin read-only user: `uaac user add NEW-USERNAME -p NEW-PASSWORD --emails NEW-EMAIL`. Replace `NEW-USERNAME`, `NEW-PASSWORD`, and `NEW-EMAIL` with appropriate information.

```
$ uaac user add Bob -p SecretPassword --emails bob@example.com
```

4. Run `uaac member add GROUP NEW-USERNAME` to add the new admin read-only account to the groups `cloud_controller.admin_read_only` and `scim.read`.

```
$ uaac member add cloud_controller.admin_read_only Bob
$ uaac member add scim.read Bob
```

## Create a Global Auditor

The global auditor account has read-only access to almost all Cloud Controller API resources but cannot access secret data such as environment variables. The global auditor account cannot view process `stats` or `logs`.

Perform the following steps to create a global auditor account.

1. Obtain the credentials of an admin client created using UAAC as above, or refer to the `uaa: scim` section of your deployment manifest for the user name and password of an admin user.
2. Run `uaac token client get admin -s ADMIN-CLIENT-SECRET` to authenticate and obtain an access token for the admin client from the UAA server. Replace `ADMIN-CLIENT-SECRET` with your admin secret. UAAC stores the token in `~/.uaac.yml`.

```
$ uaac token client get admin -s MyAdminSecret
```

- Run `uaac user add NEW-USERNAME -p NEW-PASSWORD --emails NEW-EMAIL` to create a global auditor user. Replace `NEW-USERNAME`, `NEW-PASSWORD`, and `NEW-EMAIL` with appropriate information.

```
$ uaac user add Alice -p SecretPassword --emails alice@example.com
```

- Ensure that the `cloud_controller.global_auditor` group exists.

```
$ uaac group add cloud_controller.global_auditor
```

- Run `uaac member add GROUP NEW-USERNAME` to add the new global auditor account to the `cloud_controller.global_auditor` group.

```
$ uaac member add cloud_controller.global_auditor Alice
```

## Grant Admin Permissions to an External Group (SAML or LDAP)

To grant all users under an external group admin permissions, do the following:

- Obtain the credentials of an admin client created using UAAC as above, or refer to the `uaa: scim` section of your deployment manifest for the user name and password of an admin user.
- Run `uaac token client get admin -s ADMIN-CLIENT-SECRET` to authenticate and obtain an access token for the admin client from the UAA server. Replace `ADMIN-CLIENT-SECRET` with your admin secret. UAAC stores the token in `~/.uaac.yml`.

```
$ uaac token client get admin -s MyAdminSecret
```

- Follow the procedure that corresponds to your use case:

- [Grant Admin Permissions for LDAP](#)
- [Grant Admin Permissions for SAML](#)

**Note:** The UAA will not grant scopes for users in external groups until the next time the user logs in. This means that users granted scopes from external group mappings must log out from PCF and log back in before their new scope takes effect.

### Grant Admin Permissions for LDAP

Run the commands below to grant all users under the mapped LDAP Group admin permissions. Replace `GROUP-DISTINGUISHED-NAME` with an appropriate group name.

- `uaac group map --name scim.read "GROUP-DISTINGUISHED-NAME"`
- `uaac group map --name scim.write "GROUP-DISTINGUISHED-NAME"`
- `uaac group map --name cloud_controller.admin "GROUP-DISTINGUISHED-NAME"`

### Grant Admin Permissions for SAML

- Retrieve the name of your SAML provider by navigating to the PAS tile on the Ops Manager Installation Dashboard, clicking [Authentication and Enterprise SSO](#), and recording the value under [Provider Name](#). For more information about configuring PCF for a SAML identity provider, see the [Configuring Authentication and Enterprise SSO for PAS](#) topic.
- Run the commands below to grant all users under the mapped SAML group admin permissions. Replace `GROUP-NAME` with the group name, and `SAML-PROVIDER-NAME` with the name of your SAML provider.
  - `uaac group map --name scim.read "GROUP-NAME" --origin SAML-PROVIDER-NAME`
  - `uaac group map --name scim.write "GROUP-NAME" --origin SAML-PROVIDER-NAME`
  - `uaac group map --name cloud_controller.admin "GROUP-NAME" --origin SAML-PROVIDER-NAME`

## Create Users

1. Obtain the credentials of an admin client created using UAAC as above, or refer to the `uaa: scim` section of your deployment manifest for the user name and password of an admin user.
2. Run `cf login -u NEW-ADMIN-USERNAME -p NEW-ADMIN-PASSWORD` to log in.

```
$ cf login -u Adam -p newAdminSecretPassword
```

3. Run `cf create-user NEW-USER-NAME NEW-USER-PASSWORD` to create a new user.

```
$ cf create-user Charlie aNewPassword
```

## Change Passwords

1. Obtain the credentials of an admin client created using UAAC as above, or refer to the `uaa: scim` section of your deployment manifest for the user name and password of an admin user.
2. Run `uaac token client get admin -s ADMIN-CLIENT-SECRET` to authenticate and obtain an access token for the admin client from the UAA server. Replace `ADMIN-CLIENT-SECRET` with your admin secret. UAAC stores the token in `~/.uaac.yml`.

```
$ uaac token client get admin -s MyAdminSecret
```

3. Run `uaac contexts` to display the users and applications authorized by the UAA server, and the permissions granted to each user and application.

```
$ uaac contexts

[1]*[admin]
client_id: admin
access_token: aBcdEfg0hIJKlm123.e
token_type: bearer
expires_in: 43200
scope: uaa.admin.clients.secret password.read
jti: 91b3-abed1233
```

4. In the output from `uaac contexts`, search in the `scope` section of the `client_id: admin` user for `password.write`. The value `password.write` represents sufficient permissions to change passwords.
5. If the admin user lacks permissions to change passwords, add the permissions by following these steps:

- o Run `uaac client update admin --scope "EXISTING-PERMISSIONS password.write"` to add the necessary permissions to the admin user account on the UAA server. Replace EXISTING-PERMISSIONS with the current contents of the `scope` section from `uaac contexts`.
- o Run `uaac token delete` to delete the local token.
- o Run `uaac token client get admin` to obtain an updated access token from the UAA server.

```
$ uaac contexts

[1]*[admin]
client_id: admin
...
scope: uaa.admin.clients.secret password.read
...

$ uaac client update admin --scope "uaac client get admin | \
awk '/:{e=0}/authorities/{e=1;if(e==1){$1=""};print}'" password.write"

$ uaac token delete
$ uaac token client get admin
```

6. Run `uaac password set USER-NAME -p TEMP-PASSWORD` to change an existing user password to a temporary password.

```
$ uaac password set Charlie -p ThisIsATempPassword
```

7. Provide the `TEMP-PASSWORD` to the user. Have the user use `cf target api.YOUR-DOMAIN`, `cf login -u USER-NAME -p TEMP-PASSWORD`, and `cf passwd` to change the temporary password. See the [Configuring UAA Password Policy](#) topic to configure the password policy.

```
$ cf target api.example.com
$ cf login -u Charlie -p ThisIsATempPassword
$ cf passwd
```

Current Password>ThisIsATempPassword

New Password>\*\*\*\*\*

Verify Password>\*\*\*\*\*  
Changing password...

## Retrieve User Email Addresses

Some Cloud Foundry components, like Cloud Controller, only use GUIDs for user identification. You can use the UAA to retrieve the emails of your Cloud Foundry instance users either as a list or, for a specific user, with that user's GUID.

Follow the steps below to retrieve user email addresses:

1. Run `uaac target uaa.YOUR-DOMAIN` to target your UAA server.

```
$ uaac target uaa.example.com
```

2. Record the `uaa:admin:client_secret` from your deployment manifest.

3. Run `uaac token client get admin -s ADMIN-CLIENT-SECRET` to authenticate and obtain an access token for the admin client from the UAA server. Replace `ADMIN-CLIENT-SECRET` with your admin secret. UAAC stores the token in `~/uaac.yml`.

```
$ uaac token client get admin -s MyAdminSecret
```

4. Run `uaac contexts` to display the users and applications authorized by the UAA server, and the permissions granted to each user and application.

```
$ uaac contexts

[1]*[admin]
client_id: admin
access_token: aBcdEfg0hIJKlm123.e
token_type: bearer
expires_in: 43200
scope: uaa.admin.clients.secret
jti: 91b3-abcd123
```

5. In the output from `uaac contexts`, search in the `scope` section of the `client_id: admin` user for `scim.read`. The value `scim.read` represents sufficient permissions to query the UAA server for user information.

6. If the admin user lacks permissions to query the UAA server for user information, add the permissions by following these steps:

- o Run `uaac client update admin --authorities "EXISTING-PERMISSIONS scim.write"` to add the necessary permissions to the admin user account on the UAA server. Replace EXISTING-PERMISSIONS with the current contents of the `scope` section from `uaac contexts`.
- o Run `uaac token delete` to delete the local token.
- o Run `uaac token client get admin` to obtain an updated access token from the UAA server.

```
$ uaac contexts

[1]*[admin]
client_id: admin
...
scope: uaa.admin.clients.secret
...

$ uaac client update admin --authorities "uaa.admin.clients.secret scim.read"

$ uaac token delete
$ uaac token client get admin
```

7. Run `uaac users` to list your Cloud Foundry instance users. By default, the `uaac users` command returns information about each user account including GUID, name, permission groups, activity status, and metadata. Use the `--attributes emails` or `-a emails` flag to limit the output of `uaac users` to email

addresses.

```
$ uaac users --attributes emails

resources:
 emails:
 value: user1@example.com
 emails:
 value: user2@example.com
 emails:
 value: user3@example.com
```

- Run `uaac users "id eq GUID" --attributes emails` with the GUID of a specific user to retrieve that user's email address.

```
$ uaac users "id eq 'aabbcc11-22a5-87-8056-beaf84'" --attributes emails

resources:
 emails:
 value: user1@example.com
```

## Getting Started with the Notifications Service

This topic describes how to use the Notifications Service, including how to create a client, obtain a token, register notifications, create a custom template, and send notifications to your users.

### Prerequisites

- Install [Pivotal Application Service \(PAS\)](#).
- You must have `admin` permissions on your Cloud Foundry instance. You also must configure [Application Security Groups \(ASGs\)](#).
- Install the [Cloud Foundry Command Line Interface \(cf CLI\)](#) and [User Account and Authorization Server \(UAAC\)](#) command line tools.

## Create a Client and Get a Token

To interact with the Notifications Service, you must create [UAA](#) scopes:

1. Use `uaac target uaa.YOUR-DOMAIN` to target your UAA server.

```
$ uaac target uaa.example.com
```

2. Record the **Admin Client Credentials** from the **UAA** row in the PAS **Credentials** tab.

3. Use `uaac token client get admin -s ADMIN-CLIENT-SECRET` to authenticate and obtain an access token for the admin client from the UAA server. UAAC stores the token in `~/.uaac.yml`.

```
$ uaac token client get admin -s MyAdminPassword
```

4. Create a `notifications-admin` client with the required scopes.

```
$ uaac client add notifications-admin --authorized_grant_types client_credentials --authorities \
 notifications.manage,notifications.write,notification_templates.write,notification_templates.read,critical_notifications.write
```

- `notifications.write`: send a notification. For example, you can send notifications to a user, space, or everyone in the system.
- `notifications.manage`: update notifications and assign templates for that notification.
- (Optional) `notification_templates.write`: create a custom template for a notification.
- (Optional) `notification_templates.read`: check which templates are saved in the database.

5. Log in using your newly created client:

```
$ uaac token client get notifications-admin
```

 **Note:** Stay logged in to this client to follow the examples in this topic.

## Register Notifications

 **Note:** To register notifications, you must have the `notifications.manage` scope on the client. To set critical notifications, you must have the `critical_notifications.write` scope.

You must register a notification before sending it. Using the token `notifications-admin` from the previous step, the following example registers two notifications with the following properties:

```
$ uaac curl https://notifications.user.example.com/notifications -X PUT --data '{ "source_name": "Cloud Ops Team",
 "notifications": {
 "system-going-down": { "critical": true, "description": "Cloud going down" },
 "system-up": { "critical": true, "description": "Cloud back up" }
 }
}'
```

- `source_name` has “Cloud Ops Team” set as the description.
- `system-going-down` and `system-up` are the notifications set.
- `system-going-down` and `system-up` are made `critical`, so no users can unsubscribe from that notification.

## Create a Custom Template

 **Note:** To view a list of templates, you must have the `notifications_templates.read` scope. To create a custom template, you must have the `notification_templates.write` scope.

A template is made up of a name, a subject, a text representation of the template you are sending for mail clients that do not support HTML, and an HTML version of the template.

The system provides a default template for all notifications, but you can create a custom template using the following curl command.

```
$ uaac curl https://notifications.user.example.com/templates -X POST --data \
'{"name":"site-maintenance","subject":"Maintenance: {{.Subject}}","text":"The site has gone down for maintenance. More information to follow {{.Text}}","html":"
The site has gone down for maintenance. More information to follow {{.HTML}}"}'
```

Variables that take the form `{}{{}}` interpolate data provided in the send step before a notification is sent. Data that you can insert into a template during the send step include `{}{{.Text}}`, `{}{{.HTML}}`, and `{}{{.Subject}}`.

This curl command returns a unique template ID that can be used in subsequent calls to refer to your custom template. The result looks similar to this:

```
{"template-id": "E3710280-954B-4147-B7E2-AF5BF62772B5"}
```

Check all of your saved templates by running a curl command:

```
$ uaac curl https://notifications.user.example.com/templates -X GET
```

## Associate a Custom Template with a Notification

In this example, the `system-going-down` notification belonging to the `notifications-admin` client is associated with the template ID `E3710280-954B-4147-B7E2-AF5BF62772B5`. This is the template ID of the template we created in the previous section.

Associating a template with a notification requires the `notifications.manage` scope.

```
$ uaac curl https://notifications.user.example.com/clients/notifications-admin/notifications/system-going-down/template \
-X PUT --data '{"template": "E3710280-954B-4147-B7E2-AF5BF62772B5"}'
```

Any notification that does not have a custom template applied, such as `system-up`, defaults to a system-provided template.

## Send a Notification

 **Note:** To send a critical notification, you must have the `critical_notifications.write` scope. To send a non-critical notification, you must have the `notifications_write` scope.

You can send a notification to the following recipients:

- A user
- A space
- An organization
- All users in the system
- A UAA-scope
- An email address

For details, see the [Notifications V1 Documentation](#) in GitHub.

The following example sends the `system-going-down` notification described above to all users in the system.

```
$ uaac curl https://notifications.user.example.com/everyone -X POST --data \
'{"kind_id":"system-going-down","text":"The system is going down while we upgrade our storage","html":"THE SYSTEM IS DOWN
The system is going down while we upgrade our storage
","subject":"Upgrade to Storage","reply_to":"no-reply@example.com"}'
```

## Routing

The following topics provide information about managing routes and domains in Pivotal Application Service (PAS):

- [Enabling IPv6 for Hosted Applications](#)
- [Supporting WebSockets](#)
- [Configuring Load Balancer Healthchecks for Cloud Foundry](#)
- [Securing Traffic into Cloud Foundry](#)
- [Enabling TCP Routing](#)
- [Service Mesh \(Beta\)](#)

## Enabling IPv6 for Hosted Applications

The procedure described below allows apps deployed to Pivotal Application Service (PAS) to be reached using IPv6 addresses.

 **Note:** Amazon Web Services (AWS) EC2 instances currently do not support IPv6.

Pivotal Application Service (PAS) system components use a separate DNS subdomain from hosted applications. These components currently support only IPv4 DNS resolved addresses. This means that although an IPv6 address can be used for application domains, the system domain must resolve to an IPv4 address.

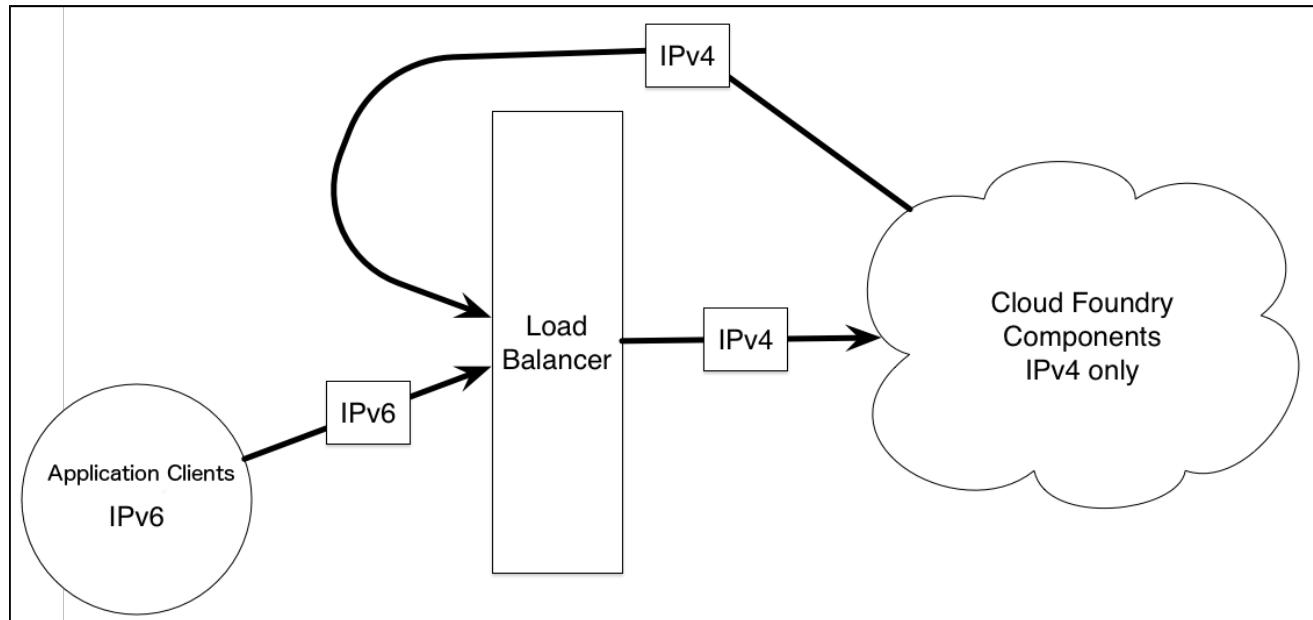
Complete the following steps to enable support for IPv6 application domains:

1. Set up an external load balancer for your Pivotal Application Service (PAS) deployment. See [Using Your Own Load Balancer](#).
2. Configure DNS to resolve application domains to an IPv6 address on your external load balancer.

 **Note:** Your IPv4 interface for the system domain and IPv6 interface for application domain can be configured on the same or different load balancers.

3. Configure the external load balancer to route requests for an IPv6 address to an IPv4 address as follows:
  - If you are using the HAProxy load balancer for SSL termination, route to its IPv4 address.
  - Otherwise, route directly to the IPv4 addresses of the Gorouters.

The following diagram illustrates how a single load balancer can support traffic on both IPv4 and IPv6 addresses for a Pivotal Application Service (PAS) installation.



See [Routes and Domains](#) for more information about domains in Pivotal Application Service (PAS).

## Supporting WebSockets

This topic explains how Cloud Foundry (CF) uses WebSockets, why developers use WebSockets in their applications, and how operators can configure their load balancer to support WebSockets.

Operators who use a load balancer to distribute incoming traffic across CF [router](#) instances must configure their load balancer for WebSockets. Otherwise, the [Loggregator](#) system cannot stream application logs to developers, or application event data and component metrics to third-party aggregation services. Additionally, developers cannot use WebSockets in their applications.

## Understand WebSockets

The WebSocket protocol provides full-duplex communication over a single TCP connection. Applications can use WebSockets to perform real-time data exchange between a client and a server more efficiently than HTTP.

CF uses WebSockets for the following metrics and logging purposes:

1. To stream all application event data and component metrics from the [Doppler](#) server instances to the [Traffic Controller](#)
2. To stream application logs from the Traffic Controller to developers using the Cloud Foundry Command Line Interface (cf CLI) or [Apps Manager](#)
3. To stream all application event data and component metrics from the Traffic Controller over the [Firehose](#) endpoint to external applications or services

For more information about these Loggregator components, see the [Overview of the Loggregator System](#) topic.

## Configure Your Load Balancer for WebSockets

To form a WebSocket connection, the client sends an HTTP request that contains an `Upgrade` header and other headers required to complete the WebSocket handshake. You must configure your load balancer to not upgrade the HTTP request, but rather to pass the `Upgrade` header through to the CF router. The procedures required to configure your load balancer depends on your IaaS and load balancer. The following list includes several possible approaches:

- Some load balancers can recognize the `Upgrade` header and pass these requests through to the CF router without returning the WebSocket handshake response. This may or may not be default behavior, and may require additional configuration.
- Some load balancers do not support passing WebSocket handshake requests containing the `Upgrade` header to the CF router. For instance, the Amazon Web Services (AWS) Elastic Load Balancer (ELB) does not support this behavior. In this scenario, you must configure your load balancer to forward TCP traffic to your CF router to support WebSockets. If your load balancer does not support TCP pass-through of WebSocket requests on the same port as other HTTP requests, you can do one of the following:
  - Configure your load balancer to listen on a non-standard port (the built-in CF load balancer listens on 8443 by default for this purpose), and forward requests for this port in TCP mode. Application clients must make WebSockets upgrade requests to this port. If you choose this strategy, you must follow the steps below in the [Set Your Loggregator Port](#) section of this topic
  - If a non-standard port is not acceptable, add a load balancer that will handle WebSocket traffic (or another IP on an existing load balancer) and configure it to listen on standard ports 80 and 443 in TCP mode. Configure DNS with a new hostname, such as `ws.cf.example.com`, to be used for WebSockets. This hostname should resolve to the new load balancer interface.

 **Note:** Regardless of your IaaS and configuration, you must configure your load balancer to send the X-Forwarded-For and X-Forwarded-Proto headers for non-WebSocket HTTP requests on ports 80 and 443. See the [Securing Traffic into Cloud Foundry](#) topic for more information.

## Set Your Loggregator Port

By default, PCF assigns port 443 for TCP/WebSocket communications. If you have configured your load balancer to use a port other than 443 for TCP/WebSocket traffic, you must edit the **Loggregator Port** field in the **Networking** pane of the PAS tile.

HTTP Headers to Log

HAProxy Request Max Buffer Size \*

HAProxy Protected Domains

HAProxy Trusted CIDRs

Loggregator Port

Default is 443. Enter a new value to override the default, for instance if port 443 on your load balancer is used for other traffic.

## Configuring Load Balancer Healthchecks for Cloud Foundry Routers

This topic describes how to configure load balancer healthchecks for Cloud Foundry (CF) routers to ensure that the load balancer only forwards requests to healthy router instances. You can also configure a healthcheck for your HAProxy if your deployment uses the HAProxy component.

In environments that require high availability, operators must configure their own redundant load balancer to forward traffic directly to the CF routers. In environments that do not require high availability, operators can skip the load balancer and configure DNS to resolve the CF domains directly to a single instance of a router.

### Add Healthcheck Endpoints for Routers

Configure your load balancer to use the following HTTP healthcheck endpoints. Add the IP addresses of all router instances along with their corresponding port and path.

- HTTP Router (Gorouter): `http://GOROUTER_IP:8080/health`
- TCP Router: `http://TCP_ROUTER_IP:80/health`

The configuration above assumes the default healthcheck ports for the CF routers. To modify these ports, see the sections below.

### Add a Healthcheck Endpoint for HAProxy

If you have deployed one or more instances of HAProxy between your infrastructure load balancer and Gorouters, configure your infrastructure load balancer to use the following HTTP healthcheck endpoint: `http://HAPROXY_IP:8080/health`.

The HAProxy is an optional component that provides some features that Gorouter does not and can be helpful for demonstrating horizontal scalability of the CF routers in environments where an infrastructure load balancer is not available.

### Set the Healthy and Unhealthy Threshold Properties for the Gorouter

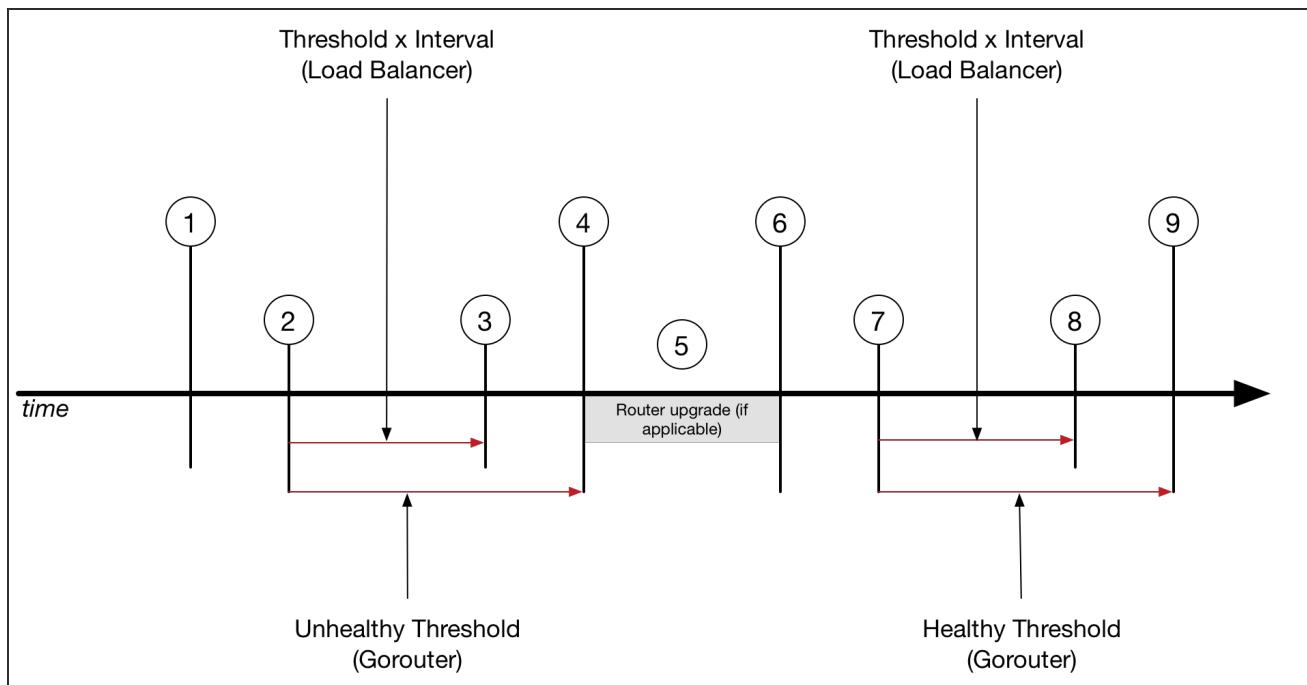
To maintain high availability during upgrades to the HTTP router, each router is upgraded on a rolling basis. During upgrade of a highly available environment with multiple routers, each router is shutdown, upgraded, and restarted before the next router is upgraded. This ensures that any pending HTTP request passed to the HTTP router are handled correctly.

Pivotal Application Service (PAS) uses the following properties:

- **Unhealthy Threshold:** Specifies the amount of time, in seconds, that the Router continues to accept connections before shutting down. During this period, the healthcheck reports `unhealthy` to cause load balancers to fail over to other routers. You should set this value greater than or equal to the maximum amount of time it could take your load balancer to consider a router instance unhealthy, given contiguous failed healthchecks.
- **Healthy Threshold:** Specifies the amount of time, in seconds, to wait until declaring the router instance started. This allows an external load balancer time to register the instance as `healthy`.

You can configure these properties from the **Settings > Network** tab.

The image and table below describe the behavior of the load balancer health checks when a router shuts down and is restarted.



| Step | Description                                                                                                                                                                                                                                                                                                                                                            |
|------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1    | A shutdown request is sent to the router.                                                                                                                                                                                                                                                                                                                              |
| 2    | The router receives shutdown request, which causes the following: <ul style="list-style-type: none"> <li>The router begins sending Service Unavailable responses to the load balancer health checks.</li> <li>The load balancer continues sending HTTP request to the router</li> </ul>                                                                                |
| 3    | The load balancer considers the router to be in an unhealthy state, which causes the load balancer to stop sending HTTP requests to the router.                                                                                                                                                                                                                        |
| 4    | The router shuts down.                                                                                                                                                                                                                                                                                                                                                 |
| 4    | The interval between step 2 and 4 is defined by the Unhealthy Threshold property of the Gorouter. In general, the value of this property should be longer than the value of the interval and threshold values (interval x threshold) of the load balancer. This additional interval ensures that any remaining HTTP requests are handled before the router shuts down. |
| 5    | If the router shutdown is initiated by an upgrade, the Gorouter software is upgraded.                                                                                                                                                                                                                                                                                  |
| 6    | The router restarts. The router will return Service Unavailable responses for load balancer health checks for 20 seconds; during this time the routing table is preloaded.                                                                                                                                                                                             |
| 7    | The routers begins returning Service Available responses to the load balancer health check.                                                                                                                                                                                                                                                                            |
| 8    | The load balancer considers the router to be in a healthy state. The time between step 7 and 8 is specified by the health check interval and threshold configured for your load balancer (health check threshold x health check interval).                                                                                                                             |
| 9    | Shutdown and upgrade of the other router begins.                                                                                                                                                                                                                                                                                                                       |

## Securing Traffic into Cloud Foundry

This topic describes the options for securing HTTP traffic into your Pivotal Application Service (PAS) deployment with TLS certificates. You can configure the location where your deployment terminates TLS depending on your needs and certificate restrictions.

### Protocol Support

The Gorouter supports HTTP/HTTPS requests only. For more information about features supported by the Gorouter, see the [HTTP Routing](#) topic.

You can force HTTPS-only connections by enabling HSTS on HAProxy. For more information, see [Secure Apps Domain with HAProxy](#).

To secure non-HTTP traffic over TCP Routing, terminate TLS at your load balancer or at the application. See [TCP Routing](#) for details.

### TLS Termination Options for HTTP Routing

There are several options for terminating TLS for HTTP traffic. You can terminate TLS at the Gorouter, your load balancer, or at both.

The following table summarizes TLS termination options and which option to choose for your deployment.

 **Note:** To ensure traffic is sent to the platform securely, Pivotal recommends as a minimum that you terminate TLS at the Gorouter. You can optionally terminate TLS at the Load Balancer.

| If the following applies to you:                                                                                                                                                                                                                                                                                           | Then configure TLS termination at:             | Related topic and configuration procedure                         |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------|-------------------------------------------------------------------|
| <ul style="list-style-type: none"> <li>You want the optimum balance of performance and security, and</li> <li>You want to make minimum changes to your load balancer, or</li> <li>You are deploying CF to AWS. For information about AWS limitations, see <a href="#">TLS Cipher Suite Support by AWS ELBs</a>.</li> </ul> | Gorouter only                                  | <a href="#">Terminating TLS at the Gorouter Only</a>              |
| <ul style="list-style-type: none"> <li>You require TLS termination at a load balancer, or</li> <li>You want the highest level of security, and</li> <li>You do not mind a slightly less performant deployment.</li> </ul>                                                                                                  | Load Balancer and Gorouter                     | <a href="#">Terminating TLS at the Load Balancer and Gorouter</a> |
| <ul style="list-style-type: none"> <li>You require TLS termination at a load balancer, and</li> <li>You prefer unencrypted traffic between the Load Balancer and the Gorouter.</li> </ul>                                                                                                                                  | Load Balancer only                             | <a href="#">Terminating TLS at the Load Balancer Only</a>         |
| <b>Optionally, if you are deploying HAProxy, and</b>                                                                                                                                                                                                                                                                       | <b>Then in addition, terminate SSL/TLS at:</b> | <b>Related topic and configuration procedure</b>                  |
| <ul style="list-style-type: none"> <li>You would like to secure traffic to the HAProxy.</li> </ul>                                                                                                                                                                                                                         | HAProxy                                        | <a href="#">Terminating SSL/TLS at HAProxy</a>                    |

### Certificate Requirements

The following requirements apply to the certificates you use to secure traffic into PAS

- You must obtain at least one TLS certificate for your environment.
  - In a production environment, use a signed TLS certificate (trusted) from a known certificate authority (CA).
  - In a development or testing environment, you can use a trusted CA certificate or a self-signed certificate. You can generate a self-signed certificate with `openssl` or a similar tool.

 Alternately, you can use the PAS Ops Manager interface to [generate a certificate](#) for you. Certificates generated in PAS are signed by the Ops Manager Certificate Authority. They are not technically self-signed, but they are sometimes referred to as “Self-Signed Certificates” in the Ops Manager UI and throughout this documentation.

- Certificates used in CF must be encoded in the PEM format.

- The Gorouter supports mutual TLS, and validates a client provided certificate chain against its CA certificates if one is provided in the TLS handshake, but does not require it. Depending on whether you choose to terminate at both the Load Balancer and the Gorouter, or at the Gorouter alone, the client certificate may be that of the load balancer or of the originating client.
- The certificate on the Gorouter must be associated with the correct hostname so that HTTPS can validate the request.
- If wildcard certificates are not supported for some or all of your domains, then configure termination requests at the [load balancer only](#). In this type of deployment, the load balancer passes unencrypted traffic to the Gorouter. As a result, you avoid having to reissue and reinstall certificates on the Gorouter for every app or UAA security zone.
- Extended Validation (EV) certificates support multiple hostnames, like SAN, but do not support wildcards. As the Gorouter has not been tested with EV certificates, if EV certificates are required, then terminate TLS at the [load balancer only](#).
- Given the dynamic and multi-tenant nature of PAS, use of wildcard domains is highly recommended to avoid the need for adding an additional certificate for each application.

## Multiple Certificates

In order to support custom domains on CF, an operator has to configure the Gorouter with a certificate that represents the domain. It is recommended that operators add a new certificate instead of reissuing a single certificate when adding TLS support for an additional domain. Using multiple certificates provides a security benefit in that it prevents clients from discovering all the custom domains of applications running on a CF platform.

The Gorouter supports SNI and can be configured with multiple certificates, each which may optionally include wildcard and alternative names. The Gorouter uses SNI to determine the correct certificate to present in a TLS handshake. It requires clients to support the SNI protocol by sending a server name outside the encrypted request payload. For clients that do not support SNI, the Gorouter presents a default certificate. The default is the first certificate keypair in the Gorouter's configuration.

The Gorouter decides which certificate to provide in the TLS handshake as follows:

- If a client provides an SNI header with a ServerName that matches to a configured certificate keypair, the Gorouter returns the matching certificate.
- If a client provides an SNI header with a ServerName that does not match a configured certificate keypair, the Gorouter returns the default certificate.

The first certificate keypair listed is used as the default.

The Gorouter supports both RSA and ECDSA certificates in PEM encoding. In the case that a certificate chain is required, the order should be as follows: primary certificate, intermediate certificate, then root certificate.

## How to Configure Multiple Certificate Keypairs

To configure multiple HTTPS certificate keypairs for PAS, add each keypair along with a meaningful name in the applicable **Certificates and Private Keys for HAProxy and Router** fields of the **Networking** configuration screen in PAS. For more information, see [Configuring PAS](#).

In PCF, multiple certificates configured for the Gorouter are also configured for HAProxy.

## TLS Cipher Suite Support

Some CF components like the Gorouter support additional TLS cipher suites to accommodate older clients. As a security best practice, only configure the TLS cipher suites that you need for your deployment.

## Default Gorouter Cipher Suites

By default, the Gorouter supports the following TLS cipher suites, both of which require TLS v1.2:

| RFC                                   | OpenSSL                     |
|---------------------------------------|-----------------------------|
| TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256 | ECDHE-RSA-AES128-GCM-SHA256 |
| TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384 | ECDHE-RSA-AES256-GCM-SHA384 |

You can override the default cipher suites in the **TLS Cipher Suites for Router** and **Minimum version of TLS** fields in the **Networking** tab of the PAS tile. See the following procedures for either [Gorouter Only](#) or [Load Balancer and Gorouter](#) for more information about using custom SSL ciphers.

## TLS Cipher Suite Support by AWS Load Balancers

AWS Classic Load Balancers (formerly referred to as ELBs) support configuration of cipher suites for front-end connections with clients only. When configuring Classic Load Balancers to forward requests to Gorouters over TLS, operators may encounter a “Cipher Suite Mismatch” error. This is because the cipher suites supported by Classic Load Balancers for TLS handshakes with backends (Gorouters in this case) are hardcoded, undocumented, and do not support the Gorouter default cipher suites.

Operators have two options:

- Configure Classic Load Balancer listeners in TCP mode so that TCP connections from clients are passed through the Classic Load Balancer to Gorouters on port 443. Then Gorouters are the first point of TLS termination.
- If you require TLS termination at an AWS load balancer in addition to terminating at the Gorouter, use [AWS Application Load Balancers \(ALBs\)](#) that support the Gorouter default cipher suites.

## TLS v1.2

The following cipher suites are optionally supported for TLS v1.2 only:

| RFC                                     | OpenSSL                       |
|-----------------------------------------|-------------------------------|
| TLS_RSA_WITH_AES_128_GCM_SHA256         | AES128-GCM-SHA256             |
| TLS_RSA_WITH_AES_256_GCM_SHA384         | AES256-GCM-SHA384             |
| TLS_ECDHE_ECDSA_WITH_RC4_128_SHA        | ECDHE-ECDSA-RC4-SHA           |
| TLS_ECDHE_ECDSA_WITH_AES_128_CBC_SHA    | ECDHE-ECDSA-AES128-SHA        |
| TLS_ECDHE_ECDSA_WITH_AES_256_CBC_SHA    | ECDHE-ECDSA-AES256-SHA        |
| TLS_ECDHE_RSA_WITH_RC4_128_SHA          | ECDHE-RSA-RC4-SHA             |
| TLS_ECDHE_RSA_WITH_3DES_EDE_CBC_SHA     | ECDHE-RSA-DES-CBC3-SHA        |
| TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA      | ECDHE-RSA-AES128-SHA          |
| TLS_ECDHE_RSA_WITH_AES_256_CBC_SHA      | ECDHE-RSA-AES256-SHA          |
| TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256   | ECDHE-RSA-AES128-GCM-SHA256   |
| TLS_ECDHE_ECDSA_WITH_AES_128_GCM_SHA256 | ECDHE-ECDSA-AES128-GCM-SHA256 |
| TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384   | ECDHE-RSA-AES256-GCM-SHA384   |
| TLS_ECDHE_ECDSA_WITH_AES_256_GCM_SHA384 | ECDHE-ECDSA-AES256-GCM-SHA384 |
| TLS_RSA_WITH_AES_128_CBC_SHA256         | AES128-SHA256                 |
| TLS_ECDHE_ECDSA_WITH_AES_128_CBC_SHA256 | ECDHE-ECDSA-AES128-SHA256     |
| TLS_ECDHE_RSA_WITH_AES_128_CBC_SHA256   | ECDHE-RSA-AES128-SHA256       |
| TLS_ECDHE_RSA_WITH_CHACHA20_POLY1305    | ECDHE-RSA-CHACHA20-POLY1305   |
| TLS_ECDHE_ECDSA_WITH_CHACHA20_POLY1305  | ECDHE-ECDSA-CHACHA20-POLY1305 |

## TLS v1.0 and v1.1

The following cipher suites are optionally supported for TLS v1.0 and TLS v1.1 only:

| RFC                           | OpenSSL      |
|-------------------------------|--------------|
| TLS_RSA_WITH_RC4_128_SHA      | RC4-SHA      |
| TLS_RSA_WITH_3DES_EDE_CBC_SHA | DES-CBC3-SHA |
| TLS_RSA_WITH_AES_128_CBC_SHA  | AES128-SHA   |
| TLS_RSA_WITH_AES_256_CBC_SHA  | AES256-SHA   |

You can override the default cipher suites in the [TLS Cipher Suites for Router](#) and [Minimum version of TLS](#) fields in the [Networking](#) tab of the PAS tile. See the following procedures for either [Gorouter Only](#) or [Load Balancer and Gorouter](#) for more information about using custom SSL ciphers.

See [Golang Constants](#) and [OpenSSL Cipher Suites](#) for more information about supported ciphers.

**Note:** ECDSA ciphers require a certificate and key for DSA, as opposed to RSA.

## Mutual Authentication with Clients

Gorouter supports validation of client certificates in TLS handshakes with clients, also known as mutual authentication. Operators can choose whether Gorouter requests client certificates and when requesting certificates, whether or not to require them.

By default, Gorouter requests but does not require client certificates in TLS handshakes.

To configure Gorouter behavior for handling client certificates, select one of the options in the **Router behavior for Client Certificates** field of the **Networking** configuration screen in PAS.

- **Router does not request client certificates.** The Gorouter does not request client certificates in TLS handshakes so clients will not provide them and validation of client certificates does not occur. This option is incompatible with the XFCC configuration options **TLS terminated for the first time at HAProxy** and **TLS terminated for the first time at the Router** in PAS because these options require mutual authentication.
- **Router requests but does not require client certificates.** The Gorouter requests client certificates in TLS handshakes. The handshake will fail if the client certificate is not signed by a CA configured for the router. This is the default configuration.
- **Router requires client certificates.** The Gorouter requests and requires client certificates in TLS handshakes. The handshake will fail if a client cert is not provided or if the client certificate is not signed by a CA configured for the router.

The behavior controlled by this property is global; it applies to all requests received by Gorouters so configured.

If Gorouter is the first point of TLS termination (your load balancer does not terminate TLS, and passes the request through to Gorouter over TCP), consider the following:

- Only option **Router does not request client certificates** should be used with PAS, as the Gorouters are in that product receive requests for the system domain. Many clients of CF platform APIs do not present client certificates in TLS handshakes, so the first point of TLS termination for requests to the system domain must not request them.
- All options may be used for routers deployed with the Isolation Segment tile, as these only receive requests for app domains.
- Options **Router requests but does not require client certificates** and **Router requires client certificates** will trigger browsers to prompt users to select a certificate if the browser is not already configured with a certificate signed by one of the CAs configured for the router.

If Gorouter is not the first point of TLS termination, this property can be used to secure communications between the Load Balancer and Gorouter. The router must be configured with the CA used to sign the client certification the load balancer will present.

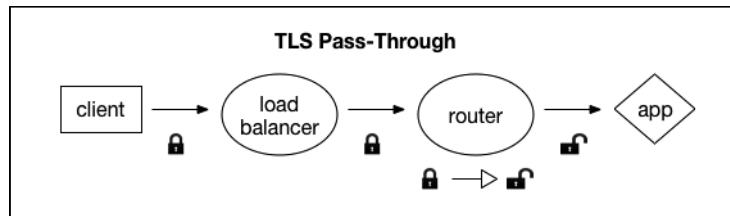
**warning:** Requests to the platform will fail upon upgrade if your load balancer is configured to present a client certificate in the TLS handshake with Gorouter but Gorouter has not been configured with the certificate authority used to sign it. To mitigate this issue, select **Router does not request client certificates** for **Router behavior for Client Certificate Validation** in the **Networking** pane or configure the router with the appropriate CA.

## Terminating TLS at the Gorouter Only

In this configuration, the load balancer does not terminate TLS for CF domains at all. Instead, it passes through the underlying TCP connection to the Gorouter.

This option is the recommended and more performant option, establishing and terminating a single TLS connection.

The following diagram illustrates communication between the client, load balancer, Gorouter, and app.



Traffic between the load balancer and the Gorouter is encrypted only if the client request is encrypted.

## About HTTP Header Forwarding

If you terminate TLS at the Gorouter only, your load balancer does not send HTTP headers.

The Gorouter appends the `X-Forwarded-For` and `X-Forwarded-Proto` headers to requests forwarded to applications and platform system components.

`X-Forwarded-For` is set to the IP address of the source. Depending on the behavior of your load balancer, this may be the IP address of your load balancer. For Gorouter to deliver the IP address of the client to applications, configure your load balancer to forward the IP address of the client or configure your load balancer to send the client IP address using the PROXY protocol.

`X-Forwarded-Proto` provides the scheme of the HTTP request from the client. The scheme is HTTP if the client made a request on port 80 (unencrypted) or HTTPS if the client made a request on port 443 (encrypted). Gorouter sanitizes the `X-Forwarded-Proto` header based on the settings of the `router.sanitize_forwarded_proto` and `router.force_forwarded_proto_https` manifest properties as follows:

- `router.sanitize_forwarded_proto: true` and `router.force_forwarded_proto_https: true` :  
Gorouter sets the value of `X-Forwarded-Proto` header to `HTTPS` in requests forwarded to backends.
- `router.sanitize_forwarded_proto: true` and `router.force_forwarded_proto_https: false` :  
Gorouter strips the `X-Forwarded-Proto` header when present in requests from front-end clients. When the request is received on port 80 (unencrypted), Gorouter sets the value of this header to `HTTP` in requests forwarded to backends. When the request is received on port 443 (encrypted), Gorouter sets the value of this header to `HTTPS`.
- `router.sanitize_forwarded_proto: false` and `router.force_forwarded_proto_https: true` :  
Gorouter passes through the header as received from the load balancer, without modification.
- `router.sanitize_forwarded_proto: false` and `router.force_forwarded_proto_https: false` :  
Gorouter passes through the header as received from the load balancer, without modification.

 **Note:** If Gorouter is the first point of TLS, we recommend setting `router.sanitize_forwarded_proto: true` and `router.force_forwarded_proto_https: false` to secure against header spoofing.

For more information about HTTP headers in Cloud Foundry, see the [HTTP Headers](#) section of [HTTP Routing](#). For information about configuring the forwarding of client certificates, see the [Forward Client Certificate to Applications](#) section of [HTTP Routing](#).

## Procedure: Gorouter Only

Perform the following steps to configure SSL termination on the Gorouter in Pivotal Cloud Foundry (PCF):

1. Configure your load balancer to pass through TCP requests from the client to the Gorouter.
2. Navigate to the Ops Manager Installation Dashboard.
3. Click the Pivotal Application Service (PAS) tile in the Installation Dashboard.
4. Click **Networking**.
5. For PCF deployments on OpenStack or vSphere, choose IP addresses for the Gorouters from the subnet configured for Ops Manager and enter them in the **Router IPs** field. Then configure your load balancer to forward requests for the above domains to these IP addresses. For more information, see [Configuring PAS](#).
6. In the **Certificates and Private Keys for HAProxy and Routerfield**, click the **Add** button to define at least one certificate keypair for HAProxy and Router. For each certificate keypair that you add, assign a name, enter the PEM-encoded certificate chain and PEM-encoded private key. You can either upload your own certificate or generate an RSA certificate in PAS. For options and instructions on creating a certificate for your wildcard domains, see [Creating a Wildcard Certificate for PCF Deployments](#).
7. In the **Minimum version of TLS supported by HAProxy and Router**, select the minimum version of TLS to use in Gorouter communications. The Gorouter uses TLS v1.2 by default. If you need to accommodate clients that use an older version of TLS, select a lower minimum version. For a list of TLS ciphers supported by the Gorouter, see [Cipher Suites](#).
8. Under **HAProxy forwards requests to Router over TLS**, select **Disable**.
9. If you want to use a specific set of TLS ciphers for the Gorouter, configure **TLS Cipher Suites for Router**. Enter an ordered, colon-separated list of TLS cipher suites in the OpenSSL format. For example, if you have selected support for an earlier version of TLS, you can enter cipher suites supported by this version. For a list of TLS ciphers supported by the Gorouter, see [Cipher Suites](#). Otherwise, leave the default values in this field.
10. (Optional) If you are not using SSL encryption or if you are using self-signed certificates, you can select **Disable SSL certificate verification for this environment**. Selecting this checkbox also disables SSL verification for route services.

 Use this checkbox only for development and testing environments. Do not select it for production environments.

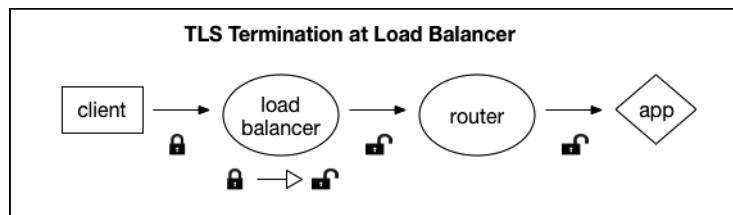
11. (Optional) If you do not want the Gorouter to accept any non-encrypted HTTP traffic, select the **Disable HTTP on HAProxy and Router** checkbox.
12. In the **Configure the CF Router support for the X-Forwarded-Client-Cert headerfield**, select the third option, **Strip the XFCC header when present and set it to the client certificate**.
13. Click **Save**.
14. In the PAS tile, click **Resource Config**.
15. In the **Instances** drop down for the **HAPRoxy** job, select **0** instances.
16. Click **Save**.

## Terminating TLS at the Load Balancer Only

In this configuration, your load balancer terminates TLS, and passes unencrypted traffic to the Gorouter, which routes it to your app. Traffic between the load balancer and the Gorouter is not encrypted.

This option is recommended if you cannot use SAN certificates and if you do not require traffic to be encrypted between the load balancer and the Gorouter.

The following diagram illustrates communication between the client, load balancer, Gorouter, and app.



## About HTTP Header Forwarding

If you terminate TLS at your load balancer, then you must also configure the load balancer to append the **X-Forwarded-For** and **X-Forwarded-Proto** HTTP headers to the HTTP traffic it passes to the Gorouter.

For more information about HTTP headers in CF, see [HTTP Headers](#). If you are configuring the forwarding of client certificates, see [Forward Client Certificate to Applications](#).

## Procedure: Load Balancer Only

Perform the following steps to configure SSL termination on the load balancer only in Pivotal Cloud Foundry (PCF):

1. Create an A record in your DNS that points to your load balancer IP address. The A record associates the **System Domain** and **Apps Domain** that you configure in the **Domains** section of the Pivotal Application Service (PAS) tile with the IP address of your load balancer.

For example, with **cf.example.com** as the main subdomain for your Cloud Foundry deployment and a load balancer IP address **198.51.100.1**, you must create an A record in your DNS that serves **example.com** and points **\*.cf** to **198.51.100.1**.

| Name        | Type | Data         | Domain      |
|-------------|------|--------------|-------------|
| <b>*.cf</b> | A    | 198.51.100.1 | example.com |

2. Navigate to the Ops Manager Installation Dashboard.
3. Click the PAS tile in the Installation Dashboard.
4. Click **Networking**.
5. For PCF deployments on OpenStack or vSphere, choose IP addresses for the Gorouters from the subnet configured for Ops Manager and enter them

in the **Router IPs** field. Then configure your load balancer to forward requests for the above domains to these IP addresses. For more information, see [Configuring PAS](#)

6. In the **Certificates and Private Keys for HAProxy and Routerfield**, click the **Add** button to define one certificate keypair for HAProxy and Router. Since you have opted for unencrypted traffic behind the load balancer, then you can generate an RSA certificate in PAS.
7. In the **Minimum version of TLS supported by HAProxy and Router**, select the minimum version of TLS to use in HAProxy communications. HAProxy use TLS v1.2 by default. If you need to accommodate clients that use an older version of TLS, select a lower minimum version. For a list of TLS ciphers supported by the HAProxy, see [Cipher Suites](#).
8. Under **HAProxy forwards requests to Router over TLS**, select **Disable**.
9. If you want to use a specific set of TLS ciphers for HAProxy, configure **TLS Cipher Suites for HAProxy**. Enter an ordered, colon-separated list of TLS cipher suites in the OpenSSL format. For example, if you have selected support for an earlier version of TLS, you can enter cipher suites supported by this version. Otherwise, leave the default values in this field.
10. (Optional) If you are not using SSL encryption or if you are using self-signed certificates, you can select **Disable SSL certificate verification for this environment**. Selecting this checkbox also disables SSL verification for route services.

 Use this checkbox only for development and testing environments. Do not select it for production environments.

11. (Optional) If you do not want HAProxy or the Gorouter to accept any non-encrypted HTTP traffic, select the **Disable HTTP on HAProxy and Router** checkbox.
12. In the **Configure the CF Router support for the X-Forwarded-Client-Cert header** field, select **Always forward the XFCC header in the request, regardless of the whether the client connection is mTLS**.
13. Click **Save**.
14. After you complete the configuration in PCF, add your certificate or certificates to your load balancer, and configure its listening port. The procedures vary depending on your IaaS.
15. Configure your load balancer to append the **X-Forwarded-For** and **X-Forwarded-Proto** headers to client requests.

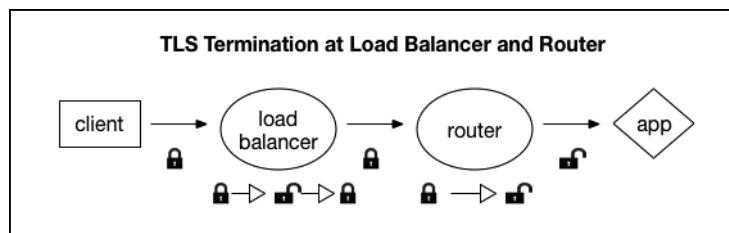
If the load balancer cannot be configured to provide the **X-Forwarded-For** header, the Gorouter will append it in requests forwarded to applications and system components, set to the IP address of the load balancer.

 **Note:** If the load balancer accepts unencrypted requests, it **must** provide the X-Forwarded-Proto header. Conversely, if the load balancer cannot be configured to send the X-Forwarded-Proto header, it should not accept unencrypted requests. Otherwise, applications and platform system components that require encrypted client requests will accept unencrypted requests when they should not accept them.

## Terminating TLS at the Load Balancer and Gorouter

In this configuration two TLS connections are established: one from the client to the load balancer, and another from the load balancer to the Gorouter. This configuration secures all traffic between the load balancer and the Gorouter.

The following diagram illustrates communication between the client, load balancer, Gorouter, and app.



This option is less performant, but allows for termination at a load balancer, as well as secure traffic between the load balancer and the Gorouter.

## Certificate Guidelines

In this deployment scenario, the following guidelines apply:

- Certificates for the PAS domains must be stored on the load balancer, as well as on the Gorouter.
- Generate certificates for your load balancer and the Gorouter with different keys. If the key for the certificate on the Gorouter is compromised, then the certificate on the load balancer is not at risk, and vice versa.
- If you choose to host only one certificate on the Gorouter and many on your load balancer, configure your load balancer with the CA and hostname with which to validate the certificate hosted by the Gorouter.

## About Hostname Verification

Hostname verification between the load balancer and the Gorouter is unnecessary when the load balancer is already configured with the Gorouter's IP address to correctly route the request.

If the load balancer uses DNS resolution to route requests to the Gorouters, then you should enable hostname verification.

## About HTTP Header Forwarding

If you terminate TLS at your load balancer, then you must configure the load balancer to append the `X-Forwarded-For` and `X-Forwarded-Proto` HTTP headers to requests it sends to the Gorouter.

If you terminate TLS at your load balancer but it does not support HTTP, such that it cannot append HTTP headers, a workaround exists. We recommend you use this workaround **only if your load balancer does not accept unencrypted requests**. Configure your load balancer to send the client IP address using the PROXY protocol, and enable PROXY in the Gorouter. As the `X-Forwarded-Proto` header will not be present, configure the Gorouter to force-set this header to 'HTTPS'.

For more information about HTTP headers in CF, see [HTTP Headers](#). If you are configuring the forwarding of client certificates, see [Forward Client Certificate to Applications](#).

## Procedure: Load Balancer and Gorouter

Perform the following steps to configure SSL termination on the Gorouter and load balancer in Pivotal Cloud Foundry (PCF):

1. Create an A record in your DNS that points to your load balancer IP address. The A record associates the **System Domain** and **Apps Domain** that you configure in the **Domains** section of the Pivotal Application Service (PAS) tile with the IP address of your load balancer.

For example, with `cf.example.com` as the main subdomain for your Cloud Foundry (CF) deployment and a load balancer IP address `198.51.100.1`, you must create an A record in your DNS that serves `example.com` and points `*.cf` to `198.51.100.1`.

| Name              | Type | Data                      | Domain                   |
|-------------------|------|---------------------------|--------------------------|
| <code>*.cf</code> | A    | <code>198.51.100.1</code> | <code>example.com</code> |

2. Navigate to the Ops Manager Installation Dashboard.
3. Click the PAS tile in the Installation Dashboard.
4. Click **Networking**.
5. For PCF deployments on OpenStack or vSphere, choose IP addresses for the Gorouters from the subnet configured for Ops Manager and enter them in the **Router IPs** field. Then configure your load balancer to forward requests for the above domains to these IP addresses. For more information, see [Configuring PAS](#).
6. In the **Certificates and Private Keys for HAProxy and Router** field, click the **Add** button to define at least one certificate keypair for HAProxy and Router. For each certificate keypair that you add, assign a name, enter the PEM-encoded certificate chain and PEM-encoded private key. You can either upload your own certificate or generate an RSA certificate in PAS. For options and instructions on creating a certificate for your wildcard domains, see [Creating a Wildcard Certificate for PCF Deployments](#).
7. In the **Minimum version of TLS supported by HAProxy and Router**, select the minimum version of TLS to use in HAProxy and Gorouter communications. The Gorouter use TLS v1.2 by default. If you need to accommodate clients that use an older version of TLS, select a lower minimum version. For a list of TLS ciphers supported by the Gorouter, see [Cipher Suites](#).
8. If you are using **HAProxy**, complete the following steps:
  - a. Under **HAProxy forwards requests to Router over TLS**, select **Enable**.
  - b. In the **Certificate Authority for HAProxy Backend** field, specify the Certificate Authority (CA) that signed the certificate you configured in the

Certificate and Private Key for HAProxy and Router field.

 If you used the [Generate RSA Certificate](#) link to generate a self-signed certificate, then the CA to specify is the Ops Manager CA, which you can locate at the `/api/v0/certificateAuthorities` endpoint in the Ops Manager API.

- c. If you want to use a specific set of TLS ciphers for HAProxy, configure **TLS Cipher Suites for HAProxy**. Enter an ordered, colon-separated list of TLS cipher suites in the OpenSSL format. For example, if you have selected support for an earlier version of TLS, you can enter cipher suites supported by this version. Otherwise, leave the default values in this field.
  - d. In the **Configure the CF Router support for the X-Forwarded-Client-Cert header** field, select **Always forward the XFCC header in the request, regardless of the whether the client connection is mTLS**.
  - e. Proceed to step 11.
9. If you want to use a specific set of TLS ciphers for the Gorouter, configure **TLS Cipher Suites for Router**. Enter an ordered, colon-separated list of TLS cipher suites in the OpenSSL format. For example, if you have selected support for an earlier version of TLS, you can enter cipher suites supported by this version. For a list of TLS ciphers supported by the Gorouter, see [Cipher Suites](#). Otherwise, leave the default values in this field.
  10. If you are not using HAProxy, complete the following steps:
    - a. Under **HAProxy forwards requests to Router over TLS**, select **Disable**.
    - b. In the **Configure the CF Router support for the X-Forwarded-Client-Cert header** field, select any of the available options depending on your client application needs. For more information about XFCC header forwarding, see [Forward Client Certificate to Applications](#).
    - c. In the PAS tile, click **Resource Config**.
    - d. In the **Instances** drop down for the **HAProxy** job, select **0** instances.
    - e. Click **Save**.
  11. (Optional) If you are not using SSL encryption or if you are using self-signed certificates, you can select **Disable SSL certificate verification for this environment**. Selecting this checkbox also disables SSL verification for route services.

 Use this checkbox only for development and testing environments. Do not select it for production environments.
  12. (Optional) If you do not want HAProxy or the Gorouter to accept any non-encrypted HTTP traffic, select the **Disable HTTP on HAProxy and Router** checkbox.
  13. Click **Save**.
  14. After you complete the configuration in PCF, add your certificate or certificates to your load balancer, and configure its listening port. The procedures vary depending on your IaaS.
  15. Configure your load balancer to append the **X-Forwarded-For** and **X-Forwarded-Proto** headers to client requests.

If you cannot configure the load balancer to provide the **X-Forwarded-For** header, the Gorouter appends it in requests forwarded to applications and system components, set to the IP address of the load balancer.

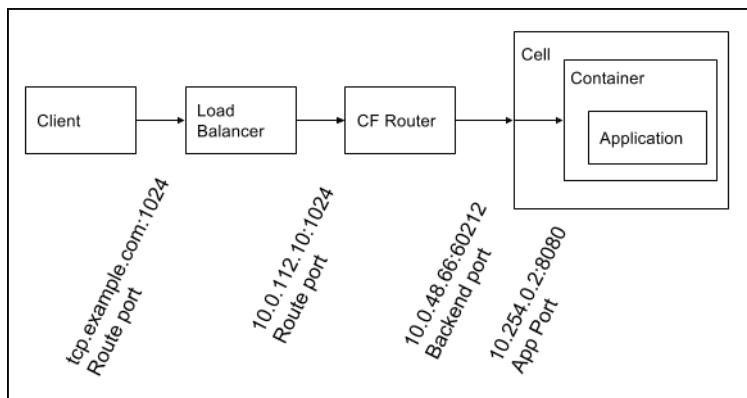
 Note: If the load balancer accepts unencrypted requests, it **must** provide the **X-Forwarded-Proto** header. Conversely, if the load balancer cannot be configured to send the **X-Forwarded-Proto** header, it should not accept unencrypted requests. Otherwise, applications and platform system components that require encrypted client requests will accept unencrypted requests when they should not accept them.

## Enabling TCP Routing

This topic describes enabling TCP Routing for your Cloud Foundry (CF) deployment. This feature enables developers to run applications that serve requests on non-HTTP TCP protocols. You can use TCP Routing to comply with regulatory requirements that require your organization to terminate the TLS as close to your apps as possible so that packets are not decrypted before reaching the application level.

## Route Ports

The diagram below shows the layers of network address translation that occur in Cloud Foundry in support of TCP Routing. The descriptions step through an example work flow that covers route ports, backend ports, and app ports.



- A developer creates a TCP route for their application based on a TCP domain and a route port, and maps this route to one or more applications. See the [Creating Routes](#) topic for more information.
- Clients make requests to the route. DNS resolves the domain name to the load balancer.
- The load balancer listens on the port and forwards requests for the domain to the TCP routers. The load balancer must listen on a range of ports to support multiple TCP route creation. Additionally, Cloud Foundry must be configured with this range, so that the platform knows what ports can be reserved when developers create TCP routes.
- The TCP router can be dynamically configured to listen on the port when the route is mapped to an application. The domain the request was originally sent to is no longer relevant to the routing of the request to the application. The TCP router keeps a dynamically updated record of the backends for each route port. The backends represent instances of an application mapped to the route. The TCP Router chooses a backend using a round-robin load balancing algorithm for each new TCP connection from a client. As the TCP Router is protocol agnostic, it does not recognize individual requests, only TCP connections. All client requests transit the same connection to the selected backend until the client or backend closes the connection. Each subsequent connection triggers the selection of a backend.
- Because containers each have their own private network, the TCP router does not have direct access to application containers. When a container is created for an application instance, a port on the Cell VM is randomly chosen and iptables are configured to forward requests for this port to the internal interface on the container. The TCP router then receives a mapping of the route port to the Cell IP and port.
- By default, the Diego Cell only routes requests to port 8080, the App Port, on the container internal interface. The App Port is the port on which applications must listen. Developers can use the Cloud Controller API to update the ports an app can receive requests on. For more information, see [Configuring Apps to Listen on Custom Ports \(Beta\)](#).

## Pre-Deployment Steps

**Note:** If you have mutual TLS app identity verification enabled, app containers accept incoming communication only from the Gorouter. This disables TCP routing.

Before enabling TCP Routing, you must complete the following steps to set up networking requirements.

- Choose a domain name from which your developers will create TCP routes for their applications. For example, create a domain which is similar to your app domain but prefixed by the TCP subdomain: `tcp.APPS-DOMAIN.com`
- Configure DNS to resolve this domain name to the IP address of a highly-available load balancer that will forward traffic for the domain to the TCP routers. For more information, view the [Domains](#) topic. If you are operating an environment that does not require high-availability, configure DNS to resolve the TCP domain name you have chosen directly to a single instance of the TCP Router.

3. (Optional) Choose IP addresses for the TCP routers and configure your load balancer to forward requests for the domain you chose in the step above to these addresses. Skip this step if you have configured DNS to resolve the TCP domain name to an instance of the TCP Router. Review the Enable TCP Routing steps in the [Configuring PAS](#) topic to configure your IP addresses for your PCF deployment.

4. (Optional) Decide on the number of TCP routes you want to support. For each TCP route, you must reserve a port. Configure your load balancer to forward the range of ports to the TCP routers. Skip this step if you have configured DNS to resolve the TCP domain name to an instance of the TCP Router.

5. Review the “Enable TCP Routing” steps in the [Configuring PAS](#) topic for your IaaS to configure your ports for your PCF deployment.

## Post-Deployment Steps

In the following steps you use the [Cloud Foundry Command Line Interface](#) (cf CLI) to add the TCP shared domain and configure organization quotas to grant developers the ability to create TCP routes. This requires an admin user account.

### Configure CF with Your TCP Domain

After deployment, you must configure Cloud Foundry with the domain that you configured in the Pre-Deployment step above so that developers can create TCP routes from it.

1. Run `cf router-groups`. You should see `default-tcp` as a response.

2. Run `cf create-shared-domain` to create a shared domain and associate it with the router group.

```
$ cf create-shared-domain tcp.APPS-DOMAIN.com --router-group default-tcp
```

3. Run `cf domains`. Verify that next to your TCP domain, `TCP` appears under `type`.

### Configure a Quota for TCP Routes

Since TCP route ports are a limited resource in some environments, quotas are configured to allow creation of zero TCP routes by default. After you deploy Cloud Foundry, you can increase the maximum number of TCP routes for all organizations or for particular organizations and spaces. Because you reserve a route port for each TCP route, the quota for this resource is managed with the cf CLI command option `--reserved-route-ports`. See the [Creating and Modifying Quota Plans](#) topic for more information.

If you have a default quota that applies to all organizations, you can update it to configure the number of route ports that can be reserved by each organization.

```
$ cf update-quota QUOTA --reserved-route-ports X
```

To create a new organization quota that can be allocated to particular organizations, provide the following required quota attributes in addition to the number of reserved route ports:

```
$ cf create-quota QUOTA --reserved-route-ports X
```

You can also create a quota that governs the number of route ports that can be created in a particular space.

```
$ cf create-space-quota QUOTA --reserved-route-ports X
```

## Create a TCP Route

For instructions about creating a TCP Route, see the [Create a TCP Route with a Port](#) topic.

## Router Groups

In [Post-Deployment Steps](#), we describe that in order to create a domain from which to create TCP routes, it must be associated with the TCP Router Group. A router group is a cluster of identically configured routers. Router Groups were introduced as mechanism to support reservation of the same port for multiple TCP routes, thus increasing the capacity for TCP routes. However, only one router group is currently supported. In the [Pre-Deployment Steps](#), we describe how an admin user can configure the port range available for TCP routes in preparation for deployment.

## Modify your TCP ports

After deployment, you can modify the range of ports available for TCP routes using `cf curl` commands, as demonstrated with the commands below. These commands require an admin user with the `routing.router_groups.read` and `routing.router_groups.write` scopes.

1. In a terminal window, run `cf curl /routing/v1/router_groups` to view the `reservable_ports`:

```
$ cf curl /routing/v1/router_groups
[
 {
 "guid": "9d1c1da9-ed63-45e8-45ee-256f8579455c",
 "name": "default-tcp",
 "type": "tcp",
 "reservable_ports": "60000-60098"
 }
]
```

2. Modify the `reservable_ports`:

```
$ cf curl /routing/v1/router_groups/f7392031-a488-4890-8835-c4a038a3bded -X PUT -d '{"reservable_ports":"1024-1199"}'
```

## Service Mesh (Beta)

This topic describes service mesh for Pivotal Application Service (PAS).

To deploy service mesh, see [Deploying Service Mesh \(Beta\)](#).

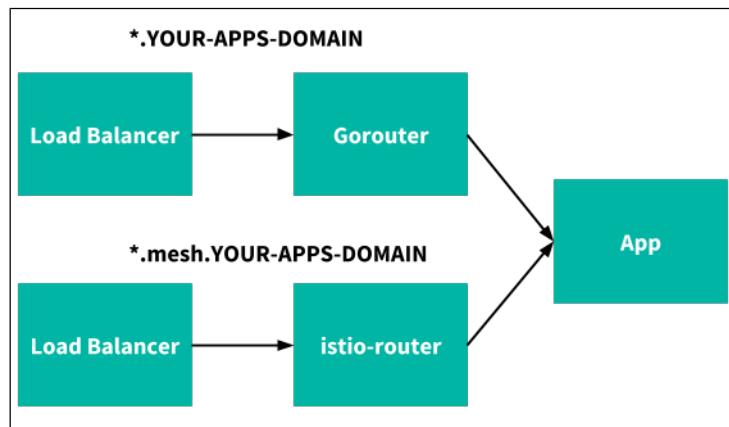
## Overview

PAS includes an optional, beta routing plane that uses a service mesh. A service mesh provides traffic management, security, and observability for microservices. For more information, see [What is a service mesh?](#) in the Istio documentation.

Service mesh in PAS uses Istio Pilot and Envoy. The Cloud Foundry [istio-release](#) packages these components into a BOSH release. For more information, see the following:

- The [Pilot](#) section in Istio documentation.
- The [What is Envoy](#) topic in the Envoy documentation.
- The [istio-release](#) repository in GitHub.

Service mesh deploys an additional router and runs as a parallel routing plane as illustrated in the following diagram:



## Features

Service mesh currently supports configuring routing weights for app. For more information, see [Using Weighted Routing \(Beta\)](#).

## Limitations

Consider the following when deploying service mesh:

- It does not have feature parity with the existing routing plane in PAS.
- It is for deployments with fewer than 20,000 routes. At greater scale, it may impact core platform functions.
- The control plane is not HA and registration of new routes may be delayed during upgrade.
- The domain for routes is [\\*.mesh.YOUR-APPS-DOMAIN](#) and is not configurable.

## Component VMs

The following table describes each component VM deployed as part of service mesh in PAS, along with their function.

| VM            | Processes                | Function                                                                                                 |
|---------------|--------------------------|----------------------------------------------------------------------------------------------------------|
| istio-router  | envoy                    | A reverse proxy to forward HTTP/HTTPS requests external to the platform to applications on the platform. |
| istio-control | copilot, pilot-discovery | Propagates PAS external routes to all service mesh routers.                                              |

|              |                 |                                                                                          |
|--------------|-----------------|------------------------------------------------------------------------------------------|
| route-syncer | cc-route-syncer | Syncs routes created through the Cloud Controller API to the service mesh control plane. |
|--------------|-----------------|------------------------------------------------------------------------------------------|

For information about the communication ports and protocols used by these components, see [Service Mesh](#) in the *Routing Network Communications* topic.

## Isolation Segments

The following topics provide information about managing isolation segments in Pivotal Application Service (PAS):

- [Managing Isolation Segments](#)
- [Routing for Isolation Segments](#)

## Managing Isolation Segments

This topic describes how operators can isolate deployment workloads into dedicated resource pools called isolation segments.

## Requirements

You must have the v.6.26.0  version or later of the [Cloud Foundry Command Line Interface \(cf CLI\)](#) installed to manage isolation segments.

Target the API endpoint of your deployment with `cf api` and log in with `cf login` before performing the procedures in this topic. For more information, see the [Identifying the API Endpoint for your PAS Instance](#) .

## Overview

To enable isolation segments, an operator must install the PCF Isolation Segment tile by performing the procedures in the [Installing PCF Isolation Segment](#) topic. Installing the tile creates a single isolation segment.

After an admin creates a new isolation segment, the admin can then create and manage relationships between the orgs and spaces of a Cloud Foundry deployment and the new isolation segment.

## Create an Isolation Segment

Before you create an isolation segment in PCF, you must install the PCF Isolation Segment tile by performing the procedures in the [Installing PCF Isolation Segment](#) topic.

To register an isolation segment with Cloud Controller, use the cf CLI.

 **Note:** The isolation segment name used in the cf CLI command must match the value specified in the **Segment Name** field of the PCF Isolation Segment tile. If the names do not match, PCF fails to place apps in the isolation segment when apps are started or restarted in the space assigned to the isolation segment.

The following command creates an isolation segment named `my_segment`:

```
$ cf create-isolation-segment my_segment
```

If successful, the command returns an `OK` message:

```
Creating isolation segment my_segment as admin...
OK
```

## Retrieve Isolation Segment Information

The `cf isolation-segments`, `cf org`, and `cf space` commands retrieve information about isolation segments. The isolation segments you can see depends on your role, as follows:

- **Admins** see all isolation segments in the system.
- **Other users** only see the isolation segments that their orgs are entitled to.

## List Isolation Segments

The following request returns a list of the isolation segments that are available to you:

```
$ cf isolation-segments
```

For example, the command returns results similar to the following:

```
Getting isolation segments as admin...
OK
```

```
name orgs
my_segment org1, org2
```

## Display Isolation Segments Enabled for an Org

An admin can entitle an org to multiple isolation segments.

Run `cf org ORG-NAME` command to view the isolation segments that are available to an org. Replace `ORG-NAME` with the name of your org.

For example:

```
$ cf org my-org
```

The command returns results similar to the following:

```
Getting info for org my-org as user@example.com...
```

```
name: my-org
domains: example.com, apps.example.com
quota: paid
spaces: development, production, sample-apps, staging
isolation segments: my_segment, my_other_segment
```

## Show the Isolation Segment Assigned to a Space

Only one isolation segment can be assigned to a space.

Run `cf space SPACE-NAME` to view the isolation segment assigned to a space. Replace `SPACE-NAME` with the name of the space.

For example:

```
$ cf space staging
```

The command returns results similar to the following:

```
name: staging
org: my-org
apps:
services:
isolation segment: my_segment
space quota:
security groups: dns, p-mysql, p.mysql, public_networks, rabbitmq, ssh-logging
```

## Delete an Isolation Segment

 **Note:** An isolation segment with deployed apps cannot be deleted.

Only admins can delete isolation segments.

Run `cf delete-isolation-segment SEGMENT-NAME` to delete an isolation segment. Replace `SEGMENT-NAME` with the name of the isolation segment. If successful, the command returns an `OK` message.

For example:

```
$ cf delete-isolation-segment my_segment
Deleting isolation segment my_segment as admin...
OK
```

## Manage Isolation Segment Relationships

The commands listed in the sections below manage the relationships between isolation segments, orgs, and spaces.

### Enable an Org to Use Isolation Segments

Only admins can enable orgs to use isolation segments. Run `cf enable-org-isolation ORG-NAME SEGMENT-NAME` to enable the use of an isolation segment. Replace `ORG-NAME` with the name of your org, and `SEGMENT-NAME` with the name of the isolation segment.

For example:

```
$ cf enable-org-isolation org2 my_segment
```

If an org is entitled to use only one isolation segment, that isolation segment does not automatically become the default isolation segment for the org. You must explicitly [set the default isolation segment](#) of an org.

### Disable an Org from Using Isolation Segments

 **Note:** You cannot disable an org from using an isolation segment if a space within that org is assigned to the isolation segment. Additionally, you cannot disable an org from using an isolation segment if the isolation segment is configured as the default for that org.

Run `cf disable-org-isolation ORG-NAME SEGMENT-NAME` to disable an org from using an isolation segment. Replace `ORG-NAME` with the name of your org, and `SEGMENT-NAME` with the name of the isolation segment.

For example:

```
$ cf disable-org-isolation org1 my_segment
```

If successful, the command returns an `OK` message:

```
Removing entitlement to isolation segment my_segment from org org1 as admin...
OK
```

### Set the Default Isolation Segment for an Org

This section requires [cf CLI v6.29.0](#) or later.

Only admins and org managers can set the default isolation segment for an org.

When an org has a default isolation segment, apps in its spaces belong to the default isolation segment unless you assign them to another isolation segment. You must restart running applications to move them into the default isolation segment.

Run `cf set-org-default-isolation-segment ORG-NAME SEGMENT-NAME` to set the default isolation segment for an org. Replace `ORG-NAME` with the name of your org, and `SEGMENT-NAME` with the name of the isolation segment.

For example:

```
$ cf set-org-default-isolation-segment org1 my_segment
Setting isolation segment my_segment to default on org org1 as admin...
OK
```

To display the default isolation segment for an org, use the `cf org` command.

## Assign an Isolation Segment to a Space

Admins and org managers can assign an isolation segment to a space. Apps in that space start in the specified isolation segment.

To assign an isolation segment to a space, you must first enable the space's org to use the isolation segment. See [Enable an Org to Use Isolation Segments](#).

Run `cf set-space-isolation-segment SPACE-NAME SEGMENT-NAME` to assign an isolation segment to a space. Replace `SPACE-NAME` with the name of the space, and `SEGMENT-NAME` with the name of the isolation segment.

For example:

```
$ cf set-space-isolation-segment space1 my_segment
```

## Reset the Isolation Segment Assignment for a Space

Admins can reset the isolation segment assigned to a space to use the org's default isolation segment.

Run `cf reset-space-isolation-segment SPACE-NAME` to the assign the default isolation segment for an org to a space. Replace `SPACE-NAME` with the name of the space.

For example:

```
$ cf reset-space-isolation-segment space1
```

## Routing for Isolation Segments

This topic describes how operators can configure and manage routing for isolation segments. Operators can deploy an additional set of routers for each isolation segment to handle requests for applications within the segment. This topic includes the following sections:

- [Overview](#)
- [Step 1: Create Networks](#)
- [Step 2: Configure Networks for Routers](#)
- [Step 3: Configure Additional Routers](#)
- [Step 4: Add Routers to Load Balancers](#)
- [Step 5: Configure DNS and Load Balancers](#)
- [Step 6: Configure Firewall Rules](#)
- [Additional GCP Information](#)
- [Sharding Routers for Isolation Segments](#)
- [Metrics for Routers Associated with Isolation](#)

For more information about how isolation segments work, see the [Isolation Segments](#) section of the *Cloud Foundry Security* topic. For more information about creating isolation segments, see the [Installing PCF Isolation Segment](#) topic.

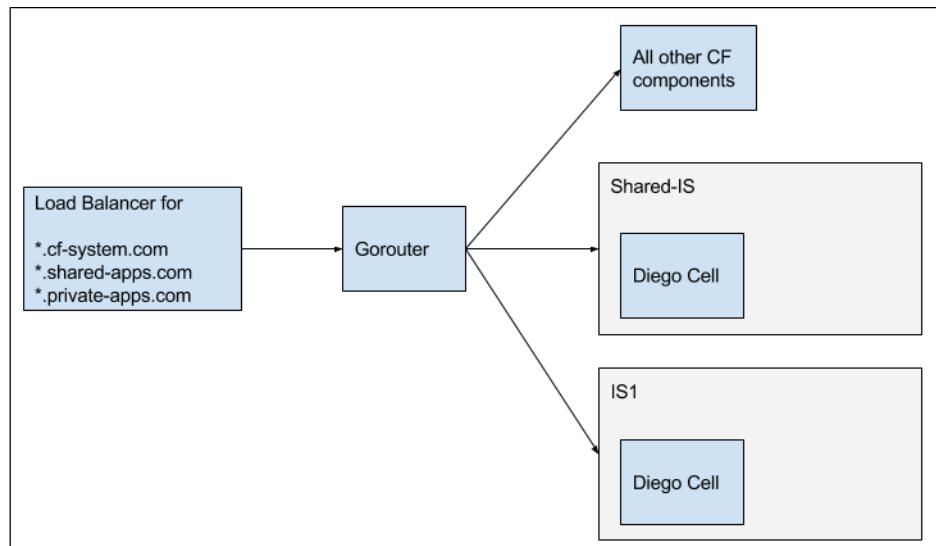
 **Note:** The instructions in this topic assume you are using Google Cloud Platform (GCP). The procedures may differ on other IaaSes, but the concepts should be transferable.

## Overview

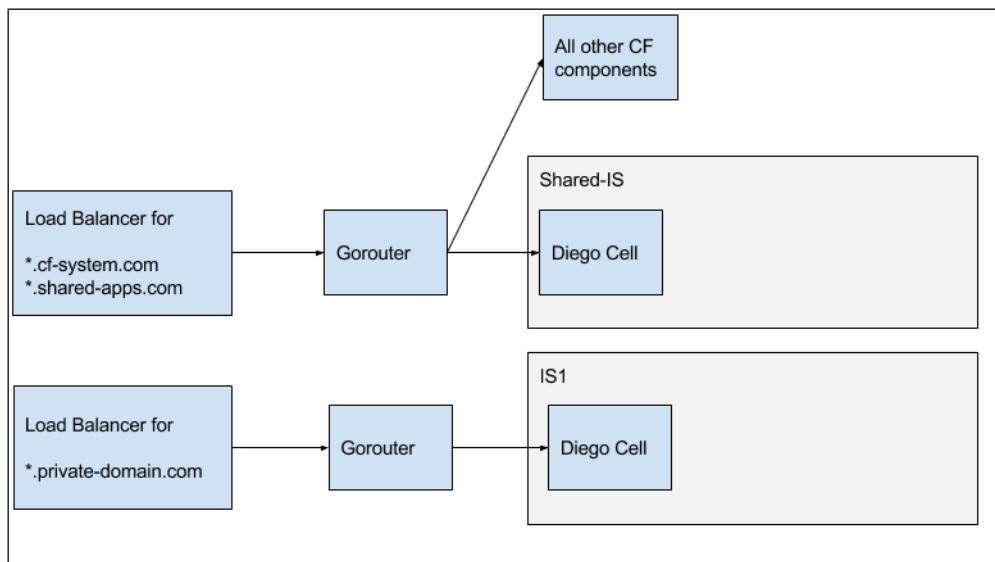
Isolation segments isolate the compute resources for one group of applications from another. However, these applications still share the same network resources. Requests for applications on all isolation segments, as well as for system components, transit the same load balancers and Cloud Foundry routers.

A shared Isolation Segment is the default isolation segment assigned to every org and space. This can be overwritten by assigning an explicit default for an organization. For more information about creating isolation segments, see the [Installing PCF Isolation Segment](#) topic.

The illustration below shows isolation segments sharing the same network resources.



Operators who want to prevent all isolation segments and system components from using the same network resources can deploy an additional set of routers for each isolation segment:



Use cases include:

- Requests for applications in an isolation segment must not share networking resources with requests for other applications.
- The Cloud Foundry management plane should only be accessible from a private network. As multiple IaaS load balancers cannot typically share the same pool of backends, such as Cloud Foundry routers, each load balancer requires an additional deployment of routers.

## Step 1: Create Networks

Create a network or subnet for each isolation segment on your infrastructure. For example, an operator who wants one shared isolation segment and one private segment could create one network named `sample-network` with two subnets named `sample-subnet-shared`, `sample-subnet-is1`.

The following diagram describes the network topology:

```

IaaS network: sample-network
 |
 |____ IaaS subnet: sample-subnet-shared
 |
 |____ IaaS subnet: sample-subnet-is1

```

Subnets do not generally span IaaS availability zones, so the same operator with two availability zones will need four subnets

```

IaaS network: sample-network
 |
 |____ IaaS subnet: sample-subnet-shared-az1
 |
 |____ IaaS subnet: sample-subnet-shared-az2
 |
 |____ IaaS subnet: sample-subnet-is1-az1
 |
 |____ IaaS subnet: sample-subnet-is1-az2

```

For more information about networks and subnets in GCP, see the [Using Networks and Firewalls](#) topic in the GCP documentation.

## Step 2: Configure Networks for Routers

To configure the subnets with Bosh, use [Bosh Cloud Config](#) subnets. Each subnet in the IaaS should correspond to a Bosh subnet that is labeled with the correct isolation segment.

Navigate to the **Assign AZs and Networks** section of the PCF Isolation Segment tile to assign your isolation segment to the network you created in Step 1. See the [Installing PCF Isolation Segment](#) topic for more information.

## Step 3: Configure Additional Routers

Navigate to the **Resource Config** section of the PCF Isolation Segment tile and use the dropdown menu to set your **Router** instances to a number greater than zero. See the [Installing PCF Isolation Segment](#) topic for more information.

## Step 4: Add Routers to Load Balancer

If your IaaS supports it, navigate to the **Resource Config** section of the PCF Isolation Segment tile and enter the name of your load balancer under **Load Balancers**. See the documentation specific to your IaaS in [Installing Pivotal Cloud Foundry](#) for more information. If your IaaS does not support this configuration, you must create static IP addresses and assign them to your load balancer out of band.

## Step 5: Configure DNS and Load Balancers

Create a separate domain name for each router instance group, and configure DNS to resolve these domain names to a load balancer that routes requests to the matching routers.

 **Note:** You must configure your load balancers to forward requests for a given domain to one router instance group only.

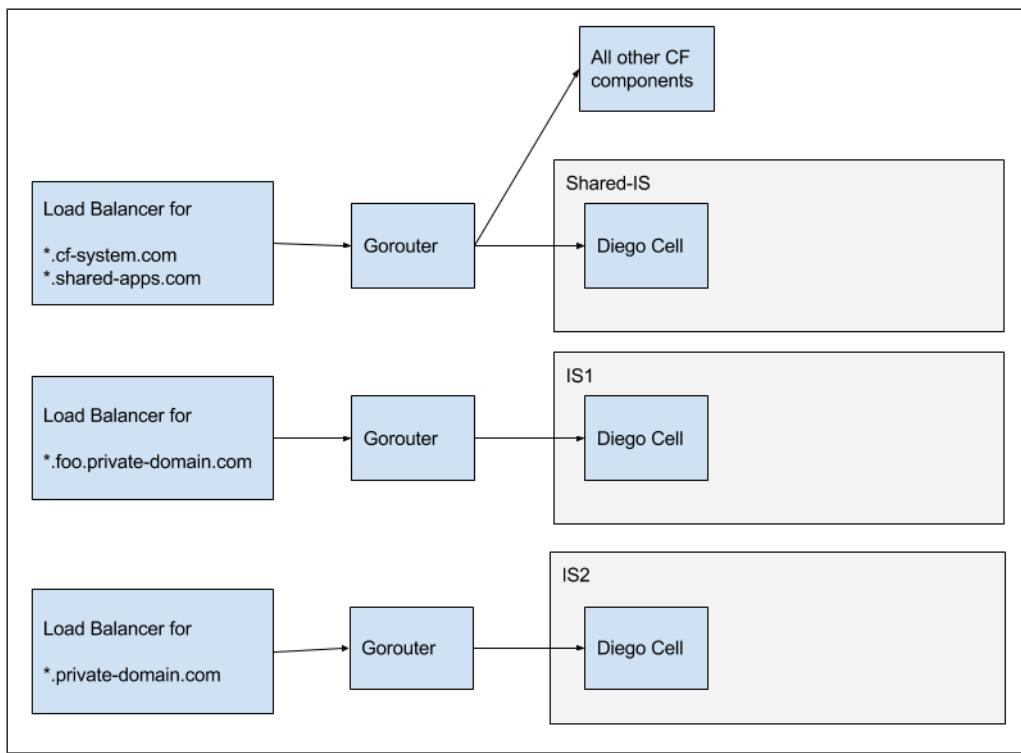
As router instance groups may be responsible for separate isolation segments, and an application may be deployed to only one isolation segment, requests should only reach a router that has access to the applications for that domain name. Load balancing requests for a domain across more than router instance group can result in request failures unless all the router instance groups have access to the isolation segments where applications for that domain are deployed.

### Shared Domain Name

It is a common requirement for applications on separate isolation segments to be accessible at domain names that share a domain, such as `private-domain.com`. To achieve this configuration while also obeying the guideline for forwarding requests for a domain to only one router instance group, create a new Cloud Foundry domain for a needed subdomain, such as `*.foo.private-domain.com`.

The diagrams illustrate a topology with separate load balancers, but you could also use one load balancer with multiple interfaces. In this configuration:

- Requests for system domain `*.cf-system.com` and the shared domain `*.shared-apps.com` are forwarded to the routers for the shared isolation segment.
- Requests for private domain `*.foo.private-domain.com` are forwarded to the routers for IS1. Requests for private domain `*.private-domain.com` are forwarded to the routers for IS2.



## Step 6: Configure Firewall Rules

Configure firewall rules to allow for necessary ingress and egress traffic for private and shared isolation segments. Assuming a default deny-all rule, properly configuring firewall rules prevents a request with a spoofed Host header from being forwarded by a router to an application in a different isolation segment.

To configure firewall rules for isolation segment traffic, do the following:

1. Configure the firewall rules in the table below:

**Note:** Firewall rules are specific to each IaaS, so the exact definition of **Source** and **Destination** depends on the IaaS. For example, on GCP, a **Source** is a subnet and a **Destination** is a tag. On AWS, both **Source** and **Destination** are security groups.

| Rule Name       | Source                    | Allowed Protocols/Ports | Destination               | Reason                                                                            |
|-----------------|---------------------------|-------------------------|---------------------------|-----------------------------------------------------------------------------------|
| shared-to-bosh  | Shared isolation segment  | tcp                     | BOSH Director             | BOSH Agent on VMs in the shared isolation segment to reach BOSH Director          |
| bosh-to-shared  | BOSH Director             | tcp                     | Shared isolation segment  | BOSH director to control VMs in the shared isolation segment                      |
| shared-internal | Shared isolation segment  | tcp                     | Shared isolation segment  | VMs within the shared isolation segment to reach one another                      |
| shared-to-is1   | Shared isolation segment  | tcp:1801,8853,9100      | Private isolation segment | Diego BBS in shared isolation segment to reach Cells in private isolation segment |
| is1-to-bosh     | Private isolation segment | tcp:4222,25250,25777    | BOSH Director             | BOSH agent on VMs in private isolation segment to reach BOSH Director             |
| is1-internal    | Private isolation segment | tcp                     | Private isolation segment | VMs within private isolation segment to reach one another                         |
|                 |                           |                         |                           |                                                                                   |

|                            |                           |                                                                                                                                                                                                                                                                                              |                          |                                                                                                                                                                                                        |
|----------------------------|---------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>isl-to-shared</code> | Private isolation segment | <p><code>tcp:3000,3457,4003<br/>4103,4222,4443,8080<br/>8082,8083,8443,8844<br/>8853,8889,8891<br/>9000,9022,9023,9090,9091</code></p> <p>See <a href="#">Port Reference Table</a> for information about the processes that use these ports and their corresponding manifest properties.</p> | Shared isolation segment | Diego Cells in private isolation segment to reach BBS, Auctioneer, and CredHub in shared isolation segment. Loggregator Agent to reach Traffic Controller. Routers to reach NATS, UAA, and Routing API |
|----------------------------|---------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

2. (Optional) Configure the firewall rules in the table below:

| Rule Name                      | Source                                     | Allowed Protocols/Ports | Destination               | Reason                                                                                                                                                           |
|--------------------------------|--------------------------------------------|-------------------------|---------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <code>jumpbox-to-isl</code>    | Jumpbox VM                                 | <code>tcp:22</code>     | Private isolation segment | Jumpbox VMs to reach private isolation segment through SSH or BOSH SSH.                                                                                          |
| <code>isl-to-jumpbox</code>    | Private isolation segment                  | <code>tcp</code>        | Jumpbox VM                | Private isolation segment to reach jumpbox VM. Opens reverse SSH or BOSH SSH tunnel from jumpbox VM to private isolation segment.                                |
| <code>diego-cell-egress</code> | Diego cell VM on private isolation segment | <code>tcp</code>        | Internet                  | If Diego Cells must download buildpacks to stage applications, allow egress traffic from all Diego Cell VMs on private isolation segments to reach the Internet. |

For more information about ports used by agents to communicate with BOSH, see [bosch-deployment](#) in GitHub.

For more information about networks and firewall rules for GCP, see [Using Subnetworks](#) in the GCP documentation.

## Port Reference Table

See the following table to understand which protocols and ports map to which processes and manifest properties for the `isl-to-shared` rule above.

| Protocol         | Port              | Process                                   | Manifest Property                                                           |
|------------------|-------------------|-------------------------------------------|-----------------------------------------------------------------------------|
| <code>tcp</code> | <code>3000</code> | Routing API                               | <code>routing_api.port</code>                                               |
| <code>tcp</code> | <code>3457</code> | Doppler                                   | <code>metron_endpoint.dropsonde_port</code>                                 |
| <code>tcp</code> | <code>4003</code> | VXLAN Policy Agent                        | <code>cf_networking.policy_server.internal_listen_port</code>               |
| <code>tcp</code> | <code>4103</code> | Silk Controller                           | <code>cf_networking.silk_controller.listen_port</code>                      |
| <code>tcp</code> | <code>4222</code> | NATS                                      | <code>router.nats.port</code>                                               |
| <code>tcp</code> | <code>8080</code> | Diego file server                         | <code>diego.file_server.listen_addr</code>                                  |
| <code>tcp</code> | <code>8082</code> | Doppler gRPC                              | <code>loggregator.doppler.grpc_port</code>                                  |
| <code>tcp</code> | <code>8083</code> | Reverse Log Proxy Gateway (cf auth proxy) | <code>loggregator.reverse_log_proxy_gateway_cf_auth_proxy.proxy_port</code> |
| <code>tcp</code> | <code>8443</code> | UAA                                       | <code>uaa.ssl.port</code>                                                   |
| <code>tcp</code> | <code>8844</code> | CredHub                                   | <code>credhub.port</code>                                                   |
| <code>tcp</code> | <code>8853</code> | BOSH DNS health                           | <code>health.server.port</code> from <code>bosh-dns-release</code>          |
| <code>tcp</code> | <code>8889</code> | Diego BBS                                 | <code>diego.rep.bbs.api_location</code>                                     |
| <code>tcp</code> | <code>8891</code> | Diego Database (Locket)                   | <code>diego.locket.listen_addr</code>                                       |
| <code>tcp</code> | <code>9000</code> | Loggregator Syslog Binding Cache          | <code>loggr-syslog-binding-cache.external_port</code>                       |
| <code>tcp</code> | <code>9022</code> | CC stager                                 | <code>capi.stager.cc.external_port</code>                                   |

|                  |                   |                       |                                                                  |
|------------------|-------------------|-----------------------|------------------------------------------------------------------|
| <code>tcp</code> | <code>9023</code> | CC TPS<br>CC uploader | <code>capi.tps.cc.external_port</code><br><code>http_port</code> |
| <code>tcp</code> | <code>9091</code> | CC uploader           | <code>https_port</code>                                          |

## Additional GCP Information

For more information, see the following:

- “Backend Services” in the GCP [documentation](#)
- BOSH Google Compute Engine CPI GitHub [repository](#)

## Sharding Routers for Isolation Segments

You can configure router sharding for isolation segments depending on your use case:

| Use Case                                       | Description                                                                                                                                                                                                                                                                            | How to Configure                                                                                                                                                                                                                                                                                                                             |
|------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Securing apps that run in an isolation segment | To provide security guarantees in addition to the firewall rules described above, you can configure sharding of the Gorouter’s routing table, resulting in a router dedicated for an isolation segment having knowledge only of routes for applications in the same isolation segment. | <ol style="list-style-type: none"> <li>In the <b>Networking</b> configuration pane of the Pivotal Application Service (PAS) tile, select the checkbox labeled <b>Routers reject requests for Isolation Segments</b>.</li> <li>Set the <b>Router Sharding Mode</b> in the isolation segment tile to <b>Isolation Segment Only</b>.</li> </ol> |
| Deploying additional routers for PAS           | The flexibility of the configuration also supports deployment of a router that excludes all isolation segments.                                                                                                                                                                        | <ol style="list-style-type: none"> <li>In the <b>Networking</b> configuration pane of the PAS tile, select the checkbox labeled <b>Routers reject requests for Isolation Segments</b>.</li> <li>Set the <b>Router Sharding Mode</b> in the isolation segment tile to <b>No isolation Segment</b>.</li> </ol>                                 |

**Note:** For compute isolation only, you can leave the **Routers reject requests for isolation segments** checkbox unselected in the PAS **Networking** pane. This is the default setting, which does not require any additional routers for the Isolation Segment tile.

## Metrics for Routers Associated with Isolation Segments

For metrics emitted by the Gorouter, metrics can be distinguished by the name of the job. For example, the following line is a metric emitted or `uptime`:

```
origin:"gorouter" eventType:ValueMetric timestamp:1491338040750977602 deployment:"superman.cf-app.com" job:"router_is1" index:"9a4b639c-8f0e-4b2b-b332-4161ee4646e6" ip:"10.0.16.23" valueMetric:<r
```

## Monitoring PAS

This guide describes how Pivotal Cloud Foundry (PCF) operators can monitor their Pivotal Application Service (PAS) deployments. For information about monitoring Pivotal Container Service (PKS) deployments, see [Logging and Monitoring PKS](#).

For more information about logging and metrics in PCF, see [Overview of Logging and Metrics](#).

## Overview

This guide includes the following topics:

- [Key Performance Indicators](#): A list of Key Performance Indicators (KPIs) that operators may want to monitor with their PAS deployment to help ensure it is in a good operational state.
- [Key Capacity Scaling Indicators](#): A list of capacity scaling indicators that operators may want to monitor to determine when they need to scale their PAS deployments.
- [Selecting and Configuring a Monitoring System](#): Guidance for setting up PAS with monitoring platforms to continuously monitor component metrics and trigger health alerts.

## KPI and Scaling Changes from PAS v2.5 to v2.6

This table highlights new and changed KPIs and scaling indicators in PAS v2.6.

| Status   | Metric                          | Description                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | Reference                                       |
|----------|---------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------|
| Modified | Doppler Egress Dropped Messages | <p>Within PAS, logs and metrics enter Loggregator for transport and then egress through Doppler. Doppler drops messages when consumers of the RLP, such as monitoring tool nozzles, ingest the exiting stream of logs and metrics too slowly.</p> <p> <b>Note:</b> The <code>doppler.dropped</code> metric includes both <code>ingress</code> and <code>egress</code> directions, but <code>egress</code> is a newly added distinction. To differentiate between <code>ingress</code> and <code>egress</code>, see the <code>direction</code> tag on the metric.</p> | <a href="#">Doppler Egress Dropped Messages</a> |

## Monitor PCF Services

For information about KPIs and metrics for PCF services, see the following topics:

- MySQL for PCF: [Monitoring MySQL for PCF](#).
- Pivotal Cloud Cache: See [Monitoring Pivotal Cloud Cache](#).
- Redis for PCF: [Monitoring Redis for PCF](#).
- RabbitMQ for PCF (pre-provisioned): See [Monitoring and KPIs for Pre-Provisioned RabbitMQ for PCF](#).
- RabbitMQ for PCF (on-demand): See [Monitoring and KPIs for On-Demand RabbitMQ for PCF](#).

## Key Performance Indicators

This topic describes Key Performance Indicators (KPIs) that operators may want to monitor with their Pivotal Application Service (PAS) deployment to help ensure it is in a good operational state.

The following PAS v2.6 KPIs are provided for operators to give general guidance on monitoring a PAS deployment using platform component and system (BOSH) metrics. Although many metrics are emitted from the platform, the following PAS v2.4 KPIs are high-signal-value metrics that can indicate emerging platform issues.

This alerting and response guidance has been shown to apply to most deployments. Pivotal recommends that operators continue to fine-tune the alert measures to their deployment by observing historical trends. Pivotal also recommends that operators expand beyond this guidance and create new, deployment-specific monitoring metrics, thresholds, and alerts based on learning from their deployments.

 **Note:** Thresholds noted as “dynamic” in the tables below indicate that while a metric is highly important to watch, the relative numbers to set threshold warnings are specific to a given PAS deployment and its use cases. These dynamic thresholds should be occasionally revisited because the PCF foundation and its usage continue to evolve. See [Determine Warning and Critical Thresholds](#) for more information.

## Diego Auctioneer Metrics

### Auctioneer App Instance (AI) Placement Failures

| auctioneer.AuctioneerLRPAuctionsFailed |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|----------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                            | <p>The number of Long Running Process (LRP) instances that the auctioneer failed to place on Diego cells. This metric is cumulative over the lifetime of the auctioneer job.</p> <p><b>Use:</b> This metric can indicate that PAS is out of container space or that there is a lack of resources within your environment. This indicator also increases when the LRP is requesting an isolation segment, volume drivers, or a stack that is unavailable, either not deployed or lacking sufficient resources to accept the work.</p> <p>This metric is emitted on event, and therefore gaps in receipt of this metric can be normal during periods of no app instances being scheduled.</p> <p>This error is most common due to capacity issues, for example, if cells do not have enough resources, or if cells are going back and forth between a healthy and unhealthy state.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Counter (Integer)<br/> <b>Frequency:</b> During each auction</p> |
| Recommended measurement                | Per minute delta averaged over a 5-minute window                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| Recommended alert thresholds           | <p><b>Yellow warning:</b> <math>\geq 0.5</math><br/> <b>Red critical:</b> <math>\geq 1</math></p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| Recommended response                   | <ol style="list-style-type: none"> <li>To best determine the root cause, examine the Auctioneer logs. Depending on the specific error and resource constraint, you may also find a failure reason in the Cloud Controller (CC) API.</li> <li>Investigate the health of your Diego cells to determine if they are the resource type causing the problem.</li> <li>Consider scaling additional cells using Ops Manager.</li> <li>If scaling cells does not solve the problem, pull Diego brain logs and BBS node logs and contact Pivotal Support telling them that LRP auctions are failing.</li> </ol>                                                                                                                                                                                                                                                                                                                                                                                             |

### Auctioneer Time to Fetch Cell State

auctioneer.AuctioneerFetchStateDuration

## auctioneer.AuctioneerFetchStateDuration

|                              | auctioneer.AuctioneerFetchStateDuration                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                  | <p>Time in ns that the auctioneer took to fetch state from all the Diego cells when running its auction.</p> <p><b>Use:</b> Indicates how the cells themselves are performing. Alerting on this metric helps alert that app staging requests to Diego may be failing.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge, integer in ns<br/> <b>Frequency:</b> During event, during each auction</p>                                                                                                                                     |
| Recommended measurement      | Maximum over the last 5 minutes divided by 1,000,000,000                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Recommended alert thresholds | <p><b>Yellow warning:</b> <math>\geq 2\text{ s}</math></p> <p><b>Red critical:</b> <math>\geq 5\text{ s}</math></p>                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Recommended response         | <ol style="list-style-type: none"> <li>1. Check the health of the cells by reviewing the logs and looking for errors.</li> <li>2. Review IaaS console metrics.</li> <li>3. Inspect the Auctioneer logs to determine if one or more cells is taking significantly longer to fetch state than other cells. Relevant log lines will have wording like `fetched cell state`.</li> <li>4. Pull Diego brain logs, cell logs, and auctioneer logs and contact Pivotal Support telling them that fetching cell states is taking too long.</li> </ol> |

## Auctioneer App Instance Starts

### auctioneer.AuctioneerLRPAuctionsStarted

|                              | auctioneer.AuctioneerLRPAuctionsStarted                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                  | <p>The number of LRP instances that the auctioneer successfully placed on Diego cells. This metric is cumulative over the lifetime of the auctioneer job.</p> <p><b>Use:</b> Provides a sense of running system activity levels in your environment. Can also give you a sense of how many app instances have been started over time. The recommended measurement, below, can help indicate a significant amount of container churn. However, for capacity planning purposes, it is more helpful to observe deltas over a long time window.</p> <p>This metric is emitted on event, and therefore gaps in receipt of this metric can be normal during periods of no app instances being scheduled.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Counter (Integer)<br/> <b>Frequency:</b> During event, during each auction</p> |
| Recommended measurement      | Per minute delta averaged over a 5-minute window                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| Recommended alert thresholds | <p><b>Yellow warning:</b> Dynamic</p> <p><b>Red critical:</b> Dynamic</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Recommended response         | <p>When observing a significant amount of container churn, do the following:</p> <ol style="list-style-type: none"> <li>1. Look to eliminate explainable causes of temporary churn, such as a deployment or increased developer activity.</li> <li>2. If container churn appears to continue over an extended period, pull logs from the Diego Brain and BBS node before contacting Pivotal support.</li> </ol> <p>When observing extended periods of high or low activity trends, scale up or down CF components as needed.</p>                                                                                                                                                                                                                                                                                                   |

## Auctioneer Task Placement Failures

### auctioneer.AuctioneerTaskAuctionsFailed

|  | auctioneer.AuctioneerTaskAuctionsFailed                                                                             |
|--|---------------------------------------------------------------------------------------------------------------------|
|  | The number of Tasks that the auctioneer failed to place on Diego cells. This metric is cumulative over the lifetime |

|                                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|-------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                     | <p>of the auctioneer job.</p> <p><b>Use:</b> Failing Task auctions indicate a lack of resources within your environment and that you likely need to scale. This indicator also increases when the Task is requesting an isolation segment, volume drivers, or a stack that is unavailable, either not deployed or lacking sufficient resources to accept the work.</p>                                                                                                                                                                                                                        |
| <b>Description</b>                  | <p>This metric is emitted on event, and therefore gaps in receipt of this metric can be normal during periods of no tasks being scheduled.</p> <p>This error is most common due to capacity issues, for example, if cells do not have enough resources, or if cells are going back and forth between a healthy and unhealthy state.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Counter (Float)</p>                                                                                                                                                                                      |
| <b>Recommended measurement</b>      | <p><b>Frequency:</b> During event, during each auction</p> <p>Per minute delta averaged over a 5-minute window</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| <b>Recommended alert thresholds</b> | <p><b>Yellow warning:</b> <math>\geq 0.5</math></p> <p><b>Red critical:</b> <math>\geq 1</math></p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| <b>Recommended response</b>         | <ol style="list-style-type: none"> <li>1. In order to best determine the root cause, examine the Auctioneer logs. Depending on the specific error or resource constraint, you may also find a failure reason in the CC API.</li> <li>2. Investigate the health of Diego cells.</li> <li>3. Consider scaling additional cells using Ops Manager.</li> <li>4. If scaling cells does not solve the problem, pull Diego brain logs and BBS logs for troubleshooting and contact Pivotal Support for additional troubleshooting. Inform Pivotal Support that Task auctions are failing.</li> </ol> |

## Diego BBS Metrics

### BBS Time to Run LRP Convergence

| bbs.ConvergenceLRPDuration          |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                  | <p>Time in ns that the BBS took to run its LRP convergence pass.</p> <p><b>Use:</b> If the convergence run begins taking too long, apps or Tasks may be crashing without restarting. This symptom can also indicate loss of connectivity to the BBS database.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (Integer in ns)<br/> <b>Frequency:</b> During event, every 30 seconds when LRP convergence runs, emission should be near-constant on a running deployment</p>                                                                                   |
| <b>Recommended measurement</b>      | Maximum over the last 15 minutes divided by 1,000,000,000                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| <b>Recommended alert thresholds</b> | <p><b>Yellow warning:</b> <math>\geq 10</math> s</p> <p><b>Red critical:</b> <math>\geq 20</math> s</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| <b>Recommended response</b>         | <ol style="list-style-type: none"> <li>1. Check BBS logs for errors.</li> <li>2. Try vertically scaling the BBS VM resources up. For example, add more CPUs or memory depending on its <code>system.cpu</code> / <code>system.memory</code> metrics.</li> <li>3. Consider vertically scaling the PAS backing database, if <code>system.cpu</code> and <code>system.memory</code> metrics for the database instances are high.</li> <li>4. If that does not solve the issue, pull the BBS logs and contact Pivotal Support for additional troubleshooting.</li> </ol> |

## BBS Time to Handle Requests

| bbs.RequestLatency           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                  | <p>The maximum observed latency time over the past 60 seconds that the BBS took to handle requests across all its API endpoints.</p> <p>Diego is now aggregating this metric to emit the max value observed over 60 seconds.</p> <p><b>Use:</b> If this metric rises, the PAS API is slowing. Response to certain cf CLI commands is slow if request latency is high.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (Integer in ns)<br/> <b>Frequency:</b> 60 s</p>                                                                                                                                                                                                                       |
| Recommended measurement      | Average over the last 15 minutes divided by 1,000,000,000                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Recommended alert thresholds | <p><b>Yellow warning:</b> <math>\geq 5</math> s<br/> <b>Red critical:</b> <math>\geq 10</math> s</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| Recommended response         | <ol style="list-style-type: none"> <li>1. Check CPU and memory statistics in Ops Manager.</li> <li>2. Check BBS logs for faults and errors that can indicate issues with BBS.</li> <li>3. Try scaling the BBS VM resources up. For example, add more CPUs/memory depending on its <code>system.cpu</code> / <code>system.memory</code> metrics.</li> <li>4. Consider vertically scaling the PAS backing database, if <code>system.cpu</code> and <code>system.memory</code> metrics for the database instances are high.</li> <li>5. If the above steps do not solve the issue, collect a sample of the cell logs from the BBS VMs and contact Pivotal Support to troubleshoot further.</li> </ol> |

## Cloud Controller and Diego in Sync

| bbs.Domain.cf-apps           |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                  | <p>Indicates if the <code>cf-apps</code> Domain is up-to-date, meaning that CF App requests from Cloud Controller are synchronized to <code>bbs.LRPsDesired</code> (Diego-desired ALs) for execution.</p> <ul style="list-style-type: none"> <li>• <code>1</code> means <code>cf-apps</code> Domain is up-to-date</li> <li>• No data received means <code>cf-apps</code> Domain is not up-to-date</li> </ul> <p><b>Use:</b> If the <code>cf-apps</code> Domain does not stay up-to-date, changes requested in the Cloud Controller are not guaranteed to propagate throughout the system. If the Cloud Controller and Diego are out of sync, then apps running could vary from those desired.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (Float)<br/> <b>Frequency:</b> 30 s</p> |
| Recommended measurement      | Value over the last 5 minutes                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| Recommended alert thresholds | <p><b>Yellow warning:</b> N/A<br/> <b>Red critical:</b> <math>\neq 1</math></p> <p>The recommended threshold value represents a state where an up-to-date metric <code>1</code> has not been received for the entire 5-minute window.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Recommended response         | <ol style="list-style-type: none"> <li>1. Check the BBS and Clock Global (Cloud Controller clock) logs.</li> <li>2. If the problem continues, pull the BBS logs and Clock Global (Cloud Controller clock) logs and contact Pivotal Support to say that the <code>cf-apps</code> domain is not being kept fresh.</li> </ol>                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |

## More App Instances Than Expected

| bbs.LRPsExtra                       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|-------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                  | <p>Total number of LRP instances that are no longer desired but still have a BBS record. When Diego wants to add more apps, the BBS sends a request to the auctioneer to spin up additional LRPs. LRPsExtra is the total number of LRP instances that are no longer desired but still have a BBS record.</p> <p><b>Use:</b> If Diego has more LRPs running than expected, there may be problems with the BBS.</p> <p>Deleting an app with many instances can temporarily spike this metric. However, a sustained spike in <code>bbs.LRPsExtra</code> is unusual and should be investigated.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (Float)<br/> <b>Frequency:</b> 30 s</p> |
| <b>Recommended measurement</b>      | Average over the last 5 minutes                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| <b>Recommended alert thresholds</b> | <p><b>Yellow warning:</b> <math>\geq 5</math><br/> <b>Red critical:</b> <math>\geq 10</math></p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| <b>Recommended response</b>         | <ol style="list-style-type: none"> <li>Review the BBS logs for proper operation or errors, looking for detailed error messages.</li> <li>If the condition persists, pull the BBS logs and contact Pivotal Support.</li> </ol>                                                                                                                                                                                                                                                                                                                                                                                                                                                              |

## Fewer App Instances Than Expected

| bbs.LRPsMissing                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|-------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                  | <p>Total number of LRP instances that are desired but have no record in the BBS. When Diego wants to add more apps, the BBS sends a request to the auctioneer to spin up additional LRPs. LRPsMissing is the total number of LRP instances that are desired but have no BBS record.</p> <p><b>Use:</b> If Diego has less LRP running than expected, there may be problems with the BBS.</p> <p>An app push with many instances can temporarily spike this metric. However, a sustained spike in <code>bbs.LRPsMissing</code> is unusual and should be investigated.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (Float)<br/> <b>Frequency:</b> 30 s</p> |
| <b>Recommended measurement</b>      | Average over the last 5 minutes                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| <b>Recommended alert thresholds</b> | <p><b>Yellow warning:</b> <math>\geq 5</math><br/> <b>Red critical:</b> <math>\geq 10</math></p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| <b>Recommended response</b>         | <ol style="list-style-type: none"> <li>Review the BBS logs for proper operation or errors, looking for detailed error messages.</li> <li>If the condition persists, pull the BBS logs and contact Pivotal Support.</li> </ol>                                                                                                                                                                                                                                                                                                                                                                                                                                      |

## Crashed App Instances

| bbs.CrashedActualLRPs |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|-----------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>    | <p>Total number of LRP instances that have crashed.</p> <p><b>Use:</b> Indicates how many instances in the deployment are in a crashed state. An increase in <code>bbs.CrashedActualLRPs</code> can indicate several problems, from a bad app with many instances associated, to a platform issue that is resulting in app crashes. Use this metric to help create a baseline for your deployment. After you have a baseline, you can create a deployment-specific alert to notify of a spike in crashes above the</p> |

|                                     |                                                                                                                                                                                                                                                                                                                                                                |
|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                     | trend line. Tune alert values to your deployment.<br><br><b>Origin:</b> Firehose<br><b>Type:</b> Gauge (Float)<br><b>Frequency:</b> 30 s                                                                                                                                                                                                                       |
| <b>Recommended measurement</b>      | Average over the last 5 minutes                                                                                                                                                                                                                                                                                                                                |
| <b>Recommended alert thresholds</b> | <b>Yellow warning:</b> Dynamic<br><b>Red critical:</b> Dynamic                                                                                                                                                                                                                                                                                                 |
| <b>Recommended response</b>         | <ol style="list-style-type: none"> <li>1. Look at the BBS logs for apps that are crashing and at the cell logs to see if the problem is with the apps themselves, rather than a platform issue.</li> <li>2. Before contacting Pivotal Support, pull the BBS logs and, if particular apps are the problem, pull the logs from their Diego cells too.</li> </ol> |

## Running App Instances, Rate of Change

| 1hr average of bbs.LRPsRunning – prior 1hr average of bbs.LRPsRunning |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|-----------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                                                    | Rate of change in app instances being started or stopped on the platform. It is derived from <code>bbs.LRPsRunning</code> and represents the total number of LRP instances that are running on Diego cells.<br><br><b>Use:</b> Delta reflects upward or downward trend for app instances started or stopped. Helps to provide a picture of the overall growth trend of the environment for capacity planning. You may want to alert on delta values outside of the expected range.<br><br><b>Origin:</b> Firehose<br><b>Type:</b> Gauge (Float)<br><b>Frequency:</b> During event, emission should be constant on a running deployment |
| <b>Recommended measurement</b>                                        | derived=(1-hour average of <code>bbs.LRPsRunning</code> – prior 1-hour average of <code>bbs.LRPsRunning</code> )                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| <b>Recommended alert thresholds</b>                                   | <b>Yellow warning:</b> Dynamic<br><b>Red critical:</b> Dynamic                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| <b>Recommended response</b>                                           | Scale components as necessary.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |

## Diego Cell Metrics

### Remaining Memory Available — Cell Memory Chunks Available

| rep.CapacityRemainingMemory |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|-----------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>          | Remaining amount of memory in MiB available for this Diego cell to allocate to containers.<br><br><b>Use:</b> Indicates the available cell memory. Insufficient cell memory can prevent pushing and scaling apps.<br><br>The strongest operational value of this metric is to understand a deployment's average app size and monitor/alert on ensuring that at least some cells have large enough capacity to accept standard app size pushes. For example, if pushing a 4 GB app, Diego would have trouble placing that app if there is no one cell with sufficient capacity of 4 GB or greater.<br><br>As an example, Pivotal Cloud Ops uses a standard of 4 GB, and computes and monitors for the number of cells with at least 4 GB free. When the number of cells with at least 4 GB falls below a defined threshold, this is a scaling indicator alert to increase capacity. This <i>free chunk</i> count threshold should be tuned to the deployment size and the standard size of apps being pushed to the deployment.<br><br><b>Origin:</b> Firehose |

|                                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|-------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                     | <p><b>Type:</b> Gauge (Integer in MiB)</p> <p><b>Frequency:</b> 60 s</p> <p><b>For alerting:</b></p> <ol style="list-style-type: none"> <li>1. Determine the size of a standard app in your deployment. This is the suggested value to calculate <i>free chunks</i> of Remaining Memory by.</li> <li>2. Create a script/tool that can iterate through each Diego Cell and do the following:             <ol style="list-style-type: none"> <li>a. Pull the <code>rep.CapacityRemainingMemory</code> metric for each cell.</li> <li>b. Divide the values received by 1000 to get the value in Gigabytes (if desired threshold is GB-based).</li> <li>c. Compare recorded values to your minimum capacity threshold, and count the number of cells that have equal or greater than the desired amount of <i>free chunk</i> space.</li> </ol> </li> <li>3. Determine a desired scaling threshold based on the minimum amount of <i>free chunks</i> that are acceptable in this deployment given historical trends.</li> <li>4. Set an alert to indicate the need to scale cell memory capacity when the value falls below the desired threshold number.</li> </ol> <p><b>For visualization purposes:</b><br/>Looking at this metric (<code>rep.CapacityRemainingMemory</code>) as a minimum value per cell has more informational value than alerting value. It can be an interesting heatmap visualization, showing average variance and density over time.</p> |
| <b>Recommended measurement</b>      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| <b>Recommended alert thresholds</b> | <p><b>Yellow warning:</b> Dynamic</p> <p><b>Red critical:</b> Dynamic</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| <b>Recommended response</b>         | <ol style="list-style-type: none"> <li>1. Assign more resources to the cells or assign more cells.</li> <li>2. Scale additional Diego cells using Ops Manager.</li> </ol>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| <b>Alternative Metric</b>           | If you are using <a href="#">PCF Healthwatch</a> , Pivotal recommends the metric <a href="#">healthwatch.Diego.AvailableFreeChunks</a> for this purpose.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |

## Remaining Memory Available — Overall Remaining Memory Available

| rep.CapacityRemainingMemory<br>(Alternative Use) |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|--------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                               | <p>Remaining amount of memory in MiB available for this Diego cell to allocate to containers.</p> <p><b>Use:</b> Can indicate low memory capacity overall in the platform. Low memory can prevent app scaling and new deployments. The overall sum of capacity can indicate that you need to scale the platform. Observing capacity consumption trends over time helps with capacity planning.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (Integer in MiB)<br/> <b>Frequency:</b> 60 s</p> |
| <b>Recommended measurement</b>                   | Minimum over the last 5 minutes divided by 1024 (across all instances)                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| <b>Recommended alert thresholds</b>              | <p><b>Yellow warning:</b> ≤ 64 GB</p> <p><b>Red critical:</b> ≤ 32 GB</p>                                                                                                                                                                                                                                                                                                                                                                                                                              |
| <b>Recommended response</b>                      | <ol style="list-style-type: none"> <li>1. Assign more resources to the cells or assign more cells.</li> <li>2. Scale additional Diego cells via Ops Manager.</li> </ol>                                                                                                                                                                                                                                                                                                                                |
| <b>Alternative Metric</b>                        | If you are using <a href="#">PCF Healthwatch</a> , Pivotal recommends the metric <a href="#">healthwatch.Diego.AvailableFreeChunks</a> for this purpose.                                                                                                                                                                                                                                                                                                                                               |

## Remaining Disk Available — Cell Disk Chunks Available

| rep.CapacityRemainingDisk    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                  | <p>Remaining amount of disk in MiB available for this Diego cell to allocate to containers.</p> <p><b>Use:</b> Indicates the available cell disk. Insufficient free disk on cell(s) will prevent the staging or starting of apps or tasks, resulting in error messages like <code>ERR Failed to stage application: insufficient resources</code>.</p> <p>Because Diego will fail to stage without at least 6 GB free, unreserved disk space on a given cell, the strongest operational value of this metric is to ensure that at least some cells have a large enough disk capacity to support the staging of apps and tasks.</p> <p>Pivotal recommends computing and monitoring for the number of cells with at least 6 GB Disk free. When the number of cells with at least 6 GB falls below a defined threshold, this is a scaling indicator alert to increase capacity. The alerting threshold value for the amount of <i>free chunks</i> of Disk should be tuned to the deployment size and the standard size of apps being pushed to the deployment. <b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (Integer in MiB)<br/> <b>Frequency:</b> 60 s</p>                                                                                                                                                                                                                  |
| Recommended measurement      | <p><b>For alerting:</b></p> <ol style="list-style-type: none"> <li>Because Diego will fail to stage without at least 6 GB free, this is the suggested minimum value to calculate <i>free chunks</i> of Remaining Disk by.</li> <li>Create a script/tool that can iterate through each Diego Cell and do the following:             <ol style="list-style-type: none"> <li>Pull the <code>rep.CapacityRemainingDisk</code> metric for each cell.</li> <li>Divide the values received by 1000 to get the value in Gigabytes (if desired threshold is GB-based).</li> <li>Compare recorded values to your minimum capacity threshold, and count the number of cells that have equal or greater than the desired amount of <i>free chunk</i> space.</li> </ol> </li> <li>Determine a desired scaling threshold based on the minimum amount of <i>free chunks</i> that are acceptable in this deployment given historical trends.</li> <li>Set an alert to indicate the need to scale cell disk capacity when the value falls below the desired threshold number.</li> </ol> <p><b>For visualization purposes:</b><br/>         Looking at this metric (<code>rep.CapacityRemainingDisk</code>) as a minimum value per cell has more informational value than alerting value. It can be an interesting heatmap visualization, showing average variance and density over time.</p> |
| Recommended alert thresholds | <b>Yellow warning:</b> Dynamic<br><b>Red critical:</b> Dynamic                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
| Recommended response         | <ol style="list-style-type: none"> <li>Assign more resources to the cells or assign more cells.</li> <li>Scale additional Diego cells using Ops Manager.</li> </ol>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Alternative Metric           | If you are using <a href="#">PCF Healthwatch</a> , Pivotal recommends the metric <a href="#">healthwatch.Diego.AvailableFreeChunksDisk</a> for this purpose.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |

## Remaining Disk Available - Overall Remaining Disk Available

| rep.CapacityRemainingDisk<br>(Alternative Use) |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                                    | <p>Remaining amount of disk in MiB available for this Diego cell to allocate to containers.</p> <p><b>Use:</b> Low disk capacity can prevent app scaling and new deployments. Because Diego staging Tasks can fail without at least 6 GB free, the recommended red threshold is based on the minimum disk capacity across the deployment falling below 6 GB in the previous 5 minutes.</p> <p>It can also be meaningful to assess how many chunks of free disk space are above a given threshold, similar to <code>rep.CapacityRemainingMemory</code>.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (Integer in MiB)</p> |

|                                     |                                                                                                                                                                     |
|-------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Recommended measurement</b>      | <b>Frequency:</b> 60 s<br>Minimum over the last 5 minutes divided by 1024 (across all instances)                                                                    |
| <b>Recommended alert thresholds</b> | <b>Yellow warning:</b> $\leq 12$ GB<br><b>Red critical:</b> $\leq 6$ GB                                                                                             |
| <b>Recommended response</b>         | <ol style="list-style-type: none"> <li>1. Assign more resources to the cells or assign more cells.</li> <li>2. Scale additional cells using Ops Manager.</li> </ol> |
| <b>Alternative Metric</b>           | If you are using <a href="#">PCF Healthwatch</a> , Pivotal recommends the metric <a href="#">healthwatch.Diego.AvailableFreeChunks</a> for this purpose.            |

## Cell Rep Time to Sync

| rep.RepBulkSyncDuration             |                                                                                                                                                                                                                                                                                                             |
|-------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                  | <p>Time in ns that the Diego Cell Rep took to sync the ActualLRPs that it claimed with its actual garden containers.</p> <p><b>Use:</b> Sync times that are too high can indicate issues with the BBS.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (Float in ns)<br/> <b>Frequency:</b> 30 s</p> |
| <b>Recommended measurement</b>      | Maximum over the last 15 minutes divided by 1,000,000,000                                                                                                                                                                                                                                                   |
| <b>Recommended alert thresholds</b> | <b>Yellow warning:</b> $\geq 5$ s<br><b>Red critical:</b> $\geq 10$ s                                                                                                                                                                                                                                       |
| <b>Recommended response</b>         | <ol style="list-style-type: none"> <li>1. Investigate BBS logs for faults and errors.</li> <li>2. If a particular cell or cells appear problematic, pull logs for the cells and the BBS logs before contacting Pivotal Support.</li> </ol>                                                                  |

## Unhealthy Cells

| rep.UnhealthyCell                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|-------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                  | <p>The Diego cell periodically checks its health against the garden backend. For Diego cells, 0 means healthy, and 1 means unhealthy.</p> <p><b>Use:</b> Set an alert for further investigation if multiple unhealthy Diego cells are detected in the given time window. If one cell is impacted, it does not participate in auctions, but end-user impact is usually low. If multiple cells are impacted, this can indicate a larger problem with Diego, and should be considered a more critical investigation need.</p> <p>Suggested alert threshold based on multiple unhealthy cells in the given time window.</p> <p>Although end-user impact is usually low if only one cell is impacted, this should still be investigated. Particularly in a lower capacity environment, this situation could result in negative end-user impact if left unresolved.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (Float, 0-1)<br/> <b>Frequency:</b> 30 s</p> |
| <b>Recommended measurement</b>      | Maximum over the last 5 minutes                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| <b>Recommended alert thresholds</b> | <b>Yellow warning:</b> = 1<br><b>Red critical:</b> > 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|                                     | <ol style="list-style-type: none"> <li>1. Investigate Diego cell servers for faults and errors.</li> <li>2. If a particular cell or cells appear problematic:</li> </ol>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |  |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| <p><b>Recommended response</b></p> <ol style="list-style-type: none"> <li>a. Determine a time interval during which the metrics from the cell changed from healthy to unhealthy.</li> <li>b. Pull the logs that the cell generated over that interval. The Cell ID is the same as the BOSH instance ID.</li> <li>c. Pull the BBS logs over that same time interval.</li> <li>d. Contact Pivotal Support.</li> </ol> <ol style="list-style-type: none"> <li>3. As a last resort, if you cannot wait for Pivotal Support, it sometimes helps to recreate the cell by running <code>bosh recreate</code>. See the <a href="#">BOSH documentation</a> for <code>bosh recreate</code> command syntax.</li> </ol> <div style="background-color: #fce4ec; border-radius: 5px; padding: 10px; margin-top: 10px;"> <p><b>⚠️ warning:</b> Recreating a cell destroys its logs. To enable a root cause analysis of the cell's problem, save out its logs before running <code>bosh recreate</code>.</p> </div> |  |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|

## Diego Locket Metrics

### Active Locks

| locket.ActiveLocks                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                  | Total count of how many locks the system components are holding.<br><br><b>Use:</b> If the ActiveLocks count is not equal to the expected value, there is likely a problem with Diego.<br><br><b>Origin:</b> Firehose<br><b>Type:</b> Gauge<br><b>Frequency:</b> 60 s                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| <b>Recommended measurement</b>      | Maximum over the last 5 minutes                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| <b>Recommended alert thresholds</b> | <b>Yellow warning:</b> N/A<br><b>Red critical:</b> ≠ 5                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |
| <b>Recommended response</b>         | <ol style="list-style-type: none"> <li>1. Run <code>monit status</code> to inspect for failing processes.</li> <li>2. If there are no failing processes, then review the logs for the components using the Locket service: BBS, Auctioneer, TPS Watcher, Routing API, and Clock Global (Cloud Controller clock). Look for indications that only one of each component is active at a time.</li> <li>3. Focus triage on the BBS first: <ul style="list-style-type: none"> <li>o A healthy BBS shows obvious activity around starting or claiming LRPCs.</li> <li>o An unhealthy BBS leads to the Auctioneer showing minimal or no activity. The BBS sends work to the Auctioneer.</li> <li>o Reference the BBS-level Locket metric <a href="#">Locks Held by BBS</a>. A value of 0 indicates Locket issues at the BBS level.</li> </ul> </li> <li>4. If the BBS appears healthy, then check the Auctioneer to ensure it is processing auction payloads. <ul style="list-style-type: none"> <li>o Recent logs for Auctioneer should show all but one of its instances are currently waiting on locks, and the active Auctioneer should show a record of when it last attempted to execute work. This attempt should correspond to app development activity, such as <code>cf push</code>.</li> <li>o Reference the Auctioneer-level Locket metric <a href="#">Locks Held by Auctioneer</a>. A value of 0 indicates Locket issues at the Auctioneer level.</li> </ul> </li> <li>5. The TPS Watcher is primarily active when app instances crash. Therefore, if the TPS Watcher is suspected, review the most recent logs.</li> <li>6. If you are unable to resolve on-going excessive active locks, pull logs from the Diego BBS and Auctioneer VMs, which includes the Locket service component logs, and contact Pivotal Support.</li> </ol> |

### Locks Held by BBS

## bbs.LockHeld

|                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                  | <p>Whether a BBS instance holds the expected BBS lock (in Locket). 1 means the active BBS server holds the lock, and 0 means the lock was lost.</p> <p><b>Use:</b> This metric is complimentary to <a href="#">Active Locks</a>, and it offers a BBS-level version of the Locket metrics. Although it is emitted per BBS instance, only 1 active lock is held by BBS. Therefore, the expected value is 1. The metric may occasionally be 0 when the BBS instances are performing a leader transition, but a prolonged value of 0 indicates an issue with BBS.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge<br/> <b>Frequency:</b> Periodically</p> |
| Recommended measurement      | Maximum over the last 5 minutes                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| Recommended alert thresholds | <p><b>Yellow warning:</b> N/A<br/> <b>Red critical:</b> ≠ 1</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| Recommended response         | <ol style="list-style-type: none"> <li>1. Run <code>monit status</code> on the Diego database VM to check for failing processes.</li> <li>2. If there are no failing processes, then review the logs for BBS. <ul style="list-style-type: none"> <li>◦ A healthy BBS shows obvious activity around starting or claiming LRPs.</li> <li>◦ An unhealthy BBS leads to the Auctioneer showing minimal or no activity. The BBS sends work to the Auctioneer.</li> </ul> </li> <li>3. If you are unable to resolve the issue, pull logs from the Diego BBS, which include the Locket service component logs, and contact Pivotal Support.</li> </ol>               |

## Locks Held by Auctioneer

|                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                              | <h3>auctioneer.LockHeld</h3>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Description                  | <p>Whether an Auctioneer instance holds the expected Auctioneer lock (in Locket). 1 means the active Auctioneer holds the lock, and 0 means the lock was lost.</p> <p><b>Use:</b> This metric is complimentary to <a href="#">Active Locks</a>, and it offers an Auctioneer-level version of the Locket metrics. Although it is emitted per Auctioneer instance, only 1 active lock is held by Auctioneer. Therefore, the expected value is 1. The metric may occasionally be 0 when the Auctioneer instances are performing a leader transition, but a prolonged value of 0 indicates an issue with Auctioneer.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge<br/> <b>Frequency:</b> Periodically</p>                                                  |
| Recommended measurement      | Maximum over the last 5 minutes                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| Recommended alert thresholds | <p><b>Yellow warning:</b> N/A<br/> <b>Red critical:</b> ≠ 1</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| Recommended response         | <ol style="list-style-type: none"> <li>1. Run <code>monit status</code> on the Diego Database VM to check for failing processes.</li> <li>2. If there are no failing processes, then review the logs for Auctioneer. <ul style="list-style-type: none"> <li>◦ Recent logs for Auctioneer should show all but one of its instances are currently waiting on locks, and the active Auctioneer should show a record of when it last attempted to execute work. This attempt should correspond to app development activity, such as <code>cf push</code>.</li> </ul> </li> <li>3. If you are unable to resolve the issue, pull logs from the Diego BBS and Auctioneer VMs, which includes the Locket service component logs, and contact Pivotal Support.</li> </ol> |

## Active Presences

| locket.ActivePresences       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                  | <p>Total count of active presences. Presences are defined as the registration records that the cells maintain to advertise themselves to the platform.</p> <p><b>Use:</b> If the Active Presences count is far from the expected, there might be a problem with Diego.</p> <p>The number of active presences varies according to the number of cells deployed. Therefore, during purposeful scale adjustments to PAS, this alerting threshold should be adjusted.</p> <p>Establish an initial threshold by observing the historical trends for the deployment over a brief period of time, Increase the threshold as more cells are deployed. During a rolling deploy, this metric shows variance during the BOSH lifecycle when cells are evacuated and restarted. Tolerable variance is within the bounds of the BOSH max inflight range for the instance group.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge<br/> <b>Frequency:</b> 60 s</p> |
| Recommended measurement      | Maximum over the last 15 minutes                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Recommended alert thresholds | <p><b>Yellow warning:</b> Dynamic<br/> <b>Red critical:</b> Dynamic</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| Recommended response         | <ol style="list-style-type: none"> <li>1. Ensure that the variance is not the result of an active rolling deploy. Also ensure that the alert threshold is appropriate to the number of cells in the current deployment.</li> <li>2. Run <code>monit status</code> to inspect for failing processes.</li> <li>3. If there are no failing processes, then review the logs for the components using the Locket service itself on Diego BBS instances.</li> <li>4. If you are unable to resolve the problem, pull the logs from the Diego BBS, which include the Locket service component logs, and contact Pivotal Support.</li> </ol>                                                                                                                                                                                                                                                                                                                       |

## Diego Route Emitter Metrics

### Route Emitter Time to Sync

| route_emitter.RouteEmitterSyncDuration |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|----------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                            | <p>Time in ns that the active Route Emitter took to perform its synchronization pass.</p> <p><b>Use:</b> Increases in this metric indicate that the Route Emitter may have trouble maintaining an accurate routing table to broadcast to the Gorouters. Tune alerting values to your deployment based on historical data and adjust based on observations over time. The suggested starting point is <math>\geq 5</math> for the yellow threshold and <math>\geq 10</math> for the critical threshold. Pivotal has observed on its Pivotal Web Services deployment that above 10 s, the BBS may be failing.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (Float in ns)<br/> <b>Frequency:</b> 60 s</p> |
| Recommended measurement                | Maximum, per job, over the last 15 minutes divided by 1,000,000,000                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
| Recommended alert thresholds           | <p><b>Yellow warning:</b> Dynamic<br/> <b>Red critical:</b> Dynamic</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Recommended response                   | <p>If all or many jobs showing as impacted, there is likely an issue with Diego.</p> <ol style="list-style-type: none"> <li>1. Investigate the Route Emitter and Diego BBS logs for errors.</li> <li>2. Verify that app routes are functional by making a request to an app, pushing an app and pinging it, or if applicable, checking that your smoke tests have passed.</li> </ol>                                                                                                                                                                                                                                                                                                                             |

If one or a few jobs showing as impacted, there is likely a connectivity issue and the impacted job should be investigated further.

## PAS MySQL KPIs

When PAS uses an internal MySQL database, as configured in the PAS tile **Settings** tab > **Databases** pane, the database cluster generates KPIs as described below.

**Note:** This section assumes you are using the **Internal Databases - MySQL - MariaDB Galera Cluster** option as your system database.

### MySQL Server Availability

|                              | /mysql/available                                                                                                                                                                                                                                                                                                                                                                                                 |
|------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                  | <p>The MySQL Server is currently responding to requests, which indicates that the server is running.</p> <p><b>Use:</b> This metric is especially useful in single-node mode, where cluster metrics are not relevant. If the server does not emit heartbeats, it is offline.</p> <p><b>Origin:</b> Firehose<br/> <b>Envelope Type:</b> Gauge<br/> <b>Unit:</b> boolean<br/> <b>Frequency:</b> 30 s (default)</p> |
| Recommended measurement      | Average over the last 5 minutes                                                                                                                                                                                                                                                                                                                                                                                  |
| Recommended alert thresholds | <p><b>Yellow warning:</b> N/A<br/> <b>Red critical:</b> &lt; 1</p>                                                                                                                                                                                                                                                                                                                                               |
| Recommended response         | Run <code>mysql-diag</code> and check the MySQL Server logs for errors.                                                                                                                                                                                                                                                                                                                                          |

### Galera Cluster Node Readiness

|                              | /mysql/galera/wsrep_ready                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                  | <p>Shows whether each cluster node can accept queries. Returns only 0 or 1. When this metric is 0, almost all queries to that node fail with the error:</p> <pre>ERROR 1047 (08501) Unknown Command</pre> <p><b>Use:</b> Discover when nodes of a cluster have been unable to communicate and, thus, unable to accept transactions.</p> <p><b>Origin:</b> Firehose<br/> <b>Envelope Type:</b> Gauge<br/> <b>Unit:</b> boolean<br/> <b>Frequency:</b> 30 s (default)</p> |
| Recommended measurement      | Average of values of each cluster node, over the last 5 minutes                                                                                                                                                                                                                                                                                                                                                                                                         |
| Recommended alert thresholds | <p><b>Yellow warning:</b> &lt; 1.0<br/> <b>Red critical:</b> 0 (cluster is down)</p>                                                                                                                                                                                                                                                                                                                                                                                    |
| Recommended response         | <ul style="list-style-type: none"> <li>- Run <code>mysql-diag</code> and check the MySQL Server logs for errors.</li> <li>- Make sure there has been no infrastructure event that affects intra-cluster communication.</li> <li>- Ensure that <code>wsrep_ready</code> has not been set to off by using the query:</li> </ul> <pre>SHOW STATUS LIKE 'wsrep_ready';</pre>                                                                                                |

### Galera Cluster Size

|  | /mysql/galera/wsrep_cluster_size |
|--|----------------------------------|
|  |                                  |

|                              |                                                                                                                                                                                                                                                                                                  |
|------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                              | <code>/mysql/galera/wsrep_cluster_size</code><br>The number of cluster nodes with which each node is communicating normally.                                                                                                                                                                     |
| Description                  | <p><b>Use:</b> When running in a multi-node configuration, this metric indicates if each member of the cluster is communicating normally with all other nodes.</p> <p><b>Origin:</b> Firehose<br/> <b>Envelope Type:</b> Gauge<br/> <b>Unit:</b> count<br/> <b>Frequency:</b> 30 s (default)</p> |
| Recommended measurement      | (Average of the values of each node / cluster size), over the last 5 minutes                                                                                                                                                                                                                     |
| Recommended alert thresholds | <b>Yellow warning:</b> < 3.0 (availability compromised)<br><b>Red critical:</b> < 1.0 (cluster unavailable)                                                                                                                                                                                      |
| Recommended response         | Run <code>mysql-diag</code> and check the MySQL Server logs for errors.                                                                                                                                                                                                                          |

## Galera Cluster Status

|                              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                              | <code>/mysql/galera/wsrep_cluster_status</code>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| Description                  | <p>Shows the primary status of the cluster component that the node is in.<br/>         Values are:<br/>         - Primary = 1<br/>         - Non-primary = 0<br/>         - Disconnected = -1<br/>         See: <a href="https://mariadb.com/kb/en/mariadb/galera-cluster-status-variables/">https://mariadb.com/kb/en/mariadb/galera-cluster-status-variables/</a></p> <p><b>Use:</b> Any value other than “Primary” indicates that the node is part of a nonoperational component. This occurs in cases of multiple membership changes that result in a loss of quorum.</p> <p><b>Origin:</b> Firehose<br/> <b>Envelope Type:</b> Gauge<br/> <b>Unit:</b> integer (see above)<br/> <b>Frequency:</b> 30 s (default)</p> |
| Recommended measurement      | Sum of each of the nodes, over the last 5 minutes                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Recommended alert thresholds | <b>Yellow warning:</b> < 3<br><b>Red critical:</b> < 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Recommended response         | <ul style="list-style-type: none"> <li>- Check node status to ensure that they are all in working order and able to receive write-sets.</li> <li>- Run <code>mysql-diag</code> and check the MySQL Server logs for errors.</li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

## Connections per Second

|                              |                                                                                                                                                                                                                                                                                                                                                                                                    |
|------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                              | <code>/mysql/net/connections</code>                                                                                                                                                                                                                                                                                                                                                                |
| Description                  | <p>Connections per second made to the server.</p> <p><b>Use:</b> If the number of connections drastically changes or if apps are unable to connect, there might be a network or app issue.</p> <p><b>Origin:</b> Firehose<br/> <b>Envelope Type:</b> Gauge<br/> <b>Unit:</b> count<br/> <b>Frequency:</b> 30 s (default)</p>                                                                       |
| Recommended measurement      | (Average of all nodes / max connections), over last 1 minute                                                                                                                                                                                                                                                                                                                                       |
| Recommended alert thresholds | <b>Yellow warning:</b> > 80%<br><b>Red critical:</b> > 90%                                                                                                                                                                                                                                                                                                                                         |
| Recommended response         | <ul style="list-style-type: none"> <li>- Run <code>mysql-diag</code> and check the MySQL Server logs for errors.</li> <li>- When approaching 100% of max connections, Apps may be experiencing times when they cannot connect to the database. The connections per second for the cluster vary based on application instances and app utilization. If this threshold is met or exceeded</li> </ul> |

for an extended period of time, monitor app usage to ensure everything is behaving as expected.

## CPU Utilization Percent

|                              | <a href="#">/mysql/performace/cpu_utilization_percent</a>                                                                                                                                                                                                                                                                                                          |
|------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                  | <p>CPU time being consumed by the MySQL service.</p> <p><b>Use:</b> A node that experiences context switching or high CPU usage will become unresponsive. This also affects the ability of the node to report metrics.</p> <p><b>Origin:</b> Doppler/Firehose<br/> <b>Envelope Type:</b> Gauge<br/> <b>Unit:</b> percent<br/> <b>Frequency:</b> 30 s (default)</p> |
| Recommended measurement      | Average over last 10 minutes                                                                                                                                                                                                                                                                                                                                       |
| Recommended alert thresholds | <b>Yellow warning:</b> > 80%<br><b>Red critical:</b> > 90%                                                                                                                                                                                                                                                                                                         |
| Recommended response         | <ul style="list-style-type: none"> <li>- Run <a href="#">mysql-diag</a> and check the MySQL Server logs for errors.</li> <li>- Determine what is using so much CPU. If it is from normal processes, update the service instance to use a plan with larger CPU capacity.</li> </ul>                                                                                 |

## Queries Delta

|                              | <a href="#">/mysql/performace/queries_delta</a>                                                                                                                                                                                                                                                                                                                |
|------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                  | <p>The number of statements executed by the server over the last 30 seconds.</p> <p><b>Use:</b> The server should always be processing some queries. If the server does not process any queries, the server is non-functional.</p> <p><b>Origin:</b> Doppler/Firehose<br/> <b>Envelope Type:</b> Gauge<br/> <b>Unit:</b> count<br/> <b>Frequency:</b> 30 s</p> |
| Recommended measurement      | Average over last 2 minutes                                                                                                                                                                                                                                                                                                                                    |
| Recommended alert thresholds | <b>Red critical:</b> 0                                                                                                                                                                                                                                                                                                                                         |
| Recommended response         | <ul style="list-style-type: none"> <li>- Run <a href="#">mysql-diag</a> and check the MySQL Server logs for errors.</li> <li>- Investigate the MySQL server logs, such as the audit log, to understand why query rate changed and determine appropriate action.</li> </ul>                                                                                     |

## Gorouter Metrics

### Router File Descriptors

| <a href="#">gorouter.file_descriptors</a> |                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|-------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                           | <p>The number of file descriptors currently used by the Gorouter job.</p> <p><b>Use:</b> Indicates an impending issue with the Gorouter. Without proper mitigation, it is possible for an unresponsive app to eventually exhaust available Gorouter file descriptors and cause route starvation for other apps running on PAS. Under heavy load, this unmitigated situation can also result in the Gorouter losing its connection to NATS and all routes being pruned.</p> |

|                                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|-------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                     | <p>While a drop in <code>gorouter.total_routes</code> or an increase in <code>gorouter.ms_since_last_registry_update</code> helps to surface that the issue may already be occurring, alerting on <code>gorouter.file_descriptors</code> indicates that such an issue is impending.</p> <p>The Gorouter limits the number of file descriptors to 100,000 per job. Once the limit is met, the Gorouter is unable to establish any new connections.</p> <p>To reduce the risk of DDoS attacks, Pivotal recommends doing one or both of the following:</p> <ul style="list-style-type: none"> <li>Within PAS, set <b>Max Connections Per Backend</b> to define how many requests can be routed to any particular app instance. This prevents a single app from using all Gorouter connections. The value specified should be determined by the operator based on the use cases for that foundation. For example, Pivotal sets the number of connections to 500 for Pivotal Web Services.</li> <li>Add rate limiting at the load balancer level.</li> </ul> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge<br/> <b>Frequency:</b> 5 s</p> |
| <b>Recommended measurement</b>      | Maximum, per Gorouter job, over the last 5 minutes                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| <b>Recommended alert thresholds</b> | <p><b>Yellow warning:</b> 50,000 per job<br/> <b>Red critical:</b> 60,000 per job</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| <b>Recommended response</b>         | <ol style="list-style-type: none"> <li>Identify which app(s) are requesting excessive connections and resolve the impacting issues with these apps.</li> <li>If the above recommended mitigation steps have not already been taken, do so.</li> <li>Consider adding more Gorouter VM resources to increase the number of available file descriptors.</li> </ol>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |

## Router Exhausted Connections

| <code>gorouter.backend_exhausted_conns</code> |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|-----------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                            | <p>The lifetime number of requests that have been rejected by the Gorouter VM due to the <code>Max Connections Per Backend</code> limit being reached across all tried backends. The limit controls the number of concurrent TCP connections to any particular app instance and is configured within PAS.</p> <p><b>Use:</b> Indicates that PAS is mitigating risk to other applications by self-protecting the platform against one or more unresponsive applications. Increases in this metric indicate the need to investigate and resolve issues with potentially unresponsive applications. A rapid rate of change upward is concerning and should be assessed further.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Counter (Integer)<br/> <b>Frequency:</b> 5 s</p>                                                                                                                                                         |
| <b>Recommended measurement</b>                | Maximum delta per minute, per Gorouter job, over a 5-minute window                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| <b>Recommended alert thresholds</b>           | <p><b>Yellow warning:</b> Dynamic<br/> <b>Red critical:</b> Dynamic</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| <b>Recommended response</b>                   | <ol style="list-style-type: none"> <li>If <code>gorouter.backend_exhausted_conns</code> spikes, first look to the Router Throughput metric <code>gorouter.total_requests</code> to determine if this measure is high or low in relation to normal bounds for this deployment.</li> <li>If Router Throughput appears within normal bounds, it is likely that <code>gorouter.backend_exhausted_conns</code> is spiking due to an unresponsive application, possibly due to application code issues or underlying application dependency issues. To help determine the problematic application, look in access logs for repeated calls to one application. Then proceed to troubleshoot this application accordingly.</li> <li>If Router Throughput also shows unusual spikes, the cause of the increase in <code>gorouter.backend_exhausted_conns</code> spikes is likely external to the platform. Unusual increases in load</li> </ol> |

may be due to expected business events driving additional traffic to applications. Unexpected increases in load may indicate a DDoS attack risk.

## Router Throughput

| gorouter.total_requests      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                  | <p>The lifetime number of requests completed by the Gorouter VM, emitted per Gorouter instance</p> <p><b>Use:</b> The aggregation of these values across all Gorouters provide insight into the overall traffic flow of a deployment. Unusually high spikes, if not known to be associated with an expected increase in demand, could indicate a DDoS risk. For performance and capacity management, consider this metric a measure of router throughput per job, converting it to requests-per-second, by looking at the delta value of <code>gorouter.total_requests</code> and deriving back to 1s, or <code>gorouter.total_requests.delta)/5</code>, per Gorouter instance. This helps you see trends in the throughput rate that indicate a need to scale the Gorouter instances. Use the trends you observe to tune the threshold alerts for this metric.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Counter (Integer)<br/> <b>Frequency:</b> 5 s</p> |
| Recommended measurement      | Average over the last 5 minutes of the derived per second calculation                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| Recommended alert thresholds | <p><b>Yellow warning:</b> Dynamic<br/> <b>Red critical:</b> Dynamic</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| Recommended response         | <p>For optimizing the Gorouter, consider the requests-per-second derived metric in the context of router latency and Gorouter VM CPU utilization. From performance and load testing of the Gorouter, Pivotal has observed that at approximately 2500 requests per second, latency can begin to increase.</p> <p>To increase throughput and maintain low latency, scale the Gorouters either horizontally or vertically and watch that the <code>system.cpu.user</code> metric for the Gorouter stays in the suggested range of 60-70% CPU Utilization.</p>                                                                                                                                                                                                                                                                                                                                                                                                        |

## Router Handling Latency

| gorouter.latency             |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                  | <p>The time in milliseconds that the Gorouter takes to handle requests to backend endpoints, which include both applications routable platform system APIs like Cloud Controller and UAA. This is the average round trip response time, which includes router handling.</p> <p><b>Use:</b> Indicates how Gorouter jobs in PAS are impacting overall responsiveness. Latencies above 100 ms can indicate problems with the network, misbehaving backends, or the need to scale the Gorouter itself due to ongoing traffic congestion. An alert value on this metric should be tuned to the specifics of the deployment and its underlying network considerations; a suggested starting point is 100 ms.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (Float in ms)<br/> <b>Frequency:</b> Emitted per Gorouter request, emission should be constant on a running deployment</p> |
| Recommended measurement      | Average over the last 30 minutes                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Recommended alert thresholds | <p><b>Yellow warning:</b> Dynamic<br/> <b>Red critical:</b> Dynamic</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| Recommended response         | <p>Extended periods of high latency can point to several factors. The Gorouter latency measure includes network and backend latency impacts as well.</p> <ol style="list-style-type: none"> <li>First inspect logs for network issues and indications of misbehaving backends.</li> <li>If it appears that the Gorouter needs to scale due to ongoing traffic congestion, do not scale on the latency metric alone. You should also look at the CPU utilization of the Gorouter VMs and keep it within a maximum 60-70% range.</li> </ol>                                                                                                                                                                                                                                                                                                                                                |

3. Resolve high utilization by scaling the Gorouter.

## Time Since Last Route Register Received

| gorouter.ms_since_last_registry_update |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|----------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                     | Time in milliseconds since the last route register was received, emitted per Gorouter instance<br><br>Use: Indicates if routes are not being registered to apps correctly.<br><br>Origin: Firehose<br>Type: Gauge (Float in ms)<br>Frequency: 30 s                                                                                                                                                                                                                                        |
| <b>Recommended measurement</b>         | Maximum over the last 5 minutes                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| <b>Recommended alert thresholds</b>    | <b>Yellow warning:</b> N/A<br><b>Red critical:</b> > 30,000<br>This threshold is suitable for normal platform usage. It alerts if it has been at least 30 seconds since the Gorouter last received a message from an app.                                                                                                                                                                                                                                                                 |
| <b>Recommended response</b>            | <ol style="list-style-type: none"> <li>1. Search the Gorouter and Route Emitter logs for connection issues to NATS.</li> <li>2. Check the BOSH logs to see if the NATS, Gorouter, or Route Emitter VMs are failing.</li> <li>3. Look more broadly at the health of all VMs, particularly Diego-related VMs.</li> <li>4. If problems persist, pull the Gorouter and Route Emitter logs and contact Pivotal Support to say there are consistently long delays in route registry.</li> </ol> |

## Router Error: 502 Bad Gateway

| gorouter.bad_gateways               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|-------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                  | The lifetime number of bad gateways, or 502 responses, from the Gorouter itself, emitted per Gorouter instance. The Gorouter emits a 502 bad gateway error when it has a route in the routing table and, in attempting to make a connection to the backend, finds that the backend does not exist.<br><br>Use: Indicates that route tables might be stale. Stale routing tables suggest an issue in the route register management plane, which indicates that something has likely changed with the locations of the containers. Always investigate unexpected increases in this metric.<br><br>Origin: Firehose<br>Type: Count (Integer, Lifetime)<br>Frequency: 5 s |
| <b>Recommended measurement</b>      | Maximum delta per minute over a 5-minute window                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| <b>Recommended alert thresholds</b> | <b>Yellow warning:</b> Dynamic<br><b>Red critical:</b> Dynamic                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| <b>Recommended response</b>         | <ol style="list-style-type: none"> <li>1. Check the Gorouter and Route Emitter logs to see if they are experiencing issues when connecting to NATS.</li> <li>2. Check the BOSH logs to see if the NATS, Gorouter, or Route Emitter VMs are failing.</li> <li>3. Look broadly at the health of all VMs, particularly Diego-related VMs.</li> <li>4. If problems persist, pull Gorouter and Route Emitter logs and contact Pivotal Support to say there has been an unusual increase in Gorouter bad gateway responses.</li> </ol>                                                                                                                                      |

## Router Error: Server Error

| gorouter.responses.5xx              |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|-------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                  | <p>The lifetime number of requests completed by the Gorouter VM for HTTP status family 5xx, server errors, emitted per Gorouter instance.</p> <p><b>Use:</b> A repeatedly crashing app is often the cause of a big increase in 5xx responses. However, response issues from apps can also cause an increase in 5xx responses. Always investigate an unexpected increase in this metric.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Counter (Integer)<br/> <b>Frequency:</b> 5 s</p> |
| <b>Recommended measurement</b>      | Maximum delta per minute over a 5-minute window                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| <b>Recommended alert thresholds</b> | <p><b>Yellow warning:</b> Dynamic<br/> <b>Red critical:</b> Dynamic</p>                                                                                                                                                                                                                                                                                                                                                                                                                   |
| <b>Recommended response</b>         | <ol style="list-style-type: none"> <li>Look for out-of-memory errors and other app-level errors.</li> <li>As a temporary measure, ensure that the troublesome app is scaled to more than one instance.</li> </ol>                                                                                                                                                                                                                                                                         |

## Number of Gorouter Routes Registered

| gorouter.total_routes               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|-------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                  | <p>The current total number of routes registered with the Gorouter, emitted per Gorouter instance</p> <p><b>Use:</b> The aggregation of these values across all Gorouters indicates uptake and gives a picture of the overall growth of the environment for capacity planning.</p> <p>Pivotal also recommends alerting on this metric if the number of routes falls outside of the normal range for your deployment. Dramatic decreases in this metric volume may indicate a problem with the route registration process, such as an app outage, or that something in the route register management plane has failed.</p> <p>If visualizing these metrics on a dashboard, <code>gorouter.total_routes</code> can be helpful for visualizing dramatic drops. However, for alerting purposes, the <code>gorouter.ms_since_last_registry_update</code> metric is more valuable for quicker identification of Gorouter issues. Alerting thresholds for <code>gorouter.total_routes</code> should focus on dramatic increases or decreases out of expected range.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (Float)<br/> <b>Frequency:</b> 30 s</p> |
| <b>Recommended measurement</b>      | 5-minute average of the per second delta                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| <b>Recommended alert thresholds</b> | <p><b>Yellow warning:</b> Dynamic<br/> <b>Red critical:</b> Dynamic</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| <b>Recommended response</b>         | <ol style="list-style-type: none"> <li>For capacity needs, scale up or down the Gorouter VMs as necessary.</li> <li>For significant drops in current total routes, see the <code>gorouter.ms_since_last_registry_update</code> metric value for additional context.</li> <li>Check the Gorouter and Route Emitter logs to see if they are experiencing issues when connecting to NATS.</li> <li>Check the BOSH logs to see if the NATS, Gorouter, or Route Emitter VMs are failing.</li> <li>Look broadly at the health of all VMs, particularly Diego-related VMs.</li> <li>If problems persist, pull the Gorouter and Route Emitter logs and contact Pivotal Support.</li> </ol>                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |

## Number of Route Registration Messages Sent and Received

| <b>gorouter.registry_message.route-emitter<br/>route_emitter.HTTPRouteNATSMessagesEmitted</b> |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|-----------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                                                                               | <p>This KPI is based on the following metrics:</p> <ul style="list-style-type: none"> <li>• <code>route_emitter.HTTPRouteNATSMessagesEmitted</code> reports the lifetime number of route registration messages sent by the Route Emitter component. The metric is emitted for each Route Emitter.</li> <li>• <code>gorouter.registry_message.route-emitter</code> reports the lifetime number of route registration messages received by the Gorouter. The metric is emitted for each Gorouter instance.</li> </ul> <p>Dynamic configuration that enables the Gorouter to route HTTP requests to apps is published by the Route Emitter component colocated on each Diego cell to the NATS clustered message bus. All router instances subscribed to this message bus receive the same configuration. (Router instances within an isolation segment receive configuration only for cells in the same isolation segment.)</p> <p>As Gorouters prune app instances from the route when a TTL expires, each Route Emitter periodically publishes the routing configuration for the app instances on the same cell.</p> <p>Therefore, the aggregate number of route registration messages published by all the Route Emitters should be equal to the number of messages received by each Gorouter instance.</p> <p><b>Use:</b> A difference in the rate of change of these metrics is an indication of an issue in the control plane responsible for updating the routers with changes to the routing table.</p> <p>Pivotal recommends alerting when the number of messages received per second for a given router instance falls below the sum of messages emitted per second across all Route Emitters.</p> <p>If visualizing these metrics on a dashboard, look for increases in the difference between the rate of messages received and sent. If the number of messages received by a Gorouter instance drops below the sum of messages sent by the Route Emitters, this is an indication of a problem in the control plane.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Counter<br/> <b>Frequency:</b> With each event</p> |
| <b>Recommended measurement</b>                                                                | Difference of 5-minute average of the per second deltas for <code>gorouter.registry_message.route-emitter</code> and sum of <code>route_emitter.HTTPRouteNATSMessagesEmitted</code> for all Route Emitters                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
| <b>Recommended alert thresholds</b>                                                           | <b>Yellow warning:</b> Dynamic<br><b>Red critical:</b> Dynamic                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| <b>Recommended response</b>                                                                   | <ol style="list-style-type: none"> <li>1. Check the Gorouter and Route Emitter logs to see if they are experiencing issues when connecting to NATS.</li> <li>2. Check the BOSH logs to see if the NATS, Gorouter, or Route Emitter VMs are failing.</li> <li>3. Look broadly at the health of all VMs, particularly Diego-related VMs.</li> <li>4. If problems persist, pull the Gorouter and Route Emitter logs and contact Pivotal Support.</li> </ol>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |

## UAA Metrics

### UAA Throughput

| <b>uaa.requests.global.completed.count</b> |                                                                                                                                                                                                                                                       |
|--------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                                            | <p>The lifetime number of requests completed by the UAA VM, emitted per UAA instance. This number includes health checks.</p> <p><b>Use:</b> For capacity planning purposes, the aggregation of these values across all UAA instances can provide</p> |

|                                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|-------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                  | <p>insight into the overall load that UAA is processing. It is recommended to alert on unexpected spikes per UAA instance. Unusually high spikes, if they are not associated with an expected increase in demand, could indicate a DDoS risk and should be investigated.</p> <p>For performance and capacity management, look at the UAA Throughput metric as either a requests-completed-per-second or requests-completed-per-minute rate to determine the throughput per UAA instance. This helps you see trends in the throughput rate that may indicate a need to scale UAA instances. Use the trends you observe to tune the threshold alerts for this metric.</p> <p>From performance and load testing of UAA, Pivotal has observed that while UAA endpoints can have different throughput behavior, once throughput reaches its peak value per VM, it stays constant and latency increases.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (Integer), emitted value increments over the lifetime of the VM like a counter<br/> <b>Frequency:</b> 5 s</p> |
| <b>Recommended measurement</b>      | Average over the last 5 minutes of the derived requests-per-second or requests-per-minute rate, per instance                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| <b>Recommended alert thresholds</b> | <p><b>Yellow warning:</b> Dynamic<br/> <b>Red critical:</b> Dynamic</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| <b>Recommended response</b>         | <p>For optimizing UAA, consider this metric in the context of <a href="#">UAA Request Latency</a> and <a href="#">UAA VM CPU Utilization</a>. To increase throughput and maintain low latency, scale the UAA VMs horizontally by editing the number of your UAA VM instances in the <b>Resource Config</b> pane of the PAS tile and ensure that the <code>system.cpu.user</code> metric for UAA is not sustained in the suggested range of 80-90% maximum CPU utilization.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |

## UAA Request Latency

| gorouter.latency.uaa                |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|-------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                  | <p>Time in milliseconds that UAA took to process a request that the Gorouter sent to UAA endpoints.</p> <p><b>Use:</b> Indicates how responsive UAA has been to requests sent from the Gorouter. Some operations may take longer to process, such as creating bulk users and groups. It is important to correlate latency observed with the endpoint and evaluate this data in the context of overall historical latency from that endpoint. Unusual spikes in latency could indicate the need to scale UAA VMs.</p> <p>This metric is emitted only for the routers serving the UAA system component and is not emitted per isolation segment even if you are using isolated routers.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (Float in ms)<br/> <b>Frequency:</b> Emitted per Gorouter request to UAA</p>                                                                                                                                     |
| <b>Recommended measurement</b>      | Maximum, per job, over the last 5 minutes                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| <b>Recommended alert thresholds</b> | <p><b>Yellow warning:</b> Dynamic<br/> <b>Red critical:</b> Dynamic</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
| <b>Recommended response</b>         | <p>Latency depends on the endpoint and operation being used. It is important to correlate the latency with the endpoint and evaluate this data in the context of the historical latency from that endpoint.</p> <ol style="list-style-type: none"> <li>Inspect which endpoints requests are hitting. Use historical data to determine if the latency is unusual for that endpoint. A list of UAA endpoints is available in the <a href="#">UAA API documentation</a>.</li> <li>If it appears that UAA needs to be scaled due to ongoing traffic congestion, do not scale based on the latency metric alone. You should also ensure that the <code>system.cpu.user</code> metric for UAA stays in the suggested range of 80-90% maximum CPU utilization.</li> <li>Resolve high utilization by scaling UAA VMs horizontally. To scale UAA, navigate to the <b>Resource Config</b> pane of the PAS tile and edit the number of your UAA VM instances.</li> </ol> |

## UAA Requests In Flight

| uaa.server.inflight.count    |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                  | <p>The number of requests UAA is currently processing (in-flight requests), emitted per UAA instance.</p> <p><b>Use:</b> Indicates how many concurrent requests are currently in flight for the UAA instance. Unusually high spikes, if they are not associated with an expected increase in demand, could indicate a DDoS risk.</p> <p>From performance and load testing of the UAA component, Pivotal has observed that the number of concurrent requests impacts throughput and latency. The UAA Requests In Flight metric helps you see trends in the request rate that may indicate the need to scale UAA instances. Use the trends you observe to tune the threshold alerts for this metric.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (Integer)<br/> <b>Frequency:</b> 5 s</p> |
| Recommended measurement      | Maximum, per job, over the last 5 minutes                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Recommended alert thresholds | <p><b>Yellow warning:</b> Dynamic<br/> <b>Red critical:</b> Dynamic</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| Recommended response         | <p>To increase throughput and maintain low latency when the number of in-flight requests is high, scale UAA VMs horizontally by editing the UAA VM field in the <b>Resource Config</b> pane of the PAS tile. Ensure that the <code>system.cpu.user</code> metric for UAA is not sustained in the suggested range of 80-90% maximum CPU utilization.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                            |

## System (BOSH) Metrics

### VM Health

| system.healthy               |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                  | <p><code>1</code> means the system is healthy, and <code>0</code> means the system is not healthy.</p> <p><b>Use:</b> This is the most important BOSH metric to monitor. It indicates if the VM emitting the metric is healthy. Review this metric for all VMs to estimate the overall health of the system.</p> <p>Multiple unhealthy VMs signals problems with the underlying IAAS layer.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (Float, 0-1)<br/> <b>Frequency:</b> 60 s</p> |
| Recommended measurement      | Average over the last 5 minutes                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| Recommended alert thresholds | <p><b>Yellow warning:</b> N/A<br/> <b>Red critical:</b> &lt; 1</p>                                                                                                                                                                                                                                                                                                                                                                                                                              |
| Recommended response         | Investigate CF logs for the unhealthy component(s).                                                                                                                                                                                                                                                                                                                                                                                                                                             |

### VM Memory Used

| system.mem.percent      |                                                                                                                                                                                                                                                   |
|-------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description             | <p>System Memory — Percentage of memory used on the VM</p> <p><b>Use:</b> Set an alert and investigate if the free RAM is low over an extended period.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (%)<br/> <b>Frequency:</b> 60 s</p> |
| Recommended measurement | Average over the last 10 minutes                                                                                                                                                                                                                  |

|                              |                                                                                                                                     |
|------------------------------|-------------------------------------------------------------------------------------------------------------------------------------|
| Recommended alert thresholds | <b>Yellow warning:</b> $\geq 80\%$<br><b>Red critical:</b> $\geq 90\%$                                                              |
| Recommended response         | The response depends on the job the metric is associated with. If appropriate, scale affected jobs out and monitor for improvement. |

## VM Disk Used

| system.disk.system.percent   |                                                                                                                                                                                                    |
|------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                  | System disk — Percentage of the system disk used on the VM<br><br>Use: Set an alert to indicate when the system disk is almost full.<br><br>Origin: Firehose<br>Type: Gauge (%)<br>Frequency: 60 s |
| Recommended measurement      | Average over the last 30 minutes                                                                                                                                                                   |
| Recommended alert thresholds | <b>Yellow warning:</b> $\geq 80\%$<br><b>Red critical:</b> $\geq 90\%$                                                                                                                             |
| Recommended response         | Investigate what is filling the jobs system partition.<br>This partition should not typically fill because BOSH deploys jobs to use ephemeral and persistent disks.                                |

## VM Ephemeral Disk Used

| system.disk.ephemeral.percent |                                                                                                                                                                                                                                                        |
|-------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                   | Ephemeral disk — Percentage of the ephemeral disk used on the VM<br><br>Use: Set an alert and investigate if the ephemeral disk usage is too high for a job over an extended period.<br><br>Origin: Firehose<br>Type: Gauge (%)<br>Frequency: 60 s     |
| Recommended measurement       | Average over the last 30 minutes                                                                                                                                                                                                                       |
| Recommended alert thresholds  | <b>Yellow warning:</b> $\geq 80\%$<br><b>Red critical:</b> $\geq 90\%$                                                                                                                                                                                 |
| Recommended response          | <ol style="list-style-type: none"> <li>Run <code>bosh vms --details</code> to view jobs on affected deployments.</li> <li>Determine cause of the data consumption, and, if appropriate, increase disk space or scale out the affected jobs.</li> </ol> |

## VM Persistent Disk Used

| system.disk.persistent.percent |                                                                                                                                                                                                                                                           |
|--------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                    | Persistent disk — Percentage of persistent disk used on the VM<br><br>Use: Set an alert and investigate further if the persistent disk usage for a job is too high over an extended period.<br><br>Origin: Firehose<br>Type: Gauge (%)<br>Frequency: 60 s |
| Recommended measurement        | Average over the last 30 minutes                                                                                                                                                                                                                          |
|                                | <b>Yellow warning:</b> $\geq 80\%$                                                                                                                                                                                                                        |

|                                     |                                                                                                                                                                                                             |
|-------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Recommended alert thresholds</b> | <b>Red critical:</b> $\geq 90\%$                                                                                                                                                                            |
| <b>Recommended response</b>         | <p>1. Run <code>bosh vms --details</code> to view jobs on affected deployments.</p> <p>2. Determine cause of the data consumption, and, if appropriate, increase disk space or scale out affected jobs.</p> |

## VM CPU Utilization

| system.cpu.user                     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                  | <p>CPU utilization — The percentage of CPU spent in user processes</p> <p><b>Use:</b> Set an alert and investigate further if the CPU utilization is too high for a job.</p> <p>For monitoring Gorouter performance, CPU utilization of the Gorouter VM is the recommended key capacity scaling indicator. For more information, see <a href="#">Gorouter Latency and Throughput</a>.</p> <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (%)<br/> <b>Frequency:</b> 60 s</p> |
| <b>Recommended measurement</b>      | Average over the last 5 minutes                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| <b>Recommended alert thresholds</b> | <p><b>Yellow warning:</b> <math>\geq 85\%</math></p> <p><b>Red critical:</b> <math>\geq 95\%</math></p>                                                                                                                                                                                                                                                                                                                                                                          |
| <b>Recommended response</b>         | <p>1. Investigate the cause of the spike.</p> <p>2. If the cause is a normal workload increase, then scale up the affected jobs.</p>                                                                                                                                                                                                                                                                                                                                             |

## Key Capacity Scaling Indicators

This topic describes key capacity scaling indicators that operators monitor to determine when they need to scale their Pivotal Application Service (PAS) deployments.

Pivotal provides these indicators to operators as general guidance for capacity scaling. Each indicator is based on platform metrics from different components. This guidance is applicable to most PAS v2.4 deployments. Pivotal recommends that operators fine-tune the suggested alert thresholds by observing historical trends for their deployments.

## Diego Cell Capacity Scaling Indicators

There are three key capacity scaling indicators recommended for Diego cell:

- [Diego Cell Memory Capacity](#) is a measure of the percentage of remaining memory capacity
- [Diego Cell Disk Capacity](#) is a measure of the percentage of remaining disk capacity
- [Diego Cell Container Capacity](#) is a measure of the percentage of remaining container capacity

### Diego Cell Memory Capacity

| rep.CapacityRemainingMemory / rep.CapacityTotalMemory |                                                                                                                                                                                                                                                                                                                                                                                                                                         |
|-------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                                    | Percentage of remaining memory capacity for a given cell. Monitor this derived metric across all cells in a deployment.<br><br>The metric <code>rep.CapacityRemainingMemory</code> indicates the remaining amount in MiB of memory available for this cell to allocate to containers.<br>The metric <code>rep.CapacityTotalMemory</code> indicates the total amount in MiB of memory available for this cell to allocate to containers. |
| <b>Purpose</b>                                        | A best practice deployment of Cloud Foundry includes three availability zones (AZs). For these types of deployments, Pivotal recommends that you have enough capacity to suffer failure of an entire AZ.<br><br>The <i>Recommended threshold</i> assumes a three-AZ configuration. Adjust the threshold percentage if you have more or fewer AZs.                                                                                       |
| <b>Recommended thresholds</b>                         | < avg(35%)                                                                                                                                                                                                                                                                                                                                                                                                                              |
| <b>How to scale</b>                                   | Scale up your Diego Cells                                                                                                                                                                                                                                                                                                                                                                                                               |
| <b>Additional details</b>                             | Origin: Firehose<br>Type: Gauge (%)<br>Frequency: Emitted every 60 s<br>Applies to: cf:diego_cells                                                                                                                                                                                                                                                                                                                                      |
| <b>Alternative Metric</b>                             | PCF Healthwatch expresses this indicator with the metric <code>healthwatch.Diego.TotalPercentageAvailableMemoryCapacity.5M</code> .                                                                                                                                                                                                                                                                                                     |

### Diego Cell Disk Capacity

| rep.CapacityRemainingDisk / rep.CapacityTotalDisk |                                                                                                                                                                                                                                                                                                                                                                                                                               |
|---------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                                | Percentage of remaining disk capacity for a given cell. Monitor this derived metric across all cells in a deployment.<br><br>The metric <code>rep.CapacityRemainingDisk</code> indicates the remaining amount in MiB of disk available for this cell to allocate to containers.<br>The metric <code>rep.CapacityTotalDisk</code> indicates the total amount in MiB of disk available for this cell to allocate to containers. |
|                                                   | A best practice deployment of Cloud Foundry includes three availability zones (AZs). For these types of deployments, Pivotal recommends that you have enough capacity to suffer failure of an entire AZ.                                                                                                                                                                                                                      |

|                               |                                                                                                                                                        |
|-------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Purpose</b>                | The <i>Recommended threshold</i> assumes a three-AZ configuration. Adjust the threshold percentage if you have more or fewer AZs.                      |
| <b>Recommended thresholds</b> | < avg(35%)                                                                                                                                             |
| <b>How to scale</b>           | Scale up your Diego Cells                                                                                                                              |
| <b>Additional details</b>     | <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (%)<br/> <b>Frequency:</b> Emitted every 60 s<br/> <b>Applies to:</b> cf:diego_cells</p>            |
| <b>Alternative Metric</b>     | PCF Healthwatch expresses this indicator with the metric<br><code>healthwatch.Diego.TotalPercentageAvailableDiskCapacity.5M</code> <a href="#">🔗</a> . |

## Diego Cell Container Capacity

| rep.CapacityRemainingContainers / rep.CapacityTotalContainers |                                                                                                                                                                                                                                                                                                                                                                                             |
|---------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                                            | <p>Percentage of remaining container capacity for a given cell. Monitor this derived metric across all cells in a deployment.</p> <p>The metric <code>rep.CapacityRemainingContainers</code> indicates the remaining number of containers this cell can host.</p> <p>The metric by <code>rep.CapacityTotalContainer</code> indicates the total number of containers this cell can host.</p> |
| <b>Purpose</b>                                                | <p>A best practice deployment of Cloud Foundry includes three availability zones (AZs). For these types of deployments, Pivotal recommends that you have enough capacity to suffer failure of an entire AZ.</p> <p>The <i>Recommended threshold</i> assumes a three-AZ configuration. Adjust the threshold percentage if you have more or fewer AZs.</p>                                    |
| <b>Recommended thresholds</b>                                 | < avg(35%)                                                                                                                                                                                                                                                                                                                                                                                  |
| <b>How to scale</b>                                           | Scale up your Diego Cells                                                                                                                                                                                                                                                                                                                                                                   |
| <b>Additional details</b>                                     | <p><b>Origin:</b> Firehose<br/> <b>Type:</b> Gauge (%)<br/> <b>Frequency:</b> Emitted every 60 s<br/> <b>Applies to:</b> cf:diego_cells</p>                                                                                                                                                                                                                                                 |
| <b>Alternative Metric</b>                                     | PCF Healthwatch expresses this indicator with the metric<br><code>healthwatch.Diego.TotalPercentageAvailableContainerCapacity.5M</code> <a href="#">🔗</a> .                                                                                                                                                                                                                                 |

## Firehose Performance Scaling Indicators

Pivotal recommends two key capacity scaling indicators for monitoring Firehose performance.

### Log Transport Loss Rate

| loggregator.doppler.dropped{direction=ingress} / loggregator.doppler.ingress |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                                                           | <p>This derived value represents the loss rate occurring as messages are transported from the Loggregator Agent components through the Doppler components to the Firehose endpoints.</p> <p>Metric <code>loggregator.doppler.ingress</code> represents the number of messages entering Doppers for transport through the firehose, and <code>loggregator.doppler.dropped</code> represents the number of messages dropped without delivery.</p> <p>Messages include the combined stream of logs from all apps and the metrics data from Cloud Foundry components.</p> <p>For more information about Loggregator components, see <a href="#">Loggregator Architecture</a>.</p> |
|                                                                              | Excessive dropped messages can indicate the Doppers or Traffic Controllers are not processing messages quickly enough.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |

|                               |                                                                                                                                                                                                                                                                                                                                                                                    |
|-------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Purpose</b>                | The recommended scaling indicator is a dropped message rate greater than <code>0.01</code> . This scaling indicator is calculated by expressing the total number of dropped messages as a percentage of the total throughput and scale.                                                                                                                                            |
|                               | Doppler emits two separate dropped metrics, one for ingress and one for egress. The envelopes have a <code>direction</code> . For this indicator, use the metric with a <code>direction</code> tag with a value of <code>ingress</code> .                                                                                                                                          |
| <b>Recommended thresholds</b> | <p><b>Scale indicator:</b> <math>\geq 0.01</math></p> <p>If alerting:</p> <ul style="list-style-type: none"> <li><b>Yellow warning:</b> <math>\geq 0.005</math></li> <li><b>Red critical:</b> <math>\geq 0.01</math></li> </ul>                                                                                                                                                    |
| <b>How to scale</b>           | <p>Scale up the number of Traffic Controller and Doppler instances.</p> <p><b>Note:</b> At approximately 40 Doppler instances and 20 Traffic Controller instances, horizontal scaling is no longer useful for improving Firehose performance. To improve performance, add vertical scale to the existing Doppler and Traffic Controller instances by increasing CPU resources.</p> |
| <b>Additional details</b>     | <p><b>Origin:</b> Firehose</p> <p><b>Type:</b> Gauge (float)</p> <p><b>Frequency:</b> Base metrics are emitted every 5 s</p> <p><b>Applies to:</b> cf:doppler</p>                                                                                                                                                                                                                  |
| <b>Alternative Metrics</b>    | PCF Healthwatch expresses this indicator with the metrics <code>healthwatch.Firehose.LossRate.1H</code> and <code>healthwatch.Firehose.LossRate.1M</code> .                                                                                                                                                                                                                        |

## Doppler Message Rate Capacity

| <code>loggregator.doppler.ingress (sum across instances) / current number of Doppler instances</code> |                                                                                                                                                                                                                                                                                |
|-------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                                                                                    | This derived value represents the average rate of envelopes (messages) per Doppler instance. Deriving this into a per-Doppler envelopes-per-second, or envelopes-per-minute, rate can indicate the need to scale when Doppler instances are at their recommended maximum load. |
| <b>Purpose</b>                                                                                        | The recommended scaling indicator is to look at the average load on the Doppler instances, and increase the number of Doppler instances when the derived rate is 16,000 envelopes per second, or 1 million envelopes per minute.                                               |
| <b>Recommended thresholds</b>                                                                         | <b>Scale indicator:</b> $\geq 16,000$ envelopes per second (or 1 million envelopes per minute)                                                                                                                                                                                 |
| <b>How to scale</b>                                                                                   | Increase the number of Doppler VMs in the <b>Resource Config</b> pane of the PAS tile.                                                                                                                                                                                         |
| <b>Additional details</b>                                                                             | <p><b>Origin:</b> Firehose</p> <p><b>Type:</b> Gauge (float)</p> <p><b>Frequency:</b> Emitted every 15 s</p> <p><b>Applies to:</b> cf:doppler</p>                                                                                                                              |
| <b>Alternative Metric</b>                                                                             | PCF Healthwatch expresses this indicator with the metric <code>healthwatch.Doppler.MessagesAverage.1M</code> .                                                                                                                                                                 |

## Reverse Log Proxy Loss Rate

| <code>loggregator.rlp.dropped / loggregator.rlp.ingress</code> |                                                                                                                                                                                                                                                                                                                                                                                                                                             |
|----------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                                             | <p>The loss rate of the reverse log proxies (RLP), that is, the total messages dropped as a percentage of the total traffic coming through the <a href="#">reverse log proxy</a>. Total messages include only logs for bound applications.</p> <p>This loss rate is specific to the RLP and does not impact the Firehose loss rate. For example, you can suffer lossiness in the RLP while not suffering any lossiness in the Firehose.</p> |
| <b>Purpose</b>                                                 | <p>Excessive dropped messages can indicate that the RLP is overloaded and that the Traffic Controllers need to be scaled.</p> <p>The recommended scaling indicator is to look at the maximum per minute loss rate over a 5-minute window and scale if the derived loss rate value grows greater than <code>0.1</code>.</p>                                                                                                                  |
| <b>Recommended thresholds</b>                                  | <p><b>Scale indicator:</b> <math>\geq 0.1</math></p> <p>If alerting:</p> <ul style="list-style-type: none"> <li><b>Yellow warning:</b> <math>\geq 0.01</math></li> <li><b>Red critical:</b> <math>\geq 0.1</math></li> </ul>                                                                                                                                                                                                                |
| <b>How to scale</b>                                            | Scale up the number of traffic controller instances to further balance log load.                                                                                                                                                                                                                                                                                                                                                            |
|                                                                | <b>Origin:</b> Firehose                                                                                                                                                                                                                                                                                                                                                                                                                     |

|                    |                                                                                                                                                                                                     |
|--------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Additional details | Type: Counter (Integer)<br>Frequency: Emitted every 60 s<br>Applies to: cf:loggregator                                                                                                              |
| Alternative Metric | PCF Healthwatch expresses this indicator with the metric <code>healthwatch.SyslogDrain.RLP.LossRate.1M</code>  . |

## Firehose Consumer Scaling Indicator

Pivotal recommends the following scaling indicator for monitoring the performance of consumers of the Firehose.

### Slow Consumer Drops

| doppler_proxy.slow_consumer |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
|-----------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                 | Within PAS, metrics and logs enter the Firehose for transport and exit out of the platform via a consumer nozzle. If the consuming downstream system fails to keep up with the exiting stream of metrics, the Firehose is forced to close the connection to protect itself from back-pressure. The Firehose increments <code>slow_consumer</code> with each connection that it closes because a consumer could not keep up.                                                                                                                                                                                                                                                                                       |
| Purpose                     | This metric indicates that a Firehose consumer, such as a monitoring tool nozzle, is ingesting too slowly. If this number is anomalous, it may result in the downstream monitoring tool not having all expected data, even though that data was successfully transported through the Firehose.<br><br>This can also be confirmed via the Loggregator emitted log message <code>TB: Output channel too full. Dropped N messages</code> , where N is the number of dropped messages.                                                                                                                                                                                                                                |
| Recommended thresholds      | <b>Scale indicator:</b> It is recommended to scale when the rate of Firehose Slow Consumer Drops is anomalous for a given environment.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| How to scale                | Scale up the number of nozzle instances. You can scale a nozzle using the subscription ID specified when the nozzle connects to the Firehose. If you use the same subscription ID on each nozzle instance, the Firehose evenly distributes data across all instances of the nozzle. For example, if you have two nozzle instances with the same subscription ID, the Firehose sends half of the data to one nozzle instance and half to the other. Similarly, if you have three nozzle instances with the same subscription ID, the Firehose sends one-third of the data to each instance. If you want to scale a nozzle, the number of nozzle instances should match the number of Traffic Controller instances. |
| Additional details          | Origin: Firehose<br>Type: Counter<br>Frequency: Emitted every 5 s<br>Applies to: cf:doppler                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |

### Reverse Log Proxy Egress Dropped Messages

| rlp.dropped, direction: egress |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|--------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                    | Within PAS, logs and metrics enter Loggregator for transport and then egress through the Reverse Log Proxy (RLP). The RLP drops messages when consumers of the RLP, such as monitoring tool nozzles, ingest the exiting stream of logs and metrics too slowly.<br><br> <b>Note:</b> The <code>rlp.dropped</code> metric includes both <code>ingress</code> and <code>egress</code> directions. To differentiate between <code>ingress</code> and <code>egress</code> , refer to the <code>direction</code> tag on the metric.                                               |
| Purpose                        | This metric indicates that a consumer of logs and metrics from the RLP, such as a monitoring tool nozzle, is ingesting RLP messages too slowly.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
| Recommended thresholds         | <b>Scale indicator:</b> Scale when the rate of <code>rlp.dropped, direction: egress</code> metrics is continuously increasing.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
| How to scale                   | Scale up the number of nozzle instances. The number of nozzle instances should match the number of Traffic Controller instances. You can scale a nozzle using the subscription ID specified when the nozzle connects to the RLP. If you use the same subscription ID on each nozzle instance, the RLP evenly distributes data across all instances of the nozzle. For example, if you have two nozzle instances with the same subscription ID, the RLP sends half of the data to one nozzle instance and half to the other. Similarly, if you have three nozzle instances with the same subscription ID, the RLP sends one-third of the data to each instance. |
| Additional details             | Origin: Reverse Log Proxy<br>Type: Counter<br>Frequency: Emitted every 5 s                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

Applies to: cf:loggregator\_trafficcontroller

## Doppler Egress Dropped Messages

| doppler.dropped, direction: egress |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |
|------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                        | Within PAS, logs and metrics enter Loggregator for transport and then egress through Doppler. Doppler drops messages when consumers of the RLP, such as monitoring tool nozzles, ingest the exiting stream of logs and metrics too slowly.<br><br>💡 Note: The <code>doppler.dropped</code> metric includes both <code>ingress</code> and <code>egress</code> directions. To differentiate between <code>ingress</code> and <code>egress</code> , refer to the <code>direction</code> tag on the metric.                                                                                                                                                        |
| Purpose                            | This metric indicates that a consumer of logs and metrics from the RLP, such as a monitoring tool nozzle, is ingesting too slowly.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| Recommended thresholds             | <b>Scale indicator:</b> Scale when the rate of <code>doppler.dropped, direction: egress</code> metrics is continuously increasing.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| How to scale                       | Scale up the number of nozzle instances. The number of nozzle instances should match the number of Traffic Controller instances. You can scale a nozzle using the subscription ID specified when the nozzle connects to the RLP. If you use the same subscription ID on each nozzle instance, the RLP evenly distributes data across all instances of the nozzle. For example, if you have two nozzle instances with the same subscription ID, the RLP sends half of the data to one nozzle instance and half to the other. Similarly, if you have three nozzle instances with the same subscription ID, the RLP sends one-third of the data to each instance. |
| Additional details                 | Origin: Doppler<br>Type: Counter<br>Frequency: Emitted every 5 s<br>Applies to: cf:doppler                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |

## CF Syslog Drain Performance Scaling Indicators

There are three key capacity scaling indicators recommended for CF Syslog Drain performance.

💡 Note: These CF Syslog Drain scaling indicators are only relevant if your deployment contains apps using the CF syslog drain binding feature.

💡 Note: If you enable agent-based syslog on your deployment, the Adapter Loss Rate and CF Syslog Drain Bindings Count indicators are not relevant and should be ignored. For more information about the Adapter Loss Rate and CF Syslog Drain Bindings Count indicators, see [Adapter Loss Rate](#) and [CF Syslog Drain Bindings Count](#). For more information about enabling agent-based syslog, see [Loggregator Syslog Agent Increases Scale For Syslog Drains](#).

## Adapter Loss Rate

| cf-syslog-drain.adapter.dropped / cf-syslog-drain.adapter.ingress |                                                                                                                                                                                                                                                                                                                                                                                                                                      |
|-------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                                                       | The loss rate of the Syslog Adapters, that is, the total messages dropped as a percentage of the total traffic coming through the <a href="#">Syslog Adapters</a> . Total messages include only logs for bound applications.<br><br>This loss rate is specific to the Syslog Adapters and does not impact the Firehose loss rate. For example, you can suffer lossiness in syslog while not suffering any lossiness in the Firehose. |
| Purpose                                                           | Indicates that the syslog drains are not keeping up with the number of logs that a syslog-drain-bound app is producing. This likely means that the syslog-drain consumer is failing to keep up with the incoming log volume.<br><br>The recommended scaling indicator is to look at the maximum per minute loss rate over a 5-minute window and scale if the derived loss rate value grows greater than <code>0.1</code> .           |
| Recommended thresholds                                            | <b>Scale indicator:</b> $\geq 0.1$<br>If alerting:<br><b>Yellow warning:</b> $\geq 0.01$<br><b>Red critical:</b> $\geq 0.1$                                                                                                                                                                                                                                                                                                          |
| How to scale                                                      | Performance test your syslog server, review the logs of the syslog consuming system for intake and other performance issues that indicate a need to scale the consuming system.                                                                                                                                                                                                                                                      |

|                    |                                                                                                                        |
|--------------------|------------------------------------------------------------------------------------------------------------------------|
| Additional details | Origin: Firehose<br>Type: Counter (Integer)<br>Frequency: Emitted every 60 s<br>Applies to: cf:cf-syslog               |
| Alternative Metric | PCF Healthwatch expresses this indicator with the metric <a href="#">healthwatch.SyslogDrain.Adapter.LossRate.1M</a> . |

## CF Syslog Drain Bindings Count

| cf-syslog-drain.drain_adapter.drain_bindings / number of Syslog Adapters |                                                                                                                                                                                                                                                            |
|--------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                                                              | The number of CF syslog drain bindings. CF syslog drain bindings enable app syslog drains by managing the connections between the Syslog Adapters and the Reverse Log Proxies (RLPs).                                                                      |
| Purpose                                                                  | Each syslog drain binding requires two Syslog Adapters. A configuration with two Syslog Adapters allows for approximately 500 syslog drain bindings. Pivotal recommends adding one Syslog Adapter instance for every 250 additional syslog drain bindings. |
| Recommended thresholds                                                   | <b>Scale indicator:</b> $\geq 450$ syslog drain bindings<br><br>Pivotal recommends scaling to three Syslog Adapters if a maximum of 450 syslog drain bindings is reached in a one hour window.                                                             |
| How to scale                                                             | Increase the number of Syslog Adapter VMs in the <b>Resource Config</b> pane of the PAS tile.                                                                                                                                                              |
| Additional details                                                       | Origin: Firehose<br>Type: Gauge (float)<br>Frequency: Emitted every 60 s<br>Applies to: cf:cf-syslog                                                                                                                                                       |
| Alternative Metric                                                       | PCF Healthwatch expresses this indicator with the metric <a href="#">healthwatch.SyslogDrain.Adapter.BindingsAverage.5M</a> .                                                                                                                              |

## Syslog Agent Loss Rate

**Note:** The Syslog Agent Loss Rate indicator is only relevant if agent-based syslog is enabled on your deployment. Agent-based syslog is disabled by default. For more information about enabling agent-based syslog, see [Loggregator Syslog Agent Increases Scale For Syslog Drains](#).

| loggregator-agent.loggr-syslog-agent.dropped{direction:egress} / loggregator-agent.loggr-syslog-agent.dropped{direction:ingress} |                                                                                                                                                                                                                                                                                                                                                                                                  |
|----------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                                                                                                                      | The loss rate of Syslog Agents. The loss rate of Syslog Agents is the messages dropped as a percentage of total message traffic through Syslog Agents. The message traffic through Syslog Agents includes logs for bound apps.<br><br>The Syslog Agent loss rate does not affect the Firehose loss rate. Message loss can occur in Syslog Agents without message loss occurring in the Firehose. |
| Purpose                                                                                                                          | This metric indicates that the syslog drain consumer is ingesting logs from a syslog-drain-bound app too slowly.                                                                                                                                                                                                                                                                                 |
| Recommended thresholds                                                                                                           | The recommended scaling indicator is the maximum Syslog Agent loss rate per minute within a five-minute window. You should scale up if the maximum loss rate is greater than <a href="#">0.1</a> .<br><br><b>Scale indicator:</b> $\geq 0.1$<br>If alerting:<br><b>Yellow warning:</b> $\geq 0.01$<br><b>Red critical:</b> $\geq 0.1$                                                            |
| How to scale                                                                                                                     | Review the logs of the syslog server for intake issues and other performance issues. Scale the syslog server if necessary.                                                                                                                                                                                                                                                                       |
| Additional details                                                                                                               | Origin: Syslog Agent<br>Type: Counter (Integer)<br>Frequency: Emitted every 60 s                                                                                                                                                                                                                                                                                                                 |

## Log Cache Scaling Indicator

Pivotal recommends the following scaling indicator for monitoring the performance of Log Cache.

## Log Cache Caching Duration

| log_cache.cache_period |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |
|------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description            | This metric indicates the age in milliseconds of the oldest data point stored in Log Cache.                                                                                                                                                                                                                                                                                                                                                                                                                            |
| Purpose                | <p>Log Cache stores all messages that are passed through the Firehose in an ephemeral in-memory store. The size of this store and the cache duration are dependent on the amount of memory available on the VM on which Log Cache runs. Some features of PAS rely on data being available in Log Cache, such as App Autoscaler.</p> <p>Pivotal recommends scaling the VM on which Log Cache runs, so Log Cache can hold all messages that pass through Loggregator in the last 15 minutes, or 900000 milliseconds.</p> |
| Recommended thresholds | <b>Scale indicator:</b> Scale the VM on which Log Cache runs when the cache period drops below 15 minutes, or 900000 milliseconds.<br>Typically, Log Cache runs on the Doppler VM.                                                                                                                                                                                                                                                                                                                                     |
| How to scale           | Scale up the number of Doppler VMs or chose a VM type for Doppler that provides more memory.                                                                                                                                                                                                                                                                                                                                                                                                                           |
| Additional details     | <b>Origin:</b> log-cache<br><b>Type:</b> Gauge<br><b>Frequency:</b> Emitted every 15 s<br><b>Applies to:</b> cf:log-cache                                                                                                                                                                                                                                                                                                                                                                                              |

## Router Performance Scaling Indicator

There is one key capacity scaling indicator recommended for Router performance.

### Router VM CPU Utilization

| system.cpu.user of the Gorouter VM(s) |                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|---------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                           | CPU utilization of the Gorouter VM(s)                                                                                                                                                                                                                                                                                                                                                                                                                                         |
| Purpose                               | <p>High CPU utilization of the Gorouter VMs can increase latency and cause throughput, or requests per/second, to level-off. Pivotal recommends keeping the CPU utilization within a maximum range of 60-70% for best Gorouter performance.</p> <p>If you want to increase throughput capabilities while also keeping latency low, Pivotal recommends scaling the Gorouter while continuing to ensure that CPU utilization does not exceed the maximum recommended range.</p> |
| Recommended thresholds                | <b>Scale indicator:</b> $\geq 60\%$<br>If alerting:<br><b>Yellow warning:</b> $\geq 60\%$<br><b>Red critical:</b> $\geq 70\%$                                                                                                                                                                                                                                                                                                                                                 |
| How to scale                          | Resolve high utilization by scaling the Gorouters horizontally or vertically by editing the <b>Router VM</b> in the <b>Resource Config</b> pane of the PAS tile.                                                                                                                                                                                                                                                                                                              |
| Additional details                    | <b>Origin:</b> Firehose<br><b>Type:</b> Gauge (float)<br><b>Frequency:</b> Emitted every 60 s<br><b>Applies to:</b> cf:router                                                                                                                                                                                                                                                                                                                                                 |

## UAA Performance Scaling Indicator

There is one key capacity scaling indicator recommended for UAA performance.

### UAA VM CPU Utilization

| system.cpu.user of the UAA VM(s) |                                                                                                                                                                                                                                                                                                                                                                             |
|----------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Description                      | CPU utilization of the UAA VM(s)                                                                                                                                                                                                                                                                                                                                            |
| Purpose                          | <p>High CPU utilization of the UAA VMs can increase latency and cause throughput, or requests per/second, to level-off. Pivotal recommends keeping the CPU utilization within a maximum range of 80-90% for best UAA performance.</p> <p>If you want to increase throughput capabilities while keeping latency low, Pivotal recommends scaling the UAA VMs and ensuring</p> |

|                               |                                                                                                                                                                                                                                                                 |
|-------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|                               | that CPU utilization does not exceed the maximum recommended range.                                                                                                                                                                                             |
| <b>Recommended thresholds</b> | <b>Scale indicator:</b> $\geq 80\%$<br>If alerting:<br><b>Yellow warning:</b> $\geq 80\%$<br><b>Red critical:</b> $\geq 90\%$                                                                                                                                   |
| <b>How to scale</b>           | Resolve high utilization by scaling UAA horizontally or vertically. To scale UAA, navigate to the <b>Resource Config</b> pane of the PAS tile and edit the number of your <b>UAA VM</b> instances or change the VM type to a type that utilizes more CPU cores. |
| <b>Additional details</b>     | <b>Origin:</b> Firehose<br><b>Type:</b> Gauge (float)<br><b>Frequency:</b> Emitted every 60 s<br><b>Applies to:</b> cf:uaa                                                                                                                                      |

## NFS/WebDAV Backed Blobstore

There is one key capacity scaling indicator for external S3 external storage.

 **Note:** This metric is only relevant if your deployment does not use an external S3 repository for external storage with no capacity constraints.

| system.disk.persistent.percent of NFS server VM(s) |                                                                                                                                                                                                                                                                                                    |
|----------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Description</b>                                 | <i>If applicable:</i> Monitor the percentage of persistent disk used on the VM for the NFS Server job.                                                                                                                                                                                             |
| <b>Purpose</b>                                     | If you do not use an external S3 repository for external storage with no capacity constraints, you must monitor the PAS object store to push new app and buildpacks.<br><br>If you use an internal NFS/WebDAV backed blobstore, consider scaling the persistent disk when it reaches 75% capacity. |
| <b>Recommended thresholds</b>                      | $\geq 75\%$                                                                                                                                                                                                                                                                                        |
| <b>How to scale</b>                                | Give your NFS Server additional persistent disk resources.                                                                                                                                                                                                                                         |
| <b>Additional details</b>                          | <b>Origin:</b> Firehose<br><b>Type:</b> Gauge (%)<br><b>Applies to:</b> cf:nfs_server                                                                                                                                                                                                              |

## Selecting and Configuring a Monitoring System

This topic describes considerations for selecting and configuring a system to continuously monitor Pivotal Cloud Foundry (PCF) component performance and health.

### Selecting a Monitoring Platform

Pivotal recommends using [PCF Healthwatch](#) to monitor your deployment. PCF Healthwatch is a service tile developed and supported by Pivotal and available on [Pivotal Network](#).

Many third-party systems can also be used to monitor a PCF deployment.

### Monitoring Platform Types

Monitoring platforms support two types of monitoring:

- A *dashboard* for active monitoring when you are at a keyboard and screen
- Automated *alerts* for when your attention is elsewhere

Some monitoring solutions offer both in one package. Others require putting the two pieces together.

### Monitoring Platforms

There are many monitoring options available, both open source and commercial products. Some commonly-used platforms among PCF customers include:

- [PCF Healthwatch](#) by Pivotal
- PCF Partner Services available on [Pivotal Network](#):
  - [AppDynamics](#)
  - [Datadog](#)
  - [Dynatrace](#)
  - [New Relic](#)
  - [SignalFx](#)
  - [WaveFront by VMware](#)
- Other Commercial Services
  - [VMware vRealize Operations \(vROPS\)](#)
- Open Source Tooling
  - [Prometheus + Grafana](#)
  - [OpenTSDB](#)

### Pivotal Cloud Ops Tools

The Pivotal Cloud Ops Team manages two types of deployments for internal Pivotal use: open-source Cloud Foundry, and Pivotal Cloud Foundry (PCF).

For [Cloud Foundry](#), Pivotal Cloud Ops uses several monitoring tools. The [Datadog Config repository](#) provides an example of how the Pivotal Cloud Ops team uses a customized Datadog dashboard to monitor the health of its open-source Cloud Foundry deployments.

To monitor its PCF deployments, Pivotal Cloud Ops leverages a combination of [PCF Healthwatch](#) and [Google Stackdriver](#).

### Key Inputs for Platform Monitoring

## BOSH VM and PCF Component Health Metrics

Most monitoring service tiles for PCF come packaged with the Firehose nozzle necessary to extract the BOSH and CF metrics leveraged for platform monitoring. Nozzles are programs that consume data from the Loggregator Firehose. Nozzles can be configured to select, buffer, and transform data, and to forward it to other apps and services.

The nozzles gather the component logs and metrics streaming from the Loggregator Firehose endpoint. For more information about the Firehose, see [Loggregator Architecture](#).

As of PCF v2.0, both BOSH VM Health metrics and Cloud Foundry component metrics stream through the Firehose by default.

- PCF component metrics originate from the Metron agents on their source components, then travel through Dopplers to the Traffic Controller.
- The Traffic Controller aggregates both metrics and log messages system-wide from all Dopplers, and emits them from its Firehose endpoint.

The following topic list high-signal-value metrics and capacity scaling indicators in a PCF deployment:

- [Key Performance Indicators](#)
- [Key Capacity Scaling Indicators](#)

## Continuous Functional Smoke Tests

PCF includes [smoke tests](#), which are functional unit and integration tests on all major system components. By default, whenever an operator upgrades to a new version of PAS, these smoke tests run as a post-deploy errand.

Pivotal recommends additional higher-resolution monitoring by the execution of continuous smoke tests, or Service Level Indicator tests, that measure user-defined features and check them against expected levels.

- [PCF Healthwatch](#) automatically executes these tests for PAS Service Level Indicators.
- The Pivotal Cloud Ops [CF Smoke Tests](#) repository offers additional testing examples.

See the [Metrics](#) topic in the Concourse documentation for how to set up Concourse to generate custom component metrics.

## Warning and Critical Thresholds

To properly configure your monitoring dashboard and alerts, you must establish what thresholds should drive alerting and red/yellow/green dashboard behavior.

Some key metrics have more fixed thresholds, with similar threshold numbers recommended across different foundations and use cases. These metrics tend to revolve around the health and performance of key components that can impact the performance of the entire system.

Other metrics of operational value are more dynamic in nature. This means that you must establish a baseline and yellow/red thresholds suitable for your system and its use cases. You can establish initial baselines by watching values of key metrics over time and noting what seems to be a good starting threshold level that divides acceptable and unacceptable system performance and health.

## Continuous Evolution

Effective platform monitoring requires continuous evolution.

After you establish initial baselines, Pivotal recommends that you continue to refine your metrics and tests to maintain the appropriate balance between early detection and reducing unnecessary alert fatigue. The dynamic measures recommended in [Key Performance Indicators](#) and [Key Capacity Scaling Indicators](#) should be revisited on occasion to ensure they are still appropriate to the current system configuration and its usage patterns.

## Pivotal Application Service for Windows

### Overview

Pivotal Application Service (PAS) for Windows enables operators to provision, operate, and manage Windows cells on Pivotal Cloud Foundry (PCF).

After operators install the PAS for Windows tile on the Ops Manager Installation Dashboard, developers can push .NET applications to Windows cells using the [Cloud Foundry Command Line Interface](#) (cf CLI).

For more information about how PAS for Windows works, see the [Product Architecture](#) topic.

For information about deploying .NET apps to PAS for Windows, see the [Tips for .NET Developers](#) topic.

For the PAS for Windows release notes, see the [PAS for Windows v2.6 Release Notes](#) topic.

### Requirements

To install the PAS for Windows tile, you must have the following:

- Ops Manager v2.6 and PAS v2.6 deployed to vSphere, Amazon Web Services (AWS), Google Cloud Platform (GCP), or Azure
- A Windows stemcell, which you can obtain by following the directions in [Downloading or Creating Windows Stemcells](#)

The minimum resource requirements for each Windows cell are as follows:

- Disk size: 64 GB (100 GB recommended)
- Memory: 16 GB
- CPUs: 4

### Contents

#### About PAS for Windows

- [Product Architecture](#)

#### Installation Guide

- [Downloading or Creating Windows Stemcells](#)
- [Creating a Windows Stemcell for vSphere Manually](#)
- [Creating a Windows Stemcell for vSphere Using stembuild \(Beta\)](#)
- [Installing and Configuring PASW](#)
- [Windows Cells in Isolation Segments](#)
- [Migrating Apps to PAS for Windows](#)

#### Admin Guide

- [Upgrading Windows Cells](#)
- [Troubleshooting Windows Cells](#)

#### Developer Guide

- [Deploying .NET Apps](#)
- [Using SMB Volumes in .NET Apps](#)
- [Changing Windows Stacks](#)
- [Tips for .NET Developers](#)

## Limitations

PAS for Windows has the following limitations:

- Developers cannot push Docker or other OCI-compatible images to Windows Diego cells.
- Container-to-container networking is not available for Windows-hosted applications.
- OpenStack is currently not supported for PAS for Windows. Contact your Pivotal representative for information about OpenStack deployments.
- SMB is supported, but Volume Services is not supported.
- Due to [Known Issue #244 ↗](#) in Windows Server OS, applications hosted on PAS for Windows cannot route traffic when deployed with the IPsec add-on for PCF.

## Product Architecture

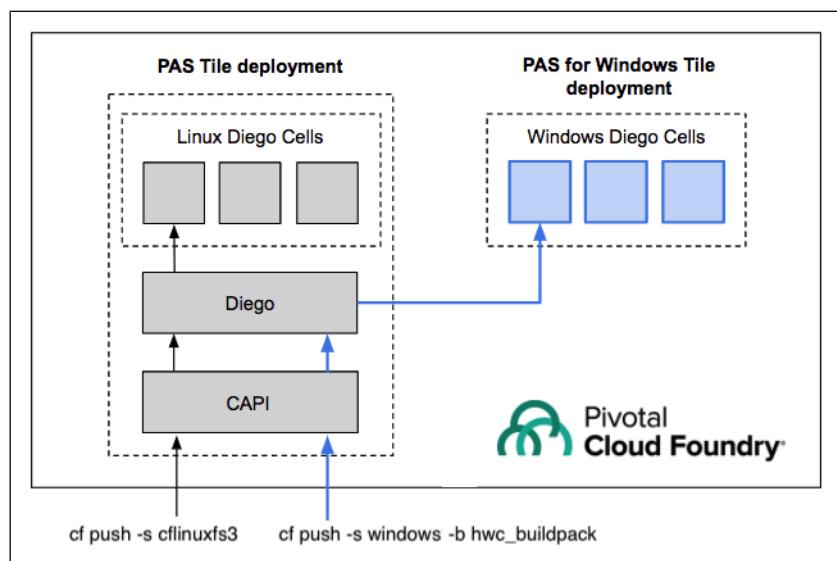
This topic describes the architecture of Windows cells that PAS for Windows deploys to run containerized .NET apps, and the stemcells that it supplies to BOSH as the operating system for the Windows cell VMs.

### Overview

Operators who want to run Windows cells in PCF to enable developers to push .NET apps can deploy the PAS for Windows tile.

Deploying this tile creates a separate BOSH deployment populated with the Garden Windows release, which runs on a Windows cell built from a Windows stemcell.

Once the Windows cell is running, developers can specify a Windows stack when pushing .NET apps from the command line. PCF passes the app to the Windows cell in the PAS for Windows BOSH deployment. The diagram below illustrates the process.



## About Windows Cells

App instances in PCF run inside containers. [Garden](#) is the API that creates and manages these containers. An implementation equivalent to that on Linux cells provides this infrastructure on Windows cells, utilizing native Windows Server Containers.

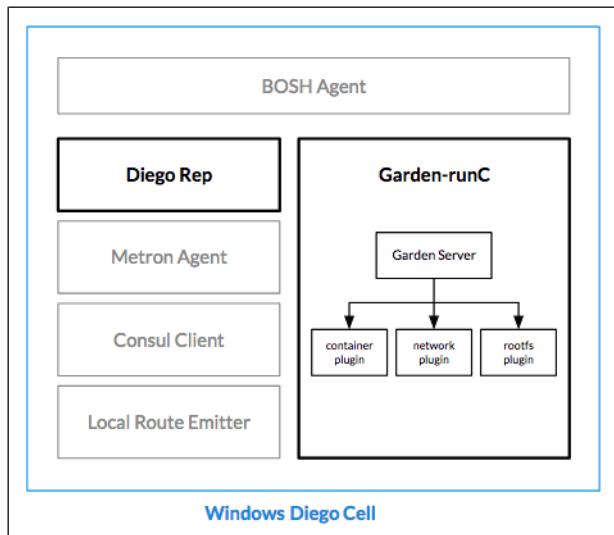
By installing the [PAS for Windows](#) tile, operators create a Windows [cell](#) from a [stemcell](#) that contains the Windows Server operating system. Garden on Windows uses [Windows Containers](#) to isolate resources on Windows cells that Cloud Foundry manages alongside Linux cells.

## Components

A Windows cell includes the following components:

- **Guardian:** Implements the [Garden](#) API on Windows
- [Metron Agent](#): Forwards app logs, errors, and metrics to the [Loggregator](#) system
- [BOSH Agent](#): Executes instructions from the BOSH Director
- [Consul Client](#): Registers the cell as a service in a Consul cluster
- [Diego Rep](#): Runs and manages [Tasks and Long Running Processes](#)

The following diagram illustrates the architecture of a Windows cell:



## Container Implementation

Garden on Windows uses the following runtime plugins to create and manage [Windows Containers](#) for PAS:

- Container plugin `winc`: Creates [OCI](#)-compliant containers, executes processes in the containers, and sets their CPU and RAM limits.
- Network plugin `winc-network`: Creates a network compartment for the container, applies its DNS settings, and defines its inbound/outbound network access rules.
- Rootfs image plugin `groot`: Sets up the container filesystem volume and uses the [FSRM](#) API to define its disk usage quotas.

## About Windows Stemcells

A “stemcell” is a customized operating system image containing the filesystem for BOSH-managed virtual machines. When deployed, the operating system includes the *BOSH Agent* process, which is dedicated to communicating with the orchestrating VM, the *BOSH Director*. The BOSH Agent executes and monitors BOSH jobs on its VM.

Deployments of Windows Server on PCF currently use a stemcell containing Windows Server 2019.

See [Downloading or Creating Windows Stemcells](#) for documentation about how to obtain or create a stemcell for PAS for Windows.

## Downloading or Creating Windows Stemcells

This topic describes how to download or create the stemcell that Pivotal Application Service (PAS) for Windows needs to create VMs on an infrastructure.

### Windows Stemcells

BOSH needs a stemcell to supply the basic operating system for any Windows VM it creates. Depending on your IaaS, you can create or download a stemcell as follows:

- **Azure:** [See instructions below](#) on how to activate and download the Azure Windows Stemcell.
- **Google Cloud Platform (GCP):** Download the [GCP light Stemcell for Windows 2019 Server](#) from the [Stemcells for PCF \(Windows Server\)](#) section of [Pivotal Network ↗](#).
- **vSphere:** Create a stemcell by following the directions in [Creating a Windows Stemcell for vSphere Manually](#) or [Creating a Windows Stemcell for vSphere Using stembuild \(Beta\)](#).
- **Amazon Web Services (AWS):** Download the [AWS light Stemcell for Windows 2019 Server](#) from the [Stemcells for PCF \(Windows Server\)](#) section of [Pivotal Network ↗](#).

### Configuring the Azure Light Stemcell

On Azure, the stemcell exists as an offering in the Azure Marketplace. To start using the Azure light stemcell, perform the steps in the following sections:

- [Enable the Azure Light Stemcell](#)
- [Download the Azure Light Stemcell](#)

#### Enable the Azure Light Stemcell

To use the Azure light stemcell, you must first accept the corresponding Microsoft license agreement and enable the stemcell for your non-trial Azure Subscription through the Azure Marketplace as follows:

1. Navigate to <https://portal.azure.com> ↗ and log in.
2. From the options on the left side of the page, click + **Create a resource**.
3. In the **Search the Marketplace** bar, search for **BOSH Stemcell for Windows Server 2019**.
4. Select **BOSH Stemcell for Windows Server 2019** from your search results. A description of the Azure stemcell offering appears.
5. Below the description, at the bottom of the page, click the blue banner that reads **Want to deploy programmatically? Get started ➔**.
6. Review the Terms of Use in the page “Configure Programmatic Deployment” that appears.
7. Under **Choose the subscriptions**, click **Enable** for each Azure subscription with which you want to use the stemcell.
8. Click **Save**.

#### Download the Azure Light Stemcell

To download the Azure light stemcell for use in Ops Manager, perform the following steps:

1. Navigate to [Pivotal Network ↗](#).
2. Select **Azure Light Stemcell for Windows 2019 Server** from the **Release Download Files** section of the [Stemcells for PCF \(Windows\)](#) page.

For information about how to deploy and configure the PAS for Windows tile, see [Installing and Configuring PAS for Windows](#).

## Creating a Windows Stemcell for vSphere Manually

This topic describes how to create the stemcell that Pivotal Application Service for Windows (PASW) needs to create VMs on vSphere.

 **Note:** The instructions in this topic are based on vSphere 6.0 using vSphere Web Client.

### Overview

To create a Windows stemcell for vSphere, you create a base Windows VM from a volume-licensed ISO and subsequently maintain that base template with all Windows recommended security updates, but without the BOSH dependencies.

 **Note:** The stemcell you create in this topic is based on Windows Server 2019.

The VM with security updates serves as the base for all future stemcells, produced from clones of that base VM. This enables you to build new stemcells without having to run Windows Updates from scratch each time. You can also use a “snapshot” feature to maintain an updated Windows image that does not contain the BOSH dependencies.

Pivotal recommends installing any available critical updates, and then rebuilding the stemcell from a clone of the original VM.

### Prerequisites

Before you create a vSphere Windows stemcell, you must have the following:

- A Windows Server 2019 ISO, from [Microsoft Developer Network \(MSDN\)](#) or [Volume Licensing Service Center \(VLSC\)](#). You can use an evaluation copy for testing, but Pivotal does not recommend an evaluation copy for production, as the licensing expires.
-  **Note:** Pivotal recommends maintaining a separate, updated Windows VM based on this ISO to serve as the basis for the installation steps below. This enables you to apply Windows Updates and create new stemcells without having to reinstall all updates from scratch.
- A vSphere/vCenter account granted sufficient permissions to perform all of the following tasks:
  - Create a VM.
  - Configure a VM.
  - Open a VM in VM Remote Console on a local desktop.
  - Export a VM.
- The ability to download/transfer files and software to a vCenter Windows VM.

### Files on Local Machine

As part of completing the procedures in this topic, you download the following files to your local machine:

- The latest release of [stembuild](#).
- [ovftool](#).

### Files on Windows VM

As part of completing the procedures in this topic, you download the following files to your Windows VM:

- [lgpo.exe](#) from the Microsoft Security Toolkit.
- [OpenSSH v7.7.2.0p1](#).
- The [BOSH PS Modules and BOSH Agent](#) for the 2019 stemcell version you want to build.

 **Note:** You must choose a stemcell version to build. Stemcells are versioned as MAJOR.MINOR, such as 2019.2. For more information about 2019 stemcells, see the [Stemcell v2019.x \(Windows Server 2019\) Release Notes](#).

## Step 1: Create Base VM for Stemcell

This section describes how to create, configure, and verify a base Windows VM from a volume-licensed ISO.

### Upload the Windows ISO

Perform the following steps to upload the Windows ISO:

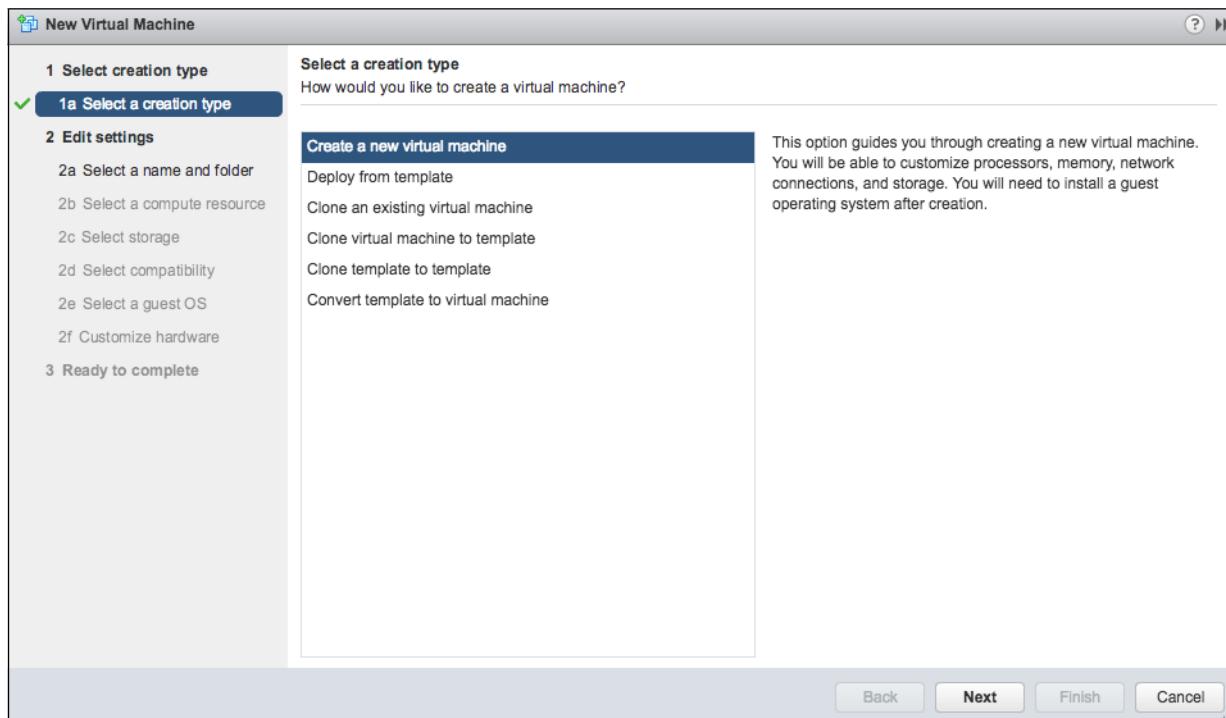
1. Log in to vCenter.
2. Click **Storage** in the vCenter menu.
3. Choose a datastore and click or create the directory where you want the Windows ISO.
4. Click **Upload a file to datastore**, and upload the Windows ISO.

 **Note:** You might need to install the vSphere client web plugin to upload through your browser, or [SCP](#) the file directly to the datastore server. For more information, see the VMware vSphere [documentation](#).

### Create and Customize a New VM

Perform the following steps to create and customize a new VM:

1. In the vSphere client, click the **VMs and Templates** view to display the inventory objects.
2. Right-click an object and select **New Virtual Machine > New Virtual Machine....**
3. On the **Select a creation type** page, select **Create a new virtual machine** and click **Next**.



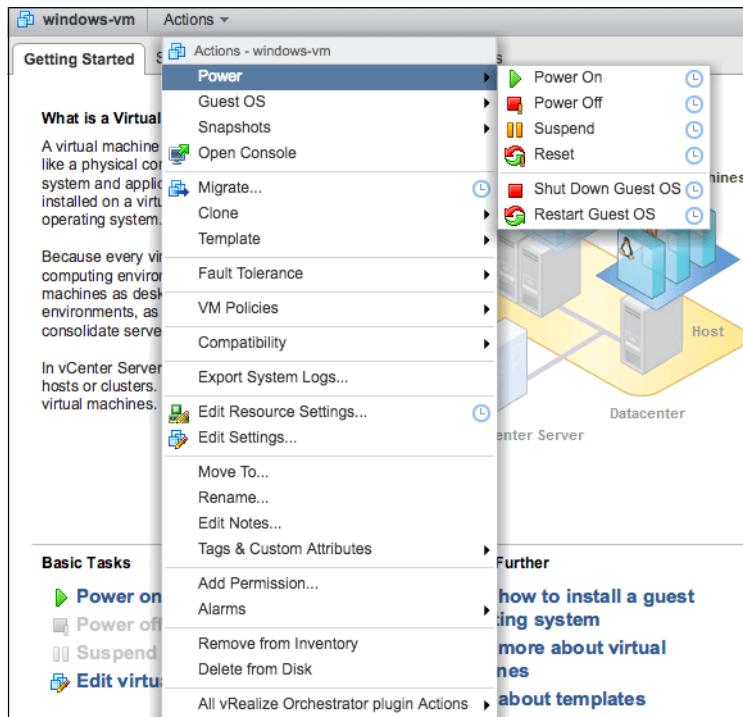
4. On the **Select a name and folder** page, perform the following steps:
  - a. Enter a name for the VM.
  - b. Select a location for the VM.
  - c. Click **Next**.
5. On the **Select a compute resource** page, select a compute resource to run the VM and click **Next**.
6. On the **Select storage** page, perform the following steps:

- a. Select a VM Storage Policy.
  - b. Select the destination datastore for the VM configuration files and virtual disks.
  - c. Click Next.
7. On the **Select compatibility** page, for the **Compatible with** configuration setting, select **ESXi 6.0 and later** and click **Next**.
8. On the **Select a guest OS** page, perform the following steps:
- a. For **Guest OS Family**, select **Windows**.
  - b. For **Guest OS Version**, select **Microsoft Windows Server 2016**.
  - c. Click **Next**.
9. On the **Customize hardware** page, configure the VM hardware and click **Next**. When configuring the VM hardware, select the following settings for **New Hard disk** and **New CD\DVD Drive**:
- a. For **New Hard disk**, specify 30 GB or greater.
  - b. For **New CD\DVD Drive**, perform the following steps:
    - i. Select **Datastore ISO File**.
    - ii. Select the ISO file you uploaded to your datastore and click **OK**.
    - iii. Enable the **Connect At Power On** checkbox.
10. Review the configuration settings on the **Ready to complete** page and click **Finish**.

## Install Windows Server

Perform the following steps to install Windows Server on the base VM:

1. After creating the VM, click **Power On** in the **Actions** tab for your VM.



2. Select **Windows Server Standard**.
3. Select **Custom installation**.
4. Complete the installation process, and enter a password for the Administrator user. BOSH later randomizes this password.

## Verify OS

**warning:** You must complete the following procedure to verify your OS version before continuing.

Ensure you are using the correct the OS version by running the following PowerShell command on the Windows VM:

```
Get-CimInstance Win32_OperatingSystem | Select-Object
Caption, Version, ServicePackMajorVersion, OSArchitecture, CSName, WindowsDirectory
```

The output includes `Version: 10.0.17763`.

## Install VMware Tools

Perform the following steps to install VMware Tools on the base VM:

1. Under the VM **Summary** tab, select **Install VMware Tools**.
  2. Navigate to the `D:` drive and run `setup64.exe`.
-  **Note:** The VMware Tools install window might appear behind the Command Prompt window.
3. Restart the VM as required to finish the install.

## Step 2: Install Windows Updates

This section describes how to install Windows updates on your base Windows VM.

### Install Windows Updates

Install Windows updates on the Windows VM using your preferred procedure.

One way to install Windows updates on the Windows VM is by using the **SConfig** utility. Perform the following steps:

1. On the Windows VM, run the **SConfig** utility.
2. Select option number 6, **Download and Install Updates**.
3. Select **A** for **(A)ll updates**.
4. For the **Select an option**, select **(A)ll updates**.

You might need to restart the Windows VM while installing updates.

### Enable Meltdown Mitigation

 **warning:** You must enable Meltdown mitigation. Not enabling Meltdown mitigation can lead to timeout issues while deploying the PASW tile.

Windows Server 2019 should receive the update containing the Meltdown mitigation automatically when you install Windows updates.

After installing Windows update, ensure that the following registry keys are set to enable Meltdown mitigation:

```
reg add "HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Session Manager\Memory Management"
/v FeatureSettingsOverride /t REG_DWORD /d 0 /f

reg add "HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Session Manager\Memory Management"
/v FeatureSettingsOverrideMask /t REG_DWORD /d 3 /f

reg add "HKLM\SOFTWARE\Microsoft\Windows NT\CurrentVersion\Virtualization"
/v MinVmVersionForCpuBasedMitigations /t REG_SZ /d "1.0" /f

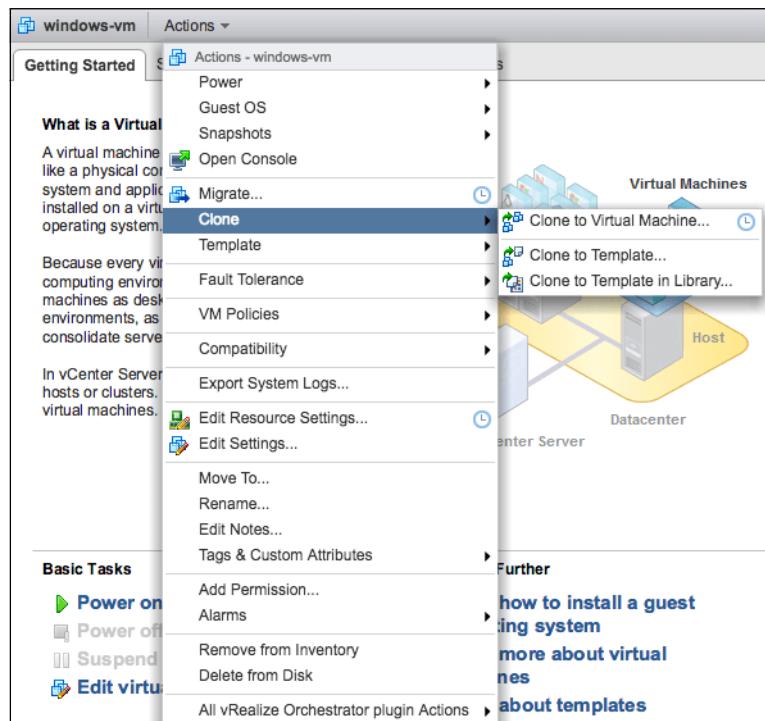
reg add "HKLM\SOFTWARE\Microsoft\Windows\CurrentVersion\QualityCompat"
/f /v cadca5fe-87d3-4b96-b7fb-a231484277cc /t REG_DWORD /d 0
```

## Step 3: Clone the VM

Clone the VM that has the Windows updates installed. Save the original VM so that you can run updates on it in the future.

Perform the following steps:

1. In the vSphere client, right-click the current Windows VM.



2. Select **Clone > Clone to Virtual Machine....**

3. Ensure that you can create the VM that can be used to create a stemcell for the next Patch Tuesday Monthly Updates.

## Step 4: Install Required Software

You might need to specify an explicit execution policy for all of the PowerShell commands in the *Step 4: Install Required Software* section. You specify an execution policy with the `-ExecutionPolicy` flag.

For example:

```
powershell -ExecutionPolicy Bypass -Command "Install-CFFeatures"
```

### Transfer Files to a Windows VM

Some of the procedures described in the sections below require transferring files to a Windows VM. Many different methods exist to transfer files to a Windows VM, such as folder sharing or the PowerShell `Invoke-WebRequest` cmdlet. Use whatever method that you prefer.

As an example, the following PowerShell `Invoke-WebRequest` command uses TLS v1.2 to transfer `filename.zip` from `EXAMPLE-URL` to the current location on the Windows VM:

```
[Net.ServicePointManager]::SecurityProtocol = [Net.SecurityProtocolType]::Tls12
Invoke-WebRequest -Uri "EXAMPLE-URL/filename.zip" -OutFile ".\filename.zip"
```

### Install the BOSH PS Modules

Perform the following steps to install the BOSH PS Modules:

1. Locate the [BOSH PS Modules](#) download for the 2019 stemcell version you want to build, such as 2019.2.
2. Transfer the `bosh-psmodules.zip` file to your Windows VM.
3. Start PowerShell in the Windows VM and run the following command:

```
Unblock-File PATH-TO-BOSH-PSMODULES.ZIP
```

Where `PATH-TO-BOSH-PSMODULES.ZIP` is the full path to the location of `bosh-psmodules.zip` on your Windows VM.

4. Unzip the archive with the following command:

```
Expand-Archive PATH-TO-BOSH-PSMODULES.ZIP C:\Program Files\WindowsPowerShell\Modules
```

## Install the Cloud Foundry Diego Cell Requirements

Perform the following steps to install the Cloud Foundry Diego cell requirements:

1. Start PowerShell in the Windows VM and run the following command:

```
Install-CFFeatures
```

The machine restarts automatically.

2. Apply the recommended ingress and service configuration with the following command:

```
Protect-CFCell
```

## Install the BOSH Agent

Perform the following steps to install the BOSH Agent:

1. Locate the [BOSH Agent](#) download for the 2019 stemcell version you want to build, such as 2019.2.
2. Transfer the `agent.zip` file to your Windows VM.
3. Start PowerShell in the Windows VM and run the following command:

```
Unblock-File PATH-TO-AGENT.ZIP
```

Where `PATH-TO-AGENT.ZIP` is the full path to the location of the `agent.zip` file on your Windows VM.

4. Install the BOSH Agent with the following command:

```
Install-Agent -IaaS vsphere -agentZipPath PATH-TO-AGENT.ZIP
```

## Install OpenSSH

You can use the `bosh ssh` command on BOSH-deployed Windows VMs if you install the OpenSSH dependency on the Windows VM and then enable it during deploy time. This lets an operator enter into a CMD or PowerShell session on the VM as a user with admin privileges.

Perform the following steps to install OpenSSH:

1. Transfer the [OpenSSH-Win64.zip](#) file to the Windows VM and place it in `C:\provision`.
2. Start PowerShell in the Windows VM and run the following command:

```
Unblock-File 'C:\provision\OpenSSH-Win64.zip'
```

3. Install OpenSSH with the following command:

```
Install-SSHD -SSHZipFile 'C:\provision\OpenSSH-Win64.zip'
```

- When configuring the PAS for Windows tile, you must select the **BETA: Enable BOSH-native SSH support on all VMs checkbox**. For more information, see [Installing and Configuring PAS for Windows](#).

## Optimize and Compress the Disk

**Note:** Windows Server stemcells can be large, and can exceed the 10GB upload limit imposed by default by the BOSH Director.

Perform the following steps to reduce the stemcell size:

- Restart the VM.
- Start PowerShell in the Windows VM and run the following command to use `dism` to clear unnecessary files:

```
Optimize-Disk
```

- Run the following command to defragment and zero out the disk:

```
Compress-Disk
```

## Step 5: Sysprep the System

This step “syspreps” the system, which ensures that each BOSH VM has a unique identity and applies the appropriate startup configuration at boot time.

The included policies help ensure the uptime and secure operations of the stemcell’s VMs, especially when deployed on PCF.

**Note:** This step disables services that could cause restarts, such as Windows Automatic Updates. OS restarts are not supported on BOSH-deployed Windows VMs, and the BOSH Director resurrects the VM by destroying and repaving it.

Perform the following steps:

- Transfer the `LGPO.ZIP` file to the Windows VM.
- Start PowerShell in the Windows VM and run the following command:

```
Expand-Archive PATH-TO-LGPO.ZIP C:\Windows
```

- Run the following command to sysrep the system:

```
Invoke-Sysprep -IaaS vsphere
[-NewPassword PASSWORD]
[-Owner OWNER] [-Organization ORGANIZATION]
```

**Note:** All of the flags of `Invoke-Sysprep` except for `-IaaS` are optional.

Where:

- `PASSWORD` is an optional flag that enables you to set a password of your choice. Do not use any special character in the password other than `!`. For example, `Example12!` is permitted but `Example#12` is not. This is a known issue.
- `OWNER` and `ORGANIZATION` are optional flags. Set them if your organization requires it.

The sysrep command powers off the VM.

**warning:** Do not turn the VM back on before completing the procedure in [Step 6: Export the VMDK File](#).

## Step 6: Export the VMDK File

Perform the following steps to export the .VMDK file associated with the VM you powered off:

1. In vCenter, right-click the VM and select **Template > Export to OVF Template**.
2. Download the OVA to your local machine. You do not need to include files in the floppy or CD Drive.

 **Note:** You can also download the standalone vSphere client and select **File > Export > Export OVF Template**.

3. Rename the downloaded OVA file to have a `.tar` extension.
4. Expand the TAR archive and locate the VMDK file.

## Step 7: Convert the VMDK File to a BOSH Stemcell

 **Note:** This final step typically takes about ten to twenty minutes to complete.

Perform the following steps to convert the VMDK file to a BOSH stemcell:

1. Download the latest release of the [stembuild](#) utility to your local machine and place the executable in your command-line path.
2. Download [ovftool](#) to your local machine and place the executable in your command-line path.

 **Note:** On the Windows desktop, `ovftool` is installed by default in `C:\Program Files\VMware\VMware OVF Tool`.

`stembuild` invokes `ovftool` to convert the disk image to the appropriate stemcell format and apply the proper configuration.

3. Build the stemcell with the following command:

```
stembuild package -vmdk PATH-TO-VDMK -stemcell-version STEMCELL-VERSION -os 2019
```

Where:

- `PATH-TO-VMDK` is the path to the VMDK file.
- `STEMCELL-VERSION` is the 2019 stemcell version you want to build. For example, if you downloaded the [BOSH PS Modules and BOSH Agent](#) for the 2019.2 release, then specify `2019.2`.

`stembuild` creates the stemcell in the directory where you execute it. The file has a `.tgz` extension and a name similar to `bosh-stemcell-2019.2-vsphere-esxi-windows2019-go_agent.tgz`.

The stemcell is ready for use in conjunction with your BOSH deployment.

## Step 8: Apply Monthly Patch Tuesday Updates

On Patch Tuesday, run Windows Updates on the base image, and then repeat [Step 3: Clone the VM](#) through [Step 7: Convert the VMDK File to a BOSH Stemcell](#).

## Troubleshooting

### Garden Windows Logs Suggest Windows Features Not Installed

#### Symptom

You see the following error in your `garden-windows` job while deploying Windows Server 2019:

Missing required Windows Features:  
Web-Webserver, Web-WebSockets, AS-Web-Support,  
AS-NET-Framework, Web-WHC, Web-ASP.  
Please use the most recent stemcell.

## Explanation

`Install-CFFeatures` might not have run successfully.

## Solution

Run the following commands in PowerShell on your Windows VM to verify whether `Install-CFFeatures` ran successfully:

```
Get-WindowsFeature "Containers" | Where InstallState -Eq "Installed"
Get-WindowsFeature "Windows-Defender-Features" | Where InstallState -Eq "Removed"
```

## Creating a Windows Stemcell for vSphere Using stembuild (Beta)

This topic describes how to create a BOSH stemcell for Pivotal Application Service (PAS) for Windows on VMware vSphere using `stembuild`.

### Overview

Before you can deploy PAS for Windows on vSphere, you must create a BOSH stemcell for Windows, a versioned operating system image. The BOSH stemcell that you create in this topic is based on Windows Server 2019.

To create a BOSH stemcell for Windows on vSphere, do the following:

1. [Create a Base VM for the BOSH Stemcell](#)
2. [Install Windows Updates](#)
3. [Clone the Base VM](#)
4. [Construct the BOSH Stemcell](#)
5. [Package the BOSH Stemcell](#)
6. [Upload the BOSH Stemcell to Ops Manager](#)

If you already have a BOSH stemcell for Windows on vSphere, see [Monthly Stemcell Upgrades](#).

### Prerequisites

- A vSphere environment
- A Windows Server 2019 ISO
- The `stembuild` command line interface (CLI) from a release in [Stemcells for PCF \(Windows\)](#) on Pivotal Network that corresponds to the operating system of your local host and the stemcell version that you want to build
- Microsoft [Local Group Policy Object Utility \(LGPO\)](#) downloaded to the same folder as your `stembuild` CLI
- The minimum vCenter user permissions required to use stembuild for vSphere stemcells, specifically:
  - `VirtualMachine.GuestOperations.Modify`
  - `VirtualMachine.GuestOperations.Execute`
  - `VirtualMachine.GuestOperations.Query`
  - `VirtualMachine.Config.AddRemoveDevice`
  - `VirtualMachine.Interact.SetCDMedia`
  - `VApp.Export`
  - `System.Anonymous`\*
  - `System.Read`\*
  - `System.View`\*

Permissions marked with an `*` are generated upon creating a new user in vCenter and cannot be set within the vCenter UI.

### Step 1: Create a Base VM for the BOSH Stemcell

This section describes how to create, configure, and verify a base VM for Windows from a volume-licensed ISO.

#### Upload the Windows Server 2019 ISO

To upload the Windows Server 2019 ISO to vSphere, do the following:

1. Log in to the vSphere Web Client.

 **Note:** The instructions in this topic are based on vSphere 6.0.

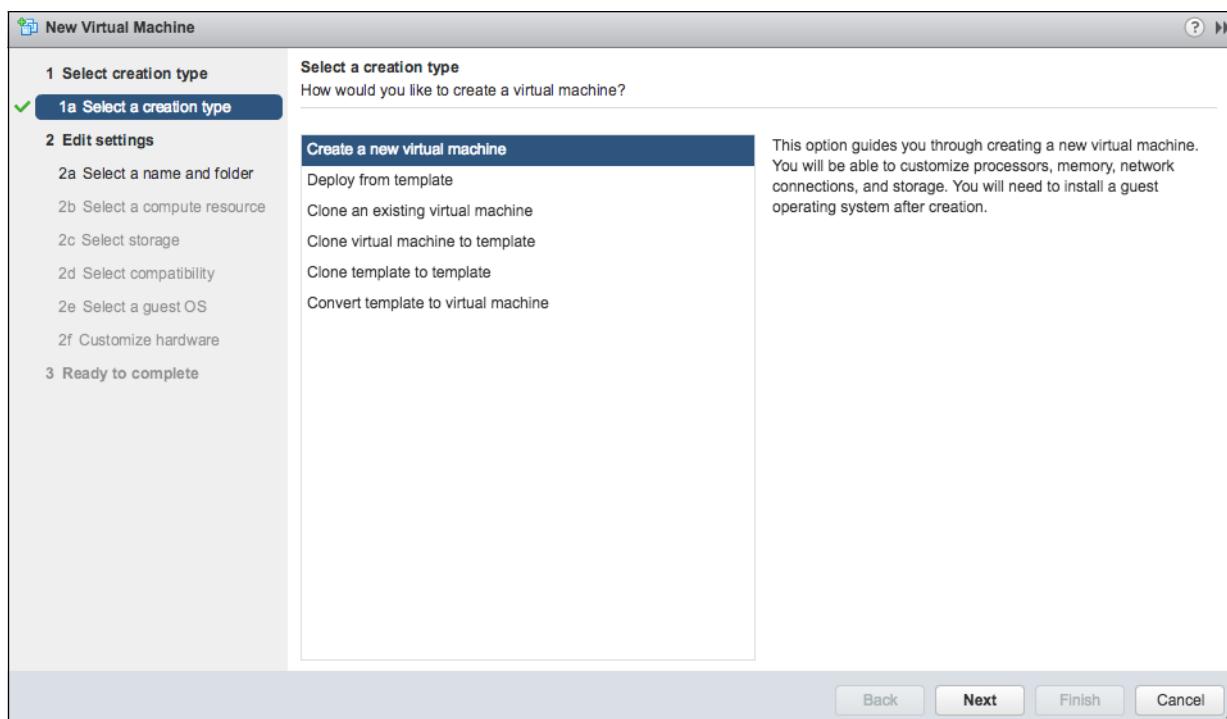
2. Click **Storage** and select a datastore.
3. Select or create a folder where you want to upload the Windows Server 2019 ISO.
4. Click **Upload a File** and select the Windows Server 2019 ISO.

You can use the `scp` utility instead of the vSphere Web Client to copy the file directly to the datastore server.

## Create and Customize a Base VM

To create and customize a base VM, do the following:

1. In the vSphere Web Client, click the **VMs and Templates** view to display the inventory objects.
2. Right-click an object and select **New Virtual Machine > New Virtual Machine**.
3. On the **Select a creation type** page, select **Create a new virtual machine** and click **Next**.



4. On the **Select a name and folder** page, do the following:
  - a. Enter a name for the VM.
  - b. Select a location for the VM.
  - c. Click **Next**.
5. On the **Select a compute resource** page, select a compute resource to run the VM and click **Next**.
6. On the **Select storage** page, do the following:
  - a. Select a **VM Storage Policy**.
  - b. Select the destination datastore for the VM configuration files and virtual disks.
  - c. Click **Next**.
7. On the **Select compatibility** page, for the **Compatible with** configuration setting, select **ESXi 6.0 and later** and click **Next**.
8. On the **Select a guest OS** page, do the following step:
  - a. For **Guest OS Family**, select **Windows**.
  - b. For **Guest OS Version**, select **Microsoft Windows Server 2019**. If **Microsoft Windows Server 2019** is not available, select **Microsoft Windows Server 2016**.
  - c. Click **Next**.

9. On the **Customize hardware** page, configure the VM hardware using the information below and click **Next**.

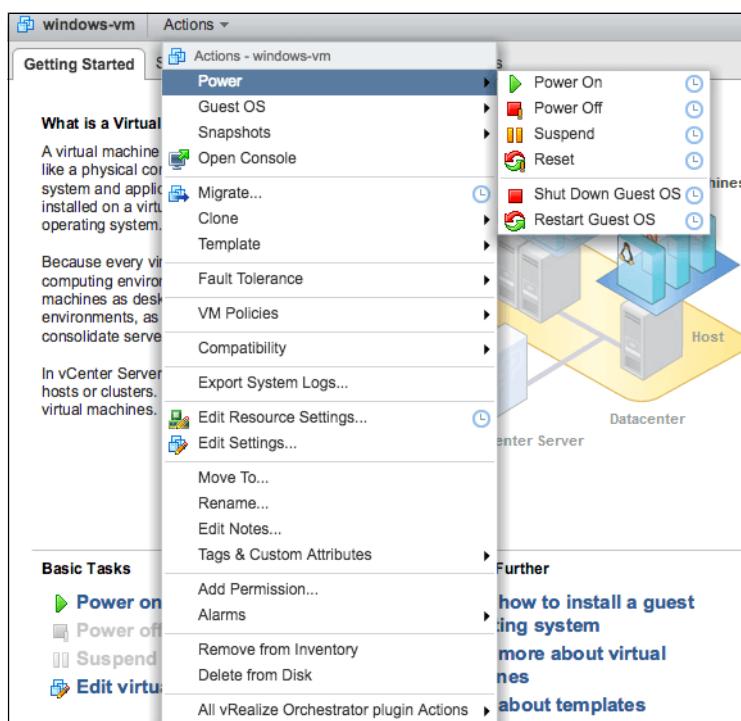
- For **New Hard disk**, specify 30 GB or greater.
- For **New CD\DVD Drive**, do the following:
  - Select **Datastore ISO File**.
  - Select the Windows Server 2019 ISO file you uploaded to your datastore and click **OK**.
  - Enable the **Connect At Power On** checkbox.

10. Review the configuration settings on the **Ready to complete** page and click **Finish**.

## Install Windows Server

To install Windows Server on the base VM, do the following:

1. After creating the VM, click **Power > Power On** in the **Actions** tab for your VM.



2. Select **Windows Server Standard**.

3. Select **Custom installation**.

4. Complete the installation process and enter a password for the Administrator user.

## Verify OS

To verify that you are using the correct OS version, run the following PowerShell command on the base VM:

```
[System.Environment]::OSVersion.Version
```

The output should display the following:

```
[System.Environment]::OSVersion.Version
```

| Major | Minor | Build | Revision |
|-------|-------|-------|----------|
|-------|-------|-------|----------|

|    |   |       |   |
|----|---|-------|---|
| 10 | 0 | 17763 | 0 |
|----|---|-------|---|

## Install VMware Tools

To install VMware Tools on the base VM, do the following:

1. In the vSphere Web Client, right-click the base VM and select **Guest OS > Install VMware Tools**.
2. Navigate to the **D:** drive and run **setup64.exe**.
3. Restart the VM to complete the installation.

## Step 2: Install Windows Updates

Install Windows updates on the base VM using your preferred procedure. For example, you can install Windows updates by following the steps below. This procedure requires internet access.

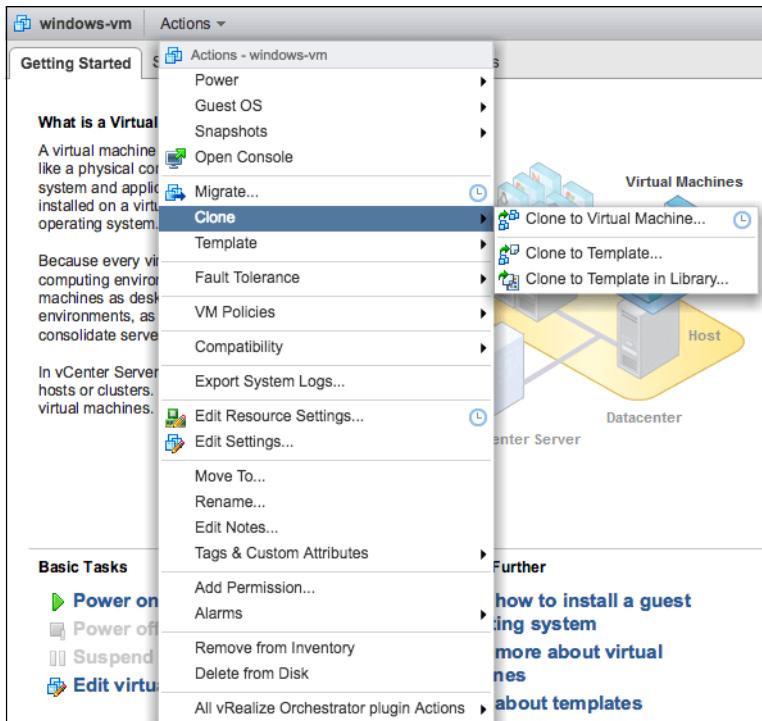
1. On the base VM, run the **SConfig** utility.
2. Select **Download and Install Updates**.
3. Enter **A** to search for all updates.
4. For **Select an option**, enter **A** to install all updates.

You may need to restart the base VM while installing the updates.

## Step 3: Clone the Base VM

To clone the base VM, do the following in the vSphere Web Client:

1. Power down the base VM.
2. Right-click the base VM.
3. Select **Clone > Clone to Virtual Machine**. This clone is your target VM.



4. Save the base VM. You will run Windows updates on this VM for future stemcells.

## Step 4: Construct the BOSH Stemcell

Before constructing the BOSH stemcell, collect the following information:

- Target VM IP address
- Target VM username
- Target VM password
- vCenter inventory path to the target VM in the `/MY-DATA-CENTER/vm/MY-FOLDER/MY-VM` format where:
  - `MY-DATA-CENTER` is the name of the data center.
  - `vm` is a static string.
  - `MY-FOLDER` is the name of the folder that contains the VM.
  - `MY-VM` is the name of the target VM.
- vCenter username
- vCenter password
- vCenter URL

 **Note:** The target VM must be routable from your local host. Before running the `construct` command, ensure you are logged out of the target VM.

To construct the BOSH stemcell, run the following command from your local host:

```
./STEMBUILD-BINARY construct -vm-ip TARGET-VM-IP -vm-username TARGET-USERNAME -vm-password TARGET-VM-PASSWORD -vcenter-url VCENTER-URL -vcenter-username VCENTER-USI
```

Where:

- `STEMBUILD-BINARY` is the `stembuild` file for the version of your local host operating system and the version of the stemcell that you want to build. For example, `stembuild-windows-2019-2`.
- `TARGET-VM-IP` is the IP address of your target VM.
- `TARGET-USERNAME` is the username of an account with Administrator privileges.
- `TARGET-VM-PASSWORD` is the password for the Administrator account. The password must be enclosed in single quotes.
- `VCENTER-URL` is the URL of your vCenter.
- `VCENTER-USERNAME` is the username of your account in vCenter.
- `VCENTER-PASSWORD` is your password. The password must be enclosed in single quotes.
- `INVENTORY-PATH` is the vCenter inventory path to the target VM.

 **Note:** This operation may take up to an hour to complete and results in a powered-off target Windows VM in your vSphere environment. During `construct` execution, the WinRM connection terminates. This behavior is expected, and the `construct` command is still being executed. Do not attempt to re-run the `construct` command.

If you want to view the status of `construct`, you can log in to the target VM and do the following:

1. Start PowerShell.
2. Run the following command:

```
Get-Content -Path "C:\provision\log.log" -Wait
```

## Step 5: Package the BOSH Stemcell

Before packaging the BOSH stemcell, collect the following information:

- vCenter inventory path to the target VM in the `/MY-DATA-CENTER/vm/MY-FOLDER/MY-VM` format where:
  - `MY-DATA-CENTER` is the name of the data center.
  - `vm` is a static string.
  - `MY-FOLDER` is the name of the folder that contains the VM.
  - `MY-VM` is the name of the target VM.

- vCenter username
- vCenter password
- vCenter URL

To package the BOSH stemcell, run the following command from your local host:

```
./STEMBUILD-BINARY package -vcenter-url VCENTER-URL -vcenter-username VCENTER-USERNAME -vcenter-password VCENTER-PASSWORD -vm-inventory-path INVENTORY-PATH
```

Where:

- `STEMBUILD-BINARY` is the `stembuild` file for the version of your local host operating system and the version of the stemcell that you want to build. For example, `stembuild-windows-2019-2`.
- `VCENTER-URL` is the URL of your vCenter.
- `VCENTER-USERNAME` is the username of your account in vCenter.
- `VCENTER-PASSWORD` is your password. The password must be enclosed in single quotes.
- `INVENTORY-PATH` is the vCenter inventory path to the target VM.

 **Note:** This command creates a stemcell on your local host in the folder where you ran the command and may take up to 30 minutes to complete.

## Step 6: Upload the BOSH Stemcell to Ops Manager

To upload the BOSH stemcell to Ops Manager, do the following:

1. In Ops Manager, navigate to [Stemcell Library](#).
2. Upload your BOSH stemcell.
3. Deploy the PAS for Windows tile.

## Monthly Stemcell Upgrades

After Microsoft releases operating system updates, you should upgrade your BOSH stemcell. Microsoft typically releases Windows updates on the second Tuesday of each month.

To upgrade your BOSH stemcell, do the following:

1. [Install Windows Updates](#) on the base VM.
2. [Clone the Base VM](#).
3. [Construct the BOSH Stemcell](#).
4. [Package the BOSH Stemcell](#).
5. Replace the existing stemcell in the Ops Manager stemcell library with this new stemcell.
6. Deploy the PAS for Windows tile.

## Known Issues

### Authentication Error with Special Characters in stembuild Commands

#### Symptom

```
./out/stembuild: ServerFaultCode: Cannot complete login due to an incorrect user name or password.
vcenter_client - unable to validate url: vcenter.example.com
```

## Explanation

stembuild uses govc libraries. These libraries cannot parse the special characters /, #, and :. This results in errors when authenticating with vCenter.

## Workaround

If your vCenter username or password contains /, #, or :, set the following environment variables:

```
export GOVC_USERNAME=VCENTER-USERNAME
export GOVC_PASSWORD=VCENTER-PASSWORD
```

Where:

- VCENTER-USERNAME is your vCenter account username. For example, `johndoe`.
- VCENTER-PASSWORD is your vCenter account password. For example, `pass#word`.

If you are using other special characters, add single quotes around the input parameters. For example:

```
./STEMBUILD-BINARY package -vcenter-url VCENTER-URL -vcenter-username 'admin@' -vcenter-password VCENTER-PASSWORD -vm-inventory-path INVENTORY-PATH
```

## Installing and Configuring PASW

This topic describes how to install and configure the Pivotal Application Service for Windows (PASW) tile. The PASW tile installs Windows cells in your Pivotal Cloud Foundry (PCF) deployment.

### Step 1: Confirm Shared PAS Tile Settings

Windows cells installed by the **Pivotal Application Service for Windows** tile are affected by two settings in the **Pivotal Application Service** tile. Configure these **Pivotal Application Service** tile settings as desired:

- **System Logging:** In the **System Logging** pane, if you configure an external syslog aggregator, logs are drained from Windows cells as well as non-Windows cells. See [Logging and Metrics](#) for additional information.
- **Disable SSL certificate verification for this environment:** In the **Networking** pane, if you select the **Disable SSL certificate verification for this environment** checkbox, SSL certificate verification is disabled for Windows cells. See [Procedure: Terminate SSL/TLS at HAProxy](#) for additional information.

**Note:** Only disable SSL certificate verification for development and testing environments. Do not select **Disable SSL certificate verification for this environment** for production environments.

### Step 2: Install the Tile

1. Download the **Pivotal Application Service for Windows** product file from [Pivotal Network](#).
2. Within [Pivotal Application Service for Windows](#), download the **Windows FS Injector** tool for your workstation OS.

The Injector tool, `winfs-injector`, is an executable binary that adds the Windows Server container base image into the product file. This step requires internet access and can take up to 20 minutes.

**Note:** You need the `git` and `tar` (BSD) executables on your `%PATH%` to run the `winfs-injector` bin. For example, to use `winfs-injector.exe`, `tar.exe` must be copied to a directory in your `%PATH%`.

3. To add the Windows Server container base image to the product file, run the following command:

```
winfs-injector --input-tile PASW-DOWNLOAD-PATH --output-tile PASW-IMPORTABLE-PATH
```

Where:

- `PASW-DOWNLOAD-PATH` is the path and filename to the downloaded PASW product file.
- `PASW-IMPORTABLE-PATH` is the desired output path for the importable product file.

For example:

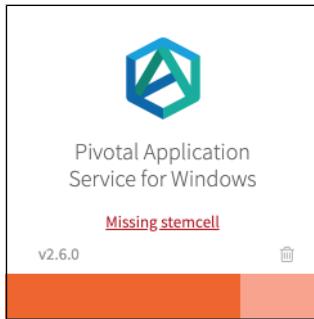
```
C:\Users\admin> winfs-injector --input-tile ^
c:\temp\pas-windows-2.6.0-build.1.pivotal ^
--output-tile c:\temp\pas-windows-2.6.0-build.1-INJECTED.pivotal
```

For troubleshooting the `winfs-injector`, see [Missing Local Certificates for Windows File System Injector](#).

4. Navigate to the Ops Manager Installation Dashboard and click **Import a Product**.
5. To add **PASW** to the **Import a Product** product list, select the importable `PASW-IMPORTABLE-PATH` file on your workstation.
6. To add **PASW** to your staging area, click **+** under the **PASW** product listing.

### Step 3: Configure the Tile

1. Click the newly added **PASW** tile.



2. Click **Assign AZs and Networks** or **Assign Networks**. The name of the pane varies depending on your IaaS.
3. Assign your AZs and networks and click **Save**.
4. Click **VM Options**.

The screenshot shows a configuration dialog for "Manage Administrator Password". It includes two radio button options: "Use random password" (selected) and "Set the password". There are also two checkboxes: "BETA: Enable BOSH-native SSH support on all VMs" (unchecked) and "Enable Remote Desktop Protocol" (unchecked). Below these are sections for "Key Management Service" with "Enable" and "Disable" radio buttons (disabled), and a "Save" button.

5. Specify your selection for **Manage Administrator Password**.
  - o **Use random password** randomizes the admin password. With this selection, the admin password is not retrievable by an operator. This is the default selection.
  - o **Set the password** sets the same admin password for every Windows cell. As a result, this password can be used to access any Windows cell, including, for example, Remote Desktop Protocol (RDP) sessions.

6. (Optional) Select the **BETA: Enable BOSH-native SSH support on all VMs** checkbox to start the Microsoft beta port of the OpenSSH daemon on port 22 on all VMs. Users can SSH onto Windows VMs with the `bosh ssh` command, and enter a CMD terminal as an admin user. They can then run `powershell.exe` to start a PowerShell session.

**Note:** This feature is beta and not considered production-ready.

7. (Optional) If you want all VMs to support connection through Remote Desktop Protocol (RDP), click **Enable Remote Desktop Protocol**.
8. (Optional) If you want to configure a Key Management Service (KMS) that your volume-licensed Windows cell can register with, perform the following steps:
  - a. Click **Enable**
  - b. For the **Host** field, enter the KMS hostname.
  - c. For the **Port** field, enter the port number. The default port number is `1688`.

Key Management Service\*

Enable

Host \*

Port \*

1688

9. Click **Save**.

10. (Optional) To deploy your PASW application workloads to an isolation segment, click **Application Containers** and perform the steps in the [Assign a Tile to an Isolation Segment](#) section below.

11. (Optional) To configure Windows cells to send Windows Event logs to an external syslog server, click **System Logging** and perform the steps in the [Send Cell Logs to a Syslog Server](#) section.

12. Click **Smoke Tests**.

Specify a Cloud Foundry organization and space where smoke tests can run if in the future you delete your Application Service deployment domains.

Choose where to deploy applications when running the smoke tests \*

Temporary space within the system organization (This is deleted after smoke tests finish.)  
 Specified org and space (The org and space must have a domain available for routing.)

**Save**

13. Choose an org and space where your smoke tests can run.

- Choosing **Temporary space within the system organization** deletes the org after smoke tests finish.
- Choosing **Specified org and space** requires a domain available for routing and that domain must be accessible to the organization.
  - In the **Organization** field, enter the Application Service org to use when running tests.
  - In the **Space** field, enter the Application Service space to use when running tests.
  - In the **Domain** field, enter the domain that the org has access to for running smoke tests.

14. (Optional) To configure memory and disk overcommit for your Windows Diego cells, click **Advanced Features** and perform the following steps:

- a. Enter the total desired amount of Diego cell memory in the **Cell Memory Capacity (MB)** field. For the current cell memory capacity settings, see the Windows Diego cell row on the **Resource Config** pane.
- b. Enter the total desired amount of Diego cell disk capacity in the **Cell Disk Capacity (MB)** field. For the current cell disk capacity settings, see the Windows Diego cell row on the **Resource Config** pane.



**Note:** Due to the risk of app failure and the deployment-specific nature of disk and memory use, Pivotal has no recommendation about how much, if any, memory or disk space to overcommit.

|                                      |
|--------------------------------------|
| Cell Memory Capacity (MB) ( min: 1 ) |
| <input type="text"/>                 |
| Cell Disk Capacity (MB) ( min: 1 )   |
| <input type="text"/>                 |

c. Click **Save**.

15. Click **Errands**.

- a. Pivotal recommends that you set the **Install HWC Buildpack Errand** to **On**. This ensures that you receive the most up-to-date HWC Buildpack.
- b. Enabling the **Smoke Test Errand** ensures that a smoke test is run against your Application Service installation.

**Errands**

Errands are scripts that run at designated points during an installation.

**Post-Deploy Errands**

Install HWC Buildpack Errand      Install the HWC buildpack to your Elastic Runtime for Windows installation

Smoke Test Errand      Runs Smoke Tests against your Application Service installation

There are no pre-delete errands for this product.

**Save**

16. Click **Save**.

## Step 4: Configure Tile Resources

To configure your tile resources, perform the following steps:

1. Navigate to the **Resource Config** pane of the PASW tile.
2. Use the dropdown menus to configure **Windows Diego Cell**. For more information, see [Disk Size of Windows Diego Cells](#).
3. Click **Save**.

**Note:** Provision your **Master Compilation Job** with at least 100 GB of disk space.

## Disk Size of Windows Diego Cells

Windows stemcells in the v2019.x line support ephemeral disks.

The recommended disk size for your Windows Diego cells is as follows:

| IaaS | Recommended Disk size of Windows Diego cell |
|------|---------------------------------------------|
|------|---------------------------------------------|

|         |        |
|---------|--------|
| AWS     | 100 GB |
| Azure   | 150 GB |
| GCP     | 150 GB |
| vSphere | 100 GB |

**Note:** If you use vSphere, you must create your own stemcell. The default root disk size of Windows stemcells v2019.x line is 30 GB. Pivotal recommends setting the root disk size of your Windows stemcell for vSphere to 30 GB. For more information, see [Creating a Windows Stemcell for vSphere Manually](#) or [Creating a Windows Stemcell for vSphere Using stembuild \(Beta\)](#).

## Step 5: Upload the Stemcell

1. Go to [Stemcell Library](#).
2. Retrieve the stemcell that you downloaded or created in [Downloading or Creating a Windows Stemcell](#).
3. Follow the steps in [Importing and Managing Stemcells](#) to upload the Windows stemcell to **Pivotal Application Service for Windows**.

## Step 6: Deploy the Tile

1. Return to the Ops Manager Installation Dashboard.
2. Click **Review Pending Changes**.
3. Select the PASW tile and review the changes. For more information, see [Reviewing Pending Product Changes](#).
4. Click **Apply Changes** to install the **PASW** tile.

## Step 7: (Optional) Create More Tiles

To run Windows cells in multiple isolation segments, you must create and configure additional PASW tiles. For more information, see [Windows Cells in Isolation Segments](#).

## Windows Cells in Isolation Segments

This topic explains how to run Windows cells in isolation segments, which are compartmentalized resource pools for Diego cells.

Cells in one isolation segment share routing, computing, and logging resources with other cells in the same segment, and do not use resources from other isolation segments.

### Prerequisites

To run a Windows cell in a particular isolation segment, the isolation segment must first already exist in the Cloud Controller Database (CCDB) before the Windows cell can be assigned to it.

To create an isolation segment in the CCDB, see: [Create an Isolation Segment](#) section of the *Managing Isolation Segments* topic.

### Overview

To run Windows cells in multiple isolation segments, you need to create and install multiple PAS for Windows tiles and configure each to run in a different isolation segment.

See [Replicate a Tile](#) for how to create multiple copies of the PAS for Windows tile.

See [Assign a Tile to an Isolation Segment](#) for how to associate a PAS for Windows tile with an isolation segment, so that its cells run in that segment.

### Replicate a Tile

To make multiple copies of the **PAS for Windows** tile that you can assign to different isolation segments, use the Replicator tool as follows:

1. Download the Replicator tool from the **Pivotal Cloud Foundry Runtime for Windows** section of [Pivotal Network](#).
2. Navigate to the directory where you downloaded the Replicator tool.
3. To replicate the tile, run the following command:

```
replicator-windows -name "TILE-NAME" -path PATH-TO-ORIGINAL -output PATH-TO-COPY
```

Where:

- **TILE-NAME** is a unique name for the new PAS for Windows tile. The name must be ten or fewer characters in length and contain only alphanumeric characters, dashes, underscores, and spaces.
- **PATH-TO-ORIGINAL** is the absolute path to the original **.pivotal** PAS for Windows tile you downloaded from Pivotal Network.
- **PATH-TO-COPY** is the absolute path for the copy of the **.pivotal** PAS for Windows tile that the Replicator tool produces.

For example:

```
C:\Users\admin> replicator-windows -name "Shiny Tile" ^
-path c:\downloads\TileOriginal.pivotal ^
-output c:\temp\TileCopy.pivotal
```

4. Install and configure the Windows isolation segment, using the new **.pivotal** file and following the procedures in this topic, starting with the **Import a Product** step of [Step 2: Install the Tile](#).

### Assign a Tile to an Isolation Segment

To assign a PAS for Windows tile to an isolation segment, perform the following steps:

1. Open the **Application Containers** pane.

Segment Name

Save

2. Under **Segment Name**, enter the name for the isolation segment to associate the tile with. If you are creating a new isolation segment, ensure that this name is unique across your PCF deployment.
3. Click **Save**.

## Migrating Apps to PAS for Windows

This topic describes the process of migrating apps running on Pivotal Application Service (PAS) for Windows 2012R2 cells to run on PAS for Windows cells, which run Windows Server 2019.

 **Note:** PAS for Windows 2012R2 has reached its End of Availability (EoA) and is no longer generally available.

Pivotal recommends you use the blue-green deployment method for high availability. For more information about blue-green deployments, see [Using Blue-Green Deployment to Reduce Downtime and Risk](#).

### Step 1: Install and Deploy PAS for Windows Tile

To install and deploy the PAS for Windows tile, follow steps 1 and 2 of [Installing and Configuring PAS for Windows](#).

### Step 2: Push App to PAS for Windows Cells

Perform the following steps to redeploy a running app with zero downtime using the blue-green method:

1. To log in to the Cloud Foundry Command Line Interface (cf CLI), run the following command:

```
cf login
```

2. Choose your org and space.

3. Navigate to the location of your app.

4. To find the name of the existing PAS for Windows 2012R2 app you are migrating to PAS for Windows, run the following command:

```
cf apps
```

5. Create a name for the replacement PAS for Windows app. Pivotal recommends you append `-green` to your existing app name, using the following format:

```
APP-NAME-green
```

Where `APP-NAME` is the existing PAS for Windows 2012R2 app's name.

6. To push your app to a PAS for Windows Diego cell, using your newly created name, run the following command:

```
cf push APP-NAME-GREEN -s windows -b BUILDPACK -n HOSTNAME --no-start --no-route
```

Where:

- `APP-NAME-GREEN` is the newly created “green” name for your app.
- `BUILDPACK` is your custom buildpack. Specify the buildpack either by name or GitHub URL with an optional branch or tag.
- `HOSTNAME` is the name of your app’s subdomain. For example, if `example.com` is your domain and you want the URL to your app to be `http://my-app.example.com`, then specify `my-app` as the `HOSTNAME`.

For example:

```
C:\Users\admin> cf push ExampleApp-green -s windows ^
-b https://github.com/cloudfoundry/binary-buildpack.git ^
-n my-app --no-start --no-route
```

 **Note:** The `--no-start` and `--no-route` parameters included in this `cf push` command are required for this procedure: `--no-start` is used to create the instance VMs and not start the app, and `--no-route` is used to prevent the push command from automatically mapping a route to the app. For additional information about `cf push`, see [cf push](#).

7. To configure the router so all incoming requests go to both `APP-NAME` and `APP-NAME-green` run the following command:

```
cf map-route APP-NAME-green DOMAIN -n HOSTNAME
```

Where:

- `APP-NAME-green` is the name of the new “green” version of your app.
- `DOMAIN` is your domain name, for example, `example.com`.
- `HOSTNAME` is the name of your app’s subdomain. For example, if `example.com` is your domain and you want the URL to your app to be `http://my-app.example.com`, then specify `my-app` as the `HOSTNAME`.

 **Note:** When using the `cf map-route` command, you must specify domain name after specifying app name.

For additional information about `cf map-route`, see [cf map-route](#).

8. To start the green app, run the following command:

```
cf start APP-NAME-green
```

Where `APP-NAME-green` is the name of the new “green” version of your app.

9. To confirm that both your `APP-NAME` and `APP-NAME-green` apps are running, run the following command:

```
cf apps
```

If you experience a problem, see [Troubleshooting Application Deployment and Health](#).

10. To unmap the original app’s route, run the following command:

```
cf unmmap-route APP-NAME DOMAIN -n HOSTNAME
```

Where:

- `APP-NAME` is the name of the existing version of your app you want to replace with `APP-NAME-green`.
- `DOMAIN` is your domain name, for example, `example.com`.
- `HOSTNAME` is the name of your app’s subdomain. For example, if `example.com` is your domain and the existing app is currently accessed using the URL `http://my-app.example.com`, then specify `my-app` as the `HOSTNAME`.

For additional information about `cf unmmap-route`, see [cf unmmap-route](#).

## Step 3: Delete App from Windows 2012R2 Server Cells

1. To delete the original app, run the following command:

```
cf delete APP-NAME
```

Where `APP-NAME` is the name of the app that you have replaced with `APP-NAME-green`.

 **Note:** To also delete any mapped routes, run the command with the `-r` flag.

For additional information about `cf delete`, see [cf delete](#).

## Step 4: (Optional) Uninstall Old Tile

Once you have migrated all of your apps and you are no longer using the Windows 2012R2 tile, a PCF operator can perform the following steps:

1. From the Installation Dashboard, click the trash icon on the tile to remove that product. In the **Delete Product** dialog box that appears, click **Confirm**.
2. Click **Review Pending Changes**.
3. Select the PAS for Windows tile and review the changes. For more information, see [Reviewing Pending Product Changes](#).
4. Click **Apply Changes**.

 **Note:** After you delete a product, the product tile is removed from the installation and the Installation Dashboard. However, the product appears in the Available Products view.

## Upgrading PAS for Windows and Windows Cells

This topic describes how to upgrade the Pivotal Application Service (PAS) for Windows tile and update the Windows stemcell.

For how developers can migrate apps to PAS for Windows from either PAS for Windows 2012R2 or PCF Runtime for Windows, see [Migrating Apps to PAS for Windows](#).

### Rotating Credentials

The PAS tile handles all of the credentials for PAS for Windows.

To rotate your credentials in PAS for Windows, re-deploy PAS and PAS for Windows following the steps below:

1. Navigate to the Ops Manager Installation Dashboard.
2. Click **Review Pending Changes**.
3. Select the PAS and PAS for Windows tiles and review the changes. For more information, see [Reviewing Pending Product Changes](#).
4. Click **Apply Changes**.

### Upgrade PAS for Windows

To upgrade PAS for Windows:

1. Follow the instructions [Upgrading PCF Products](#) with the latest PAS for Windows product download from [Pivotal Network](#).
2. If necessary, configure the product. For more information about configuring PAS for Windows, see the [Installing and Configuring PAS for Windows](#) topic.
3. Click **Review Pending Changes**.
4. Select the PAS for Windows tile and review the changes. For more information, see [Reviewing Pending Product Changes](#).
5. Click **Apply Changes**.

### Upgrade the Windows Stemcell

1. Retrieve the stemcell by following the steps below for your IaaS:
  - o For vSphere, you must [build your own stemcell](#).
  - o For AWS, GCP, and Azure, navigate to the **Stemcells for PCF (Windows Server)** section of [Pivotal Network](#).
2. Go to **Stemcell Library**.
3. Click **Import Stemcell** to import the stemcell file.
4. When prompted, enable the Ops Manager product checkbox to stage your stemcell.
5. Click **Apply Stemcell to Products**.

For more information about Windows stemcells, see [Downloading or Creating Windows Stemcells](#).

## Troubleshooting Windows Cells

This topic describes how to troubleshoot Windows Diego cells deployed by Pivotal Application Service (PAS) for Windows.

### Installation Issues

This section describes issues that may occur during the installation process.

#### Missing Local Certificates for Windows File System Injector

##### Symptom

You run the `winfs-injector` and see the following error about certificates:

```
Get https://auth.docker.io/token?service=registry.docker.io&
scope=repository:cloudfoundry/windows2016fs:pull: x509:
failed to load system roots and no roots provided
```

##### Explanation

Local certificates are needed to communicate with Docker Hub.

##### Solution

Install the necessary certificates on your local machine. On Ubuntu, you can install certificates with the `ca-certificates` package.

#### Outdated Version for Windows File System Injector

##### Symptom

You run the `winfs-injector` and see the following error about a missing file or directory:

```
open ...windows2016fs-release/VERSION: no such file or directory
```

##### Explanation

You are using an outdated version of the `winfs-injector`.

##### Solution

From [Pivotal Network](#), download the recommended version of **PAS for Windows File System Injector** tool for the tile.

#### Missing Container Image

##### Symptom

You click the + icon in Ops Manager to add the PAS for Windows tile to the Installation Dashboard, and you see the error:

**⚠️ is invalid: has an invalid release 'windows2016fs' specified for job type 'Windows Diego Cell'**

##### Explanation

The product file you are trying to upload does not contain the Windows Server container base image.

##### Solution

1. Delete the product file listing from Ops Manager by clicking its trash can icon under **Import a Product**.
2. Follow the PAS for Windows [installation instructions](#) to run the `winsf-injector` tool locally on the product file. This step requires internet access, can take up to 20 minutes, and adds the Windows Server container base image to the product file.
3. Click **Import a Product** to upload the “injected” product file.
4. Click the + icon next to the product listing to add the PAS for Windows tile to the Installation Dashboard.

## Windows Cell Log Types

Windows cells generate two types of logs:

- **BOSH job logs**, such as `rep_windows` and `consul_agent_windows`. These logs stream to the syslog server configured in the PAS tile **System Logging** pane, along with other PCF component logs. The names of these BOSH job logs correspond to the names of the logs emitted by Linux Diego cells.
- **Windows Event logs**. These stream to the syslog server configured in the PAS for Windows **System Logging** pane and are downloadable through Ops Manager, as described [below](#).

## Access Windows Event Logs

PCF operators can access log messages from Windows Diego cells in two ways:

- Configure PAS for Windows to [send](#) all Windows cell logs to an external syslog server.
- [Download](#) archived logs from each Windows cell individually.

## Send Cell Logs to a Syslog Server

To forward Windows cell log messages to an external syslog server, complete the following steps:

1. Navigate to the Ops Manager Installation Dashboard.
2. Click the **PAS for Windows** tile.

The dialog box contains the following fields:

- Enable Syslog for VM logs?\***: A question mark icon followed by a required field indicator (\*).
- Address \***: An input field for the IP address of the syslog server.
- Port \***: An input field for the port number.
- Protocol\***: A dropdown menu set to `UDP protocol`.
- Enable**: A selected radio button.
- Disable**: An unselected radio button.
- Enable System Metrics**: A checked checkbox.
- Save**: A blue button at the bottom right.

3. Under the **Settings** tab, select the **System Logging** pane.
4. Under **Enable Syslog for VM logs?**, click **Enable**.
5. Under **Address**, enter the IP address of your syslog server.

6. Under **Port**, enter the port of your syslog server. The typical port for a syslog server is `514`.

 **Note:** The host must be reachable from the PAS network. Ensure your syslog server listens on external interfaces.

7. Under **Protocol**, select the transport protocol to use when forwarding logs.
8. Select **Enable System Metrics**. For a list of the VM metrics that the System Metric Agent emits, see [VM Metrics](#) in GitHub.
9. Click **Save**.

## Download Cell Logs

Perform the following steps to retrieve the logs from a Windows cell:

1. Navigate to the Ops Manager Installation Dashboard.
2. Click the **PAS for Windows** tile.
3. Click the **Status** tab.
4. Under the **Logs** column, click the download icon for the Windows cell you want to retrieve logs from.
5. Click the **Logs** tab.
6. When the logs are ready, click the filename to download them.
7. Unzip the file to examine the contents. Each component on the cell has its own logs directory:

- o `/consul_agent_windows/`
- o `/garden-windows/`
- o `/metron_agent_windows/`
- o `/rep_windows/`

## Connect to a Windows Cell

Perform the following steps to connect to a Windows cell to run diagnostics:

1. Download and install a Remote Desktop Protocol (RDP) client.
  - o For Mac OS X, download the Microsoft Remote Desktop app from the [Mac App Store](#).
  - o For Windows, download the Microsoft Remote Desktop app from [Microsoft](#).
  - o For Linux/UNIX, download a RDP client like [rdesktop](#).
2. Follow the steps in the [Log into BOSH](#) section of the *Advanced Troubleshooting with the BOSH CLI* topic to log in to your BOSH Director.
  - o The steps vary slightly depending on whether your PCF deployment uses internal authentication or an external user store.
  - o In PAS for Windows v2.6 and later, use of the `bosh ssh` command is enabled by default.
3. To retrieve the IP address of your Windows cell using the BOSH CLI, run the following command:

```
bosh -e ENV-NAME -d DEPLOYMENT-NAME
```

Where:

- o `ENV-NAME` is the alias you assigned to your BOSH Director.
- o `DEPLOYMENT-NAME` is your deployment's name.

For example:

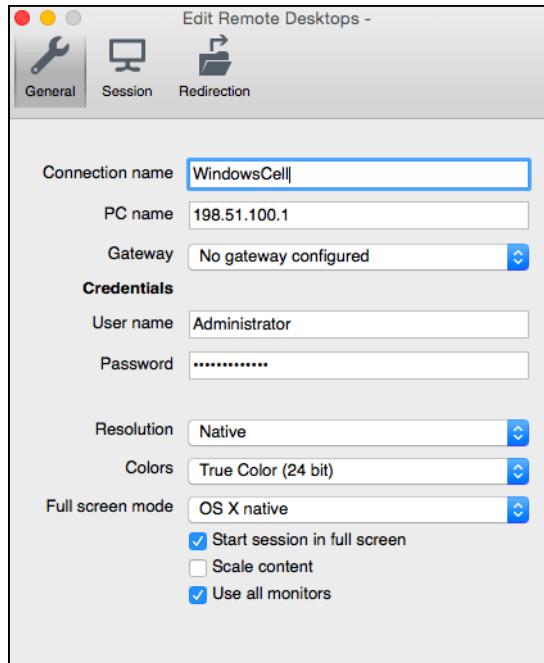
```
c:\Users\admin> bosh -e my-environ -d garden-windows
Using environment '192.0.2.6' as client 'admin'

Name Release(s) Stemcell(s) Team(s) Cloud Config
garden-windows ... - latest
```

4. Retrieve the Administrator password for your Windows cell by following the steps for your IaaS:

- On vSphere, this is the value of `WINDOWS_PASSWORD` in the `consumer-vars.yml` file you used to previously build a stemcell.
- On Amazon Web Services (AWS), navigate to the AWS EC2 console. Right-click on your Windows cell and select **Get Windows Password** from the drop-down menu. Provide the local path to the `ops_mgr.pem` private key file you used when installing Ops Manager and click **Decrypt password** to obtain the Administrator password for your Windows cell.
- On Google Cloud Platform (GCP), navigate to the Compute Engine Dashboard. Under **VM Instances**, select the instance of the Windows VM. At the top of the page, click on **Create or reset Windows password**. When prompted, enter “Administrator” under **Username** and click **Set**. You will receive a one-time password for the Windows cell.
- You cannot RDP into Windows cells on Azure.

5. Open your RDP client. The examples below use the Microsoft Remote Desktop app.

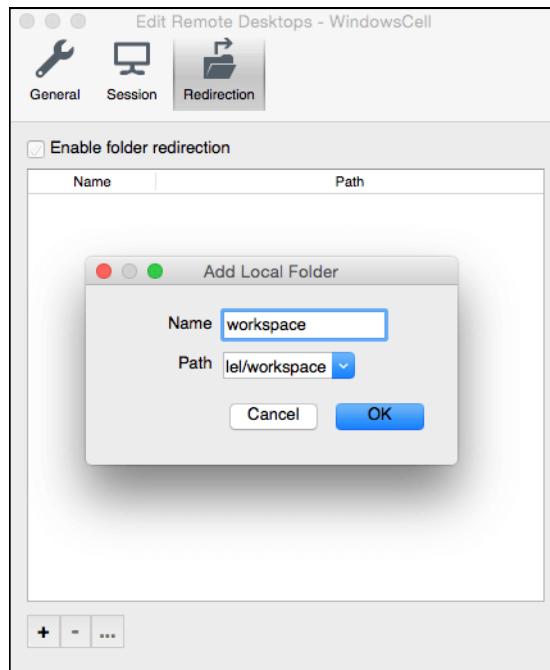


6. Click **New** and enter your connection information:

- **Connection name:** Enter a name for this connection.
- **PC name:** Enter the IP address of your Windows cell.
- **User name:** Enter `Administrator`.
- **Password:** Enter the password of your Windows cell that you obtained above.

7. To mount a directory on your local machine as a drive in the Windows cell, perform the following steps:

- a. From the same **Edit Remote Desktops** window as above, click **Redirection**.



- b. Click the plus icon at the bottom left.
  - c. For **Name**, enter the name of the drive as it will appear in the Windows cell. For **Path**, enter the path of the local directory.
  - d. Click **OK**.
8. Close the **Edit Remote Desktops** window and double-click the newly added connection under **My Desktops** to open a RDP connection to the Windows cell.
9. In the RDP session, you can use the [Consul CLI](#) to diagnose problems with your Windows cell.

## Consul CLI

Perform the following steps to use the Consul CLI on your Windows cell to diagnose problems with your Consul cluster:

1. In your RDP session, open a PowerShell window.
2. To change to the directory containing the Consul CLI binary, run the following:

```
cd CONSUL-CLI-DIR\bin
```

Where `CONSUL-CLI-DIR` is the Consul CLI package's directory path.

For example:

```
PS C:\Users\admin> cd C:\var\vcap\packages\consul-windows\bin\
```

3. To list the members of your Consul cluster, run the following:

```
consul.exe members
```

For example:

```
PS C:\var\vcap\packages\consul-windows\bin> consul.exe members
Node Address Status Type Build Protocol DC
cell-windows-0 10.0.0.111:8301 alive client 0.6.4 2 dc1
cloud-controller-0 10.0.0.94:8301 alive client 0.6.4 2 dc1
cloud-controller-worker-0 10.0.0.99:8301 alive client 0.6.4 2 dc1
consul-server-0 10.0.0.96:8301 alive server 0.6.4 2 dc1
diego-brain-0 10.0.0.109:8301 alive client 0.6.4 2 dc1
diego-cell-0 10.0.0.103:8301 alive client 0.6.4 2 dc1
diego-cell-1 10.0.0.104:8301 alive client 0.6.4 2 dc1
diego-cell-2 10.0.0.107:8301 alive client 0.6.4 2 dc1
diego-database-0 10.0.0.92:8301 alive client 0.6.4 2 dc1
ha-proxy-0 10.0.0.254:8301 alive client 0.6.4 2 dc1
nfs-server-0 10.0.0.100:8301 alive client 0.6.4 2 dc1
router-0 10.0.0.105:8301 alive client 0.6.4 2 dc1
uaa-0 10.0.0.93:8301 alive client 0.6.4 2 dc1
```

- Examine the output to ensure that the `cell-windows-0` service is registered in the Consul cluster and is `alive`. Otherwise, your Windows cell cannot communicate with your PCF deployment and developers cannot push .NET apps to the Windows cell. Check the configuration of your Consul cluster, and ensure that your certificates are not missing or misconfigured.

## Deploying .NET Apps

This topic lists resources for .NET app developers.

For general information about app deployment, see [Pushing an App](#).

### .NET Core on Linux

[.NET Core Buildpack](#) describes how to push .NET Core apps to Linux cells on Cloud Foundry using the .NET Core buildpack. For supported ASP.NET Core versions, see [.NET Core buildpack release notes](#) in GitHub.

### .NET Framework on Windows

[HWC Buildpack](#) describes how to push .NET Framework apps to Pivotal Application Service for Windows (PASW) using the HWC buildpack. It also describes what features of IIS (Internet Information Services) and HWC (Hostable Web Core) are supported, and how to configure your .NET Framework apps to run on PASW.

[Console Applications](#) in the *.NET Cookbook* provides information about pushing compiled .NET Framework Console apps to PASW using the binary buildpack.

### .NET Core on Windows

[Binary Buildpack](#) describes how to push compiled .NET Core apps to Windows cells on PASW using the binary buildpack.

## Other Resources

[Tips for .NET Framework Developers](#) describes common errors and challenges faced when pushing .NET Framework apps to PASW.

The [.Net Cookbook](#) provides code samples, guidelines, tips, and other recipes for making changes to both new and existing .NET Framework and .NET Core apps to run successfully on PASW.

## Using SMB Volumes in .NET Apps

This topic describes how to use Server Message Block (SMB) Volumes in a .NET Application. In this example, we create an application that reads and writes to a `note.txt` file within an SMB Volume share.

### Prerequisites

The following is required to use SMB Volumes in your .NET App:

- Visual Studio
- An IaaS account
- Pivotal Application Service for Windows (PASW)
- An accessible SMB Volume in an IaaS with a UNC path
- An existing .NET Framework application
- SMB Username
- SMB Password

### Using SMB Volumes in .NET Apps with Steeltoe

Steeltoe is a set of libraries that help your team write cloud native applications. Steeltoe contains functionality for mounting SMB Volumes. Using Steeltoe requires modifying your application code. For more information, see the [Steeltoe Documentation](#).

If you are using Steeltoe, add the following environment variables to your Config Server Provider:

- `SMB_PATH`
- `SMB_USERNAME`
- `SMB_PASSWORD`

Where:

- `SMB_PATH` is the UNC path to your SMB.
- `SMB_USERNAME` is your IaaS account username.
- `SMB_PASSWORD` is your IaaS account password.

For more information about adding environment variables to your Config Server Provider, see [Config Server Provider](#) in the Steeltoe documentation.

After adding your environment variables to your Config Server Provider, proceed to [SMB Mounting](#).

### Using SMB Volumes in .NET Apps with Your Batch Profile

If you do not want to modify your application code to include Steeltoe, you can utilize SMB Mounting without Steeltoe.

1. Retrieve the UNC of your existing SMB share. The UNC is your machine's FQDN.
2. Create a `.profile.bat` file in your .NET app's root directory.
3. Use Apps Manager or the Cloud Foundry Command Line Interface (cf CLI) to set the following environment variables:

- `SMB_PATH`
- `SMB_USERNAME`
- `SMB_PASSWORD`

4. Add the following command to your `.profile.bat` file:

```
net use z: %SMB-PATH% %SMB-PASSWORD% /USER:%SMB-USERNAME%
```

5. Proceed to [SMB Mounting](#).

## SMB Mounting

1. Using CF CLI or Apps Manager, add `SMB_PATH`, `SMB_USERNAME`, and `SMB_PASSWORD` as environment variables to your local computer or the computer you are deploying using for deploying your application.
  - o If you are using SMB Volumes with your batch profile and have already added these environment variables, ignore this step.

 **Note:** If Visual Studio does not detect these new environment variables, restart Visual Studio.

2. In Visual Studio, create a new file in your solution named `SMBConfiguration.cs`. This file is the single representation of your SMB Volume configuration and reads the connection data from the environment variables you established previously.

```
// SMBConfiguration.cs
using System;

namespace NetFrameworkApp.Controllers
{
 public class SMBConfiguration
 {
 public String GetSharePath()
 {
 return Environment.GetEnvironmentVariable("SMB_PATH");
 }

 public String GetUserName()
 {
 return Environment.GetEnvironmentVariable("SMB_USERNAME");
 }

 public String GetPassword()
 {
 return Environment.GetEnvironmentVariable("SMB_PASSWORD");
 }
 }
}
```

3. In Visual studio, create a new MVC Controller named `NoteController`. This file creates a controller endpoint that reads a the example note file. For more information about creating a controller, see [Add a controller to an ASP.NET Core MVC app](#) in the Microsoft Documentation.
4. In the command line, add `Steeltoe.Common` and `Steeltoe.Common.Net` to your application with a Package Manager of your choice. If you are not using Steeltoe, ignore this step.
5. Edit `NoteController.cs` to read from a file named `note.txt`, which does not exist yet but the `FileMode.OpenOrCreate` method creates that file. Refer to the example code snippet below, which reads the contents of the note file and stashes the `note.txt` content in the Viewbag. If you are not using Steeltoe, ignore the reference to `Steeltoe.Common.Net` in the following code snippet.

```
// NoteController.cs
using System;
using System.Collections.Generic;
using System.IO;
using System.Linq;
using System.Net;
using System.Web;
using System.Web.Mvc;
using Steeltoe.Common.Net;

namespace NetFrameworkApp.Controllers
{
 public class NoteController : Controller
 {
 SMBConfiguration configuration = new SMBConfiguration();
 public ActionResult Index()
 {

 var credential = new NetworkCredential(configuration.GetUserName(), configuration.GetPassword());

 using (var share = new WindowsNetworkfileShare(configuration.GetSharePath(), credential))
 using (var inputStream = new FileStream(Path.Combine(configuration.GetSharePath(), "note.txt"), FileMode.OpenOrCreate))
 using (var streamReader = new StreamReader(inputStream))
 {
 // Never display raw user input as HTML. Do not do this in production code.
 }
 }
 }
}
```

```
 ViewBag.Note = streamReader.ReadToEnd();
 }

 return View();
}

}
}
```

6. In Visual Studio, create a subdirectory in **Views** named `Note`.
  7. In Visual Studio, create a new View named `Index`. For more information about Views, see [Views in ASP.NET Core MVC](#) and [Add a view to an ASP.NET Core MVC app](#) in the Microsoft Documentation.
  8. In Visual Studio, create a `Index.cshtml` file that contains the following:

```
// Index.cshtml
```

Running the application now would show an empty page with no errors.

9. Modify the `Index.cshtml` file to contain a form. This form posts to a yet to be created update endpoint and also displays our note inside a text area.

```
// Index.cshtml
...
<form action="https://docs.pivotal.io/note/update" method="post">
 <textarea name="note">@ViewBag.Note</textarea>
 <div>
 <button type="submit">Update</button>
 </div>
</form>
```

10. Modify the `NoteController.cs` to have an update endpoint. This endpoint saves the data posted to the endpoint back into the `note.txt`. Refer to the example code snippet in the following step.
  11. Modify the `NoteController.cs` to include the  `ControllerBase` class `RedirectToAction` method, which redirects to the index page so the user can see what was just saved. For more information about the  `ControllerBase` class, see the [Microsoft .NET API Browser Documentation](#).

Refer to the example code snippet below.

```
// NoteController.cs
namespace NetFrameworkApp.Controllers
{
 public class NoteController : Controller
 {
 ...
 [HttpPost]
 public ActionResult Update(String note)
 {
 var credential = new NetworkCredential(configuration.GetUserName(), configuration.GetPassword());

 using (var share = new WindowsNetworkFileShare(configuration.GetSharePath(), credential))
 using (var outputStream = new FileStream(Path.Combine(configuration.GetSharePath(), "note.txt"), FileMode.Create))
 using (var streamWriter = new StreamWriter(outputStream))
 {
 streamWriter.Write(note);
 }

 return RedirectToAction("Index");
 }
 }
}
```

## Known Issue

### Inability to SSH Into App Instance

Despite an SMB mapping being successfully created, trying to `cf ssh` into that app instance to access the created mapping results in an unspecified path error.

## Changing Windows Stacks

Page last updated:

This topic explains how to restage apps on a new Windows stack. It also provides a description of stacks and lists the supported Windows stacks on Pivotal Application Service for Windows (PASW) v2.6.

To restage a Windows app on a new Linux stack, see [Changing Stacks](#).

You can also use the Stack Auditor plugin for the Cloud Foundry Command Line Interface (cf CLI) when changing stacks. See [Using the Stack Auditor Plugin](#).

## Overview

A stack is a prebuilt root file system (rootfs) that supports a specific operating system. For example, Linux-based systems need `/usr` and `/bin` directories at their root and Windows needs `/windows`. The stack works in tandem with a buildpack to support apps running in compartments. Under Diego architecture, cell VMs can support multiple stacks.

 **Note:** Docker apps do not use stacks.

## Available Stacks

If you push your PASW v2.6 app to a Windows stack, you must use `windows`. The `windows2016` stack is not supported on PASW.

## Restaging Apps on a New Stack

For security, stacks receive regular updates to address Common Vulnerabilities and Exposures ([CVEs](#)). Apps pick up on these stack changes through new releases of PASW. However, if your app links statically to a library provided in the rootfs, you may have to manually restage it to pick up the changes.

It can be difficult to know what libraries an app statically links to, and it depends on the languages you are using. One example is an app that uses a Ruby or Python binary, and links out to part of the C standard library. If the C library requires an update, you may need to recompile the app and restage it.

To restage an app on a new stack, do the following:

1. Use the `cf stacks` command to list the stacks available in a deployment.

```
$ cf stacks
Getting stacks in org MY-ORG / space development as developer@example.com...
OK

name description
windows2016 Windows Server 2016
windows Windows Server
```

2. To change your stack and restage your app, run the following command:

```
cf push MY-APP -s STACK-NAME
```

Where:

- MY-APP is the name of the app.
- STACK-NAME is the name of the new stack.

For example, to restage your app on the `windows` stack, run `cf push MY-APP -s windows`:

```
$ cf push MY-APP -s windows
Using stack windows...
OK
Creating app MY-APP in org MY-ORG / space development as developer@example.com...
OK
...
requested state: started
instances: 1/1
usage: 1G x 1 instances
urls: MY-APP.cfapps.io
last uploaded: Wed Apr 8 23:40:57 UTC 2015
 state since cpu memory disk
#0 running 2015-04-08 04:41:54 PM 0.0% 57.3M of 1G 128.8M of 1G
```

## Stacks API

For API information, see the *Stacks* section of the [Cloud Foundry API Documentation](#).

## Using Apps Manager

The web-based Apps Manager application helps you manage users, organizations, spaces, and applications.

## Table of Contents

- [Getting Started with Apps Manager](#)
- [Managing Orgs and Spaces Using Apps Manager](#)
- [Managing User Roles with Apps Manager](#)
- [Managing Apps and Service Instances Using Apps Manager](#)
- [Viewing ASGs in Apps Manager](#)
- [Monitoring Instance Usage with Apps Manager](#)
- [Configuring Spring Boot Actuator Endpoints for Apps Manager](#)
- [Using Spring Boot Actuator Endpoints with Apps Manager](#)

## Getting Started with Apps Manager

This topic describes the functions and scope of Apps Manager, a web-based tool for managing Pivotal Application Service (PAS) organizations, spaces, applications, services, and users.

### Scope

Apps Manager provides a visual interface for performing the following subset of functions available through the Cloud Foundry Command Line Interface (cf CLI):

- **Orgs:** You can create and manage orgs.
- **Spaces:** You can create, manage, and delete spaces.
- **Apps:** You can scale apps, bind apps to services, manage environment variables and routes, view logs and usage information, start and stop apps, and delete apps.
- **Services:** You can bind services to apps, unbind services from apps, choose and edit service plans, and rename and delete service instances.
- **Users:** You can invite new users, manage user roles, and delete users.

To access Apps Manager as the Admin user, see the [Logging in to Apps Manager](#) topic.

### Browser Support

Apps Manager is compatible with the latest major versions of the following browsers:

- [Apple Safari ↗](#)
- [Google Chrome ↗](#)
- [Microsoft Edge ↗](#)
- [Microsoft Internet Explorer ↗](#)
- [Mozilla Firefox ↗](#)

Pivotal recommends using Chrome, Firefox, Edge, or Safari for the best Apps Manager experience.

### About Permissions

Your ability to perform actions in Apps Manager depends on your user role and the [feature flags](#) that the Admin sets.

The table below shows the relationship between specific org and space management actions and the non-Admin user roles who can perform them. A non-Admin user must be a member of the org and space to perform these actions.

Admin users can perform all of these actions using either the cf CLI or by logging into Apps Manager as an Org Manager, using the UAA Admin credentials.

Space Managers assign and remove users from spaces by setting and unsetting their roles within the space.

Action	CLI command	Org Manager	Space Manager	Org Auditor, Space Developer, or Space Auditor
Create an org	<code>create-org</code>	†	†	†
Delete an org	<code>delete-org</code>	No	No	No
Rename an org	<code>rename-org</code>	Yes	No	No
View org members	<code>org-users</code>	Yes	Yes	Yes
Assign user a role in org	<code>set-org-role</code>	‡	‡	No
Remove org role from user	<code>unset-org-role</code>	‡	‡	No
View space members	<code>space-users</code>	Yes	Yes	Yes
Assign user a role in space	<code>set-space-role</code>	‡	‡	No
Remove space role from	<code>unset-space-role</code>	‡	‡	No

user

<sup>†</sup>Defaults to no. Yes if `feature flag user_org_creation` is set to `true`.

<sup>‡</sup>Defaults to no. Yes if `feature_flags` `set_roles_by_username` and `unset_roles_by_username` are set to `true`.

## Managing Orgs and Spaces Using Apps Manager

This topic discusses how to view and manage orgs and spaces in Apps Manager.

 **Note:** To manage a space, you must have Space Manager permissions in that space.

To perform the following steps, you must first log in to Apps Manager with an account that has adequate permissions. See the [About Permissions](#) topic for more information.

### Manage an Org

The following sections describe how to manage your orgs using the org management interface in Apps Manager. You can navigate to your org from the [Home](#) page or through the global search.

For details about each tab in the org page, see the following sections.

#### Spaces

In the **Spaces** tab, you can do the following:

- See the following information for each space:
  - Name
  - Number of apps
  - Number of apps in running, stopped, and crashed states
  - Number of services
  - Memory quota usage
- Create a new space

#### Domains

In the **Domains** tab, you can do the following:

- See each **Domain** in your org.
- See the **Type** associated with each domain. A type is either **Globally Shared** or **Privately Shared**. For more information about globally and locally shared domains, see [Shared Domains](#) and [Private Domains](#).
- Delete a domain by clicking the vertical three-dot icon and selecting **Delete**. You must have permission to manage the domain to delete it.
- Create a new private domain by clicking **Add a Domain**. You must have Org Manager permissions to create a private domain.

 **Note:** To create globally shared domains, you must have Admin permissions and use the Cloud Foundry Command Line Interface (cf CLI). For more information, see [Create a Shared Domain](#).

See the following screenshot for an example **Domains** interface with Org Manager permissions:

The screenshot shows the Pivotal Apps Manager interface. At the top, there is a green banner with a checkmark icon and the text "Domain 'dev.private' was successfully created". Below this, the header includes "ORG test-org", "QUOTA 768 MB / 10 GB 8%", and a "Usage Report" link. The navigation tabs are "Spaces (3)", "Domains (3)" (which is selected), "Members (10)", and "Settings". A "Domains" section displays a table with three rows: "apps.internal" (Global), "apps.test.cf-app.com" (Global), and "dev.private" (Private). A blue "ADD A DOMAIN" button is located at the top right of this section.

## Members

In the **Members** tab, you can do the following:

- See member email addresses
- Toggle member permissions
- Remove member from org
- Invite a new member

## Settings

In the **Settings** tab, you can do the following:

- Change the org name
- See spaces assigned to isolation segments. For more information, see [Isolation Segments](#).
- Delete the org. You must have admin permissions to delete an org.

## Add Metadata

You can add metadata, including labels and annotations, to orgs using Apps Manager.

For more information about adding metadata to objects in PAS, see [Using Metadata](#).

To add labels or annotations to a org using Apps Manager, do the following:

1. Click the **Settings** tab of the desired org.
2. For **Labels**, provide a name and value pair. Or, select **Enter JSON** to enter the label in JSON.
3. For **Annotations**, provide a name and value pair. Or, select **Enter JSON** to enter the annotation in JSON.
4. To add more labels or annotations, click **+** next to the **Labels** or **Annotations** field.
5. Click **Update Metadata**.

## Manage a Space

The space page displays the apps, service instances, and routes associated with the selected space.

SPACE      RUNNING    STOPPED    CRASHED

**development**    ● 5    ● 0    ● 0

[Apps \(5\)](#)    [Services \(2\)](#)    [Routes \(2\)](#)    [Settings](#)

Apps						
Status	Name	Instances	Memory	Last Push	Route	
● Running	hello-spring-cloud	1	256 MB	4 hours ago	<a href="https://docs-team-spring-app.apps.oozie-boo...">https://docs-team-spring-app.apps.oozie-boo...</a> ↗	
● Running	my-super-app	1	20 MB	4 hours ago	no bound route	
● Running	pong-matcher	1	64 MB	4 hours ago	<a href="https://docs-team-broker-app.apps.oozie-boo...">https://docs-team-broker-app.apps.oozie-boo...</a> ↗	
● Running	spring-music	1	20 MB	4 hours ago	no bound route	
● Running	the-other-app	1	20 MB	4 hours ago	no bound route	

## Apps

The **Apps** tab shows the **Status**, the **Name**, the number of **Instances**, the amount of **Memory** available, the time since the **Last Push**, and the **Route** for each app.

Apps						
Status	Name	Instances	Memory	Last Push	Route	
● Running	hello-spring-cloud	1	256 MB	4 hours ago	<a href="https://docs-team-spring-app.apps.oozie-boo...">https://docs-team-spring-app.apps.oozie-boo...</a> ↗	
● Running	my-super-app	1	20 MB	4 hours ago	no bound route	

## Services

The **Services** list shows the **Service**, the **Name**, the number of **Bound Apps**, and the **Plan** for each service instance. If you want to add a service to your space, click **Add Service**. For more information about configuring services, see the [Services Overview](#) ↗ topic.

Services					<a href="#">ADD A SERVICE</a>
Service	Name	Bound Apps	Plan		
 App Autoscaler	my-autoscaler	0	free - Standard		

## Routes

From the **Routes** tab, you can view and delete routes associated with your space. For each **Route**, the tab shows **Route Service** and **Bound Apps**. Refer to the [Route Services](#) topic to learn more about route services.

Routes			
Route	Route Service	Bound Apps	
docs-team-broker-app.apps.oogie-boogie... ↗	my-service	● pong-matcher	X
docs-team-spring-app.apps.oogie-boogie... ↗	no bound service	● hello-spring-cloud	X

## Settings

From the **Settings** tab, you can do the following:

- Modify the space name by entering a new name and clicking **Update**.
- View the Application Security Groups (ASGs) associated with the space in the **Security Groups** section.
- View the space's isolation segment assignment in the **Isolation Segment Assignment** section. For more information, see [Isolation Segments](#).
- Add metadata to the space. See [Add Metadata](#).
- Delete the space by clicking **Delete Space**.

Space Name

my-space

UPDATE CANCEL

Metadata

Name	Value	Enter JSON
		+

Name	Value	Enter JSON
		+

CANCEL UPDATE METADATA

Security Groups

A collection of egress rules that specify one or more individual protocols, ports, and destinations. [Learn more](#)

my-network  
running, staging

Isolation Segment Assignment

none

## Add Metadata

You can add metadata, including labels and annotations spaces using Apps Manager.

For more information about adding metadata to objects in PAS, see [Using Metadata](#).

To add labels or annotations to a space using Apps Manager, do the following:

1. Click the **Settings** tab of the desired space.
2. For **Labels**, provide a name and value pair. Or, select **Enter JSON** to enter the label in JSON.
3. For **Annotations**, provide a name and value pair. Or, select **Enter JSON** to enter the annotation in JSON.
4. To add more labels or annotations, click **+** next to the **Labels** or **Annotations** field.

5. Click **Update Metadata**.

## Managing User Roles with Apps Manager

 Note: The procedures described here are not compatible with using SAML or LDAP for user identity management. To create and manage user accounts in a SAML- or LDAP-enabled Cloud Foundry deployment, see [Adding Existing SAML or LDAP Users to a Pivotal Cloud Foundry Deployment](#).

Cloud Foundry uses role-based access control, with each role granting the permissions in either an org or an application space.

A user account can be assigned one or more roles.

The combination of these roles defines the actions a user can perform in an org and within specific app spaces in that org.

To view the actions that each role allows, see the [Organizations, Spaces, Roles, and Permissions](#) topic. For example, to assign roles to user accounts in a space, you must have Space Manager role assigned to the user in that space.

You can also modify permissions for existing users by adding or removing the roles associated with the user account. User roles are assigned on a per-space basis, so you must modify the user account for each space that you want to change.

Admins, Org Managers, and Space Managers can assign user roles with Apps Manager or with the Cloud Foundry Command Line Interface (cf CLI). For more information, see the [Users and Roles](#) section of the *Getting Started with the cf CLI* topic.

### Manage Org Roles

Valid [org roles](#) are Organization Manager and Organization Auditor.

To grant or revoke org roles, follow the steps below.

1. Go to the [Home](#) page
2. Select an org.
3. Click the **Members** tab. Edit the roles assigned to each user by selecting or clearing the checkboxes under each user role. Apps Manager saves your changes automatically.
4. The **Members** panel displays all members of the org. Select a checkbox to grant an org role to a user, or clear a checkbox to revoke a role from a user.

### Manage App Space Roles

Valid [app space roles](#) are Space Manager, Space Developer, and Space Auditor.

To grant or revoke app space roles, do the following:

1. Go to the page for a space
2. Select the **Members** tab. The **Members** panel displays all members of the space.
3. Select a checkbox to grant an app space role to a user, or clear a checkbox to revoke a role from a user.
  - o **Space Managers** can invite and manage users and enable features for a given space. Assign this role to managers or other users who need to administer the account.
  - o **Space Developers** can create, delete, and manage applications and services, and have full access to all usage reports and logs. Space Developers can also edit applications, including the number of instances and memory footprint. Assign this role to app developers or other users who need to interact with applications and services.
  - o **Space Auditors** have view-only access to all space information, settings, reports, and logs. Assign this role to users who need to view but not edit the application space.

### Invite New Users

 Note: The **Enable Invitations** checkbox in the Apps Manager section of the PAS tile must be selected to invite new users.

1. On the Org dashboard, click the **Members** tab.
2. Click **Invite New Members**. The **Invite New Team Member(s)** form appears.

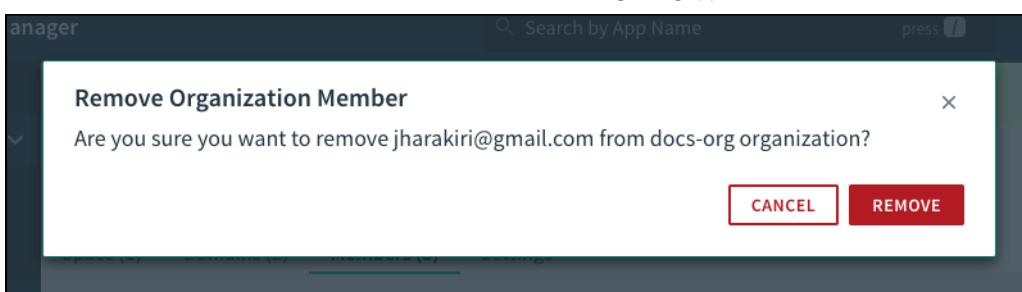
The screenshot shows the 'Invite New Team Member(s)' form. At the top, it displays the organization details: 'docs-org' (Space 1), '1.62 GB / 10 GB' (QUOTA 16%). Below this, tabs for 'Space (1)', 'Domains (2)', 'Members (6)', and 'Settings' are visible. The 'Members (6)' tab is selected. The main area contains two sections: 'Add Email Addresses' (with a text input placeholder 'Use commas to separate emails') and 'Assign Org Roles' (with a table showing checkboxes for 'Org Manager' and 'Org Auditor' for 'docs-org'). There's also a 'Select All' checkbox. Below this is another table for 'Assign Space Roles' (checkboxes for 'Space Manager', 'Space Developer', and 'Space Auditor' for 'docs-space'), with a 'Select All' checkbox. At the bottom right are 'CANCEL' and 'SEND INVITE' buttons.

3. In the **Add Email Addresses** text field, enter the email addresses of the users that you want to invite. Enter multiple email addresses as a comma-delimited list.
4. The **Assign Org Roles** and **Assign Space Roles** tables list the current org and available spaces with checkboxes corresponding to each possible user role. Select the checkboxes that correspond to the permissions that you want to grant to the invited users.
5. Click **Send Invite**. The Apps Manager sends an email containing an invitation link to each email address that you specified.

## Remove a User From an Org

Removing a user from org also removes them from all spaces in the org.

1. On the Org dashboard, click the **Members** tab.
2. Locate the user account that you want to remove.
3. Under the user's email address, click on the **Remove User** link. A warning dialog appears.



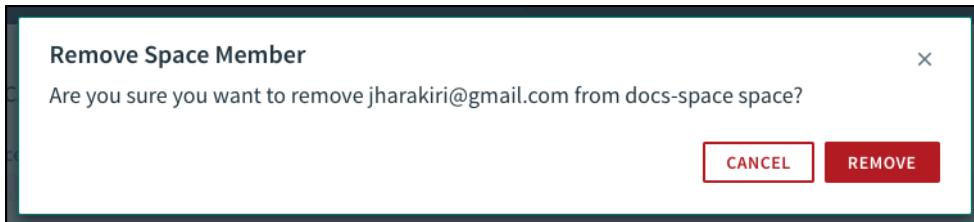
4. Click **Remove** to confirm user account deletion from the org.

## Remove a User From a Space

1. Go to the page for a space
2. Select the **Members** tab. The **Members** panel displays all members of the space
3. Locate the user account that you want to remove.

Member	Space Manager ⓘ	Space Developer ⓘ	Space Auditor ⓘ
admin	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
space-auditor-0019	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
space-auditor-0003	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
space-auditor-0013	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
space-auditor-0006	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

4. Under the user's email address, click on the **Remove User** link. A warning dialog appears.



5. Click **Remove** to confirm user account deletion from the space.

## Managing Apps and Service Instances Using Apps Manager

This topic discusses how to view and manage apps and service instances associated with a space using Apps Manager.

To perform the following steps, you must first log in to Apps Manager with an account that has adequate permissions. See the [About Permissions](#) topic for more information.

### Manage an App

To manage an app, navigate to an app page with one of the following methods:

- On the space page, click the app you want to manage.
- Search for an app by entering its name in the search bar.

From the app page, you can do the following:

- Scale apps
- Bind apps to services
- Manage environment variables and routes
- View logs and usage information
- Start and stop apps
- Delete apps
- Terminate specific instances of apps

APP  
cf-nodejs ● Running

[VIEW APP](#)

Overview Services Route (1) Logs Tasks Settings Buildpack: N/A

Events Last Push: 12:25 PM 02/07/18 App Summary

No recent events

Instances / Allocated	Memory / Allocated	Disk / Allocated
3 / 4	0.05 / 2.63 GB	0.10 / 1.63 GB

Processes and Instances

**web**

Instances	Memory Allocated	Disk Allocated	SCALE
2	1 GB	512 MB	<a href="#">SCALE</a>

Autoscaling

#	CPU	Memory	Disk	Uptime
0	0%	25.61 MB	51.07 MB	4 d 21 hr 41 min
1	0%	27.83 MB	51.07 MB	4 d 21 hr 40 min

**worker**

Instances	Memory Allocated	Disk Allocated	SCALE
1	512 MB	512 MB	<a href="#">SCALE</a>

#	CPU	Memory	Disk	Uptime
0	0%	0 Bytes	0 Bytes	4 d 21 hr 42 min

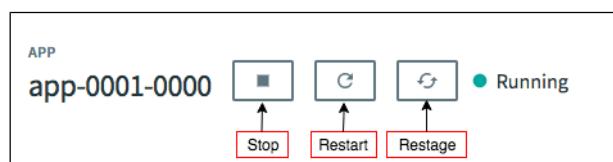
**scheduler**

Instances	Memory Allocated	Disk Allocated	SCALE
1	130 MB	128 MB	<a href="#">SCALE</a>

#	CPU	Memory	Disk	Uptime
0	CRASHED			

## Start, Stop, or Restage an App

The start, stop, and restage buttons are located next to the name of your app. See the diagram below:



To perform actions with the three buttons, see the table below:

To...	Do the following:
Stop an app	Click the  (stop) button
Restart a stopped or running app	Click the  (restart) button
Restage an app	Click the  (restage) button.

For more information about starting, stopping, and restaging apps, but with the Cloud Foundry Command Line Interface (cf CLI), see [Starting, Restarting, and Restaging Applications](#).

## Scale an App

From the app Overview pane, Space Developers can scale an app [manually](#) or configure [App Autoscaler](#) to scale it automatically.

### Scale an App Manually

#### Processes and Instances

**web**

Instances	Memory Allocated	Disk Allocated	SCALE
2	1 GB	512 MB	<input type="button" value="SCALE"/>

**Autoscaling**

#	CPU	Memory	Disk	Uptime
0	0%	6.63 MB	51.07 MB	4 d 21 hr 58 min
1	0%	27.98 MB	51.07 MB	4 d 21 hr 56 min

1. Click **Scale** to open the “Scale App” dialog.

**Scale app**

Process	Instances	Memory Limit	Disk Limit
web	2	1 GB	512 MB
worker	1	512 MB	512 MB
scheduler	1	130 MB	128 MB
<b>Usage Total</b>	4	2.63 GB	1.63 GB

**APPLY CHANGES**

2. Adjust the number of **Instances**, the **Memory Limit**, and the **Disk Limit** as desired.

3. Click **Apply Changes**.

## Configure App Autoscaler

**Processes and Instances**

Process	Instances	Memory Allocated	Disk Allocated	Action
web	2	1 GB	512 MB	<b>SCALE</b>
<b>Autoscaling</b>	<input checked="" type="checkbox"/>			<a href="#">Manage Autoscaling</a>
#	CPU	Memory	Disk	Uptime
0	0%	6.63 MB	51.07 MB	4 d 21 hr 58 min
1	0%	27.98 MB	51.07 MB	4 d 21 hr 56 min

1. Use **Autoscaling** to enable App Autoscaler.
2. Click **Manage Autoscaling** to open App Autoscaler.
3. See the [Configure Autoscaling for an App](#) section of the *Scaling an Application Using Autoscaler* topic for how to configure your App Autoscaler to scale automatically based on rules or a schedule.

## Bind or Unbind a Service

**Note:** For services that use asynchronous bindings, Apps Manager displays the status of the service while the bind is still pending. Asynchronous bindings provide more flexibility for services that require additional time before returning a successful bind response.

1. Click **Services**.
2. To bind your app to a service, click **Bind Service**.

The screenshot shows a list titled 'Bound Services'. It contains one item: 'App Autoscaler' (free - Standard) with the identifier 'my-autoscaler'. There are 'BIND SERVICE' and 'NEW SERVICE' buttons at the top right, and a three-dot menu icon on the far right.

3. To bind your app to an existing service instance, do the following:

- Click **Bind Service**.

The dialog box has a title 'Bind Service'. It contains fields for 'Service to Bind' (set to 'usr-serv'), 'Binding Name (Optional)', 'Add Parameters (Optional)', and 'Enter JSON' (disabled). At the bottom are 'CLOSE' and 'BIND' buttons.

- Under **Service to Bind**, select the service instance from the dropdown menu.
- (Optional) Enter a binding name under **Binding Name** in all lowercase letters. For example, `bind-usr-serv`.
- (Optional) Specify additional parameters under **Add Parameters**.
- Click **Bind**.

4. To bind your app to a new service instance, do the following:

- Click **New Service**.

The dialog box has a title 'Create a new service' and a 'View in Marketplace' link. It includes a search bar and two service options: 'App Autoscaler' (selected) and 'User Provided Service'.

Service Type	Description
App Autoscaler	Scales bound applications in response to load (beta)
User Provided Service	Add an external service to your apps

**Note:** If you prefer to create the new service instance in the Marketplace, you can click [View in Marketplace](#) at any time.

- b. Click the service.

X Create a new service [View in Marketplace](#)

**App Autoscaler**  
Scales bound applications in response to load (beta)

Standard  
free

Show Details

[BACK](#) [SELECT PLAN](#)

- c. Select a plan and click **Select Plan**.

X Create a new service [View in Marketplace](#)

**App Autoscaler**  
standard - free

Instance Name  
my-autoscaler

Add to Space  
development

Bind to App (Optional)  
cf-spring

Binding Name (Optional)  
bind-autoscaler

Add Parameters ⓘ (Optional) Enter JSON   
Name Value +

[BACK](#) [CREATE](#)

- d. Under **Instance Name**, enter a name for the instance.
  - e. (Optional) Enter a binding name under **Binding Name** in all lowercase letters. For example, `bind-autoscaler`.
  - f. (Optional) Specify additional parameters under **Add Parameters**. For a list of supported configuration parameters, consult the documentation for the service.
  - g. Click **Create**.
5. To unbind your app from a service instance, locate the service instance in the **Bound Services** list and click the three-dot icon on the far right. Select **Unbind** from the dropdown menu.

## Map or Unmap Routes

1. Click **Routes**.
2. The page displays the routes associated with the app. To add a new route, click **Map a Route**.

Routes	Map a Route
http://spring-music-ageless-hydrazoate.cfapps.io	

3. Enter the route and click **Map**.
4. To unmap a route, locate the route from the list and click the red . Click **Unmap** in the pop-up to confirm.

## Create Container-to-Container Networking Policies

Container networking policies enable app instances to communicate with each other directly. You can create container networking policies in the **Networking** tab.

To view and use the **Networking** tab, you must have either the `network.write` or `network.admin` UAA scope. If you do not see the **Networking** tab, request one of the above scopes from your PCF administrator.

1. Click **Networking**. The page displays any container networking policies associated with the app.

Overview	Services (2)	Route (1)	Networking	Logs	Tasks	Settings	Buildpack: N/A
Container Networking Policies							<b>CREATE POLICY</b>
No network policies exist for this app. <a href="#">Add a network policy.</a>							

2. To add a new network policy, click **Create Policy** or [Add a network policy](#).

## Add Policy

**Org**  
[select an org] 

**Space**  
[select a space] 

**App**  
[select an app] 

**Protocol**  
 TCP  UDP

**Ports**  
e.g. 4000 or 4000-5000

**CANCEL** **SAVE**

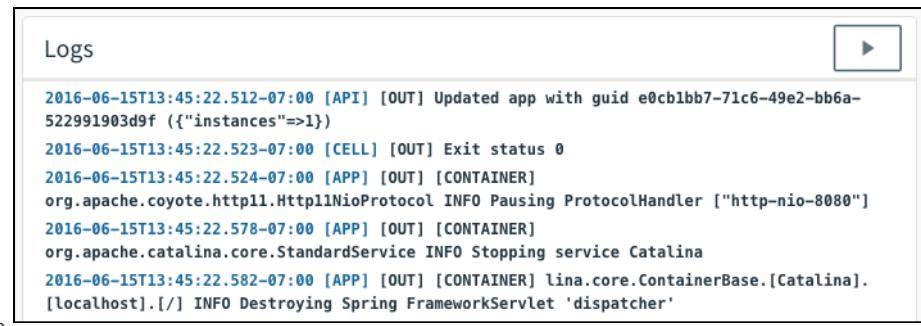
3. In the **Add Policy** window, configure the following:

- For **Org**, select the org of the destination app.
- For **Space**, select the space of the destination app.
- For **App**, select the destination app.
- For **Protocol**, select **TCP** or **UDP**.
- For **Ports**, enter the ports at which to connect to the destination app. The allowed range is from `1` to `65535`. You can specify a single port, such as `8080`, or a range of ports, such as `8080-8090`.

4. Click **Save**.

For more information about container-to-container networking in PCF, see [Container-to-Container Networking](#).

[View Logs](#)



1. Click **Logs** to view the logs for the app.

2. Click the play button to view a live version of the logs.

## View Tasks

1. Click the **Tasks** tab within Apps Manager.
2. This page displays a table containing **Task ID**, **State**, **Start Time**, **Task Name**, and **Command**.

The screenshot shows the 'pong-matcher' app details page. The 'Tasks' tab is selected. The table has columns: Task ID, State, Start Time, Task Name, and Command. There is a 'RUN TASK' button at the top right of the table area.

Task ID	State	Start Time	Task Name	Command
---------	-------	------------	-----------	---------

## Run a Task

1. Click **Run Task** to create a task.

The dialog box has two input fields: 'Task Name (optional)' and 'Task Command'. At the bottom are 'RUN' and 'CANCEL' buttons.

2. (Optional) Enter a **Task Name**.
3. Enter the **Task Command**.
4. Click **Run**.

## Enable Task Scheduling

In the **Tasks** tab, click **Enable Scheduling** to bind the PCF Scheduler service to your app. For more about the PCF Scheduler, see [Scheduling Jobs](#).

The screenshot shows the Pivotal app dashboard for an application named "dev-app". The status is "Stopped". The navigation tabs include Overview, Services, Route (1), Logs, Tasks (which is underlined in blue), and Settings. The Buildpack is listed as "N/A".

**Jobs**

The scheduler service has not been enabled yet.

**ENABLE SCHEDULING**

**Tasks**

No tasks have run for this application.

**RUN TASK**

## Schedule a Task

1. Navigate to the **Tasks** tab.

The screenshot shows the Tasks tab with the following interface elements:

- Overview, Service (1), Route (1), Logs, Tasks (underlined in blue), Settings
- Buildpack: N/A
- Jobs**
- CREATE JOB**
- No tasks are currently scheduled.

2. Click **Create Job** to schedule a task.

3. Enter a **Job Name**.

**Create Job**

**Job Name**:

**Command**:

**Cron Expressions**

**Value**:  **+**

Syntax is described in the [docs](#). Example: 0 12 \* \* ?

**CREATE JOB**

4. Enter a **Command**.

5. Enter one or more **Cron Expressions** for your desired task schedule or schedules. See [Schedule a Job](#) for more information on cron expression syntax.

6. Click **Create Job**.

## View Settings

Click the **Settings** tab. In this tab, you can do the following:

- Rename the app.
- View information about the buildpack(s), start command, and stack.
- Configure health checks.
- View or add Environment Variables associated with the app.
- Add metadata to the app.
- View the Application Security Groups (ASGs) associated with the app.
- Delete the app. When you click **Delete App**, you also have the option to delete the app's routes.

Overview Services Route (1) Logs Tasks Settings

Buildpack: <https://github.com/cloudfoundry/ruby-buildpack#v1.6.28>

**App Name** my-app **UPDATE** **CANCEL**

**Info**

Buildpack: <https://github.com/cloudfoundry/ruby-buildpack#v1.6.28>  
Start Command: `bundle exec rackup config.ru -p $PORT`  
Stack: cflinuxfs3 (Cloud Foundry Linux-based filesystem)

**Health Check**

Type: port  
Invocation timeout (seconds) (Optional):

**Metrics Forwarder**

Metrics Forwarder allows applications to emit metrics into loggregator and consume those metrics from the firehose.

**Metadata**

Labels:

Name	Value	Enter JSON
------	-------	------------

Annotations:

Name	Value	Enter JSON
------	-------	------------

**User Provided Environment Variables**

**REVEAL USER PROVIDED ENV VARS**

**Environment Variables**

Defined by the runtime and buildpack. [Learn more](#)

**REVEAL ENV VARS**

**Security Groups**

A collection of egress rules that specify one or more individual protocols, ports, and destinations. [Learn more](#)

my-dns  
running, staging

**Delete App**

This will permanently delete the app and all of its data.

**DELETE APP**

## Configure Health Checks

To configure health checks for your app, do the following:

1. Click the **Settings** tab.
2. In the **Health Check** section, click the **Type** dropdown and select `port`, `http`, or `process`. Depending on which option you select, additional fields may appear.
3. Complete the fields that appear with your selection.
  - o For `port`, optionally complete the **Invocation timeout (seconds)** field.

Health Check

Type	<input type="text" value="port"/>
Invocation timeout (seconds) (Optional)	<input type="text"/>
<input type="button" value="CANCEL"/> <input type="button" value="UPDATE"/>	

- For `http`, complete the **Endpoint** field. Optionally, complete the **Timeout (seconds)** and **Invocation timeout (seconds)** fields.

Health Check

Type	<input type="text" value="http"/>
Endpoint	<input type="text"/>
Timeout (seconds) (Optional)	<input type="text"/>
Invocation timeout (seconds) (Optional)	<input type="text"/>
<input type="button" value="CANCEL"/> <input type="button" value="UPDATE"/>	

- For `process`, there are no additional fields.

Health Check

Type	<input type="text" value="process"/>
<input type="button" value="CANCEL"/> <input type="button" value="UPDATE"/>	

4. Click **Update**.

For more information, see the [Using Application Health Checks](#) topic.

## View or Add Environment Variables

To add a user-provided environment variable, do the following:

- Click the **Settings** tab.
- Click **Reveal User Provided Env Vars**.
- Enter the **Name** and **Value** of the variable. Alternatively, enter your variable using the **Enter JSON** toggle.
- Click **Save**.

User Provided Environment Variables

<input type="text" value="Name"/>	<input type="text" value="Value"/>	<input checked="" type="checkbox" value="Enter JSON"/> +
<input type="button" value="SAVE"/> <input type="button" value="CANCEL"/>		

To view all environment variables, click **Reveal Env Vars**.

## Environment Variables

Defined by the runtime and buildpack. [Learn more](#)

```
{
 "staging_env_json": {},
 "running_env_json": {},
 "system_env_json": {
 "VCAP_SERVICES": {
 "fake-service-a455f60c-f166-42a5-bbab-dfadbf0531": [
 {
 "credentials": {
 "username": "myapp",
 "password": "myapp123"
 }
 }
]
 }
 }
}
```

 **Note:** Changes to environment variables, service bindings, and service unbindings require restarting the app to take effect. You can restart the app from the Apps Manager or with the Cloud Foundry Command Line Interface `cf restage` command.

## Add Metadata

You can add metadata, including labels and annotations, to apps and spaces using Apps Manager.

For more information about adding metadata to objects in PAS, see [Using Metadata](#).

To add labels or annotations to an app using Apps Manager, do the following:

1. Click the **Settings** tab of the desired app.
2. For **Labels**, provide a name and value pair. Or, select **Enter JSON** to enter the label in JSON.
3. For **Annotations**, provide a name and value pair. Or, select **Enter JSON** to enter the annotation in JSON.
4. To add more labels or annotations, click + next to the **Labels** or **Annotations** field.
5. Click **Update Metadata**.

## Manage a Service Instance

From the **Services** tab on the space page, you can bind or unbind apps, bind or unbind routes, view or change your service plan, manage service keys, and rename or delete your service instance.

For services that use on-demand brokers, the service broker will create, update, or delete the service instance in the background and notify you when it finishes.

## Bind an App

 **Note:** For services that use asynchronous bindings, Apps Manager displays the status of the service while the bind is still pending. Asynchronous bindings provide more flexibility for services that require additional time before returning a successful bind response.

1. From the space page **Services** tab, click the service instance you want to bind to an app.

docs-autoscale  
SERVICE: App Autoscaler PLAN: Standard

Overview Plan Settings Docs Manage

**Bound Apps**

docs-bindable-app ×

**Bound Routes**

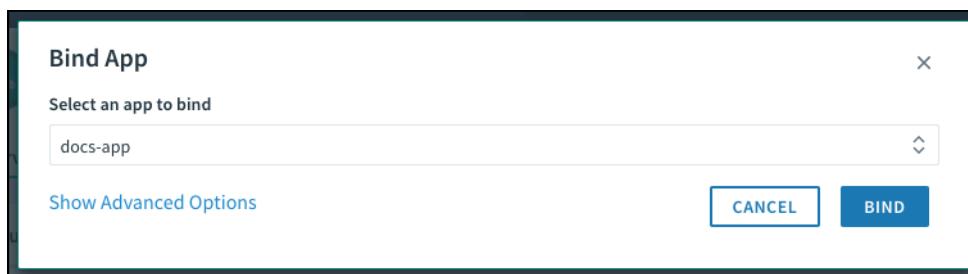
This service does not support route binding.

**Service Key Credentials**

CREATE SERVICE KEY

No Service Keys are associated with this Service

2. Click **Bind App**. A popup appears.
3. In the **Bind App** popup, select the app you want to bind to your service instance.



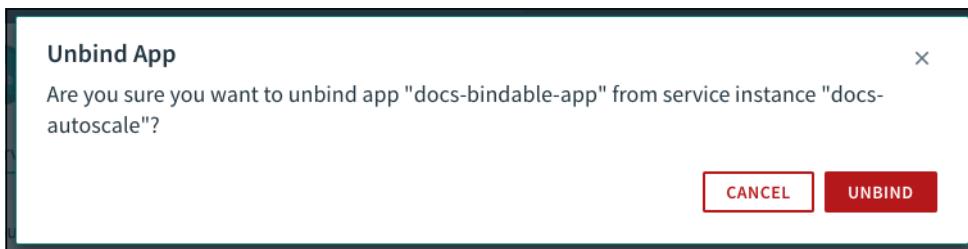
4. (Optional) To attach parameters to the binding, click **Show Advanced Options**. Under **Arbitrary Parameters**, enter any additional service-specific configuration.
5. Click **Bind**.

## Unbind an App

1. From the space page **Services** tab, click the service instance you want to unbind from an app.

The screenshot shows the 'docs-autoscale' service instance in the Pivotal Service Catalog. At the top, it displays the service name and its plan: 'SERVICE: App Autoscaler PLAN: Standard'. Below this, there are tabs for 'Overview', 'Plan', and 'Settings', with 'Overview' being the active tab. In the 'Overview' section, there are three main sections: 'Bound Apps', 'Bound Routes', and 'Service Key Credentials'. The 'Bound Apps' section contains a list with one item: 'docs-bindable-app', which has a red 'X' icon to its right. A 'BIND APP' button is located at the top right of this section. The 'Bound Routes' section contains a message: 'This service does not support route binding.' The 'Service Key Credentials' section contains a message: 'No Service Keys are associated with this Service' and a 'CREATE SERVICE KEY' button.

- Locate the app under **Bound Apps** and click the red × on the right. An **Unbind App** popup appears.



- Click **Unbind** to confirm.

## Share Service Instance

From the **Overview** tab, you can share your service instance across spaces. You can also view and delete service instances.

To share your service instance, do the following:

- From the **Overview** tab of your service instance, click **Share Service Instance**.

Home / fake-pivotal / space-0000 / ngfsgdsf

 **ngfsgdsf**  
SERVICE: MySQL for Pivotal Cloud Foundry v2 PLAN: db-small

Overview Plan Settings Docs

**Bound Apps**

No Apps have been bound to this Service

**Bound Routes**

This service does not support route binding.

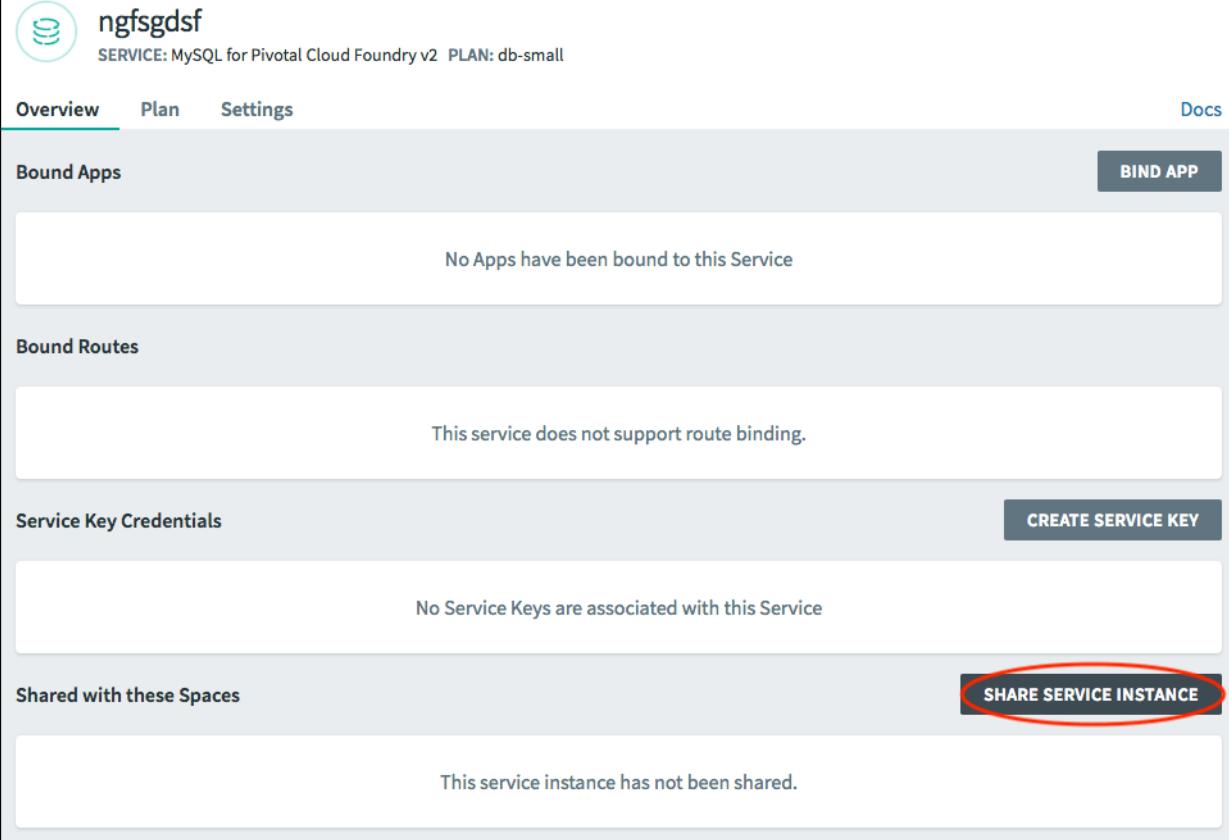
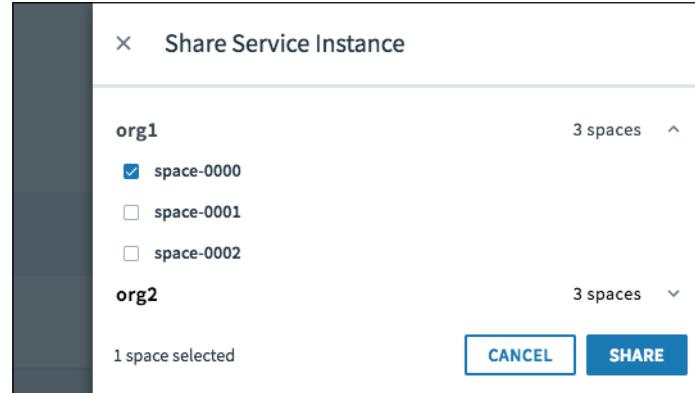
**Service Key Credentials**

No Service Keys are associated with this Service

**Shared with these Spaces**

This service instance has not been shared.

**SHARE SERVICE INSTANCE**

2. Select the spaces you want to share your service instance with.

3. Click **Share**.

## View or Change Your Service Plan

- From the space page **Services** tab, click the service instance you want to view or change the plan for.
- Click **Plan**.

**rails-blog-db**  
SERVICE: ElephantSQL PLAN: turtle

Overview Plan Settings Docs Support Manage

<b>turtle</b> free	<b>turtle free</b>
<b>panda</b> \$19.00/MONTH	<ul style="list-style-type: none"> <li>• Shared high performance cluster</li> <li>• 20 MB data</li> <li>• 4 concurrent connections</li> </ul>
<b>hippo</b> \$99.00/MONTH	
<b>elephant</b> \$499.00/MONTH	

3. Review your current plan information.
4. To change your plan, select a new plan from the list and click **Select This Plan** or **Upgrade Your Account**.

**Note:** Not all services support upgrading. If your service does not support upgrading, the service plan page only displays the selected plan.

## Rename or Delete Your Service Instance

1. From the space page **Services** tab, click the service instance you want to rename or delete.
2. Click **Settings**.

**rails-blog-db**  
SERVICE: ElephantSQL PLAN: Tiny Turtle

Overview Plan Settings Docs Support Manage

Service Instance Name: rails-blog-db **UPDATE** **CANCEL**

Configure Instance  
For a list of supported configuration parameters, see documentation for the particular service offering.

Name	Value	Enter JSON <input checked="" type="checkbox"/>
------	-------	------------------------------------------------

**DELETE SERVICE INSTANCE**  
This will permanently delete the service and all of its data.

- To change the service instance name, enter the new name and click **Update**.
- To add configuration parameters to the service instance, enter the parameters in the **Name** and **Value** fields and then click **Update**. Alternatively, enter your configuration parameters using the **Enter JSON** toggle and then click **Update**.
- To delete the service instance, click **Delete Service Instance**.

**Note:** The service broker supports creating, updating, and deleting service instances asynchronously. When the service broker completes one of these operations, a status banner appears in Apps Manager.

## Update Your User-Provided Service Instance

Follow the steps below to update an existing user-provided service instance. You can create a user-provided service instance from the Marketplace. For more information, see the [User-Provided Service Instances](#) topic.

1. Click **Configuration**. This tab only appears for user-provided service instances.

**Configuration**

Credential Parameters ⓘ (Optional)

Name	Value	Enter JSON <input checked="" type="checkbox"/>
Name	Value	+ <input type="button"/>

Syslog Drain Url ⓘ (Optional)

Route Service Url ⓘ (Optional)

**CANCEL** **UPDATE SERVICE**

2. Enter your Credential Parameters, Syslog Drain Url, and Route Service Url, and click **Update Service**.

## Manage Service Keys

On the space page, click **Services**, then click the service instance that you want to manage service keys for. This directs you to the service instance **Overview** page, where you can generate a new service key, get the credentials for a service key, and delete a service key.

**mysql**  
SERVICE: ClearDB MySQL Database PLAN: Spark DB

**Overview** **Plan** **Settings** **Docs** **Support** **Manage**

**Bound Apps**

- springpong
- spring-music

**Bind App**

**Bound Routes**

This service does not support route binding.

**Service Key Credentials**

- external-app

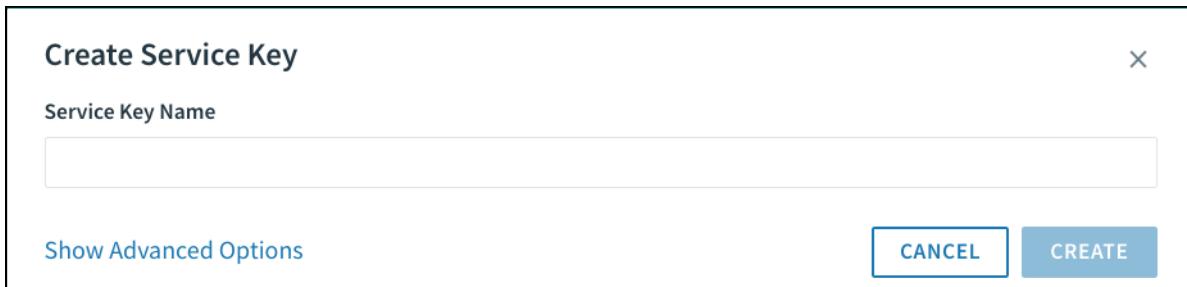
**Create Service Key**

## Generate a Service Key

Follow the steps below to generate a service key.

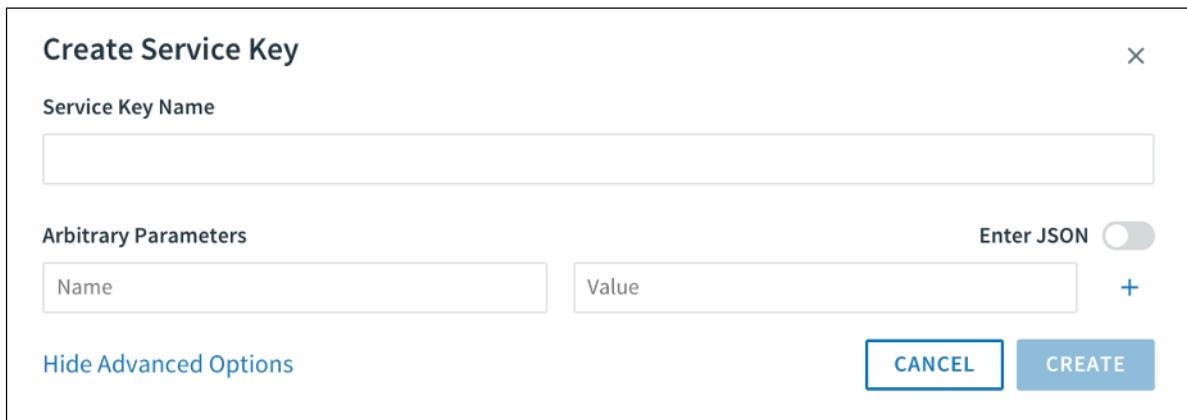
1. In the **Service Key Credentials** section, click **Create Service Key**.

2. Edit the **Service Key Name**.



The dialog box has a title bar "Create Service Key" with a close button "X". It contains a "Service Key Name" input field, a "Show Advanced Options" link, and two buttons: "CANCEL" and "CREATE".

3. (Optional) Click **Show Advanced Options**. Under **Arbitrary Parameters**, enter any additional service-specific configuration in the **Name** and **Value** fields.



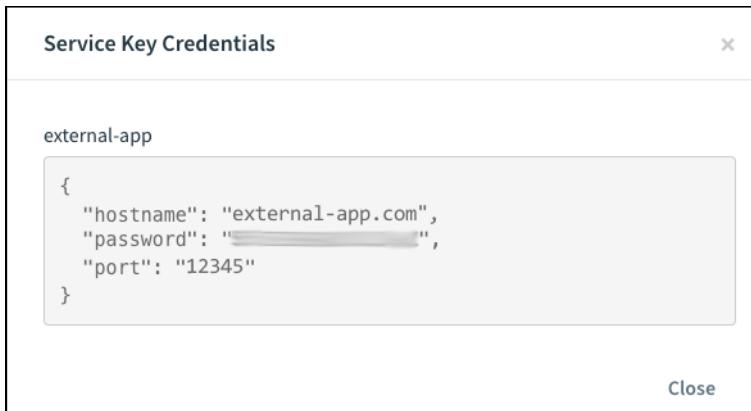
The dialog box includes an "Arbitrary Parameters" section with "Name" and "Value" input fields, an "Enter JSON" toggle switch, and a "+" button. It also features a "Hide Advanced Options" link and the same "CANCEL" and "CREATE" buttons as the first dialog.

4. Click **Create** to generate the service key.

## View Credentials for a Service Key

Follow the steps below to view the credentials for a service key.

1. To view the credentials for a particular service instance, click the service instance name under **Service Key Credentials**. The JSON object containing the credentials appears.



The dialog box shows a "Service Key Credentials" header and a "external-app" entry. The entry displays a JSON object:

```
{
 "hostname": "external-app.com",
 "password": "REDACTED",
 "port": "12345"
}
```

A "Close" button is at the bottom right.

2. Click **Close**.

## Delete Service Key

To delete a service key, click the red **x** next to the service instance name.

Service Key Credentials	CREATE SERVICE KEY
external-app	✖

## Manage Route Services

For more information about route services, see the [Route Services](#) topic.

You can bind a [new service instance](#) to a route when creating the instance in the Marketplace, or you can manage route services for an [existing service instance](#) on the service instance page.

### Bind a New Service Instance to a Route

Follow the steps below to bind a new service instance to a route.

**Configure Instance**

**Instance Name**

**Add to Space**

**Bind to App**  
 [do not bind]

**Bind to Route**  
 [do not bind] **CREATE ROUTE**

**SHOW ADVANCED OPTIONS** **CANCEL** **ADD**

1. Select the service from the Marketplace.
2. Under **Bind to Route**, either bind the service instance to an existing route or click **Create Route** to create a new custom route.

**Note:** You must choose a Marketplace service compatible with route services for the **Bind to Route** field to appear.

3. Complete the remaining fields and click **Add** to create the service instance.

### Bind an Existing Service Instance to a Route

Follow the steps below to bind an existing service instance to a route.

1. On the space page, click **Services**.
2. Click the service instance that you want to manage route services for.

**Note:** If the service is not compatible with route services, the text “This service does not support route binding” appears under **Bound Routes**.

3. To bind the service instance to a route, click **Bind Route**.



4. Select an existing route under **Select a route to bind** or enter a new route under **Create Custom Route**.

5. Click **Bind**.

To unbind a route from a service instance, click the red next to the name of the route under **Bound Routes**.

## Scaling an App Using App Autoscaler

This topic describes how Pivotal Cloud Foundry (PCF) users with the Space Developer role can set up and configure the App Autoscaler service in the Apps Manager UI to automatically scale apps based on rules that they set.

You can use the App Autoscaler command-line interface (CLI) plugin to configure App Autoscaler rules from the command line. For more information, see [Using the App Autoscaler CLI](#).

 **Note:** Space Managers, Space Auditors, and all Org roles do not have permission to use App Autoscaler. For help managing user roles, see [Managing User Accounts and Permissions Using the Apps Manager](#).

## App Autoscaler Overview

App Autoscaler is a marketplace service that helps control the cost of running apps while maintaining app performance.

To balance app performance and cost, Space Developers can use App Autoscaler to do the following:

- Configure rules that adjust instance counts based on metrics thresholds, such as CPU Usage
- Modify the maximum and minimum number of instances for an app, either manually or following a schedule

 **Breaking Change:** App Autoscaler relies on API endpoints from Loggregator's Log Cache. If you disable Log Cache, App Autoscaler will fail. For more information about Log Cache, see [Loggregator Introduces Log Cache](#).

## Set Up App Autoscaler

To use App Autoscaler, you must create an instance of the App Autoscaler service and bind it to any app you want to autoscale. You can do this using either the Apps Manager or from the Cloud Foundry Command Line Interface (cf CLI):

- **Apps Manager:**
  1. [Create an instance of the service](#).
  2. [Bind the service to an app](#).
- **cf CLI:**
  1. [Create an instance of the service](#).
  2. [Bind the service to an app](#).

To enable App Autoscaler for the app, do the following:

1. In Apps Manager, select an app from the space in which you created the App Autoscaler service.
2. Under **Processes and Instances**, enable **Autoscaling**.

**Processes and Instances**

**web**

Instances	Memory Allocated	Disk Allocated	SCALE	
2	1 GB	512 MB	<a href="#">SCALE</a>	
<input checked="" type="checkbox"/> Autoscaling		<a href="#">Manage Autoscaling</a>		
#	CPU	Memory	Disk	Uptime
0	0%	6.63 MB	51.07 MB	4 d 21 hr 58 min
1	0%	27.98 MB	51.07 MB	4 d 21 hr 56 min

- Click [Manage Autoscaling](#) to configure instance limits, scaling rules, and scheduled limit changes for the app. For more information, see [Configure Autoscaling for an App](#).

[x Manage Autoscaling](#)

[Download Autoscaler CLI](#)

**INSTANCE LIMITS**  
Currently in use: 2

Minimum	2	Maximum	5	<a href="#">APPLY CHANGES</a>
---------	---	---------	---	-------------------------------

**SCALING RULES**

Metric	Low	High
CPU Utilization	20%	80%

**SCHEDULED LIMITS**

Next: Mon June 11th, 2018 3:12:00 PM

**EVENT HISTORY**

Most Recent: Mon June 4, 2018 at 3:12 PM

Rule Applied: Scaling Limits set to 2 to 5 instances

[VIEW MORE](#)

## Manual Scaling Overrides Autoscaler

If you manually scale an app bound to an App Autoscaler service instance, the App Autoscaler instance automatically unbinds from that app, and the app scales to the manual setting. For more information, see [Managing Apps and Service Instances Using Apps Manager](#).

## Configure Autoscaling for an App

App Autoscaler keeps instance counts within an allowable range defined by minimum and maximum values, or *instance limits*.

Follow the procedures in the sections below to set any of the following:

- [Instance Limits](#)
- [Scaling Rules](#)
- [Scheduled Limit Changes](#)

## Instance Limits

Follow these steps to manually modify instance limits:

**Note:** You can also schedule changes to your instance limits for a specific date and time. For more information, see [Scheduled Limit Changes](#).

1. In Apps Manager, navigate to the [Overview](#) page for your app and click **Manage Autoscaling** under **Processes and Instances**.
2. In the **Instance Limits** section, set values for **Minimum** and **Maximum**.

The screenshot shows a modal window titled "Manage Autoscaling". Inside, there's a "Download Autoscaler CLI" link. Below it, the "INSTANCE LIMITS" section is shown with the text "Currently in use: 2". There are two input fields: "Minimum" set to 2 and "Maximum" set to 5. A large blue "APPLY CHANGES" button is at the bottom right.

3. Click **Apply Changes**.

## Scaling Rules

To keep your apps available without wasting resources, App Autoscaler increments or decrements instance counts based on how current metrics compare with configurable minimum and maximum thresholds.

### How App Autoscaler Determines When to Scale

Every 35 seconds, App Autoscaler makes a decision about whether to scale up, scale down, or keep the same number of instances. To make a scaling decision, App Autoscaler averages the values of a given metric for the most recent 120 seconds.

App Autoscaler scales apps as follows:

- Increment by one instance when any metric exceeds its maximum threshold
- Decrement by one instance only when all metrics fall below their minimum thresholds

For example, an app may have a maximum threshold of 200 milliseconds and a minimum threshold of 80 milliseconds for an HTTP latency metric. If HTTP latency averages 300 milliseconds over the past 120 seconds, App Autoscaler scales the app up one instance. If HTTP latency then averages 70 milliseconds over the next 120 second window and the app's other scaling metrics also fall below their minimum thresholds, App Autoscaler scales the app down one instance.

You can also set a maximum and minimum number of instances. For example, if an app exceeds the maximum threshold of a given metric, but the number of instances is already at the maximum number of allowed instances, App Autoscaler does not scale up the app. For more information about setting a maximum and minimum number of instances, see [Instance Limits](#).

### Scaling Rule Metrics

The table below lists the metrics that you can base App Autoscaler rules on:

Metric	Description
CPU Utilization	Average CPU percentage for all instances of the app.
Container Memory Utilization	Average memory percentage for all instances of the app.
HTTP Throughput	Total HTTP requests per second (divided by the total number of app instances).
HTTP Latency	Average latency of apps response to HTTP requests. This does not include Gorouter processing time or other network latency. Average is calculated on the middle 99% or middle 95% of all HTTP requests.

RabbitMQ Depth	The queue length of the specified queue.
Custom	Users can configure their apps to emit custom metrics out of the <a href="#">Loggregator Firehose</a> . For steps on how to set up your apps to emit custom metrics, see <a href="#">Registering Custom App Metrics</a> with Metric Registrar.
Compare	One custom metric divided by another. When you add a rule, the <b>Metric</b> field is the dividend and the <b>Comparison Metric</b> field is the divisor.

## Add a Scaling Rule

1. In the **Manage Autoscaling** pane, click **Edit** next to **Scaling Rules**. The **Edit Scaling Rules** pane appears.
2. Click **Add Rule**.
3. In the **Select type** dropdown, select the metric for the new scaling rule.

The screenshot shows the 'Edit Scaling Rules' interface. At the top is a header with a back arrow and the title 'Edit Scaling Rules'. Below the header is a descriptive text block: 'Apps scale by 1 instance per event. Apps will scale up when any metric maximum is met and scale down only when all metric minimums are met.' Underneath this text is a dropdown menu labeled 'Select type' with a red 'x' icon to its right. Below the dropdown is a blue 'ADD RULE' button. At the bottom of the pane are two buttons: 'CANCEL' and 'SAVE'.

4. Set the minimum and maximum thresholds for the metric using the table in the [Scaling Rule Metrics](#) section above as a guide.
5. Select or fill in any other fields that appear under the threshold fields:
  - o If you are adding an **HTTP Latency** rule, configure **Percent of traffic to apply**.
  - o If you are adding a **RabbitMQ** depth rule, provide the name of the queue to measure.
  - o If you are adding a **Custom** rule, enter your custom **Metric**.
  - o If you are adding a **Compare** rule, enter values in the **Metric** and **Comparison Metric** fields.
6. Click **Save**.

## Delete a Scaling Rule

1. Click the **x** icon next to the rule you want to delete.
2. Click **Save**.

## Scheduled Limit Changes

Because app demand often follows a weekly, daily, or hourly schedule, you can schedule App Autoscaler to change the allowable instance range to track expected surges or quiet periods.

### Create or Modify a Scheduled Limit Change

1. Click **Edit** next to **Scheduled Limits**.
2. Click **Add New** to add a new scheduled limit or select an existing entry to modify by clicking **EDIT** next to the entry.
3. Edit the following values:
  - o **Date and Time (local)**: Set the date and time of the change.
  - o **Repeat (Optional)**: Set the day of the week for which you want to repeat the change.
  - o **Min and Max**: Set the allowable range within which App Autoscaler can change the instance count for an app.

< Scheduled Limit Changes: rails-docs-sample

**Configure**

Date	Time (local)			
June 4, 2018	4	:	32	PM

**Repeat (Optional)**

Su	Mo	Tu	We	Th	Fr	Sa
<input type="checkbox"/>						

**Scheduled**

**ADD NEW**

Date	Time	Repeat	Min	Max
Jun 11, 2018	3:12 PM	Mo	2	5

**SAVE**

4. Click **Save**.

To delete an existing entry, click the x icon next to the entry you want to delete.

### Example: Scale Down for the Weekend

To schedule an app to scale down for a weekend, you can enter two rules as follows:

1. Scale down to a single instance on Friday evening:

- o Date and Time (local): Dec 2, 2018 and 7:00 PM
- o Repeat (Optional): Fr
- o Min and Max: 1 and 1

2. Increase instances to between 3 and 5 on Monday morning:

- o Date and Time (local): Dec 5, 2018 and 7:00 AM
- o Repeat (Optional): M
- o Min and Max: 3 and 5

## App Autoscaler Events and Notifications

App Autoscaler logs all autoscaling events.

### View Event History

To view all autoscaling events in the past 24 hours, click **View More** in the **Event History** section of the **Manage Autoscaling** pane.

<b>EVENT HISTORY</b>	
Most Recent:	Mon June 4, 2018 at 3:12 PM
Rule Applied: Scaling Limits set to 2 to 5 instances	
<a href="#">VIEW MORE</a>	

### Manage App Autoscaler Notifications

Autoscaler emails or texts its event notifications to all space users by default.

To subscribe or unsubscribe from autoscaling event notifications, do the following:

1. Navigate to the **Manage Notifications** page of PCF.

 **Note:** If installed, Notifications Management should be available at <https://notifications-ui.YOUR-SYSTEM-DOMAIN/preferences>.

2. Choose which notifications you want to receive from App Autoscaler:

## Manage Notifications

You will receive email notifications from checked items

- Cloud Foundry Autoscaling Service
  - Scaling Down
  - Manual Scaling Detected
  - Maximum Instance Limit Reached
  - Quota Limit Reached
  - Scaling Up

## Using the App Autoscaler CLI

This topic explains how to use the App Autoscaler command-line interface (CLI).

The App Autoscaler automatically scales Cloud Foundry apps in response to demand. The App Autoscaler CLI lets you control the App Autoscaler from your local command line by extending the Cloud Foundry command-line interface (cf CLI).

### Install the App Autoscaler CLI Plugin

Before you can run App Autoscaler CLI commands on your local machine, you must install the App Autoscaler CLI plugin.

To install the plugin, do the following:

1. Download the plugin from [Pivotal Network](#).

 **Note:** Ensure that you download the App Autoscaler CLI plugin v2.0 or later. In addition, the plugin requires Pivotal Application Service v2.2 or later..

2. To install the App Autoscaler CLI plugin, run the following command:

```
cf install-plugin LOCATION-OF-PLUGIN
```

Where `LOCATION-OF-PLUGIN` is the path to the binary file you downloaded from Pivotal Network. For example:

```
$ cf install-plugin ~/Downloads/autoscaler-for-pcf-cliplugin-macosx64-binary-2.0.91
```

For more information about installing cf CLI plugins, see [Installing a Plugin](#).

### View Apps

Run `cf autoscaling-apps` to view all the apps that are bound to an autoscaler service instance in a space, their instance limits, and whether or not they are enabled.

```
$ cf autoscaling-apps

Presenting autoscaler apps in org my-org / my-space autoscaling as user
Name Guid Enabled Min Instances Max Instances
test-app guid true 1 4
test-app-2 guid-2 false 10 40
OK
```

### Enable Autoscaling

Run `cf enable-autoscaling APP-NAME` to enable autoscaling on your app. Replace `APP-NAME` with the name of your app.

```
$ cf enable-autoscaling test-app-2

Enabled autoscaling for app test-app-2 for org my-org / my-space testing as admin
OK
```

### Disable Autoscaling

Run `cf disable-autoscaling APP-NAME` to disable autoscaling on your app. Replace `APP-NAME` with the name of your app.

```
$ cf disable-autoscaling test-app
```

Disabled autoscaling for app test-app for org my-org / my-space testing as admin  
OK

## Update Instance Limits

Run `cf update-autoscaling-limits APP-NAME MIN-INSTANCE-LIMIT MAX-INSTANCE-LIMIT` to update the upper and lower app instance limits. Replace `APP-NAME` with the name of your app. Replace `MIN-INSTANCE-LIMIT` with the minimum number of apps, and `MAX-INSTANCE-LIMIT` with the maximum number of apps.

```
$ cf update-autoscaling-limits test-app 10 40
```

Updated autoscaling instance limits for app test-app for org my-org / my-space testing as admin  
OK

## View Rules

Run `cf autoscaling-rules APP-NAME` to view the rules that the App Autoscaler uses to determine when to scale your app. Replace `APP-NAME` with the name of your app.

```
$ cf autoscaling-rules test-app
```

Presenting autoscaler rules for app test-app for org my-org / my-space autoscaling as user  
Rule Guid    Rule Type    Rule Sub Type    Min Threshold    Max Threshold  
guid            cpu                10            20  
guid-2        http\_throughput        20            30  
OK

## Create a Rule

Run `create-autoscaling-rule APP-NAME RULE-TYPE MIN-THRESHOLD MAX-THRESHOLD [--subtype SUBTYPE] [--metric METRIC] [--comparison-metric COMPARISON-METRIC]`

to create a new autoscaling rule.

Replace the placeholders as follows:

- `APP-NAME` is the name of your app.
- `RULE-TYPE` is the type of your scaling rule.
- `MIN-THRESHOLD` is the minimum threshold for the metric.
- `MAX-THRESHOLD` is the maximum threshold for the metric.

You can use the following command options:

- `--metric`, `-m` is the metric for a Custom rule.
- `--comparison-metric`, `-c` is the comparison metric for a Compare rule.
- `--subtype`, `-s` is the rule subtype.

For example:

```
$ cf create-autoscaling-rule test-app http_latency 10 20 -s avg_99th
```

Created autoscaler rule for app test-app for org my-org / space my-space as user  
Rule Guid    Rule Type    Rule Sub Type    Min Threshold    Max Threshold  
guid-3        http\_latency        avg\_99th    10            20

```
$ cf create-autoscaling-rule test-app custom 9380 9381 --metric jvm.classes.loaded
Created autoscaler rule for app test-app in org my-org / space my-space as user
OK
Guid Type Metric Sub Type Min Threshold Max Threshold
guid-7 custom jvm.classes.loaded 9380.00 9381.00
```

```
$ cf create-autoscaling-rule test-app compare .79 .8 --metric jvm.memory.used --comparison-metric jvm.memory.max
Created autoscaler rule for app test-app in org my-org / space my-space as user
OK
Guid Type Metric Sub Type Min Threshold Max Threshold
guid-8 compare jvm.memory.used / jvm.memory.max 0.79 0.80
```

## Valid Rule Types and Subtypes

For a list of valid types and subtypes, see the following:

- type `CPU`
- type `memory`
- type `http_throughput`
- type `http_latency`
- type `rabbit-mq`
  - queue\_name `YOUR-QUEUE-NAME`
- type `custom`
  - metric `METRIC-NAME`
- type `compare`
  - metric `METRIC-NAME`
  - comparison\_metric `METRIC-NAME`

## Delete a Rule

Run `cf delete-autoscaling-rule APP-NAME RULE-GUID [-force]` to delete a single autoscaling rule. Replace `APP-NAME` with the name of your app, and replace `RULE-GUID` with the GUID.

```
$ cf delete-autoscaling-rule test-app guid-2
Really delete rule guid-2 for app test-app?> [yN]:y
Deleted rule guid-2 for autoscaler app test-app for org my-org / space my-space as admin
OK
```

## Delete All Rules

Run `cf delete-autoscaling-rules APP-NAME [-force]` to delete all autoscaling rules. Replace `APP-NAME` with the name of your app.

```
$ cf delete-autoscaling-rules test-app
Really delete ALL rules for app test-app?> [yN]:y
Deleted rules for autoscaler app test-app for org my-org / space my-space as admin
OK
```

## View Autoscaling Events

Run `cf autoscaling-events APP-NAME` to view recent events related to autoscaling for your app. Replace `APP-NAME` with the name of your app.

```
$ cf autoscaling-events test-app
Time Description
2032-01-01T00:00:00Z Scaled up from 3 to 4 instances. Current cpu of 20 is above upper threshold of 10.
```

## View Autoscaler Scheduled Instance Limit Changes

Run `cf autoscaling-scls APP-NAME` to view all scheduled instance limit changes. Replace `APP-NAME` with the name of your app.

For example:

```
$ cf autoscaling-scls test-app
Guid First Execution Min Instances Max Instances Recurrence
guid-5 2018-06-12T22:00:00Z 0 1 Mo,Tu,We,Th,Fr
```

## Create Autoscaler Scheduled Instance Limit Change

Run `create-autoscaling-slc APP-NAME DATE-TIME MIN-INSTANCES MAX-INSTANCES [--recurrence RECURRENCE]` to create a new scheduled instance limit change.

Replace the placeholders as follows:

- `APP-NAME` is the name of your app.
- `DATE-TIME` is the date and time of the change.
- `MIN-INSTANCES` and `MAX-INSTANCES` are the minimum and maximum values of the range within which App Autoscaler can change the instance count for an app.
- `RECURRENCE` (optional) is the day of the week for which you want to repeat the change.

For example:

```
$ cf create-autoscaling-slc test-app 2018-06-14T15:00:00Z 1 2 --recurrence Sa
Created scheduled autoscaler instance limit change for app test-app in org my-org / space my-space as user
OK
Guid First Execution Min Instances Max Instances Recurrence
guid-6 2018-06-14T15:00:00Z 1 2 Sa
```

## Delete Autoscaler Scheduled Instance Limit Change

Run `delete-autoscaling-slc APP-NAME SLC-GUID [-force]` to delete a scheduled instance limit change. Replace `APP-NAME` with the name of your app and `SLC-GUID` with the GUID of your scheduled instance limit change.

For example:

```
$ cf delete-autoscaling-slc test-app slc-guid
Really delete scheduled limit change slc-guid for app test-app?> [yN]:y
Deleted scheduled limit change slc-guid for app test-app in org my-org / space my-space as user
OK
```

## Configure with a Manifest

Run `configure-autoscaling APP-NAME MANIFEST-FILE-PATH` to use a service manifest to configure your rules, add instance limits, and set scheduled limit changes at the same time. Replace `APP-NAME` with the name of your app, and replace `MANIFEST-FILE-PATH` with the path and name of your Autoscaler manifest.

An example manifest:

```

instance_limits:
 min: 1
 max: 2
rules:
- rule_type: "http_latency"
 rule_sub_type: "avg_99th"
 threshold:
 min: 10
 max: 20
scheduled_limit_changes:
- recurrence: 10
 executes_at: "2032-01-01T00:00:00Z"
 instance_limits:
 min: 10
 max: 20
```

\$ cf configure-autoscaling test-app autoscaler-manifest.yml

Setting autoscaler settings for app test-app for org my-org / space my-space as user  
OK

A `rules` block must be present in your Autoscaler manifest. If your app does not require any rules changes, include an empty block:

```

instance_limits:
 min: 1
 max: 1
rules: []
scheduled_limit_changes:
- recurrence: 365
 executes_at: "2032-01-01T00:00:00Z"
 instance_limits:
 min: 0
 max: 0
```

A `scheduled_limit_changes` block must be present in your Autoscaler manifest. If your app does not require any scheduled instance limit changes, include an empty block:

```

instance_limits:
 min: 1
 max: 2
rules:
- rule_type: "http_latency"
 rule_sub_type: "avg_99th"
 threshold:
 min: 10
 max: 20
scheduled_limit_changes: []
```

## App Autoscaler CLI Known Issues

The App Autoscaler CLI has the following known issues:

- The CLI returns an error message if you have more than one instance of the App Autoscaler service running in the same space.
  - To prevent this error, delete all but one App Autoscaler service instance from any space that the App Autoscaler service runs in.
- The CLI may output odd characters in Windows shells that do not support text color.
  - To prevent this error, run `SET CF_COLOR=false` in your Windows shell pane before you run App Autoscaler CLI commands.
  - Note that some Windows shells do not support the `CF_COLOR` setting.

## Viewing ASGs in Apps Manager

### About ASGs

Application Security Groups (ASGs) are a collections of egress rules that specify the protocols, ports, and IP address ranges where app or task instances send traffic. The platform sets up rules to filter and log outbound network traffic from app and task instances. ASGs apply to both buildpack-based and Docker-based apps and tasks.

When apps or tasks begin staging, they need traffic rules permissive enough to allow them to pull resources from the network. After an app or task is running, the traffic rules can be more restrictive and secure. To distinguish between these two security requirements, administrators can define one ASG for app and task staging, and another for app and task runtime. For more information about staging and running apps, see [Application Container Lifecycle](#).

To provide granular control when securing a deployment, an administrator can assign ASGs to apply to all app and task instances for the entire deployment, or assign ASGs to spaces to apply only to apps and tasks in a particular space.

Only admin users can create and modify ASGs. For information about creating and configuring ASGs, see [Application Security Groups](#).

### Displaying ASGs for a Space

To view the ASGs associated with a space, perform the following steps.

1. Log in to Apps Manager.
2. From the **Home** page, select the **Org** that contains the space you want to view.
3. Select the **Space** you want to view.
4. Click on the **Settings** tab.
5. In the **Security Groups** section, Apps Manager displays ASGs associated with the selected space.
6. Click on an ASG to expand its egress rules.

The screenshot shows the 'Security Groups' section of the Apps Manager settings for a specific space. On the left, there's a summary of what Security Groups are: 'A collection of egress rules that specify one or more individual protocols, ports, and destinations.' Below this, there's a link to 'Learn more'. The main area lists ASGs with their status (running, staging) and detailed egress rules:

- public\_networks**: running, staging
  - destination: 0.0.0.0-9.255.255.255  
protocol: all
  - destination: 11.0.0.0-169.253.255.255  
protocol: all
  - destination: 169.255.0.0-172.15.255.255  
protocol: all
  - destination: 172.32.0.0-192.167.255.255  
protocol: all
  - destination: 192.169.0.0-255.255.255.255  
protocol: all
- dns**: running, staging
- sshfs-service**: running
- p-mysql**: running



## Configuring Spring Boot Actuator Endpoints for Apps Manager

The Apps Manager UI supports several production-ready endpoints from Spring Boot Actuator. This topic describes the Actuator endpoints and how you can configure your app to display data from the endpoints in Apps Manager.

For more information about Spring Boot Actuator, see the [Spring Boot Actuator documentation](#).

 **Note:** This feature requires Spring Boot v1.5 or later.

### Overview

The Apps Manager integration with Spring Boot does not use the standard Spring Boot Actuators. Instead, it uses a specific set of actuators that are secured using the Space Developer role for the space that the application runs in. Authentication and authorization are automatically delegated to the [Cloud Controller](#) and the [User Account and Authentication](#) server without any configuration from the user.

By default, actuators are secure and cannot be accessed without explicit configuration by the user, even if [Spring Security](#) is not included. This allows users to take advantage of the Spring Boot Apps Manager integration without accidentally exposing their actuators without security.

### Actuator Endpoints

The table below describes the Spring Boot Actuator endpoints supported by Apps Manager. To integrate these endpoints with Apps Manager, you must first [Activate Spring Boot Actuator for Your App](#).

Endpoint	About
/info	<ul style="list-style-type: none"> <li><b>Description:</b> Exposes details about app environment, git, and build. To send build and Git information to this endpoint, see <a href="#">Configure the Info Actuator</a>.</li> <li><b>How to use in Apps Man:</b> See <a href="#">View Build and Git Information for Your App</a>.</li> </ul>
/health	<ul style="list-style-type: none"> <li><b>Description:</b> Shows health status or detailed health information over a secure connection. Spring Boot Actuator includes the auto-configured health indicators specified in the <a href="#">Auto-configured HealthIndicators</a> section of the Spring Boot documentation. If you want to write custom health indicators, see the <a href="#">Writing custom HealthIndicators</a> section of the Spring Boot documentation.</li> <li><b>How to use in Apps Man:</b> See <a href="#">View App Health</a>.</li> </ul>
/loggers	<ul style="list-style-type: none"> <li><b>Description:</b> Lists and allows modification of the levels of the loggers in an app.</li> <li><b>How to use in Apps Man:</b> See <a href="#">Manage Log Levels</a>.</li> </ul>
/dump	<ul style="list-style-type: none"> <li><b>Description:</b> Generates a thread dump.</li> <li><b>How to use in Apps Man:</b> See <a href="#">View Thread Dump</a>.</li> </ul>
/trace	<ul style="list-style-type: none"> <li><b>Description:</b> Displays trace information from your app for each of the last 100 HTTP requests. For more information, see the <a href="#">Tracing</a> section of the Spring Boot documentation.</li> <li><b>How to use in Apps Man:</b> See <a href="#">View Request Traces</a>.</li> </ul>
/heapdump	<ul style="list-style-type: none"> <li><b>Description:</b> Generates a heap dump and provides a compressed file containing the results.</li> <li><b>How to use in Apps Man:</b> See <a href="#">Download Heap Dump</a>.</li> </ul>
/mappings	<ul style="list-style-type: none"> <li><b>Description:</b> Displays the endpoints an app serves and other related details.</li> <li><b>How to use in Apps Man:</b> See <a href="#">View Mappings</a>.</li> </ul>

### Activate Spring Boot Actuator for Your App

You must add a `spring-boot-starter-actuator` dependency to your app project for the production-ready HTTP endpoints to return values. For more information,

see the [Enabling production-ready features](#) section of the Spring Boot documentation.

1. Follow the instructions below that correspond to your project type.

- o **Maven:** If you use Maven, add the following to your project:

```
<dependencies>
 <dependency>
 <groupId>org.springframework.boot</groupId>
 <artifactId>spring-boot-starter-actuator</artifactId>
 </dependency>
</dependencies>
```

- o **Gradle:** If you use Gradle, add the following to your project:

```
dependencies {
 compile("org.springframework.boot:spring-boot-starter-actuator")
}
```

2. If you use self-signed certificates in your PCF deployment for UAA or the Cloud Controller, specify in your `application.properties` file to skip SSL validation:

```
management.cloudfoundry.skip-ssl-validation=true
```

See [Cloud Foundry support](#) in the Spring Boot Actuator documentation for more information.

## Configure the Info Actuator

The `/info` endpoint provides information about the project build for your app, as well as its git details.

### Add Build Information

To add build information to the `/info` endpoint, follow the instructions below that correspond to your project type.

#### Maven

Add the following to your app project:

```
<build>
 <plugins>
 <plugin>
 <groupId>org.springframework.boot</groupId>
 <artifactId>spring-boot-maven-plugin</artifactId>
 <version>1.4.2.RELEASE</version>
 <executions>
 <execution>
 <goals>
 <goal>build-info</goal>
 </goals>
 </execution>
 </executions>
 </plugin>
 </plugins>
</build>
```

#### Gradle

Add the following to your app project:

```
springBoot {
 buildInfo()
}
```

## Add Git Information

To add git information to the `/info` endpoint, follow these instructions:

1. Add the following property to your `application.properties` file:

```
management.info.git.mode=full
```

2. Follow the instructions below that correspond to your project type.

### Maven

Add the following plugin to your project:

```
<build>
 <plugins>
 <plugin>
 <groupId>pl.project13.maven</groupId>
 <artifactId>git-commit-id-plugin</artifactId>
 </plugin>
 </plugins>
</build>
```

### Gradle

Add the following plugin to your project:

```
plugins {
 id "com.gorylenko.gradle-git-properties" version "1.4.17"
}
```

## Using Spring Boot Actuators with Apps Manager

This document describes how to view and manage app information from Spring Boot Actuator in Apps Manager.

### Prerequisites

The Apps Manager integration with Spring Boot Actuator requires the following:

- A PCF user with the `SpaceDeveloper` role. See [App Space Roles](#).
- Spring Boot v1.5 or later.
- Completing the procedures in [Configure Spring Boot Actuator Endpoints for Apps Manager](#).

After you configure your app, Apps Manager displays the Spring Boot logo next to the name of your app on the app page:

The screenshot shows the Apps Manager interface for the 'hello-spring-cloud' application. At the top, there's a header with the app name 'APP hello-spring-cloud'. To the right of the name are three small icons: a grey square, a blue square with a white circle, and a green circle with a white dot. Next to the green circle is the text 'Running'. Below the header is a navigation bar with tabs: 'Overview' (which is underlined), 'Services', 'Route (1)', 'Logs', 'Tasks', 'Trace', 'Threads', and 'Settings'. The 'Overview' tab is currently selected.

### View Build and Git Information for Your App

To view the data that your app sends to its `/info` Actuator endpoint, select the **Settings** tab:

**App Name**  **UPDATE** **CANCEL**

---

<b>Info</b>	<pre>Buildpack: java_buildpack_offline Start Command: CALCULATED_MEMORY=\$((\$PWD/.java-buildpack/open_jdk_jre/bin/java-buildpack-memory-calculator-2.0.2_RELEASE -memorySizes=metaspace:64m..,stack:228k.. -memoryWeights=heap:65,metaspace:10,native:15,stack:10 -memoryInitials=heap:100%,metaspace:100% -stackThreads=300 -totMemory=\$MEMORY_LIMIT) &amp;&amp; JAVA_OPTS="-Djava.io.tmpdir=\$TMPDIR -XX:OnOutOfMemoryError=\$PWD/.java-buildpack/open_jdk_jre/bin/killjava.sh \$CALCULATED_MEMORY -Djavax.net.ssl.trustStore=\$PWD/.java-buildpack/container_certificate_trust_store/truststore.jks -Djavax.net.ssl.trustStorePassword=java-buildpack-trust-store-password" &amp;&amp; SERVER_PORT=\$PORT eval exec \$PWD/.java-buildpack/open_jdk_jre/bin/java \$JAVA_OPTS -cp \$PWD/. org.springframework.boot.loader.JarLauncher Stack: cflinuxfs2 (Cloud Foundry Linux-based filesystem) Health check type: port</pre>
-------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

---

<b>Spring Info</b>	<b>VIEW RAW JSON</b>
--------------------	----------------------

---

<b>Git</b>	SHA: <b>c91dff0</b> Message: "Skip SSL validation [#137189187]" Date: 01/06/17 08:50PM UTC Email: jberney@users.noreply.github.com Remote: git@github.com:pivotal-cf/spring-actuator-acceptance.git Time: 04/21/17 09:03PM UTC Name: User Email: Pivotal Default
------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

---

<b>User Provided Environment Variables</b>	<b>REVEAL USER PROVIDED ENV VARS</b>
--------------------------------------------	--------------------------------------

---

<b>Environment Variables</b> Defined by the runtime and buildpack. <a href="#">Learn more</a>	<b>REVEAL ENV VARS</b>
--------------------------------------------------------------------------------------------------	------------------------

---

<b>Security Groups</b> A collection of egress rules that specify one or more individual protocols, ports, and destinations. <a href="#">Learn more</a>	> default_security_group running, staging
-----------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------

---

<b>Delete App</b> This will permanently delete the app and all of its data.	<b>DELETE APP</b>
--------------------------------------------------------------------------------	-------------------

In the upper right of the app page, Apps Manager also displays the SHA of your app code repository from the latest build:



## View App Health

To view the health-check data that your app sends to its `/health` Actuator endpoints, select the **Overview** tab and click an instance under the **Instances** section:

#	App Health	CPU	Memory	Disk	Uptime
0	Up	0%	209.97 MB	137.75 MB	1 min

**Health Check**

```

status: UP
diskSpace
status: UP
free: 912412672
threshold: 10485760
total: 1056858112
mySql
status: UP
database: MySQL
timestamp: Thu May 11 18:44:46 UTC 2017
postgres
status: UP
database: Postgres

```

## View Thread Dump

To trigger and view a thread dump from your app to its `/dump` Actuator endpoint, select the **Threads** tab and click **Refresh**.

APP amjs-test-spring-app ■ C ● Running

VIEW APP c91dff0

Buildpack: java\_buildpack\_offline

Overview Services Route (1) Logs Tasks Trace **Threads** Settings

Last refresh: 05/8/17 16:48 PM REFRESH

Instance 0 Show ALL DOWNLOAD

**container-0** TIMED\_WAITING

```

"container-0" #15
java.lang.Thread.State: TIMED_WAITING (sleeping)
at java.lang.Thread.sleep(Native Method)
at org.apache.catalina.core.StandardServer.await(StandardServer.java:427)
at org.springframework.boot.context.embedded.tomcat.TomcatEmbeddedServletContainer$1.run(TomcatEmbeddedServletContainer.java:100)

```

Locked ownable synchronizers:
- None

> ContainerBackgroundProcessor[StandardEngine[Tomcat]] TIMED\_WAITING

> DestroyJavaVM RUNNABLE

> Finalizer WAITING

> http-nio-8080-Acceptor-0 RUNNABLE

You can click each thread to expand and view its details. You can also modify which threads appear on the page using the **Instance** and **Show** drop-down menus.

## View Request Traces

To retrieve and view tracing information from the `/trace` Actuator endpoint of your app, select the **Trace** tab and click **Refresh**.

APP amjs-test-spring-app ■ C ● Running

[VIEW APP](#) Git: c91dff0

Overview Services Route (1) Logs Tasks **Trace** Threads Settings Buildpack: java\_buildpack\_offline

Instance ALL  Hide Pivotal Apps Manager Requests Last refresh: 05/8/17 16:39 PM REFRESH

16:39:17.939 200 OPTIONS /cloudfoundryapplication/trace  
0ms

```

Request:
host: amjs-test-spring-app.apps.oogie-boogie.gcp.appspot.cf-app.com
user-agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_12_1) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/57.0.2987.133 Safari.
accept: */*
accept-encoding: gzip, deflate, sdch, br
accept-language: en-US,en;q=0.8
access-control-request-headers: authorization,x-cf-app-instance
access-control-request-method: GET
origin: https://apps.sys.oogie-boogie.gcp.appspot.cf-app.com
via: 1.1 google
x-b3-spanid: bbf508bcdd3118db
x-b3-traceid: bbf508bcdd3118db
x-cf-applicationid: 2d2028b0-33ab-4e9a-b575-d63a6f2f8582
x-cf-instanceid: 29f8bf49-5b12-4131-475e-68360548cb6a
x-cloud-trace-context: 39945641373be76ef67beb185e09575/10770927836782894347
x-forwarded-for: 209.234.137.222, 35.186.230.30
x-forwarded-proto: https
x-request-start: 1494286757937
x-vcap-request-id: 75012830-8a64-432e-792b-d88ba0cb55e3
connection: close
Response:
X-Application-Context: amjs-test-spring-app:cloud:0
Access-Control-Allow-Origin: *
Vary: Origin
Access-Control-Allow-Methods: GET,POST
Access-Control-Allow-Headers: authorization, x-cf-app-instance
Allow: GET, HEAD, POST, PUT, DELETE, OPTIONS, PATCH
status: 200

```

> 16:39:14.845 200 GET /cloudfoundryapplication/info  
> 16:39:14.842 200 GET /cloudfoundryapplication/health  
> 16:39:14.734 200 OPTIONS /cloudfoundryapplication/info  
> 16:39:14.732 200 OPTIONS /cloudfoundryapplication/health  
> 16:39:14.653 200 GET /cloudfoundryapplication  
> 16:39:14.542 200 OPTIONS /cloudfoundryapplication  
> 16:38:50.619 200 GET /favicon.ico  
> 16:38:50.359 404 GET /

This page displays the last 100 requests from your app. You can click each individual request to expand and view its trace details. You can modify which requests appear on the page using the **Instance** drop-down menu.

By default, the **Trace** tab does not show requests and responses from Apps Manager polling app instances for data. To include these requests, clear the **Hide Pivotal Apps Manager Requests** checkbox next to the **Instance** drop-down menu.

## Download Heap Dump

To trigger and view a heap dump from your app to its `/heapdump` endpoint, select the settings drop-down menu for an instance of your app and click **Heap Dump**. This downloads a `.zip` file.

#	App Health	CPU	Memory	Disk	Uptime
> 0	Up	27%	206.21 MB	137.75 MB	1 hr 11 min

## View Mappings

To view a collated list of the endpoints an app serves, select the **Settings** tab and click **View Mappings**.

Overview Services Route (1) Logs Tasks Trace Threads **Settings**

**App Name** docs-spring-app **UPDATE** **CANCEL**

**Info**

Buildpack: java\_buildpack\_offline  
Start Command:  
Stack: cflinuxfs2 (Cloud Foundry Linux-based filesystem)  
Health check type: port

**Metrics Forwarder**   
Metrics Forwarder allows applications to emit metrics into loggregator and consume those metrics from the firehose.

**Spring Info**

SHA: [c91dff0](#) **VIEW RAW JSON**  
Message: "Skip SSL validation [#137189187]"  
Date: 01/06/17 08:50PM UTC  
Email: [jberney@users.noreply.github.com](mailto:jberney@users.noreply.github.com)  
Remote: <git@github.com:pivotal-cf/spring-actuator-acceptance.git>  
Time: 04/21/17 09:03PM UTC  
Name:  
User Email: Pivotal Default

**Mappings** **VIEW MAPPINGS**

## Manage Log Levels

Spring Boot apps include *loggers* for many provided and user components of the app. You can set the log level for each logger in Apps Manager.

To view the **Configure Logging Levels** screen, select the **Logs** tab and click **Configure Logging Levels**.

Overview Services Route (1) **Logs** Tasks Trace Threads Settings Git: c91dff0 Buildpack: java\_buildpack\_offline

Logs CONFIGURE LOGGING LEVELS

```
7210 4027 037c ab5a4b6a7ca Response_time:0.045521239 app_2af_b9552ecf-100_4030-3ebd-927177f858423 app_id:ca7... v
x b3_traceid:"31af4c2016156a89" x b3_spanid:"31af4c2016156a89" x b3_parentspanid:"-"
```

Apps Manager displays the default log level for each logger in gray.

**Configure Logging Levels** X

Filter Loggers 418 / 418

LOGGER	OFF	FATAL	ERROR	WARN	INFO	DEBUG	TRACE
io	●	●	●	●	●	●	●
io.pivotal	●	●	●	●	●	●	●
io.pivotal.AppsManagerDemoApplication	●	●	●	●	●	●	●
io.pivotal.TimedLogger	●	●	●	●	●	●	●
org	●	●	●	●	●	●	●
org.apache	●	●	●	●	●	●	●
org.apache.catalina	●	●	●	●	●	●	●
org.apache.catalina.core	●	●	●	●	●	●	●
org.apache.catalina.core.ContainerBase	●	●	●	●	●	●	●

**CLOSE**

You can modify the log level for a logger by clicking the desired level in the logger row, as shown in the image below. Whenever you set a log level, the following happens:

- The log level displays in blue to indicate that it is user-configured.
- Each child namespace of the logger inherits the log level.

💡 **Note:** You can manually set any of the child loggers to override this inheritance.

**Configure Logging Levels** ⓘ

Filter Loggers 418 / 418

LOGGER	OFF	FATAL	ERROR	WARN	INFO	DEBUG	TRACE
org.springframework.boot.actuate.autoconfigure.CrshAu...	●	●	●	●	●	●	●
io.pivotal	●	●	●	●	●	●	●
io.pivotal.AppsManagerDemoApplication	●	●	●	●	●	●	●
io.pivotal.TimedLogger	●	●	●	●	●	●	●
org	●	●	●	●	●	●	●
org.apache	●	●	●	●	●	●	●
org.apache.catalina	●	●	●	●	●	●	●

All of the loggers with user-configured logging levels float to the top of the list.

**Configure Logging Levels** ⓘ

Filter Loggers 418 / 418

LOGGER	OFF	FATAL	ERROR	WARN	INFO	DEBUG	TRACE
ROOT	●	●	●	●	●	●	●
io.pivotal	●	●	●	●	●	●	●
io.pivotal.AppsManagerDemoApplication	●	●	●	●	●	●	●
org.apache.catalina.startup.DigesterFactory	●	●	●	●	●	●	●
org.apache.catalina.util.LifecycleBase	●	●	●	●	●	●	●
org.apache.coyote.http11.Http11NioProtocol	●	●	●	●	●	●	●
org.apache.sshd.common.util.SecurityUtils	●	●	●	●	●	●	●
org.apache.tomcat.util.net.NioSelectorPool	●	●	●	●	●	●	●
org.crsh.plugin	●	●	●	●	●	●	●

Close

You can reset log levels by clicking the white dot displayed on the current log level.

You can also filter which loggers you see using the **Filter Loggers** textbox.

## Configure Logging Levels ⓘ

```
= org|
LOGGER
org
org.apache
org.apache.catalina
org.apache.catalina.core
```

## Troubleshoot Spring Boot Actuator Integration

This section describes how to troubleshoot common issues with the integration of Apps Manager and Spring Boot Actuator.

### /cloudfoundryapplication Failed Request

#### Symptom

You see the following failed request message in your app logs:

```
Could not find resource for relative : /cloudfoundryapplication of full path: http://example.com/cloudfoundryapplication
```

#### Explanation

Apps Manager uses the `/cloudfoundryapplication` endpoint as the root for Spring Boot Actuator integrations. It calls this endpoint for an app when you view the app in the Apps Manager UI, regardless of whether you have [configured Spring Boot Actuator endpoints for Apps Manager](#).

#### Solution

If you are not using the Spring Boot Actuator integrations for Apps Manager, you can ignore this failed request message.

## Configuring Multi-Foundation Support in Apps Manager

This topic describes how to configure multi-foundation support in Apps Manager.

### Overview

Configuring multi-foundation support in Apps Manager allows you to search, view, and manage orgs, spaces, apps, and service instances across multiple foundations from a single interface.

### Configure Multi-Foundation Support

This section describes the procedures for configuring multi-foundation support in Apps Manager.

To configure multi-foundation support in Apps Manager, do the following:

1. Configure PAS. See [Configure PAS](#).
2. Add redirects. See [Add Redirects in UAA](#).
3. Add trusted Certificate Authorities (CAs). See [Add Trusted Certificate Authorities](#).
4. (Optional) Configure the SAML Identity Provider. See [\(Optional\) Configure the SAML Identity Provider](#).

### Configure PAS

This section describes how to configure Pivotal Application Service (PAS) on a foundation to enable multi-foundation support in Apps Manager.

To configure multi-foundation support in Apps Manager, do the following:

1. In the PAS tile of one of your foundations, select **Apps Manager**. The following image shows the **Apps Manager** pane.

## Configure Apps Manager

Enable Invitations \*

Display Marketplace Service Plan Prices

Supported currencies as JSON \*

{ "usd": "\$", "eur": "€" }



Product Name

Marketplace Name

Customize Sidebar Links \*

Add

You may configure up to 10 links in the Apps Manager sidebar.

► Marketplace



► Docs



► Tools



Apps Manager Memory Usage (MB) ( min: 128 )

Invitations Memory Usage (MB) ( min: 256 )

Memory limit with which to deploy the Invitations app. Leave blank to use the default value.

Apps Manager Polling Interval \*

30

App Details Polling Interval \*

10

Multi-Foundation Configuration (BETA)

Save

- For **Multi-foundation Configuration (BETA)**, enter a JSON object for each additional foundation that you want to manage. Use the following format for the JSON object:

```
{
 "FOUNDATION_NAME": {
 "ccUrl": "https://api.FOUNDATION_SYSTEM_DOMAIN.com",
 "name": "My Foundation",
 "url": "https://foundation.mydomain.com"
 }
}
```

```

 "systemDomain": "FOUNDATION_SYSTEM_DOMAIN.com",
 "usageServiceUrl": "https://app-usage.FOUNDATION_SYSTEM_DOMAIN.com",
 "invitationsServiceUrl": "https://p-invitations.FOUNDATION_SYSTEM_DOMAIN.com",
 "logoutUrl": "https://login.FOUNDATION_SYSTEM_DOMAIN.com/logout.do",
 "metricsUrl": "https://metrics.FOUNDATION_SYSTEM_DOMAIN.com",
 "uaaUrl": "https://login.FOUNDATION_SYSTEM_DOMAIN.com"
 },
}

```

Where:

- `FOUNDATION_NAME` is a name for the foundation.
- `FOUNDATION_SYSTEM_DOMAIN` is the system domain of the foundation. The system domain is listed in the PAS tile **Domains** pane.
- `metricsUrl` is an optional field. If PCF Metrics is installed on the foundation, provide the PCF Metrics URL.

## Add Redirects in UAA

The Apps Manager UAA clients on each foundation you want to manage must include a redirect to the foundation on which you enabled multi-foundation support. This allows Apps Manager to redirect to the other foundations and back when authenticating with UAA. For more information about UAA clients and configuring valid redirect URLs, see [Client.redirect-uri](#) in *UAA Concepts*.

To add redirects, do the following:

1. Install the UAA CLI. See [CloudFoundry UAA Command Line Client](#) in GitHub.
2. Record the UAA admin client credentials password for a foundation you want to manage with Apps Manager:
  - a. In the PAS tile **Credentials** tab, for **Admin Client Credentials**, click **Link to Credentials**.
  - b. Record the value in the `password` field.
3. Use the following command to target UAA on the foundation:

```
uaac target uaa.FOUNDATION_SYSTEM_DOMAIN.com
```

Where `FOUNDATION_SYSTEM_DOMAIN` is the system domain of the foundation.

4. Use the following command to generate a token for the UAA admin client:

```
uaac token client get admin -s ADMIN_CLIENT_PASSWORD
```

Where `ADMIN_CLIENT_PASSWORD` is the **Admin Client Credentials** password you recorded in step 2.

5. Record the UAA redirect URIs. Use the following command to view the redirect URIs on the UAA `apps_manager_js` client:

```
uaac client get apps_manager_js
```

6. Add the URI of the foundation on which you enabled multi-foundation support to the `apps_manager_js` client:

```
uaac client update apps_manager_js --redirect_uri \
'PREVIOUS_REDIRECT_URIS,https://apps.FOUNDATION_SYSTEM_DOMAIN.com/**'
```

Where:

- `PREVIOUS_REDIRECT_URIS` is one or more redirect URIs that you recorded in the previous step. Include the previous redirect URIs in the command to ensure they are not overwritten.
- `FOUNDATION_SYSTEM_DOMAIN` is the system domain of foundation on which you enabled multi-foundation support.

7. Repeat steps 2 through 6 for each foundation you want to manage with Apps Manager.

## Add Trusted Certificate Authorities

Apps Manager must be able to validate the CAs used by all foundations.

If any foundation uses a certificate for TLS termination that is not signed by a globally-trusted CA, add the CA that signed the TLS certificate as a trusted CA on the foundation on which you enabled multi-foundation support.

For more information, see [Setting Trusted Certificates](#).

## (Optional) Configure the SAML Identity Provider

Pivotal recommends that all foundations use the same external SAML identity provider. This allows Apps Manager to automatically authenticate with the identity provider.

To configure all foundations to use the same SAML identity provider, do the following for each foundation:

1. Navigate to the PAS tile UAA pane.
2. Verify that the foundation uses the correct provider for **Provider Name**.

For more information on UAA and identity providers, see [Identity Providers in UAA](#).

For more information on how to setup identity providers in UAA, see [Adding Existing SAML or LDAP Users to a PCF Deployment](#)

## Using the Cloud Foundry Command Line Interface (cf CLI)

This guide explains the Cloud Foundry Command Line Interface (cf CLI), a tool you use to deploy and manage your applications.

Contents in this section:

- [Installing the cf CLI](#)
- [Getting Started with the cf CLI](#)
- [Using the cf CLI with an HTTP Proxy Server ↗](#)
- [Using the cf CLI with a Self-Signed Certificate](#)
- [Using cf CLI Plugins](#)
- [Developing cf CLI Plugins](#)
- [Cloud Foundry CLI Reference Guide](#)
- [Using Experimental cf CLI Commands](#)

## Installing the cf CLI

This topic describes how to install and uninstall the Cloud Foundry Command Line Interface (cf CLI).

### Overview

The cf CLI is the official command line client for Cloud Foundry.

Follow the instructions in this topic to install the cf CLI on your operating system. You can install the cf CLI with a package manager, an installer, or a compressed binary.

For more information about how to use the cf CLI, see [Getting Started with cf CLI](#).

To learn when cf CLI updates are released and to download a new binary or installer, see [Releases](#) in the GitHub.

### Prerequisites

If you previously used the cf CLI Ruby gem, you must uninstall the gem before installing the cf CLI.

To uninstall the gem, do the following:

1. Run `gem uninstall cf`.
2. Verify that your Ruby environment manager uninstalled the gem by closing and reopening your terminal.

## Install the cf CLI Using a Package Manager

This section describes how to install the cf CLI using a package manager. You can install the cf CLI using a package manager on Mac OS X and Linux operating systems.

### Linux Installation

There are two ways to install the cf CLI using a package manager, depending on your Linux distribution.

To install the cf CLI on Debian and Ubuntu-based Linux distributions, do the following:

1. Add the Cloud Foundry Foundation public key and package repository to your system:

```
wget -q -O - https://packages.cloudfoundry.org/debian/cli.cloudfoundry.org.key | sudo apt-key add -

echo "deb https://packages.cloudfoundry.org/debian stable main" | sudo tee /etc/apt/sources.list.d/cloudfoundry-cli.list
```

2. Update your local package index:

```
sudo apt-get update
```

3. Install the cf CLI:

```
sudo apt-get install cf-cli
```

To install the cf CLI on Enterprise Linux and Fedora RHEL6/CentOS6 and later distributions, do the following:

1. Configure the Cloud Foundry Foundation package repository:

```
sudo wget -O /etc/yum.repos.d/cloudfoundry-cli.repo https://packages.cloudfoundry.org/fedora/cloudfoundry-cli.repo
```

2. Install the cf CLI. The following command also downloads and adds the public key to your system.

```
sudo yum install cf-cli
```

## Mac OS X Installation

You can install the cf CLI on Mac OS X operating systems using the Homebrew package manager.

To install the cf CLI for Mac OS X using Homebrew, do the following:

1. Install Homebrew. For instructions, see [Install Homebrew](#) on the Homebrew website.
2. Install the Cloud Foundry Homebrew package:

```
brew tap cloudfoundry/tap
```

3. Install the cf CLI:

```
brew install cf-cli
```

## Install the cf CLI Using an Installer

This section describes how to install the cf CLI using an installer. You can install the cf CLI using an installer on Windows, Mac OS X, and Linux operating systems.

## Linux Installation

This section describes how to install the cf CLI using an installer on the Debian, Ubuntu, and Red Hat Linux distributions.

To install the cf CLI for Linux using an installer, do the following:

1. Download the Linux installer.
  - o **Debian or Ubuntu:** Paste the following URL in a browser to download the installer:

```
https://cli.run.pivotal.io/stable?release=debian64&source=github
```

- o **Red Hat:** Paste the following URL in a browser to download the installer:

```
https://cli.run.pivotal.io/stable?release=redhat64&source=github
```

2. Install the cf CLI using your system's package manager.

 **Note:** The following commands may require `sudo`.

- o **Debian or Ubuntu:** Run the following command:

```
dpkg -i path/to(cf-cli-*.deb) && apt-get install -f
```

- o **Red Hat:** Run the following command:

```
rpm -i path/to(cf-cli-*.rpm)
```

## Mac OS X Installation

This section describes how to install the cf CLI on Mac OS X operating systems using an installer.

To install the cf CLI for Mac OS X using an installer, do the following:

1. Paste the following URL in a browser to download the installer:

```
https://cli.run.pivotal.io/stable?release=macosx64&source=github
```

2. Open the `.pkg` file.
3. In the installer wizard, click **Continue**.
4. Select an install destination and click **Continue**.
5. Click **Install**.

## Windows Installation

This section describes how to install the cf CLI for use on the Windows command line. For instructions on how to install the cf CLI for use on the Windows Subsystem for Linux (WSL), also known as Bash, see [Linux Installation](#) above.

To install the cf CLI for Windows using an installer, do the following:

1. Paste the following URL in a browser to download the installer:

```
https://cli.run.pivotal.io/stable?release=windows64&source=github
```

2. Unzip the file.
3. Right-click on the `cf_installer` executable and select **Run as Administrator**.
4. When prompted, click **Install**. Then, click **Finish**.

## Install the cf CLI Using a Compressed Binary

This section describes how to install the cf CLI using a compressed binary. You can install the cf CLI using a compressed binary on Windows, Mac OS X, and Linux operating systems.

For more information about downloading and installing a compressed binary, see the [Installers and compressed binaries](#) section of the cf CLI repository on GitHub.

## Verify Installation

To verify the installation of the cf CLI, do the following:

1. Close and reopen the command prompt. Or, open a new tab in the command prompt.
2. Run the following command:

```
cf --help
```

If your installation was successful, the cf CLI help listing appears.

## Uninstall the cf CLI

This section describes how to uninstall the cf CLI. There are different methods for uninstalling the cf CLI depending on the installation method.

### Package Manager

This section describes how to uninstall the cf CLI if you installed the cf CLI with a package manager.

To uninstall the cf CLI, follow the instructions specific to your package manager.

## Installer

This section describes how to uninstall the cf CLI if you installed the cf CLI with an installer. Follow the instructions in this section that are specific to your operating system.

To uninstall the cf CLI on Mac OS X, do the following:

1. Delete the binary `/usr/local/bin/cf`.
2. Delete the directory `/usr/local/share/doc/cf-cli`.

To uninstall the cf CLI on Windows, do the following:

1. Navigate to the **Control Panel** and click **Programs and Features**.
2. Select `Cloud Foundry CLI VERSION` and click **Uninstall**.

## Binary

This section describes how to uninstall the cf CLI if you installed the cf CLI with a binary.

To uninstall the cf CLI, do the following:

1. Navigate to the location where you copied the binary.
2. Delete the binary.

## Getting Started with the cf CLI

This topic describes configuring and getting started with the Cloud Foundry Command Line Interface (cf CLI). This page assumes you have the latest version of the cf CLI. See the [Installing the Cloud Foundry Command Line Interface](#) topic for installation instructions.

### Localize

The cf CLI translates terminal output into the language that you select. The default language is `en-US`. The cf CLI supports the following languages:

- Chinese (simplified): `zh-Hans`
- Chinese (traditional): `zh-Hant`
- English: `en-US`
- French: `fr-FR`
- German: `de-DE`
- Italian: `it-IT`
- Japanese: `ja-JP`
- Korean: `ko-KR`
- Portuguese (Brazil): `pt-BR`
- Spanish: `es-ES`

Use [cf config](#) to set the language. To set the language with `cf config`, use the syntax: `$ cf config --locale YOUR_LANGUAGE`.

For example, to set the language to Portuguese and confirm the change by running `cf help`:

```
$ cf config --locale pt-BR
$ cf help
NOME:
 cf - Uma ferramenta de linha de comando para interagir com Cloud Foundry

USO:
 cf [opções globais] comando [argumentos...] [opções de comando]

VERSAO:
 6.14.1+dc6adf6-2015-12-22
 ...
```

 **Note:** Localization with `cf config --locale` affects only messages that the cf CLI generates.

### Login

Use [cf login](#) to log in to PAS. The `cf login` command uses the following syntax to specify a target API endpoint, an org (organization), and a space:

```
$ cf login [-a API_URL] [-u USERNAME] [-p PASSWORD] [-o ORG] [-s SPACE]
```

- `API_URL`: This is your API endpoint, [the URL of the Cloud Controller in your PAS instance](#).
- `USERNAME`: Your username.
- `PASSWORD`: Your password. Use of the `-p` option is discouraged as it may record your password in your shell history.
- `ORG`: The org where you want to deploy your apps.
- `SPACE`: The space in the org where you want to deploy your apps.

The cf CLI prompts for credentials as needed. If you are a member of multiple orgs or spaces, `cf login` prompts you for which ones to log into. Otherwise it targets your org and space automatically.

```
$ cf login -a https://api.example.com -u username@example.com
API endpoint: https://api.example.com
```

```
Password>
Authenticating...
OK
```

```
Select an org (or press enter to skip):
```

1. example-org
2. example-other-org

```
Org> 1
Targeted org example-org
```

```
Select a space (or press enter to skip):
```

1. development
2. staging
3. production

```
Space> 1
Targeted space development
```

Alternatively, you can write a script to log in and set your target using the non-interactive [cf api](#), [cf auth](#), and [cf target](#) commands.

Upon successful login, the cf CLI saves a `config.json` file containing your API endpoint, org, space values, and access token. If you change these settings, the `config.json` file is updated accordingly.

By default, `config.json` is located in your `~/.cf` directory. The `CF_HOME` environment variable allows you to locate the `config.json` file wherever you like.

## Users and Roles

The cf CLI includes commands that list users and assign roles in orgs and spaces. See the [Orgs, Spaces, Roles, and Permissions](#) topic.

### Commands for Listing Users

These commands take an org or space as an argument:

- [cf org-users](#)
- [cf space-users](#)

For example, to list the users who are members of an org:

```
$ cf org-users example-org
Getting users in org example-org as username@example.com...

ORG MANAGER
username@example.com

BILLING MANAGER
huey@example.com
dewey@example.com

ORG AUDITOR
louie@example.com
```

### Commands for Managing Roles

These commands require PAS admin permissions and take username, org or space, and role as arguments:

- [cf set-org-role](#)
- [cf unset-org-role](#)
- [cf set-space-role](#)
- [cf unset-space-role](#)

Available roles are `OrgManager`, `BillingManager`, `OrgAuditor`, `SpaceManager`, `SpaceDeveloper`, and `SpaceAuditor`. For example, to grant the Org Manager role to

a user within an org:

```
$ cf set-org-role huey@example.com example-org OrgManager
Assigning role OrgManager to user huey@example.com in org example-org as username@example.com...
OK
```

 **Note:** If you are not a PAS admin, you see this message when you try to run these commands:

```
error code: 10003, message: You are not authorized to perform the requested
action
```

## Identical Usernames in Multiple Origins

If a username corresponds to multiple accounts from different user stores, such as both the internal UAA store and an external SAML or LDAP store, the `set` and `unset` role commands above return an error: The user exists in multiple origins. Specify an origin for the requested user from: 'uaa', 'other'

To resolve this ambiguity, the operator can construct a `curl` command that uses the CF API to perform the desired role-management function. For an example, see the [PUT v2/organizations/:guid/auditors](#) API function.

## Push

The `cf push` command pushes a new app or syncs changes to an existing app.

If you do not provide a hostname (also known as subdomain), `cf push` routes your app to a URL of the form `APPNAME.DOMAIN` based on the name of your app and your default domain. If you want to map a different route to your app, see the [Routes and Domains](#) topic for information about creating routes.

The `cf push` command supports many options that determine how and where the app instances are deployed. For details about the `cf push` command, see the [push](#) page in the Cloud Foundry CLI Reference Guide.

The following example pushes an app called `my-awesome-app` to the URL `http://my-awesome-app.example.com` and specifies the Ruby buildpack with the `-b` flag.

 **Note:** When you push an app and specify a buildpack with the `-b` flag, the app remains permanently linked to that buildpack. To use the app with a different buildpack, you must delete the app and re-push it.

```
$ cf push my-awesome-app -b ruby_buildpack
Creating app my-awesome-app in org example-org / space development as username@example.com...
OK

Creating route my-awesome-app.example.com...
OK
...

1 of 1 instances running

App started
...

requested state: started
instances: 1/1
usage: 1G x 1 instances
urls: my-awesome-app.example.com
last uploaded: Wed Jun 8 23:43:15 UTC 2016
stack: cflinuxfs3
buildpack: ruby_buildpack

 state since cpu memory disk details
#0 running 2016-06-08 04:44:07 PM 0.0% 0 of 1G 0 of 1G
```

For more information about available buildpacks, see the [Buildpacks](#) topic.

## Always Provide an App Name to cf push

`cf push` requires an app name, which you provide either in a manifest or at the command line.

`cf push` locates the `manifest.yml` in the current working directory by default, or in the path provided by the `-f` option.

If you do not use a manifest, the minimal `cf push` command looks like this:

```
$ cf push my-app
```

 **Note:** When you provide an app name at the command line, `cf push` uses that app name whether or not there is a different app name in the manifest. If the manifest describes multiple apps, you can push a single app by providing its name at the command line; the cf CLI does not push the others. Use these behaviors for testing.

## How cf push Finds the App

By default, `cf push` recursively pushes the contents of the current working directory. Alternatively, you can provide a path using either a manifest or a command line option.

- If the path is to a directory, `cf push` recursively pushes the contents of that directory instead of the current working directory.
- If the path is to a file, `cf push` pushes only that file.

 **Note:** If you want to push more than a single file, but not the entire contents of a directory, consider using a `.cfignore` file to tell `cf push` what to exclude.

## User-Provided Service Instances

To create or update a user-provided service instance, you need to supply basic parameters. For example a database service might require a username, password, host, port, and database name.

The cf CLI has three ways of supplying these parameters to create or update an instance of a service: interactively, non-interactively, and in conjunction with third-party log management software as described in [RFC 6587](#). When used with third-party logging, the cf CLI sends data formatted according to [RFC 5424](#).

You create a service instance with `cf cups` and update one with `cf uups` as described below.

### The cf create-user-provided-service (cups) Command

Use [cf create-user-provided-service](#) (alias `cf cups`) creates a new service instance.

**To supply service instance parameters interactively:** Specify parameters in a comma-separated list after the `-p` flag. This example command-line session creates a service instance for a database service.

```
$ cf cups sql-service-instance -p "host, port, dbname, username, password"
host> mysql.example.com
port> 1433
dbname> mysqldb
username> admin
password> Pa55w0rd
Creating user provided service sql-service-instance in org example-org / space development as username@example.com...
OK
```

**To supply service instance parameters to cf cups non-interactively:** Pass parameters and their values in as a JSON hash, bound by single quotes, after the `-p` tag. This example is a non-interactive version of the `cf cups` session above.

```
$ cf cups sql-service-instance -p '{"host":"mysql.example.com", "port":"1433", "dbname":"mysqldb", "username":"admin", "password":"pa55w0RD"}'
Creating user provided service sql-service-instance in org example-org / space development as username@example.com...
OK
```

**To create a service instance that sends data to a third-party:** Use the `-l` option followed by the external destination URL. This example creates a service

instance that sends log information to the syslog drain URL of a third-party log management service. For specific log service instructions, see the [Service-Specific Instructions for Streaming Application Logs](#) topic.

```
$ cf cups mylog -l syslog://logs4.example.com:25258
Creating user provided service mylog in org example-org / space development as username@example.com...
OK
```

After you create a user-provided service instance, you bind it to an app with [cf bind-service](#), unbind it with [cf unbind-service](#), rename it with [cf rename-service](#), and delete it with [cf delete-service](#).

## The cf update-user-provided-service (uups) Command

Use [cf update-user-provided-service](#) (alias `cf uups`) to update one or more of the parameters for an existing user-provided service instance. The `cf uups` command uses the same syntax as `cf cups` [above](#) to set parameter values. The `cf uups` command does not update any parameter values that you do not supply.

## cf CLI Return Codes

The cf CLI uses exit codes, which help with scripting and confirming that a command has run successfully. For example, after you run a cf CLI command, you can retrieve its return code by running `echo $?` (on Windows, `echo %ERRORLEVEL%`). If the return code is `0`, the command was successful.

## The cf help Command

The [cf help](#) command lists the cf CLI commands and a brief description of each. Passing the `-h` flag to any command lists detailed help, including any aliases. For example, to see detailed help for `cf delete`, run:

```
$ cf delete -h
NAME:
 delete - Delete an app

USAGE:
 cf delete APP_NAME [-f -r]

ALIAS:
 d

OPTIONS:
 -f Force deletion without confirmation
 -r Also delete any mapped routes
```

## Using the cf CLI with a Proxy Server

If you have an HTTP or SOCKS5 proxy server on your network between a host running the cf CLI and your Cloud Foundry API endpoint, you must set `https_proxy` with the hostname or IP address of the proxy server.

The `https_proxy` environment variable holds the hostname or IP address of your proxy server.

`https_proxy` is a standard environment variable. Like any environment variable, the specific steps you use to set it depends on your operating system.

### Format of https\_proxy

`https_proxy` is set with hostname or IP address of the proxy server in URL format: `https_proxy=http://proxy.example.com`

If the proxy server requires a user name and password, include the credentials: `https_proxy=http://username:password@proxy.example.com`

If the proxy server uses a port other than 80, include the port number: `https_proxy=http://username:password@proxy.example.com:8080`

If the proxy server is a SOCKS5 proxy, specify the SOCKS5 protocol in the URL: `https_proxy=socks5://socks_proxy.example.com`

### Using SOCKS5 with v3-ssh

The `cf v3-ssh` command supports SOCKS5 proxies. To specify the SOCKS5 proxy server, set the `ALL_PROXY` environment variable using the following format: `ALL_PROXY=socks5://socks_proxy.example.com`.

 **Note:** `cf ssh` does not work through a SOCKS5 proxy.

### Setting https\_proxy in Mac OS or Linux

Set the `https_proxy` environment variable using the command specific to your shell. For example, in bash, use the `export` command.

Example:

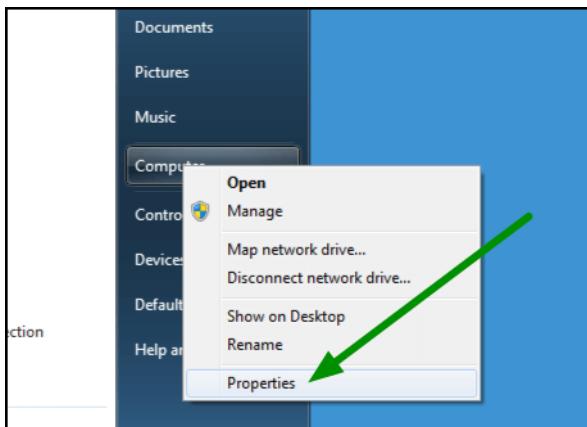
```
$ export https_proxy=http://my.proxyserver.com:8080
```

To make this change persistent, add the command to the appropriate profile file for the shell. For example, in bash, add a line like the following to your `.bash_profile` or `.bashrc` file:

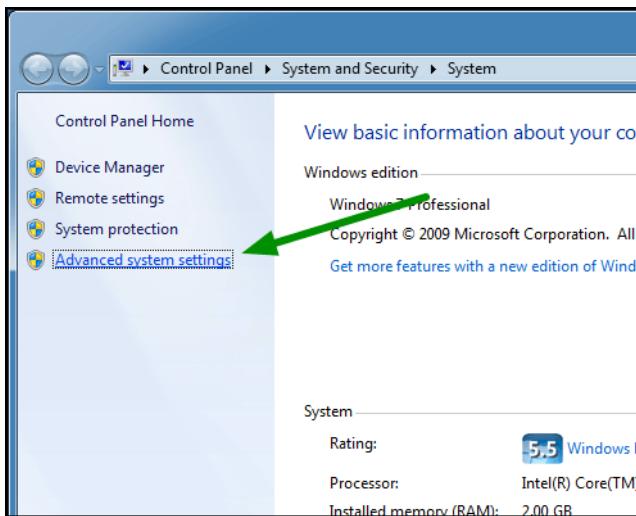
```
https_proxy=http://username:password@hostname:port
export $https_proxy
```

### Setting https\_proxy in Windows

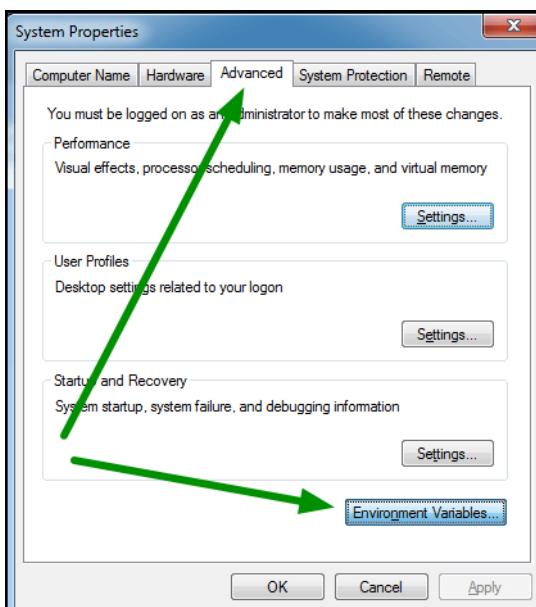
1. Open the Start menu. Right-click **Computer** and select **Properties**.



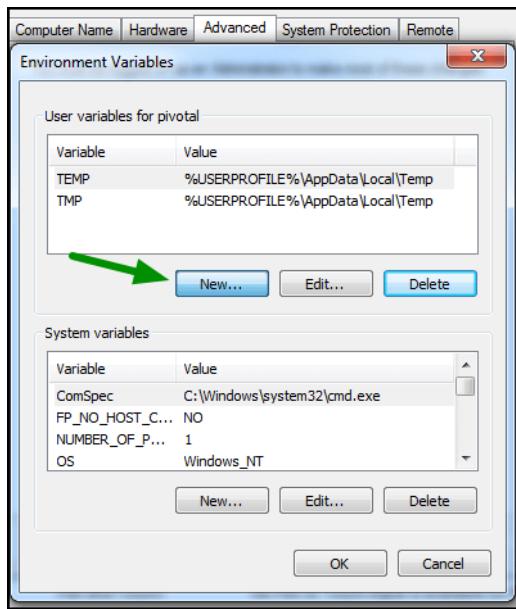
- In the left pane of the System window, click Advanced system settings.



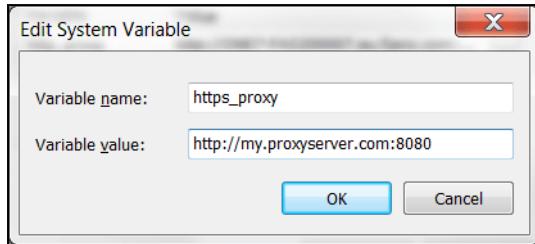
- In the System Properties window, select the Advanced tab, then click Environment Variables.



- In the Environment Variables window, under User variables, click New.



5. In the Variable name field, input `https_proxy`. In the Variable value field, input your proxy server information.



6. Click **OK**.

## Using the cf CLI with a Self-Signed Certificate

This topic describes how developers can use the cf CLI to communicate securely with a Pivotal Cloud Foundry (PCF) deployment without specifying `--skip-ssl-validation` under the following circumstances:

- The deployment uses a self-signed certificate.
- The deployment uses a certificate that is signed by a self-signed certificate authority (CA), or a certificate signed by a certificate that's signed by a self-signed CA.

Before following the procedure below, the developer must obtain either the self-signed certificate or the intermediate and CA certificate(s) used to sign the deployment's certificate. The developer can obtain these certificates from the PCF operator.

### Install the Certificate on Local Machines

The certificates that developers must insert into their local truststore vary depending on the configuration of the deployment.

- If the deployment uses a self-signed certificate, the developer must insert the self-signed certificate into their local truststore.
- If the deployment uses a certificate that is signed by a self-signed certificate authority (CA), or a certificate signed by a certificate that's signed by a self-signed CA, the developer must insert the self-signed certificate and any intermediate certificates into their local truststore.

#### Installing the Certificate on Mac OS X

Enter the following command to place a certificate file `server.crt` into your local truststore:

```
$ sudo security add-trusted-cert -d -r trustRoot -k /Library/Keychains/System.keychain server.crt
```

#### Installing the Certificate on Linux

Perform the following steps specific to your distribution to place the certificate file `server.crt` into your truststore:

- Debian/Ubuntu/Gentoo:

```
$ cat server.crt >> /etc/ssl/certs/ca-certificates.crt
```

- Fedora/RHEL:

```
$ cat server.crt >> /etc/pki/tls/certs/ca-bundle.crt
```

The above example will set certificate permanently on your machine accross all users and requires sudo permissions. You can also run the following command to set certificate in your current terminal/script:

```
$ export SSL_CERT_FILE=/path/to/server.crt
```

or

```
$ export SSL_CERT_DIR=/path/to/server/dir
```

#### Installing the Certificate on Windows

1. Right-click on the certificate file and click **Install Certificate**.
2. Choose to install the certificate as the **Current User** or **Local Machine**. Choose the **Trusted Root Certification Authorities** as the certification store.

## Using cf CLI Plugins

The Cloud Foundry Command Line Interface (cf CLI) includes plugin functionality. These plugins enable developers to add custom commands to the cf CLI. You can install and use plugins that Cloud Foundry developers and third-party developers create. You can review the [Cloud Foundry Community CLI Plugin page](#) for a current list of community-supported plugins. You can find information about submitting your own plugin to the community in the [Cloud Foundry CLI plugin repository](#) on GitHub.

 **warning:** Plugins are not vetted in any way, including for security or functionality. Use plugins at your own risk.

The cf CLI identifies a plugin by its binary filename, its developer-defined plugin name, and the commands that the plugin provides. You use the binary filename only to install a plugin. You use the plugin name or a command for any other action.

 **Note:** The cf CLI uses case-sensitive commands, but plugin management commands accept plugin and repository names irrespective of their casing.

## Changing the Plugin Directory

By default, the cf CLI stores plugins on your workstation in `$CF_HOME/.cf/plugins`, which defaults to `$HOME/.cf/plugins`. To change the root directory of this path from `$CF_HOME`, set the `CF_PLUGIN_HOME` environment variable. The cf CLI appends `.cf/plugins` to the `CF_PLUGIN_HOME` path that you specify and stores plugins in that location.

For example, if you set `CF_PLUGIN_HOME` to `/my-folder`, cf CLI stores plugins in `/my-folder/.cf/plugins`.

## Installing a Plugin

1. Download a binary or the source code for a plugin from a trusted provider.

 **Note:** The cf CLI requires a binary file compiled from source code written in Go. If you download source code, you must compile the code to create a binary.

2. Run `cf install-plugin BINARY-FILENAME` to install a plugin. Replace `BINARY-FILENAME` with the path to and name of your binary file.

 **Note:** You cannot install a plugin that has the same name or that uses the same command as an existing plugin. You will be prompted to uninstall the existing plugin.

 **Note:** The cf CLI prohibits you from implementing any plugin that uses a native cf CLI command name or alias. For example, if you attempt to install a third-party plugin that includes the command `cf push`, the cf CLI halts the installation.

## Running a Plugin Command

Use the contents of the `cf help` CLI plugin management and Commands offered by installed plugins sections to manage plugins and run plugin commands.

1. Run `cf plugins` to list all installed plugins and all commands that the plugins provide.
2. Run `cf PLUGIN-COMMAND` to execute a plugin command.

## Checking for Plugin Updates

Run `cf plugins --outdated` to check all registered plugin repositories for newer versions of currently installed plugins.

Example:

```
$ cf plugins --outdated
Searching CF-Community, company-repo for newer versions of installed plugins...
plugin version latest version
coffeemaker 1.1.2 1.2.0
Use 'cf install-plugin' to update a plugin to the latest version.
```

For more information about the `cf plugins` command, see [cf plugins](#).

## Uninstalling a Plugin

Use the `PLUGIN-NAME` to remove a plugin, not the `BINARY-Filename`.

1. Run `cf plugins` to view the names of all installed plugins.
2. Run `cf uninstall-plugin PLUGIN-NAME` to remove a plugin.

## Adding a Plugin Repository

Run `cf add-plugin-repo REPOSITORY-NAME-URL` to add a plugin repository.

Example:

```
$ cf add-plugin-repo CF-Community https://plugins.cloudfoundry.org
https://plugins.cloudfoundry.org added as CF-Community
```

## Listing Available Plugin Repositories

Run `cf list-plugin-repos` to view your available plugin repositories.

Example:

```
$ cf list-plugin-repos
OK
Repo Name Url
CF-Community https://plugins.cloudfoundry.org
```

## Listing All Plugins by Repository

Run `cf repo-plugins` to show all plugins from all available repositories.

## Troubleshooting

The cf CLI provides the following error messages to help you troubleshoot installation and usage issues. Third-party plugins can provide their own error messages.

### Permission Denied

If you receive a `permission denied` error message, you lack required permissions to the plugin. You must have `read` and `execute` permissions to the plugin binary file.

### Plugin Command Collision

Plugin names and commands must be unique. The CLI displays an error message if you attempt to install a plugin with a non-unique name or command.

If the plugin has the same name or command as a currently installed plugin, you must first uninstall the existing plugin to install the new plugin.

If the plugin has a command with the same name as a native cf CLI command or alias, you cannot install the plugin.

## Developing cf CLI Plugins

Users can create and install Cloud Foundry Command Line Interface (cf CLI) plugins to provide custom commands. These plugins can be submitted and shared to the [CF Community repository](#).

### Requirements

Using plugins requires cf CLI v.6.7 or higher. Refer to the [Installing the Cloud Foundry Command Line Interface](#) topic for information about downloading, installing, and uninstalling the cf CLI.

### Installing the Architecture

1. Implement the [predefined plugin interface](#).
2. Clone the [template repository](#). You will need the [basic GO plugin](#).

### Initializing the Plugin

To initialize a plugin, call `plugin.Start(new(MyPluginStruct))` from within the `main()` method of your plugin. The `plugin.Start(...)` function requires a new reference to the `struct` that implements the defined interface.

### Invoking cf CLI Commands

Invoke cf CLI commands with `cliConnection.CliCommand([]args)` from within a plugin's `Run(...)` method. The `Run(...)` method receives the `cliConnection` as its first argument. The `cliConnection.CliCommand([]args)` returns the output printed by the command and an error.

The output is returned as a slice of strings. The error will be present if the call to the cf CLI command fails.

For more information, see the [API commands documentation](#).

### Installing a Plugin

To install a plugin, run `cf install-plugin PATH_TO_PLUGIN_BINARY`.

For additional information about developing plugins, see the [plugin development guide](#).

## Cloud Foundry CLI Reference Guide

### Name

cf - A command line tool to interact with Cloud Foundry

### Usage

cf [global options] command [arguments...] [command options]

### Version

6.45.0+5f9ff16f9.2019-06-03

## Getting Started

Command	Description
<a href="#">help ↗</a>	Show help
<a href="#">version ↗</a>	Print the version
<a href="#">login ↗</a>	Log user in
<a href="#">logout ↗</a>	Log user out
<a href="#">passwd ↗</a>	Change user password
<a href="#">target ↗</a>	Set or view the targeted org or space
<a href="#">api ↗</a>	Set or view target api url
<a href="#">auth ↗</a>	Authenticate non-interactively

## Apps

Command	Description
<a href="#">apps ↗</a>	List all apps in the target space
<a href="#">app ↗</a>	Display health and status for an app
<a href="#">push ↗</a>	Push a new app or sync changes to an existing app
<a href="#">scale ↗</a>	Change or view the instance count, disk space limit, and memory limit for an app
<a href="#">delete ↗</a>	Delete an app
<a href="#">rename ↗</a>	Rename an app
<a href="#">start ↗</a>	Start an app
<a href="#">stop ↗</a>	Stop an app
<a href="#">restart ↗</a>	Stop all instances of the app, then start them again. This causes downtime.
<a href="#">restage ↗</a>	Recreate the app's executable artifact using the latest pushed app files and the latest environment (variables, service bindings, buildpack, stack, etc.). This action will cause app downtime.
<a href="#">restart-app-instance ↗</a>	Terminate, then restart an app instance
<a href="#">run-task ↗</a>	Run a one-off task on an app
<a href="#">tasks ↗</a>	List tasks of an app

Command	Description
<a href="#">terminate-task</a>	Terminate a running task of an app
<a href="#">events</a>	Show recent app events
<a href="#">files</a>	Print out a list of files in a directory or the contents of a specific file of an app running on the DEA backend
<a href="#">logs</a>	Tail or show recent logs for an app
<a href="#">env</a>	Show all env variables for an app
<a href="#">set-env</a>	Set an env variable for an app
<a href="#">unset-env</a>	Remove an env variable from an app
<a href="#">stacks</a>	List all stacks (a stack is a pre-built file system, including an operating system, that can run apps)
<a href="#">stack</a>	Show information for a stack (a stack is a pre-built file system, including an operating system, that can run apps)
<a href="#">copy-source</a>	Copies the source code of an application to another existing application (and restarts that application)
<a href="#">create-app-manifest</a>	Create an app manifest for an app that has been pushed successfully
<a href="#">get-health-check</a>	Show the type of health check performed on an app
<a href="#">set-health-check</a>	Change type of health check performed on an app
<a href="#">enable-ssh</a>	Enable ssh for the application
<a href="#">disable-ssh</a>	Disable ssh for the application
<a href="#">ssh-enabled</a>	Reports whether SSH is enabled on an application container instance
<a href="#">ssh</a>	SSH to an application container instance

## Services

Command	Description
<a href="#">marketplace</a>	List available offerings in the marketplace
<a href="#">services</a>	List all service instances in the target space
<a href="#">service</a>	Show service instance info
<a href="#">create-service</a>	Create a service instance
<a href="#">update-service</a>	Update a service instance
<a href="#">delete-service</a>	Delete a service instance
<a href="#">rename-service</a>	Rename a service instance
<a href="#">create-service-key</a>	Create key for a service instance
<a href="#">service-keys</a>	List keys for a service instance
<a href="#">service-key</a>	Show service key info
<a href="#">delete-service-key</a>	Delete a service key
<a href="#">bind-service</a>	Bind a service instance to an app
<a href="#">unbind-service</a>	Unbind a service instance from an app
<a href="#">bind-route-service</a>	Bind a service instance to an HTTP route
<a href="#">unbind-route-service</a>	Unbind a service instance from an HTTP route
<a href="#">create-user-provided-service</a>	Make a user-provided service instance available to CF apps
<a href="#">update-user-provided-service</a>	Update user-provided service instance
<a href="#">share-service</a>	Share a service instance with another space
<a href="#">unshare-service</a>	Unshare a shared service instance from a space

Command	Description
---------	-------------

## Orgs

Command	Description
<a href="#">orgs</a>	List all orgs
<a href="#">org</a>	Show org info
<a href="#">create-org</a>	Create an org
<a href="#">delete-org</a>	Delete an org
<a href="#">rename-org</a>	Rename an org

## Spaces

Command	Description
<a href="#">spaces</a>	List all spaces in an org
<a href="#">space</a>	Show space info
<a href="#">create-space</a>	Create a space
<a href="#">delete-space</a>	Delete a space
<a href="#">rename-space</a>	Rename a space
<a href="#">allow-space-ssh</a>	Allow SSH access for the space
<a href="#">disallow-space-ssh</a>	Disallow SSH access for the space
<a href="#">space-ssh-allowed</a>	Reports whether SSH is allowed in a space

## Domains

Command	Description
<a href="#">domains</a>	List domains in the target org
<a href="#">create-domain</a>	Create a domain in an org for later use
<a href="#">delete-domain</a>	Delete a domain
<a href="#">create-shared-domain</a>	Create a domain that can be used by all orgs (admin-only)
<a href="#">delete-shared-domain</a>	Delete a shared domain
<a href="#">router-groups</a>	List router groups

## Routes

Command	Description
<a href="#">routes</a>	List all routes in the current space or the current organization
<a href="#">create-route</a>	Create a url route in a space for later use
<a href="#">check-route</a>	Perform a simple check to determine whether a route currently exists or not
<a href="#">map-route</a>	Add a url route to an app
<a href="#">unmap-route</a>	Remove a url route from an app
<a href="#">delete-route</a>	Delete a route
<a href="#">delete-orphaned-routes</a>	Delete all orphaned routes in the currently targeted space (i.e. those that are not mapped to an app)

## Network Policies

Command	Description
<a href="#">network-policies</a>	List direct network traffic policies
<a href="#">add-network-policy</a>	Create policy to allow direct network traffic from one app to another
<a href="#">remove-network-policy</a>	Remove network traffic policy of an app

## Buildpacks

Command	Description
<a href="#">buildpacks</a>	List all buildpacks
<a href="#">create-buildpack</a>	Create a buildpack
<a href="#">update-buildpack</a>	Update a buildpack
<a href="#">rename-buildpack</a>	Rename a buildpack
<a href="#">delete-buildpack</a>	Delete a buildpack

## User Admin

Command	Description
<a href="#">create-user</a>	Create a new user
<a href="#">delete-user</a>	Delete a user
<a href="#">org-users</a>	Show org users by role
<a href="#">set-org-role</a>	Assign an org role to a user
<a href="#">unset-org-role</a>	Remove an org role from a user
<a href="#">space-users</a>	Show space users by role
<a href="#">set-space-role</a>	Assign a space role to a user
<a href="#">unset-space-role</a>	Remove a space role from a user

## Org Admin

Command	Description
<a href="#">quotas</a>	List available usage quotas
<a href="#">quota</a>	Show quota info
<a href="#">set-quota</a>	Assign a quota to an org
<a href="#">create-quota</a>	Define a new resource quota
<a href="#">delete-quota</a>	Delete a quota
<a href="#">update-quota</a>	Update an existing resource quota
<a href="#">share-private-domain</a>	Share a private domain with an org
<a href="#">unshare-private-domain</a>	Unshare a private domain with an org

## Space Admin

Command	Description

Command	Description
<a href="#">space-quotas</a>	List available space resource quotas
<a href="#">space-quota</a>	Show space quota info
<a href="#">create-space-quota</a>	Define a new space resource quota
<a href="#">update-space-quota</a>	Update an existing space quota
<a href="#">delete-space-quota</a>	Delete a space quota definition and unassign the space quota from all spaces
<a href="#">set-space-quota</a>	Assign a space quota definition to a space
<a href="#">unset-space-quota</a>	Unassign a quota from a space

## Service Admin

Command	Description
<a href="#">service-auth-tokens</a>	List service auth tokens
<a href="#">create-service-auth-token</a>	Create a service auth token
<a href="#">update-service-auth-token</a>	Update a service auth token
<a href="#">delete-service-auth-token</a>	Delete a service auth token
<a href="#">service-brokers</a>	List service brokers
<a href="#">create-service-broker</a>	Create a service broker
<a href="#">update-service-broker</a>	Update a service broker
<a href="#">delete-service-broker</a>	Delete a service broker
<a href="#">rename-service-broker</a>	Rename a service broker
<a href="#">migrate-service-instances</a>	Migrate service instances from one service plan to another
<a href="#">purge-service-offering</a>	Recursively remove a service and child objects from Cloud Foundry database without making requests to a service broker
<a href="#">purge-service-instance</a>	Recursively remove a service instance and child objects from Cloud Foundry database without making requests to a service broker
<a href="#">service-access</a>	List service access settings
<a href="#">enable-service-access</a>	Enable access to a service or service plan for one or all orgs
<a href="#">disable-service-access</a>	Disable access to a service or service plan for one or all orgs

## Security Group

Command	Description
<a href="#">security-group</a>	Show a single security group
<a href="#">security-groups</a>	List all security groups
<a href="#">create-security-group</a>	Create a security group
<a href="#">update-security-group</a>	Update a security group
<a href="#">delete-security-group</a>	Deletes a security group
<a href="#">bind-security-group</a>	Bind a security group to a particular space, or all existing spaces of an org
<a href="#">unbind-security-group</a>	Unbind a security group from a space
<a href="#">bind-staging-security-group</a>	Bind a security group to the list of security groups to be used for staging applications
<a href="#">staging-security-groups</a>	List security groups in the staging set for applications
<a href="#">unbind-staging-security-group</a>	Unbind a security group from the set of security groups for staging applications

Command	Description
<a href="#">bind-running-security-group ↗</a>	Bind a security group to the list of security groups to be used for running applications
<a href="#">running-security-groups ↗</a>	List security groups in the set of security groups for running applications
<a href="#">unbind-running-security-group ↗</a>	Unbind a security group from the set of security groups for running applications

## Environment Variable Groups

Command	Description
<a href="#">running-environment-variable-group ↗</a>	Retrieve the contents of the running environment variable group
<a href="#">staging-environment-variable-group ↗</a>	Retrieve the contents of the staging environment variable group
<a href="#">set-staging-environment-variable-group ↗</a>	Pass parameters as JSON to create a staging environment variable group
<a href="#">set-running-environment-variable-group ↗</a>	Pass parameters as JSON to create a running environment variable group

## Isolation Segments

Command	Description
<a href="#">isolation-segments ↗</a>	List all isolation segments
<a href="#">create-isolation-segment ↗</a>	Create an isolation segment
<a href="#">delete-isolation-segment ↗</a>	Delete an isolation segment
<a href="#">enable-org-isolation ↗</a>	Entitle an organization to an isolation segment
<a href="#">disable-org-isolation ↗</a>	Revoke an organization's entitlement to an isolation segment
<a href="#">set-org-default-isolation-segment ↗</a>	Set the default isolation segment used for apps in spaces in an org
<a href="#">reset-org-default-isolation-segment ↗</a>	Reset the default isolation segment used for apps in spaces of an org
<a href="#">set-space-isolation-segment ↗</a>	Assign the isolation segment for a space
<a href="#">reset-space-isolation-segment ↗</a>	Reset the space's isolation segment to the org default

## Feature Flags

Command	Description
<a href="#">feature-flags ↗</a>	Retrieve list of feature flags with status
<a href="#">feature-flag ↗</a>	Retrieve an individual feature flag with status
<a href="#">enable-feature-flag ↗</a>	Allow use of a feature
<a href="#">disable-feature-flag ↗</a>	Prevent use of a feature

## Advanced

Command	Description
<a href="#">curl ↗</a>	Executes a request to the targeted API endpoint
<a href="#">config ↗</a>	Write default values to the config
<a href="#">oauth-token ↗</a>	Retrieve and display the OAuth token for the current session
<a href="#">ssh-code ↗</a>	Get a one time password for ssh clients

## Add/remove Plugin Repository

Command	Description

Command	Description
<a href="#">add-plugin-repo ↗</a>	Plugin repository
<a href="#">remove-plugin-repo ↗</a>	Remove a plugin repository
<a href="#">list-plugin-repos ↗</a>	List all the added plugin repositories
<a href="#">repo-plugins ↗</a>	List all available plugins in specified repository or in all added repositories

## Add/remove Plugin

Command	Description
<a href="#">plugins ↗</a>	List commands of installed plugins
<a href="#">install-plugin ↗</a>	Install CLI plugin
<a href="#">uninstall-plugin ↗</a>	Uninstall CLI plugin

## Environment Variables

Variable	Description
CF_COLOR=false	Do not colorize output
CF_DIAL_TIMEOUT=5	Max wait time to establish a connection, including name resolution, in seconds
CF_HOME=path/to/dir/	Override path to default config directory
CF_PLUGIN_HOME=path/to/dir/	Override path to default plugin config directory
CF_TRACE=true	Print API request diagnostics to stdout
CF_TRACE=path/to/trace.log	Append API request diagnostics to a log file
all_proxy=proxy.example.com:8080	Specify a proxy server to enable proxying for all requests
https_proxy=proxy.example.com:8080	Enable proxying for HTTP requests

## Global Options

Option	Description
-help, -h	Show help
-v	Print API request diagnostics to stdout

## Apps (experimental)

Command	Description
<a href="#">v3-apps ↗</a>	List all apps in the target space
<a href="#">v3-create-app ↗</a>	Create a V3 App
<a href="#">v3-push ↗</a>	Push a new app or sync changes to an existing app
<a href="#">v3-scale ↗</a>	Change or view the instance count, disk space limit, and memory limit for an app
<a href="#">v3-delete ↗</a>	Delete a V3 App
<a href="#">v3-start ↗</a>	Start an app
<a href="#">v3-stop ↗</a>	Stop an app
<a href="#">v3-restart ↗</a>	Stop all instances of the app, then start them again. This causes downtime.
<a href="#">v3-stage ↗</a>	Create a new droplet for an app
<a href="#">v3-restart-app-instance ↗</a>	Terminate, then instantiate an app instance
<a href="#">v3-apply-manifest ↗</a>	Applies manifest properties to an application

Command	Description
<a href="#">v3-droplets</a>	Lists of an app
<a href="#">v3-set-droplet</a>	Set the droplet used to run an app
<a href="#">v3-env</a>	Show all env variables for an app
<a href="#">v3-set-env</a>	Set an env variable for an app
<a href="#">v3-unset-env</a>	Remove an env variable from an app
<a href="#">v3-get-health-check</a>	Show the type of health check performed on an app
<a href="#">v3-set-health-check</a>	Change type of health check performed on an app's process
<a href="#">v3-packages</a>	List packages of an app
<a href="#">v3-create-package</a>	Uploads a V3 Package
<a href="#">v3-ssh</a>	SSH to an application container instance

## Using Experimental cf CLI Commands

This topic describes how to use the experimental Cloud Foundry Command Line Interface (cf CLI) commands offered by the Cloud Controller V3 API.

These commands provide developers with the ability to better orchestrate app deployment workflows. New features include the deployment and management of apps with multiple processes and uploading and staging multiple versions of a single app.

The experimental commands described in this topic require Pivotal Cloud Foundry (PCF) 1.12+ and the cf CLI v6.30.0+.

For more information about cf CLI commands, see the [Cloud Foundry CLI Reference Guide](#). For more information about the Cloud Controller V3 API, see the [API documentation](#).

 **Note:** Because these commands are experimental, they are not guaranteed to be available or compatible in subsequent cf CLI releases.

## Overview

The new commands include a `v3-` prefix. While the syntax of some experimental commands is based on the existing cf CLI, these commands call the V3 API and support new flags to unlock additional features. Other commands expose the new primitives of apps, such as by performing operations on an app's packages and droplets.

In the V2 APIs, running and staging an app are tightly coupled operations. As a result, an app cannot be staging and running at the same time. The V3 APIs offer developers more granular control over the uploading, staging, and running of an app.

## Commands

Consult the following table for a description of the experimental commands.

Command	Description
<code>v3-apply-manifest</code>	Applies manifest properties to an app
<code>v3-apps</code>	Lists all apps in the target space
<code>v3-create-app</code>	Creates a V3 app
<code>v3-create-package</code>	Uploads a V3 package
<code>v3-delete</code>	Deletes a V3 app
<code>v3-droplets</code>	Lists droplets of an app
<code>v3-env</code>	Shows all environment variables for an app
<code>v3-get-health-check</code>	Shows the type of health check performed on an app
<code>v3-packages</code>	Lists packages of an app
<code>v3-push</code>	Pushes a new app or syncs changes to an existing app
<code>v3-restart</code>	Stops all instances of an app and then starts them again, which may cause downtime
<code>v3-restart-app-instance</code>	Terminates and then instantiates an app instance
<code>v3-set-droplet</code>	Sets the droplet used to run an app
<code>v3-set-env</code>	Sets an environment variable for an app
<code>v3-set-health-check</code>	Changes type of health check performed on an app's process
<code>v3-ssh</code>	SSH to an app container instance
<code>v3-stage</code>	Creates a new droplet for an app
<code>v3-start</code>	Starts an app
<code>v3-stop</code>	Stops an app
<code>v3-unset-env</code>	Removes an environment variable from an app

## Considerations

Keep in mind the following considerations when using the experimental commands:

- `v3-push` supports only a subset of features of `push`. In particular, it does not support app manifests.
- For some commands, such as `set-env`, `ssh`, and `bind-service`, no new V3 version exists. In those cases, use the old commands.
- You can use V3 and old commands together, but some combinations may give unexpected results. For example, if you use V3 commands to create an app with a package but it is not staged, or you use `v3-push` to push an app but it fails to stage, the old `apps` command does not return the app.

## Declaring Multiple App Processes with a Procfile

Developers can supply a Procfile with their app in order to run multiple long-lived processes using a single codebase.

To use a Procfile, include it in the root of your app directory and push your app.

For more information about Procfiles, see the [About Procfiles](#) section of the *Production Server Configuration* topic.

## Developer Guide

This guide provides instructions for deploying and managing apps and services.

 Check out the 15-minute [Getting Started with PCF](#) tutorial for learning Pivotal Cloud Foundry app deployment concepts.

See the following topics:

- [cf push](#)
- [SSH for Apps and Services](#)
- [Routes and Domains](#)
- [Managing Services](#)
- [Streaming App Logs](#)
- [Managing Apps with the cf CLI](#)
- [Cloud Foundry Environment Variables](#)
- [Cloud Controller API Client Libraries](#)
- [Considerations for Designing and Running an App in the Cloud](#)
- [Revisions](#)

## cf push

This topic provides an overview of the documentation available for `cf push`.

## How to Use cf push

The following topics provide procedures for deploying apps with `cf push`:

- [Pushing an App](#)
- [Deploying with App Manifests](#)
  - [App Manifest Attribute Reference](#)
- [Deploying an App with Docker](#)
- [Deploying a Large App](#)
- [Starting, Restarting, and Restaging Apps](#)
- [Pushing an Application with Multiple Buildpacks](#)
- [Pushing an App with Multiple Processes \(Beta\)](#)
- [Running cf push Sub-Step Commands \(Beta\)](#)
- [Rolling App Deployments \(Beta\)](#)
- [Pushing Apps with Sidecar Processes \(Alpha\)](#)

## How to Troubleshoot

For information about troubleshooting when running `cf push`, see [Troubleshooting App Deployment and Health](#).

## How cf push Works

The following topics provide information about how `cf push` works:

- The [Push](#) section of *Getting Started with the cf CLI*.
- [App Container Lifecycle](#)
- [How Apps Are Staged](#)

## Pushing an App

This topic describes how to use the Cloud Foundry Command Line Interface (cf CLI) to push an app to Pivotal Application Service (PAS).

## Overview

The cf CLI command `cf push` pushes apps to PAS. There are two main ways to run the `cf push` command:

- Run `cf push APP-NAME` to push an app the easiest way, using default settings.
- Run the `cf push` command with flags and helper files to customize:
  - How the pushed app runs, including its route (URL), instance count, and memory limits.
  - How the push process works: whether it's configured with a manifest, runs a startup script, or limits files uploaded to the Cloud Controller.

The [Push with Defaults](#) and [Push with Custom Settings](#) sections below detail both of these options.

For an explanation of what PAS does when you run `cf push`, see [How Apps are Staged](#).

For information about the lifecycle of an app, see [App Container Lifecycle](#).

## Prerequisites

Before you push your app to PAS, make sure that:

- Your app is *cloud-ready*. PAS behaviors related to file storage, HTTP sessions, and port usage may require modifications to your app. For help preparing your app to be pushed to PAS, see:
  - [Considerations for Designing and Running an App in the Cloud](#)
  - Any [Buildpacks](#) guides specific to your app language or framework, such as [Getting Started Deploying Ruby on Rails Apps](#)
- Your PAS installation supports the type of app you are going to push, or you have the URL of an externally-available buildpack that can stage the app.
- All required app resources are uploaded. For example, you may need to include a database driver.
- You have your target and credentials:
  - The API endpoint for your PAS installation. Also known as the target URL, this is [the URL of the Cloud Controller in your PAS instance](#).
  - Your username and password for your PAS installation.
- You are logged into your app's target org and space.
  1. Decide the org and space where you want to push your app. You may have access to one or more org and space.
  2. Log into this target org and space with `cf login`.
- Your app can access every service that it uses, because an instance of the service runs in, or is shared with, the app's space.
  - See [Sharing Service Instances](#) for how to share service instances across spaces.
  - Typical services that cloud apps use include databases, message queues, and key-value stores.

## Push with Defaults

To push an app with default settings, do the following:

1. Choose a name for the app.
  - The app name must consist of alphanumeric characters.
  - The app name must be unique to your PAS installation.
    - To use an app name that is not unique, customize the app's route as described in [Customize the Route](#) below.
    - Apps running at their default routes require unique names because default routes are based on app names, and all routes must be globally unique.
2. Run the following command:

```
cf push APP-NAME
```

Where `APP-NAME` is the name of the app.

## Default Route

An app's route is the URL that it runs at. PAS assembles the route for a pushed app from a hostname and a domain.

By default, PAS sets the hostname and domain as follows:

- **Hostname:** The name of the app, as specified in the `cf push` command.
  - If the app name contains underscores, PAS converts them to hyphens when creating the app's route.
- **Domain:** The default apps domain for the PAS installation.

For example, an app named `my-app-1234` running on PAS with an apps domain `apps.example.com` would, by default, run at the URL `https://my-app-1234.apps.example.com`.

For more information about routes and domains, see [Routes and Domains](#).

## Example Session

The following session illustrates how PAS assigns default values to app when given a `cf push` command.

```
$ cf push my-app-1234
Creating app my-app-1234 in org example-org / space development as a.user@shared-domain.example.com...
OK

Creating route my-app-1234.shared-domain.example.com...
OK

Binding my-app-1234.shared-domain.example.com to my-app-1234...
OK

Uploading my-app-1234...
Uploading app: 560.1K, 9 files
OK

Starting app my-app-1234 in org example-org / space development as a.user@shared-domain.example.com...
----> Downloaded app package (552K)
OK
----> Using Ruby version: ruby-1.9.3
----> Installing dependencies using Bundler version 1.3.2
 Running: bundle install --without development:test --path vendor/bundle --binstubs vendor/bundle/bin --deployment
 Installing rack (1.5.1)
 Installing rack-protection (1.3.2)
 Installing tilt (1.3.3)
 Installing sinatra (1.3.4)
 Using bundler (1.3.2)
 Updating files in vendor/cache
 Your bundle is complete! It was installed into ./vendor/bundle
 Cleaning up the bundler cache.
----> Uploading droplet (23M)

1 of 1 instances running

App started

Showing health and status for app my-app-1234 in org example-org / space development as a.user@shared-domain.example.com...
OK

requested state: started
instances: 1/1
usage: 1G x 1 instances
urls: my-app-1234.shared-domain.example.com

 state since cpu memory disk
#0 running 2014-01-24 05:07:18 PM 0.0% 18.5M of 1G 52.5M of 1G
```

## Push with Custom Settings

Pushing an app with custom settings typically proceeds as follows:

1. [\(Optional\) Customize Basic App Settings](#)
2. [\(Optional\) Customize the Route](#)
3. [\(Optional\) Limit the Upload Files](#)
4. [\(Optional\) Configure App Initialization](#)
5. [Custom Push the App](#)
6. [\(Optional\) Configure App Services](#)

The sections below detail these steps.

### (Optional) Customize Basic App Settings

Basic settings to customize when pushing an app include:

- **Name:** You can use any series of alphanumeric characters as the name of your app.
- **Instances:** Generally speaking, the more app instances you run, the less downtime your app will experience. If your app is still in development, running a single instance can simplify troubleshooting. For any production app, Pivotal recommends a minimum of two instances.
- **Memory Limit:** The maximum amount of memory that each instance of your app can consume. If an instance exceeds this limit, PAS restarts the instance. If an instance exceeds its memory limit repeatedly in a short period of time, PAS delays restarting the instance.
- **Start Command:** This is the command that PAS uses to start each instance of your app. This start command differs by app framework.

### (Optional) Customize the Route

To customize an app's route, do the following:

1. [\(Optional\) Customize the Hostname](#)
  - Pass the custom hostname into `cf push` with the `-n` flag.
2. [\(Optional\) Customize the Domain](#)
  - Pass the custom domain into `cf push` with the `-d` flag. The custom domain must be registered, and mapped to the org that contains the app's target space
3. [Ensure Route Uniqueness](#)
  - The app's route must be globally unique, whether you customize its host or domain, or let it use the [Default Route](#) described above.
  - To help ensure route uniqueness, pass the `--random-route` option into `cf push`. `--random-route` creates a route that includes the app name and random words

### (Optional) Limit the Upload Files

By default, PAS uploads all app files except version control files and folders with names such as `.svn`, `.git`, and `_darcs`.

Pivotal recommends that you explicitly exclude extraneous files residing within your app directory, particularly if your app is large. For example, if you build your app locally and push it as a binary, you can save resources by not uploading any of the app's source code.

To exclude files from upload:

1. Create a `.cfignore` file that lists the files to exclude.
2. Save the `.cfignore` file to the directory where you run the `cf push` command.

For more information, see the [Ignore Unnecessary Files When Pushing](#) section of the [Considerations for Designing and Running an App in the Cloud](#) topic.

## (Optional) Configure App Initialization

 **Note:** The Java buildpack does not support an initialization script.

You can configure `cf push` to run custom initialization tasks for an app.

These tasks run after PAS loads the app droplet but before it launches the app itself to let the initialization script access the app language runtime environment. For example, your script can map values from `$VCAP_SERVICES` into other environment variables or a config file that the app uses.

To run initialization tasks:

1. Create a `.profile` script that contains the initialization tasks.
2. Save the `.profile` script to the directory where you run the `cf push` command.

As an example, the following `.profile` file uses `bash` to set a value for the environment variable `LANG`. Setting this value at the operating system level lets the app determine which language to use for error messages and instructions, collating sequences, and date formats:

```
Set the default LANG for your apps
export LANG=en_US.UTF-8
```

Your app root directory may also include a `.profile.d` directory that contains bash scripts that perform initialization tasks for the buildpack. Developers should not edit these scripts unless they are using a [custom buildpack](#).

 **Note:** If you are using a PHP buildpack version prior to v4.3.18, the buildpack does not execute your PHP app's `.profile` script. Your PHP app will host the `.profile` script's contents. This means that any PHP app staged using the affected PHP buildpack versions can leak credentials placed in the `.profile` script.

Initialization tasks as described here are also called “pre-runtime hooks” and “`.profile` tasks”.

## Custom Push the App

To specify custom options when pushing an app with `cf push`, you can:

- Include them in the `cf push APP-NAME` command itself.
- Include them in a manifest file.
  - The manifest file must be named `manifest.yml` and reside in the directory where you run `cf push`.
  - The manifest can include the app name, which lets you run `cf push` with no arguments.
  - The manifest can also include a `Services` block that lists service instances for the app to bind to automatically.
- Both of the above, command-line options and a manifest.

See [Deploying with App Manifests](#) to learn how app settings change from push to push, and how command-line options, manifests, and commands like `cf scale` interact.

See the [Cloud Foundry CLI Reference Guide](#) for a full list of `cf push` options.

## (Optional) Configure App Services

If a newly-pushed app has the same name and route as an older app version, the new app retains the service bindings and service configuration of the previously-pushed version.

For apps that are not already set up for the services that they use, you need to:

1. Bind the services to the app. For more information about services, see the [Services Overview](#) topic.
2. (Optional) Configure the app with the service URL and credentials, if needed. For more information, see the specific documentation for your app framework, such as:
  - [Ruby](#)
  - [Node.js](#)

- [Spring](#)
- [Grails](#)

## App Updates and Downtime

When you push an app that is already running, PAS stops all existing instances of that app. Users who try to access the app get a “404 Not Found” message while `cf push` runs.

With some app updates, old and new versions of your code should never run at the same time. A worst-case example is if your app update migrates a database schema, causing old app instances to fail and lose user data. To prevent this, you need to stop all running instances of your app before you push the new version.

When old and new versions of your app can run simultaneously, you can avoid app downtime by using the [blue-green method](#) to swap routes between app versions running in parallel.

## Troubleshoot App Push Problems

If your app does not start on PAS, first ensure that the app can run locally.

For how to troubleshoot your app in the cloud using the cf CLI, see [Troubleshoot App Deployment and Health](#).

## Deploying with App Manifests

This topic provides basic procedures and guidance for deploying apps with a manifest file.

### Overview

Manifests provide consistency and reproducibility, and can help you automate deploying apps. Both manifests and command line options allow you to override the default attribute values of `cf push`. These attributes include things like number of instances, disk space limit, and memory limit.

`cf push` follows rules of precedence when setting attribute values:

- Manifests override most recent values, including defaults and values set by commands such as `cf scale`.
- Command line options override manifests.

For a full list of attributes you can specify in an app manifest, see [App Manifest Attribute Reference](#).

### Deploy an App with a Manifest

The following procedure describes how to deploy an app with a minimal manifest. For more information about manifest format and attributes, see [App Manifest Attribute Reference](#).

To deploy an app with a manifest, do the following:

1. Create a `manifest.yml` file in the root directory of your app.

 **Note:** By default, the `cf push` command looks for a manifest file in the app directory. However, you can create a manifest file in a different location and use the `-f` flag when you run `cf push`

2. Add the following content to the file, replacing `MY-APP` with the name of your app:

```

applications:
- name: MY-APP
```

3. Run `cf push`. If you specify any command line options, they override the values specified in the manifest. See the following section.

### Deploy Multiple Apps with One Manifest

This procedure describes how to deploy multiple apps with a minimal manifest. For more information about manifest format and attributes, see [App Manifest Attribute Reference](#).

### General Rules

Follow these general rules when deploying multiple apps with one manifest:

- Use a `no-route` line in the description of any app that provides background services to another app.
- You cannot use any command line options with `cf push` except for `-f` and `--no-start`.
  - If your manifest is not named `manifest.yml` or not in the current working directory, use the `-f` command line option.
- To push a single app rather than all of the apps described in the manifest, provide the desired app name by running `cf push my-app`.

### Procedure

To deploy multiple apps with a manifest, do the following:

 **Note:** Each app must be in a subdirectory under the same parent directory.

1. Create a `manifest.yml` file in the directory that contains the apps.
2. Add each app and its directory to the file. CF pushes the apps in the order specified in the manifest.

 **Note:** If you push multiple apps using a manifest and one fails to deploy, CF does not attempt to push apps specified after the app that failed.

```

applications:
- name: APP-ONE
 path: ./APP-ONE-DIRECTORY
- name: APP-TWO
 path: ./APP-TWO-DIRECTORY
```

Where:

- o `APP-ONE` is the name of the first app you want CF to push
- o `APP-ONE-DIRECTORY` is the directory containing the first app
- o `APP-TWO` is the name of the second app you want CF to push
- o `APP-TWO-DIRECTORY` is the directory containing the second app

3. Run `cf push` from the directory that contains the apps and the manifest.

## App Manifest Attribute Reference

This topic describes manifest formatting and provides a full list of attributes available for app manifests. You can use it alongside the [Deploying with App Manifests](#) topic, which provides basic procedures and guidance for deploying apps with manifests.

### Manifest Format

Manifests are written in YAML. The manifest below illustrates some YAML conventions, as follows:

- The manifest begins with three dashes.
- The `applications` block begins with a heading followed by a colon.
- The app `name` is preceded by a single dash and one space.
- Subsequent lines in the block are indented two spaces to align with `name`.

```

applications:
- name: my-app
 memory: 512M
 instances: 2
```

 **Note:** If your app name begins with the dash character (-), you cannot interact with the app using the cf CLI. This is because the cf CLI interprets the dash as a flag.

### Add Variables to a Manifest

You can use variables to create app manifests with values shared across all applicable environments in combination with references to environment-specific differences defined in separate files.

In addition, using variables enables you to store sensitive data in a separate file that the app manifest would reference, making the security sensitive data easier to manage and keep secure.

To add variables to an app manifest, do the following:

1. Create a file called `vars.yml`.
2. Add attributes to your `vars.yml` file. See the following example:

```
instances: 2
memory: 1G
```

3. Add the variables to your app manifest file using the following format: `((VARIABLE-NAME))`. See the following example:

```

applications:
- name: test-app
 instances: ((instances))
 memory: ((memory))
 buildpacks:
 - go_buildpack
 env:
 GOPACKAGENAME: go_calls_ruby
 command: go_calls_ruby
```

 **Note:** You can also use variables for partial values. For example, you can specify `host` in your variables file and `- route: ((host)).env.com` in your manifest file.

4. Run `cf push`:

```
cf push --vars-file /PATH/vars.yml
```

Where `PATH` is the path to the file you created.

## Minimize Duplication with YAML Anchors

 **Note:** Top-level attributes have been deprecated in favor of YAML anchors. For more information, see [Deprecated App Manifest Features](#).

In manifests where multiple apps share settings or services, you may see duplicated content. While the manifests still work, duplication increases the risk of typographical errors, which cause deployments to fail.

You can declare shared configuration using a YAML anchor, which the manifest refers to in app declarations by using an alias.

```

```

```
defaults: &defaults
 buildpacks:
 - staticfile_buildpack
 memory: 1G

applications:
- name: bigapp
 <<: *defaults
- name: smallapp
 <<: *defaults
 memory: 256M
```

This manifest pushes two apps, smallapp and bigapp, with the staticfile buildpack but with 256M memory for smallapp and 1G for bigapp.

## Application Attributes

This section explains how to describe optional app attributes in manifests. Each of these attributes can also be specified by a command line option. Command line options override the manifest.

 **Note:** The route component attributes `domain`, `domains`, `host`, `hosts`, and `no-hostname` have been deprecated in favor of the `routes` attribute. For more information, see [Deprecated App Manifest Features](#).

## buildpacks

You can refer to a buildpack by name in a manifest or a command line option. The `cf buildpacks` command lists the buildpacks that you can use.

See below for information on referencing buildpacks in a manifest. The command line option that overrides this attribute is `-b`.

- **Custom buildpacks:** If your app requires a custom buildpack, you can use the `buildpacks` attribute to specify it in a number of ways:
  - By name: `MY-BUILDPACK`.
  - By GitHub URL: `https://github.com/cloudfoundry/java-buildpack.git`.
  - By GitHub URL with a branch or tag: `https://github.com/cloudfoundry/java-buildpack.git#v3.3.0` for the `v3.3.0` tag.

```

```

```
...
buildpacks:
- buildpack_URL
```

- **Multiple buildpacks:** If you are using multiple buildpacks, you can provide an additional `-b` flag or add an additional value to your manifest:

```

```

```
...
buildpacks:
- buildpack_URL
- buildpack_URL
```

 **Note:** This feature does not work with [Deprecated App Manifest Features](#).

 **Note:** You must specify multiple buildpacks in the correct order: the buildpack will use the app start command given by the final buildpack. See [the multi-buildpack repository](#) for more information.

Also see [Pushing an App with Multiple Buildpacks](#) for more information.

## command

Some languages and frameworks require that you provide a custom command to start an app. Refer to the [buildpack documentation](#) to determine if you need to provide a custom start command.

You can provide the custom start command in your app manifest or on the command line. See [Starting, Restarting, and Restaging Apps](#) for more information about how Cloud Foundry determines its default start command.

To specify the custom start command in your app manifest, add it in the `command: START-COMMAND` format as the following example shows:

```

...
command: bundle exec rake VERBOSE=true
```

The start command you specify becomes the default for your app. To return to using the original default start command set by your buildpack, you must explicitly set the attribute to `null` as follows:

```

...
command: null
```

On the command line, use the `-c` option to specify the custom start command as the following example shows:

```
$ cf push my-app -c "bundle exec rake VERBOSE=true"
```

 **Note:** The `-c` option with a value of 'null' forces `cf push` to use the buildpack start command. See [Forcing cf push to use the Buildpack Start Command](#) for more information.

If you override the start command for a Buildpack app, Linux uses `bash -c YOUR-COMMAND` to invoke your app. If you override the start command for a Docker app, Linux uses `sh -c YOUR-COMMAND` to invoke your app. Because of this, if you override a start command, you should prefix `exec` to the final command in your custom composite start command.

An app needs to catch [termination signals](#) and clean itself up appropriately. Because of the way that shells manage process trees, the use of custom composite shell commands, particularly those that create child processes using `&`, `&&`, `||`, etc., can prevent your app from receiving signals that are sent to the top level bash process.

To resolve this issue, you can use `exec` to replace the bash process with your own process. For example:

- `bin/rake cf:on_instance db:migrate && bin/rails server -p $PORT -e $RAILS_ENV` The process tree is bash -> ruby, so on graceful shutdown only the bash process receives the TERM signal, not the ruby process.
- `bin/rake cf:on_instance db:migrate && exec bin/rails server -p $PORT -e $RAILS_ENV` Because of the `exec` prefix included on the final command, the ruby process invoked by rails takes over the bash process managing the execution of the composite command. The process tree is only ruby, so the ruby web server receives the TERM signal and can shutdown gracefully for 10 seconds.

In more complex situations, like making a custom buildpack, you may want to use bash `trap`, `wait`, and backgrounded processes to manage your process tree and shut down apps gracefully. In most situations, however, a well-placed `exec` should be sufficient.

## disk\_quota

Use the `disk_quota` attribute to allocate the disk space for your app instance. This attribute requires a unit of measurement: `M`, `MB`, `G`, or `GB`, in upper case or lower case.

```

```

```
...
disk_quota: 1024M
```

The command line option that overrides this attribute is `-k`.

## docker

If your app is contained in a Docker image, then you may use the `docker` attribute to specify it and an optional Docker username.

This attribute is a combination of `push` options that include `--docker-image` and `--docker-username`.

```

...
docker:
 image: docker-image-repository/docker-image-name
 username: docker-user-name
```

The command line option `--docker-image` or `-o` overrides `docker.image`. The command line option `--docker-username` overrides `docker.username`.

The manifest attribute `docker.username` is optional. If it is used, then the password must be provided in the environment variable `CF_DOCKER_PASSWORD`. Additionally, if a Docker username is specified, then a Docker image must also be specified.

 **Note:** Using the `docker` attribute in conjunction with the `buildpacks` or `path` attributes will result in an error.

## health-check-http-endpoint

Use the `health-check-http-endpoint` attribute to customize the endpoint for the `http` health check type. If you do not provide a `health-check-http-endpoint` attribute, it uses endpoint '/'.

```

...
health-check-type: http
health-check-http-endpoint: /health
```

## health-check-type

Use the `health-check-type` attribute to set the `health_check_type` flag to either `port`, `process` or `http`. If you do not provide a `health-check-type` attribute, it defaults to `port`.

```

...
health-check-type: port
```

The command line option that overrides this attribute is `-u`.

The value of `none` is deprecated in favor of `process`.

## memory

Use the `memory` attribute to specify the memory limit for all instances of an app. This attribute requires a unit of measurement: `M`, `MB`, `G`, or `GB`, in upper case or lower case. For example:

```

...
memory: 1024M
```

The default memory limit is 1G. You might want to specify a smaller limit to conserve quota space if you know that your app instances do not require 1G of memory.

The command line option that overrides this attribute is `-m`.

## no-route

By default, `cf push` assigns a route to every app. But, some apps process data while running in the background and should not be assigned routes.

You can use the `no-route` attribute with a value of `true` to prevent a route from being created for your app.

```

...
no-route: true
```

The command line option that overrides this attribute is `--no-route`.

In the Diego architecture, `no-route` skips creating and binding a route for the app, but does not specify which type of health check to perform. If your app does not listen on a port because it is a worker or a scheduler app, then it does not satisfy the port-based health check and Cloud Foundry marks it as crashed. To prevent this, disable the port-based health check with `cf set-health-check APP_NAME process`.

To remove a route from an existing app, perform the following steps:

1. Remove the route using the `cf unmap-route` command.
2. Push the app again with the `no-route: true` attribute in the manifest or the `--no-route` command line option.

For more information, see [Deploy Multiple Apps with One Manifest](#).

## path

You can use the `path` attribute to tell Cloud Foundry the directory location where it can find your app.

The directory specified as the `path`, either as an attribute or as a parameter on the command line, becomes the location where the buildpack `Detect` script executes.

The command line option that overrides this attribute is `-p`.

```

...
path: /path/to/app/bits
```

For more information, see [How cf push Finds the App](#).

## random-route

If you push your app without specifying any route-related CLI options or app manifest flags, the `cf` CLI attempts to generate a route based on the app name, which can cause collisions.

You can use the `random-route` attribute to generate a unique route and avoid name collisions. When you use `random-route`, the `cf` CLI generates an HTTP route with a random host (if `host` is not set) or a TCP route with an unused port number.

See the following example use cases:

- You deploy the same app to multiple spaces for testing purposes. In this situation, you can use `random-route` to randomize routes declared with the `route` attribute in the app manifest.
- You use an app manifest for a classroom training exercise in which multiple users deploy the same app to the same space.

The command line option that overrides this attribute is `--random-route`.

```

...
random-route: true
```

## routes

Use the `routes` attribute to provide multiple HTTP and TCP routes. Each route for this app is created if it does not already exist.

This attribute is a combination of `push` options that include `--hostname`, `-d`, and `--route-path`.

```

...
routes:
- route: example.com
- route: www.example.com/foo
- route: tcp-example.com:1234
```

## Manifest Attributes

The `routes` attribute cannot be used in conjunction with the following attributes: `host`, `hosts`, `domain`, `domains`, and `no-hostname`. An error will result.

## Push Flag Options

This attribute has unique interactions with different command line options.

Push Flag Option	Resulting Behaviour
<code>--no-route</code>	All declared routes are ignored.
<code>-d</code>	Overrides DOMAIN part of all declared HTTP and TCP routes.
<code>--hostname</code> , <code>-n</code>	Sets or overrides HOSTNAME in all HTTP routes. It has no impact on TCP routes.
<code>--route-path</code>	Sets or overrides the PATH in all HTTP routes. It has no impact on TCP routes.
<code>--random-route</code>	Sets or overrides the HOSTNAME in all HTTP routes. Sets or overrides the PORT in all TCP routes. The PORT and HOSTNAME will be randomly generated.

## stack

Use the `stack` attribute to specify which stack to deploy your app to.

To see a list of available stacks, run `cf stacks` from the cf CLI.

```

...
stack: cflinuxfs3
```

The command line option that overrides this attribute is `-s`.

## timeout

The `timeout` attribute defines the number of seconds that Cloud Foundry allocates for starting your app. It is related to the `health-check-type` attribute.

For example:

```

...
timeout: 80
```

You can increase the timeout length for very large apps that require more time to start. The `timeout` attribute defaults to 60, but you can set it to any value up to the Cloud Controller's `cc.maximum_health_check_timeout` property.

`cc.maximum_health_check_timeout` defaults to 180.

The command line option that overrides the timeout attribute is `-t`.

**warning:** Cloud Controller reports a validation error with the maximum limit if you to configure `timeout` with a value greater than it.

## Environment Variables

The `env` block consists of a heading, then one or more environment variable/value pairs.

For example:

```

...
env:
 RAILS_ENV: production
 RACK_ENV: production
```

`cf push` deploys the app to a container on the server. The variables belong to the container environment.

**Note:** You must name variables with alphanumeric characters and underscores. Non-conforming variable names may cause unpredictable behavior.

While the app is running, you can modify environment variables.

- View all variables: `cf env my-app`
- Set an individual variable: `cf set-env my-app my-variable_name my-variable_value`
- Unset an individual variable: `cf unset-env my-app my-variable_name my-variable_value`

Environment variables interact with manifests in the following ways:

- When you deploy an app for the first time, Cloud Foundry reads the variables described in the environment block of the manifest and adds them to the environment of the container where the app is staged, and the environment of the container where the app is deployed.
- When you stop and then restart an app, its environment variables persist.

## Services

Apps can bind to services such as databases, messaging, and key-value stores.

Apps are deployed into App Spaces. An app can only bind to services instances that exist in the target App Space before the app is deployed.

The `services` block consists of a heading, then one or more service instance names.

Whoever creates the service chooses the service instance names. These names can convey logical information, as in `backend_queue`, describe the nature of the service, as in `mysql_5.x`, or do neither, as in the example below.

```

...
services:
 - instance_ABC
 - instance_XYZ
```

Binding to a service instance is a special case of setting an environment variable, namely `VCAP_SERVICES`. See the [Bind a Service](#) section of the [Delivering Service Credentials to an App](#) topic.

## Deprecated App Manifest Features

These app manifest features have been deprecated in favor of other options, as described below.

**warning:** Running `cf push app -f manifest.yml` fails if your manifest uses any of these deprecated features along with the feature that replaces it.

## YAML Anchors Replace Top-Level Attributes

Previously, you could declare top-level attributes, also known as global attributes. For example, you could move an attribute above the `applications` block, where it need appear only once.

The following example illustrates how this was used to manage duplicated settings.

```

all apps use these settings and services
domain: shared-domain.example.com
memory: 1G
instances: 1
services:
- clockwork-mysql
applications:
- name: springtock
 host: tock09876
 path: ./spring-music/build/libs/spring-music.war
- name: springtick
 host: tick09875
 path: ./spring-music/build/libs/spring-music.war
```

Now, you can use YAML aliases instead.

The following example illustrates how to declare shared configuration using a YAML anchor, which the manifest refers to in app declarations by using an alias.

```

defaults: &defaults
 buildpacks:
 - staticfile_buildpack
 memory: 1G

applications:
- name: bigapp
 <<: *defaults
- name: smallapp
 <<: *defaults
 memory: 256M
```

When pushing the app, make explicit the attributes in each app's declaration. To do this, assign the anchors and include the app-level attributes with YAML aliases in each app declaration.

## Attribute routes Replaces domain, domains, host, hosts, and no-hostname

Previously, you could specify routes by listing them all at once using the `routes` attribute, or by using their hosts and domains as shown below.

```

applications
- name: webapp
 host: www
 domains:
 - example.com
 - example.io
```

The following route component attributes have been deprecated:

- `domain`
- `domains`
- `host`
- `hosts`
- `no-hostname`

Now you can only specify routes by using the `routes` attribute:

```

applications
- name: webapp
 routes:
 - route: www.example.com/foo
 - route: tcp.example.com:1234
```

This app manifest feature has been deprecated, and a replacement option is under consideration.

## Inheritance

This feature has been deprecated, and has been replaced by [Variable Substitution](#).

With inheritance, child manifests inherited configurations from a parent manifest, and the child manifests could use inherited configurations as provided, extend them, or override them. This feature has been deprecated, and has been replaced by [Variable Substitution](#).

## Buildpack Field in Manifest Is Deprecated

The singular `buildpack` field in manifests is deprecated. It has been replaced by `buildpacks`, which is now an array which takes as a value [multiple buildpacks](#).

## Deploying an App with Docker

This topic describes how to use the [Cloud Foundry Command Line Interface \(cf CLI\)](#) to push an app with a new or updated Docker image. Cloud Foundry then uses the Docker image to create containers for the app.

See [Using Docker in Cloud Foundry](#) in the Cloud Foundry documentation for an explanation of how Docker works in Cloud Foundry.

## Requirements

To push apps with Docker, you need the following:

- A Cloud Foundry deployment that has Docker support enabled. To enable Docker support, see the [Enable Docker](#) section of [Using Docker in Cloud Foundry](#) in the Cloud Foundry documentation.
- A Docker image that meets the following requirements:
  - The Docker image must contain an `/etc/passwd` file with an entry for the `root` user. In addition, the home directory and the shell for that `root` user must be present in the image file system.
  - The total size of the Docker image file system layers must not exceed the disk quota for the app. The maximum disk allocation for apps is set by the Cloud Controller. The default maximum disk quota is 2048 MB per app.

 **Note:** If the total size of the Docker image file system layers exceeds the disk quota, the app instances do not start.

- The location of the Docker image on [Docker Hub](#) or another Docker registry.
- A registry that supports the [Docker Registry HTTP API V2](#) and presents a valid certificate for HTTPS traffic.

## Requirement for cf ssh Support

If you want to log in to your app container using the `cf ssh` command, you must make a shell such as `sh` or `bash` available in the container.

The SSH server in the container looks for the following executables in absolute locations or the `PATH` environment variable:

- `/bin/bash`
- `/usr/local/bin/bash`
- `/bin/sh`
- `bash`
- `sh`

## Benefits of Specifying Tags

If you want your app container to be consistent after platform updates and code changes, specify a tag when you push your Docker image. Otherwise, the platform applies the `latest` tag without respecting changes to `PORT` or `ENTRYPOINT`.

If you push your Docker image without specifying a tag, you must run `cf restage` for the changes to take effect.

## Port Configuration

By default, apps listen for connections on the port specified in the `PORT` environment variable for the app. Cloud Foundry allocates this value dynamically.

When configuring a Docker image for Cloud Foundry, you can control the exposed port and the corresponding value of `PORT` by specifying the `EXPOSE` directive in the image Dockerfile. If you specify the `EXPOSE` directive, then the corresponding app pushed to Cloud Foundry listens on that exposed port. For example, if you set `EXPOSE` to `7070`, then the app listens for connections on port 7070.

If you do not specify a port in the `EXPOSE` directive, then the app listens on the value of the `PORT` environment variable as determined by Cloud Foundry.

If you set the `PORT` environment variable via an `ENV` directive in a Dockerfile, Cloud Foundry overrides the value with the system-determined value.

Cloud Foundry supports only one exposed port on the image.

## Start Command

By default, Docker uses the start command specified in the Docker image. You can override the start command either by using a command-line parameter or by specifying it in a manifest file.

For more information about command-line parameters for `docker start`, see [docker start](#) in the Docker Documentation.

## Push a Docker Image From Docker Hub

To deploy a Docker image from a Docker Hub repository, run `cf push APP-NAME --docker-image REPO/IMAGE:TAG`. Replace the placeholder values in the

command as follows:

- `APP-NAME` : The name of the app being pushed
- `REPO` : The name of the repository where the image is stored
- `IMAGE` : The name of an image from Docker Hub
- `TAG` : (Optional, recommended) The tag or version for the image

For example, the following command pushes the `my-image` image from Docker Hub to a Cloud Foundry app:

```
$ cf push my-app --docker-image cloudfoundry/my-image
```

## Push a Docker Image from a Private Registry

As an alternative to Docker Hub, you can use any Docker image registry that presents a valid certificate for HTTPS traffic, such as a company-internal Docker registry.

To deploy a Docker image using a specified Docker registry, run `cf push APP-NAME --docker-image MY-PRIVATE-REGISTRY.DOMAIN:PORT/REPO/IMAGE:TAG`.

Replace the placeholder values in the command as follows:

- `APP-NAME` : The name of the app being pushed
- `MY-PRIVATE-REGISTRY.DOMAIN` : The path to the Docker registry
- `PORT` : The port where the registry serves traffic
- `REPO` : The name of the repository where the image is stored
- `IMAGE` : The name of the image being pushed
- `TAG` : (Optional, but recommended) The tag or version for the image

For example, the following command pushes the `v2` version of the `my-image` image from the `my-repo` repository of the `internal-registry.example.com` registry on port `5000`:

```
$ cf push my-app --docker-image internal-registry.example.com:5000/my-repo/my-image:v2
```

## Push a Docker Image From a Registry with Authentication

Many Docker registries control access to Docker images by authenticating with a username and password. Follow the steps below to deploy a Docker image with registry authentication:

1. Make sure the `CF_DOCKER_PASSWORD` environment variable is set to the Docker registry user password.

2. Run the following command:

```
$ CF_DOCKER_PASSWORD=YOUR-PASSWORD cf push APP-NAME --docker-image REPO/IMAGE:TAG --docker-username USER
```

Replace the placeholder values in the command as follows:

- `YOUR-PASSWORD` : The password to use for authentication with the registry
- `APP-NAME` : The name of the app being pushed
- `REPO` : The repository where the image is stored:
  - For Docker Hub, this is just the repository name
  - For a private registry, this includes the registry address and port, as described in [Push a Docker Image from a Private Registry](#), in the format: `MY-PRIVATE-REGISTRY.DOMAIN:PORT/REPO`
- `IMAGE` : The name of the image being pushed
- `TAG` : (Optional, but recommended) The tag or version for the image
- `USER` : The username to use for authentication with the registry

## Push a Docker Image From Google Container Registry (GCR)

PCF supports pushing apps from images hosted on Google Container Registry (GCR) service. This feature requires that you use [json\\_key based authentication](#).

### Step 1: Authenticate with GCR

To authenticate with GCR using PCF, you must create a JSON key file and associate it with your project:

1. Create a GCP service account. See [Creating and Enabling Service Accounts for Instances](#).

```
$ gcloud iam service-accounts create MY-ACCOUNT --display-name "MY DISPLAY NAME"
```

2. Create a JSON key file and associate it with the service account:

```
$ gcloud iam service-accounts keys create key.json --iam-account=MY-ACCOUNT@MY-PROJECT-ID.iam.gserviceaccount.com
```

3. Set your project ID:

```
$ gcloud config set project MY-PROJECT-ID
```

4. Add the IAM policy binding for your project and service account:

```
gcloud projects add-iam-policy-binding MY-PROJECT --member serviceAccount:MY-ACCOUNT@MY-PROJECT-ID.iam.gserviceaccount.com --role roles/storage.objectViewer
```

### Step 2: Deploy the GCP Image

Run the following command to deploy your GCR image using the cf CLI:

```
CF_DOCKER_PASSWORD="$(cat key.json)" cf push APP-NAME --docker-image docker://MY-REGISTRY-URL/MY-PROJECT/MY-IMAGE-NAME --docker-username _json_key
```

Replace the placeholder values in the command as follows:

- `APP-NAME` : The name of the app being pushed
- `MY-REGISTRY-URL` : The URL of your registry
- `MY-PROJECT` : The name of your project
- `MY-IMAGE-NAME` : The name of your image

 **Note:** The `key.json` file must point to the file you created in the previous step.

 **Note:** For information about specifying `MY-REGISTRY-URL`, see [Pushing and Pulling Images](#) on the Google Cloud documentation.

## Docker Volume Support

You can use volume services with Docker apps. For more information about enabling volume support, see the [Using an External File System \(Volume Services\)](#) topic.

## Deploying a Large App

This topic describes constraints and recommended settings for deploying apps above 750 MB.

## Deployment Considerations and Limitations

The deployment process involves uploading, staging, and starting the app. See the [Deployment](#) section of the App Container Lifecycle topic for more information about the default time limits for uploading, staging, and starting an app.

To deploy large apps to PAS, ensure the following:

- The total size of the files to upload for your app does not exceed the maximum app file size that an admin sets in [Ops Manager > PAS > Application Developer Controls](#).
- Your network connection speed is sufficient to upload your app within the 15 minute limit. We recommend a minimum speed of 874 KB/s.

 **Note:** PAS provides an authorization token that is valid for a minimum of 20 minutes.

- You allocate enough memory for all instances of your app. Use either the `-m` flag with `cf push` or set an app memory value in your `manifest.yml` file.
- You allocate enough disk space for all instances of your app. Use either the `-k` flag with `cf push` or set a disk space allocation value in your `manifest.yml` file.
- If you use an app manifest file, `manifest.yml`, be sure to specify adequate values for your app for attributes such as app memory, app start timeout, and disk space allocation.  
For more information about using manifests, refer to the [Deploying with App Manifests](#) topic.
- You push only the files that are necessary for your app.  
To meet this requirement, push only the directory for your app, and remove unneeded files or use the `.cfignore` file to [specify excluded files](#).
- You configure Cloud Foundry Command Line Interface (cf CLI) staging, startup, and timeout settings to override settings in the manifest, as necessary.
  - `CF_STAGING_TIMEOUT` : Controls the maximum time that the cf CLI waits for an app to stage after Cloud Foundry successfully uploads and packages the app. Value set in minutes.
  - `CF_STARTUP_TIMEOUT` : Controls the maximum time that the cf CLI waits for an app to start. Value set in minutes.
  - `cf push -t TIMEOUT` : Controls the maximum time that Cloud Foundry allows to elapse between starting an app and the first healthy response from the app. When you use this flag, the cf CLI ignores any app start timeout value set in the manifest. Value set in seconds.

For more information about using the cf CLI to deploy apps, refer to the [Push section](#) of the [Getting Started with the cf CLI](#) topic.

 **Note:** Changing the timeout setting for the cf CLI does not change the timeout limit for Cloud Foundry server-side jobs such as staging or starting apps. You must change server-side timeouts in the manifest. Because of the differences between the Cloud Foundry and cf CLI timeout values, your app might successfully start even though the cf CLI reports `App failed`. Run `cf apps APP_NAME` to review the actual status of your app.

## Default Settings and Limitations Summary Table

This table provides summary information of constraints and default settings to consider when you deploy a large app to PAS.

Setting	Note
App Package Size	Maximum: Set in <a href="#">Ops Manager &gt; PAS &gt; Application Developer Controls</a>
Authorization Token Grace Period	Default: 20 minutes, minimum
<code>CF_STAGING_TIMEOUT</code>	cf CLI environment variable Default: 15 minutes
<code>CF_STARTUP_TIMEOUT</code>	cf CLI environment variable Default: 5 minutes
<code>cf push -t TIMEOUT</code>	App start timeout maximum Default: 60 seconds
Disk Space Allocation	Default: 1024 MB
Internet Connection Speed	Recommended Minimum: 874 KB/s



## Starting, Restarting, and Restaging Apps

This topic describes how to start, restart, and restage apps in Cloud Foundry.

### Start Your App

To start your app, run the following command from your app root directory:

```
$ cf push YOUR-APP
```

For more information about pushing apps, see the [Pushing an App](#) topic.

Cloud Foundry determines the start command for your app from one of the three following sources:

- The `-c` command-line option in the Cloud Foundry Command Line Interface (cf CLI). See the following example:

```
$ cf push YOUR-APP -c "node YOUR-APP.js"
```

- The `command` attribute in the app manifest. See the following example:

```
command: node YOUR-APP.js
```

- The buildpack, which provides a start command appropriate for a particular type of app.

The source that Cloud Foundry uses depends on factors explained below.

### How Cloud Foundry Determines its Default Start Command

The first time you deploy an app, `cf push` uses the buildpack start command by default. After that, `cf push` defaults to whatever start command was used for the previous push.

To override these defaults, provide the `-c` option, or the `command` attribute in the manifest. When you provide start commands both at the command line and in the manifest, `cf push` ignores the command in the manifest.

### Forcing Cloud Foundry To Use the Buildpack Start Command

To force Cloud Foundry to use the buildpack start command, specify a start command of `null`.

You can specify a null start command in one of two ways.

- Using the `-c` command-line option in the cf CLI:

```
$ cf push YOUR-APP -c "null"
```

- Using the `command` attribute in the app manifest:

```
command: null
```

This can be helpful after you have deployed while providing a start command at the command line or the manifest. At this point, a command that you provided, rather than the buildpack start command, has become the default start command. In this situation, if you decide to deploy using the buildpack start command, the `null` command makes that easy.

### Start Commands When Migrating a Database

Start commands are used in special ways when you migrate a database as part of an app deployment. For more information, see [Services Overview](#).

### Restart Your App

To restart your app, run the following command:

```
$ cf restart YOUR-APP
```

Restarting your app stops your app and restarts it with the already compiled droplet. A droplet is a tarball that includes:

- stack
- [buildpack ↗](#)
- app source code

The Diego [cell ↗](#) unpacks, compiles, and runs a droplet on a container.

Restart your app to refresh the app's environment after actions such as binding a new service to the app or setting an environment variable that only the app consumes. However, if your environment variable is consumed by the buildpack in addition to the app, then you must [restage](#) the app for the change to take effect.

## Restage Your App

To restage your app, run the following command:

```
$ cf restage YOUR-APP
```

Restaging your app stops your app and restages it, by compiling a new droplet and starting it.

Restage your app if you have changed the environment in a way that affects your staging process, such as setting an environment variable that the buildpack consumes. The staging process has access to environment variables, so the environment can affect the contents of the droplet.

Restaging your app compiles a new droplet from your app without updating your app source. If you must update your app source, re-push your app by following the steps in the section [above](#).

## Pushing an App with Multiple Processes (Beta)

This topic describes how to push and manage apps with multiple processes.

 **Note:** This is a beta feature. Because the CLI may change, this feature is not recommended for use in scripts until it is generally available.

### About Processes

The CAPI v3 API supports using a single command to push apps that run multiple processes, such as a web app that has a UI process and a worker process. To push an app that creates multiple processes from the same codebase, you first define its processes in a Procfile. See [Push an App with Multiple Processes](#).

For more information about processes, see the [Processes](#) section of the CAPI v3 documentation.

### Push an App with Multiple Processes

You can push an app with multiple processes using a Procfile. Follow the procedure below.

1. Create a file named `Procfile` in the root of your app directory.

 **Note** For more information about Procfiles, see the [Procfiles](#) section of the CAPI v3 documentation.

2. Add each process and its start command to the Procfile. See the following example:

```
web: bundle exec rackup config.ru -p $PORT
worker: bundle exec rake worker:start
```

3. Run the `v3-push` command:

```
cf v3-push myapp
```

By default, the web process has a route and one instance. Other processes have zero instances by default.

### Scale a Process

To scale an app process, run the following command:

```
cf v3-scale APP-NAME --process PROCESS-NAME -i INSTANCE-COUNT
```

### View Processes

To view the processes running as part of an app, run `cf APP-NAME`. See the following example in which the app has a `web` and a `worker` process.

```
$ cf v3-app myapp
Showing health and status for app myapp in org test / space test as admin...
```

```
name: myapp
requested state: started
processes: web:1/1, worker:2/2
memory usage: 256M x 1, 256M x 2
routes: myapp.cloudfoundry.example.com
stack: cflinuxfs3
buildpacks: ruby

web:1/1
state since cpu memory disk
#0 running 2017-09-25 15:43:26 PM 0.2% 18.9M of 256M 84.4M of 1G
```

```
worker:2/2
state since cpu memory disk
#0 running 2017-09-25 15:49:46 PM 0.1% 5M of 256M 73M of 1G
#1 running 2017-09-25 15:49:46 PM 0.1% 5M of 256M 73M of 1G
```

## Running cf push Sub-Step Commands (Beta)

This topic describes how to run beta Cloud Foundry Command Line Interface (cf CLI) commands that provide granular control over the `cf push` process. These commands break down the `cf push` process into sub-steps that can run independently.

 **Note:** This is a beta feature. Because the CLI may change, this feature is not recommended for use in scripts until it is generally available.

## Overview

The cf CLI includes commands to provide granular control over app pushes. With these commands, you can choose to perform only some steps of the `cf push` process or perform specific actions between the steps normally executed as part of running `cf push`.

Here are some example use cases for the sub-step commands:

- Updating a third party system before staging an app
- Retrying failed stagings without incurring downtime
- Calling external services to report audit data during push
- Scanning a droplet before deploy
- Integrating with a change request system

To support these custom push workflows, Cloud Foundry divides apps into smaller building blocks. The following table describes these building blocks as *Resources* and lists the command associated with each one.

For information on using these commands, see the [Example Workflows](#) section below.

Resource	Description	Command
App	The top-level resource that represents an app and its configuration. For more information, see <a href="#">Apps</a> .	<code>v3-create-app</code>
Package	The source code that makes up an app. For more information, see <a href="#">Packages</a> .	<code>v3-create-package</code>
Build	The staging process. Creating a build combines a Package with a Buildpack and builds it into an executable resource. For more information, see <a href="#">Builds</a> .	<code>v3-stage</code>
Droplet	An executable resource that results from a Build. For more information, see <a href="#">Droplet</a> .	<code>v3-set-droplet</code>

## Example Workflows

The following sections describe example workflows for working with the `cf push` sub-step commands.

### Push an App Using Sub-Step Commands

This example workflow describes how to push an app using sub-step commands instead of `cf push`.

1. Create your app with the cf CLI:

```
cf v3-create-app MY-APP
```

2. From your app directory, create a package for your app.

```
cf v3-create-package MY-APP
```

3. Locate and copy the `package guid` from the output of the previous step. See the following example output:

```
Uploading and creating bits package for app MY-APP in org test / space test as admin...
package guid: 0dfca85a-8ed4-4f00-90d0-3ab08852dba8
OK
```

#### 4. Stage the package you created:

```
cf v3-stage MY-APP --package-guid PACKAGE-GUID
```

#### 5. Locate and copy the `droplet guid` from the output of the previous step. See the following example output:

```
Staging package for MY-APP in org test / space test as admin...
...
Package staged
droplet guid: f60d3464-415a-4202-9d40-26a70373a487
state: staged
created: Mon 25 Sep 16:37:45 PDT 2018
```

#### 6. Assign the droplet to your app:

```
cf v3-set-droplet MY-APP -d DROPLET-GUID
```

#### 7. Start your app:

```
cf v3-start MY-APP
```

## Roll Back to a Previous Droplet

This example workflow describes how to roll back to a previous droplet used by your app. You may want to use this, for example, if you update your app and it has a bug that causes it to crash.

#### 1. List the droplets for your app:

```
cf v3-droplets MY-APP
```

#### 2. From the output, locate and copy the second-to-last GUID. In the following example, this is `66524145-5502-40e6-b782-47fe68e13c49`.

```
Listing droplets of app MY-APP in org test / space test as admin...

guid state created
66524145-5502-40e6-b782-47fe68e13c49 staged Mon 25 Sep 16:37:34 PDT 2018
0677ad93-9f77-4aaa-9a6b-44da022ded58 staged Mon 25 Sep 16:44:55 PDT 2018
```

#### 3. Stop your app:

```
cf v3-stop MY-APP
```

#### 4. Set the app to use the previous droplet:

```
cf v3-set-droplet MY-APP -d PREVIOUS-DROPLET-GUID
```

#### 5. Start your app:

```
cf v3-start MY-APP
```

## Rolling App Deployments (Beta)

**⚠ warning:** This CLI feature is currently experimental, use at your own risk. Further development of this feature will occur on CLI V7 beta.

This topic describes how developers use beta Cloud Foundry Command Line Interface (cf CLI) commands to push their apps using a rolling deployment.

For information about the traditional method for addressing app downtime while pushing app updates, see [Using Blue-Green Deployment to Reduce Downtime and Risk](#).

### Prerequisites

The procedures in this topic require the following:

- You must have cf CLI v6.40 or later.
- The rolling deployment feature must be enabled for your installation.
  - Ensure that the **Disable Zero Downtime App Deployments** checkbox is not selected in the **Advanced Features** pane of the Pivotal Application Service (PAS) tile.

### Commands

This section describes the beta commands for working with rolling app deployments.

#### Deploy an App

To deploy an app without incurring downtime, run the following command:

```
cf v3-zdt-push APP-NAME
```

**⚠ warning:** Ensure that you understand the [Limitations](#) of this feature before running the command.

For more information about this command, see [How it Works](#).

#### Cancel a Deployment

To stop the deployment of an app that you pushed, run the following command:

**Note:** The `v3-cancel-zdt-push` command is designed to revert the app to its original state as quickly as possible, and does not guarantee zero downtime.

```
cf v3-cancel-zdt-push APP-NAME
```

This reverts the app to its state from before the deployment started by doing the following:

- Scaling up the original web process
- Removing any deployment artifacts
- Resetting the `current_droplet` on the app

#### Restart an App

Restart your app without downtime using the following command. This is useful when applying configuration updates that require a restart, such as environment variables or service bindings.

cf v3-zdt-restart APP-NAME

## How It Works

This section describes the beta `cf v3-zdt-push` command and its limitations.

### cf v3-zdt-push Process

This section describes the process of pushing an app with `cf v3-zdt-push`.

1. The `cf v3-zdt-push` command does the following:
  - a. Stages the updated app package
  - b. Creates a droplet with the updated app package
  - c. Creates a deployment with the new droplet and any new configuration.
    - This starts a new process with one instance that shares the route with the old process.
    - At this point, if you run `cf app` on your app, you see a `web` process and a `web-deployment-DEPLOYMENT-UUID` process.
    - For more information about the deployment object, see the [Deployments](#) section of the CAPI v3 docs.
2. Once the command creates the deployment, the `cc_deployment_updater` BOSH job runs in the background, updating deployments as follows:
  - a. Adds another instance of the new web process and removes an instance from the old web process.
 

 **Note:** This only happens if all instances of the new web process are currently running.
  - b. Repeats the above step until the new web process reaches the desired number of instances.
  - c. Removes the old web process. The new web process now fully replaces the old web process.
  - d. Restarts all non-web processes of the app.
  - e. Sets the deployment to `DEPLOYED`.

## Limitations

The following table describes the limitations to consider when using the beta `v3-zdt-push` command.

Limitation	Description
Multiple app versions	During a deployment, Cloud Foundry serves both the old and new version of your app at the same route. This could lead to user issues if you push backwards-incompatible API changes.
Database migrations	Deployments do not handle database migrations. If you are migrating an app database and the old app is not compatible with the migration, you may still have downtime.
Non-web processes	<p>The <code>v3-zdt</code> commands only run web processes through the rolling update sequence described above. The commands restart worker and other non-web processes in bulk after updating all web processes.</p> <p>The CAPI v3 API introduces the concept of processes as runnable units of an app. Each app has a web process by default. You can specify additional processes with a procfile, and in some cases buildpacks create additional processes. For more information about processes, see the <a href="#">Processes</a> section of the CAPI v3 documentation.</p>
Quotas	Pushing updates to your app with a <code>v3-zdt</code> command creates an extra instance of your app. If there is not enough quota available, the deployment does not work. Administrators may need to increase quotas to account for rolling deployments.
Simultaneous apps when interrupting a push	If you run <code>v3-zdt-push</code> for an app before your previous <code>v3-zdt-push</code> command for the same app has completed, your first push gets interrupted. Until the last deployment completes, there may be many versions of the app running at once. Eventually, the app runs the code from your most recent push.
SSH to app instances	Pushing updates to your app with a <code>v3-zdt</code> command causes the new web process and app GUID to mismatch. <code>cf ssh</code> does not handle this scenario. You must use the <code>cf v3-ssh</code> command instead.



## Pushing Apps with Sidecar Processes (Beta)

This topic describes sidecar processes and how to include them when you push your app.

### Overview

You can run additional processes in the same container as your app. These are called sidecar processes, or sidecars. An example of a sidecar is an Application Performance Monitoring (APM) tool.

When you provide a sidecar for your app, Cloud Foundry packages the required code and configuration needed to run the sidecar and app in the same droplet. It deploys this droplet in a single container on Diego. Both processes within the container are health checked independently.

For additional information about sidecars, see the [Sidecars section of the CAPI v3 documentation](#).

### Use Cases

You can use sidecars for processes that depend on each other or must run in the same container.

For example, you can use sidecars for processes that must do the following:

- Communicate over a Unix socket or through localhost
- Share the same filesystem
- Be scaled and placed together
- Have fast interprocess communication

### Limitations

Sidecars have the following limitations:

- The start and stop order of app processes and their sidecars is undefined.
- App processes and sidecars are codependent. If either crashes or exits, the other will as well.
- Sidecars are currently not independently scalable. Sidecars share resources with the main app process and other sidecars within the container.
- Sidecars only support PID-based health checks. HTTP health-checks for sidecars are not currently supported.

## Push an App with a Sidecar

This section provides instructions for pushing an app with a sidecar. For an example that you can try yourself, see the [Sidecar Tutorial](#) section below.

### Prerequisites

Before you can push an app with a sidecar, you must have the following:

- An app that is currently running or ready to be pushed.
- A file that Cloud Foundry can execute inside the app container as a sidecar process. For example, an executable binary, a Java .jar file, or Ruby scripts.

### Procedure

To push an app with a sidecar, do the following:

 **Note:** This procedure uses `v3` commands because it requires the server-side manifest support in CAPI v3.

1. Create an app. Or, use an existing app. Run the following command to create an app:

```
cf v3-create-app MY-APP
```

Where `MY-APP` is a name that you define for your app.

2. Create a manifest file in the root directory of your app, such as `manifest.yml`. Or, use an existing manifest file for your app. For more information, see [Deploying with App Manifests](#).

3. Add the following content to your app manifest file under the `applications` key:

```
sidecars:
- name: MY-SIDECAR
 process_types: ['MY-PROCESS-TYPES']
 command: MY-START-COMMAND
```

Where:

- o `MY-SIDECAR` is a name that you define for your sidecar.
- o `MY-PROCESS-TYPES` is a list of app processes for the sidecar to attach to, such as `web` or `worker`. You can attach multiple sidecars to each process type your app uses.
- o `MY-START-COMMAND` is the command used to start the sidecar. For example, `./my-binary` or `java -jar my-java-file.jar`.

See the following example manifest file that includes multiple sidecars:

```

applications:
- name: my-app
 sidecars:
 - name: authenticator
 process_types: ['web', 'worker']
 command: bundle exec run-authenticator
 - name: performance monitor
 process_types: ['web']
 command: bundle exec run-performance-monitor
```

4. Apply the manifest file to your app:

```
cf v3-apply-manifest -f PATH-TO-MANIFEST
```

5. Push your app:

```
cf v3-push MY-APP
```

## Sidecar Tutorial

You can explore sidecars using the app in the [capi-sidecar-samples](#) repository. The sections below describe the app, how to build and push the app, and some ways to observe the app and its processes after pushing.

### About the Sample App

The [capi-sidecar-samples](#) repository contains the following:

- **A simple Ruby app**

This app is named `sidecar-dependent-app`. It includes a `/config` endpoint that calls to the sidecar and prints the response. See the following code snippet:

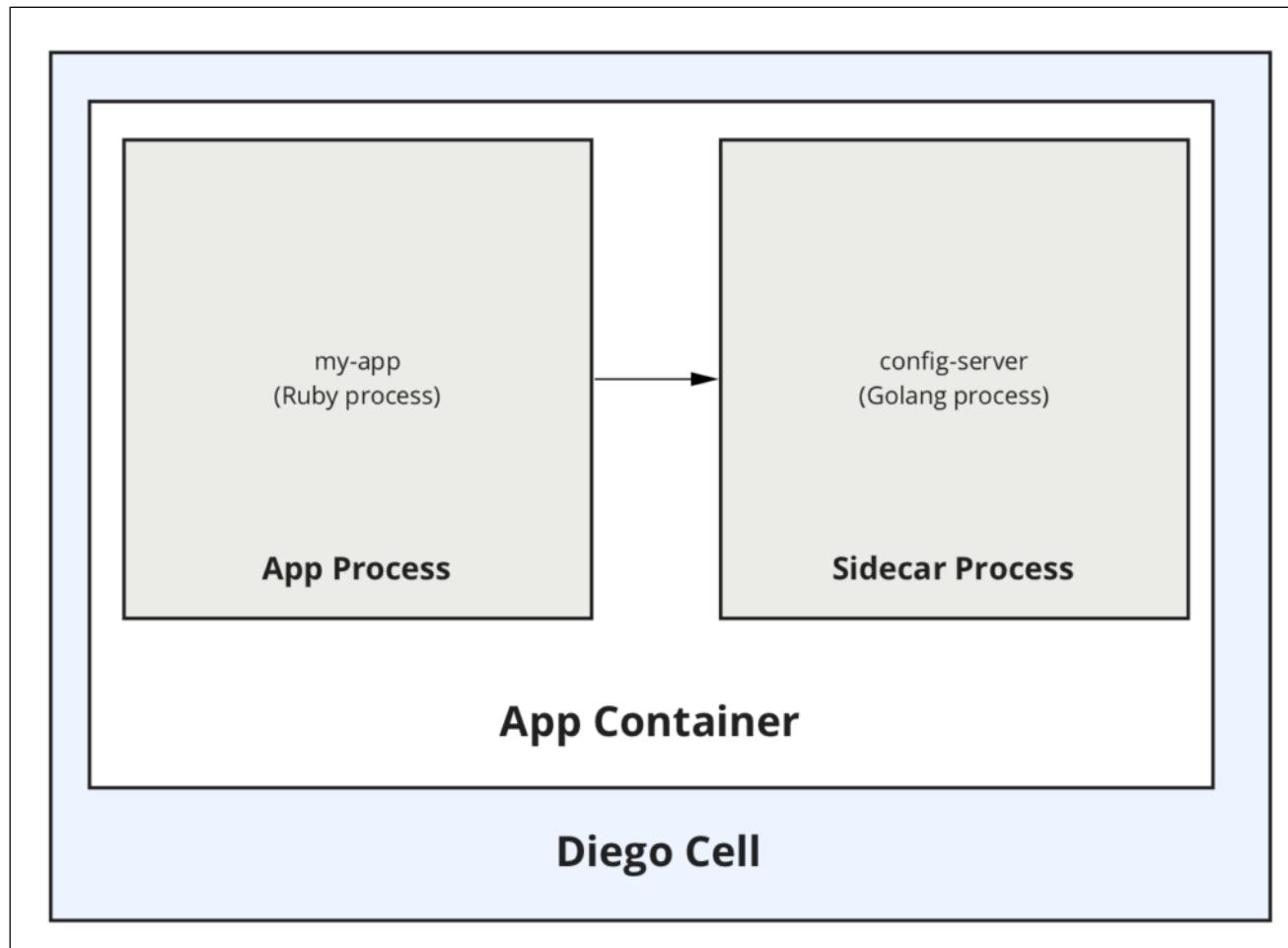
```
get '/config' do
 puts "Sending a request to the config-server sidecar at localhost:#{ENV['CONFIG_SERVER_PORT']}/config/"
 response = Typhoeus.get("localhost:#{ENV['CONFIG_SERVER_PORT']}/config/")
 puts "Received #{response.body} from the config-server sidecar"
 response.body
end
```

- **A Golang sidecar**

The `config-server-sidecar` produces a `config-server` binary. It provides apps with their required configuration over its `/config` endpoint. It also accepts

connections only over localhost on the `CONFIG_SERVER_PORT` port. This means the sidecar must be co-located in the same container as the app, so that it shares the same network namespace as the main app.

The diagram below illustrates the app architecture:



## Push the App and Sidecar

To push the app and sidecar, do the following:

1. In a terminal window, clone the git repository to your workspace:

```
git clone https://github.com/cloudfoundry-samples/capi-sidecar-samples.git
```

2. Navigate to the `config-server-sidecar` directory.

3. Build the binary for the sidecar by running the following command:

```
GOOS=linux GOARCH=amd64 go build -o config-server .
```

**Note:** If you do not have Go installed, download the `config-server_linux_x86-64` binary from the [Releases](#) section of the `capi-sidecar-samples` repository in GitHub.

4. Create the app:

```
cf v3-create-app sidecar-dependent-app
```

5. Navigate to the `sidecar-dependent-app` directory.

6. Open and review the `manifest.yml` file. Under `sidecars`, the sidecar is specified with a name, process type, and start command. Under `env`, there is

an environment variable that defines the port on which the app and sidecar communicate.

7. Apply the manifest to the app:

```
cf v3-apply-manifest
```

8. Push the app:

```
cf v3-push sidecar-dependent-app
```

After you push the app, you can further explore it in the following sections: \* [View the Processes Running in the Container](#) \* [View the Web URL and App Logs](#)

## View the Processes Running in the Container

To view the app and sidecar process running in the container, do the following:

1. SSH into the app container:

```
cf ssh sidecar-dependent-app
```

2. Run `ps aux` to see both processes running: The `rackup` process for the main app and `config-server` process for the sidecar. For example:

```
vcap@f00949bd-6601-4731-6f7e-e859:~$ ps aux
USER PID %CPU %MEM VSZ RSS TTY STAT START TIME COMMAND
root 1 0.0 0.0 1120 0 ? S 22:17 0:00 /tmp/garden-init
vcap 7 0.0 0.0 106716 4508 ? S 22:17 0:00 ./config-server
vcap 13 0.0 0.1 519688 35412 ? S 22:17 0:00 /home/vcap/deps/0/vendor_bundle/ruby/2.4.0/bin/rackup config.ru -p 8080
vcap 24 0.0 0.0 116344 10792 ? S 22:17 0:00 /tmp/lifecycle/diego-sshd --allowedKeyExchanges= --address=0.0.0.0:2222 --allowUnauthenticatedClients=false --inhe
root 82 0.0 0.0 108012 4548 ? S 22:17 0:00 /etc/cf-assets/healthcheck/healthcheck -port=8080 -timeout=1000ms -liveness-interval=30s
vcap 215 0.3 0.0 70376 3756 pts/0 S 23:12 0:00 /bin/bash
vcap 227 0.0 0.0 86268 3116 pts/0 R 23:12 0:00 ps aux
```

3. Run `lsof -i | grep $CONFIG_SERVER_PORT` to see that the sidecar is listening on the port specified by `CONFIG_SERVER_PORT` and that the main `ruby` process is connected to it. For example:

```
vcap@f00949bd-6601-4731-6f7e-e859:~$ lsof -i | grep $CONFIG_SERVER_PORT
config-se 7 vcap 3u IPv4 17265901 0t0 TCP *:8082 (LISTEN)
config-se 7 vcap 5u IPv4 17265992 0t0 TCP localhost:8082->localhost:42266 (ESTABLISHED)
uby 13 vcap 11u IPv4 17274965 0t0 TCP localhost:42266->localhost:8082 (ESTABLISHED)
```

## View the Web URL and App Logs

To view the Web URL and logs for the app, do the following:

1. In a browser, navigate to the `config` endpoint of the `sidecar-dependent-app`. For example: <https://sidecar-dependent-app.example.com/config>.
2. See that the browser displays `Scope` and `Password` information. This is the configuration that the app fetches from the `config-server` sidecar.
3. In a terminal window, run the following command to begin streaming logs for the app:

```
cf logs sidecar-dependent-app
```

4. In your browser, refresh the `/config` endpoint page and observe that the log stream in your terminal displays logs for both the sidecar and the main app process.
5. In a separate terminal window from your log stream, SSH into the app container:

```
cf ssh sidecar-dependent-app
```

6. Terminate the sidecar process:

```
kill -9 $(pgrep config-server)
```

7. View the output in the terminal window where you are streaming the app logs. The app logs indicate that the sidecar process crashed and that Diego restarted the app container. For example:

```
2019-04-17T16:48:55.41-0700 [API/0] OUT App instance exited with guid
21df1eb8-f25d-43b2-990b-c1a417310553 payload:
{"instance":>"a8db0eed-7371-4805-5ad3-4596", "index":>0,
"cell_id":>"86808ce7-afc2-47da-9e79-522a62a48cf9", "reason":>"CRASHED",
"exit_description":>"APP/PROC/WEB/SIDECAR/CONFIG-SERVER: Exited with status 137",
"crash_count":>1, "crash_timestamp":>155544935367052708,
"version":>"50892dc8-274d-4cf6-b944-3eda1e000283"}
```

## Using Blue-Green Deployment to Reduce Downtime and Risk

Blue-green deployment is a technique that reduces downtime and risk by running two identical production environments called Blue and Green.

At any time, only one of the environments is live, with the live environment serving all production traffic. For this example, Blue is currently live and Green is idle.

As you prepare a new version of your software, deployment and the final stage of testing takes place in the environment that is *not* live: in this example, Green. Once you have deployed and fully tested the software in Green, you switch the router so all incoming requests now go to Green instead of Blue. Green is now live, and Blue is idle.

This technique can eliminate downtime due to app deployment. In addition, blue-green deployment reduces risk: if something unexpected happens with your new version on Green, you can immediately roll back to the last version by switching back to Blue.

**Note:** If your app uses a relational database, blue-green deployment can lead to discrepancies between your Green and Blue databases during an update. To maximize data integrity, configure a single database for backward and forward compatibility.

**Note:** You can adjust the route mapping pattern to display a static maintenance page during a maintenance window for time-consuming tasks, such as migrating a database. In this scenario, the router switches all incoming requests from Blue to Maintenance to Green.

## Blue-Green Deployment with Cloud Foundry Example

For this example, we'll start with a simple app: "demo-time." This app is a web page that displays the words "Blue time" and the date/time on the server.

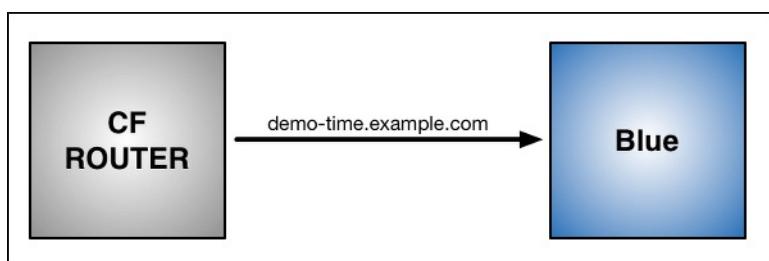
### Step 1: Push an App

Use the Cloud Foundry Command Line Interface (cf CLI) to push the app. Name the app "Blue" with the subdomain "demo-time."

```
$ cf push Blue -n demo-time
```

As shown in the graphic below:

- Blue is now running on Cloud Foundry.
- The CF Router sends all traffic for `demo-time.example.com` traffic to Blue.



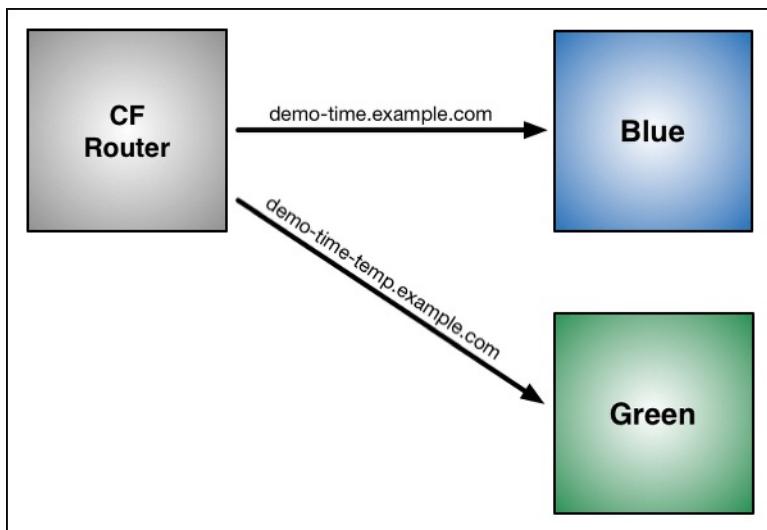
### Step 2: Update App and Push

Now make a change to the app. First, replace the word "Blue" on the web page with "Green," then rebuild the source file for the app. Run `cf push` again, but use the name "Green" for the app and provide a different subdomain to create a temporary route:

```
$ cf push Green -n demo-time-temp
```

After this push:

- Two instances of our app are now running on Cloud Foundry: the original Blue and the updated Green.
- The CF Router continues sending all traffic for `demo-time.example.com` to Blue. The router now also sends any traffic for `demo-time-temp.example.com` to Green.



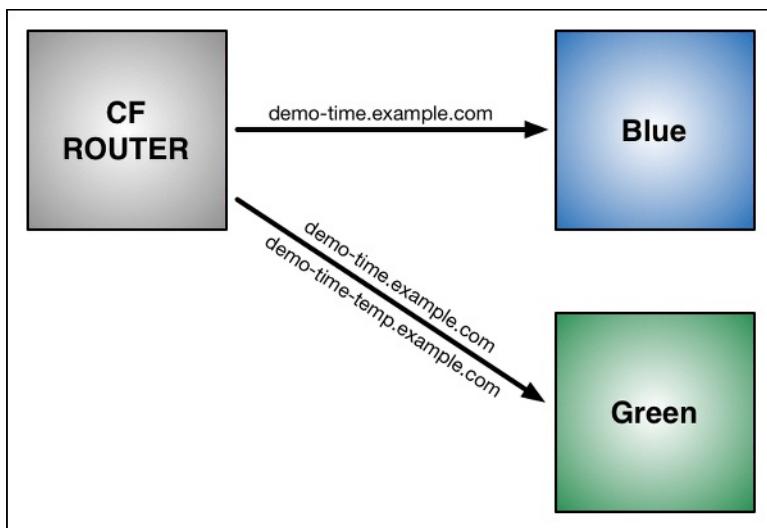
### Step 3: Map Original Route to Green

Now that both apps are up and running, switch the router so all incoming requests go to the Green app *and* the Blue app. Do this by mapping the original URL route (`demo-time.example.com`) to the Green app using the [cf map-route](#) command.

```
$ cf map-route Green example.com -n demo-time
Binding demo-time.example.com to Green... OK
```

After the `cf map-route` command :

- The CF Router continues sending traffic for `demo-time-temp.example.com` to Green.
- Within a few seconds, the CF Router begins load balancing traffic for `demo-time.example.com` between Blue and Green.



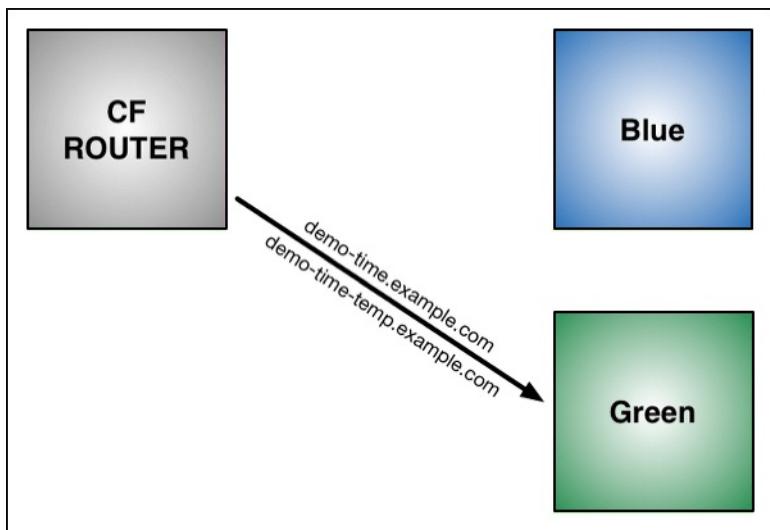
### Step 4: Unmap Route to Blue

Once you verify Green is running as expected, stop routing requests to Blue using the [cf unmap-route](#) command:

```
$ cf unmap-route Blue example.com -n demo-time
Unbinding demo-time.example.com from blue... OK
```

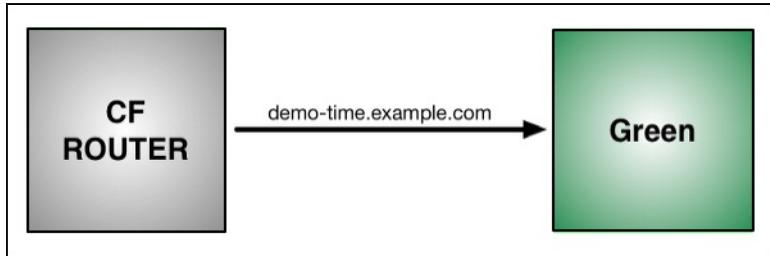
After `cf unmap-route` command:

- The CF Router stops sending traffic to Blue. Now all traffic for `demo-time.example.com` is sent to Green:



## Step 5: Remove Temporary Route to Green

You can now use `cf unmap-route` to remove the route `demo-time-temp.example.com` from Green. The route can be deleted using `cf delete-route` or reserved for later use. You can also decommission Blue, or keep it in case you need to roll back your changes.



## Implementation

Cloud Foundry community members have written a plugin to automate the blue-green deployment process:

- [BlueGreenDeploy](#): cf-blue-green-deploy is a plugin, written in Go, for the Cloud Foundry Command Line Interface (cf CLI) that automates a few steps involved in zero-downtime deploys.

## Troubleshooting app Deployment and Health

Refer to this topic for help diagnosing and resolving common issues when you deploy and run apps on Cloud Foundry.

### Common Issues

The following sections describe common issues you might encounter when attempting to deploy and run your app, and possible resolutions.

#### cf push Times Out

If your deployment times out during the upload or staging phase, you may receive one of the following error messages:

- `504 Gateway Timeout`
- `Error uploading application`
- `Timed out waiting for async job JOB-NAME to finish`

If this happens, do the following:

- **Check your network speed.** Depending on the size of your app, your `cf push` could be timing out because the upload is taking too long. We recommend an Internet connection speed of at least 768 KB/s (6 Mb/s) for uploads.
- **Make sure you are pushing only needed files.** By default, `cf push` will push all the contents of the current working directory. Make sure you are pushing only the directory for your app. If your app is too large, or if it has many small files, Cloud Foundry may time out during the upload. To reduce the number of files you are pushing, ensure that you push only the directory for your app, and remove unneeded files or use the `.cfignore` file to [specify excluded files](#).
- **Set the CF\_STAGING\_TIMEOUT and CF\_STARTUP\_TIMEOUT environment variables.** By default your app has 15 minutes to stage and 5 minutes to start. You can increase these times by setting `CF_STAGING_TIMEOUT` and `CF_STARTUP_TIMEOUT`. Type `cf push -h` at the command line for more information.
- **If your app contains a large number of files, try pushing the app repeatedly.** Each push uploads a few more files. Eventually, all files have uploaded and the push succeeds. This is less likely to work if your app has many small files.

#### App Too Large

If your app is too large, you may receive one of the following error messages on `cf push`:

- `413 Request Entity Too Large`
- `You have exceeded your organization's memory limit`

If this happens, do the following:

- **Make sure your org has enough memory for all instances of your app.** You will not be able to use more memory than is allocated for your organization. To view the memory quota for your org, use `cf org ORG_NAME`. Your total memory usage is the sum of the memory used by all apps in all spaces within the org. Each app's memory usage is the memory allocated to it multiplied by the number of instances. To view the memory usage of all the apps in a space, use `cf apps`.
- **Make sure your app is less than 1 GB.** By default, Cloud Foundry deploys all the contents of the current working directory. To reduce the number of files you are pushing, ensure that you push only the directory for your app, and remove unneeded files or use the `.cfignore` file to [specify excluded files](#). The following limits apply:
  - The app files to push cannot exceed 1 GB.
  - The droplet that results from compiling those files cannot exceed 1.5 GB. Droplets are typically a third larger than the pushed files.
  - The combined size of the app files, compiled droplet, and buildpack cache cannot total more than 4 GB of space during staging.

#### Unable to Detect a Supported App Type

If Cloud Foundry cannot [identify an appropriate buildpack](#) for your app, you will see an error message that states `Unable to detect a supported app type`.

You can view what buildpacks are available with the `cf buildpacks` command.

If you see a buildpack that you believe should support your app, refer to the [buildpack documentation](#) for details about how that buildpack detects apps it supports.

If you do not see a buildpack for your app, you may still be able to push your app with a [custom buildpack](#) using `cf push -b` with a path to your buildpack.

## App Deploy Fails

Even when the deploy fails, the app might exist on PAS. Run `cf apps` to review the apps in the currently targeted org and space. You might be able to correct the issue using the CLI or [Apps Manager](#), or you might have to delete the app and redeploy.

If you push multiple apps using a manifest and one fails to deploy, CF does not attempt to push apps specified after the app that failed.

Common reasons deploying an app fails include the following:

- The app does not successfully bind to a service, such as a PostgreSQL or MongoDB. The service may not exist, may be down, or may reject the binding.
- The app does not have a unique URL. Refer to the troubleshooting tip App Requires Unique URL.

## App Requires Unique URL

PAS requires that each app that you deploy has a unique URL. Otherwise, the new app URL collides with an existing app URL and PAS cannot successfully deploy the app. You can resolve this issue by running `cf push` with either of the following flags to create a unique URL:

- `-n` to assign a different HOST name for the app.
- `--random-route` to create a URL that includes the app name and random words. Using this option might create a long URL, depending on the number of words that the app name includes.

## App Fails to Start

After `cf push` stages the app and uploads the droplet, the app may fail to start, commonly with a pattern of starting and crashing similar to the following example:

```
----> Uploading droplet (23M)
...
0 of 1 instances running, 1 starting
0 of 1 instances running, 1 down
...
0 of 1 instances running, 1 failing
FAILED
Start unsuccessful
```

If this happens, try the following:

- **Find the reason app is failing and modify your code.** Run `cf events APP-NAME` and `cf logs APP-NAME --recent` and look for messages similar to this:

```
2018-07-20T15:53:26.00-0700 app.crash app-name index: 2, reason: CRASHED, cell_id: d874ad05-d0ca-4c63-9f5a-0b1ddd90dd5d, instance: f406c53e-b1ca-4a0f-6140-dddd, e
```

These messages may identify a memory or port issue. If they do, take that as a starting point when you re-examine and fix your app code.

- **Make sure your app code uses the `PORT` environment variable.** Your app may be failing because it is listening on the wrong port. Instead of hard coding the port on which your app listens, use the `PORT` environment variable.  
For example, this Ruby snippet assigns the port value to the `listen_here` variable: `listen_here = ENV['PORT']`  
For more examples specific to your app framework, see the appropriate [buildpacks documentation](#) for your app's language.
- **Make sure your app adheres to the principles of the [Twelve-Factor App](#) and [Prepare to Deploy an App](#).** These texts explain how to prevent situations where your app builds locally but fails to build in the cloud.
- **Verify the timeout configuration of your app.** App health checks use a timeout setting when an app starts up for the first time. See [Using Application Health Checks](#). If an app fails to start up due to health check timeout, you might see messages in the logs similar to the following:

```
2017-01-30T14:07:20.39-0800 [CELL/0] OUT Creating container
2017-01-30T14:07:20.65-0800 [CELL/0] OUT Successfully created container
2017-01-30T14:07:22.30-0800 [CELL/0] OUT Starting health monitoring of container
2017-01-30T14:08:23.52-0800 [CELL/0] ERR Timed out after 1m0s:
health check never passed.
2017-01-30T14:08:23.57-0800 [CELL/0] OUT Destroying container
2017-01-30T14:08:23.59-0800 [API/0] OUT Process has crashed with type: "web"
2017-01-30T14:08:23.59-0800 [CELL/0] OUT Creating container
2017-01-30T14:08:23.60-0800 [API/0] OUT App instance exited with guid 91086440-bac0-44f0-808f-a034a1ec2ed0
payload: {"instance":>"", "index":>0, "reason":>"CRASHED",
"exit_description":>"2 error(s) occurred:\n\n* 1 error(s)
occurred:\n\n* Exited with status 6\n\n* 2 error(s)
occurred:\n\n* cancelled\n* cancelled", "crash_count":>1,
"crash_timestamp":>1485814103565763172,
"version":>"3e6e4232-7e19-4168-9583-1176833d2c71"};
2017-01-30T14:08:23.83-0800 [CELL/0] OUT Successfully destroyed container
2017-01-30T14:08:23.84-0800 [CELL/0] OUT Successfully created container
2017-01-30T14:08:25.41-0800 [CELL/0] OUT Starting health monitoring of container
```

## App consumes too much memory, then crashes

An app that `cf push` has uploaded and started can crash later if it uses too much memory.

**Make sure your app is not consuming more memory than it should.** When you ran `cf push` and `cf scale`, that configured a limit on the amount of memory your app should use. Check your app's actual memory usage. If it exceeds the limit, modify the app to use less memory.

## Routing Conflict

PAS allows multiple apps, or versions of the same app, to be mapped to the same route. This feature enables Blue-Green deployment. For more information see [Using Blue-Green Deployment to Reduce Downtime and Risk](#).

Routing multiple apps to the same route may cause undesirable behavior in some situations by routing incoming requests randomly to one of the apps on the shared route.

If you suspect a routing conflict, run `cf routes` to check the routes in your installation.

If two apps share a route outside of a Blue-Green deploy strategy, choose one app to re-assign to a different route and follow the procedure below:

1. Run `cf unmap-route YOUR-APP-NAME OLD-ROUTE` to remove the existing route from that app.
2. Run `cf map-route YOUR-APP-NAME NEW-ROUTE` to map the app to a new, unique route.

## Gather Diagnostic Information

Use the techniques in this section to gather diagnostic information and troubleshoot app deployment issues.

### Examine Environment Variables

`cf push` deploys your app to a container on the server. The environment variables in the container govern your app.

You can set environment variables in a manifest created before you deploy. See [Deploying with app Manifests](#).

You can also set an environment variable with a `cf set-env` command followed by a `cf push` command. You must run `cf push` for the variable to take effect in the container environment.

Use the `cf env` command to view the environment variables that you have set using the `cf set-env` command and the variables in the container environment:

```
$ cf env my-app
Getting env variables for app my-app in org My-Org / space development as admin...
OK

System-Provided:
{
 "VCAP_SERVICES": {
 "p-mysql-n/a": [
 {
 "credentials": {
 "uri": "postgres://lraa:c6B-X@p-mysql[provider.example.com]:5432/lraa"
 },
 "label": "p-mysql-n/a",
 "name": "p-mysql",
 "syslog_drain_url": "",
 "tags": ["postgres", "postgresql", "relational"]
 }
]
 }
}

User-Provided:
my-env-var: 100
my-drain: http://drain.example.com
```

## View Logs

To view app logs streamed in real-time, use the `cf logs APP-NAME` command.

To aggregate your app logs to view log history, bind your app to a syslog drain service. For more information, see [Streaming app Logs to Log Management Services](#).

 **Note:** The Diego architecture does not support the `cf files` command.

## Trace Cloud Controller REST API Calls

If a command fails or produces unexpected results, re-run it with HTTP tracing enabled to view requests and responses between the Cloud Foundry Command Line Interface (cf CLI) and the Cloud Controller REST API.

For example:

- Re-run `cf push` with `-v`:  
`cf push APP-NAME -v`
- Re-run `cf push` while appending API request diagnostics to a log file:  
`CF_TRACE=PATH-TO-TRACE.LOG cf push APP-NAME`

These examples enable HTTP tracing for a single command only. To enable it for an entire shell session, set the variable first:

```
export
CF_TRACE=true

export CF_TRACE=PATH-TO-TRACE.LOG
```

 **Note:** `CF_TRACE` is a local environment variable that modifies the behavior of the cf CLI. Do not confuse `CF_TRACE` with the [variables in the container environment](#) where your apps run.

## Analyze Zipkin Trace IDs

When the Zipkin feature is enabled in Cloud Foundry, the Gorouter adds or forwards Zipkin trace IDs and span IDs to HTTP headers. For more information about what the Gorouter provides in the HTTP header, see the [HTTP Headers](#) section of the [HTTP Routing](#) topic.

After adding Zipkin HTTP headers to app logs, developers can use `cf logs` to correlate the trace and span ids logged by the Gorouter with the trace

ids logged by their app. To correlate trace ids for a request through multiple apps, each app must forward appropriate values for the headers with requests to other apps.

## Use Troubleshooting Commands

You can investigate app deployment and health using the cf CLI.

Some cf CLI commands may return connection credentials. Remove credentials and other sensitive information from command output before you post the output a public forum.

- `cf apps` : Returns a list of the apps deployed to the current space with deployment options, including the name, current state, number of instances, memory and disk allocations, and URLs of each app.
- `cf app APP-NAME` : Returns the health and status of each instance of a specific app in the current space, including instance ID number, current state, how long it has been running, and how much CPU, memory, and disk it is using.

 **Note:** CPU values returned by `cf app` show the total usage of each app instance on all CPU cores on a host VM, where each core contributes 100%. For example, the CPU of a single-threaded app instance on a Diego cell with one core cannot exceed 100%, and four instances sharing the cell cannot exceed an average CPU of 25%. A multi-threaded app instance running alone on a cell with eight cores can draw up to 800% CPU.

- `cf env APP-NAME` : Returns environment variables set using `cf set-env` and variables existing in the container environment.
- `cf events APP-NAME` : Returns information about app crashes, including error codes. Shows that an app instance exited. For more detail, look in the app logs. See <https://github.com/cloudfoundry/errors> for a list of Cloud Foundry errors.
- `cf logs APP-NAME --recent` : Dumps recent logs. See [Viewing Logs in the Command Line Interface](#).
- `cf logs APP-NAME` : Returns a real-time stream of the app STDOUT and STDERR. Use **Ctrl-C (^C)** to exit the real-time stream.
- `cf files APP-NAME` : Lists the files in an app directory. Given a path to a file, outputs the contents of that file. Given a path to a subdirectory, lists the files within. Use this to [explore](#) individual logs.

 **Note:** Your app should direct its logs to STDOUT and STDERR. The `cf logs` command also returns messages from any `log4j` facility that you configure to send logs to STDOUT.

## Access Apps with SSH

If you need to troubleshoot an instance of an app, you can gain SSH access to the app with the SSH proxy and daemon. See the [app SSH Overview](#) topic for more information.

## Send Requests to App Instance

To obtain debug data without SSH, you can make HTTP requests to a specific instance of an app by using the `X-CF-APP-INSTANCE` HTTP header. See the [App Instance Routing](#) section of the *HTTP Routing* topic for more information.

## SSH for Apps and Services

The following list provides information about configuring and using SSH for apps and services:

- [App SSH Overview](#)
- [Accessing Apps with SSH](#)
- [Accessing Services with SSH](#)

## App SSH Overview

This topic introduces SSH configuration for apps in your Pivotal Application Service (PAS) deployment.

If you need to troubleshoot an instance of an app, you can gain SSH access to the app using the SSH proxy and daemon.

For example, one of your app instances may be unresponsive, or the log output from the app may be inconsistent or incomplete. You can SSH into the individual VM that runs the problem instance to troubleshoot.

## SSH Access Control Hierarchy

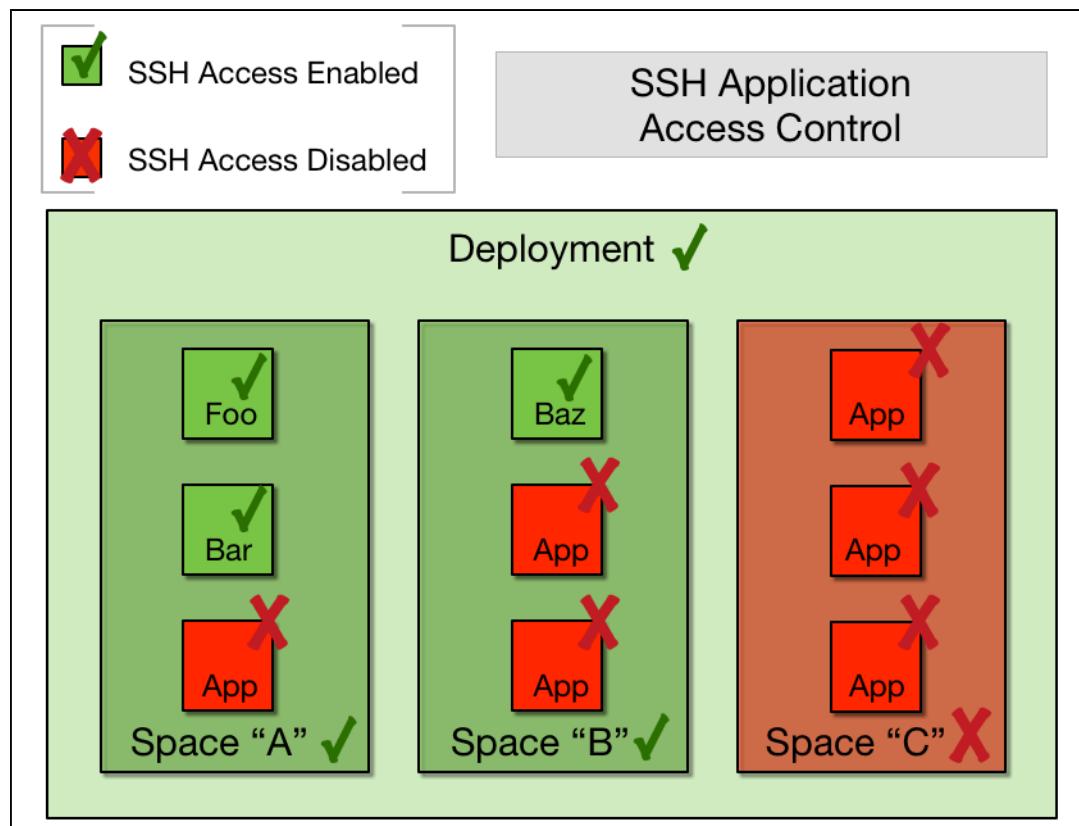
Operators, space managers, and space developers can configure SSH access for PAS, spaces, and apps as described in this table:

User Role	Scope of SSH Permissions Control	How They Define SSH Permissions
Operator	Entire deployment	Configure the deployment to allow or prohibit SSH access (one-time). For more information, see <a href="#">Configuring SSH Access for PCF</a> .
Space Manager	Space	cf CLI <a href="#">allow-space-ssh</a> and <a href="#">disallow-space-ssh</a> commands
Space Developer	App	cf CLI <a href="#">enable-ssh</a> and <a href="#">disable-ssh</a> commands

An app is SSH-accessible only if operators, space managers, and space developers all grant SSH access at their respective levels. For example, the image below shows a deployment where:

- An operator allowed SSH access at the deployment level.
- A space manager allowed SSH access for apps running in spaces “A” and “B” but not “C.”
- A space developer enabled SSH access for apps that include “Foo,” “Bar,” and “Baz.”

As a result, apps “Foo,” “Bar,” and “Baz” accept SSH requests.



## SSH Access for Apps and Spaces

Space managers and space developers can configure SSH access from the command line. The Cloud Foundry Command Line Interface (cf CLI) also includes commands to return the value of the SSH access setting. See the [Accessing Apps with Diego SSH](#) topic to use and configure SSH at both the app level and the space level.

## Configuring SSH Access for Pivotal Application Service (PAS)

Pivotal Cloud Foundry deployments control SSH access to apps at the PAS level. Additionally, Cloud Foundry supports load balancing of SSH sessions with your load balancer. The [Configuring SSH Access](#) topic describes how to set SSH access for your deployment.

## About SSH Access

The SSH system components include the SSH proxy and daemon, and the system also supports authentication, and load balancing of incoming SSH traffic. The [Application SSH Components and Processes](#) topic provides a conceptual overview.

## Accessing Apps with SSH

This page assumes you are using cf CLI v6.13.0 or later.

The Cloud Foundry Command Line Interface (cf CLI) lets you securely log into remote host virtual machines (VMs) running Pivotal Application Service (PAS) app instances. This topic describes the commands that enable SSH access to apps, and enable, disable, and check permissions for such access.

The cf CLI looks up the `app_ssh_oauth_client` identifier in the Cloud Controller `/v2/info` endpoint, and uses this identifier to query the UAA server for an SSH authorization code. On the target VM side, the SSH proxy contacts the Cloud Controller through the `app_ssh_endpoint` listed in `/v2/info` to confirm permission for SSH access.

## App SSH Commands

cf CLI Command	Purpose
<code>cf enable-ssh</code> <code>cf disable-ssh</code> <code>cf allow-space-ssh</code> <code>cf disallow-space-ssh</code>	<a href="#">Enable and Disable SSH Access</a>
<code>cf ssh-enabled</code> <code>cf space-ssh-allowed</code>	<a href="#">Check SSH Access Permissions</a>
<code>cf ssh</code>	<a href="#">Securely log into an app container</a>
<code>cf ssh-code</code>	<a href="#">Enable secure log in to an app container using non-CF SSH tools like <code>ssh</code>, <code>scp</code>, and <code>sftp</code></a>

## Enabling and Disabling SSH Access

A cloud operator can deploy Pivotal Application Service (PAS) to either allow or prohibit app SSH across the entire deployment. For more information, see [Configuring SSH Access for PCF](#).

Within a deployment that permits SSH access to apps, Space Developers can enable or disable SSH access to individual apps, and Space Managers can enable or disable SSH access to all apps running within a space.

You must restart your app after enabling SSH access.

### Configuring SSH Access at the App Level

`cf enable-ssh` [enables](#) SSH access to all instances of an app:

```
$ cf enable-ssh MY-AWESOME-APP
```

`cf disable-ssh` [disables](#) SSH access to all instances of an app:

```
$ cf disable-ssh MY-AWESOME-APP
```

### Configuring SSH Access at the Space Level

`cf allow-space-ssh` [allows](#) SSH access into all apps in a space:

```
$ cf allow-space-ssh SPACE-NAME
```

`cf disallow-space-ssh` [disallows](#) SSH access into all apps in a space:

```
$ cf disallow-space-ssh SPACE-NAME
```

## Checking SSH Permissions

`cf ssh-enabled` checks whether an app is accessible with SSH:

```
$ cf ssh-enabled MY-AWESOME-APP
ssh support is disabled for 'MY-AWESOME-APP'
```

`cf space-ssh-allowed` checks whether all apps running within a space are accessible with SSH:

```
$ cf space-ssh-allowed SPACE-NAME
ssh support is enabled in space 'SPACE-NAME'
```

## Logging Into an App Container with cf SSH

If SSH access is allowed at the deployment, space, and app level, you can run the `cf ssh APP-NAME` command to start an interactive SSH session with a VM hosting an app. By default, the command accesses the container running the first instance of the app, the instance with index 0.

```
$ cf ssh MY-AWESOME-APP
```

When logged into a VM hosting an app, you can use tools like the Cloud Foundry Diego Operator Toolkit (cfdot) to run app status diagnostics. For more information, see [How to use CF Diego Operator Toolkit](#).

## Common cf SSH Flags

You can tailor `cf ssh` commands with the following flags, most of which mimic flags for the Unix or Linux `ssh` command. Run the `cf ssh --help` command for more details.

- The `-i` flag targets a specific instance of an app. To log into the VM container hosting the third instance, `index=2`, of MY-AWESOME-APP, run:

```
$ cf ssh MY-AWESOME-APP -i 2
```

- The `-L` flag enables local port forwarding, binding an output port on your machine to an input port on the app VM. Pass in a local port, and your app VM port and port number, all colon delimited. You can prepend your local network interface, or use the default `localhost`.

```
$ cf ssh MY-AWESOME-APP -L [LOCAL-NETWORK-INTERFACE:]LOCAL-PORT:REMOTE-HOST-NAME:REMOTE-HOST-PORT
```

- The `-N` flag skips returning a command prompt on the remote machine. This sets up local port forwarding if you do not need to execute commands on the host VM.
- The `--request-pseudo-tty` and `--force-pseudo-tty` flags let you run an SSH session in pseudo-tty mode rather than generate terminal line output.

## SSH Session Environment

If you want the environment of your interactive SSH session to match the environment of your buildpack-based app, with the same environment variables and working directory, run the following commands after starting the session:

```
/tmp/lifecycle/shell
```

After running the above command, the value of the `VCAP_APPLICATION` environment variable differs slightly from its value in the environment of the app process, as it will not have the `host`, `instance_id`, `instance_index`, or `port` fields set. These fields are available in other environment variables, as described in the [VCAP\\_APPLICATION](#) documentation.

## App SSH Access without cf CLI

In addition to `cf ssh`, you can use other SSH clients such as `ssh`, `scp`, or `sftp` to access your app, if you have SSH permissions.

Follow the steps below to securely connect to an app instance by logging in with a specially-formed username that passes information to the SSH proxy running on the host VM. For the password, use a one-time SSH authorization code generated by `cf ssh-code`.

- Run `cf app MY-AWESOME-APP --guid` and record the GUID of your target app.

```
$ cf app MY-AWESOME-APP --guid
abcdefab-1234-5678-abcd-1234abcd1234
```

- Query the `/v2/info` endpoint of the Cloud Controller in your deployment. Record the domain name and port of the `app_ssh_endpoint` field, and the `app_ssh_host_key_fingerprint` field. You will compare the `app_ssh_host_key_fingerprint` with the fingerprint returned by the SSH proxy on your target VM.

```
$ curl /v2/info
{
...
"app_ssh_endpoint": "ssh.MY-DOMAIN.com:2222",
"app_ssh_host_key_fingerprint": "a6:14:c0:ea:42:07:b2:f7:53:2c:0b:60:e0:00:21:6c",
...
}
```

- Run `cf ssh-code` to obtain a one-time authorization code that substitutes for an SSH password. You can run `cf ssh-code | pbcopy` to automatically copy the code to the clipboard.

```
$ cf ssh-code
E1x89n
```

- Run your `ssh` or other command to connect to the app instance. For the username, use a string of the form `cf:APP-GUID/APP-INSTANCE-INDEX@SSH-ENDPOINT`, where `APP-GUID` and `SSH-ENDPOINT` come from the previous steps. For the port number, use the `SSH-PORT` recorded above. `APP-INSTANCE-INDEX` is the index of the instance you want to access.

With the above example, you `ssh` into the container hosting the first instance of your app by running the following command:

```
$ ssh -p 2222 cf:abcdefab-1234-5678-abcd-1234abcd1234/0@ssh.MY-DOMAIN.com
```

Or you can use `scp` to transfer files by running one of the following commands:

- o `scp -P 2222 -o User=cf:abcdefab-1234-5678-abcd-1234abcd1234/0 ssh.MY-DOMAIN.com:REMOTE-FILE-TO-RETRIEVE LOCAL-FILE-DESTINATION`
- o `scp -P 2222 -o User=cf:abcdefab-1234-5678-abcd-1234abcd1234/0 LOCAL-FILE-TO-COPY ssh.MY-DOMAIN.com:REMOTE-FILE-DESTINATION`

- When the SSH proxy reports its RSA fingerprint, confirm that it matches the `app_ssh_host_key_fingerprint` recorded above. When prompted for a password, paste in the authorization code returned by `cf ssh-code`.

```
$ ssh -p 2222 cf:abcdefab-1234-5678-abcd-1234abcd1234/0@ssh.MY-DOMAIN.com
The authenticity of host '[ssh.MY-DOMAIN.com]:2222 ([203.0.113.5]:2222)' can't be established.
RSA key fingerprint is a6:14:c0:ea:42:07:b2:f7:53:2c:0b:60:e0:00:21:6c.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added '[ssh.MY-DOMAIN.com]:2222 ([203.0.113.5]:2222)' (RSA) to the list of known hosts.
cf:d0a2e11d-e6ca-4120-b32d-140@ssh.ketchup.cf-app.com's password:
vcap@ce4l5164kws:~$
```

You have now securely connected to the app instance.

## SSH Proxy Security Configuration

The SSH proxy has the following SSH security configuration by default:

Security Parameter	Values
Ciphers	<code>chacha20-poly1305@openssh.com</code> <code>aes128-gcm@openssh.com</code> <code>aes256-ctr</code> <code>aes192-ctr</code> <code>aes128-ctr</code>
	<code>hmac-sha2-256-etm@openssh.com</code>

MACs	hmac-sha2-256
Key Exchanges	curve25519-sha256@libssh.org

The `cf ssh` command is compatible with this security configuration. If you use a different SSH client to access apps over SSH, you should ensure that you configure your client to be compatible with these ciphers, MACs, and key exchanges. For more information about other SSH clients, see the [App SSH Access without cf CLI](#) section of this topic.

## Proxy to Container Authentication

A second layer of SSH security runs within each container. When the SSH proxy attempts to handshake with the SSH daemon inside the target container, it uses the following fields associated with the `diego-ssh` key in its route to the app instance. This inner layer works invisibly and requires no user action, but is described here to complete the SSH security picture.

### COUNTAINER\_PORT (required)

`container_port` indicates which port inside the container the SSH daemon is listening on. The proxy attempts to connect to host side mapping of this port after authenticating the client.

### HOST\_FINGERPRINT (optional)

When present, `host_fingerprint` declares the expected fingerprint of the SSH daemon's host public key. When the fingerprint of the actual target's host key does not match the expected fingerprint, the connection is terminated. The fingerprint should only contain the hex string generated by `ssh-keygen -l`.

### USER (optional)

`user` declares the user ID to use during authentication with the container's SSH daemon. While this is not a required part of the routing data, it is required for password authentication and may be required for public key authentication.

### PASSWORD (optional)

`password` declares the password to use during password authentication with the container's ssh daemon.

### PRIVATE\_KEY (optional)

`private_key` declares the private key to use when authenticating with the container's SSH daemon. If present, the key must be a PEM encoded RSA or DSA public key.

## Example App Process

```
{
 "process_guid": "ssh-process-guid",
 "domain": "ssh-experiments",
 "rootfs": "preloaded:cflinuxfs3",
 "instances": 1,
 "start_timeout": 30,
 "setup": {
 "download": {
 "artifact": "diego-sshd",
 "from": "http://file-server.service.cf.internal.example.com:8080/v1/static/diego-sshd/diego-sshd.tgz",
 "to": "/tmp",
 "cache_key": "diego-sshd"
 }
 },
 "action": {
 "run": {
 "path": "/tmp/diego-sshd",
 "args": [
 "-address=0.0.0.0:2222",
 "-authorizedKey=ssh-rsa ..."
],
 "env": [],
 "resource_limits": {}
 }
 },
 "ports": [2222],
 "routes": {
 "diego-ssh": {
 "container_port": 2222,
 "private_key": "PEM encoded PKCS#1 private key"
 }
 }
}
```

## Daemon Discovery

To be accessible via the SSH proxy, containers must host an SSH daemon, expose it via a mapped port, and advertise the port in a `diego-ssh` route. If a proxy cannot find the target process or a route, user authentication fails.

```
"routes": {
 "diego-ssh": { "container_port": 2222 }
}
```

The Diego system generates the appropriate process definitions for Pivotal Application Service (PAS) apps which reflect the policies that are in effect.

## Accessing Services with SSH

This page assumes you are using Cloud Foundry Command Line Interface (cf CLI) v6.15.0 or later.

This topic describes how to gain direct command line access to your deployed service instance. For example, you may need access to your database to execute raw SQL commands to edit the schema, import and export data, or debug app data issues.

To establish direct command line access to a service, you deploy a host app and use its SSH and port forwarding features to communicate with the service instance through the app container. The technique outlined below works with TCP services such as MySQL or Redis.

 **Note:** The procedure in this topic requires use of a service key, and not all services support service keys. Some services support credentials through [app binding](#) only.

### Create a Service Instance

1. In your terminal window, log in to your deployment with `cf login`.
2. List the marketplace services installed as product tiles on your Pivotal Cloud Foundry (PCF) Ops Manager. See the [Adding and Deleting Products](#) topic if you need to add the service as a tile. In this example, we create a p-mysql service instance.

```
$ cf marketplace
p-mysql 100mb MySQL databases on demand
```

3. Create your service instance. As part of the [create-service](#) command, indicate the service name, the service plan, and the name you choose for your service instance.

```
$ cf create-service p-mysql 100mb MY-DB
```

### Push Your Host App

To push an app that will act as the host for the SSH tunnel, push any app that will successfully deploy to Pivotal Application Service (PAS).

 **Note:** Your app must be prepared before you push it. See the [Pushing an App](#) topic for details on preparing apps for pushing.

1. Push your app.

```
$ cf push YOUR-HOST-APP
```

2. Enable SSH for your app.

```
$ cf enable-ssh YOUR-HOST-APP
```

 **Note:** To enable SSH access to your app, SSH access must also be enabled for both the space that contains the app and Pivotal Application Service (PAS). See the [App SSH Overview](#) topic for more details.

### Create Your Service Key

To establish SSH access to your service instance, you must create a service key that contains critical information for configuring your SSH tunnel.

1. Create a service key for your service instance using the [cf create-service-key](#) command.

```
$ cf create-service-key MY-DB EXTERNAL-ACCESS-KEY
```

2. Retrieve your new service key using the [cf service-key](#) command.

```
$ cf service-key MY-DB EXTERNAL-ACCESS-KEY
Getting key EXTERNAL-ACCESS-KEY for service instance MY-DB as user@example.com

{
 "hostname": "us-cdbr-iron-east-01.p-mysql.net",
 "jdbcUrl": "jdbc:mysql://us-cdbr-iron-east-03.p-mysql.net/ad_b2fca6t49704585d?user=b5136e448be920&password=231f435o05",
 "name": "ad_b2fca6t49704585d",
 "password": "231f435o05",
 "port": "3306",
 "uri": "mysql://b5136e448be920:231f435o05@us-cdbr-iron-east-03.p-mysql.net:3306/ad_b2fca6t49704585d?reconnect=true",
 "username": "b5136e448be920"
}
```

## Configure Your SSH Tunnel

Configure an SSH tunnel to your service instance using [cf ssh](#). Tailor the example command below with information from your service key.

```
$ cf ssh -L 63306:us-cdbr-iron-east-01.p-mysql.net:3306 YOUR-HOST-APP
```

- Use any available local port for port forwarding. For example, `63306`.
- Replace `us-cdbr-iron-east-01.p-mysql.net` with the address provided under `hostname` in the service key retrieved above.
- Replace `3306` with the port provided under `port` above.
- Replace `YOUR-HOST-APP` with the name of your host app.

After you enter the command, open another terminal window and perform the steps below in [Access Your Service Instance](#).

## Access Your Service Instance

To establish direct command-line access to your service instance, use the relevant command line tool for that service. This example uses the MySQL command line client to access the p-mysql service instance.

```
$ mysql -u b5136e448be920 -h 0 -p -D ad_b2fca6t49704585d -P 63306
```

- Replace `b5136e448be920` with the username provided under `username` in your service key.
- `-h 0` instructs `mysql` to connect to your local machine (use `-h 127.0.0.1` for Windows).
- `-p` instructs `mysql` to prompt for a password. When prompted, use the password provided under `password` in your service key.
- Replace `ad_b2fca6t49704585d` with the database name provided under `name` in your service key.
- `-P 63306` instructs `mysql` to connect on port `63306`.

## Routes and Domains

The following list provides information about configuring routes and domains:

- [Configuring Routes and Domains](#)
- [Using Weighted Routing \(Beta\)](#)
- [Configuring CF to Route Traffic to Apps on Custom Ports](#)

## Configuring Routes and Domains

This topic describes how routes and domains work in Pivotal Application Service (PAS), and how developers and administrators configure routes and domains for their apps using the Cloud Foundry Command Line Interface (cf CLI).

For more information about routing capabilities in PAS, see the [HTTP Routing](#) topic.

### Routes

The PAS Gorouter routes requests to apps by associating an app with an address, known as a route. We call this association a **mapping**. Use the cf CLI [cf map-route](#) command to associate an app and route.

The routing tier compares each request with a list of all the routes mapped to apps and attempts to find the best match. For example, the Gorouter would make the following matches for the two routes `myapp.shared-domain.example.com` and `myapp.shared-domain.example.com/products`:

Request	Matched Route
<code>http://myapp.shared-domain.example.com</code>	<code>myapp.shared-domain.example.com</code>
<code>http://myapp.shared-domain.example.com/contact</code>	<code>myapp.shared-domain.example.com</code>
<code>http://myapp.shared-domain.example.com/products</code>	<code>myapp.shared-domain.example.com/products</code>
<code>http://myapp.shared-domain.example.com/products/123</code>	<code>myapp.shared-domain.example.com/products</code>
<code>http://products.shared-domain.example.com</code>	No match; 404

The Gorouter does not use a route to match requests until the route is mapped to an app. In the above example, `products.shared-domain.example.com` may have been created as a route in Cloud Foundry, but until it is mapped to an app, requests for the route receive a 404.

The routing tier knows the location of instances for apps mapped to routes. Once the routing tier determines a route as the best match for a request, it makes a load-balancing calculation using a round-robin algorithm, and forwards the request to an instance of the mapped app.

Developers can map many apps to a single route, resulting in load-balanced requests for the route across all instances of all mapped apps. This approach enables the blue/green zero-downtime deployment strategy. Developers can also map an individual app to multiple routes, enabling access to the app from many URLs. The number of routes that can be mapped to each app is approximately 1000 (128 KB).

Routes belong to a space, and developers can only map apps to a route in the same space.

 **Note:** Routes are globally unique. Developers in one space cannot create a route with the same URL as developers in another space, regardless of which orgs control these spaces.

### HTTP vs. TCP Routes

 **Note:** By default, PAS only supports routing of HTTP requests to apps.

Routes are considered HTTP if they are created from HTTP domains, and TCP if they are created from TCP domains. See [HTTP vs. TCP Shared Domains](#).

HTTP routes include a domain, an optional hostname, and an optional context path. `shared-domain.example.com`, `myapp.shared-domain.example.com`, and `myapp.shared-domain.example.com/products` are all examples of HTTP routes. Apps should listen to the `localhost` port defined by the `$PORT` environment variable, which is `8080` on Diego. As an example, requests to `myapp.shared-domain.example.com` would be routed to the app container at `localhost:8080`.

 **Note:** Developers can use the Cloud Controller API to update the ports an app can receive requests on. For more information, see [Configuring Apps to Listen on Custom Ports \(Beta\)](#).

- Requests to HTTP routes must be sent to ports 80 or 443.
- Ports cannot be reserved for HTTP routes.

TCP routes include a domain and a route port. A route port is the port clients make requests to. This is not the same port as what an app pushed to Cloud Foundry listens on. `tcp.shared-domain.example.com:60000` is an example of a TCP route. Just as for HTTP routes, apps should listen to the `localhost` port defined by the `$PORT` environment variable, which is `8080` on Diego. As an example, requests to `tcp.shared-domain.example.com:60000` would be routed to the app container at `localhost:8080`.

- Once a port is reserved for a route, it cannot be reserved for another route.
- Hostname and path are not supported for TCP routes.

## Internal Routes

Apps can communicate without leaving the platform on the container network using internal routes. You can create an internal route using the [cf map-route](#) command with an [internal domain](#).

After an internal route is mapped to an app, the route resolves to IP addresses of the app instances. The IP addresses are visible in the app container.

```
$ cf map-route app apps.internal --hostname app
$ cf ssh app
vcap@1234:~$ host app.apps.internal
app.apps.internal has address 10.255.169.200
app.apps.internal has address 10.255.49.7
app.apps.internal has address 10.255.49.77
```

To resolve individual instances, prepend the index to the internal route.

```
vcap@1234:~$ host 1.app.apps.internal
1.app.apps.internal has address 10.255.49.7
```

By default, apps cannot communicate with each other on the container network. To allow apps to communicate with each other you must create a network policy. For more information, see [add-network-policy](#) in the *Cloud Foundry CLI Reference Guide*.

## Create a Route

When a developer creates a route using the cf CLI, PAS determines whether the route is an HTTP or a TCP route based on the domain. To create a HTTP route, a developer must choose an HTTP domain. To create a TCP route, a developer must choose a TCP domain.

Domains in PAS provide a namespace from which to create routes. To list available domains for a targeted organization, use the [cf domains](#) command. For more information about domains, see the [Domains](#) section.

The following sections describe how developers can create HTTP and TCP routes for different use cases.

### Create an HTTP Route with Hostname

In PAS, a hostname is the label that indicates a subdomain of the domain associated with the route. Given a domain `shared-domain.example.com`, a developer can create the route `myapp.shared-domain.example.com` in space `my-space` by specifying the hostname `myapp` with the [cf create-route](#) command as shown in this example:

```
$ cf create-route my-space shared-domain.example.com --hostname myapp
Creating route myapp.shared-domain.example.com for org my-org / space my-space as username@example.com...
OK
```

This command instructs PAS to only route requests to apps mapped to this route for the following URLs:

- `http://myapp.shared-domain.example.com`
- `https://myapp.shared-domain.example.com`
- Any path under either of the above URLs, such as `http://myapp.shared-domain.example.com/bar`

### Create an HTTP Route without Hostname

This approach creates a route with the same address as the domain itself and is permitted for private domains only. For more information, see the [Private Domains](#) section.

A developer can create a route in space `my-space` from the domain `private-domain.example.com` with no hostname with the [cf create-route](#) command:

```
$ cf create-route my-space private-domain.example.com
Creating route private-domain.example.com for org my-org / space my-space as username@example.com...
OK
```

If DNS has been configured correctly, this command instructs PAS to route requests to apps mapped to this route from the following URLs:

- `http://private-domain.example.com`
- `https://private-domain.example.com`
- Any path under either of the above URLs, such as `http://private-domain.example.com/foo`

If there are no other routes for the domain, requests to any subdomain, such as `http://foo.private-domain.example.com`, will fail.

A developer can also create routes for subdomains with no hostnames. The following command creates a route in space `my-space` from the subdomain `foo.private-domain.example.com`:

```
$ cf create-route my-space foo.private-domain.example.com
Creating route foo.private-domain.example.com for org my-org / space my-space as username@example.com...
OK
```

Assuming DNS has been configured for this subdomain, this command instructs PAS to route requests to apps mapped to this route from the following URLs:

- `http://foo.private-domain.example.com`
- `https://foo.private-domain.example.com`
- Any path under either of the above URLs, such as `http://foo.private-domain.example.com/foo`

## Create an HTTP Route with Wildcard Hostname

An app mapped to a wildcard route acts as a fallback app for route requests if the requested route does not exist. To create a wildcard route, use an asterisk for the hostname.

A developer can create a wildcard route in space `my-space` from the domain `foo.shared-domain.example.com` with the following command:

```
$ cf create-route my-space foo.shared-domain.example.com --hostname '*'
Creating route *.foo.shared-domain.example.com for org my-org / space my-space as username@example.com...
OK
```

If a client sends a request to `http://app.foo.shared-domain.example.com` by accident, attempting to reach `myapp.foo.shared-domain.example.com`, PAS routes the request to the app mapped to the route `*.foo.shared-domain.example.com`.

## Create an HTTP Route with a Path

Developers can use paths to route requests for the same hostname and domain to different apps.

A developer can create three routes using the same hostname and domain in the space `my-space` with the following commands:

```
$ cf create-route my-space shared-domain.example.com --hostname store --path products
Creating route store.shared-domain.example.com/products for org my-org / space my-space as username@example.com...
OK

$ cf create-route my-space shared-domain.example.com --hostname store --path orders
Creating route store.shared-domain.example.com/orders for org my-org / space my-space as username@example.com...
OK

$ cf create-route my-space shared-domain.example.com --hostname store
Creating route store.shared-domain.example.com for org my-org / space my-space as username@example.com...
OK
```

The developer can then map the new routes to different apps by following the steps in the [Map a Route to Your App](#) section below.

If the developer maps the first route with path `products` to the `products` app, the second route with path `orders` to the `orders` app, and the last route to the `storefront` app. After this, the following occurs:

- PAS routes requests to `http://store.shared-domain.example.com/products` to the `products` app.

- PAS routes requests to `http://store.shared-domain.example.com/orders` to the `orders` app.
- PAS routes requests to `http://store.shared-domain.example.com` to the `storefront` app.

PAS attempts to match routes with a path first, and then attempts to match host and domain.

 **Note:** Routes with the same domain and hostname but different paths can only be created in the same space. Private domains do not have this limitation.

 **Note:** PAS does not route requests for context paths to the root context of an app. Apps must serve requests on the context path.

## Create a TCP Route with a Port

A developer can create a TCP route for `tcp.shared-domain.example.com` on an arbitrary port with the following command. If the clients of the app can accommodate addressing an arbitrary port, then developers should use the `--random-port` to instruct PAS to pick a port for your route.

```
$ cf create-route my-space tcp.shared-domain.example.com --random-port
Creating route tcp.shared-domain.example.com for org my-org / space my-space as user@example.com...
OK
Route tcp.shared-domain.example.com:60034 has been created
```

In this example, PAS routes requests to `tcp.shared-domain.example.com:60034` to apps mapped to this route.

To request a specific port, a developer can use the `--port` flag, so long as the port is not reserved for another space. The following command creates a TCP route for `tcp.shared-domain.example.com` on port 60035:

```
$ cf create-route my-space tcp.shared-domain.example.com --port 60035
Creating route tcp.shared-domain.example.com:60035 for org my-org / space my-space as user@example.com...
OK
```

## List Routes

Developers can list routes for the current space with the [cf routes](#) command. A route is uniquely identified by the combination of hostname, domain, port, and path.

```
$ cf routes
Getting routes as user@private-domain.example.com ...

space host domain port path type apps
my-space myapp shared-domain.example.com myapp
my-space myapp private-domain.example.com myapp
my-space store shared-domain.example.com /products products
my-space store shared-domain.example.com /orders orders
my-space store shared-domain.example.com storefront
my-space shared-domain.example.com 60000 tcp tcp-app
```

Developers can only see routes in spaces where they are members.

## Check Routes

Developers cannot create a route that is already taken. To check whether a route is available, developers can use the [cf check-route](#) command.

The following command checks whether a route with the hostname `store` and the domain `shared-domain.example.com` and the path `products` exists:

```
$ cf check-route store shared-domain.example.com --path /products
Checking for route...
OK
Route store.shared-domain.example.com/products does exist
```

## Map a Route to Your App

For an app to receive requests to a route, developers must map the route to the app with the `cf map-route` command. If the route does not already exist, this command creates it.

**Note:** Any app that is not routed to port 80 or port 443 must be explicitly mapped using the `cf map-route` command. Otherwise, the route is automatically mapped to port 443.

Developers can create and reserve routes for later use by following the steps in the [Manually Map a Route](#) section. Or they can map routes to their app immediately as part of a push by following the steps in the [Map a Route with App Push](#) section.

**Note:** Changes to route mappings are executed asynchronously. On startup, an app will be accessible at its route within a few seconds. Similarly, upon mapping a new route to a running app, the app will be accessible at this route within a few seconds of the CLI exiting successfully.

## Manually Map a Route

Given the following routes and apps:

Route	Apps
store.shared-domain.example.com/products	products
store.shared-domain.example.com/orders	orders
store.shared-domain.example.com	storefront
tcp.shared-domain.example.com:60000	tcp-app

The following commands map the above routes to their respective apps. Developers use hostname, domain, and path to uniquely identify a route to map their apps to.

```
$ cf map-route products shared-domain.example.com --hostname store --path products
$ cf map-route orders shared-domain.example.com --hostname store --path orders
$ cf map-route storefront shared-domain.example.com --hostname store
$ cf map-route tcp-app tcp.shared-domain.example.com --port 60000
```

The following command maps the wildcard route \*.foo.shared-domain.example.com to the app myfallbackapp.

```
$ cf map-route myfallbackapp foo.shared-domain.example.com --hostname '*'
```

## Map a Route with App Push

Developers can map a route to their app with the `cf push` command.

If a domain or hostname is not specified, then a route will be created using the app name and the default shared domain (see [Shared Domains](#)). The following command pushes the app myapp, creating the route myapp.shared-domain.example.com from the default shared domain shared-domain.example.com. If the route has not already been created in another space this command also maps it to the app.

```
$ cf push myapp
```

To customize the route during `push`, specify the domain using the `-d` flag and the hostname with the `--hostname` flag. The following command creates the foo.private-domain.example.com route for myapp:

```
$ cf push myapp -d private-domain.example.com --hostname foo
```

To map a TCP route during `push`, specify a TCP domain and request a random port using `--random-route`. To specify a port, push the app without a route, then create and map the route manually by following the steps in the [Create a TCP Route with a Port](#) section.

```
$ cf push tcp-app -d tcp.shared-domain.example.com --random-route
```

## Map a Route Using App Manifest

Developers can map a route to their app with a manifest by editing the `route` attribute to specify the host, domain, port and/or path components of the route. For more information, see the [Deploying with App Manifests](#) topic.

## Map a Route to Multiple Apps

PAS allows multiple apps, or versions of the same app, to be mapped to the same route. This feature enables Blue-Green deployment. For more information see [Using Blue-Green Deployment to Reduce Downtime and Risk](#).

Routing multiple apps to the same route may cause undesirable behavior in some situations by routing incoming requests randomly to one of the apps on the shared route.

See the [Routing Conflict](#) section of the Troubleshooting App Deployment and Health topic for more information about troubleshooting this problem.

## Map Multiple Routes to One App

You can have multiple routes to an app, but those routes cannot have different context paths.

The following routes are valid for a single app:

Route 1	Route 2
<code>myapp.example.com</code>	<code>myapp.apps.cf.example.com</code>
<code>myapp.example.com/foo</code>	<code>myapp.apps.cf.example.com/foo</code>

The following routes are **not** valid for a single app:

Route 1	Route 2
<code>myapp.example.com/foo</code>	<code>myapp.apps.cf.example.com/bar</code>
<code>myapp.apps.cf.example.com/foo</code>	<code>myapp.example.com/bar</code>

## Map an Internal Route to an App

You can map an internal route to any app. This internal route allows your app to communicate with other apps without leaving the platform. Once mapped, this route becomes available to all other apps on the platform.

This example creates a `foo.apps.internal` internal route for `myapp`:

```
$ cf map-route myapp apps.internal --hostname foo
```

## Unmap a Route

Developers can remove a route from an app using the `cf unmap-route` command. The route remains reserved for later use in the space where it was created until the route is deleted.

To unmap an HTTP route from an app, identify the route using the hostname, domain, and path:

```
$ cf unmap-route tcp-app private-domain.example.com --hostname myapp --path mypath
```

To unmap a TCP route from an app, identify the route using the domain and port:

```
$ cf unmap-route tcp-app tcp.shared-domain.example.com --port 60000
```

## Delete a Route

Developers can delete a route from a space using the `cf delete-route` command.

To delete a HTTP route, identify the route using the hostname, domain, and path:

```
$ cf delete-route private-domain.example.com --hostname myapp --path mypath
```

To delete a TCP route, identify the route using the domain and port.

```
$ cf delete-route tcp.private-domain.example.com --port 60000
```

## Routing Requests to a Specific App Instance

 **Note:** Usage of `X-CF-APP-INSTANCE` is supported only for Diego.

Users can route HTTP requests to a specific app instance using the header `X-CF-APP-INSTANCE`. The format of the header should be `X-CF-APP-INSTANCE: APP_GUID:APP_INDEX`.

`APP_GUID` is an internal identifier for your app. Use the `cf APP-NAME --guid` command to discover the `APP_GUID` for your app.

```
$ cf app myapp --guid
```

`APP_INDEX`, for example `0`, `1`, `2`, or `3`, is an identifier for a particular app instance. Use the CLI command `cf app APP-NAME` to get statistics on each instance of a particular app.

```
$ cf app myapp
```

The following example shows a request made to instance `9` of an app with GUID `5cdc7595-2e9b-4f62-8d5a-a86b92f2df0e` and mapped to route `myapp.private-domain.example.com`.

```
$ curl myapp.private-domain.example.com -H "X-Cf-App-Instance: 5cdc7595-2e9b-4f62-8d5a-a86b92f2df0e:9"
```

If the cf CLI cannot find the instance the format is incorrect, a `404` status code is returned.

## Domains

 **Note:** The term domain in this topic differs from its common use and is specific to Cloud Foundry. Likewise, shared domain and private domain refer to resources with specific meaning in Cloud Foundry. The use of domain name, root domain, and subdomain refers to DNS records.

Domains indicate to a developer that requests for any route created from the domain will be routed to PAS. This requires DNS to be configured out-of-band to resolve the domain name to the IP address of a load balancer configured to forward requests to the CF routers. For more information about configuring DNS, see the [DNS for Domains](#) section.

### List Domains for an Org

When creating a route, developers will select from domains available to them. Use the `cf domains` command to view a list of available domains for the targeted org:

```
$ cf domains
Getting domains in org my-org as user@example.com... OK
name status type
shared-domain.example.com shared
tcp.shared-domain.example.com shared tcp
private-domain.example.com owned
```

This example displays three available domains: a shared HTTP domain `shared-domain.example.com`, a shared TCP domain `tcp.shared-domain.example.com`, and a private domain `private-domain.example.com`. See [Shared Domains](#) and [Private Domains](#).

## HTTP vs. TCP Domains

HTTP domains indicate to a developer that only requests using the HTTP protocol will be routed to apps mapped to routes created from the domain. Routing for HTTP domains is layer 7 and offers features like custom hostnames, sticky sessions, and TLS termination.

TCP domains indicate to a developer that requests over any TCP protocol, including HTTP, will be routed to apps mapped to routes created from the domain. Routing for TCP domains is layer 4 and protocol agnostic, so many features available to HTTP routing are not available for TCP routing. TCP domains are defined as being associated with the TCP Router Group. The TCP Router Group defines the range of ports available to be reserved with TCP Routes. Currently, only Shared Domains can be TCP.

 **Note:** By default, PAS only supports routing of HTTP requests to apps.

## Shared Domains

Admins manage shared domains, which are available to users in all orgs of a PAS deployment. An admin can offer multiple shared domains to users. For example, an admin may offer developers the choice of creating routes for their apps from `shared-domain.example.com` and `cf.some-company.com`.

There is not technically a default shared domain. If a developer pushes an app without specifying a domain (see [Map a Route with App Push](#)), a route will be created for it from the first shared domain created in the system. All other operations involving route require the domain be specified (see [Routes](#)).

Shared domains are HTTP by default, but can be configured to be TCP when associated with the TCP Router Group.

### Create a Shared Domain

Admins can create an HTTP shared domain with the `cf create-shared-domain` command:

```
$ cf create-shared-domain shared-domain.example.com
```

To create a TCP shared domain, first discover the name of the TCP Router Group.

```
$ cf router-groups
Getting router groups as admin ...
name type
default-tcp tcp
```

Then create the shared domain using the `--router-group` option to associate the domain with the TCP router group.

```
$ cf create-shared-domain tcp.shared-domain.example.com --router-group default-tcp
```

### Delete a Shared Domain

Admins can delete a shared domain from PAS with the `cf delete-shared-domain` command:

```
$ cf delete-shared-domain example.com
```

### Internal Domain

The internal domain is a special type of shared domain used for app communication internal to the platform. When you enable service discovery, the internal domain `apps.internal` becomes available for route mapping.

Admins can configure multiple internal domains. First add a custom internal domain name to the `internal_domains` property on the `bosh-dns-adapter` job. Then create an internal domain using the `--internal` option:

```
$ cf create-shared-domain shared-domain.example.com --internal
```

The `--router-group` option is not used with internal domains.

## Private Domains

Org managers can add private domains, or custom domains, and give members of the org permission to create routes for privately registered domain names. Private domains can be shared with other orgs, enabling users of those orgs to create routes from the domain.

Private domains can be HTTP or HTTPS only. TCP Routing is supported for shared domains only.

### Create a Private Domain

Org managers can create a private domain with the following command:

```
$ cf create-domain my-org private-domain.example.com
```

Org managers can create a private domain for a subdomain with the following command:

```
$ cf create-domain my-org foo.private-domain.example.com
```

### Sharing a Private Domain with One or More Orgs

Org managers can grant or revoke access to a private domain to other orgs if they have permissions for these orgs with the following commands:

```
$ cf share-private-domain test-org private-domain.example.com
```

```
$ cf unshare-private-domain test-org private-domain.example.com
```

### Delete a Private Domain

Org managers can delete a domain from PAS with the `cf delete-domain` command:

```
$ cf delete-domain private-domain.example.com
```

## Requirements for Parent and Child Domains

In the domain `myapp.shared-domain.example.com`, `shared-domain.example.com` is the parent domain of subdomain `myapp`. Note the following requirements for domains:

- You can only create a private domain that is parent to a private subdomain.
- You can create a shared domain that is parent to either a shared or a private subdomain.

The domain `foo.myapp.shared-domain.example.com` is the child subdomain of `myapp.shared-domain.example.com`. Note the following requirements for subdomains:

- You can create a private subdomain for a private parent domain only if the domains belong to the same org.
- You can create a private subdomain for a shared parent domain.
- You can only create a shared subdomain for a shared parent domain.
- You cannot create a shared subdomain for a private parent domain.

## DNS for Domains

To create customized access to your apps, you can map specific or wildcard custom domains to PAS by using your DNS provider.

## Mapping Domains to Your Custom Domain

To associate a registered domain name with a domain on PAS, configure a CNAME record with your DNS provider, pointing at any shared domain offered in PAS.

### Mapping a Single Domain to Your Custom Domain

To map a single domain to a custom domain to PAS, configure a CNAME record with your DNS provider.

The following table provides some example CNAME record mappings.

Record Set in Custom Domain	Type	Target in PAS
myapp.yourcustomdomain.com.	CNAME	myapp.shared-domain.example.com
www.yourcustomdomain.com.	CNAME	myapp.shared-domain.example.com

After you create the CNAME mapping, your DNS provider routes your custom domain to `myapp.shared-domain.example.com`.

 **Note:** Refer to your DNS provider documentation to determine whether the trailing `.` is required.

### Mapping Multiple Subdomains to Your Custom Domain

Use a wildcard CNAME record to point all of the subdomains in your custom domain to shared-domain.example.com.

Each separately configured subdomain has priority over the wildcard configuration.

The following table provides some example wildcard CNAME record mappings.

Record Set in Custom Domain	Type	Target in PAS
*.yourcustomdomain.com.	CNAME	*.shared-domain.example.com
*.yourcustomdomain.com.	CNAME	*.myapp.shared-domain.example.com

If you use a wildcard as the subdomain name, then your DNS provider can route from `*.YOURCUSTOMDOMAIN` to any of the following:

- `*.shared-domain.example.com`
- `foo.myapp.shared-domain.example.com`
- `bar.foo.myapp.shared-domain.example.com`

### Configuring DNS for Your Registered Root Domain

To use your root domain (for example, `example.com`) for apps on PAS you can either use custom DNS record types like ALIAS and ANAME, if your DNS provider offers them, or subdomain redirection.

 **Note:** Root domains are also called zone apex domains.

If your DNS provider supports using an ALIAS or ANAME record, configure your root domain with your DNS provider to point at a shared domain in PAS.

Record	Name	Target	Note
ALIAS or ANAME	empty or @	private-domain.example.com.	Refer to your DNS provider documentation to determine whether to use an empty or @ value for the Name entry.

If your DNS provider does not support ANAME or ALIAS records you can use subdomain redirection, also known as domain forwarding, to redirect requests for your root domain to a subdomain configured as a CNAME.

 **Note:** If you use domain forwarding, SSL requests to the root domain may fail if the SSL certificate only matches the subdomain. For more information about SSL certificates, see [Configuring Trusted System Certificates for Apps](#).

Configure the root domain to point at a subdomain such as `www`, and configure the subdomain as a CNAME record pointing at a shared domain in PAS.

Record	Name	Target	Note
URL or Forward	private-domain.example.com	www.private-domain.example.com	This method results in a <a href="#">301 permanent redirect</a> to the subdomain you configure.
CNAME	www	myapp.shared-domain.example.com	

+— title: Using Weighted Routing (Beta)

## owner: Routing

This topic provides information about weighted routing and how to configure weighted routing for your apps.

### Overview

The weighted routing feature is available for Pivotal Application Service (PAS) deployments that use service mesh. For more information, see [Service Mesh \(Beta\)](#).

Weighted routing allows you to map multiple apps to the same route and control the amount of traffic sent to each of them. Some use cases include blue-green deployments, A/B testing, canary releases, or rolling out features over time to incremental user-bases.

### How Weights are Calculated

PAS uses numeric weight values to calculate the percentage of requests that each app receives. For a route with multiple apps mapped, the percentage of requests for a given app is calculated as follows:

*weight of app / sum of the weights of all apps mapped to route*

Consider an example in which two apps are mapped to the same route: `app-a` and `app-b`. The percentage of requests for `app-b` is calculated as follows:

`weight of app-b / ( weight of app-a + weight of app-b )`

The following table uses example values to show how the percentage of requests is calculated for the two apps:

App	Weight	Formula	Percentage of Requests
<code>app-a</code>	<code>1</code>	<code>1 / ( 1 + 3 )</code>	<code>25%</code>
<code>app-b</code>	<code>3</code>	<code>3 / ( 1 + 3 )</code>	<code>75%</code>

In this example, you can also unmap the route from one of the apps and observe `100%` of the requests going to just one app, which simulates a migration.

### Considerations for Working with Routing Weights

Consider the following when working with routing weights:

- The default weight is `1`. This means that if multiple apps are mapped to the same route and their weights are not updated, traffic is equally distributed between them.
- The percentage of requests described above is not applied exactly, but over time the requests reflect the percentage.
- You cannot set a weight to `0`. You can stop weighting traffic by deleting apps or unmapping apps from the route.
- You cannot exceed `100` routes mappings for a particular route. This could result in some apps not receiving any requests.

### Configure Routing Weights

 **Note:** This feature does not yet have cf CLI integration. The steps for updating a routing weight require you to `curl` the Cloud Controller v3 API.

You can configure routing weights when creating or updating a route mapping. Follow the procedure below that corresponds to your use case.

You can repeat the steps for each app that you want to receive a portion of requests at the same route.

## Create a New Route Mapping

This procedure describes how to use `cf curl` to map a route to an app and assign a weight.

1. If you do not already have one, create a route. The route must use the domain created for service mesh:

```
cf create-route SPACE-NAME mesh.YOUR-APPS-DOMAIN -n HOSTNAME
```

Where:

- `SPACE-NAME` is the space in which you want to create the route.
- `YOUR-APPS-DOMAIN` is the shared domain for apps configured by your operator.
- `HOSTNAME` is the hostname to use for creating the route. The format of the route is `HOSTNAME.mesh.YOUR-APPS-DOMAIN`.

2. Run the following command and record the GUID of your route:

```
cf curl /v2/routes?q=host:HOSTNAME
```

Where `HOSTNAME` is the hostname you specified when creating the route.

The GUID of the route is located after `"metadata": {"guid": }`. See the following example in which the GUID is highlighted in blue:

```
$ cf curl /v2/routes?q=host:my-hostname
{
 "total_results": 1,
 "total_pages": 1,
 "prev_url": null,
 "next_url": null,
 "resources": [
 {
 "metadata": {
 "guid": "311d34d1-c045-4853-845f-05132377ad7d",
 "url": "/v2/routes/311d34d1-c045-4853-845f-05132377ad7d",
 "created_at": "2016-06-08T16:41:44Z",
 "updated_at": "2016-06-08T16:41:26Z"
 },
 ...
 }
]
}
```

3. Run the following command and record the GUID of the app that you want to map the route to:

```
cf app APP-NAME --guid
```

4. Create a route mapping and assign a weight:

```
cf curl /v3/route_mappings -X POST -d \
'{
 "relationships": {
 "app": {
 "guid": "APP-GUID"
 },
 "route": {
 "guid": "ROUTE-GUID"
 }
 },
 "weight": WEIGHT
}'
```

Where:

- `APP-GUID` is the app GUID you recorded previously
- `ROUTE-GUID` is the route GUID you recorded previously
- `WEIGHT` is an integer used to calculate the percentage of requests the app receives. You can enter a value from `1` to `128`. See [How Weights are Calculated](#) above.

## Update an Existing Route Mapping

This procedure requires that you have mapped an app to a route in `*.mesh.YOUR-APPS-DOMAIN` either using `cf map-route`, or `cf curl` as described in [Create a New Route Mapping](#) above.

To update an existing route mapping with a new weight, do the following:

1. Run the following command and record the GUID of the app:

```
cf app APP-NAME --guid
```

2. Run the following command and record the GUID of the route:

```
cf curl /v2/routes?q=host:HOSTNAME
```

Where `HOSTNAME` is the hostname you specified when creating the route. This is generally the same as the app name.

The GUID of the route is located after `"metadata": {"guid": ...}`. See the following example in which the GUID is highlighted in blue:

```
$ cf curl /v2/routes?q=host:my-hostname
{
 "total_results": 1,
 "total_pages": 1,
 "prev_url": null,
 "next_url": null,
 "resources": [
 {
 "metadata": {
 "guid": "311d34d1-c045-4853-845f-05132377ad7d",
 "url": "/v2/routes/311d34d1-c045-4853-845f-05132377ad7d",
 "created_at": "2016-06-08T16:41:44Z",
 "updated_at": "2016-06-08T16:41:26Z"
 },
 ...
 }
]
}
```

3. List the route mappings associated with the route:

```
cf curl /v2/routes/ROUTE_GUID/route_mappings
```

Where `ROUTE_GUID` is the route GUID you got from the previous step.

4. In the output from the previous step, find the route mapping that contains the `app_guid` from when you ran `cf app --guid`. Record the GUID of the route mapping. See the following example output in which both the route mapping and app guids are highlighted in blue:

```
$ cf curl /v2/routes/521c375d-a7e2-4f87-9527-7fd1db1b2010/route_mappings
{
 "total_results": 1,
 "total_pages": 1,
 "prev_url": null,
 "next_url": null,
 "resources": [
 {
 "metadata": {
 "guid": "9feb9667-9249-44b7-9c4e-92157a2aaefb",
 "url": "/v2/route_mappings/9feb9667-9249-44b7-9c4e-92157a2aaefb",
 "created_at": "2016-06-08T16:41:28Z",
 "updated_at": "2016-06-08T16:41:26Z"
 },
 "entity": {
 "app_port": null,
 "app_guid": "bf65b03d-5416-4603-9de2-ef74491d29b6",
 "route_guid": "521c375d-a7e2-4f87-9527-7fd1db1b2010",
 "app_url": "/v2/apps/bf65b03d-5416-4603-9de2-ef74491d29b6",
 "route_url": "/v2/routes/521c375d-a7e2-4f87-9527-7fd1db1b2010"
 }
 }
]
}
```

5. Update the route mapping with a new weight:

```
cf curl /v3/route_mappings/ROUTE-MAPPING-GUID -X PATCH -d '{"weight": WEIGHT}'
```

Where:

- `ROUTE_MAPPING_GUID` is the route mapping GUID you got from the previous step.
- `WEIGHT` is an integer used to calculate the percentage of requests the app receives. You can enter a value from `1` to `128`. See [How Weights are Calculated](#) above.

## Confirm Routing Weights

If you have two or more apps mapped to the same route, one way you can confirm routing weights is to use `curl` to view the amount of requests to each app.

To view the amount of requests to each app, do the following:

1. For each app mapped to the route, open a terminal window and run the following command. This allows you to see streaming logs for the apps.

```
cf logs APP-NAME
```

2. In another terminal window, run the following command to send a `curl` request to the route every `N` seconds.

```
while sleep N; do curl ROUTE; done
```

Where:

- `N` is the frequency, in seconds, with which you want to `curl` the route.
- `ROUTE` is the route that you have multiple apps mapped to.

3. Observe each window in which you are running `cf logs APP-NAME`. The output displays a log message each time there is a request. The amount of requests each app receives corresponds to its configured routing weight.

4. When finished, enter `Ctrl + C` to stop sending requests with the `while` loop.

## Configuring CF to Route Traffic to Apps on Custom Ports

This topic describes how to configure Pivotal Cloud Foundry (PCF) to route traffic to your apps listening on custom ports.

### Overview

By default, apps only receive requests on port 8080 for both HTTP and TCP routing, and so must be configured, or hard coded, to listen on this port. Configuring custom app ports allows developers to bring workloads onto PAS that listen on ports other than 8080. Here are some example use cases:

- Serving web client requests on one port and offering stats/debug on another
- Using TCP protocols that require multiple ports
- Running Docker images on PAS

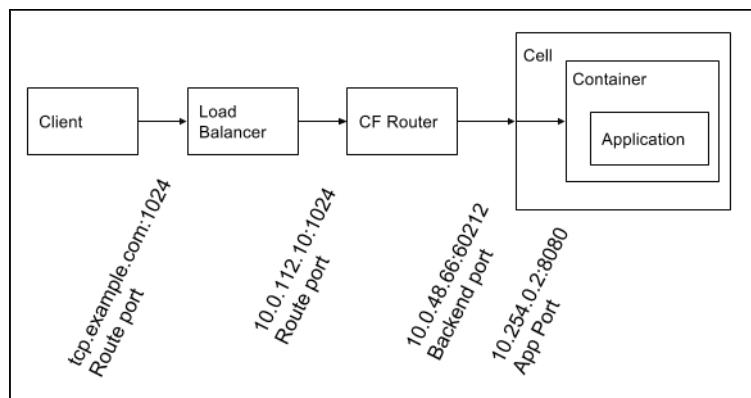
The [Procedure](#) below describes how to use the `apps` and `route_mappings` Cloud Controller API endpoints to update the ports the app can receive requests on.

### Flow of a Request to an App

The following table describes the Network Address Translation that occurs in the data path of a client request. The procedure in this document discusses configuring *app ports* specifically.

Port Type	Description	Network Hop
Route port	The port a client sends a request to	Client to load balancer, load balancer to Gorouter
Backend port	The port on the VM where an app container is hosted, which is unique to the container	Gorouter to Diego Cell
App port	The port on the container; this must match a port the app is configured to listen on	Diego Cell to app container

The following diagram provides an example data path and Network Address Translation for TCP Routing. For HTTP Routing, the Route Port is always 80 or 443.



### Prerequisites

Before following the procedure to configure routing to your app on custom ports, you must have the following:

- An app pushed to PAS that can listen on one or more custom ports.
- Routes for which you want traffic forwarded to your app on custom ports.
  - If your app listens on two ports and you want clients to be able to send requests to both of them, create two routes. These can be from HTTP or TCP domains (see [Routes and Domains](#) for more info). Consider an example in which you have two routes: `foo.example.com` and `bar.example.com`. In the following procedure, you use API endpoints to map these routes to your app on the ports it is listening.

## Procedure

To configure your app to receive HTTP or TCP traffic on custom ports, do the following:

1. Retrieve the GUID of your app:

```
cf app APP-NAME --guid
```

Where `APP-NAME` is the name of your app

2. Configure PAS with the ports your app is listening on:

```
cf curl /v2/apps/APP-GUID -X PUT -d '{"ports": [PORT1, PORT2, PORT3...]}'
```

Where:

- o `APP-GUID` is the GUID of your app
- o `PORT1, PORT2, PORT3...` is a comma-separated list of the ports on which you want your app to receive traffic.

3. Retrieve the GUID of the route to which clients will make requests, and for which Cloud Foundry will route requests to the app on a custom port:

```
cf curl /v2/routes?q=host:HOST-NAME
```

Where `HOST-NAME` is the hostname for the route; by default this is the name of your app.

4. Update the route mapping for your app:

```
cf curl /v2/route_mappings -X POST -d '{"app_guid": "APP-GUID", "route_guid": "ROUTE-GUID", "app_port": PORT1}'
```

Where:

- o `APP-GUID` is the GUID of your app
- o `ROUTE-GUID` is the GUID of the route at which the app serves.
- o `PORT1` is the app port, or one of the app ports, that you added in the previous step.

5. Repeat the previous two steps for each port that you want your app to receive requests on.

## Additional Resources

Here are additional resources related to configuring custom app ports:

- For more information about making requests to the Cloud Controller `apps` endpoint, see [Updating an App](#) in the Cloud Controller API documentation.
- For more information about making requests to the Cloud Controller `route_mappings` endpoint, see [Mapping an App and a Route](#).
- For an example multi-port app, see the [cf-acceptance-tests](#) repository.
- For a demo procedure written by an open source CF user, see [“Multiple App Ports” Demo on Cloud Foundry](#).

## Managing Services

The following list provides information about managing service instances:

- [Services Overview](#)
- [Managing Service Instances](#)
- [Sharing Service Instances](#)
- [Delivering Service Credentials to an App](#)
- [Managing Service Keys](#)
- [Configuring Play Framework Service Connections](#)
- [Using an External File System \(Volume Services\)](#)
- [User-Provided Service Instances](#)

## Managing Service Instances with the cf CLI

This topic describes lifecycle operations for service instances, including creating, updating, and deleting. For an overview of services, and documentation about other service management operations, see [Services Overview](#). If you are interested in building services for Cloud Foundry and making them available to end users, see the [Custom Services](#) documentation.

To run the commands in this topic, you must first install the Cloud Foundry Command Line Interface (cf CLI). See the [Cloud Foundry Command Line Interface](#) topics for more information.

### List Marketplace Services

After targeting and logging into Cloud Foundry, run the `cf marketplace` cf CLI command to view the services available to your targeted organization.

Available services may differ between organizations and between Cloud Foundry marketplaces.

```
$ cf marketplace
Getting services from marketplace in org my-org / space test as user@example.com...
OK

service plans description broker
p-mysql 100mb, 1gb A DBaaS mysql-broker
p-niakcs developer An S3-compatible object store object-store-broker
```

### Creating Service Instances

You can create a service instance with the following command:

```
cf create-service SERVICE PLAN
SERVICE_INSTANCE
```

Use the information in the list below to replace `SERVICE`, `PLAN`, and `SERVICE_INSTANCE` with appropriate values.

- `SERVICE`: The name of the service you want to create an instance of.
- `PLAN`: The name of a plan that meets your needs. Service providers use `plans` to offer varying levels of resources or features for the same service.
- `SERVICE_INSTANCE`: The name you provide for your service instance. You use this name to refer to your service instance with other commands. Service instance names can include alpha-numeric characters, hyphens, and underscores, and you can rename the service instance at any time.

```
$ cf create-service rabbitmq small-plan my-rabbitmq
Creating service my-rabbitmq in org console / space development as user@example.com...
OK
```

User Provided Service Instances provide a way for developers to bind apps with services that are not available in their Cloud Foundry marketplace. For more information, see the [User Provided Service Instances](#) topic.

**Note:** When multiple brokers provide two or more services with the same name, you must specify the broker by including the `-b BROKER` flag in the `cf create-service` command.

### Arbitrary Parameters

*Arbitrary parameters require cf CLI v6.12.1+*

Some services support providing additional configuration parameters with the provision request. Pass these parameters in a valid JSON object containing service-specific configuration parameters, provided either in-line or in a file. For a list of supported configuration parameters, see the documentation for the particular service offering.

Example providing service-specific configuration parameters in-line:

```
$ cf create-service my-db-service small-plan my-db -c '{"storage_gb":4}'
```

```
Creating service my-db in org console / space development as user@example.com...
OK
```

Example providing service-specific configuration parameters in a file:

```
$ cf create-service my-db-service small-plan my-db -c /tmp/config.json
```

```
Creating service my-db in org console / space development as user@example.com...
OK
```

## Instance Tags

*Instance tags require cf CLI v6.12.1+*

Some services provide a list of tags that Cloud Foundry delivers in the [VCAP\\_SERVICES Environment Variable](#). These tags provide developers with a more generic way for apps to parse `|VCAP_SERVICES|` for credentials. Developers may provide their own tags when creating a service instance by including the `-t` flag followed by a comma-separated list of tags.

Example providing a comma-separated list of tags:

```
$ cf create-service my-db-service small-plan my-db -t "prod, workers"
```

```
Creating service my-db in org console / space development as user@example.com...
OK
```

## List Service Instances

Run the `cf services` command to list the service instances in your targeted space. The output from running this command includes any bound apps and the state of the last requested operation for the service instance.

```
$ cf services
Getting services in org my-org / space test as user@example.com...
OK
```

```
name service plan bound apps last operation broker
mybucket p-riakcs developer myapp create succeeded object-store-broker
mydb p-mysql 100mb create succeeded mysql-broker
```

## Get Details for a Particular Service Instance

Details include dashboard urls, if applicable, and operation start and last updated timestamps.

```
$ cf service mydb
service instance: mydb
service: p-mysql
plan: 100mb
description: mysql databases on demand
documentation url:
dashboard: https://p-mysql.example.com/manage/instances/abcd-ef12-3456
service broker: mysql-broker
```

This service is not currently shared.

Showing status of last operation from service mydb...

```
status: create succeeded
message:
started: 2019-02-13T12:02:19Z
updated: 2019-02-13T12:02:19Z
```

There are no bound apps for this service.

## Bind a Service Instance

Depending on the service, you can bind service instances to apps and/or routes.

Not all services support binding, as some services deliver value to users directly without integration with Cloud Foundry, such as SaaS apps.

### Bind a Service Instance to an App

Depending on the service, binding a service instance to your app may deliver credentials for the service instance to the app. See the [Delivering Service Credentials to an App](#) topic for more information. Binding a service instance to an app may also trigger app logs to be streamed to the service instance. For more information, see [Streaming App Logs to Log Management Services](#).

```
$ cf bind-service my-app mydb
Binding service mydb to my-app in org my-org / space test as user@example.com...
OK
TIP: Use 'cf push' to ensure your env variable changes take effect

$ cf restart my-app
```

 **Note:** You must restart or in some cases re-push your app for changes to be applied to the `VCAP_SERVICES` environment variable and for the app to recognize these changes.

### Binding with App Manifest

As an alternative to binding a service instance to an app after pushing an app, you can use the app manifest to bind the service instance during push. As of cf CLI v6.12.1, [Arbitrary Parameters](#) are not supported in app manifests. Using the manifest to bind service instances to routes is also not supported.

The following excerpt from an app manifest binds a service instance called `test-mysql-01` to the app on push.

```
services:
- test-mysql-01
```

The following excerpt from the `cf push` command and response demonstrates that the cf CLI reads the manifest and binds the service instance to an app called `test-msg-app`.

```
$ cf push
Using manifest file /Users/Bob/test-apps/test-msg-app/manifest.yml

...
Binding service test-mysql-01 to test-msg-app in org My-Org / space development as user@example.com
OK
```

For more information about app manifests, see [Deploying with App Manifests](#).

### Bind a Service Instance to a Route

Binding a service instance to a route will cause app requests and responses to be proxied through the service instance, where it may be used to transform or intermediate requests. For more information, see [Manage App Requests with Route Services](#).

```
$ cf bind-route-service shared-domain.example.com --hostname my-app my-service-instance
Binding route my-app.shared-domain.example.com to service instance my-service-instance in org my-org / space test as user@example.com...
OK
```

Restaging your app is not required.

### Arbitrary Parameters

*Arbitrary parameters require cf CLI v6.12.1+*

Some services support additional configuration parameters with the bind request. These parameters are passed in a valid JSON object containing service-specific configuration parameters, provided either in-line or in a file. For a list of supported configuration parameters, see documentation for the particular service offering.

```
$ cf bind-service rails-sample my-db -c '{"role":"read-only"}'
Binding service my-db to app rails-sample in org console / space development as user@example.com...
OK
```

```
$ cf bind-service rails-sample my-db -c /tmp/config.json
Binding service my-db to app rails-sample in org console / space development as user@example.com... OK
```

## Unbind a Service Instance

### Unbind a Service Instance from an App

Unbinding a service instance from an app removes the credentials created for your app from the [VCAP\\_SERVICES](#) environment variable.

```
$ cf unbind-service my-app mydb
Unbinding app my-app from service mydb in org my-org / space test as user@example.com...
OK
```

 **Note:** You must restart or in some cases re-push your app for changes to be applied to the [VCAP\\_SERVICES](#) environment variable and for the app to recognize these changes.

### Unbind a Service Instance from a Route

Unbinding a service instance from a route will result in requests and responses no longer being proxied through the service instance. For more information, see [Manage App Requests with Route Services](#).

 **Note:** If your bound service instance is providing security features, like authorization, unbinding the service instance may leave your app vulnerable.

```
$ cf unbind-route-service shared-domain.example.com --hostname my-app my-service-instance
Unbinding may leave apps mapped to route my-app.shared-domain.example.com vulnerable; e.g. if service instance my-service-instance provides authentication. Do you want to proceed?> y
Unbinding route my-app.shared-domain.example.com from service instance my-service-instance n org my-org / space test as user@example.com...
OK
```

Restaging your app is not required.

## Rename a Service Instance

You can change the name given to a service instance. Keep in mind that upon restarting any bound apps, the name of the instance will change in the [VCAP\\_SERVICES](#) environment variable. If your app depends on the instance name for discovering credentials, changing the name could break your app's use of the service instance.

```
$ cf rename-service mydb mydb1
Renaming service mydb to mydb1 in org my-org / space test as user@example.com...
OK
```

## Update a Service Instance

## Upgrade/Downgrade Service Plan

*Changing a plan requires cf CLI v6.7+ and cf-release v192+*

By updating the service plan for an instance, users can effectively upgrade and downgrade their service instance to other service plans. Though the platform and CLI now support this feature, services must expressly implement support for it so not all services will. Further, a service might support updating between some plans but not others. For instance, a service might support updating a plan where only a logical change is required, but not where data migration is necessary. In either case, users can expect to see a meaningful error when plan update is not supported.

```
$ cf update-service mydb -p new-plan
Updating service instance mydb as user@example.com...
OK
```

## Arbitrary Parameters

*Arbitrary parameters require cf CLI v6.12.1+*

Some services support additional configuration parameters with the update request. These parameters are passed in a valid JSON object containing service-specific configuration parameters, provided either in-line or in a file. For a list of supported configuration parameters, see documentation for the particular service offering.

```
$ cf update-service mydb -c '{"storage_gb":4}'
Updating service instance mydb as me@example.com...
```

```
$ cf update-service mydb -c /tmp/config.json
Updating service instance mydb as user@example.com...
```

## Instance Tags

*Instance tags require cf CLI v6.12.1+*

Some services provide a list of tags that Cloud Foundry delivers in the [VCAP\\_SERVICES Environment Variable](#). These tags provide developers with a more generic way for apps to parse `VCAP_SERVICES` for credentials. Developers may provide their own tags when creating a service instance by including a comma-separated list of tags with the `-t` flag.

```
$ cf update-service my-db -t "staging, web"
Updating service my-db in org console / space development as user@example.com...
OK
```

## Delete a Service Instance

Deleting a service instance deprovisions the service instance and deletes all data associated with the service instance.

```
$ cf delete-service mydb
Are you sure you want to delete the service mydb ? y
Deleting service mydb in org my-org / space test as user@example.com...
OK
```

## Sharing Service Instances

This topic explains how to use service instance sharing.

### About Service Instance Sharing

Sharing a service instance between spaces allows apps in different spaces to share databases, messaging queues, and other types of services. This eliminates the need for development teams to use service keys and user-provided services to bind their apps to the same service instance that was provisioned using the `cf create-service` command. Sharing service instances improves security, auditing, and provides a more intuitive user experience.

- Service instances can be shared into multiple spaces and across orgs.
- Developers and administrators can share service instances between spaces in which they have the Space Developer role.
- Developers who have a service instance shared with them can only bind and unbind apps to that service instance. They cannot update, rename, or delete it.
- Developers who have a service instance shared with them can view the values of any configuration parameters that were used to provision or update the service instance.

For example, if two development teams have apps in their own spaces, and both of those apps want to send messages to each other using a messaging queue, you can do the following:

1. The development team in space A can create a new instance of a messaging queue service, bind it to their app, and share that service instance into space B.
2. A developer in space B can then bind their app to the same service instance, and the two apps can begin publishing and receiving messages from one another.

### Enabling Service Instance Sharing in Cloud Foundry

To enable service instance sharing, the platform operator must enable the `service_instance_sharing` flag in Cloud Foundry.

```
$ cf enable-feature-flag service_instance_sharing
```

### Sharing a Service Instance

You can share a service instance from one space to another if you have the Space Developer role in both spaces.

To share a service instance to another space, run the following Cloud Foundry Command Line Interface (cf CLI) command:

```
$ cf share-service SERVICE-INSTANCE -s OTHER-SPACE [-o OTHER-ORG]
```

- You cannot share a service instance into a space where a service instance with the same name already exists.
- To share a service instance into a space, the space must have access to the service and service plan of the service instance that you are sharing. Run the `cf enable-service-access` command to set this access.
- If you no longer have access to the service or service plan used to create your service instance, you cannot share that service instance.

### Unsharing a Service Instance

**A warning:** Unsharing a service instance automatically deletes all bindings to apps in the spaces it was shared into. This may cause apps to fail. Before unsharing a service instance, run the `cf service SERVICE-INSTANCE` command to see how many bindings exist in the spaces the service instance is shared into.

You can unshare a service instance if you have the Space Developer role in the space where this service instance was shared from.

Developers cannot delete or rename a service instance until it is unshared from all spaces.

To unshare a service instance, run the following cf CLI command:

```
$ cf unshare-service SERVICE-INSTANCE -s OTHER-SPACE [-o OTHER-ORG] [-f]
```

The optional `-f` flag forces unsharing without confirmation.

## Security Considerations

- [Service keys](#) cannot be created from a space that a service instance has been shared into. This ensures that developers in the space where a service instance has been shared from have visibility into where and how many times the service instance is used.
- Sharing service instances does not automatically update app security groups (ASGs). The network policies defined in your ASGs may need to be updated to ensure that apps using shared service instances can access the underlying service.
- Access to a service must be enabled using the `cf enable-service-access` command for a service instance to be shared into a space.
- Not all services are enabled for sharing instances functionality. Contact the service vendor directly if you are unable to share instances of their service. If you are a service author, see [Enabling Service Instance Sharing](#).

## Disabling Service Instance Sharing in Cloud Foundry

To disable service instance sharing, run the following command:

```
$ cf disable-feature-flag service_instance_sharing
```

This only prevents new shares from being created. To remove existing shares, see [Deleting All Shares](#).

## Deleting All Shares

The script below finds all service instances that are shared, and for each space that the service instance is shared into, all service bindings to that service instance are deleted, and all shares are deleted.

If a service binding is not successfully deleted, the script continues trying to unshare subsequent service instances.

To use this script, you must be logged in as an administrator and have jq installed.

 **Note:** This script has been tested on macOS Sierra 10.12.4 and Ubuntu 14.04.5. Use the script at your own risk.

```
#!/usr/bin/env bash

set -u
set -e

refresh auth token
cf oauth-token >/dev/null

for instance_guid in $(cf curl /v3/service_instances | jq -r '.resources[].guid'); do
 for space_guid in $(cf curl /v2/service_instances/$instance_guid/shared_to | jq -r '.resources[].space_guid'); do
 echo "Unsharing service instance $instance_guid from space $space_guid"

 set +e
 cf curl -X DELETE "/v3/service_instances/$instance_guid/relationships/shared_spaces/$space_guid"
 set -e
 done
done
```

## Delivering Service Credentials to an App

This topic describes binding apps to service instances for the purpose of generating credentials and delivering them to apps. For an overview of services, and documentation on other service management operations, see [Using Services](#). If you are interested in building services for Cloud Foundry and making them available to end users, see the [Custom Services](#) documentation.

### Bind a Service Instance

Binding a service instance to your app triggers credentials to be provisioned for the service instance and delivered to the app runtime in the `VCAP_SERVICES` environment variable. For details on consuming these credentials with your app, see [Using Bound Service Instances](#).

Not all services support binding, as some services deliver value to users directly without integration with an app. In many cases binding credentials are unique to an app, and another app bound to the same service instance would receive different credentials; however this depends on the service.

```
$ cf bind-service my-app mydb
Binding service mydb to my-app in org my-org / space test as me@example.com...
OK
TIP: Use 'cf push' to ensure your env variable changes take effect

$ cf restart my-app
```

 **Note:** You must restart or in some cases re-push your app for changes to be applied to the `VCAP_SERVICES` environment variable and for the app to recognize these changes.

### Arbitrary Parameters

*Arbitrary parameters require Cloud Foundry Command Line Interface (cf CLI) v6.12.1 or later.*

Some services support additional configuration parameters with the bind request. These parameters are passed in a valid JSON object containing service-specific configuration parameters, provided either in-line or in a file. For a list of supported configuration parameters, see documentation for the particular service offering.

```
$ cf bind-service rails-sample my-db -c '{"role":"read-only"}'
Binding service my-db to app rails-sample in org console / space development as user@example.com...
OK

$ cf bind-service rails-sample my-db -c /tmp/config.json
Binding service my-db to app rails-sample in org console / space development as user@example.com... OK
```

### Binding with App Manifest

As an alternative to binding a service instance after pushing an app, you can use the app manifest to bind the service instance during push. As of cf CLI v6.12.1, [Arbitrary Parameters](#) are not supported in app manifests.

The following excerpt from an app manifest would bind a service instance called `test-mysql-01` to the app on push.

```
services:
- test-mysql-01
```

The following excerpt from the `cf push` command and response demonstrates that the cf CLI reads the manifest and binds the service instance to an app called `test-msg-app`.

```
$ cf push
Using manifest file /Users/Bob/test-apps/test-msg-app/manifest.yml

...
Binding service test-mysql-01 to test-msg-app in org My-Org / space development as Bob@shared-domain.example.com
OK
```

For more information about app manifests, see [Deploying with App Manifests](#).

## Named Service Bindings

Service offering and service instance names vary across environments. App authors can discover the service instances their apps require, without having to know environment-specific service names. Developers can specify a service binding name to be included in [VCAP\\_SERVICES](#).

To specify the service binding name, provide the `--binding-name` flag when binding an app to a service instance:

```
$ cf bind-service my-app my-service --binding-name postgres-database
OK
```

The provided name will be available in the `name` and `binding_name` properties in the [VCAP\\_SERVICES](#) environment variable:

```
"VCAP_SERVICES": {
 "service-name": [
 {
 "name": "postgres-database",
 "binding_name": "postgres-database",
 ...
 }
]
}
```

## Using Bound Service Instances

Once you have a service instance created and bound to your app, you need to configure the app to dynamically fetch the credentials for your service instance. The [VCAP\\_SERVICES](#) environment variable contains credentials and additional metadata for all bound service instances. There are two methods developers can leverage to have their apps consume binding credentials.

- **Parse the JSON yourself:** See the documentation for [VCAP\\_SERVICES](#). Helper libraries are available for some frameworks.
- **Auto-configuration:** Some buildpacks create a service connection for you by creating additional environment variables, updating config files, or passing system parameters to the JVM.

For details on consuming credentials specific to your development framework, refer to the Service Binding section in the documentation for [your framework's buildpack](#).

## Update Service Credentials

To update your service credentials, perform the following steps:

1. [Unbind the service instance](#) using the credentials you are updating with the following command:

```
$ cf unbind-service YOUR-APP YOUR-SERVICE-INSTANCE
```

2. [Bind the service instance](#) with the following command. This adds your credentials to the [VCAP\\_SERVICES](#) environment variable.

```
$ cf bind-service YOUR-APP YOUR-SERVICE-INSTANCE
```

3. Restart or re-push the app bound to the service instance so that the app recognizes your environment variable updates.

## Unbind a Service Instance

Unbinding a service removes the credentials created for your app from the [VCAP\\_SERVICES](#) environment variable.

```
$ cf unbind-service my-app mydb
Unbinding app my-app from service mydb in org my-org / space test as me@example.com...
OK
```

 **Note:** You must restart or in some cases re-push your app for changes to be applied to the [VCAP\\_SERVICES](#) environment variable and for the app to recognize these changes.

## Managing Service Keys

This topic describes managing service instance credentials with service keys.

Service keys generate credentials for manually configuring consumers of marketplace services. Once you configure them for your service, local clients, apps in other spaces, or entities outside your deployment can access your service with these keys.

 **Note:** Some service brokers do not support service keys. If you want to build a service broker that supports service keys, see [Services](#). If you want to use a service broker that does not support service keys, see [Delivering Service Credentials to an App](#).

### Create a Service Key

To generate credentials for a service instance, use the `cf create-service-key` command:

```
$ cf create-service-key MY-SERVICE MY-KEY
Creating service key MY-KEY for service instance MY-SERVICE as me@example.com...
OK
```

Use the `-c` flag to provide service-specific configuration parameters in a valid JSON object, either in-line or in a file.

To provide the JSON object in-line, use the following format:

```
$ cf create-service-key MY-SERVICE MY-KEY -c '{"read-only":true}'
Creating service key MY-KEY for service instance MY-SERVICE as me@example.com...
OK
```

To provide the JSON object as a file, give the absolute or relative path to your JSON file:

```
$ cf create-service-key MY-SERVICE MY-KEY -c PATH-TO-JSON-FILE
Creating service key MY-KEY for service instance MY-SERVICE as me@example.com...
OK
```

### List Service Keys for a Service Instance

To list service keys for a service instance, use the `cf service-keys` command:

```
$ cf service-keys MY-SERVICE
Getting service keys for service instance MY-SERVICE as me@example.com...

name
mykey1
mykey2
```

### Get Credentials for a Service Key

To retrieve credentials for a service key, use the `cf service-key` command:

```
$ cf service-key MY-SERVICE MY-KEY
Getting key MY-KEY for service instance MY-SERVICE as me@example.com...

{
 uri: foo://user2:pass2@example.com/mydb,
 servicename: mydb
}
```

Use the `--guid` flag to display the API GUID for the service key:

```
$ cf service-key -guid MY-SERVICE MY-KEY
Getting key MY-KEY for service instance MY-SERVICE as me@example.com...
```

```
e3696feb-7a8f-437f-8692-436558e45c7b
```

OK

## Configure Credentials for a Service Key

Once these credentials are obtained, you can use a local CLI or utility to connect to the service instance, configure an app running outside the platform to connect to the service instance, or create a user-provided service instance so that apps in another space can connect to the service instance. How you configure these credentials will depend on what local client, app, or entity is used to access your service instance.

For more information on configuring credentials with a user-provided service instance, see [User-Provided Service Instances](#).

## Delete a Service Key

To delete a service key, use the `cf delete-service-key` command:

```
$ cf delete-service-key MY-SERVICE MY-KEY
```

```
Are you sure you want to delete the service key MY-KEY ? y
Deleting service key MY-KEY for service instance MY-SERVICE as me@example.com...
```

OK

Add option `-f` to force deletion without confirmation.

```
$ cf delete-service-key -f MY-SERVICE MY-KEY
```

```
Deleting service key MY-KEY for service instance MY-SERVICE as me@example.com...
```

OK

## Configuring Play Framework Service Connections

Cloud Foundry provides support for connecting a Play Framework app to services such as MySQL, and Postgres. In many cases, a Play Framework app running on Cloud Foundry can automatically detect and configure connections to services.

### Auto-Configuration

By default, Cloud Foundry will detect service connections in a Play Framework app and configure them to use the credentials provided in the Cloud Foundry environment. Auto-configuration will only happen if there is a single service of any of the supported types - MySQL or Postgres.

## Using an External File System (Volume Services)

This topic describes how Pivotal Cloud Foundry (PCF) app developers can read and write to a mounted file system from their apps. In PCF, a volume service provides a volume so your app can read or write to a reliable, non-ephemeral file system.

 **Note:** NFS and SMB volume services are available for Linux cells only. These services are not available for Windows cells.

### Prerequisite

Before you can use a volume service with your app, your Cloud Foundry administrator must add a volume service to your deployment. See the [Enabling Volume Services](#) topic for more information.

You can run the Cloud Foundry Command Line Interface (cf CLI) `cf marketplace` command to determine if any volume services are available. See the following example output of the NFS volume service:

```
$ cf marketplace
service plans description
nfs Existing Service for connecting to NFS volumes
```

If no volume service that fits your requirements exists, contact your Cloud Foundry administrator.

### Mount an External Filesystem

The sections below describe how to mount an external filesystem to your app.

#### Create and Bind a Service Instance

To use a volume service deployed by your Cloud Foundry administrator, you must first create an instance of the specific volume service that you need. Follow the instructions below to create this service instance.

 **Note:** For NFS-specific instructions and information, see [NFS Volume Service](#).

1. In a terminal window, run `cf create-service SERVICE-NAME PLAN SERVICE-INSTANCE -c SHARE-JSON` to create a service instance. Replace the following with the specified values:
  - `SERVICE` : The name of the volume service that you want to use.
  - `PLAN` : The name of the service plan. Service plans are a way for providers to offer varying levels of resources or features for the same service.
  - `SERVICE-INSTANCE` : A name you provide for your service instance. Use any series of alpha-numeric characters, hyphens, and underscores. You can rename the instance at any time.
  - `SHARE-JSON` (NFS Only): If you create an instance of the NFS volume service, you must supply an extra parameter, `share`, by using the `-c` flag with a JSON string, in-line or in a file. This parameter forwards information to the broker about the NFS server and share required for the service.

The following example shows creating an instance of the “Existing” NFS service plan, passing an in-line JSON string:

```
$ cf create-service nfs Existing nfs_service_instance -c '{"share": "10.10.10.10/export/myshare"}'
```

2. Run `cf bind-service YOUR-APP SERVICE-NAME -c GID-AND-UID-JSON MOUNT-PATH READ-ONLY-TRUE` to bind your service instance to an app. Replace the following with the specified values:
  - `YOUR-APP` : The name of the PCF app for which you want to use the volume service.
  - `SERVICE-NAME` : The name of the volume service instance you created in the previous step.
  - `GID-AND-UID-JSON` (NFS only): If you bind an instance of the NFS volume service, you must supply two extra parameters, `gid` and `uid`. You can specify these parameters with the `-c` flag and a JSON string, in-line or from a file. This parameter specifies the `gid` and `uid` to use when mounting the share to the app.
  - `MOUNT-PATH` (Optional): To mount the volume to a particular path within your app rather than the default path, you supply the `mount` parameter. Choose a path with a root-level folder that already exists in the container, such as `/home`, `/usr`, or `/var`.

 **Note:** Do not specify a `MOUNT-PATH` within the `/app` directory. For more information, see [Mount Path Limitation](#).

- o `READ-ONLY-TRUE` (Optional): When you issue the `cf bind-service` command, Volume Services mounts a read-write file system by default. You can specify a read-only mount by adding `"readonly":true` to the bind configuration JSON string.

The following example shows binding `my-app` to the `nfs_service_instance` and specifying a read-only volume to be mounted to `/var/volume1`, passing an in-line JSON string:

```
$ cf bind-service my-app nfs_service_instance -c '{"uid":"1000","gid":"1000","mount":"/var/volume1","readonly":true}'
```

If you use an LDAP server, you must specify `username` and `password` instead of a UID and GID in this command. For example:

```
$ cf bind-service my-app nfs_service_instance -c '{"username":"user1000","password":"secret","mount":"/var/volume1","readonly":true}'
```

3. Run `cf restage YOUR-APP` to complete the service binding by restaging your app. Replace `YOUR-APP` with the name of your app.

```
$ cf restage my-app
```

## Access the Volume Service from your App

To access the volume service from your app, you must know which file path to use in your code. You can view the file path in the details of the service binding, which are available from the `VCAP_SERVICES` environment variable. Follow the steps below.

1. Run `cf env YOUR-APP` to view environment variables for your app. Replace `YOUR-APP` with the name of your app.

```
$ cf env my-app
"VCAP_SERVICES": {
 "nfs": [
 {
 "credentials": {},
 "label": "nfs",
 "name": "nfs_service_instance",
 "plan": "Existing",
 "provider": null,
 "syslog_drain_url": null,
 "tags": [
 "nfs"
],
 "volume_mounts": [
 {
 "container_dir": "/var/vcap/data/153e3c4b-1151-4cf7-b311-948dd77fce64",
 "device_type": "shared",
 "mode": "rw"
 }
]
 }
]
}
```

2. Use the properties under `volume_mounts` for any information your app needs. Refer to the following table:

Property	Description
<code>container_dir</code>	String containing the path to the mounted volume that you bound to your app.
<code>device_type</code>	The NFS volume release. This currently only supports <code>shared</code> devices. A <code>shared</code> device represents a distributed file system that can mount on all app instances simultaneously.
<code>mode</code>	String that informs what type of access your app has to NFS, either read-only, <code>ro</code> , or read and write, <code>rw</code> .

## Bind with an App Manifest

You can also bind volume services using an app manifest as described in the [Services](#) section of the [Deploying with App Manifests](#) topic. However, app manifests do not support bind configuration. If you want to bind a volume service using an app manifest, you must specify bind configuration when you create the service instance. The releases that support this are `nfs-volume` v1.3.1 and later and `smb-volume` v1.0.0 and later.

## Mount Path Limitation

Do not specify a `MOUNT-PATH` within the `/app` directory, which is where PAS unpacks the droplet. If you specify a mount inside the `/app` directory, the app may fail to start and parts of the app droplet may be written to the remote file share. This is because PAS mounts the volume before moving your compiled app into the droplet.

If your app requires the shared volume to be placed within the `/app` directory, do the following:

1. Specify a mount volume in a location outside of the `/app` directory.
2. Create a symbolic link at app startup time, prior to launching the app. For example:

```
cf push my-app -c "ln -s /var/volume1 /app/volume1 && \$HOME/boot.sh"
```

## NFS Volume Service

This section describes how to use the NFS volume service.

### Create an NFS Volume Service

Cloud Foundry offers two NFS volume services:

- `nfs`: This volume service provides support for NFS volumes using both v3 and v4.x protocols.
  - If you do not specify a `version` parameter in the create configuration, the protocol version used will be negotiated between client and server at mount time. This usually results in the latest available version being used.
  - If you want to request a specific protocol version, include a `version` parameter in the create configuration.
  - For read-only mounts, the driver enables attribute caching. This results in fewer attribute RPCs and better performance.
  - When you omit `uid` and `gid` or `username` and `password` in bind configuration, the driver skips `mapfs` mounting and performs just the normal kernel mount of the NFS file system without the overhead associated with FUSE mounts.
- `nfs-legacy` (deprecated): Although it is deprecated, this volume service is still available due to the difficulty of retiring services. If you use this service, it performs exactly the same mount as the `nfs` service.

Both services offer a single plan called `Existing`.

To create an NFS volume service, follow the procedure below that corresponds to your use case.

#### Create an `nfs` Service (auto-negotiated version)

You can create a NFS volume service using the `Existing` plan of the `nfs` service. Run the following command:

```
$ cf create-service nfs Existing SERVICE-INSTANCE-NAME -c '{"share":"SERVER/SHARE"}'
```

Where:

- `SERVICE-INSTANCE-NAME` is a name you provide for this NFS volume service instance.
- `SERVER/SHARE` is the NFS address of your server and share.

 **Note:** Ensure you omit the `:` that usually follows the server name in the address.

You can run the `cf services` command to confirm that your newly-created NFS volume service displays in the output.

#### Create an `nfs` Service (explicit version)

You can create a NFS volume service with a specific NFS protocol version using the `Existing` plan of the `nfs` service. Run the following command:

```
$ cf create-service nfs Existing SERVICE-INSTANCE-NAME -c '{"share":"SERVER/SHARE", "version":"NFS-PROTOCOL"}'
```

Where:

- `SERVICE-INSTANCE-NAME` is a name you provide for this NFS volume service instance.
- `SERVER/SHARE` is the NFS address of your server and share.
- `VERSION` is the NFS protocol you want to use. For example, if you want to use NFSv4, set the version to `4.1`. Valid values are `3.0`, `4.0`, `4.1` or `4.2`.

 Note: Ensure you omit the `:` that usually follows the server name in the address.

You can run the `cf services` command to confirm that your newly-created NFS volume service displays in the output.

### Existing `nfs-legacy` Services:

With the release of NFS Volume Service 1.5.4, the original fuse-based NFS service has been deprecated in favor of the newer kernel-mount based NFS service and will eventually be removed. Existing NFS volume service bindings will now be listed as `nfs-legacy`. To switch over from `nfs-legacy` to the newer `nfs` service, we recommend that you re-create and re-bind your `nfs` service instances.

With the release of NFS Volume Service 2.0.0, the `nfs-legacy` service has been rewritten to use the `nfs` service. To avoid being affected when the `nfs-legacy` service is retired, re-create and re-bind your service instances using the `nfs` service.

## Deploy and Bind a Sample App

This section describes how to deploy a sample app and bind it to the NFS volume service.

1. Clone the github repo and push the `pora` test app:

```
a. cd ~/workspace
b. git clone https://github.com/cloudfoundry/porsi-acceptance-tests.git
c. cd ~/workspace/porsi-acceptance-tests/assets/pora
d. cf push pora --no-start
```

2. To bind the service to your app, run the following command:

```
$ cf bind-service pora SERVICE-INSTANCE-NAME -c '{"uid":"UID","gid":"GID"}'
```

Where:

- `SERVICE-INSTANCE-NAME`: The name of the volume service instance you created previously.
- `UID` and `GID`: The `gid` and `uid` to use when mounting the share to the app. The NFS driver uses these values in the following ways:
  - When sending traffic to the NFS server, the NFS driver translates the app user id and group id to the `UID` and `GID` values.
  - When returning attributes from the NFS server, the NFS driver translates the `UID` and `GID` back to the running user uid and default gid.

This allows you to interact with your NFS server as a specific user while allowing Cloud Foundry to run your app as an arbitrary user.

- **Optional parameters:**

- `mount`: Use this option to specify the path at which volumes mount to the app container. The default is an arbitrarily-named folder in `/var/vcap/data`. You may need to modify this value if your app has specific requirements. For example:

```
cf bind-service pora myVolume -c '{"uid":0,"gid":0,"mount":"/var/path"}'
```

- `readonly`: When you issue the `cf bind-service` command, Volume Services mounts a read-write file system by default. You can specify a read-only mount by adding `"readonly":true` to the bind configuration JSON string.

3. Start the app:

```
$ cf start pora
```

4. Use the following `curl` command to confirm the app is running. The command returns an instance index for your app.

```
$ curl http://pora.YOUR-CF-DOMAIN.com
```

5. Use the following `curl` command to confirm the app can access the shared volume. The command writes a file to the share and then reads it back out again.

```
$ curl http://pora.YOUR-CF-DOMAIN.com/write
```

## Additional Information

This section provides additional information about using the NFS Volume Service.

### Binding with LDAP Credentials

If your Cloud Foundry deployment has LDAP enabled, you can bind using LDAP credentials.

To bind an app to your volume using LDAP credentials, specify `username` and `password` instead of `uid` and `gid`. See the following example:

```
$ cf bind-service pora myVolume -c '{"username":"USERNAME","password":"PASSWORD"}'
```

 **Note:** If your LDAP server password changes, you must re-bind your app to the service and restage. If you do not, your app will fail during restart or scaling. This is because user credentials are stored as part of the service binding and checked whenever an app is placed on a cell.

### Specifying Bind Parameters During Service Instance Creation

As of `nfs-volume-release` v1.3.1, you can specify bind parameters in advance, when you create a service instance. Use this option if you bind the service to your app in an app manifest, where bind configuration is not supported.

### File Locking with `flock()` and `lockf()` / `fcntl()`

Apps that use file locking through unix system calls such as `flock()` and `fentl()` or script commands such as `flock` may use the `nfs` service. The `nfs-legacy` service uses a fuse mounting process that does not enforce locks across Diego cells.

## SMB Volume Service

This section describes how to use an SMB ([Server Message Block](#)) volume service.

### Create an SMB Volume Service

Cloud Foundry offers one SMB volume service:

- `smb`: This volume service provides support for existing SMB shares.

The service offers a single plan called `Existing`.

To create an SMB volume service, follow the procedure below.

### Create with `smb` Service

You can create an SMB volume service using the `Existing` plan of the `smb` service. Run the following command:

```
$ cf create-service smb Existing SERVICE-INSTANCE-NAME -c '{"share":"//SERVER/SHARE"}'
```

Where:

- `SERVICE-INSTANCE-NAME` is a name you provide for this SMB volume service instance.

- `//SERVER/SHARE` is the SMB address of your server and share.

You can run the `cf services` command to confirm that your newly-created SMB volume service displays in the output.

## Deploy and Bind a Sample App

This section describes how to deploy a sample app and bind it to the SMB volume service.

1. Clone the github repo and push the `pora` test app:

```
a. cd ~/workspace
b. git clone https://github.com/cloudfoundry/porsi-acceptance-tests.git
c. cd ~/workspace/porsi-acceptance-tests/assets/pora
d. cf push pora --no-start
```

2. To bind the service to your app, run the following command:

```
$ cf bind-service pora SERVICE-INSTANCE-NAME -c '{"username":"USERNAME","password":"PASSWORD"}'
```

Where:

- `SERVICE-INSTANCE-NAME`: The name of the volume service instance you created previously.
- `USERNAME` and `PASSWORD`: The username and password to use when mounting the share to the app. This allows you to interact with your SMB server as a specific user while allowing Cloud Foundry to run your app as an arbitrary user.
- **Optional parameters:**

- `mount`: Use this option to specify the path at which volumes mount to the app container. The default is an arbitrarily-named folder in `/var/vcap/data`. You may need to modify this value if your app has specific requirements. For example:

```
cf bind-service pora myVolume -c '{"username":"some-user","password":"some-password","mount":"/var/path"}'
```

- `readonly`: When you issue the `cf bind-service` command, Volume Services mounts a read-write file system by default. You can specify a read-only mount by adding `"readonly":true` to the bind configuration JSON string.
- `domain`: In case if your user is in windows domain you can specify `domain` parameter.

3. Start the app:

```
$ cf start pora
```

4. Use the following `curl` command to confirm the app is running. The command returns an instance index for your app.

```
$ curl http://pora.YOUR-CF-DOMAIN.com
```

5. Use the following `curl` command to confirm the app can access the shared volume. The command writes a file to the share and then reads it back out again.

```
$ curl http://pora.YOUR-CF-DOMAIN.com/write
```

## User-Provided Service Instances

This topic describes how to create and update user-provided service instances.

 **Note:** The procedures in this topic use the Cloud Foundry Command Line Interface (cf CLI). You can also create user-provided service instances in Apps Manager from the Marketplace. To update existing user-provided service instances, navigate to the service instance page and select the Configuration tab.

## Overview

User-provided service instances enable developers to use services that are not available in the marketplace with their apps running on Cloud Foundry.

User-provided service instances can be used to deliver service credentials to an app, and/or to trigger streaming of app logs to a syslog compatible consumer. These two functions can be used alone or at the same time.

Once created, user-provided service instances behave like service instances created through the marketplace; see [Managing Service Instances](#) and [App Binding](#) for details on listing, renaming, deleting, binding, and unbinding.

## Create a User-Provided Service Instance

The alias for [cf create-user-provided-service](#) is `cf cups`.

### Deliver Service Credentials to an App

Suppose a developer obtains a URL, port, username, and password for communicating with an Oracle database managed outside of Cloud Foundry. The developer could manually create custom environment variables to configure their app with these credentials (of course you would never hard code these credentials in your app!).

User-provided service instances enable developers to configure their apps with these using the familiar [App Binding](#) operation and the same app runtime environment variable used by Cloud Foundry to automatically deliver credentials for marketplace services ([VCAP\\_SERVICES](#)).

```
cf cups SERVICE_INSTANCE -p '{"username":"admin", "password":"pa55woRD"}'
```

To create a service instance in interactive mode, use the `-p` option with a comma-separated list of parameter names. The Cloud Foundry Command Line Interface (cf CLI) prompts you for each parameter value.

```
$ cf cups my-user-provided-route-service -p "host, port"
host> rdb.local
port> 5432
Creating user provided service my-user-provided-route-service in org my-org / space my-space as user@example.com...
OK
```

Once the user-provided service instance is created, to deliver the credentials to one or more apps see [App Binding](#).

### Stream App Logs to a Service

User-provided service instances enable developers to stream app logs to a syslog compatible aggregation or analytics service that isn't available in the marketplace. For more information about the syslog protocol see [RFC 5424](#) and [RFC 6587](#).

Create the user-provided service instance, specifying the URL of the service with the `-l` option.

```
cf cups SERVICE_INSTANCE -l syslog://example.log-aggregator.com
```

To stream app logs to the service, bind the user-provided service instance to your app.

## Proxy App Requests to a Route Service

User-provided service instances enable developers to proxy app requests to [route services](#) for preprocessing. To create a user-provided service instance for a route service, specify the url for the route service using the `-r` option.

```
$ cf create-user-provided-service my-user-provided-route-service -r https://my-route-service.example.com
Creating user provided service my-user-provided-route-service in org my-org / space my-space as user@example.com...
OK
```

 **Note:** When creating the user-provided service, the route service url specified must be https.

To proxy requests to the user-provided route service, you must bind the service instance to the route. For more information, see [Manage App Requests with Route Services](#).

## Update a User-provided Service Instance

You can use [cf update-user-provided-service](#) to update the attributes of an instance of a user-provided service. New credentials overwrite old credentials, and parameters not provided are deleted.

The alias for `update-user-provided-service` is `uups`.

## Streaming App Logs

The following list provides information about streaming app logs:

- [Streaming App Logs to Log Management Services](#)
- [Service-Specific Instructions for Streaming App Logs](#)
- [Streaming App Logs to Splunk](#)
- [Streaming App Logs with Fluentd](#)
- [Streaming App Logs to Azure OMS Log Analytics](#)

## Streaming App Logs to Log Management Services

This topic describes how to drain logs from Cloud Foundry to a third-party log management service.

Cloud Foundry aggregates logs for all instances of your apps as well as for requests made to your apps through internal components of Cloud Foundry. For example, when the Cloud Foundry Router forwards a request to an app, the Router records that event in the log stream for that app. Run the following command to access the log stream for an app in the terminal:

```
$ cf logs YOUR-APP-NAME
```

If you want to persist more than the limited amount of logging information that Cloud Foundry can buffer, drain these logs to a log management service.

For more information about the systems responsible for log aggregation and streaming in Cloud Foundry, see [App Logging in Cloud Foundry](#).

## Using Services from the Cloud Foundry Marketplace

Your Cloud Foundry marketplace may offer one or more log management services. To use one of these services, create an instance of the service and bind it to your app with the following commands:

```
$ cf create-service SERVICE PLAN SERVICE-INSTANCE
$ cf bind-service YOUR-APP YOUR-LOG-STORE
```

For more information about service instance lifecycle management, see the [Managing Service Instances](#) topic.

 **Note:** Not all marketplace services support syslog drains. Some services implement an integration with Cloud Foundry that enables automated streaming of app syslogs. If you are interested in building services for Cloud Foundry and making them available to end users, see the [Custom Services](#) documentation.

## Using Services Not Available in Your Marketplace

If a compatible log management service is not available in your Cloud Foundry marketplace, you can use [user-provided service instances](#) to stream app logs to a service of your choice.

You can install and use the [CF Drain CLI Plugin](#) to create and manage user-provided syslog drains from the CF command-line interface (cf CLI).

You may need to prepare your log management service to receive app logs from Cloud Foundry. For specific instructions for several popular services, see [Service-Specific Instructions for Streaming App Logs](#). If you cannot find instructions for your service, follow the generic instructions below.

### Step 1: Configure the Log Management Service

Complete the following steps to set up a communication channel between the log management service and your Cloud Foundry deployment:

1. Obtain the external IP addresses that your Cloud Foundry administrator assigns to outbound traffic.
2. Provide these IP addresses to the log management service. The specific steps to configure a third-party log management service depend on the service.
3. Whitelist these IP addresses to ensure unrestricted log routing to your log management service.
4. Record the syslog URL provided by the third-party service. Third-party services typically provide a syslog URL to use as an endpoint for incoming log data. You use this syslog URL in Step 2: Create a User-provided Service Instance.

Cloud Foundry uses the syslog URL to route messages to the service. The syslog URL has a scheme of `syslog`, `syslog-tls`, or `https`, and can include a port number. For example:

```
syslog://logs.example.com:1234
```

 **Note:** PAS does not support using `syslog-tls` or `https` with self-signed certificates. If you are running your own syslog server and want to use `syslog-tls` or `https`, you must have an SSL certificate signed by a well-known certificate authority.

## Step 2: Create and Bind a User-Provided Service Instance

You can create a syslog drain service and bind apps to it using either generic Cloud Foundry Command Line Interface (cf CLI) commands, or drain-specific commands enabled by the CF Drain plugin for the cf CLI.

Each option is described below.

### With the CF Drain CLI Plugin

1. If the CF Drain CLI Plugin is not installed on your local workstation, follow the [Installing Plugin](#) instructions in the plugin source repository.

2. Decide whether to bind the drain to a single app or all apps in a space, and run the corresponding command:

- o Single app:

```
cf drain APP-NAME SYSLOG-DRAIN-URL
```

Where:

- APP-NAME : is name of the app to stream logs from
- SYSLOG-DRAIN-URL : is the syslog URL from [Step 1: Configure the Log Management Service](#)

- o All apps in a space:

```
cf drain-space --drain-name DRAIN-NAME --drain-url SYSLOG-DRAIN-URL --username USERNAME
```

Where:

- DRAIN-NAME : is the name of the app to stream logs from.
- SYSLOG-DRAIN-URL : The syslog URL from [Step 1: Configure the Log Management Service](#).
- USERNAME : Username to use when pushing the app. If you do not specify a username, you must have admin permissions because the plugin will create a user.

After a short delay, logs begin to flow automatically.

Refer to the [Usage](#) section of the CF Drain plugin source repository for CF Drain commands, and [Managing Service Instances with the CLI](#) for general CF service commands.

### With General cf CLI Service Commands

 **Note:** To bind a drain to all apps in a space with a single command, you must use the CF Drain CLI Plugin as described in the previous section.

1. To create the service instance, run `cf create-user-provided-service` (or `cf cups`) with the `-l` flag, filling in values as follows:

- o DRAIN-NAME: A name to use for your syslog drain service instance
- o SYSLOG-DRAIN-URL: The syslog URL from [Step 1: Configure the Log Management Service](#).

```
$ cf create-user-provided-service DRAIN-NAME -l SYSLOG-URL
```

See [User-Provided Service Instances](#) for more information.

2. To bind an app to the service instance, do one of the following:

- o Run `cf push` with a manifest. The services block in the manifest must specify the service instance that you want to bind.
- o Run `cf bind-service`

```
$ cf bind-service YOUR-APP-NAME DRAIN-NAME
```

After a short delay, logs begin to flow automatically.

Refer to [Managing Service Instances with the CLI](#) for more information.

## Step 3: Verify Logs Are Draining

To verify that logs are draining correctly to a third-party log management service:

1. Take actions that produce log messages, such as making requests of your app.
2. Compare the logs displayed in the CLI against those displayed by the log management service.

For example, if your app serves web pages, you can send HTTP requests to the app. In Cloud Foundry, these generate Router log messages, which you can view in the CLI. Your third-party log management service should display corresponding messages.

 **Note:** For security reasons, Cloud Foundry apps do not respond to `ping`. You cannot use `ping` to generate log entries.

## CF Drain CLI Plugin

The CF Drain CLI plugin extends the cf CLI by adding simple commands for user-provided syslog drains. Also, you can use the plugin to bind all apps in a space to a syslog drain. This option includes app, space, and org names in the drain. It also binds any new apps pushed to the space.

**Installation:** To install the CF Drain CLI plugin see the [Installing Plugin](#) instructions in the plugin source repository.

**Commands:** The plugin adds commands for creating, deleting, and listing syslog drains, and for binding apps to the drains. See the [Usage](#) section of the plugin source repository for details.

## Service-Specific Instructions for Streaming App Logs

This topic provides instructions for configuring some third-party log management services.

Once you have configured a service, refer to the [Third-Party Log Management Services](#) topic for instructions on binding your app to the service.

### Logit.io

From your Logit.io dashboard:

1. Identify the Logit ELK stack you want to use.
2. Click Logstash **Configuration**.
3. Note your Logstash **Endpoint**.
4. Note your TCP-SSL, TCP, or UDP **Port** (not the syslog port).
5. Create the log drain service in Cloud Foundry.

```
$ cf cups logit-ssl-drain -l syslog-tls://ENDPOINT:PORT
```

or

```
$ cf cups logit-drain -l syslog://ENDPOINT:PORT
```

6. Bind the service to an app.

```
$ cf bind-service YOUR-CF-APP-NAME logit-ssl-drain
```

or

```
$ cf bind-service YOUR-CF-APP-NAME logit-drain
```

7. Restage or push the app using one of the following commands:

```
$ cf restage YOUR-CF-APP-NAME
```

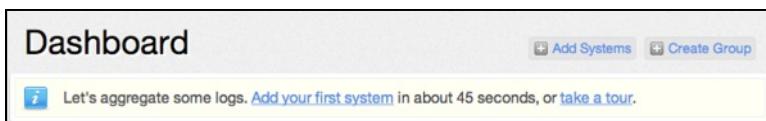
```
$ cf push YOUR-CF-APP-NAME
```

After a short delay, logs begin to appear in Kibana.

### Papertrail

From your Papertrail account:

1. Click **Add System**.



2. Click the **Other** link.

## Setup Systems

Your systems will log to **logs2.papertrailapp.com:14608**.

» Other situations: [Port 514](#) | [Other](#)

3. Select **I use Cloud Foundry**, enter a name, and click **Save**.

Choose your situation:

- A** My syslog only uses the default port  
GNU syslogd and some embedded devices will only log to port 514. A few old Linux distro versions use GNU syslogd (mostly CentOS and Gentoo).
- B** I use Cloud Foundry  
Register each app separately. Use Heroku? [Here's how.](#)
- C** My system's hostname changes  
In rare cases, one system may change hostnames frequently. For example, a roaming laptop which sets its hostname based on DHCP (and roams across networks).

We'll provide an app-specific syslog drain and step-by-step setup for [Cloud Foundry](#).

Let's create a destination for this app.

What should we call it?

CloudFoundry

Alphanumeric. Does not need to match app name.

**Save →**

4. Record the URL with port that is displayed after creating the system.

CloudFoundry will log to **logs.papertrailapp.com:36129**.

5. Create the log drain service in Cloud Foundry.

```
$ cf cups my-logs -l syslog-tls://logs.papertrailapp.com:PORT
```

6. Bind the service to an app.

```
$ cf bind-service APPLICATION-NAME my-logs
```

7. Restage the app.

```
$ cf restage APPLICATION-NAME
```

After a short delay, logs begin to flow automatically.

8. Once Papertrail starts receiving log entries, the view automatically updates to the logs viewing page.

All Systems

Dashboard Events Account Help Me

```

Skipping auto-reconfiguration.
Mar 05 14:57:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} Mar 05, 2014 10:57:36 PM
org.cloudfoundry.reconfiguration.AbstractServiceConfigure.configure
Mar 05 14:57:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} INFO: No beans of type org.springframework.data.mongodb.MongoDbFactory found
in application context. Skipping auto-reconfiguration.
Mar 05 14:57:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} Mar 05, 2014 10:57:36 PM
org.cloudfoundry.reconfiguration.AbstractServiceConfigure.configure
Mar 05 14:57:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} INFO: No beans of type
org.springframework.beans.factory.config.PropertyPlaceholderConfigurer found
in application context. Skipping auto-reconfiguration.
Mar 05 14:57:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} Mar 05, 2014 10:57:36 PM
org.cloudfoundry.reconfiguration.AbstractServiceConfigure.configure
Mar 05 14:57:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} INFO: class org.springframework.ampq.rabbit.connection.ConnectionFactory not
in classpath. Skipping auto-reconfiguration for it
Mar 05 14:57:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} 22:57:36,680 INFO RequestMappingHandlerMapping:181 - Mapped
{@albums({id}),methods={DELETE},params=[],headers=[],consumes=[],produces[],custom[]} onto public void
org.cloudfoundry.samples.music.web.controllers.AlbumController.deleteById(java.lang.String)
Mar 05 14:57:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} 22:57:36,681 INFO RequestMappingHandlerMapping:181 - Mapped
{@albums({id}),methods={PUT},params=[],headers[],consumes[],produces[],custom[]} onto public org.cloudfoundry.samples.music.domain.Album
org.cloudfoundry.samples.music.web.controllers.AlbumController.update(org.cloudfoundry.samples.music.domain.Album)
Mar 05 14:57:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} 22:57:36,681 INFO RequestMappingHandlerMapping:181 - Mapped
{@albums(),methods={GET},params[],headers[],consumes[],produces[],custom[]} onto public
java.lang.Iterable<org.cloudfoundry.samples.music.domain.Album> org.cloudfoundry.samples.music.web.controllers.AlbumController.albums()
Mar 05 14:57:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} 22:57:36,681 INFO RequestMappingHandlerMapping:181 - Mapped
{@albums(),methods={PUT},params[],headers[],consumes[],produces[],custom[]} onto public org.cloudfoundry.samples.music.domain.Album
org.cloudfoundry.samples.music.web.controllers.AlbumController.update(org.cloudfoundry.samples.music.domain.Album)
Mar 05 14:57:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} 22:57:36,682 INFO RequestMappingHandlerMapping:181 - Mapped
{@albums(),methods={POST},params[],headers[],consumes[],produces[],custom[]} onto public org.cloudfoundry.samples.music.domain.Album
org.cloudfoundry.samples.music.web.controllers.AlbumController.create(org.cloudfoundry.samples.music.domain.Album)
Mar 05 14:57:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} 22:57:36,684 INFO RequestMappingHandlerMapping:181 - Mapped
{@info(),methods=[],params[],headers[],consumes[],produces[],custom[]} onto public org.cloudfoundry.samples.music.domain.ApplicationInfo
org.cloudfoundry.samples.music.web.controllers.InfoController.info()
Mar 05 14:57:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} 22:57:36,684 INFO RequestMappingHandlerMapping:181 - Mapped
{@info(),methods=[],params[],headers[],consumes[],produces[],custom[]} onto public
java.util.List<org.springframework.web.ServiceInfo> org.cloudfoundry.samples.music.web.controllers.InfoController.showServiceInfo()
Mar 05 14:57:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} 22:57:36,684 INFO RequestMappingHandlerMapping:181 - Mapped
{@{},methods=[],params[],headers[],consumes[],produces[],custom[]} onto public java.util.Map<java.lang.String, java.lang.String>
org.cloudfoundry.samples.music.web.controllers.InfoController.showServiceInfo()
Mar 05 14:57:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} 22:57:36,704 INFO SimpleUrlHandlerMapping:382 - Root mapping to handler of
type [class org.springframework.web.servlet.mvc.ParameterizableViewController]
Mar 05 14:57:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} 22:57:36,726 INFO SimpleUrlHandlerMapping:315 - Mapped URL path [/assets/**]
onto handler of type [class org.springframework.web.servlet.resource.ResourceHttpRequestHandler]
Mar 05 14:57:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} 22:57:36,726 INFO SimpleUrlHandlerMapping:315 - Mapped URL path [/**] onto
handler of type [class org.springframework.web.servlet.resource.DefaultServletHttpRequestHandler]
Mar 05 14:57:37 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} 22:57:37,040 INFO DispatcherServlet:480 - FrameworkServlet 'appServlet':
initialization completed in 989 ms
Mar 05 14:57:37 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} Mar 05, 2014 10:57:37 PM org.apache.coyote.AbstractProtocol start
Mar 05 14:57:37 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} INFO: Starting ProtocolHandler ["http-bio-0-6399"]
Mar 05 14:57:37 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} Mar 05, 2014 10:57:37 PM org.apache.catalina.startup.Catalina start:
Mar 05 14:57:37 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{App[1]} INFO: Server startup in 11278 ms
Mar 05 14:59:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{API} Mar 05, 2014 10:59:37 PM org.apache.coyote.AbstractProtocol stop
Mar 05 14:59:36 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{API} INFO: Stopping ProtocolHandler ["http-bio-0-6399"]
Mar 05 14:59:37 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{API} Mar 05, 2014 10:59:37 PM org.apache.catalina.startup.Catalina stop:
Mar 05 14:59:37 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{API} Stopping the VirtualHost [Index 1] with grpid 4008c908-c093-4bbb-8718-186f147f3e73
Mar 05 14:59:37 CloudFoundry $odc908-c093-4bbb-8718-186f147f3e73/{DEA} Stopped app instance [Index 1] with grpid 4008c908-c093-4bbb-8718-186f147f3e73

```

Example: "access denied" (1.2.3.4 OR redis) -sshd    Search    ⌂    Contrast    ⏹    PAUSE

[https://papertrailapp.com/groups/55380/events/entered\\_on\\_id=3780034028458926114q-program%3Aaadc960-c093-4bbb-8718-186f147f3e73%2f%58App%2f1NSD](https://papertrailapp.com/groups/55380/events/entered_on_id=3780034028458926114q-program%3Aaadc960-c093-4bbb-8718-186f147f3e73%2f%58App%2f1NSD)

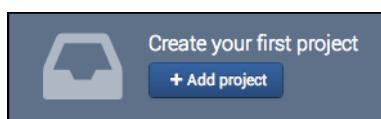
## Splunk

See [Streaming App Logs to Splunk](#)for details.

## Splunk Storm

From your Splunk Storm account:

1. Click **Add project**.



2. Enter the project details.

**Add project**

\* Project name

\* Project time zone

The Storm interface will use this time zone. If data you send to this project does not have a time zone already, it will be assigned this timezone by default [Learn More](#)

[Cancel](#) [Continue](#)

3. Create a new input for Network data.

**Network data**

Learn more [»](#)

- Sends directly from your servers
- Accepts syslog, syslog-ng, rsyslog, snare, netcat, etc.
- Works with Heroku drain

[Select](#)

4. Manually enter the external IP addresses your Cloud Foundry administrator assigns to outbound traffic.

**Add network data**

Storm can receive data from any network port. For example, Storm can accept remote data from [syslog](#), [rsyslog](#), [syslog-ng](#), [snare](#), [netcat](#) or any other application that transmits via TCP or UDP. [Learn more](#)

Authorize your IP address ▾

Automatically

Manually

5. Note the host and port provided for TCP input.

**Authorized network inputs**

Storm can receive data from any network port. For example, Storm can accept remote data from [syslog](#), [rsyslog](#), [syslog-ng](#), [snare](#), [netcat](#) or any other application that transmits via TCP or UDP.

1. Authorize your IP address

2. Send data to these ports for this project only

Network Data ▾

Authorize automatically

Authorize all IP addresses sending data to this project within the next 15 min.

Authorize manually

Specify the IP address

TCP tcp.k22g-wt6r.data.splunkstorm.com:15486

UDP udp.k22g-wt6r.data.splunkstorm.com:15486

6. Create the log drain service in Cloud Foundry using the displayed TCP host and port.

```
$ cf cups my-logs -l syslog://HOST:PORT
```

7. Bind the service to an app

```
$ cf bind-service APPLICATION-NAME my-logs
```

8. Restage the app

```
$ cf restage APPLICATION-NAME
```

After a short delay, logs begin to flow automatically.

9. Wait for some events to appear, then click **Data Summary**.

**What to Search**

266 Events INDEXED      a minute ago EARLIEST EVENT      a few seconds ago LATEST EVENT

Data Summary

10. Click the **loggregator** link to view all incoming log entries from Cloud Foundry.

**Data Summary**

Host	Count	Last Update
loggregator	26	3/7/14 3:26:01.000 PM

## SumoLogic

**Note:** SumoLogic uses HTTPS for communication. HTTPS is supported in Cloud Foundry v158 and later.

From your SumoLogic account:

1. Click the **Add Collector** link.

## Manage Collectors and Sources

[Upgrade Collectors](#) [Add Collector](#) [Access Keys](#)

- Choose Hosted Collector and fill in the details.

Add Collector

Select a type of collector:

 Installed Collector	 Hosted Collector
Select to install a Collector in your deployment.	Select to set up a Collector in the Sumo Logic Cloud.

FAQs

- ▶ What's the difference between an Installed and Hosted Collector?
- ▶ Where should I install an Installed Collector?
- ▶ How do I know if I need more than one Installed Collector?
- ▶ Where does my data go?

Add Collector

Name \*

Description

Category

Unless overwritten by Source metadata, the Collector will set the Source category of all messages to this value.

| [Cancel](#)

- In the new collector's row of the collectors view, click the Add Source link.

Manage Collectors and Sources					
<a href="#">Upgrade Collectors</a> <a href="#">Add Collector</a> <a href="#">Access Keys</a>					
Show: All Collectors   Running Collectors   Stopped Collectors   Expand: All   None					
Name	Type	Status	Source Category	Sources	Last Hour
Cloud Foundry	Hosted			1	None

[Add Source](#) | [Edit](#) | [Delete](#)

- Select HTTP source and fill in the details. Note that you'll be provided an HTTPS url

Select a type of Source:

 Amazon S3	 HTTP
Collects logs from an Amazon S3 bucket.	HTTP receiver that collects logs sent to a specific address.

Name \*   
Maximum name length is 128 characters

Description

Source Host   
Host name for the system from which the log files are being collected, e.g. LDAP\_Server

Source Category   
Log category metadata to use later for querying, e.g. OS\_Security

[Advanced](#)  
[Filters](#)

| [Cancel](#)

- Once the source is created, a URL should be displayed. You can also view the URL by clicking the Show URL link beside the created source.

Manage Collectors and Sources					
<a href="#">Upgrade Collectors</a> <a href="#">Add Collector</a> <a href="#">Access Keys</a>					
Show: All Collectors   Running Collectors   Stopped Collectors   Expand: All   None					
Name	Type	Status	Source Category	Sources	Last Hour
Cloud Foundry	Hosted			1	None
CloudFoundry	HTTP				

[Show URL](#) | [Edit](#) | [Delete](#)

6. Create the log drain service in Cloud Foundry using the displayed URL.

```
$ cf cups my-logs -l HTTPS-SOURCE-URL
```

7. Bind the service to an app.

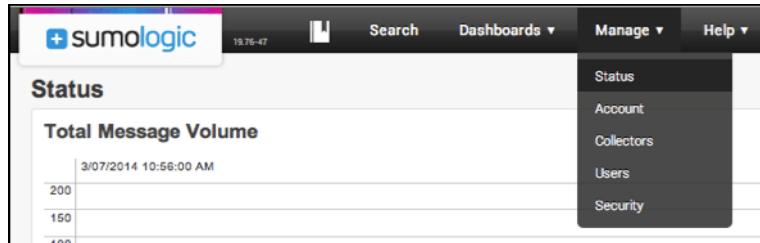
```
$ cf bind-service APPLICATION-NAME my-logs
```

8. Restage the app.

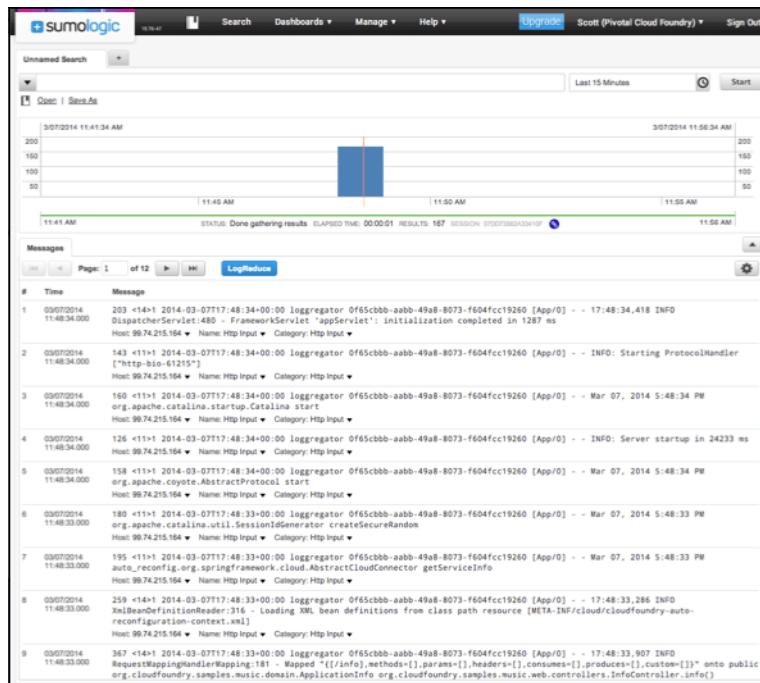
```
$ cf restage APPLICATION-NAME
```

After a short delay, logs begin to flow automatically.

9. In the SumoLogic dashboard, click **Manage**, then click **Status** to see a view of log messages received over time.



10. In the SumoLogic dashboard, click **Search**. Place the cursor in the search box, then press **Enter** to submit an empty search query.



## Logsene

**Note:** Logsene uses HTTPS for communication. HTTPS is supported in Cloud Foundry v158 and later.

From your Sematext account:

- Click the [Create App / Logsene App](#) menu item. Enter a name and click **Add Application** to create the Logsene App.
- Create the log drain service in Cloud Foundry using the displayed URL.

```
$ cf cups logsene-log-drain -l https://logsene-cf-receiver.sematest.com/YOUR_LOGSENE_TOKEN
```

3. Bind the log drain to an app. You could optionally bind multiple apps to one log drain.

```
$ cf bind-service YOUR-CF-APP-NAME logsene-log-drain
```

4. Restage the app.

```
$ cf restage APPLICATION-NAME
```

After a short delay, logs begin to flow automatically and appear in the [Logsene UI](#).

## Logentries is Not Supported

Cloud Foundry distributes log messages over multiple servers to handle load. Currently, we do not recommend using Logentries as it does not support multiple syslog sources.

## Streaming App Logs to Splunk

To integrate Cloud Foundry with Splunk Enterprise, complete the following process.

### 1. Create a Cloud Foundry Syslog Drain for Splunk

In Cloud Foundry, create a syslog drain user-provided service instance as described in [Using Third-Party Log Management Services](#).

Choose one or more apps whose logs you want to drain to Splunk through the service.

Bind each app to the service instance and restart the app.

Note the GUID for each app, the IP address of the Loggregator host, and the port number for the service. Locate the port number in the syslog URL. For example:

```
syslog://logs.example.com:1234
```

### 2. Prepare Splunk for Cloud Foundry

For detailed information about the following tasks, see the [Splunk documentation](#).

#### Install the RFC5424 Syslog Technology Add-On

The Cloud Foundry Loggregator component formats logs according to the Syslog Protocol defined in [RFC 5424](#). Splunk does not parse log fields according to this protocol. To allow Splunk to correctly parse RFC 5424 log fields, install the Splunk [RFC5424 Syslog Technical Add-On](#).

#### Patch the RFC5424 Syslog Technology Add-On

1. SSH into the Splunk VM
2. Replace `/opt/splunk/etc/apps/rfc5424/default/transforms.conf` with a new `transforms.conf` file that consists of the following text:

```
[rfc5424_host]
DEST_KEY = MetaData:Host
REGEX = <(d+)>d{1}s{1}S+s{1}(S+
FORMAT = host::$1

[rfc5424_header]
REGEX = <(d+)>d{1}s{1}S+s{1}S+s{1}(S+)s{1}(S+)s{1}(S+
FORMAT = prival::$1 appname::$2 procid::$3 msgid::$4
MV_ADD = true
```

3. Restart Splunk

#### Create a TCP Syslog Data Input

Create a TCP Syslog Data Input in Splunk, with the following settings:

- **TCP port** is the port number you assigned to your log drain service
- **Set sourcetype** is `Manual`
- **Source type** is `rfc5424_syslog` (type this value into text field)
- **Index** is the index you created for your log drain service

Your Cloud Foundry syslog drain service is now integrated with Splunk.

## 3. Verify that Integration was Successful

Use Splunk to execute a query of the form:

```
sourcetype=rfc5424_syslog index=--THE-INDEX-YOU-CREATED appname=APP-GUID
```

To view logs from all apps at once, you can omit the `appname` field.

Verify that results rows contain the three Cloud Foundry-specific fields:

- **appname:** The GUID for the Cloud Foundry app
- **host:** The IP address of the Loggregator host
- **procid:** The Cloud Foundry component emitting the log

If the Cloud Foundry-specific fields appear in the log search results, integration is successful.

If logs from an app are missing, make sure that the following are true:

- The app is bound to the service and was restarted after binding
- The service port number matches the TCP port number in Splunk

## Streaming App Logs with Fluentd

Fluentd [🔗](#) is an open source log collector that allows you to implement unified logging layers. With Fluentd, you can stream app logs to different backends or services like Elasticsearch, HDFS and Amazon S3. This topic explains how to integrate Fluentd with Cloud Foundry apps.

### Step 1: Create a Cloud Foundry Syslog Drain for Fluentd

1. In Cloud Foundry, create a syslog drain user-provided service instance as described in [Using Third-Party Log Management Services](#).
2. Choose one or more apps whose logs you want to drain to Fluentd through the service.
3. Bind each app to the service instance, and restart the app.
4. Note the GUID for each app, the IP address of the Loggregator host, and the port number for the service.
5. Locate the port number in the syslog URL. For example:

```
syslog://logs.example.com:5140
```

### Step 2: Set up Fluentd for Cloud Foundry

This section assumes you have an active Fluentd instance running. If you do not have an active Fluentd instance, refer to the [Fluentd Documentation/Install](#) [🔗](#) steps for more details.

Fluentd comes with native support for syslog protocol. To set up Fluentd for Cloud Foundry, configure the syslog input of Fluentd as follows.

1. In your main Fluentd configuration file, add the following `source` entry:

```
<source>
 @type syslog
 port 5140
 bind 0.0.0.0
 tag cf.app
 protocol_type udp
</source>
```

2. Restart the Fluentd service.

**Note:** The Fluentd syslog input plugin supports `udp` and `tcp` options. Make sure to use the same transport that Cloud Foundry is using.

Fluentd will start listening for Syslog message on port 5140 and tagging the messages with `cf.app`, which can be used later for data routing. For more details about the full setup for the service, refer to the [Config File](#) [🔗](#) article.

If your goal is to use an Elasticsearch or Amazon S3 backend, read the following guide:<http://www.fluentd.org/guides/recipes/elasticsearch-and-s3> [🔗](#)

## Streaming App Logs to Azure OMS Log Analytics (Beta)

**⚠ warning:** The OMS Log Analytics Firehose Nozzle is currently intended for evaluation and test purposes only. Do not use this product in a production environment.

This topic explains how to integrate your Cloud Foundry (CF) apps with [OMS Log Analytics](#).

Operations Management Suite (OMS) Log Analytics is a monitoring service for Microsoft Azure. The OMS Log Analytics Firehose Nozzle is a CF component that forwards metrics from the Loggregator Firehose to OMS Log Analytics.

This topic assumes you are using the latest version of the Cloud Foundry Command Line Interface (cf CLI) and a working Pivotal Application Service (PAS) deployment on Azure.

### Step 1: Create an OMS Workspace in Azure

See [Get started with Log Analytics](#) in the Microsoft Azure documentation to create an OMS workspace.

### Step 2: Deploy the Nozzle to Cloud Foundry

- Run `cf login -a https://api.YOUR-DOMAIN -u YOUR-USERNAME --skip-ssl-validation`, replacing `YOUR-DOMAIN` with your domain and `YOUR-USERNAME` with your CF username, to authenticate to your CF instance. For example:

```
$ cf login -a https://api.example.com -u admin --skip-ssl-validation
```

- Follow the steps below to create a new Cloud Foundry user and grant it access to the Loggregator Firehose using the UAA CLI (UAAC). For more information, see [Creating and Managing Users with the UAA CLI \(UAAC\)](#) and [Orgs, Spaces, Roles, and Permissions](#).

- Use `uaac target uaa.YOUR-DOMAIN` to target your UAA server:

```
$ uaac target uaa.example.com --skip-ssl-validation
```

- Run the following command to obtain an access token for the admin client:

```
$ uaac token client get admin
```

- Run `cf create-user USERNAME PASSWORD`, replacing `USERNAME` with a new username and `PASSWORD` with a password, to create a new user. For example:

```
$ cf create-user firehose-user firehose-password
```

- Run `uaac member add cloud_controller.admin USERNAME`, replacing `USERNAME` with the new username, to grant the new user admin permissions. For example:

```
$ uaac member add cloud_controller.admin firehose-user
```

- Run `uaac member add doppler.firehose USERNAME`, replacing `USERNAME` with the new username, to grant the new user permission to read logs from the Loggregator Firehose endpoint. For example:

```
$ uaac member add doppler.firehose firehose-user
```

- Download the [OMS Log Analytics Firehose Nozzle BOSH release](#) from Github. Clone the repository and navigate to the `oms-log-analytics-firehose-nozzle` directory:

```
$ git clone https://github.com/Azure/oms-log-analytics-firehose-nozzle.git
$ cd oms-log-analytics-firehose-nozzle
```

- Set the following environment variables in the [OMS Log Analytics Firehose manifest](#):

Environment Variable	Description

<pre> applications: - name: oms_nozzle ... env:   OMS_WORKSPACE: YOUR-WORKSPACE-ID   OMS_KEY: YOUR-OMS-KEY </pre>	Enter the ID and key value for your OMS workspace.
OMS_POST_TIMEOUT: 10s	(Optional) Set the HTTP post timeout for sending events to OMS Log Analytics. The default value is 10 seconds.
OMS_BATCH_TIME: 10s	(Optional) Set the interval for posting a batch to OMS. The default value is 10 seconds.  For more information, see the <a href="#">Configure Additional Logging</a> section below.
OMS_MAX_MSG_NUM_PER_BATCH: 1000	(Optional) Set the maximum number of messages to include in an OMS batch. The default amount is 1000.  For more information, see the <a href="#">Configure Additional Logging</a> section below.
FIREHOSE_USER: YOUR-FIREHOSE-USER FIREHOSE_USER_PASSWORD: YOUR-FIREHOSE-PASSWORD	Enter the username and password for the Firehose user you created in Step 2c.
API_ADDR: https://api.YOUR-DOMAIN	Enter the URL of your API endpoint.
DOPPLER_ADDR: wss://doppler.YOUR-DOMAIN:443	Enter the URL of your Loggregator traffic controller endpoint.
EVENT_FILTER: YOUR-LIST	(Optional) Enter the event types you want to filter out in a comma-separated list. The valid event types are <code>METRIC</code> , <code>LOG</code> , and <code>HTTP</code> .
IDLE_TIMEOUT: 60s	(Optional) Set the duration for the Firehose keepalive connection. The default time is 60 seconds.
SKIP_SSL_VALIDATION: TRUE-OR-FALSE	Set this value to <code>TRUE</code> to allow insecure connections to the UAA and the traffic controller. To block insecure connections to the UAA and traffic controller, set this value to <code>FALSE</code> .
LOG_LEVEL: INFO	(Optional) Change this value to increase or decrease the amount of logs. Valid log levels in increasing order include <code>INFO</code> , <code>ERROR</code> , and <code>DEBUG</code> . The default value is <code>INFO</code> .
LOG_EVENT_COUNT: TRUE-OR-FALSE	Set this value to <code>TRUE</code> to log the total count of events that the nozzle has sent, received, and lost. OMS logs this value as <code>CounterEvents</code> .  For more information, see the <a href="#">Configure Additional Logging</a> section below.
LOG_EVENT_COUNT_INTERVAL: 60s	(Optional) Set the time interval for logging the event count to OMS. The default interval is 60 seconds.  For more information, see the <a href="#">Configure Additional Logging</a> section below.

5. Push the app:

```
$ cf push
```

## Step 3: View Logs in OMS Portal

Import the Cloud Foundry OMS view to your OMS Portal to view visualized logs and metrics. You can also create alert rules for specific events.

 **Note:** The OMS view of Cloud Foundry is not yet available in the OMS Solutions Gallery. You can add it manually to view your logs in OMS Portal.

## Import the OMS View

1. From the main OMS Overview page, navigate to [View Designer](#).
2. Click **Import**.
3. Click **Browse**.
4. Select the **Cloud Foundry (Preview).omsview** file.
5. Save the view. The main OMS Overview page displays the **Tile**.
6. Click the **Tile** to view visualized metrics.

See the [OMS Log Analytics View Designer documentation](#) for more information.

## Create Alert Rules

See [Overview of alerts in Microsoft Azure](#) for more information about OMS Log Analytics alerts.

## Set Alert Queries

This section includes example queries that operators can set in the OMS Portal.

- The following query alerts the operator when the nozzle sends a `slowConsumerAlert` to OMS:

```
Type=CF_ValueMetric_CL Name_s=slowConsumerAlert
```

- The following query alerts the operator when Loggregator sends an `LGR` to indicate problems with the logging process:

```
Type=CF_LogMessage_CL SourceType_s=LGR MessageType_s=ERR
```

- The following query alerts the operator when the number of lost events reaches a certain threshold, specified in the OMS Portal:

```
Type=CF_CounterEvent_CL Job_s=nozzle Name_s=eventsLost
```

- The following query alerts the operator when the nozzle receives the `TruncatingBuffer.DroppedMessages` **CounterEvent**:

```
Type=CF_CounterEvent_CL Name_s="TruncatingBuffer.DroppedMessages"
```

## (Optional) Step 4: Configure Additional Logging

OMS Log Analytics Firehose Nozzle forwards metrics from the Loggregator Firehose to OMS with minimal processing, but the nozzle can push additional metrics to OMS.

### Log Sent, Received, and Lost Events

If you set the `LOG_EVENT_COUNT` environment variable to `TRUE` in the [manifest](#), the nozzle periodically sends the count of sent, received, and lost events to OMS. The value you set for the `LOG_EVENT_COUNT_INTERVAL` determines how frequently the nozzle sends the count.

 **Note:** The nozzle does not count **CounterEvents** themselves in the sent, received, or lost event count.

The nozzle sends the count as a **CounterEvent** with a **CounterKey** of one of the following:

CounterEvent	CounterKey
<code>nozzle.stats.eventsReceived</code>	The number of events the Firehose has received during the interval
<code>nozzle.stats.eventsSent</code>	The number of events the nozzle has successfully sent to OMS during the interval

`nozzle.stats.eventsLost` | The number of events the nozzle has tried to send to OMS during the interval, but failed to send after 4 attempts |  
In most cases, the total count of `eventsSent` plus `eventsLost` is less than the total `eventsReceived` at the same time. The nozzle buffers some messages and posts them in a batch to OMS. Operators can adjust the buffer size by adjusting the `OMS_BATCH_TIME` and `OMS_MAX_MSG_NUM_PER_BATCH` environment variables in the [manifest](#).

## Log Slow Consumer Alerts

 Note: The nozzle does not count **ValueMetrics** in the sent, received, or lost event count.

Loggregator sends the nozzle a `slowConsumerAlert` in the following situations:

- WebSocket sends the error code `ClosePolicyViolation (1008)`
- The nozzle receives a **CounterEvent** with the value `TruncatingBuffer.DroppedMessages`

In either case, the nozzle sends the `slowConsumerAlert` event to OMS as the following **ValueMetric**:

ValueMetric	MetricKey
<code>nozzle.alert.slowConsumerAlert</code>	1

See the *Slow Nozzle Alerts* section of the [Loggregator Guide for Cloud Foundry Operators](#) for more information.

## (Optional) Step 5: Scale the Deployment

### Scale the Nozzle

If the nozzle is unable to keep up with processing logs from the Firehose, Loggregator alerts the nozzle. When the nozzle receives the alert, it sends a `slowConsumerAlert` to OMS. If this happens, scaling up the nozzle minimizes data loss.

If an operator chooses to scale up their deployment, the Firehose evenly distributes events across all instances of the nozzle. See the *Scaling Nozzles* section of the [Loggregator Guide for Cloud Foundry Operators](#) for more information.

Operators can [create an alert rule](#) for the `slowConsumerAlert` message.

### Scale Loggregator

Loggregator sends `LGR` log messages to indicate problems with the logging process. See the *Scaling Loggregator* section of the [Loggregator Guide for Cloud Foundry Operators](#) for more information.

Operators can [create an alert rule](#) for the `LGR` message.

## Managing Apps with the cf CLI

The following list provides information about managing apps with the Cloud Foundry Command Line Interface (cf CLI):

- [Running Tasks](#)
- [Scaling an App Using cf scale](#)
- [Using Application Health Checks](#)

## Running Tasks

This topic describes how to run tasks in Cloud Foundry. A task is an app or script whose code is included as part of a deployed app, but runs independently in its own container.

## About Tasks

In contrast to a long running process (LRP), tasks run for a finite amount of time, then stop. Tasks run in their own containers and are designed to use minimal resources. After a task runs, Cloud Foundry destroys the container running the task.

As a single-use object, a task can be checked for its state and for a success or failure message.

 **Note:** Running a task consumes an application instance, and will be billed accordingly.

## Use Cases for Tasks

Tasks are used to perform one-off jobs, which include the following:

- Migrating a database
- Sending an email
- Running a batch job
- Running a data processing script
- Processing images
- Optimizing a search index
- Uploading data
- Backing-up data
- Downloading content

## How Tasks Are Run

Tasks are always executed asynchronously, meaning that they run independently from the parent app or other tasks that run on the same app.

The life-cycle of a task is as follows:

1. A user initiates a task in Cloud Foundry using one of the following mechanisms:
  - `cf run-task APPNAME "TASK"` command. See the [Running Tasks](#) section of this topic for more information.
  - Cloud Controller v3 API call. See the [Tasks](#) API reference page for more information.
  - Cloud Foundry Java Client. See the [Cloud Foundry Java Client Library](#) and [Cloud Foundry Java Client](#) topics for more information.
2. Cloud Foundry creates a container specifically for the task.
3. Cloud Foundry runs the task on the container using the value passed to the `cf run-task` command.
4. Cloud Foundry destroys the container.

The container also inherits environment variables, service bindings, and security groups bound to the app.

 **Note:** You cannot SSH into the container running a task.

## Task Logging and Execution History

Any data or messages the task outputs to STDOUT or STDERR is available on the app's firehose logs. A syslog drain attached to the app receives the task log output.

The task execution history is retained for one month.

## Manage Tasks

At the system level, a user with admin-level privileges can use the Cloud Controller v3 API to view all tasks that are running within an org or space. For more information, see the documentation for the [Cloud Controller v3 API](#).

In addition, admins can set the default memory and disk usage quotas for tasks on a global level. Initially, tasks use the same memory and disk usage defaults as apps. However, the default memory and disk allocations for tasks can be defined separately from the default app memory and disk allocations.

The default memory and disk allocations are defined using the **Default App Memory** and **Default Disk Quota per App** fields. These fields are available in the **Application Developer Controls** configuration screen of Pivotal Application Service (PAS).

## Run a Task on an App

You can use the Cloud Foundry Command Line Interface (cf CLI) to run a task in the context of an app.

**Note:** To run tasks with the cf CLI, you must install cf CLI v6.23.0 or later. See the [Installing the Cloud Foundry Command Line Interface](#) topic for information about downloading, installing, and uninstalling the cf CLI.

**Note:** To run a task without starting the app, push the app with `cf push -i 0` and then run the task. You can run the app later by scaling up its instance count.

To run a task on an app, perform the following steps:

1. Push your app:

```
$ cf push APP-NAME
```

2. Run your task on the deployed app:

```
$ cf run-task APP-NAME "TASK" --name TASK-NAME
```

The following example runs a database migration as a task on the `my-app` app:

```
$ cf run-task my-app "bin/rails db:migrate" --name my-task
Creating task for app my-app in org jdoe-org / space development as jdoe@pivotal.io...
OK
Task 1 has been submitted successfully for execution.
```

**Note:** To re-run a task, you must run it as a new task using the above command.

Use the `cf logs APP-NAME --recent` command to display the recent logs of the app and all its tasks.

The following example displays the logs of a successful task:

```
$ cf logs my-app --recent
2017-01-03T15:58:06.57-0800 [APP/TASK/my-task/0]OUT Creating container
2017-01-03T15:58:08.45-0800 [APP/TASK/my-task/0]OUT Successfully created container
2017-01-03T15:58:13.32-0800 [APP/TASK/my-task/0]OUT D, [2017-01-03T23:58:13.322258 #7] DEBUG -- : (15.9ms) CREATE TABLE "schema_migrations" ("version" character varying)
2017-01-03T15:58:13.33-0800 [APP/TASK/my-task/0]OUT D, [2017-01-03T23:58:13.337723 #7] DEBUG -- : (11.9ms) CREATE TABLE "ar_internal_metadata" ("key" character varying)
2017-01-03T15:58:13.34-0800 [APP/TASK/my-task/0]OUT D, [2017-01-03T23:58:13.340234 #7] DEBUG -- : (1.6ms) SELECT pg_try_advisory_lock(3720865444824511725);
2017-01-03T15:58:13.35-0800 [APP/TASK/my-task/0]OUT D, [2017-01-03T23:58:13.351853 #7] DEBUG -- : ActiveRecord::SchemaMigration Load (0.7ms) SELECT "schema_migrations"
2017-01-03T15:58:13.35-0800 [APP/TASK/my-task/0]OUT I, [2017-01-03T23:58:13.357294 #7] INFO -- : Migrating to CreateArticles (20161118225627)
2017-01-03T15:58:13.35-0800 [APP/TASK/my-task/0]OUT D, [2017-01-03T23:58:13.359565 #7] DEBUG -- : (0.5ms) BEGIN
2017-01-03T15:58:13.35-0800 [APP/TASK/my-task/0]OUT == 20161118225627 CreateArticles: migrating =====
2017-01-03T15:58:13.50-0800 [APP/TASK/my-task/0]OUT Exit status 0
2017-01-03T15:58:13.56-0800 [APP/TASK/my-task/0]OUT Destroying container
2017-01-03T15:58:15.65-0800 [APP/TASK/my-task/0]OUT Successfully destroyed container
```

The following example displays the logs of a failed task:

```
$ cf logs my-app --recent
2016-12-14T11:09:26.09-0800 [APP/TASK/my-task/0]OUT Creating container
2016-12-14T11:09:28.43-0800 [APP/TASK/my-task/0]OUT Successfully created container
2016-12-14T11:09:28.85-0800 [APP/TASK/my-task/0]ERR bash: bin/rails: command not found
2016-12-14T11:09:28.85-0800 [APP/TASK/my-task/0]OUT Exit status 127
2016-12-14T11:09:28.89-0800 [APP/TASK/my-task/0]OUT Destroying container
2016-12-14T11:09:30.50-0800 [APP/TASK/my-task/0]OUT Successfully destroyed container
```

If your task name is unique, you can `grep` the output of the `cf logs` command for the task name to view task-specific logs.

## List Tasks Running on an App

To list the tasks for a given app, run the `cf tasks APP-NAME`. For example:

```
$ cf tasks my-app
Getting tasks for app my-app in org jdoe-org / space development as jdoe@pivotal.io...
OK

id name state start time command
2 339044ef FAILED Wed, 23 Nov 2016 21:52:52 UTC echo foo; sleep 100; echo bar
1 8d0618cf SUCCEEDED Wed, 23 Nov 2016 21:37:28 UTC bin/rails db:migrate
```

Each task has one of the following states:

State	Description
RUNNING	The task is currently in progress.
FAILED	The task did not complete. This state occurs when a task does not work correctly or a user cancels the task.
SUCCEEDED	The task completed successfully.

## Cancel a Task

After running a task, you may be able to cancel it before it finishes. To cancel a running task, use the `cf terminate-task APP-NAME TASK-ID` command. For example:

```
$ cf terminate-task my-app 2
Terminating task 2 of app my-app in org jdoe-org / space development as jdoe@pivotal.io...
OK
```

## Scaling an App Using cf scale

Factors such as user load, or the number and nature of tasks performed by an app, can change the disk space and memory the app uses. For many apps, increasing the available disk space or memory can improve overall performance. Similarly, running additional instances of an app can allow the app to handle increases in user load and concurrent requests. These adjustments are called **scaling** an app.

Use [cf scale](#) to scale your app up or down to meet changes in traffic or demand.

 **Note:** You can configure your app to scale automatically based on rules that you set. See the [Scaling an Application Using Autoscaler](#) and [Using the App Autoscaler CLI](#) topics for more information.

### Scaling Horizontally

Horizontally scaling an app creates or destroys instances of your app.

Incoming requests to your app are automatically load balanced across all instances of your app, and each instance handles tasks in parallel with every other instance. Adding more instances allows your app to handle increased traffic and demand.

Use `cf scale APP -i INSTANCES` to horizontally scale your app. Cloud Foundry will increase or decrease the number of instances of your app to match `INSTANCES`.

```
$ cf scale myApp -i 5
```

### Scaling Vertically

Vertically scaling an app changes the disk space limit or memory limit that Cloud Foundry applies to all instances of the app.

Use `cf scale APP -k DISK` to change the disk space limit applied to all instances of your app. `DISK` must be an integer followed by either an **M**, for megabytes, or **G**, for gigabytes.

```
$ cf scale myApp -k 512M
```

Use `cf scale APP -m MEMORY` to change the memory limit applied to all instances of your app. `MEMORY` must be an integer followed by either an **M**, for megabytes, or **G**, for gigabytes.

```
$ cf scale myApp -m 1G
```

## Using Application Health Checks

This topic describes how to configure health checks for your apps in Cloud Foundry.

### Overview

An application health check is a monitoring process that continually checks the status of a running Cloud Foundry app.

Developers can configure a health check for an app using the Cloud Foundry Command Line Interface (cf CLI) or by specifying the `health-check-http-endpoint` and `health-check-type` fields in an app manifest.

To configure a health check using the cf CLI, follow the instructions in the [Configure Health Checks](#) section below. For more information about using an app manifest to configure a health check, see the `health-check-http-endpoint` and `health-check-type` sections of the [Deploying with App Manifest](#) topic.

Application health checks function as part of the app lifecycle managed by [Diego architecture](#).

### Configure Health Checks

To configure a health check while creating or updating an app, use the [cf push](#) command:

```
$ cf push YOUR-APP -u HEALTH-CHECK-TYPE -t HEALTH-CHECK-TIMEOUT
```

Replace the placeholders in the example command above as follows:

- `HEALTH-CHECK-TYPE`: Valid health check types are `port`, `process`, and `http`. See the [Health Check Types](#) section below for more information.
- `HEALTH-CHECK-TIMEOUT`: The timeout is the amount of time allowed to elapse between starting up an app and the first healthy response. See the [Health Check Timeouts](#) section for more information.

 **Note:** The health check configuration you provide with `cf push` overrides any configuration in the app manifest.

To configure a health check for an existing app or to add a custom HTTP endpoint, use the [cf set-health-check](#) command:

```
$ cf set-health-check YOUR-APP HEALTH-CHECK-TYPE --endpoint CUSTOM-HTTP-ENDPOINT
```

Replace the placeholders in the example command above as follows:

- `HEALTH-CHECK-TYPE`: Valid health check types are `port`, `process`, and `http`. See the [Health Check Types](#) section below for more information.
- `CUSTOM-HTTP-ENDPOINT`: A `http` health check defaults to using `/` as its endpoint, but you can specify a custom endpoint. See the [Health Check HTTP Endpoints](#) section below for more information.

 **Note:** You can change the health check configuration of a deployed app with `cf set-health-check`, but you must restart the app for the changes to take effect.

 **Note:** You can also use the `v3-set-health-check` CLI command to change the individual health check invocation timeout for an app. This option also requires restarting the app. See the [Cloud Foundry CLI Reference Guide](#) for more information.

### Understand Health Checks

#### Health Check Lifecycle

The following table describes how application health checks work in Cloud Foundry.

Stage	Description
1	App developer deploys an app to Cloud Foundry.

2	When deploying the app, the developer specifies a health check type for the app and, optionally, a timeout. If the developer does not specify a health check type, then the monitoring process defaults to a <code>port</code> health check.
3	Cloud Controller stages, starts, and runs the app.
4	Based on the type specified for the app, Cloud Controller configures a health check that runs periodically for each app instance.
5	When Diego starts an app instance, the application health check runs every 2 seconds until a response indicates that the app instance is healthy or until the health check timeout elapses. The 2-second health check interval is not configurable.
6	When an app instance becomes healthy, its route is advertised, if applicable. Subsequent health checks are run every 30 seconds once the app becomes healthy. The 30-second health check interval is not configurable.
7	If a previously healthy app instance fails a health check, Diego considers that particular instance to be unhealthy. As a result, Diego stops and deletes the app instance, then reschedules a new app instance. This stoppage and deletion of the app instance is reported back to the Cloud Controller as a crash event.
8	When an app instance crashes, Diego immediately attempts to restart the app instance several times. After three failed restarts, Cloud Foundry waits 30 seconds before attempting another restart. The wait time doubles each restart until the ninth restart, and remains at that duration until the 200th restart. After the 200th restart, Cloud Foundry stops trying to restart the app instance.

## Health Check Types

The following table describes the types of health checks available for apps and recommended circumstances in which to use them:

Health Check Type	Recommended Use Case	Explanation
<code>http</code>	The app can provide an <code>HTTP 200</code> response.	The <code>http</code> health check performs a GET request to the configured HTTP endpoint on the app's default port. When the health check receives an <code>HTTP 200</code> response, the app is declared healthy. We recommend using the <code>http</code> health check type whenever possible. A healthy HTTP response ensures that the web app is ready to serve HTTP requests. The configured endpoint must respond within 1 second to be considered healthy. <b>WARNING:</b> To prevent false negatives, use a dedicated endpoint for healthcheck where response time and result do not depend on business logic.
<code>port</code>	The app can receive TCP connections (including HTTP web apps).	A health check makes a TCP connection to the port or ports configured for the app. For apps with multiple ports, a health check monitors each port. If you do not specify a health check type for your app, then the monitoring process defaults to a <code>port</code> health check. The TCP connection must be established within 1 second to be considered healthy.
<code>process</code>	The app does not support TCP connections (for example, a worker).	For a <code>process</code> health check, Diego ensures that any process declared for the app stays running. If the process exits, Diego stops and deletes the app instance.

## Health Check Timeouts

The value configured for the health check timeout is the amount of time allowed to elapse between starting up an app and the first healthy response from the app. If the health check does not receive a healthy response within the configured timeout, then the app is declared unhealthy.

In Pivotal Cloud Foundry, the default timeout is 60 seconds and the maximum configurable timeout is 600 seconds. You can modify the timeout in the **Application Containers** pane of the Pivotal Application Service (PAS) tile.

## Health Check HTTP Endpoints

Only used by `http` type, the `--endpoint` flag of the `cf set-health-check` command specifies the path portion of a URI that must be served by the app and return `HTTP 200` when the app is healthy.

 **Note:** This command will only check the health of the default port of the app.

 **Note:** For HTTP apps, we recommend setting the health check type to `http` instead of a simple port check.



## Configuring Container-to-Container Networking

This topic describes how to configure the Container-to-Container Networking feature, which allows direct network traffic between apps. For an overview of how Container-to-Container Networking works, see the [Container-to-Container Networking](#) topic.

Container-to-Container Networking enables PCF to generate logs whenever containers communicate or attempt to communicate with each other. See [App Traffic Logging](#) for how to manage app traffic logging.

### Configure the Overlay Network

Container-to-Container Networking uses an overlay network to manage communication between app instances. By default, each Diego cell in the overlay network is allocated a /24 range that supports 254 containers per cell, one container for each of the usable IP addresses, .1 through .254. For more information about the overlay network, see [Overlay Network](#) in [Container-to-Container Networking](#).

### Configure the Number of Diego Cells

If you want to modify the number of Diego cells supported by the overlay network, follow the steps below:

1. In Ops Manager, select the PAS tile.
2. Select **Networking**.
3. Under **Overlay Subnet**, enter an IP range for the overlay network. By default, Ops Manager uses `10.255.0.0/16`. Modifying the subnet range allocated to the overlay network changes the number of Diego cells supported in your deployment. Use the table below as a reference.

Overlay subnet mask	Number of cells	Containers per cell
/20	15	254
/16	255	254
/12	4,095	254

**A** **warning:** The overlay network IP address range must not conflict with any other IP addresses in the network. If a conflict exists, Diego cells cannot reach any endpoint that has a conflicting IP address.

### Create and Manage Networking Policies

This section describes how to create and modify Container-to-Container Networking policies using the Cloud Foundry Command Line Interface (cf CLI).

**Note:** You can also create and modify container-to-container networking policies using Apps Manager. For more information, see [Create Container-to-Container Networking Policies](#).

**Note:** With the NSX-T integration, container networking policies and ASGs continue to work as normal. Advanced ASG logging is not supported with NSX-T.

### Prerequisites

- Ensure that you are using cf CLI v6.42 or later:

```
$ cf version
```

For more information about updating the cf CLI, see the [Installing the cf CLI](#) topic.

### Grant Permissions

CF admins use the following UAA scopes to grant specific users or groups permissions to configure network policies:

UAA Scope	Suitable for...	Allows users to create policies...
<code>network.admin</code>	operators	for any apps in the CF deployment
<code>network.write</code>	space developers	for apps in spaces that they can access

If you are a CF admin, you already have the `network.admin` scope. An admin can also grant the `network.admin` scope to a space developer.

For more information, see [Creating and Managing Users with the UAA CLI \(UAAC\)](#) and [Orgs, Spaces, Roles, and Permissions](#).

To grant all Space Developers permissions to configure network policies, open the **Application Developer Controls** pane in your PAS tile and enable the **Allow Space Developers to manage network policies** checkbox.

## Add a Network Policy

To add a policy that allows direct network traffic from one app to another, run the following command:

```
cf add-network-policy SOURCE_APP --destination-app DESTINATION_APP -s DESTINATION_SPACE_NAME -o DESTINATION_ORG_NAME --protocol (tcp | udp) --port RANGE
```

Replace the placeholders in the above command as follows:

- `SOURCE_APP` is the name of the app that sends traffic.
- `DESTINATION_APP` is the name of the app that will receive traffic.
- `DESTINATION_SPACE_NAME` is the space of the destination app. The default is the targeted space.
- `DESTINATION_ORG_NAME` is the org of the destination app. The default is the targeted org.
- `PROTOCOL` is one of the following: `tcp` or `udp`.
- `RANGE` are the ports at which to connect to the destination app. The allowed range is from `1` to `65535`. You can specify a single port, such as `8080`, or a range of ports, such as `8080-8090`.

The following example command allows access from the `frontend` app to the `backend` app over TCP at port 8080:

```
$ cf add-network-policy frontend --destination-app backend --protocol tcp --port 8080
Adding network policy to app frontend in org my-org / space dev as admin...
OK
```

## List Policies

You can list all the policies in your space, or just the policies for which a single app is the source:

- To list the all the policies in your space, run `cf network-policies`.

```
$ cf network-policies
```

- To list the policies for an app, run `cf network-policies --source MY-APP`. Replace `MY-APP` with the name of your app.

```
$ cf network-policies --source example-app
```

The following example command lists policies for the app `frontend`:

```
$ cf network-policies --source frontend
Listing network policies in org my-org / space dev as admin...
source destination protocol ports destination space destination org
frontend backend tcp 8080 example-space example-org
```

## Remove a Network Policy

To remove a policy that allows direct network traffic from an app, run the following command:

```
cf remove-network-policy SOURCE_APP --destination-app DESTINATION_APP -s DESTINATION_SPACE_NAME -o DESTINATION_ORG_NAME --protocol PROTOCOL --port RANGE
```

Replace the placeholders in the above command to match an existing policy, as follows:

- `SOURCE_APP` is the name of the app that sends traffic.
- `DESTINATION_APP` is the name of the app that receives traffic.
- `DESTINATION_SPACE_NAME` is the space of the destination app. The default is the targeted space.
- `DESTINATION_ORG_NAME` is the org of the destination app. The default is the targeted org.
- `PROTOCOL` is either `tcp` or `udp`.
- `PORTS` are the ports connecting the apps. The allowed range is from `1` to `65535`. You can specify a single port, such as `8080`, or a range of ports, such as `8080-8090`.

The following command deletes the policy that allowed the `frontend` app to communicate with the `backend` app over TCP on port 8080:

```
$ cf remove-network-policy frontend --destination-app backend --protocol tcp --port 8080
Removing network policy to app frontend in org my-org / space dev as admin...
OK
```

## Disable Network Policy Enforcement

You can disable Silk network policy enforcement between apps. Disabling network policy enforcement allows all apps to send network traffic to all other apps in the foundation despite no policy specifically allowing it.

To disable network policy enforcement between apps, do the following:

1. Go to the **Networking** pane of your PAS tile.
2. Disable the **Enable Silk Network Policy Enforcement** checkbox.
3. Click **Save**.

## App Service Discovery

With app service discovery, apps pushed to Pivotal Application Service (PAS) can establish container-to-container communications through a known route served by internal BOSH DNS. This allows front end apps to easily connect with back end apps.

**Note:** The internal domain used for service discovery is `apps.internal` by default. Operators can modify this value in **Application Developer Controls** pane of the PAS tile.

To establish container-to-container communications between a front end and back end app, a developer:

1. Launches a back end app that publishes a local endpoint.
2. Maps a named route to the endpoint.
3. Creates a network policy that allows direct traffic from the front end to the back end app.
4. Launches the front end app.

See [Cats and Dogs with Service Discovery](#) in GitHub for an example, written in Go, that demonstrates communication between front end and back end apps.

## Administering Dynamic Egress Policies (Beta)

 Note: This is a beta feature.

This topic describes how to administer dynamic egress policies for the apps in your deployment.

### Overview

You can create dynamic egress policies to allow your apps to communicate with external services, such as a MySQL database. The workflow is as follows:

- Create a destination object with details about the external service that your app or space need access to, including an IP range.
- Create an egress policy from the app or space to the destination object.

Dynamic egress policies provide the following benefits over [Application Security Groups \(ASGs\)](#):

- You do not have to restart your apps when applying these policies, so there is no downtime.
- The policies include an additional level of granularity: you can apply them to specific apps.

### Enable the Dynamic Egress Feature

The dynamic egress feature is disabled by default. To enable this feature, you must select **Enable Beta Dynamic Egress Enforcement** in the PAS **Networking** pane. Additionally, you must have **Silk** selected for your **Container Network Interface Plugin**.

### API Authorization

To administer dynamic egress policies, you must have the `network.admin` UAA scope. If you are a CF admin, you already have the `network.admin` scope. An admin can also grant the `network.admin` scope to a space developer. For more information, see [Creating and Managing Users with the UAA CLI \(UAAC\)](#) and [Orgs, Spaces, Roles, and Permissions](#).

If you have the `network.admin` scope, you can interact with the API in the following ways:

#### Option 1: cf curl

When you use [cf curl](#), you do not have to provide a token or the full API URL because you are already logged in and targeting your deployment. See the following example:

```
$ cf curl /networking/v1/external/egress_policies
{"total_egress_policies":2,"egress_policies":[{"source":{...}}]}
```

#### Option 2: curl

When using `curl`, you must provide the token `Authorization` header and use the full API URL. See the following example:

```
$ export TOKEN='cf oauth-token' # as CF admin
$ curl http://api.bosh-lite.com/networking/v1/external/egress_policies -H "Authorization: $TOKEN"
{"total_egress_policies":2,"egress_policies":[{"source":{...}}]}
```

## Typical Workflows

Dynamic egress policies depend on a destination object. See the following workflow examples:

## Add an Egress Policy

To create a policy, you must create a destination first:

1. Create an egress destination. See [Create Egress Destinations](#).
2. Create an egress policy linking the destination and an app or space. [Create Egress Policies](#).
3. See policy apply. No app restarts needed.

## Update an Egress Policy when an IP Changes

To update a policy, you update the destination:

1. List all egress destinations to find the one you want to change. See [List Egress Destinations](#).
2. Update the egress destination. See [Update Egress Destinations](#).
3. See updated policy apply. No app restarts needed.

## Egress Destination API

Method	Path	Description
GET	/networking/v1/external/destinations	List Destinations
POST	/networking/v1/external/destinations	Create Destinations
PUT	/networking/v1/external/destinations	Update Destinations
DELETE	/networking/v1/external/destinations/GUID	Delete Destinations

### List Egress Destinations

GET /networking/v1/external/destinations

Arguments:

- (optional) `id`: comma-separated id values. This cannot be used with `name`.
- (optional) `name`: comma-separated name values. This cannot be used with `id`.

Response Body:

This endpoint returns all egress destinations

```
{
 "total_destinations": 2,
 "destinations": [
 {
 "name": "oracle database",
 "id": "90be9c1fb694-4463-9ff1-6ce71904440d",
 "description": "db for user accounts",
 "ips": [{"start": "1.9.9.9", "end": "1.9.9.20"}],
 "ports": [{"start": 8000, "end": 9000}],
 "protocol": "tcp"
 },
 {
 "name": "AWS",
 "id": "72813418-bd38-49e0-ace0-7bf5b7c54687",
 "ips": [{"start": "1.8.8.8", "end": "1.8.8.8"}],
 "ports": [{"start": 8000, "end": 9000}],
 "protocol": "udp"
 }
]
}
```

## Create Egress Destinations

POST /networking/v1/external/destinations

Request Body:

```
{
 "destinations": [
 {
 "name": "oracle database",
 "description": "db for user accounts",
 "ips": [{"start": "1.9.9.9", "end": "1.9.9.20"}],
 "ports": [{"start": 8000, "end": 9000}],
 "protocol": "tcp"
 },
 {
 "name": "AWS",
 "ips": [{"start": "1.8.8.8", "end": "1.8.8.8"}],
 "ports": [{"start": 8000, "end": 9000}],
 "protocol": "udp"
 }
]
}
```

## Parameters

Field	Required?	Description
destinations.name	Y	The name of the destination. Must be globally unique.
destinations.description	N	A description of the destination.
destinations.ips.start*	Y	The start of the destination ip range. Must be IPv4.
destinations.ips.end*	Y	The end of the destination ip range. Must be IPv4. May be equal to the start ip.
destinations.ports.start*	Y	The destination start port (1 - 65535). Ports are not applicable for ICMP protocol.
destinations.ports.end*	Y	The destination end port (1 - 65535). Ports are not applicable for ICMP protocol.
destinations.protocol	Y	The protocol (tcp, udp, or icmp)
destinations.icmp_type	N	The icmp type to allow when using the icmp protocol. Default is all icmp types, represented by -1.
destinations.icmp_code	N	The icmp code to allow when using the icmp protocol. Default is all icmp codes, represented by -1.

\*Currently only one IP range and one port range is supported.

## Update Egress Destinations

PUT /networking/v1/external/destinations

Request Body:

```
{
 "destinations": [
 {
 "id": "90be9c1f-b694-4463-9f1f-6ce71904440d",
 "name": "oracle database",
 "description": "db for user accounts",
 "ips": [{"start": "1.9.9.9", "end": "1.9.9.20"}],
 "ports": [{"start": 8000, "end": 9000}],
 "protocol": "tcp"
 },
 {
 "id": "72813418-bd38-49e0-ace0-7bf5b7c54687",
 "name": "AWS",
 "description": null,
 "ips": [{"start": "1.8.8.8", "end": "1.8.8.8"}],
 "ports": [{"start": 8000, "end": 9000}],
 "protocol": "udp"
 }
]
}
```

Parameters

Field	Required?	Description
destinations.id	Y	The id of the destination. This id is returned in the destinations create response, as well as in the destinations index response.
destinations.name	Y	The name of the destination. Must be globally unique.
destinations.description	N	A description of the destination.
destinations.ips.start*	Y	The start of the destination ip range. Must be IPv4.
destinations.ips.end*	Y	The end of the destination ip range. Must be IPv4. May be equal to the the start ip.
destinations.ports.start*	Y	The destination start port (1 - 65535). Ports are not applicable for ICMP protocol.
destinations.ports.end*	Y	The destination end port (1 - 65535). Ports are not applicable for ICMP protocol.
destinations.protocol	Y	The protocol (tcp, udp, or icmp)
destinations.icmp_type	N	The icmp type to allow when using the icmp protocol. Default is all icmp types, represented by -1.
destinations.icmp_code	N	The icmp code to allow when using the icmp protocol. Default is all icmp codes, represented by -1.

\*Note: Currently only one IP range and one port range is supported.

## Delete an Egress Destination

DELETE /networking/v1/external/destinations/GUID

Response Body:

This endpoint returns the JSON of the deleted destination object.

```
{
 "total_destinations": 1,
 "destinations": [
 {
 "name": "oracle database",
 "id": "90be9c1fb694-4463-9ff1-6ce71904440d",
 "description": "db for user accounts",
 "ips": [{"start": "1.9.9.9", "end": "1.9.9.20"}],
 "ports": [{"start": 8000, "end": 9000}],
 "protocol": "tcp"
 }
]
}
```

## Egress Policy API

Method	Path	Description
GET	/networking/v1/external/egress_policies	List Egress Policies
POST	/networking/v1/external/egress_policies	Create EgressPolicies
DELETE	/networking/v1/external/egress_policies/GUID	Delete Egress Policy

### List Egress Policies

GET /networking/v1/external/egress\_policies

Arguments:

None

Response Body:

This endpoint returns all egress policies.

```
{
 "total_egress_policies": 1,
 "egress_policies": [
 {
 "id": "dynamic-egress-guid",
 "source": {
 "type": "app",
 "id": "SOURCE-APP-GUID"
 },
 "destination": {
 "id": "guid-abc-123",
 "name": "AWS",
 "description": "AWS",
 "id": "72813418-bd38-49e0-ace0-7bf5b7c54687",
 "ips": [{"start": "1.8.8.8", "end": "1.8.8.8"}],
 "ports": [{"start": 8000, "end": 9000}],
 "protocol": "udp"
 }
 }
]
}
```

### Create Egress Policies

POST /networking/v1/external/egress\_policies

## Request Body:

```
{
 "egress_policies": [
 {
 "source": {
 "type": "space",
 "id": "SOURCE-SPACE-GUID"
 },
 "destination": {
 "id": "EGRESS-DESTINATION-GUID"
 }
 }
]
}
```

## Parameters

Field	Required?	Description
source.type	N	The type of source. Must be 'app' or 'space'. Defaults to 'app'.
source.id	Y	The guid of the source app or space.
destination.id	Y	The guid of the egress destination.

## Delete an Egress Policy

DELETE /networking/v1/external/egress\_policies/GUID

## Response Body:

This endpoint returns the JSON of the deleted egress policy object.

```
{
 "total_egress_policies": 1,
 "egress_policies": [
 {
 "id": "dynamic-egress-guid",
 "source": {
 "type": "app",
 "id": "SOURCE-APP-GUID"
 },
 "destination": {
 "name": "AWS",
 "description": "AWS",
 "id": "72813418-bd38-49e0-ace0-7bf5b7c54687",
 "ips": [{"start": "1.8.8.8", "end": "1.8.8.8"}],
 "ports": [{"start": 8000, "end": 9000}],
 "protocol": "udp"
 }
 }
]
}
```

## Cloud Foundry Environment Variables

Environment variables are the means by which the Cloud Foundry (CF) runtime communicates with a deployed app about its environment. This page describes the environment variables that the runtime and buildpacks set for app.

For information about setting your own app-specific environment variables, see the [Environment Variable](#) section of the *Deploying with App Manifests* topic.

## View Environment Variables

Install the Cloud Foundry Command Line Interface (cf CLI), and use the `cf env` command to view the Cloud Foundry environment variables for your app. The `cf env` command displays the following environment variables:

- The `VCAP_APPLICATION` and `VCAP_SERVICES` variables provided in the container environment
- The user-provided variables set using the `cf set-env` command

```
$ cf env my-app
Getting env variables for app my-app in org my-org / space my-space as
admin...
OK
System-Provided:

{
 "VCAP_APPLICATION": {
 "application_id": "fa05c1a9-0fc1-4fb1-bae1-139850dec7a3",
 "application_name": "my-app",
 "application_uris": [
 "my-app.192.0.2.34.xip.io"
],
 "application_version": "fb8fbcc6-8d58-479e-bcc7-3b4ce5a7f0ca",
 "cf_api": "https://api.example.com",
 "limits": {
 "disk": 1024,
 "fds": 16384,
 "mem": 256
 },
 "name": "my-app",
 "space_id": "06450c72-4669-4dc6-8096-45f9777db68a",
 "space_name": "my-space",
 "uris": [
 "my-app.192.0.2.34.xip.io"
],
 "users": null,
 "version": "fb8fbcc6-8d58-479e-bcc7-3b4ce5a7f0ca"
 }
}

User-Provided:
MY_DRAIN: http://drain.example.com
MY_ENV_VARIABLE: 100
```

## App-Specific System Variables

The subsections that follow describe the environment variables that Cloud Foundry makes available to your app container. Some of these variables are the same across instances of a single app, and some vary from instance to instance.

You can access environment variables programmatically, including variables defined by the buildpack. For more information, refer to the buildpack documentation for [Java](#), [Node.js](#), and [Ruby](#).

Env Var	Running	Staging	Task
<code>CF_INSTANCE_ADDR</code>	x	x	x
<code>CF_INSTANCE_GUID</code>	x		x
<code>CF_INSTANCE_INDEX</code>	x		
<code>CF_INSTANCE_INTERNAL_IP</code>	x	x	x
<code>CF_INSTANCE_IP</code>	x	x	x

CF_INSTANCE_PORT	X	X	X
CF_INSTANCE_PORTS	X	X	X
CF_STACK		X	
DATABASE_URL	X		X
HOME	X	X	X
INSTANCE_GUID	X		
INSTANCE_INDEX	X		
LANG	X	X	X
MEMORY_LIMIT	X	X	X
PATH	X	X	X
PORT	X		
PWD	X	X	X
TMPDIR	X		X
USER	X	X	X
VCAP_APP_HOST	X		
VCAP_APP_PORT	X		
VCAP_APPLICATION	X	X	X
VCAP_SERVICES	X	X	X

## CF\_INSTANCE\_ADDR

The [CF\\_INSTANCE\\_IP](#) and [CF\\_INSTANCE\\_PORT](#) of the app instance in the format `IP:PORT`.

Example: `CF_INSTANCE_ADDR=1.2.3.4:5678`

## CF\_INSTANCE\_GUID

The UUID of the particular instance of the app.

Example: `CF_INSTANCE_GUID=41653aa4-3a3a-486a-4431-ef258b39f042`

## CF\_INSTANCE\_INDEX

The index number of the app instance.

Example: `CF_INSTANCE_INDEX=0`

## CF\_INSTANCE\_IP

The external IP address of the host running the app instance.

Example: `CF_INSTANCE_IP=1.2.3.4`

## CF\_INSTANCE\_INTERNAL\_IP

The internal IP address of the container running the app instance.

Example: `CF_INSTANCE_INTERNAL_IP=5.6.7.8`

## CF\_INSTANCE\_PORT

The external, or *host-side*, port corresponding to the internal, or *container-side*, port with value `PORT`. This value is generally different from the `PORT` of the app instance.

Example: `CF_INSTANCE_PORT=61045`

## CF\_INSTANCE\_PORTS

The list of mappings between internal, or *container-side*, and external, or *host-side*, ports allocated to the instance's container. Not all of the internal ports are necessarily available for the app to bind to, as some of them may be used by system-provided services that also run inside the container. These internal and external values may differ.

Example: `CF_INSTANCE_PORTS=[{external:61045,internal:8080},{external:61046,internal:2222}]`

## DATABASE\_URL

For apps bound to certain services that use a database, CF creates a `DATABASE_URL` environment variable based on the `VCAP_SERVICES` environment variable at runtime.

CF uses the structure of the `VCAP_SERVICES` environment variable to populate `DATABASE_URL`. CF recognizes any service containing a JSON object with the following form as a candidate for `DATABASE_URL` and uses the first candidate it finds.

```
{
 "some-service": [
 {
 "credentials": {
 "uri": "SOME-DATABASE-URL"
 }
 }
]
}
```

For example, consider the following `VCAP_SERVICES`:

```
VCAP_SERVICES =
{
 "elephantsql": [
 {
 "name": "elephantsql-c6c60",
 "label": "elephantsql",
 "credentials": {
 "uri": "postgres://exampleuser:examplepass@babar.elephantsql.com:5432/exampledbs"
 }
 }
]
}
```

Based on this `VCAP_SERVICES`, CF creates the following `DATABASE_URL` environment variable:

```
DATABASE_URL = postgres://exampleuser:examplepass@babar.elephantsql.com:5432/exampledbs
```

## HOME

Root folder for the deployed app.

Example: `HOME=/home/vcap/app`

## LANG

LANG is required by buildpacks to ensure consistent script load order.

Example: `LANG=en_US.UTF-8`

## MEMORY\_LIMIT

The maximum amount of memory that each instance of the app can consume. You specify this value in an app manifest or with the cf CLI when pushing an app. The value is limited by space and org quotas.

If an instance exceeds the maximum limit, it will be restarted. If Cloud Foundry is asked to restart an instance too frequently, the instance will instead be terminated.

Example: `MEMORY_LIMIT=512M`

## PORt

The port on which the app should listen for requests. The Cloud Foundry runtime allocates a port dynamically for each instance of the app, so code that obtains or uses the app port should refer to it using the `PORT` environment variable.

Example: `PORT=8080`

## PWD

Identifies the present working directory, where the buildpack that processed the app ran.

Example: `PWD=/home/vcap/app`

## TMPDIR

Directory location where temporary and staging files are stored.

Example: `TMPDIR=/home/vcap/tmp`

## USER

The user account under which the app runs.

Example: `USER=vcap`

## VCAP\_APP\_PORT

Deprecated name for the `PORT` variable defined above.

## VCAP\_APPLICATION

This variable contains the associated attributes for a deployed app. Results are returned in JSON format. The table below lists the attributes that are returned.

Attribute	Description
<code>application_id</code>	GUID identifying the app.
<code>application_name</code>	The name assigned to the app when it was pushed.
<code>application_uris</code>	The URLs assigned to the app.

<code>application_version_cf_api</code>	GUID identifying a version of the app. Each time an app is pushed or restarted, this value is updated.
<code>host</code>	Location of the Cloud Controller API for the CF Deployment where the app runs.
<code>limits</code>	Deprecated. IP address of the app instance.
<code>name</code>	The limits to disk space, number of files, and memory permitted to the app. Memory and disk space limits are supplied when the app is deployed, either on the command line or in the app manifest. The number of files allowed is operator-defined.
<code>space_id</code>	Identical to <code>application_name</code> .
<code>space_name</code>	GUID identifying the app's space.
<code>start</code>	Human-readable name of the space where the app is deployed.
<code>started_at</code>	Human-readable timestamp for the time the instance was started. Not provided on Diego Cells.
<code>started_at_timestamp</code>	Identical to <code>started_at</code> . Not provided on Diego Cells.
<code>state_timestamp</code>	Unix epoch timestamp for the time the instance was started. Not provided on Diego Cells.
<code>uris</code>	Identical to <code>application_uris</code> . You must ensure that both <code>application_uris</code> and <code>uris</code> are set to the same value.
<code>users</code>	Deprecated. Not provided on Diego Cells.
<code>version</code>	Identical to <code>application_version</code> .

The following example shows how to set the `VCAP_APPLICATION` environment variable:

```
VCAP_APPLICATION={"instance_id":"fe98dc76ba549876543210abcd1234",
"instance_index":0,"host":"0.0.0.0","port":61857,"started_at":"2013-08-12
00:05:29 +0000","started_at_timestamp":1376265929,"start":"2013-08-12 00:05:29
+0000","state_timestamp":1376265929,"limits":{"mem":512,"disk":1024,"fds":16384}
,"application_version":"ab12cd34-5678-abcd-0123-abcdef987654","application_name"
:"styx-james","application_uris":["my-app.example.com"],"version":"ab1
2cd34-5678-abcd-0123-abcdef987654","name":"my-app","uris":["my-app.example.com"]
,"users":null}
```

## VCAP\_SERVICES

For [bindable services](#), Cloud Foundry adds connection details to the `VCAP_SERVICES` environment variable when you restart your app, after binding a service instance to your app.

The results are returned as a JSON document that contains an object for each service for which one or more instances are bound to the app. The service object contains a child object for each service instance of that service that is bound to the app. The attributes that describe a bound service are defined in the table below.

The key for each service in the JSON document is the same as the value of the “label” attribute.

Attribute	Description
<code>binding_name</code>	The name assigned to the service binding by the user.
<code>instance_name</code>	The name assigned to the service instance by the user.
<code>name</code>	The <code>binding_name</code> if it exists; otherwise the <code>instance_name</code> .
<code>label</code>	The name of the service offering.
<code>tags</code>	An array of strings an app can use to identify a service instance.
<code>plan</code>	The service plan selected when the service instance was created.
<code>credentials</code>	A JSON object containing the service-specific credentials needed to access the service instance.

To see the value of `VCAP_SERVICES` for an app pushed to Cloud Foundry, see [View Environment Variable Values](#).

The example below shows the value of `VCAP_SERVICES` for bound instances of several services available in the [Pivotal Web Services Marketplace](#).

```
VCAP_SERVICES=
{
 "elephantsql": [
 {
 "name": "elephantsql-binding-c6c60",
 "binding_name": "elephantsql-binding-c6c60",
 "instance_name": "elephantsql-c6c60",
 "label": "elephantsql",
 "tags": [
 "postgres",
 "postgresql",
 "relational"
],
 "plan": "turtle",
 "credentials": {
 "uri": "postgres://exampleuser:examplepass@babar.elephantsql.com:5432/exampleuser"
 }
 }
],
 "sendgrid": [
 {
 "name": "mysendgrid",
 "binding_name": null,
 "instance_name": "mysendgrid",
 "label": "sendgrid",
 "tags": [
 "smtp"
],
 "plan": "free",
 "credentials": {
 "hostname": "smtp.sendgrid.net",
 "username": "QvsXMbJ3rK",
 "password": "HCHMOYluTv"
 }
 }
]
}
```

## Environment Variable Groups

Environment variable groups are system-wide variables that enable operators to apply a group of environment variables to all running apps and all staging apps separately.

An environment variable group consists of a single hash of name-value pairs that are later inserted into an app container at runtime or at staging. These values can contain information such as HTTP proxy information. The values for variables set in an environment variable group are case-sensitive.

When creating environment variable groups, consider the following:

- Only the Cloud Foundry operator can set the hash value for each group.
- All authenticated users can get the environment variables assigned to their app.
- All variable changes take effect after the operator restarts or restages the apps.
- Any user-defined variable takes precedence over environment variables provided by these groups.

The table below lists the commands for environment variable groups.

CLI Command	Description
<code>running-environment-variable-group</code> or <code>revg</code>	Retrieves the contents of the running environment variable group.
<code>staging-environment-variable-group</code> or <code>sevg</code>	Retrieves the contents of the staging environment variable group.
<code>set-staging-environment-variable-group</code> or <code>ssevg</code>	Passes parameters as JSON to create a staging environment variable group.
<code>set-running-environment-variable-group</code> or <code>srevg</code>	Passes parameters as JSON to create a running environment variable group.

The following examples demonstrate how to retrieve the environment variables:

```
$ cf revg
Retrieving the contents of the running environment variable group as
sampledeveloper@example.com...
OK
Variable Name Assigned Value
HTTP Proxy 198.51.100.130

$ cf sevg
Retrieving the contents of the staging environment variable group as
sampledeveloper@example.com...
OK
Variable Name Assigned Value
HTTP Proxy 203.0.113.105
EXAMPLE-GROUP 2001

$ cf apps
Getting apps in org SAMPLE-ORG-NAME / space dev as
sampledeveloper@example.com...
OK

name requested state instances memory disk urls
my-app started 1/1 256M 1G my-app.com

$ cf env APP-NAME
Getting env variables for app APP-NAME in org SAMPLE-ORG-NAME / space dev as
sampledeveloper@example.com...
OK

System-Provided:

{
 "VCAP_APPLICATION": {
 "application_name": "APP-NAME",
 "application_uris": [
 "my-app.example.com"
],
 "application_version": "7d0d64be-7f6f-406a-9d21-504643147d63",
 "limits": {
 "disk": 1024,
 "fds": 16384,
 "mem": 256
 },
 "name": "APP-NAME",
 "space_id": "37189599-2407-9946-865e-8ebd0e2df89a",
 "space_name": "dev",
 "uris": [
 "my-app.example.com"
],
 "users": null,
 "version": "7d0d64be-7f6f-406a-9d21-504643147d63"
 }
}

Running Environment Variable Groups:
HTTP Proxy: 198.51.100.130

Staging Environment Variable Groups:
EXAMPLE-GROUP: 2001
HTTP Proxy: 203.0.113.105
```

The following examples demonstrate how to set environment variables:

```
$ cf ssevg '{"test":"198.51.100.130","test2":"203.0.113.105"}'
Setting the contents of the staging environment variable group as admin...
OK
$ cf sevg
Retrieving the contents of the staging environment variable group as admin...
OK
Variable Name Assigned Value
test 198.51.100.130
test2 203.0.113.105

$ cf srevg '{"test3":"2001","test4":"2010"}'
Setting the contents of the running environment variable group as admin...
OK
$ cf revg
Retrieving the contents of the running environment variable group as admin...
OK
Variable Name Assigned Value
test3 2001
test4 2010
```



## Cloud Controller API Client Libraries

This topic describes the client libraries available for developers who want to consume the Cloud Controller API (CAPI).

### Overview

CAPI is the entry point for most operations within the Cloud Foundry (CF) platform. You can use it to manage orgs, spaces, and apps, which includes user roles and permissions. You can also use CAPI to manage the services provided by your CF deployment, including provisioning, creating, and binding them to apps.

For more information, see the [CAPI documentation](#).

### Client Libraries

While you can develop apps that consume CAPI by calling it directly as in the API documentation, you may want to use an existing client library. See the available client libraries below.

#### Supported

CF currently supports the following clients for CAPI:

- [Java](#)
- [Scripting](#) with the Cloud Foundry Command Line Interface (cf CLI)

#### Experimental

The following client is experimental and a work in progress:

- [Golang](#)

#### Unofficial

CF does not support the following clients, but they may be supported by third-parties:

- [Golang](#)
- [Golang](#)
- [Node.js](#)
- Python:
  - [hsdp/python-cf-api](#)
  - [cloudfoundry-community/cf-python-client](#)

## Considerations for Designing and Running an App in the Cloud

### App Design for the Cloud

Apps written in supported frameworks often run unmodified on Cloud Foundry if the app design follows a few simple guidelines. Following these guidelines facilitates app deployment to Cloud Foundry and other cloud platforms.

The following guidelines represent best practices for developing modern apps for cloud platforms. For more detailed reading about good app design for the cloud, see [The Twelve-Factor App](#).

For more information about the features of HTTP routing handled by the Cloud Foundry router, see the [HTTP Routing](#) topic. For more information about the lifecycle of app containers, see the [App Container Lifecycle](#) topic.

### Avoid Writing to the Local File System

Apps running on Cloud Foundry should not write files to the local file system for the following reasons:

- **Local file system storage is short-lived.** When an app instance crashes or stops, the resources assigned to that instance are reclaimed by the platform including any local disk changes made since the app started. When the instance is restarted, the app will start with a new disk image. Although your app can write local files while it is running, the files will disappear after the app restarts.
- **Instances of the same app do not share a local file system.** Each app instance runs in its own isolated container. Thus a file written by one instance is not visible to other instances of the same app. If the files are temporary, this should not be a problem. However, if your app needs the data in the files to persist across app restarts, or the data needs to be shared across all running instances of the app, the local file system should not be used. We recommend using a shared data service like a database or blobstore for this purpose.

For example, instead of using the local file system, you can use a Cloud Foundry service such as the MongoDB document database or a relational database like MySQL or Postgres. Another option is to use cloud storage providers such as [Amazon S3](#), [Google Cloud Storage](#), [Dropbox](#), or [Box](#). If your app needs to communicate across different instances of itself, consider a cache like Redis or a messaging-based architecture with RabbitMQ.

If you must use a file system for your app because for example, your app interacts with other apps through a network attached file system, or because your app is based on legacy code that you cannot rewrite, consider using [Volume Services](#) to bind a network attached file system to your app.

### Cookies Accessible across Apps

In an environment with shared domains, cookies might be accessible across apps.

Many tracking tools such as Google Analytics and Mixpanel use the highest available domain to set their cookies. For an application using a shared domain such as `example.com`, a cookie set to use the highest domain has a `Domain` attribute of `.example.com` in its HTTP response header. For example, an app at `my-app.shared-domain.example.com` might be able to access the cookies for an app at `your-app.shared-domain.example.com`.

You should decide whether or not you want your apps or tools that use cookies to set and store the cookies at the highest available domain.

### Port Considerations

Clients connect to apps running on Cloud Foundry by making requests to URLs associated with the app. Cloud Foundry allows HTTP requests to apps on ports 80 and 443. For more information, see the [Routes and Domains](#) topic.

Cloud Foundry also supports WebSocket handshake requests over HTTP containing the `Upgrade` header. The Cloud Foundry router handles the upgrade and initiates a TCP connection to the app to form a WebSocket connection.

To support WebSockets, the operator must configure the load balancer correctly. Depending on the configuration, clients may have to use a different port for WebSocket connections, such as port 4443, or a different domain name. For more information, see the [Supporting WebSockets](#) topic.

### Cloud Foundry Updates and Your App

For app management purposes, Cloud Foundry may need to stop and restart your app instances. If this occurs, Cloud Foundry performs the following steps:

1. Cloud Foundry sends a single `termination signal` to the root process that your start command invokes.
2. Cloud Foundry waits 10 seconds to allow your app to cleanly shut down any child processes and handle any open connections.
3. After 10 seconds, Cloud Foundry forcibly shuts down your app.

Your app should accept and handle the termination signal to ensure that it shuts down gracefully. To achieve this, the app is expected to follow the steps below when shutting down:

1. App receives termination signal
2. App closes listener so that it stops accepting new connections
3. App finishes serving in-flight requests
4. App closes existing connections as their requests complete
5. App shuts down or is killed

See the [Sample HTTP App ↗](#) GitHub repository for an implementation of the expected shutdown behavior in Golang.

## Ignore Unnecessary Files When Pushing

By default, when you push an app, all files in the app's project directory tree are uploaded to your Cloud Foundry instance, except version control and configuration files or folders with the following names:

- `.cfignore`
- `_darcs`
- `.DS_Store`
- `.git`
- `.gitignore`
- `.hg`
- `manifest.yml`
- `.svn`

In addition to these, if API request diagnostics are directed to a log file and the file is within the project directory tree, it is excluded from the upload. You can direct these API request diagnostics to a log file using `cf config --trace` or the `CF_TRACE` environment variable.

If the app directory contains other files, such as `temp` or `log` files, or complete subdirectories that are not required to build and run your app, you might want to add them to a `.cfignore` file to exclude them from upload. Especially with a large app, uploading unnecessary files can slow app deployment.

To use a `.cfignore` file, create a text file named `.cfignore` in the root of your app directory structure. In this file, specify the files or file types you wish to exclude from upload. For example, these lines in a `.cfignore` file exclude the “tmp” and “log” directories.

```
tmp
log
```

The file types you will want to exclude vary, based on the app frameworks you use. For examples of commonly-used `.gitignore` files, see [https://github.com/github/gitignore ↗](https://github.com/github/gitignore).

## Run Multiple Instances to Increase Availability

Singleton apps may become temporarily unavailable for reasons that include:

- During an upgrade, Pivotal Cloud Foundry (PCF) gracefully shuts down the apps running on each Diego cell and then restarts them on another Diego cell. Single app instances may become temporarily unavailable if the replacement instance does not become healthy within the cell's evacuation timeout, which defaults to 10 minutes.
- Unexpected faults in PCF system components or underlying infrastructure, such as container-host VMs or IaaS Availability Zones, may cause lone app instances to disappear or become unroutable for a minute or two.

To avoid the risk of an app becoming temporarily unavailable, developers can run more than one instance of the app.

## Using Buildpacks

A buildpack consists of bundles of detection and configuration scripts that provide framework and runtime support for your apps. When you deploy an app that needs a buildpack, Cloud Foundry installs the buildpack on the Diego cell where the app runs.

For more information, see the [Buildpacks](#) topic.

## App Revisions

This topic describes app revisions and provides procedures for working with them.

For additional information, see the [Revisions section of the CAPI documentation](#).

 **Note:** CAPI v3 is the recommended API version for revisions. While revisions work with v2, there are several inconsistencies. For example, revision descriptions for apps with multiple processes may be inaccurate because v2 does not support apps with multiple processes. Additionally, pushing an app for the first time with revisions in v2 creates two revisions.

## Overview

A revision represents code and configuration used by an app at a specific time. It is a Cloud Foundry API (CAPI) object that can contain references to a droplet, custom start command, and environment variables. The most recent revision for a running app represents code and configuration currently running in Cloud Foundry.

Some use cases for revisions include:

- **Viewing revisions for an app:** This can help you understand how your app has changed over time.
- **Rolling back to a previous revision:** This allows you to deploy a version of the app that you had running previously without needing to track that previous state yourself or have multiple apps running. When you create a deployment and reference a revision, the revision deploys as the current version of your app.

## Events that Trigger Revisions

Revisions are generated automatically through the following events:

- A new droplet is created for an app
- An app's environment variables are changed
- The custom start command for an app is added or changed
- An app rolls back to a prior revision

By default, the Cloud Foundry API retains at most 100 revisions per app.

## Revision Descriptions

Each revision includes a description of what changed in your app at the time the revision was created. The description includes one or more of the following:

- Process type removed
- New process type added
- Rolled back to revision X
- Custom start command removed
- Custom start command updated
- Custom start command added
- New environment variables deployed
- New droplet deployed

## Droplet Storage Considerations

By default, Cloud Foundry retains the 5 most recent staged droplets in its droplets bucket. This means that you can roll back to revisions as long as they are using one of those 5 droplets. Note that not all revisions include a change in droplet.

Operators can configure Cloud Foundry to retain more droplets if necessary using the **Max Staged Droplets per App** field in the **File Storage** pane of the

Pivotal Application Service (PAS) tile.

## Enable Revisions for an App

To enable revisions for an app, do the following:

1. Get the GUID of the app:

```
cf app APP-NAME --guid
```

2. Run the following command:

```
cf curl /v3/apps/GUID/features/revisions -X PATCH -d '{ \"enabled\": true }'
```

Where `GUID` is the GUID you got in the previous step.

## View Revisions

This section describes how to use CAPI endpoints for viewing revisions.

### List Revisions for an App

To list revisions for an app, do the following:

1. Get the GUID of the app:

```
cf app APP-NAME --guid
```

2. Run the following command:

```
cf curl /v3/apps/GUID/revisions
```

Where `GUID` is the GUID you got in the previous step.

### List Deployed Revisions for an App

Deployed revisions are revisions linked to started processes in an app. To list deployed revisions, do the following:

1. Get the GUID of the app:

```
cf app APP-NAME --guid
```

2. Run the following command:

```
cf curl /v3/apps/GUID/revisions/deployed
```

Where `GUID` is the GUID you got in the previous step.

### Get a Revision

To get a revision, do the following:

1. Run the following command:

```
cf curl /v3/revisions/GUID
```

Where `[GUID]` is the GUID of the revision.

## Rollback to a Previous Revision

To roll back to a previous revision, do the following:

1. Get the GUID of the app:

```
cf app APP-NAME --guid
```

2. Get the GUID of the revision. See [View Revisions](#).

3. Create a deployment using CAPI:

```
cf curl v3/deployments \
-X POST \
-d '{
 "revision": {
 "guid": "REVISION-GUID"
 },
 "relationships": {
 "app": {
 "data": {
 "guid": "APP-GUID"
 }
 }
 }
}'
```

Where:

- o `REVISION-GUID` is the GUID of the revision.
- o `APP-GUID` is the GUID of the app.

## Add Metadata to a Revision

To add metadata to a revision, follow the steps to [Add Metadata to an Object](#).

## PCF Security and Compliance Guide

### For Security Professionals and PCF Users

This guide explains how Pivotal Cloud Foundry (PCF) manages network access, roles and permissions, internal communications, container hardening, and other security issues. It is intended to give security professionals a complete view of PCF security, and to help all PCF users, not just the security experts, keep the platform secure.

In addition, this guide provides information about PCF compliance with published control standards and regulations such as [NIST Special Publication 800-53\(r4\)](#) and [GDPR](#).

### Pivotal Security Processes and CVE Reports

Pivotal publishes security updates regularly in response to privately- and publicly-reported Common Vulnerabilities and Exposures (CVEs).

- See the latest CVEs on the [Pivotal Application Security Team](#) page.
- To learn about Pivotal's vulnerability reporting and responsible disclosure process, read [PCF Security Overview and Policy](#).
- To learn about the testing, release and security lifecycle of PCF, see [PCF Testing, Release, and Security Lifecycle](#).

### Security

- [Security Concepts](#): Provides links to conceptual documentation about how security is implemented in PCF.
- [PCF Infrastructure Security](#): Provides guidance and procedures for securing PCF infrastructure such as hardening stemcells and managing the certificates that enable TLS communication.
- [Network Security](#): Covers the security aspects of PCF networking such as the paths, ports, and protocols that components use to communicate.
- [Credential and Identity Management](#): Describes how PCF manages permissions and trust for PCF user accounts. Also provides documentation about CredHub, the credential management system that BOSH uses to store deployment credentials and that PCF runtimes use to create and manage app and service credentials.
- [Security for Apps and Services](#): Collects documentation about the security mechanisms that surround apps and services running on PCF.
- [Certificates on PCF](#): How certificates are utilized in PCF to secure both internal and external networking calls.
- [Security Processes and Stemcells](#): How Pivotal responds to security vulnerabilities, and how it tests and updates the versioned operating systems that its products run on.

### Compliance

- [NIST Controls and PCF](#): Provides a dedicated site that assesses Pivotal Cloud Foundry against NIST SP 800-53(r4) Controls.
- [General Data Protection Regulation](#): Provides an overview of the General Data Protection Regulation (GDPR) and where Pivotal Cloud Foundry (PCF) may store personal data

## Security Processes

This section explains how Pivotal responds to security vulnerabilities, and how it tests and updates its stemcells, the versioned operating systems that its products run on.

- [Pivotal Cloud Foundry Security Overview and Policy](#): Covers Pivotal's responsible disclosure and vulnerability response procedures for the Pivotal Cloud Foundry (PCF) platform.
- [PCF Testing, Release, and Security Lifecycle](#): Explains how Pivotal's practices, tools, and organizational structures work together to create and support stable releases of PCF.
- [PKS Security Disclosure and Release Process](#): Describes the processes for disclosing security issues and releasing related fixes for Pivotal Container Service (PKS), Kubernetes, Cloud Foundry Container Runtime (CFCR), VMware NSX, and VMware Harbor.

## Pivotal Cloud Foundry Security Overview and Policy

This document outlines our security policy and is addressed to operators deploying [Pivotal Cloud Foundry](#) (PCF) using Pivotal Cloud Foundry Operations Manager.

For a comprehensive overview of the security architecture of each PCF component, refer to the [Cloud Foundry Security](#) topic.

## How Pivotal Monitors for Security Vulnerabilities

Pivotal receives private reports on vulnerabilities from customers and from field personnel via our secure disclosure process. We also monitor public repositories of software security vulnerabilities to identify newly discovered vulnerabilities that might affect one or more of our products.

## How to Report a Vulnerability

Pivotal encourages users who become aware of a security vulnerability in our products to contact Pivotal with details of the vulnerability. Please send descriptions of any vulnerabilities found to [security@pivotal.io](mailto:security@pivotal.io). Please include details on the software and hardware configuration of your system so that we can reproduce the issue.

 **Note:** We encourage use of encrypted email. Our public PGP key is located at <http://www.pivotal.io/security>.

## Notification Policy

PCF has many customer stakeholders who need to know about security updates. When there is a possible security vulnerability identified for a PCF component, we do the following:

1. Assess the impact to PCF.
2. If the vulnerability would affect a PCF component, we schedule an update for the impacted component(s).
3. Update the affected component(s) and perform system tests.
4. Announce the fix publicly via the following channels:
  - a. Automated notification to end users who have downloaded or subscribed to a PCF product on [Pivotal Network](#) when a new, fixed version is available.
  - b. Adding a new post to <http://www.pivotal.io/security>.

## Classes of Vulnerabilities

Attackers can exploit vulnerabilities to compromise user data and processing resources. This can affect data confidentiality, integrity, and availability to different degrees. For vulnerabilities related to Ubuntu provided packages, Pivotal follows [Canonical's priority levels](#). For other vulnerabilities, Pivotal follows [Common Vulnerability Scoring System v3.0 standards](#) when assessing severity.

Pivotal uses Canonical's Ubuntu distribution of Linux for PCF Ubuntu stemcells and rootfs. Canonical provides Pivotal with support services allowing us to escalate CVEs that we determine may affect PCF. In general, Pivotal does not escalate to upstream open source software components or vendors for Medium or Low CVEs that are not yet patched. PCF may escalate on behalf of a customer for High or Critical CVEs. PCF customers who are interested in addressing CVEs in Ubuntu that are not yet patched can establish their own support relationship with [Canonical](#).

Pivotal reports the severity of vulnerabilities using the following severity classes:

### High

High severity vulnerabilities are those that can be exploited by an unauthenticated or authenticated attacker, from the Internet or those that break the guest/host Operating System isolation. The exploitation could result in the complete compromise of confidentiality, integrity, and availability of user data and/or processing resources without user interaction. Exploitation could be leveraged to propagate an Internet worm or execute arbitrary code between Virtual Machines and/or the Host Operating System. This rating also applies to those vulnerabilities that could lead to the complete compromise of

availability when the exploitation is by a remote unauthenticated attacker from the Internet or through a breach of virtual machine isolation.

## Moderate

Moderate vulnerabilities are those in which the ability to exploit is mitigated to a significant degree by configuration or difficulty of exploitation, but in certain deployment scenarios could still lead to the compromise of confidentiality, integrity, or availability of user data and/or processing resources.

## Low

Low vulnerabilities are all other issues that have a security impact. These include vulnerabilities for which exploitation is believed to be extremely difficult, or for which successful exploitation would have minimal impact.

## Release Policy

PCF schedules regular monthly releases of software in the PCF Suite to address Low / Medium severity vulnerability exploits. When High severity vulnerability exploits are identified, PCF releases fixes to software in the PCF Suite on-demand, with as fast a turnaround as possible.

## Alerts/Actions Archive

<http://www.pivotal.io/security> ↗

## PCF Testing, Release, and Security Lifecycle

This topic explains how Pivotal's development practices, automated build tools, and organizational structures work together to create and support stable releases of Pivotal Cloud Foundry (PCF).

### Summary

- PCF teams building system components receive frequent feedback, which helps to secure code from exposure to vulnerability.
- Every PCF release follows a strict workflow and passes through numerous quality and compliance checks before distribution.
- Teams build tests into the product consistently and run them automatically with any code change.

### Release Mechanics

Pivotal releases, patches, and supports multiple versions of PCF simultaneously. This section explains the versioning and support conventions Pivotal follows.

### Versioning

Pivotal numbers PCF releases following a [semantic versioning](#) style format, X.Y.Z, where X and Y indicate major and minor releases, and Z designates patch releases. Major and minor releases change PCF functionality; patch releases are backward-compatible security patches or bug fixes.

### Support

As of PCF 1.8, Pivotal supports each major and minor PCF release according to the following criteria:

- Pivotal supports the release for at least 9 months following its first publication date.
- Pivotal supports the last three major or minor releases, even if this extends coverage beyond 9 months.

Support includes maintenance updates and upgrades, bug and security fixes, and technical assistance. The [Pivotal Support Policy](#) describes support standards, technical guidance, and publication phases in more detail. The Pivotal Support Services [Terms and Conditions](#) defines Pivotal support in legal terms.

### Patch Releases

Patch releases are more frequent and less predictable than major/minor releases. The v1.6.x line provides a good example of their frequency. PCF 1.6.1 was released on October 26, 2015. Through August 2016, 36 additional patches of Elastic Runtime 1.6.x and 18 patches of Ops Manager 1.6.x provided security and bug fixes to customers.

Pivotal identifies security issues using standard nomenclature from [Common Vulnerabilities and Exposures \(CVE\)](#), [Ubuntu Security Notices \(USN\)](#), and other third party sources. Read about security fixes in core Cloud Foundry code or packaged dependencies in the release notes for [Ops Manager](#) and [Pivotal Application Service \(PAS\)](#).

[Pivotal.io/security](#) maintains a running list of security fixes in PCF and PCF dependencies. Consult that page to see the most recent findings from Pivotal's security team.

### Upgrading

All PCF releases pass through extensive test suites that include automated unit, integration, and acceptance tests on multiple IaaSes. Regardless of this extensive testing, Pivotal recommends that you test major and minor releases in a non-production environment before implementing them across your deployment. Upgrade your production environment as soon as possible after you validate the new release on your test environment.

### Release Testing, Integration, and Validation

This section describes Pivotal's software development processes and explains compliance and regulatory standards to which Pivotal software adheres.

## Test-Driven Development

Pivotal's development process relies on a strict workflow with continuous automated testing. Pivotal R&D does not separate engineering and testing teams. Rather, every Pivot on each engineering team is responsible for ensuring the quality of their code. They write tests for all of the software components that they develop, often before writing the software itself.

With every software change, automated build pipelines trigger these tests for the new software component and for everything it touches. If a new code check-in does not pass its tests or causes a failure elsewhere, it pauses the build pipeline for the entire team, or sometimes all of Pivotal R&D. The transparency of this process encourages developers to work together to address code issues quickly.

Pivotal applies the following automated testing approaches, scenarios, and frameworks to PCF components and to the release as a whole:

- **Unit tests:** Development teams write unit tests to express and validate desired functional behavior of product components. Typical frameworks used are [RSpec](#) and [Ginkgo](#). These tests run continuously throughout the development cycle.
- **OSS integration tests:** The Release Integration team exercises a full deployment of open-source Cloud Foundry to validate all end-user features. They maintain the [Cloud Foundry Acceptance Test](#) (CATs) suite alongside the OSS cf-release. Cloud Foundry component teams also contribute acceptance test suites at the OSS Integration Test level. These tests exercise and validate their components' functional, performance, and integration health.
- **PCF integration tests:** The PCF Release Engineering (RelEng) team validates the quality and cross-product integration health of the commercial PCF release. RelEng runs OSS Acceptance Tests against all supported releases. These tests run on full PCF instances configured to represent diverse real-world customer scenarios on various IaaSes and using both internal and external load balancer, database, blobstore, and user store solutions.

## Additional Pre-Release Gates: Internal, PWS, and Compliance

In addition to its automated unit and integration testing, Pivotal deploys all upgrades slated for upcoming PCF releases on at-scale test environments. Prior to each Major or Minor commercial release, Pivotal runs the entire Pivotal Cloud Foundry Suite of services on several internally-managed large integration environments that run customer-like data and workloads.

Pivotal also pushes upcoming PCF feature upgrades and patches to its [Pivotal Web Services](#) platform, where customers continually deploy and host hundreds of mission-critical applications at scale, 24/7. The PWS environment gives Pivotal a continuous source of real-world usage and performance metrics that inform product development teams.

All PCF product teams participate in go-to-market steps for each release, as is often required for shipping a legally compliant product. Examples include [Open Source License File](#) attribution and an Export Compliance classification.

## Patch Releases: Security and Bug Fixes

Pivotal uses established processes to track, disclose, and remediate vulnerabilities in PCF and related dependent components. This section explains how Pivotal identifies vulnerabilities and implements fixes for them.

### Identifying Security Vulnerabilities

Pivotal has an established process to track and patch vulnerabilities in software dependencies and PCF software. Additionally, [pivotal.io/security](#) describes a responsible disclosure process for reporting vulnerabilities identified in Pivotal software by 3rd parties.

Pivotal uses multiple methods to identify security vulnerabilities in Pivotal software and dependencies internally, including:

- Security notifications from [Canonical](#) for their Ubuntu operating system, provided through Pivotal's commercial relationship with Canonical
- Software component scans several times per day, using 3rd party security software which updates continuously from external security vulnerability sources
- Dependency analysis software that identifies and catalogs software dependencies
- Security vulnerability notifications from known software dependencies

Pivotal also monitors externally-reported vulnerabilities from many sources, including:

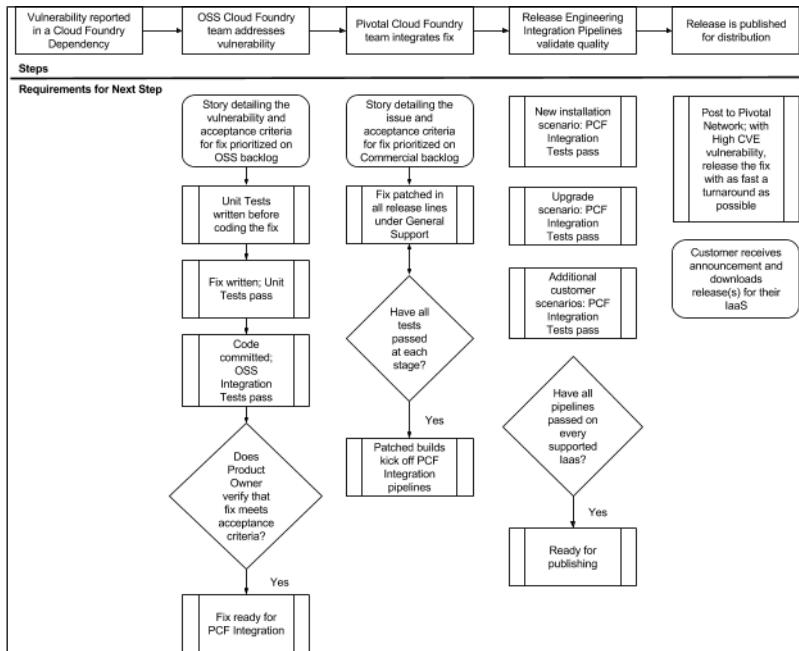
- 3rd party security analysis requested by Pivotal

- Cloud Foundry Foundation security notifications from member companies
- Customer, prospect and other 3rd party security reports

When Pivotal discovers a potential security vulnerability in PCF, the security team opens an issue to assess it. If it confirms the vulnerability exists, Pivotal identifies and updates affected components with plans to backport the fix to stable releases. Fixes are implemented on a target timeline based on the [severity level](#) of the vulnerability.

## Fix, Test, and Release Lifecycle

This flowchart details the steps that Pivotal performs on a typical high-priority CVE, to publish a patch release fix on <https://network.pivotal.io>:



## Security Concepts

This section provides links to conceptual documentation that describes how Pivotal Cloud Foundry (PCF) implements security at different levels of the platform.

- [Cloud Foundry Security ↗](#): Provides an overview of the measures Cloud Foundry implements to minimize security risks.
- [Container Security ↗](#): Describes how PCF isolates containers and limits privileges for containers.
- [Container-to-Container Networking ↗](#): Illustrates how the Container-to-Container Networking feature enables and secures internal app-to-app communication.
- [Application Security Groups ↗](#): Explains the different types of application security groups, how they work and how to apply them.
- [Application SSH Components and Processes ↗](#): Provides an overview of the components that support SSH access to apps.
- [Floating Stemcells ↗](#): Describes how PCF automatically upgrades all compatible products when a new stemcell is available.

## PCF Infrastructure Security

This section provides links to topics about security infrastructure in Pivotal Cloud Foundry (PCF).

- [Stemcell Security ↗](#): Details how Pivotal tests and updates its stemcells, which supply the base versioned operating system for all PCF products.
- [Certificates and TLS in PCF ↗](#): Provides information on how to manage the certificates that enable TLS communications in PCF.
- [Disk Encryption ↗](#): Describes how to encrypt disks or rotate disk encryption keys in a PCF deployment.
- [Security Event Logging ↗](#): Provides a reference of security-related events logged by the Cloud Controller, CredHub and UAA components in PCF.

## Stemcell Security

This section provides links to topics about how Pivotal implements security in the stemcell layer of Pivotal Cloud Foundry (PCF).

- [Floating Stemcells ↗](#): How PCF automatically upgrades all compatible products when a new stemcell is available.
- [Linux Stemcell Hardening ↗](#): How Pivotal secures Linux stemcells through regular testing and minimizing their surface of vulnerability.

For more information about stemcell availability and related security fixes, see [Stemcell Release Notes ↗](#).

## Linux Stemcell Hardening

 **Note:** This document applies to stemcell v3263 and later.

Customers and prospects often ask for details on stemcell hardening, i.e., the process by which we secure Pivotal Cloud Foundry by reducing its vulnerability surface from outside access. This document provides responses to some commonly-asked questions regarding the security configuration enhancements and hardening tests that Pivotal applies to the Cloud Foundry (“CF”) stemcell. This information will be helpful to customer accreditation teams who are responsible for running configuration scans of a Cloud Foundry deployment, and also to auditors who need a documentation artifact to feed into the customers’ existing security assessment processes.

**1. WHAT IS A STEMCELL?** A stemcell is a versioned Operating System (“OS”) image wrapped with IaaS specific packaging. A typical stemcell contains a bare minimum OS skeleton with a few common utilities pre-installed, a BOSH Agent, and a few configuration files to securely configure the OS by default. For example: with vSphere, the official stemcell for Ubuntu Trusty is an approximately 500MB VMDK file. With AWS, official stemcells are published as MIs that can be used in an AWS account. Stemcells do not contain any specific information about any software that will be installed once that stemcell becomes a specialized machine in the cluster; nor do they contain any sensitive information which would make them unable to be shared with other BOSH users. This clear separation between base OS and later-installed software is what makes stemcells a powerful concept. In addition to being generic, stemcells for one OS (e.g. all Ubuntu Trusty and Xenial stemcells) are exactly the same for all infrastructures. This property of stemcells allows BOSH users to quickly and reliably switch between different infrastructures without worrying about the differences between OS images. The CF BOSH team is responsible for producing and maintaining an official set of stemcells. Cloud Foundry currently supports Ubuntu Trusty and Xenial on vSphere, AWS, OpenStack, Google, and Azure infrastructures.

**2. WHAT IS STEMCELL HARDENING?** Stemcell hardening is the process of securing a stemcell by reducing its surface of vulnerability, which is larger when a system performs more functions; in principle a single-function system is more secure than a multipurpose one. There are various methods of hardening Linux systems. Common techniques include reducing available methods of attack by implementing more restrictive and/or conservative configurations of the OS kernel and system services, changing default passwords, the removal of unnecessary software, unnecessary usernames and logins, and the disabling or removal of unnecessary services.

**3. WHAT IS OUR GENERAL APPROACH TO STEMCELL HARDENING?** The CF stemcell is essentially a distinct Linux distribution. As such, industry-standard benchmarks are not entirely appropriate when assessing the security posture of the stemcell, but Pivotal has considered and incorporated hardening guidance from various sources both commercial and government. Some parts of the existing recommended industry-standard hardening configurations will certainly apply, but some other parts do not apply. In addition, because each stemcell is a unique Linux distribution, existing industry-standard benchmarks are silent on some important aspects of hardening the stemcell configurations. The following paragraphs describe the different categories of stemcell hardening configurations, and provide a count of the number of tests currently in each category. **Note:** The most current description of what has been delivered is always available in the BOSH public Pivotal Trackers.

- a. **Baseline Passing:** common hardening tests that pass without any changes to the stemcell or to test procedures. (*130 tests*)
- b. **Test Amended:** Stemcells are optimized for cloud deployment and some configuration settings are not stored in traditionally-expected locations. The industry standard test was changed to conform with stemcell design to accurately check the recommended setting. This new test reflects the changes to the industry standard test but the stemcell adheres to commonly accepted guidance. (*36 tests*)
- c. **Additional Hardening:** Configuration hardening improvements that have been made to the stemcell. As with most software, a stemcell’s security improves over time and every stemcell release is tested to ensure that it is suitable for use with its associated CF release. Later releases of a stemcell may include additional security features that were not present in earlier releases. (*86 tests*)
- d. **New CF-specific Tests:** New tests that have been added to check CF stemcell-specific configurations. These tests are not yet part of any industry standard Ubuntu benchmark. This category of tests is still under development and additional tests will be added over time. (*20 tests*)

### 4. WHAT ARE THE MAJOR FOCUS AREAS FOR OUR STEMCELL HARDENING APPROACH?

- a. **Maintenance, Updates, and Patching**
  - i. Regular patches and feature enhancements are delivered via routine BOSH deployments of updated stemcells (obviates apt-get upgrade).
- b. **File System Hardening**
  - i. The /tmp directory is configured to be on a separate partition.
  - ii. Users cannot create character or block special devices in the /tmp filesystem.
  - iii. Users cannot create set userid files in the /tmp filesystem.
  - iv. Users cannot run executable binaries from the /tmp filesystem.
  - v. The temporary storage directories such as /tmp and /var/tmp are mounted on a dedicated partition, and configured with appropriately limiting options such as nodev, nosuid, and noexec.
  - vi. Each of the following directories is in a separate partition, with mount options managed via BOSH agent:
    - /var
    - /var/log
    - /var/log/audit
    - /home

- /tmp

- vii. File system mount options for users' home directories are limited via appropriate mount options including nodev.
- viii. Removable media may not be mounted as character or block special device.
- ix. Executable programs may not run from removable media.
- x. setuid and setgid are not allowed on removable media.
- xi. Users cannot create special devices in shared memory partitions.
- xii. Users cannot put privileged programs onto shared memory partitions.
- xiii. Users cannot execute programs from shared memory partitions.
- xiv. Users cannot delete or rename files in world-writable directories such as /tmp that are owned by other users.
- xv. Supplementary and exotic Linux file systems that are unused in CF have been disabled.
- xvi. Additional supplementary and exotic Linux file systems that are unused in CF have been disabled.
- xvii. Automount of USB drives or disks is not permitted.

#### c. Boot Security

- i. The owner and group for the bootloader config (/boot/grub/grub.cfg) is set to root. Only root has read and write access to this file.
- ii. Boot loader has been configured so that a password is required to reboot the system.
- iii. Unauthorized users cannot reboot the system into single user mode.

#### d. Process Security

- i. Users cannot override the soft limit for core dumps.
- ii. Randomized virtual memory region placement is enabled.
- iii. Prelinking of shared libraries is disabled.

#### e. Minimization of Attack Surface

- i. The Network Information Service ("NIS") is not used in CF and is not installed.
- ii. The Berkeley rsh-server package is not used in CF and is disabled.
- iii. Classic rsh-related tools are not used in CF and are not installed.
- iv. The following servers are not used on CF stemcells and are disabled:

- talk server
- telnet server
- tftp-server
- Avahi
- print
- DHCP
- DNS
- FTP
- IMAP
- POP
- HTTP
- SNMP

- v. The talk client is not used in CF and is not installed.
- vi. The eXtended InterNET Daemon (xinetd) is not used in CF and is disabled.
- vii. The following network services are not used in CF and are disabled:

- chargen
- daytime
- echo
- discard
- time

- viii. The X Window system is not used in CF and is not installed.
- ix. NTP time setting is synchronized on the stemcell via the ntpdate utility.
- x. The Samba daemon is not used in CF and is disabled.
- xi. The Mail Transfer Agents (MTA) process only local mail.
- xii. The rsync service is not used in CF and is disabled.
- xiii. The biosdevname tool is disabled.

#### f. Network Security

- i. IPv4 networking is configured such that IP forwarding is disabled.
- ii. The IPv4 networking has been configured such that the host cannot send ICMP redirects.
- iii. IPv4 networking has been configured such that the system does not accept source routed packets.
- iv. IPv4 networking is configured such that ICMP redirects are not accepted.
- v. ICMP echo and timestamp requests with broadcast or multicast destinations will be ignored.

- vi. The stemcell will ignore malformed ICMP error responses.
- vii. IPv4 networking is configured for source route validation.
- viii. TCP SYN cookies are enabled.
- ix. Stemcells are set to refuse IPv6 router advertisements.
- x. The /etc/hosts.allow file exists and is empty.
- xi. The /etc/hosts.allow and /etc/hosts.deny files are protected from unauthorized write access.
- xii. The /etc/hosts.deny file exists and is empty.
- xiii. The following protocols are not used in CF and are disabled:

- SCTP
- DCCP
- TIPC
- LDAP
- RDS

- xiv. Wireless interfaces are disabled.
- xv. IPv6 is not used in CF deployments and the IPv6 protocol is disabled.

## g. Auditing

- i. Audit log file size is configured for a manageable maximum size of 6 MB.
- ii. The system auditd logs have been configured such that the system is resilient in the event of a denial of service attack on the auditd daemon.
- iii. Auditd daemon is configured such that all auditd logs are kept after rotation.
- iv. The auditd service is enabled.
- v. Auditing of successful and failed login/logout events is enabled.
- vi. The Linux auditing subsystem has been configured in accordance with best practice industry guidance to capture all security-relevant events. The /etc/audit/audit.rules configuration now contains more than 50 monitoring rules.
- vii. Audit records are created for loading and unloading of kernel modules and for system calls.
- viii. File Integrity Monitoring can be done on the stemcell (via a BOSH Add-on).

## h. Authentication and Authorization

- i. The cron daemon is enabled.
- ii. Access to the /etc/crontab file is limited to root.
- iii. Access to the cron utility configuration via the hourly, daily, weekly, and monthly directories is limited.
- iv. User authorization to schedule cron jobs is limited.
- v. Only the vcap user is whitelisted to use the cron and at utilities.
- vi. Password requirements follow industry best practice guidance and enforce a minimum length of 14 characters, with at least one each of: digit, uppercase, lowercase and special characters.
- vii. Password reuse: users cannot reuse their twenty most recent passwords.
- viii. SSH protocol version is configured for SSH-2.
- ix. Logging level for SSH event is INFO.
- x. Minimum permissions are set on /etc/ssh/sshd\_config.
- xi. SSH X11 forwarding is disabled.
- xii. The MaxAuthTries parameter for SSH is set to 3 attempts per connection.
- xiii. SSH is configured to require passwords and ignore host-based authentication.
- xiv. Root logins are not allowed over SSH.
- xv. Users cannot set environment variables through the SSH daemon.
- xvi. SSH has been configured to use strong ciphers:
  - aes128-ctr
  - aes192-ctr
  - aes256-ctr
- xvii. Idle SSH sessions are terminated after 15 minutes, and no client “keep alive” messages are sent.
- xviii. Idle SSH sessions are terminated after 15 minutes. No client “keep alive” messages are sent.
- xix. The SSH login banner may be configured to display site-specific text before user authentication is permitted (via BOSH Add-on).
- xx. Root login is only permitted via console, not via tty devices.
- xxi. Only the vcap user is authorized in the sudo group.
- xxii. Only users in the root group (a.k.a. wheel) are authorized to run the su command.

### i. Compliance

- i. Contents of /etc/issue and /etc/issue.net have been configured to the phrase: “Unauthorized use is strictly prohibited. All access and activity is subject to logging and monitoring.” This may be amended if and as necessary via a BOSH Add-on.
- ii. The Message of the Day file /etc/motd is not used, but may be populated via a BOSH Add-on if needed.
- iii. Identification of the OS and/or version information about the OS does not appear in any login banners.

j. **File System Permissions**

- i. The /etc/passwd, /etc/shadow, and /etc/group files are protected from unauthorized write access.
- ii. Use and/or presence of any world-writable files has been audited, and minimized to the extent possible for CF.
- iii. By default, all stemcell files are owned by a known user and group, and may not belong to a non-existent user or group.
- iv. Use of SUID and GID is restricted, and only the /usr/bin/sudo and /bin/su programs are authorized as SUID and/or GID programs.

k. **User Account Management**

- i. Users cannot change their password more than once a day.
- ii. Users are notified 7 days before their passwords expire.
- iii. Interactive logins are disabled for system accounts.
- iv. The GID for the root account is 0.
- v. User accounts may not have empty passwords.
- vi. NIS is not used in CF, and integration of OS security configuration with legacy NIS permissioning is not enabled (e.g., for /etc/passwd, shadow, and group).
- vii. By default, the only UID 0 account present is root.
- viii. By default, the root PATH does not include any risky directory such as the current working (.) or any writable directory.
- ix. Minimum privileges are applied to all users' hidden configuration ("dot") files.
- x. The .netrc and .rhosts and .forward files are not used in CF and are not present in any user home directory.
- xi. Any group present in the /etc/passwd file must also exist in the /etc/group file.
- xii. Users defined in /etc/password must have a valid home directory.
- xiii. Users must own their home directories.
- xiv. All references to user and group names, as well as UID and/or GID identifiers, are self consistent, with no duplicates or orphans allowed.
- xv. By default, the shadow group is not used in CF and must be empty.

## Certificates and TLS in PCF

This section provides links to topics about certificate and TLS infrastructure in Pivotal Cloud Foundry (PCF).

- [TLS Connections in PCF ↗](#): Provides information on TLS is used in PCF deployments, including supported TLS cipher-suites.
- [Securing Traffic into Cloud Foundry ↗](#): Provides instructions on how to how to configure Transport Layer Security (TLS) termination for HTTP traffic into PCF.
- [Providing a Certificate for Your TLS Termination Point ↗](#): Describes how to configure TLS certificates for Pivotal Application Service (PAS).
- [Managing Certificates with the Ops Manager API ↗](#): Describes how to manage internal certificate authorities (CAs) and certificates in PCF that are visible to the Ops Manager API.
- [Rotating Certificates ↗](#): Describes how to rotate internal CAs and certificates in PCF.
- [Adding a Custom Certificate Authority ↗](#): Describes how to add a custom CA to issue digital certificates in a Pivotal Cloud Foundry (PCF) deployment
- [Trusted System Certificates ↗](#): Discusses where apps deployed to PAS can find trusted system certificates.

## TLS Connections in PCF

This topic classifies the different paths through which external clients, internal components, app containers, and app services communicate in Pivotal Cloud Foundry (PCF), and how the platform uses Transport Layer Security (TLS) protocols to secure these communications.

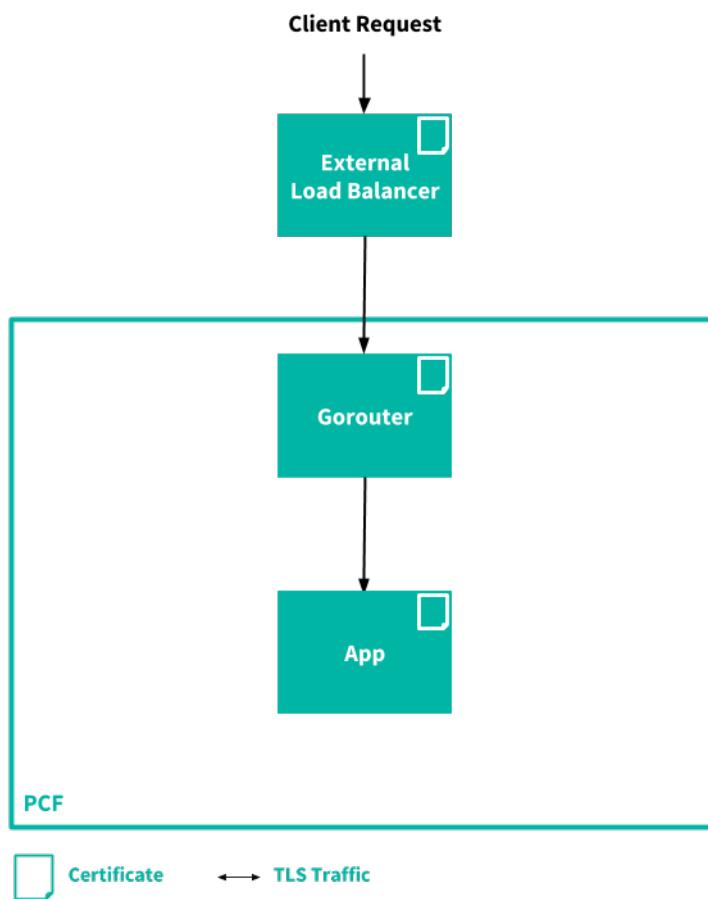
## Types of Network Communication in PCF

This section classifies the different types of network communication in PCF and how they are secured with TLS.

Within a PCF deployment, TLS secures connections between components like the BOSH Director and service tiles. PCF components also use TLS connections to secure communications with external hardware, such as customer load balancers.

### Between an External Client and an App

The following diagram illustrates the flow of communication from a client making a request to an app:

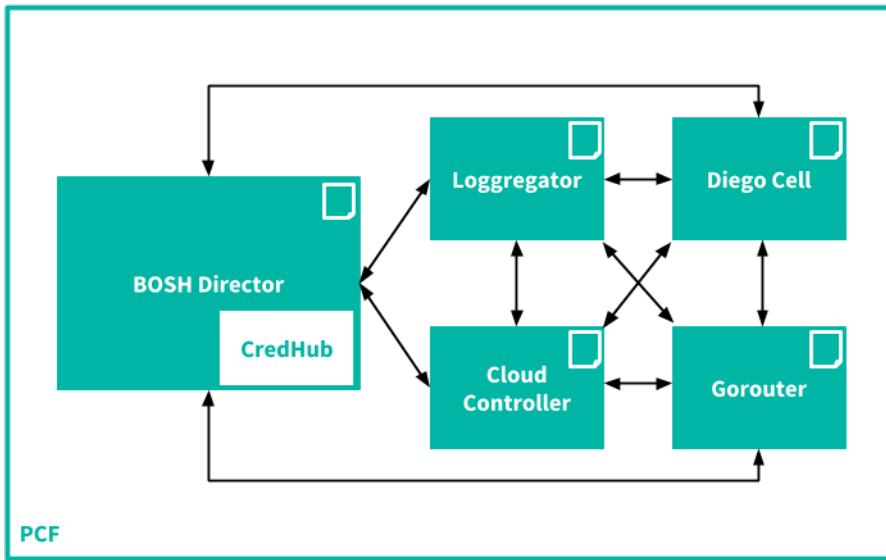


The following table describes each component involved in receiving a client request and where their certificates for TLS termination originate.

Component	Certificate Source
External Load Balancer	Enterprise Root CA.
Gorouter	Enterprise Root CA.
App	PCF root CA dedicated to app instance identity. For more information, see <a href="#">App Instance Container Identity Credentials</a> .

### Between Platform Components

The following diagram illustrates communication between platform components, secured with TLS.



The CredHub instance in BOSH generates certificates for all components in PCF. The certificates are self signed by default. To issue certificates signed by your enterprise, you can add a custom CA to CredHub.

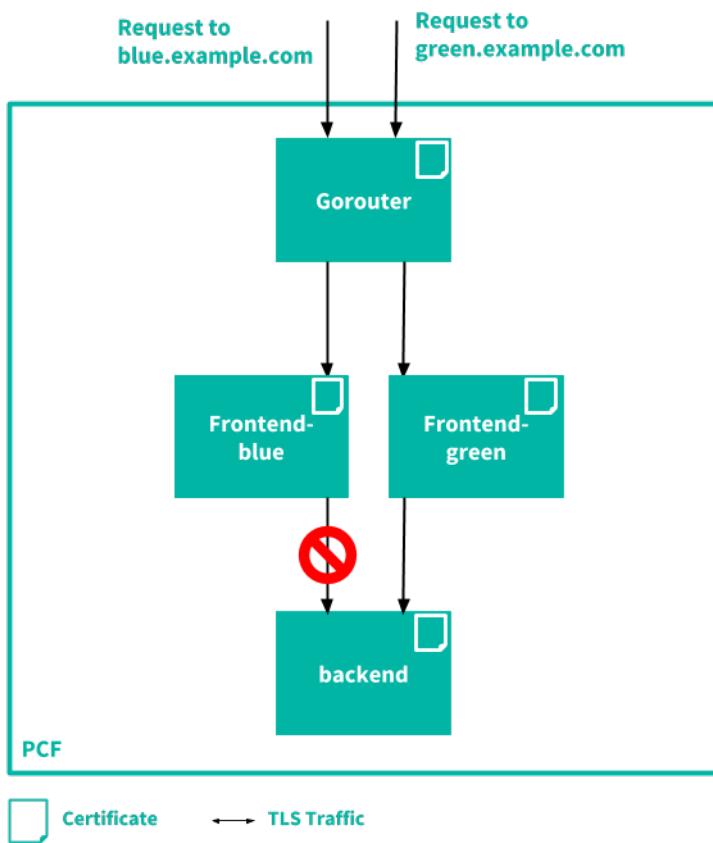
## Component Communication Details

The following topics list the paths, ports, and protocols that subsystems within Pivotal Application Service (PAS) use to communicate.

- [Cloud Controller Network Communications ↗](#)
- [Diego Network Communications ↗](#)
- [Loggregator Network Communications ↗](#)
- [MySQL Network Communications ↗](#)
- [NATS Network Communications ↗](#)
- [Routing Network Communications ↗](#)
- [UAA Network Communications ↗](#)

## Between Apps

The following diagram illustrates TLS communications between apps running on PCF. In this example, the `frontend-blue` and `frontend-green` apps both receive client requests, but only the `frontend-green` app is allowed to communicate with the `backend` app.

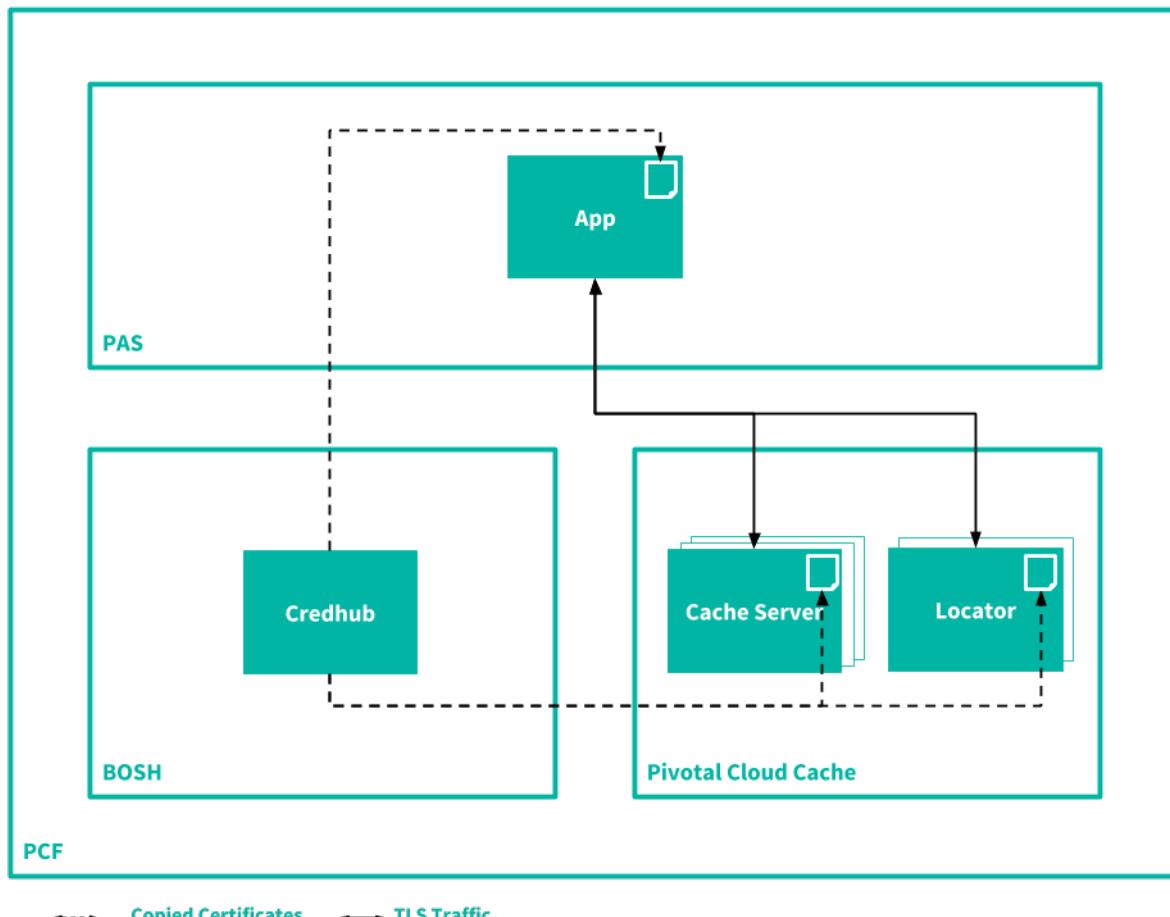


Apps can communicate with each other over TLS using certificates generated by a PCF root CA dedicated to app instance identity. For more information, see [App Instance Container Identity Credentials](#).

Developers specify which apps are allowed to communicate using container networking policies. For more information, see [Configuring Container-to-Container Networking](#).

## Between Apps and On-Platform Services

The following diagram illustrates TLS communication between apps and managed, on-platform services. It uses Pivotal Cloud Cache as an example of a managed service.

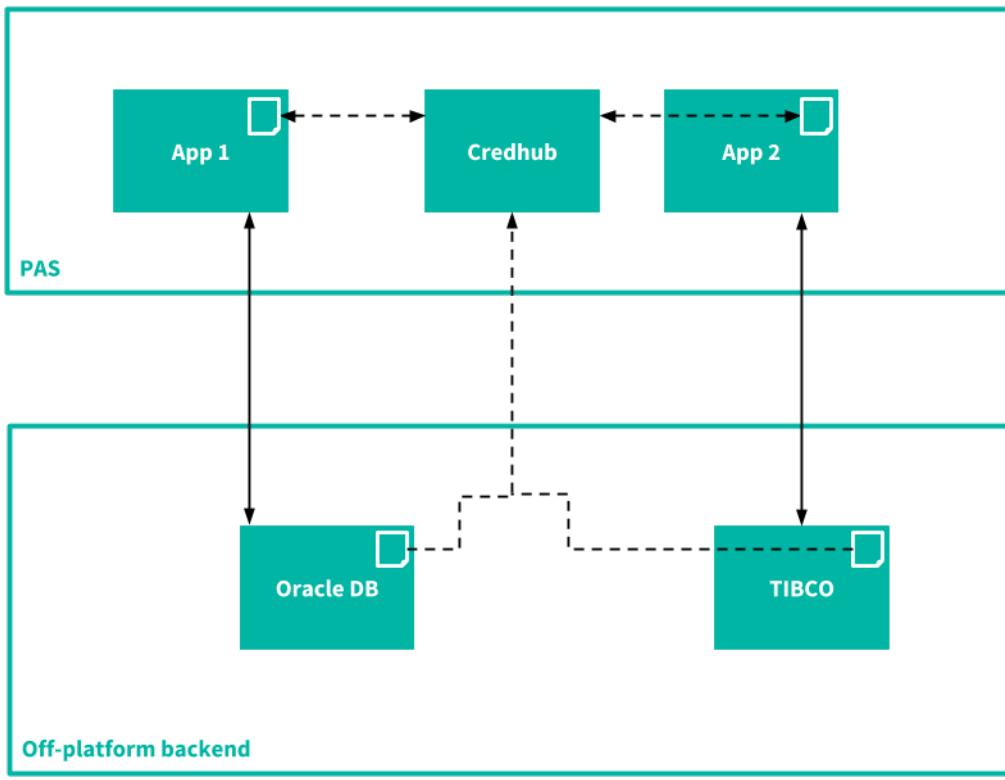


BOSH CredHub issues certificates to the Pivotal Cloud Cache components. For the app, the developer must retrieve a copy of this certificate using the CredHub API and place it in the truststore for the app. For more information, see the [Developing an App Under TLS](#) document in the Pivotal Cloud Cache documentation.

Separately, PAS Runtime CredHub might store credentials for the app to access a service over the TLS connection, adding a second layer of security. For more information, see [Securing Services Instance Credentials with Runtime CredHub](#).

## Between Apps and External Services

The following diagram illustrates communications between apps and external, brokered services secured with TLS.



The developer must retrieve the certificate from the external service and provide it to their app. One way to do this is by placing the certificate in Runtime CredHub and modifying your app to consume the certificate through the [CredHub Service Broker for PCF](#).

## App Instance Container Identity Credentials

Each app instance container in PCF has its own identity credentials. This section is meant to help PCF operators and developers [understand](#) and [use](#) these credentials.

### About App Instance Identity Credentials

See the following table to learn about app instance identity credentials.

Attribute	Description
Purpose	<ul style="list-style-type: none"> <li>For app developers to enable secure TLS communications from their apps.</li> <li>For PCF to use internally to validate the identities of app instances.</li> </ul>
Type	<ul style="list-style-type: none"> <li>A PEM-encoded <a href="#">X.509</a> certificate and <a href="#">PKCS #1 RSA</a> private key.</li> </ul>
Location	<ul style="list-style-type: none"> <li>PCF presents the certificate and private key to the app instance through the container filesystem.</li> </ul>
Properties of certificate	<ul style="list-style-type: none"> <li>The <b>Common Name</b> is the app instance GUID.</li> <li>The <b>Subject</b> of the certificate contains an <i>Organizational Unit</i> in the form of <code>app:APP-GUID</code>.</li> <li>The certificate contains a <b>Subject Alternative Name</b> (SAN) with the IP address for the app instance container.</li> <li>The certificate is valid for 24 hours after being issued.</li> </ul>
Contents of certificate file	<ul style="list-style-type: none"> <li>A chain of PEM-encoded certificates, with the instance-specific certificate first in the list and any intermediates following it.</li> </ul>
	<ul style="list-style-type: none"> <li>PCF includes a root Certificate Authority (CA) dedicated to app instance identity. This CA is saved in the system trust</li> </ul>

## Issuing authority

store for buildpack-based apps and in a file in `/etc/cf-system-certificates` in all app instance containers.

## Using the Credentials

If you want to enable secure TLS communications from an app using container instance identity credentials, ensure that you do the following:

- **Add the credentials to your development stack configuration:**
  - The credentials must be present in your development stack configuration for your app to use them. You can retrieve the credentials through following environment variables, which PCF sets to the locations of key and certificate files.
- **Reload the credential files before they expire:**
  - PCF rotates the credentials shortly before the current certificate expires. Apps that use these credentials must reload the certificate and key file contents either periodically or in reaction to filesystem watcher events.
- **Configure external clients or servers to trust the root CA:**
  - To enable secure TLS communication between an app and a client or server external to PCF, you must configure the external client or server to trust the CA that issues app instance container identity credentials. See the *Issuing Authority* row of the table in [About App Instance Identity Credentials](#).

## Additional Information

For more information about instance identity credentials, see the [Instance Identity](#) document in the diego-release repository.

## TLS Cipher Suites

By default, PCF uses a limited set of cipher suites to secure its internal communications. However, some components used in PCF, like the Gorouter and HAProxy, may support additional TLS cipher suites to accommodate older clients outside of PCF.

 The AWS Classic load balancer does not support the recommended TLS cipher suites. See [Securing Traffic into Cloud Foundry](#) for details and mitigations.

For components that allow you to configure TLS cipher suites, only specify the TLS cipher suites that you need.

### TLS Cipher Suite Recommendations

The default and recommended version of TLS to use is TLS v1.2.

The recommended TLS cipher suites to use within PCF are the following:

- `TLS_DHE_RSA_WITH_AES_128_GCM_SHA256`
- `TLS_DHE_RSA_WITH_AES_256_GCM_SHA384`
- `TLS_ECDHE_RSA_WITH_AES_128_GCM_SHA256`
- `TLS_ECDHE_RSA_WITH_AES_256_GCM_SHA384`

### Gorouter Configuration

As part of your PAS networking configuration, you must specify the TLS cipher suites that the Gorouter uses to secure its communications. Only specify the cipher suites that you need.

The recommended TLS cipher suites for the Gorouter are:

- ECDHE-RSA-AES128-GCM-SHA256
- ECDHE-RSA-AES256-GCM-SHA384

You can specify other cipher suites and a different minimum version of TLS support if your deployment requires it. For a list of other cipher suites and other versions of TLS that are optionally supported by the Gorouter, see [Securing Traffic into Cloud Foundry](#).

For instructions on how to configure the TLS cipher suites for the Gorouter, see [Configuring PAS](#).

 **Note:** In PAS v2.3, traffic between the Gorouter and Cloud Controller is encrypted. To ensure there is no downtime while upgrading from v2.2 to v2.3, download the latest patch for v2.2. This patch contains the configuration `router.backends.enable_tls: true` in the Gorouter manifest.

## HAProxy Configuration

As part of your PAS networking configuration, you must specify the TLS cipher suites that HAProxy uses to secure its communications. Only specify the cipher suites that you need.

The recommended TLS cipher suites for HAProxy are:

- DHE-RSA-AES128-GCM-SHA256
- DHE-RSA-AES256-GCM-SHA384
- ECDHE-RSA-AES128-GCM-SHA256
- ECDHE-RSA-AES256-GCM-SHA384

You can specify other cipher suites and a different minimum version of TLS support if your deployment requires it. For a list of other cipher suites and other versions of TLS that are optionally supported by HAProxy, see [ciphers - Cipher Suite Names](#) in the OpenSSL documentation.

If you use the default and recommended Gorouter TLS cipher suites in PAS, then ensure you have included these Gorouter TLS cipher suites in your HAProxy TLS cipher suite configuration.

If you change the default Gorouter TLS cipher suites in PAS, and you change the TLS cipher suites for HAProxy, ensure that you have at least one overlapping TLS cipher suite within the two sets.

For instructions on how to configure the TLS cipher suites for HAProxy, see the PAS installation documentation for the IaaS of your deployment. For example, if you are deploying PAS on GCP, see see [Configuring PAS](#).

## Custom Certificate Authorities

This topic provides an overview of using custom certificate authorities (CAs) in a Pivotal Cloud Foundry (PCF) deployment.

### Overview

To secure traffic in your PCF deployment, you must provide a CA to issue digital certificates. This can be either a Pivotal-generated or custom CA. When you add and activate a new CA, a digital certificate is issued to BOSH Director. BOSH Director then passes the certificate to other components in your PCF deployment.

Pivotal recommends you supply a CA from a trusted provider when using a production environment. While you can create your own custom CAs if necessary, a trusted CA is more secure because it has been authenticated by the trusted entities permitted to issue them.

 **Note:** Elliptic Curve Digital Signature Algorithm (ECDSA) certificates are not supported in PCF.

### Add a Custom CA

You can add a new custom CA as part of the procedure for rotating CAs and other certificate types in PCF. To add and activate a new custom CA in PCF, see [Rotate Root and Leaf Certificates](#).

## Certificate Types

This topic describes the types of certificates used in Pivotal Cloud Foundry (PCF) that require planned rotation.

## Overview

PCF uses a root Certificate Authority (CA) and various leaf certificates. Root CAs are self-signed certificates that issue leaf certificates. Root CAs can be generated by Pivotal or custom.

Leaf certificates are signed by a CA and are used to identify resources in PCF. Both root CAs and leaf certificates require planned rotation in PCF.

## Certificate Types

The following types of PCF certificates require planned rotation:

- **Ops Manager Root CA:** The Ops Manager root CA issues other certificates that PCF uses. The root CA can be a Pivotal-generated CA or your own custom CA. For more information about viewing the root CAs for Ops Manager, see [Listing the Root Certificate Authorities](#).
- **Non-configurable Certificates:** Non-configurable certificates are leaf certificates either created by a CA stored in Ops Manager, or created and stored by CredHub and managed by Ops Manager calls to the CredHub API. Non-configurable certificates are issued directly by the Ops Manager root CA, or by intermediate CAs in a chain of trust originated by the root CA. Non-configurable certificates expire after two years. For more information about viewing non-configurable leaf certificates, see [Getting Information About Certificates for Products](#). For more information about generating non-configurable leaf certificates, see [Generating New Certificates](#).
- **Configurable Certificates:** Configurable certificates are leaf certificates supplied by the user and pasted into configuration fields in Ops Manager. Some configuration panes include a **Generate RSA Certificate** button that supplies valid certificates, but users can obtain configurable certificates from elsewhere. Configurable certificates generated by Ops Manager typically expire after two years. For more information about viewing configurable leaf certificates, see [Getting Information About Certificates for Products](#).
- **Non-rotatable Certificates:** Non-rotatable certificates are leaf certificates that, like non-configurable certificates, are issued by the root CA. Unlike non-configurable certificates, non-rotatable certificates cannot be rotated by the Ops Manager API. For more information about viewing non-rotatable leaf certificates, see [Getting Information About Certificates for Products](#).

In addition to the types of certificates listed above, some Pivotal products issue their own tile certificates that are not managed by or visible to the Ops Manager API. These tile certificates do not require planned rotation because they rotate automatically with product upgrades.

Pivotal Application Service (PAS) and Pivotal Container Service (PKS) both use tile certificates in addition to their Ops Manager certificates.

## Managing Certificates with the Ops Manager API

This topic describes how to manage and retrieve information about certificates in PCF using the Ops Manager API.

### Overview

The Ops Manager API includes endpoints for managing and retrieving information about certificates in a PCF deployment.

For more information about Ops Manager API endpoints for managing certificates, see [Certificate Authorities](#) in the Ops Manager API documentation.

### Prerequisites

To use the Ops Manager API, you must generate an access token by authenticating with the Ops Manager User Account and Authentication (UAA) server.

For more information about authenticating with UAA, see [Using Ops Manager API](#).

## Generate a Single RSA Certificate

To generate and return a new RSA certificate signed by the root certificate authority (CA), use `curl` to make the following API call:

```
curl "https://OPS-MAN-FQDN/api/v0/certificates/generate \
-X POST \
-H "Authorization: Bearer YOUR-UAA-ACCESS-TOKEN"
```

Where `YOUR-UAA-ACCESS-TOKEN` is your Ops Manager access token without any newline characters such as `\n`.

## Retrieve the Ops Manager Root CA

You can view the Ops Manager root CA as a file or in JSON format.

### Retrieve the Ops Manager Root CA as a File

To return the Ops Manager root CA as a file, use `curl` to make the following API call:

```
curl "https://OPS-MAN-FQDN/download_root_ca_cert \
-X GET \
-H "Authorization: Bearer YOUR-UAA-ACCESS-TOKEN"
```

Where `YOUR-UAA-ACCESS-TOKEN` is your Ops Manager access token without any newline characters such as `\n`.

### Retrieve the Ops Manager Root CA as JSON

To return the Ops Manager root CA as JSON, use `curl` to make the following API call:

```
curl "https://OPS-MAN-FQDN/api/v0/security/root_ca_certificate \
-X GET \
-H "Authorization: Bearer YOUR-UAA-ACCESS-TOKEN"
```

Where `YOUR-UAA-ACCESS-TOKEN` is your Ops Manager access token without any newline characters such as `\n`.

## List all RSA Certificates

To return metadata from all deployed RSA certificates visible to Ops Manager, except the root CAs, use `curl` to make the following API call:

```
curl "https://OPS-MAN-FQDN/api/v0/deployed/certificates \
-X GET \
-H "Authorization: Bearer YOUR-UAA-ACCESS-TOKEN"
```

Where `YOUR-UAA-ACCESS-TOKEN` is your Ops Manager access token without any newline characters such as `\n`.

## Rotating Certificates

This topic describes how to check the expiration date of and rotate the root Certificate Authorities (CAs) and leaf certificates in Pivotal Cloud Foundry (PCF) that are visible to the Ops Manager API.

 **Note:** The rotation procedures described in this topic do not work if the certificates have already expired. If the certificates have expired, contact [Pivotal Support](#) for guidance.

For information about rotating IPsec certificates, see [Rotating IPsec Certificates](#).

For information about using trusted third-party certificates for both apps hosted on PCF and internal PCF components, see [Setting Trusted Certificates](#).

## Overview

The Ops Manager API manages and lists internal Certificate Authorities (CAs) and leaf certificates that enable PCF components to communicate with each other securely using TLS. It can also list certificates used externally, such as SAML certificates that authenticate to an external identity provider.

For more information about root CAs and leaf certificates, see [Certificate Types](#).

Rotate root CAs and leaf certificates before they expire to avoid downtime for your foundation.

## Procedure

This section provides an overview of the procedure for rotating the root CA and leaf certificates in PCF.

To rotate certificates in PCF, you first check the expiration dates of all certificates. Then, based on the types of certificates that expire soon, you follow a certificate rotation procedure to replace expiring certificates and redeploy BOSH to apply changes.

To rotate root CAs and leaf certificates in PCF, do the following:

1. Check the expiration dates of the root CA and leaf certificates, and identify the types of leaf certificates that require rotation. See [Check Expiration Dates and Certificate Types](#).
2. Rotate the certificates that expire soon. See [Rotate Certificates](#).

## Check Expiration Dates and Certificate Types

This section describes how to check the expiration dates of the root CA and leaf certificates that the Ops Manager API lists and manages. It also explains how to identify the types of leaf certificates that expire soon and require manual rotation.

After identifying the types of certificates that expire soon, you can determine which certificate rotation procedure to follow.

### Check Root CA Expiration Date

This procedure describes how to check the expiration date for the Ops Manager root CA. The Ops Manager root CA expires four years after creation.

To check the Ops Manager root CA expiration date, do the following:

1. Perform the steps in the [Using Ops Manager API](#) topic to target and authenticate with the Ops Manager User Account and Authentication (UAA) server. Record your Ops Manager access token, and use it for `YOUR-UAA-ACCESS-TOKEN` in the steps below.

 **Note:** When you record your Ops Manager access token, remove any newline characters such as `\n`.
2. To retrieve the Ops Manager root CA, use `curl` to make an Ops Manager API call to the `https://OPS-MAN-FQDN/api/v0/certificateAuthorities` endpoint. For example:

```
curl "https://OPS-MAN-FQDN/api/v0/certificateAuthorities" \
-H "Authorization: Bearer YOUR-UAA-ACCESS-TOKEN"
```

3. To make the JSON output more readable, you can pipe it to `jq` or another text editor with JSON formatting.
4. In the `certificateAuthorities` list returned, if there is more than one, find the CA with `"active": true`.
5. To determine its expiration date of the active CA, refer to its `expiresOn` value. For example, the root CA shown below expires on September 5, 2019:

```
{
 "certificateAuthorities": [
 {
 "guid": "9c9a110c8f82a1e4aaca",
 "issuer": "Pivotal",
 "createdOn": "2017-09-05T22:47:53Z",
 "expiresOn": "2019-09-05T22:47:53Z",
 "active": true,
 "certPem": "-----BEGIN CERTIFICATE-----\n[...]\n-----END CERTIFICATE-----\n"
```

## Check Leaf Certificate Expiration Dates and Types

This procedure describes how to check the expiration dates of leaf certificates using the `deployed/certificates` endpoint. It also describes how to identify non-configurable, configurable, and non-rotatable leaf certificates in the output from the `deployed/certificates` endpoint.

To check the expiration dates of leaf certificates and identify the types of leaf certificates that expire soon, do the following:

1. If you haven't already, perform the steps in the [Using Ops Manager API](#) topic to target and authenticate with the Ops Manager User Account and Authentication (UAA) server. Record your Ops Manager access token, and use it for `YOUR-UAA-ACCESS-TOKEN` in the steps below.

 **Note:** When you record your Ops Manager access token, remove any newline characters such as `\n`.

2. To check the system for certificates that expire within a given time interval, use `curl` to call the `https://OPS-MAN-FQDN/api/v0/deployed/certificates?expiresWithin=TIME` API endpoint, replacing `TIME` with an integer-letter code. Valid letter codes are `d` for days, `w` for weeks, `m` for months, and `y` for years.

For example, the following command searches for certificates expiring within six months:

```
curl "https://OPS-MAN-FQDN/api/v0/deployed/certificates?expiresWithin=6m" \
-H "Authorization: Bearer YOUR-UAA-ACCESS-TOKEN"
```

Where `YOUR-UAA-ACCESS-TOKEN` is the `accessToken` value you recorded in the previous step.

3. To make the JSON output more readable, pipe it to `jq` or another text editor with JSON formatting.
4. To determine the expiration date of each certificate, refer to its `expiresOn` value.
5. Identify the types of leaf certificates requiring rotation. The following rules identify the type of each leaf certificate:
  - o **Non-Rotatable Certificates:** Non-Rotatable leaf certificates have the following property value: `variablePath` is `/opsmgr/bosh_dns/tls_ca`
  - o **Non-Configurable Certificates:** Non-Configurable leaf certificates have the following property values and are not non-rotatable as identified above:
    - `configurable` is `false`
    - `location` is either `ops_manager` or `credhub`
  - o **Configurable Certificates:** Configurable leaf certificates have the following property value: `configurable` is `true`

## Rotate Certificates

This section includes procedures for rotating root CAs and leaf certificates. There are different rotation procedures for each type of certificate that requires rotation.

You can determine which rotation procedure to follow based on the types of certificates in your foundation that must be rotated. For more information about identifying the types of certificates in your foundation that expire soon, see [Check Expiration Dates and Certificate Types](#).

To rotate certificates, do one of the following procedures:

- To rotate non-rotatable certificates, call [Pivotal Support](#).
- To rotate the root CA, follow the [Rotate the Root CA and Leaf Certificates](#) procedure. This procedure automatically rotates all non-configurable leaf certificates. You can also optionally rotate configurable leaf certificates as part of this procedure.
- To rotate non-configurable leaf certificates, but not the root CA, follow the [Rotate Non-Configurable Leaf Certificates](#) procedure.
- To rotate configurable leaf certificates, but not the root CA, follow the [Rotate Configurable Leaf Certificates](#) procedure.

## Rotate the Root CA and Leaf Certificates

This procedure uses the Ops Manager API to rotate the Ops Manager root CA and non-configurable leaf certificates. Rotating the Ops Manager root CA automatically rotates all non-configurable leaf certificates.

To prevent system downtime, this procedure includes two BOSH deploys. The first BOSH deploy applies new certificates to PCF components. The second BOSH deploy deletes the old certificates.

**A** **warning:** You must complete these steps in the exact order specified. Otherwise, you may damage your deployment.

To rotate the root CA and leaf certificates, do the following:

1. Add a new root CA. See [Add a New Root CA](#).
2. Activate the root CA. [Activate the Root CA](#).
3. (Optional) Rotate leaf configurable certificates. See [Rotate Configurable Leaf Certificates](#).
4. Rotate non-configurable leaf certificates. See [Rotate Non-Configurable Leaf Certificates from the New Root](#).
5. (Optional) Delete the old CA. See [Delete the Old CA](#).

### Step 1: Add a New Root CA

Follow this procedure to add a new root CA for Ops Manager. The new root CA can be a Pivotal-generated CA or your own custom CA.

**!** **Note:** This procedure regenerates the NATS CA in unison with the Ops Manager root CA.

To add a new root CA for Ops Manager, do the following:

1. If you haven't already, perform the steps in the [Using Ops Manager API](#) topic to target and authenticate with the Ops Manager User Account and Authentication (UAA) server. Record your Ops Manager access token, and use it for `YOUR-UAA-ACCESS-TOKEN` in the following steps.

**!** **Note:** When you record your Ops Manager access token, remove any newline characters such as `\n`.

2. Use Ops Manager to generate a new CA, or else add your own custom CA:

**!** **Note:** Elliptic Curve Digital Signature Algorithm (ECDSA) certificates are not supported in PCF.

- To use your own custom CA, call the Ops Manager API `generate` endpoint as follows:

```
curl "https://OPS-MAN-FQDN/api/v0/certificateAuthorities" \
-X POST \
-H "Authorization: Bearer YOUR-UAA-ACCESS-TOKEN" \
-H "Content-Type: application/json" \
-d '{"cert_pem": "-----BEGIN CERTIFICATE-----\YOUR-CERTIFICATE", "private_key_pem": "-----BEGIN RSA PRIVATE KEY-----\YOUR-KEY"}'
```

Where:

- `YOUR-CERTIFICATE` is your custom CA.
- `YOUR-KEY` is your RSA key.
- `YOUR-UAA-ACCESS-TOKEN` is your UAA access token.

If the command succeeds, the API returns a response that includes the new CA certificate:

```
HTTP/1.1 200 OK
{
 "certificateAuthorities": [
 {
 "guid": "f7bc18f34f2a7a9403c3",
 "issuer": "YOUR-CA",
 "createdOn": "2017-01-09",
 "expiresOn": "2021-01-09",
 "active": true,
 "certPem": "-----BEGIN CERTIFICATE-----\nMIIC+zCC AeOgAwIBAgI....etc"
 }
]
}
```

- To use a Pivotal-generated CA, call the Ops Manager API `generate` endpoint as follows:

```
curl "https://OPS-MAN-FQDN/api/v0/certificateAuthorities/generate" \
-X POST \
-H "Authorization: Bearer YOUR-UAA-ACCESS-TOKEN" \
-H "Content-Type: application/json" \
-d '{}'
```

Where `YOUR-UAA-ACCESS-TOKEN` is your UAA access token.

If the command succeeds, the API returns a response that includes the new CA certificate:

```
HTTP/1.1 200 OK
{
 "guid": "f7bc18f34f2a7a9403c3",
 "issuer": "Pivotal",
 "createdOn": "2017-01-19",
 "expiresOn": "2021-01-19",
 "active": false,
 "certPem": "-----BEGIN EXAMPLE CERTIFICATE-----\nMIIC+zCC AeOgAwIBAgIBADANBgkqhkiG9w0BAQsFADAfMQswCQYDVQQGEwJVUzEQ\nEXAMPEoCgwHUG12b3RhbaEfw0xNDarthyMTQyMjVaFw0yMTAxMTkyMTQyMjVa\nEXAMPEoBgnVBAYTAIVTMRAwDgYDVVaderdQaXZvdGFsMIIbjANBgkqhkiG9w0B\nEXAMPEoAQ8AMIIbCgKCAQEayV4OhPIZTEym9OcdcNVip9Ev0ijPPLo9WPLUMzT\nEXAMPEo/TgD+DP09mwVXfqwBJm0j9DqRED1x/6bc0K/BFAo/P4MmOKm3QnDCt\nEXAMPEoogA++2HYrNTKWJ5fXmERs8IK9AXXT7RKXhktvWU3oNGf7zo0e3YKp\nEXAMPEoh1NwlbNcGT1AurIDxyOZyIHVzBLTisMyDogJmSCLsOw3qUDQjatjXKw\nEXAMPEoqG3nv2hvD4/aTOiHuKM3+AGbnaS2MdIoVFOh/7Y79iUp89csK0gsuOd\nEXAMPEoe4DcKw5CzUTfhKNXgHyeoVOBPCvQTp4JpliQIDAQABo0lwQDAdBgNV\nEXAMPEoyH4y7VEuImLStXM0CKR8uVqxX/gwDwYDVR0TAQH/BAUwAwEB/zAOBgNV\nEXAMPEoBAMCAQYwDQYJKoZIhvCNQELBQAAdggEBALmHOPxdyBGnuROHgR9V4TwJ\nEXAMPEoGLKV7am5z6G2Oq5ewACFHWAFrPG4W9Jm577QtewiY/Rad/PbkY0YSY\nEXAMPEokrfNjxjxI0H2sr7qLBfjJ0wBZhVmDsO6A9PkfAPu4eJvqRMuL/xGmSQ\nEXAMPEoCynMnz7FgHyFbd9D9X5YW8fWGSeVBPPikcONDrvjw9aEcAtbGEh8eZCP\nEXAMPEob33RuR+CTNqThXY9k8d7/7ba4KVdd4g8synFgwvnDQOjcJZ6Go5QY5HA\nEXAMPEoPFW8pAYcyWrXKR0rE8fL5o9qTyjmO+5yyvWIYrKPqqIUIvMCdNr84=\n-----END EXAMPLE CERTIFICATE-----\n"
```

- Confirm that your new CA has been added by listing all of the root CAs for Ops Manager:

```
curl "https://OPS-MAN-FQDN/api/v0/certificateAuthorities" \
-X GET \
-H "Authorization: Bearer YOUR-UAA-ACCESS-TOKEN"
```

The API call returns the GUID of the newly added CA. For example:

```
HTTP/1.1 200 OK
{
 "certificateAuthorities": [
 {
 "guid": "f7bc18f34f2a7a9403c3",
 "issuer": "Pivotal",
 "createdOn": "2017-01-09",
 "expiresOn": "2021-01-09",
 "active": true,
 "certPem": "-----BEGIN CERTIFICATE-----\nMIIC+zCC AeOgAwIBAgI....etc"
 },
 {
 "guid": "a8ee01e33e3e3e3303e3",
 "issuer": "Pivotal",
 "createdOn": "2017-04-09",
 "expiresOn": "2021-04-09",
 "active": false,
 "certPem": "-----BEGIN CERTIFICATE-----\zBBC +eAAAelgAwAAAeZ....etc"
 }
]
}
```

Identify your newly added CA, which has `active` set to `false`. Record its GUID.

4. Navigate to `https://OPS-MAN-FQDN` in a browser and log in to Ops Manager.
5. Click the **BOSH Director** tile in the **Installation Dashboard**.
6. Select the **Director Config** pane.
7. Select **Recreate All VMs**. This propagates the new CA to all VMs to prevent downtime.
8. Go back to the Installation Dashboard. For each service tile you have installed, do the following:
  - a. Click the tile.
  - b. Click the **Errands** tab.
  - c. Enable the **Recreate All Service Instances** errand if provided. If you do not see this errand, contact [Pivotal Support](#).
9. Click **Review Pending Changes**, then **Apply Changes**. When the deploy finishes, continue to the next section.

## Step 2: Activate the Root CA

To activate the new root CA, do the following:

1. Use `curl` to make an Ops Manager API call that activates the new CA:

```
curl "https://OPS-MAN-FQDN/api/v0/certificateAuthorities/CERT-GUID/activate" \
-X POST \
-H "Authorization: Bearer YOUR-UAA-ACCESS-TOKEN" \
-H "Content-Type: application/json" \
-d '{}'
```

Where `CERT-GUID` is the GUID of your CA that you retrieved in the previous section.

The API returns a successful response:

```
HTTP/1.1 200 OK
```

2. List your root CAs again to confirm that the new CA is active:

```
curl "https://OPS-MAN-FQDN/api/v0/certificateAuthorities" \
-X GET \
-H "Authorization: Bearer YOUR-UAA-ACCESS-TOKEN"
```

Examine the response to ensure that your new CA has `active` set to `true`.

## (Optional) Step 3: Rotate Configurable Leaf Certificates

If there are any configurable certificates in your foundation expiring soon, Pivotal recommends rotating them now. Rotating configurable leaf certificates now ensures that your next BOSH deploy includes both new configurable and non-configurable leaf certificates.

To rotate configurable leaf certificates, do the following for each configurable certificate that expires soon:

1. Find the text field where the certificate is configured within the Ops Manager interface. You might have to look through multiple configuration panes to identify the location of the certificate configuration text field. Use the following fields to identify the location of the certificate configuration text field:

- o The `product_guid` field in the API output can help identify in which tile the certificate is configured. For example, the prefix `p-bosh-` refers to the BOSH Director tile, and the prefix `cf-` refers to the PAS tile.
- o The `property_reference` field in the API output can help identify in which **Settings** pane the certificate is configured. For example, the `uaa.service_provider_key_credentials` property is configured in the PAS tile > UAA pane.

2. Paste a new value for the certificate into the text field.

3. Click **Save** at the bottom of each pane in which you added new certificates.

#### Step 4: Rotate Non-Configurable Leaf Certificates from the New Root

After activating the new root CA, you must rotate non-configurable leaf certificates from the root CA.

To rotate non-configurable leaf certificates from the new root CA, do the following:

1. Use `curl` to make an API call to regenerate all non-configurable certificates and apply the new CA to your existing BOSH Director:

```
curl "https://OPS-MAN-FQDN/api/v0/certificateAuthorities/active/regenerate" \
-X POST \
-H "Authorization: Bearer YOUR-UAA-ACCESS-TOKEN" \
-H "Content-Type: application/json" \
-d '{}'
```

The API returns a successful response:

```
HTTP/1.1 200 OK
```

2. Click the **BOSH Director** tile in the **Installation Dashboard**.

3. Select the **Director Config** pane.

4. Select **Recreate All VMs**. This propagates the new CA to all VMs to prevent downtime.

5. Go back to the Installation Dashboard. For each service tile you have installed, do the following:

- a. Click the tile.
- b. Click the **Errands** tab.
- c. Enable the **Recreate All Service Instances** errand if provided. If you do not see this errand, contact [Pivotal Support](#).

6. Navigate to Ops Manager, click **Review Pending Changes**, and click **Apply Changes** to perform a second redeploy.

#### (Optional) Step 5: Delete the Old CA

After activating a new root CA and rotating leaf certificates from the new root CA, Pivotal recommends deleting the old root CA from your foundation.

**⚠ warning:** Include the **Review Pending Changes** and **Apply Changes** from [Step 4: Rotate Non-Configurable Leaf Certificates from the New Root](#) before you delete the old CA.

To delete the old root CA, do the following:

1. List your root CAs to retrieve the GUID of your old, inactive CA:

```
curl "https://OPS-MAN-FQDN/api/v0/certificateAuthorities" \
-X GET \
-H "Authorization: Bearer YOUR-UAA-ACCESS-TOKEN"
```

2. Use `curl` to make an API call to delete your old CA:

```
curl "https://OPS-MAN-FQDN/api/v0/certificateAuthorities/OLD-CERT-GUID" \
-X DELETE \
-H "Authorization: Bearer YOUR-UAA-ACCESS-TOKEN"
```

Where `OLD-CERT-GUID` is the GUID of your old, inactive CA.

The API returns a successful response.

```
HTTP/1.1 200 OK
```

3. Navigate to Ops Manager, click **Review Pending Changes**, and click **Apply Changes**.

## Rotate Non-Configurable Leaf Certificates

This procedure rotates non-configurable leaf certificates visible to the Ops Manager API, whether they are managed and stored by Ops Manager directly, or by CredHub at Ops Manager request.

**⚠ Warning:** This procedure does not rotate the Ops Manager root CA. To rotate both the root CA and non-configurable leaf certificates, see [Rotate the Root CA and Leaf Certificates](#).

To rotate non-configurable leaf certificates, do the following:

1. Use `curl` to make an API call to regenerate all non-configurable certificates and apply the new CA to your existing BOSH Director:

```
curl "https://OPS-MAN-FQDN/api/v0/certificateAuthorities/active/regenerate" \
-X POST \
-H "Authorization: Bearer YOUR-UAA-ACCESS-TOKEN" \
-H "Content-Type: application/json" \
-d '{}'
```

The API returns a successful response:

```
HTTP/1.1 200 OK
```

2. Navigate to Ops Manager, click **Review Pending Changes**, and click **Apply Changes** to perform a second redeploy.

## Rotate Configurable Leaf Certificates

Configurable certificates are generated by the user and pasted into Ops Manager configuration panes where needed. Examples include certificates that terminate SSL traffic into PAS, or authenticate a Single Sign-On (SSO) service plan to an external SAML server.

For instructions on how to rotate SAML certificates for both PAS and the SSO service, see [Rotate Identity Provider SAML Certificates](#).

**⚠ Warning:** This procedure does not rotate the Ops Manager root CA. To rotate both the root CA and non-configurable leaf certificates, see [\[Rotate the Root CA and Leaf Certificates\]\(#rotate-ca\)](#).

To rotate configurable leaf certificates, do the following:

1. If you haven't already, use the Ops Manager API `/deployed/certificates` endpoint to retrieve information about your expiring configurable certificates, as described in the procedures [Check Expiration Dates and Certificate Types](#).

The information for each configurable certificate looks like this:

```
{
 "configurable": true,
 "property_reference": ".properties.networking_poe_ssl_certs[0].certificate",
 "property_type": "rsa_cert_credentials",
 "product_guid": "cf-36066831c119c39736d3",
 ...
 "valid_until": "2019-01-25T22:07:58Z"
},
```

2. For each configurable certificate that expires soon:

- a. Find the text field the certificate is configured within the Ops Manager interface.

- The `product_guid` field in the API output can help identify which tile the certificate is configured in. For example, the prefix `p-bosh-` refers to the BOSH Director tile, and the prefix `cf-` refers to the PAS tile.
- The `property_reference` field in the API output can often help identify which **Settings** pane the certificate is configured in. For example, the `uaa.service_provider_key_credentials` property is configured in the PAS tile > **UAA** pane.
- You might have to look through multiple configuration panes to identify where a certificate is configured.

- b. Paste a new value for the certificate into the field
  - c. Click **Save** at the bottom of each pane in which you have provided new certificates.
3. If you are rotating configurable certificates within the [Rotate Root and Leaf Certificates](#) procedure, continue to the next step. Otherwise, if you are rotating configurable leaf certificates only, return to the [Installation Dashboard](#), click **Review Pending Changes**, and click **Apply Changes**.

## Rotate Identity Provider SAML Certificates

SAML service provider credentials are one example of configurable certificates in PCF. When PAS is configured to use SAML as an identity provider, it uses a configurable CA certificate to authenticate to an external SAML server, by generating ephemeral certificates that PAS includes in its outbound request message headers. This CA has a two-year expiration period.

In addition, the [SSO](#) service shares the use of PAS SAML certificates for every SAML external Identity Provider (IdP) integration, such as trust, partnership, or Federation. You must rotate these in lockstep with PAS.

The following procedure provides an example of how to rotate certificates for each IdP, including temporarily disabling certificate validation on the IdP side during the rotation.

The Knowledge Base article [PCF Advisory - SAML Service Provider Credential Certificates Expire after 2 Years](#) provides more information about rotating SAML certificates.

SAML service provider credentials are only required for your PAS deployment if all of these conditions are met:

- You are using SSO in production for login to PAS or using the SSO service for login to apps.
- You are using SAML identity providers for PAS or SSO service plans.
- You had Ops Manager generate a certificate for you by clicking the **Generate RSA Certificate** button.
- You are validating the signature of SAML authentication request with your identity provider.

To regenerate and rotate SAML service provider certificates without disrupting PAS or your apps using the SSO service, do the following:

1. Disable certificate validation in your identity provider.
2. For PAS, follow the procedure in the table below that corresponds to your use case. This includes downloading and importing a new certificate and updated SAML metadata in your identity provider.

Solution Name	Procedure
<a href="#">CA Single Sign-On aka CA SiteMinder</a>	<a href="#">Configuring CA as an Identity Provider</a>
<a href="#">PingFederate</a>	<a href="#">Configuring PingFederate as an Identity Provider</a>
<a href="#">Active Directory Federation Services</a>	<a href="#">Configuring ADFS as an Identity Provider</a>

3. For the SSO service, follow the procedure in the table below that corresponds to your use case. This includes downloading the SAML Service Provider metadata for each SAML identity provider integration, such as trust, partnership, or Federation, and importing the updated SAML Service Provider metadata in your identity provider.

Solution Name	Procedure
ADFS	<a href="#">Configuring a Single Sign-On Service Provider</a>
CA SSO	<a href="#">Configuring a Single Sign-On Service Provider</a>
Okta	<a href="#">Configure Okta as an Identity Provider</a>
PingFederate	<a href="#">Configure PingFederate as an Identity Provider</a>
Additional Documentation	<a href="#">Integration Guides</a>

4. Re-enable certificate validation in your identity provider.

## Trusted System Certificates

A Cloud Foundry Administrator can deploy a set of trusted system certificates. These trusted certificates are available in Linux-based app instances running on the Diego backend. Such instances include buildpack-based apps using the `cflinuxfs3` stack and Docker-image-based apps.

If the administrator configures these certificates, they are available inside the instance containers as files with extension `.crt` in the read-only `/etc/cf-system-certificates` directory.

For `cflinuxfs3`-based apps, these certificates are also installed directly in the `/etc/ssl/certs` directory, and are available automatically to libraries such as `openssl` that respect that trust store. If the administrator configure these certificates, the location of the certificates is provided in the environment variable `CF_SYSTEM_CERT_PATH` on the instance container.

## Disk Encryption

This topic describes how to secure Pivotal Cloud Foundry (PCF) VMs by encrypting their disks or rotating their disk encryption keys.

## Introduction

Disk encryption protects data integrity if computing resources are stolen physically.

Disk encryption for VMs works at the IaaS level. An IaaS encrypts disks when it first creates them, or re-encrypts them when it rotates encryption keys. To encrypt disks in PCF, you must:

1. Configure the IaaS to encrypt disks when it creates or re-creates them.
2. Trigger BOSH to re-create the existing VMs that use the disks, and create encrypted disks from now on for new VMs.

The procedures below detail how to do this for each IaaS.

### Disks You Can Encrypt on a PCF VM

- The root file system for the VM. For BOSH-created VMs, this comes from the stemcell.
- Ephemeral disk for the VM.
- Persistent disk for the VM.

### Which VMs Each Procedure Encrypts

For each IaaS, there are two disk encryption procedures, which encrypt different VMs:

- The **BOSH Director** procedure encrypts the disks used by the BOSH Director VM when you first create a PCF environment.
- The **BOSH-Deployed VM** procedure encrypts disks for the VMs that the BOSH Director creates, after BOSH has been deployed.

## Encrypt Disks or Rotate Keys

You can use the same procedure to either encrypt disks for the first time or rotate encryption keys.

For BOSH-deployed VMs, some IaaSes let you associate a policy with the BOSH process that automatically encrypts all disks BOSH creates. On AWS, BOSH must explicitly tell the IaaS to encrypt each disk that it creates, and passes in an encryption key. The following table summarizes these differences:

IaaS	How configured	How encrypted	User can supply key	BOSH stores key ID
AWS	User pastes key ARN (ID) into Ops Manager	BOSH tells IaaS to encrypt disks it creates	Yes	Yes
Azure (with Managed Disks)	User configures IaaS to associate encrypt policy with BOSH	IaaS automatically encrypts disks it creates for BOSH	No	No
Azure (with Storage Accounts)	User configures IaaS to associate encrypt policy with BOSH	IaaS automatically encrypts disks it creates for BOSH	Yes	No
vSphere	User configures IaaS to associate encrypt policy with BOSH	IaaS automatically encrypts disks it creates for BOSH	Yes	No

### Azure

Azure provides virtual disk space through [Azure Storage](#) accounts. In some regions, Azure offers a [Managed Disks](#) service for storage accounts, which allocates disk space flexibly on demand.

## Managed Disks vs Unmanaged Storage Accounts

For disk encryption, Pivotal recommends Managed Disks storage where available. With Managed Disks, encryption keys are managed by the IaaS, so you need not (and cannot) supply your own keys. You also do not need to re-create VMs after encrypting disks or rotating encryption keys, because the IaaS propagates the change to all VMs automatically.

## Encrypt Azure Disks

Follow the steps below to initiate or rotate disk encryption for BOSH-deployed VMs on Azure:

1. Log in to Azure Portal and follow the [Encryption workflow](#) to encrypt new and existing PCF VMs.
2. For unmanaged Storage Account disks, follow the [Recreate BOSH-Deployed Disks](#) procedure below to propagate the change to existing VMs.
  - o With Managed Disks, you can skip this step.

For more information about how BOSH integrates with IaaS-level disk encryption on Azure, see [Encryption](#) under *Microsoft Azure* in the BOSH documentation.

## vSphere v6.5 or Later

vSphere supports disk encryption in versions 6.5 and later for encrypted VMs. Follow the steps below to initiate or rotate disk encryption for BOSH-deployed VMs on vSphere v6.5+:

1. Log in to vCenter and follow the [Encrypt an Existing Virtual Machine or Virtual Disk](#) procedure in the *VMware Docs*.
2. Follow the [Recreate BOSH-Deployed Disks](#) procedure below to propagate the change to existing VMs.

For more information about how BOSH integrates with IaaS-level disk encryption on vSphere, see [Encryption](#) under *vSphere* in the BOSH documentation.

## AWS (Ops Manager v2.0 and later)

On AWS, you can use your Amazon account key to encrypt Linux EBS volumes, or supply your own key.

For instructions on how to encrypt BOSH-deployed VMs and the Ops Manager VM on AWS, see [Configuring Amazon EBS Encryption](#).

For more information about how BOSH integrates with IaaS-level disk encryption on AWS, see [Encryption](#) under *Amazon Web Services* in the BOSH documentation.

## Recreate BOSH-Deployed Disks

Unless you are using Azure Managed Disks, you need to manually recreate disks on BOSH-deployed VMs after you have added or rotated disk encryption keys. To manually recreate disks, do the following:

1. Configure Ops Manager to encrypt VM root, ephemeral disk, and persistent disk on next deploy:
  - o **Root File System:** To recreate the root file system for VMs, you must upload a new stemcell. If you are already running the latest stemcell, you can:
    - Wait until a new stemcell comes out, typically less than two weeks.
    - Call [Pivotal Support](#) if propagating disk encryption is urgent.
  - o **Ephemeral Disks:** In the **Director Config** pane of the Ops Manager tile, enable the **Recreate All VMs** checkbox.
  - o **Persistent Disks**
    - PCF v2.3 and later: In the **Director Config** pane of the Ops Manager tile, enable the **Recreate All Persistent Disks** checkbox.
    - PCF v2.2 and earlier: In the **Resource Config** pane of all tiles, change the disk or VM sizes of all VMs that you need to encrypt.
2. Click **Review Pending Changes** and **Apply Changes**.



## Security Event Logging

This topic describes how to enable and interpret security event logging for the Cloud Controller, the User Account and Authentication (UAA) server, and CredHub. Operators can use these logs to retrieve information about a subset of requests to the Cloud Controller, UAA server, and CredHub for security or compliance purposes.

## Cloud Controller Logging

The Cloud Controller logs security events to syslog. You must configure a syslog drain to forward your system logs to a log management service.

See the [Configuring System Logging in PAS](#) topic for more information.

### Format for Log Entries

Cloud Controller logs security events in the [Common Event Format](#) (CEF). CEF specifies the following format for log entries:

```
CEF:Version|Device Vendor|Device Product|Device Version|Signature ID|Name|Severity|Extension
```

Entries in the Cloud Controller log use the following format:

```
CEF:CEF_VERSION|cloud_foundry|cloud_controller_ng|CC_API_VERSION|
SIGNATURE_ID|NAME|SEVERITY|rt=TIMESTAMP user=USERNAME uid=USER_GUID
cs1Label=userAuthenticationMechanism cs1=AUTH_MECHANISM
cs2Label=vcapRequestId cs2=VCAP_REQUEST_ID request=REQUEST
requestMethod=REQUEST_METHOD cs3Label=result cs3=RESULT
cs4Label=httpStatusCode cs4=HTTP_STATUS_CODE src=SOURCE_ADDRESS
dst=DESTINATION_ADDRESS cs5Label=xForwardedFor cs5=X_FORWARDED_FOR_HEADER
```

Refer to the following list for a description of the properties shown in the Cloud Controller log format:

- `CEF_VERSION` : The version of CEF used in the logs.
- `CC_API_VERSION` : The current Cloud Controller API version.
- `SIGNATURE_ID` : The method and path of the request. For example, `GET /v2/app:GUID`.
- `NAME` : The same as `SIGNATURE_ID`.
- `SEVERITY` : An integer that reflects the importance of the event.
- `TIMESTAMP` : The number of milliseconds since the Unix epoch.
- `USERNAME` : The name of the user who originated the request.
- `USER_GUID` : The GUID of the user who originated the request.
- `AUTH_MECHANISM` : The user authentication mechanism. This can be `oauth-access-token`, `basic-auth`, or `no-auth`.
- `VCAP_REQUEST_ID` : The VCAP request ID of the request.
- `REQUEST` : The request path and parameters. For example, `/v2/info?MY-PARAM=VALUE`.
- `REQUEST_METHOD` . The method of the request. For example, `GET`.
- `RESULT` : The meaning of the HTTP status code of the response. For example, `success`.
- `HTTP_STATUS_CODE` . The HTTP status code of the response. For example, `200`.
- `SOURCE_ADDRESS` : The IP address of the client who originated the request.
- `DESTINATION_ADDRESS` : The IP address of the Cloud Controller VM.
- `X_FORWARDED_FOR_HEADER` : The contents of the X-Forwarded-For header of the request. This is empty if the header is not present.

### Example Log Entries

The following list provides several example requests with the corresponding Cloud Controller log entries.

- An anonymous `GET` request:

```
CEF:0|cloud_foundry|cloud_controller_ng|2.54.0|GET /v2/info|GET
/v2/info|0|rt=1460690037402 suser= suid= request=/v2/info
requestMethod=GET src=127.0.0.1 dst=192.0.2.1
cs1Label=userAuthenticationMechanism cs1=no-auth cs2Label=vcapRequestId
cs2=c4bac383-7cc9-4d9f-b1c0-1iq8e0baa000 cs3Label=result cs3=success
cs4Label=httpStatusCode cs4=200 cs5Label=xForwardedFor
cs5=198.51.100.1
```

- A `GET` request with basic authentication:

```
CEF:0|cloud_foundry|cloud_controller_ng|2.54.0|GET /v2/syslog_drain_urls|GET
/v2/syslog_drain_urls|0|rt=1460690165743 suser=bulk_api suid=
request=/v2/syslog_drain_urls?batch_size=1000 requestMethod=GET
src=127.0.0.1 dst=192.0.2.1 cs1Label=userAuthenticationMechanism
cs1=basic-auth cs2Label=vcapRequestId cs2=79187189-e810-33dd-6911-b5d015bbc999
::eat1234d-4004-4622-ad11-9iaai88e3ae9 cs3Label=result cs3=success
cs4Label=httpStatusCode cs4=200 cs5Label=xForwardedFor cs5=198.51.100.1
```

- A `GET` request with OAuth access token authentication:

```
CEF:0|cloud_foundry|cloud_controller_ng|2.54.0|GET /v2/routes|GET
/v2/routes|0|rt=1460689904925 suser=admin suid=c7ca208f-8a9e-4aab-
92f5-28795f86d62a request=/v2/routes?inline-relations-depth=1&q=
host%3Adora%3Bdomain_guid%3B777-1o9f-f55n-i888-o2025cb2dfc3
requestMethod=GET src=127.0.0.1 dst=192.0.2.1
cs1Label=userAuthenticationMechanism cs1=oauth-access-token
cs2Label=vcapRequestId cs2=79187189-990i-8930-52b2-9090b2c5poz
::5a265621-b223-4520-afae-ab7d0ee7c75b cs3Label=result cs3=success
cs4Label=httpStatusCode cs4=200 cs5Label=xForwardedFor cs5=198.51.100.1
```

- A `GET` request that results in a 404 error:

```
CEF:0|cloud_foundry|cloud_controller_ng|2.54.0|GET /v2/apps/7f310103-
39aa-4a8c-b92a-9ff8a6a2fa6b|GET /v2/apps/7f310103-39aa-4a8c-b92a-
9ff8a6a2fa6b|0|rt=1460691002394 suser=bob suid=a00i2026-55io-398s-
555o-40e611410aec request=/v2/apps/7f310103-39aa-4a8c-b92a-9ff8a6a2fa6b
requestMethod=GET src=127.0.0.1 dst=192.0.2.1
cs1Label=userAuthenticationMechanism cs1=oauth-access-token cs2Label=vcapRequestId
cs2=49f21579-9eb5-4bdf-6e49-c77d2de647a2::9f8841e6-e04a-498b-b3ff-d59cf7cb7ea
cs3Label=result cs3=clientError cs4Label=httpStatusCode cs4=404
cs5Label=xForwardedFor cs5=198.51.100.1
```

- A `POST` request that results in a 403 error:

```
CEF:0|cloud_foundry|cloud_controller_ng|2.54.0|POST /v2/apps|POST
/v2/apps|0|rt=1460691405564 suser=bob suid=4f9a33f9-fb13-4774-a708-
f60c939625cd request=/v2/apps?async=true requestMethod=POST
src=127.0.0.1 dst=192.0.2.1 cs1Label=userAuthenticationMechanism
cs1=oauth-access-token cs2Label=vcapRequestId cs2=booc03111-9999-4003-88ab-
2019r33333ou::5a4993fc-722f-48bc-aff4-99b200519bb5 cs3Label=result
cs3=clientError cs4Label=httpStatusCode cs4=403 cs5Label=xForwardedFor
cs5=198.51.100.1
```

## UAA Logging

UAA logs security events to a file located at `/var/vcap/sys/log/uua/uua.log` on the UAA virtual machine (VM). Because these logs are automatically rotated, you must configure a syslog drain to forward your system logs to a log management service.

See the [Configuring System Logging in PAS](#) topic for more information.

## Log Events

UAA logs identify the following categories of events:

- Authorization and Password Events
- SCIM Administration Events
- Token Events
- Client Administration Events

- UAA Administration Events

To learn more about the names of the events included in these categories and the information they record in the UAA logs, see [User Account and Authentication Service Audit Requirements](#).

## Example Log Entries

The following sections provide several example requests with the corresponding UAA log entries.

### Successful User Authentication

```
Audit: TokenIssuedEvent (["openid","scim.read","uaa.user",
"cloud_controller.read","password.write","cloud_controller.write",
"scim.write"]): principal=a42026d6-5533-1884-eef2-838abcd0i3e3,
origin=[client=cf, user=bob], identityZoneId=[uaa]
```

- This entry records a `TokenIssuedEvent`.
- UAA issued a token associated with the scopes `"openid", "scim.read", "uaa.user", "cloud_controller.read", "password.write", "cloud_controller.write", "scim.write"` to the user `bob`.

### Failed User Authentication

```
Audit: UserAuthenticationFailure ('bob@example.com'):
principal=61965469-c821-46b7-825f-630e12a51d6c,
origin=[remoteAddress=198.51.100.1, clientId=cf],
identityZoneId=[uaa]
```

- This entry records a `UserAuthenticationFailure`.
- The user `bob@example.com` originating at `198.51.100.1` failed to authenticate.

### Successful User Creation

```
Audit: UserCreatedEvent (["user_id=61965469-c821-
46b7-825f-630e12a51d6c","username=bob@example.com"]):
principal=91220262-d901-44c0-825f-633i33b55d6c,
origin=[client=cf, user=admin, details=(198.51.100.1,
tokenType=beartokenValue=<TOKEN>,
sub=20i03423-dd8e-33e1-938d-e9999e30f500,
iss=https://uaa.example.com/oauth/token)], identityZoneId=[uaa]
```

- This entry records a `UserCreatedEvent`.
- The `admin` user originating at `198.51.100.1` created a user named `bob@example.com`.

### Successful User Deletion

```
Audit: UserDeletedEvent (["user_id=61965469-c821-
46b7-825f-630e12a51d6c","username=bob@example.com"]):
principal=61965469-c821-46b7-825f-630e12a51d6c,
origin=[client=admin, details=(remoteAddress=198.51.100.1,
tokenType=beartokenValue=<TOKEN>,
sub=admin, iss=https://uaa.example.com/oauth/token)], identityZoneId=[uaa]
```

- This entry records a `UserDeletedEvent`.
- The `admin` user originating at `198.51.100.1` deleted a user named `bob@example.com`.

## CredHub Logging

CredHub logs security events to a file located at `/var/vcap/sys/log/credhub/credhub_security_events.log` on the CredHub VM. Because these logs are automatically rotated, you must configure a syslog drain to forward your system logs to a log management service.

See the [Configuring System Logging in PAS](#) topic for more information.

## Format for Log Entries

CredHub logs security events in the [Common Event Format \(CEF\)](#) (CEF). CEF specifies the following format for log entries:

```
CEF:Version|Device Vendor|Device Product|Device Version|Signature ID|Name|Severity|Extension
```

Entries in the CredHub log use the following format:

```
CEF:0|cloud_foundry|credhub|CREDHUB_SERVER_VERSION|
SIGNATURE_ID|NAME|0|rt=TIMESTAMP user=USERNAME uid=USER_GUID
cs1Label=userAuthenticationMechanism cs1=AUTH_MECHANISM
request=REQUEST requestMethod=REQUEST_METHOD
cs3Label=versionUuid cs3=VERSION_UUID
cs4Label=httpStatusCode cs4=HTTP_STATUS_CODE src=SOURCE_ADDRESS dst=DESTINATION_ADDRESS
cs2Label=resourceName cs2=RESOURCE_NAME
cs5Label=resourceUuid cs5=RESOURCE_UUID deviceAction=OPERATION
cs6Label=requestDetails cs6=REQUEST_DETAILS
```

Refer to the following list for a description of the properties shown in the CredHub log.

- `CEF_VERSION` : The version of CEF used in the logs.
- `CREDHUB_SERVER_VERSION` : The current CredHub server version.
- `SIGNATURE_ID` : The method and path of the request. For example, `GET /v2/app:GUID`.
- `NAME` : The same as `SIGNATURE_ID`.
- `TIMESTAMP` : The number of milliseconds since the Unix epoch.
- `USERNAME` : The name of the user who originated the request, as defined by UAA. In the case of mTLS, no-auth, or not defined, this value is empty.
- `USER_GUID` : The “actor” identifier. For example, `mtls-app:GUID`. If there is no authenticated user, this value is empty.
- `AUTH_MECHANISM (cs1)` : The user authentication mechanism. This can be `oauth-access-token`, `mtls-auth`, or `no-auth`.
- `REQUEST` : The request path and parameters. For example, `/v2/info?MY-PARAM=VALUE`.
- `REQUEST_METHOD` : The method of the request. For example, `GET`.
- `HTTP_STATUS_CODE (cs4)` : The HTTP status code of the response. For example, `200`.
- `SOURCE_ADDRESS` : The IP address of the client who originated the request.
- `DESTINATION_ADDRESS` : The IP address of the CredHub VM.
- `RESOURCE_NAME (cs2)` : The credential path name. For example, `/my/awesome/credential`.
- `RESOURCE_UUID (cs5)` : The top-level credential UUID. This is not the credential version.
- `VERSION_UUID (cs3)` : The credential version UUID.
- `OPERATION (deviceAction)` : The device action. The possible operations include the following:
  - `get`
  - `set`
  - `generate`
  - `regenerate`
  - `bulk-regenerate`
  - `delete`
  - `find`
  - `get-permissions`
  - `add-permission`
  - `delete-permission`
  - `interpolate`
  - `info`
  - `version`
  - `health`
  - `key-usage`

- `update-transitional-version`
- `REQUEST_DETAILS (cs6)`: A JSON blob that differs per request.

**Note:** The CredHub logs include descriptions for each custom label. For example, the logs include `cs2Label=resourceName` to define the `cs2` label. The value for `resourceName` appears in the log next to `cs2=/path/to/credential`.

## Example Log Entries

The following sections provide several example requests with the corresponding CredHub log entries.

### Accessing a Credential

```
CEF:0|cloud_foundry|credhub|2.0.0-beta.28|GET /api/v1/data|
GET /api/v1/data|0|rt=1530901816757
suser=app:3c538393-d192-4e23-8c83-456654b3fa6c
suid=mtls-app:3c538393-d192-4e23-8c83-456654b3fa6c
cs1Label=userAuthenticationMechanism
cs1=mutual_tls request=/api/v1/data?path=0b353167-0d5a-48c7-5036-7eaa
requestMethod=GET
cs3Label=versionUuid cs3=null
cs4Label=httpStatusCode cs4=200 src=10.0.0.1 dst=credhub.service.cf.internal
cs2Label=resourceName cs2=null
cs5Label=resourceUuid cs5=null deviceAction=FIND
cs6Label=requestDetails
cs6={"nameLike":null,"path":"0b353167-0d5a-48c7-5036-7eaa","paths":null}
```

- A user authenticated to the CredHub instance using `mutual_tls` from `10.0.0.1`.
- The authenticated user accessed a CredHub credential.

### Setting a Credential

```
CEF:0|cloud_foundry|credhub|2.0.0-beta.28|PUT /api/v1/data|
PUT /api/v1/data|0|rt=1530901097921
suser=cc_service_key_client suid=uaa-client:cc_service_key_client
cs1Label=userAuthenticationMechanism
cs1=uaa request=/api/v1/data requestMethod=PUT
cs3Label=versionUuid cs3=32b3d5bf-463a-4045-a6b5-65c97e61059a
cs4Label=httpStatusCode cs4=200 src=10.0.0.1 dst=credhub.service.cf.internal
cs2Label=resourceName cs2=/1530901097842989110
cs5Label=resourceUuid cs5=cda25c5-a40a-4512-b6f5-a42f8c3b5c4c deviceAction=SET
cs6Label=requestDetails
cs6={"name":"/1530901097842989110","type":"json","mode":null,
"additionalPermissions":[{"actor": "uaa-client:cc_service_key_client",
"allowedOperations":["READ","WRITE","DELETE","WRITE_ACL","READ_ACL"]}]}
```

- A user authenticated to the CredHub instance with UAA from `10.0.0.1`.
- The authenticated user set a CredHub credential.

### Generating a Credential

```
CEF:0|cloud_foundry|credhub|2.0.0-beta.28|POST /api/v1/data|
POST /api/v1/data|0|rt=1530901403532
suser=app:3c538393-d192-4e23-8c83-456654b3fa6c
suid=mtls-app:3c538393-d192-4e23-8c83-456654b3fa6c
cs1Label=userAuthenticationMechanism
cs1=mutual_tls request=/api/v1/data requestMethod=POST
cs3Label=versionUuid cs3=8c21b7b3-d4bf-4d8a-a0c5-64c8c207cb40
cs4Label=httpStatusCode cs4=200 src=10.0.0.1 dst=credhub.service.cf.internal
cs2Label=resourceName cs2=/0b353167-0d5a-48c7-5036-7eaa/2
cs5Label=resourceUuid cs5=1d55cff3-5264-434c-a127-0810f341cb2b deviceAction=GENERATE
cs6Label=requestDetails cs6={"name":"/my/awesome/credential","type":"password",
"mode":null,"additionalPermissions":[{"actor": "mtls-app:3c538393-d192-4e23-8c83-456654b3fa6c",
"allowedOperations":["READ","WRITE","DELETE","WRITE_ACL","READ_ACL"]}]}
```

- A user authenticated to the CredHub instance using `mutual_tls` from `10.0.0.1`.
- The authenticated user generated a password credential named `/my/awesome/credential`.

## Network Security

This section introduces some of the networking and routing security options for your Pivotal Cloud Foundry (PCF) deployment.

## Securing Traffic and Controlling Routes

You can enable and configure a number of customization options to secure traffic in and out of your PCF deployment.

- [TLS Connections in PCF Deployments](#)
- [Securing Traffic into Cloud Foundry](#)
- [Providing a Certificate for Your SSL/TLS Termination Point](#)
- [Enabling TCP Routing](#)

## Using the IPsec Add-On

The IPsec add-on for PCF provides additional security to the network layer for each BOSH-deployed virtual machine (VM).

The PCF IPsec add-on secures network traffic within a Cloud Foundry deployment and provides internal system protection if a malicious actor breaches your firewall.

- [Securing Data in Transit with the IPsec Add-on ↗](#)
- [Rotating IPsec Credentials ↗](#)
- [Installing the Pivotal Cloud Foundry IPsec Add-On ↗](#)

## Network Communication Paths in PCF

- [BOSH DNS Network Communications](#)
- [Cloud Controller Network Communications](#)
- [Container-to-Container Network Communications](#)
- [CredHub Network Communications](#)
- [Diego Network Communications](#)
- [Loggregator Network Communications](#)
- [MySQL Network Communications](#)
- [NATS Network Communications](#)
- [Routing Network Communications](#)
- [UAA Network Communications](#)

## BOSH DNS Network Communications

This topic describes [BOSH DNS](#) internal network communication paths with other Pivotal Application Service (PAS) components.

### BOSH DNS Communications

The following table lists network communication paths for BOSH DNS. For information about configuring BOSH DNS in the BOSH Director, see [BOSH DNS Config Pane in the BOSH Director Tile](#).

**Note:** Port 8853 is the destination port for communications between BOSH DNS health processes. You must allow TCP traffic on 8853 for all VMs running BOSH DNS.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
Any VM running BOSH DNS	backup-prepare	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	ccdb	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	clock_global	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	cloud_controller	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	cloud_controller_worker	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	consul_server	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	diego_brain	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	diego_cell	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	diego_database	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	doppler	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	ha_proxy	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	loggregator_trafficcontroller	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	mysql_proxy*	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	mysql_monitor*	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	nats	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	nfs_server†	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	router	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	syslog_adapter	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	syslog_scheduler	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	tcp_router	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS					Unencrypted. This communication happens inside the VM.

Any VM running BOSH DNS	uaa	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	uaadb	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	Service tile VMs	53	TCP and UDP	DNS	Unencrypted. This communication happens inside the VM.
Any VM running BOSH DNS	backup-prepare	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	ccdb	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	clock_global	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	cloud_controller	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	cloud_controller_worker	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	consul_server	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	diego_brain	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	diego_cell	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	diego_database	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	doppler	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	ha_proxy	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	loggregator_trafficcontroller	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	mysql_proxy*	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	mysql_monitor*	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	nats	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	nfs_server†	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	router	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	syslog_adapter	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	syslog_scheduler	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	tcp_router	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	uaa	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	uaadb	8853	TCP	HTTPS	Mutual TLS
Any VM running BOSH DNS	Service tile VMs	8853	TCP	HTTPS	Mutual TLS

\* Applies only to deployments where internal MySQL is selected as the database.

† Applies only to deployments where the internal NFS server is selected for file storage.



## Cloud Controller Network Communications

This topic describes [Cloud Controller](#) internal network communication paths with other Pivotal Application Service (PAS) components.

### Inbound Communications

The following table lists network communication paths that are inbound to the Cloud Controller.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
cloud_controller	cloud_controller (Routing API)	443	TCP	HTTPS	OAuth 2.0
diego_brain	cloud_controller	9023	TCP	HTTPS	Mutual TLS
diego_brain (SSH Proxy)	cloud_controller	9024	TCP	HTTPS	OAuth 2.0
diego_cell (Rep)	cloud_controller	9023	TCP	HTTPS	Mutual TLS
diego_database (BBS)	cloud_controller	9023	TCP	HTTPS	Mutual TLS
doppler (Syslog Drain Binder)	cloud_controller	9023	TCP	HTTPS	Mutual TLS
loggregator_trafficcontroller	cloud_controller	9023	TCP	HTTPS	Mutual TLS
router	cloud_controller	9024	TCP	HTTPS	OAuth 2.0

### Outbound Communications

The following table lists network communication paths that are outbound from the Cloud Controller.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
cloud_controller	mysql_proxy*	3306	TCP	MySQL	MySQL authentication**
cloud_controller	nfs_server or other blobstore†	4443	TCP	HTTPS	TLS and basic authentication
cloud_controller	uaa	8443	TCP	HTTPS	OAuth 2.0 or none‡
cloud_controller	diego_database (BBS)	8889	TCP	HTTPS	Mutual TLS
cloud_controller (Route Registrar)	nats	4222	TCP	NATS	Basic authentication
cloud_controller (Routing API)	diego_database (Locket)	8891	TCP	HTTPS	Mutual TLS
cloud_controller_worker	mysql_proxy*	3306	TCP	MySQL	MySQL authentication**
cloud_controller_worker	nfs_server or other blobstore†	4443	TCP	HTTPS	TLS and basic authentication
clock_global	mysql_proxy*	3306	TCP	MySQL	MySQL authentication**

\*Applies only to deployments where internal MySQL is selected as the database.

\*\*MySQL authentication uses the MySQL native password method.

†The destination depends on your file storage or blobstore configuration.

‡The authentication method depends on the type of request.

### BOSH DNS Communications

By default, PAS components and app containers look up services using the BOSH DNS service discovery mechanism. To support this lookup, BOSH Director collocates a BOSH DNS server on every deployed VM. For more information, see [BOSH DNS Network Communications](#).



## Container-to-Container Networking Communications

This topic describes [Container-to-Container Networking](#) internal network communication paths with other Pivotal Application Service (PAS) components.

### Inbound Communications

The following table lists network communication paths that are inbound to Container-to-Container Networking.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
diego_cell (Silk CNI)	diego_cell (Silk Daemon)	23954	TCP	HTTP	None
diego_cell (Silk Daemon)	diego_api (Silk Controller)	4103	TCP	HTTP	Mutual TLS
diego_cell (VXLAN Policy Agent)	diego_database (api - Policy Server Internal)	4003	TCP	HTTP	Mutual TLS
diego_cell (BOSH DNS Adapter)	diego_brain (Service Discovery Controller)	8054	TCP	HTTP	Mutual TLS

### Outbound Communications

The following table lists network communication paths that are outbound from Container-to-Container Networking.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
diego_database (api - Policy Server)	uaa	8443	TCP	HTTPS	TLS
diego_database (api - Policy Server)	cloud_controller (api - Cloud Controller)	9022	TCP	HTTP	OAuth 2.0
diego_database (api - Policy Server)	mysql_proxy*	3306	TCP	MySQL	MySQL authentication*
diego_brain (Service Discovery Controller)	nats (NATS)	4222	TCP	HTTP	Basic authentication
diego_cell (BOSH DNS)	diego_cell (BOSH DNS Adapter)	8053	TCP	HTTP	None
diego_cell (VXLAN Policy Agent)	mysql_proxy*	3306	TCP	MySQL	MySQL authentication*

\*Applies only to deployments where internal MySQL is selected as the database.

### BOSH DNS Communications

By default, PAS components and app containers look up services using the BOSH DNS service discovery mechanism. To support this lookup, BOSH Director collocates a BOSH DNS server on every deployed VM. For more information, see [BOSH DNS Network Communications](#).

## CredHub Network Communications

This topic describes [CredHub](#) internal network communication paths with other Pivotal Application Service (PAS) components.

### Inbound Communications

The following table lists network communication paths that are inbound to the CredHub.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
cloud_controller (api)	credhub	8844	TCP	HTTPS	OAuth 2.0
diego_cell	credhub	8844	TCP	HTTPS	Mutual TLS†
windows_cell	credhub	8844	TCP	HTTPS	Mutual TLS†
windows2016_cell	credhub	8844	TCP	HTTPS	Mutual TLS†

### Outbound Communications

The following table lists network communication paths that are outbound from the CredHub.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
credhub	uaa	8443	TCP	HTTPS	n/a
credhub	mysql_proxy*	3306	TCP	MySQL	MySQL authentication**

\*Applies only to deployments where internal MySQL is selected as the database.

\*\*MySQL authentication uses the MySQL native password method.

†Diego cells use the certificate pairs generated for individual containers to authenticate with CredHub on behalf of applications.

### BOSH DNS Communications

By default, PAS components and app containers look up services using the BOSH DNS service discovery mechanism. To support this lookup, BOSH Director collocates a BOSH DNS server on every deployed VM. For more information, see [BOSH DNS Network Communications](#).

## Diego Network Communications

This topic describes [Diego](#) internal network communication paths with other Pivotal Application Service (PAS) components.

### Inbound Communications

The following table lists network communication paths that are inbound to Diego.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
cloud_controller	diego_database (BBS)	8889	TCP	HTTPS	Mutual TLS
cloud_controller (Routing API)	diego_database (Locket)	8891	TCP	HTTPS	Mutual TLS

### Diego Internal Communications

The following table lists network communication paths that are internal for Diego.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
diego_brain (Auctioneer)	diego_cell (Rep)	1801	TCP	HTTPS	Mutual TLS
diego_brain (Auctioneer)	diego_database (BBS)	8889	TCP	HTTPS	Mutual TLS
diego_brain (Auctioneer)	diego_database (Locket)	8891	TCP	HTTPS	Mutual TLS
diego_brain (SSH Proxy)	diego_database (BBS)	8889	TCP	HTTPS	Mutual TLS
diego_brain (SSH Proxy)	diego_cell (App instances)	Varies <sup>‡</sup>	TCP	SSH	SSH
diego_brain (TPS Watcher)	diego_database (Locket)	8891	TCP	HTTPS	Mutual TLS
diego_cell (local Route Emitter)	diego_database (BBS)	8889	TCP	HTTPS	Mutual TLS
diego_cell (Rep)	diego_brain (CC Uploader)	9091	TCP	HTTPS	Mutual TLS
diego_cell (Rep)	diego_brain (File Server) <sup>*</sup>	8080	TCP	HTTP	None
diego_cell (Rep)	diego_database (BBS)	8889	TCP	HTTPS	Mutual TLS
diego_cell (Rep)	diego_database (Locket)	8891	TCP	HTTPS	Mutual TLS
diego_database (BBS)	diego_brain (Auctioneer)	9016	TCP	HTTPS	Mutual TLS
diego_database (BBS)	diego_cell (Rep)	1801	TCP	HTTPS	Mutual TLS
diego_database (BBS)	diego_database (Locket)	8891	TCP	HTTPS	Mutual TLS

<sup>\*</sup>The Diego File Server is responsible for distributing non-sensitive, static platform assets to internal platform components.

### Outbound Communications

The following table lists network communication paths that are outbound from Diego.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
diego_brain	cloud_controller	9023	TCP	HTTPS	Mutual TLS
diego_brain (SSH Proxy)	cloud_controller	9024	TCP	HTTPS	OAuth 2.0
diego_brain (SSH Proxy)	uaa	443	TCP	HTTPS	TLS and OAuth 2.0
diego_cell (local Route Emitter)	nats	4222	TCP	NATS	Basic authentication
diego_cell (Rep)	cloud_controller	9023	TCP	HTTPS	Mutual TLS
diego_cell (Rep)	nfs_server or other blobstore <sup>*</sup>	Varies	TCP	HTTP	Signed URLs/TLS

diego_database (BBS)	cloud_controller	9023	TCP	HTTPS	Mutual TLS
diego_database (BBS)	mysql_proxy <sup>†</sup>	3306	TCP	MySQL	MySQL authentication <sup>**</sup>
diego_database (Locket)	mysql_proxy <sup>†</sup>	3306	TCP	MySQL	MySQL authentication <sup>**</sup>

<sup>\*</sup>The destination depends on your PAS blobstore configuration. If you use the internal blobstore, the Diego Cell communicates to the blobstore using TLS on port 4443.

<sup>\*\*</sup>MySQL authentication uses the MySQL native password method.

<sup>†</sup>Applies only to deployments where internal MySQL is selected as the database.

<sup>‡</sup>These are the host-side ports that map to port 2222 in app instance containers and are typically within the range 61001 to 65534.

## BOSH DNS Communications

By default, PAS components and app containers look up services using the BOSH DNS service discovery mechanism. To support this lookup, BOSH Director collocates a BOSH DNS server on every deployed VM. For more information, see [BOSH DNS Network Communications](#).

## Loggregator Network Communications

This topic describes [Loggregator](#) internal network communication paths with other Pivotal Application Service (PAS) components.

### Loggregator Communications

The following table lists network communication paths for Loggregator.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
Any*	loggregator_trafficcontroller	8081	TCP	HTTP/WebSocket	OAuth
Any VM running Metron	doppler	8082	TCP	gRPC over HTTP/2	Mutual TLS
loggregator_trafficcontroller	doppler	8082	TCP	gRPC over HTTP/2	Mutual TLS
loggregator_trafficcontroller	uaa	8443	TCP	HTTPS	TLS
loggregator_trafficcontroller	cloud_controller	9023	TCP	HTTPS	Mutual TLS
loggregator_trafficcontroller (Reverse Log Proxy)	doppler	8082	TCP	gRPC over HTTP/2	Mutual TLS
Any*	loggregator_trafficcontroller (cf auth proxy)	8083	TCP	HTTP	OAuth
loggregator_trafficcontroller (Route Registrar)	nats	4222	TCP	NATS	Basic authentication
loggregator_trafficcontroller (metrics_forwarder)	BOSH Director (metrics_server)	25595 and 8443	TCP	gRPC over HTTP/2	Mutual TLS
loggregator_trafficcontroller	doppler (Log Cache)	8080	TCP	gRPC over HTTP/2	Mutual TLS
syslog_adapter	loggregator_trafficcontroller (Reverse Log Proxy)	8082	TCP	gRPC over HTTP/2	Mutual TLS
syslog_adapter	Any**	Any***	TCP	TCP, TCP w/ TLS, HTTPS	Basic authentication****
syslog_scheduler	syslog_adapter	4443	TCP	gRPC over HTTP/2	Mutual TLS
syslog_scheduler	cloud_controller	9023	TCP	HTTPS	Mutual TLS
loggregator_trafficcontroller (Reverse Log Proxy Gateway)	cloud_controller	9023	TCP	HTTPS	Mutual TLS
Any*	loggregator_trafficcontroller (Reverse Log Proxy Gateway)	8088	TCP	HTTP/Server Sent Events	OAuth

\*Any source VM can send requests to the specified destination within its subnet.

\*\* Any host configured through a user-provided service binding with a syslog URL.

\*\*\* Any port configured through a user-provided service binding with syslog URL.

\*\*\*\* Basic authentication only supported for HTTPS syslog drains.

### Log Cache Communications

The following table lists network communication paths for Log Cache.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
loggregator_trafficcontroller (Reverse Log Proxy)	log-cache (Nozzle)	8082	TCP	gRPC over HTTP/2	Mutual TLS
Any*	log-cache	8080	TCP	gRPC over HTTP/2	Mutual TLS
gorouter	log-cache (Auth)	8083	TCP	HTTP	Oauth

	Proxy)				
log-cache (Auth Proxy)	uaa	8443	TCP	HTTPS	TLS
log-cache (Auth Proxy)	cloud_controller	9024	TCP	HTTPS	TLS

\*Any source VM can send requests to the specified destination within its subnet.

## BOSH DNS Communications

By default, PAS components and app containers look up services using the BOSH DNS service discovery mechanism. To support this lookup, BOSH Director collocates a BOSH DNS server on every deployed VM. For more information, see [BOSH DNS Network Communications](#).

## MySQL Network Communications

This topic describes MySQL internal network communication paths with other Pivotal Application Service (PAS) components.

 **Note:** These communications only apply to deployments where internal MySQL is selected as the PAS database.

### Inbound Communications

The following table lists network communication paths that are inbound to MySQL VMs.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
cloud_controller	mysql_proxy	3306	TCP	MySQL	MySQL authentication*
cloud_controller_worker	mysql_proxy	3306	TCP	MySQL	MySQL authentication*
clock_global	mysql_proxy	3306	TCP	MySQL	MySQL authentication*
credhub	mysql_proxy	3306	TCP	MySQL	MySQL authentication*
diego_cell (VXLAN Policy Agent)	mysql_proxy	3306	TCP	MySQL	MySQL authentication*
diego_database (Policy Server)	mysql_proxy	3306	TCP	MySQL	MySQL authentication*
diego_database (BBS)	mysql_proxy	3306	TCP	MySQL	MySQL authentication*
diego_database (Locket)	mysql_proxy	3306	TCP	MySQL	MySQL authentication*
uaa	mysql_proxy	3306	TCP	MySQL	MySQL authentication*

(\*) MySQL authentication uses the MySQL native password method.

### Internal Communications

The following table lists network communication paths that are internal to MySQL VMs.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
mysql	mysql (Galera)	4567	TCP	MySQL	MySQL authentication*
mysql_monitor	mysql (MySQL Server)	3306	TCP	HTTP	Basic authentication
mysql_monitor	mysql_proxy (Proxy health check)	443/8080**	TCP	HTTP	Basic authentication
mysql_proxy	mysql (MySQL Server)	3306	TCP	HTTP	MySQL authentication*
mysql_proxy	mysql (Galera health check)	9200	TCP	HTTP	Basic authentication

(\*) MySQL authentication uses the MySQL native password method.

(\*\*) Port 443 is used if mysql\_proxy is registered with Gorouter. If not registered, mysql\_proxy uses port 8080 instead.

### Outbound Communications

The following table lists network communication paths that are outbound from MySQL.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
mysql_monitor	uaa	8443	TCP	HTTPS	OAuth
mysql_proxy (Route Registrar)	nats	4222	TCP	NATS	Basic authentication

### BOSH DNS Communications

By default, PAS components and app containers look up services using the BOSH DNS service discovery mechanism. To support this lookup, BOSH Director collocates a BOSH DNS server on every deployed VM. For more information, see [BOSH DNS Network Communications](#).



## NATS Network Communications

This topic describes [NATS](#) internal network communication paths with other Pivotal Application Service (PAS) components.

### Publish Communications

The following table lists network communications that are published to NATS.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
cloud_controller (Route Registrar)	nats	4222	TCP	NATS	Basic authentication
loggregator_trafficcontroller (Route Registrar)	nats	4222	TCP	NATS	Basic authentication
mysql_proxy (Route Registrar)*	nats	4222	TCP	NATS	Basic authentication
nfs_server (Route Registrar)†	nats	4222	TCP	NATS	Basic authentication
uaa (Route Registrar)	nats	4222	TCP	NATS	Basic authentication
diego_cell (local Route Emitter)	nats	4222	TCP	NATS	Basic authentication

\*Applies only to deployments where internal MySQL is selected as the database.

†Applies only to deployments where the internal NFS server is selected for file storage.

### Subscribe Communications

The following table lists network communications that are subscribed to NATS.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
router	nats	4222	TCP	NATS	Basic authentication

### BOSH DNS Communications

By default, PAS components and app containers look up services using the BOSH DNS service discovery mechanism. To support this lookup, BOSH Director collocates a BOSH DNS server on every deployed VM. For more information, see [BOSH DNS Network Communications](#).

## Routing Network Communications

This topic describes the internal network communication paths of the routing subsystem with other Pivotal Application Service (PAS) components.

### HTTP Routing

The following table lists network communication paths for HTTP routing.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
diego_cell (local Route Emitter)	nats	4222	TCPs	NATS	Basic authentication
Load balancer	router (Gorouter)	80	TCP	HTTP	None
Load balancer	router (Gorouter)	443	TCP	HTTPS	TLS
router (Gorouter)	nats	4222	TCP	NATS	Basic authentication
router (Gorouter)	System components	Varies	TCP	Varies	None
router (Gorouter)	App containers	Varies	TCP	Varies	Optional TLS
haproxy	router (Gorouter)	80	TCP	HTTP	None
haproxy	router (Gorouter)	443	TCP	HTTPS	TLS
Load balancer	haproxy	80	TCP	HTTP	None
Load balancer	haproxy	443	TCP	HTTPS	TLS

### TCP Routing (Optional)

The following table lists network communication paths for TCP routing.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
cloud_controller	cloud_controller (Routing API)*	443	TCP	HTTPS	TLS and OAuth 2.0
cloud_controller (Routing API)	diego_database (Locket)	8891	TCP	HTTPS	Mutual TLS
cloud_controller (Routing API)	mysql_proxy	3306	TCP	MySQL	MySQL authentication**
cloud_controller (Routing API)	uaa	8443	TCP	HTTPS	TLS
diego_brain (global TCP Emitter)	cloud_controller (Routing API)	3000	TCP	HTTP	OAuth 2.0
diego_brain (global TCP Emitter)	uaa	8443	TCP	HTTPS	TLS
diego_cell (local Route Emitter)	cloud_controller (Routing API)	3000	TCP	HTTP	OAuth 2.0
diego_cell (local Route Emitter)	uaa	8443	TCP	HTTPS	TLS
Load balancer	tcp_router	1024-65535†	TCP	TCP	None
router (Gorouter)	cloud_controller (Routing API)	3000	TCP	HTTP	OAuth 2.0
router (Gorouter)	uaa	8443	TCP	HTTPS	TLS
tcp_router	cloud_controller (Routing API)	3000	TCP	HTTP	OAuth 2.0
tcp_router	uaa	8443	TCP	HTTPS	TLS

\* This communication happens through a load balancer and a Gorouter. Requests are received by Routing API on port 3000.

<sup>†</sup> You can use this port range to configure the port in the PAS tile.

<sup>\*\*</sup> MySQL authentication uses the MySQL native password method.

## Service Mesh (Optional)

The following table lists network communication paths for service mesh.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
cloud_controller (cloud_controller_ng)	istio_control (copilot)	9001	TCP	GRPC	Mutual TLS
istio_control (copilot)	diego_database (bbs)	8889	TCP	HTTP	Mutual TLS
istio_control(pilot-discovery)	istio_control (copilot)	9009	TCP	GRPC	Mutual TLS
istio_router (envoy)	App Containers	Varies	TCP	HTTP/HTTPS	Optional TLS
istio_router(envoy)	istio_control(pilot-discovery)	15010	TCP	GRPC	None
Load balancer	istio_router (envoy)	80	TCP	HTTP	None
Load balancer	istio_router (envoy)	443	TCP	HTTPS	TLS
Load balancer (health_check)	istio_router (envoy)	8002	TCP	HTTP	None
route_syncer (cc_route_syncer)	istio_control (copilot)	9001	TCP	GRPC	Mutual TLS
route_syncer (cc_route_syncer)	mysql_proxy*	3306	TCP	MySQL	MySQL authentication*
n/a (admin)	istio_router (envoy)	8001	TCP	HTTP	None
n/a (for envoy secure GRPC communication)	istio_control(pilot-discovery)	15012	TCP	GRPC	Mutual TLS
n/a (for HTTP discovery service)	istio_control(pilot-discovery)	8080	TCP	HTTP	None
n/a (for pilot's self-monitoring)	istio_control(pilot-discovery)	9093	TCP	HTTP	None

\* Applies only to deployments where internal MySQL is selected as the database.

## BOSH DNS Communications

By default, PAS components and app containers look up services using the BOSH DNS service discovery mechanism. To support this lookup, BOSH Director collocates a BOSH DNS server on every deployed VM. For more information, see [BOSH DNS Network Communications](#).

## UAA Network Communications

This topic describes [User Account and Authentication \(UAA\)](#) internal network communication paths with other Pivotal Application Service (PAS) components.

### Inbound Communications

The following table lists network communication paths that are inbound to UAA.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
cloud_controller	uaa	8443	TCP	HTTPS	OAuth 2.0 or none*
diego_brain (SSH Proxy)	uaa	443	TCP	HTTPS	OAuth 2.0
loggregator_trafficcontroller	uaa	8443	TCP	HTTPS	TLS
mysql_monitor	uaa	8443	TCP	HTTPS	OAuth
router	uaa	8443	TCP	HTTPS	OAuth 2.0

\*The authentication method depends on the type of request.

### Outbound Communications: Internal to PCF

The following table lists network communication paths that are outbound from UAA.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Security and Authentication
uaa	mysql_proxy*	3306	TCP	MySQL	MySQL authentication**
uaa (Route Registrar)	nats	4222	TCP	NATS	Basic authentication

\*Applies only to deployments where internal MySQL is selected as the database.

\*\* MySQL authentication uses the MySQL native password method.

### Outbound Communications: External to PCF

The following table lists network communication paths from UAA that are outbound to external systems.

Source VM	Destination VM	Port	Transport Layer Protocol	App Layer Protocol	Authentication
uaa	LDAP	LDAP server communication port	TCP	LDAP/LDAPS	Basic authentication (LDAP bind)
uaa	SAML/OIDC	80 or 443 (HTTP port)	TCP	HTTP/HTTPS	Key

### BOSH DNS Communications

By default, PAS components and app containers look up services using the BOSH DNS service discovery mechanism. To support this lookup, BOSH Director collocates a BOSH DNS server on every deployed VM. For more information, see [BOSH DNS Network Communications](#).

## Credential and Identity Management

This section provides links to different aspects of identity management, including credential management handled by CredHub, user creation and permissions management, and authentication for Pivotal Cloud Foundry (PCF).

### General Identity Management

The following topics provide general information about credential and identity management in PCF.

- [Pivotal Cloud Foundry User Types ↗](#)
- [Retrieving Credentials from Your Deployment ↗](#)

### CredHub Documentation

CredHub provides centralized credential management in Pivotal Cloud Foundry (PCF). Credentials can include passwords, certificates, and SSH keys.

CredHub centralizes and secures credential generation, storage, lifecycle management, and system access.

For more information about CredHub, see the following topics.

- [CredHub ↗](#): Provides an overview of CredHub.
- [CredHub Credential Types ↗](#): Provides a reference of credential types supported in CredHub.

### UAA Documentation

PCF uses UAA to manage account roles and permissions in PCF runtimes. UAA supports access control as OAuth2 services and can store user information internally, or connect to external user stores through LDAP or SAML.

For more information about UAA, see the following topics.

- [UAA Overview ↗](#)
- [UAA Concepts ↗](#)
- [UAA Architecture ↗](#)
- [Identity Providers in UAA ↗](#)

### PCF Roles and User Accounts

The following topics describe how to manage PCF roles and user accounts in PCF.

- [Creating and Managing Users with the UAA CLI \(UAAC\) ↗](#)
- [Adding Existing SAML or LDAP Users to a PCF Deployment ↗](#)

### Ops Manager Roles and Permissions

- [Creating UAA Clients for BOSH Director ↗](#)
- [Configuring Role-Based Access Control \(RBAC\) in Ops Manager ↗](#)
- [Creating and Managing Ops Manager User Accounts ↗](#)

### PAS Roles and Permissions

This section provides links to Pivotal Application Service (PAS) roles and permissions documentation.

- [Orgs, Spaces, Roles, and Permissions ↗](#):
- [Creating New PAS User Accounts ↗](#)
- [Managing User Roles with Apps Manager ↗](#)

## PKS Roles and Permissions

This section provides links to Pivotal Container Service (PKS) roles and permissions documentation.

- [Managing Users in PKS with UAA ↗](#)

## Pivotal Cloud Foundry User Types

This topic describes the types of users in a Pivotal Cloud Foundry (PCF) deployment, their roles and permissions, and who creates and manages their user accounts.

The users who run a PCF deployment and have admin privileges are [operators](#). With Pivotal Application Service (PAS) installed to host apps, you add two more user types: [PAS users](#) who develop the apps and manage the development environment, and [end users](#) who just run the apps.

PCF distinguishes between these three user types and multiple user roles that exist within a single user type. Roles are assigned categories that more specifically define functions that a user can perform. A user may serve in more than one role at the same time.

## Operators

Operators have the highest, admin-level permissions. We also refer to operators as Ops Manager admins and PAS admins because they perform an admin role within these contexts.

## Tools and Tasks

Operators fulfill system administrator roles covering the entire PCF deployment. They work primarily with their IaaS and Ops Manager, to configure and maintain PAS component VMs. The component VMs, in turn, support the VMs that host applications. Typical operator tasks include:

- Deploying and configuring Ops Manager, PAS, and other product and service tiles.
- Maintaining and upgrading PCF deployments.
- Creating user accounts for PAS users and the orgs that PAS users work within.
- Creating service plans that define the access granted to end users.

## User Accounts

When Ops Manager starts up for the first time, the operator specifies one of the following authentication systems for operator user accounts:

- Internal authentication, using a new UAA database that Ops Manager creates.
- External authentication, through an existing identity provider accessed through SAML protocol.

The operator can then use the UAAC to [create more](#) operator accounts.

## PAS Users

PAS users are app developers, managers, and auditors who work within orgs and spaces, the virtual compartments within a deployment where PAS users can run apps and locally manage their roles and permissions.

A Role-Based Access Control (RBAC) system defines and maintains the different PAS user roles:

- Org Manager, Org Auditor, Org Billing Manager
- Space Manager, Space Developer, Space Auditor

The [Orgs, Roles, Spaces, Permissions](#) topic describes the PAS user roles, and what actions they can take within the orgs and spaces they belong to. Some of these permissions depend on the values of [feature flags](#).

## Tools

Space Developer users work with their software development tools and the apps deployed on host VMs.

All PAS users use system tools such as the Cloud Foundry Command Line Interface (cf CLI), PCF Metrics, and [Apps Manager](#), a dashboard for managing PAS users, orgs, spaces, and apps.

## User Accounts

When an operator configures PAS for the first time, they specify one of the following authentication systems for PAS user accounts:

1. Internal authentication, using a new UAA database created for PAS. This system-wide UAA differs from the Ops Manager internal UAA, which only stores Ops Manager Admin accounts.
2. External authentication, through an existing identity provider accessed through SAML or LDAP protocol.

In either case, PAS user role settings are saved internally in the Cloud Controller Database, separate from the internal or external user store.

Org and Space Managers then use Apps Manager to invite and manage additional PAS users within their orgs and spaces. PAS users with proper permissions can also use the cf CLI to assign user roles.

Operators can log into Apps Manager by using the **UAA Administrator User** credentials under the **Credentials** tab of the PAS tile. These UAA Admin credentials grant them the role of Org Manager within all orgs in the deployment. The UAA Admin can also use the UAAC to create new user accounts and the cf CLI to assign user roles.

## End Users

End users are the people who log into and use the apps hosted on PAS. They do not interact directly with PAS components or interfaces. Any interactions or roles they perform within the apps are defined by the apps themselves, not Pivotal Application Service (PAS).

## User Accounts and SSO

App developers can configure apps any way they want to grant end user access individually. In a deployment with [Single Sign-On Service for Pivotal Cloud Foundry](#) installed, they can also offer end users a single login that accesses multiple apps.

The Single Sign-On (SSO) service can save user account information in an external database accessed through SAML or LDAP, or in the internal PAS user store, along with PAS User accounts.

To make the SSO service available to developers, an operator creates service plans that give login access to specific groups of end users. A Space Manager then creates a local instance of the service plan, and registers apps with it. Apps registered to the plan instance then become available through SSO to all end users covered by the plan.

## User Types Summary

The following table summarizes PCF user types, their roles, the tools they use, the System of Record (SOR) that stores their accounts, and what accounts they can provision.

User Type	Available Roles	Tools They Use	Account SOR	Accounts They Can Provision
Operators	Admin (UAA Admin, SSO Plan Admin, other system admins)	<ul style="list-style-type: none"> <li>• IaaS UI</li> <li>• PivNet</li> <li>• Ops Manager</li> <li>• cf CLI</li> <li>• UAA CLI (UAAC)</li> <li>• SSO Dashboard</li> <li>• Marketplace</li> </ul>	Ops Manager user store through UAA or External store through SAML	Operators and PAS Users
PAS Users	<ul style="list-style-type: none"> <li>• UAA Administrator</li> <li>• Org Manager</li> <li>• Org Auditor</li> <li>• Org Billing Manager</li> <li>• Space Manager</li> <li>• Space Developer</li> <li>• Space Auditor</li> </ul>	<ul style="list-style-type: none"> <li>• cf CLI</li> <li>• CAPI</li> <li>• Apps Manager</li> <li>• PCF Metrics</li> <li>• Marketplace</li> </ul>	PAS user store through UAA or External store through SAML or LDAP	PAS Users within permitted orgs and spaces, and End Users

End Users	Defined by apps they use	Hosted apps	Individual apps <i>or</i> PAS user store through SSO	
-----------	--------------------------	-------------	------------------------------------------------------------	--

## Retrieving Credentials from Your Deployment

This topic describes how the credentials for your Pivotal Cloud Foundry (PCF) deployment are stored and how you can access them.

- **What credentials does PCF store?**
  - Many PCF components use credentials to authenticate connections, and PCF installations often have hundreds of active credentials. This includes certificates, virtual machine (VM) credentials, and credentials for jobs running on the VMs.
- **Where does PCF store these credentials?**
  - PCF stores credentials in either the Ops Manager database or [BOSH CredHub](#). In PCF v1.11 and later, the BOSH Director VM includes a co-located CredHub instance. Ops Manager, Pivotal Application Service (PAS), and service tiles running on PCF can use this CredHub instance to store their credentials. For example, in PCF v1.12, PAS began migrating its credentials to CredHub. See the [PAS Release Notes](#) for a full list.
- **When do I need to access these credentials?**
  - You may need to access credentials for Ops Manager, PAS, and service tiles as part of regular administrative tasks in PCF, including troubleshooting. Many procedures in this documentation require you to retrieve credentials.
- **How can I retrieve credentials?**
  - The workflow for retrieving credentials depends on where they are stored. See the procedures below.

### Retrieve Credentials Stored in BOSH CredHub

To retrieve credentials from CredHub using the Ops Manager API, do the following:

1. Perform the procedures in the [Using the Ops Manager API](#) topic to authenticate and access the Ops Manager API.
2. Use the Ops Manager API to retrieve a list of deployed products:

```
$ curl "https://OPS-MAN-FQDN/api/v0/deployed/products" \
-X GET \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
```

Replace `UAA-ACCESS-TOKEN` with the access token recorded in the previous step.

3. In the response to the above request, locate the `guid` for the product from which you want to retrieve credentials. For example, if you want to retrieve PAS credentials, find the `installation_name` starting with `cf-` and copy its `guid`.
4. Run the following `curl` command to list the names of the credentials stored in CredHub for the product you selected. If you already know the name of the credential, you can skip this step.

```
$ curl "https://OPS-MAN-FQDN/api/v0/deployed/products/PRODUCT-GUID/variables" \
-X GET \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
```

Replace `PRODUCT-GUID` with the value of `guid` from the previous step.

5. Run the following `curl` command to view the credential:

```
$ curl "https://OPS-MAN-FQDN/api/v0/deployed/products/PRODUCT-GUID/variables?name=VARIABLE-NAME" \
-X GET \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
```

Replace `VARIABLE-NAME` with the name of the credential you want to retrieve.

### Retrieve Credentials Stored in the Ops Manager Database

To retrieve credentials stored in the Ops Manager database and not CredHub, use the Ops Manager UI or API as outlined in the procedures below.

## Retrieve Credentials Using the Ops Manager UI

1. From Ops Manager, select the product tile for which you want to retrieve credentials.
2. Click the **Credentials** tab.
3. Locate the credential that you need and click **Link to Credential**.

## Retrieve Credentials Using the Ops Manager API

1. Perform the procedures in the [Using the Ops Manager API](#) topic to authenticate and access the Ops Manager API.
2. Use the Ops Manager API to retrieve a list of deployed products:

```
$ curl "https://OPS-MAN-FQDN/api/v0/deployed/products" \
-X GET \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
```

Replace `UAA-ACCESS-TOKEN` with the access token recorded in the previous step.

3. In the response to the above request, locate the `guid` for the product from which you want to retrieve credentials. For example, if you want to retrieve PAS credentials, find the `installation_name` starting with `cf-` and copy its `guid`.
4. Run the following `curl` command to list references for the credentials stored in Ops Manager for the product you selected. If you already know the reference for the credential, you can skip this step.

```
$ curl "https://OPS-MAN-FQDN/api/v0/deployed/products/PRODUCT-GUID/credentials" \
-X GET \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
```

Replace `PRODUCT-GUID` with the value of `guid` from the previous step.

5. Run the following `curl` command to view the credential:

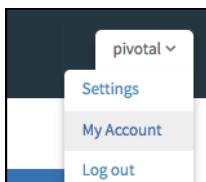
```
$ curl "https://OPS-MAN-FQDN/api/v0/deployed/products/PRODUCT-GUID/credentials/CREDENTIAL-REFERENCE" \
-X GET \
-H "Authorization: Bearer UAA-ACCESS-TOKEN"
```

Replace `CREDENTIAL-REFERENCE` with the name of the credential you want to retrieve.

## Changing Ops Manager Credentials

### Ops Manager Password

1. Log in to Ops Manager and navigate to **My Account**. You can access this at <https://OPS-MAN-FQDN/uaa/profile>.



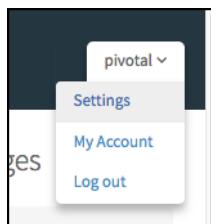
2. Navigate to **Change Password**. You can access this at [https://OPS-MAN-FQDN/uaa/change\\_password](https://OPS-MAN-FQDN/uaa/change_password).
3. Enter your current password and a new password.

The screenshot shows the Pivotal account settings interface. At the top is the Pivotal logo. Below it is a section titled "Account Settings". Under "Account Settings" is a "Profile" section containing an email address (admin@test.org) and five asterisks for a password. To the right of the email are "Change Email" and "Change Password" links. Below the profile is a "Third Party Access" section stating "You have not yet authorized any third party applications."

## Ops Manager Decryption Passphrase

You must have the existing passphrase to update the decryption passphrase.

1. Log in to Ops Manager, and navigate to **Settings**. You can access this at [https://OPS-MAN-FQDN/encryption\\_passphrase/edit](https://OPS-MAN-FQDN/encryption_passphrase/edit).



2. In the **Decryption Passphrase** panel, enter your current decryption passphrase and the new decryption passphrase, then click **Save**.

The screenshot shows the 'PCF Ops Manager' interface with a dark header bar. Below it, a navigation bar includes a back arrow to 'Installation Dashboard' and a lock icon for 'Settings'. A sidebar on the left lists several settings categories: 'Decryption Passphrase' (selected), 'Authentication Method', 'External API Access', 'Proxy Settings', 'Export Installation Settings', and 'Advanced'. The main content area is titled 'Update Decryption Passphrase' and contains three input fields with asterisks indicating required fields: 'Current decryption passphrase\*', 'New decryption passphrase\*', and 'New decryption passphrase confirmation\*'. A blue 'Update' button is at the bottom right of the form.

## S3 Compatible Blobstore Credentials

If you use an S3 compatible blobstore, you can rotate your blobstore credentials from the **Director Config** panel of the BOSH Director tile. After entering new credentials and clicking **Review Pending Changes**, then **Apply Changes**, BOSH recreates the VMs in your deployment to apply the new credentials.

1. From a browser, navigate to <https://OPS-MAN-FQDN/> and log in to Ops Manager.
2. From the Installation Dashboard, click the BOSH Director tile.
3. From the **Director Config** panel, select the **Recreate all VMs** checkbox.

## Director Config

NTP Servers (comma delimited)\*



JMX Provider IP Address

Bosh HM Forwarder IP Address

Enable VM Resurrector Plugin

Enable Post Deploy Scripts

Recreate all VMs

This will force BOSH to recreate all VMs on the next deploy. Persistent disk will be preserved

4. For **Access Key**, enter a new access key.

Blobstore Location

Internal  
 S3 Compatible Blobstore

S3 Endpoint\*

Bucket Name\*

Access Key\*

Secret Key\*

V2 Signature  
 V4 Signature

Region\*

5. For **Secret Key**, enter a new secret key.
6. Click **Review Pending Changes**, then **Apply Changes**.
7. Clear the **Recreate all VMs** checkbox.

## CredHub

### Overview

CredHub is a component designed for centralized credential management in Pivotal Cloud Foundry (PCF). It is a single component that can address several scenarios in the PCF ecosystem. At the highest level, CredHub centralizes and secures credential generation, storage, lifecycle management, and access.

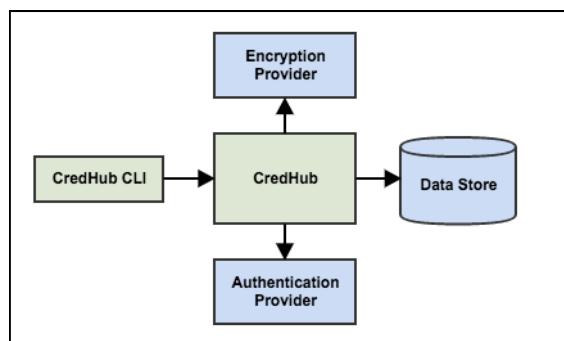
### What Can CredHub Do?

CredHub performs a number of different functions to help generate and protect the credentials in your PCF deployment.

- Securing data for storage
- Authentication
- Authorization
- Access and change logging
- Data typing
- Credential generation
- Credential metadata
- Credential versioning

### Application Architecture

CredHub consists of a REST API and a CLI. The REST API conforms to the Config Server API spec. CredHub is an OAuth2 resource server that integrates with User Account Authentication (UAA) to provide core authentication and federation capabilities.



### CredHub in PCF

A PCF deployment stores credentials in the following locations:

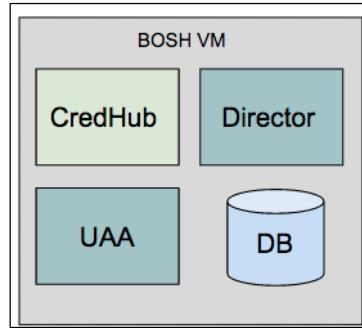
- **BOSH CredHub:** Colocated with the BOSH Director on a single VM. This CredHub instance stores credentials for the BOSH Director.
- **Runtime CredHub:** Deployed as an independent service and stores service instance credentials.

### BOSH CredHub

In PCF, BOSH Director VM includes a CredHub job. This provides a lightweight credential storage instance for the BOSH Director. The BOSH Director, Pivotal Application Service (PAS), and other tiles store credentials in BOSH CredHub. For more information, see [Retrieving Credentials from Your Deployment](#).

 Note: This configuration does not provide high availability.

In this colocated deployment architecture, the BOSH Director, CredHub, UAA, and the Director database are all installed on a single BOSH VM, as shown in the following diagram:

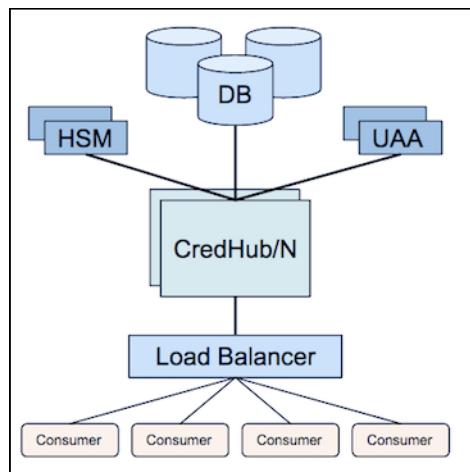


## Runtime CredHub

The PAS tile deploys CredHub as an independent service on its own VM. This provides a highly available credential storage instance for securing service instance credentials. For more information, see [Securing Service Instance Credentials with Runtime CredHub](#).

CredHub is a stateless application, so you can scale it to multiple instances that share a common database cluster and encryption provider.

With CredHub as a service, the load balancer and external databases communicate directly with the CredHub VMs, as shown in the following diagram:



## Using CredHub to Store Credentials for Service Tiles

If you develop a service tile for PCF and want to store its credentials in BOSH CredHub, see the [CredHub](#) section of the *Tile Developer Guide*.

## CredHub Credential Types

Credentials exist in multiple places in the PCF ecosystem. PCF components use credentials to authenticate connections between components. PCF installations often have hundreds of active credentials. Leaked credentials are common causes of data and security breaches, so managing them securely is very important.

For more information, read [CredHub Credential Types](#).

## Backing Up and Restoring CredHub Instances

The CredHub application does not hold state, but you must ensure its dependent components are backed up. Redundant backups can help prevent data loss if an individual component fails. For more information, read [Backing Up and Restoring CredHub Instances](#).



## CredHub Credential Types

This topic describes the different credential types supported by CredHub.

CredHub supports different types of credentials to simplify generating and managing multi-part credentials. For example, a TLS certificate contains three parts: the root certificate authority (CA), the certificate, and the private key. CredHub supports all three parts, which helps keep connection requests from being rejected erroneously.

CredHub supports the following credential types:

Type	Description
value	A single string value for arbitrary configurations and other non-generated or validated strings.
json	An arbitrary JSON object for static configurations with many values.
user	Three string values for username, password, and password hash.
password	A single string value for passwords and other random string credentials. Values for this type can be automatically generated.
certificate	An object containing a root CA, certificate, and private key. Use this type for key pair applications that utilize a certificate, such as TLS connections. Values for this type can be automatically generated.
rsa	An object containing an RSA public key and private key without a certificate. Values for this type can be automatically generated.
ssh	An object containing an SSH-formatted public key and private key. Values for this type can be automatically generated.

Each credential type supports distinct parameters for customizing how credentials are generated. These include minimum password lengths, required characters, and certificate fields. For more information, see the [Generate Credentials](#) section of the CredHub API documentation.

For every credential type, secret values are encrypted before storage. For instance, the private key of a certificate-type credential and the password of a user-type credential are encrypted before storage. For JSON and Value type credentials, the full contents are encrypted before storage.

## Security for Apps and Services

This topic links to topics that describe how PCF and PCF users manage security for apps and service instances.

- [Application Security Groups ↗](#): Describes how ASGs work and how to manage them in all versions of Cloud Foundry, including PCF.
- [Configuring SSH Access for PCF ↗](#): Explains how to configure PCF to allow SSH access to app instances, for debugging.
- [Restricting App Access to Internal PCF Components ↗](#): Details how to create Application Security Groups (ASGs), rules that allow internal outgoing communications from all apps in PCF, or the apps running in the same space.
- [Configuring Application Security Groups for Email Notifications ↗](#): Describes how to define an ASG to enable app-generated notifications.
- [Trusted System Certificates ↗](#): Explains where applications can find trusted system certificates.
- [Managing Access to Service Plans ↗](#): Describes how to enable or disable access to service plans for a subset of users.
- [Delivering Service Credentials to an Application ↗](#): Provides documentation on how to bind apps to service instances, which generate the credentials that enable the apps to use the service.
- [Managing Service Keys ↗](#): Explains to create and manage service keys that enable apps to use service instances.

## Compliance and Other Security-Related Topics

This section contains topics related to compliance and securing Pivotal Cloud Foundry (PCF) deployments.

### Compliance

- [NIST Controls and PCF](#): Provides a dedicated site that assesses Pivotal Cloud Foundry against NIST SP 800-53(r4) Controls.
- [General Data Protection Regulation](#): Provides an overview of the General Data Protection Regulation (GDPR) and where Pivotal Cloud Foundry (PCF) may store personal data

### Other Security-Related Topics

- [Security-Related PCF Tiles and Add-ons](#): Provides links to other security-related services available for PCF deployments.
- [Security Guidelines for Your IaaS Provider](#): Provides links to security-related documentation related to different IaaS providers. The topics may come from third party providers, and are not necessarily owned or maintained by Pivotal. Pivotal provides them for your convenience, but cannot guarantee the accuracy or currency of these documents.

## General Data Protection Regulation

This topic provides an overview of the General Data Protection Regulation (GDPR) and where Pivotal Cloud Foundry (PCF) may store personal data.

### Overview

GDPR came into effect on May 25, 2018 and impacts any company processing the data of EU citizens or residents, even if the company is not EU-based. The GDPR sets forth how companies should handle privacy issues, securely store data, and respond to security breaches.

### Understand Personal Data

The GDPR grants data subjects certain rights, such as the right to obtain a copy of their personal data, object to the processing of personal data, and the right to have their personal data erased. Organizations subject to GDPR need to ensure that they can address and respond to requests by data subjects if they are processing their personal data.

Article 4, Section 1 of the GDPR defines personal data as follows:

'personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;

For more information, see the GDPR [text ↗](#).

Personal data can be collected, stored, and processed in a PCF deployment. Pivotal has performed a review of PCF components and determined that personal data may reside in the following areas:

- [User Account and Authentication \(UAA\)](#)
- [Cloud Foundry API](#)
- [Routing](#)
- [Diego](#)
- [Notifications Service](#)

### Where Personal Data May Reside

The following sections explain how different PCF components collect personal data.

#### User Account and Authentication (UAA)

UAA is an open-source Cloud Foundry component that provides identity management features and identity-based security for applications and APIs. For more information, see [User Account and Authentication ↗](#).

GDPR	Workflow	What personal data is collected?	When is it collected?	Where is it stored?	How is it processed?	Who has access to it?
	User registers	<ul style="list-style-type: none"><li>• Username</li><li>• Email address</li><li>• First name (optional)</li><li>• Last name (optional)</li><li>• User ID (UAA GUID, generated)</li></ul>	User registration submission	UAA DB	Stored in UAA DB	<ul style="list-style-type: none"><li>• End user</li><li>• UAA administrators</li></ul>
		<ul style="list-style-type: none"><li>• Username</li><li>• Email address</li></ul>				

Business Initiation	Just-in-time provisioning: create user on user login	<ul style="list-style-type: none"> <li>First name (optional)</li> <li>Last name (optional)</li> <li>User ID (UAA GUID, generated)</li> <li>Additional attributes as defined by the organization</li> </ul>	User login	UAA DB	Stored in UAA DB	UAA administrators
	Admin user makes a creation API call	<ul style="list-style-type: none"> <li>Username</li> <li>Email address</li> <li>First name (optional)</li> <li>Last name (optional)</li> <li>User ID (UAA GUID, generated)</li> <li>Additional attributes as defined by the organization</li> </ul>	Admin API call	UAA DB	Stored in UAA DB	UAA administrators
Business Execution	User self-updates profile	<ul style="list-style-type: none"> <li>Email address</li> <li>First name (optional)</li> <li>Last name (optional)</li> </ul>	User registration submission	UAA DB	Stored in UAA DB	<ul style="list-style-type: none"> <li>End user</li> <li>UAA administrators</li> </ul>
	Just-in-time provisioning: user update	<ul style="list-style-type: none"> <li>Email address</li> <li>First name (optional)</li> <li>Last Name (optional)</li> <li>Additional attributes as defined by the organization</li> </ul>	User login	UAA DB	Stored in UAA DB	UAA administrators
	User logs in	<ul style="list-style-type: none"> <li>Current account cookie (generated)</li> <li>Saved account cookie (generated)</li> </ul>	User login	User browser	By UAA	<ul style="list-style-type: none"> <li>End user</li> <li>UAA login page</li> </ul>
	Admin user makes an update API call	<ul style="list-style-type: none"> <li>Email address</li> <li>First name (optional)</li> <li>Last name (optional)</li> <li>Additional attributes as defined by the organization</li> </ul>	Admin API call	UAA DB	Stored in UAA DB	UAA administrators
Delete User Flow	Admin user makes a hard delete API call	n/a	n/a	n/a	Deleted from UAA DB	UAA administrators
	Admin user makes a deactivation API call	n/a	n/a	n/a	Soft delete (records still held in database but user cannot login)	UAA administrators
Reports/Logs	Event or debug logs	Any information	When event happens	UAA logs	Depends on setup of <a href="#">Loggregator</a> and log forwarding	BOSH administrators

## Cloud Foundry API

The Cloud Foundry API release contains several components, including the Cloud Controller. For more information, see the Cloud Foundry API release [README](#).



GDPR	Workflow	What personal data is collected?	When is it collected?	Where is it stored?	How is it processed?	Who has access to it?	How long is it kept?
Business Initiation	User makes a request for the first time	User ID	The first time a user makes a request to the API	Cloud Controller DB	It is used to identify permissions for the user	PCF operator	As long as the user is part of the system
Business Execution	Troubleshooting API requests	<ul style="list-style-type: none"> <li>• User ID</li> <li>• User agent</li> <li>• IP address</li> </ul>	On each request	<ul style="list-style-type: none"> <li>• Local VM: component and logs</li> <li>• Log aggregator used by PCF operator</li> </ul>	n/a	PCF operator	<ul style="list-style-type: none"> <li>• Local VM: 4 week maximum by default</li> <li>• Log aggregator as configured by PCF operator</li> </ul>
Audit Trails	Audit what changes a user makes	<ul style="list-style-type: none"> <li>• Name</li> <li>• User ID</li> <li>• Email address</li> </ul>	On specific API requests that mutate the state of resources	Audit Event Table in the Cloud Controller DB	n/a	<ul style="list-style-type: none"> <li>• PCF operator</li> <li>• Users that can view the resource that had an audited change</li> </ul>	31 days
	Audit what changes a user makes	<ul style="list-style-type: none"> <li>• IP Address</li> <li>• Email address</li> <li>• User ID</li> <li>• Username</li> </ul>	On each request	<ul style="list-style-type: none"> <li>• Local VM: CEF logs</li> <li>• Log aggregator used by PCF operator</li> </ul>	n/a	PCF operator	<ul style="list-style-type: none"> <li>• Local VM: 4 week maximum by default</li> <li>• Log aggregator as configured by PCF operator</li> </ul>
	Audit what user created a resource	<ul style="list-style-type: none"> <li>• Name</li> <li>• User ID</li> <li>• Email address</li> </ul>	When API resources are created	As part of the resource row in Cloud Controller DB	n/a	<ul style="list-style-type: none"> <li>• PCF operator</li> <li>• Users that can view the resource</li> </ul>	As long as the resource exists

## Routing

By default, the [Gorouter](#) logs include the `X-Forwarded-For` header, which may include the originating client IP. Under GDPR, client IP addresses should be considered personal data.

### Disable Client IP Logging

In Pivotal Application Service (PAS) v2.0 and later and Elastic Runtime v1.12, operators can disable logging of client IP addresses in the Gorouter.

To disable logging of client IP addresses, do the following:

1. Navigate to the Ops Manager Installation Dashboard and click the PAS or Elastic Runtime tile.
2. Click **Networking**.
3. Under **Logging of Client IPs in CF Router**, select one of the two options:
  - If the source IP address exposed by your load balancer is its own IP address, select **Disable logging of X-Forwarded-For header only**.
  - If the source IP address exposed by your load balancer belongs to the downstream client, select **Disable logging of both source IP and X-Forwarded-For header**.
4. Click **Save**.

5. Return to the Ops Manager Installation Dashboard, click **Review Pending Changes**, and click **Apply Changes** to redeploy.

## Diego

Diego is the container management system for PCF. For more information, see [Diego Components and Architecture](#).

GDPR	Workflow	What personal data is collected?	When is it collected?	Where is it stored?	How is it processed?	Who has access to it?	How can I delete it?
Business Execution	Executing apps and tasks	No personal data is collected explicitly, but personal data may be encoded in app file contents or runtime metadata such as environment variables or start commands.	Runtime metadata is collected when Cloud Controller submits work specification to the Diego BBS API. File contents are collected when Diego schedules an app or a task on a Diego cell.	Runtime metadata is stored in the Diego BBS DB. App file contents are cached on Diego cells.	Runtime metadata is used to start processes inside app instance containers and to configure their environment. App file contents are presented as part of the app instance container filesystem.	Platform operators and other developers with access to the Cloud Controller space containing that app can view the data.	<ul style="list-style-type: none"> <li>To delete the runtime metadata stored in the Diego BBS DB, stop the app or cancel the task that includes that data.</li> <li>To delete the app file contents stored in the running app and task containers, stop the app or cancel the task to destroy the containers. To destroy the app file contents stored in the download cache on the Diego cells, recreate the Diego cell VMs.</li> </ul>
Reports/Logs	SSH proxy logs Cloud Foundry user access.	UAA user name and ID	When the user authenticates for SSH access to an app.	The data is stored in a log file collocated with the SSH proxy instance handling the authentication request. This log file may also have its contents forwarded to a remote syslog destination.	No processing of the local log file is done automatically. If the log file contents are forwarded to a log aggregation service, they may be parsed and processed arbitrarily.	Only platform operators have access to the local log file. Platform operators or auditors may have access to these log contents in a downstream log aggregation service.	<p>To delete the log lines containing the user ID, perform the following steps:</p> <ol style="list-style-type: none"> <li>Run <code>bosh recreate</code> on the VMs hosting the SSH proxy processes to remove all the logs on ephemeral disk.</li> <li><code>bosh ssh</code> into the VMs hosting the SSH proxy processes and remove specific log lines containing user IDs.</li> <li>Scrub corresponding log lines from any log aggregation service.</li> </ol>

## Notifications Service

The Notifications Service enables operators to configure components of Cloud Foundry to send emails to end users. For more information, see [Getting Started with the Notifications Service](#).

GDPR	Workflow	What personal data is collected?	When is it collected?	Where is it stored?	How is it processed?	Who has access to it?
Business Execution	Sending email to UAA users	User ID	First email sent	The <code>receipts</code> table in the Notifications database	Stored in the Notifications database	Notifications operator making a database query
	UAA user unsubscribes globally	User ID	When the UAA user unsubscribes	The <code>global_unsubscribes</code> table in the Notifications database	Stored in the Notifications database	Notifications operator making a database query
	UAA user unsubscribes from a specific kind of email	User ID	When the UAA user unsubscribes	The <code>unsubscribes</code> table in the Notifications database	Stored in the Notifications database	Notifications operator making a database query
	UAA user unsubscribes from a campaign in the v2 API	User ID	When the UAA user unsubscribes	The <code>unsubscribes</code> table in the Notifications database	Stored in the Notifications database	Notifications operator making a database query
Reports/Logs	UAA user unsubscribes	User email address	When the UAA user unsubscribes	Log output	<a href="#">Loggregator</a>	Loggregator Firehose users

## Security-Related PCF Tiles and Add-ons

This section provides links to other security-related services available for Pivotal Cloud Foundry (PCF) deployments.

For information about other security-related tiles provided by Pivotal partners, refer to [Pivotal Network](#) and [Pivotal Documentation](#).

### Tiles

This section provides links to security-related tiles available for Pivotal Cloud Foundry (PCF) deployments.

- [Single Sign-On \(SSO\)](#): Provides support for native authentication, federated single sign-on, and authorization.
- [CredHub Service Broker](#): Allows apps running on Pivotal Application Service (PAS) to access secure credentials in CredHub.
- [Compliance Scanner](#): Enables operators and auditors to conduct compliance scanning of BOSH-managed Xenial VMs against configuration hardening guidelines.

### Add-ons

This section provides links to security-related BOSH add-ons available for Pivotal Cloud Foundry (PCF) deployments.

- [File Integrity Monitoring Add-on for PCF](#): Provides File Integrity Monitoring (FIM) protection within the PCF environment. This monitoring is sometimes required for compliance purposes.
- [IPsec Add-on for PCF](#): Secures data transmissions inside Pivotal Cloud Foundry (PCF).
- [ClamAV Add-on for PCF](#): Provides antivirus protection within the PCF environment. This protection is sometimes required for compliance purposes.

## Security Guidelines for Your IaaS Provider

Pivotal Cloud Foundry supports a variety of Infrastructure as a Service (IaaS) providers. Different IaaS providers require different configuration steps to secure user data, identity information, and credentials.

Security requirements can vary broadly based on the unique configuration and infrastructure of each organization. Rather than provide specific guidance that may not apply to all use cases, Pivotal has collected links to IaaS providers' security and identity management documentation. The documents below may help you understand how your IaaS' security requirements impact your PCF deployment.

Pivotal does not endorse these documents for accuracy or guarantee that their contents apply to all PCF installations.

### How to Use This Topic

Find your IaaS provider in the list below. The documentation items linked for each IaaS may help you configure and secure your installation infrastructure.

## Amazon Web Services (AWS)

- [AWS Identity and Access Management guide ↗](#)

This guide is a reference for AWS' Identity and Access Management (IAM) features. If you're new to AWS, start here.

- [AWS identity documentation ↗](#)

- [AWS credential documentation ↗](#)

This documentation provides a general definition of IAM terms and provide best practices to help you manage IaaS users and permissions.

## Google Cloud Platform (GCP)

- [GCP authentication documentation ↗](#)

This developer-facing documentation explains general authentication guidelines for GCP.

## Microsoft Azure

- [Azure security documentation ↗](#)

This site has documentation on Azure security tools. It provides a general guide to how to manage IaaS users and credentials.

## OpenStack

- [OpenStack credential configuration ↗](#)

- [OpenStack credential creation ↗](#)

- [OpenStack deployment configuration ↗](#)

These documents provide a general reference for OpenStack service credential management.

## VMware vSphere

- [vSphere Security guide \(PDF\) ↗](#)

This guide contains best practices for securing and managing a vSphere installation.

## Buildpacks

Buildpacks provide framework and runtime support for apps. Buildpacks typically examine your apps to determine what dependencies to download and how to configure the apps to communicate with bound services.

When you push an app, Cloud Foundry automatically detects an appropriate buildpack for it. This buildpack is used to compile or prepare your app for launch.

 **Note:** Cloud Foundry deployments often have limited access to dependencies. This limitation occurs when the deployment is behind a firewall, or when administrators want to use local mirrors and proxies. In these circumstances, Cloud Foundry provides a [Buildpack Packager](#) app.

## Using Buildpacks

The following topics discuss different usage scenarios for buildpacks in Cloud Foundry:

- [How Buildpacks Work in Cloud Foundry](#)
- [Stack Association](#)
- [Pushing an Application with Multiple Buildpacks](#)
- [Using a Proxy](#)
- [Supported Binary Dependencies](#)
- [Production Server Configuration](#)

## System Buildpacks

For information about the existing buildpacks that Cloud Foundry supports, see [System Buildpacks](#).

## Community Buildpacks

You can find a list of unsupported, community-created buildpacks here: [cf-docs-contrib](#).

## Customizing and Developing Buildpacks

For information about customizing existing buildpacks and developing new buildpacks, see [Customizing and Developing Buildpacks](#).

## Using Continuous Integration for Buildpacks

For information about updating and releasing a new version of a Cloud Foundry buildpack through the Cloud Foundry Buildpacks Team Concourse pipeline, see [Using CI for Buildpacks](#). You can use this as a model when working with Concourse to build and release new versions of your own buildpacks.

## Using Buildpacks

This topic provides links to additional information about using buildpacks. Each of the following are applicable to all supported buildpack languages and frameworks:

- [How Buildpacks Work in Cloud Foundry](#)
- [Stack Association](#)
- [Pushing an Application with Multiple Buildpacks](#)
- [Using a Proxy](#)
- [Supported Binary Dependencies](#)
- [Production Server Configuration](#)

## How Buildpacks Work in Cloud Foundry

This topic describes how buildpacks are structured and detected in Cloud Foundry.

### Buildpack Scripts

A buildpack repository may contain the following five scripts in the `bin` directory:

- `bin/detect` determines whether or not to apply the buildpack to an app.
- `bin/supply` provides dependencies for an app.
- `bin/finalize` prepares the app for launch.
- `bin/release` provides feedback metadata to Cloud Foundry indicating how the app should be executed.
- `bin/compile` is a deprecated alternative to `bin/supply` and `bin/finalize`.

The `bin/supply` and `bin/finalize` scripts replace the deprecated `bin/compile` script. Older buildpacks may still use `bin/compile` with the latest version of Cloud Foundry. In this case, applying multiple buildpacks to a single app is not supported. Similarly, newer buildpacks may still provide `bin/compile` for compatibility with Heroku and older versions of Cloud Foundry.

The `bin/supply` script is required for non-final buildpacks. The `bin/finalize` (or `bin/compile`) script is required for final buildpacks.

 **Note:** In this document, the terms *non-final buildpack* and *final buildpack*, or *last buildpack*, are used to describe the process of applying multiple buildpacks to an app. See the following example: `cf push APP-NAME -b FIRST-BUILDPACK -b SECOND-BUILDPACK -b FINAL-BUILDPACK`.

 **Note:** If you use only one buildpack for your app, this buildpack behaves as a final, or last, buildpack.

 **Note:** When using multi-buildpack support, the last buildpack in order is the final buildpack, and is able to make changes to the app and determine a start command. All other specified buildpacks are non-final and only supply dependencies.

### bin/detect

The `detect` script determines whether or not to apply the buildpack to an app. The script is called with one argument, the `build` directory for the app. The `build` directory contains the app files uploaded when a user performs a `cf push`.

The `detect` script returns an exit code of `0` if the buildpack is compatible with the app. In the case of system buildpacks, the script also prints the buildpack name, version, and other information to `STDOUT`.

The following is an example `detect` script that checks for a Ruby app based on the existence of a `Gemfile`:

```
#!/usr/bin/env ruby

gemfile_path = File.join ARGV[0], "Gemfile"

if File.exist?(gemfile_path)
 puts "Ruby"
 exit 0
else
 exit 1
end
```

Optionally, the buildpack `detect` script can output additional details provided by the buildpack developer. This includes buildpack versioning information and a list of configured frameworks and their associated versions.

The following is an example of the detailed information returned by the Java buildpack:

```
java-buildpack=v3.0-https://github.com/cloudfoundry/java-buildpack.git#3bd15e1 open-jdk-jre=1.8.0_45 spring-auto-reconfiguration=1.7.0_RELEASE tomcat-access-logging-support=2.4.0_RELEASE tomcat-in-
```

 **Note:** Cloud Foundry detects only one buildpack by default. When multiple buildpacks are desired, you must explicitly specify them.

For more information, see the [Buildpack Detection](#) section below.

## bin/supply

The `supply` script provides dependencies for the app and runs for all buildpacks. All output sent to `STDOUT` is relayed to the user through the Cloud Foundry Command Line Interface (cf CLI).

The script is run with four arguments:

- The `build` directory for the app
- The `cache` directory, which is a location the buildpack can use to store assets during the build process
- The `deps` directory, which is where dependencies provided by all buildpacks are installed
- The `index`, which is a number that represents the ordinal position of the buildpack

The `supply` script stores dependencies in `deps` / `index`. It may also look in other directories within `deps` to find dependencies supplied by other buildpacks.

The `supply` script must not modify anything outside of the `deps` / `index` directory. Staging may fail if such modification is detected.

The `cache` directory provided to the `supply` script of the final buildpack is preserved even when the buildpack is upgraded or otherwise changes. The `finalize` script also has access to this cache directory.

The `cache` directories provided to the `supply` scripts of non-final buildpacks are cleared if those buildpacks are upgraded or otherwise change.

The following is an example of a simple `supply` script:

```
#!/usr/bin/env ruby

#sync output

$stdout.sync = true

build_path = ARGV[0]
cache_path = ARGV[1]
deps_path = ARGV[2]
index = ARGV[3]

install_ruby

private

def install_ruby
 puts "Installing Ruby"

 #!!! build tasks go here !!!
 #download ruby
 #install ruby
end
```

## bin/finalize

The `finalize` script prepares the app for launch and runs only for the last buildpack. All output sent to `STDOUT` is relayed to the user through the cf CLI.

The script is run with four arguments:

- The `build` directory for the app
- The `cache` directory, which is a location the buildpack can use to store assets during the build process
- The `deps` directory, which is where dependencies provided by all buildpacks are installed
- The `index`, which is a number that represents the ordinal position of the buildpack

The `finalize` script may find dependencies installed by the `supply` script of the same buildpack in `deps` / `index`. It may also look in other directories within `deps` to find dependencies supplied by other buildpacks.

The `cache` directory provided to the `finalize` script is preserved even when the buildpack is upgraded or otherwise changes. The `supply` script of the

same buildpack also has access to this cache directory.

The following is an example of a simple `finalize` script:

```
#!/usr/bin/env ruby

#sync output

$stdout.sync = true

build_path = ARGV[0]
cache_path = ARGV[1]
deps_path = ARGV[2]
index = ARGV[3]

setup_ruby

private

def setup_ruby
 puts "Configuring your app to use Ruby"

 # !!! build tasks go here !!!
 # setup ruby
end
```

## bin/compile (Deprecated)

The `compile` script is deprecated. It encompasses the behavior of the `supply` and `finalize` scripts for single buildpack apps by using the `build` directory to store dependencies.

The script is run with two arguments:

- The `build` directory for the app
- The `cache` directory, which is a location the buildpack can use to store assets during the build process

During the execution of the `compile` script, all output sent to `STDOUT` is relayed to the user through the cf CLI.

## bin/release

The `release` script provides feedback metadata to Cloud Foundry indicating how the app should be executed. The script is run with one argument, the `build` directory. The script must generate a YAML file in the following format:

```
default_process_types:
 web: start_command.filename
```

`default_process_types` indicates the type of app being run and the command used to start it. This start command is used if a start command is not specified in the `cf push` or in a Procfile.

At this time, only the `web` type of apps is supported.

**Note:** To define environment variables for your buildpack, add a Bash script to the `.profile.d` directory in the root folder of your app.

The following example shows what a `release` script for a Rack app might return:

```
default_process_types:
 web: bundle exec rackup config.ru -p $PORT
```

**Note:** The `web` command runs as `bash -c COMMAND` when Cloud Foundry starts your app. Refer to [the command attribute](#) section for more information about custom start commands.

## Droplet Filesystem

The buildpack staging process extracts the droplet into the `/home/vcap` directory inside the instance container and creates the following filesystem tree:

```
app/
deps/
logs/
tmp/
staging_info.yml
```

The `app` directory includes the contents of the `build` directory, and `staging_info.yml` contains the staging metadata saved in the droplet.

## Buildpack Detection

When you push an app, Cloud Foundry uses a detection process to determine a single buildpack to use. For general information about this process, see [How Apps Are Staged](#).

During staging, each buildpack has a position in a priority list. You can retrieve this position by running `cf buildpacks`.

Cloud Foundry checks if the buildpack in position 1 is a compatible buildpack. If the position 1 buildpack is not compatible, Cloud Foundry moves on to the buildpack in position 2. Cloud Foundry continues this process until the correct buildpack is found.

If no buildpack is compatible, the `cf push` command fails with the following error:

```
None of the buildpacks detected a compatible application
Exit status 222
Staging failed: Exited with status 222

FAILED
NoAppDetectedError
```

For a more detailed account of how Cloud Foundry interacts with the buildpack, see the [Sequence of Interactions](#) section below.

## Sequence of Interactions

This section describes the sequence of interactions between the Cloud Foundry platform and the buildpack. The sequence of interactions differs depending on whether the platform [skips](#) or [performs](#) buildpack detection.

### No Buildpack Detection

Cloud Foundry skips buildpack detection if the developer specifies one or more buildpacks in the app manifest or in the `cf push APP-NAME -b BUILDPACK-NAME` CLI command.

If you explicitly specify buildpacks, Cloud Foundry performs the following interactions:

1. For each buildpack except the last buildpack, the platform does the following:
  - a. Creates the `deps / index` directory
  - b. Runs `/bin/supply` with the `build`, `cache`, and `deps` directories and the buildpack `index`
  - c. Accepts any modification of the `deps / index` directory
  - d. Accepts any modification of the `cache` directory
  - e. May disallow modification of any other directories
2. For the last buildpack, the platform does the following:
  - a. If `/bin/finalize` is present:
    - i. Creates the `deps / index` directory if it does not exist
    - ii. If `/bin/supply` is present, runs `/bin/supply` with the `build`, `cache`, and `deps` directories and the buildpack `index`
    - iii. Accepts any modification of the `deps / index` directory

- iv. May disallow modification of the `build` directory
  - v. Runs `/bin/finalize` with the `build`, `cache`, and `deps` directories and the buildpack `index`
  - vi. Accepts any modification of the `build` directory
- b. If `/bin/finalize` is not present:
    - i. Runs `/bin/compile` with the `build` and `cache` directories
    - ii. Accepts any modification of the `build` directory
  - c. Runs `/bin/release` to determine staging information

At the end of this process, the `deps` directory is included at the root of the droplet, adjacent to the `app` directory.

## Buildpack Detection

Cloud Foundry performs buildpack detection if the developer does not specify one or more buildpacks in the app manifest or in the

```
cf push APP-NAME -b BUILDPACK-
NAME
```

 **Note:** Cloud Foundry detects only one buildpack to use with the app.

If the platform performs detection, it does the following:

1. Runs `/bin/detect` for each buildpack
2. Selects the first buildpack with a `/bin/detect` script that returns a zero exit status
3. If `/bin/finalize` is present:
  - a. Creates the `deps / index` directory if it does not exist
  - b. If `/bin/supply` is present, runs `/bin/supply` with the `build`, `cache`, and `deps` directories and the buildpack `index`
  - c. Accepts any modification of the `deps / index` directory
  - d. May disallow modification of the `build` directory
  - e. Runs `/bin/finalize` on the `build`, `cache`, and `deps` directories
  - f. Accepts any modification of the `build` directory
4. If `/bin/finalize` is not present:
  - a. Runs `/bin/compile` on the `build` and `cache` directories
  - b. Accepts any modification of the `build` directory
5. Runs `/bin/release` to determine staging information

At the end of this process, the `deps` directory is included at the root of the droplet, adjacent to the `app` directory.

## Stack Association

This topic describes the stack association feature for Pivotal Cloud Foundry (PCF) buildpacks.

 **Note:** This functionality is supported as of CAPI [v1.58.0](#) and cf CLI [v6.39.0](#).

 **warning:** To avoid security exposure, ensure that you migrate your apps and custom buildpacks to use the `cflinuxfs3` stack based on Ubuntu 18.04 (Bionic Beaver). `cflinuxfs2` is based on Ubuntu 14.04 (Trusty Tahr), which reaches end of general support (EOGS) in April 2019.

## Overview

Each buildpack in your PCF deployment is associated with a stack. You can see this when you run `cf buildpacks`: there is a `stack` column in the output that shows a corresponding stack for each buildpack. See the following example:

```
$ cf buildpacks
Getting buildpacks...
buildpack position enabled locked filename stack
staticfile_buildpack 1 true false staticfile_buildpack-cached-cflinuxfs3-v1.4.29.zip cflinuxfs3
java_buildpack_offline 2 true false java-buildpack-offline-cflinuxfs3-v4.12.1.zip cflinuxfs3
ruby_buildpack 3 true false ruby_buildpack-cached-cflinuxfs3-v1.7.21.zip cflinuxfs3
...
```

Because of this stack association, buildpacks do not have to be uniquely named. This helps in managing similar buildpacks that are compatible with different stacks.

The [buildpack packager](#) includes a `-stack` option. If you use this option and upload a buildpack to PCF, the Cloud Controller detects the stack association and creates a stack record for the buildpack.

## Buildpacks without a Stack Record

Some buildpacks may have a missing stack record. For example, if you uploaded a custom buildpack before PCF introduced stack association. The output of `cf buildpacks` shows a blank `stack` column if the buildpack does not have a stack record.

In this case, you must manually assign a stack to the buildpack. To do this, run `cf update-buildpack BUILDPACK-NAME --assign-stack stack`.

Buildpacks with a missing stack record will continue to work, but are more manageable when the stack record is present.

Consider the following behavior when pushing apps to a deployment that has buildpacks with a missing stack record:

- If you push an app and specify a stack with `cf push app-name -s stack`, PCF uses that stack. Otherwise, it uses the system default, `cflinuxfs3`.
- You may see additional logging in the buildpack detection output of the `cf push` command when PCF detects buildpacks without a stack record.

## Managing Stack Association with the cf CLI

The cf CLI commands for managing buildpacks include functionality to support association between buildpacks and stacks. The `update-buildpack`, `rename-buildpack`, and `delete-buildpack` commands all include a `-s` flag for specifying a stack.

When operating on buildpacks with the cf CLI, consider the following:

- You cannot upload a buildpack with `cf create-buildpack` if a buildpack of the same name already exists and it has a missing stack record.
- When using `cf create-buildpack`, you may inadvertently create a duplicate buildpack with a `nil` stack. `cf create-buildpack` does not disallow creation of buildpacks with no stack association.
- The `-s` flag is required when there are two buildpacks with the same name. If you are operating on a uniquely named buildpack, you do not need to specify its stack.

- If you have two buildpacks of the same name, one with a stack record and one without, running cf CLI commands without `-s` operates on the buildpack with the missing stack record.

## Example Scenarios

See the following examples of managing buildpacks with the cf CLI, which are applicable when running `cf update-buildpack`, `cf rename-buildpack`, or

```
cf delete-buildpack :
```

- Updating, renaming, or deleting a uniquely-named buildpack:**
  - You have a single buildpack named `my-buildpack`, and it is associated with `stack_a`. If you want to delete the buildpack, you can run`cf delete-buildpack my-buildpack`.  
You can also provide `-s stack_a`, but the option is not required if you have a uniquely-named buildpack.
- Updating, renaming, or deleting a uniquely-named buildpack that has a `nil` stack:**
  - You have a single buildpack named `my-buildpack`, and it is not associated with a stack. If you want to delete the buildpack, you can run`cf delete-buildpack my-buildpack`.
- Updating, renaming, or deleting a buildpack when another buildpack exists with the same name, and both buildpacks have stack associations:**
  - You have two buildpacks named `my-buildpack`, one that is associated with `stack_a` and the other associated with `stack_b`. If you want to delete the buildpack that uses `stack_a`, you can run `cf delete-buildpack my-buildpack -s stack_a`.
- Updating, renaming, or deleting a buildpack when another buildpack exists with the same name. One buildpack has a stack association, and the other buildpack has a `nil` stack:**
  - You have two buildpacks named `my-buildpack`, one associated with `stack_a` and the other associated with no (`nil`) stack association:
    - If you want to delete the buildpack that uses `stack_a`, you can run `cf delete-buildpack my-buildpack -s stack_a`.
    - If you want to delete the buildpack that is associated with the `nil` stack, run `cf delete-buildpack my-buildpack`.

## Pushing an Application with Multiple Buildpacks

This topic describes how developers can push an application with multiple buildpacks.

 **Note:** As an alternative to the cf CLI procedure below, you can specify multiple buildpacks in your [application manifest](#). This is not compatible with [Deprecated App Manifest Features](#).

For more information about pushing applications to Pivotal Application Service (PAS), see the [Pushing an App](#) topic.

### Specifying Buildpacks with the cf CLI

To push an application with multiple buildpacks using the Cloud Foundry Command Line Interface (cf CLI), perform the following procedure:

 **Note:** You must use cf CLI v6.38 or later.

1. Run the following command to ensure that you are using the cf CLI v6.38 or later:

```
$ cf version
```

For more information about upgrading the cf CLI, see [Installing the cf CLI](#).

2. To push your app with multiple buildpacks, specify each buildpack with a `-b` flag:

```
$ cf push YOUR-APP -b BUILDPACK-NAME-1 -b BUILDPACK-NAME-2 ... -b FINAL-BUILDPACK-NAME
```

The last buildpack you specify is the **final buildpack**, which can modify the launch environment and set the start command.

To see a list of available buildpacks, run `cf buildpacks`.

For more information on multi-buildpack order, see the [How Buildpacks Work in Cloud Foundry](#) topic.

For more information about using the cf CLI, see the [Cloud Foundry Command Line Interface](#) topic.

## Using a Proxy Server

This topic describes how developers can use a proxy server with the buildpacks for their application.

### Use a Proxy

You can assign proxy servers to environment variables. Proxy servers can be used to monitor your application's traffic or to fetch your application's dependencies.

A buildpack uses a proxy server if that buildpack contacts the internet during staging.

 **Note:** The binary buildpack does not use a proxy server because it does not access the internet during staging.

### Setting environment variables

 **Note:** If you are using the Java buildpack, the `http_proxy` and `https_proxy` environment variables are not supported at runtime. The Java buildpack does not add the functionality to make proxies work at runtime.

1. Add the following section to the `env` block of your application manifest YAML file:

```

env:
 http_proxy: http://YOUR-HTTP-PROXY:PORT
 https_proxy: https://YOUR-HTTPS-PROXY:PORT
```

Where:

- `YOUR-HTTP-PROXY` is the address of your proxy server for HTTP requests.
- `YOUR-HTTPS-PROXY` is the address of your proxy server for HTTPS requests.
- `PORT` is the port number you are using for your proxy server.

2. Set the environment variables with the Cloud Foundry Command Line Interface's (cf CLI) `cf set-env` command:

```
$ cf set-env YOUR-APP http_proxy "http://YOUR-HTTP-PROXY:PORT"
$ cf set-env YOUR-APP https_proxy "https://YOUR-HTTPS-PROXY:PORT"
```

Where:

- `YOUR-APP` is the name of your application.
- `YOUR-HTTP-PROXY` is the address of your proxy server for HTTP requests.
- `YOUR-HTTPS-PROXY` is the address of your proxy server for HTTPS requests.
- `PORT` is the port number you are using for your proxy server.

## Supported Binary Dependencies

Each buildpack only supports the stable patches for each dependency listed in the buildpack's `manifest.yml` and also in its GitHub releases page. For example, see the [php-buildpack releases page](#).

If you try to use an unsupported binary, staging your app fails with the following error message:

```
Could not get translated url, exited with: DEPENDENCY_MISSING_IN_MANIFEST:
...
!
! exit
!
Staging failed: Buildpack compilation step failed
```

## Production Server Configuration

This topic describes how to configure a production server for your apps.

When you deploy an app, PAS determines the command used to start the app through the following process:

1. If the developer uses the command `cf push -c COMMAND`, then PAS uses `COMMAND` to start the app.
2. If the developer creates a file called a Procfile, PAS uses the Procfile to configure the command that launches the app. See the [About Procfiles](#) section below for more information.
3. If the developer does not use `cf push -c COMMAND` and does not create a Procfile, then PAS does one of the following, depending on the buildpack:
  - Uses a default start command.
  - Fails to start the app and shows a warning that the app is missing a Procfile.

## About Procfiles

One reason to use a Procfile is to specify a start command for buildpacks where a default start command is not provided. Some buildpacks, such as Python, that work on a variety of frameworks, do not attempt to provide a default start command.

Another reason to use a Procfile is to configure a production server for web apps.

A Procfile enables you to declare required runtime processes, called process types, for your web app. Process managers in a server use the process types to run and manage the workload. In a Procfile, you declare one process type per line and use the following syntax:

`PROCESS_TYPE: COMMAND`

- `PROCESS_TYPE` is `web`. A `web` process handles HTTP traffic.
- `COMMAND` is the command line to launch the process.

For example, a Procfile with the following content starts the launch script created by the build process for a Java app:

`web: build/install/MY-PROJECT-NAME/bin/MY-PROJECT-NAME`

## Specify a Web Server

Follow these steps to specify a web server using a Procfile. For more information about configuring a web server for Rails apps, see the [Configure a Ruby Web Server](#) section of this topic.

1. Create a blank file with a command line for a `web` process type.
2. Save it as a file named `Procfile` with no extension in the root directory of your app.
3. Push your app.

## Configure a Ruby Web Server

PAS uses the default standard Ruby web server library WEBrick for Ruby and Ruby on Rails apps. However, PAS can support a more robust production web server, such as Phusion Passenger, Puma, Thin, or Unicorn.

To instruct PAS to use a web server other than WEBrick, perform the following steps:

1. Add the gem for the web server to your Gemfile.
2. In the `config` directory of your app, create a new configuration file or modify an existing file. Refer to your web server documentation for how to configure this file. The following example uses the Puma web server:

```
config/puma.rb
threads 8,32
```

```
workers 3

on_worker_boot do
 # things workers do
end
```

3. In the root directory of your app, create a Procfile and add a command line for a `web` process type that points to your web server. For information about configuring the specific command for a process type, see your web server documentation.

The following example shows a command that starts a Puma web server and specifies the app runtime environment, TCP port, and paths to the server state information and configuration files:

```
web: bundle exec puma -e $RAILS_ENV -p 1234 -S ~/puma -C config/puma.rb
```

## System Buildpacks

This topic describes the existing system buildpacks that Cloud Foundry supports.

Cloud Foundry includes a set of system buildpacks for common languages and frameworks. This table lists the system buildpacks.

Name	Supported Languages, Frameworks, and Technologies	GitHub Repository
Binary	n/a	<a href="#">Binary source</a>
Go	Go	<a href="#">Go source</a>
HWC	HWC	<a href="#">HWC source</a>
Java	Grails, Play, Spring, or any other JVM-based language or framework	<a href="#">Java source</a>
.NET Core	.NET Core	<a href="#">.NET Core source</a>
NGINX	NGINX	<a href="#">NGINX source</a>
Node.js	Node or JavaScript	<a href="#">Node.js source</a>
PHP	Cake, Symfony, Zend, NGINX, or HTTPD	<a href="#">PHP source</a>
Python	Django or Flask	<a href="#">Python source</a>
R	R	<a href="#">R source</a>
Ruby	Ruby, JRuby, Rack, Rails, or Sinatra	<a href="#">Ruby source</a>
Staticfile	HTML, CSS, JavaScript, or NGINX	<a href="#">Staticfile source</a>

## Binary Buildpack

Page last updated:

Use the binary buildpack for running arbitrary binary web servers.

### Push an App

Specify the binary buildpack to stage an app as a binary file. On a command line, use `cf push APP-NAME` with the `-b` option to specify the buildpack.

For example:

```
$ cf push my_app -b https://github.com/cloudfoundry/binary-buildpack.git
```

You can provide Cloud Foundry with the shell command to execute your binary in the following two ways:

- **Procfile:** In the root directory of your app, add a `Procfile` that specifies a `web` task:

```
web: ./app
```

- **Command line:** Use `cf push APP-NAME` with the `-c` option:

```
$ cf push my_app -c './app' -b binary_buildpack
```

### Compile your Binary

Cloud Foundry expects your binary to bind to the port specified by the `PORT` environment variable.

The following example in [Go](#) binds a binary to the `PORT` environment variable:

```
package main

import (
 "fmt"
 "net/http"
 "os"
)

func handler(w http.ResponseWriter, r *http.Request) {
 fmt.Fprintf(w, "Hello, %s", "world!")
}

func main() {
 http.HandleFunc("/", handler)
 http.ListenAndServe(": "+os.Getenv("PORT"), nil)
}
```

Your binary should run without any additional runtime dependencies on the cflinuxfs3 or lucid64 root filesystem (rootfs). Any such dependencies should be statically linked to the binary.

To boot a Docker container running the cflinuxfs3 filesystem, run the following command:

```
$ docker run -it cloudfoundry/cflinuxfs3 bash
```

To boot a Docker container running the lucid64 filesystem, run the following command:

```
$ docker run -it cloudfoundry/lucid64 bash
```

To compile the above Go application on the rootfs, golang must be installed. `apt-get install golang` and `go build app.go` will produce an `app` binary.

When deploying your binary to Cloud Foundry, use `cf push` with the `-s` option to specify the root filesystem it should run against.

```
$ cf push my_app -s (cflinuxfs3|lucid64)
```

## BOSH Configured Custom Trusted Certificate Support

Your platform operator can configure the platform to add the custom certificates into the application container. The custom trusted certificates are added to the `/etc/ssl/certs` directory and can be used by binary applications.

For more information, see [Configuring Trusted System Certificates for Applications](#).

## .NET Apps

You can use the Binary buildpack to push .NET Core apps, .NET Console apps, and self-hosted apps to the Windows stack. Self-hosted apps are apps that include their own web server.

To push a .NET Console app using the binary buildpack, see [Console Applications](#) in the *.NET Cookbook* for more information.

To deploy .NET Core apps to Windows using the binary buildpack, build it as a self-hosted app and push it with the binary buildpack. Note that you will need to specify the `windows` stack and use a custom start command in your manifest file. Here is an example manifest:

```

```

```
applications:
```

```
- name: mydemoapp
```

```
random-route: true
```

```
stack: windows
```

```
health-check-type: none
```

```
command: cmd /c ./CloudFoundry --server.urls http://*:PORT%
```

```
env:
```

```
ASPNETCORE_ENVIRONMENT: production
```

For information about deploying different types of .NET apps, follow the links in the table below.

Type of .NET App	Buildpack
ASP.NET MVC	
ASP.NET Web Forms	
ASP.NET WebAPI Apps	
Windows Communication Foundation (WCF)	<a href="#">HWC</a>
.NET Core pushed to Linux stack	<a href="#">.NET Core</a>

## Help and Support

Join the `#buildpacks` channel in our [Slack community](#) if you need any further assistance.

For more information about using and extending the binary buildpack in Cloud Foundry, see the [binary-buildpack GitHub repository](#).

You can find current information about this buildpack on the binary buildpack [release page](#) in GitHub.

## Go Buildpack

### Supported Versions

Supported Go versions can be found [in the release notes](#).

### Push an App

The Go buildpack will be automatically detected in the following circumstances:

- Your app has been packaged with [godep](#) using `godep save`.
- Your app has a `vendor/` directory and has any files ending with `.go`.
- Your app has a `GOPACKAGENAME` environment variable specified and has any files ending with `.go`.
- Your app has a `glide.yml` file and is using [glide](#), starting in buildpack version [1.7.9](#).
- Your app has a `Gopkg.toml` file and is using [dep](#), starting in buildpack version [1.8.9](#).

If your Cloud Foundry deployment does not have the Go Buildpack installed, or the installed version is out of date, you can use the latest version with the command:

```
$ cf push my_app -b https://github.com/cloudfoundry/go-buildpack.git
```

When specifying versions, specify only major/minor versions, such as Go 1.6, rather than Go 1.6.0. This ensures you receive the most recent patches.

### Start Command

When pushing Go apps, you can specify a start command for the app. You can place the start command in the `Procfile` file in root directory of your app. For example, if the binary generated by your Go project is `my-go-server`, your `Procfile` could contain the following:

```
web: my-go-server
```

For more information about Procfiles, see the [Configuring a Production Server](#) topic.

You can also specify the start command for your app in the `manifest.yml` file in the root directory. For example, your `manifest.yml` could contain the following:

```
--
applications:
- name: my-app-name
 command: my-go-server
```

If you do not specify a start command in a `Procfile`, in the manifest, or with the `-c` flag for `cf push`, the generated binary will be used as the start command. Example: `my-go-server`

### Push an App with godep

If you are using [godep](#) to package your dependencies, make sure that you have created a valid `Godeps/Godeps.json` file in the root directory of your app by running `godep save`.

When using godep, you can fix your Go version in `GoVersion` key of the `Godeps/Godeps.json` file.

### Go Sample App

- [Sample Go app](#)

An example `Godeps/Godeps.json`:

```
{
 "ImportPath": "go_app",
 "GoVersion": "go1.6",
 "Deps": []
}
```

## Push an App with Glide

If you use [glide](#) to specify or package your dependencies, make sure that you have created a valid `glide.yml` file in the root directory of your app by running `glide init`.

To vendor your dependencies before pushing, run `glide install`. This will generate a `vendor` directory and a `glide.lock` file specifying the latest compatible versions of your dependencies. You must have a `glide.lock` file when pushing a vendored app. You do not need a `glide.lock` file when deploying a non-vendored app.

### Glide

- [Sample Go app with Glide](#)

An example `glide.yml` file:

```
package: go_app_with_glide
import:
- package: github.com/ZiCog/shiny-thing
 subpackages:
 - foo
```

You can specify Go version in the `manifest.yml` file:

```
--
applications:
- name: my-app-name
env:
 GOVERSION: go1.8
```

## Push an App with dep

If you use [dep](#) to specify or package your dependencies, make sure that you have created a valid `Gopkg.toml` file in the root directory of your app by running `dep init`.

To vendor your dependencies before pushing, run `dep ensure`. This will generate a `vendor` directory and a `Gopkg.lock` file specifying the latest compatible versions of your dependencies. You must have a `Gopkg.lock` file when pushing a vendored app. You do not need a `Gopkg.lock` file when deploying a non-vendored app.

### dep

- [Sample Go app with dep](#)

An example `Gopkg.toml` file:

```
[[constraint]]
branch = "master"
name = "github.com/ZiCog/shiny-thing"
```

You can specify Go version in the `manifest.yml` file:

```
--
applications:
- name: my-app-name
env:
GOVERSION: go1.8
```

## Push an App with Native Go Vendoring

If you use the native Go vendoring system, which packages all local dependencies in the `vendor/` directory, you must specify your app's package name in the `GOPACKAGENAME` environment variable.

An example `manifest.yml`:

```
--
applications:
- name: my-app-name
command: go-online
env:
GOPACKAGENAME: example.com/user/app-package-name
```

### Go 1.6

- [Sample Go 1.6 app with native vendoring ↗](#).

Go 1.6 has vendoring enabled by default. Set the `GO15VENDOREXPERIMENT` environment variable to `0` to disable vendoring.

An example `manifest.yml` file:

```
--
applications:
- name: my-app-name
command: example-project
env:
GOVERSION: go1.6
GOPACKAGENAME: example.com/user/app-package-name
```

### Go 1.7 and Later

- [Sample Go 1.7 app with native vendoring ↗](#).

Go 1.7 and later always has vendoring enabled, and you cannot disable it with an environment variable.

An example `manifest.yml`:

```
--
applications:
- name: my-app-name
command: example-project
env:
GOVERSION: go1.8
GOPACKAGENAME: example.com/user/app-package-name
```

## Pass a Symbol and String to the Linker

The Go buildpack supports the Go [linker's ↗](#) ability, `-X symbol value`, to set the value of a string at link time. Set the `GO_LINKER_SYMBOL` and `GO_LINKER_VALUE` in the application's configuration before pushing code.

This can be used to embed the commit SHA or other build-specific data directly into the compiled executable.

For a sample Go app, see the [go-buildpack ↗](#) repository on GitHub.

## C Dependencies

The Go buildpack supports building with C dependencies using [cgo](#). You can set config vars to specify cgo flags to, for example, specify paths for vendored dependencies. As an example, to build [gopgsqldriver](#), add the config var `CGO_CFLAGS` with the value `-I/app/code/vendor/include/postgresql` and include the relevant Postgres header files in `vendor/include/postgresql/` in your app.

## Proxy Support

If you need to use a proxy to download dependencies during staging, you can set the `http_proxy` and/or `https_proxy` environment variables. For more information, see [Using a Proxy](#).

## BOSH Configured Custom Trusted Certificate Support

Go uses certificates stored in `/etc/ssl/certs`. Your platform operator can configure the platform to [add the custom certificates into the application container](#).

## Help and Support

Join the `#buildpacks` channel in our [Slack community](#) if you need any further assistance.

For more information about using and extending the Go buildpack in Cloud Foundry, see the [go-buildpack GitHub repository](#).

You can find current information about this buildpack on the Go buildpack [release page](#) in GitHub.

## HWC Buildpack

This topic describes how to configure your .NET Framework apps for use with the HWC buildpack and how to push your .NET Framework apps to Pivotal Application Service for Windows (PASW).

### Prerequisites

Using the HWC buildpack requires deploying Windows cells with PASW.

The HWC buildpack supports the following common app types by default:

- ASP.NET MVC
- ASP.NET Web Forms
- ASP.NET WebAPI Apps
- Windows Communication Foundation (WCF)

For information about deploying different types of .NET apps, follow the links in the table below.

Type of .NET App	Buildpack
.NET Console	<a href="#">Binary</a>
.NET Core pushed to Linux stack	<a href="#">.NET Core</a>
.NET Core pushed to Windows stack	<a href="#">Binary</a>

### Overview

The HWC buildpack provides a runtime server that uses the Hosted Web Core API for running .NET Framework applications in Windows Server containers. For more information, see [Hosted Web Core API Reference](#) in the Microsoft documentation.

The HWC buildpack provides access to .NET Framework 4.5.1 and later, made available by the Windows root file system (rootfs).

Before you push your first app using the HWC buildpack, see the [Getting Started](#) guide in the *.NET Cookbook*.

### Step 1. Configure HWC

HWC relies on `Web.config` and `applicationHost.config` configuration files for configuring the .NET applications.

Most `Web.config` files work out of the box with PASW, with the following constraints:

- Integrated Windows Authentication (IWA) is not yet supported on PASW.
- SQL server connection strings must use fully qualified domain names.
- Place connection string values in environment variables or [user-provided service instances](#).

Additionally, [Shadow Copy Setting](#), and [Dynamic and Static HTTP Compression](#) `Web.config` settings can be customized as needed.

The HWC buildpack includes a default configuration for the `applicationHost.config`, similar to IIS.

### Step 2. Add a Global Error Handler

Before you push your app for the first time, add a global error handler to receive log information from your app if it crashes on startup.

To configure a global error handler that logs to `stdout`, see [Application Error Handling](#) in the *.NET Cookbook*.

### Step 3. Push an App

Follow the steps below to push your application.

1. Build your HWC app in Visual Studio.
2. On the command line, navigate to the directory containing the app files.
3. To push your HWC app, run the following `cf push` command:

```
cf push APP-NAME -s windows -b hwc_buildpack
```

Where `APP-NAME` is the name you want to give your app.

For example:

```
$ cf push my-app -s windows -b hwc_buildpack
Creating app my-app in org sample-org / space sample-space as username@example.com...
OK
...
requested state: started
instances: 1/1
usage: 1 GB x 1 instances
urls: my-app.example.com
```

4. Confirm your application is running by finding your app's URL in the push command output and navigating to it. In the example above, `my-app.example.com` is the URL of your app.

## Features

The following features can be used with HWC buildpack:

- [Context Path Routing](#)
- [Shadow Copy Setting](#)
- [Dynamic and Static HTTP Compression](#)
- [URL Rewrite](#)
- [Profile Scripts](#)

## Context Path Routing

Context path routing enables multiple apps to share the same route hostname, such as `app1.example.com/app2`. The context path routing feature is analogous to IIS virtual directories.

Making an application accessible under another app's URL requires pushing both apps and applying a map-route between them. To define a context path route, such as `app1.example.com/app2`, run the following commands:

1. To push the primary app, run the following command:

```
cf push TOP-LEVEL-APP-NAME -s windows -b hwc_buildpack
```

Where:

- `TOP-LEVEL-APP-NAME` is your top-level app's name.

2. To push the secondary app and disable the app's starting and default routing, run the following command:

```
cf push LOWER-LEVEL-APP-NAME --no-start --no-route -s windows -b hwc_buildpack
```

Where:

- `LOWER-LEVEL-APP-NAME` is your lower-level app's name.

3. To map routes between the primary and secondary apps, run the following command:

```
cf map-route LOWER-LEVEL-APP-NAME APP-DOMAIN --hostname TOP-LEVEL-APP-NAME --path LOWER-LEVEL-APP-NAME
```

Where:

- `TOP-LEVEL-APP-NAME` is your top-level app's name.
- `LOWER-LEVEL-APP-NAME` is your lower-level app's name.
- `APP-DOMAIN` is your site's public domain name.

4. To start the secondary app, run the following command:

```
cf start LOWER-LEVEL-APP-NAME
```

Where: `LOWER-LEVEL-APP-NAME` is your lower-level app's name.

For example, the following commands define context path routing for two HWC apps, `app1` and `app2`, where `app2` is made accessible under `app1` as `app1.example.com/app2`:

```
$ cf push app1 -s windows -b hwc_buildpack
$ cf push app2 --no-start --no-route -s windows -b hwc_buildpack
$ cf map-route app2 example.com --hostname app1 --path app2
$ cf start app2
```

 **Note:** HWC-hosted apps use the `VCAP_APPLICATION` environment variable to read out the bound app URIs. Any context path that exists underneath the root in the app's bound route corresponds to the `applicationHost.config`.

## Shadow Copy Setting

[Shadow Copy](#) is a hosting option that copies assemblies for an app in the `bin` directory to the app's temporary files directory. This feature is turned off and unnecessary for apps running under Cloud Foundry. An app can override this setting in its `Web.config` file.

## Dynamic and Static HTTP Compression

The HWC buildpack enables dynamic and static HTTP compression by default. You can disable HTTP compression in your app's `Web.config` file.

Dynamic HTTP compression is hardcoded at level 4. Static HTTP compression is hardcoded at level 9.

## URL Rewrite

The HWC buildpack supports the URL Rewrite module. It is preinstalled in the Windows file system.

## Profile Scripts

The HWC buildpack allows developers to provide `.profile.bat` scripts with their applications. You can use a `.profile.bat` script to perform app-specific initialization tasks, such as setting custom environment variables.

For information about configuring `.profile.bat` scripts, see the [Configure Pre-Runtime Hooks](#) section of *Pushing an App*.

## Buildpack Support

A number of channels exist to assist you with using the HWC buildpack, or when developing your own HWC buildpack.

- **HWC Buildpack Repository in GitHub:** Find more information about using and extending the HWC buildpack in the [HWC buildpack repository](#) in GitHub.
- **Release Notes:** Find current information about this buildpack on the [HWC buildpack release page](#) in GitHub.
- **Slack:** Join the `#buildpacks` channel in the [Cloud Foundry Slack community](#).



## Creating an Extension Buildpack for .NET Apps

This topic explains how to write extension buildpacks for .NET apps.

### Prerequisites

- Linux or MacOS development machine or VM
- [golang](#) v1.10 or later programming language
- [direnv](#) environment variable manager for your shell

### Overview

Most .NET Framework apps are pushed with the [Hosted Web Core \(HWC\) buildpack](#).

The HWC buildpack contains the commonly required .NET dependencies used by .NET apps. This set of HWC buildpack dependencies is a subset entire gamut of possible .NET dependencies and does not contain all the DLLs that every app could need.

An extension buildpack supplies a custom set of .NET dependencies to an app's container.

You cannot add modules or extensions directly to the HWC buildpack. If the HWC buildpack lacks dependencies your app requires, we recommend that you create an extension buildpack containing the missing dependencies required by your app.

**Note:** The HWC buildpack contains a number of built-in HWC features, for example, the URL Rewrite and HTTP compression modules. Extensions should only provide additional functionality to your app. For more information about existing HWC features, see [Features](#) in *HWC Buildpack*.

For additional extension buildpack configuration and creation commands see: [Creating a Custom Buildpack](#) in the *Pivotal Engineering Journal*.

### Step 1. Initialize an Extension Buildpack

The [buildpack-packager](#) provides boilerplate code that you can use to start writing your buildpack.

1. To install the `buildpack-packager` CLI tool, run the commands below.

```
go get github.com/cloudfoundry/libbuildpack/
go install github.com/cloudfoundry/libbuildpack/tree/master/packager
```

2. To create your buildpack boilerplate, run the following command:

```
buildpack-packager init --name BUILDPACK-NAME --path=BUILDPACK-DIRECTORY
```

Where:

- `BUILDPACK-NAME` is the name of the buildpack you are creating.
- `BUILDPACK-DIRECTORY` is the directory where the boilerplate code will be written. If the directory does not exist, it is created by the `buildpack-packager init` process.

For example:

```
$ buildpack-packager init --name my-example --path=my-example-buildpack
```

3. To activate direnv in your buildpack directory, run the following commands:

```
cd BUILDPACK-DIRECTORY
direnv allow
```

Where: `BUILDPACK-DIRECTORY` is the directory created by the `buildpack-packager` command, containing the boilerplate code.

## Step 2. Create an Extension Buildpack to Supply Additional DLLs to a .NET Framework App

An extension buildpack provides the dependencies required by your app that are missing in HWC buildpack. The extension buildpack references and supplies needed dependencies to the app's container when you push the app.

To create an extension buildpack containing additional DLLs, follow the steps below:

1. Create a list of the dependencies that your published app requires.
2. Copy the DLLs for the dependencies into a ZIP file.
3. On the command line, browse to the directory containing the new ZIP file.
4. To generate a SHA for your ZIP file, run the following command:

```
shasum -a256 DLL-ZIP-FILE
```

Where `DLL-ZIP-FILE` is the name of your DLL ZIP file.

5. Upload the ZIP file containing your dependencies to a private web server that is accessible by Cloud Foundry. You can also use an S3 bucket or Azure Storage.
6. On the command line, browse to the directory created by the `buildpack-packager init` command above.
7. Edit the `manifest.yml` file by adding a `dependency` section that references the DLL ZIP file as follows:

```
dependencies:
- name: BINARY-NAME
 uri: http://s3.amazonaws.com/BUCKET-NAME/DLL-ZIP-FILE
 version: 0.0.1
 sha256: GENERATED-SHA-256
 cf_stacks:
 - windows
```

Where:

- o `BINARY-NAME` is the name of your binary application.
- o `BUCKET-NAME` is your S3 bucket name.
- o `DLL-ZIP-FILE` is the name of your DLL ZIP file.
- o `GENERATED-SHA-256` is the generated SHA.

8. Ensure the `include_files` section of the manifest contains `bin/supply.exe`.
9. To navigate to the `BUILDPACK-DIRECTORY/src/BUILDPACK-NAME/` directory, use either Windows Explorer, or your text editor's 'File - Open' option.
10. Edit the `supply.go` script.
11. Add dependencies to the container's `PATH` as follows:
  - a. Add a `stager.InstallDependency` for the `dependency` and `version` you have in your `manifest.yml` file.
  - b. Add a `stager.AddBinDependency` for every DLL file you want your app to be able to access.

For example:

```
dep := libbuildpack.Dependency{Name: "my-binary", Version: "0.0.1"}
if err := s.Installer.InstallDependency(dep, s.Stager.DepDir()); err != nil {
 return err
}
if err := s.Stager.AddBinDependencyLink(filepath.Join(s.Stager.DepDir(), "MyBinary.dll"), "MyBinary.dll"); err != nil {
 return err
}
```

12. To configure the `build.sh` script to cross-compile, edit `scripts/build.sh` to include the following:

```
GOOS=windows go build -ldflags="-s -w" -o bin/supply.exe mysql/supply/cli
```

## Step 3. Build the Extension Buildpack

The following steps build a cached extension buildpack artifact.

1. To build a cached buildpack artifact, run the following command:

```
buildpack-packager build -cached -stack STACK-NAME
```

Where:

- o `STACK-NAME` is the pre-built operating system that can run apps.

 **Note:** If the buildpack's `supply.go` script contains a mistake, `buildpack-packager` throws an error. You should not push your app with the new extension `buildpack` until after all local errors have been corrected. You must push your app to determine whether the new extension buildpack functions as intended.

## Step 4. Validate the Extension Buildpack

The following steps push an app with the cached extension buildpack artifact.

1. Upload the buildpack to a private web server accessible to Cloud Foundry. You can upload the finished buildpack to the same web server as your dependencies. You can also use an S3 bucket or Azure Storage.
2. To deploy your app with the extension buildpack and determine whether the new buildpack functions as intended, run the following command:

```
cf push APP-NAME -s windows -b EXT-BUILDPACK-URL -b hwc_buildpack
```

Where:

- o `APP-NAME` is your app's name.
- o `EXT-BUILDPACK-URL` is the URL of the uploaded extension buildpack zip file.

For example:

```
$ cf push my-app -s windows -b http://s3.amazonaws.com/my-bucket/my-example-buildpack.zip -b hwc_buildpack
```

If your extension buildpack does not include the correct dependencies, you will receive an error message.

## Step 5. Combine Extension Buildpack with Other Features

If your extension buildpack has executables or scripts that need to be run, you can reference them either in the start command or in profile scripts.

- For more information about the start command, see [command](#) in *Deploying with Application Manifests*.
- For more information about profile scripts, see [Configure Pre-Runtime Hooks](#) in *Deploying an Application*.

## Tips for .NET Framework Developers

This topic describes how to push .NET Framework apps to Pivotal Application Service (PAS) for Windows cells.

For how to develop microservices for .NET Framework and .NET Core using Steeltoe, see the [Steeltoe documentation](#).

## Prerequisites

The PAS for Windows tile must be installed and configured. For documentation about installing the PAS for Windows tile, see the [Installing and Configuring PAS for Windows](#) topic.

Operators must also install the Cloud Foundry Command Line Interface (cf CLI) to run the commands on this topic. For information about installing the cf CLI, see the [Installing the cf CLI](#) topic.

## Overview

After operators install and configure the PAS for Windows tile, developers can push .NET Framework apps to the Windows cell. Developers can push both [OWIN](#) and non-OWIN apps, and can push apps that are served by [Hostable Web Core](#) or [self-hosted](#).

If you have upgraded to PAS for Windows and have apps that you want to migrate, see the [Upgrading Cells](#) topic.

## Develop .NET Framework Apps

### .NET on PCF Cookbook

The [.NET Cookbook](#) has useful tips and recipes for migrating and developing .NET Framework apps to run on PAS for Windows.

## Push a .NET Framework App

By default, PCF serves .NET Framework apps with Hostable Web Core (HWC). HWC is a lighter version of the Internet Information Services (IIS) server that contains the core IIS functionality.

Perform the following steps to push a .NET Framework app to a Windows cell:

1. To target the Cloud Controller of your PCF deployment, run the following command:

```
cf api APP-DOMAIN
```

Where `APP-DOMAIN` is your application's public domain name, for example: `example.com`.

2. Run one of the following commands to deploy your .NET app:

- To deploy an app with `.bat` or `.exe` files, run the following command:

```
cf push -s windows -b binary_buildpack
```

- To deploy a .NET Framework app, run the following command:

```
cf push APP-NAME -s windows -b hwc_buildpack
```

Where `APP-NAME` is the name of your application.

 **Note:** The `-s windows` option instructs PCF to run the app in the Windows cell.

 **Note:** If you are not pushing your app from its directory, use the `-P` option and specify the path to the directory that contains the app.

3. Wait for your app to stage and start. If you see an error message, see the [Troubleshoot App Errors](#) section of this topic.

## Context Path Routing Support for ASP.NET Apps

Context path routing enables multiple apps to share the same route hostname, such as `app1.example.com/app2`. ASP.NET developers can host apps under a route path. Within Windows cells, you can have multiple routes to an app, but those routes cannot have different context paths.

Making an app accessible under another app's URL requires a pair of commands. To define a context path route, such as `app1.example.com/app2`, run the following commands:

1. To push the top-level app, run the following command:

```
cf push TOP-LEVEL-APP-NAME
```

Where `TOP-LEVEL-APP-NAME` is the name of your top-level app.

2. To push the lower-level app, run the following command:

```
cf push LOWER-LEVEL-APP-NAME -d APP-DOMAIN --hostname TOP-LEVEL-APP-NAME --route-path LOWER-LEVEL-APP-NAME
```

Where:

- `TOP-LEVEL-APP-NAME` is the name of your top-level app.
- `LOWER-LEVEL-APP-NAME` is the name of your lower-level app.
- `APP-DOMAIN` is your application's public domain name, for example: `example.com`.

 **Note:** The `-d` parameter is only needed when pushing an app to a non-default domain.

## Push a Self-Hosted App

Developers can choose to push a self-hosted app instead of using Hostable Web Core. Self-hosted apps combine server code with the app code.

Perform the following steps to push a self-hosted app:

1. To target the Cloud Controller of your PCF deployment, run the following command:

```
cf api api.APP-DOMAIN
```

Where `APP-DOMAIN` is your application's public domain name, for example: `example.com`.

2. To push your .NET app from the app root, run the following command:

```
cf push APP-NAME -s windows -b binary_buildpack -c PATH-TO-BINARY
```

Where:

- `APP-NAME` is the name of your app.
- `PATH-TO-BINARY` is the path to your executable.

3. Wait for your app to stage and start. If you see an error message, see the [Troubleshoot App Errors](#) section of this topic.

## Push a SOAP Service

Developers can push Simple Object Access Protocol (SOAP) web services to their PCF deployment by following the procedures in the sections below.

## Step 1: Deploy Your Web Service

Perform the following steps to deploy a SOAP web service:

1. Develop the service as an ASMX web service in Microsoft Visual Studio.
2. Publish the service to your local file system.
3. Open a command line to the directory containing the published web service.
4. To push your service, run the following command:

```
cf push SOAP-SERVICE-NAME -s windows -b hwc_buildpack -p WEB-SERVICE-DIRECTORY -u none
```

Where

- `SOAP-SERVICE-NAME` is the name of your service.
- `WEB-SERVICE-DIRECTORY` is the path to the directory containing the published web service.

For example:

```
$ cf push webservice -s windows -b hwc_buildpack -u none

requested state: started
instances: 1/1
usage: 1G x 1 instances
urls: webservice.example.com
```

 **Note:** The push command must include the `-s` flag to specify the stack, which instructs PCF to run the app in the Windows cell.

 **Note:** The `-p` and `-u` parameters are optional parameters. The `-p` parameter is needed only when pushing your service from a directory that does not contain the published web service. The `-u` parameter is needed only when disabling the health check when you do not have a route serving `/`.

5. Confirm your service is running by finding your service's URL in the push command's output and browsing to it. In the example above, the URL for the service would be: <http://webservice.example.com>.

## Step 2: Modify the WSDL File

Your SOAP web service is now deployed on PCF, but the service's WSDL file contains incorrect port information. The WSDL file must be modified to enable an app to consume your web service. Either you or the service developer can perform the needed modification.

Examine the following portion of an example WSDL file:

```
- <wsdl:service name="WebService1">
- <wsdl:port name="WebService1Soap" binding="tns:WebService1Soap">
 <soap:address location="http://webservice.example.com:62492/WebService1.asmx"/>
</wsdl:port>
- <wsdl:port name="WebService1Soap12" binding="tns:WebService1Soap12">
 <soap12:address location="http://webservice.example.com:62492/WebService1.asmx"/>
</wsdl:port>
- </wsdl:service>
```

The WSDL file provides the port number for the SOAP web service as `62492`. This is the port that the web service listens on in the [Garden container](#), but external applications cannot access the service on this port. Instead, external applications must use port `80`, and the [Gorouter](#) routes requests to the web service in the container.

The URL of the web service in the WSDL file must be modified to remove `62492`. With no port number, the URL defaults to port `80`. In the example above, the modified URL would be:

```
http://webservice.example.com/WebService1.asmx
```

SOAP web service developers can resolve this problem in one of two ways:

- Modify the WSDL file by following the instructions in [Modify a Web Service's WSDL Using a SoapExtensionReflector](#) from the Microsoft Developers

Network.

- Instruct the developers of external applications that consume the web service to perform the steps in the [Consume the SOAP Web Service](#) section of this topic.

## Consume the SOAP Web Service

Developers of external applications that consume the SOAP web service can perform the following steps to use a modified version of the WSDL file:

- In a browser, navigate to the WSDL file of the web service, using the following URL:

```
https://SOAP-SERVICE-NAME.APP-DOMAIN/ASMX-FILE?wsdl
```

Where:

- `SOAP-SERVICE-NAME` is the name of your service.
- `APP-DOMAIN` is your site's public domain name.
- `ASMX-FILE` is the filename of your asmx file.

For example:

```
https://webservice.example.com/WebService1.asmx?wsdl
```

- Download the WSDL file to your local machine.
- Edit the WSDL file to eliminate the container port, as described in the [Modify the WSDL File](#) section of this topic.
- In Microsoft Visual Studio, right-click on your application in the **Solution Explorer** and select **Add > Service Reference**.
- Under **Address**, enter the local path to the modified WSDL file.

For example:

```
C:\Users\example\wsdl.xml
```

- Click **OK**. Microsoft Visual Studio generates a client from the WSDL file that you can use in your codebase.

## Context Path Routing Support for SOAP Web Services

Developers can push SOAP web services to their PCF deployment with context path routing. For more information, see the [Context Path Routing Support for ASP.NET Apps](#) section.

## Troubleshoot App Errors

If a .NET app fails to start, consult the following list of errors and their possible solutions:

- `NoCompatibleCell`: Your PCF deployment cannot connect to your Windows cell. For information about troubleshooting your Windows cell configuration, see the [Troubleshooting Windows Cells](#) topic.
- `Start unsuccessful`: Your app may be misconfigured or lacks the required DLL files and dependencies.
  - Ensure that your app directory contains either a valid `.exe` binary or a valid `Web.config` file.
  - Ensure that you are pushing from a directory containing your app dependencies. If it does not, specify the app dependency directory with the `-p` flag.

## Java Buildpack

You can use the Java buildpack with apps written in Grails, Play, Spring, or any other JVM-based language or framework.

See the following topics for more information:

- [Tips for Java Developers](#)
- [Getting Started Deploying Apps](#)
- [Configuring Service Connections](#)
- [Cloud Foundry Java Client Library](#)

See the [Java Buildpack Release Notes](#) for information about specific versions. You can find the source for the Java buildpack in the [Java buildpack repository](#) on GitHub:

## Buildpack and Application Logging

The Java buildpack only runs during the staging process, and therefore only logs staging information such as the downloaded components, configuration data, and work performed on your application by the buildpack.

The Java buildpack source documentation states the following:

- The Java buildpack logs all messages, regardless of severity, to `APP-DIRECTORY/.java-buildpack.log`. The buildpack also logs messages to `$stderr`, filtered by a configured severity level.
- If the buildpack fails with an exception, the exception message is logged with a log level of `ERROR`. The exception stack trace is logged with a log level of `DEBUG`. This prevents users from seeing stack traces by default.

Once staging completes, the buildpack stops logging. The Loggregator handles application logging.

Your application must write to STDOUT or STDERR for its logs to be included in the Loggregator stream. For more information, see the [Application Logging in Cloud Foundry](#) topic.

## BOSH Custom Trusted Certificate Support

Versions 3.7 and later of the Java buildpack support BOSH-configured custom trusted certificates. For more information, see [Configuring Trusted Certificates](#) in the BOSH documentation.

The Java buildpack pulls the contents of `/etc/ssl/certs/ca-certificates.crt` and `SCF_INSTANCE_CERT/$CF_INSTANCE_KEY` by default.

The log output for Diego Instance Identity-based `KeyStore` appears as follows:

```
Adding System Key Manager
Adding Key Manager for /etc/cf-instance-credentials/instance.key and /etc/cf-instance-credentials/instance.crt
Start watching /etc/cf-instance-credentials/instance.crt
Start watching /etc/cf-instance-credentials/instance.key
Initialized KeyManager for /etc/cf-instance-credentials/instance.key and /etc/cf-instance-credentials/instance.crt
```

The log output for Diego Trusted Certificate-based `TrustStore` appears as follows:

```
Adding System Trust Manager
Adding TrustManager for /etc/ssl/certs/ca-certificates.crt
Start watching /etc/ssl/certs/ca-certificates.crt
Initialized TrustManager for /etc/ssl/certs/ca-certificates.crt
```

## Memory Constraints in Java Buildpack 4.0

The memory calculator in Java buildpack 4.0 accounts for the following memory regions:

- `-Xmx` : Heap
- `-XX:MaxMetaspaceSize` : Metaspace

- `-Xss`: Thread Stacks
- `-XX:MaxDirectMemorySize`: Direct Memory
- `-XX:ReservedCodeCacheSize`: Code Cache
- `-XX:CompressedClassSpaceSize`: Compressed Class Space

Applications which previously ran in 512 MB or smaller containers may no longer be able to. Most applications will run if they use the Cloud Foundry default container size of 1 G without any modifications. However, you can configure those memory regions directly as needed.

The Java buildpack optimizes for all non-heap memory regions first and leaves the remainder for the heap.

The Java buildpack prints a histogram of the heap to the logs when the JVM encounters a terminal failure.

The Cloud Foundry default Java buildpack is currently 3.x to allow time for apps to be upgraded to 4.x.

For more information, see [Java buildpack 4.0](#).

## Tips for Java Developers

Cloud Foundry can deploy a number of different JVM-based artifact types. For a more detailed explanation of what it supports, see the [Java Buildpack documentation](#).

## Java Buildpack

For detailed information about using, configuring, and extending the Cloud Foundry Java buildpack, see [Java Buildpack documentation](#).

## Design

The Java Buildpack is designed to convert artifacts that run on the JVM into executable applications. It does this by identifying one of the supported artifact types (Grails, Groovy, Java, Play Framework, Spring Boot, and Servlet) and downloading all additional dependencies needed to run. The collection of services bound to the application is also analyzed and any dependencies related to those services are also downloaded.

As an example, pushing a WAR file that is bound to a PostgreSQL database and New Relic for performance monitoring would result in the following:

```
Initialized empty Git repository in /tmp/buildpacks/java-buildpack/.git/
--> Java Buildpack source: https://github.com/cloudfoundry/java-buildpack#0928916a2dd78e9faf9469c558046cef09f60e5d
--> Downloading Open Jdk JRE 1.7.0_51 from
 http://.../openjdk/lucid/x86_64/openjdk-1.7.0_51.tar.gz (0.0s)
 Expanding Open Jdk JRE to java-buildpack/open_jdk_jre (1.9s)
--> Downloading New Relic Agent 3.4.1 from
 http://.../new-relic/new-relic-3.4.1.jar (0.4s)
--> Downloading Postgresql JDBC 9.3.1100 from
 http://.../postgresql-jdbc/postgresql-jdbc-9.3.1100.jar (0.0s)
--> Downloading Spring Auto Reconfiguration 0.8.7 from
 http://.../auto-reconfiguration/auto-reconfiguration-0.8.7.jar (0.0s)
 Modifying /WEB-INF/web.xml for Auto Reconfiguration
--> Downloading Tomcat 7.0.50 from
 http://.../tomcat/tomcat-7.0.50.tar.gz (0.0s)
 Expanding Tomcat to java-buildpack/tomcat (0.1s)
--> Downloading Buildpack Tomcat Support 1.1.1 from
 http://.../tomcat-buildpack-support/tomcat-buildpack-support-1.1.1.jar (0.1s)
--> Uploading droplet (57M)
```

## Configuration

In most cases, the buildpack should work without any configuration. If you are new to Cloud Foundry, we recommend that you make your first attempts without modifying the buildpack configuration. If the buildpack requires some configuration, use [a fork of the buildpack](#).

## Java Client Library

The Cloud Foundry Client Library provides a Java API for interacting with a Cloud Foundry instance. This library, `cloudfoundry-client-lib`, is used by the Cloud Foundry Maven plugin, the Cloud Foundry Gradle plugin, and other Java-based tools.

For information about using this library, see the [Java Cloud Foundry Library](#) page.

## Grails

Grails packages applications into WAR files for deployment into a Servlet container. To build the WAR file and deploy it, run the following:

```
$ grails prod war
$ cf push my-application -p target/my-application-version.war
```

## Groovy

Groovy applications based on both [Ratpack](#) and a simple collection of files are supported.

## Ratpack

Ratpack packages applications into two different styles; Cloud Foundry supports the `distZip` style. To build the ZIP and deploy it, run the following:

```
$ gradle distZip
$ cf push my-application -p build/distributions/my-application.zip
```

## Raw Groovy

Groovy applications that are made up of a [single entry point](#) plus any supporting files can be run without any other work. To deploy them, run the following:

```
$ cf push my-application
```

## Java Main

Java applications with a `main()` method can be run provided that they are packaged as [self-executable JARs](#).

 **Note:** If your application is not web-enabled, you must suppress route creation to avoid a “failed to start accepting connections” error. To suppress route creation, add `no-route: true` to the application manifest or use the `--no-route` flag with the `cf push` command.

For more information about the `no-route` attribute, see the [Deploying with Application Manifests](#) topic.

## Maven

A Maven build can create a self-executable JAR. To build and deploy the JAR, run the following:

```
$ mvn package
$ cf push my-application -p target/my-application-version.jar
```

## Gradle

A Gradle build can create a self-executable JAR. To build and deploy the JAR, run the following:

```
$ gradle build
$ cf push my-application -p build/libs/my-application-version.jar
```

## Play Framework

The [Play Framework](#) packages applications into two different styles. Cloud Foundry supports both the `staged` and `dist` styles. To build the `dist` style and deploy it, run the following:

```
$ play dist
$ cf push my-application -p target/universal/my-application-version.zip
```

## Spring Boot CLI

Spring Boot [can run applications comprised entirely of POGOs](#). To deploy then, run the following:

```
$ spring grab *.groovy
$ cf push my-application
```

## Servlet

Java applications can be packaged as Servlet applications.

### Maven

A Maven build can create a Servlet WAR. To build and deploy the WAR, run the following:

```
$ mvn package
$ cf push my-application -p target/my-application-version.war
```

### Gradle

A Gradle build can create a Servlet WAR. To build and deploy the JAR, run the following:

```
$ gradle build
$ cf push my-application -p build/libs/my-application-version.war
```

## Binding to Services

Information about binding apps to services can be found on the following pages:

- [Service Bindings for Grails Applications](#)
- [Service Bindings for Play Framework Applications](#)
- [Service Bindings for Spring Applications](#)

## Java and Grails Best Practices

### Provide JDBC driver

The Java buildpack does not bundle a JDBC driver with your application. If your application will access a SQL RDBMS, include the appropriate driver in your application.

### Allocate Sufficient Memory

If you do not allocate sufficient memory to a Java application when you deploy it, it may fail to start, or PAS may terminate it. You must allocate enough memory to allow for the following:

- Java heap
- Metaspace, if using Java 8
- PermGen, if using Java 7 or earlier
- Stack size per Thread

- JVM overhead

The `config/open_jdk_jre.yml` file of the [Cloud Foundry Java buildpack](#) contains default memory size and weighting settings for the JRE. See the [Open JDK JRE README](#) on GitHub for an explanation of JRE memory sizes and weightings and how the Java buildpack calculates and allocates memory to the JRE for your app.

To configure memory-related JRE options for your app, you either create a custom buildpack and specify this buildpack in your deployment manifest or you override the default memory settings of your buildpack as described in the <https://github.com/cloudfoundry/java-buildpack#configuration-and-extension> with the properties listed in the [Open JDK JRE README](#). For more information about configuring custom buildpacks and manifests, refer to the [Custom Buildpacks](#) and [Deploying with Application Manifests](#) topics.

When your app is running, you can use the `cf app APP-NAME` command to see memory utilization.

## Troubleshoot Out Of Memory

A Java app may crash because of insufficient memory on the Garden container or the JVM on which it runs. See the following sections for help diagnosing and resolving such issues.

### JVM

- **Error:** `java.lang.OutOfMemoryError`. See the following example:

```
$ cf logs APP-NAME --recent
2016-06-20T09:18:51.00+0100 [APP/0] OUT java.lang.OutOfMemoryError: Metaspace
```

- **Cause:** If the JVM cannot garbage-collect enough space to ensure the allocation of a data-structure, it fails with [java.lang.OutOfMemoryError](#). In the example above, JVM has an under-sized `metaspace`. You may see failures in other memory pools, such as `heap`.
- **Solution:** Configure the JVM correctly for your app. See [Allocate Sufficient Memory](#).

### Garden Container

**Note:** The solutions in this section require configuring the memory calculator, which is a sub-project of the CF Java buildpack that calculates suitable memory settings for Java apps when you push them. See the [java-buildpack-memory-calculator](#) repository for more information. If you have questions about the memory calculator, you can ask them in the `#java-buildpack` channel of the [Cloud Foundry Slack organization](#).

- **Error:** The Garden container terminates the Java process with the `out of memory` event. See the following example:

```
$ cf events APP-NAME
time event actor description
2016-06-20T09:18:51.00+0100 app.crash app-name index: 0, reason: CRASHED, exit_description: out of memory, exit_status: 255
```

This error appears when the JVM allocates more OS-level memory than the quota requested by the app, such as [through the manifest](#).

- **Cause 1 - Insufficient native memory:** This error commonly means that the JVM requires more `native` memory. In the scope of the Java buildpack and the memory calculator, the term `native` means the memory required for the JVM to work, along with forms of memory not covered in the other classifications of the memory calculator. This includes the memory footprint of OS-level threads, [direct NIO buffers](#), code cache, program counters, and others.
- **Solution 1:** Determine how much `native` memory a Java app needs by [measuring it](#) with realistic workloads and fine-tuning it accordingly. You can then configure the Java buildpack using the `native` [setting](#) of the memory calculator, as in the example below:

```

applications:
- name: app-name
 memory: 1G
 env:
 JBP_CONFIG_OPEN_JDK_JRE: '[memory_calculator: {memory_sizes: {native: 150m}}]'
```

This example shows that `150m` of the overall available `1G` is reserved for anything that is not `heap`, `metaspace`, or `permgen`. In less common cases, this may come from companion processes started by the JVM, such as the [Process API](#).

- **Cause 2 - High thread count:** Java threads in the JVM can cause memory errors at the Garden level. When an app is under heavy load, it uses a high number of threads and consumes a significant amount of memory.

- **Solution 2:** Set the reserved memory for stack traces to the correct value for your app.

You can use the `stack` [setting](#) of the memory calculator to configure the amount of space the JVM reserves for each Java thread. You must multiply this value by the number of threads your app requires. Specify the number of threads in the `stack_threads` [setting](#) of the memory calculator. For example, if you estimate the max thread count for an app at `800` and the amount of memory needed to represent the deepest stacktrace of a Java thread is `512KB`, configure the memory calculator as follows:

```
--
applications:
- name: app-name
 memory: 1G
 env:
 JBP_CONFIG_OPEN_JDK_JRE: '[memory_calculator: {stack_threads: 800, memory_sizes: {stack: 512k}}]'
```

In this example, the overall memory amount reserved by the JVM for representing the stacks of Java threads is `800 * 512k = 400m`.

The correct settings for `stack` and `stack_threads` depend on your app code, including the libraries it uses. Your app may technically have no upper limit, such as in the case of [cavalier usage](#) of [CachedThreadPool](#) [executors](#). However, you still must calculate the depth of the thread stacks and the amount of space the JVM should reserve for each of them.

## Troubleshoot Failed Upload

If your application fails to upload when you push it to Cloud Foundry, it may be for one of the following reasons:

- **WAR is too large:** An upload may fail due to the size of the WAR file. Cloud Foundry testing indicates WAR files as large as 250 MB upload successfully. If a WAR file larger than that fails to upload, it may be a result of the file size.
- **Connection issues:** Application uploads can fail if you have a slow Internet connection, or if you upload from a location that is very remote from the target Cloud Foundry instance. If an application upload takes a long time, your authorization token can expire before the upload completes. A workaround is to copy the WAR to a server that is closer to the Cloud Foundry instance, and push it from there.
- **Out-of-date cf CLI client:** Upload of a large WAR is faster and hence less likely to fail if you are using a recent version of the cf CLI. If you are using an older version of the cf CLI client to upload a large WAR, and having problems, try updating to the latest version of the cf CLI.
- **Incorrect WAR targeting:** By default, `cf push` uploads everything in the current directory. For a Java application, `cf push` with no option flags uploads source code and other unnecessary files, in addition to the WAR. When you push a Java application, specify the path to the WAR:

```
$ cf push MY-APP -p PATH/TO/WAR-FILE
```

You can determine whether or not the path was specified for a previously pushed application by examining the application deployment manifest, `manifest.yml`. If the `path` attribute specifies the current directory, the manifest includes a line like the following:

```
path: .
```

To re-push just the WAR, do one of the following:

- Delete `manifest.yml` and run `cf push` again, specifying the location of the WAR using the `-p` flag.
- Edit the `path` argument in `manifest.yml` to point to the WAR, and re-push the application.

## Debug Java Apps on Cloud Foundry

Because of the way that Cloud Foundry deploys your applications and isolates them, it is not possible to connect to your application with the remote Java debugger. Instead, instruct the application to connect to the Java debugger on your local machine.

Here are the instructions for setting up remote debugging when using BOSH Lite or a Cloud Foundry installation.

1. Open your project in Eclipse.
2. Right-click on your project, go to **Debug as** and pick **Debug Configurations**.
3. Create a new **Remote Java Application**.
4. Make sure your project is selected, pick **Standard (Socket Listen)** from the **Connection Type** drop down and set a port. Make sure this port is open if you are running a firewall.
5. Click **Debug**.

The debugger should now be running. If you switch to the Debug perspective, you should see your application listed in the Debug panel and it should say

Waiting for vm to connect at .  
port

Next, push your application to Cloud Foundry and instruct Cloud Foundry to connect to the debugger running on your local machine using the following instructions:

1. Edit your `manifest.yml` file. Set the instances count to 1. If you set this greater than one, multiple applications try to connect to your debugger.
2. Also in `manifest.yml`, add an `env` block and create a variable named `JAVA_OPTS`. For more information about the `env` block, see [Environment Variables](#) in *Deploying with Application Manifests*.
3. Add the remote debugger configuration to the `JAVA_OPTS` variable:  
`-agentlib:jdwp=transport=dt_socket,address=YOUR-IP-ADDRESS:YOUR-PORT`.
4. Save the `manifest.yml` file.
5. Run `cf push`.

Upon completion, you should see that your application has started and is now connected to the debugger running in your IDE. You can now add breakpoints and interrogate the application just as you would if it were running locally.

## Slow Starting Java or Grails Apps

Some Java and Grails applications do not start quickly, and the health check for an application can fail if an application starts too slowly.

The current Java buildpack implementation sets the Tomcat `bindOnInit` property to `false`. This prevents Tomcat from listening for HTTP requests until an application has fully deployed.

If your application does not start quickly, the health check may fail because it checks the health of the application before the application can accept requests. By default, the health check fails after a timeout threshold of 60 seconds.

To resolve this issue, use `cf push APP-NAME` with the `-t TIMEOUT-THRESHOLD` option to increase the timeout threshold. Specify `TIMEOUT-THRESHOLD` in seconds.

```
$ cf push my-app -t 180
```

 **Note:** The timeout threshold cannot exceed 180 seconds. Specifying a timeout threshold greater than 180 seconds results in the following error:

```
Server error, status code: 400, error code: 100001, message: The app is invalid: health_check_timeout
maximum_exceeded
```

## Extension

The Java Buildpack is also designed to be easily extended. It creates abstractions for [three types of components](#) (containers, frameworks, and JREs) in order to allow users to easily add functionality. In addition to these abstractions, there are a number of [utility classes](#) for simplifying typical buildpack behaviors.

As an example, the New Relic framework looks like the following:

```

class NewRelicAgent < JavaBuildpack::Component::VersionedDependencyComponent

@macro base_component_compile
def compile
 FileUtils.mkdir_p logs_dir

 download_jar
 @droplet.copy_resources
end

@macro base_component_release
def release
 @droplet.java_opts
 add_javaagent(@droplet.sandbox + jar_name)
 add_system_property('newrelic.home', @droplet.sandbox)
 add_system_property('newrelic.config.license_key', license_key)
 add_system_property('newrelic.config.app_name', "#{application_name}")
 add_system_property('newrelic.config.log_file_path', logs_dir)
end

protected

@macro versioned_dependency_component_supports
def supports?
 @application.services.one_service? FILTER, 'licenseKey'
end

private

FILTER = /newrelic/.freeze

def application_name
 @application.details['application_name']
end

def license_key
 @application.services.find_service(FILTER)['credentials']['licenseKey']
end

def logs_dir
 @droplet.sandbox + 'logs'
end

end

```

## Environment Variables

You can access environments variable programmatically.

For example, you can obtain `VCAP_SERVICES` as follows:

```
System.getenv("VCAP_SERVICES");
```

See the [Cloud Foundry Environment Variables](#) topic for more information.

## Getting Started Deploying Java Apps

This topic provides links to additional information about getting started deploying apps using the Java buildpack. See the following topics for deployment guides specific to your app framework:

- [Grails](#)
- [Ratpack](#)
- [Spring](#)

## Getting Started Deploying Grails Apps

This guide is intended to walk you through deploying a Grails app to Pivotal Application Service (PAS). If you experience a problem following the steps below, refer to the [Known Issues](#) or [Troubleshooting Application Deployment and Health](#) topics for more information.

### Sample App Step

If you want to go through this tutorial using the sample app, run `git clone https://github.com/cloudfoundry-samples/pong_matcher_grails.git` to clone the `pong_matcher_grails` app from GitHub, and follow the instructions in the Sample App Step sections.

 **Note:** Ensure that your Grails app runs locally before continuing with this procedure.

## Deploy a Grails Application

This section describes how to deploy a Grails application to PAS.

### Prerequisites

- A Grails app that runs locally on your workstation
- Intermediate to advanced Grails knowledge
- The [Cloud Foundry Command Line Interface \(cf CLI\)](#)
- JDK 1.7 or 1.8 for Java 7 or 8 configured on your workstation

 **Note:** You can develop Grails applications in Groovy, Java 7 or 8, or any JVM language. The Cloud Foundry Java buildpack uses JDK 1.8, but you can modify the buildpack and the manifest for your app to compile to JDK 1.7 as described in *Step 8: Configure the Deployment Manifest* of this topic.

## Step 1: Declare App Dependencies

Declare all the dependency tasks for your app in the build script of your chosen build tool. The table lists build script information for Gradle, Grails, and Maven, and provides documentation links for each build tool.

Build Tool	Build Script	Documentation
Gradle	<code>build.gradle</code>	<a href="#">Gradle User Guide</a>
Grails	<code>BuildConfig.groovy</code>	<a href="#">Grails: Configuration - Reference Documentation</a>
Maven	<code>pom.xml</code>	<a href="#">Apache Maven Project Documentation</a>

### Sample App Step

You can skip this step. The `pong_matcher_grails/app/grails-app/conf/BuildConfig.groovy` file contains the dependencies for the `pong_matcher_grails` sample app, as the example below shows.

```
dependencies {
 // specify dependencies here under either 'build', 'compile', 'runtime', 'test' or 'provided' scopes e.g.
 // runtime 'mysql:mysql-connector-java:5.1.29'
 // runtime 'org.postgresql:postgresql:9.3-1101-jdbc41'
 test "org.grails:grails-datastore-test-support:1.0-grails-2.4"
 runtime 'mysql:mysql-connector-java:5.1.33'
}
```

## Step 2: Allocate Sufficient Memory

Run the Cloud Foundry Command Line Interface (cf CLI) `cf push -m` command to specify the amount of memory that should be allocated to the application. Memory allocated this way is done in preset amounts of 64M, 128M, 256M, 512M, 1G, or 2G. For example:

```
$ cf push -m 128M
```

When your app is running, you can use the `cf app APP-NAME` command to see memory utilization.

#### Sample App Step

You can skip this step. In the `manifest.yml` of the `pong_matcher_grails` sample app, the `memory` sub-block of the `applications` block allocates 1 GB to the app.

## Step 3: Provide a JDBC Driver

The Java buildpack does not bundle a JDBC driver with your application. If your application accesses a SQL RDBMS, you must do the following:

- Include the appropriate driver in your application.
- Create a dependency task for the driver in the build script for your build tool or IDE.

#### Sample App Step

You can skip this step. The `pong_matcher_grails` sample app declares a MySQL JDBC driver in the `pong_matcher_grails/app/grails-app/conf/DataSource.groovy` file because the app uses ClearDB, which is a database-as-service for MySQL-powered apps. The example below shows this declaration.

```
dataSource {
 pooled = true
 jmxExport = true
 driverClassName = "com.mysql.jdbc.Driver"
 dialect = org.hibernate.dialect.MySQL5InnoDBDialect
 uri = new URI(System.getenv("DATABASE_URL") ?: "mysql://foo:bar@localhost")
 username = uri.userInfo ? uri.userInfo.split(":")[0] : ""
 password = uri.userInfo ? uri.userInfo.split(":")[1] : ""
 url = "jdbc:mysql://" + uri.host + uri.path

 properties {
 dbProperties {
 autoReconnect = true
 }
 }
}
```

## Step 4: (Optional) Configure a Procfile

Use a Procfile to declare required runtime processes for your web app and to specify your web server. For more information, see the [Configuring a Production Server](#) topic.

#### Sample App Step

You can skip this step. The `pong_matcher_grails` app does not require a Procfile.

## Step 5: Create and Bind a Service Instance for a Grails Application

This section describes using the cf CLI to configure a ClearDB managed service instance for an app. You can use either the CLI or [Apps Manager](#) to perform this task.

PAS supports two types of service instances:

- Managed services integrate with PAS through service brokers that offer services and plans and manage the service calls between PAS and a service provider.
- User-provided service instances enable you to connect your application to pre-provisioned external service instances.

For more information about creating and using service instances, refer to the [Services Overview](#) topic.

## Create a Service Instance

Run the cf CLI `cf marketplace` command to view managed and user-provided services and plans available to you.

The example shows two of the available managed database-as-a-service providers and their offered plans: `cleardb` database-as-a-service for MySQL-powered apps and `postgresql-10-odb` PostgreSQL as a Service.

```
$ cf marketplace
Getting services from marketplace in org Cloud-Apps / space development as cloudunder@example.com...
OK

service plans description
cleardb spark, boost, amp Highly available MySQL for your apps
postgresql-10-odb standalone, standalone-replica, general PostgreSQL as a Service
```

Run `cf create-service SERVICE PLAN SERVICE-INSTANCE` to create a service instance for your app. Choose a SERVICE and PLAN from the list, and provide a unique name for the SERVICE-INSTANCE.

### Sample App Step

Run `cf create-service cleardb spark mysql`. This creates a service instance named `mysql` that uses the `cleardb` service and the `spark` plan, as the example below shows.

```
$ cf create-service cleardb spark mysql
Creating service mysql in org Cloud-Apps / space development as cloudunder@example.com....
OK
```

## Bind a Service Instance

When you bind an app to a service instance, PAS writes information about the service instance to the VCAP\_SERVICES app environment variable. The app can use this information to integrate with the service instance.

Most services support bindable service instances. Refer to your service provider's documentation to confirm if they support this functionality.

You can bind a service to an application with the command `cf bind-service APPLICATION SERVICE-INSTANCE`.

Alternately, you can configure the deployment manifest file by adding a `services` sub-block to the `applications` block and specifying the service instance. For more information and an example on service binding using a manifest, see the Sample App step.

You can also bind a service using the [Apps Manager](#).

### Sample App Step

You can skip this step because the service instance is already bound. Open the `manifest.yml` file in a text editor to view the bound service instance information. Locate the file in the app root directory and search for the `services` sub-block in the `applications` block, as the example below shows.

```

applications:
...
 services:
 - mysql
```

## Step 6: Configure the Deployment Manifest

You can specify deployment options in the `manifest.yml` that the `cf push` command uses when deploying your app.

Refer to the [Deploying with Application Manifests](#) topic for more information.

### Sample App Step

You can skip this step. The `manifest.yml` file for the `pong_matcher_grails` sample app does not require any additional configuration to deploy the app.

## Step 7: Log in and Target the API Endpoint

Run `cf login -a API-ENDPOINT`, enter your login credentials, and select a space and org. The API endpoint is [the URL of the Cloud Controller in your PAS instance](#).

### Sample App Step

You must do this step to run the sample app.

## Step 8: Deploy the Application

 **Note:** You must use the cf CLI to deploy apps.

From the root directory of your application, run `cf push APP-NAME -p PATH-TO-FILE.war` to deploy your application.

 **Note:** You must deploy the `.war` artifact for a Grails app, and you must include the path to the `.war` file in the `cf push` command using the `-p` option if you do not declare the path in the `applications` block of the manifest file. For more information, refer to the [Grails section](#) in the Tips for Java Developers topic.

`cf push APP-NAME` creates a URL route to your application in the form `HOST.DOMAIN`, where `HOST` is your `APP-NAME` and `DOMAIN` is specified by your administrator. Your `DOMAIN` is `shared-domain.example.com`. For example: `cf push my-app` creates the URL `my-app.shared-domain.example.com`.

The URL for your app must be unique from other apps that PAS hosts or the push will fail. Use the following options to help create a unique URL:

- `-n` to assign a different HOST name for the app
- `--random-route` to create a URL that includes the app name and random words
- `cf help push` to view other options for this command

If you want to view log activity while the app deploys, launch a new terminal window and run `cf logs APP-NAME`.

Once your app deploys, browse to your app URL. Search for the `urls` field in the `App started` block in the output of the `cf push` command. Use the URL to access your app online.

### Sample App Step

1. Change to the `app` directory, and run `./grails war` to build the app.

2. Run `cf push pong_matcher_grails -n HOST-NAME` to push the app.

Example: `cf push pong_matcher_grails -n my-grails-app`

**Note:** You do not have to include the `-p` flag when you deploy the sample app. The sample app manifest declares the path to the archive that `cf push` uses to upload the app files.

The example below shows the terminal output of deploying the `pong_matcher_grails` app. `cf push` uses the instructions in the manifest file to create the app, create and bind the route, and upload the app. It then binds the app to the `mysql` service and follows the instructions in the manifest to start two instances of the app, allocating 1 GB of memory between the instances. After the instances start, the output displays their health and status.

```
$ cf push pong_matcher_grails -n my-grails-app
Using manifest file /Users/example/workspace/pong_matcher_grails/app/manifest.yml

Creating app pong_matcher_grails in org Cloud-Apps / space development as cloudunder@example.com...
OK

Creating route my-grails-app.cfapps.io...
OK

Binding my-grails-app.cfapps.io to pong_matcher_grails...
OK

Uploading pong_matcher_grails...
Uploading app files from: /Users/example/workspace/pong_matcher_grails/app/target/pong_matcher_grails-0.1.war
Uploading 4.8M, 704 files
OK
Binding service mysql to app pong_matcher_grails in org Cloud-Apps / space development as cloudunder@example.com...
OK

Starting app pong_matcher_grails in org Cloud-Apps / space development as cloudunder@example.com...
OK
----> Downloaded app package (38M)
----> Java Buildpack Version: v2.5 | https://github.com/cloudfoundry/java-buildpack.git#840500e
----> Downloading Open Jdk JRE 1.8.0_25 from https://download.run.pivotal.io/openjdk/lucid/x86_64/openjdk-1.8.0_25.tar.gz (1.5s)
 Expanding Open Jdk JRE to java-buildpack/open_jdk_jre (1.1s)
----> Downloading Spring Auto Reconfiguration 1.5.0_RELEASE from https://download.run.pivotal.io/auto-reconfiguration/auto-reconfiguration-1.5.0_RELEASE.jar (0.0s)
 Modifying /WEB-INF/web.xml for Auto Reconfiguration
----> Downloading Tomcat Instance 8.0.14 from https://download.run.pivotal.io/tomcat-8.0.14.tar.gz (0.4s)
 Expanding Tomcat to java-buildpack/tomcat (0.1s)
----> Downloading Tomcat Lifecycle Support 2.4.0 RELEASE from https://download.run.pivotal.io/tomcat-lifecycle-support/tomcat-lifecycle-support-2.4.0_RELEASE.jar (0.0s)
----> Downloading Tomcat Logging Support 2.4.0_RELEASE from https://download.run.pivotal.io/tomcat-logging-support/tomcat-logging-support-2.4.0_RELEASE.jar (0.0s)
----> Downloading Tomcat Access Logging Support 2.4.0_RELEASE from https://download.run.pivotal.io/tomcat-access-logging-support/tomcat-access-logging-support-2.4.0_RELEASE.jar (0.0s)
----> Uploading droplet (83M)

0 of 2 instances running, 2 starting
0 of 2 instances running, 2 starting
0 of 2 instances running, 2 starting
2 of 2 instances running

App started

Showing health and status for app pong_matcher_grails in org Cloud-Apps / space development as cloudunder@example.com...
OK

requested state: started
instances: 2/2
usage: 1G x 2 instances
urls: my-grails-app.cfapps.io

 state since cpu memory disk
#0 running 2014-11-10 05:07:33 PM 0.0% 686.4M of 1G 153.6M of 1G
#1 running 2014-11-10 05:07:36 PM 0.0% 677.2M of 1G 153.6M of 1G
```

## Step 9: Test Your Deployed App

You've deployed an app to PAS!

Use the cf CLI or [Apps Manager](#) to review information and administer your app and your PAS account. For example, you can edit the `manifest.yml` to increase the number of app instances from 1 to 3, and redeploy the app with a new app name and host name.

See the [Manage Your Application with the cf CLI](#) section for more information. See also [Using the Apps Manager](#).

### Sample App Step

To test the sample app, do the following:

1. To export the test host, run `export HOST=SAMPLE-APP-URL`, substituting the URL for your app for `SAMPLE-APP-URL`.

2. To clear the database from any previous tests, run:

```
curl -v -X DELETE
$HOST/all
```

You should get a response of 200.

3. To request a match as “andrew”, run:

```
curl -v -H "Content-Type: application/json" -X PUT $HOST/match_requests/firstrequest -d '{"player":
"andrew"}'
```

You should again get a response of 200.

4. To request a match as a different player, run:

```
curl -v -H "Content-Type: application/json" -X PUT $HOST/match_requests/secondrequest -d '{"player":
"navratilova"}'
```

5. To check the status of the first match request, run:

```
curl -v -X GET
$HOST/match_requests/firstrequest
```

The last line of the output shows the `match_id`.

6. Replace `MATCH_ID` with the `match_id` value from the previous step in the following command:

```
curl -v -H "Content-Type: application/json" -X POST $HOST/results -d '{ "match_id":'MATCH_ID", "winner": "andrew", "loser": "navratilova"
}'
```

You should receive a 201 response.

## Manage Your Application with the cf CLI

Run `cf help` to view a complete list of commands, grouped by task categories, and run `cf help COMMAND` for detailed information about a specific command.

For more information about using the cf CLI, refer to the Cloud Foundry Command Line Interface (cf CLI) topics, especially the [Getting Started with cf CLI](#) topic.

 **Note:** You cannot perform certain tasks in the CLI or [Apps Manager](#) because these are commands that only a PAS administrator can run. If you are not a PAS administrator, the following message displays for these types of commands:  
`error code: 10003, message: You are not authorized to perform the requested action` For more information about specific Admin commands you can perform with the Apps Manager, depending on your user role, refer to the [Getting Started with the Apps Manager](#) topic.

## Troubleshooting

If your application fails to start, verify that the application starts in your local environment. Refer to the [Troubleshooting Application Deployment and Health](#) topic to learn more about troubleshooting.

## App Deploy Fails

Even when the deploy fails, the app might exist on PAS. Run `cf apps` to review the apps in the currently targeted org and space. You might be able to correct the issue using the CLI or [Apps Manager](#), or you might have to delete the app and redeploy.

## App Requires a Unique URL

PAS requires that each app that you deploy have a unique URL. Otherwise, the new app URL collides with an existing app URL and PAS cannot successfully deploy the app. You can fix this issue by running `cf push` with either of the following flags to create a unique URL:

- `-n` to assign a different HOST name for the app.

- `--random-route` to create a URL that includes the app name and random words. Using this option might create a long URL, depending on the number of words that the app name includes.

## Getting Started Deploying Ratpack Apps

This guide is intended to walk you through deploying a Ratpack app to Pivotal Application Service (PAS). If you experience a problem following the steps below, check the [Known Issues](#) topic or refer to the [Troubleshooting Application Deployment and Health](#) topic.

### Sample App Step

If you want to go through this tutorial using the sample app, run `git clone https://github.com/cloudfoundry-samples/pong_matcher_groovy.git` to clone the `pong_matcher_groovy` app from GitHub, and follow the instructions in the Sample App Step sections.

 **Note:** Ensure that your Ratpack app runs locally before continuing with this procedure.

## Deploy a Ratpack Application

This section describes how to deploy a Ratpack application to PAS.

### Prerequisites

- A Ratpack app that runs locally on your workstation
- Intermediate to advanced Ratpack knowledge
- The [Cloud Foundry Command Line Interface \(cf CLI\)](#)
- JDK 1.7 or 1.8 for Java 7 or 8 configured on your workstation

 **Note:** You can develop Ratpack applications in Java 7 or 8 or any JVM language. The Cloud Foundry Java buildpack uses JDK 1.8, but you can modify the buildpack and the manifest for your app to compile to JDK 1.7. Refer to Step 8: Configure the Deployment Manifest.

## Step 1: Declare App Dependencies

Declare all the dependency tasks for your app in the build script of your chosen build tool. The table lists build script information for Gradle and Maven and provides documentation links for each build tool.

Build Tool	Build Script	Documentation
Gradle	<code>build.gradle</code>	<a href="#">Gradle User Guide</a>
Maven	<code>pom.xml</code>	<a href="#">Apache Maven Project Documentation</a>

### Sample App Step

You can skip this step. The `build.gradle` file contains the dependencies for the `pong_matcher_groovy` sample app, as the example below shows.

```
dependencies {
 // SpringLoaded enables runtime hot reloading.
 // It is not part of the app runtime and is not shipped in the distribution.
 springloaded "org.springframework:springloaded:1.2.0.RELEASE"

 // Default SLF4J binding. Note that this is a blocking implementation.
 // See here for a non blocking appender http://logging.apache.org/log4j/2.x/manual/async.html
 runtime 'org.slf4j:slf4j-simple:1.7.7'

 compile group: 'redis.clients', name: ' jedis', version: '2.5.2', transitive: true

 testCompile "org.spockframework:spock-core:0.7-groovy-2.0"
}
```

## Step 2: Allocate Sufficient Memory

Use the `cf push -m` command to specify the amount of memory that should be allocated to the application. Memory allocated this way is done in preset amounts of `64M`, `128M`, `256M`, `512M`, `1G`, or `2G`. For example:

```
$ cf push -m 128M
```

When your app is running, you can use the `cf app APP-NAME` command to see memory utilization.

### Sample App Step

You can skip this step. In the `manifest.yml` of the `pong_matcher_groovy` sample app, the `memory` sub-block of the `applications` block allocates 512 MB to the app.

## Step 3: Provide a JDBC Driver

The Java buildpack does not bundle a JDBC driver with your application. If your application accesses a SQL RDBMS, you must do the following:

- Include the appropriate driver in your application.
- Create a dependency task for the driver in the build script for your build tool or IDE.

### Sample App Step

You can skip this step. The `pong_matcher_groovy` sample app does not require a JDBC driver.

## Step 4: (Optional) Configure a Procfile

Use a Procfile to declare required runtime processes for your web app and to specify your web server. For more information, see the [Configuring a Production Server](#) topic.

### Sample App Step

You can skip this step. The `pong_matcher_groovy` app does not require a Procfile.

## Step 5: Create and Bind a Service Instance for a Ratpack Application

This section describes using the CLI to configure a Redis managed service instance for an app. You can use either the CLI or [Apps Manager](#) to perform this task.

PAS supports two types of service instances:

- Managed services integrate with PAS through service brokers that offer services and plans and manage the service calls between PAS and a service provider.
- User-provided service instances enable you to connect your application to pre-provisioned external service instances.

For more information about creating and using service instances, refer to the [Services Overview](#) topic.

### Create a Service Instance

Run `cf marketplace` to view managed and user-provided services and plans available to you.

The example shows two of the available managed database-as-a-service providers and their offered plans: `postgresql-10-odb` PostgreSQL as a Service and `rediscloud` Enterprise-Class Redis for Developers.

```
$ cf marketplace
Getting services from marketplace in org Cloud-Apps / space development as clouduser@example.com...
OK
```

service	plans	description
postgresql-10-odb	standalone, standalone-replica, general	PostgreSQL as a Service
rediscloud	30mb, 100mb, 1gb, 10gb, 50gb	Enterprise-Class Redis for Developers

Run `cf create-service SERVICE PLAN SERVICE-INSTANCE` to create a service instance for your app. Choose a SERVICE and PLAN from the list, and provide a unique name for the SERVICE-INSTANCE.

#### Sample App Step

Run `cf create-service rediscloud 30mb baby-redis`. This creates a service instance named `baby-redis` that uses the `rediscloud` service and the `30mb` plan, as the example below shows.

```
$ cf create-service rediscloud 30mb baby-redis
Creating service baby-redis in org Cloud-Apps / space development as clouduser@example.com....
OK
```

## Bind a Service Instance

When you bind an app to a service instance, PAS writes information about the service instance to the `VCAP_SERVICES` app environment variable. The app can use this information to integrate with the service instance.

Most services support bindable service instances. Refer to your service provider's documentation to confirm if they support this functionality.

You can bind a service to an application with the command `cf bind-service APPLICATION SERVICE_INSTANCE`.

Alternately, you can configure the deployment manifest file by adding a `services` sub-block to the `applications` block and specifying the service instance. For more information and an example on service binding using a manifest, see the Sample App step.

You can also bind a service using the [Apps Manager](#).

#### Sample App Step

You can skip this step because the service instance is already bound. Open the `manifest.yml` file in a text editor to view the bound service instance information. Locate the file in the app root directory and search for the `services` sub-block in the `applications` block, as the example below shows.

```

applications:
...
services:
- baby-redis
```

## Step 6: Configure the Deployment Manifest

You can specify deployment options in the `manifest.yml` that the `cf push` command uses when deploying your app.

Refer to the [Deploying with Application Manifests](#) topic for more information.

#### Sample App Step

You can skip this step. The `manifest.yml` file for the `pong_matcher_groovy` sample app does not require any additional configuration to deploy the app.

## Step 7: Log in and Target the API Endpoint

Run `cf login -a API-ENDPOINT`, enter your login credentials, and select a space and org. The API endpoint is [the URL of the Cloud Controller in your PAS instance](#).

### Sample App Step

You must perform this step to run the sample app.

## Step 8: Deploy the Application

 **Note:** You must use the cf CLI to deploy apps.

From the root directory of your application, run `cf push APP-NAME -p PATH-TO-FILE.distZip` to deploy your application.

 **Note:** You must deploy the `.distZip` artifact for a Ratpack app, and you must include the path to the `.distZip` file in the `cf push` command using the `-p` option if you do not declare the path in the `applications` block of the manifest file. For more information, refer to the [Tips for Java Developers](#) topic.

`cf push APP-NAME` creates a URL route to your application in the form `HOST.DOMAIN`, where `HOST` is your `APP-NAME` and `DOMAIN` is specified by your administrator. Your `DOMAIN` is `shared-domain.example.com`. For example: `cf push my-app` creates the URL `my-app.shared-domain.example.com`.

The URL for your app must be unique from other apps that PAS hosts or the push will fail. Use the following options to help create a unique URL:

- `-n` to assign a different HOST name for the app
- `--random-route` to create a URL that includes the app name and random words
- `cf help push` to view other options for this command

If you want to view log activity while the app deploys, launch a new terminal window and run `cf logs APP-NAME`.

Once your app deploys, browse to your app URL. Search for the `urls` field in the `App started` block in the output of the `cf push` command. Use the URL to access your app online.

### Sample App Step

1. Change to the `app` directory, and run `./gradlew distZip` to build the app.

2. Run `cf push pong_matcher_groovy -n HOST-NAME` to push the app.

Example: `cf push pong_matcher_groovy -n groovy-ratpack-app`

 **Note:** You do not have to include the `-p` flag when you deploy the sample app. The sample app manifest declares the path to the archive that `cf push` uses to upload the app files.

The example below shows the terminal output of deploying the `pong_matcher_groovy` app. `cf push` uses the instructions in the manifest file to create the app, create and bind the route, and upload the app. It then binds the app to the `baby-redis` service and follows the instructions in the manifest to start one instance of the app with 512 MB. After the app starts, the output displays the health and status of the app.

```
$ cf push pong_matcher_groovy -n groovy-ratpack-app
Using manifest file /Users/example/workspace/pong_matcher_groovy/app/manifest.yml

Creating app pong_matcher_groovy in org Cloud-Apps / space development as cloudunder@example.com...
OK

Creating route groovy-ratpack-app.cfapps.io...
OK

Binding groovy-ratpack-app.cfapps.io to pong_matcher_groovy...
OK

Uploading pong_matcher_groovy...
Uploading app files from: /Users/example/workspace/pong_matcher_groovy/app/build/distributions/app.zip
Uploading 138.2K, 18 files
OK
Binding service baby-redis to app pong_matcher_groovy in org Cloud-Apps / space development as cloudunder@example.com...
OK

Starting app pong_matcher_groovy in org Cloud-Apps / space development as cloudunder@example.com...
OK
----> Downloaded app package (12M)
Cloning into '/tmp/buildpacks/java-buildpack'...
----> Java Buildpack Version: 9e096be | https://github.com/cloudfoundry/java-buildpack#9e096be
 Expanding Open Jdk JRE to .java-buildpack/open_jdk_jre (1.3s)
----> Uploading droplet (49M)

0 of 1 instances running, 1 starting
1 of 1 instances running

App started

Showing health and status for app pong_matcher_groovy in org Cloud-Apps / space development as cloudunder@example.com...
OK

requested state: started
instances: 1/1
usage: 512M x 1 instances
urls: groovy-ratpack-app.cfapps.io

 state since cpu memory disk
#0 running 2014-10-28 04:48:58 PM 0.0% 193.5M of 512M 111.7M of 1G
```

## Step 9: Test Your Deployed App

You've deployed an app to PAS!

Use the cf CLI or [Apps Manager](#) to review information and administer your app and your PAS account. For example, you can edit the `manifest.yml` to increase the number of app instances from 1 to 3, and redeploy the app with a new app name and host name.

See the [Manage Your Application with the cf CLI](#) section for more information. See also [Using the Apps Manager](#).

### Sample App Step

To test the sample app, do the following:

1. To export the test host, run `export HOST=SAMPLE-APP-URL`, substituting the URL for your app for `SAMPLE-APP-URL`.

2. To clear the database from any previous tests, run:

```
curl -v -X DELETE
$HOST/all
```

You should get a response of 200.

3. To request a match as "andrew", run:

```
curl -v -H "Content-Type: application/json" -X PUT $HOST/match_requests/firstrequest -d '{"player": "andrew"}'
```

You should again get a response of 200.

4. To request a match as a different player, run:

```
curl -v -H "Content-Type: application/json" -X PUT $HOST/match_requests/secondrequest -d '{"player": "navratilova"}'
```

5. To check the status of the first match request, run:

```
curl -v -X GET
$HOST/match_requests/firstrequest
```

The last line of the output shows the `match_id`.

6. Replace `MATCH_ID` with the `match_id` value from the previous step in the following command:

```
curl -v -H "Content-Type: application/json" -X POST $HOST/results -d '{ "match_id": "MATCH_ID", "winner": "andrew", "loser": "navratilova" }'
```

You should receive a `201 Created` response.

## Manage Your Application with the cf CLI

Run `cf help` to view a complete list of commands, grouped by task categories, and run `cf help COMMAND` for detailed information about a specific command. For more information about using the cf CLI, refer to the Cloud Foundry Command Line Interface (cf CLI) topics, especially the [Getting Started with cf CLI](#) topic.

 **Note:** You cannot perform certain tasks in the CLI or [Apps Manager](#) because these are commands that only a PAS administrator can run. If you are not a PAS administrator, the following message displays for these types of commands:

error code: 10003, message: You are not authorized to perform the requested action For more information about specific Admin commands you can perform with the Apps Manager, depending on your user role, refer to the [Getting Started with the Apps Manager](#) topic.

## Troubleshooting

If your application fails to start, verify that the application starts in your local environment. Refer to the [Troubleshooting Application Deployment and Health](#) topic to learn more about troubleshooting.

## App Deploy Fails

Even when the deploy fails, the app might exist on PAS. Run `cf apps` to review the apps in the currently targeted org and space. You might be able to correct the issue using the CLI or [Apps Manager](#), or you might have to delete the app and redeploy.

## App Requires a Unique URL

PAS requires that each app that you deploy have a unique URL. Otherwise, the new app URL collides with an existing app URL and PAS cannot successfully deploy the app. You can fix this issue by running `cf push` with either of the following flags to create a unique URL:

- `-n` to assign a different HOST name for the app.
- `--random-route` to create a URL that includes the app name and random words. Using this option might create a long URL, depending on the number of words that the app name includes.

## Getting Started Deploying Spring Apps

This guide is intended to walk you through deploying a Spring app to Pivotal Application Service (PAS). You can choose whether to push a sample app, your own app, or both.

If you experience a problem following the steps below, see the [Known Issues](#) topic, or refer to the [Troubleshooting Application Deployment and Health](#) topic.

### Sample App Step

If you want to go through this tutorial using the sample app, run `git clone https://github.com/cloudfoundry-samples/pong_matcher_spring` to clone the `pong_matcher_spring` app from GitHub, and follow the instructions in the Sample App Step sections.

 **Note:** Ensure that your Spring app runs locally before continuing with this procedure.

## Deploy a Spring Application

This section describes how to deploy your Spring application to PAS.

### Prerequisites

- A Spring app that runs locally on your workstation
- Intermediate to advanced Spring knowledge
- The [Cloud Foundry Command Line Interface \(cf CLI\)](#)
- JDK 1.6, 1.7, or 1.8 for Java 6, 7, or 8 configured on your workstation

 **Note:** The Cloud Foundry Java buildpack uses JDK 1.8, but you can modify the buildpack and the manifest for your app to compile to an earlier version. For more information, refer to the [Custom Buildpacks](#) topic.

## Step 1: Declare App Dependencies

Be sure to declare all the dependency tasks for your app in the build script of your chosen build tool.

The [Spring Getting Started Guides](#) demonstrate features and functionality you can add to your app, such as consuming RESTful services or integrating data. These guides contain Gradle and Maven build script examples with dependencies. You can copy the code for the dependencies into your build script.

The table lists build script information for Gradle and Maven and provides documentation links for each build tool.

Build Tool	Build Script	Documentation
Gradle	<code>build.gradle</code>	<a href="#">Gradle User Guide</a>
Maven	<code>pom.xml</code>	<a href="#">Apache Maven Project Documentation</a>

### Sample App Step

You can skip this step. The `pom.xml` file contains the dependencies for the `pong_matcher_spring` sample app, as the example below shows.

```
<dependencies>
 <dependency>
 <groupId>mysql</groupId>
 <artifactId>mysql-connector-java</artifactId>
 </dependency>
 <dependency>
 <groupId>org.flywaydb</groupId>
 <artifactId>flyway-core</artifactId>
 </dependency>
 <dependency>
 <groupId>org.springframework.boot</groupId>
 <artifactId>spring-boot-starter-data-jpa</artifactId>
 </dependency>
 <dependency>
 <groupId>org.springframework.boot</groupId>
 <artifactId>spring-boot-starter-test</artifactId>
 </dependency>
 <dependency>
 <groupId>org.springframework.boot</groupId>
 <artifactId>spring-boot-starter-web</artifactId>
 </dependency>
 <dependency>
 <groupId>com.h2database</groupId>
 <artifactId>h2</artifactId>
 <scope>test</scope>
 </dependency>
 <dependency>
 <groupId>com.jayway.jsonpath</groupId>
 <artifactId>json-path</artifactId>
 <scope>test</scope>
 </dependency>
</dependencies>
```

 **Note:** Make sure you are not building [fully executable jars](#) because application push may fail.

## Step 2: Allocate Sufficient Memory

Use the `cf push -m` command to specify the amount of memory that should be allocated to the application. Memory allocated this way is done in preset amounts of `64M`, `128M`, `256M`, `512M`, `1G`, or `2G`. For example:

```
$ cf push -m 128M
```

When your app is running, you can use the `cf app APP-NAME` command to see memory utilization.

### Sample App Step

You can skip this step. The Cloud Foundry Java buildpack uses settings declared in the sample app to allocate 1 GB of memory to the app.

## Step 3: Provide a JDBC Driver

The Java buildpack does not bundle a JDBC driver with your application. If your application accesses a SQL RDBMS, you must do the following:

- Include the appropriate driver in your application.
- Create a dependency task for the driver in the build script for your build tool or IDE.

### Sample App Step

You can skip this step. In the `pong_matcher_spring` sample app, the `src/main/resources/application.yml` file declares the JDBC driver, and the `pom.xml` file includes the JDBC driver as a dependency.

## Step 4: Configure Service Connections for a Spring App

PAS provides extensive support for creating and binding a Spring application to services such as MySQL, PostgreSQL, MongoDB, Redis, and RabbitMQ. For more information about creating and binding a service connection for your app, refer to the [Configure Service Connections for Spring](#) topic.

#### Sample App Step: Create a Service Instance

Run `cf create-service cleardb spark mysql`. This creates a service instance named `mysql` that uses the `cleardb` service and the `spark` plan, as the example below shows.

```
$ cf create-service cleardb spark mysql
Creating service mysql in org Cloud-Apps / space development as a.user@example.com....
OK
```

#### Sample App Step: Bind a Service Instance

You can skip this step because the service instance is already bound. Open the `manifest.yml` file in a text editor to view the bound service instance information. Locate the file in the app root directory and search for the `services` sub-block in the `applications` block, as the example below shows.

```

applications:
...
 services:
 - mysql
```

## Step 5: Configure the Deployment Manifest

You can specify deployment options in a manifest file `manifest.yml` that the `cf push` command uses when deploying your app.

Refer to the [Deploying with Application Manifests](#) topic for more information.

#### Sample App Step

You can skip this step. The `manifest.yml` file for the `pong_matcher_spring` sample app does not require any additional configuration to deploy the app.

## Step 6: Log in and Target the API Endpoint

Run `cf login -a API-ENDPOINT`, enter your login credentials, and select a space and org. The API endpoint is [the URL of the Cloud Controller in your PAS instance](#).

#### Sample App Step

You must do this step to run the sample app.

## Step 7: Deploy Your Application

 **Note:** You must use the cf CLI to deploy apps.

From the root directory of your application, run `cf push APP-NAME -p PATH-TO-FILE.jar` to deploy your application.

 **Note:** Most Spring apps include an artifact, such as a `.jar`, `.war`, or `.zip` file. You must include the path to this file in the `cf push` command using the `-p` option if you do not declare the path in the `applications` block of the manifest file. The example shows how to specify a path to the `.jar` file

for a Spring app. Refer to the [Tips for Java Developers](#) topic for CLI examples for specific build tools, frameworks, and languages that create an app with an artifact.

`cf push APP-NAME` creates a URL route to your application in the form HOST.DOMAIN, where HOST is your APP-NAME and DOMAIN is specified by your administrator. Your DOMAIN is `shared-domain.example.com`. For example: `cf push my-app` creates the URL `my-app.shared-domain.example.com`.

The URL for your app must be unique from other apps that PAS hosts or the push will fail. Use the following options to help create a unique URL:

- `-n` to assign a different HOST name for the app
- `--random-route` to create a URL that includes the app name and random words
- `cf help push` to view other options for this command

If you want to view log activity while the app deploys, launch a new terminal window and run `cf logs APP-NAME`.

Once your app deploys, browse to your app URL. Search for the `urls` field in the `App started` block in the output of the `cf push` command. Use the URL to access your app online.

### Sample App Step

1. Run `brew install maven`.
2. Change to the `app` directory, and run `mvn package` to build the app.
3. Run `cf push pong_matcher_spring -n` `HOSTNAME` to push the app.

Example: `cf push pong_matcher_spring -n my-spring-app`

 **Note:** You do not have to include the `-p` flag when you deploy the sample app. The sample app manifest declares the path to the archive that `cf push` uses to upload the app files.

The example below shows the terminal output of deploying the `pong_matcher_spring` app. `cf push` uses the instructions in the manifest file to create the app, create and bind the route, and upload the app. It then binds the app to the `mysql` service and starts one instance of the app with 1 GB of memory. After the app starts, the output displays the health and status of the app.

```
$ cf push pong_matcher_spring -n spring1119
Using manifest file /Users/example/workspace/pong_matcher_spring/manifest.yml

Creating app pong_matcher_spring in org Cloud-Apps / space development as a.user@example.com...
OK

Creating route spring1119.cfapps.io...
OK

Binding spring1119.cfapps.io to pong_matcher_spring...
OK

Uploading pong_matcher_spring...
Uploading app files from: /Users/example/workspace/pong_matcher_spring/target/pong-matcher-spring-1.0.0.BUILD-SNAPSHOT.jar
Uploading 797.5K, 116 files
OK
Binding service mysql to app pong_matcher_spring in org Cloud-Apps / space development as a.user@example.com...
OK

Starting app pong_matcher_spring in org Cloud-Apps / space development as a.user@example.com...
OK
----> Downloaded app package (25M)
----> Downloading Open Jdk JRE 1.8.0_25 from https://download.run.pivotal.io/openjdk/lucid/x86_64/openjdk-1.8.0_25.tar.gz (1.2s)
 Expanding Open Jdk JRE to .java-buildpack/open_jdk_jre (1.1s)
----> Downloading Spring Auto Reconfiguration 1.5.0 RELEASE from https://download.run.pivotal.io/auto-reconfiguration/auto-reconfiguration-1.5.0_RELEASE.jar (0.1s)

----> Uploading droplet (63M)

0 of 1 instances running, 1 starting
1 of 1 instances running

App started

Showing health and status for app pong_matcher_spring in org Cloud-Apps / space development as a.user@example.com...
OK

requested state: started
instances: 1/1
usage: 1G x 1 instances
urls: spring1119.cfapps.io

 state since cpu memory disk
#0 running 2014-11-19 12:29:27 PM 0.0% 553.6M of 1G 127.4M of 1G
```

## Step 8: Test Your Deployed App

You've deployed an app to PAS!

Use the cf CLI or [Apps Manager](#) to review information and administer your app and your PAS account. For example, you can edit the `manifest.yml` to increase the number of app instances from 1 to 3, and redeploy the app with a new app name and host name.

See the [Manage Your Application with the cf CLI](#) section for more information. See also [Using the Apps Manager](#).

### Sample App Step

To test the sample app, do the following:

1. To export the test host, run `export HOST=SAMPLE-APP-URL`, substituting the URL for your app for `SAMPLE-APP-URL`.

2. To clear the database from any previous tests, run:

```
curl -v -X DELETE
$HOST/all
```

You should get a response of 200.

3. To request a match as "andrew", run:

```
curl -v -H "Content-Type: application/json" -X PUT $HOST/match_requests/firstrequest -d '{"player": "andrew"}'
```

You should again get a response of 200.

4. To request a match as a different player, run:

```
curl -v -H "Content-Type: application/json" -X PUT $HOST/match_requests/secondrequest -d '{"player": "navratilova"}'
```

5. To check the status of the first match request, run:

```
curl -v -X GET
$HOST/match_requests/firstrequest
```

The last line of the output shows the `match_id`.

6. Replace `MATCH_ID` with the `match_id` value from the previous step in the following command:

```
curl -v -H "Content-Type: application/json" -X POST $HOST/results -d '{ "match_id": "MATCH_ID", "winner": "andrew", "loser": "navratilova" }'
```

You should receive a `201 Created` response.

## Manage Your App with the cf CLI

Run `cf help` to view a complete list of commands, grouped by task categories, and run `cf help COMMAND` for detailed information about a specific command. For more information about using the cf CLI, refer to the Cloud Foundry Command Line Interface (cf CLI) topics, especially the [Getting Started with cf CLI](#) topic.

 **Note:** You cannot perform certain tasks in the CLI or [Apps Manager](#) because these are commands that only a PAS administrator can run. If you are not a PAS administrator, the following message displays for these types of commands:

`error code: 10003, message: You are not authorized to perform the requested action` For more information about specific Admin commands you can perform with the Apps Manager, depending on your user role, refer to the [Getting Started with the Apps Manager](#) topic.

## Troubleshooting

If your application fails to start, verify that the application starts in your local environment. Refer to the [Troubleshooting Application Deployment and Health](#) topic to learn more about troubleshooting.

### App Deploy Fails

Even when the deploy fails, the app might exist on PAS. Run `cf apps` to review the apps in the currently targeted org and space. You might be able to correct the issue using the CLI or [Apps Manager](#), or you might have to delete the app and redeploy.

### App Requires a Content-Type

If you specify a `Content-Encoding` header of `gzip` but do not specify a `Content-Type` within your application, PAS might send a `Content-Type` of `application/x-gzip` to the browser. This scenario might cause the deploy to fail if it conflicts with the actual encoded content of your app. To avoid this issue, be sure to explicitly set `Content-Type` within your app.

### App Requires a Unique URL

PAS requires that each app that you deploy have a unique URL. Otherwise, the new app URL collides with an existing app URL and PAS cannot successfully deploy the app. You can fix this issue by running `cf push` with either of the following flags to create a unique URL:

- `-n` to assign a different HOST name for the app.
- `--random-route` to create a URL that includes the app name and random words. Using this option might create a long URL, depending on the number of words that the app name includes.

## Configuring Service Connections

This topic provides links to additional information about configuring service connections for Java apps. See the specific documentation for your app framework:

- [Grails](#)
- [Play](#)
- [Spring](#)

## Configuring Service Connections for Grails

Cloud Foundry provides extensive support for connecting a Grails application to services such as MySQL, Postgres, MongoDB, Redis, and RabbitMQ. In many cases, a Grails application running on Cloud Foundry can automatically detect and configure connections to services. For more advanced cases, you can control service connection parameters yourself.

### Auto-Configuration

Grails provides plugins for accessing SQL (using [Hibernate](#)), [MongoDB](#), and [Redis](#) services. If you install any of these plugins and configure them in your `Config.groovy` or `DataSource.groovy` file, Cloud Foundry reconfigures the plugin with connection information when your app starts.

If you use all three types of services, your configuration might look like this:

```
environments {
 production {
 dataSource {
 url = 'jdbc:mysql://localhost/db?useUnicode=true&characterEncoding=utf8'
 dialect = org.hibernate.dialect.MySQLInnoDBDialect
 driverClassName = 'com.mysql.jdbc.Driver'
 username = 'user'
 password = "password"
 }
 grails {
 mongo {
 host = 'localhost'
 port = 27107
 databaseName = "foo"
 username = 'user'
 password = 'password'
 }
 redis {
 host = 'localhost'
 port = 6379
 password = 'password'
 timeout = 2000
 }
 }
 }
}
```

The `url`, `host`, `port`, `databaseName`, `username`, and `password` fields in this configuration will be overridden by the Cloud Foundry auto-reconfiguration if it detects that the application is running in a Cloud Foundry environment. If you want to test the application locally against your own services, you can put real values in these fields. If the application will only be run against Cloud Foundry services, you can put placeholder values as shown here, but the fields must exist in the configuration.

### Manual Configuration

If you do not want to use auto-configuration, you can configure the Cloud Foundry service connections manually.

Follow the steps below to manually configure a service connection.

1. Add the `spring-cloud` library to the `dependencies` section of your `BuildConfig.groovy` file.

```
repositories {
 grailsHome()
 mavenCentral()
 grailsCentral()
 mavenRepo "http://repo.spring.io/milestone"
}

dependencies {
 compile "org.springframework.cloud:spring-cloud-cloudfoundry-connector:1.0.0.RELEASE"
 compile "org.springframework.cloud:spring-cloud-spring-service-connector:1.0.0.RELEASE"
}
```

Adding the `spring-cloud` library allows you to disable auto-configuration and use the `spring-cloud` API in your `DataSource.groovy` file to set the connection parameters.

2. Add the following to your `grails-app/conf/spring/resources.groovy` file to disable auto-configuration:

```
beans = {
 cloudFactory(org.springframework.cloud.CloudFactory)
}
```

3. Add the following `imports` to your `DataSource.groovy` file to allow `spring-cloud` API commands:

```
import org.springframework.cloud.CloudFactory
import org.springframework.cloud.CloudException
```

4. Add the following code to your `DataSource.groovy` file to enable Cloud Foundry's `getCloud` method to function locally or in other environments outside of a cloud.

```
def cloud = null
try {
 cloud = new CloudFactory().cloud
} catch (CloudException) {}
```

5. Use code like the following to access the cloud object:

```
def dbInfo = cloud?.getServiceInfo('myapp-mysql')
url = dbInfo?.jdbcUrl
username = dbInfo?.userName
password = dbInfo?.password
```

`myapp-mysql` is the name of the service as it appears in the `name` column of the output from `cf services`. For example, `mysql` or `rabbitmq`.

The example `DataSource.groovy` file below contains the following:

- The `imports` that allow `spring-cloud` API commands
- The code that enables the `getCloud` method to function locally or in other environments outside of a cloud
- Code to access the cloud object for SQL, MongoDB, and Redis services

```

import org.springframework.cloud.CloudFactory
import org.springframework.cloud.CloudException

def cloud = null
try {
 cloud = new CloudFactory().cloud
} catch (CloudException) {}

dataSource {
 pooled = true
 dbCreate = 'update'
 driverClassName = 'com.mysql.jdbc.Driver'
}

environments {
 production {
 dataSource {
 def dbInfo = cloud?.getServiceInfo('myapp-mysql')
 url = dbInfo?.jdbcUrl
 username = dbInfo?.userName
 password = dbInfo?.password
 }
 grails {
 mongo {
 def mongoInfo = cloud?.getServiceInfo('myapp-mongodb')
 host = mongoInfo?.host
 port = mongoInfo?.port
 databaseName = mongoInfo?.database
 username = mongoInfo?.userName
 password = mongoInfo?.password
 }
 redis {
 def redisInfo = cloud?.getServiceInfo('myapp-redis')
 host = redisInfo?.host
 port = redisInfo?.port
 password = redisInfo?.password
 }
 }
 }
 development {
 dataSource {
 url = 'jdbc:mysql://localhost:5432/myapp'
 username = 'sa'
 password = ''
 }
 grails {
 mongo {
 host = 'localhost'
 port = 27107
 databaseName = 'foo'
 username = 'user'
 password = 'password'
 }
 redis {
 host = 'localhost'
 port = 6379
 password = 'password'
 }
 }
 }
}
}

```

## Configuring Service Connections for Play Framework

Cloud Foundry supports running Play Framework applications and the Play JPA plugin for auto-configuration for Play versions up to and including v2.4.x.

Cloud Foundry provides support for connecting a Play Framework application to services such as MySQL and Postgres. In many cases, a Play Framework application running on Cloud Foundry can automatically detect and configure connections to services.

### Auto-Configuration

By default, Cloud Foundry detects service connections in a Play Framework application and configures them to use the credentials provided in the Cloud Foundry environment. Note that auto-configuration happens only if there is a single service of either of the supported types—MySQL or Postgres.

## Configuring Service Connections for Spring

Cloud Foundry provides extensive support for connecting a Spring application to services such as MySQL, PostgreSQL, MongoDB, Redis, and RabbitMQ. In many cases, Cloud Foundry can automatically configure a Spring application without any code changes. For more advanced cases, you can control service connection parameters yourself.

### Auto-Reconfiguration

If your Spring application requires services such as a relational database or messaging system, you might be able to deploy your application to Cloud Foundry without changing any code. In this case, Cloud Foundry automatically re-configures the relevant bean definitions to bind them to cloud services.

For information about connecting to services from a Spring application, see [Spring Cloud Spring Service Connector ↗](#).

Cloud Foundry auto-reconfigures applications only if the following items are true for your application:

- Only one service instance of a given service type is bound to the application. In this context, different relational databases services are considered the same service type. For example, if both a MySQL and a PostgreSQL service are bound to the application, auto-reconfiguration does not occur.
- Only one bean of a matching type is in the Spring application context. For example, you can have only one bean of type `javax.sql.DataSource`.

With auto-reconfiguration, Cloud Foundry creates the `DataSource` or connection factory bean itself, using its own values for properties such as host, port, and username. For example, if you have a single `javax.sql.DataSource` bean in your application context that Cloud Foundry auto-reconfigures and binds to its own database service, Cloud Foundry does not use the username, password, and driver URL you originally specified. Instead, Cloud Foundry uses its own internal values. This is transparent to the app, which really only cares about having a `DataSource` where it can write data but does not really care what the specific properties are that created the database. Also, if you have customized the configuration of a service, such as the pool size or connection properties, Cloud Foundry auto-reconfiguration ignores the customizations.

For more information about auto-reconfiguration of specific services types, see the [Service-Specific Details](#) section.

### Manual Configuration

Use manual configuration if you have multiple services of a given type or you want to have more control over the configuration than auto-reconfiguration provides.

To use manual configuration, include the `spring-cloud` library in your list of application dependencies. Update your application Maven `pom.xml` or Gradle `build.gradle` file to include dependencies on the `org.springframework.cloud:spring-cloud-spring-service-connector` and `org.springframework.cloud:spring-cloud-cloudfoundry-connector` artifacts.

For example, if you use Maven to build your application, the following `pom.xml` snippet shows how to add this dependency.

```
<dependencies>
 <dependency>
 <groupId>org.springframework.cloud</groupId>
 <artifactId>spring-cloud-spring-service-connector</artifactId>
 <version>1.2.3.RELEASE</version>
 </dependency>
 <dependency>
 <groupId>org.springframework.cloud</groupId>
 <artifactId>spring-cloud-cloudfoundry-connector</artifactId>
 <version>1.2.3.RELEASE</version>
 </dependency>
</dependencies>
```

You also need to update your application build file to include the Spring Framework Milestone repository. The following `pom.xml` snippet shows how to do this for Maven:

```
<repositories>
 <repository>
 <id>repository.springsource.milestone</id>
 <name>SpringSource Milestone Repository</name>
 <url>http://repo.springsource.org/milestone</url>
 </repository>
 ...
</repositories>
```

## Java Configuration

Typical use of Java config involves extending the `AbstractCloudConfig` class and adding methods with the `@Bean` annotation to create beans for services. Apps migrating from [auto-reconfiguration](#) might first try [Scanning for Services](#) until they need more explicit control. Java config also offers a way to expose application and service properties. Use this for debugging or to create service connectors using a lower-level access.

### Create a Service Bean

In the following example, the configuration creates a `DataSource` bean connecting to the only relational database service bound to the app. It also creates a `MongoDbFactory` bean, again, connecting to the only MongoDB service bound to the app. Check Javadoc for `AbstractCloudConfig` for ways to connect to other services.

```
class CloudConfig extends AbstractCloudConfig {
 @Bean
 public DataSource inventoryDataSource() {
 return connectionFactory().dataSource();
 }
 ... more beans to obtain service connectors
}
```

The bean names match the method names unless you specify an explicit value to the annotation such as `@Bean("inventory-service")`, following Spring's Java configuration standards.

If you have more than one service of a type bound to the app or want to have an explicit control over the services to which a bean is bound, you can pass the service names to methods such as `dataSource()` and `mongoDbFactory()` as follows:

```
class CloudConfig extends AbstractCloudConfig {

 @Bean
 public DataSource inventoryDataSource() {
 return connectionFactory().dataSource("inventory-db-service");
 }

 @Bean
 public MongoDbFactory documentMongoDbFactory() {
 return connectionFactory().mongoDbFactory("document-service");
 }

 ... more beans to obtain service connectors
}
```

Method such as `dataSource()` come in an additional overloaded variant that offer specifying configuration options such as the pooling parameters. See Javadoc for more details.

### Connect to Generic Services

Java config supports access to generic services through the `service()` method. Generic services do not have a directly mapped method. This is typical for a newly introduced service or when connecting to a private service in private PaaS. The generic `service()` method follows the same pattern as the `dataSource()`, except it allows supplying the connector type as an additional parameters.

### Scan for Services

You can scan for each bound service using the `@ServiceScan` annotation as shown below. This is conceptually similar to the `@ComponentScan` annotation in Spring:

```
@Configuration
@ServiceScan
class CloudConfig {
```

Here, one bean of the appropriate type (`DataSource` for a relational database service, for example) is created. Each created bean will have the `id` matching the corresponding service name. You can then inject such beans using auto-wiring:

```
@Autowired DataSource inventoryDb;
```

If the app is bound to more than one services of a type, you can use the `@Qualifier` annotation supplying it the name of the service as in the following code:

```
@Autowired @Qualifier("inventory-db") DataSource inventoryDb;
@Autowired @Qualifier("shipping-db") DataSource shippingDb;
```

## Access Service Properties

You can expose raw properties for all services and the app through a bean as follows:

```
class CloudPropertiesConfig extends AbstractCloudConfig {

 @Bean
 public Properties cloudProperties() {
 return properties();
 }
}
```

## Cloud Profile

Spring Framework versions 3.1 and above support bean definition profiles as a way to conditionalize the application configuration so that only specific bean definitions are activated when a certain condition is true. Setting up such profiles makes your application portable to many different environments so that you do not have to manually change the configuration when you deploy it to, for example, your local environment and then to Cloud Foundry.

See the Spring Framework documentation for additional information about using Spring bean definition profiles.

When you deploy a Spring application to Cloud Foundry, Cloud Foundry automatically enables the `cloud` profile.

 **Note:** Cloud Foundry auto-reconfiguration requires the Spring application to be version 3.1 or later and include the Spring context JAR. If you are using an earlier version, update your framework or use a custom buildpack.

## Profiles in Java Configuration

The `@Profile` annotation can be placed on `@Configuration` classes in a Spring application to set conditions under which configuration classes are invoked. By using the `default` and `cloud` profiles to determine whether the application is running on Cloud Foundry or not, your Java configuration can support both local and cloud deployments using Java configuration classes like these:

```

public class Configuration {
 @Configuration
 @Profile("cloud")
 static class CloudConfiguration {
 @Bean
 public DataSource dataSource() {
 CloudFactory cloudFactory = new CloudFactory();
 Cloud cloud = cloudFactory.getCloud();
 return cloud.getSingletonServiceConnector(DataSource.class, null);
 }
 }

 @Configuration
 @Profile("default")
 static class LocalConfiguration {
 @Bean
 public DataSource dataSource() {
 BasicDataSource dataSource = new BasicDataSource();
 dataSource.setUrl("jdbc:postgresql://localhost/db");
 dataSource.setDriverClassName("org.postgresql.Driver");
 dataSource.setUsername("postgres");
 dataSource.setPassword("postgres");
 return dataSource;
 }
 }
}

```

## Property Placeholders

Cloud Foundry exposes a number of application and service properties directly into its deployed applications. The properties exposed by Cloud Foundry include basic information about the application, such as its name and the cloud provider, and detailed connection information for all services currently bound to the application.

Service properties generally take one of the following forms:

```

cloud.services.{service-name}.connection.{property}
cloud.services.{service-name}.{property}

```

In this form, `{service-name}` refers to the name you gave the service when you bound it to your application at deploy time, and `{property}` is a field in the credentials section of the `VCAP_SERVICES` environment variable.

For example, assume that you created a Postgres service called `my-postgres` and then bound it to your application. Assume also that this service exposes credentials in `VCAP_SERVICES` as discrete fields. Cloud Foundry exposes the following properties about this service:

```

cloud.services.my-postgres.connection.hostname
cloud.services.my-postgres.connection.name
cloud.services.my-postgres.connection.password
cloud.services.my-postgres.connection.port
cloud.services.my-postgres.connection.username
cloud.services.my-postgres.plan
cloud.services.my-postgres.type

```

If the service exposed the credentials as a single `uri` field, then the following properties would be set up:

```

cloud.services.my-postgres.connection.uri
cloud.services.my-postgres.plan
cloud.services.my-postgres.type

```

For convenience, if you have bound just one service of a given type to your application, Cloud Foundry creates an alias based on the service type instead of the service name. For example, if only one MySQL service is bound to an application, the properties takes the form `cloud.services.mysql.connection.{property}`. Cloud Foundry uses the following aliases in this case:

- `mysql`
- `postgresql`
- `mongodb`
- `redis`

- rabbitmq

A Spring application can take advantage of these Cloud Foundry properties using the property placeholder mechanism. For example, assume that you have bound a MySQL service called `spring-mysql` to your application. Your application requires a `c3p0` connection pool instead of the connection pool provided by Cloud Foundry, but you want to use the same connection properties defined by Cloud Foundry for the MySQL service - in particular the username, password and JDBC URL.

The following table lists all the application properties that Cloud Foundry exposes to deployed applications.

Property	Description
<code>cloud.application.name</code>	The name provided when the application was pushed to Cloud Foundry.
<code>cloud.provider.url</code>	The URL of the cloud hosting the application, such as <code>cloudfoundry.com</code> .

The service properties that are exposed for each type of service are listed in the [Service-Specific Details](#) section.

## Service-Specific Details

The following sections describe Spring auto-reconfiguration and manual configuration for the services supported by Cloud Foundry.

### MySQL and Postgres

#### Auto-Reconfiguration

Auto-reconfiguration occurs if Cloud Foundry detects a `javax.sql.DataSource` bean in the Spring application context. The following snippet of a Spring application context file shows an example of defining this type of bean which Cloud Foundry will detect and potentially auto-reconfigure:

```
<bean class="org.apache.commons.dbcp.BasicDataSource" destroy-method="close" id="dataSource">
<property name="driverClassName" value="org.h2.Driver" />
<property name="url" value="jdbc:h2:mem:" />
<property name="username" value="sa" />
<property name="password" value="" />
</bean>
```

The relational database that Cloud Foundry actually uses depends on the service instance you explicitly bind to your application when you deploy it: MySQL or Postgres. Cloud Foundry creates either a commons DBCP or Tomcat datasource depending on which datasource implementation it finds on the classpath.

Cloud Foundry internally generates values for the following properties: `driverClassName`, `url`, `username`, `password`, `validationQuery`.

#### Manual Configuration in Java

To configure a database service in Java configuration, create a `@Configuration` class with a `@Bean` method to return a `javax.sql.DataSource` bean. The bean can be created by helper classes in the `spring-cloud` library, as shown here:

```
@Configuration
public class DataSourceConfig {
 @Bean
 public DataSource dataSource() {
 CloudFactory cloudFactory = new CloudFactory();
 Cloud cloud = cloudFactory.getCloud();
 return cloud.getSingletonServiceConnector(DataSource.class, null);
 }
}
```

## MongoDB

#### Auto-Reconfiguration

You must use Spring Data MongoDB 1.0 M4 or later for auto-reconfiguration to work.

Auto-reconfiguration occurs if Cloud Foundry detects an `org.springframework.data.document.mongodb.MongoDbFactory` bean in the Spring application context. The following snippet of a Spring XML application context file shows an example of defining this type of bean which Cloud Foundry will detect and potentially auto-reconfigure:

```
<mongo:db-factory
 id="mongoDbFactory"
 dbname="pwdtest"
 host="127.0.0.1"
 port="1234"
 username="test_user"
 password="test_pass" />
```

Cloud Foundry creates a `SimpleMongoDbFactory` with its own values for the following properties: `host`, `port`, `username`, `password`, `dbname`.

## Manual Configuration in Java

To configure a MongoDB service in Java configuration, create a `@Configuration` class with a `@Bean` method to return an `org.springframework.data.mongodb.MongoDbFactory` bean from Spring Data MongoDB. The bean can be created by helper classes in the `spring-cloud` library, as shown here:

```
@Configuration
public class MongoConfig {

 @Bean
 public MongoDbFactory mongoDbFactory() {
 CloudFactory cloudFactory = new CloudFactory();
 Cloud cloud = cloudFactory.getCloud();
 return cloud.getSingletonServiceConnector(MongoDbFactory.class, null);
 }

 @Bean
 public MongoTemplate mongoTemplate() {
 return new MongoTemplate(mongoDbFactory());
 }
}
```

## Redis

### Auto-Configuration

You must be using Spring Data Redis 1.0 M4 or later for auto-configuration to work.

Auto-configuration occurs if Cloud Foundry detects a `org.springframework.data.redis.connection.RedisConnectionFactory` bean in the Spring application context. The following snippet of a Spring XML application context file shows an example of defining this type of bean which Cloud Foundry will detect and potentially auto-configure:

```
<bean id="redis"
 class="org.springframework.data.redis.connection.jedis.JedisConnectionFactory"
 p:hostName="localhost" p:port="6379" />
```

Cloud Foundry creates a `JedisConnectionFactory` with its own values for the following properties: `host`, `port`, `password`. This means that you must package the Jedis JAR in your application. Cloud Foundry does not currently support the JRedis and RJC implementations.

## Manual Configuration in Java

To configure a Redis service in Java configuration, create a `@Configuration` class with a `@Bean` method to return an `org.springframework.data.redis.connection.RedisConnectionFactory` bean from Spring Data Redis. The bean can be created by helper classes in the `spring-cloud` library, as shown here:

```

@Configuration
public class RedisConfig {

 @Bean
 public RedisConnectionFactory redisConnectionFactory() {
 CloudFactory cloudFactory = new CloudFactory();
 Cloud cloud = cloudFactory.getCloud();
 return cloud.getSingletonServiceConnector(RedisConnectionFactory.class, null);
 }

 @Bean
 public RedisTemplate redisTemplate() {
 return new StringRedisTemplate(redisConnectionFactory());
 }

}

```

## RabbitMQ

### Auto-Configuration

You must be using Spring AMQP 1.0 or later for auto-configuration to work. Spring AMQP provides publishing, multi-threaded consumer generation, and message conversion. It also facilitates management of AMQP resources while promoting dependency injection and declarative configuration.

Auto-configuration occurs if Cloud Foundry detects an `org.springframework.amqp.rabbit.connection.ConnectionFactory` bean in the Spring application context. The following snippet of a Spring application context file shows an example of defining this type of bean which Cloud Foundry will detect and potentially auto-configure:

```

<rabbit:connection-factory
 id="rabbitConnectionFactory"
 host="localhost"
 password="testpwd"
 port="1238"
 username="testuser"
 virtual-host="virthost" />

```

Cloud Foundry creates an `org.springframework.amqp.rabbit.connection.CachingConnectionFactory` with its own values for the following properties: `host`, `virtual-host`, `port`, `username`, `password`.

### Manual Configuration in Java

To configure a RabbitMQ service in Java configuration, create a `@Configuration` class with a `@Bean` method to return an `org.springframework.amqp.rabbit.connection.ConnectionFactory` bean from the Spring AMQP library. The bean can be created by helper classes in the `spring-cloud` library, as shown here:

```

@Configuration
public class RabbitConfig {

 @Bean
 public ConnectionFactory rabbitConnectionFactory() {
 CloudFactory cloudFactory = new CloudFactory();
 Cloud cloud = cloudFactory.getCloud();
 return cloud.getSingletonServiceConnector(ConnectionFactory.class, null);
 }

 @Bean
 public RabbitTemplate rabbitTemplate(ConnectionFactory connectionFactory) {
 return new RabbitTemplate(connectionFactory);
 }
}

```

## Cloud Foundry Java Client Library

### Introduction

This is a guide to using the [Cloud Foundry Java Client Library](#) to manage an account on a Cloud Foundry instance.

 **Note:** The 1.1.x versions of the Cloud Foundry Java Client Library work with apps using Spring 4.x, and the 1.0.x versions of the Cloud Foundry Java Client Library work with apps using Spring 3.x. Both versions are available in the [source repository](#) on GitHub.

### Adding the Library

Visit the [Cloud Foundry Java Client Library](#) GitHub page to obtain the correct components.

Most projects need two dependencies: the Operations API and an implementation of the Client API. Refer to the following sections for more information about how to add the Cloud Foundry Java Client Library as dependencies to a Maven or Gradle project.

### Maven

Add the `cloudfoundry-client-reactor` dependency (formerly known as `cloudfoundry-client-spring`) to your `pom.xml` as follows:

```
<dependencies>
 <dependency>
 <groupId>org.cloudfoundry</groupId>
 <artifactId>cloudfoundry-client-reactor</artifactId>
 <version>2.0.0.BUILD-SNAPSHOT</version>
 </dependency>
 <dependency>
 <groupId>org.cloudfoundry</groupId>
 <artifactId>cloudfoundry-operations</artifactId>
 <version>2.0.0.BUILD-SNAPSHOT</version>
 </dependency>
 <dependency>
 <groupId>io.projectreactor</groupId>
 <artifactId>reactor-core</artifactId>
 <version>2.5.0.BUILD-SNAPSHOT</version>
 </dependency>
 <dependency>
 <groupId>io.projectreactor</groupId>
 <artifactId>reactor-netty</artifactId>
 <version>2.5.0.BUILD-SNAPSHOT</version>
 </dependency>
 ...
</dependencies>
```

The artifacts can be found in the Spring release and snapshot repositories:

```
<repositories>
 <repository>
 <id>spring-releases</id>
 <name>Spring Releases</name>
 <url>http://repo.spring.io/release</url>
 </repository>
 ...
</repositories>
```

```
<repositories>
 <repository>
 <id>spring-snapshots</id>
 <name>Spring Snapshots</name>
 <url>http://repo.spring.io/snapshot</url>
 <snapshots>
 <enabled>true</enabled>
 </snapshots>
 </repository>
 ...
</repositories>
```

## Gradle

Add the `cloudfoundry-client-reactor` dependency to your `build.gradle` file as follows:

```
dependencies {
 compile 'org.cloudfoundry:cloudfoundry-client-reactor:2.0.0.BUILD-SNAPSHOT'
 compile 'org.cloudfoundry:cloudfoundry-operations:2.0.0.BUILD-SNAPSHOT'
 compile 'io.projectreactor:reactor-core:2.5.0.BUILD-SNAPSHOT'
 compile 'io.projectreactor:reactor-netty:2.5.0.BUILD-SNAPSHOT'
 ...
}
```

The artifacts can be found in the Spring release and snapshot repositories:

```
repositories {
 maven { url 'http://repo.spring.io/release' }
 ...
}
```

```
repositories {
 maven { url 'http://repo.spring.io/snapshot' }
 ...
}
```

## Sample Code

The following is a very simple sample application that connects to a Cloud Foundry instance, logs in, and displays some information about the Cloud Foundry account. When running the program, provide the Cloud Foundry target API endpoint, along with a valid user name and password as command-line parameters.

```

import org.cloudfoundry.client.lib.CloudCredentials;
import org.cloudfoundry.client.lib.CloudFoundryClient;
import org.cloudfoundry.client.lib.domain.CloudApplication;
import org.cloudfoundry.client.lib.domain.CloudService;
import org.cloudfoundry.client.lib.domain.CloudSpace;

import java.net.MalformedURLException;
import java.net.URI;
import java.net.URL;

public final class JavaSample {

 public static void main(String[] args) {
 String target = args[0];
 String user = args[1];
 String password = args[2];

 CloudCredentials credentials = new CloudCredentials(user, password);
 CloudFoundryClient client = new CloudFoundryClient(credentials, getTargetURL(target));
 client.login();

 System.out.printf("%nSpaces:%n");
 for (CloudSpace space : client.getSpaces()) {
 System.out.printf(" %s%n", space.getName(), space.getOrganization().getName());
 }

 System.out.printf("%nApplications:%n");
 for (CloudApplication application : client.getApplications()) {
 System.out.printf(" %s%n", application.getName());
 }

 System.out.printf("%nServices:%n");
 for (CloudService service : client.getServices()) {
 System.out.printf(" %s%n", service.getName(), service.getLabel());
 }
 }

 private static URL getTargetURL(String target) {
 try {
 return URI.create(target).toURL();
 } catch (MalformedURLException e) {
 throw new RuntimeException("The target URL is not valid: " + e.getMessage());
 }
 }
}

```

For more details about the Cloud Foundry Java Client Library, visit the [source repository](#) in GitHub. The [domain package](#) shows the objects that you can query and inspect.

## .NET Core Buildpack

Page last updated:

This topic describes how to push Cloud Foundry apps to Linux cells using the .NET Core buildpack. You can find supported ASP.NET Core versions in the [.NET Core buildpack release notes](#).

 **Note:** The .NET Core buildpack can only be used to deploy apps to Linux cells. To deploy .NET Core apps to Windows cells, use the [Binary buildpack](#).

## Overview

Buildpacks provide needed dependencies to Pivotal Application Service for Windows (PASW) apps. PASW automatically uses the .NET Core buildpack when one or more of the following conditions are met:

- The pushed app contains one or more `*.csproj` or `*.fsproj` files.
- The app is pushed from the output directory of the `dotnet publish` command.

For information about deploying different types of .NET apps, follow the links in the table below.

Type of .NET App	Buildpack
ASP.NET MVC	
ASP.NET Web Forms	
ASP.NET WebAPI Apps	
Windows Communication Foundation (WCF)	<a href="#">HWC</a>
.NET Console	<a href="#">Binary</a>
.NET Core pushed to Linux stack	<a href="#">.NET Core</a>
.NET Core pushed to Windows stack	<a href="#">Binary</a>

## Push a .NET Core App

To push a .NET Core app to a Linux cell, do the following:

- Run the `cf push` command:

```
cf push APP-NAME
```

Where `APP-NAME` is the name you want to give your app. For example:

```
$ cf push my-app
Pushing app my-app to org my-org / space my-space as username@example.com...
...
name: my-app
requested state: started
routes: my-app.example.com
last uploaded: Mon 13 May 15:48:00 PDT 2020
stack: cflinuxfs3
buildpacks: dotnet-core
```

If your PASW deployment does not have the .NET Core buildpack installed or the installed version is out of date, run the push command with the `-b` option to specify the buildpack:

```
cf push APP-NAME -b https://github.com/cloudfoundry/dotnet-core-buildpack.git
```

Where `APP-NAME` is the name you want to give your app.

- Find the URL of your app in the output of the `push` command. In the example above, `my-app.example.com` is the URL of the app.
- Open a browser and navigate to the URL to see your app running.

For a basic sample app, see [ASP.NET Core Getting Started App](#) in GitHub.

## Source-Based, Non-Published Deployments

For a source-based, non-published deployment, you push your app's source code, not the output directory of the `dotnet publish` command.

The source-based, non-published workflow ensures the buildpack can keep all your dependencies in sync and up to date. For additional information about using source-based, non-published deployments, see the following sections:

- [Deploy Apps with Multiple Projects](#)
- [Use Non-Default Package Sources](#)
- [Disable and Clear Your NuGet Package Cache](#)

 **Note:** The source-based deployment workflow also uses the `cf push` command to push source-based apps to Cloud Foundry.

### Deploy Apps with Multiple Projects

If you are deploying an app containing multiple projects, you must specify which of the app's projects is the main project.

To specify the main project in a multi-project deployment perform the following steps:

1. Create a `.deployment` file in your app's root folder and open the new file in a text editor.
2. Designate the main project's path by configuring the file, using the following format:

```
[config]
project = PATH-TO-YOUR-MAIN-PROJECT
```

Where `PATH-TO-YOUR-MAIN-PROJECT` is the location of your main project's `*.csproj` or `*.fsproj` file.

For example:

```
[config]
project = src/MyApp.Web/MyApp.Web.csproj
```

In this example, by pointing to the `MyApp.Web` `*.csproj` file, `MyApp.Web` is configured as the main project.

3. Save the revised `.deployment` file.

When deployed, the buildpack attempts to execute the main project using the `dotnet run` command, `dotnet run -p PATH-TO-YOUR-MAIN-PROJECT`, and automatically compiles all projects listed as dependencies in the main project's `*.csproj` or `*.fsproj` file.

For example: Suppose your `.deployment` file is configured as above and you have an app `src` folder containing three projects: `MyApp.Web`, `MyApp.DAL` and `MyApp.Services`. If your `MyApp.Web.csproj` file lists the `MyApp.DAL` and `MyApp.Services` projects as dependencies, the two additional projects will be compiled by the buildpack.

### Use Non-Default Package Sources

If you want to deploy an app that uses non-default package sources, you must specify those package sources in the `NuGet.Config` file. For information about `NuGet.Config`, see [nuget.config reference](#) in the Microsoft documentation.

### Disable and Clear Your NuGet Package Cache

You may need to disable NuGet package caching or clear NuGet packages cached in the staging environment in one of the following scenarios:

- Your app fails to stage because it runs out of space, exceeding the maximum allowable disk quota.
- You have added pre-release packages to test a new feature and then decided to revert back to the main NuGet feed. You may need to remove the

packages you changed from the cache to avoid conflicts.

Disabling NuGet caching clears any existing NuGet dependencies from the staging cache and prevents the buildpack from adding NuGet dependencies to the staging cache.

NuGet package caching is disabled by default. If the default is not explicitly overridden, no additional NuGet caching configuration is required.

To disable NuGet package caching, do the following:

1. Confirm the `CACHE_NUGET_PACKAGES` environment variable is not set to `true` in your app manifest by locating the `manifest.yml` file and confirming `CACHE_NUGET_PACKAGES` is not set to true.
2. If needed, set `CACHE_NUGET_PACKAGES` to `false` in the `manifest.yml` file by setting the `CACHE_NUGET_PACKAGES` environment variable to `false`.

For example:

```
--
applications:
- name: sample-aspnetcore-app
 memory: 512M
 env:
 CACHE_NUGET_PACKAGES: false
```

3. To alternatively configure the setting to false in the environment variables settings, run the following on the command line:

```
cf set-env APP-NAME CACHE_NUGET_PACKAGES false
```

Where `APP-NAME` is the name of your app.

For more information, see the [Environment Variables](#) section of the [Deploying with Application Manifests](#) topic.

## Framework-Dependent Deployments

For a framework-dependent deployment (FDD), you deploy only your app and third-party dependencies. Cloud Foundry recommends using this workflow if you deploy an app in an offline setting. For information about deploying FDDs, see [Framework-dependent deployments \(FDD\)](#) in the Microsoft documentation.

To deploy an FDD using the buildpack, do the following:

1. Publish the app by running the `dotnet publish` command:

```
dotnet publish [-f FRAMEWORK-NAME] [-c Release]
```

Where `FRAMEWORK-NAME` is your target framework.

2. Prepare to push your app by doing one of the following:

- o If your app uses a `manifest.yml`, specify a path to the output folder of `dotnet publish`. This allows you to push your app from any directory.
- o If not, navigate to the `bin/Debug|Release/YOUR-FRAMEWORK/YOUR-RUNTIME/publish` directory.

3. Push your app.

## Self-Contained Deployments

For a self-contained deployment (SCD), you deploy your app, third-party dependencies, and the version of .NET Core that you used to build your app. For information about SCDs, see [Self-contained deployments \(SCD\)](#) in the Microsoft documentation.

 **Note:** Cloud Foundry does not recommend using the SCD workflow. The buildpack is unable to keep dependencies in sync and up to date for workflows that deploy a pre-published binary.

When using the SCD workflow for deploying your app, you must do the following:

- Specify a runtime in the `dotnet publish` command. For example:

```
$ dotnet publish -r ubuntu14.04-x64
```

- Include the specified runtime in the `Runtimeldentifiers` section of the project file.

## Specify .NET Core SDKs

To pin the .NET Core SDK to a specific version or version line, create a `buildpack.yml` file at the app root and add your SDK version in one of the following formats:

```
dotnet-core:
sdk: 2.1.201
```

```
dotnet-core:
sdk: 2.1.x
```

```
dotnet-core:
sdk: 2.x
```

The buildpack chooses what SDK to install based on the files present at the app root in the following order of precedence:

- `buildpack.yml`
- `global.json`
- `*.fsproj`

 **Note:** The app respects the SDK version specified in `global.json` at runtime. If you provide versions in both `global.json` and `buildpack.yml` files, ensure you specify the same versions in both files.

## Specify .NET Runtime Versions

This section explains how to specify a .NET Runtime version for source-based and framework-dependent apps.

### Source-Based Apps

If you want to lock the .NET Runtime version, configure your `.csproj` or `.fsproj` file to lock your app to the desired version.

For example, the following configuration locks the runtime to 2.1:

```
<PropertyGroup>
 <TargetFramework>netcoreapp2.1</TargetFramework>
 <RuntimeFrameworkVersion>2.1.*</RuntimeFrameworkVersion>
</PropertyGroup>
```

 **Note:** For source-based apps, specify a minor version of the .NET Runtime. Do not specify a patch version because buildpacks contain only the two most recent patch versions of each minor version.

### Framework-Dependent Apps

Your app is configured to use the latest .NET Runtime patch version by default.

If you want your app to maintain a specific .NET Runtime version, you must modify your app's `.runtimeconfig.json` file to include the `applyPatches` property and set the property to `false`.

For example:

```
{
 "runtimeOptions": {
 "tfm": "netcoreapp2.0",
 "framework": {
 "name": "Microsoft.NETCore.App",
 "version": "2.0.0"
 },
 "applyPatches": false
 }
}
```

 **Note:** Set `applyPatches: false` in `*.runtimeconfig.json` only if you want to pin your .NET Framework to a specific version. This prevents your app from receiving updates to the runtime version and assemblies.

## Push an App in a Disconnected Environment

For offline environments, Cloud Foundry recommends using the Framework-Dependent Deployment workflow. This workflow enables the deployed app to use the latest runtime provided by the offline buildpack. For more information, see: [Framework-Dependent Deployments](#).

## Maintain ASP.NET Core Assemblies

 **Note:** This section applies only to source-based and framework-dependent deployments.

For maintaining ASP.NET Core assemblies, it is recommended you do one of the following

- Configure your app as a fully vendored app requiring fewer buildpack updates. Modify your `.csproj` file with the following:

```
<PropertyGroup>
 <PublishWithAspNetCoreTargetManifest>false</PublishWithAspNetCoreTargetManifest>
</PropertyGroup>
```

- Keep your SDK up to date, by setting `buildpack.yml` to the .NET SDK line you want to use. For example:

```
--
dotnet-core:
 sdk: 2.0.x
```

 **Note:** `2.0.x` ASP.NET Core assemblies are released in the `2.1.200 - 2.1.299` SDK versions, and `2.1.x` assemblies are released in the `2.1.300` and above SDK versions.

## Configure the Listen Port

PASW sets the `$PORT` environment variable automatically. For your .NET Core app to work on PASW, you must configure the app to listen on the environment's specified port.

For C# apps, the following modifications allow the buildpack to pass the correct port from `$PORT` environment variable to the app when running the initial startup command:

- Open the file that contains your `Main` method.
- Add a `using` statement to the top of the file:

```
using Microsoft.Extensions.Configuration;
```

- Add the following lines before the line `var host = new WebHostBuilder()`:

```
var config = new ConfigurationBuilder()
 .AddCommandLine(args)
 .Build();
```

4. Add the following line after `.UseKestrel()`:

```
.UseConfiguration(config)
```

The `Main` method should now resemble the following example:

```
public static void Main(string[] args)
{
 var config = new ConfigurationBuilder()
 .AddCommandLine(args)
 .Build();
 var host = new WebHostBuilder()
 .UseKestrel()
 .UseConfiguration(config)
 .UseContentRoot(Directory.GetCurrentDirectory())
 .UseStartup<Startup>()
 .Build();
 host.Run();
}
```

5. Save your changes.

6. Add `Microsoft.Extensions.Configuration.CommandLine` as a dependency in `*.csproj`:

```
<PackageReference Include="Microsoft.Extensions.Configuration.CommandLine">
 <Version>VERSION</Version>
</PackageReference>
```

Where `VERSION` is the version of the package to use. To find a list of valid versions, navigate to <https://www.nuget.org>.

7. If your app requires any other files at runtime, such as JSON configuration files, add them to the `include` section of `copyToOutput`.

8. Save your changes.

With these changes, the `dotnet run` command copies your app `Views` to the build output where the .NET CLI can find them.

## Add Custom Libraries

If your app requires external shared libraries that are not provided by the rootfs or the buildpack, you must place the libraries in an `ld_library_path` directory at the app root.

 **Note:** You must keep these libraries up to date. They do not update automatically.

The .NET Core buildpack automatically adds the directory `<app-root>/ld_library_path` to `LD_LIBRARY_PATH` so that your app can access these libraries at runtime.

## NGINX Buildpack

This topic describes how to push your NGINX app to Cloud Foundry and how to configure your NGINX app to use the NGINX buildpack.

### Push an App

If your app contains an `nginx.conf` file, Cloud Foundry automatically uses the NGINX buildpack when you run `cf push` to deploy your app.

If your Cloud Foundry deployment does not have the NGINX buildpack installed or the installed version is outdated, run

`cf push YOUR-APP -b https://github.com/cloudfoundry/nginx-buildpack.git` to deploy your app with the current buildpack. Replace `YOUR-APP` with the name of your app.

For example:

```
$ cf push my-app -b https://github.com/cloudfoundry/nginx-buildpack.git
```

### Configure NGINX

We recommend that you use the default NGINX directory structure for your NGINX webserver. You can view this directory structure in the [nginx-buildpack](#) repository in GitHub.

The NGINX webserver setup includes the following:

- A root folder for all static web content
- A MIME type configuration file
- An NGINX configuration file
- A `buildpack.yml` YAML file that defines the version of NGINX to use. As an example, see [buildpack.yml](#) in the Cloud Foundry NGINX Buildpack repository in GitHub.

You should make any custom configuration changes based on these default files to ensure compatibility with the buildpack.

### Create the `nginx.conf` File

Use the templating syntax when you create an `nginx.conf` file. This templating syntax loads modules and binds to ports based on values known at launch time.

#### Port

Use  `{{port}}` to set the port to listen on. At launch time,  `{{port}}` will interpolate in the value of `$PORT`.

 **Note:** You must use  `{{port}}` in your `nginx.conf` file.

For example, to set an NGINX server to listen on `$PORT`, include the following in your `nginx.conf` file:

```
server {
 listen {{port}};
}
```

### Environment Variables

To use an environment variable, include  `{{env "YOUR-VARIABLE"}}` . Replace `YOUR-VARIABLE` with the name of an environment variable. At staging and at launch, the current value of the environment variable is retrieved.

For example, include the following in your `nginx.conf` file to enable or disable GZipping based on the value of `GZIP_DOWNLOADS`:

```
gzip {{env "GZIP_DOWNLOADS"}};
```

- If you set `GZIP_DOWNLOADS` to `off`, NGINX does not GZip files.
- If you set `GZIP_DOWNLOADS` to `on`, NGINX GZips files.

## Loading Dynamic Modules

To load an NGINX module, use the following syntax in your app's `nginx.conf` file:

```
{{module "MODULE-NAME"}}
```

If you have provided a module in a `modules` directory located at the root of your app, the buildpack instructs NGINX to load that module. If you have not provided a module, the buildpack instructs NGINX to search for a matching built-in module.

As of v0.0.5 of the buildpack, the `ngx_stream_module` is available as a built-in module.

For example, to load a custom module named `ngx_hello_module`, provide a `modules/ngx_hello_module.so` file in your app directory and add the following to the top of your `nginx.conf` file:

```
{{module "ngx_hello_module"}}
```

To load a built-in module like `ngx_stream_module`, add the following to the top of your `nginx.conf` file. You do not need to provide an `ngx_stream_module.so` file:

```
{{module "ngx_stream_module"}}
```

 **Note:** To name your modules directory something other than `modules`, use the NGINX `load_module` directive, providing a path to the module relative to the location of your `nginx.conf` file. For example:  
`load_module  
some_module_dir/my_module.so`

## Buildpack Support

The following resources can assist you when using the NGINX buildpack or when developing your own NGINX buildpack:

- **NGINX Buildpack Repository in GitHub:** Find more information about using and extending the NGINX buildpack in the [NGINX buildpack](#) GitHub repository.
- **Release Notes:** Find current information about this buildpack on the [NGINX buildpack release page](#) in GitHub.
- **Slack:** Join the #buildpacks channel in the [Cloud Foundry Slack community](#).

## Node.js Buildpack

Use the Node.js buildpack with Node or JavaScript apps.

You must install the [Cloud Foundry Command Line Interface tool](#) (cf CLI) to run some of the commands listed in this topic.

### Push Node.js Apps

Cloud Foundry automatically uses the Node.js buildpack if it detects a `package.json` file in the root directory of your project.

The `-b` option lets you specify a buildpack to use with the `cf push` command. If your Cloud Foundry deployment does not have the Node.js buildpack installed or the installed version is out of date, run `cf push APP-NAME -b https://github.com/cloudfoundry/nodejs-buildpack`, where `APP-NAME` is the name of your app, to push your app with the latest version of the buildpack.

For example:

```
$ cf push my-nodejs-app -b https://github.com/cloudfoundry/nodejs-buildpack ↗
```

For more detailed information about deploying Node.js apps, see the following topics:

- [Tips for Node.js Developers](#)
- [Environment Variables Defined by the Node Buildpack](#)
- [Configuring Service Connections for Node](#)
- [Node.js Buildpack Source Code on GitHub ↗](#)

### Supported Versions

For a list of supported Node versions, see the Node.js Buildpack [release notes ↗](#) on GitHub.

### Specify a Node.js Version

To specify a Node.js version, set the `engines.node` in the `package.json` file to the semantic versioning specification (semver) range or the specific version of Node you are using.

Example showing a semver range:

```
"engines": {
 "node": "6.9.x"
}
```

Example showing a specific version:

```
"engines": {
 "node": "6.9.0"
}
```

If you try to use a version that is not currently supported, staging your app fails with the following error message:

```
Could not get translated url, exited with: DEPENDENCY_MISSING_IN_MANIFEST...
!
! exit
!
Staging failed: Buildpack compilation step failed
```

## Specify an npm Version

To specify an npm version, set `engines.npm` in the `package.json` file to the semantic versioning specification (semver) range or the specific version of npm you are using:

Example showing a semver range:

```
"engines": {
 "node": "6.9.x",
 "npm": "2.15.x"
}
```

Example showing a specific version:

```
"engines": {
 "node": "6.9.0",
 "npm": "2.15.1"
}
```

If you do not specify an npm version, your app uses the default npm packaged with the Node.js version used by your app, as specified on the [Node.js releases](#) page.

If your environment cannot connect to the Internet and you specified a non-supported version of npm, the buildpack fails to download npm and you see the following error message:

```
We're unable to download the version of npm you've provided (...).
Please remove the npm version specification in package.json (...)
Staging failed: Buildpack compilation step failed
```

## Using NPM or Yarn

By default, the Node.js buildpack assumes you are using NPM. If you want to use [Yarn](#) instead, you must provide a `yarn.lock` in your top-level app directory. For more information on the Yarn lock file, see [yarn.lock](#) in the Yarn documentation.

## Vendorizing App Dependencies

To vendor dependencies for an app using the Node.js buildpack, run `npm install` (or `yarn install`, if you are using Yarn) from your app directory. This command vendors dependencies into the `node_modules` directory of your app directory.

For example, the following example vendors dependencies into the `my-nodejs-app/node_modules` directory:

```
$ cd my-nodejs-app
$ npm install
```

The `cf push` command uploads the vendored dependencies with the app.

 **Note:** For an app to run in a disconnected environment, it must vendor its dependencies and provide a lock file.

## Usage in Offline Environments

When vendorizing apps for usage in offline environments, you must supply a lock file. For information on npm lockfiles, see [npm-package-locks](#) in the NPM documentation. For information on `npm-package-lock.json` files, see [package-lock.json](#) in the NPM documentation.

This lock file informs the package manager of the exact versions of dependencies and transitive dependencies to look for when running `npm install` or `yarn install`. For this reason, you must keep `package.json`, the vendored `node_modules` directory, and your lock file synchronized to avoid network calls.

 **Note:** `package-lock.json` is only supported by NPM 5.x and later. For earlier versions of NPM, provide a `npm-shrinkwrap.json` file instead.

## Yarn Offline Mode

Versions 1.5.28 and later of the Node.js buildpack include the ability to use Yarn in offline mode. To do so, you must mirror the Yarn registry locally by providing an `npm-packages-offline-cache` directory:

```
$ cd APP-DIR
$ yarn config set yarn-offline-mirror ./npm-packages-offline-cache
$ yarn config set yarn-offline-mirror-pruning true
$ cp ~/.yarnrc .
$ rm -rf node_modules/ yarn.lock # if they were previously generated
$ yarn install
```

When you push the app, the buildpack looks for an `npm-packages-offline-cache` directory at the top level of the app directory. If this directory exists, the buildpack runs Yarn in offline mode. Otherwise, it runs Yarn normally, which may require an Internet connection. You do not have to provide a `node_modules` directory when running Yarn in offline mode, as the offline cache will provide the dependencies.

For more information about using an offline mirror with Yarn, see the [Yarn Blog](#).

## Integrity Check

By default, the Node.js buildpack uses npm to download dependencies. Note that npm does not perform integrity checks of the downloaded packages, which is a [security risk](#).

If missing dependencies are detected, the buildpack runs `npm install` for non-vendored dependencies or `npm rebuild` for dependencies that are already vendored. This may result in code being downloaded and executed without verification.

You can use Yarn as an alternative that verifies dependencies.

## OpenSSL Support

The [nodejs-buildpack](#) packages binaries of Node.js with OpenSSL that are statically linked. The Node.js buildpack supports Node.js 4.x and later, which relies on the Node.js release cycle to provide OpenSSL updates. The [binary-builder](#) enables [static OpenSSL compilation](#).

## Proxy Support

If you need to use a proxy to download dependencies during staging, you can set the `http_proxy` and/or `https_proxy` environment variables. For more information, see [Using a Proxy](#).

## BOSH Configured Custom Trusted Certificate Support

Node.js hardcodes root CA certs in its source code. To use [BOSH configured custom trusted certificates](#), a developer must pass the specified CAs to the `tls.connect` function as extra arguments.

## Help and Support

Join the `#buildpacks` channel in our [Slack community](#) if you need any further assistance.

For more information about using and extending the Node.js buildpack in Cloud Foundry, see the [Node.js GitHub repository](#).

You can find current information about this buildpack in the Node.js buildpack [release page](#) in GitHub.

## Tips for Node.js Applications

This topic provides Node-specific information to supplement the general guidelines in the [Pushing an App](#) topic.

### About the Node.js Buildpack

For information about using and extending the Node.js buildpack in Cloud Foundry, see the [nodejs-buildpack repository](#) in GitHub.

You can find current information about this buildpack on the Node.js buildpack [release page](#) in GitHub.

The buildpack uses a default Node.js version. To specify the versions of Node.js and npm an app requires, edit the app's `package.json`, as described in "node.js and npm versions" in the [nodejs-buildpack repository](#).

### Application Package File

Cloud Foundry expects a `package.json` in your Node.js app. You can specify the version of Node.js you want to use in the `engines` node of your `package.json` file.

In general, Cloud Foundry supports the two most recent versions of Node.js. See the GitHub [Node.js buildpack page](#) for current information.

Example `package.json` file:

```
{
 "name": "first",
 "version": "0.0.1",
 "author": "Demo",
 "dependencies": {
 "express": "3.4.8",
 "consolidate": "0.10.0",
 "swig": "1.3.2"
 },
 "engines": {
 "node": "0.12.7",
 "npm": "2.7.4"
 }
}
```

### Application Port

You must use the `PORT` environment variable to determine which port your app should listen on. To also run your app locally, set the default port as `3000`.

```
app.listen(process.env.PORT || 3000);
```

### Low Memory Environments

When running node apps, you might notice that instances are occasionally restarted due to memory constraints. Node does not know how much memory it is allowed to use, and thus sometimes allows the garbage collector to wait past the allowed amount of memory. To resolve this issue, set the `OPTIMIZE_MEMORY` environment variable to `true` (requires node v6.12.0 or greater). This sets `max_old_space_size` based on the available memory in the instance.

```
$ cf set-env my-app OPTIMIZE_MEMORY true
```

### Application Start Command

Node.js apps require a start command. You can specify the web start command for a Node.js app in a Procfile or in the app deployment manifest. For more

information about Procfiles, see the [Configuring a Production Server](#) topic.

The first time you deploy, you are asked if you want to save your configuration. This saves a `manifest.yml` in your app with the settings you entered during the initial push. Edit the `manifest.yml` file and create a start command as follows:

```
--
applications:
- name: my-app
 command: node my-app.js
... the rest of your settings ...
```

Alternately, specify the start command with `cf push -c`.

```
$ cf push my-app -c "node my-app.js"
```

## Application Bundling

You do not need to run `npm install` before deploying your app. Cloud Foundry runs it for you when your app is pushed. You can, if you prefer, run `npm install` and create a `node_modules` folder inside of your app.

## Solve Discovery Problems

If Cloud Foundry does not automatically detect that your app is a Node.js app, you can override auto-detection by specifying the Node.js buildpack.

Add the buildpack into your `manifest.yml` and re-run `cf push` with your manifest:

```
--
applications:
- name: my-app
 buildpack: https://github.com/cloudfoundry/nodejs-buildpack
... the rest of your settings ...
```

Alternately, specify the buildpack on the command line with `cf push -b`:

```
$ cf push my-app -b https://github.com/cloudfoundry/nodejs-buildpack
```

## Bind Services

Refer to [Configure Service Connections for Node.js](#).

## Environment Variables

You can access environments variable programmatically.

For example, you can obtain `VCAP_SERVICES` as follows:

```
process.env.VCAP_SERVICES
```

Environment variables available to you include both those [defined by the system](#) and those defined by the Node.js buildpack, as described below.

### BUILD\_DIR

Directory into which Node.js is copied each time a Node.js app is run.

## CACHE\_DIR

Directory that Node.js uses for caching.

## PATH

The system path used by Node.js.

```
PATH=/home/vcap/app/bin:/home/vcap/app/node_modules/.bin:/usr/bin
```

## Environment Variables Defined by the Node Buildpack

Pivotal Application Service (PAS) provides configuration information to applications through environment variables. This topic describes the additional environment variables provided by the Node buildpack.

For more information about the standard environment variables provided by Pivotal Application Service (PAS), see the [Cloud Foundry Environment Variables](#) topic.

### Node Buildpack Environment Variables

The following table describes the environment variables provided by the Node buildpack.

Environment Variable	Description
BUILD_DIR	The directory where Node.js is copied each time a Node.js application runs.
CACHE_DIR	The directory Node.js uses for caching.
PATH	The system path used by Node.js: PATH=/home/vcap/app/bin:/home/vcap/app/node_modules/.bin:/bin:/usr/bin

## Configuring Service Connections for Node.js

This guide is for developers who wish to bind a data source to a Node.js application deployed and running on Cloud Foundry.

### Parse VCAP\_SERVICES for Credentials

You must parse the VCAP\_SERVICES environment variable in your code to get the required connection details such as host address, port, user name, and password.

For example, if you are using PostgreSQL, your VCAP\_SERVICES environment variable might look something like this:

```
{
 "mypostgres": [
 {
 "name": "myinstance",
 "credentials": {
 "uri": "postgres://myusername:mypassword@host.example.com:5432/serviceinstance"
 }
 }
]
}
```

This example JSON is simplified; yours may contain additional properties.

### Parse with cfenv

The `cfenv` package provides access to Cloud Foundry application environment settings by parsing all the relevant environment. The settings are returned as JavaScript objects. `cfenv` provides reasonable defaults when running locally, as well as when running as a Cloud Foundry application.

- <https://www.npmjs.org/package/cfenv>

### Manual Parsing

First, parse the VCAP\_SERVICES environment variable.

For example:

```
var vcap_services = JSON.parse(process.env.VCAP_SERVICES)
```

Then pull out the credential information required to connect to your service. Each service packages requires different information. If you are working with Postgres, for example, you will need a `uri` to connect. You can assign the value of the `uri` to a variable as follows:

```
var uri = vcap_services.mypostgres[0].credentials.uri
```

Once assigned, you can use your credentials as you would normally in your program to connect to your database.

## Connecting to a Service

You must include the appropriate package for the type of services your application uses. For example:

- Rabbit MQ via the [amqp](#) module
- Mongo via the [mongodb](#) and [mongoose](#) modules
- MySQL via the [mysql](#) module
- Postgres via the [pg](#) module
- Redis via the [redis](#) module

## Add the Dependency to package.json

Edit `package.json` and add the intended module to the `dependencies` section. Normally, only one would be necessary, but for the sake of the example we will add all of them:

```
{
 "name": "hello-node",
 "version": "0.0.1",
 "dependencies": {
 "express": "*",
 "mongodb": "*",
 "mongoose": "*",
 "mysql": "*",
 "pg": "*",
 "redis": "*",
 "amqp": "*"
 },
 "engines": {
 "node": "0.8.x"
 }
}
```

You must run `npm shrinkwrap` to regenerate your `npm-shrinkwrap.json` file after you edit `package.json`.

## PHP Buildpack

Use the PHP buildpack with PHP or HHVM runtimes.

## Supported Software and Versions

The [release notes page](#) has a list of currently supported modules and packages.

- **PHP Runtimes**
  - php-cli
  - php-cgi
  - php-fpm
- **Third-Party Modules**
  - New Relic, in connected environments only.

## Push an App

### 30 Second Tutorial

Getting started with this buildpack is easy. With the [Cloud Foundry Command Line Interface](#) installed, open a shell, change directories to the root of your PHP files and push your application using the argument `-b https://github.com/cloudfoundry/php-buildpack.git`.

Example:

```
$ mkdir my-php-app
$ cd my-php-app
$ cat << EOF > index.php
<?php
phpinfo();
?>
EOF
$ cf push -m 128M -b https://github.com/cloudfoundry/php-buildpack.git my-php-app
```

Change `my-php-app` in the above example to a unique name on your target Cloud Foundry instance to prevent a hostname conflict error and failed push.

The example above creates and pushes a test application, `my-php-app`, to Cloud Foundry. The `-b` argument instructs CF to use this buildpack. The remainder of the options and arguments are not specific to the buildpack, for questions on those consult the output of `cf help push`.

Here's a breakdown of what happens when you run the example above.

- On your PC:
  - It will create a new directory and one PHP file, which just invokes `phpinfo()`
  - Run `cf` to push your application. This will create a new application with a memory limit of 128M (more than enough here) and upload our test file.
- Within Cloud Foundry:
  - The buildpack is executed.
  - Application files are copied to the `htdocs` folder.
  - Apache HTTPD & PHP are downloaded, configured with the buildpack defaults and run.
  - Your application is accessible at the URL `http://my-php-app.example.com` (Replacing `example.com` with the domain of your public CF provider or private instance).

## More information about deploying

While the *30 Second Tutorial* shows how quick and easy it is to get started using the buildpack, it skips over quite a bit of what you can do to adjust, configure and extend the buildpack. The following sections and links provide a more in-depth look at the buildpack.

## Features

Here are some special features of the buildpack.

- Supports running commands or migration scripts prior to application startup.
- Supports an extension mechanism that allows the buildpack to provide additional functionality.
- Allows for application developers to provide custom extensions.
- Easy troubleshooting with the `BP_DEBUG` environment variable.
- Download location is configurable, allowing users to host binaries on the same network (i.e. run without an Internet connection)
- Smart session storage, defaults to file w/sticky sessions but can also use redis for storage.

## Examples

Here are some example applications that can be used with this buildpack.

- [php-info](#) ↗ This app has a basic index page and shows the output of `phpinfo()`.
- [PHPMyAdmin](#) ↗ A deployment of PHPMyAdmin that uses bound MySQL services.
- [PHPPgAdmin](#) ↗ A deployment of PHPPgAdmin that uses bound PostgreSQL services.
- [Drupal](#) ↗ A deployment of Drupal that uses bound MySQL service.
- [CodeIgniter](#) ↗ CodeIgniter tutorial application running on CF.
- [Stand Alone](#) ↗ An example which runs a standalone PHP script.
- [pgbouncer](#) ↗ An example which runs the PgBouncer process in the container to pool database connections.
- [phalcon](#) ↗ An example which runs a Phalcon based application.
- [composer](#) ↗ An example which uses Composer.

## Advanced Topics

See the following topics:

- [Tips for PHP Developers](#)
- [Getting Started Deploying PHP Apps](#)
- [PHP Buildpack Configuration](#)
- [Composer](#)
- [Sessions](#)
- [New Relic](#)

You can find the source for the buildpack on GitHub: <https://github.com/cloudfoundry/php-buildpack> ↗

## Proxy Support

If you need to use a proxy to download dependencies during staging, you can set the `http_proxy` and/or `https_proxy` environment variables. For more information, see the [Proxy Usage Docs](#).

## BOSH Configured Custom Trusted Certificate Support

For versions of PHP 5.6.0 and later, the default certificate location is `/usr/lib/ssl/certs`, which symlinks to `/etc/ssl/certs`. Your platform operator can configure the platform to [add the custom certificates into the application container](#) ↗.

## Help and Support

Join the #buildpacks channel in our [Slack community](#) ↗ if you need any further assistance.

For more information about using and extending the PHP buildpack in Cloud Foundry, see the [php-buildpack GitHub repository](#).

You can find current information about this buildpack on the PHP buildpack [release page](#) in GitHub.

## License

The Cloud Foundry PHP Buildpack is released under version 2.0 of the [Apache License](#).

## Tips for PHP Developers

### About the PHP Buildpack

For information about using and extending the PHP buildpack in Cloud Foundry, see the [php-buildpack GitHub repository](#).

You can find current information about this buildpack on the PHP buildpack [release page](#) in GitHub.

The buildpack uses a default PHP version specified in [.defaults/options.json](#) under the `PHP_VERSION` key.

To change the default version, specify the `PHP_VERSION` key in your app's [.bp-config/options.json](#) file.

## Getting Started Deploying PHP Apps

This topic is intended to guide you through the process of deploying PHP apps to Pivotal Application Service (PAS). If you experience a problem deploying PHP apps, review the [Troubleshooting](#) section below.

### Getting Started

#### Prerequisites

- Basic PHP knowledge
- The [Cloud Foundry Command Line Interface \(cf CLI\)](#) installed on your workstation

### A First PHP Application

```
$ mkdir my-php-app
$ cd my-php-app
$ cat << EOF > index.php
<?php
 phpinfo();
?>
EOF
$ cf push my-php-app -m 128M
```

Change “my-php-app” to a unique name or you may see an error and a failed push.

The example above creates and pushes a test application to Cloud Foundry.

Here is a breakdown of what happens when you run the example above:

- On your workstation...
  - It creates a new directory and one PHP file, which calls `phpinfo()`
  - Run `cf push` to push your application. This will create a new application with a memory limit of 128M and upload our test file.
- On Cloud Foundry...
  - The buildpack detects that your app is a php app
  - The buildpack is executed.
    - Application files are copied to the `htdocs` folder.
  - Apache HTTPD & PHP are downloaded, configured with the buildpack defaults, and run.
  - Your application is accessible at the default route. Use `cf app my-php-app` to view the url of your new app.

### Folder Structure

The easiest way to use the buildpack is to put your assets and PHP files into a directory and push it to PAS. This way, the buildpack will take your files and automatically move them into the `WEBDIR` (defaults to `htdocs`) folder, which is the directory where your chosen web server looks for the files.

### URL Rewriting

If you select Apache as your web server, you can include `.htaccess` files with your application.

Alternatively, you can [provide your own Apache or Nginx configurations](#).

### Prevent Access To PHP Files

The buildpack will put all of your files into a publicly accessible directory. In some cases, you might want to have PHP files that are not publicly accessible but are on the [include\\_path](#). To do that, create a `lib` directory in your project folder and place your protected files there.

For example:

```
$ ls -lRh
total 0
-rw-r--r-- 1 daniel staff 0B Feb 27 21:40 images
-rw-r--r-- 1 daniel staff 0B Feb 27 21:39 index.php
drwxr-xr-x 3 daniel staff 102B Feb 27 21:40 lib

./lib:
total 0
-rw-r--r-- 1 daniel staff 0B Feb 27 21:40 my.class.php <-- not public, http://app.cfapps.io/lib/my.class.php == 404
```

This comes with a catch. If your project legitimately has a `lib` directory, these files will not be publicly available because the buildpack does not copy a top-level `lib` directory into the `WEBDIR` folder. If your project has a `lib` directory that needs to be publicly available, then you have two options as follows:

## Option #1

In your project folder, create an `htdocs` folder (or whatever you've set for `WEBDIR`). Then move any files that should be publicly accessible into this directory. In the example below, the `lib/controller.php` file is publicly accessible.

Example:

```
$ ls -lRh
total 0
drwxr-xr-x 7 daniel staff 238B Feb 27 21:48 htdocs

./htdocs: <-- create the htdocs directory and put your files there
total 0
-rw-r--r-- 1 daniel staff 0B Feb 27 21:40 images
-rw-r--r-- 1 daniel staff 0B Feb 27 21:39 index.php
drwxr-xr-x 3 daniel staff 102B Feb 27 21:48 lib

./htdocs/lib: <-- anything under htdocs is public, including a lib directory
total 0
-rw-r--r-- 1 daniel staff 0B Feb 27 21:48 controller.php
```

Given this setup, it is possible to have both a public `lib` directory and a protected `lib` directory. The following example demonstrates this setup:

Example:

```
$ ls -lRh
total 0
drwxr-xr-x 7 daniel staff 238B Feb 27 21:48 htdocs
drwxr-xr-x 3 daniel staff 102B Feb 27 21:51 lib

./htdocs:
total 0
-rw-r--r-- 1 daniel staff 0B Feb 27 21:40 images
-rw-r--r-- 1 daniel staff 0B Feb 27 21:39 index.php
drwxr-xr-x 3 daniel staff 102B Feb 27 21:48 lib

./htdocs/lib: <-- public lib directory
total 0
-rw-r--r-- 1 daniel staff 0B Feb 27 21:48 controller.php

./lib: <-- protected lib directory
total 0
-rw-r--r-- 1 daniel staff 0B Feb 27 21:51 my.class.php
```

## Option #2

The second option is to pick a different name for the `LIBDIR`. This is a configuration option that you can set (it defaults to `lib`). Thus if you set it to something else such as `include`, your application's `lib` directory would no longer be treated as a special directory and it would be placed into `WEBDIR` (i.e. become public).

## Other Folders

Beyond the `WEBDIR` and `LIBDIR` directories, the buildpack also supports a `.bp-config` directory and a `.extensions` directory.

The `.bp-config` directory should exist at the root of your project directory and it is the location of application-specific configuration files. Application-specific configuration files override the default settings used by the buildpack. This link explains [application configuration files](#) in depth.

The `.extensions` directory should also exist at the root of your project directory and it is the location of application-specific custom extensions. Application-specific custom extensions allow you, the developer, to override or enhance the behavior of the buildpack. See the [php-buildpack README](#) in GitHub for more information about extensions.

## Troubleshooting

To troubleshoot problems using the buildpack, you can do one of the following:

1. Review the output from the buildpack. The buildpack writes basic information to stdout, for example the files that it downloads. The buildpack also writes information in the form of stack traces when a process fails.
2. Review the logs from the buildpack. The buildpack writes logs to disk. Follow the steps below to access these logs:
  - a. Run `cf ssh APP-NAME` to ssh into the app container, replacing `APP-NAME` with the name of your app.
  - b. Run `cat app/.bp/logs/bp.log` to view the logs.

```
$ cf ssh my-app
$ cat app/.bp/logs/bp.log
```

By default, log files contain more detail than output to stdout. Set the `BP_DEBUG` environment variable to `true` for more verbose logging. This instructs the buildpack to set its log level to `DEBUG`, and to output logs to stdout. Follow [Environment Variables documentation](#) to set `BP_DEBUG`.

## Increase Log Output with fpm.d

If you use `fpm.d`, follow the steps below to configure `fpm` to redirect worker stdout and stderr into the logs.

1. Create a file in the `.bp-config/php/fpm.d/` directory of your app.
2. Add `catch_workers_output=yes` to the file you created.
3. Push your app.

For more information about allowed configuration settings in the `.bp-config/php/fpm.d` directory, see the [List of global php-fpm.conf directives](#).

You can see an example [fpm fixture and configuration file](#) in the php-buildpack GitHub repository.

## PHP Buildpack Configuration

### Defaults

The PHP buildpack stores all of its default configuration settings in the [defaults](#) directory.

### options.json

The `options.json` file is the configuration file for the buildpack itself. It instructs the buildpack what to download, where to download it from, and how to install it. It allows you to configure package names and versions (i.e. PHP, HTTPD, or Nginx versions), the web server to use (HTTPD, Nginx, or None), and the PHP extensions that are enabled.

The buildpack overrides the default `options.json` file with any configuration it finds in the `.bp-config/options.json` file of your application.

Below is an explanation of the common options you might need to change.

Variable	Explanation
WEB_SERVER	Sets the web server to use. Must be one of <code>httpd</code> , <code>nginx</code> , or <code>none</code> . This value defaults to <code>httpd</code> .
HTTPD_VERSION	Sets the version of Apache HTTPD to use. Currently the build pack supports the latest stable version. This value will default to the latest release that is supported by the build pack.
ADMIN_EMAIL	The value used in HTTPD's configuration for <a href="#">ServerAdmin</a>
NGINX_VERSION	Sets the version of Nginx to use. By default, the buildpack uses the latest stable version.
PHP_VERSION	Sets the version of PHP to use. Set to a minor instead of a patch version, such as <code>"{PHP_70_LATEST}"</code> . See <a href="#">options.json</a> .
PHP_EXTENSIONS	(DEPRECATED) A list of the <a href="#">extensions</a> to enable. <code>bz2</code> , <code>zlib</code> , <code>curl</code> , and <code>mcrypt</code> are enabled by default.
PHP_MODULES	A list of the <a href="#">modules</a> to enable. No modules are explicitly enabled by default, but the buildpack automatically chooses either <code>fpm</code> or <code>cli</code> . You can explicitly enable any or all of the following: <code>fpm</code> , <code>cli</code> , <code>cgi</code> , and <code>pear</code> .
ZEND_EXTENSIONS	A list of the Zend extensions to enable. Nothing is enabled by default.
APP_START_CMD	When the <code>WEB_SERVER</code> option is set to <code>none</code> , this command is used to start your app. If <code>WEB_SERVER</code> and <code>APP_START_CMD</code> are not set, then the buildpack searches, in order, for <code>app.php</code> , <code>main.php</code> , <code>run.php</code> , or <code>start.php</code> . This option accepts arguments.
WEBDIR	The root directory of the files served by the web server specified in <code>WEB_SERVER</code> . Defaults to <code>htdocs</code> . Other common settings are <code>public</code> , <code>static</code> , or <code>html</code> . The path is relative to the root of your application.
LIBDIR	This path is added to PHP's <code>include_path</code> . Defaults to <code>lib</code> . The path is relative to the root of your application.
HTTP_PROXY	The buildpack downloads uncached dependencies using HTTP. If you are using a proxy for HTTP access, set its URL here.
HTTPS_PROXY	The buildpack downloads uncached dependencies using HTTPS. If you are using a proxy for HTTPS access, set its URL here.
ADDITIONAL_PREPROCESS_CMDS	A list of additional commands that run prior to the application starting. For example, you might use this command to run migration scripts or static caching tools before the application launches.

For details about supported versions, see the [release notes](#) for your buildpack version.

### HTTPD, Nginx, and PHP configuration

The buildpack automatically configures HTTPD, Nginx, and PHP for your application. This section explains how to modify the configuration.

The `.bp-config` directory in your application can contain configuration overrides for these components. Name the directories `httpd`, `nginx`, and `php`. We recommend that you use `php.ini.d` or `fpm.d`.

 **NOTE:** If you override the `php.ini` or `php-fpm.conf` files, many other forms of configuration will not work.

For example: `.bp-config httpd nginx php`

Each directory can contain configuration files that the component understands.

For example, to change HTTPD logging configuration:

```
$ ls -l .bp-config/httpd/extra/
total 8
-rw-r--r-- 1 daniel staff 396 Jan 3 08:31 httpd-logging.conf
```

In this example, the `httpd-logging.conf` file overrides the one provided by the buildpack. We recommend that you copy the default from the buildpack and modify it.

You can find the default configuration files in the [PHP Buildpack /defaults/config directory ↗](#).

You should be careful when modifying configurations, as doing so can cause your application to fail, or cause Cloud Foundry to fail to stage your application.

You can add your own configuration files. The components will not know about these, so you must ensure that they are included. For example, you can add an include directive to the [httpd configuration ↗](#) to include your file:

```
ServerRoot "${HOME}/httpd"
Listen ${PORT}
ServerAdmin "${HTTPD_SERVER_ADMIN}"
ServerName "0.0.0.0"
DocumentRoot "${HOME}/#${WEBDIR}"
Include conf/extra/httpd-modules.conf
Include conf/extra/httpd-directories.conf
Include conf/extra/httpd-mime.conf
Include conf/extra/httpd-logging.conf
Include conf/extra/httpd-mpm.conf
Include conf/extra/httpd-default.conf
Include conf/extra/httpd-remoteip.conf
Include conf/extra/httpd-php.conf
Include conf/extra/httpd-my-special-config.conf # This line includes your additional file.
```

## .bp-config/php/php.ini.d/

The buildpack adds any `.bp-config/php/php.ini.d/FILE-NAME.ini` files it finds in the application to the PHP configuration. You can use this to change any value acceptable to `php.ini`. See [http://php.net/manual/en/ini.list.php ↗](http://php.net/manual/en/ini.list.php) for a list of directives.

For example, adding a file `.bp-config/php/php.ini.d/something.ini` to your app, with the following contents, overrides both the default charset and mimetype.

```
default_charset="UTF-8"
default_mimetype="text/xhtml"
```

## Precedence

In order of highest precedence, php configuration values come from the following sources:

- php scripts using `ini_set()` to manually override config files
- user.ini files for local values
- `.bp-config/php/php.ini.d` to override master value, but not local values from user.ini files

## .bp-config/php/fpm.d/

The buildpack adds any files it finds in the application under `.bp-config/php/fpm.d` that end with `.conf` (i.e `my-config.conf`) to the PHP-FPM configuration. You can use this to change any value acceptable to `php-fpm.conf`. See [http://php.net/manual/en/install.fpm.configuration.php ↗](http://php.net/manual/en/install.fpm.configuration.php) for a list of directives.

PHP FPM config snippets are included by the buildpack into the global section of the configuration file. If you need to apply configuration settings for a PHP FPM worker, that needs to be indicated in your configuration file as well.

For example:

```
; This option is specific to the `www` pool
[www]
catch_workers_output = yes
```

## PHP Extensions

The buildpack adds any `.bp-config/php/php.ini.d/FILE-NAME.ini` files it finds in the application to the PHP configuration. You can use this to enable PHP or Zend extensions. For example:

```
extension=redis.so
extension=gd.so
zend_extension=opcache.so
```

If an extension is already present and enabled in the compiled php, for example `intl`, you do not need to explicitly enable it to use that extension.

### PHP\_EXTENSIONS vs. ZEND\_EXTENSIONS

PHP has two kinds of extensions, *PHP extensions* and *Zend extensions*. These hook into the PHP executable in different ways. See <https://wiki.php.net/internals/extensions> for more information about the way extensions work internally in the engine.

Because they hook into the PHP executable in different ways, they are specified differently in ini files. Apps fail if a Zend extension is specified as a PHP extension, or a PHP extension is specified as a Zend extension.

If you see the following error, move the `example` extension from `extension` to `zend_extension`, then re-push your app:

```
php-fpm | [warn-ioncube] The example Loader is a Zend-Engine extension and not a module (pid 40)
php-fpm | [warn-ioncube] Please specify the Loader using 'zend_extension' in php.ini (pid 40)
php-fpm | NOTICE: PHP message: PHP Fatal error: Unable to start example Loader module in Unknown on line 0
```

If you see the following error, move the `example` extension from `zend_extension` to `extension`, then re-push your app.

```
NOTICE: PHP message: PHP Warning: example MUST be loaded as a Zend extension in Unknown on line 0
```

## PHP Modules

You can include the following modules by adding them to the `PHP_MODULES` list:

- `cli`, installs `php` and `phar`
- `fpm`, installs `PHP-FPM`
- `cgi`, installs `php-cgi`
- `pear`, installs Pear

By default, the buildpack installs the `cli` module when you push a standalone app, and installs the `fpm` module when you push a web app. You must specify `cgi` and `pear` if you want them installed.

## Buildpack Extensions

The buildpack comes with extensions for its default behavior. These are the [HTTPD](#), [Nginx](#), [PHP](#), and [NewRelic](#) extensions.

The buildpack is designed with an extension mechanism, allowing app developers to add behavior to the buildpack without modifying the buildpack code.

When you push an app, the buildpack runs any extensions found in the `.extensions` directory of your app.

The [Developer Documentation](#) explains how to write extensions.



## Composer

Composer is activated when you supply a `composer.json` or `composer.lock` file. A `composer.lock` is not required, but is strongly recommended for consistent deployments.

You can require dependencies for packages and extensions. Extensions must be prefixed with the standard `ext-`. If you reference an extension that is available to the buildpack, it will automatically be installed. See the main [README](#) for a list of supported extensions.

The buildpack uses the version of PHP specified in your `composer.json` or `composer.lock` file. Composer settings override the version set in the `options.json` file.

The PHP buildpack supports a subset of the version formats supported by Composer. The buildpack supported formats are:

Example	Expected Version
<code>5.3.*</code>	latest 5.4.x release (5.3 is not supported)
<code>&gt;=5.3</code>	latest 5.4.x release (5.3 is not supported)
<code>5.4.*</code>	latest 5.4.x release
<code>&gt;=5.4</code>	latest 5.4.x release
<code>5.5.*</code>	latest 5.5.x release
<code>&gt;=5.5</code>	latest 5.5.x release
<code>5.4.x</code>	specific 5.4.x release that is listed
<code>5.5.x</code>	specific 5.5.x release that is listed

## Configuration

The buildpack runs with a set of default values for Composer. You can adjust these values by adding a `.bp-config/options.json` file to your application and setting any of the following values in it.

Variable	Explanation
<code>COMPOSER_VERSION</code>	The version of Composer to use. It defaults to the latest bundled with the buildpack.
<code>COMPOSER_INSTALL_OPTIONS</code>	A list of options that should be passed to <code>composer install</code> . This defaults to <code>--no-interaction --no-dev --no-progress</code> . The <code>--no-progress</code> option must be used due to the way the buildpack calls Composer.
<code>COMPOSER_VENDOR_DIR</code>	Allows you to override the default value used by the buildpack. This is passed through to Composer and instructs it where to create the <code>vendor</code> directory. Defaults to <code>{BUILD_DIR}/{LIBDIR}/vendor</code> .
<code>COMPOSER_BIN_DIR</code>	Allows you to override the default value used by the buildpack. This is passed through to Composer and instructs it where to place executables from packages. Defaults to <code>{BUILD_DIR}/php/bin</code> .
<code>COMPOSER_CACHE_DIR</code>	Allows you to override the default value used by the buildpack. This is passed through to Composer and instructs it where to place its cache files. Generally you should not change this value. The default is <code>{CACHE_DIR}/composer</code> which is a subdirectory of the cache folder passed in to the buildpack. Composer cache files will be restored on subsequent application pushes.

By default, the PHP buildpack uses the `composer.json` and `composer.lock` files that reside inside the root directory, or in the directory specified as `WEBDIR` in your `options.json`. If you have composer files inside your app, but not in the default directories, use a `COMPOSER_PATH` environment variable for your app to specify this custom location, relative to the app root directory. Note that the `composer.json` and `composer.lock` files must be in the same directory.

## GitHub API Request Limits

Composer uses GitHub's API to retrieve zip files for installation into the application folder. If you do not vendor dependencies before pushing an app, Composer will fetch dependencies during staging using the GitHub API.

GitHub's API is request-limited. If you reach your daily allowance of API requests (typically 60), GitHub's API returns a `403` error and staging fails.

There are two ways to avoid the request limit:

1. Vendor dependencies before pushing your application.

2. Supply a GitHub OAuth API token.

## Vendor Dependencies

To vendor dependencies, you must run `composer install` before you push your application. You might also need to configure `COMPOSER_VENDOR_DIR` to “vendor”.

## Supply a GitHub Token

Composer can use [GitHub API OAuth tokens](#), which increase your request limit, typically to 5000 per day.

During staging, the buildpack looks for this token in the environment variable `COMPOSER_GITHUB_OAUTH_TOKEN`. If you supply a valid token, Composer uses it. This mechanism does not work if the token is invalid.

To supply the token, you can use either of the following methods:

1. `cf set-env YOUR_APP_NAME COMPOSER_GITHUB_OAUTH_TOKEN "OAUTH_TOKEN_VALUE"`
2. Add the token to the `env` block of your application manifest.

## Buildpack Staging Environment

Composer runs in the buildpack staging environment. Variables set with `cf set-env` or with [a manifest.yml ‘env’ block](#) are visible to Composer.

For example:

```
$ cf push a_symfony_app --no-start
$ cf set-env a_symfony_app SYMFONY_ENV "prod"
$ cf start a_symfony_app
```

In this example, `a_symfony_app` is supplied with an environment variable, `SYMFONY_ENV`, which is visible to Composer and any scripts started by Composer.

## Non-configurable Environment Variables

User-assigned environment variables are applied to staging and runtime. Unfortunately, `LD_LIBRARY_PATH` and `PHPRC` must be different for staging and runtime. The buildpack takes care of setting these variables, which means user values for these variables are ignored.

## Sessions

### Usage

When your application has one instance, it's mostly safe to use the default session storage, which is the local file system. You would only see problems if your single instance crashes as the local file system would go away and you'd lose your sessions. For many applications, this will work just fine but please consider how this will impact your application.

If you have multiple application instances or you need a more robust solution for your application, then you'll want to use Redis or Memcached as a backing store for your session data. The build pack supports both and when one is bound to your application it will detect it and automatically configure PHP to use it for session storage.

By default, there's no configuration necessary. Create a Redis or Memcached service, make sure the service name contains `redis-sessions` or `memcached-sessions` and then bind the service to the application.

Example:

```
$ cf create-service redis some-plan app-redis-sessions
$ cf bind-service app app-redis-sessions
$ cf restage app
```

If you want to use a specific service instance or change the search key, you can do that by setting either `REDIS_SESSION_STORE_SERVICE_NAME` or `MEMCACHED_SESSION_STORE_SERVICE_NAME` in `.bp-config/options.json` to the new search key. The session configuration extension will then search the bound services by name for the new session key.

## Configuration Changes

When detected, the following changes will be made.

### Redis

- the `redis` PHP extension will be installed, which provides the session save handler
- `session.name` will be set to `PHPSESSIONID` which disables sticky sessions
- `session.save_handler` is configured to `redis`
- `session.save_path` is configured based on the bound credentials, for example `tcp://host:port?auth=pass`

### Memcached

- the `memcached` PHP extension will be installed, which provides the session save handler
- `session.name` will be set to `PHPSESSIONID` which disables sticky sessions
- `session.save_handler` is configured to `memcached`
- `session.save_path` is configured based on the bound credentials (i.e. `PERSISTENT=app_sessions host:port`)
- `memcached.sess_binary` is set to `On`
- `memcached.use_sasl` is set to `On`, which enables authentication
- `memcached.sess_sasl_username` and `memcached.sess_sasl_password` are set with the service credentials

## New Relic

[New Relic](#) collects analytics about application and client-side performance.

## Configuration

You can configure New Relic for the PHP buildpack in one of two ways:

- With a license key
- With a Cloud Foundry service

### With a License Key

Use this method if you already have a New Relic account,

1. In a web browser, navigate to the New Relic website to find your [license key](#).
2. Set the value of the environment variable `NEWRELIC_LICENSE` to your New Relic license key, either through the [manifest.yml file](#) or by running the `cf set-env` command.

For more information, see <https://github.com/cloudfoundry/php-buildpack#supported-software>

### With a Cloud Foundry Service

To configure New Relic for the PHP buildpack with a Cloud Foundry service, bind a New Relic service to the app. The buildpack automatically detects and configures New Relic.

Your `VCAP_SERVICES` environment variable must contain a service named `newrelic`, the `newrelic` service must contain a key named `credentials`, and the `credentials` key must contain named `licenseKey`.

 **NOTE:** You cannot configure New Relic for the PHP buildpack with user-provided services.

## Python Buildpack

### Push an App

Cloud Foundry automatically uses this buildpack if it detects a `requirements.txt` or `setup.py` file in the root directory of your project.

If your Cloud Foundry deployment does not have the Python Buildpack installed, or the installed version is out of date, you can use the latest version by specifying it with the `-b` option when you push your app. For example:

```
$ cf push my_app -b https://github.com/cloudfoundry/python-buildpack.git
```

### Supported Versions

You can find the list of supported Python versions in the [Python buildpack release notes](#).

### Specify a Python Version

You can specify a version of the Python runtime by including it within a `runtime.txt` file. For example:

```
$ cat runtime.txt
python-3.5.2
```

The buildpack only supports the stable Python versions, which are listed in the `manifest.yml` and [Python buildpack release notes](#).

To request the latest Python version in a patch line, replace the patch version with `x : 3.6.x`. To request the latest version in a minor line, replace the minor version: `3.x`.

If you try to use a binary that is not currently supported, staging your app fails and you see the following error message:

```
Could not get translated url, exited with: DEPENDENCY_MISSING_IN_MANIFEST: ...
!
! exit
!
Staging failed: Buildpack compilation step failed
```

### Specify a Start Command

The Python buildpack does not generate a [default start command](#) for your applications.

To stage with the Python buildpack and start an application, do one of the following:

- Supply a Procfile. For more information about Procfiles, see the [Configuring a Production Server](#) topic. The following example Procfile specifies `gunicorn` as the start command for a web app running on Gunicorn:

```
web: gunicorn SERVER-NAME:APP-NAME
```

- Specify a start command with `-c`. The following example specifies `waitress-serve` as the start command for a web app running on waitress:

```
$ cf push python-app -c "waitress-serve --port=$PORT DJANGO-WEB-APP.wsgi:MY-APP"
```

- Specify a start command in the application manifest by setting the `command` attribute. For more information, see the [Deploying with Application Manifests](#) topic.

### Vendor App Dependencies

If you are deploying in an environment that is disconnected from the Internet, your application [must vendor its dependencies](#).

For the Python buildpack, use `pip`:

```
$ cd YOUR-APP-DIR
$ mkdir -p vendor

vendors all the pip *.tar.gz into vendor/
$ pip download -r requirements.txt --no-binary=:all: -d vendor
```

`cf push` uploads your vendored dependencies. The buildpack installs them directly from the `vendor/` directory.

 **Note:** To ensure proper installation of dependencies, we recommend non-binary vendored dependencies. The above `pip install` command achieves this.

## Private Dependency Repository

To deploy apps in an environment that needs to use a private dependency repository, you have the following options:

- [PIP](#)
- [Conda](#)

### PIP

To install dependencies using PIP, add the URL of the repository to the `requirements.txt` file in the following format:

```
--index-url=https://example.com/api/pypi/ext_pypi/simple
fixtures==2.0.0
```

If the private repository uses a custom SSL certificate that is installed on the platform, you may see an error similar to the following:

```
Could not fetch URL https://example.com/api/pypi/ext_pypi/simple/fixtures/:
There was a problem confirming the ssl certificate:
[SSL: CERTIFICATE_VERIFY_FAILED] certificate verify failed (_ssl.c:777) - skipping
```

This error occurs because `pip` does not use system certificates by default. To resolve this issue, set the `PIP_CERT` environment variable in the `manifest.yml` file to point to the system certificate store.

For example:

```

env:
 PIP_CERT: /etc/ssl/certs/ca-certificates.crt
```

### Conda

To install dependencies using Conda, add a `channels` block to the `environment.yml` file.

In the `channels` block, list custom channels and add `nodelfaults`. Specifying `nodelfaults` tells Conda to only use the channels in the channels block.

For example:

```
channels:
 - https://conda.example.com/repo
 - nodelfaults
```

If the private repository uses a custom SSL certificate that is installed on the platform, you may see an error similar to the following:

Error: Connection error: [SSL: CERTIFICATE\_VERIFY\_FAILED] certificate verify failed (\_ssl.c:581):

This error occurs because `conda` does not use system certificates by default. To resolve this issue, set the `CONDA_SSL_VERIFY` environment variable in the `manifest.yml` file to point to the system certificate store.

For example:

```

env:
 CONDA_SSL_VERIFY: /etc/ssl/certs/ca-certificates.crt
```

## Parse Environment Variables

The `cfenv` package provides access to Cloud Foundry application environment settings by parsing all the relevant environment variables. The settings are returned as a class instance. See <https://github.com/jmcarp/py-cfenv> for more information.

## Miniconda Support (starting in buildpack version 1.5.6)

To use miniconda instead of pip for installing dependencies, place an `environment.yml` file in the root directory.

For examples, see our sample apps:

- [Using Python 2 with miniconda](#)
- [Using Python 3 with miniconda](#)

## Pipenv Support (starting in buildpack version 1.5.19)

To use [Pipenv](#) instead of pip (directly) for installing dependencies, place a `Pipfile` in the root directory. Easiest to let Pipenv generate this for you.

## NLTK Support

To use NLTK corpora in your app, you can include an `nltk.txt` file in the root of your application. Each line in the file specifies one dataset to download. The full list of data sets available this way can be found [on the NLTK website](#). The `id` listed for the corpora on that page is the string you should include in your app's `nltk.txt`.

Example `nltk.txt` : brown wordnet

Having an `nltk.txt` file only causes the buildpack to download the corpora. You still must specify NLTK as a dependency of your app if you want to use it to process the corpora files.

## Proxy Support

If you need to use a proxy to download dependencies during staging, you can set the `http_proxy` and `https_proxy` environment variables. For more information, see [Using a Proxy](#).

## BOSH Configured Custom Trusted Certificate Support

Versions of Python 2.7.9 and later use certificates stored in `/etc/ssl/certs`. Your platform operator can configure the platform to [add the custom certificates into the application container](#).

## Help and Support

Join the #buildpacks channel in our [Slack community](#) if you need any further assistance.

For more information about using and extending the Python buildpack in Cloud Foundry, see the [python-buildpack GitHub repository](#).

You can find current information about this buildpack on the [Python buildpack release page](#) in GitHub.

## R Buildpack

This topic describes how to push your R app to Cloud Foundry and how to configure your R app to use the R buildpack.

### Push an App

Cloud Foundry automatically uses the R buildpack if it detects a `r.yml` file in the root directory of your project.

If your Cloud Foundry deployment does not have the R buildpack installed, or the installed version is out of date, you can use the latest version by specifying it with the `-b` option when you push your app. For example:

```
$ cf push my_app -b https://github.com/cloudfoundry/r-buildpack.git
```

### Supported Versions

You can find the list of supported R versions in the [R buildpack release notes](#).

### Start Command

The R buildpack does not generate a default start command for your applications. Instead, you must specify a start command for your app.

To stage an app with the R buildpack and start the app, do one of the following:

- **Option 1:** Supply a Procfile. For more information about Procfiles, see [Production Server Configuration](#). The following example Procfile specifies `R -f` as the start command for a web app with the entrypoint `simple.r`:

```
web: R -f simple.r
```

- **Option 2:** Specify a start command with `-c`:

```
$ cf push r-app -c "R -f simple.r"
```

- **Option 3:** Specify a start command in the application manifest by setting the `command` attribute. For more information, see the [Deploying with Application Manifests](#) topic.

For more information about starting apps, see [Starting, Restarting, and Restaging Applications](#).

### Specifying App Dependencies

As of v0.0.5, the following packages are provided by the buildpack:

- Rserve
- forecast
- shiny

To specify additional packages needed by your app, provide the CRAN mirror and names of the packages inside your `r.yml` file. You can also specify the number of threads to use for parallel installation. For example:

```

packages:
- cran_mirror: https://cran.r-project.org
 num_threads: 2
 packages:
 - name: string
 - name: jsonlite
```

## Vendorizing App Dependencies

If you are deploying in an environment that is disconnected from the Internet, you must *vendor* your app's dependencies, which means you must make these packages available for offline use. You can vendor dependencies by using a package manager.

To set up your own custom, local CRAN-like repository to vendor your packages, create the `src/contrib` directories and populate them with your package source tarballs. For more information, see [How to Set Up a Custom CRAN-like Repository](#) published on the *Packrat Documentation* site.

Add the `src/contrib` directories containing your package tarballs to a `vendor_r` directory at the root of your app. This directory is named `vendor_r` so as not to conflict with vendor directories of other languages, such as python.

Then, inside your `r.yml`, provide the names of your vendored packages in the `packages` list:

```

packages:
- packages:
- name: string
- name: jsonlite
```

`cf push` uploads your vendored dependencies. The buildpack installs them directly from the `vendor_r/` directory.

Example app directory tree:

```
└── r.yml
└── simple-app.r
 └── vendor_r
 └── src
 └── contrib
 ├── PACKAGES
 ├── PACKAGES.gz
 ├── jsonlite_1.5.tar.gz
 └── stringr_1.3.1.tar.gz
```

For more information about using buildpacks in disconnected environments, see [Disconnected environments](#).

## Proxy Support

If you need to use a proxy to download dependencies during staging, you can set the `http_proxy` and `https_proxy` environment variables. For more information, see [Using a Proxy](#).

## BOSH Configured Custom Trusted Certificate Support

R uses certificates stored in `/etc/ssl/certs`. Your platform operator can configure the platform to add the custom certificates into the application container. For more information, see [Configuring Trusted System Certificates for Applications](#).

## Help and Support

Join the `#buildpacks` channel in our [Slack community](#) if you need any further assistance.

For more information about using and extending the R buildpack in Cloud Foundry, see the [R-buildpack GitHub repository](#).

You can also find current information about this buildpack on the [R buildpack release page](#) in GitHub.

## Ruby Buildpack

### Push Apps

Cloud Foundry uses the Ruby buildpack if your app has a `Gemfile` and `Gemfile.lock` in the root directory. The Ruby buildpack uses Bundler to install your dependencies.

If your Cloud Foundry deployment does not have the Ruby buildpack installed or the installed version is out of date, push your app with the `-b` option to specify the buildpack:

```
$ cf push MY-APP -b https://github.com/cloudfoundry/ruby-buildpack.git
```

For more detailed information about deploying Ruby applications see the following topics:

- [Tips for Ruby Developers](#)
- [Getting Started Deploying Apps](#)
- [Configuring Rake Tasks for Deployed Apps](#)
- [Environment Variables Defined by the Ruby Buildpack](#)
- [Configuring Service Connections for Ruby](#)
- [Support for Windows Gemfiles](#)

You can find the source for the Ruby buildpack in the [Ruby buildpack repository](#) on GitHub.

### Supported Versions

You can find supported Ruby versions in the [release notes for the Ruby buildpack](#) on GitHub.

### Specify a Ruby Version

Specify specific versions of the Ruby runtime in the `Gemfile` for your app as described in the sections below.

 **Note:** If you don't specify a version of the Ruby runtime, we will use the default Ruby version listed in the [buildpack manifest.yml file](#).

### MRI

For MRI, specify the version of Ruby in your `Gemfile` as follows:

```
ruby '~> 2.2.3'
```

With this example declaration in the `Gemfile`, if Ruby versions `2.2.4`, `2.2.5`, and `2.3.0` are present in the Ruby buildpack, the app uses Ruby version `2.2.5`.

For more information about the `ruby` directive for Bundler Gemfiles, see the [Bundler documentation](#).

 **Note:** If you use v1.6.18 or earlier, you must specify an exact version, such as `ruby '2.2.3'`. In [Ruby Buildpack v1.6.18](#) and earlier, Rubygems do not support version operators for the `ruby` directive.

### JRuby

For JRuby, specify the version of Ruby in your `Gemfile` based on the version of JRuby your app uses.

JRuby version 1.7.x supports either 1.9 mode or 2.0 mode.

- For 1.9 mode, use the following: `ruby ruby '1.9.3', :engine => 'jruby', :engine_version => '1.7.25'`
- For 2.0 mode, use the following: `ruby ruby '2.0.0', :engine => 'jruby', :engine_version => '1.7.25'`

For JRuby version >= 9.0, use the following: `ruby ruby '2.2.3', :engine => 'jruby', :engine_version => '9.0.5.0'`

The Ruby buildpack only supports the stable Ruby versions listed in the `manifest.yml` and [release notes for the Ruby buildpack](#) on GitHub.

If you try to use a binary that is not supported, staging your app fails with the following error message:

```
Could not get translated url, exited with: DEPENDENCY_MISSING_IN_MANIFEST: ...
!
! exit
!
Staging failed: Buildpack compilation step failed
```

 **Note:** The Ruby buildpack does not support the pessimistic version operator `>~` on the Gemfile `ruby` directive for JRuby.

## Vendor App Dependencies

As stated in the [Disconnected Environments documentation](#), your application must ‘vendor’ its dependencies.

For the Ruby buildpack, use the `bundle package --all` command in Bundler to vendor the dependencies.

Example:

```
$ cd my-app-directory
bundle package --all
```

The `cf push` command uploads your vendored dependencies. The Ruby buildpack compiles any dependencies requiring compilation while staging your app.

## Buildpack Logging and Application Logging

The Ruby buildpack only runs during the staging process, and only logs what is important to staging, such as what is being downloaded, what the configuration is, and work that the buildpack does on your application.

The buildpack stops logging when the staging process finishes. The Loggregator handles application logging.

Your application must write to STDOUT or STDERR for its logs to be included in the Loggregator stream. For more information, see the [Application Logging in Cloud Foundry](#) topic.

If you are deploying a Rails application, the buildpack may or may not automatically install the necessary plugin or gem for logging, depending on the Rails version of the application:

- Rails 2.x: The buildpack automatically installs the `rails_log_stdout` plugin into the application. For more information about the `rails_log_stdout` plugin, refer to the [GitHub README](#).
- Rails 3.x: The buildpack automatically installs the `rails_12factor` gem if it is not present and issues a warning message. You must add the `rails_12factor` gem to your `Gemfile` to quiet the warning message. For more information about the `rails_12factor` gem, refer to the [GitHub README](#).
- Rails 4.x: The buildpack only issues a warning message that the `rails_12factor` gem is not present, but does not install the gem. You must add the `rails_12factor` gem to your `Gemfile` to quiet the warning message. For more information about the `rails_12factor` gem, refer to the [GitHub README](#).

For more information about the `rails_12factor` gem, refer to the [GitHub README](#).

## Proxy Support

If you need to use a proxy to download dependencies during staging, you can set the `http_proxy` and/or `https_proxy` environment variables. For more

information, see [Using a Proxy](#).

## BOSH Configured Custom Trusted Certificate Support

Ruby uses certificates stored in `/etc/ssl/certs`. Your platform operator can configure the platform to add the custom certificates into the app container.

## Help and Support

Join the `#buildpacks` channel in our [Slack community](#) if you need any further assistance.

For more information about using and extending the Ruby buildpack in Cloud Foundry, see the [Ruby buildpack repository](#) on GitHub.

You can find current information about the Ruby buildpack in the [release notes for the Ruby buildpack](#) on GitHub.

## Tips for Ruby Developers

This page has information specific to deploying Rack, Rails, or Sinatra apps.

## App Bundling

You must run [Bundler](#) to create a `Gemfile` and a `Gemfile.lock`. These files must be in your app before you push to Cloud Foundry.

## Rack Config File

For Rack and Sinatra, you must have a `config.ru` file. For example:

```
require './hello_world'
run HelloWorld.new
```

## Asset Precompilation

Cloud Foundry supports the Rails asset pipeline. If you do not precompile assets before deploying your app, Cloud Foundry precompiles them when staging the app. Precompiling before deploying reduces the time it takes to stage an app.

Use the following command to precompile assets before deployment:

```
$ rake assets:precompile
```

Note that the Rake precompile task reinitializes the Rails app. This could pose a problem if initialization requires service connections or environment checks that are unavailable during staging. To prevent reinitialization during precompilation, add the following line to `application.rb`:

```
config.assets.initialize_on_precompile = false
```

If the `assets:precompile` task fails, Cloud Foundry uses live compilation mode, the alternative to asset precompilation. In this mode, assets are compiled when they are loaded for the first time. You can force live compilation by adding the following line to `application.rb`.

```
Rails.application.config.assets.compile = true
```

## Running Rake Tasks

Cloud Foundry does not provide a mechanism for running a Rake task on a deployed app. If you need to run a Rake task that must be performed in the Cloud Foundry environment, rather than locally before deploying or redeploying, you can configure the command that Cloud Foundry uses to start the app to invoke the Rake task.

An app start command is configured in the app manifest file, `manifest.yml`, using the `command` attribute.

For more information about app manifests and supported attributes, see the [Deploying with Application Manifests](#) topic.

### Example: Invoking a Rake database migration task at app startup

The following is an example of migrating a database schema using a Rake task. For more information about migrating database schemas, see [Services Overview](#).

1. If a Rakefile does not exist, create one and add it to your app directory.
2. In your Rakefile, add a Rake task to limit an idempotent command to the first instance of a deployed app:

```
namespace :cf do
 desc "Only run on the first application instance"
 task :on_first_instance do
 instance_index = JSON.parse(ENV["VCAP_APPLICATION"])[“instance_index”] rescue nil
 exit(0) unless instance_index == 0
 end
end
```

- Add the task to the `manifest.yml` file, referencing the idempotent command `rake db:migrate` with the `command` attribute.

```

```

```
applications:
- name: my-rails-app
 command: bundle exec rake cf:on_first_instance db:migrate && bundle exec rails s -p $PORT -e $RAILS_ENV
```

- Update the app using `cf push`.

## Rails 3 Worker Tasks

This section shows you how to create and deploy an example Rails app that uses a worker library to defer a task that a separate app executes.

The guide also describes how to scale the resources available to the worker app.

**Note:** Most worker tasks do not serve external requests. Use the `--no-route` flag with the `cf push` command, or `no-route: true` in the app manifest, to suppress route creation and remove existing routes.

### Choose a Worker Task Library

You must choose a worker task library. The table below summarizes the three main libraries available for Ruby / Rails:

Library	Description
<a href="#">Delayed::Job</a>	A direct extraction from <a href="#">Shopify</a> where the job table is responsible for a multitude of core tasks.
<a href="#">Resque</a>	A Redis-backed library for creating background jobs, placing those jobs on multiple queues, and processing them later.
<a href="#">Sidekiq</a>	Uses threads to handle many messages at the same time in the same process. It does not require Rails, but integrates tightly with Rails 3 to simplify background message processing. This library is Redis-backed and semi-compatible with Resque messaging.

For other alternatives, see [https://www.ruby-toolbox.com/categories/Background\\_Jobs](https://www.ruby-toolbox.com/categories/Background_Jobs).

### Create an Example App

For the purposes of the example app, we use Sidekiq.

First, create a Rails app with an arbitrary model named “Things”:

```
$ rails create rails-sidekiq
$ cd rails-sidekiq
$ rails g model Thing title:string description:string
```

Add `sidekiq` and `uuidtools` to the Gemfile:

```
source 'https://rubygems.org'
```

```
gem 'rails', '3.2.9'
gem 'mysql2'
```

```
group :assets do
 gem 'sass-rails', '~> 3.2.3'
 gem 'coffee-rails', '~> 3.2.1'
 gem 'uglifier', '>= 1.0.3'
end

gem 'jquery-rails'
gem 'sidekiq'
gem 'uuidtools'
```

Install the bundle.

```
$ bundle install
```

In `app/workers`, create a worker for Sidekiq to carry out its tasks:

```
$ touch app/workers/thing_worker.rb
```

```
class ThingWorker
 include Sidekiq::Worker

 def perform(count)
 count.times do
 thing_uuid = UUIDTools::UUID.random_create.to_s
 Thing.create(:title => "New Thing (#{$thing_uuid})", :description =>
 "Description for thing #{$thing_uuid}")
 end
 end
end
```

This worker creates `n` number of things, where `n` is the value passed to the worker.

Create a controller for “Things”:

```
$ rails g controller Thing
```

```
class ThingController < ApplicationController

 def new
 ThingWorker.perform_async(2)
 redirect_to '/thing'
 end

 def index
 @things = Thing.all
 end
end
```

Add a view to inspect our collection of “Things”:

```
$ mkdir app/views/things
$ touch app/views/things/index.html.erb
```

```
nil
```

## Deploy the App

This app needs to be deployed twice for it to work, once as a Rails web app and once as a standalone Ruby app. One way to do this is to keep separate

Cloud Foundry manifests for each app type:

Web Manifest: Save this as `web-manifest.yml`:

```
--
applications:
- name: sidekiq
 memory: 256M
 instances: 1
 host: sidekiq
 domain: ${target-base}
 path: .
 services:
 - sidekiq-mysql:
 - sidekiq-redis:
```

Worker Manifest: Save this as `worker-manifest.yml`:

```
--
applications:
- name: sidekiq-worker
 memory: 256M
 instances: 1
 path: .
 command: bundle exec sidekiq
 no-route: true
 services:
 - sidekiq-redis:
 - sidekiq-mysql:
```

Since the url “sidekiq.cloudfoundry.com” is probably already taken, change it in `web-manifest.yml` first, then push the app with both manifest files:

```
$ cf push -f web-manifest.yml
$ cf push -f worker-manifest.yml
```

If the cf CLI asks for a URL for the worker app, select **none**.

## Test the App

Test the app by visiting the new action on the “Thing” controller at the assigned url. In this example, the URL would be <http://sidekiq.cloudfoundry.com/thing/new>

This creates a new Sidekiq job which is queued in Redis, then picked up by the worker app. The browser is then redirected to `/thing` which shows the collection of “Things”.

## Scale Workers

Use the `cf scale` command to change the number of Sidekiq workers.

Example:

```
$ cf scale sidekiq-worker -i 2
```

## Use `rails_serve_static_assets` on Rails 4

By default Rails 4 returns a 404 if an asset is not handled via an external proxy such as Nginx. The `rails_serve_static_assets` gem enables your Rails server to deliver static assets directly, instead of returning a 404. You can use this capability to populate an edge cache CDN or serve files directly from your web app. The gem enables this behavior by setting the `config.serve_static_assets` option to `true`, so you do not need to configure it manually.

## Add Custom Libraries

If your app requires external shared libraries that are not provided by the rootfs or the buildpack, you must place the libraries in an `ld_library_path` directory at the app root.

 **Note:** You must keep these libraries up-to-date. They do not update automatically.

The Ruby buildpack automatically adds the directory `<app-root>/ld_library_path` to `LD_LIBRARY_PATH` so that your app can access these libraries at runtime.

## Environment Variables

You can access environments variable programmatically. For example, you can obtain `VCAP_SERVICES` as follows:

```
ENV['VCAP_SERVICES']
```

Environment variables available to you include both those defined by the system and those defined by the Ruby buildpack, as described below. For more information about system environment variables, see the [Application-Specific System Variables](#) section of the *Cloud Foundry Environment Variables* topic.

### BUNDLE\_BIN\_PATH

Location where Bundler installs binaries.

Example: `BUNDLE_BIN_PATH:/home/vcap/app/vendor/bundle/ruby/1.9.1/gems/bundler-1.3.2/bin/bundle`

### BUNDLE\_GEMFILE

Path to app Gemfile.

Example: `BUNDLE_GEMFILE:/home/vcap/app/Gemfile`

### BUNDLE\_WITHOUT

The `BUNDLE_WITHOUT` environment variable instructs Cloud Foundry to skip gem installation in excluded groups.

Use this with Rails applications, where “assets” and “development” gem groups typically contain gems that are not needed when the app runs in production.

Example: `BUNDLE_WITHOUT=assets`

### DATABASE\_URL

Cloud Foundry examines the `database_uri` for bound services to see if they match known database types. If known relational database services are bound to the app, the `DATABASE_URL` environment variable is set using the first match in the list.

If your app depends on `DATABASE_URL` to be set to the connection string for your service and Cloud Foundry does not set it, use the `cf set-env` command to can set this variable manually.

Example:

```
$ cf set-env my-app-name DATABASE_URL mysql://example-database-connection-string
```

### GEM\_HOME

Location where gems are installed.

Example: `GEM_HOME:/home/vcap/app/vendor/bundle/ruby/1.9.1`

## GEM\_PATH

Location where gems can be found.

Example: `GEM_PATH=/home/vcap/app/vendor/bundle/ruby/1.9.1:`

## RACK\_ENV

This variable specifies the Rack deployment environment. Valid values are `development`, `deployment`, and `none`. This governs which middleware is loaded to run the app.

Example: `RACK_ENV=development`

## RAILS\_ENV

This variable specifies the Rails deployment environment. Valid values are `development`, `test`, and `production`. This controls which of the environment-specific configuration files governs how the app is executed.

Example: `RAILS_ENV=production`

## RUBYOPT

This Ruby environment variable defines command-line options passed to Ruby interpreter.

Example: `RUBYOPT: -I/home/vcap/app/vendor/bundle/ruby/1.9.1/gems/bundler-1.3.2/lib -rbundler/setup`

## Getting Started Deploying Ruby Apps

This topic provides links to additional information about getting started deploying apps using the Ruby buildpack. See the following topics for deployment guides specific to your app language and framework:

- [Ruby](#)
- [Ruby on Rails](#)

## Getting Started Deploying Ruby Apps

This guide is intended to walk you through deploying a Ruby app to Pivotal Application Service (PAS). If you experience a problem following the steps below, check the [Known Issues](#) topic, or refer to the [Troubleshooting Application Deployment and Health](#) topic.

### Sample App Step

If you want to go through this tutorial using the sample app, run `git clone https://github.com/cloudfoundry-samples/pong_matcher_ruby.git` to clone the `pong_matcher_ruby` app from GitHub, and follow the instructions in the Sample App Step sections.

 **Note:** Ensure that your Ruby app runs locally before continuing with this procedure.

## Deploy a Ruby Application

This section describes how to deploy a Ruby application to PAS, and uses output from a sample app to show specific steps of the deployment process.

### Prerequisites

- A Ruby 2.x application that runs locally on your workstation
- [Bundler](#) configured on your workstation
- Basic to intermediate Ruby knowledge
- The [Cloud Foundry Command Line Interface \(cf CLI\)](#) installed on your workstation

### Step 1: Create and Bind a Service Instance for a Ruby Application

This section describes using the cf CLI to configure a Redis Cloud managed service instance for an app. You can use either the CLI or [Apps Manager](#) to perform this task.

PAS supports two types of service instances:

- Managed services integrate with PAS through service brokers that offer services and plans and manage the service calls between PAS and a service provider.
- User-provided service instances enable you to connect your application to pre-provisioned external service instances.

For more information about creating and using service instances, refer to the [Services Overview](#) topic.

### Create a Service Instance

Run `cf marketplace` to view managed and user-provided services and plans that are available to you.

The example shows three of the available managed database-as-a-service providers and the plans that they offer: `cleardb MySQL` and `postgresql-10-odb PostgreSQL` as a Service.

```
$ cf marketplace
Getting services from marketplace in org Cloud-Apps / space development as clouduser@example.com...
OK

service plans description
...
cleardb spark, boost, amp, shock Highly available MySQL for your Apps
...
postgresql-10-odb standalone, standalone-replica, general PostgreSQL as a Service
...
```

Run `cf create-service SERVICE PLAN SERVICE_INSTANCE` to create a service instance for your app. Choose a SERVICE and PLAN from the list, and provide a

unique name for the SERVICE\_INSTANCE.

#### Sample App Step

Run `cf create-service rediscloud 30mb redis`. This creates a service instance named `redis` that uses the `rediscloud` service and the `30mb` plan, as the example below shows.

```
$ cf create-service rediscloud 30mb redis
Creating service redis in org Cloud-Apps / space development as cloudunder@example.com....
OK
```

## Bind a Service Instance

When you bind an app to a service instance, PAS writes information about the service instance to the VCAP\_SERVICES app environment variable. The app can use this information to integrate with the service instance.

Most services support bindable service instances. Refer to your service provider's documentation to confirm if they support this functionality.

You can bind a service to an application with the command `cf bind-service APPLICATION SERVICE_INSTANCE`.

Alternately, you can configure the deployment manifest file by adding a `services` block to the `applications` block and specifying the service instance. For more information and an example on service binding using a manifest, see the Sample App Step.

You can also bind a service using the [Apps Manager](#).

#### Sample App Step

You can skip this step. The manifest for the sample app contains a `services` sub-block in the `applications` block, as the example below shows. This binds the `redis` service instance that you created in the previous step.

```
services:
 - redis
```

## Step 2: Configure Deployment Options

### Configure the Deployment Manifest

You can specify app deployment options in a manifest that the `cf push` command uses. For more information about application manifests and supported attributes, refer to the [Deploying with Application Manifests](#) topic.

### Configure a Production Server

PAS uses the default standard Ruby web server library, WEBrick, for Ruby and RoR apps. However, PAS can support a more robust production web server, such as Phusion Passenger, Puma, Thin, or Unicorn. If your app requires a more robust web server, refer to the [Configuring a Production Server](#) topic for help configuring a server other than WEBrick.

#### Sample App Step

You can skip this step. The `manifest.yml` file for `pong_matcher_ruby` does not require any additional configuration to deploy the app.

## Step 3: Log in and Target the API Endpoint

Run `cf login -a API_ENDPOINT`, enter your login credentials, and select a space and org. The API endpoint is [the URL of the Cloud Controller in your PAS instance](#).

### Sample App Step

You must do this step to run the sample app.

## Step 4: Deploy an App

 **Note:** You must use the cf CLI to deploy apps.

From the root directory of your application, run `cf push APP_NAME` to deploy your application.

`cf push APP_NAME` creates a URL route to your application in the form HOST.DOMAIN, where HOST is your APP\_NAME and DOMAIN is specified by your administrator. Your DOMAIN is `shared-domain.example.com`. For example: `cf push my-app` creates the URL `my-app.shared-domain.example.com`.

The URL for your app must be unique from other apps that PAS hosts or the push will fail. Use the following options to help create a unique URL:

- `-n` to assign a different HOST name for the app.
- `--random-route` to create a URL that includes the app name and random words.
- `cf help push` to view other options for this command.

If you want to view log activity while the app deploys, launch a new terminal window and run `cf logs APP_NAME`.

Once your app deploys, browse to your app URL. Search for the `urls` field in the `App started` block in the output of the `cf push` command. Use the URL to access your app online.

### Sample App Step

Run `cf push pong_matcher_ruby -n HOST_NAME`.

Example: `cf push pong_matcher_ruby -n pongmatch-ex12`

The example below shows the terminal output of deploying the `pong_matcher_ruby` app. `cf push` uses the instructions in the manifest file to create the app, create and bind the route, and upload the app. It then binds the app to the `redis` service and follows the instructions in the manifest to start one instance of the app with 256M. After the app starts, the output displays the health and status of the app.

 **Note:** The `pong_matcher_ruby` app does not include a web interface. To interact with the `pong_matcher_ruby` app, see the interaction instructions on GitHub: [https://github.com/cloudfoundry-samples/pong\\_matcher\\_ruby](https://github.com/cloudfoundry-samples/pong_matcher_ruby).

```
$ cf push pong_matcher_ruby -n pongmatch-ex12
Using manifest file /Users/clouduser/workspace/pong_matcher_ruby/manifest.yml

Creating app pong_matcher_ruby in org Cloud-Apps / space development as clouduser@example.com...
OK

Creating route pongmatch-ex12.shared-domain.example.com
 Binding pongmatch-ex12.shared-domain.example.com to pong_matcher_ruby...
OK

Uploading pong_matcher_ruby...
Uploading app files from: /Users/clouduser/workspace/pong_matcher_ruby
Uploading 8.8K, 12 files
OK
Binding service redis to app pong_matcher_ruby in org Cloud-Apps / space development as clouduser@example.com...
OK

Starting app pong_matcher_ruby in org Cloud-Apps / space development as clouduser@example.com...
OK
...
0 of 1 instances running, 1 starting
1 of 1 instances running

App started

Showing health and status for app pong_matcher_ruby in org Cloud-Apps / space development as clouduser@example.com...
OK

requested state: started
instances: 1/1
usage: 256M x 1 instances
urls: pongmatch-ex12.cfapps.io

 state since cpu memory disk
#0 running 2014-12-09 10:04:40 AM 0.0% 35.2M of 256M 45.8M of 1G
```

## Step 5: Test a Deployed App

You've deployed an app to PAS!

Use the cf CLI or [Apps Manager](#) to review information and administer your app and your PAS account. For example, you could edit the `manifest.yml` to increase the number of app instances from 1 to 3, and redeploy the app with a new app name and host name.

See the [Manage Your Application with the cf CLI](#) section for more information. See also [Using the Apps Manager](#).

## Manage Your Application with the cf CLI

Run `cf help` to view a complete list of commands, grouped by task categories, and run `cf help COMMAND` for detailed information about a specific command.

For more information about using the cf CLI, refer to the Cloud Foundry Command Line Interface (cf CLI) topics, especially the [Getting Started with cf CLI](#) topic.

**Note:** You cannot perform certain tasks in the CLI or [Apps Manager](#) because these are commands that only a PAS administrator can run. If you are not a PAS administrator, the following message displays for these types of commands:  
error code: 10003, message: You are not authorized to perform the requested action  
For more information about specific Admin commands you can perform with the Apps Manager, depending on your user role, refer to the [Getting Started with the Apps Manager](#) topic.

## Troubleshooting

If your application fails to start, verify that the application starts in your local environment. Refer to the [Troubleshooting Application Deployment and Health](#) topic to learn more about troubleshooting.

## App Deploy Fails

Even when deploying an app fails, the app might exist on PAS. Run `cf apps` to review the apps in the currently targeted org and space. You might be able to correct the issue using the CLI or [Apps Manager](#), or you might have to delete the app and redeploy.

Common reasons deploying an app fails include:

- You did not successfully create and bind a needed service instance to the app, such as a PostgreSQL service instance. Refer to Step 2: Create and Bind a Service Instance for a Ruby Application.
- You did not successfully create a unique URL for the app. Refer to the troubleshooting tip [App Requires Unique URL](#).

## App Requires Unique URL

PAS requires that each app that you deploy has a unique URL. Otherwise, the new app URL collides with an existing app URL and PAS cannot successfully deploy the app. You can resolve this issue by running `cf push` with either of the following flags to create a unique URL:

- `-n` to assign a different HOST name for the app.
- `--random-route` to create a URL that includes the app name and random words. Using this option might create a long URL, depending on the number of words that the app name includes.

## Getting Started Deploying Ruby on Rails Apps

This guide walks you through deploying a Ruby on Rails app to Pivotal Cloud Foundry (PCF).

### Prerequisites

In order to deploy a sample Ruby on Rails app, you must have the following:

- A Cloud Foundry deployment or a Pivotal Web Services account
- The [Cloud Foundry Command Line Interface](#)
- A Cloud Foundry username and password with [Space Developer permissions](#). See your [Org Manager](#) if you require permissions.

### Step 1: Clone the App

Run the following terminal command to create a local copy of the cf-sample-app-rails.

```
$ git clone https://github.com/cloudfoundry-samples/cf-sample-app-rails.git
```

The newly created directory contains a `manifest.yml` file, which assists CF with deploying the app. See [Deploying with Application Manifests](#) for more information.

### Step 2: Log in and Target the API Endpoint

1. Run the following terminal command to log in and target the API endpoint of your deployment. For more information, see the [Identifying the API Endpoint for your PAS Instance](#) topic.

```
$ cf login -a YOUR-API-ENDPOINT
```

2. Use your credentials to log in, and to select a [Space and Org](#).

### Step 3: Create a Service Instance

Run the following terminal command to create a PostgreSQL service instance for the sample app. Our service instance is `rails-postgres`. It uses the `postgresql-10-odb` service and the `standalone` plan. For more information about the `postgresql-10-odb` service, see [Crunchy PostgreSQL](#).

```
$ cf create-service postgresql-10-odb standalone rails-postgres
Creating service rails-postgres in org YOUR-ORG / space development as clouduser@example.com....
OK
```

### Step 4: Deploy the App

Make sure you are in the `cf-sample-app-rails` directory. Run the following terminal command to deploy the app:

```
$ cf push cf-sample-app-rails
```

`cf push cf-sample-app-rails` creates a URL route to your application in the form HOST.DOMAIN. In this example, HOST is `cf-sample-app-rails`. Administrators specify the DOMAIN. For example, for the DOMAIN `shared-domain.example.com`, running `cf push cf-sample-app-rails` creates the URL `cf-sample-app-rails.shared-domain.example.com`.

The example below shows the terminal output when deploying the `cf-sample-app-rails`. `cf push` uses the instructions in the manifest file to create the app, create and bind the route, and upload the app. It then follows the information in the manifest to start one instance of the app with 256M of RAM. After the

app starts, the output displays the health and status of the app.

```
$ cf push cf-sample-app-rails
Using manifest file ~/workspace/cf-sample-app-rails/manifest.yml

Creating app cf-sample-app-rails in org my-org / space dev as cloudunder@example.com...
OK

Creating route cf-sample-app-rails.cfapps.io...
OK

Binding cf-sample-app-rails.cfapps.io to cf-sample-app-rails...
OK

Uploading cf-sample-app-rails...
Uploading app files from: ~/workspace/cf-sample-app-rails
Uploading 746.6K, 136 files
Done uploading
OK

Starting app cf-sample-app-rails in org my-org / space dev as cloudunder@example.com...
...
0 of 1 instances running, 1 starting
1 of 1 instances running

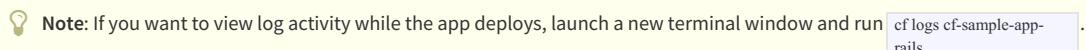
App started
OK

App cf-sample-app-rails was started using this command 'bundle exec rails server -p $PORT'

Showing health and status for app cf-sample-app-rails in org my-org / space dev as cloudunder@example.com...
OK

requested state: started
instances: 1/1
usage: 512M x 1 instances
urls: cf-sample-app-rails.cfapps.io
last uploaded: Fri Dec 22 18:08:32 UTC 2017
stack: cflinuxfs3
buildpack: ruby

 state since cpu memory disk details
#0 running 2018-8-17 10:09:57 AM 0.0% 20.7M of 512M 186.8M of 1G
```

 **Note:** If you want to view log activity while the app deploys, launch a new terminal window and run `cf logs cf-sample-app-rails`.

## Step 5: Bind the Service Instance

- Run the command below to bind the service instance to the sample app. Once bound, environment variables are stored that allow the app to connect to the service after a `cf push`, `cf restage`, or `cf restart` command.

```
$ cf bind-service cf-sample-app-rails rails-postgres
Binding service rails-postgres to app cf-sample-app-rails in org my-org / space dev
OK
TIP: Use 'cf restage cf-sample-app-rails' to ensure your env variable changes take effect
```

- Run the following command to restage the sample app.

```
$ cf restage cf-sample-app-rails
```

- Run the following command to verify the service instance is bound to the sample app.

```
$ cf services
Getting services in org my-org / space dev
OK
name service plan bound apps last operation
rails-postgres postgresql-10-odb standalone cf-sample-app-rails create succeeded
```

## Step 6: Verify the App

Verify that the app is running by browsing to the URL generated in the output of the previous step. In this example, navigating to `cf-sample-app-rails.shared-domain.example.com` verifies that the app is running.

You've now pushed an app to PAS! For more information about this topic, see the [Pushing an App](#) topic.

## What to Do Next

You have deployed an app to PCF. Consult the sections below for information about what to do next.

### Test a Deployed App

Use the cf CLI or [Apps Manager](#) to review information and administer your app and your PCF account. For example, you could edit the `manifest.yml` file to increase the number of app instances from 1 to 3 or redeploy the app with a new app name.

### Manage Your App with the cf CLI

Run `cf help` to view a complete list of commands and run `cf help COMMAND` for detailed information about a specific command. For more information about using the cf CLI, refer to the cf CLI topics, especially the [Getting Started with cf CLI](#) topic.

### Troubleshooting

If your app fails to start, verify that the app starts in your local environment. Refer to the [Troubleshooting Application Deployment and Health](#) topic to learn more about troubleshooting.

## Configure Rake Tasks for Deployed Apps

For PAS to automatically invoke a Rake task while a Ruby or Ruby on Rails app is deployed, you must do the following:

- Include the Rake task in your app
- Configure the application start command using the `command` attribute in the application manifest

The following is an example that shows how to invoke a Rake database migration task at application startup.

1. Create a file with the Rake task name and the extension `.rake`, and store it in the `lib/tasks` directory of your application.
2. Add the following code to your rake file:

```
namespace :cf do
 desc "Only run on the first application instance"
 task :on_first_instance do
 instance_index = JSON.parse(ENV["VCAP_APPLICATION"])["instance_index"] rescue nil
 exit(0) unless instance_index == 0
 end
end
```

This Rake task limits an idempotent command to the first instance of a deployed application.

3. Add the task to the `manifest.yml` file with the `command` attribute, referencing the idempotent command `rake db:migrate` chained with a start command.

```
applications:
- name: my-rails-app
 command: bundle exec rake cf:on_first_instance db:migrate && rails s -p $PORT
```

## Environment Variables Defined by the Ruby Buildpack

Pivotal Application Service (PAS) provides configuration information to apps through environment variables. This topic describes the additional environment variables provided by the Ruby buildpack.

For more information about the standard environment variables provided by Pivotal Application Service (PAS), see the [Cloud Foundry Environment Variables](#) topic.

## Ruby Buildpack Environment Variables

The following table describes the environment variables provided by the Ruby buildpack.

Environment Variable	Description
BUNDLE_BIN_PATH	The directory where Bundler installs binaries. Example: <pre>BUNDLE_BIN_PATH:/home/vcap/app/vendor/bundle/ruby/1.9.1/gems/bundler-1.3.2/bin/bundle</pre>
BUNDLE_GEMFILE	The path to the Gemfile for the app. Example: <pre>BUNDLE_GEMFILE:/home/vcap/app/Gemfile</pre>
BUNDLE_WITHOUT	Instructs Cloud Foundry to skip gem installation in excluded groups. Use this with Rails applications, where “assets” and “development” gem groups typically contain gems that are not needed when the app runs in production. Example: <pre>BUNDLE_WITHOUT=assets</pre>
DATABASE_URL	Cloud Foundry examines the <code>database_uri</code> for bound services to see if they match known database types. If known relational database services are bound to the app, then the <code>DATABASE_URL</code> environment variable is set to the first services in the list.  If your application requires that <code>DATABASE_URL</code> is set to the connection string for your service, and Cloud Foundry does not set it, use the Cloud Foundry Command Line Interface (cf CLI) <code>cf set-env</code> command to set this variable manually. Example:  <pre>\$ cf set-env my-app DATABASE_URL mysql://example-database-connection-string</pre>
GEM_HOME	The directory where gems are installed. Example: <pre>GEM_HOME:/home/vcap/app/vendor/bundle/ruby/1.9.1</pre>
GEM_PATH	The directory where gems can be found. Example: <pre>GEM_PATH=/home/vcap/app/vendor/bundle/ruby/1.9.1:</pre>
RACK_ENV	The Rack deployment environment, which governs the middleware loaded to run the app. Valid value are <code>development</code> , <code>deployment</code> , and <code>none</code> . Example: <pre>RACK_ENV=none</pre>
RAILS_ENV	The Rails deployment environment, which controls which environment-specific configuration file governs how the app is executed. Valid value are <code>development</code> , <code>test</code> , and <code>production</code> . Example: <pre>RAILS_ENV=production</pre>
RUBYOPT	Defines command-line options passed to Ruby interpreter. Example: <pre>RUBYOPT: -I/home/vcap/app/vendor/bundle/ruby/1.9.1/gems/bundler-1.3.2/lib -rbundler/setup</pre>

## Configure Service Connections for Ruby

After you create a service instance and bind it to an application, you must configure the application to connect to the service.

### Query VCAP\_SERVICES with cf-app-utils

The `cf-app-utils` gem allows your application to search for credentials from the `VCAP_SERVICES` environment variable by name, tag, or label.

- [cf-app-utils-ruby ↗](#)

### VCAP\_SERVICES defines DATABASE\_URL

At runtime, Cloud Foundry creates a `DATABASE_URL` environment variable for every application based on the `VCAP_SERVICES` environment variable.

Example `VCAP_SERVICES`:

```
VCAP_SERVICES =
{
 "postgresql-10-odb": [
 {
 "name": "postgresql-10-odb-c6c60",
 "label": "postgresql-10-odb",
 "credentials": {
 "uri": "postgres://exampleuser:examplepass@babar.postgresql-10-odb.com:5432/exampledbs"
 }
 }
]
}
```

Based on this `VCAP_SERVICES`, Cloud Foundry creates the following `DATABASE_URL` environment variable:

```
DATABASE_URL = postgres://exampleuser:examplepass@babar.postgresql-10-odb.com:5432/exampledbs
```

Cloud Foundry uses the structure of the `VCAP_SERVICES` environment variable to populate `DATABASE_URL`. Any service containing a JSON object with the following form will be recognized by Cloud Foundry as a candidate for `DATABASE_URL`:

```
{
 "some-service": [
 {
 "credentials": {
 "uri": "<some database URL>"
 }
 }
]
}
```

Cloud Foundry uses the first candidate found to populate `DATABASE_URL`.

## Configure Non-Rails Applications

Non-Rails applications can also access the `DATABASE_URL` variable.

If you have more than one service with credentials, only the first will be populated into `DATABASE_URL`. To access other credentials, you can inspect the `VCAP_SERVICES` environment variable.

```
vcap_services = JSON.parse(ENV['VCAP_SERVICES'])
```

Use the hash key for the service to obtain the connection credentials from `VCAP_SERVICES`.

- For services that use the [v2 format ↗](#), the hash key is the name of the service.

- For services that use the [v1 format](#), the hash key is formed by combining the service provider and version, in the format PROVIDER-VERSION.

For example, the service provider “p-mysql” with version “n/a” forms the hash key `p-mysql-n/a`.

## Seed or Migrate Database

Before you can use your database the first time, you must create and populate or migrate it. For more information about migrating database schemas, see [Services Overview](#).

## Troubleshooting

To aid in troubleshooting issues connecting to your service, you can examine the environment variables and log messages Cloud Foundry records for your application.

## View Environment Variables

Use the `cf env` command to view the Cloud Foundry environment variables for your application. `cf env` displays the following environment variables:

- The `VCAP_SERVICES` variables existing in the container environment
- The user-provided variables set using the `cf set-env` command

```
$ cf env my-app
Getting env variables for app my-app in org My-Org / space development as admin...
OK

System-Provided:
{
 "VCAP_SERVICES": {
 "p-mysql-n/a": [
 {
 "credentials": {
 "uri": "postgres://lrra:e6B-X@p-mysql/provider.example.com:5432/lraa",
 "label": "p-mysql-n/a",
 "name": "p-mysql",
 "syslog_drain_url": "",
 "tags": ["postgres", "postgresql", "relational"]
 }
]
 }
 }
}

User-Provided:
my-env-var: 100
my-drain: http://drain.example.com
```

## View Logs

Use the `cf logs` command to view the Cloud Foundry log messages for your application. You can direct current logging to standard output, or you can dump the most recent logs to standard output.

Run `cf logs APPNAME` to direct current logging to standard output:

```
$ cf logs my-app
Connected, tailing logs for app my-app in org My-Org / space development as admin...
1:27:19.72 [App/0] OUT [CONTAINER] org.apache.coyote.http11.Http11Protocol INFO Starting ProtocolHandler ["http-bio-61013"]
1:27:19.77 [App/0] OUT [CONTAINER] org.apache.catalina.startup.Catalina INFO Server startup in 10427 ms
```

Run `cf logs APPNAME --recent` to dump the most recent logs to standard output:

```
$ cf logs my-app --recent
Connected, dumping recent logs for app my-app in org My-Org / space development as admin...
1:27:15.93 [App/0] OUT 15,935 INFO EmbeddedDatabaseFactory:124 - Creating embedded database 'SkyNet'
1:27:16.31 [App/0] OUT 16,318 INFO LocalEntityManagerFactory:287 - Building TM container EntityManagerFactory for unit 'default'
1:27:16.50 [App/0] OUT 16,505 INFO Version:37 - HCANN001: Hibernate Commons Annotations {4.0.1.Final}
1:27:16.51 [App/0] OUT 16,517 INFO Version:41 - HHH412: Hibernate Core {4.1.9.Final}
1:27:16.95 [App/0] OUT 16,957 INFO SkyNet-Online:73 - HHH268: Transaction strategy: org.hibernate.internal.TransactionFactory
1:27:16.96 [App/0] OUT 16,963 INFO InitiateTerminatorT1000Deployment:48 - HHH000397: Using TranslatorFactory
1:27:17.02 [App/0] OUT 17,026 INFO Version:27 - HV001: Hibernate Validator 4.3.0.Final
```

If you encounter the error, “A fatal error has occurred. Please see the Bundler troubleshooting documentation,” update your version of bundler and run `bundle install`.

```
$ gem update bundler
$ gem update --system
$ bundle install
```

## Support for Windows Gemfiles

This topic describes how the Ruby buildpack handles dependencies on Windows machines.

### Windows Gemfiles

When a `Gemfile.lock` is generated on a Windows machine, it often contains gems with Windows-specific versions. This results in versions of gems such as `mysql2`, `thin`, and `pg` containing “`-x86-mingw32`.” For example, the `Gemfile` may contain the following:

```
gem 'sinatra'
gem 'mysql2'
gem 'json'
```

When you run `bundle install` with the above `Gemfile` on a Windows machine, it results in the following `Gemfile.lock`:

```
GEM remote: http://rubygems.org/
specs:
 json (1.7.3)
 mysql2 (0.3.11-x86-mingw32)
 rack (1.4.1)
 rack-protection (1.2.0)
 rack sinatra (1.3.2)
 rack (~> 1.3, >= 1.3.6)
 rack-protection (~> 1.2)
 tilt (~> 1.3, >= 1.3.3)
 tilt (1.3.3)
PLATFORMS x86-mingw32
DEPENDENCIES
 json
 mysql2
 sinatra
```

Notice the “`-x86-mingw32`” in the version number of `mysql2`. Since Cloud Foundry runs on Linux machines, this would fail to install. To mitigate this, the Ruby Buildpack removes the `Gemfile.lock` and uses Bundler to resolve dependencies from the `Gemfile`.

 **Note:** Removing the `Gemfile.lock` will cause dependency versions to be resolved from the `Gemfile`. This could result in different versions being installed than those listed in the `Gemfile.lock`.

## Staticfile Buildpack

### Overview

This topic describes how to configure and use the Staticfile buildpack.

 **Note:** BOSH configured custom trusted certificates [are not supported by the Staticfile buildpack.](#)

### Definitions

**Staticfile app:** An app or content that requires no backend code other than the NGINX webserver, which the buildpack provides. Examples of staticfile apps are front-end JavaScript apps, static HTML content, and HTML/JavaScript forms.

**Staticfile buildpack:** The buildpack that provides runtime support for staticfile apps and apps with backends hosted elsewhere. To find which version of NGINX the current Staticfile buildpack uses, see the [Staticfile buildpack release notes](#).

### Staticfile Detection

If you create a file named `Staticfile` in the root directory of your app, Cloud Foundry automatically uses the Staticfile buildpack when you push your app.

The `Staticfile` file can be an empty file, or it can contain configuration settings for your app. For more information, see [Configuring the Buildpack](#) and [Pushing an App](#).

### Memory Usage

NGINX requires 20 MB of RAM to serve static assets. When using the Staticfile buildpack, we recommend pushing apps with the `-m 64M` option to reduce RAM allocation from the default 1 GB allocated to containers by default.

### Configure the Buildpack

This section describes configuration options available for the Staticfile buildpack. If you need to make configuration changes to NGINX that are not listed in the table below, use the NGINX Buildpack instead of the Staticfile buildpack. For more information, see [NGINX Buildpack](#).

### Staticfile File Configuration

To configure these options, add the configuration property as a new line in your `Staticfile` file as described in the following table.

Staticfile Configuration Property	Description
<code>root: YOUR-DIRECTORY-NAME</code>  Example: <code>root: public</code>	<b>Alternative root:</b> Specifies a root directory other than the default. Use this option to serve <code>index.html</code> and other assets, such as HTML, CSS, or JavaScript files, from a location other than the root directory.  For example, you can specify an alternate folder such as <code>dist</code> or <code>public</code> .
<code>directory: visible</code>	<b>Directory list:</b> Displays an HTML page that shows a directory index for your site. A sample of a directory list is shown below.

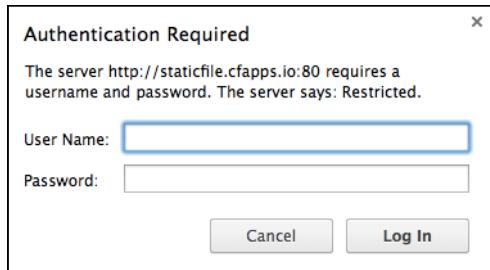
<h2 style="text-align: center;">Index of /</h2>	
<pre>ssi: enabled</pre>	<p><b>SSI</b>: Enables Server Side Includes (SSI), which allow you to embed the contents of one or more files into a web page on a web server. For general information about SSI, see the <a href="#">Server Side Includes</a> entry on Wikipedia.</p>
<pre>pushstate: enabled</pre>	<p><b>Pushstate routing</b>: Keeps browser-visible URLs clean for client-side JavaScript apps that serve multiple routes. For example, pushstate routing allows a single JavaScript file to route to multiple anchor-tagged URLs that look like <code>/some/path1</code> instead of <code>/some#path1</code>.</p>
<pre>gzip: off</pre>	<p><b>GZip file serving and compression</b>: Disables <a href="#">gzip_static</a> and <a href="#">gunzip</a> modules, which are enabled by default. These modules allow NGINX to serve files stored in compressed GZ format and to uncompress them for clients that do not support compressed content or responses.</p> <p>You may want to disable compression under particular circumstances such as serving content to very old browser clients.</p>
<pre>http_proxy: HTTP-URL https_proxy: HTTPS-URL</pre> <p>Example:</p> <pre>http_proxy: http://www.example.com/ https_proxy: https://www.example.com/</pre>	<p><b>Proxy support</b>: Allows you to use a proxy when downloading dependencies during the staging of your app.</p>
<pre>force_https: true</pre> <p>Alternatively, set the <code>FORCE_HTTPS</code> environment variable to <code>true</code>.</p>	<p><b>Force HTTPS</b>: Forces all requests to be sent through HTTPS. This redirects non-HTTPS requests as HTTPS requests.</p> <div style="border: 1px solid #ccc; padding: 10px;"> <p><b>Note:</b> Do not enable <code>FORCE_HTTPS</code> if you have a proxy server or load balancer that terminates SSL/TLS. Doing so can cause infinite redirect loops, for example, if you use <a href="#">Flexible SSL</a> with CloudFlare.</p> </div>
<pre>host_dot_files: true</pre>	<p><b>Dot files</b>: By default, hidden files, which are those starting with a <code>.</code>, are not served by this buildpack.</p>
<pre>status_codes:   404: /404.html   500: /500.html</pre>	<p><b>Status Codes</b>: Allows to define custom pages for HTTP STATUS codes instead of the default NGINX pages.</p>
<pre>http_strict_transport_security: true</pre>	<p><b>HTTP Strict Transport Security (HSTS)</b>: Causes NGINX to respond to all requests with the header <code>Strict-Transport-Security: max-age=31536000</code>. This forces receiving browsers to make all subsequent requests over HTTPS. This setting defaults to a <code>max-age</code> of one year.</p> <div style="border: 1px solid #ccc; padding: 10px;"> <p><b>Note:</b> Because this setting persists in browsers for a long time, only enable this setting after you ensure that you have completed your app configuration.</p> </div>
<pre>http_strict_transport_security_include_subdomains: true</pre>	<p><b>HSTS includes subdomains</b>: Causes NGINX to respond to all requests with the following header: <code>Strict-Transport-Security: max-age=31536000; includeSubDomains</code>. This forces browsers to make all subsequent requests over HTTPS including subdomains. This setting defaults to a <code>max-age</code> of one year.</p> <div style="border: 1px solid #ccc; padding: 10px;"> <p><b>Note:</b> Setting this property to <code>true</code> also makes <code>http_strict_transport_security</code> default to true.</p> </div>
<pre>http_strict_transport_security_preload:   true</pre>	<p><b>HSTS preload</b>: Causes NGINX to respond to all requests with the following header: <code>Strict-Transport-Security: max-age=31536000; includeSubDomains; preload</code>. This forces browsers to make all subsequent requests over HTTPS including subdomains and requests inclusion in browser-managed HSTS preload lists. For more information, see <a href="https://hstspreload.org">https://hstspreload.org</a>. This setting defaults to a <code>max-age</code> of one year. This setting defaults to a <code>max-age</code> of one year.</p> <div style="border: 1px solid #ccc; padding: 10px;"> <p><b>Note:</b> Setting this property to <code>true</code> also makes <code>http_strict_transport_security</code> and <code>http_strict_transport_security_include_subdomains</code> default to true.</p> </div>

## Other Configuration

Follow the instructions below to make other configuration changes.

### Basic Authentication

Allows you to enable basic authentication for your app or website. A sample basic authentication dialog box is shown below:



You can create a hashed username and password pair for each user by using a site like [Htpasswd Generator](#).

#### Configuration

Add a new `Staticfile.auth` file to the root or alternative root directory. In the new file, add one or more username and password entries using the following format: `USERNAME:HASHED_PASSWORD`

Example:

```
pat:$example1$DuUQEQp8$ZccZCHQEINSjrgerwSFC0
stevie:$example1$22derfaecZSJRW4rKE$KKEWKS
```

### Custom Location

Allows you to specify custom location definitions with additional directives. For information about NGINX directives, see [Creating NGINX Plus and NGINX Configuration Files](#) and [Alphabetical index of directives](#) in the *NGINX documentation*.

#### Configuration

To customize the `location` block of the NGINX configuration file, follow the steps below.

1. Set an alternative `root` directory. The `location_include` property only works in conjunction with an alternative `root`.
2. Create a file with location-scoped NGINX directives. See the following example, which causes visitors of your site to receive the `X-MySiteName` HTTP header:
  - File: `nginx/conf/includes/custom_header.conf`
  - Content: `add_header X-MySiteName BestSiteEver;`
3. Set the `location_include` variable in your `Staticfile` to the path of the file from the previous step. This path is relative to `nginx/conf`.

Example:

```
...
root: public
location_include: includes/*.conf
...
```

### Additional MIME Type Support

Allows you to configure additional MIME types on your NGINX server.

#### Configuration

To add MIME types, add a `mime.types` file to your root folder, or to the alternate root folder if you specified one.

For more information about the `mime.types` file, see [mime.types](#) in the *NGINX documentation*.

Example MIME types file:

```
types {
 text/html html htm shtml;
 text/css css;
 text/xml xml rss;
 image/gif gif;
 ...
}
```

## Push an App

To push your app, you can use either the system Staticfile buildpack or specify a Staticfile buildpack.

### Option 1: Use the System Staticfile Buildpack

To use the Staticfile buildpack installed in your deployment, run the command below in the root directory of the app. The root directory of the app must contain a file named `Staticfile`.

```
cf push APP-NAME
```

Where `APP-NAME` is the name you want to give your app.

For example:

```
$ cf push my-app
Creating app my-app in org sample-org / space sample-space as username@example.com...
OK
...
requested state: started
instances: 1/1
usage: 1G x 1 instances
urls: my-app.example.com
last uploaded: Wed Apr 1 12:55:41 UTC 2020
stack: cflinuxfs3
buildpack: staticfile

 state since cpu memory disk details
#0 running 2020-04-01 12:55:57 PM 0.0% 12.6M of 1G 5.9M of 1G
```

### Option 2: Specify a Staticfile Buildpack

To explicitly specify a Staticfile buildpack, run the command below in the root directory of the app. You may want to specify a buildpack if your deployment does not have the Staticfile buildpack installed or the installed version is outdated.

```
cf push APP-NAME -b BUILDPACK-NAME-OR-PATH
```

Where:

- `APP-NAME` is the name you want to give your app.
- `BUILDPACK-NAME-OR-PATH` is either the name of a buildpack that is installed in your deployment or the path to a buildpack. You can find the Cloud Foundry Staticfile buildpack in the [Staticfile repository](#) on GitHub.

For example:

```
$ cf push my-app -b https://github.com/cloudfoundry/staticfile-buildpack.git
Creating app my-app in org sample-org / space sample-space as username@example.com...
OK
...
requested state: started
instances: 1/1
usage: 1G x 1 instances
urls: my-app.example.com
last uploaded: Wed Apr 1 12:55:41 UTC 2020
stack: cflinuxfs3
buildpack: https://github.com/cloudfoundry/staticfile-buildpack.git

 state since cpu memory disk details
#0 running 2020-04-01 12:55:57 PM 0.0% 12.6M of 1G 5.9M of 1G
```

## Verifying the Push

After you push the app, follow the steps below to verify that the push was successful:

1. Find the URL of your app in the output of the `cf push` command. For example, the URL in the terminal output above is `my-app.example.com`.
2. In a browser, navigate to the URL to view your app.

## Example Staticfile Buildpack Apps

For different examples of apps that use the Staticfile buildpack, see the [fixtures](#) directory in the Staticfile buildpack GitHub repo.

## Help and Support

A number of channels exist where you can get more help when using the Staticfile buildpack, or with developing your own Staticfile buildpack.

- **Staticfile Buildpack Repository in GitHub:** Find more information about using and extending the Staticfile buildpack in [GitHub repository](#).
- **Release Notes:** Find current information about this buildpack on the Staticfile buildpack [release page](#) in GitHub.
- **Slack:** Join the #buildpacks channel in the [Cloud Foundry Slack community](#).

## Customizing and Developing Buildpacks

Buildpacks enable you to packaging frameworks and/or runtime support for your application. Cloud Foundry provides with system buildpacks out-of-the-box and provides an interface for customizing existing buildpacks and developing new ones.

### Customizing and Creating Buildpacks

If your application uses a language or framework that the Cloud Foundry system buildpacks do not support, do one of the following:

- Use a [Cloud Foundry Community Buildpack ↗](#).
- Use a [Heroku Third-Party Buildpack ↗](#).
- Customize an existing buildpack or create your own [custom buildpack](#). A common development practice for custom buildpacks is to fork existing buildpacks and sync subsequent patches from upstream. For information about customizing an existing buildpack or creating your own, see the following:
  - [Creating Custom Buildpacks](#)
  - [Packaging Dependencies for Offline Buildpacks](#)

### Maintaining Buildpacks

After you have modified an existing buildpack or created your own, it is necessary to maintain it. Refer to the following when maintaining your own buildpacks:

- [Merging from Upstream Buildpacks](#)
- [Upgrading Dependency Versions](#)

 **Note:** To configure a production server for your web app, see the [Configuring a Production Server](#) topic.

### Using CI for Buildpacks

For information about updating and releasing a new version of a Cloud Foundry buildpack through the Cloud Foundry Buildpacks Team Concourse pipeline, see [Using CI for Buildpacks](#). You can use this as a model when working with Concourse to build and release new versions of your own buildpacks.

## Creating Custom Buildpacks

This topic describes how to create custom buildpacks for Pivotal Application Service (PAS).

For more information about how buildpacks work, see the [How Buildpacks Work in Cloud Foundry](#) topic.

## Package Custom Buildpacks

PAS buildpacks can work with limited or no Internet connectivity. The [buildpack-packager](#) gives the same flexibility to custom buildpacks, enabling them to work in partially or completely disconnected environments.

### Use the Buildpack Packager

1. Ensure that you have installed the [buildpack-packager](#).
2. Create a manifest.yml in your buildpack.
3. Run the packager in cached mode:

```
$ buildpack-packager build -cached -any-stack
```

The packager will add (almost) everything in your buildpack directory into a zip file. It will exclude anything marked for exclusion in your manifest.

In cached mode, the packager downloads and adds dependencies as described in the manifest.

For more information, see the documentation in the [buildpack-packager GitHub repository](#).

### Use and Share the Packaged Buildpack

After you have packaged your buildpack using `buildpack-packager` you can use the resulting `.zip` file locally, or share it with others by uploading it to any network location that is accessible to the CLI. Users can then specify the buildpack with the `-b` option when they push apps. See [Deploying Apps with a Custom Buildpack](#) for details.

 **Note:** Offline buildpack packages may contain proprietary dependencies that require distribution licensing or export control measures. For more information about offline buildpacks, refer to [Packaging Dependencies for Offline Buildpacks](#).

You can also use the `cf create-buildpack` command to upload the buildpack into your Cloud Foundry deployment, making it accessible without the `-b` flag:

```
$ cf create-buildpack BUILDPACK PATH POSITION [--enable|--disable]
```

You can find more documentation in the [Managing Custom Buildpacks](#) topic.

### Specify a Default Version

As of [buildpack-packager version 2.3.0](#), you can specify the default version for a dependency by adding a `default_versions` object to the `manifest.yml` file. The `default_versions` object has two properties, `name` and `version`. For example:

```
default_versions:
- name: go
 version: 1.6.3
- name: other-dependency
 version: 1.1.1
```

To specify a default version:

1. Add the `default_version` object to your manifest, following the `rules` below. For a complete example, see [manifest.yml](#) in the `go-buildpack` repository in GitHub.

2. Run the `default_version_for` script from the [compile-extensions](#) repository, passing the path of your `manifest.yml` and the dependency name as arguments. The following command uses the example manifest from step 1:

```
$./compile-extensions/bin/default_version_for manifest.yml go 1.6.3
```

## Rules for Specifying a Default Version

The `buildpack-packager` script validates this object according to the following rules:

- You can create at most one entry under `default_versions` for a single dependency. The following example causes `buildpack-packager` to fail with an error because the manifest specifies two default versions for the same `go` dependency.

```
Incorrect; will fail to package
default_versions:
- name: go
 version: 1.6.3
- name: go
 version: 1.7.5
```

- If you specify a `default_version` for a dependency, you must also list that dependency and version under the `dependencies` section of the manifest. The following example causes `buildpack-packager` to fail with an error because the manifest specifies `version: 1.9.2` for the `go` dependency, but lists `version: 1.7.5` under `dependencies`.

```
Incorrect; will fail to package
default_versions:
- name: go
 version: 1.9.2

dependencies:
- name: go
 version: 1.7.5
 uri: https://storage.googleapis.com/golang/go1.7.5.linux-amd64.tar.gz
 md5: c8cb76e2308c792e2705c2eb1b55de95
 cf_stacks:
- cflinuxfs3
```

## Core Buildpack Communication Contract

This section describes the communication contract followed by the PAS core buildpacks. This contract enables buildpacks to interact with one another, so that developers can use multiple buildpacks with their applications.

Buildpack developers must ensure their custom buildpacks follow the contract.

This section uses the following placeholders:

- `IDX` is the zero-padded index matching the position of the buildpack in the priority list.
- `MD5` is the MD5 checksum of the buildpack's URL.

For all buildpacks that supply dependencies via `/bin/supply`:

- The buildpack must create `/tmp/deps/IDX/config.yml` to provide a name to subsequent buildpacks. This file may also contain miscellaneous configuration for subsequent buildpacks.
- The `config.yml` file should be formatted as follows, replacing `BUILDPACK` with the name of the buildpack providing dependencies and `YAML-OBJECT` with the YAML object that contains buildpack-specific configuration: `name: BUILDPACK config: YAML-OBJECT`
- The following directories may be created inside of `/tmp/deps/IDX/` to provide dependencies to subsequent buildpacks:
  - `/bin`: Contains binaries intended for \$PATH during staging and launch
  - `/lib`: Contains libraries intended for \$LD\_LIBRARY\_PATH during staging and launch
  - `/include`: Contains header files intended for compilation during staging
  - `/pkgconfig`: Contains `pkgconfig` files intended for compilation during staging
  - `/env`: Contains environment vars intended for staging, loaded as `FILENAME=FILECONTENTS`
  - `/profile.d`: Contains scripts intended for `/app/.profile.d`, sourced before launch
- The buildpack may make use of previous non-final buildpacks by scanning `/tmp/deps/` for index-named directories containing `config.yml`.

For the last buildpack:

- To make use of dependencies provided by the previously applied buildpacks, the last buildpack must scan `/tmp/deps/` for index-named directories containing `config.yml`.
- To make use of dependencies provided by previous buildpacks, the last buildpack:
  - May use `/bin` during staging, or make it available in `$PATH` during launch
  - May use `/lib` during staging, or make it available in `$LD_LIBRARY_PATH` during launch
  - May use `/include`, `/pkgconfig`, or `/env` during staging
  - May copy files from `/profile.d` to `/tmp/app/.profile.d` during staging
  - May use the supplied config object in `config.yml` during the staging process

## Deploy Apps with a Custom Buildpack

Once a custom buildpack has been created and pushed to a public git repository, the git URL can be passed via the cf CLI when pushing an app.

For example, for a buildpack that has been pushed to GitHub:

```
$ cf push my-new-app -b git://github.com/johndoe/my-buildpack.git
```

Alternatively, you can use a private git repository, with https and username/password authentication, as follows:

```
$ cf push my-new-app -b https://username:password@github.com/johndoe/my-buildpack.git
```

By default, PAS uses the default branch of the buildpack's git repository. You can specify a different branch using the git url as shown in the following example:

```
$ cf push my-new-app -b https://github.com/johndoe/my-buildpack.git#my-branch-name
```

Additionally, you can use tags in a git repository, as follows:

```
$ cf push my-new-app -b https://github.com/johndoe/my-buildpack#v1.4.2
```

The app will then be deployed to PAS, and the buildpack will be cloned from the repository and applied to the app.

 **Note:** If a buildpack is specified using `cf push -b` the `detect` step will be skipped and as a result, no buildpack `detect` scripts will be run.

## Disable Custom Buildpacks

Operators can choose to disable custom buildpacks. For more information, see [Disabling Custom Buildpacks](#).

 **Note:** A common development practice for custom buildpacks is to fork existing buildpacks and sync subsequent patches from upstream. To merge upstream patches to your custom buildpack, use the approach that GitHub recommends for [Syncing a fork](#).

## Packaging Dependencies for Offline Buildpacks

This topic describes the dependency storage options available to developers creating offline buildpacks.

### About Offline Buildpacks

Online, or uncached, buildpacks require an Internet connection to download dependencies, such as language interpreters and compilers. Alternatively, you can create offline, or cached, buildpacks that are packaged with their dependencies. These offline buildpacks do not connect to the Internet when they are used to deploy Cloud Foundry apps.

 **Note:** Offline buildpacks may contain proprietary dependencies that require distribution licensing or export control measures.

You can find instructions for building the offline packages in the `README.md` file of each buildpack repository. For example, see the [Java buildpack](#).

### Package Dependencies in the Buildpack

The simplest way to package dependencies in a custom buildpack is to keep the dependencies in your buildpack source. However, this is strongly discouraged. Keeping the dependencies in your source consumes unnecessary space.

To avoid keeping the dependencies in source control, load the dependencies into your buildpack and provide a script for the operator to create a zipfile of the buildpack.

For example, the operator might complete the following process:

```
$ # Clones your buildpack
$ git clone http://YOUR-GITHUB-REPOSITORY.example.com/repository
$ cd SomeBuildPackName

$ # Creates a zipfile using your script
$./SomeScriptName
---> downloading-dependencies.... done
---> creating zipfile: ZippedBuildPackName.zip

$ # Adds the buildpack zipfile to the Cloud Foundry instance
$ cf create-buildpack SomeBuildPackName ZippedBuildPackName.zip 1
```

### Pros

- Least complicated process for operators
- Least complicated maintenance process for buildpack developers

### Cons

- Cloud Foundry admin buildpack uploads are limited to 1 GB, so the dependencies might not fit
- Security and functional patches to dependencies require updating the buildpack

### Package Selected Dependencies in the Buildpack

This is a variant of the [package dependencies in the buildpack](#) method described above. In this variation, the administrator edits a configuration file such as `dependencies.yml` to include a limited subset of the buildpack dependencies, then packages and uploads the buildpack.

 **Note:** This approach is strongly discouraged. Please see the Cons section below for more information.

The administrator completes the following steps:

```
$ # Clones your buildpack
$ git clone http://YOUR-GITHUB-REPOSITORY.example.com/repository
$ cd SomeBuildPackName

$ # Selects dependencies
$ vi dependencies.yml # Or copy in a preferred config

$ # Builds a package using your script
$./package
----> downloading-dependencies... done
----> creating zipfile: cobol_buildpack.zip

$ # Adds the buildpack to the Cloud Foundry instance
$ cf create-buildpack cobol-buildpack cobol_buildpack.zip 1

$ # Pushes an app using your buildpack
$ cd ~/my_app
$ cf push my-cobol-webapp -b cobol-buildpack
```

## Pros

- Possible to avoid the Cloud Foundry admin buildpack upload size limit in one of two ways:
  - If the administrator chooses a limited subset of dependencies
  - If the administrator maintains different packages for different dependency sets

## Cons

- More complex for buildpack maintainers
- Security updates to dependencies require updating the buildpack
- Proliferation of buildpacks that require maintenance:
  - For each configuration, there is an update required for each security patch
  - Culling orphan configurations may be difficult or impossible
  - Administrators need to track configurations and merge them with updates to the buildpack
  - May result in with a different config for each app

## Rely on a Local Mirror

In this method, the administrator provides a compatible file store of dependencies. When running the buildpack, the administrator specifies the location of the file store. The buildpack should handle missing dependencies gracefully.

The administrator completes the following process:

```
$ # Clones your buildpack
$ git clone http://YOUR-GITHUB-REPOSITORY.example.com/repository
$ cd SomeBuildPackName

$ # Builds a package using your script
$./package https://dependency/repository
----> creating zipfile: cobol_buildpack.zip

$ # Adds the buildpack to the Cloud Foundry instance
$ cf create-buildpack cobol-buildpack cobol_buildpack.zip 1

$ # Pushes an app using your buildpack
$ cd ~/my_app
$ cf push my-cobol-webapp -b cobol-buildpack
----> deploying app
----> downloading dependencies:
https://OUR-INTERNAL-SITE.example.com/dependency/repository/dep1.tgz.... done
https://OUR-INTERNAL-SITE.example.com/dependency/repository/dep2.tgz.... WARNING: dependency not found!
```

## Pros

- Avoids the Cloud Foundry admin buildpack upload size limit
- Leaves the administrator completely in control of providing dependencies
- Security and functional patches for dependencies can be maintained separately on the mirror given the following conditions:
  - The buildpack is designed to use newer semantically versioned dependencies
  - Buildpack behavior does not change with the newer functional changes

## Cons

- The administrator needs to set up and maintain a mirror
- The additional config option presents a maintenance burden

## Merging from Upstream Buildpacks

This topic describes how to maintain your forked buildpack by merging it with the upstream buildpack. This allows you to keep your fork updated with changes from the original buildpack, providing patches, updates, and new features.

The following procedure assumes that you are maintaining a custom buildpack that was forked from a Cloud Foundry system buildpack. However, you can use the same procedure to update a buildpack forked from any upstream buildpack.

To sync your forked buildpack with an upstream Cloud Foundry buildpack:

1. Navigate to your forked repository on GitHub and click **Compare** in the upper right to display the **Comparing changes** page. This page shows the unmerged commits between your forked buildpack and the upstream buildpack.
2. Inspect the unmerged commits and confirm that you want to merge them all.
3. In a terminal window, navigate to the forked repository and set the upstream remote as the Cloud Foundry buildpack repository.

```
$ cd ~/workspace/ruby-buildpack
$ git remote add upstream git@github.com:cloudfoundry/ruby-buildpack.git
```

4. Pull down the remote upstream changes.

```
$ git fetch upstream
```

5. Merge the upstream changes into the intended branch. You may need to resolve merge conflicts. This example shows merging the `master` branch of the upstream buildpack into the `master` branch of the forked buildpack.

```
$ git checkout master
$ git merge upstream/master
```

 **Note:** When merging upstream buildpacks, do not use `git rebase`. This approach is not sustainable because you confront the same merge conflicts repeatedly.

6. Run the buildpack test suite to ensure that the upstream changes do not break anything.

```
$ BUNDLE_GEMFILE=cf.Gemfile buildpack-build
```

7. Push the updated branch.

```
$ git push
```

Your forked buildpack is now synced with the upstream Cloud Foundry buildpack.

For more information about syncing forks, see the GitHub topic [Syncing a Fork](#).

## Upgrading Dependency Versions

This topic describes how to upgrade a dependency version in a custom buildpack. These procedures enable Cloud Foundry (CF) operators to maintain custom buildpacks that contain dependencies outside of the dependencies in the CF system buildpacks.

## Cloud Foundry Buildpacks Team Process

**Note:** The procedures in this topic refer to the tools used by the CF buildpacks team, but they do not require the specific tools listed below. You can use any continuous integration (CI) system and workflow management tool to update dependencies in custom buildpacks.

The CF buildpacks team uses the following tools to update dependencies:

- A Concourse deployment of the [buildpacks-ci](#) pipelines
- The [public-buildpacks-ci-robots](#) GitHub repository
- [Pivotal Tracker](#) for workflow management

When the [New Releases](#) job in the [notifications pipeline](#) detects a new version of a tracked dependency in a buildpack, it creates a [Tracker](#) story about building and including the new version of the dependency in the buildpack manifests. It also posts a message as the `dependency-notifier` to the [#buildpacks channel](#) in the Cloud Foundry Slack channel.

## Build the Binaries

For all dependencies, you must build the binary from source or acquire the binary as a tarball from a trusted source. For most dependencies, the CF buildpacks team builds the binaries from source.

**Note:** The steps below assume you are using a Concourse deployment of the `buildpacks-ci` pipelines and Pivotal Tracker.

To build the binary for a dependency, perform the following steps:

1. Navigate to the `public-buildpacks-ci-robots` directory and verify no uncommitted changes exist.

```
$ cd ~/workspace/public-buildpacks-ci-robots
$ git status
```

2. Run the `git pull` command in the directory to ensure it contains most recent version of the contents.

```
$ git pull -r
```

3. Navigate to the `binary-builds` directory.

```
$ cd binary-builds
```

4. Locate the YAML file for the buildpack you want to build a binary for. The directory contains YAML files for all the packages and dependencies tracked by the CF buildpacks team. Each YAML file correlates to the build queue for one dependency or package, and uses the naming format `DEPENDENCY-NAME.yml`. For example, the YAML file tracking the build queue for Ruby is named `ruby-builds.yml` and contains the following contents:

```
--
ruby: []
```

5. Different buildpacks use different signatures for verification. Determine which signature your buildpack requires by consulting the list in the [buildpacks](#) section of this topic and run follow the instructions to locate the SHA256, MD5, or GPG signature for the binary:

- For the SHA256 of a file, run `shasum -a 256 FILE-NAME`.
- For the MD5 of a file, run `md5 FILE-NAME`.
- For the GPG signature (for Nginx), see the [Nginx Downloads](#) page.

6. Add the version and verification for the new binary to the YAML file as attributes of an element under the dependency name. For example, to build the Ruby 2.3.0 binary verified with SHA256, add the following to the YAML file:

```
--
ruby:
- version: 2.3.0
sha256: ba5ba60e5f1aa21b4ef8e9bf35b9ddb57286cb546aac4b5a28c71f459467e507
```

 **Note:** Do not preface the version number with the name of the binary or language. For example, specify `2.3.0` for `version` instead of `ruby-2.3.0`.

You can enqueue builds for multiple versions at the same time. For example, to build both the Ruby 2.3.0 binary and the Ruby 2.3.1 binary, add the following to the YAML file:

```
--
ruby:
- version: 2.3.0
sha256: ba5ba60e5f1aa21b4ef8e9bf35b9ddb57286cb546aac4b5a28c71f459467e507
- version: 2.3.1
sha256: b87c738cb2032bf4920fef8e3864dc5cf8eae9d89d8d523ce0236945c5797dcd
```

7. Use the `git add` command to stage your changes:

```
$ git add .
```

8. Use the `git commit -m "YOUR-COMMIT-MESSAGE [#STORY-NUMBER]"` command to commit your changes using the Tracker story number. Replace `YOUR-COMMIT-MESSAGE` with your commit message, and `STORY-NUMBER` with the number of your Tracker story.

```
$ git commit -m "make that change [#1234567890]"
```

9. Run `git push` to push your changes to the remote origin.

```
$ git push
```

10. Pushing your changes triggers the binary building process, which you can monitor at the [binary-builder pipeline](#) of your own `buildpacks-ci` Concourse deployment. When the build completes, it adds a link to the Concourse build run to the Tracker story specified in the commit message for the new release.

 **Note:** Binary builds are executed by the Cloud Foundry [Binary Builder](#) and the [binary-builder pipeline](#).

## Update Buildpack Manifests

After you build the binary for a dependency that you can access and download from a URL, follow the instructions below to add the dependency version to the buildpack manifest.

-  **Note:** The steps below assume you are using a Concourse deployment of the `buildpacks-ci` pipelines and Pivotal Tracker.

1. Navigate to the directory of the buildpack for which you want to update dependencies and run `git checkout develop` to check out the `develop` branch.

```
$ cd ~/workspace/ruby-buildpack
$ git checkout develop
```

2. Edit the `manifest.yml` file for the buildpack to add or remove dependencies.

```
dependencies:
- name: ruby
version: 2.3.0
md5: 535342030a11abeb11497824bf642bf2
uri: https://pivotal-buildpacks.s3.amazonaws.com/concourse-binaries/ruby/ruby-2.3.0-linux-x64.tgz
cf_stacks:
- cflinuxfs3
```

- Follow the current structure of the manifest. For example, if the manifest includes the two most recent patch versions for each minor version of the language, you should also include the two most recent patch versions for each minor version of the language, such as both `ruby-2.1.9` and `ruby-2.1.8`.

- o Copy the `uri` and the `md5` from the `build-BINARY-NAME` job that ran in the Concourse binary-builder pipeline and add them to the manifest.

**Note:** In the PHP buildpack, you may see a `modules` line for each PHP dependency in the manifest. Do not include this `modules` line in your new PHP dependency entry. The `modules` line will be added to the manifest by the `ensure-manifest-has-modules` Concourse job in the `php-buildpack` when you commit and push your changes. You can see this in the output logs of the `build-out` task.

3. Replace any other mentions of the old version number in the buildpack repository with the new version number. The CF buildpack team uses [Ag ↗](#) for text searching.

```
$ ag OLD-VERSION
```

4. Run the following command to package and upload the buildpack, set up the org and space for tests in the specified CF deployment, and run the CF buildpack tests.

```
$ BUNDLE_GEMFILE=cf.Gemfile buildpack-build
```

If the command fails, you may need to fix or change the tests, fixtures, or other parts of the buildpack.

5. Once the test suite completely passes, use git commands to stage, commit, and push your changes:

```
$ git add .
$ git commit -m "YOUR-MESSAGE[#TRACKER-STORY-ID]"
$ git push
```

6. Monitor the `LANGUAGE-buildpack` pipeline in Concourse. Once the test suite builds, the `specs-lts-develop` job and `specs-edge-develop` job, pass for the buildpack, you can deliver the Tracker story for the new Dependency release. Copy and paste links for the successful test suite builds into the Tracker story.

## Buildpacks

The following list contains information about the buildpacks maintained by the CF buildpacks team.

### Go Buildpack

Go:

- **Built from:** A tarred binary, `GO-VERSION.linux-amd64.tar.gz`, provided by [Google on the Go Downloads ↗](#) page
- **Verified with:** The MD5 of the tarred binary
- **Example:** [Using the Google Tarred Binary for Go 1.6.2 ↗](#)

Godep:

- **Built from:** A source code `.tar.gz` file from the [Godep GitHub releases ↗](#) page
- **Verified with:** The SHA256 of the source **Example:** [Automated enqueueing of binary build for Godep 72 ↗](#)

**Note:** The [buildpacks-ci ↗](#) [binary-builder ↗](#) pipeline automates the process of detecting, uploading, and updating Godep in the manifest.

### Node.js Buildpack

Node:

- **Verified with:** The SHA256 of the `node-vVERSION.tar.gz` file listed on <https://nodejs.org/dist/vVERSION/SHASUMS256.txt> For example, for Node version 4.4.6, the CF buildpacks team verifies with the SHA256 for `node-v4.4.6.tar.gz` on its [SHASUMS256 ↗](#) page.
- **Example:** [Enqueuing binary builds for Node 4.4.5 and 6.2.0 ↗](#)

## Python Buildpack

Python:

- **Verified with:** The MD5 of the `Gzipped source tarball`, listed on <https://www.python.org/downloads/release/python-VERSION/>, where `VERSION` has no periods. For example, for Python version `2.7.12`, use the MD5 for the `Gzipped source tarball` on its [downloads](#) page.
- **Example:** [Enqueuing binary build for Python 2.7.12](#)

## Java Buildpack

OpenJDK:

- **Built from:** The tarred [OpenJDK files](#) managed by the CF Java Buildpack team
- **Verified with:** The MD5 of the tarred OpenJDK files

## Ruby Buildpack

JRuby:

- **Verified with:** The MD5 of the Source `.tar.gz` file from the [JRuby Downloads](#) page
- **Example:** [Enqueuing binary build for JRuby 9.1.2.0](#)

Ruby:

- **Verified with:** The SHA256 of the source from the [Ruby Downloads](#) page
- **Example:** [Enqueuing binary builds for Ruby 2.2.5 and 2.3.1](#)

Bundler:

- **Verified with:** The SHA256 of the `.gem` file from [Rubygems](#)
- **Example:** [Enqueuing binary build for Bundler 1.12.5](#)

## PHP Buildpack

PHP:

- **Verified with:** The SHA256 of the `.tar.gz` file from the [PHP Downloads](#) page.
- To enqueue builds for PHP, you need to edit a file in the `public-buildpacks-ci-robots` repository. For PHP5 versions, the CF buildpacks team enqueues builds in the `binary-builds/php-builds.yml` file. For PHP7 versions, the CF buildpacks team enqueues builds in the `binary-builds/php7-builds.yml` file.
- **Example:** [Enqueuing binary builds for PHP 7.2.5 and 7.0.30](#)

Nginx:

- **Verified with:** The `gpg-rsa-key-id` and `gpg-signature` of the version. The `gpg-rsa-key-id` is the same for each version/build, but the `gpg-signature` will be different. This information is located on the [Nginx Downloads](#) page.
- **Example:** [Enqueuing binary build for Nginx 1.11.0](#)

HTTPD:

- **Verified with:** The MD5 of the `.tar.bz2` file from the [HTTPD Downloads](#) page
- **Example:** [Enqueuing binary build for HTTPD 2.4.20](#)

Composer:

- **Verified with:** The SHA256 of the `composer.phar` file from the [Composer Downloads](#) page
- For Composer, there is no build process as the `composer.phar` file is the binary. In the manual process, connect to the appropriate S3 bucket using the correct AWS credentials. Create a new directory with the name of the composer version, for example `1.0.2`, and put the appropriate `composer.phar` file into that directory. For Composer `v1.0.2`, connect and create the `php/binaries/trusty/composer/1.0.2` directory. Then place the `composer.phar` file into that directory so the binary is available at `php/binaries/trusty/composer/1.0.2/composer.phar`.

 Note: The [buildpacks-ci](#) ↗ [binary-builder](#) ↗ pipeline automates the process of detecting, uploading, and updating Composer in the manifest.

- Example: [Automated enqueueing of binary build for Composer 1.1.2](#) ↗

## Staticfile Buildpack

Nginx:

- **Verified with:** The `gpg-rsa-key-id` and `gpg-signature` of the version. The `gpg-rsa-key-id` is the same for each version/build, but the `gpg-signature` will be different. This information is located on the [Nginx Downloads](#) ↗ page.
- Example: [Enqueuing binary build for Nginx 1.11.0](#) ↗

## Binary Buildpack

The Binary buildpack has no dependencies.

## Using CI for Buildpacks

The Cloud Foundry (CF) Buildpacks team and other CF buildpack development teams use [Concourse continuous integration \(Concourse CI\)](#) pipelines to integrate new buildpack releases. This topic provides links to information that describes how to release new versions of Cloud Foundry buildpacks using Concourse CI, and how to update Ruby gems used for CF buildpack development.

Each of the following are applicable to all supported buildpack languages and frameworks:

- [Releasing a New Buildpack Version](#)
- [Updating Buildpack-Related Gems](#)

## Releasing a New Buildpack Version

This topic describes how to update and release a new version of a Cloud Foundry (CF) buildpack through the CF Buildpacks Team Concourse [pipeline](#). Concourse is a continuous integration (CI) tool for software development teams. This is the process used by the CF Buildpacks Team and other CF buildpack development teams. You can use this process as a model for using Concourse to build and release new versions of your own buildpacks.

The Concourse pipelines for Cloud Foundry buildpacks are located in the [buildpacks-ci](#) GitHub repository.

## Release a New Buildpack Version

To release a new buildpack version, perform the following:

1. Ensure you have downloaded the [buildpacks-ci](#) repository:

```
$ git clone https://github.com/cloudfoundry/buildpacks-ci.git
```

2. From the buildpack directory, check out the `develop` branch of the buildpack:

```
$ cd /system/path/to/buildpack
$ git checkout develop
```

3. Ensure you have the most current version of the repository:

```
$ git pull -r
```

4. Run `bump` to update the version in the buildpack repository:

```
$ /system/path/to/buildpacks-ci/scripts/bump
```

5. Modify the `CHANGELOG` file manually to condense recent commits into relevant changes. For more information, see [Modify Changelogs](#).

6. Add and commit your changes:

```
$ git add VERSION CHANGELOG
$ git commit -m "Bump version to $(cat VERSION) [{insert story #}]"
```

7. Push your changes to the `develop` branch:

```
$ git push origin develop
```

## Concourse Buildpack Workflow

If [buildpacks-ci](#) is not deployed to Concourse, manually add a Git tag to the buildpack and mark the tag as a release on GitHub.

If [buildpacks-ci](#) is deployed to Concourse, the buildpack update passes through the following life cycle:

1. Concourse triggers the `buildpack-to-master` job in the pipeline for the updated buildpack. This job merges `develop` onto the `master` branch of the buildpack.
2. The `detect-new-buildpack-and-upload-artifacts` job triggers in the pipeline for the updated buildpack. This job creates a cached and uncached buildpack and uploads them to an AWS S3 bucket.
3. The `specs-lts-master` and `specs-edge-master` jobs trigger and run the buildpack test suite and the buildpack-specific tests of the [Buildpack Runtime Acceptance Tests \(BRATS\)](#).
4. If you are using [Pivotal Tracker](#), paste the links for the `specs-edge-master` and `specs-lts-master` builds in the related buildpack release story and deliver that story.
5. Your project manager can manually trigger the `buildpack-to-github` job on Concourse as part of the acceptance process. This releases the buildpack to GitHub.

6. After the buildpack has been released to GitHub, the `cf-release` pipeline is triggered using the manual trigger of the `recreate-bosh-lite` job on that pipeline. If the new buildpack has been released to GitHub, the CF that is deployed for testing in the `cf-release` pipeline is tested against that new buildpack.
7. After the `cats` job has successfully completed, your project manager can ship the new buildpacks to the `cf-release` repository and create the new buildpack BOSH release by manually triggering the `ship-it` job.

 **Note:** If errors occur during this workflow, you may need to remove unwanted tags. For more information, see [Handle Unwanted Tags](#).

## Modify Changelogs

The [Ruby Buildpack changelog](#) shows an example layout and content of a changelog. In general, changelogs follow these conventions:

- Reference public tracker stories whenever possible.
- Exclude unnecessary files
- Combine and condense commit statements into individual stories containing valuable changes.

## Handle Unwanted Tags

If you encounter problems with the commit that contains the new version, change the target of the release tag by performing the following:

1. Ensure the repository is in a valid state and is building successfully.
2. Remove the tag from your local repository and from GitHub.
3. Start a build. The pipeline build script should re-tag the build if it is successful.

## Updating Buildpack-Related Gems

This topic describes how to update [buildpack-packager](#) and [machete](#), used for CF system buildpack development.

`buildpack-packager` packages buildpacks and `machete` provides an integration test framework.

The CF Buildpacks team uses the [gems-and-extensions pipeline](#) to:

1. Run the integration tests for `buildpack-packager` and `machete`
2. Update the gems in the buildpacks managed by the team

## Running the Update Process

 **Note:** The steps below assume you are using a Concourse deployment of the `buildpacks-ci` pipelines

At the end of the process, there will be a new GitHub release and updates will be applied to the buildpacks.

To update the version of either gem in a buildpack:

1. Confirm that the test job `<gemname>-specs` for the gem to be updated successfully ran on the commit you plan to update.
2. Manually trigger the `<gemname>-tag` job to update (“bump”) the version of the gem.
3. The `<gemname>-release` job will trigger. This will create a new GitHub release of the gem.
4. Each of the buildpack pipelines (e.g. the [go-buildpack pipeline](#)) has a job which watches for new releases of the gem. When a new release is detected, the buildpack’s `cf.Gemfile` is updated to that release version.
5. The commit made to the buildpack’s `cf.Gemfile` triggers the full integration test suite for that buildpack.

 **Note:** The final step will trigger all buildpack test suites simultaneously, causing contention for available shared BOSH-lite test environments.

## Services

The documentation in this section is intended for developers and operators interested in creating Managed Services for Cloud Foundry. Managed Services are defined as having been integrated with Cloud Foundry via APIs, and enable end users to provision reserved resources and credentials on demand. For documentation targeted at end users, such as how to provision services and integrate them with applications, see [Services Overview ↗](#).

To develop Managed Services for Cloud Foundry, you'll need a Cloud Foundry instance to test your service broker with as you are developing it. You must have admin access to your CF instance to manage service brokers and the services marketplace catalog. For local development, we recommend using the [BOSH guide ↗](#) to deploy your own local instance of Cloud Foundry using VirtualBox.

## Table of Contents

- [Overview](#)
- [Service Broker API](#)
  - [Open Service Broker API ↗](#)
  - [Platform Profiles ↗](#)
  - [Catalog Metadata ↗](#)
  - [Volume Services ↗](#)
  - [Release Notes ↗](#)
- [Managing Service Brokers](#)
- [Access Control](#)
- [Dashboard Single Sign-On](#)
- [Example Service Brokers](#)
- [Binding Credentials](#)
- [Enabling Service Instance Sharing](#)
- [Application Log Streaming](#)
- [Route Services](#)
- [Supporting Multiple Cloud Foundry Instances](#)

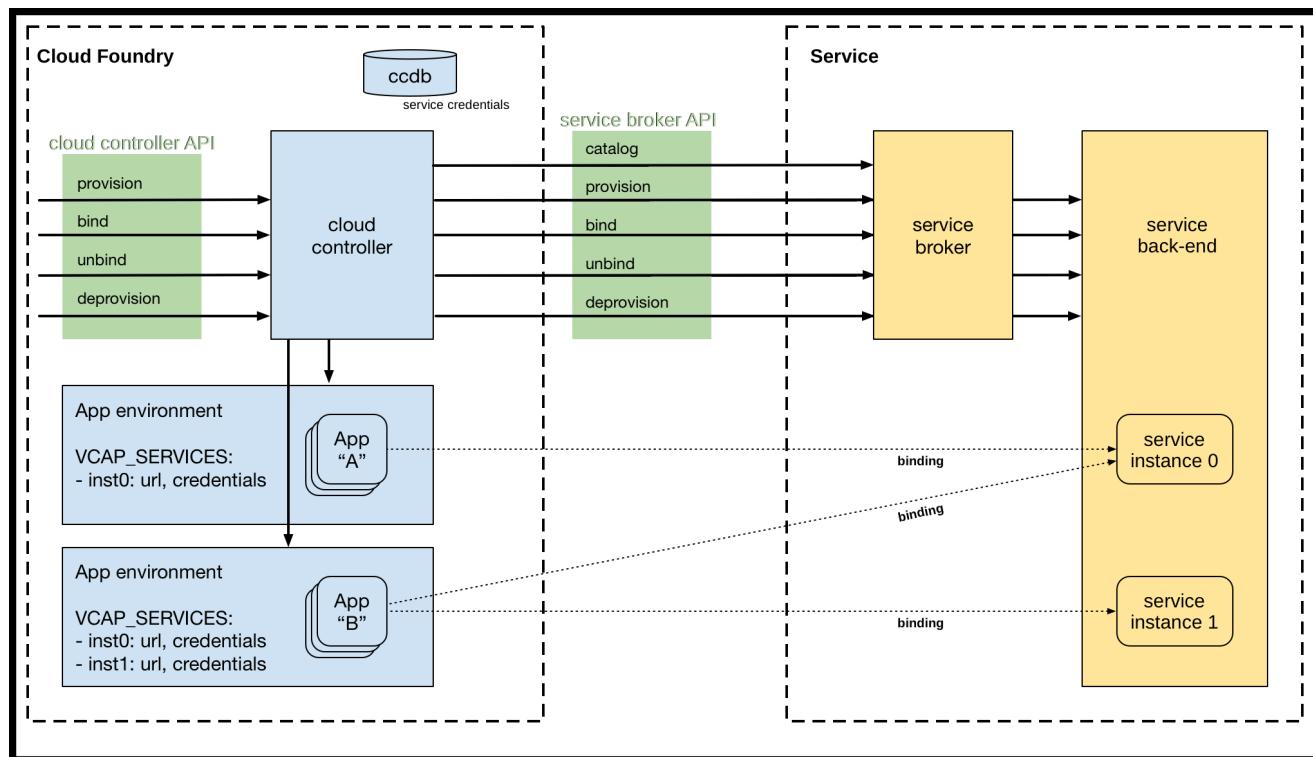
## Overview

### Architecture & Terminology

Services are integrated with Cloud Foundry by implementing a documented API for which the cloud controller is the client; we call this the Service Broker API. This should not be confused with the cloud controller API, often used to refer to the version of Cloud Foundry itself; when one refers to “Cloud Foundry v2” they are referring to the version of the cloud controller API. The services API is versioned independently of the cloud controller API.

Service Broker is the term we use to refer to a component of the service which implements the service broker API. This component was formerly referred to as a Service Gateway, however as traffic between applications and services does not flow through the broker we found the term gateway caused confusion. Although gateway still appears in old code, we use the term broker in conversation, in new code, and in documentation.

Service brokers advertise a catalog of service offerings and service plans, as well as interpreting calls for provision (create), bind, unbind, and deprovision (delete). What a broker does with each call can vary between services; in general, ‘provision’ reserves resources on a service and ‘bind’ delivers information to an application necessary for accessing the resource. We call the reserved resource a Service Instance. What a service instance represents can vary by service; it could be a single database on a multi-tenant server, a dedicated cluster, or even just an account on a web application.



[View a larger version of this image.](#)

### Implementation & Deployment

How a service is implemented is up to the service provider/developer. Cloud Foundry only requires that the service provider implement the service broker API. A broker can be implemented as a separate application, or by adding the required http endpoints to an existing service.

Because Cloud Foundry only requires that a service implements the broker API in order to be available to Cloud Foundry end users, many deployment models are possible. The following are examples of valid deployment models.

- Entire service packaged and deployed by BOSH alongside Cloud Foundry
- Broker packaged and deployed by BOSH alongside Cloud Foundry, rest of the service deployed and maintained by other means
- Broker (and optionally service) pushed as an application to Cloud Foundry user space
- Entire service, including broker, deployed and maintained outside of Cloud Foundry by other means



## Managing Service Brokers

This page assumes you are using cf CLI v6.16 or later.

In order to run many of the commands below, you must be authenticated with Cloud Foundry as an admin user or as a space developer.

## Quick Start

Given a service broker that has implemented the [Service Broker API](#), two steps are required to make its services available to end users in all orgs or a limited number of orgs by service plan.

1. [Register a Broker](#)

2. [Make Plans Public](#)

## Register a Broker

Registering a broker causes Cloud Controller to fetch and validate the catalog from your broker, and save the catalog to the Cloud Controller database. The basic auth username and password which are provided when adding a broker are encrypted in Cloud Controller database, and used by the Cloud Controller to authenticate with the broker when making all API calls. Your service broker should validate the username and password sent in every request; otherwise, anyone could curl your broker to delete service instances. When the broker is registered with an URL having the scheme `https`, Cloud Controller will make all calls to the broker over HTTPS.

As of cf-release 229, CC API 2.47.0, Cloud Foundry supports two types of brokers:*standard brokers* and *space-scoped brokers*. A list of their differences follows:

### Standard Brokers

- Publish service plans to specific orgs or all orgs in the deployment. Can also keep plans unavailable, or private.
- Created by admins, with the command `cf create-service-broker`

```
$ cf create-service-broker mybrokername someuser somethingsecure https://mybroker.example.com
```

- Managed by admins
- Service plans are created private. Before anyone can use them, an admin must explicitly [make them available](#) within an org or across all orgs.

### Space-Scoped Brokers

- Publish service plans only to users within the space they are created. Plans are unavailable outside of this space.
- Created by space developers using the command `cf create-service-broker` with the `--space-scoped` flag

```
$ cf create-service-broker mybrokername someuser somethingsecure https://mybroker.example.com --space-scoped
```

 **Note:** If a space developer runs `cf create-service-broker` without the `--space-scoped` flag, they receive an error.

- Managed by space developers
- Newly-created plans automatically publish to all users in their space.

## Make Plans Public

After an admin creates a new service plan from a standard broker, no one can use it until the admin explicitly makes it available to users within a specific org or all orgs in the deployment.

See the [Access Control](#) topic for how to make standard broker service plans available to users.

## Multiple Brokers, Services, Plans

Many service brokers may be added to a Cloud Foundry instance, each offering many services and plans. The following constraints should be kept in mind:

- It is not possible to have multiple brokers with the same name
- The service ID and plan IDs of each service advertised by the broker must be unique across Cloud Foundry. GUIDs are recommended for these fields.

**Note:** If your deployment uses Cloud Foundry API (CAPI) v1.71 or later, you can add multiple brokers with the same URL. In this case, the brokers must have different names. CAPI v1.70 and earlier do not support this feature.

See [Possible Errors](#) below for error messages and what do to when you see them.

## List Service Brokers

```
$ cf service-brokers
Getting service brokers as admin...
OK

Name URL
my-service-broker https://mybroker.example.com
```

## Update a Broker

Updating a broker is how to ingest changes a broker author has made into Cloud Foundry. Similar to adding a broker, update causes Cloud Controller to fetch the catalog from a broker, validate it, and update the Cloud Controller database with any changes found in the catalog.

Update also provides a means to change the basic auth credentials cloud controller uses to authenticate with a broker, as well as the base URL of the broker's API endpoints.

```
$ cf update-service-broker mybrokernome someuser somethingsecure https://mybroker.example.com
```

## Rename a Broker

A service broker can be renamed with the `rename-service-broker` command. This name is used only by the Cloud Foundry operator to identify brokers, and has no relation to configuration of the broker itself.

```
$ cf rename-service-broker mybrokernome mynewbrokernome
```

## Remove a Broker

Removing a service broker will remove all services and plans in the broker's catalog from the Cloud Foundry Marketplace.

```
$ cf delete-service-broker mybrokernome
```

**Note:** Attempting to remove a service broker will fail if there are service instances for any service plan in its catalog. When planning to shut down or delete a broker, make sure to remove all service instances first. Failure to do so will leave [orphaned service instances](#) in the Cloud Foundry database. If a service broker has been shut down without first deleting service instances, you can remove the instances with the CLI; see [Purge a Service](#).

## Purge a Service

If a service broker has been shut down or removed without first deleting service instances from Cloud Foundry, you will be unable to remove the service broker or its services and plans from the Marketplace. In development environments, broker authors often destroy their broker deployments and need a way to clean up the Cloud Controller database.

The following command will delete a service offering, all of its plans, as well as all associated service instances and bindings from the Cloud Controller database, without making any API calls to a service broker. Once all services for a broker have been purged, the broker can be removed normally.

```
$ cf purge-service-offering service-test
Warning: This operation assumes that the service broker responsible for this
service offering is no longer available, and all service instances have been
deleted, leaving orphan records in Cloud Foundry's database. All knowledge of
the service will be removed from Cloud Foundry, including service instances and
service bindings. No attempt will be made to contact the service broker; running
this command without destroying the service broker will cause orphan service
instances. After running this command you may want to run either
delete-service-auth-token or delete-service-broker to complete the cleanup.
```

```
Really purge service offering service-test from Cloud Foundry? y
OK
```

 **Note:** To purge services from v1 brokers, you must append your `cf purge-service-offering` command with `-p PROVIDER`.

## Purge a Service Instance

The following command will delete a single service instance, its service bindings and its service keys from the Cloud Controller database, without making any API calls to a service broker. This can be helpful in instances a Service Broker is not conforming to the Service Broker API and not returning a 200 or 410 to requests to delete the service instance.

```
$ cf purge-service-instance mysql-dev
WARNING: This operation assumes that the service broker responsible for this
service instance is no longer available or is not responding with a 200 or 410,
and the service instance has been deleted, leaving orphan records in Cloud
Foundry's database. All knowledge of the service instance will be removed from
Cloud Foundry, including service bindings and service keys.
```

```
Really purge service instance mysql-dev from Cloud Foundry?> y
Purging service mysql-dev...
OK
```

`purge-service-instance` requires cf-release v218 and cf CLI 6.14.0.

 **Note:** When multiple brokers provide two or more service instances with the same name, you must specify the broker by including the `-b BROKER` flag in the `cf purge-service-offering` command.

## Catalog Validation Behaviors

When Cloud Foundry fetches a catalog from a broker, it will compare the broker's id for services and plans with the `unique_id` values for services and plans in the Cloud Controller database.

Event	Action
The catalog fails to load or validate.	Cloud Foundry will return a meaningful error that the broker could not be reached or the catalog was not valid.
A service or plan in the broker catalog has an ID that is not present among the <code>unique_id</code> values in the marketplace database.	A new record must be added to the marketplace database.
A service or plan in the marketplace database are found with a <code>unique_id</code> that matches the broker catalog's ID.	The marketplace must update the records to match the broker's catalog.
The database has plans that are not found in the broker catalog, and there are no associated service instances.	The marketplace must remove these plans from the database, and then delete services that do not have associated plans from the database.
The database has plans that are not found in the broker catalog, but there are provisioned instances.	The marketplace must mark the plan inactive and no longer display it or allow it to be provisioned.

## Possible Errors

If incorrect basic auth credentials are provided:

Server error, status code: 500, error code: 10001, message: Authentication failed for the service broker API.  
Double-check that the username and password are correct:  
<https://github-broker.a1-app.example.com/v2/catalog>

If you receive the following errors, check your broker logs. You may have an internal error.

Server error, status code: 500, error code: 10001, message:  
The service broker response was not understood

Server error, status code: 500, error code: 10001, message:  
The service broker API returned an error from  
<https://github-broker.a1-app.example.com/v2/catalog>: 404 Not Found

Server error, status code: 500, error code: 10001, message:  
The service broker API returned an error from  
<https://github-broker.primo.example.com/v2/catalog>: 500 Internal Server Error

If your broker's catalog of services and plans violates validation of presence, uniqueness, and type, you will receive meaningful errors.

Server error, status code: 502, error code: 270012, message: Service broker catalog is invalid:  
Service service-name-1  
service id must be unique  
service description is required  
service "bindable" field must be a boolean, but has value "true"  
Plan plan-name-1  
plan metadata must be a hash, but has value [ {"bullets":>["bullet1", "bullet2"]} ]

## Managing Access to Service Plans

All new service plans from standard brokers are private by default. This means that when adding a new broker, or when adding a new plan to an existing broker's catalog, service plans won't immediately be available to end users. This lets an admin control which service plans are available to end users, and manage limited service availability.

Space-scoped brokers are registered to a specific space, and all users within that space can automatically access the broker's service plans. With space-scoped brokers, service visibility is not managed separately.

### Prerequisites

- CLI v6.4.0
- Cloud Controller API v2.9.0 (cf-release v179)
- Admin user access; the following commands can be run only by an admin user

### Display Access to Service Plans

The `service-access` CLI command enables an admin to see the current access control setting for every service plan in the marketplace, across all service brokers.

```
$ cf service-access
getting service access as admin...
broker: elasticsearch-broker
 service plan access orgs
 elasticsearch standard limited

broker: p-mysql
 service plan access orgs
 p-mysql 100mb-dev all
```

The `access` column shows values `all`, `limited`, or `none`, defined as follows:

- `all` - The service plan is available to all users, or *public*.
- `none` - No one can use the service plan; it is *private*.
- `limited` - The plan is available only to users within the orgs listed.

The `-b`, `-e`, and `-o` flags let you filter by broker, service, and org.

```
$ cf help service-access
NAME:
 service-access - List service access settings

USAGE:
 cf service-access [-b BROKER] [-e SERVICE] [-o ORG]

OPTIONS:
 -b access for plans of a particular broker
 -e access for plans of a particular service offering
 -o plans accessible by a particular org
```

### Enable Access to Service Plans

Admins use the `cf enable-service-access` command to give users access to service plans. The command grants access at the org level or across all orgs.

When an org has access to a plan, its users see the plan in the services marketplace (`cf marketplace`) and its Space Developer users can provision instances of the plan in their spaces.

#### Enable Access to a Subset of Users

The `-p` and `-o` flags to `cf enable-service-access` let the admin limit user access to specific service plans or orgs as follows:

- `-p PLAN` grants all users access to one service plan (access: `all`)
- `-o ORG` grants users in a specified org access to all plans (access: `limited`)
- `-p PLAN -o ORG` grants users in one org access to one plan (access: `limited`)

For example, the following command grants the org dev-user-org access to the p-mysql service.

```
$ cf enable-service-access p-mysql -o dev-user-org
Enabling access to all plans of service p-mysql for the org dev-user-org as admin...
OK

$ cf service-access
getting service access as admin...
broker: p-mysql
 service plan access orgs
 p-mysql dev-user-org
```

Run `cf help enable-service-access` to review these options from the command line.

Running `cf enable-service-access SERVICE-NAME` without any flags lets all users access every plan carried by the service. For example, the following command grants all-user access to all `p-mysql` service plans:

```
$ cf enable-service-access p-mysql
Enabling access to all plans of service p-mysql for all orgs as admin...
OK

$ cf service-access
getting service access as admin...
broker: p-mysql
 service plan access orgs
 p-mysql 100mb-dev all
```

 **Note:** When multiple brokers provide two or more services with the same name, you must specify the broker by including the `-b BROKER` flag in the `cf enable-service-access` command.

## Disable Access to Service Plans

Admins use the `cf disable-service-access` command to disable user access to service plans. The command denies access at the org level or across all orgs.

### Disable Access to All Plans for All Users

Running `cf disable-service-access SERVICE-NAME` without any flags disables all user access to all plans carried by the service. For example, the following command denies any user access to all `p-mysql` service plans:

```
$ cf disable-service-access p-mysql
Disabling access to all plans of service p-mysql for all orgs as admin...
OK

$ cf service-access
getting service access as admin...
broker: p-mysql
 service plan access orgs
 p-mysql 100mb-dev none
```

### Disable Access for Specific Orgs or Plans

The `-p` and `-o` flags to `cf disable-service-access` let the admin deny access to specific service plans or orgs as follows:

- `-p PLAN` disables user access to one service plan

- `-o ORG` disables access to all plans for users in a specified org
- `-p PLAN -o ORG` prevents users in one org from accessing one plan

Run `cf help disable-service-access` to review these options from the command line.

 **Note:** When multiple brokers provide two or more services with the same name, you must specify the broker by including the `-b BROKER` flag in the `cf disable-service-access` command.

## Limitations

- You cannot disable access to a service plan for an org if the plan is currently available to all orgs. You must first disable access for all orgs; then you can enable access for a particular org.

## Dashboard Single Sign-On

### Introduction

Single Sign-On (SSO) enables Pivotal Cloud Foundry (PCF) users to authenticate with third-party service dashboards using their PCF credentials. Service dashboards are web interfaces which enable users to interact with some or all of the features the service offers. SSO provides a streamlined experience for users, limiting repeated logins and multiple accounts across their managed services. The user's credentials are never directly transmitted to the service since the OAuth protocol handles authentication.

### Service Broker Responsibilities

#### Registering the Dashboard Client

1. A service broker must include the `dashboard_client` field in the JSON response from its [catalog endpoint](#) for each service implementing this feature. A valid response would appear as follows:

```
{
 "services": [
 {
 "id": "44b26033-1f54-4087-b7bc-da9652c2a539",
 ...
 "dashboard_client": {
 "id": "p-mysql-client",
 "secret": "p-mysql-secret",
 "redirect_uri": "http://p-mysql.example.com"
 }
 }
]
}
```

The `dashboard_client` field is a hash containing three fields:

- `id` is the unique identifier for the OAuth client that will be created for your service dashboard on the token server (UAA), and will be used by your dashboard to authenticate with the token server (UAA). If the client id is already taken, PCF will return an error when registering or updating the broker.
- `secret` is the shared secret your dashboard will use to authenticate with the token server (UAA).
- `redirect_uri` is used by the token server as an additional security precaution. UAA will not provide a token if the callback URL declared by the service dashboard doesn't match the domain name in `redirect_uri`. The token server matches on the domain name, so any paths will also match; e.g. a service dashboard requesting a token and declaring a callback URL of `http://p-mysql.example.com/manage/auth` would be approved if `redirect_uri` for its client is `http://p-mysql.example.com/`.

2. When a service broker which advertises the `dashboard_client` property for any of its services is [added or updated](#), Cloud Controller will create or update UAA clients as necessary. This client will be used by the service dashboard to authenticate users.

#### Dashboard URL

A service broker should return a URL for the `dashboard_url` field in response to a [provision request](#). Cloud Controller clients should expose this URL to users. `dashboard_url` can be found in the response from Cloud Controller to create a service instance, enumerate service instances, space summary, and other endpoints.

Users can then navigate to the service dashboard at the URL provided by `dashboard_url`, initiating the OAuth login flow.

### Service Dashboard Responsibilities

#### OAuth Flow

When a user navigates to the URL from `dashboard_url`, the service dashboard should initiate the OAuth login flow. A summary of the flow can be found in [section 1.2 of the OAuth RFC](#). OAuth expects the presence of an [Authorization Endpoint](#) and a [Token Endpoint](#). In PCF, these endpoints are provided by the UAA. Clients can discover the location of UAA from Cloud Controller's info endpoint; in the response the location can be found in the `token_endpoint` field.

```
$ curl api.example.com/info
{"name":"vcap","build":"2222","support":"http://support.example.com","version":2,"description":"Cloud Foundry sponsored by Example Company","authorization_endpoint":"https://login.example.com","token_endpoint":"https://uaa.example.com","allow_debug":true}
```

 To enable service dashboards to support SSO for service instances created from different PCF instances, the `/v2/info` url is sent to service brokers in the `X-Api-Info-Location` header of every API call. A service dashboard should be able to discover this URL from the broker, and enabling the dashboard to contact the appropriate UAA for a particular service instance.

A service dashboard should implement the OAuth Authorization Code Grant type ([UAA documentation](#), [RFC documentation](#)).

1. When a user visits the service dashboard at the value of `dashboard_url`, the dashboard should redirect the user's browser to the Authorization Endpoint and include its `client_id`, a `redirect_uri` (callback URL with domain matching the value of `dashboard_client.redirect_uri`), and list of requested scopes.  
Scopes are permissions included in the token a dashboard client will receive from UAA, and which Cloud Controller uses to enforce access. A client should request the minimum scopes it requires. The minimum scopes required for this workflow are `cloud_controller_service_permissions.read` and `openid`. For an explanation of the scopes available to dashboard clients, see [On Scopes](#).
2. UAA authenticates the user by redirecting the user to the Login Server, where the user then approves or denies the scopes requested by the service dashboard. The user is presented with human readable descriptions for permissions representing each scope. After authentication, the user's browser is redirected back to the Authorization endpoint on UAA with an authentication cookie for the UAA.
3. Assuming the user grants access, UAA redirects the user's browser back to the value of `redirect_uri` the dashboard provided in its request to the Authorization Endpoint. The `Location` header in the response includes an authorization code.

```
HTTP/1.1 302 Found
Location: https://p-mysql.example.com/manage/auth?code=F45jH
```

4. The dashboard UI should then request an access token from the Token Endpoint by including the authorization code received in the previous step. When making the request the dashboard must authenticate with UAA by passing the client `id` and `secret` in a basic auth header. UAA will verify that the client id matches the client it issued the code to. The dashboard should also include the `redirect_uri` used to obtain the authorization code for verification.
5. UAA authenticates the dashboard client, validates the authorization code, and ensures that the redirect URI received matches the URI used to redirect the client when the authorization code was issued. If valid, UAA responds back with an access token and a refresh token.

## Checking User Permissions

UAU is responsible for authenticating a user and providing the service with an access token with the requested permissions. However, after the user has been logged in, it is the responsibility of the service dashboard to verify that the user making the request to manage an instance currently has access to that service instance.

The service can accomplish this with a GET to the `/v2/service_instances/:guid/permissions` endpoint on the Cloud Controller. The request must include a token for an authenticated user and the service instance guid. The token is the same one obtained from the UAA in response to a request to the Token Endpoint, described above..

Example Request:

```
curl -H 'Content-Type: application/json' \
-H 'Authorization: bearer eyJ0eXAiOiJKV1QiLCJhbGciOiJIUzI1NiJ9.eyJc2VyX2lkIjoid' \
http://api.cloudfoundry.com/v2/service_instances/44b26033-1f54-4087-b7bc-da9652c2a539/permissions
```

Response:

```
{
 "manage": true,
 "read": true
}
```

The response includes the following fields which indicate the various user permissions for the given service instance: - `manage` – a `true` value indicates that the user has sufficient permissions to make changes to and update the service instance; `false` indicates insufficient permissions. - `read` – a `true` value indicates that the user has permission to access read-only diagnostic and monitoring information for the given service instance (e.g. permission to view a read-only dashboard); `false` indicates insufficient permissions.

Since administrators may change the permissions of users at any time, the service should check this endpoint whenever a user uses the SSO flow to access the service's UI.

## On Scopes

Scopes let you specify exactly what type of access you need. Scopes limit access for OAuth tokens. They do not grant any additional permission beyond that which the user already has.

### Minimum Scopes

The following two scopes are necessary to implement the integration. Most dashboard shouldn't need more permissions than these scopes enabled.

Scope	Permissions
<code>openid</code>	Allows access to basic data about the user, such as email addresses
<code>cloud_controller_service_permissions.read</code>	Allows access to the CC endpoint that specifies whether the user can manage a given service instance

### Additional Scopes

Dashboards with extended capabilities may need to request these additional scopes:

Scope	Permissions
<code>cloud_controller.read</code>	Allows read access to all resources the user is authorized to read
<code>cloud_controller.write</code>	Allows write access to all resources the user is authorized to update / create / delete

## Reference Implementation

The [MySQL Service Broker](#) is an example of a broker that also implements a SSO dashboard. The login flow is implemented using the [OmniAuth library](#) and a custom [UAA OmniAuth Strategy](#). See this [OmniAuth wiki page](#) for instructions on how to create your own strategy.

The UAA OmniAuth strategy is used to first get an authorization code, as documented in [this section](#) of the UAA documentation. The user is redirected back to the service (as specified by the `callback_path` option or the default `auth/cloudfoundry/callback` path) with the authorization code. Before the application / action is dispatched, the OmniAuth strategy uses the authorization code to [get a token](#) and uses the token to request information from UAA to fill the `omniauth.auth` environment variable. When OmniAuth returns control to the application, the `omniauth.auth` environment variable hash will be filled with the token and user information obtained from UAA as seen in the [Auth Controller](#).

## Restrictions

- UAA clients are scoped to services. There must be a `dashboard_client` entry for each service that uses SSO integration.
- Each `dashboard_client_id` must be unique across the CloudFoundry deployment.

## Resources

- [OAuth](#)

- [Example broker with SSO implementation ↗](#)
- [Cloud Controller API Docs ↗](#)
- [User Account and Authentication \(UAA\) Service APIs ↗](#)

## Example Service Brokers

The following example service broker applications have been developed - these are a great starting point if you are developing your own service broker.

### Ruby

- [GitHub Repository service](#) - this is designed to be an easy-to-read example of a service broker, with complete documentation, and comes with a demo app that uses the service. The broker can be deployed as an application to any Cloud Foundry instance or hosted elsewhere. The service broker uses GitHub as the service back end.
- [MySQL database service](#) - this broker and its accompanying MySQL server are designed to be deployed together as a [BOSH](#) release. BOSH is used to deploy or upgrade the release, monitors the health of running components, and restarts or recreates unhealthy VMs. The broker code alone can be found [here](#).

### Java

- [Spring Cloud - Cloud Foundry Service Broker](#) - This implements the REST contract for service brokers and the artifacts are published to the Spring Maven repository. This greatly simplifies development: include a single dependency in Gradle, implement interfaces, and configure. A sample implementation has been provided for [MongoDB](#).
- [MySQL Java Broker](#) - a Java port of the Ruby-based [MySQL broker](#) above.

### Go

- [Asynchronous Service Broker for AWS EC2](#) - This broker implements support for [Asynchronous Service Operations](#), and calls AWS APIs to provision EC2 VMs.

## Binding Credentials

A bindable service returns credentials that an application can consume in response to the `cf bind` API call. Cloud Foundry writes these credentials to the `VCAP_SERVICES` environment variable. In some cases, buildpacks write a subset of these credentials to other environment variables that frameworks might need.

Choose from the following list of credential fields if possible, though you can provide additional fields as needed. Refer to the [Using Bound Services](#) section of the *Managing Service Instances with the CLI* topic for information about how these credentials are consumed.

 **Note:** If you provide a service that supports a connection string, provide the `uri` key for buildpacks and application libraries to use.

CREDENTIALS	DESCRIPTION
uri	Connection string of the form <code>DB-TYPE://USERNAME:PASSWORD@HOSTNAME:PORT/NAME</code> , where <code>DB-TYPE</code> is a type of database such as mysql, postgres, mongodb, or amqp.
hostname	FQDN of the server host
port	Port of the server host
name	Name of the service instance
vhost	Name of the messaging server virtual host - a replacement for a <code>name</code> specific to AMQP providers
username	Server user
password	Server password

The following is an example output of `ENV['VCAP_SERVICES']`.

 **Note:** Depending on the types of databases you are using, each database might return different credentials.

```
VCAP_SERVICES=
{
 cleardb: [
 {
 name: "cleardb-1",
 label: "cleardb",
 plan: "spark",
 credentials: {
 name: "ad_c6f4446532610ab",
 hostname: "us-cdbr-east-03.cleardb.com",
 port: "3306",
 username: "b5d435f40dd2b2",
 password: "ebfc00ac",
 uri: "mysql://b5d435f40dd2b2:ebfc00ac@us-cdbr-east-03.cleardb.com:3306/ad_c6f4446532610ab",
 jdbcUrl: "jdbc:mysql://b5d435f40dd2b2:ebfc00ac@us-cdbr-east-03.cleardb.com:3306/ad_c6f4446532610ab"
 }
 }
],
 cloudamqp: [
 {
 name: "cloudamqp-6",
 label: "cloudamqp",
 plan: "lemur",
 credentials: {
 uri: "amqp://ksvyjmiv:IwN6dCdZmeQD400ZPKpu1Y0aLx1he8wo@lemur.cloudamqp.com/ksvyjmiv"
 }
 }
],
 {
 name: "cloudamqp-9dbc6",
 label: "cloudamqp",
 plan: "lemur",
 credentials: {
 uri: "amqp://vhuklnxa:9lNFxpTuJsAdTts98vQIdKHW3MojyMyV@lemur.cloudamqp.com/vhuklnxa"
 }
 }
],
 rediscloud: [
 {
 name: "rediscloud-1",
 label: "rediscloud",
 plan: "20mb",
 credentials: {
 port: "6379",
 host: "pub-redis-6379.us-east-1-2.3.ec2.redislabs.com",
 password: "1M5zd3QfWi9nUyya"
 }
 }
]
}
```

```
}
```

## Enabling Service Instance Sharing

This topic provides information about enabling service instance sharing in managed services.

### Overview

Service authors can allow instances of their services to be shared across spaces and orgs within Cloud Foundry. This enables apps running in different spaces and orgs to use the same service instance. Developers with Space Developer permissions in the space where the service instance was created are responsible for sharing, unsharing, updating, and deleting the service instance.

For more information about the developer and administrator tasks related to service instance sharing, see [Sharing Service Instances](#).

### Enabling Service Instance Sharing

Service brokers must explicitly enable service instance sharing by setting a flag in their service-level metadata object. This allows service instances, of any service plan, to be shared across orgs and spaces. The `"shareable"` flag must be set to `true` in the service-level metadata to enable service instance sharing. If the flag is set to `false` or is absent, sharing is disabled.

An example catalog is shown below:

```
{
 "services": [
 {
 "id": "766fa866-a950-4b12-adff-c11fa4cf8fdc",
 "name": "example-service",
 "metadata": {
 "shareable": true
 }
 }
]
}
```

### Binding Permissions Based on Instance Sharing

When a service instance is created in one space and shared into another, developers can bind their apps to the service instance from both spaces.

You may want to have the service broker return credentials with different permissions depending on which space an app is bound from. For example, a messaging service may permit writes from the originating space and only reads from any spaces that the service is shared into.

To determine whether the space of the app is the same as the originating space of the service instance, the service broker can compare the `context.space_guid` and `bind_resource.space_guid` fields in the binding request. The `context.space_guid` field represents the space where the service instance was created, and `bind_resource.space_guid` represents the space of the app involved in the binding.

### Security Considerations

Service authors must consider the following before enabling service instance sharing:

- [Service keys](#) can only be generated by users who have access to the space where the service instance was created. This ensures that only developers with access to this space have visibility into where and how many times the service instance is used.
- Consider the impact of giving out excessive permissions for service bindings, especially bindings that originate from spaces that the service instance has been shared into. For example, a messaging service may permit writes from the originating space and only reads from any shared spaces. For more information, see [Binding Permissions Based on Instance Sharing](#).
- You should generate unique credentials for each binding. This ensures that developers can unshare a service instance at any time. Unsharing an instance deletes any service bindings and revokes access for those credentials. Unsharing an instance prevents unauthorized future access from developers and apps that saved the credentials they were previously provided using the service binding.
- Consider the impact of a service instance dashboard being accessed by users of shared service instances. If authenticating through SSO, see [Dashboard Single Sign-On](#).



## Application Log Streaming

By binding an application to an instance of an applicable service, Cloud Foundry will stream logs for the bound application to the service instance.

- Logs for all apps bound to a log-consuming service instance will be streamed to that instance
- Logs for an app bound to multiple log-consuming service instances will be streamed to all instances

To enable this functionality, a service broker must implement the following:

1. In the [catalog](#) endpoint, the broker must include `requires: syslog_drain`. This minor security measure validates that a service returning a `syslog_drain_url` in response to the [bind](#) operation has also declared that it expects log streaming. If the broker does not include `requires: syslog_drain`, and the bind request returns a value for `syslog_drain_url`, Cloud Foundry will return an error for the bind operation.
2. In response to a [bind](#) request, the broker should return a value for `syslog_drain_url`. The syslog URL has a scheme of syslog, syslog-tls, or https and can include a port number. For example:  
`"syslog_drain_url": "syslog://logs.example.com:1234"`

## How does it work?

1. Service broker returns a value for `syslog_drain_url` in response to bind
2. Loggregator periodically polls an internal Cloud Controller endpoint for updates
3. Upon discovering a new `syslog_drain_url`, Loggregator identifies the associated app
4. Loggregator streams app logs for that app to the locations specified by the service instances' `syslog_drain_url`s

Users can manually configure app logs to be streamed to a location of their choice using User-provided Service Instances. For details, see [Using Third-Party Log Management Services](#).

## Route Services

This documentation is intended for service authors who are interested in offering a service to a Cloud Foundry (CF) services marketplace. Developers interested in consuming these services can read the [Manage Application Requests with Route Services](#) topic.

## Introduction

Cloud Foundry application developers may wish to apply transformation or processing to requests before they reach an application. Common examples of use cases include authentication, rate limiting, and caching services. Route Services are a kind of Marketplace Service that developers can use to apply various transformations to application requests by binding an application's route to a service instance. Through integrations with service brokers and, optionally, with the Cloud Foundry routing tier, providers can offer these services to developers with a familiar, automated, self-service, and on-demand user experience.

**Note:** The procedures in this topic use the Cloud Foundry Command Line Interface (cf CLI). You can also manage route services using Apps Manager. For more information, see the [Manage Route Services](#) section of the *Managing Apps and Service Instances Using Apps Manager* topic.

## Architecture

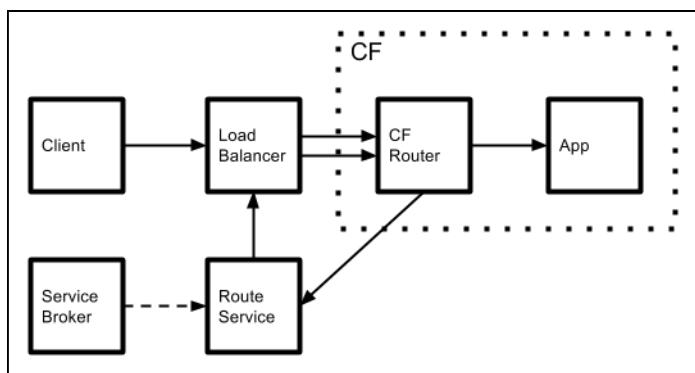
Cloud Foundry supports the following three models for Route Services:

- [Fully-brokered services](#)
- [Static, brokered services](#)
- [User-provided services](#)

In each model, you configure a route service to process traffic addressed to an app.

### Fully-Brokered Service

In the fully-Brokered Service model, the CF router receives all traffic to apps in the deployment before any processing by the route service. Developers can bind a route service to any app, and if an app is bound to a route service, the CF router sends its traffic to the service. After the route service processes requests, it sends them back to the load balancer in front of the CF router. The second time through, the CF router recognizes that the route service has already handled them, and forwards them directly to app instances.



The route service can run inside or outside of CF, so long as it fulfills the [Service Instance Responsibilities](#) to integrate it with the CF router. A service broker publishes the route service to the CF marketplace, making it available to developers. Developers can then create an instance of the service and bind it to their apps with the following commands:

```
cf create-service BROKER-SERVICE-PLAN SERVICE-INSTANCE
```

```
cf bind-route-service YOUR-APP-DOMAIN SERVICE-INSTANCE [--hostname HOSTNAME] [--path PATH]
```

Developers configure the service either through the service provider's web interface or by passing [arbitrary parameters](#) to their `cf create-service` call through the `-c` flag.

**Advantages:**

- Developers can use a Service Broker to dynamically configure how the route service processes traffic to specific applications.
- Adding route services requires no manual infrastructure configuration.
- Traffic to apps that do not use the service makes fewer network hops because requests for those apps do not pass through the route service.

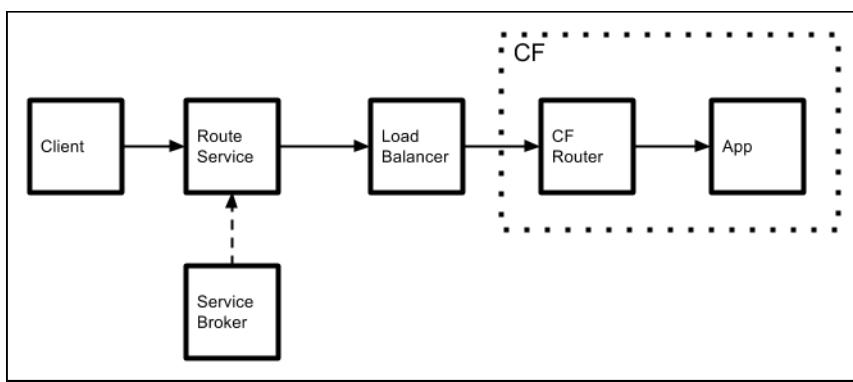
#### Disadvantages:

- Traffic to apps that use the route service makes additional network hops, as compared to the static model.

## Static, Brokered Service

In the static, brokered service model, an operator installs a static routing service, which might be a piece of hardware, in front of the load balancer. The routing service runs outside of Cloud Foundry and receives traffic to all apps running in the CF deployment. The service provider creates a service broker to publish the service to the CF marketplace. As with a [fully-brokered service](#), a developer can use the service by instantiating it with `cf create-service` and

binding it to an app with `cf bind-route-service`.



In this model, you configure route services on an app-by-app basis. When you bind a service to an app, the service broker directs the routing service to process that app's traffic rather than pass the requests through unchanged.

#### Advantages:

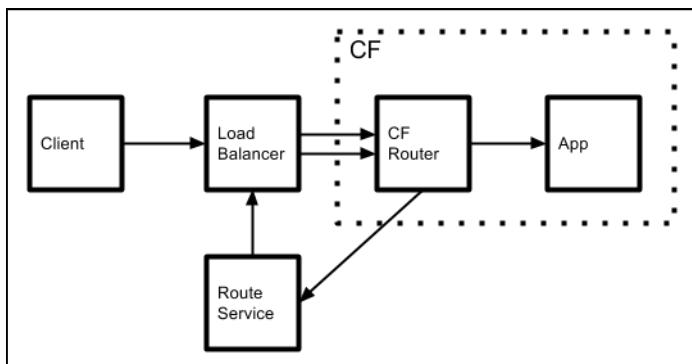
- Developers can use a Service Broker to dynamically configure how the route service processes traffic to specific applications.
- Traffic to apps that use the route service takes fewer network hops.

#### Disadvantages:

- Adding route services requires manual infrastructure configuration.
- Traffic to apps that do not use the route service make unnecessary network hops. Requests for all apps hosted by the deployment pass through the route service component.

## User-Provided Service

If a route service is not listed in the CF marketplace by a broker, a developer can still bind it to their app as a user-provided service. The service can run anywhere, either inside or outside of CF, but it must fulfill the integration requirements described in [Service Instance Responsibilities](#). The service also needs to be reachable by an outbound connection from the CF router.



This model is identical to the [fully-brokered service](#) model, except without the broker. Developers configure the service manually, outside of Cloud Foundry. They can then create a user-provided service instance and bind it to their application with the following commands, supplying the URL of their route service:

```
cf create-user-provided-service SERVICE-INSTANCE -r ROUTE-SERVICE-URL
```

```
cf bind-route-service DOMAIN SERVICE-INSTANCE [--hostname HOSTNAME]
```

#### Advantages:

- Adding route services requires no manual infrastructure configuration.
- Traffic to apps that do not use the service makes fewer network hops because requests for those apps do not pass through the route service.

#### Disadvantages:

- Developers must manually provision and configure route services out of the context of Cloud Foundry because no service broker automates these operations.
- Traffic to apps that use the route service makes additional network hops, as compared to the static model.

## Architecture Comparison

The models above require the [broker](#) and [service instance](#) responsibilities summarized in the following table:

Route Services Architecture	Fulfils CF Service Instance Responsibilities	Fulfils CF Broker Responsibilities
Fully-Brokered	Yes	Yes
Static Brokered	No	Yes
User-Provided	Yes	No

## Enabling Route Services in Pivotal Cloud Foundry

You configure Route Services for your deployment in the PAS tile, under [Settings > Networking](#). For more information, see [Configuring PAS](#).

## Service Instance Responsibilities

The following applies only when a broker returns `route_service_url` in the bind response.

### How It Works

Binding a service instance to a route associates the `route_service_url` with the route in the CF router. All requests for the route are proxied to the URL specified by `route_service_url`.

Once a route service completes its function it may choose to forward the request to the originally requested URL or to another location, or it may choose to reject the request; rejected requests will be returned to the originating requestor. The CF router includes a header that provides the originally requested URL, as well as two headers that are used by the router itself to validate the request sent by the route service. These headers are described below in [Headers](#).

### Headers

The following HTTP headers are added by the Gorouter to requests forwarded to route services.

#### X-CF-Forwarded-Url

The `X-CF-Forwarded-Url` header contains the originally requested URL. The route service may choose to forward the request to this URL or to another.

## X-CF-Proxy-Signature

The `X-CF-Proxy-Signature` header contains an encrypted value which only the Gorouter can decrypt.

The header contains the originally requested URL and a timestamp. When this header is present, the Gorouter will reject the request if the requested URL does not match that in the header, or if a timeout has expired.

`X-CF-Proxy-Signature` also signals to the Gorouter that a request has transited a route service. If this header is present, the Gorouter will not forward the request to a route service, preventing recursive loops. For this reason, route services should not strip off the `X-CF-Proxy-Signature` and `X-CF-Proxy-Metadata` headers.

If the route service forwards the request to a URL different from the originally requested one, and the URL resolves to a route for an application on Cloud Foundry, the route must not have a bound route service or the request will be rejected, as the requested URL will not match the one in the `X-CF-Proxy-Signature` header.

## X-CF-Proxy-Metadata

The `X-CF-Proxy-Metadata` header aids in the encryption and description of `X-CF-Proxy-Signature`.

## SSL Certificates

When Cloud Foundry is deployed in a development environment, certificates hosted by the load balancer are self-signed, and not signed by a trusted certificate authority. When the route service finishes processing an inbound request and makes a call to the value of `X-CF-Forwarded-Url`, be prepared to accept the self-signed certificate when integrating with a non-production deployment of Cloud Foundry.

## Timeouts

Route services must forward the request to the application route within 60 seconds.

In addition, all requests must respond in 900 seconds.

## Broker Responsibilities

### Catalog Endpoint

Brokers must include `requires: ["route_forwarding"]` for a service in the catalog endpoint. If this is not present, Cloud Foundry will not permit users to bind an instance of the service to a route.

### Binding Endpoint

When users bind a route to a service instance, Cloud Foundry sends a [bind request](#) to the broker, including the route address with `bind_resource.route`. A route is an address used by clients to reach apps mapped to the route. The broker may return `route_service_url`, containing a URL where Cloud Foundry should proxy requests for the route. This URL must have an `https` scheme, otherwise the Cloud Controller rejects the binding. `route_service_url` is optional, and not returning this field enables a broker to dynamically configure a network component already in the request path for the route, requiring no change in the CF router.

## Example Route Services

- [Logging Route Service](#): This route service can be pushed as an app to Cloud Foundry. It fulfills the service instance responsibilities above and logs requests received and sent. It can be used to see the route service integration in action by tailing its logs.
- [Rate Limiting Route Service](#): This example route service is a simple Cloud Foundry app that provides rate limiting to control the rate of traffic to an application.
- [Spring Boot Example](#): Logs requests received and sent; written in Spring Boot

## Tutorial

The following instructions show how to use the [Logging Route Service](#) described in [Example Route Services](#) to verify that when a route service is bound to a route, requests for that route are proxied to the route service.

A video of this tutorial is available on [Youtube](#).

These commands require the Cloud Foundry Command Line Interface (cf CLI) version 6.16 or later.

1. Push the [Logging Route Service](#) as an app.

```
$ cf push logger
```

2. Create a user-provided service instance, and include the route of the [Logging Route Service](#) you pushed as `route_service_url`. Be sure to use `https` for the scheme.

```
$ cf create-user-provided-service mylogger -r https://logger.cf.example.com
```

3. Push a sample app like [Spring Music](#). By default this creates a route `spring-music.cf.example.com`.

```
$ cf push spring-music
```

4. Bind the user-provided service instance to the route of your sample app. The `bind-route-service` command takes a route and a service instance; the route is specified in the following example by domain `cf.example.com` and hostname `spring-music`.

```
$ cf bind-route-service cf.example.com mylogger --hostname spring-music
```

5. Tail the logs for your route service.

```
$ cf logs logger
```

6. Send a request to the sample app and view in the route service logs that the request is forwarded to it.

```
$ curl spring-music.cf.example.com
```

## Supporting Multiple Cloud Foundry Instances

It is possible to register a service broker with multiple Cloud Foundry instances. It may be necessary for the broker to know which Cloud Foundry instance is making a given request. For example, when using [Dashboard Single Sign-On](#), the broker is expected to interact with the authorization and token endpoints for a given Cloud Foundry instance.

There are two strategies that can be used to discover which Cloud Foundry instance is making a given request.

## Routing & Authentication

The broker can use unique credentials and/or a unique url for each Cloud Foundry instance. When registering the broker, different Cloud Foundry instances can be configured to use different base urls that include a unique id. For example:

- On Cloud Foundry instance 1, the service broker is registered with the url `broker.example.com/123`
- On Cloud Foundry instance 2, the service broker is registered with the url `broker.example.com/456`

## X-Api-Info-Location Header

All calls to the broker from Cloud Foundry include an `X-Api-Info-Location` header containing the `/v2/info` url for that instance. The `/v2/info` endpoint will return further information, including the location of that Cloud Foundry instance's UAA.

Support for this header was introduced in cf-release v212.

## Logging and Metrics

This documentation describes logging and metrics in Cloud Foundry (CF). It includes topics related to monitoring, event logging, CF data sources, and viewing logs and metrics. It also includes information about the Loggregator system, which aggregates and streams logs and metrics from apps and platform components in Pivotal Application Service (PAS).

## Contents

### Logging and Metrics

- [Overview of Logging and Metrics](#): This topic provides an overview of logging and metrics in CF.
- [Application Logging in Cloud Foundry](#): This topic describes log types and their messages. It also explains how to view logs from the cf CLI.
- [Customizing Platform Log Forwarding](#): This topic describes how to modify log forwarding rules.
- [Configuring the Metric Registrar](#): This topic provides an overview of the Metric Registrar. It also explains how to configure the Metric Registrar for exporting custom app metrics to the Loggregator.
- [Identifying the Source Deployments of Metrics](#): This topic describes how to identify which runtime deployment is the source of metrics.

### Loggregator

- [Loggregator Architecture](#): This topic describes the architecture of the Loggregator system.
- [Loggregator Guide for Cloud Foundry Operators](#): The topic provides information about configuring Loggregator to avoid data loss with high volumes of logging and metrics data.
- [Deploying a Nozzle to the Loggregator Firehose](#): This topic describes deploying a nozzle application to the Loggregator Firehose.
- [Installing the Loggregator Firehose Plugin for cf CLI](#): This topic describes how to use the Loggregator Firehose Plugin for cf CLI to access the output of the Firehose.

## See Also

For more information about logging, monitoring, and reporting, see the topics below.

- [Configuring Logging in PAS](#): This topic describes how to forward system logs to an external service, scale Loggregator component VMs, and manage app traffic logging.
- [Security Event Logging](#): This topic describes how to enable and interpret security event logging for the Cloud Controller, the User Account and Authentication (UAA) server, and CredHub.
- [Monitoring Pivotal Cloud Foundry](#): This topic describes how operators can monitor their deployments. It includes information about selecting and configuring a monitoring system.
- [Reporting Instance Usage with Apps Manager](#): This topic describes how to retrieve app, task, and service instance usage information.
- [Reporting App, Task, and Service Instance Usage](#): This topic describes how to use the cf CLI to retrieve system and org level usage information about apps, tasks, and service instances.
- [Monitoring PCF VMs from Ops Manager](#): This topic describes how to check current VM status.
- [Access Windows Event Logs](#): This topic section describes how to access log messages from Windows Diego cells.

## Overview of Logging and Metrics

This topic provides an overview of logging and metrics in Cloud Foundry (CF). It includes information about logs and metrics sources and transport systems. It also lists products for viewing logs and metrics.

## Sources of Logs and Metrics

There are two sources of CF logs and metrics:

- CF platform components, such as a Diego Cell, MySQL Server, or Cloud Controller
- Apps and app containers deployed on CF

The table below describes the data included in logs and metrics from each source.

Source	Logs Data	Metrics Data
Platform components	Logs from CF components	<ul style="list-style-type: none"> <li>• Health metrics from BOSH-deployed VMs*</li> <li>• Platform metrics from CF components. For example, cell capacity remaining and router throughput.</li> <li>• Metrics for any service tile that self-publishes to the Loggregator Firehose. For example, Redis and MySQL.</li> </ul>
Apps	Logs from apps **	<ul style="list-style-type: none"> <li>• App container metrics. For example, CPU, memory, and disk usage.</li> <li>• Custom metrics ***</li> </ul>

\* For more information about using the BOSH Health Monitor to collect health metrics on VMs, see [Configuring a Monitoring System](#).

\*\* For more information about app logging, see [Application Logging in Cloud Foundry](#).

\*\*\* For more information about configuring an application to stream custom metrics to Loggregator, see [Metrics Forwarder for PCF](#).

## Transport Systems for Logs and Metrics

The following transport systems deliver logs and metrics from their source to an observability product for viewing:

- **Loggregator:** Loggregator is the transport system for both logs and metrics on apps deployed on CF as well as metrics on CF platform components. For more information about the Loggregator system, including Loggregator architecture and components, see [Overview of the Loggregator System](#).
- **rsyslogd on CF component VMs:** rsyslogd is the transport system for CF component logs. Users can configure rsyslogd to transport component logs to a third-party syslog server.

The table below lists the transport system for logs and metrics on CF platform components and apps.

Source	Logs Transport System	Metrics Transport System
Platform components	rsyslogd on CF component VMs	Loggregator
Apps	Loggregator	Loggregator

## Viewing Logs and Metrics

The table below lists the products and tools for viewing CF logs and metrics.

Source	Products and Tools for Viewing Logs	Products and Tools for Viewing Metrics
Platform components	To view system logs from CF components, configure rsyslogd to transport logs to a third-party product.	<p>You can use the following products or tools to view platform component and VM metrics:</p> <ul style="list-style-type: none"> <li>• Loggregator Firehose CLI Plugin. See <a href="#">Installing the Loggregator Firehose Plugin for CLI</a>.</li> <li>• Loggregator Log Cache CLI Plugin. See <a href="#">Cloud Foundry</a></li> </ul>

		<p><a href="#">Community of CLI Plugins ↗</a>.</p> <ul style="list-style-type: none"> <li>• PCF Healthwatch. See <a href="#">PCF Healthwatch ↗</a>.</li> </ul>
Apps	<p>You can use the following products or tools to view app logs:</p> <ul style="list-style-type: none"> <li>• cf CLI cf logs command. See <a href="#">Cloud Foundry CLI Reference Guide. ↗</a></li> <li>• Apps Manager. See <a href="#">Managing Apps and Service Instances Using Apps Manager ↗</a>.</li> <li>• Syslog forwarding. See <a href="#">Streaming Application Logs to Log Management Services ↗</a>.</li> <li>• Loggregator Firehose CLI Plugin. See <a href="#">Installing the Loggregator Firehose Plugin for CLI.</a></li> <li>• Loggregator Log Cache CLI Plugin. See <a href="#">Cloud Foundry Community of CLI Plugins ↗</a>.</li> </ul>	<p>You can use the following products or tools to view app metrics:</p> <ul style="list-style-type: none"> <li>• Loggregator Firehose CLI Plugin. See <a href="#">Installing the Loggregator Firehose Plugin for CLI.</a></li> <li>• Loggregator Log Cache CLI Plugin. See <a href="#">Cloud Foundry Community of CLI Plugins ↗</a>.</li> <li>• PCF Metrics. See <a href="#">PCF Metrics ↗</a>.</li> </ul>

## Loggregator Architecture

This topic describes the Loggregator architecture and components. It also describes the Pivotal Cloud Foundry (PCF) components that send BOSH-reported VM metrics to Loggregator.

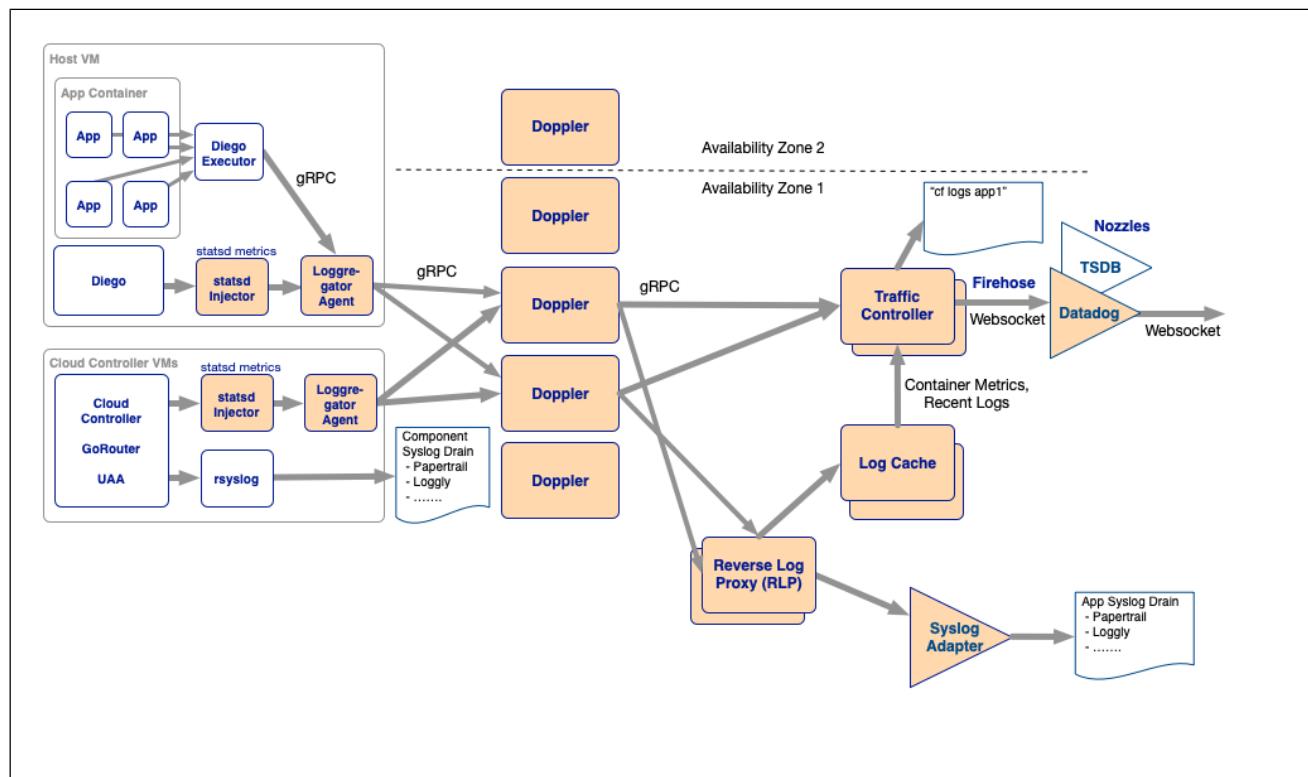
### Overview

Loggregator gathers and streams logs and metrics from user apps in a PCF deployment. It also gathers and streams metrics from PCF components and health metrics from BOSH-deployed VMs. Loggregator allows you to view these logs and metrics through the Loggregator CLI plugins or through a third-party service. For more information, see [Loggregator](#) in GitHub.

The Loggregator architecture includes components for collecting, storing, and forwarding logs and metrics.

### Loggregator Architecture and Components

The diagram below shows the architecture of Loggregator, including the PCF components with which it interacts.



[View a larger version of this image.](#)

The following are Loggregator components, as shown in the diagram above:

- **Loggregator Agent:** Loggregator Agents run on both PCF component VMs and Diego cell VMs. They receive logs and metrics from the apps and PCF components located on those VMs. Loggregator Agents then forward the logs and metrics to Dopplers.
- **System Metrics Agent:** A standalone agent to provide VM system metrics using a Prometheus-scrapable endpoint.
- **Doppler:** Dopplers receive logs and metrics from Loggregator Agents, store them in temporary buffers, and forward them to Traffic Controllers.
- **Traffic Controller:** Traffic Controllers poll Doppler servers for logs and metrics, then translate these messages from the Doppler servers as necessary for external and legacy APIs. The Loggregator Firehose is located on the Traffic Controller.
- **Reverse Log Proxy:** Reverse Log Proxies (RLPs) collect logs and metrics from Dopplers and forward them to Syslog Adapters for syslog drains. Operators can scale up the number of RLPs based on overall log volume.
- **Syslog Adapter:** Syslog Adapters manage connections with and write to syslog drains. Operators can scale the Syslog Adapter component based on the number of syslog drains.
- **Syslog Agents:** When used instead of Syslog Adapters, Syslog Agents increase the number of syslog drain service bindings supported by the

Loggregator system and reduce the workload for Loggregator VMs. Agents run on the Loggregator Trafficcontroller VM, which removes the need for the Syslog Adapter and Syslog Scheduler VM.

- **Syslog Binding Cache:** Syslog Agents can overwhelm CAPI when querying for existing bindings. This component acts as a proxy for the CAPI Binding query.
- **Firehose:** The Firehose is a WebSocket endpoint that streams all the event data from a PCF deployment. The data stream includes HTTP events, app logs, container metrics from apps, and metrics from PCF platform components. The Firehose cf CLI plugin allows you to view the output of the Firehose. For more information about the Firehose plugin, see [Installing the Loggregator Firehose Plugin for cf CLI](#).
- **Log Cache:** The Log Cache allows you to view logs and metrics from the Firehose over a specified period of time. The Log Cache includes API endpoints and a CLI plugin to query and filter logs and metrics. To download the Log Cache CLI plugin, see [Cloud Foundry Plugins](#). The Log Cache API endpoints are available by default. For more information about using the Log Cache API, see [Log Cache](#) on GitHub.
- **Nozzles:** Nozzles are programs that consume data from the Loggregator Firehose. Nozzles can be configured to select, buffer, and transform data, and to forward it to other apps and services. For more information about creating and deploying nozzles, see [Deploying a Nozzle to the Loggregator Firehose](#).

## Related BOSH Components

This section describes the PCF components that forward BOSH-reported VM metrics to Loggregator. BOSH-reported VM metrics measure the health of BOSH-deployed VMs on which apps and PCF components are deployed. Loggregator streams BOSH-reported VM metrics through the Firehose.

The following are the PCF components that send BOSH-reported VM metrics to Loggregator:

- **BOSH Agent:** BOSH Agents are located on component VMs and Diego cell VMs. They collect metrics, such as cell capacity remaining, from the VM and forward them to the BOSH Health Monitor.
- **BOSH Health Monitor:** The BOSH Health Monitor receives metrics from the BOSH Agents. It then forwards the metrics to a third-party service or to the BOSH System Metrics Forwarder.
- **BOSH System Metrics Plugin:** This plugin reads health events, such as VM heartbeats and alerts from the BOSH Health Monitor, and streams them to the BOSH System Metrics Server.
- **BOSH System Metrics Server:** The BOSH System Metrics Server streams metrics and heartbeat events to the BOSH System Metrics Forwarder over gRPC.
- **BOSH System Metrics Forwarder:** The BOSH System Metrics Forwarder is located on the Loggregator Traffic Controller. It forwards heartbeat events from the BOSH System Metrics Server as envelopes to Loggregator through a Loggregator Agent.

## Loggregator Guide for Cloud Foundry Operators

This topic contains information for Cloud Foundry deployments operators about how to configure the [Loggregator system](#) to avoid data loss with high volumes of logging and metrics data.

## Loggregator Message Throughput and Reliability

For determining the message throughput and reliability rates of your Loggregator system, see the section below.

### Measuring Message Throughput

To measure the message throughput of the Loggregator system, you can monitor the total number of egress messages from all Metrons on your platform using the `metron.egress` metric.

If you do not use a monitoring platform, you can follow the instructions below to measure the overall message throughput of your Loggregator system:

1. Log in to the Cloud Foundry Command Line Interface (cf CLI) with your admin credentials:

```
$ cf login
```

2. [Install](#) the Cloud Foundry Firehose plugin.

3. Install Pipe Viewer:

```
$ apt-get install pv
```

4. Run the following command:

```
$ cf nozzle -n | pv -l -i 10 -r > /dev/null
```

### Measuring Message Reliability

To measure the message reliability rate of your Loggregator system, you can run black-box tests. If you want to use this method, see the open-source [cf-logmon](#) app and the configuration instructions provided in the README.md file.

## Scaling Loggregator

Most Loggregator configurations use preferred resource defaults. For more information about customizing these defaults and planning the capacity of your Loggregator system, see [Key Capacity Scaling Indicators](#).

### Scaling Nozzles

Nozzles are programs that consume data from the Loggregator Firehose. Nozzles can be configured to select, buffer, and transform data, and to forward it to other apps and services. You can scale a nozzle using the subscription ID specified when the nozzle connects to the Firehose. If you use the same subscription ID on each nozzle instance, the Firehose evenly distributes data across all instances of the nozzle.

For example, if you have two nozzle instances with the same subscription ID, the Firehose sends half of the data to one nozzle instance and half to the other. Similarly, if you have three nozzle instances with the same subscription ID, the Firehose sends one-third of the data to each instance.

If you want to scale a nozzle, the number of nozzle instances should match the number of Traffic Controller instances:

*Number of nozzle instances = Number of Traffic Controller instances*

Stateless nozzles should handle scaling gracefully. If a nozzle buffers or caches the data, the nozzle author must test the results of scaling the number of nozzle instances up or down.

## Slow Nozzle Alerts

The [Traffic Controller](#) alerts nozzles if they consume events too slowly. If a nozzle falls behind, Loggregator alerts the nozzle in two ways:

- **TruncatingBuffer** alerts: If the nozzle consumes messages more slowly than they are produced, the Loggregator system may drop messages. In this case, Loggregator sends the log message, `TB: Output channel too full. Dropped N messages`, where `N` is the number of dropped messages. Loggregator also emits a `CounterEvent` with the name `doppler_proxy.slow_consumer`. The nozzle receives both messages from the Firehose, alerting the operator to the performance issue.

## Forwarding Logs to an External Service

You can configure Pivotal Application Service (PAS) to forward log data from apps to an external aggregator service. [Using Log Management Services](#) explains how to bind apps to the external service and configure it to receive logs from Pivotal Application Service (PAS).

## Log Message Size Constraints

When a Diego Cell emits app logs to Metron, Diego breaks up log messages greater than approximately 60 KiB into multiple envelopes.

## Deploying a Nozzle to the Loggregator Firehose

This topic describes deploying a nozzle app to the Cloud Foundry (CF) Loggregator Firehose. For more information about nozzles and the Loggregator Firehose, see [Loggregator Architecture](#). The Cloud Foundry Loggregator team created an example nozzle application for use with this tutorial.

The procedure described below deploys this example nozzle to the Firehose of a Cloud Foundry installation deployed locally with BOSH Lite. For more information about BOSH Lite, see the [BOSH Lite GitHub repository](#).

To lighten the load on custom nozzles you develop, you can request Firehose subscriptions that emit only whitelisted metrics. For examples, see [rlpreader](#) and [rltypereader](#) in the [loggregator-tools](#) repository, and see the *V2 Subscriptions* page of the [loggregator-release](#) repository for more information.

## Prerequisites

Before you deploy a nozzle to the Loggregator Firehose, you must have the following:

- BOSH CLI. See [BOSH CLI](#) installed locally.
- Spiff installed locally and added to the load path of your shell. See [Spiff on GitHub](#).
- BOSH Lite deployed locally using VirtualBox. See [BOSH Lite on GitHub](#).
- A working Cloud Foundry deployment, including Loggregator, deployed with your local BOSH Lite. This serves as the source of data. See [Deploying Cloud Foundry using BOSH Lite](#), or use the `provision_cf` script included in the BOSH Lite release. See [BOSH Lite](#) on GitHub.

 **Note:** Deploying Cloud Foundry can take several hours depending on your internet bandwidth, even when using the automated `provision_cf` script.

## Step 1: Download Cloud Foundry BOSH Manifest

To download the BOSH manifest, do the following:

1. Run `bosh -e MY-ENV deployments` to identify the names of all deployments in the environment you specify. Replace `MY-ENV` with the alias you set for your BOSH Director. For example:

```
$ bosh -e dev deployments
Using environment '192.168.15.4' as client 'admin'

Name Release(s) Stemcell(s) Team(s) Cloud Config
cf-example binary-buildpack/1.0.9 bosh-warden-boshlite-ubuntu-trusty-go_agent/3363.9 - latest
 capi/1.21.0
 cf-mysql/34
 cf-smoke-tests/11
 cflinuxfs3-rootfs/1.52.0
 consul/155
 diego/1.8.1
 etcd/94
 garden-runc/1.2.0
 loggregator/78
 nats/15
 routing/0.145.0
 statsd-injector/1.0.20
 uaa/25
service-instance_0d4140a-42b7-... mysql/0.6.0 bosh-warden-boshlite-ubuntu-trusty-go_agent/3363.9 - latest

2 deployments

Succeeded
```

2. Run `bosh -e MY-ENV -d MY-DEPLOYMENT manifest > MY-MANIFEST.yml` to download and save the current BOSH deployment manifest. Replace `MY-ENV` with your BOSH Director alias, `MY-DEPLOYMENT` with the deployment name from the output of the previous step, and `MY-MANIFEST.yml` with a name you choose for the saved manifest file. You need this manifest to locate information about your databases.

```
$ bosh -e dev -d cf-example manifest > cf.yml
```

## Step 2: Add UAA Client

You must authorize the example nozzle as a UAA client for your CF deployment. To do this, add an entry for the example nozzle as `client` for `uaa` under the `properties` key in your CF deployment manifest YAML file. You must enter the example nozzle object in the correct location in the manifest, and with the correct indentation.

To add the nozzle as a UAA client for your deployment, do the following:

1. Open the deployment manifest in a text editor.
2. Locate the left-aligned `properties` key.
3. Under the `properties` key, locate `uaa` at the next level of indentation.
4. Under the `uaa` key, locate the `clients` key at the next level of indentation.
5. Enter properties for the `example-nozzle` at the next level of indentation, exactly as shown below. The `...` in the text below indicate other properties that may populate the manifest at each level in the hierarchy.

```
properties:
...
uaa:
...
clients:
...
example-nozzle:
 access-token-validity: 1209600
 authorized-grant-types: client_credentials
 override: true
 secret: example-nozzle
 authorities: oauth.login,doppler.firehose
```

6. Save the deployment manifest file.

## Step 3: Redeploy Cloud Foundry

To deploy Cloud Foundry with BOSH, run the following command:

```
bosh -e MY-ENV deploy
```

WHERE `MY-ENV` is the alias you set for your BOSH Director.

For example:

```
$ bosh -e dev deploy
Acting as user 'admin' on deployment 'cf-warden' on 'Bosh Lite Director'
Getting deployment properties from director...
Detecting deployment changes

Releases
No changes

Compilation
No changes

Update
No changes

Resource pools
No changes

Disk pools
No changes

Networks
No changes

Jobs
No changes

Properties
uaa
clients
example-nozzle
+ access-token-validity: 1209600
+ authorized-grant-types: authorization_code,client_credentials,refresh_token
+ override: true
+ secret: example-nozzle
+ scope: openid,oauth.approvals,doppler.firehose
+ authorities: oauth.login,doppler.firehose

Meta
No changes

Please review all changes carefully
Deploying

Are you sure you want to deploy? (type 'yes' to continue):yes
```

## Step 4: Clone Example Release

The Cloud Foundry Loggregator team created an example nozzle app for use with this tutorial.

To clone the example nozzle release, do the following:

1. Run `git clone` to clone the main release repository from [GitHub](#).

```
$ git clone https://github.com/cloudfoundry-incubator/example-nozzle-release.git
Cloning into 'example-nozzle-release'...
```

2. Navigate to the `example-nozzle-release` directory.

```
$ cd example-nozzle-release
```

3. Run `git submodule update --init --recursive` to update all of the included submodules.

```
$ git submodule update --init --recursive
Submodule 'src/github.com/cloudfoundry-incubator/example-nozzle' (git@github.com:cloudfoundry-incubator/example-nozzle.git) registered for path 'src/github.com/cloudfoundry-incubator/example-nozzle'
Submodule 'src/github.com/cloudfoundry-incubator/uaago' (git@github.com:cloudfoundry-incubator/uaago.git) registered for path 'src/github.com/cloudfoundry-incubator/uaago'
...
Cloning into 'src/github.com/cloudfoundry-incubator/example-nozzle'...
...
```

## Step 5: Prepare Nozzle Manifest

To prepare the nozzle deployment manifest, do the following:

1. In the `example-nozzle-release` directory, navigate to the `templates` directory.

```
$ cd templates
```

Within this directory, examine the two YAML files. `bosh-lite-stub.yml` contains the values used to populate the missing information in `template.yml`. By combining these two files we create a deployment manifest for our nozzle.

2. Create a `tmp` directory for the compiled manifest.
3. Use [Spiff](#) to compile a deployment manifest from the template and stub, and save this manifest.

```
$ spiff merge templates/template.yml templates/bosh-lite-stub.yml > tmp/manifest_bosh_lite.yml
```

4. Run `bosh -e MY-ENV deployments` to identify the names of all deployments in the environment you specify. Replace `MY-ENV` with the alias you set for your BOSH Director.
5. Run `bosh -e MY-ENV env --column=uuid` to obtain your BOSH Director UUID. Replace `MY-ENV` with the alias you set for your BOSH Director. For example:

```
$ bosh -e dev env --column=uuid
```

6. In the compiled nozzle deployment manifest, locate the `director_uuid` property. Replace `PLACEHOLDER-DIRECTOR-UUID` with your BOSH Director UUID.

```
compilation:
cloud_properties:
 name: default
network: example-nozzle-net
reuse_compilation_vms: true
workers: 1
director_uuid: PLACEHOLDER-DIRECTOR-UUID # replace this
```

**Note:** If you do not want to see the complete deployment procedure, run the following command to automatically prepare the manifest:  
`scripts/make_manifest_spiff_bosh_lite`

## Step 6: Create Nozzle BOSH Release

To create a nozzle BOSH release, run the following command:

```
bosh -e MY-ENV create-release --name RELEASE-NAME
```

WHERE:

- `MY-ENV` is the alias you set for your BOSH Director.
- `RELEASE-NAME` is `example-nozzle` to match the UAA client that you created in the CF deployment manifest.

For example:

```
$ bosh -e dev create-release --name example-nozzle
Syncing blobs...
...
```

## Step 7: Upload Nozzle BOSH Release

Upload the nozzle BOSH release that you created as part of [Step 6: Create Nozzle BOSH Release](#).

To upload the BOSH release, run the following command:

```
bosh -e MY-ENV upload-release
```

WHERE `MY-ENV` is the alias you set for your BOSH Director.

For example:

```
$ bosh -e dev upload-release
Acting as user 'admin' on 'Bosh Lite Director'

Copying packages

example-nozzle
golang1.7

Copying jobs

example-nozzle

Generated /var/folders/4n/qs1rjbmd1c5gb78m3_06j6r0000gn/T/d20151009-71219-17a5m49/d20151009-71219-rts928/release.tgz
Release size: 59.2M

Verifying release...
...
Release info

Name: nozzle-test
Version: 0+dev.2

Packages
-
- example-nozzle (b0944f95cb5a332c9be2adfb4db1bc88f9755894)
- golang1.7 (b68dc9557ef296cb21e577c31ba97e2584a5154b)

Jobs
-
- example-nozzle (112e01c6ee91e8b268a42239e58e8e18e0360f58)

License
-
- none

Uploading release
```

## Step 8: Deploy Nozzle

To deploy the nozzle, run the following command:

```
bosh -e MY-ENV deploy
```

WHERE `MY-ENV` is the alias you set for your BOSH Director.

For example:

```
$ bosh -e dev deploy
Acting as user 'admin' on deployment 'example-nozzle-lite' on 'Bosh Lite Director'
Getting deployment properties from director...
Unable to get properties list from director, trying without it...
Cannot get current deployment information from director, possibly a new deployment
Please review all changes carefully

Deploying

Are you sure you want to deploy? (type 'yes' to continue):yes
```

## Step 9: View Nozzle Output

The example nozzle outputs all of the data originating coming from the Firehose to its log files. To view this data, SSH into the example-nozzle VM and examine the logs.

To view nozzle output, do the following:

1. Run `bosh -e MY-ENV ssh` to access the nozzle VM at the IP configured in the nozzle's manifest template stub `./templates/bosh-lite-stub.yml`. Replace `MY-ENV` with the alias you set for your BOSH Director. For example:

```
$ bosh -e dev ssh example-nozzle
Welcome to Ubuntu 14.04.1 LTS (GNU/Linux 3.19.0-25-generic x86_64)

Documentation: https://help.ubuntu.com/
Last login: Wed Sep 23 21:29:50 2015 from 192.0.2.1
```

2. Use the `cat` command to output the `stdout` log file.

```
$ cat /var/vcap/sys/log/example-nozzle/example-nozzle.stdout.log
===== Streaming Firehose (will only succeed if you have admin credentials)
origin:"bosh-system-metrics-forwarder" eventType:ValueMetric timestamp:154109185100000000 deployment:"cf-c42ae2c4dfb6f67b6c27" job:"loggregator_trafficcontroller" index:"d10" ip:"192.0.2.1" type:"firehose"
origin:"gorouter" eventType:ValueMetric timestamp:1541091851218590916 deployment:"cf-c56ab7c4dfb6f67b6c28" job:"router" index:"d5d1b5a4-2497-4679-8d3b-66ffc978d829" ip:"192.0.2.1" type:"firehose"
origin:"netmon" eventType:ValueMetric timestamp:1541091851234217334 deployment:"cf-c56ab7c4dfb6f67b6c28" job:"diego_cell" index:"8007afda-3bff-4856-857f-a47a43cbf994" ip:"192.0.2.1" type:"firehose"
origin:"netmon" eventType:ValueMetric timestamp:1541091851234129669 deployment:"cf-c56ab7c4dfb6f67b6c28" job:"diego_cell" index:"8007afda-3bff-4856-857f-a47a43cbf994" ip:"192.0.2.1" type:"firehose"
origin:"netmon" eventType:ValueMetric timestamp:1541091851234292367 deployment:"cf-c56ab7c4dfb6f67b6c28" job:"diego_cell" index:"8007afda-3bff-4856-857f-a47a43cbf994" ip:"192.0.2.1" type:"firehose"
origin:"netmon" eventType:ValueMetric timestamp:1541091851234279470 deployment:"cf-c56ab7c4dfb6f67b6c28" job:"diego_cell" index:"8007afda-3bff-4856-857f-a47a43cbf994" ip:"192.0.2.1" type:"firehose"
...
...
```

## Installing the Loggregator Firehose Plugin for cf CLI

The Loggregator Firehose plugin for the Cloud Foundry Command Line Interface (cf CLI) allows Cloud Foundry (CF) administrators to access the output of the Loggregator Firehose. The output of the Firehose includes logs and metrics from apps deployed on CF as well as metrics from CF platform components. For more information about the Firehose, see [Overview of the Loggregator System](#).

For more information about using plugins with the cf CLI, see [Using cf CLI Plugins](#).

### Prerequisites

- Administrator access to the Cloud Foundry deployment that you want to monitor
- Cloud Foundry Command Line Interface (cf CLI) 6.12.2 or later

Refer to the [Installing the cf CLI](#) topic for information about downloading, installing, and uninstalling the cf CLI.

### Install the Plugin

1. Run `cf add-plugin-repo REPOSITORY-NAME URL` to add the CF Community plugin repository to your cf CLI plugins.

```
$ cf add-plugin-repo CF-Community https://plugins.cloudfoundry.org
```

2. Run `cf install-plugin -r PLUGIN-REPOSITORY PLUGIN-NAME` to install the Firehose plugin from the CF Community plugin repository.

```
$ cf install-plugin -r CF-Community "Firehose Plugin"
```

### View the Firehose

Run `cf nozzle --debug` to view the streaming output of the Firehose, which includes logging events and metrics from CF system components. For more information about logging and metrics in CF, see [Overview of the Loggregator System](#).

```
$ cf nozzle --debug
```

 **Note:** You must be logged in as a Cloud Foundry administrator to access the Firehose.

### Uninstall the Plugin

Run `cf plugins` to see a list of installed plugins.

```
$ cf plugins
Listing Installed Plugins...
OK
Plugin Name Version Command Name Command Help
FirehosePlugin 0.6.0 nozzle Command to print out messages from the firehose
```

Run `cf uninstall-plugin PLUGIN-NAME` to uninstall the plugin.

```
$ cf uninstall-plugin FirehosePlugin
```

## App Logging in Cloud Foundry

Loggregator, the Cloud Foundry component responsible for logging, provides a stream of log output from your app and from Cloud Foundry system components that interact with your app during updates and execution.

By default, Loggregator streams logs to your terminal. If you want to persist more than the limited amount of logging information that Loggregator can buffer, you can drain logs to a third-party log management service. See [Third-Party Log Management Services](#).

Cloud Foundry gathers and stores logs in a best-effort manner. If a client cannot consume log lines quickly enough, the Loggregator buffer may need to overwrite some lines before the client has consumed them. A syslog drain or a CLI tail can usually keep up with the flow of app logs.

## Contents of a Log Line

Every log line contains four fields:

- Timestamp
- Log type (origin code)
- Channel: either `OUT`, for logs emitted on `stdout`, or `ERR`, for logs emitted on `stderr`
- Message

Loggregator assigns the timestamp when it receives log data. The log data is opaque to Loggregator, which simply puts it in the message field of the log line. Apps or system components sending log data to Loggregator may include their own timestamps, which then appear in the message field.

Origin codes distinguish the different log types. Origin codes from system components have three letters. The app origin code is `APP` followed by slash and a digit that indicates the app instance.

Many frameworks write to an app log that is separate from `stdout` and `stderr`. This is not supported by Loggregator. Your app must write to `stdout` or `stderr` for its logs to be included in the Loggregator stream. Check the buildpack your app uses to determine whether it automatically ensures that your app correctly writes logs to `stdout` and `stderr` only. Some buildpacks do this, and some do not.

## Log Types and Their Messages

Different types of logs have different message formats, as shown in the examples below. The digit appended to the code indicates the instance index: 0 is the first instance, 1 is the second, and so on.

### API

Users make API calls to request changes in app state. Cloud Controller, the Cloud Foundry component responsible for the API, logs the actions that Cloud Controller takes in response.

For example:

```
2016-06-14T14:10:05.36-0700 [API/0] OUT Updated app with guid cdabc600-0b73-48e1-b7d2-26af2c63f933 ({ "name":>"spring-music", "instances":>1, "memory":>512, "environment_js
```

### STG

The Diego cell or the Droplet Execution Agent emits STG logs when staging or restaging an app. These actions implement the desired state requested by the user. After the droplet has been uploaded, STG messages end and CELL messages begin. For STG, the instance index is almost always 0.

For example:

```
2016-06-14T14:10:27.91-0700 [STG/0] OUT Staging...
```

### RTR

The Router emits RTR logs when it routes HTTP requests to the app. Router messages include the app name followed by a Router timestamp and then selections from the HTTP request.

For example:

```
2016-06-14T10:51:32.51-0700 [RTR/1] OUT www.example.com - [14/06/2016:17:51:32.459 +0000] "GET /user/ HTTP/1.1" 200 0 103455 "-" "Mozilla/5.0 (Windows NT 6.1; WOW64) Ap
```

## Zipkin Trace Logging

If Zipkin trace logging is enabled in Cloud Foundry, then Gorouter access log messages contain Zipkin HTTP headers.

The following is an example access log message containing Zipkin headers:

```
2016-11-23T16:04:01.49-0800 [RTR/0] OUT www.example.com - [24/11/2016:00:04:01.227 +0000] "GET / HTTP/1.1" 200 0 109 "-" "curl/7.43.0" 10.0.2.150:4070 10.0.48.66:60815 x_forv
```

For more information about Zipkin tracing, see [Zipkin Tracking in HTTP Headers](#).

## LGR

Loggregator emits LGR to indicate problems with the logging process. Examples include “can’t reach syslog drain url” and “dropped log messages due to high rate.”

## APP

Every app emits logs according to choices by the developer.

For example:

```
2016-06-14T14:10:15.18-0700 [APP/0] OUT Exit status 0
```

## SSH

The Diego cell emits SSH logs when a user accesses an app container through SSH by using the Cloud Foundry Command Line Interface (cf CLI)[cf ssh](#) command.

For example:

```
2016-06-14T14:16:11.49-0700 [SSH/0] OUT Successful remote access by 192.0.2.33:7856
```

## CELL

The Diego cell emits CELL logs when it starts or stops the app. These actions implement the desired state requested by the user. The Diego cell also emits messages when an app crashes.

For example:

```
2016-06-14T13:44:38.14-0700 [CELL/0] OUT Successfully created container
```

## Writing to the Log from Your App

Your app must write logs to `stderr` or `stdout`. Both are typically buffered, and you should flush the buffer before delivering the message to Loggregator.

Alternatively, you can write log messages to `stderr` or `stdout` synchronously. This approach is mainly used for debugging because it may affect app

performance.

## Including Container Metrics in Syslog Drains

By default, app logs are included in syslog drains. To include container metrics in your syslog drain, an operator must first set `scalablesyslog.adapter.metrics_to_syslog_enabled` to `true`. In Pivotal Application Service (PAS) 2.1 and later, this property is set to true by default.

Follow the instructions below to include container metrics in your syslog drain.

1. Log in to Ops Manager.
2. Go to the **System Logging** pane, in the PAS tile.
3. Enable **Include container metrics in Syslog Drains**.
4. Run the following cf CLI command:

```
cf drain APP-NAME DRAIN-URL --type metrics --drain-name YOUR-SYSLOG-DRAIN
```

Where:

- o `APP-NAME` is your app name
- o `YOUR-SYSLOG-DRAIN` is the name of your syslog drain
- o `DRAIN-URL` is the URL of your syslog drain

## Viewing Logs in the Command Line Interface

Use the cf CLI [cf logs](#) command to view logs. You can tail, dump, or filter log output.

### Tailing Logs

Run `cf logs APP-NAME` to stream Loggregator output to the terminal. Replace `APP-NAME` with the name of your app.

For example:

```
$ cf logs spring-music
Connected, tailing logs for app spring-music in org example / space development as admin@example.com...
2016-06-14T15:16:12.70-0700 [RTR/4] OUT www.example.com - [14/06/2016:22:16:12.582 +0000] "GET / HTTP/1.1" 200 0 103455 "-" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10_10_5
2016-06-14T15:16:20.06-0700 [RTR/4] OUT www.example.com - [14/06/2016:22:16:20.034 +0000] "GET /test/ HTTP/1.1" 200 0 6879 "http://www.example.com/" "Mozilla/5.0 (Macintosh
2016-06-14T15:16:22.44-0700 [RTR/4] OUT www.example.com - [14/06/2016:22:17:22.415 +0000] "GET /test5/ HTTP/1.1" 200 0 5461 "http://www.example.com/test5" "Mozilla/5.0 (Mac
...

```

Use **Ctrl-C (^C)** to exit the real-time stream.

### Dumping Logs

Run `cf logs APP-NAME --recent` to display all the lines in the Loggregator buffer. Replace `APP-NAME` with the name of your app.

### Filtering Logs

To view some subset of log output, run `cf logs APP-NAME` in conjunction with filtering commands of your choice. Replace `APP-NAME` with the name of your app. In the example below, `grep -v RTR` excludes all Router logs:

```
$ cf logs spring-music --recent | grep -v RTR
2016-06-14T14:10:05.36-0700 [API/0] OUT Updated app with guid cdabc604-0b73-47e1-a7d5-24af2c63f723 ({ "name": "spring-music", "instances": 1, "memory": 512, "environment_js": null })
2016-06-14T14:10:14.52-0700 [APP/0] OUT - Gracefully stopping, waiting for requests to finish
2016-06-14T14:10:14.52-0700 [CELL/0] OUT Exit status 0
2016-06-14T14:10:14.54-0700 [APP/0] OUT === puma shutdown: 2016-06-14 21:10:14 +0000 ===
2016-06-14T14:10:14.54-0700 [APP/0] OUT - Goodbye!
2016-06-14T14:10:14.56-0700 [CELL/0] OUT Creating container
...
...
```

## Log Ordering

Ensuring log ordering in drains can be an important consideration for both operators and developers.

- Diego uses a nanosecond-based timestamp that can be ingested properly by Splunk. For more information, see [TIME\\_FORMAT and subseconds](#) in the Splunk documentation.
- The Elastic Stack can ingest the nanosecond timestamps but only supports millisecond precision, so timestamps are truncated. For more information, see the [Date datatype](#) in the Elasticsearch documentation and the [Date type has not enough precision for the logging use case](#) GitHub issue.

If you are developing a client that displays a stream of Cloud Foundry logs to users, you can order the logs to improve the debugging experience for your user. The following are general tips for ordering logs:

- For CLIs, batch the logs and display the logs you have in that timeframe, sorted by timestamp.
- For web clients, use dynamic HTML to insert older logs into the sorting as they appear. This creates complete, ordered logs.
- Java app developers may want to convert stack traces into a single log entity. To simplify log ordering for Java apps, use the [multi-line Java message workaround](#) to convert your multi-line stack traces into a single log entity.

By modifying the Java log output, you can force your app to reformat stack trace messages, replacing newline characters with a token. Set your log parsing code to replace that token with newline characters again to display the logs properly in Kibana.

## Customizing Platform Log Forwarding

You can configure Pivotal Application Service (PAS) to forward logs to remote endpoints using the syslog protocol defined in [RFC 5424](#). For more information about enabling log forwarding, see the *Configure System Logging* section of [Configuring PAS](#).

PAS annotates forwarded messages with structured data. This structured data identifies the originating BOSH Director, deployment, instance group, availability zone, and instance ID. All logs forwarded from BOSH jobs have their PRI set to 14, representing **Facility: user-level messages** and **Warning: warning conditions**, as defined in [RFC 5424 Section 6.2.1](#). This PRI value may not reflect the originally intended PRI of the log.

Logs forwarded from other sources, such as kernel logs, retain their original PRI value.

The following table describes the log line Structured Data:

Structured Data	Description
ENTERPRISE_NUMBER	Cloud Foundry's private enterprise number, 47450, as defined in <a href="#">RFC 5424 Section 7.2.2</a> .
DIRECTOR	The name of the BOSH Director managing the deployment.
DEPLOYMENT	The name of the BOSH deployment.
INSTANCE_GROUP	The name of the BOSH instance_group.
AVAILABILITY_ZONE	The name of the BOSH availability zone.
ID	The BOSH GUID.

Log lines use the format below:

```
<$PRI>$VERSION $TIMESTAMP $HOST $APP_NAME $PROC_ID $MSG_ID
[instance@ENTERPRISE_NUMBER director="$DIRECTOR" deployment="$DEPLOYMENT"
group="$INSTANCE_GROUP" az="$AVAILABILITY_ZONE" id="$ID"] $MESSAGE
```

Example log messages:

```
<14>1 2017-01-25T13:25:03.18377Z 192.0.2.10 etcd -- [instance@47450
director="test-env" deployment="cf-c42ae2c4dfb6f67b6c27" group="diego_database" az="us-west1-a"
id="83bd66e5-3fdf-44b7-bdd6-508deae7c786"] [INFO] the leader is
[https://diego-database-0.etcd.service.cf.internal:4001]
<14>1 2017-01-25T13:25:03.184491Z 192.0.2.10 bbs -- [instance@47450
director="test-env" deployment="cf-c42ae2c4dfb6f67b6c27" group="diego_database" az="us-west1-a"
id="83bd66e5-3fdf-44b7-bdd6-508deae7c786"]
{"timestamp": "1485350702.539694548", "source": "bbs", "message":
"bbs.request.start-actual-lrp.starting", "log_level": 1, "data":
{"actual_lrp_instance_key": {"instance_guid":
"271f71c7-4119-4490-619f-4f44694717c0", "cell_id":
"diego_cell-2-41f21178-d619-4976-901c-325bc2d0d11d"}, "actual_lrp_key":
{"process_guid": "1545ce89-01e6-4b8f-9cb1-5654a3cae10-137e7eb4-12de-457d-
8e3e-1258e5a74687", "index": 5, "domain": "cf-apps"}, "method": "POST", "net_info":
{"address": "192.0.2.12", "ports": [{"container_port": 8080, "host_port": 61532}, {"container_port": 2222, "host_port": 61533}], "request":
"/v1/actual_lrps/start", "session": "418.1"}}}
```

## Modify Which Logs PAS Forwards

When you enable log forwarding, PAS forwards all log lines written to the /var/vcap/sys/log directories on all Cloud Foundry virtual machines (VMs) to your configured External Syslog Aggregator endpoint by default.

You can configure PAS to forward a subset of logs instead of forwarding all logs as follows.

- In the PAS tile, select **System Logging**.
- In the **Custom rsyslog Configuration** textbox, enter a custom syslog rule. See the example custom syslog rules below.
- Click **Save**.

The custom rsyslog rules shown below are written in [RainerScript](#). The custom rules are inserted before the rule that forwards logs. The stop command, stop, prevents logs from reaching the forwarding rule. This filters out these logs.

Logs filtered out before forwarding remain on the local disk, where the BOSH job originally wrote them. These logs remain on the local disk only until

BOSH Director recreates the VMs. You can access these logs from Ops Manager or through SSH.

 **Note:** PAS requires a valid custom rule to forward logs. If your custom rule contains syntax errors, PAS forwards no logs.

## Forward Only Logs From a Certain Job

This rule filters out logs unless they originate from the `uaa` job:

```
if ($app-name != "uaa") then stop
```

## Exclude Logs With Certain Content

This rule filters out logs that contain “DEBUG” in the body.

```
if ($msg contains "DEBUG") then stop
```

 **Note:** The above example contains “DEBUG” in the message body. Not all logs intended for debugging purposes contain the string “DEBUG” in the message body.

## Configuring the Metric Registrar

This topic describes the Metric Registrar for Pivotal Application Service (PAS). It also includes information about enabling and configuring the Metric Registrar.

### Overview

The Metric Registrar allows app developers to export custom app metrics in a format that Loggregator can consume. App developers can then use the custom metrics to monitor apps with PCF Metrics and configure autoscaling rules with PCF Autoscaler.

App developers can export custom metrics to Loggregator by configuring their apps in one of the following ways:

- **Metrics Endpoint**—Publish and register a Prometheus Exposition metrics endpoint to an app. The Metric Registrar will then poll this endpoint every 35 seconds and convert the metrics found in the response to Loggregator metrics.
- **Structured Log**—Modify your app to emit metrics using a specific JSON or DogStatsD format. The Metric Registrar then converts all matching log messages into Loggregator metrics or events.

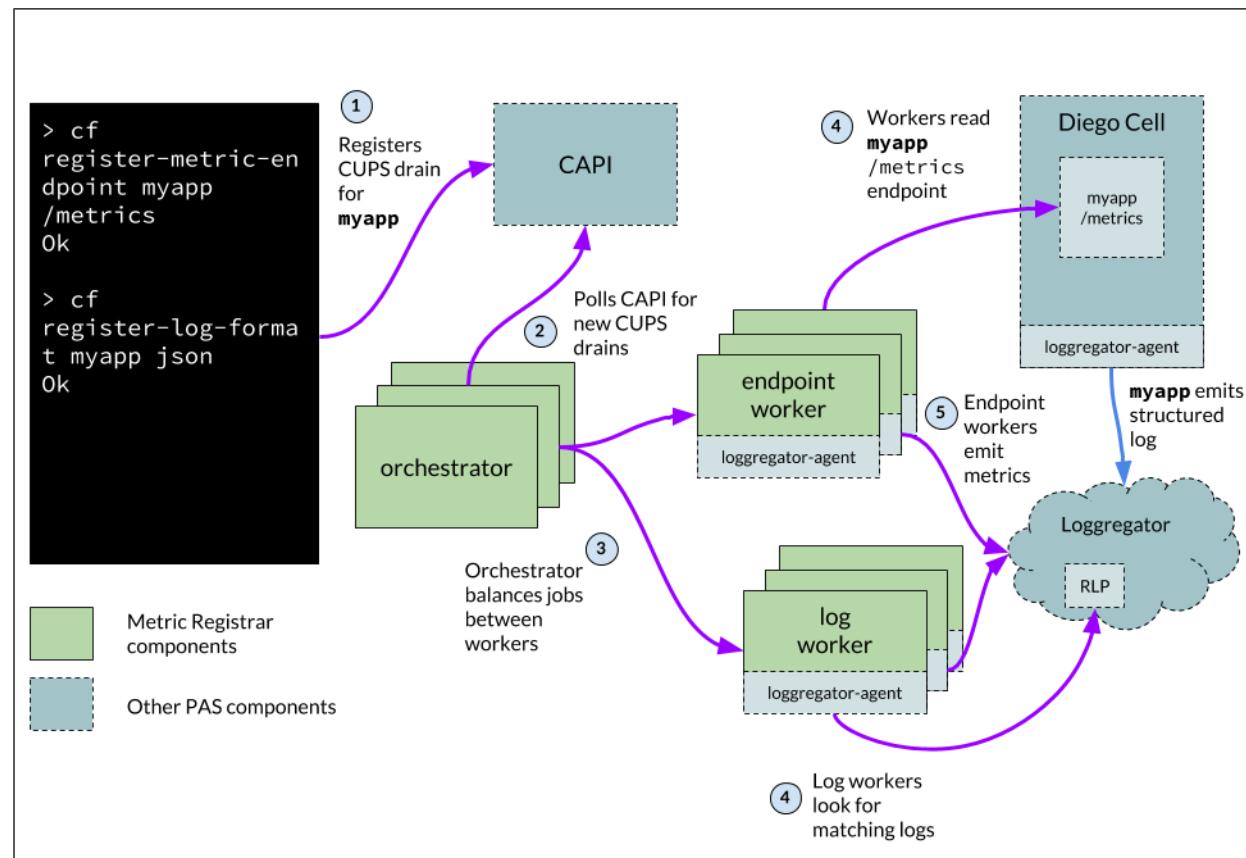
For more information about installing the Metric Registrar Plugin and registering your app, see [Emitting Custom App Metrics to the Metric Registrar](#).

For more information about the components and products mentioned, see the following: - [Loggregator](#) - [PCF Metrics](#) - [PCF Autoscaler](#)

### Architecture

The following diagram illustrates how the Metric Registrar sends your custom app metrics to Loggregator. The components of the Metric Registrar are as follows:

- The cf CLI plugin
- The `metric_registrar_endpoint_worker` and `metric_registrar_log_worker` jobs running on the Doppler VM of the PAS deployment
- The `metric_registrar_orchestrator` and `metric_registrar_smoke_test` jobs running on the Clock Global VM of the PAS deployment



Click the image for a larger representation.

## Configure the Metric Registrar

The PAS tile includes a **Metric Registrar** configuration pane. You can configure the Metric Registrar from this pane in the following ways:

- Enable / Disable
- Edit default scraping interval
- Add blacklisted tags

### Enable and Disable the Metric Registrar

The PAS tile deploys the Metric Registrar by default.

To disable the Metric Registrar, do the following:

1. In the PAS tile, click **Metric Registrar**.
2. Clear the **Enable Metric Registrar** checkbox.

### Edit Default Scraping Interval

The scraping interval defines how often the Metric Registrar polls custom metric endpoints. The default is 35 seconds.

To edit the Metric Registrar scraping interval, do the following:

1. In the PAS tile, click **Metric Registrar**.
2. Edit the **Endpoint Scraping Interval** field.

### Add Blacklisted Tags

To prevent the Metric Registrar from consuming the value of a metric or event tag, you can add the tag to the **Blacklisted Tags** field. For example, if you tag your metrics with a `customer_id`, you may want to add `customer_id` to the list of blacklisted tags. The field accepts a comma-separated list.

By default, the following tags are blacklisted to prevent interference with other products like PCF Metrics that use and rely on such tags.

- `deployment`
- `job`
- `index`
- `id`

To prevent the Metric Registrar from consuming the value of a metric or event tag, do the following:

1. In the PAS tile, click **Metric Registrar**.
2. Add the desired tag to the **Blacklisted Tags** field.

## Registering and Emitting Custom App Metrics

This topic describes how to install the Metric Registrar CLI plugin and emit custom app metrics to the Metric Registrar by registering your app as a metric source.

### Overview

Registering your app as a metric source allows you to see your custom metrics in PCF Metrics and configure autoscaling rules with PCF Autoscaler.

For more information, see [PCF Metrics](#) and [PCF Autoscaler](#).

### Install the Plugin

To install the Metric Registrar CLI plugin, do the following:

1. Log in to the Cloud Foundry Command Line Interface (cf CLI).
2. Run the following command:

```
cf install-plugin -r CF-Community "metric-registrar"
```

### Register Your App

To register your app as a metric source, do one of the following:

- **Register a public endpoint:** Register a public endpoint for custom metrics to be parsed and emitted to Loggregator. See [Register a Public Metrics Endpoint](#) below.
- **Register a structured log format:** Register a structured log format that can be emitted to Loggregator. See [Register a Structured Log Format](#).

 **Note:** If you are migrating from and manually send data to Metrics Forwarder for PCF, Pivotal recommends registering a structured log format. See [Register a Structured Log Format](#).

### Register a Public Metrics Endpoint

The Metric Registrar supports custom metrics created with the open-source tool, Prometheus. Prometheus uses a text-based exposition format common in many open-source libraries. It also provides several out-of-the-box metrics for different programming languages.

For more information about Prometheus, see [What is Prometheus?](#) in the Prometheus documentation.

For examples of apps that use Prometheus to publish metrics to an endpoint, see [metric-registrar-examples](#) in GitHub.

 **Note:** These endpoints must be public for the Metric Registrar to retrieve the data. If you do not want to expose public endpoints for your app, see [Register a Structured Log Format](#) below.

### Prerequisites

Before registering a public metrics endpoint, you must do the following:

- For all Spring apps, update the `application.yml` file to include one or more Prometheus endpoints. For example:

```
management:
 endpoints:
 web:
 exposure:
 include: "metrics,prometheus"
 endpoint:
 metrics:
 enabled: true
 prometheus:
 enabled: true
```

- For all Spring apps, update the security configuration file to permit access to the Prometheus endpoints. For an example, see [metric-registrar-examples](#) in GitHub.
- For all other apps, see [client libraries] provided by the open-source community.

## Register a Public Metrics Endpoint

To register a public metrics endpoint for an app, do the following:

- Log in to the cf CLI.
- For each Prometheus endpoint in your app, run the following command to register the endpoint as a metric source:

```
cf register-metrics-endpoint APP-NAME PATH
```

Where:

- `APP-NAME` is the name of the app.
- `PATH` is the path to the Prometheus endpoint.

For example, if the app name is `example` and the metrics endpoint lives at `https://example.com/metrics` the command would be:

```
...
cf register-metrics-endpoint example /metrics
...
```

## Verify a Public Metrics Endpoint

Pivotal recommends that you install the LogCache cf CLI plugin to view metrics.

To install and use the Log Cache cf CLI plugin, do the following:

- To install Log Cache, run the following command:

```
cf install-plugin -r CF-Community "log-cache"
```

- To view metrics for an app, run the following command:

```
cf tail --envelope-class=metrics <APP_NAME>
```

For example, if you provided a custom metric called `users_per_cpu` as a gauge, you should see the following output:

```
2019-06-18T10:49:28.27-0600 [app-name/1] GAUGE cpu:0.154763 percentage disk:14000128.000000 bytes disk_quota:33554432.000000 bytes memory:10190848.000000 bytes memor
2019-06-18T10:49:26.42-0600 [app-name/1] GAUGE users_per_cpu:557.500000
```

Where the first line is default container metrics and the second line is custom metrics `users_per_cpu`.

 **Note:** The Metric Registrar produces the following benign error message at every polling interval: [LGR/] ERR Invalid syslog drain URL: parse failure

## Register a Structured Log Format

The Metric Registrar supports metrics emitted in JSON or DogStatsD formats. For more information about these formats, see the [JSON](#) and [DogStatsD](#) sections below.

To register your app as a metric source, do the following:

1. Log in to the cf CLI.
2. Run the following command:

```
cf register-log-format APP-NAME FORMAT
```

Where:

- o `APP-NAME` is the name of the app.
- o `FORMAT` is either `json` or `DogStatsD`.

3. In your app, log a structured `json` or `DogStatsD` message to represent the custom metric.

## JSON

The table below shows the supported JSON format for event, gauge, and counter log types.

Type	Format
Events	<pre>{   "type": "event",   "title": "title",   "body": "body",   "tags": {     "tag1": "tag value"   } }</pre>
Gauges	<pre>{   "type": "gauge",   "name": "some-counter",   "value": ,   "tags": {     "tag1": "tag value"   } }</pre>
Counters	<pre>{   "type": "counter",   "name": "some-counter",   "delta": ,   "tags": {     "tag1": "tag value"   } }</pre>

## DogStatsD

The table below shows the supported DogStatsD format for event, gauge, and counter log types. It also lists the supported fields. For more information about DogStatsD, see the [DogStatsD](#) topic in the Datadog documentation.

Type	Format	Supported Fields
Events	<code>_e{title.length,text.length}:title text d:timestamp  h:hostname p:priority t:alert_type #tag1,tag2</code>	<code>title</code> <code>text</code>
Gauges	<code>gauge.name:value g @sample_rate #tag1:value,tag2</code>	<code>gauge.name</code>

Counters	counter.name:value c @sample_rate #tag1:value,tag2	value counter.name value
----------	----------------------------------------------------	--------------------------------

## Identifying the Source Deployment of Metrics

This topic describes operators can identify which runtime deployment the metrics in their Pivotal Cloud Foundry (PCF) foundation are coming from. For example, an operator may want to know whether metrics are coming from Pivotal Application Service (PAS) or PCF Isolation Segment.

### Tile GUIDs

As the value for the `deployment` tag, metrics use `cf-` and then the GUID for your PAS tile. With a GUID, you can uniquely identify your metrics by tile.

### Human-friendly Metadata

PAS and other runtime tiles tag metrics with additional metadata to help operators better parse the metrics coming from their different deployments. Also, downstream monitoring products, such as [PCF Healthwatch](#), use this metadata to display human readable names.

The tags are as follows:

- `product` : The value of this tag is always `Pivotal Application Service` for the PAS tile. The tags for other products are: `PCF Isolation Segment`, `PCF Small Footprint`, `Pivotal Application Service for Windows 2012R2`, and `Pivotal Application Service for Windows`.
- `system_domain` : The value of this tag corresponds to what you set in the **System Domain** field in the **Domains** tab of a given tile.
- `placement_tag` : The value of this tag is always `null` for PAS. However, for PAS for Windows and PCF Isolation Segment tiles, you can configure this value using the **Segment Name** field in the **Application Containers** pane. An Operator can display capacity and other relevant metrics using the `placement_tag` name. This makes it easier to reason about the importance of a given segment when issues arise.

These tags are properties of the [metron](#) agent running on each VM in a deployment.

## Troubleshooting and Diagnostics

Navigate these topics to troubleshoot and diagnose issues you may encounter when installing or using Pivotal Cloud Foundry (PCF).

### Table of Contents

- [Diagnosing Problems in PCF](#)
- [Troubleshooting Problems in PCF](#)
- [Advanced Troubleshooting with the BOSH CLI](#)
- [Troubleshooting Slow Requests in Cloud Foundry](#)
- [Troubleshooting TCP Routing](#)
- [Recovering MySQL from PAS Downtime](#)
- [Troubleshooting BBR](#)
- [Troubleshooting PCF on Azure](#)
- [Troubleshooting PCF on GCP](#)
- [Troubleshooting Ops Manager for VMware vSphere](#)
- [Troubleshooting Windows Cells](#)
- [Running mysql-diag](#)

## Diagnosing Problems in PCF

This guide provides help with diagnosing issues encountered during a [Pivotal Cloud Foundry](#) (PCF) installation.

Besides whether products install successfully or not, an important area to consider when diagnosing issues is communication between VMs deployed by Pivotal Cloud Foundry. Depending on what products you install, communication takes the form of messaging, routing, or both. If they go wrong, an installation can fail. For example, in an Pivotal Application Service (PAS) installation the PCF VM tries to push a test application to the cloud during post-installation testing. The installation fails if the resulting traffic cannot be routed to the HA Proxy load balancer.

## Viewing the Debug Endpoint

The debug endpoint is a web page that provides information useful for diagnostics. If you have superuser privileges and can view the Ops Manager Installation Dashboard, you can access the debug endpoint.

- In a browser, open the URL:

```
https://OPS-MANAGER-FQDN/debug
```

The debug endpoint offers three links:

- *Files* allows you to view the YAML files that Ops Manager uses to configure products that you install. The most important YAML file, `installation.yml`, provides networking settings and describes `microbosh`. In this case, `microbosh` is the VM whose BOSH Director component is used by Ops Manager to perform installations and updates of PAS and other products.
- *Components* describes the components in detail.
- *Rails log* shows errors thrown by the VM where the Ops Manager web application (a Rails application) is running, as recorded in the `production.log` file. See the next section to learn how to explore other logs.

## Logging Tips

### Identifying Where to Start

This section contains general tips for locating where a particular problem is called out in the log files. Refer to the later sections for tips regarding specific logs (such as those for PAS Components).

- Start with the largest and most recently updated files in the job log
- Identify logs that contain 'err' in the name
- Scan the file contents for a "failed" or "error" string

## Viewing Logs for PAS Components

To troubleshoot specific PAS components by viewing their log files, browse to the Ops Manager interface and follow the procedure below.

1. In Ops Manager, browse to the **PAS Status** tab. In the **Job** column, locate the component of interest.
2. In the **Logs** column for the component, click the download icon.

[Installation Dashboard](#)

## Pivotal Elastic Runtime

Settings Status Credentials Logs

JOB	INDEX	IPS	CID	LOAD AVG15	CPU	MEMORY	SWAP	SYSTEM DISK	EPHEM. DISK	PERS. DISK	LOGS
HAProxy	0	10.0.0.254	vm-9985a13c-106a-48d1-a3de-d0e0e816c857	0.08%	0.1%	8.7%	0.0%	41%	5%	N/A	
NATS	0	10.0.0.5	vm-dee49615-aea8-4f4f-bf0f-b1060083dddef	0.05%	0.2%	8.9%	0.0%	41%	19%	N/A	
			vm-6d43e59e-								

3. Browse to the PAS Logs tab.

[Installation Dashboard](#)

## Pivotal Elastic Runtime

Settings Status Credentials Logs

FILENAME	UPDATED AT
<b>Downloaded:</b>	
<a href="#">/tmp/jobs_logs/diego_cell-partition-ee0b66b1415c8591855d-0-150e690af6dd.zip</a>	2016-04-12 19:06:47 UTC
<a href="#">/tmp/jobs_logs/diego_database-partition-ee0b66b1415c8591855d-0-ab7bd3378ef2.zip</a>	2016-04-12 19:09:30 UTC
<b>Pending:</b>	
<a href="#">/tmp/jobs_logs/diego_brain-partition-ee0b66b1415c8591855d-0-196edf23631b.zip</a>	2016-04-12 19:09:34 UTC

4. Once the zip file corresponding to the component of interest moves to the **Downloaded** list, click the linked file path to download the zip file.

5. Once the download completes, unzip the file.

The contents of the log directory vary depending on which component you view. For example, the Diego cell log directory contains sub-directories for the `metron_agent`, `rep`, `monit`, and `garden` processes. To view the standard error stream for `garden`, download the Diego cell logs and open `diego.0.job > garden > .garden.stderr.log`.

## Viewing Web Application and BOSH Failure Logs in a Terminal Window

You can obtain diagnostic information from the Operations Manager by logging in to the VM where it is running. To log in to the Operations Manager VM, you need the following information:

- The IP address of the PCF VM shown in the `Settings` tab of the Ops Manager Director tile.
- Your **import credentials**. Import credentials are the username and password used to import the PCF `.ova` or `.ovf` file into your virtualization system.

Complete the following steps to log in to the Operations Manager VM:

1. Open a terminal window.

2. Run `ssh IMPORT-USERNAME@PCF-VM-IP-ADDRESS` to connect to the PCF installation VM.

3. Enter your import password when prompted.

4. Change directories to the home directory of the web application:

```
cd /home/tempest-web/tempest/web/
```

5. You are now in a position to explore whether things are as they should be within the web application.

You can also verify that the `microbosh` component is successfully installed. A successful MicroBOSH installation is required to install PAS and any products like databases and messaging services.

6. Change directories to the BOSH installation log home:

```
cd /var/tempest/workspaces/default/deployments/micro
```

7. You may want to begin by running a tail command on the `current` log:

```
cd /var/tempest/workspaces/default/deployments/micro
```

If you are unable to resolve an issue by viewing configurations, exploring logs, or reviewing common problems, you can troubleshoot further by running BOSH diagnostic commands with the BOSH Command Line Interface (CLI).

 **Note:** Do not manually modify the deployment manifest. Operations Manager will overwrite manual changes to this manifest. In addition, manually changing the manifest may cause future deployments to fail.

## Viewing the VMs in Your Deployment

To view the VMs in your PCF deployment, perform the following steps specific to your IaaS.

### Amazon Web Services (AWS)

1. Log in to the [AWS Console](#).

2. Navigate to the EC2 Dashboard.

3. Click **Running Instances**.

4. Click the gear icon in the upper right.

5. Select the following: **job**, **deployment**, **director**, **index**.

6. Click **Close**.

### OpenStack

1. Install the [novaclient](#).

2. Point novaclient to your OpenStack installation and tenant by exporting the following environment variables:

```
$ export OS_AUTH_URL=YOUR_KEYSTONE_AUTH_ENDPOINT
$ export OS_TENANT_NAME=TENANT_NAME
$ export OS_USERNAME=USERNAME
$ export OS_PASSWORD=PASSWORD
```

3. List your VMs by running the following command:

```
$ nova list --fields metadata
```

### vSphere

1. Log into vCenter.

2. Select **Hosts and Clusters**.
3. Select the top level object that contains your PCF deployment. For example, select **Cluster**, **Datastore** or **Resource Pool**.
4. In the top tab, click **Related Objects**.
5. Select **Virtual Machines**.
6. Right click on the **Table** heading and select **Show/Hide Columns**.
7. Select the following boxes: **job**, **deployment**, **director**, **index**.

## Viewing Apps Manager Logs in a Terminal Window

The [Apps Manager](#) provides a graphical user interface to help manage organizations, users, applications, and spaces.

When troubleshooting Apps Manager performance, you might want to view the Apps Manager application logs. To view the Apps Manager application logs, follow these steps:

1. Run `cf login -a api.MY-SYSTEM-DOMAIN -u admin` from a command line to log in to PCF using the UAA Administrator credentials. In Pivotal Ops Manager, refer to [PAS Credentials](#) for these credentials.

```
$ cf login -a api.example.com -u admin
API endpoint: api.example.com

Password>*****
Authenticating...
OK
```

2. Run `cf target -o system -s apps-manager` to target the `system` org and the `apps-manager` space.

```
$ cf target -o system -s apps-manager
```

3. Run `cf logs apps-manager` to tail the Apps Manager logs.

```
$ cf logs apps-manager
Connected, tailing logs for app apps-manager in org system / space apps-manager as
admin...
```

## Changing Logging Levels for the Apps Manager

The Apps Manager recognizes the `LOG_LEVEL` environment variable. The `LOG_LEVEL` environment variable allows you to filter the messages reported in the Apps Manager log files by severity level. The Apps Manager defines severity levels using the Ruby standard library [Logger class](#).

By default, the Apps Manager `LOG_LEVEL` is set to `info`. The logs show more verbose messaging when you set the `LOG_LEVEL` to `debug`.

To change the Apps Manager `LOG_LEVEL`, run `cf set-env apps-manager LOG_LEVEL` with the desired severity level.

```
$ cf set-env apps-manager LOG_LEVEL debug
```

You can set `LOG_LEVEL` to one of the six severity levels defined by the Ruby Logger class:

- **Level 5:** `unknown` – An unknown message that should always be logged
- **Level 4:** `fatal` – An unhandleable error that results in a program crash
- **Level 3:** `error` – A handleable error condition
- **Level 2:** `warn` – A warning
- **Level 1:** `info` – General information about system operation
- **Level 0:** `debug` – Low-level information for developers

Once set, the Apps Manager log files only include messages at the set severity level and above. For example, if you set `LOG_LEVEL` to `fatal`, the log includes `fatal` and `unknown` level messages only.

## Analyzing Disk Usage on Containers and Diego Cell VMs

To obtain disk usage statistics by Diego Cell VMs and containers, see [Examining GrootFS Disk Usage](#).

## Troubleshooting Problems in PCF

This guide provides help with resolving issues encountered during a [Pivotal Cloud Foundry](#) (PCF) installation.

### Retrying the Deployment

Although an install or update can fail for many reasons, the system is self-healing, and can often automatically correct or work around hardware or network faults.

Click **Install** or **Review Pending Changes**, then **Apply Changes** again, and the system may resolve a problem on its own.

Some failures only produce generic errors like `Exited with 1`. In cases like this, where a failure is not accompanied by useful information, click **Install** or **Review Pending Changes**, then **Apply Changes** to retry.

When the system does provide informative evidence, review the [Common Problems](#) section at the end of this guide to see if your problem is covered there.

## Common Issues

Compare evidence that you have gathered to the descriptions below. If your issue is covered, try the recommended remediation procedures.

### BOSH Does Not Reinstall

You might want to reinstall BOSH for troubleshooting purposes. However, if PCF does not detect any changes, BOSH does not reinstall. To force a reinstall of BOSH, select **BOSH Director > Resource Sizes** and change a resource value. For example, you could increase the amount of RAM by 4 MB.

### Creating Bound Missing VMs Times Out

This task happens immediately following package compilation, but before job assignment to agents. For example:

```
cloud_controller/0: Timed out pinging to f690db09-876c-475e-865f-2cece06aba79 after 600 seconds (00:10:24)
```

This is most likely a NATS issue with the VM in question. To identify a NATS issue, inspect the agent log for the VM. Since the BOSH director is unable to reach the BOSH agent, you must access the VM using another method. You will likely also be unable to access the VM using TCP. In this case, access the VM using your virtualization console.

To diagnose:

1. Access the VM using your virtualization console and log in.
2. Navigate to the **Credentials** tab of the Pivotal Application Service (PAS) tile and locate the VM in question to find the **VM credentials**.
3. Become root.
4. Run `cd /var/vcap/bosh/log`.
5. Open the file `current`.
6. First, determine whether the BOSH agent and director have successfully completed a handshake, represented in the logs as a “ping-pong”:

```
2013-10-03_14:35:48.58456 #[608] INFO: Message: {"method":>"ping", "arguments":>[], "reply_to":>"director.f4b7df14-cb8f.19719508-e0dd-4f53-b755-58b6336058ab"}
2013-10-03_14:35:48.60182 #[608] INFO: reply_to: director.f4b7df14-cb8f.19719508-e0dd-4f53-b755-58b6336058ab:
payload: {value:>"pong"}
```

This handshake must complete for the agent to receive instructions from the director.

7. If you do not see the handshake, look for another line near the beginning of the file, prefixed `[INFO: loaded new infrastructure settings]`. For example:

```
2013-10-03_14:35:21.83222 #[608] INFO: loaded new infrastructure settings:
{"vm"=>{"name"=>"vm-4d80ede4-b0a5-4992-aea6a0386e18e", "id"=>"vm-360"},
 "agent_id"=>"56aca4ef-6aa9-4c39-8019-7024cefdd4",
 "networks"=>{"default"=>{"ip"=>"192.0.2.19",
 "netmask"=>"255.255.255.0", "cloud_properties"=>{"name"=>"VMNetwork",
 "default"=>["dns", "gateway"],
 "dns"=>["192.0.2.2", "192.0.2.17"], "gateway"=>"192.0.2.2",
 "dns_record_name"=>"0.nats.default.cf-d729343071061.microbosh",
 "mac"=>"00:50:56:9b:71:67"}, "disks"=>{"system"=>0, "ephemeral"=>1,
 "persistent"=>{}}, "ntp"=>[], "blobstore"=>{"provider"=>"dav",
 "options"=>{"endpoint"=>"http://192.0.2.17:25250",
 "user"=>"agent", "password"=>"agent"},
 "mbus"=>"nats://nats:nats@192.0.2.17:4222",
 "env"=>{"bosh"=>{"password"=>"$6$40ftQ9K4rvvC/8ADZHW0"}}}
```

This is a JSON blob of key/value pairs representing the expected infrastructure for the BOSH agent. For this issue, the following section is the most important:

```
"mbus"=>"nats://nats:nats@192.0.2.17:4222"
```

This key/value pair represents where the agent expects the NATS server to be. One diagnostic tactic is to try pinging this NATS IP address from the VM to determine whether you are experiencing routing issues.

## Install Exits With a Creates/Updates Deletes App Failure or With a 403 Error

**Scenario 1:** Your PCF install exits with the following 403 error when you attempt to log in to the Apps Manager:

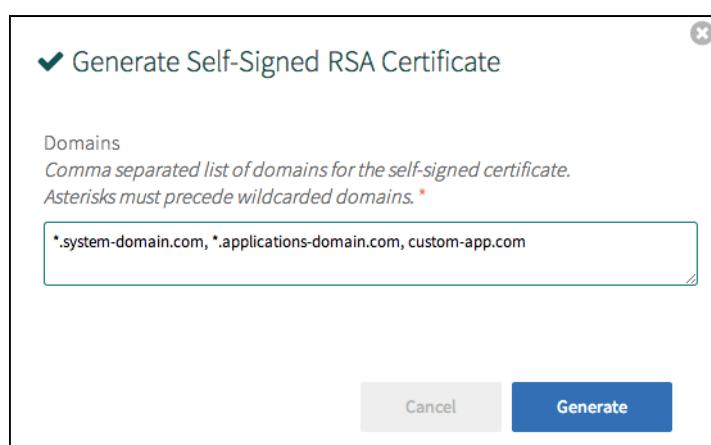
```
{"type": "step_finished", "id": "apps-manager.deploy"}
/home/tempest-web/tempest/web/vendor/bundle/ruby/1.9.1/gems/mechanize-2.7.2/lib/mechanize/http/agent.rb:306:in
`fetch': 403 => Net::HTTPForbidden for https://login.api.example.net/oauth/authorizeresponse_type=code&client_id=portal&redirect_uri=https%3..
-- unhandled response (Mechanize::ResponseCodeError)
```

**Scenario 2:** Your PCF install exits with a creates/updates/deletes an app (FAILED - 1) error message with the following stack trace:

```
1) App CRUD creates/updates/deletes an app
Failure/Error: Unable to find matching line from backtrace
CFoundry::TargetRefused:
 Connection refused - connect(2)
```

In either of the above scenarios, ensure that you have correctly entered your domains in wildcard format:

1. Browse to the Operations Manager fully qualified domain name (FQDN).
2. Click the PAS tile.
3. Select **HAProxy** and click **Generate Self-Signed RSA Certificate**.
4. Enter your system and app domains in wildcard format, as well as optionally any custom domains, and click **Save**. Refer to **PAS Cloud Controller** for explanations of these domain values.



## Install Fails When Gateway Instances Exceed Zero

If you configure the number of Gateway instances to be greater than zero for a given product, you create a dependency on PAS for that product installation. If you attempt to install a product tile with an PAS dependency before installing PAS, the install fails.

To change the number of Gateway instances, click the product tile, then select **Settings > Resource sizes > INSTANCES** and change the value next to the product Gateway job.

To remove the PAS dependency, change the value of this field to  0.

## Out of Disk Space Error

PCF displays an **Out of Disk Space** error if log files expand to fill all available disk space. If this happens, rebooting the PCF installation VM clears the tmp directory of these log files and resolves the error.

If users receive **Out of Disk Space** errors when trying to push apps, this can mean that Diego cells may be running out of disk space capacity.

To perform a detailed analysis of disk usage by containers and host VMs in your PAS deployment, see [Examining GrootFS Disk Usage](#).

## Installing BOSH Director Fails

If the DNS information for the PCF VM is incorrectly specified when deploying the PCF .ova file, installing BOSH Director fails at the “Installing Micro BOSH” step.

To resolve this issue, correct the DNS settings in the PCF Virtual Machine properties.

## Deleting Ops Manager Fails

Ops Manager displays an error message when it cannot delete your installation. This scenario might happen if the BOSH Director cannot access the VMs or is experiencing other issues. To manually delete your installation and all VMs, you must do the following:

1. Use your IaaS dashboard to manually delete the VMs for all installed products, with the exception of the Ops Manager VM.
2. SSH into your Ops Manager VM and remove the `installation.yml` file from `/var/tempest/workspaces/default/`.

 **Note:** Deleting the `installation.yml` file does not prevent you from reinstalling Ops Manager. For future deploys, Ops Manager regenerates this file when you click **Save** on any page in the BOSH Director.

Your installation is now deleted.

## Installing PAS Fails

The following sections describe errors that may occur when installing PAS and how to resolve them.

### Ops Manager Cannot Verify App Push

If the DNS information for the PCF VM becomes incorrect after BOSH Director has been installed, installing PAS with Pivotal Operations Manager fails at the “Verifying app push” step.

To resolve this issue, correct the DNS settings in the PCF Virtual Machine properties.

### MySQL Monitor Not Running After Update

When MySQL cannot communicate with UAA, Ops Manager shows the following error:

Error: ‘mysql\_monitor/12a3b456-cd7e-8fgh-9012-345678b90ijk (0)’ is not running after update. Review logs for failed jobs: replication-canary.

If you see this error, create firewall rules that allow MySQL to reach UAA, using the [MySQL Network Communications](#) topic as a reference.

## Ops Manager Hangs During MicroBOSH Install or HAProxy States “IP Address Already Taken”

During an Ops Manager installation, you might receive the following errors:

- The Ops Manager GUI shows that the installation stops at the “Setting MicroBOSH deployment manifest” task.
- When you set the IP address for the HAProxy, the “IP Address Already Taken” message appears.

When you install Ops Manager, you assign it an IP address. Ops Manager then takes the next two consecutive IP addresses, assigns the first to MicroBOSH, and reserves the second. For example:

```
203.0.113.1 - Ops Manager (User assigned)
203.0.113.2 - MicroBOSH (Ops Manager assigned)
203.0.113.3 - Reserved (Ops Manager reserved)
```

To resolve this issue, ensure that the next two subsequent IP addresses from the manually assigned address are unassigned.

## Poor PCF Performance

If you notice poor network performance by your PCF deployment and your deployment uses a Network Address Translation (NAT) gateway, your NAT gateway may be under-resourced.

### Troubleshoot

To troubleshoot the issue, set a custom firewall rule in your IaaS console to route traffic originating from your private network directly to an S3-compatible object store. If you see decreased average latency and improved network performance, perform the solution below to scale up your NAT gateway.

### Scale Up Your NAT Gateway

Perform the following steps to scale up your NAT gateway:

1. Navigate to your IaaS console.
2. Spin up a new NAT gateway of a larger VM size than your previous NAT gateway.
3. Change the routes to direct traffic through the new NAT gateway.
4. Spin down the old NAT gateway.

The specific procedures will vary depending on your IaaS. Consult your IaaS documentation for more information.

## Common Issues Caused by Firewalls

This section describes various issues you might encounter when installing PAS in an environment that uses a strong firewall.

### DNS Resolution Fails

When you install PCF in an environment that uses a strong firewall, the firewall might block DNS resolution. To resolve this issue, refer to the [Troubleshooting DNS Resolution Issues](#) section of the Preparing Your Firewall for Deploying PCF topic.

## Advanced Troubleshooting with the BOSH CLI

This topic describes using the BOSH CLI to help diagnose and resolve issues with your [Pivotal Cloud Foundry](#) (PCF) deployment. Before using the information and techniques in this topic, review [Diagnosing Problems in PCF](#).

To follow the steps in this topic, you must log in to the BOSH Director VM. The BOSH Director runs on the virtual machine (VM) that Ops Manager deploys on the first install of the BOSH Director tile.

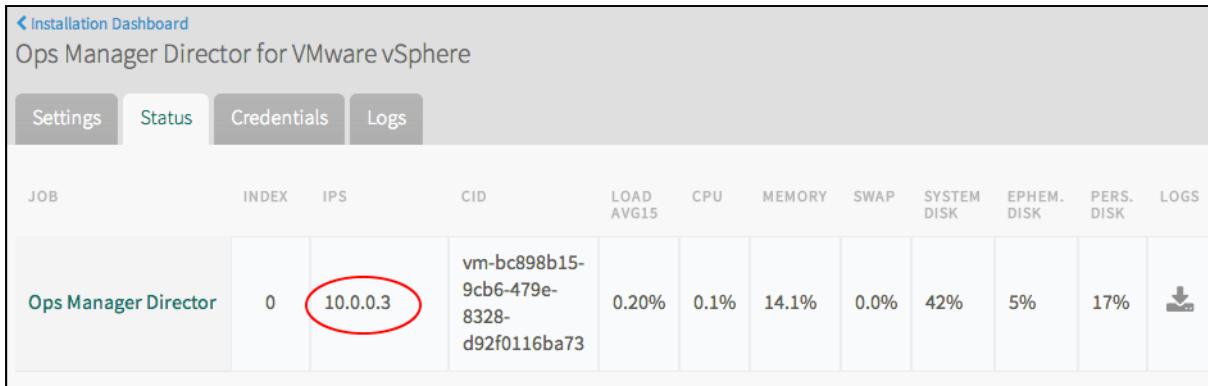
After authenticating into the BOSH Director, you can run specific commands using the BOSH Command Line Interface (BOSH CLI). BOSH Director diagnostic commands have access to information about your entire [Pivotal Cloud Foundry](#) (PCF) installation.

**Note:** Before running any BOSH CLI commands, verify that no BOSH Director tasks are running on the Ops Manager VM. See the [Tasks](#) section of [BOSH CLI commands](#) for more information.

## Gather Credential and IP Address Information

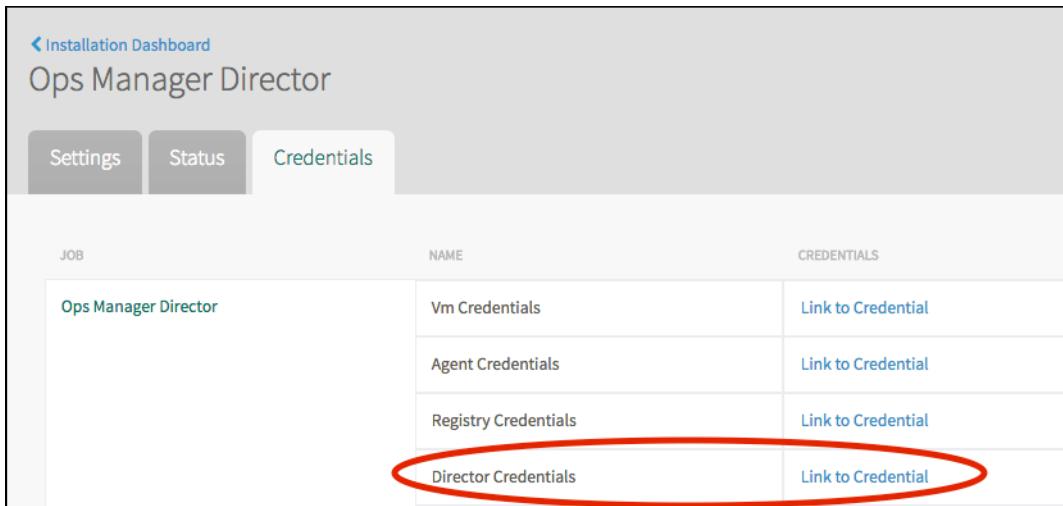
Before you begin troubleshooting with the BOSH CLI, follow the instructions below to collect the information you need from the Ops Manager interface.

1. Open the Ops Manager interface by navigating to the Ops Manager fully qualified domain name (FQDN) in a web browser.
2. Click the **BOSH Director** tile and select the **Status** tab.
3. Record the IP address for the Director job. This is the IP address of the VM where the BOSH Director runs.



Ops Manager Director for VMware vSphere											
JOB	INDEX	IPS	CID	LOAD AVG15	CPU	MEMORY	SWAP	SYSTEM DISK	EPHEM. DISK	PERS. DISK	LOGS
Ops Manager Director	0	10.0.0.3	vm-bc898b15-9cb6-479e-8328-d92f0116ba73	0.20%	0.1%	14.1%	0.0%	42%	5%	17%	

4. Select the **Credentials** tab.
5. Click [Link to Credential](#) to view the **Director Credentials**. Record these credentials.



Ops Manager Director		
SETTINGS	STATUS	CREDENTIALS
<a href="#">Installation Dashboard</a>		
<a href="#">Ops Manager Director</a>		
JOB	NAME	CREDENTIALS
Ops Manager Director	Vm Credentials	<a href="#">Link to Credential</a>
	Agent Credentials	<a href="#">Link to Credential</a>
	Registry Credentials	<a href="#">Link to Credential</a>
	Director Credentials	<a href="#">Link to Credential</a>

6. Return to the **Installation Dashboard**.
7. (Optional) To prepare to troubleshoot the job VM for any other product, click the product tile and repeat the procedure above to record the IP

address and VM credentials for that job VM.

8. Log out of Ops Manager.

 **Note:** Ensure that there are no Ops Manager installations or updates in progress while using the BOSH CLI.

## Log in to the Ops Manager VM with SSH

Use SSH to connect to the Ops Manager VM. Follow the instructions in one of the sections below to log in to the Ops Manager VM with SSH.

### AWS

To log in to the Ops Manager VM with SSH in AWS, you need the key pair you used when you created the Ops Manager VM. To see the name of the key pair, click on the Ops Manager VM and locate the `key pair name` in the properties.

To log in to the Ops Manager VM with SSH in AWS, do the following:

1. Locate the Ops Manager FQDN on the AWS EC2 instances page.
2. Run `chmod 600 ops_mgr.pem` to change the permissions on the `.pem` file to be more restrictive. For example:

```
$ chmod 600 ops_mgr.pem
```

3. Run `ssh -i ops_mgr.pem ubuntu@OPS-MANAGER-FQDN` to log in to the Ops Manager VM with SSH. Replace `OPS-MANAGER-FQDN` with the fully qualified domain name of Ops Manager. For example:

```
$ ssh -i ops_mgr.pem ubuntu@my-opsmanager-fqdn.example.com
```

### Azure

To log in to the Ops Manager VM with SSH in Azure, you need the key pair you used when creating the Ops Manager VM. If you need to reset the SSH key, locate the Ops Manager VM in the Azure portal and click **Reset Password**.

To log in to the Ops Manager VM with SSH in Azure, do the following:

1. Locate the Ops Manager FQDN by selecting the VM in the Azure portal.
2. Run `chmod 600 ops_mgr.pem` to change the permissions on the `.pem` file to be more restrictive. For example:

```
$ chmod 600 ops_mgr.pem
```

3. Run `ssh -i ops_mgr.pem ubuntu@OPS-MANAGER-FQDN` to log in to the Ops Manager VM with SSH. Replace `OPS-MANAGER-FQDN` with the fully qualified domain name of Ops Manager. For example:

```
$ ssh -i ops_mgr.pem ubuntu@my-opsmanager-fqdn.example.com
```

### GCP

To log in to the Ops Manager VM with SSH in GCP, do the following:

1. Confirm that you have installed the [Google Cloud SDK and CLI](#). See the [Google Cloud Platform documentation](#) for more information.
2. Initialize Google Cloud CLI, using a user account with Owner, Editor, or Viewer permissions to access the project. Ensure that the Google Cloud CLI can login to the project by running the command `gcloud auth login`.
3. From the GCP web console, navigate to **Compute Engine**.
4. Locate the Ops Manager VM in the **VM Instances** list.

5. Under **Remote access**, click the **SSH** dropdown and select **View gcloud command**.

6. Copy the SSH command that appears in the popup window.

7. Paste the command into your terminal window to SSH to the VM. For example:

```
$ gcloud compute ssh "YOUR-VM" --zone "YOUR-ZONE-ID"
```

8. Run `sudo su - ubuntu` to switch to the `ubuntu` user.

## OpenStack

To log in to the Ops Manager VM with SSH in OpenStack, you need the key pair that you created in Step 2: *Configure Security* of [Deploying Ops Manager to OpenStack](#). If you need to reset the SSH key, locate the Ops Manager VM in the OpenStack console and boot it in recovery mode to generate a new key pair.

To log in to the Ops Manager VM with SSH in OpenStack, do the following:

1. Locate the Ops Manager FQDN on the **Access & Security** page.

2. Run `chmod 600 ops_mgr.pem` to change the permissions on the `.pem` file to be more restrictive. For example:

```
$ chmod 600 ops_mgr.pem
```

3. Run `ssh -i ops_mgr.pem ubuntu@OPS-MANAGER-FQDN` to log in to the Ops Manager VM with SSH. Replace `OPS-MANAGER-FQDN` with the fully qualified domain name of Ops Manager. For example:

```
$ ssh -i ops_mgr.pem ubuntu@my-opsmanager-fqdn.example.com
```

## vSphere

To log in to the Ops Manager VM with SSH in vSphere, you need the credentials used to import the PCF .ova or .ovf file into your virtualization system. You set these credentials when you installed Ops Manager.

 **Note:** If you lose your credentials, you must shut down the Ops Manager VM in the vSphere UI and reset the password. See the [vSphere documentation](#) for more information.

1. From a command line, run `ssh ubuntu@OPS-MANAGER-FQDN` to log in to the Ops Manager VM with SSH. Replace `OPS-MANAGER-FQDN` with the fully qualified domain name of Ops Manager.

2. When prompted, enter the password that you set during the .ova deployment into vCenter. For example:

```
$ ssh ubuntu@my-opsmanager-fqdn.example.com
Password: *****
```

## Log in to the BOSH Director VM

Follow the steps below to log in to the BOSH Director VM.

### Create a Local BOSH Director Alias

1. Run the following command to create a local alias for the BOSH Director using the BOSH CLI: `bosh alias-env MY-ENV -e DIRECTOR-IP-ADDRESS --ca-cert /var/tempest/worksheets/default/root_ca_certificate`

Replace the placeholder text with the following:

- o `MY-ENV` : Enter an alias for the BOSH Director, such as `gep`.
- o `DIRECTOR-IP-ADDRESS` : Enter the IP address of your BOSH Director VM. For example:

```
$ bosh alias-env gcp -c 10.0.0.3 --ca-cert /var/tempest/workspaces/default/root_ca_certificate
```

2. Log in to the BOSH Director VM using one of the following options:

- o [Internal User Store Login through UAA](#): Log in to the BOSH Director VM using BOSH.
- o [External User Store Login through SAML](#): Use an external user store to log in to the BOSH Director VM.

## Log in to the BOSH Director VM with UAA

1. Retrieve the Director password from the **BOSH Director > Credentials** tab. Alternatively, launch a browser and visit `https://OPS-MANAGER-FQDN/api/v0/deployed/director/credentials/director_credentials` to obtain the password. Replace `OPS-MANAGER-FQDN` with the fully qualified domain name of Ops Manager.

2. Run `bosh -e MY-ENV log-in` to log in to the BOSH Director VM. Replace `MY-ENV` with the alias for your BOSH Director. For example:

```
$ bosh -e gcp log-in
```

Follow the BOSH CLI prompts and enter the BOSH Director credentials to log in to the BOSH Director VM.

## Log in to the BOSH Director VM with SAML

1. Log in to your identity provider and use the following information to configure SAML Service Provider Properties:

- o **Service Provider Entity ID:** `bosh-uaa`
- o **ACS URL:** `https://DIRECTOR-IP-ADDRESS:8443/saml/SSO/alias/bosh-uaa`
- o **Binding:** HTTP Post
- o **SLO URL:** `https://DIRECTOR-IP-ADDRESS:8443/saml/SSO/alias/bosh-uaa`
- o **Binding:** HTTP Redirect
- o **Name ID:** Email Address

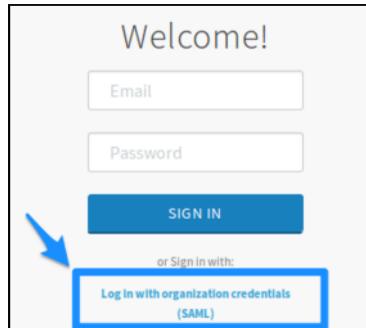
2. Run `bosh -e MY-ENV log-in` to log in to the BOSH Director VM. Replace `MY-ENV` with the alias for your BOSH Director. For example:

```
$ bosh -e gcp log-in
```

Follow the BOSH CLI prompts and enter your SAML credentials to log in to the BOSH Director VM.

**Note:** Your browser must be able to reach the BOSH Director in order to log in with SAML.

3. Click **Log in with organization credentials (SAML)**.



4. Copy the **Temporary Authentication Code** that appears in your browser.

**Temporary Authentication Code**  
94Ao1L

5. You see a login confirmation. For example:

Logged in as admin@example.org

## Log in to the BOSH Director VM with SSH

Do the following steps to log in to the BOSH Director VM with SSH:

1. From Ops Manager, open the BOSH Director tile.
2. Select the **Credentials** tab.
3. Next to **Bbr Ssh Credentials**, click **Link to Credential**. A tab opens containing a JSON credential structure.
4. Copy the **RSA PRIVATE KEY** and paste it into a file named **bbr.pem**. Include **-----BEGIN RSA PRIVATE KEY-----** and **-----END RSA PRIVATE KEY-----**.

**⚠ warning:** Pivotal recommends you keep the key secure. The key provides full access to the entire PCF environment.

5. Replace all **\n** characters in **bbr.pem** with a line break.
6. Copy **bbr.pem** to the **~/.ssh/** directory on your machine.
7. Modify the permissions of the file by running **chmod 600 ~/.ssh/bbr.pem**.
8. Log in to the BOSH Director VM with SSH from your machine.

```
ssh bbr@BOSH-DIRECTOR-IP -i ~/.ssh/bbr.pem
```

**💡 Note:** If you are using GCP, ensure SSH port **22** is open for your BOSH Director VM in your GCP console. If the SSH port is not open, open it by creating a firewall rule.

9. Run **sudo -i** to get the root privilege.

## Use the BOSH CLI for Troubleshooting

This section describes three BOSH CLI commands commonly used during troubleshooting.

- **VMs:** Lists the VMs in a deployment
- **Cloud Check:** Runs a cloud consistency check and interactive repair
- **SSH:** Starts an interactive session or executes commands with a VM

### BOSH VMs

The **bosh vms** command provides an overview of the virtual machines that BOSH manages.

To use this command, run **bosh -e MY-ENV vms** to see an overview of all virtual machines managed by BOSH, or **bosh -e MY-ENV -d MY-DEPLOYMENT vms** to

see only the virtual machines associated with a particular deployment. Replace **MY-ENV** with your environment, and, if using the **-d** flag, also replace **MY-DEPLOYMENT** with the name of a deployment.

When troubleshooting an issue with your deployment, **bosh vms** may show a VM in an **unknown** state. Run **bosh cloud-check** on a VM in an **unknown** state to instruct BOSH to diagnose problems with the VM.

You can also run **bosh vms** to identify VMs in your deployment, then use the **bosh ssh** command to log in to an identified VM with SSH for further troubleshooting.

**bosh vms** supports the following arguments:

- **--dns** : Report also includes the DNS A record for each VM
- **--vitals** : Report also includes load, CPU, memory usage, swap usage, system disk usage, ephemeral disk usage, and persistent disk usage for each

 Note: The **Status** tab of the Pivotal Application Service (PAS) product tile displays information similar to the `bosh vms` output.

## BOSH Cloud Check

Run the `bosh cloud-check` command to instruct BOSH to detect differences between the VM state database maintained by the BOSH Director and the actual state of the VMs. For each difference detected, `bosh cloud-check` can offer the following repair options:

- `Reboot VM`: Instructs BOSH to reboot a VM. Rebooting can resolve many transient errors.
- `Ignore problem`: Instructs BOSH to do nothing. You may want to ignore a problem in order to run `bosh ssh` and attempt troubleshooting directly on the machine.
- `Reassociate VM with corresponding instance`: Updates the BOSH Director state database. Use this option if you believe that the BOSH Director state database is in error and that a VM is correctly associated with a job.
- `Recreate VM using last known apply spec`: Instructs BOSH to destroy the server and recreate it from the deployment manifest that the installer provides. Use this option if a VM is corrupted.
- `Delete VM reference`: Instructs BOSH to delete a VM reference in the Director state database. If a VM reference exists in the state database, BOSH expects to find an agent running on the VM. Select this option only if you know that this reference is in error. Once you delete the VM reference, BOSH can no longer control the VM.

To use this command, run `bosh -e MY-ENV -d MY-DEPLOYMENT cloud-check`. Replace `MY-ENV` with your environment, and `MY-DEPLOYMENT` with your deployment.

## Example Scenarios

### Unresponsive Agent

```
$ bosh -e example-env -d example-deployment cloud-check
cdb/0 (vm-3e37133c-bc33-450e-98b1-f86d5b63502a) is not responding:

- Ignore problem
- Reboot VM
- Recreate VM using last known apply spec
- Delete VM reference (DANGEROUS!)
```

### Missing VM

```
$ bosh -e example-env -d example-deployment cloud-check
VM with cloud ID 'vm-3e37133c-bc33-450e-98b1-f86d5b63502a' missing:

- Ignore problem
- Recreate VM using last known apply spec
- Delete VM reference (DANGEROUS!)
```

### Unbound Instance VM

```
$ bosh -e example-env -d example-deployment cloud-check
VM 'vm-3e37133c-bc33-450e-98b1-f86d5b63502a' reports itself as 'cdb/0' but does not have a bound instance:

- Ignore problem
- Delete VM (unless it has persistent disk)
- Reassociate VM with corresponding instance
```

### Out of Sync VM

```
$ bosh -e example-env -d example-deployment cloud-check
VM 'vm-3e37133c-bc33-450e-98b1-f86d5b63502a' is out of sync:
expected 'cf-d7293430724a2c421061: cedb/0', got 'cf-d7293430724a2c421061: nats/0':

- Ignore problem
- Delete VM (unless it has persistent disk)
```

## BOSH SSH

Use `bosh ssh` to log in to the VMs in your deployment with SSH.

Follow the steps below to use `bosh ssh`:

1. Identify a VM to log in to with SSH. Run `bosh -e MY-ENV -d MY-DEPLOYMENT vms` to list the VMs in the given deployment. Replace `MY-ENV` with your environment alias and `MY-DEPLOYMENT` with the deployment name.
2. Run `bosh -e MY-ENV -d MY-DEPLOYMENT ssh VM-NAME/GUID`. For example:

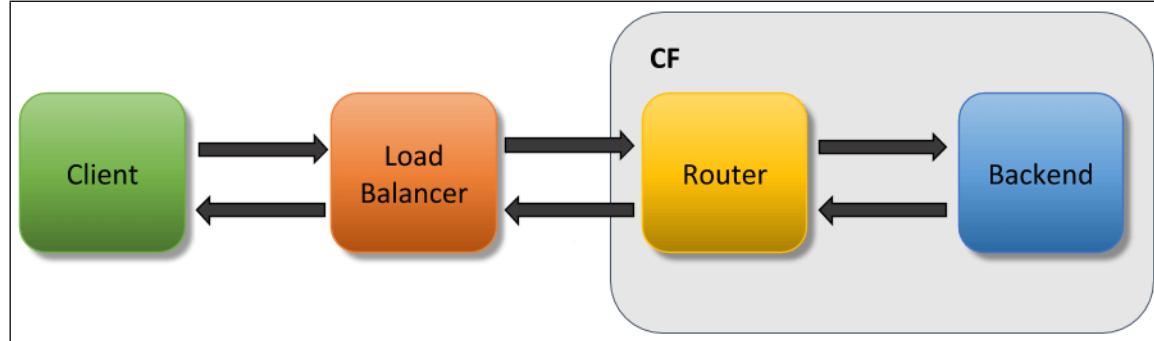
```
$ bosh -e example-env -d example-deployment ssh diego-cell/abcd0123-a012-b345-c678-9def01234567
```

## Troubleshooting Slow Requests in Cloud Foundry

This topic suggests ways that an operator of Cloud Foundry (CF) can diagnose the location of app request delays.

### App Request Path

App requests typically transit the following components. Only the router (Gorouter) and app are within the scope of Cloud Foundry. Operators may have the HAProxy load balancer that comes with CF deployed, instead of or in addition to, an infrastructure load balancer.



You can use `time` to measure a request's full round trip time from the client and back, examine `cf logs` output to measure the time just within Cloud Foundry, and add log messages to your app for fine-grained measurements of where the app itself takes time. By comparing these times, you can determine whether your delay comes from outside CF, inside the Gorouter, or in the app.

The following sections describe a scenario of diagnosing the source of delay for an app `app1`.

### Measure Total Round-Trip App Requests

On a command line, run `time curl -v APP-ENDPOINT` to measure the total round-trip time for deployed app `app1`. For example:

```
$ time curl -v http://app1.app_domain.com
GET /hello HTTP/1.1
Host: app1.app_domain.com
User-Agent: curl/7.43.0
Accept: */*

HTTP/1.1 200 OK
Content-Type: application/json; charset=utf-8
Date: Tue, 14 Dec 2016 00:31:32 GMT
Server: nginx
X-Content-Type-Options: nosniff
X-Vcap-Request-Id: c30fad28-4972-46eb-7da6-9d07dc79b109
Content-Length: 602
hello world!
real 2m0.707s
user 0m0.005s
sys 0m0.007s
```

The `real` time output shows that the request to `http://app1.app_domain.com` took approximately 2 minutes, round-trip. This seems like an unreasonably long time, so it makes sense to find out where the delay is occurring. To narrow it down, the next step measures the part of that request response time that comes from within Cloud Foundry.

**Note:** If your `curl` outputs an error like `Could not resolve host: NONEXISTENT.com` then DNS failed to resolve. If `curl` returns normally but lacks a `X-Vcap-Request-Id`, the request from the Load Balancer did not reach Cloud Foundry.

### Measure App Requests within Cloud Foundry

The `cf logs` command streams log messages from the Gorouter as well as from apps. To see the timestamps of Gorouter messages to and from your app, do the following:

1. If necessary, run `cf apps` to determine the name of the app.
2. Run `cf logs APP-NAME`. Replace `APP-NAME` with the name of the app.
3. From another terminal window, send a request to your app.
4. After your app returns a response, enter `Ctrl-C` to stop streaming `cf logs`.

For example:

```
$ cf logs app1
2016-12-14T00:33:32.35-0800 [RTR/0] OUT app1.app_domain.com - [14/12/2016:00:31:32.348 +0000] "GET /hello HTTP/1.1" 200 0 60 "-" "HTTPClient/1.0 (2.7.1, ruby 2.3.3 (2016-11-21
2016-12-14T00:32:32.35-0800 [APP/PROC/WEB/0]OUT app1 received request at [14/12/2016:00:32:32.348 +0000] with "vcap_request_id": "01144146-1e7a-4c77-77ab-49ae3e286fe9"
^C
```

In the example above, the first line contains timestamps from the Gorouter for both when it received the request and what was its response time processing the request:

- `14/12/2016:00:31:32.348` : Gorouter receives request
- `response_time:120.00641734` : Gorouter round-trip processing time

This output shows that it took 120 seconds for the Gorouter to process the request, which means that the 2-minute delay above takes place within CF, either within the Gorouter or within the app.

To determine whether the app is responsible, [add logging](#) to your app to measure where it is spending time.

 **Note:** Every incoming request should generate an access log message. If a request does not generate an access log message, it means the Gorouter did not receive the request.

## Use App Logs to Locate Delays in CF

To gain a more detailed picture of where delays exist in your request path, augment the logging that your app generates. For example, call your logging library from the request handler to generate log lines when your app receives a request and finishes processing it:

```
2016-12-14T00:33:32.35-0800 [RTR/0] OUT app1.app_domain.com - [14/12/2016:00:31:32.348 +0000] "GET /hello HTTP/1.1" 200 0 60 "-" "HTTPClient/1.0 (2.7.1, ruby 2.3.3 (2016-11-21
2016-12-14T00:32:32.35-0800 [APP/PROC/WEB/0]OUT app1 received request at [14/12/2016:00:32:32.348 +0000] with "vcap_request_id": "01144146-1e7a-4c77-77ab-49ae3e286fe9"
2016-12-14T00:32:32.50-0800 [APP/PROC/WEB/0]OUT app1 finished processing req at [14/12/2016:00:32:32.500 +0000] with "vcap_request_id": "01144146-1e7a-4c77-77ab-49ae3e286fe9"
```

Comparing the router access log messages from the [previous section](#) with the new app logs above, we can construct the following timeline:

- `14/12/2016:00:31:32.348` : Gorouter receives request
- `2016-12-14T00:32:32.35` : App receives request
- `2016-12-14T00:32:32.50` : App finishes processing request
- `2016-12-14T00:33:32.35` : Gorouter finishes processing request

The timeline indicates that the Gorouter took close to 60 seconds to send the request to the app and another 60 seconds to receive the response from the app. This suggests a delay either with the Gorouter, or in network latency between the Gorouter and Diego cells hosting the app.

## Time the Gorouter Processing

To determine whether a Gorouter delay comes from the Gorouter itself or network latency between the Gorouter and the app, log into the Gorouter and compare the response times from calling the app two different ways:

- Call the app through the router proxy to process requests through the Gorouter
- Directly call a specific instance of the app, bypassing Gorouter processing

To make this comparison, do the following:

1. Log in to the `router` using `bosh ssh`. See the [BOSH SSH documentation](#) for more information.
2. Use `time` and `curl` to measure the response time of an app request originating from and processing through the Gorouter, but not running through the client network or load balancer:

```
$ time curl -H "Host: app1.app_domain.com" http://IP-GOROUTER-VM:80"
```

3. Obtain the IP address and port of a specific app instance by running the following and recording associated `host` and `port` values:

```
$ cf curl /v2/apps/$(cf app app1 --guid)/stats
{
 "0": {
 "state": "RUNNING",
 "stats": {
 [...]
 "host": "10.10.148.39",
 "port": 60052,
 [...]
 },
 [...]
 }
}
```

4. Use the IP address and port values to measure response time calling the app instance directly, bypassing Gorouter processing:

```
$ time curl http://APP-HOST-IP:APP-PORT
```

If the Gorouter and direct response times are similar, it suggests network latency between the Gorouter and the Diego cell. If the Gorouter time is much longer than the direct-to-instance time, the Gorouter is slow processing requests. The next section explains why the Gorouter might be slow.

## Potential Causes for Gorouter Latency

- Routers are under heavy load from incoming client requests.
- Apps are taking a long time to process requests. This increases the number of concurrent threads held open by the Gorouter, reducing capacity to handle requests for other apps.

## Operations Recommendations

- Monitor CPU load for Gorouters. At high CPU (70%+), latency increases. If the Gorouter CPU reaches this threshold, consider adding another Gorouter instance.
- Monitor latency of all routers using metrics from the Firehose. Do not monitor the average latency across all routers. Instead, monitor them individually on the same graph.
- Consider using [Pingdom](#) against an app on your Cloud Foundry deployment to monitor latency and uptime.
- Consider enabling access logs on your load balancer. See your load balancer documentation for how. Just as we used Gorouter access log messages above to determine latency from the Gorouter, you can compare load balancer logs to identify latency between the load balancer and the Gorouter. You can also compare load balancer response times with the client response times to identify latency between client and load balancer.
- [Deploy a nozzle to the Loggregator Firehose](#) to track metrics for the Gorouter. Available metrics include:
  - CPU utilization
  - Latency
  - Requests per second

## Troubleshooting TCP Routes

The following topic provides steps to determine whether TCP routing issues are related to DNS and load balancer misconfiguration, the TCP routing tier, or the routing subsystem.

### Rule Out the App

If you are having TCP routing issues with an app, follow the procedure below to determine what to troubleshoot.

**Note:** If you have mutual TLS app identity verification enabled, app containers accept incoming communication only from the Gorouter. This disables TCP routing.

### Prerequisites

This procedure requires that you have the following:

- An app with TCP routing issues.
- A TCP domain. See [Routes and Domains](#) for more information about creating a TCP domain.
- A simple HTTP web app that you can use to curl.

### Procedure

1. Push a simple HTTP app using your TCP domain by entering the following command:

```
$ cf push MY-APP -d tcp.MY-DOMAIN --random-route
```

2. Curl your app on the port generated for the route. For example, if the port is 1024:

```
$ curl tcp.MY-DOMAIN:1024
```

3. If the curl request fails to reach the app, proceed to the next section: [Rule Out DNS and the Load Balancer](#).
4. If the curl request to your simple app succeeds, curl the app you are having issues with.
5. If you cannot successfully curl your problem app, TCP routing is working correctly. There is an issue with the app you cannot successfully curl.

### Rule Out DNS and the Load Balancer

1. Curl the TCP router healthcheck endpoint:

```
$ curl tcp.MY-DOMAIN:80/health -v
```

2. If you receive a `200 OK` response, proceed to the next section: [Rule Out the Routing Subsystem](#).
3. If you do not receive a `200 OK`, your load balancer may not be configured to pass through the healthcheck port. Continue following this procedure to test your load balancer configuration.
4. Confirm that your TCP domain name resolves to your load balancer:

```
$ dig tcp.MY-DOMAIN
...
tcp.MY-DOMAIN. 300 IN A 123.456.789.123
```

5. As an admin user, list the reservable ports configured for the `default-tcp` router group:

```
$ cf curl /routing/v1/router_groups
[
{
 "guid": "d9b1db52-ea78-4bb9-7473-ec8e5d411b14",
 "name": "default-tcp",
 "type": "tcp",
 "reservable_ports": "1024-1123"
}
]
```

6. Choose a port from the `reservable_ports` range and curl the TCP domain on this port. For example, if you chose port 1024:

```
$ curl tcp.MY-DOMAIN:1024 -v
```

7. If you receive an immediate rejection, then the TCP router likely rejected the request because there is no route for this port.

8. If your connection times out, then you need to configure your load balancer to route all ports in `reservable_ports` to the TCP routers.

## Rule Out the Routing Subsystem

Send a direct request to the TCP router to confirm that it routes to your app.

1. SSH into your TCP router.
2. Curl the port of your route using the IP address of the router itself. For example, if the port reserved for your route is `1024`, and the IP address of the TCP router is `10.0.16.18`:

```
$ curl 10.0.16.18:1024
```

3. If the curl is successful, then the load balancer either:

- o cannot reach the TCP routers, or
- o is not configured to route requests to the TCP routers from the `reservable_ports`

4. If you cannot reach the app by curling the TCP router directly, perform the following steps to confirm that your TCP route is in the routing table.

- a. Record the **Tcp Emitter Credentials** from the **UAA** row in the **PAS Credentials** tab. This OAuth client has permissions to read the routing table.
- b. Install the UAA CLI `uaac`:

```
$ gem install cf-uaac
```

- c. Obtain a token for this OAuth client from UAA by providing the client secret:

```
$ uaac token client get tcp_emitter
Client secret:
```

- d. Obtain an `access_token`:

```
$ uaac context
```

- e. Use the [routing API](#) to list TCP routes:

```
$ curl api.MY-DOMAIN/routing/v1/tcp_routes -H "Authorization: bearer TOKEN"
[{"router_group_guid": "f3518f7d-d8a0-4279-4e89-e058040d0000",
 "backend_port": 60000, "backend_ip": "10.244.00.0", "port": 60000, "modification_tag": {"guid": "d4cc3bbe-e838-4857-7360-19f034440000",
 "index": 1}, "ttl": 120}]
```

- f. In this output, each route mapping has the following:

- **port**: your route port
- **backend\_ip**: an app instance mapped to the route
- **backend\_port**: the port number for the app instance mapped to the route

- g. If your route port is not in the response, then the `tcp_emitter` may be unable to register TCP routes with the routing API. Look at the logs for `tcp_emitter` to see if there are any errors or failures.

- h. If the route is in the response, but you were not able to curl the port on the TCP route directly, then the TCP router may be unable to reach the routing API. Look at the logs for `tcp_router` to see if there are any errors or failures.

## Troubleshooting Router Error Responses

This topic helps operators to better understand if 502's are a result of the Pivotal Application Service (PAS) Platform or an application.

### Points of Failure

There are different points of failure in which 502's can come from:

#### 1. Infrastructure

- Load Balancer
- Network

#### 2. Platform - PCF

- Gorouter
- Diego Cell(s)

#### 3. Application

In the **Infrastructure**, 502's can occur in the following way:

- From the Load Balancer, 502's can surface when the Gorouters are not receiving traffic at all.
  - This can be observed if the Load Balancer is logging 502's but the Gorouters are not.

In the **Platform**, 502's can occur in the following ways:

- If the Gorouter is unable to connect to the application container:
  - TCP dial issues (can't make an initial connection to the [backend](#)). The Gorouter will retry TCP dial errors up to three times, if it still fails then a 502 will be returned to the client and logged to the `access.log`. This may be due to:
    - An application that is unresponsive (which indicates an issue with the application)
    - The Gorouter has a stale route (which indicates an issue with the platform)
    - The application container is corrupted (which indicates a problem with the platform)
    - These types of errors may look like this within the `gorouter.log`:

```
[2018-07-05 17:59:10+0000] {"log_level":3,"timestamp":1530813550.92134,"message": "backend-endpoint-failed","source":"vcap.gorouter","data":{"route-endpoint": {"ApplicationId":"","Addr":"10.0.32.15:60099","Tags":null,"RouteServiceUrl":""}, "error":"dial tcp 10.0.32.15:60099: getsockopt: connection refused"}}
```

- If the Gorouter successfully dials the endpoint but an error occurs:
  - `read: connection reset by peer` errors can occur when the application closes the connection abruptly with a TCP RST packet and not the expected FIN-ACK. This will cause the Gorouter to retry the next endpoint. Note, Gorouter does not currently retry on `write: connection reset by peer` failures.
  - TLS Handshake errors. When these errors occur, the Gorouter will retry up to three times and if it's still failing then a 502 may be returned. These errors appear similar to the following in the `gorouter.log` (and a 502 will be logged in the `access.log`):

```
[2018-07-05 18:20:54+0000] {"log_level":3,"timestamp":1530814854.4359834,"message": "backend-endpoint-failed","source":"vcap.gorouter","data":{"route-endpoint": {"ApplicationId":"","Addr":"10.0.16.17:61002","Tags":null,"RouteServiceUrl":""}, "error":"x509:certificate is valid for 53079ca3-c4fe-4910-78b9-c1a6, not xxx"}}
```

- If the Gorouter successfully connects to the endpoint, but an error occurs while the request is in transport (i.e. Gorouter has not received a response from the endpoint):
- Prior to PCF 2.0, there was a bug that logged a 502 for requests canceled by clients before the server responded with headers. PCF 2.0 and beyond, if the same situation occurs, a 499 is returned.

In an **Application**, 502's can occur in the following ways (Note: the Gorouter will not retry any error response that is returned by the application):

- If 502's are only occurring from a particular application instance(s) and not all of the applications on the platform, then it is likely an application-related error (i.e. application is overloaded, unresponsive, can't connect to database, etc.).
- If all applications are experiencing 502's, then it could either be a platform issue (possible misconfiguration) or an application issue (i.e. all applications are unable to connect to an upstream database).

## General Debugging Steps

Here are general debugging steps for any issue resulting with 502 error codes:

- Gather the Gorouter logs & Diego Cell logs at the time of the incident. To SSH into the router VM, see [Advanced Troubleshooting with the BOSH CLI](#). To download the router VM logs from Ops Manager, see [Monitoring PCF VMs from Ops Manager](#).
- Review the logs and consider the following questions:
  - Which errors are the Gorouters returning?
  - Is the Gorouter's [routing table](#) accurate (are the endpoints for the route as expected)?
  - Do the Diego Cell logs have anything interesting about unexpected app crashes and/or restarts?
  - Is the application healthy and handling requests successfully? (try using [request tracing headers](#) to verify)
- Was there a recent platform change/upgrade that caused an increase in 502's?
- Are there any suspicious [metrics](#) spiking? How is CPU & Memory utilization?

## Gorouter Error Classification Table

Error Type	Status Code	Source of Issue	Evidence
Dial	502	Application or Platform	logs with error <code>dial tcp</code>
AttemptedTLSWithNonTLSBackend	525	Platform*	- logs with error <code>tls: first record does not look like a TLS handshake</code> - <code>backend_tls_handshake_failed</code> metric increments
HostnameMismatch	503	Platform	- logs with error <code>x509: certificate is valid for &lt;x&gt; not &lt;y&gt;</code> - <code>backend_invalid_id</code> metric increments
UntrustedCert	526	Platform	- logs with an error prefix <code>x509: certificate signed by unknown authority</code> - <code>backend_invalid_tls_cert</code> metric increments
RemoteFailedCertCheck	496	Platform	logs with error remote <code>error: tls: bad certificate</code>
ContextCancelled	499	Client/App	logs with error <code>context canceled</code>  <div style="background-color: #ffffcc; padding: 10px;"> <span style="color: #808000;">💡</span> Note: This status code is never returned to clients, only logged, as it occurs when the downstream client closes the connection before Gorouter responds.         </div>
RemoteHandshakeFailure	525	Platform	- logs with error remote <code>error: tls: handshake failure</code> - <code>backend_tls_handshake_failed</code> metric increments

\*Note: any platform issue could be the result of a misconfiguration

For each of the above errors, there will be a `backend-endpoint-failure` log line in `gorouter.log` and an error message in `gorouter.err.log`. Additionally, the `access.log` will record the request status codes.

## Troubleshooting Ops Manager for VMware vSphere

This guide provides help with diagnosing and resolving issues that are specific to [Pivotal Cloud Foundry](#) (PCF) deployments on VMware vSphere.

For infrastructure-agnostic troubleshooting help, refer to [Diagnosing Problems in PCF](#).

## Common Issues

The following sections list common issues you might encounter and possible resolutions.

### PCF Installation Fails

If you modify the vCenter Statistics Interval Duration setting from its default setting of 5 minutes, the PCF installation might fail at the MicroBOSH deployment stage, and the logs might contain the following error message: `The specified parameter is not correct, [redacted]. This failure happens because Ops interval`

Manager expects a default value of 5 minutes, and the call to this method fails when the retrieved value does not match the expected default value.

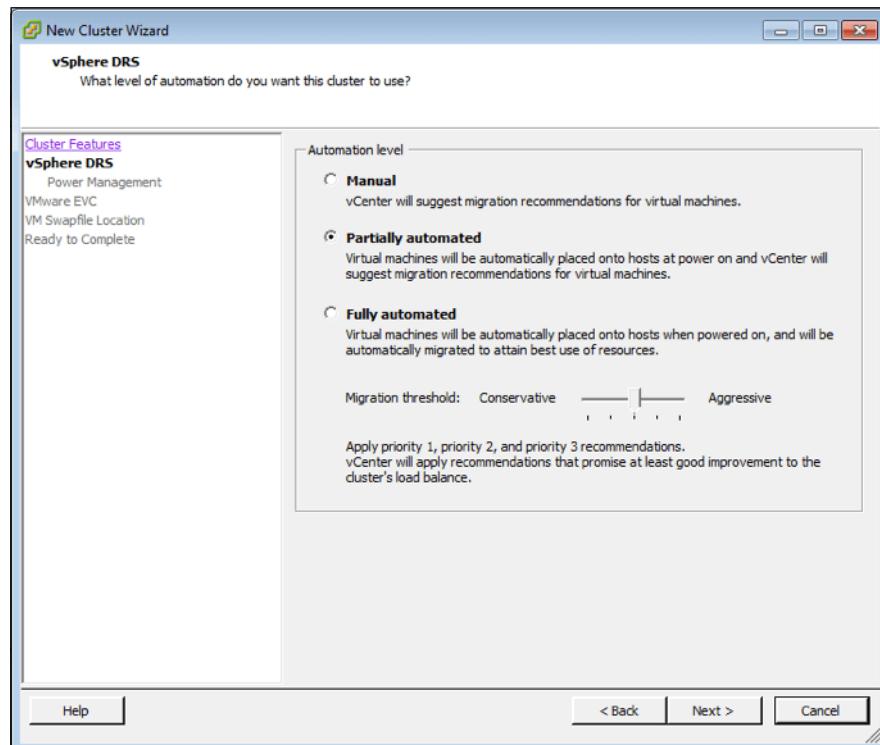
To resolve this issue, launch vCenter, navigate to **Administration > vCenter Server Settings > Statistics**, and reset the vCenter Statistics Interval Duration setting to 5 minutes.

### BOSH Automated Installation Fails

Before starting an Pivotal Application Service (PAS) deployment, you must set up and configure a vSphere cluster.

If you enable vSphere DRS (Distributed Resource Scheduler) for the cluster, you must set the Automation level to **Partially automated** or **Fully automated**.

If you set the Automation level to **Manual**, the BOSH automated installation will fail with a `[power_on_vm]` error when BOSH attempts to create virtual VMs.



### Ops Manager Loses Its IP Address After HA or Reboot

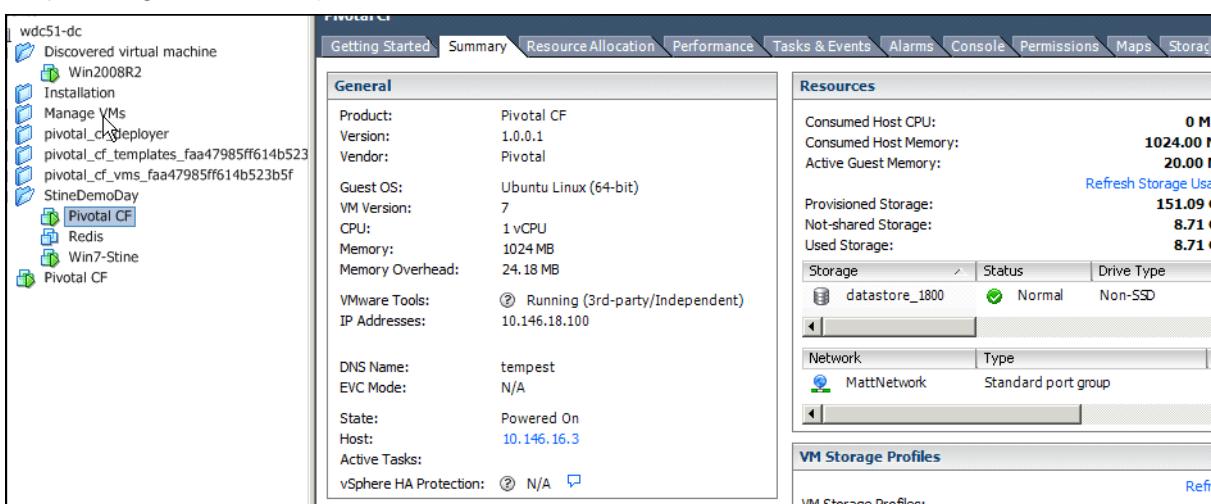
Ops Manager can lose its IP address and use DHCP due to an issue in the open source version of VMware Tools. For more information to troubleshoot this

issue, see [IP is not Getting Assigned in vSphere after Ops Manager Restart](#) in the Pivotal Support Knowledge Base.

## Cannot Connect to the OVF in a Browser

If you deployed the OVF file but cannot connect to it in a browser, check that the network settings you entered in the wizard are correct.

1. Access the PCF installation VM using the vSphere Console. If your network settings are configured incorrectly, you cannot SSH into the installation VM.
2. Log in using the credentials you provided when you imported the PCF .ova in vCenter.
3. Confirm that the network settings are correct by checking that the ADDRESS, NETMASK, GATEWAY, and DNS-NAMESERVERS entries are correct in `/etc/network/interfaces`.
4. If any of the settings are incorrect, run `sudo vi /etc/network/interfaces` and correct the incorrect entries.
5. In vSphere, navigate to the **Summary** tab for the VM and confirm that the network name is correct.



6. If the network name is incorrect, right click on the VM, select **Edit Settings > Network adapter 1**, and select the correct network.

7. Reboot the installation VM.

## Installation Fails with Failed Network Connection

If you experience a communication error while installing Ops Manager or MicroBOSH Director, check the following settings.

- Ensure that the routes are not blocked. vSphere environments use [NSX](#) for firewall and NAT/SNAT translation and load balancing. All communication between PCF VMs and vCenter or ESXi hosts route through the NSX firewall and are blocked by default.
- Open port 443. Ops Manager and MicroBOSH Director VMs require access to vCenter and all ESX through port 443.
- Allocate more IP addresses. BOSH requires that you allocate a sufficient number of additional dynamic IP addresses when configuring a reserved IP address range during installation. BOSH uses these IP addresses during installation to compile and deploy VMs, install PAS, and connect to services. We recommend that you allocate at least 36 dynamic IP addresses when deploying Ops Manager and PAS.

## Insufficient External Database Permissions

Upgrade issues can be caused when the external database user used for the network policy DB is given insufficient permissions. To avoid this upgrade issue, ensure that the `networkpolicyserver` database user has the `ALL PRIVILEGES` permission.