Red Hat OpenShift Container Platform on the AWS Cloud

Quick Start Reference Deployment

September 2017

Tony Vattathil, Andrew Glenn, David Duncan
Amazon Web Services (AWS)

Contents

Quick Starts are automated reference deployments for key technologies on the Amazon Web Services (AWS Cloud), based on AWS best practices for security and high availability.

Overview

This Quick Start reference deployment guide provides step-by-step instructions for deploying Red Hat OpenShift Container Platform on the AWS Cloud.

Red Hat OpenShift Container Platform is a platform as a service (PaaS) solution that is based on Docker-formatted Linux containers, Google Kubernetes orchestration, and Red Hat Enterprise Linux (RHEL) 7.

Red Hat OpenShift Container Platform gives application development and IT operations teams the ability to accelerate application delivery. It provides the following features:

- Support for standardized containers through the Red Hat APIs for Docker

- Container orchestration, scheduling, and management at scale with Kubernetes

- Integration with container-optimized RHEL 7 operating system

- Extensive selection of programming languages, frameworks, and services

- Rich set of tools and interfaces, including a web console and collaboration features, for development and operations

- Distributed application platform with container networking, streamlined deployment, and administration

For more information about Red Hat OpenShift Container Platform, see the OpenShift documentation.

This Quick Start is primarily for developers, engineers, architects, or DevOps or systems engineering staff who want to deploy OpenShift on the AWS Cloud. As an optional feature, the Quick Start also allows deployments using the upstream version of Ansible Playbook, for development purposes only.

OpenShift Container Platform on AWS

The Quick Start includes AWS CloudFormation templates that build the AWS infrastructure using AWS best practices, and then pass that environment to Ansible playbooks to build out the OpenShift environment. The AWS CloudFormation templates use AWS Lambda to generate a dynamic SSH key pair that is loaded into an Auto Scaling group. The Ansible inventory file is auto-generated accordingly. The combination of AWS CloudFormation and Ansible enables you to deploy and tear down your OpenShift environment by using CloudFormation stacks.

OpenShift Components

The Quick Start deploys the following OpenShift Container Platform components:

- **Master** provides master components such as the following:
    - API server (responsible for handling requests from clients, including nodes, users, administrators, and other infrastructure systems deployed to OpenShift)
    - Controller manager server (includes the scheduler and replication controller)
    - OpenShift client tools (`oc` and `oadm`).
- `etcd` stores the persistent master state while other components watch `etcd` for changes to bring themselves into the desired state.
- **Nodes** provide the runtime environments for containers.

Costs and Licenses

You are responsible for the cost of the AWS services used while running this Quick Start reference deployment. There is no additional cost for using the Quick Start.

The AWS CloudFormation template for this Quick Start includes configuration parameters that you can customize. Some of these settings, such as instance type, will affect the cost of deployment. For cost estimates, see the pricing pages for each AWS service you will be using. Prices are subject to change.

This Quick Start requires a Red Hat subscription.

Architecture

Deploying this Quick Start for a new virtual private cloud (VPC) with **default parameters** builds the following OpenShift Container Platform environment in the AWS Cloud.
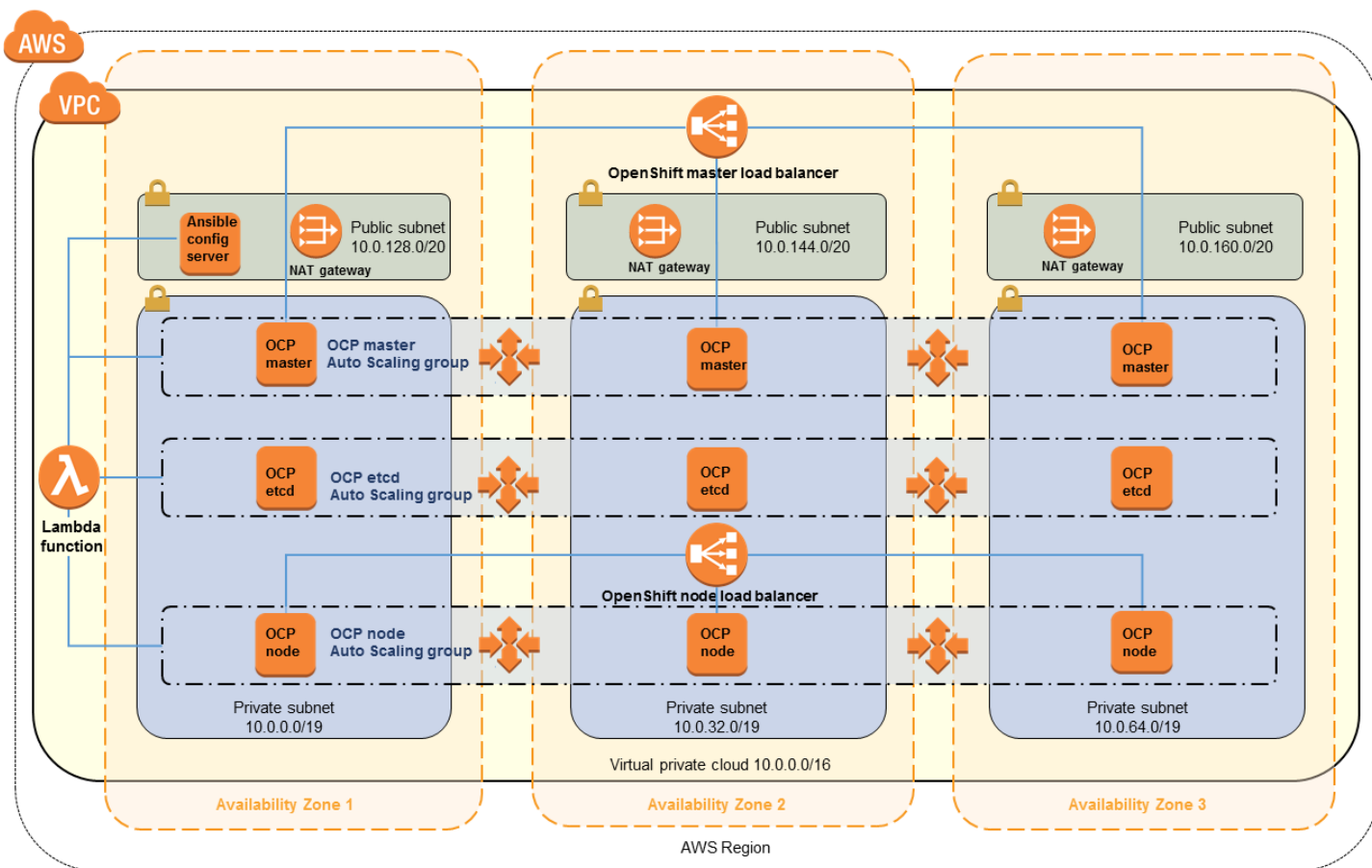
Figure 1: Quick Start architecture for OpenShift Container Platform on AWS

The Quick Start sets up the following:

- A single virtual private cloud (VPC) that spans three Availability Zones, with one private and one public subnet in each Availability Zone.*

- An Internet gateway to provide Internet access to each subnet.*

- In one of the public subnets, an Ansible config server instance.

- In the private subnets:

  – Three OpenShift master instances in an Auto Scaling group

  – Three OpenShift `etcd` instances in an Auto Scaling group

  – A variable number of OpenShift node instances in an Auto Scaling group

> **Note**   The template that deploys the Quick Start into an existing VPC skips the tasks marked by asterisks.

Auto Scaling Use

In this initial release, the Quick Start places the OpenShift instances in Auto Scaling groups, but doesn't enable scaling. The number of master and `etcd` instances are fixed at three (one per Availability Zone). Nodes can be set to a variable number and will be distributed among the selected Availability Zones.

Prerequisites

Specialized Knowledge

Before you deploy this Quick Start, we recommend that you become familiar with the following AWS services. (If you are new to AWS, see Getting Started with AWS.)

- Amazon VPC
- Amazon EC2
- Amazon EBS
- Amazon Lambda

Deployment Options

This Quick Start provides two deployment options:

- **Deploy OpenShift Container Platform into a new VPC** (end-to-end deployment). This option builds a new AWS environment consisting of the VPC, subnets, NAT gateways, security groups, and other infrastructure components, and then deploys OpenShift Container Platform into this new VPC.

- **Deploy OpenShift Container Platform into an existing VPC**. This option provisions OpenShift Container Platform in your existing AWS infrastructure.

The Quick Start provides separate templates for these options. It also lets you configure CIDR blocks, instance types, and OpenShift Container Platform settings, as discussed later in this guide.

Deployment Steps

Step 1. Sign up for a Red Hat Subscription

This Quick Start requires a valid Red Hat subscription. During the deployment of the Quick Start, you'll need to provide your Red Hat subscription user name, password, and pool ID. You can sign up for a subscription at https://www.redhat.com/wapps/ugc/register.html.

Step 2. Prepare Your AWS Account

1. If you don't already have an AWS account, create one at https://aws.amazon.com by following the on-screen instructions.

2. Use the region selector in the navigation bar to choose the AWS Region where you want to deploy OpenShift Container Platform on AWS.

3. Create a key pair in your preferred region.

4. If necessary, request a service limit increase for the Amazon EC2 **M4** instance type. You might need to do this if you already have an existing deployment that uses this instance type, and you think you might exceed the default limit with this reference deployment.

Step 3. Launch the Quick Start

> **Note**   You are responsible for the cost of the AWS services used while running this Quick Start reference deployment. There is no additional cost for using this Quick Start. For full details, see the pricing pages for each AWS service you will be using in this Quick Start. Prices are subject to change.

1. Choose one of the following options to launch the AWS CloudFormation template into your AWS account. For help choosing an option, see deployment options earlier in this guide.

Option 1	Option 2
Deploy OpenShift Container Platform into a new VPC	**Deploy OpenShift Container Platform into an existing VPC**
Launch	**Launch**

aws

> **Important**   If you're deploying OpenShift into an existing VPC, make sure that your VPC has three private subnets in different Availability Zones for the OpenShift instances. These subnets require NAT gateways or NAT instances in their route tables, to allow the instances to download packages and software without exposing them to the Internet. You'll also need the domain name option configured in the DHCP options as explained in the [Amazon VPC documentation](). You'll be prompted for your VPC settings when you launch the Quick Start.

Each deployment takes about 1.5 hours to complete.

2. Check the region that's displayed in the upper-right corner of the navigation bar, and change it if necessary. This is where the network infrastructure for OpenShift will be built. The template is launched in the US East (Ohio) Region by default.

3. On the **Select Template** page, keep the default setting for the template URL, and then choose **Next**.

4. On the **Specify Details** page, change the stack name if needed. Review the parameters for the template. Provide values for the parameters that require input. For all other parameters, review the default settings and customize them as necessary. When you finish reviewing and customizing the parameters, choose **Next**.

In the following tables, parameters are listed by category and described separately for the two deployment options:

– [Parameters for deploying OpenShift Container Platform into a new VPC]()

– [Parameters for deploying OpenShift Container Platform into an existing VPC]()

• **Option 1: Parameters for deploying OpenShift into a new VPC**

[View template]()

VPC Network Configuration:

Parameter label (name)	Default	Description
Availability Zones (AvailabilityZones)	*Requires input*	The list of Availability Zones to use for the subnets in the VPC. The Quick Start uses three Availability Zones from your list and preserves the logical order you specify.
VPC CIDR (VPCCIDR)	10.0.0.0/16	CIDR block for the VPC.
Private Subnet 1 CIDR (PrivateSubnet1CIDR)	10.0.0.0/19	CIDR block for the private subnet located in Availability Zone 1.

Parameter label (name)	Default	Description
Private Subnet 2 CIDR (PrivateSubnet2CIDR)	10.0.32.0/19	CIDR block for the private subnet located in Availability Zone 2.
Private Subnet 3 CIDR (PrivateSubnet3CIDR)	10.0.64.0/19	CIDR block for the private subnet located in Availability Zone 3.
Public Subnet 1 CIDR (PublicSubnet1CIDR)	10.0.128.0/20	CIDR block for the public (DMZ) subnet located in Availability Zone 1.
Public Subnet 2 CIDR (PublicSubnet2CIDR)	10.0.144.0/20	CIDR block for the public (DMZ) subnet located in Availability Zone 2.
Public Subnet 3 CIDR (PublicSubnet3CIDR)	10.0.160.0/20	CIDR block for the public (DMZ) subnet located in Availability Zone 3.
Allowed External Access CIDR (RemoteAccessCIDR)	*Requires input*	The CIDR IP range that is permitted to access the OpenShift instances. We recommend that you set this value to a trusted IP range. For example, you might want to grant only your corporate network access to the software.

Amazon EC2 Configuration:

Parameter label (name)	Default	Description
SSH Key Name (KeyPairName)	*Requires input*	Public/private key pair, which allows you to connect securely to your instance after it launches. When you created an AWS account, this is the key pair you created in your preferred region. All instances will launch with this key pair.

OpenShift Nodes Configuration:

Parameter label (name)	Default	Description
Number of Masters (NumberOfMaster)	3	The number of OpenShift master instances to provision. This deployment requires three OpenShift master instances.
Number of Etcds (NumberOfEtcd)	3	The number of OpenShift `etcd` instances to provision. This deployment requires three OpenShift `etcd` instances.
Number of Nodes (NumberOfNodes)	3	The number of OpenShift node instances to provision. You can choose any number of instances. **Warning** If the number of node instances exceeds your Red Hat entitlement limits or AWS instance limits, the stack will fail. Choose a number that is within your limits.
Master Instance Type (MasterInstanceType)	m4.xlarge	EC2 instance type for the OpenShift master instances.
Etcd Instance Type (EtcdInstanceType)	m4.xlarge	EC2 instance type for the OpenShift `etcd` instances.

Parameter label (name)	Default	Description
Nodes Instance Type (NodesInstanceType)	m4.xlarge	EC2 instance type for the OpenShift node instances.

Red Hat Subscription Information:

Parameter label (name)	Default	Description
Red Hat Subscription User Name (RedhatSubscription UserName)	*Requires input*	Your Red Hat (RHN) user name, from step 1.
Red Hat Subscription Password (RedhatSubscription Password)	*Requires input*	Your Red Hat (RHN) password, from step 1.
Red Hat Pool ID (RedhatSubscription PoolID)	*Requires input*	Your Red Hat (RHN) subscription pool ID, from step 1.

Ansible Playbook Configuration:

Parameter label (name)	Default	Description
Ansible Playbook Mode (AnsiblePlayBookType)	Subscription-Version	The Ansible Playbook version to use. If you're using this deployment for production, keep the default setting. If you're using this deployment for development purposes, you can choose **OpenSource-Version**.
Git Repo Release Version (AnsiblePlayBookGitRepo Tag)	3.6.173.0.5-5	This parameter is used only if you choose **OpenSource-Version** for the **Ansible Playbook Mode** parameter. You can specify one of the development releases available at https://github.com/openshift/openshift-ansible/releases. (For more information, see the Ansible Playbook Releases section.)

AWS Quick Start Configuration:

Parameter label (name)	Default	Description
Quick Start S3 Bucket Name (QSS3BucketName)	quickstart-reference	S3 bucket where the Quick Start templates and scripts are installed. Use this parameter to specify the S3 bucket name you've created for your copy of Quick Start assets, if you decide to customize or extend the Quick Start for your own use. The bucket name can include numbers, lowercase letters,

aws

Parameter label (name)	Default	Description
		uppercase letters, and hyphens, but should not start or end with a hyphen.
Quick Start S3 Key Prefix (QSS3KeyPrefix)	redhat/openshift/ latest/	The S3 key name prefix used to simulate a folder for your copy of Quick Start assets, if you decide to customize or extend the Quick Start for your own use. This prefix can include numbers, lowercase letters, uppercase letters, hyphens, and forward slashes.

- **Option 2: Parameters for deploying OpenShift into an existing VPC**

  View template

  Network Configuration:

Parameter label (name)	Default	Description
VPC ID (VPCID)	*Requires input*	ID of your existing VPC (e.g., vpc-0343606e).
VPC CIDR (VPCCIDR)	10.0.0.0/16	CIDR block for the VPC.
Private Subnet 1 ID (PrivateSubnet1ID)	*Requires input*	ID of the private subnet in Availability Zone 1 in your existing VPC (e.g., subnet-a0246dcd).
Private Subnet 2 ID (PrivateSubnet2ID)	*Requires input*	ID of the private subnet in Availability Zone 2 in your existing VPC (e.g., subnet-b58c3d67).
Private Subnet 3 ID (PrivateSubnet3ID)	*Requires input*	ID of the private subnet in Availability Zone 3 in your existing VPC (e.g., subnet-b1f4a2cd).
Public Subnet 1 ID (PublicSubnet1ID)	*Requires input*	ID of the public subnet in Availability Zone 1 in your existing VPC (e.g., subnet-9bc642ac).
Public Subnet 2 ID (PublicSubnet2ID)	*Requires input*	ID of the public subnet in Availability Zone 2 in your existing VPC (e.g., subnet-e3246d8e).
Public Subnet 3 ID (PublicSubnet3ID)	*Requires input*	ID of the public subnet in Availability Zone 3 in your existing VPC (e.g., subnet-e3246d7f).
Allowed External Access CIDR (RemoteAccessCIDR)	*Requires input*	The CIDR IP range that is permitted to access the OpenShift instances. We recommend that you set this value to a trusted IP range. For example, you might want to grant only your corporate network access to the software.

Amazon EC2 Configuration:

Parameter label (name)	Default	Description
SSH Key Name (KeyPairName)	*Requires input*	Public/private key pair, which allows you to connect securely to your instance after it launches. When you created an AWS account, this is the key pair you created in your preferred region. All instances will launch with this key pair.

OpenShift Nodes Configuration:

Parameter label (name)	Default	Description
Number of Masters (NumberOfMaster)	3	The number of OpenShift master instances to provision. This deployment requires three OpenShift master instances.
Number of Etcds (NumberOfEtcd)	3	The number of OpenShift `etcd` instances to provision. This deployment requires three OpenShift `etcd` instances.
Number of Nodes (NumberOfNodes)	3	The number of OpenShift node instances to provision. You can choose any number of instances. **Warning** If the number of node instances exceeds your Red Hat entitlement limits or AWS instance limits, the stack will fail. Choose a number that is within your limits.
Master Instance Type (MasterInstanceType)	m4.xlarge	EC2 instance type for the OpenShift master nodes.
Etcd Instance Type (EtcdInstanceType)	m4.xlarge	EC2 instance type for the OpenShift etcd nodes.
Nodes Instance Type (NodesInstanceType)	m4.xlarge	EC2 instance type for the OpenShift nodes.

Ansible Playbook Configuration:

Parameter label (name)	Default	Description
Ansible Playbook Mode (AnsiblePlayBookType)	Subscription-Version	The Ansible Playbook version to use. If you're using this deployment for production, keep the default setting. If you're using this deployment for development purposes, you can choose **OpenSource-Version**.
Git Repo Release Version (AnsiblePlayBookGitRepo Tag)	3.6.173.0.5-5	This parameter is used only if you choose **OpenSource-Version** for the **Ansible Playbook Mode** parameter. You can specify one of the development releases available at https://github.com/openshift/openshift-ansible/releases. (For more information, see the Ansible Playbook Releases section.)

Red Hat Subscription Information:

Parameter label (name)	Default	Description
Red Hat Subscription User Name (RedhatSubscription UserName)	*Requires input*	Your Red Hat (RHN) user name, from step 1.
Red Hat Subscription Password (RedhatSubscription Password)	*Requires input*	Your Red Hat (RHN) password, from step 1.
Red Hat Pool ID (RedhatSubscription PoolID)	*Requires input*	Your Red Hat (RHN) subscription pool ID, from step 1.

AWS Quick Start Configuration:

Parameter label (name)	Default	Description
Quick Start S3 Bucket Name (QSS3BucketName)	quickstart-reference	S3 bucket where the Quick Start templates and scripts are installed. Use this parameter to specify the S3 bucket name you've created for your copy of Quick Start assets, if you decide to customize or extend the Quick Start for your own use. The bucket name can include numbers, lowercase letters, uppercase letters, and hyphens, but should not start or end with a hyphen.
Quick Start S3 Key Prefix (QSS3KeyPrefix)	redhat/openshift/ latest/	The S3 key name prefix used to simulate a folder for your copy of Quick Start assets, if you decide to customize or extend the Quick Start for your own use. This prefix can include numbers, lowercase letters, uppercase letters, hyphens, and forward slashes.

5. On the **Options** page, you can specify tags (key-value pairs) for resources in your stack and set advanced options. When you're done, choose **Next**.

6. On the **Review** page, review and confirm the template settings. Under **Capabilities**, select the check box to acknowledge that the template will create IAM resources.

7. Choose **Create** to deploy the stack.

8. Monitor the status of the stack. When the status is **CREATE_COMPLETE**, the OpenShift Container Platform cluster is ready.

9. Use the URLs displayed in the **Outputs** tab for the stack to view the resources that were created.

Step 4. Test the Deployment

Verify that OpenShift services are running

OpenShift components are deployed into multiple private subnets. You can access the OpenShift web console by using the OpenShiftMasterELB on port 8443. You can also connect to one of the OpenShift master nodes and use the OpenShift command line interface (CLI). To log in, you'll use SSH agent forwarding to hop from the Ansible config server to the master node. (The SSH agent will provide your private key on connection.)

> **Important**   Do not copy your private key to the Ansible config server.

For more information on SSH agents, see the [GitHub documentation.](#)

1. Use an SSH agent to access the Ansible config server environment on MacOS or Linux, by using the command:

    ```
    ssh-add ~/.ssh/id_rsa
    ```

2. At the prompt, type your passphrase or press **Enter** for no passphrase.

    ```
    Enter passphrase (empty for no passphrase): [Hit Enter Again or
    Enter passphrase]
    Enter same passphrase again: [Hit Enter Again or Enter passphrase]
    ```

3. Open the Amazon EC2 console at https://console.aws.amazon.com/ec2/, select the instance tagged ansible-configserver, and note its public address.

Figure 2: Finding the public DNS of the `ansible-configserver` instance

4. SSH as `ec2-user` with your key pair to the config server, and then enter **yes** to connect.

5. From the config server, connect to a master node:

     a. From the EC2 console select one of the instances tagged `openshift-master` and note the private DNS name or IP of the node

Figure 3: OpenShift master nodes

b. Connect to the OpenShift master node's private DNS name or IP address and **sudo** to become `root`:

```
$ sudo -s
```

c. On the Master node **as root**, run `oc get pods` and verify that services are in the running state:

```
$ oc get pods
```


Connect to the OpenShift Web Console

1. Open the AWS CloudFormation console at https://console.aws.amazon.com/cloudformation/, and then select the OpenShift stack.
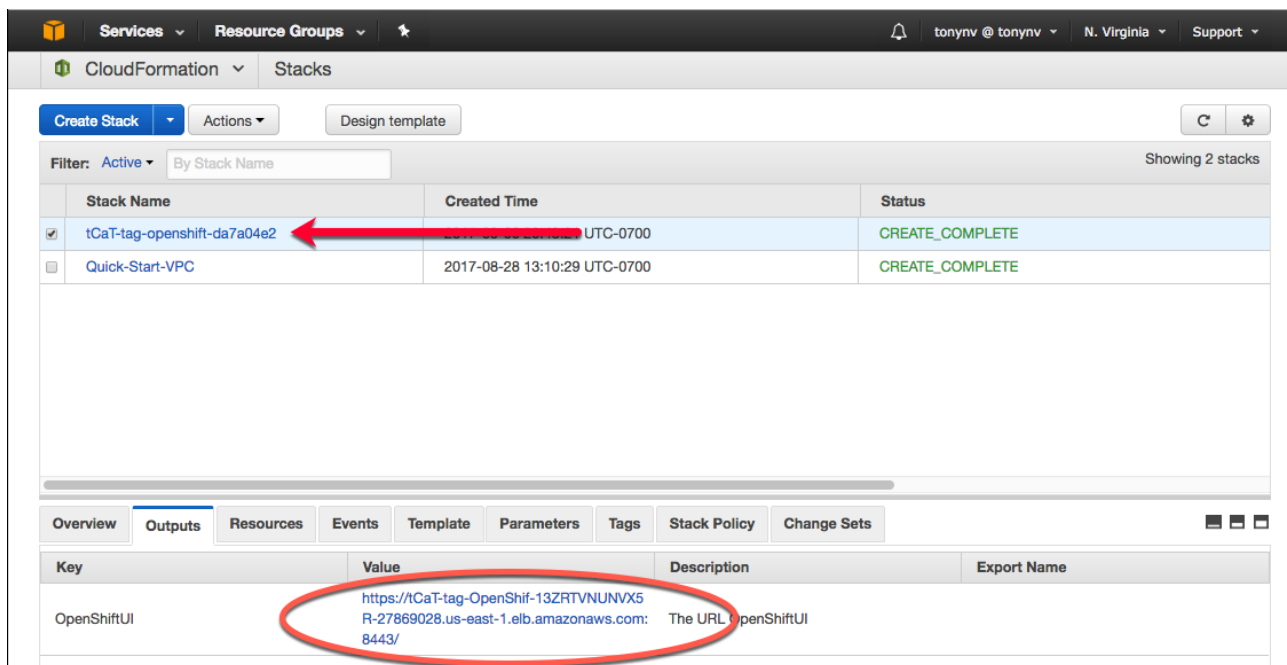

Figure 4: OpenShift stack and console link

2. Point your browser to the value for the `OpenShiftUI` key to connect to the console.

3. Accept the self-signed certificate warnings (there will be a few redirects on the initial connection) to reach the OpenShift user interface.

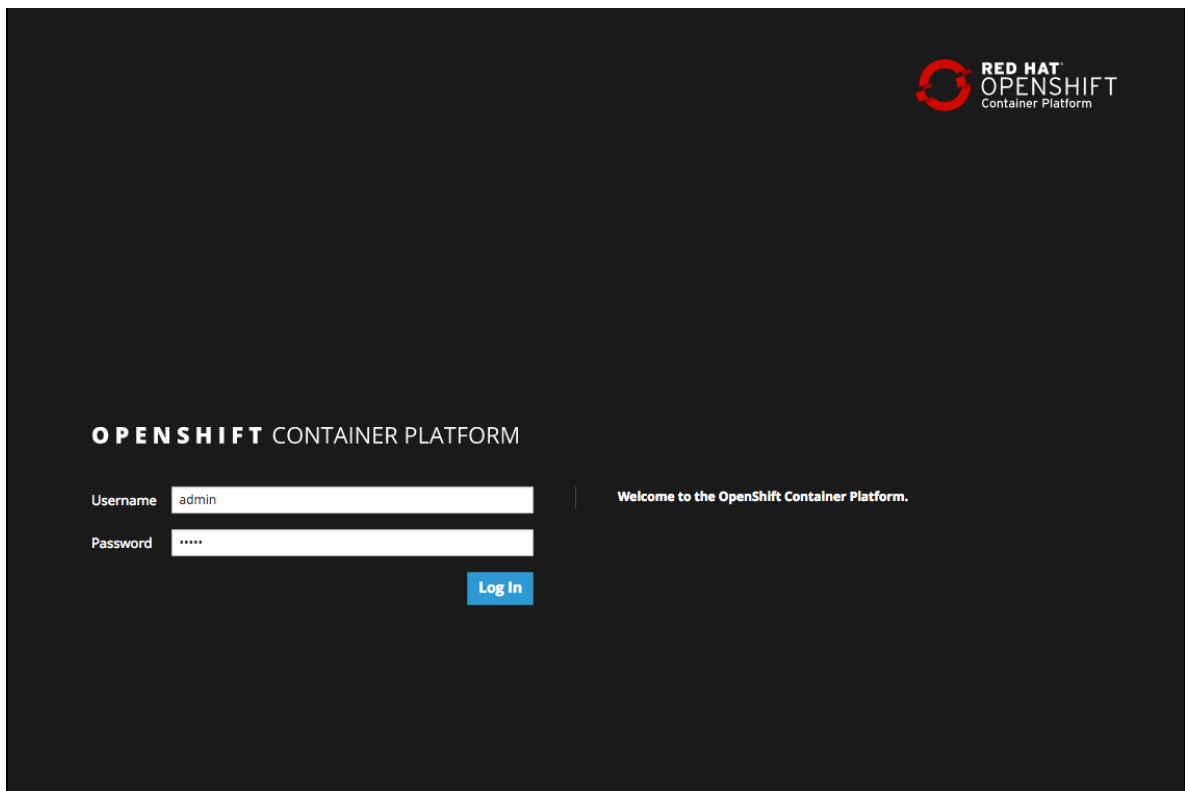> **Note**   The default login *(username/password)* is `admin/admin`.

Figure 5: OpenShift login screen

> **Important**   Please change your password upon login. For more information, see
> [Managing Users in the OpenShift documentation.](#)

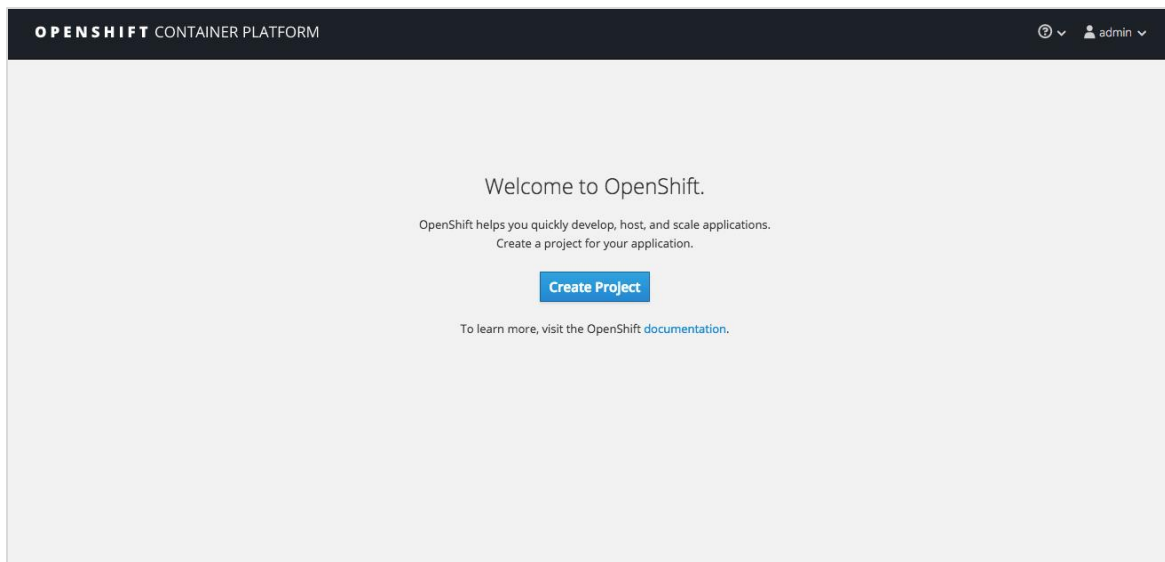Upon login, you will reach the **Create Project** screen.

Figure 6: OpenShift Create Project screen

4. Follow the instructions in the [OpenShift documentation](#) to create projects and applications.

If you're using OpenShift Container Platform for production use, we recommend switching to a CA-signed certificate. For details, see [Certificate Management](#) in the OpenShift documentation.

For further customization and additional functionality, see the links in the [Additional Resources](#) section.

Best Practices for Using OpenShift Container Platform on AWS

This Quick Start deploys the AWS infrastructure for OpenShift and generates Ansible inventory files based on your selected instance types and VPC configuration. The AWS CloudFormation templates follow Quick Start best practices for AWS resource management and the best practices dictated by Red Hat for Ansible Playbook and OpenShift.

Security

By default, this Quick Start does not allow direct access to OpenShift nodes and limits access to ports 22 and 8443. If you want to expose additional ports for added functionality, you can adjust the OpenShift security group and ELB load balancer accordingly.

Ansible Playbook Releases

This Quick Start has been tested with the officially released version of Ansible Playbook for production environments, as specified by the default setting (**Subscription-Version**) of the **Ansible Playbook Mode** parameter. If you're using this deployment in a development environment, you can select **OpenSource-Version** for the **Ansible Playbook Mode** parameter, and specify a development release from https://github.com/openshift/openshift-ansible/releases. However, note that this Quick Start hasn't been tested with all development releases, which may include changes to core components.

Troubleshooting

Q. I encountered a CREATE_FAILED error when I launched the Quick Start. What should I do?

A. If AWS CloudFormation fails to create the stack, we recommend that you relaunch the template with **Rollback on failure** set to **No**. (This setting is under **Advanced** in the AWS CloudFormation console, **Options** page.) With this setting, the stack's state will be retained and the instance will be left running, so you can troubleshoot the issue. (You'll want to look at the log files in `%ProgramFiles%\Amazon\EC2ConfigService` and `C:\cfn\log`.)

> **Important**   When you set **Rollback on failure** to **No**, you'll continue to incur AWS charges for this stack. Please make sure to delete the stack when you've finished troubleshooting.

For additional information, see Troubleshooting AWS CloudFormation on the AWS website.

Q. I encountered a size limitation error when I deployed the AWS Cloudformation templates.

A. We recommend that you launch the Quick Start templates from the location we've provided or from another S3 bucket. If you deploy the templates from a local copy on your computer or from a non-S3 location, you might encounter template size limitations when you create the stack. For more information about AWS CloudFormation limits, see the AWS documentation.

Additional Resources

AWS services

- Amazon EC2
  https://docs.aws.amazon.com/AWSEC2/latest/WindowsGuide/
- AWS CloudFormation
  https://aws.amazon.com/documentation/cloudformation/
- Amazon VPC
  https://aws.amazon.com/documentation/vpc/

OpenShift Container Platform

- Getting Started with the CLI
  https://docs.openshift.com/container-platform/3.6/cli_reference/get_started_cli.html
- Web console walkthrough
  https://docs.openshift.com/container-platform/3.6/getting_started/developers_console.html

Quick Start reference deployments

- AWS Quick Start home page
  https://aws.amazon.com/quickstart/

Send Us Feedback

You can visit our GitHub repository to download the templates and scripts for this Quick Start, to post your comments, and to share your customizations with others.

Document Revisions

Date	Change	In sections
September 2017	Initial publication	—

© 2017, Amazon Web Services, Inc. or its affiliates. All rights reserved.