VMware Validated Design for Cloud Providers: Scale and Performance Guidelines

for VMware Cloud Director 10.2 environments

Published January 2021

Table of contents

List of Tables

List of Figures

Introduction

The VMware Validated Designs (VVD) for Cloud Providers: Scale and Performance Guidelines is a pre-validated set of software components that simplify the deployment of a VMware Cloud Director® - based multitenant cloud in a predictable and efficient manner. The intent of the Scale and Performance initiative is to document a verified stack and provide scale and performance benchmarking. It also helps reduce the complexity of figuring out dependencies between the VMware components required for a Cloud Director-based service. While this initiative does not yet involve software automation for software upgrades, it aims to present clearly what components are needed, which versions must be used, and what kind of scale and performance VMware Cloud Providers can expect. Partners who deploy the core technologies outlined in this VVD for Cloud Provider architecture and who meet other programmatic requirements are also eligible to attain the VMware Cloud Verified designation.

Partners who display the VMware Cloud Verified badge are promoted to VMware customers as specialists who can deliver the full value of a VMware software defined datacenter delivered as a service. In order to achieve VMware Cloud Verified, VMware Cloud Providers must meet requirements and apply directly via the *VMware Technology Validation* page. For complete requirements and detailed application process, please refer to *How to Achieve Cloud Verified* and *Cloud Verified Application Process* documents.

VMware Cloud Providers get clarity and predictability about which version of each software component of the stack is recommended at a given time. Each Scale and Performance version also includes a predictable support timeframe for all underlying components, typically 12 – 18 months from the launch of the corresponding Scale and Performance release. This reduces the expense and time involved in determining what components to upgrade when and to which version, so that the entire software stack stays in support and incompatible combinations are avoided.

VMware Cloud Providers also benefit from clear guidelines for sizing hardware and software components to match their expected tenant load. While the Scale and Performance does not cover every cloud configuration and size, it provides a sizing recommendation for a "typical" cloud (a cloud size representative of a broad set of VMware Cloud Providers). Future versions of the Scale and Performance might address larger and less common environment configurations as well as more specialized use cases.

It is not the current intent of Scale and Performance to push VMware Cloud Director to its absolute limits. For configuration maximums and limits, see VMware Cloud Director Configuration Maximums.

This document also includes the expected performance as observed by tenant users and VMware Cloud Provider administrators interacting with the VMware Cloud Director user interface and API.

A VMware Cloud Director-based platform can be properly sized by following the sizing guidelines for hardware and scale based on anticipated tenant demand.

Audience

This document is intended for VMware Cloud Provider architects and technical leads responsible for planning and performing the deployment and upgrades of a VMware-based cloud environment.

Scope

This document addresses the following aspects:

Interoperability stack

Provides a list of certified versions of all the component software comprising the software stack. Using the recommended versions guarantees known support life of the stack as well as performance characteristics.

Sizing guidelines and software requirements

Performance characteristics of the solution

The certified solution stack provides known performance and scale characteristics and includes recommendations and guidelines for hardware and scale based on anticipated tenant demand.

See the complimentary documents that are part of the *VMware Cloud® Architecture Toolkit™ for Service Providers* and *VMware Validated Design Documentation*.

The *VMware Product Interoperability Matrices* is the authoritative resource for interoperability between the VMware software components.

A compliant solution must comply with all relevant security guidelines outlined in the product-specific documentation as well as security recommendations in the VMware Cloud Architecture Toolkit for Service Providers document.

Interoperability Stack (Bill of Materials)

The Bill of Materials table lists the pre-validated set of software components for Cloud Providers at the time of the Scale and Performance launch. While VMware Cloud Providers are free to choose and pick other versions or different combinations of VMware Cloud Provider Program software products, the specified stack guarantees a known predictable support time and specific performance and scaling characteristics. Performance and scaling information is provided later this document. Products marked "Core" are required to officially achieve VVD for Cloud Providers compliance.

TABLE 1. BILL OF MATERIALS

BILL OF MATERIALS [1]			
COMPONENT	VERSION AND BUILD	CORE/ OPTIONAL	NOTES
VMware vCenter Server®	7.0 U1 [2]	Core	Bundled as part of VMware Cloud Foundation 4.1
VMware ESXi™	7.0 U1 [2]	Core	Bundled as part of VMware Cloud Foundation 4.1
VMware NSX-T	3.0.2 [2]	Core	Bundled as part of VMware Cloud Foundation 4.1
VMware vSAN	7.0 U1	Core [3]	Bundled as part of VMware Cloud Foundation 4.1
VMware Cloud Director	10.2	Core	10.2 Virtual Appliance with an embedded database.
VMware Cloud Director Availability	4.1	Optional	
VMware Cloud Director Container Service Extension	3.0.0	Optional	Container Services Extension
VMware vRealize® Log Insight™	8.1.1	Optional	Bundled as part of VMware Cloud Foundation 4.1
VMware vRealize® Network Insight™	5.3	Optional	Bundled as part of VMware Cloud Foundation 4.1
VMware vRealize® Orchestrator™	8.2	Optional	Apply the latest available patch.
VMware vCloud Usage Meter	4.3	Core	
VMware vRealize Operations™	8.1.1	Optional	Bundled as part of VMware Cloud Foundation 4.1
VMware vRealize Operations Manager Tenant App for VMware Cloud Director	2.5	Optional	
VMware Cloud Foundation	4.1	Optional	
VMware Cloud Foundation SDDC Manager	4.1	Optional	Bundled as part of VMware Cloud Foundation 4.1
VMware vRealize Suite Lifecycle Manager	8.1 patch 1	Optional	Bundled as part of VMware Cloud Foundation 4.1
VMware Cloud Director App Launchpad	2.0	Optional	

[1] These are the recommended set of products, but this is not a full interoperability matrix. VMware Cloud Director 10.2 is supported with multiple versions of NSX-T but in the current benchmarking we used a specific NSX-T version. Test results generally apply to all patches within the specified major version of each component. See the *VMware Product Interoperability Matrix* for full Cloud Director interoperability information

[2] See *TABLE 8* for patch level tested

[3] vSAN-based storage must be deployed in at least one cluster (either management or capacity).

BILL OF MATERIALS [1]			
COMPONENT	VERSION AND BUILD	CORE/ OPTIONAL	NOTES
VMware Cloud Director Object Storage Extension	2.0	Optional	
VMware NSX Advanced Load Balancer	20.1.1	Optional	

Scale and Performance

The Scale Profile B table represents a common environment similar to the environments of approximately 60% of all VMware Cloud Providers. While VMware Cloud Director is capable of a larger scale, the following profile is what is validated and benchmarked in the current Scale and Performance.

TABLE 2. SCALE PROFILE B

SCALE PROFILE B	
PARAMETER	VALUE
Number of tenants (Organizations in Cloud Director)	400
Number of powered-on tenant virtual machines (with an OS installed)	10,000
Number of physical data centers	1
Number of VMware Cloud Director cells	5 (1 primary, 2 stand-by cells, 2 application cells)
Number of vCenter Server instances managed by VMware Cloud Director	3 vCenter Server instances for resource capacity
Maximum network latency from VMware Cloud Director to VMware vCenter Server and VMware NSX-T Manager™	Network RTT latency up to 150 ms
Concurrent API operations	Up to 128 concurrent users running operations against the VMware Cloud Director API
Concurrent virtual machine migrations to VMware Cloud Director from tenant environments by VMware Cloud Director Availability	100

Performance Characteristics

Environment Setup

The multitenant cloud environment is set up based on Scale Profile B.

Testing is performed at different levels of network latency from VMware Cloud Director cells to vCenter Server and NSX-T Manager to measure the impact of network latency on performance.

Performance and Throughput

The test throughput is measured as the number of operations performed over 30 minutes. The test was run with different test concurrency (32, 64, and 128) and network latency (0.3 ms, 40 ms, and 150 ms). During this test, a representative random sample of operations from the *List of Operations* (Table 9) is used.

TABLE 3. PERFORMANCE AND THROUGHPUT

PERFORMANCE AND THROUGHPUT			
CONCURRENCY (NUMBER OF CONCURRENT USERS)	THROUGHPUT AT RTT = 0.3 MS (SUCCESSFULLY COMPLETED OPERATIONS PER MINUTE)	THROUGHPUT AT RTT = 40 MS	THROUGHPUT AT RTT = 150 MS
32	144	136	115
64	268	242	215
128	431	380	356

API Latency

The API Operations Latency table shows average user observed latency (in seconds) for a selection of API operations at RTT = 0.3 ms. See the *List of Operations* (Table 9) for the full list of operations invoked during this test.

TABLE 4. API OPERATIONS LATENCY

API OPERATIONS LATENCY			
OPERATION	CONCURRENCY (SECONDS)		
	32	64	128
Instantiate a 150 MB vApp from a template	17 s	18 s	25 s
Create an edge gateway	6 s	7 s	9 s
Create an independent disk	9 s	10 s	13 s

Increasing the network RTT from 0.3 ms to 150 ms affects these numbers with the size of the effect varying significantly depending on the operation. With most API operations RTT increase from 0.3 ms to 150 ms caused the latency to increase by a factor of 2 or less.

Upload and Download Performance

The OVF upload and download times observed in the test environment vary depending on the different network latencies.

TABLE 5. OVF UPLOAD AND DOWNLOAD TIMES

OVF UPLOAD AND DOWNLOAD TIMES			
	RTT = 0.3 MS	RTT = 40 MS	RTT= 150 MS
OVF upload time in seconds (4 GB)	306	322	325
OVF download time in seconds (4 GB)	233	234	238

VMware Cloud Director Availability

The Time to Protect a VM in VMware Cloud Director Availability represents the time to establish replication of virtual machines of various sizes between a VMware Cloud Director and a DR target vCenter environment. 10 GB uplinks were configured between vCenter Server and VMware Cloud Director. Network throughput was stable around 910 Mbps.

TABLE 6. TIME TO PROTECT A VM

TIME TO PROTECT A VM	
VM SIZE	TIME TO PROTECT
1 GB	35 seconds
10 GB	4 minutes
100 GB	20 minutes
500 GB	2 hours 4 minutes
1 TB	3 hours 56 minutes

The Network Latency Impact on Migration Performance table displays how the network latency from VMware Cloud Director cells to vCenter Server and NSX-T Manager impacts cold migration for virtual machines of 100 GB size.

TABLE 7. NETWORK LATENCY IMPACT ON TIME TO PROTECT (VM SIZE = 100 GB)

NETWORK LATENCY IMPACT ON TIME TO PROTECT (VM SIZE = 100 GB)	
RTT LATENCY BETWEEN VCENTER SERVER AND VMWARE CLOUD DIRECTOR	TIME TO PROTECT
0.3 ms	20 minutes
40 ms	27 minutes
150 ms	30 minutes

Sizing Guidelines

Many environment variables influence the number of hosts and CPU and memory resources required to run a cloud service based on VVD for Cloud Providers. It is impossible to give a precise formula for how much of each resource is required. The current effort focuses on demonstrating how our deployment setup behaves in terms of scale and performance under the defined test load.

The Scale Profile B captures the parameters of the load on the system in terms of number of tenants, organizations, VMs, network latency, and cloud management operations load. The results in terms of average response time, throughput, and uptime under these controlled conditions provide a starting point for Cloud Providers to estimate how much capacity is needed for their use cases. We strongly recommend that Cloud Providers extensively test each environment prior to production use to ensure that the performance meets the business SLAs.

VMware Cloud Director Virtual Appliance

The current VVD for Cloud Providers recommends the use of a VMware Cloud Director 10.2 virtual appliance. The virtual appliance includes an embedded PostgreSQL database fully managed by VMware Cloud Director with built-in replication for maintaining consistency between cells. A properly configured virtual appliance cluster is also resilient to failures of individual cells.

The recommended virtual appliance deployment includes 1 primary cell, 2 stand-by cells and 0 or more application cells. See the *VMware Cloud Director Installation, Configuration, and Upgrade guide* for more details about the deployment, configuration, and operation of the virtual appliance.

For the purposes of VVD scale and performance benchmarking, the following setup was used: 1 primary cell, 2 standby cells, 2 application cells. The Table 8 Management and Resource Component Sizing gives the details of memory and CPU capacity on each VMware Cloud Director appliance node.

Note that the amount of hardware and software resources deployed is meant to eliminate any resource-related constraints and allow to benchmark performance and stability of the software.

Determining how much hardware and software components are appropriate for a production deployment requires testing. There is no simple formula. However, the following VMware Cloud Director sizing guideline can be used as a starting point:

- Start with a Large VMware Cloud Director virtual appliance with 1 primary and 2 stand-by nodes for environments of scale similar to **Profile B**.
- Benchmark the environment performance under a production-like load. Add application cells as needed to achieve desired throughput and response times.

The embedded PostgreSQL database is automatically configured. However, some post-deployment tuning might be needed for best performance. See the section VMware Cloud Director Appliance PostgreSQL Tuning for the database parameters used in this benchmarking.

Management and Resource Component Sizing

The following table summarizes sizing choices made for various management and resource components.

TABLE 8. MANAGEMENT AND RESOURCE COMPONENT SIZING

MANAGEMENT AND RESOURCE COMPONENT SIZING				
COMPONENT	VERSION	SIZE	RESOURCES	NOTES
Management VMware vCenter Server (vCenter Server with an embedded Platform Services Controller)	7.0 U1	Small	RAM: 19 GB CPU: 4 Storage: 480 GB	1 management vCenter Server

MANAGEMENT AND RESOURCE COMPONENT SIZING				
COMPONENT	VERSION	SIZE	RESOURCES	NOTES
Resource VMware vCenter Server (vCenter Server with an embedded Platform Services Controller)	7.0 U1	Large	RAM: 37 GB CPU: 16 Storage: 1065 GB	3 resource vCenter Server instances. VMware Cloud Foundation deploys medium vCenter Server instances by default. The size is increased to Large
VMware ESXi	7.0 U1		Cisco UCSC-C240-M5SX servers	27 hosts for 3 resource clusters, 13 hosts for 1 management cluster
VMware vSAN (deployed in Management vCenter Server instance)	7.0 U1		21 TB	1 vSAN Cluster for Management vCenter Server instance
VMware vSAN (deployed in Resource vCenter Server instances)	7.0 U1		15 TB (For each Resource vCenter Server instance)	1 vSAN Cluster per Resource vCenter Server instance (3 Resource vCenters)
VMware NSX-T	3.0.2	**Manager** Size: Large	**Manager** RAM: 48 GB CPU: 12 Storage: 300 GB	1 Management NSX-T Manager Cluster (3 NSX-T nodes) 3 Workload NSX-T Manager Clusters (3 NSX-T nodes for each Workload NSX-T Manager Cluster)
		Edge Size: Large	**Edge** RAM: 32 GB CPU: 8 Storage: 200 GB	1 Management NSX-T Edge Cluster (2 NSX-T Edge nodes) 3 Workload NSX-T Edge Clusters (2 NSX-T Edge nodes for each Workload NSX-T Edge Cluster)
VMware Cloud Director	10.2	Virtual appliance 1 Primary (Large) 2 Standby cells (Large) 2 Application Cells (Large)	**Each Primary Cell**: RAM: 32 GB CPU: 24 Storage: 202 GB (increased root partition to 50 GB for logs, increased PostgreSQL database disk to 150 GB) **Each Standby Cell:** RAM: 32 GB CPU: 24 Storage: 202 GB (increased root partition to 50 GB for logs, increased PostgreSQL database disk to 150 GB) **Each Application Cell:** RAM: 8 GB CPU: 8 Storage: 132 GB (increased root partition to 50 GB for logs)	1 TB vSAN File Service NFS storage for the VMware Cloud Director transfer service.

MANAGEMENT AND RESOURCE COMPONENT SIZING				
COMPONENT	VERSION	SIZE	RESOURCES	NOTES
VMware Cloud Director Database	PostgreSQL 10.12.0			Database is embedded with the VMware Cloud Director virtual appliance
VMware Cloud Director Availability	4.1	1 Cloud Replication Management Appliance	CPU: 2 RAM: 4 GB Storage: 10 GB	
		1 Cloud Replicator Appliance	CPU: 4 RAM: 6 GB Storage: 10 GB	
		1 Cloud Tunnel Appliance	CPU: 2 RAM: 2 GB Storage: 10 GB	
		5 On-Premises Appliances	CPU: 4 RAM: 4 GB Storage: 10 GB	
VMware Cloud Director Container Service Extension	3.0.0			
VMware Cloud Director AMQP	RabbitMQ 3.8.9			Cluster with 2 nodes
VMware Cloud Director Metrics Database	Cassandra 3.11.6		CPU: 4 RAM: 8 GB Storage: 120 GB (100GB Dedicated for Keyspace data)	Cluster with 4 nodes
VMware Cloud Foundation SDDC Manager	4.1		RAM: 16 GB CPU: 4 Storage: 816 GB	
VMware vRealize Log Insight	8.1.1	**Master:** Large (1 node) **Worker:** Large (2 nodes)	**Master**: RAM: 32 GB CPU: 16 Storage: 1030 GB **Worker:** RAM: 32 GB CPU: 16 Storage: 1030 GB	Use the vRealize Log Insight sizing calculator: *http://www.vmware.com/go/loginsight/calculator*

MANAGEMENT AND RESOURCE COMPONENT SIZING				
COMPONENT	VERSION	SIZE	RESOURCES	NOTES
VMware vRealize Network Insight	5.3	**Platform:** Large (3 nodes) **Collector:** Large (3 nodes)	**Platform**: RAM: 48 GB CPU: 12 Storage: 1 TB **Collector:** RAM: 8 GB CPU: 16 Storage: 215 GB	
VMware vCloud Usage Meter	4.3	Standard	RAM: 8 GB CPU: 2 Storage: 60 GB	
VMware Identity Manager	3.3.2		RAM: 6 GB CPU: 2 Storage: 60 GB	
VMware vRealize Orchestrator	8.2		RAM: 12 GB CPU: 4 Storage: 200 GB	
VMware vRealize Orchestrator plug-in for VMware Cloud Director	10.0.0.1			
VMware vRealize Operations Manager	8.1.1	**Master**: Size: Large **Master Replica:** Size: Large **Data:** Size: Large	**Master** RAM: 48 GB CPU: 16 Storage: 274 GB **Master Replica** RAM: 48 GB CPU: 16 Storage: 274 GB **Data** RAM: 48 GB CPU: 16 Storage: 274 GB	1 Master, 1 Master Replica, 1 Data node Use the vRealize Operations sizing guidelines: *https://kb.vmware.com/s/article/78495*
VMware vRealize Lifecycle Manager	8.1 Patch 1		RAM: 6 GB CPU: 2 Storage: 148 GB	
Management Pack for VMware NSX-T	8.1			Native Management Pack
Management Pack for VMware vSphere	8.1			Native Management Pack

MANAGEMENT AND RESOURCE COMPONENT SIZING				
COMPONENT	VERSION	SIZE	RESOURCES	NOTES
Management Pack for VMware vRealize Log Insight	8.1			Native Management Pack
Management Pack for VMware Cloud Director	5.4			
Management Pack for VMware vSAN	8.1			Native Management Pack
VMware vRealize Operations Tenant App for VMware Cloud Director	2.5			
VMware Cloud Director App Launchpad	2.0			
VMware Cloud Director Object Storage Extension	2.0			
VMware NSX Advanced Load Balancer	20.1.1			

Appendix A – Test Environment and Benchmarking Methods

Test Environment

The test environment is broadly divided into three main setups:

- Management cluster
- Resource cluster (70% of workloads on vSAN, 30% on iSCSI storage)
- Test driver

Management Cluster

This is where all the management components were deployed.

- 13 physical servers (Cisco UCSC-C240-M5SX)
  Each server is with 192 GB RAM and 28 cores, and with vSAN supported SSDs
- 21 TB vSAN, 14 TB iSCSI

Figure 1. Management Component Deployment

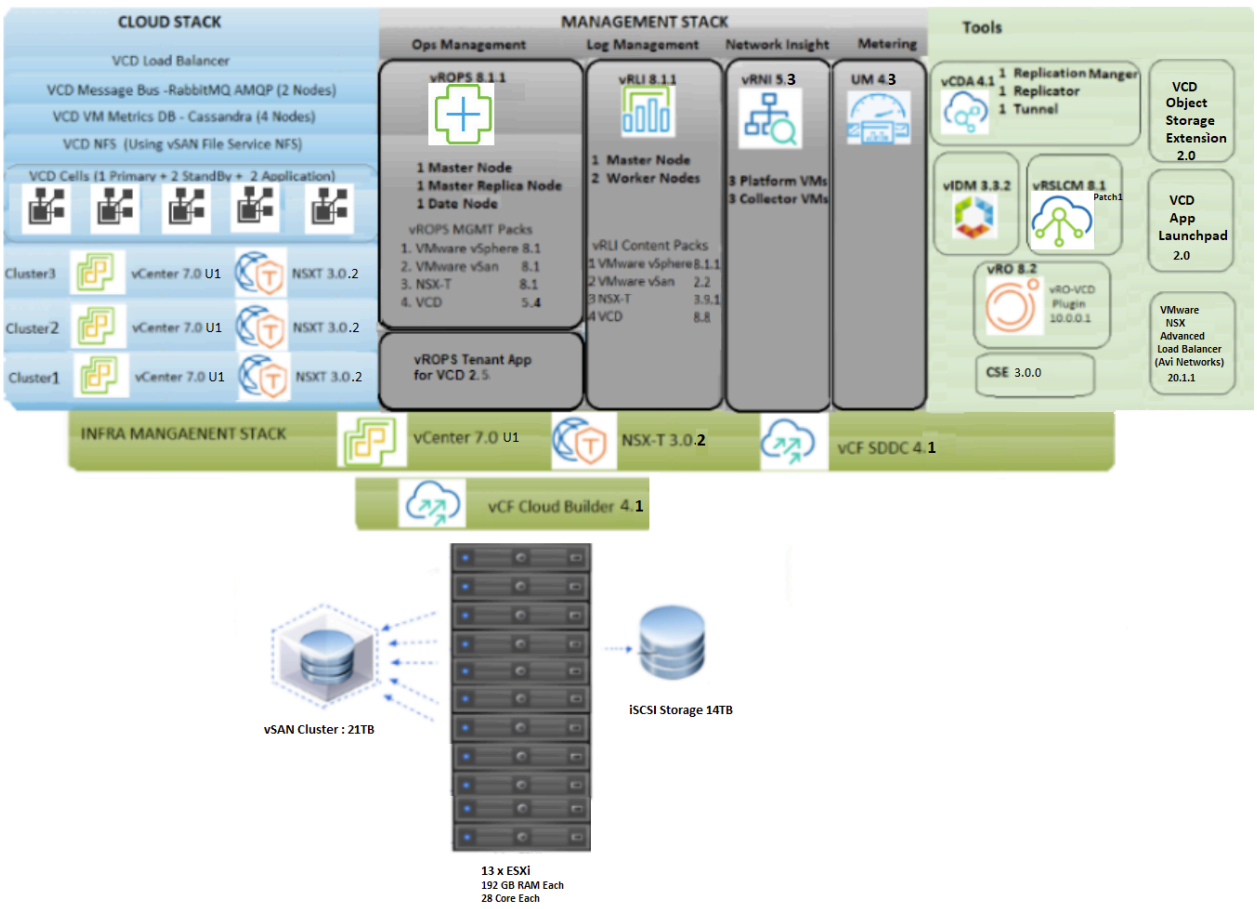
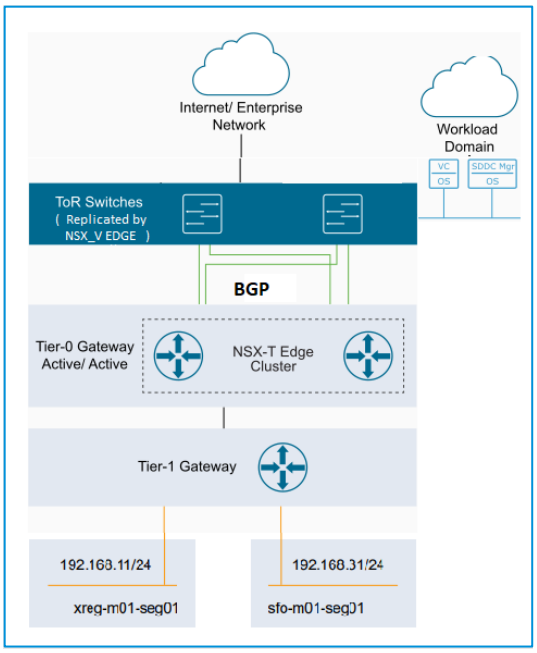


Figure 2. Management Cluster Networking



Resource Cluster

This is where the Tenant Organizations and workload virtual machines were created.

- 27 physical servers (Cisco UCSC-C240-M5SX)
  Each physical server is with 192 GB RAM and 28 cores, each with vSAN supported SSDs
- 15.72 TB vSAN, 10 TB iSCSI (Shared) and 4TB iSCSI for each cluster

Figure 3. Resource Cluster Setup

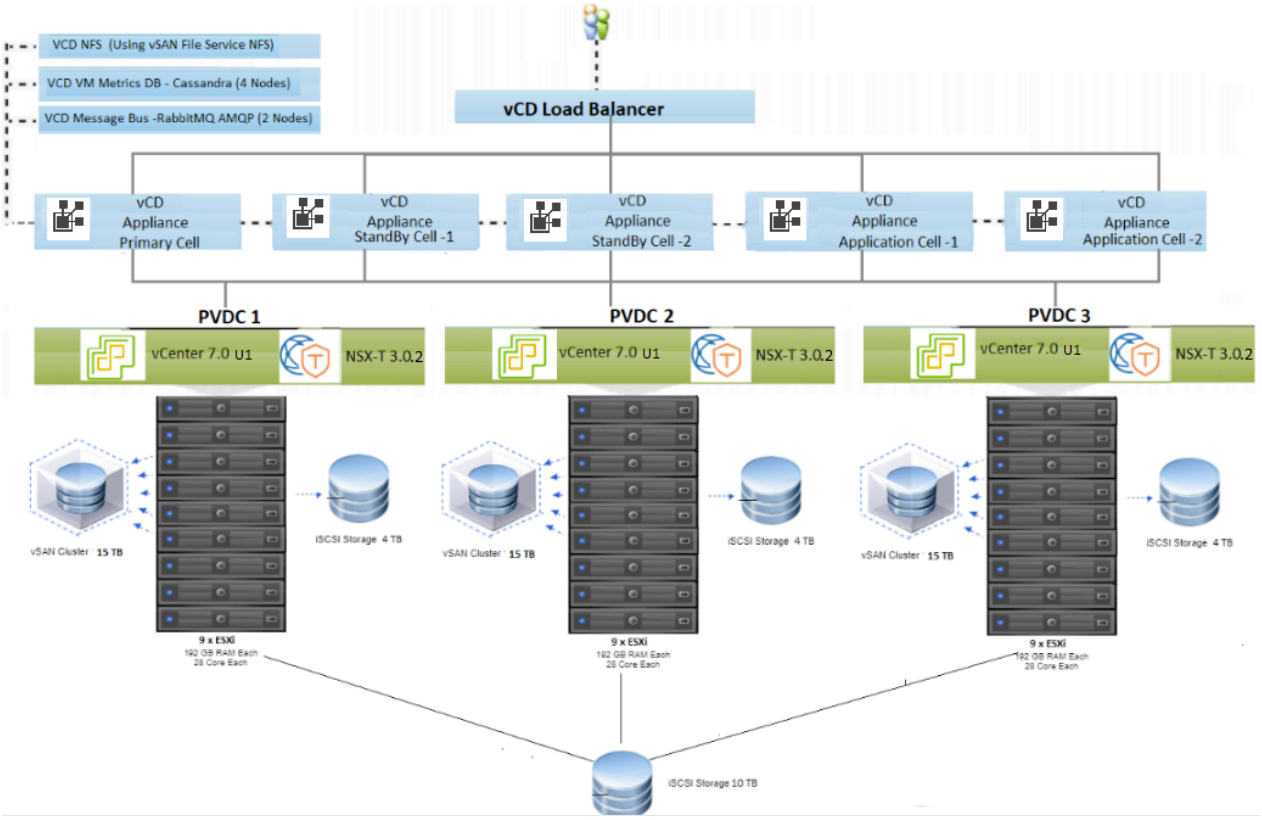
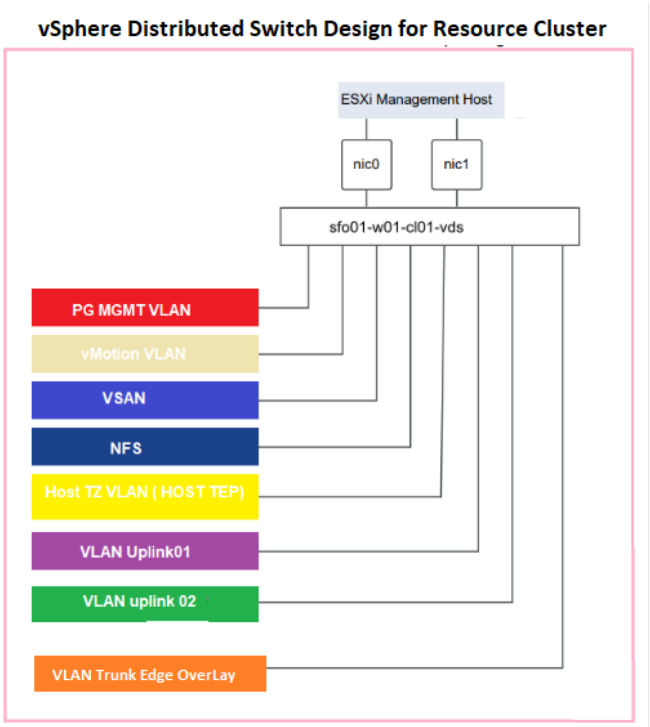


Figure 4. Resource Cluster Networking



VMware Cloud Director Appliance PostgreSQL Tuning

PostgreSQL database parameters were set as follows:

shared_buffers = '7GB';"
effective_cache_size = '21GB';"
work_mem = '8MB' ;"
max_worker_processes = '24';"
maintenance_work_mem = '1GB';"
wal_buffers = '16MB';"
max_wal_size = '2GB';"
min_wal_size = '2GB';"
checkpoint_timeout = '5min';"
checkpoint_completion_target = '0.9';"

See *How to Modify PostgreSQL Configuration* from the VMware Cloud Director documentation.

vCenter Server Sizing

The default size for a workload vCenter Server per VMware Cloud Foundation deployment is medium. To support the current scale, the size is changed from medium to large.

NSX-T Tuning

The http properties value of API rate limit, global-api-concurrency, client-api-concurrency-limit on NSX-T Manager must be increased to:

- 500 for global-api-concurrency-limit
- 450 for client-api-concurrency-limit

- 500 for client-api-rate-limit

To update the http properties values on NSX-T Manager:

1. Connect to NSX-T manager as admin user.
2. Run the following commands on NSX-T Manager:

```
set service http global-api-concurrency-limit 500
set service http client-api-concurrency-limit 450
set service http client-api-rate-limit 500
```

Data Plane Rx Ring Size value for all Workload Domain NSX-T Edge Transport Nodes must be increased to 2048. To Update Data Plane Rx Ring Size:

1. Connect to NSX-T Edge Transport Node as admin user.
2. Run the following commands on NSX-T Edge Transport Node:

```
set dataplane ring-size rx 2048
restart service dataplane
```

vRealize Operations vCloud Director Configuration Properties Tuning

The vCloud thread pool properties value must be decreased to as mentioned below:

- RESOURCE_COLLECT_THREAD_POOL_SIZE=10
- RELATIONSHIP_COLLECT_THREAD_POOL_SIZE=30
- METRICS_COLLECT_THREAD_POOL_SIZE=40

See *How to Modify vCloud Properties Configuration* from Management-Packs-for-vRealize-Operations-Manager documentation. For more info follow the latest KB article at *https://kb.vmware.com/s/article/81977*.

Test Driver

The test driver suite is run from this environment.

- 4 CPU, 8 GB memory, CentOS 7.3

Benchmarking Methods

The testing process is focused primarily on verifying and measuring environment behavior for:

- Scale
  Verify whether the environment meets the Scale Profile B requirement of 10,000 powered-on virtual machines.
- Performance
  Measure the operation latency and throughput when the environment is running at scale (10,000 powered-on virtual machines).
- Uptime
  Verify that the environment can operate at scale with reasonable performance for a long time.

The remainder of this section details the exact methods used for testing and measurement.

Scale Test

Scale was carried out with a mix of manual operations and JMeter test tool-based script operations by using the following steps:

1. Create 400 Tenant Organizations in VMware Cloud Director.
2. Create 15,000 virtual machines and power on 10,000 virtual machines across these 400 tenant organizations. All virtual machines were running CentOS with 2 GB disk, 0.5 GB memory.
3. A sample of VMware Cloud Director operations were carried out to verify that system behaves normally at this scale.

Performance Test

Performance tests were done by performing a well-known distribution of VMware Cloud Director operations with the help of an internal test tool. For the complete operation list, see *List of Operations* (Table 9).

The following were the key steps in performing and measuring the operations:

1. Scaled up the environment as outlined in the previous section.
2. After the environment was at scale, run a continuous stream of operations for 30 minutes with following distribution:
   – 35-40% vApp operations such as instantiate, deploy, edit, clone, and delete.
   – 25% storage-centric operations such as create, attach, detach, and delete disk.
   – 15% networking-related operations, such as create and delete gateway, routed networks, and firewall configurations.
   – 5% create and delete Orgs, users, catalogs, and virtual data centers.
3. Operations were performed by using VMware Cloud Director local users of different roles (such as **vApp Author**, **Org Admin**, and **System Admin**) with 10% admin roles and 90% user operation roles.
4. Given that most of the operations are asynchronous, the test tool monitors the task returned by VMware Cloud Director to get a completion status and performance time details.
5. Steps 2 to 4 were repeated with 32, 64, and 128 concurrent users to ascertain the ability of the system to deal with concurrent operation invocation.
6. Step 5 was repeated for following latency (between VMware Cloud Director and vCenter Server / NSX-T Manager) values (achieved by artificial latency injection with a tool):
   – 0.3 ms (default)
   – 40 ms
   – 150 ms

Uptime Tests

Uptime tests involved running the environment (based on Scale Profile B) for 5 days and running a constant stream of API calls covering a representative set of operations. The purpose of the test is to establish the API call success rate and system uptime.

1. Tests ran continuously for 5 days.
2. API workflows were triggered by 100 concurrent clients, each client would invoke an operation roughly every 20 seconds. 10,000 powered on VMs
3. No artificial latency injection was done.

Results

Total Operations completed over 5 days: **1,807,200**

Average test throughput over 5 days: **251 Ops/min**

Success rate: **99.95%**

List of Operations

For performance benchmarking, API test clients ran a predetermined distribution across different types of VMware Cloud Director operations as described in the following tables.

TABLE 9. VMWARE CLOUD DIRECTOR OPERATIONS

VMWARE CLOUD DIRECTOR OPERATIONS	
COMPONENT	**OPERATIONS**
vApp	• Instantiate a vApp • Deploy (power on) • Edit a vApp • Compose a vApp • Clone a vApp • Power off a vApp • Delete a vApp
Network	• Create a gateway • Create a routed org network • Create an isolated network • Delete an isolated network • Instantiate a vApp • Deploy a vApp • Undeploy a vApp • Delete a gateway • Delete a routed Org network • Delete a vApp
Management	• Create an Org • Create a user • Create an Org VDC • Create a VDC network • Create a catalog • Delete a catalog
NSX Management	• Add DNAT/SNAT • Add Firewall Rule • Delete Firewall Rule • Delete DNAT/SNAT
Datastore	• Create a disk • Instantiate a vApp • Attach a disk to a vApp • Detach a disk from a vApp • Delete a disk • Delete a vApp
OVF	• OVF upload • OVF download

Appendix B – FAQ

How frequently will the Scale and Performance be updated?

- We expect to release an updated Scale and Performance with every major VMware Cloud Director release.

How is this document related to the VMware interoperability matrix?

- The benchmarked stack is a subset of the full interoperability matrix and reflects the exact components we validated and benchmarked in this exercise. The full interoperability includes many more products and versions than what is tested in this exercise

How is the Scale and Performance related to VMware Cloud Foundation?

- VMware Cloud Foundation is being used to deploy vCenter Server, vSAN, and NSX-T. It is not required that partners use VMware Cloud Foundation, but it is a supported and recommended option.

Is NSX-V supported as part of the current VVD?

- Either NSX-T and NSX-V or a mix of both can be deployed in a Cloud Provider environment. Refer to VMware Product Interoperability for supported versions of NSX-V. For our testing environment we deployed NSX-T only.

Is Scale and Performance suitable for greenfield environments or brownfield environments?

- Any environment can be made compliant by simply upgrading all its components to the versions listed in the Scale and Performance Bill of Materials. There is no other qualification.

How can we provide input/recommendations for future versions of this doc?

- Contact the VMware Cloud Director team at *vcd-feedback@vmware.com* or reach out to your VMware account team and pass your feedback through them.

What is the support model for an environment configured according to these guidelines?

- Each component of the Cloud Provider Platform stack is supported according to its support lifecycle. A cloud deployment compliant with the Bill of Materials is in support for at least 12 months after the Scale and Performance release date.

VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 vmware.com Item No: vmw-wp-tech-temp-uslet-word-101-proof 6/20