



VMware vCloud[®] Architecture Toolkit

Architecting a VMware vCloud

Version 2.0.1

October 2011

© 2011 VMware, Inc. All rights reserved. This product is protected by U.S. and international copyright and intellectual property laws. This product is covered by one or more patents listed at <http://www.vmware.com/download/patents.html>.

VMware is a registered trademark or trademark of VMware, Inc. in the United States and/or other jurisdictions. All other marks and names mentioned herein may be trademarks of their respective companies.

VMware, Inc
3401 Hillview Ave
Palo Alto, CA 94304
www.vmware.com

Contents

1. Overview	7
1.1 Audience.....	7
1.2 Scope.....	7
1.3 Document Topics.....	8
2. vCloud Architecture	9
2.1 System Architecture	9
2.2 vCloud Components	10
2.3 vCloud Infrastructure Logical Design	11
3. Management Cluster	13
3.1 Component Sizing	14
3.2 Compute	15
3.3 Network	16
3.4 Storage	16
3.5 vCenter Linked Mode	16
3.6 Cell Load Balancing.....	17
4. Resource Groups	18
4.1 Compute	18
4.2 Network	19
4.3 Storage	22
4.4 vCloud Resource Sizing	24
5. vCloud Resource Design	28
5.1 vCloud Director Constructs	28
5.2 Organizations	30
5.3 Provider Virtual Datacenter	31
5.4 Organization Virtual Datacenters	33
5.5 vCloud Networking	39
5.6 Networking – Public vCloud Example	51
5.7 Networking – Private vCloud Example	54
5.8 vApp.....	55

6.	vCloud Metering	57
6.1	vCenter Chargeback.....	57
6.2	Maximums	59
6.3	Cost Calculation	60
7.	Orchestration and Extension	62
7.1	vCloud API.....	62
7.2	vCloud Messages	63
7.3	vCenter Orchestrator	64
7.4	vCenter Orchestrator Examples	70
8.	Multi-Site Considerations.....	72
9.	Hybrid vCloud Considerations	75
9.1	vCloud Connector Considerations.....	75
10.	References.....	78
	Appendix A: Availability Considerations	79
	vCloud Director Cell Load Balancing.....	82
	Appendix B: Security	85
	Network Access Security	85
	DMZ Considerations	88
	Port Requirements	89

List of Figures

Figure 1. System Architecture	9
Figure 2. vCloud Components.....	11
Figure 3. vCloud Logical Architecture Overview	12
Figure 4. vCloud Management Cluster.....	13
Figure 5. Three-Host Management Cluster	15
Figure 6. vCloud Resource Groups	18
Figure 7. Auto Deploy First Boot	19
Figure 8. Physical, Virtual, and vCloud Abstraction Mapping	28
Figure 9. Extending Virtual Datacenters.....	32
Figure 10. Reservation Pool.....	33
Figure 11. Allocation Pool.....	34
Figure 12. Pay-As-You-Go	34
Figure 13. vApp Placement Algorithm.....	37
Figure 14. External Organization Network (Direct).....	41
Figure 15. External Organization Network (Routed)	42
Figure 16. Internal Organization Network (Isolated).....	42
Figure 17. vApp Network (Direct) → Organization Network (Direct)	43
Figure 18. vApp Network (Direct) → Organization Network (Routed)	44
Figure 19. vApp Network (Direct) → Organization Network (Isolated)	44
Figure 20. vApp Network (Fenced) → Organization Network (Direct).....	45
Figure 21. vApp Network (Fenced) → Organization Network (Routed)	45
Figure 22. vApp Network (Fenced) → Organization Network (Isolated)	45
Figure 23. vApp Network (Routed) → Organization Network (Direct)	46
Figure 24. vApp Network (Routed) → Organization Network (Routed).....	46
Figure 25. vApp Network (Routed) → Organization Network (Isolated).....	47
Figure 26. vApp Network (Isolated).....	47
Figure 27. Organization Network Static Routing Use Case 1	48
Figure 28. Organization Network Static Routing Use Case 2	49
Figure 29. vApp Network Static Routing Use Case.....	50
Figure 30. Example of Public vCloud Networking	53
Figure 31. Example of Private vCloud Networking.....	54
Figure 32. vCenter Chargeback Clustering	58
Figure 33. vCloud Messages.....	63

Figure 34. vCenter Orchestrator Architecture 64

Figure 35. vCenter Orchestrator as a vCloud Director Extension 71

Figure 36. Two Sites with Local VCD Instances Managing Two Local vCenter Servers 72

Figure 37. Remote Console Flow 73

Figure 38. Two Sites with Isolated vCloud Director Instances 74

Figure 39. Hybrid vCloud Example 75

Figure 40. vCloud Connector Architecture 76

Figure 41. Site-to-Site VPN connectivity 87

Figure 42. vCloud Director Port Requirements Illustrated 90

List of Tables

Table 1. Document Topics 8

Table 2. vCloud Components 10

Table 3. Component Requirements for a Management Cluster 14

Table 4. Definition of Resource Pool and Virtual Machine Split 24

Table 5. Memory, CPU, Storage, and Networking 25

Table 6. Example Consolidation Ratios 25

Table 7. vCloud Maximums 26

Table 8. vCloud Director Constructs 29

Table 9. Linked Clone Deployment 35

Table 10. Public vCloud Virtual Datacenter Requirements 38

Table 11. Private vCloud Virtual Datacenter Requirements 39

Table 12. Network Pool Options 40

Table 13. Public vCloud Network Requirements 51

Table 14. Private vCloud Network Requirements 54

Table 15. vCloud Hierarchy Allocation Units 60

Table 16. Reference Documentation 78

Table 17. vCloud Availability Considerations 79

Table 18. Load Balancer Considerations 83

Table 19. Network Access Security Use Cases 85

Table 20. vCloud Director Port Requirements 89

Table 21. vCenter Orchestrator Port Requirements 91

1. Overview

Architecting a VMware vCloud provides guidance to architect an Infrastructure-as-a-Service (IaaS) cloud based on VMware vCloud® Director™ (VCD). Simplifying the delivery of resources to end users requires the coordination of various infrastructure platforms. Both service providers and enterprises can use the guidelines in this document, with some variations depending on point of view.

This document, combined with the private or public service definition, can help you navigate through the applicable design considerations for architecting a vCloud solution. The documents, *Architecting a VMware vCloud*, *Operating a VMware vCloud*, and *Consuming a VMware vCloud* are designed to work together throughout the lifecycle of a VMware vCloud computing implementation with VMware technologies.

- *Architecting a VMware vCloud* provides best practices, design considerations, and design patterns for constructing a vCloud environment from its constituent components.
- *Operating a VMware vCloud* includes best practices and considerations for operating and maintaining a vCloud environment. It covers the people, process, and technology involved in running a vCloud environment.
- *Consuming a VMware vCloud* covers the various considerations for the consumer when choosing to leverage vCloud computing resources.

This document is not a substitute for VMware product documentation, nor does it provide detailed implementation procedures for installing a vCloud.

1.1 Audience

This document is intended for, but not limited to, those involved in planning, designing, and implementing VMware vCloud solutions. The target audience is VMware Certified Professionals (VCP) familiar with VMware products, particularly VMware vSphere®, VMware vCloud Director, VMware vShield Manager™ (VSM), and VMware vCenter Chargeback™. It is assumed that the reader has knowledge of and familiarity with vSphere concepts.

1.2 Scope

This document includes design considerations and design patterns for building a vCloud.

1.3 Document Topics

This document is structured into the sections shown in the following table.

Table 1. Document Topics

Section	Description
Section 2, vCloud Architecture	Introduces the core concepts of the vCloud solution stack.
Section 3, Management Cluster	Describes the components required to stand up a vCloud.
Section 4, Resource Groups	Provides guidance for configuring resources reserved for end-user workloads.
Section 5, vCloud Resource Design	Offers best practices for partitioning and delivering vCloud resources relative to customer requirements.
Section 6, vCloud Metering	Covers how to meter and charge for resources with vCenter Chargeback.
Section 7, Orchestration and Extension	Walks through extending VCD automation via orchestration.
Section 8, Multi-Site Considerations	Covers multi-site considerations.
Section 9, Hybrid vCloud Considerations	Discusses extending VCD into the hybrid vCloud model.
Appendix A: Availability	Covers design considerations for availability.
Appendix B: Security	Covers design considerations for security.

2. vCloud Architecture

VMware vCloud is a common set of cloud computing services for enterprises and service providers. Cloud computing is a style of computing that leverages the efficient pooling of on-demand, self-managed virtual infrastructure to provide resources consumable as a service.

This document focuses strictly on the IaaS layer, using vCloud Director to extend the capabilities of the vSphere virtualization platform.

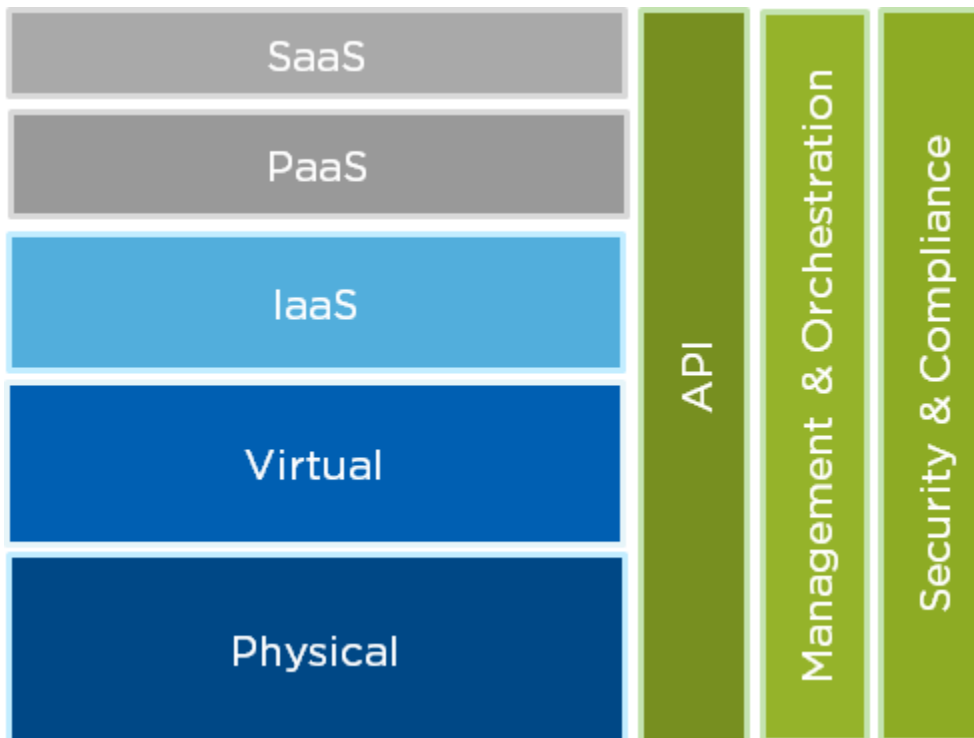
2.1 System Architecture

VMware vCloud provides an open and modular architecture that offers choice and flexibility for running applications in public and private vCloud instances. vCloud Director implements the vCloud API, which provides compatibility and interoperability with other vCloud instances.

A cloud architecture is a strategic design that involves devising a conceptual framework that supports primary business requirements, deciding on the discrete functions of the system, organizing elements into distinct components, and defining boundaries and connections between interdependent components. Focus on clearly defining architecture goals and analyzing elements in a systematic and sufficient manner to facilitate design decisions that cut through the complexity found in today's technology.

Figure 1 shows the fundamental structure and components of the cloud computing stack.

Figure 1. System Architecture



2.2 vCloud Components

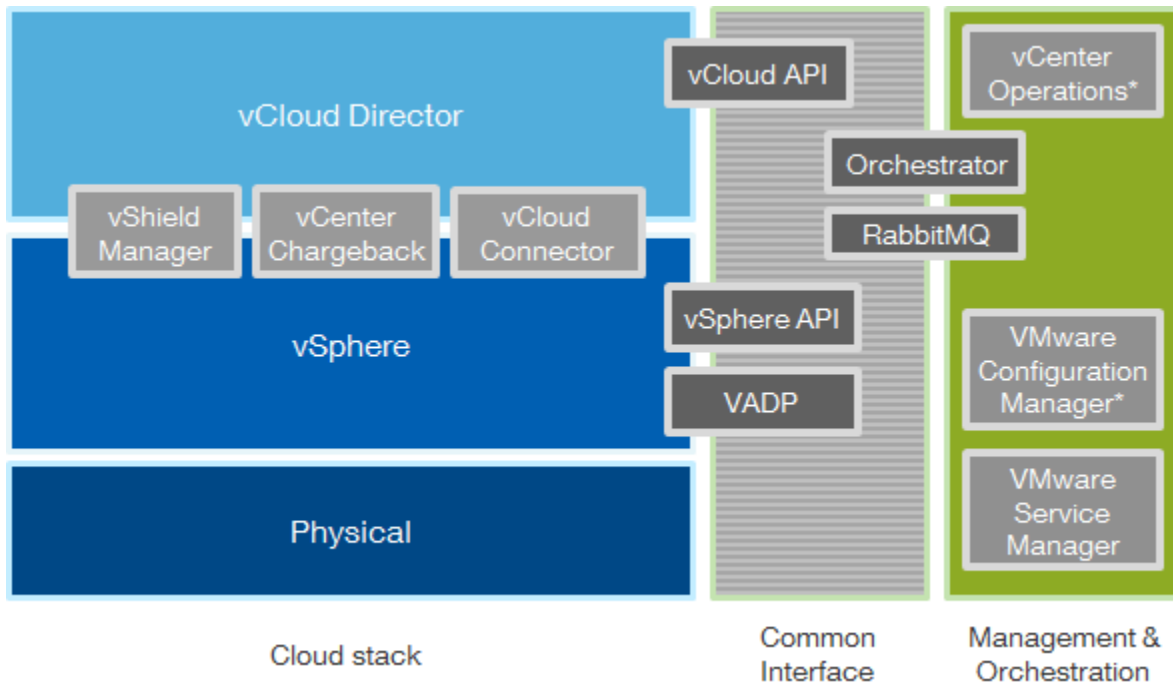
Table 2 describes the components that comprise a VMware vCloud.

Table 2. vCloud Components

vCloud Component	Description
VMware vCloud Director vCloud API	Layer of software that abstracts virtual resources and exposes vCloud components to consumers. Includes: <ul style="list-style-type: none"> • vCloud Director Server (also referred to as a <i>cell</i>). • vCloud Director Database. • VMware vCloud API, used to manage vCloud objects programmatically.
VMware vSphere	Virtualization platform providing abstraction of physical infrastructure layer for vCloud. Includes: <ul style="list-style-type: none"> • VMware ESXi™ hosts. • VMware vCenter™ Server. • vCenter Server database.
VMware vShield	Provides perimeter network security for virtual datacenters using VMware vShield Edge™. Includes: <ul style="list-style-type: none"> • vShield Manager. • vShield Edge.
VMware vCenter Chargeback	<ul style="list-style-type: none"> • Provides dynamic resource metering, cost modeling, and report generation. Includes: • vCenter Chargeback Server. • vCenter Chargeback database. • Chargeback data collector. • vCloud data collector. • VSM data collector.
VMware vCenter Orchestrator™	VMware vCenter Orchestrator enables the automation of provisioning and operational tasks across VMware and third-party applications through an open and flexible plug-in architecture.
VMware vCloud Connector	vSphere Client plug-in that allows users to connect to vSphere-based or vCloud Director-based clouds and manage them through a single interface.

Figure 2 shows the relationship of vCloud components.

Figure 2. vCloud Components



* Denotes current management products without vCloud Director integration.

2.3 vCloud Infrastructure Logical Design

VMware recommends decoupling resources allocated for management functions from resources dedicated to user-requested workloads. Partition resources into:

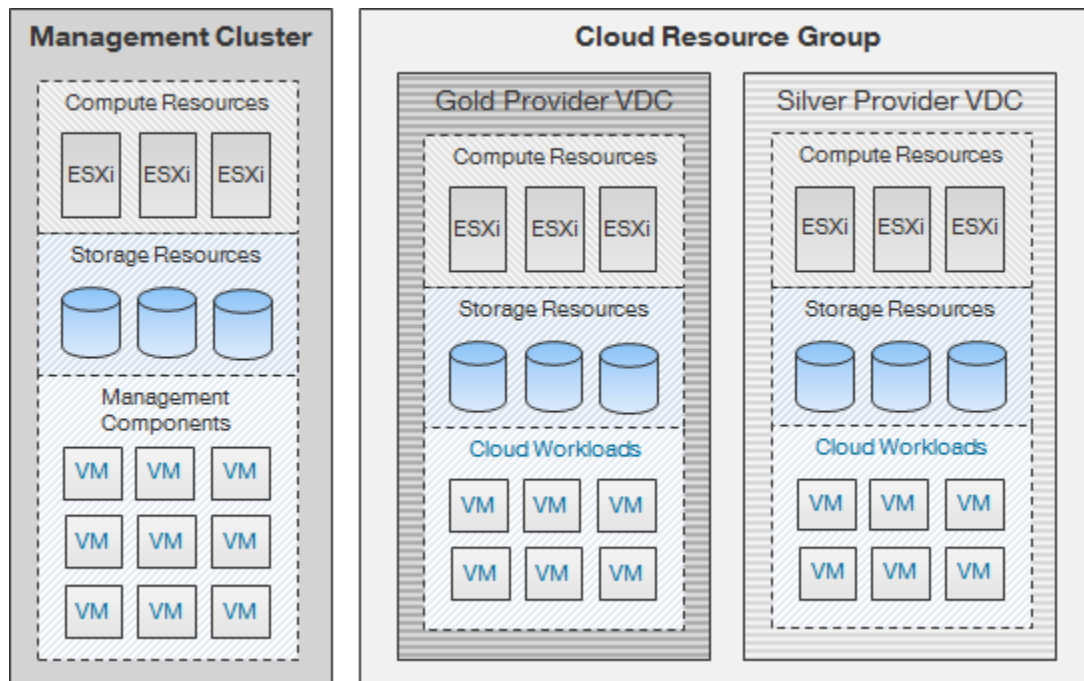
- A *management cluster* containing core components and services needed to run the vCloud. This includes core vCloud components such as VMware vCenter Server, vCloud Director, and vCenter Chargeback Server.
- *Resource groups* represent dedicated resources for end-user consumption. Each resource group consists of VMware ESXi hosts managed by a vCenter Server, and is under the control of vCloud Director. vCloud Director can manage the resources of multiple resource groups.

Reasons for separating virtual resources along these lines are:

- To facilitate quicker troubleshooting and problem resolution. Management components are strictly contained in a relatively small and manageable cluster. Running management components on a large cluster with mixed environments can be time-consuming and make it difficult to track down and manage such workloads.
- Separation of management components from the resources they are managing. This helps avoid inadvertent changes to vCloud Director-created entities through the vSphere Client.

- Resources allocated for vCloud use have minimal overhead reserved. For example, vCloud resource groups would not host vCenter Server virtual machines. The exception is for vShield Edge devices that run on resource group hosts to facilitate traffic isolation.
- Enables consistent and transparent management of infrastructure resources, which is critical for scaling vCloud environments. Increases upgrade flexibility because management cluster upgrades are not tied to resource group upgrades.
- Prevent security attacks or intensive provisioning activity on the resource groups from affecting management component availability.

Figure 3. vCloud Logical Architecture Overview

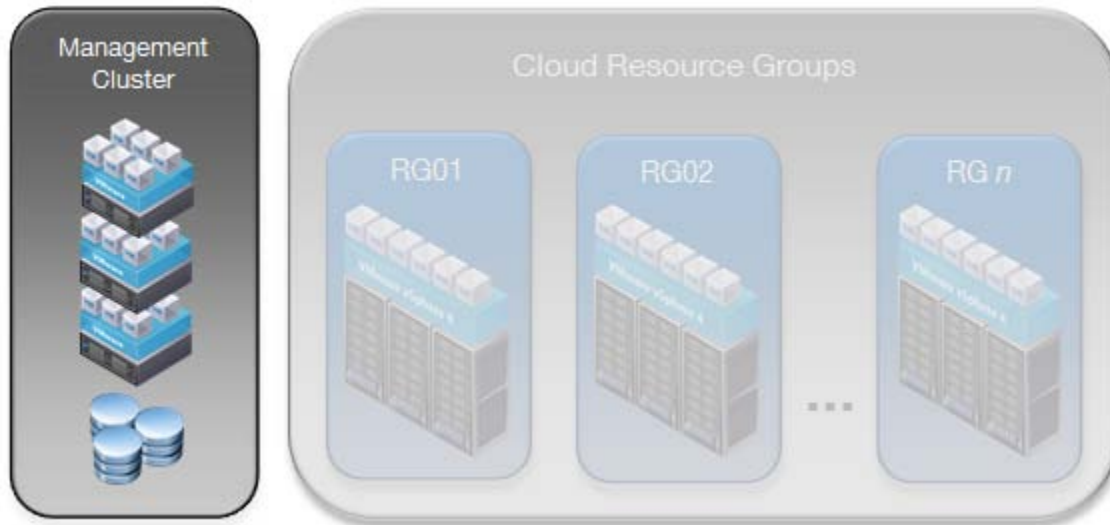


Achieving economies of scale means scaling vCloud resources in a consistent and predictable manner. Follow applicable best practices when deploying the underlying vSphere infrastructure and other vCloud components.

3. Management Cluster

The management cluster hosts all the necessary vCloud infrastructure components. Separating infrastructure components from resources used for end-user workloads improves manageability of the vCloud infrastructure.

Figure 4. vCloud Management Cluster



Management components include:

- vCenter Server or vCenter Server appliance.
- vCenter Server database.
- vCloud Director cells.
- vCloud Director database.
- vShield Manager (one per resource group vCenter Server.)
- vCenter Chargeback Server.
- vCenter Chargeback database.
- VMware vCenter™ Update Manager.
- vCenter Orchestrator.

Note vShield Edge appliances are deployed automatically by vCloud Director through vShield Manager as needed and reside in the resource groups, not in the management cluster.

Optional management cluster components include:

- VMware vCenter Server Heartbeat™.
- VMware Service Manager.
- VMware vCloud Connector.
- VMware vFabric™ RabbitMQ™.
- VMware vFabric™ Hyperic® HQ.

- Databases for optional components.
- VMware vSphere Management Assistant.
- VMware vCenter Operations (currently not VCD-aware).
- VMware vCenter Configuration Manager (currently not VCD-aware).

Optional components are not required by the service definition, but are highly recommended to increase the operational efficiency of the solution.

The management cluster may also include virtual machines or have access to servers that provide infrastructure services such as directory (LDAP), timekeeping (NTP), networking (DNS, DHCP), logging (syslog), and security. Review Section 4.4.1, Public vCloud Sizing Example, for detailed sizing considerations.

Component databases, if running on the same platform, can be placed on the same database server if sized properly. For example, the databases used by vCloud Director, vCenter Server, and vCenter Chargeback can run on the same database server.

Both the management cluster and resource groups reside in the same physical site to provide a consistent level of service. This minimizes latency issues, which could arise in a multi-site environment if workloads move between sites over a slower or less reliable network. See Section 8, Multi-Site Considerations for considerations on connecting clouds residing in different sites.

3.1 Component Sizing

Table 3 lists the requirements for each of the components that run in the management cluster. For the number of virtual machines and organizations listed in the private or public service definitions, you do not have to worry about scaling too far beyond the provided values.

Table 3. Component Requirements for a Management Cluster

Item	vCPU	Memory	Storage	Networking
vCenter Server	2	4GB	20GB	1 GigE
Database server	4	16GB	100GB	1 GigE
vCloud Director cell 1	2	4GB	30GB	1 GigE
vCloud Director cell 2	2	4GB	30GB	1 GigE
vCenter Chargeback	2	4GB	30GB	1 GigE
vShield Manager	1	4GB	8GB	100Mb
TOTAL	13	36GB	218GB*	4 GigE*

* Numbers rounded up or down do not affect overall sizing.

The database server hosts databases for vCenter Server, vCloud Director, and vCenter Chargeback Server. Use different users and instances for each database per VMware best practices. VMware vCloud Director 1.5 supports both Oracle and SQL Server databases.

In addition to the storage requirements in Table 3, a shared storage volume must be configured and made accessible to all cells in a vCloud Director server group to facilitate file transfers in a multicell environment. The size needed for this volume varies depending on the expected number of concurrent uploads. Once an upload completes, the vApp data moves to the designated organization virtual datacenter and the data no longer resides on the NFS volume. The recommended starting size for the NFS transfer volume is 250GB. Transferred images can be large, so monitor this volume and increase the size if necessary.

For additional installation prerequisites, see the *vCloud Director Installation and Configuration Guide* (http://www.vmware.com/support/pubs/vcd_pubs.html).

3.2 Compute

The compute layer encompasses the CPU, memory, and hypervisor technology components. Follow vSphere best practices where possible when configuring and sizing compute resources.

Figure 5. Three-Host Management Cluster



A three-host cluster is used to support vCloud management components. Given advances in technology, three hosts are sufficient for typical vCloud environments. Add additional hosts if the management cluster becomes resource constrained. Enable vSphere High Availability (HA) and Distributed Resource Scheduling (DRS) on the management cluster to provide availability for all management components. For vSphere HA, use percentage-based admission control policy in an “N+1” fashion instead of defining the amount of host failures a cluster can tolerate or specifying failover hosts. This allows management workloads to run evenly across the hosts in the cluster without the need to dedicate a host strictly for host failure situations. For higher availability, you can add an additional host for an N+2 cluster, although this is not a requirement of the vCloud private or public service definitions.

vCenter Server plays an integral role in end-user self-service provisioning by handling all virtual machine deployment requests from vCloud Director. VMware recommends increasing the availability of vCenter Server through solutions such as VMware vCenter Server Heartbeat.

vSphere Fault Tolerance can be used for continuous virtual machine protection only if all FT requirements are met. vCenter Site Recovery Manager™ (SRM) can be used to protect most components of the management cluster, but at this time, SRM cannot be used to protect vCloud Director cells because a secondary (DR) site is out of scope. Changes to IP addresses and schemas in recovered vCloud Director cells can result in complications.

vCloud Director 1.5 supports vSphere 4.0 U2 and later. Deploy vSphere 5.0 if possible to take advantage of the new features introduced in vCloud Director 1.5. When deciding on vSphere licensing, keep in mind that some functionality in vCloud Director requires specific features tied to particular vSphere editions. For example, automated deployment of vCloud networks requires the use of a distributed switch, a feature that is available in VMware vSphere Enterprise Plus edition.

3.3 Network

The network configuration for the management cluster includes, but is not limited to, the following best practices:

- Logical separation of network traffic for security and load considerations by type (management, virtual machine, VMware vSphere® vMotion®, FT, IP storage).
- Network component and path redundancy.
- At least 10 GigE or GigE network speeds, if available.
- Use vSphere Distributed Switches (VDS) where possible for network management simplification. The architecture calls for the use of VDS in the resource group vCenter Servers, so it is a best practice to standardize on the VDS across all clusters, including the management cluster.

3.4 Storage

Use vSphere storage best practices where applicable for the management cluster. Examples include:

- Redundancy at the host (connector), switch, and storage array levels.
- All hosts in a cluster have access to the same datastores.

3.5 vCenter Linked Mode

vCenter linked mode provides a *single pane-of-glass* to allow a common administrative state to manage multiple vCenter instances. With linked mode configured, users can view and manage the inventories of all participating vCenter Server systems. Tasks invoked on a linked mode object are executed by the vCenter Server that manages the corresponding resource. Using linked mode in the vCloud Director context is useful because you can view all vCenter Servers that manage vCloud resources.

vCloud Director maximums for powered on virtual machines and registered virtual machines are substantially less than the vCenter linked mode maximums. Therefore, the number of linked mode objects in a vCloud environment will not approach linked mode maximums unless multiple vCloud instances are involved.

Additional considerations:

- The vCenter Server appliance does not support linked mode.
- A vCenter instance can only link with other vCenter instances that have the same version. This has upgrade implications when upgrading all vCenter Servers in a vCloud instance.
- Upgrading a linked vCenter breaks the link and the instance becomes independent.

3.6 Cell Load Balancing

vCloud Director cells are stateless front-end processors for the vCloud. Each cell has a variety of purposes and self-manages various functions among cells while connecting to a central database. The cell manages connectivity to the vCloud and provides API and UI endpoints, or clients.

To improve availability and scale, implement a vCloud Director server group comprised of multiple vCloud Director cells. A multicell configuration requires load balancing or content switching of the front-end portal. Load balancers present a consistent address for services regardless of the underlying, responding node. They can spread session load across cells, monitor cell health, and add/remove cells from the active service pool. The cell architecture is not a true cluster since there is no failover from one cell to another.

Any load balancer that supports SSL session persistence with network connectivity to the “public” facing Internet or internal service network can perform load balancing of vCloud Director cells. Refer to general best practices regarding performance, security, manageability, and so forth when deciding to share or dedicate load balancing resources.

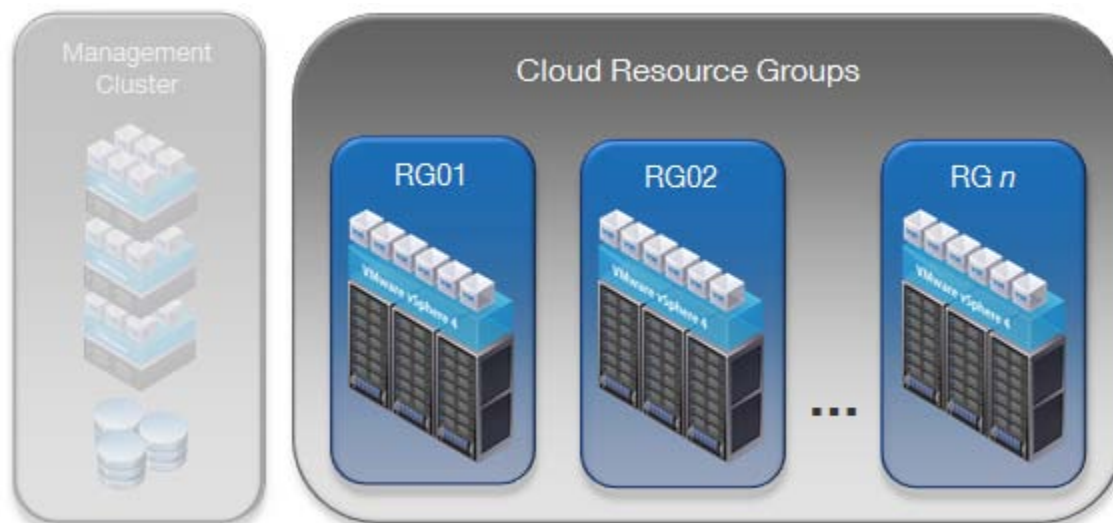
Note SSL offloading will not work with console proxy connections (VMRC).

See Appendix A: Availability for additional load balancing considerations.

4. Resource Groups

A *resource group* is a set of resources dedicated to end-user workloads and managed by a single vCenter Server. vCloud Director manages the resources of all attached resource group vCenter Servers. All provisioning tasks are initiated through vCloud Director and are passed down to the appropriate vCenter Server instance.

Figure 6. vCloud Resource Groups



Provisioning resources in standardized groupings promotes a consistent approach for scaling vCloud environments. At a minimum, place all vCloud resource workloads in a separate cluster if you are using a single vCenter Server to manage both management components and vCloud resources.

Avoid making changes to resource group objects through the vSphere Client. Changing the state of vCloud Director-created objects may cause side effects and unpredictability because these objects are owned by vCloud Director.

4.1 Compute

Configure resource group ESXi hosts per vSphere best practices. Enable vSphere HA appropriately to protect against host and virtual machine failures.

The shift to Fault Domain Manager-based HA (FDM) in vSphere 5 is transparent to vCloud Director. The total number of hosts in an HA/DRS cluster remains 32, so cluster sizing guidance for vCloud environments does not change. FDM requires a single master host, as opposed to legacy HA's five primary nodes. If the master host fails, the remaining slave hosts participate in an election to choose a new master.

Do not exceed eight hosts in a cluster if using fast provisioning (linked clones) and VMFS datastores.

Provider virtual datacenters represent a service offering. When building clusters, group similar servers together (number of hosts, number of cores, amount of memory, CPU type) to support differentiation of compute resources by capacity or performance.

4.1.1 Stateless ESXi

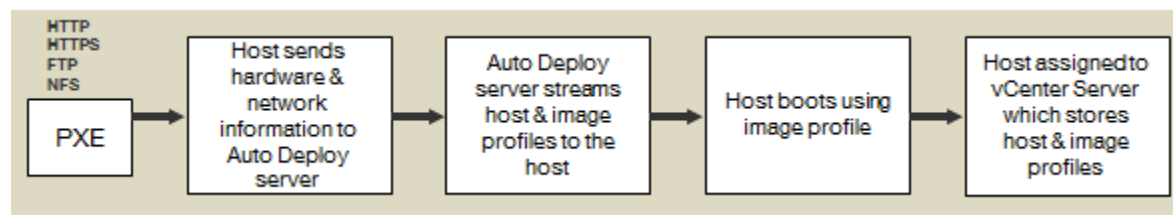
Stateless ESXi refers to running ESXi software on a host entirely in memory, with no local persistence data. Centralizing management of host state enables consistent configuration over large sets of similar hosts, as well as rapid provisioning of ESXi hosts. These aspects help improve operational efficiency in large-scale vCloud environments.

Stateless ESXi requires vSphere Auto Deploy, a deployment server that applies the image profile and host profile to the PXE-booted hosts. Install Auto Deploy on a standalone host or on the vCenter Server. The vCenter Server virtual appliance has Auto Deploy installed by default. Install vSphere PowerCLI in a location reachable by both vCenter and Auto Deploy. The host profile is essential to the stateless environment, as every reboot of a server clears the host of any local configuration data.

Configure all stateless ESXi hosts for DHCP. The DHCP server requires configuration changes to direct the ESXi host to a TFTP server. This can be a separate DHCP server or existing organization's DHCP server. The vCenter Server virtual appliance includes DHCP and TFTP services.

Identify an image profile to use for vCloud hosts. This can be a profile stored in a public depot or a zipped file stored locally. If using host profiles, save a copy of the host profile to a location accessible by Auto Deploy and add rules to the rules engine using Image Builder PowerCLI.

Figure 7. Auto Deploy First Boot



vCloud Director can manage stateful or stateless hosts. If you choose the stateless option, add the vCloud Director vSphere Installation Bundle (VIB) (which contains the agent) to the image profile. (Currently, this is a manual process as there is no API call to register VIBs in an image profile.) The vCloud Director VIB is loaded automatically when the host boots up. For preparation and un-preparation of stateless hosts, vCloud Director configures the agent using a host profile with an associated answer file.

If the host is rebooted, the appropriate image profile is reloaded when the host starts back up. vCloud Director detects the state change, and the configuration is re-pushed to the host.

If using stateless mode, avoid creating designs that require host-specific configuration. When converting a prepared stateful host to stateless, unprepare hosts prior to the conversion.

4.2 Network

For the vCloud resource groups, configure networking with vSphere best practices in mind. In addition, increase the number of vSphere Distributed Switch ports per host from the default value to the maximum of 4096. This improves the scale at which vCloud Director can dynamically create port groups for vCloud networks. See the *vSphere Administrator Guide* for more information on increasing this value.

Increase the maximum transmission unit (MTU) size to 1600 for all devices residing in the transport network for vCloud Director network isolation. This includes physical network devices as well as the vSphere Distributed Switches. Failure to increase the MTU size causes packet fragmentation, negatively affecting network throughput performance of end-user workloads

vCloud networking considerations are covered in Section 5, vCloud Resource Design.

4.2.1 I/O Controls

vCloud Director offers controls to guard against the misuse of resources by consumers. These include:

- Quotas for running and stored virtual machines determine how many virtual machines each user in the organization can store and power on in the organization's virtual datacenters. The quotas specified act as the default for all new users added to the organization.
- Limits for resource-intensive operations prevent these operations from affecting all users in an organization and provide defense against denial-of-service attacks.
- Simultaneous VMware Remote Console (VMRC) connections limits the number of simultaneous connections for performance or security reasons.

4.2.2 IPv6

Internet Protocol version 6 (IPv6) is the latest version of IP addressing, designed to succeed IPv4 as the standard protocol for the Internet. One of the key drivers for transitioning to IPv6 is that it supports a much larger address space of 2^{64} , as opposed to the 2^{32} addresses for IPv4.

The vCloud Director components required to support IPv6 are:

- Static IP pools.
- DHCP Server.
- Static IP assignments.
- NAT rules.
- Firewall rules.

vSphere infrastructure components supporting IPv6 include:

- vCenter Server.
- ESXi.
- vSwitches (standard and distributed).
- VMkernel.
- VMware vSphere® vMotion®.
- Virtual machines (guest customization available for Windows and Linux).

vSphere virtual machines support IPv6 addressing and can be configured with:

- Static IPv6 address.
- Autoconfigure, using a prefix announcement from a router.
- DHCP, from a DHCP6 server.
- Local net addresses, for internal communication.

Virtual machines managed by vCloud Director using IPv6 can only communicate to endpoints that are not behind vShield Edge devices. Currently, vShield Edge does not support IPv6. Virtual machines that communicate on the same directly attached vApp or organization network can use IPv6. To communicate with the outside world using IPv6, connect the organization's virtual machines to a direct external organization network.

Run virtual machines in dual stack IPv4 and IPv6. This is necessary because many destinations do not currently support IPv6.

If the underlying physical infrastructure does not support IPv6, another option is to establish a 6to4 tunnel using a router to provide connectivity into an IPv6 vCloud. Terminate the tunnel on a relay router that has a pure IPv6 interface as well as an IPv4 interface to move traffic between the two environments.

vCloud Director does not support IPv6 addressing for the cell network interfaces.

4.2.3 vShield Edge

VMware vShield Edge is a virtual firewall router that provides the perimeter security needed to support multi-tenancy. vShield Edge devices deploy automatically when routed or isolated organization or vApp networks are created from vCloud Director. For vApp networks, vShield Edge devices dynamically deploy and undeploy based on the power state of the vApp.

The license for vShield Edge included with vCloud Director (*vShield for vCloud Director*) does not include features such as VPN and load balancing capabilities, which are part of the fully licensed vShield Edge.

4.2.4 vShield App

VMware vShield App is a hypervisor-based, vNIC-level application firewall that controls and monitors all flows between virtual machines in a virtual datacenter. Firewall policies can be applied to vCenter containers or security groups, which are custom containers created through the vShield Manager user interface. Container policies enable the creation of mixed trust zone clusters without requiring an external physical firewall. vShield App also supports classic five tuple firewall rules.

Note Currently, vCloud Director does not have integration with vShield App. Using vShield App in conjunction with vCloud Director is a supported configuration, but requires careful design of the vCloud infrastructure.

4.2.5 vShield Endpoint

VMware vShield Endpoint offloads antivirus functions to a hardened security virtual machine delivered by partners such as Trend Micro. Endpoint uses EPSec APIs to peer into the file system to scan and remediate viruses. This removes the need for agents in the guest and prevents antivirus storms from consuming precious CPU cycles during scanning or AV update activities. Offloading antivirus provides enhanced security, as often the first task of malware is to disable AV agents. vShield Endpoint's efficient AV architecture provides antivirus as a service for large-scale vCloud environments.

Currently vCloud Director does not have integration with vShield Endpoint.

4.3 Storage

Designing storage resources for vCloud differs from traditional vSphere approach. Self-service provisioning shifts the responsibility of virtual machine storage placement from the provider to the consumer. Though this restricts the ability to optimize storage for applications, platform improvements have emerged that assist in balancing workloads across storage resources. The goal shifts to provisioning flexible pools of storage resources and maintaining consistent performance for end users.

VMware recommends the following:

- Perform a current state analysis for storage usage and trends.
- Define the range of storage SLAs needed and appropriate pricing models.
- Create multiple tiers of storage based on SLAs, workloads, and cost.
- Map tiers to provider virtual datacenters in vCloud Director.
- Storage design provides optimal availability.
- Physical storage is modular and scalable.
- Monitor storage usage and trends using capacity analysis tools.
- Use storage performance tools to tune vApp storage workloads.

Currently, vCloud Director does not support tiering storage within a virtual datacenter, vApp, or virtual machine. If multiple types of storage resources are available, supply tiered pools of storage grouped by shared characteristics to provide differentiated service offerings. Configure shared storage in the resource groups per vSphere best practices.

Sizing considerations include:

- Datastore storage expectations
 - What is the optimal size for datastores based on the physical storage and vCloud workload expectations?
 - What is the average vApp size x number of vApps x spare capacity? For example: Average virtual machine size * # virtual machines * (1+ % headroom)
 - What is the average virtual machine disk size?
 - On average, how many virtual machines are in a vApp?
 - What is the expected number of virtual machines?
 - How much reserved spare capacity is needed for growth?
 - Will expected workloads be transient or static?
 - Is Fast Provisioning utilized?
- Datastore performance characteristics
 - Will expected workloads be disk intensive?
 - What are the performance characteristics of the associated cluster?

Note vCloud Director 1.5 does not support Raw Device Mappings (RDM).

4.3.1 Storage vMotion

Storage vMotion enables live migration of virtual machine disk files between and across shared storage locations. Relocating vApp disks is possible using the vCloud API or vSphere Client if the following conditions apply:

- The target datastore is part of the same organization virtual datacenter as the vApp.
- All virtual disks for an individual virtual machine are migrated to the same datastore.
- Use the vCloud API to initiate storage vMotion for linked clones to preserve the linked clone state.

Note It is not recommended to invoke Storage vMotion migration of linked clones using the vSphere Client since this may cause undesirable effects such as the inflation of delta disks. If a storage vMotion of datastore *and* host is attempted, the operation may fail.

4.3.2 Storage I/O Control

Storage I/O Control (SIOC) manages storage resources across hosts through storage device latency monitoring and disk shares that are enforced at the datastore level. Preventing imbalances of storage resource allocation during times of contention protects virtual machine performance in highly consolidated, virtualized environments.

Enabling SIOC on all datastores in a cluster ensures lower worst-case device latency by maintaining a balance between workload isolation/prioritization and storage I/O throughput. For more information, refer to the *Storage I/O Control Technical Overview and Considerations for Deployment* (<http://www.vmware.com/files/pdf/techpaper/VMW-vSphere41-SIOC.pdf>).

SIOC does not support raw device mappings (RDM) or datastores with multiple extents. If you are using datastores backed by arrays with automated storage tiering, validate compatibility with SIOC.

4.3.3 vSphere Storage APIs – Array Integration

vSphere Storage APIs – Array Integration (formally VAAI) is a set of protocol interfaces between ESXi and storage arrays. These ESXi extensions enable storage-based hardware acceleration by allowing vSphere to pass storage primitives to supported arrays. VAAI improves storage task execution times, network traffic utilization, and CPU host utilization during heavy storage operations. In vSphere 5, new thin provisioning primitives enhance vSphere management of thin-provisioned block-based storage.

In vSphere 5.0, array integration extensions are implemented as T10 SCSI-based commands. Devices that support the T10 SCSI standard do not require a VAAI plug-in to offload storage functions such as full copy, block zeroing, and locking. See the *VMware Compatibility Guide* (<http://www.vmware.com/resources/compatibility/search.php>) for more details.

vCloud Director 1.5 supports block-based primitives (FC, iSCSI), but not the new NAS primitives introduced in vSphere 5.

4.3.4 Storage DRS

Storage DRS provides initial placement and on-going balancing recommendations for datastores in a Storage DRS-enabled *datastore cluster*. A datastore cluster represents an aggregation of datastore resources, analogous to clusters and hosts.

Currently datastore clusters cannot serve as a storage source to provider virtual datacenters. *Do not* enable Storage DRS for datastore clusters used with vCloud Director, as this is not a supported configuration.

4.4 vCloud Resource Sizing

Resource sizing for a vCloud depends on the corresponding service definition. A private vCloud service definition may not specifically call out a required number of workloads to support. In that case, use the initial sizing for a public vCloud as guidance.

For a public vCloud, initial sizing for the vCloud consumer resources can be difficult to predict due to lack of data points on expected consumer uptake. The provider is also not aware of existing usage statistics for vCloud workloads.

The sizing examples in the next section come from the *Service Definition for a Public VMware vCloud* and can assist with initial sizing of the vCloud environment.

Note VMware recommends that you engage your local VMware representative for detailed sizing of your vCloud environment.

4.4.1 Public vCloud Sizing Example

The service definition states that 50% of the virtual machines will use the Reservation Pool model and 50% will use the Pay-As-You-Go allocation model. The Reservation Pool model is applied to small, medium, and large pools with a respective split of 75%, 20%, and 5%. Therefore, *small* represents 37.5% of the total, *medium* represents 10% of the total, and *large* represents 2.5% of the total number of virtual machines in the environment.

Table 4 lists the virtual machine count for the various virtual datacenters. The total virtual machine count of 1,500 reflects the specifications outlined in the *Service Definition for a Public VMware vCloud*. Change this total to reflect your own target virtual machine count.

Table 4. Definition of Resource Pool and Virtual Machine Split

Type of Resource Pool	Total Percentage	Total Virtual Machines
Pay-As-You-Go	50%	750
Small Reservation Pool	37.5%	563*
Medium Reservation Pool	10%	150
Large Reservation Pool	2.5%	37*
TOTAL	100%	1,500

Note Some total virtual machines values rounded up or down due to percentages.

The *Service Definition for a Public VMware vCloud* also calls out the distribution for virtual machines in the vCloud with 45% small, 35% medium, 15% large, and 5% extra large. Table 5 shows the total amount of CPU, memory, storage, and networking needed.

Table 5. Memory, CPU, Storage, and Networking

Item	# of VM	Percent	vCPU	Memory	Storage	Networking
Small	675	45%	675	675GB	40.5TB	400Gb
Medium	525	35%	1,050	1,050GB	31.5TB	300Gb
Large	225	15%	900	900GB	54TB	400Gb
Extra Large	75	5%	600	600GB	4.5TB	200Gb
TOTAL	1500	100%	3,225	3,225GB	130.5	1,300Gb

Before determining your final sizing numbers, refer to VMware best practices for common consolidation ratios. Table 6 shows what the final numbers might look like using typical consolidation ratios seen in field deployments.

Table 6. Example Consolidation Ratios

Resource	Before	Ratio	After
CPU	3,225	8:1	403 vCPUs
Memory	3,225GB	1.6:1	2,016GB
Storage	130.5TB	2.5:1	52TB
Network	1,300Gb	6:1	217Gb

Sixteen hosts with the following configuration can support the required capacity:

- Socket count: 4
- Core count: 6
- Hyper threading: Yes
- Memory: 144GB
- Networking: Dual 10 GigE

These calculations do not factor in storage consumed by consumer or provider templates, nor do they take into account the resources consumed by vShield Edge appliances. A vShield Edge device backs each private organization network and external routed organization network. The service definition target of 25 organizations requires 275 vShield Edge appliances.

The specifications for each vShield Edge appliance are.

- CPU: 1 vCPU
- Memory: 64MB
- Storage: 16MB
- Network: 1 GigE (this is already calculated in the throughput of the workloads and should not be added again)

4.4.2 vCloud Maximums

Scalability in vCloud infrastructures reflects the ability of the platform to manage increasing numbers of vCloud consumers and workloads with minimal impact on manageability, performance, and reliability. From the consumer's perspective, scalability refers to the ability to consume infrastructure resources on-demand in a responsive fashion.

When designing for scale, consider the maximums of the vCloud platform as well as the underlying vSphere platform. vCloud Director 1.5 is tied to the release of vSphere 5.0, which brings many platform improvements and enhancements. vCloud Director also introduces a number of new features that can impact scalability, including fast provisioning, extensions, SQL Server support, third-party distributed switch integration, and UUIDs.

vCloud Director web console maximums are the primary constraint, followed by vSphere platform maximums. The choice of the vCloud Director database platform (Oracle or SQL Server) may result in slight performance differences.

The maximum number of vShield Edge devices per vCenter Server is 1000. Table 7 lists vCloud Maximums based on a 10-cell configuration.

Table 7. vCloud Maximums

Constraint	Limit	Explanation
Virtual machines per vCloud Director	20000	The maximum number of virtual machines that may be resident in a vCloud instance.
Powered on Virtual Machines per vCloud Director	10000	Number of concurrently powered on virtual machines permitted per vCloud instance.
Virtual machines per vApp	64	The maximum number of virtual machines that can reside in a single vApp.
Hosts per vCloud Director	2000	Number of hosts that can be managed by a single vCloud instance.
vCenter Servers per vCloud Director	25	Number of vCenter servers that can be managed by a single vCloud instance.
Users per vCloud Director	10000	The maximum number of users that can be managed by a single vCloud instance.
Organizations per vCloud Director	10000	The maximum number of organizations that can be created in a single vCloud instance.

vApps per organization	500	The maximum number of vApps that can be deployed in a single organization.
Virtual datacenters per vCloud Director	10000	The maximum number of virtual datacenters that can be created in a single vCloud instance.
Datastores per vCloud Director	1024	Number of datastores that can be managed by a single vCloud instance.
Networks per vCloud Director	7500	The maximum number of logical networks that can be deployed in a single vCloud instance.
Catalogs per vCloud Director	1000	The maximum number of catalogs that can be created in a single vCloud instance.
Media Items per vCloud Director	1000	The maximum number of media items that can be created in a single vCloud instance.

See *Configuration Maximums for VMware vSphere 5.0* for more information (<https://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-pubs.html>).

5. vCloud Resource Design

Resource design for vCloud involves examining requirements to determine how best to partition and organize resources. With the commoditization of infrastructure resources, the ability to scale these fungible units up and down becomes increasingly important.

When designing for vCloud, keep in mind that the ultimate consumers of the product are the end-users of system. These users have a varying range of technical skills and experience, typically less than that of the architects and administrators of the vCloud environment. To encourage the use of vCloud computing as an effective tool, simplify user decision points where possible. If complexity is unavoidable, document all required steps to guide the end-users through a particular process.

Taking a top-down approach to vCloud design necessitates understanding of the new abstractions introduced in the vCloud API and how they map to traditional vSphere objects.

5.1 vCloud Director Constructs

VMware vCloud Director introduces logical constructs to facilitate multi-tenancy and provide interoperability between vCloud instances built to the vCloud API standard.

Figure 8 shows the logical constructs within vCloud Director that abstract underlying vSphere resources.

Figure 8. Physical, Virtual, and vCloud Abstraction Mapping

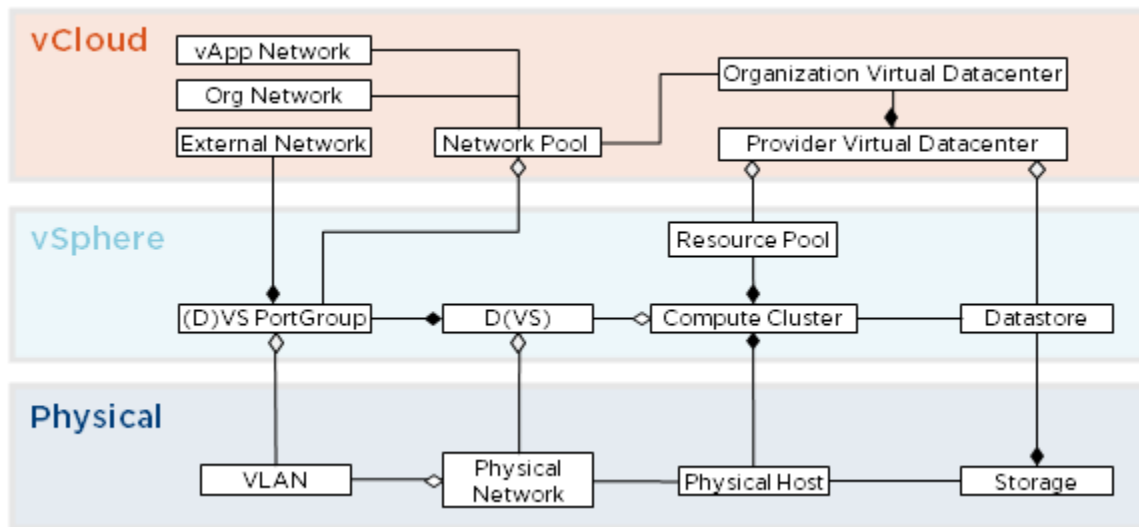


Table 8 provides descriptions of each construct.

Table 8. vCloud Director Constructs

Construct	Definition
Organization	The unit of multi-tenancy that represents a single logical security boundary. An organization contains users, virtual datacenters, and networks.
Provider virtual datacenter	A grouping of compute and storage resources from a single vCenter Server. A provider virtual datacenter consists of a single resource pool and one or more datastores. Multiple organizations can share provider virtual datacenter resources.
Organization virtual datacenter	<p>A sub-grouping of compute and storage resources allocated from a provider virtual datacenter and assigned to a single organization. A virtual datacenter is a deployment environment where vApps can be instantiated, deployed, and powered on.</p> <p>An organization virtual datacenter allocates resources using one of the following models:</p> <ul style="list-style-type: none"> • Pay-As-You-Go • Reservation • Allocation
Catalog	A repository of vApp templates and media available to users for deployment. Catalogs can be published to all organizations in the same vCloud environment.
vApp	A container for a software solution in the vCloud, and the standard unit of deployment for workloads in vCloud Director. vApps contain one or more virtual machines, have power-on operations, and can be imported or exported as an OVF.
External network	External networks provide external connectivity to organization networks and are backed by port groups configured for Internet accessibility.
Organization network	Organization networks are instantiated through network pools and bound to a single organization. Organization networks map to a vSphere port group and can be isolated, routed, or directly connected to an external network.
vApp network	A network that connects virtual machines within a vApp, deployed by a consumer from a network pool. vApp networks can be directly connected or routed to an organization network.
Network pool	A network pool is a collection of isolated Layer 2 virtual networks available to vCloud Director for the automated deployment of private and NAT-routed networks.

Use the vSphere Client to observe how creating entities through vCloud Director translate into vCenter Server tasks.

5.2 Organizations

Organizations are the unit of multi-tenancy within vCloud Director and represent a single logical security boundary. Each organization contains a collection of end users, computing resources, catalogs, and vCloud workloads. For a public vCloud, vCloud Director organizations typically represent different customers. In a private vCloud, organizations can map to different department or business units. Each department or business unit may have several environments, such as development or production.

Organization users can be local users or imported from an LDAP server. LDAP integration can be specific to an organization or inherit the system LDAP configuration defined by the vCloud system administrator. For information about how to configure LDAP, see the *vCloud Installation and Configuration Guide* (http://www.vmware.com/support/pubs/vcd_pubs.html). Create a local organization administrator for each organization to mitigate loss of administrative control due to LDAP authentication or connectivity issues.

The name of the organization, specified during creation time, maps to a unique URL that allows access to the GUI for that organization. For example, an organization named NewCo maps to the following URL: `https://<hostname>/cloud/org/NewCo`. Use a standard naming convention for organization names and avoid using special characters or spaces because they affect the URL in undesirable ways.

Use system defaults for most of the other organization settings, with the exception of leases, quotas, and limits. There are no specific requirements called out by the service definitions for these values—set them as needed.

5.2.1 Administrative Organization

A common best practice is to create an administrative organization. This organization provides a sandbox for system administrators and maintains a master catalog of vApp templates published to all other organizations in the vCloud environment. Users in an organization typically consume resources by deploying vApps from a predefined catalog. The master catalog provides a global library of standardized vApp templates to promote reusability of common assets built to provider standards.

Administrators assigned to the administrative organization are responsible for creating standardized gold master vApp templates for inclusion in the master catalog. Place non-finalized vApps in a non-published internal catalog.

Configure the administrative organization to allow catalog publishing. Create a Pay-As-You-Go organization virtual datacenter to minimize the amount of resources reserved.

5.2.2 Standard Organizations

Create an organization for each tenant of the vCloud with the following considerations:

- Cannot publish catalogs.
- Use leases, quotas, and limits that meet the provider's requirements.

5.2.3 Policies

Policies govern end-user behavior in vCloud environments. During the creation of an organization, policies can be set for the total number of running and stored virtual machines:

- *Running VM quota* refers to the maximum number of powered on virtual machines.
- *Stored VM quota* refers to the maximum number of all virtual machines, including powered off virtual machines.

Lease policies govern the persistence of vApps and vApp templates in an organization virtual datacenter. Specify the maximum length of time vApps can run and be stored in the organization virtual datacenters:

- The maximum runtime lease specifies the amount of time vApps can run before vCloud Director automatically stops them.
- The storage lease specifies the amount of time vApps or vApp templates are stored before vCloud Director automatically performs storage cleanup.
- Lease policies can also be set to **never expire**.

When any option for storage lease except **never expire** is selected, the storage is automatically cleaned up. Storage cleanup options include:

- **Permanently deleted** – After the lease expires, the vApp or vApp template is automatically deleted.
- **Moved to expired items** – This flags the vApps or vApp templates for deletion. Items move to the expired items view where they are unusable unless the lease is renewed.

5.3 Provider Virtual Datacenter

The *virtual datacenter* is a new construct that represents the standard container for a pool of compute and storage resources. There are two types of virtual datacenters: provider and organization. Provider virtual datacenters are composed of resource pools and datastores from a single vCenter Server. When creating a provider virtual datacenter, observe the following guidelines:

- Define standard units of consumption. Variance in virtual datacenter allocations decreases manageability. Look at existing trends to determine common container sizes.
- Resource pools can map to a single provider virtual datacenter.
- If enough capacity exists, map the root resource pool of the cluster to provider virtual datacenters. This simplifies resource management. If the cluster expands, the backed provider virtual datacenter automatically grows as well. This is not the case if a standard resource pool is used. Multiple parent-level resource pools can add unnecessary complexity and lead to unpredictable results or inefficient use of resources if the reservations are not set appropriately.
- Create multiple provider virtual datacenters to differentiate computing levels or performance characteristics of a service offering. An example of differentiating by availability would be N+1 for a Bronze provider virtual datacenter versus N+2 for a Silver provider virtual datacenter.
- One or more datastores can be attached to a provider virtual datacenter. Multiple provider virtual datacenters can share the same datastore. For isolation and predictable storage growth, do not attach the same datastore to multiple provider virtual datacenters.
- Storage tiering is not possible within a provider virtual datacenter. Instead, supply tiered pools of storage through multiple provider virtual datacenters.
- As the level of expected consumption increases for a given provider virtual datacenter, add additional hosts to the cluster from vCenter Server and attach more datastores.

- As the number of hosts backing a provider virtual datacenter approaches the halfway mark of cluster limits, implement controls to preserve headroom and avoid reaching the cluster limits. For example, restrict the creation of additional tenants for this virtual datacenter and add additional hosts to accommodate increased resource demand for the existing tenants.
- If the cluster backing a provider virtual datacenter has reached the maximum number of hosts, create a new provider virtual datacenter associated with a separate cluster.

Refer to the private or public service definition for provider virtual datacenter sizing guidance. Consider:

- Expected number of virtual machines.
- Size of virtual machines (CPU, memory, storage).

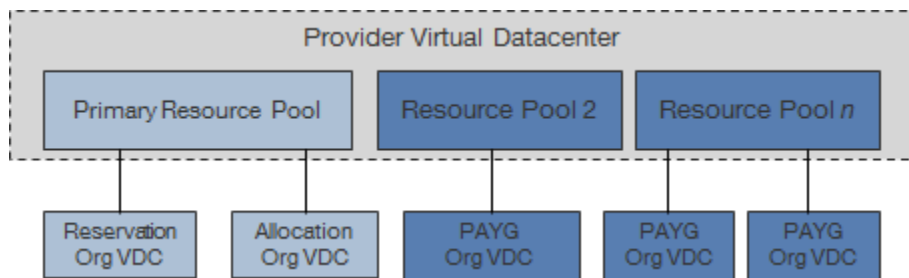
There are cases where a “special purpose” provider virtual datacenter is viewed as “special purpose.” Special use-case provider virtual datacenters are a great example of what makes vCloud computing so flexible and powerful. The primary driver behind this need for a special purpose virtual datacenter is to satisfy the license restrictions imposed by software vendors that stipulate that all the processors that *could* run specific software must be licensed for it, regardless of whether or not they actually *are* running that software.

To keep licensing costs down while meeting the EULA requirements of such a software vendor, create a purpose-specific provider virtual datacenter backed by the minimum number of CPU sockets needed to achieve performance requirements. Create corresponding organization virtual datacenter per tenant, and provide prescriptive naming to guide users to deploy workloads accordingly.

5.3.1 Extending Virtual Datacenters

Rapid elasticity is one of the primary characteristics of vCloud computing. This involves quickly adding and releasing resources based on customer usage demands. vCloud Director 1.5 introduces the ability to expand a provider virtual datacenter by adding multiple resource pools.

Figure 9. Extending Virtual Datacenters



Considerations when extending a provider virtual datacenter:

- The primary resource pool is the resource pool used in the initial creation of the provider virtual datacenter.
- Reservation Pool and Allocation Pool virtual datacenters are bound to the primary resource pool and cannot draw resources from multiple resource pools.
- After creating a provider virtual datacenter, system administrators can add additional resource pools through the web console or vCloud API. This allows a provider virtual datacenter to draw resources from multiple resource pools.

- Resource pools added to existing provider virtual datacenters supply resources to Pay-As-You-Go organization virtual datacenter. Initially only Pay-As-You-Go virtual datacenters can draw from multiple resource pools. Elasticity is limited to a single vCenter datacenter. A provider virtual datacenter can only draw resources from resource pools created in the same vCenter datacenter as the primary resource pool.
- Newly added resource pools may connect to datastores that have not been added to the provider virtual datacenter. Add all visible datastores to the provider virtual datacenter.
- Do not add extra resource pools from the same compute cluster that is already backing a provider virtual datacenter. Instead, increase the size of the existing resource pool that is mapped to the virtual datacenter.
- For elastic Pay-As-You-Go virtual datacenters, vCloud Director places the vApp in the resource pool with the most available constrained capacity.

5.4 Organization Virtual Datacenters

An organization virtual datacenter allocates resources from a provider virtual datacenter and makes it available for use for a given organization. Multiple organization virtual datacenters can share the resources of the same provider virtual datacenter.

Network pools provide network resources to organization virtual datacenters. When creating an organization virtual datacenter, select a network pool and specify the maximum allowable number of provisioned networks to allow users to self-provision vApp networks.

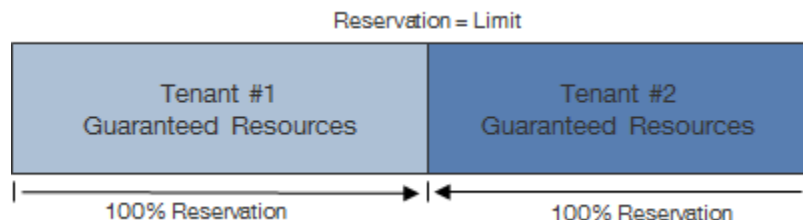
5.4.1 Allocation Models

Organizations can draw resources from multiple organization virtual datacenters using one of three resource allocation models: Reservation Pool, Allocation Pool, or Pay-As-You-Go.

5.4.1.1. Reservation Pool

Reservation Pool resources allocated to the organization virtual datacenter are completely dedicated. This is identical to an allocation pool with all guarantees set to 100%. Reservation pool virtual datacenters map to resource pools with the reservations set equivalent to the limits.

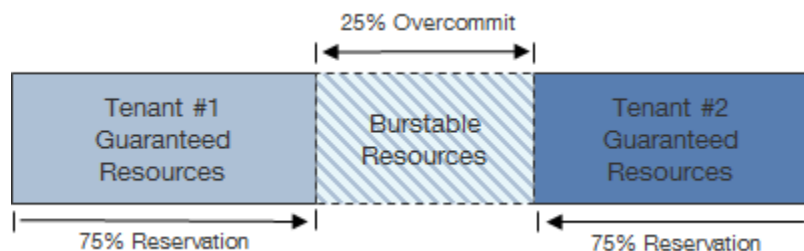
Figure 10. Reservation Pool



5.4.1.2. Allocation Pool

An Allocation Pool is a pool of allocated resources with a certain percentage guaranteed. The percentage guaranteed directly translates into reservations configured on the sub-resource pool. The difference between the reservation and the limit are resources that can be oversubscribed. In Figure 11, two tenants have organization virtual datacenters with 75% guaranteed. Resource usage of the two tenants cannot exceed the combined total of the reserved resources (75% for each) plus the resources available for overcommitment (25%). The percentage of resources guaranteed is not visible to end-consumers, who only see the total resources allocated.

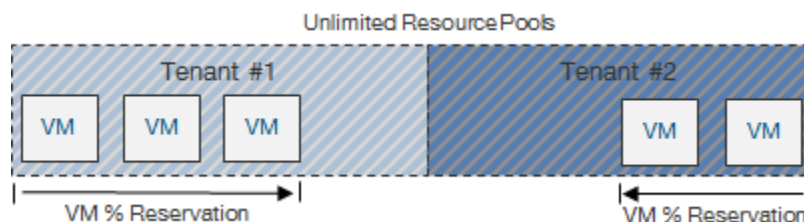
Figure 11. Allocation Pool



5.4.1.3. Pay-As-You-Go

The Pay-As-You-Go model provides the illusion of an unlimited resource pool. This model maps to a sub-resource pool with no configured reservations or limits. Resources are committed only when vApps are deployed in the organization virtual datacenter.

Figure 12. Pay-As-You-Go



During the creation of an organization virtual datacenter, vCenter Server creates child resource pools with corresponding resource reservations and limits under the resource pool that represent the provider virtual datacenter.

For each vCloud tenant, review the applicable service definition to determine the number and types of organization virtual datacenters to create. Consider expected use cases, workload types, future capacity, and the maximum number of virtual machines per organization virtual datacenter.

Use prescriptive naming for organization virtual datacenters to guide expected user behavior. All users in an organization can view all allocated organization virtual datacenters.

Note Pay-As-You-Go virtual datacenters have an additional setting called **vCPU speed** that translates into the limit of the vCPU on each virtual machine. VMware recommends increasing the default vCPU speed to a minimum of 1GHz.

5.4.2 Thin Provisioning

Thin Provisioning allows oversubscription of datastores by presenting a virtual machine with more capacity than is physically allocated. For applications with predictable capacity growth, thin provisioning may provide a more efficient way of allocating capacity. When using thin provisioning, additional management processes are required. Configure vCenter Server alarms to alert when approaching an out-of-space condition, providing for sufficient time to source and provision additional storage.

Thin provisioning is an available option when configuring organization virtual datacenters. vApps created after enabling thin provisioning use thin provisioned virtual disks.

5.4.3 Fast Provisioning

Fast provisioning is a new feature that enables rapid provisioning of vApps through vSphere 5 linked clone technology. A linked clone uses the same base disk as the original, with a chain of delta disks to keep track of the differences between the original and the clone. By default, fast provisioning is enabled when allocating storage to an organization virtual datacenter. Disabling fast provisioning on organization virtual datacenters results in full clones for subsequent vApp deployments.

Fast provisioning benefits include:

- Increased elasticity – The ability to quickly provision vApps from a catalog using linked technology enable vCloud applications to scale up as needed.
- Increased operational efficiency – Usage of linked clones typically result in significant improvement in storage utilization.

Fast provisioning components are:

- Linked clone – Virtual machine created as a result of a copy operation, leveraging a redo-log based linked clone from the parent.
- Shadow VM – Full copy of the primary virtual machine used as the source for linked clone creation. A shadow VM allows cross-datastore provisioning, and is transparent to end-users. Shadow VMs are created for vApp templates only, not MyCloud vApps.

During fast provisioning, vApp files can land on the same virtual datacenter as the primary virtual machine or a different virtual datacenter. The choice of destination virtual datacenter impacts fast provisioning deployment based on the associated datastores and vCenter Servers, as shown in Table 9.

Table 9. Linked Clone Deployment

Source vCenter	Target vCenter	Source Datastore	Target Datastore	Shadow VM
VC1	VC1	DS1	DS1	Not created until linked clone depth limit is reached (default = 31).
VC1	VC1	DS1	DS2	Created on DS2 and registered on VC1.
VC1	VC2	DS1	DS1	Created on DS1 and registered on VC2.
VC1	VC2	DS1	DS2	Created on DS2 and registered on VC2.

Both source and target virtual datacenters have fast provisioning enabled. Linked clones created from VC1 use the primary virtual machine as the base disk. Linked clones created from VC2 use the shadow virtual machine as the base disk.

Fast Provisioning considerations:

- There is an 8-host cluster limit when using VMFS datastores.
- For clusters larger than eight hosts that require linked clones, use NFS datastores.
- Separate datastores reserved for fast provisioning from datastores reserved for full clones to improve performance and manageability.
- Fast provisioning requires vSphere 5 (vCenter Server 5 and ESXi 5).
- Provisioning time is nearly instantaneous when provisioning to the same datastore.
- Provisioning a virtual machine to a different datastore triggers creation of shadow VMs if one does not already exist on the target datastore.
- Shadow VMs are full copies of the source virtual machines, which factors into sizing considerations for pre-provisioning shadow VMs across datastores.
- Storage array caching can boost linked clone performance. Ample storage array cache greatly benefits an environment that utilizes linked clones.
- Although there is no limit to the width of a tree, datastores can fill up if a tree gets too wide. Use cross-datastore linked clones to mitigate this issue.
- The maximum linked clone chain length is 30. Further clones of the vApp result in full clones.
- Shadow VMs are treated differently from normal virtual machines and can be referenced through the vCloud API by the `SHADOW_VM` entity type.
- Only invoke Storage vMotion migration of linked clones through the vCloud API (`Relocate_VM` call). The target virtual datacenter must have visibility to the datastore that contains the source disks.
- Avoid invoking Storage vMotion operations on linked clones through the vSphere Client as this consolidates the linked-clones and may result in inconsistent behavior.

Caution *Do not* enable storage DRS on datastore clusters used with vCloud Director. If storage DRS enabled datastore is used as part of a provider virtual datacenter, linked clones may be migrated to another datastore in the provider virtual datacenter.

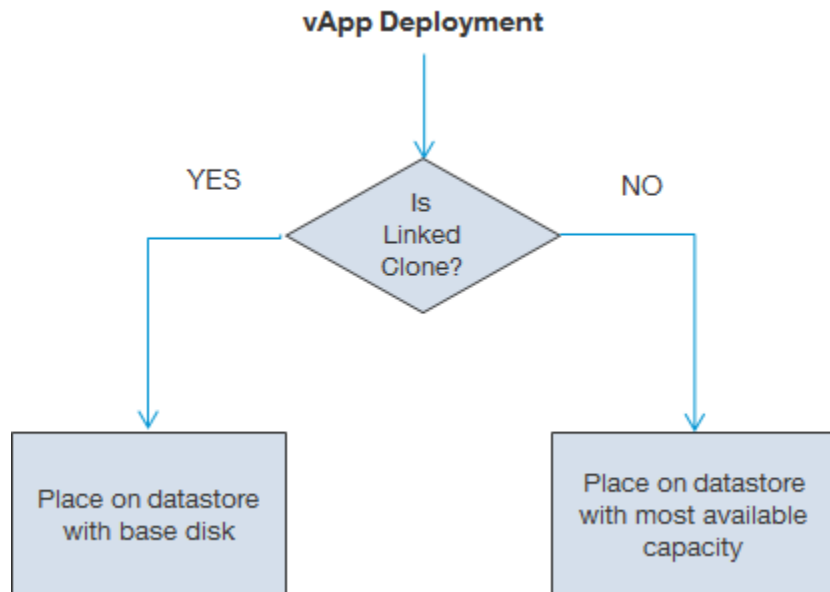
5.4.4 vApp Placement Algorithm

During vApp deployments, the virtual machine storage placement algorithm is as follows:

1. For fast provisioning-enabled virtual datacenters, find a datastore containing a base disk. If a base disk for the virtual machine exists, place a virtual machine on that datastore. The following conditions apply if the target datastore is reaching yellow or red disk thresholds.
 - If base disk exists but target datastore exceeds red threshold, look for a normal or yellow-threshold datastore. If no suitable datastores are available, the operation will fail.
 - If base disk exists but target datastore exceeds yellow threshold, look for a datastore that has not reached its yellow threshold. If none exists, deploy on the target datastore if capacity is sufficient.
2. If no base disk exists, place the virtual machine on the datastore with the most available capacity that does not exceed yellow threshold.

Figure 13 provides a flowchart of the vApp placement algorithm.

Figure 13. vApp Placement Algorithm



vApp creation fails if the vApp contains multiple virtual machines that cannot fit on a single datastore in the target virtual datacenter. Consider the following scenario:

- Virtual Datacenter1:
- Datastore1 – 30GB free space
- Datastore2 – 30GB free space
- vApp1 :
- VM1 – 20GB
- VM2 – 30GB

Because the total size required for vApp1 exceeds the maximum available capacity of all datastores, the vApp deployment task fails. To mitigate this risk, follow VMware best practices for datastore utilization through proactive monitoring and storage maintenance.

5.4.5 Public vCloud Considerations

The *Service Definition for a Public VMware vCloud* provides some of the requirements used in this example.

Table 10. Public vCloud Virtual Datacenter Requirements

Requirement
Three different service offerings are required: Basic (Pay-As-You-Go), Committed (Allocation Pool), Dedicated (Reservation Pool).
vCloud infrastructure to support a minimum of 1500 virtual machines across the three service offerings.
Split reservation pool into small, medium, and large pools with a split of 75%, 20%, and 5%.

- The *Basic* service offering uses the Pay-As-You-Go allocation model allowing customers to vary their resource usage while being charged for the resources that they consume.
- The *Committed* service offering uses the Allocation Pool model, which specifies a resource container size that has a certain percentage reserved.
- The *Dedicated* service offering uses the Reservation Pool model because this offering requires dedicated and guaranteed resources for the consumer.

The *Service Definition for the Public VMware vCloud* has specific requirements for the maximum number of virtual machines each organization can have based on size. Refer to the service definition for the maximum virtual machine count for each virtual datacenter type

The service definition provides detailed and descriptive guidance on how much a provider should charge for each service tier. VMware vCenter Chargeback integrates with vCloud Director to provide metering and cost calculation functionality. See the *VMware vCenter Chargeback User's Guide* (https://www.vmware.com/support/pubs/vcbm_pubs.html) for information.

Using vCenter Chargeback with vCloud Director (http://www.vmware.com/support/pubs/vcbm_pubs.html) details how to set up vCloud Director and vCenter Chargeback to accommodate instance-based pricing (Pay-As-You-Go), reservation-based pricing, and allocation-based pricing.

5.4.6 Private vCloud Considerations

The *Service Definition for a Private VMware vCloud* provides the requirements used in this example.

Table 11. Private vCloud Virtual Datacenter Requirements

Description
Three different service offerings are required: Basic (Pay-As-You-Go), Committed (Allocation Pool), Dedicated (Reservation Pool).
vCloud infrastructure to support a minimum of 1500 virtual machines across the three service offerings.
Split reservation pool into small, medium, and large pools with a split of 75%, 20%, and 5%.

Each organization virtual datacenter has a storage limit specified unless using the Pay-As-You-Go allocation model, which can be set to unlimited. For this example, no storage limit is set because we are providing static values for the individual virtual machine storage and are limiting the number of virtual machines in an organization. To improve storage efficiency, enable thin provisioning on organization virtual datacenters.

5.5 vCloud Networking

Workloads for vCloud consumers require network connectivity at the following levels:

- External networks connect vApps to outside networks. An external networks maps to a vSphere port group with external connectivity.
- Internal or routed networks, used to facilitate virtual machine-to-virtual machine communication within a vCloud. These are backed by vCloud Director network pools.
- Network design complexity depends on vCloud workload requirements. A vApp with a large number of upstream dependencies is more complex to deploy than a vApp with a self-contained application.
- vCloud Director coordinates with vShield Manager to provide automated network security for a vCloud environment. vShield Edge devices are automatically deployed during the provisioning of routed or private networks. Each virtual router runs a firewall service that allows or blocks inbound traffic to virtual machines that are connected to a public access organization network. The vCloud Director web console exposes the ability to create five-tuple firewall rules that are comprised of source address, destination address, source port, destination port, and protocol.

5.5.1 External Networks

- An external network provides connectivity “outside” an organization through an existing, preconfigured vSphere port group. These can be a standard port group, distributed port group, or a third-party distributed switch port group construct such as the Cisco Nexus 1000V port profile.
- In a public vCloud, external networks can provide access through the Internet to customer networks, typically using VPN or MPLS termination. Before creating external networks, provision the requisite number of vSphere port groups with external connectivity.

5.5.2 Network Pools

- Network pools contain network definitions used to instantiate private/routed organization and vApp networks. Networks created from network pools must be isolated at Layer 2.
- Three types of network pools are available:
 - vSphere port group-backed network pools are backed by pre-provisioned port groups, distributed port groups, or third-party distributed switch port groups.
 - VLAN-backed network pools are backed by a range of pre-provisioned VLAN IDs. This assumes all VLANs specified are trunked into the vCloud environment (requires distributed switch).
 - vCloud Director Network Isolation-backed (VCD-NI) network pools are backed by vCloud isolated networks. A vCloud isolated network is an overlay network uniquely identified by a fence ID implemented through encapsulation techniques which span hosts and provides traffic isolation from other networks (requires distributed switch).

Table 12 compares the options for a network pool.

Table 12. Network Pool Options

Consideration	vSphere Port Group Backed	VLAN Backed	vCloud Network Isolation Backed
How it works	Isolated port groups must be created and exist on all hosts in cluster.	Uses range of available, VLANs dedicated for vCloud.	Creates an overlay network (with fence ID) within a shared transport network.
Advantages	N/A	Best network performance. vCloud Director creates port groups as needed.	Best scalability, able to create thousands of networks per transport network. More secure than VLAN backed option due to VCD enforcement. vCloud Director creates port groups as needed.
Disadvantages	Requires manual creation and management of port groups. Possible to use a port group that is in fact not isolated.	Potential for VLAN exhaustion in larger vCloud environments.	Overhead required to perform encapsulation.

vSphere Port Group-backed considerations:

- Standard or distributed virtual switches may be used.
- vCloud Director does not automatically create port groups. Manually provision port groups ahead of time for vCloud Director to use.

VLAN-backed considerations:

- Distributed switches are required.
- vCloud Director creates port groups automatically as needed.

vCloud Network Isolation-backed considerations:

- Distributed switches are required
- Increase the MTU size of network devices in the transport VLAN to at least 1600 to accommodate the additional information needed for VCD-NI. This includes all physical switches and vSphere Distributed Switches. Failure to increase the MTU size causes packet fragmentation, negatively affecting network throughput performance of vCloud workloads.
- Specify a VLAN ID for the VCD-NI transport network (this is optional, but recommended for security). Leaving this blank defaults to VLAN 0.
- The maximum number of VCD-NI backed network pools per vCloud instance is 10.
- vCloud Director creates port groups automatically on distributed switches as needed.

5.5.3 Organization Networks

Organization networks provide network connectivity to vApp workloads within an organization. Users in an organization have no visibility into external networks and connect to outside networks through external organization networks. This is analogous to users in an organization connecting to a corporate network that is uplinked to a service provider for Internet access.

Connectivity options for organization networks include:

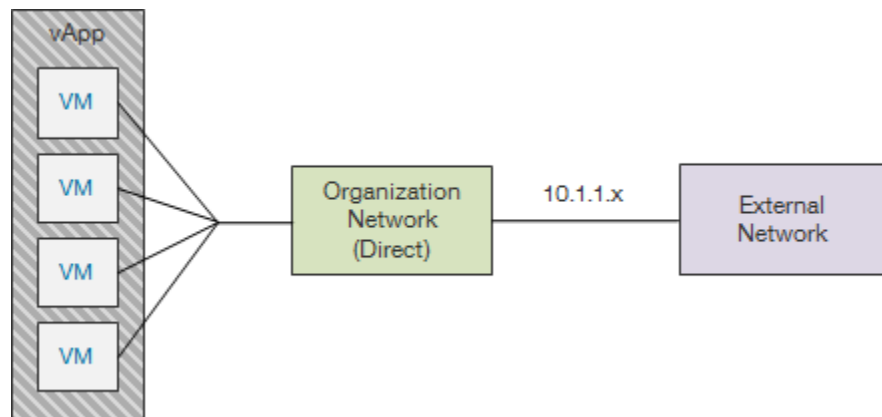
- External organization network – direct connection.
- External organization network – routed connection.
- Internal organization network – isolated.

Internal and routed organization networks are instantiated through network pools by vCloud system administrators. Organization administrators do not have the ability to provision organization networks, but can configure network services such as firewall, NAT, DHCP, VPN, and static routing.

5.5.3.1. Direct

A directly connected external organization network places the vApp virtual machines in the port group of the external network. IP address assignments for vApps follow the external network IP addressing.

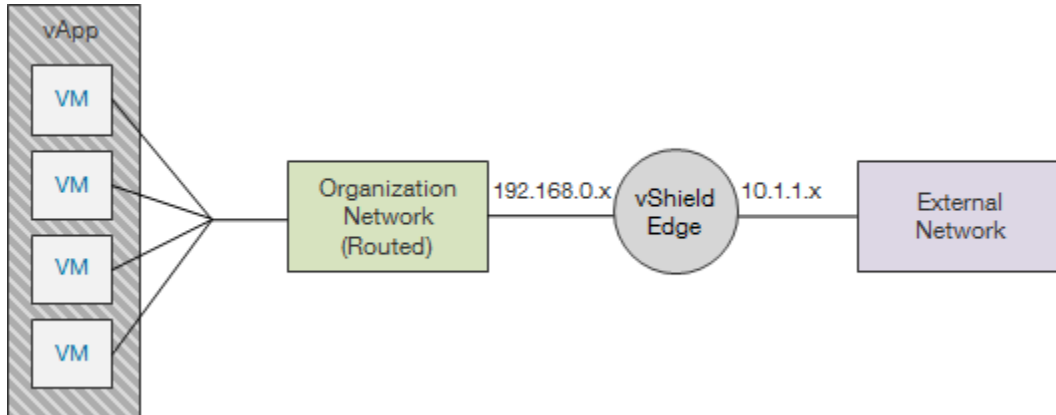
Figure 14. External Organization Network (Direct)



5.5.3.2. Routed

A routed external organization network is protected by a vShield Edge device which provides DHCP, Firewall, NAT, VPN, and static routing services. vShield Edge connects to the organization network and the external network port groups.

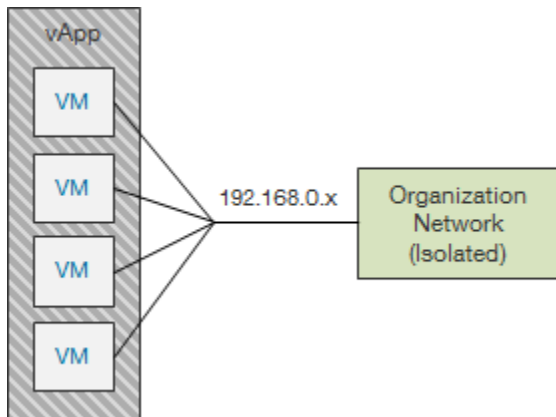
Figure 15. External Organization Network (Routed)



5.5.3.3. Isolated

An internal organization network is isolated from all other networks.

Figure 16. Internal Organization Network (Isolated)



5.5.4 vApp Networks

vApp networks are created by vCloud consumers and connect multiple virtual machines in a vApp together. vApp networks segment vApp virtual machines from the workloads in the organization network. The effect is similar to placing a router in front of a group of systems (vApp), gaining the ability to shield those systems from the rest of the corporate network. vApp networks are instantiated from a network pool and consume vSphere resources.

Connectivity options for vApp networks include:

- Direct– vApps connect directly to the organization network.
- Fenced – Allows identical virtual machines to exist in different vApps by using a virtual router to provide isolation and proxy ARP.
- Routed – Define a new network and use a virtual router to provide NAT and firewall functionality.
- Isolated – No connection to an organization network, with communication restricted to the virtual machines in the vApp.

There are two ways to create vApp networks:

- Fencing vApps directly connected to an organization network. Choosing the **fence option** automatically creates a vApp network that is not visible from the vCloud Director web console. Firewall and NAT services are configurable on a fenced network.
- Manually creating vApp networks through the **Add Network** wizard. Connecting the vApp network to an organization network creates a routed connection, with configurable NAT and firewall services.

5.5.4.1. Direct

Connecting virtual machines in a vApp directly to an organization network places vApp virtual machines in the port group of the organization network. IP address assignments for vApps follow the organization network IP addressing.

Figure 17 shows a vApp network directly connected to a direct external organization network.

Figure 17. vApp Network (Direct) → Organization Network (Direct)

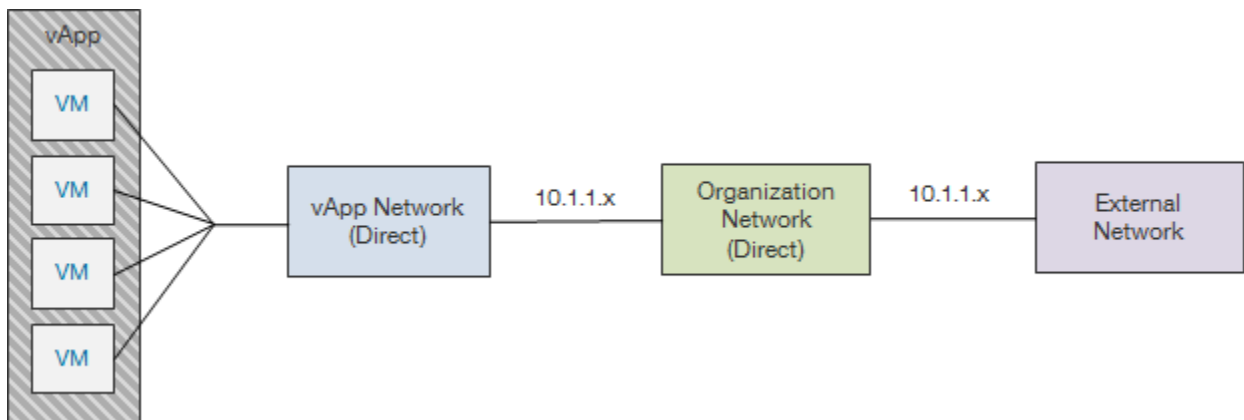


Figure 18 shows a vApp network directly connected to a routed external organization network. vShield Edge provides DHCP, Firewall, NAT, and static routing services to the organization network.

Figure 18. vApp Network (Direct) → Organization Network (Routed)

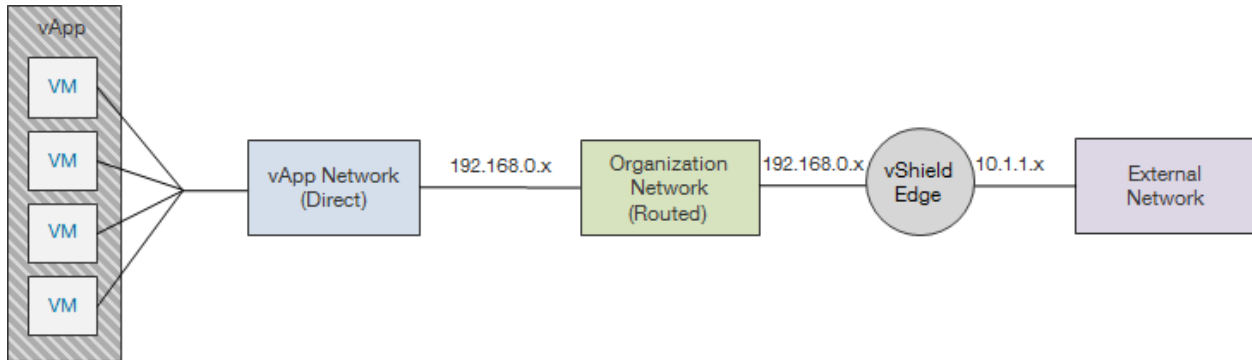
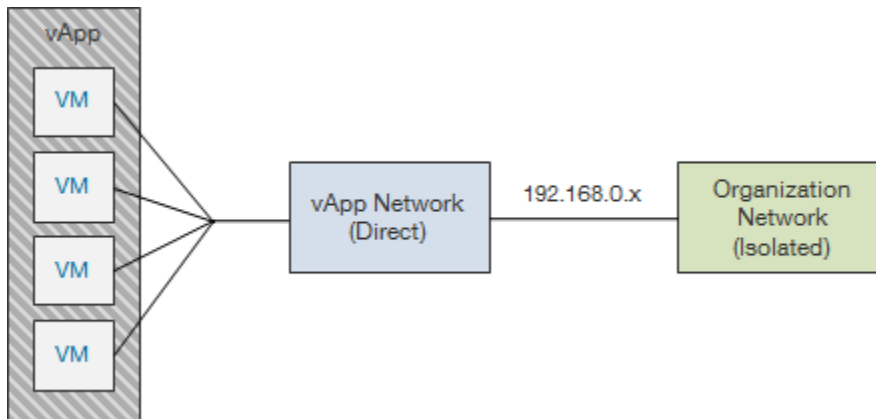


Figure 19 shows a vApp network directly connected to an isolated organization network. A vShield Edge automatically deploys only if using DHCP services.

Figure 19. vApp Network (Direct) → Organization Network (Isolated)



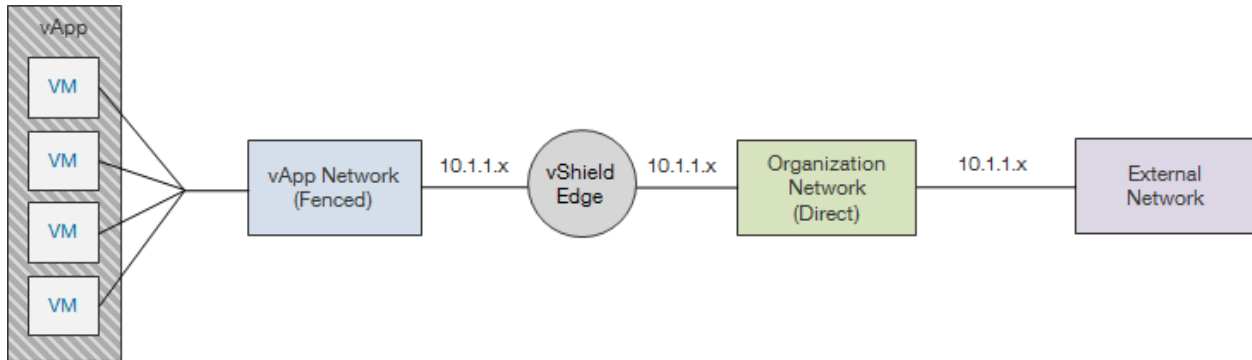
5.5.4.2. Fenced

For a fenced network, the external and internal IP subnet is the same with proxy ARP used to move traffic. vShield Edge provides the network fencing functionality for vCloud environments. The option to fence a vApp is available if the vApp directly connects to an organization network.

Depending on the organization network connection, NAT or double NAT may take place for incoming or outgoing traffic from a vApp network perspective. The following scenarios describe a single and double NAT situation.

Figure 22 illustrates a scenario where a vApp network connected to a direct organization network is fenced.

Figure 20. vApp Network (Fenced) → Organization Network (Direct)



If fencing a vApp network connected to a routed organization network, double NAT occurs with two vShield Edges deployed. Figure 20 illustrates this scenario.

Figure 21. vApp Network (Fenced) → Organization Network (Routed)

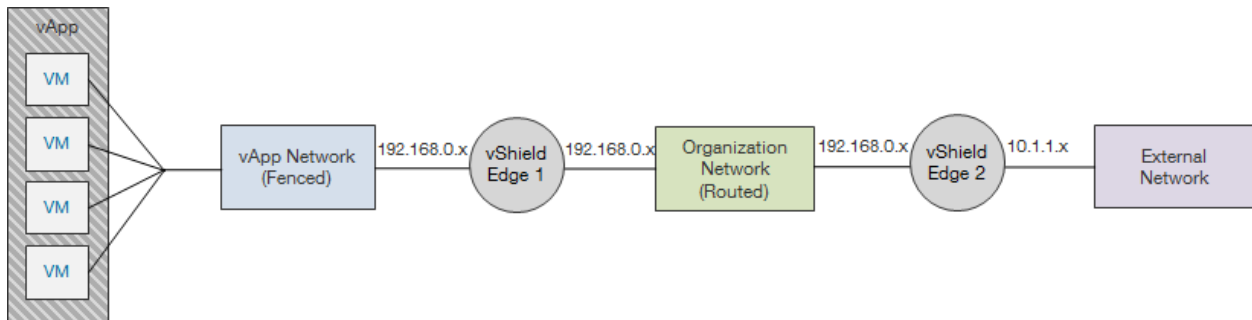
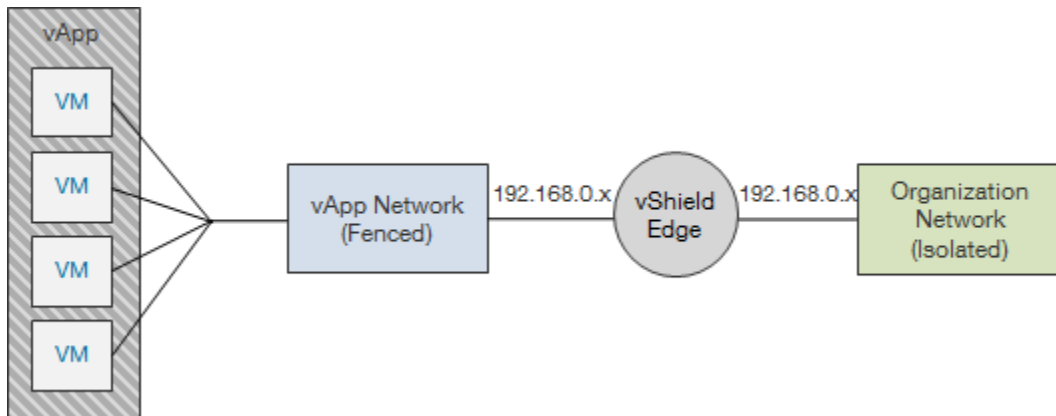


Figure 21 shows a fenced vApp network connected to an isolated organization network. There is only one NAT.

Figure 22. vApp Network (Fenced) → Organization Network (Isolated)



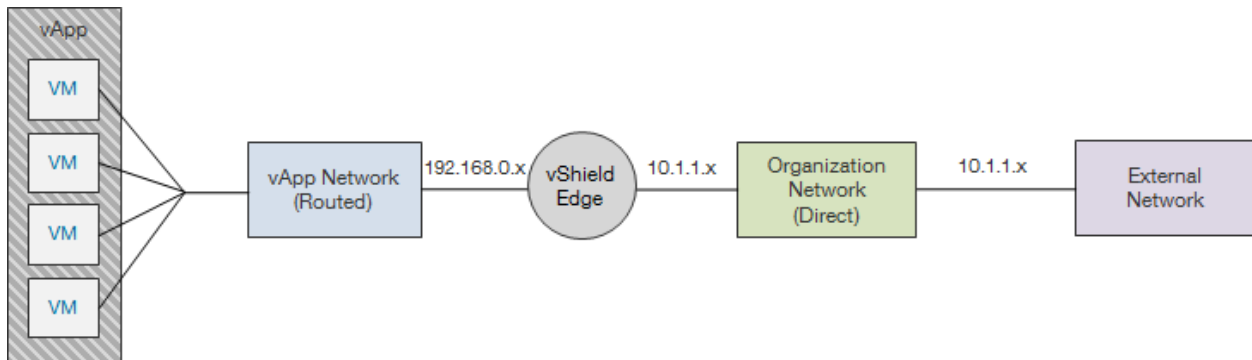
5.5.4.3. Routed

A routed vApp network is a vApp network connected to an organization network where the IP address space differs between the two networks. vShield Edge provides the DHCP, NAT, and firewall services.

Depending on the organization network connection, NAT or double NAT may take place for incoming or outgoing traffic from a vApp network perspective. The following scenarios describe a single and double NAT situation.

Figure 22 illustrates a scenario where a routed vApp network connects to a direct organization network.

Figure 23. vApp Network (Routed) → Organization Network (Direct)



If a routed vApp network connects to a routed organization network, double NAT occurs with two vShield Edges deployed. Figure 20 illustrates this scenario.

Figure 24. vApp Network (Routed) → Organization Network (Routed)

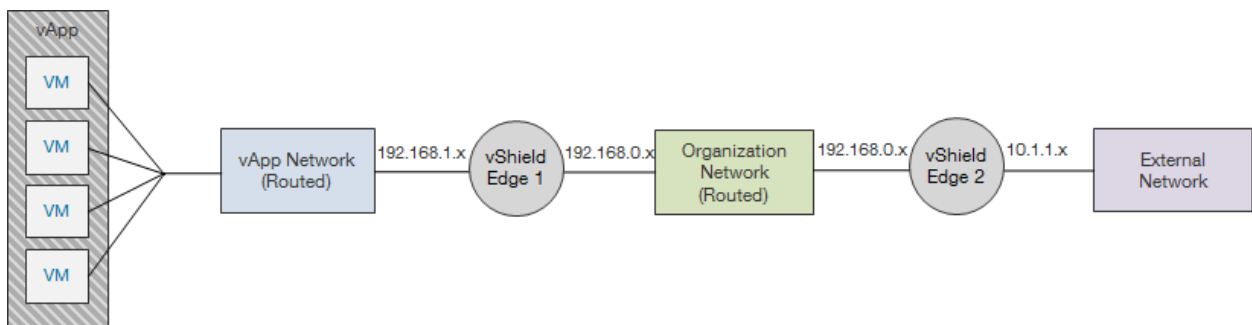
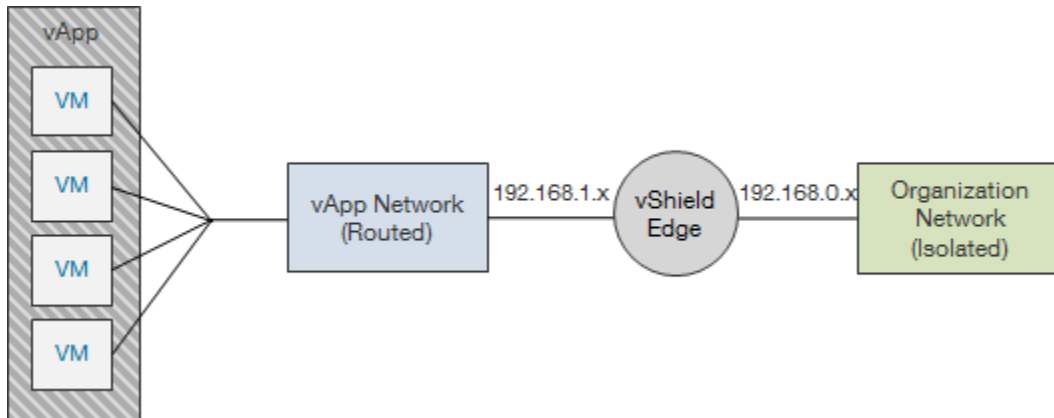


Figure 21 shows a routed vApp network connected to an isolated organization network.

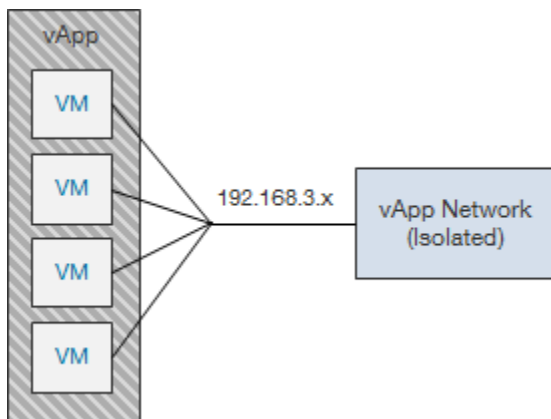
Figure 25. vApp Network (Routed) → Organization Network (Isolated)



5.5.4.4. Isolated

A vApp network configured to “none” is completely isolated and the virtual switch of the corresponding port group is the endpoint for this network. This network is isolated on Layer 2 and only intra-vApp communication is possible.

Figure 26. vApp Network (Isolated)



5.5.5 Static Routing

Another feature in vCloud Director 1.5 is support for static routing. This provides the ability to route between network segments without the use of NAT and enables increased flexibility in implementing network connectivity within a vCloud environment.

Though most networks will have a directly connected default gateway, it is possible for networks to have more than one router (such as multiple vShield Edge devices). Static routing provides a way to manually configure routing tables so that traffic can be forwarded to these remote networks while still using the default gateway for all remaining traffic.

In vCloud Director, static routing can be configured at both the routed organization network level and vApp network level.

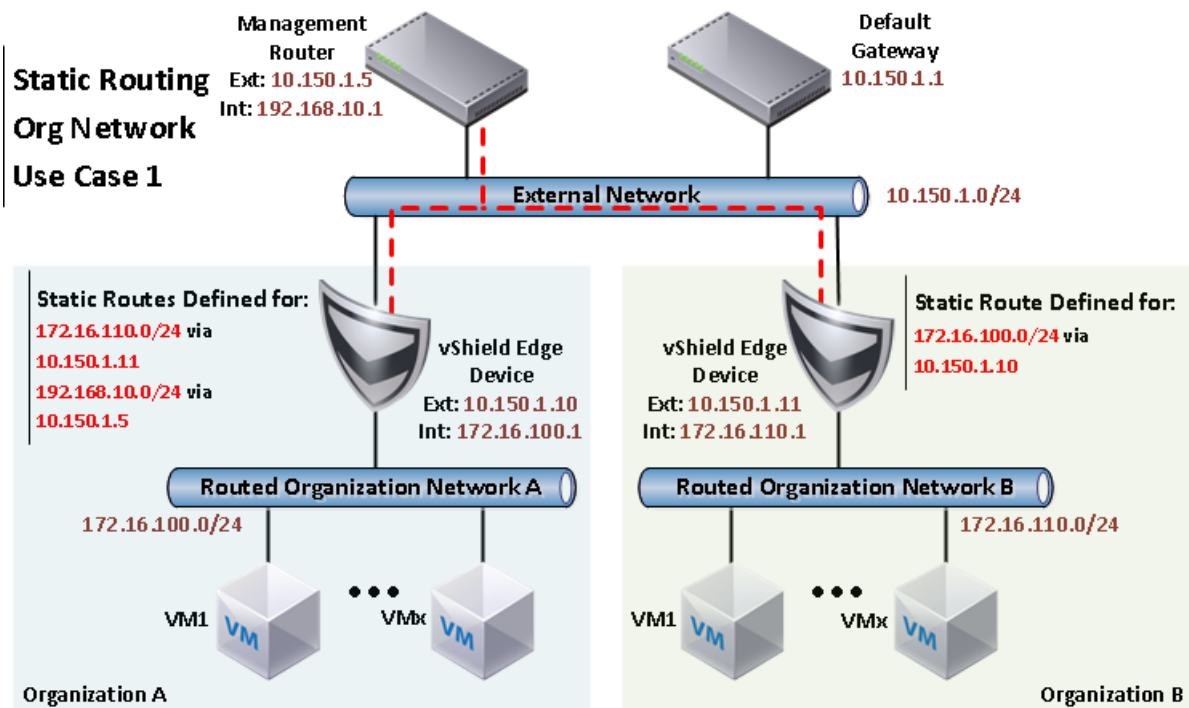
- For organization networks, routes can be defined within the organization network or to an external network.
- For vApp networks, the static routing configuration is simplified as routes are only applied on the external interface.

To demonstrate the different options for static routing with vCloud Director, several common use cases are covered:

5.5.5.1. Organization Network use cases

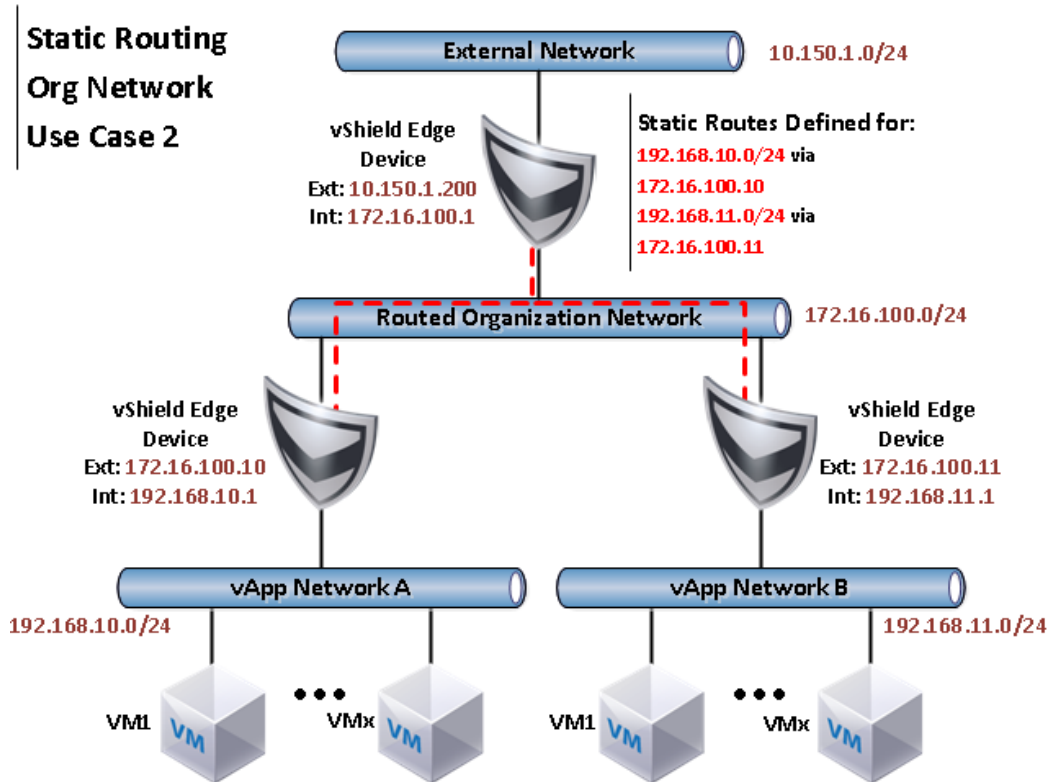
- Accessing network resources on an external network – This use case applies to scenarios where there is a requirement for connectivity to network resources through a different next hop address than the default external gateway. An example might be access to a remote management network via a VPN/proxy or accessing services in another organization.

Figure 27. Organization Network Static Routing Use Case 1



- Enabling vApp networks connected to an organization network to communicate directly – Allows virtual machines connected to different vApp networks (but a common organization network) to communicate without the use of Network Address Translation. This reduces the operational overhead of maintaining Port Forwarding or IP Translation NAT rules for connectivity within the organization.

Figure 28. Organization Network Static Routing Use Case 2

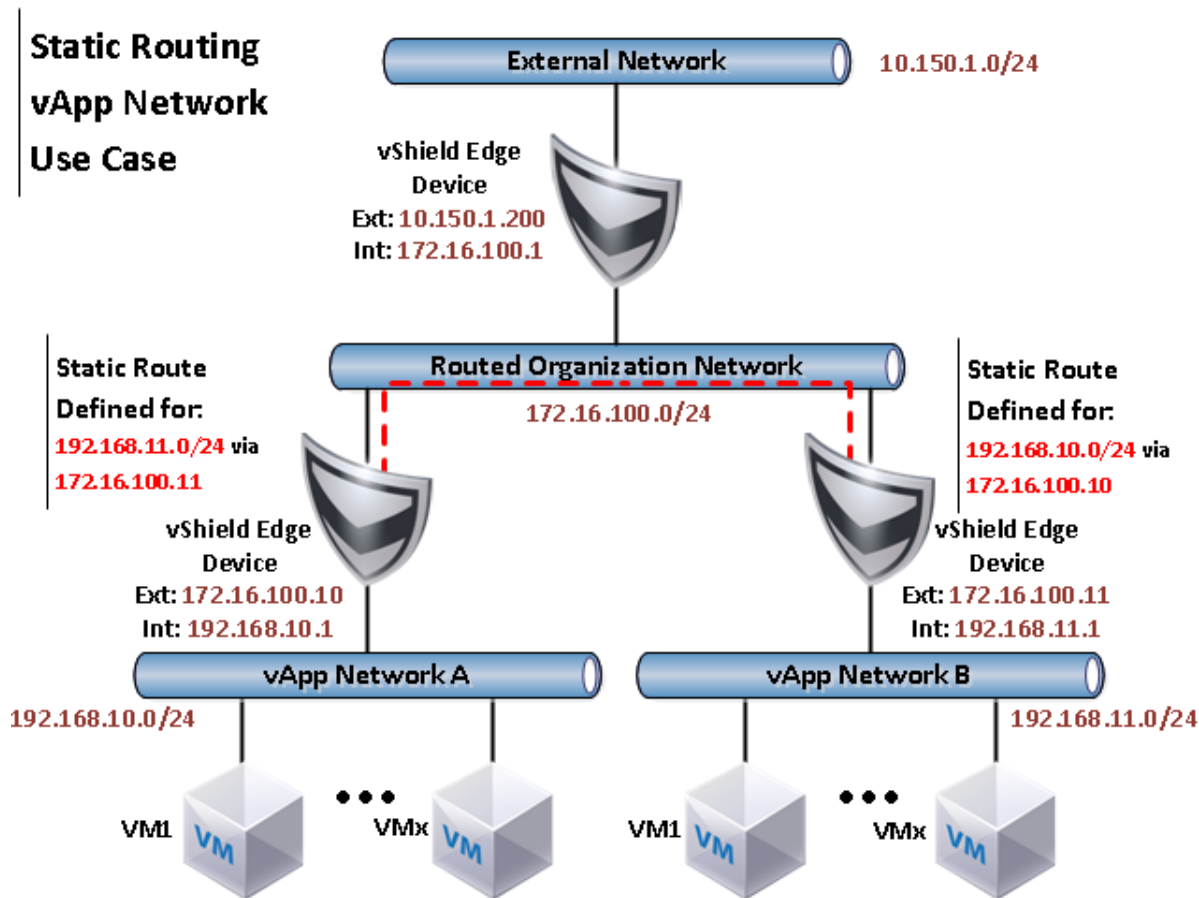


- Reducing layers of NAT from external networks to vApp networks – In vCloud Director 1.0 for a system outside the vCloud environment to access services on a virtual machine connected to a vApp network this required up to two levels of NAT (one if the organization network is directly connected, or two if the organization network is routed). Static routing significantly simplifies connectivity to external systems required for services such as monitoring and patch management or for integration into centralized services such as authentication and logging. Importantly as these routing capabilities are delivered through vShield Edge, the capability of self-service firewall management is still maintained. This is particularly important in private vCloud deployments where networks are typically flatter to support these centralized services, and static routing provides an alternative to directly connecting virtual machines to the external networks.

5.5.5.2. vApp Network use cases

- Enabling vApp networks connected to an organization network to communicate directly – This scenario provides similar connectivity to organization network use case 2.

Figure 29. vApp Network Static Routing Use Case



Note that if vApp level Static Routing is configured it is advised that the **Always use assigned IP addresses until this vApp or associated networks are deleted.** setting is enabled to make sure the next hop addresses for static routes does not change while vApps are powered off.

As there is an overlap between organization network use case 2 and the vApp network use case, it is important to understand the advantages and disadvantages of both configurations:

- Applying static routes at the organization network consolidates management to a common view, but requires all traffic to pass through the organization vShield Edge.
- vApp network static routes allow traffic directly between vApps that will provide the highest performance.
- Static routing at the vApp network layer also supports scenarios where the organization network is directly connected.

Therefore though it is required to provide connectivity between vApps without address translation, it is recommended to apply static routes at the vApp network vShield Edge device. In addition, unlike NAT, static routing does not support overlapping network ranges. This means that vCloud admins and users need to take care when allocating IP addresses for organization and vApp networks to make sure they are unique if there are plans to leverage static routing within the vCloud environment.

The static routing and NAT features are not mutually exclusive and can be used together. For example, NAT may provide external connectivity, while static routing enables direct access to other vApps within an organization.

The following limitations need to be taken into consideration when using static routing with vCloud Director:

- Only supported with vShield Edge 5.0.
- Limited to a maximum of 64 static routes per vShield Edge device.
- Dynamic Routing Protocols are not currently supported.
- Does not apply to fenced vApps.

5.5.6 Third-Party Distributed Switch Considerations

vCloud Director 1.5 enhances third-party distributed switch integration by extending support for all three network pool types. VLAN-backed and VCD-NI-based network pools are available for creation with a supported third-party distributed switch.

Environments that require increased network visibility, segregated network management, and advanced Layer 2 security capabilities may require the capabilities of third-party distributed switches such as the Cisco Nexus 1000V.

5.6 Networking – Public vCloud Example

The Service Definition for a Public VMware vCloud provides the requirements used in this example.

Table 13. Public vCloud Network Requirements

Description
Each tenant receives a pool of 8 public routable IP addresses.
Minimum of one routed organization network protected by a firewall service.
Ability to create up to 10 vApp networks.

5.6.1 External Networks

All service tiers use a shared public Internet connection. When establishing the external network:

- Map to a vSphere port group that is configured for Internet connectivity.
- Provide the network configuration details, including netmask, default gateway, and DNS.
- Reserve the static IP address range available for this network. vCloud Director automatically assigns IP addresses to devices directly connecting to external networks.
- Give the network a descriptive name, such as “Shared-Internet.”

For sizing purposes, create an IP address pool large enough to support Internet connectivity for all organizations in the vCloud. The estimated number of organizations for 1500 virtual machines is 25, so provide at least 25 IP addresses in your static IP pool. Each organization requires at least eight public IP addresses to allow inbound access to virtual machines.

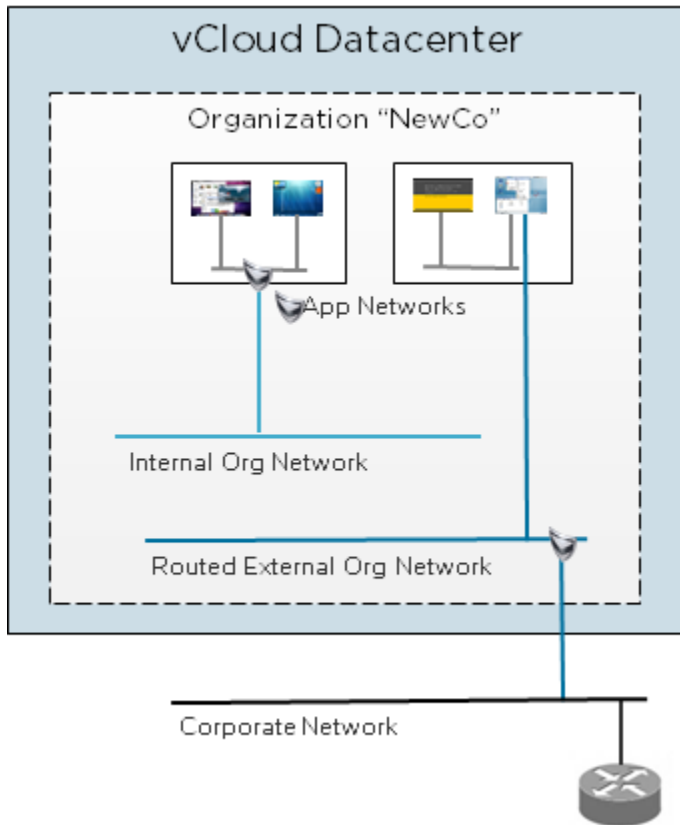
5.6.2 Network Pools

Each organization in a public vCloud requires individual private networks. vCloud Director instantiates Isolated L2 networks through the use of network pools.

Create a single vCloud Director VCD-NI network pool for all organization network deployment. VCD-NI requires the use of a distributed switch and VLAN ID that specifies the isolated transport network.

Network pools handle the automated creation of organization networks and vApp networks. A minimum of 12 networks are required in the network pool per organization, with 10 reserved for vApp networks and two used for organization networks. Given the estimate of 25 organizations, the network pool should contain at least 300 networks. vCloud Director creates auto-expandable static port groups for organization and vApp networks. The maximum number of networks in a network pool is limited to 1000.

Figure 30. Example of Public vCloud Networking



5.6.3 Organization Networks

Create two different organization networks for each organization: one routed external organization network, and one internal organization network. The **Create Organization Network Wizard** provides the option of creating these two organization networks in one workflow. When naming an organization network, start with the organization name and a hyphen, for example, "NewCo-Internet."

The routed external organization network leverages vShield Edge for firewall and NAT services to isolate organization traffic from other organizations that share the same external provider network. Both the external organization network and the internal organization networks are instantiated from the previously established vCloud Director Network Isolation network pool. Each organization network requires network configuration settings and a pool of IP addresses. Because both networks are private networks, you can use RFC 1918 addresses for each static IP address pools. The static IP address pool can be as large as desired, typically a RFC 1918 class C is used.

The last step is to add external public IP addresses to the vShield Edge configuration on the external organization network. Using the **Configure Services** interface, add eight public IP addresses to an external organization network. The IP addresses listed come from the external network static IP pool.

5.7 Networking – Private vCloud Example

The *Service Definition for a Private VMware vCloud* provides the requirements used in this example.

Table 14. Private vCloud Network Requirements

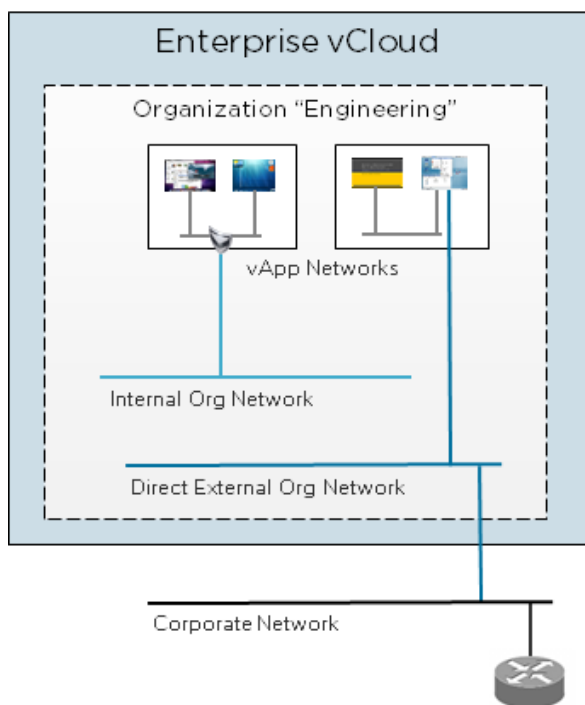
Description
vApps require a direct connection to the external network due to upstream dependencies.
An isolated network is needed for dev/test and pre-production workloads.
Users have the ability to self-provision networks.

5.7.1 External Networks

Private vCloud networking requirements tend to vary depending on the primary use cases driving the project. Enterprises acting as service providers to their internal customers tend to have comparable network requirements to that of a public vCloud. Enterprises using vCloud for development or pre-production environments will have different requirements.

Enterprises commonly require direct connections from inside the vCloud environment into the networking backbone. This is analogous to “extending a wire” from the network switch that contains the network or VLAN to be used all the way through the vCloud layers into the vApp. Each organization in the private vCloud has an internal organization network and a direct connect external organization network.

Figure 31. Example of Private vCloud Networking



At least one external network is required to enable external organization networks to access resources outside of the vCloud Director—the Internet for public vCloud deployments, and an internal (local) network for private vCloud deployments.

To establish this network, use the **New External Network** wizard and specify external network settings and static IP address ranges. For the static IP address pool a good starting range is 30 reserved IP addresses for use as static assignments.

5.7.2 Network Pools

The requirements call for one internal organization network and the ability for consumers to create private vApp networks. No minimum number of vApp networks defined, but typically, organizations start with around 10. Size the network pool to be the number organizations times 11. VMware recommends setting the maximum number of networks per network pool to 1000.

5.7.3 Organization Networks

At least one organization external network is required to connect organization vApps to other vApps and/or the networking layers beyond the private vCloud.

To accomplish this, create an external organization network using the **Create Organization Network Wizard**, and select **direct connection** from the drop-down menu. vApps that connect to this organization network are dropped directly on the vSphere port group that corresponds to the external network.

Implementing routed networking may add complexity to the networking design. For more information on adding additional network options, refer to the *vCloud Director Administrator's Guide* (https://www.vmware.com/support/pubs/vcd_pubs.html).

Catalogs are the primary deployment mechanism in vCloud Director, serving as a centralized repository for vApp templates and media. Users self-provision vApps from vApp templates located in internal catalogs or global published catalogs.

The administrative organization virtual datacenter has two catalogs:

- Internal catalog – Staging area for developing new vApp templates.
- Master catalog – Contains gold master vApp templates that are published to all organizations.

Organizations leverage the published master catalog to deploy standardized vApp templates. Each organization also has a private catalog created by the organization administrator. This private catalog is used for to upload new vApps or media to an individual organization.

Guest customization changes the identity of the vApp and can also perform post-deployment steps, such as the joining of vApps to domains.

There are no additional configuration requirements for the catalogs or vApp templates in this vCloud architecture. Refer to the private or public service definition for a full listing of recommended templates. Usually vApp templates include base operating system templates with no applications installed, or application-specific vApp templates.

5.8 vApp

A *vApp* is a container for a distributed software solution and is the standard unit of deployment in vCloud Director. It has power on operations, consists of one or more virtual machines, and can be imported or exported as an OVF package. Although similarly named, vCloud vApps have subtle differences compared to vSphere vApps. For example, vCloud vApps can contain additional constructs such as vApp networks, but do not offer the resource controls found in vSphere vApps.

5.8.1 General Design Considerations

Some general design considerations for vApps are:

- Default to one vCPU unless requirements call for more (such as a multi-threaded application).
- Always install the latest version of VMware Tools.
- Deploy virtual machines using default shares, reservations, and limits settings unless a clear requirement exists for doing otherwise.
- For virtual network adaptors, use VMXNET3, if supported.
- Secure virtual machines as you would physical machines.
- Hardware version 7 and 8 are supported, depending on the ESXi version backing the hosts in the provider virtual datacenter. vSphere 5 introduces hardware version 8.
- Use standard virtual machine naming conventions.

5.8.2 vSphere vApp Differences

5.8.2.1. OVF

An OVF section is an XML fragment that contains data for a specific functionality or aspect, such as resource settings, startup and shutdown sequence, or operating system type. The general format of an OVF section is as follows:

```
<myns:MyOvfSection ovf:required="true or false">
  <Info>A description of the purpose of the section</Info>
  ... section specific content ...
</myns:MyOvfSection>
```

Because vCloud Director does not currently support all of the OVF sections that vSphere does, the following sections of the vSphere vApp OVF representation are not visible to vCloud Director:

- AnnotationSection
- DeploymentOptionSection
- InstallSection
- ProductSection
- ResourceAllocationSection

vCloud Director and vSphere support all other OVF sections. When vCloud Director ignores a section, vSphere may interpret the contents differently than if it was a native vSphere vApp. This can result in differences in behavior when operating the imported vApp in a virtual datacenter. vCloud Director removes the ignored OVF sections during a vApp download.

6. vCloud Metering

For vCloud environments, resource metering is essential to accurately measure consumer usage and shape consumer behavior through chargeback policies. VMware vCenter Chargeback provides the metering capability to enable cost transparency and accountability in vCloud environments.

In running a private vCloud, enterprises do not necessarily have the same cost pressures as a public vCloud service provider. The requisite chargeback procedures or policies also may not exist. An alternative to chargeback is *showback*, which attempts to raise awareness of consumption usage and cost without involving formal accounting procedures to bill the usage back to the consumer's department.

vCenter Chargeback provides the cost transparency and accountability to align consumer behavior with the actual cost of the consumed resources. Without showback or chargeback, consumers are not aware of the actual cost of the resources they have consumed and thus have little incentive to change their consumption patterns. vCloud computing resources can be easily spun up, and with the exception of deployment policies that dictate resource leases, there are no disincentives or penalties to curb excessive use. Metering exposes heavy or demanding users who may monopolize vCloud resources.

6.1 vCenter Chargeback

VMware vCenter Chargeback provides the metering capability to measure, analyze, and report on resources used in private and public cloud environments. Cloud providers can configure and associate various cost models to vCloud Director entities. The cost transparency enabled by vCenter Chargeback allows cloud providers to validate and adjust financial models based on the demand for resources.

6.1.1 Chargeback Server

The vCenter Chargeback Server is a Windows server that runs the vCenter Chargeback web application, load balancer, and data collector services. This server can be virtual or physical and has the following recommended specifications:

- 2.0 GHz or faster Intel/AMD x86 processor
- 4GB or more of RAM
- 3GB disk storage
- 1Gbps Ethernet adapter

Chargeback Servers can be clustered together to provide improved performance and availability for the user interface. A cluster configuration leverages the Apache load balancer, which is bundled with the Chargeback software. All instances in a cluster must run the same version of Chargeback. A vCenter Chargeback cluster can include up to three vCenter Chargeback servers. Sizing for chargeback instances in a cluster depends on number of simultaneous users.

Load balancing is active/active. Each user request, whether it comes from user interface or API, routes through the load balancer. The load balancer forwards the request to a vCenter Chargeback instance in the cluster based on the number of requests currently serviced by each instance in the cluster. With multiple instances, vCenter Chargeback also load-balances the report processing load by leveraging the internal Quartz scheduler. If the load balancer service goes down, you can restart the service in Windows. The built-in load balancer cannot be replaced with a third-party load balancer. All vCenter Chargeback instances in a cluster connect to the same vCenter Chargeback database.

If the load balancer service becomes unavailable, the vCenter Chargeback application will not work. If the tomcat server on a cluster instance dies, the load balancer redirects requests to other cluster instances.

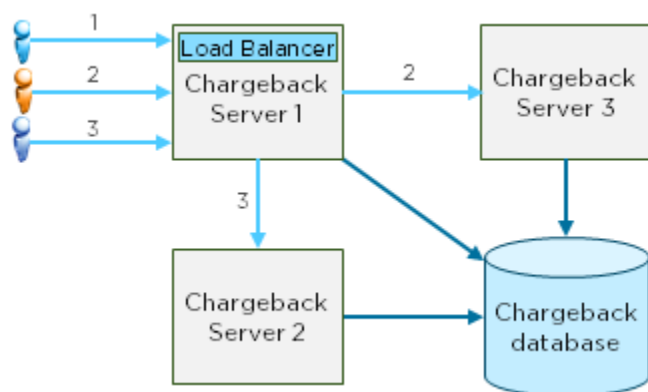
For a load balanced session, *stickiness* is enabled. Therefore, the session always sticks to one vCenter Chargeback server. If there are multiple sessions, then the following algorithm is used:

1. The load balancer uses number of requests to find the best worker.
2. Access is distributed according to the lbfactor (it is the same for all the servers in the cluster) in a sliding time window.

For more information, see *The Apache Tomcat Reference Guide* (<http://tomcat.apache.org/connectors-doc/reference/workers.html>) for the following properties:

- sticky_session = 1 (true)
- method=R

Figure 32. vCenter Chargeback Clustering



Multiple Chargeback environments (separate Chargeback Server and Database) can point to a single vCloud Director database, but this increases the load on the VCD database.

The vCenter Chargeback Database stores organization hierarchies, cost/rate plans, and global Chargeback configuration data. Supported databases include Microsoft SQL Express, Microsoft SQL Server, and Oracle.

6.1.2 Data Collectors

vCenter Chargeback integration with vCloud Director is handled through data collectors:

- Chargeback data collector – Connects to vCenter Server to gather virtual machine metrics. Add all vCenter Servers imported into vCloud Director to vCenter Chargeback in order to see virtual machine-level details. Virtual machines are absent in the vCloud hierarchies until their respective vCenter Servers are registered with vCenter Chargeback.
- vCloud data collector – Connects to the vCloud Director database and monitors all VCD chargeback-related events. The vCloud data collector populates the Chargeback database with vCloud hierarchies and allocation unit information.
- vShield Manager data collector – Connects to vCloud-associated vShield Managers to collect network statistics for networks included in vCloud hierarchy.

Install additional vCloud Director or vShield Manager data collectors on separate servers for increased availability. Multiple data collectors act in an active/passive manner. When one instance goes down, the other instance takes ownership and starts processing. A Chargeback environment may have multiple vCloud data collectors, but can only connect to one vCloud Director instance.

6.1.3 User Roles

The default superuser role has access to entire Chargeback application. The administrator role has access and permissions to resources that are assigned by the super user. Similarly, users created in less privileged roles by administrators are visible only by those administrators. For example, administrator A1 does not have access to users created by administrator A2. With this in mind, administrators should carefully create and assign roles and privileges. This extends to LDAP users and groups as well.

6.2 Maximums

Constraint	Limit	Explanation
vCenter Servers in a Chargeback system	10	The maximum number of vCenter Servers supported by a single Chargeback system.
vCenter Servers per data collector	5	The maximum of vCenter Servers supported by a single Chargeback data collector.
Virtual machines per data collector	15000	Number of virtual machines supported by a single Chargeback data collector.
Virtual machines/entities in a Chargeback system	35000	The maximum number of entities per Chargeback system.
Virtual machines/entities per hierarchy	1000	The maximum number of entities per Chargeback hierarchy.
Hierarchies in a Chargeback system	5000	The maximum number of hierarchies per Chargeback system.
Concurrent reports (~3000 pages) per Chargeback system	5	The maximum number of concurrent reports per Chargeback system.

6.3 Cost Calculation

To track resource metrics for vCloud entities, vCenter Chargeback sets allocation units on vCloud hierarchies based on the parameters of the allocation model configured in vCloud Director. Allocation units are variables associated with chargeback metrics that represent the allocated size of the resource. The table below shows which allocations units are set.

Table 15. vCloud Hierarchy Allocation Units

Entity	Pay-as-you-Go	Allocation Pool	Reservation Pool
Organization virtual datacenter	None	CPU	CPU
		Memory	Memory
		Storage	Storage
vApp	None	None	None
Virtual machine	vCPU	vCPU	vCPU
	Memory	Memory	Memory
	Storage	Storage	Storage
Template	Storage	Storage	Storage
Media file	Storage	Storage	Storage
Network	DHCP	DHCP	DHCP
	NAT	NAT	NAT
	Firewall	Firewall	Firewall
	Count of Networks	Count of Networks	Count of Networks

6.3.1 Cost Models

Installing vCloud and vShield Manager data collectors also creates default cost models and billing policies that integrate with vCloud Director and vShield Manager. Billing policies control costs assessed to resources used. Default vCloud billing policies charge based on allocation for vCPU, memory, and storage. Cost time intervals include hourly, daily, weekly, monthly, quarterly, half-yearly or yearly.

Instead of modifying default billing copies and cost models, make copies and modify the duplicates. For more information, refer to the *vCenter Chargeback User's Guide* (http://www.vmware.com/support/pubs/vcbm_pubs.html) for the version of chargeback that you are using.

Rate factors allow the scaling of base costs for a specific chargeable entity. Example use cases include:

- Promotional rate – A service provider offers new clients a 10% discount. Instead of modifying base rates in the cost model, apply a 0.9 rate factor to reduce the base costs for client by 10%.
- Rates for unique configurations – A service provider decides to charge clients for special infrastructure configurations using a rate factor to scale costs.

VM instance costing assigns a fixed cost to a hard bundle of vCPU and memory. VM instance matrices are linked with a cost model and consist of the VDC selection criteria, a fixed cost table, and a default fixed cost. Selection criteria options include name pattern matching, custom attribute matching, or no criteria. VM Instance uses a stepping function – if there is no entry for a particular VM size, the charge is based on the next largest instance size.

VM instance is only available with the Pay-As-You-Go allocation model. Use VM instance costing to create a fixed cost matrix for different virtual machine bundles.

6.3.2 Reporting

Chargeback can generate cost, usage, and comparison reports for hierarchies and entities. Match the entity or hierarchy with the appropriate cost model when generating reports.

The Chargeback API provides the capability to export reports to XML. Developers can use XSLT to transform the raw XML into a format supported by the customer's billing system. Reports run from the Chargeback user interface are available in PDF and XLS format. Create service accounts with read-only privileges to run reports from the UI or Chargeback API.

7. Orchestration and Extension

The vCloud environment is composed of several components that expose Web Services. A vCloud orchestration platform provides the ability to tie services together into a logical workflow. VMware has different management applications supporting workflow process definition and execution.

- *VMware vCenter Orchestrator* is a technical orchestration authoring platform part of vCenter that enables administrators to automate repetitive tasks by creating workflows that leverage extensive integrations with VMware and third-party vCloud components.
- *VMware Service Manager* is a configurable ITIL platform that features service desk, automated configuration and change management, IT asset management, self-service, and request fulfillment. As part of the service request, it supports a configurable portal using high-level business workflow modeling for approvals, notifications, and tasks integration.

7.1 vCloud API

The vCloud API provides an interface for managing resources in vCloud instances and is the cornerstone to federation and ecosystem support. All current federation tools talk to the vCloud environment through the vCloud API. It is important that a vCloud environment expose the vCloud API to vCloud consumers.

The vCloud API can be used to facilitate communication to vCloud resources using a user interface other than the vCloud Director web console. For example, VMware Service Manager communicates with vCloud Director using the vCloud API.

Currently, vCloud Director is the only software package that exposes the vCloud API. In some environments, vCloud Director is behind another portal or in a location that is not accessible to the vCloud consumer. In this case, use an API proxy or relay to expose the vCloud API to the end consumer.

Due to the value of the vCloud API, some environments may want to meter and charge for API usage. Protecting the vCloud API through audit trails and API inspection is also recommended. Lastly, there are several cases where vCloud providers may want to extend the vCloud API with new features.

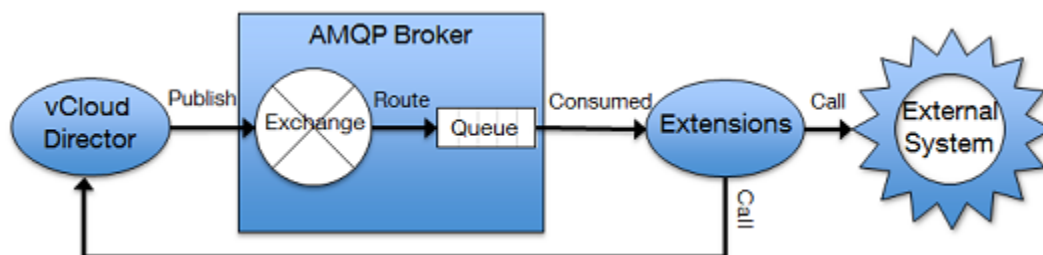
To assist with the vCloud API use cases, the vCloud provider may want to implement an API proxy. The vCloud API is a REST-based service that contains XML payloads. For this reason, any suitable XML gateway can be used to proxy the vCloud API. Several third-party solutions on the market today excel in XML gateway services. VMware collaborates with some of these vendors to develop joint guidance on how to deploy their solutions in a vCloud Director environment. For the latest information on these efforts and collateral, contact your local VMware vCloud specialist.

For more information about the vCloud API and SDKs, visit the developer communities at <http://communities.vmware.com/community/vmtn/developer/forums/vcloudapi>.

7.2 vCloud Messages

vCloud messages is a new feature that introduces the capability to connect vCloud Director with external systems. vCloud Director can be configured to post notifications or messages to AMQP-based enterprise messaging brokers. vCloud messages provide visibility through non-blocking and blocking notifications allowing for end-to-end integration.

Figure 33. vCloud Messages



7.2.1 Message publication

The system administrator can configure vCloud Director to enable the publication of messages for all event notifications and/or for specific blocking tasks:

- Notifications are published on user-initiated events (for example, creation, deployment, and deletion of a vApp) as well as system-initiated events (for example, vApp lease expiration) containing the new state of the corresponding vCloud Director entity.
- Blocking tasks suspend long running operations started as a task before publishing messages and wait until a system administrator approves or rejects the request.

Message publication is enabled for operations started in the vCloud Director GUI or vCloud API.

vCloud Director publishes notification messages to an Advanced Message Queuing Protocol (AMQP) exchange (AMQP version 0.9.1 supported by RabbitMQ version 2.0 and later).

7.2.2 Routing

The AMQP broker uses routing as an effective way to filter vCloud notification messages and dispatch them to different queues for one or multiple extensions.

The exchange routes notifications to its bound queues according to their queue routing key and exchange type. The vCloud notification messages routing key has the following syntax format:

```
<operationSuccess>.<entityUUID>.<orgUUID>.<userUUID>.<subType1>.<subType2>...
<subTypeN>.[taskName]
```

7.2.3 Extension

An extension is a script or an application with the following capabilities:

- Subscribe to an AMQP queue for receiving new messages.
- Triage the received messages.
- Process messages into operations (internal or external calls).
- Call the vCloud API to get more information on the objects involved in an operation and take action on blocked task.

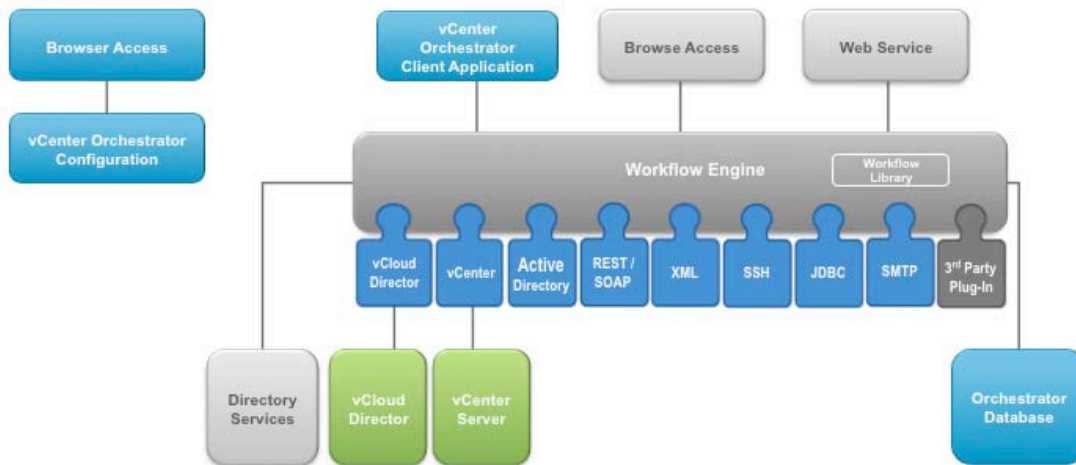
7.3 vCenter Orchestrator

vCenter Orchestrator (vCO) is a system for assembling operational workflows. The primary benefit of vCenter Orchestrator is to coordinate multiple systems to achieve a composite operation that would have otherwise taken several individual operations on different systems. In general, if an operation uses only one underlying system, consider providing direct access to that system for efficiency and reduction of complexity. In a vCloud environment, vCenter Orchestrator can automate highly repetitive tasks to avoid manual work and potential errors.

vCenter Orchestrator consists of the following applications:

- *vCenter Orchestrator Client* enables the workflow developer to author, assemble, test, package workflows, actions, policies, resources, and configurations.
- *vCenter Orchestrator Server Web configuration* is an independent application that runs side-by-side with a web front-end that allows administrators to configure the vCO Server and its plug-ins, and perform maintenance operations.
- *vCenter Orchestrator Server* is the runtime orchestration service, including its interfaces and its pluggable adapters.

Figure 34. vCenter Orchestrator Architecture



vCenter Orchestrator has a plug-in framework and plug-ins are available for vCenter Server, vCloud Director, and vCenter Chargeback. This enables vCenter Orchestrator to orchestrate workflows at the VIM API, VIX API, vCloud API, and Chargeback API levels.

There are three main categories of orchestration use cases:

- vCloud administration operations
- Organization administration operations
- Organization consumer operations

7.3.1 Design Considerations

Depending on the overall architecture and how orchestration is leveraged, orchestrating a vCloud can require one or more vCenter Orchestrator servers. vCenter Orchestrator manages vCloud Director and vCenter using their web services.

vCenter Orchestrator can manage a variable number of hosts per plug-in. Actual limits are subject to a number of determining factors such as bandwidth, number of objects, and concurrent workflows. For example, a single vCenter Orchestrator can manage:

- Multiple vCloud Director hosts
- Multiple vCenter hosts
- Multiple other host types (UCSM, REST, SOAP, VUM)

Note Plug-ins designed for a given version are designed to work for the same version of the product. If managing a mix of host versions, keep the plug-in version at the earliest common version to leverage backward compatibility of the product (that is, use plug-in 1.0.1 if managing a mixed VCD 1.0.1 and VCD 1.5 environment). Avoid mixing host versions where possible—if versions are mixed the operations need to be thoroughly tested. Using the latest version of a plug-in to support an older version of the product is not supported.

Multiple vCenter Orchestrator servers can manage:

- The same vCloud Director host (or load balanced cells)
- The same vCenter server

vCloud Director uses a stateless RESTful web service. There is no session maintained between vCloud Director and vCenter Orchestrator. This minimizes resource usage on both servers. When updates are needed (for example, when starting a workflow using vCloud Director objects), the resources used are proportional to the number of objects updated. This means sending several HTTP GET/PUT/POST/DELETE requests to the vCloud Director server and, on getting replies, creating or updating objects in vCenter Orchestrator. Using multiple sessions (*Per user* mode in the plug-in configuration) multiplies the number of objects. vCloud Director can be load balanced to avoid having a single point of failure and using too many resources on a single cell.

vCenter uses a stateful SOAP web service that supports a very large service definition and advanced mechanisms, such as a notification service, that are extensively used by vCenter Orchestrator. Sessions are maintained between vCenter and vCenter Orchestrator all the time. This has an important impact on resource consumption on both servers even when there is no workflow activity.

The session activity and associated resource consumption on both servers is proportional to the number of objects loaded in the vCenter Orchestrator vCenter inventory that multiply the number of sessions opened. For this reason, configure the vCenter plug-in using a shared session instead of a session per user, and limit the number of vCenter Orchestrator servers that manage a single vCenter. Workflow activity also consumes resources for objects that are not in the inventory cache.

Additional considerations:

- Although the *vCenter Orchestrator Installation and Configuration Guide* mentions handling a maximum of 10 vCenter Servers, 100 ESXi hosts, and 3000 virtual machines, these are not hard limitations. The 3000 virtual machine number reflects intensive, concurrent workflow tasks that may not reflect typical usage.
- If a vCenter *Orchestrator* is overloaded by a high level of objects to manage, first attempt to tune the server for higher scalability. Alternatively, design the solution to use different vCenter Orchestrators that manage different vCenter Servers, or connect to a large vCenter using different vCenter Orchestrators that are configured with accounts to access different zones of vCenter.

7.3.2 Scalability

When configuring vCenter Orchestrator to run a high number of concurrent workflows, it is necessary to understand how the Orchestration engine works.

The vCenter Orchestrator Workflow Engine default configuration allows for running up to 300 concurrent workflows. When the running queue exceeds this number, the workflows are placed in an execution queue and moved back to the running queue as soon as one or more workflows has completed its execution run. Completed workflow states may be “completed successfully,” “failed”, “canceled” or passivated (“waiting-for-signal” state). The execution queue has a default size of 10000 workflows. After the execution queue size is exceeded, the workflow engine marks subsequent workflows as failed.

A running workflow consumes at least one running thread (either running the workflow or updating the workflow state) and from 1MB to a few MB of memory (varies depending on the number of enabled plug-ins and plug-in objects). Limiting the number of workflows makes sure that threads and memory can be allocated, with the maximum depending on the JVM settings, the operating system, and the underlying hardware.

You can change the default value by changing the following properties in the `Orchestrator\appserver\server\vm\conf\vm.properties` configuration file:

- `com.vmware.vco.workflow-engine.executors-count`
- `com.vmware.vco.workflow-engine.executors-max-queue-size`

Note VMware recommends following the instructions in the rest of this document before increasing the default settings for the concurrent workflows, because it requires expanding the resources for the vCenter Orchestrator Java Virtual Machine, the host operating system, the host virtual machine, and possibly the vCenter Orchestrator Database.

Each active plug-in has an impact on the workflow engine performance. A plug-in loads classes, runs update threads, logs information to disk, provides objects to the scripting engine, and maintains the inventory. Even if the plug-in is unused, it consumes unnecessary resources and increases the memory footprint of each running workflow. Disable all the plug-ins that are not in use to increase the workflow engine capacity.

7.3.3 Workflow Design

Workflow design impacts duration and usage of resources. Best practices are:

- Effective scripting – Use scripting development best practices to avoid unnecessary highly resource-demanding operations such as active wait loops, repetitive expensive calls to the same resources, and ineffective algorithms. Perform extensive testing on a vCO test server before running new or updated workflows on a production system.
- Workflow threading control – Having many distinct running workflows increases the amount of resources are used.
- Workflows started individually and workflows started using the `Asynchronous workflow` or `Nested workflow` palette elements run in different workflow instances.
- A sub-workflow in a master workflow is still running within the same workflow instance, but using fewer resources. It is recommended to link workflows in higher-level workflows instead of calling individual workflows in sequence.
- Reducing the number of workflows waiting – If the reason for the high concurrency is due to a high number of workflows waiting on external systems, there are methods to avoid consuming resources while waiting:
 - The `Wait Until date` workflow palette element and the `System.Sleep()` methods keep the workflow in a running state in the execution queue. Even if the thread is in Sleep mode, it still consumes memory. For long running workflows, these can be replaced by the `Waiting timer` or `Waiting Event` palette elements. Using one of these elements passivates the workflow execution and saves its state in the vCenter Orchestrator database. The workflow is then removed from the running queue and memory is freed. The vCloud Director library long running workflows make extensive use of `Waiting Event`.
- When workflow activity needs to be suspended until a determined time, programmatically schedule a workflow task.

Though saving active resources, each passivation and activation consumes CPU resources and database access. Best practices for using the `Waiting Timer` or `Waiting Event` are:

- Do not trigger a large number of these at the same time.
- Do not set very short timers in loops.

7.3.4 Solution Guidelines

In addition to the server configuration and the workflow design, you must have a well-controlled overall solution that includes the upper management layers and the orchestrated systems.

- Misuse of orchestration – An orchestration engine provides automation and integration to manage complex cross-domain processes. It provides several facilities for versatility, resiliency, and auditing that are excessive for simple operations that do not require this level of service. vCenter Orchestrator should not be used to replace single calls to the vCloud Director API.
- Control of the workflows – The systems calling a vCenter Orchestrator should have a workflow throttling mechanism adjusted according to vCO-tested maximums to avoid resource starvation.
- Load balancing – When going over the maximums, it may be necessary to design the system to load balance the workflows across different vCenter Orchestrator servers.

- Orchestrated systems bottleneck – vCenter orchestrator workflows should have logic that prevents starting too many operations at once on the orchestrated systems. Design this logic to support the defined load. The parameters that have an influence on the started workload should be exposed as configuration elements to be adjusted by the Orchestration Administrator (a parameter that determines the number of vApp clones to be processed in parallel).

7.3.5 Orchestrator Client

The vCenter Orchestrator Server has a client application that is used to develop workflows and actions. During server installation, install the client on the same system as the server to have a client version available that matches the server. In production environments, this local installation of the client software is only used in emergency cases where a matching client is not available via developers' workstations.

Developers should install the client on their workstations for daily development purposes. This allows developers to connect to their Test/Dev servers as needed. Only use the client to connect to servers on the same LAN. If connecting to a remote server, use remote desktop to run the client from the same LAN.

7.3.6 vCloud Director Plug-in

When specifying the **Host** field of the plug-in, the value specified must be exactly what is specified by the vCloud Director server. This value is determined as follows:

1. If a value is specified under the vCloud Director *Administration – Public Addresses – External REST API Base URI*, use this value in the plug-in configuration. For example, using a load balanced vCloud Director requires changing the public address to the one specified for the virtual server in the load balancer configuration. Forward and reverse DNS should be functional for the address specified.
2. If a hostname or fully qualified domain name is specified, make sure that forward and reverse DNS is functional and use that name in the plug-in configuration.
3. If no hostname is specified and the vCloud Director server is only configured to use an IP address, use the same IP address for the plug-in configuration.

Note Failure to configure the plug-in as specified will result in undesired effects.

After specifying the **Host** field, choose a strategy for managing the user logins. The possible options are **share a unique session** and **per user session**.

- When **Share a unique session** is configured, a single session is created between vCenter Orchestrator and vCloud Director based on the configured organization and credentials. The vCenter Orchestrator user inherits the rights of those credentials for any workflow executed. From an auditing perspective, a shared session shifts the auditing responsibility from vCloud Director to vCenter Orchestrator. The workflows developed for such integration need to have an appropriate level of logging set up to meet the organization's audit requirements.
- When **Session per user** is configured, the user authenticated in vCenter Orchestrator is used to authenticate in vCloud Director. This creates for each user a session between vCenter Orchestrator and vCloud Director that is associated with an inventory based on this user role and permissions. This requires having the organization use an LDAP host synchronized with the LDAP host configured in vCO.

Notes

- In the case of organizations that use different LDAP hosts, one dedicated instance of vCO is required per organization.
- Multiple sessions can strain CPU, memory and bandwidth.

In addition, setting an organization is required. This organization defines the scope of the operations that vCenter Orchestrator can perform.

- **SYSTEM** is set when requiring create, read, update, and delete access to all organizations and to their associated virtual infrastructure resources.
- A specific organization is set when requiring create, read, update, and delete access to all elements that belong to a given organization.

The most common use cases of the plug-in usually correspond to one of the following scenarios:

- As a (public or private) vCloud provider using a vCenter Orchestrator server as part of the vCloud management cluster:
 - Tasks such as managing provider resources and on-boarding new organizations require system level administrative permission to vCloud Director. This scenario uses a **Share a unique session**, an organization set to **SYSTEM**, and the system administrator credentials.
 - Use **Session per user** if the administrative tasks require different roles and permissions. In this case, the SYSTEM organization has to be setup to synchronize with the vCloud provider LDAP host configured with vCenter Orchestrator.

Note If configuring more than one vCloud Director connection, using a combination of Shared Session and Per user session grants vCO workflows users the shared access session permissions for the configured organization. For example, if the plug-in is set with a System Shared Session and there is a requirement to grant vCO users access to a given organization, both connections should use **Session per user** and permissions should be set differently to avoid all users to have wide access to all organizations.

- As a public vCloud tenant of one organization or more, using vCenter Orchestrator in the tenant premise or as part of the organization vApps:
 - For organization administrative tasks, use **Share a unique session** with organization administrator credentials. If administering more than one organization, one new vCloud Director Connection can be added per organization.
 - Configure the plug-in as **Session per user** for delegating workflows operations tasks that are not covered by the vCloud Director interface to organization users having different roles and permissions. In this configuration, set up the organization to synchronize with the tenant LDAP host configured in vCenter Orchestrator.
- As a private vCloud organization tenant using a vCenter Orchestrator server as part of the vCloud management cluster, and a single LDAP host.
 - The vCloud provider configures a new connection using this specific organization and **Session per user**. Set up the organization to synchronize with the LDAP host that is configured with vCenter Orchestrator. All other organizations configured in other connections also synchronize with the same LDAP HOST server.

7.4 vCenter Orchestrator Examples

Orchestration brings automation to vCloud administration, organization administration, and self-service consumer operations.

7.4.1 vCloud Administration Orchestration Examples

The following examples highlight the value of vCenter Orchestrator to the vCloud system administrator. The use case focuses on infrastructure management and the resource provisioning process.

- A provider wants to onboard a new customer. The main steps are to add a new organization, users (possibly from LDAP), networks, virtual datacenters, and catalogs. The provider may also want to schedule a recurring chargeback report for billing, and send an email notification to the tenant advising them that their vCloud environment is ready.
- Another example is when a tenant requests additional external network capacity. The provider wants to automate the creation of the network, which includes name generation, identification, and allocation of available VLAN and IP address range; configuration of the network switch and vCloud perimeter firewall, creation of the external network in vCenter, and finally, allocation of the external network to the tenant's organization.

7.4.2 Organization Administration Orchestration Examples

Operational tasks within the tenant's organization can benefit from automation as well. Typically, these tasks address the vApp lifecycle management, such as vApp creation, configuration, maintenance, and decommissioning.

- Consider the case of virtual machine creation in an environment using Active Directory to identify services such as authentication and printing. After deployment, it is required that the virtual machine join the Active Directory domain. Usually, it is preferable to use an organization unit (OU) other than the default `Computers` container. vCenter Orchestrator can create the virtual machine's computer account in the proper OU prior to virtual machine deployment so that the computer account name is unique and residing in the proper OU. Similarly, when the virtual machine is decommissioned, vCenter Orchestrator can remove the entry in the OU as part of the same workflow.
- Another example is the case where an organization administrator wants to manage recurring updates to a software package or configuration element across several virtual machines in a single operation. A workflow could accept a list of systems and a source for the software or configuration as parameters, then perform the update on each system.

7.4.3 Cloud Consumer Operation Orchestration Examples

Cloud consumer operations are tasks that the organization administrator wants to offload as a self-service operation. Performing the operation as a vCenter Orchestrator workflow provides an easy way to expose the operation to a customer via the built-in portal or a customized portal leveraging the web-services API. Many operations in this category can be satisfied directly via the vCloud Director web console; however, some operations affect multiple systems or fit better into a customer portal. These operations are a natural candidate for an orchestration workflow. vCloud consumers do not have visibility into orchestration components, which makes it somewhat difficult. The vCloud provider must initiate the workflow using the vCenter Orchestrator Client unless the provider creates a portal to front-end vCenter Orchestrator.

Some examples of these types of use cases include resetting of user account passwords on virtual machines using the VIX plug-in, placing a load balanced service into maintenance mode (stopping the service, removing it from the load balancing pool, and disabling monitors), loading certificates into virtual machines, and deploying instances of custom applications from the organization's catalog.

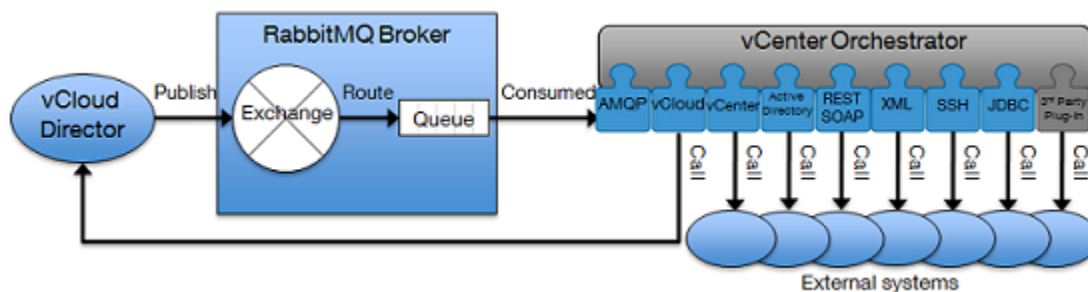
vCenter Orchestrator can be used to create custom workflows at the vCloud API and vSphere levels. VMware Service Manager has built-in workflow functionality that integrates with vCloud Director through the vCloud API and is an alternative to vCenter Orchestrator.

See the vCenter Orchestrator documentation set for additional information on vCO installation, configuration, and workflow development: http://www.vmware.com/support/pubs/orchestrator_pubs.html.

7.4.4 Using Orchestrator as a VCD extension

vCenter Orchestrator fully supports consuming blocked tasks and notifications messages, callbacks, and calls to external systems via the vCloud Director, AMQP, and other specific product plug-ins.

Figure 35. vCenter Orchestrator as a vCloud Director Extension



The AMQP plug-in comes with workflows, and requires a onetime setup. Provide values for the following:

- **Add a broker** – Add an AMQP broker by providing hostname and credentials.
- **Declare an exchange** – Declare an exchange for the configured broker.
- **Declare a queue** – Declare a queue.
- **Bind** – Bind a queue to an exchange by providing a routing key.
- **Subscribe to queues** – Allow vCO to receive message updates on new messages.

Restarting the vCO server automatically saves and reloads the configuration.

The plug-in supports adding a policy element of type `subscription` having an `onMessage` trigger event. A policy can be setup to start a workflow processing new messages.

Workflows are provided to triage and process the message to output vCloud Director objects. These can provide all of the information necessary for audit purposes and for designing custom logic before calling external systems. There are two ways to call external systems:

- Specific vCenter Orchestrator plug-ins adapters such as vCloud Director, vCenter, Update Manager, and Active Directory.
- Generic plug-ins adapters such as REST, SOAP, XML, SSH, and JDBC.

vCloud Director Workflows can abort, resume, or fail blocked task objects. See *Operating a VMware vCloud* for example workflows using vCloud messages.

8. Multi-Site Considerations

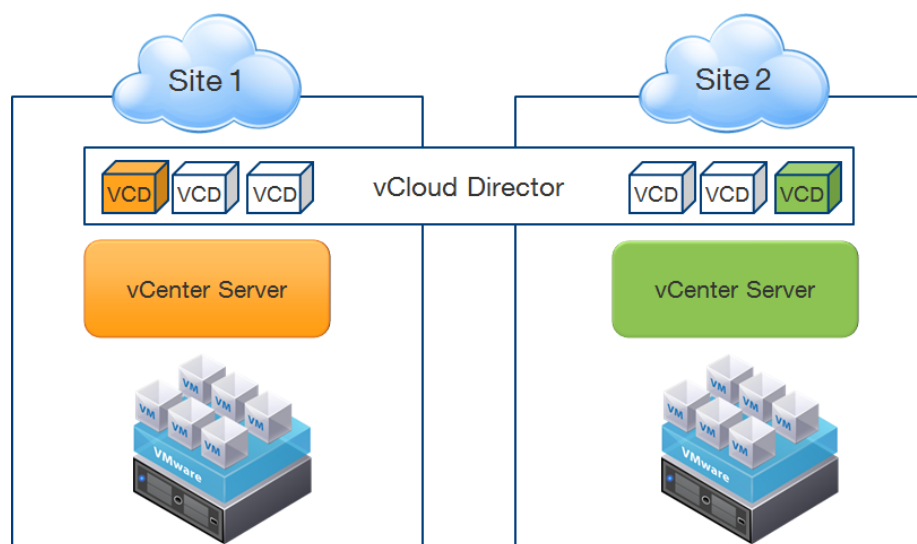
vCloud Director is neither designed nor supported for multi-site deployments in the current version of the product due to potential issues with network latency and reliability. In an environment with multiple sites, design each site as separate vCloud instances with considerations for future interconnectivity instead of having a single vCloud that spans the sites.

Multi-site means different things to different audiences. Some providers would like to have a single interface that encompasses all of their sites. Other providers do not mind having multiple interfaces, but would like to have the same services available in each location.

8.1.1 Scenario #1 – Common User Interface

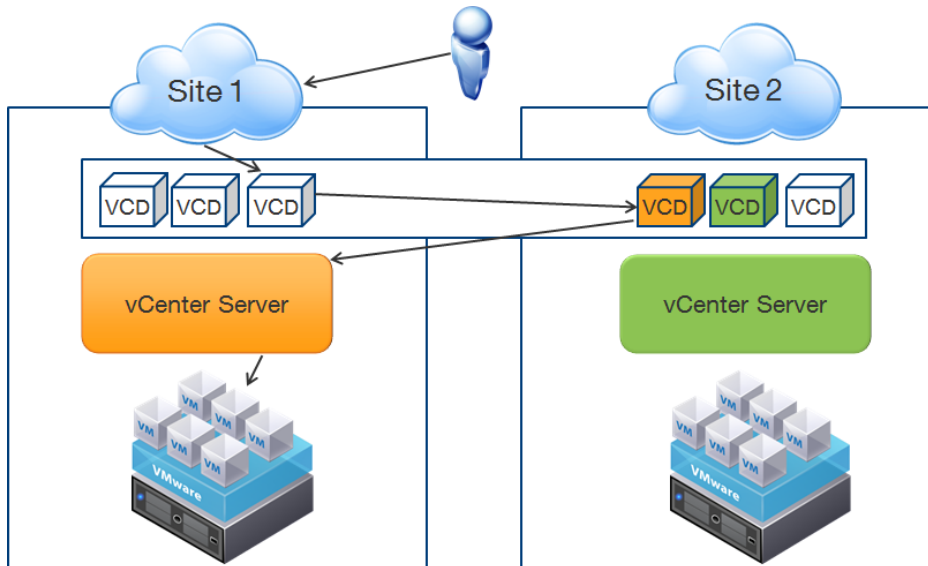
Scenario 1 depicts a use case where one vCloud Director instance supports two locations. vCloud Director cells provide the web console and can sit in either or in both locations. As illustrated in Figure 36, one vCloud Director cell serves as the proxy for a vCenter Server in one of the sites.

Figure 36. Two Sites with Local VCD Instances Managing Two Local vCenter Servers



The local vCenter Servers control resources local to each site. This seems like a very logical infrastructure setup until you examine some of the user flows.

If a user comes in through site #1 and requests remote console access to a virtual machine in site #1, all traffic is *not* guaranteed to stay in site #1. This is because it is not possible to control which vCloud Director cell acts as the proxy for a particular vCenter Server. A user could initiate a session through a vCloud Director cell in site #1, which then communicates to the proxy for vCenter server #1 in site #2. That vCloud Director cell talks back to the vCenter server in site #1 to finish establishing the remote console connection to the local ESXi host hosting the workload in site #1. Traffic flows through the vCloud Director cell that initiated the request in site #1. Figure 37 illustrates the flow of events.

Figure 37. Remote Console Flow

Another problem with this setup is controlling the vCloud Director cell that a user is terminated on based on virtual machine and site-specific data. It is nearly impossible to figure this out and provide that logic to a load balancer. In addition, a centralized vCloud Director database is needed to support all vCloud Director cells from both sites. This creates even more traffic on the link between the two sites because the message bus in vCloud Director uses the vCloud Director database for communication. Overall, this solution is less than optimal for most use cases, with the exception of cross-campus multi-site configurations where site-to-site communication will not overwhelm the network and where network availability is highly reliable.

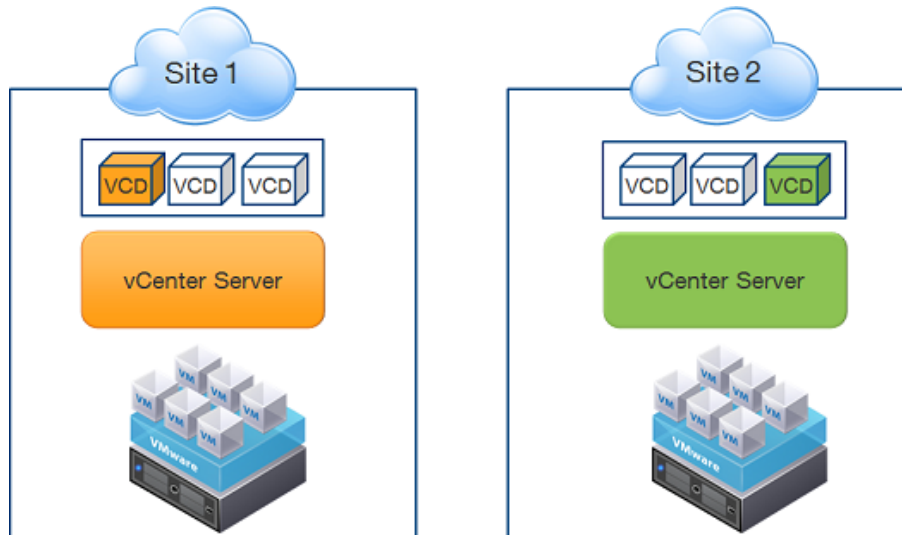
8.1.2 Scenario #2 – Common Set of Services

A more pragmatic approach to multi-site setups is to configure isolated vCloud instances at each site. This solves the network cross-talk issue, but it introduces other problems. For example, how do you provide a common set of services across the different sites? How do you keep organization names and rights as well as catalogs, networks, storage, and other information common across the different sites? Currently, there is no mechanism to do this in vCloud Director. Using other VMware technologies included in the vSphere suite of products, it is possible to synchronize vCloud deployments using automation scripts to provide common sets of services across locations.

In an enterprise, a private vCloud maps to a single site. Multiple vCloud instances can be connected using vCloud Connector for offline vApp migrations.

8.1.3 Suggested Deployment

VMware does not officially support multi-site deployments at this time. The recommended approach is to set up an isolated vCloud Director instance in each location. As illustrated in Figure 38, this isolated vCloud Director instance includes local vCloud Director cells, vShield Manager instances, vCenter Servers, a vCloud Director database instance, a vCenter Chargeback instance, and local vSphere resources.

Figure 38. Two Sites with Isolated vCloud Director Instances

To keep the sites synchronized with organization and resource information, create a set of onboarding scripts and workflows. Use these workflows when creating a new organization or a new resource for an organization to drive creation across all vCloud sites. VMware Professional Services can assist with the creation of customer-specific workflows based on using vCenter Orchestrator to keep multiple vCloud instances synchronized with organization resources.

8.1.4 Other Multi-Site Considerations

When creating multi-site configurations, consider the physical resources outside of the vCloud environment. How do you set up networking between the sites? How is IP addressing handled between sites? Are stretched L2 networks an option? Guidelines for these issues are currently beyond the scope of the VMware vCloud Architecture Toolkit.

8.1.5 Merging Chargeback Reports

In our reference multi-site setup, two vCenter Chargeback instances were included. To provide one common bill or usage report to the consumer, aggregate all associated chargeback reports into one report. Leverage the vCenter Chargeback API and vCenter Orchestrator to periodically pull chargeback reports from each vCenter Chargeback server and consolidate them into one master report.

8.1.6 Synchronizing Catalogs

Synchronizing catalogs between sites is a time-consuming task. When setting up multiple vCloud sites designate one site as the master site for vApp template creation, and designate all other sites to be replication peers. If possible, leverage native storage array replication to replicate the vApp template storage in each catalog. Array replication can provide several benefits for long distance data movement, including data de-duplication and compression. After synchronizing the data, leverage the catalog synchronization workflows provided by VMware vCloud API to import the replicated templates into the appropriate catalogs in vCloud Director.

Synchronizing templates added at remote sites is out of scope for this version of the reference architecture. VMware Professional Services can assist with the creation of these workflows.

9. Hybrid vCloud Considerations

A hybrid vCloud incorporates a combination of vCloud instances and may include both on-premise and off-premise resources—applications can be located on-premise, off-premise, or a combination of both. Enterprises with an existing private vCloud may choose to provide and manage public vCloud resources in a secure and scalable way. Connectivity between different vCloud instances that enables data and application portability indicates a hybrid vCloud solution.

Figure 39. Hybrid vCloud Example



9.1 vCloud Connector Considerations

With the emergence of cloud computing, private enterprises may soon be managing multiple vCloud instances, private and public. The ease of migrating workloads between vCloud instances becomes increasingly important.

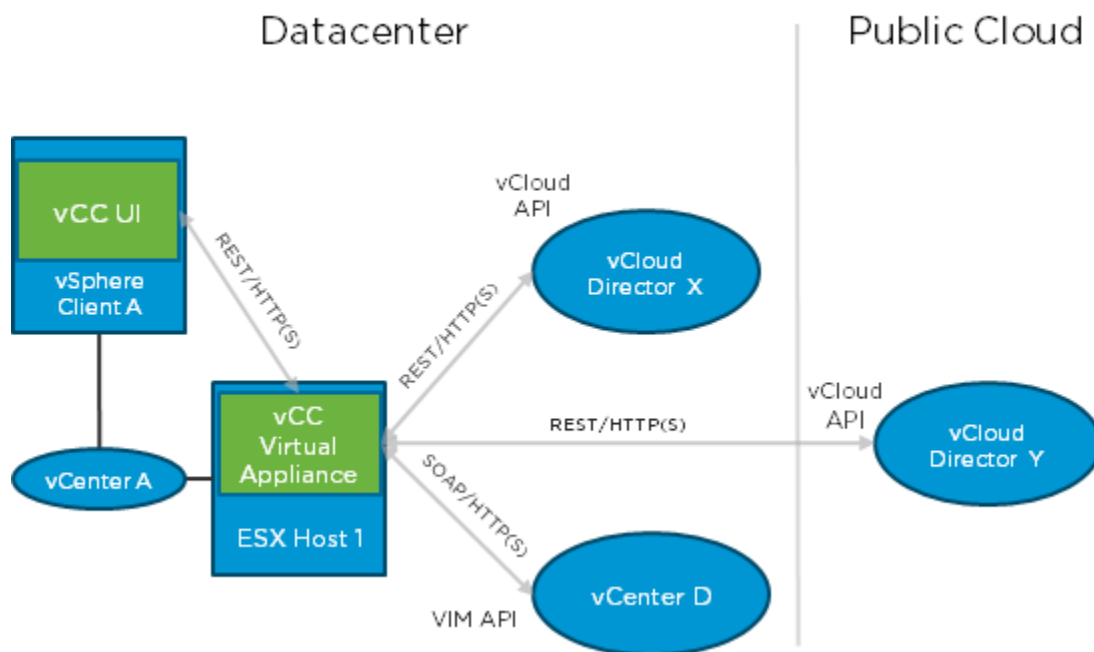
vCloud Connector (vCC) is a vSphere Client plug-in that allows users to connect to vSphere or vCloud Director-based vCloud instances and manage them under a single interface. Through the vCloud Connector single pane of glass view, users can view, copy, and operate workloads across internal datacenters and private or public vCloud instances. vCloud Connector is installed by vCloud administrators, but can be used by both administrators and end users to view and manage workloads. vCloud Connector is delivered as a virtual appliance with the UI instantiated as a vSphere Client plug-in.

9.1.1 vCloud Connector Placement

The following are considerations regarding where to place your vCloud Connector virtual appliance:

- Deploy the virtual appliance to a vCenter Server that can be accessed by target users. The only user access is via the vSphere Client, so apply the appropriate vCenter roles to vCC users.
- Workload copy operations use the vCloud Connector appliance as a middleman so consider network latency and bandwidth between vCloud instances. For some use cases, it may be preferable to run multiple instances of vCloud Connector across multiple vCenter Servers to avoid network latency or consuming excessive bandwidth.

Figure 40. vCloud Connector Architecture



9.1.2 vCloud Connector Example Usage Scenarios

vCloud Connector can support a number of workload migration use cases involving virtual machines, virtual machine templates, vApps, and vApp templates. Migrations are possible between:

- vSphere <--> vCloud
- vSphere <--> vSphere
- vCloud <--> vCloud

9.1.3 vCloud Connector Limitations

vCloud Connector has the following restrictions:

- Currently, there is no way to have predefined clouds appear in vCloud Connector. Each user must manually add all clouds that they intend to access to vCloud Connector. There are no clouds defined by default.
- Traffic to and from the vCloud Connector appliance is not WAN optimized, so migrating workloads over WAN links is not ideal even if sufficient bandwidth exists. Avoid traversing WAN links where possible by installing vCloud Connector appliances in optimal locations. Currently, there is no way to limit which clouds can be added to a vCloud Connector instance, so instruct users to only use the proper vCloud Connector instance for their needs.
- The transfer process caches virtual machines in two different locations. To facilitate successful transfers, size the vCloud Connector staging storage and vCloud Director transfer storage appropriately. The staging storage is 40GB by default, so the largest virtual machine vCloud Connector can transfer is around 40GB.

- vCloud Connector is designed to give you a consistent view of your workloads across multiple clouds and migrate those workloads. vCloud Connector cannot perform all of the operations vCloud Director can handle, so use the vCloud Director web console to manage your workloads.
- All workload transfers are cold migrations. Power off vApps and virtual machines prior to migration. Hot migrations are currently not available. Also, vApp networking configuration needs to be modified before powering on the virtual machines.
- vCloud Connector can handle up to ten concurrent transfers. Subsequent requests are queued. The maximum number of cloud connections for a single vCloud Connector is five (VCD or vSphere).

10. References

Table 16 lists documents you can refer to for additional information.

Table 16. Reference Documentation

Topic	Referenced Document
vCloud Director	<p><i>vCloud Director Security Hardening Guide</i> http://www.vmware.com/files/pdf/techpaper/VMW_10Q3_WP_vCloud_Director_Security.pdf</p> <p>Go to the VMware vCloud Director documentation site for the following vCloud Director documentation (http://www.vmware.com/support/pubs/vcd_pubs.html):</p> <ul style="list-style-type: none"> • <i>vCloud Director Installation and Configuration Guide</i> • <i>vCloud Director Administrator's Guide</i> <p><i>What's New in VMware vCloud Director 1.5 Technical Whitepaper</i> http://www.vmware.com/resources/techresources/10192</p>
vCloud API	<p>Go to the VMware vCloud Director documentation site for the following vCloud Director documentation (http://www.vmware.com/support/pubs/vcd_pubs.html):</p> <ul style="list-style-type: none"> • <i>vCloud API Specification</i> • <i>vCloud API Programming Guide</i>
vSphere	<p>VMware vSphere documentation:</p> <ul style="list-style-type: none"> • VMware vSphere 5 documentation: https://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-pubs.html <p><i>Performance Best Practices for VMware vSphere 5.0</i> http://www.vmware.com/resources/techresources/10199</p>
vShield	<p><i>vShield Administration Guide</i> https://www.vmware.com/support/pubs/vshield_pubs.html</p>
vCenter Chargeback	<ul style="list-style-type: none"> • <i>vCenter Chargeback User's Guide</i> https://www.vmware.com/support/pubs/vcbm_pubs.html • <i>Using vCenter Chargeback with VMware Cloud Director Technical Note</i> http://www.vmware.com/files/pdf/techpaper/vCenterChargeback_v_1_5_Tech_Note.pdf
vCenter Orchestrator (vCO)	<ul style="list-style-type: none"> • <i>vCenter Orchestrator Developer's Guide</i> https://www.vmware.com/pdf/vco_410_developers_guide.pdf • <i>VMware vCenter Orchestrator Administration Guide</i> https://www.vmware.com/pdf/vco_410_admin_guide.pdf • <i>vCenter Server 4.1 Plug-In API Reference for vCenter Orchestrator</i> https://www.vmware.com/support/orchestrator/doc/vco_vsphere41_api/index.html
VMware Service Manager	<ul style="list-style-type: none"> • VMware Service Manager Installation and Configuration Guide • VMware Service Manager User's Guide

Appendix A: Availability Considerations

vCloud availability depends on elimination of single points of failure (SPOF) in the underlying infrastructure, personnel with the appropriate skills being available, and suitable operational processes being in place and followed.

Table 17. vCloud Availability Considerations

Maintaining Running Workload		
Component	Availability	Failure Impact
vSphere ESXi hosts	Configure all ESXi hosts in highly available clusters with a minimum of n+1 redundancy. This provides protection for the customer's virtual machines, the virtual machines hosting the platform portal/management applications, and all of the vShield Edge appliances.	<p>In the event of a failure of a host, vSphere HA detects the failure within 13 seconds and begins to power on the host's virtual machines on other hosts within the cluster.</p> <p>vSphere HA Admission Control makes sure sufficient resources are available in the cluster to restart the virtual machines. The admission control policy Percentage of cluster resources is recommended as it is flexible while guaranteeing resource availability.</p> <p>For a description of best practices about increasing availability and resiliency, see the white paper <i>VMware High Availability: Deployment Best Practices: VMware vSphere 4.1</i> (http://www.vmware.com/files/pdf/techpaper/VMW-Server-WP-BestPractices.pdf.)</p> <p>It is also recommended that vCenter is configured to proactively migrate virtual machines off a host in the event the host's health become unstable. Rules can be defined in vCenter to monitor host system health.</p>
Virtual machine resource consumption	<p>vSphere DRS automatically migrate virtual machines between hosts to balance the cluster and reduce the risk of a "noisy neighbor" virtual machine monopolizing CPU and memory resources within a host at the expense of other virtual machines running on the same host.</p> <p>vSphere Storage I/O Control automatically throttles hosts and virtual machines when detecting I/O contention and preserves fairness of disk shares across virtual machines in a datastore. This makes sure that a noisy neighbor virtual machine does not monopolize storage I/O resources. Storage I/O Control makes sure that each virtual machine receives the resources it is entitled to by leveraging the shares mechanism.</p>	<p>No impact. Virtual machines are automatically migrated between hosts with no downtime by vSphere DRS.</p> <p>No impact. Virtual machines and ESXi hosts are throttled by Storage I/O Control automatically based on their entitlement relative to the amount of shares or the maximum amount of IOPS configured. For more information on Storage I/O Control, see the white paper <i>Storage I/O Control Technical Overview and Considerations for Deployment</i> (http://www.vmware.com/files/pdf/techpaper/VMW-vSphere41-SIOC.pdf).</p>

vSphere ESXi host network connectivity	Configure port groups with a minimum of two physical paths to make sure a single link failure does not impact platform or virtual machine connectivity. This includes management and vMotion networks. The Load Based Teaming mechanism is used to avoid oversubscribed network links.	No impact. Failover occurs with no interruption to service. Configuration of failover and failback as well as corresponding physical settings such as PortFast are a requirement.
vSphere ESXi host storage connectivity	ESXi hosts are configured with a minimum of two physical paths to each LUN or NFS share to make sure a single storage path failure does not result in an impact to service. Path selection plug-in is selected based on the storage vendor's best practices.	No impact. Failover occurs with no interruption to service.

Maintaining Workload Accessibility

Component	Availability	Failure Impact
VMware vCenter Server	vCenter Server runs as a virtual machine and makes use of vCenter Server Heartbeat.	vCenter Server Heartbeat provides a clustered solution for vCenter Server with fully automated failover between nodes, thereby providing near zero downtime.
VMware vCenter Database	VMware vCenter Database resiliency is provided with vCenter Heartbeat if MS SQL is used or Oracle RAC. If using Oracle RAC, refer to the VMware KB article on Oracle RAC configuration.	vCenter Heartbeat or Oracle RAC provides a clustered solution for a vCenter database with fully automated failover between nodes, thereby providing zero downtime.
vCloud component databases (vCloud Director and Chargeback)	VMware vCloud component database resiliency is provided through database clustering. If using Oracle RAC, refer to the VMware KB article on Oracle RAC configuration.	Oracle RAC supports the resiliency of the vCloud Director and Chargeback databases as it maintains vCloud Director state information and the critical Chargeback data required for customer billing respectively. While not required for maintaining workload accessibility, clustering the Chargeback database makes sure providers can accurately produce customer billing information.
VMware vCenter Chargeback	Multiple Chargeback, vCloud, and vShield Manager data collectors are installed for active/passive protection.	In the event that one of the data collectors goes offline, the other picks up the load such that transactions are captured by vCenter Chargeback.

vCloud Infrastructure Protection		
Component	Availability	Failure Impact
vShield Manager	<p>vShield Manager can receive the additional protection of vSphere FT, resulting in continuous protection in the event of a host failure.</p> <p>VM Monitoring is enabled on a cluster level within HA and uses the VMware Tools heartbeat to verify that virtual machines are alive. When a virtual machine fails and the VMware Tools heartbeat is not updated, VM Monitoring verifies if any storage or networking I/O has occurred over the last 120 seconds before restarting the virtual machine.</p> <p>It is highly recommended to configure scheduled backups of vShield Manager to an external FTP or SFTP server.</p>	<p>Infrastructure availability is impacted, but service availability is not. vShield Edge devices continue to run without the management control, but no additional edge appliances or modifications can occur until the service comes back online.</p>
vCenter Chargeback	<p>vCenter Chargeback virtual machines can be deployed in a cluster configuration. Multiple Chargeback data collectors can be deployed to avoid a single point of failure.</p>	<p>There is no impact on Infrastructure availability or customer virtual machines. However, it is important to keep Chargeback available to preserve all resource metering data.</p> <p>Clustering the vCenter Chargeback servers makes sure providers can accurately produce customer billing information and usage reports.</p>
vCloud Director	<p>The vCloud Director cell virtual machines are deployed as a load balanced, highly available clustered pair in an N+1 redundancy set up, with the option to scale out when needed.</p>	<p>Session state of users connected via the portal to failed instance is lost. Users can to reconnect immediately.</p> <p>No impact to customer virtual machines.</p>
vShield Edge	<p>vShield Edge can be deployed through the API and vCloud Director web console. To provide network reliability, VM Monitoring is enabled. In case of a vShield Edge guest OS failure, VM Monitoring restarts the vShield Edge device. vShield Edge appliances have a custom version of VMware Tools and thus are not monitored by vSphere HA guest OS monitoring.</p>	<p>Partial temporary loss of service. vShield Edge is a possible connection into organization.</p> <p>No impact to customer virtual machines or Virtual Machine Remote Console (VMRC) access.</p> <p>All external network routed connectivity is lost if a the corresponding vShield Edge appliance is lost.</p>

vCenter Orchestrator	<p>Plan for high availability of all systems involved in the orchestration workflow. Design the workflows to remediate the non-availability of orchestrated systems (for example, by alerting and retrying periodically).</p> <p>High availability for vCO can be provided by vSphere HA and vSphere FT in addition to application-based clustering.</p> <p>As long as a copy of the database is available, a vCenter Orchestrator Application Server with the appropriate configuration can resume workflow operations. An active-passive node configuration best suits vCenter Orchestrator.</p>	<p>Temporary loss of access to end users interacting directly with vCenter Orchestrator.</p> <p>Disruption to workflows executed by vCenter Orchestrator. This includes workflows started by vCenter Orchestrator and workflows started by external applications.</p>
----------------------	--	---

vCloud Director Cell Load Balancing

A load balanced, multicell vCloud Director architecture provides the following benefits:

- Scalability, by distributing session load across cells.
- Improved availability by monitoring cell server health and adding or removing cells from service based on status.
- Enables non-disruptive operating system patching and maintenance of the cell servers.
- Reduced impact to vCloud Director application upgrades.

Load balancing improves scalability in the following areas:

- Number of concurrent operations.
- Number of active and concurrent console sessions via the console proxy service.
- Number of concurrent users.
- Number of vCenter Server operations (in the case that multiple vCenter servers are attached to the vCloud Director instance).

Table 18 highlights the design considerations for load balancing of vCloud Director cells.

Table 18. Load Balancer Considerations

Consideration	Detail
Security	<p>A front-end firewall is typically deployed before the load balancer. In some environments additional firewalls may be located between vCloud Director cells and the resource tiers managed by vCenter.</p> <p>Load balancers may also provide NAT/SNAT (source network address translation) and is typically configured to provide this for the clustered cells.</p> <p>VMware recommends that access be secured between cells and the other management and resource group components. Refer to the <i>vCloud Director Installation and Configuration Guide</i> for ports that must be opened.</p>
Single vCloud Director site and scope	This architecture covers load balancing of a single vCloud Director site or instance. It does not cover client application load balancing or global load balancing.
Sizing recommendations for number of cells	VMware recommends that the number of vCloud Director cell instances = $n + 1$, where n is the number of vCenter Server instances providing compute resources for vCloud consumption. Based on the service definition requirements, two vCloud Director cell instances are sufficient to increase availability and upgradability (upgrading one vCloud Director cell then the other).
Requirements for multicell configurations	<p>Multiple vCloud Director cells require NTP (Network Time Protocol), which is a best practice for all elements of the vCloud infrastructure.</p> <p>Consult the white paper, <i>Timekeeping in VMware Virtual Machines</i> (www.vmware.com/files/pdf/Timekeeping-In-VirtualMachines.pdf) for more information on how to set up NTP.</p>
Load balancer availability	At least two load balancers in a HA configuration should be used to reduce single points of failure. There are multiple strategies for this depending on vendor or software used.
Proxy configuration	Each load-balanced vCloud Director cell requires setting a proxy console IP address which is typically provided by the load balancer.
Rest API URL configuration	The vCloud service URL should map to the address provided via the load balancer. This is configured in the vCloud Director administrator GUI as well as in the load balancer configuration. This is the address that should be used to check the health status of the vCloud Director cell.
Awareness of Multicell Roles	Some vCloud Director cell roles (image transfer) may consume high resources. All cells can perform the same set of tasks, but it is possible to set policies that affect which ones are used. See the advanced configuration settings.
Load balancer session persistence	Sessions are generally provided in secure methods and are terminated at the cells. Because of this, session persistence should be enabled using SSL.
Load balancing algorithm	Least connections or round-robin is generally acceptable.

Consideration	Detail
vCloud Director cell status health checks	<p>Configure the load balancer service to check the health of individual vCloud Director cells. Because each cell responds via HTTPS, this can be configured quickly via the IP and API end point URL. Load balancers may support other types of health checks.</p> <p>Check services periodically based on load. A good starting point is every 5 seconds.</p> <p>Example GUI URL – https://my.cloud.com/cloud/</p> <p>Example API URL – https://my.cloud.com/api/versions</p> <p>In the second example, the versions supported by this end point are returned as XML.</p>
Public IP/port	Specify the service IP appropriately before adding cells to the service group. Typically, port 443 is the only port exposed – standard HTTPS.
Web Application Firewall	Can be used to apply URL restrictions on vCloud Director access to Admin or organization portals based on source address. Requires SSL sessions to be terminated on the load balancer.
SSL Initiation	Used when SSL is terminated on the load balancer to initiate an SSL session to the vCloud Director cells (which only accept HTTPS).
Advanced configurations	Load balancers can also provide Layer 7 content switching or direction, which may allow a vCloud Director configuration to send certain types of client traffic to “dedicated” cells. While cells can perform any function, it is possible to separate functions by directing certain types of requests to specific cells.
Connection mapping	When a cell joins an existing vCloud Director server group, it may try and load balance sessions. This may impact connection mapping through the load balancer as it is unaware of the balancing happening within the server group.

Appendix B: Security

Network Access Security

vShield Edge VPN functionality allows the creation of site-to-site tunnels using IPSEC. It supports NAT-T traversal for using IPSEC through network address translation (NAT) devices.

Table 19. Network Access Security Use Cases

Category	Description
Multi-site vCloud deployment	<p>The vShield VPN can connect multiple vCloud deployments. For example, an organization's virtual datacenter at a public vCloud provider can be securely connected with the organization's internal private vCloud. Or virtual datacenters hosted at a vCloud service provider in Europe can be connected to a vCloud service in Asia.</p> <p>Note that because vShield also provides address translation, it is possible to deploy multiple organization virtual datacenters at different providers using the same RFC1918 address space as long as unique subnets are used.</p>
Single-site vCloud deployment	<p>vShield VPNs can be created between either different organizations in the same vCloud Director instance, or different networks within the same organization.</p> <p>In this use case the site-to-site VPN is used to secure sensitive traffic between networks over shared infrastructure.</p>
Remote Site to vCloud VPN	<p>A permanent secure connection from a router or firewall based VPN; for example, Cisco/Juniper devices at a remote site to a vCloud environment with the vShield Edge. As the vShield VPN is a standard IPsec implementation, a wide range of devices can be used at the remote site (Commercial or Open Source).</p>
Client to cloud VPN	<p>Client software is generally not used with IPsec VPNs (as it is typically a permanent network-to-network tunnel), although clients with static IP addresses that implement pre-shared key authentication are supported.</p>

Site-to-site IPsec VPN configuration is now available to organization administrators directly from the vCloud Director web console. VPN functionality is implemented using integration with vShield Edge, which provides per-tenant Layer 3 network security and routing. It currently supports pre-shared key mode, IP unicast traffic, and NAT-T traversal with no dynamic routing protocols between the vShield Edge and peers. Behind each remote VPN endpoint multiple subnets can be configured to connect to the network behind a vShield Edge device over IPsec tunnels. These networks *must* have non-overlapping address ranges.

When configuring a site-to-site VPN between different organization networks in a vCloud environment (either across different vCloud environments or within an organization), much of the configuration complexity is abstracted from the vCloud consumer. After the appropriate networks are selected, both ends of the VPN tunnel are configured, automatically providing compatibility between the vShield Edge peers. In comparison, configuring remote devices to connect to a vShield Edge-based VPN requires an understanding of IPsec and the supported policies to successfully establish an encrypted tunnel.

The IKE Phase 1 parameters used by the vShield Edge VPN are:

- Main Mode.
- Pre-Shared Key Authentication Mode.
- 3DES or AES128 encryption.
- SHA1 authentication.
- MODP group 2 (1024 bits).
- SA lifetime of 28800 seconds (eight hours).
- Disable ISAKMP aggressive mode.

Additional parameters for IKE Phase 2:

- Quick Mode.
- Diffie-Helman Group 2/5 (1024 bit/1536 bit, respectively).
- PFS (Perfect Forward Secrecy).
- ESP Tunnel Mode.
- SA lifetime of 3600 seconds (one hour).

vShield Edge VPN proposes a policy that requires 3DES or AES128 (configurable although AES is recommended), SHA1, PSK and DH Group 2/5.

To allow IPsec VPN traffic, following ports need to be opened on firewalls in between the two endpoints:

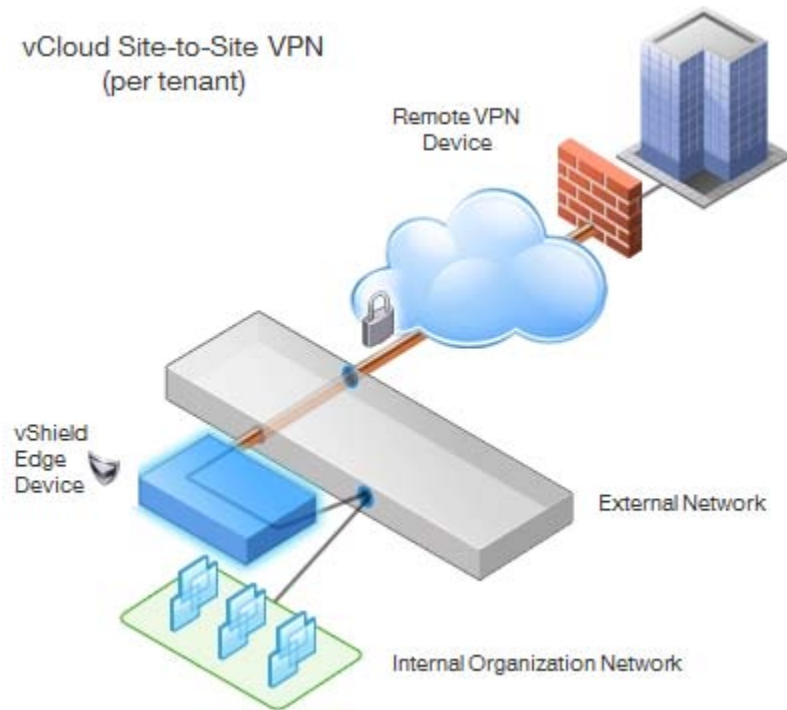
- Protocol 50 ESP.
- Protocol 51 AH.
- UDP port 500 IKE.
- UDP port 4500.

The external IP address for the vShield Edge device must be accessible to the remote endpoint, either directly or using NAT. In a NAT deployment, the external address of the vShield Edge must be translated into a publicly accessible address. Remote VPN endpoints then use this public address to access the vShield Edge.

It is also possible for the remote VPN endpoints to be located behind an NAT device as well, although on both ends a static one-to-one NAT is required for the external IP address.

As VPNs are used to provide secure access to an organization's remote networks, consumers should be aware of any security implications. A best practice for VPN configuration is to make sure that any VPN traffic is filtered and restricted to only those destinations that are absolutely necessary. vCloud Director 1.5 also introduces the ability to apply firewall rules to VPN traffic, whereas filtering was previously restricted to the remote end of a VPN tunnel only.

The vCloud Director IPsec VPN has a maximum of 10 sites per vShield Edge devices.

Figure 41. Site-to-Site VPN connectivity

The following features are not currently supported in the vShield Edge VPN implementation:

- Remote endpoints with dynamic IP addresses.
- Site-to-site VPNs at the vApp network level (available to organization networks only).
- SSL VPNs. These typically support per-user tunnels as opposed to network tunnels with IPsec VPNs, work over HTTPS and are often based on vendor specific implementations.
- IPv6 support.
- Authentication types other than Pre-Shared Keys. For example, certificates.
- Fenced vApps (VPN can only be enabled on routed networks).

DMZ Considerations

In general, standard firewall best practices should be followed in a vCloud environment. However, there are some areas which require special consideration. A number of vCloud Director operations involve sessions that remain open for a long period of time to management infrastructure, which is protected by the back-end firewall.

- Idle session timeouts – Depending on the level of activity within the vCloud environment some connections, such as the sessions to ESXi hosts to retrieve thumbnails via the vsld agent and to vCenter Server for inventory may require adjustment to default TCP timeout policies. This is also a consideration for ONS connections required for Fast Connection Failover support in Oracle RAC environments.
- Dead Connection Detection or equivalent – Many firewalls support functionality to allow idle but still valid connections to persist. This modifies the idle timeout behavior by probing endpoints of the connection and making sure the session is not terminated.
- Logging – Make sure firewall logs are collected by a centralized syslog server.
- SMTP filtering – Many firewalls filter email connections, restricting ESMTP commands. In some cases this feature may need to be disabled to permit vCloud Director to send mail notifications.
- Bandwidth – Some vCloud operations require either high throughput or low latency (examples of this are NFS transfer access and database access). Therefore, it is important to make sure the firewall is appropriately specified and does not become a performance bottleneck.
- Availability – Deploy firewalls and load balancers in highly available pairs where possible.
- Secure Administrative Access – Tightly control access to the management networks using strong authentication, logging and encryption.
- Scalability – vCloud environments are typically architected to scale and support a large number of workloads and users. Make sure that the firewalls can scale along with the vCloud to help avoid future downtime.

Port Requirements

Table 20. vCloud Director Port Requirements

Description	Ports	Protocol	Direction
vCloud Director Portal and Console Proxy Access	443	TCP	Inbound
SSH (back-end management access only)	22	TCP	Inbound
JDBC access to Oracle database	1521 (default)	TCP	Outbound
ONS connections for Oracle RAC	6200 (default)	TCP	Outbound
Microsoft SQL database port	1433 (default)	TCP	Outbound
vSphere Web Access to vCenter Server	443	TCP	Outbound
VM Console to vCenter Server	902, 903	TCP	Outbound
vSphere Web Access to ESX/ESXi host	443	TCP	Outbound
VM Console to ESXi host	902	TCP	Outbound
REST API access to vShield Manager	443	TCP	Outbound
SMTP	25	TCP	Outbound
DNS client	53	TCP, UDP	Outbound
NTP client	123	TCP, UDP	Outbound
LDAP	389	TCP	Outbound
LDAPS	636	TCP	Outbound
Syslog	514	UDP	Outbound
NFS Portmapper (optional)	111	TCP, UDP	Inbound and Outbound
NFS <code>rpc.statd</code> (optional)	920	TCP, UDP	Inbound and Outbound
ActiveMQ	61611, 61616	TCP	Inbound and Outbound

Figure 42. vCloud Director Port Requirements Illustrated

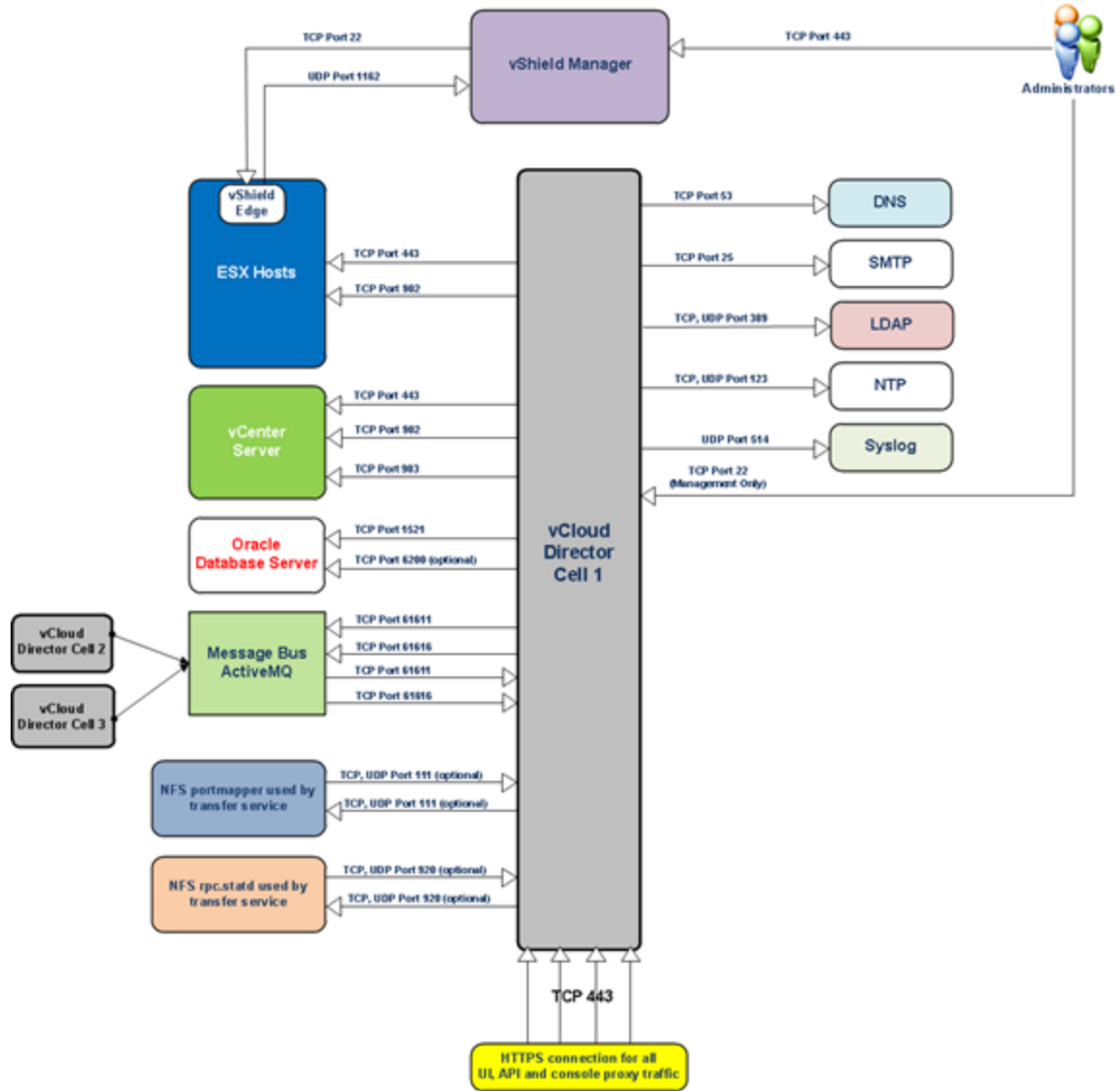


Table 21. vCenter Orchestrator Port Requirements

Name	Protocol	Hostname	Default Port
Database	Oracle	Oracle Database Server	1521
	MSSQL	Microsoft SQL Server	1433
Directory Service	LDAP/LDAP SSL /GC	Microsoft Active Directory Server	389/636/3268
	LDAP/LDAP SSL	Novell eDirectory	389/636
	LDAP/LDAP SSL	Sun Java Directory Server	389/636
Domain Name System	DNS	DNS Server	53
vCenter Server	HTTPS	vCenter Server	443
vCloud	HTTPS	vCloud Server or vCloud load balancer if configured	443
SSH	SSH	SSH Server	22
Mail	SMTP	SMTP Server	25
Net	POP3	POP3 Server	110
JDBC	Oracle	Oracle Database Server	1521
	MSSQL	Microsoft SQL Server	1433
Cisco UCS Manager	HTTP	UCS Manager Server	80
SOAP	HTTP	SOAP Server	80
	HTTPS		443
REST	HTTP	Rest Server	80
	HTTPS		443
Microsoft Active Directory	LDAP msft-gc	Active Directory Domain Controller Server	3268
		Active Directory Global Catalog Domain Controller Server	389
VIX	HTTPS	vCenter Server	443