vmware | vCLOUD® ARCHITECTURE TOOLKIT™

VMware vCloud® Architecture Toolkit
Architecting a VMware vCloud

Version 3.1

January 2013

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

Contents

List of Figures

List of Tables

1. Overview

Architecting a VMware vCloud provides guidance to architect an Infrastructure-as-a-Service (IaaS) cloud based on the VMware® vCloud® Suite. The vCloud Suite dramatically simplifies IT operations, delivering enhanced agility and better economics. At the heart of the suite are a set of *software-defined datacenter* services. These represent the application of the virtualization principles of pooling, abstraction, and automation to the domains of storage, networking, security, and availability.

The vCloud Suite components addressed in this guide include VMware vSphere®, VMware vCloud Director®, VMware vCloud Networking and Security™ (formerly vShield), the VMware vCenter™ Operations Management Suite™, VMware vFabric™ Application Director™, VMware vCenter Site Recovery Manager™, and VMware vCloud Connector™. Simplifying the delivery of resources to end users requires the architectural integration and coordination of these components. Both service providers and enterprises can use the design guidelines, with some variations depending on the point of view.

This document, combined with a service definition, can help you evaluate the design considerations for architecting a vCloud solution. Use the following vCloud documents together throughout the lifecycle of a VMware vCloud computing implementation.

- *Architecting a VMware vCloud* provides design guidelines, design considerations, and design patterns for constructing a vCloud environment from its constituent components.
- *Operating a VMware vCloud* includes design guidelines 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.

Additionally, *VMware vCloud Implementation Examples* provides modular examples that show how to use VMware component software to implement a vCloud instance.

1.1 Audience

This document is for people involved in planning, designing, and implementing VMware vCloud solutions. The target audience is architects, engineers, and IT professionals who have achieved VMware Certified Professional (VCP) or higher certification and are familiar with VMware products. The reader is assumed to be familiar with vSphere and vCloud concepts.

1.2 Scope

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

1.3 Document Topics

The remainder of this document is divided into the sections listed 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, vCloud Management Architecture	Describes the components required to build a vCloud solution.
Section 4, Resource Group Architecture	Provides guidance for configuring resources reserved for end-user workloads.
Section 5, vCloud Resource Design	Offers design guidelines for partitioning and delivering vCloud resources relative to customer requirements.
Section 6, vCloud Metering	Covers how to meter and charge for resources with VMware vCenter Chargeback Manager™ (a component of the VMware vCenter Operations Management Suite).
Section 7, Orchestration and Extension	Provides information about extending vCloud Director automation through orchestration.
Section 8, Multisite Considerations	Covers multisite considerations.
Section 9, Hybrid vCloud Considerations	Provides information about extending vCloud Director into the hybrid vCloud model.
Appendix A: Availability	Design considerations for availability.
Appendix B: Security	Design considerations for security.
Appendix C: vCloud Suite Disaster Recovery	Design considerations for disaster recovery.
Appendix D: vCloud Director Upgrade Considerations	Design considerations for upgrading to vCloud Director 5.1.

2. vCloud Architecture

Cloud computing is a model that allows ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources. vCloud resources are provisioned rapidly and released with minimal management effort.

The VMware vCloud suite delivers a complete, integrated cloud infrastructure suite that simplifies IT operations while delivering the best service level agreements (SLAs) for all applications. The vCloud Suite includes the entire set of vCloud infrastructure capabilities: virtualization, software-defined datacenter services, policy-based provisioning, disaster recovery, application management, and operations management.

The vCloud solution encompasses the vCloud Suite along with an architecture defined in the *VMware vCloud Architecture Toolkit* (vCAT) and a set of recommended guidelines for organization, process design, and instrumentation. These all feed a CIO scorecard enabled by the VMware IT Business Management (ITBM) product suite. It is an all-encompassing approach to maximizing the benefits of the software-defined datacenter.

Architecting a vCloud focuses on the IaaS layer, detailing use of the vCloud Suite to extend the capabilities of the vSphere virtualization platform.

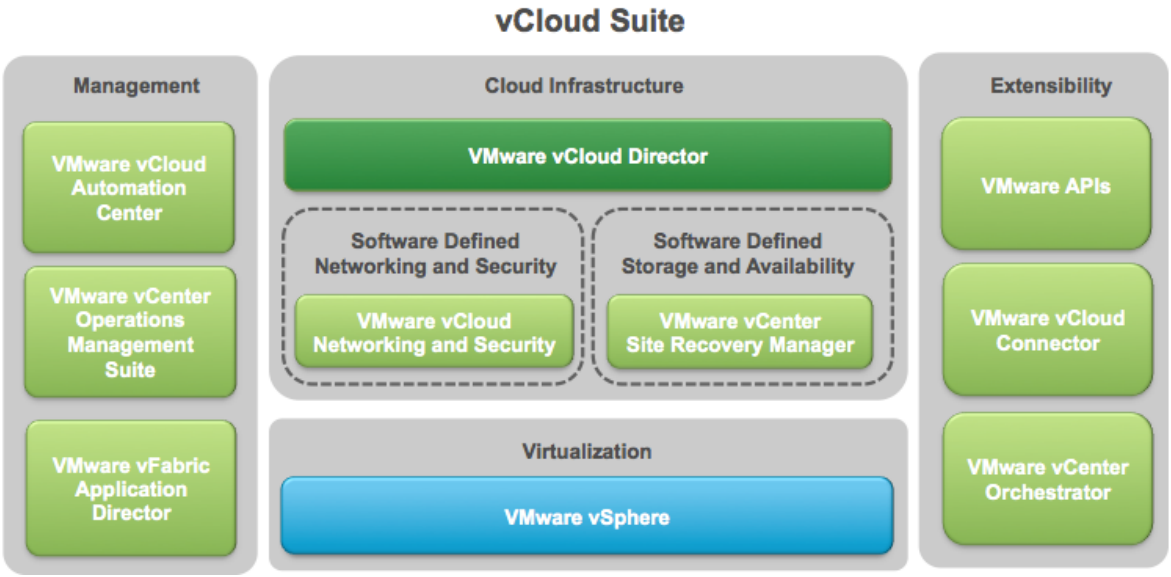
2.1 Technology Mapping

The VMware vCloud Solution is 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, interoperability, and extensibility with other vCloud instances. VMware vCloud Automation Center™ extends vCloud Networking and Security, virtualizes the network, and creates agile, extensible, secure logical networks that meet the performance and scale requirements of virtualized applications and data. The vCenter Operations Management Suite provides the capabilities necessary to achieve an integrated approach to performance, capacity, and configuration across a vCloud infrastructure.

A vCloud architecture provides a conceptual framework to support primary business requirements, determine system functions, organize elements into distinct components, and define boundaries and connections. The focus is on clearly defined goals, analysis, and design decisions that cut through the complexity in today's technology.

The following figure shows the components of the vCloud Suite solution stack.

Figure 1. Technology Mapping



2.2 vCloud Suite Components

The following table describes the components that comprise the vCloud Suite.

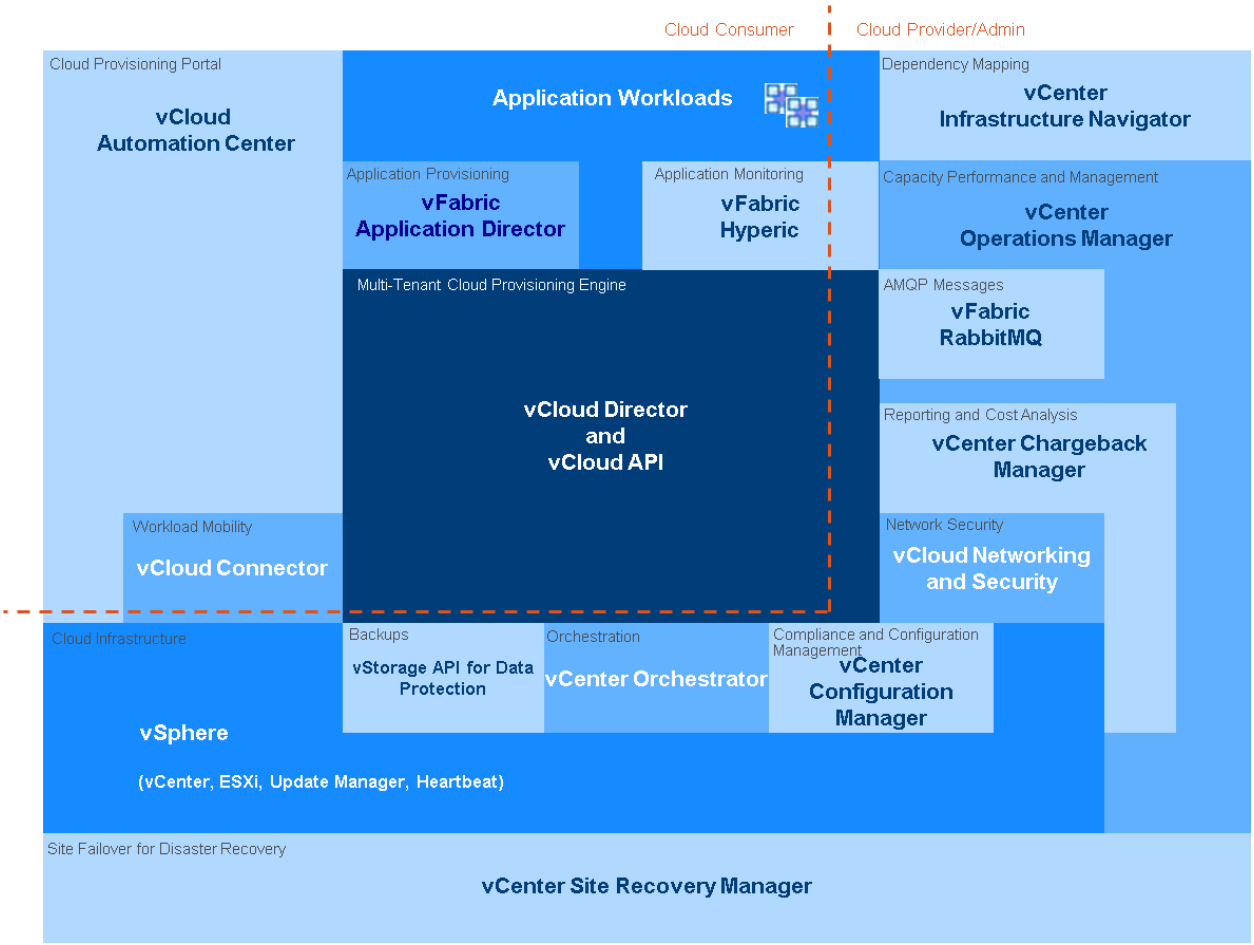
Table 2. vCloud Components

vCloud Component	Description
VMware vCloud Director	Layer of software that abstracts virtual resources and exposes vCloud components to consumers. This includes: • vCloud Director server (also referred to as a cell). • vCloud Director database.
vCloud API	VMware vCloud API, used to manage vCloud objects programmatically.
VMware vCloud Automation Center	Enables rapid deployment and provisioning of vCloud services across private and public vClouds, physical infrastructures, hypervisors, and public vCloud providers through a secure self-service portal.
VMware vSphere	Virtualization platform providing abstraction of physical infrastructure layer for vCloud. This includes: • vSphere hosts. • VMware vCenter Server™. • vCenter Server database.

VMware vCloud Networking and Security	Decouples network and security from the underlying physical network hardware through software-defined networking and security. This includes: • VXLAN support. • vCloud Networking and Security Edge™ gateway. • vCloud Networking and Security App™ and vCloud Networking and Security Data Security™. • vCloud Networking and Security Manager™.
VMware vCenter Operations Management Suite	Provides predictive capacity and performance planning, compliance and configuration management, dynamic resource metering, cost modeling, and report generation using the following components: • vCenter Operations Manager™. • vCenter Configuration Manager™. • vCenter Infrastructure Navigator™. • vCenter Chargeback Manager™.
VMware vFabric Application Director	Part of the VMware vFabric Cloud Application Platform family of products that provide automated provisioning of application infrastructure.
VMware vCenter Orchestrator™	Enables the automation of provisioning and operational tasks across VMware and third-party applications using an open and flexible plug-in architecture.
vCloud Connector	VMware vSphere Client™ plug-in that enables users to connect to vClouds based on vSphere or vCloud Director, and manage them through a single interface.

The following figure shows the relationship among vCloud Suite components. Except for gray components, components that touch are integrated with each other.

Figure 2. vCloud Suite Components



2.3 vCloud Infrastructure Logical Design

When architecting a VMware vCloud infrastructure logical design, VMware recommends using a *building block* approach to provide a scalable, resilient architecture. The following top-level logical building blocks are used to segregate resources that are allocated for management functions from resources dedicated to user-requested workloads.

- *vSphere virtual management cluster* – Contains the core and optional components and services needed to run the vCloud instance. This includes core vCloud components such as VMware vCenter Server, vCloud Director, vCenter Chargeback Manager, vCenter Orchestrator, and optional components such as the vCenter Operations Management Suite and vFabric Application Director.
- *Resource group* – Represents vCloud-dedicated resources for end-user consumption. Each resource group consists of vSphere clusters (vSphere hosts managed by a vCenter Server) and is under the control of vCloud Director. vCloud Director can manage the resources of multiple resource groups.

Separate management and resource clusters are important for the following reasons:

- Separation of duties – A vCloud infrastructure typically has at least two types of administrator: infrastructure (vSphere) administrator and vCloud administrator. Separating the virtual management cluster from resource groups allows separation of duties and enforcement of administrative boundaries, limiting the actions that can be performed in the vSphere clusters of a resource group.

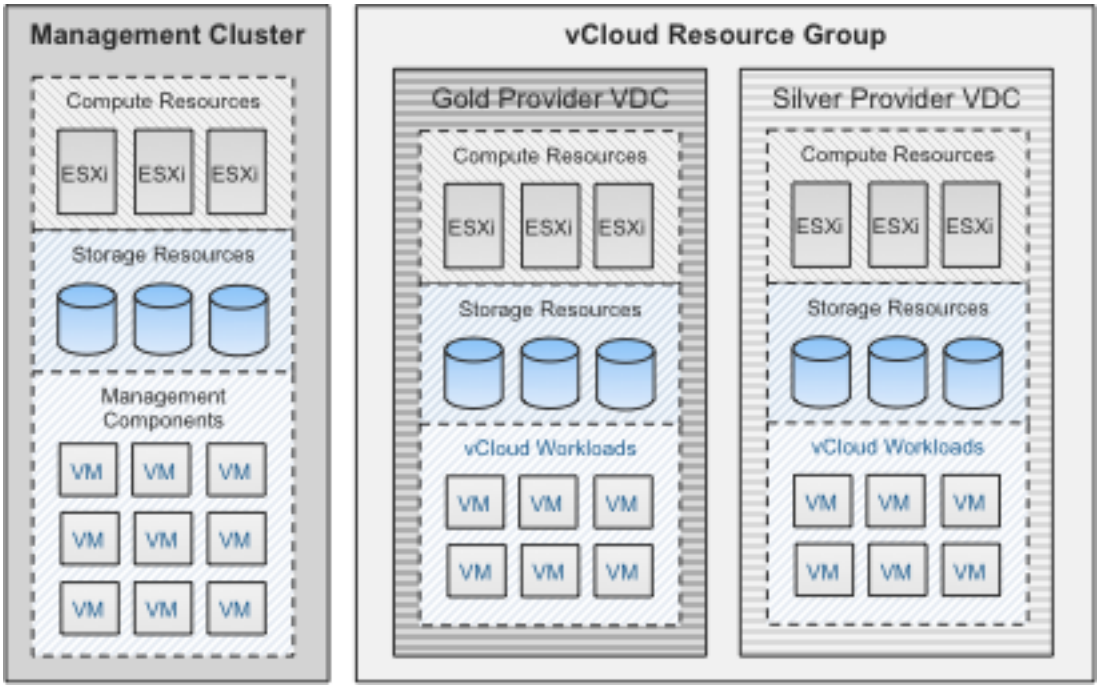
  An administrator should not perform the following actions on a resource group through the vSphere Client:

  - Editing virtual machine properties.
  - Renaming virtual machines.
  - Disabling VMware vSphere Distributed Resource Scheduler™ (DRS).
  - Deleting or renaming resource pools.
  - Changing networking properties.
  - Renaming datastores.
  - Changing or renaming folders.

  This is not an exhaustive list, but it covers some of the detrimental actions a vCenter administrator could perform on a vCloud resource group.

- Resource consumption – Virtual machines deployed into resource groups that are not managed by vCloud Director consume resources that are allocated for a particular vCloud virtual datacenter. This skews the resource utilization and consumption metrics available to the vCloud.
- Scalability and configuration maximums – Having separate vSphere clusters to manage compute resources consumed by end users increases resource group scalability. A vCloud environment must conform to vSphere scalability and configuration maximums. Having dedicated resource group vSphere clusters means that the scalability of vCloud user resources is not affected by management workloads.
- Availability – A virtual management cluster allows the use of VMware vSphere High Availability (HA) and DRS to provide enhanced availability to all management components. A separate management cluster enables this protection in a granular fashion to satisfy management-specific SLAs. It also increases upgrade flexibility because management cluster upgrades are not tied to resource group upgrades.
- Denial-of-service attacks or intensive provisioning – Having separate management clusters and resource groups keeps this type of activity on the resource groups from affecting management component availability.
- Disaster recovery facilitation – Having separate management clusters and resource groups simplifies design and implementation of vCloud disaster recovery. The vCloud disaster recovery solution uses a vSphere cluster managed by vCenter Site Recovery Manager that contains the vCloud infrastructure management components. For more information, see Appendix C: vCloud Suite Disaster Recovery.
- Support and troubleshooting – Running management components in large clusters that contain mixed resource and management components makes it difficult to diagnose issues with the management components. To facilitate troubleshooting and problem resolution, place the management components in a small and manageable cluster.
- Separation of management components from managed resources – Separation helps to prevent inadvertent changes through the vSphere Client to entities created with vCloud Director.

Figure 3. vCloud Logical Architecture Overview



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

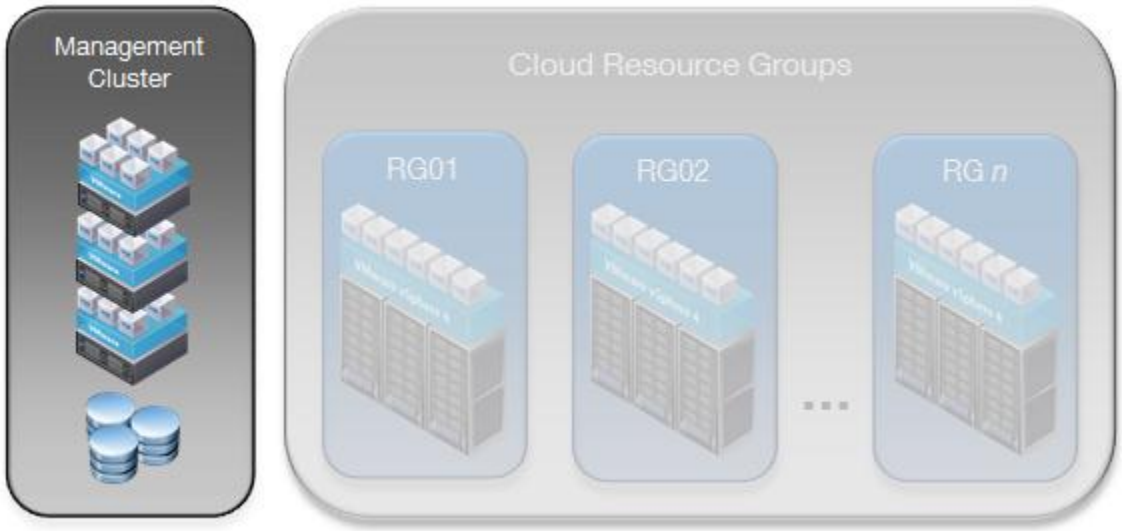
3. vCloud Management Architecture

The design and architecture of the vCloud management infrastructure is critical to support availability and scalability of the vCloud solution.

3.1 Management Cluster

The management cluster hosts 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



Core management cluster components include the following:

- vCenter Server or VMware vCenter Server Appliance™.
- vCenter Server database.
- vCloud Director cells.
- vCloud Director database.
- vCloud Networking and Security Manager (one per resource group vCenter Server).
- vCenter Chargeback Manager.
- vCenter Chargeback database.
- VMware vCenter Update Manager™.
- vCenter Orchestrator.
- VMware vCloud Networking and Security Edge gateway appliances deployed by vCloud Director through vCloud Networking and Security Manager as needed, residing in the resource groups, not in the management cluster.

The following management cluster components are optional:

- VMware vCenter Server Heartbeat™.
- vCloud Automation Center.
- vCloud Connector.
- VMware vFabric RabbitMQ™.
- vFabric Application Director.
- VMware vFabric Hyperic® HQ.
- VMware vSphere Management Assistant.
- vCenter Operations Manager.
- vCenter Configuration Manager.
- vCenter Infrastructure Navigator.
- vCenter Site Recovery Manager.
- Databases for optional components.

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

The management cluster can 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. See *Service Definitions* for detailed sizing considerations.

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

Both the management cluster and resource groups reside at the same physical site to provide a consistent level of service. This minimizes latency issues that might arise in a multisite environment if workloads move between sites over a slower or less reliable network. See Section 8, Multisite Considerations for considerations associated with connecting vCloud instances at different sites.

3.1.1 Component Sizing

The following table lists the requirements for each of the components that run in the management cluster. The following recommendations will scale to accommodate the number of virtual machines and organizations listed in the private or public service definitions.

Table 3. Component Requirements for a Management Cluster

Item	vCPU	Memory	Storage	Networking
vCenter Server	2	4GB	20GB	1GigE
Database server	4	16GB	100GB	1GigE
vCloud Director cell 1	2	4GB	30GB	1GigE
vCloud Director cell 2	2	4GB	30GB	1GigE
vCenter Chargeback Manager	2	4GB	30GB	1GigE
vCloud Networking and Security Manager	2	8GB	8GB	100Mb
Total	14	40GB	218GB*	4GigE*

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

The database server hosts databases for vCenter Server, VMware vCenter Single Sign-On, vCloud Director, and vCenter Chargeback Manager. Use different users and instances for each database based on VMware design guidelines. VMware vCloud Director 5.1 supports both Oracle and Microsoft SQL Server databases.

To facilitate file transfers in a multicell environment, a shared storage volume must be configured and made accessible to all cells in a vCloud Director server group. The necessary volume size varies on the expected number of concurrent uploads. Following an upload, 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 Upgrade Guide* in the vCloud Director documentation (http://www.vmware.com/support/pubs/vcd_pubs.html).

3.2 Compute Layer

The management cluster compute layer encompasses the CPU, memory, and hypervisor technology components. Follow vSphere design guidelines when configuring and sizing compute resources.

Figure 5. Three-Host Management Cluster

Use a three-host cluster to support vCloud management components. Add additional hosts if the management cluster becomes resource-constrained.

Enable VMware vSphere High Availability and DRS on the management cluster to provide availability for all management components. For vSphere HA, use the **Percentage as cluster resources reserved** admission control policy in an N+1 fashion instead of defining the number 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.

The vCloud Director-managed vCenter Server instances play 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 using solutions such as VMware vCenter Server Heartbeat.

VMware vSphere Fault Tolerance can be used for continuous virtual machine protection only if all FT requirements are met. vCenter Site Recovery Manager can be used to protect components of the management cluster against site failure. See Appendix C: vCloud Suite Disaster Recovery for details.

vCloud Director 5.1 supports vSphere 5.0 and later. Deploy vSphere 5.1, if possible, to take advantage of the new features. Some functionality in vCloud Director requires specific features and requires particular vSphere editions. For example, automated deployment of vCloud networks requires a distributed switch, which is supported in the VMware vSphere Enterprise Plus Edition™.

3.3 Network Layer

The following design guidelines apply to network configuration for the management cluster:

- Separate network traffic logically for security and load according to traffic type (management, virtual machine, VMware vSphere vMotion®, FT, IP storage).
- Implement network component and path redundancy.
- Implement network speeds of least 1GigE–10GigE, if possible.
- Standardize on VMware vSphere Distributed Switch™ across all clusters, including the management cluster.

3.4 Storage Layer

Use vSphere storage design guidelines where applicable for the management cluster. Examples include the following:

- Configure redundancy at the host (connector), switch, and storage array levels.
- Give all hosts in a cluster access to the same datastores.
- Enable VMware vSphere Storage APIs – Array Integration (VAAI).
- Use single-initiator storage fabric zoning for vSphere hosts.

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. Linked mode in the vCloud Director context allows viewing of 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. The number of linked mode objects in a vCloud environment does not reach the linked mode maximums unless multiple vCloud instances are involved.

Additional considerations include the following:

- The vCenter Server appliance does not support linked mode.
- A vCenter instance can link only with other vCenter instances of the same version. Keep this in mind when upgrading all vCenter Server instances in a vCloud instance.
- Upgrading a linked vCenter instance breaks the link and the instance becomes independent.

3.6 Cell Load Balancing

vCloud Director cells are stateless front-end processors for the vCloud instance. Each cell 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 with 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 or remove cells from the active service pool. The cell architecture is not a true cluster because 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, such as the vCloud Networking and Security Edge gateway, can perform load balancing of vCloud Director cells. See the general design guidelines regarding performance, security, and manageability when deciding to share or dedicate load balancing resources.

Note SSL offloading does not work with virtual machine remote console (VMRC) proxy connections.

See Appendix A: Availability for additional load balancing considerations.

3.7 vCenter Operations Manager

An embedded adaptor handles integration between vCloud Director and vCenter Operations Manager. The vCloud Director Adapter discovers and creates the mapping for the following vCloud entities:

- Organization.
- Provider virtual datacenter.
- Organization virtual datacenter.
- vApp.

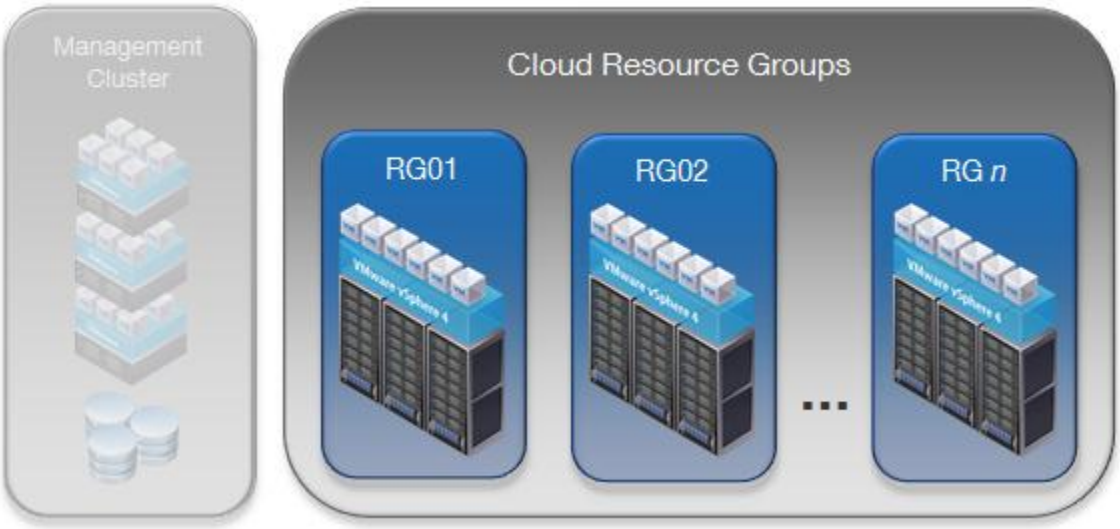
After the mapping is performed, vCloud Director objects can be incorporated into vCenter Operations dashboards. Optionally, the adapter can import change events related to vCloud entities.

For details on installing and configuring the adapter, see the *VMware vCloud Director Adapter Installation and Configuration Guide* (ftp://ftp.integrien.com/VCOPS-VMware_Product_Adapters/VMware-vCD/vCloud-Director-Adapter-Install-and-Config-Guide.pdf).

4. Resource Group Architecture

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 Server instances. 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 provides a consistent approach for scaling vCloud environments. A separate vCenter Server instance is recommended to manage Cloud Resource Groups. 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 Cloud Resource Groups.

Caution Do *not* make changes to resource group objects using the vSphere Client. Changing the state of vCloud Director-created objects using the vSphere Client can cause unpredictable side effects because these objects are owned by vCloud Director.

4.1 Compute Resources

Configure resource group vSphere hosts per vSphere design guidelines. Enable vSphere HA appropriately to protect against host and virtual machine failures.

The shift to Fault Domain Manager (FDM)-based HA in vSphere 5 is transparent to vCloud Director. The total number of hosts in an HA/DRS cluster remains at 32, so cluster sizing guidelines for vCloud environments do not change. FDM requires a single master host instead of five primary nodes for legacy HA. If the master host fails, the remaining slave hosts select a new master.

The eight hosts per cluster limitation for fast provisioning (linked clones) and VMFS datastores does not apply to vSphere 5.1-backed resource groups. Fast provisioning on VMFS5 datastores supports up to 32 hosts.

Provider virtual datacenters represent a service offering. When building clusters, group similar servers together (based on 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

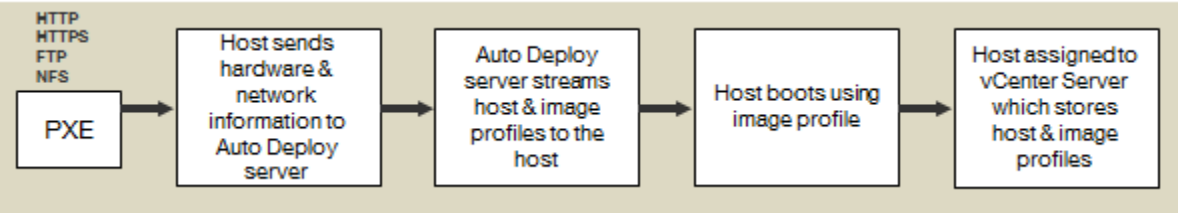
Stateless ESXi refers to running VMware ESXi™ software on a host entirely in memory, with no local persistence data. Centralizing management of the host state enables consistent configuration over large sets of similar hosts and rapid provisioning of vSphere hosts. This helps to improve operational efficiency in large-scale vCloud environments.

Stateless ESXi requires VMware vSphere Auto Deploy™, a deployment server that applies the image profile and host profile to the PXE-booted hosts. Install vSphere Auto Deploy on a standalone host or on the vCenter Server instance. vSphere Auto Deploy is installed by default on the vCenter Server virtual appliance. Install VMware vSphere PowerCLI™ in a location that vCenter and vSphere Auto Deploy can both reach. 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 vSphere hosts for DHCP. The DHCP server requires configuration changes to direct the vSphere host to a TFTP server. The server can be a separate DHCP server or it can be the 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 vSphere Auto Deploy and add rules to the rules engine using VMware vSphere® ESXi™ Image Builder CLI.

Figure 7. Auto Deploy First Boot



vCloud Director can manage stateful or stateless vSphere hosts. If you choose the stateless option, add the vCloud Director vSphere Installation Bundle (VIB) (which contains the agent) to the image profile. The vCloud Director VIB is loaded automatically when the host boots up. For preparation and unpreparation 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 pushed again 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 Resources

For the vCloud resource groups, configure networking with vSphere design guidelines in mind. Increase the number of vSphere Distributed Switch ports per host to the maximum of 4096 to improve the scale at which vCloud Director can dynamically create port groups for vCloud networks. For more information about increasing this value, see the *vCenter Server and Host Management Guide* in the *VMware vSphere Documentation* (http://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-pubs.html).

Increase the maximum transmission unit (MTU) size to 1600 for all physical network devices and vSphere Distributed Switches in the transport network to support VXLAN or vCloud Director network isolation. Failure to increase the MTU size causes packet fragmentation, negatively affecting throughput of end-user workloads.

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

4.2.1 I/O Controls

vCloud Director offers the following controls to guard against the misuse of resources by consumers:

- 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 act as the default for all new users added to the organization.
- Limits for resource-intensive operations prevents consumers from affecting all users in an organization and provide a defense against denial-of-service attacks.
- Limit the number of simultaneous VMRC 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. A key driver for transitioning to IPv6 is the much larger supported address space of 2^{64} addresses as opposed to the 2^{32} addresses for IPv4.

The following vCloud Director components are required to support IPv6:

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

The following vSphere infrastructure components support IPv6:

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

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

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

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

Many destinations do not currently support IPv6, so operate virtual machines in dual stack IPv4 and 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 Virtual eXtensible LAN (VXLAN)

Virtual eXtensible LAN (VXLAN) is an IETF submitted protocol that uses an encapsulation mechanism to enable Layer 2 overlay on Layer 3 networks. VXLAN is used to support elastic virtual datacenters across different networks.

VXLAN is designed to be deployed seamlessly on existing networks, requiring few changes on the physical network. VXLAN requires deployment of IP multicast across the physical network infrastructure by enabling IGMP (v1, v2, and v3) snooping on physical switches and PIM for multicast routing.

4.2.4 vCloud Networking and Security Edge

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

The license for vCloud Networking and Security Edge that is included with vCloud Director does not include features such as SSLVPN and load balancing capabilities, which are part of the fully licensed VMware vCloud Networking and Security Advanced Edition.

4.2.5 vCloud Networking and Security App

VMware vCloud Networking and Security 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 security groups, which are custom containers created through the vCloud Networking and Security Manager UI. Container policies allow the creation of mixed trust zone clusters without requiring an external physical firewall. vCloud Networking and Security App also supports classic five tuple firewall rules.

4.2.6 vShield Endpoint

VMware vShield Endpoint offloads antivirus functions to a hardened security virtual machine supplied by partners such as Trend Micro. vShield Endpoint uses VMware endpoint security (EPSEC) APIs to access the file system to scan and remediate viruses. This removes the need for agents in the guest operating system and prevents antivirus storms from consuming CPU cycles during scanning or antivirus update activities. Offloading antivirus functions provides enhanced security, as a malware attack often begins by disabling antivirus agents. The efficient antivirus architecture of vShield Endpoint provides antivirus as a service for large-scale vCloud environments.

4.2.7 vCloud Networking and Security Data Security

VMware vCloud Networking and Security Data Security™ provides visibility into sensitive data stored within your organization's virtualized and vCloud environments. Violations data reported by vCloud Networking and Security Data Security can provide information needed to protect sensitive data and achieve regulatory compliance.

Note Currently, vCloud Director 5.1 is not integrated with vCloud Network and Security App, vCloud Networking and Security Data Security, or vShield Endpoint. Using these features in conjunction with vCloud Director requires careful design of the vCloud infrastructure.

4.3 Storage Resources

Designing storage resources for vCloud differs from the traditional vSphere approach. Platform features such as VMware vSphere Storage DRS™ and storage profiles assist in balancing workloads across storage resources, allowing providers to offer differentiated storage. This allows the provisioning of flexible pools of storage resources while maintaining consistent performance for end users. Users can choose the right storage tier for a particular type of workload.

VMware recommends the following when designing storage resources:

- Perform a current state analysis for storage usage and trends.
- Define the range of storage SLAs needed and appropriate pricing models.
- Create multiple storage profiles in vSphere, based on SLAs, workloads, and cost.
- Map storage profiles to the provider virtual datacenter.
- Select a subset of storage profiles provided by the provider virtual datacenter to the organization virtual datacenter.
- Design for optimal availability (redundant paths from vSphere hosts to storage fabric).
- Deploy modular and scalable physical storage.
- Monitor storage usage and trends using capacity analysis tools.
- Use storage performance tools to tune vApp storage workloads.

vCloud Director supports tiering storage within a virtual datacenter using storage profiles. Configure shared storage and storage profiles in the resource groups per vSphere design guidelines.

Datastore sizing considerations include both capacity and performance:

- Datastore capacity considerations:
  - 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 x number of virtual machines x (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 used?
- Datastore performance considerations:
  - Will expected workloads be disk intensive?
  - What are the performance characteristics of the associated cluster?

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

4.3.1 Storage Tiering

Storage tiering in vCloud Director 5.1 is enabled on a per virtual machine basis through storage profiles:

- Authoring, renaming, and deletion of storage profiles are performed through vSphere.
- Storage profiles can be added, disabled, or removed at the provider virtual datacenter level.
- All available storage profiles across selected clusters are listed at provider virtual datacenter creation.
- Organization virtual datacenter storage profiles are based on a subset of storage profiles provided by the provider virtual datacenter.
- Each organization virtual datacenter has an associated default storage profile.
- All virtual machines have an associated storage profile that defaults to the organization virtual datacenter storage profile.
- Virtual machine placement is based on storage profiles.

Other entities that support storage profiles:

- Templates.
- Media.
- Independent disks.

OVF storage profiles support:

- vSphere does not export storage profile association when exporting a virtual machine to OVF.
- vCloud Director template download exports a template virtual machine's default instantiation profile.
- vCloud Director template upload applies OVF-specified template virtual machine's default instantiation profile.

Disks independent of virtual machine:

- Are associated with an organization virtual datacenter storage profile.
- Allow selection of datastore to place the disk accounts for storage profile.
- Can have their storage profile changed.
- Are allowed to be on a different storage profile than the virtual machine to which the disk is attached.

The following are vSphere changes that affect vCloud Director storage profiles:

- Changed datastore labels.
- Deleted storage profiles.
- Changed virtual machine storage profile association.
- Virtual machine migration to a new datastore using VMware vSphere Storage vMotion.

Storage profile compliance checks are performed:

- When initiated through the REST API.
- When automatically performed at set time intervals.
- When a storage profile in use by an organization virtual datacenter is deleted in vCenter.
- When a virtual machine is migrated using vSphere Storage vMotion.

Non-compliance shows up in the form of system alerts on the virtual machine.

4.3.2 vSphere Storage vMotion

vSphere 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 the 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.
- The vCloud API is used to initiate vSphere Storage vMotion for linked clones to preserve the linked clone state.

Caution Do not invoke vSphere Storage vMotion migration of linked clones using the vSphere Client because doing so might cause undesirable effects such as the inflation of delta disks. A vSphere Storage vMotion operation involving a datastore *and* host might fail.

4.3.3 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 results in lower worst-case device latency by maintaining a balance between workload isolation/prioritization and storage I/O throughput. For more information, see *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 mapping (RDM) or datastores with multiple extents. If you are using datastores backed by arrays with automated storage tiering, validate compatibility with SIOC.

4.3.4 vSphere Storage APIs – Array Integration

vSphere Storage APIs – Array Integration (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.

In vCloud environments, cloning and snapshot operations stemming from provisioning tasks can quickly overwhelm the system. VAAI improves storage task execution times, network traffic utilization, and CPU host utilization during heavy storage operations.

For block-based storage systems, 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.

Hardware acceleration for NAS is enabled through the installation of vendor plug-ins. VAAI NAS plug-ins are developed by storage vendors and validated by VMware.

vCloud Director 5.1 supports the following offload through VAAI integration:

- Block (FC, iSCSI) – Full copy offload to array (ESXi 4.1 or later, with supported storage array firmware listed in the *VMware Compatibility Guide*).
- NFS – Full copy offload to array (ESXi 5.0 or later only, with vendor supplied VIB (Virtual Infrastructure Bundle) installed on ESXi server).
- NFS – Linked clone offload to array for storage arrays supporting clones of clones.

See the *VMware Compatibility Guide* (http://www.vmware.com/resources/compatibility/search.php) for more details.

4.3.5 vSphere Storage DRS

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

- vCloud Director 5.1 supports vSphere Storage DRS when using vSphere 5.1 hosts. vSphere Storage DRS also supports fast provisioning (linked clones) in vCloud Director 5.1.
- vSphere Storage DRS continuously balances storage space usage and storage I/O load, avoiding resource bottlenecks to meet service levels and increase manageability of storage at scale.
- vCloud Director 5.1 recognizes storage clusters. The member datastore clusters are visible in vCloud Director, but cannot be modified from vCloud Director.
- vCloud Director 5.1 utilizes vSphere Storage DRS for initial placement of virtual machines.
- vCloud Director uses vSphere Storage DRS to manage space utilization and I/O load balancing. vSphere Storage DRS can help rebalance virtual machines, media, and virtual machine disks within the storage pod.
- As in vCloud Director 1.x, vCloud Director 5.1 determines optimal placement between datastore clusters and standalone datastores across all vSphere instances assigned within vCloud Director.
- There is a new VIM object type in the REST API named `DATASTORE_CLUSTER`. The datastore properties now include the member datastore list when the VIM object type is `DATASTORE_CLUSTER`.

4.3.5.1. vSphere Storage DRS and Fast Provisioning

- vSphere Storage DRS supports fast provisioning in vCloud Director.
- vSphere Storage DRS supports linked clones only with vCloud Director 5.1.
- Linked clone configurations that span across datastores are not supported in vCloud Director 5.1.
  - vSphere Storage DRS will not recommend placement of a linked clone that would span datastores from the base disk.
  - vSphere Storage DRS will migrate a clone to a datastore containing a shadow virtual machine and relink the clone to the existing shadow virtual machine.
- Linked clones can be migrated between VMFS3 and VMFS5 and are supported in vSphere Storage DRS. The format conversions are handled automatically at the platform level.
- The logic for migrating a virtual machine is influenced by factors such as the following:
  - Amount of data being moved.
  - Amount of space reduction in the source datastore.
  - Amount of additional space on the destination datastore.
- Linked clone decisions also depend on whether the destination datastore has a copy of a base disk or whether a shadow virtual machine must be instantiated:
  - Putting a linked clone on a datastore without the base disk results in more space used on the datastore as opposed to placing the clone on a disk where a shadow virtual machine already exists.
  - During the initial placement, vSphere Storage DRS selects a datastore that contains a shadow virtual machine so that placement results in maximum space savings. If necessary, initial placement recommendations can include evacuating existing virtual machines from the destination datastore.
- If a datastore that already contains the base or a shadow virtual machine is not available, vCloud Director makes a full clone to create a shadow virtual machine in a selected datastore, and then makes linked clones in the selected datastore.
- The latest model in vSphere Storage DRS takes linked clone sharing into account when calculating the effects of potential moves.
- Linked clones and virtual machines that are not linked clones can reside on the same datastores.

4.3.5.2. vSphere Storage DRS Limitations

- vCloud Director does not support creation, deletion, or modification of storage pods. These tasks must be performed at the vSphere level.
- vCloud Director does not support member datastore operations.
- Enabling vSphere Storage DRS for the datastore clusters used with vCloud Director is not supported if vSphere hosts are pre-vSphere 5.1.

4.4 vCloud Resource Sizing

Resource sizing for a vCloud depends on the corresponding service definition. A private vCloud service definition might not explicitly call out a required number of workloads to support. In this 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 sections come from *Service Definitions* and can assist with initial sizing of the vCloud environment.

Note Contact 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 use the reservation pool model and 50% use the pay-as-you-go allocation model. The 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.

The following table lists the virtual machine counts for the various virtual datacenters. The total virtual machine count of 1,500 reflects the specifications outlined in *Service Definitions* for the public vCloud service definition. 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 machine values are rounded up or down due to percentages.

Service Definitions also calls out the distribution for virtual machines in the vCloud with 45% small, 35% medium, 15% large, and 5% extra large. The following table shows the total amount of CPU, memory, storage, and networking needed.

Table 5. Memory, CPU, Storage, and Networking

Item	# of virtual machines	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 design guidelines for common consolidation ratios. The following table 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.
- Hyperthreading: Yes.
- Memory: 144GB.
- Networking: Dual 10GigE.

These calculations do not consider storage consumed by consumer or provider templates, nor do they take into account the resources consumed by vCloud Networking and Security Edge appliances. A vCloud Networking and Security Edge device backs each private organization virtual datacenter network and external routed organization virtual datacenter network.

The following are specifications for each vCloud Networking and Security Edge appliance:

- CPU: 1 vCPU compact, 2 vCPU large.
- Memory: 256MB compact, 1GB large.
- Storage: 200MB compact, 256MB large.
- Network: 1GigE (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 responsively, on demand.

When designing for scale, consider the maximums of the vCloud platform and the underlying vSphere platform. vCloud Director 5.1 requires vSphere 5.1 and has many platform improvements and enhancements. vCloud Director also introduces a number of 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) can result in slight performance differences.

The following table lists vCloud maximums based on a 10-cell configuration.

Table 7. vCloud Maximums

Constraint	Limit	Explanation
Virtual machines per vCloud Director	30000	Maximum number of virtual machines that can be resident in a vCloud instance.
Powered on per vCloud Director	10000	Number of concurrently powered on virtual machines permitted per vCloud instance.
Virtual machines per vApp	128	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	Maximum number of users supported by a single vCloud instance.
Concurrent users per vCloud Director	1500	Maximum number of current uses that can be logged into a single vCloud instance.
Organizations per vCloud Director	10000	Maximum number of organizations that can be created in a single vCloud instance.

Constraint	Limit	Explanation
vApps per organization	3000	Maximum number of vApps that can be deployed in a single organization.
Virtual datacenters per vCloud Director	10000	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	10000	Maximum number of logical networks that can be deployed in a single vCloud instance.
Routed Networks per vCloud Director	2000	Maximum number of routed networks that can be deployed in a single vCloud instance.
Catalogs per vCloud Director	10000	Maximum number of catalogs that can be created in a single vCloud instance.
Media Items per vCloud Director	1000	Maximum number of media items that can be created in a single vCloud instance.

See *Configuration Maximums for VMware vSphere 5.1* in the *VMware vSphere Documentation* (https://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-pubs.html), and the VMware Knowledge Base article, *vCloud Director 5.1 Configuration Maximums* (http://kb.vmware.com/kb/2036392), for more information on configuration maximums.

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. End 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 wherever possible. If complexity is unavoidable, document all required steps to guide the end users.

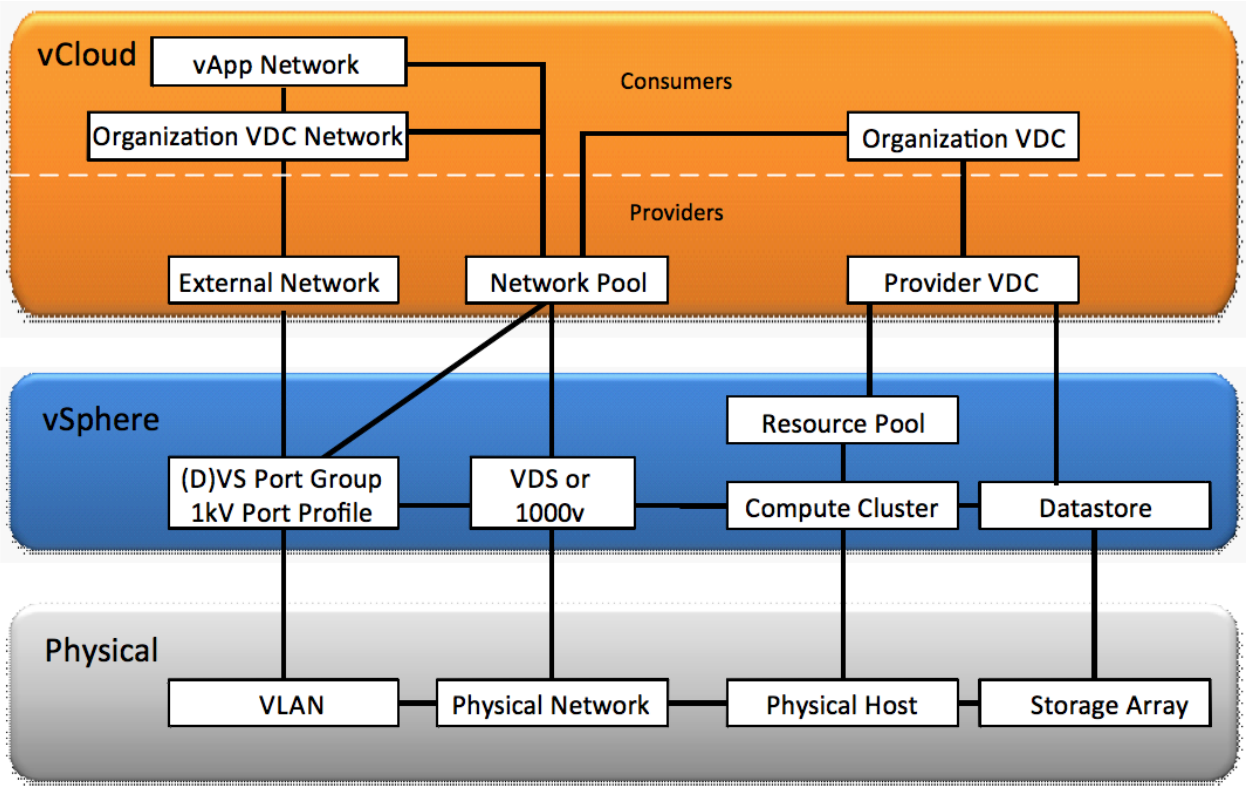
Taking a top-down approach to vCloud design necessitates understanding of the new abstractions introduced in the vCloud API and how they map to traditional vSphere objects.

5.1 vCloud Director Constructs

vCloud Director introduces logical constructs to facilitate multitenancy and provide interoperability between vCloud instances built to the vCloud API standard.

The following figure shows the abstraction mapping for vCloud Director.

Figure 8. Physical, Virtual, and vCloud Abstraction Mapping



The following table describes the logical constructs for vCloud Director.

Table 8. vCloud Director Constructs

Construct	Definition
Organization	The unit of multitenancy 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	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. An organization virtual datacenter allocates resources using one of the following models: • Pay As You Go. • Reservation Pool. • Allocation Pool.
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 virtual datacenter networks and are backed by port groups configured for Internet accessibility.
Organization virtual datacenter network	Organization virtual datacenter networks are instantiated through network pools and bound to a single organization. Organization virtual datacenter 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 virtual datacenter 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 translates into vCenter Server tasks.

5.2 Organizations

Organizations are the unit of multitenancy 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 might have multiple environments, such as development and 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 Director Installation and Upgrade 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 when the organization is created, maps to a unique URL that allows access to the UI for that organization. For example, an organization named Company1 maps to `https://<hostname>/cloud/org/Company1`. Use a standard naming convention for organization names and avoid using special characters or spaces because they can 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 design guideline is to create an administrative organization to provide a sandbox for system administrators and maintain 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:

- Do not allow the organization to 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. When creating an organization, specify policies for the total number of running and stored virtual machines according to the following definitions:

- *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.

*Lea*se 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, according to the following definitions:

- 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 the never expire option is selected, the storage is automatically cleaned up. Storage cleanup options include the following:

- Permanently deleted – After the lease expires, the vApp or vApp template is automatically deleted.
- Moved to expired items – 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 the 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.

See *Service Definitions* for guidance on provider virtual datacenter sizing. Consider the following:

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

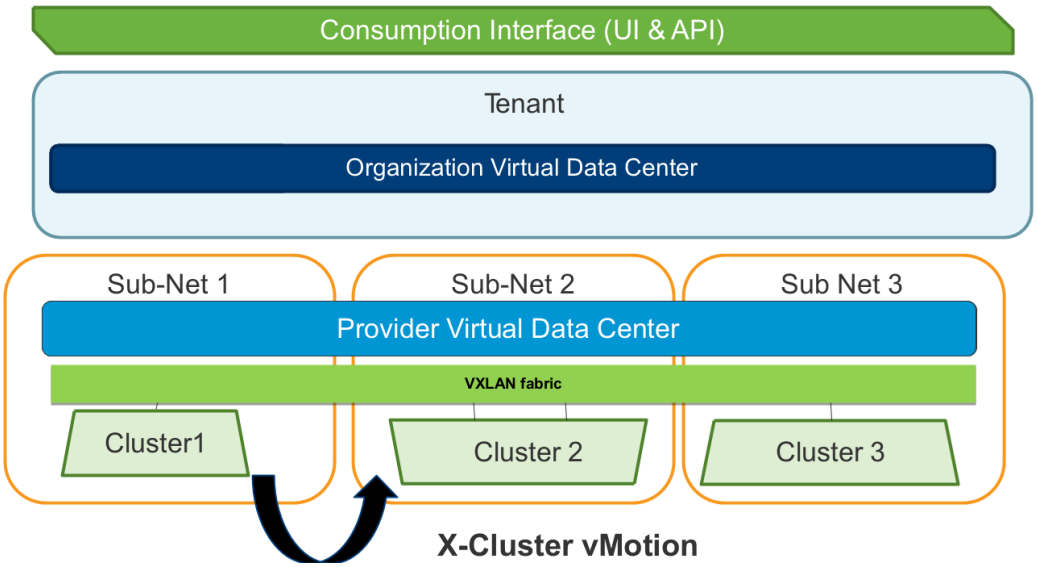
In some cases, a “special purpose" provider virtual datacenter dedicated to a single workload type might be needed. Special use case provider virtual datacenters are an example of what makes vCloud computing so flexible and powerful. The primary driver behind the need for a special purpose virtual datacenter is to satisfy the license restrictions imposed by software vendors that stipulate that all the processors that *might* run specific software must be licensed for it, regardless of whether they actually *are* running that software.

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

5.3.1 Elastic Virtual Datacenter

Rapid elasticity is a primary characteristic of vCloud computing. It involves quickly adding and releasing resources based on customer usage demands. vCloud Director supports compute elasticity by allowing provider virtual datacenters to span multiple clusters and by providing automatic placement of vApps. Aggregating capacity across multiple clusters into a single shared buffer offers potential for greater efficiency and utilization of the hardware.

Figure 9. Elastic Virtual Datacenters



Expansion of a provider virtual datacenter can occur in the following scenarios:

- Creation of an organization virtual datacenter.
- Increase in the size of an organization virtual datacenter.
- Powering on of a virtual machine or vApp.
- Resuming or unsuspending of a virtual machine or vApp.

The requested operation succeeds if enough non-fragmented compute capacity exists in the underlying provider virtual datacenter and network requirements are met.

The primary resource pool is the resource pool used in the initial creation of the provider virtual datacenter. Reservation 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. Adding the resource pools allows the virtual datacenter to draw resources from multiple sources.

The following are considerations for an elastic virtual datacenter:

- The datacenter can support Pay As You Go and allocation pool organization virtual datacenter types.
- Elasticity is limited to a single vCenter datacenter. A provider virtual datacenter can draw resources from resource pools created in the same vCenter datacenter as the primary resource pool.
- Existing provider virtual datacenters and organization virtual datacenters are upgraded to elastic automatically after upgrading to VMware vCloud Director 5.1.
- Organization virtual datacenters expand automatically in response to user consumption. Pay As You Go grows as needed and allocation pool grows to the allocated size.
- Clusters in a provider virtual datacenter can connect to a common network, which can be the same network or different networks connected through a VXLAN fabric.
- Newly added resource pools might connect to datastores that have not been added to the provider virtual datacenter. Add all visible datastores to the provider virtual datacenter.
- Use elastic virtual datacenter functionality to mitigate the eight-host cluster limit for fast provisioning on VMFS3 datastores. (Fast provisioning on VMFS5 datastores supports up to 32 hosts.)
- Do not add extra resource pools from the same compute cluster if it is already backing a provider virtual datacenter. Instead, increase the size of the existing resource pool that is mapped to the virtual datacenter.
- 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.

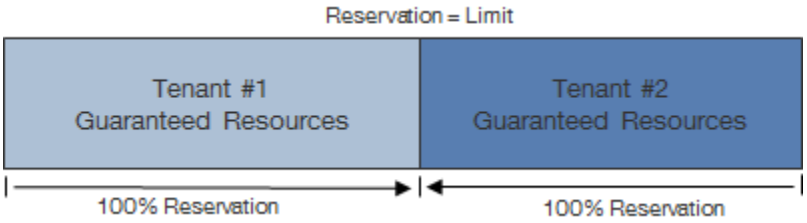
5.4.1 Allocation Models

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

5.4.1.1. Reservation Pool Model

Reservation Pool resources allocated to the organization virtual datacenter are completely dedicated. All guarantees are set to 100%. Reservation Pool virtual datacenters map to resource pools with the reservations set equivalent to the limits.

Figure 10. Reservation Pool

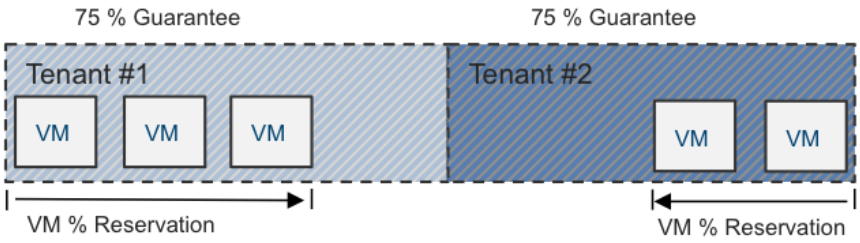


5.4.1.2. Allocation Pool Model

An allocation pool is a pool of allocated resources with a certain percentage of resources guaranteed. When an organization virtual datacenter is created using the Allocation Pool model, a dynamic resource pool is instantiated. This resource pool automatically adjusts available resources as new workloads are powered on based on the values specified within the virtual datacenter. Each value has a direct impact on how the related resource pool dynamically changes as new virtual machines are deployed. The following values are required when defining the size of an organization virtual datacenter.

- CPU allocation – The maximum amount of CPU resources available to the virtual machines running within the organization virtual datacenter.
- CPU resources guaranteed – The percentage of CPU resources guaranteed to be available to the running virtual machines.
- vCPU speed – The maximum speed in GHz that each vCPU can consume. A virtual machine with two vCPUs can consume twice this value.
- Memory allocation – The maximum amount of memory available to the virtual machines running within the organization virtual datacenter.
- Memory resources guaranteed – The percentage of the memory resources guaranteed to the running virtual machines.

Figure 11. Allocation Pool



An organization virtual datacenter is created with the following values:

- CPU allocation = 100 GHz.
- CPU resources guaranteed =75%.
- vCPU speed = 1 GHz.
- Memory allocation = 100GB.
- Memory resources guaranteed = 75%.

After the organization virtual datacenter is created, a corresponding resource pool is created. Unlike versions of vCloud Director earlier than version 5.1, a reservation is not initially set on the resource pool. The resource settings are dynamically changed as each new virtual machine is powered on within the organization virtual datacenter. The following formulas can be used to calculate resource pools reservations as virtual machines are deployed.

- CPU resources guaranteed * vCPU speed * # of virtual machine CPUs allocated = CPU reservation.
- CPU allocation = CPU limit.
- Memory resources guaranteed * virtual machine memory allocated = memory reservation.

Using the example values, when a virtual machine with 1 vCPU and 2GB of memory is powered on, the corresponding resource pool is updated based on the values defined at the organization virtual datacenter. In this example, the resource pool is dynamically set with the following values:

- CPU reservation = 750MHz.
- CPU limit = organization virtual datacenter CPU allocation.
- Memory reservation = 1536MB. This corresponds to 1 vCPU at 1GHz and 2GB of memory with a 75% guarantee.

Note A memory limit is not required because it not possible for a virtual machine to consume more memory than allocated.

When an additional virtual machine with the same configuration is powered on, the resource pool is again updated. In this example, the resource pool is set to the following configuration:

- CPU reservation = 1500MHz.
- CPU limit = organization virtual datacenter CPU allocation.
- Memory reservation = 3000MB.

This corresponds to 2 vCPUs at 2GHz and 4GB of memory with a 75% guarantee.

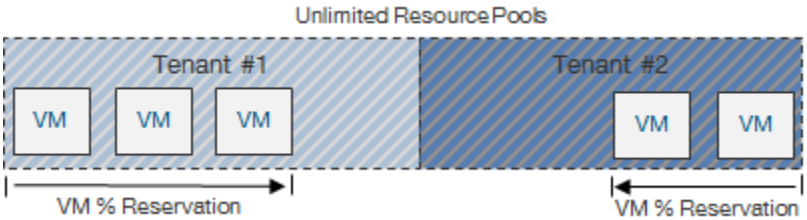
As additional virtual machines are deployed, the resource pool is dynamically updated until 100 vCPUs (100GHz) or 100GB of memory is allocated. After the CPU or memory allocation is reached, vCloud Director prohibits powering on additional virtual machines.

In vCloud Director 5.1.1, the CPU limit applied to the resource pool is set to the same value as the allocation of the virtual datacenter. This allows an organization virtual datacenter that has only a single virtual machine to consume as much as possible. In this case, it would be limited to the core speed of the physical CPU, as a 100GHz physical core currently is not available. As additional virtual machines are powered on, only the reservation changes on the resource pool. The limit remains equal to the virtual datacenter allocation. After the allocation is reached, vCloud Director prevents new virtual machines from powering on.

5.4.1.3. Pay As You Go Model

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



When an organization virtual datacenter is created, vCenter Server creates child resource pools with corresponding resource reservations and limits under the resource pool representing the organization 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 vCloud Director 5.1 introduces the ability to limit the resources allocated within a Pay As You Go organization virtual datacenter through CPU and memory resource limits. Previous versions of vCloud Director allowed capping only the number of virtual machines.

5.4.1.4. Mixed Allocation Models in a Provider Virtual Datacenter

A single provider virtual datacenter mapped to the cluster level can be configured with multiple allocation models for consumers based on their functional requirements. Creating a provider virtual datacenter model (Pay As You Go) does not guarantee that the same settings are applied across the organization virtual datacenter. The model changes the vSphere resource distribution in a similar manner to using multiple allocation models.

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 can provide an efficient way of allocating capacity. When using thin provisioning, additional management processes are required. Configure vCenter Server alarms to issue an alert when approaching an out-of-space condition, and provide for sufficient time to source and provision additional storage.

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

5.4.3 Fast Provisioning

Fast provisioning allows 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 means that full clones are created for subsequent vApp deployments.

Fast provisioning benefits include the following:

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

Fast provisioning includes the components:

- *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 virtual machine allows cross-datastore provisioning and is transparent to end users. Shadow virtual machines are created for vApp templates only, not for MyCloud vApps.

During fast provisioning, vApp files can reside 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 Server instances, as shown in the following table.

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.

The following are considerations for fast provisioning:

- Separate the datastores reserved for fast provisioning from the datastores reserved for full clones to improve performance and manageability.
- Fast provisioning requires vSphere 5.x.
- Fast provisioning is supported by vSphere Storage DRS with vCloud Director 5.1.
- Datastore selection is determined by vSphere Storage DRS when using VMware vCenter 5.1.
- Provisioning time is nearly instantaneous when provisioning to the same datastore.
- Using VMFS5 datastores removes the 8-host limit for fast provisioning (32-host maximum).
- Using VMFS3 datastores enforces an 8-host limit for fast provisioning.
- Provisioning a virtual machine to a different datastore triggers creation of shadow virtual machines if one does not already exist on the target datastore.
- Shadow virtual machines are full copies of the source virtual machines, which factors into sizing considerations for pre-provisioning shadow virtual machines across datastores.
- Storage array caching can boost linked clone performance. Ample storage array cache greatly benefits an environment that uses 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 virtual machines are treated differently from normal virtual machines and can be referenced through the vCloud API by the `SHADOW_VM` entity type.
- Invoke vSphere Storage vMotion migration of linked clones only through the vCloud API (`Relocate_VM` call). The target virtual datacenter must have visibility to the datastore that contains the source disks.
- Do not invoke vSphere Storage vMotion operations on linked clones through the vSphere Client, as doing so consolidates the linked-clones and might result in inconsistent behavior.

5.4.4 vApp Placement

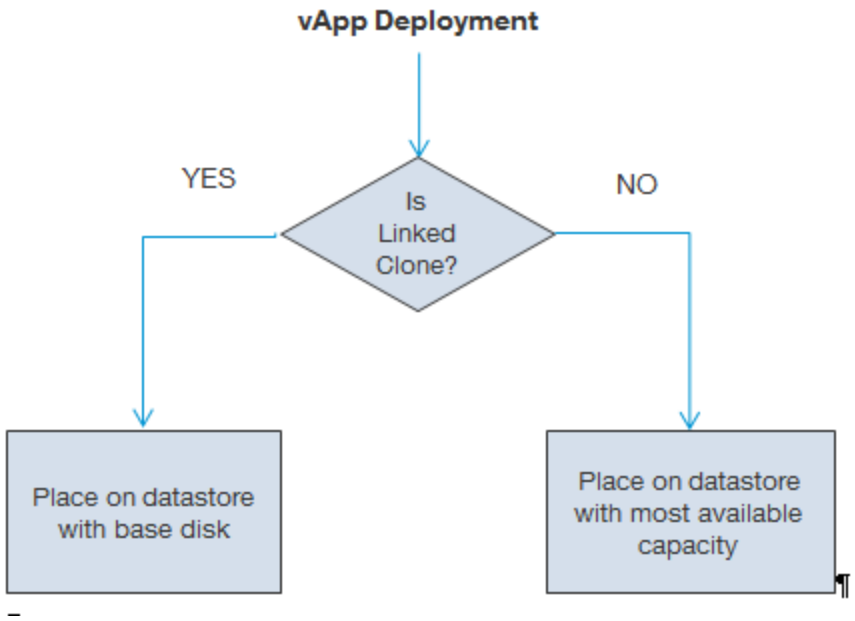
During vApp deployments, the vCloud Director virtual machine storage placement algorithm works as follows:

1. For fast provisioning-enabled virtual datacenters, identify 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 a base disk exists but the target datastore exceeds red threshold, look for a normal or yellow-threshold datastore. If no suitable datastores are available, the operation fails.
    - If a base disk exists but the 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.

The following figure charts the vApp placement algorithm used by the vCloud Director Placement Engine.

Figure 13. vCloud Director Placement Engine vApp Placement Algorithm



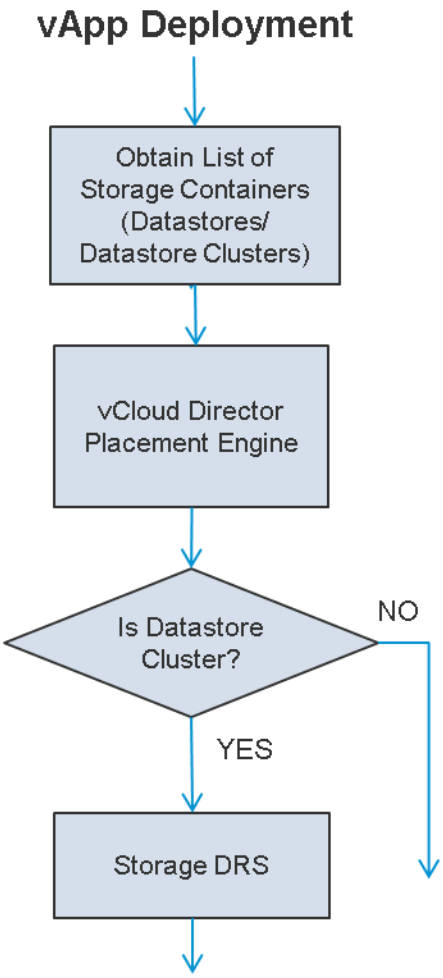
vApp creation fails if the vApp contains multiple virtual machines that cannot fit on a single datastore in the target virtual datacenter. Consider the following scenario:

- Virtual datacenter1:
  - Datastore1 – 20GB free space.
  - Datastore2 – 30GB free space.
- vApp1:
  - VM1 – 30GB.
  - 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 design guidelines for datastore utilization through proactive monitoring and storage maintenance.

When vCenter 5.1 is used in combination with vCloud Director 5.1, the vCenter 5.1 vSphere Storage DRS datastore placement engine is used in lieu of the vCloud Director placement engine when datastore clusters are available as a deployment target.

Figure 14. vCloud Director Storage Placement



5.4.5 Public vCloud Considerations

The public service definition requirements used in this example are taken from *Service Definitions*.

Table 10. Public vCloud Virtual Datacenter Requirements

Requirements
Three different service offerings are required: Basic (Pay As You Go), Committed (Allocation Pool), and 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 service definitions are as follows:

- 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 has specific requirements for the maximum number of virtual machines each organization can have based on size. Refer to the public 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. Chargeback integrates with vCloud Director to provide metering and cost calculation functionality. For more information, see the *User's Guide* in the *vCenter Chargeback Manager Documentation* (https://www.vmware.com/support/pubs/vcbm_pubs.html).

Best Practices and Troubleshooting Guide (http://www.vmware.com/support/pubs/vcbm_pubs.html) describes recommended practices around vCloud Director and vCenter Chargeback integration to accommodate instance-based pricing (pay-as-you-go), reservation-based pricing, and allocation-based pricing.

5.4.6 Private vCloud Considerations

The private service definition requirements used in this example are from service definitions.

Table 11. Private vCloud Virtual Datacenter Requirements

Requirements
Three different service offerings are required: Basic (Pay As You Go), Committed (Allocation Pool), and 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 specified storage limit except when using the Pay As You Go allocation model, for which the storage limit can be unlimited. For this example, no storage limit is set because the static storage values that are provided for individual virtual machines limit the number of virtual machines in the 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 network maps to a vSphere port group with external connectivity.
- Internal or routed networks are used to facilitate communication between virtual machines within a vCloud instance. 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 vCloud Networking and Security Manager to provide automated network security for a vCloud environment. vCloud Networking and Security Edge gateway devices are deployed during the provisioning of routed or private networks. Each vCloud Networking and Security Edge gateway runs a firewall service that allows or blocks inbound traffic to virtual machines that are connected to a public access organization virtual datacenter 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 or routed organization and vApp networks. Networks created from network pools must be isolated at Layer 2.

The following 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.
- Virtual eXtensible LAN (VXLAN) network pools use a Layer 2 over Layer 3 MAC in UDP encapsulation to provide scalable, standards based traffic isolation across Layer 3 boundaries (requires distributed switch).
- VLAN-backed network pools are backed by a range of pre-provisioned VLAN IDs. For this arrangement, all specified VLANs 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 that is implemented through encapsulation techniques that span hosts and provides traffic isolation from other networks (requires distributed switch).

The following table compares the options for a network pool.

Table 12. Network Pool Options

Consideration	vSphere Port Group Backed	VXLAN Backed	VLAN-Backed	vCloud Network Isolation-Backed
How it works	Isolated port groups must be created and exist on all hosts in cluster.	• Multicast address is mapped to a VXLAN segment ID for isolation. • Virtual machine to virtual machine traffic is tunneled over a Layer 3 network by a VTEP (ESXi hosts). • Node learning done through multicast, not broadcast.	• Uses range of available VLANs dedicated for vCloud. • Network isolation relies on inherent VLAN isolation.	Creates an overlay network (with fence ID) within a shared transport network.
Advantages	N/A	• Does not rely on VLAN IDs for isolation. • Works over any Layer 3 multicast-enabled network. • No “distance” restrictions, managed by multicast radius.	• Best network performance. • vCloud Director creates port groups as needed.	• Scalable to create thousands of networks per transport network. • More secure than VLAN-backed option due to vCloud Director 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.	End-to-end multicast required.	• VLANs are a limited commodity (4096 maximum). • Requires used VLANs to be configured on all associated physical switches. • Scoped to a single virtual datacenter and vCenter Server.	Overhead required to perform encapsulation.

5.5.2.1. vSphere Port Group-Backed Considerations

- Use standard or distributed virtual switches.
- vCloud Director does not automatically create port groups. Manually provision port groups for vCloud Director to use ahead of time.

5.5.2.2. VXLAN-Backed Considerations

- Distributed switches are required.
- Configure the MTU to be at least 1600 at ESXi and on the physical switches to avoid IP fragmentation.
- Map the guest MTU size to accommodate the VXLAN header insertion at the ESXi level.
- Use explicit failover or “route based on IP hash” as the load balancing policy.
- If VXLAN transport is traversing routers, multicast routing must be enabled (PIM – BIDIR or SM).
    - More multicast groups are better.
    - Multiple segments can be mapped to a single multicast group.
    - If VXLAN transport is contained to a single subnet, IGMP Querier must be enabled on the physical switches.
    - Use BIDIR-PIM where available so any sender can be a receiver as well. If BIDIR-PIM is not available, use PIM-SM.
- If VXLAN traffic is traversing a router, enable proxy ARP on the first hop router.
- Use five-tuple hash distribution for uplink and interswitch LACP.

5.5.2.3. VLAN-Backed Considerations

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

5.5.2.4. 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. The information 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 (optional, but recommended for security). If no VLAN ID is specified, it defaults to VLAN 0.
- The maximum number of VCD-NI-backed network pools per vCloud instance is 10.
- vCloud Director automatically creates port groups on distributed switches as needed.

5.5.3 vCloud Networking and Security Edge Gateway

vCloud Networking and Security Edge gateways provide external network connectivity to vCloud consumers. The gateways are first class entities that are associated with an organization virtual datacenter. They cannot be shared across other organization virtual datacenters. Each has up to 10 interfaces and can be connected to multiple external networks.

5.5.4 Organization Virtual Datacenter Networks

Organization virtual datacenter networks provide network connectivity to vApp workloads within an organization. Users in the organization connect to outside networks through external organization virtual datacenter networks, similar to how users in an organization connect to a corporate network that is uplinked to an Internet service provider. During creation, you can specify whether organization virtual datacenter networks are specific to a virtual datacenter or shared with all of the organization's virtual datacenters (as in vCloud Director 5.1).

Connectivity options for organization virtual datacenter networks include:

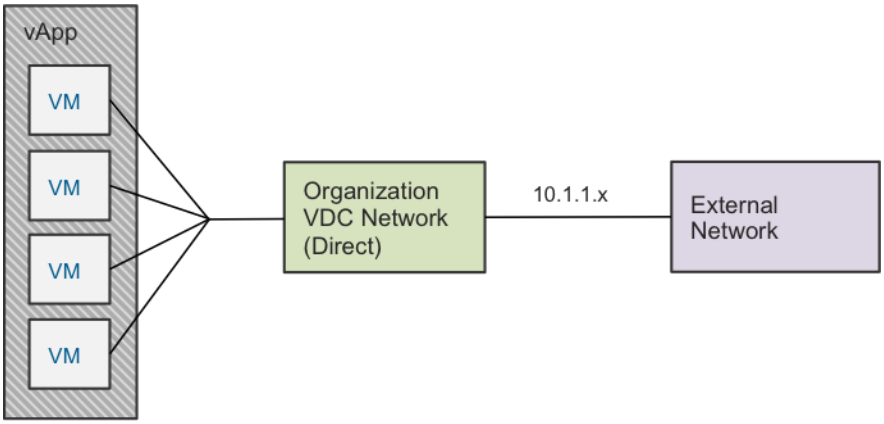
- External direct connect organization virtual datacenter network.
- External routed organization virtual datacenter network.
- Internal isolated organization virtual datacenter network.

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

5.5.4.1. Direct

In a directly connected external organization virtual datacenter network, the vApp virtual machines are in the port group of the external network. IP address assignments for vApps follow the external network IP addressing.

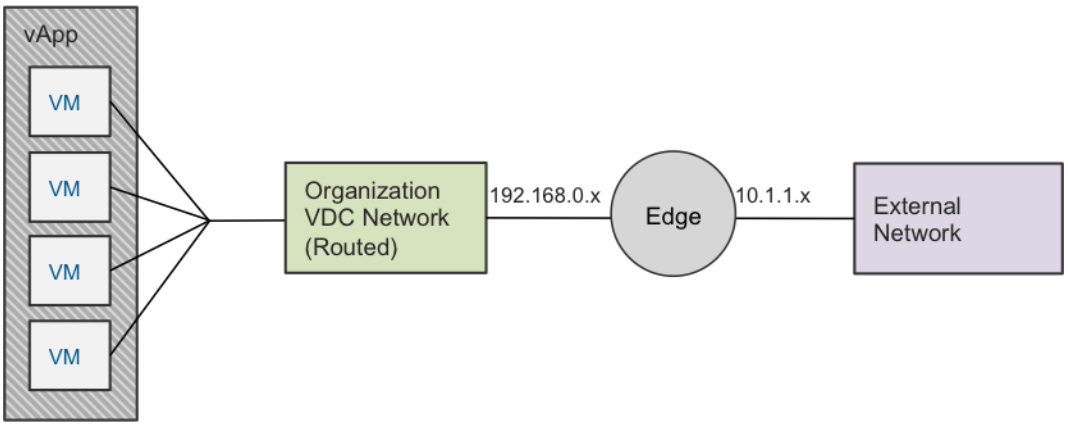
Figure 15. External Organization Virtual Datacenter Network (Direct)



5.5.4.2. Routed

A routed external organization virtual datacenter network is protected by a vCloud Networking and Security Edge device that provides DHCP, firewall, NAT, VPN, and static routing services. The vCloud Networking and Security Edge device connects to the organization virtual datacenter network and the external network port groups.

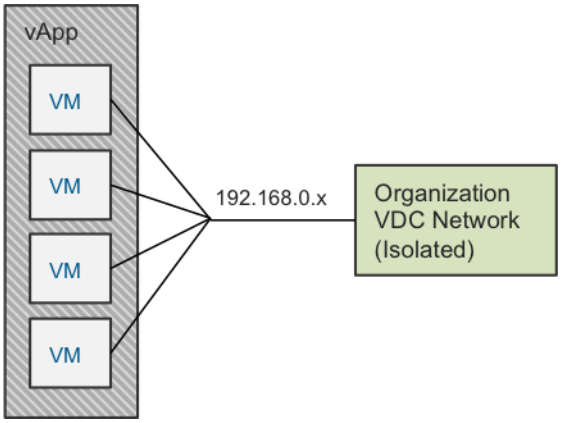
Figure 16. External Organization Virtual Datacenter Network (Routed)



5.5.4.3. Isolated

An internal organization virtual datacenter network is isolated from all other networks.

Figure 17. Internal Organization Virtual Datacenter Network (Isolated)



5.5.5 vApp Networks

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

Connectivity options for vApp networks include the following:

- Direct – vApps connect directly to the organization virtual datacenter network.
- Fenced – Identical virtual machines can exist in different vApps. A virtual router provides isolation and proxy ARP.
- Routed – A new network is defined. A virtual router provides NAT and firewall functionality.
- Isolated – Communication is restricted to the virtual machines in the vApp. No connection exists to an organization virtual datacenter network.

vApp networks are created as follows:

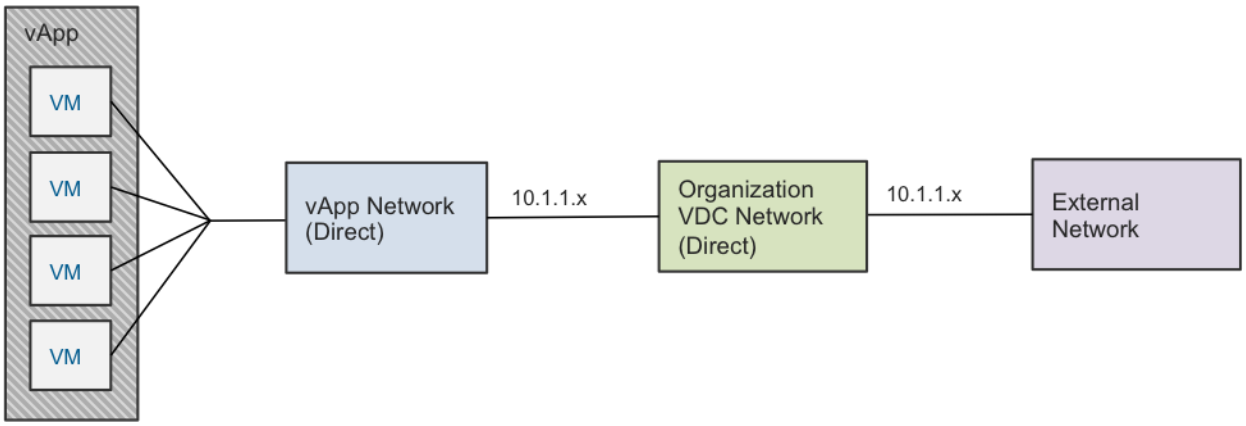
- Manually create vApp networks using the **Add Network** wizard. Connecting the vApp network to an organization virtual datacenter network creates a routed connection, with configurable NAT and firewall services.
- Fencing a vApp directly connected to an external or organization virtual datacenter network. Choosing the **fence** option associates an implicit vApp network to the vApp. Firewall and NAT services are configurable on a fenced network.

5.5.5.1. Direct

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

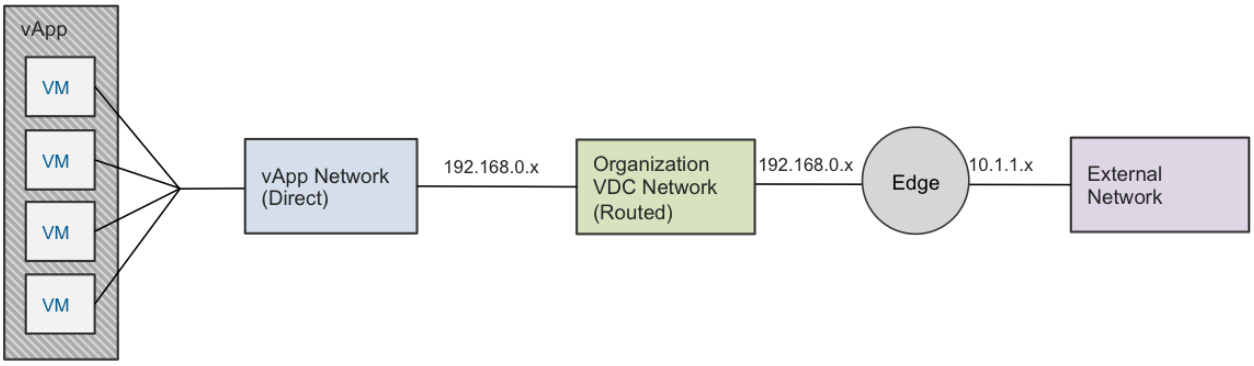
The following figure shows a vApp network directly connected to a direct external organization virtual datacenter network.

Figure 18. vApp Network (Direct) for Organization Virtual Datacenter Network (Direct)



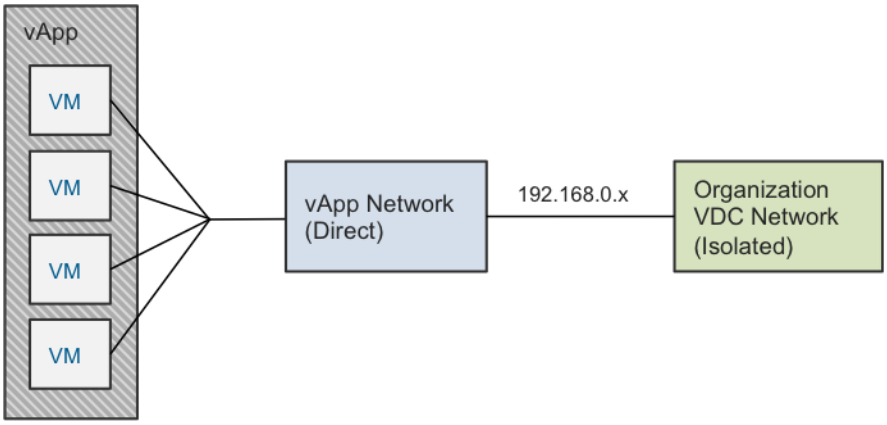
The following figure shows a vApp network directly connected to a routed external organization virtual datacenter network. vCloud Networking and Security Edge provides DHCP, firewall, NAT, and static routing services to the organization virtual datacenter network.

Figure 19. vApp Network (Direct) for Organization Virtual Datacenter Network (Routed)



The following figure shows a vApp network directly connected to an isolated organization virtual datacenter network. A vCloud Networking and Security Edge automatically deploys only if using DHCP services.

Figure 20. vApp Network (Direct) for Organization Virtual Datacenter Network (Isolated)



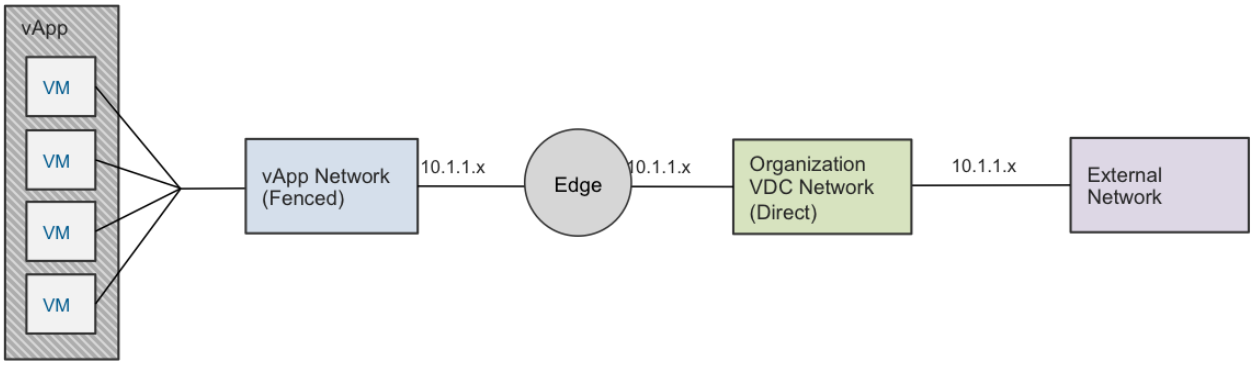
5.5.5.2. Fenced

For a fenced network, the external and internal IP subnet is the same, with proxy ARP used to move traffic. vCloud Networking and Security 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 virtual datacenter network.

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

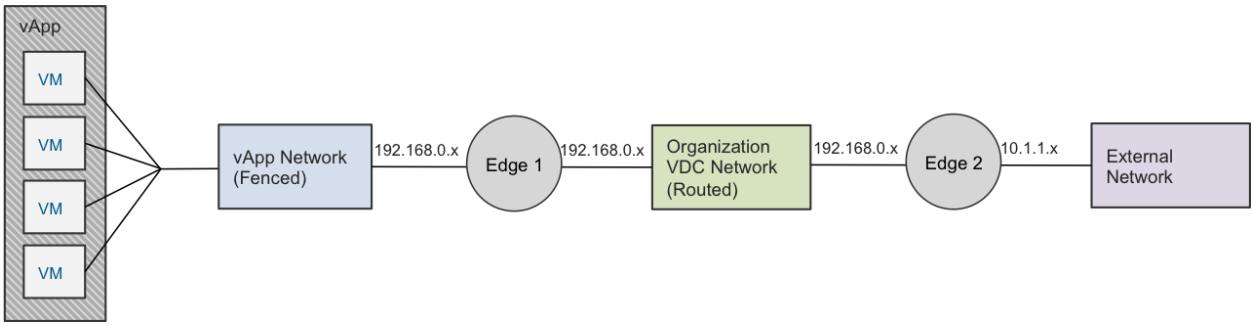
The following figure illustrates a scenario where a vApp network connected to a direct organization virtual datacenter network is fenced.

Figure 21. vApp Network (Fenced) for Organization Virtual Datacenter Network (Direct)



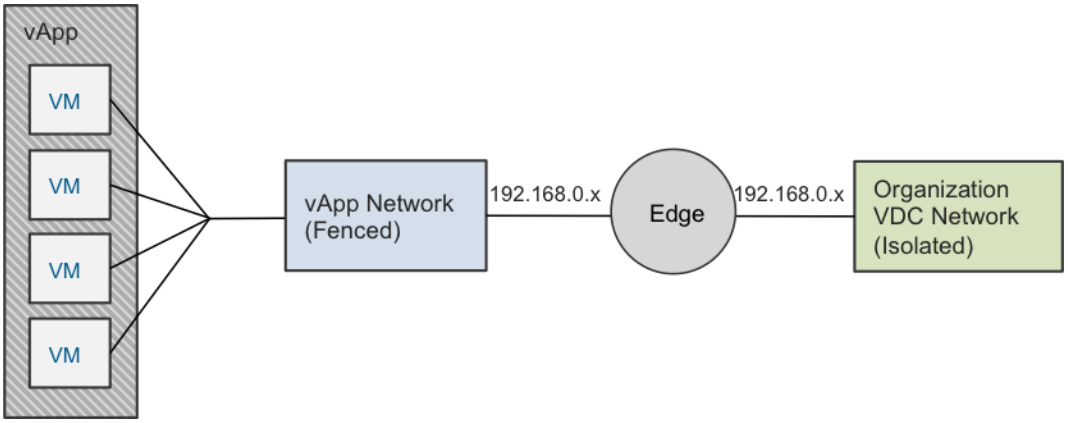
If you are fencing a vApp network connected to a routed organization virtual datacenter network, double NAT occurs with two vCloud Networking and Security Edge instances deployed. The following figure illustrates this scenario.

Figure 22. vApp Network (Fenced) for Organization Virtual Datacenter Network (Routed)



The following figure shows a fenced vApp network connected to an isolated organization virtual datacenter network. There is only one NAT.

Figure 23. vApp Network (Fenced) for Organization Virtual Datacenter Network (Isolated)



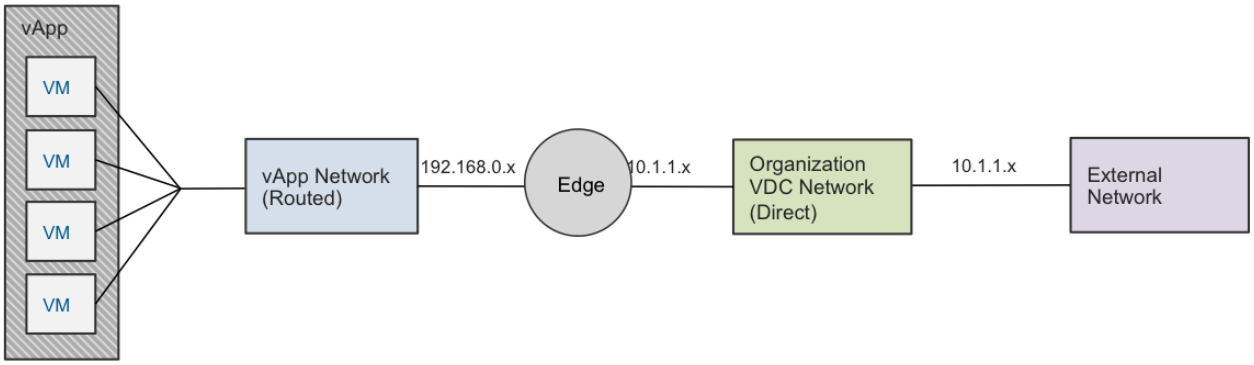
5.5.5.3. Routed

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

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

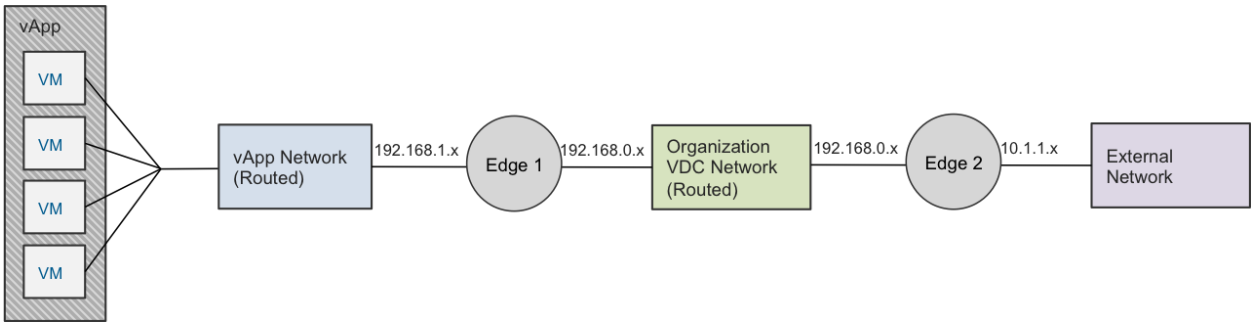
The following figure illustrates a scenario where a routed vApp network connects to a direct organization virtual datacenter network.

Figure 24. vApp Network (Routed) for Organization Virtual Datacenter Network (Direct)



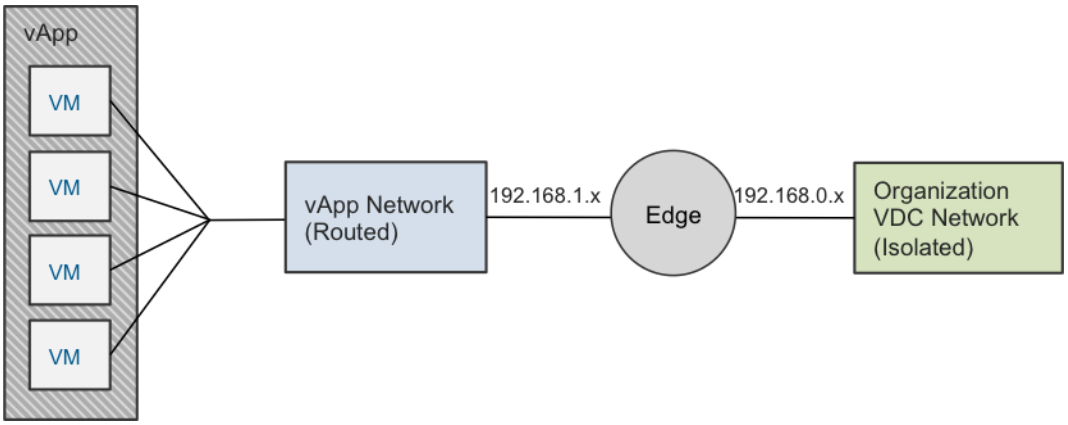
If a routed vApp network connects to a routed organization virtual datacenter network, double NAT occurs with two vCloud Networking and Security Edge instances deployed. The following figure illustrates this scenario.

Figure 25. vApp Network (Routed) for Organization Virtual Datacenter Network (Routed)



The following figure shows a routed vApp network connected to an isolated organization virtual datacenter network.

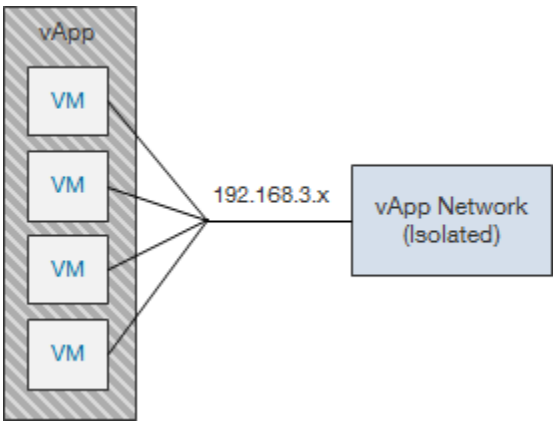
Figure 26. vApp Network (Routed) for Organization Virtual Datacenter Network (Isolated)



5.5.5.4. 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 27. vApp Network (Isolated)



5.5.6 Static Routing

Static routing support in vCloud Director provides the ability to route between network segments without using NAT and allows increased flexibility in implementing network connectivity within a vCloud environment.

Although most networks have a directly connected default gateway, it is possible for networks to have more than one route (for example, when using multiple interfaces on vCloud Networking and Security 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 virtual datacenter network level and vApp network level.

- For vCloud Networking and Security Edge gateway instances that are connected to multiple external networks and organization virtual datacenter networks, routes can be applied on the entire vCloud Networking and Security Edge gateway or on any one of the external networks connected to the gateway.
- For vApp networks that are route-connected to an external network, the static routing configuration is simplified as routes are applied only on the external interface.

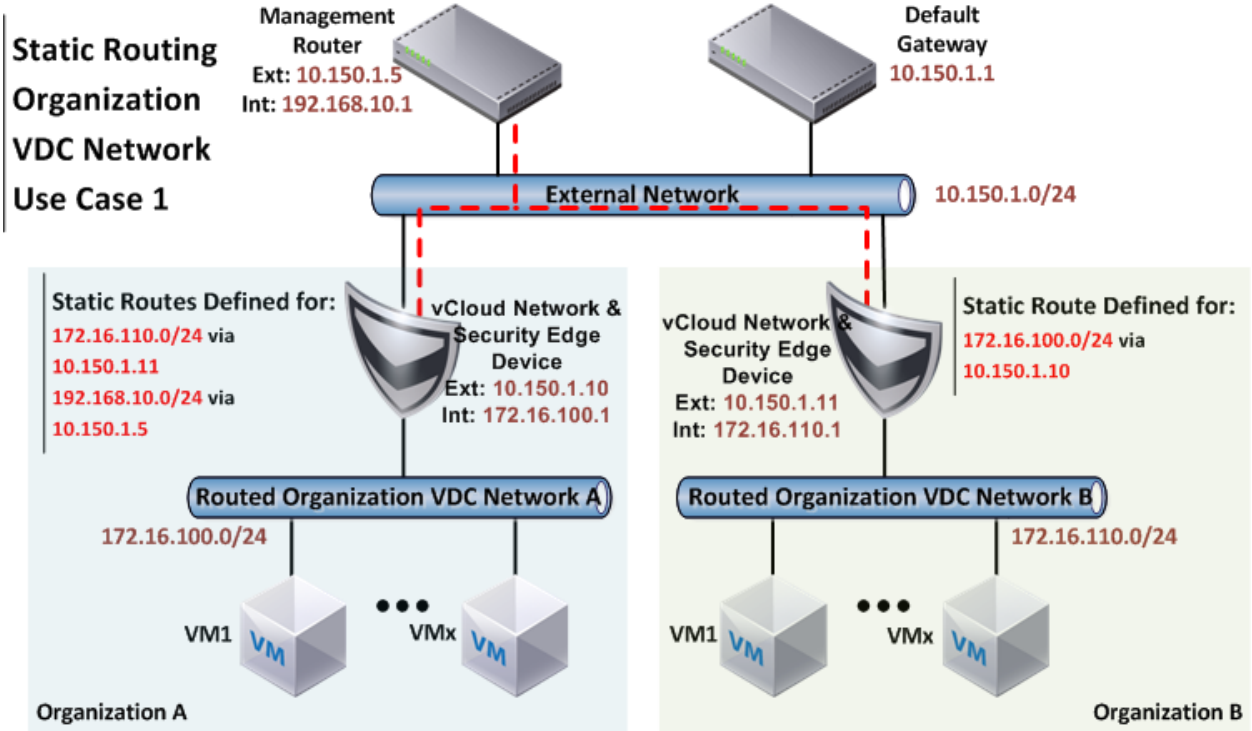
5.5.7 Static Routing Organization Virtual Datacenter Network Use Cases

The following use cases demonstrate the different options for static routing with vCloud Director.

5.5.7.1. 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 involves access to a remote management network through a VPN or proxy, or by accessing services in another organization.

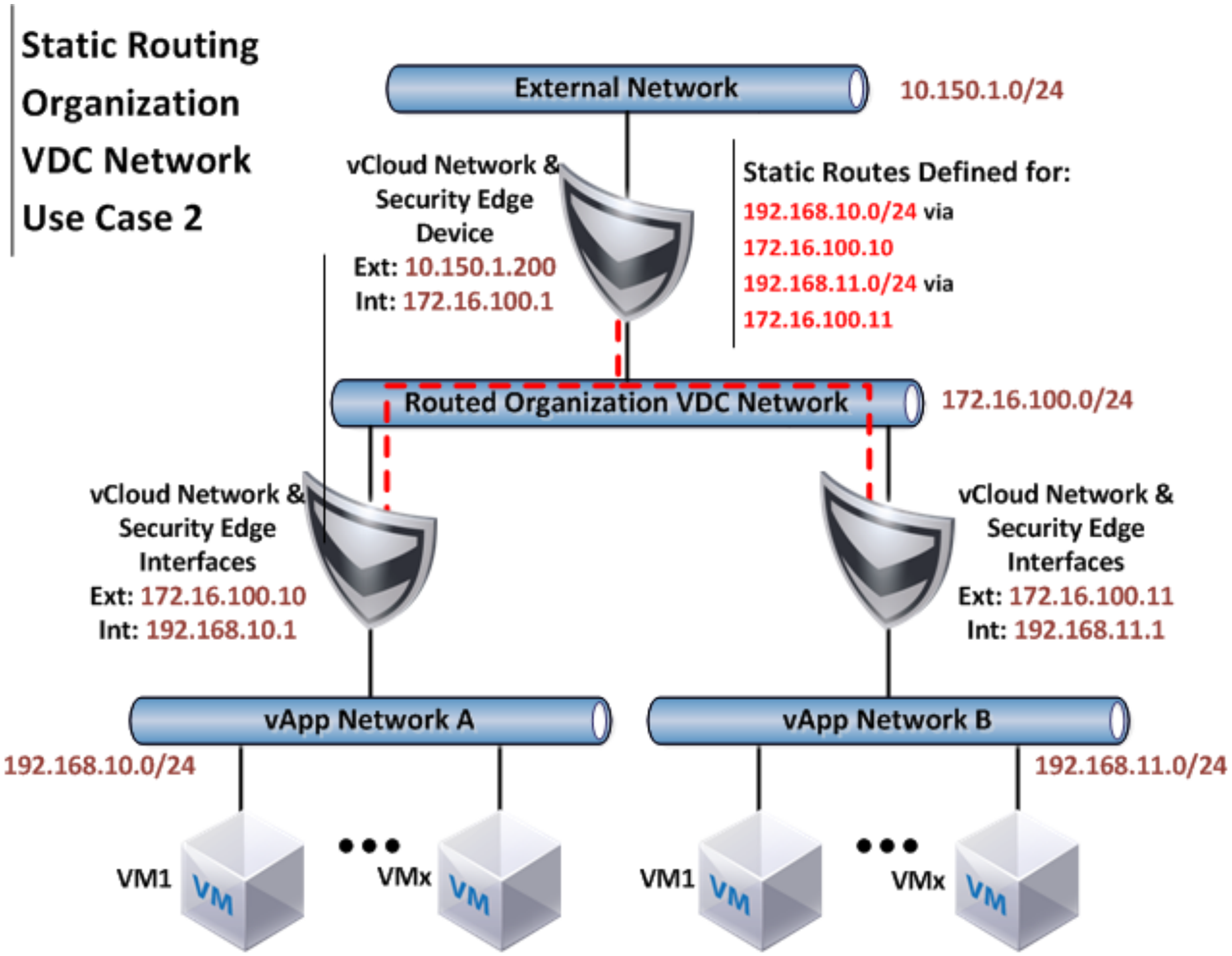
Figure 28. Organization Virtual Datacenter Network Static Routing Use Case 1



5.5.7.2. Enabling vApp Networks Connected to an Organization Virtual Datacenter Network to Communicate Directly

This use case allows virtual machines connected to different vApp networks (in a common organization virtual datacenter network) to communicate without NAT. This configuration reduces the operational overhead of maintaining port forwarding or IP translation NAT rules for connectivity within the organization.

Figure 29. Organization Virtual Datacenter Network Static Routing Use Case 2



5.5.7.3. Reducing Layers of NAT from External Networks to vApp Networks

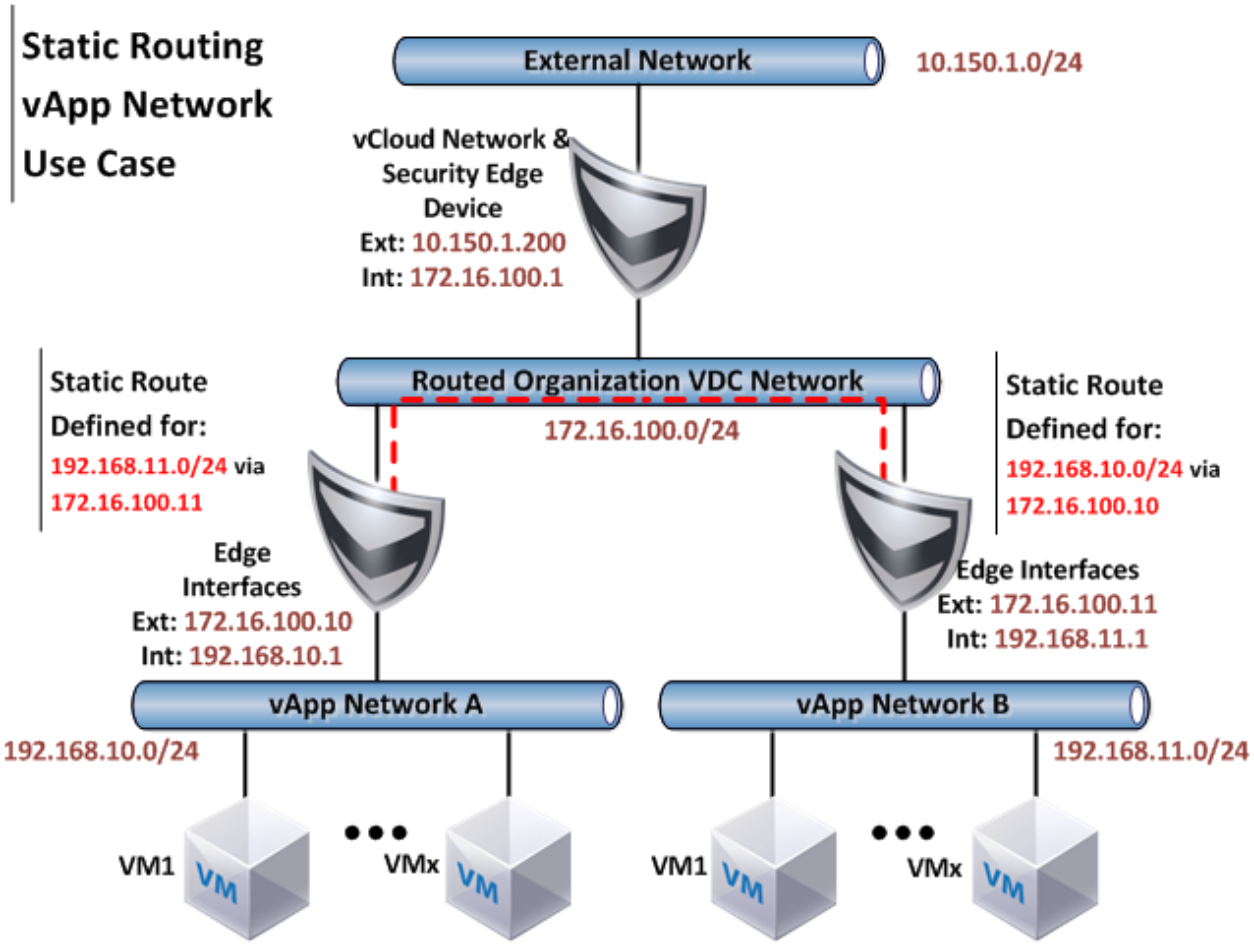
In vCloud Director 1.0, up to two levels of NAT are required for a system outside the vCloud environment to access services on a virtual machine connected to a vApp network. One NAT level is required if the organization virtual datacenter network is directly connected, and two are required if the organization virtual datacenter 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. Because these routing capabilities are delivered through vCloud Networking and Security Edge, the self-service firewall management is still maintained. This is important in private vCloud deployments where networks are typically flatter to support the centralized services, and static routing is an alternative to directly connecting virtual machines to the external networks.

5.5.8 Static Routing vApp Network Use Cases

5.5.8.1. Enabling vApp Networks Connected to an Organization Virtual Datacenter Network to Communicate Directly

This scenario provides connectivity similar to the use case shown in the following figure.

Figure 30. vApp Network Static Routing Use Case



If vApp level static routing is configured, enable **Always use assigned IP addresses until this vApp or associated networks are deleted** so that the next hop addresses for static routes does not change while vApps are powered off.

There is an overlap between organization virtual datacenter network Use Case 2 (Figure 29) and the vApp network use case, so it is important to understand the advantages and disadvantages of both configurations:

- Applying static routes at the organization virtual datacenter network consolidates management to a common view, but requires all traffic to pass through the organization's vCloud Networking and Security Edge device.
- vApp network static routes allow traffic to flow directly between the vApps that provide the highest performance.
- Static routing at the vApp network layer supports scenarios where the organization virtual datacenter network is directly connected.

Although it is required to provide connectivity between vApps without address translation, VMware recommends that you apply static routes in the vApp network vCloud Networking and Security Edge device. Unlike NAT, static routing does not support overlapping network ranges. If there are plans to leverage static routing within the vCloud environment, allocated IP addresses for organization and vApp networks must be unique.

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

Consider the following limitations when using static routing with vCloud Director:

- Static routing is supported only with vCloud Networking and Security Edge 5.0 (or later).
- Static routing is limited to a maximum of 64 static routes per vCloud Networking and Security Edge device.
- Dynamic routing protocols are not currently supported.
- Static routing does not apply to fenced vApps.

5.5.9 Third-Party Distributed Switch Considerations

vCloud Director 5.1 enhances third-party distributed switch integration by extending support for all the network pool types. Port-group backed, VXLAN backed, VLAN backed, and VCD-NI based network pools are available for creation with a supported third-party distributed switch.

5.6 Networking – Public vCloud Example

The public service definition requirements used in this example are from *Service Definitions*.

Table 13. Public vCloud Network Requirements

Requirements
Pool of eight public routable IP addresses for each tenant.
Minimum of one routed organization virtual datacenter 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, do the following:

- Map to a vSphere port group that is configured for Internet connectivity.
- Provide the network configuration details, including subnet mask, 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.

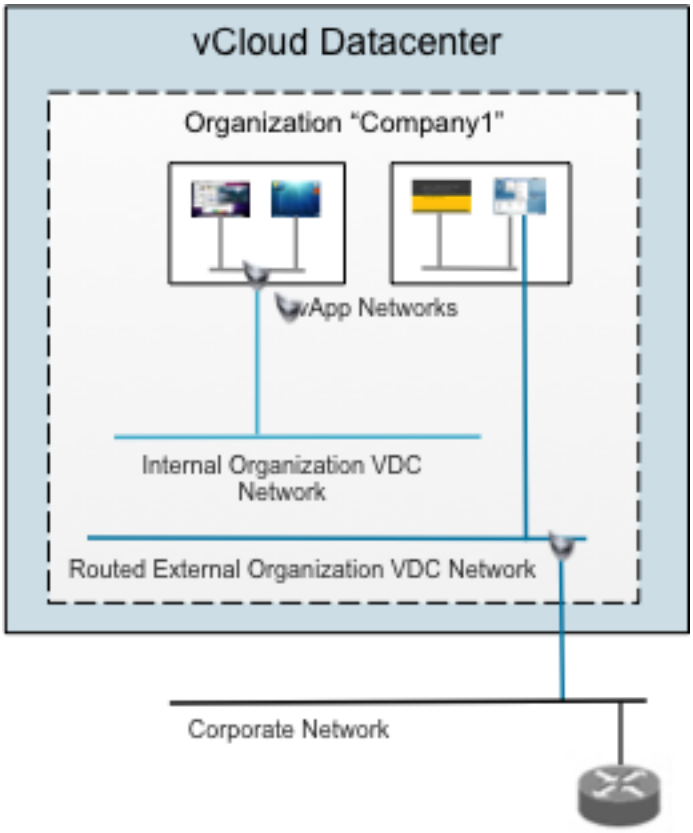
5.6.2 Network Pools

Each organization in a public vCloud requires individual private networks. vCloud Director instantiates isolated Layer 2 networks using network pools.

Create a single vCloud Director VXLAN network pool for all organization virtual datacenter network deployments. VXLAN requires the use of a distributed switch.

Network pools handle the automated creation of organization virtual datacenter networks and vApp networks. A minimum of 12 networks is required in the network pool per organization, with 10 reserved for vApp networks and 2 used for organization virtual datacenter 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 10,000 direct connect vCloud datacenter networks or 2,000 routed vCloud datacenter networks.

Figure 31. Example of Public vCloud Networking



5.6.3 Organization Virtual Datacenter Networks

Create two different organization virtual datacenter networks for each organization: one routed external organization virtual datacenter network, and one internal organization virtual datacenter network. The **Create Organization Network** wizard provides the option of creating these two organization virtual datacenter networks in one workflow. When naming an organization virtual datacenter network, start with the organization name and a hyphen, for example, “Company1-Internet.”

The routed external organization virtual datacenter network uses vCloud Networking and Security Edge for firewall and NAT services to isolate organization traffic from other organizations that share the same external provider network. Both the external organization virtual datacenter network and the internal organization virtual datacenter networks are instantiated from the previously established vCloud Director Network Isolation network pool. Each organization virtual datacenter 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 pool. The static IP address pool can be as large as desired. Typically an RFC 1918 class C is used.

The last step is to add external public IP addresses to the vCloud Networking and Security Edge configuration on the external organization virtual datacenter network. Using the **Configure Services** interface, add eight public IP addresses to an external organization virtual datacenter network. The IP addresses listed come from the external network static IP address pool.

5.7 Networking – Private vCloud Example

The private service definition requirements used in this example are from *Service Definitions*.

Table 14. Private vCloud Network Requirements

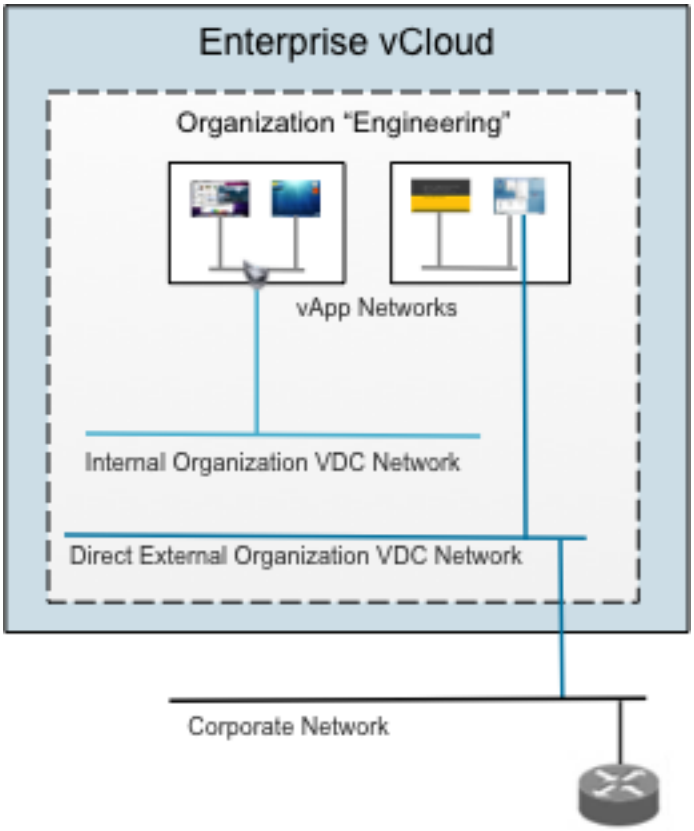
Requirements
vApps require a direct connection to the external network due to upstream dependencies.
An isolated network is needed for development, test, and pre-production workloads.
Users have the ability to self-provision networks.

5.7.1 External Networks

Private vCloud networking requirements tend to vary depending on the primary use cases driving the project. Enterprises that act as service providers to their internal customers tend to have comparable network requirements to that of a public vCloud. Enterprises that use vCloud for development or pre-production environments 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 virtual datacenter network and a direct connect external organization virtual datacenter network.

Figure 32. Example of Private vCloud Networking



At least one external network is required for external organization virtual datacenter networks to access resources outside of 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 static assignment.

5.7.2 Network Pools

The requirements call for one internal organization virtual datacenter network and the ability for consumers to create private vApp networks. No minimum number of vApp networks is defined, but typically organizations start with around 10. Size the network pool to be the number of organizations multiplied by 11. VMware recommends setting the maximum number of networks per network pool to 2,000 routed or 10,000 direct connect networks.

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 virtual datacenter network using the **Create Organization Network** wizard, and select **direct connection** from the drop-down menu. vApps that

connect to this organization virtual datacenter network are dropped directly on the vSphere port group that corresponds to the external network.

Implementing routed networking can add complexity to the network design. For more information about adding network options, see the *vCloud Director Administrator's Guide* in the *VMware vCloud Director Documentation* (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 to upload new vApps or media to an individual organization.

Guest customization changes the identity of the vApp and can be used for 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 list of recommended templates. vApp templates usually 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, VMware vSphere vApps™ and vCloud vApps have subtle differences. 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

The following are general design considerations for vApps:

- Use a default of one vCPU unless requirements call for more (such as a multithreaded 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.
- Virtual hardware versions 7, 8, and 9 are supported, depending on the vSphere host version backing the hosts in the provider virtual datacenter. Virtual hardware version 9 is supported in vSphere 5.1.
- Verify that the virtual machine virtual hardware version matches the highest required version within the provider virtual datacenter. The highest version chosen is the highest available with the provider virtual datacenter.
- Use standard virtual machine naming conventions.

5.8.1.1. Virtual Hardware Version 9

vCloud Director 5.1 exposes the highest version of virtual hardware available in the provider virtual datacenter. Users can choose the virtual hardware version desired up to the latest version supported by the provider virtual datacenter for their organization virtual datacenter. vSphere 5.1 supports the use of virtual hardware version 9.

Virtual hardware version 9 provides capabilities to vCloud vApps such as Windows 8 XP mode, 64-bit nested virtualization, and CPU-intensive workloads.

- Windows 8 XP mode – XP mode allows a virtualized XP instance to run for compatibility with older applications that do not natively run on Windows 8. Users who need to run XP mode in Windows 8 must choose an organization virtual datacenter that is backed by a provider that allows version 9 virtual hardware. After specifying virtual hardware version 9, the user must also enable the Nested HV feature.
- 64-bit nested virtualization – Hyper-V and virtualized ESXi nested virtualization can be helpful for non-production use cases such as training and demonstration environments. Virtualized Hyper-V or virtualized ESXi running nested 64-bit virtual machines requires virtual hardware version 9 with the Nested HV feature enabled.
- CPU-intensive workloads – Users who need to run an extremely CPU-intensive workload in a virtual machine that requires 32 to 64 vCPUs must use virtual hardware version 9.

5.8.2 Differences between vSphere and vCloud Director vApps

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

```
<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 supports, 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 might 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.

5.9 Snapshots

vCloud Director 5.1 provides full support for snapshot functionality. This section discusses snapshots, the impact they have on the underlying infrastructure, and the considerations to take into account before enabling snapshot functionality in a vCloud environment.

5.9.1 Snapshot Architecture

A snapshot preserves the state and data of a virtual machine at a specific point in time:

- The state includes the virtual machine's power state (powered-on, powered-off, suspended).
- The data includes all of the files that make up the virtual machine.

Snapshots work by creating delta copies (point-in-time) of the specified virtual machine files. The following figure provides a high-level illustration of how the process works.

Figure 33. Snapshot Processing

Each snapshot is composed of the following files, where <vm> is the name of the virtual machine and <number> identifies the specific snapshot:

- `<vm>-<number>.vmdk` and `<vm>-<number>-delta.vmdk`.

  A collection of `.vmdk` and `-delta.vmdk` files for each virtual disk is connected to the virtual machine at the time of the snapshot. These files are referred to as *child disks*, *redo logs*, or *delta links*. Child disks can later become parent disks for future child disks. From the original parent disk, each child constitutes a redo log that points back from the present state of the virtual disk, one step at a time, to the original one.

  Note The <number> value might not be consistent across all child disks from the same snapshot. The file names are chosen based on filename availability.

- `<vm>.vmsd`

  The `.vmsd` file is a database of the virtual machine's snapshot information and the primary source of information for the snapshot manager. The file contains line entries that define the relationships between snapshots as well as the child disks for each snapshot.

- `<vm>Snapshot<number>.vmsn`

  These files record the memory state at the time of the snapshot.

5.9.2 Snapshot Use Cases

The following are primary use cases for using snapshots in a vCloud environment:

- Production backups.
- Development/test environments.
- Third-party backup integration.

5.9.2.1. Production Backups

Do not use snapshots as a long-term production backup solution. A snapshot is a copy of files stored within the same datastore, and if the datastore is lost, the virtual machine and snapshot are also lost. However, snapshots do allow consumers to quickly take temporary near-line backups of the current virtual machine state to mitigate risk during change management windows and then quickly restore when needed to return to a previous configuration.

5.9.2.2. Development and Test Environments

Snapshots allow easy in-place upgrades with minimal risk and are an excellent solution for version control when a vCloud environment is used for development. Developers can make changes to the virtual machine and then, if a failure occurs, they can easily roll back to the previous version (state).

5.9.2.3. Third-Party Backup Integration

Some backup vendors use snapshots to create a copy of the virtual machine and then export the snapshots to a storage location outside of the vCloud infrastructure. For more information, see the individual vendor solution briefs on the VMware Solutions Exchange (https://solutionexchange.vmware.com/).

5.9.3 Design Considerations

This section describes the design considerations to take into account when enabling snapshot functionality in a vCloud environment.

5.9.3.1. Security

Consumers must have the user right `vAPP_Clone` to create snapshots.

5.9.3.2. Storage

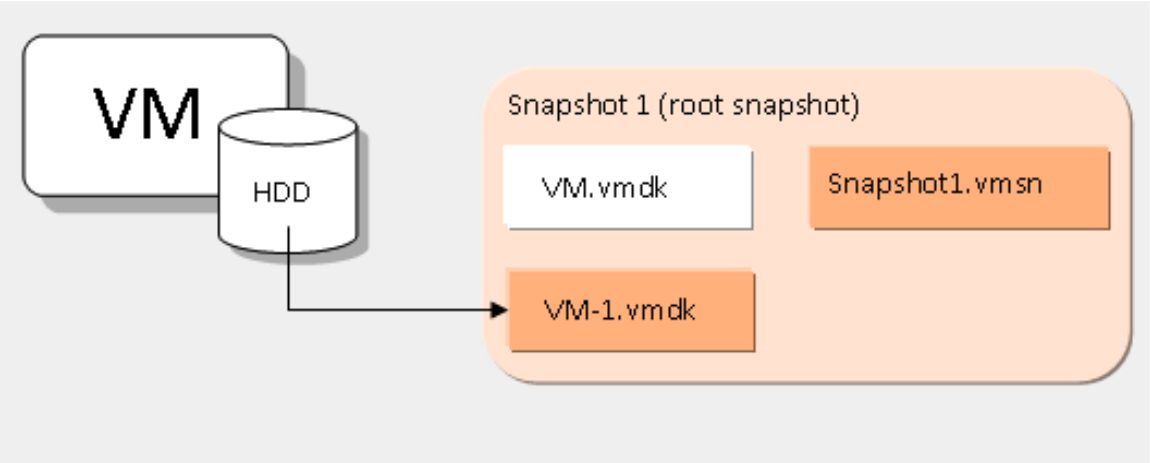
For each snapshot, the total consumed disk space includes the sizes of the files needed to capture the state of the virtual machine at the time of the snapshot (for example, hard disk and memory).

For example:

vmdk file + memory size = total consumed disk space

This is illustrated in the following figure.

Figure 34. Snapshot Sizing



vCloud administrators must take into account the number of consumers that they will permit to take snapshots. Because a vCloud virtual machine can only create one snapshot, this calculation is relatively easy.

Datastore free space monitoring is critical to the success of any vCloud environment, and even more so in an environment that allows snapshots. Allowing multiple virtual machines and snapshots to consume a datastore can impact the ability of consumers to start their virtual machines. To mitigate this, consider using vSphere Storage DRS, which allows for the redistribution of virtual machines if a datastore violates a free space threshold. However, vSphere Storage DRS is not a replacement for careful datastore sizing and monitoring because it does not stop a snapshot from writing to the datastore when performing migrations.

5.9.3.3. Performance

To reduce the impact of storage performance issues when creating snapshots, configure the storage array that serves the vCloud infrastructure to support VAAI. VAAI provides hardware-assisted locking. Hardware-assisted locking allows offloading of the lock mechanism to the arrays and does so with much less granularity than for an entire LUN. Therefore, the VMware cluster can provide significant scalability without compromising the integrity of the VMFS shared storage-pool metadata.

5.9.4 vCloud Director Snapshot Characteristics

vCloud Director 5.1 snapshot capabilities include the following:

- One snapshot per virtual machine is permitted.
- NIC settings are marked read-only after a snapshot is taken.
- Editing of NIC settings is disabled through the API after a snapshot is taken.
- To take a snapshot, the user must have the `vAPP_Clone` user right.
- Snapshot storage allocation is added to Chargeback.
- vCloud Director performs a storage quota check.
- REST API support is provided to perform snapshots.
- Virtual machine memory can be included in the snapshot.
- Full clone is forced on copy or move of the virtual machine, resulting in the deletion of the snapshot (shadow VMDK).

5.10 Storage Independent of Virtual Machines

The use of independent disks with vCloud Director 5.1 allows updates of virtual machines without impacting the underlying data. For example, you can detach the data disk from the existing virtual machine, delete the existing virtual machine, recreate the virtual machine, and reattach the original disk. This feature is a key enabler to enhance the deployment of a Cloud Foundry PaaS cloud within a vCloud environment.

5.10.1 Independent Disk Architecture

The independent disk feature consists of the following:

- A database schema to represent independent disks in vCloud Director and their associated backing in vSphere.
- A set of methods that implement the external vCloud Director API by manipulating the database schema and invoking the VIM API.
- A set of event handlers invoked by the VC Listener (Inventory Service) that allows vCloud Director to keep track of relevant vCenter Server activity (for example, vSphere Storage vMotion initiated by vSphere Storage DRS, or the vSphere Client).

Virtual disks in vSphere do not necessarily have unique IDs. For example, when a virtual disk is cloned (virtual machine clone) in vSphere, the new virtual machine receives a unique ID, but the disk IDs are reused. Also, the vSphere disk ID might be changed at any point using the vSphere API, which would break the vCloud Director reference pointer if it were the unique ID.

Therefore, vCloud Director generates and uses its own identifier for independent disks, persisted in the vCloud Director database. vCloud Director does not currently have the API infrastructure to support adding the vCloud Director disk ID to the disk metadata in the VMDK files.

A disk becomes detached in vCenter when a virtual machine using that disk is deleted in vCloud Director, but the disk must be saved for future virtual machines. Because detached disks are not known objects in vSphere, features such as vSphere Storage DRS cannot be used to migrate the detached independent disks. To aid in this situation, vCloud Director creates a virtual machine shell for each detached virtual disk and attaches the disks to the new shell. If the independent disk must be attached to a new virtual machine, the shell is deleted.

If a delete action takes place before an attach action, vCloud Director performs a check to verify that the disk is attached to a virtual machine object before completing the delete request to avoid inadvertently deleting the independent disk.

If actions are taken against the vCenter (through the UI or API), certain update actions are either safe or unsafe to perform:

- vSphere Storage vMotion or virtual machine relocate actions are safe actions to perform. vCenter updates vCloud Director with the revised locations of the disk files.
- Disk add and disk remove actions and the associated disk locations are unknown to vCloud Director and are therefore unsafe to perform.

5.10.2 Design Considerations

Independent disk limitations and usage considerations are as follows:

- Only SCSI controller types are supported.
- The disk size counts against the organization virtual datacenter quota.
- If the class of storage is not specified, the organization virtual datacenter default is used.
- If you delete a virtual machine, the independent disk is first detached from the virtual machine.
- When exporting a virtual machine with an independent disk, the disk is not tagged to identify that its source was an independent disk.
- The following operations cannot be performed if the virtual machine currently has an attached independent disk:
    - Clone vApp.
    - Copy vApp virtual machine.
    - Capture vApp to catalog.
    - Move virtual machine.
    - Change owner.
- When using the elastic virtual datacenter feature and allowing a provider virtual datacenter to span multiple clusters, it might be necessary to move an independent disk to a different datastore to attach it to a virtual machine in a different cluster. To avoid such a move, use the *locality* argument to create a disk in the destination cluster for the virtual machine (not necessarily on the same datastore).
- The scalability maximum for this feature is one independent disk per virtual machine up to the maximum number of virtual machines for vCloud Director.

5.11 vApp Load Balancing

vApp load balancing is used to increase availability through redundancy, increase throughput, minimize response time by redirecting to servers based on load or proximity, and avoid overloading resources.

5.11.1 Background

The vCloud Director environment is compatible with traditional IP based load balancing schemes. You can even run multicast-based load balancing schemes with some caveats. Global load balancers can also be used with vCloud Director-hosted virtual machines, as back end servers do not need any special configuration options.

VMware vCloud Director 5.1 offers options for self-service load balancing using the vCloud Networking and Security Edge built-in load balancer.

5.11.2 Load Balancing Architecture Options

In a vCloud Director environment there are various options for implementing load balanced vApps. Differences in architecture are based on the type of load balancer. Use cases include the following:

- External hardware-based load balancer.
- Third-party virtual appliance load balancer.
- vCloud Networking and Security Edge used as a load balancer.

5.11.3 vApp Load Balancing Examples

This section provides examples for each type of load balancer use case.

5.11.3.1. Example: External Hardware-Based Load Balancer Appliance

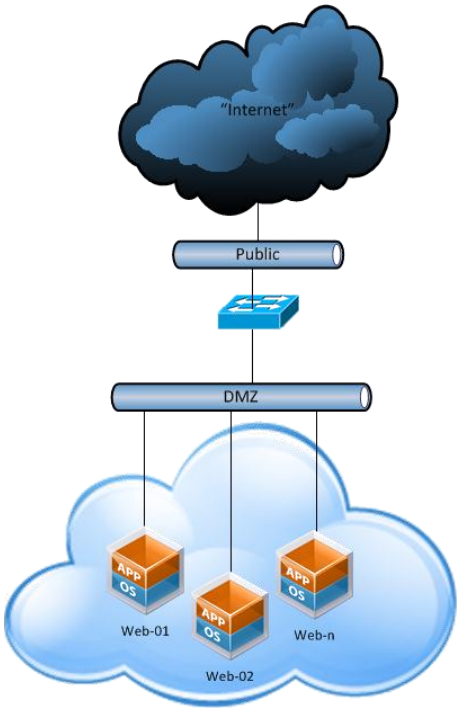
Third-party hardware load balancers provide options to control exactly how the load is to be balanced or distributed. Load balancers are not restricted to web traffic only—they can often be configured to handle arbitrary protocols. When you have esoteric workloads that you need to put behind a load balancer, these hardware boxes are still the most feature-rich option available.

In a vCloud Director environment, the most straightforward way to use hardware load balancers is by putting the back end (load to be balanced) virtual machines on a directly attached organization virtual datacenter network that is shared with the back end connection of the load balancer. This is usually thought of as a DMZ network. The load balancing logic is contained in the load balancer, and the virtual machines based on vCloud Director are used as pure compute resources.

In the following figure, the DMZ network is a vApp or organization virtual datacenter network that is bridged to the external network. The public network can be any physical networking that routes to the client location.

When evaluating the use of hardware based load balancers, weight the higher per-port cost against the availability of multiprotocol support and other advanced load balancing options.

Figure 35. Hardware-Based Load Balancer



5.11.3.2. Example: Third-Party Virtual Appliance as a Load Balancer

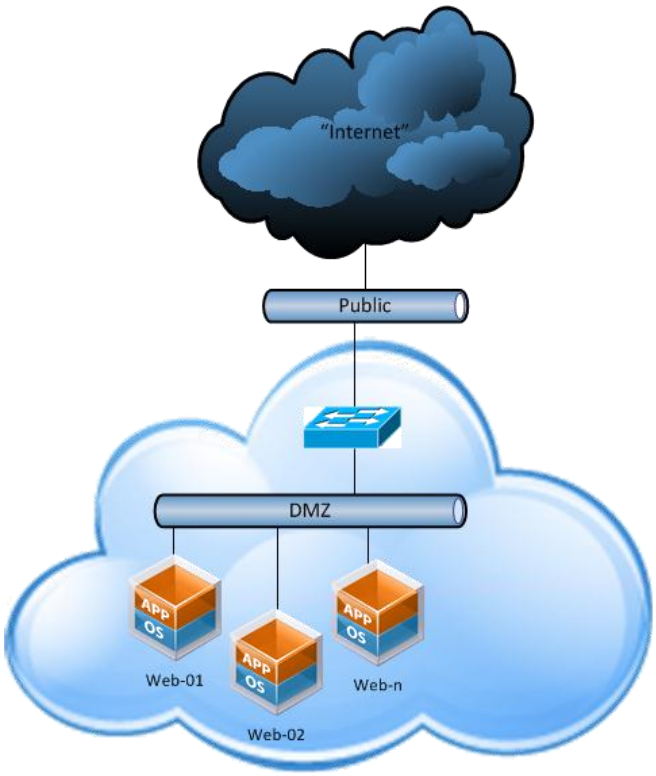
There are many third-party virtual load balancers available with varying degrees of multiprotocol support and advanced features. The third-party configuration works with all virtualization-supported networking protocols that the virtual load balancer supports.

When using a virtual appliance as a load balancer, protect the security of the vApp workloads upstream by using a firewall. In this configuration, vCloud Director does not provide security or isolation for the back end workloads other than what the load balancer provides.

In the following figure, the DMZ network is an isolated vApp or organization virtual datacenter network, and the public network is a vApp or organization virtual datacenter network that is bridged to an external network able to route to clients.

A major advantage in running a virtual appliance instead of a hardware appliance is that the network port can scale up to the bandwidth that is available on the vSphere host (usually 10Gbps per port). In some implementations, having up to 10Gbps of bandwidth available is a significant advantage over the bandwidth available with a physical appliance. Hardware appliances usually are limited to 1Gbps ports.

Figure 36. Third-Party Virtual Load Balancer



5.11.3.3. Example: vCloud Networking and Security Edge as a Load Balancer

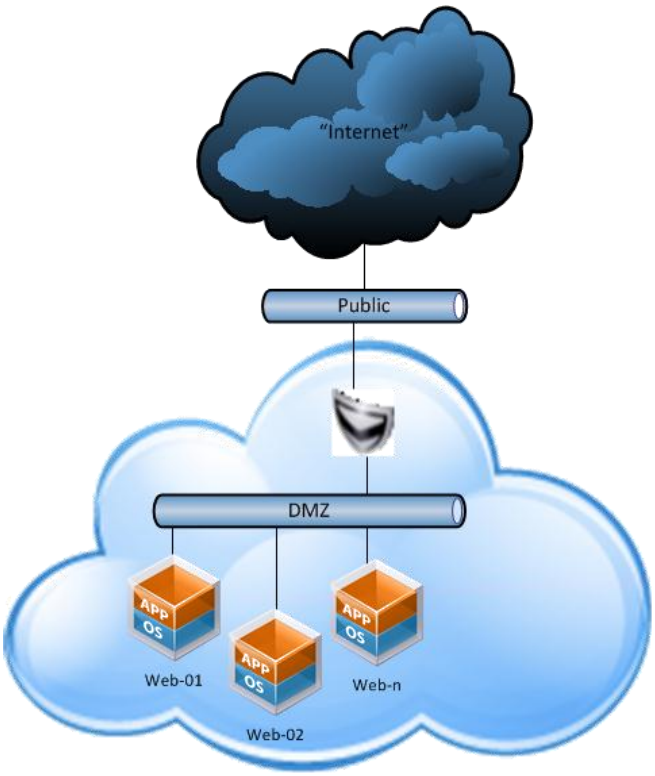
vCloud Network and Security Edge offers basic HTTP (port 80) and HTTPS (port 443) load balancing and can be used for applications that need one or both of these protocols. vCloud Networking and Security Edge can be used to load balance vCloud Director cells and the basic web server configuration.

vCloud Networking and Security Edge currently has limited load balancing advanced features such as SSL termination and stickiness. If advanced features are critical to the operation of the application being load balanced, consider evaluating a third-party virtual or physical load balancing appliance.

In the following figure, vCloud Networking and Security Edge provides the load balancing functionality and firewall needed to secure the vApp workloads.

As in the third-party virtual appliance as a load balancer example, the DMZ network should be an isolated vApp or organization virtual datacenter network, but the public network should be a vApp or organization virtual datacenter network that is bridged to the external network that routes to client locations.

Figure 37. vCloud Networking and Security Edge as a Load Balancer



5.11.4 Load Balancing Design Implications

VMware vSphere High Availability protects against physical host failures and restarts failed virtual machines using a third-party load balancing virtual appliance or the solution based on vCloud Networking and Security Edge. This affords about 99.9% uptime for the load balancing functionality (based on vSphere HA availability numbers).

You can improve availability for the load balancer by running it in native high availability mode. This affords the edge an almost instantaneous failover, with session preservation.

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 Manager provides the metering capability to enable cost transparency and accountability in vCloud environments.

When running a private vCloud, enterprises do not necessarily have the same cost pressures as a public vCloud service provider. Required chargeback procedures or policies might not exist. An alternative to chargeback is *showback*, which attempts to raise awareness of consumption usage and cost by showing the consumer what the services would cost without involving formal accounting procedures to bill the usage back to the consumer's department.

Chargeback provides 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 consume, 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 Manager

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

6.1.1 vCenter Chargeback Manager Architecture

The vCenter Chargeback Manager is based on a Windows server that runs the vCenter Chargeback web application, load balancer, and data collector services. Services may run on separate servers for scalability and resiliency. The server can be virtual or physical and has the following recommended specifications:

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

vCenter Chargeback Manager instances 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 Chargeback cluster can include up to three 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 the user interface or an API, routes through the load balancer. The load balancer forwards the request to a Chargeback instance in the cluster based on the number of requests currently serviced by each instance in the cluster. With multiple instances, Chargeback also load balances the report processing load by leveraging the internal Quartz scheduler. If the load balancer service terminates, the Windows service can be restarted. The built-in load balancer cannot be replaced with a third-party load balancer. All Chargeback instances in a cluster connect to the same Chargeback database.

If the load balancer service becomes unavailable, the Chargeback Manager application does not work. If the Tomcat server on a cluster instance terminates, the load balancer redirects requests to other cluster instances.

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

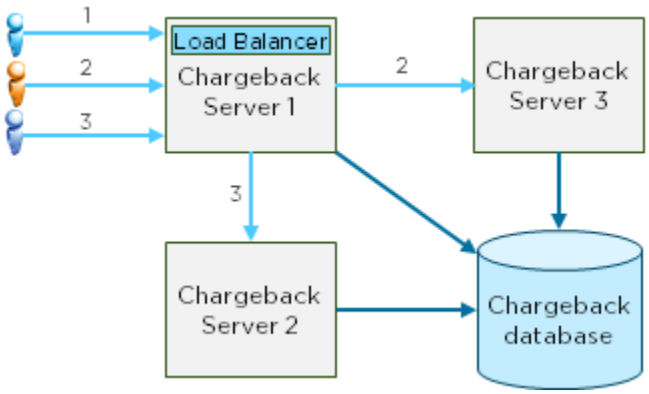
1. The load balancer uses the number of requests to find the best worker.
2. Access is distributed according to the lbfactor (it is the same for all the servers in the cluster) in a sliding time window.

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

- sticky_session = 1 (true)
- method = R

The following figure shows a vCenter Chargeback cluster.

Figure 38. vCenter Chargeback Cluster



Multiple Chargeback environments (separate vCenter Chargeback Manager and database) can work with a single vCloud Director instance, but this increases the load on the vCloud Director instance.

The vCenter Chargeback Database stores organization hierarchies, cost/rate plans, and global chargeback configuration data. Supported databases include Microsoft SQL Server Express, Microsoft SQL Server, and Oracle. Database partitioning helps to improve the performance of vCenter Chargeback Manager. vCenter Chargeback Manager does not support the database (DB2 or PostgreSQL) included in the vCenter appliance.

6.1.2 Data Collectors

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

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

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

6.1.3 User Roles

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

6.2 Maximums

The following table lists maximums for Chargeback Manager.

Table 15. Maximums

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

6.3 Cost Calculation

To track resource metrics for vCloud entities, vCenter Chargeback Manager 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 following table lists these allocation units.

Table 16. vCloud Hierarchy Allocation Units

Entity	Pay As You Go	Allocation Pool	Reservation Pool
Organization virtual datacenter	None	• CPU • Memory • Storage	• CPU • Memory • Storage
vApp	None	None	None
Virtual machine	• vCPU • Memory • Storage	• vCPU • Memory • Storage	• vCPU • Memory • Storage
Template	Storage	Storage	Storage
Media file	Storage	Storage	Storage
Independent Disk	Storage	Storage	Storage
Network	• DHCP • NAT • Firewall Count of networks	• DHCP • NAT • Firewall Count of networks	• DHCP • NAT • Firewall Count of networks

6.3.1 Pricing Models

Installing vCloud and vCloud Networking and Security Manager data collectors creates default cost models and billing policies that integrate with vCloud Director and vCloud Networking and Security 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 policies and pricing models, make copies and modify the duplicates. For more information, see the *User's Guide* in the *vCenter Chargeback Manager Documentation* (http://www.vmware.com/support/pubs/vcbm_pubs.html).

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

- 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 pricing assigns a fixed price to a hard bundle of vCPU and memory. Virtual machine instance matrixes are linked with a pricing model. The pricing model includes the hierarchy selection criteria, a fixed pricing table, and a default fixed price. Selection criteria options include name pattern matching, custom attribute matching, or no criteria. VM instance uses a step function—if there is no entry for a particular virtual machine size, the charge is based on the next larger instance size.

vCenter Chargeback Manager 2.5 introduces VM instance pricing for all allocation models. Use VM instance pricing to create a fixed price 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 can 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 can tie services together into a logical workflow. VMware has different management applications supporting workflow process definition and execution.

- *vCenter Orchestrator* is a technical orchestration authoring platform within vCenter that enables administrators to automate repetitive tasks by creating workflows that leverage extensive integrations with VMware and third-party vCloud components. See *Workflow Examples* for detailed examples of orchestrated workflows.
- *vFabric Application Director* automates the deployment of multitier applications to the vCloud. vFabric Application Director can simplify virtual machine template management by providing a catalog of services used to install, configure, and start software services on virtual machines. vFabric Application Director uses the vCloud API to issue provisioning requests to a vCloud provider and can be deployed in public, private, and hybrid vCloud environments.
- *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 of federation and ecosystem support. All current federation tools communicate with 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, provisioning portals communicate 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 might want to meter and charge for API usage. VMware also recommends protecting the vCloud API through audit trails and API inspection. In some cases, vCloud providers might want to extend the vCloud API with new features.

To assist with the vCloud API use cases, the vCloud provider might 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 as a proxy for 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 Cloud Provisioning with vFabric Application Director

vFabric Application Director is an entry point into the vCloud for creating and deploying multitier applications. vFabric Application Director consumes vCloud resources by defining a vCloud *provider* that is associated with a vCloud Director organization and associated catalogs.

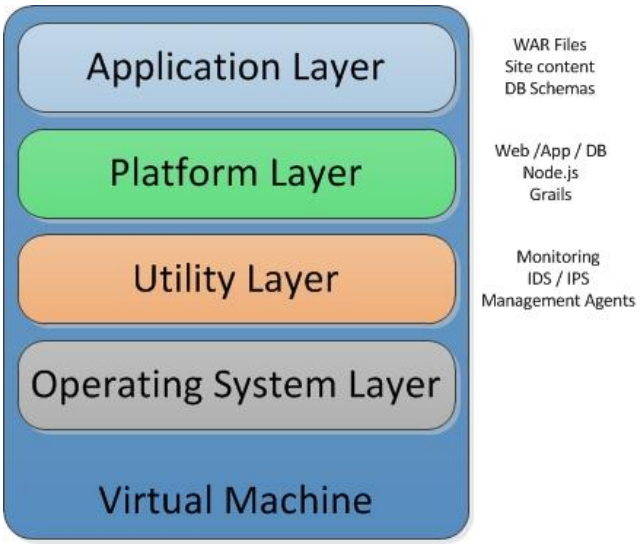
- vFabric Application Director uses the vCloud API and requires access to the vCloud Director servers to issue provisioning requests.
- vFabric Application Director has a catalog of services that define how to install, configure, and start services on a virtual machine.
- vFabric Application Director can assemble virtual machines and services into a multitier application that is deployed to a vCloud provider.

7.2.1 Simplifying vApp Template Management

Catalog services can be constructed for each software component that is normally installed on virtual machines that are deployed to the vCloud environment. Consider a virtual machine as a collection of software packages and services running on a guest operating system. Most software components fit into a layered model where administrative duties might fall to different departments for maintaining software at each layer.

In the following figure, multiple layers of software and services define the characteristics of the virtual machine. By creating services for each component in the vFabric Application Director catalog, each department can maintain a service component in the catalog. This simplifies base virtual machine template creation and management process because the templates need to contain only the base operating system and appropriate patch level. All other services can be installed, configured, and started by vFabric Application Director.

Figure 39. Software Component Layers

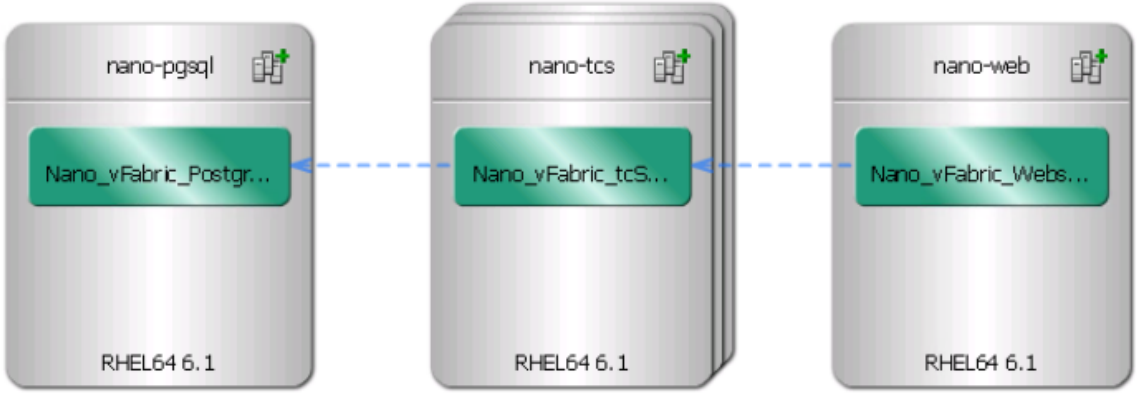


7.2.2 Simplifying vApp Template Management

To build a multitier vApp, vFabric Application Director uses a blueprint to construct a vCloud vApp that contains multiple virtual machines. Each virtual machine in the vApp can be based on different vApp templates and each virtual machine can be customized by the services selected from the vFabric Application Director catalog.

In the following figure, a three-tier application consisting of a presentation (web) tier, application tier, and database tier is modeled in vFabric Application Director as a blueprint. At deployment time, vFabric Application Director creates a corresponding vApp in the vCloud provider based on the virtual machine templates specified in the blueprint.

Figure 40. Three-Tier Application Modeled in vFabric Application Director



7.2.3 Guest Customization and the vFabric Application Director Agent

Virtual machine templates consumed by vFabric Application Director must have the vFabric Application Director agent installed. The interaction works as follows:

1. On first boot, virtual machines deployed by vFabric Application Director in the vCloud provider environment go through the vCloud guest customization process.
2. At the end of guest customization, the vFabric Application Director agent on each deployed virtual machine initiates contact with the vFabric Application Director server and downloads the latest version of the agent software.
3. The agent downloads the service scripts and creates environment variables that correspond to properties created in the service or blueprint.
4. Service scripts can then be executed to install, configure, and start software on each deployed virtual machine.

The vFabric Application Director agent in each virtual machine establishes the connection to the vFabric Application Director server. This reduces the complexity of firewall management.

7.2.4 vCloud Networks and vFabric Application Director

When provisioning a vApp, vFabric Application Director does not create any vApp internal networks. Application Director connects provisioned vApps directly to a vCloud organization virtual datacenter network. This removes the ability to provision fenced vApps with vFabric Application Director. Properties can be dynamically updated at deployment time, so service scripts can be written to modify relevant configuration parameters for software being installed or configured.

As an example, a property can be created to acquire the IP address of a new virtual machine at deployment time. This IP address property can be used by a service script to properly configure an application based on the new IP address of the newly provisioned virtual machine. This property can be exposed across multiple services and across multiple virtual machines deployed by vFabric Application Director through dependency mapping in the blueprint.

- vFabric Application Director-deployed vApps that are directly connected to an organization virtual datacenter network must allow for the agent service in each virtual machine to contact the vFabric Application Director server.
- vFabric Application Director does not connect a vApp to an isolated organization virtual datacenter network as that removes the ability for the agent to contact the vFabric Application Director server.
- vFabric Application Director connects vApps only to an organization virtual datacenter network that is “direct” or “routed” to an external network.

7.2.5 Building a Software Repository

Building a central software repository or depot simplifies the service development process. Locate the software repository in the same environment or datacenter as the target vCloud provider where vFabric Application Director provisioned applications are deployed.

Data downloads from the software repository can be large in complex deployments, so consider bandwidth and latency between the software repository and provisioned virtual machines.

vFabric Application Director can optionally place content on a provisioned virtual machine using a special *content* type property. To support this feature, the software repository must allow HTTP access for file downloads. Other access methods require service authors to write their own methods to retrieve data from a software repository.

7.2.6 Design Implications

vFabric Application Director server is supported only when deployed to an environment based on vCloud Director. Often the vCloud environment on which vFabric Application Director is deployed is the same environment where applications are being provisioned. Because vFabric Application Director uses the vCloud API to issue provisioning requests, the vFabric Application Director server must be able to issue API calls to the vCloud Director servers that are managing the vCloud environment. This has the following security implications for some consumers:

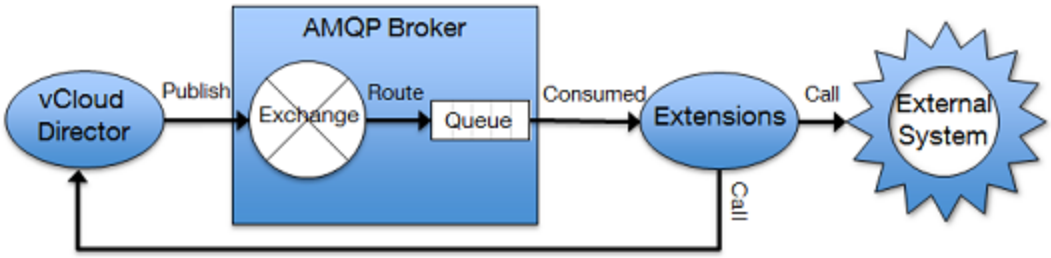
- In public vCloud deployments, vCloud consumers often have access to only one vCloud organization. In this scenario, the vFabric Application Director vCloud provider organization is the same as the organization housing the vFabric Application Director server. If access to multiple organizations is available, it may be beneficial to deploy the vFabric Application Director server and software repository to one organization and have provisioned workloads deployed to another organization. Network access must be available from the deployed virtual machines to the vFabric Application Director server and software repository.

- In private vCloud deployments, deploy the vFabric Application Director server to a vCloud organization designated for management systems. This provides isolation for administrative purposes and can simplify Chargeback administration. The software repository can also be deployed to the management organization. vCloud providers can be defined in vFabric Application Director based on organization separation policies. Network access must be available from the deployed virtual machines to the vFabric Application Director server and software repository.
- In a hybrid vCloud deployment, the vFabric Application Director server might not be local to the vCloud provider where applications are deployed. vFabric Application Director uses the vCloud API to make provisioning requests to the vCloud provider. The agent installed in each vFabric Application Director-provisioned virtual machine must also be able to establish a network connection to the vFabric Application Director server. VMware recommends locating the software repository in the same environment or datacenter as the target vCloud provider due to bandwidth and latency considerations.
- Deploying vFabric Application Director servers into a vSphere environment is not currently a supported configuration.

7.3 vCloud Messages

vCloud messages provides 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 41. vCloud Messages



7.3.1 Message Publication

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

- Notifications are published on user-initiated events (for example, creation, deployment, and deletion of a vApp) and system-initiated events (for example, vApp lease expiration) that contain 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 UI or vCloud API.

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

7.3.2 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]
```

7.3.3 Extension

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

- Subscribes to an AMQP queue for receiving new messages.
- Triages the received messages.
- Processes messages into operations (internal or external calls).
- Calls the vCloud API to get more information on the objects involved in an operation and takes action on blocked tasks.

7.3.4 Design Considerations

The following applies for notifications and blocking tasks:

- Notifications and blocking tasks are separate mechanisms that are implemented over the same AMQP message bus.
- When a task is blocked, an extension is responsible for delivering a message to query the status of the related object or take action on the blocked task.
- Resume, Progress (that was made), Abort, and Continue are valid calls against a blocking task.
- Configure the timeout for a blocking task globally in vCloud Director.
- You can abort a waiting blocking task directly from the VMware vCloud Director UI.

7.4 vCenter Orchestrator

vCenter Orchestrator 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 required several individual operations on different systems. See *Workflow Examples* for detailed examples of orchestrated workflows.

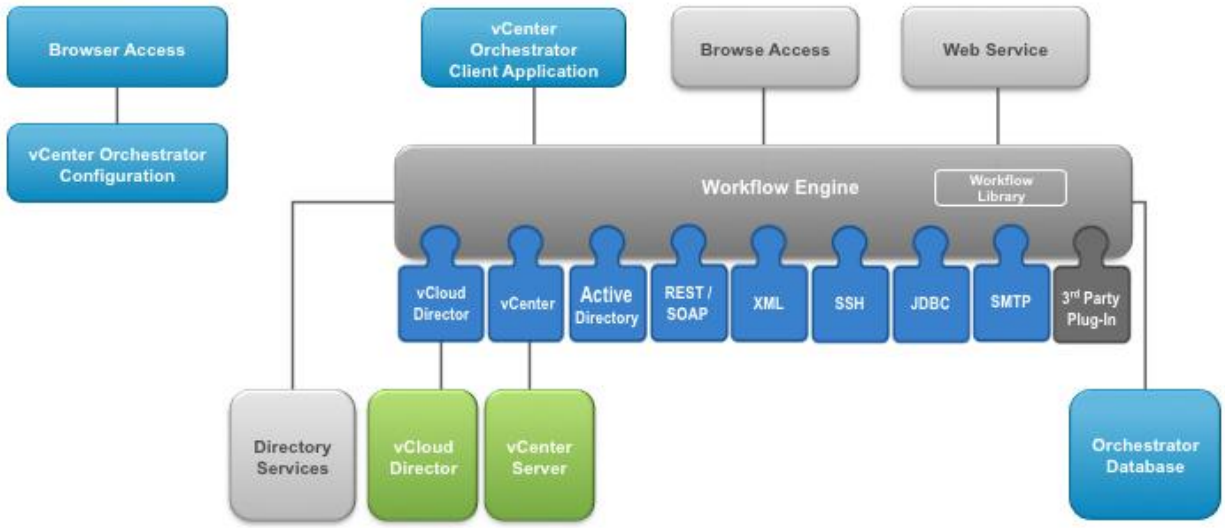
In general, if an operation uses only one underlying system, consider providing direct access to that system for efficiency and reduction of complexity. In a vCloud environment, vCenter Orchestrator can automate highly repetitive tasks to avoid manual work and errors.

vCenter Orchestrator consists of the following applications:

- vCenter Orchestrator Client – Enables the workflow developer to create, assemble, test, and package workflows, actions, policies, resources, and configurations.
- vCenter Orchestrator Server Web configuration – Runs as an independent application side-by-side with a web front-end to allow administrators to configure the vCenter Orchestrator Server and its plug-ins and perform maintenance operations.

- vCenter Orchestrator Server – Provides runtime orchestration service, including its interfaces and its pluggable adapters.

Figure 42. 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.

Orchestration use cases include the following:

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

7.4.1 Design Considerations

Depending on the overall architecture and how orchestration is used, 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 the following:

- Multiple vCloud Director hosts.
- Multiple vCenter hosts.
- Multiple other host types (UCSM, REST, SOAP, vCenter Update Manager).

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 for backward compatibility (for example, use plug-in 5.1 if managing a mixed vCloud Director 1.5 and vCloud Director 5.1 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 involves sending several HTTP GET/PUT/POST/DELETE requests to the vCloud Director server and, upon reply, 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 large service definition and advanced mechanisms, such as a notification service, that are used extensively by vCenter Orchestrator. Sessions are continually maintained between vCenter and vCenter Orchestrator. 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 include the following:

- vCenter Orchestrator 5.1 introduced new per-node maximums of 20 vCenter Server instances, 1280 vSphere hosts, and up to 35000 virtual machines in inventory.
- vCenter Orchestrator scalability can be increased with the use of the VMware vCenter Orchestrator Multi-Node Plug-In. See the *Multi-Node Plug-In* blog (http://blogs.vmware.com/orchestrator/2012/01/vco-multi-node-plug-in.html) for more information.
- If a vCenter Orchestrator is overloaded by a large level of objects to manage, attempt to tune the server for higher scalability. Alternatively, design the solution to use different vCenter Orchestrator instances that manage different vCenter Servers, or connect to a large vCenter using different vCenter Orchestrator instances that are configured with accounts to access different zones of vCenter.

7.4.2 Scalability

When configuring vCenter Orchestrator to run a numerous 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 have completed an execution run. Completed workflow states can be “completed successfully,” “failed”, “canceled” or "passivated" (waiting-for-signal state). The execution queue has a default size of 10000 workflows. If 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 megabytes of memory (varies depending on the number of enabled plug-ins and plug-in objects). Limiting the number of workflows allows allocation of threads and memory, with the maximum depending on the JVM settings, the operating system, and the underlying hardware.

To change the default value, change the following properties in the `Orchestrator\appserver\server\vmo\conf\vmo.properties` configuration file:

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

Note VMware recommends following the guidelines in the rest of this document before increasing the default settings for the concurrent workflows because doing so 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 resources and increases the memory footprint of each running workflow. Disable all plug-ins that are not in use to increase the workflow engine capacity.

7.4.3 Workflow Design

Workflow design affects duration and use of resources. The following are design guidelines for workflow design:

- Effective scripting – Use scripting development design guidelines to avoid unnecessary and highly resource-demanding operations such as active wait loops, repetitive expensive calls to the same resources, and ineffective algorithms. Perform extensive testing on a vCenter Orchestrator 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 still runs within the same workflow instance, but uses fewer resources. Link the workflows in higher-level workflows instead of calling individual workflows in sequence.
- Reduce the number of waiting workflows – If the reason for the high concurrency is due to a high number of workflows waiting on external systems, the following methods can help 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's long running workflows make extensive use of the waiting event palette element.

- When workflow activity must be suspended until a determined time, programmatically schedule a workflow task.

Although they save active resources, each passivation and activation consumes CPU resources and database access. The following are design guidelines for using the waiting timer or waiting event:

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

7.4.4 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 would be excessive for simple operations that do not require this level of service. Do not use vCenter Orchestrator 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 tested maximums for vCenter Orchestrator to avoid resource starvation.
- Load balancing – If maximums are exceeded, it may be necessary to design the system to load balance the workflows across different vCenter Orchestrator servers.
- Orchestrated systems bottleneck – Use vCenter Orchestrator workflows to prevent starting too many operations at once on the orchestrated systems. Design this logic to support the defined load. Expose the parameters that have an influence on the started workload as configuration elements to be adjusted by the orchestration administrator (a parameter that determines the number of vApp clones to be processed in parallel).

7.4.5 Orchestrator Client

The vCenter Orchestrator Server has a client application to develop workflows and actions. During server installation, install the client on the same system as the server. In production environments, the local installation of the client software is used only in emergency cases when a matching client is not available through developer workstations.

Have developers install the client on their workstations to connect to their test or development servers on the same LAN. If connecting to a remote server, use Remote Desktop to run the client from the same LAN.

7.4.6 vCloud Director Plug-In

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

- 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. Verify that forward and reverse DNS are working for the specified address.
- If a hostname or fully qualified domain name (FQDN) is specified, verify that forward and reverse DNS are working and use the FQDN in the plug-in configuration.
- If no hostname is specified and the vCloud Director server is configured only to use an IP address, use the same IP address for the plug-in configuration.

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

After specifying the **Host** field, choose a strategy for managing the user logins. The available 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 must 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 a session for each user 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 vCenter Orchestrator.

Also consider the following:

- For organizations that use different LDAP hosts, one dedicated instance of vCenter Orchestrator is required per organization.
- Multiple sessions can strain CPU, memory, and bandwidth.

In addition, an organization setting is required. The 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 restricting create, read, update, and delete access to all elements that belong to the given organization.

The most common use cases for 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 must be set up to synchronize with the vCloud provider LDAP host that is configured with vCenter Orchestrator.

  If configuring more than one vCloud Director connection, use a combination of shared session and per user session to grant vCenter Orchestrator 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 vCenter Orchestrator users access to a given organization, have both connections use **Session per user** and set permissions differently for the sessions to avoid all users having wide access to all organizations.

- As a public vCloud tenant of one or more organizations, 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.

7.5 vCenter Orchestrator Examples

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

7.5.1 vCloud Administration Orchestration Examples

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

- A provider wants to begin working with a new customer. The main steps are to add a new organization, users (possibly from LDAP), networks, virtual datacenters, and catalogs. The provider might also want to schedule a recurring chargeback report for billing and send an email notification to tenants advising them that their vCloud environment is ready.
- 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 allocation of the external network to the tenant's organization.

7.5.2 Organization Administration Orchestration Examples

Operational tasks within the tenant's organization can also benefit from automation. The following examples address vApp lifecycle management, such as vApp creation, configuration, maintenance, and decommission.

- Virtual machines are created in an environment using Active Directory to identify services such as authentication and printing. After deployment, the virtual machine must join the Active Directory

domain. It is usually 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 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.

- 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 can accept a list of systems and a source for the software or configuration as parameters, and then perform the update on each system.

7.5.3 vCloud Consumer Operation Orchestration Examples

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

Example 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.

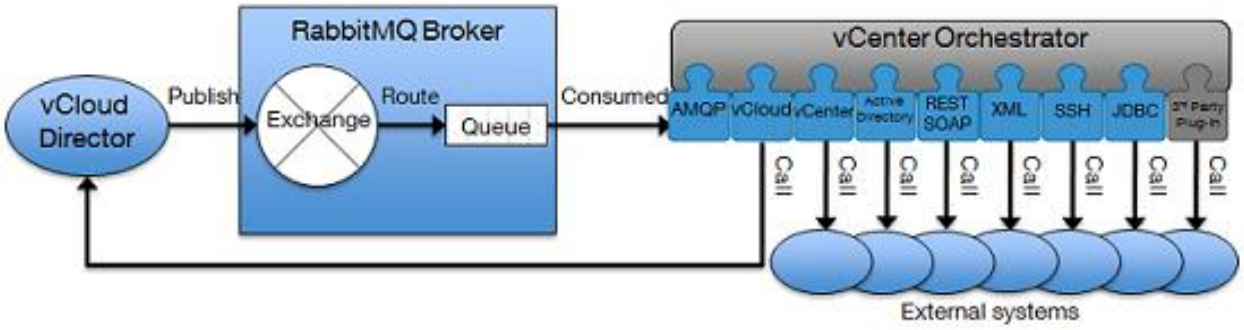
vCenter Orchestrator can be used to create custom workflows at the vCloud API and vSphere levels. Other vCloud provisioning solutions frequently have built-in workflow functionality that integrates with vCloud Director through the vCloud API and is an alternative to vCenter Orchestrator.

See the *VMware vCenter Orchestrator Documentation* (http://www.vmware.com/support/pubs/orchestrator_pubs.html) for additional information on installation, configuration, and workflow development:. Also see *Workflow Examples* for detailed examples of orchestrated workflows.

7.5.4 Using Orchestrator as a vCloud Director Extension

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

Figure 43. 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 vCenter Orchestrator to receive message updates on new messages.

Restarting the vCenter Orchestrator 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.

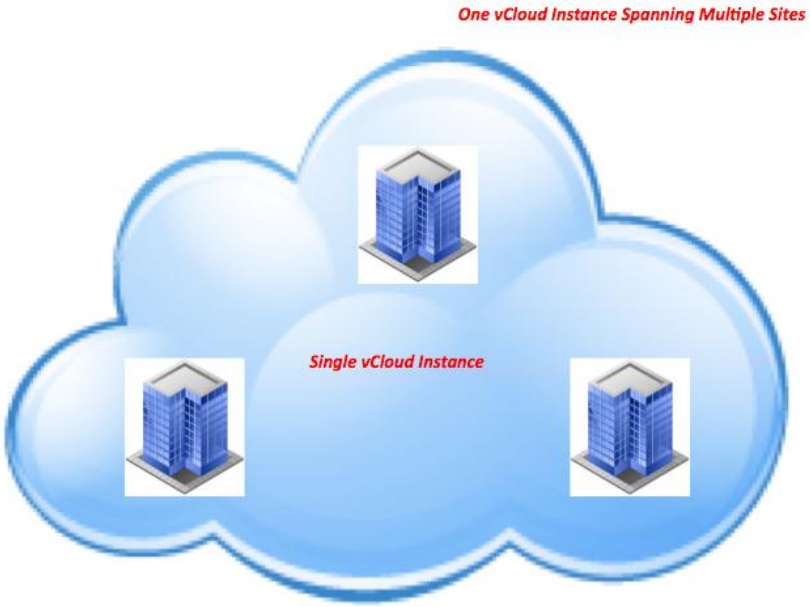
8. Multisite Considerations

Initial vCloud Director deployments were targeted at traditional test and development, scale-out infrastructure, and tier-3 workloads where advanced infrastructure features were typically not required. As customers started to adopt vCloud Director for a different set of workloads, a new set of requirements arose. One of these requirements is the ability to deploy vCloud Director to manage resources that span more than a single site.

Note Multisite hybrid vCloud scenarios (defined as a combination of private and public cloud resources) are not discussed.

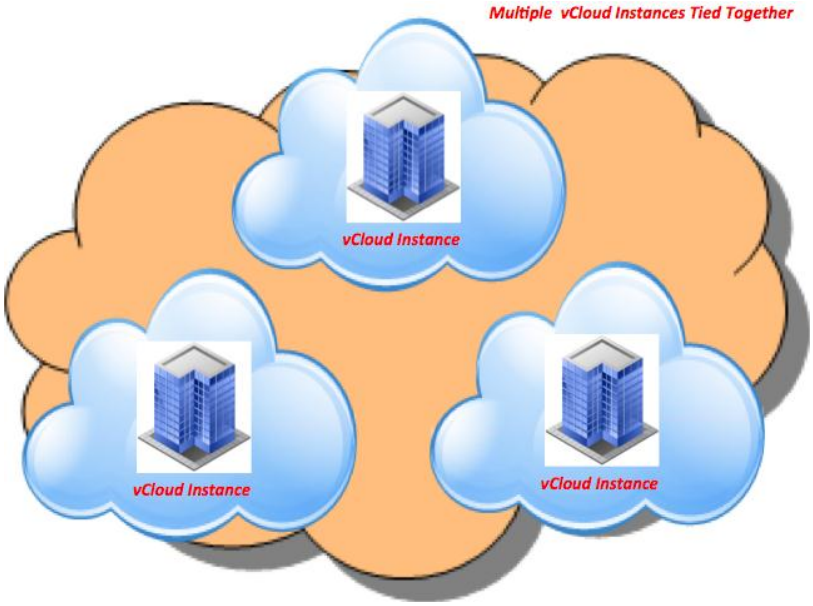
This section provides guidelines and discusses various options, limitations, and supported scenarios to deploy vCloud Director in a distributed scenario, that focuses on a specific vCloud distributed model. The goal is to take a private or public vCloud instance and describe the options for stretching it across multiple locations—this model takes a single vCloud Director instance with two or more vCloud Director cells and determines how the different components can be deployed separately in different locations. From a vCloud Director perspective this is considered a single vCloud. The following figure illustrates this concept.

Figure 44: Single vCloud, Multiple Sites



There are additional models that enable enterprise customers and service providers to create a single vCloud that spans multiple remote sites. One of the alternative models is to deploy traditional vCloud Director instances in each of the available locations and layer an additional level of management on top of them to create a single entry point into dispersed vCloud instances. This additional layer can be implemented with an additional software layer such as VMware vCloud Automation Center. The following figure illustrates this concept.

Figure 45: Multiple vCloud Instances Tied Together



8.1 Multisite Availability Considerations

Various distributed scenarios can enable vCloud capacity to be spread across different premises located around the world. Some of the distributed design options lead to resiliency advantages. Others may be prerequisites for DR scenarios (however, these, are more of an implicit outcome of distributing the compute farm rather than an explicit design goal).

Creating a distributed vCloud model is foundationally required, but is not sufficient to address high availability and disaster recovery for vCloud workloads. The focus of this section is on how to distribute resources, not how to make workloads highly available on those resources.

8.2 Distributed Cloud Deployments Use Cases

The following are some major use cases for spanning a vCloud across multiple locations (other use cases are possible.):

- Better and more uniform usage and management of distributed resources – Many customers and service providers want to build one single vCloud that contains resources that are distributed across cities, countries and continents. That is the way they operate their IT. They prefer to install and operate one vCloud out of the box rather than building two or more vCloud instances that would require additional integration.

- The second use case is similar to the first, but the business driver is different. Though there are customers and service providers that distribute resources because that is how they operate as a global company, there are situations where they (specifically, service providers) require that the resources are distributed in various countries and geographies. This is due to data regulations and compliance requirements, and because their customers cannot take their assets outside of a certain country or geography. In this case, service providers must distribute locations where they are selling their services. These service providers want to manage these datacenters under the same single vCloud umbrella.
- The third use case is a variant of the first two use cases and it is specific to service providers. Many service providers are interested in offering vCloud services to their customers where the service is managed centrally on a shared management platform, but it is delivered at the customer premise, where a dedicated physical environment is deployed. This may be done for various reasons, from security and compliance to network requirements. Think about a customer subscribing to a public vCloud service where the service provider assigns an entire provider virtual datacenter to that organization and that provider virtual datacenter happens to be physically deployed at the customer premises.
- The fourth use case involves public or private service providers that have vCloud consumers distributed across the globe and want to guarantee the lowest possible latency and best possible experience. The best way to achieve this is to move the user workloads and the systems where they run physically as close as possible to the consumer. These service providers also want to manage these resources as a single vCloud.
- The fifth use case is to provide a mechanism to allow end users to consume resources that are physically distributed in different locations for increased resiliency. In scenarios where the resiliency of the end-user workload is managed at the application level, the end users can instantiate loosely coupled virtual machines in standalone provider virtual datacenters distributed at remote locations, thus achieving scalability and resiliency. In this case, the end user is responsible for managing the resiliency of the application.
- The sixth use case enables a vCloud provider to increase resiliency of end-user workloads by failing them over to different sites if anything fails at the datacenter where the workload is originally instantiated. In this scenario, application resiliency is achieved through recovery mechanisms implemented at the infrastructure level, not at the application level. This is a resiliency service that the vCloud administrator offers to the end user regardless of the application resiliency attributes.

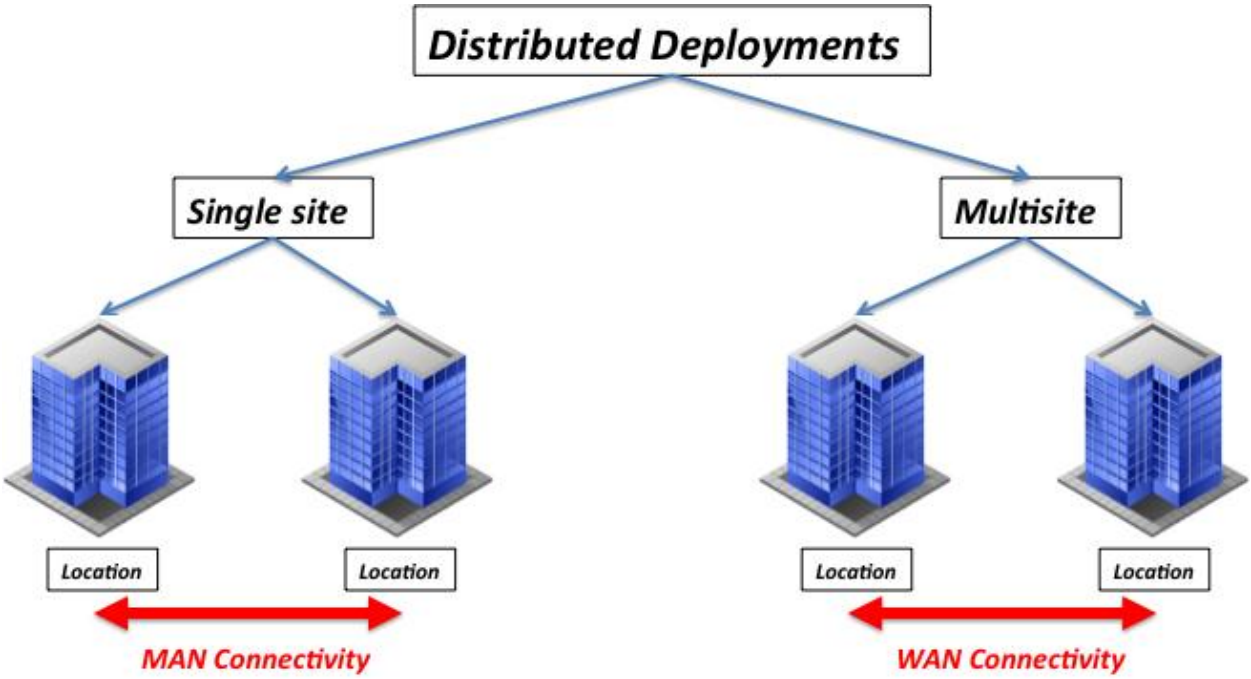
This section focuses on describing the various options for stretching standalone compute resources in different locations under the governance of a single vCloud.

8.3 Multisite Terminology

This section uses the following terminology:

- *Distributed vCloud* – The generic concept of spreading vCloud resources and components across different locations.
- *Location* – A physical location, a building, or an entire physical datacenter with LAN connectivity, where vCloud components are deployed.
- *Single Site vCloud* – A vCloud that spans one or more locations that are connected with MAN connectivity.
- *Multisite vCloud* – A vCloud that spans one or more locations that are connected with WAN connectivity.

Figure 46. Distributed Deployment Options



Historically, a vCloud Director deployment was supported only in a single site or in a single location. However, this statement led to some confusion because it is not very deterministic. In fact, it is not unusual to find connectivity between different locations at large corporations that is better than the connectivity in a single site at smaller organizations.

For this reason this statement is clarified with a more deterministic approach and *a single site* is considered to be any local or distributed IT deployment where connectivity between any of the deployed components has a latency round trip time (RTT) of 20 milliseconds (or less).

This does not call out bandwidth requirements. This is because bandwidth is more of a problem from the perspective of an end-user experience than it is from a functional perspective. We assume that bandwidth in a MAN scenario is sufficient to not cause connectivity problems. However, we realize that, depending on the usage patterns of the vCloud, a relatively low bandwidth can result in higher response time for the user. The vCloud architect is responsible for planning according to the expected result and projected usage patterns.

These network characteristics are referred to as *MAN connectivity*. A single-site deployment is where one or more locations are used to host all of the vCloud Director components, and the RTT connectivity within a location or across two or more locations is less than 20ms.

- If the vCloud Director deployment has all of the components that fall within these connectivity characteristics, it is considered to be a single site and the deployment is fully supported.
- If the distributed vCloud Director deployment has components that fall outside of these network characteristics, then it is a *multisite* implementation.

8.4 Deployment Options

There are an infinite number of ways to distribute a vCloud platform. This is due to the number of vCloud components that must be deployed and the various connectivity options.

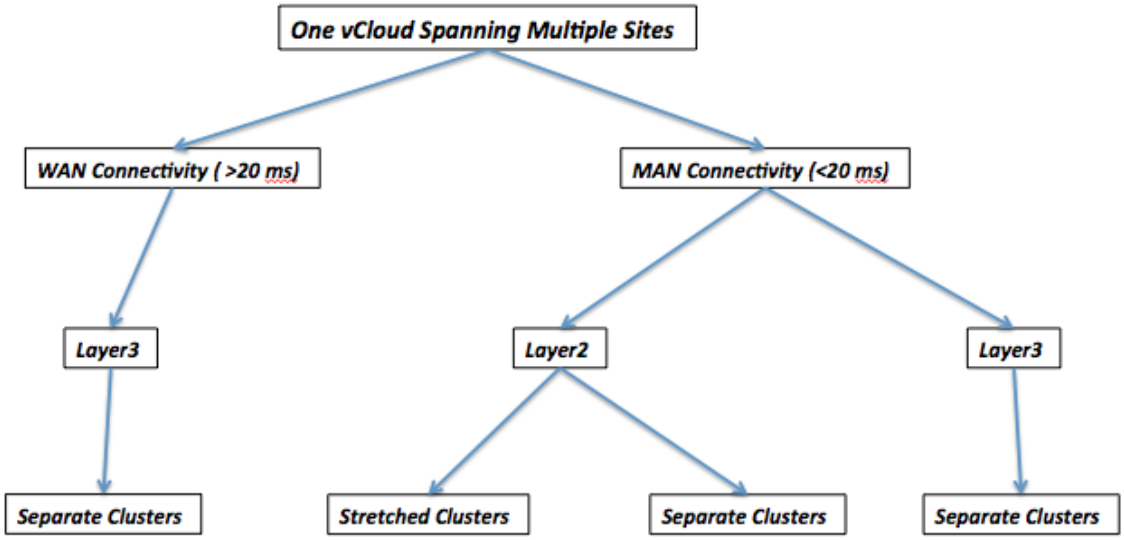
The following options lead to different combinations and layouts:

- Connectivity between locations (MAN/WAN).
- Network layer (Layer 2/Layer 3).
- End-user workloads clusters configurations (stretched/separate).

Note The combination of some of these options may not be viable. For example, a vSphere stretched cluster configuration requires and is deployable only in conjunction with a Layer 2 stretched network.

The following diagram shows the scenarios covered later in this section.

Figure 47. Summary of Deployment Scenarios



A slightly different view of the same options is given in the following table:

Table 17. Summary of Deployment Scenarios

Connectivity	**Network Layer**	**Clusters Configuration**
MAN	Layer 2	Stretched Clusters
MAN	Layer 2	Separate Clusters
MAN	Layer 3	Separate Clusters
WAN	Layer 3	Separate Clusters

The following four figures show the logical architecture of the four different deployment models. The connectivity type is indicated in each figure caption.

Figure 48. MAN Connectivity – Stretched Layer 2 Clusters

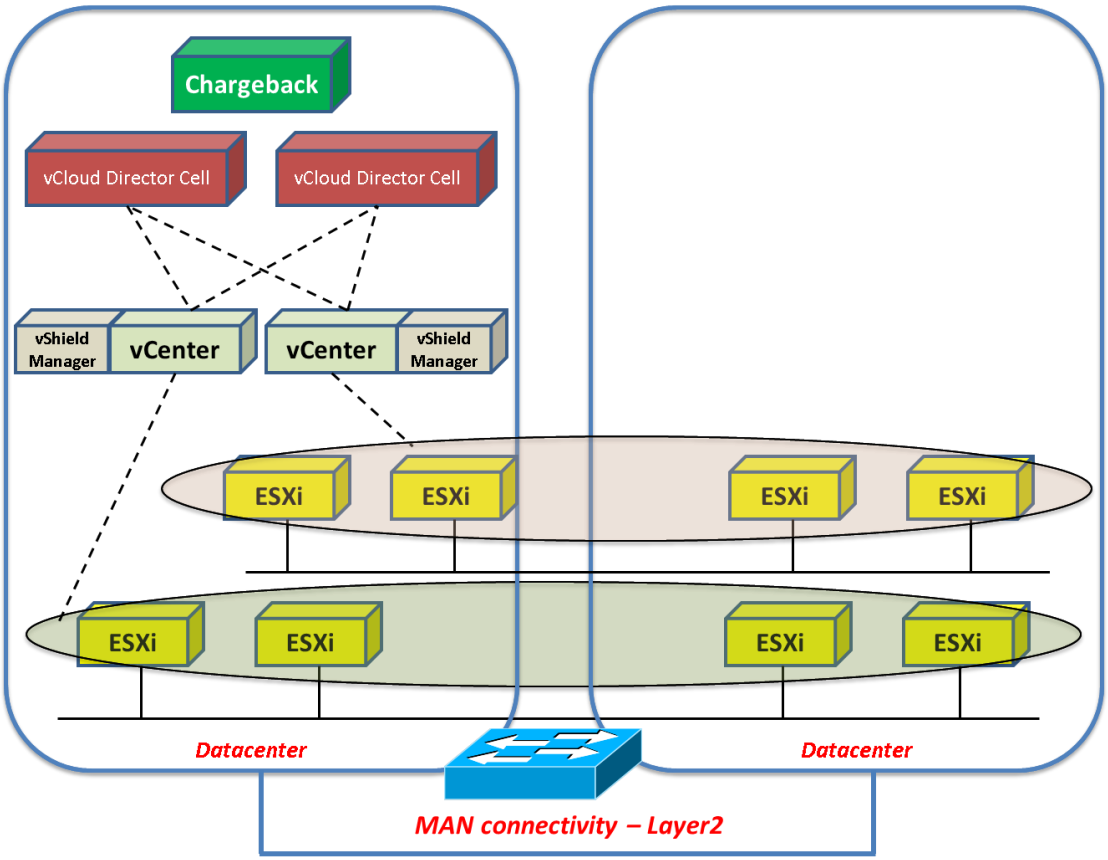


Figure 49. MAN Connectivity – Separate Layer 2 Clusters

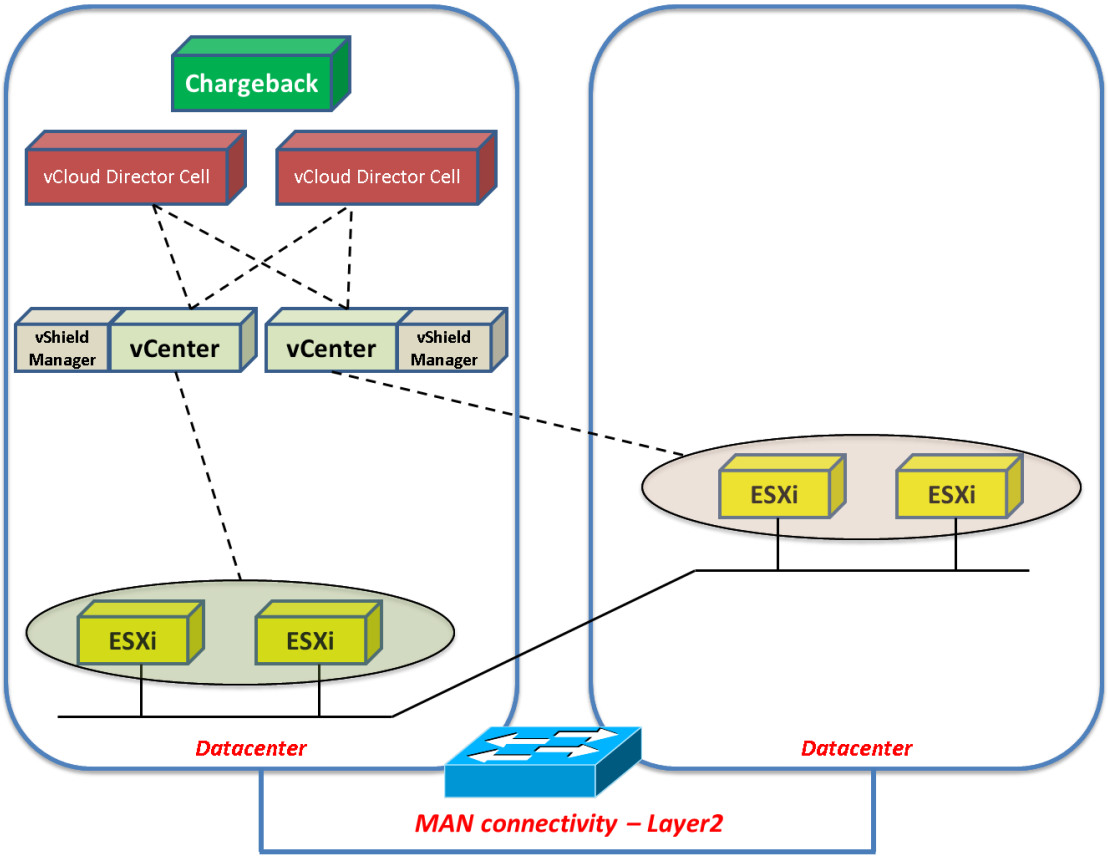


Figure 50. MAN Connectivity – Separate Layer 3 Clusters

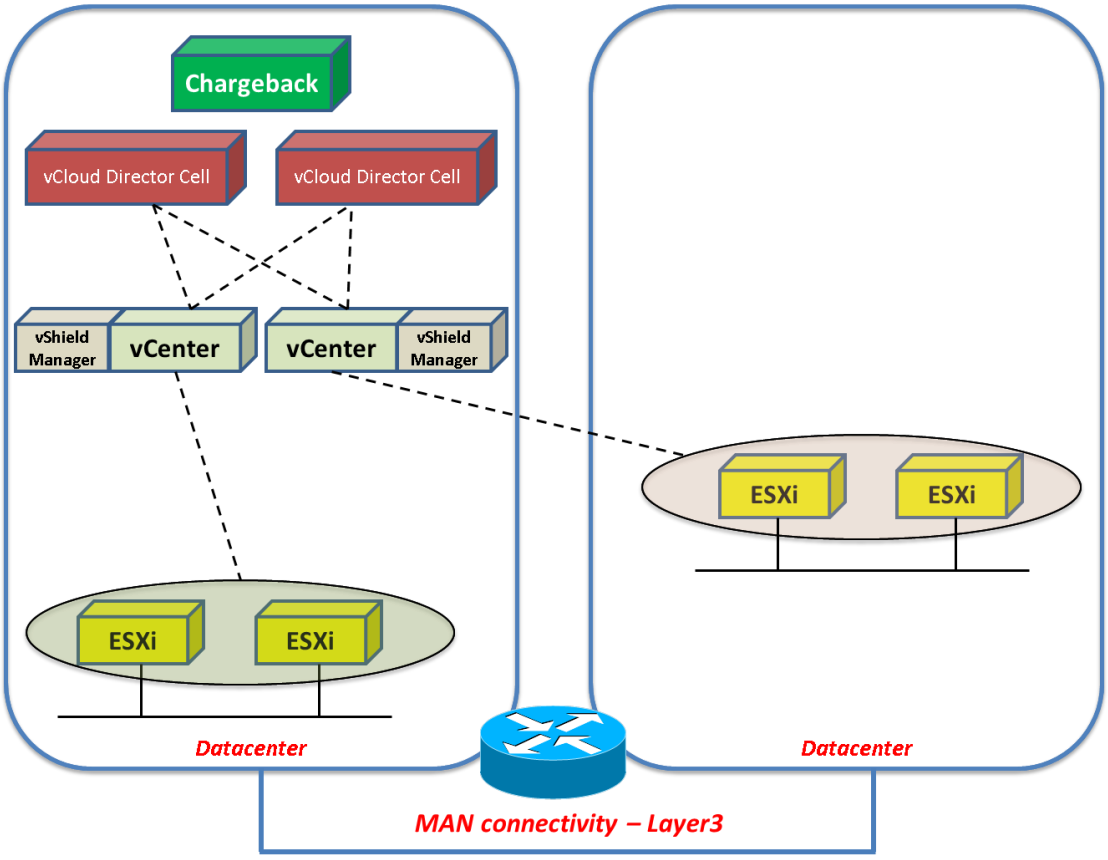
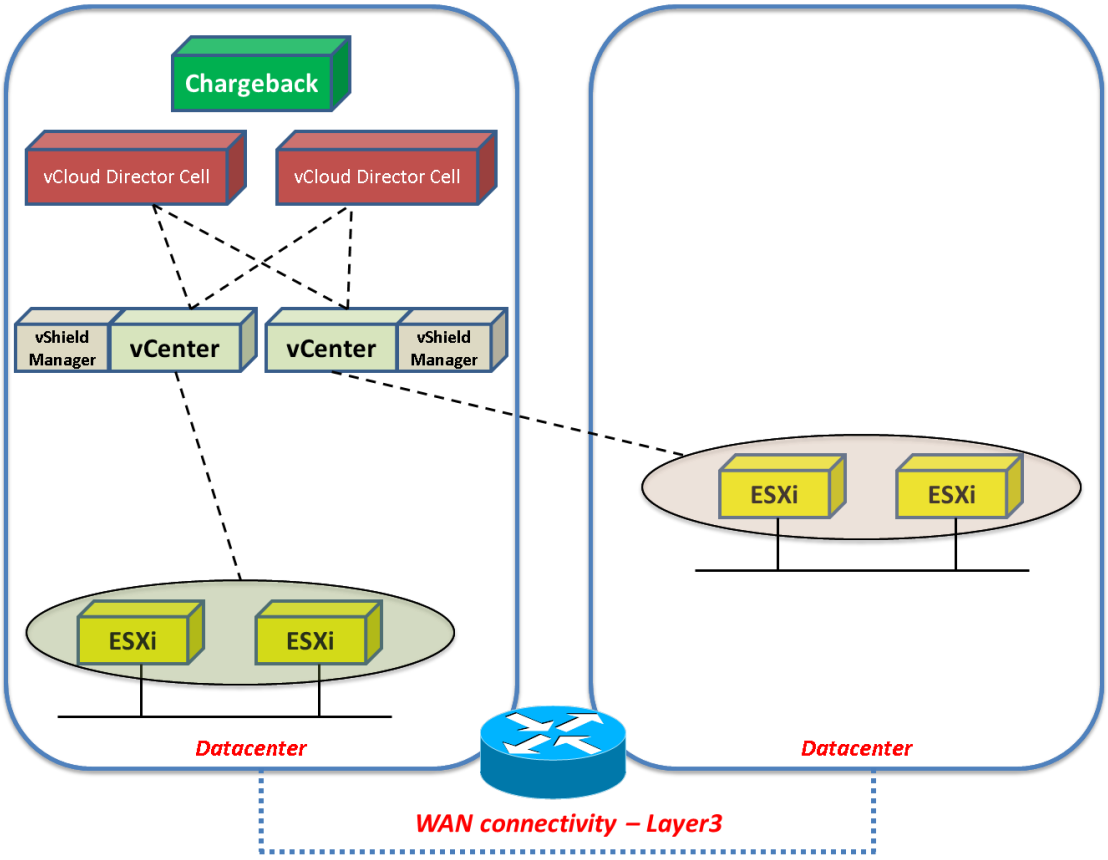


Figure 51. WAN Connectivity – Layer 3 Clusters



Note Though the diagrams show two vCenter Servers managing two different clusters, the same concepts apply with a single vCenter that manages two clusters. For convenience, vCenter servers are always shown located close to the vCloud Director cells and far from the ESXi hosts. The same supportability considerations apply in scenarios where the vCenter servers are located close to the ESXi hosts and far from the vCloud Director cells.

8.5 Supportability Considerations for Single Site Deployments

VMware supports vCloud Director 5.1 in MAN scenarios (as described in section 8.3, Multisite Terminology). Some supportability considerations are as follows:

- All provider workloads, with the exception of vCenter Server and vCloud Networking and Security Manager instances, must be deployed in a single location.
- Clusters backing provider virtual datacenters can be deployed in different locations if connectivity between locations has latency requirements as described in section 8.3, Multisite Terminology. Minimize the likelihood of a path failure across datacenters that might partition the provider workloads.
- vCenter Server and vCloud Networking and Security Manager instances managing and servicing clusters in distributed locations can be deployed either close to the vCloud Director core components (vCloud Director cells and vCloud Director database) or close to the clusters they manage.

Architects implementing a single site vCloud across different locations should deploy the various components with consideration given to sensitive operations such as vApp copies so that the deployment is fully optimized and the architecture takes into account network chokepoints (especially in terms of bandwidth) that can exist even in a MAN scenario. This has to do more with optimization than supportability.

Stretched clusters (which include stretched vSphere DRS clusters and stretched storage) are fully supported when implemented with the storage vendor-neutral guidance documented in this section. Stretched clusters (which require 10ms or lower latency) can enable increased flexibility in both tenant and provider workload placement.

Note The generic single site considerations in this section apply to tenant deployments that have latency within 20ms. Stretched clusters for tenant workloads are only supported when sites have latency within 5ms or 10ms (depending on the vSphere release and underlying storage technology used). At 5ms or 10ms latency (depending on the vSphere release and underlying storage technology used), the location of provider infrastructure components is more flexible. It is recommended that you:

- Follow the above-recommended guidance for any single site deployment within 20ms.
- Follow specific recommended guidance for vCloud Director deployments on top of vSphere stretched clusters within 10ms (these practices are dictated by the underlying storage solution supporting the stretched cluster, and may override the vendor neutral stretch cluster recommended practices in this document).

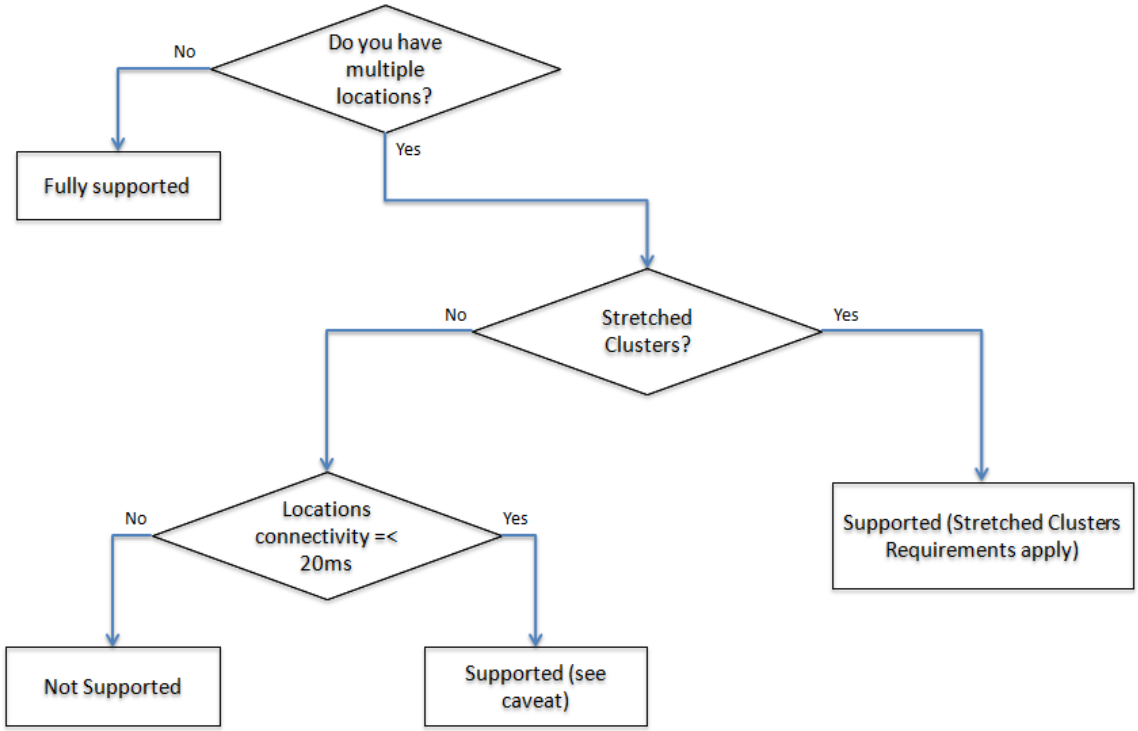
See the VMware vCloud Blog article, *Stretched vCloud Director Infrastructure* (http://blogs.vmware.com/vcloud/2013/01/stretched-vcloud-director-infrastructure.html) for more information.

8.6 Multisite Supportability Considerations

VMware does not currently support distributed vCloud Director 5.1 deployments in multisite scenarios. It is not possible to instantiate provider virtual datacenters that are located across a WAN (as described in Section 8.3, Multisite Terminology).

The following diagram summarizes the supportability options with associated constrains and requirements.

Figure 52. Supportability Flowchart



9. Hybrid vCloud Considerations

A hybrid vCloud incorporates a combination of vCloud instances and can 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 or be required 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.

Note This section focuses on workload mobility within a hybrid vCloud enabled by vCloud Connector. It does not discuss the hybrid vCloud governance enabled by vCloud Automation Center.

Figure 53. Hybrid vCloud Example

9.1 vCloud Connector

With the emergence of cloud computing, private enterprises must manage multiple vCloud instances, both private and public. Given these public and private options, ease of migrating workloads between vCloud instances becomes increasingly important.

vCloud Connector allows users to connect to vCloud instances based on vSphere or vCloud Director and manage them under a single interface. Through the vCloud Connector interface, users can view, copy, and operate workloads across internal datacenters and private or public vCloud instances.

vCloud Connector provides point-to-point reliable transfers between vCloud instances by using a checkpoint restart mechanism. If a transfer between nodes fails, vCloud Connector can restart the task and continue the copy from where it stopped, rather than having to start at the beginning of the file as required by a standard HTTP upload. vCloud Connector also uses HTTPS so that transfers are secure.

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 Web Client.

9.1.1 vCloud Connector Placement

Workload copy operations use the vCloud Connector appliance as an intermediary, so you must consider network latency and bandwidth between vCloud instances. For some use cases, it might be preferable to run multiple instances of vCloud Connector across multiple vCenter Server instances to avoid network latency or consuming excessive bandwidth.

Figure 54. vCloud Connector Basic Transfer Path

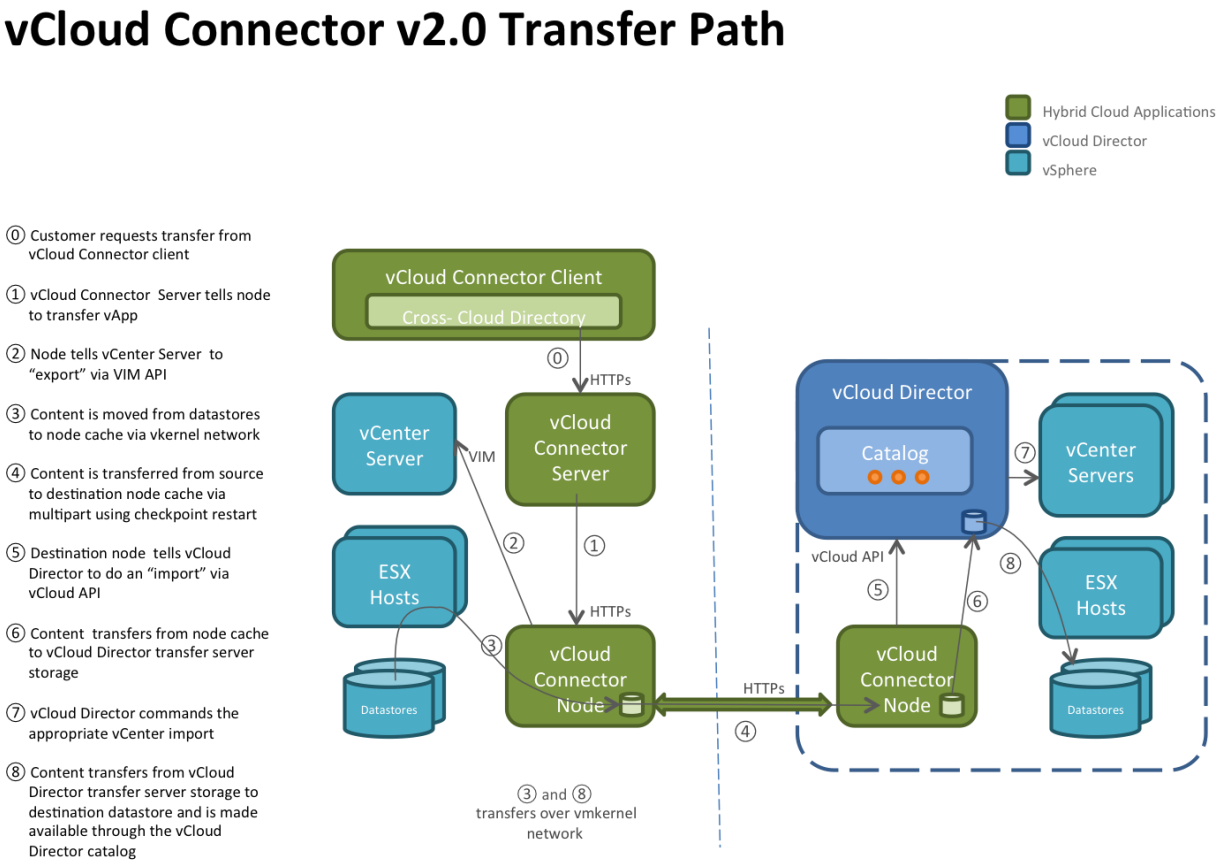
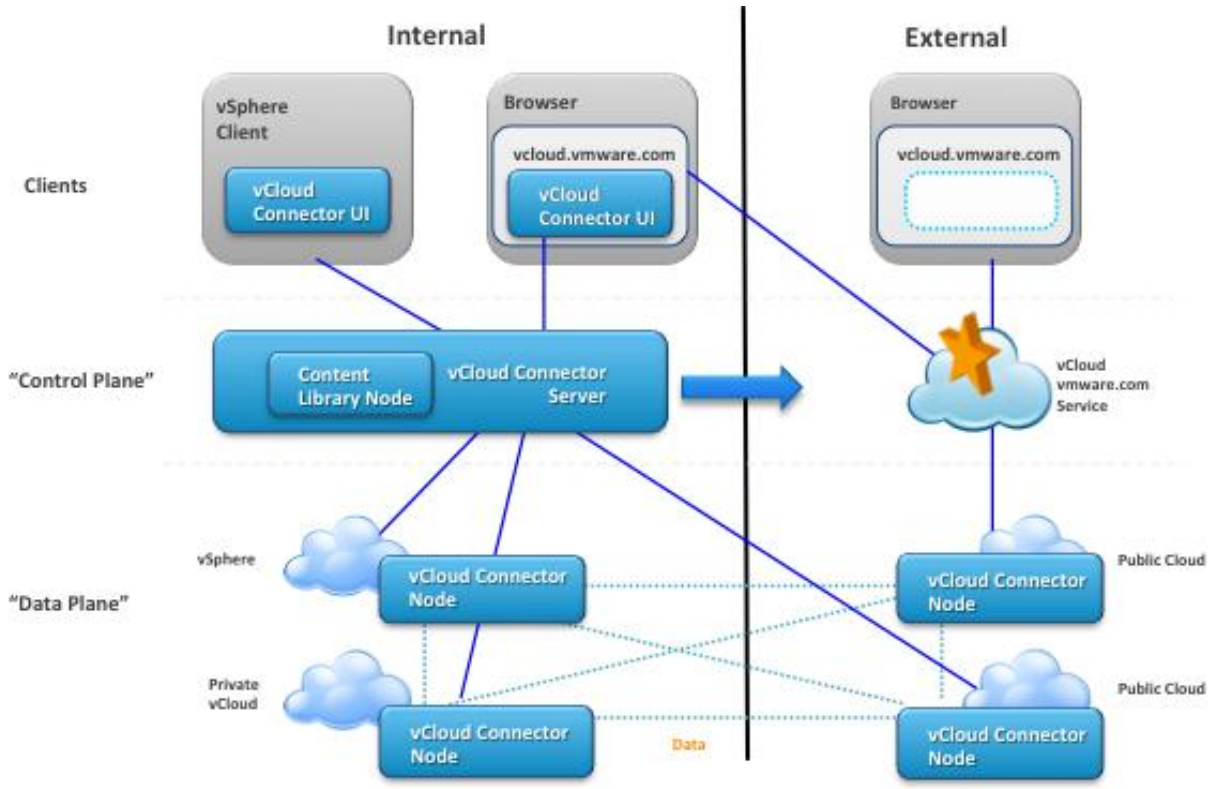


Figure 55. vCloud Connector Architecture



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. The following migrations are possible:

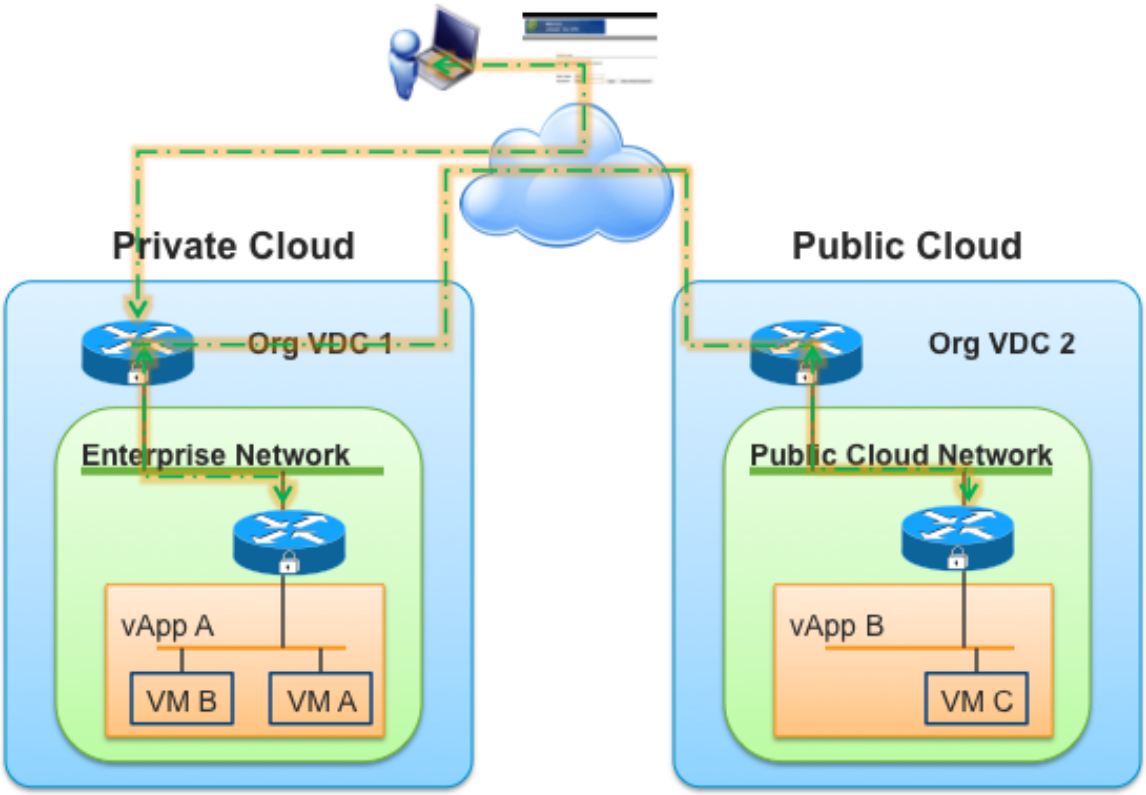
- vSphere to vCloud:
  - Migrate vSphere workloads into private or public vCloud instances.
  - Populate vCloud Director with templates from vSphere.
- vSphere to vSphere: Balance workloads across multiple vSphere instances.
- vCloud to vCloud: Move workloads between private or public vCloud providers.

9.1.3 Additional vCloud Connector 2.0 Features

vCloud Connector has the following additional features:

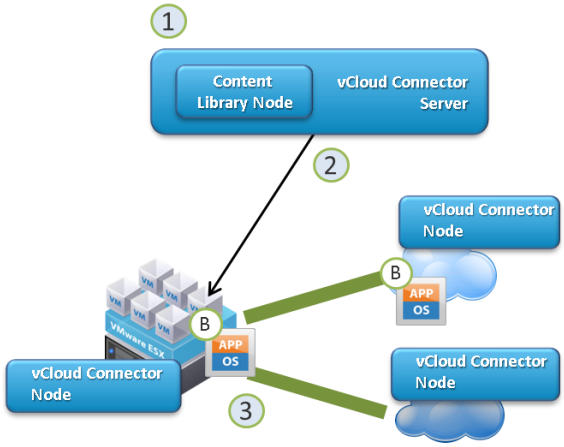
- Datacenter extension:
  - Layer 2 extension from existing enterprise network to a public vCloud over secure SSL VPN tunnel.
  - Ability to move a virtual machine from an enterprise network (vSphere or vCloud Director) to the public vCloud and retain the same IP and MAC addresses.
  - Requires version 5.1 of vSphere, vCloud Networking and Security Manager, and vCloud Director.

Figure 56. vCloud Connector Datacenter Extension



- Content sync:
  - Public vSphere folder, or private or public vCloud catalog.
  - Subscribe to the published folder or catalog from one or more vSphere instances, or private or public vCloud instances.
  - New or modified templates in the published folder or catalog are automatically copied to the subscribing vCloud instances that leverage the vCloud Connector secure transfer mechanism.
  - Folders and catalogs are kept synchronized across vCloud instances.
  - Supports vSphere 4.x (and later), and vCloud Director 1.5 (and later).

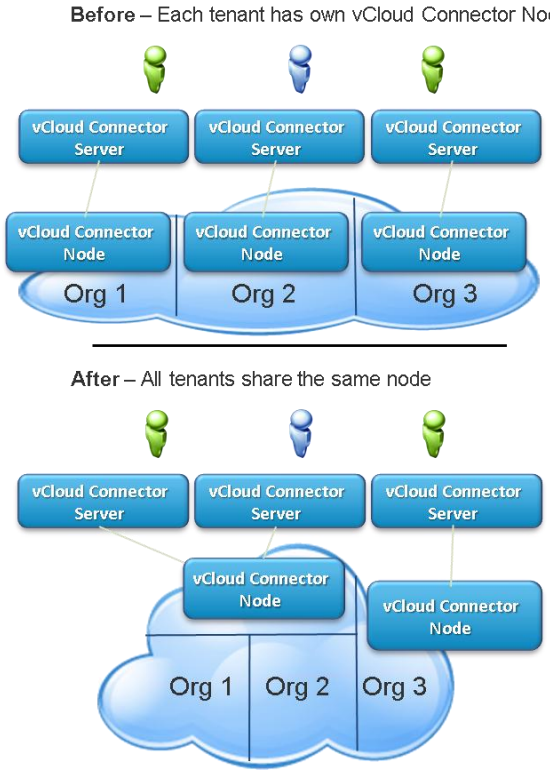
Figure 57. vCloud Connector Content Sync



1. vCloud Connector stores info on published folders/catalogs in its Content Library database
2. vCloud Connector periodically polls for changes in published folders/catalogs
3. vCloud Connector triggers copy of modified / new templates to the subscribing folders/catalogs across clouds

- Multitenant vCloud Connector node:
  - Multitenant node is deployed and managed by the service provider.
  - Choice of single or multitenant node.
  - Secure multitenancy with isolation for logging and troubleshooting.
  - Supports vCloud Director 1.5 (or later).

Figure 58. vCloud Connector Multitenant vCloud Connector Node



- Multi-tenant node deployed/managed by service provider
- Choice of single **or** multi-tenant node
- Secure multi-tenancy with isolation for logging & troubleshooting
- Ease of vCloud Connector adoption helps customers get to Hybrid Cloud faster
- Support vCloud Director 1.5+

9.1.4 vCloud Connector Limitations

vCloud Connector has the following restrictions:

- Currently, there is no way to have predefined vCloud instances display in vCloud Connector. Each user must manually add all vCloud instances that they intend to access to vCloud Connector. There are no vCloud instances 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 vCloud instances can be added to a vCloud Connector instance, so instruct users to use only the proper vCloud Connector instance for their needs.
- The transfer process caches virtual machines in two different storage 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 a consistent view of workloads across multiple vCloud instances 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 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 must 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 vCloud connections for a single vCloud Connector is five (vCloud Director or vSphere).

Note vCloud Connector 1.5 does not support the vCloud 5.1 Suite. vCloud Director 5.1 requires vCloud Connector 2.0 or later.

10. References

For additional information, see the documents listed in the following table.

Table 18. Reference Documentation

Topic	Document
vCloud Director	*vCloud Director Security Hardening Guide* (http://www.vmware.com/files/pdf/techpaper/VMW_10Q3_WP_vCloud_Director_Security.pdf) Go to the VMware vCloud Director documentation site for the following vCloud Director documentation (http://www.vmware.com/support/pubs/vcd_pubs.html): • *vCloud Director Installation and Configuration Guide* • *vCloud Director Administrator's Guide* *What's New in VMware vCloud Director 1.5 Technical Whitepaper* (http://www.vmware.com/resources/techresources/10192)
vCloud Automation Center	Go to the VMware vCloud Automation Center documentation site for the following vCloud Automation Center documentation (https://www.vmware.com/support/pubs/vcac-pubs.html): • *vCloud Automation Center Installation Guide* • *vCloud Automation Center Reference Architecture* • *vCloud Automation Center Operating Guide* • *vCloud Automation Center Self-Service Portal* • *vCloud Automation Center Extensibility Guide* *What's New in VMware vCloud Automation Center 5.1* (http://www.vmware.com/resources/techresources/10340)
vCloud API	Go to the VMware vCloud Director documentation site for the following vCloud Director documentation (http://www.vmware.com/support/pubs/vcd_pubs.html): • *vCloud API Specification* • *vCloud API Programming Guide*

Topic	Document
vSphere	VMware vSphere documentation: • VMware vSphere 5 documentation: (https://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-pubs.html) • *What's New in VMware vSphere 5.1* (http://www.vmware.com/files/pdf/products/vsphere/vmware-what-is-new-vsphere51.pdf) • *Performance Best Practices for VMware vSphere 5.0* (http://www.vmware.com/resources/techresources/10199) • *VMware vCenter Server 5.1 Database Performance Improvements and Best Practices for Large-Scale Environments* (http://www.vmware.com/files/pdf/techpaper/VMware-vCenter-DBPerfBestPractices.pdf)
vCloud Networking and Security	• Administration Guide (https://www.vmware.com/support/pubs/vshield_pubs.html) • VMware vCloud Networking poster (http://www.vmware.com/files/pdf/techpaper/VMware-vCloud-Networking-Poster.pdf) • *VXLAN Performance Evaluation on VMware vSphere 5.1* (http://www.vmware.com/files/pdf/techpaper/VMware-vSphere-VXLAN-Perf.pdf) • *Replacing Default vCenter 5.1 and ESXi Certificates* (http://www.vmware.com/files/pdf/techpaper/vsp_51_vcserver_esxi_certificates.pdf)
vCenter Chargeback	• *vCenter Chargeback User's Guide* (https://www.vmware.com/support/pubs/vcbm_pubs.html)
vCenter Orchestrator	• *vCenter Orchestrator Developer's Guide* (https://www.vmware.com/pdf/vco_410_developers_guide.pdf) • *VMware vCenter Orchestrator Administration Guide* (https://www.vmware.com/pdf/vco_410_admin_guide.pdf) • *vCenter Server 4.1 Plug-In API Reference for vCenter Orchestrator* (https://www.vmware.com/support/orchestrator/doc/vco_vsphere41_api/index.html)

Appendix A: Availability Considerations

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

Table 19. vCloud Availability Considerations

Component	Availability	Failure Impact
Maintaining Running Workload		
vSphere hosts	Configure all vSphere 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 vCloud Networking and Security Edge appliances.	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. vSphere HA Admission Control makes sure sufficient resources are available in the cluster to restart the virtual machines. VMware recommends the admission control policy *Percentage of cluster resources* , as it is flexible and also provides resource availability. For a description of design guidelines about increasing availability and resiliency, see the white paper *VMware High Availability: Deployment Best Practices: VMware vSphere 4.1* (http://www.vmware.com/files/pdf/techpaper/VMW-Server-WP-BestPractices.pdf.) VMware also recommends configuring vCenter to proactively migrate virtual machines off a host in the event that the host's health becomes unstable. Rules can be defined in vCenter to monitor host system health.

Component	Availability	Failure Impact
Virtual machine resource consumption	vSphere DRS and vSphere Storage DRS migrate virtual machines between hosts to balance the cluster and reduce the risk of a “noisy neighbor” virtual machine monopolizing CPU, memory, and storage resources within a host at the expense of other virtual machines running on the same host.	No impact. Virtual machines are migrated between hosts with no downtime by vSphere DRS or vSphere Storage DRS.
	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.	No impact. Virtual machines and vSphere hosts are throttled by Storage I/O Control 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 *Storage I/O Control Technical Overview and Considerations for Deployment* (http://www.vmware.com/files/pdf/techpaper/VMW-vSphere41-SIOC.pdf).
vSphere host network connectivity	Configure port groups with a minimum of two physical paths to prevent a single link failure from impacting platform or virtual machine connectivity. This includes management and vMotion networks. Use the load-based teaming mechanism 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 required.
vSphere host storage connectivity	vSphere hosts are configured with a minimum of two physical paths to each LUN or NFS share to prevent a single storage path failure from resulting in an impact to service. Path selection plug-in is selected based on the storage vendor’s design guidelines.	No impact. Failover occurs with no interruption to service.
Maintaining Workload Accessibility		
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 providing near zero downtime.

Component	Availability	Failure Impact
VMware vCenter Database	VMware vCenter Database resiliency is provided with vCenter Heartbeat if MS SQL is used or Oracle RAC.	vCenter Heartbeat or Oracle RAC provides a clustered solution for a vCenter database with fully automated failover between nodes providing zero downtime.
vCloud component databases (vCloud Director and Chargeback)	VMware vCloud component database resiliency is provided through database clustering. Microsoft Cluster Service for SQL and Oracle RAC are supported.	Microsoft Cluster Service and 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. Though not required to maintain workload accessibility, clustering the chargeback database protects the ability to collect chargeback transactions so that providers can accurately produce customer billing information.
VMware vCenter Chargeback	Multiple Chargeback, vCloud, and vCloud Networking and Security Manager data collectors are installed for active/passive protection.	If one of the data collectors goes offline, the other picks up the load so that transactions continue to be captured by vCenter Chargeback.
vCloud Infrastructure Protection		
Component	**Availability**	**Failure Impact**
Manager	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. VMware highly recommends scheduling backups of vCloud Networking and Security Manager to an external FTP or SFTP server.	Infrastructure availability is impacted, but service availability is not. vCloud Networking and Security Edge devices continue to run without the management control, but no additional edge appliances can be added and no modifications can occur until the service comes back online.
vCenter Chargeback	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.	There is no impact on infrastructure availability or customer virtual machines. However, it is important to keep vCenter Chargeback available to preserve all resource metering data. Clustering the vCenter Chargeback servers protects the ability to collect chargeback transactions so that providers can accurately produce customer billing information and usage reports.

Component	Availability	Failure Impact
vCloud Director	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.	• Session state of users connected via the portal to failed instance is lost. Users can reconnect immediately. • No impact to customer virtual machines.
vCloud Networking and Security Edge	vCloud Networking and Security Edge can be deployed through the API and vCloud Director web console. To provide network reliability, VM Monitoring is enabled. In case of an edge guest OS failure, VM Monitoring restarts the edge device. vCloud Networking and Security Edge appliances use a custom version of VMware Tools and are not monitored by vSphere HA guest OS monitoring. vCloud Networking and Security Edge Gateway 5.1 provides the following HA capabilities: • Network HA – Customer can choose to deploy two appliances working in an active-passive configuration. A stateful failover occurs if the active dies. Then, a second appliance is deployed, and it becomes the new passive. • VMware HA – If the vSphere host dies, taking an appliance down with it, the appliance is restarted on another vSphere host • Application HA – The appliance internals are monitored for process lock-up, and so on, and VMware HA failover is triggered if problems are detected.	• Partial temporary loss of service. vCloud Networking and Security Edge is a possible connection into organization. • No impact to customer virtual machines or Virtual Machine Remote Console (VMRC) access. • All external network routed connectivity is lost if the corresponding edge appliance is lost.

Component	Availability	Failure Impact
vCenter Orchestrator	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). High availability for vCenter Orchestrator can be provided by vSphere HA and vSphere FT in addition to application-based clustering. If 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.	Temporary loss of access to end users interacting directly with vCenter Orchestrator. Disruption to workflows executed by vCenter Orchestrator. This includes workflows started by vCenter Orchestrator and workflows started by external applications.

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.
- 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).

vCloud Networking and Security Edge can be used to load balance vCloud Director cells, in addition to third-party external hardware or virtual appliances as load balancers.

The following table lists the design considerations for load balancing of vCloud Director cells.

Table 20. Load Balancer Considerations

Consideration	Detail
Security	A front-end firewall is typically deployed in front of the load balancer. In some environments, additional firewalls can be located between vCloud Director cells and the resource tiers managed by vCenter. Load balancers can also provide NAT/SNAT (source network address translation) for the clustered cells. VMware recommends securing access between cells and the other management and resource group components. Refer to the *vCloud Director Installation and Configuration Guide* for ports that must be opened.
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 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 (first upgrade one vCloud Director cell, then the other).
Requirements for multicell configurations	Multiple vCloud Director cells require NTP (Network Time Protocol), which is a design guideline for all elements of the vCloud infrastructure. See the white paper, *Timekeeping in VMware Virtual Machines* (www.vmware.com/files/pdf/Timekeeping-In-VirtualMachines.pdf) for more information on how to set up NTP.
Load balancer availability	Use at least two load balancers in a HA configuration 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 that is typically provided by the load balancer.
Rest API URL configuration	Map the vCloud service URL to the address that the load balancer provides. This is configured in the vCloud Director administrator GUI and in the load balancer configuration. Use this address to check the health status of the vCloud Director cell.
Awareness of multicell roles	Some vCloud Director cell tasks (such as image transfer) can consume significant 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.

Consideration	Detail
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.
vCloud Director cell status health checks	• Configure the load balancer service to check the health of individual vCloud Director cells. Because each cell responds by way of HTTPS, this can be configured through the IP and API end point URL. Load balancers might support other types of health checks. Example UI URL – https://my.cloud.com/cloud/ • Example API URL – https://my.cloud.com/api/versions In the second example, the versions supported by this endpoint are returned as XML. Check services periodically based on load. A good starting point is to check every five seconds.
Public IP/port	Specify the service IP appropriately before adding cells to the service group. Typically, port 443 (standard HTTPS) is the only port exposed.
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 can allow a vCloud Director configuration to send certain types of client traffic to dedicated cells. Although each cell 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 might try and load balance sessions. This can affect connection mapping through the load balancer as it is unaware of the balancing that occurring within the server group.

Appendix B: Security

This appendix provides security guidance relating to VMware security certifications, network access, Single Sign-On (SSO), and demilitarized zone (DMZ).

VMware Security Certifications

Third-party certifications such as Common Criteria (CC), Federal Information Processing Standards (FIPS), and the National Institute of Standards and Technology (NIST) Security Content Automation Protocol (SCAP) provide independent validation of the security of VMware products.

The Common Criteria (CC) certification is an industry recognized standard for computer security certification, which provides a set of requirements used to evaluate and certify the security of a system.

Federal Information Processing Standards (FIPS) is a standard for U.S. Government computer systems used by all nonmilitary agencies. The Federal Information Processing Standards (FIPS) 140 series specifies hardware and software requirements for cryptographic modules.

The Security Content Automation Protocol (SCAP) certification determines the ability of a product to implement the Security Content Automation Protocol for configuration and vulnerability scanning and remediation.

Organizations across a variety of verticals either require or recommend the use of products adhering to validation under Common Criteria (CC), Federal Information Processing Standards (FIPS) and SCAP.

For the latest information on the state of VMware security certifications, go to https://www.vmware.com/support/support-resources/certifications.html.

Common Criteria

The National Institute of Standards and Technology (NIST) and the National Security Agency (NSA) established a program under the National Information Assurance Partnership (NIAP) to evaluate IT product conformance to international standards.

Characteristics of the Common Criteria certification include the following:

- Internationally recognized standard (between 26 member nations).
- Mutually recognized by all nations (up to EAL4).
- ISO standard (ISO15408).

The Common Criteria program, officially known as the NIAP Common Criteria Evaluation and Validation Scheme for IT Security (CCEVS) is a partnership between the public and private sectors. This program is intended to help consumers select commercial off-the-shelf information technology (IT) products that meet necessary security requirements and to help manufacturers of those products gain acceptance in the global marketplace. VMware has participated in Common Criteria (CC) evaluation of products beginning with VMware ESX Server 2.5 and VMware VirtualCenter 1.2 in March 2006.

As of 1 Aug 2012, the NIAP have instituted multiple changes to the Common Criteria certification processes, including changes to the certification levels offered, and eliminating the “In Evaluation List”. The highest level of certification now available is EAL 2+ (Evaluation Assurance Level 2). This new designation is more robust than the previous EAL4+ certification level, which was the highest level previously attainable. Each successive level of the Common Criteria is harder to achieve and requires additional validation, testing, and documentation.

vSphere 5.1, including the new Single-Sign on components is currently under evaluation under the new EAL2+ certification. EAL2+ is now the highest level of certification available, and is at least equivalent to the previous EAL4+ designation.

Other CC evaluations are as follows:

- VMware vCloud Networking and Security v5.1.2 is undergoing Common Criteria Certification evaluation for EAL4+ under the old program.
- VMware vFabric™ tc Server 2.8.0 is going through CC certification under the new EAL2+ scheme.

Federal Information Processing Standards

The Federal Information Processing Standards (FIPS) publications are guidelines that set best practices for software and hardware computer security products.

FIPS 140-2 is the standard for Security Requirements for Cryptographic Modules. VMware products that are evaluated for Common Criteria also achieve the FIPS 140-2 security designation.

Security Content Automation Protocol

VMware vCenter Configuration Manager has been validated by NIST to have achieved the following SCAP certifications:

- Authenticated Configuration Scanner – The capability to audit and assess a target system to determine its compliance with a defined set of configuration requirements using target system logon privileges.
- Authenticated Vulnerability and Patch Scanner – The capability to scan a target system to locate and identify the presence of known vulnerabilities and evaluate the software patch status to determine compliance with a defined patch policy using target system logon privileges.

For more information on SCAP validated products, go to http://nvd.nist.gov/scapproducts.cfm.

Network Access Security

vCloud Networking and Security 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 21. Network Access Security Use Cases

Category	Description
Multisite vCloud deployment	The vCloud Networking and Security 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. Because vCloud Networking and Security also provides address translation, it is possible to deploy multiple organization virtual datacenters at different providers using the same RFC 1918 address space as long as unique subnets are used.
Single-site vCloud deployment	vCloud Networking and Security VPNs can be created between different organizations in the same vCloud Director instance or different networks within the same organization. For these deployments, the site-to-site VPN is used to secure sensitive traffic between networks over shared infrastructure.

Remote Site to vCloud VPN	A permanent secure connection from a router or firewall based VPN, for example, from Cisco or Juniper devices at a remote site to a vCloud environment with vCloud Networking and Security Edge. vCloud Networking and Security VPN is a standard IPsec implementation, and a wide range of devices can be used at the remote site (commercial or open source).
Client to vCloud VPN	Client software is generally not used with IPsec VPNs (an IPsec VPN is typically a permanent network-to-network tunnel), although clients with static IP addresses that implement pre-shared key authentication are supported.

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

When configuring a site-to-site VPN between different organization virtual datacenter networks in a vCloud environment (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 to provide compatibility between the edge peers. In comparison, configuring remote devices to connect to a VPN based on vCloud Networking and Security Edge requires an understanding of IPsec and the supported policies to successfully establish an encrypted tunnel.

The following IKE Phase 1 parameters are used by the vCloud Networking and Security Edge VPN:

- Main mode.
- Pre-shared key authentication mode.
- 3DES or AES128 encryption.
- SHA1 authentication.
- MODP Group 2 (1024 bits).
- SA lifetime of 28800 seconds (8 hours).
- Disable ISAKMP aggressive mode.

The following additional parameters for IKE Phase 2 are supported:

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

vCloud Networking and Security 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, the following ports must be open on firewalls between the two endpoints:

- Protocol 50 ESP.

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

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

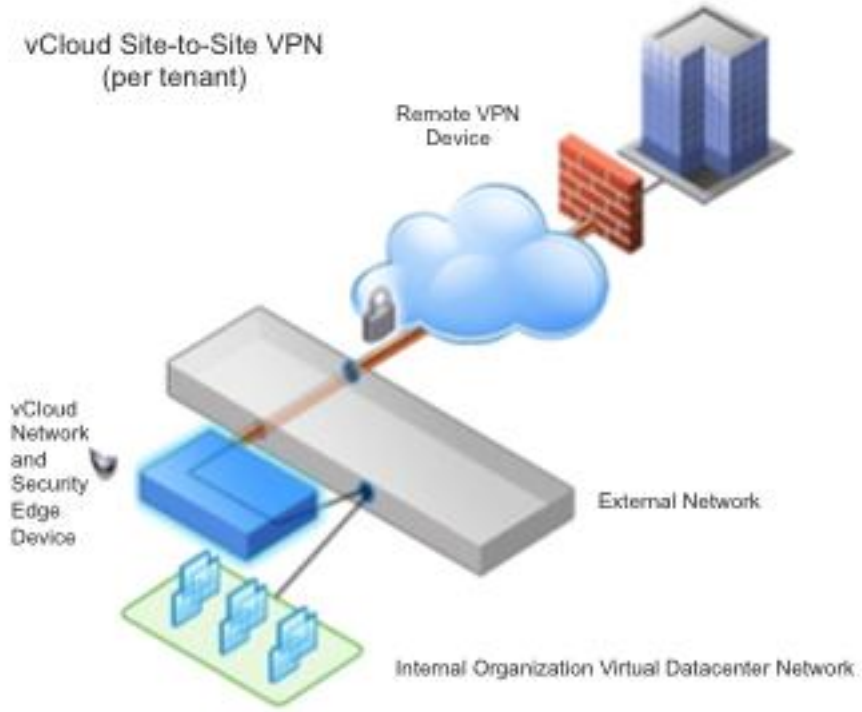
It is also possible for the remote VPN endpoints to be located behind an NAT device, although on both ends a static one-to-one NAT is required for the external IP address.

VPNs are used to provide secure access to an organization's remote networks, and consumers should be aware of any security implications. A guideline for VPN configuration is to filter and restrict VPN traffic to destinations that are necessary. vCloud Director 1.5 (and later) can also apply firewall rules to VPN traffic (filtering was previously restricted only to the remote end of a VPN tunnel).

The vCloud Director IPsec VPN has a maximum of 10 sites per vCloud Networking and Security Edge device.

The following figure shows a sample configuration for site-to-site VPN connectivity.

Figure 59. Site-to-Site VPN connectivity



The following features are not currently supported in the any VPN traffic is Edge VPN implementation:

- Remote endpoints with dynamic IP addresses.
- Site-to-site VPNs at the vApp network level (available only to organization virtual datacenter networks).
- 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 be enabled only on routed networks).

Two-Factor Authentication

The following are options for providing two-factor authentication to a vCloud Solution:

- Enable SSPI support in vCloud Director 5.1 and delegate authentication to Active Directory, which has a number of two-factor solutions.
- Implement a third-party solution (for example, HyTrust Cloud Control).

vCloud Director 5.1 adds support for Security Support Provider Interface (SSPI), which is Microsoft's proprietary implementation of GSSAPI. SSPI is an API for obtaining numerous security services, including integrated Windows authentication. Using SSPI to delegate identity verification to Windows and Active Directory allows for the use of a number of authentication mechanisms such as secure token or two-factor authentication.

The following are two-factor authentication design implications:

- The authentication method must be set to Kerberos to enable SSPI.
- The Service Principal Name (SPN) must be specified. The SPN is a name that a client uses to uniquely identify an instance of a service.
- A KeyTab file is needed to enable authentication for the SPN.
- Using SSPI implies that the workstation must be a member of an Active Directory domain.
- By using SSPI, vCloud Director is allowing a trust relationship to Active Directory to perform the authentication on behalf of vCloud Director.
- Using native support for two-factor authentication solutions through SSPI enables service providers and enterprise organizations to achieve strong authentication without requiring manual configuration or integration of each individual virtualization host.
- Combining technologies from VMware and third parties such as RSA, Symantec, and HyTrust enables end-to-end security of vCloud infrastructure and accelerates time to market.
- VMware is continually evolving and adding new security components to its security framework, including capabilities such as controlling identities enterprise-wide, supporting more secure authentication methods, and providing interoperability with future vCloud Director releases.

Secure Certificates

To provide security for a vCloud service based on VMware vCloud Director, VMware requires the implementation of certificates and key management for secure access and authentication to the vCloud Director server during its installation.

vCloud Director uses symmetric encryption to protect sensitive data from eavesdroppers and unwanted guests, uses public-key encryption to exchange keys securely over an insecure transport, and supports certificates and their digital signatures to establish a trust relationship. This makes it possible to create a secure protocol and channel between the vCloud Director service and end-tenant that functions over an insecure connection without any previous interaction between the parties. This enables secure data transmission in a shared, multitenant environment.

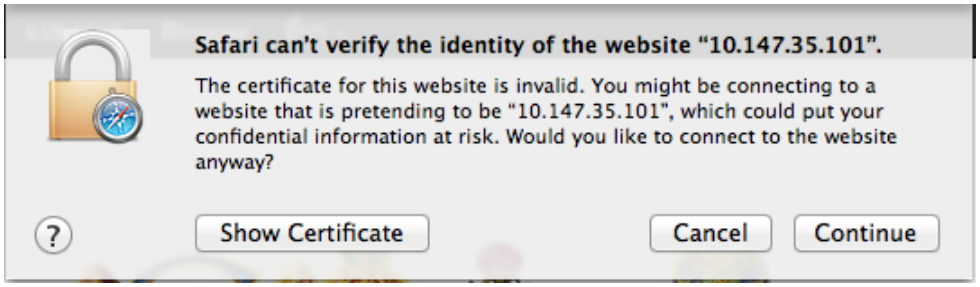
Secure Certificates Example

Deployment models: private, public, hybrid.

In the vCloud environment Secure Sockets Layer (SSL) and Transport Layer Security (TLS) protocols provide secure communication between the end tenant (client) and vCloud Director cell (server). The secure communication includes confidentiality and privacy of communication, message integrity and hashing, and authentication.

Web browsers display a warning message indicating that a site's identity cannot be trusted if a certificate has expired, or the certificate was issued by a certificate authority that is not trusted. It is the primary role of SSL/TLS to provide *confidentiality and privacy* of the communication, and to prevent MITM (man-in-the-middle) attacks, side channel attacks, and tax intended to compromise your privacy and security.

Figure 60. Example Error Message



Message Integrity and Hashing is the ability to guarantee that the data has not been modified during the protocol exchange and transmission.

Using *certificates for authentication* is a process of confirming identity. In the context of network interactions, authentication is the identification of one party by another party, and certificates are one way of supporting authentication.

Certificates or digital certificates are collections of data that uniquely identify and verify an individual, company, or other entity on the Internet. Certificates also enable secure, confidential communication between two entities. In the context of vCloud Director, server certificates are used to establish secure sessions between the cell server and clients using SSL and TLS.

The following figure shows the address bar for a web site that has been secured using SSL/TLS. The URL begins with https, and a padlock symbol is shown in the top right far corner of the browser.

Figure 61. Web Site Address Bar

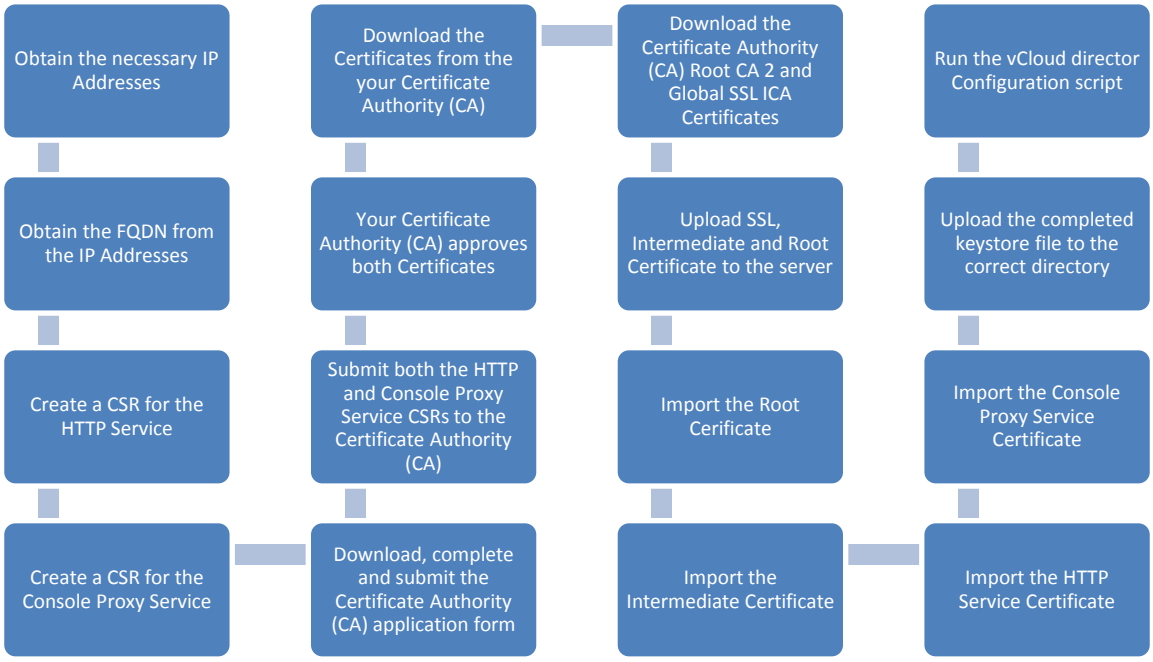
Types of SSL certificates are the following:

- Self-Signed certificate – Generated for internal purposes, not issued by a certificate authority (CA).
- Domain-signed certificate:
  - An entry level SSL certificate that can be issued quickly.
  - The only check performed is to verify that the applicant owns the domain where they plan to use the certificate.
  - No additional checks are done to confirm that the owner of the domain is a valid business entity.
- Fully authenticated SSL certificate:
  - First step to true online security and confidence building.
  - Takes slightly longer to issue because these certificates are granted only after the organization passes a number of validation procedures and checks to confirm the existence of the business, the ownership of the domain, and the user's authority to apply for the certificate.
- Server Gated Cryptography (SGC)-enabled SSL Certificate – Used for old browsers or clients that do not support 128/256 bit encryption.
- Wildcard certificate – Allows full SSL security to any host in domain.
- SAN (Subject Alternative Name) SSL certificate – Allows more than one domain to be added to a single SSL certificate.
- Code signing certificate – Specifically designed to make sure that downloaded software was not tampered with during the download.
- Extended Validation (EV) SSL certificates – Offers the highest industry standard for authentication and the highest level of customer trust.

For a private, hybrid, or public vCloud provider, VMware recommends implementing SSL certificates from a Trusted CA.

The following process flow outlines all of the steps to request, configure, obtain, and install an SSL certificate from a CA that can be used as for a CA for vCloud Director.

Figure 62. Requesting, Configuring, Obtaining and Installing an SSL Certificate from a CA



The following guidelines apply when considering SSL certificates:

- Understand and evaluate the different types of SSL certificates that are available and use one that matches your requirements.
- In a production environment, do not configure vCloud Director to use self-signed certificates. This is an insecure practice. Self-signed certificates are certificates that are digitally signed by the private key corresponding to the public key included in the certificate. This is done in place of a CA signing the certificate. By self-signing a certificate, you are attesting that you are who you say you are. No trusted third-party validation is involved.
- Self-signed certificates do not have a valid chain of signatures leading to a trusted root certificate and provide a weaker form of security. Although you can verify such a certificate is internally consistent, anyone can create one, so by examining the certificate, you cannot know if it is safe to trust the issuer or the site from which the certificate is coming. Nevertheless, self-signed certificates are common. For example, vCenter installations use a self-signed certificate by default.
- The server keystore highly sensitive because a compromise of the server key allows impersonation of the server and/or access to the encrypted traffic. Java keystores provide a password-protected method of securely storing private keys and their associated certificates. vCloud Director supports only the JCEKS format for keystore files. (Other formats that Java supports include PKCS12 and JKS. JKS is less secure and not recommended).

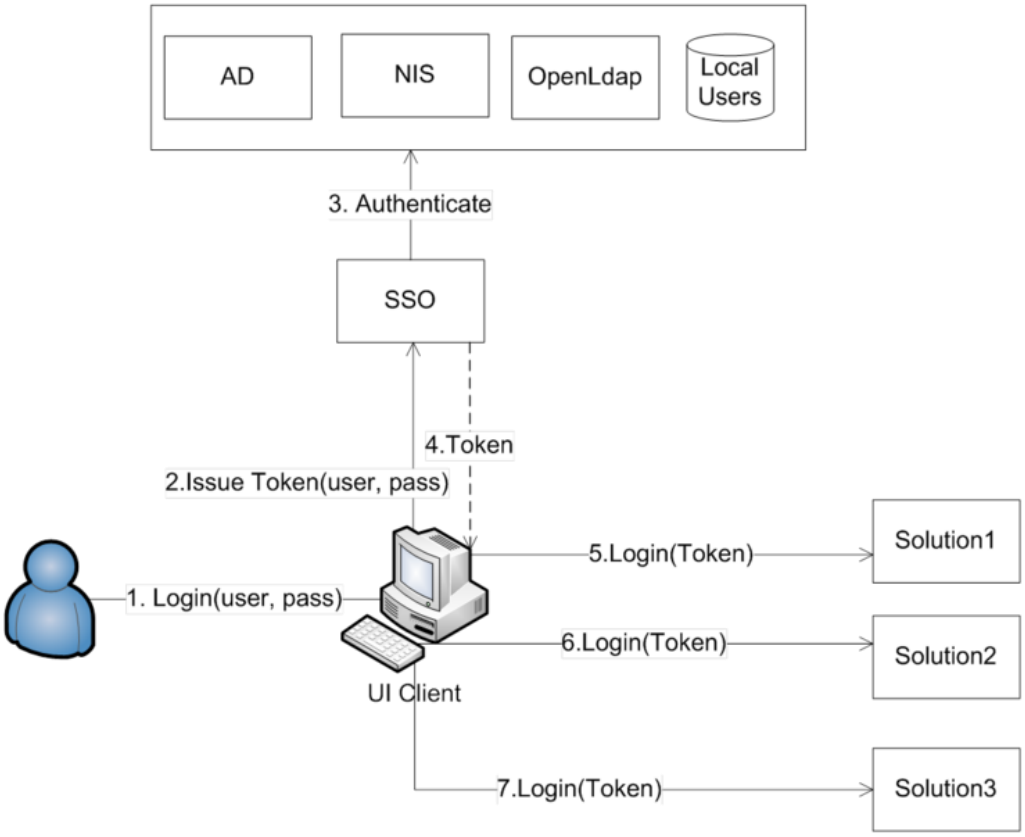
Single Sign-On

The use cases in this section show how the Web Single Sign-On (SSO) feature and configuration are exposed through vCloud Director 5.1 for both service provider and consumer architectures.

Use Case 1

In this use case, SSO applies to a single client and multiple back end services.. A user accesses multiple back end servers through a single UI client. The user provides credentials to the UI client, and the client validates them against the SSO server. If the validation is successful, the SSO server issues a Security Assertion Markup Language (SAML) token, which is then used by the UI client to access the different back end servers. The following diagram shows this use case.

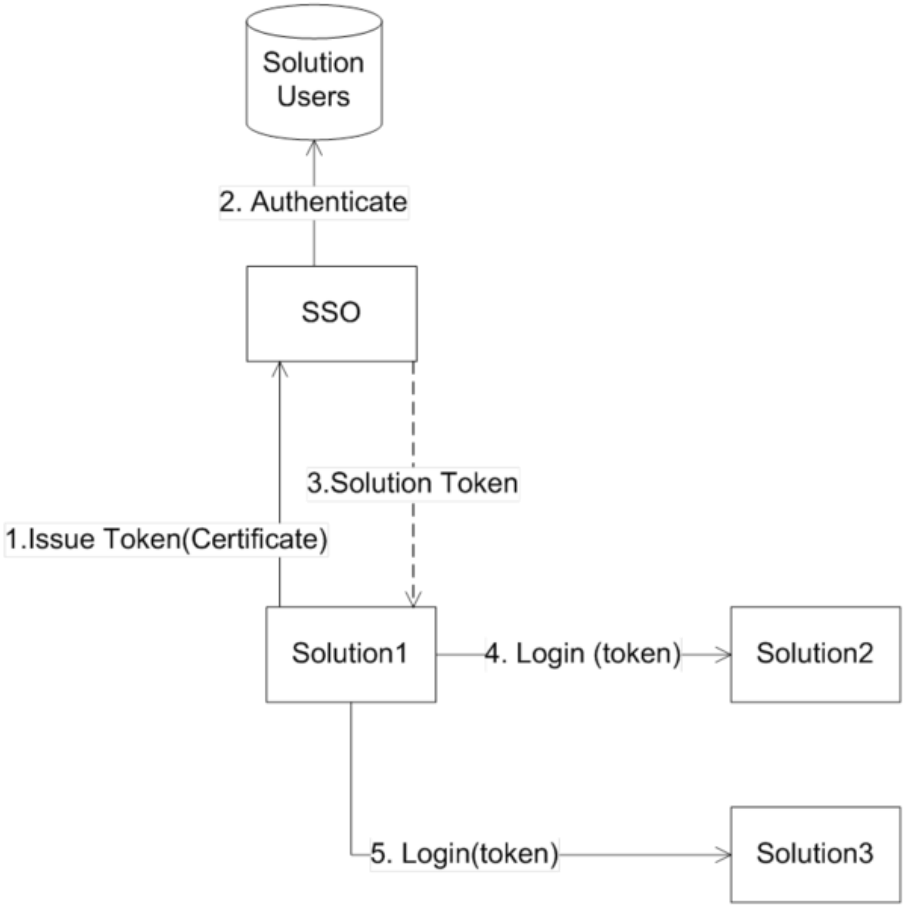
Figure 63. SSO Between a Single Client and Multiple Back End Services



Use Case 2

This use case illustrates solution-to-solution authentication in which an SSO user is assigned to each solutions. In the following figure, two solutions need to communicate with each other. Before they start to communicate, they must verify each other's identity. To do so, each solution initiates a request from the SSO server to issue a SAML token that asserts its identity. As part of this request, the solution proves its identity using its own private key. After the SSO server has issued a token, the solution can use that token to access any other solution as if it were a normal user. For this use case to work, each solution must be registered with its public key in the SSO server.

Figure 64. SSO Solution-to-Solution Authentication

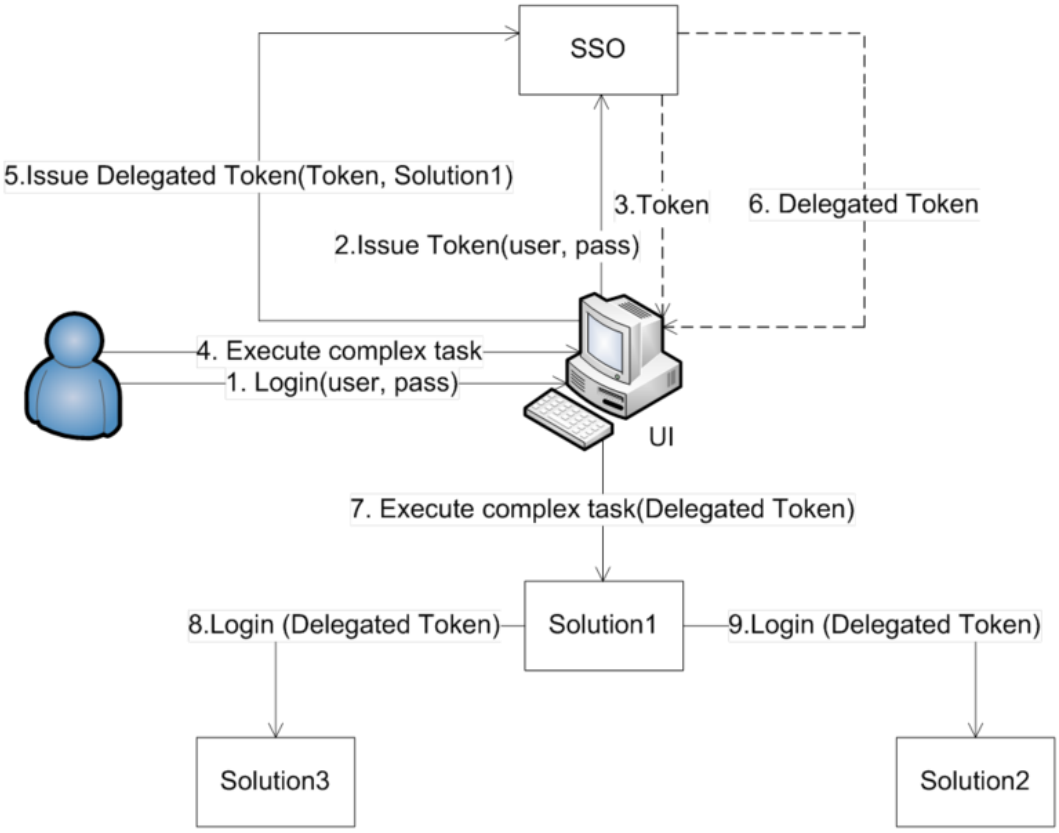


Use Case 3

In this use case, tasks are executed on behalf of a user (referred to as delegation). Some workflows that an end user initiates might require multiple solutions to communicate with each other, and SSO can support such workflows. Before the user can initiate the workflow through a given UI, the user must provide credentials. The UI validates the credentials against the SSO server, which issues a SAML token. The user then initiates a workflow.

In the following figure, the workflow requires Solution-1 to access Solution-2 and Solution-3 on behalf of the end user. As part of this process, the UI requests a *delegated* token from the SSO server for Solution-1 by providing the SAML token of the end user. The delegated token asserts that the user has granted Solution-1 the privileges to execute tasks on the user's behalf. After the UI has the delegated token, it gives it to Solution-1 to use to log in to Solution-2 and Solution-3.

Figure 65. Executing Tasks on Behalf of a User

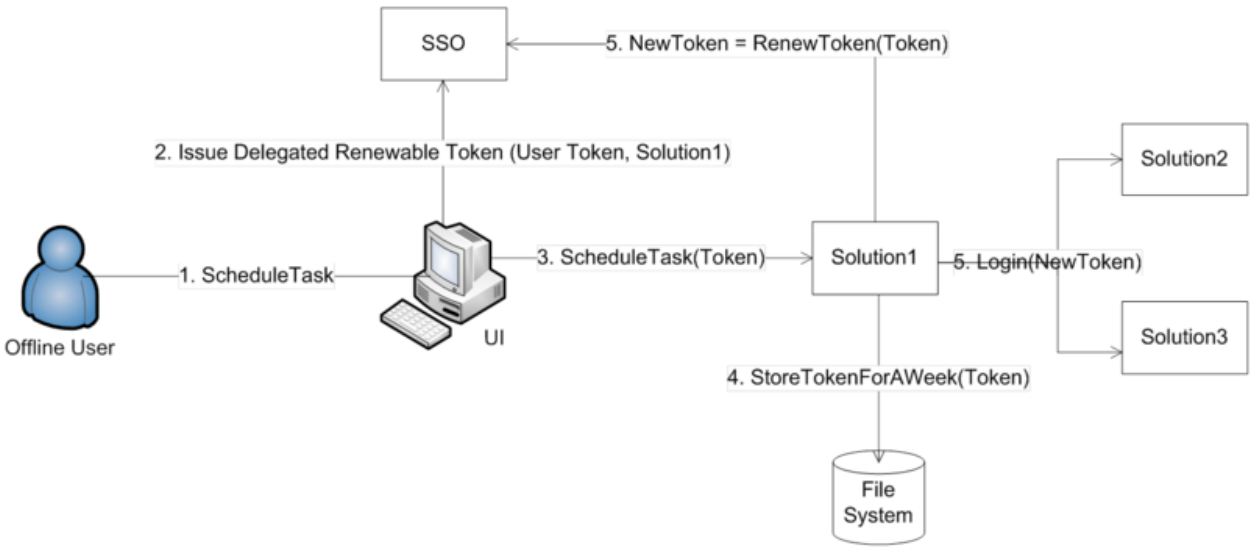


Use Case 4

This use case involves the scheduling of long-lived tasks and is referred to as *delegation and renew*. Some long running operations in the infrastructure require execution of long running tasks in the absence of the end user who initiated them. The SSO server supports such tasks by means of delegated and renewable tokens.

After a long running task is identified, the UI obtains a delegated and renewable token from the SSO server. It then passes the token to the solution, which performs the long running task. The solution persists the token in a non-secured way, as the token is self-secured. Every time the task is activated, the solution reads the token from the disk and makes a request to the SSO server to renew it. The user is not deleted from the system during this process.

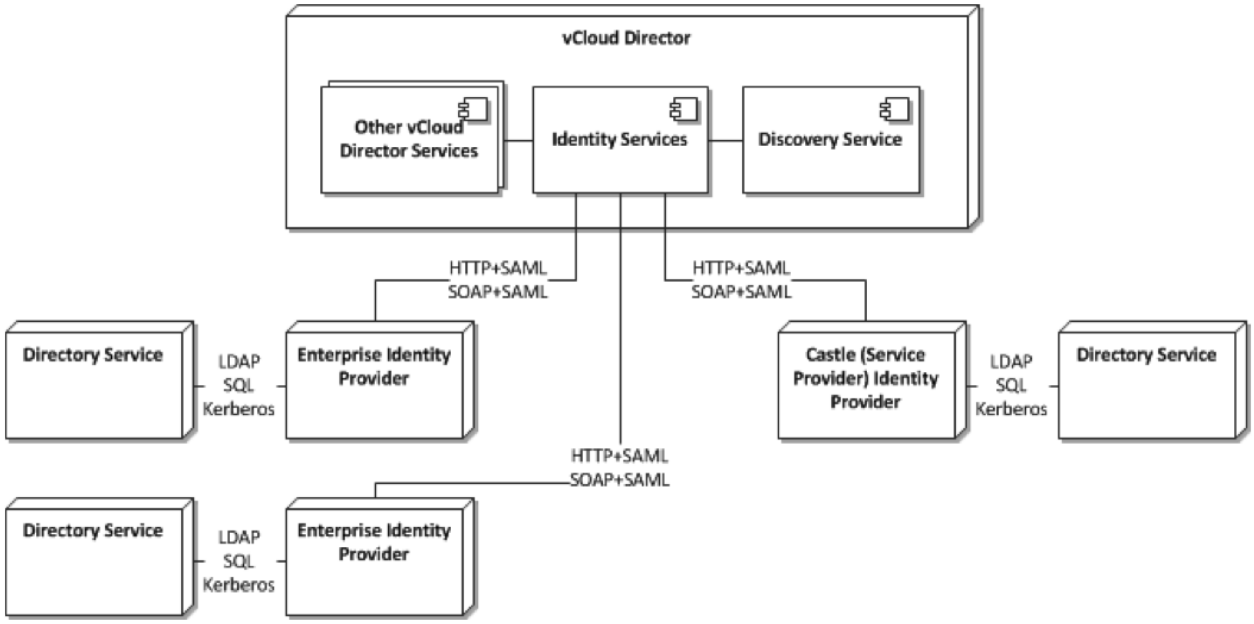
Figure 66. Scheduling Long-Lived Tasks



Consumer SSO Architecture Example

The following figure shows a consumer logical SSO deployment architecture.

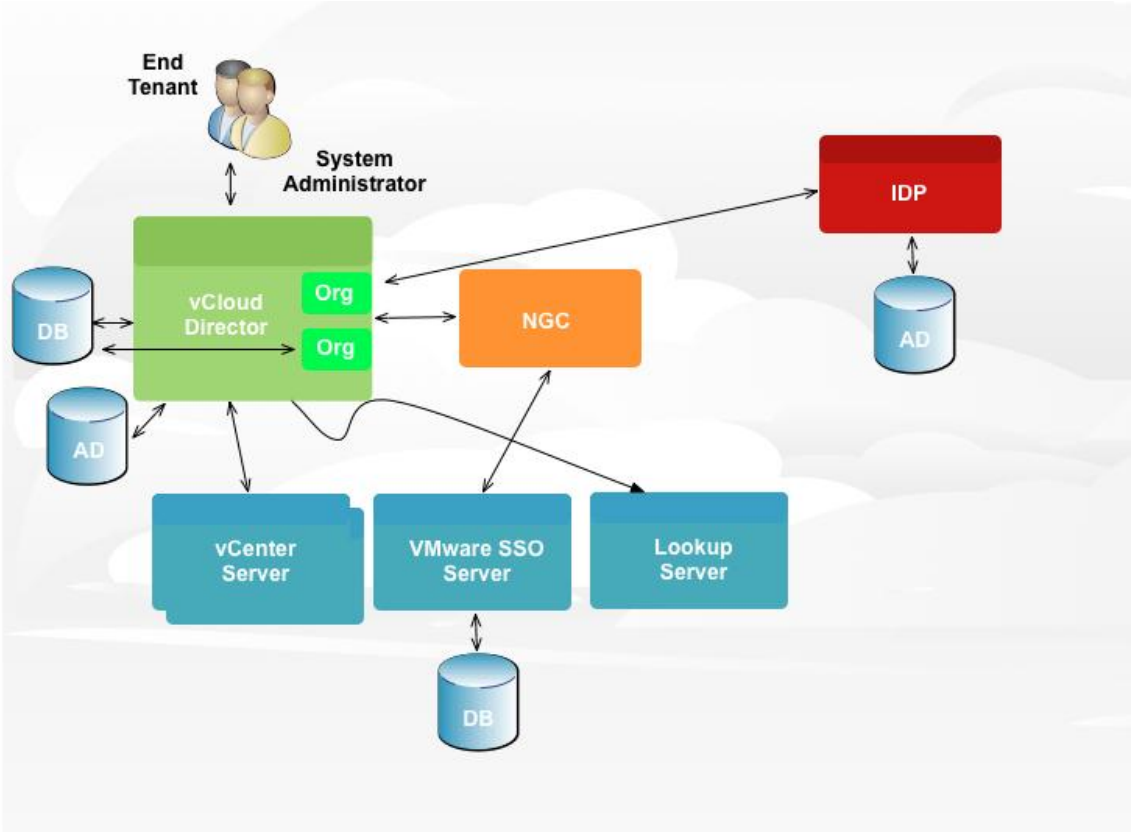
Figure 67. Consumer Logical SSO Deployment Architecture



vCloud Provider SSO Architecture Example

The following figure shows a vCloud provider SSO architecture example.

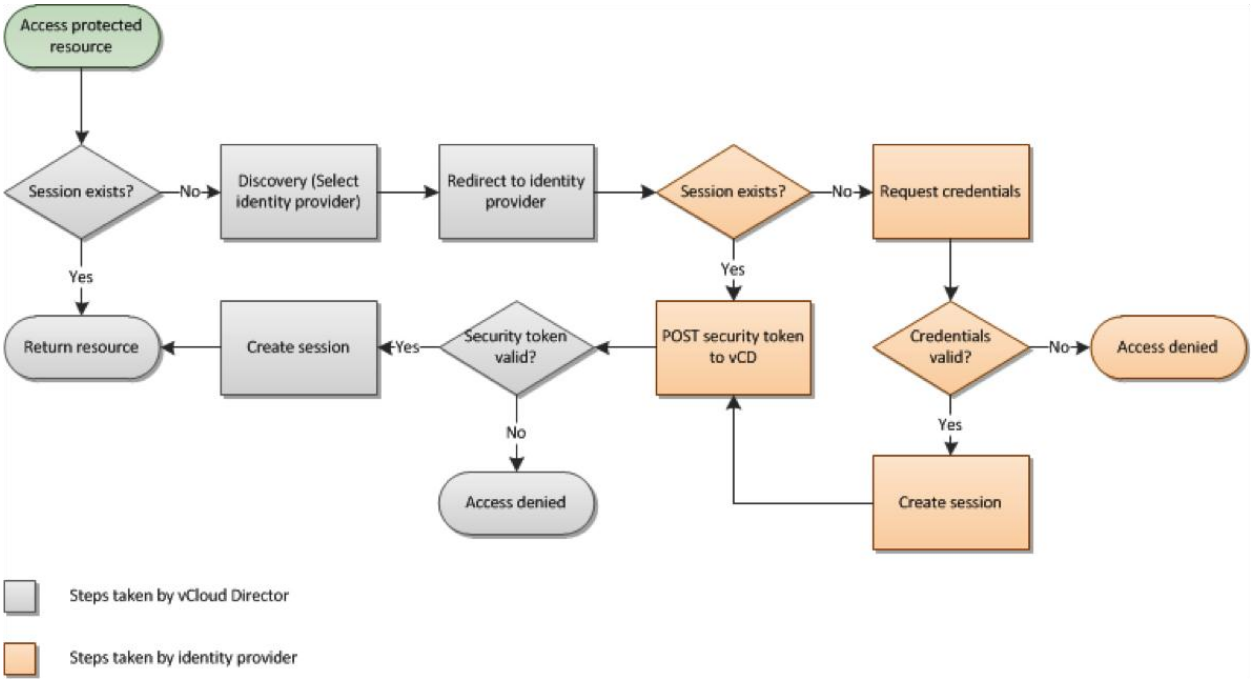
Figure 68. vCloud Provider SSO Architecture Example



SSO Authentication Workflow

The following figure shows an SSO authentication workflow.

Figure 69. SSO Authentication Workflow



You can use SSO to authenticate with the vCloud API in the following ways:

- Use the POST/sessions vCloud API, which accepts security tokens as the request body:
  - HTTP-Basic authentication – Logs in using user name and password to integrated identity provider for backwards-compatibility with vCloud Director v1.5.
  - SAML assertion – Verifies assertion is trusted.
  - Proprietary token – Verifies token from integrated identity provider is valid.
- Use the vCloud API `GET /org/{id}/hostedIdentityProvider/token`, which returns the security token for the integrated identity provider.
  - HTTP-Basic authentication logs in using the user name and password.
  - Kerberos – Verifies a Kerberos token using the Active Directory settings.
- Use the vCloud API `GET /org/{id}/identityProviders` which returns a list of identity providers (IdPs) federated with vCloud (currently integrated identity provider and possibly external identity provider) can be called anonymously.
- Use the vCloud API `GET /org/{id}/saml/authnRequest`, which returns the signed SAML `AuthnRequest`.

SSO Design Considerations

- Use SSO to provide a common service, both internally and externally.
- Use a supported IdP from VMware.
- The SAML assertion must contain attributes that vCloud Director can interpret.
- vCloud Director and the IdP must be time synchronized to within a few seconds.
- vCloud Director and the IdP must have valid endpoint certificates.
- Use consistent hostname (or IP address) when registering with the LookupService.
- If the SSO Server is not accessible and the accessibility issue cannot be resolved, use the SSO repoint tools that are that are packaged with SSO clients such as vCenter and the Web client.
- Consider the following identity sources: OpenAM, Active Directory Federation Services, Shibboleth.
- Provide a highly available SSO service.
- Deploying vCenter Single Sign-On as a cluster means that two or more instances of vCenter Single Sign-On are installed in high availability (HA) mode. vCenter Single Sign-On HA mode is not the same as vSphere HA. All instances of vCenter Single Sign-On use the same database and must point to the same identity sources. vCenter Single Sign-On administrator users, when connected to vCenter Server through the VMware vSphere Web Client, see the primary SSO instance. In this deployment scenario, the installation process grants admin@System-Domain vCenter Server privileges by default, and the installation process creates the user admin@System-Domain to manage vCenter Single Sign-On.
- ESXi 5.1 is not integrated with vCenter Single Sign-On, and you cannot create ESXi users with the vSphere Web Client. You must create and manage ESXi users with the vSphere Client. vCenter Server is not aware of users that are local to ESXi, and ESXi is not aware of vCenter Server users. However, you can configure vCenter Single Sign-On to use an Active Directory domain as an identity source, and configure ESXi to point to the same Active Directory domain to obtain user and group information.

DMZ Considerations

VMware recommends that you follow standard DMZ firewall design guidelines in a vCloud environment. However, the following aspects require special consideration. Some vCloud Director operations involve sessions that remain open to management infrastructure, which is protected by the back end firewall, for an extended period.

- Idle session timeouts – Depending on the level of activity within the vCloud environment, some connections, such as sessions to vSphere hosts to retrieve thumbnails by way of the vslad agent and to vCenter Server for inventory, might require adjustment to default TCP timeout policies. This also applies to the Oracle Notification Service (ONS) connections needed 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 verifying that the session is not terminated.
- Logging – Send firewall logs to a centralized syslog server.
- SMTP filtering – Many firewalls filter email connections, restricting ESMTP commands. It might be necessary to disable this capability 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). The firewall must be correctly specified so that it 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. Scale firewalls along with the vCloud to help avoid future downtime.

Port Requirements

Table 22. 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
Virtual machine console to vCenter Server	902, 903	TCP	Outbound
vSphere Web access to ESX/vSphere host	443	TCP	Outbound
Virtual machine console to vSphere host	902	TCP	Outbound
REST API access to vCloud Networking and Security 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 `rpc.statd` (optional)	920	TCP, UDP	Inbound and Outbound
ActiveMQ	61611, 61616	TCP	Inbound and Outbound

Figure 70. vCloud Director Port Requirements

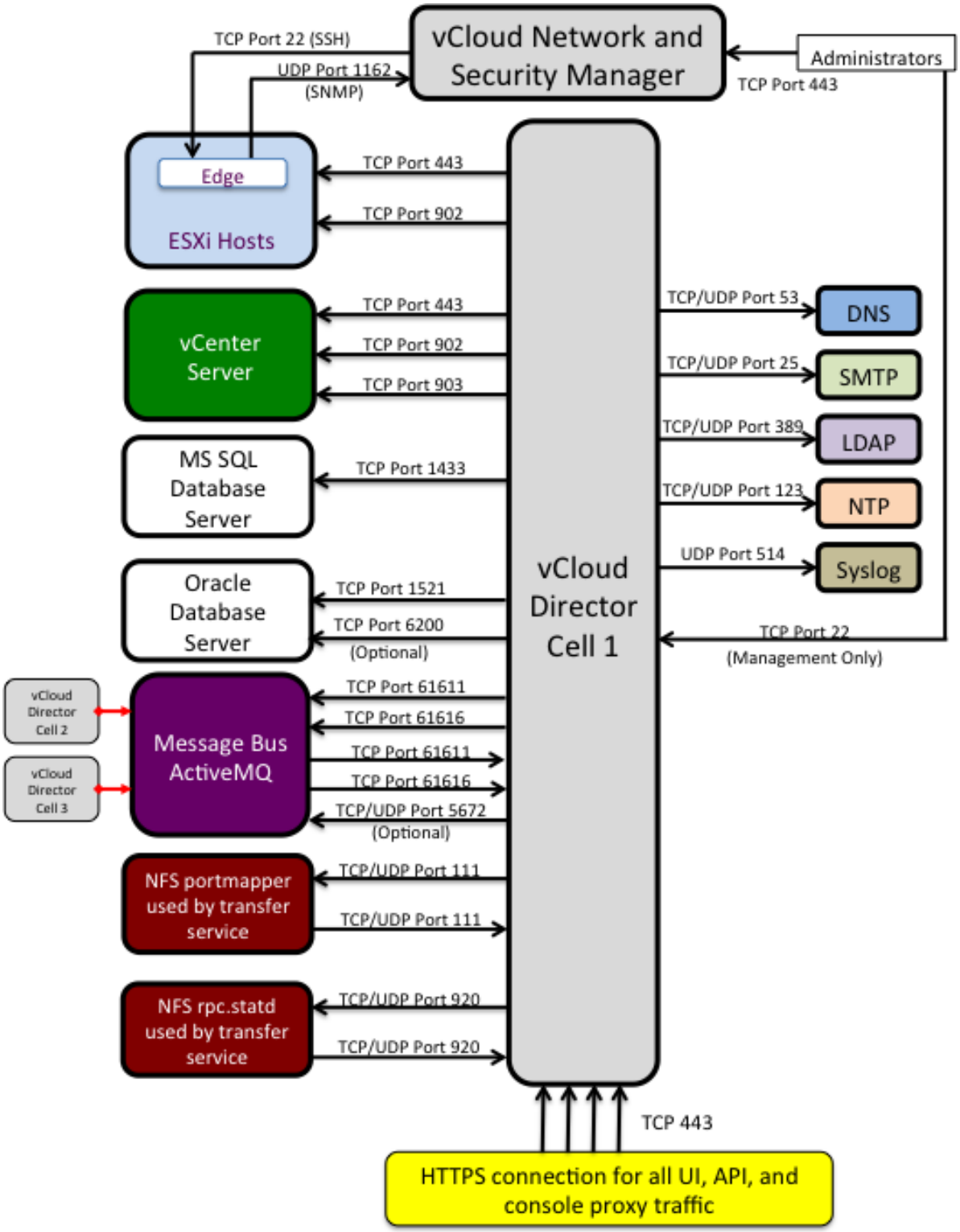


Table 23. vCenter Orchestrator Port Requirements

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

Appendix C: vCloud Suite Disaster Recovery

Disaster recovery for vCloud Director is described as “DR of the Cloud.” It is full-site-based failover and recovery of the entire vCloud infrastructure, including associated vApps.

Because vCloud Director does not currently integrate with vCenter Site Recovery Manager, there is no obvious way to use vCenter Site Recovery Manager to protect a vCloud environment from a disaster scenario by failing the site over to a recovery site.

The VMware vCloud Suite assembles existing products together to facilitate disaster recovery of the vCloud from one site to another. See the following for more information about how this architecture supports disaster recovery:

- *Overview of Disaster Recovery in vCloud Director*
  http://blogs.vmware.com/vcloud/2012/02/overview-of-disaster-recovery-in-vcloud-director.html
- *VMware vCloud Director Infrastructure Resiliency Case Study*
  http://www.vmware.com/files/pdf/techpaper/vcloud-director-infrastructure-resiliency.pdf

The following are the most important considerations:

- Stretched Layer 2 networking (see *Using VXLAN to Simplify vCloud Disaster Recovery*).
- IP address changes to applications.
- Force mounting of LUNs.
- Order of management startup after vCenter Site Recovery Manager failover.
- vApp startup with HA.
- Failover process steps, order of operations.
- Manual versus automated steps.

This vCloud solution as described covers only the main use case of complete site-based failover. It also requires the configuration to handle the failover of an entire provider virtual datacenter. It does not prevent a provider from having unprotected virtual datacenters

Dealing with vCloud disaster recovery has some design implications. Refer to your internal architecture documentation to understand vCloud disaster recovery design implications. If a vCloud design already exists, changes may need to be made to the design to support the current disaster recovery solution.

Using VXLAN to Simplify vCloud Disaster Recovery

When architecting a resilient multisite VMware virtual infrastructure, always consider the use of stretched Layer 2 networks to simplify solution design and the associated recovery process. The following are the main benefits of implementing stretched Layer 2 networks:

- Ability to run workloads in more than one geographical location.
- Migration of virtual machine workloads between geographic locations.
- No need for virtual machine IP address changes when migrating between environments.
- Simplified disaster recovery when not using VMware vCenter Site Recovery Manager.
- When used with vCenter Site Recovery Manager, simplified disaster recovery by not having to change IP addresses.

Even with the simplification afforded by stretched Layer 2 networks, people still tend to avoid them. The reason for this has to do with network instability that is introduced when there is a lot of latency between switching nodes on the network. Stretched Layer 2 networks also increase the failure domain radius by encompassing multiple locations. Most people do not opt for stretched Layer 2 due to the higher cost usually associated with implementation.

Background

It has been demonstrated how vCenter Site Recovery Manager, in conjunction with some complementary custom automation, can be used to offer a vCloud DR solution that enables recovery of a vCloud Director solution at a recovery site.

In cases where stretched Layer 2 networks are present, the recovery of vApps is greatly simplified because vApps can remain connected to the same logically defined virtual networks, regardless of the physical location in which the vApps are running.

The existing vCloud disaster recovery process, while theoretically capable of supporting designs that do not include stretched Layer 2 networks, does not lend itself well to this configuration. The primary issue is the requirement to update the network configuration of all vApps. The complexity associated with the reconfiguration of vApp networking is influenced by a number of factors including:

- Type of networks to which a vApp is connected (organization virtual datacenter, organization external, or vApp).
- Routing configuration of the networks to which the vApp is connected (NAT routed or direct).
- Firewall and/or NAT configuration defined on vCloud Networking and Security Edge devices (NAT routed).
- Quantity of networks to which the vApp is connected.

When connected to an organization virtual datacenter network, there is little or no impact. The vApp can retain its initial configuration, as there are no dependencies upon the physical network. This is not the case for organization external networks.

In the case of vApps connected to an organization external network that is directly connected, the current vCloud disaster recovery process involves disconnection of vApps from the network for the production site and connection to an equivalent network for the recovery site. For this to work, site-specific network configuration parameters such as netmask, gateway, and DNS must be defined. Following reconfiguration, external references to the vApps also need updating. This situation is further complicated when an organization external network has a routed connection. The complication arises from the multiple IP address changes taking place:

1. The vApp is allocated a new IP address from the new organization virtual datacenter network.
2. The associated external network has a different IP address.

The introduction of vApp networks can further complicate the process.

VXLAN makes it possible to the disaster recovery and multi-location implementation of vCloud Director. This is achieved by creating a Layer 2 overlay network without changing the Layer 3 interconnects that are already in place.

This section describes a test scenario in which a vCloud Director implementation based on vCenter Site Recovery Manager fails over without the need to reassign IP addresses to the virtual machines, and describes the scripted changes that must be done to simplify the process.

VXLAN for DR Architecture

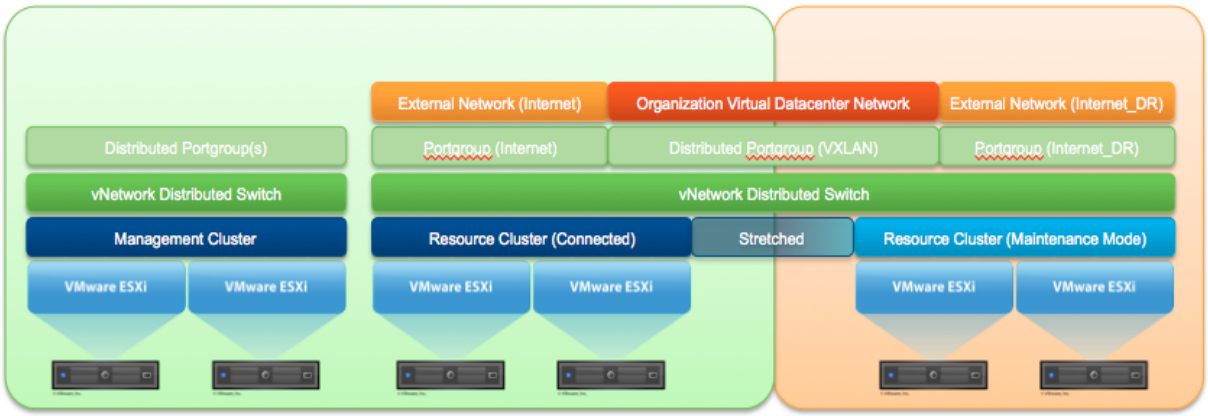
To conduct the required testing, a sample architecture was deployed to simulate the process. In keeping with the reference infrastructure and methodology defined in the *vCloud DR Solution Tech Guide*, the test infrastructure constitutes a cluster that has ESXi members in both the primary and the recovery site. The premise is the workloads run in the primary site where the vSphere hosts are Connected.

At the recovery site, the vSphere hosts are in maintenance mode, but configured in the same cluster and attached to all of the same vSphere Distributed Switches. The solution approach considered in the following sections is developed based on *Overview of Disaster Recovery in vCloud Director*, (http://communities.vmware.com/docs/DOC-18861). All prerequisites for this solution continue to apply.

Logical Infrastructure

To address the complexities of recovering a vApp from the production site to the recovery site in the absence of stretched Layer 2 networking, a mechanism is required to abstract the underlying IP address changes from the recovered vApps. The following diagram provides a logical overview of the deployment infrastructure.

Figure 71. Logical View of Infrastructure



In the resource cluster, all vSphere hosts are connected to a common vSphere Distributed Switch with site-specific port groups defined for the Internet and Internet_DR networks. In vCloud Director, the Internet and Internet_DR port groups are defined as external networks. An organization virtual datacenter network is defined in conjunction with this, and a port group from the VXLAN network pool is deployed.

The vSphere hosts deployed in the production site are connected to a common Layer 3 management network. Similarly, the vSphere hosts deployed in the recovery site are connected to a common Layer 3 management network, albeit in a different Layer 3 than that of the network for the production site. The Internet external networks are the primary networks that will be used for vApp connectivity, and they are also in a different Layer 3 than the Internet network available in the recovery site. These are attached to vCloud Director as two distinct external networks.

vCloud Networking and Security Edge firewall rules, NAT translations, load balancer configurations, and VPN configurations must be duplicated to cover the disparate production and failover address spaces. There are two options for keeping the configurations in sync:

Option 1: Maintain the configuration for both sites at the same time.

- Advantages – Simplifies failover as configurations are already in place.
- Disadvantages:
  - Requires the organization administrator to be diligent in maintaining the configurations.
  - Difficult to troubleshoot if there is a configuration mismatch.
  - Primary interface needs to be removed if hosts on the original Layer 2 primary network need to be reachable.

Option 2: Use the API upon failover to duplicate the primary site configuration to the failover site.

- Advantages:
  - No maintenance after the initial failover address space metadata has been populated.
  - Address mapping can be done and allocated in advance.
- Disadvantages:
  - Must have failover address metadata specified to work.
  - Address size needs match to simplify mapping.
  - Address pool size needs to match to simplify mapping.

Leveraging VXLAN can greatly simplify the deployment of vCloud Director DR solutions in the absence of stretched Layer 2 networking. Furthermore, this type of networking topology is complimentary to the solution defined in the *vCloud DR Solution Tech Guide* and can be implemented with relatively few additions to the existing vCloud DR recovery process.

Following the successful recovery of a vCloud Director management cluster, some additional steps need to be included in the recovery of resource clusters to facilitate the recovery of edge gateway appliances and vApps. See the VXLAN Example in *Implementation Examples*.

VXLAN for DR Design Implications

Recovery hosts must be in maintenance mode so that virtual machines do not end up running in the recovery site and generating traffic between the recovery site and the primary site. The reason for this is that the vCloud Networking and Security Edge device is available in only one site at a time. Having the hosts in maintenance mode also keeps them in sync, with all the changes that happen in the primary set.

If an organization has organization virtual datacenter networks that are directly attached the process is as outlined in the existing vCloud DR recovery process. All of the vApps on that network need to be re-IPed to the correct addressing used in the recovery site Layer 3 network.

However, if the organization is using isolated networks that are VLAN-backed, they need to be recreated on the recovery site using the associated VLAN IDs available in the recovery site and the vApps that are reconnected to the new network. If they were port group-backed, the port groups still exist in the recovery

site, but their definitions need to be revisited to verify that they were valid from a configuration point of view. Ease of recovery is afforded by using VXLAN backed networks outlined in this scenario, be they NAT routed or isolated.

References

vCloud Director 5.1 Documentation Center, http://tpub-review.eng.vmware.com:8080/vcd-20/index.jsp.

Appendix D: vCloud Director Upgrade Considerations

The upgrade process from vCloud Director 1.5.x to 5.1 requires thorough planning. This document focuses on the impact, considerations, and advantages when performing a phased upgrade of vCloud Director. Several upgrade phases are described, with guidance on phase one of the upgrade process. For further guidance on all upgrade phases, see the vCloud Suite 5.1 product documentation.

Background

The upgrade process can be divided into four phases. After completion of a phase, the next phase can be started immediately or can be deferred until later without having a major effect on the vCloud infrastructure. However, new features are not available until all components in each phase are fully upgraded.

This document focuses on the considerations for Phase 1, Upgrade considerations for moving from VMware vCloud Director 1.5 to 5.1.

The following table lists the steps for each phase.

Table 24 Upgrade Phases

Phase	Steps
I	• Upgrade vCloud Director cells from 1.5.x to 5.1. • Upgrade vCloud Networking and Security Manager and deployed vCloud Networking and Security Edge from 5.0 to 5.1. • Upgrade Chargeback from 2.0.1 to 2.5. • (Optional) Upgrade the Oracle/SQL Database versions on database servers.
Deferring next phase affect	vCloud Director 5.1 can manage existing vSphere 5.0, however new features are not available until the components in phases II, III, and IV are upgraded.
II	• Upgrade vCenter Server from 5.0 to 5.1. • Upgrade vCenter Orchestrator 5.0 to 5.1 (equivalent).
Deferring next phase affect	vCenter Server 5.1 and vCenter Orchestrator 5.1 can manage vSphere ESXi 5.0, however new features are not available until the components below are upgraded.
III	Upgrade vSphere hosts from 5.0 to 5.1.
Deferring next phase affect	New features are not available until the components below are upgraded.

Phase	Steps
IV	• Upgrade vSphere Distributed Switches. • Update vCloud Director configuration. • Upgrade vApp hardware levels to hardware version 9. • Upgrade VMware Tools.

Phase I Impact

Phase 1 requires downtime of the following components:

- vCenter Chargeback Manager – Version 2.5 is required for vCloud Director 5.1. vCenter Chargeback Manager 2.5 is backwards compatible with VMware vCloud Director 1.5x. VMware recommends stopping vCenter Chargeback Manager services and upgrading only after vCloud Director is fully upgraded.
- vCloud Networking and Security Manager – This usually requires a specific build to work with vCloud Director. vCloud Director 5.1 supports vCloud Networking and Security Manager 5.1 but does not support vCloud Networking and Security Manager 5.0 when deploying new vCloud Networking and Security Edge devices.
- vCloud Networking and Security Manager – Although vCloud Networking and Security Manager 5.0 will continue to work with older edge devices after vCloud Director is upgraded, new vCloud Networking and Security Edge devices cannot be deployed until vCloud Networking and Security Manager is upgraded from 5.0 to 5.1.
- vCloud Portal and API – The vCloud portal and API are not available during this phase. It is difficult to determine specific downtime, because downtime depends on the number of cells and size of each customer's database. The VMRC is not available to users.
- vCloud Director database – The upgrade changes the schema, so a database backup is important. vCloud Director 5.1 does not support Oracle 10 (all release versions). vCloud Director 5.1 provides support for Microsoft SQL Server 2008 SP3 Standard/Enterprise.
- Rollback – Rollback is complex because the database changes are irreversible.
    - Back up the database after stopping the vCloud Director Services before proceeding.
    - Back up the vCloud Networking and Security Manager database using the UI FTP commands. This is the only method of possibly restoring the vCloud Networking and Security Manager for a redeployment.
    - Perform a backup of the vCenter database at this time, as well as during future steps, as the vCloud Director services get started multiple times throughout the process and can affect the vCenter and vCloud Networking and Security Manager databases.

Upgrade Considerations

The following tables list some of the general things to consider before starting an upgrade of vCloud Director components.

Back up the following components of vCloud before making any changes. VMware recommends that all backups occur at the same time while all vCloud components are shut down. This has a major impact on availability, but maintains data consistency between all components in the event of a rollback.

Table 25. Components to Back Up

Component	Backup Considerations	Resources
vCloud Director database	Create a full backup of the vCloud Director database after all cells are shut down.	Database administrator
vCloud Networking and Security Manager database	Create a full backup of the vCloud Networking and Security Manager database.	vCloud administrator
vCenter Chargeback database	Create a full backup of the vCenter Chargeback Database after all vCenter Chargeback servers are shut down.	Database administrator

VMware strongly recommends performing a full virtual machine backup. If this is not possible, take a snapshot while the virtual machine is powered off or while creating a full clone of the virtual machine.

Table 26. Backup or Snapshot Considerations

Virtual Machine	Backup and/or Snapshot Considerations	Resources
vCloud Director cells	Suspend vCloud Director scheduler, stop the vCloud Director services and shut down the cells. Then take a backup, snapshot, or full clone of the virtual machine.	vCloud administrator
vCloud Networking and Security Manager	After creating a full backup of the vCloud Networking and Security Manager Database, shut down the virtual machine and take a backup, snapshot, or full clone of the virtual machine.	vSphere administrator
Chargeback Managers	Shutdown the virtual machines, then take a backup, snapshot, or full clone of the virtual machine.	vSphere administrator

For non vCloud components, consider the following guidelines.

Table 27. Non-vCloud Considerations

Component	Consideration	Resources
Red Hat patches	Run Red Hat patch updates prior to running the vCloud Director installer. The package dependencies have updates that may be required by vCloud Director 5.1. Do not update kernel or other packages that would bring the system to an unsupported version of RHEL. vCloud Director 5.1 supports the following Red Hat releases: • Red Hat Enterprise Linux 5 (64 bit) Updates 4,5,6, and 8. • Red Hat Enterprise Linux 6 (64 bit) Updates 1 and 2. Update the packages only when necessary as detailed in the *vCloud Director Installation and Configuration Guide* (https://www.vmware.com/support/pubs/vcd_pubs.html).	Linux Administrator
DNS for load balancer VIPs	Consider lowering the TTL (Time to Life) on the DNS for the load balanced VIPs for HTTP and console proxy a day or two prior to upgrading. Lowering these allows clients to update their DNS cache quicker when resolving the portal name. Because the DNS name is directed to temporary maintenance pages, then back to the original pages, a lower TTL prevents the need for users to manually flush their DNS cache for updates. Redirect the DNS prior to upgrading to a custom maintenance page on a completely separate web server outside of the vCloud Director cells. Verify that all users are redirected to the maintenance page before shutting down cells.	DNS Administrator

Phase 1 Process

The following sections cover upgrade preparation and execution.

Pre-Upgrade Considerations

The following process assumes that all vCloud components are turned off so that backups and snapshots are data consistent prior to any upgrade work.

Table 28. Pre-Upgrade Considerations

#	Component	Consideration	Resource
1	DNS	• Lower TTL on DNS. • Redirect to maintenance page.	DNS administrator
2	vCloud Director Cells	• Suspend vCloud Director scheduler and shutdown first cell. • Repeat for subsequent cells. • Back up or take a snapshot of virtual machines.	vCloud administrator vSphere administrator
3	VMware Service Manager	• Back up the Service Manager database. • Shut down Service Manager. • Backup and/or snapshot of virtual machine.	vCloud administrator vSphere administrator
4	vCloud Director database	Back up the vCloud Director database.	Database administrator
5	Chargeback Managers	Shut down vCenter Chargeback servers and perform backup or take a snapshot of virtual machines.	vSphere administrator
6	Chargeback database	Back up the Chargeback database.	Database administrator

Upgrade Considerations

The following guidelines list the key steps to perform during an upgrade.

Table 29. Upgrade Procedure

#	Component	Consideration	Resource
1	vCloud Director Cells	• Suspend vCloud Director scheduler and stop the vCloud Director service. • Repeat for subsequent cells. • Perform Red Hat patches. • Un-mount the NFS transfer share on each cell. • Run the vCloud Director installer but do not start the services. • Run the database upgrade script on the first cell only. Do not repeat this on the other cells. • Run the vCloud Director installer on the other cells. • Reboot each vCloud Director cell one at a time to make sure that all services start up correctly and that the NFS transfer volume is successfully mounted on the first cell before rebooting subsequent cells. • Validate that vCloud Director has started by checking /opt/vmware/vcloud-director/logs/cell.log. • Validate that the portal is working on each cell by connecting directly to the cell's HTTP interface. • Do not redirect the load balancer or allow users back onto the system yet.	vCloud administrator Linux administrator vCloud administrator
2	Update vCloud Director Agent on vSphere hosts	Update the host agent on all connected hosts. Check that connected hosts are still showing **Available** and **Prepared**.	vCloud administrator

#	Component	Consideration	Resource
3	vCloud Director Validation 1	• Validate basic functionality of vCloud Director by deploying a new vApp. • Validate basic functionality by deploying a NAT routed network (either a routed organization network or fence a vApp that is deployed). • Troubleshoot any issues before moving on. • Rollback is possible at this stage by restoring the vCloud Director database and restoring the vCloud Director cells virtual machine backup or deleting the virtual machine snapshot.	vCloud administrator vCloud administrator vSphere administrator or database administrator
4	vCloud Networking and Security Manager Server	Upgrade of the vCloud Networking and Security Manager Server requires the use of an upgrade package. This file is usually named `VMware--Manager-upgrade-bundle-5.1.0-.tar.gz`. Do not deploy a new Service Manager appliance (OVA). Removing the existing Service Manager appliance breaks all connections and management to any deployed vCloud Networking and Security Manager Edge devices, resulting in errors. After a Service Manager is deployed, upgrade it only using a `tar.gz` file (in place upgrade). This preserves the local Service Manager database.	vCloud administrator
5	Edge devices	After vCloud Networking and Security Manager is upgraded, wait at least 15 minutes for it to update information with vCloud Director. Upgrade any organization vCloud Networking and Security Manager devices that are connected to an organization network that is routed. These devices are identified in vCloud Director 5.1 as *edge gateways* and can be upgraded by performing **Re-Deploy**. Upgrade any vApp network vCloud Networking and Security Manager devices connected to an vApp network that is routed. These devices are identified in vCloud Director 5.1 as *edge gateways* and can be upgraded by performing **Re-Deploy**.	vCloud administrator vCloud administrator or organization administrator

#	Component	Consideration	Resource
6	vCloud Director Validation 2	• Validate basic functionality of vCloud Director by deploying a new vApp. • Validate basic functionality by deploying a NAT-routed network (either a routed organization network or fence a vApp that is deployed). • Troubleshoot any issues before proceeding. • Rollback is not possible at this stage as VSEs have been upgraded to the latest compatible versions.	vCloud administrator vCloud administrator
7	Chargeback Managers	The installer requires uninstallation of the previous version. Select **Do not empty the database**.	vCloud administrator

Post-Upgrade Considerations

The following table lists post-upgrade considerations that apply after a successful upgrade of the vCloud environment.

Table 30. Post-Upgrade Considerations

#	Component	Consideration	Resource
1	Local datastores	vCloud Director 5.1 automatically adds local datastores currently presented to ESXi hosts. Disable these datastores from vCloud Director to prevent local datastores from being used by vCloud Director.	vCloud administrator
2	Storage profiles	All datastores that were used by vCloud Director 1.5 are placed into the * (Any) storage profile. VMware recommends that these not be changed at this stage. Storage profiles must be configured in vCenter 5.1 before they can be used in vCloud Director.	vCloud administrator vSphere administrator
3	Upgrade VMRC	vCloud Director 5.1 requires a reinstallation of the VMRC plug-in.	vCloud VMRC users

Upgrade Advantages

The following is a list of advantages that customers have cited as their main reasons for upgrading to vCloud Director 5.1. However, not all new features are available upon completing Phase I.

- User/tenant usability Improvements – The user/tenant usability improvements in vCloud Director 5.1 are targeted at enabling enterprises and service providers to appeal to less tech-savvy vCloud consumers, and expand to users who may not necessarily work in traditional infrastructure management roles.
- Elastic virtual datacenter – Customers can purchase a virtual datacenter of arbitrary size from a robust set of offerings and grow it at will. vCloud Director and vSphere intelligently manage capacity below a robust virtual datacenter abstraction and prevent virtual datacenters from hitting boundaries unless physical capacity is exhausted.
- Multiple classes of capacity – vApps can be deployed as multitier applications with differing infrastructure performance requirements across different tiers (for example, DB on fast storage, web tier on standard storage) within a single application construct.
- New features enabled by upgrade to vCloud Director 5.1:
    - vSphere Storage DRS.
    - Storage profiles.
    - Virtual hardware version 9.
    - Windows 8 guest OS support.
    - Snapshot and revert.
    - Multi-interface vCloud Networking and Security Edge supports ten interfaces that can be configured as either uplinks (external networks) or internal interfaces (facing internal networks).
    - Fast provisioning support for more than eight hosts.
    - Support for Google Chrome browser.

This is not an exhaustive list. Completion of all four upgrade phases is required to enable all new features.