www.vce.com

DESIGNING VMWARE VCLOUD DIRECTOR ON VBLOCK™ SYSTEMS

May 2013

VCE believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED "AS IS." VCE MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OR MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Contents

Introduction

VMware vCloud Director gives enterprises and service providers the ability to build secure public and private clouds that dramatically increase data center efficiency and business agility through Infrastructure as a Service (IaaS). VMware vCloud Director delivers cloud computing for existing data centers by pooling virtual infrastructure resources and delivering them to users as catalog-based services. End users can consume infrastructure without the burden of manual configuration and provisioning.

The Vblock™ System from VCE is the world's most advanced converged infrastructure—one that optimizes infrastructure, lowers costs, secures the environment, simplifies management, speeds deployment, and promotes innovation. The Vblock System is designed as one architecture that spans the entire portfolio, includes best-in-class components, offers a single point of contact from initiation through support, and provides the industry's most robust range of configurations.

Successful deployment of VMware vCloud Director on Vblock Systems requires careful planning and meeting specific design requirements. This paper outlines design practices for deploying VMware vCloud Director on Vblock Systems.

In the design considerations described, Vblock Systems manage the physical and virtualization layers of the infrastructure while VMware vCloud Director provides a layer of abstraction between the underlying infrastructure and the tenants in a multi-tenant, cloud-computing environment. The design philosophy allows you to leverage the standardization, flexibility, and scalability of Vblock Systems while reducing the complexity of the VMware vCenter and vCloud Director environment.

vCloud Director can use every component in a Vblock System as a provider virtual data center (vDC). The provider vDC is set to establish SLAs based on physical characteristics and criteria.

Multiple Vblock System families can be deployed within a single instance of vCloud Director. For instance, there could be more than one Vblock System 200, each dedicated to a department or business unit, while a core Vblock System 720 is dedicated to running mission-critical production workloads in the cloud. At the same time, a Vblock System 100 can be responsible for a dedicated DMZ environment in the cloud.

Understanding the workload to be served in the cloud is critical. vSphere design principals play a key role when architecting for vCloud Director. An architect must gather application requirements for workloads in vCloud Director and then provide a Vblock System to meet those requirements. Requirements might dictate using a single Vblock System or spreading the workload among multiple Vblock System families. vCloud Director can collect multiple types of Vblock Systems under the same cloud offering to make provisioning and administration easier.

Disclaimer

Many of the original design concepts from vCloud Director 1.5 still play a vital role when moving to vCloud Director 5.1 on top of Vblock Systems. vCloud Director 5.1 brings new concepts that can enhance a Vblock Systems functionality. These design concepts are inherited from real world experience and not all design concepts will apply to all situations. These designs are meant to be a guide and you as the architect should find a design that works well for your need.

Document objectives

This document provides guidance for building a VMware vCloud Director solution on Vblock Systems. Both technologies provide flexibility to enable an organization or service provider to successfully deploy a VMware vCloud environment on Vblock Systems. This document is intended to provide guidance to properly architect and manage infrastructure, virtual and physical networking, storage configuration, and scalability of any VMware vCloud Director on Vblock System solution.

As VMware vCloud Director is being increasingly deployed on Vblock Systems, customers and partners might need additional information specific to a combined solution, which requires some additional considerations. This document addresses these considerations in the following target areas:

- Management infrastructure
- Designing for two (or more) vCenter Servers
- Database considerations
- Networking infrastructure
- Provider vDC strategy using Vblock Systems
- VMware vCloud Director and Vblock System scalability

Scope

This document does not address additional products outside of VMware vCloud Director, including but not limited to:

- vCenter Orchestrator
- vCloud Automation Center
- vCloud Director APIs
- VCE Vision™ Intelligent Operations

Audience

This document is intended for individuals with a highly technical background who will be designing, deploying, managing, or selling a VMware Cloud Director on Vblock System solution, including, but not limited to; technical consultants, infrastructure architects, IT managers, implementation engineers, partner engineers, sales engineers, and potentially customer staff.

Feedback

To suggest documentation changes and provide feedback on this paper, send email to docfeedback@vce.com. Include the title of this paper, the name of the topic to which your comment applies, and your feedback.

Management infrastructure

VMware vCloud Director and Vblock Systems include individual management components. The availability of these components dictates the availability of the cloud and infrastructure supporting the cloud. The Vblock System management components are typically supported in the Advanced Management Pod (AMP), which runs components that support the operability and management of the Vblock System. The VMware vCloud Director management components do not have a designated spot to run. The availability of the VMware vCloud components is critical to the logical and virtual sustainability of the cloud. Each layer is equally important and has its own special requirements.

This section addresses the following considerations:

- AMP
- VMware vCloud Director management requirements
- Suggested minimum configuration
- Additional servers
- Blades
- Vblock Systems
- Using existing management structure

Advanced Management Pod

Vblock Systems include an AMP, which is a virtualized infrastructure with varying degrees of availability where the virtual machines necessary to control a Vblock System reside. The virtual machines typically (but not entirely or inclusive of) deployed are:

- VMware vCenter Server (2008 R2)
- VMware vCenter Update Manager (2008 R2)
- VMware Single Sign-On Server (2008 R2)
- Microsoft SQL Server (2008 R2)
- Cisco Nexus 1000V VSM primary (NX-OS)
- Cisco Nexus 1000V VSM standby (NS-OS)
- Array Management and PowerPath Licensing Server (2008 R2)
- EMC Secure Remote Support (ESRS) Policy Manager (2008 R2)
- EMC Secure Remote Support (ESRS) Remote Gateway (2008 R2)
- VCE Vision software

The following table describes the available AMPs:

AMP	Description	Available on
Logical	No additional physical infrastructure required. Management virtual machines reside in the Vblock System and function next to production virtual machines.	Vblock System 100 Vblock System 200 Vblock System 320 Vblock System 720
Mini	A single Cisco UCS C220 M3 rack-mount server running VMware vSphere is connected to two Cisco Catalyst 3560X Ethernet switches. The virtual machines reside on the single server and storage is supplied by internal hard drives.	Vblock System 320 Vblock System 720
High availability (HA)	The highly available AMP consists of two Cisco UCS C220 M3 rack-mount servers connected to two Cisco Catalyst 3560X Ethernet switches. The virtual machines reside on shared storage supplied by an EMC VNXe3150 storage array.	Vblock System 320 Vblock System 720

The mini and HA AMPs are pre-engineered technologies from VCE to host only the virtual machines essential to the management applications of the Vblock System and should be considered out-of-band management. VCE does not suggest running vCloud Director infrastructure virtual machines on the mini and HA AMPs, as the AMP was never intended to accommodate these extra workloads.

Additional reasons for not modifying or using the mini or HA AMPs for vCloud Director infrastructure virtual machines include:

- The amount of RAM necessary to accommodate high availability failovers in the HA AMP is insufficient after adding vCloud Management components.
- Additional Cisco Nexus C220 servers cannot be added because of cabinet space restrictions and Cisco Catalyst 3560X network ports.
- Cisco Nexus C220 AMP servers come with 2x1 GB NICs and therefore cannot handle the network bandwidth required for vCloud Director.
- Additional NICs cannot be added to the Cisco Nexus C220 server because of lack of available network ports on Cisco Catalyst 3560X switches.
- Loss of AMP infrastructure would mean the loss of the cloud management and operations.
- EMC VNXe3150 in the HA AMP ships with 2 TB NL-SAS drives and would be a constraint on I/O-intensive operations.
- HA AMP does not satisfy VMware’s N+1 availability recommendation.

VMware vCloud Director management requirements

The VMware vCloud Director Architecture Toolkit suggests using separate management and compute clusters to provide a scalable cloud infrastructure. Therefore, it is suggested that the management infrastructure components not reside next to virtual machines being provisioned by VMware vCloud Director. This is to overcome false positives in vCloud Director not appropriately estimating consumed resources.

When building a cloud, the following virtual machines migh be necessary to satisfy the management requirements (not all are required):

Virtual machine	Quantity
VMware vCenter Server for management components	1
VMware vCenter Server for vCloud Resources (vCAT suggested)	1
Database servers (SQL/Oracle)	1 (required); can have 2
vCloud network and security manager	1
VMware vCloud Director nodes	2-x (The number of nodes depends on the vCloud environment size and level of redundancy.)
VMware vCenter Chargeback server	1
VMware vCenter Orchestrator server	1
RabbitMQ server	1
Load balancer	1

For a complete cloud solution, these virtual machines might be only a fraction of what is needed to satisfy the requirements. When designing for the cloud solution, keep in mind the amount of resources needed by each virtual machine, as that dictates the necessary management infrastructure size.

The VMware vCloud Architecture Toolkit recommends adding a secondary VMware vCenter Server. One vCenter Server is responsible for the management components, while the secondary vCenter Server is responsible for the compute resources used by vCloud Director. The following are reasons to use two vCenter Servers:

- Create a separation of management domains. vSphere administrators can treat the management server as normal production workloads, while vCloud administrators can administer the compute resources server.
- Abstract vCenter another layer into vCloud Director where it is pooled into provider vDCs.
- Create functional permission boundaries so certain users cannot control vCloud resources.
- Relieve stress on vCenter, as a flood of API calls from vCloud to vCenter can render vCenter useless. A management vCenter not being prone to this keeps management operations active.

Suggested minimum configuration

There is some overlap in the virtual machines described in both the AMP and vCloud Director management stack. The following table provides a suggested minimum configuration needed for a production instance of vCloud Director:

Virtual machine	Quantity	Comments
VMware vCenter Server for management components	1	Provided by existing customer vCenter Server or can be added to bill of materials
VMware vCenter Server for vCloud Resources	1	Provided as standard Vblock System bill of materials
VMware vCenter Single Sign-On server	1	Provided as standard Vblock System bill of materials
Microsoft SQL Server	1	Provided as a standard Vblock System bill of materials; can host databases for VMware vCenter, VMware vCenter Update Manager, and VMware vCloud Director
VMware vCloud Network and Security Manager server	1	
VMware vCloud Director nodes	2	Nodes must be front-ended by a load balancer; can add additional nodes to satisfy redundancy and user requests
Cisco Nexus 1000V Virtual Supervisor Modules	2	Each logical data center instance needs a pair of Virtual Supervisor Modules

The growth of the cloud is modular. As you add more compute resources to satisfy growth demands, you must add more vCloud Director components to satisfy more users and incoming connections. It is critical to choose a management infrastructure that can grow and adapt as your cloud begins to expand and grow. This could mean additional VMware vCenter Servers, Cisco Nexus 1000V VSMs, and VMware vCloud Director Nodes.

This overlap of components makes the VMware vCloud Director management stack a critical part in the cloud's existence. Since many critical components cannot be hosted in the AMP, the choice of AMP type for a Vblock System becomes easier.

- The HA AMP might be deemed unnecessary because many critical components will need to be migrated out of the AMP. The HA AMP is overkill for many typical designs.
- There will still be some components that are non-critical to the cloud but necessary for Vblock Systems management. A mini-AMP might be a more suitable choice to run those non-critical workloads, such as ESRS.
- The Logical AMP option migrates all critical and non-critical workloads to an infrastructure management solution, as defined in the following sections.

VMware vCloud Director and Vblock Systems critical workloads need to be on a highly resilient physical infrastructure that should also have performance characteristics capable of handling a multitude of incoming requests, low-latency bandwidth for transferring vApps, and IOPS for transactional database events. The IOPS requirements for a minimal vCloud Director deployment are undefined, but any EMC array can handle the workload. Defining IOPS characteristics becomes important when additional management components, such as VMware vCenter Operations Manager, are required. The physical management infrastructure should scale as easily as the vCloud Director footprint.

Using a separate management infrastructure

Additional servers

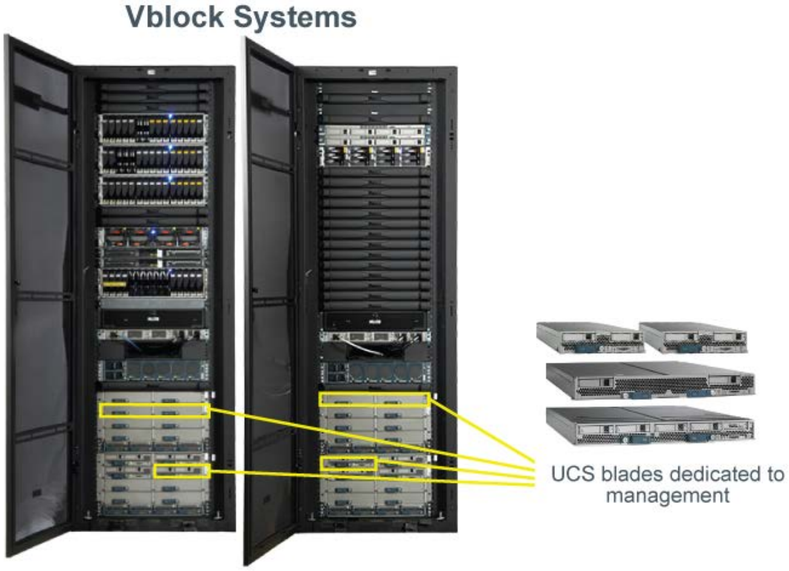
Using additional Cisco C-series servers by interconnecting into the Vblock System fabric and using the Vblock System array for storage is not a supported management infrastructure. Adding servers not supported by VCE into the networking fabric of the Vblock System brings complicated architecture and design considerations as every port in the Vblock System is reserved for certain functions. This brings uncertainty relating to performance and scalability. Additional servers also bring complexity to the delivery and integration process.

Using Vblock System server blades

A simple way to accomplish a separate management infrastructure is to use blades within the Vblock System. Keep the following considerations in mind:

- VMware recommends no less than three hosts in a single cluster to account for N+1 reliability.
- For many smaller clouds where budget is a constraint, two hosts might be a suitable option for initial deployment, with expansion later.

Design requirements dictate how many blades are needed and how big the failure domain needs to be. To keep costs to a minimum, the management infrastructure blades can be blade packs with the minimum amount of RAM. Software components of EMC PowerPath/VE and Cisco Nexus 1000V are still requirements for blades in Vblock Systems.



Using blades creates a highly resilient management infrastructure. Blades dedicated to management can be separated into different Cisco UCS chassis and even different cabinets to overcome rare chassis or cabinet failures. The Cisco UCS system is highly tuned for performance in virtual environments with Vblock Systems. Blades use 10GbE fibre channel over Ethernet (FCoE) connections for Ethernet and storage so performance capabilities are optimal.

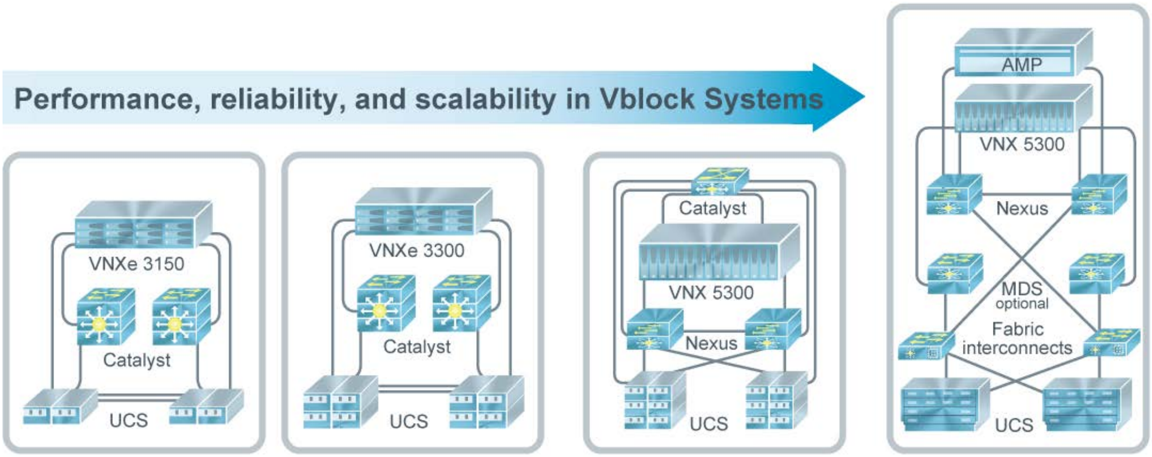
As the cloud grows, more blades can be added to the Vblock System and to the management cluster. The scalability limit is determined by the maximum number of blades that can be added to a configured Vblock System. This route is a simple quick-start to launching your vCloud Director journey. There is no additional separate infrastructure to purchase or integrate, and it is a clean setup and delivery. The constraints of this method make troubleshooting more difficult as all infrastructure is shared. Any outage or traffic spike that affects the storage array or network also affects the management footprint.

Mid-market Vblock™ Systems to manage larger Vblock Systems

Many enterprises and service providers on their cloud journey continue to see their management footprint expand. As management infrastructure continues to grow, a dedicated smaller Vblock System might be a suitable alternative. The smaller version of Vblock Systems can host everything relating to cloud infrastructure management, as well as services like Active Directory, DNS, DHCP, and so forth. It can also be the home of all operational and monitoring based services.

Having a smaller Vblock System managing your cloud creates a much more reliable and highly available footprint because the entire infrastructure is dedicated to management. Nothing is shared between the management and core workloads. The servers, network fabric, and storage are not susceptible to any I/O bursts that happen inside the cloud. Security zones are also not compromised between individuals within the IT or service provider organization. Virtual applications accessible from an external network are susceptible to different types of denial-of-service attacks. Putting the management footprint outside this zone will not compromise the managerial elements.

Owning a smaller Vblock System is a viable alternative because Vblock Systems bring operational and technical benefits. The arrival of the mid-market Vblock Systems makes acquiring a smaller dedicated infrastructure easier to accomplish. Every Vblock System has a different performance and scalability profile and must be taken into account when architecting this solution.

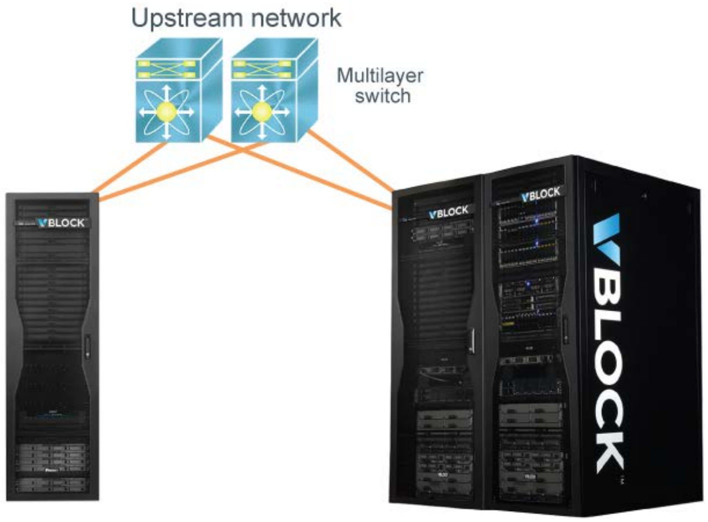


Vblock 100 BX	Vblock 100 DX	Vblock 200	Vblock 320 with EMC VNX 5300
Max. 4 C220 M2 UCS blades with 48 GB RAM	Max. 8 C220 M2 UCS blades with 96 GB RAM	Max. 12 C220 M2 UCS blades with 64-256 GB RAM	Max. 16 UCS B series blades with 64-256 GB RAM
VNXe 3150	VNXe 3300	VNX 5300	VNX 5300
			EMC PowerPath/VE
1 GbE ——	1 GbE ——	10 GbE ——	10 GbE ——

The Vblock System 100 is a solution that works well for those not concerned with extreme performance and scalability. The Vblock 100 uses a networking fabric of 1 GbE to the Cisco UCS C220 servers, which might cause a bottleneck for thousands of connections into vCloud Director. The storage of the Vblock System 100 model BX uses an EMC VNXe3150, while the Vblock System 100 model DX uses an EMC VNXe3300. The choice of which array to use depends on the I/O characteristics of your management infrastructure. The Vblock 100 has a set standard of configurable drive types and amount of usable storage. The performance drive pack uses 15k RPM SAS disk drives, which is a suitable choice for production workloads. Scalability in a Vblock 100 BX is limited to 4x Cisco UCS C220 servers, each equipped with 48 GB RAM. The largest scale of the Vblock 100 DX is 8x Cisco UCS C220 servers, each equipped with 96 GB RAM

The Vblock System 200 combines performance, reliability, and scalability. The Vblock 200 scales to 12x Cisco UCS C220 servers, equipped with anywhere from 64 GB to 256 GB RAM in each server. The servers are also equipped with 10 GbE connections to allow greater network performance. As your cloud expands, the transfer of images through catalogs or to different clusters relies heavily on network throughput. In addition, this configuration satisfies thousands of user connections into VMware vCloud Director. Storage is based on an EMC VNX5300, which is an enterprise-grade solution that is highly resilient and offers plenty of backend I/O to satisfy any management and operational workload.

Regarding management infrastructure, the Vblock System 320 with EMC VNX5300 combines performance, reliability, and scale. The Vblock 320 with EMC VNX5300 uses up to 16x Cisco UCS B-Series Blades for superior scale. The addition of EMC PowerPath/VE improves performance by gaining more I/O paths for greater VMware vSphere I/O throughput. Storage is supplied from an EMC VNX5300 to add enterprise-class functionality.



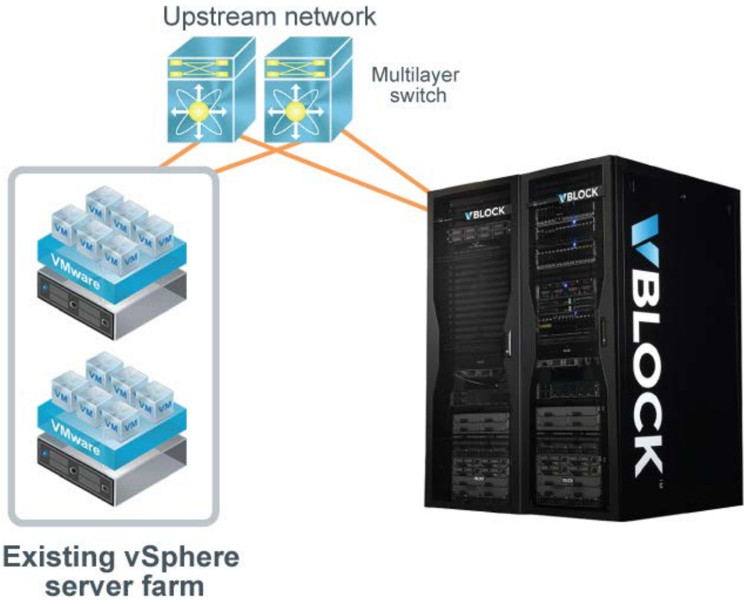
© 2013 VCE Company, LLC. All Rights Reserved.

Using existing management infrastructure

Many enterprises and service providers adopting VMware vCloud Director might already have a VMware vSphere Server farm that hosts management applications such as Active Directory, DNS, vCenter, SQL, and so forth. You could choose to use an existing vSphere Server farm to provide resources for the vCloud Director management components. Critical workloads of the Vblock System (VMware vCenter, SQL, and Cisco Nexus 1000V VSM) will require migration from the AMP. Therefore, the existing vSphere farm must:

- Provide at least three to four hosts dedicated to management to satisfy N+1 or N+2 redundancy, following the architectural guidelines presented earlier in this document
- Be fully redundant and have high bandwidth connections to the Vblock System it manages

One possible approach is to use a mini-AMP to hold all pieces for delivery and have a place to host non-critical management components, and then migrate the critical management components to the existing vSphere farm after delivery. This approach has many variables and is not recommended because it will cause disruption to the standard Vblock System delivery. It also complicates the process of data centers moving toward standardization.



Designing for two (or more) VMware vCenter Servers

Based on the architecture aligning with the vCloud Director Architecture Toolkit, we want to ensure an understanding of where each VMware vCenter instance is hosted, and which vCenter Server is managing each ESXi host and the vSphere virtual machines.

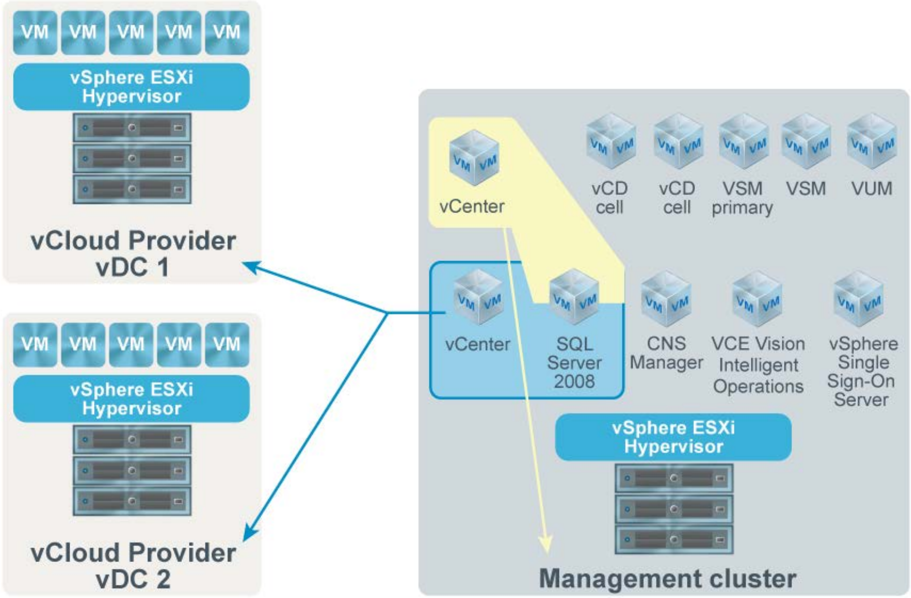
The vCloud Architecture Toolkit recommends using a secondary VMware vCenter Server for a production instance of vCloud Director. The two vCenter Servers are aligned as follows:

- Management infrastructure VMware vCenter Server
- Cloud resources VMware vCenter Server(s) (must adhere to vCloud and vCenter maximum configurations)

This section contains three examples, for which all previously mentioned management infrastructures can be used to satisfy requirements. The management infrastructure in these example designs can be blades inside the Vblock System, a dedicated Vblock System for management, or an existing VMware vSphere infrastructure.

Example 1: Using a logical AMP to move workloads

This example uses the logical AMP to move all AMP workloads into the management infrastructure.



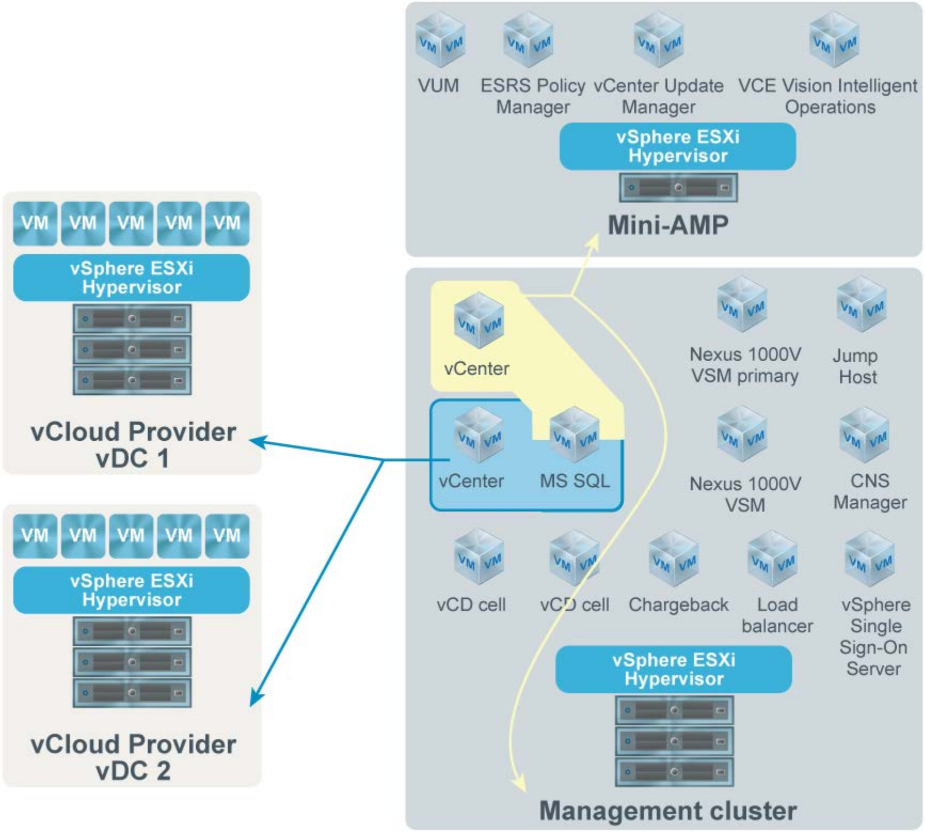
The first instance of VMware vCenter Server is the management infrastructure vCenter, which is responsible for managing the vSphere hosts in the management infrastructure. These hosts exist inside a single logical data center object. The hosts can be grouped as a single cluster or broken down into multiple clusters as N+1 availability is met.

The secondary vCenter, Cloud Resource vCenter, is responsible for managing the vCloud consumable resource hosts.

This design approach saves on costs traditionally spent on an AMP infrastructure. However, any loss of the management infrastructure renders the vCloud Director instance unusable and is much harder to troubleshoot. Therefore, having a highly resilient management infrastructure is strongly recommended.

Example 2: Using a mini-AMP for non-critical workloads

In this example, non-critical workloads remain in the mini-AMP, instead of moving all workloads to a management infrastructure.



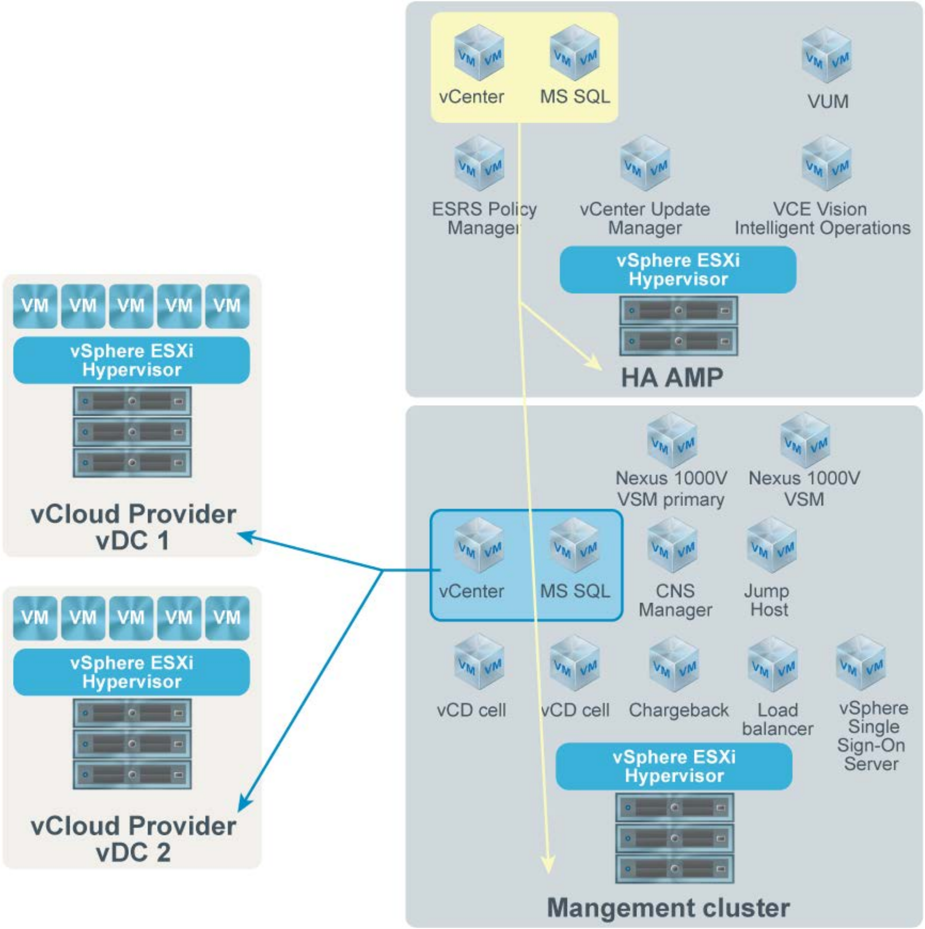
The management infrastructure vCenter is responsible for managing the vSphere hosts in the management infrastructure. There are two defined logical data center objects: the AMP data center and the management infrastructure data center. The mini-AMP server acts as a stand-alone host to the AMP data center object. The vSphere hosts dedicated for management are added to the management infrastructure data center.

The Cloud Resource vCenter is responsible for managing the vCloud consumable resource hosts.

The constraint of this approach is that any loss of the management infrastructure renders the vCloud Director instance unusable and much harder to troubleshoot. Therefore, having a highly resilient management infrastructure is strongly recommended.

Example 3: Using a mini-AMP or HA AMP for a distributed failure domain

This example uses the mini-AMP or HA AMP, as well as a management infrastructure. Both the mini-AMP and HA AMP contain standard Vblock System components, except the Cisco Nexus 1000V VSMs, which are moved to the management infrastructure. There are only two vCenter Servers in use in this example.



The management infrastructure vCenter Server resides in the mini-AMP or HA AMP. This vCenter Server controls the vSphere hosts responsible for the management infrastructure. It also controls the mini-AMP or HA-AMP cluster. There are two defined logical data center objects: AMP data center and management infrastructure data center. The mini-AMP or HA AMP is added as a cluster to the AMP data center object. The vSphere hosts dedicated for management are added as a cluster to the management infrastructure data center.

The secondary vCenter Server resides in the management infrastructure and controls the vSphere hosts responsible for consumable cloud resources. Since this vCenter instance controls the Vblock System, the Cisco Nexus 1000V VSMs should reside in this infrastructure. In addition, a separate Microsoft SQL Server is recommended to keep high-bandwidth connections between the servers.

Any loss of the management infrastructure could render your vCloud Director instance unusable; however, the vCenter Server managing that infrastructure will be accessible so troubleshooting is easier.

Although there might be other designs that satisfy all requirements, the recommended approach is to separate the two environments completely. Each vCenter Server, data center object, or cluster can have separate distinct access roles and permissions. Each recommended design provides out-of-cluster management, which is a generally accepted best practice with vSphere architecture.

VMware vCenter protection

The existence of VMware vCenter is critical in a VMware vCloud Director implementation because vCenter is a secondary layer in the vCloud Director stack. The vCloud Director servers are a layer higher in the management stack and control the vCenter Servers. To further protect any vCloud Resource vCenter Server instance hosted inside the management infrastructure, you can use vCenter Heartbeat. vCenter Heartbeat is not a required component of the vCloud Director on Vblock System design.

Database considerations

VMware vCloud Director is a scale-out application that uses a single backend database. Oracle and SQL are the two supported databases for vCloud Director. There are many design approaches as to where your vCloud Director database can reside.

Database placement after moving critical workloads to the management infrastructure

As suggested earlier in this document, the AMP is not a suitable place to run vCloud Director components. One suggested approach is to move all critical AMP workloads to a dedicated management infrastructure. When Microsoft SQL Server is delivered with a Vblock System and moved to a dedicated management infrastructure, it can support multiple databases. These include multiple VMware vCenter Server, VMware vCenter Server Update Manager, and vCloud Director databases.

Database placement in a distributed failure domain

It is a best practice to keep VMware vCenter Servers and their databases close together through high-bandwidth links to mitigate any latency. In this case, the management infrastructure VMware vCenter Server and Microsoft SQL Server remain in the AMP, but a secondary Microsoft SQL Server is needed for the Cloud Resource vCenter residing in the management infrastructure. In this scenario, the VMware vCloud Director database should use the SQL Server instance residing in the management infrastructure because it is more tolerant to failures.

Using an existing SQL Server or SQL Cluster

Every Vblock System ships with a Microsoft SQL Server to be the backend database for the Vblock System it is managing. If you are adding multiple Vblock Systems to a supported VCE management infrastructure (excluding non-VCE VMware vSphere farms), then you can use an existing SQL Server.

Many enterprises today already have SQL Clusters in use to provide software-defined availability for many applications. The Cloud Resource vCenter database cannot be part of this cluster because it is not a part of the standard Vblock System delivery process. However, the VMware vCloud Director database can be part of this cluster as long as there is a high-bandwidth connection between the two.

Using a dedicated database

Many environments might feel the necessity for VMware vCloud Director to have its own database in order to create a security boundary. There are certain modifications that must be done on a vCloud database prior to installation. Providing the vCloud administrators their own SQL Server creates a security zone where DBAs need not worry about vCloud Director.

In addition, it might be suitable for vCloud Director and all vCloud Resource vCenters to share the same database because the vCloud administrator is responsible for everything that belongs to vCloud Director. This is not a requirement of vCloud Director design, but it is a design consideration.

Networking infrastructure

VMware vCloud Director provides networking services that can be provisioned on demand and consumed by tenants in the cloud. The networks provisioned by vCloud Director are Layer-2 isolated entities that are created through network pools and used to create organization and vApp networks.

There are three types of networks available to VMware vCloud Director:

Network type	Description
External	Considered a shared services network that provides distinct network resources to the cloud.
Organization	Accessible networks to vApps within a particular organization and controlled by the cloud administrator. Can be directly connected to external networks, NAT-routed to external networks, or isolated.
vApp	Accessible networks to individual vApps and are controlled by the individual vApp owners. Can be directly connected to organization networks, NAT-routed to organization networks, or isolated.

VMware vCloud Director creates these networks using the following pool types:

- VXLAN-backed pools
- Port group-backed pools
- VLAN-backed pools
- vCloud Director Network Isolation (vCDNI)–backed pools

This discussion focuses on VXLAN, which is rapidly becoming an adopted standard. It allows greater flexibility than the other types of network pools.

Cisco Nexus 1000V Switch

The Cisco Nexus 1000V Switch is an integral part of the Vblock System, allowing for advanced feature sets of the Cisco NX-OS to live in the virtual space. Cisco NX-OS gives network administrators the ability to see deeper into network traffic and inspect traffic that traverses the network. The Cisco Nexus 1000V Switch is compatible with VMware vCloud Director and extends the benefits of Cisco NX-OS features, feature consistency, and Cisco's non-disruptive operational model to enterprise private clouds and service provider-hosted public clouds managed by VMware vCloud Director. The Cisco Nexus 1000V Switch provides all necessary features and functions for a production VMware vCloud Director deployment.

Configuring the VMware vSphere ESXi host

Every host being used for cloud resources will be attached to the Cisco Nexus 1000V Switch. The VMware vSphere ESXi hosts are configured by default with 2x 10GbE network adapters. No additional network adapters are necessary for VMware vCloud Director implementation.

The benefit of using the Cisco Nexus 1000V over the VMware vSphere Distributed Virtual Switch is that Vblock Systems are delivered in a way that uses both uplinks for VXLAN traffic while still maintaining network consistency for everything else.

To use multiple physical uplinks on the vSphere Distributed Switch with VXLAN, Link Aggregation Control Protocol (LACP) must be configured on both the host and upstream switches. This can lead to network confusion and mishandling of configurations. In the past, VMware vSphere was not optimized for LACP; therefore, LACP was not a best practice when using the vSphere Distributed Switch. A previously recommended practice was to set a port group to “Route based on Physical NIC Load” to let the ESXi host choose the preferred route. vSphere now natively supports LACP for the vSphere Distributed Switch, which changes the previous best practice. Many vSphere Distributed Switch implementations use Failover mode, which allows only one physical uplink to be used for VXLAN traffic. Vblock Systems remove this complexity and allow multiple uplinks to be used as static Etherchannel without any additional configuration.

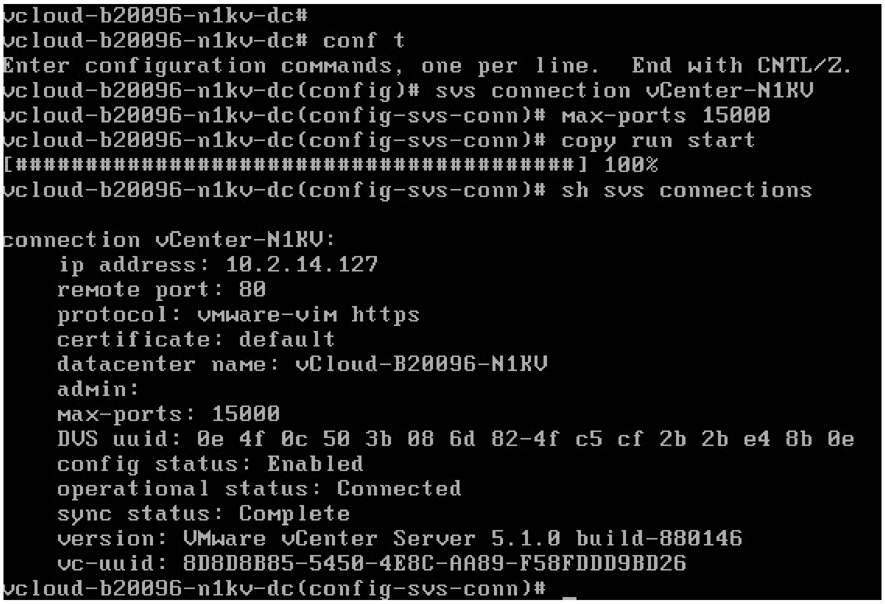
Extending network ports

After connecting the Cisco Nexus 1000V Switch to VMware vCenter, consider port allocation.

- Estimate the approximate port usage by every host. There is a maximum of 60,000 network switch ports per VMware vCenter with the Cisco Nexus 1000V Switch.
- Determine the number of Cisco Nexus 1000V switches to use to satisfy the overall requirements of the VMware vCloud Director implementation, taking into account future expansion. By default, the Cisco Nexus 1000V Switch has 8192 switch ports.

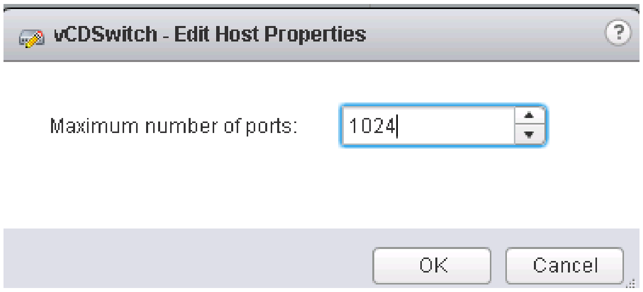
For example, if requiring four Cisco Nexus 1000V switches, use the following commands to create additional ports:

```
N1kv (config)# svs connection MYvCenter

N1kv (config-svs-conn)# max-ports 15000
```



```
vcloud-b20096-n1kv-dc#
vcloud-b20096-n1kv-dc# conf t
Enter configuration commands, one per line.  End with CNTL/Z.
vcloud-b20096-n1kv-dc(config)# svs connection vCenter-N1KV
vcloud-b20096-n1kv-dc(config-svs-conn)# max-ports 15000
vcloud-b20096-n1kv-dc(config-svs-conn)# copy run start
[########################################] 100%
vcloud-b20096-n1kv-dc(config-svs-conn)# sh svs connections

connection vCenter-N1KV:
    ip address: 10.2.14.127
    remote port: 80
    protocol: vmware-vim https
    certificate: default
    datacenter name: vCloud-B20096-N1KV
    admin:
    max-ports: 15000
    DVS uuid: 0e 4f 0c 50 3b 08 6d 82-4f c5 cf 2b 2b e4 8b 0e
    config status: Enabled
    operational status: Connected
    sync status: Complete
    version: VMware vCenter Server 5.1.0 build-880146
    vc-uuid: 8D8D8B85-5450-4E8C-AA89-F58FDDD9BD26
vcloud-b20096-n1kv-dc(config-svs-conn)# _
```

Next, go to each host and change the number of ports available. A port includes VMkernel and virtual machine usage. If you think you are going to have more than the default of 512 ports in use on each host, change it to an appropriate number. Changing the number of ports per host requires a host reboot.



vCDSwitch - Edit Host Properties

Maximum number of ports: 1024

External networks must be pre-provisioned and consumed by VMware vCloud Director. Deploying Cisco Nexus 1000V port-profiles using the default settings can create misconfigurations for deployments done on a larger scale. By default, a Cisco Nexus 1000V port-profile is deployed with 32 ports. When an attempt is made to use a 33rd port, the operation fails until another port is released. Do not change the **max-ports** variable to something higher because you run the risk of maxing out available ports.

When deploying external networks for VMware vCloud Director using the Cisco Nexus 1000V Switch, there are two options available to maximize the available ports for use:

Option	Result
Change the default port-profile to use port-profile default **port-binding static auto.**	Every port-profile is not dependent on the max-ports configuration but everything can auto-expand.
Change the port-profiles used for external networks to have **port-binding static auto.**	Port-profiles used for external networks will use auto-expand.

Using VXLAN for network pools

VXLAN was co-engineered by Cisco and VMware and has since spread to many partners. VXLAN functionality is easily enabled on the Cisco Nexus 1000V Switch and is completely supported by VCE. VXLAN allows a single Layer-3 multicast-enabled VLAN to spawn 16,000,000+ isolated networks. This far exceeds any previous technology.

To enable VXLAN in the Cisco Nexus 1000V Switch, type the following commands:

```
N1kv (config)# feature segmentation

N1kv (config)# feature network-segmentation-manager
```

For end-to-end communication of VXLAN, enable an MTU size of 1600. This means the Cisco Nexus 1000V and Cisco Nexus 55xx upstream switches must be able to transmit jumbo frame data. VXLAN Tunneling End Points (VTEP) must be created on every participating VMware vSphere host and be configured with an MTU size of 1600. The VTEP on each vSphere host are simple VMkernel ports that are created manually or through a script and are enabled during vCloud Networking and Security Manager setup.

You must also enable multicast on all the upstream switches and have IGMP Snooping enabled. This is enabled by default on all Vblock Systems. The other necessity is an IGMP Querier Address on an upstream Layer-3 router that is outside the Vblock System. If your VXLAN traffic is going to span multiple subnets, enable PIM routing on the Layer 3 router as well.

Provider vDC strategy

VMware vCloud Director provides resources, called provider virtual data centers (vDC), for organizations to use. There are many different provider vDC strategies. The cloud administrator is responsible for providing an appropriate SLA relating to provider vDCs.

A provider vDC can be one or more clusters of servers with associated data stores. A standard best practice is to associate a cluster of servers and associated data stores as a tier of service and to not use resource pools or share server clusters among different tiers. VMware vSphere storage profiles also play a major role in storage decisions.

VMware vCloud Director 5.1 uses compute resources more efficiently than previous versions. Previously, a cluster of servers and associated data stores (for example, all fibre channel (FC) or SATA data stores) were coupled as a provider vDC. Since a design principle was to never mix different data stores in the same cluster, the result was multiple provider vDC offerings. With the addition of storage profiles in vCloud Director 5.1, all servers are pooled together. Data stores can be associated with storage profiles, which can then be granted on a per-organization vDC basis.

Note: This document uses three approaches for examples: Gold, Silver, and Bronze.

Provider vDC strategies

This section describes three provider vDC strategies:

- Disk types
- RAID types
- Storage software technologies

Disk types

The simplest way to tie an SLA to a storage profile in a provider vDC is based on the disk types in the Vblock System. This should be relatively easy because differences in performance between EFD/SSD, FC/SAS, and SATA drives are discernible. For example, assign an appropriate SLA based on performance characteristics as follows:

- Gold aligns with EFD/SSD
- Silver aligns with FC/SAS
- Bronze aligns with SATA

The downside of this method is the inability to appropriately estimate the amount of each disk type needed and the associated wasted cost. To prevent over- or under-purchasing for a particular tier, it is imperative to gather tenant requirements. The wasted costs are risky if an abundance of a particular drive tier is purchased and no tenants want to pay for that sort of premium.

RAID types

A second provider vDC form relates to disks, but instead builds upon multiple tiers by using multiple types of RAID groups in the Vblock System. This is difficult to standardize, as there are many types of RAID offerings. There is also the chance of wasting money on unused disks. Different applications may warrant the need for RAID 5 versus RAID 6 versus RAID1+0 for performance characteristics.

For example, set the following:

- Gold tier to SAS/FC Raid 1+0
- Silver tier to SAS/FC in RAID5
- BronzePlus tier to SATA in RAID5
- Bronze tier to SATA in RAID6

Note: The example does not include an EFD/SSD tier; however, it could be used.

Keep the following design consideration in mind: make everything simple to the tenant of the cloud. The goal is to keep costs in mind and find the optimum spot for return on investment. Using RAID types as a differentiating factor might not be the most efficient method, as the applications hosted within vCloud Director (at this point in time) are probably not critical enough. Choosing a single standard RAID type might be a better choice to ensure that resources are not over- or under-allocated.

Storage software technologies

The third type of provider vDC strategy still uses media types as an approach, but focuses on EMC storage technology. There are two types of storage technologies that can be packaged with a Vblock System:

- EMC FAST Cache
- EMC Fully Automated Storage Tiering (FAST)

EMC FAST Cache

EMC FAST Cache is a mandatory component on all Vblock 200, Vblock 320, and Vblock 720 Systems. It contains a minimum of 100GB as Enterprise Flash Drives (EFD). In its simplest form, FAST Cache is an extension of read/write through cache used by the array to move to the EFT tier hot blocks of data that have been accessed at least three times to an EFD tier. FAST Cache is an array-wide technology that cannot be turned off for specific data stores. When using a disk-based SLA approach, FAST Cache can become an uncertain variable because tenants in a Bronze tier could be getting EFD-type responses even though they only paid for SATA. At the same time, tenants in Silver or Gold might be getting only what they paid for depending on the workload across the array and the hot blocks that were able to move up. A way to circumvent this cost is to break up the cost of FAST Cache among all tiers and have a standard fixed cost as an added bonus, even though the tenants might or might not be able to leverage it. EMC FAST Cache is still very much recommended to help general performance of the array.

EMC FAST

EMC FAST can determine a storage profile in a provider vDC strategy because different types of disks can be offered on a single data store basis. FAST allows different tiers of disks to be aggregated into a single data store and an algorithm determines where a block of data is stored. SSD, FC, and SATA can all be placed into a single data store and hot blocks are moved to a higher tier of disk, while other unused blocks can be moved to a lower tier. If those lower tier blocks start seeing action, they can potentially move up a tier or two based on how often the algorithm runs.

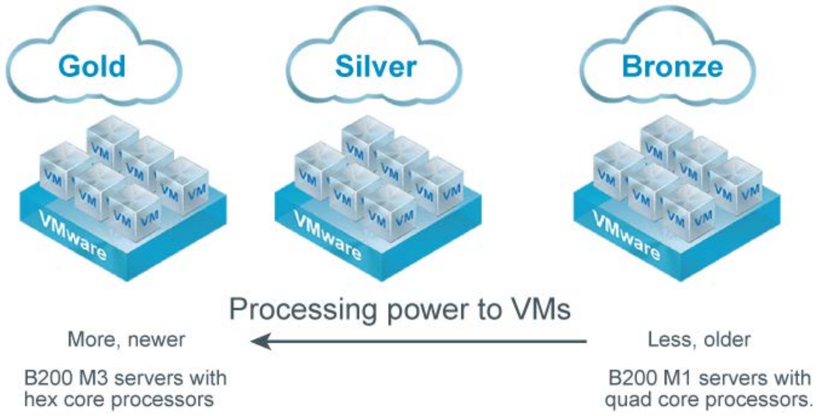
FAST allows cloud administrators to offer multiple disk-based SLA offerings. For example:

- Gold tier equals 30% EFD and 70% FC, giving Gold tenants more room to "burst" into EFD while not paying a premium for EFD drives in the short term.
- Silver tier could be tied to 5% EFD, 70% FC, and 25% SATA, giving tenants an offering that allows little bursting capability but provides good performance when needed.
- BronzePlus tier could be 25% FC and 75% SATA, allowing tenants to burst into FC-type performance while still keeping costs minimal.
- Bronze tier could be 100% SATA and no FAST to offer a predictable performance tier.

This strategy gives the cloud provider more options for the level of service to offer to tenants while also saving money on expensive EFD drives. The only downside to a FAST offering is that you cannot guarantee tenants a predictable IO pattern or performance guarantee. VMware vCloud Director sees all data stores assigned to a storage profile equally in a provider vDC. If multiple tenants use the same FAST data store, they compete for those higher grade tiers based on their workload.

Server hardware

The tiered approach can exist with servers as well. Perhaps a previous Vblock System purchase used Cisco UCS B200 M1 servers with quad-core processors. Another Vblock System purchase has Cisco UCS B200 M3 servers with hex-core processors. Both clusters still rely on the same types of back-end data stores, but the differentiating factor is the processing power given to the virtual machines.



Block versus file-based storage

Do not base an SLA on fibre channel (FC) versus network file system (NFS) storage. Both solutions achieve what is needed. Instead, think about how SLA are tied to connections on 1 GB versus 10 GB NFS and 4 GB versus 8 GB FC. For example, a mixed environment could have:

- 1 GB IP = Bronze
- 4 GB FC = Silver
- 8 GB FC = Gold or 10 GB NFS = Gold

These come into play if multiple Vblock System families and generations are being consumed by VMware vCloud Director.

It is best to remain neutral about how an SLA is tied to a type of network medium. In addition to speed, take into account reliability. What type of switch or fabric switch is in the middle? Are the fabric switches redundant, and if the loss of a switch occurs, what is the impact to the throughput? Block and file data stores can both be available to the cloud administrator for presenting VMware vSphere data stores.

Vblock System approach

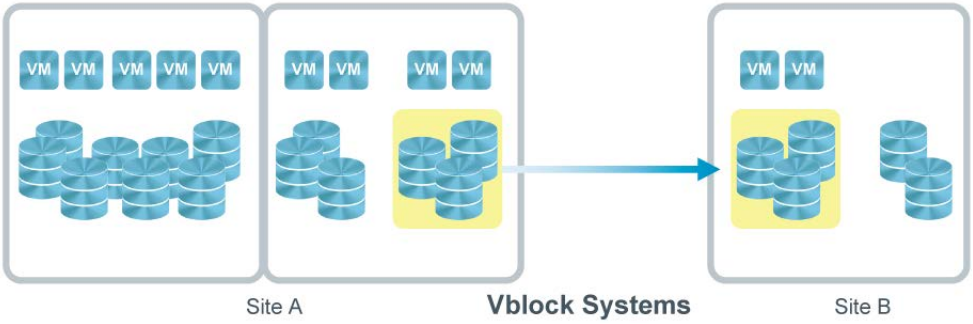
Basing a provider vDC strategy on disk is a good choice for use in a single Vblock System because it can be easily managed. VMware vCloud Director gives the cloud provider freedom and control over the provider vDC offering. Many companies have older or somewhat new VMware farms but are looking for a refresh or to expand. We can now use VMware vCloud Director provider vDCs in a pod approach instead of taking a granular disk approach. As new Vblock Systems are added to VMware vCloud Director, they can be associated with a higher tier such as Gold. Older vSphere farms with out-of-date equipment can be given a lower SLA on an as-is basis.

Replication and higher availability

Replication options include the following:

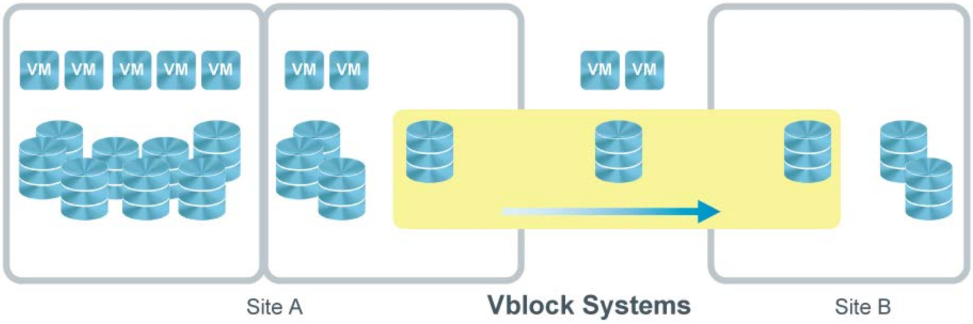
- Replicate one Vblock System in data center A to another Vblock System in data center B.
- Select certain data stores to be replicated from one site to another. From here, you can still provide the same levels of tiering based on disks and servers in each Vblock System, except that the data stores in one Vblock System can have a "Plus" appended to them because the workloads are now protected by replication. EMC RecoverPoint offers this technology.

Workload replication with EMC RecoverPoint



Many companies are also looking at implementing stretched clusters for mission-critical applications requiring little to no downtime. As VMware vCloud Director becomes an increasingly trusted platform, more mission-critical workloads are placed there. The provider vDC might need to be modified with technologies that satisfy the needs of these workloads. Technologies like EMC VPLEX along with Cisco Nexus 7000 for Overlay Transport Virtualization (OTV) and Locator/ID Separation Protocol (LISP) can create stretched-cluster scenarios to provide that level of availability.

Stretched cluster HA for mission-critical workloads



When designing replication or highly available scenarios, consider the architecture of the management infrastructure in order to achieve a successful DR failover.

Vblock™ Data Protection can implement EMC RecoverPoint and EMC VPLEX today. Both solutions can be delivered with a Vblock System to ensure a pre-defined architecture, complete compatibility and functionality, and a single line of support for all technologies.

Vblock System and VMware vCloud Director System scalability

This section describes items that affect decisions and recommendations around the scalability of VMware vCloud Director on Vblock Systems. While every Vblock System uses VMware vCenter to manage the vSphere layer, in vCloud deployments the vSphere layer is actually controlled by VMware vCloud Director.

Multiple Vblock Systems in the same vCloud

VMware vCloud Director uses elastic provider vDCs to allow the expansion of a provider vDC beyond the 32-host cluster limit. There is a one-to-one mapping of provider vDCs to organization vDCs, that is, adding provider vDCs means additional organization vDCs. Keep the number of provider vDCs to a minimum to create a cleaner experience for the user and make administration easier for the cloud administrator.

Elastic provider vDCs can be accomplished only through a single data center object within vCenter, by combining two or more clusters within vCloud Director. Keep the following considerations regarding the Cisco Nexus 1000V Switch in mind:

- The maximum allowed for each Virtual Supervisor Module (VSM) pair is 64 hosts and 2,048 virtual machines (as of the publication of this document). Therefore, the maximum elastic provider vDC with the Cisco Nexus 1000V Switch is 64 hosts.
- The suggested approach is to keep Vblock Systems together within a cluster object. It is possible to have multiple clusters in the same data center from different Vblock Systems all managed by the same pair of Cisco Nexus 1000V VSMs.
- For every 64 hosts, deploy a new pair of Cisco Nexus 1000V VSMs for a new data center object. The Cisco Nexus 1000V goes through continual updates to increase the maximum number of hosts and virtual machines supported per VSM pair. Once new releases are integrated into the VCE Release Certification Matrix, those new maximums can be reflected in the design.
- The Cisco Nexus 1000V Switch currently supports a maximum of 2,048 networks. This is a combination of both VLAN and VXLAN networks. To scale beyond 2,048 networks, create additional Cisco Nexus 1000V instances. To provide robust scalability, use a single VXLAN segment with multiple Cisco Nexus 1000V switches. The only requirement is that multiple Nexus 1000V switches must belong to the same multi-cast group.

If using VXLAN for network pools by VMware vCloud Director, a recommended approach is to have as many Vblock Systems be controlled by as few vCenter Servers as possible, preferably a single vCenter Server. VXLAN networks cannot span vCenter Servers at this time; therefore, keeping everything configured under a single vCloud Network and Security Manager makes administration easier.

Combining multiple Vblock Systems into a single vCenter instance requires communication to the delivery team as this is not a standard Vblock System delivery procedure. It is, however, supported.

When approaching vSphere and vCloud maximums for a single vCenter instance, you need a new vCenter Server for cloud resources. Keeping maximums in mind is key to any design. Keep a 50-85% maximum threshold to allow for flexibility.

Multiple vCenter Servers delivered with each Vblock System can be consumed by vCloud Director if the cloud administrator:

- Does not mind having multiple provider vDCs to manage
- Does not care that VXLAN networks are only accessible to a single provider vDC

A single vCloud Director instance can consume multiple Vblock Systems. At the time of this writing, vCloud Director can consume up to ten vCenter Servers and 10,000 powered-on virtual machines. The number of Vblock Systems that adhere to these maximums varies with design.

NFS considerations

During the deployment of a production instance of VMware vCloud Director, an NFS share must be presented to all vCloud Director nodes. This NFS share is responsible for the uploading and downloading of vApps to VMware vCloud Director, deploying vApps across the cloud from a catalog, and storing the response.properties file.

This NFS share can be presented from multiple places.

- If the storage array on the Vblock System has a unified personality, it can be presented directly from the array.
- If the Vblock System is configured as block storage only, you can use a virtual storage appliance. This appliance is a virtual machine deployed on block storage that presents an NFS share to vCloud Director cells. This virtual appliance should reside on the management infrastructure.

Another approach is to create an NFS share on one of the RHEL vCloud cells. This approach is not suggested because the NFS share can go offline during vCloud Director maintenance tasks, such as upgrading or troubleshooting.

Making responsive catalogs

As your VMware vCloud Director footprint begins to expand beyond multiple clusters, multiple data center objects, and multiple vCenter Servers, you need to address the global catalog.

The global catalog is publicly accessible by all tenants of the cloud. It contains golden images of vApps that can be deployed by any tenant in the cloud.

Problems can arise when a tenant is accustomed to deploying a vApp from the global catalog to a provider vDC that exists in the same cluster as the global catalog. The deployment is almost instantaneous. As the cloud expands to multiple clusters, data centers objects, and vCenters, a copy process is performed behind the scenes. The NFS transfer directory is used by exporting the vApp via OVF, copying to the NFS transfer directory, then importing the OVF to the appropriate location. Multiple copy processes must be performed so the deployment process is no longer instantaneous.

To increase the speed at which vApps are deployed from the global catalog, invoke the VMware ESXi copy process. The easiest way to do this is to create a data store based on NFS and present this NFS data store to every host in VMware vCloud Director. If a host can see the data store it needs to copy from, the ESXi copy process can be used without performing OVF exports and imports. If the Vblock System has only block storage, then this data store should be zoned to as many hosts as possible. Since NFS is file-based, it makes mounting data stores across all hosts much easier. It is also possible to create a storage appliance that uses block storage to present an NFS share to all ESXi hosts.

If every organization has its own catalog with customized images, this becomes much harder to police. The recommended approach is as follows:

1. Create one or more NFS data stores and assign each an appropriate storage profile such as “Catalog Use Only”.
2. Create an organization vDC within every organization that has access to this storage profile.
3. Make sure organizations use this storage profile when adding vApps to catalogs.

This approach reduces time spent on deploying vApps from catalogs. However, the cloud administrator must be conscious of maximum data store connections per host within VMware vSphere.

Summary

The following list summarizes best practices recommended by VCE for successful deployment of VMware vCloud Director on Vblock Systems:

- Do not modify the AMP.
- Deploy a VCE-supported dedicated management infrastructure that combines reliability, performance, and scalability.
- Follow VMware vCloud Director recommended practices of using two VMware vCenter Servers. Decide on where they will be placed to control the management infrastructure as well as cloud resources.
- Use the SQL Server that is packaged with the Vblock System or use an existing SQL Cluster for your vCloud Director database.
- When using the Cisco Nexus 1000V Switch, keep scale and maximums in mind.
- Use VXLAN for scaling your network footprint. Remember to check all prerequisites prior to implementation.
- Decide on a provider vDC strategy that will be attractive to your tenants while also providing return on investment.
- As more Vblock Systems are added to a single vCloud Director instance, try to consolidate them into as few vCenter Servers as possible.
- Provide NFS storage for vCloud Director cells and the catalog vApps to provide fast copy processes.

Next steps

To learn more about this and other solutions, contact a VCE representative or visit www.vce.com.

ABOUT VCE

VCE, formed by Cisco and EMC with investments from VMware and Intel, accelerates the adoption of converged infrastructure and cloud-based computing models that dramatically reduce the cost of IT while improving time to market for our customers. VCE, through the Vblock Systems, delivers the industry's only **fully integrated and fully virtualized cloud infrastructure system**. VCE solutions are available through an extensive partner network, and cover horizontal applications, vertical industry offerings, and application development environments, allowing customers to focus on business innovation instead of integrating, validating, and managing IT infrastructure.

For more information, go to www.vce.com.

ABOUT THE AUTHOR

Kendrick Coleman is a virtualization evangelist with a vision for the future of enterprise cloud and end-user computing. In his Sr. Corporate Architect role at VCE, he is responsible for architecting data center solutions for a wide variety of virtualized solutions including Tier 1 applications on VMware vSphere, VMware vCloud Director, and VMware View that have been implemented at Fortune 100 companies. Kendrick has been a breakout speaker at VMworld 2010, 2011, and 2012, holds many certifications from VMware and Cisco, and is a contributing author to VMware vSphere Design, 2nd Edition. Kendrick continues to speak at VMUGs and Partner Conferences around the United States.