VMware vCloud® Architecture Toolkit™
for Service Providers

Developing a Hyper-Converged Storage Strategy for VMware vCloud Director® with VMware vSAN™

Version 2.9

January 2018

Martin Hosken

© 2018 VMware, Inc. All rights reserved. This product is protected by U.S. and international copyright and intellectual property laws. This product is covered by one or more patents listed at http://www.vmware.com/download/patents.html.

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

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

Contents

List of Figures

vmware
CLOUD PROVIDER
PROGRAM

Introduction

VMware vCloud Director® gives service providers the ability to build secure public clouds, which can significantly increase virtual data center efficiency for delivering resources to consumers, by pooling compute (CPU and memory), storage, and networks and presenting them through catalog-based services. This allows VMware vCloud® tenants to consume infrastructure without the operational overhead of having to manually configure and provision hardware resources.

Resources such as compute, storage, and networks are the fundamental building blocks for service providers to consider when designing a solution. By employing vCloud Director, these resources can be grouped into multiple tiers, based on their specification, capacity, or performance capabilities.

Storage is of course a key resource, which vCloud Director abstracts to deliver infrastructure within the Virtual Data Center (VDC), and therefore configuring VMware vSAN™ storage resources to support a vCloud platform requires the consideration of a range of provider use cases and workload types.

Audience

This *VMware vCloud Architecture Toolkit for Service Providers* (vCAT-SP) document is intended for VMware Cloud Providers™, customers, IT planners, virtualization architects and administrators, and others involved in the evaluating, acquiring, operating, or designing of a vCloud Director infrastructure using vSAN storage technologies.

Solution Overview

Storage is probably one of the most significant areas of change for vCloud customers, and therefore is of most interest to VMware Cloud Providers. The aim of this paper is to provide guidance to service providers in developing a strategic approach to transforming storage infrastructure with vSAN for VMware Cloud Providers. The key benefits of the features discussed in this white paper include:

- Service providers can increase storage resource efficiency, simplicity, agility, and Total Cost of Ownership (TCO) in their VMware vCloud® environments by leveraging vSAN.
- vSAN integrates with and enhances vCloud Director functionality through Storage Policy-Based Management (SPBM).
- vSAN, when employed with vCloud Director, can provide the foundation for improved storage agility, service offerings, and efficiency in VMware Cloud Provider Program deployments.

3.1 vSAN Overview

vSAN is the VMware Hypervisor-Converged Software (HCS) storage solution, which comes fully integrated into the VMware ESXi™ kernel, unlike many of the other third-party options, which typically employ a storage virtual machine that runs alongside the other workloads to provide the storage platform. As a result, vSAN allows storage and compute to scale together, like building blocks, allowing for virtual machines to be consolidated into a single host device, with storage provided within the hypervisor itself.

This is achieved as vSAN aggregates locally attached disks within the VMware vSphere® cluster to create a storage solution through a distributed datastore, available exclusively to the cluster, which can be employed to provision virtual machines from VMware vCenter Server®, in the same way that vSphere administrators are accustomed to.

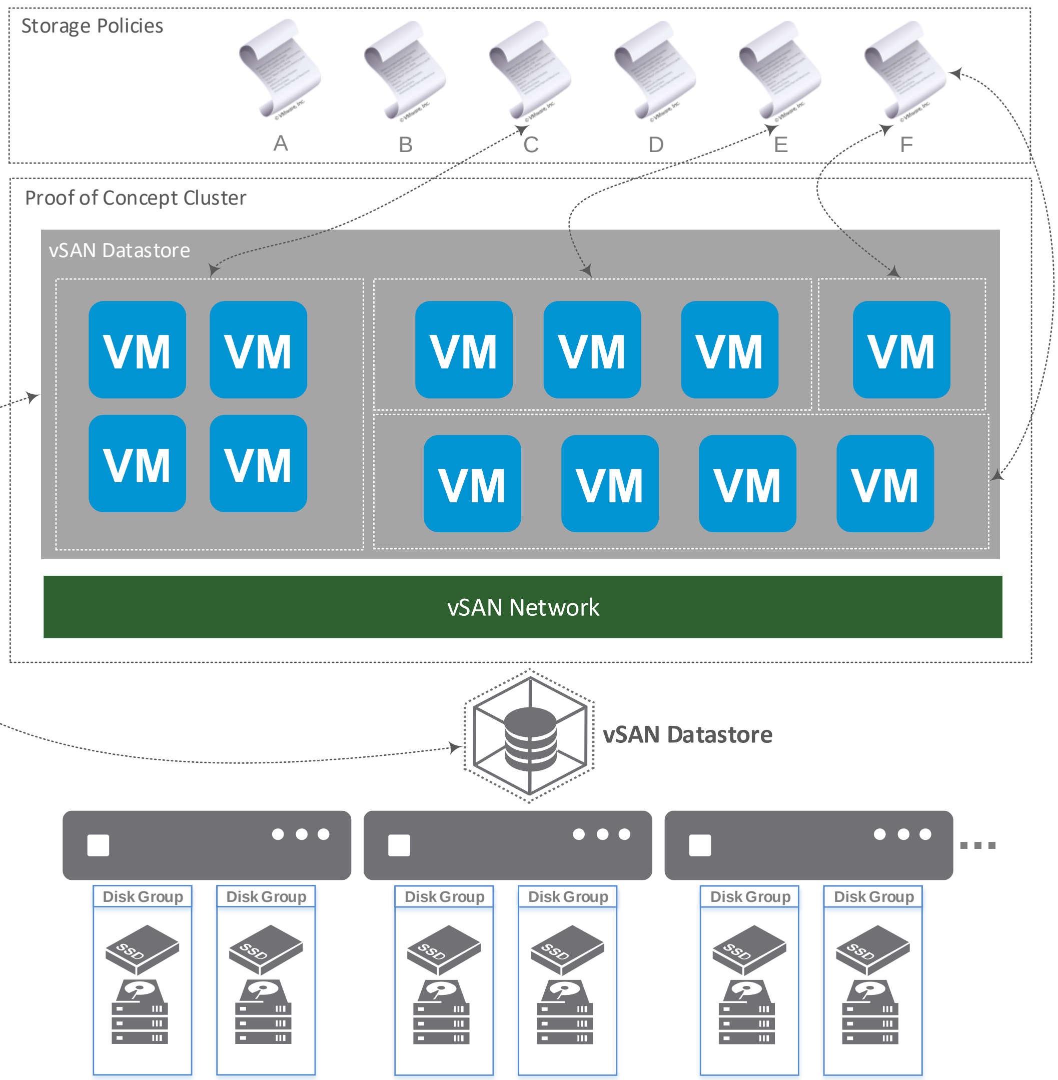
Unlike the traditional VMware vSphere VMFS datastore, vSAN provides an object-based storage datastore model, in that vSAN stores and manages data in the form of flexible data containers, referred to as objects, to provide virtual machine-centric storage services, and deliver capabilities through a Storage Policy-Based Management (SPBM) platform. The SPBM platform and virtual machine storage policies are designed to simplify virtual machine storage placement decisions for vSphere and vCloud administrators. The goal of SPBM is to provide both high availability and scale-out storage capabilities, in the context of quality of service (QoS). This is achieved through the creation of virtual machine storage policies, which are employed to define the level of availability, and in some cases performance, required on a per-virtual machine or virtual disk basis. In addition, vSAN is fully integrated with core vSphere features such as VMware vSphere vMotion®, VMware vSphere High Availability, and VMware vSphere Distributed Resource Scheduler™ (DRS).

A vSAN distributed datastore can be constructed with as few as three vSphere ESXi hosts, although for most use cases four hosts are recommended, each containing at least one disk group with at least one SSD flash drive and one mechanical drive, as illustrated in Figure 1. vSAN can also support up to seven magnetic drives per disk group, and up to five disk groups per host, or up to 35 drives per node, whichever maximum is reached first. In the vSAN hybrid model, the VMware virtual machine files are stored on the mechanical drives, with the SSD flash drive providing read caching and write buffering. Therefore, the SSD drives in each disk group are not used for capacity, but are employed for cache and buffering only, with 70 percent utilized for read cache, and 30 percent for write cache. All writes are targeted to the SSD first, before being destaged to the capacity mechanical drives. vSAN has the requirement that all of the drives are presented individually to the host. Therefore, a storage controller is required that supports JBOD or pass-through mode, so that the ESXi host can see each disk individually. vSAN 6.0 supports scalability of up to 64 nodes per cluster, 200 virtual machines per host, and up to 6400 virtual machines per vSAN cluster.

From the release of vSAN 6.0, the platform also provides support for an all-flash architecture, on flash devices that deliver high, predictable performance and low latency, with response times of less than a millisecond for some of the most storage-intensive applications. However, do not assume that an all-flash configuration always provides higher performance than a hybrid solution. The actual performance is based on a number of factors including, but not limited to, disk group design, active dataset, workload read/write profile, and storage hardware, including disk storage controller performance capabilities.

Note that for simplicity, this document focuses on the hybrid disk group solution, and not an all-flash configuration. However, the concepts discussed in this paper are equally pertinent to both vSAN models, hybrid and all-flash.

Figure 1. vSAN Hybrid Datastore Configuration



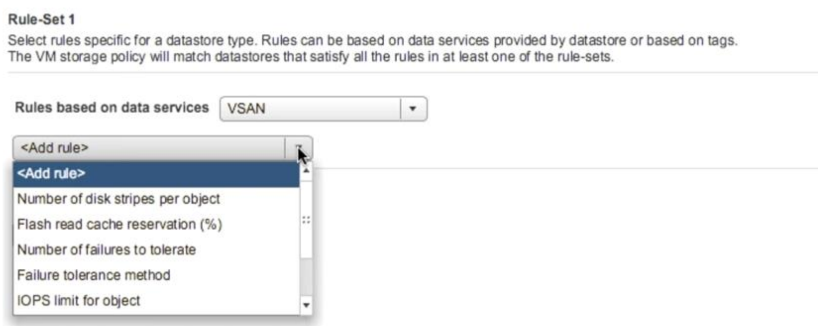
The capacity of the vSAN datastore is dictated by the number of mechanical disks per disk group in the host, and by the number of hosts in the cluster. vSAN is a scale-out solution, in which more capacity and performance can be achieved by adding more disks to a disk group, more disk groups to a host, and more hosts to the cluster.

In a vSAN design, the SPBM platform plays a major part in defining the way in which vCloud administrators can use virtual machine storage policies to specify a set of required storage capabilities for a specific workload, based on the requirements for the application running in the virtual machine.

The following vSAN datastore capabilities are available in vCenter Server, configurable under VM storage policy:

- Number of failures to tolerate
- Number of disk stripes per object
- Flash read-cache reservation
- Object-space reservation
- Force provisioning
- I/O limits (new in vSAN 6.2)

Figure 2. Storage Policy-Based Management Capabilities



vSAN leverages a RAIN (Redundant Array of Independent Nodes) mechanism and employs RAID for its object mirroring capabilities, and also what is referred to as Erasure Coding, which provides data striping in an all-flash vSAN configuration. By using parameters defined in the storage policies, administrators can assign each virtual machine to withstand one or more disk failures, or one or more host failures, providing redundancy and uptime. In addition, the storage policy also allows administrators to define the number of mechanical disks that the virtual machine resides across, by striping the object across multiple disks, potentially increasing performance for data that is not currently residing in the read cache. For more information on defining the these capabilities, find the VMware vSAN documentation pages at https://www.vmware.com/products/virtual-san.

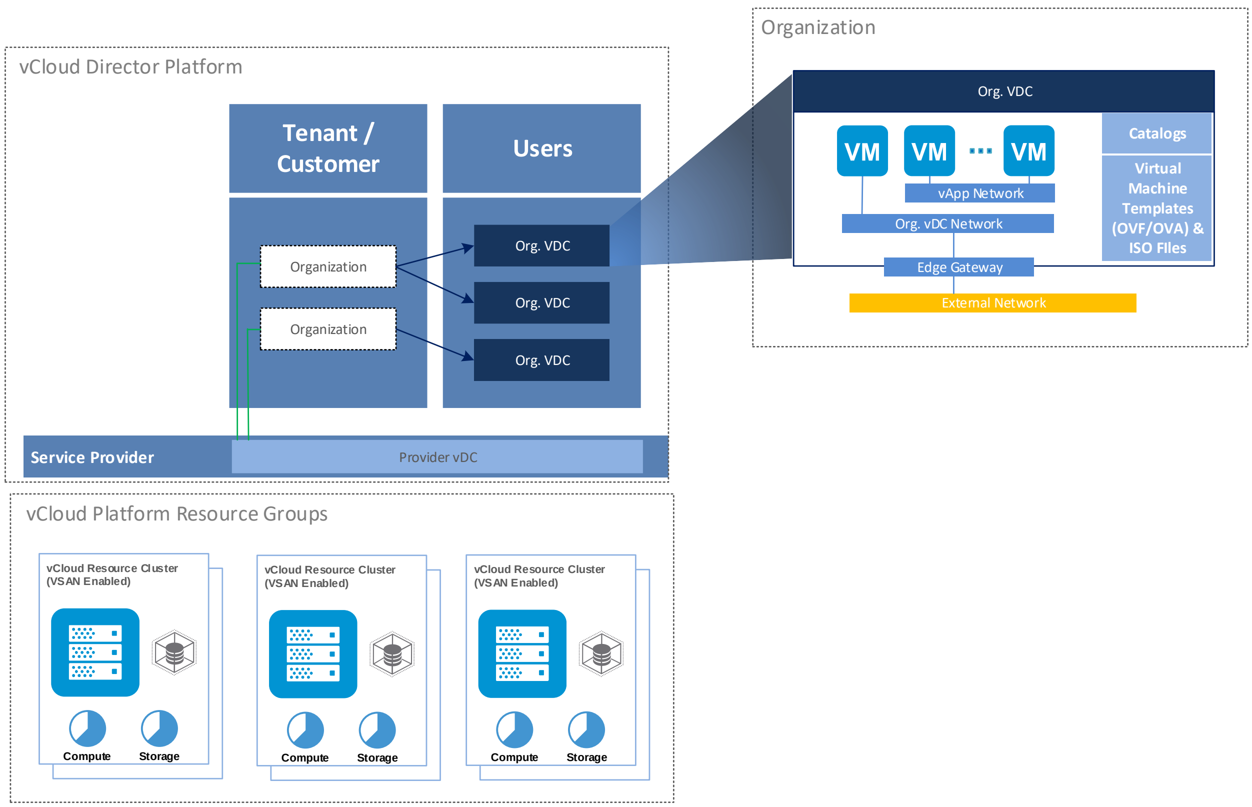
The key design factors associated with the design and implementation of vSAN include the following:

- ESXi host hardware
- vSphere 5.5 U1 or later
- vSAN network – VMkernel port group on 10-GbE (preferred) network infrastructure
- Disk group design – Collections of a single SSD and up to 7 HDDs
  - Up to five disk groups per ESXi host
  - SSD is typically at least 10 percent of HDD capacity in disk group, before the number of failures to tolerate capability is taken into account
- HA cluster
- Storage policy based management

3.2 Enabling vSAN Storage for vCloud Director

VMware vCloud Director for Service Providers creates an abstraction layer on top of vSphere that pools data center resources, such as compute, memory, networks, and storage, along with their associated policies, into Virtual Data Centers (VDCs). In a vCloud Director environment, all physical hardware resources are presented first to vCenter Server, and are subsequently carved up into logical data centers, referred to as Provider Virtual Data Centers (PVDCs). The consumer's organization, defined within vCloud Director, draws on the PVDC resources through the pooled resources assigned to their respective Organization Virtual Data Centers (Org VDCs). Each Organization VDC can only consume resources from a single Provider Virtual Data Center, although a PVDC can provide resources to multiple Organization VDCs. vApps represent the consumer's organization deployed workloads, which reside within the tenants' Org VDC. This architecture is illustrated in the following figure.

Figure 3. vCloud Director Platform



By employing these constructs, vCloud Director allows VMware Cloud Providers to create logical data centers, composed of compute, network, and storage resources, mapped from the underlining physical components. These logical data centers are referred to as Provider VDCs, and play an important part in employing vSAN as part of a vCloud Director public cloud platform. However, in addition to this, Provider VDCs pose the following design considerations:

- A vCloud Director Organization can utilize multiple Organization VDCs
- An Organization VDC can only consume resources from a single Provider VDC
- Provider VDCs can provide resources to multiple Organization VDCs, as illustrated in the previous figure.

As outlined previously, the vCloud Director consumer's vApps represent the tenants' deployed workloads, the resource requirements for which can vary significantly. As with all applications, some vApps require more compute resources, while others might require more disk-intensive operations to be processed. vCloud Director, by design, allows Organization VDCs to be categorized based on workload requirements, or tier of performance required.

CLOUD PROVIDER PROGRAM

To address these varying workload requirements, Provider VDCs must typically be configured to offer multiple service levels within the vCloud environment. These are sometimes categorized as either precious metals, such as Gold, Silver and Bronze, or as various tiers of service, which may include Tier 0, Tier 1, Tier 2, or Tier 3.

However, many service providers also avoid these types of labels. After all, vCloud consumers might not like the idea of placing workloads on lower tiers or lower values of storage. Therefore, this document categorizes storage as either Fast, Standard, or Capacity. The following table provides an example of tiered Provider VDCs.

Table 1. Example of resource tiered PVDCs within vCloud Director

PVDC Tier of Service	Host CPU	Host Memory	Storage Configurations
Fast (Tier 0)	12 Core CPUs	512 GB	vSAN all-flash configuration
Standard (Tier 1)	Hex Core CPUs	256 GB	vSAN hybrid configuration with 10K SAS capacity drives
Capacity (Tier 2)	Quad Core CPUs	256 GB	vSAN hybrid configuration with 7.2k NL-SAS capacity drives

Accurately capturing tenants' application performance requirements, which must be satisfied by each level of service, remains a challenge for VMware Cloud Providers. While guest operating systems typically have well-understood requirements, it is more difficult to quantify application performance requirements, particularly when it comes to storage. In addition, workload storage requirements can be measured in various ways. Some vCloud consumers define their requirements into read and write IOPS, maximum latency requirements, and sequential or random I/O profiles, whereas other consumers of services define their storage requirements from an application perspective, and request resources to meet specific application requirements.

Either way, capturing tenants' workload performance requirements can be challenging for VMware Cloud Providers who might have little or no control over storage tiering and pricing of the VMs being deployed. While applications and operating systems typically have well-defined requirements for CPU and memory consumption, it is often far more difficult to quantify the performance characteristics of storage resources. For instance, a customer might simply request storage to support the capabilities outlined in the following table.

vmware
CLOUD PROVIDER
PROGRAM

Table 2. Capturing Tenant Storage Workload Requirements

	Storage Profile Requirement 1	Storage Profile Requirement 2	Storage Profile Requirement 3
I/O Profile Requirement	1000 heavy Microsoft Exchange 2016 users, with 25 GB mailboxes	Oracle application (1250 IOPS) with a 60:40 read/write	Data warehouse application, (600 IOPS) with a 100% storage read profile
Capacity Requirement	25 TB capacity requirement	3 TB capacity requirement	16 TB capacity requirement

It is important to define both performance and capacity requirements from vCloud consumers, because often vApp storage profiles are not simply about performance or capacity, but are a balance between the two, which can be the storage design challenge within a vCloud Director environment.

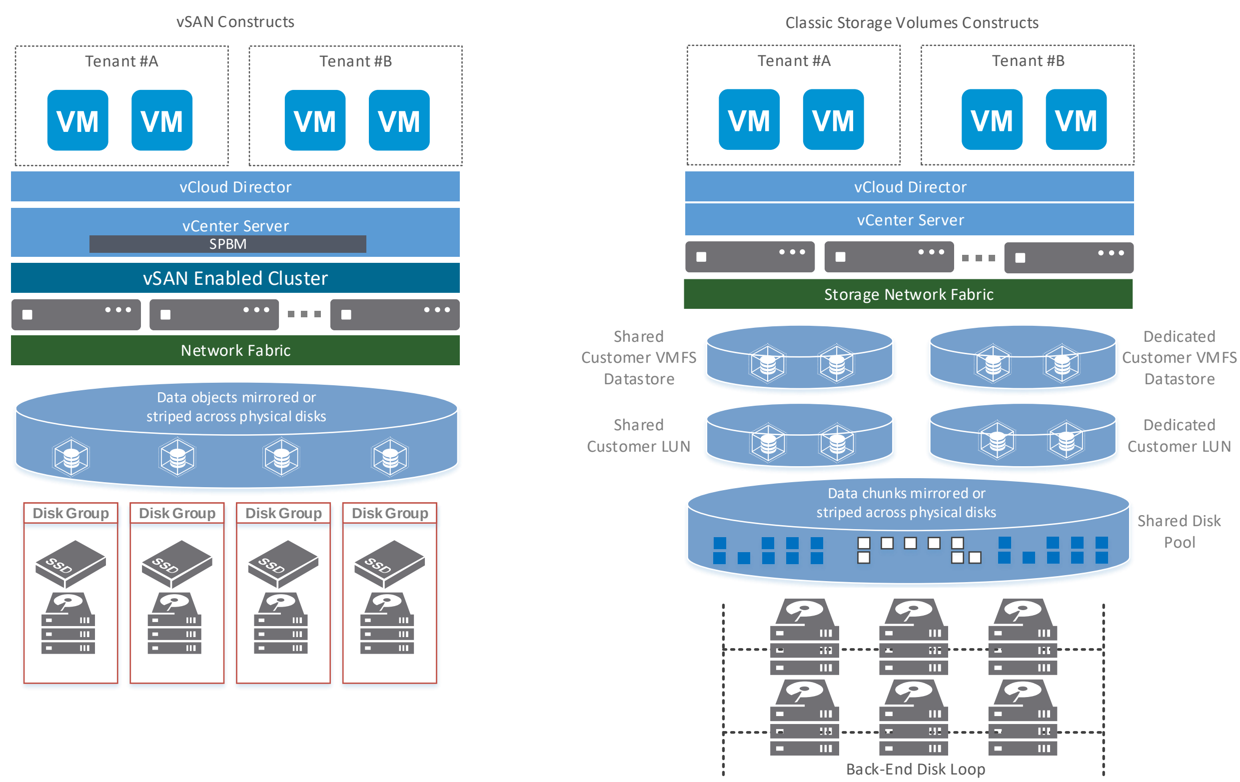
Providing storage resources that can meet all of the various performance requirements set out by tenants, while also providing the ability to scale and manage changes or spikes in demand for resources, are key design requirements for VMware Cloud Providers when implementing vCloud Director for Service Providers with vSAN storage resources. The key aim for any service provider is to allocate storage as efficiently as possible across the platform, so that resources can be utilized by Organization VDCs with a minimum of waste and optimum performance, delivering operational stability and accurate billing. This optimized configuration and allocation of storage across Provider VDCs and consumption by Organization VDCs is vital to achieving an efficient tiered storage model.

vSAN Storage Resource Isolation

vSAN offers a storage layer for multitenanted environments, and is fully capable of offering shared storage to tenants. In this shared storage model, service providers can allow tenants to reside workloads across shared storage resources. The exact location of the tenants' data on the vSAN enabled cluster cannot be specified with this model, because vSAN provides common storage resources within the Provider VDC only.

As a result of this architecture, in any multi-tenanted environment, where various different business entities are storing data across a common vSAN enabled cluster, tenant data is highly likely to be shared across all disk groups, with no way of creating tenant isolation. However, this does not differ significantly from the storage model employed with legacy shared storage systems. In the past it has not been uncommon for service providers to deliver *dedicated* storage to vCloud consumers by provisioning LUNs on a customer-by-customer basis. However, even in this model, it is highly likely that the actual tenant data is striped or mirrored across a common set of physical disks, as illustrated in the following figure.

Figure 4. vSAN Datastore Abstraction Compared with Classic Storage Volume



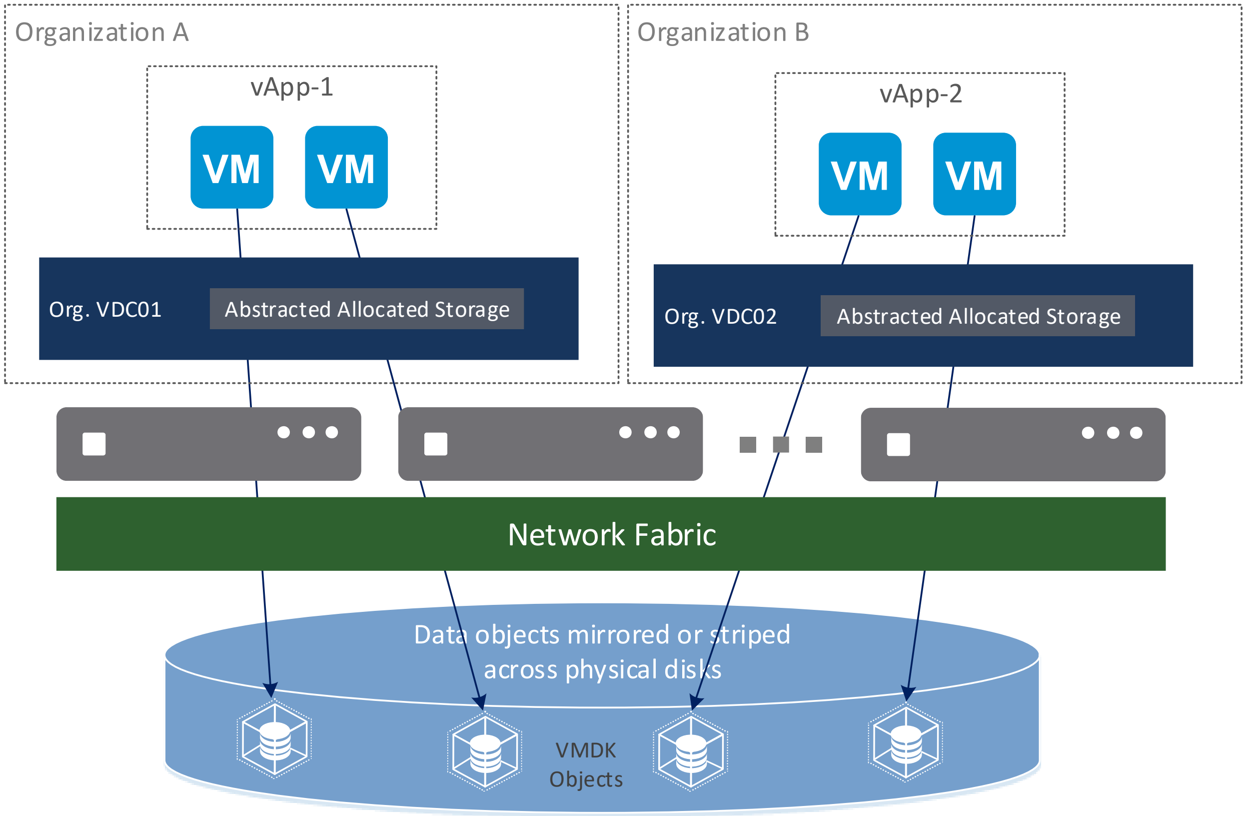
However, in a vCloud Director environment, data being striped or mirrored across a common set of physical disks is not typically of concern to tenants or vCloud administrators. The reason for this is while a vSphere administrator can see the physical storage systems, including a vSAN datastore, these datastores are created through vCenter Server, which requires specific rights, rights that are not required by a vCloud administrator or vApp owner.

Therefore, even if the vCloud Director cell is configured with a user with rights to create new datastores, that privilege is not exposed to the provider system administrator or organization users. Instead, system administrators enable datastores, disable datastores, and assign datastores to organizations.

In addition, organization users, who create and upload vApps, simply store the VM VMDKs on the vSAN datastores assigned to the VDC, without any visibility of the back end disk system providing the resource.

CLOUD PROVIDER PROGRAM

Figure 5. vCloud Director Storage Isolation



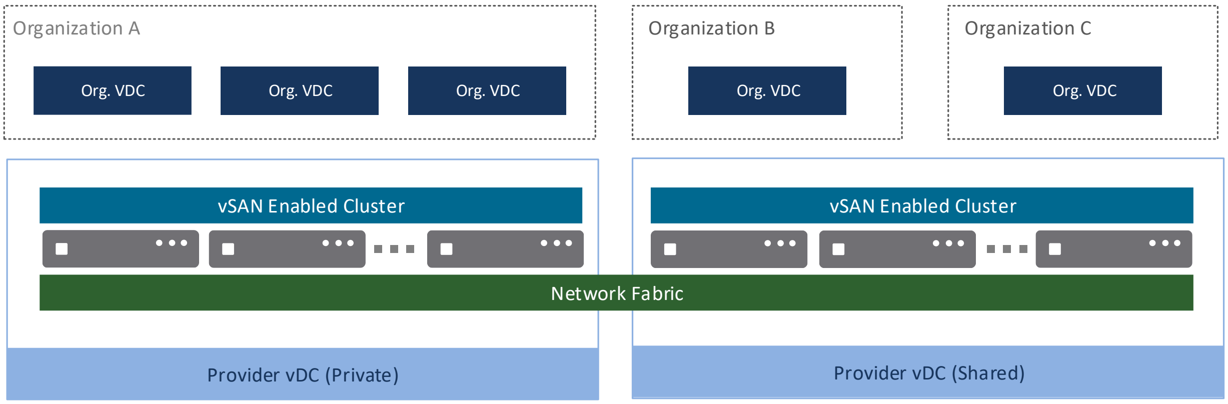
For this reason, virtual machines can never see any storage outside of their VMDKs, unless they have network connectivity to other storage systems (such as NFS shares or CIFS). While this is not recommended; a service provider could give access to external storage for vApps as a network service, but it must be separate from the vSAN datastore assigned to the vSphere hosts backing the vCloud Director platform.

Likewise, vCloud users only see the storage policies backed by one or more of the vSAN datastores assigned to them, but even that view is limited to the vCloud Director abstraction layer, and they are therefore unable to browse the datastore. They only see what is published in catalogs or that which exists within vApps they manage.

Therefore, VMware Cloud Providers can only achieve storage resource isolation when employing vSAN in a vCloud Director environment by assigning dedicated PVDC resources per-tenant, which is not typically an efficient approach to resource management.

However, this approach does enable service providers to eliminate any risk of resource contention across tenant applications, while simultaneously providing the secure separation of user data. This approach might also provide advantages to service providers where performance and capacity requirements from the tenant are dynamic and encompass a wide range of application types.

Figure 6. vSAN Storage Resource Isolation

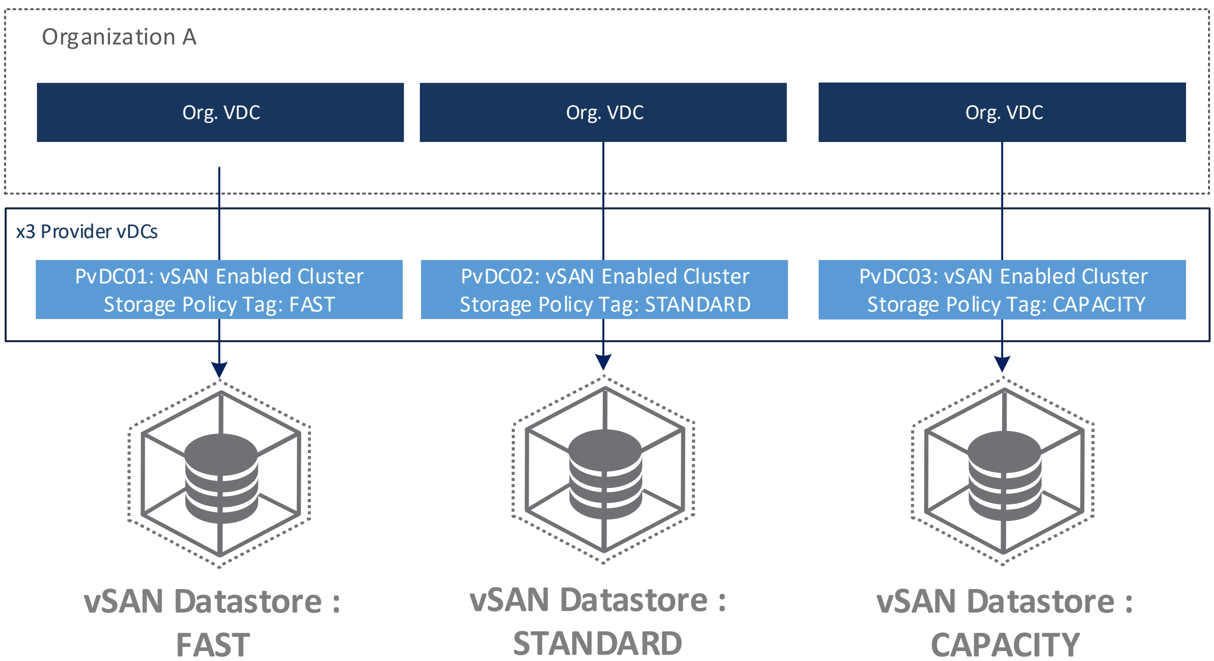


vCloud Director Storage Design with vSAN

vSAN storage policies focus on availability, with some slight exceptions such as I/O limits, read cache and striping (depending on the working dataset), read profile, configuration, and design. Therefore, to provide multiple tiers of storage performance, you must configure multiple resource clusters, each with different defined capabilities and each containing a unique vSAN configuration with the resources being consumed by Organization VDCs. This document addresses two potential design models that can be considered when employing vSAN as the primary storage platform in a vCloud Director environment:

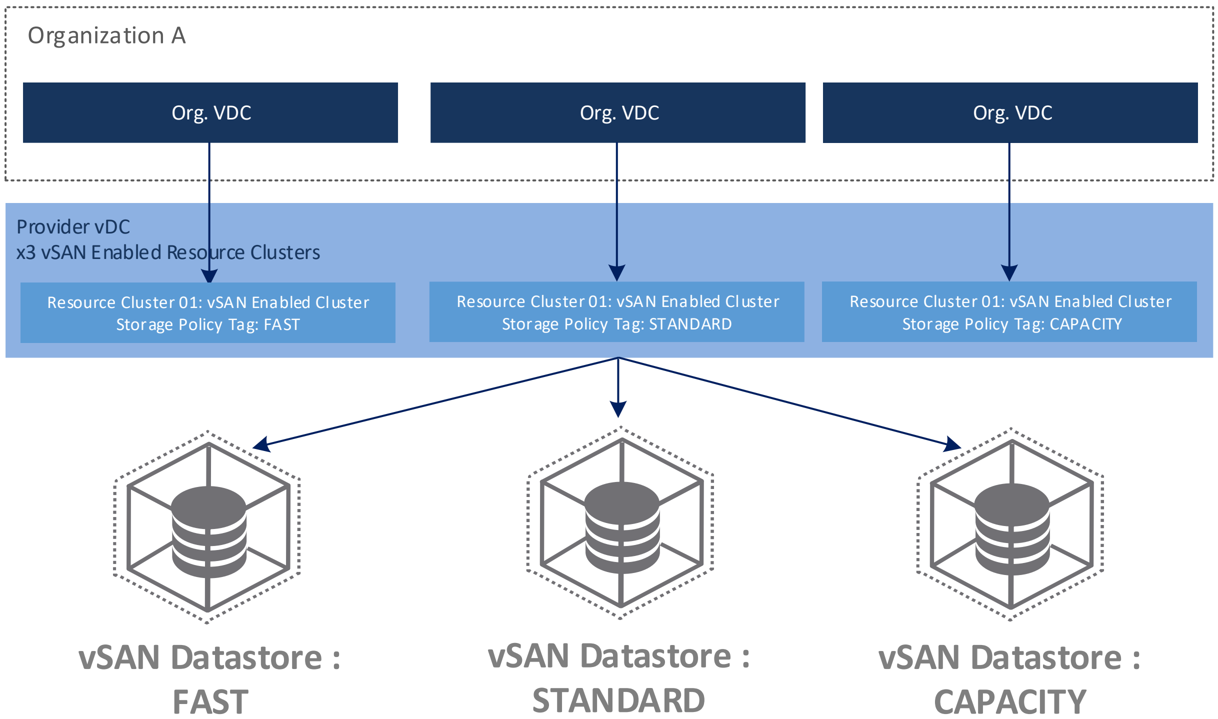
- Multiple tiered resource clusters as dedicated Provider VDCs (1:1 resource cluster tier PVDC)
- Single Provider VDC hosting multiple tiered resource clusters

Figure 7. Model 1: Multiple Provider VDC Storage Tiers Supporting Multiple Org VDCs



CLOUD PROVIDER PROGRAM

Figure 8. Model 2: Single Provider VDC Offering Tiered vSAN Storage to Org VDCs



5.1 Model 1: Multiple Tiered Resource Clusters as Dedicated Provider VDCs

In model 1, each Provider VDC (mapped 1:1 with resource cluster tier) has been configured with a vSAN datastore that meets the specific capability requirements set out by the service provider's service level agreement (SLA) for that tier of service.

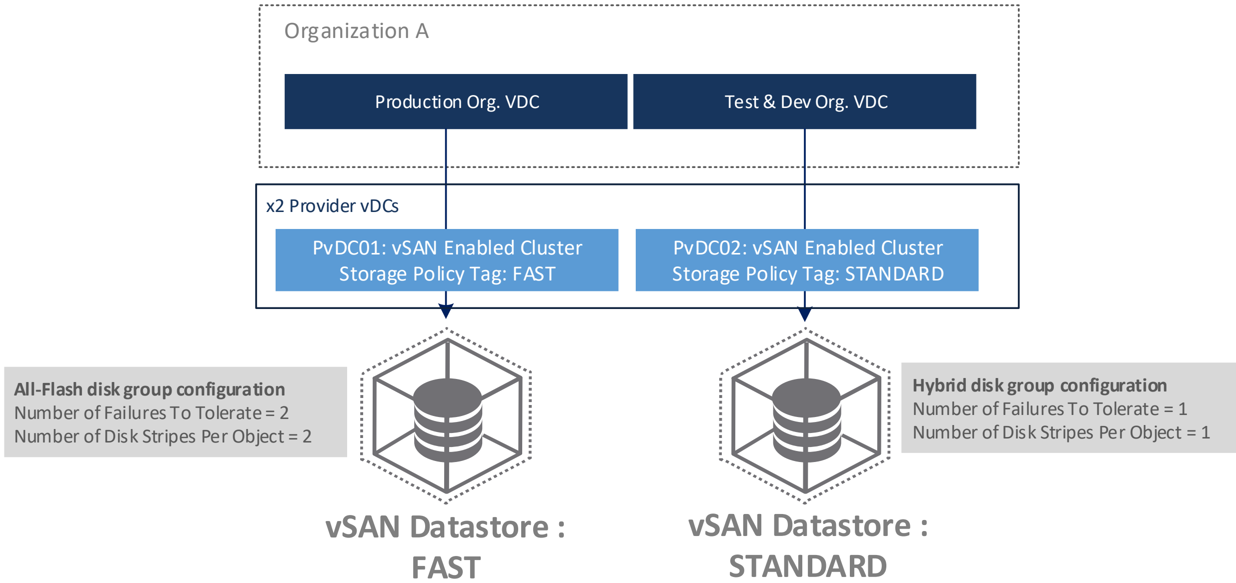
As illustrated in the following figure, the vCloud Director design contains two Provider VDCs offering different tiers of storage:

- Fast performance tier (Tier 0) is backed by an all-flash vSAN enabled cluster, which is offered up to the *Production* Org VDC through PVDC01.
- Standard performance tier (Tier 1) is backed by a hybrid vSAN enabled cluster, which is offered up to the *Test & Dev* Org VDC through PVDC02.

The compute consolidation ratio and vSAN capability, based on the disk group configuration and storage policy, defines how these offerings perform for a consumer. In addition, not shown in the figure, NIOC and QoS are being employed by the service provider so that an appropriate balance of network resources are assigned, based on tier of service.

This example illustrates how two different tiers of storage can be manually configured. However, the exact disk configuration varies depending on hardware manufacturer and provider SLAs.

Figure 9. Manual Storage Tiering within a Single Organization



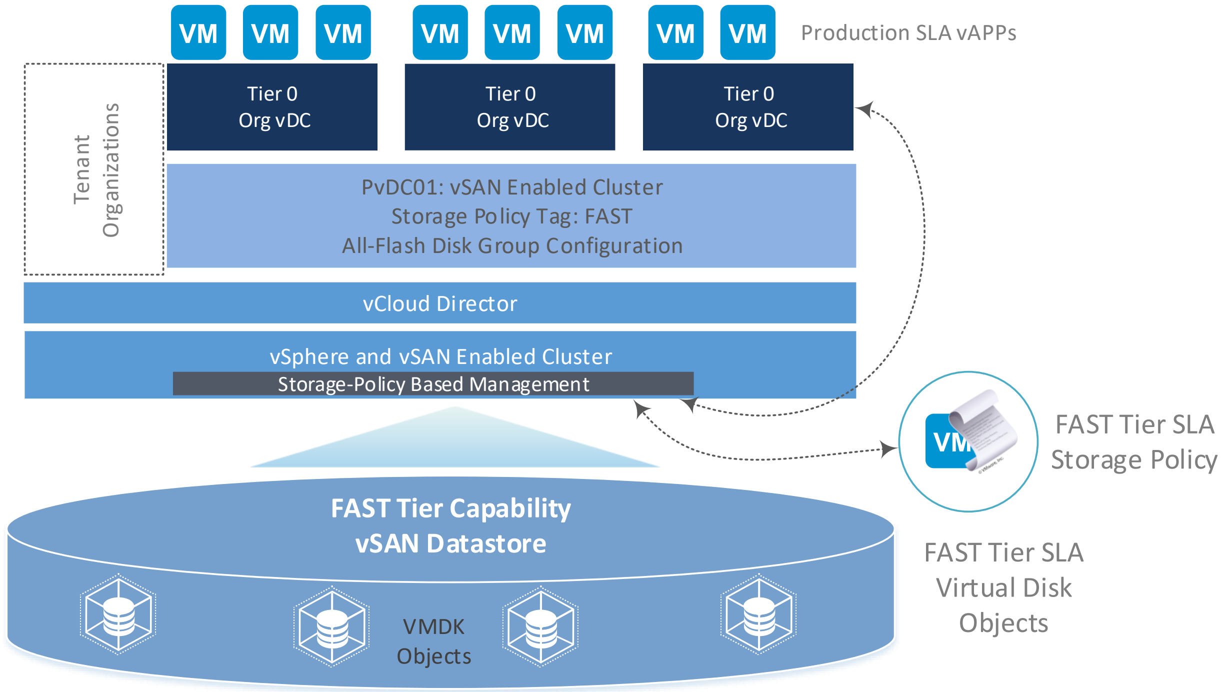
As you would expect, the VMware Cloud Provider configures individual cost models and rate factors specific to each storage tier to accurately charge for the utilization of the storage resources by the various vCloud Director vApps within the organization.

By employing this design strategy, the vCloud Director administrator can take advantage of the multiple tiers of storage offered through vSAN. Different tiers of storage can also be easily mapped to cost models, which can be configured to reflect the level of storage being offered and consumed within the organization.

Typically, tiered service offerings such as these are defined by more than storage capability. vCPU consolidation ratios, levels of guaranteed memory and network resources, backups, and so on, are employed by a service provider to define the SLAs. These aspects are addressed in more detail later in this document.

The following figure shows how the solution is constructed on VMware technologies. The core vSphere platform provides the storage capability through vSAN, which in turn is abstracted through vCloud Director. The vSAN disk group configuration, along with the storage policy, is configured at the vSphere level and automatically presented to the vCloud Director layer. These constructs define the performance, availability, and capacity capabilities of the vSAN datastore, which in turn is employed to define the SLAs for this tier of the cloud offering.

Figure 10. VMWare Technology Solution Stack



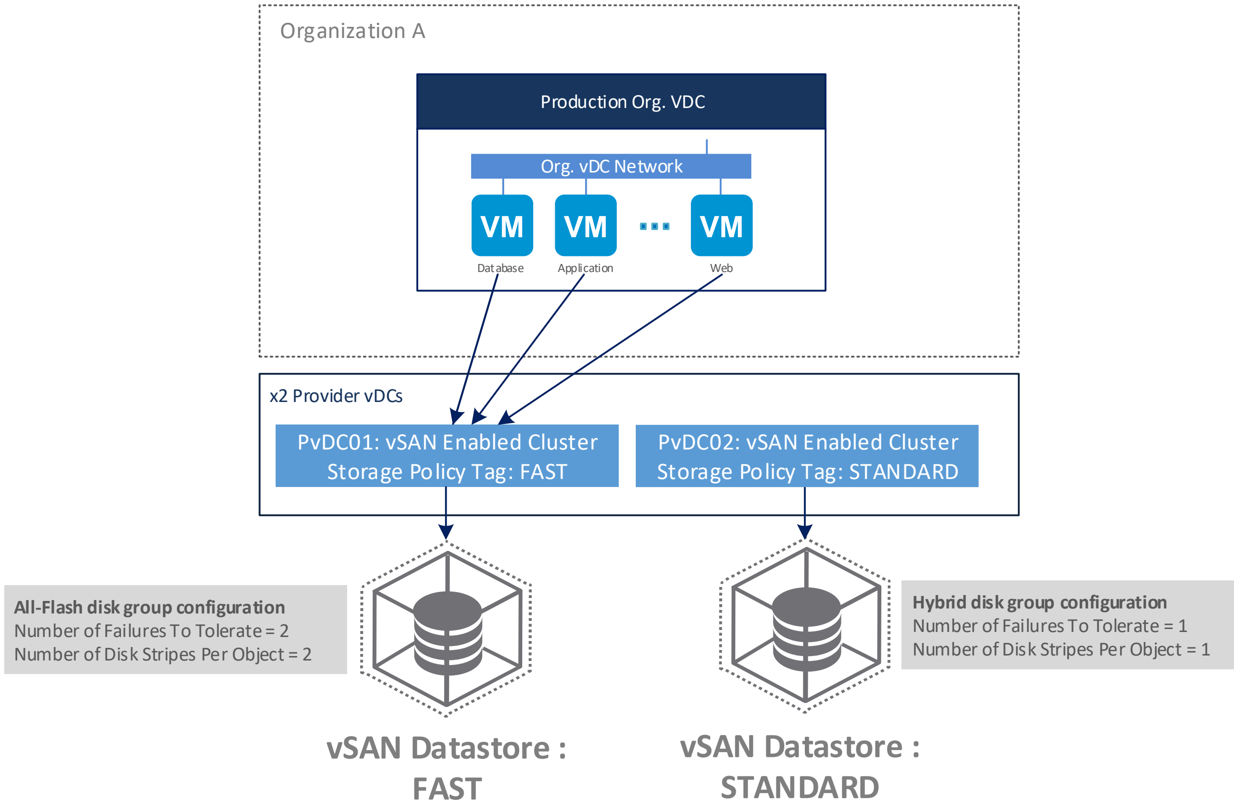
As illustrated in the figure, the vSphere resources are abstracted by vCloud Director into a Provider Virtual Data Center (PVDC). These resources are then further carved up into individual Virtual Data Centers (VDC) assigned to Organizational tenants. The overall result is that the vApps that reside within the Organizational VDCs represent the vSAN storage capability defined by the service provider.

When employing this model, storage tiering within a Provider VDC resource cluster is not possible. Where vSAN is providing the storage resources, only a single type of storage can be configured within each vSphere cluster.

However, for a vApp author, building a typical multi-tier application often requires Tier 0 storage for a database and Tier 1 storage for the web components and middle or application layer. This might present a problem that cannot be solved with this design.

The reason for this issue is, while vCloud Director has been able to employ storage policies since the 5.1 release, and storage policies can be employed to identify the performance characteristics, as defined at the vSphere layer, the storage policies associated with vSAN are predominantly focused on availability, and not performance. Therefore, this means that different performance characteristics cannot be assigned within a single resource cluster, and as discussed earlier in this paper, an Org VDC may only map to a single PVDC's resources. This lack of available storage types within the Provider VDC means that a vApp author cannot deploy their multitier application within a single Org VDC where multiple different tiers of storage are required.

Figure 11. Multi-Tiered Application



However, as illustrated in the previous figure, it is common for Org VDCs to support different vApp types or functions, such as databases and applications, or test and development. Each of these application types might have different performance requirements for their respective vApps. The vCloud Director administrator, however, cannot map the Org VDC to the most appropriate Provider VDC based on their requirements.

For instance, the vApp author might want the databases to reside on the fast tier (all-flash) vSAN storage, while the web and application workloads utilize the hybrid SAS-based standard tier vSAN storage. However, allowing the vCloud Director administrator to manually tier the storage in this way when employing this model is not possible. Therefore, for each of the Org VDCs illustrated, the vApp and associated virtual workloads are based on the single storage type supporting that Org VDC through its Provider VDC.

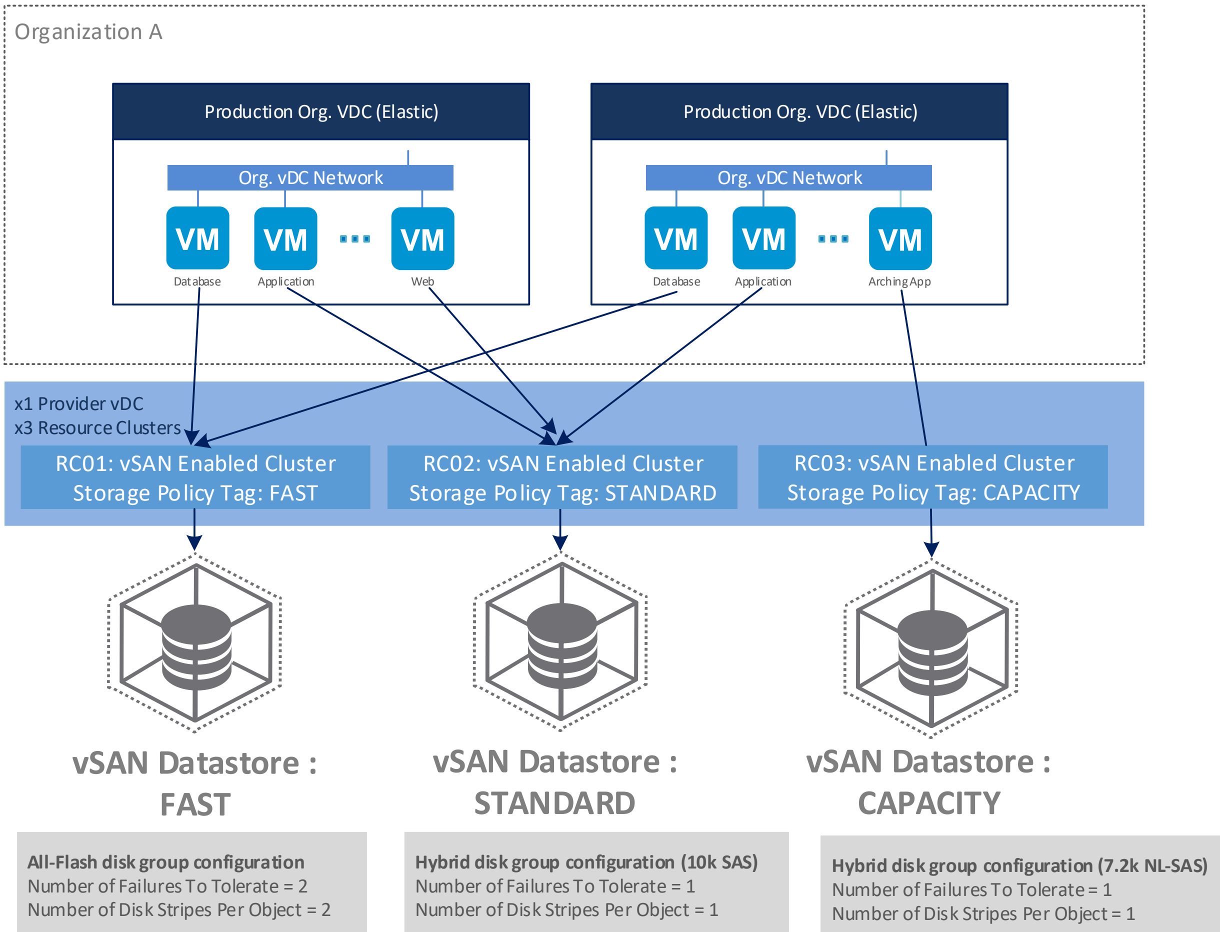
5.2 Model 2: Single Provider VDC Hosting Multiple Tiered Resource Clusters

The second design model takes advantage of elastic VDC, and the ability to span a Provider VDC across multiple vSAN enabled resource clusters.

vCloud Director gives VMware Cloud Providers the ability to create elastic VDCs, which allow an Org VDC to consume resources from multiple vSAN enabled clusters. By having the Provider VDC abstract the resources from multiple vSAN clusters, it not only becomes simpler to grow capacity when needed, but also allows providers to design multiple storage performance and availability capabilities through storage policies, and deliver them to a single Org VDC. However, in this model the compute resources remain consistent throughout the Provider VDC.

In the following figure, there are three different tiers of storage available to the Org VDC, supporting Organization A. An all-flash based vSAN cluster, a 10k SAS-backed tier, and a NL-SAS backed capacity tier.

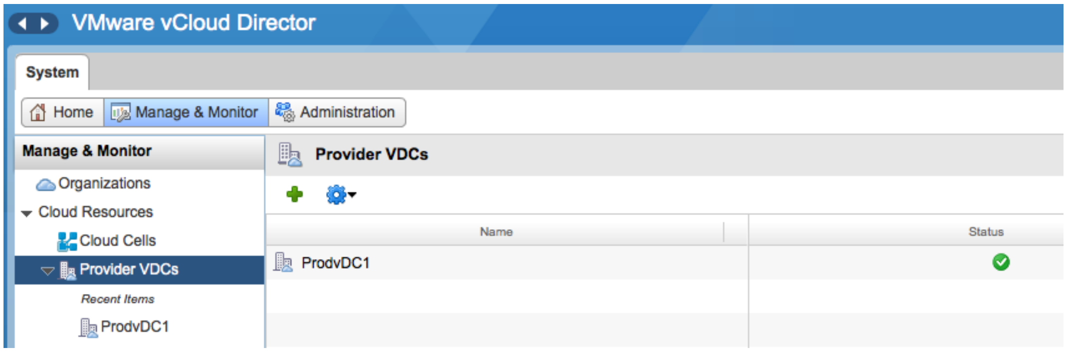
Figure 12. vCloud Elastic Provider VDC vSAN Logical Architecture



Note that the way in which vCloud administrators extend a Provider VDC requires the creation of a Provider VDC primary vSphere resource pool before additional cluster resources can be added. The additional vSphere resource pools clusters must be added in subsequent steps.

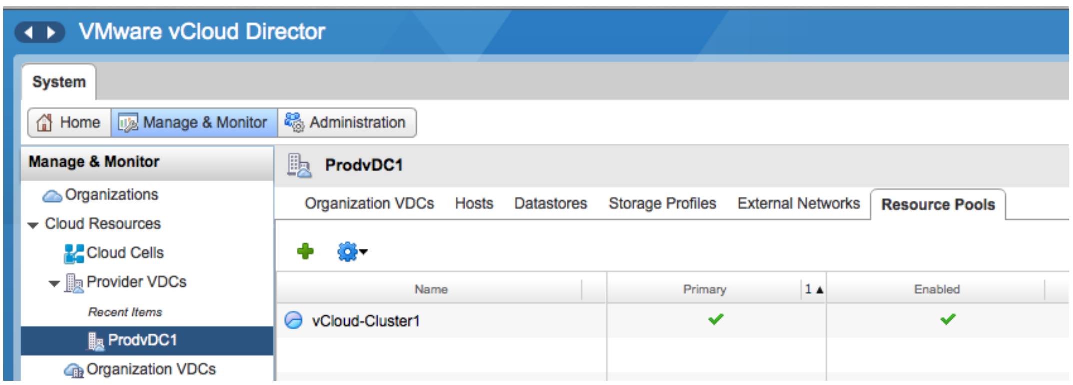
To add additional clusters, go to the **Manage & Monitor** tab and select **Provider VDCs**.

Figure 13. Adding Clusters Through vCloud Director (1 of 2)



Select the provider VDC and click the **Resource Pools** tab.

Figure 14. Adding Clusters Through vCloud Director (2 of 2)



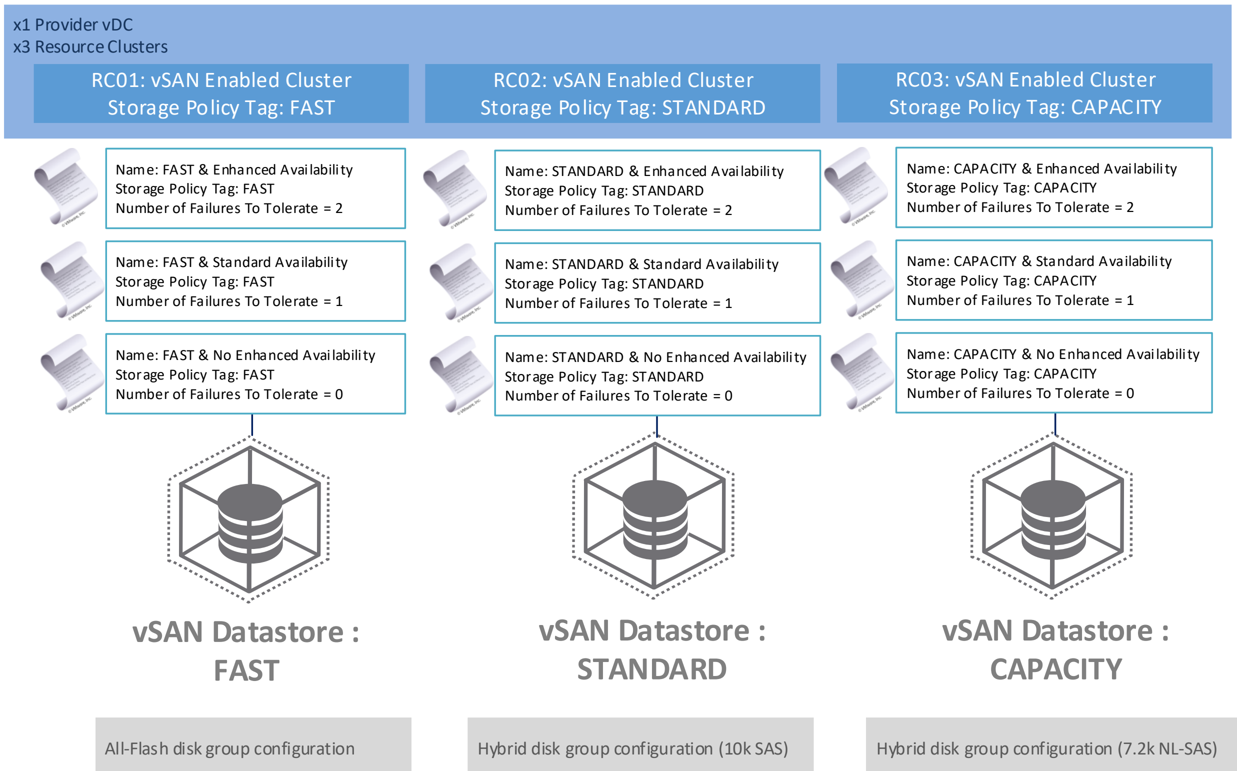
Then select the green plus icon to add another vSAN enabled cluster.

Note A provider VDC can only span clusters managed by the same vCenter Server in the same vCenter Server data center.

The Provider VDC is now able to provide compute and storage resources from multiple vSAN clusters. In vCloud Director, both the pay-as-you-go and allocation pool model Org VDCs are able to consume resources from an elastic Provider VDC.

An additional approach that can be employed with both design models is the inclusion of Storage Policy-Based Management to enable multiple levels of storage offerings within each tier based on availability and SPBM performance capabilities. For instance, if employed, the common storage devices, across the resource clusters, can be assigned multiple storage policies to reflect availability characteristics, such as the number of failures to tolerate (FTT) capability, as illustrated in Figure 12. These storage profiles can then be categorized as Enhanced Availability (FTT=2), Standard Availability (FTT=1), or even No Enhanced Availability, (FTT=0) (not illustrated), with all storage policies configured across a common shared storage class, such as a hybrid configuration backed by 10k SAS drives.

Figure 15. Provider VDC with Mixed Storage Availability Policy Capabilities



The storage policies, which must be configured at the vSphere layer, before being imported into vCloud Director at the Provider VDC level, can be manually assigned by the vSphere / vCloud Director administrator.

For instance, within the Org VDC, the vApp can be deployed on a single storage type, but by leveraging the storage policies, multiple availability capabilities can be chosen. The next section provides more in-depth information on Storage Policy-Based Management capabilities found in vSAN.

Using Storage Policy-Based Management, the vApps and associated virtual workloads operating in the vCloud Director Provider VDC have their data located in one, two or three locations across the vSAN datastore. Although vSAN, and therefore vCloud Director, offers no control over the dynamic placement of the vApp data, the underlining vSAN storage in the Provider VDC is employing its native replication technologies at the back end to meet the appropriate availability requirement defined in the policy.

vSAN Storage Policy Based Management

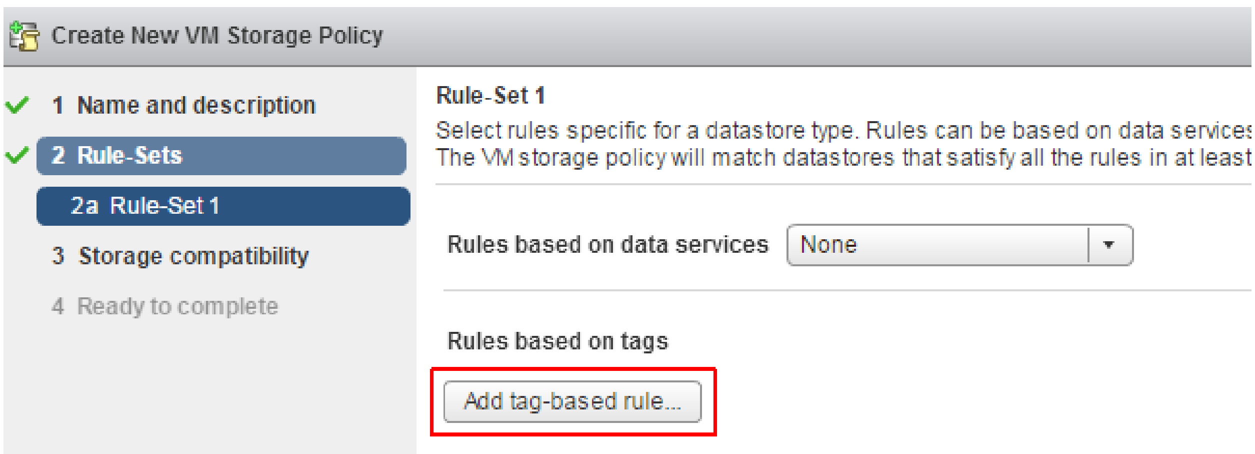
To enable efficient storage operations in vCloud Director, even at scale, when maintaining thousands of vApp workloads, vSAN uses a Storage Policy-Based Management (SPBM) framework.

Storage Policy-Based Management is the foundation of the VMware software-defined storage model control plane, and enables vCloud administrators to overcome a number of storage provisioning challenges, such as capacity planning, differentiated service levels, and maintaining capacity headroom.

By defining storage policies through vCloud Director, SPBM can optimize the vApp provisioning process and leverage the vSphere Storage Policy-Based Management API to automate storage management operations for the software-defined storage infrastructure.

By default, the *(Any) storage policy is created and employed by vCloud Director. However, do not employ the *(Any) policy, which is provided for compatibility reasons only. If the design requires storage tiering within an elastic VDC as described in Model 2, vCloud administrators must use the storage policies leveraged from vCenter Server to place a vApp on a specific storage policy within vCloud Director achieved through the use of tagging rules.

Figure 16. Adding a Tag-Based Rule



After the datastore storage capabilities are configured, the storage policies must first be created in vSphere. These policies are then automatically presented to vCloud Director where they can be viewed within the Provider VDC. These storage policies define the quality of storage that the VM disks reside on, and therefore, must be mapped to the datastores that have been classified.

By default, a storage policy is set when the Org VDC is created, which means that when provisioned by a vApp Author, all virtual machines, unless otherwise specified, utilize that storage policy, and therefore be located on that tier of storage. However, by using this architecture, where a three-tier vApp is required to be deployed, each individual virtual machine within the vApp can have its own storage policy, with the storage policies being specified individually when the vApp is provisioned.

As a result, compliance and deployment of multi-tier applications can be achieved through SPBM. Furthermore, using SPBM in this way through vCloud Director allows the provisioning in an on-demand basis, reducing the burden of storage administrators and empowering the vCloud consumers and vApp authors to take responsibility for vApps and automated provisioning tasks.

This architecture enables the vApp authors to tier the individual virtual machines in their vApps, as required by the application, in each Org VDC. This is made possible by the tiered storage resources available within the single elastic Provider VDC. In addition, the storage policy of a virtual machine can also be changed at any time. By changing the storage policy of the virtual machine, the virtual machine and each of its VMDK files will be migrated to a cluster and datastore which is compliant with the selected storage policy.

VMware vCenter Chargeback Manager™ integrates with the storage policy features presented to vCloud Director from vCenter Server, and can be employed to accurately identify and track storage policies utilized by virtual machines. This makes it possible to meter vCloud consumers based on the virtual machine storage policy, or classification of the storage on which the virtual machine resides.

Impact of vSAN on vCloud Director Compute and Network Resources

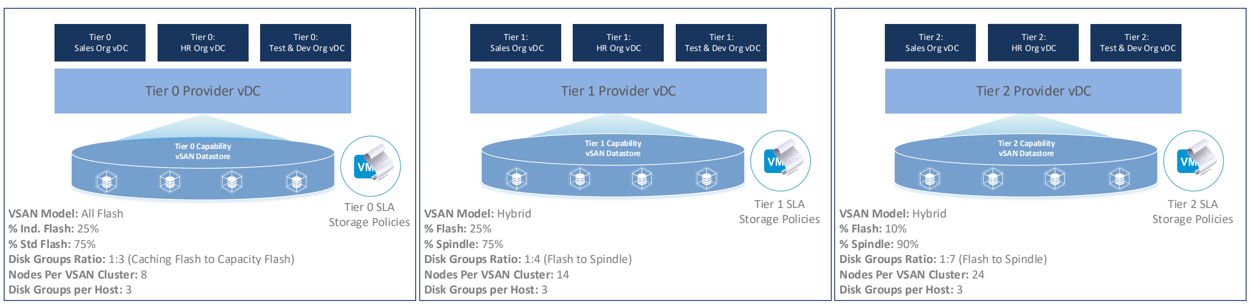
When employing vSAN, it is important to consider the impact on other resources in the vCloud Director environment because tenant organizations also require CPU and memory compute resources to operate. Organizations across a vCloud Director platform draw all of their resources from Org VDCs, whose resources are drawn from the respective Provider VDC. The Provider VDC resources are a collection of CPU, memory, network, and storage abstracted from the underlining vSphere platform.

In a hyper-converged vSAN environment, the appropriate allocation and balance of required CPU, memory and network resources is equally as important to calculate as storage to implement a successful design, and to provide for the ongoing operation of the Provider VDCs across the vCloud Director platform.

Just as a Provider VDC allows the service provider to pool infrastructure resources to create standard offerings to tenants, Provider VDCs can also be employed to group different levels of compute resources to compliment the different tiers of storage. Ensuring a balance of all three resource types is an important design calculation that must be addressed to provide optimum balance of compute, storage, and networks within the PVDC.

For instance, Tier 0 CPU, memory, and network resources can be combined with the fastest storage type to create a balanced Provider VDC, which can be priced to consumers at the appropriate level, typically at a higher cost than Tier 1 or 2, as illustrated in the following figure.

Figure 17. Balanced Provider VDC Resource Offerings



Host Configuration

Component	Specification
ESXi Host CPU	2 x Intel(R) Xeon(R) CPU E5-2690 v2 @ 3.00GHz 10C (60GHz)
Target vCPU:pCPU	1:1
ESXi Host RAM	256GB
ESXi Version	6.0 Build
Network Adaptor	2 x 10Gb SFI/SFP+
Storage Adaptor	2 x 12Gb/s HBA queue depth 600
ESXi Host Disks	SSD (3 x 200GB) Caching SSD (9x 400GB) Capacity

vSAN Disk Group Configuration

Disk	Caching	Capacity
Disk Group 1	1 x 200GB	3 x 400GB
Disk Group 2	1 x 200GB	3 x 400GB
Disk Group 3	1 x 200GB	3 x 400GB

vSAN Storage Policy Configuration

Number of Failures to Tolerate (FTT)	2
Number of Disk Stripes per Object	2
Flash Read Cache Reservation	0
Object Space Reservation	0
Force Provisioning	Disabled

Host Configuration

Component	Specification
ESXi Host CPU	2 x Intel(R) Xeon(R) CPU E5-2690 v2 @ 3.00GHz 10C (60GHz)
Target vCPU:pCPU	3:1
ESXi Host RAM	256GB
ESXi Version	6.0 Build
Network Adaptor	2 x 10Gb SFI/SFP+
Storage Adaptor	2 x 12Gb/s HBA queue depth 600
ESXi Host Disks	SSD (3 x 200GB) Caching SSD (9x 400GB) Capacity

vSAN Disk Group Configuration

Disk	Caching	Capacity
Disk Group 1	1 x 200GB	3 x 600GB
Disk Group 2	1 x 200GB	3 x 600GB
Disk Group 3	1 x 200GB	3 x 600GB

vSAN Storage Policy Configuration

Number of Failures to Tolerate (FTT)	1
Number of Disk Stripes per Object	2
Flash Read Cache Reservation	0
Object Space Reservation	0
Force Provisioning	Disabled

Host Configuration

Component	Specification
ESXi Host CPU	2 x Intel(R) Xeon(R) CPU E5-2690 v2 @ 3.00GHz 10C (60GHz)
Target vCPU:pCPU	4:1
ESXi Host RAM	256GB
ESXi Version	6.0 Build
Network Adaptor	2 x 10Gb SFI/SFP+
Storage Adaptor	2 x 12Gb/s HBA queue depth 600
ESXi Host Disks	SSD (3 x 200GB) Caching SSD (9x 400GB) Capacity

vSAN Disk Group Configuration

Disk	Caching	Capacity
Disk Group 1	1 x 480GB	7 x 900GB
Disk Group 2	1 x 480GB	7 x 900GB
Disk Group 3	1 x 480GB	7 x 900GB

vSAN Storage Policy Configuration

Number of Failures to Tolerate (FTT)	1
Number of Disk Stripes per Object	1
Flash Read Cache Reservation	0
Object Space Reservation	0
Force Provisioning	Disabled

vCloud Director NFS Storage Requirements

As part of the design of a vCloud Director platform with vSAN, the architect needs to address the location of the organization catalogs, and the type and tier of storage on which it is hosted.

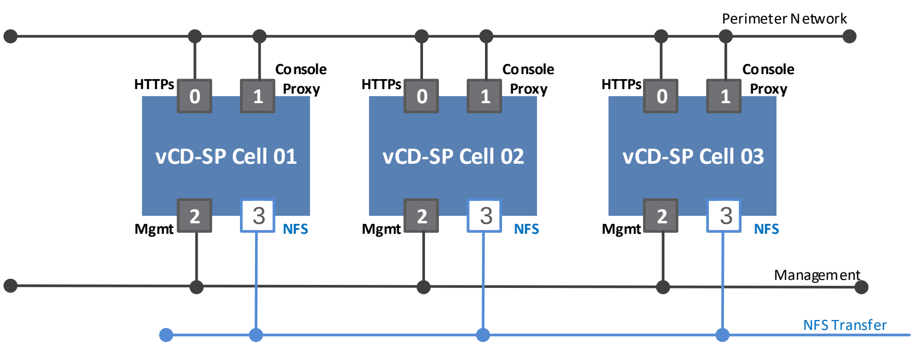
A catalog provides vCloud consumers with a library of content, such as ISO media and vApp templates. Public catalogs allow multiple organizations to share specific catalog items with authorized users and developers having the ability to upload and publish their content to these public catalogs.

Catalog media files and other content do not require the disk performance of *active* systems, and can therefore typically be located on a much lower tier of storage, with the primary design consideration being based on capacity, and ideally thin provisioned disk. NFS storage is well suited to meet this requirement, and the use of third-party independent software vendor (ISV) solutions such as NexentaConnect, can be employed to provide a NFS layer of storage offered by a virtual appliance running on top of vSAN.

In general, VMware does not recommend utilizing vSAN storage for catalogs. The reason behind this recommendation is that all catalog media images, such as ISO files, are uploaded by vCloud Director as file objects into a directory structure under a single folder. Therefore, assuming the same storage policy is used for different catalogs, the catalogs will all share one VM Home Namespace object, with a maximum size of 256 GB. For this reason, third-party virtual storage appliances, which consume vSAN storage and provide NFS file services, are better for providing catalogs across clusters, proving the ability to scale beyond this 256 GB imposed limit.

In addition, another design consideration is that a vCloud Director server group, that is a deployment made up of multiple cells, requires the use of a shared volume, which is referred to as the transfer server storage. The storage employed for the transfer server storage must be available to all vCloud Director cells in the server group, and while any type of shared storage, such as iSCSI, is possible, because NFS storage is easier to set up and more flexible than other shared storage file system types, NFS is typically the preferred choice to meet this requirement.

Figure 18. vCloud Director NFS Network Design Example



When employing NFS for the transfer server storage, a number of configuration settings must be set so that each vCloud Director cell in the vCloud Director server group can mount and use the NFS-based transfer server storage. The ability for each cell to mount the NFS-based location and use it as the transfer server storage is related to the user and group permissions settings that vCloud Director employs.

Another design consideration that supports the use of NFS is the inclusion of maintenance storage. For instance, it is common for all local disks on a host to be allocated to vSAN disk groups and, therefore, the deployment of the ESXi hypervisor is onto USB or SD devices. However, if you allocate local storage to vSAN, you will not have local disks to deploy a local datastore, which is required for persistent logging. The reason for this is because of the I/O sensitivity of USB and SD devices and the hypervisor behavior, which means the installer will not create a scratch partition on these devices. When installing on USB or SD devices, the installer attempts to allocate a scratch region on an available local disk or datastore. If no local disk or datastore is found, `/scratch` is placed on the ramdisk. After the installation, reconfigure `/scratch` to use a persistent datastore.

Therefore, an additional design option to maintain logs and vSAN trace files is to redirect them to an NFS datastore, which can be employed as a centralized maintenance location. For instance, the vSAN trace files can be redirected by employing the `esxcli vsan trace set` command: `vsantraces -> /vmfs/volumes/NFS-Extent/`.

Note This is typically done in addition to the configuration of VMware vSphere ESXi Dump Collector and VMware vSphere Syslog Collector to redirect ESXi memory dumps and system logs onto a centralized network server.

Conclusion

By introducing vCloud Director with a vSAN storage platform, VMware Cloud Providers are now able to develop new ways to reduce the cost of provisioning and resource consumption. vCloud Director layers on top of the vSphere infrastructure in a way that offers great agility in provisioning, isolation for multitenancy, and resource metering. Employing vSAN storage on a vCloud Director public cloud platform allows the further abstraction of storage resources and further simplification of the allocation of storage, with new benefits being gained from automation and reduced operational overhead, while maintaining the scalability required to build and manage a service provider storage platform.