DELLTechnologies

Datastore sharing using HCI Mesh in VMware Cloud Foundation 4.2 on Dell EMC PowerEdge Servers

Abstract

This whitepaper provides an overview of VMware HCI Mesh remote datastore sharing on Dell EMC PowerEdge hardware with VMware Cloud Foundation 4.2 and the instructions to access clusters deployed using SDDC Manager.

September 2021

Revisions

Date	Description
September 2021	Initial release

Acknowledgments

Authors: Peter Giulietti

Support: Anitha R, Sherry Keller

D¿LL Technologies

Table of contents

Executive summary

Business IT departments are challenged to deploy cloud infrastructure in a fast changing, fluid environment. Hyperconverged Infrastructure (HCI) addresses virtualization, software-defined storage, and software-defined networking, but can create silos of unused compute, memory, and storage resources. This unused capacity in deployed clusters may build up over time. One cluster might have excess storage capacity, whereas another might have a surplus of compute and memory resources but might not have the storage capacity to use those resources.

Sizing clusters is a challenge. Requirements change quickly. Redeploying clusters to meet those requirements is a possibility but presents concerns in a data center. Sizing the right balance of compute, memory, and storage capacity is time-consuming and might provide a limited return on an administrator's time investment.

The VMware HCI Mesh technology deployed by Cloud Foundation running on the Dell EMC PowerEdge or PowerEdge MX hardware provides a flexible solution that is easy to deploy, manage, and expand. HCI Mesh allows IT departments to take maximum advantage of their hardware investment.

1   Overview

The goal of virtualization and specifically HCI is to optimize the use of compute, memory, and storage resources deployed in a data center.

In a traditional data center, operating systems are installed directly on the hardware, and applications are installed on the operating system. This monolithic approach can leave untapped compute, memory, and storage resources if the applications running on the server do not consume the installed hardware resources.

Server virtualization leverages a hypervisor running on the bare metal server that can divide server resources into VMs. Management of multiple servers creates the ability to arrange servers into clusters to provide load balancing, failover, and other advantages that are not available in a monolithic server model. Network resources can be divided into multiple network segments and these segments can be exposed to the VMs.

The next evolution was the hyperconverged infrastructure. HCI provided all the advantages of server virtualization, clustering and also allowed for the creation of a vSAN. A vSAN is made up of disks from servers in the cluster. The disks are combined into disk groups to generate a tiered storage hierarchy. These deployed disk groups also manage the redundancy and performance. Storage policies are created to provision the storage capabilities required by different applications.

The HCI is the cornerstone of the Software-Defined Data Center (SDDC). VMware Cloud Foundation is the epitome of the SDDC. An SDDC incorporates a complete HCI in an end to end managed cloud. This provides simplified management, monitoring, flexible deployment options, and security.

Even with all the advantages of these technologies, cumbersome the possibility of clusters running compute and memory intensive VMs that do not require a great deal of storage exists. This can lead to large clusters with a surplus of storage stranded within the cluster. There may also be clusters running VMs that may be starved for storage space yet limited by the storage installed on the physical servers. Until now, the only option was to add servers to the cluster that will add storage to the vSAN, but adding more servers comes at the cost of adding compute and memory that may still go unused. This problem is especially relevant where compute dense modular systems are deployed. These systems prioritize volume over onboard storage, making these deployments susceptible to storage capacity limitations. The Dell EMC PowerEdge MX7000 modular infrastructure compresses eight physical servers into a rack dense 7U enclosure. Each of these modular systems can contain up to six drives. This can result in a rack dense deployment that is rich in compute and memory but may be storage constrained.

Dell EMC PowerEdge rack servers such as R740xd can contain up to 24 drives but is still limited to two CPUs. An eight-node cluster of these R740xd servers would provide the capacity of 192 drives, but the same number of CPU and memory resources as the rack-dense MX7000.

This table shows the comparison of resources in an eight-node cluster that is built on the MX740c modular compute sled against the same cluster built on the rack-mounted R740xd.

Table 1   Compute and storage resource comparison

Eight node cluster	Maximum disks	CPU	Max memory	Rack units
PowerEdge MX7000 MX740c	48	16	1536 GB 6 GB(RDIMM)	7
PowerEdge R740xd	192	16	1536 GB 6 GB(RDIMM)	16

1.1 Process overview

The deployment of a VCF management domain is covered in the Dell EMC VMware Cloud Foundation 4 for PowerEdge Rack and MX7000 deployment guides. Ensure that a VCF management domain and at least one workload domain have already been deployed.

SDDC Manager is used to deploy workload domains and to add additional clusters to existing workload domains. These new clusters must be added to the same data center to share storage resources. The two requirements for enabling HCI Mesh using SDDC Manager are:

- Workload clusters must be configured as vSAN clusters.
- Workload clusters must be deployed to the same data center object in vCenter through SDDC Manager.

SDDC Manager configures a vSAN cluster for datastore sharing through an HCI Mesh. The SDDC Manager workflow allows vCenter to expose other vSAN datastores running within the same vSphere data center.
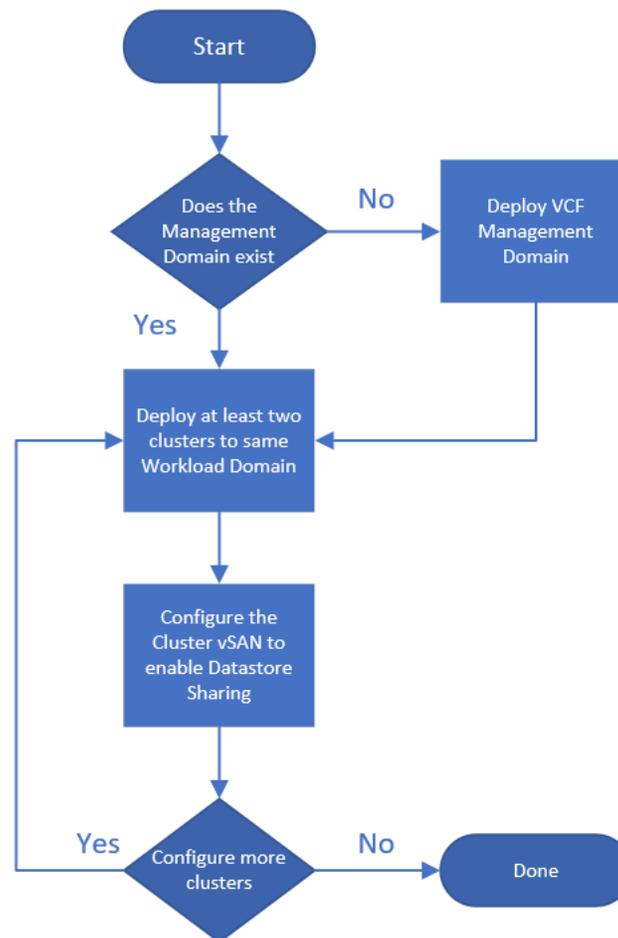
Figure 1   High-level overview of the configured HCI Mesh

2   VMware HCI Mesh and Cloud Foundation 4.2

HCI Mesh is a technology that enables clusters to access unused storage capacity. HCI Mesh allows storage hungry clusters to mount remote datastores running on other clusters within the data center. After mounting, these remote datastores appear as local datastores.

Unlike other storage fabrics such as FC, FCoE, NFS, and iSCSI, HCI Mesh communicates using the existing vSAN VMkernel ports, and vSAN transport protocols end to end. HCI Mesh requires no specialized hardware and no additional networking. The best practice for vSAN networks is to isolate each vSAN network, or VLAN, so that no two vSANs share the same network segment. The vSAN VMkernel ports are configured to route vSAN traffic by overriding the default gateway on the vSAN VMkernel ports.

2.1   VMware Cloud Foundation 4.2

VCF 4.2 added support for HCI Mesh in deployed workload domains. Clusters added to workload domains using SDDC manager have HCI Mesh configured by default. Any vSAN cluster added to a workload domain data center requires only the vCenter interface to access and mount remote vSAN datastores running on other clusters.

The vSAN is routed between the clusters in the same workload domain as shown in below figure 2.
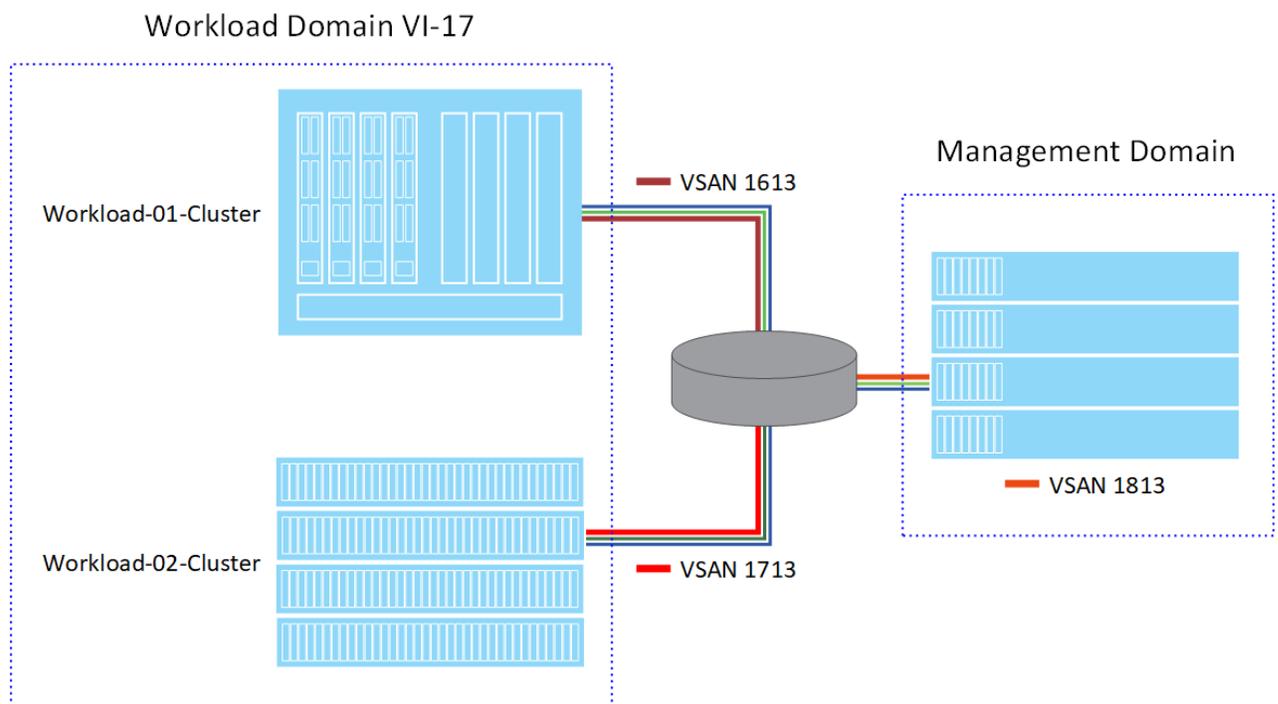

Figure 2   vSAN routing overview

DELLTechnologies

3　Hardware requirements and deployment

The scenario is a four-node, storage-constrained VCF workload domain cluster running on a Dell EMC PowerEdge MX7000 chassis and a storage-rich four-node VCF workload domain cluster running on Dell EMC PowerEdge R740xd rack server.

The rack servers are connected to a pair of Dell EMC Power Switch S5248F-ON ToR (Top of Rack) switches. The MX7000 chassis contains a pair of Dell EMC Networking MX9116n Fabric Switching engines. All these switches connect to an upstream pair of Dell EMC Networking Z9000 series switches configured as an L3, BGP enabled router.

3.1　MX740c compute sleds

- Dual Intel Xeon Gold 5118 CPU
- 256 GB DDR4 RAM
- Dell BOSS-S1 controller with two M.2 SSD disks (boot device)
- Dell HBA330 MX-integrated storage controller
  - Two 750 GB SSD disks (cache disks)
  - Four 1.9 TB SSD disks (capacity disks)
- Dual port QL41232HMKR NIC

3.2　R740xd rack servers

- Dual Intel Xeon Gold 6126 CPU
- 384 GB DDR4 RAM
- Dell BOSS-S1 controller with two M.2 SSD disks (boot device)
- Dell HBA330 integrated storage controller
  - Two 1.9 TB SSD disks (cache disks)
  - Six 3.5 TB SSD disks (capacity disks)
- Dual port Broadcom Advanced 25 Gb Ethernet

3.3　Reference infrastructure

In this paper, our attention is on the workload domain. We assume that the management domain is deployed and connected to the infrastructure. It is recommended not to use the management domain as a client or target of HCI Mesh. HCI Mesh is intended to share storage only between clusters in a workload domain.

The infrastructure deployed is shown below. vSAN requires a minimum of a 10 GbE network connection if a separate hypervisor management connection is configured. To deploy a Dell MX7000 enclosure or a rack server with a single network fabric requires a 25 GbE connection for this the MX9116n is selected as the I/O module (IOM) and the S5248F-ON is selected as the fixed port ToR switch.
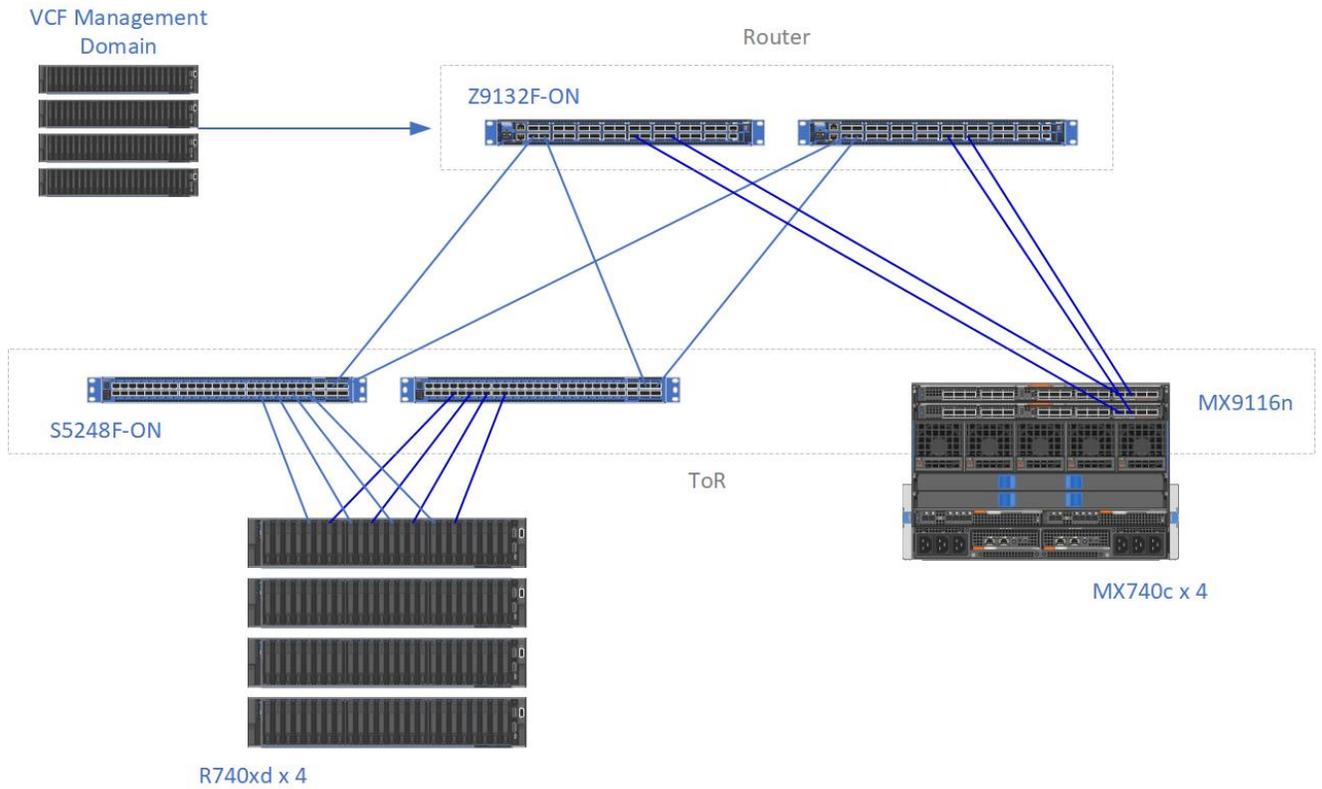
D∕ELL Technologies

Figure 3   Deployed infrastructure

---

4 SDDC Manager workflow

The workflow discussed here describes the process of adding a second cluster to a workload domain. The processes of creating an initial workload domain and that of adding a cluster are similar. The entire process demonstrates the ease of adding new clusters to the same vCenter data center object to share storage using SDDC Manager.

4.1 Prior to deployment

VCF 4.2 requires that the version of ESXi 7.0 update 01d build be updated to build 17551050.

- Prior to the deployment of the management domain, all the hosts are updated to 7.0u1d-17551050.
- Prior to deployment of the workload domain, all the hosts are updated to 7.0u1d-17551050.

Check the product documentation to determine the exact version and build of ESXi 7.0 required to support the version of VCF in use.

Note: Failure to update the host to 7.0u1d-17551050 prior to deployment results in workflow failures to validate due to inconsistent ESXi versions.

To update the hosts to that specific version, it is necessary to first install ESXi 7.0 update 01 on all hosts and then update the hosts using the ESXi-7.0u1d-17551050 depot file.

4.2 Commissioning prerequisites

Prior to commissioning the hosts, ensure that the following are configured on each host:

- Administrator username and password
- VLAN ID for management network, if needed
- IP Address, netmask, and default gateway
- Host Fully Qualified Domain Name (FQDN)
- DNS search domains
- NTP to same source as SDDC Manager, if possible
- MTU of the vSwitch to be used to carry the vSAN or vMotion traffic
- Certificate
    - Either generate a new self-signed certificate, or
    - Configure a certificate issued by your certificate authority.
- Update the ESXi version and build specified for the version of VCF that has been deployed.
- Enable both the ESXi shell and SSH in the troubleshooting section of the host configuration interface.

In addition to the host configuration settings above, a cluster network pool is required. For a vSAN cluster, the network pool includes IP address information for both the vSAN VMkernel and the vMotion VMkernel.

4.3 Commissioning hosts

Commissioning additional hosts for a new cluster is done through SDDC Manager. The process adds the hosts into inventory as unassigned hosts which can then be deployed as a Workload. The process begins at the SDDC Manager interface. On the left side navigation pane, go to **Inventory** > **Hosts**.

1. From the Hosts window, select **COMMISSION HOSTS.**

D∜LL Technologies

A checklist is displayed. Ensure that the hosts to be commissioned meet all the specified requirements. Click each requirement and then click **PROCEED.**

2. From the **Host Addition** and **Validation** panel, select **Import** and then click **BROWSE**.

Note: When adding multiple hosts, it is much more efficient to add hosts together using a json data file which is available on the **Add Hosts** page.

---

3. Select the appropriate json data file, click **Open,** and then click **Upload** to import the data. The hosts will now appear in the workflow pane.
4. Enable a checkmark in the **FQDN** checkbox at the top of the list, confirm the **FingerPrint,** and click **VALIDATE ALL.**
   If the hosts were configured correctly, the message "Host Validated Successfully" appears in the green banner. If the banner turns red, read the error messages and resolve the issues that are detected.
   When validation is successful, click **NEXT** to continue.
   Review the list of validated hosts. If they are correct, click **COMMISSION.**

The commissioning process begins and runs unattended. You can view the progress of the subtasks in the **Tasks** window at the bottom of the SDDC Manager interface. After the commissioning process is complete, the hosts are added to the list of unassigned hosts in the Cloud Foundation Inventory.

4.4   Add cluster to a workload domain

Adding a vSAN cluster to an existing workload domain is also done through SDDC Manager. The workflow allows the administrator to add a vSAN cluster to an existing data center object in vCenter. The vSAN cluster is prepared for HCI Mesh as it is deployed with no additional effort by the administrator. Once the new cluster is deployed, it can mount other vSAN clusters that exist in its data center.

In the SDDC Manager, go to **Inventory** > **Workload domains**.

1. Click the target virtual infrastructure workload domain.
2. From the workload domain, click **ACTIONS** and select **Add Cluster.**
3. From the Storage Selection window, select **vSAN.**
4. In the **General** window, enter a name for the new cluster and click **NEXT.**
5. From the **Networking** window, enter the **VLAN ID** for the new overlay network and select either **DHCP** or **Static IP Pool** for IP Allocation.
6. Reuse an existing pool or create a pool. In this example, we will create a new static IP pool**.**

   a. Enter a name for the new static IP pool.
   b. Enter the **CIDR subnet, IP range, and gateway IP**.

7. Click **NEXT.**
8. Review the object names that are created. Make any adjustments needed and click **NEXT**.
9. Select the **vSAN Storage Failures** to **Tolerate** and then select **vSAN Deduplication** and **Compression**, if required. **Failures to Tolerate** represent the number of cluster hosts that are lost without impacting the vSAN storage.
10. Click **NEXT.**
11. On the **Host Selection** window, select the hosts that make up the cluster and click **NEXT.**
12. From the drop-down list, select the licenses and click **NEXT.**
13. Verify the settings and click **FINISH.**

D∕∕LL Technologies

The workload domain window is displayed. When the cluster is added to the workload domain, the SDDC Manager displays all the configured workload clusters.

D&LL Technologies

5   Mount a remote HCI-Mesh datastore

Now that there are two clusters in the data center, one cluster can mount a vSAN datastore from the other. All the configuration necessary to mount the remote datastores is completed by SDDC Manager when the clusters are deployed.

From the vCenter interface:

1. Select the cluster that mounts the remote vSAN datastore.
2. From the top menu, select **Configure.**
3. From the side menu, scroll down to **vSAN > Datastore Sharing**.
4. In the datastore sharing pane, select **MOUNT REMOTE DATASTORE**.
5. Select the datastore from **Workload-Cluster-01** and click **NEXT**.
   The compatibility check starts and should only take a few seconds.
6. Click **FINISH**.
7. Back in vCenter select Workload-Cluster-02, and then select datastores from the top menu.
   The two vSAN datastores are now available to the cluster. Both the vSAN datastores reappear to be local vSAN datastores and can be used by Workload-Cluster-02.

6   Conclusion

The promise of a Hyper Converged Infrastructure continues to evolve. Sharing resources within a cluster has been the capability driving the adoption of HCI. The expansion of the capability to better leverage shared resources between clusters within a data center is not only a reality, but also easy to deploy and manage. The SDD provides maximum utilization of deployed hardware resources and the greatest return on the hardware investment.

HCI Mesh as a feature of Cloud Foundation SDDC Manager delivers key benefits to businesses and administrators. vSAN clusters may be quickly added to a vSphere data center through a rapid deployment process which ultimately delivers shared capacity, thanks to easily mounted datastores available to any VM within the workload domain. With ease-of-use and high ROI to enable the workload, the value delivered is another reason to leverage VMware VCF 4 on Dell EMC PowerEdge servers.

D‍ELL Technologies