vCloud NFV – ScaleIO Detailed Design

Installation Guide

Dell Networking Solutions Engineering
December 2017

A Dell EMC Technical White Paper

Revisions

Date	Description	Version
September 2017	Initial release	1.0
December 2017	Updated VMware NSX versions	2.0

Table of contents

Revisions....2
Executive summary....6
Audience....6
Document structure....6
1 Prerequisites....7
1.1 Software requirements....7
1.2 Hardware requirements....8
1.2.1 Compute resources....8
1.2.2 Storage resources....8
1.2.3 Network resources....9
1.3 Supporting components....10
2 vCloud NFV detail design....11
2.1 NFV infrastructure design....11
2.1.1 Cluster design....11
2.1.2 Network design....15
2.1.3 Storage design....25
2.2 Virtualized Infrastructure Manager Design....27
2.2.1 VMware vCenter Server....27
2.2.2 Virtual networking design using VMware NSX Manager....29
2.2.3 VMware Integrated OpenStack....29
2.2.4 VMware Integrated OpenStack design....30
2.3 Operations management design....31
2.3.1 VMware vRealize Operations Manager....31
2.3.2 VMware vRealize Log Insight....32
2.3.3 VMware vRealize Network Insight....32
3 Business continuity and disaster recovery using VMware....33
3.1 VMware vSphere Data Protection....33
3.1.1 Storage design....34
3.1.2 Backup policies....35
3.1.3 Monitoring....36
3.2 VMware Site Recovery Manager....36
3.2.1 How does it work....37

3.2.2 Key benefits 37
3.3 VMware vSphere Replication 38
3.3.1 Introduction 38
3.3.2 Architecture overview 38
3.4 Multi-site recovery in VMware Site Recover Manager 39
3.4.1 Using VMware Site Recovery Manager with multiple protected sites and shared recovery site 40
4 Capacity planning and sizing 42
4.1 Sizing guidelines 42
4.1.1 Cluster sizing 42
4.1.2 Storage sizing 43
4.2 Sizing design 43
4.2.1 Management cluster 43
4.2.2 Edge and resource cluster 45
5 VNF onboarding 47
5.1 Capacity requirements 47
5.2 Resource requirements 47
5.3 Operational requirements 47
5.4 High availability requirements 48
5.5 Security requirements 48
5.6 Network requirements 48
5.7 VMware Tools requirements 49
5.8 Onboarding process 49
6 Supporting components 51
7 Monitoring and logging 52
7.1 Logging 52
7.2 ScaleIO logging 53
7.3 Monitoring 53
7.3.1 Metrics 54
7.3.2 Dashboards 56
8 High availability 57
8.1 VMware vCloud NFV infrastructure 57
8.2 Virtualized Infrastructure Manager 57

Executive summary

This document provides the detailed design guidance for creating a second version VMware vCloud NFV™ platform for Network Functions Virtualization (NFV) based on VMware best practices and real-world scenarios. The solution uses the Dell EMC™ ScaleIO® software-defined storage (SDS) solution. The platform dramatically simplifies data center operations, delivering enhanced agility, rapid innovation, better economics, and scale.

Audience

This document is for those individuals who are responsible for the implementation of the VMware vCloud NFV™ ScaleIO®. This document is based on the reference architecture described in the *vCloud NFV Reference Architecture v2.0* document. This document assumes that the audience has some understanding of the VMware and ScaleIO components used and have access to the installation and configuration guides of the respective components.

Document structure

This document is divided into the sections listed in the following table:

Table 1 Document structure

Section	Description
VMware vCloud NFV™ detail design	This section contains the design details for all the components of the vCloud NFV platform with ScaleIO
Capacity planning and sizing	Capacity planning and sizing guidelines
Monitoring and logging	Metrics and dashboards for monitoring the platform are described in this section

1 Prerequisites

1.1 Software requirements

The following table depicts two types of components:

- Required – The solution relies on these components and will not function as planned without them.
- Recommended – These components provide more useful capabilities. These capabilities are discussed in this document. Alternative or third party components could be used where appropriate.

Table 2 VMware software requirements

Component	Version	Required in solution	Functional block
VMware vSphere®			
VMware ESXi™	6.5a	Required	NFVI
VMware vCenter™ Server Appliance™	6.5b	Required	VIM
VMware vSphere® Replication™	6.5	Recommended	NFVI
VMware vSphere® Data Protection™	6.1.4	Recommended	NFVI Ops
EMC® ScaleIO®	2.0.1.3	Recommended	NFVI
VMware vRealize® Operations Insight™			
VMware vRealize® Operations™ Advanced	6.5	Required	NFVI Ops
VMware vRealize® Orchestrator™ Appliance	7.2	Required	NFVI Ops
VMware vRealize® Log Insight™	4.3.0	Required	NFVI Ops
VMware vRealize® Network Insight	3.3.0	Required	NFVI Ops

VMware vCloud Director® for Service Providers	8.20	Required	VIM
VMware® Integrated OpenStack	3.1.0	Required	VIM
VMware Site Recovery Manager™	6.5	Recommended	NFVI Ops
VMware NSX®			
VMware NSX® for vSphere®	6.3.4	Required	NFVI
VMware NSX® Manager™	6.3.4	Required	VIM
Dell EMC™ ScaleIO®			
MDM	2.0.1.3	Required	NFVI
SDS	2.0.1.3	Required	NFVI
SDC	2.0.1.3	Required	NFVI

1.2 Hardware requirements

1.2.1 Compute resources

The compute resources are the physical servers on which the hypervisor is installed. The server nodes contribute CPU and memory capacity to a workload cluster for pooling the resources. These nodes must have sufficient bandwidth and redundancy for the network connectivity of the workloads they host. All hardware used must be on the VMware Hardware Compatibility List (HCL).

1.2.2 Storage resources

This reference architecture uses Dell EMC™ ScaleIO® as the shared storage solution.

ScaleIO is a software-only solution that uses existing local disks and LANs so that the host can realize a virtualized SAN with all the benefits of external storage. ScaleIO software turns existing local internal storage into internal shared block storage. ScaleIO software components are installed in the application hosts and intercommunicate using a standard LAN to handle the application I/O requests sent to the ScaleIO block volumes.

The ScaleIO virtual SAN software consists of three software components:

- Meta Data Manager (MDM) - Configures and monitors the ScaleIO system. The MDM can be configured in a redundant cluster Mode, with three members on three servers, or in Single Mode on a

single server.

- ScaleIO Data Server (SDS) - Manages the capacity of a single server and acts as a back-end for data access. The SDS is installed on all servers contributing storage devices to the ScaleIO system.
- ScaleIO Data Client (SDC) - SDC is a lightweight device driver situated in each host whose applications or file system requires access to the ScaleIO virtual SAN block devices. The SDC exposes block devices representing the ScaleIO volumes that are currently mapped to that host.

ScaleIO components are designed to work with a minimum of three server nodes. The physical server node, running VMware vSphere, can host other workloads beyond the ScaleIO virtual machine. ScaleIO is implemented as a software layer that takes over the existing local storage on the servers. This software layer combines the local storage with storage from the other servers in the environment, and presents logical units (LUNs) from this aggregated storage for use by the virtual environment. These LUNs are presented using the iSCSI protocol and are then usable as data stores within the environment.

The software sits between the disks and the file system at the same layer as a logical volume manager. Virtual machines continue to process I/O to VMDKs within a datastore, however the ScaleIO software now provides this datastore instead of the local disks. In a vSphere environment, ScaleIO is implemented as a separate virtual machine. The software components are installed on the ScaleIO virtual machine.

The Protection domain, which is a large ScaleIO storage pool, can be divided into multiple protection domains, each of which contains a set of SDSs. ScaleIO volumes are assigned to specific protection domains.

The storage pool is a subset of physical storage devices in a protection domain. Each storage device belongs to only one storage pool. When a protection domain is generated, by default it has one storage pool.

Note: See *EMC ScaleIO Basic Architecture Documentation* for more details.

1.2.3 Network resources

Each ESXi host in the cluster should have a network configuration to cater to the redundancy and performance needs of the platform. At a minimum, there should be no single point of failure by providing redundant network controllers and the Ethernet ports.

Connect the Ethernet ports of the ESXi hosts in a redundant configuration to the physical switches. A redundant configuration provides alternate paths if there is hardware failure. VLANs are configured to segregate network workloads such as VMware vSphere® vMotion® traffic, the ScaleIO Virtual SAN traffic, and host management traffic.

1.3 Supporting components

Table 3 Supporting components

Product	Description
Directory server	Centralized authentication source for management components
DNS server	Provide forward and reverse lookup service to all platform components
NTP server	Time sync service to all components
SMTP server	Used to send email notifications from platform as a result of events and alarms
SNMP server	Used to send SNMP alerts to external monitoring systems
SFTP/FTP server	Used for NSX Manager back ups

2 vCloud NFV detail design

2.1 NFV infrastructure design

The VMware vCloud® NFV™ infrastructure components are the ESXi hosts that provide the underlying resources for the virtualized network functions (VNFs). In addition to this, Virtual SAN is used to provide the storage resources for the platform while NSX caters to the network requirements. This section examines the design for the NFV Infrastructure (NFVI) and its components.

2.1.1 Cluster design

The two POD NFV Infrastructure platform contains two clusters, the management cluster, edge and resource (ER) cluster. This architectural best practice allows for efficient resource management, and a clear demarcation between resource providers and resource consumers. Also, this practice establishes security boundaries, and designs different levels of availability based on cluster workloads.

For efficient management of resources, Dell EMC recommends that all hosts in a cluster, have identical configuration and specifications. For better resource management, the management components are deployed in the management cluster. VMware NSX® Edge™ devices for the VNFs are deployed in the edge and resource cluster. NSX Edge devices that the management components use are deployed in the management cluster.

The management cluster consists of five nodes, however Dell EMC recommends a baseline of four nodes. The edge and resource clusters can scale up from a baseline of four nodes. In this example, five nodes are used to meet the needs of tenants. Each cluster has its own fault-tolerant, three-way MDM complex, and native protection domain. Each node has three datastores - Local datastore, ScaleIO_HDD datastore, and ScaleIO_SSD datastore. The ScaleIO_HDD datastore and the ScaleIO_SSD datastore contribute to the ScaleIO volume for the protection domain. SSD datastore serves as a high-performance pool for the ScaleIO volume.

The 2x3 vCenter servers and two NSX manager instances use a 1:1 relationship are deployed in the management cluster. A model of three VCs comprises triple modular redundancy. The first set of three VCs look after components in management cluster. The second set of three VCs manage the VNFs deployed in the edge-resource cluster. Each vCenter Server set points to a load-balanced pair of external VMware® Platform Services Controller™ (PSC) instances.

Note: See section 2.2.1 for more information about PSC design.

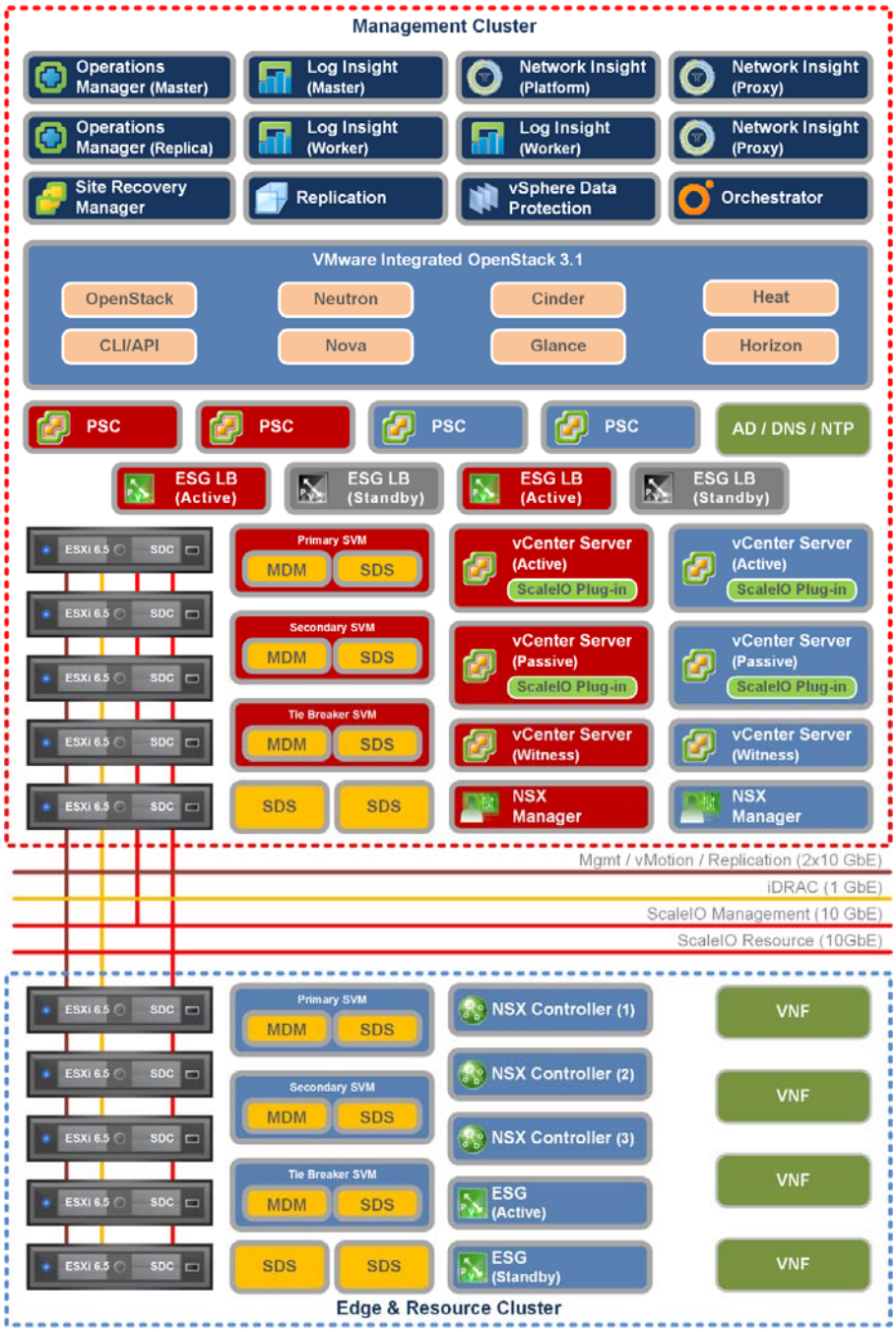

Figure 1 VMware vCloud® NFV™ 2.0 cluster design with VMware® Integrated OpenStack

Management Cluster - The management cluster uses VMware vSphere® High Availability and VMware vSphere® Distributed Resource Scheduler™, and requires specific configuration parameters. The following table lists the parameters pertaining to the management cluster:

Table 4 Management cluster settings

Parameter	Value
VMware vSphere® High Availability	
vSphere HA	Enabled
Host monitoring	Enabled
Admission control	Enabled
Admission control policy	Host failures to tolerate = 1 5-Node ScaleIO configuration supports 1 host failure
VM monitoring	Enabled
Monitoring sensitivity	High
Datastore heart-beating	Automatically select datastores accessible from the host
VMware vSphere® Distributed Resource Scheduler™	
vSphere DRS	Enabled
Automation level	Fully Automated
Migration threshold	2
Virtual machine automation	Disabled
Power management	Disabled
Enhanced vMotion Compatibility (EVC)	
EVC	Enabled
Antiaffinity and affinity rules	
Recommended antiaffinity rules (VMs should be on separate hosts)	1. NSX Edge for PSC load balancer (active, standby) 2. Management PSCs (psc1, psc2) 3. Resource PSCs (psc1, psc2) 4. VMware Integrated OpenStack 5. vRealize Operations Manager (master, replica, data) 6. vRealize Log Insight (master, worker1, worker2) 7. vRealize Network Insight Platform (x1) 8. vRealize Network Insight Proxy 9. vRealize Orchestrator Application
Recommended affinity rules	1. SVM (Master MDM) tied to a ESXi host 2. SVM (Slave MDM) tied to a ESXi host 3. SVM (Tiebreaker MDM) tied to a ESXi host

The edge and resource cluster leverages vSphere HA and vSphere DRS and requires specific configuration parameters. The following table lists the parameters pertaining to the edge and resource cluster:

Table 5 Edge and resource cluster settings

Parameter	Value
VMware vSphere® High Availability	
vSphere HA	Enabled
Host monitoring	Enabled
Admission control	Enabled
Admission control policy	Host failures to tolerate = 1 5-Node ScaleIO configuration supports 1 host failure
VM monitoring	Enabled
Monitoring sensitivity	High
Datastore heart-beating	Automatically select datastores accessible from the host
VMware vSphere® Distributed Resource Scheduler™	
vSphere DRS	Enabled
Automation level	Partially automated
Virtual machine automation	Disabled
Power management	Disabled
Enhanced vMotion Compatibility	
EVC	Enabled
Affinity and anti-affinity rules	
Recommended anti-affinity rules (VMs should be on separate hosts)	1. NSX Controllers (controller1, controller2, controller3) 2. NSX Edge VNF routing (x1 with Standby)
Recommended affinity rules	VNF workloads as defined by vendor SVM (SDS) tied to a particular host

VMware recommends evaluating the performance impact of enabling Enhanced vMotion Compatibility (EVC) on VNF workloads in the resource cluster. See *Impact of EVC on Application Performance* for more details.

VMware recommends enabling the EVC mode and setting this to the processor vendor of the CPUs of the ESXi hosts in the cluster. It is recommended that the hosts have the same CPU vendors, such as all Intel or all AMD, in the same cluster. For details on selecting the EVC mode, see *KB1003212 - Enhanced vMotion Compatibility (EVC) processor support*.

2.1.2 Network design

The VMware vCloud® NFV™ platform consists of infrastructure networks and tenant networks. The infrastructure network traffic consists of EMC ScaleIO SAN traffic (SIO1 and SIO2), and host management traffic.

Management network connects management components such as VMware vCenter Server®, VMware® Integrated OpenStack, VMware NSX® Manager®, VMware vRealize® Operations Manager™, VMware vRealize® Orchestrator™, VMware vRealize® Log Insight™, and VMware vRealize Network Insight. Tenant networks, also referred to as HostIO, provide connectivity to VNFs.

All ESXi hosts in the vCloud NFV platform are configured with two VMware vSphere® Distributed Switch™ (VDS) devices. The VDS devices provide a consistent network configuration across multiple hosts and are a Reference Architecture (RA) requirement. One VDS is used for management and tenant networks, while the other VDS facilitates SIO1 and SIO2. SIO1 and SIO2 each have unique VLANs, but management and tenant networks host multiple VLANs.

The hypervisor communicates through VMkernel-type port groups on the VDS. Virtual machines connect to virtual machine-type port groups and are labeled for easy identification.

The ESXi host's physical NICs are used as uplinks to connect the VDS to the physical network switches. Each ESX host has 4x 10Gbps Ethernet NICs (three is the baseline), each with two ports. Uplink distribution across VDSs per cluster is 3/1, where the VDS trafficking SIO data has just one NIC - one port each for SIO1 and SIO2. ScaleIO manages its own data availability using SIO1 and SIO2, but for Management and VNF traffic one needs to aggregate links explicitly - 2x 10GbE for Management, and 4x 10GbE for VNF Application Traffic, to meet the HA guidelines in vCloud v2 RA.

Lag configurations can be manual or LACP driven, but uplinks in a LAG need to spread across two physical network devices and span different NICs, wherever applicable, to meet HA objectives. There is a fifth 1x GbE port, typically a LOM port, for external connection to hosts and jumphost. For consistent performance, it is important to have network devices from same chipset family and vendor.

All ESXi physical NICs connect to Layer 2 or Layer 3 managed switches on the physical network. At least two 10-Gigabit Ethernet switches with enough ports for all the physical NICs of all the ESXi hosts are required. Port breakouts are necessary if the pairs of switches are 40GbE switches instead.

Table 6 lists the VDS configuration parameters. Since the VXLAN traffic frames are slightly larger in size because of encapsulation, the MTU for each VDS must be adjusted for use with NSX. For best performance, the same MTU size should be set throughout the network.

Table 6 VDS configuration parameters

Specification	Value
MTU	1600 Bytes
Teaming Mode	IEEE 802.3ad, LACP
Segment IDs	5000 – 7999

Network I/O Control (NIOC) prioritizes the network traffic over the two shared uplinks of each VDS. If there is contention, the NIOC share value determines the bandwidth allocation of the networks on the VDS.

Note: See *Performance Evaluation of Network I/O Control in VMware vSphere® 6 guide* for information about NIOC configuration.

Table 9 lists the recommended NIOC shares for this reference architecture.

2.1.2.1 Physical networks

Physical networks consist of physical network switches and computer servers. In the ESXi hypervisor example provided, five nodes are in the management cluster, and five nodes are in the edge resource cluster. The following image shows the physical network topology:

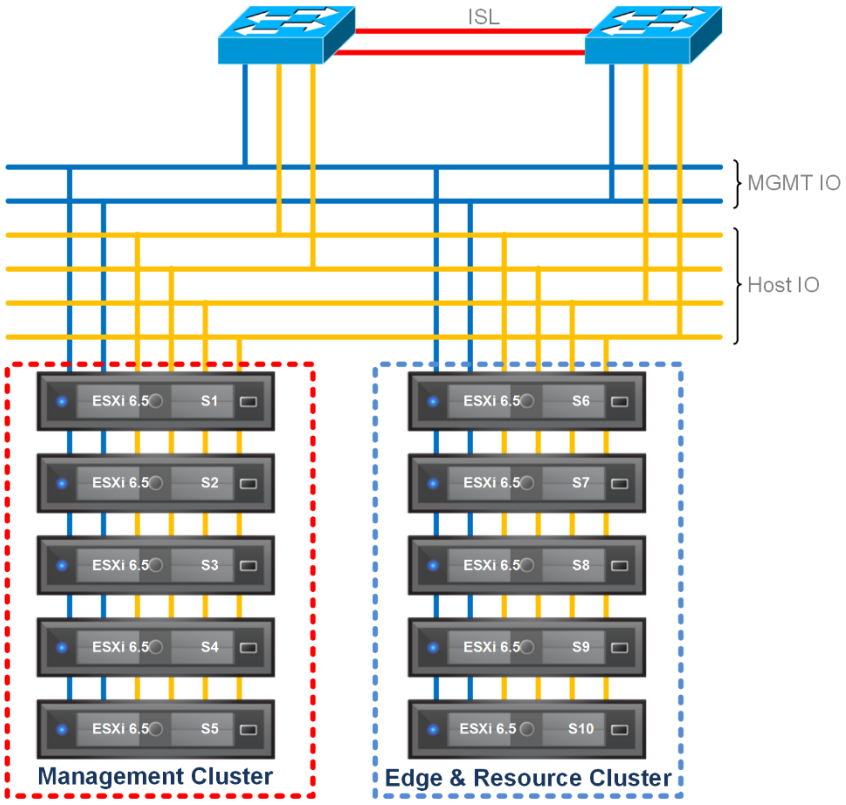

Figure 2 Physical network topology

In the following table, each server node has 4x 10GB NICs, each with 2x ports - two for management IO NIC bonding, four for Hos-tIO NIC bonding, and one each for SIO1 and SIO2. *11-*15 for Management Cluster and *16-*20 for Edge-Resource Cluster. See Table 7 for port-map details.

Table 7 Physical network connection port map

	VMNIC#	Slot::Port	Switch	Switch port	Bond name	
server1	4	1::1	sw1	Te0/65		SIO-1
iDrac: 172.16.104.10	5	1::2	sw2	Te0/65		SIO-2
	6	2::1	sw1	Te0/64	po64	MGMT
	7	2::2	sw2	Te0/64	po64	
	10	4::1	sw1	Te0/81	po80	HostIO
	11	4::2	sw2	Te0/81	po80	
	12	5::1	sw1	Te0/80	po80	
	13	5::2	sw2	Te0/80	po80	
server2	4	1::1	sw1	Te0/67		SIO-1
iDrac: 172.16.104.11	5	1::2	sw2	Te0/67		SIO-2
	6	2::1	sw1	Te0/66	po66	MGMT
	7	2::2	sw2	Te0/66	po66	
	10	4::1	sw1	Te0/83	po82	HostIO
	11	4::2	sw2	Te0/83	po82	
	12	5::1	sw1	Te0/82	po82	
	13	5::2	sw2	Te0/82	po82	
server3	4	1::1	sw1	Te0/69		SIO-1
iDrac: 172.16.104.12	5	1::2	sw2	Te0/69		SIO-2
	6	2::1	sw1	Te0/68	po68	MGMT
	7	2::2	sw2	Te0/68	po68	
	10	4::1	sw1	Te0/85	po84	HostIO
	11	4::2	sw2	Te0/85	po84	
	12	5::1	sw1	Te0/84	po84	
	13	5::2	sw2	Te0/84	po84	

server4	4	1::1	sw1	Te0/71		SIO-1
iDrac: 172.16.104.13	5	1::2	sw2	Te0/71		SIO-2
	6	2::1	sw1	Te0/70	po70	MGMT
	7	2::2	sw2	Te0/70	po70	
	10	4::1	sw1	Te0/87	po86	HostIO
	11	4::2	sw2	Te0/87	po86	
	12	5::1	sw1	Te0/86	po86	
	13	5::2	sw2	Te0/86	po86	
server5	4	1::1	sw1	Te0/73		SIO-1
iDrac: 172.16.104.14	5	1::2	sw2	Te0/73		SIO-2
	6	2::1	sw1	Te0/72	po72	MGMT
	7	2::2	sw2	Te0/72	po72	
	10	4::1	sw1	Te0/89	po88	HostIO
	11	4::2	sw2	Te0/89	po88	
	12	5::1	sw1	Te0/88	po88	
	13	5::2	sw2	Te0/88	po88	
server6	4	1::1	sw1	Te0/75		SIO-1
iDrac: 172.16.104.15	5	1::2	sw2	Te0/75		SIO-2
	6	2::1	sw1	Te0/74	po74	MGMT
	7	2::2	sw2	Te0/74	po74	
	10	4::1	sw1	Te0/91	po90	HostIO
	11	4::2	sw2	Te0/91	po90	
	12	5::1	sw1	Te0/90	po90	
	13	5::2	sw2	Te0/90	po90	
server7	4	1::1	sw1	Te0/77		SIO-1
iDrac: 172.16.104.16	5	1::2	sw2	Te0/77		SIO-2
	6	2::1	sw1	Te0/76	po76	MGMT

	7	2::2	sw2	Te0/76	po76	
	10	4::1	sw1	Te0/93	po92	HostIO
	11	4::2	sw2	Te0/93	po92	
	12	5::1	sw1	Te0/92	po92	
	13	5::2	sw2	Te0/92	po92	
server8	4	1::1	sw1	Te0/79		SIO-1
iDrac: 172.16.104.17	5	1::2	sw2	Te0/79		SIO-2
	6	2::1	sw1	Te0/78	po78	MGMT
	7	2::2	sw2	Te0/78	po78	
	10	4::1	sw1	Te0/95	po94	HostIO
	11	4::2	sw2	Te0/95	po94	
	12	5::1	sw1	Te0/94	po94	
	13	5::2	sw2	Te0/94	po94	
server9	4	1::1	sw1	Te0/97		SIO-1
iDrac: 172.16.104.18	5	1::2	sw2	Te0/97		SIO-2
	6	2::1	sw1	Te0/96	po96	MGMT
	7	2::2	sw2	Te0/96	po96	
	10	4::1	sw1	Te0/99	po98	HostIO
	11	4::2	sw2	Te0/99	po98	
	12	5::1	sw1	Te0/98	po98	
	13	5::2	sw2	Te0/98	po98	
server10	4	1::1	sw1	Te0/101		SIO-1
iDrac: 172.16.104.19	5	1::2	sw2	Te0/101		SIO-2
	6	2::1	sw1	Te0/100	po100	MGMT
	7	2::2	sw2	Te0/100	po100	
	10	4::1	sw1	Te0/107	po106	HostIO
	11	4::2	sw2	Te0/107	po106	
	12	5::1	sw1	Te0/106	po106	

	13	5::2	sw2	Te0/106	po106	
server11	4	1::1	sw1	Te0/103		SIO-1
iDrac: 172.16.104.20	5	1::2	sw2	Te0/103		SIO-2
	6	2::1	sw1	Te0/102	po102	MGMT
	7	2::2	sw2	Te0/102	po102	
	10	4::1	sw1	Te0/109	po108	HostIO
	11	4::2	sw2	Te0/109	po108	
	12	5::1	sw1	Te0/108	po108	
	13	5::2	sw2	Te0/108	po108	
server12	4	1::1	sw1	Te0/105		SIO-1
iDrac: 172.16.104.21	5	1::2	sw2	Te0/105		SIO-2
	6	2::1	sw1	Te0/104	po104	MGMT
	7	2::2	sw2	Te0/104	po104	
	10	4::1	sw1	Te0/111	po110	HostIO
	11	4::2	sw2	Te0/111	po110	
	12	5::1	sw1	Te0/110	po110	
	13	5::2	sw2	Te0/110	po110	

2.1.2.2 Virtual network

This section elaborates on the layout of the virtual network. Key component shaping the virtual network is vSphere Distributed Switches (VDS). There are two VDSs per cluster - one mapped to the SIO underlay, and one mapped to the management and HostIO underlay. Distributed switches interconnect VM kernels and VMs, following VLAN tag specifications defined on a port-group basis. VDS uplinks are mapped to physical NICs, and flow is governed by teaming policies. The following image shows both VDSs in the management cluster, which has five ESX nodes:

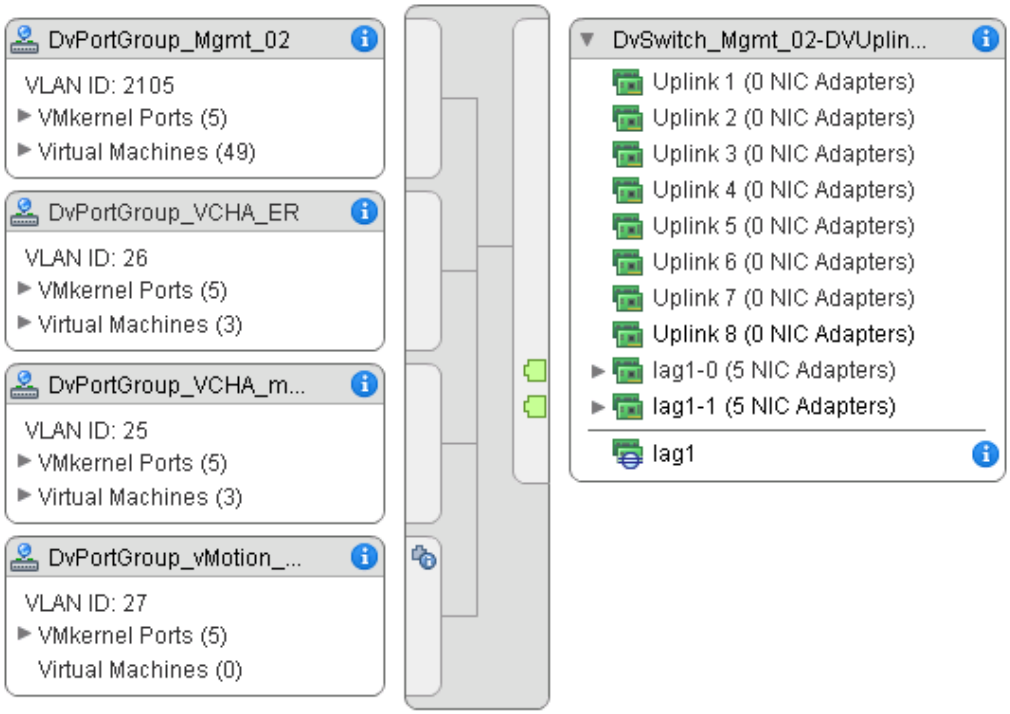

Figure 3 Management VDS and infrastructure VDS in management cluster

By default, LACP is not added automatically within VDS. For VDS, there is an option to use LACP, static LAG, or VMWare native bonding mechanisms. Dell EMC recommends Best practice is to stay steady for a cluster, and configure the switch accordingly.

Note: For assistance in manually configuring LACP, see LACP Support on a vSphere Distributed Switch

2.1.2.3 Infrastructure networks

Each ESXi host has multiple VMkernel port groups. The allocation of VMkernels are necessary for the following networks:

- vMotion Network - Network for vSphere vMotion traffic
- ScaleIO Virtual SAN Network - Network for ScaleIO Virtual SAN shared storage traffic
- ESXi Management - Network for ESXi host management traffic

VLAN trunking is necessary for management and HostIO networks to enable the realization of multiple logical networks on the underlay to meet vCloud NFV RA requirements. The following table lists the VDS configuration for management and ER clusters:

Table 8 Infrastructure VDS configuration

Post group	Type	Teaming policy	VLAN	Cluster
ESXi Management	VMkernel	Load Based	2105, 25, 26	Management

vMotion Network	Vmkernel	Explicit Failover	27	Management
ScaleIO Virtual SAN Network	Vmkernel	Explicit Failover	30	Management
ESXi Management	Vmkernel	Load Based	2105	Edge & resource
vMotion Network	Vmkernel	Explicit Failover	27	Edge & resource
ScaleIO Virtual SAN Network	Vmkernel	Explicit Failover	40	Edge & resource

The NIOC share values are configured at the VDS level. The following table lists the recommended I/O parameters for the Infrastructure VDS for each of the two clusters:

Table 9 Infrastructure VDS NIOC parameters

Network	Limit	Shares	NIC Shares	Share value	Cluster
ESXi Management traffic	Unlimited	Normal	Normal	64	Management
vMotion traffic	Unlimited	Normal	Normal	64	Management
ScaleIO Virtual SAN traffic	Unlimited	Normal	Normal	100	Management
ESXi Management traffic	Unlimited	Normal	Normal	64	Edge and resource
vMotion traffic	Unlimited	Normal	Normal	100	Edge and resource
ScaleIO Virtual SAN traffic	Unlimited	Normal	Normal	100	Edge and resource

2.1.2.4 Tenant network - secure multitenancy

Tenant networks are used to interconnect the VMs of the vCloud NFV platform in ER clusters. These are configured on a dedicated tenant VDS in each ER cluster. The tenant networks include:

- VNF Network - VXLAN based network for VNF to VNF communication
- Management VLAN - VLAN based network for management component communication

The following table lists the recommended I/O parameters for the tenant VDS for each of the three clusters:

Table 10 Infrastructure VDS NIOC parameters

Network	Limit	Shares	NIC Shares	Shares value	Cluster
Management VLAN	Unlimited	Normal	Normal	64	Management
Management VLAN	Unlimited	Normal	Normal	64	Edge and resource
VNF Network	Unlimited	Normal	High	100	Edge and resource

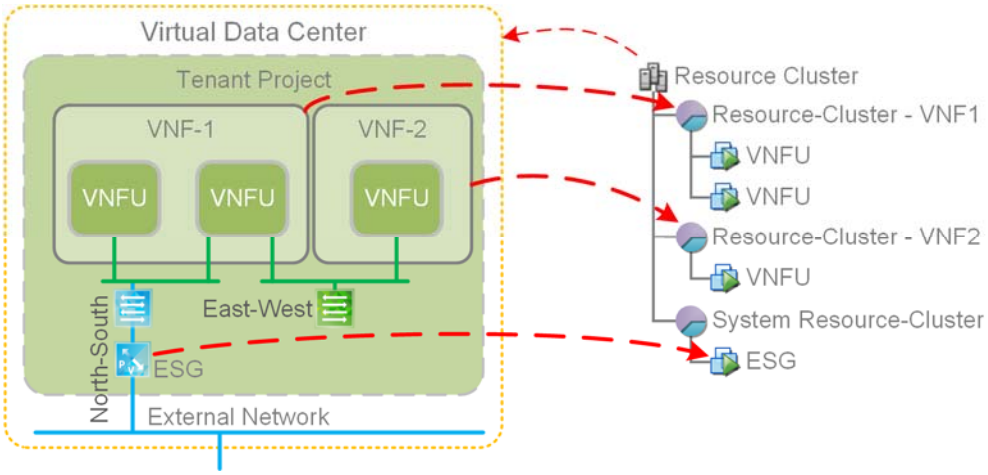

Figure 4 VMware vCloud Director® multitenant networking

The VXLAN Tunnel Endpoint (VTEP) Logical Switches handle the east-west traffic from the VNFs and can span across separate VDS instances over the entire transport zone. The Telco cloud consumes the logical switches by mapping them to vCloud Director external networks. Depending on the VNF network topology, VNFs connect to one or more Org Networks that are in turn connected to the external networks.

North-South traffic flow is implemented by connecting the logical switch to the Distributed Logical Router (DLR) and the NSX Edge for NFV traffic to external network. When stateful services such as firewall, load-balancing, VPN, and NAT are required, an active/standby NSX Edge pair is deployed. When NSX Edge is used solely to provide routing function such as for VNFs, an OSPF ECMP configuration can be deployed to provide more resilience and fault tolerance.

DLR is deployed as an active-standby HA configuration while three NSX Edge devices are deployed to provide routing services, and are configured with ECMP OSPF peering.

Since a DLR and NSX Edge cannot be connected directly to each other, a transit network is used for this purpose. Anti-affinity rules are to be configured such that the DLR active-standby pairs are on separate hosts. Anti-affinity rules are created to keep the NSX Edge devices on separate hosts as well.

Figure 4 shows the vCloud Director networks and logical switches for east-west traffic, and the NSX Edge devices deployed in the edge cluster for dynamic routing of the VNF network for north-south traffic.

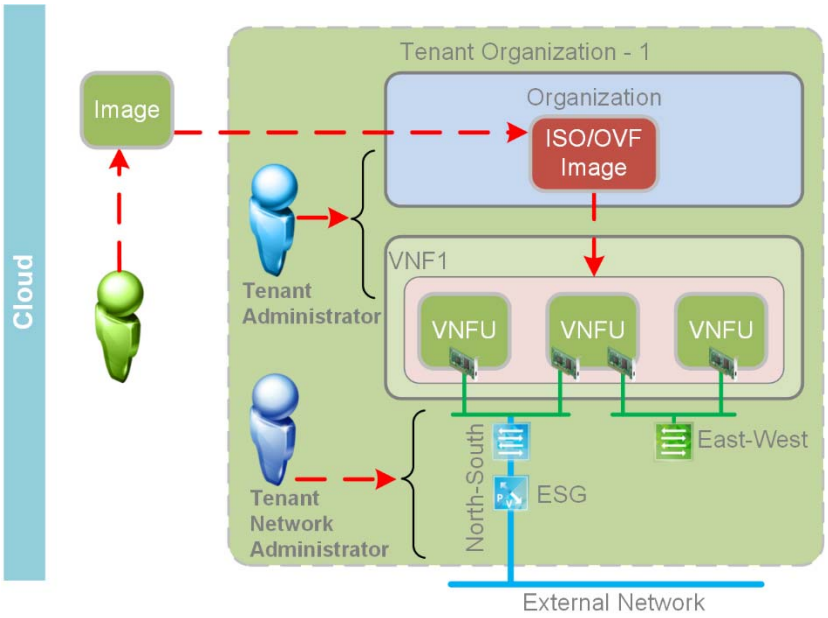

Figure 5 VNF networking

Management VLAN - All of the management nodes local to the site are interconnected using the Management VLAN network across two clusters. The following image shows the management VLAN and the management components that utilize this network:

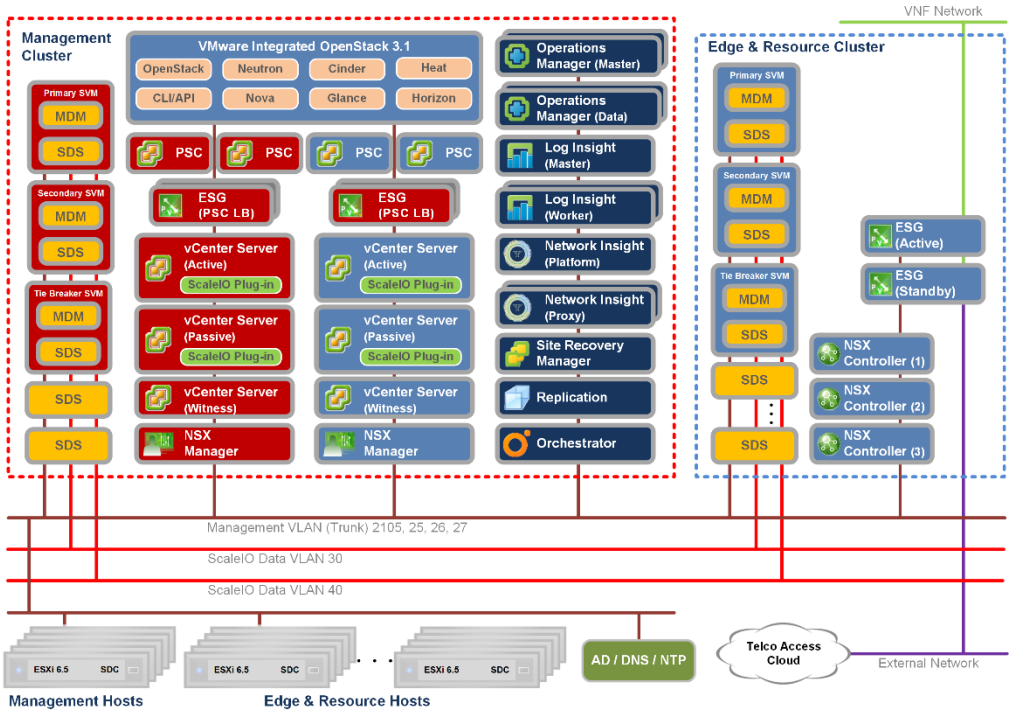

Figure 6 Management VLAN tenant network

2.1.2.5 Datacenter network

The datacenter network is the physical network on which the NFVI support services such as Active Directory for authentication, DNS for name resolution, NTP for time synchronization, and SMTP for email notifications are connected.

These are shared components used by both the management components and the ESXi hosts. This network is to be routed at the physical network with the Management VLAN and ESXI Management networks so that these services can be consumed.

2.1.2.6 Network summary

The following table lists the VLAN IDs, port groups, and their functions for the three clusters. The network names and VLAN ID are for the purpose of describing the network architecture in this document; replace these with the actual datacenter network configuration.

Table 11 VLAN IDs, function, and port groups

VLAN ID	Function	Port group	Network type
Management cluster			
2105	ESXi management	ESXi management	Management
27	vSphere vMotion	vMotion network	Management
30, 40	ScaleIO Virtual SAN	ScaleIO network	Infrastructure
2105	vSphere Replication	Replication network	Management
Edge and resource cluster			
2105	ESXi management	ESXi management	Management
27	vSphere vMotion	vMotion network	Management
30,40	ScaleIO Virtual SAN	ScaleIO network	Infrastructure
50,60,70,80	Tenant management	Management VLAN	Tenant management
1,90	Overlay network	Created when VXLAN software switch is created	Tenant

2.1.3 Storage design

This section discusses the design for shared storage solution based on Dell EMC™ ScaleIO®. The ESXi hosts in all the clusters are connected to a dedicated VLAN for ScaleIO SAN traffic. The ScaleIO components, such as MDM, SDS, and SDC, and an iSCSI target, are installed on dedicated ScaleIO virtual machines (SVMs). The SDS adds the ESXi hosts to the ScaleIO to be used for storage which enables the creation of volumes. Using iSCSI targets, the volumes are exposed to the ESXi via an iSCSI adapter. ScaleIO volumes must be mapped both to the SDC and to iSCSI initiators. This ensures that only authorized ESXs can see the targets. Reliability is enhanced by enabling multipathing, either automatically or manually. Before starting to deploy ScaleIO, ensure that the following prerequisites are satisfied:

- The management network and Virtual Machine Port Group on all the ESXs that are part of the ScaleIO system are be configured
- Devices that are to be added to SDS are free of partitions
- One datastore is created from one of the local devices for all the ESXs - this datastore is needed when deploying SVMs
- ScaleIO supports the following network configuration:
  - There are two or more data networks, each on separate IP subnets
  - Management of ScaleIO MDM/SDS is part of the management network and spans both clusters

Each host in the ScaleIO virtual SAN cluster must be configured with a VMkernel port group and a Virtual Machine port group and enabled for ScaleIO virtual SAN on the infrastructure distributed switch.

ScaleIO components are designed to work with a minimum of five server nodes. When all SDSs in a Protection Domain have one HDD and one SSD drive associated with them, then two storage pools should be defined – High-performance storage pool consisting of SSD drives for latency sensitive workloads and the capacity storage pool consisting of HDD drives for nonsensitive workloads.

VMware recommends that when the disks are connected to a RAID controller, each disk must be configured as a standalone RAID-0.

Two formats of storage can be used with the solution – virtual machine disk (VMDK), or raw device mapping (RDM). This solution recommends using VMDK or VMFS-based storage to use all of the vSphere benefits.

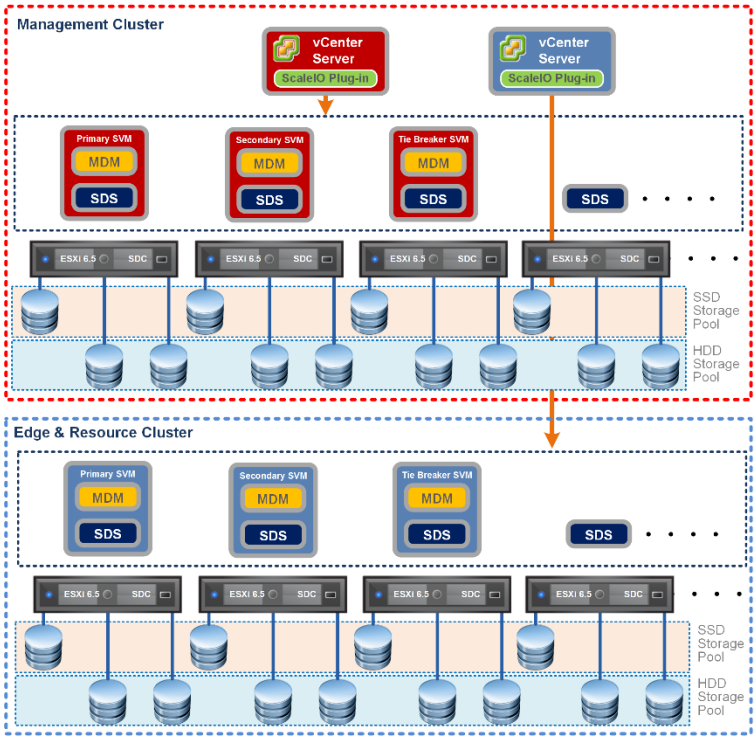

Figure 7 Storage design - management cluster, edge and resource cluster

To configure Scale IO, the following steps need to be performed:

1. Prepare the ScaleIO environment by configuring each ESXi host in the cluster
2. Register the ScaleIO plug-in to the vSphere Web Client
3. Upload the OVA template to the ESXi host
4. Deploy the ScaleIO system from the vSphere Web Client using the ScaleIO plug-in
5. Create volumes with required capacity from the ScaleIO system and map the volumes to the ESXi hosts
6. Create datastores by scanning the ScaleIO LUNs from ESXi hosts
7. Install the ScaleIO GUI to manage the system

See *EMC ScaleIO 2.0 User Guide* for more information.

Note: You must have an EMC Community Network login to access the *Dell EMC ScaleIO 2.0 User Guide*.

VMware recommends enabling the sparse VM swap files feature for efficient utilization of available usable storage capacity. The *SwapThickProvisionDisabled* feature ensures that VM swap files are created as sparse files instead of thick provisioned files. This advanced setting, which is disabled by default, must be set on each ESXi host that is in the VSAN cluster.

2.2 Virtualized Infrastructure Manager Design

2.2.1 VMware vCenter Server

In a two-pod design, the management pod is implemented as a cluster, governed by the first VMware vCenter Server® instance. The use of a cluster allows the components of the pod to benefit from cluster features such as resource management, high availability, and resiliency to form the foundation of a carrier grade virtual infrastructure management. A second vCenter Server is deployed in the Management pod to oversee the edge/resource pod.

Each vCenter Server is a virtual appliance that contains an embedded database. The vCenter Server Appliance (VCSA) is preconfigured, hardened, and fast to deploy. Use of the appliance allows for a simplified design, eases management, and reduces administrative efforts. VCSA availability is ensured using a cluster of three nodes. This consists of one active node that serves client requests, one passive node as backup in the event of failure, and one quorum node referred to as the witness node. Automatic replication between nodes ensures that VCSA data is always synchronized and up-to-date.

The Platform Services Controller (PSC) contains common infrastructure security services such as VMware vCenter® Single Sign-On, VMware Certificate Authority, licensing, and server reservation and certificate management services. The PSC handles identity management for administrators and applications that interact with the vSphere platform. Each pair of PSCs is configured to use a separate vCenter Single Sign-On domain. This approach secures the management components by maintaining administrative separation between the two pods. PSCs are deployed as load balanced appliances external to vCenter Server for high availability. An NSX ESG instance is used as the load balancer between the PSCs and their respective vCenter Servers.

Each vCenter Server instance and its PSC data retention is ensured using the native backup service built into the appliances. This backup is performed to a separate storage system using network protocols such as SFTP, HTTPS, and SCP.

Physical storage devices on ESXi hosts are pooled to logical ScaleIO datastores, which are aggregated across clusters to form a ScaleIO volume for optimum utilization of storage capacity, using tiered pools. Persistent storage for each cluster-VM lives in this abstract storage volume representation for HA considerations. Only exception to this model are persistent storage for MDM/SDS VMs, which need to live in the host's local datastore. All management components are stored in the management volume, while VNF workloads deployed by VMware® Integrated OpenStack are stored in the edge-resource volume. This delineation is essential to meet administrative, performance, and fault-tolerance objectives in the system. The image below shows the vCenter Server instances and their relationship to the two-pod design:

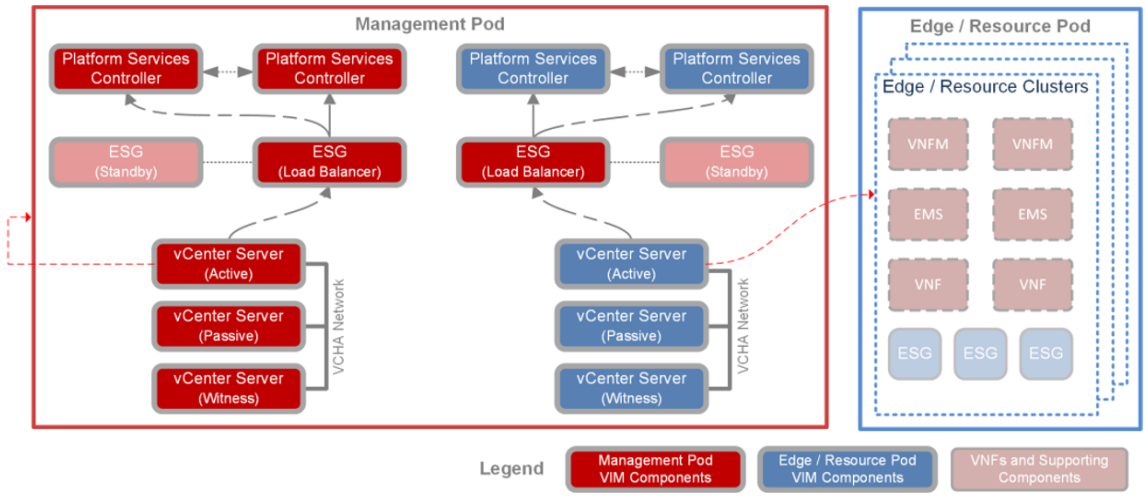

Figure 8 Management cluster VIM Components

VMware recommends using the vCenter Server virtual appliance as it is pre-configured and enables faster deployment. VMware vCloud recommends using external PSC that is not embedded within vCenter Server.

The resource cluster can be sized and scaled independently of the management cluster where the capacity, such as number and size, are more or less fixed. vCenter Server appliance supports up to 2,000 hosts or 25,000 virtual machines at full vCenter Server scale. For additional information, see *Configuration Maximums vSphere 6.5*.

The management cluster *cannot* be scaled up however the edge-resource (ER) clusters may be scaled out to increase capacity, as it is part of separate protection domain. Scaling decisions must be evaluated carefully to ensure there are no virtual or physical infrastructure bottlenecks.

The small deployment size for the management cluster vCenter Server instance supports up to 150 hosts and 3,000 virtual machines. This allows for the handling of the current management cluster workloads and accommodates future scalability. The large ER cluster vCenter Server deployment size supports up to 2,000 hosts and 25,000 virtual machines per vCenter Server instance. For additional information, see *Configuration Maximums vSphere 6.5*.

2.2.2 Virtual networking design using VMware NSX Manager

Each VMware NSX® Manager™ has a 1:1 relationship with VMware vCenter Server®. Therefore, two NSX Managers are created in the management cluster.

The first NSX Manager in the management cluster is solely responsible for the deployment and operation of the highly available ESG instances that provide load balancing functionality, for example, PSC and VCD.

The second NSX Manager in the management cluster is responsible for all ER cluster networking. It is registered with VMware Integrated OpenStack to provide networking services to tenants, including stateful firewalls and load balancers. It is used to configure east-west VNF connectivity, north-south routing, and out-of-band management access for VNFs.

Infrastructure networks, VLAN=30, 40, underlay=vmnic-4, -5, are used for ScaleIO data traffic. Management networks, VLAN=25, 26, 27, 2105, underlay=vmnic-6, -7, are used for VMware vSphere® High Availability, VMware vSphere® vMotion®, and management rendezvous. For each pod, separation between infrastructure and management networks ensures security and provides network resources where needed. This separation is realized by two distributed switches, one for infrastructure networks and the other for management networks. Each distributed switch has separate uplink connectivity to the physical data center network, completely separating its traffic from other network traffic. The uplinks are mapped to a pair of physical NICs on each ESXi host, for optimal performance and resiliency. The following image shows the NSX Manager instances and their components:

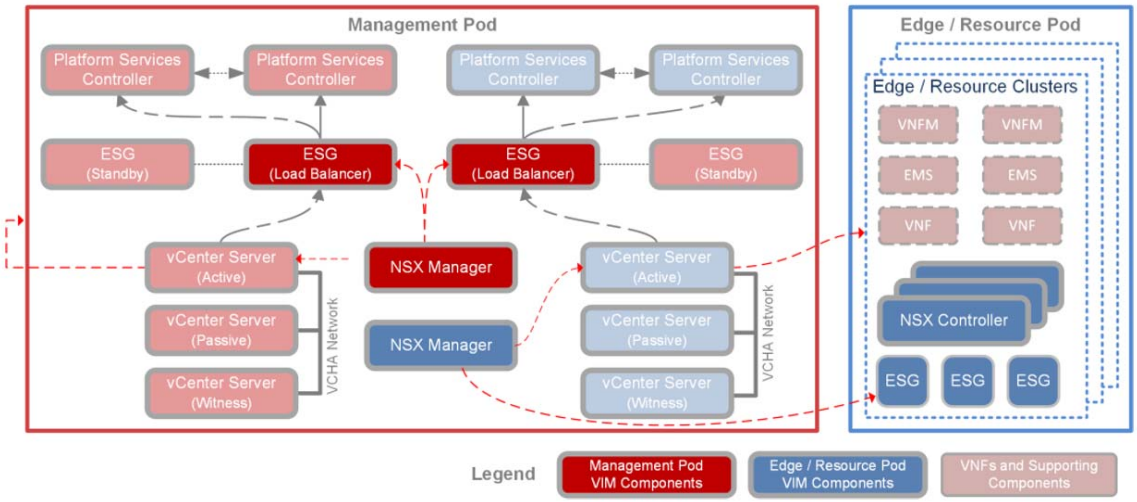

Figure 9 VMware NSX® Manager™ in a two-pod design

2.2.3 VMware Integrated OpenStack

With VMware® Integrated OpenStack, the OpenStack services can be implemented on existing VMware vSphere® implementation. VMware Integrated OpenStack is deployed through the Integrated OpenStack Manager vApp in vCenter. This Integrated OpenStack Manager provides a workflow that guides one through the deployment process, which includes specifying compute infrastructure, and storage and network

configurations. Post-deployment, one can use Integrated OpenStack Manager to modify VNF graph or underlying configurations. VMware Integrated OpenStack 3.x is based on the Mitaka release of OpenStack. See the VMware Integrated OpenStack documentation for more information.

2.2.4 VMware Integrated OpenStack design

The VMware Integrated OpenStack Management Server (OMS) connects to the vCenter Server instance that man-ages the Management pod. OMS uses a virtual machine template to rapidly deploy, administer and perform day 2 management operations of the VMware Integrated OpenStack management plane components deployed in the Management pod. OMS is used to instantiate either the two-pod or three-pod deployment of the VMware Integrated OpenStack management plane.

VMware Integrated OpenStack connects to the second vCenter Server instance that manages the collapsed Edge / Resource pod. This vCenter Server is responsible for storage and compute resources. VMware Integrated Open-Stack is also connected to the NSX Manager instance associated with the Edge / Resource pod networking. Figure 10 illustrates the VMware Integrated OpenStack management components for the two-pod design.

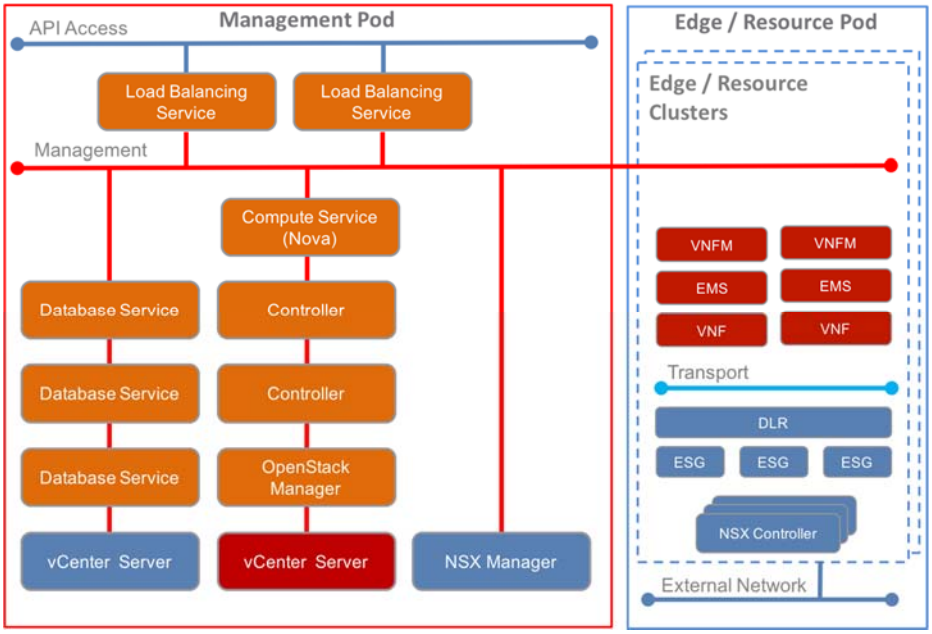

Figure 10 VMware® Integrated OpenStack in two-pod design

The VMware Integrated OpenStack management plane is deployed with redundancy for all the VMware Integrated OpenStack management components with no single point of failure. Even though this requires higher resources to be made available in the Management pod, it offers the best configuration for high availability and is the recommended topology for production environments. OMS can also be used to deploy a compact instance of VMware Integrated OpenStack with a significantly smaller resource requirement, however this topology is not as highly available or scalable as the full deployment.

OMS deploys all the necessary components for a scalable and highly available VMware Integrated OpenStack deployment, this includes clustered databases, controllers and VMware Integrated OpenStack

load balancers. All the management components have connectivity to each other through a dedicated management network. The clustered VMware Integrated OpenStack management components are shown in Figure 11.

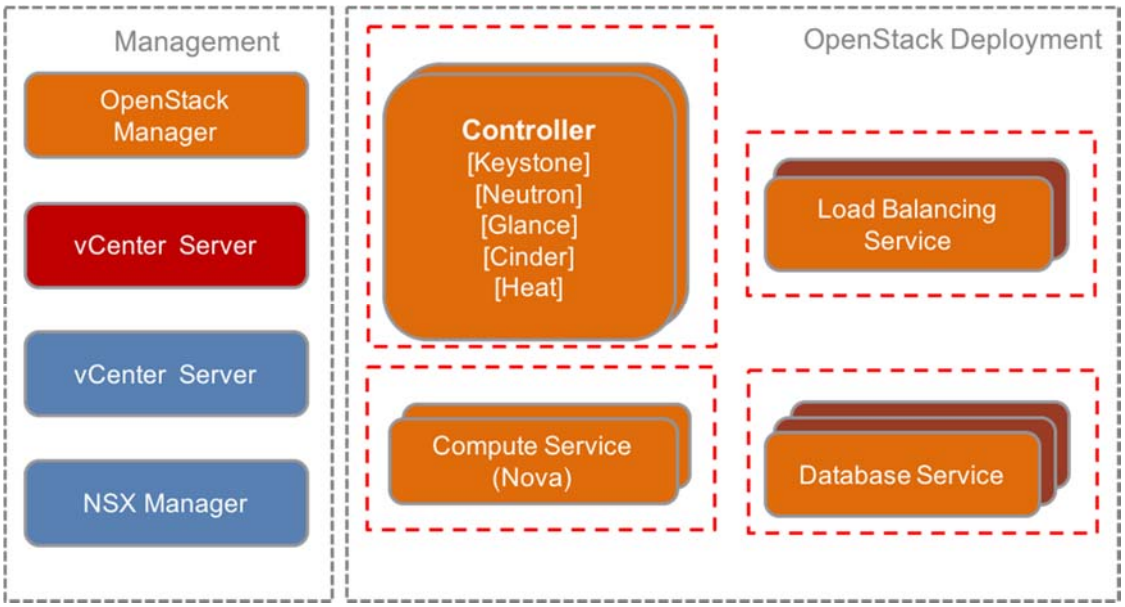

Figure 11 VMware Integrated OpenStack High Availability

VMware Integrated OpenStack is closely integrated with NSX for vSphere, which provides tenants with more features and capabilities for managing their VNF networking needs directly from within the Horizon interface and APIs. All the building blocks for creating secure multitenant VNF networks are in the hands of the tenant. These network services include firewalling, network address translation (NAT), static and dynamic routing, and load balancing. Tenants can provision VXLAN backed logical switches for east-west VNF component connectivity. At the same time, they can deploy NSX ESGs for north-south traffic, as required when connecting to other tenants or to external networks. With this integration, CSPs spend fewer administrative resources configuring and setting up VNFs, reducing the cost of managing the platform.

2.3 Operations management design

2.3.1 VMware vRealize Operations Manager

The VMware vRealize® Operations Manager™ appliance is deployed in a master-replica configuration for high availability. The appliance has all the services required by vRealize Operations Manager hence allows for an architecture that can be scaled easily by adding additional instances.

The appliance deployment size is selected as small with four vCPUs, 16GB RAM, and 84GB HDD storage space. This size assumes a data retention period of six months for 50 VMs, 12 hosts and three datastores. VMware recommends sizing the appliance as per the exact data retention requirements using the *vRealize Operations Manager 6.5 Sizing Guidelines*.

First, a single master node is deployed then a second replica node is deployed to form the cluster. The data is replicated and switch over happens automatically in case the master fails. Anti-affinity rules ensure that the nodes are always deployed on separate hosts. The vRealize Operations Manager has a pair of proxies, each gathering data from a vCenter complex.

The VMware vRealize® Operations Management Pack™ listed under the monitoring section in this document is installed to retrieve various performance and health parameters from the vCloud NFVI platform. If additional management packs are installed, the resource requirements of the vRealize Operations Manager appliance may need to be increased.

2.3.2 VMware vRealize Log Insight

VMware recommends deploying one VMware vRealize® Log Insight™ master node and two worker nodes. This gives the best performance and high availability configuration. The integrated load balancer of the cluster is enabled and used to ensure that load is balanced fairly amongst the available nodes. All the nodes should be deployed on the same Layer2 network and clients should point to the FQDN of the load balancer.

The initial vRealize Log Insight appliance deployment size is kept at default with 132GB of disk space provisioned, 100GB of the disk space is used to store raw data. The vRealize Log Insight appliance should be sized based on the IOPS, syslog connections and events per second. For more details on sizing the appliance, see *Sizing the vRealize Log Insight Virtual Appliance*.

Additional sizing considerations, such as the number of vSphere vCenter servers supported by a single instance of vRealize Log Insight are documented in the *vRealize Log Insight Configuration Limits*.

2.3.3 VMware vRealize Network Insight

The VMware vRealize® Network Insight™ is installed in the management pod of the two-pod design. In an ideal situation, vRealize Network Insight is configured to monitor all networking-related components in the NFVI. Naturally, vRealize Network Insight can connect to the vCloud NFV networking components: VMware vSphere®, and VMware NSX® for vSphere®. It can also be configured to monitor a myriad of physical devices such as Dell switches, Cisco Nexus and Catalyst switches, and Arista, Juniper Networks, Hewlett-Packard Enterprise, Brocade, and Palo Alto Networks switches.

The vRealize Network Insight architecture consists of a platform VM, pair of proxy VMs, and data sources. The role of the platform VM within the architecture is to perform analytics, storage, and to provide a user interface into the data. The proxy VM, or the collector, collects data from sources using various protocols such as HTTPS, SSH, CLI, and SNMP, depending on the source and the configuration. A variety of data sources are supported, including VMware vCenter®, VMware NSX®, firewalls, and various switch vendors. To provide a complete overview of the NFV environment, vRealize Network Insight is connected to the VMware vCenter Server® that operates the edge and resource clusters.

3 Business continuity and disaster recovery using VMware

3.1 VMware vSphere Data Protection

This section of the document covers the backup and recovery of the management components of the VMware vCloud® NFV™ platform. For the purpose of this reference architecture, this document will cover the VMware vSphere® Data Protection™ as the backup solution, however one can use supported third party backup solutions instead.

The vSphere Data Protection appliance is deployed on a separate datastore than the ScaleIO virtual SAN datastore of the protected workloads in the management cluster. The appliance is connected to the management VLAN for communication with the Management vCenter Server. Connectivity through vCenter Server provides vSphere Data Protection with visibility to all VMware ESXi™ servers, and therefore to the virtual machines that must be backed up. The VMware vSphere Web Client interface is used to select, schedule, configure, and manage backups and recoveries of virtual machines.

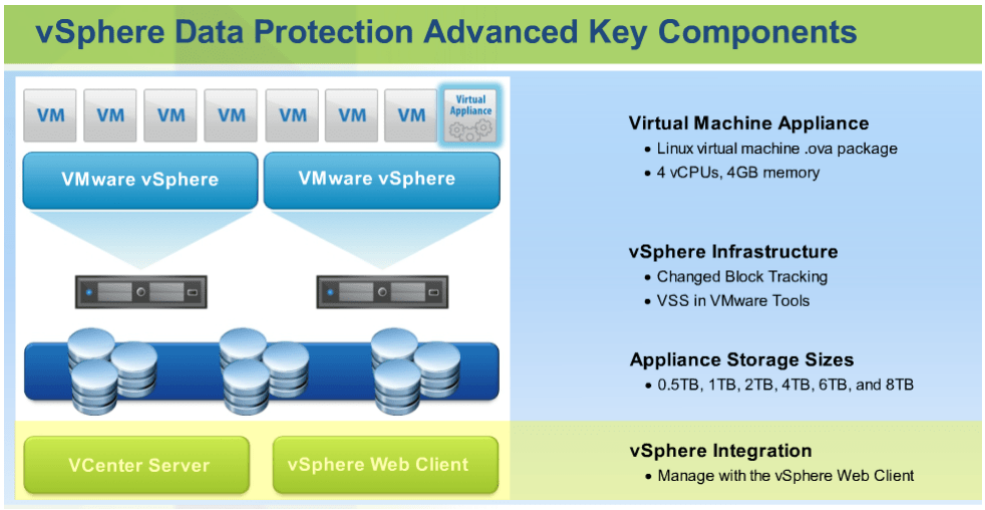

Figure 12 VMware vSphere® Data Protection™

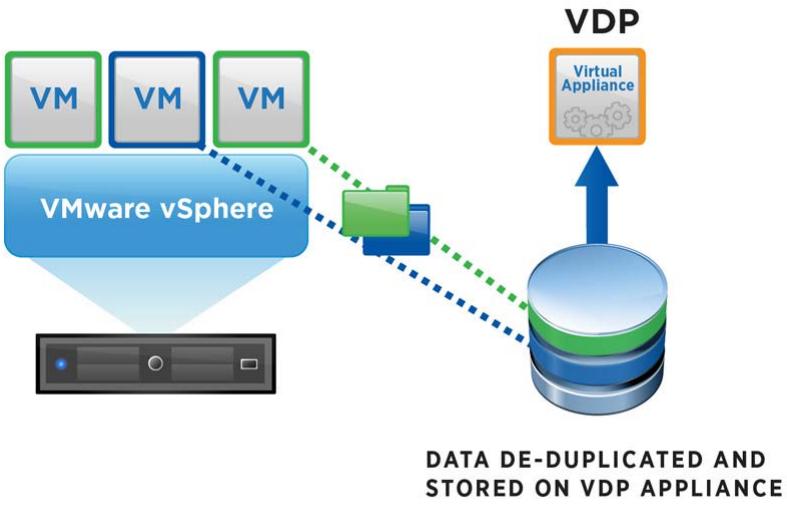

Figure 13 VM backup using VDP

The VDP appliance communicates with the vCenter server to make a snapshot of the .vmdk files within the virtual machine. Deduplication takes place within the appliance by using a variable-length deduplication technology. To increase the efficiency of image level backups, VDP utilizes the Changed Block Tracking (CBT) feature. The CBT feature reduces the backup time of a given virtual machine image, and provides the ability to process many virtual machines within a particular backup window.

3.1.1 Storage design

The backup datastore stores all the production data that is required in a disaster recovery event or data loss to recover the backup up management components based on a recovery point objective (RPO).

It is important to choose the target location and meet the minimum performance requirements to mitigate such a scenario. There are two options when choosing the target storage location.

Option 1: Store backup data on the same ScaleIO virtual SAN datastore

- Simple management with no dependency on storage administrator
- Takes full advantage of VMware vSphere capabilities
- If the destination datastore is unrecoverable, you risk losing the ability to recover your data

Option 2: Store backup data on dedicated storage

- If the ScaleIO virtual SAN storage becomes unavailable, data can be recovered because the backup data is not located on the same shared storage
- Separate management and backup workloads
- The backup schedule does not impact the management cluster storage performance, because the backup storage is separate

VMware vSphere® Data Protection™ generates a significant amount of I/O, especially when performing multiple, concurrent backups. The storage platform must be able to handle this I/O. If the storage does not meet the performance requirements, it is possible for backup failures to occur and for error messages to be generated. VMware recommends using a separate dedicated storage volume for best performance.

3.1.2 Backup policies

VMware recommends using the *HotAdd* transport mechanism for faster backups and restores and less exposure to network routing, firewall and SSL certificate issues when taking image backups of entire virtual machine.

Even when the VMware vSphere® Data Protection™ uses Changed Block Tracking (CBT) technology to optimize the success rate to back up data, it is crucial to avoid any window where the management components storage is in high demand to avoid any business impact. For more information, see *Changed Block Tracking (CBT) on virtual machines*.

The retention policies are the properties of a backup job, therefore it is important to group virtual machines by business priorities and the retention requirements set by the business level. For this reference architecture, vSphere Data Protection will only backup the management components deployed in the management cluster. The section below lists the vCloud NFV management components and their backup strategies.

- VMware ESXi™ hosts – The ESXi hosts are not backed up, instead their configuration data can be exported and imported back on a newly installed server. Alternatively host profiles may be used to restore the configuration of the hosts to their initial configured state. See *How to back up ESXi host configuration* for more information.
- VMware vCenter Server® with the External Platform Services Controller (PSC) – The VMware vCloud® NFV™ platform uses a pair of load balanced platform services controller instances for each vCenter server. The PSC instances are replicated while the vCenter server has an embedded database and points to the PSC load balancer virtual IP. The vCenter server and its corresponding PSCs must be backed up at the same time. If all the components fail at the same time, the PSC must be restored first. vSphere Data Protection is used to take a full image level backup of both the PSCs and vCenter Server. See *Backing up and restoring vCenter Server 6.0 external deployment models* for more information.
- VMware NSX® Manager™ – The VMware NSX® Manager™ has a built in backup and restore mechanism. All of the configuration data can be backed up on a schedule to an FTP server. The NSX Manager backup contains all of the NSX configuration including controllers, logical switching and routing entities, security, firewall rules, and everything else that you configure within the NSX Manager UI or API. The vCenter database and related elements like the virtual switches need to be backed up separately. See *NSX Backup and Restore* for more information.
- VMware vRealize® Operations Manager™ – With the VMware vRealize Operations Manager™, single or multi-node clusters can be backed up and restored using VMware vSphere® Data Protection™ or other backup tools. You can perform full, differential, and incremental backups and restores of virtual machines. All nodes need to be backed up and restored at the same time.
- VMware vRealize® Log Insight™ – The VMware vCloud® NFV™ platform utilizes a 3-node VMware vRealize® Log Insight™ cluster. The entire cluster needs to be backed up and restored at the same time. See *VMware vRealize Log Insight Administration Guide* for more information.
- VMware vRealize® Network Insight™ – The VMware vCloud® NFV™ platform utilizes a 3-node vRealize Network Insight cluster and 1-node vRealize Network Insight proxy.
- VMware vSphere® Replication™ – The VMware vSphere® Replication™ appliance is backed up using VMware vSphere® Data Protection™ using an image level backup of the entire appliance. When an image is restored and the appliance powered on, the data replication resumes after a few minutes.
- VMware Site Recovery Manager™ – The Site Recovery Manager™ instance is deployed on a Windows machine along with an embedded database where all the configuration information is

stored. This database can be backed up and restored. See *Back Up and Restore the Embedded vPostgres Database* for more information.

- VMware vSphere® Data Protection™ - The VMware vSphere® Data Protection™ appliance has a checkpoint and rollback mechanism built in. By default, the vSphere Data Protection appliance keeps two system checkpoints. If you roll back to a checkpoint, all backups and configuration changes taken since the checkpoint was taken are lost when the rollback is completed. The first checkpoint is created when vSphere Data Protection is installed. Subsequent checkpoints are created by the maintenance service. This service is disabled for the first 24 to 48 hours of vSphere Data Protection operation. See *vSphere Data Protection Administration Guide* for more information.

3.1.3 Monitoring

CPU, memory, network, and disk performance and capacity will be monitored by VMware vRealize® Operations Manager™. Events and log information are sent to VMware vRealize® Log Insight™. Capacity can also be viewed via the Reports tab/vSphere Data Protection capacity.

3.2 VMware Site Recovery Manager

The VMware Site Recovery Manager™ is the industry-leading disaster recovery management solution. Site Recovery Manager offers automated orchestration and non-disruptive testing of centralized recovery plans for all virtualized applications. It can integrate natively with VMware vSphere® Replication™ and support a broad range of array-based replication products available by all major VMware storage partners.

Site Recovery Manager can integrate natively with VMware vSphere® Replication™ and support a broad set of array-based replication products available by all major VMware storage partners. A deployment founded on vSphere and complemented with Site Recovery Manager dramatically lowers the cost of DR through management and testing automation that eliminates the complexity of legacy processes, while ensuring fast and highly predictable recovery time objectives (RTO) to maintain business continuity.

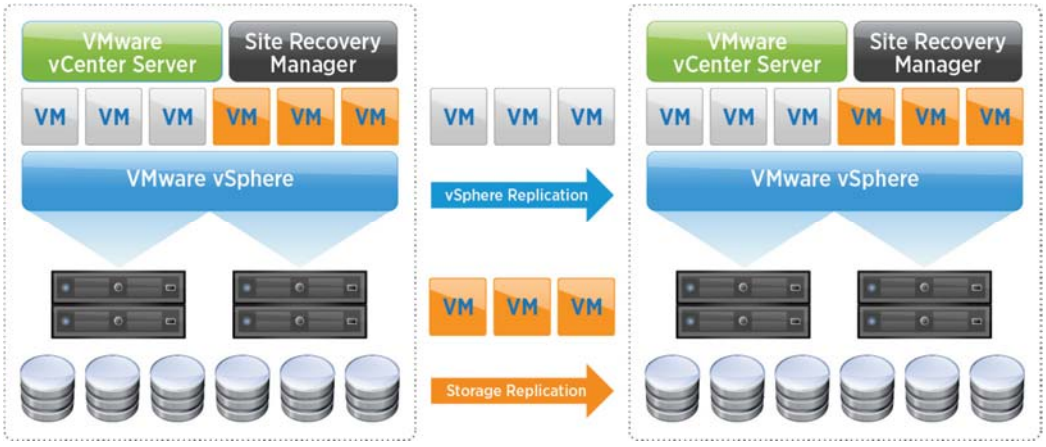

Figure 14 VMware Site Recovery Manager™

The Site Recovery Manager automates the failover and migration of virtual machines to a secondary site. The Site Recovery Manager relies on either vSphere Replication or a storage-based replication product to replicate virtual machines to the recovery site.

3.2.1 How does it work

- VMware vCenter Server™ - VMware Site Recovery Manager™ instances deployed at both production and recovery sites integrate directly with local vCenter Server instances.
- Replication - Site Recovery Manager requires an underlying replication technology to copy virtual machine (VM) data to the recovery site.
- VMware vSphere® Replication™ - VMware's hypervisor-based replication technology enables VM-centric, storage-independent replication with customizable recovery point objective (RPO) and multiple point-in-time recovery. vSphere Replication is included with most vSphere editions at no additional cost.
- Array-based replication - Site Recovery Manager™ integrates with third-party storage array-based replication products through a Storage Replication Adapter (SRA). See *VMware Compatibility Guide* for more details.

3.2.2 Key benefits

Traditional DR solutions often fail to meet business requirements because they are too expensive, complex and unreliable. Organizations using VMware Site Recovery Manager™ ensure highly predictable RTOs at a much lower cost and level of complexity.

- Lower cost for DR – Site Recovery Manager reduces the operating overhead by 50 percent by replacing complex manual runbooks with simple, automated recovery plans that can be tested without disruption. For organizations with an RPO of 15 minutes or higher, VMware vSphere® Replication™ can eliminate up to $10,000 per TB of protected data with storage-based technologies. The combined solution can save over USD $1,100 per protected virtual machine per year. These calculations were validated by a third-party global research firm. Integration with Virtual SAN reduces the DR footprint through hyper-converged, software-defined storage that runs on any standard x86 platform. Virtual SAN can decrease the total cost of ownership for recovery storage by 50 percent.
- Hardware and app independence - The combination of VMware vSphere®, Virtual SAN and vSphere® Replication™ provides a DR infrastructure that is completely hardware-independent at the compute and storage layers. Site Recovery Manager offers DR automation for any vSphere VM.
- Choice for replication - Companies can use vSphere Replication to eliminate storage lock-in and simplify data management with a VM-centric approach, or array-based technologies to leverage current storage investments or ensure zero data loss through synchronous replication.
- Simplified setup - Setting up a recovery plan can be done in a matter of minutes, instead of the weeks as required with manual runbooks. Ongoing DR provisioning to new VMs can be tenant driven through predefined policies via blueprints in vRealize Automation.
- Peace of mind - Recovery plans can be tested as frequently as required without disrupting production systems through automation. A detailed report of the testing outcomes, including RTO achieved, delivers confidence that DR objectives were met and provides a way to demonstrate compliance with regulatory requirements.
- Automated execution - Every DR workflow is automated to minimize RTOs and eliminate errors from manual processes. Automation also enables different use cases. The failover workflow enables disaster recovery with an emphasis on minimizing recovery time. The planned migration workflow enables proactive disaster avoidance and data center mobility without data loss in an application-consistent state. The failback workflow enables bi-directional migrations with ease.
- DR to the cloud services - Companies that don't have the financial resources to invest in a secondary site can use the cloud of a service provider as recovery infrastructure, for example, Amazon. VMware

has built an ecosystem of service providers that offer cloud-based DR services powered by Site Recovery Manager™.

- Find a provider - For organizations looking for a DR Service fully delivered and supported by VMware, the company offers VMware vCloud® Air™ Disaster Recovery.

3.3 VMware vSphere Replication

3.3.1 Introduction

VMware vSphere® Replication™ is a virtual machine data protection and disaster recovery solution. It is fully integrated with VMware vCenter Server® and VMware vSphere® Web Client, providing host-based, asynchronous replication of virtual machines. vSphere Replication is a proprietary replication engine developed by VMware that is included with VMware vSphere Essentials Plus Kit and higher editions of VMware vSphere, VMware vSphere® with Operations Management™ editions, and VMware vCloud Suite® editions.

3.3.2 Architecture overview

The VMware vSphere® Replication™ 6.5 requires VMware vCenter Server® 6.5, either the Microsoft® Windows® implementation or the Linux-based VMware vCenter Server® Appliance™. VMware vCenter® Single Sign-On is also required. If using vSphere Replication with VMware Site Recovery Manager™, the versions of the two must be the same. For example, VMware vSphere® Replication™ 6.5 is the only version of vSphere Replication supported with VMware Site Recovery Manager™ 6.5. For complete details on VMware feature and product interoperability, see the *VMware Compatibility Guide*.

vSphere Replication is deployed as one or more prebuilt, Linux-based virtual appliances. A maximum of 10 vSphere Replication appliances can be deployed per vCenter Server. Each appliance is deployed with 4GB of memory and either two virtual CPUs for small environments, or four virtual CPUs. A vSphere Replication virtual appliance is configured with two virtual machine disk (VMDK) files totaling 18GB in size. Because vSphere Replication is host-based replication, it is independent of the underlying storage and it works with a variety of storage types including vSAN, traditional SAN, NAS, and direct-attached storage (DAS). Unlike many array replication solutions, vSphere Replication enables virtual machine replication between heterogeneous storage types. For example, vSAN to DAS, SAN to NAS, and SAN to vSAN. vSphere Replication can, of course, replicate virtual machines between the same types of storage, such as vSAN to vSAN. vSphere Replication can also serve as the replication engine for VMware Site Recovery Manager.

In this scenario, vSphere Replication virtual appliances are deployed at both the source and target locations, as with VMware Site Recovery Manager. Replication is configured on a per-virtual machine basis that enables fine control and selection of the virtual machines that are included in VMware Server Site Recovery Manager protection groups and recovery plans. The use of VMware Site Recovery Manager is to protect virtual machines running on vSAN requires vSphere Replication.

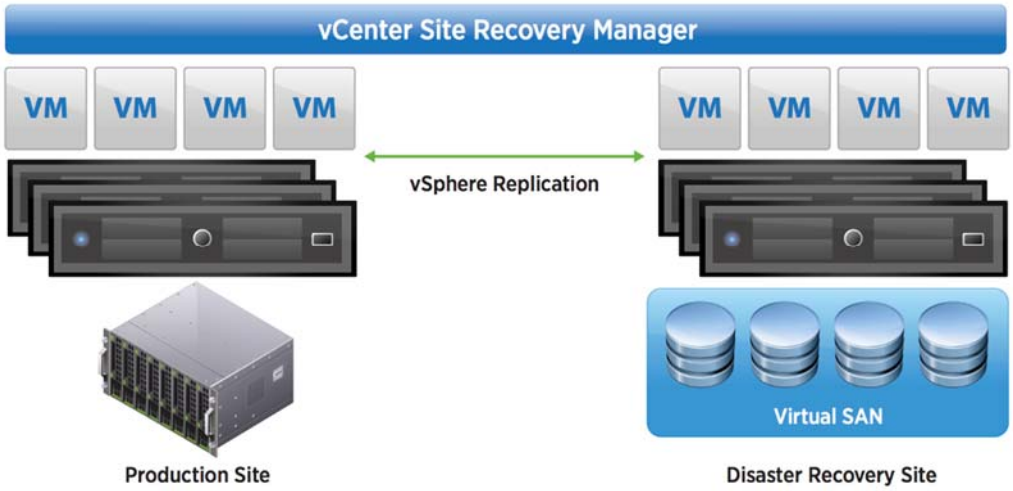

Figure 15 VMware vSphere® Replication™ with VMware® vCenter Site Recovery Manager™ and Virtual SAN

3.4 Multi-site recovery in VMware Site Recover Manager

With the VMware® Site Recovery Manager™, you can connect multiple protected sites to a single recovery site. The virtual machines on the protected sites recover to the same recovery site. This configuration is known as a shared recovery site, a many-to-one, or an N:1 configuration.

The standard one-to-one Site Recovery Manager configuration protects a specific instance of VMware vCenter Server® by pairing it with another vCenter Server instance. The first vCenter Server instance, the protected site, recovers virtual machines to the second vCenter Server instance, the recovery site.

Another example is to have multiple protected sites configured to recover to a single, shared recovery site. For example, an organization can provide a single recovery site with which multiple protected sites for remote field offices can connect. Another example for a shared recovery site is for a service provider that offers business continuity services to multiple customers.

In a shared recovery site configuration, install one Site Recovery Manager Server instance on each protected site, each of which connects to a different vCenter Server instance. On the recovery site, install multiple Site Recovery Manager Server instances to pair with each Site Recovery Manager Server instance on the protected sites. All of the Site Recovery Manager Server instances on the shared recovery site connect to a single vCenter Server instance.

Note: Each Site Recovery Manager Server instance in a pair must have the same Site Recovery Manager extension ID, which you can set when you install Site Recovery Manager Server. Consider the owner of a Site Recovery Manager Server pair to be a customer of the shared recovery site.

You can convert an existing one-to-one configuration of Site Recovery Manager into a shared recovery site configuration. To convert a one-to-one configuration to a shared recovery site configuration, you deploy additional Site Recovery Manager Server and vCenter Server instances as protected sites, and pair them with additional Site Recovery Manager Server instances that all connect to the existing vCenter Server instance on

the recovery site. Each pair of Site Recovery Manager Server instances in the shared recovery site configuration must use a different Site Recovery Manager extension ID. For example, if you installed a one-to-one configuration that uses the default Site Recovery Manager Extension ID, you must deploy all subsequent Site Recovery Manager Server pairs with different custom extension IDs.

One can use either array-based replication or vSphere Replication or a combination of both when you configure Site Recovery Manager Server to use a shared recovery site. In addition to the shared recovery site configuration, Site Recovery Manager also allows and supports shared protected site (1:N) and many-to-many (N:N) configurations.

3.4.1 Using VMware Site Recovery Manager with multiple protected sites and shared recovery site

An organization has two field offices and a head office. Each of the field offices is a protected site. The head office acts as the recovery site for both of the field offices. Each field office has a Site Recovery Manager Server instance and a vCenter Server instance. The head office has two Site Recovery Manager Server instances, each of which is paired with a Site Recovery Manager Server instance in one of the field offices. Both of the Site Recovery Manager Server instances at the head office extend a single vCenter Server instance.

- Field office 1
  - Site Recovery Manager Server A
  - vCenter Server A
- Field office 2
  - Site Recovery Manager Server B
  - vCenter Server B
- Head office
  - Site Recovery Manager Server C, that is paired with Site Recovery Manager Server A
  - Site Recovery Manager Server D, that is paired with Site Recovery Manager Server B
  - vCenter Server C, that is extended by Site Recovery Manager Server C and Site Recovery Manager Server D

The following is an example of using Site Recovery Manager in a shared recovery site configuration:

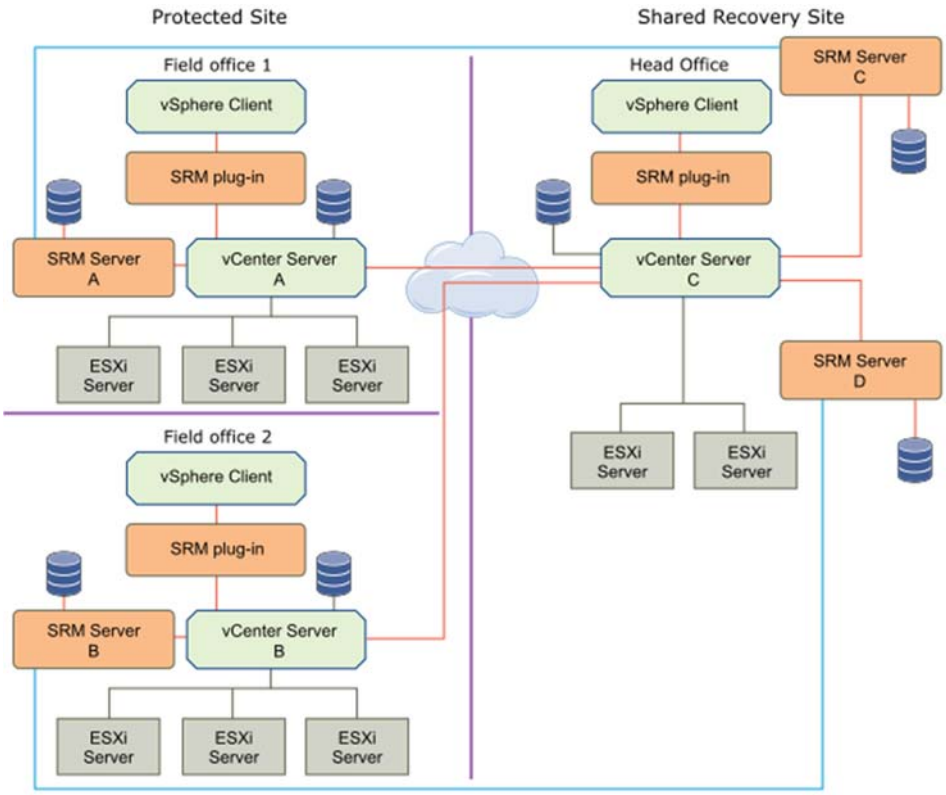

Figure 16 Example of multi-site replication

- Limitations of using VMware® Site Recovery Manager™ in shared recovery site configuration - When the Site Recovery Manager is configured to use a shared recovery site, Site Recovery Manager supports the same operations as it does in a standard one-to-one configuration. Using Site Recovery Manager with a shared recovery site is subject to some limitations.
- VMware® Site Recovery Manager™ licenses in a shared recovery site configuration - Licenses can be assigned individually on the shared recovery site if the Site Recovery Manager is configured to be used with a shared recovery site. You can also share a license between all Site Recovery Manager Server instances on the shared recovery site.
- Install VMware® Site Recovery Manager™ In a shared recovery site configuration - To install Site Recovery Manager in a shared recovery site configuration, you deploy Site Recovery Manager Server on one or more protected sites, and deploy a corresponding number of Site Recovery Manager Server instances on the shared recovery site.
- Use array-based replication in a shared recovery site configuration - You can use array-based replication with VMware® Site Recovery Manager™ in a shared recovery site configuration in the same way as you do in a standard one-to-one configuration.
- Use VMware vSphere® Replication™ in a shared recovery site configuration - You can use vSphere Replication with VMware® Site Recovery Manager™ in a shared recovery site configuration in the same way that you do in a standard one-to-one configuration.
- Upgrade VMware® Site Recovery Manager™ in a shared recovery site configuration - You can upgrade existing Site Recovery Manager Installations that use a shared recovery site.

4 Capacity planning and sizing

4.1 Sizing guidelines

4.1.1 Cluster sizing

This section lists the cluster configuration for the management and edge resource clusters. The sizing considerations for each of the management components already explained in the design sections needs to be factored into the calculations.

- Management Cluster – The management cluster sizing is based on the number of management nodes in the cluster. This has to take into account factors such as multi-node topologies of management components such as components with external databases, and internal or external Platform Services Controller™ location. The management cluster capacity needs to be proportionately increased where components have high availability requirements, such as active-active NSX controllers, active-passive NSX Edge, and HA pairs.

  Additionally, VMware vCenter Server® uses admission control to ensure that sufficient resources are available in a cluster to provide failover protection and to ensure that virtual machine resource reservations are respected. Admission control imposes constraints on resource usage. Violation of the resource usage constraints prevents virtual machines from being powered on.

  Other factors such as allowed host failures, and number of hosts allowed to be taken offline for maintenance, need to be factored when planning for capacity. It is recommended to have all the hosts in the cluster with the same specification and preferably same vendor/model/family. This results in a balanced cluster, with a balanced load across the cluster in a predictable way.

  Resource overheads of virtualization required by the hypervisor need to be taken into account. As a best practice, it is recommended that the design architecture not have resources that are overcommitted for the management nodes.

- Edge and resource cluster – The edge and resource cluster is used to deploy the VNFs. Since it is impossible to calculate the VNF sizes before knowing which VNFs will be deployed, this document suggests an approach to arrive at the resource cluster size once the VNF size is known. VNF vendors are expected to provide the sizing requirements for their VNFs.

  This guideline aims to build a scalable architecture that can be scaled at all tiers as additional VNFs are deployed without compromising the functional aspects of the platform. If the VNF sizing requirements are not known, tools such as VMware Capacity Planner™ can be used to analyze the Telco workloads and make a suitable sizing recommendation.

  The VMware vCloud® NFV™ platform uses an elastic configuration in VMware® Integrated OpenStack, which means that additional ESXi host clusters could be added when there is a need to scale the cluster's capacity. This is a key business benefit as it allows one to optimize TCO/ROI and gradually scale investment over a period of time.

The number of hosts in the cluster also needs to take into account operational aspects of the workloads for example when a host is placed in maintenance mode. The edge is used to host the NSX network infrastructure for the north-south VNF networks. The NSX components deployed in the edge are:

- VMware NSX® Controller™ – From a high availability point of view, three NSX Controllers are deployed together and capacity for all three needs to be accounted for in the clusters in which they are deployed. Note that the NSX Controller nodes are deployed by the VMware® NSX Manager™ and their size is not customizable.
- VMware NSX® Edge™ - NSX Edge instances are deployed in an active-active configuration with ECMP OSPF peering for high availability. Multiple NSX Edge instances can be deployed for performance. For stateful services such as load balancing, NSX Edge is deployed as an active-standby pair.

The number of NSX devices in the edge cluster are determined by the expected north-south VNF network traffic. This document provides guidelines for one set of such devices; these can be scaled as per the requirements of the VNF workloads.

4.1.2 Storage sizing

When sizing the ScaleIO virtual SAN datastore of each cluster, identify the workloads in each cluster and add their disk requirements. Add a buffer to the overall capacity for operational tasks such as for VM templates, cloning, swap files, log files, and snapshots. These will depend on the specific VNF configuration and requirements during VNF onboarding.

The minimum configuration required for ScaleIO virtual SAN is three ESXi hosts. However, this smallest environment has important restrictions. In a 3-node MDM cluster - Master MDM, Slave MDM, and Tiebreaker, there are two copies of the repository and can withstand one MDM cluster failure. However, to avoid a single point of failure a 5-node MDM cluster – Master MDM, two Slave MDM, and two Tiebreaker, has three copies of the repository and can withstand two MDM cluster failure.

The ScaleIO virtual SAN cluster can be configured to tolerate a number of host failures. This configuration makes a mirror copy of the data in the datastore resulting in reduced usable capacity.

For the purpose of this reference architecture, VMware recommends four nodes in each cluster. This ensures that there are enough nodes to meet the availability requirements and allows for a rebuild of the components after a failure. For detailed guidelines refer to the *ScaleIO User Guide* and *ScaleIO Installation Guide*.

As a best practice, VMware recommends having dedicated storage for backups. The storage size depends on the number of virtual machines to back up, the backup frequency, retention policy, and amount of blocks changed since the last backup.

4.2 Sizing design

4.2.1 Management cluster

Table 12 lists the management components along with the clustered nodes that are deployed in the management cluster:

Table 12 Management cluster sizing

System	VCPU	Memory	Storage	Description
Platform Services Controller	4	8GB	30 GB	For management cluster vCenter Server
Platform Services Controller	4	8GB	30 GB	For management cluster vCenter Server
Platform Services Controller	4	8GB	30 GB	For resource cluster vCenter Server
Platform Services Controller	4	8 GB	30 GB	For resource cluster vCenter Server
vCenter Server X 3 (small)	4	16 GB	82 GB	Management cluster
vCenter Server X 3 (large)	16	32 GB	82 GB	Resource and edge cluster
NSX Manager	4	12 GB	60 GB	For management cluster workloads
NSX Manager	4	12 GB	60 GB	For edge and resource cluster workloads
Edge Service Gateway (active) (Large)	6	8 GB	4.5 GB	Active Load Balancer for VCD cells
Edge Service Gateway (standby) (Large)	6	8 GB	4.5 GB	Standby Load Balancer for VCD cells
Edge Service Gateway (active) (Large)	6	8 GB	4.5 GB	Active Load Balancer for management and resource PSCs
Edge Service Gateway (standby) (Large)	6	8 GB	4.5 GB	Standby Load Balancer for management and resource PSCs
Log Insight (master)	4	8 GB	150 GB	Deployed in management cluster
Log Insight (replica)	4	8 GB	150GB	Deployed in management cluster
Log Insight (replica)	4	8 GB	150GB	Deployed in management cluster
Operations Manager (active)	4	16 GB	250 GB	Deployed in management cluster
Operations Manager (passive)	4	16 GB	250 GB	Deployed in management cluster
VMware® Integrated OpenStack	56	192GB	565GB	Deployed in management cluster
NFS Server	2	4 GB	10 GB	vCloud Director shared transfer space
vSphere Data Protection	2	4 GB	873 GB	Space includes backup data storage space
vRealize Network Insight				
vRealize Orchestrator				
Primary MDM – Mgmt	1	1 GB	32 GB	SVM deployed in management cluster
Secondary MDM – Mgmt	1	1 GB	32 GB	SVM deployed in management cluster

Tiebreaker - Mgmt	1	1 GB	32 GB	SVM deployed in management cluster
SVM – SDS	1	1 GB	32 GB	SVM deployed on edge cluster
Total	144	396 GB	3.29 TB	

4.2.2 Edge and resource cluster

Considering that VNF workloads are network intensive, deploy the X-Large configuration for the Edge Service Gateway. Three of these are deployed with ECMP configuration for high availability.

Table 13 lists the VMware NSX® Edge™ devices deployed in the edge cluster. These should cater to most VNFs however, should be sized appropriately as per the VNF workload requirements in the resource cluster. As VNF workload requirements increase, both the edge cluster capacity and the NSX Edge devices deployed in it can be scaled.

The sizing requirements for the VNFs will be captured as part of the VNF onboarding process. NEPs are expected to provide all the infrastructure requirement details of a VNF so that resource requirements can be planned and sized as required. The VMware vCloud® NFV™ platform allows scaling the resource cluster to accommodate VNF capacity requirements.

Table 13 Edge and resource cluster sizing

System	VCPU	Memory	Storage	Description
VMware NSX Controller 1	4	4 GB	20 GB	For resource cluster workloads
VMware NSX Controller 2	4	4 GB	20 GB	For resource cluster workloads
VMware NSX Controller 3	4	4 GB	20 GB	For resource cluster workloads
Edge Service Gateway (active) X-Large	6	8GB	4.5GB	For resource cluster workloads (ECMP)
Edge Service Gateway (active) X-Large	6	8 GB	4.5 GB	For resource cluster workloads (ECMP)
Edge Service Gateway (active) X-Large	6	8 GB	4.5 GB	For resource cluster workloads (ECMP)
Primary MDM – Res/Edge	1	1 GB	32 GB	SVM deployed in management cluster
Secondary MDM – Res/Edge	1	1 GB	32 GB	SVM deployed in management cluster

Tiebreaker – Res/Edge	1	1 GB	32 GB	SVM deployed in management cluster
SVM – SDS	1	1 GB	32 GB	SVM deployed on edge cluster
SVM – SDS	1	1 GB	32 GB	SVM deployed on edge cluster
Total	35	41 GB	233.5 GB	

5 VNF onboarding

This section lists the key considerations when onboarding VNFs on to the VMware vCloud® NFV™ platform. Since each VNF has its own specific requirements, these guidelines must be adjusted as required.

5.1 Capacity requirements

The first step is to identify the capacity requirements of the VNF. The target NFVI needs to have sufficient capacity for the successful deployment and operation of the VNF. Some of the points to be taken into account when planning for the NFVI capacity required for hosting a VNF are:

- Number of active nodes of the VNF
- Nodes required for high availability and redundancy
- VNFs deployed for scaling for transient and seasonal peaks
- Capacity reservation for failover nodes
- Capacity required for operational overheads
- Capacity reserved for future growth

NEPs are expected to provide the above information for their VNFs before actual onboarding so that capacity can be planned and provided for in advance.

5.2 Resource requirements

Once the capacity in terms of number of nodes has been identified, the resource requirements of these nodes need to be assessed. Each category of node may have its own resource requirement e.g. nodes used for scaling may be sized differently than the primary nodes.

In addition to the resource requirements of the VNFs, the VMware vCloud® NFV™ platform needs to be configured for the efficient allocation of resources. Resource allocation is done by configuring VMware® Integrated OpenStack to aggregate hardware resources of the NFVI. VMware Integrated OpenStack is then used to allocate the pooled resources to the VNFs.

5.3 Operational requirements

When onboarding VNFs, their operational aspects must be considered in terms of capacity and performance requirements. Some of the operational tasks that may impact the capacity at the time of VNF onboarding are:

- Space for temporary machine images such as snapshots
- Space for maintenance and backup operations
- Log retention policies
- Scale up and scale out for VNF nodes
- VNF templates in catalog

A key operation is the deployment of VNFs from catalog, which can either deploy a new VNF instance or scale an existing VNF. VMware® Integrated OpenStack allows users to create catalogs and vApp templates in the catalog. A vApp is a group of related VMs along with all configuration information required by the application deployed inside the vApp. Deploying a vApp template from catalog is the fastest way to deploy one or more VNFs. This can be automated using the VMware Integrated OpenStack APIs.

VNF Managers make use of the VMware Integrated OpenStack APIs to dynamically scale the VNFs by deploying additional VNFs from the catalog. The catalog can contain a vApp template for an entire VNF stack or vApps with individual VNF components. This gives the flexibility to scale specific nodes of the VNF instead of deploying the entire stack.

5.4 High availability requirements

The high availability requirement of the deployed VNF workloads needs to be factored when calculating capacity requirements and capacity availability of the NFVI platform. The size, quantity, and type (active/passive) of the redundant nodes influences storage, network and compute resource consumption.

VNF Workloads may have anti affinity requirements such as ensuring that nodes of a VMware vSphere® High Availability group are never placed on the same physical server. This puts additional constraints on resource management.

A new feature of VMware® Integrated OpenStack is the ability of tenants to set affinity/anti-affinity rules for VNFs deployed in their VMware vCloud® organization. As a result, service providers need to closely monitor the resource utilization and ensure both high availability requirements and resource requirements of VNFs are met.

For disaster recovery, reserve sufficient resources that allow VNFs to be powered on. This is done by ensuring that sufficient resource reservations are in place at both the edge and resource cluster, and in the VMware Integrated OpenStack.

5.5 Security requirements

When VNFs are onboarded, necessary security requirements need to be identified. Resource partitioning can be addressed by separating the VNFs in VMware® Integrated OpenStack by placing them in separate containers. For network security, the vCloud NFV platform supports network configurations such as the creation of firewall rules for both east-west and north-south traffic. Firewall rules can be configured at both the perimeter or internally by micro-segmentation using NSX. Depending on the network requirements, isolated networks may be created for VNF interconnects and also segregation of workload traffic based on role, such as management traffic and data traffic.

5.6 Network requirements

The most important factors when onboarding a VNF is its networking requirements. The VMware® vCloud NFV™ platform is designed to be flexible to meet the network needs of the VNFs being deployed. Some of the network considerations when onboarding a VNF are:

- Network routing between nodes

- External or MPLS network connectivity
- Network isolation and segregation requirements
- Firewall configuration
- WAN or VPN connectivity
- Connectivity to management systems
- Latency and bandwidth requirements

5.7 VMware Tools requirements

The VMware Tools™ suite of utilities installed in the VNF enhances the performance of the virtual machine's guest operating system, and improves management of the virtual machine. Without VMware Tools, guest performance lacks important functionality. VMware recommends installing VMware Tools to eliminate or improve these issues:

- Ability to take silent snapshots of the VNF
- Synchronize guest OS time with ESXi host
- Support for guest-bound calls created with the VMware VIX API
- Device drivers to optimize mouse operations and improve sound, graphics, and networking performance
- Guest OS customization support from within VMware® Integrated OpenStack
- Scripting that helps automate guest operating system operations
- VMware vRealize® Operations Manager™ end point operations

Although the guest operating system can run without VMware Tools, many of the VMware features cannot be accessed or used until VMware Tools is installed. For example, the shutdown or restart options are not available from the toolbar unless VMware Tools is installed. The only options available for use are the power options.

Note: For an overview of VMware Tools, see *Overview of VMware Tools (340) | VMware KB*.

5.8 Onboarding process

The following is a sequence of important steps involved in VNF onboarding process:

- Setup connectivity to external networks
- Deploy or configure NSX Edge for north-south traffic
- Configure transit networks
- Deploy or configure logical routers for north-south traffic
- Create logical switch for east-west traffic
- Create catalog for the organization

- Set OVF environment parameters, if any
- Import the VNF into the catalog
- Review VNF sizing and network parameters in the catalog
- Configure VMware® Integrated OpenStack and allocate resources
- Deploy VNF as a vApp from catalog
- Review network mapping and connectivity of VNF
- Configure anti-affinity rules if required in VMware Integrated OpenStack
- Power ON VNFs
- Review end-to-end network connectivity

6 Supporting components

The VMware vCloud® NFV™ platform relies on the following supporting components:

- Directory service - This is used to provide single sign-on authentication services to all the management components.
- DNS - This is to provide a domain naming service for all management components. Management components such as VMware® Integrated OpenStack, require the host name to be fully resolvable, either in forward or reverse lookup.
- NTP - As a best practice, all the management components, including the physical ESXi servers should have their clocks synchronized to a single time source. VMware recommends a dedicated NTP server be setup for this purpose.
- SMTP - An SMTP server is valuable in sending email notifications from the platform to various administration teams. This is best used with VMware vRealize® Operations Manager™ to send email notifications based on advanced metrics, events and health parameters.

7 Monitoring and logging

The monitoring and logging blocks of the VMware vCloud® NFVTM platform comprises of VMware vRealize® Log InsightTM and the VMware vRealize® Operations ManagerTM. CPU, memory, network, and disk performance and capacity are monitored by vRealize Operations Manager and syslog events are sent to vRealize Log Insight. The VMware vRealize® Log Insight Content PackTM gives it the ability to parse product-specific information from the respective product logs. Similarly, the vRealize Operations Manager Management Pack allows product-specific metrics, alerts, events, and dashboards to be used for monitoring. These are available for download from the VMware Solution Exchange.

7.1 Logging

VMware vRealize® Log Insight™ is the central repository for log files from all the management components. vRealize Log Insight uses content packs to parse the log files to extract information. This information from the log files can be used as a data source for VMware vRealize® Operations Manager™. Table 14 lists the VMware vCloud® NFV™ platform components and logs that are sent to vRealize Log Insight.

Note: See *vRealize Operations Management Pack for Log Insight Installation and Configuration Guide* for instructions on how to install and configure the vRealize Log Insight management pack. This management pack allows the vRealize Operations Manager to query vRealize Log Insight for the logs of the selected inventory object. For this, enable the **Launch in Context for Log Insight** feature as described in the Installation and Configuration Guide.

Table 14 Component log collection

System	Logs	Content Pack
VMware ESXi™	syslog host.d vpxa.log	vRealize Log Insight Content Pack
vCenter Server®	syslog events, tasks, alarms	vRealize Log Insight Content Pack
Dell EMC ScaleIO®	logs, traces	None
NSX® Manager™	NSX Manager syslogs Distributed firewall logs NSX Edge syslog	NSX for vSphere
VMware® Integrated OpenStack	Syslog	None
VMware vSphere® Data Protection™	Syslog	None
Microsoft® Windows Server® 2012		Microsoft Windows
Microsoft® SQL Server® or Oracle® Database		Microsoft SQL Content Pack for Oracle® Databases

VMware vRealize® Orchestrator™	Syslog	None
VMware vRealize® Operations Manager™	Syslog	vRealize Operations Manager
VMware Site Recovery Manager™	SRM Server Logs	None
VMware vSphere® Replication™	Syslog	None

If logs are to be retained for a longer period of time, an external NFS share can be mounted for log archiving. The size of this share is based on the logs to be archived and the period of retention.

7.2 ScaleIO logging

- Retrieve ESXi logs – Collect logs from these folders:
  - /var/log
  - /scratch/log
- Retrieve ScaleIO components logs – Log in to SVM and run the following script for each component
  - `/opt/emc/scaleio/<scaleio component>/diag/get_info.sh -f`
    `where <scaleio component> is mdm or sds`

See *ScaleIO User Guide* for information about collecting logs.

7.3 Monitoring

The VMware vRealize® Operations Manager™ is used as the monitoring component for the VMware vCloud® NFV™ platform. The following management packs are used to gather metrics for management components and provide out of the box events, alerts and dashboards.

Note: Detailed information about installing, configuring and using these management packs are provided in the documentation accompanying the management packs and are not repeated here.

Table 15 Monitoring management packs

Management pack	Version	Description
vCenter Server	Bundled	Monitors entire data center including clusters, hosts, storage, networking, virtual machines etc
NSX for vSphere	3.0.2	Discover, analyze and graphically represent the broad number of virtual networking services available within NSX for vSphere. Quickly identify configuration, health or capacity problems within virtual NSX networks

vRealize Log Insight	1.0.1	Access the unstructured log data for any component of RA environment by allowing to launch in context
vRealize Network Insight	3.3.0	vRealize Network Insight has three primary use cases: Micro-segmentation planning/Security analysis 360-degree Network visibility NSX (Advanced) Operations Reference: VMware vRealize Network Insight 3.3 Information Center
Network Devices	1.0.1	Not only provides insight into virtual network layer but also provides information of the physical layer network devices like switches and routers
OpenStack	1.5	Provides out of the box dashboards, reports, inventory views, and alerts complete with remediation actions for comprehensive operational capabilities for managing an OpenStack environment Reference: VMware Integrated OpenStack

The vROps monitors general health of SVM, although not at application level. In case of any issues, storage admin can drill down further by logging into the ScaleIO Graphical User Interface (GUI). The ScaleIO GUI enables users to review the system status, drill down to component level, and monitor individual components. The various screens display different views and data that are beneficial to the storage administrator. The ScaleIO GUI provides an interface to monitor the underlying ScaleIO components. There are multiple monitoring areas on the GUI:

- Dashboard
- Two protection domains
- Two storage pools

For more information, see the *EMC ScaleIO User Guide*.

7.3.1 Metrics

The following table lists the metric categories that are collected by the management packs:

Table 16 Monitoring metrics

System	Metric categories/ resource kinds
VMware vCenter Server®	• VC Server Resources - CPU usage, disk, memory, network, and summary metrics • VC Datacenter - CPU usage, disk, memory, network, storage, disk space, and summary metrics • VC Compute Resources - CPU usage, configuration, storage, disk space, disk, memory, network, power, and summary metrics • VC Resource Pools - Configuration, CPU usage, memory, and summary metrics • Host Systems - Configuration, Hardware, agent, CPU usage, datastore, disk, memory, network, storage, and summary metrics • Virtual Machines - Configuration, CPU use, memory, datastore, disk, virtual disk, guest file system, network, power, disk space, storage, and summary metrics • Datastores - Capacity, device, and summary metrics, do refer performance guide for more metrics information Reference: *vSphere Monitoring and Performance Guide*
VMware NSX® for vSphere®	• NSX Manager • NSX Controller Cluster • NSX Controllers • NSX Transport Zone • NSX Logical Router • NSX Edge • NSX Edge - DHCP Service • NSX Edge - DNS Service • NSX Edge - Firewall Service • NSX Edge - IPSec VPN Service • NSX Edge - L2 VPN Service • NSX Edge - Load Balancer Service • NSX Edge - NAT Service • NSX Edge - Routing Service • NSX Edge - SSL VPN Service • Top of Rack Switch • Physical Fabric Reference: *VMware® NSX for vSphere (NSX) Network Visualization Design Guide*
VMware® Integrated OpenStack	• Horizon (Web Portal) • OpenStack APIs/SDKs • Heat (App Templates) • Ceilometer(Telemetry) • Nova(compute)

	• Cinder(Block Storage) • Glance(Image Catalog) • Neutron(Networking) • vROPS • SRM • vSphere Replication • vOrchestrator Appliance • vRLI • vRNI • VDP

7.3.2 Dashboards

The following table lists the default dashboards that are provided out of the box with the respective component management packs:

Table 17 Monitoring dashboards

System	Dashboards
VMware NSX® for vSphere®	• NSX Main • NSX Logical Topology • NSX Object Path • NSX Edge Services
VMware® Integrated OpenStack	

8 High availability

The VMware vCloud® NFV™ platform architecture and design ensures that there is no single point of failure for any of the components that make up the platform. High availability of the individual components has already been covered in earlier sections. This section describes the high availability considerations in the key architectural blocks of the platform.

8.1 VMware vCloud NFV infrastructure

The first step in building a highly available platform is to ensure that all the hardware components are redundant. In the case of the VMware vCloud® NFV™ platform, this includes the physical switches and physical compute servers. The physical switches are connected in a redundant mesh configuration to ensure there is more than one path between two points. This ensures that physical switch failures can be tolerated.

Each of the compute servers should have redundant Ethernet ports connected to the physical network in link aggregation mode. Since this reference architecture uses ScaleIO virtual SAN for storage, this simplifies the storage network topology. If third party storage devices are deployed then similar high availability requirements of the storage network and devices needs to be taken into account.

Each cluster has four ESXi hosts at minimum to provide the requisite capacity and redundancy for both compute and ScaleIO virtual SAN storage resources. ScaleIO uses a mirroring scheme to protect data against disk and node failures. When an SDS node or SDS disk fails, applications can continue to access ScaleIO volumes and the data is still available through the remaining mirrors. ScaleIO starts a rebuild process that creates another mirror for the data chunks that were lost in the failure. The surviving SDS cluster nodes carry out the rebuild process by using the aggregated disk and network bandwidth of the cluster.

8.2 Virtualized Infrastructure Manager

All the components of the virtualized infrastructure manager have high availability configuration. The VMware NSX® Manager™ for the management cluster is used to deploy one armed load balancers to load balance management traffic for VMware® Integrated OpenStack and another load balancer for the Platform Services Controller (PSC). This allows immediate redirecting of traffic to the standby node in case the primary node fails.

In addition to this, VMware® vSphere High Availability is used to bring up the failed node on another ESXi host. Anti-affinity rules ensure that two vSphere HA nodes of the same management component are never on the same physical host. Each cluster has enough resources reserved to tolerate one host failure. The resource reservation is taken into account when planning the capacity of the clusters and ensures that there is sufficient resources to power on a failed node on another ESXi host should the need arise without impacting the performance of the nodes already running on the ESXi host. Management components such as NSX Manager and VMware vCenter Server® are protected using vSphere HA.