Architecting a vCloud NFV OpenStack Edition Platform

REFERENCE ARCHITECTURE
VERSION 2.0

vmware

Table of Contents

Figures

Tables

Executive Summary

This reference architecture provides guidance for designing and creating a greenfield Network Functions Virtualization (NFV) platform using VMware vCloud® NFV™ OpenStack Edition. VMware vCloud NFV OpenStack Edition combines a carrier grade NFV infrastructure with VMware® Integrated OpenStack as the NFV Virtualized Infrastructure Manager (VIM). This version of the VMware vCloud NFV OpenStack Edition platform combines the open Application Programming Interface (API) defined by OpenStack with stable and supportable NFVI, to create a platform to support Communication Service Providers (CSPs) in realizing the goals of Network Functions Virtualization.

The VMware vCloud® NFV™ OpenStack Edition platform is compliant with the European Telecommunications Standards Institute (ETSI) Network Functions Virtualization (NFV); Architectural Framework. The platform is based on VMware components that are tightly integrated and tested. Each of the components has numerous potentially valid configurations, but only a few of these configurations result in a cohesive and robust functional system that meets business and technical requirements, and aligns with the ETSI NFV Architectural Framework.

The VMware vCloud NFV OpenStack Edition platform delivers the following ETSI NFV architectural components:

- NFV Infrastructure
- Virtualized Infrastructure Manager (VIM)
- NFVI Operations Management

These components, their interactions with each other, and the way in which they meet Communication Service Provider requirements, are described in this reference architecture.

Audience

This document is written to guide telecommunications and solution architects, sales engineers, field consultants, advanced services specialists, and customers responsible for virtualized network services and the NFV environment on which they run.

Document Structure

This document is divided into the four chapters listed in Table 1.

SECTION	DESCRIPTION
Network Functions Virtualization Overview	This section of the document introduces the core concepts of the ETSI NFV Architectural Framework.
Communication Service Provider Requirements	This section of the document describes the requirements of an NFVI platform from the perspective of the Communication Service Provider.
Solution Overview	This section of the document provides an overview of the components used to build the vCloud NFV OpenStack platform.
Reference Architecture	This section of the document describes how components are combined to deliver an NFV platform.

Table 1: VMware vCloud NFV OpenStack Edition Reference Architecture Document Structure

1. Network Functions Virtualization Overview

NFV is an architectural framework developed by the ETSI NFV Industry Specification Group. The framework aims to transform the telecommunications industry through lower costs, rapid innovation, and scale. The framework provides a standardized model that moves away from proprietary, purpose-built hardware dedicated to a single service, and toward network functions delivered through software virtualization of VNFs with commercial off-the-shelf (COTS) hardware. The result is a network that is more agile and quick to respond. The framework identifies functional blocks and the main reference points between such blocks. Some of these are already present in current deployments, while others must be added to support the virtualization process and operation.

Figure 1 shows the ETSI NFV Architectural Framework, depicting its functional blocks. The architectural framework focuses on the functions and capabilities necessary for the virtualization and operation of a CSP's network.

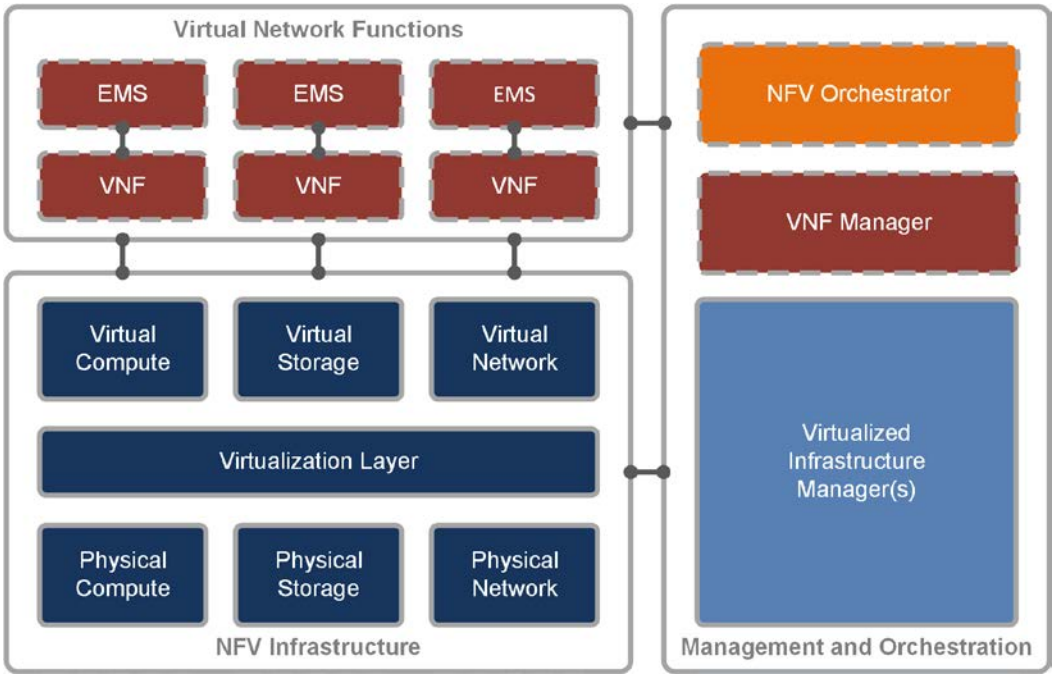

Figure 1: The ETSI NFV Architectural Framework

1.1 NFV Infrastructure

The Network Functions Virtualization Infrastructure (NFVI) is the foundation of the overall NFV architecture. It provides the physical compute, storage, and networking hardware that is used to host the virtual network functions. Each NFVI block can be thought of as an NFVI node, and many nodes can be deployed and controlled geographically. NFVI nodes are deployed at multiple sites and regions to provide service high-availability, and to support locality and workload latency requirements. The virtualization layer provided by the hypervisor allows for workloads that are agnostic to the underlying hardware. With this approach, operators choose hardware from their preferred vendors at competitive prices, and upgrade hardware independent of its workloads.

1.2 Management and Orchestration

The Management and Orchestration (MANO) functional block is responsible for the management of all resources in the infrastructure along with the orchestration and life cycle management of virtual network functions. These elements support the infrastructure virtualization and life cycle management of MANO VNFs, with a focus on the virtualization specific management tasks necessary to the NFV framework.

1.2.1 Virtualized Infrastructure Manager

The Virtualized Infrastructure Manager (VIM) is a functional block of the MANO and is responsible for controlling, managing, and monitoring the NFVI compute, storage, and network hardware, the software for the virtualization layer, and the virtualized resources. The VIM manages the allocation and release of virtual resources, and the association of virtual to physical resources, including the optimization of resources. The complete inventory of the NFVI is maintained in the VIM, including the linkage and relationship between components as they relate to an instance of a VNF workload, to allow for monitoring in the context of a single VNF.

1.2.2 Virtual Network Functions Manager

This document does not cover the Virtual Network Functions Manager (VNFM) functional block. For information about the VNFM functional block, refer to the publicly available ETSI NFV Standards.

1.2.3 Network Functions Virtualization Orchestrator

This document does not cover the NFV Orchestrator (NFVO) functional block. For information about the NFVO functional block, refer to the publicly available ETSI NFV Standards.

1.3 Virtualized Network Functions

This document does not cover the Virtualized Network Function (VNF) working domain. For information about the VNF working domain, refer to the publicly available ETSI NFV Standards.

1.4 Operations Support Systems and Business Support Systems

The vCloud NFV OpenStack Edition platform exposes application programmable interfaces (APIs) that can be consumed from one or multiple operations support systems and business support systems (OSS/BSS). These are not described in this document. For information about APIs that can be consumed from the OSS/BSS working domain, refer to the publicly available ETSI NFV Standards.

2. Communication Service Provider Requirements

More and more Communication Service Providers (CSPs) are using vCloud NFV to embark on a journey to modernize and transform networks and services with virtualized software components. Collected requirements help shape the current and future releases of the vCloud NFV solution. These key requirements are introduced in this section and will be discussed in detail in the *Reference Architecture* section of this document.

CSPs have specific requirements from the NFVI, VIM, and FCAPS elements, based on the need to demonstrate progress in an NFV deployment while generating revenue from the virtual domain. For this reason, a great deal of focus is given to the ability to easily, programmatically, and repeatedly deploy services from a service component catalog. Since CSPs deliver services that are often regulated by local governments, carrier grade aspects of these services, such as high availability and deterministic performance, are also included in this list. CSPs must ensure that managing the NFVI and the deployed virtualized network function is tightly integrated in the solution.

The following sections explain the requirements in further detail.

2.1 Automated Service Delivery

One of the benefits of virtualization is the ability to centrally orchestrate the deployment of service building blocks from a software catalog, as opposed to using proprietary appliances. Instead of sending engineers to a site to install physical devices, VNFs, also known as service components, are selected from a catalog. By clicking a button, the new service is installed. To reach this level of simplicity, the NFV platform must support the following:

- **Quick VNF Onboarding**. VNF onboarding is automated using enhanced templates for multi-vm services and declarative abstract resource requirements for underlying compute, storage, and networking resources.
- **Programmatic VNF Provisioning**. The speed and efficiency of VNF deployment is increased through automation, selecting service operations from a catalog of VNFs to deploy specific services.
- **True Multitenant Isolation**. Physical resources abstracted into virtual resource pools are shared between services and customers, referred to as tenants of the platform. The ability to partition the service and VNF from each other is key to ensure performance and quality of service (QoS) across the platform.
- **Service Life Cycle Management**. Programmatic service creation and dynamic orchestration of running VNFs are required pieces of an automation framework. Interfaces between the VIM, the VNFM, and the NFV Orchestrator (NFVO) must leverage a robust and open API. Using these interfaces, the NFV platform deploys, scales, restarts, and decommissions VNFs as needed.
- **Dynamic Optimization**. As more and more VNFs are deployed on the NFVI, NFVI resources must be able to proactively act on specific operations. Since the NFV environment is software based, the system must be able to move VNF components to balance fair and optimized resource utilization. NFVI resiliency is improved with proactive monitoring and automation – from scalability of resource pools to avoid issues, to policy based workload placement.

2.2 Operational Intelligence

Building an NFVI and managing VNFs effectively are primary requirements of all CSPs. Operation and management of the NFV environment must be tightly integrated with the other benefits of the solution. The functions CSPs require include:

- **Discovery and Reconciliation**. The NFV platform must automatically discover the network and virtual machine topologies across the physical and virtual domains, and reconcile runtime states as they change. The NFVI, VNFs, and VNF components (VNFCs) must be entirely visible to the operating personnel.
- **Performance and Capacity Monitoring**. Continuous system performance monitoring must provide a holistic view and alerting of key performance indicators such as interface utilization, data rates, capacity demand, service-level agreement (SLA) violations, and component availability. The same system must be intelligent and provide capacity and performance forecasts with actionable recommendations.
- **Fault Isolation and Remediation**. The platform must provide near real-time root cause analysis, and meaningful alerting for fast remediation and proactive issue avoidance.
- **Workflow Automation and Expansion**. The monitoring platform must be expandable to allow integration with new data source consumption and coexistence with other elements such as OSS, service orchestration, service assurance, and big data analytics. Where possible, the monitoring system must provide a way to add third-party expansion modules for higher layer monitoring, such as VoIP and video quality.

2.3 Carrier Grade

CSPs deliver certain services that are considered critical to infrastructure and are therefore tightly regulated by many governments. These regulations force a level of service quality to which over-the-top (OTT) providers do not adhere. CSPs must conform to specific service quality metrics, for example resolving service disruptions quickly and automatically without packet loss affecting service quality. The same applies to services offered from a CSP to enterprise customers. Service quality is at the core of brand protection and customer experience management. As such, SLAs require the delivery of carrier grade quality services. The following examples are essential NFV platform requirements for carrier grade systems:

- **High Availability and Resiliency**. The platform must provide integrated high availability and fault tolerance across the NFVI, virtual, and management domains. In the event of a failure, the platform must be able to self-heal to maximize service uptime. Mean Time To Failure (MTTF) must increase over the lifetime of the platform through adaptive, proactive, and automated monitoring systems.
  Mean Time To Repair (MTTR) must decrease over the lifetime of the NFV environment, as the system is optimized and proactive alerting takes place.
- **Performance**. The platform must achieve deterministic performance. The amount of resources required to deliver a certain level of performance must be well understood. Data plane intensive VNFs must be supported by the same components as control and management plane VNFs.
- **Scalability**. CSPs expect growth in the number of customers and services deployed on the NFV platform. The platform must be able to provide long term scale out capabilities, and dynamic and short term scale up and scale out functions.
- **NFVI Life Cycle (Patching and Upgrades)**. The platform must be patched and upgraded by using optimized change management approaches for zero to minimal downtime.

3. Solution Overview

The VMware vCloud NFV 2.0 OpenStack platform is an evolution of the VMware NFV solution, based on extensive customer deployment and the continued development of standards organizations such as the European Telecommunications Standards Institute (ETSI). The vCloud NFV OpenStack platform provides a comprehensive, service-oriented solution, leveraging a cloud computing model that allows ubiquitous, programmatic, on-demand access to a shared pool of compute, network, and storage resources. The solution is integrated with holistic operations management and service assurance capabilities, empowering the operator to rapidly deliver services while ensuring their quality. With a fully integrated VIM, the same vCloud NFV infrastructure delivers a myriad of telecommunications use cases, and facilitates reusability of the service catalog based VNFs.

The vCloud NFV OpenStack Edition platform delivers a complete, integrated solution that has been rigorously tested to ensure compatibility, robustness, and functionality. Components used in creating the solution are currently deployed across many industries and scenarios. vCloud NFV software components can be used in a variety of ways to construct a comprehensive, end-to-end solution that meets the business goals of CSPs. This document discusses one way components can be used to create a vCloud NFV architecture.

3.1 Technology Mapping

The vCloud NFV OpenStack Edition platform is an ETSI compliant, fully integrated, modular, multitenant NFV platform. It meets the ETSI NFV framework, which covers the virtualization layer of the NFVI and the VIM. vCloud NFV expands on the ETSI NFV framework by integrating robust operations management and intelligence components to provide the operator with complete platform visibility and monitoring functionality. The NFVI Operations Management functional block is responsible for providing and extending visibility for fault, configuration, accounting, performance, and security (FCAPS) of the NFVI. These northbound reporting components provide infrastructure alarms, events, and measurable metrics relating to the NFVI working domain and VNF workloads.

This document focuses on the NFVI layer, the VIM components, and the NFV platform operations management. Figure 2 depicts the mapping between the vCloud NFV OpenStack Edition functional elements and the ETSI NFV reference model.

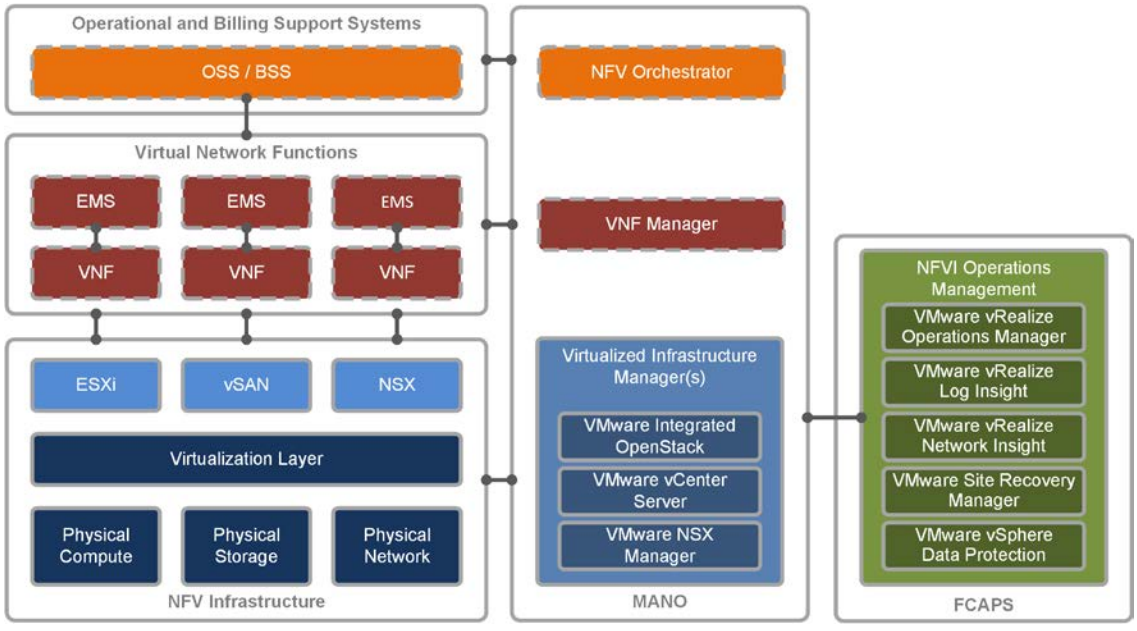

Figure 2: Mapping Functional Elements to the ETSI NFV Reference Model

The vCloud NFV OpenStack Edition bundle packages together the essential building blocks to deploy an NFVI and VIM platform, featuring the newest releases of VMware production proven solutions. Table 2 lists the components of vCloud NFV OpenStack Edition and their alignment with the ETSI NFV architectural framework.

COMPONENT	INCLUDED IN VCLOUD NFV BUNDLE	REQUIRED IN SOLUTION	ETSI FUNCTIONAL BLOCK
VMware ESXi™	Yes	Required	NFVI
VMware vCenter ® Server Appliance™	No	Required	VIM
VMware vSphere® Replication™	Yes	Recommended	NFVI Operations
VMware vSphere® Data Protection™	Yes	Recommended	NFVI Operations
VMware vSAN™ Standard Edition	Yes	Recommended	NFVI
VMware vRealize® Operations™ Advanced	Yes	Required	NFVI Operations
VMware vRealize® Network Insight™	No	Recommended	NFVI Operations
VMware vRealize® Log Insight™	Yes	Required	NFVI Operations
VMware® Integrated OpenStack Carrier Edition	Yes	Required	VIM
VMware NSX® for vSphere®	No	Required	NFVI
VMware NSX® Manager™	No	Required	VIM
VMware Site Recovery Manager™	No	Recommended	NFVI Operations

Table 2: VMware vCloud NFV OpenStack Edition Components

3.2 NFVI Components Overview

The NFV infrastructure consists of ESXi to virtualize the compute resources, NSX for vSphere to provide virtual networking, and vSAN for storage. Together these components create the virtualization layer described by the ETSI NFV framework. The virtualization layer of the NFVI provides the following functions:

- **Physical Resource Abstraction**. Using the software component layers between physical hardware and the VNFs, physical resources are abstracted. This provides a standardized software based platform for running VNFs, regardless of the underlying hardware. As long as the CSP uses certified physical components, VNFs can be deployed by the carrier at the point of presence (POP), distributed, or centralized data center.
- **Physical Resource Pooling**. Physical resource pooling occurs when vCloud NFV OpenStack Edition presents a logical virtualization layer to VNFs, combining the physical resources into one or more resource pools. Resource pooling together with an intelligent scheduler facilitates optimal resource utilization, load distribution, high availability, and scalability. This allows for fine grained resource allocation and control of pooled resources based on specific VNF requirements.
- **Physical Resource Sharing**. In order to truly benefit from cloud economies, the resources pooled and abstracted by a virtualization layer must be shared between various network functions. The virtualization layer provides the functionality required for VNFs to be scheduled on the same compute resources, collocated on shared storage, and to have network capacity divided among them. The virtualization layer also ensures fairness in resource utilization and usage policy enforcement.

The following components constitute the virtualization layer in the NFVI domain:

Compute – VMware ESXi

ESXi is the hypervisor software used to abstract physical x86 server resources from the VNFs. Each compute server is referred to as a host in the virtual environment. ESXi hosts are the fundamental compute building blocks of vCloud NFV. ESXi host resources can be grouped together to provide an aggregate set of resources in the virtual environment, called a cluster. Clusters are used to logically separate between management components and VNF components and are discussed at length in the *Reference Architecture* section of this document. ESXi hosts are managed by VMware vCenter® Server Appliance™, described as one of the VIM components in the *VIM Components* section of this document.

Storage – VMware vSAN

vSAN is the native vSphere storage component in the NFVI virtualization layer, providing a shared storage pool between hosts in the cluster. With vSAN, storage is shared by aggregating the local disks and flash drives attached to the host. Although third-party storage solutions and technologies such as NFS, iSCSI and FC-SAN with storage replication adapters that meet VMware storage compatibility guidelines are also supported, this reference architecture discusses only the vSAN storage solution.

It is a best practice recommendation that each cluster within vCloud NFV OpenStack Edition is configured to use a shared storage solution. When hosts in a cluster use shared storage, manageability and agility improve.

Network – VMware NSX for vSphere

The third component of the NFV infrastructure is the virtualized networking component, NSX for vSphere. NSX for vSphere allows CSPs to programmatically create, delete, and restore software based virtual networks. These networks are used for communication between VNF components, and to give customers dynamic control of their service environments. Dynamic control is provided through tight integration between the VIM layer and NSX for vSphere. Network multitenancy is implemented using NSX for vSphere, by assigning different customers their own virtual networking components and providing different network segments to each.

Just as ESXi abstracts the server resources, NSX for vSphere provides a layer of abstraction by supporting an overlay network with standards based protocols. This approach alleviates the limitations of traditional network segregation technologies such as VLANs, while creating strict separation between management, customer, and service networks. NSX for vSphere is designed as three independent layers: the data plane, the control plane, and the management plane. The data plane and control plane layers are described in the bullet points here, while the management plane is described in the *VIM Components* section of this document.

- **VMware NSX® Logical Switch™**. The NSX Logical Switch is a distributed data plane component within the ESXi hypervisor kernel that is used for the creation of logical overlay networks, facilitating flexible workload placement of the VNF components. The NSX Logical Switch is based on the VMware vSphere® Distributed Switch™ (VDS) and extends VDS functionality by adding distributed routing, a logical firewall, and enabling VXLAN bridging capabilities. The NSX Logical Switch is central to network virtualization, as it enables logical networks that are independent of physical constructs, such as VLANs. The NSX LogicalSwitch is a multilayer switch and therefore supports Layer 3 functionality to provide optimal routing between subnets directly within the host, for communication within the data center.

  Stateful firewall services are supported by the NSX Logical Switch through the distributed firewall service known as micro-segmentation. This functionality provides firewall policy enforcement within the hypervisor kernel at the granularity of the virtual Network Interface Card (vNIC) level on a VNF component, thereby supporting fine grained network multitenancy.

- **VMware NSX® Edge™**. The NSX Edge acts as the centralized virtual appliance for routing traffic into and out of the virtual domain, toward other virtual or physical infrastructure. This is referred to as North-South communication. In its role in vCloud NFV design, the NSX Edge is installed as an Edge Services Gateway (ESG). The ESG is used to provide routing, firewalling, network address translation (NAT), and other services to consumers of the NFVI platform. These NSX ESG instances, together with NSX Virtual Switches, provide true logical tenant isolation.

- **VMware NSX® Controller™**. The NSX Controller is the control plane responsible for the creation of the logical topology state necessary for connectivity between the components that form a VNF. Consisting of three active virtual controller appliances, the NSX Controller nodes form a cluster to maintain NSX Controller availability. The NSX Controller communicates with the ESXi hosts to maintain connectivity to the data plane components using out-of-band connectivity.

3.3 MANO Components Overview

As described in the *Management and Orchestration* section of this reference architecture, the ETSI NFV Management and Orchestration (MANO) framework consists of three functional blocks: The Virtualized Infrastructure Manager (VIM), the NFV Orchestrator (NFVO), and the VNF Manager (VNFM).

The vCloud NFV OpenStack Edition platform includes an integrated VIM, which exposes well documented northbound interfaces to VNFMs and NFVOs. VNFM components are often packaged together with VNFs. NFVO partners and independent VNFM solutions are listed on the VMware Network Functions Virtualization Telecommunication Solutions page.

3.3.1 VIM Components

The three main products that form the VIM functionality are VMware vCenter Server, VMware NSX Manager and VMware Integrated OpenStack.

- The VMware vCenter Server is the centralized management interface for compute and storage resources in the NFVI. It provides an inventory of allocated virtual to physical resources, manages inventory related information, and maintains an overview of the virtual resource catalogs. vCenter Server collects data detailing the performance, capacity, and state of its inventory objects. It exposes programmatic interfaces to other management components for fine grained control, operation, and monitoring of the underlying virtual infrastructure.
- NSX Manager provides the interface for creating, configuring, and monitoring NSX components. It is accessible through the graphical user interface (GUI) or through REST APIs and controls various virtualized network components, including the controllers, logical switches, and gateways. The NSX Manager is responsible for component deployment and management, and for the functions used to support the creation of network services by the VNFs.
- VMware Integrated OpenStack is the component exposed by vCloud NFV as the interface to the NFVI. It leverages the vCenter Server Appliance and the NSX Manager to orchestrate compute, storage, network, and imaging infrastructure services from a single, programmable interface. The VMware Integrated Openstack components used to enable these services include OpenStack projects: Horizon, Keystone, Nova, Neutron, Cinder, Glance, and Heat. More detail about VNF services and components will be provided in the next sections of this document.

3.4 Operations Management Components

The operations management solution includes six components that together provide a holistic approach to operations management functionality for the NFV infrastructure of a Communication Service Provider (CSP). Together, the vRealize Operations Manager, vRealize Log Insight, vRealize Network Insight, Site Recovery Manager, vSphere Replication, and vSphere Data Protection components monitor and manage the health of the virtualization infrastructure. They collect its logs and alarms, correlate events across multiple data sources and components to predict future issues, leverage the policy based automation framework to conduct remediation, and analyze data to help with health prediction and capacity planning.

The key operations management tasks carried out by these components are:

- **NFVI Visibility**. NFVI visibility is achieved by collecting key performance and capacity metrics from the entire virtualization layer, the physical devices, and the VIM components. When problems occur, the operator can uncover the root cause and determine its location quickly, reducing the Mean Time To Repair (MTTR).
- **Fault Collection and Reporting**. Components used in the NFV environment, in the physical infrastructure, the virtualization layer, and even the VNFs themselves, generate various log messages and alarms. vCloud NFV includes an integrated log collection system that correlates between alerts and log messages to quickly troubleshoot issues.
- **Performance Management and Capacity Planning**. Ongoing management of performance and capacity across the NFVI is required for optimal and economic use of the platform. The performance management capability helps identify degraded performance before VNFs are affected. This operator has enough time to take corrective measures, increasing the Mean Time To Failure (MTTF).
- **Optimization**. The operations management components analyze system usage and proactively provide optimization recommendations, including network topology modifications.

The specific components responsible for operations and management are:

VMware vRealize Operations Manager

VMware vRealize® Operations Manager™ delivers intelligent operations management with full stack visibility across the physical and virtual infrastructure. Through integrated performance and health monitoring functions, vRealize Operations Manager improves system performance, avoids service disruption, and helps the CSP provide proactive management of the NFVI. The key capabilities that enable these benefits include predictive analytics, smart and configurable alerts, and user guided remediation.

The vRealize Operations Manager extends to collect information through management packs. Information collected is filtered for relevancy, analyzed, and presented in the form of customizable dashboards. The monitoring solution exposes an API that retrieves performance and health data pertaining to the NFVI, and the virtual resources that make up the VNF instance, through an external system.

Out of the box, vRealize Operations Manager does not monitor VNF service availability or VNF internal key performance indicators (KPIs). The VNF Manager derives this information through direct interaction with the respective VNFs. However, VNF vendors can write their own management packs, known as plugins in vRealize Operations Manager, to extend functionality to the VNF application. In doing so, the vRealize Operations Manager becomes a single pane of glass from which the operator manages all components required to construct a virtual network service.

vRealize Operations Manager exposes the information it gathers through an API that can be consumed by OSS/BSS, or integrated directly with other MANO components.

VMware vRealize Log Insight

vRealize Log Insight delivers heterogeneous and highly scalable log management with intuitive, actionable dashboards, sophisticated analytics, and broad third-party extensibility. It provides deep operational visibility and faster troubleshooting across physical, virtual, and cloud environments. Its innovative indexing and machine learning based grouping enable fast log searches that aid in quickly troubleshooting issues.

vRealize Log Insight ingests large amounts of syslog data from the physical and virtual NFVI components to deliver near real-time monitoring, search, and log analytics. It automatically identifies structure from all types of machine generated log data including application logs, network traces, configuration files, messages, performance data, and system state dumps, to build a high-performance index used for analytics. Coupled with a highly intuitive dashboard for stored queries, reports, and alerts, vRealize Log Insight assists the operator in speedy root cause analysis and reduction in Mean Time To Repair (MTTR).

The vRealize Log Insight API provides programmatic access to vRealize Log Insight functionality and to its datastore. As a result, the OSS/BSS systems or MANO components can integrate with vRealize Log Insight to gain further insight into the system events and logs.

VMware vRealize Network Insight

vRealize Network Insight collects metrics, log data, network topology, and event data to provide a detailed view of the network configuration and its health. Information is collected on all NSX managed networks including East-West traffic between VNF components, and North-South traffic into and out of the NFV infrastructure. Broad Layer 2 to Layer 3 support means that vRealize Network Insight can visualize both the underlay and overlay networks, providing the operator with a holistic view into all relevant network layers. Using this information for visibility and analytics across all virtual and physical elements, the operator can optimize network performance and increase its availability.

VMware Site Recovery Manager

Site Recovery Manager works in conjunction with various storage replication solutions, including vSphere Replication, to automate the process of migrating, recovering, testing, and failing back virtual machine workloads for disaster recovery across multiple sites.

VMware vSphere Replication

vSphere Replication is a virtual machine disaster recovery solution. It is fully integrated with vCenter Server and VMware vSphere® Web Client, providing host based, asynchronous replication of virtual machines including their storage.

VMware vSphere Data Protection

vSphere Data Protection is used for backup and recovery. It is fully integrated with vCenter Server and vSphere Web Client, providing image-level, file-level, and guest-level application consistent backup and recovery. It conserves storage usage through standard deduplication techniques.

4. Reference Architecture

This reference architecture provides a template for creating an ETSI compliant vCloud NFV OpenStack Edition platform to support the rapid deployment of virtualized network services across different sites and regions. The architecture is designed in accordance with these principles:

- To be a carrier grade solution offering performance and high availability
- With modularity of infrastructure, VNFs, and services
- To support a service life cycle with rapid VNF onboarding and automation
- As a tenant based architecture with reusable services, policy driven service components, and resource reservation
- For integrated operations management monitoring and analysis of the entire NFV environment

4.1 Design Principles

The five architectural pillars on which vCloud NFV OpenStack Edition stands are driven by VMware customer requirements and individual component capabilities. These are described in more detail in the following sections of this document.

4.1.1 Carrier Grade

The vCloud NFV OpenStack Edition platform components are used by a variety of VMware customers from industries such as large enterprise, health care, and finance. Carrier grade capabilities are continuously added to the platform to address the requirements of VMware CSP customers. With this release, improvements in high availability and performance are fundamental to the vCloud NFV design.

Data plane forwarding performance of the platform has improved to meet carrier grade requirements for specific VNFs such as vEPC and vRouter by providing dedicated CPU resources where needed, and identifying and resolving slow data plane paths. vCloud NFV OpenStack Edition includes specific functionality to enable VNFs that require precise and dedicated resource allocation to receive it.

The carrier grade design principle of high availability (HA) is divided into two different layers in the NFV environment:

- **Platform High Availability**. Platform HA ensures that the components needed to manage the NFV environment are always configured in a redundant fashion, replicating data across multiple storage elements and databases. Ensuring that the management components in the platform are always available and self-healing allows the operations team to focus on the services and service constructs.
- **VNF High Availability**. The vCloud NFV OpenStack Edition platform provides native resiliency functions that can be consumed by VNFs to increase their availability. For VNFs that do not provide their own high availability mechanisms, the platform offers advanced support to ensure that a VNFC instance failure can be quickly recovered and the boot sequence orchestrated to meet the VNF logic.

With these two availability principles, both the NFV platform and VNFs minimize service disruption.

4.1.2 Modularity

vCloud NFV organizes distinct functions into pods. A pod is a grouping of compute, network, and storage that work together to deliver a function required by the solution. Each pod has its own characteristics in terms of role, performance, and scale. This architectural best practice allows for efficient resource management, creates a clear demarcation between resource providers and resource consumers, establishes security boundaries, and allows for the design of different levels of availability based on cluster workloads.

Architecting vCloud NFV OpenStack Edition using well defined modules allows the CSP to accelerate deployment and reliably expand it when needed. The platform components are grouped into three distinct containments:

- **Management Functions**. Management functions are required to manage the NFV infrastructure and the Virtual Network Functions (VNFs). MANO components, FCAPS functionality, and ancillary elements such as DNS and OSS/BSS are grouped into this category.
- **Edge Functions**. The edge functions provide a logical networking delineation between VNFs and external networks. Network traffic transitioning between the physical domain and the virtual domain is processed by these functions. An example of such a function is the NSX Edge Services Gateway (ESG).
- **Resource Functions**. The VNFs, and functions related to VNFs such as VNF Managers, are grouped into this category.

Pods can be used to streamline the NFV environment operations and delineate between different roles. For example, a cloud management team can operate the Management pod while a network management team is likely to oversee the Edge pod. VNFs are always deployed in the Resource pod.

Each pod is identified by its functional designation: Management pod, Edge pod, and Resource pod. Components hosted in each pod include:

- **Management Pod**. VIM components such as vCenter Server Appliance, NSX Manager, and VMware Integrated OpenStack are hosted in this pod. The plethora of FCAPS components, which include vRealize Operations Manager, vRealize Network Insight, vRealize Log Insight, Site Recovery Manager, vSphere Replication, and vSphere Data Protection are located in this pod. Other management related components such as NFV Orchestrators run in the Management pod.
- **Edge Pod**. As described in the *Solution Overview* section of this document, the virtual network NFVI building block is based on NSX for vSphere. NSX ESG, hosted in the form of a virtual machine appliance in this pod, handles all connectivity to the physical domain for the architecture. Other edge functions can also be hosted in this pod based on the operator's needs. The type of networking traffic that traverses the Edge pod is referred to as North-South traffic.
- **Resource Pod**. Virtual Network Functions (VNFs) and their supporting components are located in the Resource pod. The VNFs form the virtual network services.

Pods represent a functional and modular boundary that is well defined and therefore easy to replicate. The three functions described here can be grouped into pods giving rise to two possible designs. The first design deploys the three pods into three dedicated distinct entities, forming the basis of three-pod design. The second design, two-pod design, collapses the edge and resource functions into a single Edge / Resource pod for a smaller initial footprint.

Figure 3 illustrates the logical building blocks used to create the platform.

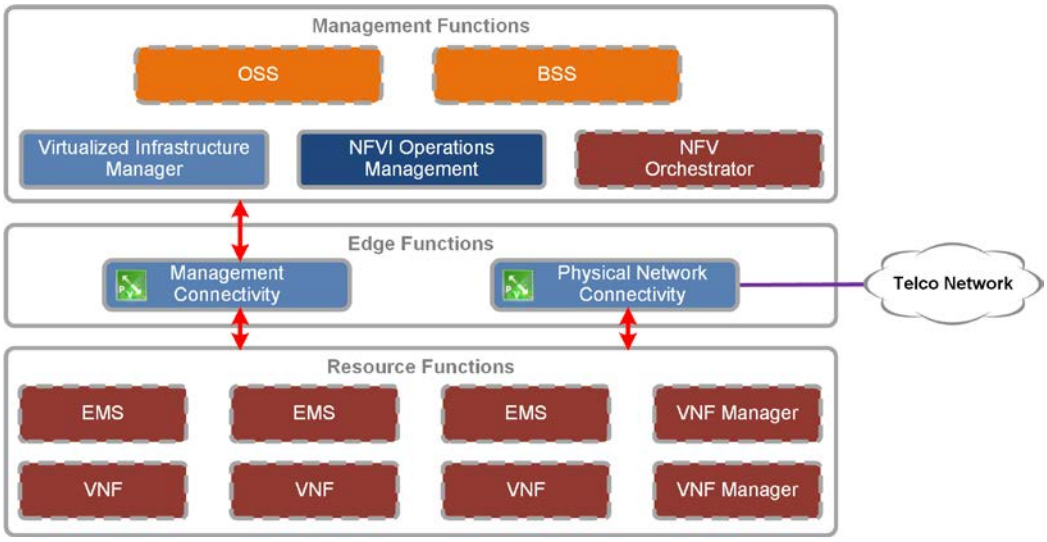

Figure 3: VMware vCloud NFV OpenStack Edition Logical Building Blocks

Service Life Cycle

The service life cycle design principle focuses on ease and the speed at which VNFs can be consumed by the NFV platform, maintained over their life time, and deployed when needed. The VIM facilitates this approach and enables the CSP to perform common tasks to benefit from virtualizing network functions.

VNF vendors package their VNFs and deliver them to the CSP in various consumable forms. CSPs can quickly onboard a packaged VNF to the NFV platform to speed up deployment, and to ensure that VNFs are consistently behaving in a predictable fashion in each deployment.

Once the VNF is onboarded, it is placed in an image library that can be consumed based on the CSP's policies. The goal of placing a VNF in the vCloud NFV OpenStack image library is to enable the NFVO, responsible for creating the VNF service, to quickly and programmatically deploy the VNFCs required.

Since many operational activities related to VNFs are performed using higher layer components such as the VNFM and VNFO, the vCloud NFV OpenStack Edition platform provides a well-documented northbound API that can be used by these components to complete the service life cycle.

VMware Integrated OpenStack also addresses life cycle activities including deployment, decommissioning, and restarting service components.

Tenant Based Architecture

The NFVI is shared between multiple entities, referred to as tenants of the NFVI. A fundamental aspect of NFVI design is ensuring that multiple tenants remain logically isolated from each other, although the physical and virtual layers they use may be shared. The design principles for multitenancy are:

- An NFVI tenant cannot interfere with the operations of another tenant, nor can one VNF interfere with another.
- Fair resource sharing must take place. When the system has available resources, and tenants require resources, available resources are split appropriately among the tenants.
- One tenant network must be isolated from another. A tenant's choice of IP allocation, default gateway, and routing cannot interfere with another tenant. In fact, other tenants may use the same networking information. Network access from one tenant to another must follow the trusted networking and security policy of the CSP.

- A tenant must be proactively monitored to ensure the health and efficiency to deliver optimal service quality.

These design principles allow multiple tenants to share resources on the operator's network, and to maintain a great deal of control and self-management. Tenants can use overlapping IP addressing and, together with the use of resource policies, the CSP can ensure that the amount of resources required by each tenant is controlled. Tenant based architecture, together with a well-defined process for VNF onboarding and VNF resource allocation, means that a CSP can offer service-level agreements (SLAs) with which to create high quality, mission critical services. With the integrated operations management principles, SLAs are monitored and ensured.

Integrated Operations Management

The multilayer, multi-vendor nature of the NFV environment can lead to increased operational management complexity. To resolve this complexity, vCloud NFV OpenStack Edition is integrated with a robust operational management system that monitors and analyzes components involved in the NFV environment. Physical servers and switches of the VMware operational management components include monitoring adaptors, which enable the entire system, including the virtualization layer and the VNFs themselves, to be automatically discovered, monitored, and analyzed.

For a monitoring system to provide visibility into the entire environment, you must be able to integrate data collection from VMware software components alongside third-party elements such as VNFs, routers, switches, servers, and storage elements. Complete visibility is achieved using the suite of software components described in the *Operations Management Components* section of this document.

Data collected by these components is continuously analyzed, which allows for near real-time monitoring. This robust performance monitoring enables the operator to perform detailed capacity management. With a monitoring system tightly integrated with the VIM, virtualization, and physical layers, proactive failure avoidance leverages vRealize Operations Manager analytics and the DRS. In cases where a problem does occur, root cause analysis is easily performed through the operator's holistic visibility into the entire system.

4.2 VIM Modularity Using VMware Integrated OpenStack

Modularity is one of the key design principles of the VMware vCloud NFV OpenStack Edition reference architecture. By integrating VMware Integrated OpenStack into the platform, the emphasis on modularity remains consistent. VMware Integrated OpenStack greatly simplifies the deployment of an OpenStack cloud infrastructure by streamlining the integration process. VMware Integrated OpenStack consists of upstream OpenStack that is hardened, tested, and preconfigured to use VMware optimized OpenStack drivers. VMware Integrated OpenStack includes tools required to install, upgrade, and operate a production-grade OpenStack cloud on top of existing VMware technologies.

OpenStack is a cloud operating system that manages pools of storage, network, and compute resources through a dashboard while allowing users to provision resources through a Web interface. Each of these resources are managed by core OpenStack services that are developed by corresponding OpenStack projects. Figure 4 provides a high-level overview of the resource abstrtaction layers and infrastructure services for the VNFs.

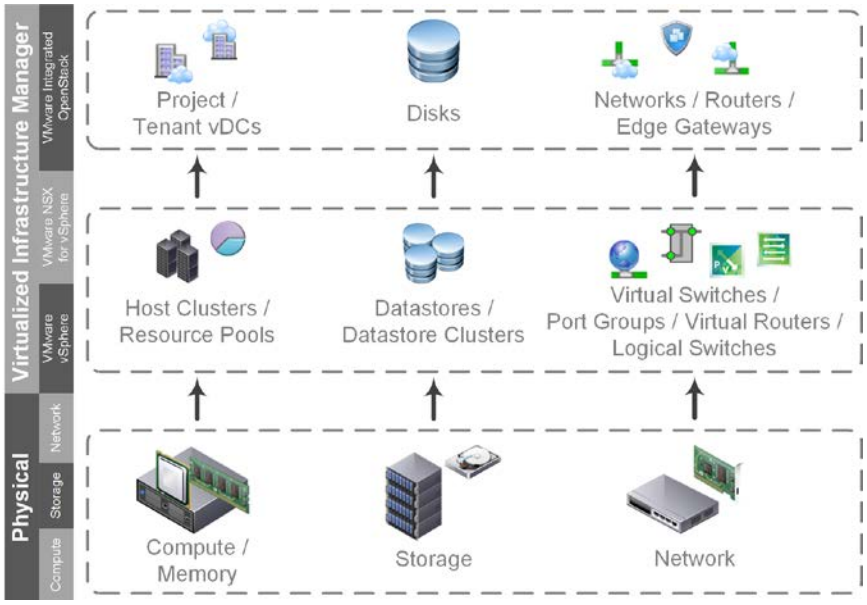

Figure 4: VIM Hierarchy in VMware vCloud NFV OpenStack Edition

VMware Integrated OpenStack packages core OpenStack projects such as Nova, Neutron, Cinder, Glance, Keystone, Heat, and Horizon. Administrators and tenants use the OpenStack user interface of Horizon, or APIs, to consume and manage resources in the OpenStack cloud. The VMware Integrated OpenStack services are run in one or more virtual machine instances managed by the OpenStack management server.

Figure 5 shows the various OpenStack components and their integration points with the vCloud NFV platform.

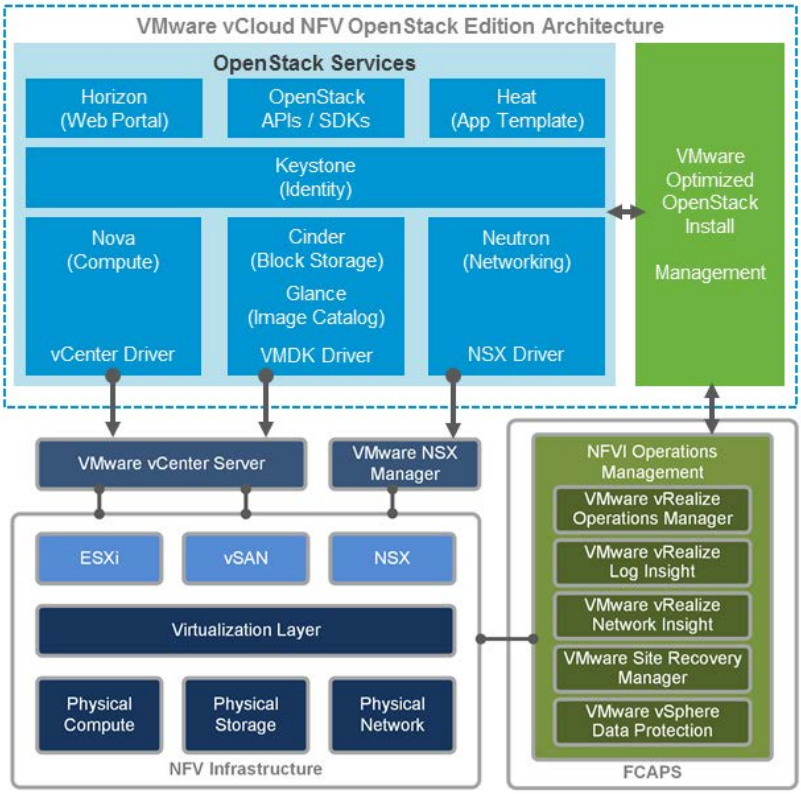

Figure 5: VMware vCloud NFV Integrated OpenStack Architecture

The following is a list of the OpenStack components and their services in the vCloud NFV Integrated OpenStack architecture:

- **Nova**. This component identifies each VMware vSphere compute cluster as a single compute node for pooling compute resources and selects the vSphere cluster in which to place the VNF workloads. Once a cluster is selected, vSphere Distributed Resource Scheduler (DRS) optimally places the VNF within the cluster to balance workloads during runtime.
- **Cinder**. This component executes block volume operations through a virtual machine disk (VMDK) driver. VMware vCenter creates the volume, which initially belongs to a shadow VM. When the volume is attached to a running VM, VMware vCenter changes the parent of the volume from the shadow VM to the actual VM.
- **Glance**. This component manages images stored in a set of dedicated image service storage containers. VMware Integrated OpenStack supports ISO, VMDK, and OVA formats, but RAW, QCOW2, VDI, and VHD formats can also be imported using the command line.
- **Neutron**. This component is used to provide the networking services for the creation of the virtual network resources that are used for tenants. It communicates with the NSX Manager to manage virtual network resources in the OpenStack cloud. An NSX Edge device functions as an OpenStack Layer 3 agent, Metadata server, DHCP Layer 2 agent, and is used for tenant networking and security group policy enforcement.
- **Horizon**. This component provides the Web based user interface for interacting with OpenStack APIs to perform common cloud operations. This interface is available to cloud administrators and users.
- **Keystone**. This component is used in OpenStack as an identity source for user authentication, authentication token management, and as a central service catalog for finding the API endpoints of OpenStack services.
- **Heat**. Heat is the orchestration service used for launching virtual machine instances that are connected together.

The VMware Integrated OpenStack Manager deploys and manages an instance of VMware Integrated OpenStack. The Integrated OpenStack Manager provides a workflow that allows CSPs to deploy VMware Integrated OpenStack in a few clicks, from within the vSphere Web Client.

4.3 Two-pod Design Overview

The vCloud NFV OpenStack Edition facilitates combining the edge and resource functionality into a single, collapsed pod that provides a small footprint. CSPs can use a two-pod design to gain operational experience with vCloud NFV OpenStack Edition. As demand grows, they scale up and scale out within the two-pod construct.

An initial two-pod deployment consists of one cluster for the Management pod and another cluster for the collapsed Edge / Resource pod. Clusters are vSphere objects for pooling virtual domain resources and managing resource allocation. Clusters scale up as needed by adding ESXi hosts, while pods scale up by adding new clusters to the existing pods. This design ensures that management boundaries are clearly defined, capacity is managed, and resources are allocated based on the functionality hosted by the pod. vCloud NFV OpenStack Edition VIM components allow for fine grained allocation and partitioning of resources to the workloads, regardless of the scaling method used.

Figure 6 shows a typical two-pod design with all management functions centrally located within the Management pod. Edge and resource functions are combined into the collapsed Edge / Resource pod. During initial deployment, two clusters of ESXi hosts are used: one for the Management pod, and the other for the collapsed Edge / Resource pod. Additional clusters can be added to each pod as the infrastructure is scaled up.

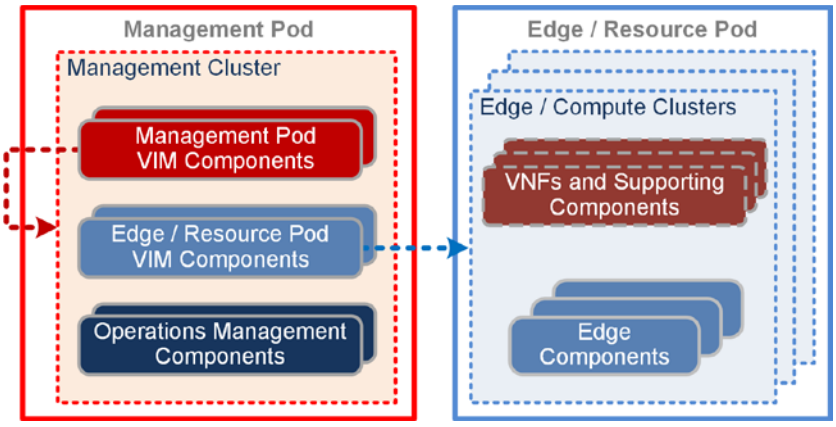

Figure 6: Two-pod Design Overview

As a best practice, begin the initial deployment with a minimum of four hosts per cluster within each pod, for a total of eight hosts. With initial four-host cluster deployment, a high degree of resiliency is enabled using vSAN storage. At the same time, four hosts allow for the placement of clustered management components such as vCenter Server active node, standby node, and witness node, on separate hosts in the Management pod creating a highly available Management pod design. This same design principle is used for clustered OpenStack components such as database nodes.

The initial number and sizing of management components in the Management pod are pre-planned. As a result, the capacity requirement of the Management pod is expected to remain steady. Considerations for planning Management pod storage capacity must include operational headroom for VNF files, snapshots, backups, virtual machine templates, operating system images, and log files.

The collapsed edge / resource cluster sizing will change based on the VNF and networking requirements. When planning the capacity of the Edge / Resource pod, tenants must work with VNF vendors to gather requirements for the VNF service to be deployed. Such information is typically available from the VNF vendors in the form of deployment guides and sizing guidelines. These guidelines are directly related to the scale of the VNF service, for example to the number of subscribers to be supported. In addition, the capacity utilization of ESGs must be taken into consideration, especially when more instances of ESGs are deployed to scale up as the number of VNFs increase.

When scaling up the Edge / Resource pod by adding hosts to the cluster, newly added resources are automatically pooled, resulting in added capacity to the compute cluster. New tenants can be provisioned to consume resources from the total available pooled capacity. Allocation settings for existing tenants must be modified before they can benefit from increased resource availability.

OpenStack compute nodes are added to scale out the Resource pod. Additional compute nodes can be added using the Integrated OpenStack Manager vSphere Web Client extension. Due to leaf-and-spine network design, additional ESGs will continue to be deployed in the initial Edge cluster.

Within the Management pod, a set of components is deployed to manage the pod itself. The components include an instance of vCenter Server Appliance, Platform Services Controllers (PSCs), an instance of NSX Manager, and VMware Integrated OpenStack management components.

Figure 7 provides an overview of the Management pod components and the Management pod's relationship to the collapsed Edge / Resource pod.

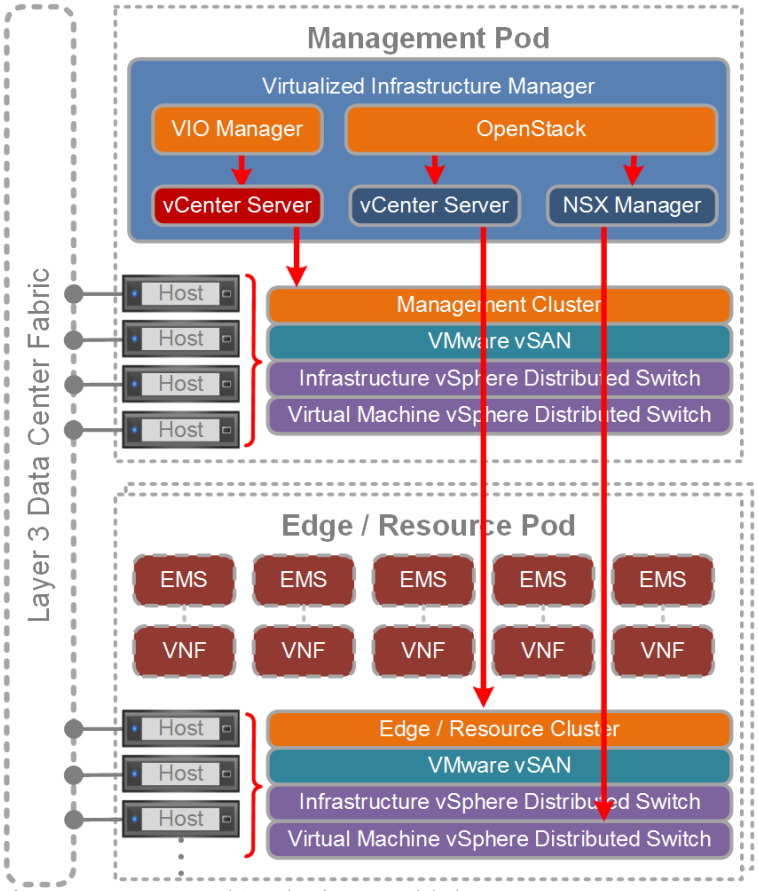

Figure 7: Management Pod Overview for Two-pod design

A 1:1 relationship is required between NSX Manager and vCenter Server. Ancillary components necessary for the healthy operation of the platform, such as a Domain Name System (DNS), can be deployed in the Management pod. The tenant-facing Virtualized Infrastructure Manager (VIM) component, VMware Integrated OpenStack, is located in the Management pod and is connected to the vCenter Server and its associated NSX Manager for networking. Also within the Management pod, a separate instance of vCenter Server is deployed to manage the Edge / Resource pod. Likewise, a separate NSX Manager is deployed to maintain the 1:1 relationship to the vCenter Server.

The Edge / Resource pod hosts all edge functions, VNFs, and VNFMs. The edge functions in the pod are NSX ESGs used to route traffic between different tenants and to provide North-South connectivity. Tenants can deploy either edge routers or distributed routers for East-West connectivity. Edge VMs deployed as part of the distributed logical router are used for firewall and dynamic routing.

Since both edge functions and VNF functions can be associated with a given tenant, resource utilization of the collapsed Edge / Resource pod must be carefully monitored. For example, an increase in the number of tenants and subsequent VNF functions deployed will inevitably expand the number of edge resources used. The *Operations Management* design section of this document discusses the approach to resource capacity management in this case. When resources are limited, collapsed Edge / Resource pod scale up operations must be carefully coordinated.

The VMware Integrated OpenStack layer of abstraction, and the ability to partition resources in vSphere, facilitate an important aspect of a shared NFV environment: secure multitenancy. Secure multitenancy ensures that more than one consumer of the shared NFV platform can coexist on the same physical infrastructure,

without an awareness of, ability to influence, or ability to harm one another. With secure multitenancy resources are oversubscribed yet fairly shared and guaranteed as necessary. This is the bedrock of the NFV business case. Implementation of secure multitenancy is described in the *Secure Multitenancy* section of this document.

4.3.1 Two-pod VMware vCenter Design

Figure 8 shows the vCenter Server instances and their relationship to two-pod design.

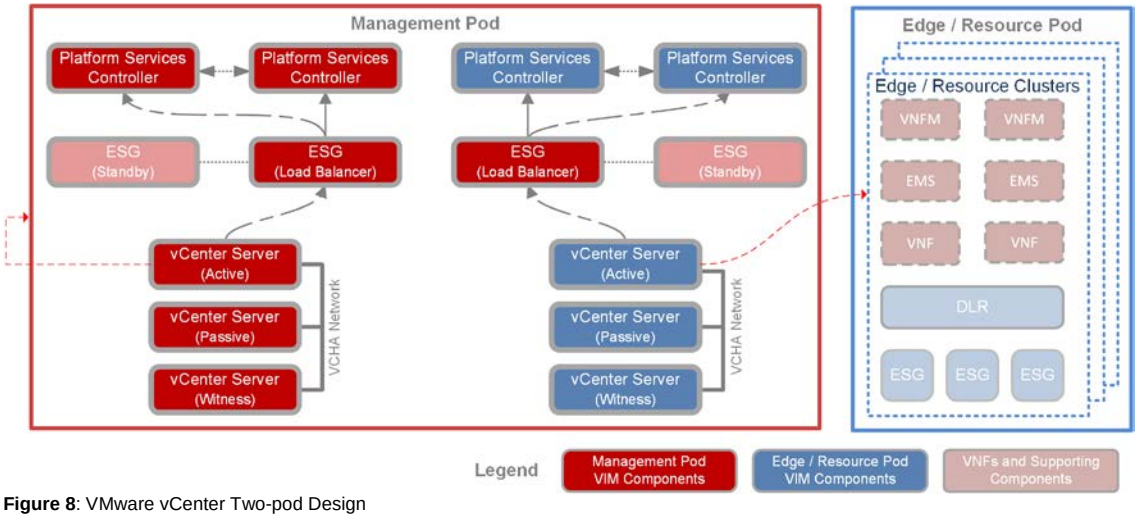

Figure 8: VMware vCenter Two-pod Design

In two-pod design, the Management pod is implemented as a cluster, governed by the first 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, in order to form the foundation of a carrier grade virtualized infrastructure. A second vCenter Server is deployed in the Management pod to oversee the Edge / Resource pod.

Each vCenter Server is a virtual appliance deployed with an embedded database. The VMware vCenter® Server Appliance™ is preconfigured, hardened, and fast to deploy. Use of the appliance allows for a simplified design, eases management, and reduces administrative efforts. vCenter Server Appliance availability is ensured using a three-node cluster. 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. Replication between nodes ensures that vCenter Server Appliance 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, service registration, 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.

Data backup and restore of each vCenter Server instance and its associated PSC 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.

Local storage drives on the ESXi hosts are pooled into a highly available shared vSAN datastore for optimum utilization of storage capacity. Each cluster has its own vSAN datastore, an abstracted representation of the storage into which virtual machine persistent data is stored. All management components are stored in the management cluster datastore, while VNF workloads deployed from VMware Integrated OpenStack are stored in the resource cluster datastore. This allows for the separation of administrative, performance, and failure storage boundaries for management and VNF workloads.

4.3.2 Two-pod Virtual Networking Design Using VMware NSX Manager

Each NSX Manager has a 1:1 relationship with vCenter Server. Therefore, two NSX Managers are created in the management cluster. Figure 9 shows how the NSX Manager is used in a two-pod design.

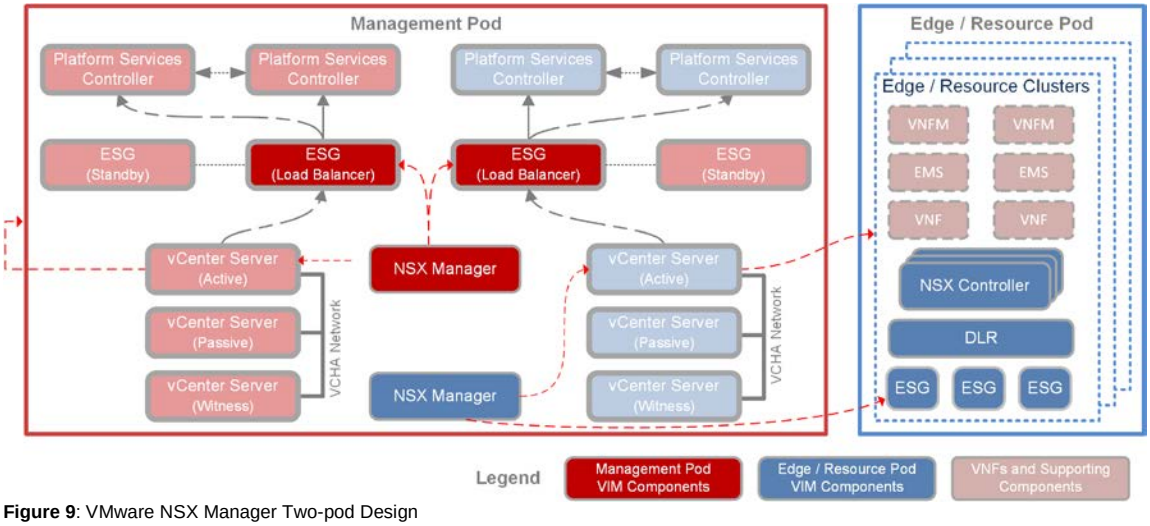

Figure 9: VMware NSX Manager Two-pod Design

The first NSX Manager in the Management pod is solely responsible for deploying and operating the highly available ESG instances that provide load balancing functionality. Every component in the management cluster, which relies on multiple external services such as PSCs, uses the ESG as a load balancer to ensure reachability should a component fail. VMware Integrated OpenStack comes with its own built-in high availability control plane implemented using HAproxy on the load balancer virtual machines (VMs).

The second NSX Manager in the Management pod is responsible for all Edge / Resource pod networking. It is registered with VMware Integrated OpenStack to provide networking services to tenants, including stateful firewalls and load balancers. The same NSX Manager is used to configure East-West VNF connectivity, North-South routing, and out-of-band management access for VNFs.

Infrastructure networks are used by the ESXi hypervisor for vMotion, VMware vSphere® Fault Tolerance, and vSAN traffic. The VM networks are used by VMs to communicate with each other. For each pod, the separation between infrastructure and VM networks ensures security and provides network resources where they are needed. This separation is implemented by two distributed switches, one for infrastructure networks and the other for VM 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. This is shown in Figure 10.

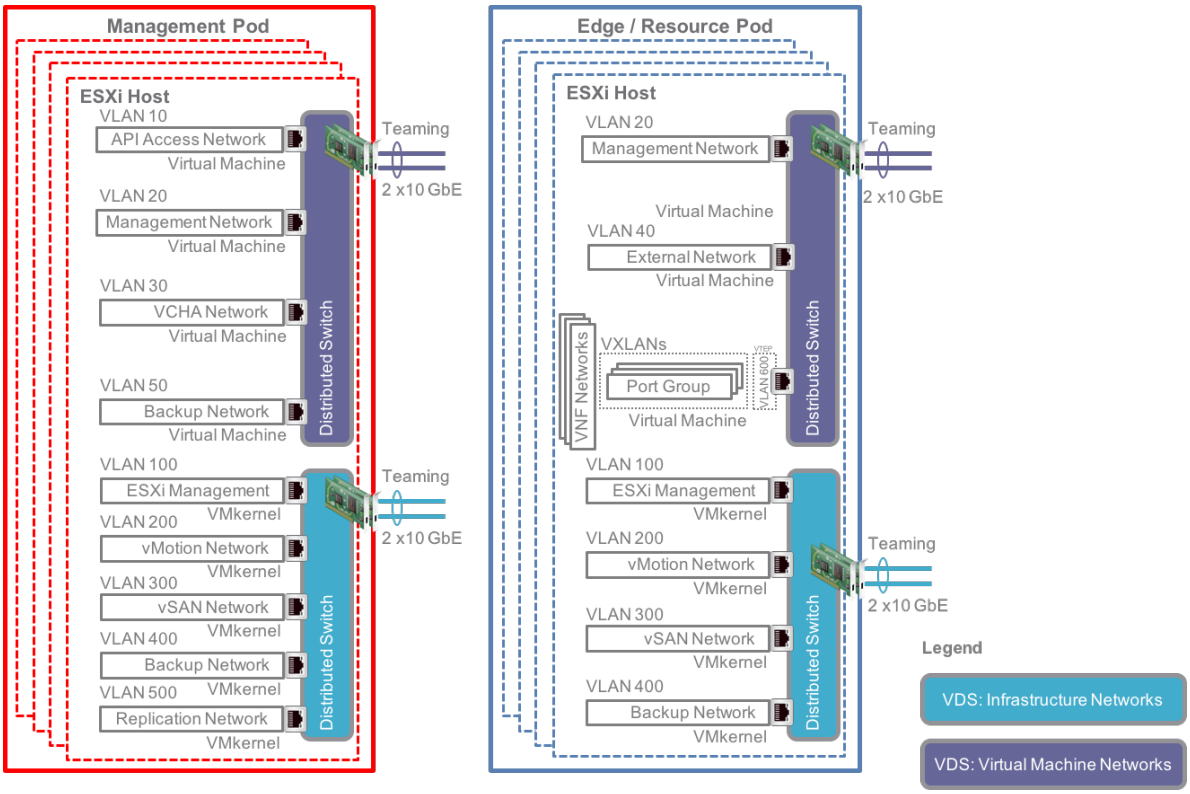

Figure 10: Distributed Virtual Switch Two-pod design

NSX for vSphere provides the network infrastructure for the East-West connectivity VNFCs require, and edge services for their North-South connectivity. When tenants deploy a logical router, they can choose from three types:

- **Centralized and Exclusive Routers**. With centralized and exclusive routers, the Neutron service in VMware Integrated OpenStack installs a dedicated instance of an ESG for the tenant. This design is the preferred way to secure tenant-level isolation, ensuring separation of VNF traffic across tenants.
- **Centralized and Shared Routers**. Centralized and shared routers share the ESG with other tenants, as long as there are no overlapping subnets. Distributed routers provide East-West routing capabilities.
- **Distributed Routers**. With distributed routers, Edge function is deployed by default and is used for firewalling and dynamic routing.

Two-pod design places NSX Controllers in the Edge / Resource pod to position them close to the data plane components of the NFVI platform. Three controllers are deployed as a cluster, ensuring control plane resiliency. NSX Controller Disconnect Operations (CDO) is a new feature introduced in NSX for vSphere 6.3.2. CDO ensures that data plane connectivity is unaffected when hosts lose connectivity with the controller.

In two-pod design, both the East-West and North-South networks are provisioned by the tenant from within VMware Integrated OpenStack.

4.3.3 Two-pod VMware Integrated OpenStack Design

The VMware Integrated OpenStack Manager connects to the vCenter Server instance that manages the Management pod. It uses a VM template to rapidly deploy, administer, and perform Day 2 management operations of the VMware Integrated OpenStack management plane components deployed in the Management pod.

Once deployed, VMware Integrated OpenStack connects to the vCenter Server instance that manages the collapsed Edge / Resource pod. This vCenter Server is responsible for storage and compute resources. VMware Integrated OpenStack also connects to the NSX Manager instance associated with Edge / Resource pod networking. Figure 11 illustrates the VMware Integrated OpenStack management components for two-pod design.

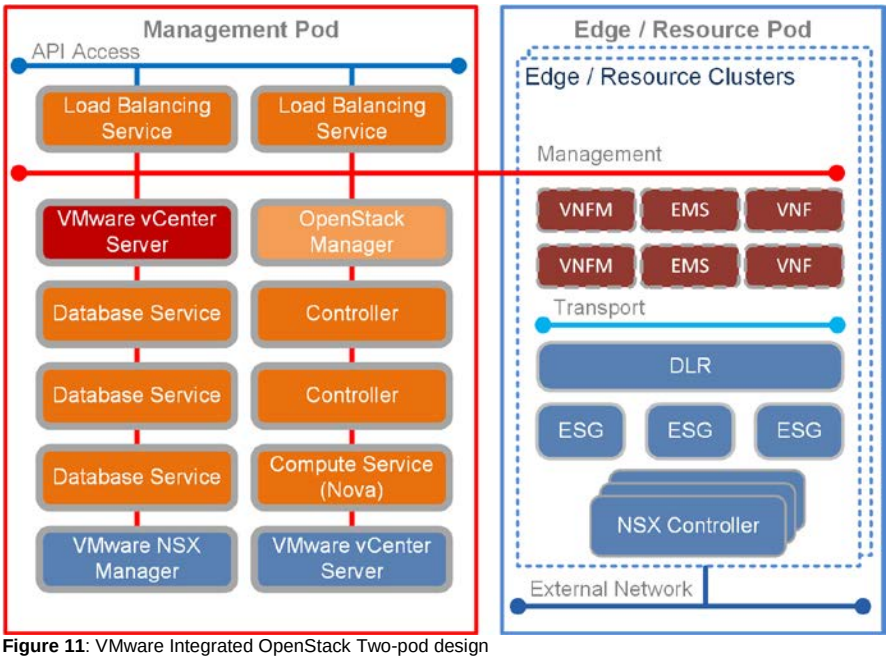

Figure 11: VMware Integrated OpenStack Two-pod design

The VMware Integrated OpenStack management plane is deployed with redundancy for all VMware Integrated OpenStack management components, ensuring that there is no single point of failure. Although this requires greater resource availability in the Management pod, it offers the best configuration for high availability and is the recommended topology for production environments.

In a VMware Integrated OpenStack HA deployment, all the components for a scalable and highly available VMware Integrated OpenStack full deployment, including clustered databases, controllers, and VMware Integrated OpenStack load balancers, can also be deployed by the Integrated OpenStack Manager. All management components have connectivity to each other through a dedicated management network. The clustered VMware Integrated OpenStack management components are shown in Figure 12.

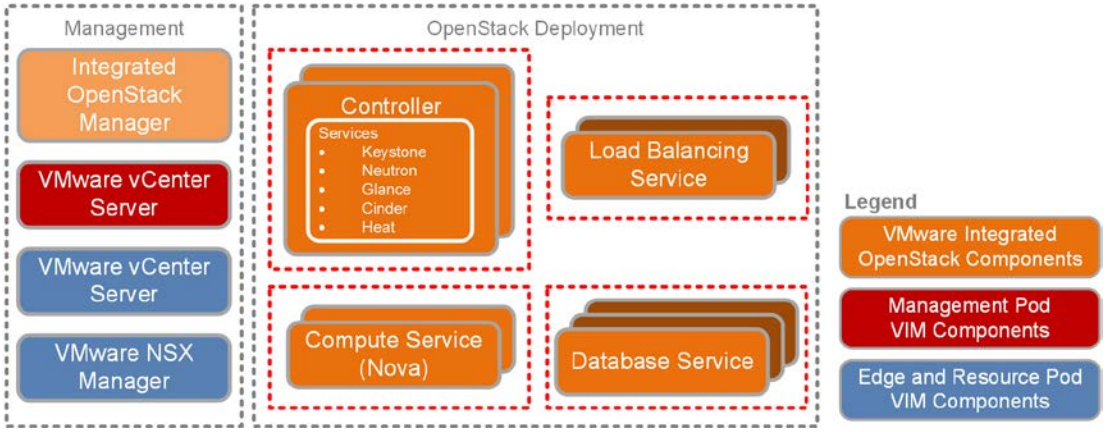

Figure 12: VMware Integrated OpenStack High Availability

VMware Integrated OpenStack is closely integrated with NSX for vSphere, providing tenants with more features and capabilities for managing their VNF networking needs using the Horizon interface and APIs. With VMware vCloud NFV OpenStack Edition, all of the building blocks for creating secure multitenant VNF networks are in the hands of the tenant. 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 and 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.

4.4 Three-pod Design Overview

The three-pod design completely separates the vCloud NFV functional blocks by using a Management pod, an Edge pod, and a Resource pod, for their respective functions. The initial deployment of a three-pod design consists of three clusters, with one cluster in each pod. Clusters can scale up as needed by adding ESXi hosts, while pods scale up by adding additional clusters. The separation of management, edge, and resource functions into individually scalable pods allows CSPs to plan capacity based on the needs of the specific function hosted by each pod. This provides greater operational flexibility.

The initial deployment of a three-pod design is more hardware intensive than that of a two-pod design, however each pod scales up independently of the others. A three-pod platform design consists of the same components that are used in a two-pod design, with differences in the way the pod functions are combined to form the solution. Regardless of the pod design used to create the NFVI, VNFs perform in the same way.

For best practices when using vSAN as the shared storage solution for all clusters, initial deployment requires a minimum of four hosts per cluster, for a total of twelve hosts. This sizing recommendation provides balance between the implementation footprint and resiliency, while maintaining the operational requirements necessary for each pod.

The resource and edge clusters are sized in accordance with the VNFs and their respective networking requirements. CSPs must work with the VNF vendors to gather requirements for the VNF service to be deployed. This information is typically available in deployment guides and sizing guideline documents.

As more tenants are provisioned, CSPs must make additional resources available to the Resource pod to support VNF growth. The increase in VNF workloads in the Resource pod can imply an increase in North-South network traffic, which in turn requires the scaling up of the ESG in the Edge pod by adding compute resources. CSPs must closely monitor and manage the resource consumption and capacity availability of the Edge pod as the ESGs are scaled. Figure 13 provides an overview of three-pod design.

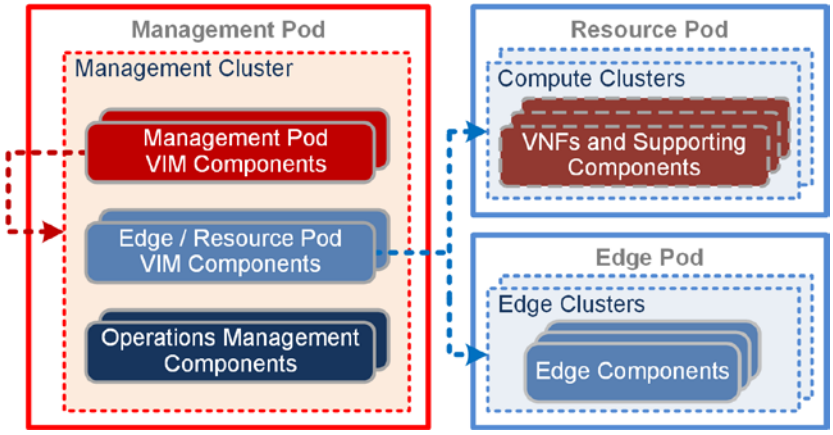

Figure 13: Three-pod Design Overview

4.4.1 Three-pod VMware vCenter Design

The three-pod VMware vCenter design separates the management, edge, and resource functions into their own dedicated pods. The three-pod vCenter Server differs from that in a two-pod design, in that the second vCenter Server instance manages separate clusters for the resource and edge functions. Figure 14 shows the three-pod VMware vCenter design.

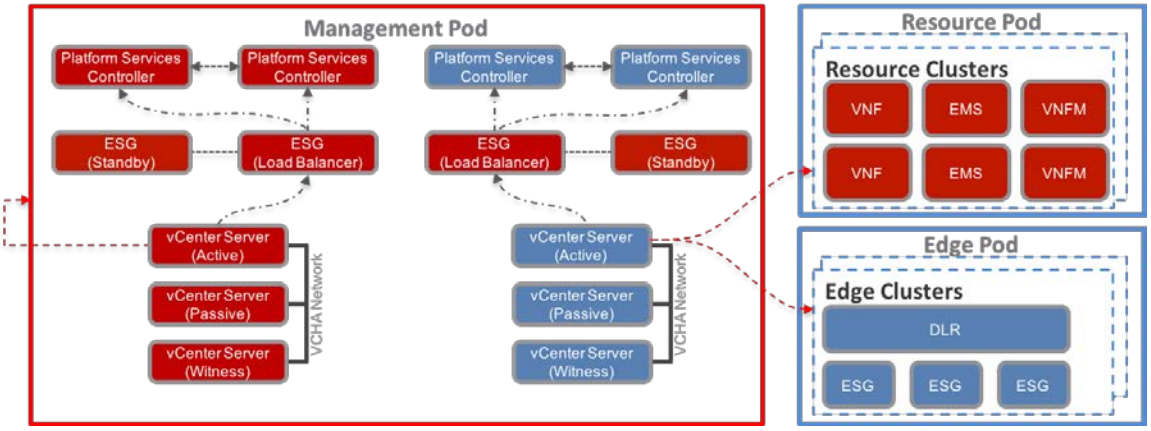

Figure 14: VMware vCenter Three-pod design

With three-pod design, each pod is implemented as a cluster. This enables the NFVI operator to specify different cluster-level configurations for the edge and compute clusters.

The Management pod of a three-pod design is almost identical to that of a two-pod design. The only exception in three-pod design is that the vCenter Server manages two clusters, one for the Resource pod and one for the Edge pod respectively, while the same vCenter Server in two-pod design manages only the collapsed Edge / Resource pod.

4.4.2 Three-pod Virtual Networking Design Using VMware NSX Manager

Figure 15 shows the role of the NSX Manager in three-pod design.

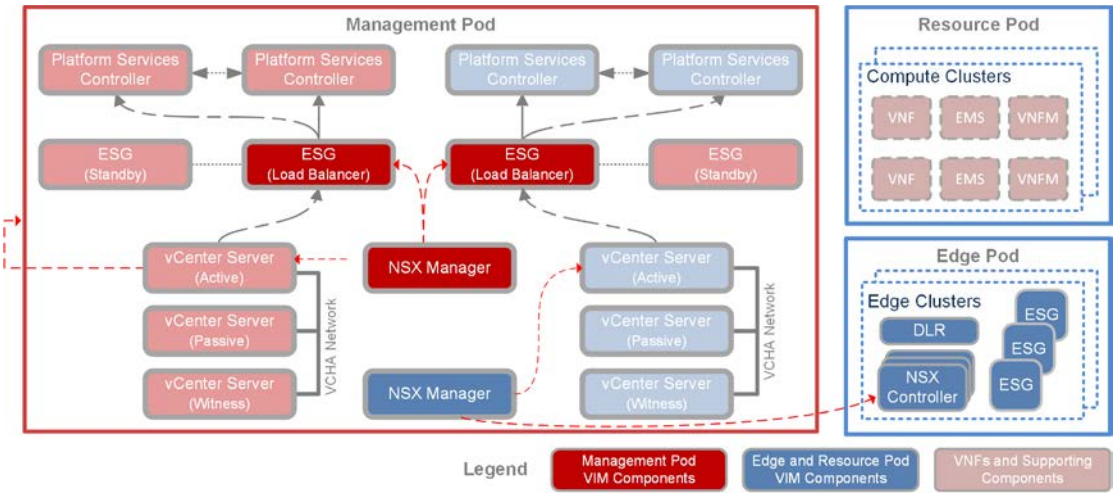

Figure 15: VMware NSX Manager NSX Manager Design in a Three-pod design

The second NSX Manager in the Management pod is associated with the vCenter Server that manages the Edge pod. This instance of NSX Manager is used by tenants to deploy tenant networks and tenant network services from within their tenancy in VMware Integrated OpenStack. Although tenants can deploy their own edge devices for North-South connectivity, typically they must coordinate their connectivity needs with the CSP network administrator during VNF onboarding.

For VNF components that require North-South connectivity, logical switches are routed to the telecommunications network through the edge services deployed in the Edge pod. A VXLAN transport zone is created between the Resource pod and the Edge pod, which allows logical switches to seamlessly interconnect the VNFs in the Resource pod to the edge networking services in the Edge pod. Figure 16 shows the design for Edge pod and Resource pod connectivity.

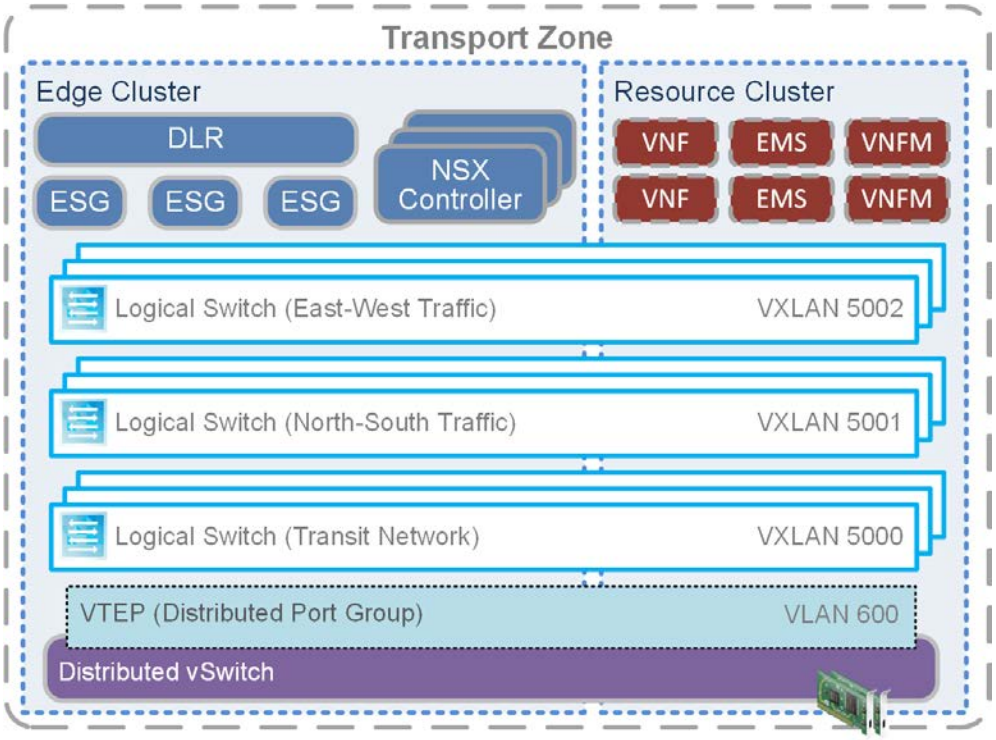

Figure 16: Virtual Networking Design for Edge Pod and Resource Pod Connectivity

Different routing scenarios are detailed in the *VNF Onboarding* section of this document. These scenarios address the connectivity between VNFCs inside a VNF, connectivity among different VNFs, and connectivity between VNFs and their management components.

Each tenant has direct access over the compute, network, and storage resources allowing them to create multiple routings, switching, and security policies per tenant. Network services such as routing, switching, load balancing, firewalling, and micro-segmentation are possible. The logical routing design is based on per VNF requirements and must be decided as part of the onboarding process. Figure 17 provides an example of the proposed port group configuration for a three-pod design.

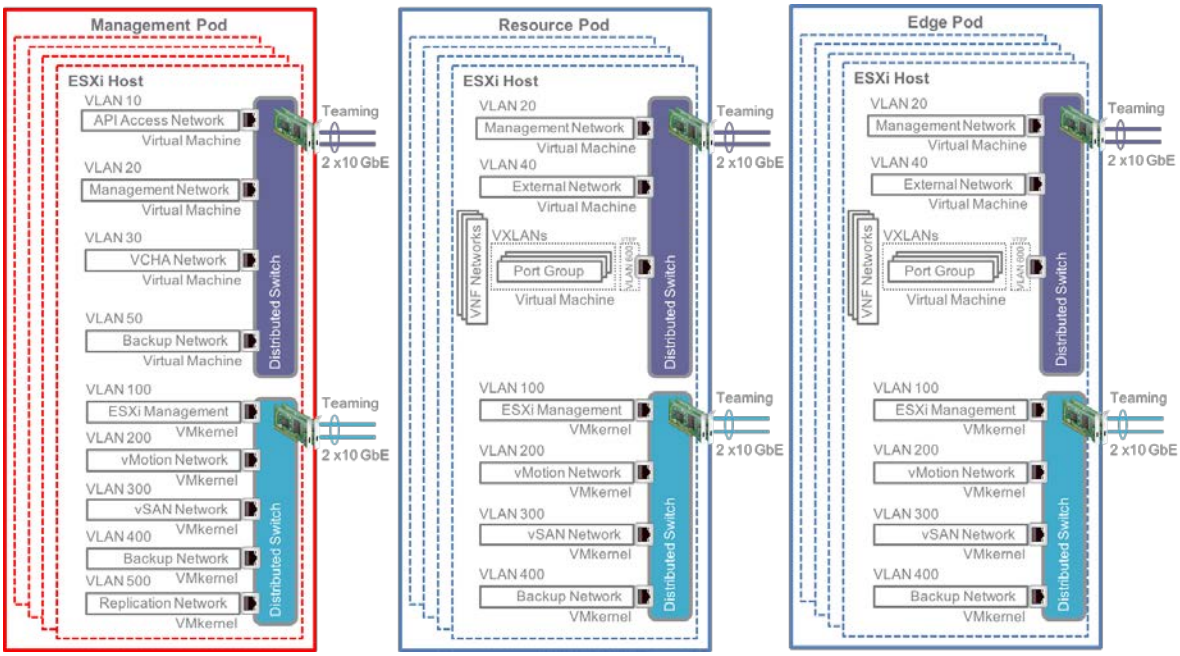

Figure 17: Distributed Virtual Switch Three-pod design

4.4.3 Three-Pod VMware Integrated OpenStack Design

Figure 18 shows the VMware Integrated OpenStack deployment of the management plane components for three-pod design.

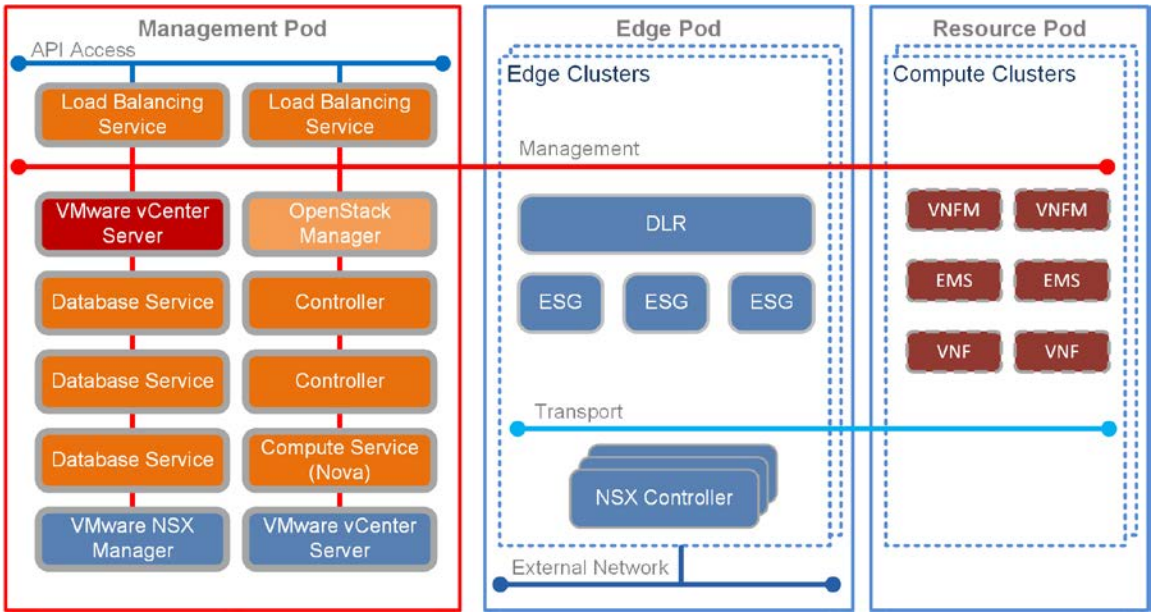

Figure 18: VMware Integrated OpenStack Three-pod design

In three-pod design, the Integrated OpenStack Manager vApp is used to instantiate three-pod deployment of the VMware Integrated OpenStack management plane. VMware Integrated OpenStack connects to the second vCenter Server instance that manages the Edge pod and the Resource pod. It connects to the NSX Manager instance associated with the Edge pod to deploy the edge networking components for North-South connectivity, and to the Resource pod for provisioning logical switches for East-West connectivity.

As with two-pod design, the Integrated OpenStack Manager vApp can be used to deploy a compact instance of the VMware Integrated OpenStack management plane for cases where a smaller footprint is required, when the functionality of the full deployment and high availability is not necessary.

Further details about edge services and their consumption by tenants is described in the *Secure Multitenancy* section of this document.

4.5 Two-pod Versus Three-pod Design

Given the two design options available, a greenfield deployment will be faced with a choice regarding the most appropriate vCloud NFV design to use. As an engineering decision, differences between the two designs and their potential use cases must first be understood. This section of the document provides considerations to guide CSPs in choosing between two-pod and three-pod design.

The vCloud NFV OpenStack Edition architecture ensures that key capabilities like secure multitenancy, integrated operational management, and carrier grade readiness are untouched by the choice of pod design. Neither option influences the ability of the NFV platform to support virtualized network services.

A key difference between two-pod and three-pod design is the footprint required for deployment. Two-pod design initially requires a smaller number of hosts, racks, and ToR switches than three-pod design. This means that distributed deployments such as those used in micro data centers or telecommunications central offices, where space and cooling is at a premium, can benefit from two-pod design. Some enterprise service use cases, such as virtual Customer Premise Equipment (vCPE), also benefit from the smaller footprint of two-pod design. In these use cases administrative boundaries can be made very clear by mapping the collapsed Edge / Resource pod to a single rack. In use cases where space is ample and virtual network functions perform

a centralized role, maintaining the functional separation between the three pods is beneficial.

Capacity planning and scale up operations are straight forward in three-pod design. With this design, each pod scales independent of the others. In two-pod design with both edge and resource functions sharing the same pod, as VNFs are added careful consideration must be taken of the resources available to edge function operations. All the tools required for capacity planning and proactive resource usage monitoring are provided with vCloud NFV OpenStack Edition. Tools to migrate VNFs to necessary resources are also available.

4.6 Secure Multitenancy

Together, the vCenter Server, NSX Manager, and VMware Integrated OpenStack form the secure multitenant platform of the vCloud NFV design. VMware Integrated OpenStack provides the abstraction layers for secure multitenancy in a three-pod design, in the same way as it does for a two-pod design. vCenter Server provides the infrastructure for fine grained allocation and partitioning of compute and storage resources, while NSX for vSphere creates the network virtualization layer. The network virtualization layer is an abstraction between physical and virtual networks. NSX for vSphere provides logical switches, firewalls, load balancers, and VPNs.

VMware Integrated OpenStack divides pooled resources among tenants to create a secure tenant virtual data center, the Tenant vDC. CSPs can improve their capacity planning for resource allocation to the tenant with resource-level tenant isolation and guaranteed resource availability. This is created for each tenant, while simultaneously securing tenants within the network. This section of the document describes how the VMware Integrated OpenStack abstraction layers, Project and Tenant vDC, are leveraged to provide a secure multitenant environment to deploy VNFs. Figure 19 provides an overview of secure multitenancy.

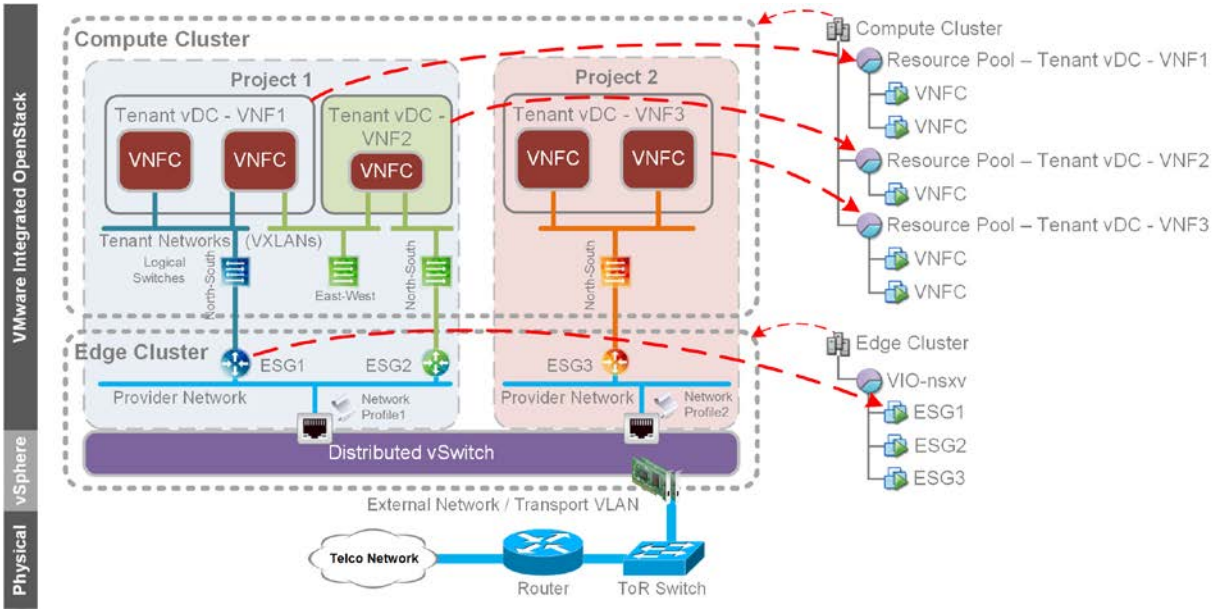

Figure 19: VMware vCloud NFV OpenStack Edition Multitenant Networking in Three-Pod Design

Physical compute, storage, and network resources are first mapped to the NFVI virtual resources – clusters for compute resources, data stores for storage resources, and virtual switches for network resources. The virtual resources are managed by VMware Integrated OpenStack for consumption by tenants.

The CSP allocates and reserves resources for tenants using the VMware Integrated OpenStack Tenant vDC construct. Every Tenant vDC is associated with a resource pool within the compute cluster. The resource settings of the resource pool are managed by the CSP from within VMware Integrated OpenStack. This ensures that every Tenant vDC is allocated the resources to which it is entitled, while not exceeding the resource limits.

Tenant edge devices that are deployed using VMware Integrated OpenStack are placed in the dedicated Edge cluster in a three-pod design, and in the compute cluster in a two-pod design. VNFs are deployed in a separate and dedicated resource pool nested within the compute cluster. This separation of edge devices from VNF workload resources prevents one from starving the other.

The separation of network access between NFVI tenants is important to support secure multitenancy on a horizontally shared platform. VMware Integrated OpenStack integrates with vCenter Server and NSX for vSphere to manage the creation and consumption of isolated Layer 2 networks. CSPs must ensure that the necessary connectivity to external networks is in place for consumption by tenants. Networks that are internal to an NFVI tenant, or to a VNF instance, can be created using the VMware Integrated OpenStack user interface or API. As described in the *Virtual Networking Design Using VMware NSX Manager* section of this document, ESG firewall rules and additional services can be configured by the tenant from within the Tenant vDC.

4.7 Operations Management

The NFVI Operations Management components are a functional block in the Management pod. These components are responsible for providing and extending full visibility to fault, configuration, accounting, performance, and security (FCAPS) of the NFVI, and when needed the Virtual Network Functions (VNFs). VMware implementation of the vCloud NFV OpenStack Edition platform expands the capabilities of this functional block by offering OpenStack content packs for vRealize Log Insight, content packs for VMware vRealize Operations Manager, and business continuity and disaster recovery capabilities. Disaster recovery is discussed in the *Business Continuity and Disaster Recovery* section of this document.

For the VMware Integrated OpenStack NFV platform, the NFVI Operations Management tasks are delivered using the components listed in Table 3. All components are deployed in the Management pod.

COMPONENT NAME	DESCRIPTION
VMware vRealize Operations Manager	VMware vRealize Operations Manager handles performance and capacity management of the NFVI and VIM components. It is the primary network operations center (NOC) NFVI management console.
VMware vRealize Log Insight	VMware vRealize Log Insight provides real-time log management and log analysis with machine learning based intelligent grouping, high-performance search, and targeted troubleshooting across physical, virtual, and cloud environments.
VMware vRealize Network Insight	VMware vRealize Network Insight provides visibility and analytics into the networking aspects of the NFVI. It monitors network performance and availability across virtual and physical networks, provides visibility into network communication between VNF components, and extends visibility into external network paths, Internet access, and VXLAN.

Table 3: NFVI Operations Management Components

4.7.1 Operations Workflow

Management of the NFV environment is driven by the three tools described in this section of the document: vRealize Operations Manager, vRealize Log Insight, and vRealize Network Insight. The NOC primarily interacts with vRealize Operations Manager as a single pane of glass, while using the other tools for issue isolation, remediation, and planning.

The vRealize Operations user interface can be configured in various ways, however the main pane informs the NOC personnel about three categories:

- **Health**. The current health of the system is displayed in this tab. Red alerts here indicate that an immediate issue is taking place.
- **Risk**. Future issues, based on deep machine learning and analytics, are displayed in this tab. Risks indicate future performance or capacity problems that can become health issues. Proactively resolving risks is the best approach to maintaining high quality services.
- **Efficiency**. This area indicates optimization opportunities based on the way the platform is used. If the operator follows these recommendations, NFVI resources used in a wasteful way or suboptimally configured can be recovered and platform efficiency increased.

The NFVI operator first focuses on maintaining the healthy state of the environment. When a vRealize Operations Manager Health Badge reports red, a critical issue is raised and an indication of the cause is provided. To resolve the issue, further detail is specified through the vRealize Operations graphical user interface. These details are collected using vRealize Log Insight. The operator correlates network information using vRealize Network Insight to speed up issue resolution. In combination, these three tools ensure that all layers of the NFVI environment are monitored, and that issues are quickly isolated and remediated.

vRealize Operations Manager monitors performance and capacity by collecting information exposed by the NFVI devices, including the number of hosts, virtual machines, physical cores, and vCPUs used. vRealize Operations Manager also collects information about networking components, including interface utilization, packet drop rate, observed throughput, storage information as read and write performance, usage rate, and total capacity. The performance and capacity data collected provide a holistic system view that can be used to manually address issues and perform capacity planning. Distributed Resource Scheduler (DRS) can be used to automatically balance VNF components based on performance needs and capacity availability, eliminating resource contention that might otherwise occur.

In vCloud NFV OpenStack Edition, future resource contention is evaluated based on continuous monitoring. Coupled with vRealize Operations Manager dynamic thresholds, which understand the behavior of VNFs throughout the day, calculations are run to create a band describing normal operations for each metric and object combination. The band includes an upper and lower boundary for each metric associated with an object, and is tailored specifically to the individual VNF component, providing data about the amount of resources the component requires throughout the day. Together with an understanding of the size of the NFVI hosts and their aggregated resources, the operator can predict where contention will occur and can balance the VNFs accordingly. Network utilization is one of the new data points added to the DRS considerations in vCloud NFV OpenStack Edition in addition to storage, CPU, and memory.

Network aware DRS is fundamental to the multitenancy characteristic of the platform. Using Network I/O Control (NIOC) it is possible to set a network bandwidth reservation for a VNF component and have DRS consider the reservation in resource placement. This means that when network capacity planning is performed for tenants, the operator can rely on DRS to ensure that tenants are not consuming other tenants' network capacity.

Figure 20 shows the overall vCloud NFV Operations Management design. In the following sections of this document, we discuss the design of each individual operations management component.

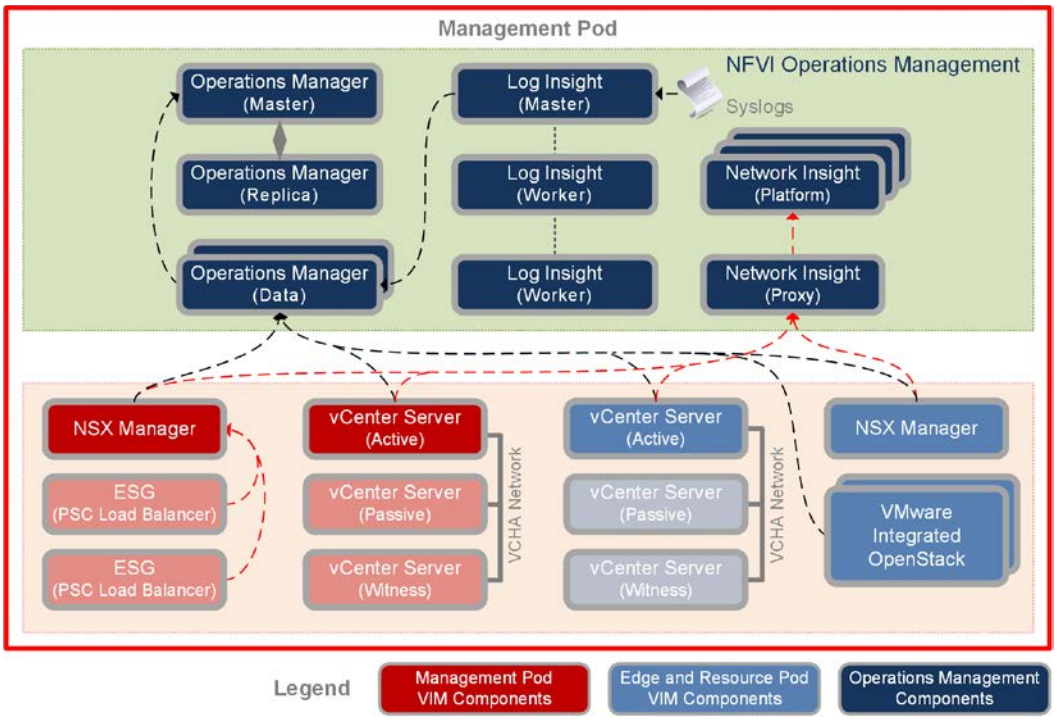

Figure 20: VMware vCloud NFV Operations Management Design

4.7.2 VMware vRealize Operations Manager

The vCloud NFV virtual infrastructure relies on a monitoring solution able to collect data regarding its health, capacity, availability, and performance. vRealize Operations Manager provides a robust and integrated monitoring platform that sits at the center of the NFV environment. As described in the *Operations Workflow* section of this document, it serves as the single pane of glass into the NFV environment.

vRealize Operations Manager is installed in the Management pod in both two-pod and three-pod design. As it collects data over time, it is possible that additional storage capacity will be required. Adding more hosts to the management cluster, or simply adding more storage, is sufficient to address the growing vRealize Operations Manager storage needs.

Since vRealize Operations Manager is the central management component in the vCloud NFV OpenStack Edition platform, its availability to the operator is essential. vRealize Operations Manager supports high availability (HA). HA creates a master replica for the vRealize Operations Manager master node and protects the analytics cluster against the loss of a node. With HA, data stored on the master node is always completely backed up on the master replica node. To enable HA, at least one other data node must be deployed in addition to the master node. Further information is provided in the vRealize Operations Manager vApp Deployment and Configuration Guide.

VMware vRealize Operations Management Packs

VMware vRealize Operations Manager collects structured data from various sources, including gathering data from adapters connected to source servers. For this mechanism to work, vRealize Operations Manager is configured to communicate with source servers using an authorized user account. If the user account has limited access to objects in the source server, it sees only the data for which the account has permissions. At a minimum, the user account must have read privileges across the objects from which it collects data. A collection of management packs is available on the VMware Solution Exchange marketplace.

To minimize traffic between the vCenter Server and the vRealize Operations Manager, the vCenter Server Adapter is installed with a five minute collection interval.

VNF vendors can create plug-ins, which are interfaces between vRealize Operations Manager and external components that require management. Plug-in development requires an understanding of the vRealize Operations Manager inventory model, the management functions that plug-ins implement. These include auto-discovery and monitoring. Additional information is available in the Endpoint Operations Management Agent Plug-in Development Kit document.

VMware vRealize Operations Management Pack for OpenStack

The VMware vRealize® Operations Management Pack™ for OpenStack allows the operator to quickly view the health of the environment, including services running within the VMware Integrated OpenStack infrastructure. The vRealize Operations Management Pack for OpenStack integrates with NSX, which allows for easy monitoring and management of the network infrastructure. vRealize Operations Management Pack for OpenStack includes dashboards to provide visibility into VMware Integrated OpenStack deployments, including:

- VMware Integrated OpenStack Management Services
- VMware Integrated OpenStack Compute Infrastructure
- VMware Integrated OpenStack Network Infrastructure
- VMware Integrated OpenStack Tenants
- VMware Integrated OpenStack Storage

Figure 21 shows an example from the OpenStack Tenants dashboard in vRealize Operations Manager. Four window panes are highlighted. The tenant “Luling” has deployed one network with one Ubuntu instance. Pane 1 shows the Tenant Quota Usage, and displays three types of information: the instances, vCPUs, and vRAM used by the tenant. Pane 2 shows the Tenants List, which displays a list of all of the tenants. Pane 3 displays alerts for the selected tenant and pane 4 shows the inventory of items for the selected tenant.

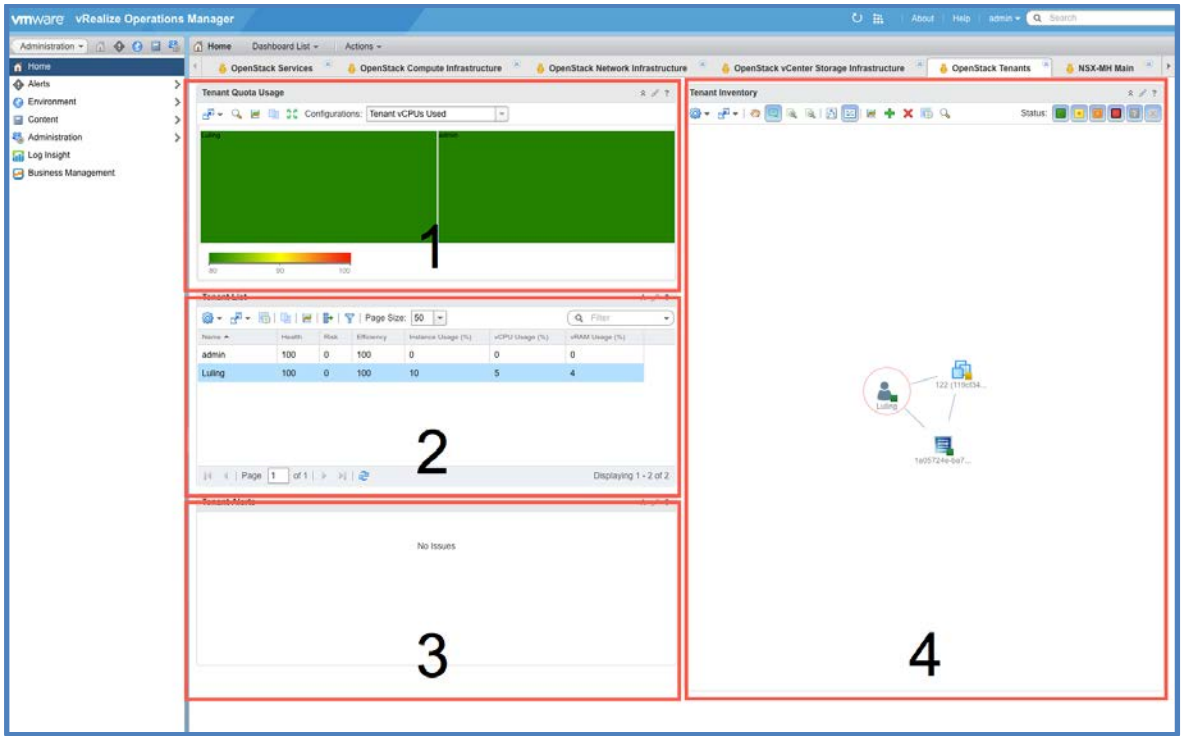

Figure 21: VMware vRealize Operations Manager VIO Dashboard

vRealize Operations Manager logging and alerts assist in quickly identifying issues and configuration changes:

- **VMware vRealize Operations Manager Logging**. vRealize Operations Manager logging compiles audit logs, the logs used to track configuration changes performed by authenticated users. They are useful in identifying which user initiated a change or scheduled the job to perform a change. All audit logs are forwarded to vRealize Log Insight.
- **VMware vRealize Operations Manager Alerts**. When alerts are generated in vRealize Operations Manager, they appear in the alert details and object details windows. Alerts can be configured and sent to other applications using one or more outbound alert options.

To configure notification options the operator must specify which alerts are sent out with standard email, REST, SNMP traps, and log file outbound alert plug-ins. The decision to use a specific alerting method is implementation specific and is typically driven by the external monitoring and management tools available.

4.7.3 VMware vRealize Log Insight

VMware vRealize Log Insight is used to collect log data from ESXi hosts, and to connect to vCenter Servers to collect server events, tasks, and alarm data. vRealize Log Insight integrates with vRealize Operations Manager to send notification events. Since vRealize Log Insight collects real-time unstructured data, all elements in the NFV environment can be configured to send it their log data for aggregation, creating a single log collector for the NFV environment.

VMware vRealize Log Insight is deployed in the Management pod using a single cluster configuration, which consists of a minimum of three nodes leveraging the vRealize Log Insight Integrated Load Balancer (ILB). Each log message is present in only one location within the cluster at a time. The cluster remains available to ingest data and serve queries during any temporary unavailability of a single node.

Data is collected using either the syslog protocol or an API. All NSX Manager syslog information, distributed firewall logs, and NSX Edge Services Gateway syslog information is sent to vRealize Log Insight. VMware Integrated OpenStack troubleshooting and API access logs are stored locally. These logs can be forwarded by creating an additional logger that sends diagnostics logs to vRealize Log Insight.

VMware vRealize Log Insight Archiving

Archiving is primarily a long term retention tool. The process copies raw data to an external NFS storage location. Archives are much smaller than indexes, but require indexing if they are to be loaded back into the vRealize Log Insight system. For additional information about vRealize Log Insight archiving, see the VMware vRealize Log Insight Information page.

VMware vRealize Log Insight Content Pack

vRealize Log Insight gathers log events from multiple sources, and through special content packs delivers solution specific dashboards to perform log analytics, using redefined alerts. For additional information about vRealize Log Insight solutions, see the VMware Solution Exchange marketplace.

VMware vRealize Log Insight Content Pack for OpenStack

VMware Integrated OpenStack is very log intensive, so trying to troubleshoot across the layers is difficult without central logging. vRealize Log Insight provides an interactive logging environment where administrators can perform real-time queries on VMware Integrated OpenStack values. A special vRealize Log Insight OpenStack Content Pack is provided for free download through the VMware Solution Exchange marketplace.

The vRealize Log Insight Content Pack for OpenStack offers pre-configured dashboards that provide quick insight into logs across the infrastructure. An overview of the events in the OpenStack environment is also available. This dashboard allows the administrator to view a list of events grouped by component and severity. When OpenStack events must be reviewed quickly, interactive analytics provide the ability to dive deep by applying filters that enable real-time queries into a specific event.

The VMware vRealize Log Insight OpenStack Content Pack provides several dashboards, including:

- OpenStack Overview dashboard shows events grouped by component and severity over time.
- The Errors dashboard shows both errors grouped by component, and the number of errors per VMware Integrated OpenStack component.
- The API Requests dashboard groups API requests in various ways, including by user ID, tenant ID, hostname, and status code.
- The API Response Time dashboard can be used to review response times from Nova, Neutron, Cinder, Glance, Keystone, and Heat.

Individual dashboards exist for each of the individual VMware Integrated OpenStack components.

Figure 22 shows an example of the output provided by the vRealize Log Insight OpenStack Content Pack overview dashboard, based on information collected by vRealize Log Insight.

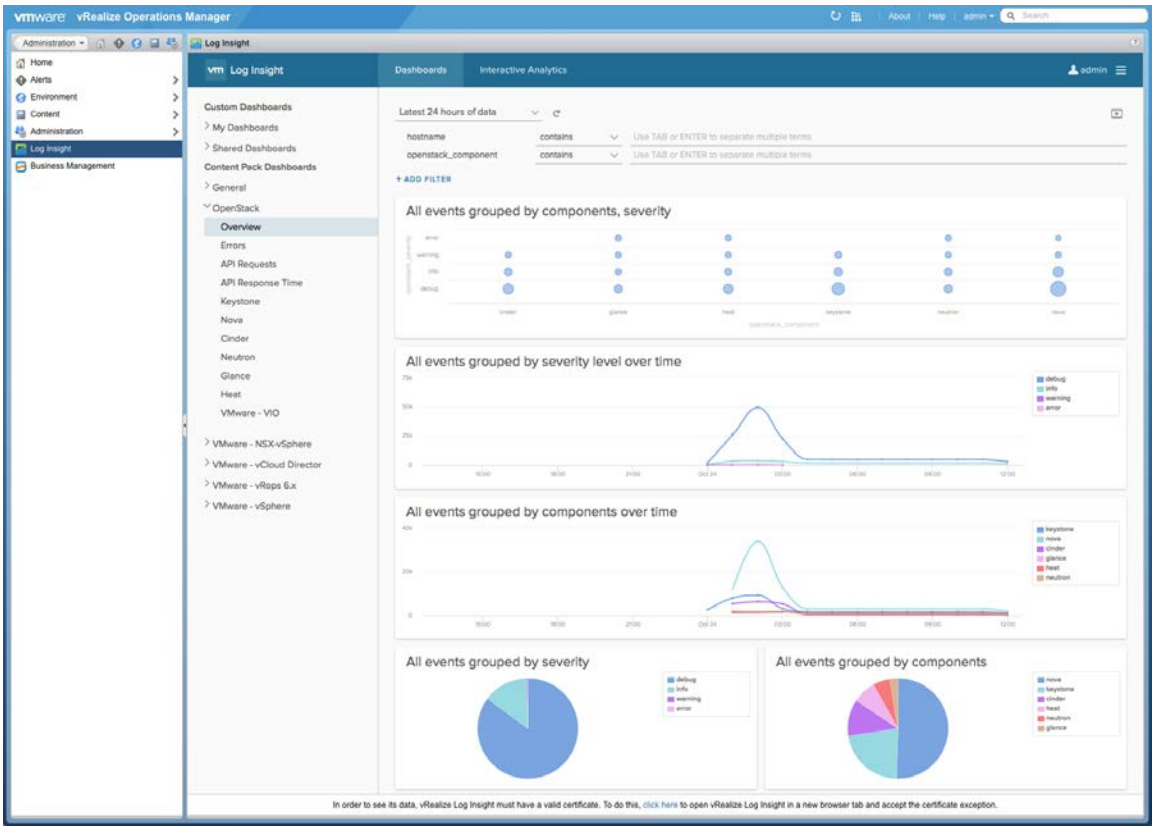

Figure 22: VMware vRealize Log Insight OpenStack Dashboard from VMware vRealize Operations.

This dashboard provides a complete overview of all VMware Integrated OpenStack components, grouped by severity. A trend line shows the different types of events over time, grouped by severity.

Figure 23 shows an example of the output from the vRealize Log Insight OpenStack Errors dashboard.

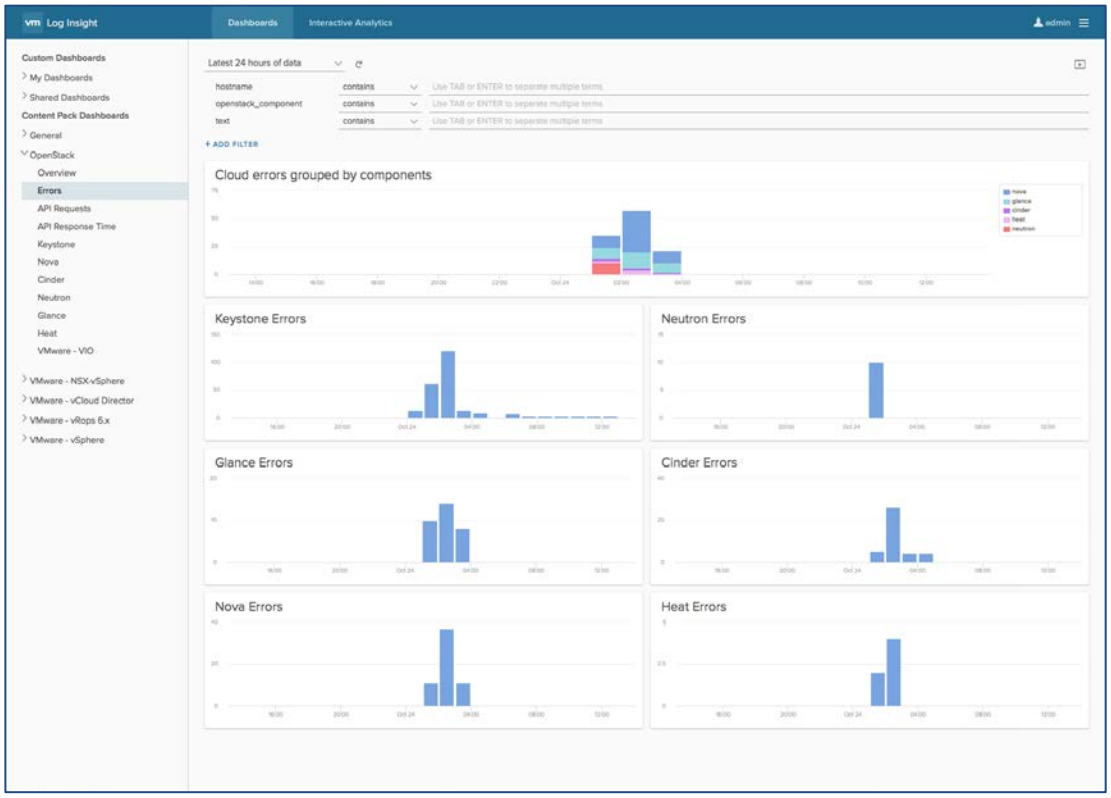

Figure 23: VMware vRealize Log Insight OpenStack Errors Dashboard

This dashboard provides an overview of all errors, grouped by component. With the interactive analytics capabilities of vRealize Log Insight, the administrator can, with a single click, retrieve additional error information. This example's output shows that Nova has a large number of errors.

Figure 24 shows an example of the detailed output of the Nova errors.

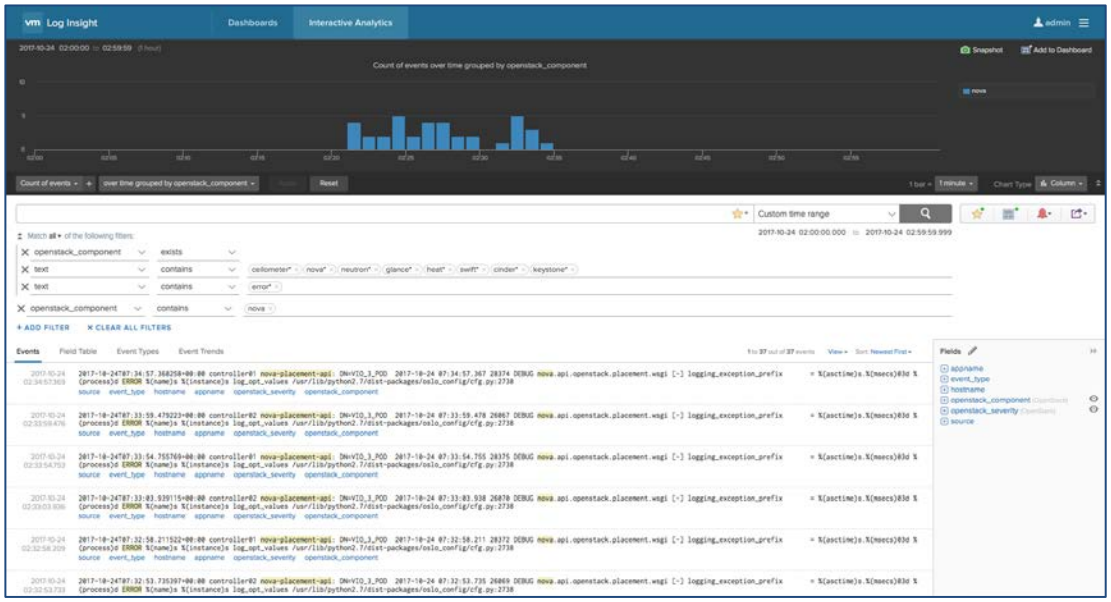

Figure 24: VMware vRealize Log Insight Interactive Analysis of Nova

The interactive analytics view of Nova shows the preset filters with which to view error information. With the Interactive Analytics dashboard, administrators can change the filters in real time to tailor output to their needs.

4.7.4 VMware vRealize Network Insight

vRealize Network Insight is part of the data center operations and analytics suite of products used to monitor the vCloud NFV platform. vRealize Network Insight offers intelligent operations for software-defined networking and security across virtual and physical infrastructure with micro-segmentation planning that can be scaled to thousands of VNFs, offering 360-degree visibility. vRealize Network Insight is installed in the Management pod in both two-pod and three-pod design.

vRealize Network Insight is configured to monitor and connect to all networking-related components in the NFVI. It provides valuable insight into the networking segements used to deploy and manage the vCloud NFV platform. Management VLANs, external VLANs, and VXLAN segments are all available for monitoring and diagnostics. vRealize Network Insight can also be configured to monitor a myriad of physical devices including Dell, Cisco Nexus, Catalyst, Arista, Juniper Networks, Hewlett Packard Enterprise, Brocade, and Palo Alto Networks switches.

The vRealize Network Insight architecture consists of a platform VM, a proxy VM, and data sources. Within the architecture the role of the platform VM is to perform analytics, storage, and to provide a user interface into the data. The proxy VM, known as the collector, collects data from sources using various protocols including HTTPS, SSH, CLI, and SNMP depending on the source and its configuration. A variety of data sources are supported, including VMware vCenter, NSX, firewalls, and various switch vendors. For complete visibility of the NFV environment, vRealize Network Insight is connected to the vCenter Server that is used to operate the edge and resource clusters.

4.7.5 Business Continuity and Disaster Recovery

Business continuity and disaster recovery solutions are an integral part of the vCloud NFV OpenStack Edition platform. The three components in Table 4 are used to achieve a robust solution.

COMPONENT NAME	DESCRIPTION
VMware vSphere Replication	VMware vSphere Replication is a hypervisor based asynchronous replication solution that provides granular replication and recovery of management components.
VMware Site Recovery Manager	VMware Site Recovery Manager is a disaster recovery management and orchestration engine for providing predictable failover of management components.
VMware vSphere Data Protection	VMware vSphere Data Protection provides data protection by performing backup and recovery of management components.

Table 4: NFVI Business Continuity and Disaster Recovery Components

The methods for using these business continuity and disaster recovery tools to ensure healthy operations in the NFV environment are described in the following sections of this document. While a multi-site design will be the subject of a future document, this reference architecture provides an overview of the business continuity capabilities built into vCloud NFV OpenStack Edition.

4.7.5.1. VMware vSphere Replication

vSphere Replication is the technology used to replicate virtual machine data between data center objects, within a single site or across sites. The two most important aspects to consider when designing or executing a disaster recovery plan are the Recovery Point Objective (RPO) and Recovery Time Objective (RTO). RPO is the duration of acceptable data loss. It is fulfilled by the replication technology. RTO is a target duration with an attached service-level agreement, during which the business process must be restored. The RTO includes the time for the recovery and service readiness, in a state for business to operate as usual.

vSphere Replication provides the ability to set the RPO, however RTO is application dependent. For an appliance deployed within the management cluster, vSphere Replication provides an RPO of five minutes to 24 hours.

vSphere Replication fully supports vSAN.

4.7.5.2. VMware Site Recovery Manager

Site Recovery Manager provides a solution for automating the recovery and execution of a disaster recovery plan in the event of a disaster in a data center. When a catastrophe occurs, components in the Management pod must be available to recover and continue the healthy operations of the NFV-based services.

To ensure robust business continuity and disaster recovery, network connectivity between the protected and recovery site is required, with enough bandwidth capacity to replicate the management components using vSphere Replication. Each site must have an instance of vCenter Server that governs the Management pod and its ESXi hosts, a Site Recovery Manager server, and a vSphere Replication appliance to orchestrate the disaster recovery workflows and replicate content across the sites. The protected site provides business critical services, while the recovery site is an alternative infrastructure on which services are recovered in the event of a disaster.

Networking Considerations

Moving a service from one site to another represents a networking challenge for maintaining IP addressing, security policies, and the bandwidth to ensure ample network capacity. Some of these challenges, such as IP addressing, are managed using NSX for vSphere.

Distributed Resource Scheduler Considerations

Some management components for the vCloud NFV OpenStack Edition platform such as NSX for vSphere, Edge Services Gateway, PSCs, vRealize Operations Manager, and vRealize Log Insight have specific affinity or anti-affinity rules configured for availability. When protected management components are recovered at a recovery site DRS rules, reservations, and limits are not carried over as part of the recovery plan. However, it is possible to manually configure rules, reservations, and limits on placeholder VMs at the recovery site during the platform build.

Inventory Mappings

Elements in the vCenter Server inventory list can be mapped from the protected site to their vCenter Server inventory counterparts on the recovery site. Such elements include VM folders, clusters or resource pools, and networks. All items within a single data center on the protected site must map to a single data center on the recovery site.

These inventory mapping details are used across both the protected and recovery sites:

- **Resource Mapping**. Resource mapping maps cluster objects on the protected site to cluster objects on the recovery site.
- **Folder Mapping**. Folder mapping maps the folder structures like data centers or VM folders on the protected site to folder structures on the recovery site.
- **Network Mapping**. Network mapping maps the management networks on the protected site to management networks on the recovery site.

VNF Recovery Considerations

Every vendor must provide a specific strategy for disaster recovery for any VNF managed directly by the VNF Managers.

Protection Groups

A protection group is a group of management components at the protected site that can fail over together to the recovery site during testing and recovery. All protected management components are placed within a single protection group.

Recovery Plans

Recovery plans are the run books associated with a disaster recovery scenario. A recovery plan determines which management components are started, what needs to be powered down, which scripts to run, the startup order, and the overall automated execution of the failover.

A complete site failure is the only scenario that invokes a disaster recovery. There is no requirement for recovery plans to handle planned migrations or to move a single failed application within the management cluster. A single recovery plan is created for the automated failover of the primary site, and the placement of management components into priority groups ensures the correct startup order.

The recovery of the resource cluster, edge cluster, vCenter Server, and NSX Manager are required to maintain management capabilities where additional physical data centers are managed within the site.

4.7.6 VMware vSphere Data Protection

vSphere Data Protection is a disk-based backup and recovery solution that is fully integrated with vCenter Server and vSphere Web Client to enable centralized management of backup and restore tasks, while storing backups in deduplicated backup storage. Managing the backup and restore tasks is accomplished through the vSphere Data Protection user interface, an add on plug-in to the vSphere Web Client.

vSphere Data Protection creates image-level backups, which are integrated with the VMware vSphere® Storage APIs - Data Protection, a feature set within VMware vSphere used to offload the backup processing overhead from a virtual machine to the vSphere Data Protection appliance. The vSphere Data Protection appliance communicates with the vCenter Server to create a snapshot of the `.vmdk` files in a virtual machine. Deduplication takes place within the appliance, using a patented variable length deduplication technology.

vSphere Data Protection is distributed in a prepackaged OVA file. The vSphere Data Protection appliance is responsible for the backup and restore of the management components residing within the management cluster. vSphere Data Protection is configured so the appliance backs up the data to a deduplicated backup datastore, which is separate from the datastore hosting management components.

vSphere Data Protection protects the management cluster through the vCenter Server management layer. Connectivity through vCenter Server provides visibility to all ESXi servers in the management clusters, and therefore to all management components that require backup.

The initial configuration of the vSphere Data Protection appliance must be set to 6 terabytes. Additional disk space above the usable capacity of the appliance is required to create and manage checkpoints. Backups can be configured to protect required components of the management cluster. In a disaster recovery or data loss event, the protected components can be restored to resume normal services, based on the RPO. The target location must meet the minimum performance requirements for mitigation.

These issues must be considered when configuring backups for the vCloud NFV platform:

RPO. vSphere Data Protection can perform daily backups at scheduled intervals for required components within the management cluster. The RPO value and the backup start time must be set based on the business need. Schedule backups during off-peak business hours.

Retention Policies. Retention policies are the properties of a backup job. It is important to group management components in a backup job by business priority and by the retention requirements set based on the business need.

Monitoring. CPU, memory, network, disk performance, and the capacity of the vSphere Data Protection appliance are monitored by vRealize Operations Manager, with syslog events sent to vRealize Log Insight. Capacity can be viewed through vSphere Data Protection reports.

For backup of VMware vCenter components and NSX Manager data, use the respective inbuilt backup mechanisms.

4.8 Carrier Grade

Carrier grade attributes are injected into every layer of the vCloud NFV OpenStack Edition platform.

High availability, a core pillar of the architecture and a crucial requirement in every CSP network, spans the entire platform. The design highlights various redundancy mechanisms used by the management components. This document also describes the ways in which networking elements can be made highly redundant with the use of routing protocols and NIC teaming. The NFVI virtualization layer also provides mechanisms to enhance and improve the availability of Virtual Network Functions. These include VMware vSphere Fault Tolerance (FT), VMware vSphere® High Availability (HA), and Orchestrated HA.

Tuning the vCloud NFV OpenStack Edition platform for performance spans the NFVI, VNFs, and VIM. However, since it is applicable to specific VNFs, performance architecture recommendations are grouped in this section of the document.

4.8.1 Performance

ETSI classifies NFV workloads into three categories: management, control, and data plane. For data plane intensive VNFs hosted on the vCloud NFV OpenStack Edition platform, specific design considerations are provided in the following section of this document. A few of the key performance capabilities available within the platform are CPU pinning, NUMA placement, HugePages support, and Direct Passthrough SR-IOV support allow CSPs to maintain high network performance.

4.8.1.1. Data Plane Intensive Design Framework

Two parties are involved in the successful deployment and operation of a data plane intensive VNF: the VNF vendor and the NFVI operator. Both parties must be able to understand the performance requirements and design of the VNF. They must also be willing to tune the entire stack from the physical layer to the VNF itself, for the demands data plane intensive workloads place on a system. The responsibilities of the two parties are described as follows:

Virtual Network Function Design and Configuration. The vendor supplying the VNF must tune the performance of the VNF components and optimize their software. Data plane intensive workloads benefit from the use of a Data Plane Development Kit (DPDK) to speed up VNFC packet processing and optimize the handling of packet off-loading to the virtual NIC. Use of the VMware VMXNET3 paravirtualized network interface card (NIC) is a best practice designed for performance demanding VNFs. VMXNET3 is the most advanced virtual NIC on the VMware platform and has been contributed to the Linux community, making it ubiquitous in Linux distributions.

Once the VNF is created by its supplier, there are several VNFC-level configurations that are essential to these types of workloads. Dedicated resource allocation, for the VNFC and the networking-related processes associated with it, can be configured and guaranteed through the use of two main parameters: Latency Sensitivity and System Contexts. Both parameters are discussed in detail in the Tuning vCloud NFV for Data Plane Intensive Workloads whitepaper.

Another aspect essential to the performance of a data plane intensive VNF is the number of virtual CPUs required by the VNFC. Modern multiprocessor server architecture is based on a grouping of resources, including memory and PCIe cards, into Non-Uniform Memory Access (NUMA) nodes. Resource usage within a

NUMA node is fast and efficient. However, when NUMA boundaries are crossed, due to the physical nature of the QPI bridge between the two nodes, speed is reduced and latency increases. VNFCs that participate in the data plane path are advised to contain the virtual CPU, memory, and physical NIC associated with them to a single NUMA node for optimal performance.

Data plane intensive VNFs tend to serve a central role in a CSP network: as a Packet Gateway in a mobile core deployment, a Provider Edge router (PE) in an MPLS network, or a media gateway in an IMS network. As a result, these VNFs are positioned in a centralized location in the CSP network: in the data center. With their crucial role, VNFs are typically static and are used by the central organization to offer services to a large customer base. For example, a virtualized Packet Gateway in a mobile core network will serve a large geographical region as the central termination point for subscriber connections. Once the VNF is deployed, it is likely to remain active for a long duration, barring any NFVI life cycle activities such as upgrades or other maintenance.

This aggregation role translates to a certain sizing requirement. The VNFs must serve many customers, which is the reason for their data plane intensive nature. Such VNFs have many components that allow them to be scaled and managed. These components include at minimum an OAM function, packet processing functions, VNF-specific load balancing, and often log collection and monitoring functions. Individual components can also require significant resources to provide large scale services.

The central position of these VNFs, their sizeable scale, and their static nature, all suggest that dedicated resources are required to achieve their expected performance goals. These dedicated resources begin with

hosts using powerful servers with high performing network interface cards. The servers are grouped together into a cluster dedicated to data plane intensive workloads. Using the same constructs introduced earlier in this document, the data plane intensive cluster is consumed by VMware Integrated OpenStack and is made into a Tenant vDC. VNFs are onboarded into the VMware Integrated OpenStack image library for deployment.

VMware Integrated OpenStack supports NUMA aware placement on the underlying vSphere platform. This feature provides low latency and high throughput to VNFs that run in telecommunications environments. To achieve low latency and high throughput, it is important that vCPUs, memory, and physical NICs used for VM traffic are aligned on same NUMA node. The specific teaming policy that must be created depends on the type of deployment you have.

With the architecture provided in this section of the document, data plane intensive workloads are ensured the resources they require to benefit from platform modularity, while meeting carrier grade performance requirements. Specific configuration and VNF design guidelines are detailed in the Tuning VMware vCloud NFV for Data Plane Intensive Workload whitepaper.

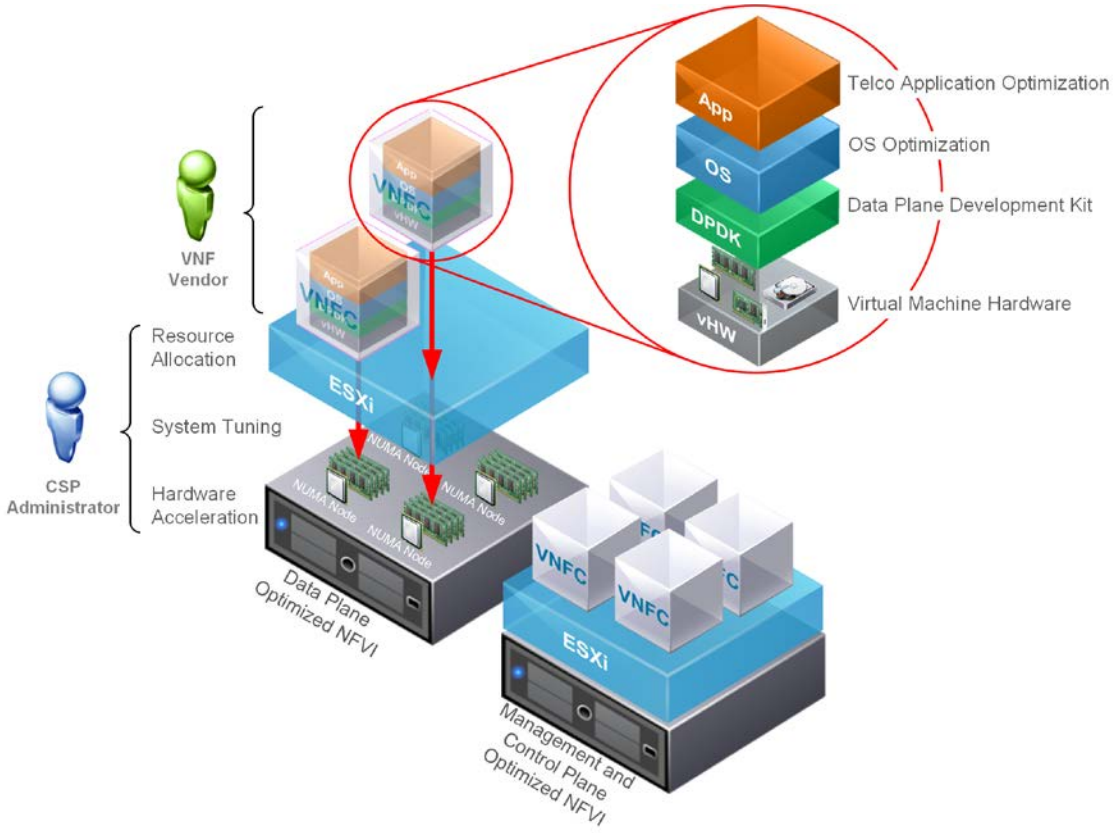

Figure 25: VMware vCloud NFV OpenStack Edition Design for Data Performance

4.9 VNF Onboarding

The VNF onboarding process is typically collaborative between the VNF vendor and the CSP. Before a VNF is onboarded, the VNF supplier must provide the CSP with all prerequisite information for the successful onboarding of the VNF. This includes configuration information such as the VNF format, number of networks required, East-West and North-South network connectivity, routing policy, security policy, IP ranges, and performance requirements. Figure 26 provides a high-level view of the roles, areas of responsibility, and tenant flow required to bring an image onboard and deploy it. This is discussed in more detail in the following section of this document.

The initial format of the VNF is taken into consideration for onboarding, as is the format of any additional components the VNF requires to function. Images are either able to be directly imported, or they can be converted. VMware Integrated OpenStack natively supports ISO, VMDK, and OVA formats, non-native formats such as RAW, QCOW2, VDI, and VHD are also supported after automatic conversion by the import process. These formats can also be imported using the command line.

After the initial VNF requirements, images, and formats are understood, a project must be created for the VNF to be deployed in an operational environment. Projects are the VMware Integrated OpenStack constructs that map to tenants. Administrators create projects and assign users to each project. Permissions are managed through user, group, and project definitions. Users have a further restricted set of rights and privileges. Users are limited to the projects to which they are assigned, although they can be assigned to more than one project. When a user logs in to a project, they are authenticated by Keystone. Once the user is authenticated, they can perform operations within the project.

One of the tasks the administrator must fulfill when building a project for the VNF is the setting of initial quota limits for the project. To guarantee resources, a Tenant vDC can be used. A Tenant vDC provides resource isolation and guaranteed resource availability for each tenant. Quotas are the operational limits that configure the amount of system resources available per project. Quotas can be enforced at project and user levels. When a user logs in to a project, they see an overview of the project including the resources they have been provided, the resources they have consumed, and the resources they have remaining. The quota of resources available to a project can be further divided using Tenant vDCs for fine grained resource allocation and control.

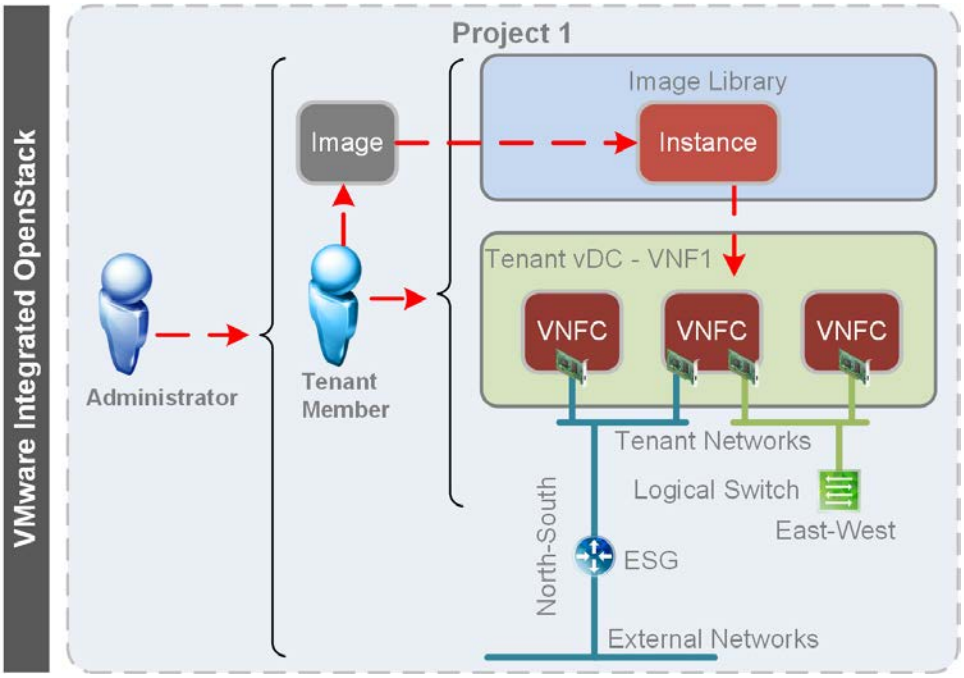

Figure 26: VMware Integrated OpenStack VNF Onboarding

Based on specific VNF deployment requirements, a tenant can provision East-West connectivity, security groups, firewalls, micro-segmentation, NAT, and LBaaS from within the VMware Integrated OpenStack user interface or commandline. VNF North-South connectivity is achieved by connecting tenant networks to external networks. External networks are created by administrators and a variety of VNF routing scenarios are possible.

Before full deployment of VNFs, tenants must consider the performance requirements for the VNFs. They must also evaluate whether VNFs must be deployed to optimized compute nodes. Optimized compute nodes are nodes that can be configured with special performance capabilities. VNFs must be placed correctly in order to take advantage of the performance capabilities in optimized compute nodes.

After the VNFs are deployed their routing, switching, and security policies must be considered. There are many different infrastructure services available that can be configured in different ways, and in the coming sections of this document a couple of options are discussed.

Figure 27 provides an example of VNFs with VLAN external connectivity. Project 1 has a logical segment built using an NSX logical switch with connectivity to an NSX ESG. The northbound interface of the NSX ESG is connected to a VLAN backed external network trunked to the CSP's router. The NSX ESG routes traffic to the SP router.

Project 2 has a similar configuration with a slight modification. The VNF is directly connected to the VLAN based external network. The VLAN is trunked through the DVS to the physical router. The first hop router for Project 2 is the CSP router. In both cases, the CSP router terminates the VLANs and the route to the telecommunications network. To maintain separation, the VLANs can connect to MPLS Layer 2 or Layer 3 VPNs.

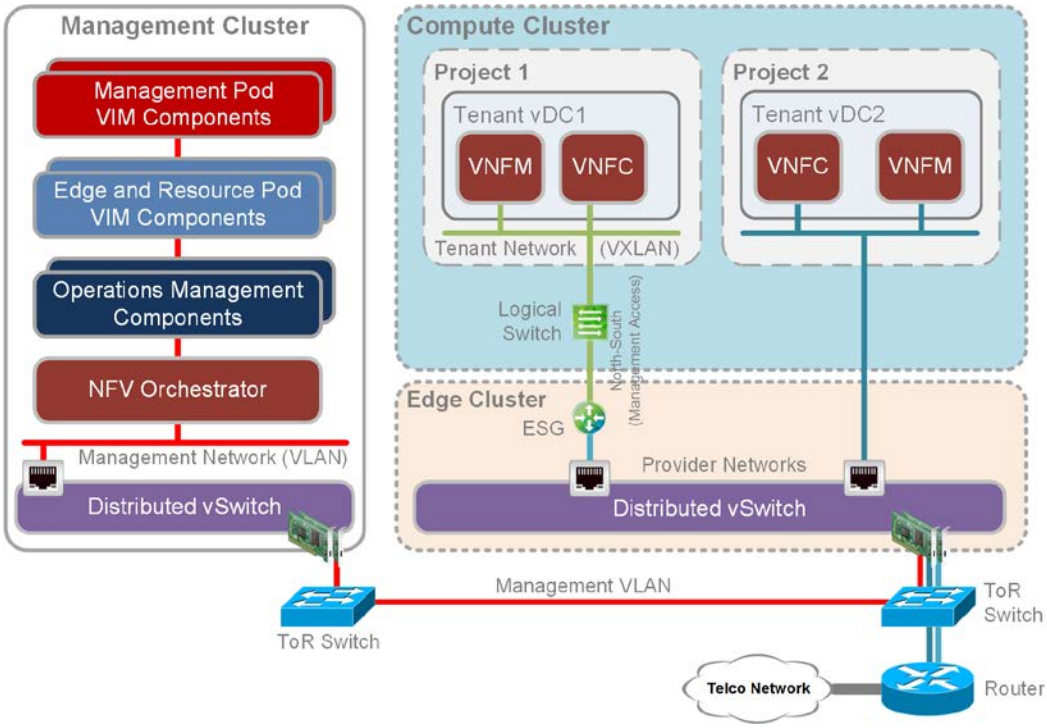

Figure 27: VNF Networking with VLAN Backed External Network

Figure 28 provides another example of VNF networking. VNFs require management access to establish the necessary network connectivity, for example between a VNFM deployed in the Resource pod and an NFVO deployed in the Management pod. Components in the Management pod are connected to the management network VLAN. This VLAN is trunked to the hosts in the Edge pod where the physical NICs are assigned as uplinks to a vSphere Distributed Switch. The CSP provisions a vSphere Distributed Switch port group used by administrators to configure the external network. In this scenario, the NSX Edge instance performs the role of a VXLAN to VLAN bridge and provides edge network services, including NAT and a stateful firewall for security.

Implementation of East-West connectivity between VNFCs in the same Tenant vDC, and connectivity between VNFs in two different Tenant vDCs belonging to the same project, is identical. This is true because tenant networks are accessible by all Tenant vDCs within the project. Tenant networks are implemented as logical switches within the project. The North-South network is a tenant network connected to the telecommunications network through an NSX for vSphere ESG. The ESG can be configured with edge services including NAT, load balancing, and firewall services, in addition to its role as a router. Additionally, different interfaces can be used for different services, such as direct passthrough SR-IOV for data plane traffic and VMXNET3 for management and control plane traffic.

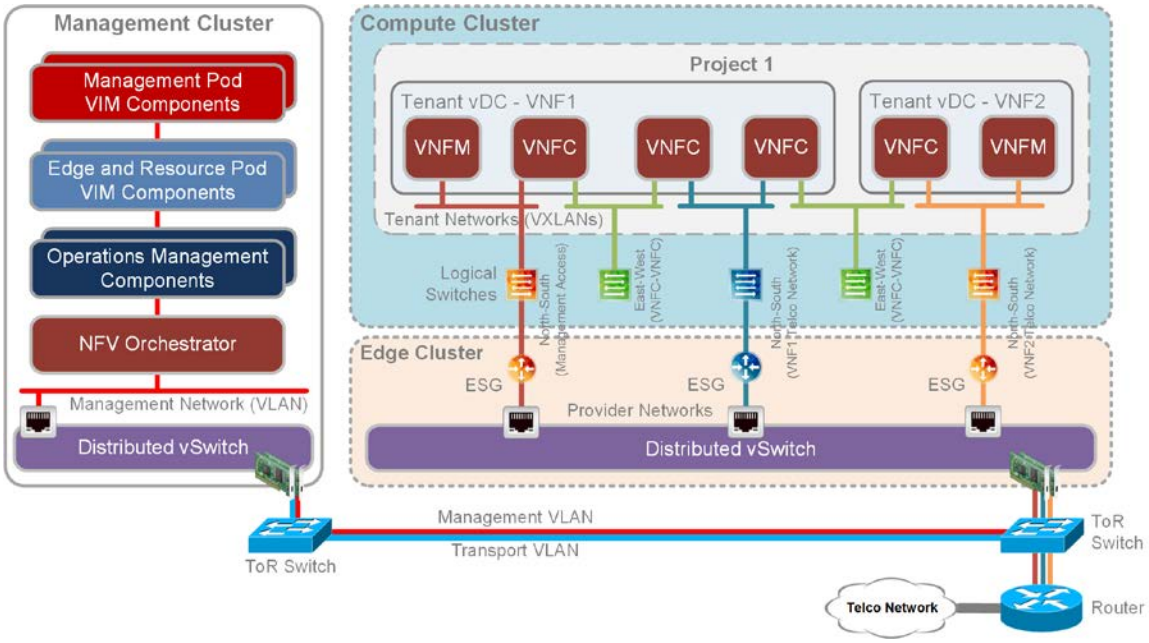

Figure 28: VNF Networking in Three-Pod Design

VMware Integrated OpenStack exposes a rich set of API calls to allow automation. You can automate the deployment of VNFs using a Heat template. With API calls, the upstream VNFM and NFVO can automate all aspects of the VNF life cycle.

Authors and Contributors

The following authors co-wrote this paper:

- Jason Sauviac, Lead Solutions Architect, NFV Solutions Engineering, VMware
- Indranil Bal, Solution Consultant, NFV Solutions Engineering, VMware
- Pradip Kadam, Solution Consultant, NFV Solutions Engineering, VMware

Many thanks for contributions from:

- Danny Lin, Senior Director, NFV Solutions Engineering, VMware
- Michelle Han, Director, Solutions Testing and Validation, NFV Solutions Engineering, VMware
- Jambi Ganbar, Senior Technical Solutions Manager, NFV Solutions Engineering, VMware

VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 www.vmware.com