vCloud NFV Reference Architecture 3.2.1

VMware vCloud NFV 3.2.1

You can find the most up-to-date technical documentation on the VMware website at:

https://docs.vmware.com/

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

Contents

About vCloud NFV Reference Architecture

1

This reference architecture provides guidance for designing and creating a Network Functions Virtualization (NFV) platform by using VMware vCloud® NFV™.

This document describes the high-level design principles and considerations when implementing an environment that is based on vCloud NFV. It also provides example scenarios to help you understand the platform capabilities.

Intended Audience

This document is intended for telecommunications and solution architects, sales engineers, field consultants, advanced services specialists, and customers who are responsible for the virtualized network functions (VNFs) and the NFV environment on which they run.

Introduction to vCloud NFV

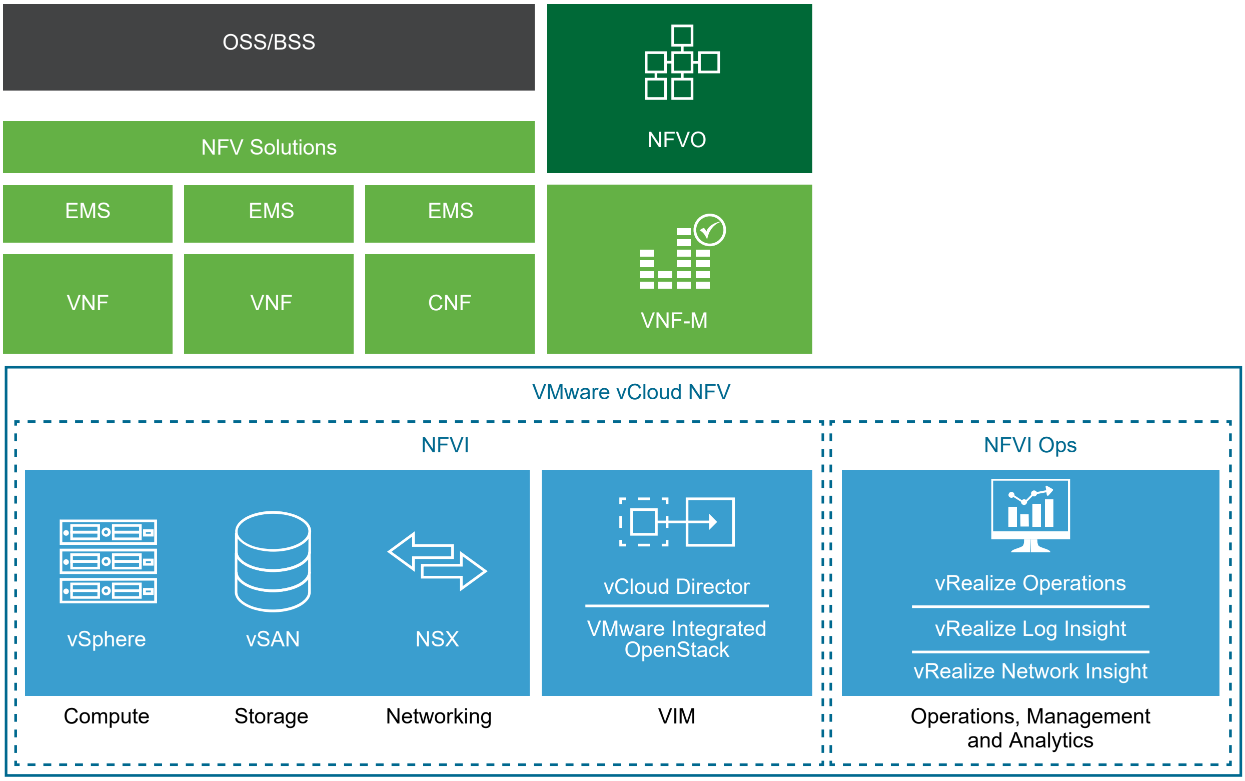
2

VMware vCloud NFV combines a carrier-grade NFV infrastructure with VMware vCloud Director® for Service Providers as the NFV Virtualized Infrastructure Manager (VIM). This version of the vCloud NFV platform combines the vCloud Director for Service Providers API with stable and supportable vCloud NFV Infrastructure (NFVI). This way, vCloud NFV provides a platform to support Communication Service Providers (CSPs) in realizing the goal for network modernization and business transformation.

The vCloud NFV platform implements a modular design with abstractions that enable multi-vendor, multi-domain, and hybrid physical, and virtual execution environments. The IaaS layer that is exposed through this version of vCloud Director for Service Providers, provides a CI/CD environment for workload life cycle management. The platform also delivers an automation framework to interoperate with external functions for service orchestration and management.

In addition to the core NFV infrastructure components for compute, storage, networking, and VIM, the vCloud NFV platform includes a fully integrated suite for operational intelligence and monitoring. This suite can be used to further enhance the runtime environments with workflows for dynamic workload optimization and proactive issue avoidance.

Figure 2-1. vCloud NFV Components



The vCloud NFV components, their interactions with each other, and how they meet CSP requirements, are described in this reference architecture.

Acronyms and Definitions

3

vCloud NFV uses a specific set of acronyms that apply to the NFV technology and the telco industry.

Table 3-1. General Acronyms

Abbreviation	Description
BFD	Bidirectional Forwarding Detection, for failure detection on the transport links.
DPDK	Data Plane Development Kit, an Intel-led packet processing acceleration technology.
EMS	Element Management System
MTTR	Mean Time to Repair
MTTU	Mean Time to Understand
NMS	Network Management System

Table 3-2. NFV Acronyms

Abbreviation	Description
CNF	Cloud-Native Network Function, executing within a Kubernetes environment.
CCP	Centralized Control Plane in the VMware NSX-T™ Data Center architecture.
EPA	Enhanced Platform Awareness
LCP	Local Control Plane in the NSX-T Data Center architecture.
MANO	Management and Orchestration components, a term originating from the ETSI NFV architecture framework.
NFVI	Network Functions Virtualization Infrastructure
NFV-OI	NFV Operational Intelligence
N-VDS (E)	Enhanced mode when using the NSX-T Data Center N-VDS switch. This mode enables DPDK for workload acceleration.
N-VDS (S)	Standard mode when using the NSX-T Data Center N-VDS switch.
VIM	Virtualized Infrastructure Manager

Table 3-2. NFV Acronyms (continued)

Abbreviation	Description
VNF	Virtual Network Function, executing in a virtual machine.
VNFM	Virtual Network Function Manager

Table 3-3. Telco Acronyms

Abbreviation	Description
HSS	Home Subscriber Server in the mobile evolved packet core 4G architecture.
MVNO	Mobile Virtual Network Operator
PCRF	Policy, Charging and Rating Function, in the mobile evolved packet core 4G architecture.
PGW	Packet Gateway in the mobile evolved packet core 4G architecture.
SGW	Service Gateway in the mobile evolved packet core 4G architecture.
SBC	Session Border Controller used in voice telephone for control and data plane communications between clients.
UPF	User Plane Function

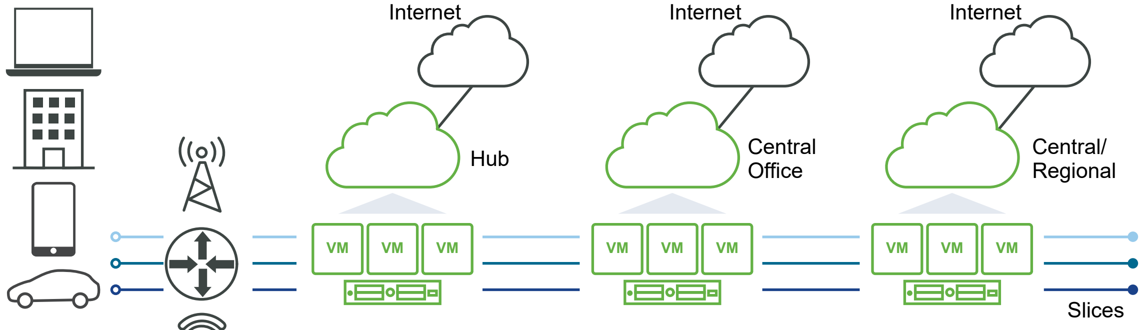
Reference Environment

4

5G services require a mixture of low-latency, high throughput, and high user densities and concurrences. The distribution of functional components requires a more sophisticated service delivery model.

The network is transforming into a mixture of highly distributed functions together with centralized functions. This way, the network is moving away from the typical centralized models in service delivery. There is also an emerging paradigm shift with employing third-party IaaS, PaaS, and SaaS offerings from public cloud providers.

Figure 4-1. Reference Environment



Customer Edge (Millions)	Far Edge (1000's)	NearEdge (100's)	Core (10's)
40-80 µs	<1-5 ms	<5-10 ms	<20-50 ms
1-3 Servers	5-10 Servers	1-5 Racks	Multiple Racks
SD-WAN uCPE	vRAN AR/VR Gaming	vEPC UP MEC Video Surveillance CDN IoT Apps	vEPC CP vIMS 5G CP Subscriber (HSS) Policy (PCRF)

The highly distributed topology of the network supports the next-generation service characteristics in distribution and composition. It also requires a new way of managing the network and infrastructure resources. The number of services that are spanning the industry verticals is exploding exponentially. Today's endpoint ranging in fixed and mobile offers will grow into the billions with IoT connections. The highly distributed edge sites are projected to be in the thousands, regional sites in the 100s, core sites in the 10s, and a large variety of public cloud provider sites.

NFV and Software-Defined Networking (SDN) transformations introduce complex interconnections from endpoints such as branches, small offices, connected cars, and IoT gateways to private data centers and public cloud providers. This reference environment and its transformation are not only a technical challenge but also impacts the business and operating processes.

Reference Environment Requirements

The reference environment places strict requirements for service placement and management to achieve optimal performance.

- **Federation options**: The reference environment topology offers a diverse set of federation options for endpoints, private and public clouds, each with distinct ownership and management domains. Virtualized endpoints provide better control and manageability, however they are not suitable for all types of use cases. Likewise, service functions can be distributed and managed across private and public clouds.
- **Disaggregated functions**: Services are highly disaggregated so that control, data, and management planes can be deployed across the distributed topology. Edge clouds offer the performance advantages of low latency and data plane intensive workloads, while control and management plane components can be centralized with a regional and global scope.
- **Functional isolation**: Multitenancy provides network and service isolation across different tenancy models in the reference environment. However, resource management considerations must be made for shared network functions such as DNS, policy, authentication, and so on.
- **Service placement**: The highly distributed topology allows for flexibility in the workload placement. Making decisions based on proximity, locality, latency, analytical intelligence, and other Enhanced Platform Awareness (EPA) criteria are critical to enable an intent-based placement model.
- **Workload life cycle management**: Each cloud is elastic with workload mobility and how applications are deployed, executed, and scaled. An integrated operations management solution can enable efficient life cycle management to ensure service delivery and Quality of Service (QoS).
- **Carrier-grade characteristics**: Because 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 important.

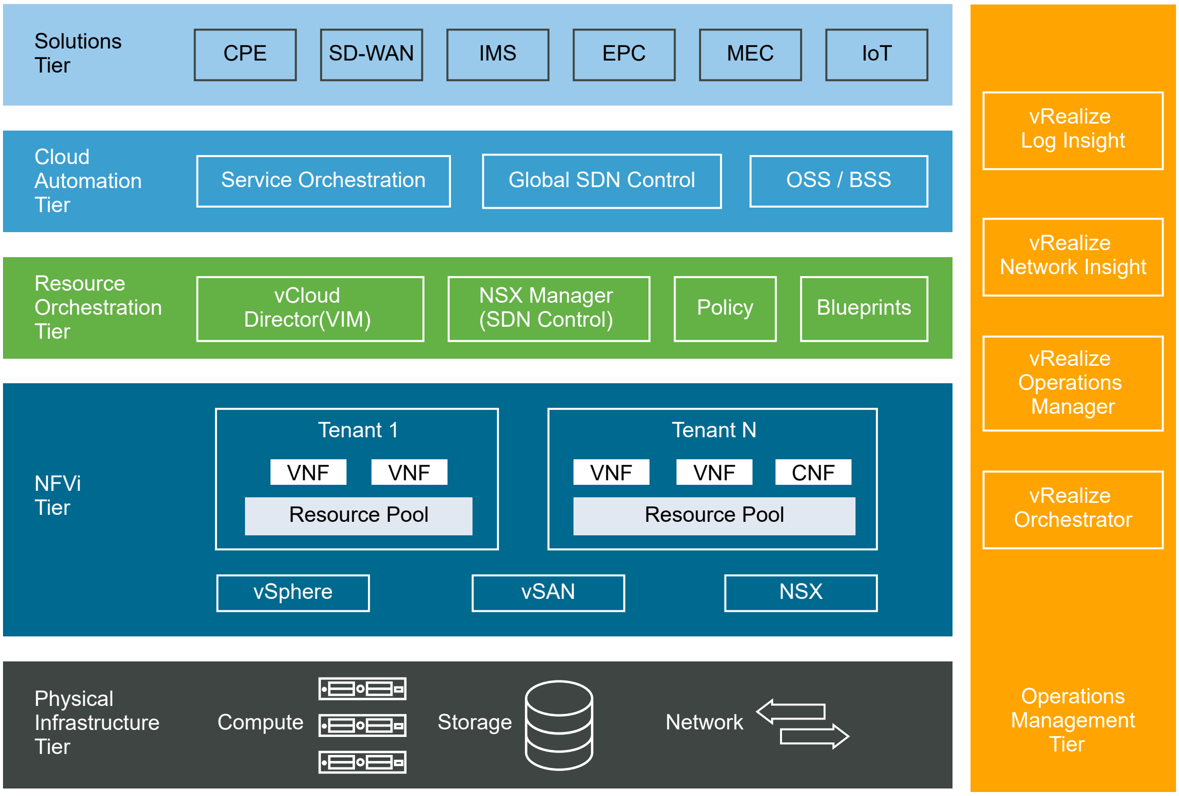
- **NFVI life cycle (patching and upgrades)**: The platform must be patched and upgraded by using optimized change management approaches for zero to minimal downtime.

NFV Reference Model

NFV is an architectural framework that was first developed by the ETSI NFV Industry Specification Group. The framework provides a reference model where network functions are delivered through software virtualization with Commercial-Off-The-Shelf (COTS) hardware. This way, NFV moves away from the proprietary, purpose-built hardware that is dedicated to a single service. The result is a network that is agile, resilient, and equipped to deliver high-quality services. The NFV framework defines functional abstractions and interactions between the building blocks. Some of these abstractions are already present in current deployments, while others must be added to support the virtualization process and operation.

The following diagram shows the reference model for an NFV environment with clear functional abstractions and interactions in a tiered approach:

Figure 4-2. Layered Abstractions of the NFV Environment



Physical Tier

Represents compute, storage, and physical networking as the underlying pool of shared resources. Also, there are numerous other physical network devices such as switches, routers, Element Management System (EMS), and so on, making the execution ecosystem a hybrid virtual and physical topology.

NFVI Tier

The lowest tier of the vCloud NFV platform. It delivers the virtualization run-time environment with network functions and resource isolation for VM workloads. In NFVI, virtualized compute, storage, and networking are delivered as an integrated solution through VMware vSphere®, VMware vSAN, VMware NSX® Data Center for vSphere®, and NSX-T Data Center. Isolated resources and networks can be assigned to a tenant slice, which is a runtime isolated partition delivering services. Tenant slices can be dedicated to a tenant or shared across tenants. The NFVI is optimized and adjusted for telco-class workloads to enable the delivery of quality and resilient services. Infrastructure high availability, performance, and scale considerations are built into this tier for performance optimization.

Resource Orchestration Tier

It provides resource management capabilities to the NFVI tier. This way, the NFVI can deliver a flexible infrastructure-as-code for life cycle management of workloads, network management, and resource management. The resource orchestration tier 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 module manages the allocation and release of virtual resources, and the association of virtual to physical resources, including resource optimization. VIM also maintains the inventory of NFVI, including the linkage and relationship between components as they relate to an instance of a VNF or CNF workload. This way, VIM allows for monitoring in the context of a single VNF.

Cloud Automation Tier

The service management and control functions that bridge the virtual resource orchestration and physical functions to deliver services and service chains. It is typically a centralized control and management function, including the embedded automation and optimization capabilities.

Solutions Tier

The multi-domain ecosystem of software virtual functions as native VM functions. Such functions are composed in complex solutions to enable service offers and business models that CSP customers consume. Solutions can range from small branch office functions to a fully evolved packet core that is delivered as tenant slices across multiple clouds.

Operations Management Tier

An integrated operational intelligence for infrastructure day 0, 1, and 2 operations that spans across all other tiers. The functional components within the operations management tier provide the topology discovery, health monitoring, alerting, issue isolation, and closed-loop automation.

This chapter includes the following topics:

- Key Customer Objectives

Key Customer Objectives

The goal of network modernization is to drive greater classes of service innovation and timely enablement. CSPs are considering the following key objectives as they transform their networks and design for new business and operational models.

Fixed Mobile Convergence

As networks evolved through 2G and 3G generations, the voice and data network architectures particularly the circuit-switched and packet-switched networks were separated. As networks evolved, the CSPs went towards an all IP network, therefore the convergence of fixed and mobile networking. The environments for voice mostly share the core networking components with different access networks. The scale, performance, and management of such converged networks are more critical now than earlier.

Data Intensive Workload Acceleration

The demand for throughput has increased exponentially with smart devices and immersive media services. The networking and compute expenditures continue to grow to meet such demands in traffic throughput. Acceleration technologies such as DPDK, SR-IOV, and hardware offload are at the forefront to reduce OpEx for data intensive applications.

Distributed Clouds

To meet the increased bandwidth and low-latency requirements, network designs are expanding the centralized compute models to distributed edge computing models. Certain level of edge distribution exists in the regional and core data centers; however further edge distribution is necessary to control traffic backhauling and to improve latencies. In conjunction, VNFs are disaggregating to distribute data plane functions at the edges of the network whereas control functions are centralized. Service distribution and elasticity are vital part of the network design consideration.

Network Slicing

Network slicing is a way for cloud infrastructure to isolate resources and networking to control the performance and security for workloads that are executing on the shared pool of physical infrastructure. With distributed topologies, the concept of network slicing furthermore stretches across multiple cloud infrastructures, including access, edge, and core virtual and physical infrastructures. Multi-tenancy leverages such resource isolation to deploy and optimize VNFs to meet customer SLAs.

Dynamic Operational Intelligence

The cloud infrastructures must be adaptive to meet the needs of workloads. Rightsizing the environment and dynamic workload optimizations, including initial placement, are part of the continuous automation orchestration. The cloud infrastructure environments require integrated operational intelligence to continuously monitor, report, and act in a timely manner with prescriptive and predictive analytics.

Policy-Based Consistency and Management

Model-driven approaches play a key role in the modern cloud infrastructures. Resource modeling, runtime operational policies, security profiles, declarative policies, movement of policies with workloads, onboarding, and so on, ensure consistency and ease of management.

Carrier-Grade Platform

The cloud infrastructure environment must meet the strict requirements for availability, fault tolerance, scale, and performance. Security is necessary across the transport, data, and workload dimensions. The mobility of workloads across distributed clouds introduces a new challenge for its authenticity and integrity.

Architectural Framework and Components

5

This section explores the overall framework for the vCloud NFV platform architecture, including the key stakeholders, conceptual architecture environment, logical architecture, and components of the vCloud NFV platform. The reference architecture design principles set the frame for the core and analytics-enabled designs of the vCloud NFV platform.

This chapter includes the following topics:

- Key Stakeholders
- Conceptual Architecture
- Logical Architecture and Components
- vCloud NFV Components
- Design Principles

Key Stakeholders

The reference architecture considers key stakeholders that are involved in the end-to-end service management, life cycle management, and operations.

Cloud provider

The CSP operations personnel who are responsible for provisioning and on-boarding all day 0 and day 1 functions to enable services for target customers and tenants.

Consumer

The end user who consumes the services that the tenants provide. For example, IoT devices, mobile handsets, API consumers, and MVNO.

Customer

The enterprise or entity who owns the business relationship with the CSP. The customer might be an internal line of business such as fixed line and mobile services and can also be an external enterprise.

Tenant

The various classes of services (offers) that a customer provides to their consumers. Each tenant is represented as a resource slice; hence a customer can have one or more tenants. A mobile line of business can offer slices to an MVNO customer, for example a tenant for voice services and a tenant for data services.

Operations Support

The operations management process and the team that ensure the services are operating to meet the promised stringent SLAs.

Network Planning

The operations management planning function that is responsible for the resource and VNF capacity and forecasting, and new data center designs.

Security Operations

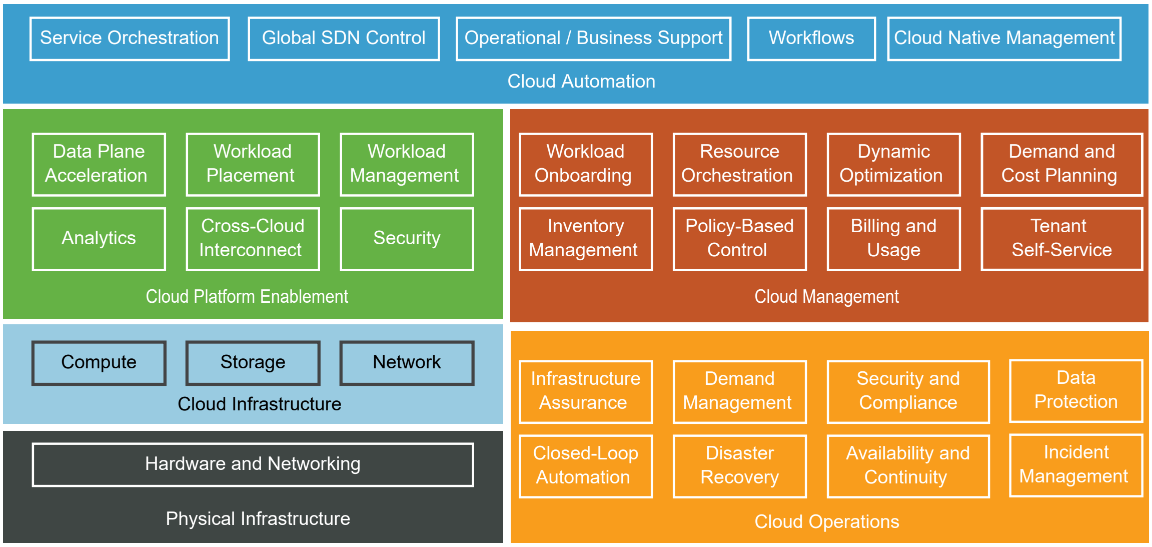
Security operations function that is responsible for all aspects of security, network, application, and data.

Conceptual Architecture

Cloud infrastructure-based computing is the next-generation standard in modernizing the CSP networks as they evolve to 5G architectures, services, and agile delivery. The shared infrastructure with complete softwarization of network functions and applications provide greater advantages in cost, performance, and agility.

The modernization of the CSP infrastructure requires a complex ecosystem of solutions and functions delivering to a pre-set business and operating model. The cloud infrastructure modernization changes not only the business model in service agility and metered revenue models, but also challenges the silo operating model. The following figure shows the conceptual view of various domains, capabilities, and their interactions that need consideration in the modernization of networks and business and operational models.

Figure 5-1. Conceptual Architecture



Cloud Automation

Centralizes the overall service management functions such as service definitions, composition, onboarding, life cycle management, and support. The Cloud Automation domain hosts functions such as an NFV-O that is responsible for service blueprinting, chaining, and orchestration across multiple cloud infrastructure environments. Besides NFV-O, the global SDN control functions are responsible for stitching and managing physical and overlay networks for cross-site services. Real-time performance monitoring can be integrated into the SDN functions to dynamically optimize network configurations, routes, capacity, and so on.

The successful cloud automation strategy implies full programmability across other functions.

Cloud Platform Enablement

Extends a set of platform capabilities from the cloud infrastructure to cloud automation, so VNFs, VNF managers, and other core components can leverage. Examples of enablement capabilities include:

- Analytics to ingest VNF metrics that can be correlated with infrastructure metrics for smarter context and insights.
- Workload placement to determine the right location for a workload depending on available resources, class of resources, and feature capabilities such as data intensive acceleration.
- Workload acceleration using DPDK and SR-IOV for data intensive VNFs.
- Security for network, data, and workloads.

Cloud Management

Cloud Management plays a critical role across many different dimensions. More fundamentally, it provides a templated and prescriptive workload management capabilities that the automation layer can use to program and orchestrate service on-demand and with agility. Service onboarding models can be turned into fully zero-touch provisioning and exposed to tenants through a self-service portal. Business models such as metered billing can be enabled as a catalog of services and tariffs.

Once services and workloads are onboarded, the cloud management functions must also ensure dynamic optimization such as workload rebalancing and capacity growth or shrink to maintain agreed SLAs. Such optimizations need to integrate with the cloud operations domain for real-time usage and performance intelligence. Policies, including platform awareness, NUMA affinity, host affinity, restart-sequences, are necessary for efficient optimization.

Cloud Operations

Ensures that the operational policies and SLAs are met by continuous data collection, correlation, and analytics. Infrastructure assurance is a key component of this Cloud Operations domain. Intelligence can be tied into a closed-loop workflow that can be integrated in the automation domain for proactive issue avoidance, for example, triggering a trouble ticket incident management system.

Also, other functions for day 2 operations such as demand and capacity planning, security and compliance, high availability, and disaster recovery are necessary to ensure availability and integrity across the cloud infrastructure environments.

Cloud Infrastructure

The core virtualization domain providing resource abstraction for compute, storage, and networking and their orchestration through a VIM to allocate, control, and isolate with full multi-tenancy and platform-awareness.

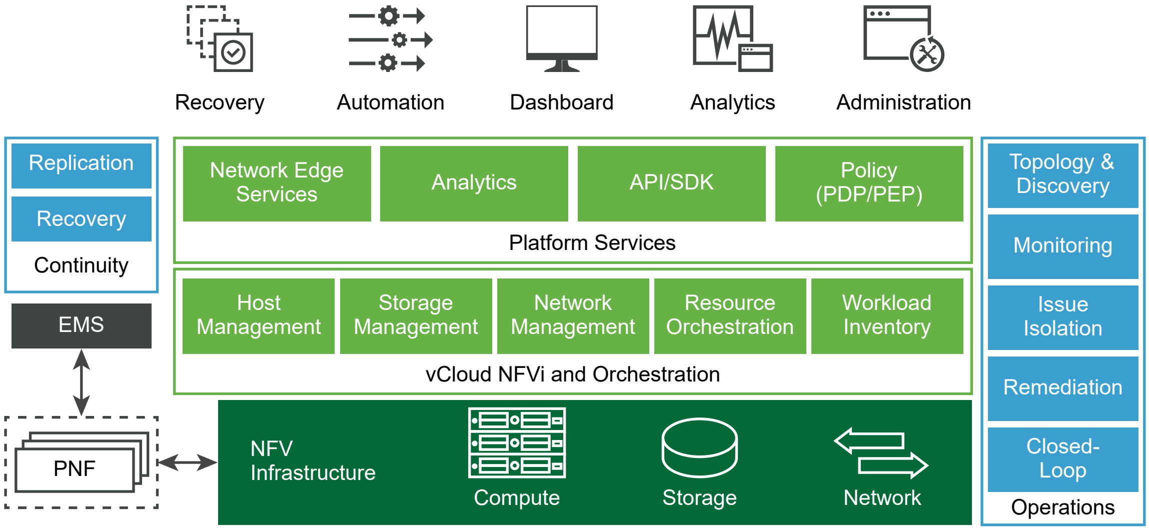
Logical Architecture and Components

The vCloud NFV platform implements the conceptual architecture that is outlined and defined at a high level through the logical building blocks and core components.

The vCloud NFV platform is an evolution of the VMware NFV solution, based on the extensive customer deployment and the continued development of standards organizations such as the European Telecommunications Standards Institute (ETSI). The vCloud NFV 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 NFV solution is integrated with holistic operations management and service assurance capabilities, empowering the CSP 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 following diagram maps the conceptual architecture to a logical view for the vCloud NFV platform:

Figure 5-2. Logical Architecture



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

Logical Architecture Components

The vCloud NFV platform consists of three core domains of functions, core NFV infrastructure, infrastructure orchestration, and operations management. At the core infrastructure, ESXi is used to virtualize the compute resources, NSX-T Data Center to provide virtual networking, and vSAN for storage. The core NFV infrastructure virtualization layer provides the following functions:

- **Physical Resource Abstraction**: By using the software component layers between the physical hardware and the VNFs, physical resources are abstracted. This provides a standardized software-based platform for running workloads, regardless of the underlying hardware. As long as the CSP uses certified physical components, workloads 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 presents a logical virtualization layer to workloads, 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 workload requirements.
- **Physical Resource Sharing**: To benefit from cloud economies, the resources that are pooled and abstracted by the virtualization layer must be shared between various network functions. The virtualization layer provides the functionality that is required for VNFs to be scheduled on the same compute resources, collocated on the shared storage, and to have the network capacity divided among them. The virtualization layer also ensures fairness in resource utilization and usage policy enforcement.

vCloud NFV Infrastructure and Orchestration

The infrastructure and orchestration domain contain the NFVI abstractions for compute, storage, and networking. It also contains the Virtualized Infrastructure Manager (VIM), which is the resource orchestration component of the NFV platform.

Compute - VMware ESXi

ESXi is the hypervisor software that abstracts the physical x86 server resources from the VNFs. Each compute server is called a host in the virtual environment. ESXi hosts are the fundamental compute building blocks of vCloud NFV. ESXi host resources can be grouped to provide an aggregate set of resources in the virtual environment that is called a cluster. Clusters logically separate the management and VNF components. ESXi hosts are managed by the VMware vCenter® Server Appliance™ that is part of the VIM components. For new features, see the vSphere documentation .

Host Management - VMware vCenter Server

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 about the performance, capacity, and state of its inventory objects. It exposes APIs to other management components for fine-grained control, operation, and monitoring of the underlying virtual infrastructure.

Networking - VMware NSX-T Data Center

NSX-T Data Center is the software-defined networking component of the vCloud NFV reference architecture. It allows CSPs to programmatically create, delete, and manage software-based virtual networks. NSX-T Data Center also leverages technologies for high-performance workloads such as DPDK. These networks serve the communication between VNF Components and provide customers with dynamic control over their service environments. Dynamic control is enabled through tight integration between the resource orchestration layer and NSX-T Data Center.

Network multitenancy is implemented with NSX-T Data Center by assigning tenants their virtual networking components and providing different network segments. A two-tiered architecture is used in the NSX-T Data Center design to implement a provider and tenant separation of control across the logical switching and routing fabric. Logical switching is supported in two modes, N-VDS Standard and N-VDS Enhanced, both of which support the overlay and VLAN-backed networks. The fully distributed routing architecture enables routing functionality closest to the source. This structure gives both provider and tenant administrators complete control over their services and policies.

NSX-T Data Center is derived by three separate but integrated planes: management, control, and data. Each of these planes is a set of process, modules, and agents residing on VMware NSX® Manager™ and transport nodes.

Storage - VMware vSAN

vSAN is the native vSphere storage component in the NFVI virtualization layer, providing a shared storage pool between hosts in the vSphere cluster. With vSAN, storage is shared by aggregating the local disks and flash drives that are attached to the host. Other third-party storage solutions with storage replication adapters that meet the VMware storage compatibility guidelines are also supported.

Resource Orchestration - VMware vCloud Director for Service Providers

VMware vCloud Director for Service Providers is the VIM component that vCloud NFV exposes as the interface for the VNF life cycle management. It uses vCenter Server and NSX Manager to orchestrate compute, storage, network, and imaging infrastructure services from a single, programmable interface.

Platform Services

The platform services domain represents the capabilities that vCloud Director for Service Providers delivers to the NFV platform. VNFs, VNF managers, and other components running within the vCloud NFV platform can use these capabilities.

Edge Services - VMware NSX-T Data Center

NSX-T Data Center provides two classes of routing capabilities: Distributed Router (DR) and Service Router (SR). The Service Router capability provides services such as NAT, firewall, load balancer, and so on. It also provides Tier-0 and Tier-1 gateways that VNFs can employ with stateful and stateless options. The Edge services cannot be distributed and require a centralized pool of capacity with high availability and scalability. The appliances that host the centralized services or SR instances are called Edge Nodes. These nodes also provide connectivity to the physical infrastructure.

Analytics – VMware vRealize Operations Manager

The vCloud NFV platform is fully integrated with an operations management solution for day 1 and day 2 operations for health monitoring, issue avoidance, and closed-loop automation. This way, the platform provides infrastructure assurance out of the box. The analytics framework can be used by the network function and application developers to ingest and correlate their services-specific data in VMware vRealize® Operations Manager™ and leverage its capabilities seamlessly. The framework provides various mechanisms through the management and content packs for data management and closed-loop automation with workflows that are custom to their operations and planning needs.

Policy Consistency

The vCloud NFV platform components use policy-based frameworks that can be defined once and applied at runtime to maintain the desired end state. The decision and enforcement split makes the management and operation of the platform and its services highly flexible and leveraged by operations and applications. Policies in compute can be used for rightsizing the infrastructure to ensure capacity and performance. Workload placement and runtime optimization for DRS and vMotion can be prescribed with platform awareness. Policies in networking range in physical networking such as teaming, network management and control, and security for East-West and perimeter traffic control. Storage policies provide services such as availability levels, capacity consumption, and stripe widths for performance. Policies in operations ensure that SLAs are met with configurable alerting, recommendation, and remediation framework.

Programmability

The vCloud NFV solution is a fully open platform supporting flexible APIs and SDKs. For more information about the VMware APIs, see the VMware API Explorer.

Continuity

This domain represents components for business continuity and disaster recovery solutions, which are an integral part of the vCloud NFV platform.

VMware Site Recovery Manager

VMware Site Recovery Manager™ works with various storage replication solutions, including VMware vSphere® Replication™, to automate the migration, recovery, testing, and failing back virtual machine workloads for disaster recovery across multiple sites.

VMware vSphere Replication

vSphere Replication is a virtual machine data protection and 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.

Operations Management

The Operations Management domain represents the day 1 and day 2 functions to ensure that the infrastructure and service components are operating in a healthy state and SLAs are met.

The operations management solution includes four components that together provide a holistic approach to the operations management for the NFVI of a CSP. Together VMware vRealize® Operations™, VMware vRealize® Log Insight™, and VMware vRealize® Network Insight™ monitor the health of the virtual environment, collect logs and alarms, correlate events across multiple data sources and components to predict future issues. These components and VMware vRealize® Orchestrator™ use policy-based automation framework to conduct remediation and analyze data to help the operator with health prediction and issue avoidance.

The key tasks of the operations management components are:

- **Topology and discovery**: NFVI visibility is achieved by collecting essential performance and fault metrics from the virtualization layer, the physical devices, and the VIM components. The elastic and dynamic nature of the NFVI layer requires tracking of objects and maintaining correlations to help with decision tree accuracy and intelligence for infrastructure assurance.
- **Monitoring**: The components of the Operations domain continuously collect and analyze health, SLA, and planning metrics to ensure that services are meeting stringent QoS and to help avoiding issues in a predictable and prescriptive cadence.
- **Issue isolation**: The components of the NFV environment in the physical infrastructure, the virtualization layer, or 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.

- **Remediation**: 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. Issues pertaining to performance, capacity, congestion, and so on, can be proactively avoided, increasing the Mean Time To Failure (MTTF).
- **Closed-loop optimization**: The operations management components analyze the system usage and proactively provide optimization recommendations, including network topology modifications. Actions can be tied to orchestrated workflows and automation to trigger Virtual Network Function Manager (VNFM), service orchestrators, and others to ensure continuous optimization, balancing, and recover in the event of failures.

VMware vRealize Operations Manager

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

vRealize Operations Manager exposes the information it gathers through an API that MANO and other components can use.

With this release, vRealize Operations Manager enables intent-based business and operational policies for dynamic optimization and capacity targets. Cost analysis is embedded into the solution, as with improved capacity planning and forecasting engines.

VMware vRealize Log Insight

VMware vRealize Log Insight delivers heterogeneous log management with dashboards, analytics, and third-party extensibility. It provides deep operational visibility and troubleshooting across physical, virtual, and cloud environments. Its indexing and machine learning based grouping provides log searches that help with troubleshooting issues.

VMware vRealize Network Insight

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

On the security aspects, vRealize Network Insight offers intelligence operations for SDN and security across the virtual and physical infrastructure by using micro-segmentation planning and policy distribution into NSX-T. The solution can be scaled, providing early warning on security policy violations.

VMware vRealize Orchestrator

The vCloud NFV platform provides closed-loop automation workflows to enable self-healing across VMs, hosts, and datastores at the infrastructure level. It also allows the CSP to create VNF-specific custom workflows for the faster time to resolution. vRealize Operations Manager integrates with vRealize Orchestrator through a management pack that provides access to the Orchestrator workflow engine for more remediation actions and the ability to run Orchestrator workflows directly from the vRealize Operations Manager user interface.

vCloud NFV Components

The vCloud NFV bundle packages together the essential building blocks to deploy an NFVI and VIM platform, featuring the newest releases of VMware solutions that are proven in production.

The components of this reference architecture are bundled as vCloud NFV Standard Edition and vCloud NFV Advanced Edition. This tiered structure provides flexibility in selecting editions and offers a simple path to customize deployments based on NFV use cases and requirements.

Table 5-1. vCloud NFV Components

Component	vCloud NFV Standard Edition Bundle	vCloud NFV Advanced Edition Bundle
VMware ESXi™	Core Component	Core Component
VMware vSphere® Replication™	Core Component	Core Component
VMware vRealize® Orchestrator™	Core Component	Core Component
VMware NSX-T® Data Center Standard	Core Component	Not Included
VMware NSX-T® Data Center Advanced	Not Included	Core Component
VMware vSAN™ Standard Edition	Not Included	Core Component
VMware vRealize® Operations™	Not Included	Core Component
VMware vRealize® Log Insight™	Not Included	Core Component
VMware vCloud Director®	Core Component	Core Component
VMware vCenter® Server Appliance™	Required Add-On	Required Add-On
VMware vRealize® Network Insight™	Optional Add-On	Optional Add-On
VMware Site Recovery Manager™	Optional Add-On	Optional Add-On

Design Principles

This reference architecture is designed considering key principle elements such as Deployment, Networking, Workload acceleration and Operational Intelligence.

Flexible Deployment Options

vCloud NFV can be deployed to support either a Three-Pod (Management, Resource, and Edge Pods) or a Two-Pod (Management and combined Resource and Edge Pods) configuration. The Three-Pod architecture provides the highest flexibility and performance, because Edge Nodes are dedicated to packet forwarding in and out of the virtual domain. CSPs can use the Two-Pod architecture for smaller starting deployments, where it is acceptable to combine the resource and Edge functionality in the same hosts.

Advanced Networking

To provide multitenant capabilities to distributed logical routers in the networking stack, vCloud NFV uses NSX-T Data Center to deploy multiple tiers of distributed routing through Tier-0 and Tier-1 gateways. Providers can use Tier-0 gateways, whereas tenants can use Tier-1 gateways. Network virtualization capabilities that are enabled through Geneve encapsulation provide a flexible capability in line with industry standards. NSX-T Data Center performance enhancements for the N-VDS and NSX Edge Node offer advanced network capabilities.

Workload Acceleration

vCloud NFV includes several features to support workloads that require high performance. These features are delivered with N-VDS with Enhanced Data Path Mode that can be used for high-performance workloads. Also, the North-South traffic forwarding between logical and physical domains can benefit from bare metal NSX Edge Nodes. This high-performance capability is available through Data Plane Development Kit (DPDK) based enhancements that come with NSX-T Data Center. Capabilities include optimizations through poll mode drivers, CPU affinity and optimization, and buffer management. DPDK provides support for workloads requiring acceleration.

Integrated Operational Intelligence

By using a framework that continuously collects data from local and distributed agents, vCloud NFV also provides the capability to correlate, analyze, and enable day 2 operations. Also, this analytical intelligence can be used with existing assurance engines for closed-loop remediation.

Core Reference Architecture 6

The VMware vCloud NFV platform is an evolution of the VMware NFV solution, based on the extensive customer deployment and the continued development of standards organizations such as the European Telecommunications Standards Institute (ETSI). The vCloud NFV 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 deliver services rapidly while ensuring their quality. With a fully integrated VIM, the same vCloud NFV infrastructure delivers a myriad of telecommunications use cases and facilitates the reusability of the service catalog based VNFs.

The vCloud NFV 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 various ways to construct a comprehensive, end-to-end solution that meets the business goals of CSPs. This document discusses how components can be used to create a vCloud NFV architecture.

This chapter includes the following topics:

Core Building Blocks

Architecting vCloud NFV by using well-defined modules allows the CSP to accelerate the deployment of the platform and reliably expand it when needed. The platform components are grouped into three distinct containments. The vCloud NFV platform uses the term Pods as a mean to streamline the NFV environment operations and delineate between different roles. For example, a cloud management team can easily operate the Management Pod, whereas 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. The Pod functions are the following:

Management Pod

Management functions are required to manage the NFV Infrastructure and the VNFs and their components. Management, orchestration, analytics functionality, and ancillary elements such as DNS, VNFM, and NFV-O are grouped into this category. Resource orchestration such as vCenter Server Appliance, NSX Manager, and vCloud Director for Service Providers are hosted in this Pod. All the analytics components (such as vRealize Operations Manager, vRealize Network Insight, vRealize Log Insight, vRealize Orchestrator) and the business continuity components (such as Site Recovery Manager and vSphere Replication) are located in this pod. Other management-related components such as NFV Orchestrators run in the Management Pod. OSS/BSS can be huge in sizing, and therefore their placement depends on the system itself.

Edge Pod

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. NSX Edge is hosted in a VM or bare metal appliance in the Edge Pod and handles all connectivity to the physical domain in the architecture. The type of networking traffic that traverses the Edge Pod is called North-South traffic.

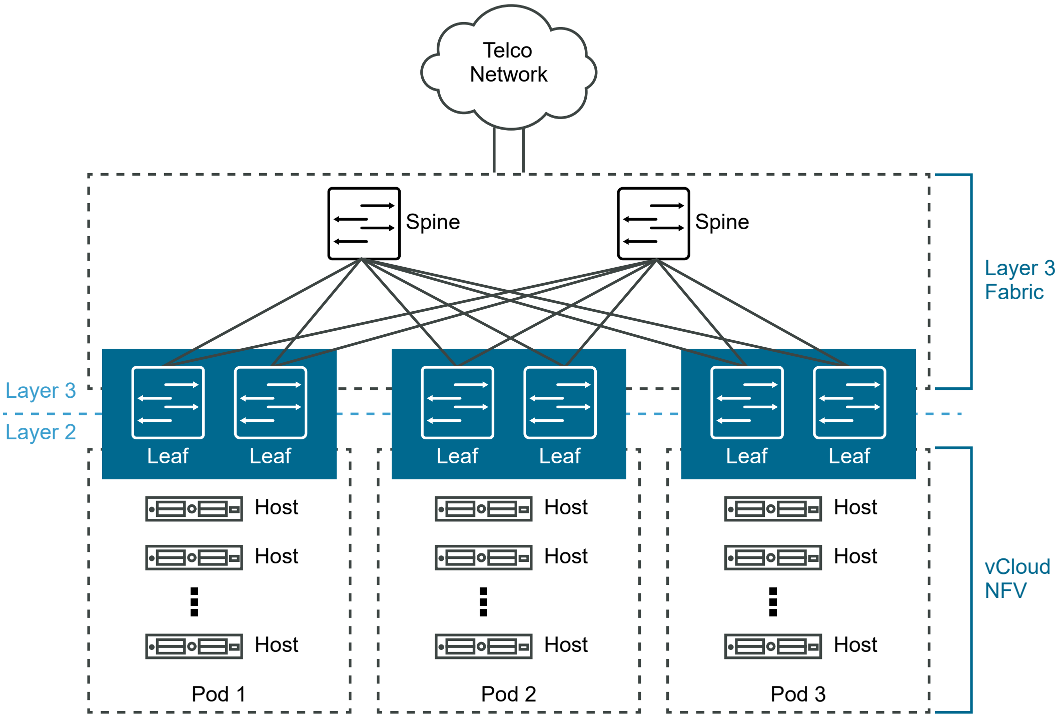
Resource Pod

VNFs and CNFs are placed in the Resource Pod; the VNFs form the virtual network service.

Physical Building Blocks

Traditional data center network fabrics are often designed with three tiers of switches: core, aggregation, and access. Access switches connect to aggregation switches, which in turn connect to the core switches. The design topology of the physical layer can impact efficiency and latencies.

Figure 6-1. Physical Network Design



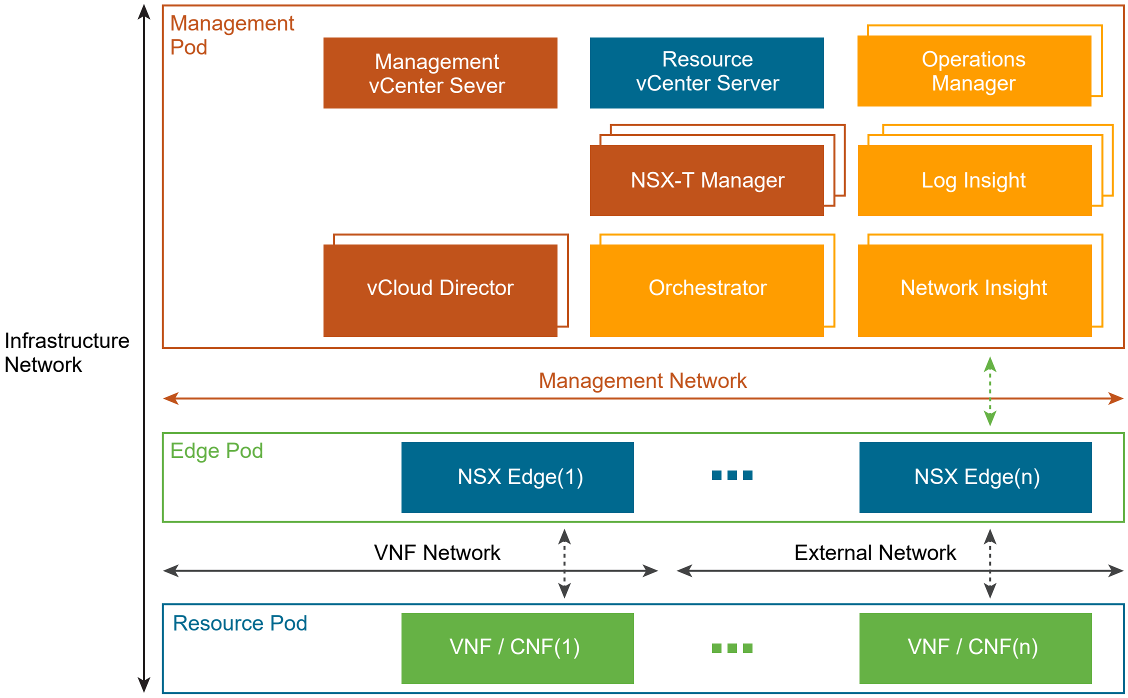
A two-tier leaf-and-spine network architecture is the preferred approach for building a new data center infrastructure. The two-tier architecture uses an access switch, or leaf, which is connected to an aggregation switch, or spine. The leaf switch provides connectivity between endpoints in the data center, while the spine switch provides high-speed interconnectivity between leaf switches. The leaf-and-spine network is connected in a full mesh, providing predictable communication and latency between endpoints. Ethernet connectivity is used from the host to the leaf switch, and the broadcast domain terminates at the leaf. External Border Gateway Protocol (eBGP) is often used for routing within the leaf-and-spine architecture.

Virtual Building Blocks

The virtual infrastructure design comprises the design of the software components that form the virtual infrastructure layer. This layer supports running telco workloads and workloads that maintain the business continuity of services. The virtual infrastructure components include the virtualization platform hypervisor, virtualization management, storage virtualization, network virtualization, and backup and disaster recovery components.

This section outlines the building blocks for the virtual infrastructure, their components, and the networking to tie all the components together.

Figure 6-2. Virtual Building Blocks



Storage Design

This reference architecture uses a shared storage design that is based on vSAN. vCloud NFV also supports certified third-party shared storage solutions, as listed in the VMware Compatibility Guide.

vSAN is a storage virtualization software that allows locally attached storage to be pooled and presented as a shared storage pool for all hosts in a vSphere cluster. This simplifies the storage configuration with a single datastore per cluster for the management and VNF workloads. With vSAN, VM data is stored as objects and components. An object consists of multiple components, which are distributed across the vSAN cluster based on the policy that is assigned to the object. The policy for the object ensures a highly available storage backend for the cluster workload, with no single point of failure.

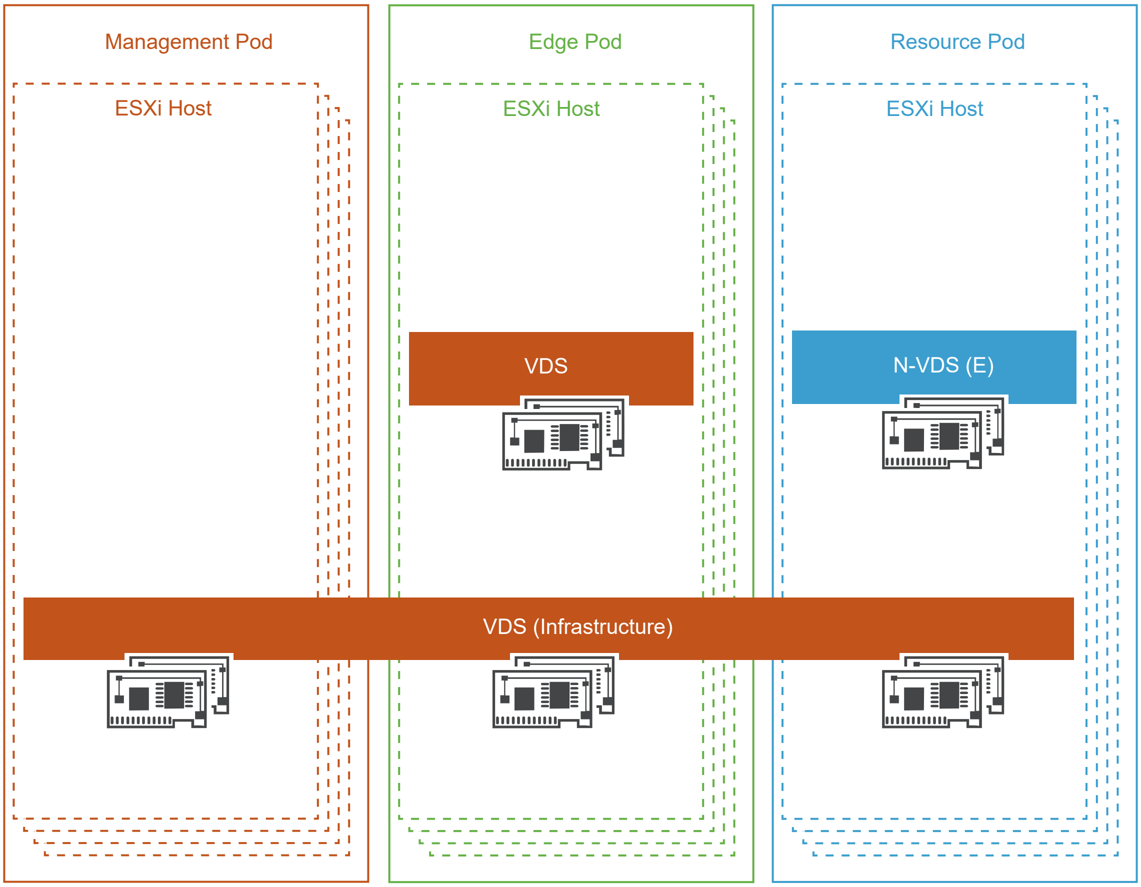
vSAN is a fully integrated hyperconverged storage solution that allows creating hybrid or all flash-based vSAN storage cluster. vSAN presents a flash-optimized, highly resilient, shared storage datastore to ESXi hosts and virtual machines. This allows for the control of capacity, performance, and availability through storage policies, on a per VM basis.

Network Design

The vCloud NFV platform consists of Infrastructure (VMkernel and VM network traffic) network. Infrastructure networks are host-level networks that connect hypervisors to physical networks. Each ESXi host has multiple port groups configured for each infrastructure network.

The hosts in each Pod are configured with VMware vSphere® Distributed Switch™ (VDS) that provide a consistent network configuration across multiple hosts. One VDS switch is used for VM networks and VMkernel networks. The N-VDS Enhanced switch is used as the transport for the telco workload traffic.

Figure 6-3. Virtual Network Design



Infrastructure networks are used by the ESXi hypervisor for vMotion, vSphere Replication, vSAN traffic, management, and backup as well as Management VMs to communicate with each other. A separate N-VDS Enhanced switch is used for workload traffic. Each N-VDS 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.

VMs can be connected to each other over a VLAN or over Geneve-based overlay tunnels. Both networks are designed according to the requirements of the workloads that are hosted by a specific Pod. The infrastructure VDS switch and networks remain the same regardless of the Pod function. However, the VM networks depend on the networks that the specific Pod requires. The VM networks are created by NSX-T Data Center to provide enhanced networking services and performance to the Pod workloads. The ESXi host's physical NICs are used as uplinks to connect the distributed switches to physical network switches. All ESXi physical NICs connect to layer 2 or layer 3 managed switches on the physical network. It is common to use two switches for connecting to the host physical NICs for redundancy purposes.

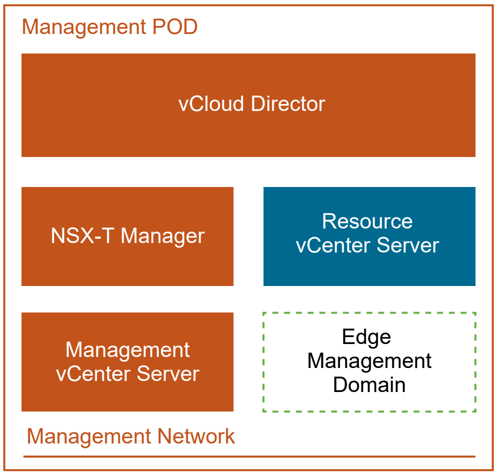
The VMkernel traffic networks used in the Pods are:

- **ESXi Management Network**: The network for the ESXi host management traffic.
- **vMotion Network**: The network for the VMware vSphere® vMotion® traffic.
- **vSAN Network**: The network for the vSAN shared storage traffic.
- **Backup Network**: The network that is dedicated to offline storage such as NFS and used for workload backup and restore as needed.
- **Replication Network**: The network that is used for replicating data for data protection.

Management Pod

This section describes the design for the Virtualized Infrastructure Management (VIM) components: vCenter Server Appliance, NSX Manager, vCloud Director for Service Providers, and Edge Management domain component.

Figure 6-4. Management Pod



In addition to these core components, the Management Pod also contains the operations management components.

Components

The Management Pod contains the components that manage the vCloud NFV runtime environment.

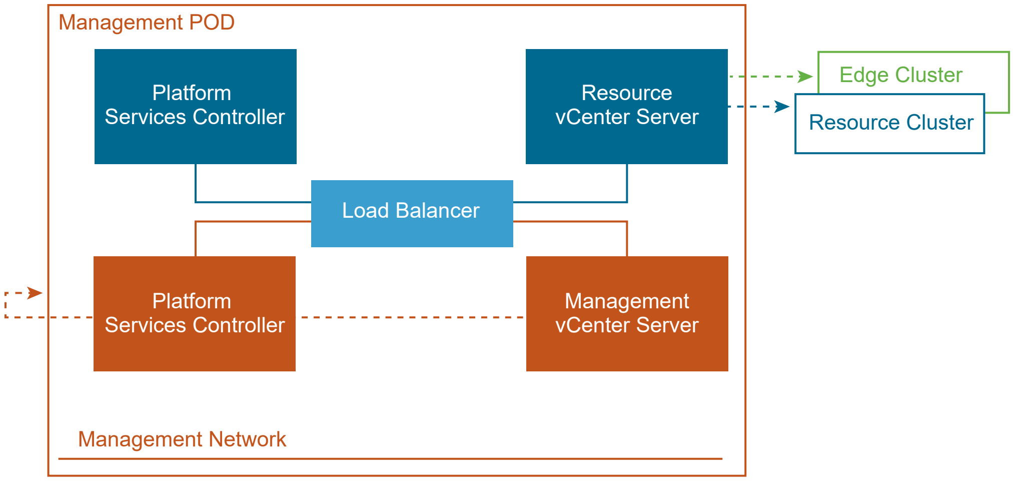
vCenter Server

The Management Pod is implemented as a cluster that is managed by the Management vCenter Server instance. The Management Pod hosts the core management domain and all the edge management domains. To form the foundation of a carrier-grade virtualized infrastructure, the components of the Management Pod benefit from the cluster features such as resource management, high availability, and resiliency. A Resource vCenter Server is deployed to oversee the Resource Pods and Edge Resource vCenter Server is deployed to manage Edge sites.

Each vCenter Server instance is a virtual appliance that is deployed with an embedded database. The vCenter® Server Appliance™ is preconfigured, hardened, and fast to deploy. The appliance allows for a simplified design, eases management, and reduces administrative efforts. vCenter Server Appliance availability is ensured by using either vSphere High Availability or vCenter High Availability (vCenter HA) cluster, which is realized through three vCenter Server Appliance instances.

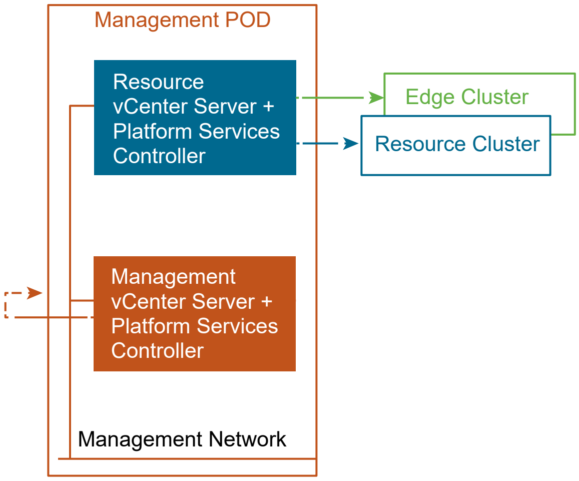
The Platform Services Controller contains common infrastructure security services such as VMware vCenter® Single Sign-On, VMware Certificate Authority, licensing, service registration, and certificate management services. The Platform Services Controller handles identity management for administrators and applications that interact with the vSphere platform. The Platform Services Controller may be deployed as a load-balanced pair of appliances per vCenter Server as shown in the following diagram:

Figure 6-5. vCenter Server with External Platform Services Controller



Alternatively, the Platform Services Controller and its related services may be embedded within the vCenter Server Appliance. This eliminates the need for separate Platform Services Controller VM instances and their corresponding load balancers, thus simplifying its deployment and administration and also reducing the management components footprint.

Figure 6-6. vCenter Server with Embedded Platform Services Controller


Management POD
Resource vCenter Server + Platform Services Controller
Edge Cluster
Resource Cluster
Management vCenter Server + Platform Services Controller
Management Network

Data backup and restore of each vCenter Server instance and its embedded Platform Services Controller is provided by using the native backup service that is built in the appliances. This backup is performed to a separate storage system by using network protocols such as SFTP, HTTPS, and SCP. VCSA has a built-in feature to configure the native vcsa scheduler for the backup.

VMware NSX-T Data Center

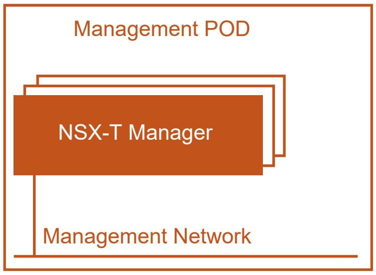
NSX Manager is a management plane component of the NSX-T system. It provides the ability to create, configure, and monitor NSX-T Data Center components, such as segments, gateways, and NSX Edge Nodes.

NSX Manager provides an aggregated system view and centralized network management of NSX-T Data Center. It provides a method for monitoring and troubleshooting workloads that are attached to the virtual networks that NSX-T Data Center creates. NSX-T Data Center provides configuration and orchestration of logical networking components such as segments and gateways, networking services, Edge services, security services, and distributed firewall capabilities.

NSX Manager contains an advanced distributed state management system that controls virtual networks and overlay transport tunnels. NSX Controller is part of the converged NSX Manager appliance and deployed as a three-node highly available cluster, responsible for the programmatic deployment of virtual networks across the entire NSX-T Data Center architecture. The control plane is split into two parts in NSX-T Data Center:

- **Central Control Plane (CCP)** that runs on the NSX Controller cluster nodes. The CCP computes some ephemeral runtime state based on configuration from the management plane and disseminates information reported through the local control plane by the data plane elements. The CCP is logically separated from all data plane traffic, therefore any failure in the control plane does not affect the existing data plane operations.
- **Local Control Plane (LCP)** that runs on the transport nodes, adjacent to the data plane it controls. The LCP monitors local link status, computes most ephemeral runtime state based on updates from the data plane and CCP, and pushes the stateless configuration to forwarding engines. The LCP shares fate with the data plane element that hosts it.

Figure 6-7. NSX Manager

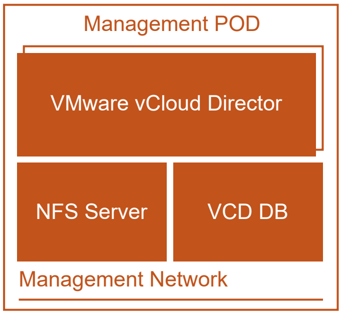


NSX Managers are deployed as a cluster of three manager nodes for high availability with anti-affinity rules configured to ensure that the NSX Managers and the CCP reside on separate hosts to protect against host failures. The LCP shares fate with the data plane element that hosts it, while the CCP inherits the same fate as the NSX Manager in terms of availability. The NSX Manager communicates with Edge clusters over a common management network. The management components of the vCloud NFV platform communicate over the same management network to request network services from the NSX Manager.

VMware vCloud Director for Service Providers

VMware vCloud Director for Service Providers is an abstraction layer that operates on top of other VIM components, vCenter Server, and NSX Manager. A highly available vCloud Director implementation that uses multiple load-balanced vCloud Director cells are deployed in a vCloud Director Server Group. All cells in the server group are stateless and use a shared highly available clustered database. Each cell contains all the software components required for vCloud Director. A cell can run on its own, but multiple cells running in an active-active cluster are used for scalability and redundancy.

Figure 6-8. vCloud Director Management Components



vCloud Director builds a secure, multitenant virtual environment by pooling virtual infrastructure resources to Virtual Data Centers (VDCs) and exposing them to users through Web-based portals and APIs as fully-automated, catalog-based services.

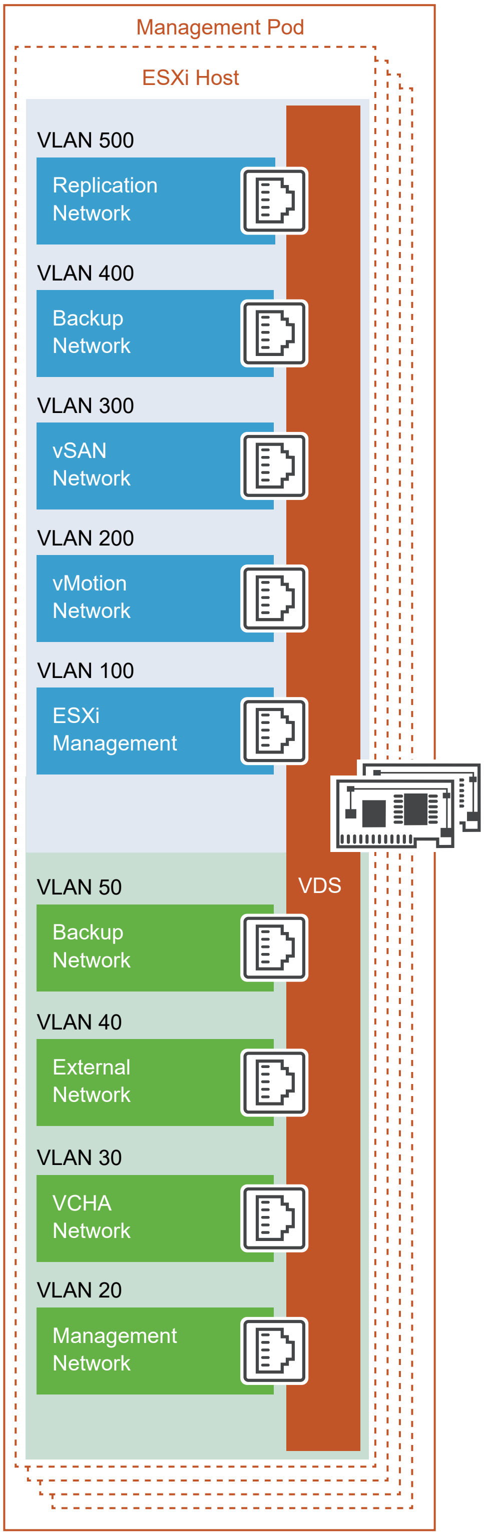
A fundamental concept in vCloud Director is that of the tenant. A tenant is a logically isolated construct representing a customer, department, network function, or service, used to carve out infrastructure resources and deploy VNF workloads. vCloud Director isolates administrative boundaries to NFVI tenants. VNF workload resource consumption is therefore segmented from other VNF workloads, even though the VNFs can share the resources.

vCloud Director implements the open and publicly available vCloud API, which provides compatibility, interoperability, and programmatic extensibility to Network Equipment Providers (NEPs) and their VNF Managers. The vCloud Director capabilities can be extended to create adapters to external systems including OSS/BSS.

Networking

The Management Pod networking consists of infrastructure and VM networks. The following diagram shows all the virtual switches and port groups of the Management Pod:

Figure 6-9. Management Pod Networking



Edge Pod

The Edge Pod provides the fabric for North-South connectivity to the provider networks.

Components

The Edge Pod components provide the fabric for North-South connectivity to the provider networks. Multiple configurations can be used for performance, scale, and Edge services.

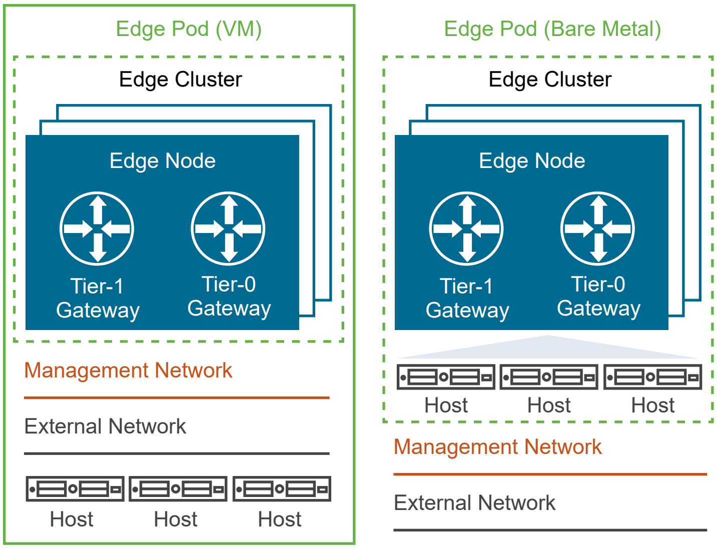
Edge Nodes

An Edge Node is the appliance that provides physical NICs to connect to the physical infrastructure and to the virtual domain. Edge Nodes serve as pools of capacity, dedicated to running network services that cannot be distributed to the hypervisors. The network functionalities of the Edge node include:

- Connectivity to the physical infrastructure.
- Edge services such as NAT, DHCP, firewall, and load balancer.

Edge nodes are available in two form-factors: VM and bare metal. Both leverage DPDK for faster packet processing and high performance. Depending on the use case, the appropriate form-factor is deployed.

Figure 6-10. Edge Pod Components



The NSX bare metal Edge runs on a physical server and is installed by using an ISO file or PXE boot. The bare metal Edge is recommended for production environments where services such as NAT, firewall, and load balancer are needed in addition to Layer 3 unicast forwarding. A bare metal Edge differs from the VM form-factor Edge in terms of performance. It provides sub-second convergence, faster failover, and throughput greater than 10 Gbps.

The VM form-factor of NSX Edge is installed by using an OVA, OVF, or ISO file. Depending on the required functionality, there are deployment-specific VM form-factors.

Edge Clusters

Edge nodes are deployed as pools of capacity (a cluster), dedicated to running network services that cannot be distributed to the hypervisors. An Edge cluster can either be all VM or all bare metal form-factors.

The Edge cluster provides scale-out, redundant, and high-throughput gateway functionality for logical networks. Scale-out from the logical networks to the Edge nodes is achieved by using ECMP. There is total flexibility in assigning gateways to any specific clusters. Tier-0 and Tier-1 gateways can be hosted on either the same or different Edge clusters. Centralized services must be enabled for the Tier-1 logical router to coexist in the same cluster.

There can be only one Tier-0 gateway per Edge node, however, multiple Tier-1 gateways can be hosted on one Edge node.

In addition to providing distributed routing capabilities, the Edge cluster enables Edge services at a provider or tenant scope. As soon as one of these Edge services is configured or an uplink is defined on the gateways to connect to the physical infrastructure, a Service Router (SR) is instantiated on the Edge Node. Similar to the compute nodes in NSX-T Data Center, the Edge Node is also a transport node and it can connect to more than one transport zone: one for overlay and another for N-S peering with external devices.

A maximum of eight Edge Nodes can be grouped in an Edge cluster. A Tier-0 gateway supports a maximum of eight equal-cost paths, thus a maximum of eight Edge Nodes are supported for ECMP. Edge Nodes in an Edge cluster run Bidirectional Forwarding Detection (BFD) on both tunnel and management networks to detect the Edge Node failure. The BFD protocol provides fast detection of failure for forwarding paths or forwarding engines, improving convergence. Bare metal form factors can support sub-second convergence.

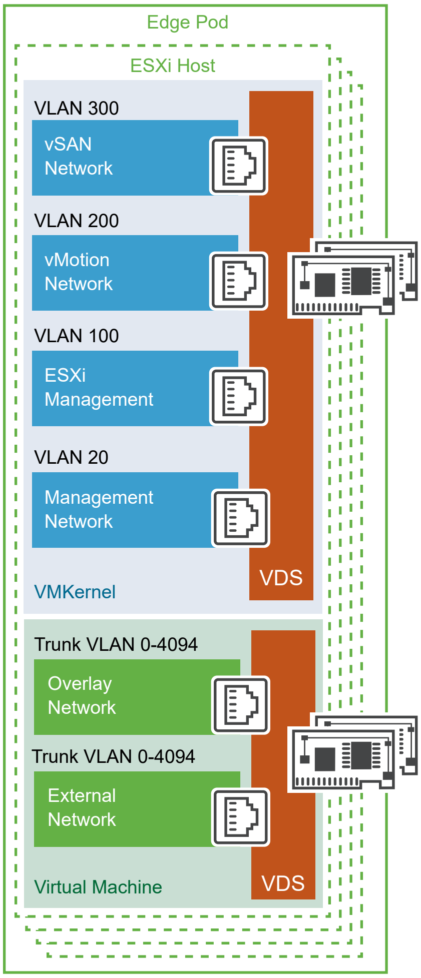
NSX-T Data Center supports static routing and the dynamic routing protocol BGP on Tier-0 gateways on interfaces connecting to upstream routers. Tier-1 gateways support static routes but do not support any dynamic routing protocols.

For more information, see the VMware® NSX-T Reference Design Guide.

Networking

The Edge Pod virtual network largely depends on the network topology that is required by the VNF workloads. In general, the Edge Pod has the infrastructure networks, networks for management and control plane connectivity, and networks for workloads.

Figure 6-11. Edge Pod Networking



Resource Pod

This section describes the components of the Resource Pod and their functions. VNFs are placed in the Resource Pod that forms the virtual network services.

Components

The Resource Pod provides the runtime environment for the network functions. The section covers the logical tenancy and networking components.

Provider VDC

The Provider VDC is a standard container for a pool of compute, storage, and network resources from a single vCenter Server instance. During the deployment, a VNF can use one or more Provider VDCs per site, or multiple VNFs can share a single Provider VDC with different organization VDCs. The requirements of each VNF are assessed as part of the onboarding process.

Organization

Organizations are the unit of multi-tenancy within vCloud Director for Service Providers. An NFVI tenant can be defined in multiple ways in the vCloud NFV platform. An NFVI tenant can, for example, be a VNF provider, but it can also be defined per Telco application consisting of multiple VNF providers who provide a joint service. The necessary level of separation between different VNFs and the creation of NFVI tenants should be identified during the onboarding process.

Organization VDCs

An organization VDC is a subgroup of compute, storage, and network resources that are allocated from a Provider VDC and mapped to an organization. Multiple organization VDCs can share the resources of the same Provider VDC. An organization VDC is the environment where VNF workloads are deployed and executed.

Quotas on organizations set limits on the vCloud Director tenant resources. Organization VDCs allow providing resource guarantees for workloads from the resource quota available to the organization and avoid noisy neighbor scenarios in a multitenant environment.

vApps

A vApp is a container for multiple virtual machines and the standard unit for workloads in vCloud Director. vApps contain one or more virtual machines and networks and can be imported or exported as an OVF file. In the vCloud NFV platform, a VNF instance can be a vApp.

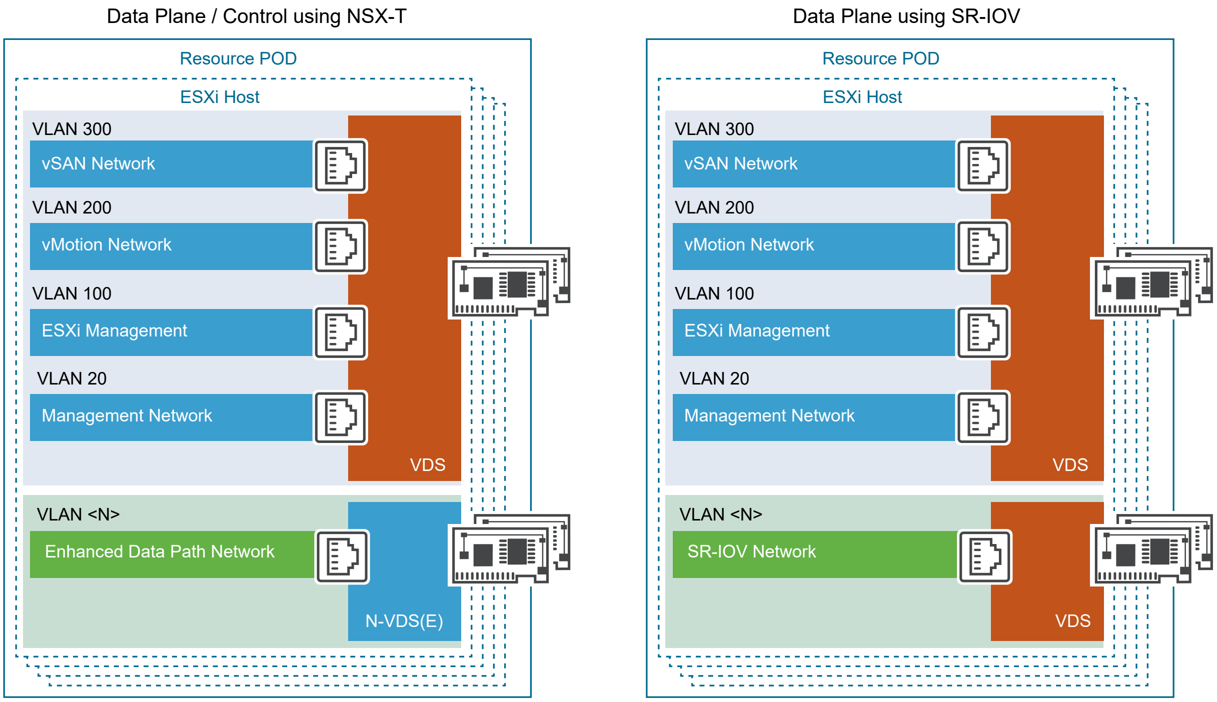
Networking

The networking of the Resource Pod is highly dependent on the network topology that is required by the telco workloads, which tenants deploy. Tenant workloads require a certain set of networking building blocks.

Segments

Segments are the layer 2 networks created by NSX-T Data Center to provide connectivity between its services and the VMs. Segments form the basis of the tenant networks in the vCloud NFV platform. The primary component in the data plane of the transport nodes is N-VDS. N-VDS forwards traffic between components running on the transport node (that is between VMs) or between VMs and the physical network. In the latter case, N-VDS must own one or more physical interfaces (physical NICs) on the transport node. As with other virtual switches, an N-VDS cannot share a physical interface with another N-VDS. It can coexist with another N-VDS each using a separate set of physical NICs.

Figure 6-12. Resource Pod Networking



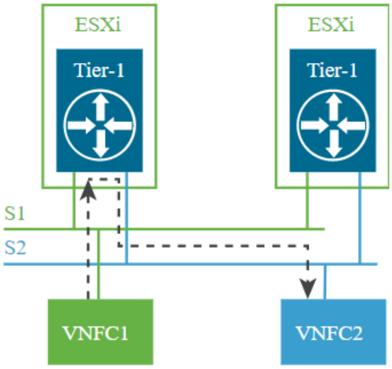
Logical Routing

The NSX-T Data Center platform provides the ability to interconnect both virtual and physical workloads that are deployed in different logical layer 2 networks. NSX-T enables the creation of network elements like segments and gateways as software logical constructs and embeds them in the hypervisor layer, abstracted from the underlying physical hardware.

East-West Traffic

Configuring a gateway through the NSX Manager instantiates a gateway on each hypervisor. For the VNFs hosted on the same hypervisor, the East-West traffic does not leave the hypervisor for routing. The gateway is also responsible for routing East-West traffic between hypervisors. The Tier-1 gateway is deployed and managed by the org VDC of the vCloud NFV platform, for routing services between their respective org VDC networks within their tenancy.

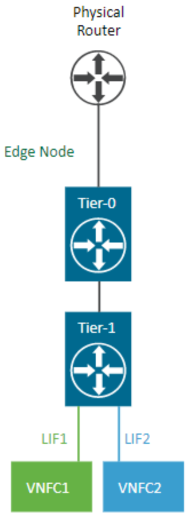
Figure 6-13. East-West Traffic



North-South Traffic

In addition to providing optimized distributed and centralized routing functions, NSX-T Data Center supports a multi-tiered routing model with logical separation between the provider routing function and the tenant routing function. This way, the concept of multitenancy is built in the routing model. The top-tier logical router is called a Tier-0 gateway, whereas the bottom-tier logical router is called a Tier-1 gateway. Northbound, the Tier-0 logical gateway connects to one or more physical routers or layer 3 switches and serves as an on/off-ramp to the physical infrastructure. Southbound, the Tier-0 gateway connects to one or more Tier-1 gateways.

Figure 6-14. North-South Traffic



This model also eliminates the dependency on a physical infrastructure administrator to configure or change anything on the physical infrastructure when a new tenant is configured in the data center. For a new tenant, the Tier-0 gateway simply advertises the new tenant routes that are learned from the tenant Tier-1 gateways on the established routing adjacency with the physical infrastructure.

Deployment Options

7

Two design configurations are optimal with the vCloud NFV platform: a compact Two-Pod configuration and a Three-Pod configuration. These configurations can be deployed in the data center to meet target design and scale objectives.

This chapter includes the following topics:

- Three-Pod Configuration
- Two-Pod Configuration

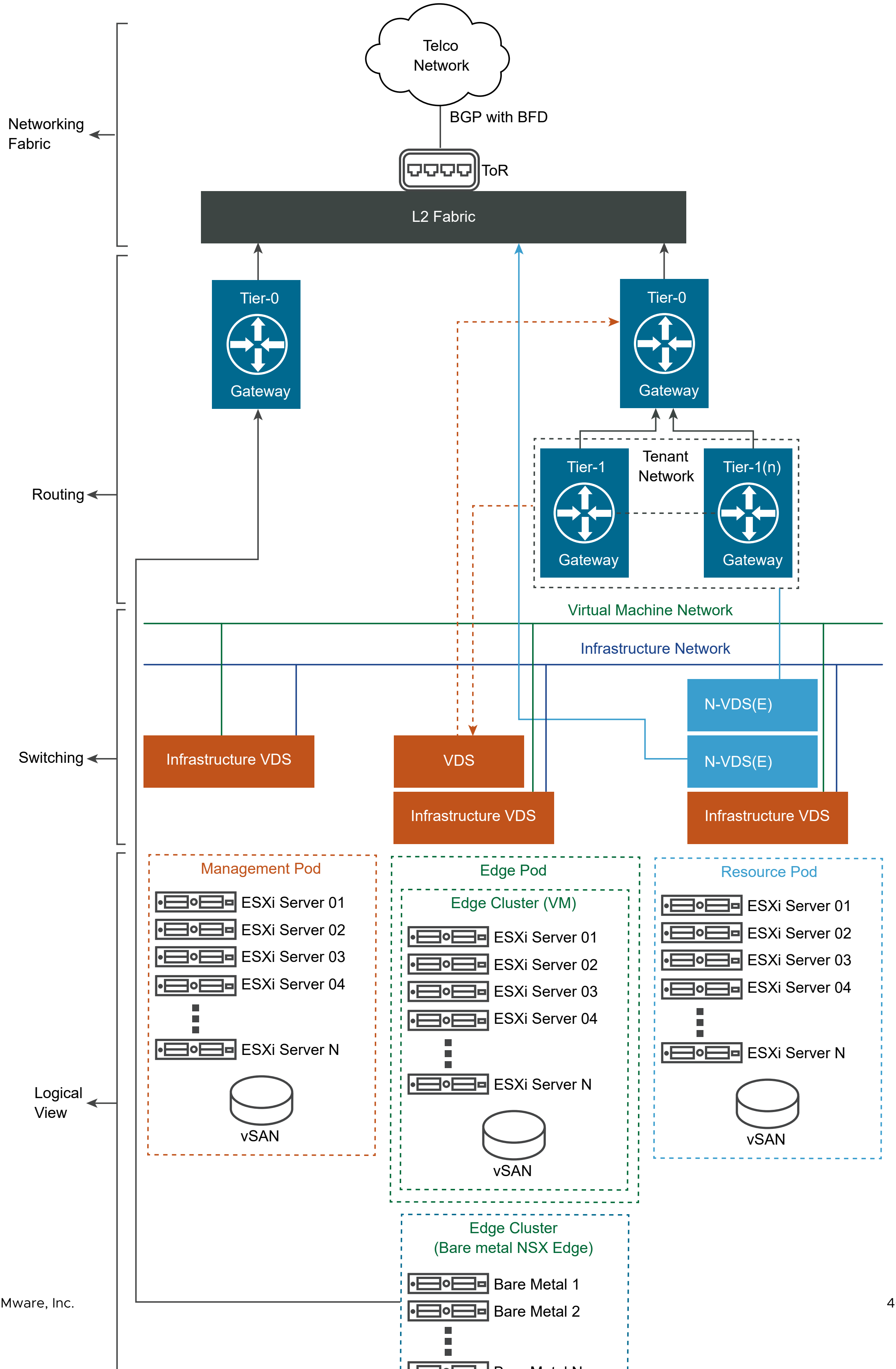
Three-Pod Configuration

The Three-Pod design separates the vCloud NFV functional blocks by using a Management Pod, Edge Pod, and Resource Pod for their functions. The initial deployment of a Three-Pod design consists of three vSphere clusters, one cluster per Pod respectively. Clusters can be scaled up by adding ESXi hosts, whereas Pods can be scaled up by adding clusters. The separation of management, Edge, and resource functions in individually scalable Pods allows the CSPs to plan capacity according to the needs of the specific function that each Pod hosts. This provides greater operational flexibility.

The following diagram depicts the physical representation of the compute and networking connectivity and the logical layers of the switching and routing fabric.

The following diagram illustrates the deployment of both the Edge form factors (VM and Bare Metal). Your deployment will include only one form factor.

Figure 7-1. Three-Pod Conceptual Design



The initial deployment of a Three-Pod design is more hardware intensive than the initial deployment of a Two-Pod design. Each Pod in the design scales up independently from the others. A Three-Pod design consists of the same components as in a Two-Pod design, as the way functions are combined to form the solution is different in each design. Regardless of the Pod design that is used to create the NFVI, VNFs perform the same way.

Logical View

- **Management Pod**: Hosts all the NFV management components. Its functions include resource orchestration, analytics, BCDR, third-party management, NFV-O, and other ancillary management.
- **Edge Pod**: Hosts the NSX-T Data Center network components, which are the NSX Edge nodes. Edge nodes participate in East-West traffic forwarding and provide connectivity to the physical infrastructure for North-South traffic management and capabilities. Edge nodes can be deployed in a VM and bare metal form-factor to meet capacity and performance needs.
- **Resource Pod**: Provides the virtualized runtime environment (compute, network, and storage) to execute workloads.

Routing and Switching

Before deploying the Three-Pod configuration, consider a VLAN design to isolate the traffic for infrastructure, VMs, and VIM.

The vSphere Distributed Switch port groups have different requirements based on the Pod profile and networking requirements. For example, a vSphere Distributed Switch can be used for both the VMkernel traffic and VM management traffic. While the switching design is standardized for the Management and Edge Pods, the Resource Pod offers flexibility with two classes of NSX-T Data Center switches (N-VDS Standard and N-VDS Enhanced). The N-VDS switch offers the overlay and VLAN networking. The N-VDS Enhanced switch offers acceleration by using DPDK to workloads.

The NSX-T Data Center's two-tiered routing fabric provides the separation between the provider routers (Tier-0) and the tenant routers (Tier-1).

The Edge nodes provide the physical connectivity to the CSP's core and external networking.

Note NSX-T Data Center requires a minimum of 1600 MTU size for overlay traffic. The recommended MTU size is 9000.

Design Considerations

A Three-Pod configuration provides flexibility, performance, and VNF distribution design choice for telco workloads.

Footprint

The best practice when using vSAN as shared storage for all clusters is to use a minimum of four hosts per cluster for the initial deployment, which sums up a total of 12 hosts for the entire configuration. This creates a balance between the implementation footprint and resiliency and maintains the operational requirements for each Pod. A mid-level class of servers can be used in this configuration because the design maximizes the scale flexibility.

Scale Flexibility

The separation of Pods provides maximum scalability to the design. Individual Pods can be scaled to meet the target services, delivery, and topology objectives.

The Resource and Edge clusters are sized according to the VNFs and their respective networking requirements. CSPs must work with the VNF vendors to gather the requirements for the VNFs to be deployed. This information is typically available in deployment and sizing guides.

As more tenants are provisioned, CSPs must provide additional resources to the Resource Pod to support the VNF growth. The increase in the VNF workloads in the Resource Pod can lead to an increase in the North-South network traffic, which in turn requires adding more compute resources to the Edge Pod so that to scale up the Edge Nodes. CSPs must closely monitor and manage the resource consumption and capacity that is available to the Edge Pod when the Edge Nodes are scaled. For higher performance, an Edge Pod can also comprise bare metal Edges that are grouped in an Edge Cluster in NSX-T Data Center.

A high level of resiliency is provided by using four-host clusters with vSAN storage in the initial deployment. Four-host clusters also ensure a highly available Management Pod design, because clustered management components such as vCenter Server active, standby, and witness nodes, can be placed on separate hosts. This design principle is used for clustered vCloud Director components such as database nodes.

The initial number and sizing of management components in the Management Pod should be planned. As a result, the capacity that is required for the Management Pod can remain steady. When planning the storage capacity of the Resource Pod, the CSP should consider the operational headroom for VNF files, snapshots, backups, VM templates, OS images, upgrade files, and log files.

When the Resource Pod is scaled up by adding hosts to the cluster, the 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 pooled capacity that is available. Allocation settings for existing tenants must be modified before they can benefit from increased resource availability. vCloud Director Provider VDCs are added to scale out the Resource Pod. Additional compute nodes can be added by using the vSphere Web Client extension of vCloud Director.

In the Edge Pod, additional Edge Nodes can be deployed in the cluster along with the physical leaf-spine fabric for higher capacity needs.

Performance

The platform is optimized for workload acceleration and the Three-Pod design offers maximum flexibility to achieve these goals. The configuration can support the needs of use cases with high-bandwidth requirements. A dedicated accelerated Resource Pod or a hybrid standard and accelerated Resource Pod can be considered when designing for workload distribution. The separate Edge Pod provides not only the separation benefit but also alternatives with VM and bare metal options for the Edge Nodes.

Function Distribution

Network function designs are evolving to the disaggregated control and data plane functions. The separation also maximizes the distribution of data plane functions closer to the edge of the network with centralized control and management planes. A separated Resource or Edge Pod design allows for having different control plane versus data plane configurations and scale designs depending on the service offers.

Operations Management

The Management Pod provides centralized operation function across the deployment topology, be it a single or distributed data centers.

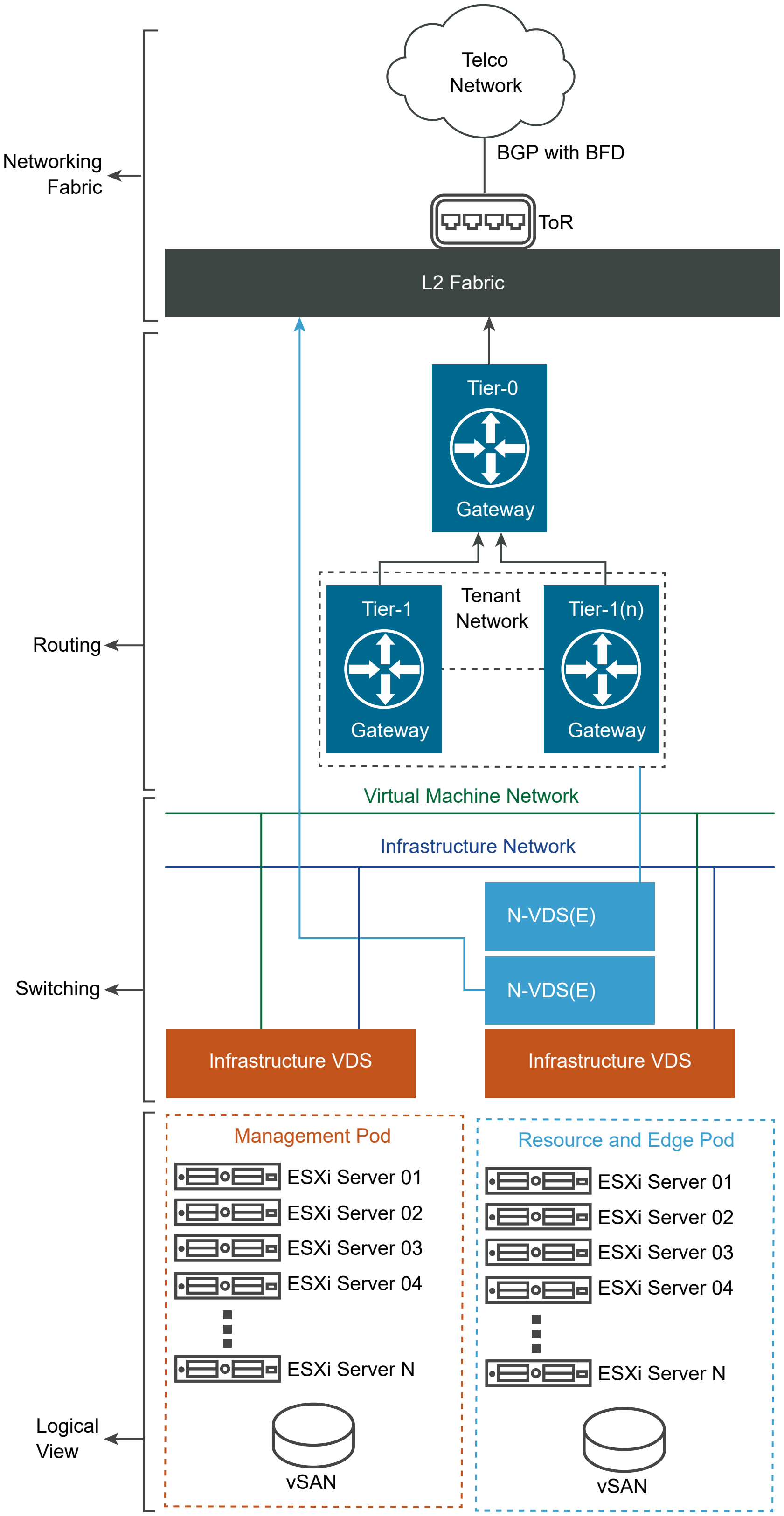
Two-Pod Configuration

The vCloud NFV facilitates combining the Edge and resource functionality in a single, collapsed Pod that provides a small footprint. CSPs can use a Two-Pod design to gain operational experience with the vCloud NFV platform. As demand grows, they can 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, whereas Pods scale up by adding new clusters to the existing Pods. This design ensures that the management boundaries are clearly defined, capacity is managed, and resources are allocated based on the functionality that the Pod hosts. The vCloud NFV VIM components allow for fine grained allocation and partitioning of resources to the workloads, regardless of the scaling method that is used.

The following diagram shows a Two-Pod design with all management functions located centrally in the Management Pod. Edge and resource functions are combined in the collapsed Edge and Resource Pod. During the initial deployment, two clusters of ESXi hosts are used: one for the Management Pod and another for the collapsed Edge and Resource Pod. Additional clusters can be added to each Pod as the infrastructure is scaled up.

Figure 7-2. Two-Pod Conceptual Design



Logical View

- **Management Pod**: This Pod hosts all NFV management components. These components include resource orchestration, analytics, BCDR, third-party management, NFVO, and other ancillary management.
- **Edge and Resource Pod**: This Pod provides the virtualized runtime environment (compute, network, and storage) to execute workloads. It also consolidates the NSX Edge Node to participate in the East-West traffic and to provide connectivity to the physical infrastructure for North-South traffic management capabilities. Edge Nodes can be deployed in a VM form factor only.

Routing and Switching

Before deploying a Two-Pod configuration, a VLAN design needs to be considered as a best practice to isolate the traffic for infrastructure, VMs, and VIM.

The general design of the routing and switching fabric is similar to the Three-Pod design, with the resource and Edge fabric converged into a single pod.

Design Considerations

A Two-Pod configuration provides a compact footprint design choice for telco workloads.

Footprint

A high level of resiliency is provided by using four-host clusters with vSAN storage in the initial deployment. Four-host clusters also ensure a highly available Management Pod design, because clustered management components such as vCenter Server active, standby, and witness nodes can be placed on separate hosts. The same design principle is used for clustered vCloud Director components such as database nodes.

With high-end class servers, this design maximizes the resource utilization and densification of workloads and Edge Nodes in the same Pod.

Scale Flexibility

The initial number and sizing of management components in the Management Pod should be planned. As a result, the capacity requirement for the Management Pod can remain steady. When planning the storage capacity of the Management Pod, the CSP should consider the operational headroom for VNF files, snapshots, backups, VM templates, OS images, upgrade files, and log files.

The sizing of the collapsed Edge and Resource cluster changes according to the VNF and networking requirements. When planning the capacity of the Edge and Resource Pod, tenants must work with VNF vendors to gather requirements for the VNF to be deployed. Such information is typically available from VNF vendors as deployment and sizing guidelines. These guidelines are directly related to the scale of the VNF service, for example, the number of subscribers to be supported. Also, the capacity utilization of Edge Nodes must be considered, especially when more instances of Edge Nodes are deployed to scale up as the number of VNFs increases.

When scaling up the Edge and Resource Pod by adding hosts to the cluster, the 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 pooled capacity that is available. Allocation settings for existing tenants must be modified before they can benefit from increased resource availability.

Operations Management

The Management Pod provides a centralized operation across the deployment topology, be it a single or distributed data center.

Next Generation Data Center Evolution

8

This section describes a set of solutions and use case scenarios to modernize the CSP cloud infrastructure environment with vCloud NFV 3.2.1.

This chapter includes the following topics:

Private Data Center NFV Transformation

Transform the cloud infrastructure in the CSP private data centers with the vCloud Director IaaS layer together with the VMware core virtualization technologies.

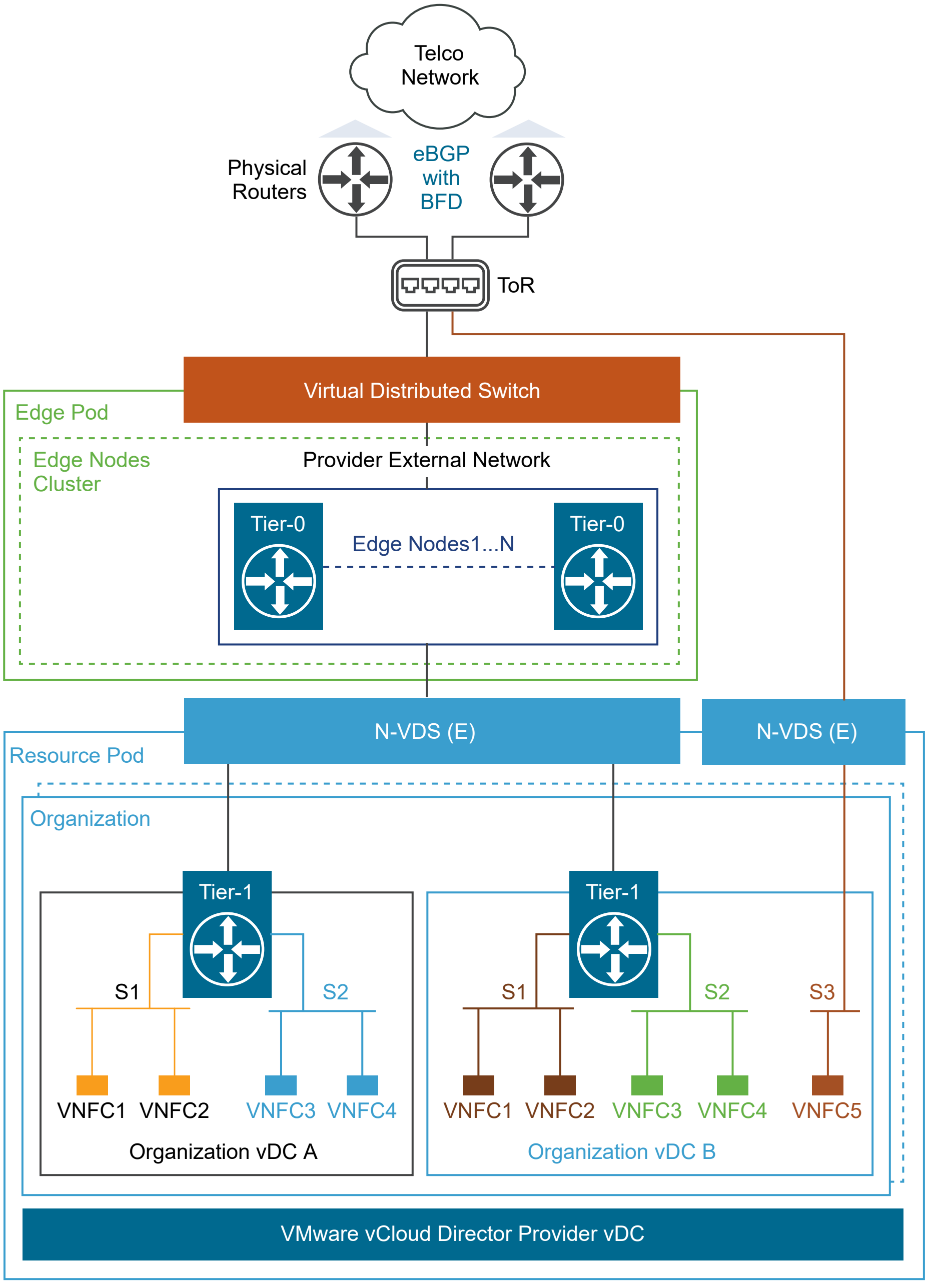
Scope

This release of vCloud NFV introduces vCloud Director together with NSX-T Data Center, which is the next-generation advanced networking stack of the NFVI. vCloud NFV explores the transformation of the CSP private cloud environments to next-generation NFV networks by using multitenant design objectives.

Considering a Three-Pod design, the following diagram illustrates how the Resource Pod enables accelerated and normal workloads with the N-VDS Enhanced mode. A separate Edge Pod provides external connectivity with different options for scale and performance.

The diagram illustrates how a fully integrated vCloud Director Provider VDC, Organization, Organization VDC, NSX-T segments, Tier-1 gateways can be used to provide a multitenant environment for deploying VNFs.

Figure 8-1. Private Cloud with Multiple Tenants



The Edge Pod hosts both the Tier-0 and Tier-1 gateways that provide North-South and East-West network connectivity for workloads. The NSX Manager is the management component that CSPs can use to provision these networks. After the North-South and East-West networks are mapped to provider networks in vCloud Director, tenants consume them for their workload connectivity requirements.

Design Objectives

You can transform the private data center by using a common shared pool of resources with multitenant compute, storage, and network isolation. Use the NSX-T Data Center advanced networking and acceleration capabilities for control and data plane workloads.

Site Design

The Management, Resource, and Edge Pods are deployed in each data center that is part of a site. Pods that belong to a site can be scaled to meet the use cases and requirements of their data centers.

Resource Pods in each data center are sized to meet the scale and traffic requirements of the VNFs and other data plane functions.

A pair of availability zones should be created in each site for redundancy and fault tolerance. A group of homogenous hosts, also called host aggregates, should be mapped across the availability zones. A single tenant (Organization and Organization VDC) map to the host aggregates to deploy workloads for control and user plane functions.

Compute Resource Isolation

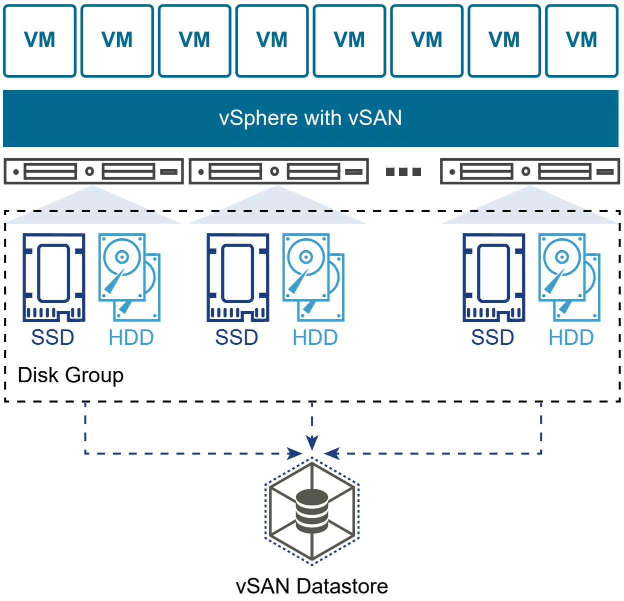
Resource pools represent a unit of CPU, memory, and storage that are portioned from the shared infrastructure pool. These resource pools are attached to Provider VDCs that in return are the abstracted pools of resources that tenants consume through Organization VDCs. The resources of a Provider VDC, and hence the resources available to tenants, can be scaled by scaling the underlying resource pool to meet future resource demands.

By mapping different Provider VDCs to different resource pools backed by different vSphere host clusters, CSPs can meet different SLAs as per tenant resource requirements.

Shared Storage

In this design, vSAN is used as a shared storage solution. Other storage options with NFS, iSCSI, and Fiber Channel are also supported. vSAN aggregates local or directly attached data storage devices to create a single storage pool that is shared across all hosts in the vSAN cluster. vSAN eliminates the need for external shared storage and simplifies the storage configuration and VM provisioning.

Figure 8-2. Shared Storage with vSAN



Using a shared datastore is recommended for high availability and fault tolerance. In case of a host failure, vSphere HA can restart the VMs on another host.

A cluster in vCenter Server provides the management of software-defined storage resources the same way as with compute resources. Instead of CPU or memory reservations, limits, and shares, storage policies can be defined and assigned to VMs. The policies specify the characteristics of the storage and can be changed as the business requirements change.

Design for Tenants

The diagram in the Scope section of this use case shows how vCloud Director Compute Cluster, Organization, Organization VDC, NSX-T Data Center segments, and Tier-1 gateways can be used to provide a multitenant environment for VNF deployments. The Edge Pod hosts the NSX-T Data Center Tier-0 gateways that are used as provider routers for external network connectivity. The NSX Manager provisions the Tier-0 gateway and other networking components needed for organization VDCs such as segment (East-West) and segment (North-South) connecting to Tier-1 gateway. NSX Manager also creates the Tenant Tier-1 gateway.

The CSP can map a vSphere compute cluster to a Provider VDC, then for each tenant the CSP can allocate and reserve resources by using organization-based constructs. Every organization VDC is associated with a resource pool within the compute cluster for resource guarantee.

The separation of network access between organizations is important for multitenancy. vCloud Director integrates with NSX Manager to create isolated layer 2 tenant networks. Tier-1 gateways allow tenants to route traffic between their tenant networks.

Dual-Mode Switching with N-VDS

The tenant external network can carry network traffic to NSX-T Data Center switches that can be either N-VDS in Standard mode, Enhanced mode, or both. This dual-mode N-VDS switching fabric design can be used for accelerated and non-accelerated workloads within the same or separated compute clusters.

The N-VDS Standard switch can be used to carry the 'tenant Tier-1 to provider Tier-0' traffic and the East-West overlay Geneve traffic between VNFCs. VNFCs that require high data path traffic can use the N-VDS Enhanced DPDK fabric. This design can be leveraged by services such as the PGW in a mobile network for broadband connectivity, streaming applications for audio and video distribution, and SBC for IP voice peer-to-peer communications. NSX-T N-VDS Standard and N-VDS Enhanced switches are used to carry the tenant traffic to the provider router and the external physical network respectively. This design allows for multiple tenant Tier-1 gateways to use the same provider Tier-0 Edge Pod.

Service Provider Edge Cluster Configuration

Before establishing the vCloud Director configuration, the CSP must create an Edge Node provider cluster. The NSX Edge Node cluster consists of Tier-0 gateways. The Edge Node cluster can consist of either the VM or bare metal form-factor. The bare metal Edge is installed on a physical server providing higher throughput data rates.

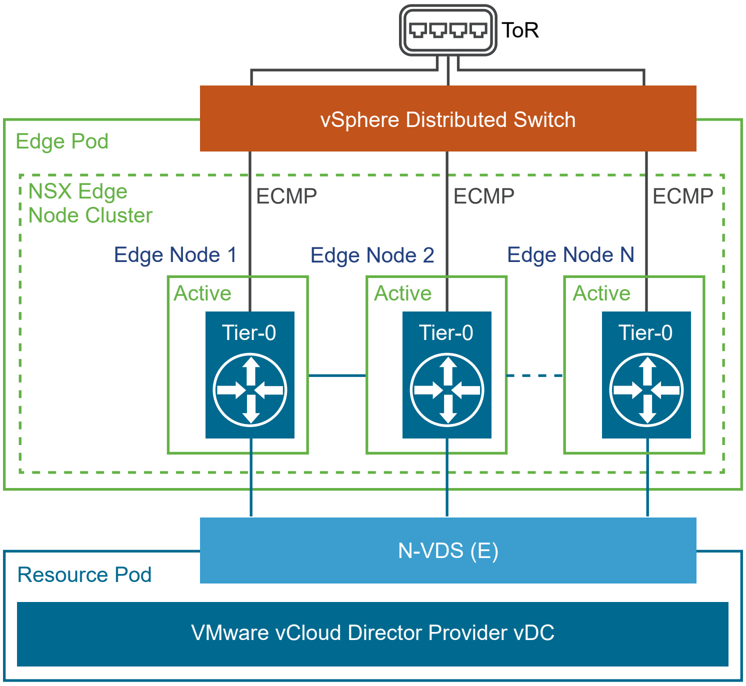
Table 8-1. Edge Node Options

Edge Node Type	Use
VM form-factor	■ Production deployment with centralized services such as NAT, Edge firewall, and load balancer. ■ Workloads that can tolerate acceptable performance degradation loss with virtual edges. ■ Can tolerate lower failure convergence by using BFD (3 seconds). ■ Low-cost options instead of dedicated bare-metal nodes ■ Test proof of concept and trial setups.
Bare metal form-factor	■ Production deployment with centralized services such as NAT, Edge firewall, and load balancer. ■ Higher throughput more than 10 Gbps. ■ Faster failure convergence using BFD (less than 1 second).

Edge Node Active-Active

In an Edge Node Active-Active configuration, Tier-0 gateways are hosted on more than one Edge Nodes at a time to provide high availability. In ECMP mode, the traffic is load balanced between the links to the external physical routers. A maximum of eight Edge Nodes can be configured in ECMP mode to provide scalable throughput that spreads across the Edge Node physical uplinks to the provider network. Stateful services such as NAT and Firewall cannot be used in this mode.

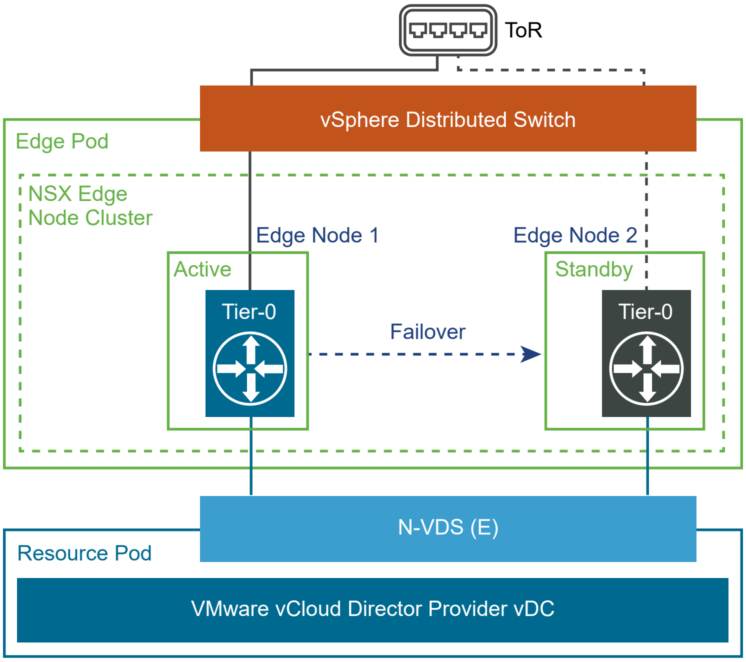
Figure 8-3. Edge Node Active-Active Design



Edge Node Active-Standby

This mode defines the high availability configuration where a Tier-0 gateway is active on a single Edge Node at a time. This mode is required when stateful services such as NAT, Firewall, and load balancer must remain in a constant state of synchronization between the active and standby Tier-0 gateways on the Edge Node pair.

Figure 8-4. Edge Node Active-Standby Design



Dynamic Routing

Tier-0 gateways can be connected to physical routers by using BGP or static routes. If static routes are used, every newly created external network must be added manually to the Tier-0 gateway that peers with the physical routers.

The NSX Edge Node also supports fast failure recovery by using Bidirectional Forwarding Detection (BFD) that is integrated with BGP. The VM form-factor edges support a minimum timeout of one second with three retries, providing a three-second failure detection time between nodes. With bare-metal nodes, the detection or convergence timeout is less than one second.

For more information about NSX-T Data Center, see the NSX-T Reference Design Guide.

Workload Life Cycle Management

Once the compute, storage, and networking environment are set for the tenant, workloads can be onboarded to the organization VDC.

For more information on workload onboarding, see the Workload On-Boarding section.

Workload Acceleration

This version of the vCloud NFV platform supports accelerated data plane. Both control and user plane workloads can take advantage of this capability and achieve higher throughput, reduced latency, and scale in performance.

Scope

VNF components that require high network throughput can achieve high data plane performance by using the advanced switching fabric that is introduced with the N-VDS logical switch of NSX-T Data Center and by using the Edge cluster acceleration in the Three-Pod configuration.

Design Objectives

The vCloud NFV platform includes NSX-T Data Center as the virtualized networking component. NSX-T Data Center leverages key DPDK features, such as Poll Mode Driver, Flow Cache, and optimized packet copy, and provides better performance for both small and large packet sizes that are applicable to data plane intensive VNFs requirements. The NSX-T Data Center networking stack increases the CPU efficiency while preserving the existing functionality of the VMware NFV infrastructure. The accelerated switching and enhanced platform deliver advanced networking architecture that is transparent to VNFs, providers kernel-based security, deterministic resource allocation, and linear scalability.

vSphere introduces support for a high-performance networking mode that is called Enhanced data path mode, which works together with NSX-T Data Center. NSX-T Data Center N-VDS provides logical switching fabric that works in two modes: Standard NSX-managed Virtual Distributed Switch (N-VDS Standard) and Enhanced NSX-managed Virtual Distributed Switch (N-VDS Enhanced). CSPs may use the accelerated N-VDS (E) without losing any of the operational benefits of virtualization such as vMotion and DRS.

Both modes of the N-VDS switch use dedicated physical NICs and can exist on the same host to carry different traffic types.

Table 8-2. NSX-T Data Center Logical Switch Mode Options

Switch Mode	**Use**
N-VDS Enhanced	■ Recommended for control and data plane intensive workloads ■ High transactional control plane VNFs. ■ Connectivity towards external networks can be overlay or VLAN backed.
N-VDS Standard	■ Suitable for control and management plane workloads. ■ VNFs that require the overlay and VLAN-backed connectivity. ■ VNFs that require stateful and stateless Edge services such as load balancer, firewall, and NAT

The section describes the design best practices using N-VDS to accelerate workloads, so that hosts and vCloud NFV platform components are configured to deliver optimal performance.

Acceleration with the N-VDS Switch

This section describes how the VNF-Components (VNF-C), the Data plane VNF-C, the Control plane VNF-C, and the Operations, Administration, Management (OAM) VNF-C can leverage N-VDS capabilities.

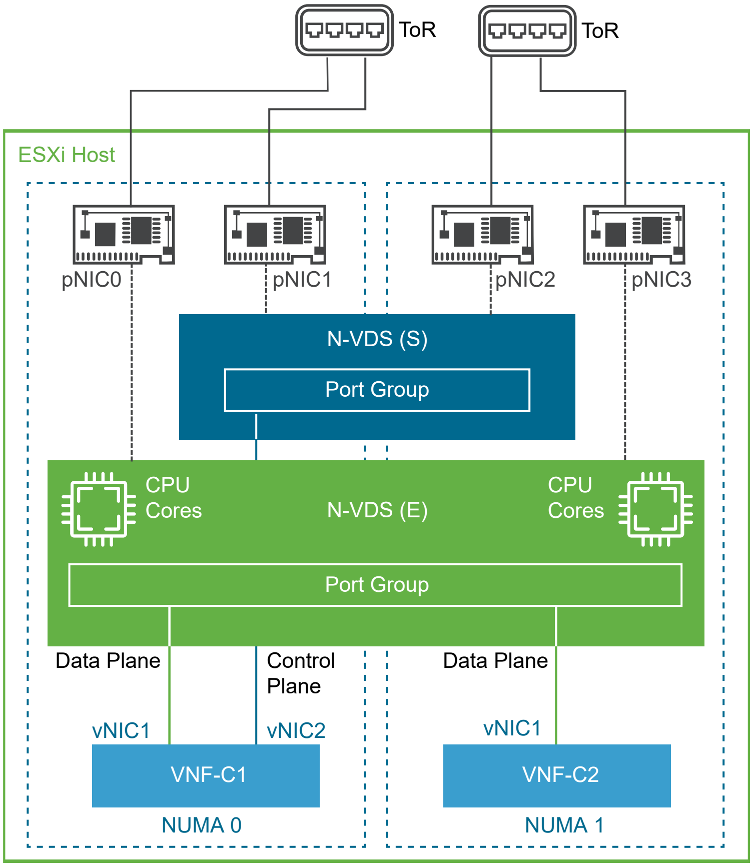
VNF Workload Using N-VDS (S) and N-VDS (E)

The Data plane VNF-C requires a large number of vCPUs to support high data rates for a large number of subscribers. To ensure optimal performance of the Data plane VNF-C, its vCPUs are mapped to CPU cores. Most of the CPU cores on each NUMA are used by the data plane VNF-C. The N-VDS Enhanced mode uses vertical NUMA alignment capabilities to ensure that pNIC, Logical Cores, and vCPUs are aligned in the same NUMA. The Control plane and OAM VMs have a small resource footprint and do not require data plane acceleration.

The Data plane VNF-C is deployed in a cluster that is dedicated to data plane components and occupies most of the CPU resources on a NUMA node in the host. The Data plane VNF-C ensures that the number of vCPUs it uses does not exceed the number of CPU cores that are available on the NUMA. Some CPU cores are left free for the virtualization layer.

After the VNF-Cs are distributed on the NUMA nodes of a host, the components are connected to the virtual networking elements as shown in the following figure:

Figure 8-5. N-VDS Dual Mode Configuration



N-VDS in Standard and Enhanced Data Path modes work side by side on the same host. vNICs that require accelerated performance can be connected directly to N-VDS (E), whereas vNICs that are used for internal communication, OAM, or control plane can use the N-VDS (S). The Control plane and Data plane VNF-Cs communicate in the East-West connectivity.

As VNF plays a vital role in the CSP network, it must also provide a high level of availability. To achieve a resilient network, the NUMA node where the Data plane VNF-C is installed has two physical NICs (preferably in each NUMA node) configured as redundant uplinks from an N-VDS (E). The Data plane VNF-C also benefits from host resiliency by leveraging anti-affinity rules. The N-VDS (E) uplinks are bonded using a teaming policy that increases availability.

The benefits of this approach are:

- *Network traffic separation*: This approach separates the bursty network traffic (management and control plane traffic) from the faster throughput and constant data plane traffic.
- *Ease of designing the network*: The physical NICs used to transport the data plane traffic can use a different path and physical components than the NICs that are used to transport the management and control plane traffic.
- *Benefits of shared resources*: As N-VDS (E) is configured with dedicated CPU resources that are not accessible by other applications, the dedicated CPU resources process the data path traffic. However, the control and management plane traffic is served by N-VDS (S).

For detailed design considerations for accelerated workloads, see the Tuning vCloud NFV for Data Plane Intensive Workloads guide.

Acceleration by Using Bare Metal Edge

NSX Edge Nodes are also available in a bare metal form-factor with improved performance compared to the VM-based Edge Nodes.

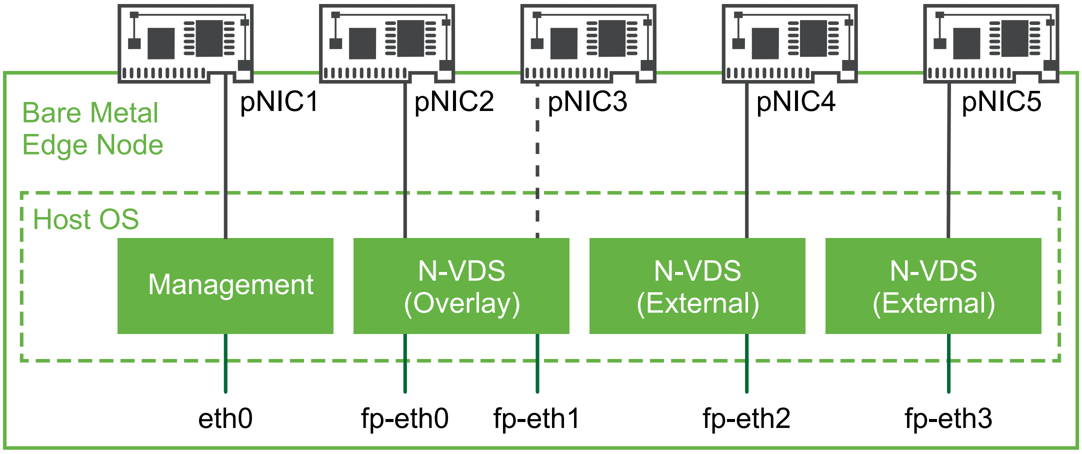
A Three-Pod configuration is required for a bare metal NSX Edge cluster deployment. The Edge cluster acts as an Edge router that connects to the CSP’s physical router.

A bare metal NSX Edge delivers improved performance, higher throughput, sub-second BFD convergence, and faster failover.

Physical Design

When a bare metal NSX Edge Node is installed, a dedicated interface is retained for management. Two physical NICs can be used for the management plane to provide high availability and redundancy.

Figure 8-6. Bare Metal Edge Physical View



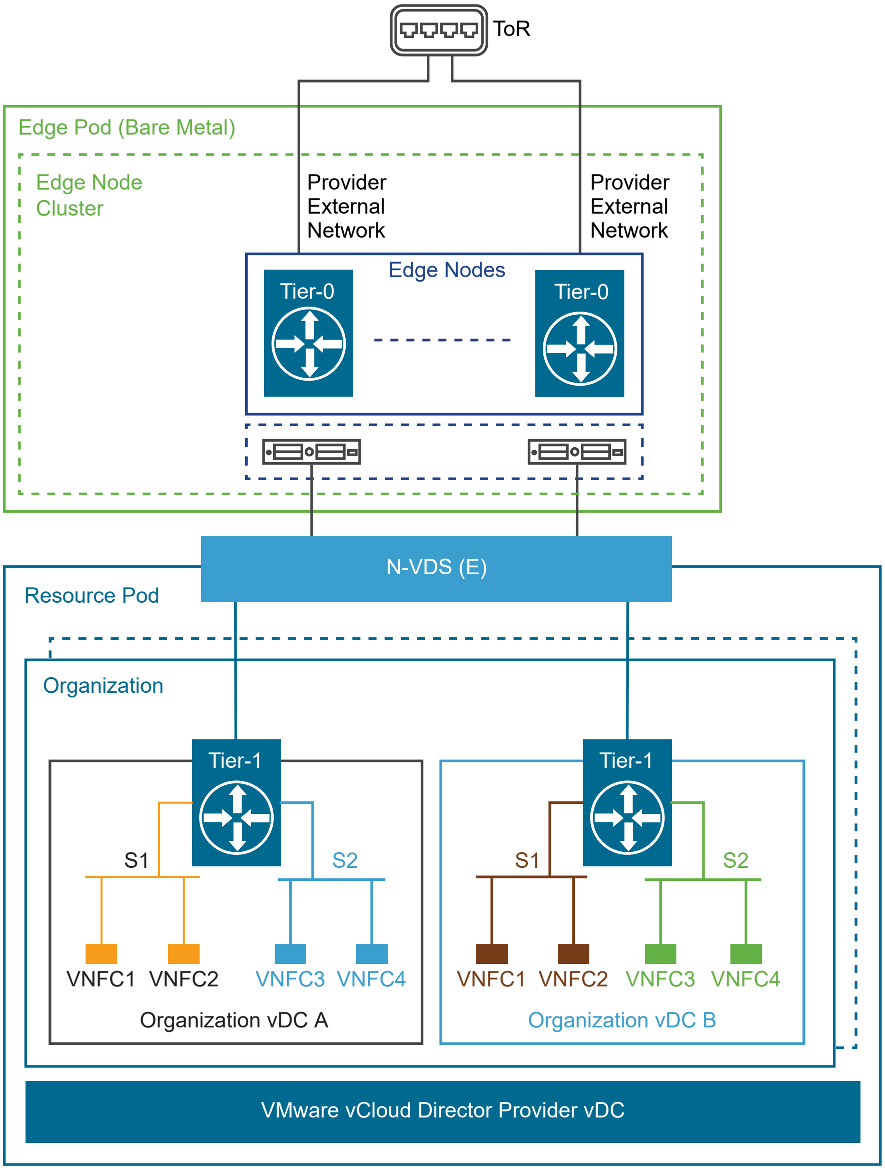
For each physical NIC on the server, an internal interface is created following the **`fp-ethX`** naming scheme. These internal interfaces are assigned to FastPath and are allocated for overlay tunneling traffic or uplink connectivity to top-of-rack (ToR) switches. There is full flexibility in assigning `fp-eth` interfaces to physical NICs for overlay or uplink connectivity. Physical NICs can be assigned to each overlay or external connectivity for network redundancy. Because there are four `fp-eth` interfaces on the bare metal NSX Edge, a maximum of four physical NICs are supported for overlay and uplink traffic in addition to the primary interface for management.

The BFD protocol provides fast failure detection for forwarding paths or forwarding engines, improving the loss of connectivity detection and therefore enabling quick response. The bare metal NSX Edges support BFD for both the interfaces towards the provider router and a BFD-like protocol operating between the NSX Edges and the resource hosts.

Logical Design

In a multitenant environment, vCloud Director for Service Providers is configured to use the N-VDS Enhanced switch with a bare metal NSX Edge for higher performance. Regardless of the workload type, both control and data plane workloads can take advantage of this configuration. Tenant workloads can take advantage of bare metal router features such as dynamic routing, firewall, NAT, and load balancer. With bare metal, there is no virtualization overhead, because it directly connects to physical NIC.

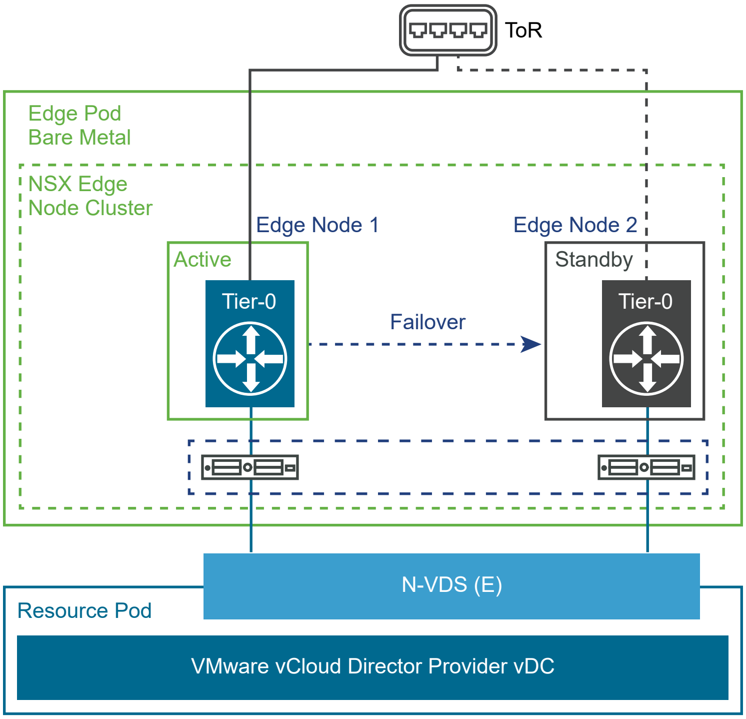
Figure 8-7. Bare Metal Edge Logical View



High Availability Options

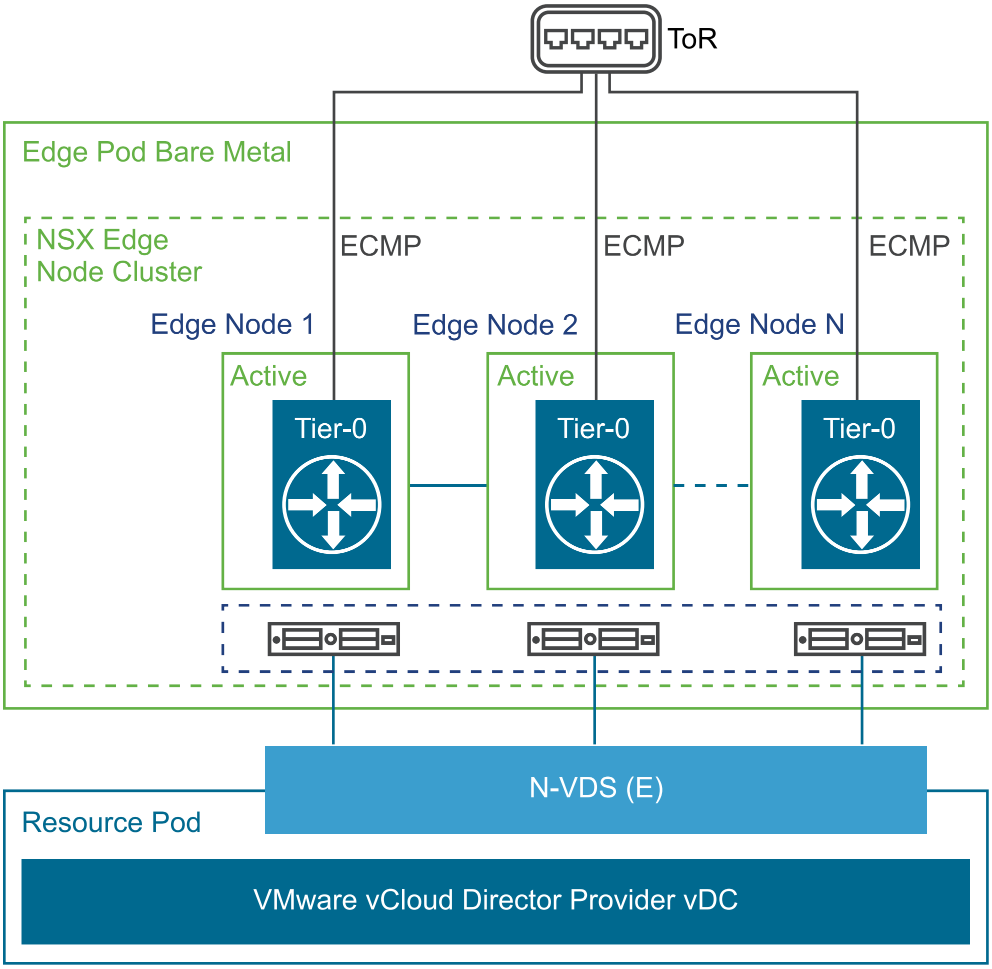
- **Active-Standby mode**: A high availability configuration that requires a minimum of two Edge Nodes hosting the Tier-0 gateways. This mode also enables stateful services such as NAT, firewall, and load balancer to be in a constant state of sync between the active and standby Tier-0 on the Edge Nodes.

Figure 8-8. Bare Metal Edge Active-Standby Mode



- **Active-Active Mode**: Edge Node Active-Active configuration provides high availability mode as Tier-0 gateways are hosted on more than one Edge Node at a time. In ECMP mode, traffic is load balanced between the links to external physical routers. A maximum of eight Edge Nodes can be configured in ECMP mode to provide scalable throughput that spreads across the Edge physical uplinks towards the provider network. Stateful services like NAT and firewall cannot be used in this mode.

Figure 8-9. Bare Metal Edge Active-Active Mode



Acceleration with SR-IOV

SR-IOV is a specification that allows a single Peripheral Component Interconnect Express (PCIe) physical device under a single root port to appear as multiple separate physical devices to the hypervisor or the guest operating system.

SR-IOV uses physical functions (PFs) and virtual functions (VFs) to manage global functions for the SR-IOV devices. PFs are full PCIe functions that can configure and manage the SR-IOV functionality. VFs are lightweight PCIe functions that support data flow but have a restricted set of configuration resources.

The number of virtual functions provided to the hypervisor or the guest operating system depends on the device. SR-IOV enabled PCIe devices require appropriate BIOS and hardware support, and SR-IOV support in the guest operating system driver or hypervisor instance.

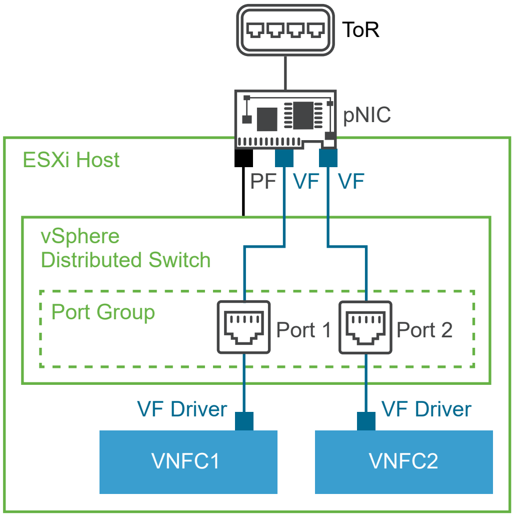
Prepare Hosts and VMs for SR-IOV

In vSphere, a VM can use an SR-IOV virtual function for networking. The VM and the physical adapter exchange data directly without using the VMkernel stack as an intermediary. Bypassing the VMkernel for networking reduces the latency and improves the CPU efficiency for a higher data transfer performance.

vSphere supports SR-IOV in an environment with a specific configuration only. Find a detailed support specification on the SR-IOV Support page.

In the topology below, the vSphere SR-IOV support relies on the interaction between the VFs and the PF of the physical NIC port for higher performance. VM network adapters directly communicate with the VFs that SR-IOV provides to transfer data. However, the ability to configure the VFs depends on the active policies for the vSphere Distributed Switch port group ports (VLAN IDs) on which the VMs reside. The VM handles incoming and outgoing external traffic through its virtual ports that reside on the host. The virtual ports are backed by physical NICs on the host. VLAN ID tags are inserted by each SR-IOV virtual function. For more information about configuring SR-IOV, see Configure a Virtual Machine to Use SR-IOV.

Figure 8-10. SR-IOV Virtual Function Configuration

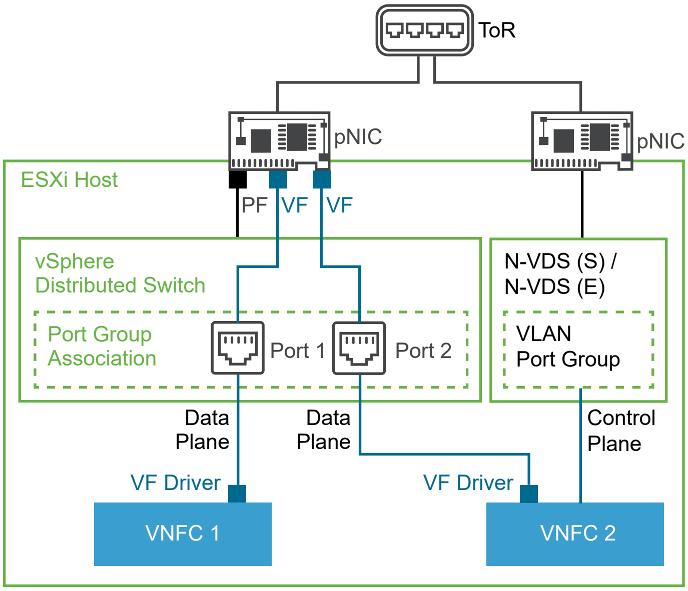


SR-IOV can be used for data-intensive traffic, but it cannot use virtualization benefits such as vMotion, DRS and so on. Hence, VNFs employing SR-IOV become static hosts. A special host aggregate can be configured for such workloads. The NSX-T Data Center fabric can be used to plumb interfaces into an N-VDS Standard switch and for the VLAN and overlay connectivities.

SR-IOV Configuration by Using vCloud Director for Service Providers

vSphere Distributed Switch and N-VDS Standard can be deployed on the same host with dedicated physical NICs to each switch. VNFs can have a VIF connected to both the vSphere Distributed Switch port group ports for North-South connectivity and N-VDS Standard switches for East-West connectivity.

Figure 8-11. SR-IOV Logical View



Multi-Tenancy

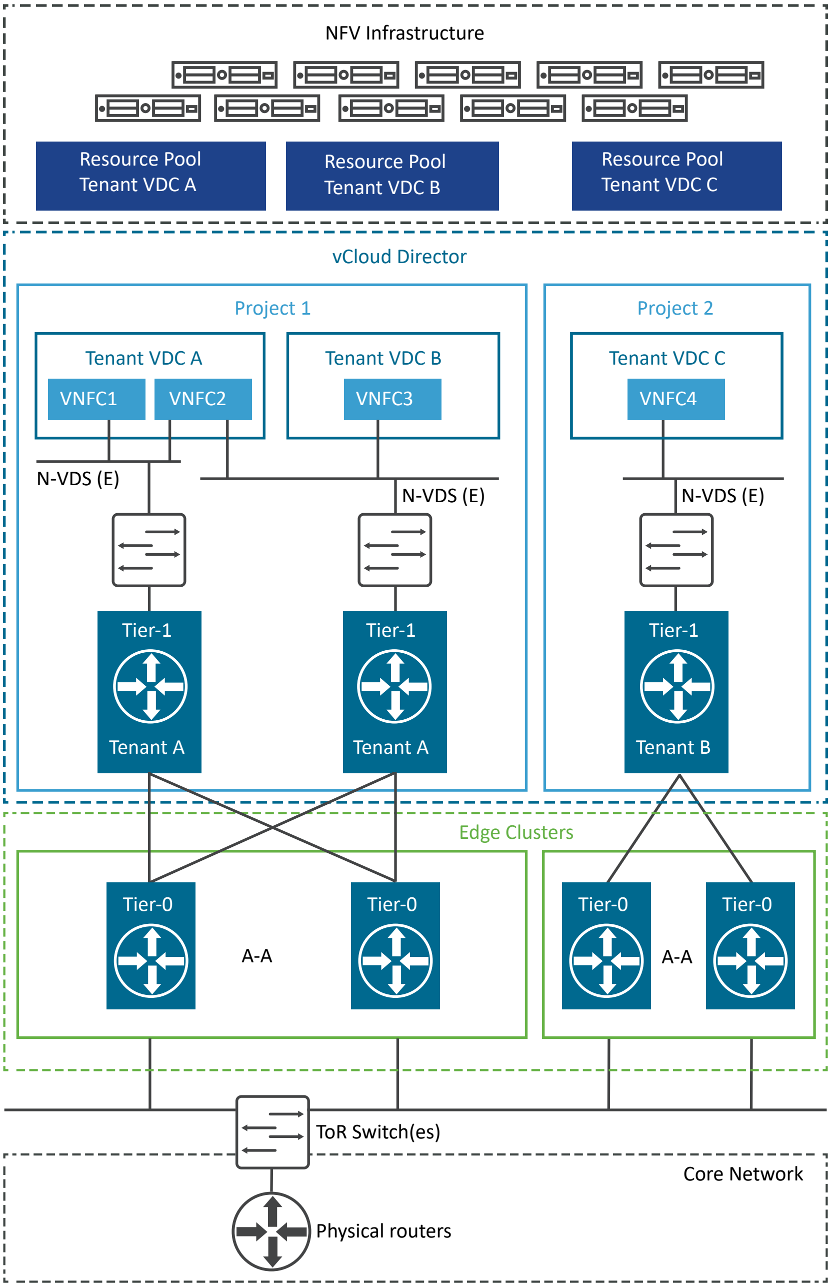
The CSP can facilitate the NFV transformation on a shared resource infrastructure environment with multitenant consumption models. The design of those tenant environments is outlined in this section.

Scope

Multitenancy defines the isolation of resources and networks to deliver applications with quality. Because multiple tenants share the resource infrastructure, secure multitenancy can be enabled by using VMware vCloud Director in a single cloud island and across distributed clouds. In addition to the built-in workload and resource optimization capabilities, predictive optimization can be enabled with analytics by using features such as vSphere DRS.

CSPs can converge their resource infrastructures across their IT and Network clouds enabling a multitenancy IaaS realization over it. Consumption models can serve both internal and external tenants over the common shared infrastructure to deploy and operate their respective workloads and services. Network, compute, and storage isolation with QoS are the design objective discussed in this section.

Figure 8-12. Multi-Tenancy with vCloud Director



Design Objectives

A unit of tenancy is called an Organization VDC within the scope of an Organization. It is defined as a composition of dedicated compute, storage, and network resources as well as workloads. The tenant is associated with a set of operational policies and SLAs. The tenant can be bound to a single Organization VDC or can be composed of many Organization VDCs. Services such as HSS and DNS are examples of shared tenancy.

The design objectives include considerations for the compute and network resource isolation and automation.

Management Plane

The management plane functions reside in the Management Pod. They are responsible for the orchestration of resources and operations. The management plane functions are local to each cloud instance providing infrastructure management, network management, and operations management capabilities.

Resource isolation for compute and networking design are enabled together with vCenter Server, NSX Manager, and VMware vCloud Director. Irrespective of the pod deployment configuration, the VMware vCloud NFV platform provides the abstraction layers for multi-tenancy. vCenter Server provides the infrastructure for fine-grained allocation and partitioning of compute and storage resources, whereas NSX-T Data Center creates the network virtualization layer.

The concept of tenancy also introduces shared administrative ownership. A cloud provider, that is the CSP admin, can create a resource pool allocation and overlay networking for a tenant who in turn would consume these resources based on the workload requirements. In vCloud Director, multiple tenants can be defined with assigned RBAC privileges to manage resources and VNF onboarding.

Compute Isolation

The allocation of compute and storage resources ensures that there is an optimal footprint available to each tenant that is used to deploy workloads, with room for expansion to meet future demand.

Organizations provide a secure multitenant environment to deploy VNFs. Compute resources are defined as resource pools when an Organization VDC is created. The resource pool is an allocation of memory and CPU from the available shared infrastructure, assignable to an Organization VDC. More resources can be added to a pool as capacity needs to grow. The Organization VDC can also stretch across multiple hosts residing in different physical racks.

Network Isolation

The advanced networking model of NSX-T Data Center provides a fully isolated and secure traffic path across workloads in tenant switch and routing fabric. Advanced security policies and rules can be applied at the VM boundary to further control unwarranted traffic. Also, for better traffic management, QoS switching profile can be used to provide high-quality and dedicated network performance for preferred traffic that requires high bandwidth using Class of Service (CoS) and Differentiated Services Code Point (DSCP) values for tenants.

NSX-T Data Center introduces a two-tiered routing architecture that enables the management of networks at the provider (Tier-0) and tenant (Tier-1) tiers. The provider routing tier is attached to the physical network for North-South traffic, while the tenant routing can connect to the provider Tier-0 and manage East-West communications. The Tier-0 provides traffic termination to the cloud physical gateways and existing CSP underlay networks for inter-cloud traffic communication.

Each Organization VDC has a single Tier-1 distributed router (DR) that provides the intra-tenant routing capabilities. It can be also enabled for stateful services such as firewall and NAT. VMs belonging to a Tenant can be plumbed to multiple logical interfaces for layer 2 and layer 3 connectivity.

Resource Allocation

To avoid contention and starvation, compute, storage, and network isolation should be applied consistently to the workloads.

The CSP admin can allocate and reserve resources for tenants by using Organization VDC. Every Organization VDC is associated with a resource pool across the Resource Pods. The resource settings of the resource pool are managed from vCloud Director. This ensures that every Organization VDC allocates the resources to which it is entitled, without exceeding the infrastructure resource limits, such as CPU clock cycles, total memory, network bandwidth, and storage.

vCloud Director supports an allocation model that determines how and when the allocated Provider Virtual Data Center (VDC) compute and memory resources are committed to the organization VDC. Every allocation model provides different levels of performance control and management. For the suggested use of the allocation model, see the vCloud Director Administrator guide.

1. **Resource Allocation Reservation**: Defines the minimum guarantee. This parameter ensures a minimum guarantee to each VM when it is launched.

2. **Resource Allocation Limit**: Defines the upper boundary. Use this parameter with caution in a production environment, because it restricts the VM from bursting utilization beyond the configured boundaries.

3. **Resource Allocation Shares**: Defines the distribution of resources under contention. Shares can be used to prioritize certain workloads over others in case of contention. If the resources are over-provisioned across VMs and there is resource contention, the VM with higher shares gets the proportional resource assignment.

vSphere resource distribution settings at the virtual machine (VM) or resource pool level based on the organization VDC allocation model.

For the control plane workload functions, a higher-order elasticity is acceptable and memory can be reserved based on the workload requirement. For the data plane intensive workloads, both CPU and memory should be fully reserved. Storage IO and network throughput reservations need to be determined based on the VNF needs.

Automation

To meet the operational policies and SLAs for workloads, closed-loop automation is necessary across the shared cloud infrastructure environment.

The vCloud NFV leverages vSphere DRS to optimize the initial and runtime placement of workloads to ensure health and performance to the cloud infrastructure. Organization VDCs and workloads are monitored to ensure that the resources are being tuned and balanced dynamically.

Capacity and demand planning, SLA violations, performance degradations, and issue isolation capabilities can be augmented with the analytics-enabled reference architecture. The analytics-enabled architecture provisions a workflow automation framework to provide closed-loop integrations with NFVO and VNFM for just-in-time optimizations.

Distributed Clouds

The NFV transformation in the private cloud spans the distributed topologies. For low latency and compute proximity, user plane functions such as Evolved Packet Core (EPC) gateways and IMS media move closer to the edge clouds. Control plane functions can be still centralized.

Scope

Distributed clouds in this reference architecture are examined from a management and operations perspective.

Design Objectives

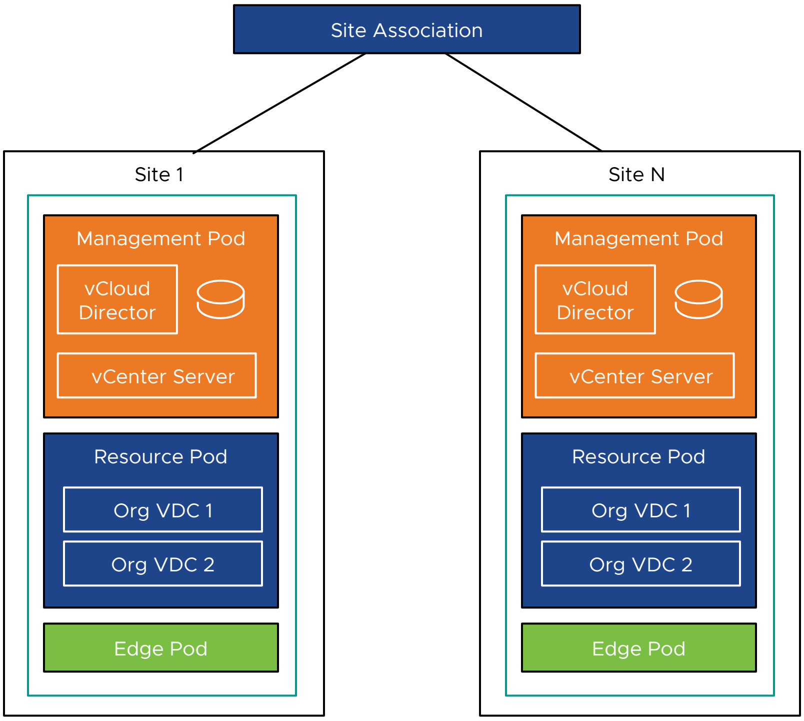
Multi-cloud sites can be used for distributing workload functions as well as user plane and control plane. Multi-cloud sites can be used to provide geographic redundancy for workloads running across the different sites.

Site Design

In a distributed cloud topology, VNFs are stitched across different cloud islands to deliver services. Each site can also have multiple instances of the NFV environment for redundancy and scale.

Each site is a standalone island with its own instance of the Management, Resource, and Edge Pods. Multi-site designs are typically used for load distribution and for high availability in case of catastrophic site failures.

Figure 8-13. Distributed Clouds



Cloud Management Plane

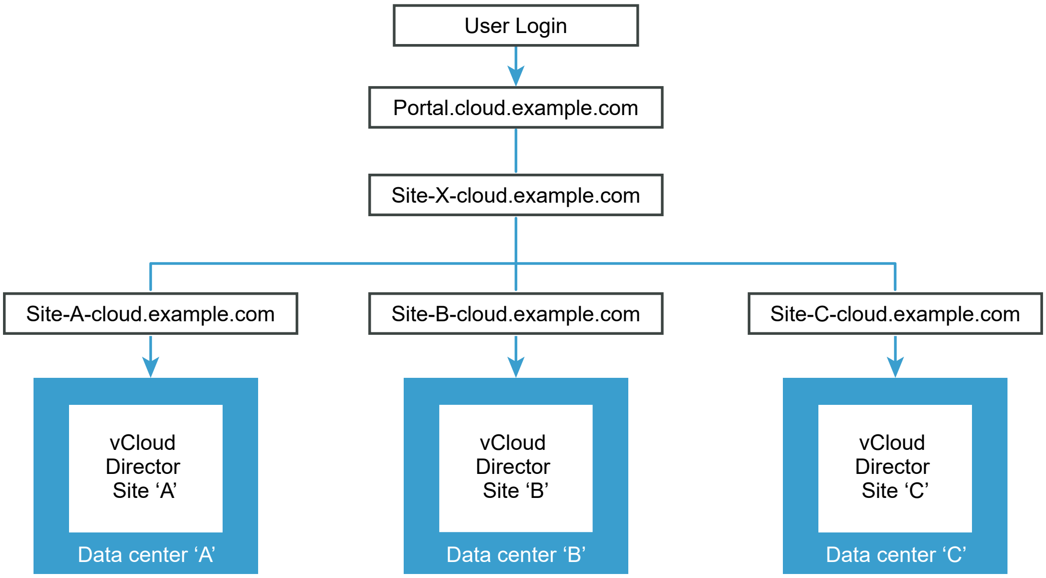
The management plane functions reside in the Management Pod. The management plane functions are local to each site providing the virtual infrastructure and host management, network management, and operations management capabilities.

Resource Orchestration

Each cloud island should have a dedicated vCloud Director instance for virtual infrastructure management of that island. Each site should also have dedicated resource and management vCenter Server instances. Optionally, this design choice could be centralized as latency and site distances should be considered.

The vCloud Director site and Organization credential association feature federates identity management across various distributed vCloud Director instances and their Organizations within each of the distributed sites. A tenant user can have an Organization VDC created across multiple data centers. A tenant administrator can then log in to any of the associated vCloud Director instances to manage their tenancy from a single pane view.

Figure 8-14. Global Site Access Conceptual Overview



For more information, see Architecting Multisite VMware vCloud Director.

Operations Management

Operations management functions by using the VMware vRealize suite can be deployed in a single centralized site to collect and process analytics. The latency between sites should be considered in the design.

Workload On-Boarding

The VNF onboarding process is typically a collaborative effort between the VNF vendor and the CSP. A prescriptive set of steps needs to be followed to onboard and deploy a VNF. The VMware Ready for NFV program is a good vehicle to onboard and certify the VNFs on the vCloud NFV platform to ensure smooth deployment in the CSP's environment.

Scope

The VNF life cycle involves provisioning of the cloud infrastructure, VNF packaging, resource assignment and configuration, deployment, and its placement.

Design Objectives

The workload onboarding process is aimed at VNFs that are deployed as native VM applications. This section covers the process of packaging, instantiation, placement.

vCloud Director Conceptual Design

The conceptual design provides a high-level view of the roles, areas of responsibility, and tenant flow that are required to upload an image, onboard, and deploy it.

The VNF onboarding process is typically a collaborative effort between the VNF vendor and the CSP. Before a VNF is onboarded, the VNF vendor must provide the CSP with all the prerequisites for the successful onboarding of the VNF. This includes information such as the VNF format, number of the required networks, East-West and North-South network connectivity, routing policy, security policy, IP ranges, and performance requirements.

VNF Format and Packaging

vCloud Director supports importing VNFs as standard OVF/OVA packages. As a best practice, the VNF vendor must author the package in a separate environment identical to the target vCloud NFV environment to match its features and capabilities.

VMware Ready for NFV is a program where VMware and VNF vendors collaborate to ensure that the VNF is interoperable with vCloud NFV. Based on experience, VMware provides best practices to help VNF vendors in preparing their VNFs for consumption by vCloud Director.

CSPs expect their VNF supplier to deliver a package that is ready for consumption by vCloud Director. The package must support:

1. **Autonomous VNF Life Cycle Operations**: CSPs must be able to install, upgrade, and manage the VNF themselves.
2. **Compact**: All components that are required to operate the VNF are packaged together. VNFCs are clearly defined, use unique and specific names, and are included in the package.
3. **Authorized**: The package provided to the CSP must be verifiable. For example, an md5 hash could be used to verify that the package provided by the vendor is the same package that is installed by the CSP.
4. **Holistic**: The packaged VNF must include all configuration that is required to deploy a healthy VNF. For example, if the VNF is data plane intensive, all performance-related configuration such as CPU and NUMA affinity, CPU and memory reservations must be included in the package.

VNF Onboarding by Using VMware vCloud Director

Onboarding a VNF in vCloud NFV includes the following steps:

1. Prepare the VNF for consumption by vCloud Director.
2. Import the VNF images to the vCloud Director catalog.
3. Ensure that the environment is configured correctly for the VNF deployment.
4. Deploy the VNF in the vCloud NFV environment.

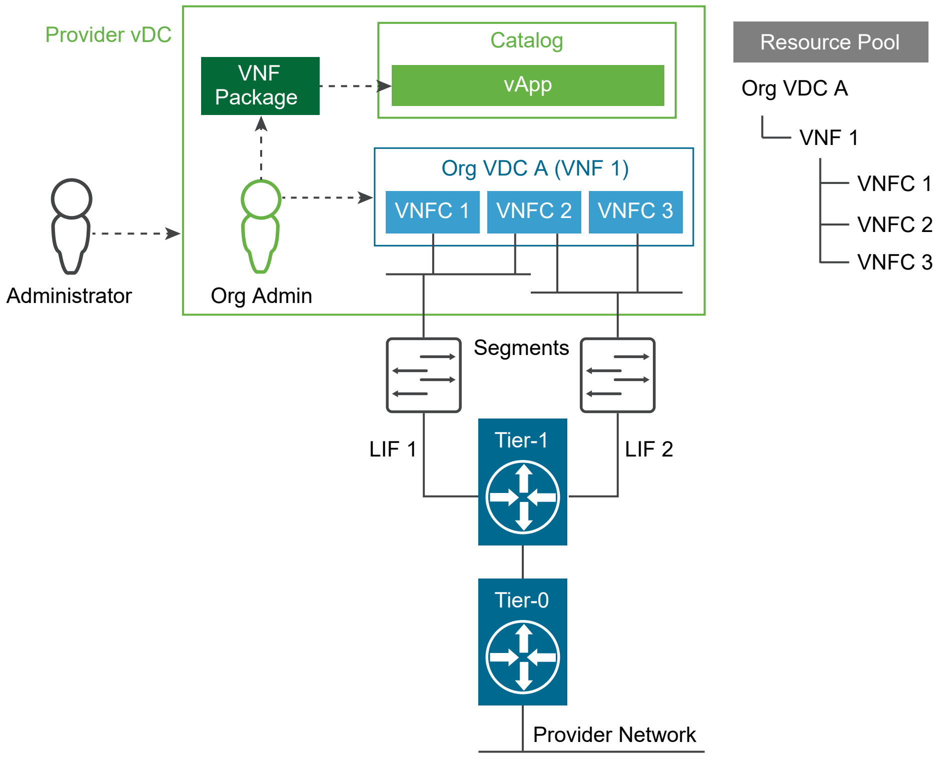
The process can be divided into CSP operations and tenant operations.

Before a VNF is onboarded, the CSP administrator must collect the prerequisite information from the VNF supplier. The prerequisites include configuration information such as the number of required networks, IP ranges, North-South network connectivity, and so on.

vCloud Director uses the concept of catalogs for storing content. Catalogs are containers for VNF templates and media images. CSPs can create global catalogs with golden VNF templates that can be shared to one or more tenants, while tenants can retain their own private catalog of VNF templates. The OVF/OVA package of a VNF is directly imported to the tenant catalog. In addition to VNF templates, the catalog can contain VNFC templates, used to scale deployed VNFs.

Tenant administrators deploy VNFs from the available templates in the self-service catalog. Tenants also provision East-West connectivity by using Organization VDC networks that they import from NSX-T. VNF North-South connectivity is established by connecting Organization VDC networks to the external network through an NSX-T Edge Router.

Figure 8-15. VNF Onboarding Conceptual Design



Resource Allocation

This process is typically collaborative between the VNF vendor and the CSP. Based on lab testing, the VNF vendor provides guidance on the amount of virtual resources needed to support a specific size or scale of telco service. For example, a virtual Evolved Packet Core (vEPC) is likely to specify that to serve a certain number of subscribers with active features such as Deep Packet Inspection (DPI) and Quality of Service (QoS), a specific number of vCPUs, memory, network bandwidth, and storage is required. The CSP accounts for near-future scale requirements, using the resource allocation mechanisms provided by vCloud Director such as resource allocation models and storage policies

VNF Networking

Based on specific VNF networking requirements, a CSP administrator can use NSX Manager to provision East-West connectivity, security groups, firewalls, micro-segmentation, NAT, and LBaaS for a tenant. Tenants import their respective networks created by the CSP administrator, to an Organization VDC. VNF North-South connectivity is established by connecting tenant networks to external networks through NSX-T Data Center routers that are deployed in Edge Nodes. External networks are created by CSP administrators and backed by physical networks.

Tenant networks are accessible by all Organization VDCs within the Organization. Therefore, the implementation of East-West connectivity between VNFCs in the same Organization VDC, and the connectivity between VNFs in two different Organization VDCs belonging to the same Organization, is identical. Tenant networks are implemented as logical switches within the Organization. The North-South network is a tenant network that is connected to the telco network through an N-VDS Enhanced switch for data-intensive workloads or by using N-VDS Standard through an NSX Edge Cluster.

vCloud Director exposes a rich set of REST API calls to enable automation. By using these API calls, the upstream VNFM and NFVO can automate all aspects of the VNF life cycle. Examples include the VNF deployment from a catalog, tenant network consumption, power operations, and VNF decommissioning.

Host Affinity and Anti-Affinity Policy

A vCloud Director system administrator can create groups of VMs in a resource pool, then use VM-Host affinity rules to specify whether members of a VM group should be deployed on members of a vSphere host DRS Group. vCloud Director VM-Host affinity rules provide vCloud Director system administrators with a way to specify how vSphere DRS should place VMs on hosts in a resource pool. VM-Host affinity rules can be useful when host-based licensing requires VMs that are running certain applications to be placed on hosts that are licensed to run those applications. They can also be useful when VMs with workload-specific configurations require placement on hosts that have certain characteristics such as acceleration.

DRS Host Groups for Placement

A host group is a vSphere host DRS group. The vSphere administrator must create host DRS groups in a resource pool mapped to a Provider VDC before they can be used in vCloud Director VM Host affinity rules. This helps in placing the workload on a particular host cluster at power-on once they are onboarded.

Availability and Disaster Recovery

Business continuity is supported by the vCloud NFV platform across its management, control, and data plane components. This section discusses the available designs and recovery considerations.

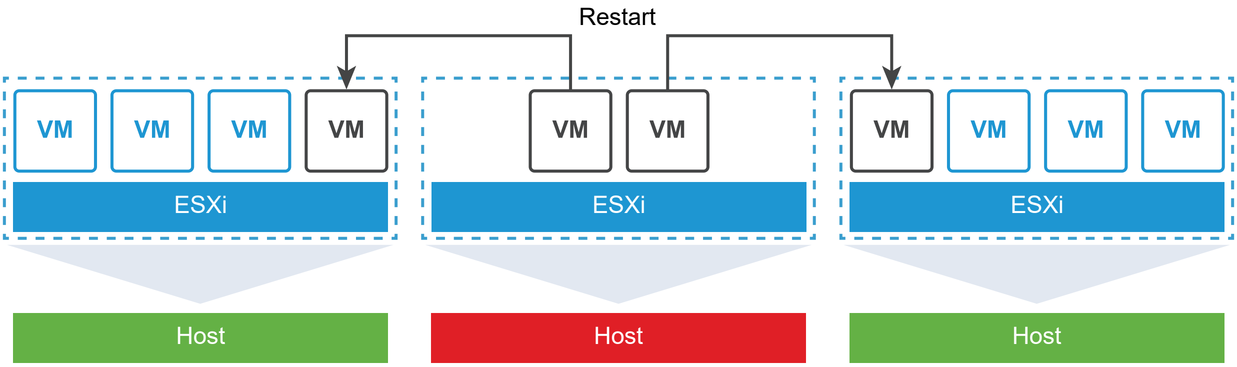
Availability

All vCloud NFV platform components implement a high availability design by default. Also, VNFs can take advantage of platform capabilities to extend their availability in compute, storage, and networking.

vSphere High Availability

Redundancy with vSphere uses VM-level replicas along with the VNF high availability architecture. vSphere HA can restart a VM on another host in the event of host failure thus providing redundancy for a VNF and VNFC pair. vSphere HA can be fully automated without the need for manual intervention for failure and recovery.

Figure 8-16. vSphere High Availability



VMware NSX-T Data Center Availability

NSX-T Data Center by default provides high availability to VNFCs in the overlay network. NIC Teaming and protocols such as Equal-Cost Multipath (ECMP), Graceful Restart, and Link Aggregation Group (LAG) provide redundant connectivity. The NSX-T architecture also separates the management, control, and data plane traffic to further optimize service availability.

VMware vSAN

vSAN is fully integrated into vSphere and provides policy-driven fault handling with platform awareness such as chassis and rack affinity to store object replicas. The virtual storage is integrated with vRealize Operations so that in case of failure, the failed VM can be cloned and spun up automatically. Storage vMotion can be used to perform the live migration of Virtual Machine Disk Files (VMDK) within and across storage arrays by maintaining continuous service availability and complete transaction integrity at the same time.

Disaster Recovery

In the event of a failure, the site is recovered by executing an automated recovery plan. Data replication across protected and failover zones is necessary to recover the state of the site.

vSphere Replication

vSphere Replication replicates virtual machine data between data center objects within a single site or across sites. vSphere Replication fully supports vSAN. It is deployed as a virtual appliance in the Management Pod to provide a Recovery Point Objective (RPO) of five minutes to 24 hours.

When executing a disaster recovery plan, RPO and Recovery Time Objective (RTO) are the most important aspects that must be considered. RPO is the duration of the acceptable data loss and it is fulfilled by the replication technology. RTO is a target duration with an attached SLA, during which the business process must be restored. It includes the time for the recovery and service readiness, in a state of normal business operation.

vSphere Replication provides the ability to set the RPO, however RTO is application dependent.

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 sites is required, with enough bandwidth capacity to replicate the management components by using vSphere Replication. Each site must have an instance of vCenter Server that governs the Management Pod and its ESXi hosts, and a Site Recovery Manager server and 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.

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 the protected and recovery sites:

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

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 that are 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.

VNF Recovery Considerations

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

NSX Data Center for vSphere Coexistence with NSX-T Data Center

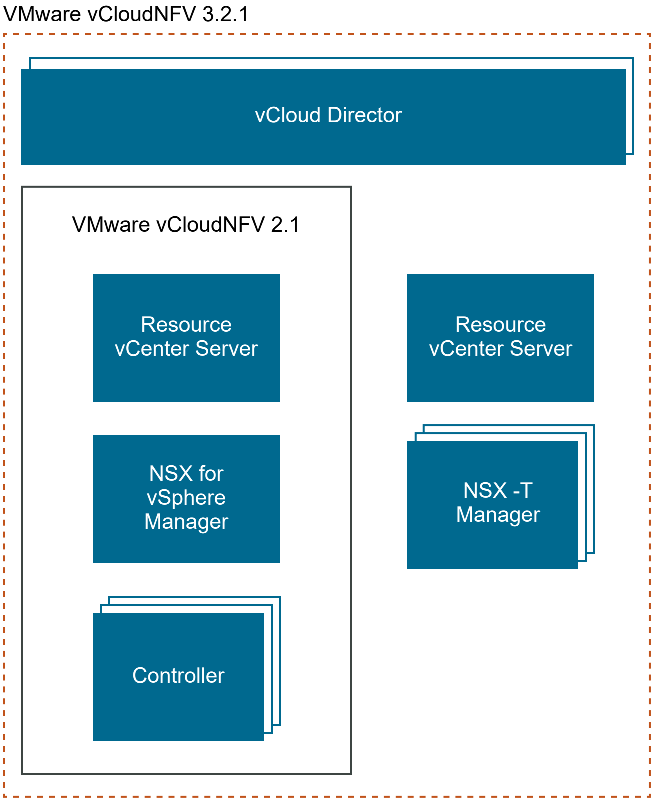
As NSX evolves from NSX Data Center for vSphere to its successor NSX-T Data Center, CSPs can deploy the vCloud NFV platform with two software-defined networking stacks simultaneously. CSPs can deploy NSX Data Center for vSphere along with NSX-T Data Center as part of vCloud NFV 3.2.1 or as a complement to an existing vCloud NFV 2.1 deployment.

NSX Data Center for vSphere Interoperating with NSX-T Data Center in Existing vCloud NFV Deployments

This deployment model can serve CSPs that have existing vCloud NFV 2.1 deployments and cannot upgrade their entire platform to vCloud NFV 3.2.1. Such CSPs can complement their vCloud NFV deployments with advanced networking capabilities available through NSX-T Data Center and features described in this reference architecture.

In this scenario, the CSP should upgrade only some of the platform components to enable the platform to support both networking stacks, the existing NSX Data Center for vSphere based networking as well as the enhanced NSX-T Data Center networking. The objective is to have a common vCloud Director instance in the existing and new vCloud NFV 3.2.1 deployment stack (NSX-T based).

Figure 8-17. NSX-T and NSX-V Coexistence in Brownfield



This deployment is also referred to as a brown field deployment model.

Table 8-3. Brown Field vCloud NFV 3.2.1 Deployment

Building Block	Design Objective
Management Pod	■ Dedicated vCenter Server instances for each NSX Data Center for vSphere and NSX-T Data Center stacks. ■ Separate management and control planes for NSX Data Center for vSphere and NSX-T Data Center. ■ Single vCloud Director instance for both NSX Data Center for vSphere and NSX-T Data Center networking and workload management.
Resource Pod	■ Dedicated Resource Pods for NSX Data Center for vSphere and NSX Data Center. ■ Disparate vSphere version for each stack. ■ Dedicated Provider VDC for the vSphere clusters of NSX Data Center for vSphere and NSX-T Data Center.
Edge Pod	■ Disparate NSX-V and NSX-T edge pods for North-South communications.

The CSP can start from existing vCloud NFV 2.1 components by upgrading only the components that are required to support the deployment of NSX-T Data Center alongside NSX Data Center for vSphere. Workloads continue to run on the vCloud NFV platform with minimal maintenance window. This is because only a subset of the components is upgraded as opposed to the entire suite. For the vCloud NFV 2.x deployment, see the respective vCloud NFV 2.x Reference Architecture guide.

Telco Edge for vCloud NFV

Telecommunication operators need a dis-aggregated and distributed virtual infrastructure that allows them to selectively place workloads closer to the subscriber, especially with the advent of 5G networks. These distributed mini or micro data centers are broadly termed Telco Edge sites. The geography of a country coupled with its population density could lead a typical Telco operator to deploy thousands of these Edge sites to cater to multiple use cases.

With high data throughput, low latency, and a large number of devices that 5G network needs to support, Telecom operators can introduce new services to the marketplace. The ability to deploy new services quickly and at scale is a key requirement to monetize this market opportunity effectively. To do this, Telecom operators need their virtual infrastructure to be distributed, scalable, and manageable.

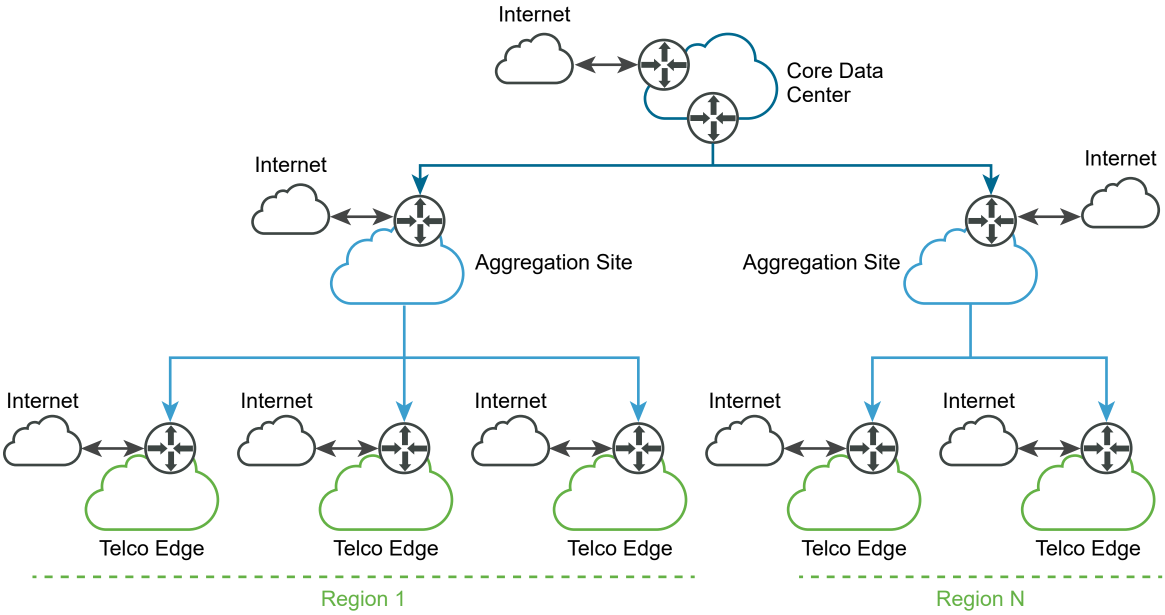
This section describes how vCloud NFV, with VMware vCloud Director as the VIM, can be deployed in a dis-aggregated and distributed fashion to meet the growing needs of the Telco Edge use case.

Conceptual Architecture

The Telco Edge is a collection of distributed virtual infrastructure deployments that can be used to run Telco workloads (such as VNFs) and other user applications. The Edge Reference Architecture provides the flexibility depending on the nature of the workloads and applications, the position in the network, the size of the deployment, and the software-defined infrastructure model.

A classic 3-layer conceptual architecture for the Telco Edge deployment is a hierarchical model that consists of a group of Telco Edges that form an Aggregation site and a group of Aggregation sites that form a Core data center.

Figure 8-18. Telco Edge Conceptual Architecture



Telco Edge

Customer edges connect at the “last mile” to cell towers or wireline aggregation points that are the Telco Edge Sites. The number of servers and the types of Telco network functions, such as virtualized RAN (vRAN) and applications (AR/VR) are constrained by deployment locations, power, cooling, or network factors. Internet “breakout” allows the applications deployed on the Telco Edge to access the Internet directly without having the traffic backhauled to the Aggregation Site or regional/core location.

Aggregation Site

The next level of the hierarchy is the Aggregation site, which aggregates traffic from multiple Telco Edges and generally has fewer constraints related to capacity. The aggregation site consists of a large number of servers with a higher capacity to run applications. A repurposed central office in the wireline scenario is an example of an Aggregation site deployment location. An Aggregation site can contain multiple racks in a typical deployment. Latencies from the user equipment to the Aggregation site are in the range of 5-10 milliseconds but can vary depending on the deployment. An Internet breakout is also present in this deployment.

The aggregation functionality can involve a separate management plane installation to manage Telco edges for scalability requirements. In some cases, the Aggregation site is only used to run applications, while the management functionality for both the Telco Edges and Aggregation site is instantiated in a Core data center.

Core Data Center

The final level of the hierarchy is the Core Data Center that acts as a centralized location for aggregating all control and management plane components for a given region. This deployment is similar to the current centralized model used in Telco networks where the core runs VNF functions and other applications. In the 5G world, the 5G control plane (CP) functions are run in the Core data center and the user plane (UP) functions are run in the edges.

The maximum number of Edge sites in a specific group is governed by the maximum scale supported by the respective management components. In addition to functioning as a traffic aggregation point, a higher layer site also functions as a management layer for the lower tiers, as appropriate. Therefore, the management component for all the Telco Edges aggregating into an Aggregate site is usually located in the specific Aggregate site, in this case, the Core Data Center.

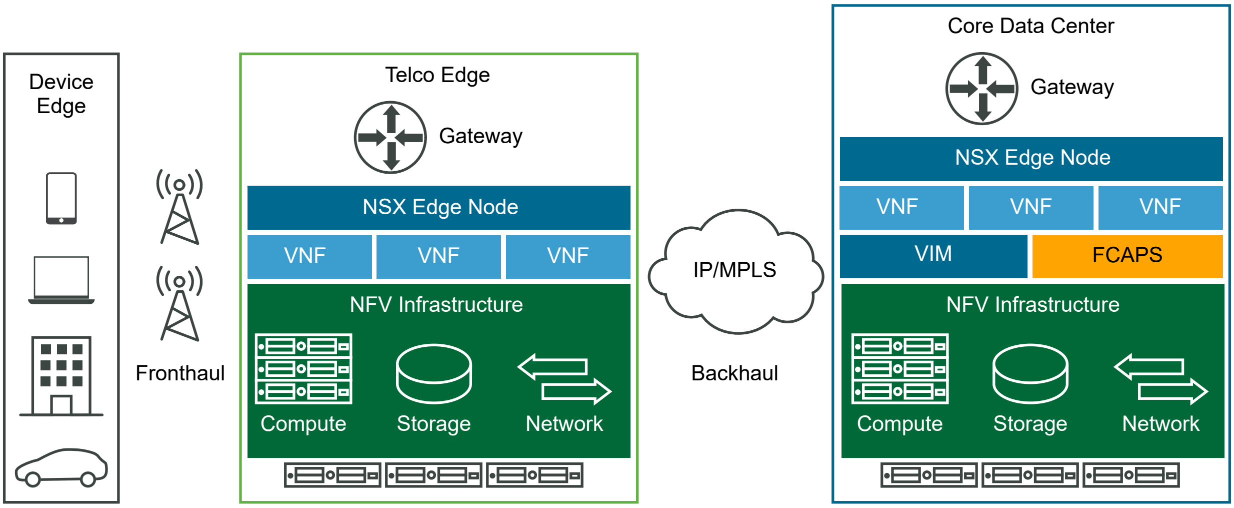
Reference Model

To optimize the Edge deployment to run workloads without significant management overhead, the management plane functionality for the Edges is usually implemented in a site remote to the Edge such as Core data center. While this imposes some constraints on networking (including link availability and end to end latency), this model of remote management is useful for constrained environments such as Telco Edges.

The benefit of this model is the ease and simplicity with which the entire Telco infrastructure can be administered. Instead of connecting to each site to configure and control the resources at that site, users can access the centralized management at the Core data center, which can give them access to all the Edge sites under its purview.

In some deployments, the number of Telco Edges can be very large, running into several thousands. In such cases, a hierarchical management approach is optimal. Irrespective of the model of aggregation, an architectural principle that an edge site is managed from a central site is used. This reference architecture considers a model where a group of sites, both Telco Edges and Aggregation sites, are managed from a single management instance at a core site. This is depicted in the following figure:

Figure 8-19. Telco Edge Reference Model



The Telco Edge and Aggregation sites are collapsed into a single Telco Edge that is managed from a Core data center. We use the term "region" to indicate a group of Edge sites. The preceding figure depicts the components of a Telco edge representing the NFV infrastructure that includes compute, networking, and storage.

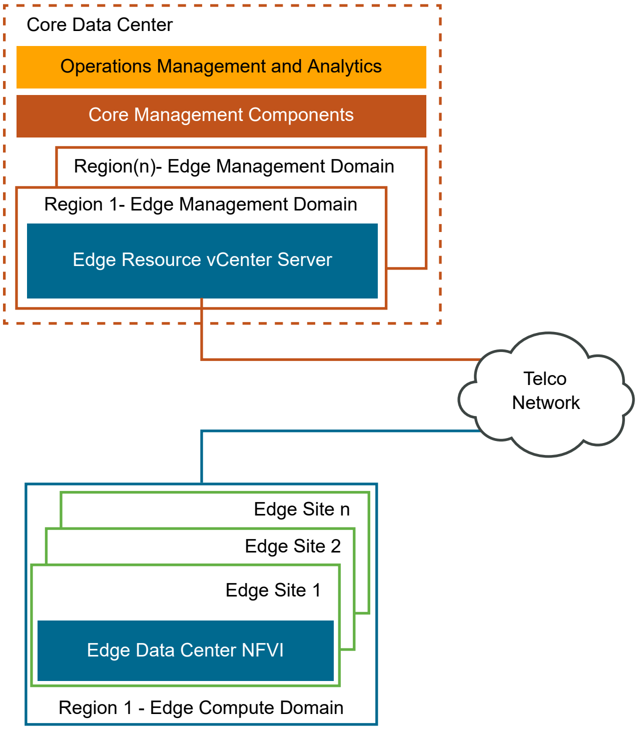
The Core site is connected to the Edge site through a Telco network generically described as metro/WAN in this reference architecture. Examples of such networks include Metro Ethernet, MPLS, and so on. However, the actual technology used is not pertinent to this reference architecture. Because the core sites and the Telco Edges are connected over Layer 3, it is important that a routed network topology exists between the sites.

Layer 2 networks are expected to be terminated at the Provider Edge (PE) routers at the Core and Edge data centers. A Layer 3 path and connection between the core PE router and the Edge PE router through the metro or WAN network is assumed to be configured already. It is important to ensure that sufficient bandwidth at low latency is ensured for this Core site to Edge site connectivity.

Logical Architecture

The vCloud NFV Edge reference architecture implements the conceptual architecture that is outlined and defined at a high level through the logical building blocks and core components. The following diagram maps the conceptual architecture to a logical view for the vCloud NFV Edge reference architecture.

Figure 8-20. Telco Edge Logical Building Blocks


Core Data Center
Operations Management and Analytics
Core Management Components
Region(n)- Edge Management Domain
Region 1- Edge Management Domain
Edge Resource vCenter Server
Telco Network
Edge Site n
Edge Site 2
Edge Site 1
Edge Data Center NFVI
Region 1 - Edge Compute Domain

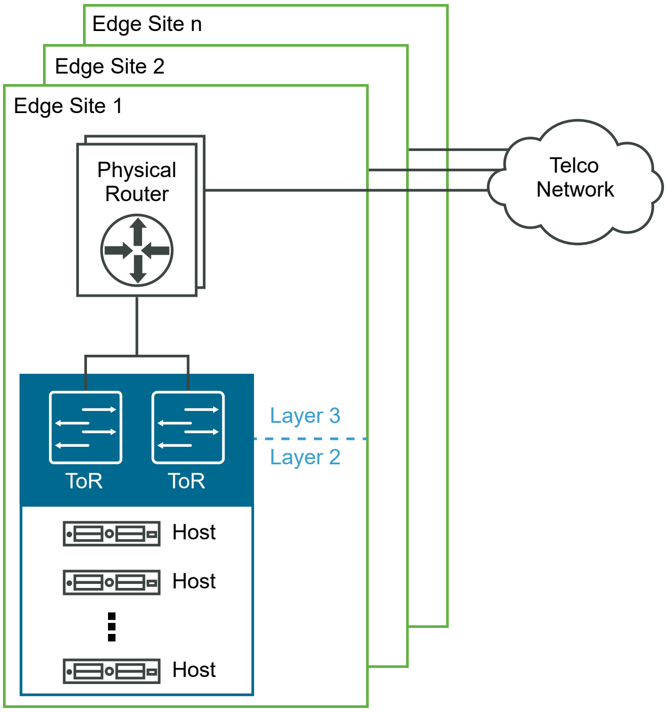
Core Data Center: The Core data center is used to house the management components and other NFV functions that need to run at a central location. For more information, see Core Data Center in the Conceptual Architecture section.

Telco Edge: The Telco Edge is the site that houses the remote workloads (VNFs and applications). It consists of NSX Edge and VNF workloads in a regional Edge compute domain. The Edge compute domain maps to a vSphere cluster managed by the Edge Resource vCenter Server in the Core data center. The number of edge servers at a remote site depends on the workload to run. The Edge Compute Domain is also used to host the NSX Edge Node VMs that forward traffic from the logical network to the physical network in a north-south direction through the edge router and to the metro/WAN network connecting the edge and core sites.

Physical Configuration

The Telco Edge site follows the collapsed edge and resource cluster design as shown in the following figure. While the management components are hosted in the Core data center, the Edge site hosts the low footprint collapsed Edge/resource cluster. The Edge site includes a minimum of three hosts (four recommended if using vSAN) that cater to both the Edge and resource workloads of the site.

Figure 8-21. Telco Edge Physical Network Design

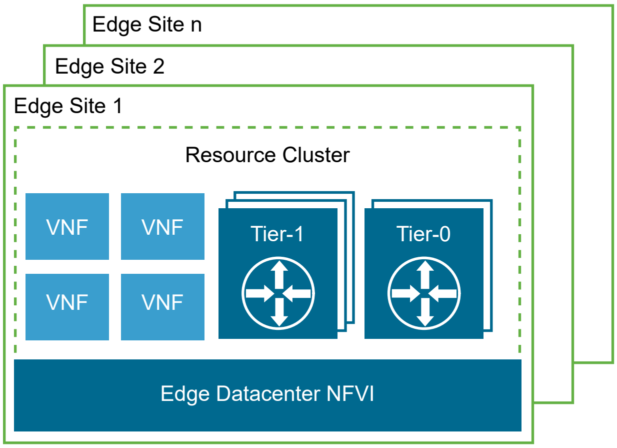


Physical network infrastructure is provided by a combination of ToR and inter-rack/spine switches. The latter is optional and depends on the server count and if all the server ports can be accommodated within a pair of ToR switches. Within a single site, intra-site networking is based on Ethernet (typically 10 Gbps or higher). Each server has at least two NICs connected to a pair of Top of Rack (ToR) switches (for redundancy). This reference architecture assumes that two switches are used with host NIC ports connected to either Switch 1 or Switch 2.

Logical Configuration

This section describes the logical components of the Edge site and their functions. The Edge components provide the fabric for connectivity to the Core data center and also to other edge sites. Multiple instances of the Edge components may be used for performance and scale.

Figure 8-22. Telco Edge Logical Components



The preceding diagram depicts the individual Edge sites with their own VNFs deployed (typically user plane VNFs in the 5G context). The Edge data center NFVI consists of vSphere, vSAN, and NSX data plane components with the management components at the central site. The Tier-1 (tenant-level gateways) and Tier-0 gateway forward the traffic both east-west and north-south into and out of the infrastructure. The Tier-0 gateways are used for north-south traffic and to connect to the provider/customer edge router in the edge site. They are instantiated in two NSX Edge Node VMs which run on the hosts in the Edge Site. Host anti-affinity rules are configured for these VMs to tolerate single host failures.

Logical Building Blocks

The platform components are grouped into Edge Management Domain and Edge Compute Domain. While the management domain is used to host the management components for edge sites, the resource domain is used to host the NSX Edge VMs and VNFs.

Telco Edge Management Domain

An Edge data center is mapped to an edge site. Edge sites (or Edge data centers) are grouped into regions and have a corresponding instance of Resource vCenter Server deployed in the Core data center to manage the Edge Compute Domains.

The number of edge sites within a region is limited by the configuration maximums of the management components for that region. When the maximum limit is reached, a new management instance of the components is deployed to accommodate the growth.

The sizing of an edge site depends on factors influenced by the inter-dependencies between management components and their configuration limits. The number of supported edge sites varies depending on the sizing of each site. All the edge sites need not be sized equally but are sized based on their respective workload requirements.

Telco Edge Compute Domain

Logically an edge site is a separate vSphere cluster with hosts connecting to and managed by the vCenter Server that is a part of the Edge Management Domain for the corresponding region.

An Edge site follows the collapsed edge/resource pod design used in the vCloud NFV reference architecture. This design entails a minimum of three servers (four recommended if using vSAN) to provide pooled compute resources to both the VNF workloads deployed at the site and to the NSX Edge nodes.

The edge storage may be provided by any supported shared storage solution. This reference architecture uses vSAN as the storage provider. Each server should have local disks for vSAN caching and capacity tiers. An All-Flash vSAN is recommended for reliability and performance.

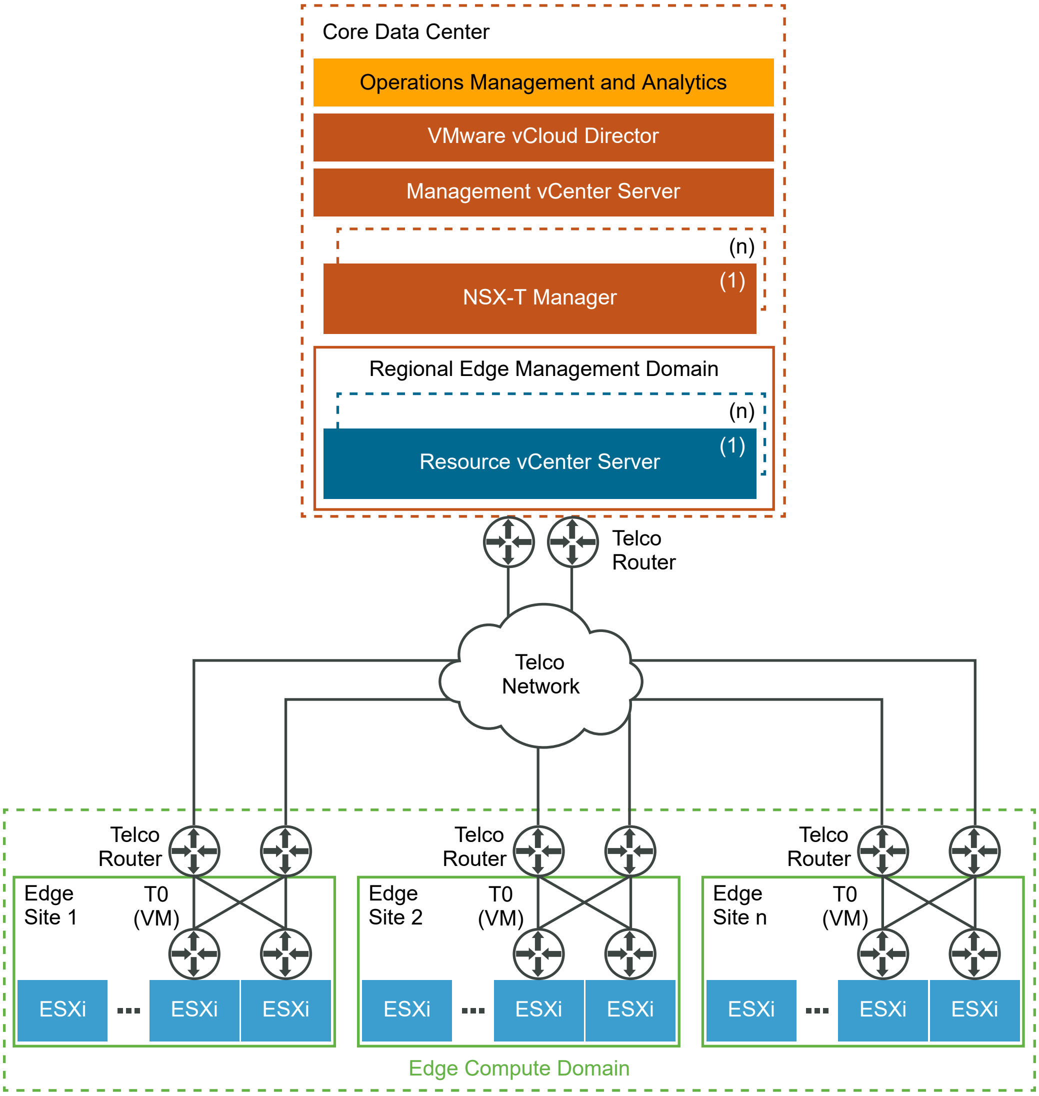
Each physical host should have a minimum of four physical NICs connected to a pair of ToR switches configured for redundancy. The ToR switches connect to an external WAN Edge physical router to transport packets for Internet breakout and backhaul to the Core site.

Virtual Building Blocks

The virtual infrastructure design comprises the software components that form the virtual infrastructure layer. This layer supports running Telco workloads and the workloads that maintain the business continuity of services. The virtual infrastructure components include the virtualization platform hypervisor, virtualization management, storage virtualization, network virtualization, and backup and disaster recovery components.

This section outlines the building blocks for the virtual infrastructure, their components, and the networking to tie all the components together.

Figure 8-23. Telco Edge Virtual Building Blocks



Compute Design

It is important to limit the distance between the core site and the edge sites to ensure that the latency is below 150 ms RTT. Also, each site is treated as a remote cluster with its own storage; HCI storage with vSAN is recommended. An NSX Edge (pair) needs to be deployed at the remote site (even though the NSX Manager and Controller reside at the Core site) for connectivity to the Core site and for Internet breakout.

The network links between the core site and the edge sites should also be redundant and path-diverse without any SRLGs (Shared Risk Link Groups) between the paths at a transport layer. Also, sufficient bandwidth between each edge site and the core site should be ensured.

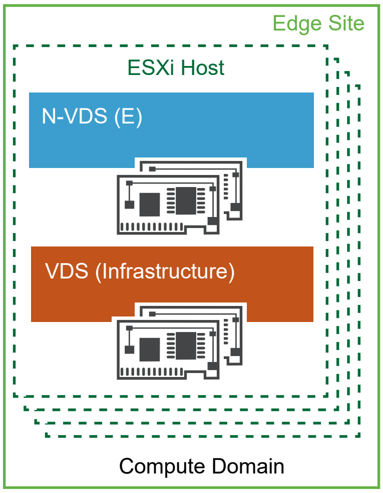
Storage Design

vSAN is a fully integrated hyper-converged storage software. By creating a cluster of server Hard Disk Drives (HDDs) or Solid-State Drives (SSDs), vSAN presents a flash-optimized, highly resilient, shared storage datastore to ESXi hosts and virtual machines. This allows for the control of capacity, performance, and availability through storage policies on a per VM basis. Certified third-party shared storage solutions as listed in the VMware Compatibility Guide are also supported.

Network Design

The vCloud NFV Edge platform consists of infrastructure networks and VM networks. The hosts in each cluster are configured with VDS switches that provide consistent network configuration across multiple hosts. VDS switch is used for VM networks and infrastructure networks while the N-VDS switch is used as the transport for Telco workload traffic.

Figure 8-24. Telco Edge Virtual Network Design



The networks on the VDS (infrastructure) switch used in the edge site include:

- **ESXi Management Network**: The network for the ESXi host management traffic.
- **vMotion Network**: The network for the VMware vSphere® vMotion® traffic.
- **vSAN Network**: The network for the vSAN shared storage traffic.
- **VM Network**: The network for the management VM traffic.

Telco Edge Management Domain

The Telco Edge Management Domain is responsible for the orchestration of resources and operations of the Edge site. This includes vCenter Server Appliance that manages the virtual infrastructure resources of the Edge Region runtime environment. The design of the edge management domain component remains identical to the core data center management domain component and is not covered in this section.

Isolation of compute resources is enabled with the Resource vCenter Server for the Edge. Irrespective of the Domain deployment configuration, abstraction layers for multi-tenancy are provided by the vCloud Director deployed in the Core data center. vCenter Server provides the infrastructure for fine-grained allocation and partitioning of compute and storage resources, whereas NSX-T deployed in the Core data center provides the virtual network resources for Edge site.

Telco Edge Compute Domain

The components of the Edge Compute Domain and their functions constitute the edge compute domain. NSX Edge node and VNFs are placed in the Edge Compute Domain cluster that forms the runtime environment for the VNFs.

Telco Edge Compute Domain Networking: The networking of the Edge Compute Domain depends on the network topology that is required by the Telco workloads deployed by the tenant. The network building blocks as required by tenant workloads is identical to the network building blocks of the Core Data Center compute domain networking.

Telco Edge Site Networking: Edge sites can be connected to two separate domains. The first domain is an Internet breakout where the tunneled traffic from the user equipment is terminated and routed as IP packets to the Internet. The second domain is where the traffic continues to be tunneled to the central site (as happens today with user traffic). In both cases, the Edge site uses a physical router as the egress device to transport traffic to the Internet or to the central site.

There are multiple options for the physical router egress connectivity, such as metro Ethernet and MPLS. The technology that is used to connect Edge to the Internet or Core site does not impact this reference architecture, except for certain latency and speed requirements.

Edge Deployment

This Reference Architecture proposes a centralized model where the VIM, Operations Management, and analytics components are all hosted in the central site. The Edge sites only have the ESXi hosts. Management components for each region are identical and follow the architecture described earlier in this document.

Network Design

The vCloud NFV Edge Reference Architecture network solution consists of separate orchestration, management, control, and data planes.

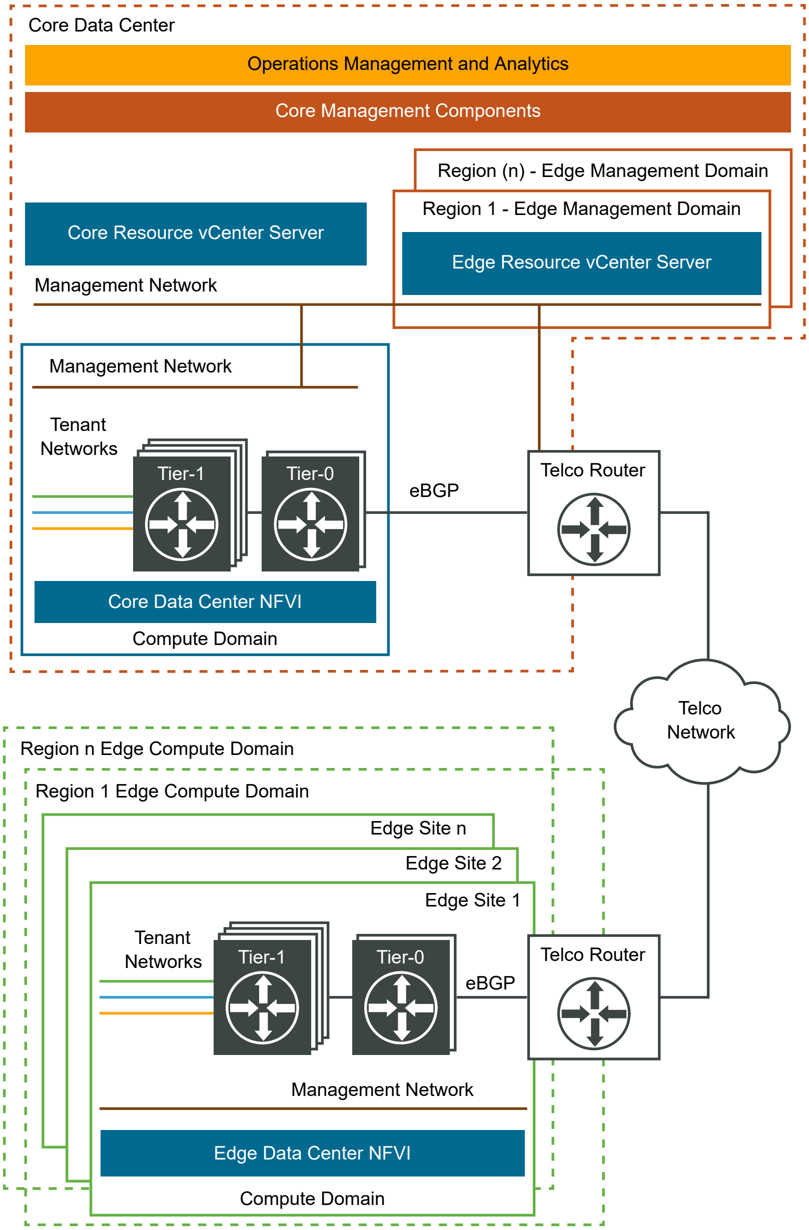
Management Plane: This plane is responsible for central configuration and monitoring. The management plane assists in the automatic onboarding of CE/PE routers into the NSX-T Edge T0/T1 overlay.

Control Plane: This plane builds and maintains the network topology and makes decisions on traffic flows.

Data Plane: This plane is responsible for forwarding packets based on decisions from the control plane.

Note The WAN connectivity between Central and Edge sites is beyond the scope of this reference architecture; customers need to ensure connectivity between the two sites. All Edge Sites are connected to its respective aggregation site which is the corresponding Core data center for the region.

Figure 8-25. Telco Edge Deployment Network Design



Design Considerations

This reference architecture assumes a separate network connection over Layer 3 for management connectivity between the management components and its edge sites. This management connectivity includes traffic between vCenter Server and Edge site ESXi hosts. NSX Manager also uses this for management of NSX Edge Nodes at the Edge site.

A pair of NSX-T Edge Nodes (in VM form factor) is used at each Edge site for the logical to the physical network and also to assist in the mapping of tenant gateways when a multi-tenant environment is needed. Note that the segmentation of tenants and QoS at the networking level may increase the number of Edge Nodes per site.

Note The end-to-end round-trip latency between any Edge site and core site should not exceed 150 ms. Recommended bandwidth between the Edge and core sites is 10 Gbps. VLAN-based network segmentation is restricted within a data center. There is no VLAN stretching between the core and Edge sites.

Network Redundancy

The vCloud NFV Edge reference architecture configuration has Edge nodes (in the VM form factor) in active/active mode to connect to the Provider Edge (PE) router at the Edge site. To define the high availability configuration for the edge node, the administrator from the Core data center must use a control plane network.

Operations Management

There are two models for placement of the operations management components such as vROps, vRNI, and vRLI:

- The central components of these products are always placed at the core site. Scaling of these products depends on the number of Edge sites under management and the total number of workloads at those Edge sites.
- The remote collector components of these components are to be placed at the Edge sites.

There are three FCAP collectors: Remote Collector for vROps, Proxy for vRNI, and Syslog collector for vRLI. For potentially large-scale deployments, consider placing the remote collectors at the Edge sites.

Network Tenancy

The vCloud NFV Edge Reference Architecture relies on NSX-T to provide network tenancy for end-to-end isolation capabilities by deploying multiple tiers of distributed routing through Tier-0 and Tier-1 gateways in the networking stack.

The uplink of a Tier-0 gateway that resides in NSX-T Edge is connected to upstream physical routers. A tenant uses a Tier-1 gateway at its Edge to connect to the Tier-0 gateway. Tier-0 gateway relays traffic to other tenants on the upstream router at each side of the Core data center or Edge site. Network virtualization capabilities with Geneve encapsulation provide flexibility in-line with industry standards. NSX-T Data Center performance enhancements for N-VDS and NSX Edge Nodes offer advanced network capabilities.

Each tenant's traffic is associated with a different VLAN behind the per-tenant WAN access. Similar to a physical switch, an N-VDS Uplink port can carry multiple VLANs encapsulated on the single connected link using IEEE 802.1q.

Architectural Realization

This section covers the multi-tenancy use case scenario applied to the Telco Edge CSP cloud infrastructure environment with vCloud Director.

Telco Edge Workload Placement

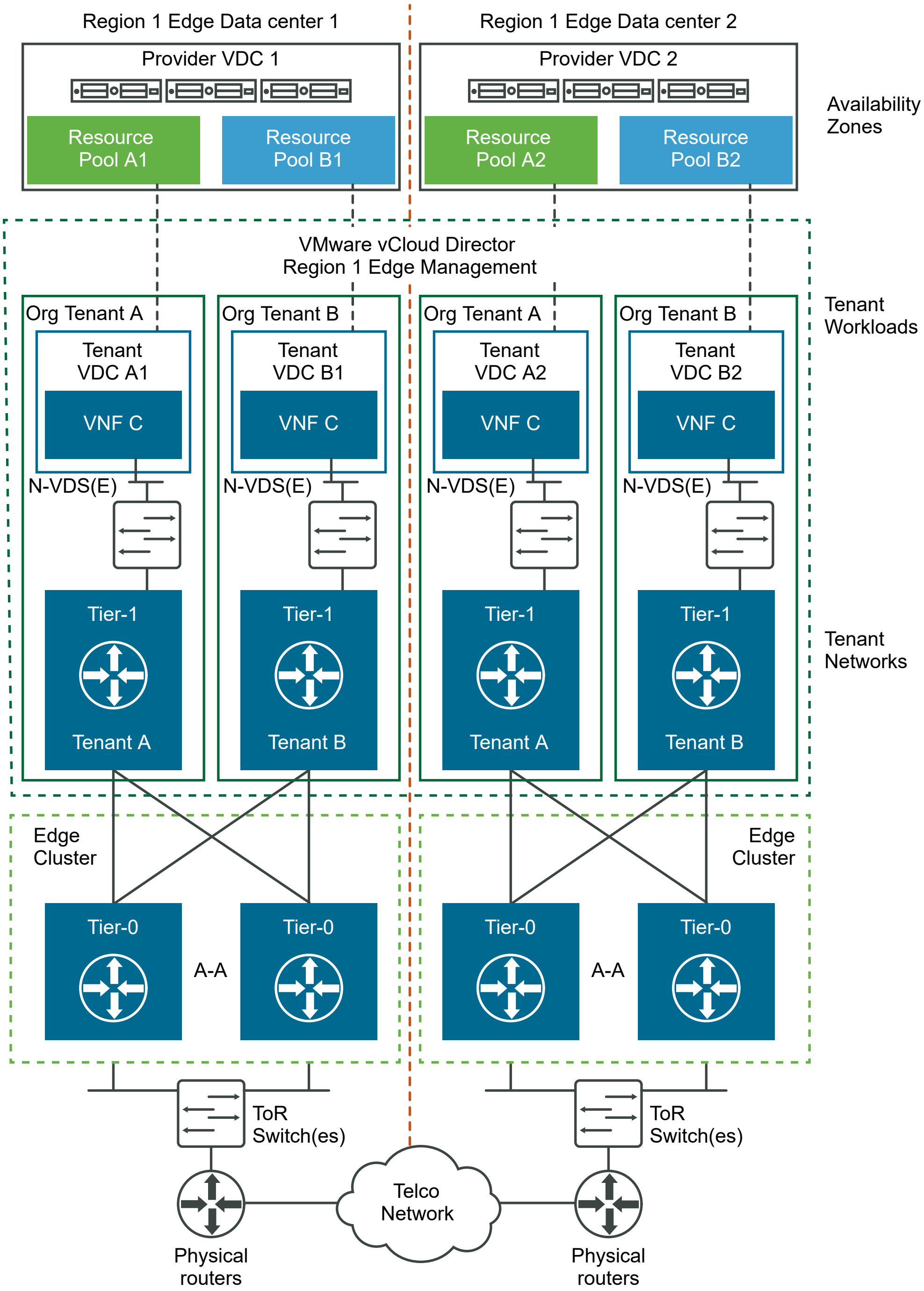
The placement of workloads is typically a subset of the VNF onboarding process and is a collaborative effort between the VNF vendor and the CSP. A prescriptive set of steps must be followed to package and deploy a VNF. The VMware Ready for NFV program is a good vehicle to pre-certify the VNFs and the onboarding compliance with the vCloud Director platform to ensure smooth deployment in the CSP environment.

Before a VNF is onboarded, the VNF vendor must provide the CSP with all the prerequisites for the successful onboarding of the VNF. This includes information such as the VNF format, number of the required networks, East-West and North-South network connectivity, routing policy, security policy, IP ranges, and performance requirements.

After the VNF is onboarded, the tenant administrator deploys the VNF to either the Core data center or the Edge data center depending on the defined policies and workload requirements. The DRS, NUMA, and Nova Schedulers ensure that the initial placement of the workload meets the target host aggregate and acceleration configurations defined in the policy. Dynamic workload balancing ensures that the policies are respected when there is resource contention. The workload balancing can be manual, semi-supervised, or fully automated.

The following diagram describes the workload placement architecture for the edge sites with vCloud Director:

Figure 8-26. Telco Edge Multi-Tenancy with vCloud Director



Analytics and Monitoring

9

CSPs can enable the vCloud NFV platform for day 1 and day 2 operations after the platform is deployed in the cloud provider topology. The platform is integrated with an operations management suite that provides capabilities for health monitoring, issue isolation, security, and remediation of the NFVI and VNFs.

The NFVI operations management framework defines and packages a five-step approach to make day 1 and day 2 workflows operational.

- Onboard service operations.
- Service launch and monitoring.
- Dynamic optimizations.
- Issue isolation.
- Demand planning and expansion.

The integrated operational intelligence adapts to the dynamic characteristics of the NFV infrastructure to ensure service quality and issue resolution. Some of the key characteristics include:

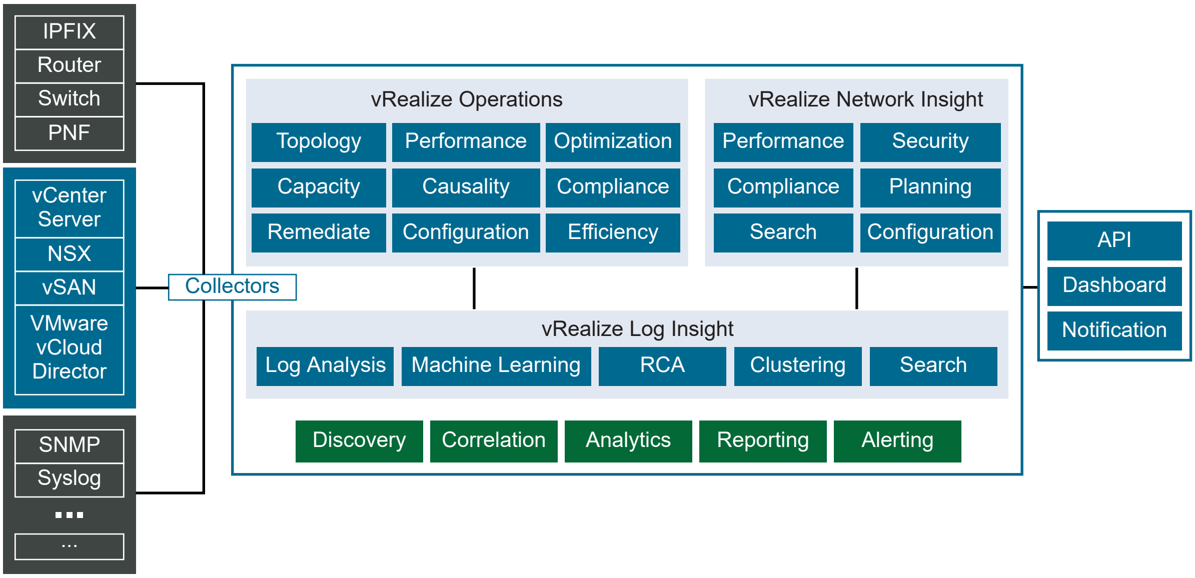
- **Dynamic resource discovery**: Distributed and complex topologies together with workloads in motion require dynamic resource and service discovery. The platform provides continuous visibility over service provisioning, workload migrations, auto-scaling, elastic networking, and network-sliced multitenancy that spans across VNFs, hosts, clusters, and sites.
- **SLA management**: Continuous operational intelligence and alert notifications enable proactive service optimizations, capacity scale-out or scale-in, SLA violations, configuration and compliance gaps, and security vulnerabilities.
- **Remediation**: Reduced MTTU and timely issue isolation for improved service reliability and availability. Prioritized alerting, recommendations, and advanced log searching enable isolation of service issues across physical and overlay networks.
- **Security and policy controls**: Multivendor services operating in a shared resource pool can create security risks within the virtual environment.
  - Ability to profile and monitor traffic segments, types, and destination to recommend security rules and policies for north-south and east-west traffic.

  - Identification of security policy and configuration violations, performance impacts, and traffic routes.

- **Capacity planning and forecasting**: New business models and flexible networks demand efficient capacity planning and forecasting abilities in contrast to the traditional approach of over-provisioning that is costly and unrealistic.

The framework continuously collects data from local and distributed agents, correlating, analyzing, and enabling day 2 operations. The analytical intelligence can be also queried and triggered by third-party components such as existing assurance engines, Network Management System (NMS), EMS, OSS/BSS, VNFM, and NFV-O for closed-loop remediation.

Figure 9-1. Analytics and Monitoring Overview



CSPs can deploy the operations management components in the Management Pod and centralize them across the cloud topology, assuming that inter-site latency constraints are met.

- vRealize Operations Manager collects compute, storage, and networking data providing performance and fault visibility over hosts, hypervisors, virtual machines, clusters, and site.

- vRealize Log Insight captures unstructured data from the environment, providing log analysis and analytics for issue isolation. Platform component logs and events are ingested, tokenized, and mined for intelligence so that they can be searched, filtered, aggregated, and alerted.

- vRealize Network Insight provides layer 2, 3, and 4 visibility into the virtual and physical networks and security policy gaps. The engine is integrated with the NFVI networking fabric, ingesting data that ranges in performance metrics, device and network configuration, IPFIX flow, and SNMP. It discovers gaps in network traffic optimization, micro-segmentation, compliance, security violations, traffic routing, and performance.

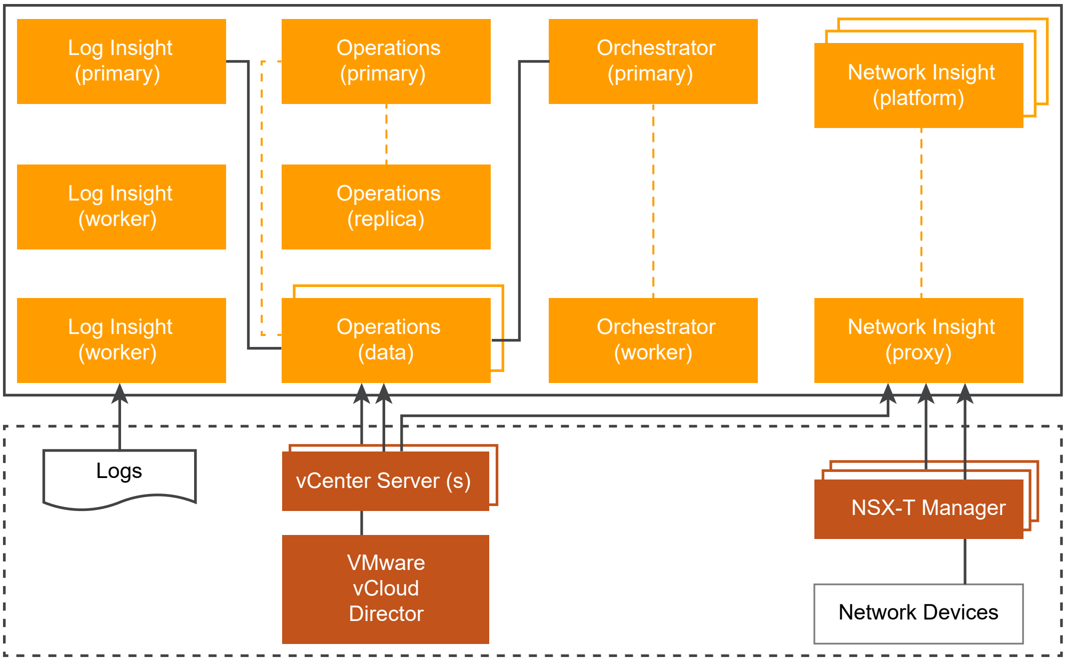
This chapter includes the following topics:

- Management Pod Extensions

Management Pod Extensions

The analytics-enhanced reference architecture enriches the Management Pod with the vRealize management components to provide infrastructure assurance capability.

Figure 9-2. Analytics Extension to Management Pod



Components

The operations management components are deployed as a centralized function that is capable of day 1 and day 2 operations spanning the CSP's deployment topology. The data collection architecture is specific to each operations management component with a centralized single pane for monitoring, reporting, troubleshooting, and closed-loop automation.

vRealize Operations Manager

The virtual environment relies on a monitoring solution that can collect data regarding its health, capacity, availability, and performance. vRealize Operations Manager provides a robust and integrated monitoring platform that resides at the center of the NFV environment. It monitors the virtual environment and collects data about its health, capacity, availability, and performance. vRealize Operations Manager serves as the single pane of glass to the NFV environment.

vRealize Operations Manager extends and collects information through management packs. The collected information is filtered for relevancy, analyzed, and presented in customizable dashboards. It exposes an API that retrieves performance and health data about NFVI and the virtual resources of the VNF instance.

The design of the operations management components is based on centralized management and collection, with an optional remote collection for a distributed topology. vRealize Operations Manager supports HA across the various components. HA creates a primary replica for the vRealize Operations Manager primary node and protects the management functions. In smaller deployments, the primary node can also act as a data node. In larger deployments, data nodes host adapters that are responsible for collecting data and can be scaled to meet additional capacity. To enable HA, at least one more data node must be deployed in addition to the primary node. Anti-affinity rules should be used to keep nodes on specific hosts.

vRealize Operations Manager is installed in the Management Pod in both Two-Pod and Three-Pod designs. Depending on the number of metrics that are collected over time, additional storage capacity and compute capacity might be required. Adding more hosts to the management cluster or more storage is sufficient to address the growing storage needs of vRealize Operations Manager.

By default, VMware offers Extra-Small, Small, Medium, Large, and Extra-Large configurations during installation. The CSP can size the environment according to the existing infrastructure to be monitored. After the vRealize Operations Manager instance outgrows the existing size, the CSP must expand the cluster to add nodes of the same size. For more information, see the vRealize Operations Manager Sizing Guidelines.

vRealize Log Insight

CSPs can use vRealize Log Insight to collect log data from ESXi hosts and data about server events, tasks, and alarms from vCenter Server systems. vRealize Log Insight integrates with vRealize Operations Manager to send notification events. Because vRealize Log Insight collects real-time unstructured data, the CSP can configure all elements in the NFV environment to send their log data to vRealize Log Insight. This log aggregation provides a single log collector for the entire NFV environment.

vRealize Log Insight ingests syslog data from the physical and virtual NFVI components to deliver monitoring, search, and log analytics. It builds an index for analytics purposes by automatically identifying structure from machine-generated log data including application logs, network traces, configuration files, messages, performance data, and system state dumps. Coupled with a dashboard for stored queries, reports, and alerts, vRealize Log Insight assists the CSP in root cause analysis and reduction in MTTR. All NSX Manager syslog information, distributed firewall logs, and NSX Edge syslog information is sent to vRealize Log Insight.

The vRealize Log Insight API provides programmatic access to the 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.

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

vRealize Log Insight provides preset VM sizes that the CSP can select from to meet the ingestion requirements of their environment, Extra-Small, Small, Medium, Large, and Extra-Large configurations. These presets are certified size combinations of compute and disk resources, though extra resources can be added afterward. For the sizing details, see the VMware vRealize Log Insight documentation.

vRealize Network Insight

vRealize Network Insight provides operations for software-defined networking and security across virtual and physical infrastructure with micro-segmentation planning that can be scaled to thousands of VNFs. vRealize Network Insight is installed in the Management Pod in both Two-Pod and Three-Pod designs.

The vRealize Network Insight architecture consists of a platform VM, a proxy VM, and data sources. The platform VM provides analytics, storage, and a user interface to the data. The proxy VM or the collector collects data by using various protocols such as HTTPS, SSH, CLI, and SNMP, depending on the source and the configuration. Various data sources are supported, including vCenter Server, NSX-T Data Center, firewalls, and various switch vendors.

The platform VM is deployed as a single cluster to provide high availability and scale. A minimum of three platform VMs are required in the cluster. The proxy VMs are used to collect data and can be deployed in a single data center or distributed across sites. Depending on the amount of data that is collected, typically CSPs need one or more proxy VMs.

vRealize Network Insight provides the following enhanced capabilities for NSX-T Data Center:

- Support for Tier-0 and Tier-1 gateways for multi-tenancy and micro-segmentation along the East-West network.
- Provides NSX-T Data Center inventory with a number of nodes, layer2 networks, firewall rules, and logical routers.
- Support for NSX Distributed Firewall (DFW) generated IPFIX flows and firewall rule recommendations to the micro-segmentation applications.
- Support for monitoring overlay and underlay virtual machine networks managed by NSX-T Data Center.
- End to End Virtual machine network visibility for multiple Equal-Cost Multipath (ECMP) paths between routers.

Ensure that the system meets the minimum hardware configurations to install vRealize Network Insight. For the sizing details, see the VMware vRealize Network Insight documentation.

vRealize Orchestrator

VMware vRealize Orchestrator is a development-and process-automation platform that provides a library of extensible workflows. Orchestrator workflows run on objects that are exposed through plug-ins and custom scripting to interact with any component that is reachable through an API. By default, a VMware vCenter plug-in is provided to orchestrate tasks in the cloud infrastructure environment. The CSP can use Orchestrator workflows to define and run automated configurable processes thus creating a framework for closed-loop automation.

vRealize Orchestrator is integrated with vRealize Operations Manager through a management pack that provides workflows for automating the cloud infrastructure environment and for orchestration of third-party components as well as management functions.

The vRealize Orchestrator is highly available and configured as a single cluster of multiple virtual appliance instances. The appliance is registered with a vCenter Single Sign-On service by using the vSphere Authentication mode. The appliance also requires a shared database instance.

Ensure that the system meets the minimum hardware configurations to install vRealize Orchestrator. For the sizing details, see the VMware vRealize Orchestrator documentation.

Enabling Analytics with vRealize

The vRealize components of the vCloud NFV platform are integrated into the platform and configured to report and trigger analytical intelligence. Various configurations and extensions that Follow help the CSP to get started with the analytics-enabled reference architecture.

vRealize Operations

vRealize Operations is configured with the adapters that are necessary to start collecting data from the infrastructure. The following solutions should be considered:

- **vCenter Adapter**: vSphere connects vRealize Operations Manager to one or more Resource and Management vCenter Server instances. The system collects data and metrics from those instances, monitors them, and runs actions in them.
- **Endpoint Operations Management**: The Endpoint Operations Management solution is used to gather operating system metrics and to monitor the availability of remote platforms and applications.
- **vSAN Adapter**: In a production environment, operations management for vSAN can be provided by using dashboards to evaluate, manage, and optimize the performance of vSAN and vSAN-enabled objects in the vCenter Server system.
- **vCloud Director Adapter**: The vRealize® Operations Management Pack™ for vCloud Director includes dashboards to provide visibility to vCloud Director deployments, such as Compute Infrastructure, Tenants, and Storage for easy monitoring and management.
- **NSX Adapter**: This provides the CSPs with insight to the health of the network infrastructure components such as NSX Manager and the network devices, both virtual and physical.

- **vRealize Orchestrator Solution**: This management pack enables an adapter to communicate with the vRealize Orchestrator workflow automation engine.

vRealize Network Insight

vRealize Network Insight can be configured to monitor all networking-related components in the NFVI. vRealize Network Insight can connect to the NFV components that are related to networking and security and provide insights into the networking segments that are used to deploy and manage the vCloud NFV Edition platform. The management VLANs, external VLANs, and Overlay segments are all available for monitoring and diagnostics.

vRealize Network Insight uses the following data sources:

- **vCenter Server**: Both Resource and Management vCenter Server instances.
- **NSX Manager**: The management plane of the NSX-T Data Center networking.
- **IPFIX**: For flow analysis. It can be enabled within vCenter Server.
- **Network devices**: Physical devices such as Dell switches, Cisco Nexus and Catalyst switches, Arista, Juniper Networks, Hewlett Packard Enterprise, Brocade, and Palo Alto Networks switches.

vRealize Orchestrator

vRealize Orchestrator is integrated into vRealize Operations Manager as a management pack to provide bidirectional communication with the Orchestrator workflow engine. The management pack is deployed in vRealize Operations Manager and configured to point and authenticate to the vRealize Orchestrator instance.

VMware vRealize Operations Management Pack

vRealize Operations Manager can be extended through management packs to monitor VNFs based on virtual machines.

vRealize Operations Manager collects structured data from various vCloud NFV Edition components, including gathering data from adapters that are connected to external components. For this mechanism to work, vRealize Operations Manager is configured to communicate with data sources by using an authorized user account for the respective components. 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 VMware Solution Exchange.

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

Out-of-the-box, vRealize Operations Manager does not monitor the VNF service availability or VNF internal key KPIs. The VNF Manager derives this information through direct interaction with the respective VNFs. To extend the operations management functionality to the VNFs, VNF vendors can create custom management packs for vRealize Operations Manager. Management pack development requires an understanding of the vRealize Operations Manager inventory model and the management functions that management packs implement. These include auto-discovery and monitoring. For more information, see Endpoint Operations Management Agent Plugin Development Kit .

VMware vRealize Log Insight Content Pack

vRealize Log Insight gathers log events natively from multiple syslog data sources and through special content packs. Specific dashboards can be customized to perform log analytics and alerts. For additional information about vRealize Log Insight solutions, see VMware Solution Exchange.

Authors and Contributors

10

The following authors co-wrote this paper:

- Arunkumar Kodagi Ramachandra, Solutions Architect, NFV Solutions Engineering, VMware
- Indranil Bal, Solution Consultant, Telco Solutions, VMware
- Ram Pratap Singh, Senior Member of Technical Staff, Telco Solutions, VMware

Many thanks for contributions from:

- Jambi Ganbar, Sr. Technical Solutions Manager, Telco Solutions Engineering, VMware
- Michelle Han, Director, Solutions Testing and Validation, Telco Solutions Engineering, VMware
- Mike Brown, Sr. Staff Solution Architect, Telco Solutions, VMware
- Ramkumar Venketaramani, Director Product Management, Telco Solutions, VMware
- Revathi Govindarajan, Senior Technical Writer - Professional, Telco Solutions, VMware
- Sudesh Tendulkar, Staff Solutions Architect, Telco Solutions, VMware
- Sumit Verdi, Sr. Director, Telco Solutions Management, Telco Solutions, VMware
- Xiao Gao, Lead Architect, NFV Lighthouse Solutions, Telco Solutions, VMware

Special thanks for their contribution and feedback to:

- Andrea Li, Ashish Chorge, Brindha Palanivel, Deepak Yadav, Eswaran K S, Kandasamy M, Piyush Kumar Singh, Prankul Bansal, Ramesh Tammana, Rohit Singh, Savvina Mitrova, Srinivas Rao Veerla, Yordan Ugrinov