# Edge Reference Architecture for vCloud NFV

VMware vCloud NFV OpenStack Edition 3.1

vmware®

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

https://docs.vmware.com/

If you have comments about this documentation, submit your feedback to

docfeedback@vmware.com

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

# Contents


**1** Edge Reference Architecture for vCloud NFV 5

- Introduction to vCloud NFV Telco Edge 5
- Acronyms and Definitions 6
- Reference Environment 7
- Telco Edge Conceptual Architecture 11

**2** Architectural Framework and Components 14

- Key Stakeholders 14
- Logical Architecture and Components 15
  - Core Data Center 15
  - Telco Edge 18
- NFV OpenStack Edition Components 20
- Design Principles 21

**3** Telco Edge Reference Architecture 23

- Telco Edge Logical Building Blocks 23
- Telco Edge Virtual Building Blocks 25
  - Telco Edge Management Domain 28
  - Telco Edge Compute Domain 34

**4** Telco Edge Deployment 39

- Telco Edge Deployment Network Design 40
- Design Considerations 41

**5** Telco Edge Analytical Architecture 44

- Introduction to Analytics for Telco Edge 46
- Analytic Components for Edge 47

**6** Architectural Realization 49

- Multi Tenancy 49
- Telco Edge Workload Placement 51
- Availability and Disaster Recovery 57
  - Availability 57
  - Disaster Recovery 58

**7** Applications for the Telco Edge 59

- Mobile User Plane Functions 59
- vCDN Realization Through MEC 63

Private LTE and Industrial Automation 64

8 Authors and Contributors 65

# Edge Reference Architecture for vCloud NFV

1

This reference architecture provides guidance for designing and creating Network Function Virtualization Infrastructure (NFVI) for distributed Telco Edge deployments by using VMware vCloud® NFV™. It also provides sample deployment scenarios for specific Telco Edge use cases highlighting the platform capabilities. This version of the reference architecture is based on vCloud NFV 3.1 (OpenStack Edition).

## 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 Edge network services and the NFV environment on which they run.

This chapter includes the following topics:


- Introduction to vCloud NFV Telco Edge
- Acronyms and Definitions
- Reference Environment
- Telco Edge Conceptual Architecture


## Introduction to vCloud NFV Telco Edge

Telecommunication operators are in need of a disaggregated 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 can 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 large number of devices that 5G networks need 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.

The VMware vCloud NFV platform is a carrier grade NFV Infrastructure with VMware® Integrated OpenStack and VMware vCloud Director as the two options for NFV Virtualized Infrastructure Manager (VIM).

This reference architecture describes how vCloud NFV, with VMware Integrated OpenStack as the VIM, can be deployed in a disaggregated and distributed fashion to meet the growing needs of the Telco Edge use case.

The VMware vCloud NFV platform is a modular multi-tenant virtual infrastructure with abstractions to enable multi-vendor, multi-domain, and multi-cloud execution environments. The centralized VMware Integrated OpenStack management instance provides the IaaS layer to manage the workload placement and management across multiple Telco Edge sites that are distributed geographically. By supporting standard OpenStack interfaces, the platform inter-operates with external service orchestration and management functions. Being based on the same virtualization infrastructure, the vCloud NFV platform for Telco Edge runs both control plane and data plane workloads, but now in a disaggregated and distributed manner.

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 is used to enhance the run-time environments further with workflows for dynamic workload optimization and proactive issue avoidance.

# Acronyms and Definitions

VMware Telco Edge Architecture uses a specific set of abbreviations that apply to the NFV technology and the Telco industry.

### Table 1-1. General Acronyms

| Abbreviation | Description |
|---|---|
| AR | Augmented Reality |
| BFD | Bidirectional Forwarding Detection, for failure detection on the transport links. |
| CPF | Control Plane Functions |
| IoT | Internet of Things |
| MTTU | Mean Time to Understand. |
| MTTR | Mean Time to Repair. |
| VR | Virtual Reality |
| UPF | User Plane Functions |

### Table 1-2. NFV Acronyms

| Abbreviation | Description |
|---|---|
| CCP | Centralized Control Plane in the NSX-T Data Center architecture.-- |
| CNF | Container Network Function, executing within a Kubernetes environment. |
| LCP | Local Control Plane r in the NSX-T Data Center architecture. |
| LIF | Logical Interfaces |
| MANO | Management and Orchestration components, a term originating from the ETSI NFV architecture framework. |

| Abbreviation | Description |
|---|---|
| NFVI | Network Functions Virtualization Infrastructure. |
| NFVO | Network Functions Virtualization Orchestrator. |
| NFV-OI | NFV Operational Intelligence. |
| N-VDS (E) | Enhanced mode when using the NSX-T Data Center N-VDS logical switch. This mode enables DPDK for workload acceleration. |
| N-VDS (S) | Standard mode when using the NSX-T Data Center N-VDS logical switch. |
| VIM | Virtualized Infrastructure Manager. |
| VNF | Virtual Network Function, executing in a virtual machine. |
| VNFC | Virtual Network Function Component. |
| VNFM | Virtual Network Function Manager. |

### Table 1-3. Telco Acronyms

| Abbreviation | Description |
|---|---|
| 5GC | 5G Core functions. |
| BSS | Business Support Systems. |
| CDN | Content Delivery Network. |
| EMS | Element Management System. |
| EPA | Environmental Protection Agency. |
| EPC | Evolved Packet Core. |
| HSS | Home Subscriber Server in the mobile evolved packet core 4G architecture. |
| MEC | Mobile edge computing. |
| MVNO | Mobile Virtual Network Operator. |
| NMS | Network Management Systems. |
| OSS | Operational Support Systems. |
| 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. |
| UPF | User Plane Function. |
| UPF (B) | Breakout User Plane Function |
| vEPC | Virtualized Evolved Packet Core |
| vRAN | Virtual Radio Access Network |

# Reference Environment

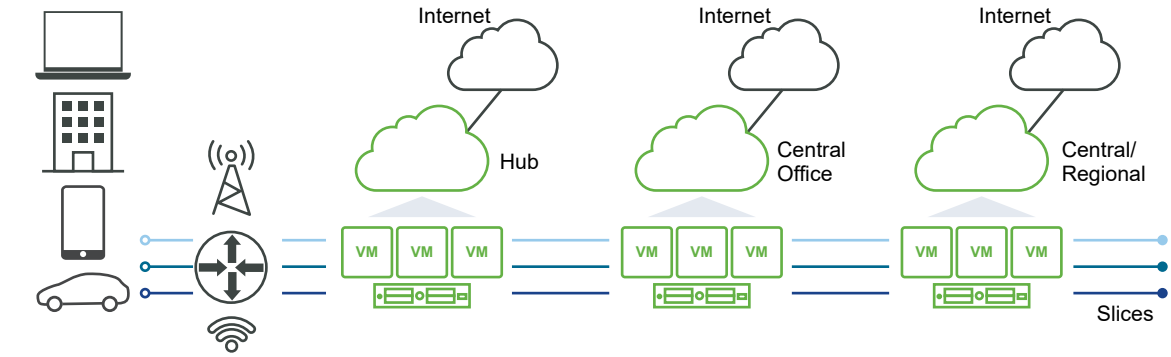
The Telco Edge is a collection of distributed virtual infrastructure deployments that can be used to run Telco specific workloads such as VNFs, and other user applications. Depending on the nature of the

workloads and applications, the position in the network, the size of the deployment, and the software defined infrastructure model, this Edge Reference Architecture provides the desired flexibility.

The mobile network is also transforming into a mixture of highly distributed network functions coupled with some centralized control plane and management plane functions. 5G services are typically comprised of a mixture of low-latency, high throughput, and high user density applications. This requires deployment of both applications and network functions at the edge. The implication for the Telco network is that it evolves from a purely centralized model of service delivery to a highly distributed one. There is also an emerging paradigm shift with employing third-party IaaS, PaaS, and SaaS offerings from public cloud providers. These changes essentially require a more sophisticated service delivery model.

The following diagram shows the reference environment for a Telco network with the addition of edges with clear functional abstractions and interactions in a tiered approach.

**Figure 1-1. 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<br>uCPE | vRAN<br>AR/VR<br>Gaming | vEPC UP<br>MEC<br>Video Surveillance<br>CDN<br>IoT Apps | vEPC CP<br>vIMS<br>5G CP<br>Subscriber (HSS)<br>Policy (PCRF) |

## Customer Edge

On the left of the reference environment in the preceding figure is the customer edge, which comprises millions of end-points (especially with the requirements for IoT). The Customer Edge is on customer premises or it may be a wireless device. The capacity and capability of the customer edge depends on the use cases. SD-WAN is a usecase for such Customer Edges and other services, such as virtual firewalls, and so on. The Customer Edge is not discussed in this Reference Architecture.

## Far Edge

Customer edges connect at the “last mile” to cell towers or wireline aggregation points that are called Far Edges. 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 and related power, cooling, or network factors. There is an Internet “breakout” to allow for applications deployed on the Far Edge to access the Internet directly without having the traffic backhauled to the Near Edge or regional/core location.

## Near Edge

The next level of the hierarchy is the Near Edge, which aggregates traffic from multiple Far Edges and generally has fewer constraints related to capacity. There are a larger number of servers with a higher capacity to run applications. A repurposed central office in the wireline scenario is an example of a Near Edge deployment location. A Near Edge can contain multiple racks in a typical deployment. Latencies from the user equipment to Near Edge are in the range of 5-10 milliseconds, but can vary depending on the deployment.

Content Delivery Network (CDN) and MEC applications are usually instantiated at the Near Edge. The Telco VNFs that are instantiated at the Near Edge include vEPC user plane functions (UPFs) that require higher performance and lower latency. An Internet breakout is also present in this deployment.

The aggregation functionality can involve a separate management plane installation to manage Far Edges. In some cases, the Near Edge is only used to run applications, while the management functionality for both Near and Far Edges is instantiated in a Core data center.

## Core Data Center

The final level of the hierarchy is the Core 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.

For more information on implementing the Core data center and its management components, see the VMware vCloud NFV OpenStack Edition Reference Architecture document.

## Reference Environment Requirements for Edge

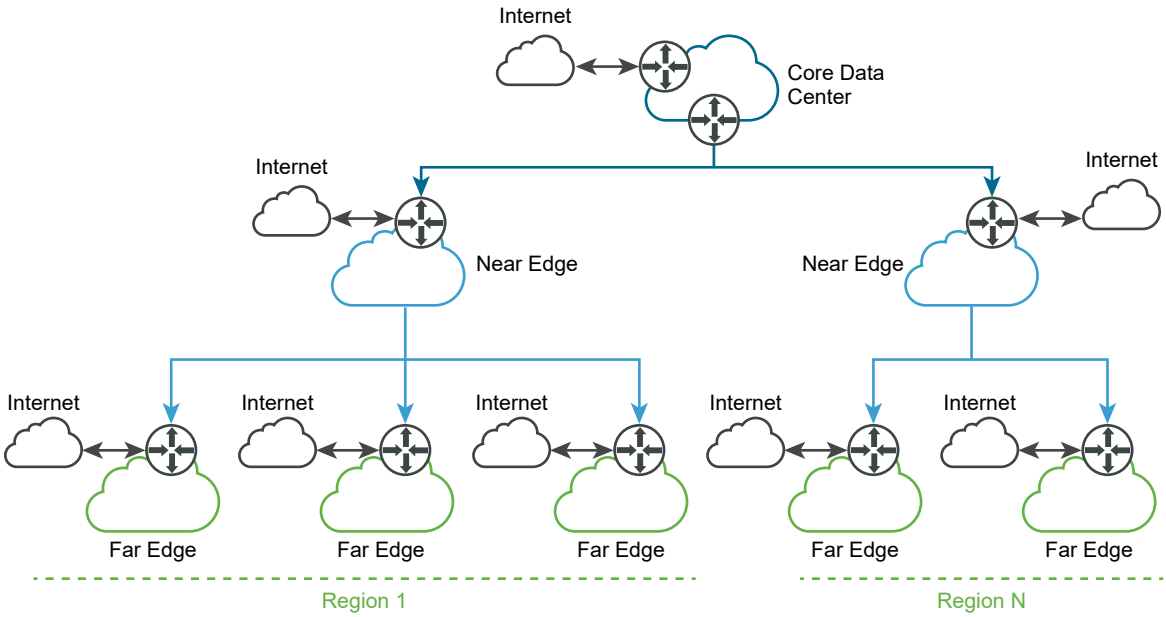
The edge infrastructure reference environment places strict requirements for service placement and management to achieve optimal performance.

| | |
|---|---|
| **Federation options** | The reference environment topology offers a diverse set of federation options for end-points, private and public clouds, each with distinct ownership and management domains. |
| **Disaggregated functions** | Services are highly disaggregated so that control, data, and management planes can be deployed across the distributed topology. Edge clouds offer performance advantages of low latency to allow for data plane intensive workloads while control and management plane components can be centralized with a regional and global scope. |
| **Functional isolation** | With the ability to isolate tenants and providing them with their own resource slices, the reference environment allows for network and service isolation. However, resource management decisions are to be made for shared network functions such as DNS, policy, authentication, and so on. Another facet of 5G technology is the ability to carve an independent slice of the end-to-end mobile network to specific use or enterprise. Each slice has its own end-to-end logical network that includes guarantees, dedicated mobile core elements such as 5G CPF/UPF, and enterprise-specific Telco networking. While the virtual assets are created in the same Telco Cloud infrastructure, it is the responsibility of the Virtualization Infrastructure to provide thecomplete resource isolation with a guarantee for each network slice. This includes compute, network, and storage resource guarantees. |
| **Service placement** | The highly distributed topology allows for flexibility in the workload placement. Making decisions based on proximity, locality, latency, analytical intelligence, and other 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 an efficient life cycle management to ensure service delivery and QoS. |
| **Carrier grade characteristics** | CSPs deliver services that are often regulated by local governments. Hence, 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. |

# Telco Edge Conceptual Architecture

A classic 3-layer conceptual architecture for the Telco Edge deployment is a hierarchical model that consists of a group of Far Edge sites that aggregate into a Near Edge site and a group of Near Edge sites that aggregate into a Core site.

**Figure 1-2. Telco Edge Conceptual Architecture**




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 Far Edge sites aggregating into a Near Edge is usually located in the specific Near Edge site. But, the management component for the Near Edge site can be located locally for expediency reasons.

Each of the Edge sites is individually addressable and the workloads can be placed in the correct Edge site. Each Edge site also has a local breakout capability to the Internet. Therefore, Internet bound traffic does not have to traverse through the Telco network before being routed to its destination.

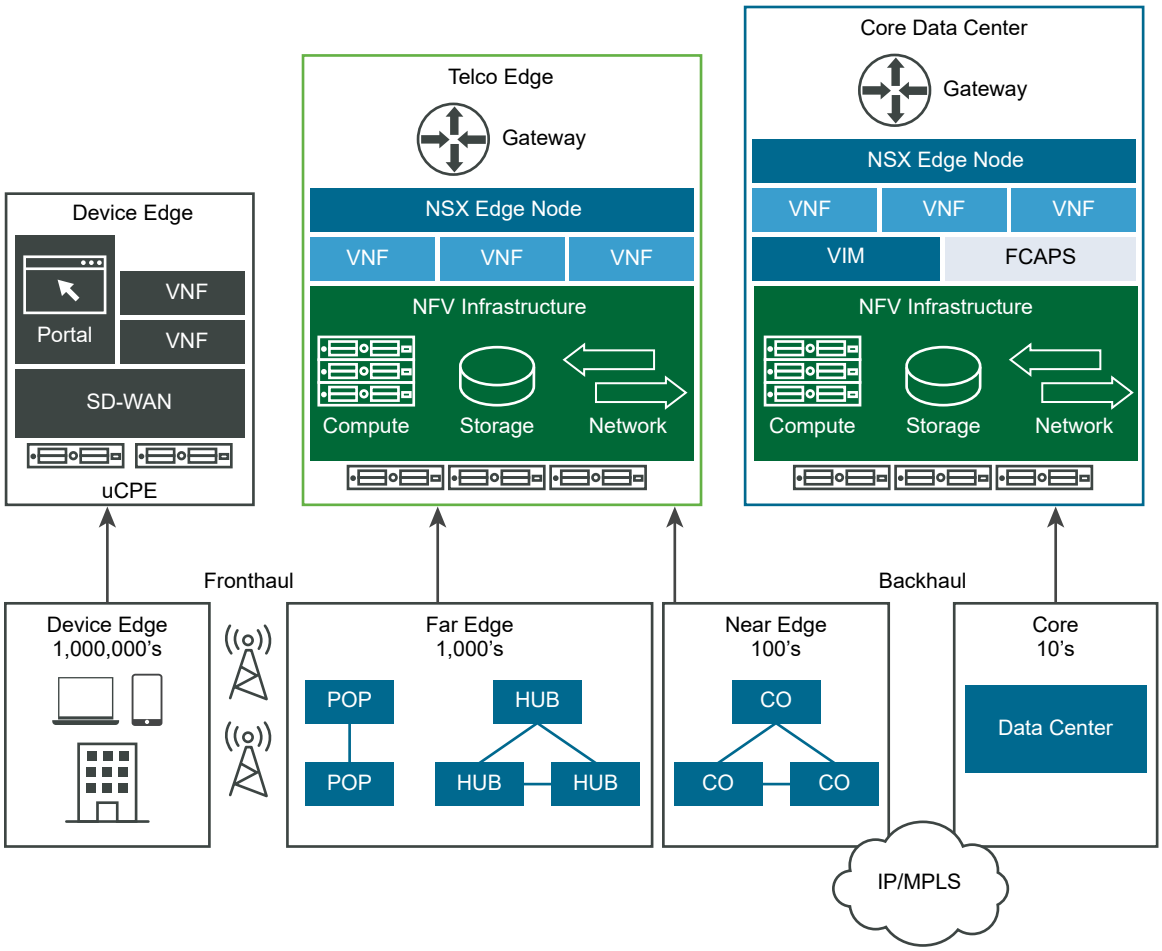
## Telco Edge Reference Model

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. 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 Far Edges.

The benefit of this model is the ease and simplicity with which the entire Telco infrastructure can be administered. Instead of having to connect to each Edge 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 Far Edge sites 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 edge sites, both Telco Near and Far edge, are managed from a single management instance at a core site. This is depicted in following figure.

**Figure 1-3. Telco Edge Reference Model**




The two edge types are collapsed into a single Telco Edge that is managed from a Core data center. The components of a Telco edge representing the NFV infrastructure that includes compute, networking, and storage are depicted in the preceding diagram.

## Scope of this Reference Architecture

vCloud NFV Edge Reference Architecture considers both near and far edges as a single edge type and one or groups of these edge sites are managed remotely from the core site. We use the term "Region" to indicate a group of Edge Sites.

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 edge sites 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 at the 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 beforehand. It is important to ensure that a sufficient bandwidth at low latency is available for the Core site and Edge site connectivity.

# Architectural Framework and Components

2

The overall framework of the VMware vCloud NFV Telco Edge Architecture includes the key stakeholders, conceptual architecture environment, logical architecture, and components of the platform.

This chapter includes the following topics:


- Key Stakeholders
- Logical Architecture and Components
- NFV OpenStack Edition 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 of the infrastructure and the applications running on it.

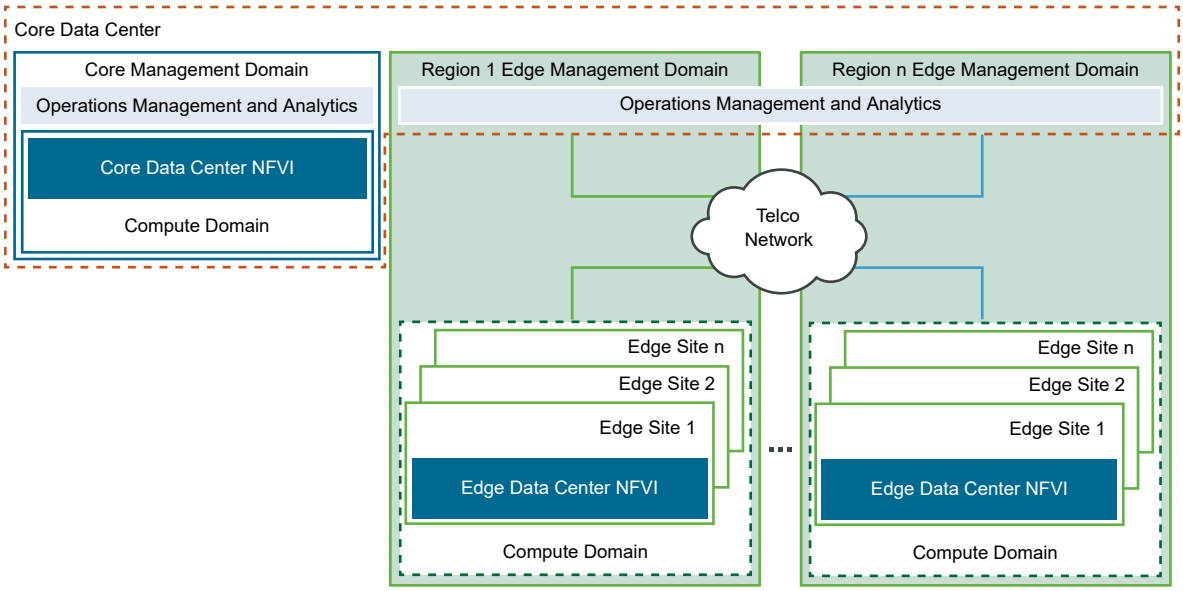
| | |
|---|---|
| **Cloud provider** | The CSP operations personnel who are responsible for provisioning and on-boarding all day 0 and day 1 functions to enable services for target customers and tenants. |
| **Consumer** | The end user who is consuming the services that the tenants provide. For example, 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** | A tenant represents a logical separation of the entity housed on and consuming infrastructure services to a provide a service. This may be a specific subset of customers for the Telco, an MVNO (Mobile Virtual Network Operator) which operates on top of the Telco infrastructure to provide their own virtual network, or a VNF vendor which provides a specific capability and so on. Each tenant is represented as a resource slice. Key resources that can be packaged in a resource slice are mobile network bandwidth, specific services enabled, number of subscribers, and so on. |

| | |
|---|---|
| **Operations Support** | The operations management process and team that ensure the services are operating to meet the promised stringent SLAs. |
| **System 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. |

# Logical Architecture and Components

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 2-1. Logical Architecture**



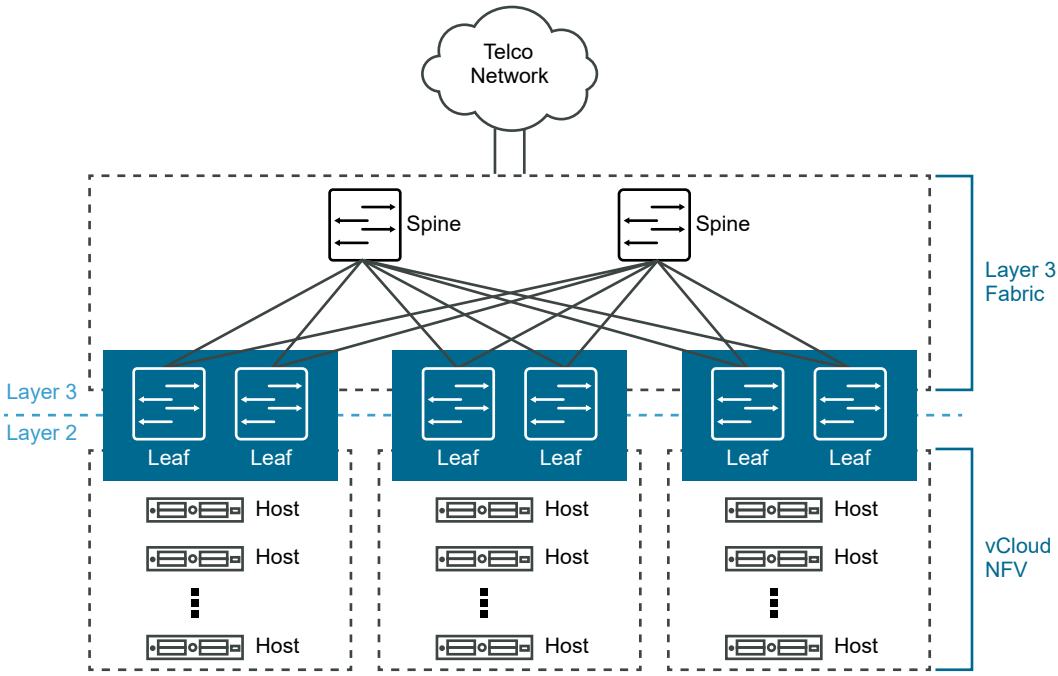

## Core Data Center

The Core data center is used to house the management components and other NFV functions that need to run at a central location. In the context of this reference architecture, a Core Data Center continues to run the core NFVi functions while also hosting the management components to manage the various telco edge sites.

### Physical Configuration

The Core site follows the three-pod design used in the vCloud NFV 3.x Reference Architecture and includes a minimum of three hosts (four recommended) in the management pod and the Edge pods. The resource pods comprise the hosts in the remote sites and any local clusters related to operations and functionality of the Edge sites.

**Figure 2-2. Core Data Center Physical Network Design**

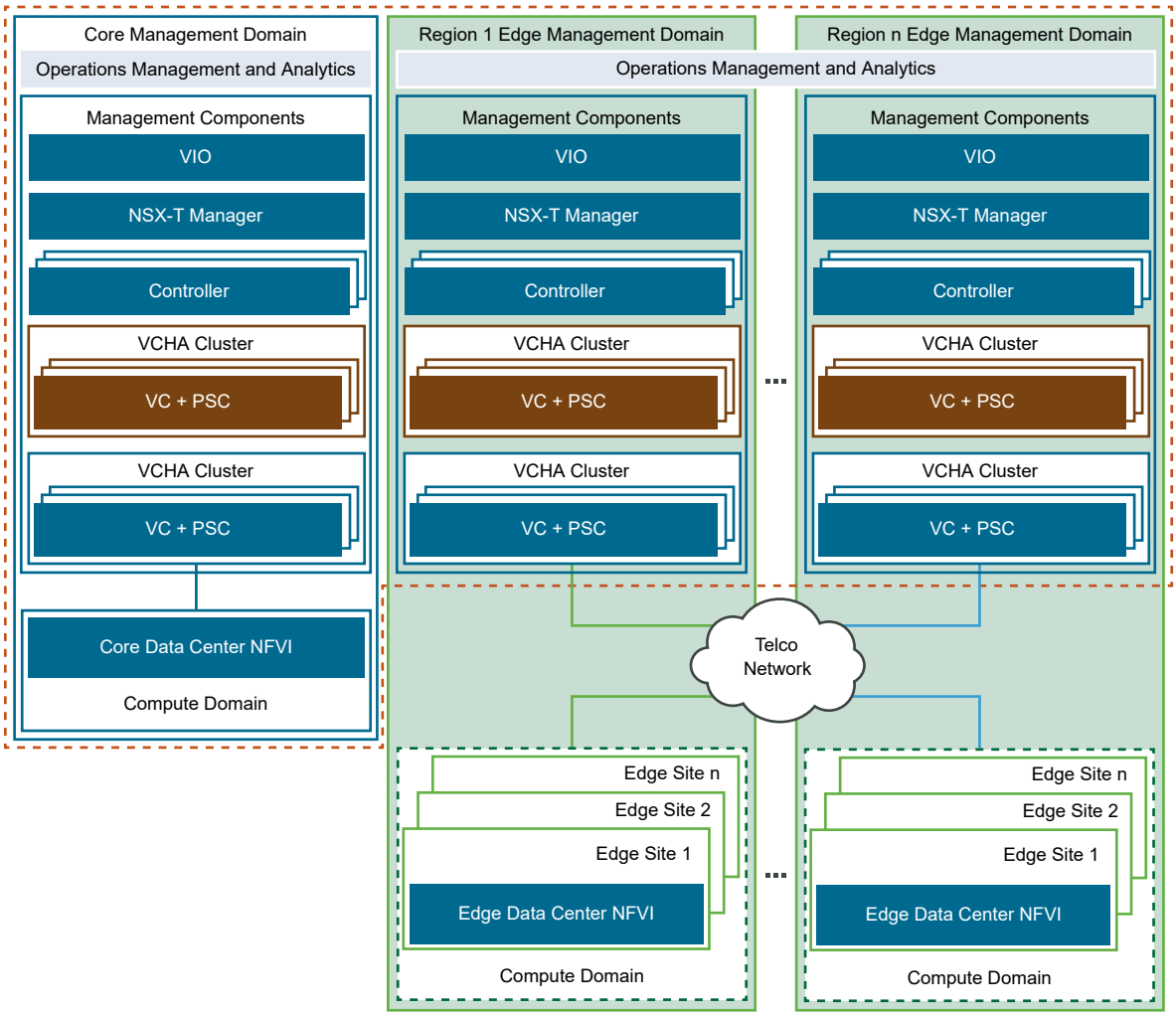



## Logical Configuration

This section describes the architecture for the Core data center to manage the Edge sites and to backhaul traffic from them.

The management domain comprises the Virtual Infrastructure Management (VIM), vCenter Server Appliance, NSX Manager, and NSX Controller. In addition to the management domain components, the Core data center also contains the operations management components. For more information of the operations management, see the Telco EdgeAnalytical Architecture section.

### Figure 2-3. Core Data Center Logical Components




In the preceding figure, the brown colored vCenter Server blocks are responsible for managing the components of the respective management domains. It is the entity used to create and spawn off the other management components in this cluster such as NSX Manager, NSX Controllers, and vRealize components. The left side of the diagram under Core VIM is infrastructure used to manage all resources in the Core data center. The right side of the diagram indicates the management components that are used to manage a group of edge sites within a region. The Region 1 Edge Management Domain manages a set of edge sites in Region 1 (shown connected through the Telco network) and Region 2 Edge Management Domain manages Region 2 edge sites and so on.

This Edge reference architecture consists of two sets of VIM components. One component is used to manage the VNF workloads deployed in the Core data center, while the other is used to manage the Edge sites within a region. Each edge site is a vSphere cluster (resource and edge pod) that can be scaled up as needed by adding ESXi hosts. The separation of Core data center and Edge data center resource functions into individually scalable pods allows CSPs to plan capacity based on the needs of the specific function hosted by each pod. This provides greater operational flexibility.

For best practices when using vSAN as the shared storage solution for the Core data center, initial deployment requires a minimum of four hosts per cluster. This sizing recommendation provides balance between the implementation footprint and resiliency, while maintaining the operational requirements necessary for the site.

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

The number of remote sites under a single Region can vary depending upon the size of the Edge sites and the number of local clusters in the resource pod. For detailed configuration maximums click here.

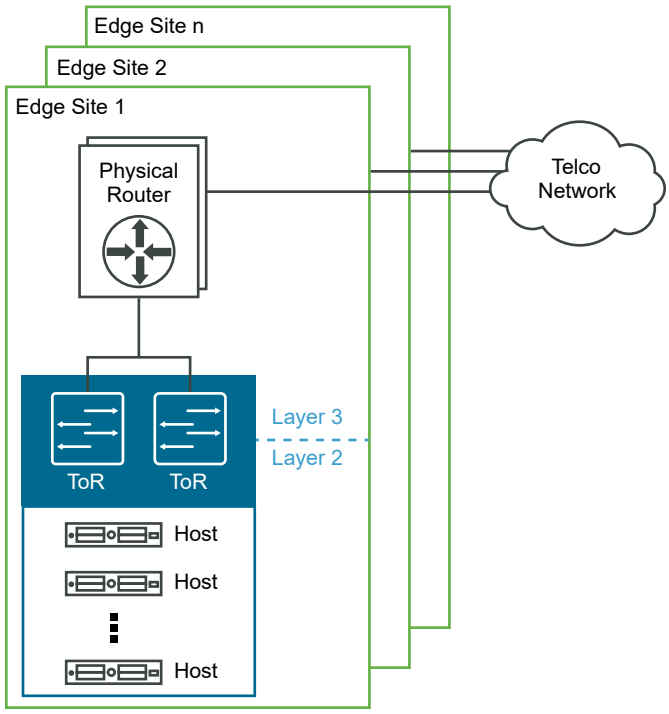
## Telco Edge

The Telco Edge is the site that houses the remote workloads (VNFs and applications). It consists of NSX Edge and VNF workloads in a compute domain. The compute domain maps to a vSphere cluster managed by the vCenter Server from the Core data center. The number of edge servers at a remote site depends on the workload to run. Two servers are 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 pod is hosted in the Core data center, the Edge site hosts the low footprint collapsed Edge/resource cluster. The Edge site includes a minimum of 3 hosts (4 recommended if using vSAN) that caters to both the Edge and resource workloads of the site.

**Figure 2-4. Telco Edge Physical Network Design**

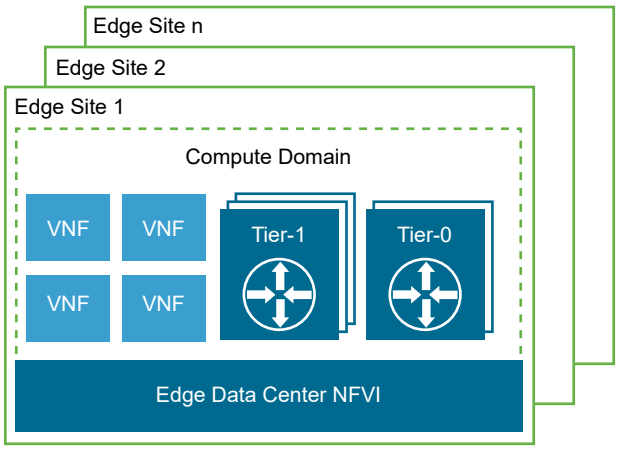



Physical network infrastructure is provided by a combination of ToR and inter-rack/spine switches. The latter are optional and depends on the server count and if all the server ports can be accommodated within a pair of ToR switches. Within a single site, intra-site networking is based on Ethernet (typically 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 2 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 2-5. 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 routers) and Tier-0 routers forward the traffic both east west and north south into and out of the infrastructure. The Tier-0 routers 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 two of the hosts in the Edge Site. Host anti-affinity rules are configured for these VMs to tolerate single host failures.

# NFV OpenStack Edition Components

The vCloud NFV Edge reference architecture is based on the vCloud NFV OpenStack Edition bundle. This package consists of essential building blocks to deploy an NFVI and VIM platform, featuring the newest releases of VMware production proven solutions.

Product version details can be found in the vCloud NFV OpenStack Edition under the VMware download page.

**Table 2-1. vCloud NFV OpenStack Edition Components**

| Component | Included in the vCloud NFV OpenStack Edition Bundle | Required for the Telco Edge Reference Architecture | Required for the Telco Edge Analytical Architecture |
|---|---|---|---|
| VMware ESXi™ | Yes | Yes | Yes |
| VMware vCenter® Server Appliance™ | No | Yes | Yes |
| VMware vSphere® Replication™ | Yes | Recommended | Recommended |
| VMware vSAN™ Standard Edition | Yes | Recommended | Recommended |
| VMware® Integrated OpenStack Carrier Edition | Yes | Yes | Yes |
| VMware NSX-T® Data Center | No | Yes | Yes |

| Component | Included in the vCloud NFV OpenStack Edition Bundle | Required for the Telco Edge Reference Architecture | Required for the Telco Edge Analytical Architecture |
|---|---|---|---|
| VMware Site Recovery Manager™ | No | Recommended | Recommended |
| VMware vRealize® Operations™ | Yes | No | Yes |
| VMware vRealize® Log Insight™ | Yes | No | Yes |
| VMware vRealize® Orchestrator | Yes | No | Yes |
| VMware vRealize® Network Insight™ | No | No | Recommended |

# Design Principles

The following design principles communicate a summary of the details that form the basis for the vCloud NFV Edge reference architecture using vCloud NFV components.

## Deployment Flexibility

Starting with a low management overhead at the Edge sites, the architecture provides flexibility for including a two vSphere cluster configuration (separate compute cluster and Edge cluster) at Edge sites requiring higher performance with NSX Edge Nodes. In cases where the number of compute nodes at the Edge site are limited, it possible to have a combined compute and Edge cluster where the NSX Edge Nodes run as VMs on the same compute cluster as the workloads

## Multi Tenancy and Advanced Networking

Within an Edge site, it is possible to house one or more tenants managed from the regional or core data center. This is realized using the multi-tenant capabilities of NSX-T (T1 and T0 routers) and the integration with VMware Integrated OpenStack (VIO).

## Workload Acceleration

To run user plane functions requiring lower latency and higher throughput at the Edge sites, the N-VDS (Enhanced) virtual switch can be configured on the hosts where these functions are to run. The UPF VMs can then be configured to run on top of the N-VDS (Enhanced) switch.

## Real-time Integrated Operational Intelligence

By using a framework that continuously collects data from local and distributed agents, vCloud NFV Edge Reference Architecture provides the capability to correlate, analyze, and enable day 2 operations. This analytics engine can be used with existing assurance engines for closed-loop. In addition, this analytics engine can be deployed in a distributed fashion in near-Edge sites towards smarter WAN bandwidth management and real-time closed-loop assurance.

## Orchestration Integration

vCloud NFV OpenStack Edition provides the flexibility to integrate the distributed management components, such as VIMs, with open-source-based Orchestrators such as ONAP, OSM, and third-party Orchestrators using industry standard interfaces.

# Telco Edge Reference Architecture

3

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.

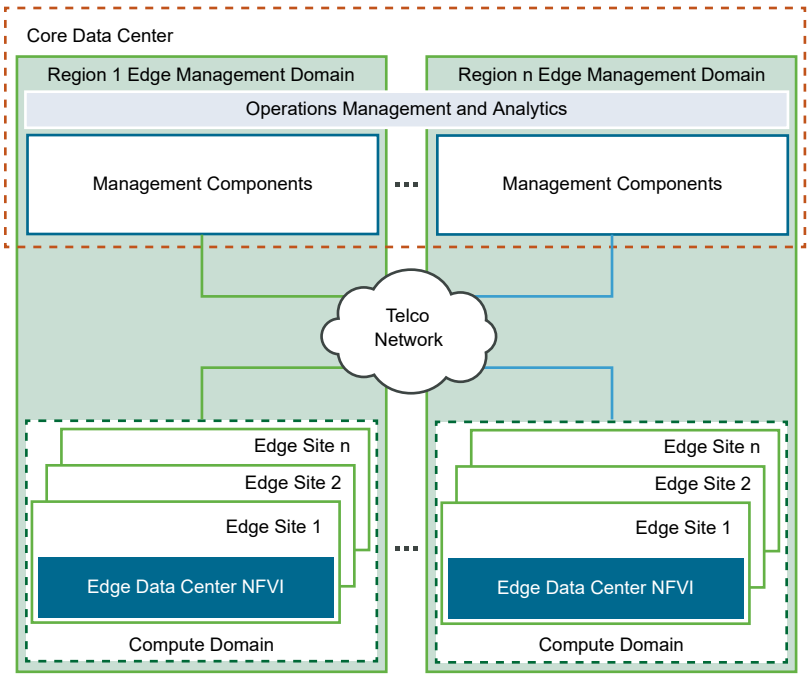
This chapter includes the following topics:


- Telco Edge Logical Building Blocks
- Telco Edge Virtual Building Blocks


## Telco Edge Logical Building Blocks

The platform components are grouped into two distinct containments: Edge Management Domain and Edge Resource 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.

**Figure 3-1. Telco Edge Logical Building Blocks**



The preceding figure shows the building blocks for the edge. An Edge data center is mapped to an edge site. Edge sites (or Edge data centers) are grouped into regions and they have a corresponding instance of VIO deployed in the Core data center. The operations management and analytics components oversee the edge management instances. This section describes the physical and logical configuration of the edge building blocks.

**Telco Edge Management Domain**

An Edge site follows the collapsed edge/resource pod design used in the vCloud NFV 3.1 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-T Data Center 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 must have local disks for vSAN caching and capacity tiers. An All Flash vSAN is recommended for reliability and performance.

Each physical host must have a minimum of six physical NICs connected to a pair of ToR switches in a LAG configuration for redundancy. The pairing of the physical NICs and their distribution across the virtual switches is covered in later sections. The ToR switches connect to an external WAN Edge physical router to transport packets for Internet breakout and backhaul to the Core site.

**Telco Edge Compute Domain**

Logically an edge site is a separate vSphere/vSAN cluster with hosts connecting to and managed by the vCenter Server that is a part of the management instance for the corresponding region. This edge cluster must host both the VNF workloads and the NSX-T Data Center edge nodes, and it must be sized accordingly.

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

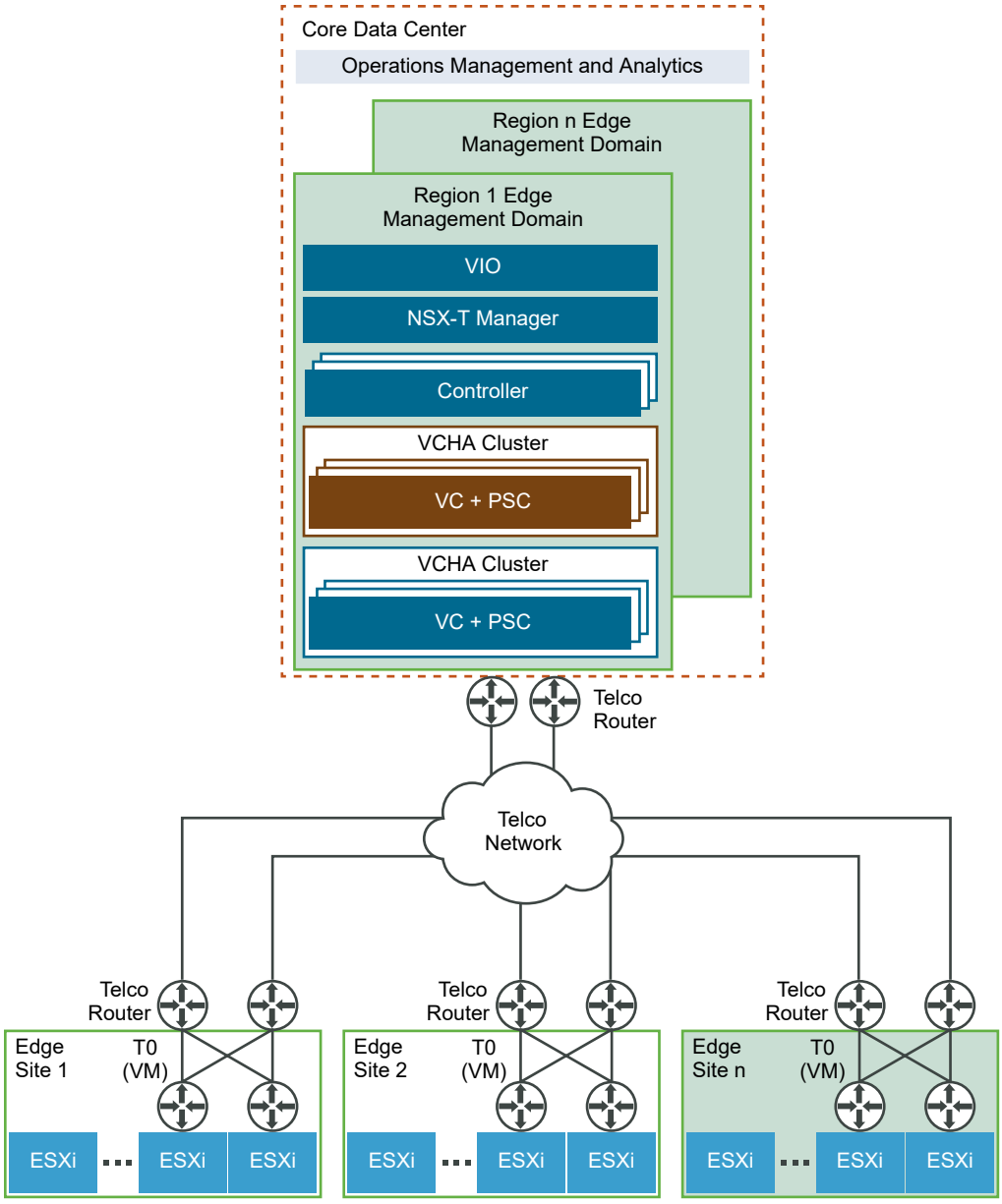
The actual sizing depends on factors influenced by the inter-dependencies between management components and their configuration limits. For example, a single NSX-T Data Center 2.3 installation can manage a maximum of 768 hosts as the limiting factor, whereas a single vCenter Server 6.7 can manage up to 2000 hosts. The actual number of supported sites vary 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 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 3-2. 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. In addition, 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 resides 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. In addition, a minimum bandwidth of 10Gbps is required between each edge site and the core site.

## Storage Design

This section outlines the building blocks for the virtual infrastructure shared storage design that is based on vSAN. vCloud NFV OpenStack Edition also supports certified third-party shared storage solutions, as listed in the VMware Compatibility Guide.

vSAN is a software feature built into the ESXi hypervisor 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 management and VNF workloads. With vSAN, VM data is stored as objects and components. One object consists of multiple components that 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.

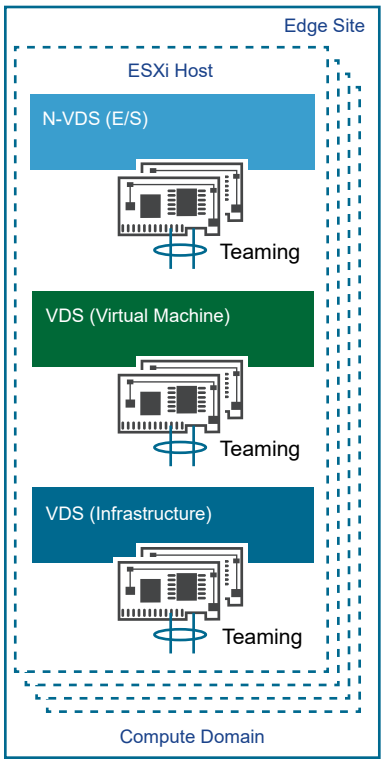
vSAN is a fully integrated hyper-converged storage software. Creating a cluster of server hard disk drives (HDDs) or solid-state drives (SSDs), vSAN presents a flash-optimized, highly resilient, shared storage datastore to ESXi hosts and virtual machines. This allows for the control of capacity, performance, and availability through storage policies, on a per VM basis.

## Network Design

The vCloud NFV Edge platform consists of infrastructure networks and VM networks. 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 cluster are configured with VMware vSphere® Distributed Switch™ (vDS) devices that provide consistent network configuration across multiple hosts. One vSphere Distributed Switch is used for VM networks and the other one maintains the infrastructure networks. Also, the N-VDS switch is used as the transport for Telco workload traffic.

### Figure 3-3. Virtual Network Design




Infrastructure networks are used by the ESXi hypervisor for vMotion, VMware vSphere Replication, vSAN traffic, management, and backup. The Virtual Machine networks are used by VMs to communicate with each other. For each cluster, the separation between infrastructure and VM networks ensures security and provides network resources where needed. This separation is implemented by two vSphere Distributed Switches, one for infrastructure networks and another for VM networks. Each distributed switch has separate uplink connectivity to the physical data center network, completely separating its traffic from other network traffic. The uplinks are mapped to a pair of physical NICs on each ESXi host for optimal performance and resiliency. In addition to the infrastructure networks, virtual machine network on the VDS is required for the NSX-T Edge for North-South traffic.

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 cluster. The infrastructure vSphere Distributed Switch and networks remain the same regardless of the cluster function. However, the VM networks depend on the networks that the specific cluster requires. The VM networks are created by NSX-T Data Center to provide enhanced networking services and performance to the workloads. The ESXi host's physical NICs are used as uplinks to connect the distributed switches to the physical network switches. All ESXi physical NICs connect to layer 2 or layer 3 managed switches on the physical network. For redundancy purposes, it is common to use two switches for connecting to the host physical NICs.

The infrastructure networks used in the edge site include:

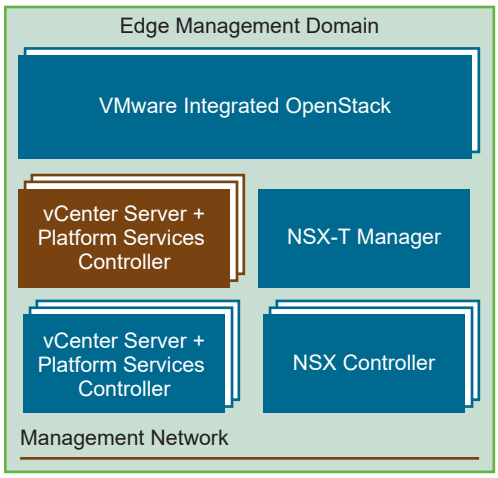
- ESXi Management Network. The network for the ESXi host management traffic.
- vMotion Network. The network for the VMware vSphere® vMotion® traffic.

- vSAN Network. The network for the vSAN shared storage traffic.

## Telco Edge Management Domain

Telco Edge Management Domain includes Virtualized Infrastructure Management (VIM) components such as vCenter Server Appliance, NSX Manager, and VMware Integrated OpenStack.

**Figure 3-4. Telco Edge Management Domain**




In addition to these components, the Management Domain also contains the operations management components. For more information, see the Chapter 5 Telco Edge Analytical Architecture section.

### Telco Edge Management Domain Components

The Management domain contains the components that manage the Edge Region runtime environment. This domain includes VMware Integrated OpenStack(VIM), vCenter Server, NSX Manager and its components.

#### vCenter Server

The Edge Management domain is implemented as a cluster that is managed by the brown vCenter Server instance shown in preceding figure. To form the foundation of a carrier grade virtualized infrastructure, the components of the Management Domain benefit from the cluster features such as resource management, high availability, and resiliency. A second vCenter Server is deployed in the Management Domain to oversee the Edge Compute Domain for the respective region.

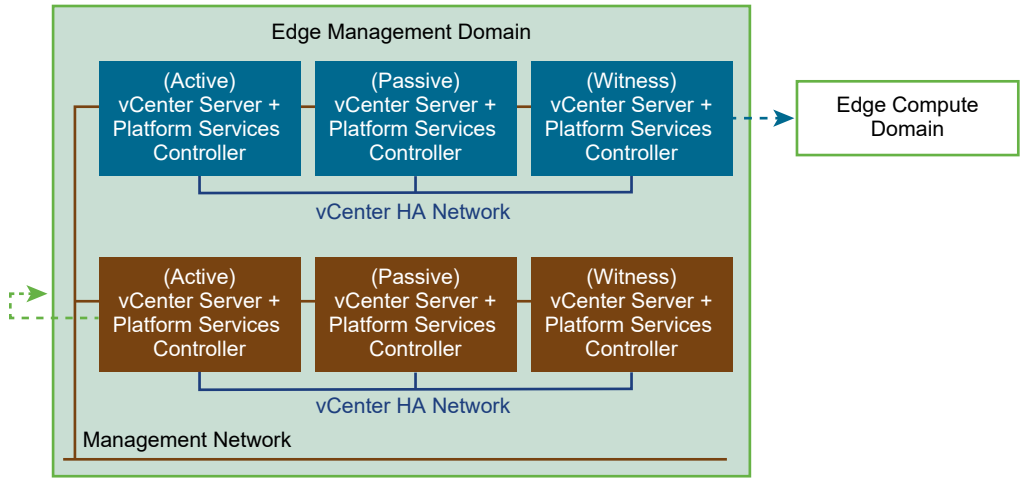
Each vCenter Server instance is a virtual appliance that is deployed with an embedded database. The vCenter® Server Appliance™ is preconfigured, hardened, and fast to deploy. The appliance allows for a simplified design, eases management, and reduces administrative efforts. vCenter Server Appliance availability is ensured by using a vCenter High Availability (vCenter HA) cluster, which is realized through three vCenter Server Appliance instances. The vCenter HA cluster consists of one active node that serves client requests, one passive node as a backup in the event of failure, and one quorum node that is called a witness node. Replication between nodes using a dedicated vCenter HA network ensures that vCenter Server Appliance data is always synchronized and up-to-date.

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 and its related services are 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 reducing the management components footprint.

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.

When vCenter HA is used with an embedded Platform Services Controller, the environment setup is as in the following figure.

**Figure 3-5. vCenter Server High Availability**



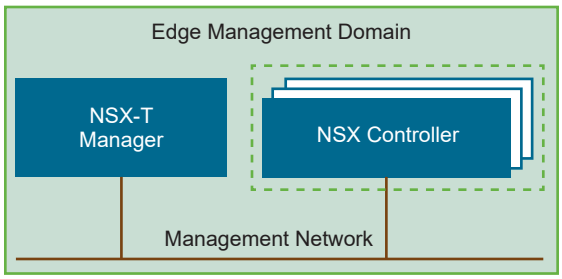

### VMware NSX-T Data Center

The NSX-T Data Center includes the NSX Manager and NSX Controller.

**NSX Manager** is the management plane for the NSX-T system. It provides the ability to create, configure, and monitor NSX-T Data Center components, such as logical switches, and NSX Edge Nodes. NSX Manager provides an aggregated system view and is the centralized network management component 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 logical switching and routing, networking services, edge services, security services, and distributed firewall capabilities.

NSX Manager is deployed as a single VM that uses vSphere HA for high availability. NSX Manager communicates with its controller and 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 NSX-T Data Center.

Figure 3-6. NSX Manager and Components



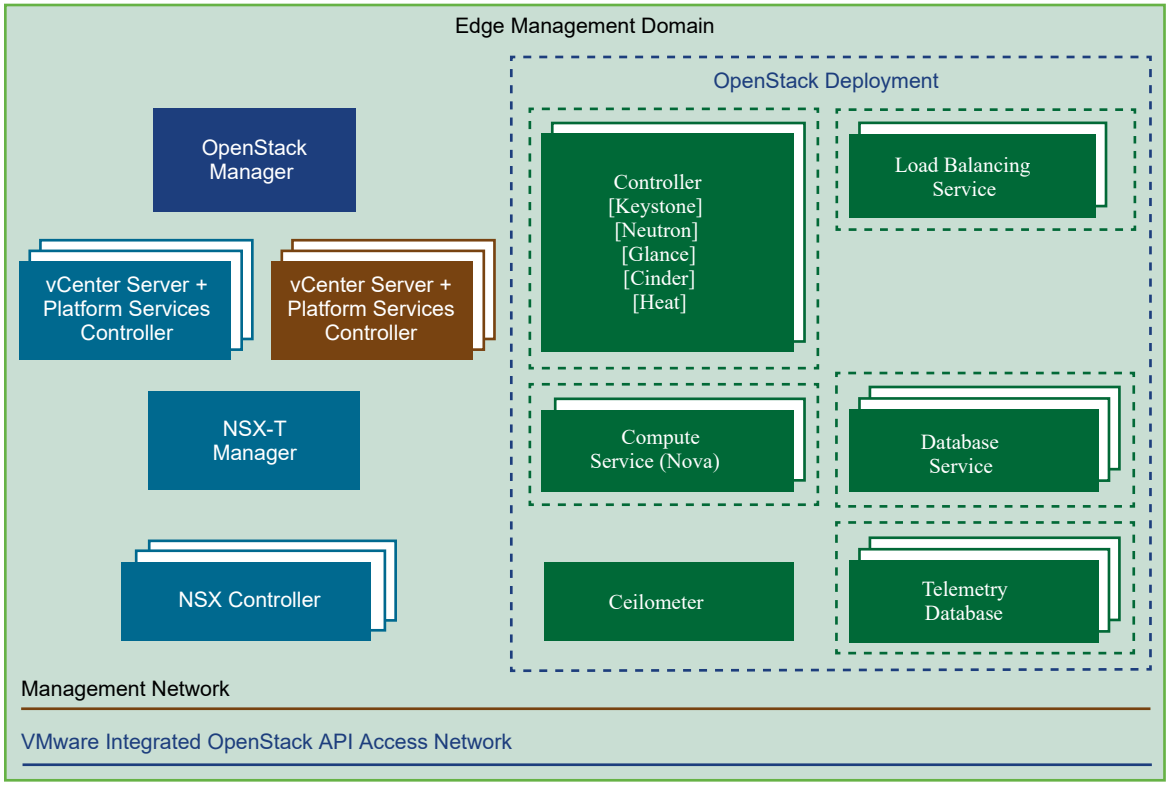

**NSX Controller** is an advanced distributed state management system that controls virtual networks and overlay transport tunnels. NSX Controller is deployed as a cluster of highly available virtual appliances that are responsible for the programmatic deployment of virtual networks across the entire NSX-T Data Center architecture. NSX Controller is responsible for providing configuration to other NSX Controller components such as the logical switches, logical routers, and edge configuration.

To enhance the high availability and scalability further, the NSX Controller is deployed in a cluster of three instances in the Edge cluster. Anti-affinity rules are configured to ensure that the controller instances reside on separate hosts to protect against host failures.

### VMware Integrated OpenStack

The VMware Integrated OpenStack Manager connects to the vCenter Server instance that manages the Management Domain. It uses a VM template to rapidly deploy, administer, and perform day 2 management operations of the VMware Integrated OpenStack management plane components that are deployed in the Management Domain. After deploying, VMware Integrated OpenStack connects to the vCenter Server instance that manages the Edge and Resource Domain. This vCenter Server instance is responsible for storage and compute resources. VMware Integrated OpenStack also connects to the NSX Manager instance that is associated with tenant networking.

### Figure 3-7. VMware Integrated OpenStack Management Components




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

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

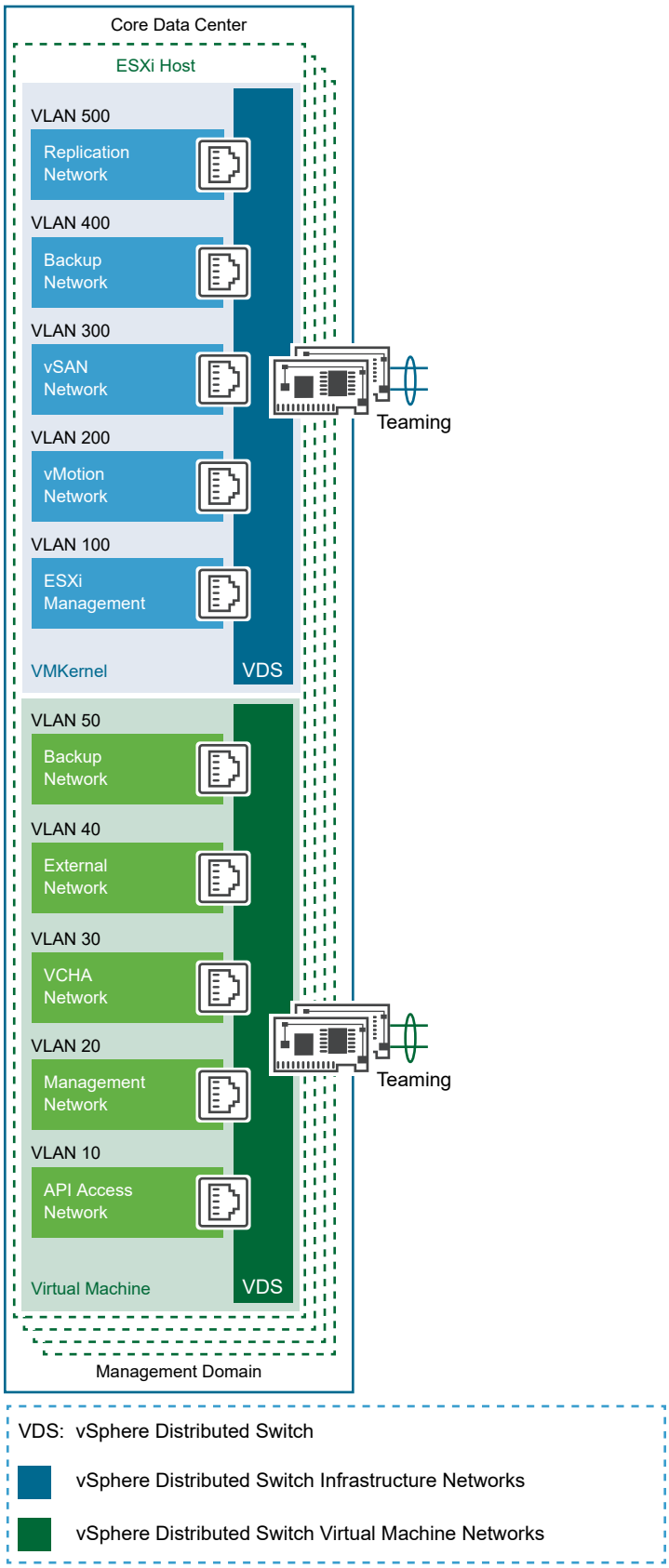
VMware Integrated OpenStack is closely integrated with NSX-T Data Center, providing tenants with enhanced features and capabilities for managing their VNF networking needs by using the Horizon interface and APIs. Network services include firewalling, network NAT, static and dynamic routing, and load balancing. Tenants can provision Geneve-backed logical switches for East-West VNF component connectivity and deploy NSX Edges for North-South traffic as required when connecting to other tenants or to external networks.

It is a best practice that each cluster is configured to use a shared storage solution. When hosts in a cluster use shared storage, manageability and agility improve.

## Telco Edge Management Domain Networking

The Telco Edge Management Domain networking consists of the infrastructure and Virtual Machine networks. The following diagram shows all the virtual switches and port groups of the Management Domain. The VLAN numbers here are for illustrative purposes only.

## Figure 3-8. Telco Edge Management Domain Networking



# Telco Edge Compute Domain

This section describes the components of the Edge Compute Domain and their functions. NSX Edge node and VNFs are placed in the Edge Compute Domain cluster that forms the virtual network services.

## Telco Edge Compute Domain Components

The Edge Compute Domain provides the runtime environment for the network functions. The section covers the logical tenancy and networking components.

### Projects

In VMware Integrated OpenStack, cloud administrators manage permissions through user, group, and project definitions. Projects in OpenStack are equal to tenants in vCloud NFV. A project is the administrative container where Telco workloads are deployed and managed.

### Tenant VDCs

A Tenant VDC allows creation of virtual data centers for tenants under different compute nodes that offer specific SLA levels for each Telco workload. While quotas on projects set limits on the OpenStack resources, Tenant VDCs allow providing resource guarantees for tenants and avoid noisy neighbor scenarios in a multitenant environment.

### VNFs

One or more VMs that are deployed in the tenancy to provide specific network functions or Telco services.

### NSX Edge

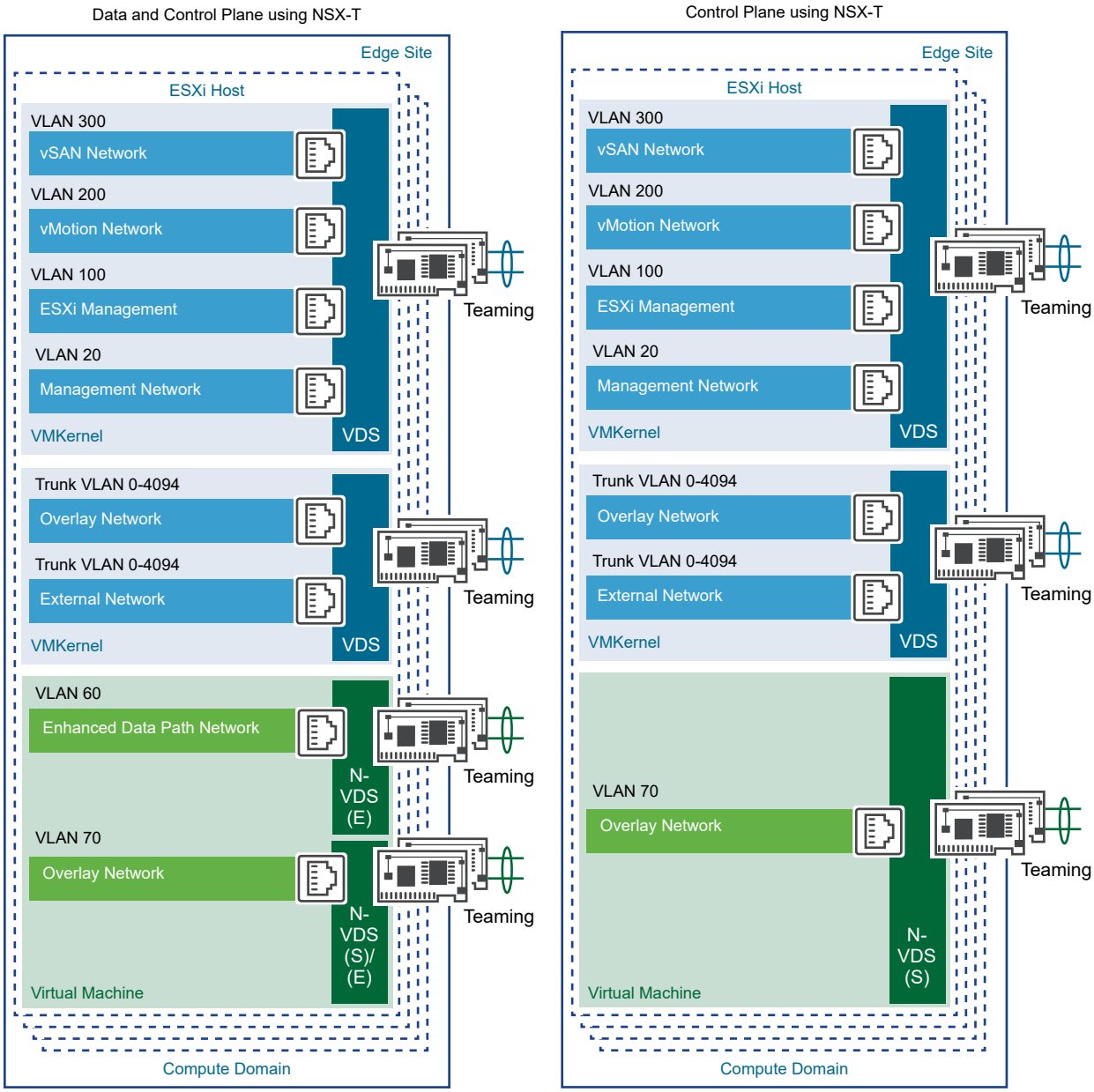
NSX Edge provides network edge security and gateway services to isolate a virtualized network. The NSX Edge logical (distributed) router provides East-West distributed routing with tenant IP address space and data path isolation. The NSX Edge gateway connects isolated, stub networks to shared (uplink) networks by providing common gateway services such as DHCP, VPN, NAT, dynamic routing, and Load Balancing.

## Telco Edge Compute Domain Networking

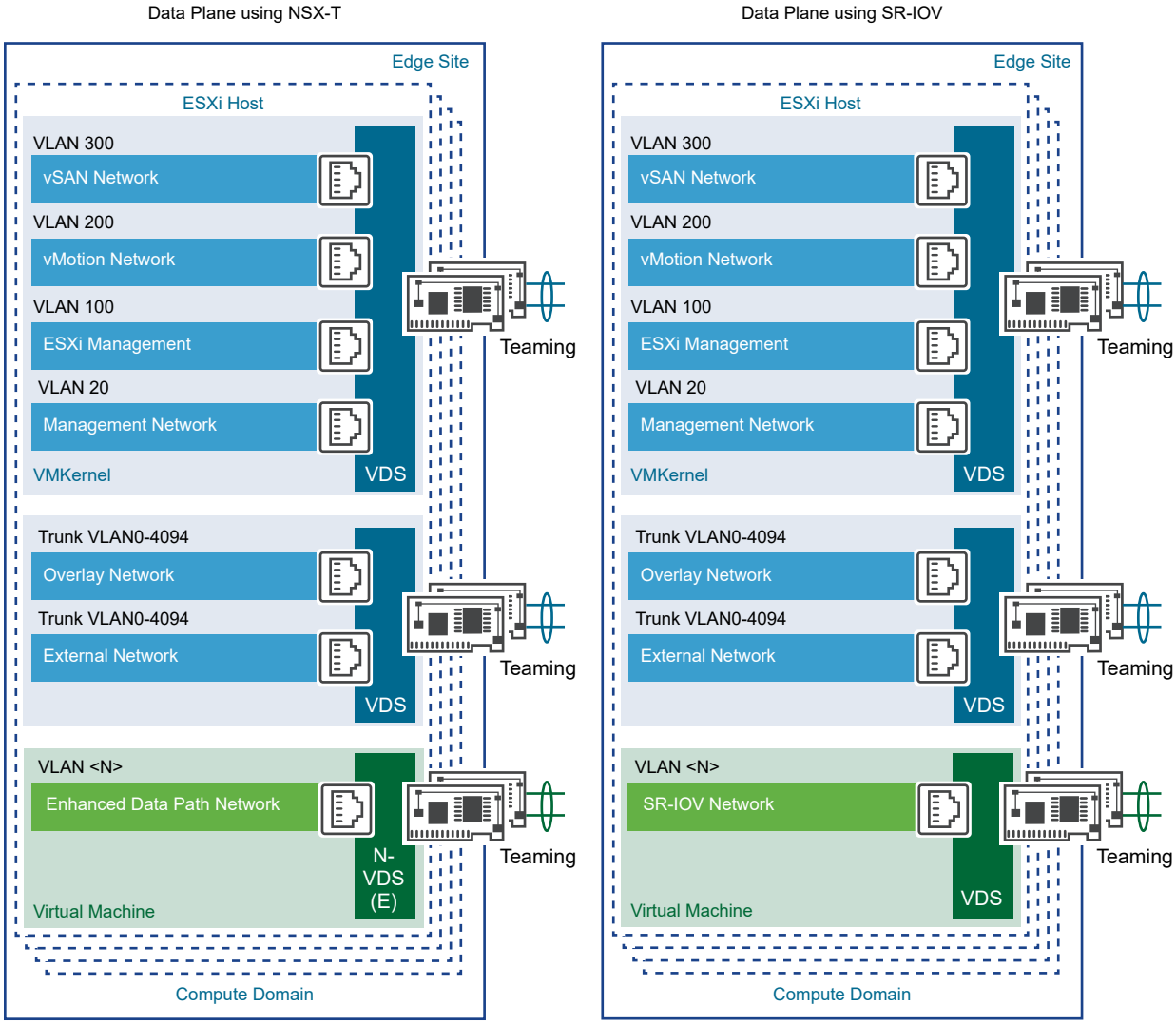
The networking of the Edge Compute Domain is highly dependent on the network topology that is required by the Telco workloads that are deployed by the tenant. This section describes the network building blocks as required by tenant workloads and is applicable to both VNF and other Telco workloads.

Following diagrams depict example scenarios and how networking components such as VDS, N-VDS, and SR-IOV can be used to provide network connectivity to the Telco workloads. The VLAN numbers in the following figure are for illustrative purposes only.

## Figure 3-9. Telco Edge Compute Domain Networking



### Figure 3-10. Telco Edge Compute Domain Networking




As shown in the preceding figure, Logical switches are the layer 2 networks created by NSX-T Data Center to provide connectivity between its services and the VMs. Logical switches form the basis of the tenant networks created by tenant administrators from with their tenancy. 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.

### 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 is where traffic remains 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 Internet or Core sitedoes not impact this reference architecture, except for certain latency and speed requirements.

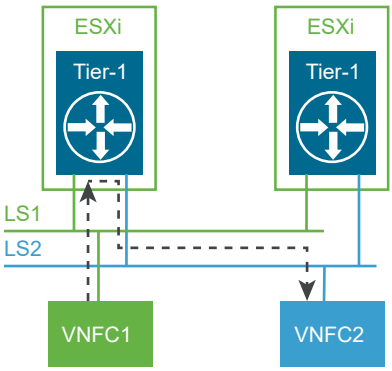
### Logical Routing

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 switches and routers as software logical constructs and embeds them in the hypervisor layer, abstracted from the underlying physical hardware.

### East-West Traffic

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

**Figure 3-11. East-West Traffic**

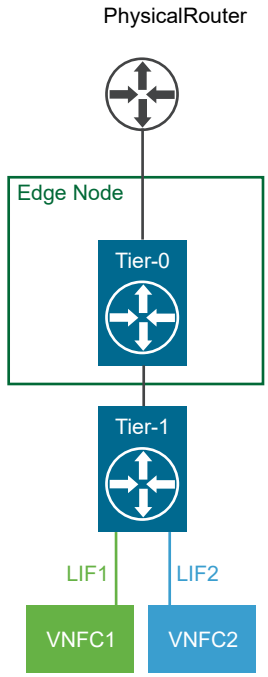



### North-South Traffic

In addition to providing optimized distributed and centralized routing functions, NSX-T Data Center supports a multi-tiered routing model with a 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 router, whereas the bottom-tier logical router is called a Tier-1 router. This structure provides both provider and tenant administrators a complete control over their services and

policies. The provider administrator controls and configures Tier-0 routing and services and the tenant administrators control and configure Tier-1 routing services. Northbound, the Tier-0 logical router 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 logical router connects to one or more Tier-1 logical routers.

**Figure 3-12. North-South Traffic**



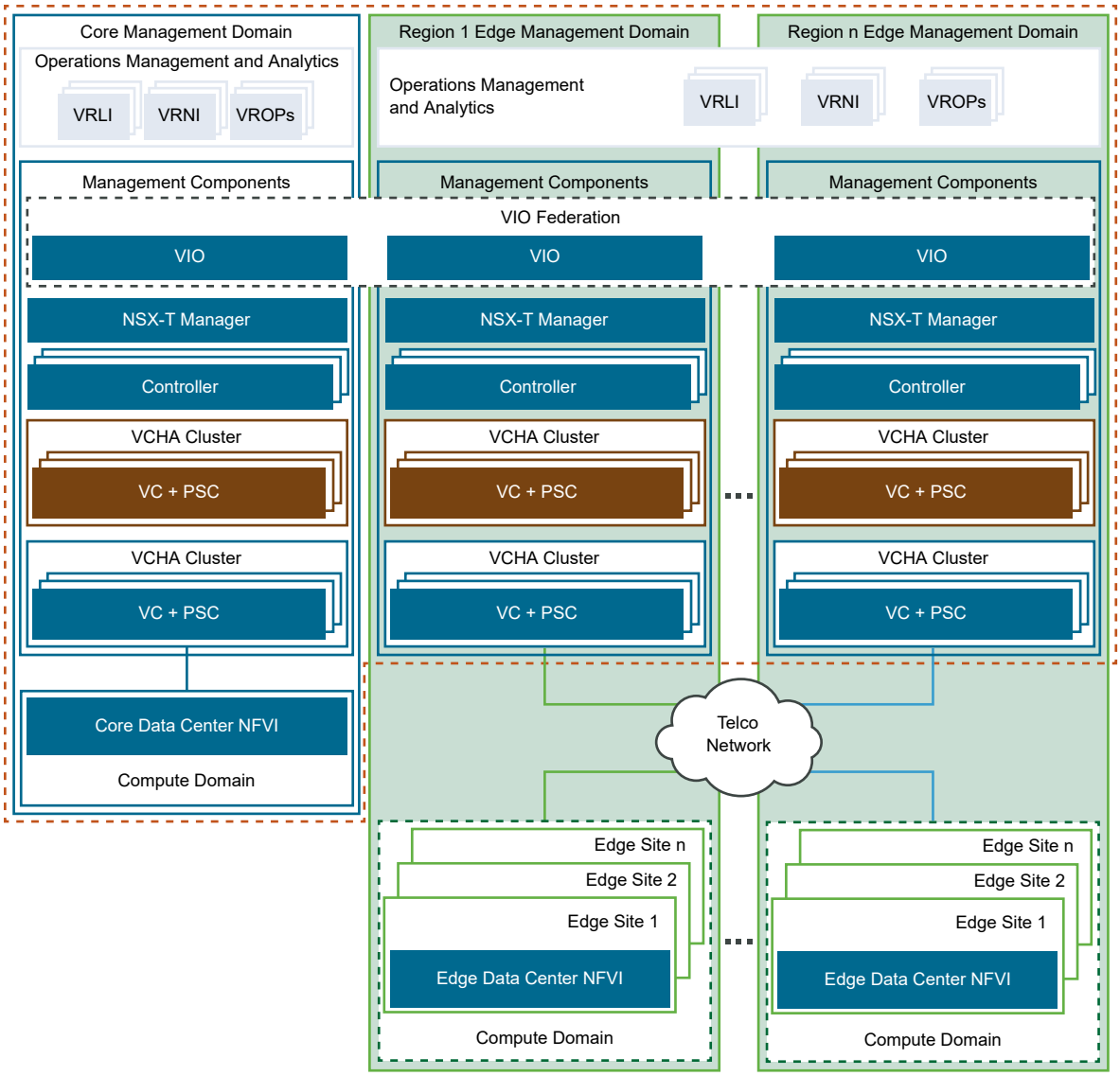

This model also eliminates the dependency on a physical infrastructure administrator to configure or change anything on the physical infrastructure when a new tenant is configured in the data center. For a new tenant, the Tier-0 logical router simply advertises the new tenant routes that are learned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

# Telco Edge Deployment

4

This Reference Architecture proposes a centralized model where the VIM, Management Components, Operations Management and Analytics components are all hosted in the central site. The Edge sites comprise of only compute resources.

**Figure 4-1. Telco Edge Deployment Conceptual Design**




**Note** Management components for each region are identical and follow the architecture described in the vCloud NFV OSE 3.1 Reference Architecture.

This chapter includes the following topics:


- Telco Edge Deployment Network Design
- Design Considerations

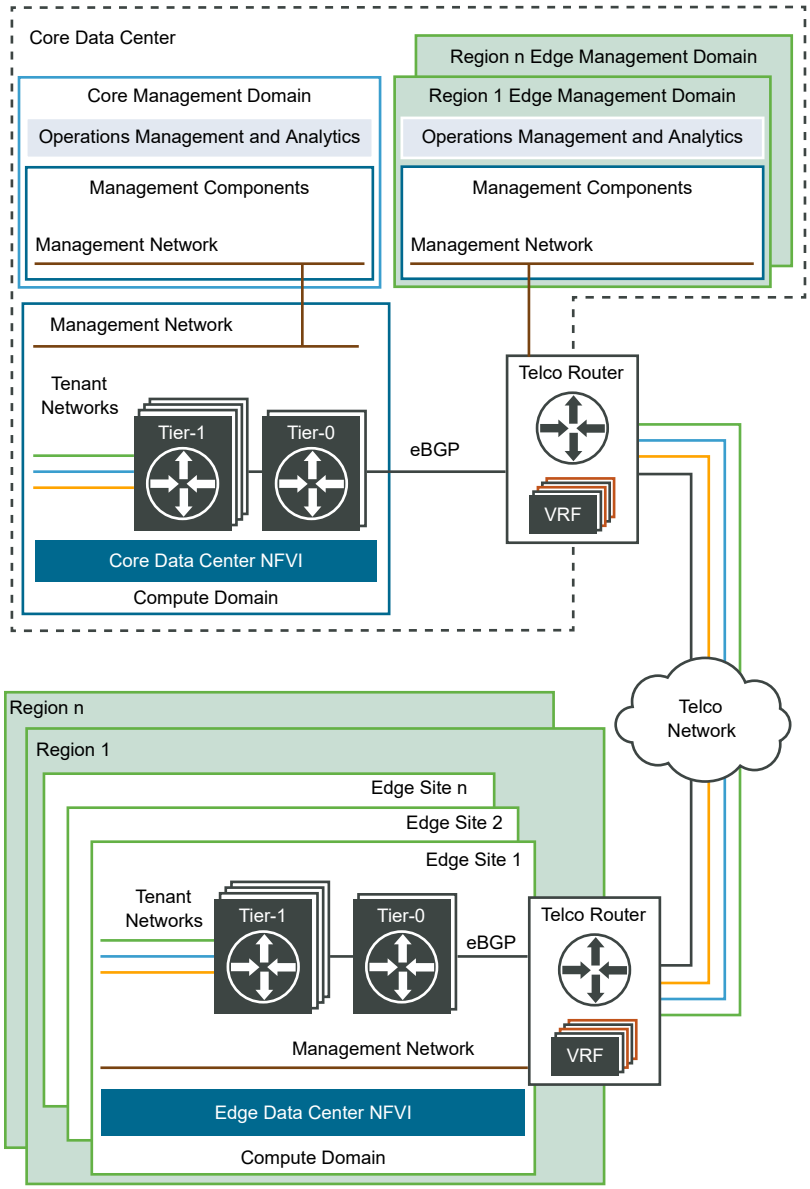

# Telco Edge Deployment Network Design

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 helps 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 L2/L3 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 4-2. Telco Edge Deployment Network Design




# Design Considerations

This reference architecture assumes a separate network connection over Layer 3 for management connectivity between the VIM components and its edge sites for traffic such as that between vCenter Server and Edge site ESXi hosts. NSX-T Manager also uses this network for management of NSX-T 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 physical network function and also to assist in mapping of tenant routers when a multi-tenant environment is needed. Note that segregation 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 Edge and core sites is 10 Gbps.

For segmentation of the individual traffic types, such as specific tenant traffic or management traffic, it is best to connect each Edge network (traffic) segment to a specific VRF (Virtual Routing & Forwarding instance) at each PE router. A similar theme of per VRF traffic forwarding is followed at the core site for the core to Edge traffic.

VLAN-based network segmentation is restricted to within a data center. There is no VLAN stretching between core and Edge sites.

## Network Redundancy

The vCloud NFV Edge reference architecture configuration has two Edge nodes (in VM form factor) in active/standby mode to connect to the Provider Edge (PE) router at the Edge site. To define the high availability configuration, the administrator from the Core data center vCenter Server must use the control plane network VRF to configure HA appliance name with the mode as primary on that vCenter cluster.

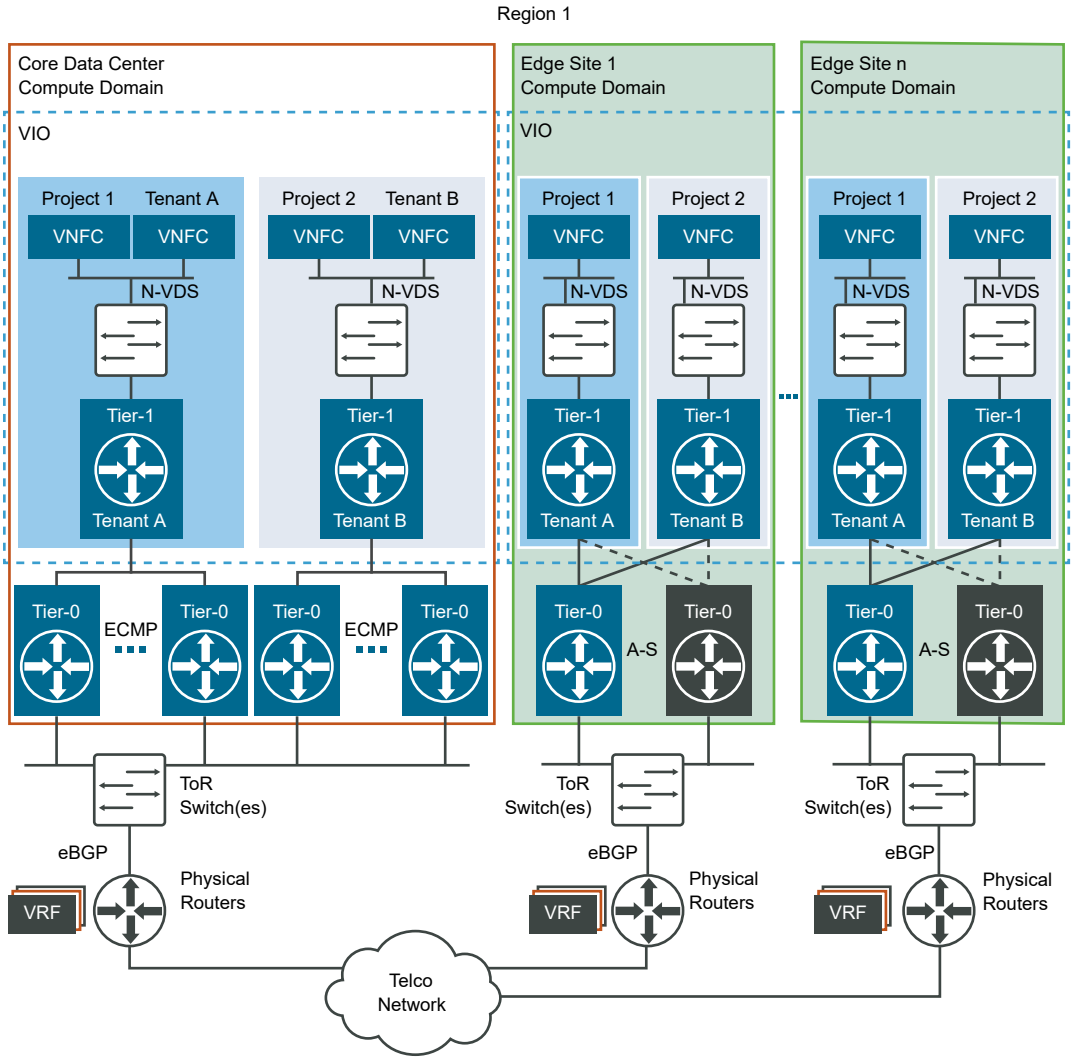
## Operations Management

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 FCAPs 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 routers in the networking stack.

### Figure 4-3. Telco Edge Reference Architecture Multi-tenancy Networking




The uplink of a Tier-0 router that resides in NSX-T Edge is connected to upstream physical routers. A tenant uses a Tier-1 router at its Edge to connect to the Tier-0 router. Tier-0 router will relay traffic to other tenants residing at the Core data center through a VRF 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.

An example of how multi-tier end-to-end routing can be applied is shown in the preceding figure. In this example, the service provider presents a WAN access by connecting one end of the Tier-0 router to upstream CE/PE router while the other end of the router connects to the customer’s Tier-1 tenant, terminating on its Edge Services Gateway.

# Telco Edge Analytical Architecture

5

CSPs can enable the vCloud NFV Edge Reference Architecture 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.

1 Onboard service operations.

2 Service launch and monitoring.

3 Dynamic optimizations.

4 Issue isolation.

5 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.

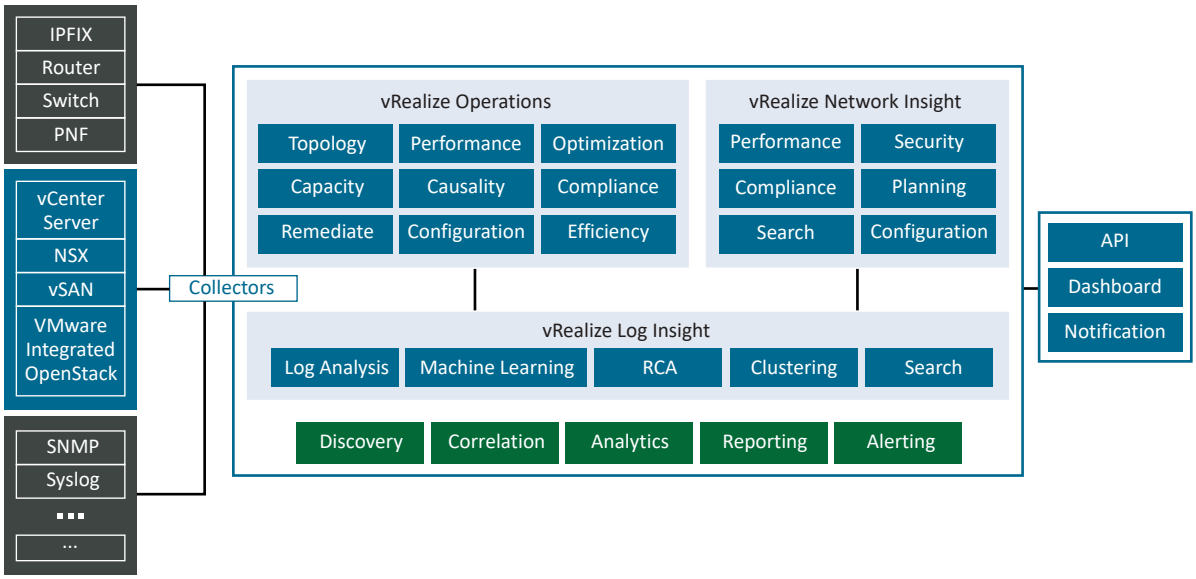
- Ability to profile and monitor traffic segments, types, and destination to recommend security rules and policies for north-south and east-west traffic.
- Identification of security policy and configuration violations, performance impacts, and traffic routes.

**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, NMS, EMS, OSS/BSS, and VNFM and NFVO for closed loop remediation.

**Figure 5-1. Analytical Reference Architecture**




CSPs can deploy the operations management components in the Telco Edge Management Domain 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 and 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 the network traffic optimization, micro-segmentation, compliance, security violations, traffic routing, and performance.

This chapter includes the following topics:


- Introduction to Analytics for Telco Edge
- Analytic Components for Edge

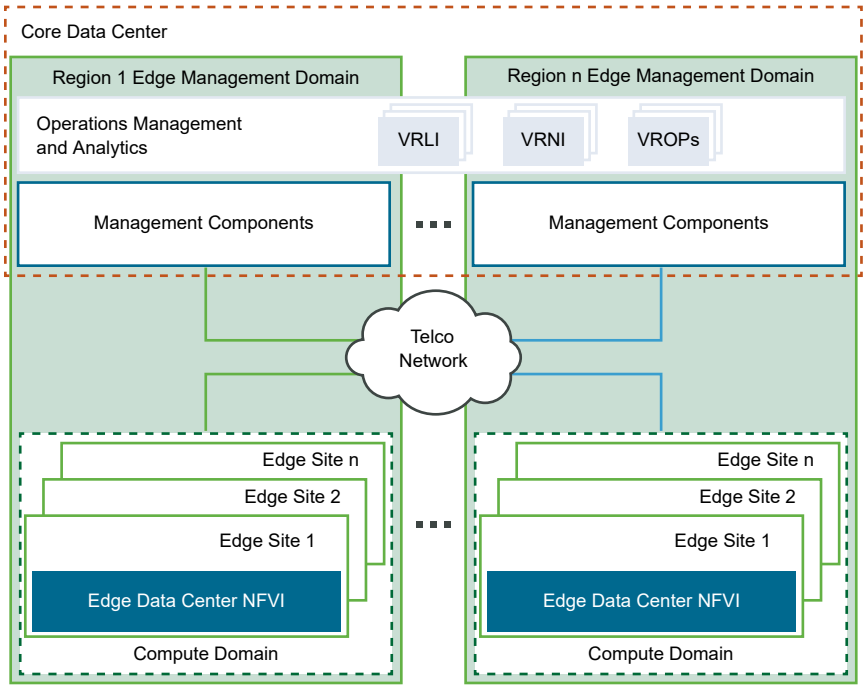

# Introduction to Analytics for Telco Edge

This Edge Reference Architecture considers two instances of Operations Management and Analytics in the Central Core data center, one associated with the Core Management Domain and another with all the Edge Management Domains.

## Telco Edge Analytics Logical Architecture

Each of the Operations Management and Analytics instances has one instance each of vROPs, vRLI, and vRNI. For very large-scale edge deployments, more Operations Management and Analytics instances may be needed.

**Figure 5-2. Telco Edge Analytics Logical Architecture**



With the Operations Management and Analytics components being placed in the central Core data center, it is important to limit the distance between the central site and the edge site to ensure that the latency is below 150 ms RTT. For large-scale deployments, remote collector components of vRNI, vRLI, or vROps may be needed to address collection end-point scalability and to facilitate bulk data export.

# Analytic Components for Edge

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

vRealize Operations Manager is configured with adapters for vCenter Server, End-Point Operations Management, vSAN, OpenStack and NSX. These adapters are necessary to start collecting data from the infrastructure.

For large scale deployments of vRealize Operations Manager, a remote collector is an optional component that allows vRealize Operations Manager to perform monitoring at scale. The remote collector is typically deployed in the Edge site(s) to reduce the frequency of updates from Core data center to the Edge sites, and/or reduce the load on the vRealize Operations Manager analytics cluster.

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. See the vRealize Operations Manager Sizing Guidelines.

## vRealize Log Insight

VMware vRealize Log Insight delivers real-time, heterogeneous and highly scalable log management with intuitive, actionable dashboards, sophisticated analytics and broad third-party extensibility, providing deep operational visibility and faster troubleshooting. The key components include built-in Syslog server and Log Insight agent which gathers log events natively from multiple syslog data sources through special management packs and intelligently group.

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 later. For sizing details, see the VMware 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.

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 using various protocols such as HTTPS, SSH, CLI, and SNMP, depending on the source and the configuration. A variety of data sources are supported, including 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 will be collected, typically CSPs need one or more proxy VMs.

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

# Architectural Realization

6

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

This chapter includes the following topics:


- Multi Tenancy
- Telco Edge Workload Placement
- Availability and Disaster Recovery


## Multi Tenancy

Multitenancy defines the isolation of resources and networks to deliver applications with quality. Because multiple tenants share the same resource infrastructure, secure multitenancy can be enabled by using VMware Integrated OpenStack 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 like vSphere DRS.

CSPs can converge their resource infrastructures across their IT and Network clouds, enabling a multi-tenant IaaS realization. Consumption models can serve both internal and external tenants over the common shared infrastructure to deploy and operate their respective workloads and services.

A unit of tenancy within the scope of a VIO deployment is a Project. It is defined as a composition of dedicated compute, storage, and network resources and as workloads. The tenant is associated with a set of operational policies and SLAs. The Project can be bound to a single site or shared across many sites.

Using the Tenant Virtual Data Center available with the VMware Integrated OpenStack Carrier Edition, a CSP can create virtual datacenters for tenants under different compute nodes that offer specific service level agreements for each telecommunication workload. By using the Tenant Virtual Data Center to allocate CPU and memory for an OpenStack project or tenant on a compute node, CSPs provide a resource guarantee for tenants and avoid noisy neighbor scenarios in a multi-tenant environment.

The design objectives include considerations for the end-to-end compute and network resource isolation from the Core data center to the Edge data centers.

## Management Plane

The management plane functions reside in the Edge Management Domain at the Core data center. They are responsible for the orchestration of resources and operations. The management plane functions are local to each cloud instance providing the 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 Integrated OpenStack. Irrespective of the Domain deployment configuration, VMware Integrated OpenStack 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 multiple administrative ownerships. A cloud provider, that is the CSP admin, can create a resource pool allocation for a tenant who in turn can manage the underlying infrastructure and overlay networking. In VMware Integrated OpenStack, multiple tenants can be defined with assigned RBAC privileges to manage compute and network resources and also the VNF onboarding.

## Compute Isolation

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.

Tenant vDCs provide a secure multitenant environment to deploy VNFs. Compute resources are defined as resource pools when a Tenant vDC is created. The resource pool is an allocation of memory and CPU from the available shared infrastructure, assignable to a Tenant vDC. More resources can be added to a pool as required. The Tenant vDC can also stretch across multiple hosts in a resource cluster residing in different physical racks. The same constructs for resource management are used to implement multitenancy in the Edge data centers.

Though the resource pools can be further sub-segmented into smaller resource pools, this is not a recommended design.

## Network Isolation

The advanced networking model of NSX-T Data Center provides a fully-isolated and secure traffic paths across workloads and tenant switch or routing fabric. Advanced security policies and rules can be applied at the VM boundary to further control unwarranted traffic.

NSX-T Data Center introduces a two-tiered routing architecture which 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 context 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 and communication to remote Edge data centers.

Each Tenant vDC has a single Tier-1 distributed router that provides the intra-tenant routing capabilities. It can also be enabled for stateful services such as firewall, NAT, load balancer, and so on. VMs belonging to Tenant A can be plumbed to multiple logical interfaces for layer 2 and layer 3 connectivity.

By using VMware Integrated OpenStack as the IaaS layer, user profile and RBAC policies can be used to enable and restrict access to the networking fabric at the Tier-1 level.

## QoS Resource Allocation

To avoid contention and starvation, QoS policies along with compute, storage, and network isolation policies must be applied consistently to the workloads.

The CSP admin can allocate and reserve compute resources for tenants by using Tenant vDCs. Every Tenant vDC is associated with a vSphere resource pool the resource settings of which are managed from VMware Integrated OpenStack. This ensures that every Tenant vDC consumes resources to which it is entitled, without exceeding the infrastructure resource limits, such as CPU clock cycles and total memory.

QoS policies can be applied to VMs so that they receive a fair share of resources across the infrastructure pool. Each VM configuration is taken from a template that is called a Flavor. QoS can be configured by using Flavor metadata to allocate CPU (MHz), memory (MB), storage (IOPS), and virtual interfaces (Mbps).

QoS can be shaped by setting boundary parameters that control the elasticity and priority of the resources that are assigned to the VNF component executing within the VM.

`Reservation` is the minimum guarantee. Reservations ensure a minimum guarantee to each VM when it is launched.

`Limit` is the upper boundary. Should be used with caution in a production environment, because it restricts the VM from bursting utilization beyond the configured boundaries.

`Shares` are 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 a resource contention, the VM with the higher shares gets the proportional resource assignment.

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

# Telco Edge Workload Placement

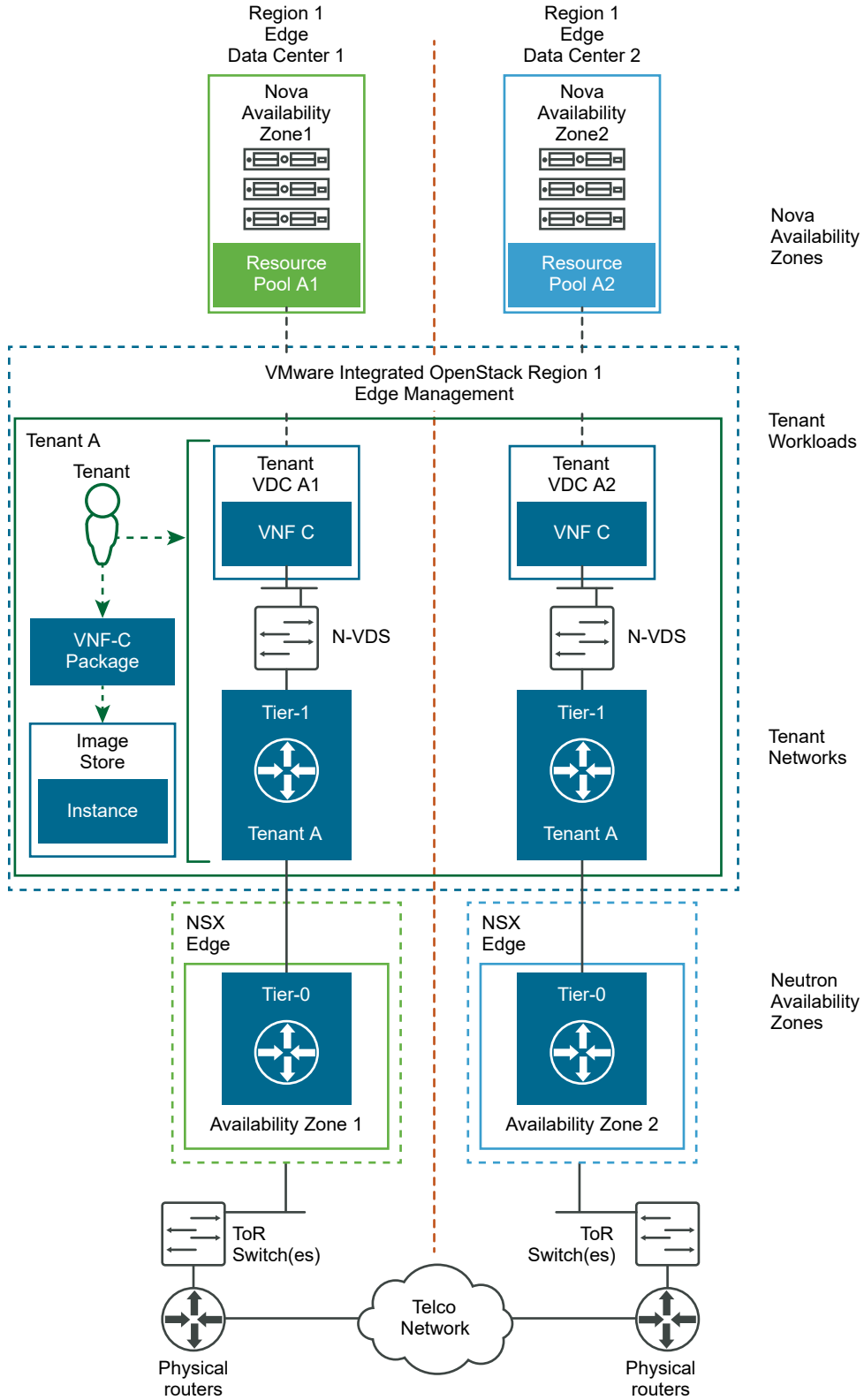
The placement of workloads is typically a subset of the VNF onboarding process and is a collaborative effort between the VNF vendor and the CSP. A prescriptive set of steps must be followed to package and deploy a VNF. The VMware Ready for NFV program is a good vehicle to pre-certify the VNFs and the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

## VNF Onboarding with VMware Integrated OpenStack

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

## Figure 6-1. VNF Onboarding Conceptual Design



## Resource Allocation

When building a project for the VNF, the administrator must set the initial quota limits for the project. For fine-grained resource allocation and control, the quota of the resources that are available to a project can be further divided using Tenant vDCs. A Tenant vDC provides resource isolation and guaranteed resource availability for each tenant. Quotas are the operational limits that configure the amount of system resources that are available per project. Quotas can be enforced at a project and user level. When a user logs in to a project, they see an overview of the project including the resources that are provided for them, the resources they have consumed, and the remaining resources.

Resource allocation at the Edge data centers leverage the same mechanism as in the core data center. Tenant VDCs are created at the Edge data center to carve resources for the tenant Edge workloads and assigned to the respective project for the tenant, ensuring consistent level of fine-grained resource allocation and control across the infrastructure.

## VNF Networking

Based on specific VNF networking requirements, a tenant can provision East-West connectivity, security groups, firewalls, micro-segmentation, NAT, and LBaaS using the VMware Integrated OpenStack user interface or command line. VNF North-South connectivity is established by connecting tenant networks to external networks through NSX-T Data Center Tier-0 routers that are deployed in Edge nodes. External networks are created by administrators and a variety of VNF routing scenarios are possible.

After the VNFs are deployed, their routing, switching, and security policies must be configured. There are many different infrastructure services available that can be configured in different ways. This document discusses few of these options in the later sections.

Tenant networks are accessible by all Tenant vDCs within the project in the same data center. Therefore, the implementation of East-West connectivity between VNF-Cs in the same Tenant vDC in the Core data center and the connectivity between VNF-Cs in Tenant vDCs in the Edge data center belonging to the same project is identical. Tenant networks are implemented as logical switches within the project. The North-South network is a tenant network that is connected to the telecommunications network through an N-VDS Enhanced for data-intensive workloads or by using NVDS standard through an NSX Edge Cluster. VNFs utilize the North-South network of their data center when connectivity to remote data centers is required, such as when workloads in the Core data center need to communicate with those in the Edge data center and vice-versa.

VMware Integrated OpenStack exposes a rich set of API calls to provide automation. The deployment of VNFs can be automated by using a Heat template. With API calls, the upstream VNF-M and NFVO can automate all aspects of the VNF life cycle.

## VNF Onboarding

The VNF workload onboarding introduced in OSE 3.1 is still valid in this Edge reference architecture, with the exception that compute nodes are now remotely located at an edge site. All the functional VNF workload onboarding principles for VIO remain the same.

Edge sites contain VNFs that are run on VIO tenant networks that can potentially connect to other tenant networks at the Core data center where other VNF components may be deployed. The VIO instances themselves are deployed only at the Core data center and are inline with vCenter Server and its associated FCAPS (vROPs, VRNI, and vRLI). From a virtual infrastructure management perspective, vCenter Server in the Core data center is the only component talking to the compute nodes in the Edge data center for workload placement and state management. The VIO instances in the Core data center relay virtual infrastructure requests to their respective vCenter Server and NSX-T Data Center VIM components through the use of nova and neutron plugins respectively. These VIM components then carry out the tasks on behalf of VIO, such as provisioning tenant networks and workload resource management.

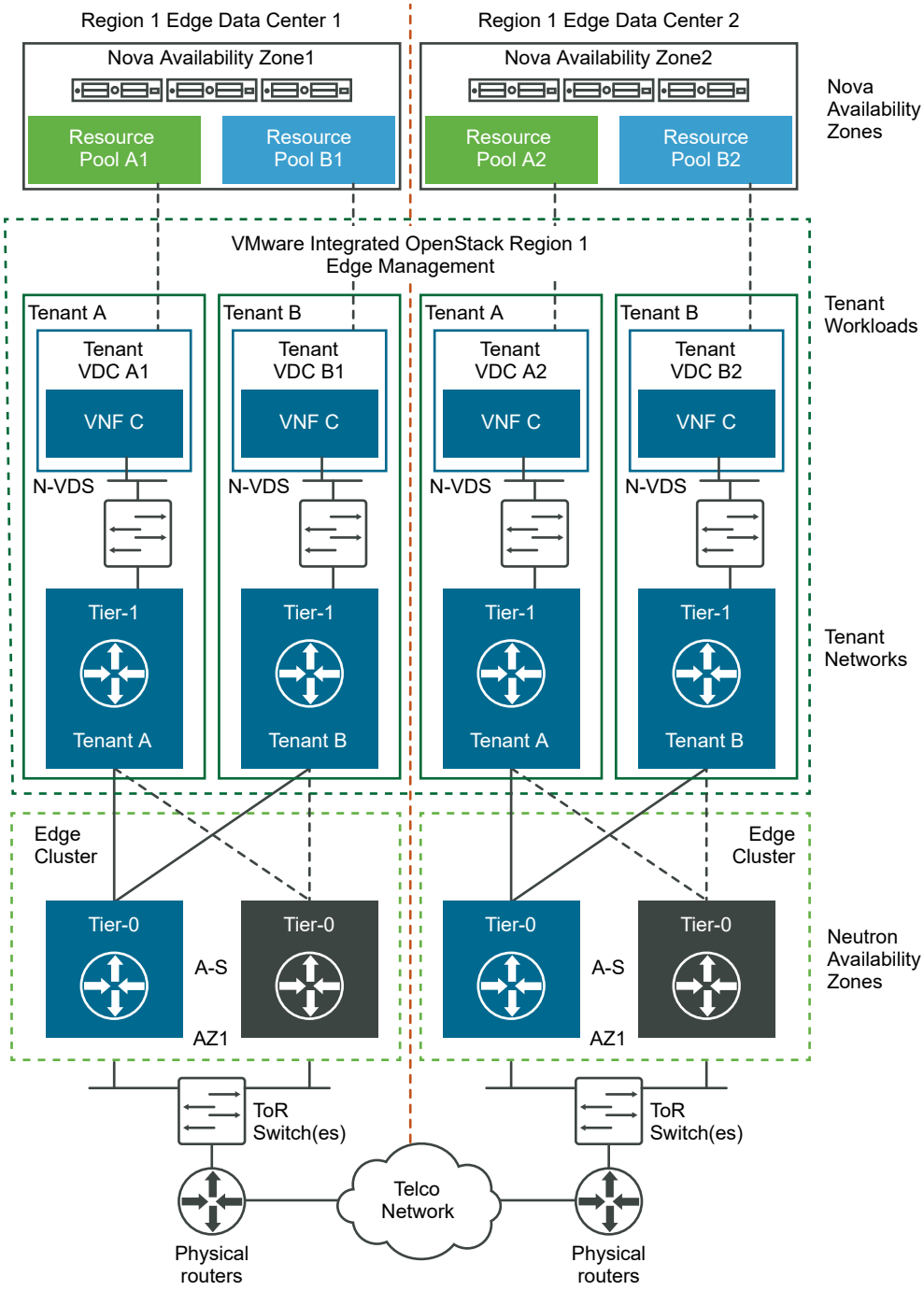
## VNF Placement

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 the polices are respected when there is resource contention. The workload balancing can be manual, semi-supervised, or fully automated.

After the host aggregates are defined and configured, policies for workload placement should be defined for the workload instances. The following diagram describes the workload placement architecture for the edge sites with VIO.

**Figure 6-2. Telco Edge Multi-Tenancy with VMware Integrated OpenStack**




## Flavor Specification for Instances

Flavors are templates with a predefined or custom resource specification that are used to instantiate workloads. A flavor can be configured with additional Extra Specs metadata parameters for workload placement.

## Automation

To meet the operational policies and SLAs for workloads, a closed-loop automation is necessary across the shared cloud infrastructure environment. This domain can host functions such as an NFV-O that is responsible for service blueprinting, chaining, and orchestration across multiple cloud infrastructure environments. Next to NFV-O are the global SDN control functions that 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.

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. The successful cloud automation strategy implies full programmability across other domains. For more information, refer to the API Documentation for Automation.

# Availability and Disaster Recovery

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

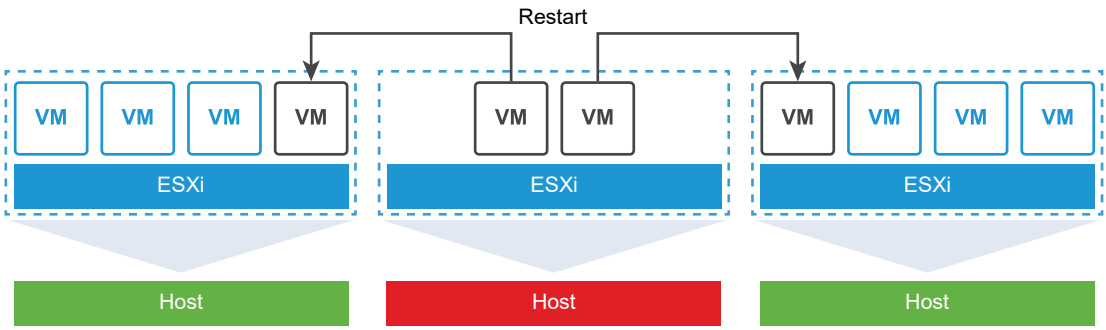
## Availability

All Edge reference architecture components implement a high availability design by default. In addition, VNF's 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 in conjunction with the VNF high availability architecture. vSphere HA can instantiate a new VM 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 6-3. 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 schemes, such as Equal-Cost Multipath (ECMP), Graceful Restart, and Link Aggregation Group (LAG), provide redundant connectivity. The NSX-T architecture also separates management, control, and data plane traffic to further optimize service availability.

### VMware vSAN

vSAN is fully integrated in 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 live migration of virtual machine disk files (VMDK) within and across storage arrays when these are deployed in the environment. This ensures 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.

### Multi-Site Operation and Disaster Recovery

In a hierarchical model, each leaf Edge site is treated as a self-contained entity with a limited number of hosts performing specific functions. When an Edge site fails due either to connectivity issues or a complete outage, such as fires, floods, and other disaster based shutdown, the next level in the hierarchy (central site) should take action. There are two options to accomplish this:

- Restart the workloads that were on the failed Edge site in another Edge site assuming that the new Edge site has additional capacity.
- Spin up these workloads on the central site itself.

  This option is similar to the DR to cloud use case and assumes that the cloud/central site has sufficient capacity to accommodate the workloads that must be recovered and restarted. It is possible to use VMware HCX as an engine for migration of workloads assuming that the destination (central site) has sufficient capacity to run these workloads.

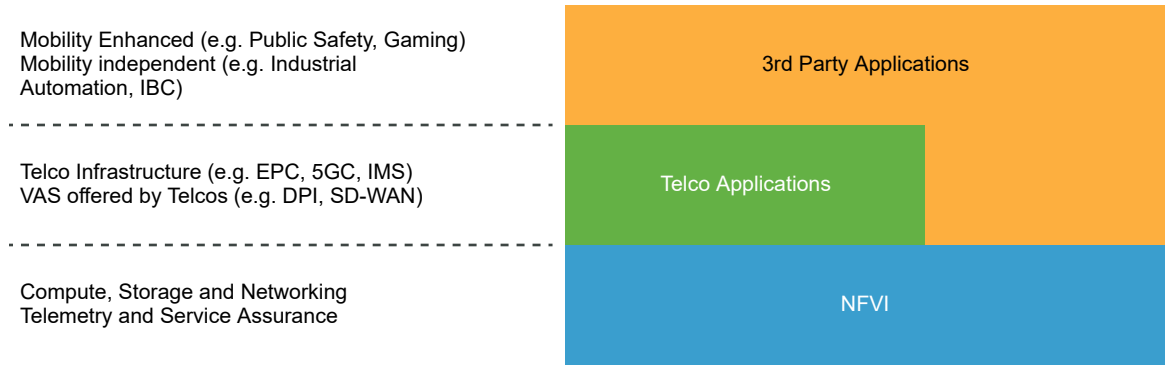
While both these options are viable with respect to restarting workloads in a software defined infrastructure environment, there are some unsolved issues with respect to networking. For example, the user equipment must now establish connectivity to the new site where the workloads are being hosted, for example VNFs. In addition, the network bandwidth requirements for Internet breakout and the additional workload traffic (either from an Edge site or central site) are higher. A related issue is the IP address management for these restarted workloads on the new site.

# Applications for the Telco Edge

7

The Telco Edge deployment enables operators not just to enhance and optimize the delivery of their telecommunication services, but also extend their network to accommodate third-party applications.

**Figure 7-1. Edge Application Layers**




A typical Telco Edge site can be viewed as a 3-layered stack consisting of Virtual Infrastructure, Telco Applications, and third-party applications. As depicted in the preceding figure, end user services can either be on top of the Telco applications or be running directly on the virtual infrastructure. By running on top of the Telco applications, the end-user applications can leverage the mobility aspects of the Telco infrastructure and provide more customized/localized services, such as location-based services. The architecture also allows for some independent applications to run on top of the virtual infrastructure directly. For example, a video broadcasting application that is performing some video pre-processing of a live event before uploading it to the central media distribution site.

This chapter includes the following topics:


- Mobile User Plane Functions
- vCDN Realization Through MEC
- Private LTE and Industrial Automation


## Mobile User Plane Functions

Mobile operators and technologists, under the auspices of 3GPP, have been working for the past several years to define the technical specification for 5G, the next generation mobile network technology. This section briefly discusses how such Telco use cases may be realized by using this VMware Telco edge architecture.

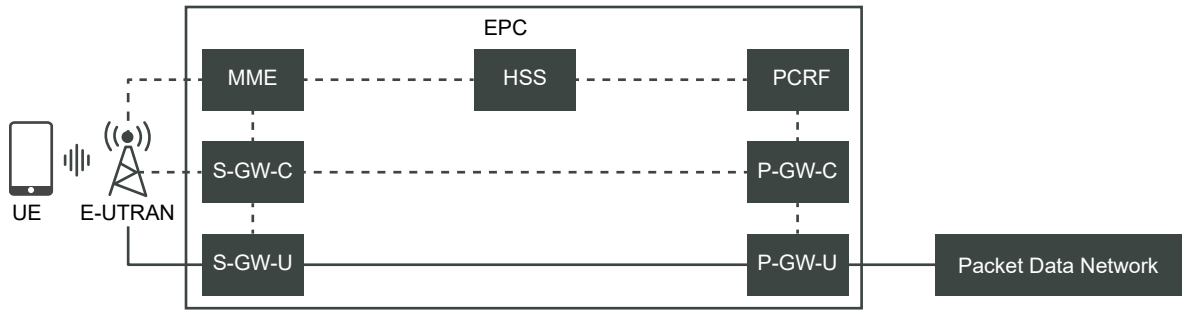
The business objective of 5G is to enable applications that consume a large amount of data (for example, such as live streaming), provide immersive experience (for example, AR/VR), and enable proliferation of network connectivity to billions of devices (for example, IoT) with low latency. These goals of 5G essentially translate to following technical requirements: 20 times the peak data rate (speed), 10 times lower latency (responsiveness), and 3 times more spectral efficiency. This needs upgrading both the Radio network and also the mobile core that handles this large amount of data. There are significant advancements made with respect to Radio networks in terms of spectral efficiency, increased data packing through advanced coding techniques, and so on. With respect to the mobile core, to handle the massive amount of data traffic, there is a push to separate the data plane from the control plane ( CUPS architecture), and to distribute the user plane as close to the device as possible. This is known as the Telco Edge architecture.

## Realization Through VMware Telco Edge

The conceptual design principles discussed in this document are applied in this section to transform the next-generation 5G ready cloud infrastructure with a highly distributed control-user plane architecture in a hub-spoke topology.

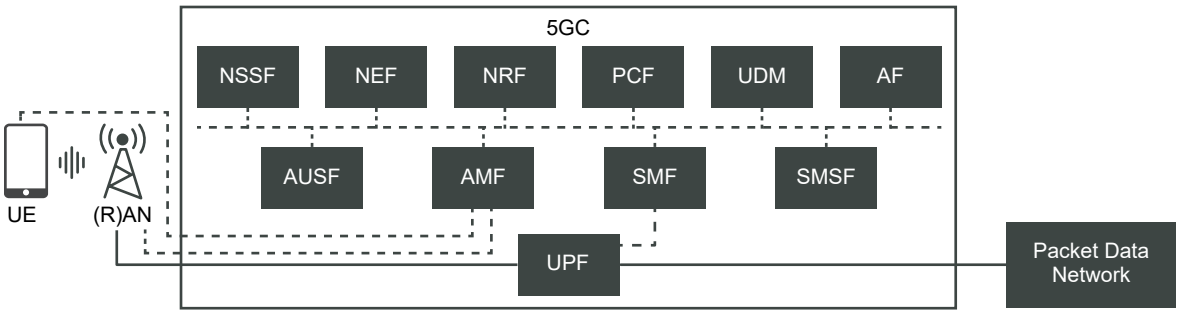
The following diagram depicts the CUPS architecture of 4G EPC and 5GC . In both cases, the CP (control plane) is positioned in an aggregation site, while the UP (user plane) is distributed to the Edges. Low latency and lower cost of front,backhaul, and core transports is a critical challenge to support massive bandwidth and number (billions) of devices. Mobile operators often look at distributing the CP functions to a few regionally distributed sites.

**Figure 7-2. 4G EPC CUPS Architecture**



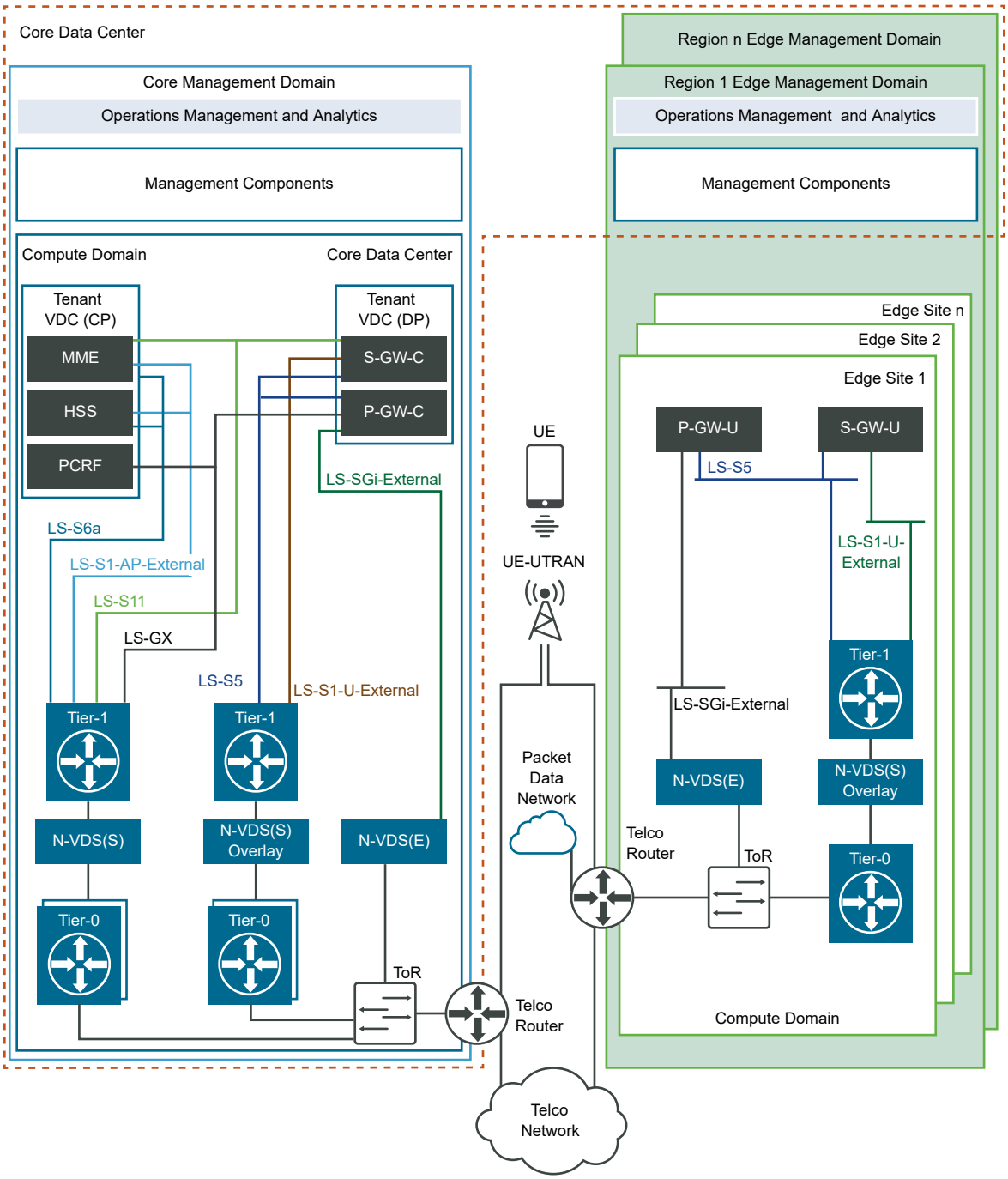

Telco data services for 4G are now delivered over the 3GPP SAE architecture for Evolved Packet Core (EPC). The EPC bridges the user equipment (UE) over a Radio Access Network (RAN) to the core network, creating traffic bearers to deliver a mobile broadband with differentiated class of services. The separation of user plane functions – SGW-u and PGW-u are created by vendors to drive this edge computing and data plane acceleration. Various technologies influence the service quality of the broadband service from the air interface to overlay networks and physical transports.

**Figure 7-3. 5GC Software Architecture**




The following figure shows an example use case for the conceptual deployment of 5GC components onto the vCloud NFV Edge infrastructure. All the management components are deployed in the Core data center, however these are bifurcated as one set of compomemts to manage the workloads deployed at the Core data center and another set to manage the workloads that are deployed at the edge site, as already described earlier in this document.

## Figure 7-4. Mobile UPF Realization Through VMware Telco Edge




This model provides operators with a good balance between simpler topology and efficient processing. Based on this theme, a typical network is likely to look like the architecture presented in preceding figure. Typically, the 5GC UPF components are located at the edge site while the core components are hosted at a Core data center.

# vCDN Realization Through MEC

An important new revenue opportunity for Telco Operators with the 5G networks is with Mobile Edge Computing Platform, alternatively known as MEC. With MEC Platform, operators can dynamically run third-party applications at the Edge of the network to deliver more value to the customer and drive higher revenues.

The MEC platform also exposes mobility characteristics of the end device (UE) to the application, thereby enabling more advanced use cases. For example, a gaming application can customize the screen based on the real-time location of the end user and provide a better user experience. A Content Delivery Network (CDN) is a distributed system positioned throughout the network that places popular content such as streaming video at locations closer to the user than are possible with a traditional centralized data center. When using infrastructure edge computing, CDN nodes operate in software at edge data centers. When instantiated as a virtualized CDN (vCDN), it can operate across a range of virtualized infrastructures from Core data center to every Edge data center and it can fully support all types of virtualized network function deployment at different operator network vantage points.

With latest processors supporting multiple VMs, a CDN function used for delivering video can run in one VM while other services can operate in different VMs on the same physical server. Because this model allows operators to better utilize server resources, a virtualized CDN can be more cost effective than a CDN running on a dedicated infrastructure. Thoughe vCDN addresses cost, it still does not solve challenges related to capacity planning, agility, or quality.

To solve these challenges, operator networks are typically structured as a hierarchy, beginning with core network and large centralized data centers, extending to regional aggregation networks and data centers, and finally reaching last mile access networks. Small PoPs are located close to subscribers that reside at the Edge DC whenever needed so that the vCDN component can be placed accordingly. The full vCDN implementation has to incorporate with four foundational components: virtualized infrastructure, policy-driven orchestration, Software Defined Networks (automation of CDN network configuration and deployment), and Analytics. Edge Reference Architecture relies on NFV stack from VMware and can address the requirements from vCDN.

In recent years, online video consumption over the mobile network has increased manifold due to several reasons:

- Availability of popular content from online video sources, such as Netflix, Hulu, and YouTube.
- Popularity and demand for live streaming of sporting events, concerts, an so on.
- Advancements in the display technology resulting in larger video streams to support increased resolution of HD, 4K, or 8K.
- Unlimited/cheap mobile data plans incentivizing people to access video content on their mobile device.

To address this increasing demand for online video consumption, CDN vendors and Telco operators are looking at a multi-layer Virtualized CDN solution that provides the following key benefits:

- Agile Capacity Allocation: With increasing consumption of live video streams, the demand for video consumption is expected to be more dynamic. Forexample, a regional concert might drive the online video consumption for that event by 3x to 10x in and around that area. However, the actual number cannot be precisely determined. Therefore, it is important that operators can dynamically allocate additional capacity depending on the instantaneous demand.
- Consolidated Delivery Platform: CDN requires specific characteristics from the virtual infrastructure such as fast and large storage and large and elastic network bandwidth. It is not always possible for Telco Operator to install temporary capacity with CDN characteristic in the Edge for live events. Therefore, to keep the CAPEX low, the operator should be able to standardize on a consistent hardware profile and have the software make the best use of the available hardware.
- Orchestrated vCDN: In a multi-layer Telco Edge architecture, the Near Edge and Far Edge can serve as anchor points for fixed and/or mobile broadband offerings. Because these Edges are going to be remotely managed, are numerous in number, and widely distributed, an Orchestration solution that leverages a policy driven lifecycle automation capability is essential to deliver the desired benefits.

## Private LTE and Industrial Automation

As real-time analytics and automation increases in manufacturing, it is becoming increasingly important that the industry premise provides a reliable and scalable wireless communication mechanism. However maintaining session continuity from inside the premise to outside, In-building coverage and uninterrupted mobility are some of the drawbacks of conventional wireless technologies such as WiFi.

To overcome these challenges, many operators are looking at deploying a dedicated mobile network (4G or 5G) at the enterprise site. To make the site completely self-sufficient, the Enterprise deployment includes all data plane and control plane components required to manage a Telco network. By this, none of the mobile session leaves the Enterprise premise unless necessary. In addition, many Enterprises may deploy additional applications in the Edge Site. Sometimes, the Edge can also be directly connected to the Enterprise IT cloud. In terms of requirements, this is similar to the distributed 4G/5G mobile core.

# 8 Authors and Contributors

The following authors co-wrote this paper:

- Arunkumar Kodagi Ramachandra, Solutions Architect, NFV Solutions Engineering, VMware
- Indranil Bal, Solution Consultant, NFV Solutions, VMware
- Ramki Krishnan, Lead Technologist, OpenSource NFV, NFV Solutions, VMware
- Ramkumar Venketaramani, Director Product Management, NFV Solutions, VMware
- Sumit Verdi, Director Lighthouse Solutions, NFV Solutions, VMware
- T. Sridhar, Principal Engineer & Chief Architect, NFV Solutions, VMware
- Tuan Nguyen, Senior Solutions Engineer, NFV Solutions, VMware

Many thanks for contributions from:

- Andrea Li, Lead Solutions Test Architect, NFV Solutions Engineering, VMware
- Danny Lin, Sr. Director Product Management and Architecture, NFV Solutions, VMware
- Jambi Ganbar, Sr. Technical Solutions Manager, NFV Solutions Engineering, VMware
- Kandasamy M, Graphic Designer, Content Strategy & Ops, VMware
- Michelle Han, Director, Solutions Testing and Validation, NFV Solutions Engineering, VMware
- Viji Subramaniam, Program Manager, Information Experience, VMware

Special thanks for their valuable feedback to:

- Arul Dharmalingam, Christian Hasner, Dharma Rajan, Gary Day, Frank Escaros-Buechsel, Gary Kotton, Giridhar Jayavelu, Henrik Blixt, Henrik Oberg, Marcos Hernandez, Mark Voelker, Mauricio Valdueza, Mikael Brihed, Mike Richer, Mustafa Bayramov, Neil Moore, Pradip Kadam, Phil Kippen, Revathi Govindarajan, Richard Boswell, Sudesh Tendulkar, Suresh Babu Nekkalapudi.