Dell EMC VxBlock™ Systems for VMware NSX-T Data Center 3.0

Architecture Overview

June 2021
Rev. 1.4

DELL Technologies

Notes, cautions, and warnings

NOTE: A NOTE indicates important information that helps you make better use of your product.

CAUTION: A CAUTION indicates either potential damage to hardware or loss of data and tells you how to avoid the problem.

WARNING: A WARNING indicates a potential for property damage, personal injury, or death.

© 2016 - 2021 Dell Inc. or its subsidiaries. All rights reserved. Dell, EMC, and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be trademarks of their respective owners.

Contents

Revision history

Date	Document revision	Description of changes
June 2021	1.4	Updated for VMware NSX-T Data Center version 3.1.2.
March 2021	1.3	Added a mention of AMP-2S.
November 2020	1.2	Added support for VxBlock Systems 350, 540, and 740.
September 2020	1.1	Added support for VMware NSX-T Data Center 3.0
May 2020	1.0	Initial release

1

Introduction

This document describes the high-level design of VMware NSX-T Data Center network virtualization technologies for single VxBlock Systems.

This document covers VMware NSX-T Data Center with VMware vSphere running on Cisco UCS C-Series Rack Servers for the physical edge host VMware vSphere cluster. See the *Release Certification Matrix* for more information about supported hardware and software with VMware NSX-T Data Center.

This document is not meant to provide an exhaustive reference guide to the VMware NSX-T Data Center. The purpose is to educate the Dell Technologies community and Dell Technologies customers about the specifics of the VMware NSX-T Data Center design for VxBlock Systems. VMware has published a document titled *NSX-T Reference Design Guide Version 2.0* which provides information about how all aspects of VMware NSX-T Data Center work.

Dell Technologies architected and engineered VMware NSX Data Center for VxBlock Systems (VMware NSX-T Data Center) where VMware NSX-V Data Center is not deployed. Although VMware NSX-T Data Center has many features and supports other hypervisors, the architecture in this guide is limited to VMware vSphere hypervisor and is only for VxBlock Systems. The architecture from Dell Technologies is currently for single site deployment of VMware NSX-T Data Center. For additional use cases (such as multi-site), contact Dell Technologies Support and Deployment Services. Dell Technologies supports the VMware NSX-T Data Center architecture explained in this guide. Support for customizations and additional VMware NSX-T Data Center features is provided by VMware.

The target audience for this document includes Dell Technologies Sales Engineers, field consultants, and advanced services specialists. Use this document to understand the architecture and components of a virtualized infrastructure using VMware NSX-T Data Center on VxBlock Systems.

The Glossary provides related terms, definitions, and acronyms.

2

VMware NSX-T Data Center network virtualization

VMware NSX-T Data Center network virtualization is part of the software-defined data center that offers cloud computing across several platforms.

Platforms include VMware virtualization technologies, bare-metal workloads, Kubernetes-managed (https://kubernetes.io/) container-based workloads, and public cloud. VMware NSX-T Data Center expands on the NSX-V Data Center product by decoupling the management interface and network virtualization capabilities from VMware vCenter Server.

VMware NSX-T Data Center programmatically provisions virtual networks and manages them independent of the underlying hardware. VMware NSX-T Data Center reproduces the entire network model in software, enabling the creation and provisioning of a network topology in seconds. Network virtualization abstracts L2 switching and L3 routing operations from the underlying hardware, just as server virtualization does for processing power and operating systems.

VMware vSphere cluster summary

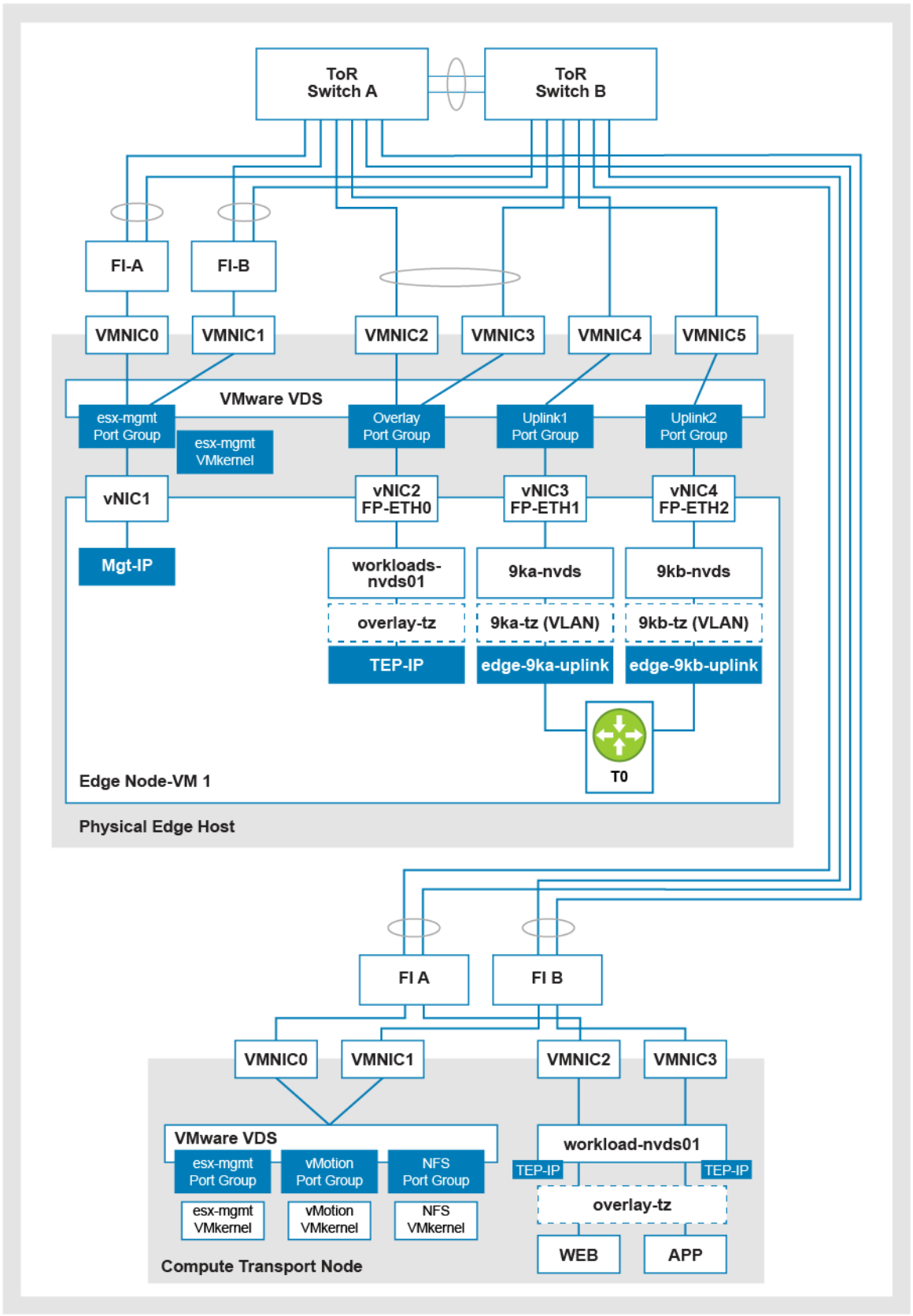
The following table describes the VMware vSphere clusters for VMware NSX-T Data Center on the VxBlock Systems:

VMware vSphere cluster	Description
Management	Consists of the AMP VMware ESXi hosts on which the VMware NSX-T Data Center manager appliance cluster resides. Supported AMP platforms for this solution are AMP-VX, AMP-3S, or AMP-Central. The VMware NSX-T Data Center manager appliance cluster consists of three VMs that host the management plane and control plane components of the VMware NSX-T Data Center system. The management cluster also contains the VMware vCenter Server, which enables the VMware NSX-T Data Center manager appliance cluster to be deployed into the management cluster.
Physical Edge	Consists of the physical edge Cisco UCS C220 servers and contains the edge node virtual machines that provide external connectivity to the physical network and various network services
Transport node	Consists of the production compute hosts and contains the production VMs. There can be more than one transport node cluster.

VMware NSX-T Data Center logical topologies

There are two logical topologies of VMware NSX-T Data Center on a VxBlock System. Topology 1 applies to VxBlock Systems that are running VMware vSphere 6.x or VxBlock Systems that are running VMware NSX-T Data Center along with VMware vSphere that has been upgraded from 6.x to 7.0. Topology 2 applies to VxBlock Systems on which VMware NSX-T Data Center and VMware vSphere 7.0 are factory installed.

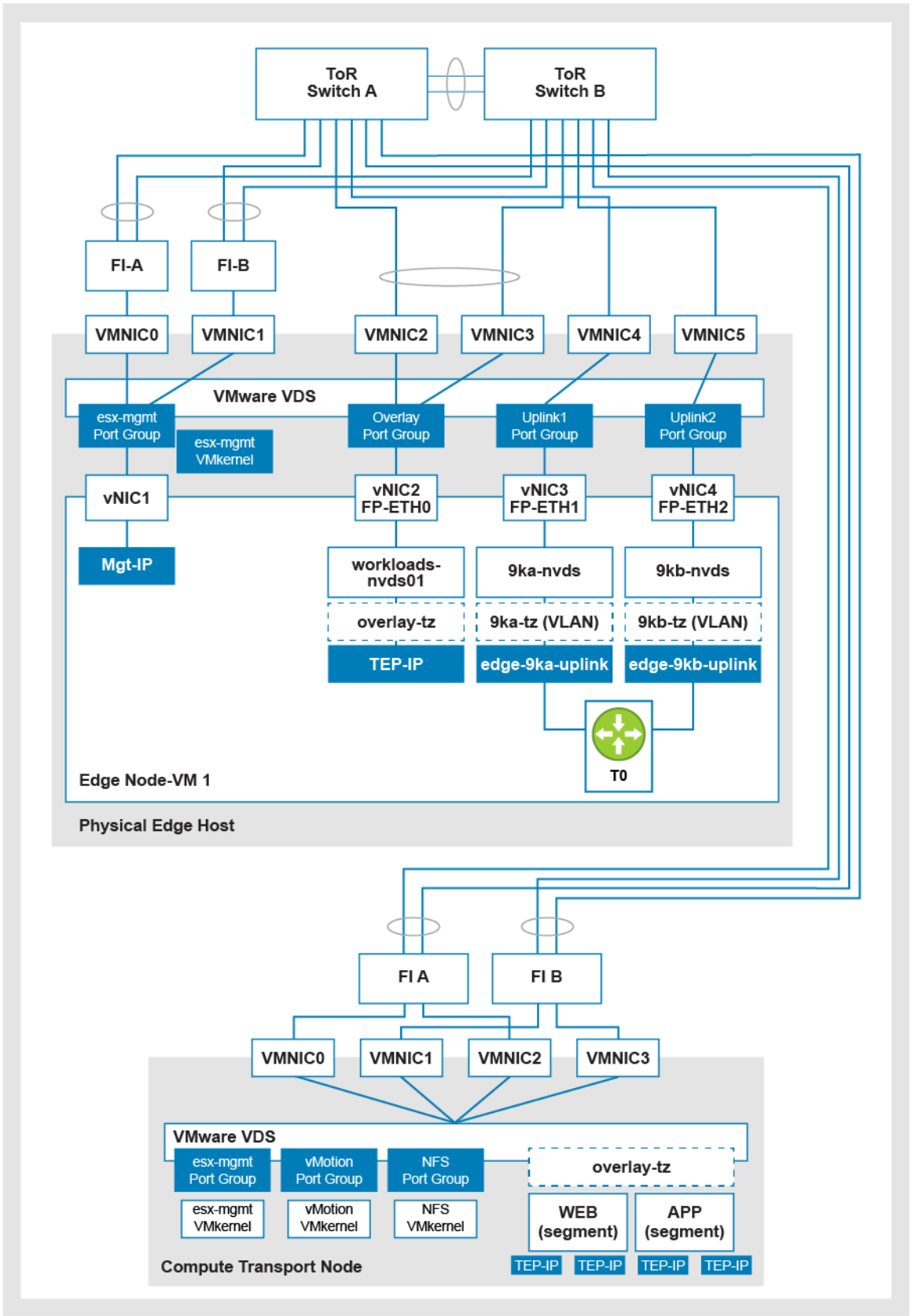
The following figure shows topology 1:



The edge node physical NIC definition includes the following:

- VMNIC0 and VMNIC1: Cisco VIC
- VMNIC2 and VMNIC3: Intel XXV710 (TEP and Overlay)
- VMNIC4 and VMNIC4: Intel XXV710 (N/S BGP Peering)

The following figure shows topology 2:



The edge node physical NIC definition includes the following:

- VMNIC0 and VMNIC1: Cisco VIC
- VMNIC2 and VMNIC3: Intel XXV710 (TEP and Overlay)
- VMNIC4 and VMNIC4: Intel XXV710 (N/S BGP Peering)

In topology 1 and topology 2, the edge node topology is the same, but the transport node topology uses VMware vSphere 7.0 VDS. Under VMware vSphere 7.0, the VMware VDS is aware of the VMware NSX-T Data Center. The VMware VDS contains the overlay-backed segments as opposed to creating a separate N-VDS to house these segments.

3

Cisco Nexus 9000 Series Switch requirements

When you deploy VMware NSX-T Data Center on a VxBlock System, some requirements apply to the Cisco Nexus 9000 Series (ToR) Switches.

Ensure that the following requirements are met:

- L3 licensing is present on the switches. For specific licensing requirements, see Chapter 12: Licensing.
- All VMware and NSX-T VLANs are deployed.
- BGP is enabled and configured.
- Uplinks to the external network are configured as L3. If the uplinks are L2, a separate services engagement is needed to convert the uplinks to L3.

4

VLANs specific to VMware NSX-T Data Center

VMware NSX-T Data Center on VxBlock Systems requires additional VLANs for the ToR switch pair.

Overlay VLAN

The Overlay VLAN carries all east-west data flows for overlay traffic between transport nodes.

For VMware NSX-T Data Center on VxBlock System deployments, this network is defined as L3 routable. If the use case requires transport nodes that are spread across multiple ToR switch pairs, the boundary between L2 and L3 is at the ToR switch pair. A VLAN cannot span that boundary. Examples include:

- A single-site multisystem deployment of VMware NSX-T Data Center
- A future multisite deployment of VMware NSX-T Data Center

In these use cases, each L2 domain (ToR switch pair) must provision a unique subnet to hold the TEP IP addresses for the connected transport nodes. The transport nodes for each transport network subnet need to reach the other transport network subnets in the VMware NSX-T Data Center deployment.

A routable transport network ensures that deployments are flexible and extensible for all use cases.

Uplink1 VLAN

The Uplink1 VLAN enables BGP peering between the Tier-0 Gateway in the edge VM and the ToR Cisco Nexus 9000 Series switch on the A side of the network fabric.

Uplink2 VLAN

The Uplink2 VLAN enables BGP peering between the Tier-0 Gateway in the edge VM and the ToR switch on the B side of the network fabric.

VRF Uplink VLANs

If the VRF-lite option is deployed, an additional pair of uplink VLANs is provisioned for each VRF that is configured. These VLANs are trunked between the corresponding T0-VRF Gateway and the ToR Cisco Nexus 9000 Series Switch. Each VRF has a unique set of BGP peerings between the Tier-0 gateway and ToR switch.

Remote Tunnel End Point VLAN

RTEP VLAN is used when VMware NSX Federation is selected. The NSX-T Global manager uses this VLAN to connect to local NSX-T managers. The VLAN resides on the VPC where the Overlay VLAN is trunked.

5

Default VLAN names and IDs

VxBlock System factory deployments with VMware vSphere 7.0 introduces a naming and numbering scheme for VLANs and port groups. All VLAN names and VLAN IDs are configurable when ordering.

See the *Dell EMC VxBlock System 1000 Architecture Overview* for full details of the naming scheme.

The following table shows the VLAN functional names that are used throughout this document, along with the default values for the VLAN name, VLAN ID, and port group name.

VLAN name	VMware vSphere 6.x VLAN ID	VMware vSphere 6.x port group name	VMware vSphere 7.0 VLAN ID	VMware vSphere 7.0 port group name	VMware vSphere 7.0 VLAN name
esx-mgmt	105	vcesys_esx_mgmt	1631	w-cl01-vds01-pg-mgmt	w-mgmt
vmotion	106	vcesys_esx_vmotion	1632	w-cl01-vds01-pg-vmotion	w-vmotion
overlay	121	vcesys-nsx-transport	1634	w-cl01-vds01-pg-overlay	w-host-overlay
uplink1	122	vcesys-nsxedge-01	2731	w-cl01-vds01-pg-uplink01	w-edge01
uplink2	123	vcesys-nsxedge-02	2732	w-cl01-vds01-pg-uplink02	w-edge02
VRF uplink1	222	vrf1-uplink01	222	vrf1-uplink01	vrf1-uplink01
VRF uplink2	223	vrf1-uplink02	223	vrf1-uplink02	vrf1-uplink02
RTEP	2833	vcesys-rtep	2833	w-cl0-vds-pg-rtep	w-retp

6

BGP routing configuration

Enable BGP routing for VMware NSX-T Data Center on VxBlock Systems ToR switches. Also, a BGP routing process must be created on the ToR switches and associated with the appropriate Autonomous System (AS) identifier.

VMware recommends the BGP dynamic routing protocol for peering Tier-0 gateways to the physical network. If OSPF is used for VxBlock System uplinks, the BGP routes from VMware NSX-T Data Center are redistributed into OSPF.

7

VMware NSX-T Data Center management cluster

The management cluster consists of the management and control planes for VMware NSX-T Data Center. The VMware NSX-T Data Center manager appliance handles both management plane and control plane functions.

AMP cluster components

The VMware NSX-T Data Center contains a management appliance cluster on AMP-VX, AMP-3S, or AMP Central that performs management and control plane functions. Three unified appliances ensure high availability for the control plane. The deployment process assigns a virtual IP address to the cluster to ensure ease of access to the VMware NSX-T Data Center HTML5 user interface.

AMP cluster specifications

Compare the cluster specifications for VMware NSX-T Data Center versions 2.5 and 3.0.

The following table provides the specifications for the VMware NSX-T Data Center manager appliance cluster:

Specification	VMware NSX-T Data Center 2.5 manager node	VMware NSX-T 3.x Data Center manager node	VMware NSX-T 3.x Global manager node
Quantity	Three VMs per VMware NSX-T Data Center manager cluster A single VMware NSX-T Data Center manager cluster manages transport nodes that are connected to: • Up to eight VMware vCenter Servers • Other compute managers like KVM or Kubernetes	Three VMs per VMware NSX-T Data Center manager cluster A single VMware NSX-T Data Center manager cluster manages transport nodes that are connected to: • Up to eight VMware vCenter Servers • Other compute managers like KVM or Kubernetes	Three VMs per VMware NSX-T Data Center Global manager cluster A single VMware NSX-T Data Center Global manager cluster • Supports 4 different locations • Supports a total of 650 Hypervisor transport nodes across all locations
Location	Management cluster	Management cluster	Management cluster
Hardware	AMP-VX, AMP-3S, or AMP Central with a minimum of four servers	AMP-VX, AMP-3S, or AMP Central with a minimum of four servers	AMP-VX, AMP-3S, or AMP Central with a minimum of four servers
Size	Large The three-node cluster consumes: • 24 vCPU • 96 GB RAM • 600 GB disk	Large The three-node cluster consumes: • 36 vCPU • 144 GB RAM • 900 GB disk	Large The three-node cluster consumes: • 36 vCPU • 144 GB RAM • 900 GB disk
Network	vcesys_amx_mgmt (VLAN 205)	vSphere 6.7: vcesys_amx_mgmt (VLAN 205) vSphere 7.0: m-mgmt (VLAN 1611)	vSphere 6.7: vcesys_amx_mgmt (VLAN 205) vSphere 7.0: m-mgmt (VLAN 1611)

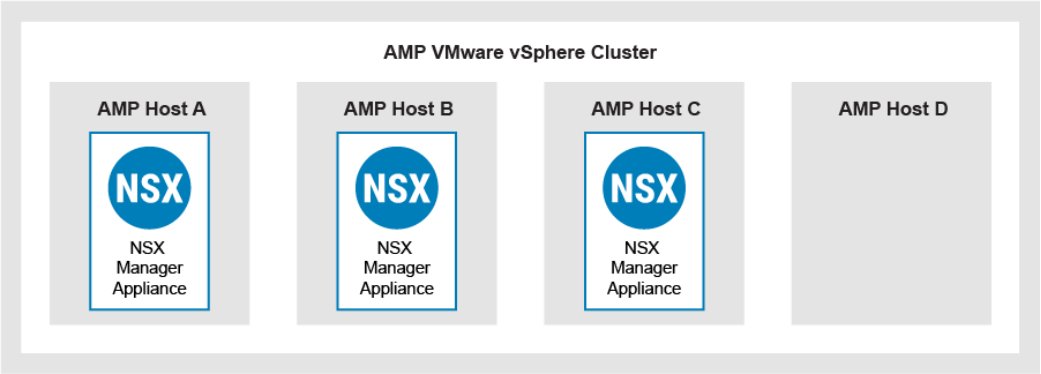
Specification	VMware NSX-T Data Center 2.5 manager node	VMware NSX-T 3.x Data Center manager node	VMware NSX-T 3.x Global manager node
Availability	VMware HA	VMware HA	VMware HA
Distribution	The first node is deployed as part of the OVA deployment. Other nodes are deployed from the VMware NSX-T Data Center manager that runs on the first node.	The first node is deployed as part of the OVA deployment. Other nodes are deployed from the VMware NSX-T Data Center manager that runs on the first node.	The first node is deployed as part of the OVA deployment. Other nodes are deployed from the VMware NSX-T Data Center manager that runs on the first node.

AMP hardware requirements

VxBlock Systems support VMware NSX-T Data Center virtual networking with AMP-VX, AMP-2S, AMP-3S, or AMP Central with a minimum of four servers.

When sizing the solution, ensure that the total AMP workload is compatible with the number of nodes. In some cases, four AMP nodes may not be able to handle the added workload for the VMware NSX-T Data Center management cluster. In this case, more nodes must be ordered. Ensure that there is sufficient storage available for the VMware NSX-T Data Center manager VMs. If necessary, upgrade the Dell EMC Unity or Dell EMC Unity XT storage array in the AMP to the larger option. No special cabling is required.

The following figure shows the layout of the VMware NSX-T Data Center manager cluster that is installed across a four-node AMP cluster:



The configuration requires at least one AMP node in the cluster that does not host a VMware NSX-T Data Center manager appliance under normal operation. This requirement ensures capacity to perform maintenance on the AMP cluster without degrading the VMware NSX-T Data Center manager appliance cluster.

VMware vSphere AMP cluster requirements

The management cluster requires VMware HA and VMware vSphere Distributed Resource Scheduler (DRS) for two reasons. The first is to provide VM protection against a VMware ESXi host failure. The second is to balance VM workloads in the cluster.

The following rules are applied to the DRS:

- Anti-affinity rules are applied to the management cluster to ensure that each VMware NSX-T Data Center manager appliance runs on its own host where possible.
- Too few hosts may prevent each VMware NSX-T Data Center manager appliance to run on its own host. If so, HA allows the VMware NSX-T Data Center manager appliances to co-exist on the same host.

AMP custom resource pool requirements

The VMware NSX-T Data Center management cluster does not require custom resource pools. However, for heavy workloads, create memory reservations for the VMware NSX-T Data Center manager. Configure the AMP with sufficient resources so that there is no competition for resources among the workloads running on the VMware NSX-T Data Center manager appliance.

AMP storage requirements

The management cluster does not require a specific disk layout other than the standard disk layout of the AMP-VX, AMP-3S, or AMP Central.

The VMware ESXi hosts that are connected to the management cluster use the AMP storage array. The VMware NSX-T Data Center manager appliances are deployed across separate data stores where possible to protect against LUN corruption while improving performance and resilience.

VMware NSX-T Data Center manager appliances require a disk latency that is less than 10 milliseconds.

AMP networking requirements

There are no special network requirements for the AMP-VX, AMP-3S, or AMP Central. The VMware NSX-T Data Center management traffic, control plane traffic, and VMware ESXi management traffic share the same network segment to improve performance. A network latency value of less than 10 milliseconds is required between the VMware NSX-T Data Center manager appliances. All AMPs meet or exceed this requirement.

VMware vLCM support for VMware NSX-T

VMware NSX-T Data Center 3.1.2 now integrates with VMware vLCM, the NSX Manager that can manage all the life cycle aspects of NSX-T via vLCM image manager APIs.

NSX Manager leverages vLCM Image manager to enable:

- Installation of NSX-T
- Upgrade of NSX-T
- Uninstallation of NSX-T
- Addition, Removal, or Move of a host in and out of vLCM-enabled clusters

There is no support for the N-VDS switch on a vSphere Lifecycle Manager (vLCM) enabled cluster. Only vDS managed NSX environments are manageable via vSphere Lifecycle Manager (vLCM). Sites that have upgraded from VMware vSphere 6.x to 7.0 cannot use this function unless they migrate NVDS to VDS. This migration is available in NSX-T 3.1.x for sites who are upgrading to VMware vSphere 7.0 U2.

8

VMware NSX-T Data Center physical edge cluster

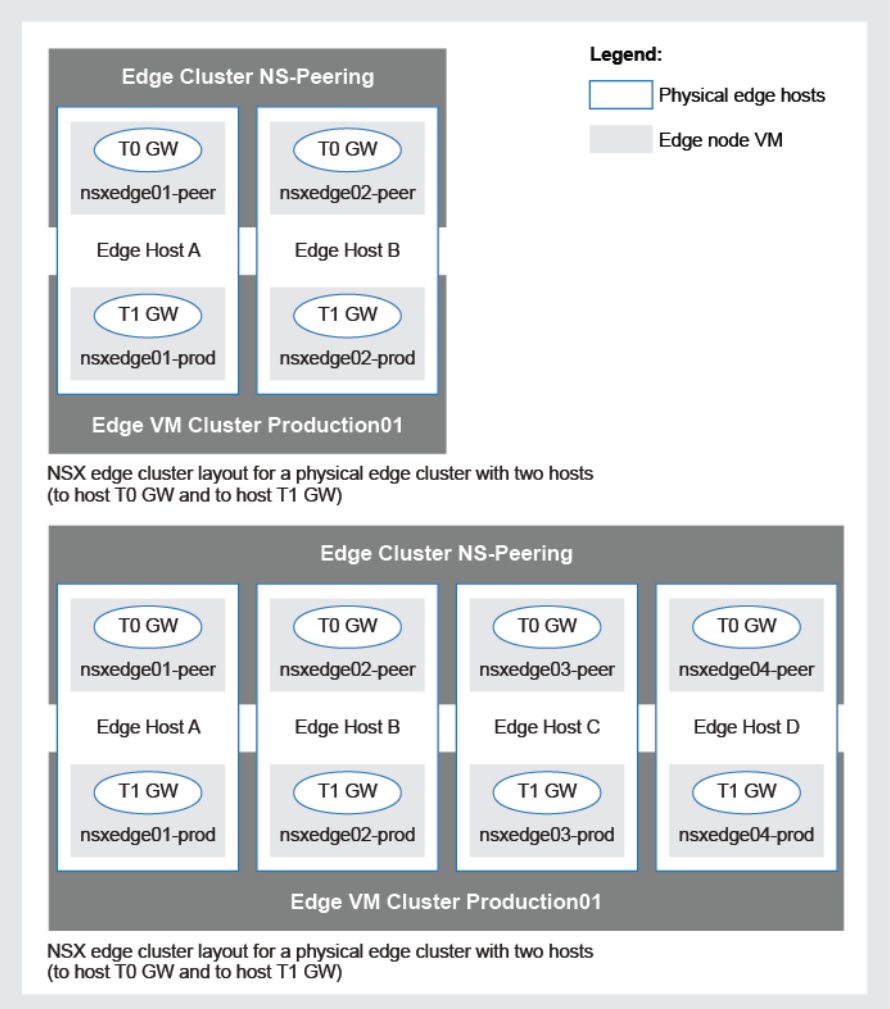
The VMware NSX-T Data Center physical edge cluster connects to the physical network. The cluster provides the physical platform for the network services the edge node virtual machines provide. These services include routing, bridging, and other network services. The Cisco UCS C-Series Rack Servers host the physical edge cluster.

Components

The VMware NSX-T Data Center physical edge cluster consists of an even number of physical servers to host edge VMs.

During initial deployment, the two VMware NSX-T Data Center logical edge clusters are built on the first two or four physical edge hosts in the cluster. Create host affinity rules in vSphere to ensure that this configuration is maintained in all situations except a component failure.

The following figure illustrates how the VMs are configured:



If the initial deployment includes more than four physical edge hosts, only the first four hosts are populated with edge VMs. Use the rest of the hosts as needed.

Specifications

NS-peering cluster and the Production01 cluster configuration values are compared.

The following table summarizes the VMware NSX-T Data Center edge cluster specifications:

Specification	Edge VMs NS-peering cluster	Edge VMs Production01 cluster
Quantity	Two or four VMs active/active with ECMP	Two or four VMs active/active with ECMP
Location	Physical edge host cluster	Physical edge host cluster
Hardware	Cisco UCS C-Series Rack Servers	Cisco UCS C-Series Rack Servers
Size	Medium 4 vCPUs 8 GB RAM 200 GB disk	Large 8 vCPUs 32 GB RAM 200 GB disk
Network	Four VMNIC interfaces: • Appliance management on the esx_mgmt VLAN • Transport (East-West/TEP) traffic on the Overlay VLAN • Fabric A edge (North-South) traffic on the Uplink1 VLAN, plus any VRF Uplink 1 VLANs • Fabric B edge (North-South) traffic on the Uplink2 VLAN, plus any VRF Uplink 2 VLANs	Four VMNIC interfaces: • Appliance management on the esx_mgmt VLAN • Transport (East-West/TEP) traffic on the Overlay VLAN • Fabric A edge (North-South) traffic on the Uplink1 VLAN, plus any VRF Uplink 1 VLANs • Fabric B edge (North-South) traffic on the Uplink2 VLAN, plus any VRF Uplink 2 VLANs
Availability	VMware HA or DRS in partially automated mode	VMware HA or DRS in partially automated mode
Distribution	Affinity rules are enabled to place the edge VMs to match the layout in the preceding figure.	Affinity rules are enabled to place the edge VMs to match the layout in the preceding figure.

Hardware requirements

This section provides the hardware requirements for the Cisco UCS C-Series Rack Servers.

VMware NSX-T Data Center hardware requirements

The selection and sizing of the hardware platform for the VMware NSX-T Data Center design considers the VMware system requirements for the various components.

The use cases for VMware NSX-T Data Center are broad and are customizable for many business use cases. Two engineered sizes are provided for physical edge hosts. The standard host has 96 GB of memory and supports most use cases. The extra large host has 192 GB of memory. It supports use cases that require many large edge node VMs such as SSL, VPN, and load-balancing today. The extra-large host size also supports upcoming features that have larger resource requirements.

Cisco UCS C220 M5 component configuration overview

Review component values for the Cisco UCS C220 M5-based edge node.

See the following table for the component values for the Cisco UCS C220 M5 based edge node.

Component	Cisco UCS C220 M5 Server based edge node
CPU	Two Intel Xeon 5218 2.3 GHz, 16 core, 22 MB cache
Memory	96 GB (6 x 16 GB DDR4-2933) standard-size host 192 GB (12 x 16 GB DDR4-2933) extra-large-size host

Component	Cisco UCS C220 M5 Server based edge node
NIC	Two Intel XXV710-DA2, dual-port 25 Gbps PCIe adapter
VIC	Cisco UCS VIC 1457 (4 x 10 Gbps or 25 Gbps SFP28 mLOM)
Storage	The Cisco UCS primary storage array provides the storage. Valid primary storage arrays are VMAX, PowerMax, Dell EMC Unity, and XtremIO X2.
Boot device	SD or boot from SAN

CPU

For VMware NSX-T Data Center edge workloads, the optimal configuration has enough CPU cores to prevent oversubscription. Having enough CPU cores provides maximum performance to the VMware NSX-T Data Center edge-node VMs that carry the workload on these nodes. The VMware NSX-T Data Center design includes one large and one medium edge node VM on each of the first four physical edge hosts in the edge cluster.

The configuration uses the Intel Xeon 5218 CPU, which has 16 cores running at 2.3 GHz. Two CPU sockets are populated per server. The Cisco UCS C220 M5 server requires a dual-socket configuration to enable the second PCIe slot that is used for the second Intel NIC. This configuration provides 32 physical cores per edge host.

The following table shows the vCPU consumption for a default configuration of the physical edge host:

Component	Quantity	CPU requirement	Notes
VMware NSX-T Data Center edge node, size medium	1	4 vCPU	This edge node is used for the NS peering edge cluster, which is used for the BGP peering between VMware NSX-T Data Center and the physical network.
VMware NSX-T Data Center edge node, size large	1	8 vCPU	This edge node is used for the Production01 edge cluster, providing T1 routing capabilities and services for applications.
Total used (no oversubscription)		12 cores	NA
Free		20 cores	Free represents the available capacity for edge requirements defined in the LCS.

This configuration consumes only 12 of the 32 available CPU cores in a physical edge host. The remaining capacity is available for customization. Use the additional CPU cores to deploy additional edge nodes or increase the size of the default edge nodes.

VMware NSX-T Data Center 3.0 introduces a new extra large edge node VM form factor, which consumes 16 vCPUs. After upgrading to NSX-T Data Center 3.x, you can expand the Production01 edge nodes to the larger size. For new deployments, you can specify the extra large size edge appliance in the LCS.

The following table shows the vCPU consumption for a default configuration of the physical edge host with the extra large sized edge node VM appliances:

Component	Quantity	CPU Requirement	Notes
VMware NSX-T Data Center edge node, size medium	1	4 vCPU	This edge node is used for the NS-Peering edge cluster, which is used for the BGP peering between the NSX Tier-0 router and the physical network.
WMware NSX-T Data Center edge node, size extra large	1	16 vCPU	This edge node is used for the Production01 edge cluster, providing T1 routing capabilities and services for applications.
Total Used (no oversubscription)		20 cores	N/A
Free		12 cores	This is the available capacity for site-defined edge requirements.

Memory

The Cisco UCS C220 M5 physical edge host uses 6 or 12 16 GB DDR4-2933 RDIMMs. These RDIMMs provide 96 GB or 192 GB of RAM.

The default deployment of a physical edge host allocates system memory as described in the following table:

Component	Quantity	Memory requirement (standard host)	Memory requirement (extra large host)	Notes
VMware NSX-T Data Center edge node VM, size medium	1	8 GB	8 GB	This edge node is used for the NS-Peering edge cluster, which is used for the BGP peering between VMware NSX-T Data Center and the physical network.
VMware NSX-T Data Center edge node VM, size large	1	32 GB	32 GB	This edge node is used for the Production01 edge cluster, providing T1 routing capabilities and services for applications.
VMware ESXi Hypervisor 6.5, 6.7, or 7.0.	1	4 GB	4 GB	This memory size represents the minimum requirement to run VMware ESXi.
Total used		44 GB	44 GB	NA
Free		52 GB	148 GB	This amount of memory is available for site-defined edge requirements, for support or recovery from failures.

The default deployment of a physical edge host allocates CPU resources as described in the following table:

Component	Qty	CPU Requirement	Notes
VMware NSX-T Data Center edge node, size medium	1	4 vCPU	This edge node is used for the NS-Peering edge cluster, which is used for the BGP peering between the NSX Tier-0 router and the physical network.
VMware NSX-T Data Center edge node, size extra large	1	16 vCPU	This edge node is used for the Production01 edge cluster, providing T1 routing capabilities and services for applications.
Total Used (no oversubscription)		20 cores	N/A
Free		12 cores	This is the available capacity for site-defined edge requirements.

VMware NSX-T Data Center 3.0 contains an extra large edge node VM form factor, which consumes 16 vCPUs. For new deployments, you can specify the extra large edge appliance size.

The following table describes the allocation of memory in a default factory deployment of a physical edge host with extra large edge node VM appliances:

Component	Quantity	Memory Requirement (standard host)	Memory Requirement (extra large host)	Notes
VMware NSX-T Data Center edge node VM, size medium	1	8 GB	8 GB	This edge node is used for the NS-Peering edge cluster, which is used for the BGP peering between NSX Tier-0 router and the physical network.
VMware NSX-T Data Center edge node VM, extra large	1	64 GB	64 GB	This edge node is used for the Production01 edge cluster, providing T1 routing capabilities and services for applications.
VMware ESXi Hypervisor 6.7	1	4 GB	4 GB	This reflects the minimum system requirement to run VMware ESXi.
Total Used		76 GB	76 GB	NA
Free		20 GB	116 GB	This is the available capacity for site-defined edge requirements, as well as extra memory that can be used for support and recovery from failures.

NICs

The choice of physical network interface cards in a VMware ESXi host is critical for maximizing performance. VMware NSX-T Data Center makes extensive use of protocol offloading to achieve line-rate performance.

The network interface card and its associated drivers for VMware ESXi must support the following:

- GENEVE offload
- RX or TX queueing

The Intel XXV710-DA2 card provides dual-port SFP28 connectivity and can support both 10 Gbps and 25 Gbps transceivers. This card has a standard, eight-lane PCI Express interface. Because each lane has a bandwidth of 8 Gbps, the maximum PCIe throughput for the PCIe interface of a single card is 64 Gbps.

Cisco VIC

The Cisco VIC 1457 connects to FI using the Cisco SingleConnect technology.

Cisco SingleConnect provides the following services to the physical edge host:

- Enables Cisco UCS Manager to provide storage to the physical edge hosts from the VxBlock System primary storage arrays
- Provides connectivity to the VxBlock System VMware ESXi management and VMware vMotion networks
- Enables server configuration using Cisco UCS service profiles to simplify deployment

Connectivity model

This section provides information on the connectivity model for VMware NSX-T Data Center.

VIC to FI or FEX connectivity

A Cisco UCS VIC 1457 to FI connection uses either 10 Gbps direct-attach cables or SFP-10G-SR optics and LC cabling.

Standardizing on 10 Gbps connectivity to the FI ensures compatibility with Cisco UCS second generation or fourth generation domains. Connect ports 1 and 3 on the VIC to the FIs. Ports 2 and 4 on the VIC are unused. If two Cisco UCS Fabric Extenders are connected to the FI, FEX ports can be used to connect the VIC to the Cisco UCS infrastructure.

For physical edge hosts that are connected to a Cisco UCS 6332-16UP Fabric Interconnect, the hosts connect to QSFP-based server ports using optical breakouts. To configure the optical breakout, do the following:

- Connect a Cisco QSFP-40G-SR4 transceiver into a QSFP port on the FI.

- In Cisco UCS Manager, configure the transceiver as a breakout port.

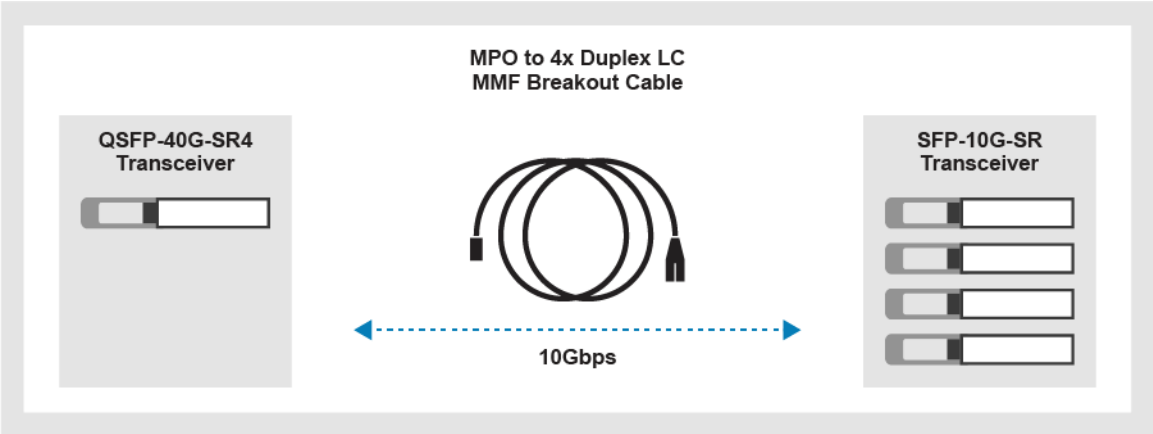
This configuration yields four 10 Gbps lanes that can be used for server connections. The number of QSFPs populated on the FI depends on the number of physical edge hosts in the cluster.

The Cisco UCS VIC 1457 card in the server is populated with Cisco SFP-10G-SR transceivers. MPO hydra and trunk cabling connect the two together.

The following table shows the quantity of QSFP ports that are populated for various physical edge cluster sizes:

Cluster size (physical edge nodes)	QSFP ports populated on each third generation FI
2 or 4	2
6 or 8	4
10 or 12	6
14 or 16	8

The 40 Gbps to 10 Gbps breakout connection is used when directly connecting a physical edge host to a Cisco UCS third-generation FI. The following figure shows the connection topology:



Intel NIC to ToR switch connectivity for VxBlock Systems 1000

Connect the Intel XXV710-DA2 card to a Cisco Nexus 9300 Series Switch with QSFP28 ports.

The Cisco PID for Cisco 25 Gbps SFP+ SR transceiver (SFP-25G-SR-S) is used in the Intel cards. Optical hydra cabling is used to connect to one of the four 25 Gbps channels on a Cisco QSFP-100G-SR4-S module in the switch.

Intel NIC to ToR switch connectivity (VxBlock Systems 340, 350, 540, and 740)

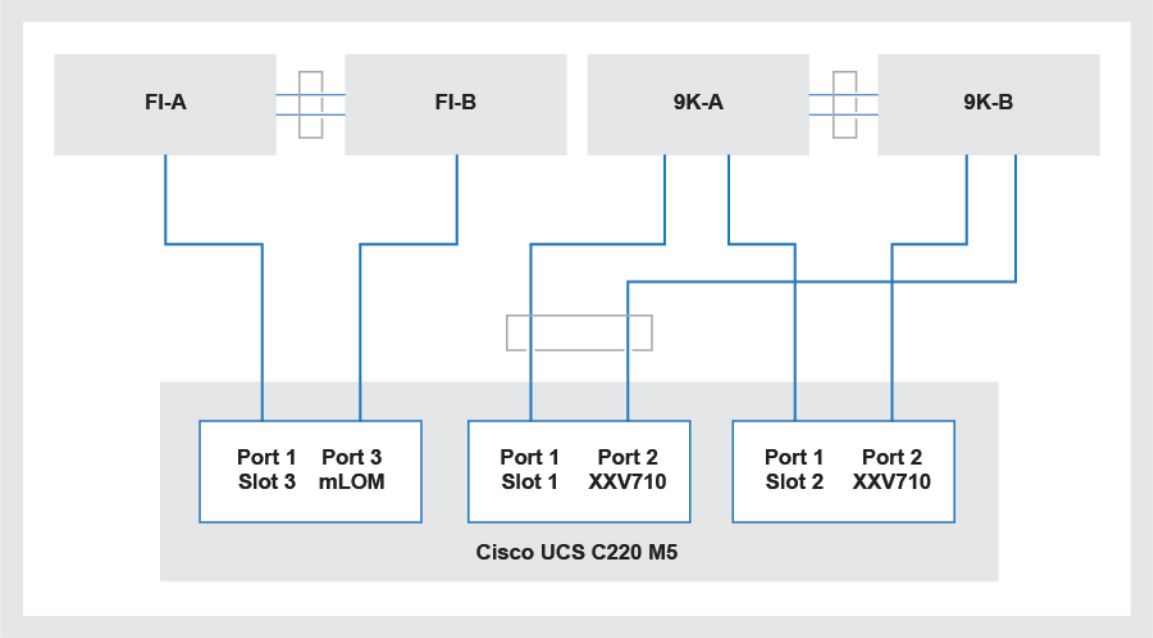
Connect the Intel XXV710-DA2 card to a Cisco Nexus 9300 Series Switch with QSFP or QSFP28 ports.

The Cisco PID for Cisco 25 Gbps SFP+ SR transceiver (SFP-25G-SR-S) is used in the Intel cards. Optical hydra cabling is used to connect to one of the four 25 Gbps channels on a Cisco QSFP-100G-SR4-S module in the switch.

Connect the Intel XXV710-DA2 card to a Cisco Nexus 93180YC-EX or 9396PX series ToR switch with SFP+ ports. The Cisco PID for Cisco 10 Gbps SFP + SR transceiver (SFP-10G-SR) is used in the Intel cards. LC fiber cabling is used to connect to one of the 10 Gbps interfaces on a Cisco SFP-10G-SR module in the switch.

Physical edge node uplink cabling topology

The following figure shows the physical topology from the Cisco UCS C220 M5 Server to Cisco FIs and to the Cisco Nexus 9300 Series ToR switches:



Storage requirements

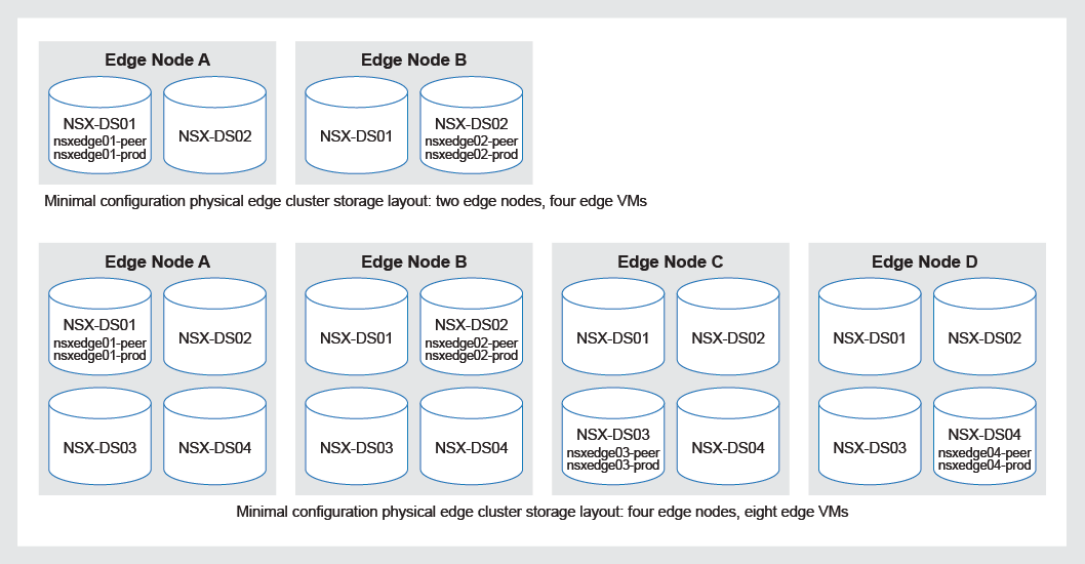
There are storage and networking requirements for the edge cluster.

Storage requirements

- Data stores: The VMware ESXi hosts that are connected to the edge cluster can host the VMware NSX-T Data Center components on any supported primary storage array. A 1.2 TB LUN is deployed for each physical edge server in the edge cluster. The edge node virtual machines that are associated with that host reside on this data store.
- Edge data stores: All edge data stores are presented to all physical edge hosts in the cluster reducing recovery time after an HA event.
- Edge node VMs: The size of each edge node VM under VMware NSX-T Data Center is 200 GB. A default deployment includes two edge node VMs on each VMware NSX-T Data Center physical edge host, which consumes 400 GB of storage. More storage space is left on the data store to support the deployment of additional edge node VMs to each edge node. The additional edge node VMs help to meet business requirements.
- Disk layout: No specific disk layout is necessary. VMware NSX-T Data Center supports all primary storage arrays that are available on the VxBlock System.

See *Network requirements for the Cisco UCS physical edge host servers*.

The following figure shows the physical edge hosts and their storage layout from the factory:



The first configuration shows the layout for a minimal edge node configuration with two physical edge hosts. The second configuration shows the configuration for a four-node physical edge cluster. Any edge nodes above four are not configured in the factory. You must deploy the data stores for these hosts at the site using the deployment model for the first four hosts.

Network requirements for the Cisco UCS physical edge host servers

The network requirements for the Cisco UCS physical edge host servers are described.

The physical edge host cluster requires three VLAN SVIs on the ToR switches:

- Two external uplink VLAN SVIs are used for external traffic for North-South traffic flows. Each VRF has a distinct set of uplink VLANs.
- One overlay VLAN is used to pass overlay traffic for East-West traffic flows.

With physical Edge host Cisco UCS C-Series Rack Servers, using the Cisco UCS Manager to create external edge traffic VLAN IDs is unnecessary. However, because the transport nodes pass overlay traffic, the transport VLAN ID must be added to the Cisco UCS Manager.

VMware virtual network for the physical edge hosts

For Cisco UCS C-Series Rack Server physical edge hosts, a single VMware VDS is created for the physical edge cluster.

The VMware VDS uses the following uplinks:

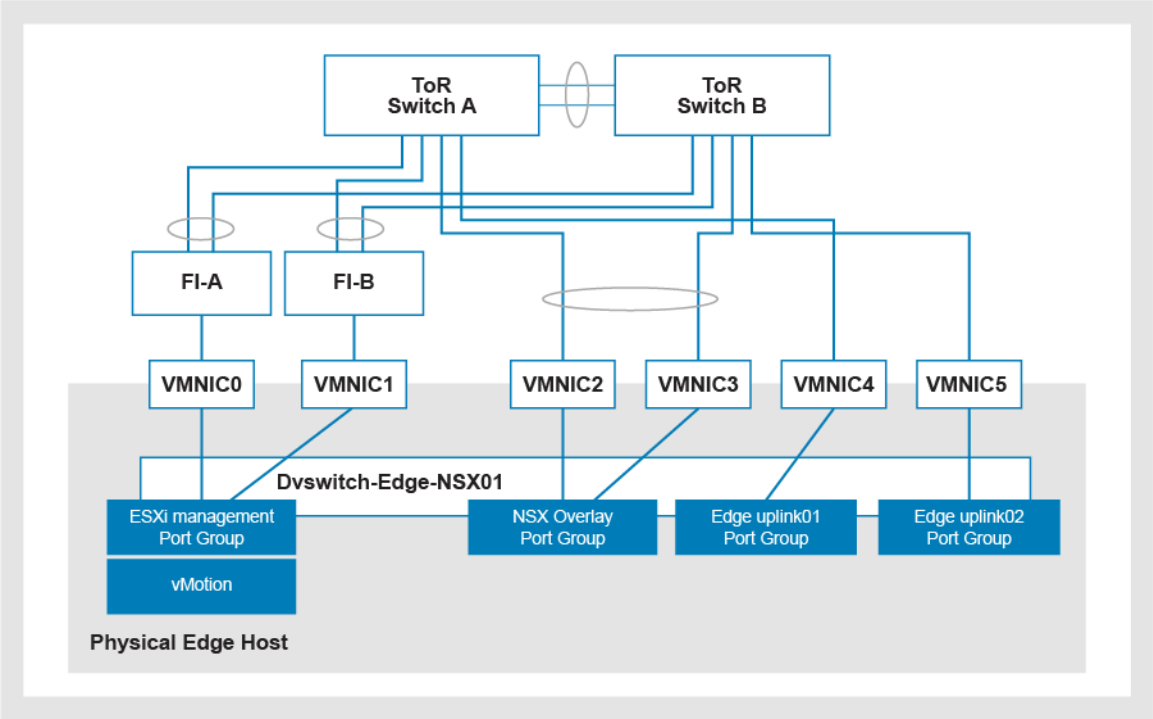
- Two uplinks from the Cisco VIC 1457 mLOM adapter to the FI
- Four total uplinks, two from each of the two Intel XXV710-DA2 adapters to the ToR switches

The port groups are pinned to the appropriate uplinks for the traffic type in the VMware VDS. The following table describes the uplink pinning for VMware vSphere 6.x and 7.0 deployments:

Distributed port group	Description	Uplink pinning
ESX-mgmt	Provides VMware ESXi management VLAN access to the host.	Cisco VIC 1457 uplinks
vMotion	Provides the vSphere vMotion network to the host.	Cisco VIC 1457 uplinks
Overlay	Carries transport traffic, which is the GENEVE-encapsulated traffic for all East-West traffic flows on overlay backed segment. On the physical edge cluster, this VLAN is used exclusively for traffic:	Intel XXV710-DA2 adapter 1 Ports 1 and 2 (vPC/LAG)

Distributed port group	Description	Uplink pinning
	• That arrives from a transport node destined to an edge service, or external network. • That arrives from an edge service or external network destined to a transport node. • Must egress through the edge cluster to an endpoint outside of the VMware NSX-T Data Center deployment.	
Uplink1	Provides the network that is used on the A side of the network fabric for BGP peering northbound from the Edge node VM cluster to ToR switch A in the VxBlock System. This VLAN is the egress point from VMware NSX-T Data Center to the rest of the external network on the A side of the fabric.	Intel XXV710-DA2 adapter 2 Port 1
Uplink2	Provides the network that is used on the B side of the network fabric for BGP peering northbound from the Edge node VM cluster to ToR switch B in the VxBlock System. This VLAN is the egress point from VMware NSX-T Data Center to the rest of the external network on the B side of the fabric.	Intel XXV710-DA2 adapter 2 Port 2

The following figure shows the uplink topology for a VMware NSX-T Data Center physical edge host:



Interlink does not carry external VLANs 122 or 123.

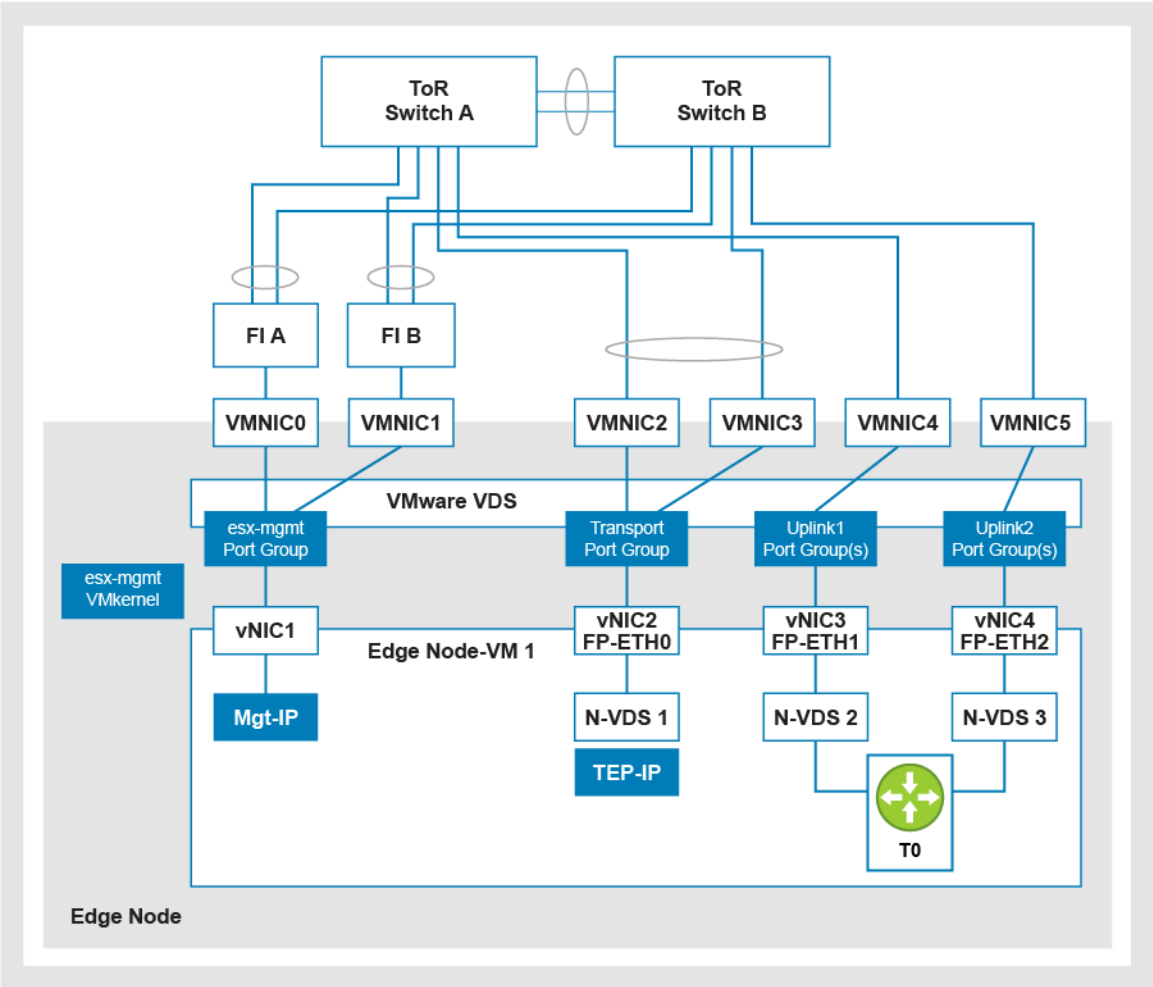
The edge node physical NIC definition includes the following:

- VMNIC0 and VMNIC1: Cisco VIC 1457
- VMNIC2 and VMNIC3: Intel XXV710 adapter 1 (TEP and Overlay)
- VMNIC4 and VMNIC4: Intel XXV710 adapter 2 (N/S BGP Peering)

Logical topology of the physical edge host

The physical edge host logical topology is presented.

The following figure shows the logical topology of the edge hosts in a VMware NSX-T Data Center configuration:



The edge node physical NIC definition includes the following

- VMNIC0 and VMNIC1: Cisco VIC 1457
- VMNIC2 and VMNIC3: Intel XXV710 adapter 1 (TEP and Overlay)
- VMNIC4 and VMNIC4: Intel XXV710 adapter 2 (N/S BGP Peering)

9

Edge cluster architecture standards

In VMware NSX-T Data Center, an edge cluster is a group of edge nodes. The edge nodes can be deployed in a virtual (VM based) or physical (bare-metal) form factor. The initial Vblock System implementation of VMware NSX-T Data Center supports only virtual edge clusters. This implementation provides some benefits for flexibility of deployment and serviceability over bare-metal edge nodes.

The VMware NSX-T Data Center for VxBlock Systems design adopts the VMware recommended edge cluster design for service providers. This design includes two edge clusters:

- Edge-Cluster-NS-Peering is a dedicated edge cluster to host a Tier-0 gateway for BGP peering and north-south traffic flows. The traffic flows need to communicate from the VMware NSX-T Data Center environment to the physical environment and the outside world. The edge node VMs that make up this cluster are deployed as medium-sized appliances.
- Edge-Cluster-Production01 is a production edge cluster that tenants or business units can use. The edge node VMs that make up this cluster are deployed as large-sized or extra-large-sized appliances. A T1 gateway should be associated with this cluster **only** if the Tier-1 gateway is hosting centralized services such as NAT or edge firewall. If the intended use case for the deployment does not require any Tier-1 services, this cluster can be safely removed post-deployment.

In a large IT shop or multitenant environment, this design provides role-based security:

- One group can secure the NS-Peering edge cluster to VMware NSX-T Data Center administrators and support personnel.
- A different group can manage the Production01 edge cluster.

VMware NSX-T Data Center enables the definition of role-based access to these clusters independently, using the API.

This design also mitigates limitations in the BGP peering capabilities of the Tier-0 gateway. The Tier-0 gateway supports eight-way ECMP peering. The VxBlock System VMware NSX-T Data Center design connects a single Tier-0 gateway in one of two ways in the edge-cluster-NS-peering cluster:

- Through two edge node VMs (4-way ECMP)
- Through four edge node VMs (8-way ECMP)

Depending on the use case, you can enhance edge services in the following ways:

- Add more edge node VMs.
- Add a Tier-0 gateway and associated edge node VMs.

Two ToR switches north of the edge cluster can support only four edge VMs peering north-bound to the ToR switches in a cluster. The Production01 edge cluster Tier-1 gateway does not need to participate in ECMP peering with the ToR switches. If necessary, the cluster can contain more than four nodes.

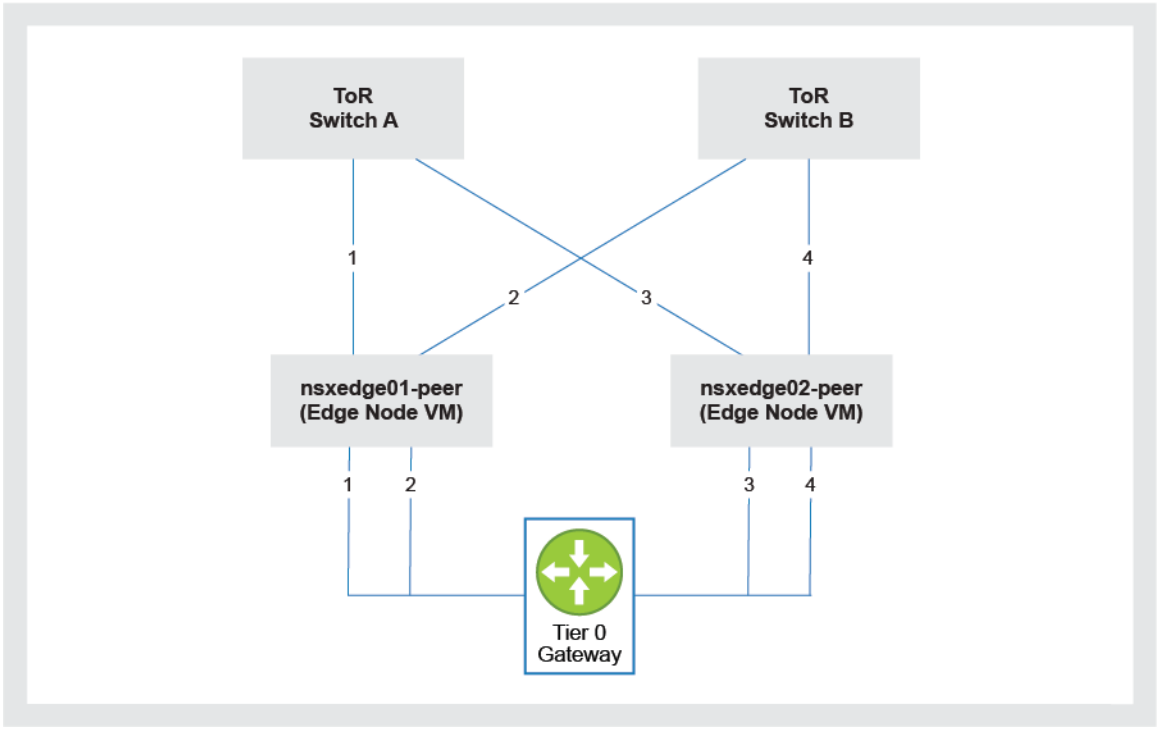
10

Edge ECMP topology

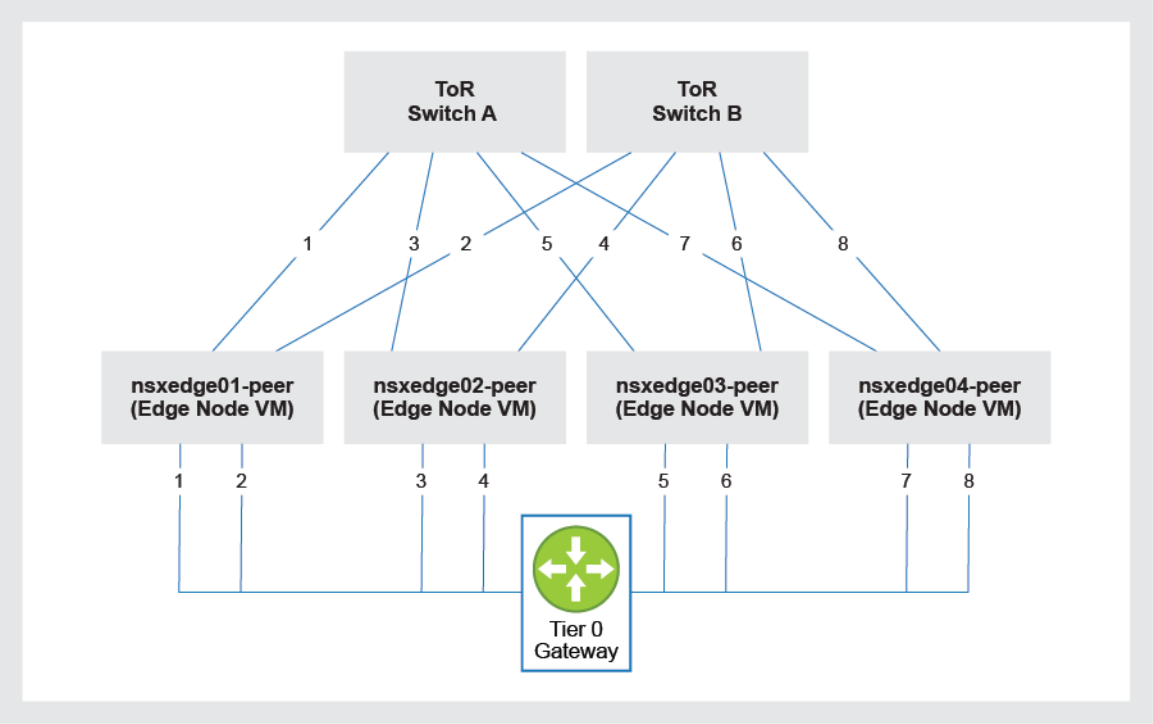
In VMware NSX-T Data Center for VxBlock Systems, ECMP provides load-balancing across multiple paths northbound from the Tier-0s to the ToR switches. VMware NSX-T Data Center supports a maximum of eight paths per edge node cluster.

Each Tier-0 gateway peers with each of the two ToR Cisco Nexus 9000 Series Switches. As a result, the maximum number of edge node VMs in the cluster is four.

The following figure shows a minimum edge cluster configuration with two edge node VMs peering with the ToR switches, consuming four ECMP paths:



The following figure shows a maximum edge node cluster configuration with four edge node VMs peering with the ToR switches, consuming eight ECMP paths:



This design peers each edge node with each ToR switch.

ECMP routing configuration

Each Tier-0 gateway peers with each ToR switch using BGP, and traffic is distributed across the available edge node VMs using ECMP.

The ECMP implementation in VMware NSX-T Data Center limits the number of paths to eight.

In the VMware NSX-T Data Center for VxBlock Systems design, each edge node VM peers with each of the two ToR switches. As a result, the maximum number of edge node VMs in the cluster is four. With two edge node VMs, the default configuration is a single Tier-0 gateway with four interfaces. Two interfaces peer with ToR switch A and two interfaces peer with ToR switch B. This configuration is the default as delivered. Each topology dictates the specific Tier-0 configuration.

For more details on the ECMP design, see *VMware NSX-T Data Center physical edge cluster*.

Bi-directional Forwarding Detection

The Bi-directional Forwarding Detection (BFD) technology detects when a send or receive channel of a network connection is not functioning correctly.

This VMware NSX-T Data Center design implements BFD in aggressive mode on the following:

- The uplinks for the Tier-0 gateway on the edge node VM
- The physical interfaces on the ToR switches

BFD is also configured on the uplinks from the ToR switches to the external network for all VRFs.

The BFD timer value is set to 500 milliseconds, and the BFD multiplier is set to 3.

The timer value means that a BFD control packet is sent every half second. When the system detects that three packets are missed, BGP automatically adjusts the routing table to exclude the problem link.

The detection and exclusion process takes less than two seconds, and conforms to the VMware recommended minimum values to achieve the quickest possible convergence time.

Edge node VM resource usage and Data Plane Development Kit

The VMware NSX-T Data Center edge node VMs implement the Data Plane Development Kit (DPDK) standard to provide high-performance packet forwarding capabilities.

The DPDK reserves CPU cores to constantly poll the NIC for packets instead of waiting for them to be processed using an interrupt. The constant polling produces the following performance on edge mode VM vCPUs:

- Some of the vCPUs run at 50 percent of capacity under no load.
- Some of the vCPUs run at 100 percent of capacity under moderate load.

There may be a high CPU utilization level on edge node VMs that are operating properly.

NS-peering edge cluster

The engineered VxBlock System implementation of VMware NSX-T Data Center deploys medium-sized edge VMs for the NS-Peering edge cluster. This cluster carries North-South traffic flows exclusively and must not run any services other than BGP peering and ECMP. The medium appliance meets these requirements.

Production01 edge cluster

The production01 edge cluster uses the large-sized edge VM. The production01 cluster hosts various services for the VMware NSX-T Data Center environment, including load balancing, VPN services, NAT, DHCP, and Edge Firewall. These services consume additional resources so use a large-sized edge VM. This cluster should be used only in cases where VMware NSX-T Data Center Tier 1 centralized services are instantiated.

Custom edge clusters

The VMware NSX-T Data Center solution for VxBlock Systems includes two edge clusters. Optionally, deploy additional edge VM clusters after initial system deployment. All VMware supported configurations are allowed provided they do not interfere with operation of the original design. The physical edge nodes have sufficient resources to allow an additional large edge node or several small or medium edge nodes per host. Ensure that each of the hosts in the cluster has some reserve capacity. Maintenance and recovery operations can be completed as needed with no warnings or errors appearing in the VMware vCenter Server UI.

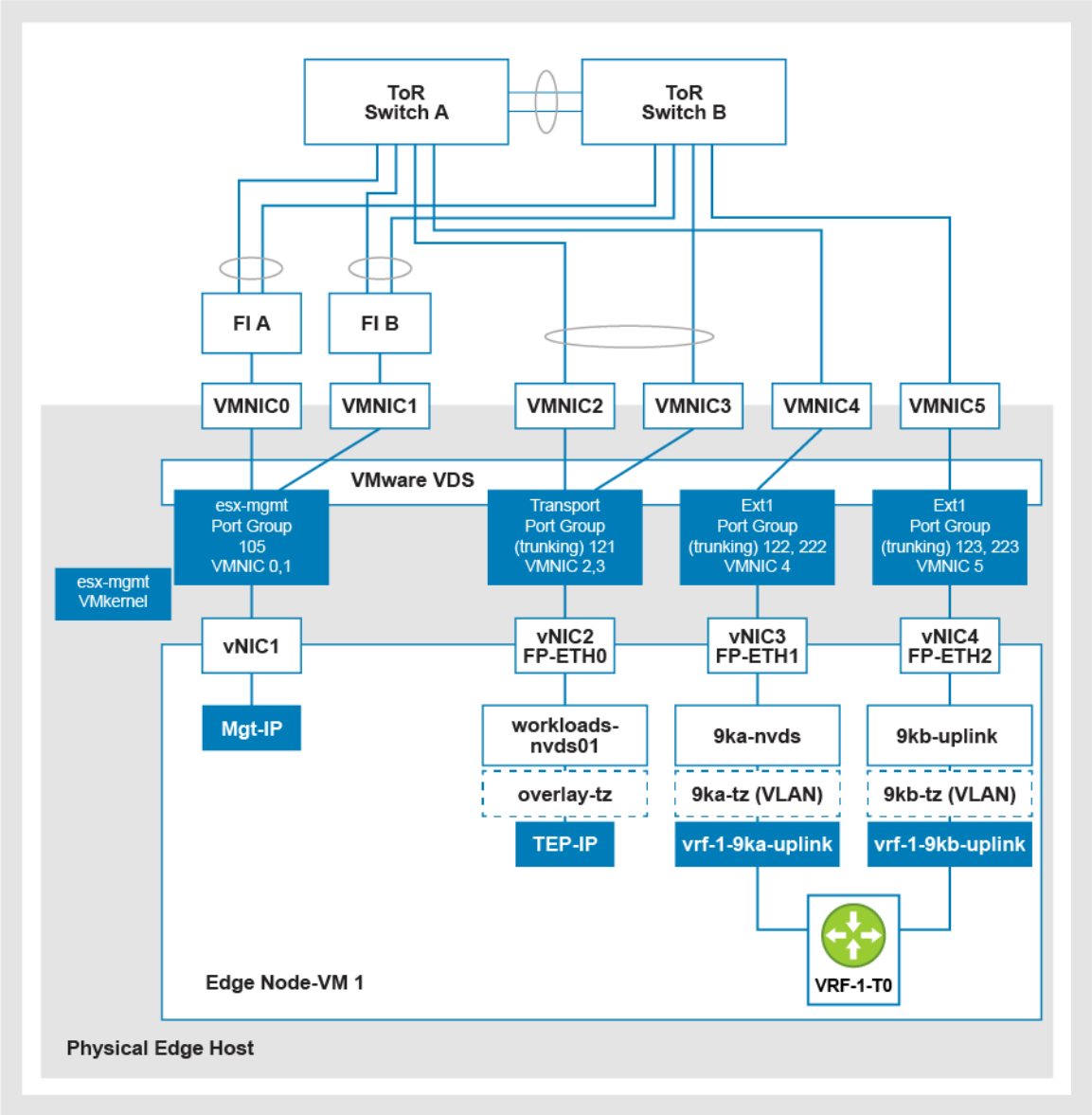
VRF Lite

VMware NSX-T Data Center 3.0 contains a new feature into the VMware NSX-T Data Center UI to configure multiple Virtual Routing and Forwarding (VRF) instances. This feature adds the ability to create separate routing tables for various environments.Using VRF makes it possible to completely separate the routing tables for production, development, and test environments. You can also separate traffic by tenant in a multitenant environment.

The VMware NSX-T Data Center for VxBlock Systems design has added the capability to deploy multiple VRFs. Each VRF must have a distinct pair of uplink VLANs to carry ingress and egress traffic. This traffic is carried on the same physical uplinks from the Edge hosts to the ToR switches. Each pair of VRF VLANs carries the BGP peering connections that are associated with that VRF.

Subinterfaces on the physical uplink ports on both sides of the connection segment the traffic by VRF. This segmentation occurs between the ToR switches and the external network. Distinct point-to-point connections are created on each subinterface to keep the traffic separated all the way to the external network.

The default VRF uses VLAN 122 and VLAN 123. The first VRF defined in the LCS uses VLAN 222 and VLAN 223. For VMware vSphere 7.0, the default VRF uses VLAN 2731 and VLAN 2732, and the first VRF defined in the LCS uses VLAN 222 and VLAN 223. The following figure shows the topology that is used for VMware vSphere 6.7 on the transport nodes with two VRFs deployed:



The edge node physical NIC definition includes the following:

- VMNIC0 and VMNIC1: Cisco VIC
- VMNIC2 and VMNIC3: Intel XXV710 (TEP and Overlay)
- VMNIC4 and VMNIC5: Intel XXV710 (North/South BGP Peering)

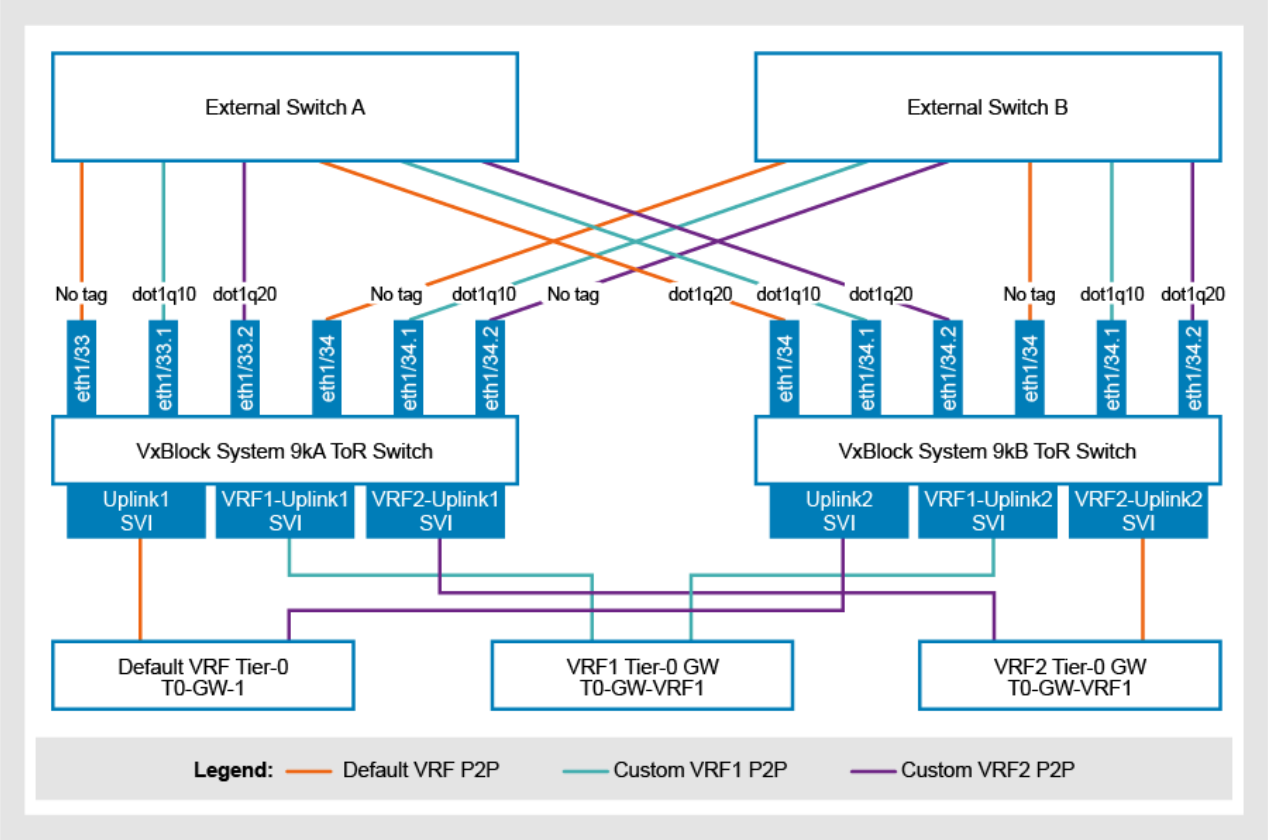
VRF deployment uplink topology

Each VRF requires dedicated BGP peerings to the ToR switch pair so you must implement a pair of uplink VLANs per VRF. Just like with the default VRF, one of the uplink VLANs is associated with ToR switch A, and the other is associated with ToR switch B.

The ports connected to physical edge hosts vmnic4 and vmnic5 are configured as trunk to provide connectivity to the SVIs associated with the uplink VLANs for each VRF.

The physical ports on the ToR switches that are connected to the external network are configured with subinterfaces, and these subinterfaces are marked with a **`dot1q`** tag to enable segregation of the VRF traffic they are carrying.

The following figure illustrates the uplink topology for a VMware NSX-T Data Center 3.0 on VxBlock Systems deployment with two custom VRFs deployed in addition to the default VRF:



Segments

VMware NSX-T Data Center uses GENEVE encapsulation to encapsulate data plane traffic. An Overlay Backed Segment is the VMware NSX-T Data Center equivalent of a port group on a VMware VDS.

VMware NSX-T Data Center also allows for VLAN backed segments, which are segments for which the traffic is not GENEVE encapsulated.

The VMware NSX-T Data Center for VxBlock System design uses segments for North-South and East-West traffic. The segments are used for uplinks to the ToR switches and overlay traffic for T1 gateway subnets. The N-VDS uses the overlay traffic segments to enable the hosts to pass the traffic through the overlay network.

- The **edge-fabric-a-seg** carries the north-south traffic to the Cisco Nexus 9000 Series ToR switch A for the default VRF.
- The **edge-fabric-b-seg** carries the north-south traffic to the Cisco Nexus 9000 Series ToR switch B for the default VRF.
- **vrf-fabric-a-seg** carries the north-south traffic to the Cisco Nexus 9000 Series ToR switch A for all custom VRFs in the solution, if configured. This segment trunks all uplink-1 VLANs associated with custom VRFs.
- **vrf-fabric-b-seg** carries the north-south traffic to the Cisco Nexus 9000 Series ToR switch B for all custom VRFs in the solution, if configured. This segment trunks all uplink-2 VLANs associated with custom VRFs.
- **T1-customer-seg1** is a sample overlay backed segment that carries the Tier-1 gateway east-west traffic flows on the overlay network.
- **T1-customer-seg2** is a sample overlay backed segment that carries the Tier-1 gateway east-west traffic flows on the overlay network.
- **customer-vlan-seg** carries production VLAN backed traffic.

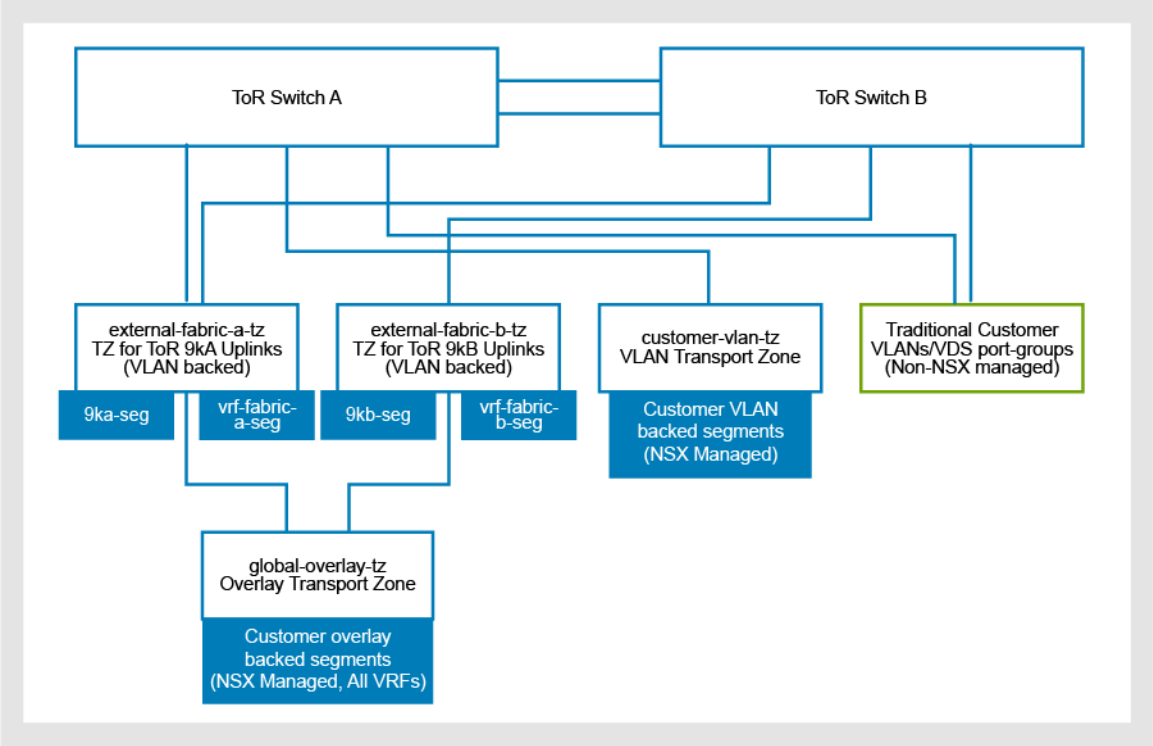
The VMware NSX-T Data Center for VxBlock System design instantiates two segments on the Tier-1 gateway. These segments enable testing of intersegment routing after the VMware NSX-T Data Center deployment is completed. If the segments are not needed, you can delete them.

When configuring a VxBlock System with VMware NSX-T Data Center, you have the option to configure each compute cluster to participate in NSX. For clusters that do not participate in VMware NSX-T Data Center, you can define custom VLANs and present them to these clusters.

Transport zones

Segments are created as part of a VMware NSX-T Data Center object called a transport zone. There are VLAN-backed transport zones and overlay-backed transport zones.

The following figure shows the default transport zones and segments that are deployed:



- VLAN backed transport zones: These zones connect to the physical infrastructure for north-south connectivity.
- VLAN backed transport zones can also provide VMware NSX-T Data Center services such as microsegmentation to workloads that do not require the GENEVE overlay.

Overlay transport zones: These zones use the VMware NSX-T Data Center domain to route GENEVE-encapsulated traffic to external devices or networks and centralized services. An edge VM can support one overlay transport zone.

The VMware NSX-T Data Center for VxBlock System design includes transport zones for uplink and overlay segments. The transport zones create the NSX-managed Virtual Distributed Switch (N-VDS) that the segments are connected to.

The uplink transport zones are VLAN-backed segments which are defined as the following in the design:

- external-fabric-a-tz for the VLAN traffic north-south to the ToR switch A
- external-fabric-b-tz for the VLAN traffic north-south to the ToR switch B

An overlay transport zone carries the east-west GENEVE traffic. The overlay transport zone is global-overlay-tz.

The VMware NSX-T Data Center for VxBlock System design provides a minimal framework to support the following:

- East-west overlay-backed data flows
- North-south egress from overlay-backed segments to VLAN backed endpoints on the physical network
- North-south ingress from the physical network that is destined to an endpoint behind the overlay
- Connectivity to NSX managed VLAN backed segments

Customize or add to this design after the solution is delivered. However, modified designs must not interfere with the operation of the system as designed.

Profiles

An uplink profile template defines how an NSX-managed Virtual Distributed Switch (N-VDS) connects to the physical network.

An uplink profile specifies the:

- Format of the uplinks of an N-VDS
- Default teaming policy that is applied to those uplinks
- Transport VLAN used for overlay traffic (if relevant)
- MTU of the uplinks
- Network I/O control profile

The VMware NSX-T Data Center for VxBlock Systems design includes two uplink profiles:

Profile	Description
Edge-9ka-uplink	This uplink profile carries north-south traffic to the ToR switch A. This profile has one uplink port that is assigned with an MTU of 9000. No VLAN is attached to this profile because tagging of this VLAN occurs at the port group of the VDS where the edge node is connected.
Edge-9kb-uplink	This uplink profile carries north-south traffic to the ToR switch B. This profile has one uplink port that is assigned with an MTU of 9000. No VLAN is attached to this profile because tagging of this VLAN occurs at the port group of the VDS where the edge node is connected.

For the compute workload environment, the VMware NSX-T Data Center for VxBlock System design includes the following profiles to connect workloads to the VMware NSX-T Data Center domain:

Profile	Description
nsx-edge-tep-profile	This profile defines the Overlay network VLAN, the teaming policy (failover order), and uplink port (uplink-1) for which the workload traffic is routed. This profile also assigns the host/vm an overlay IP address based on the TEP pool for East-West traffic in the VMware NSX-T Data Center domain. The centralized Service Port Feature requires tagging to happen at the N-VDS. Trunk (not tag) in the VDS for the Overlay VLAN on the edge nodes. The design sets the VLAN Type in the Overlay port group to **VLAN Trunking**. Also, in the VMware NSX-T Data Center Manager, the design defines the Overlay VLAN in the uplink profile for nsx-edge-tep.
nsx-compute-transport	The transport hosts use this profile to connect to the VMware NSX-T Data Center domain using the following: ● Teaming policy (load balance source) ● Uplink ports (uplink-1,uplink-2) for a VMware vSphere 6.7 deployment. In a VMware vSphere 7.0 deployment with VMware NSX-T Data Center 3.0, the uplink ports are uplink-1, uplink-2, uplink-3, and uplink-4. ● Overlay VLAN, which carries the transport host traffic

Tier-0 gateway

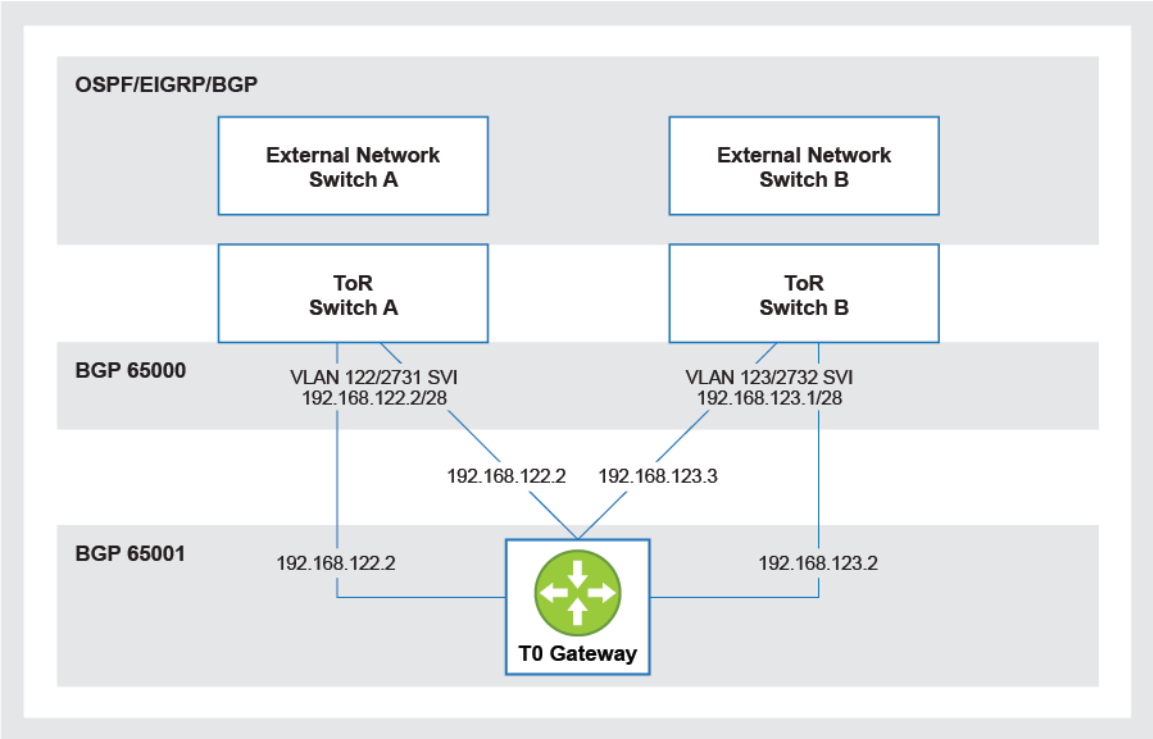
A Tier-0 gateway has downlink connections to Tier-1 gateways and uplink connections to physical networks. The Tier-0 gateway provides a gateway service between the logical and physical network. The Tier-0 gateway enables route redistribution, routing, and BGP, as well as optional T0 services.

The interfaces for this gateway connect the physical edge hosts to each ToR switch using the uplink segments described in *Segments*.

In the VMware NSX-T Data Center for VxBlock System design, by default, there is one Tier-0 gateway to handle the T0-GW-1 traffic.

T0-GW-1 uses BGP and ECMP to link the nodes to the ToR switches.

The following figure shows the north-south routing for the Tier-0 gateway in the design:

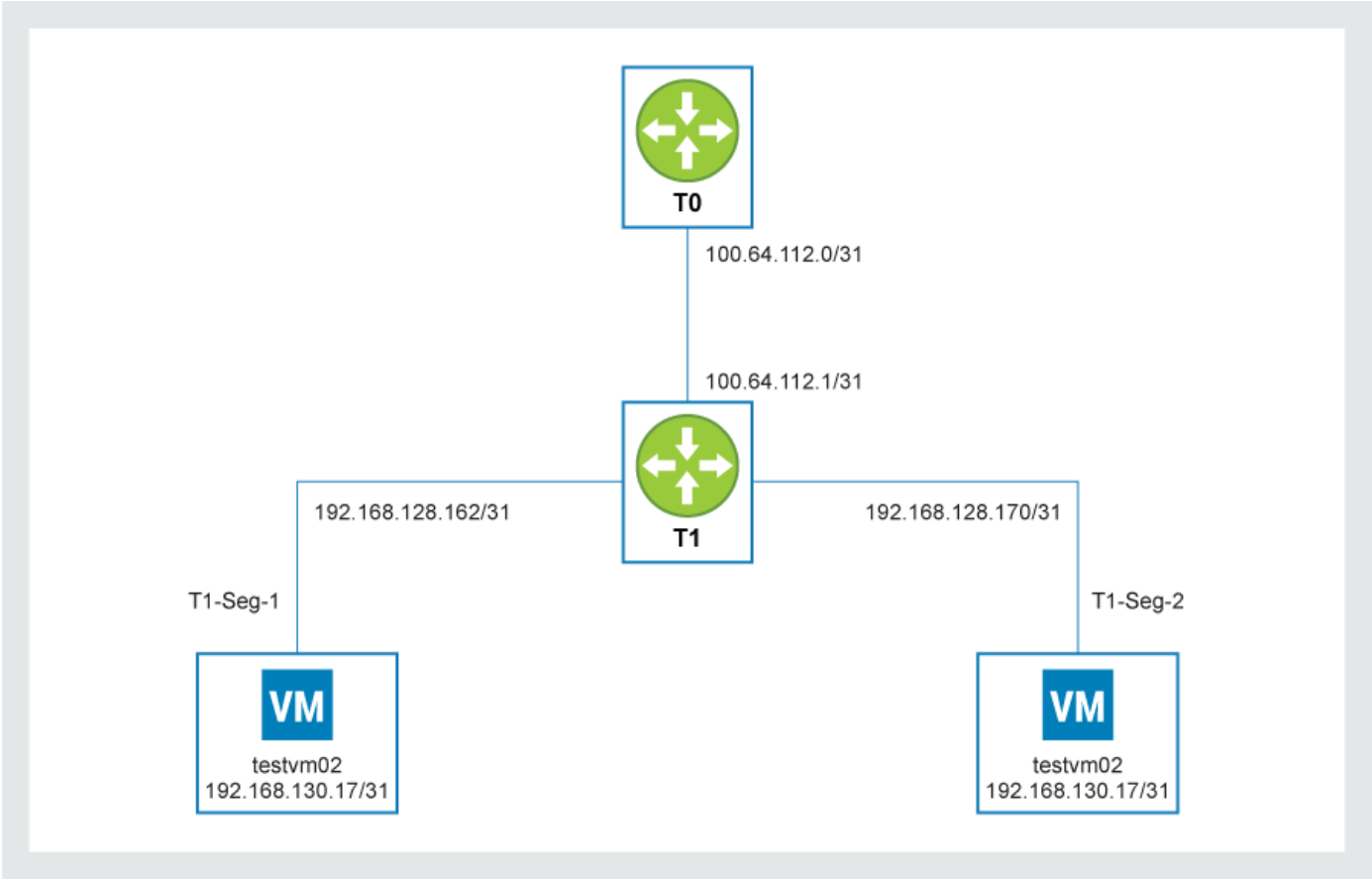


Tier-1 gateway

A Tier-1 gateway connects northbound to a Tier-0 gateway and to southbound segments. Tier-1 gateways may also provide centralized services such as load balancing, VPN services, NAT, DHCP, and so on.

The VMware NSX-T Data Center for VxBlock System design includes, by default, a single Tier-1 gateway (T1-GW-1). This gateway provides a link between the two overlay-backed segments and the Tier-0 gateway that connects to the physical network.

The following figure shows the topology for the Tier-0 and Tier-1 gateways:


T0
100.64.112.0/31
100.64.112.1/31
T1
192.168.128.162/31
192.168.128.170/31
T1-Seg-1
T1-Seg-2
VM
testvm02
192.168.130.17/31
VM
testvm02
192.168.130.17/31

11

VMware NSX-T Data Center transport nodes

Hypervisor transport nodes are prepared and configured for VMware NSX-T Data Center.

The VMware VDS provides network services to the virtual machines that are running on those hypervisors. VMware NSX-T Data Center supports VMware ESXi and KVM hypervisors. The N-VDS that is implemented for KVM is based on the Open vSwitch (OVS) and is platform-independent. The N-VDS can be ported to other hypervisors and serves as the foundation when implementing VMware NSX-T Data Center in other environments (for example, cloud and containers). For the VMware NSX-T Data Center on VxBlock Systems design, Dell Technologies Sales Engineers deploy and support only VMware ESXi based transport nodes. VMware supports other types of transport nodes.

Transport node topology 1

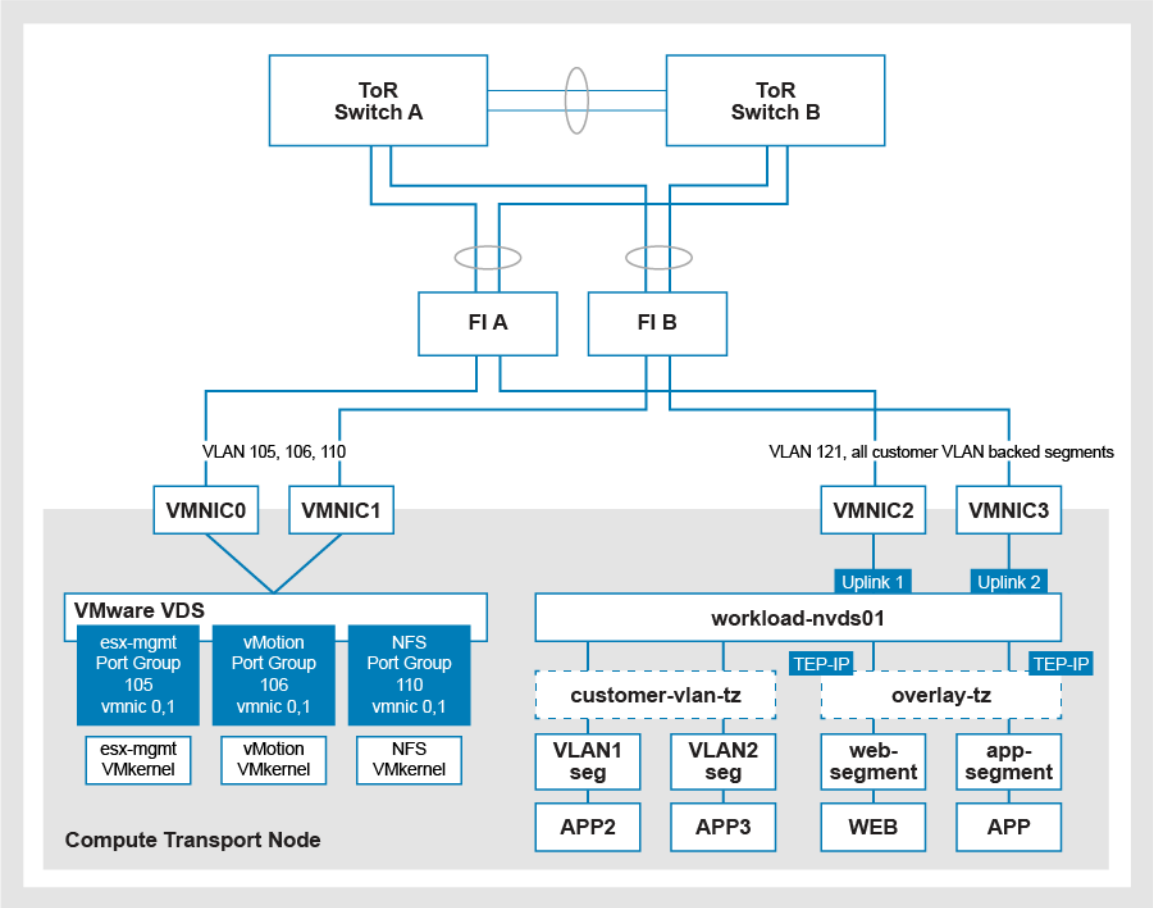
The VMware NSX-T Data Center for VxBlock System transport node topology 1 implementation uses the following criteria:

- One VMware VDS uses VMNIC0 and VMNIC1 for VMware ESXi host functions including port groups and kernels for VMware ESXi management, VMware vSphere vMotion, and NFS.
- There is one VMware N-VDS using VMNIC2 and VMNIC3 for workload VMs.

  The VMware N-VDS is used for east-west traffic with the use of TEPs to create an overlay network.

- Use uplink teaming of source port on the VMware N-VDS to ensure load-balancing.
- Use an MTU value of 9000 on these vNICs to allow for the overhead of GENEVE tunnel encapsulation.
- TEP traffic requires VLAN tagging at the uplink profile. There is no VMware VDS in front of the VMware N-VDS.

The use of the VMware VDS and VMware N-VDS separates the VMware ESXi host functions from VMware NSX-T Data Center traffic and if a failure occurs, troubleshooting and recovery of a host is easier. The following figure illustrates the topology of this design:



In this design, the transport nodes are connected to FI A and B through all four VMNICs.

The VLANs are added to the vNIC templates as follows:

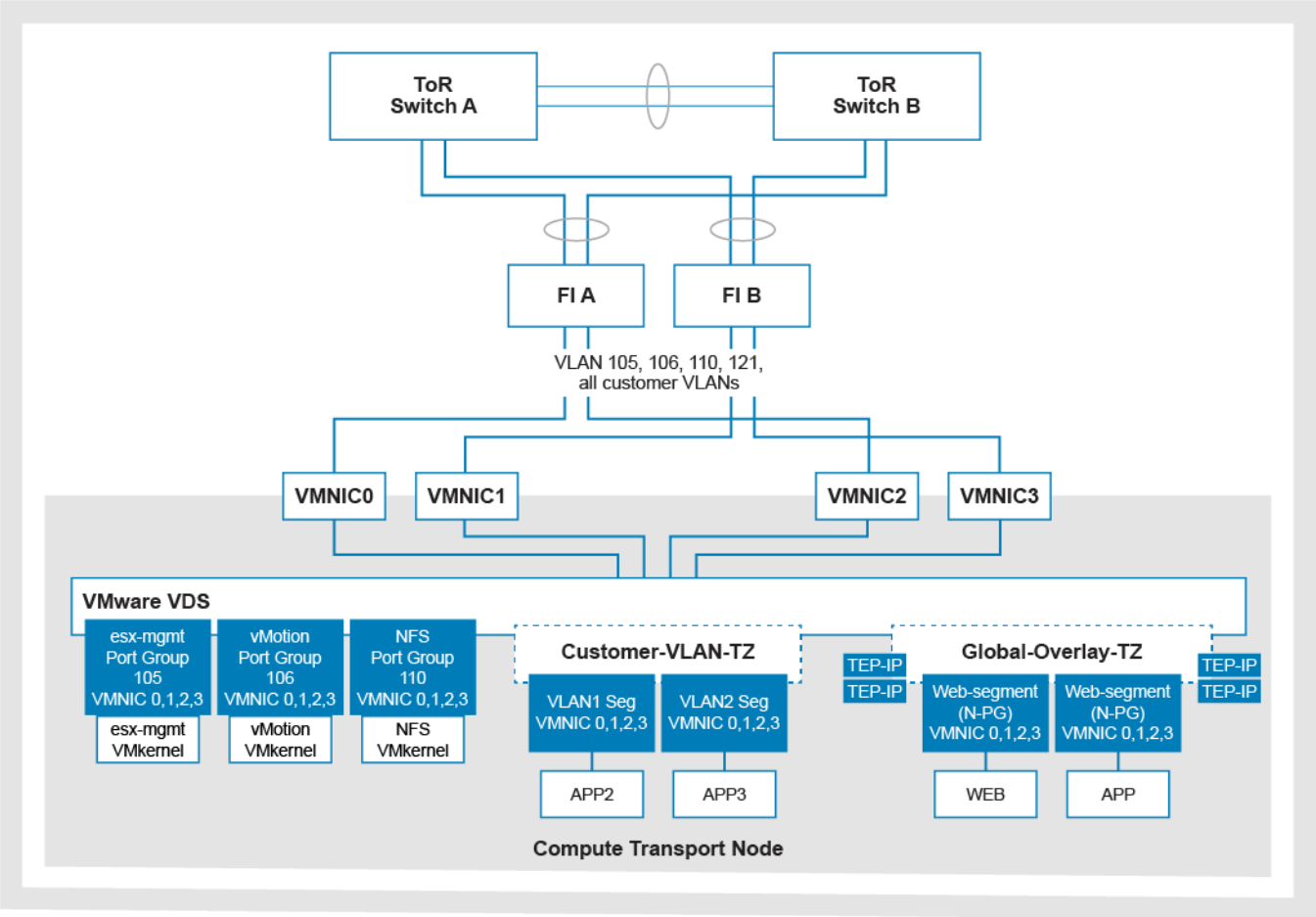
- The VLANs for Overlay and VLAN-backed segments are added to the templates for vNIC2 and vNIC3.
- The VLANs for Management, vMotion, and NFS traffic are added to the templates for vNIC0 and vNIC1.
- Non-NSX site VLANs are not deployed on NSX transport nodes, but participation in NSX can be defined at a cluster level so that some clusters carry non-NSX site VLANs and others carry NSX VLAN and overlay-backed segments.

Transport node topology 2

The VMware NSX-T Data Center for VxBlock System transport node topology 2 implementation uses the following criteria:

- In VMware vSphere 7.0, all port groups and VLANs are under the VMware VDS, including ESXi-mgmt, VMware vSphere vMotion, Uplink1, Uplink2, and Overlay.
- Uplink teaming of source ports on VDS ensures load balancing.
- An MTU value of 9000 on these vNICs allows for the overhead of GENEVE tunnel encapsulation.
- TEP traffic requires VLAN tagging at the uplink profile.

The following figure illustrates the topology that is used for VMware vSphere 7.0 on the transport nodes:



In this design, the transport nodes connect to FI A and B through all four VMNICs. The four-vNIC design aligns with the VxBlock System standard configuration for a VMware vSphere 7.0 compute host, allowing standardization of service profile templates across hosts with single and dual VIC configurations.

The overlay VLAN is added to the vNIC template for vNICs 0, 1, 2, and 3 and the ToR switch A and ToR switch B trunk ports to FIs for the overlay network.

Adapter policy settings for transport nodes

For maximum performance from the VMware NSX-T Data Center transport nodes using Cisco VIC adapters, Cisco recommends modifying the Ethernet Adapter Policy from Cisco UCS manager.

Create an adapter policy to define these settings and apply it to all VNICs on each transport node. The following table provides the settings:

Parameter	Value
Transmit queue	1
Ring size (transmit)	2048
Receive queues	8
Ring size (receive)	2048
Completion queues	9
Interrupts	11
Receive side scaling (RSS)	Enabled
VXLAN offload	Disabled

IP address pool

To provide a mechanism to transmit GENEVE overlay-backed traffic over the physical VLAN backed network, VMware NSX-T Data Center requires a TEP IP address pool. This pool is a range of IP addresses on the VMware NSX-T Data Center transport VLAN that are reserved for TEPs.

Each edge node VM requires one IP address from this pool. For Transport Node Topology 1, each transport node requires two IP addresses also from TEP IP pool. For Transport Node Topology 2, each transport node requires four IP addresses from the TEP IP pool. When sizing the pool, consider the current and future requirements for TEPs.

For VMware NSX-T Data Center deployments with more than one L3 domain (for example, multisystem cross-VMware vCenter Server), you must deploy a separate TEP IP pool for each domain. L3 routing must be configured and working between all TEP IP pools. This configuration enables transport traffic to traverse between transport nodes on different pools.

12

Licensing

Order either the Advanced or Enterprise Plus editions of VMware NSX-T Data Center on the VxBlock System. In VMware NSX-T Data Center, the physical edge hosts and all transport nodes participating in NSX need to be licensed for each physical CPU socket.

See the NSX-T data sheet.

You can license any number of transport nodes (compute hosts). Consider the following:

- The licensed transport nodes can be a subset of the physical hosts in the VxBlock System.
- The licensed transport nodes could be more hosts than physically exist in the system. Licensing more hosts in this way enables you to stretch overlay-backed segments across multiple VxBlock Systems.
- All physical edge hosts must be licensed for NSX, even though they are not prepared as transport nodes.
- All VMware NSX-T Data Center licensing must be purchased directly through VMware. Dell Technologies is unable to resell the licensing for the product due to export compliance issues.

Cisco UCS licensing

The licensing that is required to connect the physical edge hosts to the UCS domain varies depending on the following:

- The type of Cisco UCS domain to which the hosts are connecting
- Whether there is a FEX connected to the FIs

The various connectivity models and their licensing requirements are detailed in the following table:

Connectivity model	Licensing details
Cisco UCS third-generation domain (Cisco UCS 6332-16UP Fabric Interconnects), no FEX	Each QSFP port on a Cisco UCS third-generation FI used for VMware NSX-T Data Center edge direct-connect should be licensed using the direct-connect SKU UCS-LIC-6300-40GC. Two edge hosts require two SKUs. Four, six, or eight hosts require four SKUs. More than eight edge hosts require eight SKUs.
Cisco UCS fourth-generation domain (Cisco UCS 6400 Series Fabric Interconnects), no FEX	Each edge compute host that connects directly to a Cisco UCS fourth-generation FI must be licensed to connect to the Cisco UCS domain. The license is the Cisco UCS C-Series Rack Server only 25 Gbps SKU (UCS-LIC-6400-25GC), quantity two ports per server.
Any Cisco UCS domain with FEX	Any edge compute host that connects to a FEX and not to an FI does not need any additional port licensing.

VMware ESXi licensing

Each physical edge host consumes standard per-socket CPU licensing for VMware ESXi. In this configuration, each host has two physical CPU sockets that are populated, and consumes two VMware ESXi CPU licenses.

Distributed IDS and IPS

Distributed IDS and IPS is a licensed feature which is not included in the traditional NSX per CPU licenses. You must apply the *Add-On NSX Advanced Threat Prevention* license in your NSX-T Manager to enable these capabilities.

Cisco switch Layer 3 licensing

VMware NSX-T Data Center deployments on VxBlock Systems 1000, must license the ToR switch pair for L3 routing services.

Dell EMC offers two license packages that enable this capability:

- Essentials
- Advantage

These packages include the licensing that is required to run VMware NSX-T Data Center on a VxBlock System.

For VMware NSX-T Data Center deployments on VxBlock Systems 340, 350, 540, and 740 Systems, license the switch pair for L3 routing services. License types are different based on the type of ToR switch that is installed in the system.

Depending on the type of ToR switch, install the appropriate license:

Switch	License type
Cisco Nexus 9396PX	LAN Enterprise
Cisco Nexus 93180YC-EX	Essentials Advantage

VxBlock Systems with ToR Cisco Nexus 5000 Series Switches are not supported with VMware NSX-T Data Center.

Cisco switch Layer 3 licensing

VMware NSX-T Data Center deployments on VxBlock Systems 1000, must license the ToR switch pair for L3 routing services.

Dell EMC offers two license packages that enable this capability:

- Essentials
- Advantage

These packages include the licensing that is required to run VMware NSX-T Data Center on a VxBlock System.

For VMware NSX-T Data Center deployments on VxBlock 340, 350, 540, and 740 systems, the switch pair must be licensed for L3 routing services.

The VxBlock 340, 350, 540, and 740 Systems have a ToR Cisco Nexus 93180YC-EX or a ToR Cisco Nexus 9396PX Switch. One of the following licenses must be installed on the switch.

Switch	License type
Cisco Nexus 9396PX	LAN Enterprise
Cisco Nexus 93180YC-EX	Essentials Advantage

VxBlock Systems with ToR Cisco Nexus 5000 Series Switches are not supported with VMware NSX-T Data Center.

13

VMware NSX Intelligence

VMware NSX Intelligence is a distributed analytics engine that is built into VMware NSX-T Data Center. VMware NSX Intelligence provides continuous data center-wide visibility for network and application security teams. VMware NSX Intelligence delivers a more granular and dynamic security posture, simplify compliance analysis, and streamline security operations.

See Using and Managing VMware NSX Intelligence for more information around using VMware NSX Intelligence

A small appliance size can be used for lab or proof-of-concept environment, or a small-scale production environment. A large appliance size is for a large-scale production environment. See the following table for additional information:

Small appliance	Large appliance
16 vCPU	32 vCPU
64 GB RAM	128 GB RAM
2 TB storage	2 TB storage

An NSX-T Data Center Enterprise Plus license is required for VMware NSX Intelligence.

14

VMware NSX Federation

With VMware NSX Federation, you can manage multiple NSX-T Data Center environments with a single view. You can create gateways and segments that span one or more locations. You can configure and enforce firewall rules consistently across locations. After you have installed the Global Manager and have added locations, you can configure networking and security from global manager.

For information about the initial VMware NSX Federation configuration, go to https://www.vmware.com/support/pubs/ and access the *VMware NSX-T Data Center Administration Guide*:

The NSX-T Global managers are added to the Nirvana tool for proper sizing on the AMP. Three global managers are installed with a VIP for load balancing are installed on the AMP. VMware NSX Federation is deployed as a service provided by VMware.

An RTEP VLAN and SVI are added to the ToR switches. The VLAN is trunked down to the NSX-T edge nodes using the VPC links to VMNIC 2 and VMNIC 3.

The following figure shows where RTEP traffic is routed to the NSX-T edge VM: