



VMware® NSX Network Virtualization Design Guide

Deploying VMware NSX with Cisco UCS and Nexus 7000

Table of Contents

Intended Audience.....	3
Executive Summary.....	3
Why deploy VMware NSX on Cisco UCS and Nexus 7000?.....	4
Objective.....	5
Hardware assumptions.....	5
Software requirements.....	5
Design assumptions.....	6
Design constraints.....	6
VLANs.....	7
Transport VLAN.....	7
Cisco UCS hosts with a Single VTEP.....	7
Cisco UCS hosts with VXLAN Multipath and Multiple VTEPs.....	9
Edge VLAN.....	10
Host networking design.....	11
VMware NSX Traffic Flow on Cisco UCS.....	13
Cisco UCS configuration for VMware NSX.....	14
QoS System Class.....	14
Fabric minimum bandwidth.....	14
QoS Policy.....	15
VLANs.....	15
MAC Pools.....	16
Network Control Policy.....	17
vNIC Templates.....	17
LAN Connectivity Policy.....	18
Service Profile Template.....	19
Use your LAN Connectivity Policy.....	19
vNIC Placement (PCI bus ordering).....	19
NSX for vSphere configuration for Cisco UCS.....	20
Assign host VDS uplinks to adapters.....	20
VXLAN traffic Resource Pool.....	21
Virtual Machine Traffic Resource Pool.....	22
Configure VXLAN networking.....	23
Configure the new VXLAN vmknics Port Group.....	24
Verify teaming and failover.....	25
Configure QoS tagging for system traffic.....	26
NSX Edge routing considerations with Cisco Nexus 7000.....	27
Deploying applications and virtual networks.....	28
1) The software-defined data center.....	28
2) Self-service IT network provisioning.....	28
Design Poster.....	28
Addendum.....	29
Cisco UCS and NSX for vSphere design with multiple physical adapters.....	29
Additional Resources.....	29

Intended Audience

This document is targeted toward virtualization and network architects interested in deploying VMware® NSX network virtualization solutions with Cisco Nexus 7000 and Cisco UCS.

Executive Summary

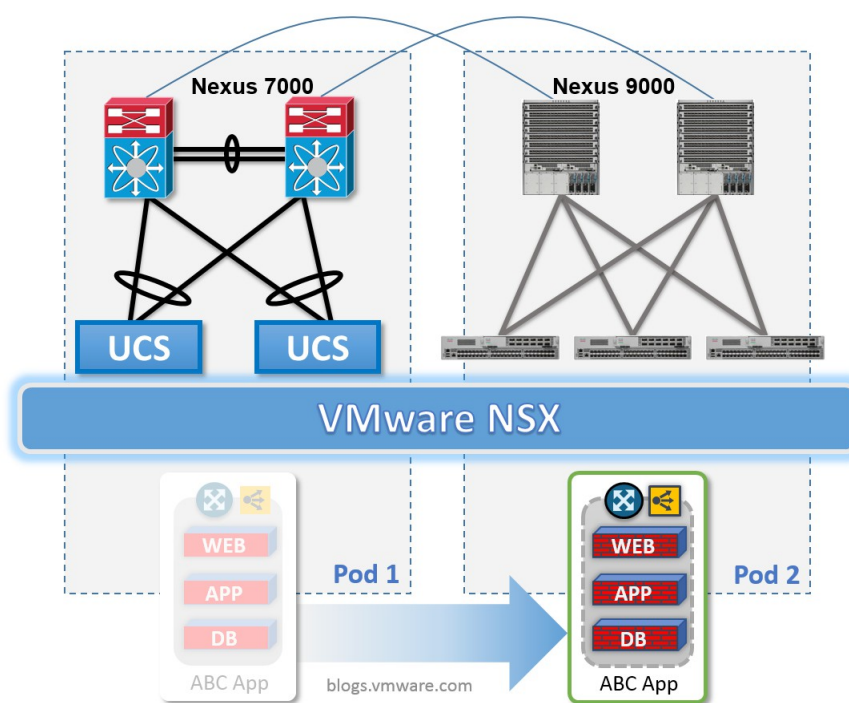
VMware NSX network virtualization software makes it possible for IT organizations to obtain new levels of business agility, allowing you to deploy advanced policy based network services for your applications as quickly as a virtual machine, today, leveraging your existing physical network infrastructure or next generation network hardware from any vendor. This article describes a simple and straight forward example of how to deploy VMware NSX for vSphere on an infrastructure consisting of Cisco's Unified Computing System (UCS) and the Cisco Nexus 7000 series switches. This basic guide should serve as a starting point from which to tailor a design for your specific environment, and the basis for additional reference architectures guides to build upon.

- VMware NSX brings business agility to existing hardware investments
- NSX is a mature software solution available today
- This article provides technical guidance for deploying NSX with Cisco UCS and Cisco Nexus 7000

Why deploy VMware NSX on Cisco UCS and Nexus 7000?

VMware NSX software is a key building block of a software-defined data center (SDDC) approach, enabling automated and policy based virtual networking services for vSphere environments, including those running on existing Cisco UCS and Cisco Nexus 7000 hardware. VMware NSX brings a comprehensive set of L2-L7 network services including Routing, Switching, Perimeter Firewalling, Load Balancing, DHCP, DNS, Monitoring, Security, VPN, and a powerful VM NIC level distributed firewalling; all of which can be programmatically deployed for your applications (in seconds) on your existing Cisco UCS infrastructure running VMware NSX for vSphere.

Moving forward, the VMware NSX virtual network platform can seamlessly extend across pods of dissimilar physical infrastructure, creating a common pool of IT resources and consistent networking service model across multiple disparate pods. When the time is right, this normalization provided by NSX enables a non-disruptive migration to any other physical network infrastructure, such as the Cisco Nexus 9000 series, or something else.



Non-disruptive migration and pooling across any network infrastructure

With network services deployed by VMware NSX in a software based virtualization layer, complex hardware configurations are simplified and by consequence made more scalable, while increasing performance, all by orders of magnitude. For more detail on this refer to Brad Hedlund's blog post: [Seven reasons VMware NSX, Cisco UCS and Nexus are orders of magnitude more awesome together](#).

Note: Other hardware choices also realize the simplicity, scale, performance, and normalization from VMware NSX. This guide is simply addressed to users who already have Cisco UCS and Cisco Nexus 7000 as part of their installed infrastructure.

- NSX brings fully automated L2-L7 networking services to vSphere on Cisco UCS
- NSX adds a powerful VM NIC level distributed firewalling to applications on Cisco UCS
- NSX simplifies the Cisco UCS and Nexus 7000 hardware configurations
- NSX abstracts and normalizes the differences between disparate infrastructure pods
- NSX facilitates a non-disruptive migration to next generation hardware platforms

Objective

Whether you're starting with a new installation or working with an existing environment, there are three basic milestones towards attaining network virtualization with VMware NSX:

1. Prepare the physical infrastructure configuration (done one time)
2. Install and configure VMware NSX network virtualization software (done one time)
3. Programmatically create virtual networks (ongoing)

The objective of this guide is to help you understand the design and configuration steps to reach milestones #1 and #2, the one-time initial configuration of your Cisco UCS, Nexus 7000, and VMware vSphere components. After a few hours of initial configuration described in this guide, you'll be ready to start creating dynamic L2-L7 virtual networks in seconds with cloud automation tools such as vCloud Automation Center (vCAC), or with self-service GUI based workflows in the vSphere Web Client without any changes to the initial infrastructure configurations.

Hardware assumptions

This guide was written specifically for environments where VMware NSX for vSphere is running on Cisco UCS servers (blade or rack), connected to current generation 61xx or 62xx series UCS fabric interconnects, connected to current generation Nexus 7000 series data center switches. The example in this guide shows (4) Cisco UCS servers, (2) Nexus 7000 switches, and (2) UCS Fabric Interconnects.

- (2) Cisco Nexus 7000 data center switches
- (4) Cisco UCS rack or blade servers with 1 or 2 physical adapters
- (2) Cisco UCS 6100 or 6200 series fabric interconnects

This guide assumes Cisco UCS servers with a single physical adapter. Configurations with multiple physical adapters are possible and shown as an addendum.

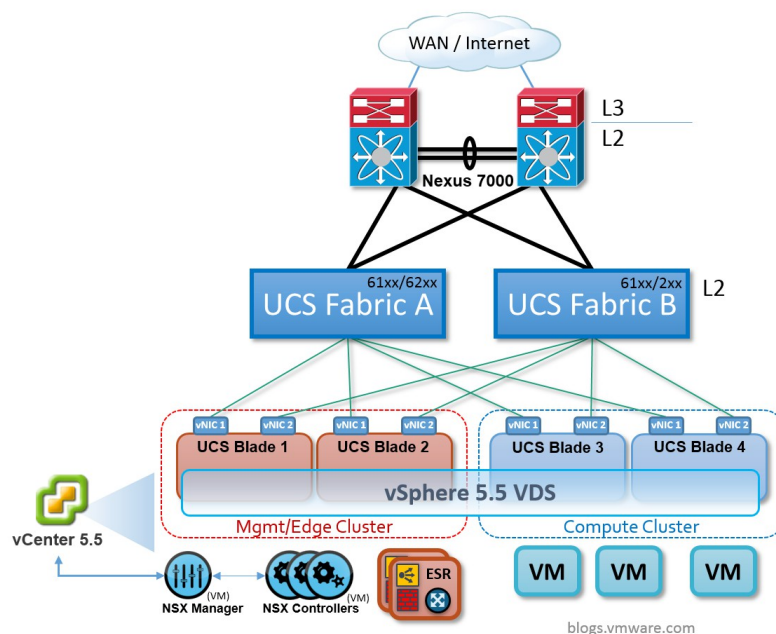
Software requirements

This guide was written specifically for VMware NSX for vSphere, whereby the only prerequisite is VMware vSphere (ESXi) 5.5 host hypervisor software on the selected Cisco UCS servers, managed by VMware vCenter 5.5. With this in place you can deploy the VMware vSphere NSX Manager OVA (version 6.0 or greater). The NSX Manager is packaged with, and will automatically install, the additional requisite NSX components, such as NSX controllers and ESXi host kernel modules. For Cisco UCS, you should be running Cisco UCS Manager 2.0 or greater on your fabric interconnects.

- vSphere 5.5
- vCenter 5.5
- vSphere NSX Manager 6.0 or greater
- Cisco UCS Manager 2.0 or greater

Design assumptions

In this guide we assume that you are starting with a pair of Cisco UCS fabric interconnects operating in a single redundant domain. One interconnect represents Fabric A, the other represents Fabric B. Each fabric interconnect is operating in End Host Mode with a virtual Port Channel (vPC) uplink landing evenly on a pair of Cisco Nexus 7000 switches.



Design assumption and starting point

The Cisco Nexus 7000 switches provide Layer 2 connectivity between Fabric A and Fabric B, as well as providing the Layer 3 boundary for the infrastructure VLANs. The Nexus 7000s might also run an IP routing protocol (such as OSPF or BGP) to both learn and advertise IP networks with the Enterprise WAN or Internet. This is a pretty common setup.

For the vSphere 5.5 environment we show a deployment with at least four Cisco UCS servers, two vSphere clusters, and one vSphere VDS. One cluster is called "Management & Edge" for deploying the NSX management and Edge Services Router virtual machines (L3-L7 services). The second cluster is called "Compute" for deploying the virtual machines for applications. Your vSphere design doesn't need to follow this exact template.

The design presumes that both FCoE and IP Storage may exist, but doesn't require it.

Design constraints

Cisco Nexus 7000 switches do not support dynamic IP routing with peers on vPC VLANs. We will need to give extra consideration to this fact when we want to design for dynamic IP routing protocols between our NSX routers and the Nexus 7000 switches.

Cisco UCS fabric interconnects do not support LACP adapter teaming for UCS server vNICs. This means that we will not be able to choose LACP as the uplink teaming policy for the Port Groups containing the vSphere vmkernel NICs for traffic types such as vMotion, IP Storage, and VXLAN. Instead, we can create a multipath design using multiple vmkernel NICs (two for each traffic type), or we can choose the fail over teaming policy with a single vmkernel NIC in a basic active-passive design.

VLANs

The Cisco UCS and Nexus 7000 switches will start with three infrastructure VLANs that you would normally deploy in a typical VMware vSphere environment, with or without NSX.

- VLAN 100: Management
- VLAN 110: vMotion
- VLAN 120: IP Storage

The three VLANs above will be forwarded to every vSphere host in both clusters, because each vSphere host will have three vmkernel NICs, one each for Management, vMotion, and IP Storage, attached to the above VLANs respectively.

VLAN 1 is also present as the default untagged VLAN on all links. The VLAN numbers are examples in this guide. Yours may vary.

To prepare for VMware NSX, we will add two additional infrastructure VLANs to the Cisco UCS and Nexus 7000 configurations.

- VLAN 200: Transport
- VLAN 300: Edge DMZ

The Cisco Nexus 7000 switches should have a Layer 3 switched virtual interface (SVI) and HSRP configured on the above five VLANs to provide the IP default gateway for the vSphere host vmkernel NICs and NSX routers attached to these networks.

Five VLANs in total is all that is necessary for the Cisco UCS fabric interconnects and Cisco Nexus 7000 switches. **This one time initial configuration is enough for VMware NSX to dynamically create thousands of logical networks each with Routing, Load Balancing, Firewalling, security, monitoring, isolation, and multi-tenancy.**

From this point, VMware NSX obviates the need to manually create ongoing configurations like Port Groups, VLANs, VRFs, default gateways, ACLs, Load Balancer policies, Firewall rules, and so on, as you add virtual machines and applications to this infrastructure. The initial configuration is sufficient as NSX dynamically creates and applies advanced networking services and policies in software at the distributed virtualization layer.

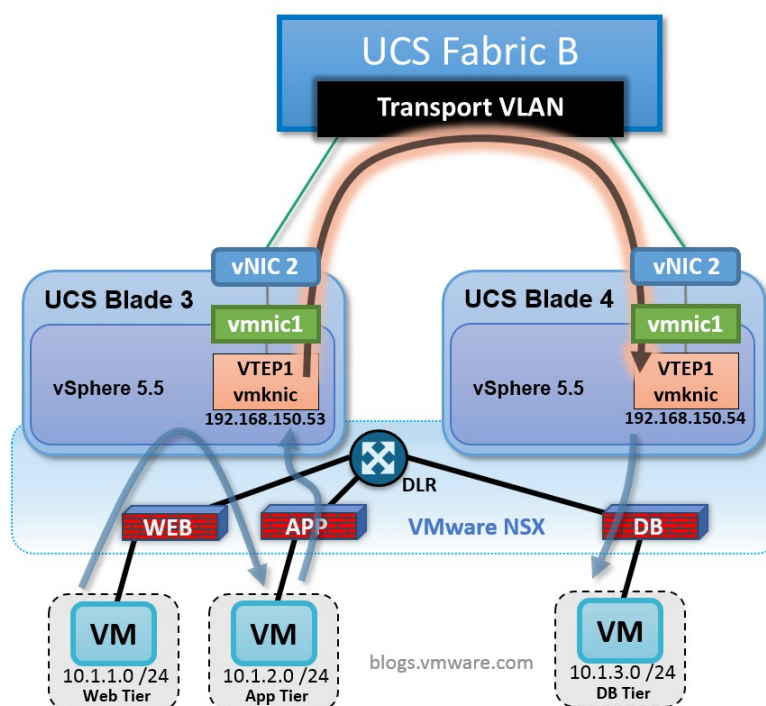
Transport VLAN

After the NSX installation is complete, each vSphere host will have a new vmkernel NIC specifically used by NSX as the VXLAN Tunnel End Point (VTEP). When virtual machines on different hosts are attached to NSX virtual networks and need to communicate, the source host VTEP will encapsulate VM traffic with a standard VXLAN header and send it to the destination host VTEP over the Transport VLAN. Each host will have one VTEP, or multiple VTEPs, depending on the VXLAN vmknics teaming policy you've chosen.

- LACP = Single VTEP using multiple uplinks
- Fail Over = Single VTEP using one uplink
- Load Balance = Multiple VTEPs each using one uplink

Cisco UCS hosts with a Single VTEP

Based on the design constraints we already know that Cisco UCS adapter vNICs cannot establish LACP connections to the UCS fabric interconnects. This means that in a design with a single VTEP vmknics per vSphere host, the teaming policy we must choose is Fail Over. As a result, all host-to-host virtual machine traffic encapsulated in VXLAN will use the bandwidth of one UCS vNIC and one UCS fabric, not both.



Host to host virtual machine traffic on the Transport VLAN with a single VTEP

In the single VTEP design, each UCS vSphere host will have only one VTEP vmknic, and we will place it on the Transport VLAN. Any virtual machine traffic between hosts will egress the source host from a single VTEP, and ingress on a single VTEP at the destination host.

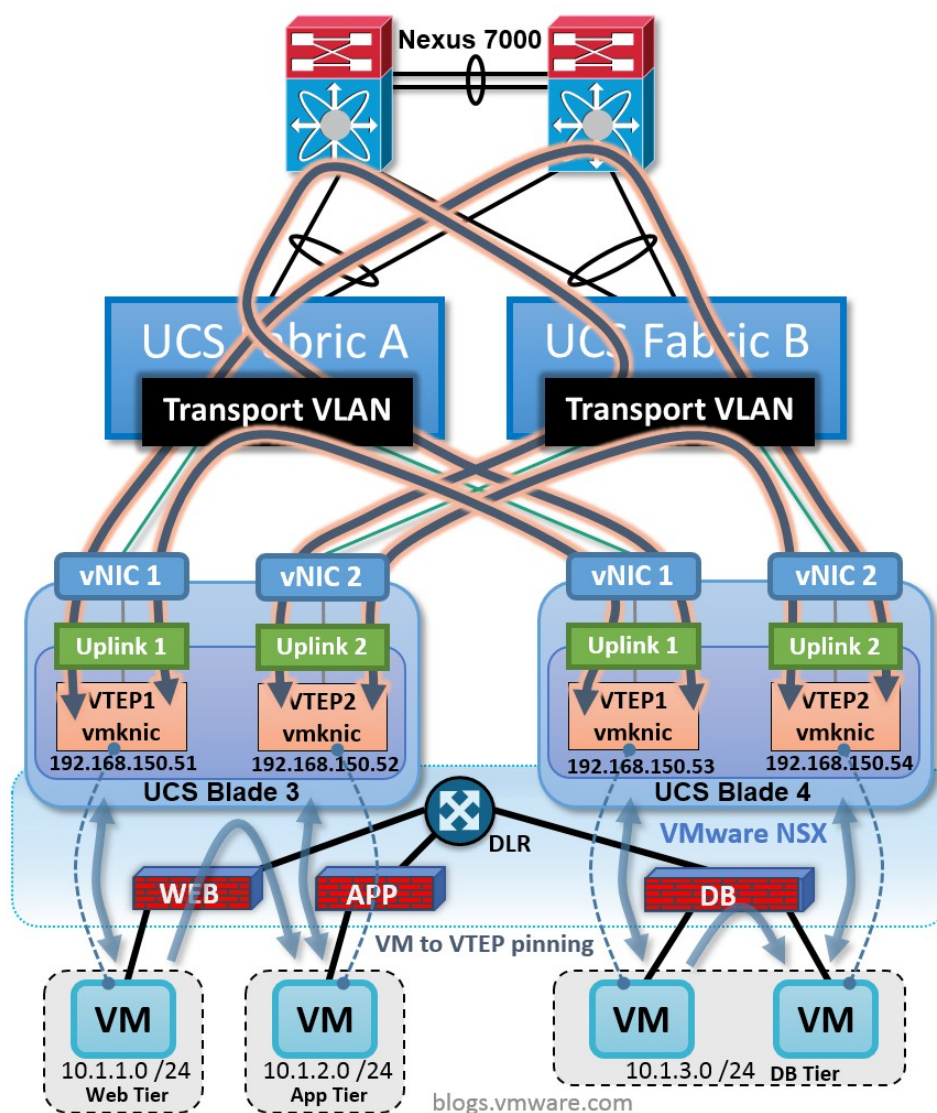
In any design it's important for the UCS adapter vNICs to consistently map to the same UCS fabric and the same VDS uplinks across all hosts, especially in situations where you have a single VTEP with fail over teaming. For example, vNIC 1 should always connect to Fabric A and correspond to VDS Uplink 1, whereas vNIC 2 should always connect to Fabric B and correspond to VDS Uplink 2. This way, *any* virtual machine traffic between any two host VTEPs will be attached to the same UCS fabric and forwarded directly by the Layer 2 fabric interconnect, the most efficient path possible, without any unnecessary cross-fabric hops through the Nexus 7000s.

To assure this consistency we can use template oriented deployment tools such as Cisco UCS vNIC Templates, Cisco UCS vNIC placement policies, and VMware vSphere Host Profiles.

Cisco UCS hosts with VXLAN Multipath and Multiple VTEPs

VMware NSX for vSphere allows a host to have multiple VTEP vmknics for situations where you want to use the bandwidth of multiple adapters for VXLAN traffic, and LACP adapter teaming is not an option.

VMware NSX on Cisco UCS with VXLAN Multipath



When you choose Load Balance as the VXLAN vmknics teaming policy, one VTEP will be created for (and pinned directly to) every VDS uplink. This means on a Cisco UCS host with two vNICs two VTEPs will be created; VTEP 1 will be pinned to VDS Uplink 1, while VTEP 2 will be pinned to VDS Uplink 2. Each VTEP will be assigned its own IP address on the Transport VLAN.

With each host having multiple virtual machines and multiple VTEPs, the obvious question becomes; how is the host-to-host virtual machine traffic assigned to a VTEP? The answer should sound familiar to anyone who knows how VMware virtual switches have always load balanced virtual machine traffic without LACP: by source MAC or source virtual port. Virtual machines are essentially pinned to a particular VTEP. Meanwhile, the NSX Controllers maintain consistency in the VM-to-VTEP mapping tables across each host.

With VXLAN Multipath, both UCS fabrics are utilized for *any* virtual machine traffic between hosts. The specific fabric used and the path taken for this traffic depends on which pair of VMs are communicating and

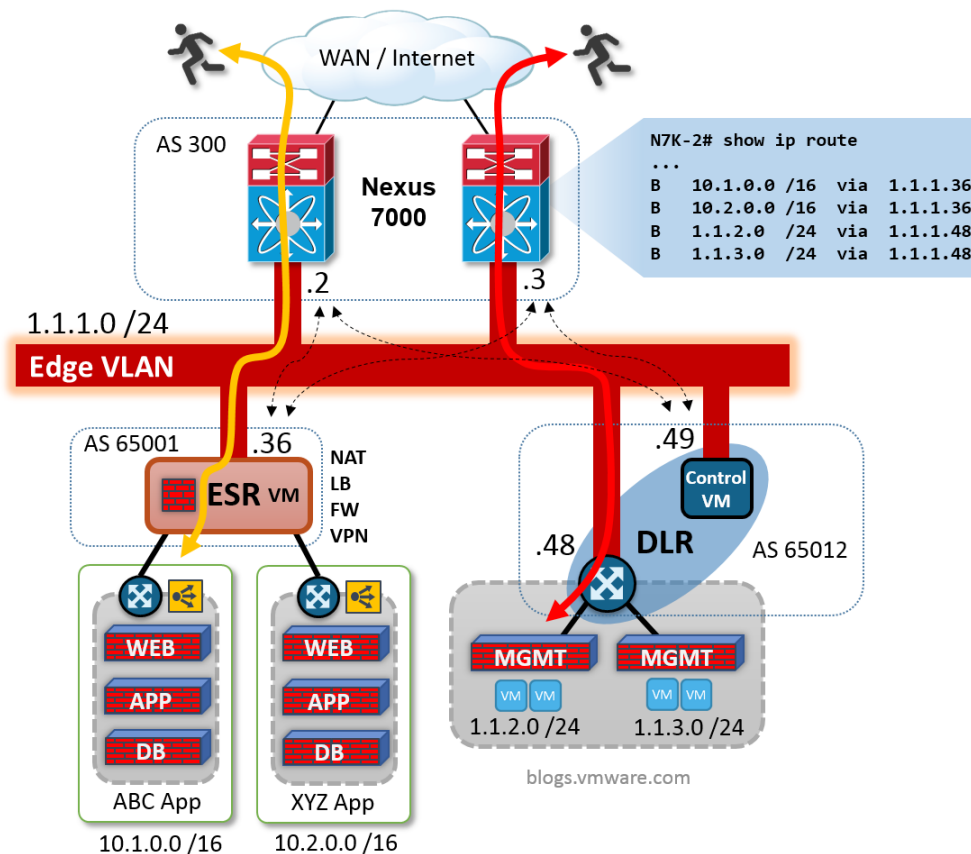
the VTEPs they're pinned to. For example, if two virtual machines are communicating, each pinned to VTEP 1 on VDS Uplink 1, this traffic will be forwarded directly by the Layer 2 fabric interconnect (Fabric A) on the Transport VLAN.

On the other hand, if two virtual machines are communicating, one pinned to VTEP 1 on VDS Uplink 1, the other pinned to VTEP 2 on VDS Uplink 2, this traffic will be forwarded between Fabric A and Fabric B by a Nexus 7000 switch on the Transport VLAN.

With either the single path or multipath design, VTEP-to-VTEP traffic is Layer 2 in the sense that both the source and destination VTEP are communicating on the same VLAN, the Transport VLAN. It's important to understand that while VMware NSX does not require VTEPs to exist on the same VLAN, in this guide we have designed it this way intentionally for the Layer 2 only Cisco UCS fabric interconnects.

Edge VLAN

The Edge VLAN will provide a common segment for the NSX Edge Services Routers (ESR) and Distributed Logical Routers (DLR) to be Layer 2 adjacent with the Cisco Nexus 7000 switches for the purpose of routing IP traffic to and from the Enterprise WAN or Internet into the NSX virtual networks. The Cisco Nexus 7000 switches will each have a Layer 3 switched virtual interface (SVI) on the Edge VLAN and from this will see the NSX routers on this segment as standard IP next-hop routing devices with which to mutually establish dynamic or static IP routing.



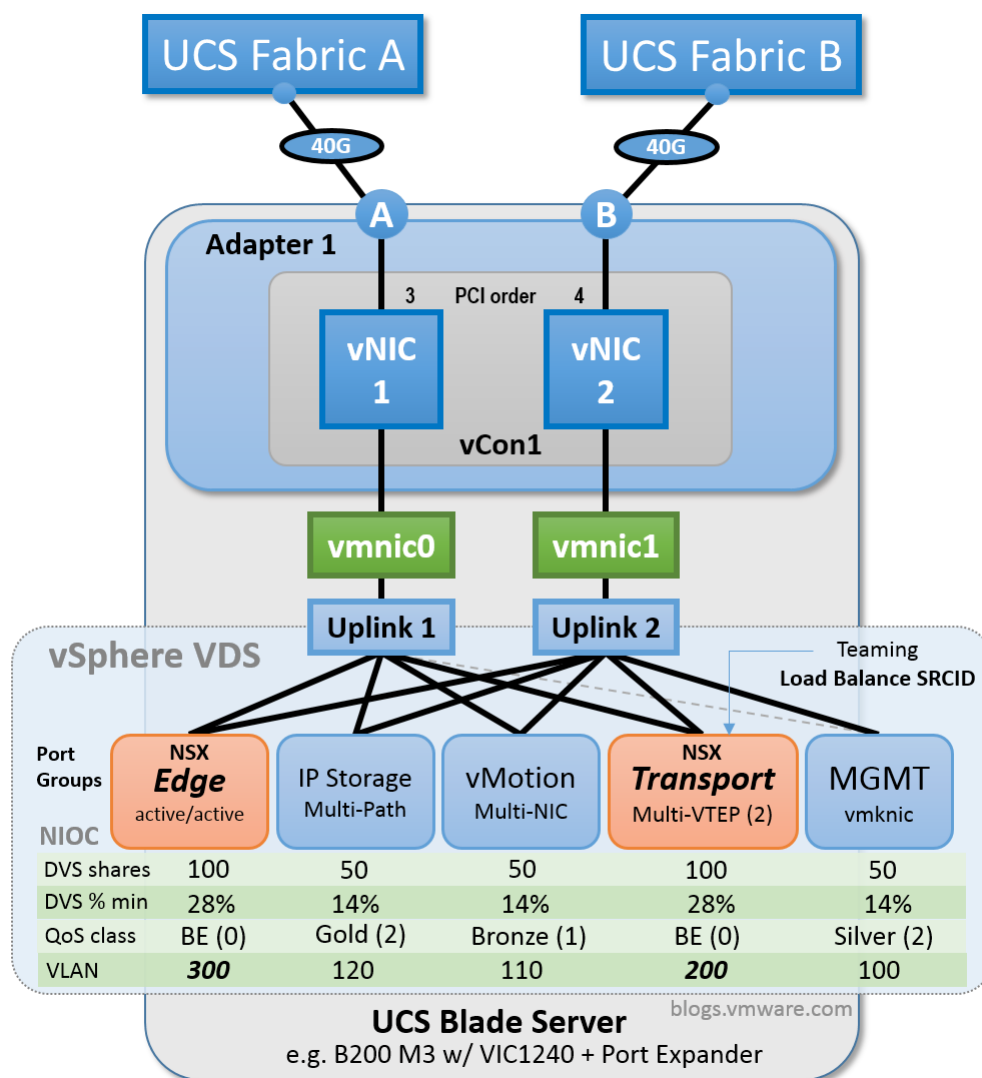
BGP routing on the Edge VLAN

Note: You can certainly deploy more than one Edge VLAN. For example, you might want to have separate Edge networks for external Internet vs internal Enterprise facing applications.

Host networking design

VMware NSX for vSphere leverages the existing VMware Virtual Distributed Switch (VDS), with NSX adding kernel modules and agents that enable the advanced networking capabilities. Those already experienced in designing infrastructure with VMware virtual switches can approach their designs much in the same way, just with an additional system traffic type and vmknic to include, VXLAN and its VTEP vmknic.

In this design our goals are to achieve Active-Active forwarding across both Cisco UCS fabrics, especially for the host-to-host VXLAN virtual machine traffic, as well as vMotion and IP Storage, with a simple and repeatable configuration for every host. To that end, we will provision the Cisco UCS host physical adapter with just two vNICs mapping to two VMware VDS uplinks. Multiple vmknics will be used for VXLAN, vMotion, and IP Storage for fabric load balancing, and NIOC will be enabled providing QoS for all traffic types. Management traffic will have a standard active-passive configuration with a single vmknic.



VMware NSX for vSphere and Cisco UCS single adapter design

The VMware VDS has two uplinks (Uplink 1, Uplink 2) mapped respectively to two vSphere host adapters (vmnic0, vmnic1).

- VDS Uplink 1 maps to vmnic0
- VDS Uplink 2 maps to vmnic1

The Cisco UCS physical adapter has two vNICs provisioned (vNIC 1, vNIC 2) and mapped respectively to two Cisco UCS fabric interconnects (Fabric A, Fabric B).

- UCS vNIC 1 maps to Fabric A
- UCS vNIC 2 maps to Fabric B

To achieve consistent, deterministic traffic forwarding, we want a given VDS uplink to always map to the same Cisco UCS fabric interconnect, with the opposite VDS uplink mapping to the opposite fabric. Meaning, any traffic transmitted on Uplink 1 should always be forwarded by Fabric A. Similarly, any traffic transmitted on Uplink 2 should always be forwarded by Fabric B.

To ensure this consistency, it's important that we define the PCI bus ordering of the UCS adapter vNICs at provisioning time in UCS Manager. Cisco UCS Manager allows you define the order in which vNICs will be found by a host operating system scanning the PCI bus for hardware. With vNIC 1 ordered before vNIC 2, we will achieve the following consistency on each vSphere host.

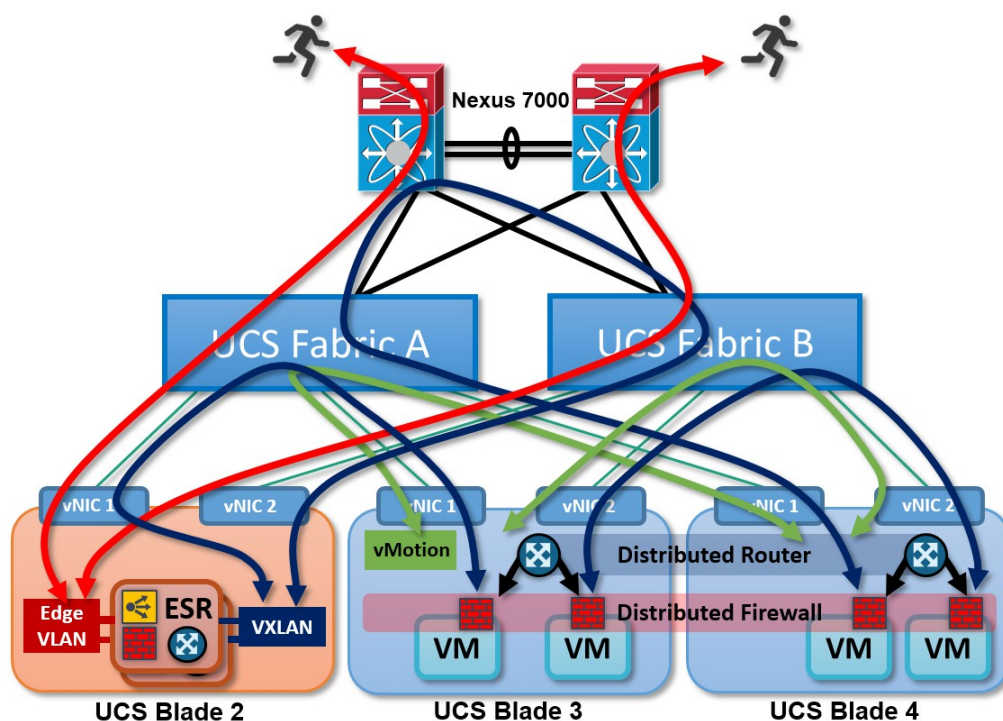
- UCS vNIC 1 recognized by vSphere as vmnic0
- UCS vNIC 2 recognized by vSphere as vmnic1

As you can see in the diagram above, we have the normal vSphere traffic types with their Port Groups including Management, vMotion, and IP Storage. The presence of NSX doesn't require that you change how these standard vSphere traffic types are handled. The vSphere vMotion and IP Storage traffic is capable of multipath and you can configure those settings as normal. You can use the suggestions provided here or follow your own preference.

We've already discussed how the installation of VMware NSX automatically adds an additional vSphere traffic type and Port Group including a new vmkernel NIC for VXLAN. We will also manually add a regular VDS Port Group called "Edge" to be used when we deploy the NSX service router virtual machines with an interface on the Edge VLAN. You can configure the "Edge" Port Group with active-active teaming (e.g. source virtual port) such that some NSX router virtual machines will be located on Fabric A for their Edge VLAN interface, others on Fabric B.

For these new Port Groups for NSX (shown in orange), both the inter-host virtual machine traffic and the external Edge Services north-south traffic will utilize all available bandwidth on each UCS fabric.

VMware NSX Traffic Flow on Cisco UCS



blogs.vmware.com

VMware NSX for vSphere multipath traffic flow on Cisco UCS and Nexus 7000

Finally, it's important to make sure each traffic type has fair access to bandwidth at the host adapters, as well as within the fabric. For this, we will use the QoS capabilities included in both the VMware VDS and Cisco UCS fabric.

In this design, VMware Network I/O Control (NIOC) is enabled on the VDS to provide fair access for all traffic types to the bandwidth of the UCS host adapter. Here, we have a simple allocation of 100 shares for virtual machine related traffic (VXLAN, and Edge), and 50 shares for the other system traffic types. Additionally, we have NIOC configured to tag the COS field of packet headers in accordance with the Cisco UCS Manager QoS policy; thus providing each traffic type fair access to bandwidth within the UCS fabric.

- VM traffic = Best Effort, Class 0
- vMotion = Bronze, Class 1
- Management = Silver, Class 2
- IP Storage = Gold, Class 4

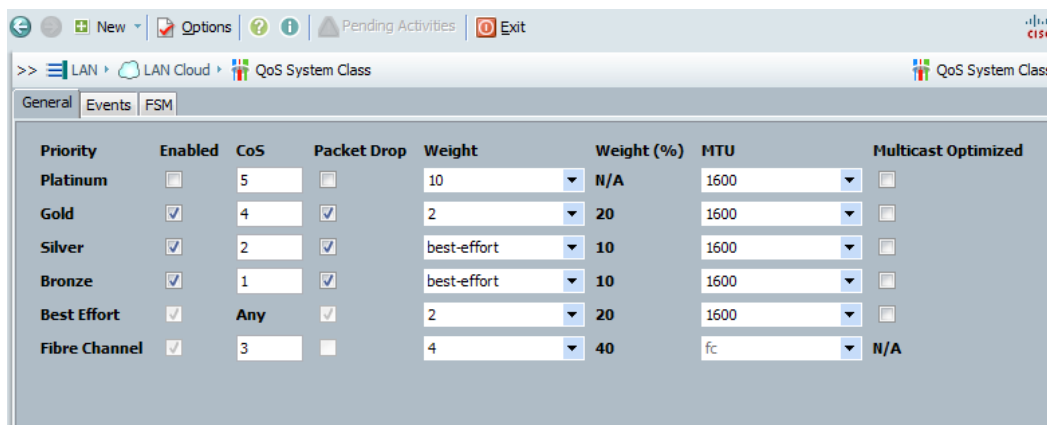
The VMware VDS is also capable of traffic classification, marking, and rate limiting policies. This makes it possible to classify some VM traffic into a QoS class other than Best Effort, if necessary.

Note: Class 3 is presumed to be used for FCoE traffic if present.

Cisco UCS configuration for VMware NSX

In this section we'll cover the important configuration steps in Cisco UCS Manager relevant to running VMware NSX on Cisco UCS. We assume that you have some familiarity with Cisco UCS Manager and have a functional environment up and running.

QoS System Class



Cisco UCS QoS System class settings for VMware NSX for vSphere

Above we can see the suggested configuration of Cisco UCS Manager QoS System class settings. Most importantly, the MTU for each class has been set to 1600 bytes to account for the added VXLAN header. While VXLAN encapsulated VM traffic should only exist within the Best Effort class, the 1600 byte MTU has been configured for all traffic classes just in case a policy is configured later on that places VXLAN encapsulated VM traffic into a different QoS class.

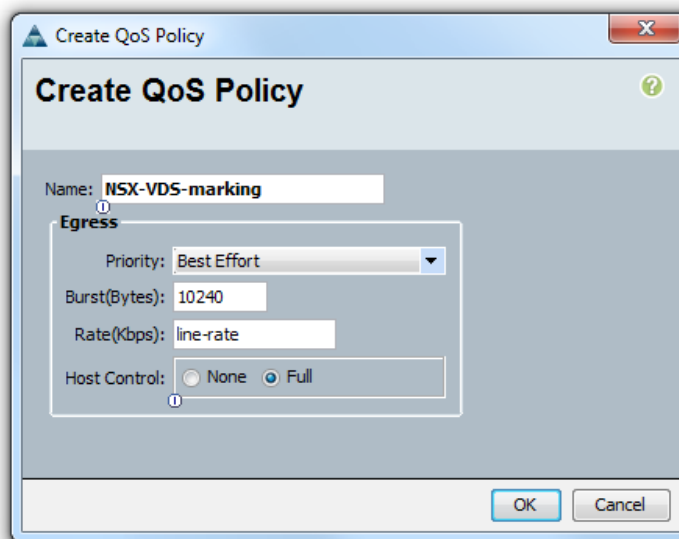
Fabric minimum bandwidth

- Best Effort (VM traffic) = 20%
- Bronze (vMotion) = 10%
- Silver (Management) = 10%
- Gold (IP Storage) = 20%
- FCoE = 40%

This configuration setting establishes the Best Effort, Bronze, Silver, and Gold traffic classes and provides a minimum bandwidth weighting for each. The weightings shown are offered here as a suggestion, not a requirement. You might decide on something different.

Note that the previously discussed VMware VDS NIOC configuration will mark traffic at the source host, which in turn will determine the QoS class to which each traffic type will be assigned when traversing the UCS fabric.

QoS Policy



Cisco UCS QoS policy for VMware NSX

The VMware VDS will be responsible for marking the vSphere traffic at the source host, specifically Management, vMotion, VM traffic, and IP Storage. In doing so, we need to configure a policy in Cisco UCS that allows for this and assign it to the VNIC Templates we will create later on. Otherwise, the default behavior of the Cisco UCS adapter will be to ignore and delete QoS markings it receives from the host operating system.

- Set Host Control = FULL

VLANs

- VLAN 100: Management
- VLAN 110: vMotion
- VLAN 120: IP Storage
- VLAN 200: Transport
- VLAN 300: Edge

Cisco UCS VLANs for VMware NSX

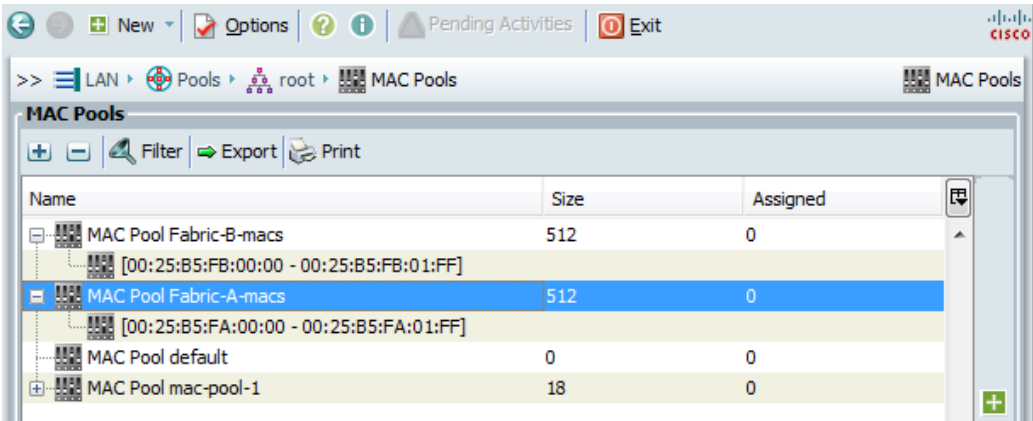
Name	ID	Type	Transport	Native
VLAN Edge_DMZ (300)	300	Lan	Ether	No
VLAN IP_Storage (120)	120	Lan	Ether	No
VLAN Management (100)	100	Lan	Ether	No
VLAN Transport (200)	200	Lan	Ether	No
VLAN default (1)	1	Lan	Ether	Yes
VLAN vMotion (110)	110	Lan	Ether	No

The five VLANs above should be configured in both the Cisco UCS and Nexus 7000 switches, and will be forwarded to every vSphere host in both clusters.

The Management VLAN carries the usual vSphere host management traffic as well as the NSX management traffic from the NSX Manager and NSX Controllers, using one IP address on the host for all management traffic.

MAC Pools

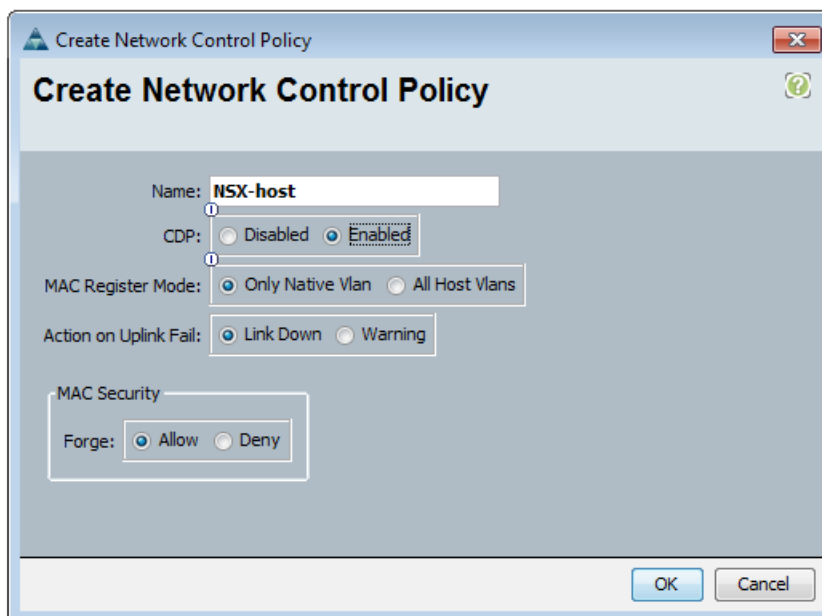
Create two custom pools of MAC addresses to be assigned to vNIC 1 and vNIC 2 as they are provisioned on the physical adapter for each host. One pool will be used to assign a MAC address for vNIC 1 on Fabric A. Similarly, the second pool will assign MAC addresses for vNIC 2 on Fabric B.



Cisco UCS MAC Pools for vNICs

The MAC addresses allocated from these pools will be visible on the vSphere host adapters vmnic0 and vmnic1. This provides a visual cue within vSphere confirming our intended mappings of vmnic0 to Fabric A, and vmnic1 to Fabric B.

Network Control Policy



Create Network Control Policy

Name: **NSX-host**

CDP: ☐ Disabled ☒ Enabled

MAC Register Mode: ☒ Only Native Vlan ☐ All Host Vlans

Action on Uplink Fail: ☒ Link Down ☐ Warning

MAC Security

Forge: ☒ Allow ☐ Deny

OK Cancel

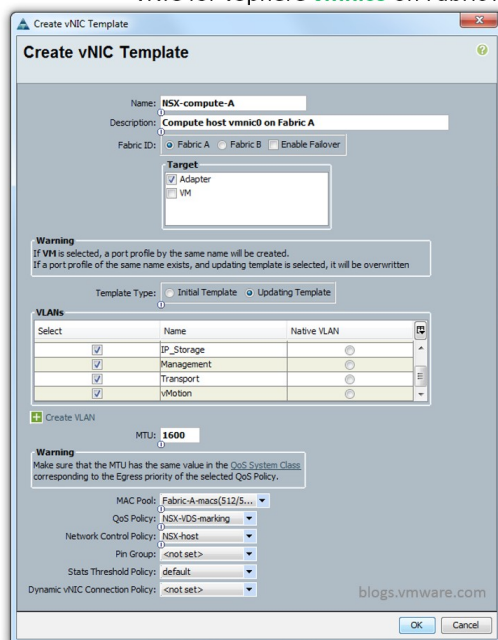
Network Control policy enabling CDP

Create a Network Control Policy that will allow the Cisco UCS fabric interconnect to transmit CDP packets to the vSphere hosts.

vNIC Templates

Create two vNIC Templates in Cisco UCS manager to be used as configuration templates for the Cisco UCS server adapter on each host. One for template for vNIC 1 on Fabric A, the other for vNIC 2 on Fabric B.

vNIC for vSphere **vmnic0** on Fabric A



Create vNIC Template

Name: **NSX-compute-A**

Description: **Compute host vmnic0 on Fabric A**

Fabric ID: ☒ Fabric A ☐ Fabric B ☐ Enable Fallover

Target: ☒ Adapter ☐ VM

Warning
If VM is selected, a port profile by the same name will be created.
If a port profile of the same name exists, and updating template is selected, it will be overwritten

Template Type: ☐ Initial Template ☐ Updating Template

Select	Name	Native VLAN
<input checked="" type="checkbox"/>	IP_Storage	
<input checked="" type="checkbox"/>	Management	
<input checked="" type="checkbox"/>	Transport	
<input checked="" type="checkbox"/>	vMotion	

Create VLAN

MTU: **1600**

Warning
Make sure that the MTU has the same value in the QoS System Class corresponding to the Egress priority of the selected QoS Policy.

MAC Pool: **Fabric-A-macs(12/S...**

QoS Policy: **NSX-VDS-marking**

Network Control Policy: **NSX-host**

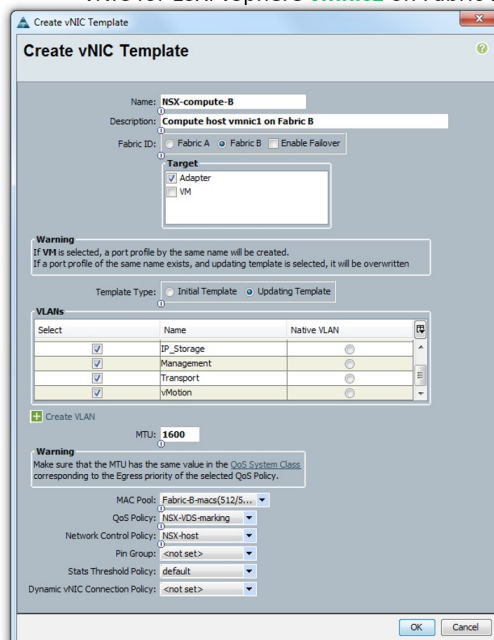
Pin Group: **<not set>**

Stats Threshold Policy: **default**

Dynamic vNIC Connection Policy: **<not set>**

OK Cancel

vNIC for ESXi vSphere **vmnic1** on Fabric B



Create vNIC Template

Name: **NSX-compute-B**

Description: **Compute host vmnic1 on Fabric B**

Fabric ID: ☐ Fabric A ☒ Fabric B ☐ Enable Fallover

Target: ☒ Adapter ☐ VM

Warning
If VM is selected, a port profile by the same name will be created.
If a port profile of the same name exists, and updating template is selected, it will be overwritten

Template Type: ☐ Initial Template ☐ Updating Template

Select	Name	Native VLAN
<input checked="" type="checkbox"/>	IP_Storage	
<input checked="" type="checkbox"/>	Management	
<input checked="" type="checkbox"/>	Transport	
<input checked="" type="checkbox"/>	vMotion	

Create VLAN

MTU: **1600**

Warning
Make sure that the MTU has the same value in the QoS System Class corresponding to the Egress priority of the selected QoS Policy.

MAC Pool: **Fabric-B-macs(12/S...**

QoS Policy: **NSX-VDS-marking**

Network Control Policy: **NSX-host**

Pin Group: **<not set>**

Stats Threshold Policy: **default**

Dynamic vNIC Connection Policy: **<not set>**

OK Cancel

Cisco UCS vNIC settings for VMware NSX for vSphere (Click to enlarge)

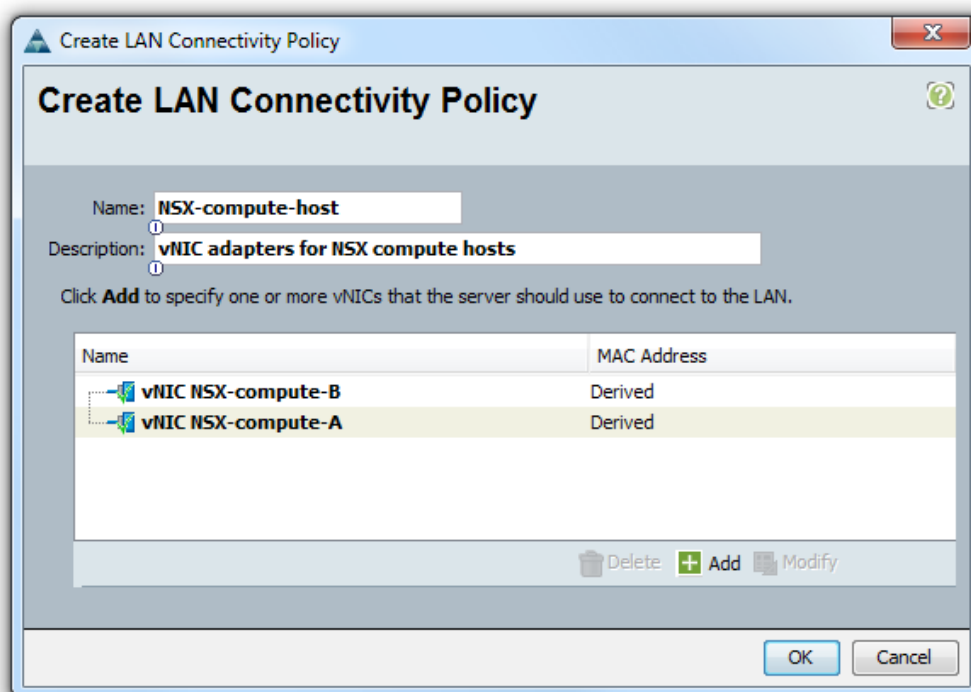
Important vNIC settings:

- Give the vNIC a name that includes the fabric it's been assigned to, such as "NSX-compute-A" for the vNIC assigned to Fabric A
- Leave "Enable Failover" unchecked
- Select the five VLANs
- Select VLAN 1 as the Native VLAN
- Set the MTU to 1600
- Select the MAC pool created for the fabric this vNIC is assigned
- Select the QoS policy
- Select the Network Control Policy

You can create one set of vNIC Templates for hosts in the Management & Edge and Compute clusters.

The vNIC Template will be where we assign the QoS policy and MAC Pools created earlier, as well as setting the MTU to 1600 Bytes to account for the VXLAN headers added to frames for virtual machine traffic.

LAN Connectivity Policy



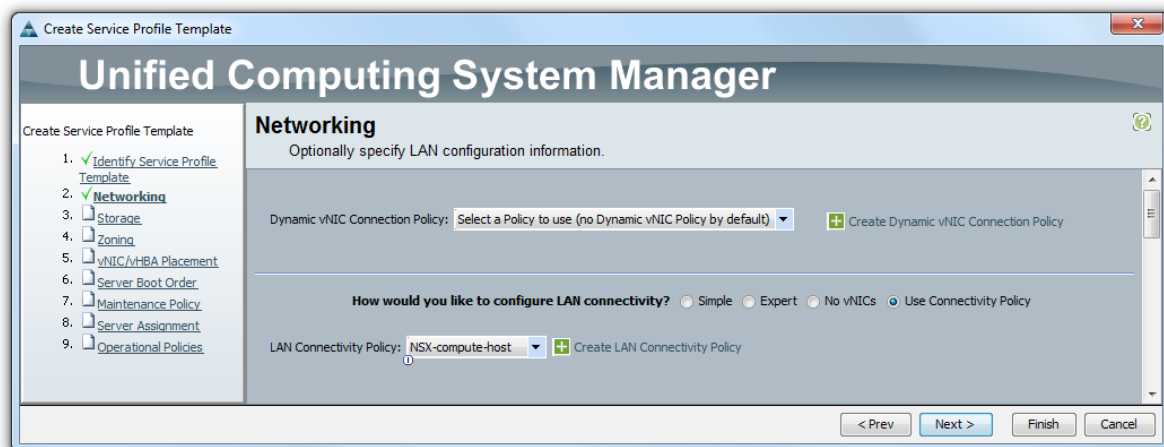
Cisco UCS LAN connectivity policy

After the two vNIC Templates are successfully created, the next step is to create a LAN Connectivity Policy that includes the two vNIC Templates. This will allow us to easily associate the vNIC Templates to the Service Profiles for the vSphere hosts.

Service Profile Template

At this point you are ready to create a Service Profile Template that will include (among other things) the network adapter configuration of the UCS servers that will become vSphere hosts.

Use your LAN Connectivity Policy

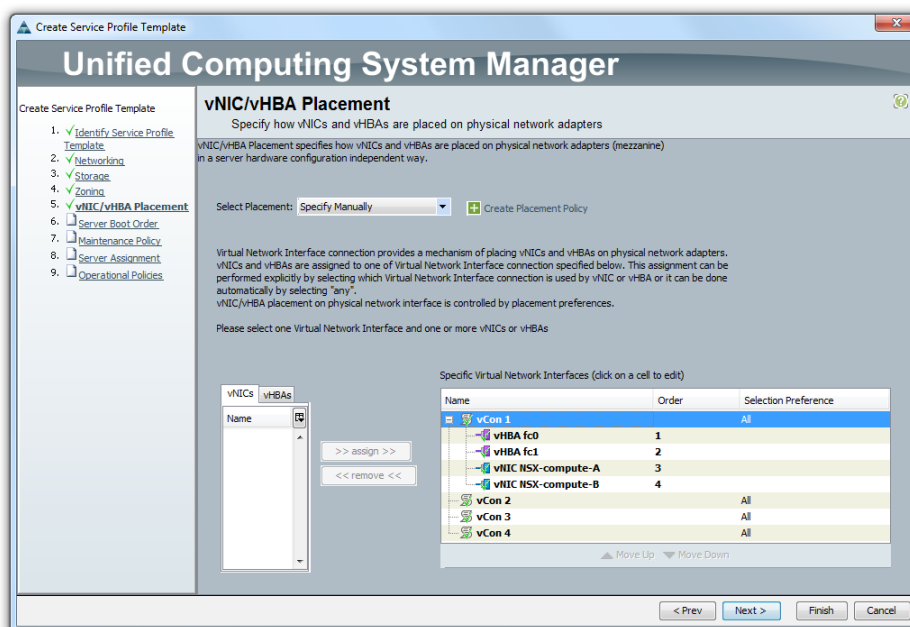


Cisco UCS service profile Networking configuration (Click to enlarge)

In the "Networking" section of your Service Profile Template, select "Use Connectivity Policy". Then select the LAN Connectivity Policy you just created in the previous step. This will link the network adapter configuration in this profile to your two vNIC Templates for any UCS servers associated with this profile.

vNIC Placement (PCI bus ordering)

In the "vNIC/vHBA Placement Policy" section of your Service Profile, select "Specify Manually".



Cisco UCS vNIC placement for vSphere hosts with NSX

In the "Specify Virtual Network Interfaces" area you will see the names of your vNIC Templates under the "vNICs" Tab. Follow these simple rules for placing the vNICs in the Order column:

- Use only vCon1, leave the other vCons empty
- If using FCoE, place your vHBA fc0 and fc1 adapters first
- Place the VNIC Template for Fabric A first (vNIC 1)
- Place the VNIC Template for Fabric B second (vNIC 2)

Complete the rest of the Service Profile Template to your individual preferences and associate it to UCS servers to be used as vSphere hosts with VMware NSX.

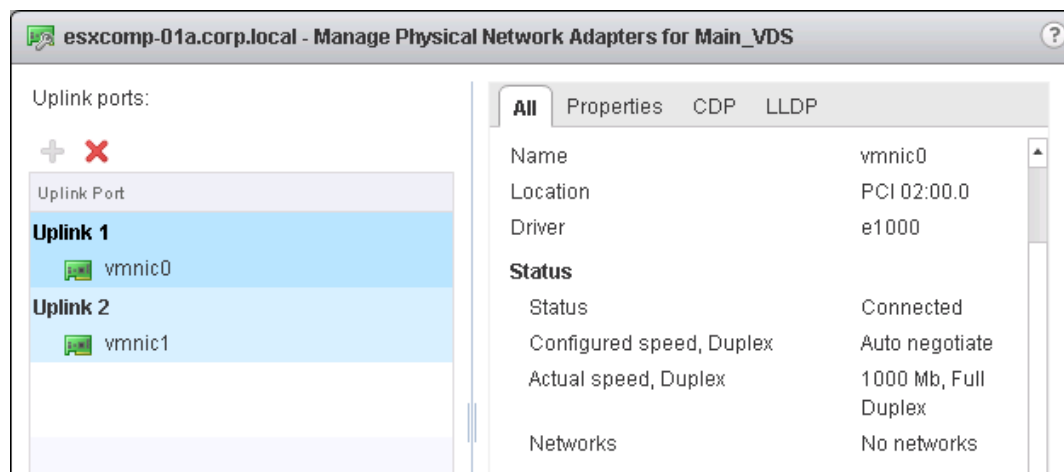
NSX for vSphere configuration for Cisco UCS

In this section we will cover the important configurations in vSphere relevant to running VMware NSX for vSphere on Cisco UCS.

At this point we assume you have successfully created a standard functional vSphere 5.5 environment managed by vCenter 5.5 with two clusters, Management & Edge and Compute, and one VMware VDS spanning both clusters. The NSX Manager and NSX Controllers have been successfully installed and registered with vCenter. You have NOT installed NSX components on any of the hosts, as some of the important installation settings will be covered below.

Assign host VDS uplinks to adapters

It's always important to make sure that each host has consistently mapped the same vmnic adapters to the same VDS uplinks. Go to the Network settings of your vSphere hosts and make sure that vmnic0 and vmnic1 adapters are properly assigned to the two VDS uplinks:



Connecting vSphere host adapters to VDS uplinks

- Uplink 1 assigned to vmnic0
- Uplink 2 assigned to vmnic1

This assignment and audit can also be performed with automation using vSphere Host Profiles.

VXLAN traffic Resource Pool

The host-to-host virtual machine traffic encapsulated into VXLAN by the vmknic is not automatically assigned to the system level "Virtual Machine Traffic" resource pool. So we will create a custom pool for VXLAN traffic and assign it adapter shares. First, make sure NIOC is enabled on your VDS. Then create a custom Network Resource Pool that will be assigned to your VXLAN vmknic Port Group later on.

Main_VDS - New Network Resource Pool

Name:

Origin:

Description:

Limit (Mbps): ☒ Unlimited

Physical adapter shares:

QoS tag:

OK **Cancel**

Creating a custom network Resource Pool for VXLAN traffic

- Set adapter shares to "High" (100)
- Set the QoS tag to "none"

All host-to-host virtual machine traffic will be encapsulated in VXLAN and given 100 shares of adapter bandwidth.

By default, the VXLAN outer header will copy the same QoS tag from the original frame. Setting the QoS tag to "none" will preserve that behavior when we apply this Network Resource Pool to the Port Group that will contain the VTEP vmknics for each host.

Note: Traffic on the Edge VLAN such as traffic sent and received by the NSX Edge Services virtual machines is operating on a normal vSphere Port Group, not encapsulated by VXLAN, and therefore will be subject to the settings of the system "Virtual Machine Traffic" resource pool.

Virtual Machine Traffic Resource Pool

We don't want a virtual machine tagging its own traffic and placing itself in a different QoS class without permission. This configuration change will prevent that by resetting the COS tag to zero (Best Effort) for all virtual machine traffic.

Edit the default system network resource pool "Virtual Machine Traffic" and change the QoS tag setting from

The screenshot shows a dialog box titled "Main_VDS - Edit Network Resource Pool Virtual Machine Traffic". It contains the following fields and values:

Field	Value
Name:	Virtual Machine Traffic
Origin:	System network resource pools
Description:	Virtual Machine Traffic Type
Limit (Mbps):	10000
Physical adapter shares:	High
QoS tag:	0

At the bottom right, there are "OK" and "Cancel" buttons.

"none" to "0".

Editing the Virtual Machine traffic pool for QoS marking

- Set QoS tag to "0"

Note: If we want some virtual machines to be in a different QoS class other than Best Effort, we can configure the VDS to classify specific VMs and set their QoS tag accordingly.

Configure VXLAN networking

Now let's configure the settings for the VXLAN vmknic on each vSphere host that will run VMware NSX. Go to the NSX Installation > Host Preparation area of the vSphere Web Client. In the VXLAN column select

Configure VXLAN networking

Configuring all hosts in cluster "Compute Cluster A" for VXLAN networking.

Switch: * Main_VDS

VLAN: * 200

MTU: * 1600

VMKnic IP Addressing: * ☐ Use DHCP
☒ Use IP Pool

IP Pool: VTEP-IP-Pool-Tran...

VMKnic Teaming Policy: * Load Balance - SRCID

VTEP: * 2

OK Cancel

"Configure" on a vSphere cluster.

Configure VXLAN networking

Choose "Load Balance - SRCID" for the VXLAN vmknic teaming policy. This will produce the VXLAN multipath configuration with (2) VTEPs per host. SRCID means that VM-to-VTEP pinning will be based on the virtual machine source virtual port ID. This is the recommended setting as all traffic from a given virtual machine will be pinned to one VTEP on the host. The other Load Balance option is SRCMAC which effectively works the same way as SRCID for virtual machines with only one MAC address. If the virtual machines have multiple MAC addresses, SRCMAC will conflict with Forged Transmit Reject settings, whereas SRCID will not.

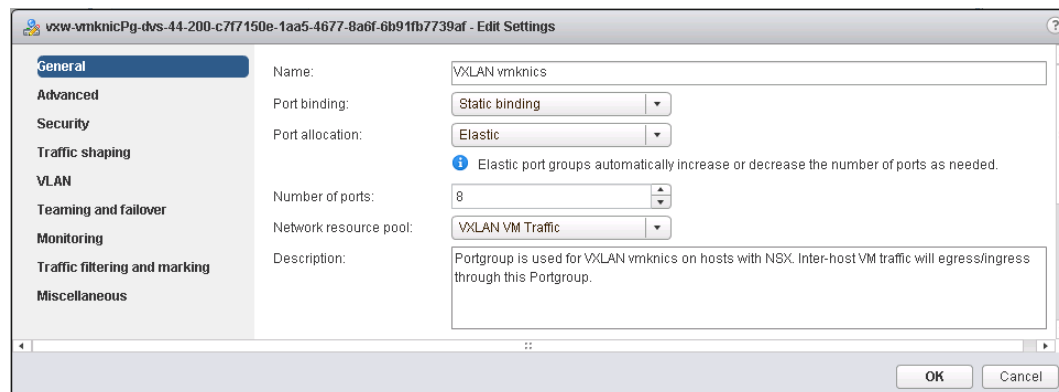
Note: for the single VTEP scenario without VXLAN multipath you would choose "Fail Over" as the Teaming Policy.

Make sure to set the VLAN to match the Transport VLAN. In our case, VLAN 200. Also make sure your IP Pool matches the IP subnet of the Transport VLAN. These will be the IP address settings assigned to the VXLAN vmknic. You can create the IP pool here at this step, or select DHCP. Verify that vmknic MTU is already set to 1600 bytes.

Configure the new VXLAN vmknic Port Group

After you've completed the VXLAN vmknic installation above, a new VDS Port Group has been automatically created with a name that begins with "vxw-vmknicPg-dvs..."

Assign the Network Resource Pool

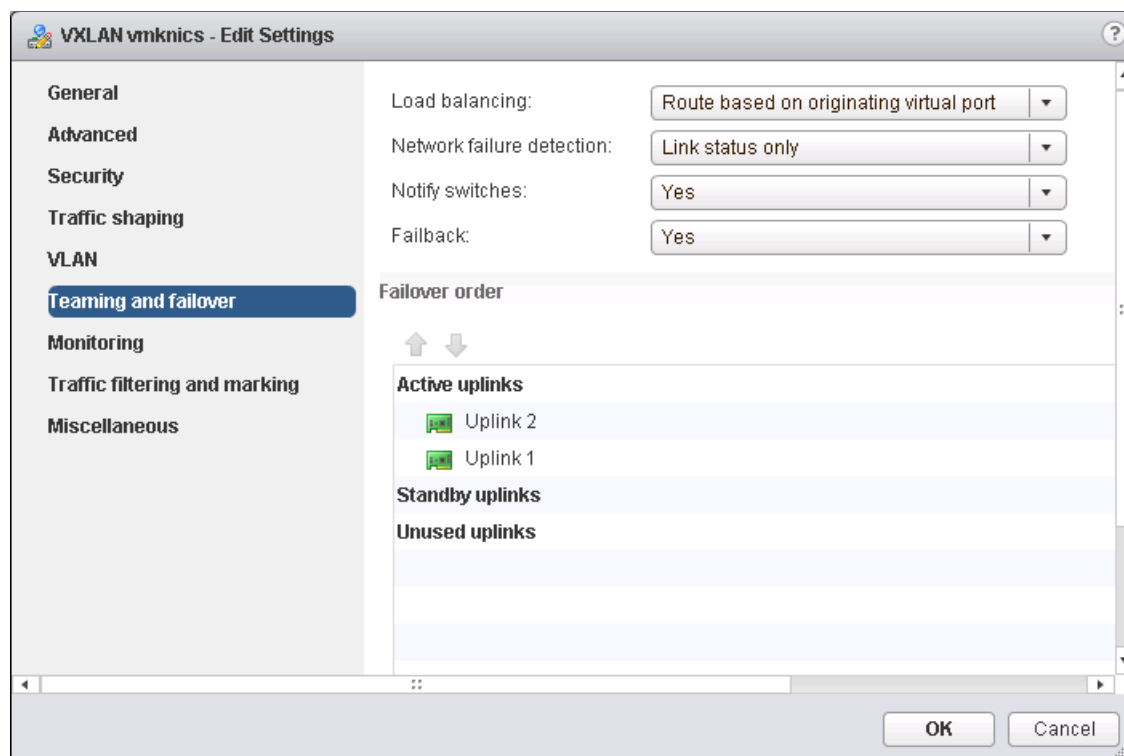


VXLAN Port Group settings (Click to enlarge)

This Port Group will contain the VXLAN vmknics for the hosts running VMware NSX. Edit this Port Group and do the following:

- Assign the Network Resource Pool to the previously created VXLAN traffic custom resource pool
- Edit the name (optional)
- Inspect the "Teaming and failover" section
- Verify that Uplink 1 is set to Active (on Fabric A)
- Verify that Uplink 2 is set to Active (on Fabric B)

Verify teaming and failover



VXLAN Teaming and failover settings for VXLAN Multipath (Click to enlarge)

The teaming and failover section should show both VDS uplinks as Active. VXLAN virtual machine traffic between hosts will be transmitted from the two VTEP vmknics in this Port Group, one VTEP pinned to Uplink 1, the other VTEP pinned to Uplink 2, utilizes both Cisco UCS adapter vNICs and both fabric interconnects.

We can also see the "Route based on originating virtual port" selection that was enabled when we chose "Load Balance - SRCID" during the initial VXLAN configuration step from earlier.

Note: for the single VTEP scenario without VXLAN multipath you should see that one VDS uplink is Active, and the other Standby.

Configure QoS tagging for system traffic

Now that VXLAN is configured and operational, edit the Resource Pool settings of your VDS and make sure that system traffic including Management, vMotion, and IP Storage are being tagged to match with the intended QoS Class in Cisco UCS.

QoS marking for system Network Resource pools

Main_VDS - Edit Network Resource Pool vMotion Traffic

Name: vMotion Traffic

Origin: System network resource pools

Description: vMotion Traffic Type

Limit (Mbps): 10000

☒ Unlimited

Physical adapter shares: Normal 50

QoS tag: 1

OK Cancel

- vMotion = Bronze class, COS 1
- Management = Silver class, COS 2
- IP Storage = Gold class, COS 4

Now might be a good time to double check your vSphere system traffic Port Groups, and verify that the "Teaming and failover" settings match your desired design. In our case the settings should look like this:

- vMotion = Multi-NIC (Fabric A & B)
- Management = Active on Uplink 2 (Fabric B)
- Management = Standby on Uplink 1 (Fabric A)
- IP Storage = Multi-Path (Fabric A & B)

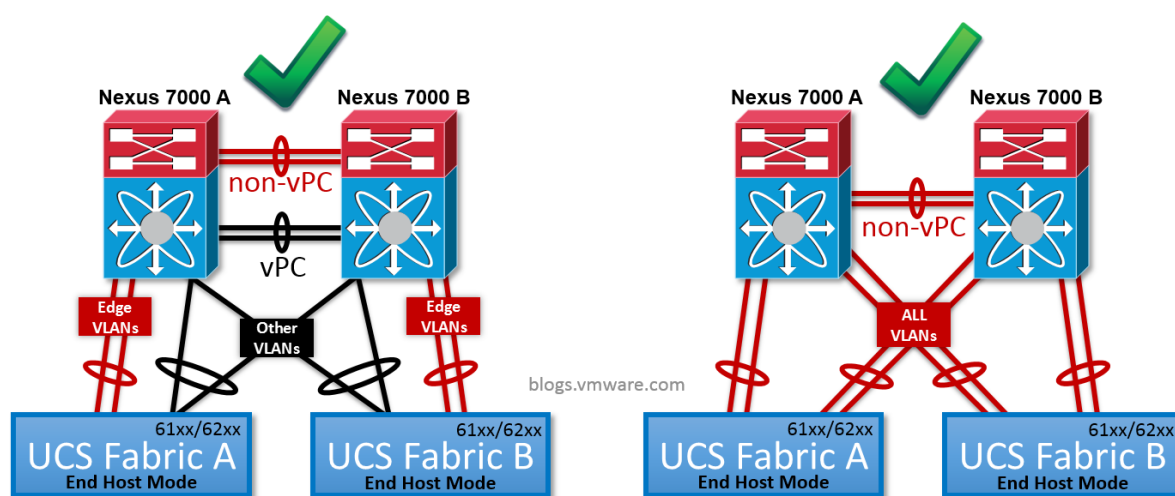
The settings above are offered here as an example for this specific design guide. You might choose something different based on your own preferences.

NSX Edge routing considerations with Cisco Nexus 7000

VMware NSX includes virtual routers, both in the form of a full service virtual machine (Edge Services Router), and in the form of distributed kernel level routing (Distributed Logical Router). Both are capable of running dynamic IP routing protocols like OSPF and BGP, and establishing neighbor relationships with other NSX routers, or any other router, such as the Cisco Nexus 7000. The virtual networks created automatically with NSX, including IP subnets, can be advertised dynamically to the physical data center routers without the need to manually provision (or deprovision) static routes.

When deploying the NSX routers in a way that will establish routing protocol relationships with the Cisco Nexus 7000, it's important that these routing protocol sessions are NOT established on VLANs that have been designated as vPC VLANs on the Cisco Nexus 7000. This is a [known hardware caveat with the Nexus 7000 switches](#).

In a deployment with Cisco UCS, this means we need pay special attention to how the Cisco UCS uplinks are configured, and the uplinks where the Edge VLANs are enabled.



Cisco UCS uplinks and the Edge VLAN with Nexus 7000 (Click to enlarge)

Above are two examples of configurations that will work just fine.

On the left, the Cisco UCS fabric interconnects have both a vPC uplink, and a non-vPC uplink. The Edge VLANs are enabled on the non-vPC uplink, while all other VLANs are enabled on the vPC uplink. The Cisco Nexus 7000 switches have non-VPC port channel between them that carries the Edge VLANs, in addition to the vPC port channel carrying the other VLANs.

On the right, all uplinks from Cisco UCS are non-vPC uplinks, connected to a set of Nexus 7000 switches that do not have a vPC enabled port channel between them. With the Cisco UCS fabric interconnects in their default "End Host" mode, spanning-tree will be disabled will all of the bandwidth still available for any VLAN on any of the uplinks.

