



# NSX Network Design

## ECMP Logical Routing

N-S Communication Option 2  
Uplinks VLANs "rack-confined"  
ToR switch serves as gateway

N-S Communication Option 1  
Number of ESG uplinks match ESXi uplinks.

Faster convergence time

if one ESG dies there is nearly zero downtime

up to 8 ESGs

VLANs spanning limited to the rack

L2 extension from the edge cluster not required

1 VLAN ID = Routing adjacency

all uplinks are active

if uplink fails the adjacency is withdraw

Separate TCP/IP stacks

VXLAN with dedicated stack

VMkernel VLANs do not extend

beyond the rack in an L3 fabric design

beyond the cluster with an L2 fabric

## VMkernel networking

Static routes are supported but should be minimized or avoided

Host profiles can simplify static routes management

Multi VTEP supports

Route Based Originating Port

Route Based Source Hash

## Physical Fabric

L2 Fabric (Access/Aggregation/Core)  
Cluster stretches across racks over same L2 domain

any network with 1600 MTU

Independent of network topology

Switch/Router Vendor Independent

Inter-POD is L3

default gateway at Aggregation

L3 Fabric (Spine/Leaf)  
Cluster stretches across racks but each rack belongs to a different L2 domain

same VLAN ID used but different IP addressing

dynamic routing protocols between leaf and spine

L3 at ToR (leaf) per VLAN (SVI)

## NSX Controllers

3 is the only supported number

4 vCPU, 4GB RAM, 2048MHz reserved

DRS anti-affinity rules are not automatically created

deployed on either Edge or Mgmt cluster

requires connectivity to ESXi vmk management interface and NSX Manager

TCP 1234 from hosts to Controllers

NSX Manager generates self-signed certs for hosts and controllers

## HA Logical Routing

4 ESG sizes (easy to change)

Stateful services

ESG Placement

Edge cluster for L3 routing, VPN

Compute cluster for LB,DHCP

Slower recovery time

Failover time

Heartbeat

vSphere HA

Routing processes restarts

## vSphere Cluster Design

Collapsed Compute/Edge/Mgmt

Small environment or POC

Dedicated

Compute

Connects to Edge via Transport VLAN

Edge

terminates 802.1Q VLANs

Mgmt

L2 for vCenter, NSX Controllers, NSX Manager and IP Storage

Production environment

## Scale Boundaries (vCenter)

10k Logical Switches

Cluster: max 32 hosts

512 hosts per vCenter

256 hosts per Transport Zone

1024 ESXi hosts overall

1:1 mapping NSX Manager

8 vCenter (Cross-VC)

## VXLAN Performance

Unicast Mode: appropriate for small deployments or L3 fabric where the number of hosts is limited

Hybrid Mode: generally recommended for Production and L2 network topologies

Consider Segment IDs overlapping across NSX domains (don't overlap!)

50 bytes VXLAN overhead  
original L2 header becomes payload + VXLAN,UDP and IP header

VXLAN Offloading (TSO)

Receive Side Scaling (RSS)

## NSX Manager

resiliency by vSphere HA

4 vCPU, 12GB RAM

config backup via S/FTP

reserve memory to ensure responsiveness from Web Client

Roles

Primary

Secondary

Standalone

Transit