

TOMORROW starts here.



Cisco *live!*

Deploying and Troubleshooting the Nexus 1000v Virtual Switch (vSphere and HyperV)

BRKVIR-3013

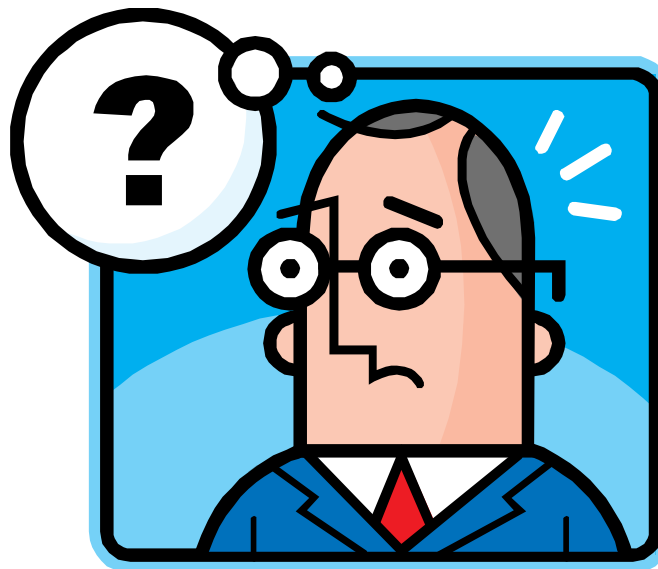
Arup Deb

Customer Support Engineer

Cisco Nexus 1000v VS Nissan GTR

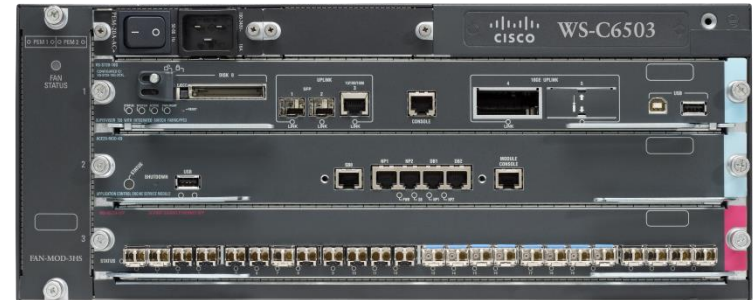


Confusion!?!?



What is the Nexus 1000v?

- Think back to traditional switches
- Relate components to:
 - Physical supervisor
 - Physical linecards
 - Physical servers



Cisco Virtual Networking Vision

Long Term Strategy

Nexus 1000V

Multi-Cloud

Multi-Service

Multi-Hypervisor

Agenda

- Current Nexus 1000V Releases and New Features
- Licensing
- Virtual Supervisor Module (VSM)
- VSM High Availability
- Virtual Ethernet Module (VEM)
- Upgrades
- Port-Profiles
- Port Channels
- VXLAN
- Cisco Nexus 1x10
- Appendix



Current Releases and Futures

Current Nexus 1000V Releases

- **5.1 Hyper-V only [5.2(1)SM1(5.1)]**
 - Support for SCVMM 2012 SP1 and Windows Server 2012
- **2.2 ESX [4.2(1)SV2(2.1)]**
 - Increase in number of VEMs, veth ports
 - VXLAN improvements
- **2.1 [4.2(1)SV2(1.1)]**
 - New Features including VCPlugin, Vtracker
 - Cisco Trustsec Support (CTS)
 - License changes for Essential(free) and Advanced
- **1.5.2 [4.2(1)SV1(5.2)]**
 - 1.5, 1.5.1a, and 1.5.2
 - 1.5.2 first version to support ESXi 5.1

2.2 Features

- Increased Scale
 - 128 VEM support per VSM
 - 300 ports per host
 - 4000+ ports per VSM
- VXLAN Evolution
 - Unicast mode
 - Unicast Mac-address distribution option
- VXLAN Gateway
 - Bridge VXLANs to VLANs

2.1 Features - Recap

- License Changes
- Cisco TrustSec Support
- vTracker
- VCPlugin
- VSM HA improvements
- VSMS split between Data Centres
- VEM remote branch support
- Enhanced Upgrades

vTracker Feature

- New show command
- Pulls data from vCenter and VEM
 - Gives “cloud” view
- Enable with “feature vtracker”
- 5 outputs:
 - **module-view**
 - **upstream-view**
 - **vlan-view**
 - **vm-view**
 - **vmotion-view**

```
nlkv-test# show vtracker vm-view info

Module 3:
  VM Name:                rsmall-linux-11-v
  Guest Os:                Red Hat Enterprise Linux 5
  (64-bit)
  Power State:            Powered On
  VM Uuid:                 42027c4a-3222-44a0-4866-
0e320aa40560
  Virtual CPU Allocated:  1
  CPU Usage:               16 %
  Memory Allocated:       384 MB
  Memory Usage:           2 %
  VM FT State:             Unknown
  Tools Running status:   Running
  Tools Version status:   supported old
  Data Store:              nas-virt
  VM Uptime:               4 hours 16 minutes 38 seconds
```

VCPlugin

- vCenter plugin that displays Nexus 1000V data in vCenter WebClient
- Works with Webclient 5.1
- Read-only
- Works with vCenter
 - Windows version
 - Linux Appliance Version

The screenshot shows the vCenter WebClient interface for a Cisco Nexus 1000V switch. The page is titled "Cx-VSM-VCP-Test" and includes a navigation menu with "Getting Started", "Summary", "Monitor", "Manage", and "Related Objects". The "Monitor" tab is active, and the "Cisco Nexus 1000V" sub-tab is selected. The main content area is divided into several sections:

- System Information:** A table with columns for Switch Name, NX-OS Version, VSM IP, DC Name, Connectivity Mode, VC Connectivity, and VSM HA. The data row shows: Cx-VSM-VCP-Test, 4.2(1)SV2(1.1) [build 4.2(1)SV2(1.0.134)] [gdb], 10.78.0.126, DC-1, L2, Connected, and true.
- Network Statistics:** A table with columns for VNICs vs Max, Hosts vs Max, Port-Groups vs Max, Veths/Host Max, and *Vlan/VxLan vs Max. The data row shows: 76(2048), 4(64), 1942(2048), 76(216), and 1944(2048) Vlan, N/A(N/A) VxLan.
- Licenses:** A table with columns for License Type, Licenses Available, Licenses Used, Earliest Expiration, and Status. The data rows are:
 - NEXUS_VSO_SERVICES_PKG: 16 Available, 0 Used, 16 Oct 2012 Expiration, Unused Status.
 - NEXUS_ASA1000V_SERVICES_PKG: 48 Available, 0 Used, Never Expiration, Unused Status.
 - NEXUS1000V_LAN_SERVICES_PKG: 2075 Available, 5 Used, 03 Sep 2012 Expiration, In use Status.

A "Logout" button is visible in the bottom right corner of the interface.



Licensing Info

Licensing – Essential

- All features but...
 - Cisco TrustSec
 - DHCP snooping
 - IP Source Guard
 - Dynamic ARP Inspection
 - VXLAN Gateway
- 512 socket perpetual license
- Support Options
 - Free – support is through the communities forum
 - <https://communities.cisco.com/community/technology/datacenter/nexus1000v>
 - Paid service contract
 - Cost is \$39 per socket per year for TAC support.

Licensing – Advanced

- For customers that want more security features
- Customers with old licenses will be considered Advanced
- Upgrade process with put VSM in Advanced Mode
- Required for VXLAN Gateway and VSG
- Licensed customers can get Virtual Security Gateway(VSG) for free
 - Account team can submit request for software
 - VSG no longer be sold separately

Licensing – New Commands

- To show switch edition

- `switch# show switch edition`

- To switch between Essential or Advanced

- `switch(config)# svs switch edition essential/advanced`

- Remember licenses per VEM are sticky

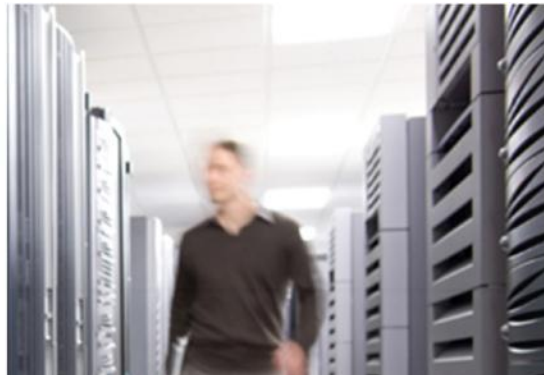
- `switch# show module vem license-info`

- `Licenses are Sticky`

Mod	Socket Count	License Usage	License Version	License Status
3	2	-	-	licensed

- VEM license transfer with

- `switch(config)# svs license transfer src-vem <module> license_pool`

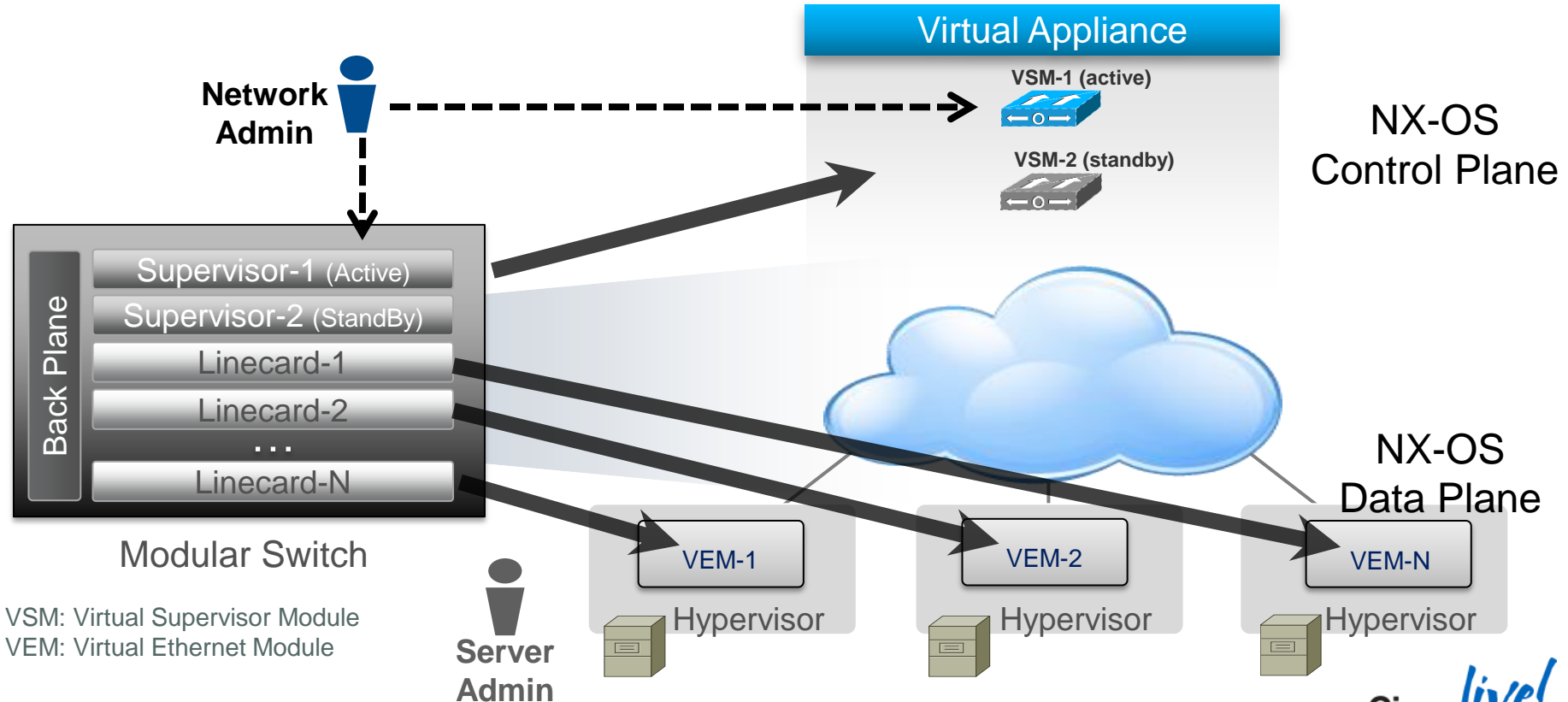


Virtual Supervisor Module Deployment and Troubleshooting

Agenda - VSM Lifecycle

- What is the Virtual Supervisor Module (VSM)
- Planning
- Installation
- Troubleshooting

Cisco Nexus 1000V Architecture



Virtual Supervisor Module (VSM)

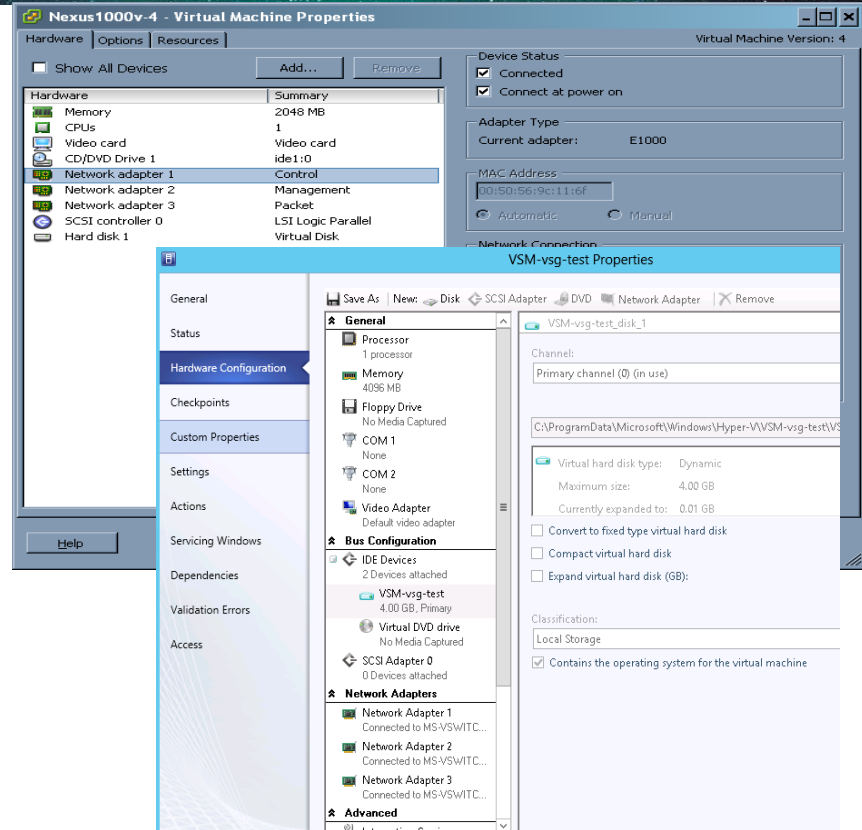
- VSM is a Virtual Machine
- Control plane for the Nexus 1000V solution
- Responsibilities:
 - Programming and managing Virtual Ethernet Modules (VEM)
 - Communicating with Management Applications
- 1 VSM HA pair can manage 128 VEMs (64 on Hyper-V)
- Coexist with VMware vSwitch, vDS, Microsoft Logical, Native Switch

VSM On Hyper-V vs. ESX

- VSM is all Hyper-V or all ESXi VEM modules
 - No mixing of VEM modules currently allowed
- VSM can run anywhere
 - VSM attached to Hyper-V VEMs can run on ESXi
 - Nothing specific to the VSM VM to require it to run on a specific hypervisor
- Configs are slightly different

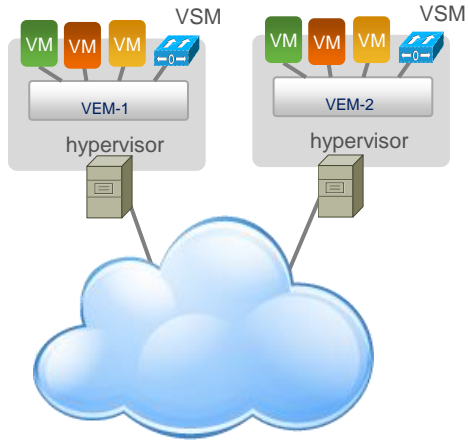
VSM Virtual Machine Requirements

- Adapter 1 is the **Control** interface
 - Heartbeat between VSMs and VEM(L2)
 - Heartbeat between VSMs
 - Control 0
- Adapter 2 is the **Management** interface
 - VSM terminal connectivity
 - Connectivity to VMware vCenter
 - Backup Heartbeat for VSM HA
 - Mgmt 0
- Adapter 3 is the **Packet** interface
 - Passes CDP and IGMP information
- 2 or 3GB of memory ESXi
- 4GB of memory Hyper-V

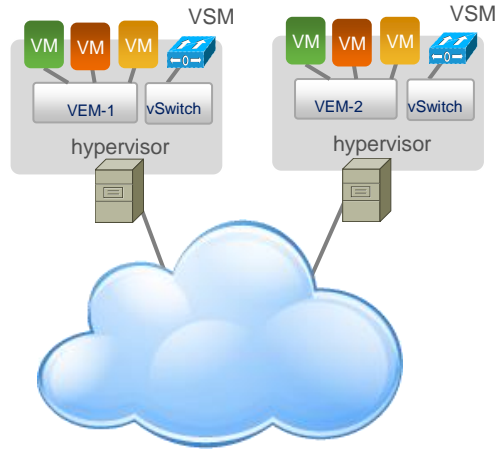


Deployment Scenarios

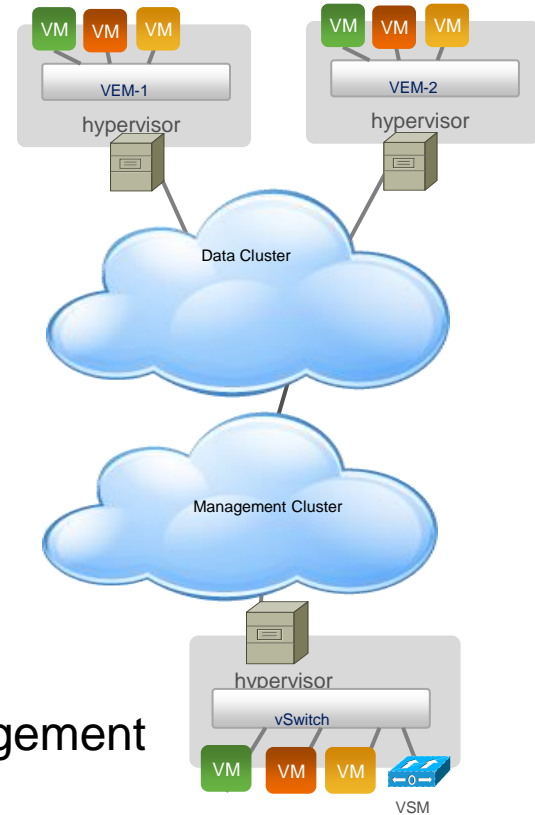
VSMs on VEM



VSMs on vSwitch

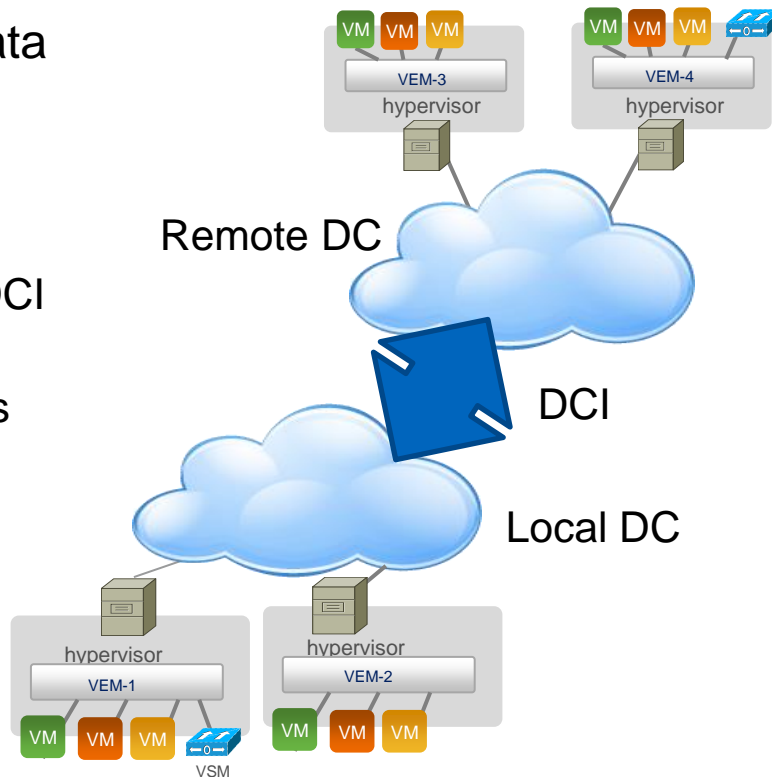


VSMs in Management Cluster



Stretched Nexus 1000V Model

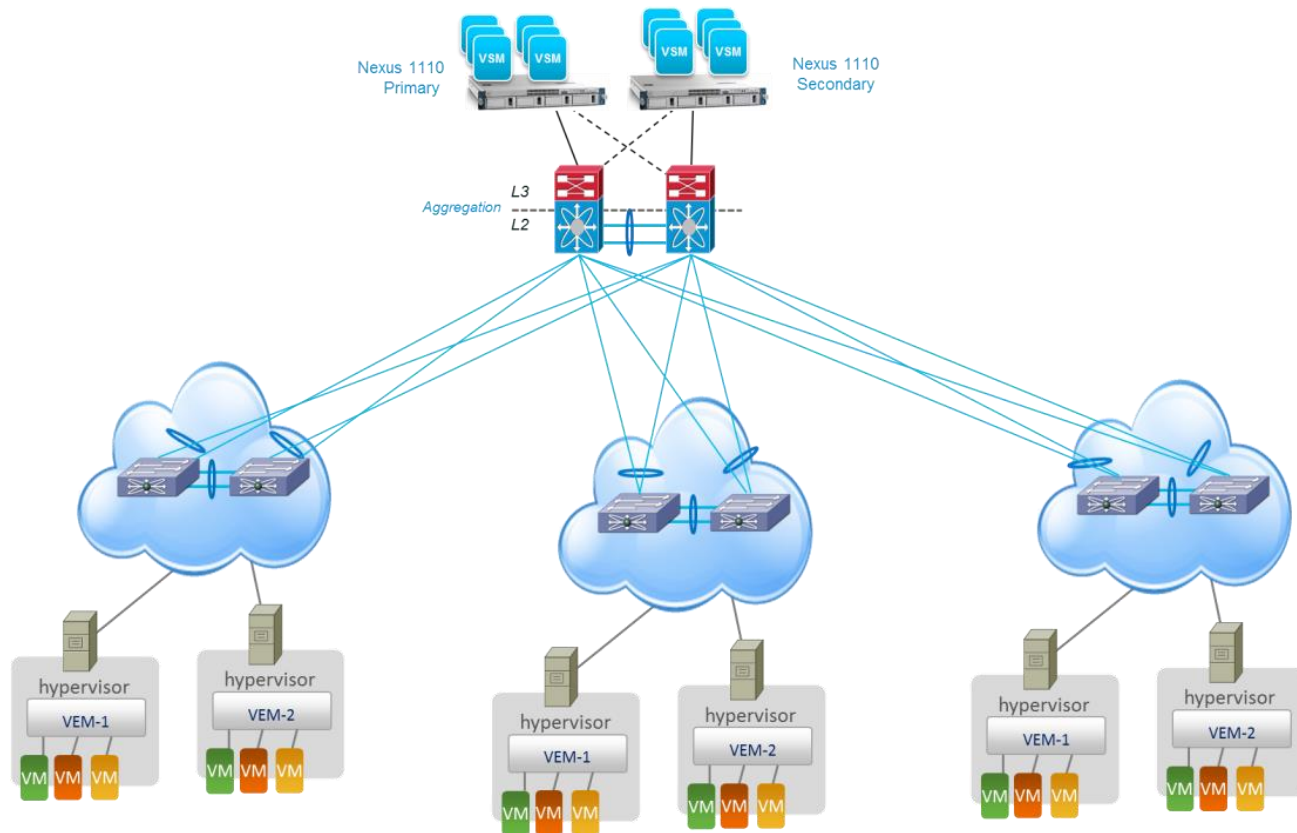
- VSMs and VEMs split across Data Centres
- VSMs can be split across DCs
 - New with 2.1
 - Requires L2 connectivity across DCI
 - 10ms latency across DCI
 - No reduction in configuration limits
- Will be supported in future for Hyper-V



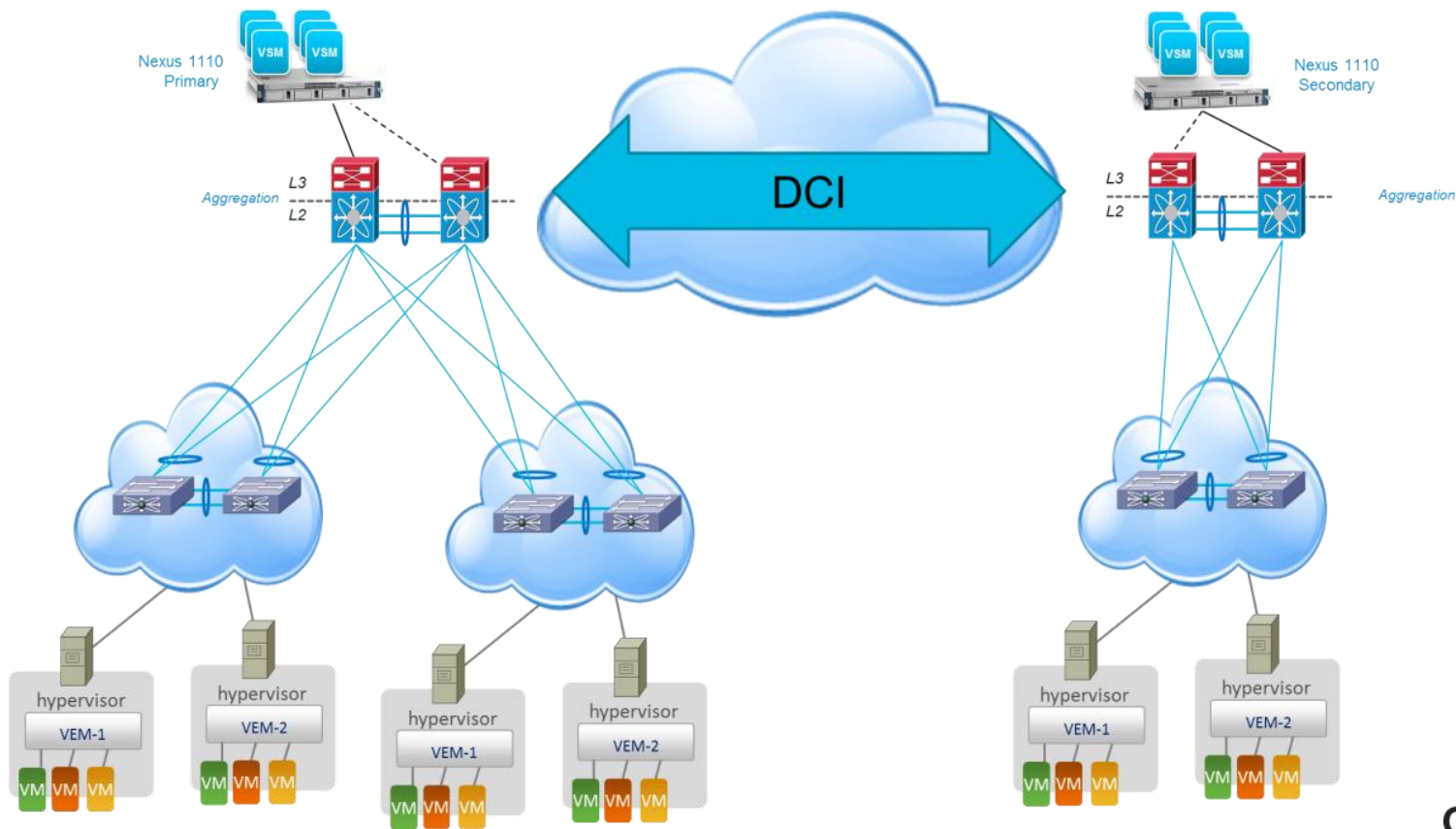
VSM Deployment Scenarios – Nexus 1110

- VSM on a Nexus 1010/X or 1110-S/X
 - It's still a Virtual Machine
 - Up to 10 VSMs on one 1110-X
- Only deployed in pairs
- 1x10 allows for Network team to own the virtualisation platform
- 1x10s should go in the Aggregation Layer
- Stretched Model requires
 - L2 Connectivity
 - 10ms latency
 - No Hyper-V support

1110-S/X Deployment Scenario



1110-S/X Deployment Scenario



VSM Control Modes

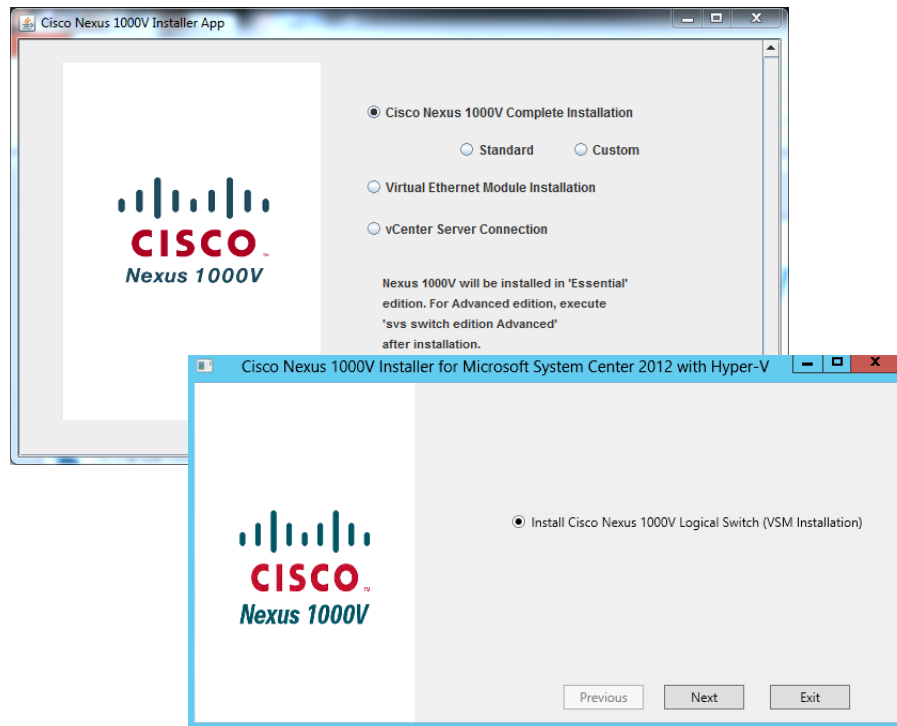
- L3 Mode
 - We recommend L3 and it's the default mode now
 - Easier to troubleshoot
 - Flexible
- L2 mode
 - Requires L2 connectivity through Control interface to all VEM modules
 - L2 still supported on ESX
 - Not supported with Hyper-V

VSM L3 Configuration And Planning

- Two options for the L3 control interface
 - Mgmt 0 (default)
 - Control 0
- Recommend using mgmt 0
- Wish to separate control and management traffic?
- Mgmt and Control use different VRF
 - Mgmt 0 uses VRF management
 - Control 0 uses VRF default
- Primary and secondary VSM still need to be L2 adjacent

VSM Installation

- For new installations use the Installer application
 - ESXi and Hyper-V
- ESXi
 - Installer can also migrate ESXi hosts to the N1KV
 - ISO and OVA still available
 - Use for established VMware vSphere deployments
- Hyper-V
 - Installer or Template/ISO install methods



VSM Connectivity To VMware vCenter

- VSM connects to vCenter using SSL connection
 - Plugin that contains SSL cert
 - Unique extension ID for the VSM
- VSM talks to vCenter using its API
 - We push and pull data to/from vCenter
- VSMS are tied to a VMware Data Centre
 - Multiple VSMS tied to same DC is allowed

Troubleshooting VSM/vCenter Connectivity

- VSM should always be connected to vCenter

```
n1000v# show svcs connections
```

```
connection VC-test:
```

```
ip address: 172.18.217.241
```

```
protocol: vmware-vim https
```

```
certificate: default
```

```
datacenter name: Harrison
```

```
DVS uuid: 72 f7 01 50 b2 01 7b 8b-55 68 cf df 10 5a db 55
```

```
config status: Enabled
```

```
operational status: Connected
```

VSM Connectivity Errors - ESXi

- If you see
“**Extension key was not registered before it's use**”
 - Re-register the Extension Key with VMware vCenter
- If you see
“**Connection refused. connect failed in tcp_connect()**”
 - VMware admin could have changed the http port
 - API communication is through port 80 with VMware vCenter
 - Find new port and change it on VSM

VSM Connectivity To SCVMM 2012 SP1

- Extension gets installed on SCVMM via a Provider Extension
 - Simple Windows installable file
 - VSM is a forwarding extension in the SCVMM logical switch infrastructure
- SCVMM talks to the VSM using our API
 - SCVMM pulls data once every 30 minutes
 - We can force a refresh in SCVMM
- VSMS get tied to a SCVMM Host Group
 - Multiple VSMS tied to same Host Group is allowed

Troubleshooting VSM And SCVMM Connectivity

- No persistent connection
- Any connectivity errors should show in the “Jobs” Screen in SCVMM
- Verify Connectivity by accessing API from the host
 - `http://<vsm-ip>/api/n1k`
 - Remember to add admin credentials to SCVMM run-as accounts
 - Currently API can use local accounts only
 - Verify any proxies
 - Firewalls
- If no extensions show up re-install the provider

VSM And vMotion/Live Migration

- Manual vMotion/Live Migration is supported
- Not recommended to allow VMware DRS/SCVMM Dynamic Operations to move primary and secondary VSM
- Aggressive settings can cause VSM to drop packets
- Best practice to keep Primary and Secondary VSM on separate hosts

Backing Up The VSM

- A running-config is not enough to restore
- VSM on ESXi and Hyper-V
 - Clone to a template
 - You can restore from a template and saved-config
- VSM on Nexus 1x10
 - Export or import VSM
- VSM on ESXi Snapshots
 - Not supported

Hyper-V Specific Gotchas

- VSM requires 4GB RAM
- No stretched deployment support
 - Even with Nexus 1010/1110
 - Coming in later release
- Recommend VSM be on a Microsoft vSwitch
- 64 VEM support, 2k veth ports
- Domain ids shrink from 4k to 1k
 - Another NXOS 5 issue
- L3 Control mode only

VSM Best Practices

- L3 control is the preferred method
- Use mgmt 0 for control
- Primary and Standby VSM in same L2 domain
 - Required even if VSMs are split between Data Centres
- VSM on VEM is supported
- VSM primary to secondary latency max 10ms
 - 10ms even for VSMs split between Data Centres
- VSM to VEM latency 5-10ms
 - For VEMs at branch locations 100ms
- Backup your config



Nexus 1000V High Availability


VSM Redundancy Manager

- HA evolved to support split Data Centre deployments
- New Redundancy Manager process polls
 - VEM Manager – polls for number of active VEMs attached to VSM
 - VMS process – retrieves which VSM has active VC connectivity
 - SNMP Library – gets the last configuration time
- Runs on primary and secondary VSM
- Heartbeat timeout values
 - VSM-VSM every second. Drop after 6 missed
 - VSM-VEM every second. Drop after 15 missed

HA Logic

- Management interface is now used for backup heartbeat
 - State changes to “degraded” when control 0 fails
 - No configuration changes are allowed
- Useful Commands
 - show system internal active-active [remote] accounting logs
 - show system internal redundancy trace

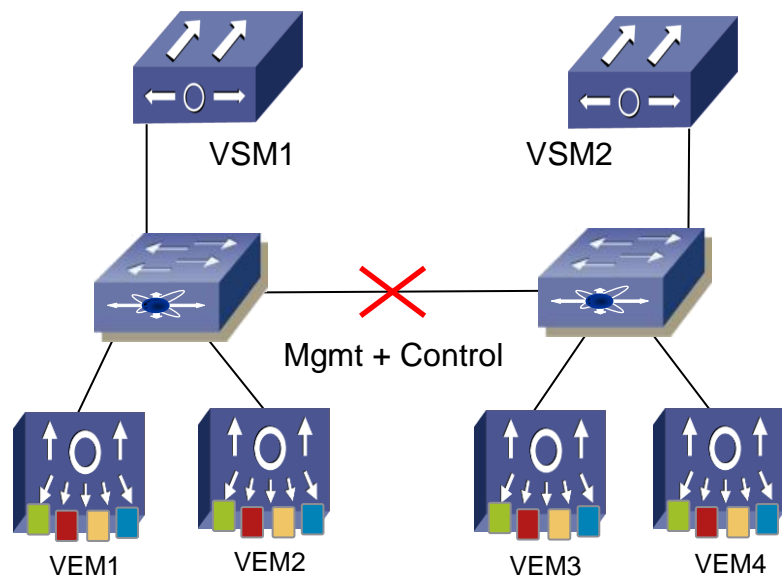
```
n1kv-13# show system internal redundancy trace
1 0s START_THREAD ST_INIT ST_NP ST_INVALID
2 0s CP_STATUS_CHG ST_INIT ST_NP ST_INIT
3 5s DEGRADED_MODE ST_INIT ST_NP ST_INIT
4 5s STATE_TRANS ST_INIT ST_NP ST_INIT EV_OS_NP ST_AC_NP
5 0s CP_STATUS_CHG ST_AC ST_NP ST_AC_NP
6 4s SET_VER_RCVD ST_AC ST_NP ST_AC_NP
7 0s STATE_TRANS ST_AC ST_INIT ST_AC_NP EV_OS_INIT ST_AC_INIT
8 0s STATE_TRANS ST_AC ST_SB ST_AC_INIT EV_OS_SB ST_AC_SB
```



When Does A VEM Register To Another VSM?

- What if we have two active VSMs
- What causes a VEM to switch?
 - Standby VSM becomes active and broadcasts to all VEMs
 - VEM will attach depending on
 1. Connectivity between VEM and VSM
 2. VEM receives the “request to switch”
- VEM goes into headless mode after 15 seconds
- If a VEM is headless vMotion/Live Migration is blocked

Failure Scenario – Split Brain



- VSM1 stays active, keeping VEM 1 & 2
- VSM2 becomes Active, taking over VEM 3 & 4
- Redundancy manager will note changes
- Important note - creating and changing configuration on VSM2 is not enough – Check VC Sync



Virtual Ethernet Module Deployment and Troubleshooting

VEM Deployment Overview

VMware

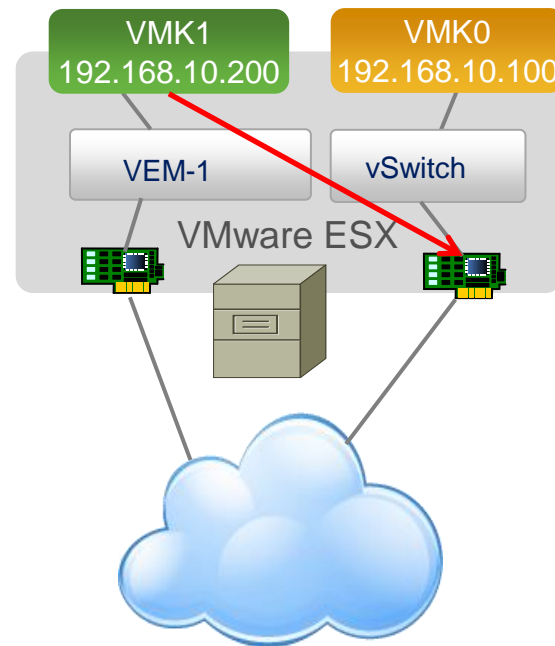
- L3 control requires a VMKernel NIC
- Recommend using the ESXi management VMKernel
 - Migrate the management interface to the VEM
 - No static routes on required
- Do not create an L3 vmk on same subnet as mgmt vmk

Microsoft

- Only L3 supported
- No special NIC required
- We talk to the Windows Server 2012 Management NIC
 - Attach virtual NIC to native or logical switch
 - Not recommended to connect Management NIC to VEM
- No special port-profile required

VEM Deployment – VMKs On Same Subnet

- Don't use multiple VMKs on the same subnet but virtual switches
- VMware uses a single TCP/IP stack for all VMK interfaces
- No way to direct traffic up a particular interface when they share the same subnet
- One interface gets picked for all traffic on that subnet
- Check out VMware KB article 2010877

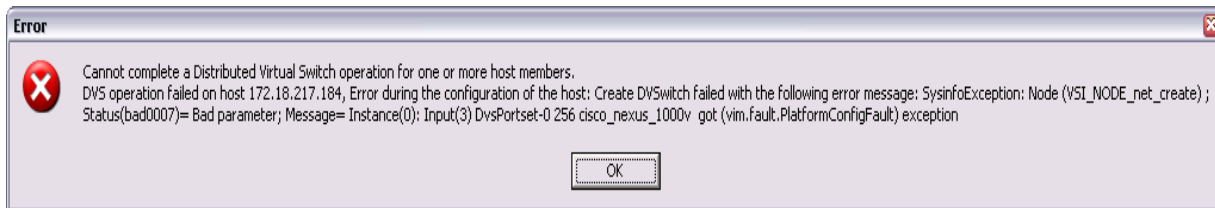


VEM Installation Options

- Vsphere Update Manager
- Manual
 - Esxcli
 - Esxupdate
- Cisco installer application
- Stateless/PXE
 - Use VMware PowerCLI
- Microsoft MSI installer

VEM Installation DVS Error

- If you are using VUM check the logs
 - VUM could not find the right VEM version
 - Check the CISCO and CSCO directories
- Make sure the VSM is connected to vCenter
- Make sure Cluster HA, DRS, and DPM are disabled



VEM Troubleshooting

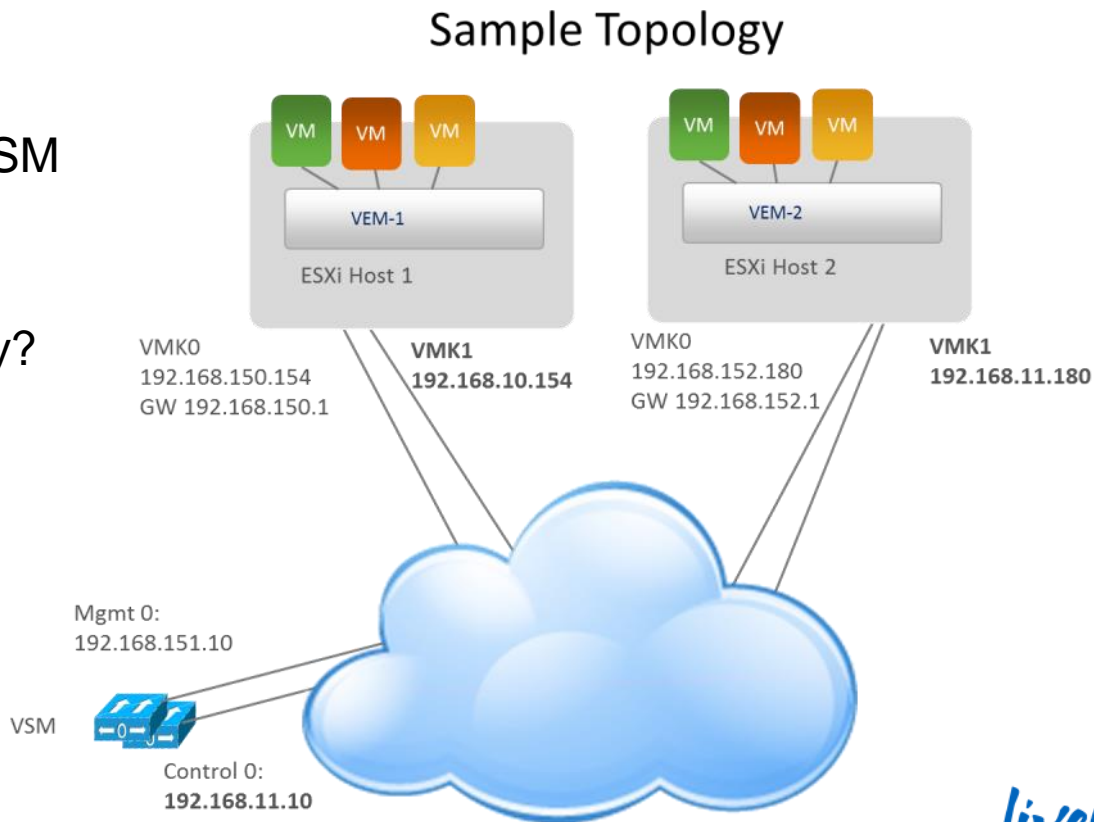
- Use vemlog on hyper-v or ESXi host
- Enable different debug options to help troubleshoot
 - LACP
 - QOS
 - VEM-VSM issues
- http://www.cisco.com/en/US/products/ps9902/products_tech_note09186a0080bed119.shtml

Troubleshooting VEM – VSM Connectivity - ESXi

- VEM adds in vCenter but does not show up on VSM “show module”
- With L3 check for IP routing problem
 - If you can ping VMK interface the VEM should connect to VSM
 - Troubleshoot as you would all VMware L3 issues
- With L2 most of the time its a Control VLAN issue
 - Verify Control VLAN connectivity
 - Check appendix for more L2 troubleshooting tips

VEM L3 Troubleshooting – Steps ESXi

- VMK interface working?
- Can the ESXi host ping the VSM control/mgmt interface?
 - Static route needed?
- Uplink profile created correctly?
- L3 veth port-profile created correctly?
- Check the opaque data



Verify VSM Settings

- Check SVS domain parameters

```
n1kv-13# show svcs domain
SVS domain config:
  Domain id:      43
  Control vlan:  1
  Packet vlan:   1
  L2/L3 Control mode: L3
  L3 control interface: control0
```

- Verify control 0

```
n1kv-13# show run int control 0
interface control0
  ip address 192.168.11.10/24
```

- Verify VRF default

```
n1kv-13# show ip route
IP Route Table for VRF "default"
0.0.0.0/0, ubest/mbest: 1/0, pending
    *via 192.168.11.1, control0, [1/0], 4d23h, static
```

- Can the VSM ping the VMK interface

```
n1kv-13# ping 192.168.11.180 vrf default
PING 192.168.11.180 (192.168.11.180): 56 data bytes
64 bytes from 192.168.11.180: icmp_seq=0 ttl=63 time=1.082 ms
64 bytes from 192.168.11.180: icmp_seq=1 ttl=63 time=0.841 ms
```

Check The Port-Profiles

Uplink port profile

- Check trunking and port channelling
- Check allowed and system vlans

```
n1kv-13# show run port-profile uplink
port-profile type ethernet uplink
  vmware port-group
  switchport mode trunk
  switchport trunk allowed vlan 10-11,150-152
  channel-group auto mode on mac-pinning
  no shutdown
  system vlan 10-11
  state enabled
```

Vethernet port profile

- Check VMK
- Check for l3control and system vlans

```
n1kv-13# show run port-profile L3-control-vlan11
port-profile type vethernet L3-control-vlan11
  capability l3control
  vmware port-group
  switchport mode access
  switchport access vlan 11
  no shutdown
  system vlan 11
  state enabled
```

Check Opaque Data

- Opaque data is bootstrap data for the VEM
 - Pushed via SCVMM or vCenter during install
- Is the right Opaque data getting pushed to the ESXi host?

```
# vemcmd show card
Card UUID type 2: 414a3031-3341-5553-4538-31384e375a37
Card name: cae-esx-186
Switch name: n1kv-13
Switch alias: DvsPortset-0
Switch uuid: 48 68 29 50 e2 ba af 6c-13 72 14 bc 25 cf 3f 86
Card domain: 43
Card slot: 4
VEM Tunnel Mode: L3 Mode
L3 Ctrl Index: 49
L3 Ctrl VLAN: 11
VEM Control (AIPC) MAC: 00:02:3d:10:2b:03
VEM Packet (Inband) MAC: 00:02:3d:20:2b:03
VEM Control Agent (DPA) MAC: 00:02:3d:40:2b:03
VEM SPAN MAC: 00:02:3d:30:2b:03
Primary VSM MAC : 00:50:56:af:49:30
Primary VSM PKT MAC : 00:50:56:af:59:a1
Primary VSM MGMT MAC : 00:50:56:af:47:51
Standby VSM CTRL MAC : 00:50:56:af:61:e7
```

Should match VLAN defined
in veth port-profile

Should match MAC of
control 0 or mgmt 0

Troubleshooting VEM – VSM Connectivity Hyper-V

- Hyper-V host won't add
 - Verify NIC MTU settings they should match
 - Uplinks should always be of type “TEAM”
 - Verify no old teams with “Get-NetSwitchTeam” and “Remove-NetSwitchTeam”
- Verify SCVMM can talk to VSM
 - Check for proxies/firewalls
- Verify Logical Switch Compliance
 - Verify compliance and remediate if necessary

View Heartbeat Messages On VEM

- Use vempkt on the ESXi host
 - vempkt capture ingress/egress
 - Let it run
 - vempkt cancel capture all
 - vempkt display detail all
- vempkt can now export to a pcap file
 - vempkt pcap export <filename>
- Look for heartbeat messages from VSM

VEM Best Practices

- Use L3 Control
- Use ESXi mgmt vmk interface for control
- Control network should have low latency and available bandwidth
 - 10ms for local DC
 - 100ms for branch office deployment
- On UCS make Service Profile does not contain “Dynamic VNICs”
 - VEM and VM-FEX are mutually exclusive



Upgrades

Prepare For Upgrade!

- Take a backup of the VSMs
 - On ESXi use the clone to template option
 - On Nexus 1x10s use the export function
 - Backup the running-config
- Follow the upgrade guides
 - Proceed in order
- Generate a tech-support before the upgrade
- Maintenance window only required in some scenarios

Upgrades To 2.2

- Scalability limits require changes to the VM settings
- For full scalability support need to change
 - CPU reservation to 2GHs
 - Memory to 3GB
- Steps
 - Shutdown Secondary
 - Make VM changes
 - Power secondary on
 - System switchover
 - Repeat steps on Primary VSM

Upgrading The VSM

- Changes from 2.1
 - You can now Upgrade VSMs and make changes without upgrading VEMs
- Upgrade is similar to other Nexus switches
 - Copy new kickstart and system images
 - Run single “install all” command to launch upgrade
- Requires no VSM outage or network outage for hosts

Troubleshooting VSM Upgrades

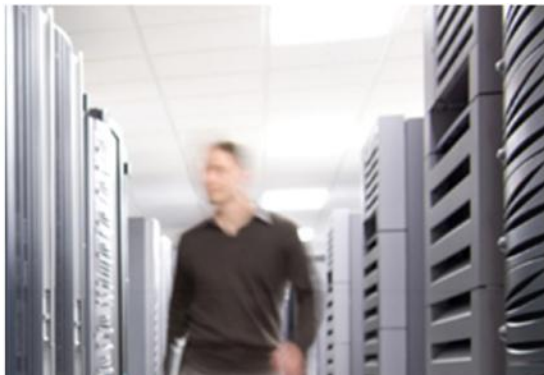
- If something is wrong after the VSM upgrade
 - Call TAC
 - Roll back using backup method
 - Shutdown the VSM VMs
 - Power-on the Clones (ESXi), Import the backup (Nexus 1x10)
- If VEM won't connect to the Standby VSM
 - “system switchover” once the old primary is upgraded
- Verify Standby VSM before upgrade
 - Make sure VEMs can connect to standby with system switchover

Upgrading The VEMs

- VEM module upgrade kicked off on VSM
 - If VUM is installed everything is automatic
 - Host is placed in maintenance mode(if DRS is installed VMs are migrated off)
 - VEM is upgraded and host exits maintenance mode
 - Moves on to the next host
 - If VUM is not installed
 - Still initiate the process on the VSM
 - User manually places ESXi hosts in maintenance mode
 - Upgrade the VEM with esxcli command
 - Exit maintenance mode and move to the next host
- Always complete the upgrade
 - Issue the “vmware vem upgrade complete” command

Troubleshooting VUM Upgrades

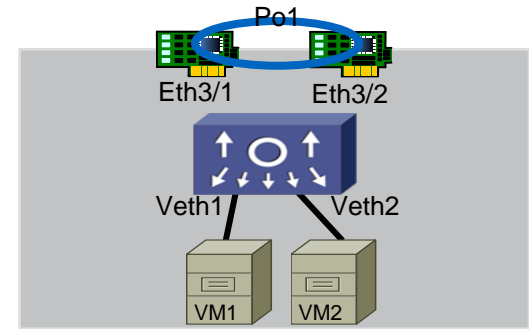
- Remember the VMware admin has to acknowledge upgrade in vCenter
- Make sure you have DRS capacity
 - Need to be able to handle one ESXi host failure
- If a particular ESXi host fails
 - Usually because the host cannot go into maintenance mode
 - From vCenter attempt to place host in maintenance mode
 - ESXi host running a vCenter VM can be the cause
- Its ok to restart the VEM upgrade after it fails
 - It will only upgrade hosts that did not succeed



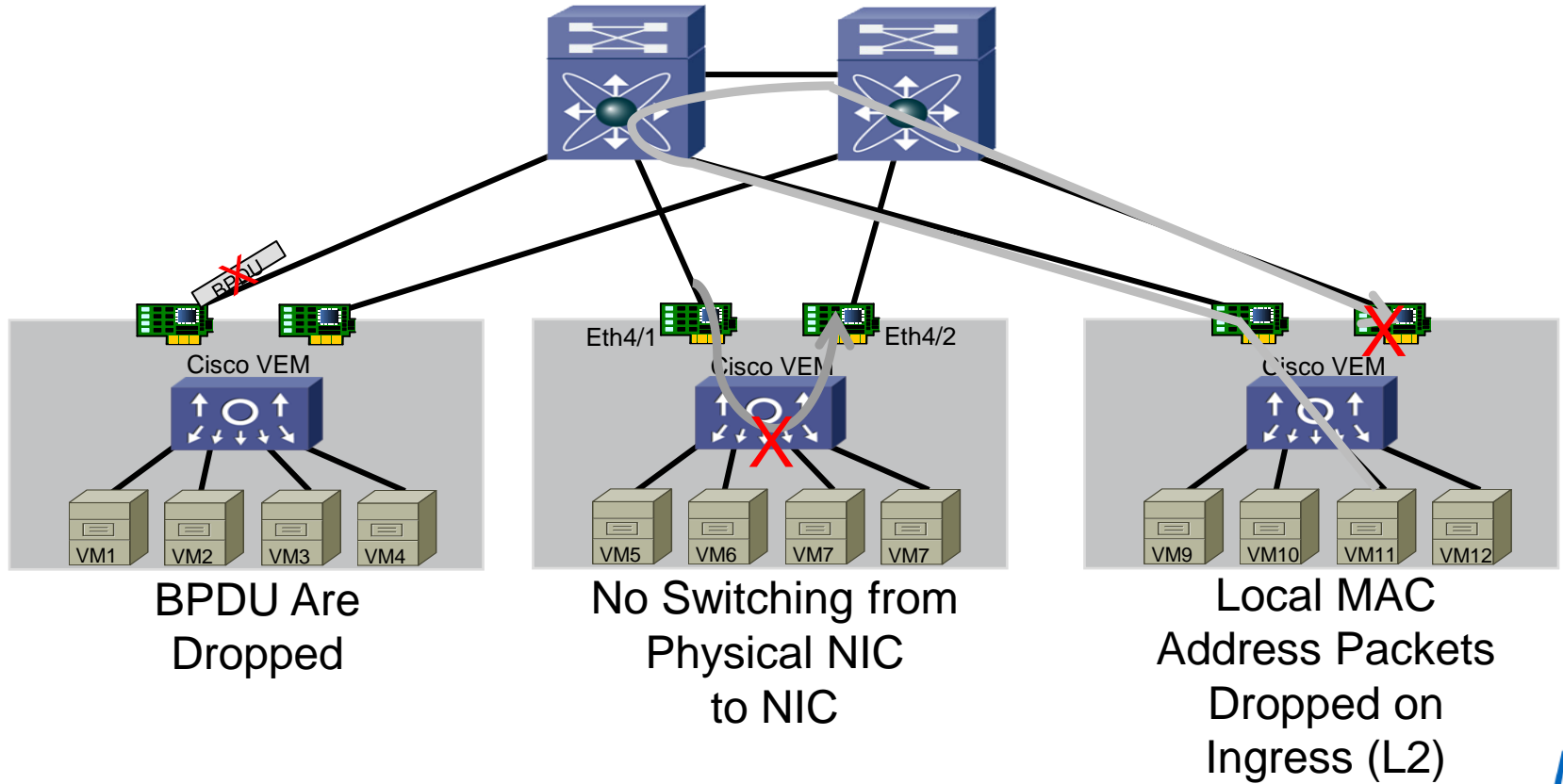
Port-Profiles Deploying and Troubleshooting

Switch Interface Types

- Ethernet Port (eth)
 - Correspond to the physical NIC interfaces leaving the server
 - Specific to each “module” or VEM
 - Up to 32 physical ports supported per host
- Port Channel (po)
 - Aggregation of physical Ethernet ports
 - Up to eight Port Channels per host
- Virtual Ethernet Port (veth)
 - One per virtual NIC interfaces (vnic) including service console and vmknics
 - Notation is VethX
 - No module number is assigned to keep naming persistent as VMs move between modules



Loop Prevention Without STP



Uplink(eth) Port-Profile Troubleshooting

- Port-Profiles with multiple NICs need a port-channel
- Duplicate packets kick in déjà vu driver
 - Requires extra CPU processing
 - Fills the logs
- Example: Eth 6/1 and Eth 6/3 added to below Port-Profile
- Also same issue if you overlap VLANs in different Port-Profiles

WRONG

```
port-profile type ethernet uplink-nopc
vmware port-group
switchport mode trunk
switchport trunk allowed vlan 1-3967,4048-4093
no shutdown
system vlan 11
state enabled
```

RIGHT

```
port-profile type ethernet uplink-nopc
vmware port-group
switchport mode trunk
switchport trunk allowed vlan 1-3967,4048-4093
channel-group auto mode on mac-pinning
no shutdown
system vlan 11
state enabled
```

Cisco Nexus 1000V System VLANs

- System VLANs enable interface connectivity before an interface is programmed
- Addresses chicken and egg issue
- Port profiles that contain system VLANs are “system port profiles”
- System port-profiles become part of the opaque data
 - VEM will load system port-profiles and pass traffic even if VSM is not up
- System vlans must be set on egress and ingress port-profiles

System VLAN Example

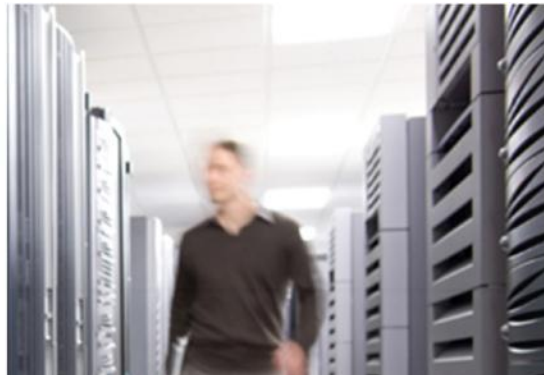
- Migrate VMware vmkernel mgmt interface to VEM
- Uplink port-profile must define VLAN 2 as system vlan

```
port-profile type ethernet uplink-pinning
  vmware port-group
  switchport mode trunk
  switchport trunk allowed vlan all
  channel-group auto mode on mac-pinning
  no shutdown
  system vlan 2,10,150-151
```

```
n1000v# show run port-profile vmk-mgmt
port-profile type vethernet vmk-mgmt
  vmware port-group
  switchport mode access
  switchport access vlan 2
  no shutdown
  system vlan 2
```


VMware DVS Max-Port Issues

- Default to 32 max-ports per port-profile
- Counts toward the maximum number of VMware DVS ports
 - pre-provisioned
 - Some ports are consumed when you add an ESX host to the DVS
- Max-ports under “svs connection <name>”
 - Allows you to increase the ports of the VMware DVS
- Port-binding “auto expand” in veth port-profiles
 - N1KV dynamically adds ports as VMs are added
 - Set port-binding as default with “port-profile default port-binding static auto expand”

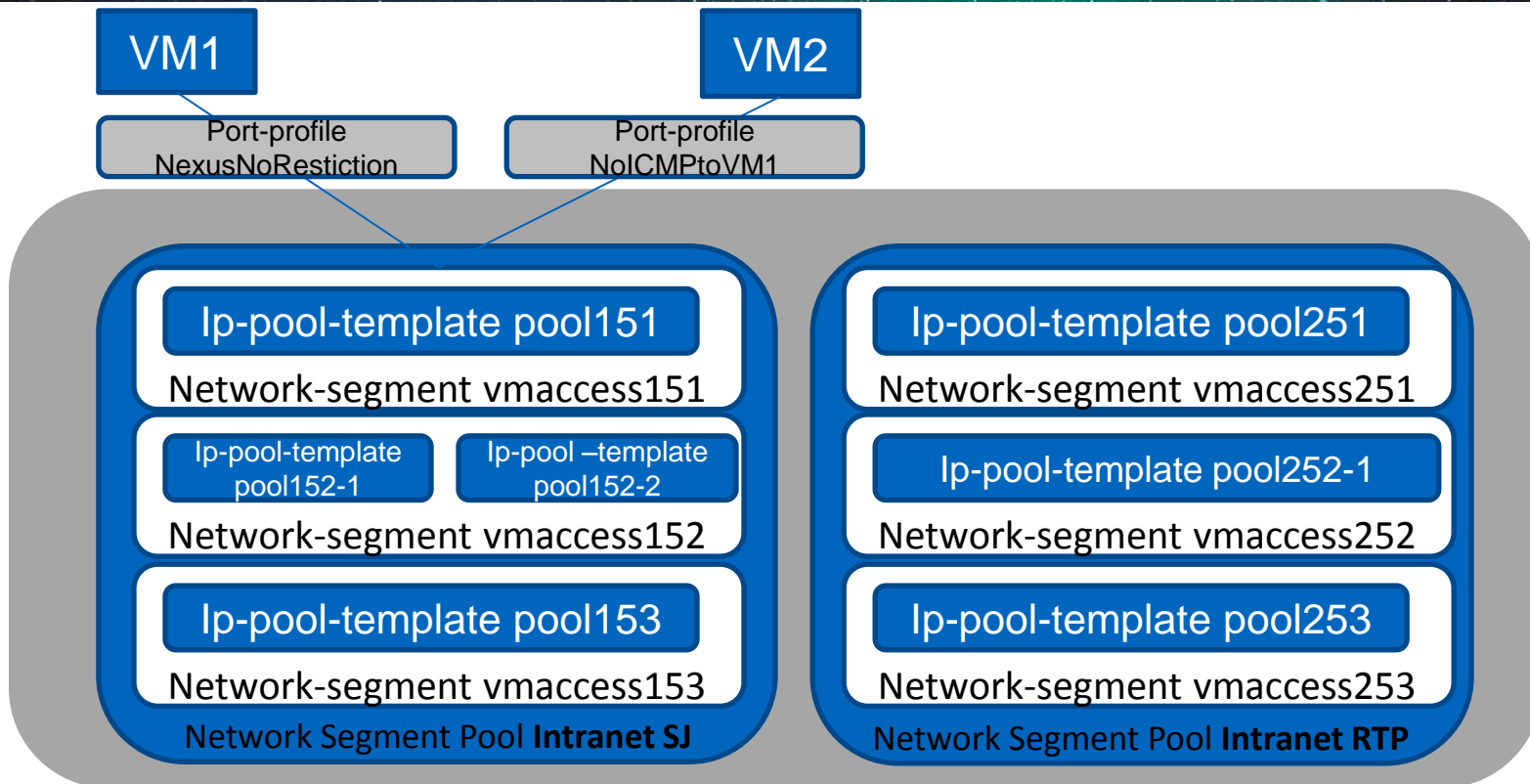


Port Profiles With Microsoft SCVMM 2012

Microsoft SCVMM Networking Concepts

- Logical Networks
 - Represents a network with a certain type of connectivity characteristics (for eg. DMZ network, intranet, isolation)
- Network Sites
 - An instantiation of a Logical network on a set of host-groups (for eg. hosts in a POD)
- VM Networks
 - What VMs get bound to. Connects to a network site and ip-pool
 - Defines connectivity
- Port Classification
 - Defines port specific behaviour like QOS or ACLs
- IP-Pools
 - Defines an IP Range, Gateway, DNS, DHCP, and WINS server

Visualising How It All Fits



Logical Network – “Intranet”

Eth Port Profiles ESX vs Hyper-V

Splitting “Network Connectivity” and “Policy”

Current N1KV/ESX Version

```
# port-profile type eth uplink
switchport mode access
switchport trunk allowed vlan all
vmware port-group
channel-group mode auto on mac-pinning
no shut
system vlan 2,10,11
state enabled
```

- VLANs
 - Created on demand
 - Added to uplink on creation of network-segments
- System VLANs are also added to uplink on creation of network-segment

N1KV/Hyper-V Version

```
#port-profile type ethernet Uplink-346
channel-group auto mode on mac-pinning
no shutdown
state enabled
```

```
# nsm network uplink nsm-uplink-346
import port-profile Uplink-346
allow network segment pool vm-nsp-data
switchport mode trunk
system network uplink
publish network uplink
```

Veth Port Profiles ESX vs Hyper-V

Splitting “Network Connectivity” and “Policy”



```
# port-profile type veth db-client
switchport mode access
switchport access vlan 10
ip port access-group dbclient in
no shut
system vlan
state enabled
```

```
# port-profile type veth db-server
switchport mode access
switchport access vlan 10
ip port access-group dbserver in
no shut
system vlan 10
state enabled
```

```
#nsm network segment db-network
switchport access vlan 10
system network segment
publish network-segment
member-of network segment pool IntranetRTP
```

```
# port-profile type veth db-client
ip port access-group dbclient in
no shut
state enabled
```

```
# port-profile type veth db-server
ip port access-group dbserver in
no shut
state enabled
```

Hyper-V – It's Still A Port Profile

- On the VSM everything still gets assigned to a Port Profile
- A Dynamic Port Profile gets created
 - Combination of the Network Segment and the Port Profile
 - One Dynamic port profile for each combination
 - Multiple VMs can be attached to the same Dynamic port profile

```
n1kv-test# show run port-profile dynpp_308ad66b-7c42-4067-90af-13f7a6e59afe_e11eeb54-431d-4310-bae2-e07583ff41a8

port-profile type vethernet dynpp_308ad66b-7c42-4067-90af-13f7a6e59afe_e11eeb54-431d-4310-bae2-e07583ff41a8
  inherit port-profile NoRest-unicast-norest
  switchport mode access
  switchport access vlan 151
  guid 930387de-d44c-4674-9f02-adce1e794c31
  description NSM created profile. Do not delete.
  state enabled

interface Vethernet12
  inherit port-profile dynpp_308ad66b-7c42-4067-90af-13f7a6e59afe_e11eeb54-431d-4310-bae2-e07583ff41a8
```



Port Channels

Port Channels

- LACP Port-channels
 - Upstream switch support and configuration
- VPC – MAC Pinning
 - Works with any upstream switch
 - Allows for pinning of veths (VM) to specific links.
- VPC – Host Mode CDP/Manual
 - NIC association is either Manual or CDP
 - Multiple connections per physical switch require a port-channel

LACP Troubleshooting

- Do not use Network State Tracking(NST) with LACP
- LACP Port-Channel configured on the upstream switches
- Port-profile created with “channel-group auto mode active”
- Beware of LACP fast/slow timers
- On the VEM
 - vemcmd show lacp
- On the VSM and Upstream Switch
 - show port-channel summary
 - show lacp counters/neighbor
 - Are you seeing LACP PDUs?

VPC - MAC Pinning

- Use when upstream switch does not support LACP
- Each Eth interface added is a unique Sub Group
 - SGID # get assigned based off vmnic#
- Use “pinning id” command under vethernet port-profile
 - Pins the VM to a particular uplink
 - Ordered list for backup

```
n1kv-13(config-port-prof)# pinning id 0 backup 1 2
```
- Default assignment is Round Robin to an SGID
- New command to make SGID # relative

```
n1kv-13(config-port-prof)# channel-group auto mode on mac-pinning relative
```

Port Channels – How To View Pinning

```
n1kv-13# show int virtual pinning module 5
```

```
-----  
Veth      Pinned      Associated PO List of  
          Sub Group id  interface      Eth interface(s)  
-----  
Veth2     0            Po5            Eth5/1  
Veth4     2            Po5            Eth5/3  
Veth5     0            Po5            Eth5/1  
Veth6     2            Po5            Eth5/3  
Veth7     0            Po5            Eth5/1
```

Port Channels – Best Practice

- If the upstream switch can be clustered (VPC, VBS Stack, VSS) use LACP
- If the upstream switch can NOT be clustered use MAC-PINNING
- For Cisco UCS only MAC-PINNING is supported for blades
- Create channel-groups in port-profile
 - VSM builds the port-channel
- Configure all physical ports in port-channel identically

Spanning-Tree And BPDU – Best Practice

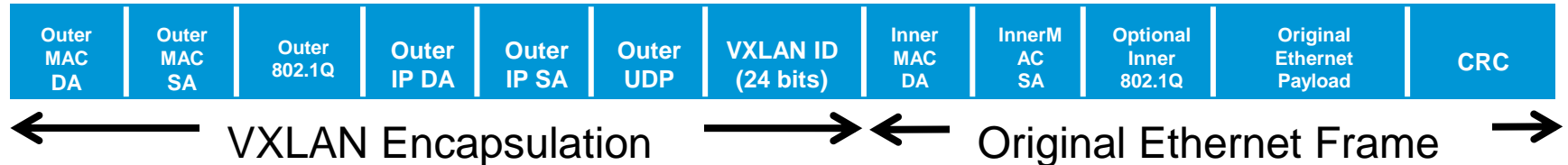
- Mandatory Spanning-tree settings per port
 - IOS set STP portfast
 - **cat65k-1(config-if)# spanning-tree portfast trunk**
 - NXOS set port type edge
 - **n5k-1(config-if)# spanning-tree port type edge trunk**
- Misconfiguration can cause unexpected blocked ports during changes in the network
- Short flaps in connectivity are telltale signs of possible misconfiguration



Virtual Extensible LAN (VXLAN)

What Is VXLAN?

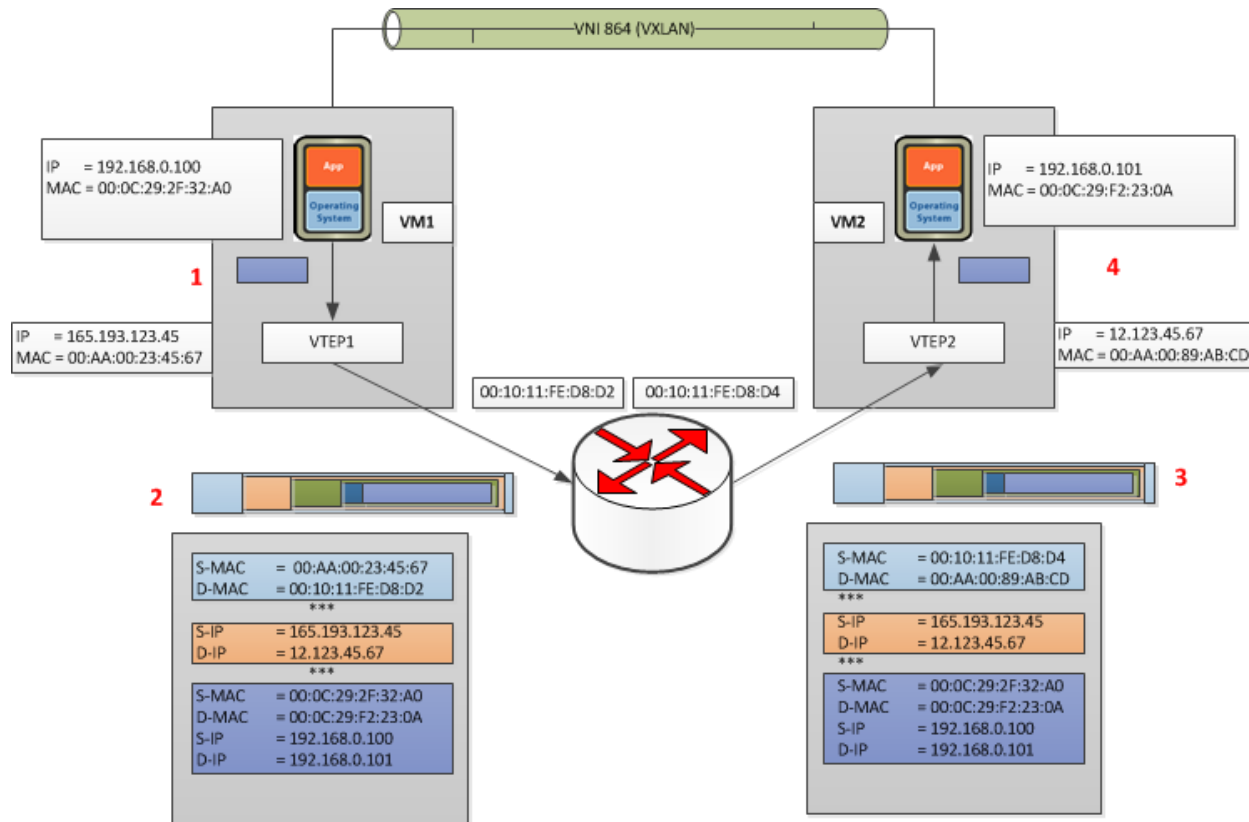
- The X stands for eXtensible
 - Scale
 - More layer 2 segments than VLANs
 - Wider stretch than VLANs
- VXLANs are an Overlay Network technology
 - MAC Over IP/UDP
- VXLAN specification submitted by multiple vendors



VXLAN Checklist

- Multicast enabled core
 - Multiple segment can be mapped to single group
 - If VXLAN transport is contained to a single VLAN, IGMP Querier must be enabled on that VLAN
 - If VXLAN transport is traversing routers, multicast routing must be enabled
- Increase MTU across end to end physical and virtual network
 - 50 bytes overhead on top of the VNIC MTU size
- 5-tuple hashing on LACP ISLs
 - Leverage inner packet for load balancing
- Proxy ARP must be enabled on first hop router
 - Only if VXLAN traffic traversing router

VXLAN on Nexus 1000v



Deployment Modes: When To Use MAC Distribution?

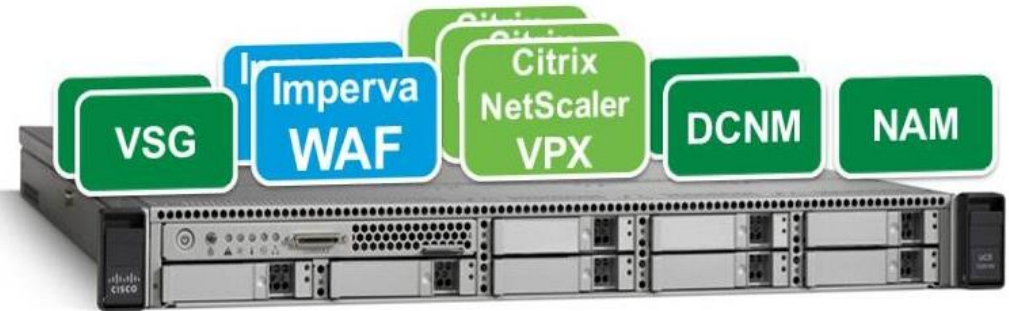
- Unicast Mac Distribution Mode
 - VEMs keep a list of where each VM is in each VXLAN
 - No Flood and Learn
- MAC distribution will provide best performance
- VXLAN traffic cannot span single Nexus 1000V
- Two caveats
 - No veth VXLAN trunk mode support with MAC distribution
 - Won't work with Microsoft NLB
- If above issues are not a problem use MAC Distribution



Nexus 1010 and 1110

Cisco Nexus 1010/1010-X/1110-S/1110-X

- Based off UCS C2x0 server
- Virtual Service Blade (VSB) Support
 - 1010/1110-S supports 6
 - 1010/1110-X supports 10
- Supported VSBs
 - Nexus 1000V VSM
 - VXLAN Gateway
 - Virtual Security Gateway (VSG)
 - Network Analysis Module (NAM)
 - Data Centre Network Manager (DCNM)
 - Imperva WAF
 - Citrix NetScaler



Cisco Nexus 1x10

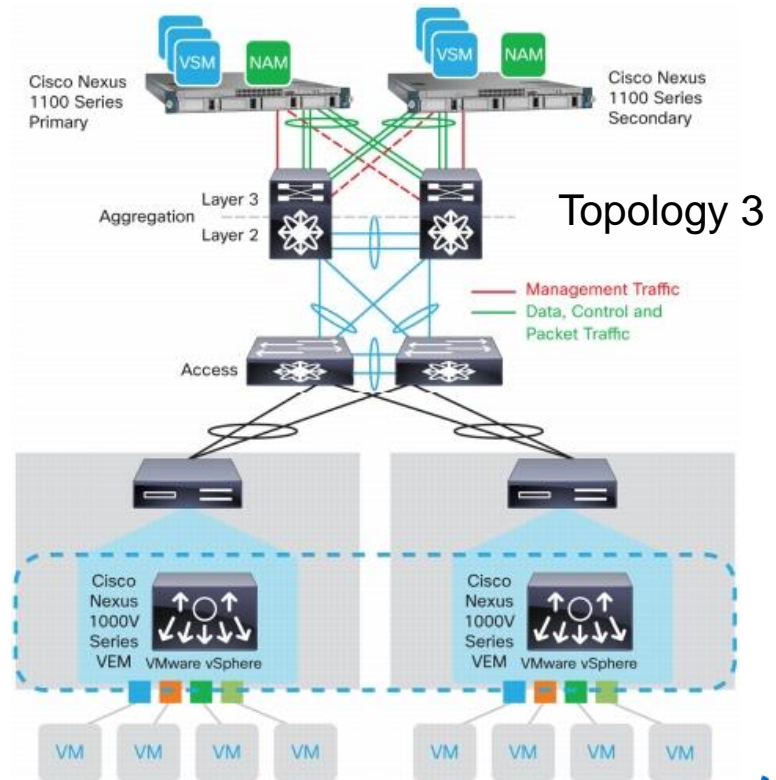
- Must be deployed in pairs
- Deploy in the Aggregation Layer
- Must be in the same L2 domain for management and control
- Uses same HA mechanism as VSM with domain id and control vlan
 - Do not overlap the domain id between a 1x10 and a VSM
- Not supported
 - Primary and Secondary VSM on same 1x10
 - Primary VSM on ESX and Secondary VSM on 1x10 or vice versa

Nexus 1x10 Network Classes And Topologies

- Management
 - Carries the mgmt 0 interface of the 1x10
 - Carries the mgmt 0 traffic for all VSMS installed
- Control
 - Carries all the control and packet traffic for the VSMS installed on the 1x10
 - Carries control traffic for HA between primary and secondary 1x10
- Data
 - Used by Virtual Service Blades (VSB) other than VSM
- 5 choices of Network Topologies

Nexus 1x10 Recommendations

- Only planning on VSM/NAM VSBs?
 - Topology 3 gives best bandwidth and redundancy for control VLAN
 - More configuration steps
- Flexible allows any configuration
 - Recommend using port-channels
 - VSM latency is key over bandwidth
- Use VPC or VSS upstream if available



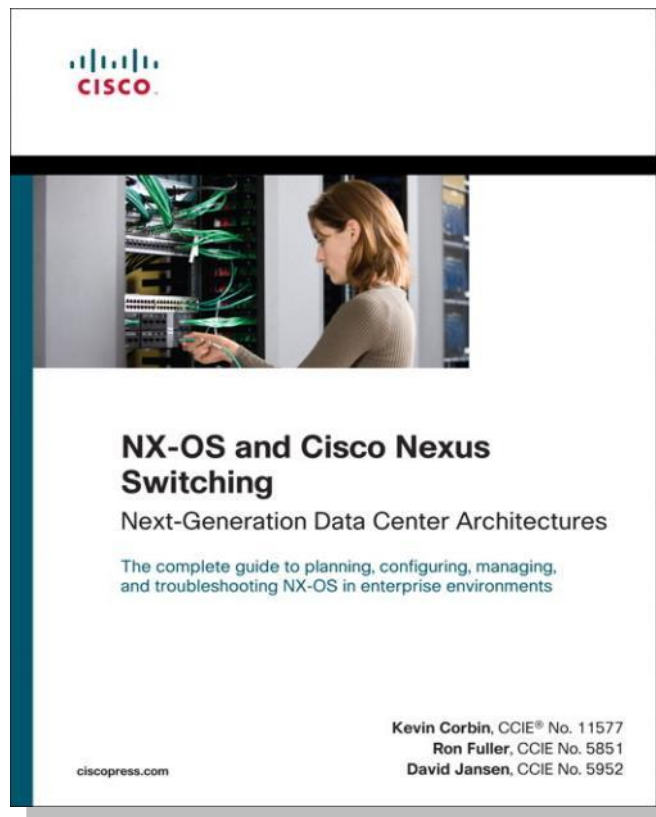


Wrap Up

N1K Public Links

- N1K Download: www.cisco.com/go/1000vdownload
- N1K Product Page: www.cisco.com/go/1000v
- N1K Community: www.cisco.com/go/1000vcommunity
- N1K Twitter www.twitter.com/official_1000V
- N1K Webinars: www.cisco.com/go/1000vcommunity
- N1K Case Studies: www.tinyurl.com/n1k-casestudy
- N1K Whitepapers www.tinyurl.com/n1k-whitepaper
- N1K Deployment Guide: www.tinyurl.com/N1k-Deploy-Guide
- N1K on UCS Best Practices: www.tinyurl.com/N1k-On-UCS-Deploy-Guide
- VXLAN Web Conference: www.brighttalk.com/webcast/279/41277
- Cisco N1K CloudLab: cloudlab.cisco.com

Recommended Reading For BRKVIR-3013





Q & A

Complete Your Online Session Evaluation

Give us your feedback and receive a Cisco Live 2014 Polo Shirt!

Complete your Overall Event Survey and 5 Session Evaluations.

- Directly from your mobile device on the Cisco Live Mobile App
- By visiting the Cisco Live Mobile Site www.ciscoliveaustralia.com/mobile
- Visit any Cisco Live Internet Station located throughout the venue

Polo Shirts can be collected in the World of Solutions on Friday 21 March 12:00pm - 2:00pm



Learn online with Cisco Live!

Visit us online after the conference for full access to session videos and presentations.

www.CiscoLiveAPAC.com





Appendix

Troubleshooting VSM to VEM Connectivity With L2 Control

L2 Control VEM – VSM Troubleshooting Steps

1. VSM MAC address
2. VSM is connected to vCenter
3. VSM has Control VLAN on right interface
4. Uplink port-profile has Control vlan
5. VEM sees control VLAN
6. VEM and VSM see each others MAC
7. Physical network sees VEM and VSM MAC
8. VSM sees heartbeat messages from VEM

Step 1: VSM MAC

- Need for L2 troubleshooting
- On VSM run show svcs neighbors
- Its the AIPC Interface MAC

```
n1kv-12# show svcs neighbors
```

```
Active Domain ID: 422
```

```
AIPC Interface MAC: 0050-56a9-2535
```

```
Inband Interface MAC: 0050-56a9-2537
```

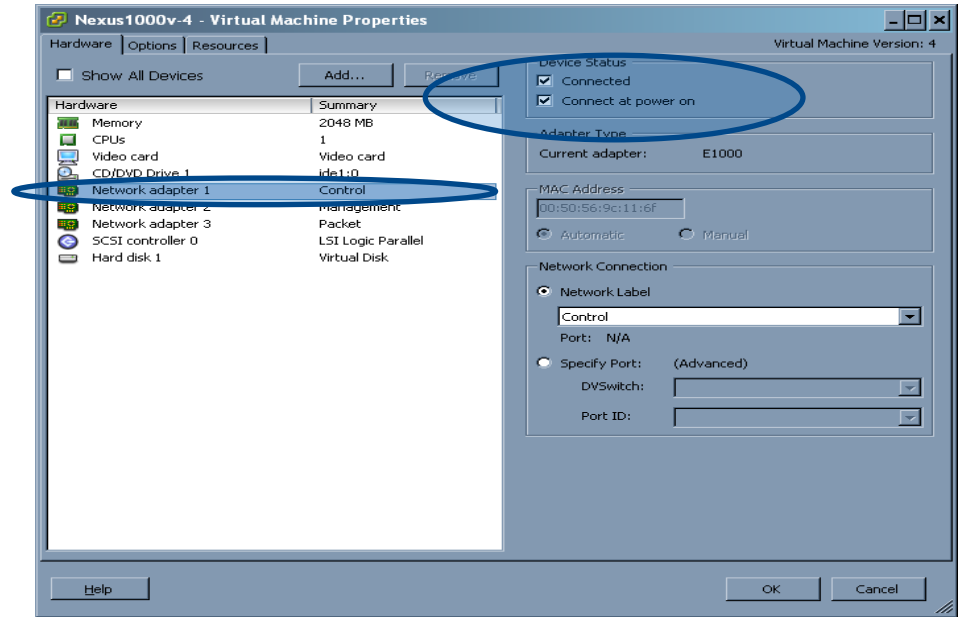
Step 2: VSM – vCenter Connectivity

- Verify VSM is connected to vCenter

```
n1kv-12# show svcs connections
connection VC:
  ip address: 172.18.217.241
  remote port: 80
  protocol: vmware-vim https
  certificate: default
  datacenter name: Harrington
  admin:
  max-ports: 8192
  DVS uuid: 3e 80 29 50 ad 9f f9 7f-43 d6 9b 6d a2 af cb 3e
  config status: Enabled
  operational status: Connected
```

Step 3: Verify VSM VM Control interface

- 1st interface listed is Control Interface
- Interface connected?



Step 4: Verify Uplink Port-Profile

- The first ESX interface added to the N1KV must have Control VLAN
- Verify uplink port-profile has Control VLAN defined and system VLAN

```
n1kv-12# show run port-profile uplink

version 4.2(1)SV1(5.1)
port-profile type ethernet uplink
  vmware port-group
  switchport mode trunk
  switchport trunk allowed vlan 1-3967,4048-4093
  no shutdown
  system vlan 2
  state enabled
```

Step 5: Verify VEM Sees Control VLAN

- Verify VEM sees control VLAN with commands
 - vemcmd show card
 - vemcmd show port
 - vemcmd show trunk

Vemcmd Show Card

- Control, packet vlans and domain-ID match with VSM

```
[~ # vemcmd show card
Card UUID type 2: 33393138-3335-5553-4537-31314e343636
Card name: cae-esx-154
Switch name: n1kv-12
Switch alias: DvsPortset-0
Switch uuid: 3e 80 29 50 ad 9f f9 7f 43 d6 9b 6d a2 af cb 3e
Card domain: 422
Card slot: 5
VEM Tunnel Mode: L2 Mode
VEM Control (AIPC) MAC: 00:02:3d:11:a6:04
VEM Packet (Inband) MAC: 00:02:3d:21:a6:04
VEM Control Agent (DPA) MAC: 00:02:3d:41:a6:04
..
..
Card control VLAN: 2
Card packet VLAN: 2
```

MAC the VSM
should learn for
VEM

Vemcmd Show Port-Old

- Ports with LTLs 8, 9,10 are UP and CBL states are 1.
- ESX Physical ports are UP and CBL states 1.

```
~ # vemcmd show port-old
```

LTL	IfIndex	Vlan/ SegId	Bndl	SG_ID	Pinned_SGID	Type	Admin	State	CBL	Mode	Name
6	0	1 T	0	32	32	VIRT	UP	UP	1	Trunk	vnc
8	0	3969	0	32	32	VIRT	UP	UP	1	Access	
9	0	3969	0	32	32	VIRT	UP	UP	1	Access	
10	0	2	0	32	32	VIRT	UP	UP	1	Access	
11	0	3968	0	32	32	VIRT	UP	UP	1	Access	
12	0	2	0	32	32	VIRT	UP	UP	1	Access	
13	0	1	0	32	32	VIRT	UP	UP	0	Access	
14	0	3971	0	32	32	VIRT	UP	UP	1	Access	
15	0	3971	0	32	32	VIRT	UP	UP	1	Access	
16	0	1 T	0	32	32	VIRT	UP	UP	1	Trunk	ar
17	25010000	1 T	0	32	32	PHYS	UP	UP	1	Trunk	vmnic0

- Local Target Logic (LTL) is an index to address a port, or group of ports. Data path lookup engine takes LTL as input, and gives LTL as output.
- LTL scheme: [0-14: internal ports] [15-271: pNICs, VMs, etc...]

Vemcmd Show Trunk

- Control and packet are CBL states 1 on the physical ports.

```
~ # vemcmd show trunk
Trunk port 6 native_vlan 1 CBL 1
vlan(1) cbl 1, vlan(3970) cbl 1, vlan(3969) cbl 1, vlan(3968) cbl 1, vlan(3971) cbl 1,
vlan(11) cbl 1, vlan(10) cbl 1, vlan(150) cbl 1, vlan(2) cbl 1, vlan(151) cbl 1,
vlan(152) cbl 1, vlan(153) cbl 1, vlan(154) cbl 1, vlan(155) cbl 1,
Trunk port 16 native_vlan 1 CBL 1
vlan(1) cbl 1, vlan(3970) cbl 1, vlan(3969) cbl 1, vlan(3968) cbl 1, vlan(3971) cbl 1,
vlan(11) cbl 1, vlan(10) cbl 1, vlan(150) cbl 1, vlan(2) cbl 1, vlan(151) cbl 1,
vlan(152) cbl 1, vlan(153) cbl 1, vlan(154) cbl 1, vlan(155) cbl 1,
Trunk port 17 native_vlan 1 CBL 1
vlan(1) cbl 1, vlan(11) cbl 1, vlan(10) cbl 1, vlan(150) cbl 1, vlan(2) cbl 1,
vlan(151) cbl 1, vlan(152) cbl 1, vlan(153) cbl 1, vlan(154) cbl 1, vlan(155) cbl 1,
```

- vemcmd show port vlans

```
~ # vemcmd show port vlans
      LTL      VSM Port  Mode      Native  VLAN  Allowed
      17      Eth5/1   T         VLAN   State  Vlans
      17      Eth5/1   T         1      FWD   2,10-11,150-155
~ #
```


Step 6: VEM And VSM See Each Other's MAC

- Is the VEM learning the MAC of the VSM?
- On VEM “**vemcmd show l2 <control-vlan>**” do you see the mac of the VSM?

```
~ # vemcmd show l2 2
Bridge domain      9 brtmax 4096, brtcnt 32, timeout 300
VLAN 2, swbd 2, ""
Flags: P - PVLAN  S - Secure  D - Drop
      Type      MAC Address  LTL  timeout  Flags  PVLAN
      Static    00:02:3d:21:a6:04  12   0        0
      Dynamic   00:50:56:a9:25:35  17   1        1
```

VEM and VSM See Each Other's MAC

- Is the VSM learning the MAC of the VEM?

```
n1kv-12# show mac address-table vlan 2
VLAN      MAC Address      Type    Age      Port
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
2         0002.3d21.a604   static  0        N1KV Internal Port      5
2         0002.3d41.a604   static  0        N1KV Internal Port      5
```

Step 7: Physical Switch Mac Table

- Check the physical switch MAC address table
- Are the MACs of the VEM and VSM getting learned by the physical switches in the right VLANs?

```
cae-cat6k-1#show mac-address-table vlan 2
```

```
Legend: * - primary entry
```

```
age - seconds since last seen
```

```
n/a - not available
```

vlan	mac address	type	learn	age	ports	
*	2	0050.5677.7770	dynamic	Yes	360	Gi3/48
*	2	0050.56a9.2535	dynamic	Yes	0	Gi4/9
*	2	3333.0000.0016	static	Yes	-	Switch,Stby-Switch
*	2	0002.3d41.a604	dynamic	Yes	0	Gi1/4

Step 8: VEM – VSM Heartbeat

- One Heartbeat per second per VEM from VSM
- Timeout for VEM from VSM is 6 seconds of missed heartbeats
- After 6 seconds VSM will drop VEM
- Use vempkt capture to view heartbeats
- SPAN physical switch ports for heartbeats

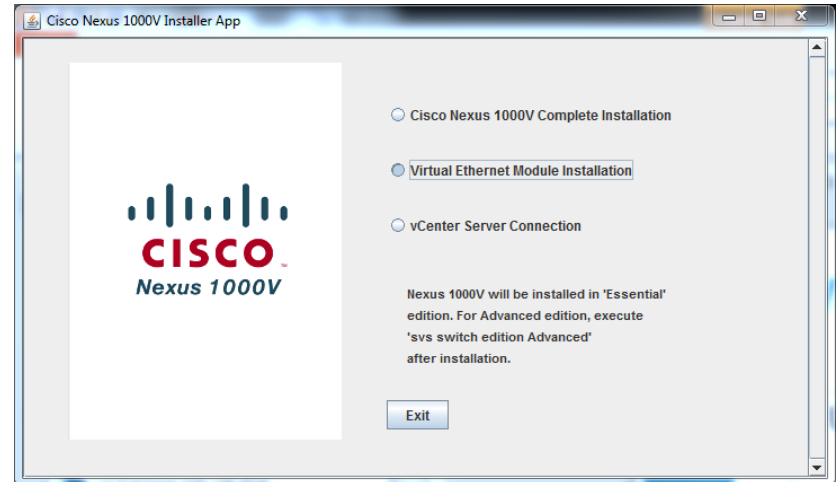


Appendix

Miscellaneous Commands

VEM Installation – Nexus 1000V Installer App

- Uses VUM or httpClient
 - Tries VUM first
- httpClient needs to be enabled on every ESXi host
- Requires administrator privileges to the ESXi host
- VSM must be connected to vCenter

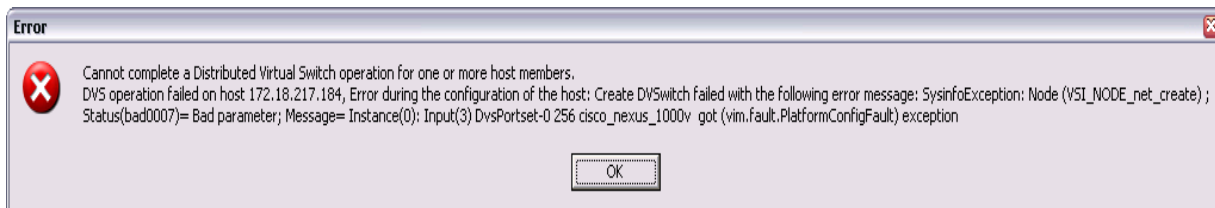


VEM Installation – ESXi Stateless

- VMware introduced Stateless ESXi with version 5
- ESXi PXE boots
- No information is stored on local disks
- VEM module has to be built into the boot image
- This is possible using VMware PowerCLI
- Instructions are in the Install and Upgrade Guide
 - http://www.cisco.com/en/US/docs/switches/datacenter/nexus1000/sw/4_2_1_s_v_1_5_1/install_upgrade/vsm_vem/guide/n1000v_installupgrade.html

VEM Installation DVS Error

- If you are using VUM check the logs
 - VUM could not find the right VEM version
 - Check the CISCO and CSCO directories
- Make sure the VSM is connected to vCenter
- Make sure Cluster HA, DRS, and DPM are disabled



VEM Installation - Microsoft

- VEM is a simple Windows installation
- Manually double click the msi file and install
- SCVMM will automatically install VEM if it's not manually installed
- Upgrades will leverage Windows Server Update Services (WSUS)

How To Hard Code VEM To Module Number

- Control network should have low latency and available bandwidth
- VEM and VSM running on the same versions
- Upstream switch ports configured identically
- On UCS make Service Profile does not contain “Dynamic VNICs”
 - VEM and VM-FEX are mutually exclusive
- Hard code VEM to module number with....

```
N1000v-mv# config t
n1000v-mv(config)# vem 12
n1000v-mv(config-vem-slot)# host vmware id 33393138-3335-5553-4537-30354E375832
```

VEM Removal

- VEM can only be removed manually
- Removing a host from the Nexus 1000V using vSphere does not remove the VEM
- ESX/ESXi 4/4.1
 - “vem-remove -s -r”
 - Will unload and remove the driver
- ESXi 5
 - Host has to be in Maintenance Mode
 - “esxcli software vib remove -n cisco-vem-v140-esx”

Assigning A Veth To A Particular VM

- By default veths are assigned in order
- To specify do the following
 - Make sure the veth you want to use is free
 - Make sure old veth is down with reason nonParticipating
 - Can be done by disconnecting vnic or powering down VM
 - Get veth dvs port # with “show int veth #”
 - Remove vmware dvport config line on the old veth
 - switch(config)# interface veth1
 - switch(config-if)# no vmware dvport 32
 - Create the veth and add dvport
 - switch(config)# interface veth100
 - switch(config-if)# vmware dvport 32
 - Power on VM or connect vnic

Port Channels – vPC HM CDP/Manual

- vPC-HM uses Service Group (SG)
 - Service Group is a collection of Ethernet interfaces from ESX host
 - One Service Group per physical path
- CDP is used to determine SG membership
 - Can be a 60 second delay while VSM determines NIC membership because of CDP
- Can configure SG membership manually for switches without CDP support
- Multiple links per physical path must be configured as a port-channel upstream

iSCSI Support

- iSCSI is supported
- Provide automatic multipathing capability for “veth” type port-profiles
 - “capability iscsi-multipath”
- When turned on it attempts to balance connections across uplinks
- Works on a per vlan basis
 - So access to iSCSI storage paths has to be via the same vlan
 - One path cannot be vlan 10 and the other vlan 11
- Verify with “vemcmd show iscsi pinning”
- Vemcmd set command to change pinning
 - vemcmd set iscsi pinning <vmk-ltl> <vmnic-ltl>

Jumbo Frames Support

- MTU setting for “eth” type port-profile
 - Simply use “mtu size” in port-profile and nothing else
 - Add system vlan directive to Port-profile if needed
- “System jumbo mtu” global setting – all versions
 - Sets the system wide jumbo mtu size
 - Generally do not need to change

Jumbo Frames Support - Example

- Eth port-profile with jumbo MTU

```
n1000v-AV(config)# port-profile type eth uplink-jumbo
n1000v-AV(config-port-prof)# switchport mode trunk
n1000v-AV(config-port-prof)# switchport trunk allowed vlan 1,3-149,151-1000
n1000v-AV(config-port-prof)# vmware port-group
n1000v-AV(config-port-prof)# mtu 9000
n1000v-AV(config-port-prof)# no shut
n1000v-AV(config-port-prof)# system vlan 3,10
n1000v-AV(config-port-prof)# state enabled
```

- Use “esxcfg-nics -l” on the ESX host to confirm

```
[root@cae-esx-133 ~]# esxcfg-nics -l
Name      PCI      Driver   Link Speed  Duplex  MAC Address      MTU   Description
vmnic0    08:00.00 enic     Up    10000Mbps  Full    00:25:b5:00:00:1e 1500   Cisco Systems Inc 10G Ethernet NIC
vmnic1    09:00.00 enic     Up    10000Mbps  Full    00:25:b5:00:00:0e 1500   Cisco Systems Inc 10G Ethernet NIC
vmnic2    0a:00.00 enic     Up    10000Mbps  Full    00:25:b5:00:00:0f 9000   Cisco Systems Inc 10G Ethernet NIC
```


Port-Profile Using Weighted QOS

- ESX 4.1 introduced NetIOC (Network I/O Control)
 - Allows for QOS to classify traffic based on type
- On Nexus 1000V its called QOS Fair Weighted Queuing
 - Works on egress uplink ports only
 - Easy to setup and configure
- Use “vemcmd show qos queue-rate ‘Itl’” to verify VEM is matching packets

Port-Profile Using Weighted QOS

- Configuration Steps to limit vMotion traffic

```
n1kv-13(config)# class-map type queuing match-all vmotion-class
n1kv-13(config-cmap-que)# match protocol ?
  n1k_control    N1K control traffic
  n1k_mgmt       N1K management traffic
  n1k_packet     N1K inband traffic
  vmw_ft        VMware fault tolerance traffic
  vmw_iscsi     VMware iSCSI traffic
  vmw_mgmt       VMware management traffic
  vmw_nfs       VMware NFS traffic
  vmw_vmotion   VMware vmotion traffic

n1kv-13(config-cmap-que)# match protocol vmw_vmotion
n1kv-13(config-cmap-que)# policy-map type queuing vmotion-policy
n1kv-13(config-pmap-que)# class type queuing vmotion-class
n1kv-13(config-pmap-c-que)# bandwidth percent 50

n1kv-13(config)# port-profile type eth uplink-vpc
n1kv-13(config-port-prof)# service-policy type queuing output vmotion-policy
```

VSM – VSM Heartbeat

My CP:

```
slot: 0
domain: 184
role: primary
status: RDN_ST_AC
state: RDN_DRV_ST_AC_SB
intr: enabled
power_off_reqs: 0
reset_reqs: 1
```

Active VSM

Other CP:

```
slot: 1
status: RDN_ST_SB
active: true
ver_rcvd: true
degraded_mode: false
```

Standby VSM

Redun Device 0:

```
name: ha0
```

pdev: bc1bb000

```
alarm: false
mac: 00:50:56:8e:5e:f5
tx_set_ver_req_pkts: 13
tx_set_ver_rsp_pkts: 2
tx_heartbeat_req_pkts: 168155
tx_heartbeat_rsp_pkts: 318
rx_set_ver_req_pkts: 2
rx_set_ver_rsp_pkts: 1
rx_heartbeat_req_pkts: 318
rx_heartbeat_rsp_pkts: 168148
rx_drops_wrong_domain: 0
rx_drops_wrong_slot: 0
rx_drops_short_pkt: 0
rx_drops_queue_full: 0
rx_drops_inactive_cp: 0
rx_drops_bad_src: 0
rx_drops_not_ready: 0
rx_unknown_pkts: 0
```



Appendix

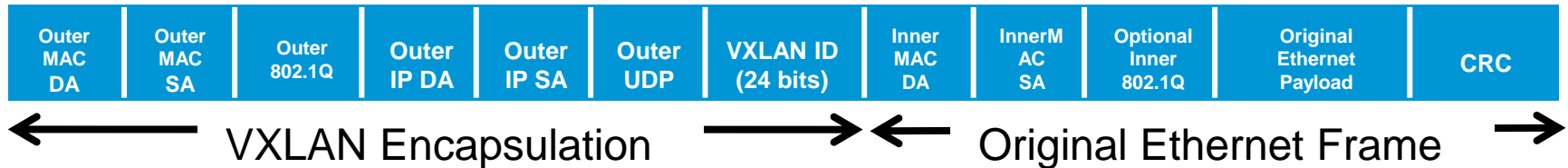
Virtual Extensible LAN (VXLAN)

What is VXLAN?

- What is it?
- Why do I need it
- What is required
- How to set it up
- Troubleshooting tips

Virtual Extensible Local Area Network (VXLAN)

- Ethernet in IP overlay network
 - Entire L2 frame encapsulated in UDP (port 8472)
 - 50 bytes of overhead
- Include 24 bit VXLAN Identifier
 - 16 M logical networks
 - Mapped into local bridge domains
 - Unique multicast group per segment
- VXLAN can cross Layer 3
- Tunnel between VEMs
 - VMs do NOT see VXLAN ID



Virtual Extensible Local Area Network (VXLAN)

- Each overlay network is known as a VXLAN segment
- Each VXLAN segment identified by a 24-bit segment ID (VNI)
- VXLAN traffic carried between VXLAN Tunnel Endpoints (VTEP)
- VEM module acts as the VTEP
- VM traffic is carried over point to point tunnels between VTEPs
 - VM to VM traffic is encapsulated in a VXLAN header

Why Do I Need VXLAN?

- Fast on-demand provisioning of large number of isolated L2 networks
- On demand networks without physical network re-configuration
 - Enables network snapshot of applications
 - Move from dev to test to production with no re-IP
- Massive scale for multi-tenant environments
 - Tenants can use the same IP addresses
- Allows virtual L2 network to stretch across physical L2 boundaries

Deployment Modes: Multicast Or Unicast?

- Multicast used to be required for unknown broadcast/unicast on VXLAN
- N1KV 2.2 introduced Unicast Mode and Unicast Mac Distribution Mode
- Multicast
 - Needs Multicast configured through complete network
 - IGMP Querier in VLAN
 - Multicast routing and proxy ARP across subnets
 - VTEPs all join multicast group
 - Works across VSMS
- Unicast Mode
 - VEMs flood each other directly for unknown broadcast/unicast
 - Keep a list of other VEMs in each VXLAN
- Unicast Mac Distribution Mode
 - VEMs keep a list of where each VM is in each VXLAN
 - No Flood and Learn

Deployment Modes: When To Use MAC Distribution?

- MAC distribution will provide best performance
- VXLAN traffic cannot span single Nexus 1000V
- Two caveats
 - No veth VXLAN trunk mode support with MAC distribution
 - Won't work with Microsoft NLB
- If above issues are not a problem use MAC Distribution

VXLAN Requirements

- VMkernel interface to act as VTEP
- VSM Control Mode should be L3
- Multicast or Unicast for Broadcast traffic?
- 1550 MTU for VXLAN encapsulation overhead

VXLAN Configuration: Multicast

- VMkernel interface to act as VTEP
- VSM Control Mode should be L3
- Multicast for Broadcast traffic
- IP Multicast forwarding is required
 - Multicast addresses
 - Multiple segments can be mapped to a single multicast group
 - If VXLAN transport is contained to a single VLAN, IGMP Querier must be enabled on that VLAN
 - If VXLAN transport is traversing routers
 - Multicast routing must be enabled.
 - Proxy ARP must also be enabled
- 1550 MTU for VXLAN encapsulation overhead

VXLAN Configuration: Multicast

- Upstream Switch Configuration
 - Enable IGMP Querier
 - Set physical switch port MTU to 1550
- ESXi Host
 - Create VMK interface for VXLAN
- Nexus 1000V
 - Enable “feature segmentation”
 - Create a Bridge Domain
 - Create a port-profile for VTEP VMK interface
 - Create a veth port-profile for the VMs

VXLAN Configuration: Multicast

- Increase the MTU on your eth port-profile

```
n1kv-13(config)# port-profile type eth uplink
n1kv-13(config-port-prof)# mtu 1550
```

- Create veth port-profile for VXLAN VMK interface

```
n1kv-13(config)# port-profile type vethernet VXLAN-VMK
n1kv-13(config-port-prof)# switchport mode access
n1kv-13(config-port-prof)# switchport access vlan 11
n1kv-13(config-port-prof)# no shutdown
n1kv-13(config-port-prof)# system vlan 11
n1kv-13(config-port-prof)# vmware port-group
n1kv-13(config-port-prof)# capability vxlan
n1kv-13(config-port-prof)# state enabled
```

VXLAN Configuration: Multicast

- Configure the Bridge Domain

- Maps a segment ID to a multicast address
- Segment ID >4096

```
n1kv-l3(config)# bridge-domain vxlan-1
n1kv-l3(config-bd)# segment id 5000
n1kv-l3(config-bd)# group 224.3.5.2
```

- Create VM port-profile

```
n1kv-l3(config)# port-profile type veth vm-vxlan-1
n1kv-l3(config-port-prof)# vmware port-group
n1kv-l3(config-port-prof)# switchport mode access
n1kv-l3(config-port-prof)# switchport access bridge-domain vxlan-1
n1kv-l3(config-port-prof)# no shut
n1kv-l3(config-port-prof)# state enabled
```

VXLAN Troubleshooting Tips

- Verify your Bridge Domains, VM port-profiles, and VXLAN VMK port-profiles
- Verify multicast on your upstream switches
 - show ip igmp snooping
 - Do you see the VTEPs
- Use vmkping on the ESXi host to verify network and MTU
 - Use 1542 to cover the addition of the ICMP header
- Verify the VEM has the right VXLAN capability

```
~ # vmkping -s 1542 -d 1.1.1.1
```

```
~ # vemcmd show vxlan interfaces
```

```
LTL      IP
```

```
-----
```

```
69 1.1.1.2
```


VXLAN Troubleshooting Tips

- Verify your VM is on the right segment id

```
~ # vemcmd show port vlans
```

LTL	VSM Port	Mode	Native VLAN/ SegID	VLAN State	Allowed Vlans/SegID
17	Eth4/1	T	1	FWD	25,626-640
18	Eth4/2	T	1	FWD	25,626-640
49	Veth2	A	25	FWD	25
52	Veth14	A	25	FWD	25
53	Veth19	A	6000	FWD	6000

- Verify the VEM was programmed correctly

```
~ # vemcmd show segment 6000
```

```
BD 23, vdc 1, segment id 6000, segment group IP 225.6.26.10, swbd 4096, 2 ports, "dvs.VCDVsvCDNI-6-26-vl634-backed-b69c1d1d-02bf-4581-9b7e-fa06c64e8c18"
```

```
Portlist:
```

```
53 vse-vCDNI-6-26-vl634-backed (b6  
68 vCDNI-2 (5ac7d73c-d1d1-4877-8ef
```

VXLAN Other Useful Commands

- `vemcmd show port`
- `vemcmd show igmp <vlan>`
- `vemcmd show l2 segment <segment-id>`
- `vemcmd show vxlan-encap [l2l/mac] <l2l/MAC address>`
- `vemcmd show vxlan-stats all`
- Detailed slides in the Appendix



Appendix

VXLAN Additional Troubleshooting Slides

VXLAN Other Useful Commands

- Verify Multicast Upstream Nexus 7K/5K
 - Verify querier is enabled for vlan VMK interfaces are on

```
switch# show run
vlan configuration 634
ip igmp snooping querier 1.1.1.161
```

VXLAN Other Useful Commands

- Verify IGMP snooping is configured

```
CWD 35 04-7000-1# show ip igmp snooping vlan 634
```

```
IGMP Snooping information for vlan 634
```

```
IGMP snooping enabled
```

```
Optimised Multicast Flood (OMF) enabled
```

```
IGMP querier present, address: 1.1.1.161, version: 3, i/f Po1
```

```
Querier interval: 125 secs
```

```
Querier last member query interval: 1 secs
```

```
Querier robustness: 2
```

```
Switch-querier enabled, address 1.1.1.161, currently running
```

```
....
```

```
IGMPv3 Report suppression disabled
```

```
Link Local Groups suppression disabled
```

```
Router port detection using PIM Hellos, IGMP Queries
```

```
Number of router-ports: 1
```

```
Number of groups: 1
```

```
VLAN vPC function enabled
```

```
Active ports:
```

```
  Po1 Po9      Po17  Po25
```

```
 Po31      Po52      Po100  Eth2/30
```

VXLAN Other Useful Commands

- Verify multicast IP address for the VXLAN is being learned

```
CWD.35.04-7000-1# show ip igmp snooping groups vlan 634
```

Type: S - Static, D - Dynamic, R - Router port, F - Fabricpath core port

Vlan	Group	Address	Ver	Type	Port list
634	*/*		-	R	Po1
634		225.6.26.10	v2	D	Po100 Eth2/30

VXLAN Other Useful Commands

- `vemcmd show port`
 - Will show ports that are on a vxlan

```
~ # vemcmd show port
```

LTL	VSM Port	Admin	Link	State	PC-LTL	SGID	Vem Port	Type
17	Eth4/1	UP	UP	F/B*	305	0	vmnic0	
18	Eth4/2	UP	UP	F/B*	305	1	vmnic1	
49	Veth2	UP	UP	FWD	0	0	vmk0	

```
...
```

53	Veth19	UP	UP	FWD	0		vse-vCDNI-6-26-v1634-backed	(b6
54	Veth16	UP	UP	FWD	0	0	vse-vCDNI-6-26-v1634-backed	(b6

```
...
```

68	Veth21	UP	UP	FWD	0		vCDNI-2 (5ac7d73c-d1d1-4877-8ef	
69	Veth22	UP	UP	FWD	0	0	vmk1	vxlan

305	Po2	UP	UP	F/B*	0			
-----	-----	----	----	------	---	--	--	--

VXLAN Other Useful Commands

- `vemcmd show igmp <vlan>`
 - Verify that multicast is enabled

```
# vemcmd show igmp 634
IGMP is ENABLED on VLAN 634
Multicast Group Table:
Group */*, Multicast LTL: 4410
```

- `vemcmd show l2 segment <segment-id>`
 - Verify the VEM is learning MAC addresses in the VXLAN

```
~ # vemcmd show l2 segment 6000
Bridge domain 23 brtmax 4096, brtcnt 3, timeout 300
Segment ID 6000, svbd 4096, "dvs.VCDVSVCDNI-6-26-vl634-backed-b69c1d1d-02bf-4581-9b7e-fa06c64e8c18"
Flags: P - PVLAN S - Secure D - Drop

  Type          MAC Address      LTL  timeout  Flags  PVLAN  Remote IP
-----
Static 00:50:56:01:02:0a 68      0      0.0.0.0
Dynamic 00:50:56:01:02:09 305     1      1.1.1.1
Static 00:50:56:01:02:15 53      0      0.0.0.0
```


VXLAN Other Useful Commands

- `vemcmd show vxlan-encap [ltl/mac] <ltl/MAC address>`

- Identify the traffic path a MAC or LTL will utilise

```
~ # vemcmd show vxlan-encap ltl 68
Encapsulation details for LTL 68 in BD "dvs.VCDVSvCDNI-6-26-vl634-backed-b69c1d1d-02bf-4581-9b7e-fa06c64e8c18":
Source MAC: 00:50:56:01:02:0a
```

```
Segment ID: 6000
Multicast Group IP: 225.6.26.10
```

```
Encapsulating L2 LISP Interface LTL: 69
Encapsulating Source IP: 1.1.1.2
Encapsulating Source MAC: 00:50:56:7e:0e:b6
```

Pinning of L2 LISP Interface to the Uplink:

LTL	IfIndex	PC LTL	VSM SGID	Eff SGID	iSCSI LTL*	Name
69	1c000150	305	32	0	0	vmk1

VXLAN Other Useful Commands

- `vemcmd show vxlan-stats all`

- Show VXLAN traffic stats

~ # `vemcmd show vxlan-stats all`

LTL	Ucast Encaps	Mcast Encaps	Ucast Decaps	Mcast Decaps	Total Drops
53	67	300	47	0	23
68	11701	47135	11690	61	12
69	11768	125793	11737	61	0

VXLAN Load Balancing

- With LACP port-channel 5-tuple hash is used
 - Use single VMK VXLAN interface
 - VEM does the hashing across all the links
 - Remember to change load balancing to 5-tuple hashing
 - On the upstream switch and on the VSM
- With VPC MAC Pinning
 - Create a VMK VXLAN interface for each available uplink
 - VEM will pin an interface to each available link
 - The VEM will distribute the VM's flows between the vmknics based on a hash of the source MAC.

Verification: Unicast

■ Verify the bridge-domain configuration on VSM

```
switch# sho bridge-domain
```

```
Global Configuration:  
Mode: Unicast-only  
MAC Distribution: Disable
```

```
Bridge-domain segment-cisco (3 ports in all)  
Segment ID: 9001 (Manual/Active)  
Mode: Unicast-only (default)  
MAC Distribution: Disable (default)  
Group IP: NULL  
State: UP           Mac learning: Enabled  
Veth2, Veth3, Veth5
```

If MAC Distribution is enabled this will be 'Enable'

If MAC Distribution is enabled this will be
"Segment Mode: Unicast, Mac-Distribution"

■ Verify the bridge-domain configuration on VEM

```
switch# module vem 4 execute vemcmd show bd bd-name segment-cisco  
BD 26, vdc 1, segment id 9001, segment group IP 0.0.0.0, swbd 4102, 2 ports, "segment-cisco"  
Segment Mode: Unicast  
VTEP DSN: 1 , MAC DSN: 1  
Portlist:  
 53 RedHat_VM1_112.eth4  
 54 RedHat_VM1_112.eth5  
~ #
```

VTEP and MAC download sequence numbers should be checked against VTEP entries (vemcmd show vxlan-vteps) and MAC entries (vemcmd show I2 bd bd-name <>) respectively

Verification (Port Configuration)

- Verify the port configuration on VSM

```
switch# sho int switchport | begin Vethernet2
Name: Vethernet2
  Switchport: Enabled
  Switchport Monitor: Not enabled
  Operational Mode: access
  Access Mode VLAN: 0 (none)
  Access BD name: segment-cisco
[SNIP]
```

Verification (Port Configuration)

- Verify the port configuration on VEM

```
switch# module vem 4 execute vemcmd show port
  LTL   VSM Port  Admin Link  State  PC-LTL  SGID  Vem Port  Type
   17   Eth4/1    UP   UP   F/B*   561    0    vmnic0
   49                   DOWN  UP   BLK    0                RedHat_VM1_112 ethernet7
   50   Veth8     DOWN  UP   BLK    0                RedHat_VM1_112.eth8
   51   Veth4     UP    UP   FWD    0    0    vmk1    VXLAN
   52                   DOWN  UP   BLK    0                RedHat_VM1_112.eth6
   53   Veth2     UP    UP   FWD    0                RedHat_VM1_112.eth4
   54   Veth3     UP    UP   FWD    0                RedHat_VM1_112.eth5
  561   Po2       UP    UP   F/B*   0
```

* F/B: Port is BLOCKED on some of the vlans.
One or more vlans are either not created or
not in the list of allowed vlans for this port.
Please run "vemcmd show port vlans" to see the details.
~ #

Verification (VTEP Distribution)

- Verify the VTEP distribution on VSM

```
switch# sho bridge-domain segment-cisco vteps
```

```
D: Designated VTEP      I:Forwarding Publish Incapable VTEP
```

```
Bridge-domain: segment-cisco
```

```
VTEP Table Version: 2
```

```
Ifindex      Module  VTEP-IP Address
```

```
-----  
-----  
Veth4        4      10.106.199.116 (D)
```

To be compared with
echo "show vxlan
version-table" output on
VEM

Compare against "vemcmd show bd bd-name <>"
VTEP DSN output

```
switch# module vem 4 execute vemcmd show vxlan-vteps
```

```
Bridge-Domain: segment-cisco Segment ID: 9001
```

```
Designated Remote VTEP IPs (*=forwarding publish incapable):
```

```
10.106.199.117 (DSN: 1),
```

Verification (MAC Table In Unicast Only Mode)

- MAC address table will display remote IP learning in the segment-cisco bridge domain

```
switch# module vem 4 execute vemcmd show l2 bd-name segment-cisco
Bridge domain 26 brtmax 4096, brtcnt 3, timeout 300
Segment ID 9001, swbd 4102, "segment-cisco"
Flags: P - PVLAN S - Secure D - Drop
  Type          MAC Address    LTL   timeout   Flags    PVLAN    Remote IP    DSN
  Dynamic       00:50:56:83:01:4e 561    1         10.106.199.117 0
  Static        00:50:56:83:01:61 54     0         0.0.0.0        0
  Static        00:50:56:83:01:60 53     0         0.0.0.0        0

switch#
```




Appendix: Upgrade Info

Upgrades – List of Changes Allowed

- Add or remove modules.
- Add or remove ports (ETH and VETH).
- Shut or no-shut a port.
- Migrate ports to or from a vswitch.
- Change port modes (trunk or access) on ports.
- Add or remove port profiles.
- Modify port profiles to add or remove specific features such as VLANs, ACLs, QoS, or PortSec.
- Change port channel modes in uplink port profiles.
- Add or delete VLANs and VLAN ranges.
- Add or delete static MACs in VEMs.

Note: Queuing configuration changes not supported on QoS.



Appendix: Nexus 1010 and 1110

Using Export To Backup A VSM

- Shutdown primary VSM
 - Secondary VSM will take over and run Nexus 1000v control plane
 - f340-33-09-n1010-1(config)# virtual-service-blade training
 - f340-33-09-n1010-1(config-vs-b-config)# shutdown primary
- Export VSB on 1x10
 - f340-33-09-n1010-1(config-vs-b-config)# export primary
 - Note: export started..
 - Note: please be patient..
 - Note: export completed...
- Verify
 - f340-33-09-n1010-1(config-vs-b-config)# dir bootflash:///export-import/4
 - 147779575 Oct 18 02:47:10 2011 Vdisk4.img.tar.00

Using Import To Restore A VSB

- Copy the backup to bootflash

- f340-33-09-n1010-1# copy
scp://root@172.18.217.165/root/Vdisk4.img.tar.00
bootflash:export-import vrf management

- Import the image

- f340-33-09-n1010-1(config)# virtual-service-blade training
 - f340-33-09-n1010-1(config-vsbc-config)# import primary
Vdisk4.img.tar.00
 - Note: import started..
 - Note: please be patient..
 - Note: Import cli returns check VSB status for completion

- Verify

- f340-33-09-n1010-1(config-vsbc-config)# show virtual-service-blade name training

