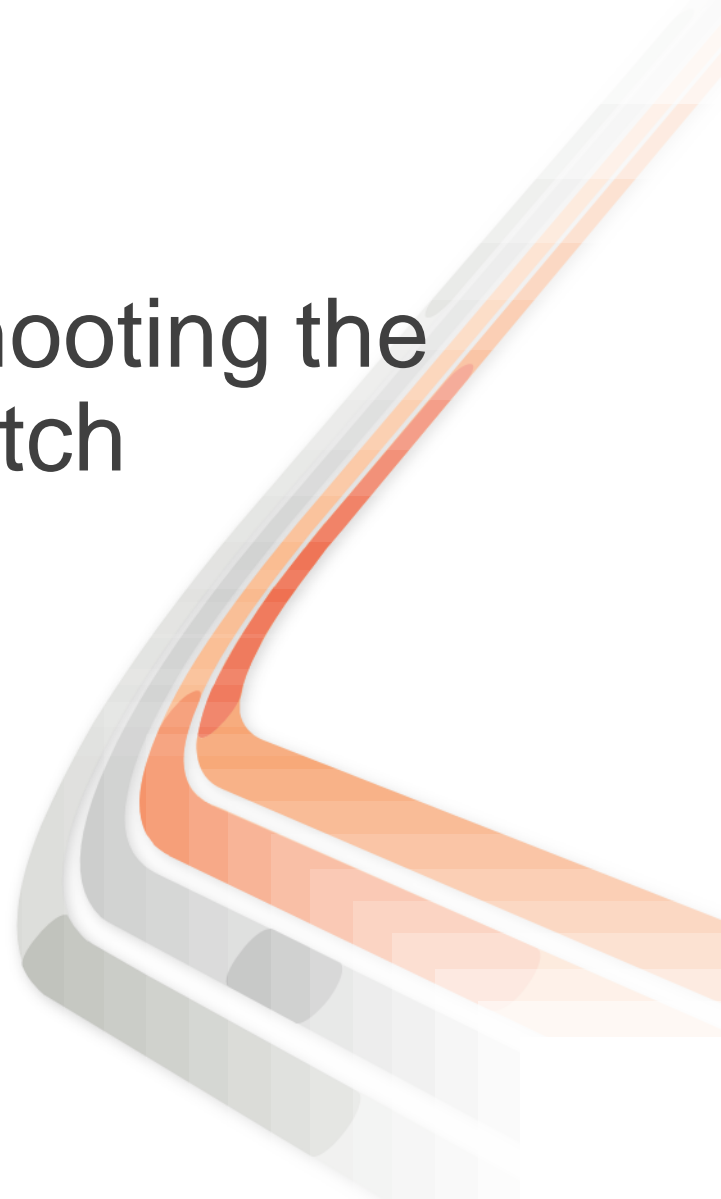


# Deploying and Troubleshooting the Nexus 1000V Virtual Switch

BRKVIR-3013



# Agenda

- Session Prerequisites
- Introduction Cisco Nexus 1000V
- Useful tool locations
- Nexus 1000V Releases
- Virtual Supervisor Module (VSM) troubleshooting
- Virtual Ethernet Module (VEM) troubleshooting
- Port-Profiles
- Port Channels
- HA
- Cisco Nexus 1010

# Session Prerequisites

A decorative graphic element at the bottom of the slide, consisting of a horizontal orange line that curves downwards and to the right, forming a ramp-like structure with a white and grey checkered pattern on its side.

# Related Presentations

- BRKVIR-1012      The Business Case for Cisco Virtualization Experience
- BRKVIR-2011      Deploying services in a virtualized environment
- BRKVIR-2012      Inside the Nexus 1000V Virtual Switch

# Prerequisites

- Understanding of VMware ESX and vCenter Server
  - Difference between ESX and ESXi
- Unix CLI
  - Running commands on ESX Service Console
- Cisco NXOS
  - Understand the CLI general switching concepts
- Cisco Nexus 1000V concepts
  - Understand VSM and VEM
  - Port-profile concepts

# Cisco Nexus 1000V Introduction

A decorative graphic element at the bottom of the slide, consisting of a thick orange line that curves from the left towards the right, then splits into a series of parallel lines in shades of orange and white, creating a sense of depth and movement.

# Cisco Nexus 1000V

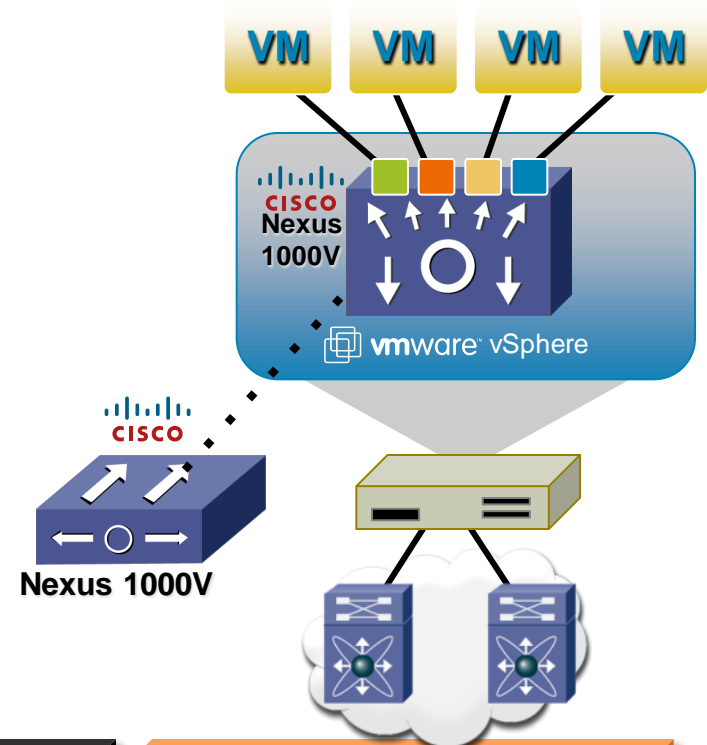
- Industry's most advanced software switch for VMware vSphere
- Built on Cisco NX-OS
- Compatible with all switching platforms
- Maintain vCenter provisioning model unmodified for server administration; allow network administration of virtual network via familiar Cisco NX-OS CLI

**Policy-Based  
VM Connectivity**

**Mobility of Network &  
Security Properties**

**Non-Disruptive  
Operational Model**

**BEST OF  
vmworld® 2008**



# Cisco Nexus 1000V

## Faster VM Deployment

### Cisco VN-Link: Virtual Network Link

Policy-Based  
VM Connectivity

Mobility of Network &  
Security Properties

Non-Disruptive  
Operational Model

#### Port Profiles

WEB Apps



HR



DB

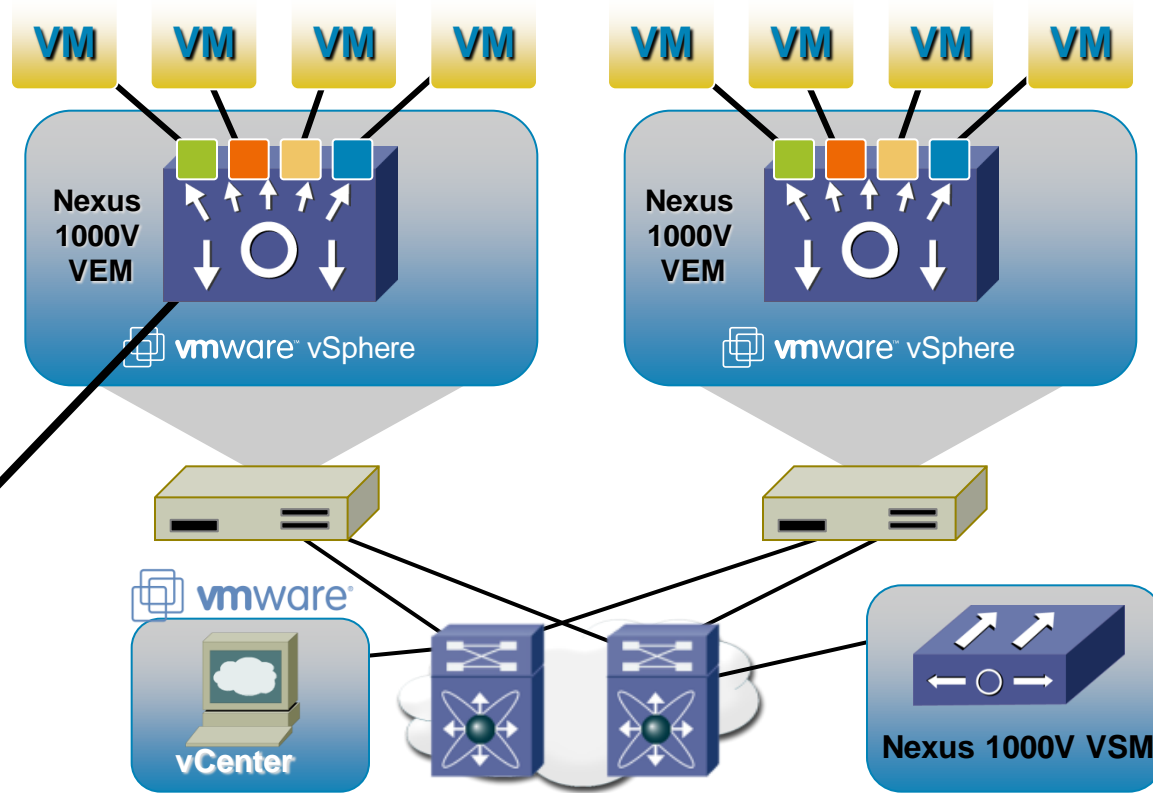


DMZ



#### VM Connection Policy

- Defined in the network
- Applied in Virtual Center
- Linked to VM UUID





# Cisco Nexus 1000V

## Richer Network Services

### Cisco VN-Link: Virtual Network Link

Policy-Based  
VM Connectivity

Mobility of Network &  
Security Properties

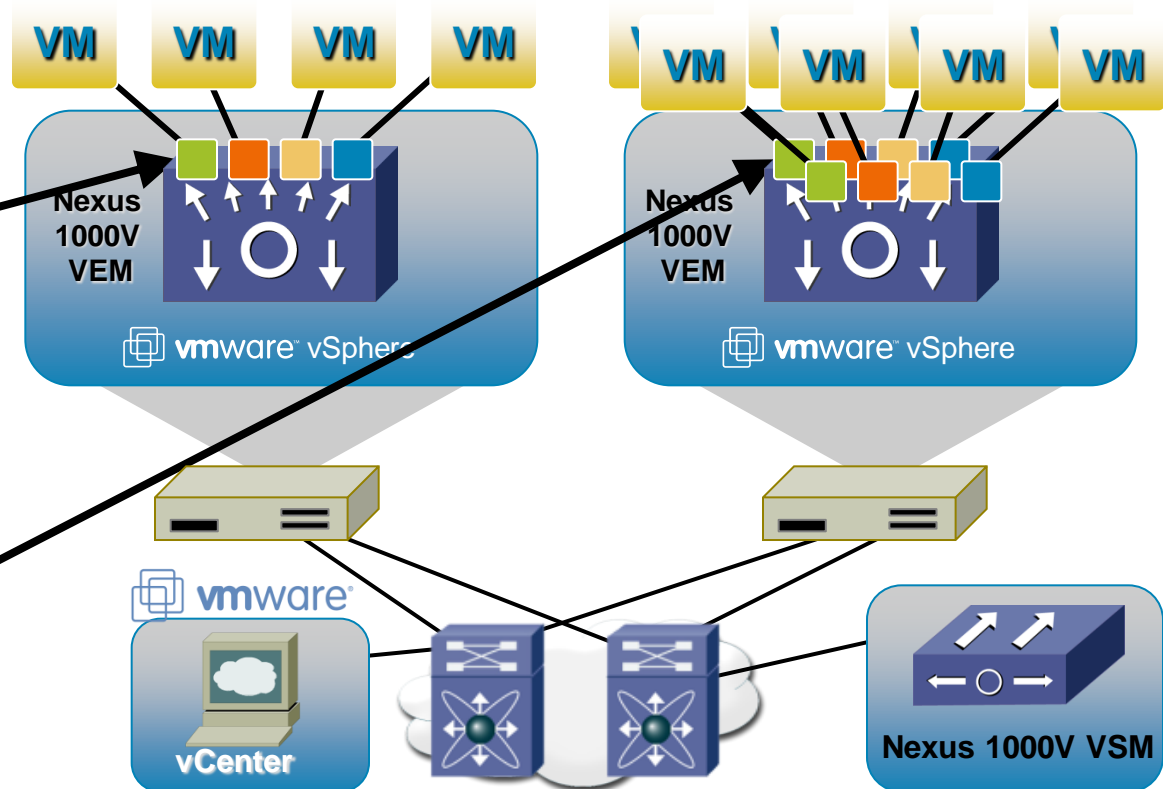
Non-Disruptive  
Operational Model

#### VMs Need to Move

- VMotion
- DRS
- SW Upgrade/Patch
- Hardware Failure

#### Property Mobility

- VMotion for the network
- Ensures VM security
- Maintains connection state



# Cisco Nexus 1000V

## Increased Operational Efficiency

### Cisco VN-Link: Virtual Network Link

Policy-Based  
VM Connectivity

Mobility of Network &  
Security Properties

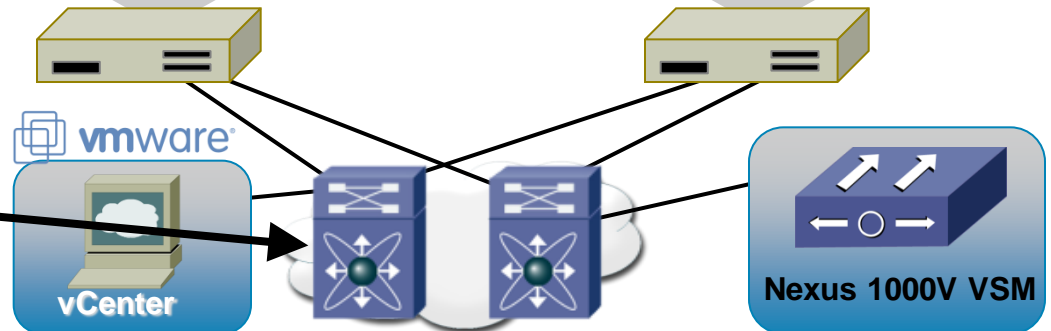
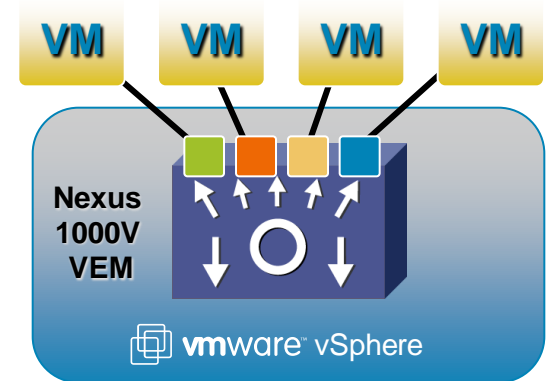
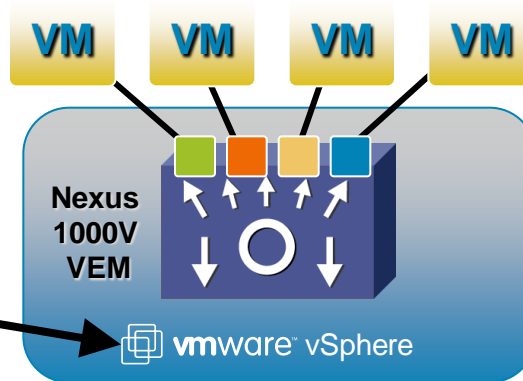
Non-Disruptive  
Operational Model

#### VI Admin Benefits

- Maintains existing VM mgmt
- Reduces deployment time
- Improves scalability
- Reduces operational workload
- Enables VM-level visibility

#### Network Admin Benefits

- Unifies network mgmt and ops
- Improves operational security
- Enhances VM network features
- Ensures policy persistence
- Enables VM-level visibility



# Tool Locations



# Locations of ESX Tools

- Cisco VEM commands on ESX

`/usr/lib/ext/cisco/nexus/vem*/sbin`

Linked in `/usr/sbin`

- VEM commands can also be run remote from VSM

```
n1000v# module vem 3 execute vemcmd show port
```

LTL	IfIndex	Vlan	Bndl	SG_ID	Pinned_SGID	Type	Admin	State	CBL	Mode	Name
8	0	3969	0	2	2	VIRT	UP	UP	4	Access	120

- For ESXi use the remote commands from the VSM

`vemcmd` through RCLI does not work for ESXi

Enable “unsupported mode”

Google unsupported ESXi

# Useful Commands

- Vemcmd on ESX host
  - Can query and set configuration
  - Improved in 1.4 to allow more “set” commands and many more show commands
- Vem-health on ESX host
  - Will try to point you in right direction if the vem is having issues
- Mping on VSM
  - Command that will broadcast out on the control network looking for VEM modules

Learn. Connect.  
Collaborate. *together.*

# Nexus 1000V Releases

A decorative graphic element at the bottom of the slide, consisting of a horizontal orange line that curves downwards and to the right, forming a ramp-like shape. The ramp is composed of several parallel lines in shades of orange and white, creating a sense of depth and movement.

# Cisco Nexus 1000V Releases


- First Release version 4.0(4)SV1(1) – 1.1
- Second Release version 4.0(4)SV1(2) - 1.2
- Third Release version 4.0(4)SV1(3) - 1.3
  - Three maintenance releases 1.3a, 1.3b, and 1.3c
- Current Release version 4.2(1)SV1(4)
- All releases work with ESX/ESXi 4.x
- All releases require VMware Enterprise plus license
- Expect VSM/VEM backwards compatibility to work one version back.

# 1.4 New Features

- NXOS 4.2 changes
  - “Feature” command for LACP, PVLAN, Netflow, and port-profile-roles
    - Must enable the feature for above commands to become active
- Change to upgrade procedure
  - Now VEMs first then VSMs
- Network State Tracking (Beaconing)
- LACP Offload
- QOS Weighted Fair Queuing
- Port-profile-roles
  - specify which port-profiles a user has access to
- vPath for VSG and vWAAS



# Virtual Supervisor Module Troubleshooting

A decorative graphic element at the bottom of the slide, consisting of a thick orange line that curves to the right and then continues as a series of parallel lines in orange and white, creating a sense of depth and movement.

# Virtual Supervisor Module (VSM)

- VSM is a Virtual Machine
  - On ESX
  - On Nexus 1010
- Control plane for the Nexus 1000V solution
- Responsible for
  - VMware vCenter communication
  - Programming and managing Virtual Ethernet Modules (VEM)
- 1 VSM HA pair can manage 64 VEMs
- Nexus 1000V can coexist with VMware vSwitch and DVS

# VSM Virtual Machine Requirements

- 3 network interfaces

Adapter 1 is always the **Control** interface – Must be on the control VLAN

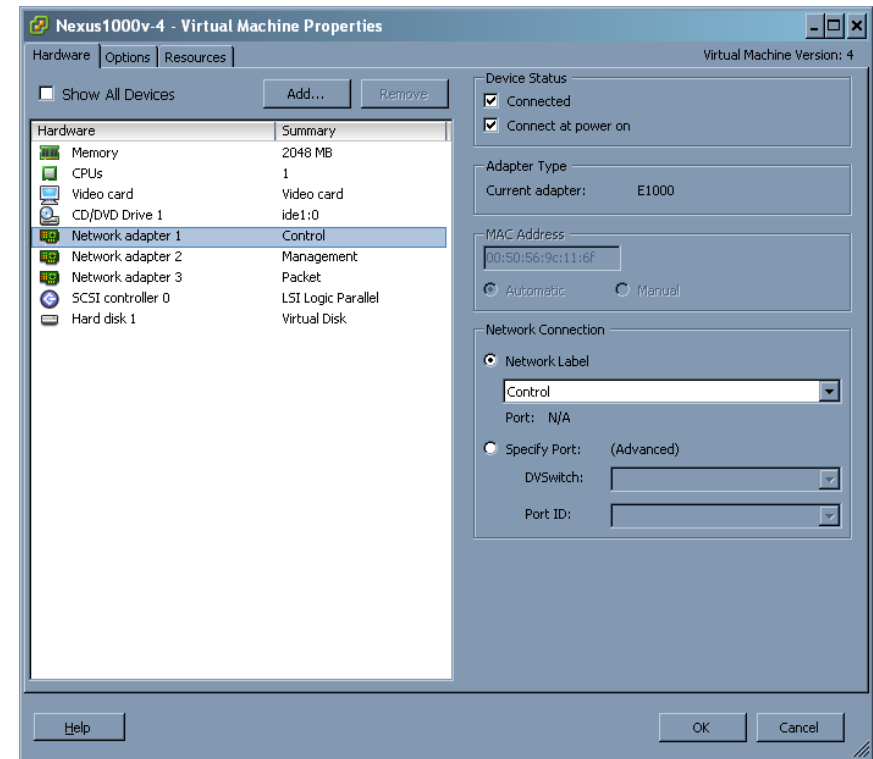
Adapter 2 is always the **Management** interface – Used for VSM connectivity

Adapter 3 is always the **Packet** interface – Must be on the packet VLAN

- Need 2GB of memory RAM reserved

VSM will do odd things when it is memory starved

- One vCPU



# VSM Required Interfaces

- **Management**

- VSM terminal connectivity
  - Connecting to VMware vCenter
  - Backup connectivity for VSM HA

- **Control**

- Heartbeat between VSM and VEM
  - Heartbeat and information passing between Active and Standby VSMs

- **Packet**

- Passes CDP and IGMP information

- **VLANs can be shared with multiple VSM installations**

- Several VSMs using same Control VLAN
  - SVS domain ID keeps them separate

# VSM Control Modes

- L2 mode

  - Default mode

  - Requires L2 connectivity through Control interface to all VEM modules

  - Uses Control Interface of VSM VM

- L3 Mode

  - Uses IP address to communicate with VEM

  - L3 uses UDP port 4785 for both source and destination

  - Uses Mgmt or Control interface of VSM VM

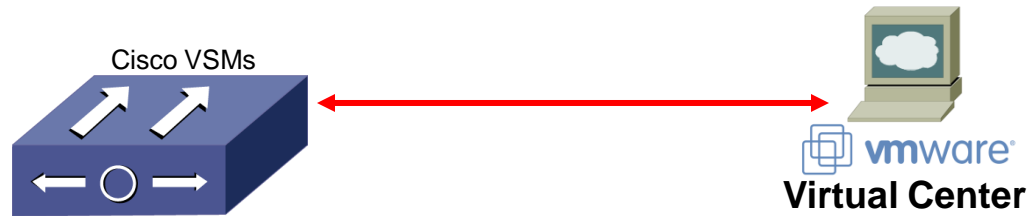
  - VSM mgmt 0 is default interface for L3

  - Can also use “control 0”

    - Ties to control adapter of the VM (Adapter 1)

    - Control 0 and mgmt 0 must be on different subnets

# VSM to VMware vCenter Communication



- VSM connects to vCenter using SSL connection.
- Management Interface
- Self-Signed certificate used for this connection
- VSM configures vCenter using its API
  - Create N1KV dv-Port-Groups in vCenter
  - Stores DVS data to be passed to ESX hosts which become member of N1KV
  - Get useful information from vCenter (DC, DVS, VM, ...)

# VSM to VMware vCenter Connectivity

- VSM registers to vCenter using extension key plug-in
  - Contains a public SSL Certificate
  - Extension key of the VSM
- Only need to register a specific VSM extension key once
- Can register multiple keys from different VSMS on same vCenter
- A VSM is tied to a VMware Datacenter
- Extension key only changes
  - Brand new VSM install
  - Change the certificate on the VSM. "install certificate" command in svcs connection mode
  - Change the extension key using "vmware vc extension-key"

# Verify VSM to vCenter Connectivity

- Verify SVS connection settings

```
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
```

- If Datacenter is underneath a folder and spaces

```
n1000v(config-svs-conn)# vmware dvs datacenter-name ?
```

```
LINE Datacenter name in VC with path (e.g. DCName,  
DCFolder/DC Name)
```



# Connectivity Error – Extension Key

- Below error means wrong key or key is not registered

```
n1000v(config-svs-conn)# connect  
ERROR: [VMware vCenter Server 4.0.0. build-162856]  
Extension key was not registered before its use
```

- Check the key on VSM

```
n1000v# show vmware vc extension-key  
Extension ID: Cisco_Nexus_1000V_165241074
```

# Check Extension Key on VMware MOB

- Check MOB to see what is registered
- <http://VMware-vCenter-IP/mob>  
Content->ExtensionManager

Managed Object Browser - Mozilla Firefox

File Edit View History Bookmarks Tools Help

172.18.217.241 https://172.18.217.241/mob/?moid=Extens

Managed Object Browser

Home

Managed Object Type: **ManagedObjectReference:ExtensionManager**  
Managed Object ID: ExtensionManager

Properties

NAME	TYPE	VALUE
extensionList	Extension[]	<ul style="list-style-type: none"><li>extensionList["Cisco_Nexus_1000V_1776281504"]</li><li>extensionList["com.vmware.vcintegrity"]</li><li>extensionList["cim-ui"]</li><li>extensionList["com.vmware.vim.sms"]</li><li>extensionList["com.vmware.vim.stats.report"]</li><li>extensionList["health-ui"]</li><li>extensionList["hostdiag"]</li><li>extensionList["integrity"]</li><li>extensionList["Cisco_Nexus_1000V_515382320"]</li><li>extensionList["Cisco_Nexus_1000V_165241074"]</li></ul>

Methods

RETURN TYPE	NAME
Extension	<a href="#">FindExtension</a>
string	<a href="#">GetPublicKey</a>
void	<a href="#">RegisterExtension</a>
void	<a href="#">SetExtensionCertificate</a>
void	<a href="#">SetPublicKey</a>
void	<a href="#">UnregisterExtension</a>
void	<a href="#">UpdateExtension</a>

Done

# Connectivity Error – Connection Refused

- Below error could indicate port mismatch

```
h1000v(config-svs-conn)# connect  
ERROR: [VMWARE-VIM] Operation could not be completed due to  
connection failure.Connection refused. connect failed in  
tcp_connect()
```

- Default port for communication is port 80
- All communication is https
- VMware accepts on port 80 and tunnels internally to port 8089

# VSM to VMware vCenter Connectivity

- Make sure VSM SVS port matches vCenter http port

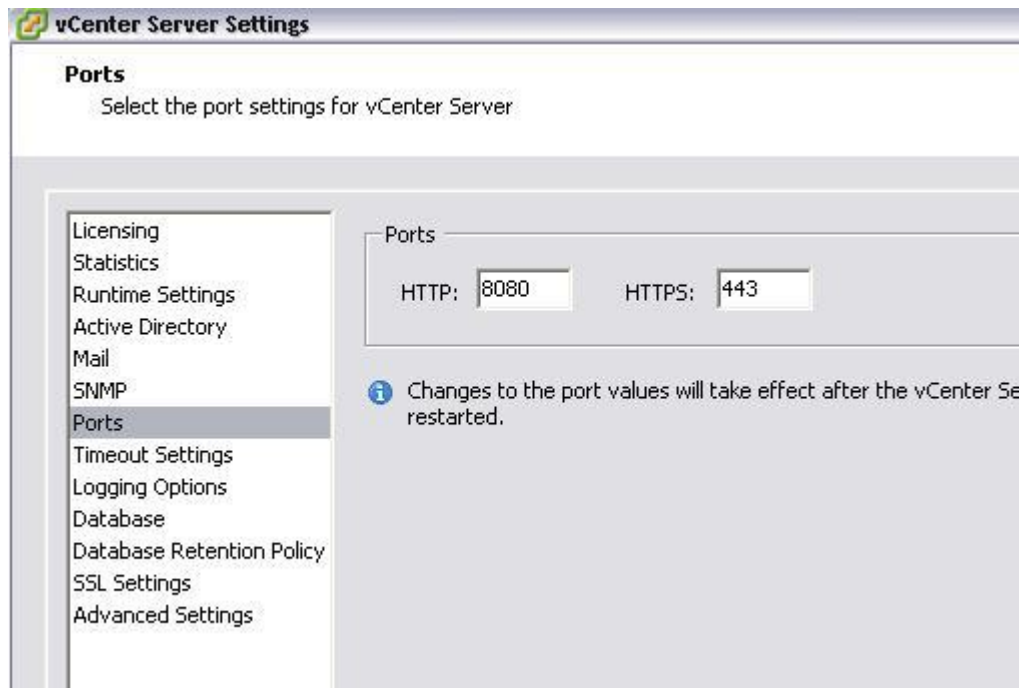
```
n1000v# show svcs connections
connection vcenter:
  ip address: 172.18.217.41
  remote port: 80
  protocol: vmware-vim https
```

- To change port

```
n1000v(config)# svcs connection vcenter
n1000v(config-svs-conn)# remote port 8080
```

# VSM to VMware vCenter Connectivity

- Verify Port number in vCenter
- Administration->vCenter Server Settings



# VSM and vMotion

- Manual vMotion of VSM is supported
- Not recommended to allow DRS to vMotion Primary and Secondary VSM
- Aggressive DRS vMotion setting can cause VSM to drop packets and loose connectivity to VEM
- Best practice to keep Primary and Secondary VSM outside of DRS

# Removing Nexus 1000V from vCenter

- Removal must be done from VSM
- Cannot remove VSM from within vCenter
- Deleting VSM without proper removal will leave orphaned Nexus 1000V in vCenter
- Delete Nexus 1000V from vCenter with following steps
  - config t
  - svs connection ConnectionName
  - no vmware dvs
- Use the “Recreating the Cisco Nexus 1000V Installation” from Troubleshooting guide to clean up failed removal

# VSM Best Practices

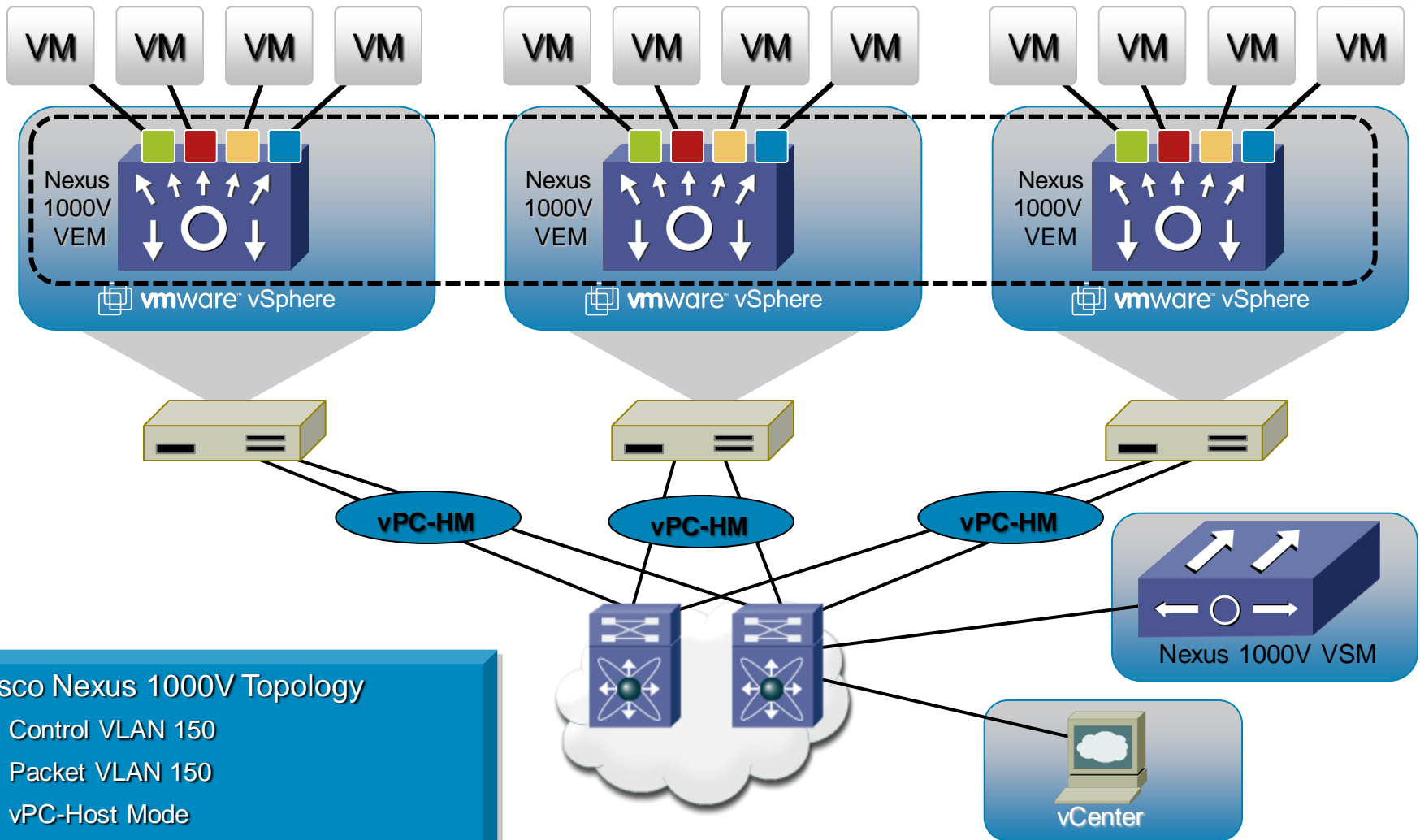
- L2 control is preferred
- Management, Control, and Packet can use same VLAN
- Do not use VLAN 1 for Control and Packet
- Primary and Standby VSM in same L2 domain!!!
- VSM on VEM is supported
- VSM primary to secondary latency  $\leq 100\text{ms}$
- VSM to VEM latency between 50ms and 100ms
- Backup your config!!!
- VMware snapshots and cloning are not supported



# Virtual Ethernet Module Troubleshooting

A decorative graphic element at the bottom of the slide, consisting of a thick orange line that curves to the right and then down, with a white line following its path. The background behind the orange line is a gradient of orange and white squares, creating a 3D effect.

# Topology for VEM Troubleshooting



## Cisco Nexus 1000V Topology

- Control VLAN 150
- Packet VLAN 150
- vPC-Host Mode

# Virtual Ethernet Module (VEM) Troubleshooting Agenda

- VEM installation
- VEM patches
- VEM removal
- VEM does not show up on VSM
- VEM useful commands

# VEM Installation

- Automatically with VMware Update Manager (VUM)

VUM does all the work user just adds the host to the N1KV

If you are having problems check the VUM logs

On vCenter Server in

`C:/Documents and Settings/Application Data/All Users/VMware/VUM/logs`

On ESX host in

`/var/log/vmware/esxupdate.log`

- Manually with `esxupdate` or `vihostupdate`

# VEM Installation Manual

## ■ ESX - Esxupdate

- Download vib file to ESX server via scp/sftp/ftp

```
Esx-1# esxupdate -b ./ cross_cisco-vem-v100-4.0.4.1.1.27-0.4.2-  
release.vib update
```

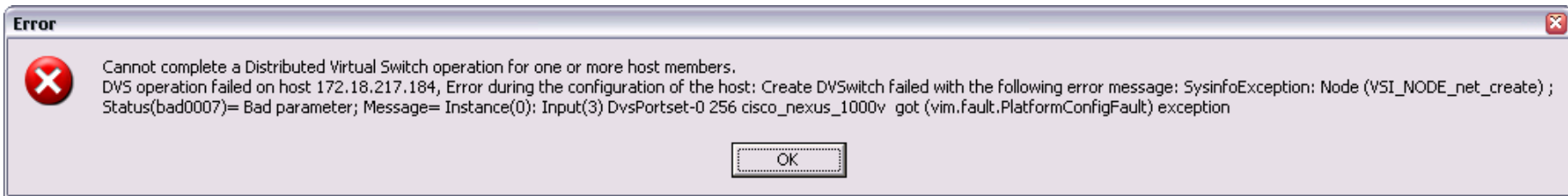
## ■ ESXi - Vihostupdate via vSphere Management Assistant VM (vMA) or RCLI

- RCLI – Remote command line package (Windows and Linux)
- vMA – Virtual appliance with RCLI installed
- [http://downloads.vmware.com/d/info/datacenter\\_downloads/vmware\\_vsphere\\_4/4#drivers\\_tools](http://downloads.vmware.com/d/info/datacenter_downloads/vmware_vsphere_4/4#drivers_tools)

```
Linux-1# vihostupdate -i -b ./cisco-vem-v100-4.0.4.1.1.nn-0.4.nn.zip  
--server hostname/ip address
```

# VEM Installation DVS Error

- DVS operation failed error
  - VUM is not installed or configured
  - VUM could not find the right VEM version
  - Manual installation of VEM was not performed



# VEM Manual Installation Issues

- Dependency error

```
[root@cae-esx-180 ~]# esxupdate -b ./cross_cisco-vem-v100-4.0.4.1.1.27-0.4.2-release.vib u
pdate
cross_cisco-vem-v100-4.0.4.1.1.27.. ##### [100%]

Unpacking cross_cisco-vem-v100-es.. ##### [100%]

The following problems were encountered trying to resolve dependencies:
No VIB provides 'vmknexus1kvapi-0-4' (required by cross_cisco-
vem-v100-esx_4.0.4.1.1.27-0.4.2)
Requested VIB cross_cisco-vem-v100-esx_4.0.4.1.1.27-0.4.2 conflicts with the
host
```

- Verify using correct VEM VIB version for ESX Kernel
- Use the N1KV compatibility matrix to identify the correct VIB to load on the host.

# VEM Patches

- ESX 4.0U1 and lower  
Every ESX patch gets a new VEM
- ESX 4.0U2 and higher  
Single VEM version for all new updates and patches
- If you are using VUM add the following repository  
<https://hostupdate.vmware.com/software/VUM/PRODUCTION/cisco-main-index.xml>  
VMware KB article  
[http://kb.vmware.com/selfservice/microsites/search.do?language=en\\_US&cmd=displayKC&externalId=1013134](http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1013134)
- If you install VEM manually you can download new VEM  
From VMWare  
<https://www.vmware.com/mysupport/download/>  
From Cisco  
<http://tools.cisco.com/support/downloads/pub/Redirect.x?mdfid=282362725>



# Determining ESX Patch Version

- How to tell what ESX patch you are running

Check /proc

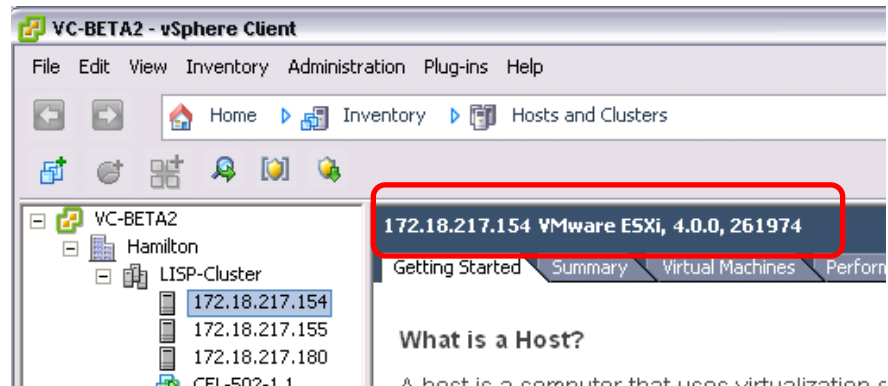
```
root@cae-cali-host6 vmware]# cat
/proc/vmware/version | grep vmkernel

vmkernel build: 193498, vmkcall: 45.0 driver
interface: 9.0 kernel: 0.4096

vmkernelID: 0xb1d912ac

vmkernel                Version Releasebuild-193498
```

From vCenter



# Matching a VEM Patch to a VIB Version

- Compatibility Matrix – from Cisco

[http://www.cisco.com/en/US/docs/switches/datacenter/nexus1000/sw/4\\_2\\_1\\_s\\_v\\_1\\_4/compatibility/information/n1000v\\_compatibility.html](http://www.cisco.com/en/US/docs/switches/datacenter/nexus1000/sw/4_2_1_s_v_1_4/compatibility/information/n1000v_compatibility.html)

VMware Software Version	Host Build <sup>1</sup>	VIB Version <sup>2</sup>	VEM Bundle <sup>3</sup>	Minimum Required Version	
				vCenter Server	Update Manager
ESX/ESXi 4.1.0 Update 1 ESX/ESXi410-201101201-SG	348481	cross_cisco-vem_v130-4.2.1.1.4.0.0-2.0.1.vib	VEM410-201101108-BG (Online) VEM410-201101407-BG (Offline)	345043	341095
ESX/ESXi 4.1.0 GA	260247	cross_cisco-vem_v130-4.2.1.1.4.0.0-2.0.1.vib	VEM410-201101108-BG (Online) VEM410-201101407-BG (Offline)	258902	256596
ESX/ESXi 4.0.0 P09 ESX/ESXi400-201103401	360236	cross_cisco-vem_v130-4.2.1.1.4.0.0-1.20.1.vib	VEM400-201101121-BG (Online) VEM400-201101406-BG (Offline)	258672	264019
				208111	282702
ESX/ESXi 4.0.0 Update 2 upgrade-from-esx/esxi4.0-4.0_update02	261974	cross_cisco-vem_v130-4.2.1.1.4.0.0-1.20.1.vib	VEM400-201101121-BG (Online) VEM400-201101406-BG (Offline)	258672 <sup>4</sup>	264019 <sup>4</sup>
				208111	282702
ESX/ESXi 4.0.0 P06 ESX/ESXi400-201005001	256968	cross_cisco-vem_v130-4.2.1.1.4.0.0-1.13.1.vib	VEM400-201101115-BG (Online) VEM400-201101405-BG (Offline)	208111	282702

# VEM VIB Versions

- 4.2.1.1.4.0.0-1.20.1
- 4.2.1.1.4 = N1KV version (show version)
  - 0.0 = Cisco NXOS build number
  - 1.20 = VEM VIB version
  - .1 = Cisco VIB build version
- VEM hardware version in “show module” correlates to VIB version

```
n1000v-AV# show module
Mod  Ports  Module-Type                Model                Status
---  -
1    0      Virtual Supervisor Module  Nexus1000V          active *
6    248    Virtual Ethernet Module    NA                  ok
7    248    Virtual Ethernet Module    NA                  ok
8    248    Virtual Ethernet Module    NA                  ok

Mod  Sw
---  -
1    4.2 (1) SV1 (4)          0.0
6    4.2 (1) SV1 (4)          VMware ESX 4.0.0 Releasebuild-332073 (1.20)
7    4.2 (1) SV1 (4)          VMware ESX 4.0.0 Releasebuild-332073 (1.20)
8    4.2 (1) SV1 (4)          VMware ESX 4.0.0 Releasebuild-261974 (1.20)
```

# VEM Removal

- VEM can only be removed manually
- Removing a host from the Nexus 1000V using vSphere does not remove the VEM
- ESX
  - `vem-remove -s -r`  
Will unload and remove the driver
- ESXi
  - `vihostupdate.pl --server 192.168.10.10 --username root -query`
  - `vihostupdate.pl --server 192.168.10.10 --username root -B VEM400-200906002-BG --remove`
  - reboot

# VEM Seeding

- How does VEM know VSM information?
- Opaque data – copied to VEM to seed during install.
- Opaque data consists of:
  - Domain-cfg (Domain ID, Control VLAN, Packet VLAN)
  - Switchname
  - VSM image version
  - System profiles [System VLANS, profile names]
  - IP address
  - MAC address

# Checking Opaque Data

- VSM stores opaque-data in vCenter as persistent data for its DVS.
- vCenter downloads this information to ESX for VEM to use, whenever a host is added to N1KV-DVS

## Checking opaque-data in VSM

```
switch-cp# show svcs domain
```

```
SVS domain config:  
Domain id: 100  
Control vlan: 150  
Packet vlan: 150  
Status: Config push
```

## Checking opaque-data in VEM

```
[root@sfish-30-119 sbin]#  
/usr/lib/ext/cisco/nexus/vem/sbin/vemcmd show  
card
```

```
Switch name: switch-cp  
Card domain: 100  
Card slot: 2  
Card control VLAN: 150  
Card packet VLAN: 150
```

## Checking opaque-data in vCenter

1. [https://vc\\_ip\\_address/mob/](https://vc_ip_address/mob/).
2. Content → rootFolder (group-dx) → childEntity (dataCenter-n) → networkFolder (group-n6) → childEntity (group-n) → childEntity (dvs-n) → config → VendorSpecificConfig

# VEM – VSM Connectivity Troubleshooting

- VEM adds in vCenter but does not show up on VSM “show module”
- With L2 most of the time its a Control VLAN issue  
Verify Control VLAN connectivity
- With L3 its usually an IP routing problem  
If you can ping VMK interface the VEM should connect to VSM  
Troubleshoot as you would all VMware L3 issues  
Is the VMK port-profile set with system VLAN?

# 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

```
n1000v# show svcs neighbors
```

```
Active Domain ID: 1254
```

```
AIPC Interface MAC: 0050-5681-0da4
```

```
Inband Interface MAC: 0050-5681-3595
```

## Step 2: VSM – vCenter Connectivity

- Verify VSM is connected to vCenter

```
h1000v# show svcs connections
```

```
connection MV-testing:
```

```
ip address: 172.18.217.241
```

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

```
certificate: default
```

```
datacenter name: Harrison
```

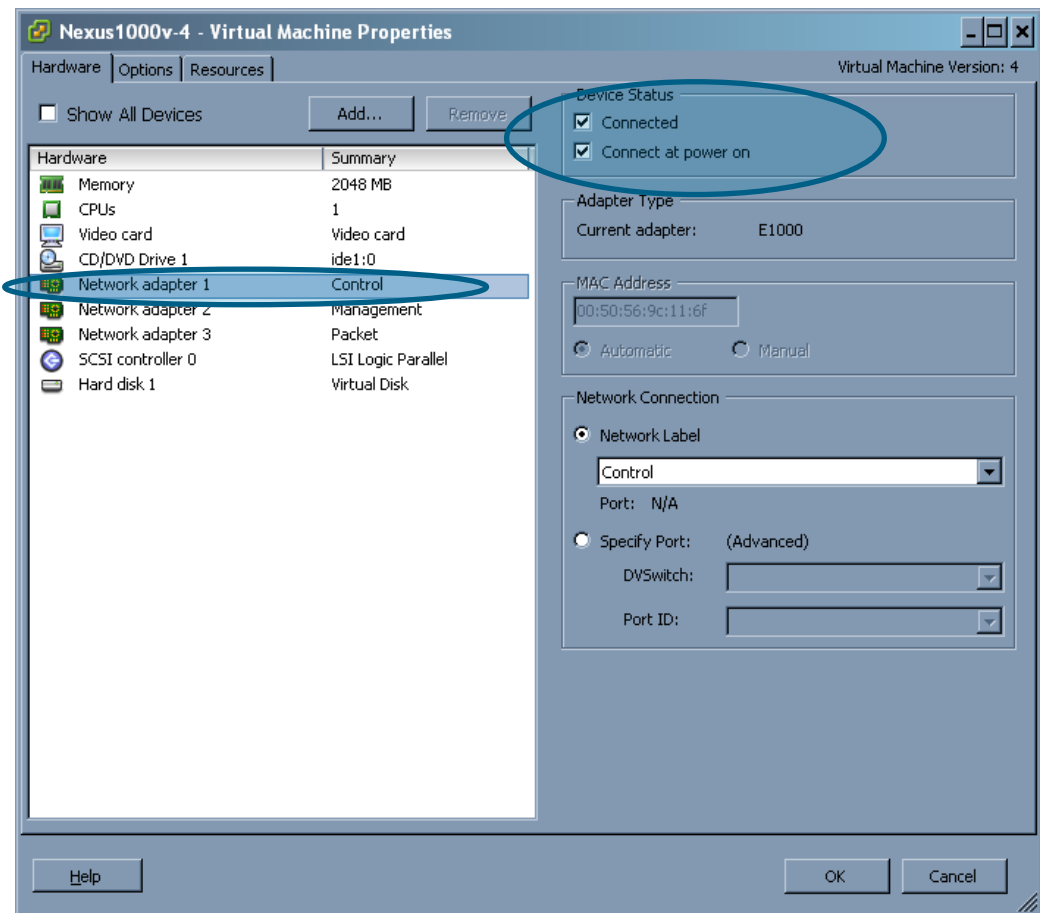
```
DVS uuid: 34 21 01 50 8e 07 cf 03-77 80 29 55 4f 77 4e 89
```

```
config status: Enabled
```

```
operational status: Connected
```

# Step 3: Verify VSM VM Control interface

- 1<sup>st</sup> interface listed is Control Interface
- Is 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

```
n1000v# show port-profile name uplink
port-profile uplink
  description:
  status: enabled
  capability uplink: yes
  capability l3control: no
  system vlans:2,10,150-152
  port-group: uplink
  max-ports: -
  inherit:
  config attributes:
    switchport mode trunk
    switchport trunk allowed vlan all
```

# 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

```
[root@cae-cali-host6 ~]# vemcmd show card
Card UUID type 2: 8c1d5178-8c02-11d9-0000-00000000000d
Card name: cae-cali-host6
Switch name: n1000v-MV
Switch alias: DvsPortset-0
Switch uuid: 72 f7 01 50 b2 01 7b 8b-55 68 cf df 10 5a db 55
Card domain: 234
Card slot: 4
VEM Control (AIPC) MAC: 00:02:3d:10:ea:03
VEM Packet (Inband) MAC: 00:02:3d:20:ea:03
VEM Control Agent (DPA) MAC: 00:02:3d:40:ea:03
VEM SPAN MAC: 00:02:3d:30:ea:03
Management IP address: 172.18.217.177
Max physical ports: 32
Max virtual ports: 216
Card control VLAN: 150
Card packet VLAN: 150
Processors: 16
```

MAC the VSM  
should learn for  
VEM

# Vemcmd show port-old

- Ports with LTLs 8(I20), 9(I21), 10(I22) are UP and CBL states are 4.
- ESX Physical ports are UP and CBL states 4.

```
[root@cae-cali-host6 ~]# vemcmd show port-old
```

LTL	IfIndex	Vlan	Bndl	SG ID	Pinned	SGID	Type	Admin	State	CBL	Mode	Name
8	0	3969	0	2		2	VIRT	UP	UP	4	Access	l20
9	0	3969	0	2		2	VIRT	UP	UP	4	Access	l21
10	0	150	0	2		0	VIRT	UP	UP	4	Access	l22
11	0	3968	0	2		2	VIRT	UP	UP	4	Access	l23
12	0	151	0	2		1	VIRT	UP	UP	4	Access	l24
13	0	1	0	2		2	VIRT	UP	UP	0	Access	l25
14	0	3967	0	2		2	VIRT	UP	UP	4	Access	l26
15	1a030000	1 T	303	0		2	PHYS	UP	UP	4	Trunk	vmnic0
16	1a030100	1 T	303	1		2	PHYS	UP	UP	4	Trunk	vmnic1
47	1b030000	10	0	2		0	VIRT	UP	UP	4	Access	vmk0
48	1b030010	10	0	2		0	VIRT	UP	UP	4	Access	vmk1
49	1b030020	2	0	2		1	VIRT	UP	UP	4	Access	vswif0
50	1b030030	152	0	2		1	VIRT	UP	UP	4	Access	linux1.eth0
303	16000002	1 T	0	2		2	VIRT	UP	UP	4	Trunk	

- 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 4 on the physical ports.

```
[root@cae-cali-host6 ~]# vemcmd show trunk
Trunk port 15 native_vlan 1 CBL 4
vlan(1) cbl 4, vlan(2) cbl 4, vlan(10) cbl 4, vlan(150) cbl 4, vlan(151)
cbl 4, vlan(152) cbl 4, vlan(153) cbl 4, vlan(154) cbl 4, vlan(155) cbl 4,

Trunk port 16 native_vlan 1 CBL 4
vlan(1) cbl 4, vlan(2) cbl 4, vlan(10) cbl 4, vlan(150) cbl 4, vlan(151)
cbl 4, vlan(152) cbl 4, vlan(153) cbl 4, vlan(154) cbl 4, vlan(155) cbl 4,
```

- Vemcmd show port vlans

```
[root@cae-esx-184 ~]# vemcmd show port vlans
```

LTl	VSM Port	Mode	Native VLAN	VLAN State	Allowed Vlans
17	Eth7/1	T	1	FWD	2,10,150-155
18	Eth7/2	T	1	FWD	2,10,150-155
49	Veth8	A	10	FWD	10
50	Veth9	A	2	FWD	2
305	Po2	T	1	FWD	2,10,150-155



# 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?

```
[root@cae-esx-180 ~]# vemcmd show l2 150
Bridge domain 150 brtmax 1024, brtcnt 3, timeout 300
Flags: P - PVLAN S - Secure
      Type          MAC Address    LTL    timeout    Flags
PVLAN
Dynamic 00:50:56:81:0d:a4 304     1
Static 00:02:3d:40:ea:03 10      0
```

# VEM and VSM See Each Other's MAC

- Is the VSM learning the MAC of the VEM?

```
n1000v-AV# show mac address-table vlan 150
```

VLAN	MAC Address	Type	Age	Port	Mod
150	<b>0002.3d40.ea03</b>	static	0	N1KV Internal Port	6
150	0002.3d47.db05	static	0	N1KV Internal Port	6
150	0017.a4a8.340a	dynamic	16	Eth6/2	6

# 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 150
```

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

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

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

vlan	mac address	type	learn	age	ports
* 150	0050.5677.7770	dynamic	Yes	360	Gi3/48
* 150	<b>0002.3d40.ea03</b>	dynamic	Yes	330	Gi3/48
* 150	3333.0000.0016	static	Yes	-	Switch,Stby-Switch
* 150	<b>0050.5681.0da4</b>	dynamic	Yes	0	Gi4/19

## 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

# Check VSM Counters for Heartbeat

- On VSM

```
N1K-VSM# show module vem counters
```

```
-----  
Mod   InNR  OutMI  InMI  OutHBeats  InHBeats  InAipcMsgs  OutTO  OutTOC  InsCnt  RemCnt  
-----  
    3     1     1     1     82243     246554     247526     0       0       1       0
```

InNR - NodeID requests received count

OutMI - Module Insert Start requests sent to VEM

InMI - Module Insert Start responses received from VEM

**OutHBeats** - Number of HBs which have been broadcast by VSM

**InHBeats** - Number of HBs received from this VEM

**InAipcMsgs** - Number of AIPC msgs received from this VEM

OutTO - Number of aipc transmit timeout errors recorded for this VEM

# View Heartbeat Messages on VEM

- On the ESX host
- Use vempkt command to view Heartbeat messages
  - vempkt capture ingress vlan 150
  - vempkt display brief all
  - vempkt display detail all
  - vempkt stop
- Look for heartbeat messages from VSM – from detail all

```
***** Entry 484 *****
-----Packet Entry Information-----
      Timestamp   : Aug 10 21:42:18.388822
-----SF Packet Information-----
      Capture Stage : Ingress
      Source LTL    : 303
      Vlan          : 150
-----Packet L2 Header Information-----
      Source MAC Address : 00:50:56:9c:11:6f
      Destination MAC Address : 00:02:3d:40:ea:03
      Length          : 98
```

# View Heartbeat Messages With L3

- On the ESX host
- If you are using L3 you can use standard Linux commands
- L3 uses the eobc0 device
- `Ifconfig -a eobc0` - to see settings
- You can also use tools like tcpdump or tethereal  
`esx1>tethereal -i eobc0`
- Look for heartbeat messages from VSM

# Verifying Modules with “vem status”

- Run “vem status -v”

```
[root@cae-cali-host6 ~]# vem status -v
Package vssnet-esx4.1.0-00000-release
Version 4.0.4.1.1.30-1.9.16
Build 16
Date Fri Nov 13 17:41:59 PST 2009

Number of PassThru NICs are 0
VEM modules are loaded

Switch Name  Num Ports  Used Ports  Configured Ports  MTU  Uplinks
vSwitch0    32         1          32                1500
DVS Name     Num Ports  Used Ports  Configured Ports  Uplinks
n1000v-MV   256        13         256                vmnic1,vmnic0

Number of PassThru NICs are 0
VEM Agent (vemdpa) is running
```

- Reboot the ESX host if not all the modules have loaded.



# Verifying Modules With esxcfg-module

- Run “esxcfg-module -l | grep vem”
- Should see 4 modules loaded

```
[root@cae-cali-host6 bin]# esxcfg-module -l | grep vem
vem-v100-l2device 0x418030938000 0x5000 0x417ff1889e40 0x1000 28 Yes
vem-v100-n1kv 0x41803093d000 0xf000 0x417ff188ce80 0x3000 29 Yes
vem-v100-vssnet 0x41803094c000 0xd0000 0x417ff188fea0 0x58b000 2a Yes
vem-v100-stun 0x418030a1c000 0x12000 0x417ff1e1aee0 0x2000 2b Yes
```

# VEM Best Practices

- Control network should have low latency (50-100ms) and available bandwidth
- Match VEM version to VSM
- Use one uplink with all vlans
  - Segregate VM traffic with port-profiles and MAC pinning
- Upstream switch ports configured identically
- 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
```

Learn. Connect.  
Collaborate. *together.*

# Port-Profiles



# Uplink(eth) Port-Profile Troubleshooting

- Do not add multiple pnics from same ESX host to same uplink port-profile if no port-channeling is configured

While it may work you will end up with the ESX host receiving duplicate packets and require extra processing from CPU to deal with this improper configuration

- Do not configure multiple uplink port-profiles to an ESX host carrying the same vlan

Uplink1 and uplink2 to same ESX host both carrying vlan 100  
vPC-Mac pinning

- If you want NIC teaming use one of the approved port-channel mechanisms

# VM(veth) Port-Profile Troubleshooting

- VM port-profiles common issues
  - Port-profiles for Service Console and VMK should have system vlan set
- Be careful modifying veths directly
  - If at all possible modify the port-profile and not the veth
  - VSM remembers VM veths until they are deleted
    - Changes to a veth will stick around until the VM nic is deleted
- VM to veth mapping does not change until
  - NIC is removed from the VM
  - NIC is reassigned to another port-profile
- Use VMware VMXNET3 NIC type over E1000

# 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

# Cisco Nexus 1000V System VLANs

- System VLANs enable interface connectivity before an interface is programmed
- Address chicken and egg issue
  - VEM needs to be programmed, but it needs a working network for this to happen
- Port profiles that contain system VLANs are “system port profiles”
  - 32 from 1.3b onward
  - 16 in all older versions
- 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 Guidelines

- The system VLAN list must be a subset of the allowed VLAN list on trunk ports
- There must be only one system VLAN on an access port (the access VLAN)
- The 'no system vlan' command can be given only when no interface is using the profile.
- Once a system profile is in use by at least one interface, you can only add to the list of system VLANs, but not delete any VLANs from the list.
- Required System VLANs
  - Control, Packet, IP Storage, Service Console, VMKernel, any Management Networks



# System VLAN Example

- Migrate VMware Service Console to VEM
- SC interface uses VLAN 2
- Uplink port-profile must define VLAN 2 as system

```
n1000v# show run port-profile uplink-pinning
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
```

- Service Console Port-profile must also define system vlan

```
n1000v# show run port-profile SC
port-profile type vethernet SC
  vmware port-group
  switchport mode access
```

```
switchport access vlan 2
```

```
no shutdown
```

```
system vlan 2
```

# How do I recover?

- As of 1.3 you can now set system vlan on LTLs on ESX host
- From ESX/ESXi console

```
[root@cae-esx-180 ~]# vemcmd show port
```

LTL	IfIndex	Vlan	Bndl	SG_ID	Pinned_SGID	Type	Admin	State	CBL	Mode	Name
8	0	3969	0	32	32	VIRT	UP	UP	4	Access	l20
9	0	3969	0	32	32	VIRT	UP	UP	4	Access	l21
10	0	150	0	32	0	VIRT	UP	UP	4	Access	l22
11	0	3968	0	32	32	VIRT	UP	UP	4	Access	l23
12	0	151	0	32	0	VIRT	UP	UP	4	Access	l24
13	0	1	0	32	32	VIRT	UP	UP	0	Access	l25
14	0	3967	0	32	32	VIRT	UP	UP	4	Access	l26
15	0	3967	0	32	32	VIRT	UP	UP	4	Access	l27
16	1a040000	1 T	304	0	32	PHYS	UP	UP	4	Trunk	vmnic0
48	1b040000	10	0	32	0	VIRT	UP	UP	4	Access	vmk0
49	1b040010	2	0	32	0	VIRT	UP	UP	4	Access	vswif0
304	16000002	1 T	0	32	32	VIRT	UP	UP	4	Trunk	

```
[root@cae-esx-180 ~]# vemcmd set system-vlan 2 ltl 49
```

# Jumbo Frames Support

- MTU setting for “eth” type port-profile – 1.4
  - Simply use “mtu size” in port-profile and nothing else
  - Add system vlan directive to Port-profile if needed
- MTU setting for “eth” type port-profile – 1.3
  - “system mtu 9000” use with “system vlan” directive
  - No need to set MTU directly on eth interfaces or port-channels
  - Will consume a system port-profile
- Modify interfaces directly - 1.3
  - If not you do no need “system vlan” or do not want to consume “system port-profile”
  - Modify eth or port-channel interfaces directly
- “System jumbo mtu” global setting – all versions
  - Sets the system wide jumbo mtu size
  - Generally do not need to change

# Jumbo Frames Support - Examples

- Uplink for jumbo MTU - 1.4

```
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
```

- Uplink for Jumbo MTU – 1.3

```
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 1-149,151-1000
n1000v-AV(config-port-prof)# vmware port-group
n1000v-AV(config-port-prof)# system mtu 9000
n1000v-AV(config-port-prof)# no shut
n1000v-AV(config-port-prof)# system vlan 2,10
n1000v-AV(config-port-prof)# state enabled
```

# Jumbo Frames Support Confirmation

- 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
```

# 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>

# Port-profiles max-ports

- We default to 32 max-ports per port-profile
  - The default number of veths that can be assigned to a port-profile
- Max-ports number counts toward the maximum number of DVS ports that VMware can support
  - Even if veths are not yet assigned the ports are pre-provisioned
  - Some ports are consumed when you add an ESX host to the DVS
- 8192 ports per DVS in ESX 4.0U1/U2
- 20000 ports per DVS in ESX 4.1
- Note if you upgrade from 4.0 to 4.1 the number stays at 8192
  - VMware has a workaround to change the setting to 20000
- Example with vCenter 4.0U1
  - Create 10 port-profiles with max-ports set to 1024
  - Limit is 8192 ports per DVS
  - Will get error from vCenter about Max Total Ports Exceeded

# Port-Profile using Weighted QOS

- vMotion in 4.1 can now take advantage of 10Gb links
  - vMotion can use more bandwidth and more vMotions can run at once
  - This mean vMotion can overrun a network
- ESX 4.1 introduced NetIOC (Network I/O Control)
  - Allows for QOS to classify traffic based on type
  - Makes it easy to classify VMware traffic and assign it QOS
- Nexus 1000V can do this with QOS Fair Weighted Queuing
  - Works on egress uplink ports only
  - Easy to setup and configure
- Use “vemcmd show qos queue-rate ‘tl’” to verify VEM is matching packets



# Port-Profile using Weighted QOS

- Configuration Steps to limit vMotion traffic

```
n1kv-bl(config)# class-map type queuing match-all vmotion-class
```

```
n1kv-bl(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-bl(config-cmap-que)# match protocol vmw_vmotion
```

```
n1kv-bl(config)# policy-map type queuing vmotion-policy
```

```
n1kv-bl(config-pmap-que)# class type queuing vmotion-class
```

```
n1kv-bl(config-pmap-c-que)# bandwidth percent 50
```

```
n1kv-bl(config)# port-profile uplink-vpc
```

```
n1kv-bl(config-port-prof)# service-policy type queuing output vmotion-  
policy
```

# Port Channels



# Port Channels

- 3 load balancing modes

  - LACP Port-channels – require support on physical switch

  - vPC – Host Mode CDP/Manual allows balancing to multiple physical switches

    - NIC association is either Manual or via CDP

    - If more than one connection per physical switch then port-channel is required

  - vPC – Host Mode MAC Pinning works with any switch

    - Preferred Channeling method over vPC-HM CDP/Manual

    - Allows for pinning of veths (VM) to specific links.

- Channeling mode must be configured on uplink port-profile

  - Not recommended to create a channel by modifying individual Ethernet interfaces

# LACP Port Channels

- Use when single upstream or clustered (vPC,VSS) switch
- Use “**channel-group auto mode active**” on N1KV
- Use “**channel-group # mode active**” on upstream switch
- Switchports must be configured with
  - spanning-tree portfast trunk
  - spanning-tree bpdudfilter enable
- LACP PC will cycle between w and h states when uplink carries control and packet
  - One uplink is held back from LACP negotiation to guarantee control and packet have a valid link path while other links are aggregated
- VSM negotiates LACP for every VEM.
  - If VSM is down VEM cannot negotiate LACP and link will not come up

# LACP Offload

- New in 1.4
- Turned on with  
n1000v-AV(config)# lacp offload  
Requires a reboot of VSM to become active
- Feature moves LACP negotiation from VSM to VEM module
- Allows a VEM to negotiate LACP if VSM is not up
- Improves LACP stability

# 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

# Port Channels – vPC HM MAC Pinning

- Each Eth interface added is a unique Service Group
- Use “pinning id” command under vethernet port-profile
- veths will failover to another interface if pinned Ethernet interface fails
- If port-profile is not “pinned” vethernet interfaces are assigned Round Robin to an SG

# Port Channels – How to Tell Pinning

- Use “vemcmd show port” on ESX host

```
[root@cae-esx-184 ~]# vemcmd show port
```

LTL	VSM Port	Admin	Link	State	PC-LTL	SGID	Vem Port
17	Eth7/1	UP	UP	FWD	305	0	vmnic0
18	Eth7/2	UP	UP	FWD	305	1	vmnic1
49	Veth1	UP	UP	FWD	0	0	VSM-1.3a-2.eth0
50	Veth2	UP	UP	FWD	0	0	VSM-1.3a-2.eth1
51	Veth3	UP	UP	FWD	0	1	VSM-1.3a-2.eth2
52	Veth5	UP	UP	FWD	0	1	vswif0
53	Veth4	UP	UP	FWD	0	1	vmk0
305	Po1	UP	UP	FWD	0		



# Port Channels – Best Practice

- If the upstream switch can be clustered (VPC, VBS Stack, VSS) use **LACP**
- If you are using LACP also use LACP Offload  
Remember requires a VSM reboot to turn on
- If the upstream switch can NOT be clustered use **MAC-PINNING**
- Create channel-groups in port-profile  
Let VSM build the port-channel
- All physical switch ports in port-channel configured identical

# 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
```

- Highly Recommended Global BPDU Filter/Guard

IOS

```
cat65k(config)# spanning-tree portfast bpduguard
```

```
cat65k(config)# spanning-tree portfast bpduguard
```

NXOS

```
n5k-1(config)# spanning-tree port type edge bpduguard default
```

```
n5k-1(config)# spanning-tree port type edge bpdufilter default
```

- BPDU Filter is mandatory for LACP port-channels
- Set per port BPDU Filter/Guard when Global is not possible

# High Availability



# VSM High Availability (HA)

- VSM has 3 modes
  - Standalone – No HA, can later be converted
  - Primary
  - Secondary
- HA is limited to 2 VSM Virtual Machines (primary and secondary)
- Standby VSM is powered up
  - Standby cannot be powered down
- They must be on the same L2 Management network
- They must be on the same Control and Packet network
- Keep Primary and Secondary on different ESX hosts

# VSM – VSM Heartbeat

n1000v-MV# show system internal redundancy info

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

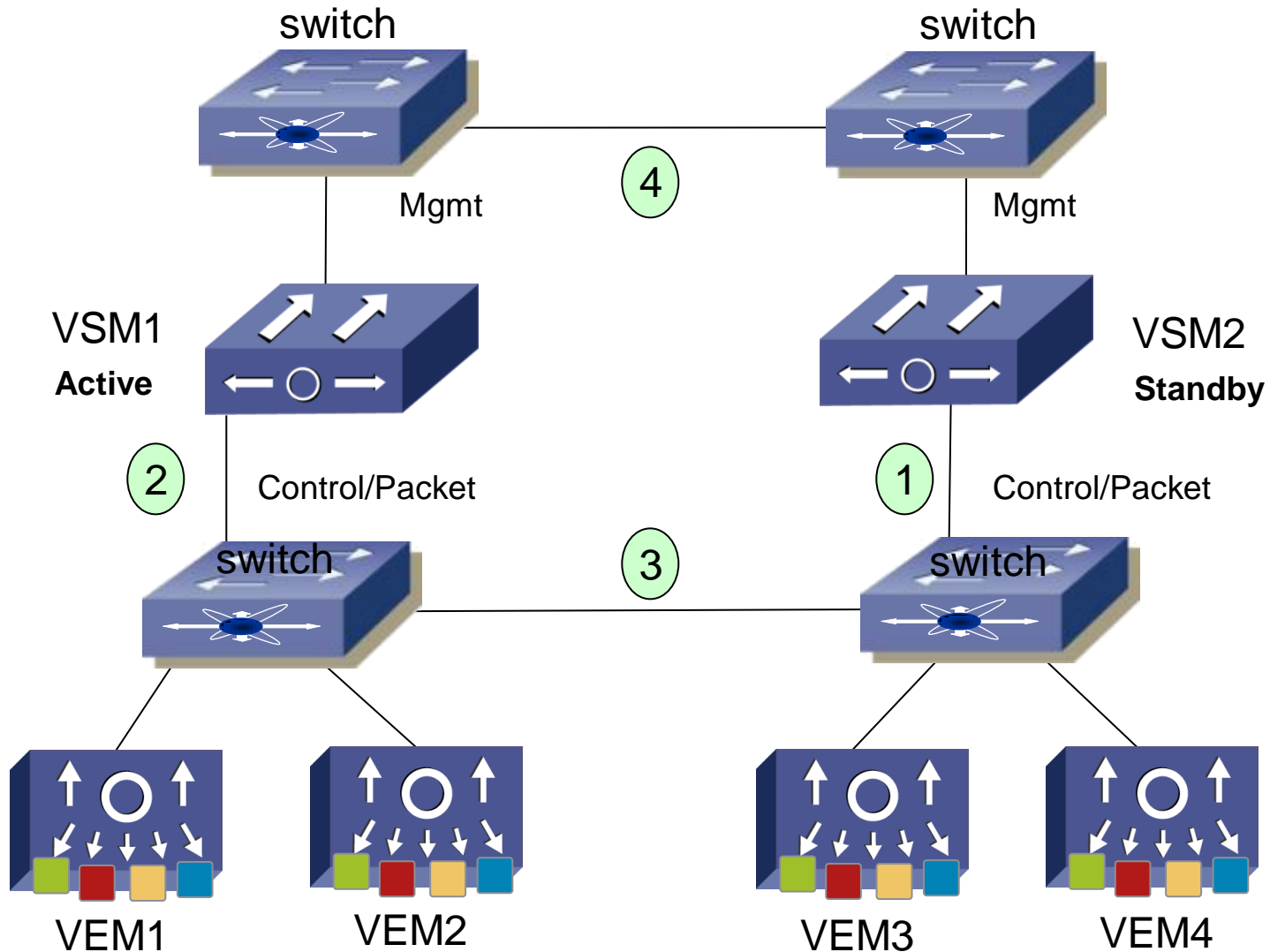
rx\_unknown\_pkts: 0

Statistics & Errors

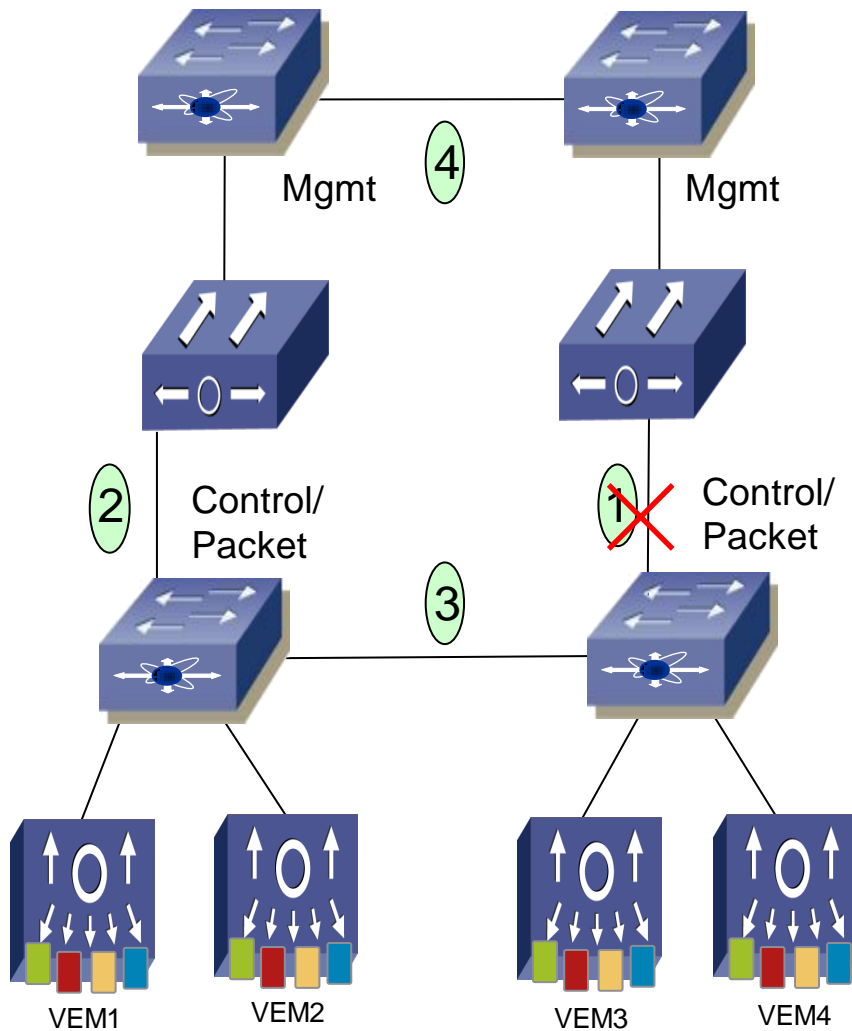
# VSM Active – Standby Connection Failure Scenarios

- Go through different link failures on VSMS
- Mgmt interface is used between VSMS to detect and prevent active/active scenario
- No Quorum Device

# Reference Topology



# Fail Scenario 1



## Failed Interface

### #1- VSM2 Control

**Effect:** (when communication lost)

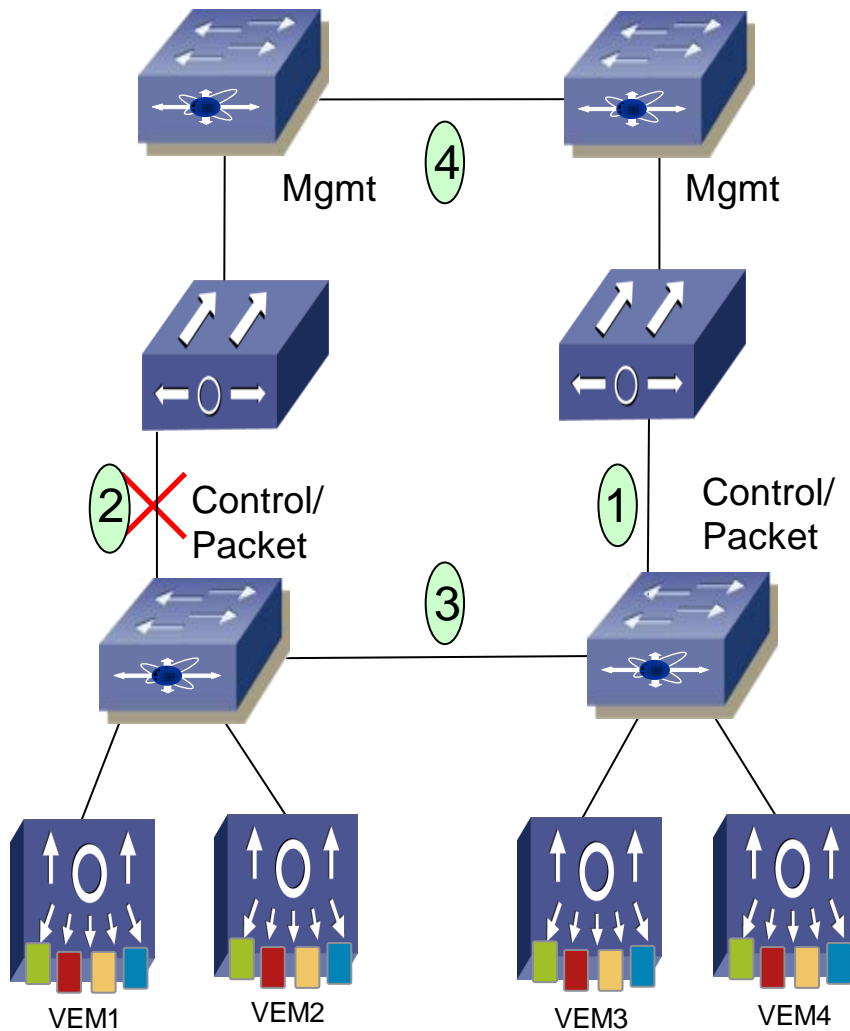
- VSM2 stays as Standby
- VSM2 resets until control is restored

**Exit:** (when communication restored)

- No VEM flap
- VSM2 resets and comes back up in HA mode



# Fail Scenario 2



## Failed Interface

### #2- VSM1 Control

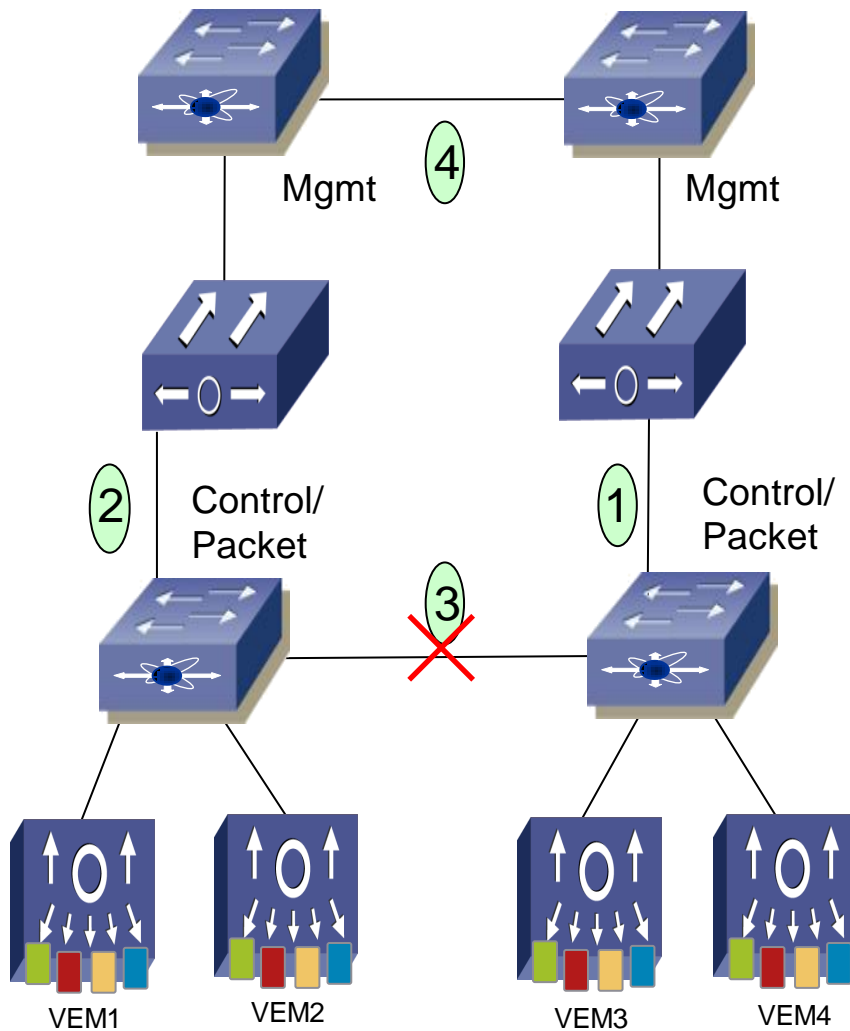
#### Effect:

- VSM1 drops VEMs.
- VSM2 stays as Standby
- VSM1 keeps resetting VSM2 until connectivity is restored

#### Exit:

- VSM2 comes back up in HA mode
- VEM Reconnect to VSM1

# Fail Scenario 3



## Failed Interface

### #3 - Split DVS

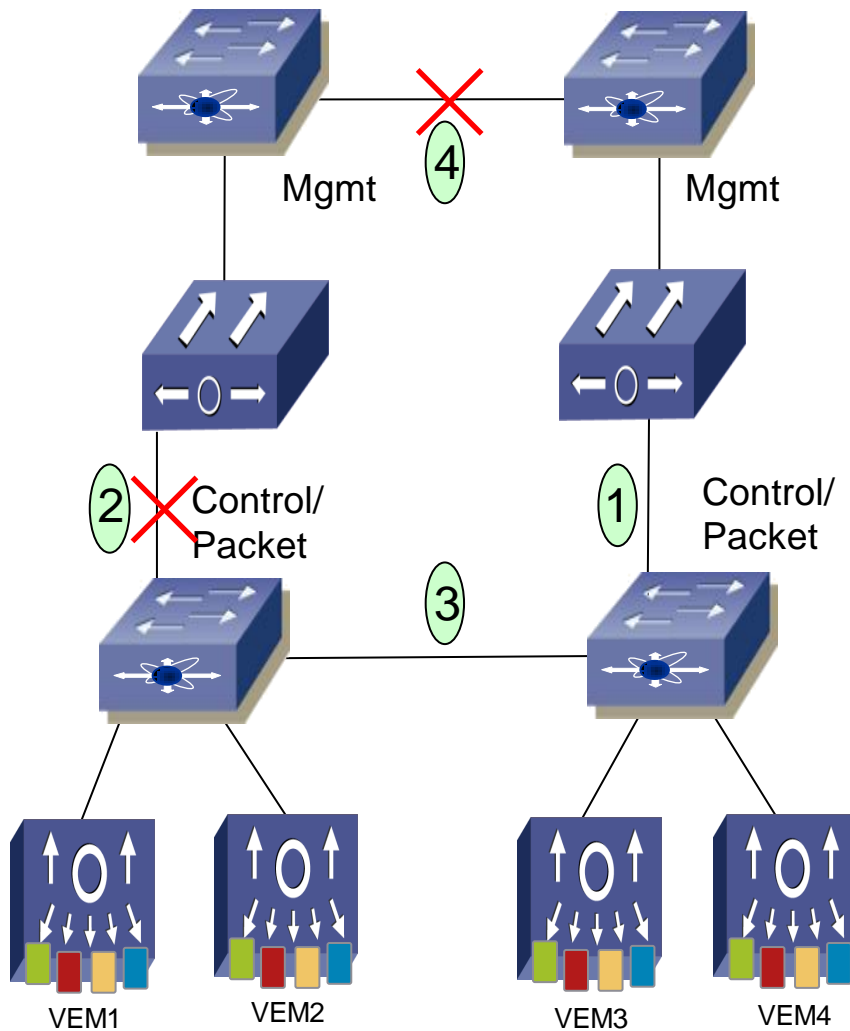
#### Effect:

- VSM1 drops VEM3,4.
- Otherwise, same as Fail Scenario 2
- VSM1 keeps resetting VSM2 until connectivity is restored

#### Exit:

- VEM3,4 reconnect to VSM1
- VSM2 comes back up in HA mode

# Fail Scenario 4



## Failed Interface

### #2 & #4 - Split Brain

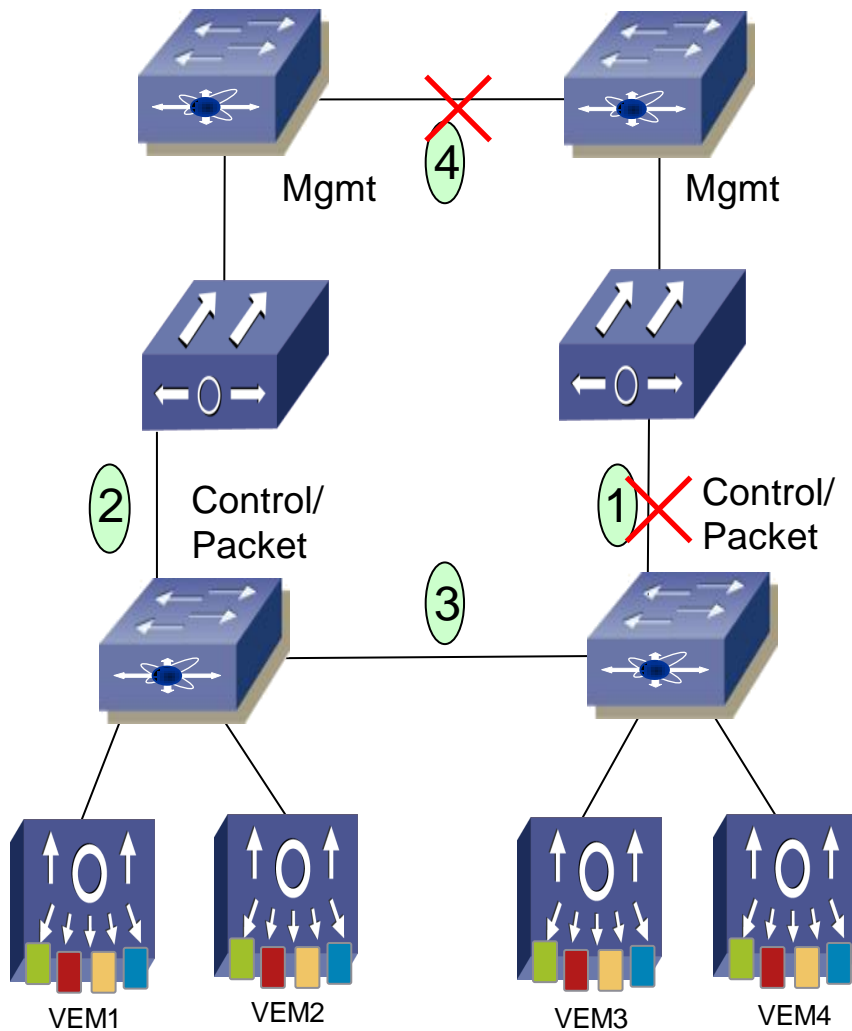
#### Effect:

- VSM2 becomes Active, taking over VEM1-4
- VSM1 stays active but drops all VEMs
- VSM1 and VSM2 use same IP address

#### Exit:

- VSM1 is reset when 1 and 4 are up again
- VEMs stay connected to VSM2

# Fail Scenario 5



## Failed Interface

### #1 & #4 - Split Brain

#### Effect:

- VSM2 becomes Active, but does not see any VEMs
- VSM1 stays active, handling all VEMs
- VSM1 and VSM2 use same IP address

#### Exit:

- VSM2 resets VSM1 when 2 and 4 are up again
- VEM1-4 connect to VSM2

# Cisco Nexus 1010

A decorative graphic element at the bottom of the slide, consisting of a horizontal orange line that curves downwards and to the right, forming a ramp-like structure with a white and grey checkered pattern on its side.

# Cisco Nexus 1010

- Runs NXOS
- Supports 4 VSMS and 1 Virtual NAM
- Troubleshooting VSMS on 1010 is identical to VSM on ESX/ESXi
- VSMS still have same restrictions
  - 64 VEMs per VSM
  - VSM tied to VMware Datacenter

# Cisco Nexus 1010

- Must be deployed in pairs
  - No option for a standalone 1010
- Must be in the same L2 domain for management and control
- Currently cannot be split across datacenters
  - Not meant for disaster recovery
- Splitting VSM primary and secondary between 1010 and ESX is not supported
- Uses same HA mechanism as VSM with domain id and control vlan
  - Do not overlap the domain id between a 1010 and a VSM

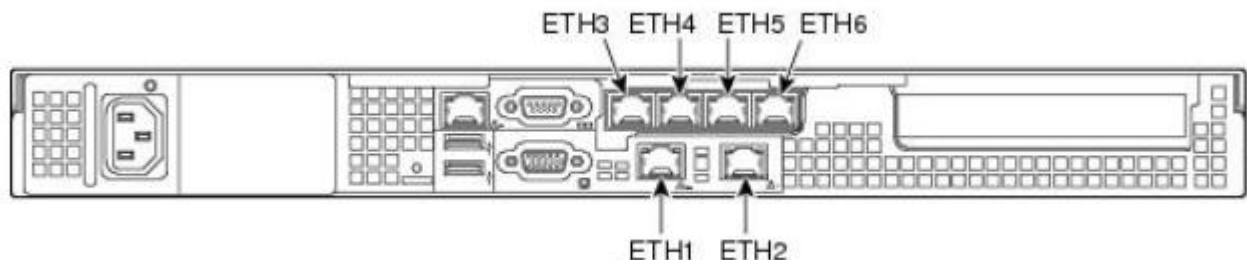
# Nexus 1010 Networking

- 1010 is based off UCS C200
- 6 x 1GB network connections
  - 2 LOM (Lan on motherboard) connections
  - 4 via Broadcom 2 port PCI card
- 4 distinct topologies supported by the 1010
- All ports must be trunk ports to 1010
  - 1010 does not support an access port nor native vlan designations
  - Native VLAN connectivity will not work with 1010



# Nexus 1010 Network Ports

- It is important to know which ports are which on the appliance
  - Picture of the back of the appliance
  - Eth1 and Eth2 are the onboard LOM ports
  - Eth3 – Eth6 are on the expansion card
- The expansion card should always be in the left slot of the appliance  
Do not move the card to the right slot



# Nexus 1010 Traffic

- Network traffic passed is classed into 3 categories
- Management
  - Carries the mgmt 0 interface of the 1010
  - Also all mgmt 0 traffic for every VSM installed on the 1010
  - Requirement that VSM mgmt and 1010 mgmt be on the same subnet
  - VSM installed on 1010 automatically inherits the mgmt vlan of the 1010 for its mgmt
- Control
  - Carries all the control and packet traffic for the VSMS installed on the 1010
  - Carries control traffic for HA between primary and secondary 1010
- Data
  - Used by Virtual Service Blades (VSB) other than VSM
  - Currently carries traffic for NAM blade

# Nexus 1010 Topologies

- 4 topologies

1. Single uplink

Uses eth1 and eth2

Management, Control and Data run over the same links

2. Two uplinks - 1

Uses all eth interfaces

Management and Control are combined

Data is separated out for maximum throughput and redundancy

3. Two uplinks - 2

Uses all eth interfaces

Management is separated

Control and Data are combined

4. Three uplinks

Uses all eth interfaces

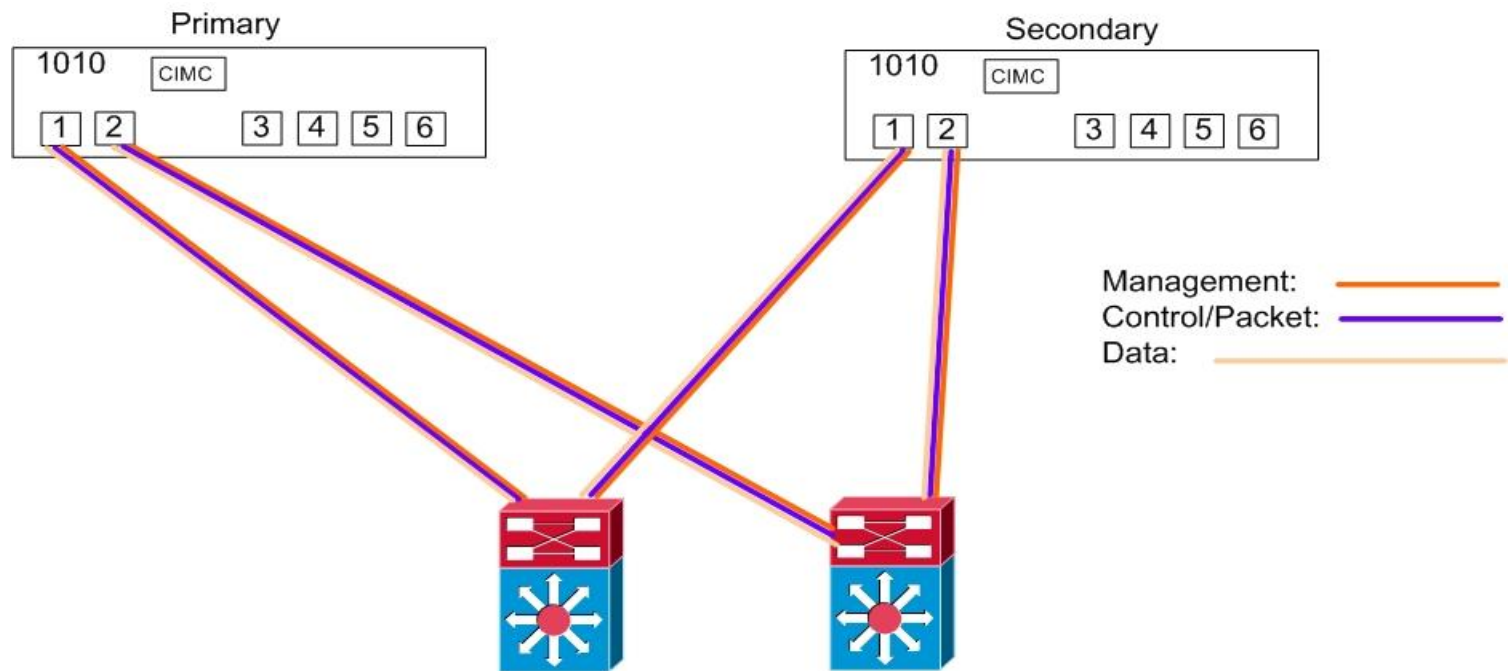
Management, Control, and Data all have dedicated eth ports

# Single Uplink

- Uses eth1 and eth2
- No port-channel can be configured or supported
- Traffic only passes over one link. Second link is for failover only
- 1GB of throughput
- Single eth connected (ex not connecting eth2) is not supported and can cause problems
- Best practice that eth1 and eth2 connect to different switches for redundancy
- Eth1 and eth2 switch ports need to be configure identically

# Examples of Valid Configuration

- Traffic will only flow over one link (ex eth1)
- Should eth1 fail traffic will fail to eth2
- Should eth2 fail services will fail to secondary 1010



# Two Uplinks Version 1

- Management and Control are still combined

  - Eth1 and Eth2 dedicated for Management and Control

  - 1GB throughput

  - Simple redundancy

- Data traffic for vNAM is broken out for best redundancy and throughput

  - Eth3 - Eth6 are used explicitly for Data traffic

  - Eth3 and Eth5 are paired** and must connect to same switch and port-channel (LACP)

  - Eth4 and Eth6 are paired** and must connect to same switch and port-channel (LACP)

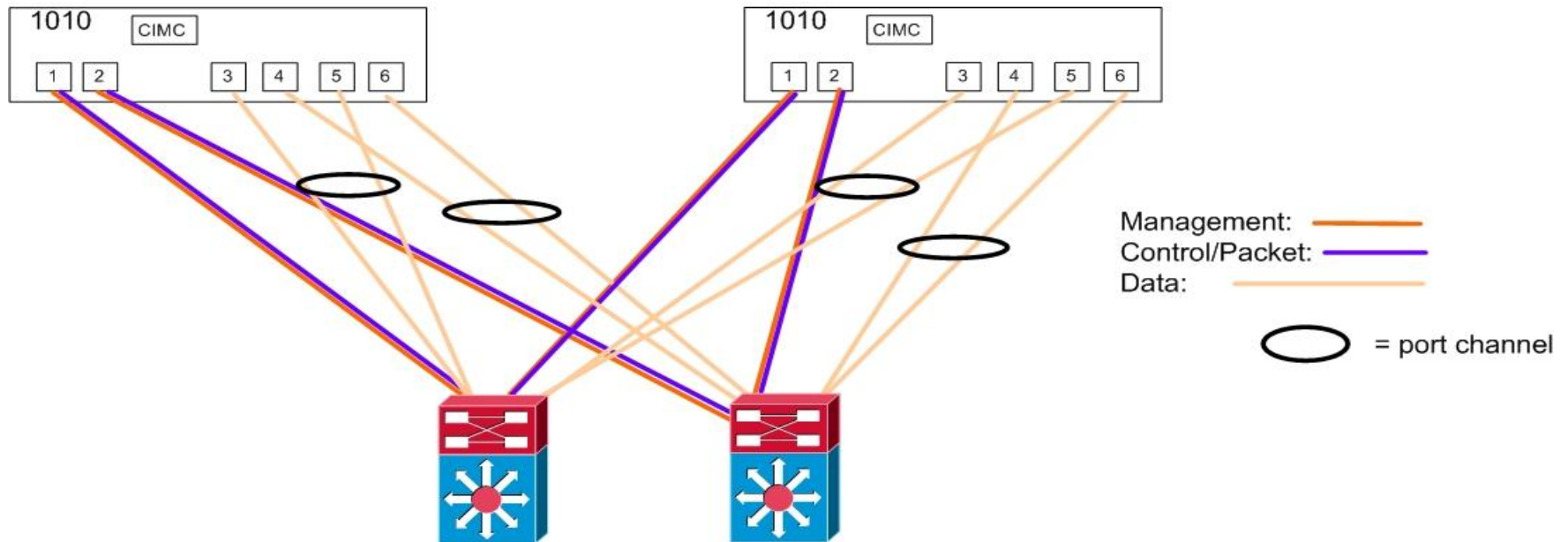
  - 2 GB of throughput for Data traffic

  - Dual redundancy

  - Best for maximum data throughput and redundancy

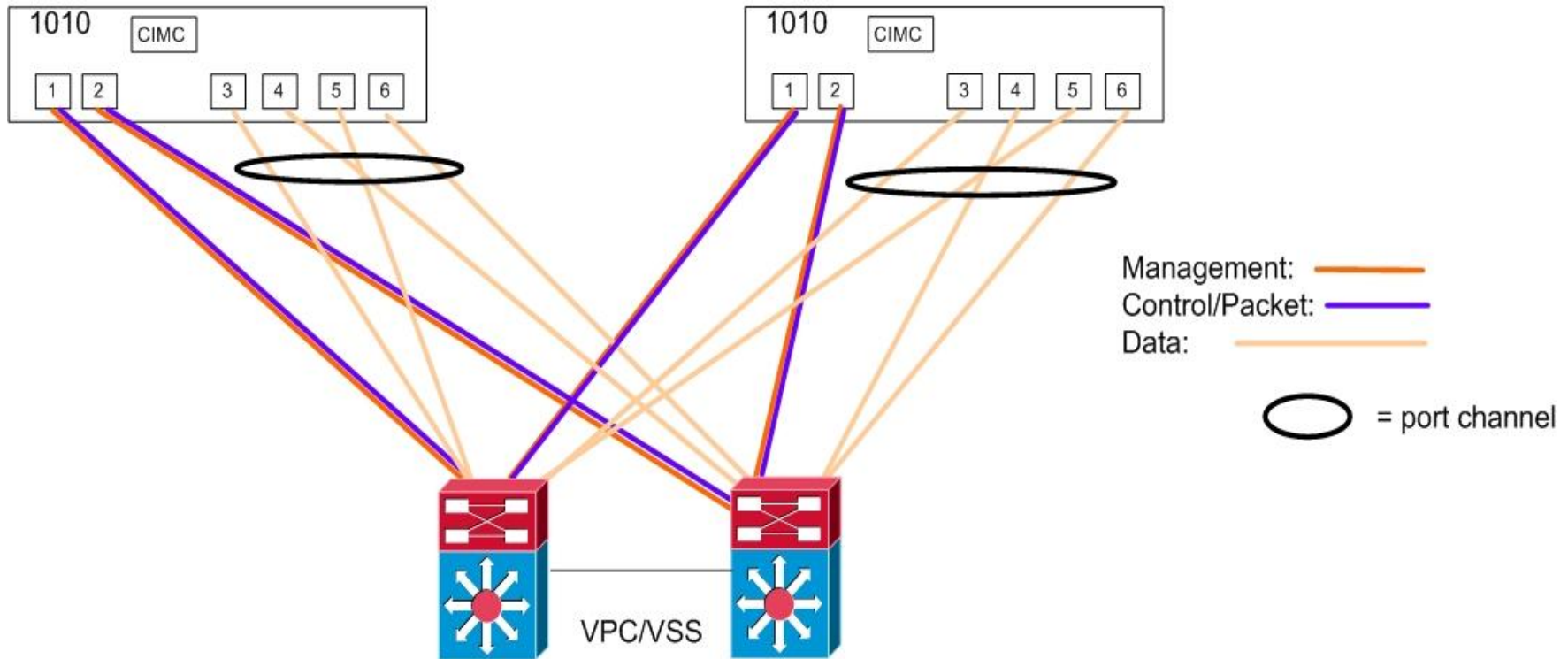
# Examples of Valid Configuration

- Paired eth interfaces must be LACP port-channelled to same switch
- Even eth interfaces to one switch and odd eth interfaces to another switch
- Traffic will only pass over one of the port-channels
- Second port-channel is only used if primary port-channel fails



# Two uplinks Version 1 and VPC/VSS

- Eth3-6 become one big port-channel
- 4GB of throughput



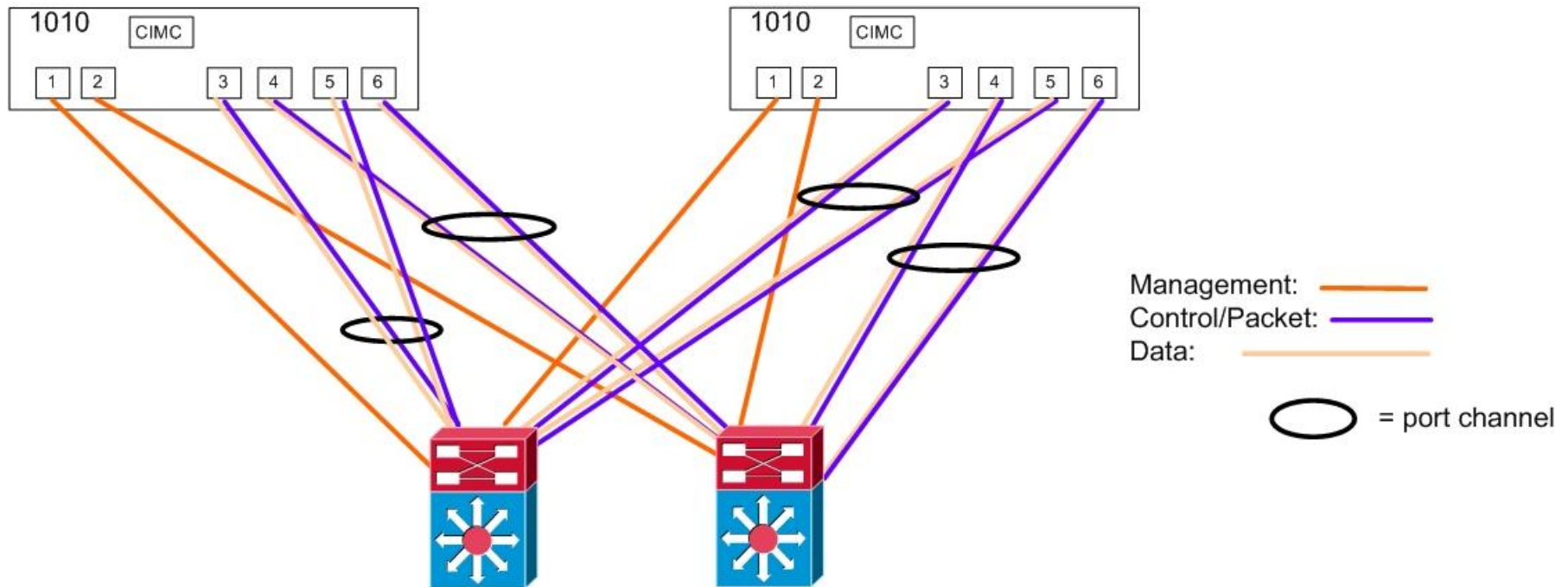


# Two Uplinks Version 2

- Management gets dedicated interfaces
  - Eth1 and Eth2 dedicated for Management
- Data and Control share interfaces Eth3-Eth6
- Best solution for Control (assuming no vNAM)
  - 2GB of throughput
  - Maximum redundancy
- Same as Uplinks Version 1 for Eth interfaces
  - Eth3 and Eth5 are paired** and must connect to same switch and port-channel (LACP)
  - Eth4 and Eth6 are paired** and must connect to same switch and port-channel (LACP)
  - Data and Control share interfaces
- Cannot define which uplink is used for specific VSMS
  - Traffic is automatically assigned and load balanced

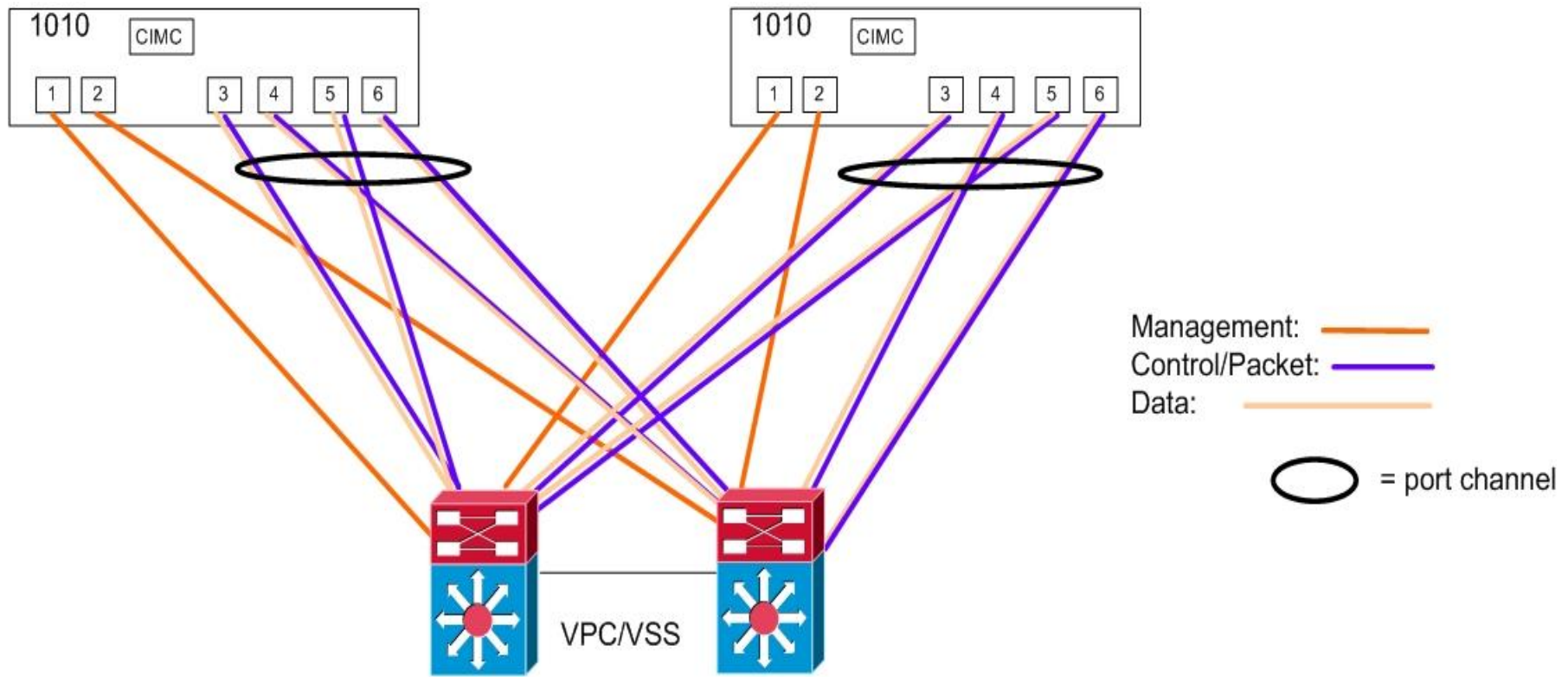
# Examples of Valid Configuration

- Paired eth interfaces must be LACP port-channelled to same switch
- Even eth interfaces to one switch and odd eth interfaces to another switch



# Two Uplinks Version 2 and VPC/VSS

- Eth3- 6 become one big port-channel
- 4GB of throughput



# Three Uplinks

- **Management gets dedicated interfaces**

- Eth1 and Eth2 dedicated for Management
  - 1GB of throughput ; single layer redundancy

- **Control gets dedicated interfaces**

- Eth3 and Eth4 dedicated for Control
  - 1GB of throughput ; single layer redundancy

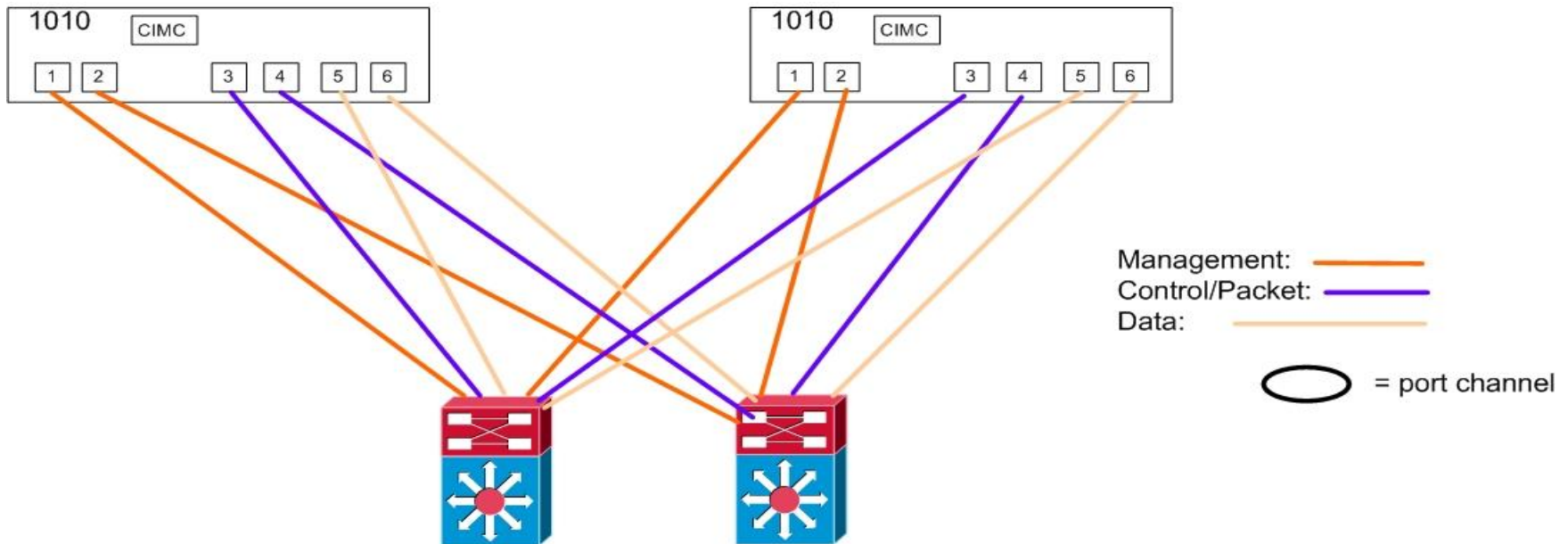
- **Data gets dedicated interfaces**

- Eth5 and Eth6 dedicated for Data
  - 1GB of throughput ; single layer redundancy

- **Best for solutions where Control and Data need dedicated links**

# Examples of Valid Configuration

- Only one link for Control or Data is active
- Ex. All traffic for Control will pass over eth3
- Eth4 is only used if eth3 fails
- If eth4 fails then services failover to secondary 1010



# Recommendations

- If you are not planning on using vNAM

Two uplinks version 2 gives best bandwidth and redundancy for control VLAN

Negative is that is harder to configure

Three uplinks would be next best scenario as it gives control dedicated NICs but no LACP

- Biggest key to deploying is to determine how much traffic flow you expect vNAM to get
- No issues using vPC or VSS
- You don't need to cable all the connections if you are not using a feature

Example if you use Three Uplinks but are not using vNAM you do not need to cable up interfaces eth5 and eth6

# Summary



# Summary

- Cisco Nexus 1000V troubleshooting document  
[http://www.cisco.com/en/US/docs/switches/datacenter/nexus1000/sw/4\\_0\\_4\\_s\\_v\\_1\\_2/troubleshooting/configuration/guide/trouble\\_n1000v.html](http://www.cisco.com/en/US/docs/switches/datacenter/nexus1000/sw/4_0_4_s_v_1_2/troubleshooting/configuration/guide/trouble_n1000v.html)
- Cisco MyCommunity Nexus 1000V Space  
<https://www.myciscocommunity.com/community/products/nexus1000v>
- Cisco Nexus 1000V FAQ  
<https://www.myciscocommunity.com/docs/DOC-14464>
- VMware ESX Networking Community  
<http://communities.vmware.com/community/vmtn/vsphere/networking?view=discussions&start=0>





# Q & A

# Complete Your Online Session Evaluation

- Give us your feedback and you could win fabulous prizes. Winners announced daily
- Receive 20 Cisco Preferred Access points for each session evaluation you complete
- Complete your session evaluation online now (open a browser through our wireless network to access our portal) or visit one of the Internet stations throughout the Convention Center



Don't forget to activate your Cisco Live and Networkers Virtual account for access to all session materials, communities, and on-demand and live activities throughout the year. Activate your account at any internet station or visit [www.ciscolivevirtual.com](http://www.ciscolivevirtual.com).

Thank you.

