
TOMORROW starts here.
cisco
Cisco live!

Hypervisors Networking: Best Practices for Interconnecting with Cisco Switches

BRKVIR-2019

Michael Ciesla
CCIE R&S #38421
Customer Support Engineer

Cisco live!

Thong?

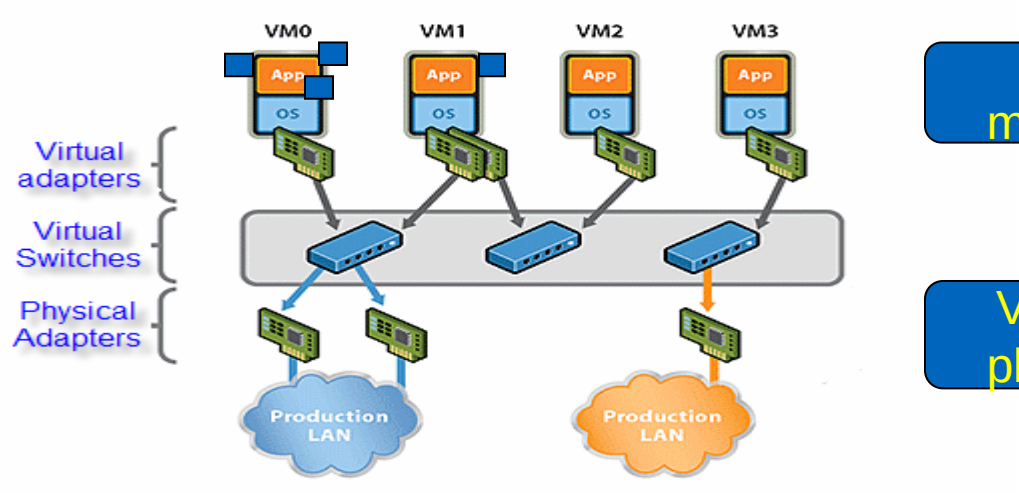
What is this Session About?

- Networking Virtualisation Concepts
- Hypervisor Overview
  - VMware vSphere ESXi 5.5
  - Microsoft Windows Server 2012 R2 (Hyper-V 3.0)
  - Citrix XenServer 6.2
- Topology Overview
- Nexus 1000v

Networking Virtualisation Concepts

Hypervisor Networking Virtualisation

- VM-to-VM and VM-to physical-host, traffic handled via software switch that lives in the hypervisor



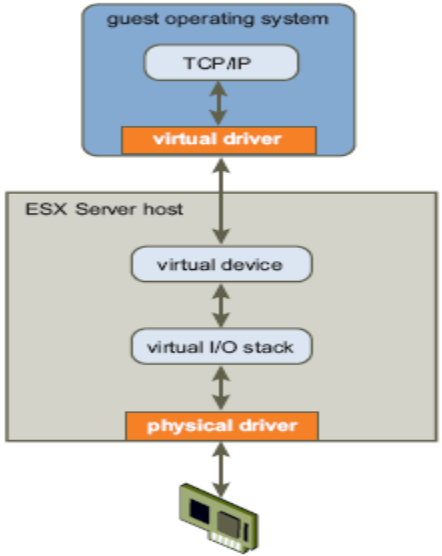
VM-to-VM: memory transfer

VM-to-physical: physical adapter

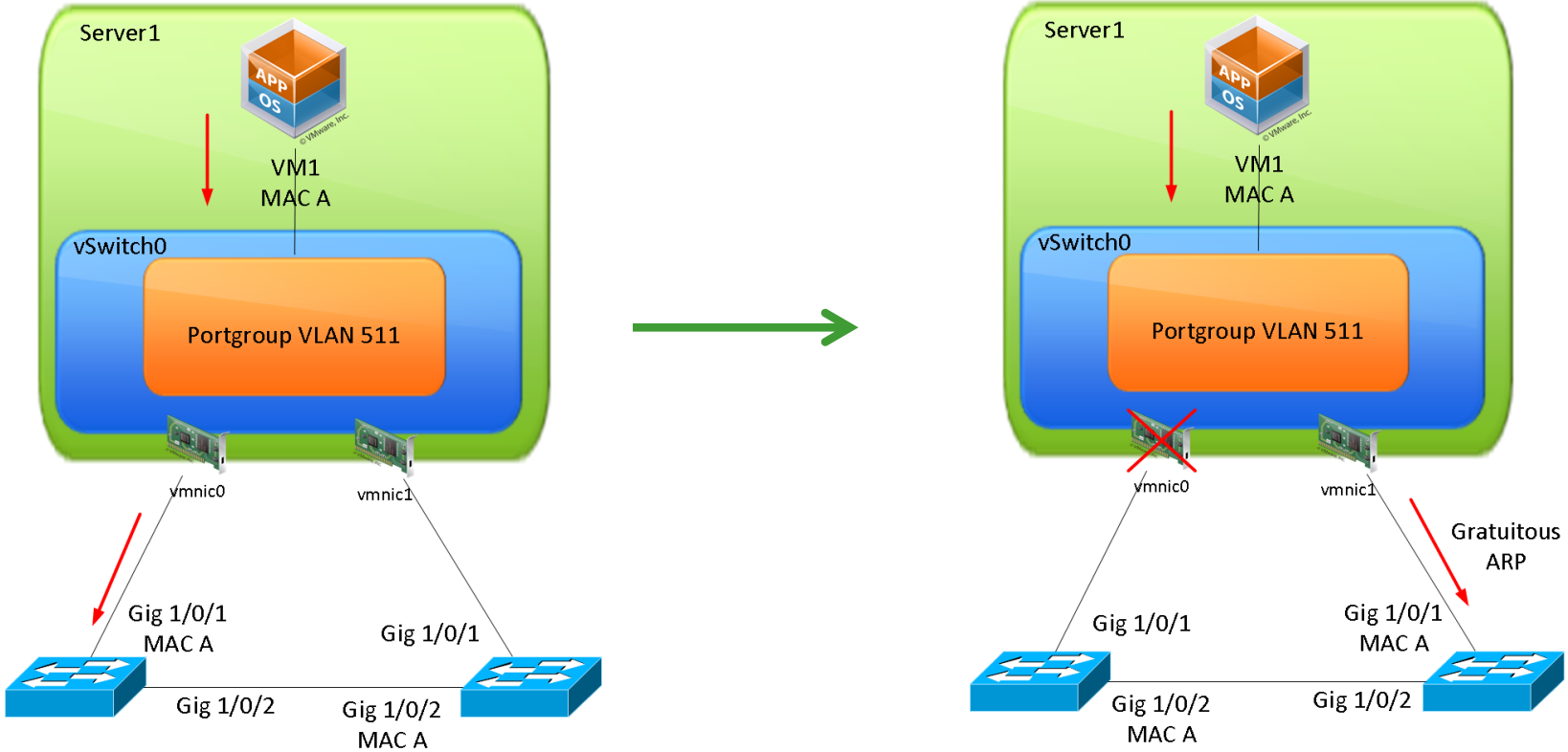
Hypervisor Networking Virtualisation

VMware vSphere ESXi example

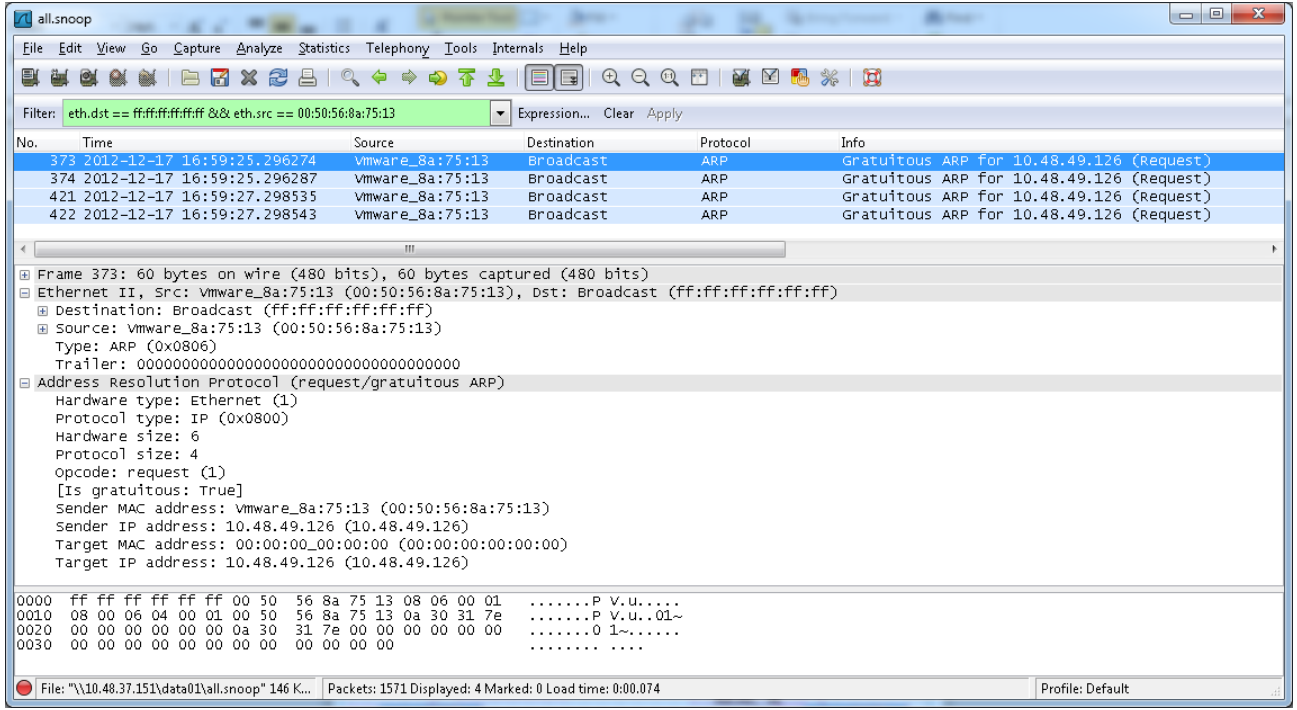
- VMs don't directly control networking hardware
  - x86 hw designed to be handled by only one device driver!
- When a VM communicates with the outside world, it:
  - … passes the packet to its local device driver …
  - … which in turns hands it to the hypervisor…
  - … which in turns passes it to the physical NIC
- ESX gives VMs several device driver options:
  - Strict emulation of Intel's e1000 / e1000e
  - VMware vmxnet/vmxnet3: paravirtualised!
- VMs have MAC addresses that appear on the wire



Gratuitous ARP



Gratuitous ARP



Spanning-Tree?

- Disable Spanning-tree
  - NX-OS: spanning-tree port type edge [trunk]
  - IOS: spanning-tree portfast

- Link flap with spanning-tree enabled = 30s network outage

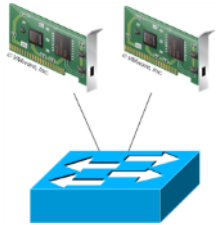
Link Aggregation Demystified

Port-channel, Etherchannel, LACP, 802.3ad, 802.1ax, LAG, Teaming, Bonding

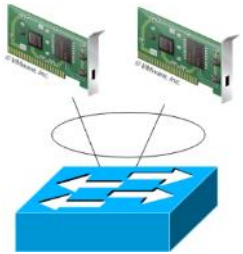
- IEEE 802.3ad predecessor to 802.1AX
- 802.1AX defines Link Aggregation Group (LAG)
- LAG can be static or dynamic
- Link Aggregation Control Protocol (LACP) is part of 802.1AX

Switch Independent / Dependent

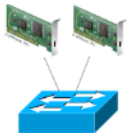
- Independent - no specific switch protocol / configuration is required



- Dependent - specific switch configuration is required (port-channel)



Switch Independent Configuration Options (NX-OS)



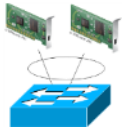
Access

```
interface Ethernet101/1/31
  switchport mode access
  switchport access vlan 511
  spanning-tree port type edge trunk
```

Trunk

```
interface Ethernet101/1/31
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree port type edge trunk
```

Switch Dependent Configuration Options (NX-OS)



Static

```
interface Ethernet101/1/31-32
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree port type edge trunk
  channel-group 300
```

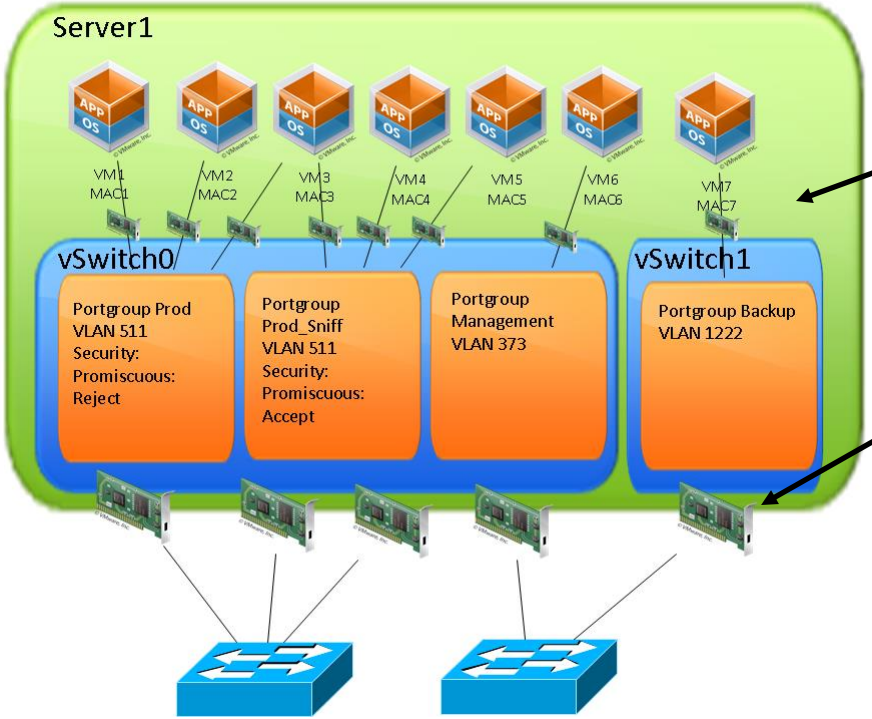
LACP

```
interface Ethernet101/1/31-32
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree port type edge trunk
  channel-group 300 mode active
```

```
interface Port-Channel300
  switchport mode trunk
  switchport trunk allowed vlan 511
```

Hypervisor Overview
VMware vSphere ESXi

vSphere Standard Switch (vSS)

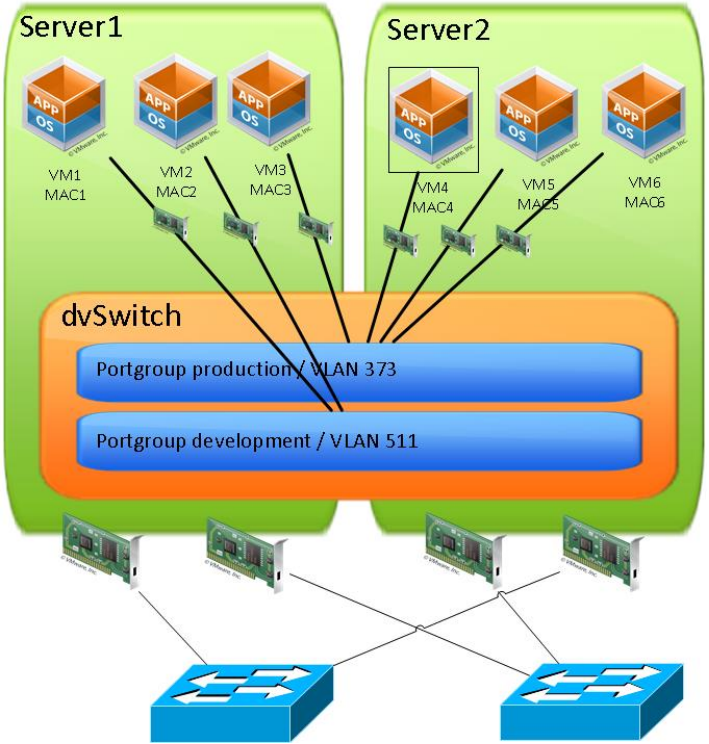


Virtual NIC (vNIC)

Physical NIC (pNIC) aka vmnic

Portgroup != VLAN

vSphere Distributed Switch (vDS)



- Spans multiple servers
- Enterprise Plus license required

VMware vSphere ESXi

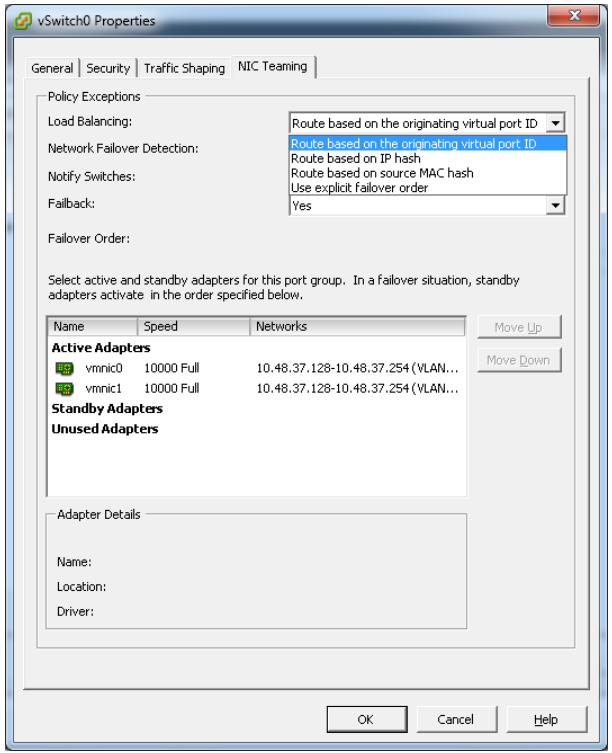
- Route based on originating virtual port
- Route based on source MAC hash
- Use explicit failover order
- Route based on physical NIC load (vDS)
- Route based on IP hash
- Route based on IP hash + LACP (vDS)

Switch Independent

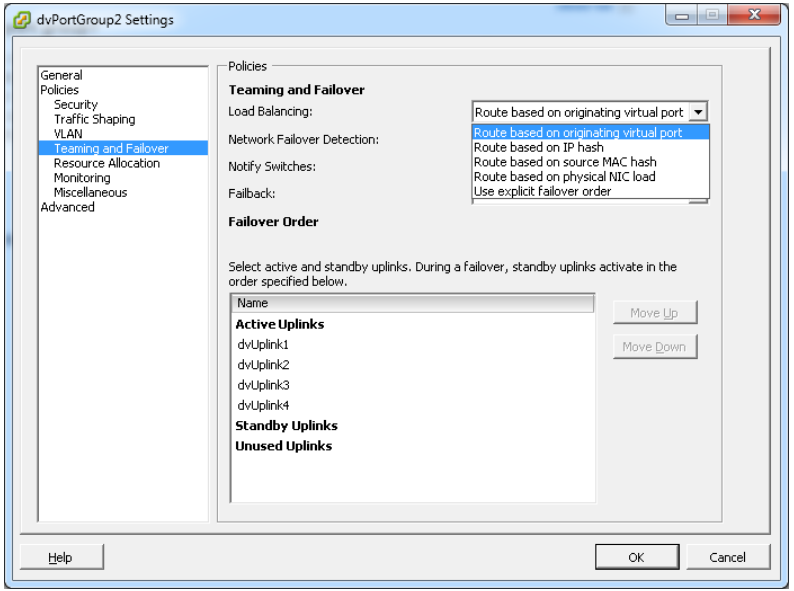
Switch Dependent

Uplink Options

vSphere Standard Switch



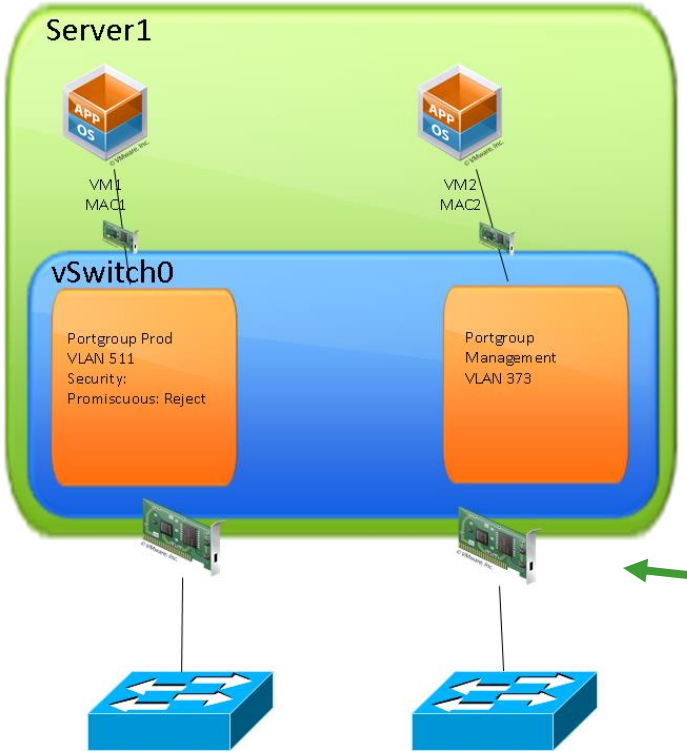
vSphere Distributed Switch (vDS)



vDS (extra):

- Route based on physical NIC load
- Route based on IP hash + LACP

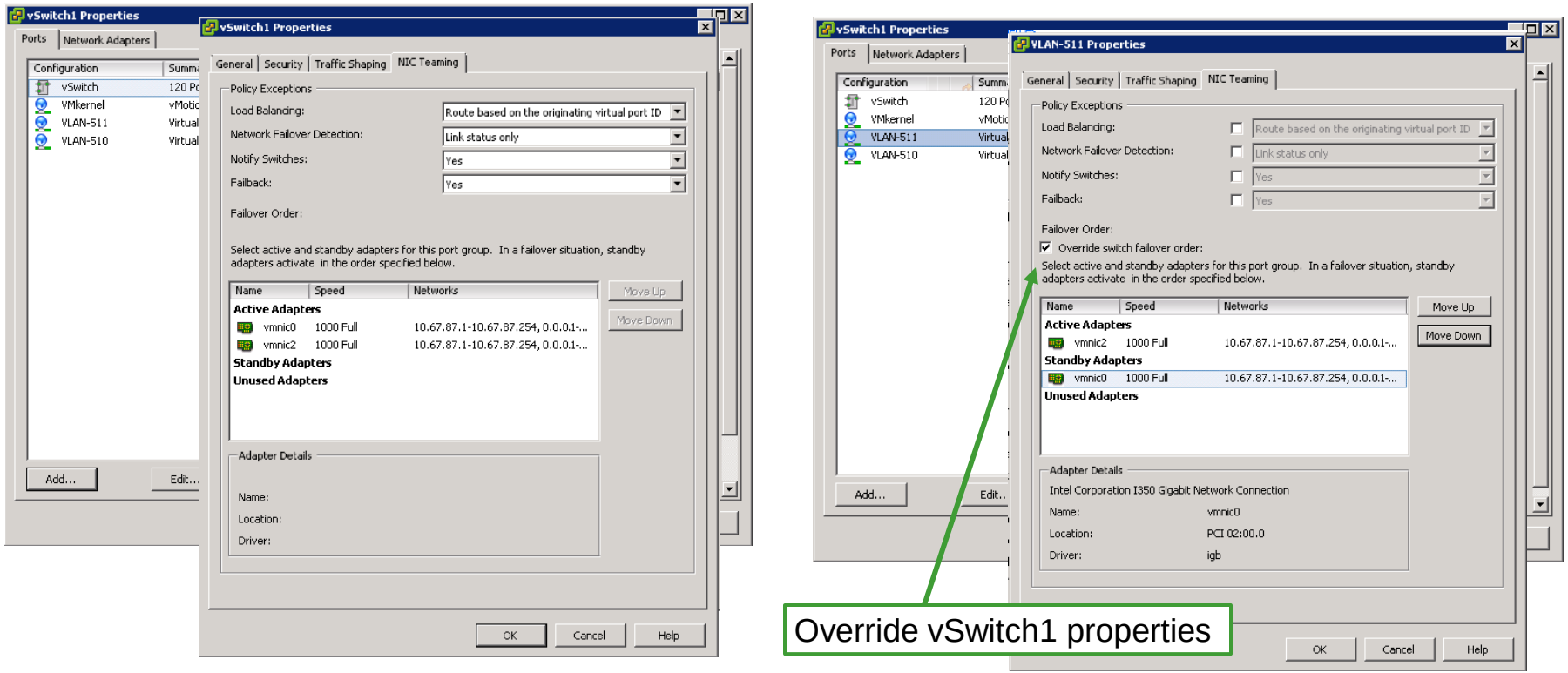
Load Balancing



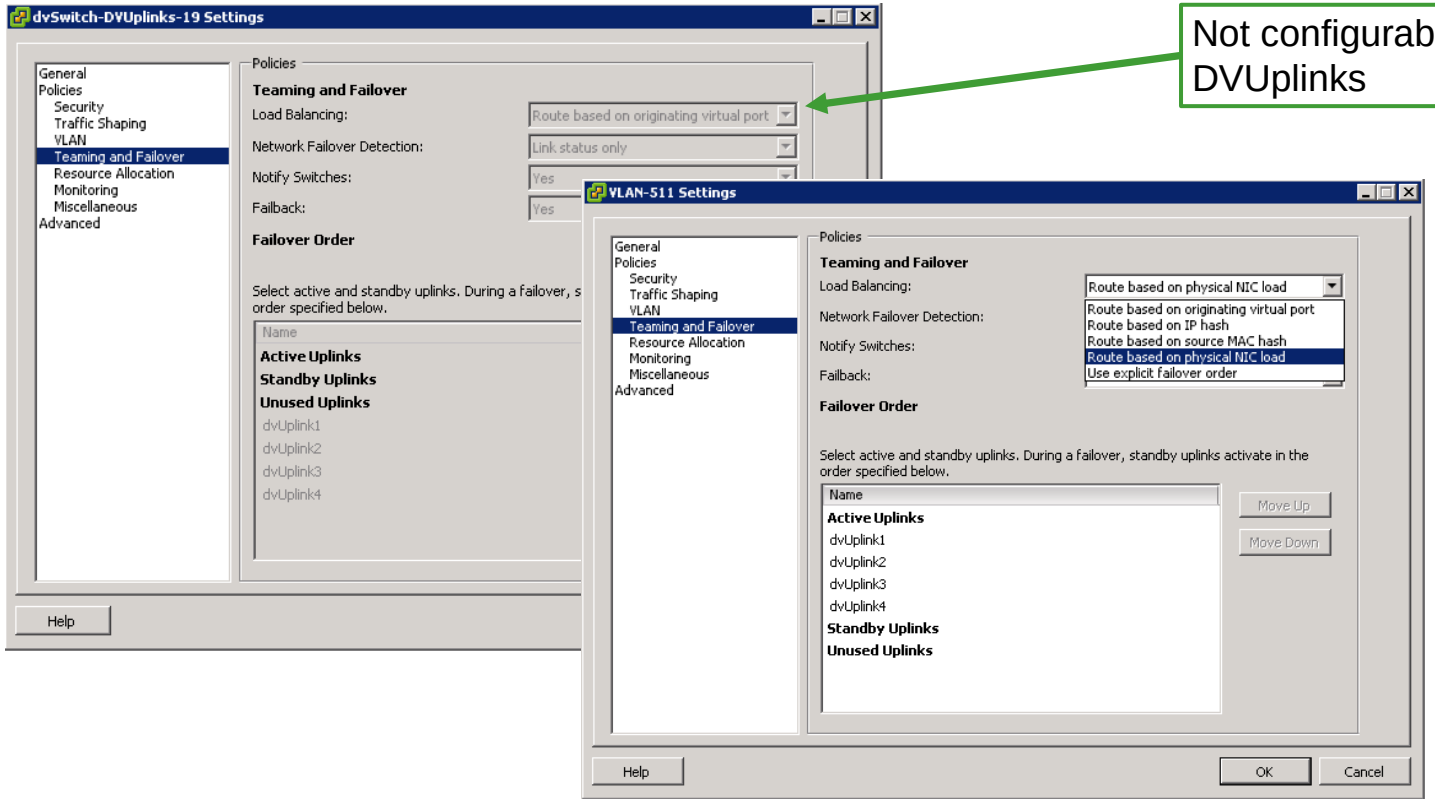
Per portgroup:
Route based on originating virtual port
Route based on IP hash
Route based on source MAC hash
Route based on physical NIC load (vDS)
Use explicit failover order

Per DVUplink:
Route based on IP hash + LACP (vDS)

Load Balancing: VMware Standard Switch

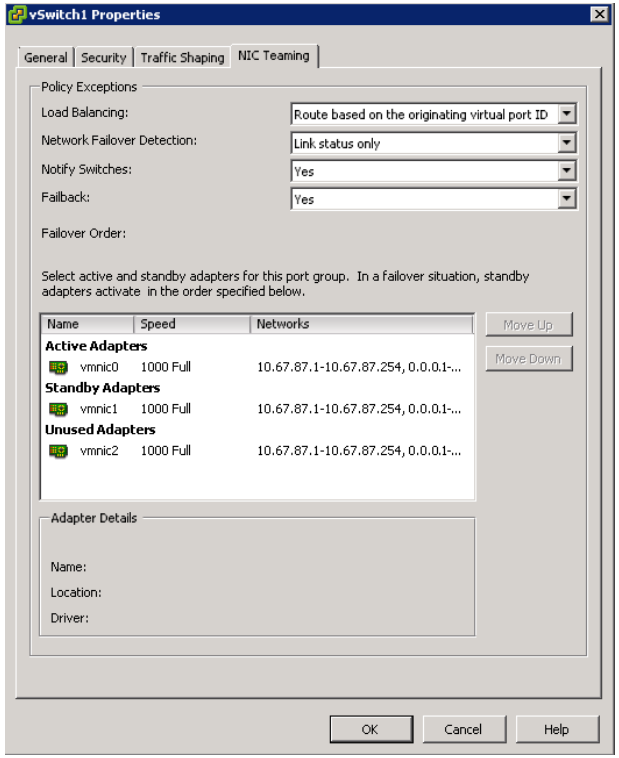


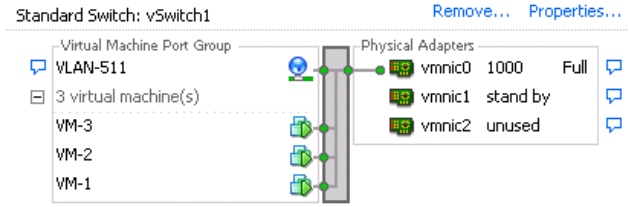
Load Balancing: VMware Distributed Switch



Not configurable on the DVUplinks

Active / Standby / Unused





vmnic0

Ethernet101/1/31 is up

vminc1

Ethernet101/1/31 is up

vmnic2

Ethernet101/1/32 is up

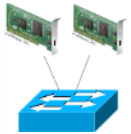
Cisco Discovery Protocol

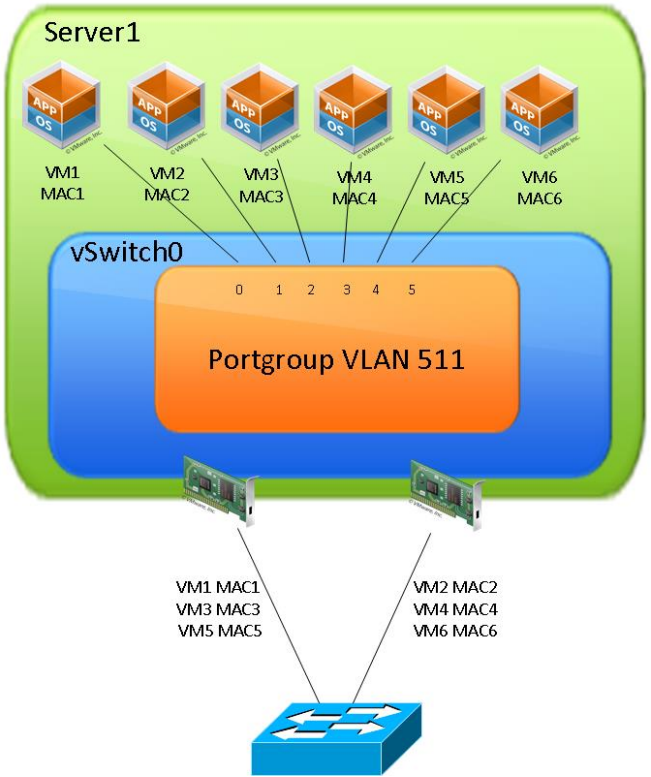
Properties	
Version:	2
Timeout:	0
Time to live:	175
Samples:	253
Device ID:	SV-5K-1(FOX1726GPAC)
IP Address:	10.67.86.4
Port ID:	Ethernet101/1/31
Software Version:	Cisco Nexus Operating System (...
Hardware Platform:	N5K-C5596UP
IP Prefix:	0.0.0.0
IP Prefix Length:	0
VLAN:	511
Full Duplex:	Enabled
MTU:	1500
System Name:	SV-5K-1
System OId:	1.3.6.1.4.1.9.12.3.1.3.1038
Management Address:	10.67.83.198
Location:	snmplocation
Peer Device Capability Enabled	
Router:	No
Transparent Bridge:	No
Source Route Bridge:	No
Network Switch:	Yes
Host:	No
IGMP:	Yes
Repeater:	No

Route Based on Originating Virtual Port

vmware

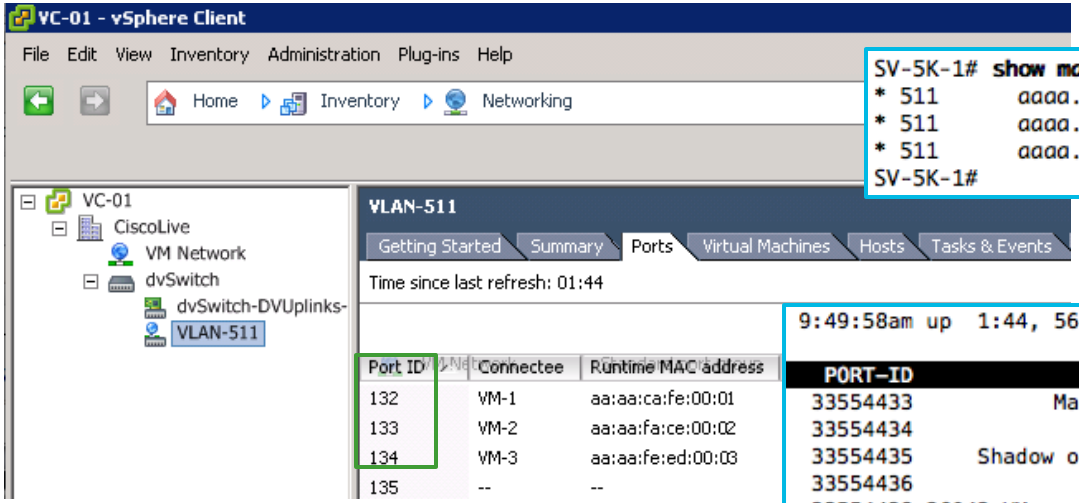
vSS + vDS





- Default configuration
- Even distribution of traffic if number of vNICs > physical adapters
- vNIC pinned to single physical NIC
- Can go to one or multiple switches

Route Based on Originating Virtual Port

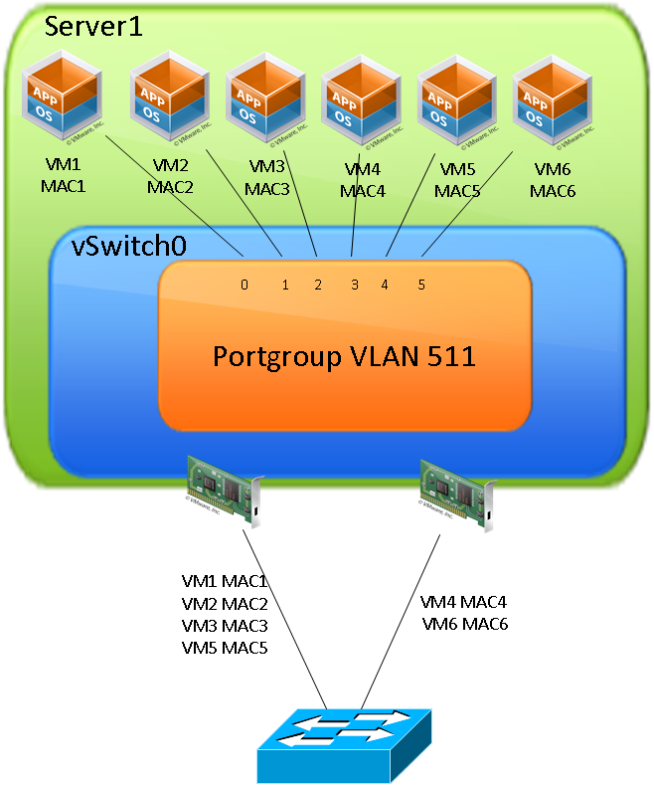


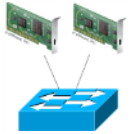

```
SV-5K-1# show mac address-table | grep aaa
* 511       aaaa.cafe.0001    dynamic   0         F    F  Eth101/1/31
* 511       aaaa.face.0002    dynamic   0         F    F  Eth101/1/31
* 511       aaaa.feed.0003    dynamic   0         F    F  Eth101/1/32
SV-5K-1#
```

```
9:49:58am up  1:44, 567 worlds, 4 VMs, 5 vCPUs; CPU load average: 0.01, 0.01,

  PORT-ID              USED-BY  TEAM-PNIC DNAME              PKTTX/s  MbTX/s
 33554433           Management       n/a vSwitch0            0.00    0.00
 33554434                vmnic3         - vSwitch0           53.10    1.52
 33554435     Shadow of vmnic3       n/a vSwitch0            0.00    0.00
 33554436                 vmk0    vmnic3 vSwitch0           53.10    1.52
 33554438 36042:VMware vCenter    vmnic3 vSwitch0            0.00    0.00
 50331649           Management       n/a DvsPortset-0        0.00    0.00
 50331650                vmnic0         - DvsPortset-0        2.16    0.00
 50331651     Shadow of vmnic0       n/a DvsPortset-0        0.00    0.00
 50331656      46658:VM-1.eth0    vmnic0 DvsPortset-0        0.98    0.00
 50331657      46777:VM-2.eth0    vmnic0 DvsPortset-0        1.18    0.00
 50331658      48692:VM-3.eth0    vmnic2 DvsPortset-0        1.18    0.00
 50331659                vmnic2         - DvsPortset-0        1.18    0.00
 50331660     Shadow of vmnic2       n/a DvsPortset-0        0.00    0.00
```

Route Based on Source MAC Hash

vSS + vDS





- vNIC pinning based on src MAC hash
- vNIC pinned to single physical NIC
- Even distribution of traffic if random MAC addresses
- Be careful with manual MAC address assignments
- Can go to one or multiple switches

Route Based on Source MAC Hash

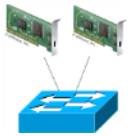
```
11:20:03am up  3:15, 567 worlds, 4 VMs, 5 vCPUs; CPU load average: 0.01, 0.01, 0
```

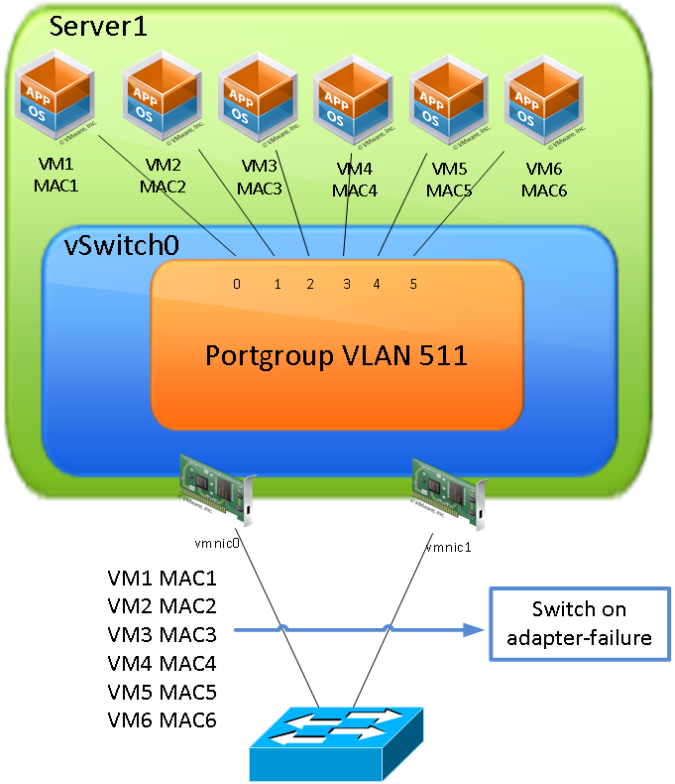
PORT-ID	USED-BY	TEAM-PNIC	DNAME	PKTTX/s	MbTX/s
33554433	Management	n/a	vSwitch0	0.00	0.00
33554434	vmnic3	-	vSwitch0	54.72	2.68
33554435	Shadow of vmnic3	n/a	vSwitch0	0.00	0.00
33554436	vmk0	vmnic3	vSwitch0	54.92	2.68
33554438	36042:VMware vCenter	vmnic3	vSwitch0	0.00	0.00
50331649	Management	n/a	DvsPortset-0	0.00	0.00
50331650	vmnic0	-	DvsPortset-0	1.18	0.00
50331651	Shadow of vmnic0	n/a	DvsPortset-0	0.00	0.00
50331659	vmnic2	-	DvsPortset-0	2.17	0.00
50331660	Shadow of vmnic2	n/a	DvsPortset-0	0.00	0.00
50331661	46658:VM-1.eth0	vmnic2*	DvsPortset-0	1.18	0.00
50331662	46777:VM-2.eth0	vmnic0*	DvsPortset-0	1.18	0.00
50331663	48692:VM-3.eth0	vmnic2*	DvsPortset-0	0.98	0.00

```
SV-5K-1# show mac address-table | grep aaa
* 511      aaaa.cafe.0001    dynamic   0          F    F  Eth101/1/32
* 511      aaaa.face.0002    dynamic   0          F    F  Eth101/1/31
* 511      aaaa.feed.0003    dynamic   0          F    F  Eth101/1/32
SV-5K-1#
```

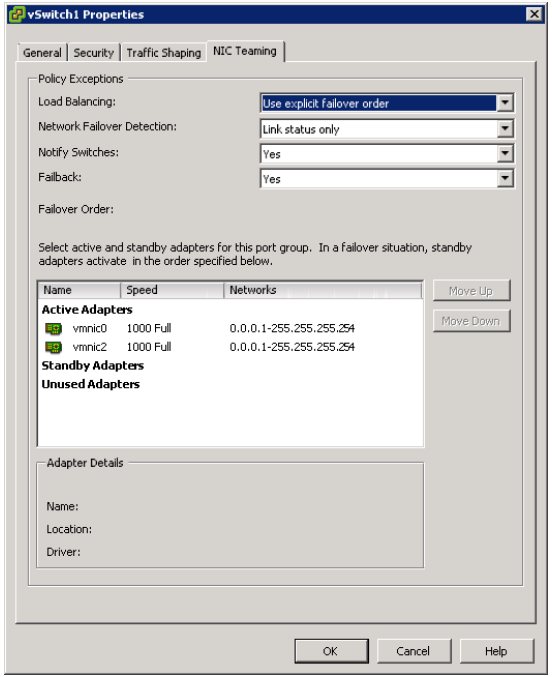
Use Explicit Failover Order

vSS + vDS





- Highest order uplink from the list of active adapters
- Can go to one or multiple switches



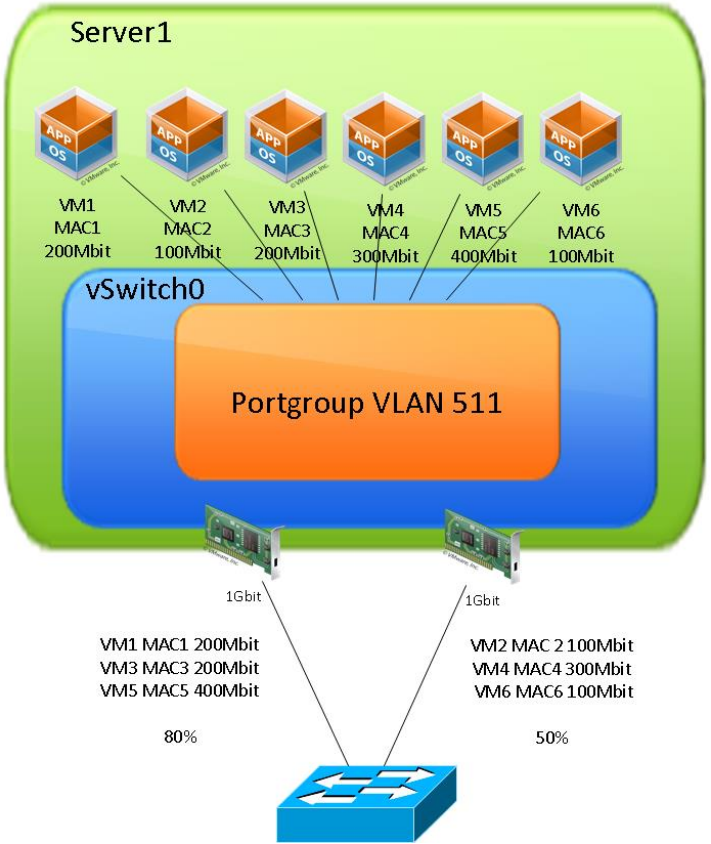
Use Explicit Failover Order

```
11:54:41am up  3:49, 567 worlds, 4 VMs, 5 vCPUs; CPU load average: 0.01, 0.01,

  PORT-ID              USED-BY  TEAM-PNIC DNAME              PKTTX/s  MbTX/s
 33554433           Management        n/a vSwitch0            0.00    0.00
 33554434                vmnic3          - vSwitch0           17.13    0.34
 33554435     Shadow of vmnic3        n/a vSwitch0            0.00    0.00
 33554436                 vmk0     vmnic3 vSwitch0           17.13    0.34
 33554438 36042:VMware vCenter     vmnic3 vSwitch0            0.00    0.00
 50331649           Management        n/a DvsPortset-0        0.00    0.00
 67108865           Management        n/a vSwitch1            0.00    0.00
 67108868                vmnic0          - vSwitch1            4.33    0.00
 67108869     Shadow of vmnic0        n/a vSwitch1            0.00    0.00
 67108870                vmnic2          - vSwitch1            0.00    0.00
 67108871     Shadow of vmnic2        n/a vSwitch1            0.00    0.00
 67108872          46658:VM-1     vmnic0 vSwitch1            0.98    0.00
 67108873          46777:VM-2     vmnic0 vSwitch1            2.36    0.00
 67108874          48692:VM-3     vmnic0 vSwitch1            0.98    0.00
```

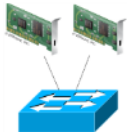
All MACs on single port

```
SV-5K-1# show int e101/1/31-32 | grep "is up"
Ethernet101/1/31 is up
Ethernet101/1/32 is up
SV-5K-1# show mac address-table | grep aaa
* 511      aaaa.cafe.0001    dynamic   10          F    F  Eth101/1/31
* 511      aaaa.face.0002    dynamic   10          F    F  Eth101/1/31
* 511      aaaa.feed.0003    dynamic   10          F    F  Eth101/1/31
SV-5K-1# show mac address-table interface e101/1/32
SV-5K-1#
```

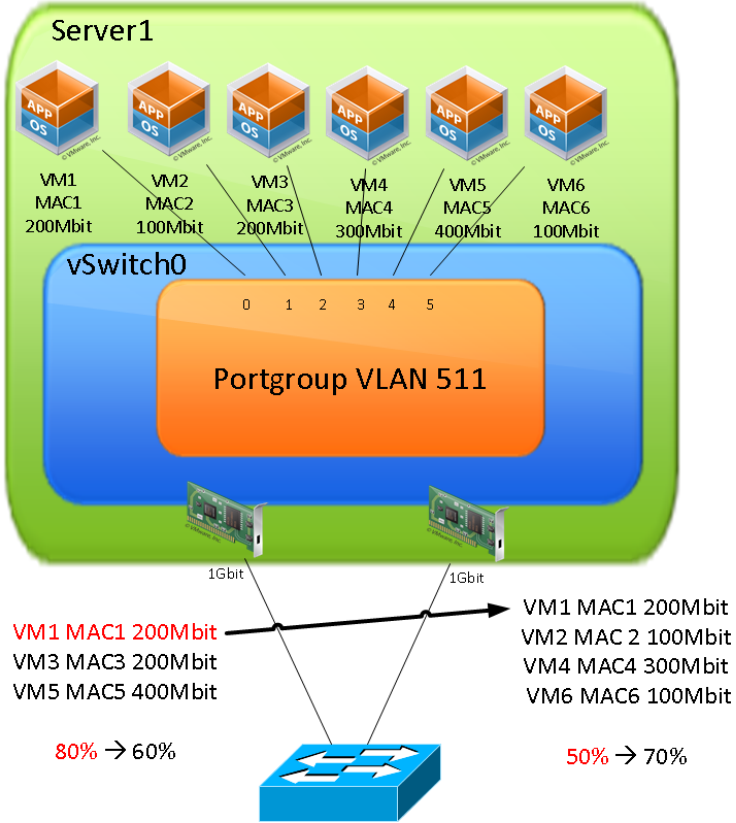
Route Based on Physical NIC Load (vDS)

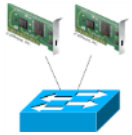


- Initial vNIC placement uses “Route based on originating virtual port”
- If pNIC RX or TX > 75% → rebalance
- VMkernel checks load every 30s
- MAC moves on switch layer
- Can go to single or multiple switches



Route Based on Physical NIC Load (vDS)





- Initial vNIC placement uses “Route based on originating virtual port”
- If pNIC RX or TX > 75% → rebalance
- VMkernel checks load every 30s
- MAC moves on switch layer
- Can go to single or multiple switches

Route Based on Physical NIC Load (vDS)

```
11:27:00am up  3:21, 565 worlds, 4 VMs, 5 vCPUs; CPU load average: 0.01, 0.01, 0.01

  PORT-ID              USED-BY  TEAM-PNIC DNAME              PKTTX/s  MbTX/s    PKTRX/s  MbRX/s
 33554433           Management        n/a vSwitch0            0.00    0.00       0.00    0.00
 33554434                vmnic3          - vSwitch0          309.68   51.35    2405.21    1.26
 33554435     Shadow of vmnic3        n/a vSwitch0            0.00    0.00       0.00    0.00
 33554436                 vmk0     vmnic3 vSwitch0          309.68   51.35    2368.92    1.21
 33554438 36042:VMware vCenter     vmnic3 vSwitch0            0.00    0.00       8.28    0.01
 50331649           Management        n/a DvsPortset-0        0.00    0.00       0.00    0.00
 50331650                vmnic0          - DvsPortset-0     1894.74    0.87    8030.26   84.93
 50331651     Shadow of vmnic0        n/a DvsPortset-0        0.00    0.00       0.00    0.00
 50331659                vmnic2          - DvsPortset-0        1.97    0.00      35.11    0.05
 50331660     Shadow of vmnic2        n/a DvsPortset-0        0.00    0.00       0.00    0.00
 50331661      46658:VM-1.eth0     vmnic0 DvsPortset-0     1894.74    0.78    8005.60   84.90
 50331662      46777:VM-2.eth0     vmnic2 DvsPortset-0        0.99    0.00       9.27    0.01
 50331663      48692:VM-3.eth0     vmnic2 DvsPortset-0        0.99    0.00       9.27    0.01
```

```
SV-5K-1# show mac address-table | grep aaa
* 511       aaaa.cafe.0001    dynamic    0          F    F  Eth101/1/31
* 511       aaaa.face.0002    dynamic    0          F    F  Eth101/1/32
* 511       aaaa.feed.0003    dynamic    0          F    F  Eth101/1/32
SV-5K-1#
```

VMware Switch Independent

Switch Configuration

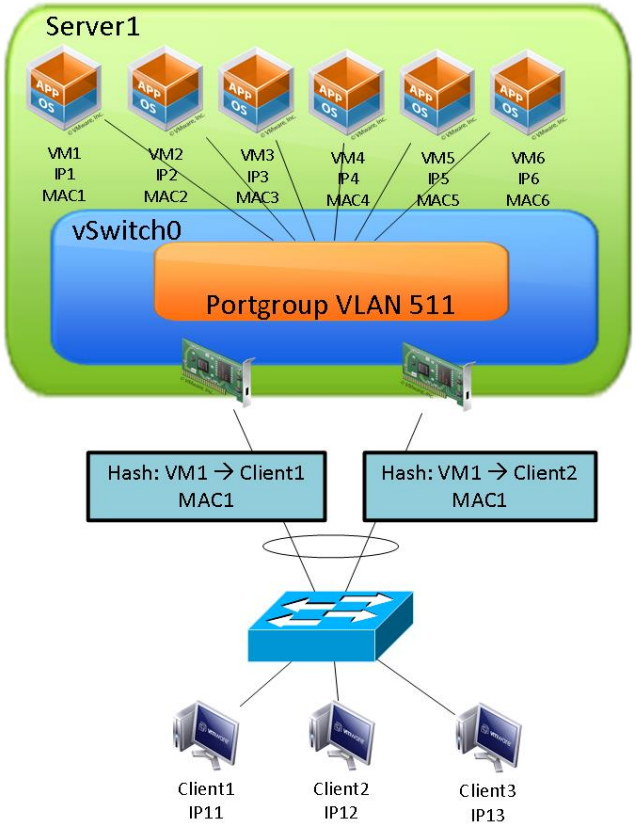
For Your Reference

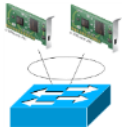
```
interface Ethernet101/1/31-32
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree port type edge trunk
```

Used for:

- Route based on originating virtual port
- Route based on source MAC hash
- Use explicit failover order
- Route based on physical NIC load (vDS)

Route Based on IP Hash





- IP source and destination hash
- Evenness of traffic distribution depends on TCP/IP attributes of flows
- Requires static 802.3ad port-channel (mode on)
- Connects to single port-channel

Route Based on IP Hash

Misconfiguration Example

```
SV-5K-1# show logging | grep mac
2014 Jan  3 13:38:46 SV-5K-1 %FWM-2-STM_LOOP_DETECT: Loops detected in the network for mac aaaa.cafe.0001 among ports Eth101/1/32 and Eth101/1/31 vlan 511
2014 Jan  3 13:42:07 SV-5K-1 %FWM-2-STM_LOOP_DETECT: Loops detected in the network for mac aaaa.cafe.0001 among ports Eth101/1/32 and Eth101/1/31 vlan 511
2014 Jan  3 13:45:27 SV-5K-1 %FWM-2-STM_LOOP_DETECT: Loops detected in the network for mac aaaa.cafe.0001 among ports Eth101/1/31 and Eth101/1/32 vlan 511
2014 Jan  3 13:48:37 SV-5K-1 %FWM-2-STM_LOOP_DETECT: Loops detected in the network for mac aaaa.cafe.0001 among ports Eth101/1/32 and Eth101/1/31 vlan 511
2014 Jan  3 13:51:53 SV-5K-1 %FWM-2-STM_LOOP_DETECT: Loops detected in the network for mac aaaa.cafe.0001 among ports Eth101/1/31 and Eth101/1/32 vlan 511
2014 Jan  3 13:54:54 SV-5K-1 %FWM-2-STM_LOOP_DETECT: Loops detected in the network for mac aaaa.cafe.0001 among ports Eth101/1/32 and Eth101/1/31 vlan 511
2014 Jan  3 13:58:01 SV-5K-1 %FWM-2-STM_LOOP_DETECT: Loops detected in the network for mac aaaa.cafe.0001 among ports Eth101/1/32 and Eth101/1/31 vlan 511
2014 Jan  3 14:01:03 SV-5K-1 %FWM-2-STM_LOOP_DETECT: Loops detected in the network for mac aaaa.cafe.0001 among ports Eth101/1/31 and Eth101/1/32 vlan 511
SV-5K-1#
```

```
SV-5K-1# show mac address-table | grep aaaa.cafe.0001
* 511      aaaa.cafe.0001    dynamic   0          F    F  Eth101/1/31
SV-5K-1# show mac address-table | grep aaaa.cafe.0001
* 511      aaaa.cafe.0001    dynamic   0          F    F  Eth101/1/32
SV-5K-1# show mac address-table | grep aaaa.cafe.0001
* 511      aaaa.cafe.0001    dynamic   0          F    F  Eth101/1/31
SV-5K-1# show mac address-table | grep aaaa.cafe.0001
* 511      aaaa.cafe.0001    dynamic   10         F    F  Eth101/1/32
SV-5K-1# show mac address-table | grep aaaa.cafe.0001
* 511      aaaa.cafe.0001    dynamic   0          F    F  Eth101/1/32
SV-5K-1# show mac address-table | grep aaaa.cafe.0001
* 511      aaaa.cafe.0001    dynamic   0          F    F  Eth101/1/31
```

VMware Switch Dependent

Route Based on IP Hash – Switch Configuration

```
interface Ethernet101/1/31-32
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree port type edge trunk
  channel-group 300

interface Port-Channel300
  switchport mode trunk
  switchport trunk allowed vlan 511
```

```
SV-5K-1(config-if)# channel-group 300 mode ?
  active     Set channeling mode to ACTIVE
  on         Set channeling mode to ON
  passive    Set channeling mode to PASSIVE

SV-5K-1(config-if)# channel-group 300 mode
```

Route Based on IP Hash

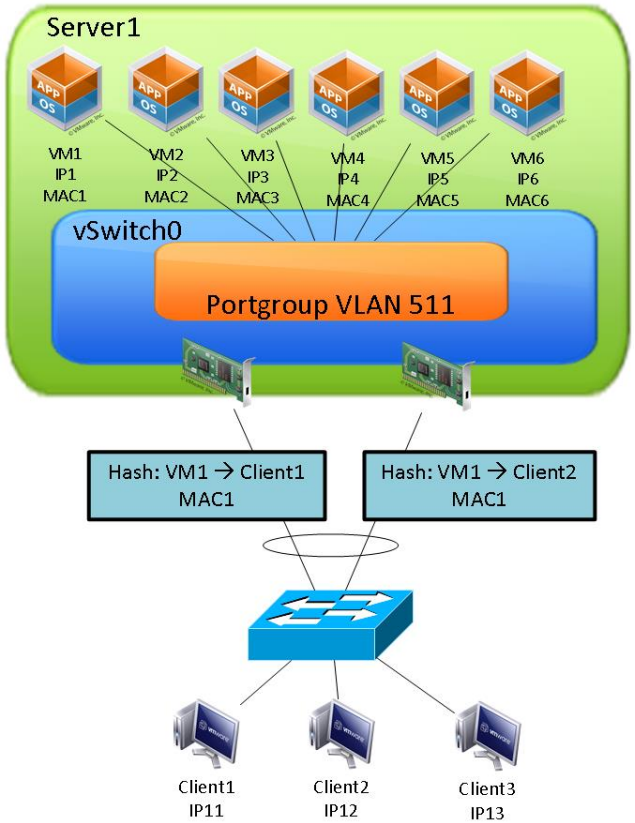
```
11:01:40am up  2:56, 567 worlds, 4 VMs, 5 vCPUs; CPU load average: 0.01, 0.01,
```

PORT-ID	USED-BY	TEAM-PNIC	DNAME	PKTTX/s	MbTX/s
33554433	Management	n/a	vSwitch0	0.00	0.00
33554434	vmnic3	-	vSwitch0	12.59	0.15
33554435	Shadow of vmnic3	n/a	vSwitch0	0.00	0.00
33554436	vmk0	vmnic3	vSwitch0	12.59	0.15
33554438	36042:VMware vCenter	vmnic3	vSwitch0	0.00	0.00
50331649	Management	n/a	DvsPortset-0	0.00	0.00
50331650	vmnic0	-	DvsPortset-0	127.26	0.12
50331651	Shadow of vmnic0	n/a	DvsPortset-0	0.00	0.00
50331659	vmnic2	-	DvsPortset-0	65.89	0.07
50331660	Shadow of vmnic2	n/a	DvsPortset-0	0.00	0.00
50331661	46658:VM-1.eth0	all(2)	DvsPortset-0	191.58	0.18
50331662	46777:VM-2.eth0	all(2)	DvsPortset-0	1.38	0.00
50331663	48692:VM-3.eth0	all(2)	DvsPortset-0	1.77	0.00

```
SV-5K-1# show mac address-table | grep aaaa.cafe.0001
* 511      aaaa.cafe.0001    dynamic   0          F    F  Po301
SV-5K-1#
```

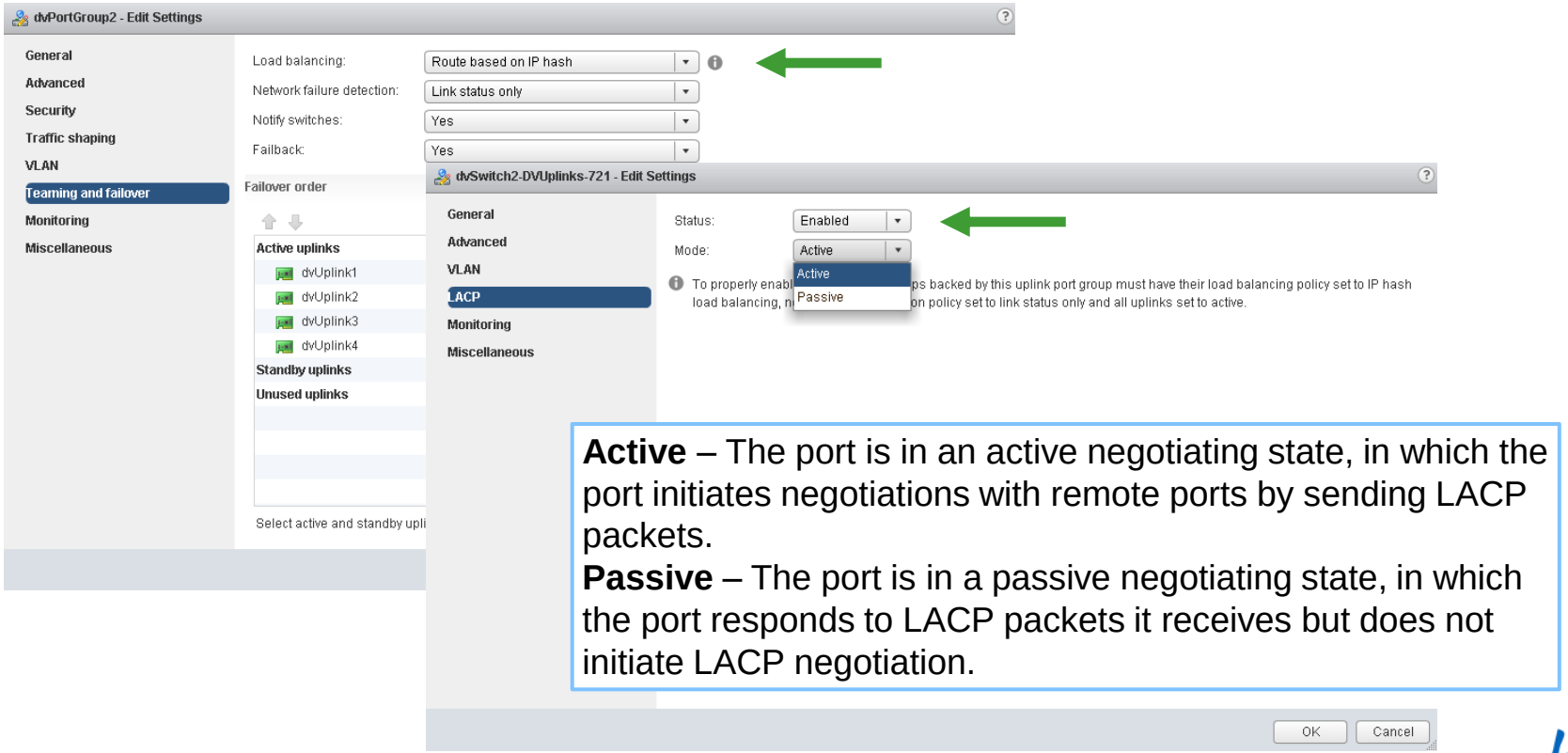
```
--------------------------------------------------------------------------------
Group Port-       Type      Protocol  Member Ports
      Channel
--------------------------------------------------------------------------------
301   Po301(SU)   Eth       NONE      Eth101/1/31(P)  Eth101/1/32(P)
SV-5K-1#
```

Route Based on IP Hash + LACP (vDS)



- IP source and destination hash
- Evenness of traffic distribution depends on TCP/IP attributes of flows
- Requires dynamic 802.3ad port-channel (mode active/passive)

Route Based on IP Hash + LACP (vDS)



Active – The port is in an active negotiating state, in which the port initiates negotiations with remote ports by sending LACP packets.

Passive – The port is in a passive negotiating state, in which the port responds to LACP packets it receives but does not initiate LACP negotiation.

Route Based on IP Hash + LACP (vDS)

```
11:12:31am up  3:07, 565 worlds, 4 VMs, 5 vCPUs; CPU load average: 0.01, 0.01,

  PORT-ID              USED-BY  TEAM-PNIC DNAME              PKTTX/s  MbTX/s
 33554433           Management        n/a vSwitch0             0.00    0.00
 33554434                vmnic3          - vSwitch0            36.21    2.12
 33554435     Shadow of vmnic3        n/a vSwitch0             0.00    0.00
 33554436                 vmk0     vmnic3 vSwitch0            36.21    2.12
 33554438 36042:VMware vCenter     vmnic3 vSwitch0             0.00    0.00
 50331649           Management        n/a DvsPortset-0         0.00    0.00
 50331650                vmnic0          - DvsPortset-0         2.36    0.00
 50331651     Shadow of vmnic0        n/a DvsPortset-0         0.00    0.00
 50331659                vmnic2          - DvsPortset-0         0.98    0.00
 50331660     Shadow of vmnic2        n/a DvsPortset-0         0.00    0.00
 50331661      46658:VM-1.eth0    all(2) DvsPortset-0         1.18    0.00
 50331662      46777:VM-2.eth0    all(2) DvsPortset-0         0.98    0.00
 50331663      48692:VM-3.eth0    all(2) DvsPortset-0         1.18    0.00
 50331664        LACP_MgmtPort        n/a DvsPortset-0         0.00    0.00
```

Extra LACP_MgmtPort

```
SV-5K-1# show mac address-table | grep aaa
* 511      aaaa.cafe.0001    dynamic   0          F     F   Po300
* 511      aaaa.face.0002    dynamic   0          F     F   Po300
* 511      aaaa.feed.0003    dynamic   0          F     F   Po300
SV-5K-1#
```

```
--------------------------------------------------------------------------------
Group Port-       Type     Protocol  Member Ports
      Channel
--------------------------------------------------------------------------------
300   Po300(SU)   Eth      LACP      Eth101/1/31(P)  Eth101/1/32(P)
SV-5K-1#
```

VMware Switch Dependent

Route based on IP hash + LACP (vDS) – Switch Configuration

```
interface Ethernet101/1/31-32
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree port type edge trunk
  channel-group 300 mode active

interface Port-Channel300
  switchport mode trunk
  switchport trunk allowed vlan 511
```

```
SV-5K-1(config-if)# channel-group 300 mode ?
  active     Set channeling mode to ACTIVE
  on         Set channeling mode to ON
  passive    Set channeling mode to PASSIVE

SV-5K-1(config-if)# channel-group 300 mode
```

ESXi Switch	Active	Passive
Active	Yes	Yes
Passive	Yes	No

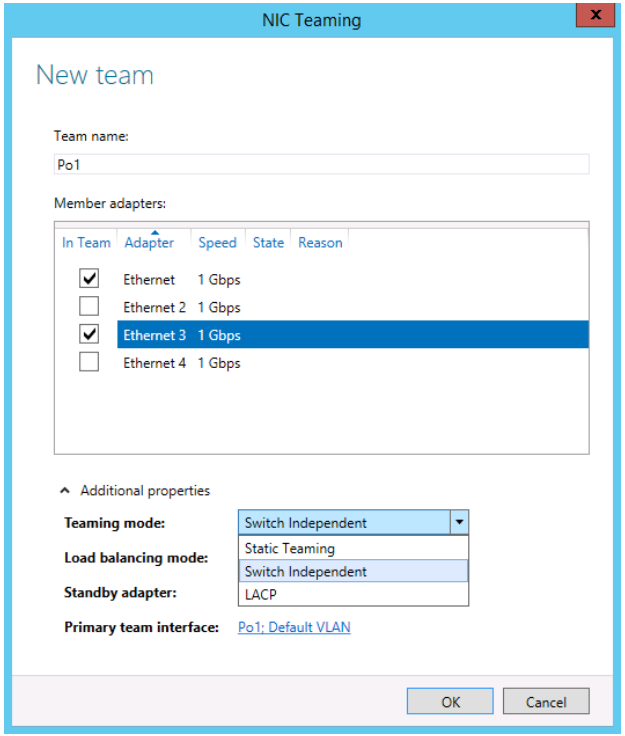
VMware Conclusion

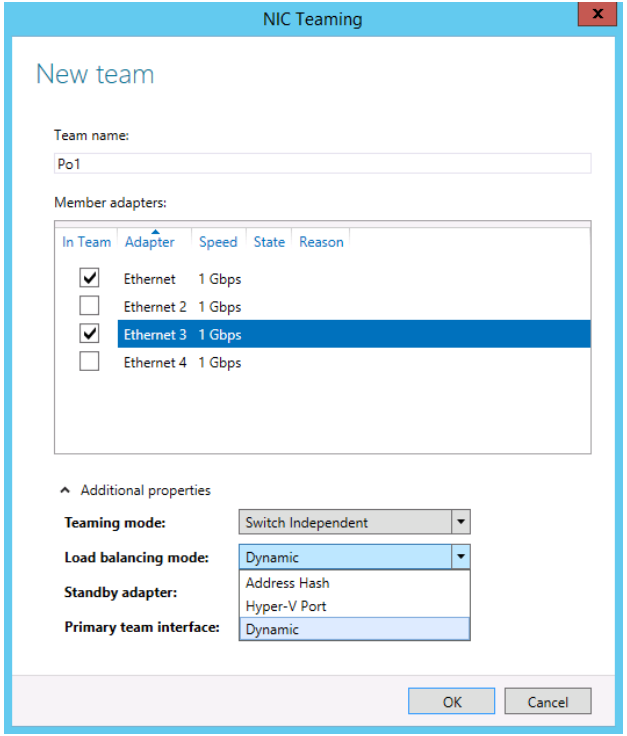
- vSphere Standard Switch (vSS)
  - Switch independent
    - Route based on originating virtual port
  - Switch dependent
    - Route based on IP hash
- vSphere Distributed Switch (vDS)
  - Switch independent
    - Route based on physical NIC load
  - Switch dependent
    - Route based on IP hash + LACP

Hypervisor Overview
Microsoft Hyper-V 3.0
Windows 2012 R2

Uplink Options

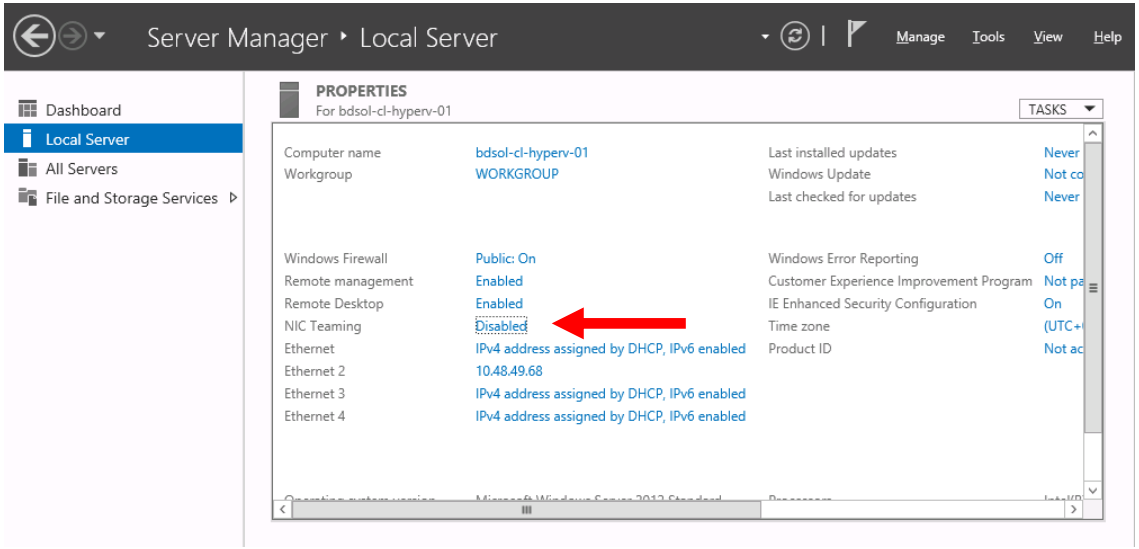
- Single physical NIC or Teaming



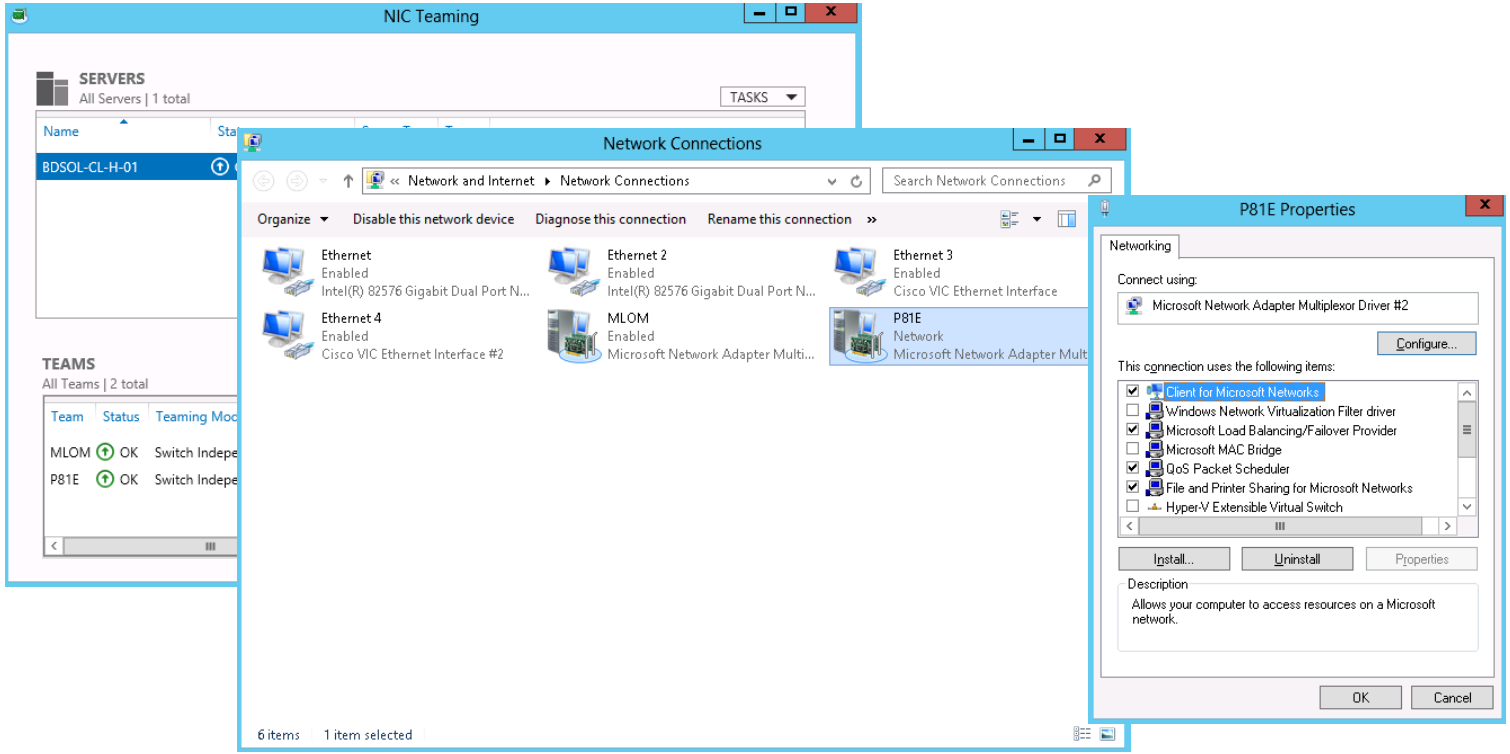


Enabling NIC Teaming

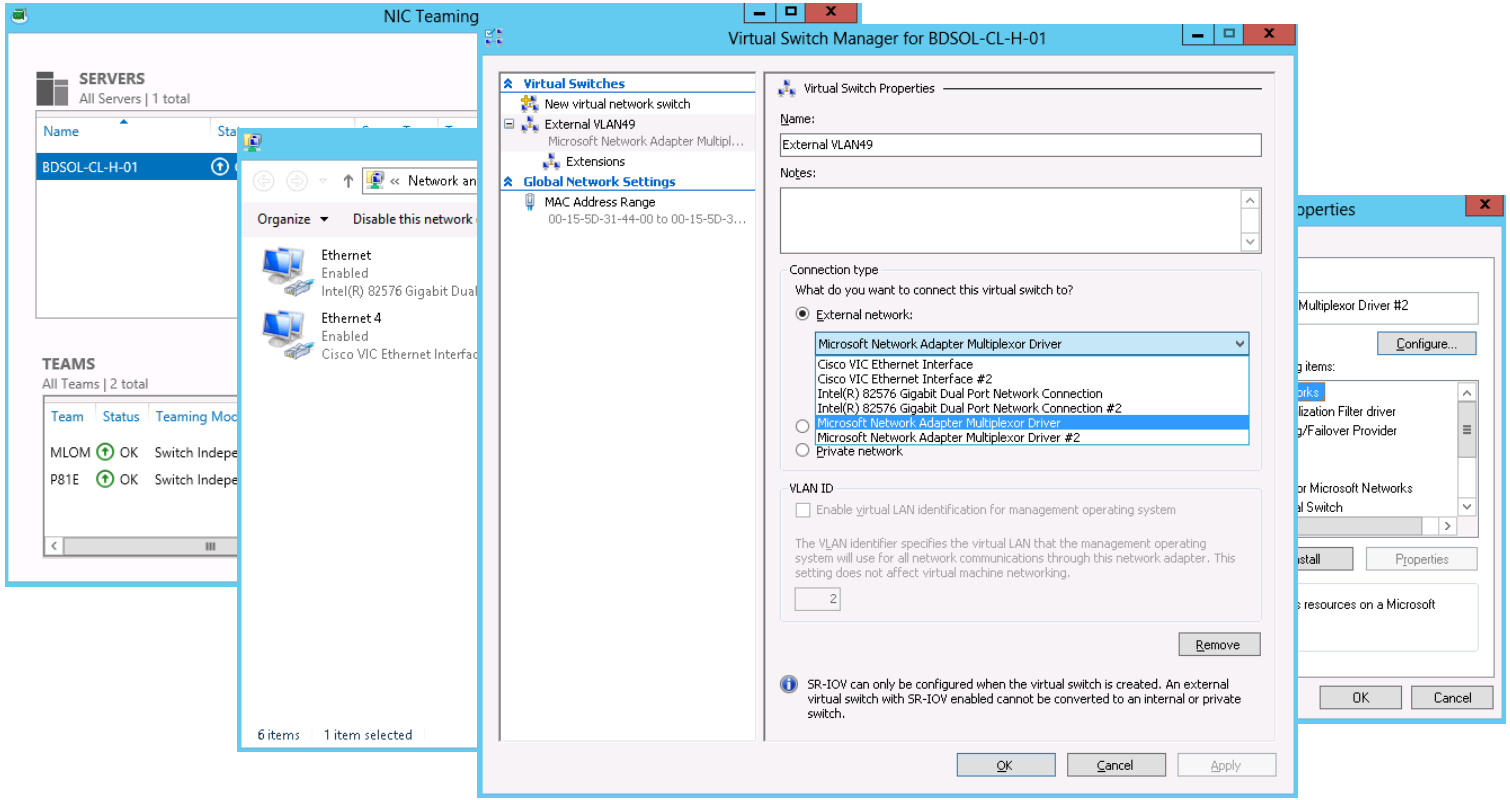
- Pre-Windows 2012 -Teaming provided by vendor device drivers
- Windows 2012 - Native teaming support



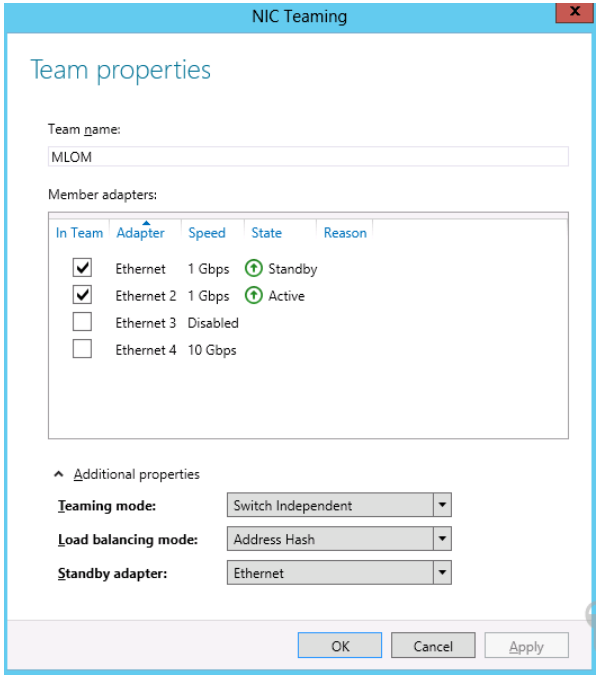
Microsoft Network Adapter Multiplexor Driver



Microsoft Network Adapter Multiplexor Driver

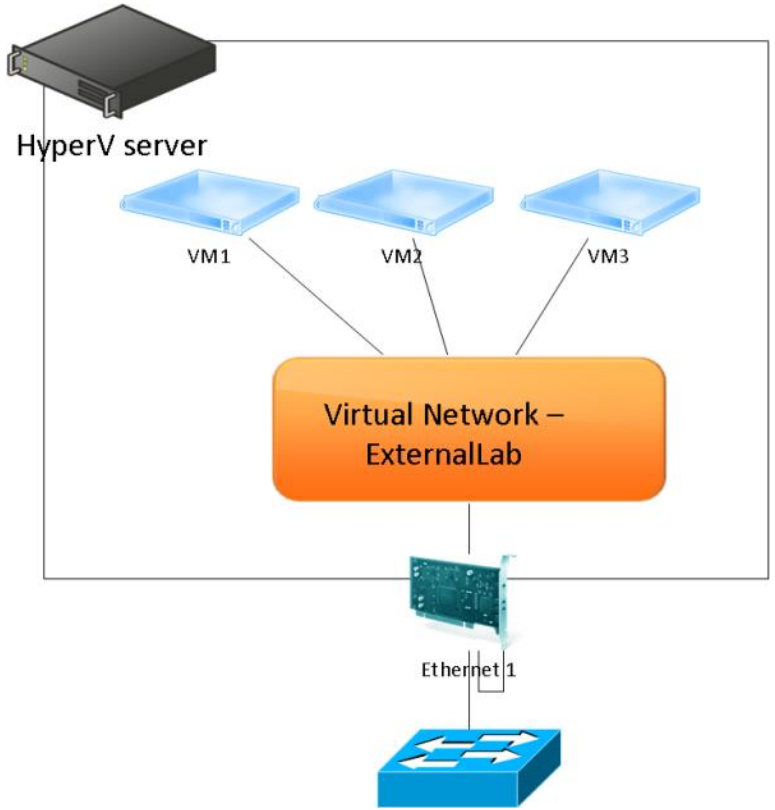


Standby Adapter



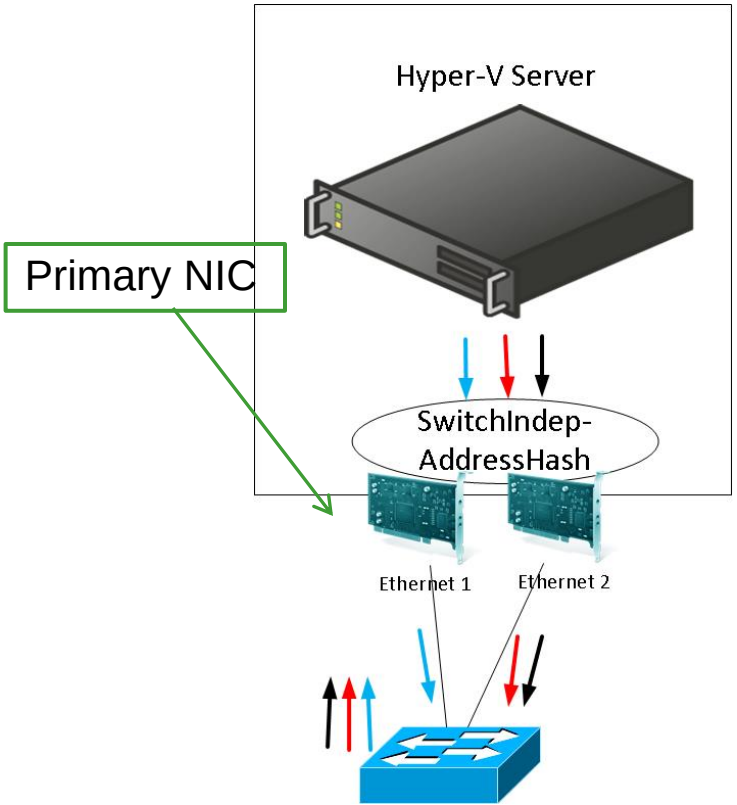

```
SV-5K-1# show int e101/1/31-32 | grep "is up"
Ethernet101/1/31 is up
Ethernet101/1/32 is up
SV-5K-1# show mac address-table | grep aaa
* 511      aaaa.cafe.0001    dynamic   10         F     F  Eth101/1/31
* 511      aaaa.face.0002    dynamic   10         F     F  Eth101/1/31
* 511      aaaa.feed.0003    dynamic   10         F     F  Eth101/1/31
SV-5K-1# show mac address-table interface e101/1/32
SV-5K-1#
```

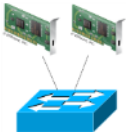
Single Physical NIC



- Normal server connection
- Access or trunk interface
- Multiple MACs on switch port

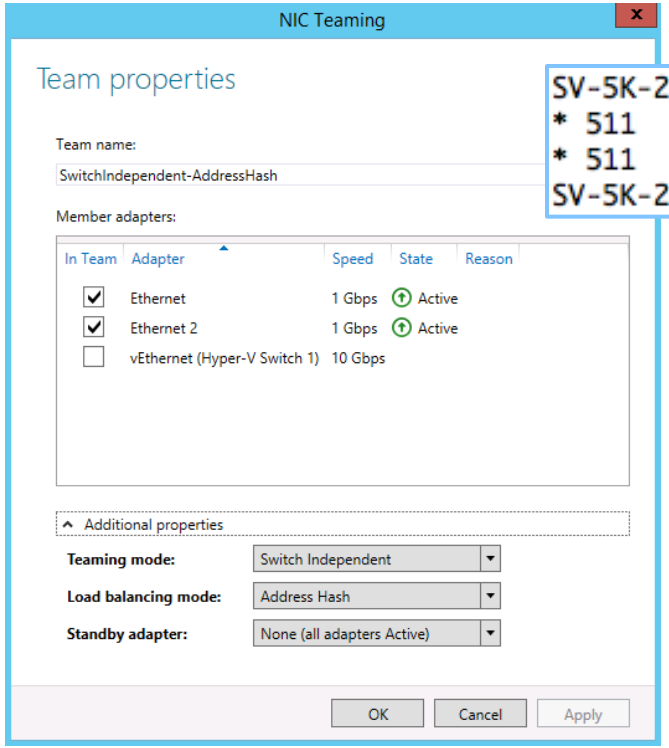
Switch Independent – Address Hash





- MAC/IP/Port hash for outbound traffic
- Outbound frame uses src MAC of egress NIC
- ARP replies contain MAC of primary NIC
- Primary NIC receives inbound traffic
- Can go to single or multiple switches

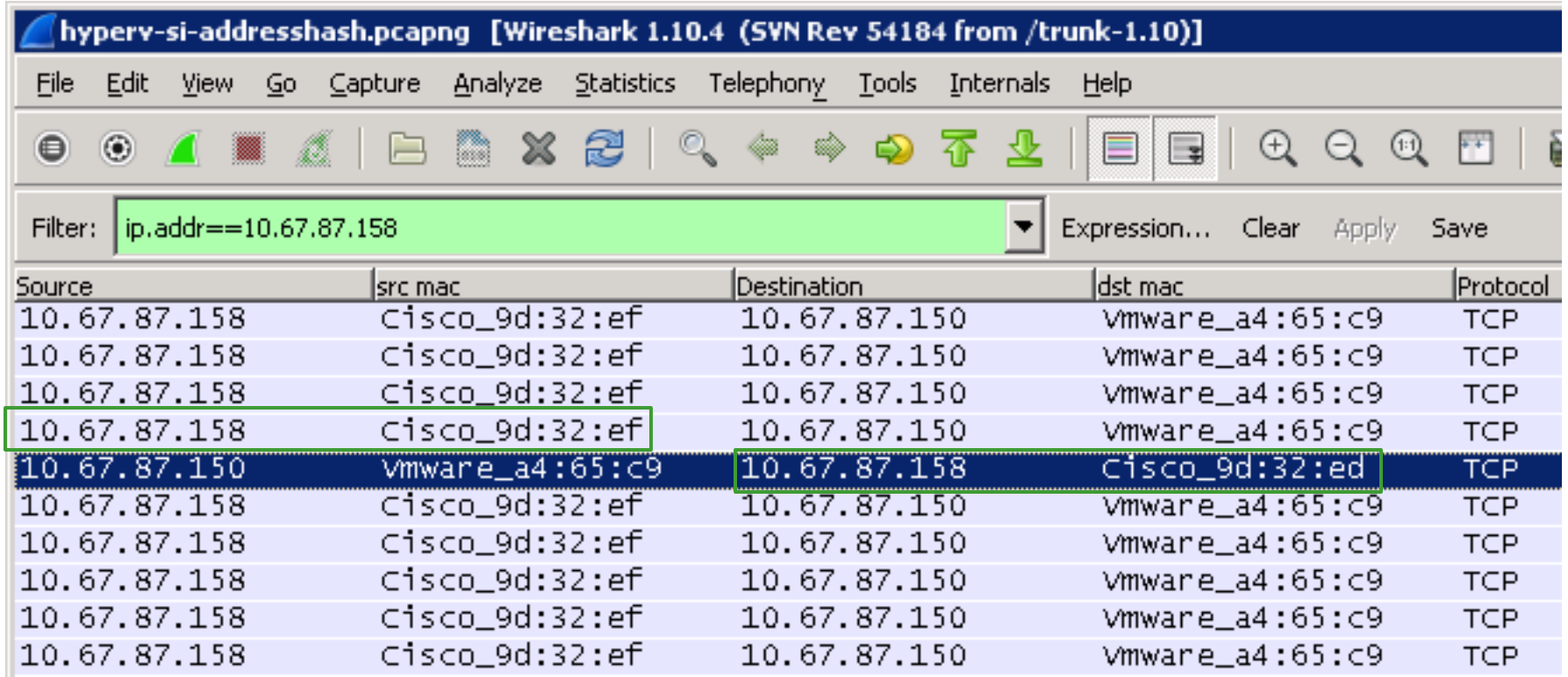
Switch Independent – Address Hash



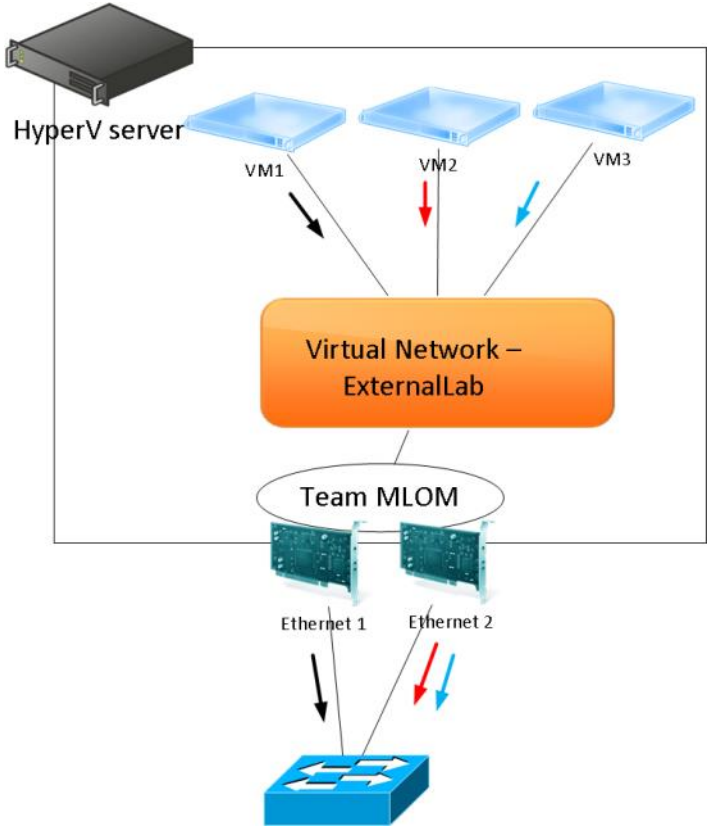

```
SV-5K-2# show mac address-table vlan 511 | egrep Eth101/1/31|Eth101/1/32
* 511      503d.e59d.32ed    dynamic   10          F    F  Eth101/1/31
* 511      503d.e59d.32ef    dynamic   10          F    F  Eth101/1/32
SV-5K-2#
```

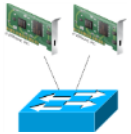
```
Ethernet adapter SwitchIndependent-AddressHash:

   Connection-specific DNS Suffix  . :
   Description . . . . . . . . . . . : Microsoft Network Adapte
   Physical Address. . . . . . . . . : 50-3D-E5-9D-32-ED
   DHCP Enabled. . . . . . . . . . . : No
   Autoconfiguration Enabled . . . . : Yes
   IPv4 Address. . . . . . . . . . . : 10.67.87.158(Preferred)
   Subnet Mask . . . . . . . . . . . : 255.255.255.0
   Default Gateway . . . . . . . . . : 10.67.87.1
   DNS Servers . . . . . . . . . . . : 64.104.123.245
   NetBIOS over Tcpip. . . . . . . . : Enabled
```

Switch Independent – Address Hash



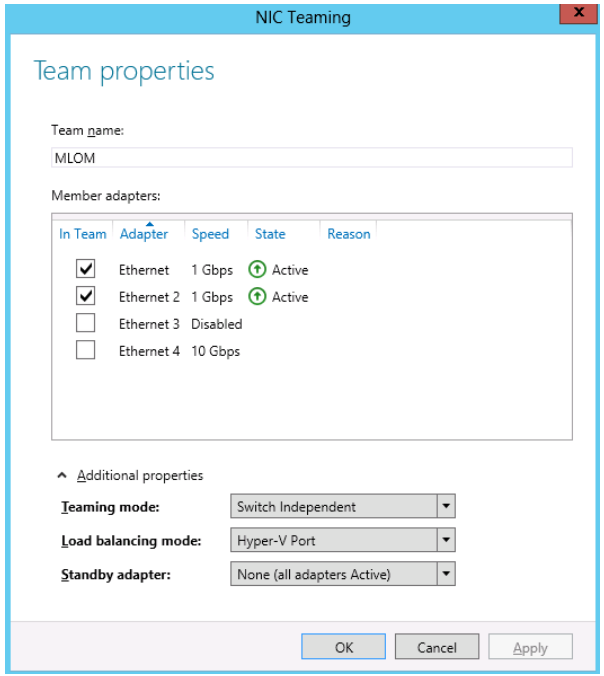
Switch Independent - Hyper-V Port





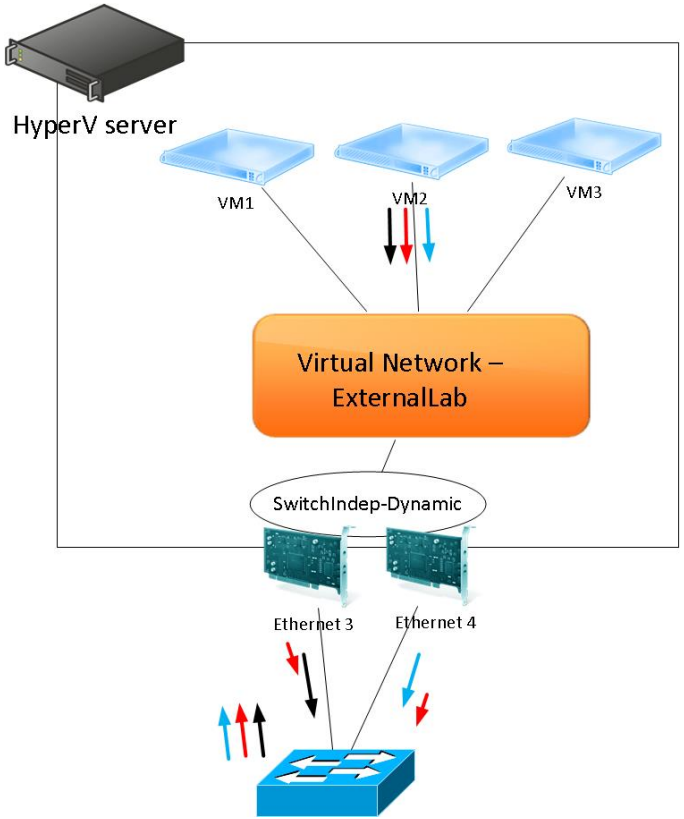
- vNICs are pinned to single team member
- vNICs pinned in round robin
- Can go to single or multiple switches

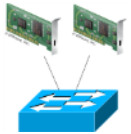
Switch Independent - Hyper-V Port




```
SV-5K-1# show mac address-table | grep aaa
* 511       aaaa.cafe.0001    dynamic   0          F     F  Eth101/1/31
* 511       aaaa.face.0002    dynamic   0          F     F  Eth101/1/31
* 511       aaaa.feed.0003    dynamic   0          F     F  Eth101/1/32
SV-5K-1#
```

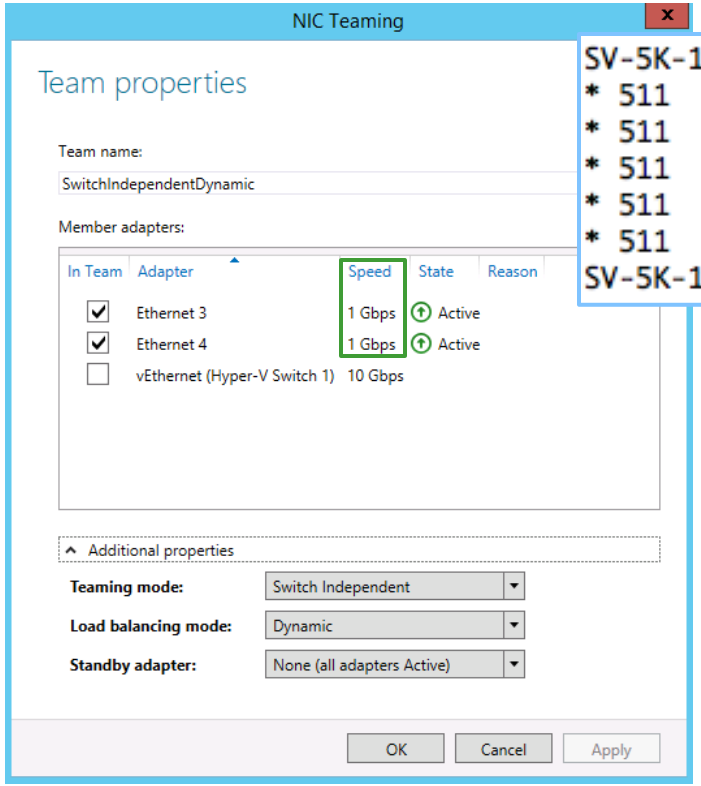
Switch Independent - Dynamic





- Outbound flows redistributed to optimise pNIC bandwidth utilisation
- Algorithm can move flows between pNICs w/o packet reordering (flowlets)
- Inbound traffic works like Hyper-V Port mode
- Can go to single or multiple switches

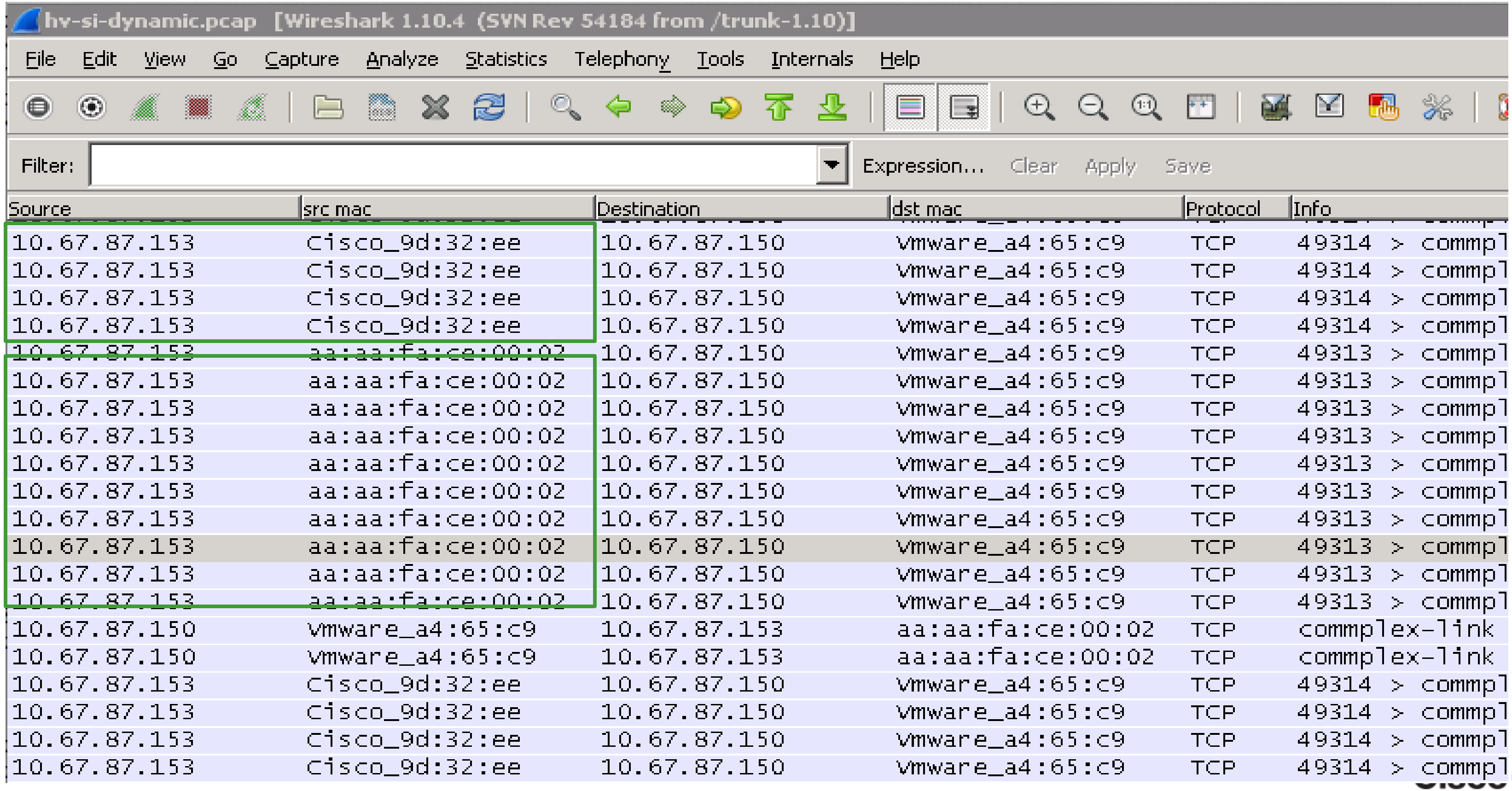
Switch Independent - Dynamic




```
SV-5K-1# show mac address-table | egrep Eth101/1/31|Eth101/1/32
* 511      503d.e59d.32ec    dynamic   180         F    F  Eth101/1/31
* 511      503d.e59d.32ee    dynamic   10          F    F  Eth101/1/32
* 511      aaaa.cafe.0001    dynamic   0           F    F  Eth101/1/31
* 511      aaaa.face.0002    dynamic   0           F    F  Eth101/1/31
* 511      aaaa.feed.0003    dynamic   0           F    F  Eth101/1/32
SV-5K-1#
```

```
[240]   5.0- 6.0 sec  5.59 MBytes  46.9 Mbits/sec
[232]   5.0- 6.0 sec  4.06 MBytes  34.1 Mbits/sec
[224]   5.0- 6.0 sec  32.8 MBytes   275 Mbits/sec
[216]   5.0- 6.0 sec  34.8 MBytes   292 Mbits/sec
[208]   5.0- 6.0 sec  10.7 MBytes  89.6 Mbits/sec
[200]   5.0- 6.0 sec  86.9 MBytes   729 Mbits/sec
[192]   5.0- 6.0 sec  36.0 MBytes   302 Mbits/sec
[248]   5.0- 6.0 sec  6.41 MBytes  53.7 Mbits/sec
[SUM]   5.0- 6.0 sec   217 MBytes  1.82 Gbits/sec
```

Switch Independent - Dynamic



Source	src mac	Destination	dst mac	Protocol	Info
10.67.87.153	Cisco_9d:32:ee	10.67.87.150	Vmware_a4:65:c9	TCP	49314 > commpl
10.67.87.153	Cisco_9d:32:ee	10.67.87.150	Vmware_a4:65:c9	TCP	49314 > commpl
10.67.87.153	Cisco_9d:32:ee	10.67.87.150	Vmware_a4:65:c9	TCP	49314 > commpl
10.67.87.153	Cisco_9d:32:ee	10.67.87.150	Vmware_a4:65:c9	TCP	49314 > commpl
10.67.87.153	aa:aa:fa:ce:00:02	10.67.87.150	Vmware_a4:65:c9	TCP	49313 > commpl
10.67.87.153	aa:aa:fa:ce:00:02	10.67.87.150	Vmware_a4:65:c9	TCP	49313 > commpl
10.67.87.153	aa:aa:fa:ce:00:02	10.67.87.150	Vmware_a4:65:c9	TCP	49313 > commpl
10.67.87.153	aa:aa:fa:ce:00:02	10.67.87.150	Vmware_a4:65:c9	TCP	49313 > commpl
10.67.87.153	aa:aa:fa:ce:00:02	10.67.87.150	Vmware_a4:65:c9	TCP	49313 > commpl
10.67.87.153	aa:aa:fa:ce:00:02	10.67.87.150	Vmware_a4:65:c9	TCP	49313 > commpl
10.67.87.153	aa:aa:fa:ce:00:02	10.67.87.150	Vmware_a4:65:c9	TCP	49313 > commpl
10.67.87.153	aa:aa:fa:ce:00:02	10.67.87.150	Vmware_a4:65:c9	TCP	49313 > commpl
10.67.87.153	aa:aa:fa:ce:00:02	10.67.87.150	Vmware_a4:65:c9	TCP	49313 > commpl
10.67.87.153	aa:aa:fa:ce:00:02	10.67.87.150	Vmware_a4:65:c9	TCP	49313 > commpl
10.67.87.150	Vmware_a4:65:c9	10.67.87.153	aa:aa:fa:ce:00:02	TCP	commplex-link
10.67.87.150	Vmware_a4:65:c9	10.67.87.153	aa:aa:fa:ce:00:02	TCP	commplex-link
10.67.87.153	Cisco_9d:32:ee	10.67.87.150	Vmware_a4:65:c9	TCP	49314 > commpl
10.67.87.153	Cisco_9d:32:ee	10.67.87.150	Vmware_a4:65:c9	TCP	49314 > commpl
10.67.87.153	Cisco_9d:32:ee	10.67.87.150	Vmware_a4:65:c9	TCP	49314 > commpl
10.67.87.153	Cisco_9d:32:ee	10.67.87.150	Vmware_a4:65:c9	TCP	49314 > commpl

Microsoft Switch Independent

Switch Configuration

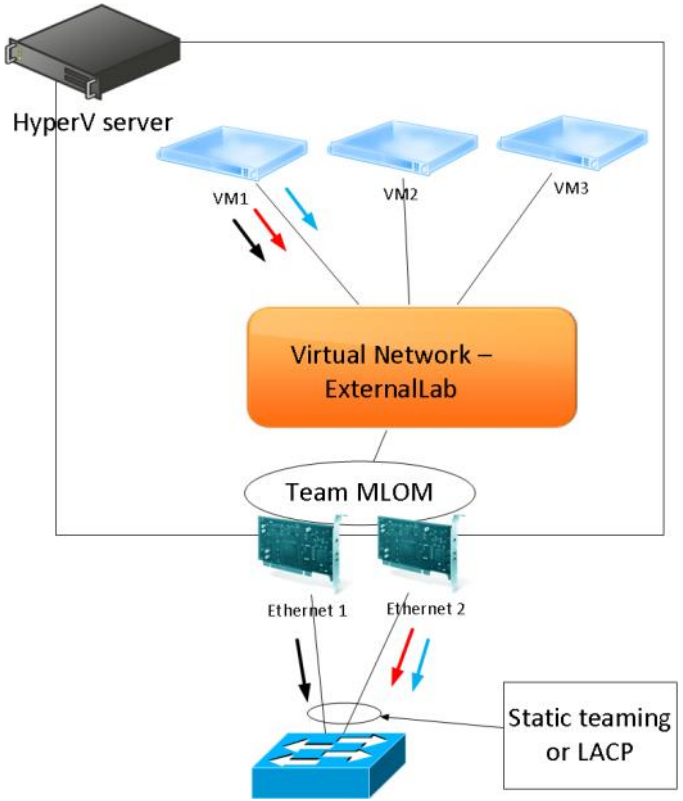
For Your Reference

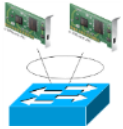
```
interface Ethernet101/1/31-32
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree port type edge trunk
```

Used for:

- **Teaming mode**: Switch Independent
- **Load balancing modes**:
  - Address Hash
  - Hyper-V Port
  - Dynamic

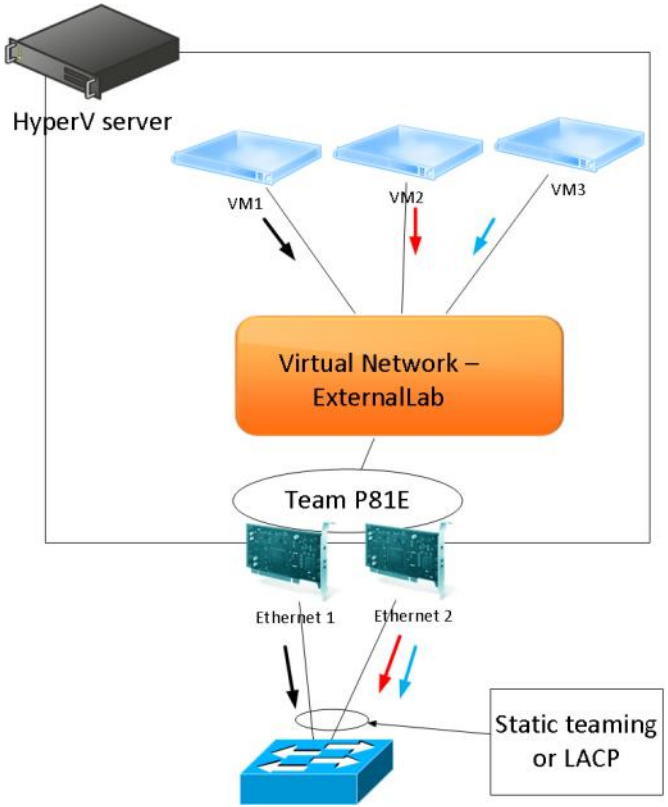
Switch Dependent – Address Hash

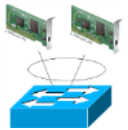




- Outbound traffic spread across all active members based on hash
- Inbound traffic distributed by switch load-balancing algorithm
- Team connects to single port-channel

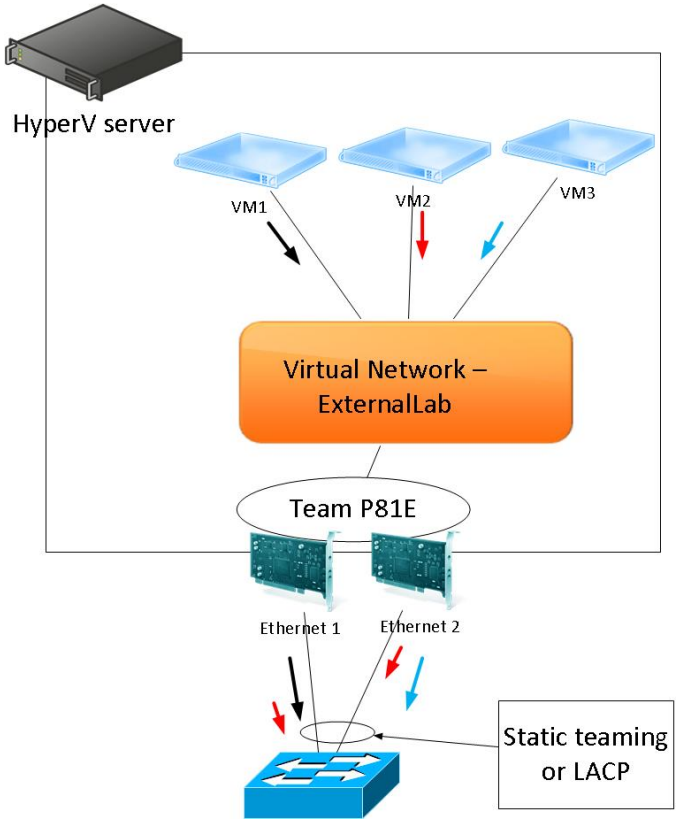
Switch Dependent - Hyper-V Port

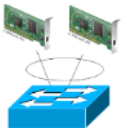




- vNICs are pinned to single team member
- vNICs pinned in round robin
- Inbound traffic distributed by switch load-balancing algorithm
- Team connects to single port-channel

Switch Dependent - Dynamic

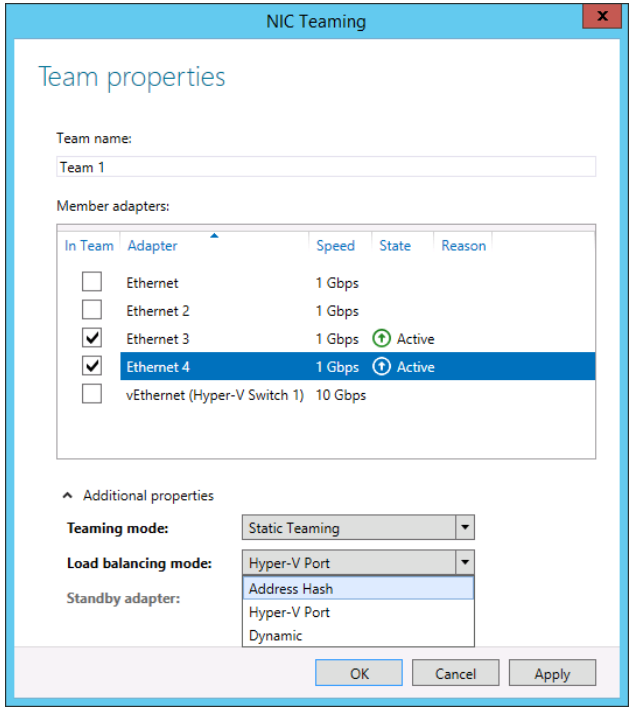


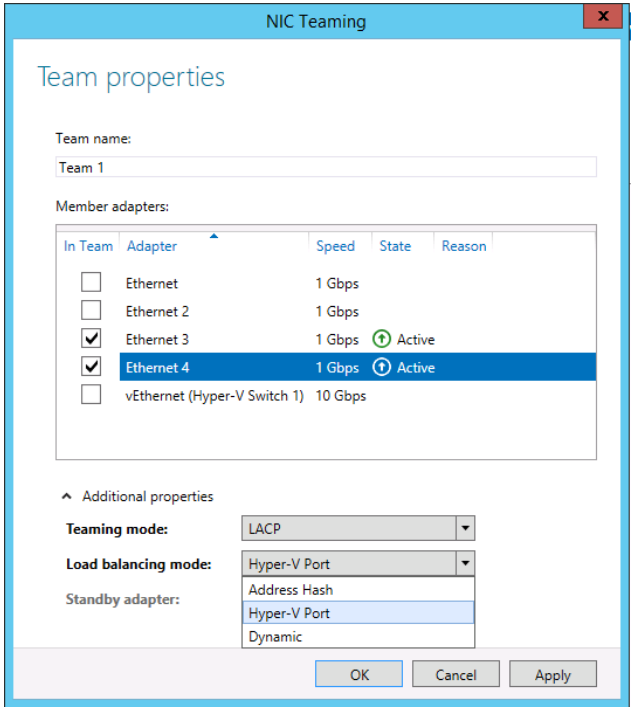


- Outbound flows redistributed to optimise pNIC bandwidth utilisation
- Algorithm can move flows between pNICs w/o packet reordering (flowlets)
- Inbound traffic distributed by switch load-balancing algorithm

Switch Dependent

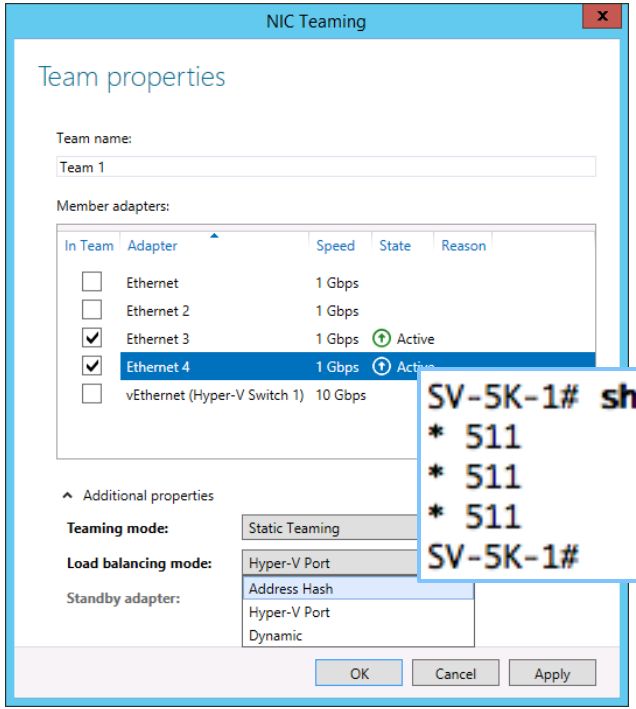
Load Balancing Mode: Address Hash, Hyper-V Port and Dynamic

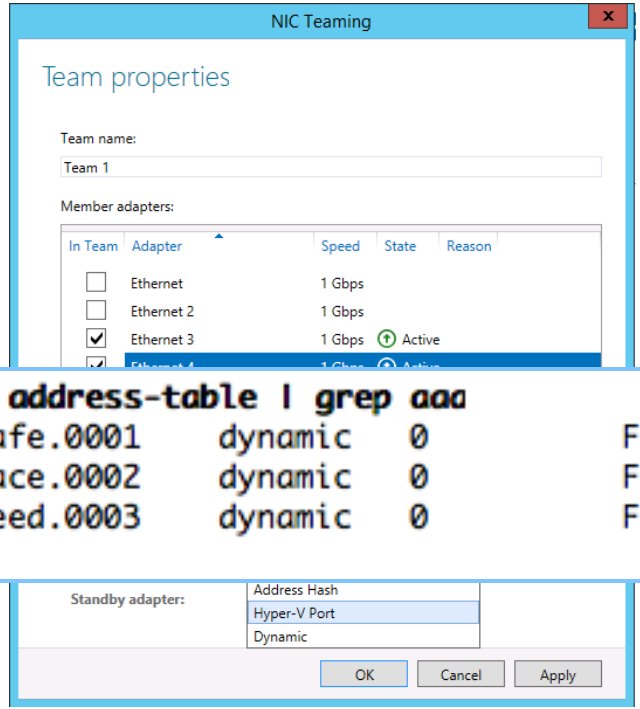




Switch Dependent

Load Balancing Mode: Address Hash, Hyper-V Port and Dynamic






```
SV-5K-1# show mac address-table | grep aaa
* 511       aaaa.cafe.0001    dynamic   0          F     F   Po300
* 511       aaaa.face.0002    dynamic   0          F     F   Po300
* 511       aaaa.feed.0003    dynamic   0          F     F   Po300
SV-5K-1#
```

Microsoft Switch Dependent

Teaming Mode: Static Teaming – Switch Configuration

```
interface Ethernet101/1/31-32
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree port type edge trunk
  channel-group 300

interface Port-Channel300
  switchport mode trunk
  switchport trunk allowed vlan 511
```

```
SV-5K-1(config-if)# channel-group 300 mode ?
  active     Set channeling mode to ACTIVE
  on         Set channeling mode to ON
  passive    Set channeling mode to PASSIVE

SV-5K-1(config-if)# channel-group 300 mode
```

- Used with all Load Balancing modes

Microsoft Switch Dependent

Teaming Mode: LACP – Switch Configuration

```
interface Ethernet101/1/31-32
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree port type edge trunk
  channel-group 300 mode active

interface Port-Channel300
  switchport mode trunk
  switchport trunk allowed vlan 511
```

```
SV-5K-1(config-if)# channel-group 300 mode ?
  active     Set channeling mode to ACTIVE
  on         Set channeling mode to ON
  passive    Set channeling mode to PASSIVE

SV-5K-1(config-if)# channel-group 300 mode
```

- Used with all Load Balancing modes

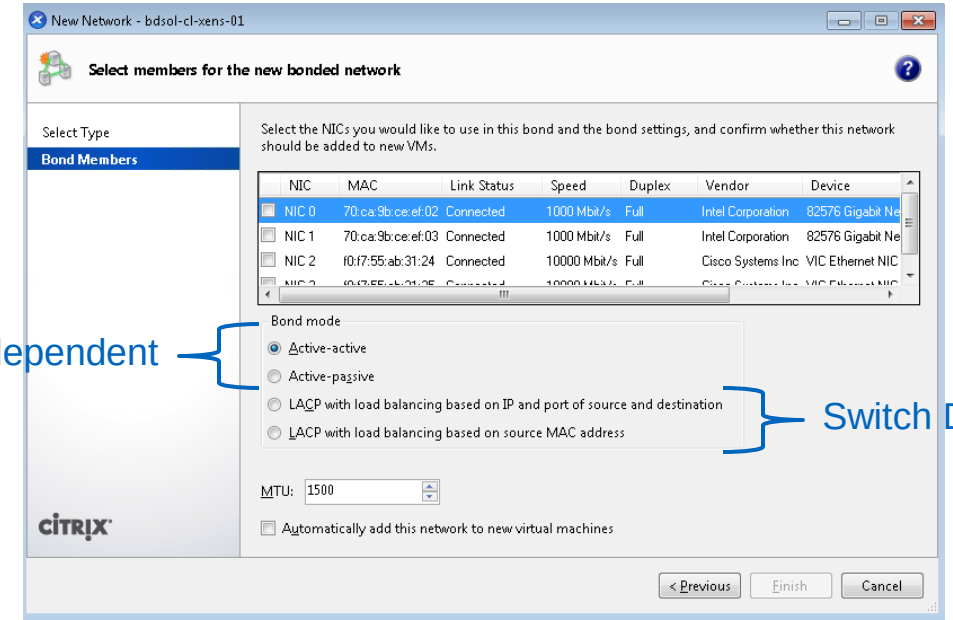
Hyper-V Conclusion

- Switch independent
  - Switch Independent – Dynamic
- Switch dependent
  - Switch Dependent - Dynamic

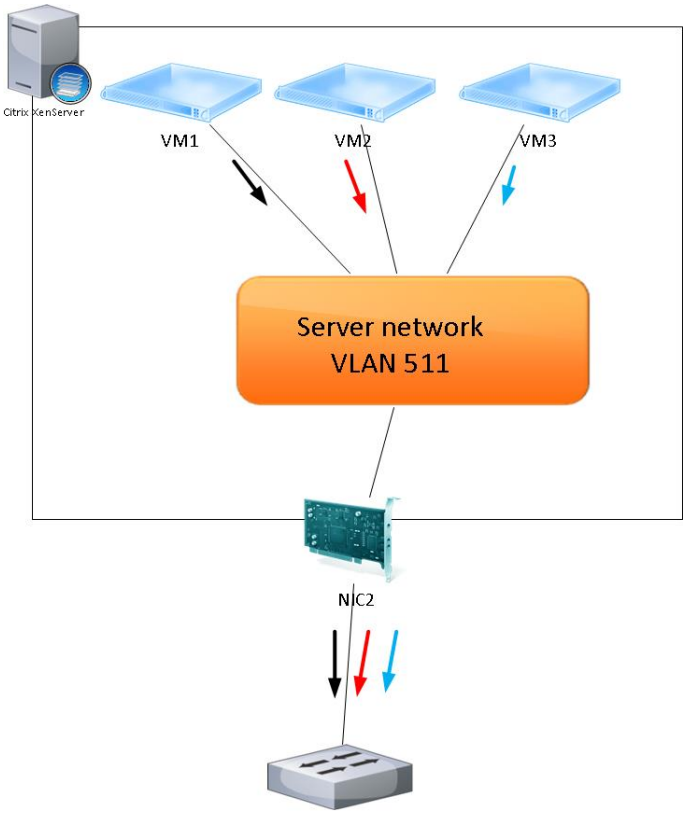
Hypervisor Overview
Citrix XenServer 6.2

Uplink Options

Single physical NIC or Bonding
LACP as of XenServer 6.1

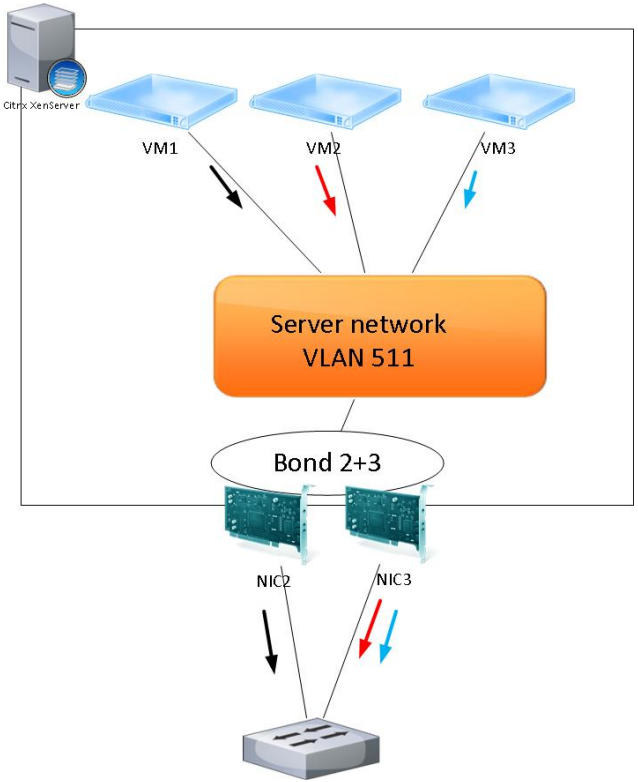


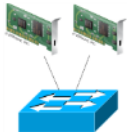
Single Physical NIC



- Normal server connection
- Access or trunk
- Multiple MACs on one switch port

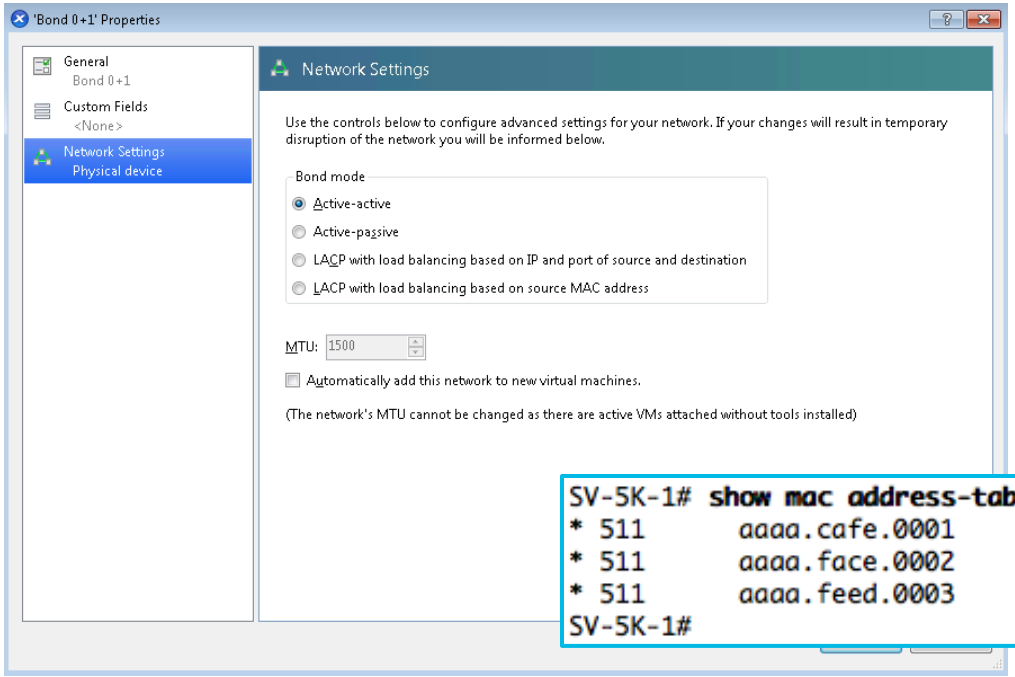
Active-Active





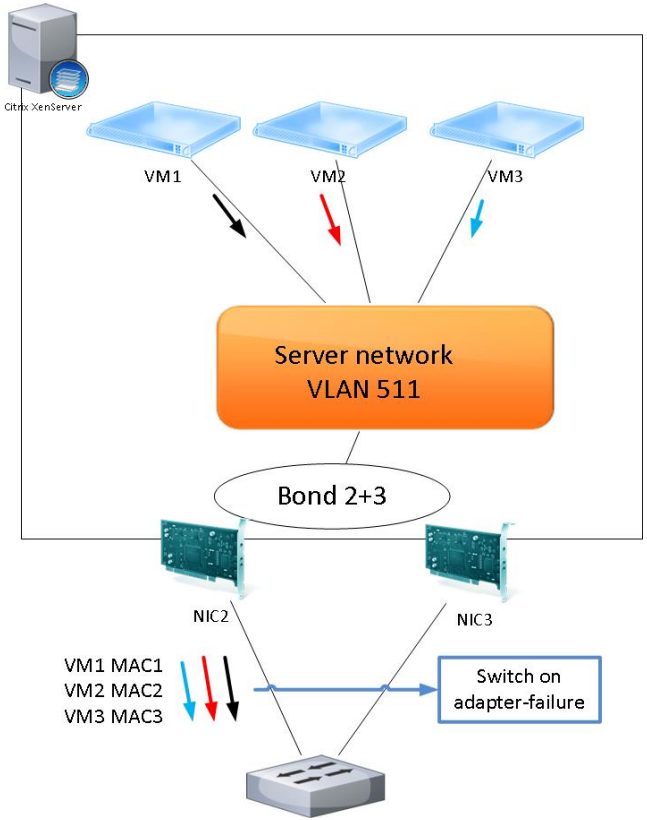
- Balance-SLB - Server Load Balancing
- vNIC pinned to single pNIC
- vNIC pinning recalculated every 10s based on pNIC load
- MAC moves
- Can go to single or multiple switches

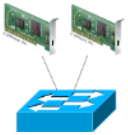
Active-Active




```
SV-5K-1# show mac address-table | grep aaaa
* 511      aaaa.cafe.0001    dynamic   10          F    F  Eth101/1/31
* 511      aaaa.face.0002    dynamic   20          F    F  Eth101/1/32
* 511      aaaa.feed.0003    dynamic   0           F    F  Eth101/1/31
SV-5K-1#
```

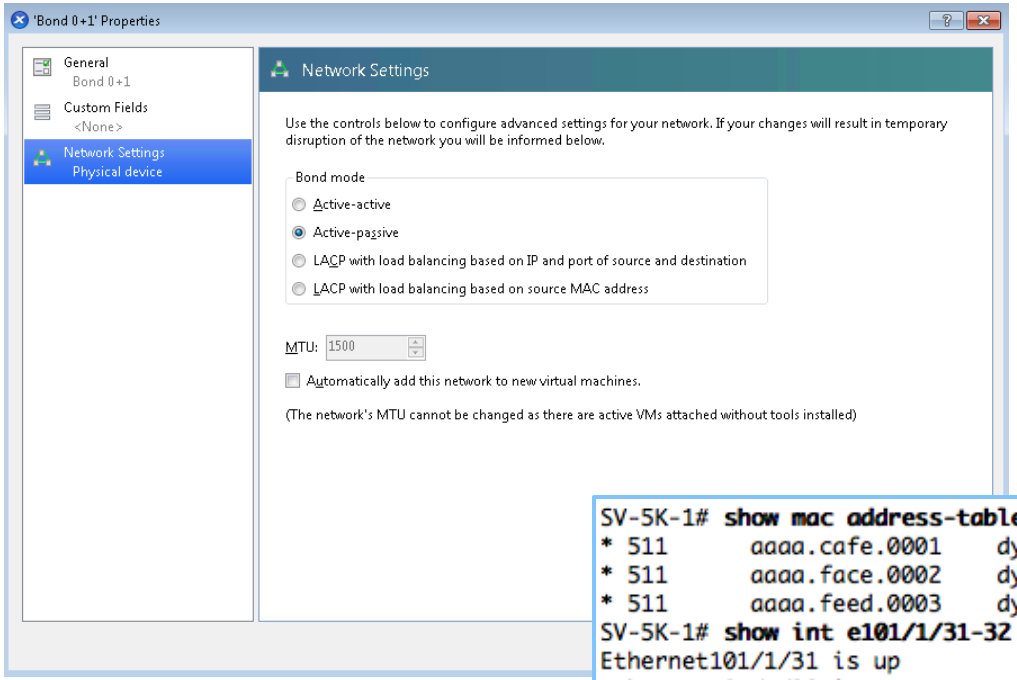
Active-Passive





- Single pNIC is active, all other pNICs are standby
- Can go to single or multiple switches

Active-Passive




```
SV-5K-1# show mac address-table | grep aaaa
* 511      aaaa.cafe.0001    dynamic   10         F    F  Eth101/1/32
* 511      aaaa.face.0002    dynamic   10         F    F  Eth101/1/32
* 511      aaaa.feed.0003    dynamic   10         F    F  Eth101/1/32
SV-5K-1# show int e101/1/31-32 | grep 101/1/3
Ethernet101/1/31 is up
Ethernet101/1/32 is up
SV-5K-1#
```

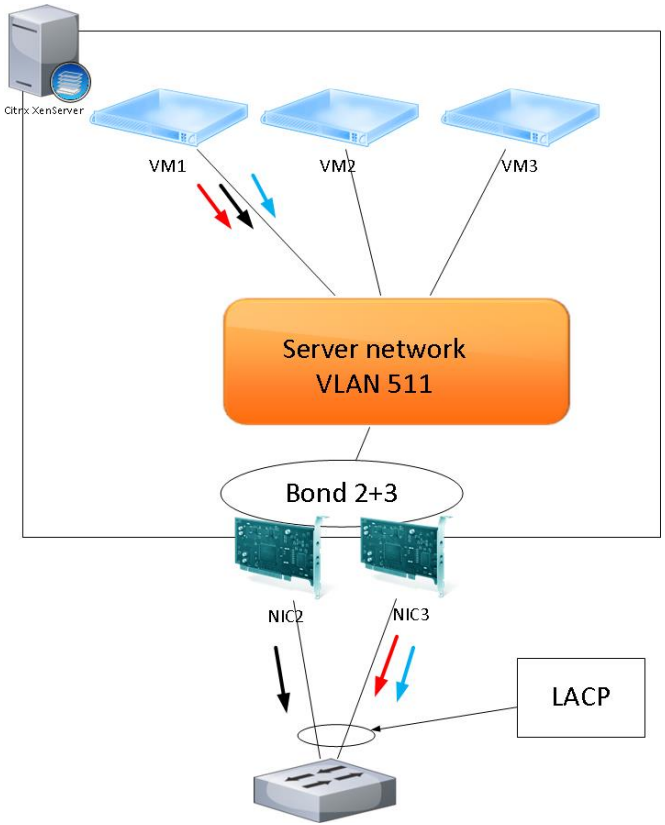
Citrix Switch Independent

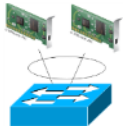
Active/Active & Active/Passive – Switch Configuration

For Your Reference

```
interface Ethernet101/1/31-32
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree port type edge trunk
```

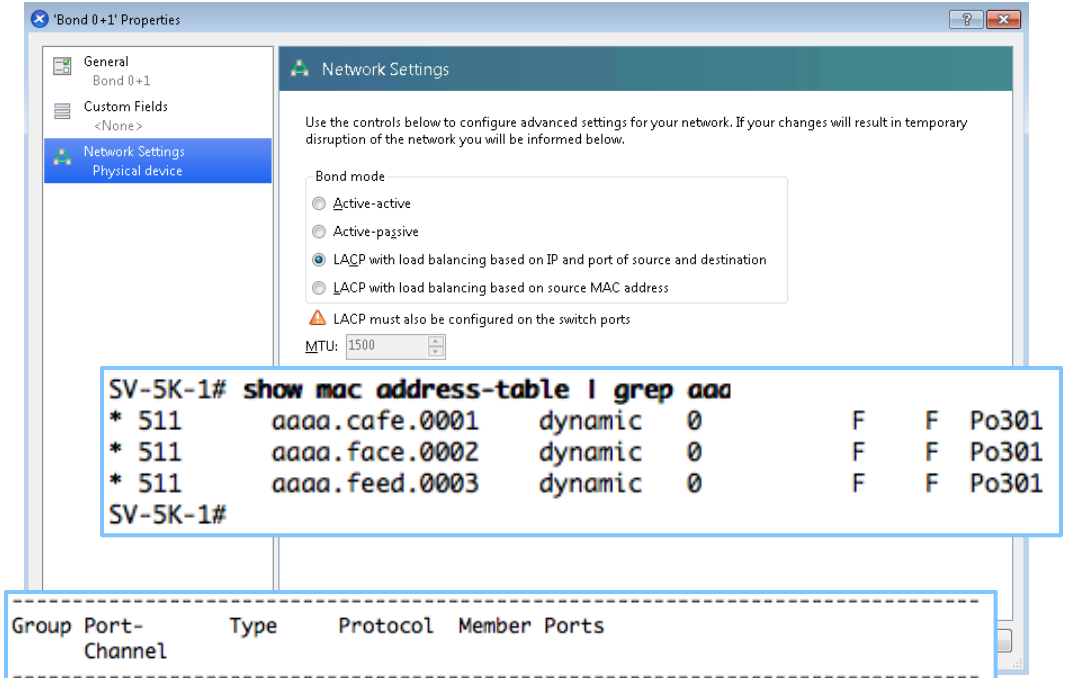
LACP w/ Load Balancing Based on IP & Port of SRC & DST





- Outgoing traffic will be hashed over multiple uplinks based on src/dst IP and port
- Inbound traffic distributed by switch load balancing algorithm
- Connect to single port-channel

LACP w/ Load Balancing Based on IP & Port of SRC & DST

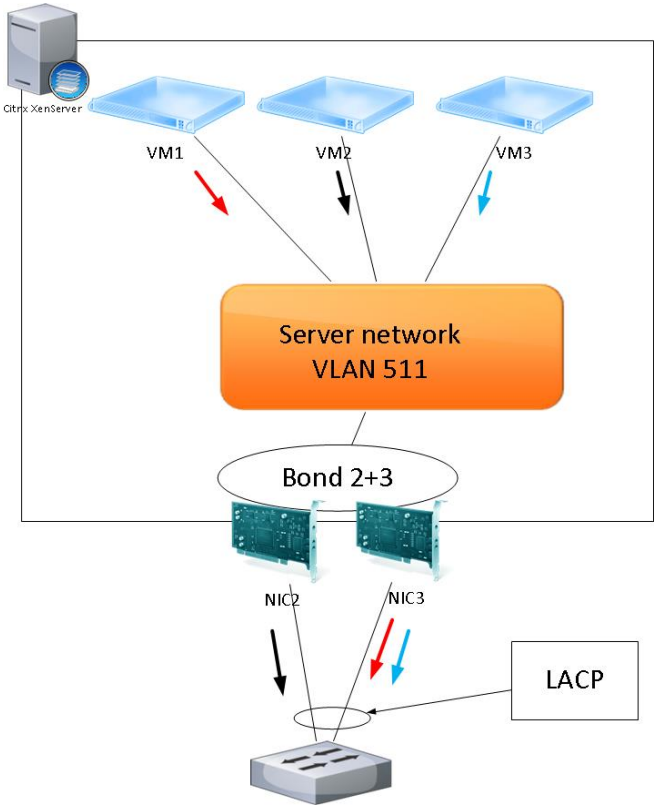


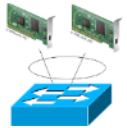

```
SV-5K-1# show mac address-table | grep aaa
* 511       aaaa.cafe.0001    dynamic   0          F     F   Po301
* 511       aaaa.face.0002    dynamic   0          F     F   Po301
* 511       aaaa.feed.0003    dynamic   0          F     F   Po301
SV-5K-1#
```

```
--------------------------------------------------------------------------------
Group Port-       Type     Protocol  Member Ports
      Channel
--------------------------------------------------------------------------------
301   Po301(SU)   Eth      LACP      Eth101/1/31(P)  Eth101/1/32(P)
SV-5K-1#
```

```
[root@xenserver-1 ~]# ovs-appctl bond/show bond0
bond_mode: balance-tcp
bond-hash-algorithm: balance-tcp
bond-hash-basis: 0
updelay: 31000 ms
downdelay: 200 ms
next rebalance: 1382936 ms
lacp_negotiated: true

slave eth0: enabled
        active slave
        may_enable: true
        hash 30: 37 kB load
        hash 70: 38 kB load
        hash 72: 0 kB load
        hash 93: 0 kB load
        hash 131: 1 kB load
        hash 139: 0 kB load
        hash 212: 0 kB load
        hash 220: 0 kB load
        hash 222: 0 kB load
        hash 250: 0 kB load

slave eth2: enabled
        may_enable: true
        hash 55: 0 kB load
        hash 69: 37 kB load
        hash 83: 0 kB load
        hash 142: 0 kB load
[root@xenserver-1 ~]#
```

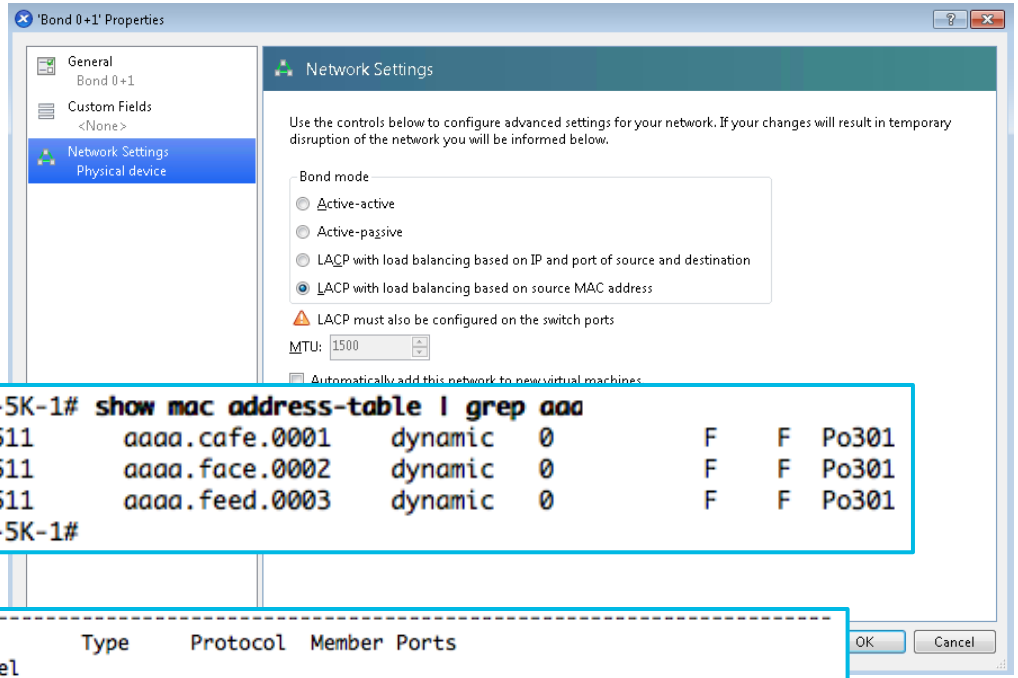
LACP w/ Load Balancing Based on SRC MAC Address





- vNIC is pinned to pNIC base on src MAC address hash
- Connect to single port-channel

LACP w/ Load Balancing Based on SRC MAC Address




```
SV-5K-1# show mac address-table | grep aaa
* 511       aaaa.cafe.0001    dynamic   0          F     F  Po301
* 511       aaaa.face.0002    dynamic   0          F     F  Po301
* 511       aaaa.feed.0003    dynamic   0          F     F  Po301
SV-5K-1#
```

```
--------------------------------------------------------------------------------
Group Port-       Type     Protocol  Member Ports
      Channel
--------------------------------------------------------------------------------
301   Po301(SU)   Eth      LACP      Eth101/1/31(P)  Eth101/1/32(P)
SV-5K-1#
```

```
[root@xenserver-1 ~]# ovs-appctl bond/show bond0
bond_mode: balance-slb
bond-hash-algorithm: balance-slb
bond-hash-basis: 0
updelay: 31000 ms
downdelay: 200 ms
next rebalance: 1764448 ms
lacp_negotiated: true

slave eth0: enabled
        active slave
        may_enable: true
        hash 93: 3 kB load
        hash 115: 3 kB load
        hash 150: 3 kB load

slave eth2: enabled
        may_enable: true
[root@xenserver-1 ~]#
```

Citrix Switch Dependent

LACP Bonds – Switch Configuration

```
interface Ethernet101/1/31-32
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree port type edge trunk
  channel-group 300 mode active

interface Port-Channel300
  switchport mode trunk
  switchport trunk allowed vlan 511
```

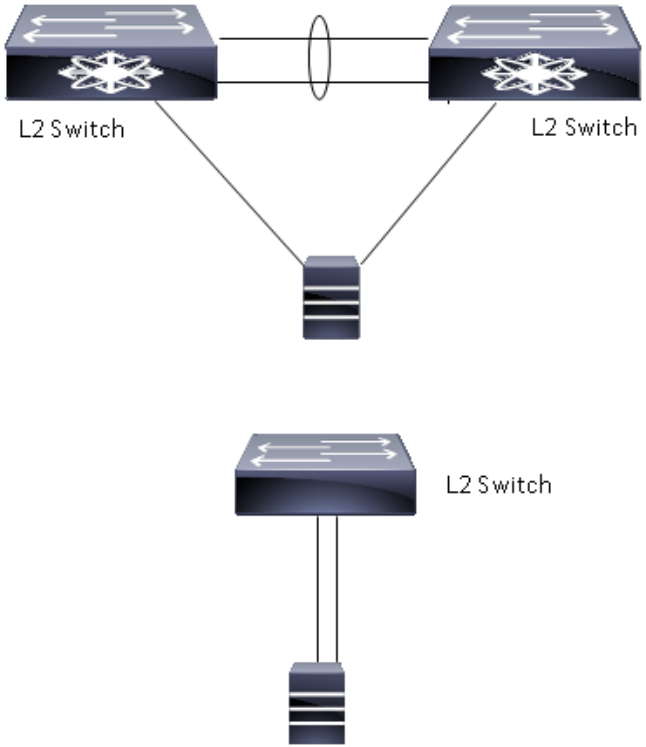
```
SV-5K-1(config-if)# channel-group 300 mode ?
  active     Set channeling mode to ACTIVE
  on         Set channeling mode to ON
  passive    Set channeling mode to PASSIVE

SV-5K-1(config-if)# channel-group 300 mode
```

XenServer Conclusion

- Switch independent
  - Active-Active
- Switch dependent
  - LACP with load balancing based on source MAC address

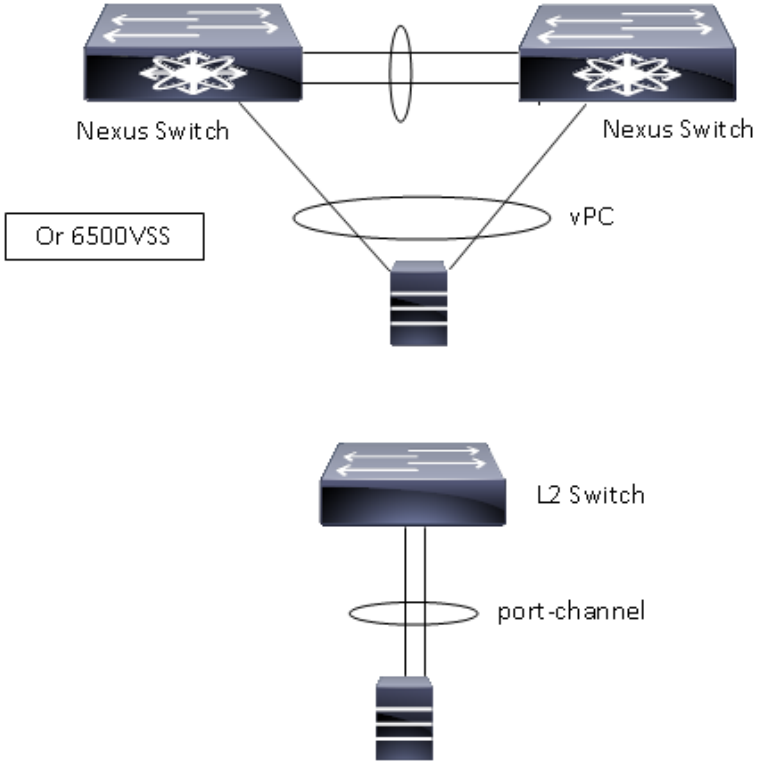
Topology Overview

Switch Independent



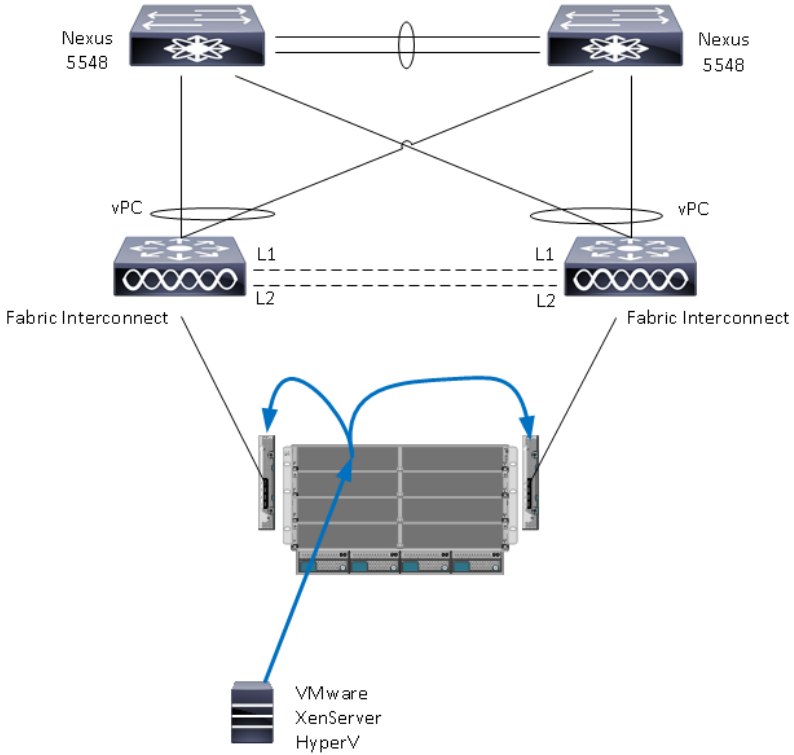
- VMware
  - Route based on originating virtual port
  - Route based on source MAC hash
  - Route based on physical NIC load (vDS)
  - Use explicit failover order
- Hyper-V
  - Switch Independent - Address hash
  - Switch Independent - Hyper-V Port mode
  - Switch Independent - Dynamic
- XenServer
  - Active-active
  - Active-passive

Switch Dependent



- VMware
  - Route based on IP hash
  - Route based on IP hash + LACP (vDS)
- Hyper-V
  - Switch Dependent - All Address hash modes
  - Switch Dependent - Hyper-V Port mode
  - Switch Dependent - Dynamic
- XenServer
  - LACP with load balancing based on IP and port of source and destination
  - LACP with load balancing based on source MAC address

Cisco UCS-B – Switch Independent

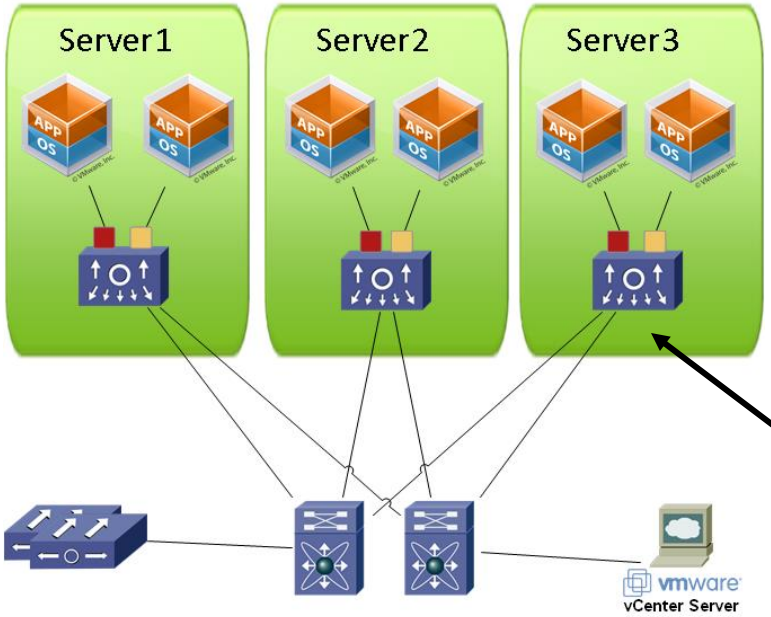


- Each Fabric Interconnect has a port-channel towards the Nexus 5000 vPC pair
- Fabric Interconnects are connected for control-plane clustering only - no data-plane traffic is exchanged
- The hypervisor running on a blade has 2 independent connections - no switch dependent protocols can be used




Nexus 1000V

Nexus 1000V



Virtual Supervisor Module(VSM)

- CLI interface into the Nexus 1000V
- Leverages NX-OS
- Controls multiple VEMs as a single network device

Virtual Ethernet Module(VEM)

- Replaces Vmware's virtual switch
- Enables advanced switching capability on the hypervisor
- Provides each VM with dedicated "switch ports"

Load Balance Options

```
Nexus1000v(config)# port-channel load-balance ethernet ?
  dest-ip-port               Destination IP address and L4 port
  dest-ip-port-vlan          Destination IP address, L4 port and VLAN
  destination-ip-vlan        Destination IP address and VLAN
  destination-mac            Destination MAC address
  destination-port           Destination L4 port
  source-dest-ip-port        Source & Destination IP address and L4 port
  source-dest-ip-port-vlan   Source & Destination IP address, L4 port and VLAN
  source-dest-ip-vlan        Source & Destination IP address and VLAN
  source-dest-mac            Source & Destination MAC address
  source-dest-port           Source & Destination L4 port
  source-ip-port             Source IP address and L4 port
  source-ip-port-vlan        Source IP address, L4 port and VLAN
  source-ip-vlan             Source IP address and VLAN
  source-mac                 Source MAC address
  source-port                Source L4 port
  source-virtual-port-id     Source Virtual Port Id
  vlan-only                  VLAN only

Nexus1000v(config)#
```

Nexus 1000V

- Multi hypervisor switch (VMware, Hyper-V* and KVM*)
- Layer 2 switching: VLANs, private VLANs, VXLAN, loop prevention, multicast, virtual PortChannels, LACP, ACLs
- Network management: SPAN, ERSPAN, Netflow 9, vTracker, vCenter Server Plug-in
- Enhanced QOS features
- Cisco vPath
- Security: DHCP Snooping, IP Source Guard, Dynamic ARP inspection, Cisco TrustSec SGA support
- Cisco Virtual Security Gateway
- Other virtual services (Cisco ASA 1000V, Cisco vWAAS, etc..)

Nexus 1000V

Sessions

- LTRVIR-2005 Deploying the Nexus 1000V on ESXi and Hyper-V
- BRKVIR-2017 The Nexus 1000V on Microsoft Hyper-V: Expanding the Virtual Edge
- BRKVIR-3013 Deploying and Troubleshooting the Nexus 1000v virtual switch
- BRKVIR-2023 Cisco Nexus 1000V InterCloud based Hybrid Cloud Architectures and Approaches

Key Take-Aways

- Understand the hypervisor's load-balancing mechanisms
- Align the configuration on hypervisor and upstream switch
- All adapters are always up
- MAC moves are possible depending on the load balancing algorithm
- Use the correct port-channel configuration (on/active/passive)




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!

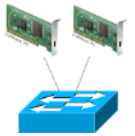
Visit us online after the conference for full access to session videos and presentations.
www.CiscoLiveAPAC.com


CISCO
TM




Appendix

Switch Independent Configuration Options (IOS)



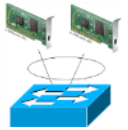
Access

```
interface range GigabitEthernet1/0/31 - 32
  switchport mode access
  switchport access vlan 511
  spanning-tree portfast
```

Trunk

```
interface range GigabitEthernet1/0/31 - 32
  switchport trunk encapsulation dot1q
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree portfast trunk
```

Switch Dependent Configuration Options (IOS)



Static

```
interface range GigabitEthernet1/0/31 - 32
  switchport trunk encapsulation dot1q
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree portfast trunk
  channel-group 300 mode on
```

LACP

```
interface range GigabitEthernet1/0/31 - 32
  switchport trunk encapsulation dot1q
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree portfast trunk
  channel-group 300 mode active
```

```
interface Port-Channel300
  switchport mode trunk
  switchport trunk allowed vlan 511
  spanning-tree portfast trunk
```


CISCO
TM