
TOMORROW
starts here.
CISCO
Cisco live!

UCS Networking Deep Dive

BRKCOM-2003

Michael Ciesla, Customer Support Engineer

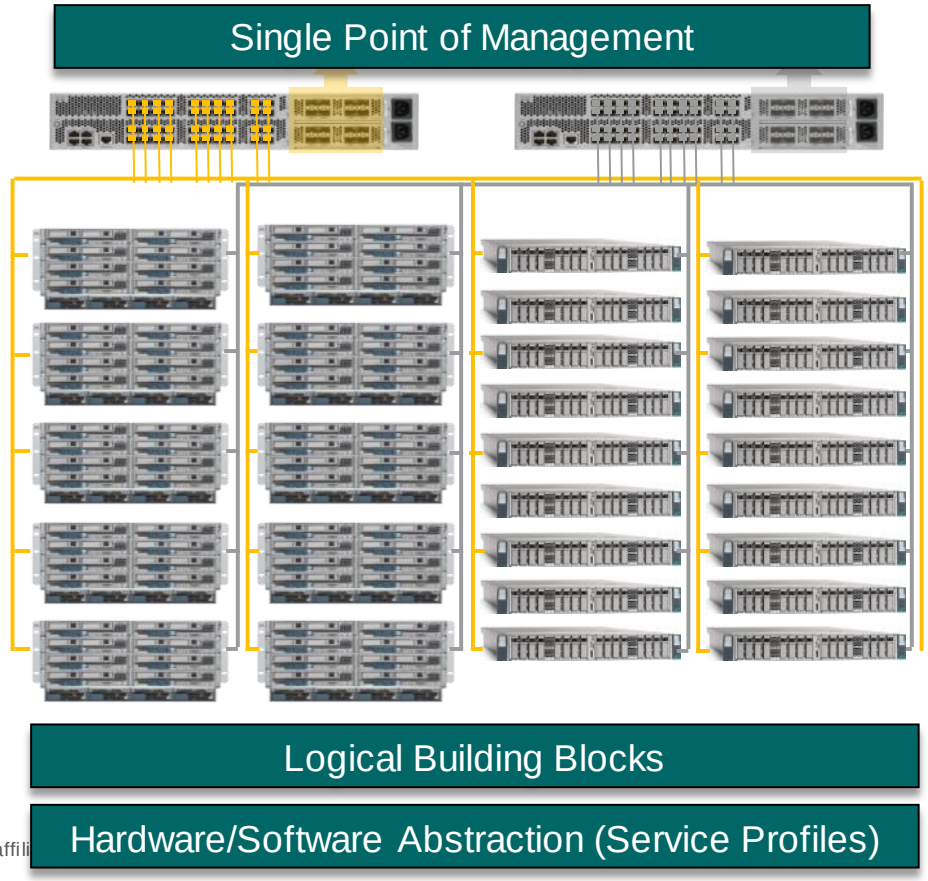
#clmel

Cisco live!

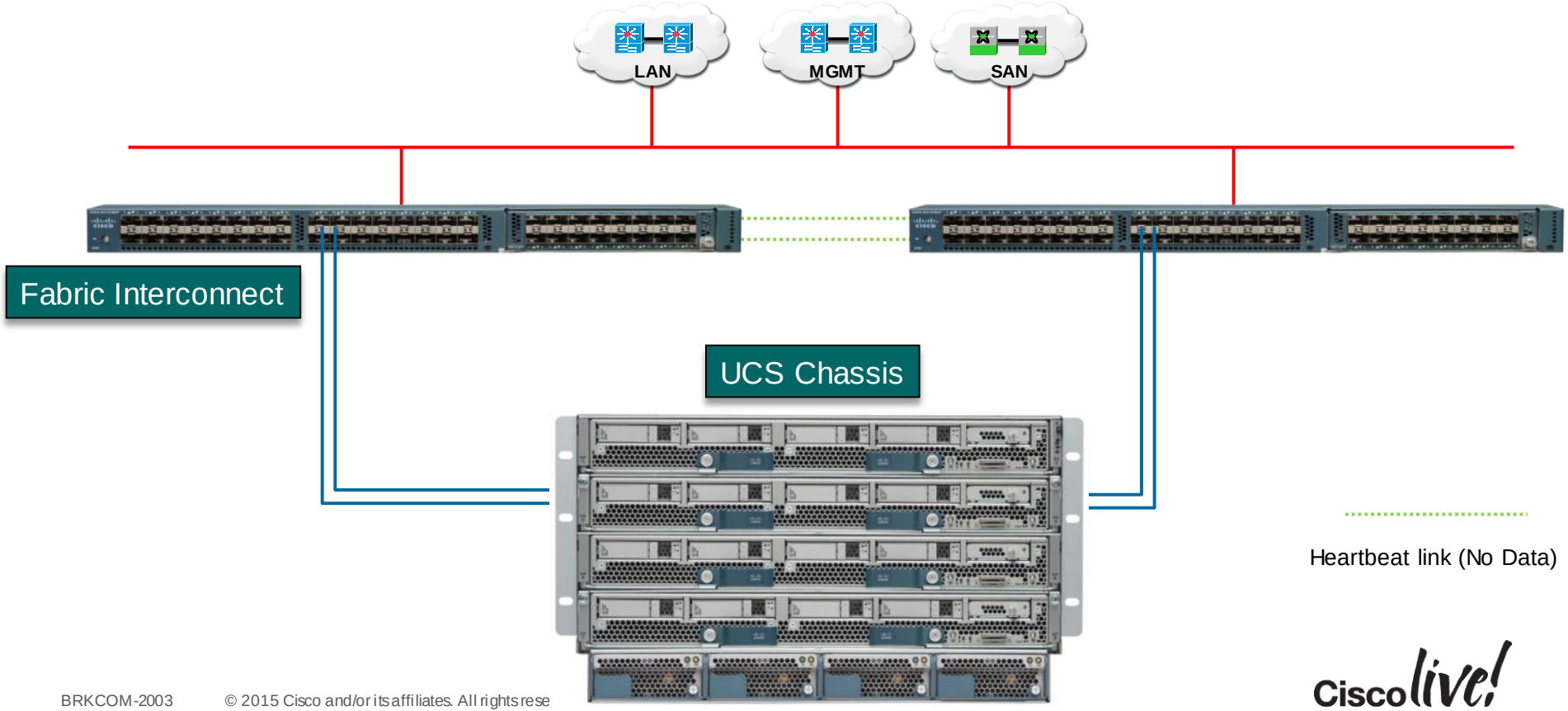
Agenda

- UCS Overview
- Chassis Connectivity
- Server Connectivity
- Fabric Forwarding
- Topologies

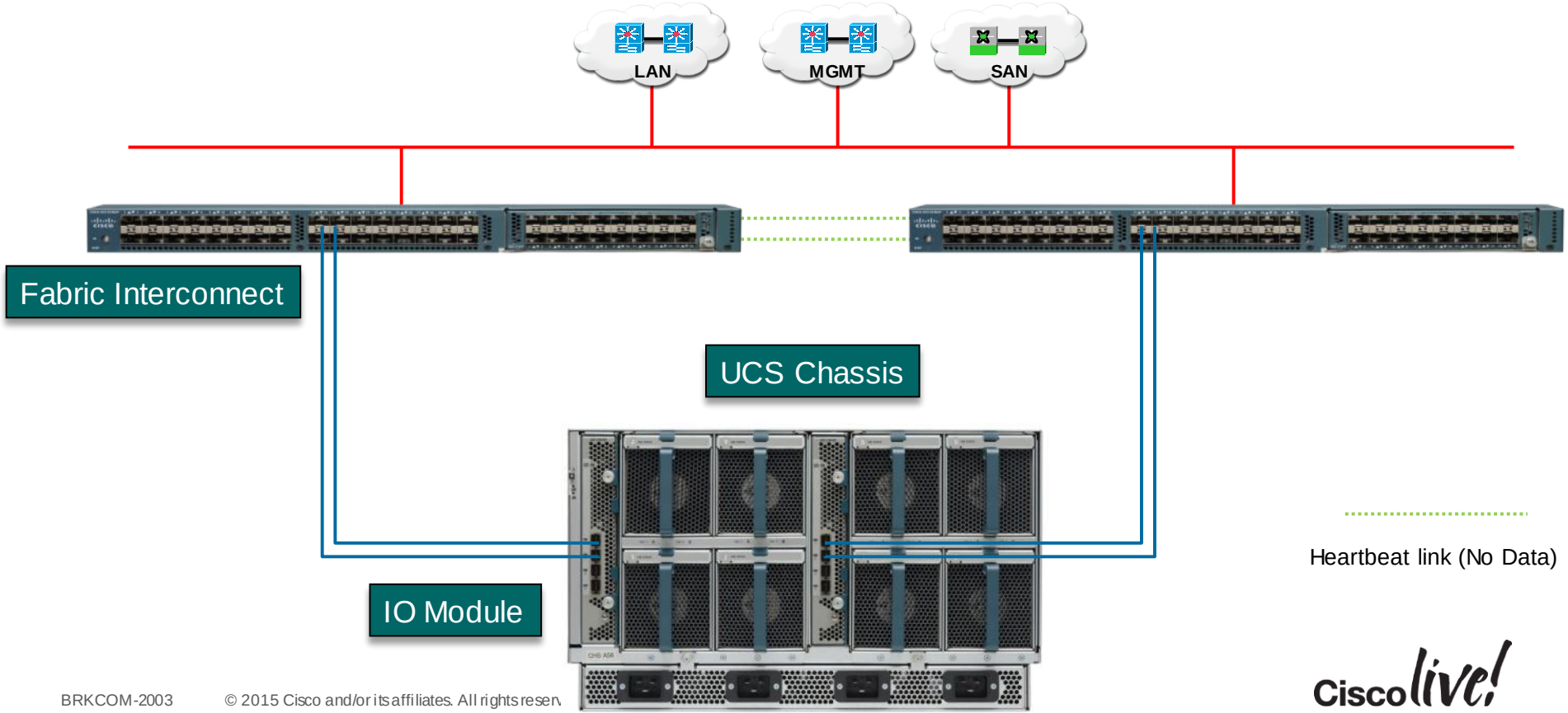
Cisco Unified Computing System (UCS)



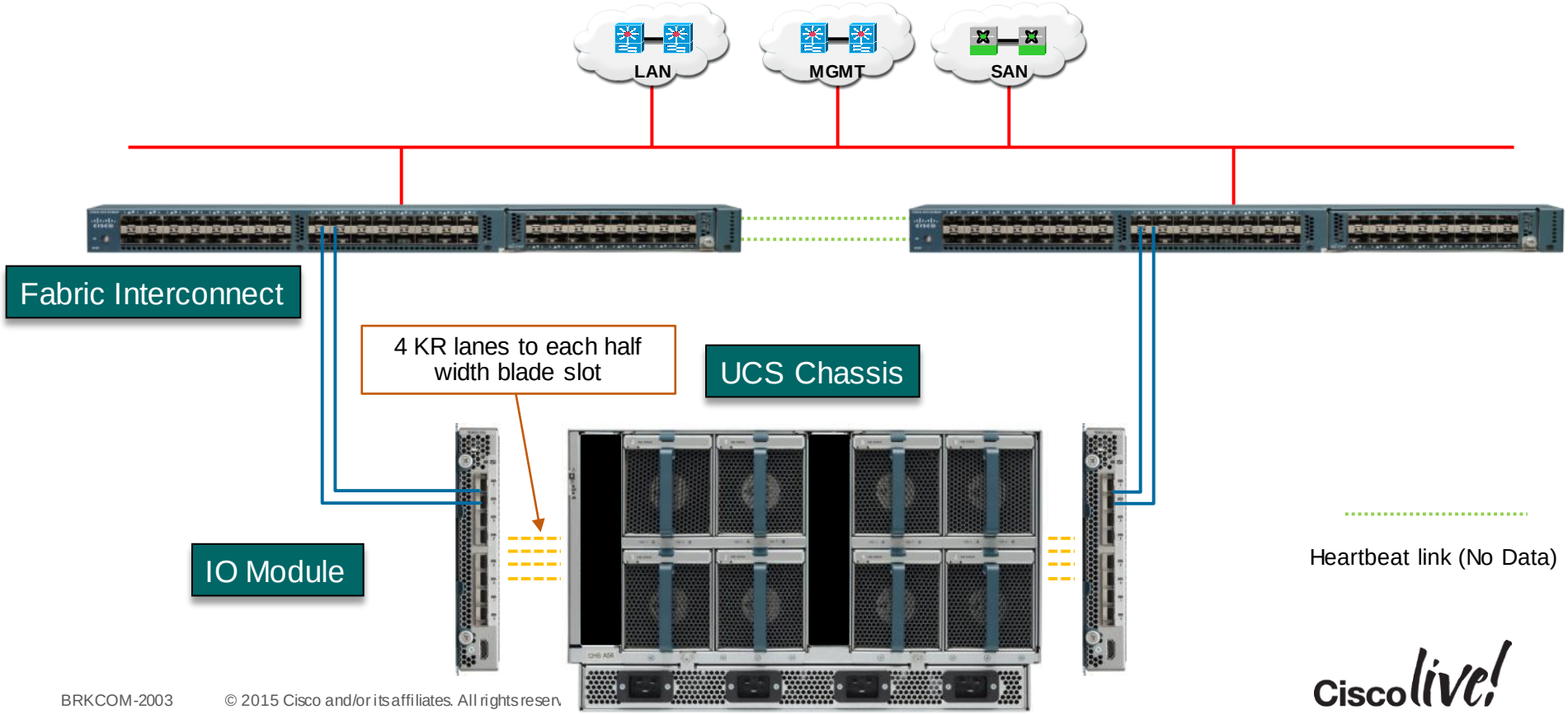
UCS Components



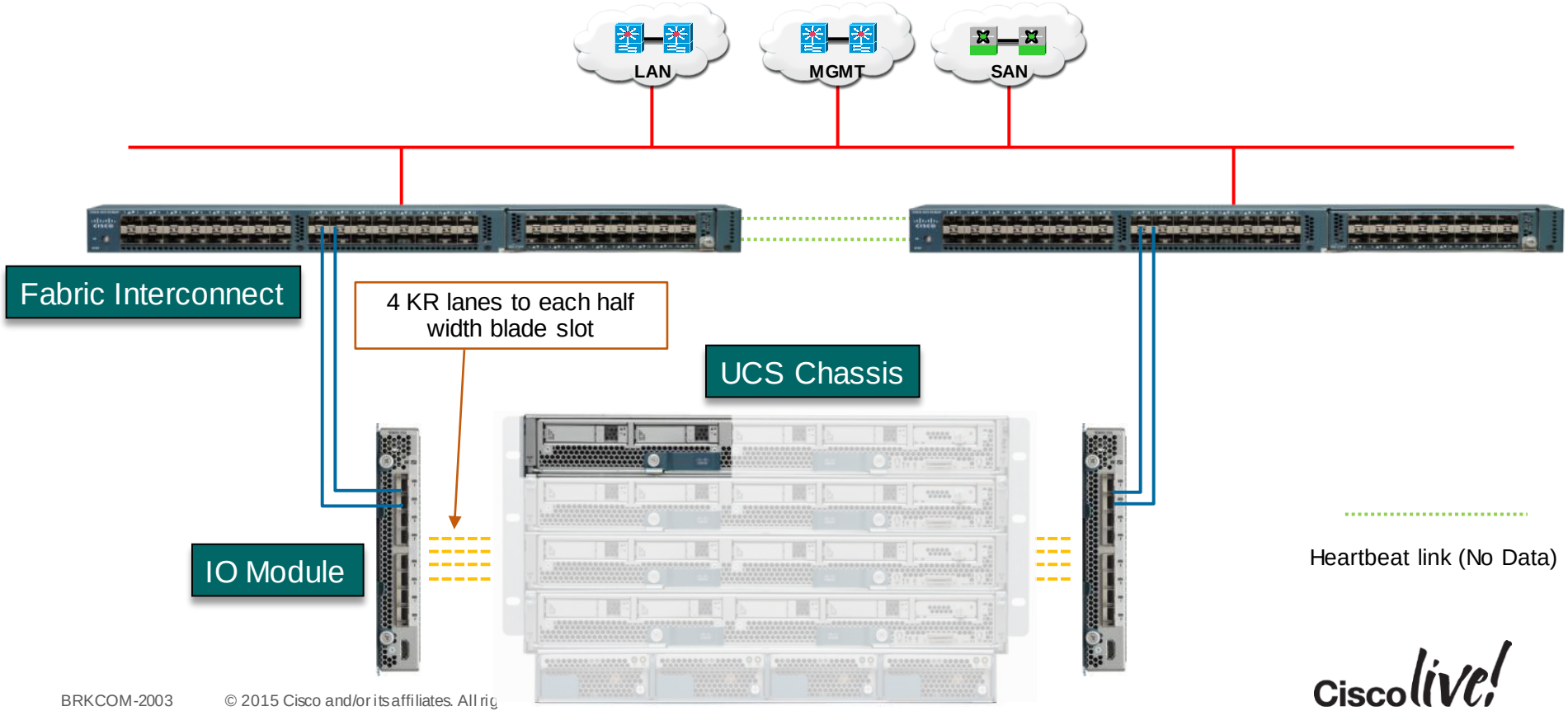
UCS Components



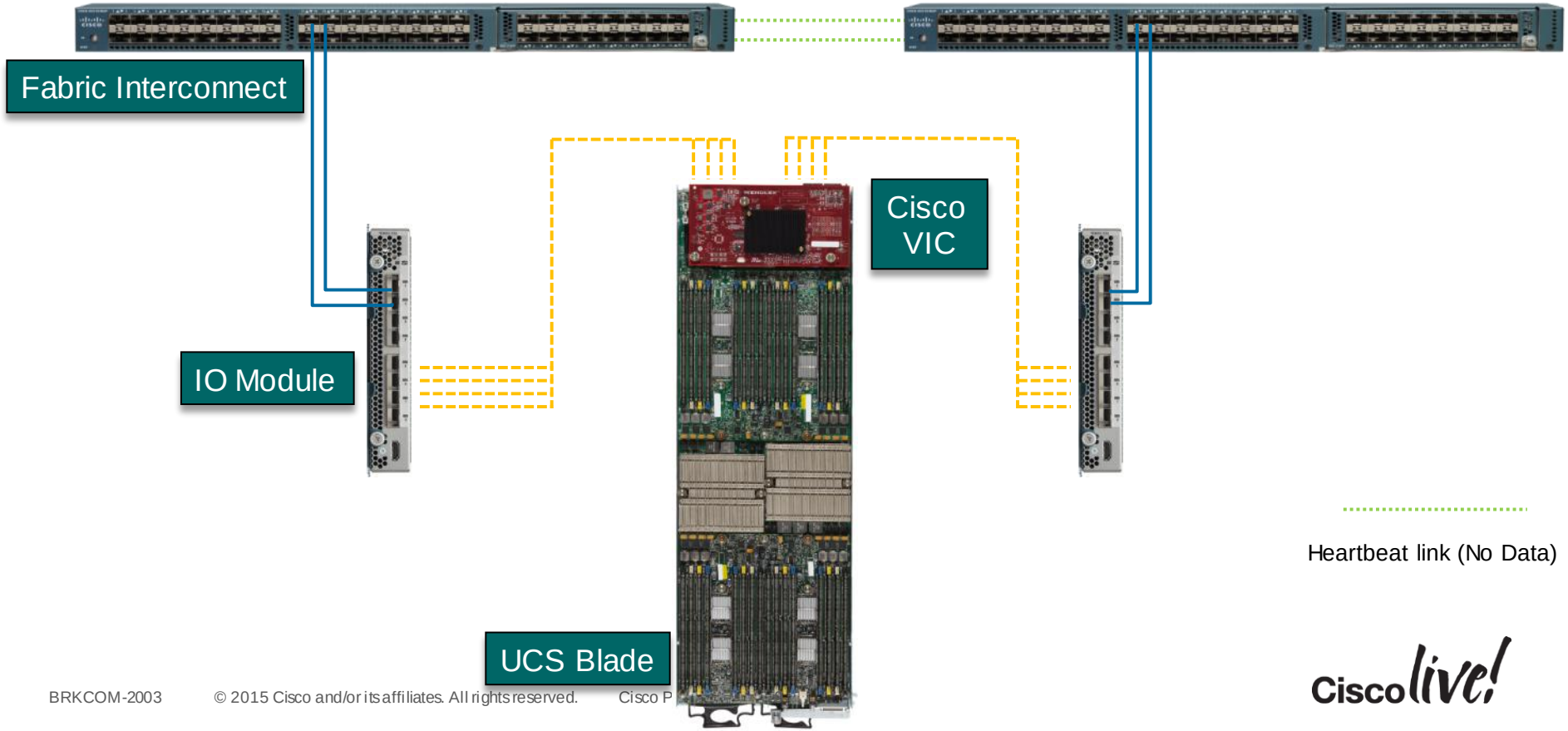
UCS Components



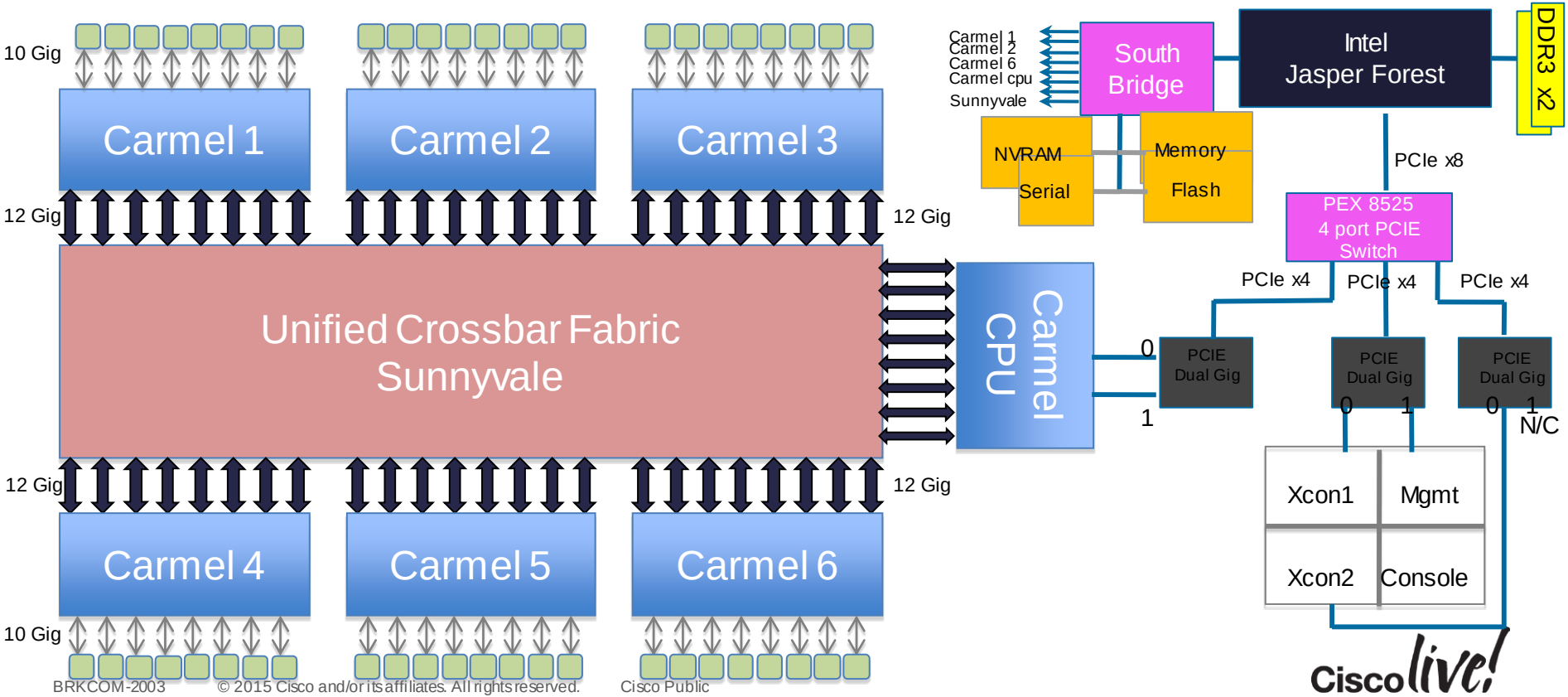
UCS Components



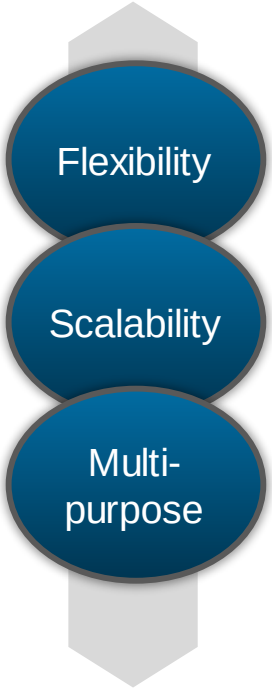
UCS Components



UCS 6248 Hardware Diagram



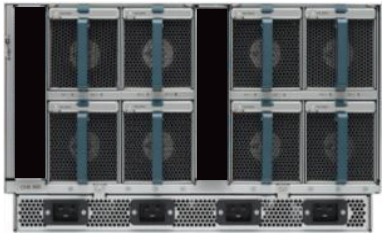
Cisco UCS 6200 Series Fabric Interconnects



Product Features and Specs	UCS 6248UP	UCS 6296UP
Switch Fabric Throughput	960 Gbps	1.92 Tbps
Switch Footprint	1RU	2RU
1 Gigabit Ethernet Port Density	48	96
10 Gigabit Ethernet Port Density	48	96
8G Native FC Port Density	48	96
Port-to-Port Latency	2.0us	2.0us
Active # of VLANs	2000	2000

UCS Mini: 6324 Fabric Interconnect

UCS B

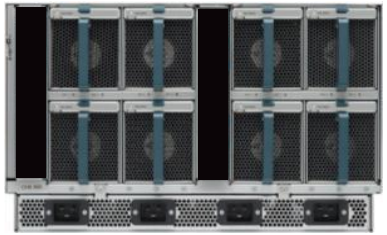


UCS 5108 Chassis
Supports existing and future blades

IO Modules

6248 or 6296 Fabric
Fabric Interconnects

UCS Mini

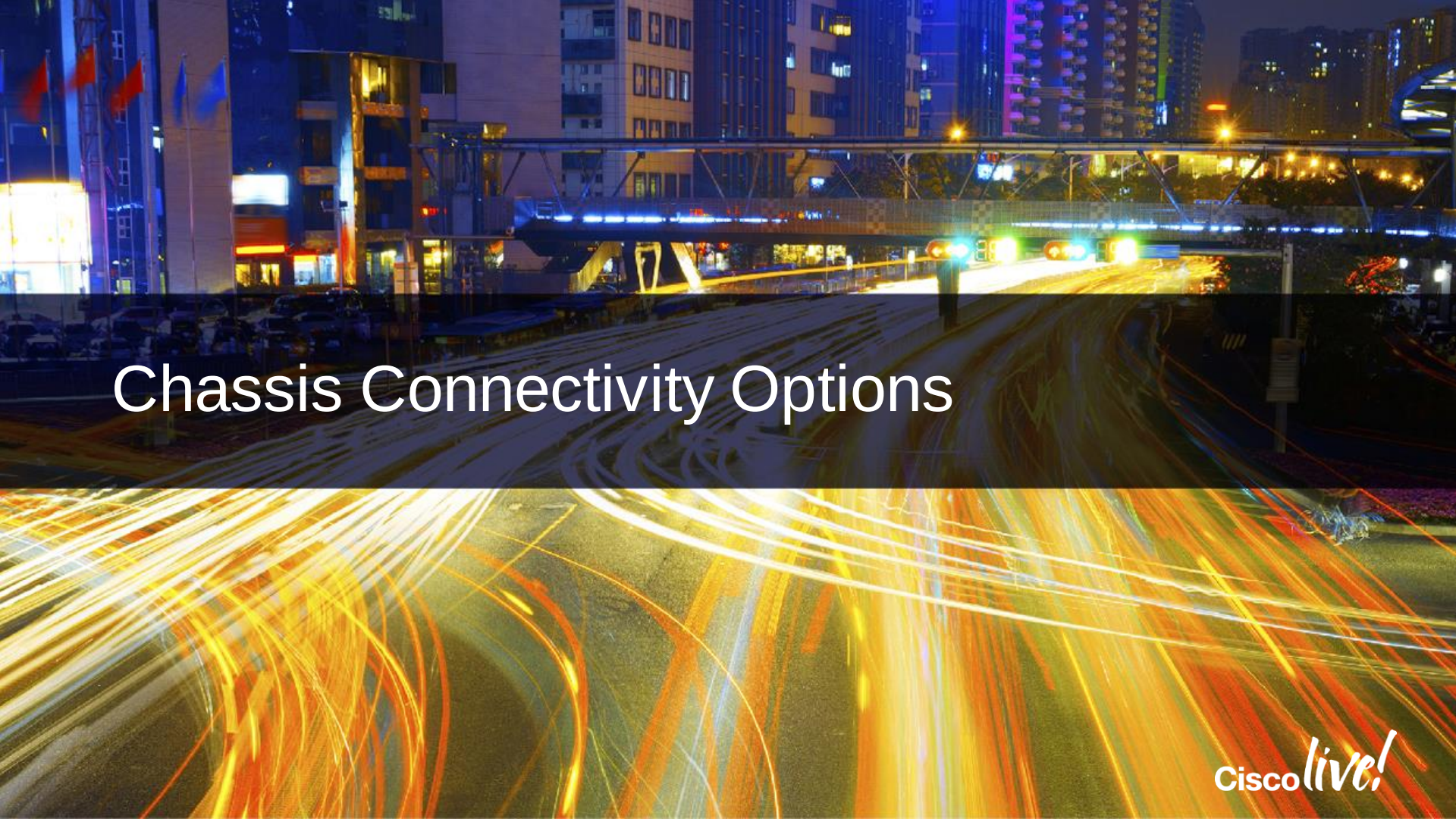


UCS 5108 Chassis
Supports existing and future blades

6324 Fabric Interconnect

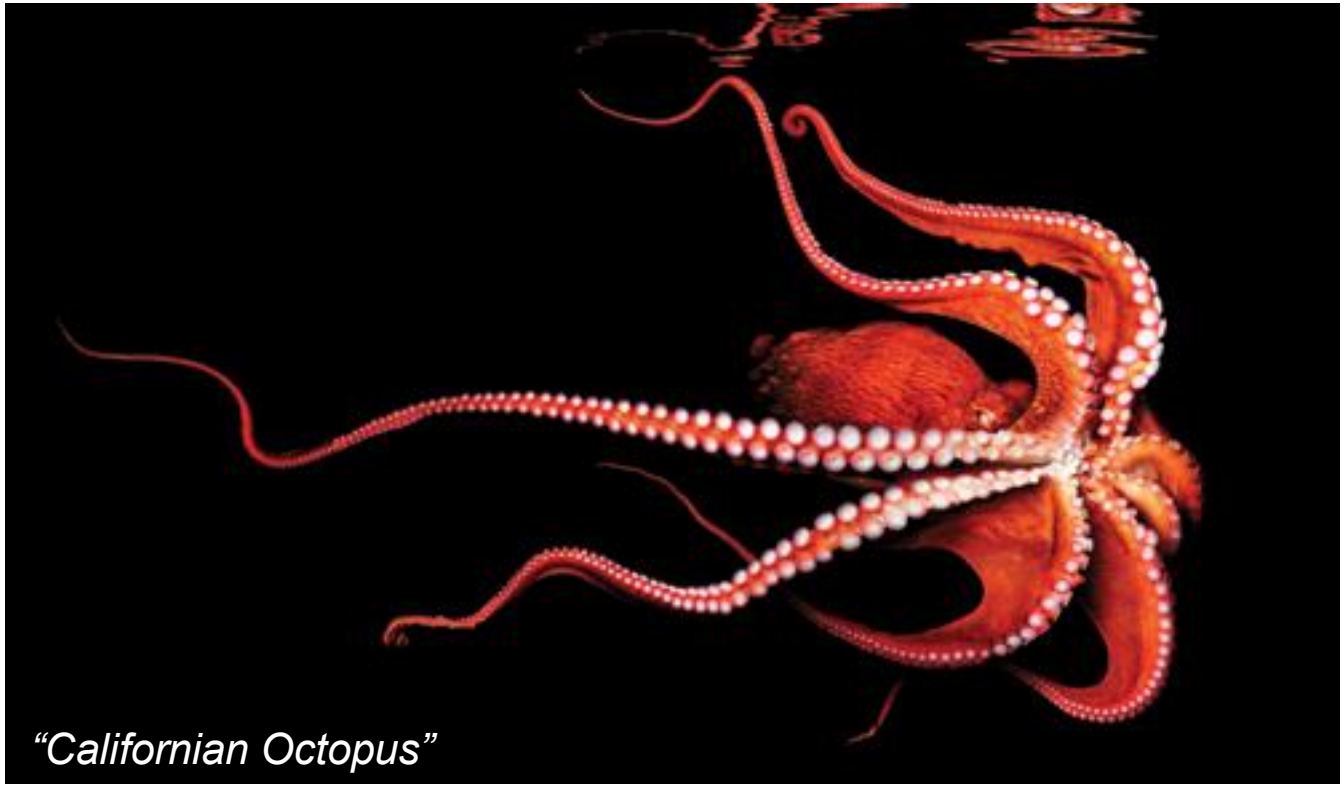
TAC Tip: Carmel ASIC Port Mapping

```
UCSB-2-A(nxos)# show hardware internal carmel all-ports

Carmel Port Info:
name    |log|car|mac|flag|adm|opr|m:s:l|ipt|fab|xcar|xpt|if_index|diag|ucVer
--------+---+---+---+----+---+---+-----+---+---+----+---+--------+----+-----
xgb1/1  |0  |0  |0 p|b7  |en |up |0:0:f|0  |92 |0   |0  |1a000000|pass| 0.00
xgb1/2  |1  |0  |1 p|b7  |en |up |1:1:f|1  |88 |0   |0  |1a001000|pass| 0.00
xgb1/3  |2  |0  |2 p|b7  |en |up |2:2:f|2  |93 |0   |0  |1a002000|pass| 0.00
xgb1/4  |3  |0  |3 p|b7  |en |up |3:3:f|3  |89 |0   |0  |1a003000|pass| 0.00
xgb1/5  |4  |0  |4 -|b7  |dis|dn |4:4:f|4  |90 |0   |0  |1a004000|pass| 0.00
xgb1/6  |5  |0  |5 -|b7  |dis|dn |5:5:f|5  |94 |0   |0  |1a005000|pass| 0.00
xgb1/7  |6  |0  |6 -|b7  |dis|dn |6:6:f|6  |95 |0   |0  |1a006000|pass| 0.00
xgb1/8  |7  |0  |7 -|b7  |dis|dn |7:7:f|7  |91 |0   |0  |1a007000|pass| 0.00
xgb1/9  |8  |1  |0 -|b7  |dis|dn |0:0:f|0  |80 |0   |0  |1a008000|pass| 0.00
xgb1/10 |9  |1  |1 -|b7  |dis|dn |1:1:f|1  |87 |0   |0  |1a009000|pass| 0.00
xgb1/11 |10 |1  |2 -|b7  |dis|dn |2:2:f|2  |81 |0   |0  |1a00a000|pass| 0.00
xgb1/12 |11 |1  |3 -|b7  |dis|dn |3:3:f|3  |86 |0   |0  |1a00b000|pass| 0.00
xgb1/13 |12 |1  |4 -|b7  |dis|dn |4:4:f|4  |82 |0   |0  |1a00c000|pass| 0.00
xgb1/14 |13 |1  |5 -|b7  |dis|dn |5:5:f|5  |85 |0   |0  |1a00d000|pass| 0.00
xgb1/15 |14 |1  |6 -|b7  |dis|dn |6:6:f|6  |83 |0   |0  |1a00e000|pass| 0.00
xgb1/16 |15 |1  |7 -|b7  |dis|dn |7:7:f|7  |84 |0   |0  |1a00f000|pass| 0.00
```



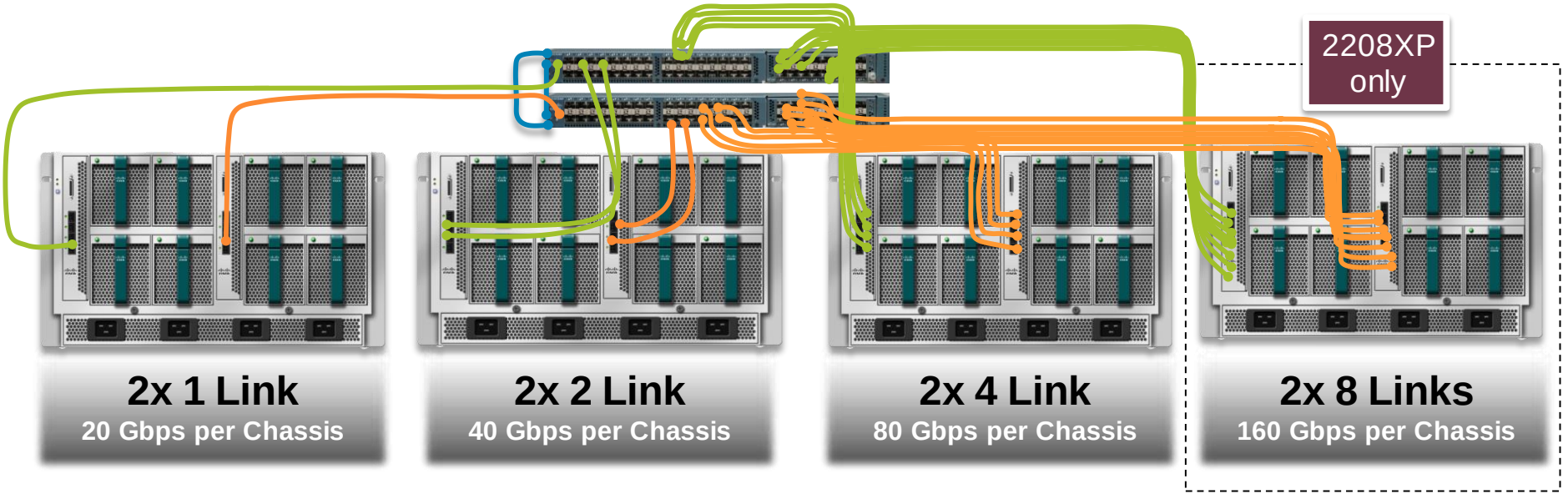
Chassis Connectivity Options

Cisco live!



UCS Fabric Topologies

Chassis Bandwidth Options



UCS 2200 IO Module (FEX)

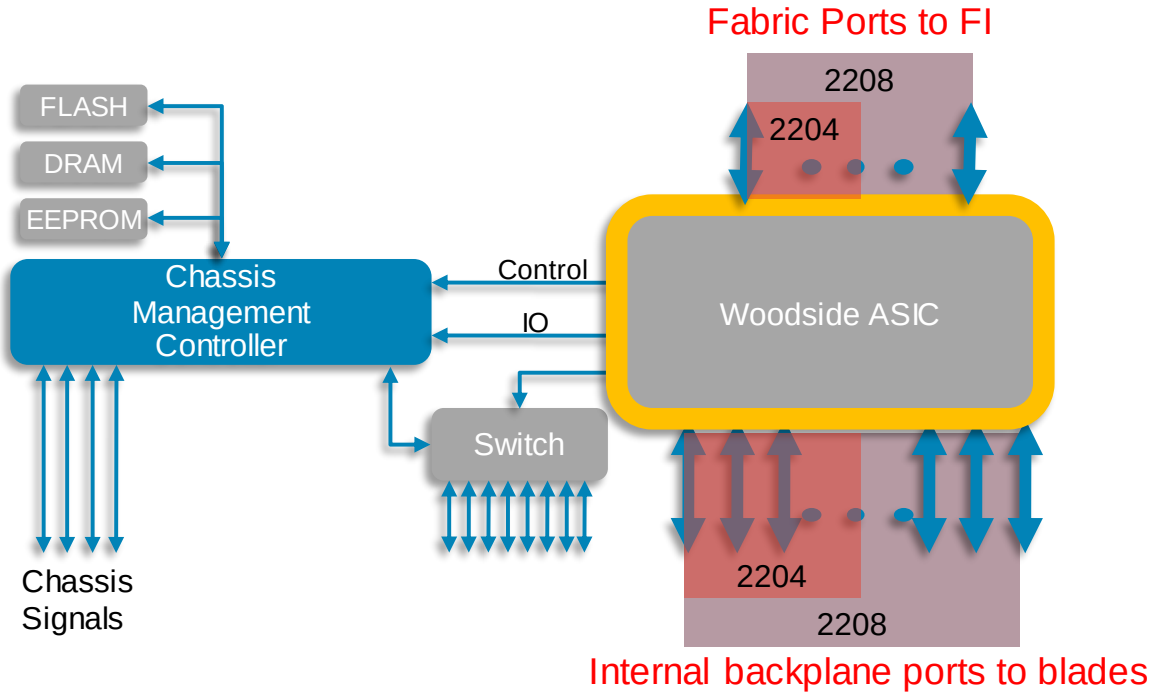
- **UCS-IOM-2204XP**

- 40G to the Network
- 160G to the Hosts
  - 2x10G Half width slot
  - 4x10G Full width slot

- **UCS-IOM-2208XP**

- 80G to the Network
- 320G to the Hosts
  - 4x10G Half width slot
  - 8x10G Full width slot

220x-XP Architecture



Feature	2204-XP	2208-XP
ASIC	Woodside	Woodside
Fabric Ports (NIF)	4	8
Host Ports (HIF)	16	32
CoS	8	8
Latency	~ 500ns	~ 500ns

No Local Switching – ever!
Traffic goes up to FI

NIF/HIF Interfaces

- Output from “show fex detail” of the NXOS shell

```
Fabric interface state:
  Po1153 - Interface Up. State: Active
  Eth1/4 - Interface Up. State: Active
Fex Port         State  Fabric Port
       Eth1/1/1  Down       Po1153
       Eth1/1/2  Down         None
       Eth1/1/3    Up       Po1153
       Eth1/1/4  Down       Po1153
       Eth1/1/5    Up       Po1153
       Eth1/1/6    Up       Po1153
       Eth1/1/7    Up       Po1153
       Eth1/1/8  Down         None
       Eth1/1/9    Up       Po1153
      Eth1/1/10  Down         None
      Eth1/1/11    Up       Po1153
      Eth1/1/12    Up       Po1153
      Eth1/1/13    Up       Po1153
      Eth1/1/14  Down         None
      Eth1/1/15    Up       Po1153
      Eth1/1/16  Down         None
      Eth1/1/17    Up       Po1153
```

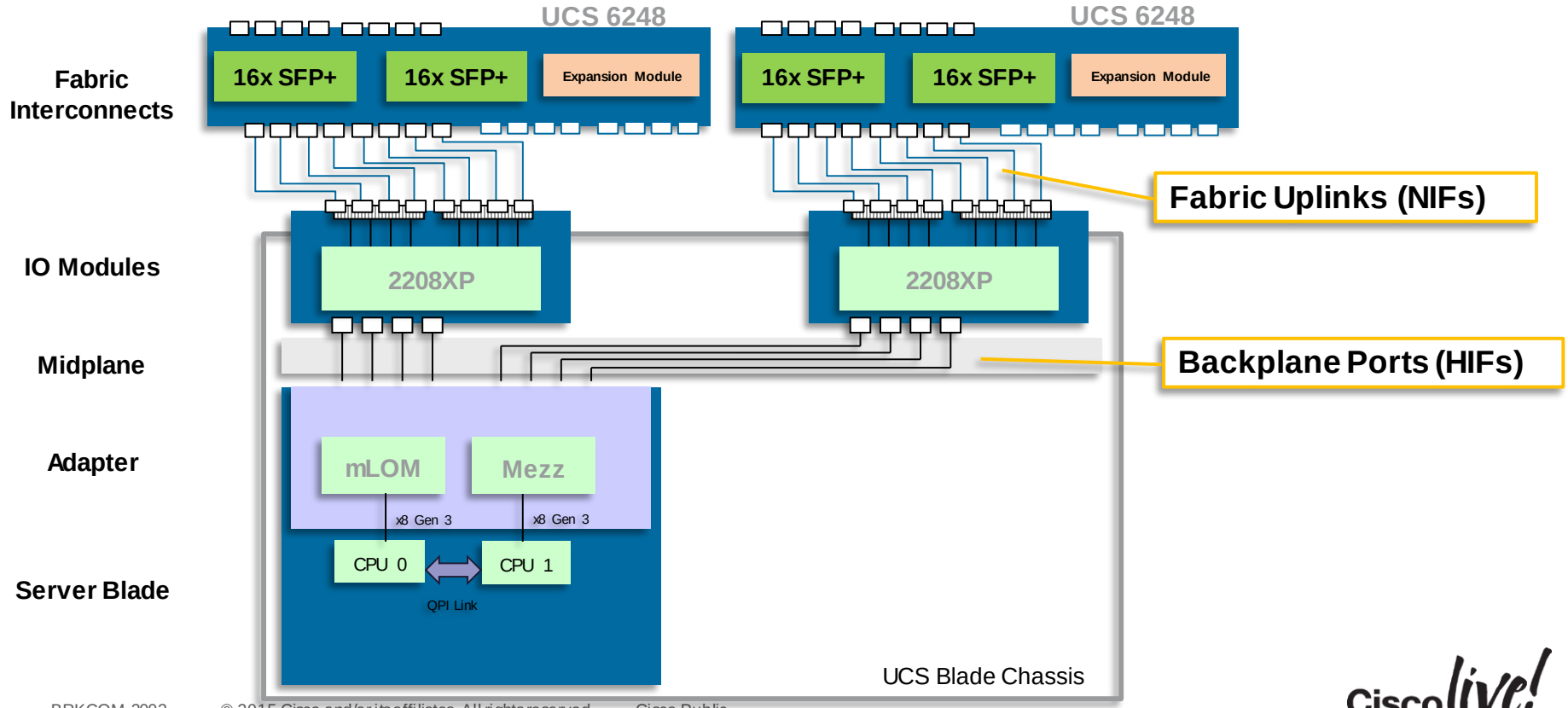
FI ports connecting to the FEX

Backplane ports connecting to server 1/3

Link between the FI and IOM/FEX which the server will be using

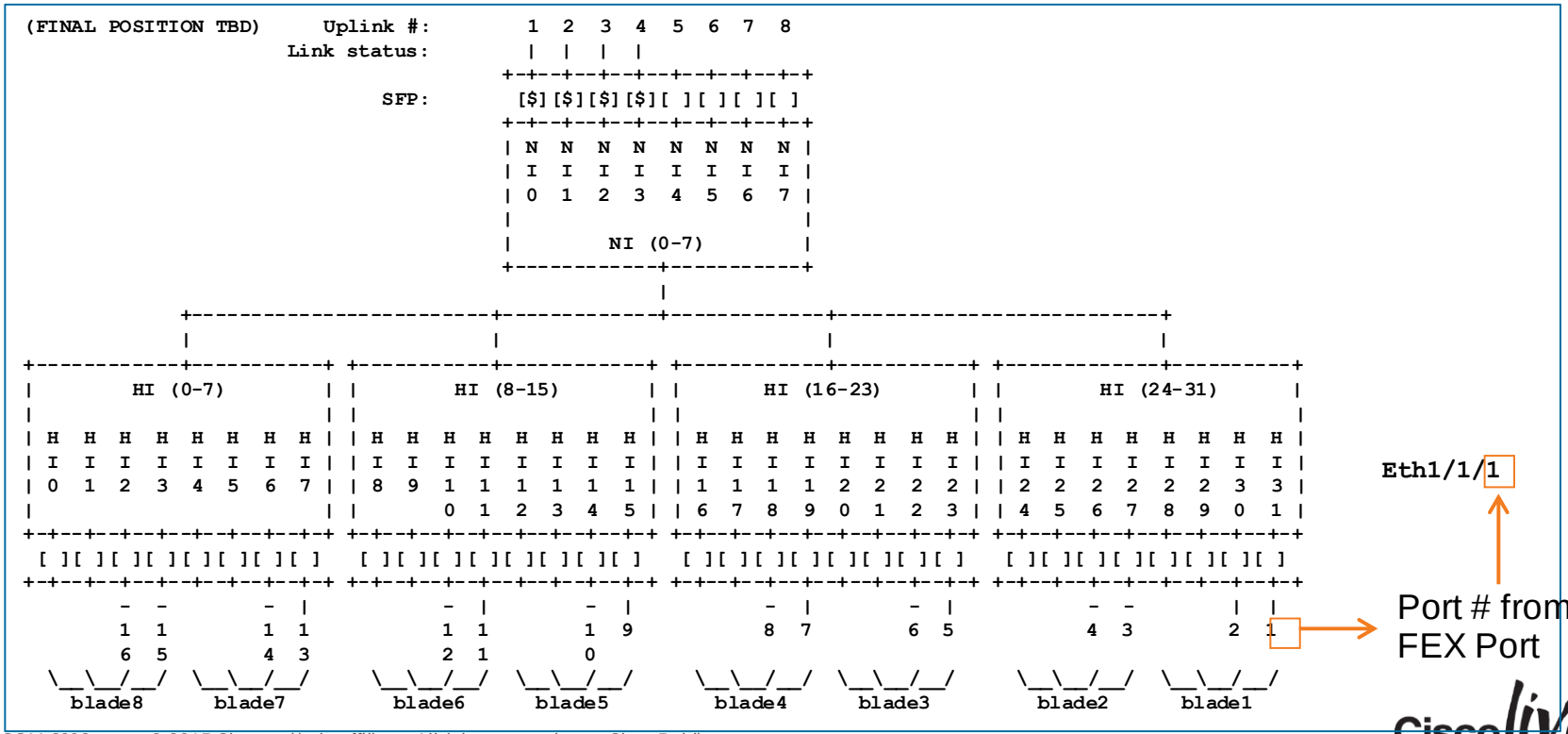
Additional interface?

UCS Internal Block Diagram



CLI Block Diagram

- Output from “show platform software woodside sts” of the IOM



CLI Command

- Output from “show platform software woodside rate” of the IOM

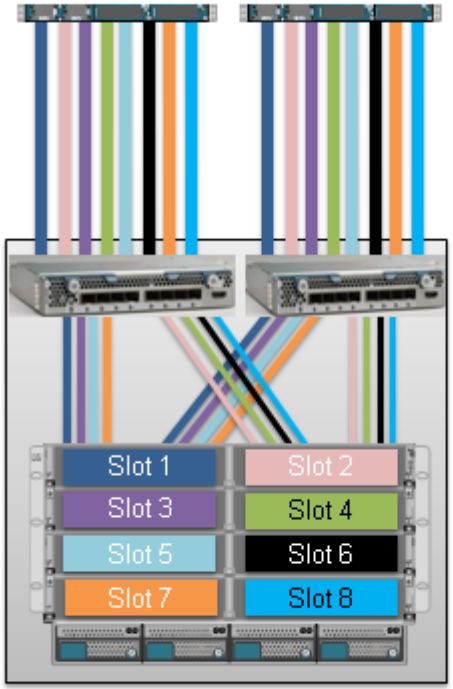
```
fex-1# sh platform software woodside rate
+--------++------------+-----------+------------++------------+-----------+------------+-------+-------+---+
| Port   || Tx Packets |  Tx Rate  |   Tx Bit   || Rx Packets |  Rx Rate  |   Rx Bit   |Avg Pkt|Avg Pkt|   |
|        ||            |  (pkts/s) |    Rate    ||            |  (pkts/s) |    Rate    | (Tx)  | (Rx)  |Err|
+--------++------------+-----------+------------++------------+-----------+------------+-------+-------+---+
| 0-BI   ||         25 |         5 |   5.04Kbps ||         14 |         2 |   3.09Kbps |   106 |   118 |   |
| 0-CI   ||         65 |        13 |  20.20Kbps ||         49 |         9 |  25.44Kbps |   174 |   304 |   |
| 0-NI3  ||          2 |         0 |   3.48Kbps ||         63 |        12 |  12.98Kbps |  1072 |   108 |   |
| 0-NI2  ||         35 |         7 |  12.05Kbps ||          6 |         1 |   2.75Kbps |   195 |   267 |   |
| 0-NI1  ||          3 |         0 |   3.69Kbps ||         14 |         2 |   4.22Kbps |   750 |   168 |   |
| 0-NI0  ||         17 |         3 |   6.86Kbps ||         20 |         4 |  12.38Kbps |   232 |   367 |   |
| 0-HI31 ||          7 |         1 |   8.00Kbps ||          1 |         0 | 768.00 bps |   695 |   464 |   |
| 0-HI30 ||          5 |         1 |   1.28Kbps ||          1 |         0 | 200.00 bps |   141 |   106 |   |
| 0-HI23 ||          5 |         1 |   1.22Kbps ||          1 |         0 | 352.00 bps |   133 |   200 |   |
| 0-HI19 ||         13 |         2 |   2.48Kbps ||          1 |         0 |   1.19Kbps |    99 |   728 |   |
| 0-HI11 ||          3 |         0 | 560.00 bps ||          1 |         0 | 128.00 bps |    98 |    64 |   |
| 0-HI7  ||          4 |         0 | 616.00 bps ||          1 |         0 | 560.00 bps |    77 |   332 |   |
+--------++------------+-----------+------------++------------+-----------+------------+-------+-------+---+
```

Blade 1

Fabric Link Connectivity

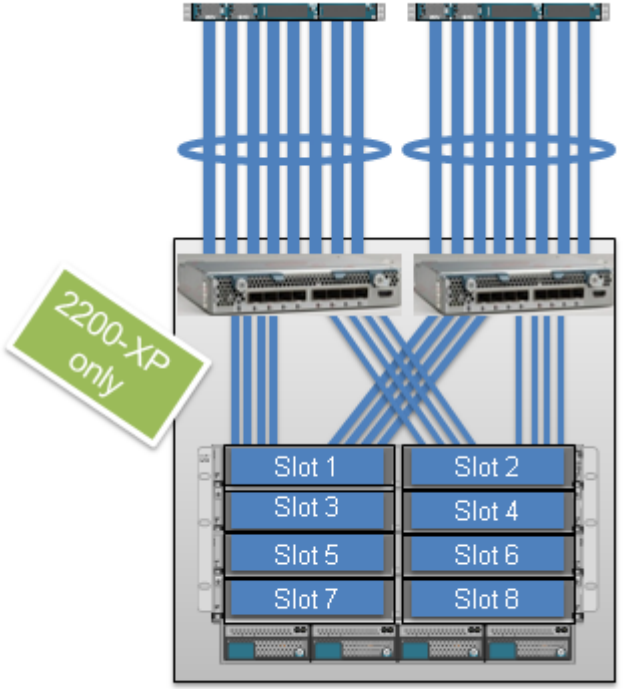
Chassis Connectivity Policy

160 Gb (Discrete Mode)



- 6100 to 2208
- 6200 to 2208

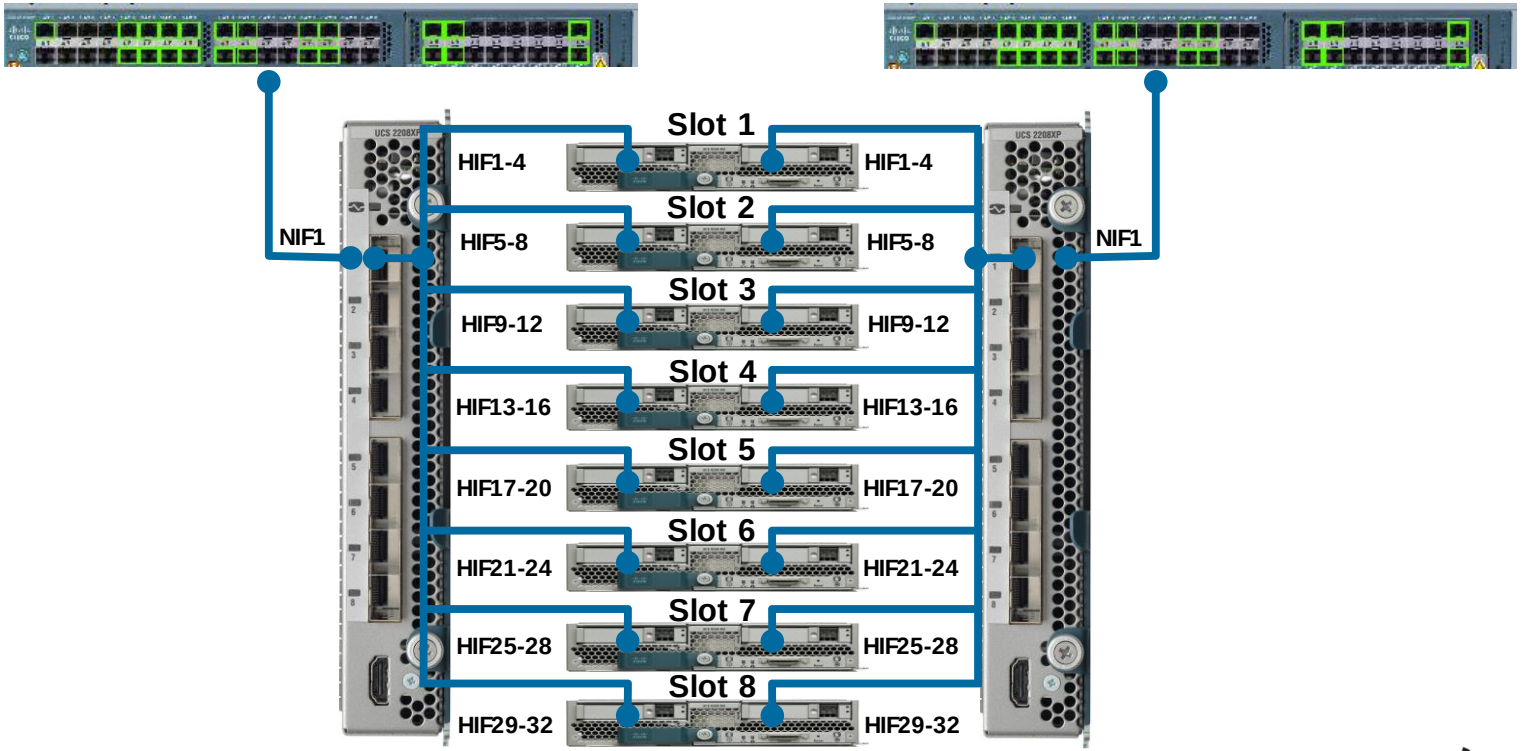
160 Gb (Port Channel Mode)



- 6200 to 2208

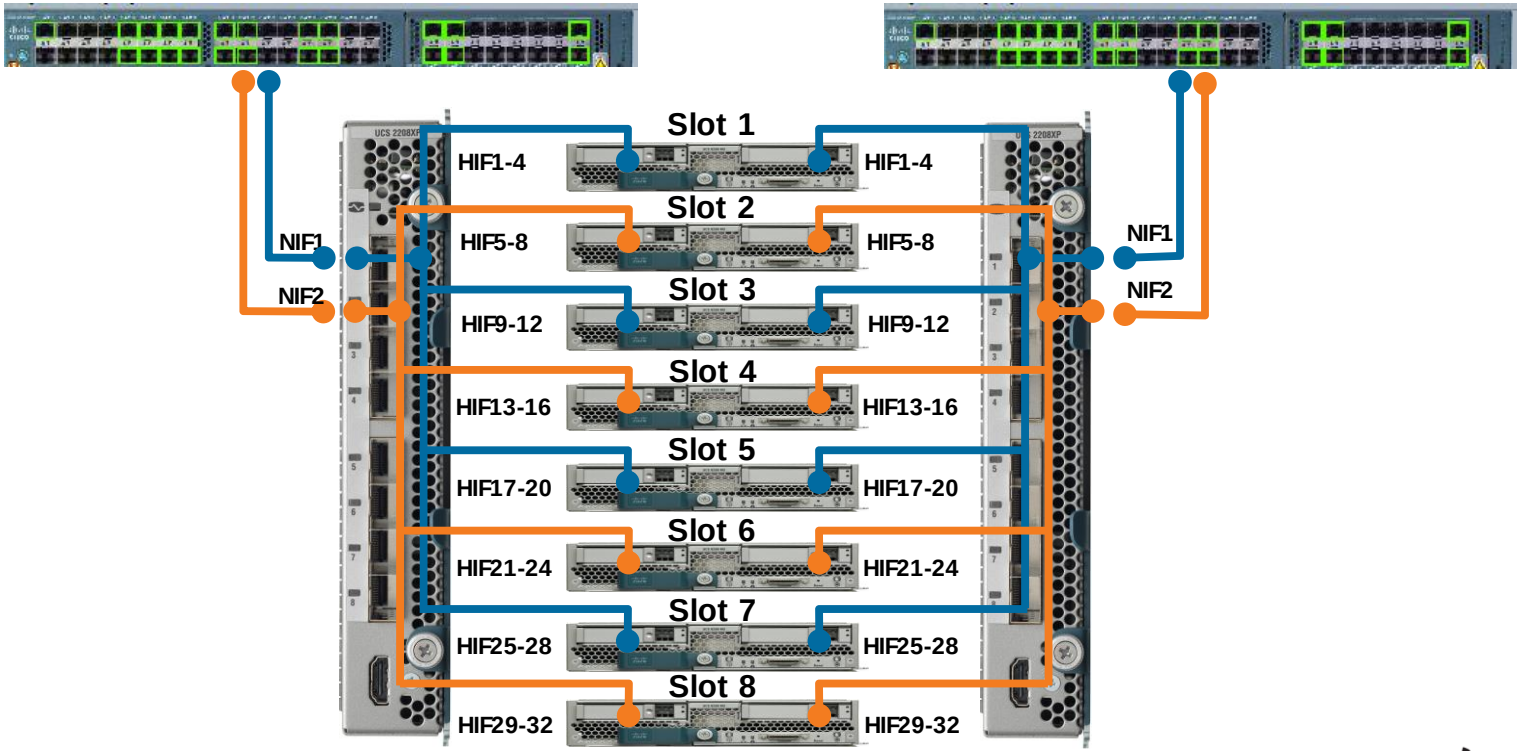
IO Module HIF to NIF Pinning

2208XP – 1 Link



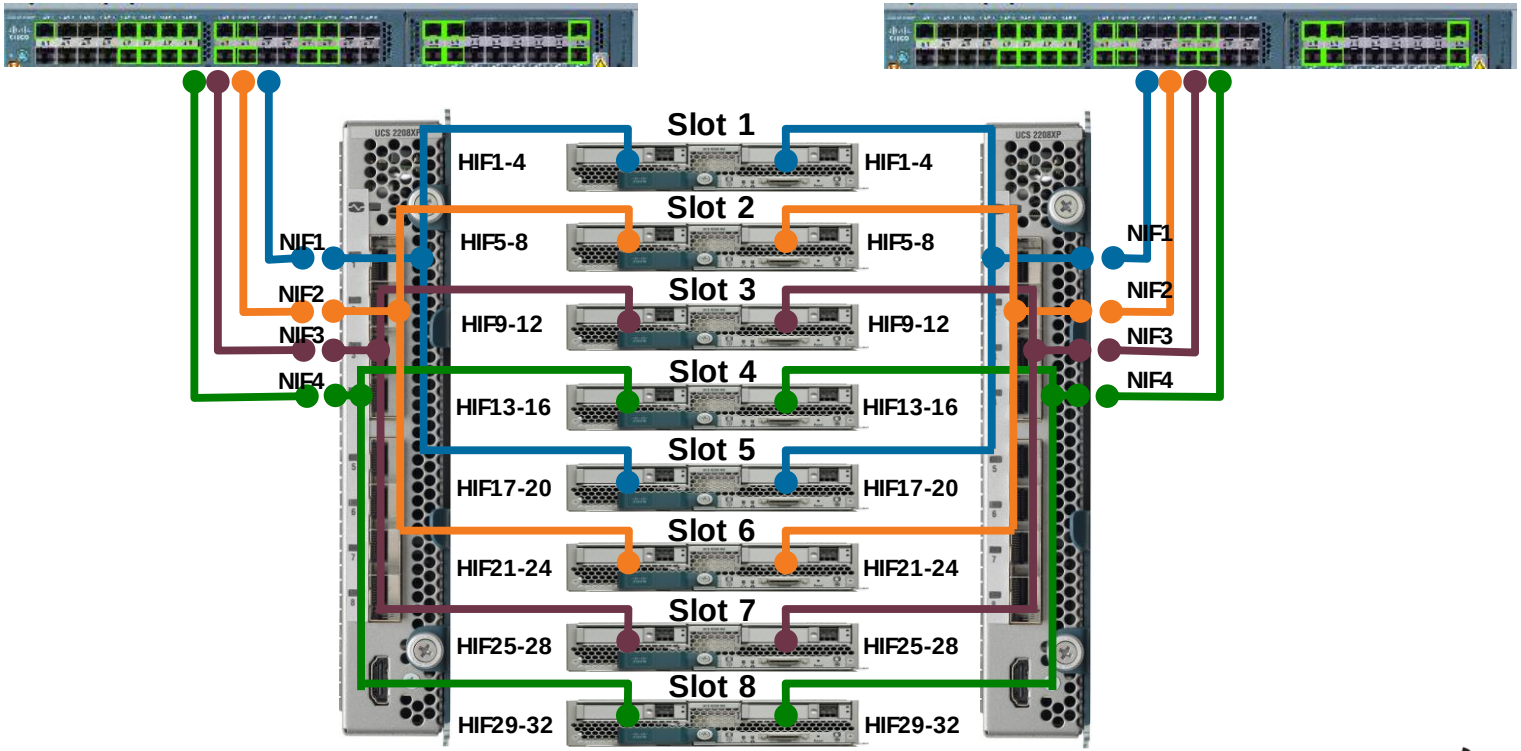
IO Module HIF to NIF Pinning

2208XP – 2 Links



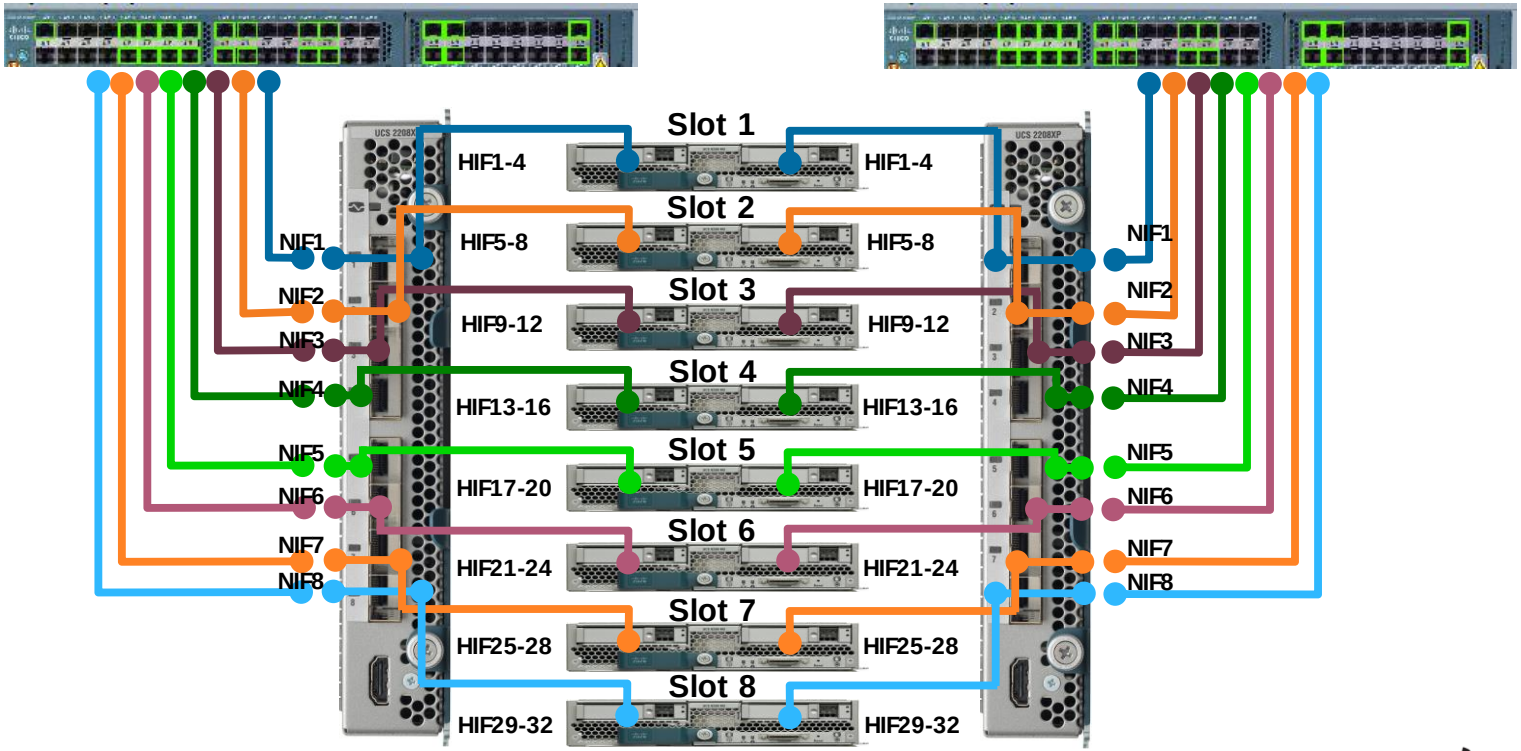
IO Module HIF to NIF Pinning

2208XP – 4 Links



IO Module HIF to NIF Pinning

2208XP – 8 Links



CLI Mapping

FEX Level command

- Output from “show fex detail” of the NXOS shell

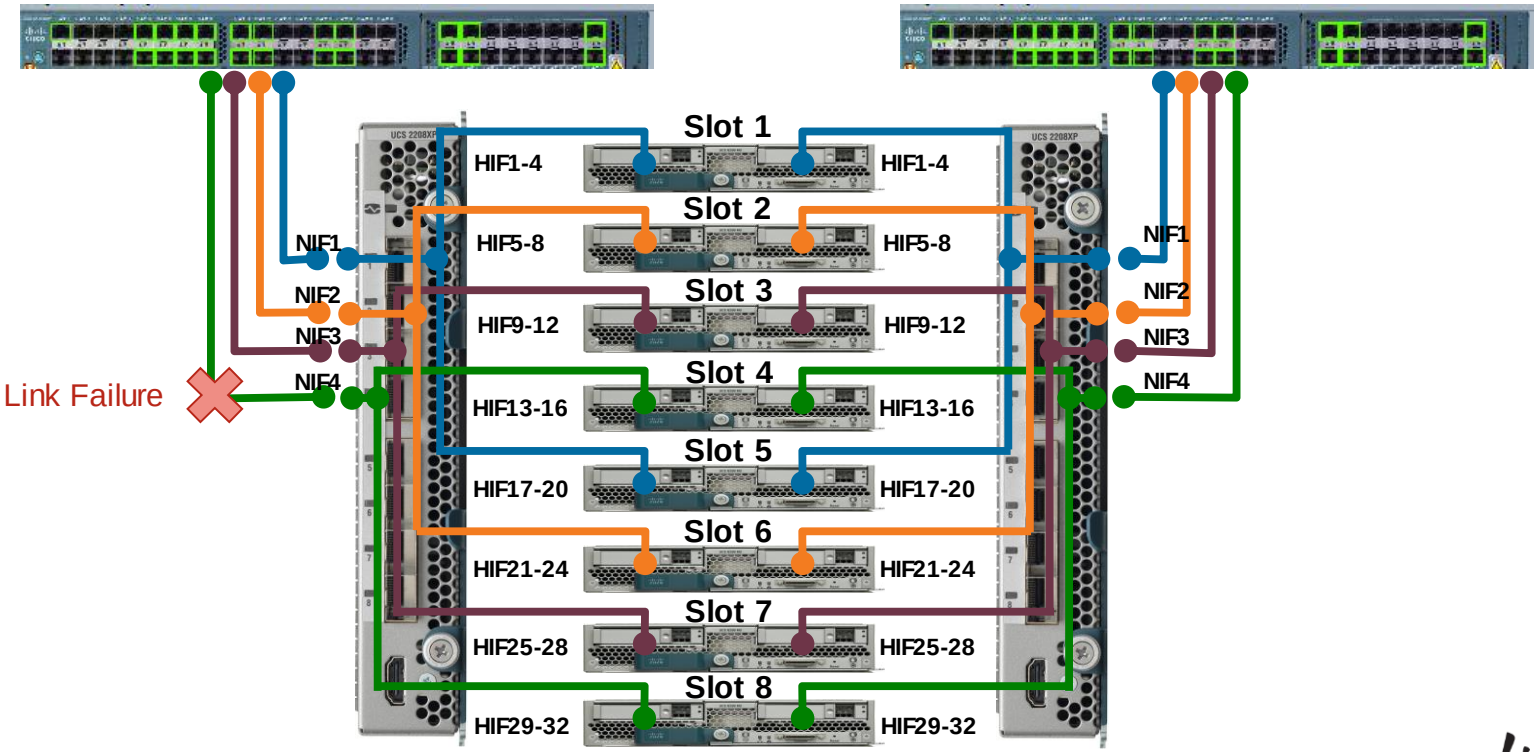
```
<output truncated>

Fabric interface state:
    Eth1/9 - Interface Up. State: Active
    Eth1/10 - Interface Up. State: Active
    Eth1/11 - Interface Up. State: Active
    Eth1/12 - Interface Up. State: Active
  Fex Port        State  Fabric Port
        Eth3/1/1    Up      Eth1/9
        Eth3/1/2    Up     Eth1/10
        Eth3/1/3  Down     Eth1/11
        Eth3/1/4    Up     Eth1/12
        Eth3/1/5    Up      Eth1/9
        Eth3/1/6  Down        None
        Eth3/1/7  Down     Eth1/11
        Eth3/1/8  Down     Eth1/12
        Eth3/1/9    Up     Eth1/12

<output truncated>
```

Interface Level command

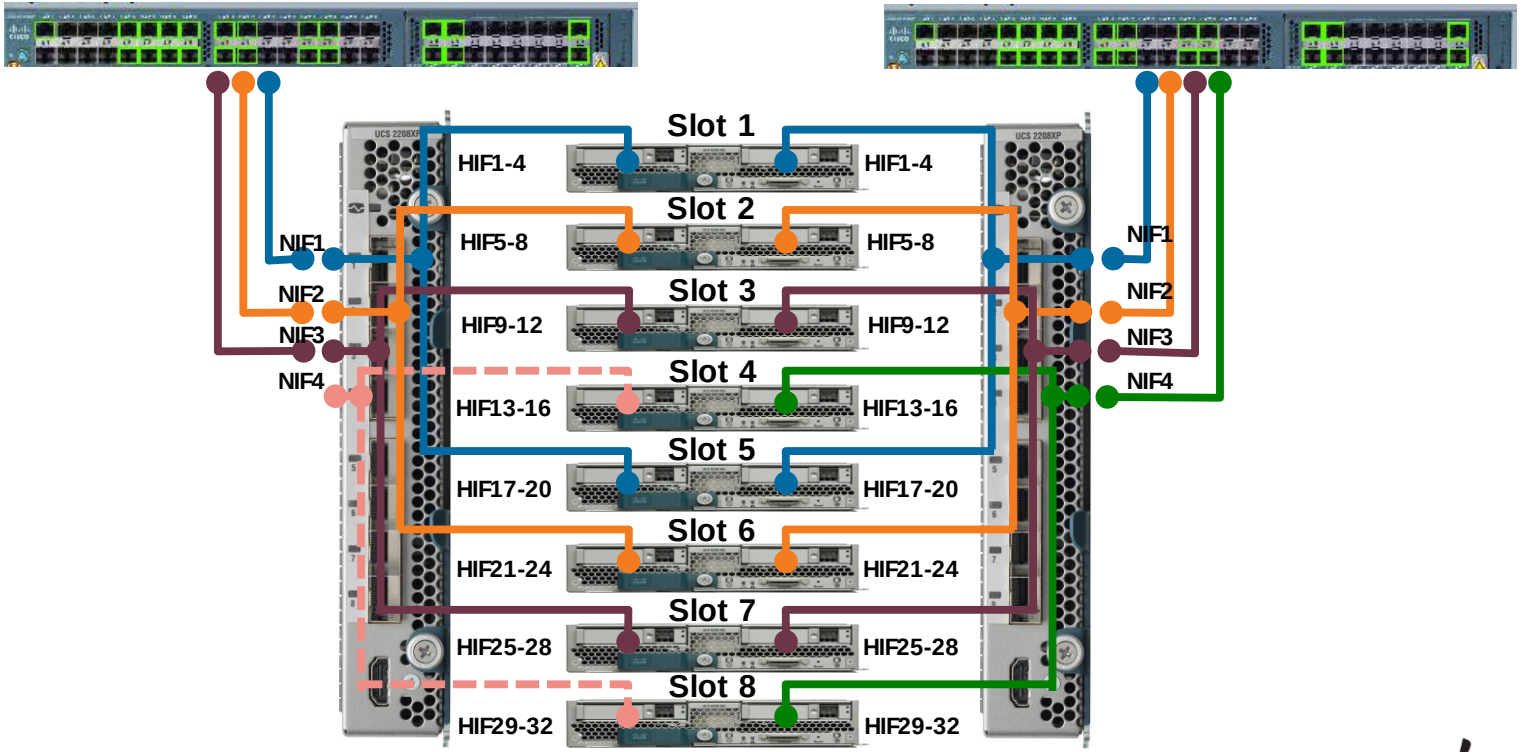
- Output from “show interface ethernet <mod/port> fex-intf

```
Fabric          FEX
Interface       Interfaces
---------------------------------------------
Eth1/9          Eth3/1/1      Eth3/1/5
```

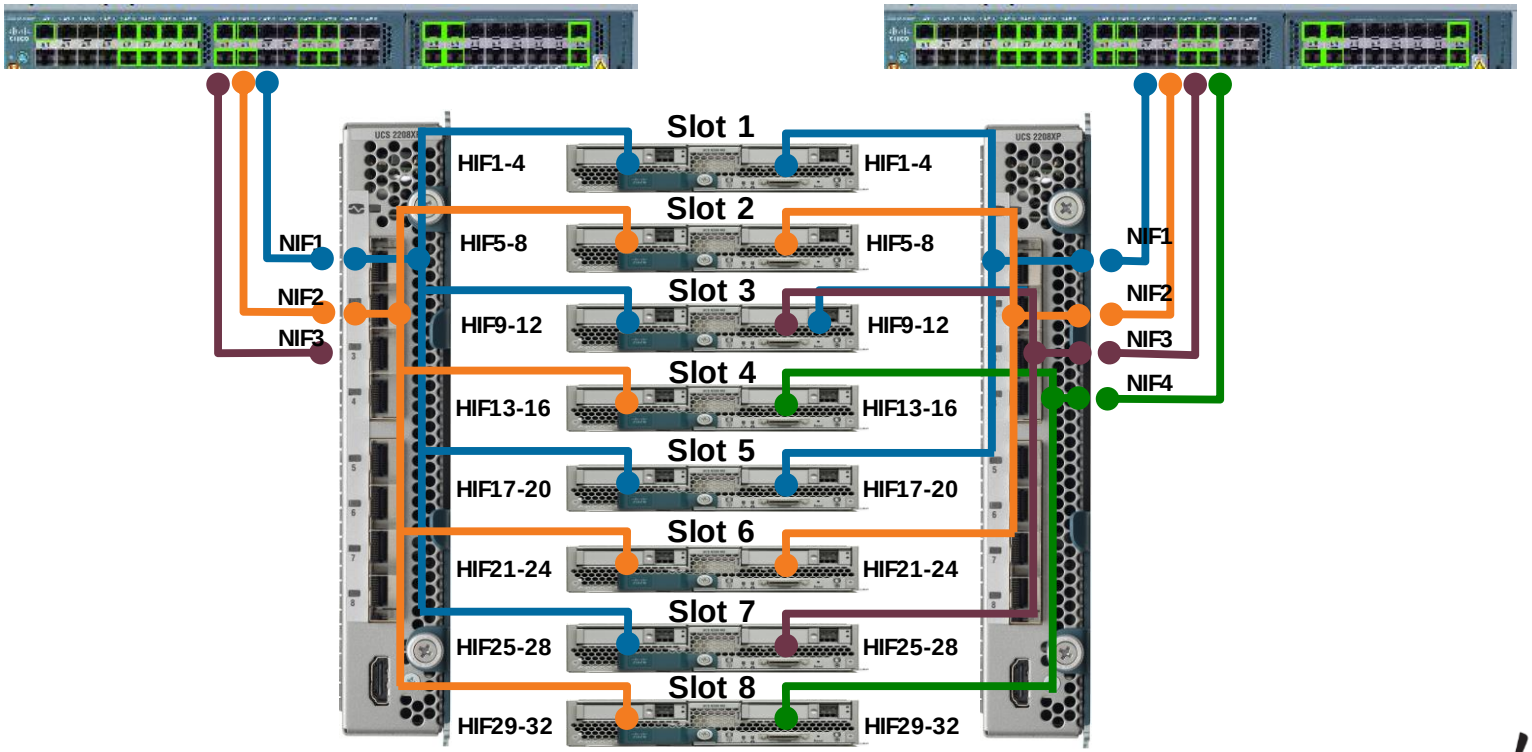
IOM Link Failure Scenario



IOM Link Failure Scenario



IOM Link Failure Scenario



Increased Bandwidth Access to Blades

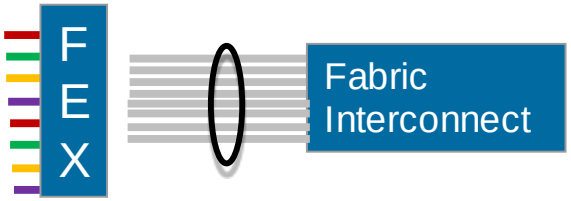
4 links, Discrete - Today

- Available bandwidth per blade – 10Gb
- Statically pinned to individual fabric links
- Deterministic Path

8 links, Discrete

- Available bandwidth per blade – 20Gb
- Statically pinned to individual fabric links
- Deterministic Path
- Guaranteed 10Gb to each blade

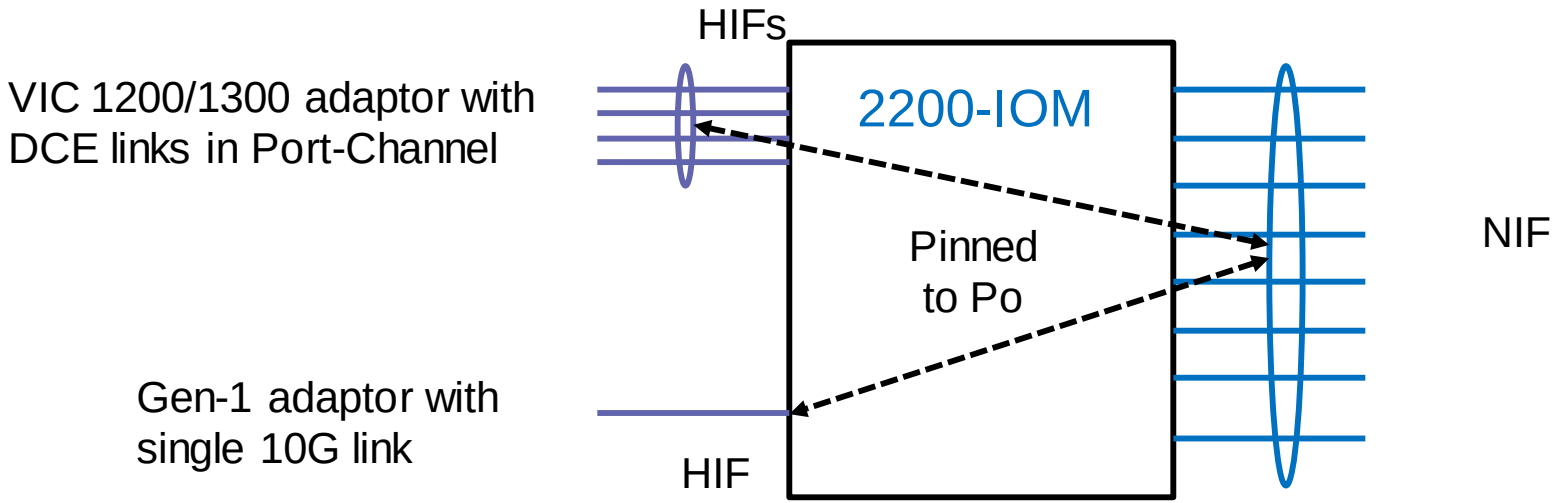
Up to 8 links, Port-channel



- Available bandwidth per blade – up to 160Gb
- Statically pinned to Port-channel
- Increased and shared bandwidth
- Higher Availability

Port-channel Pinning

- No slot based pinning
- No invalid link count for NIF ports (no “power of 2” rule)



Server Connectivity

Cisco live!

Cisco Virtual Interface Cards (VIC)

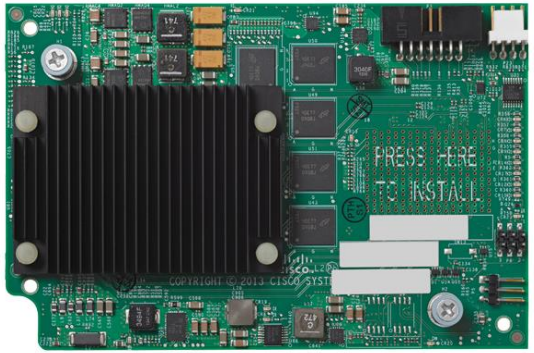
1st Gen

- M81KR, P81E
- 128 PCIe devices
- Dual 10Gb
- 16x PCIe Gen1

2nd Gen

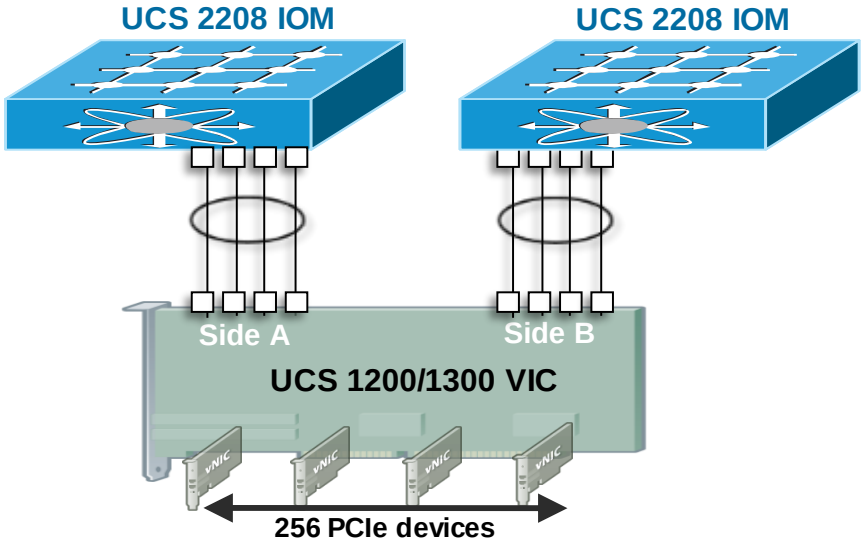
- 1240, 1280, 12xx
- 256 PCIe Device
- Dual 40Gb (4 x 10Gb)
- 16x PCIe Gen 2

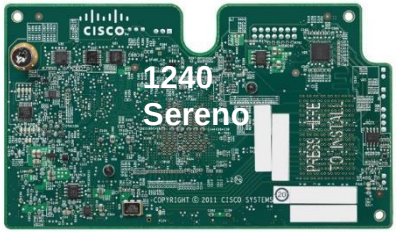
3rd Gen

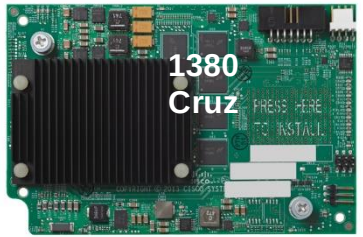


- 1340, 1380
- Dual 8x PCIe Gen 3
- Native 40Gb Support
- VXLAN & NVGRE
- RoCE

UCS Cisco 1200/1300 VIC Adapter

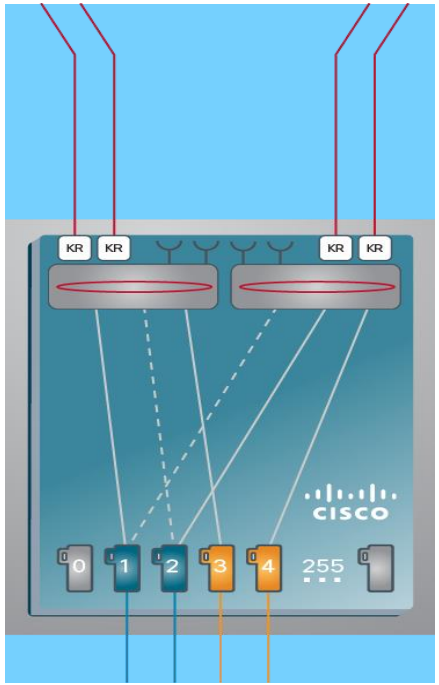


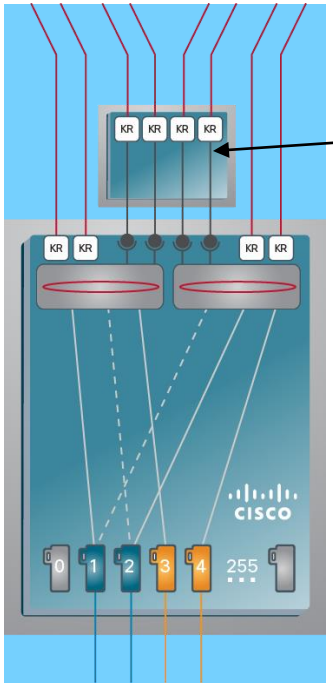




VIC 1240/1340 Plus Port Expander Card

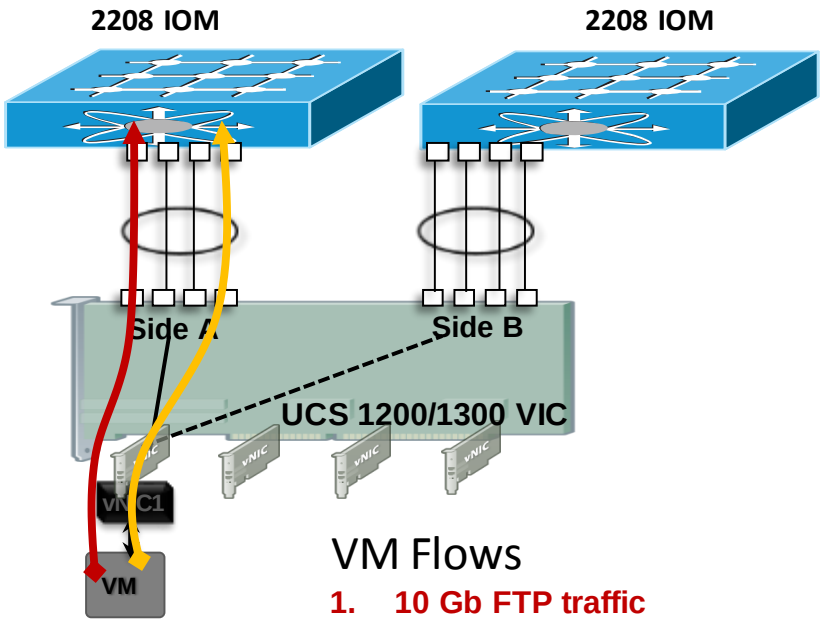
- Base option supports dual 2x10Gb





- Option to enable all port of ASIC (Sereno)
- Fits in the Mezzanine slot of B200 M3/M4
- Port Expander has no PCIe presence
- It is a “passive connector” device

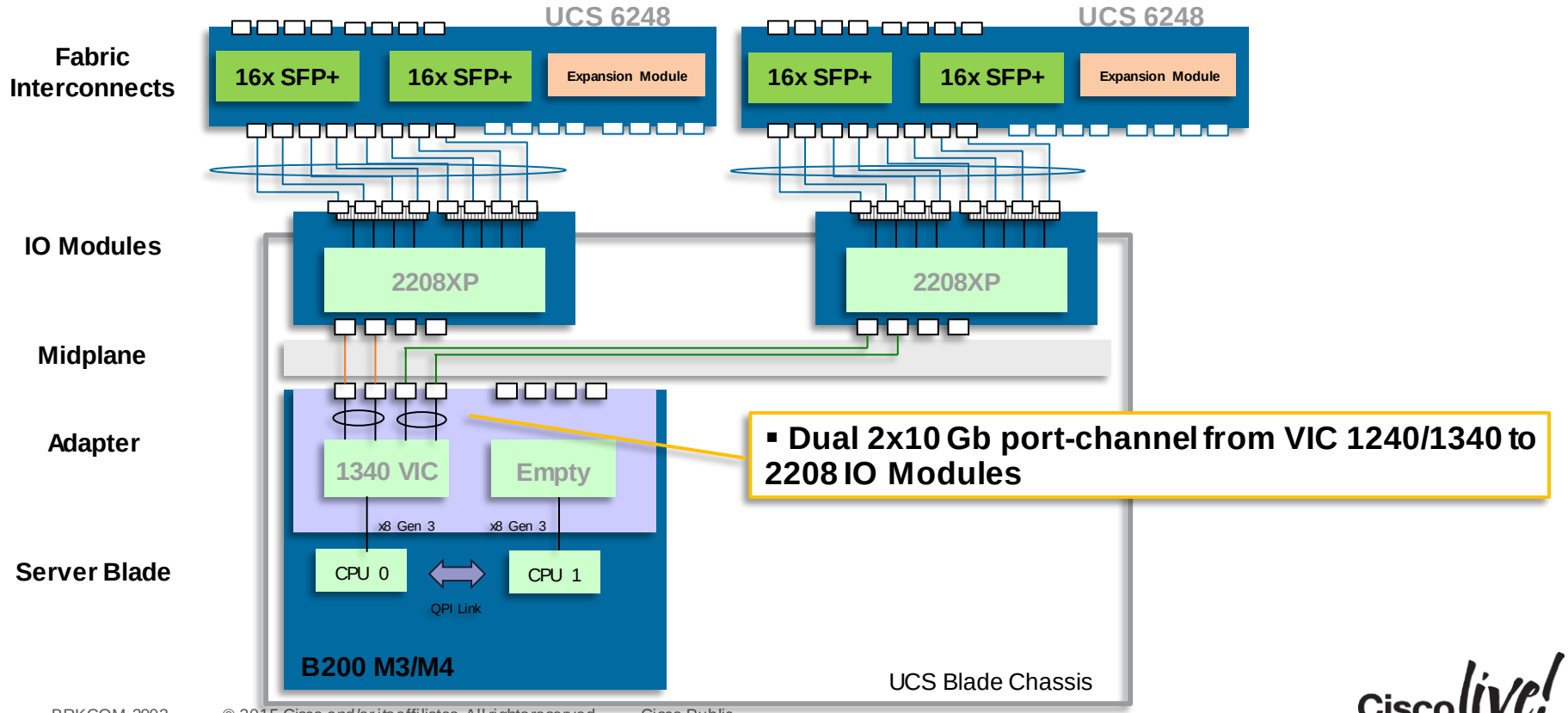
Connectivity IOM to Adapter



- Implicit Port-channel between VIC 1200/1300 and UCS 2200 IOM
- 7-Tuple Flow based hash, 10 Gbps per flow
- A vNIC is active on side A or B

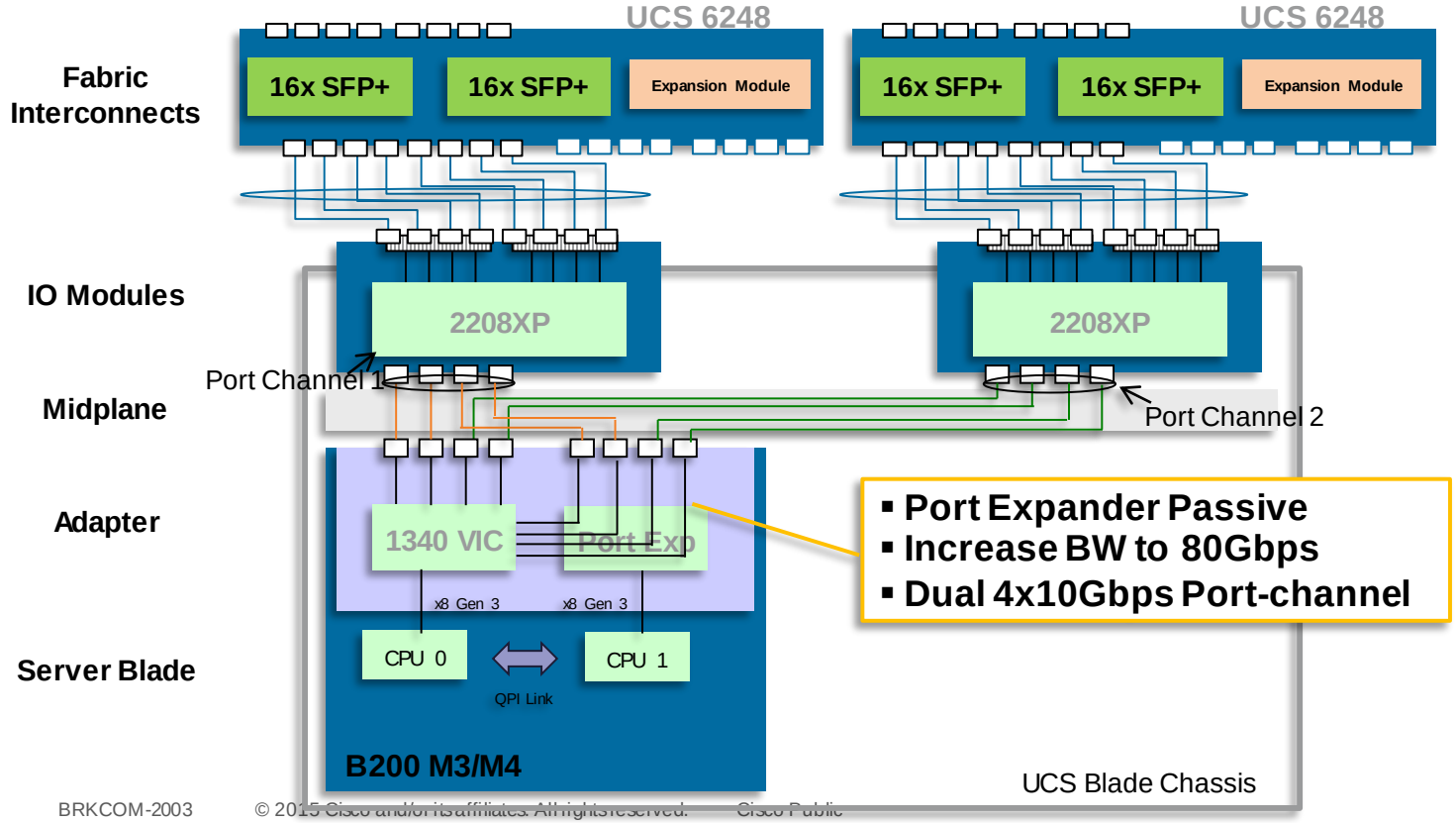
VIC 1240/1340 to IOM Connectivity

MLOM only

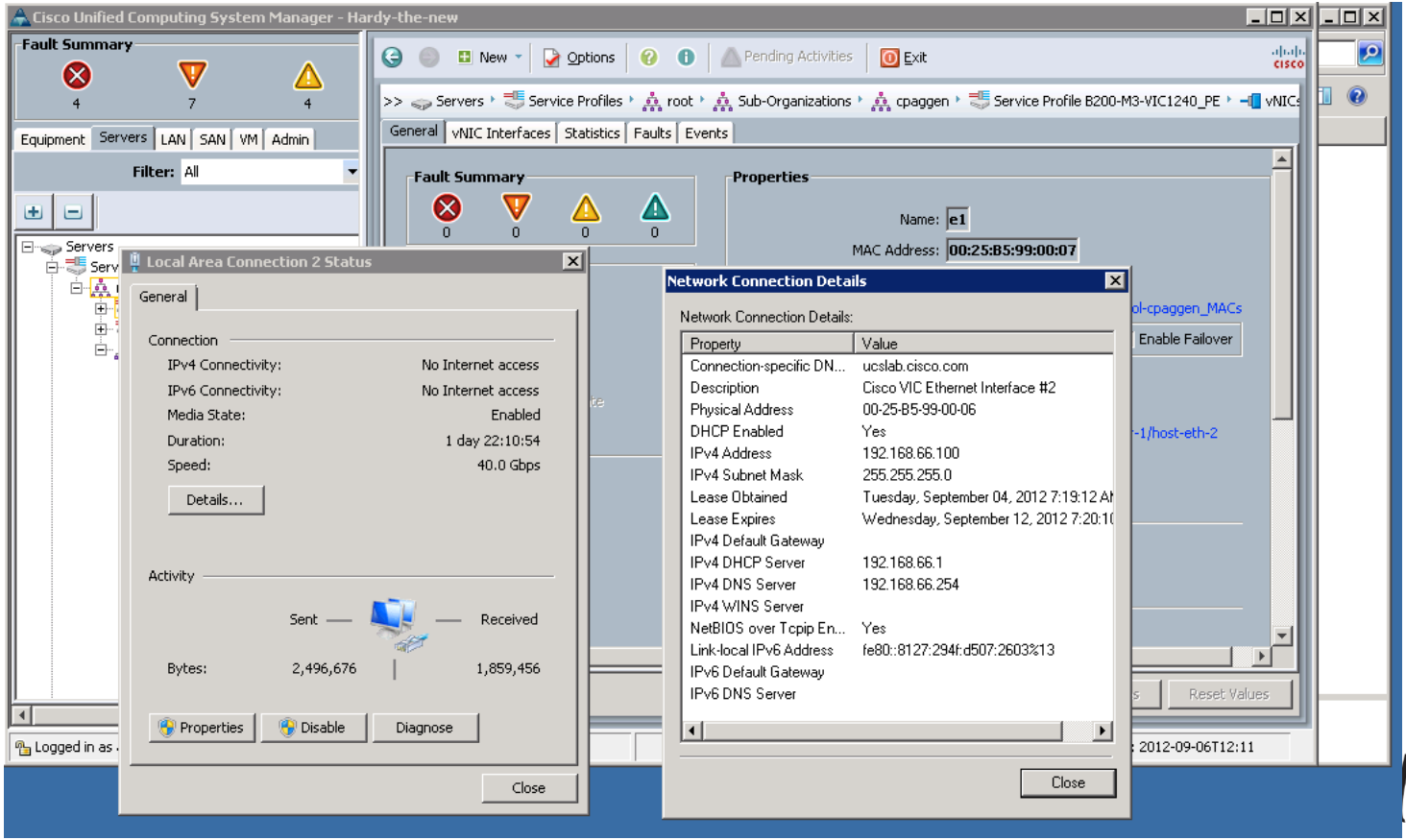


VIC 1240/1340 to IOM Connectivity

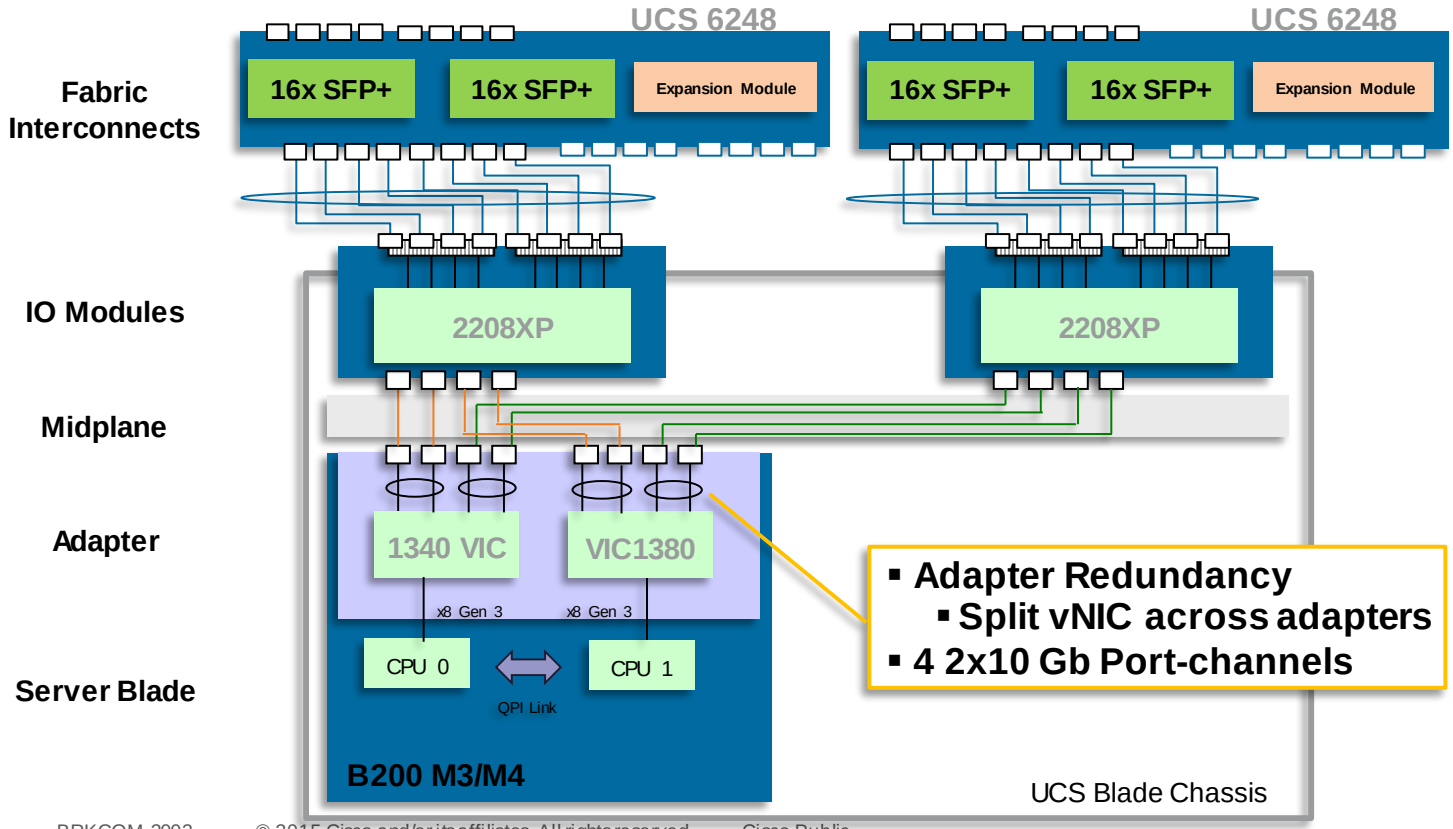
MLOM plus Port Expander



What Does The OS See?

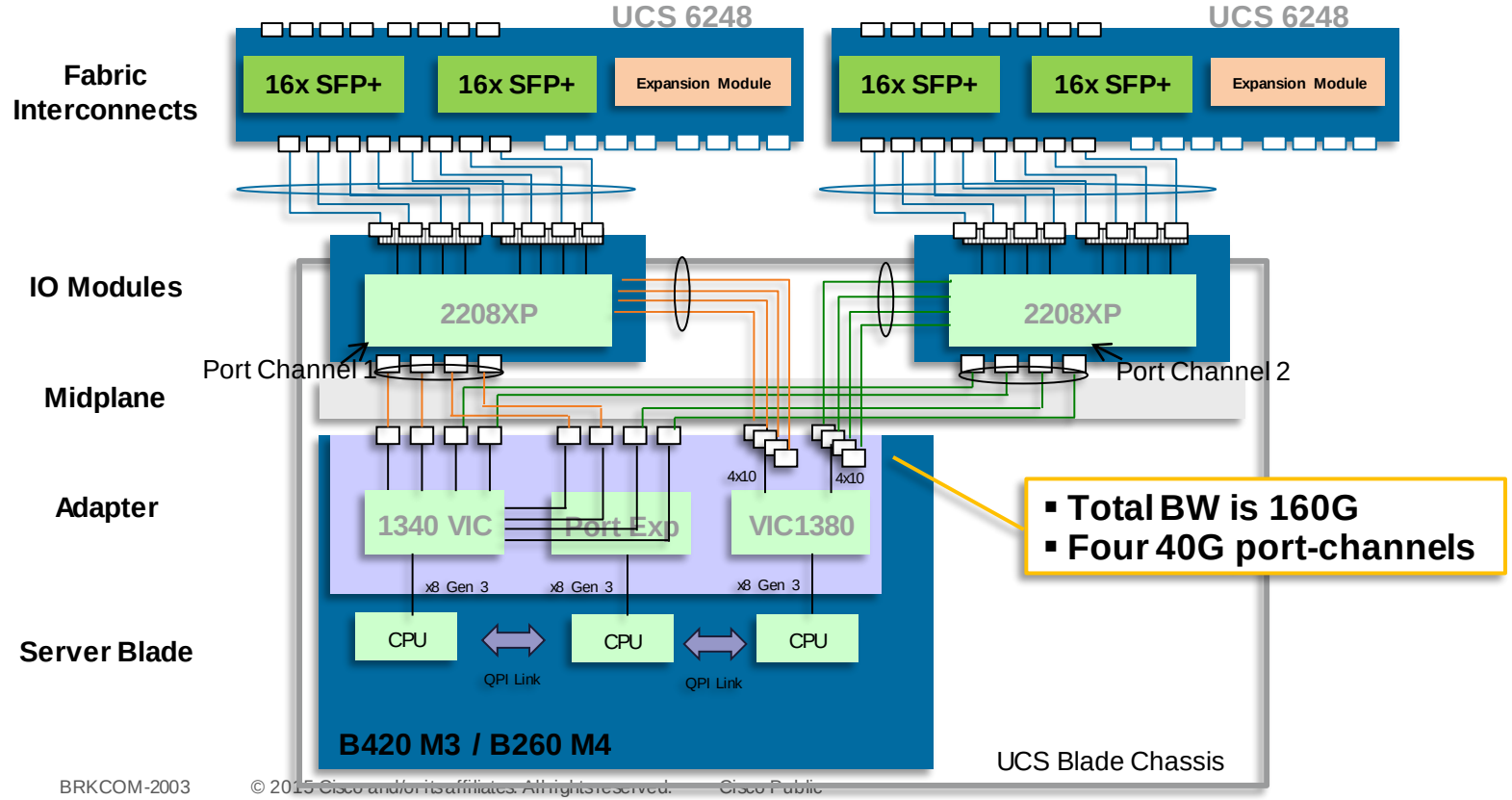


VIC 1x40 & 1x80 to IOM Connectivity



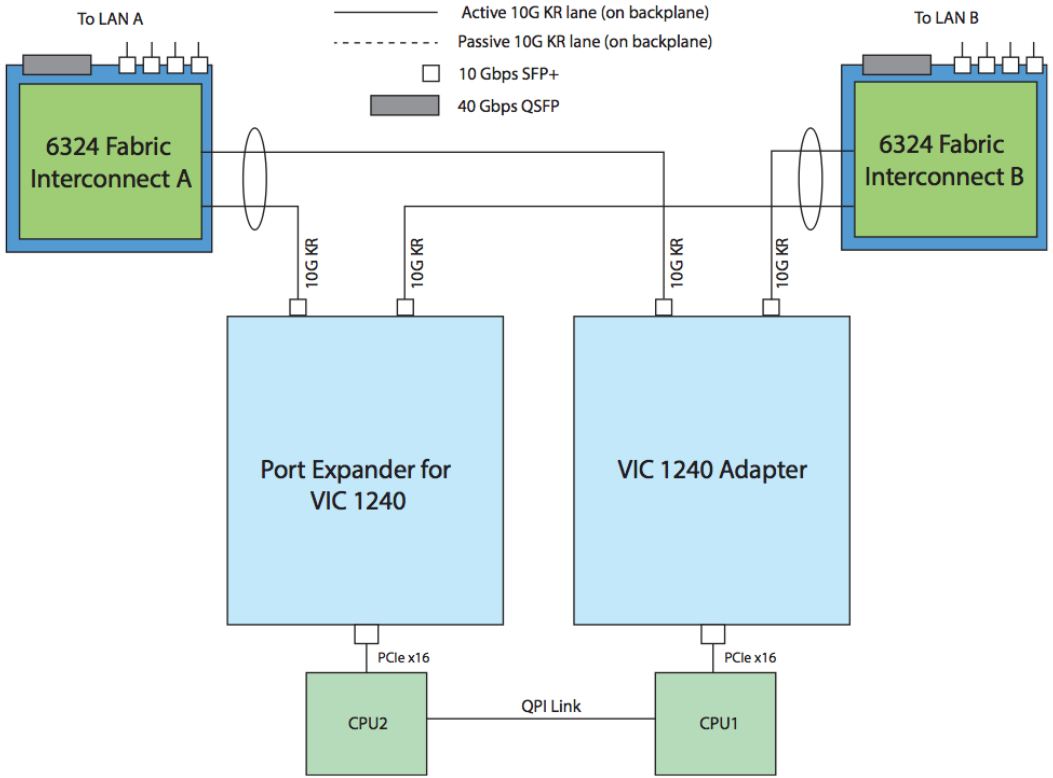
Full Width Blade to IOM Connectivity

MLOM, Port Expander, VIC1x80

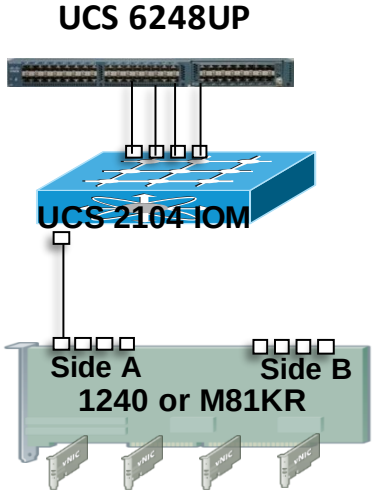


UCS Mini: Fabric to Server Connectivity

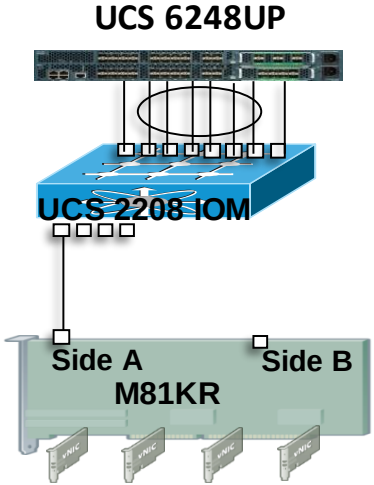
- Same server-side connectivity as the 2204XP IOM
- 20G per half width blade



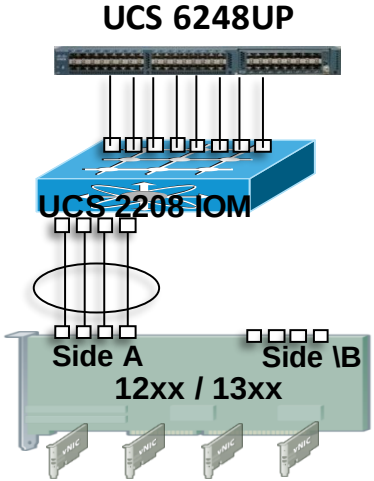
Topology Designs For Maximum Bandwidth



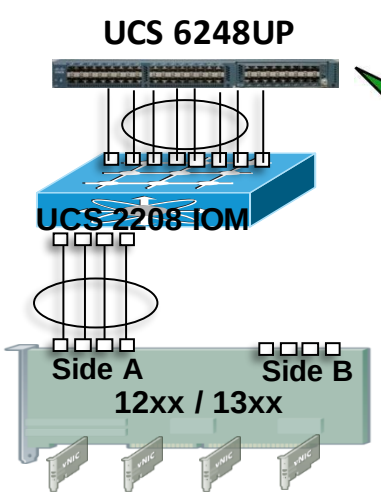
- Shared IOM uplink bandwidth of 10Gbps
- vNIC Burst up to 10Gbps



- Shared IOM uplink bandwidth of 80Gbps
- vNIC Burst up to 10Gb

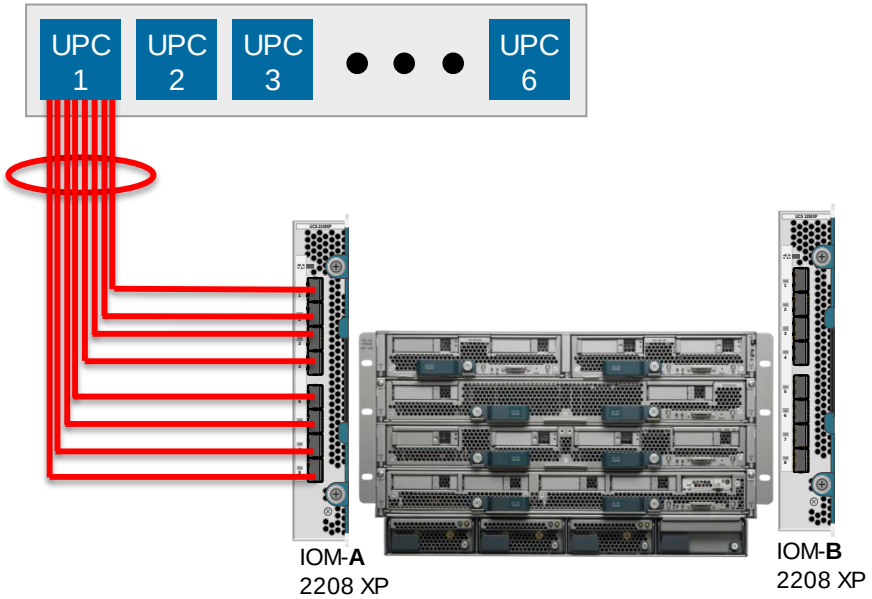


- Dedicated IOM uplink bandwidth of 10Gbps
- vNIC Burst up to 10Gbps

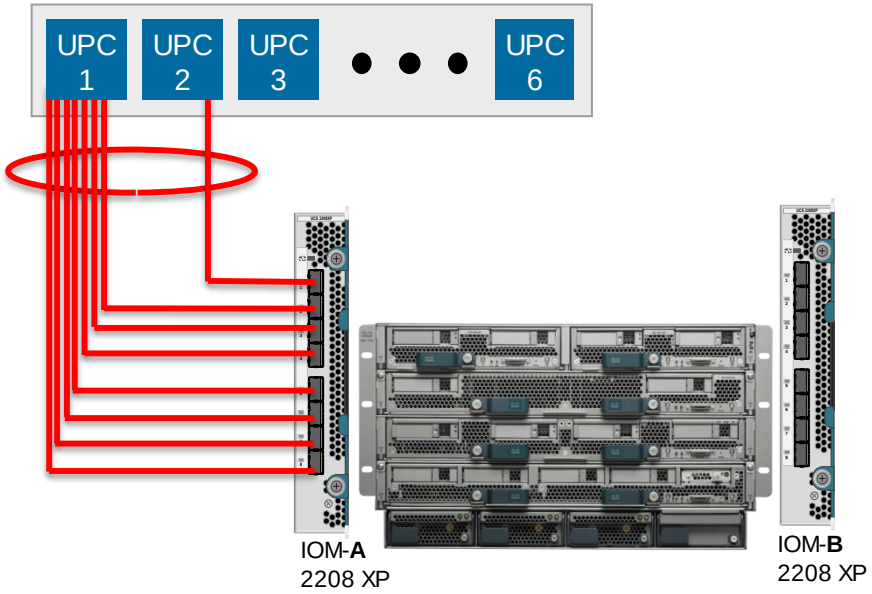


- Shared IOM uplink bandwidth of 80Gbps

Fabric Interconnect VIF Calculation



- Recommended
- Maximise number of available VIFs to the host

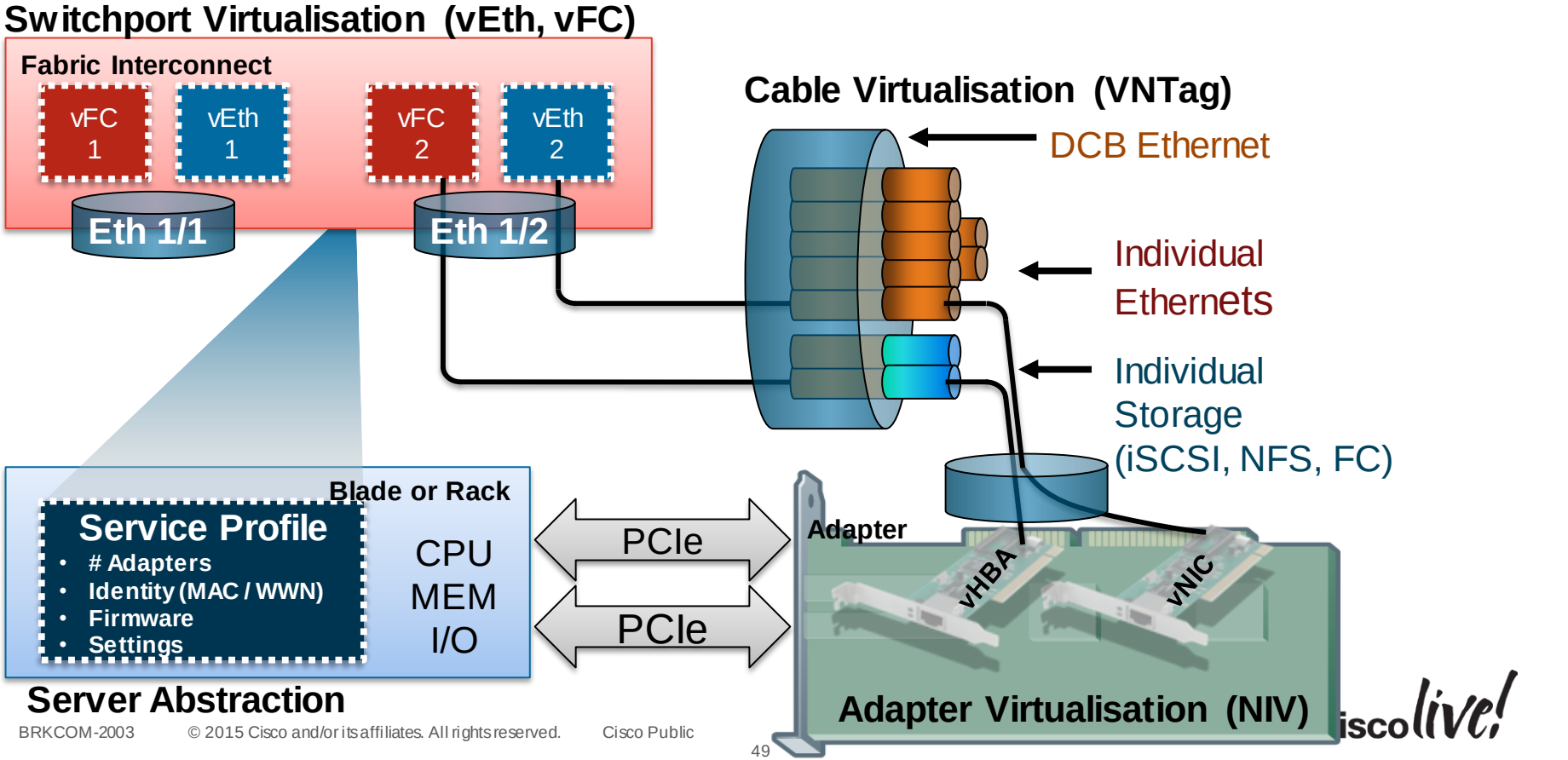


- Not recommended
- Minimal number of VIFs to the host

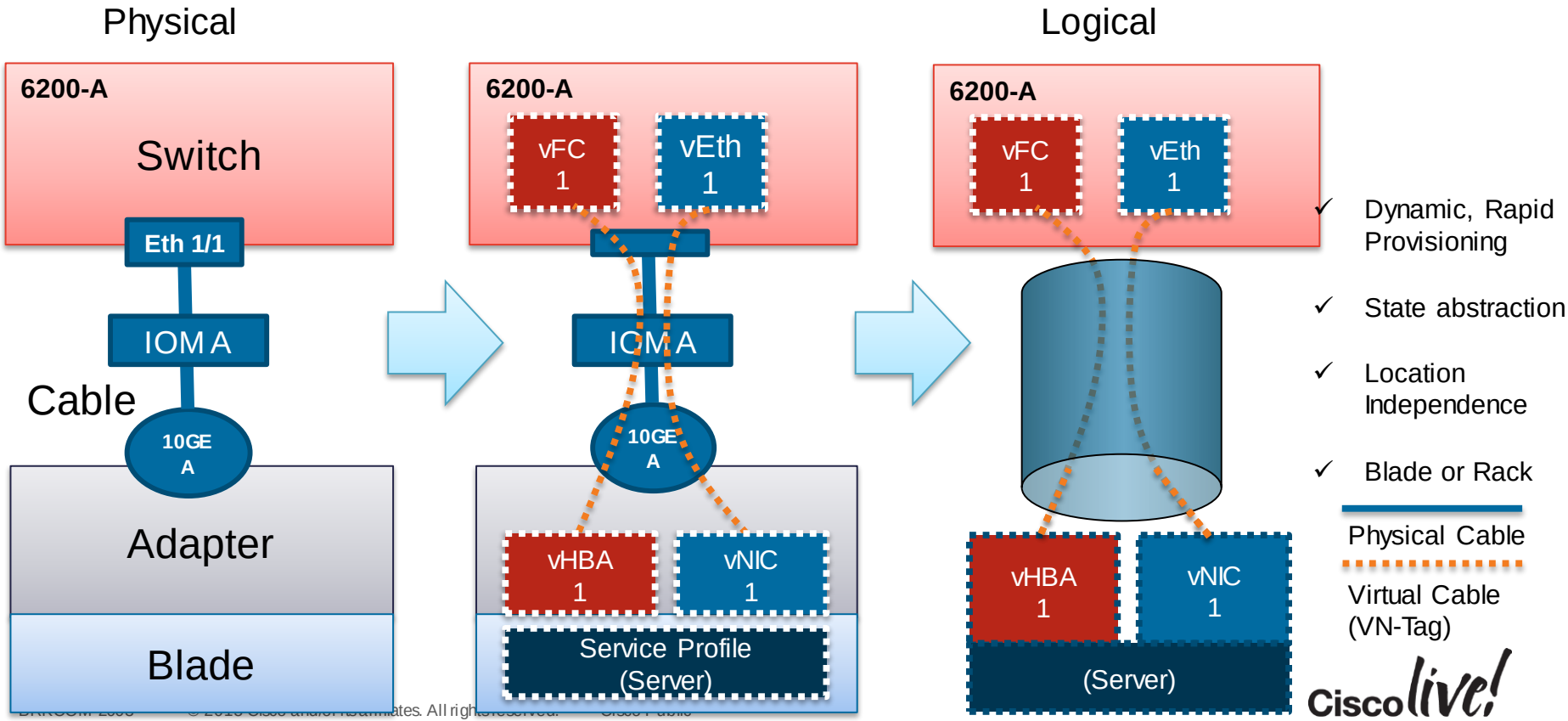
Virtual Interfaces

Cisco live!

Cisco UCS: Infrastructure Virtualisation



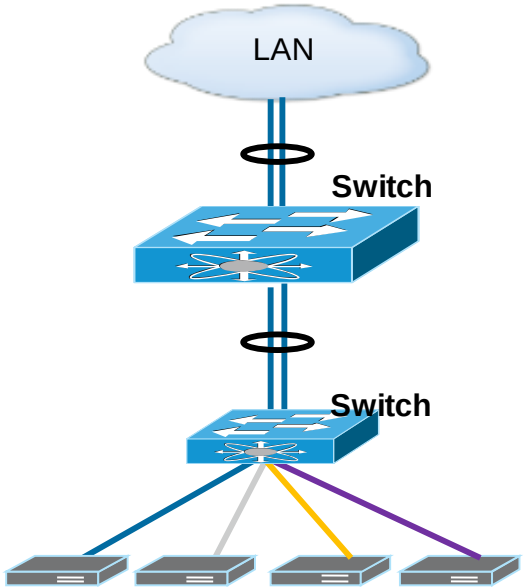
Abstracting the Logical Architecture



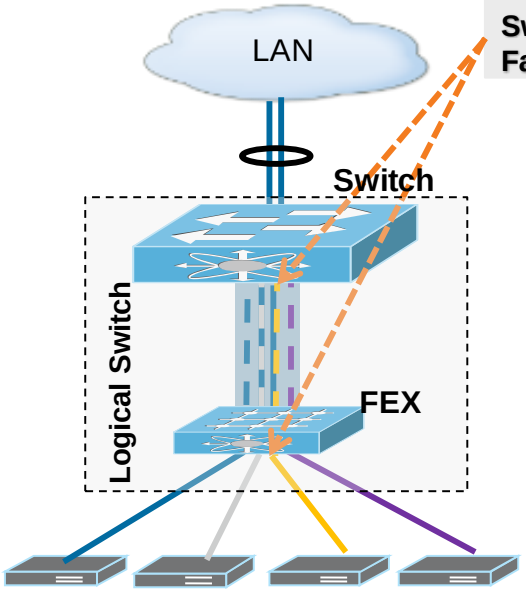
Fabric Extension (FEX) Concept

Virtualising the Network Port

Legacy multi-tier architecture

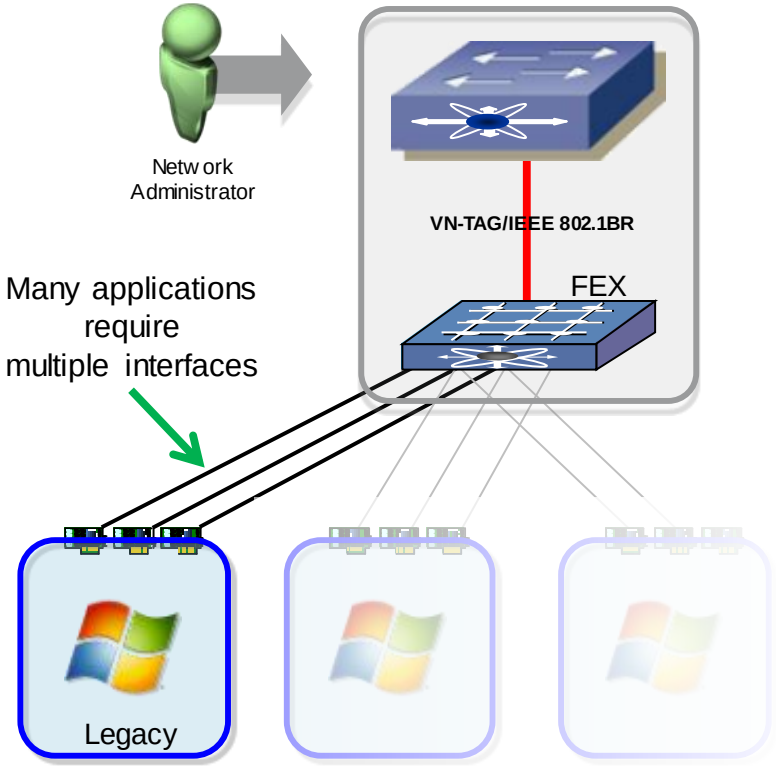


FEX architecture



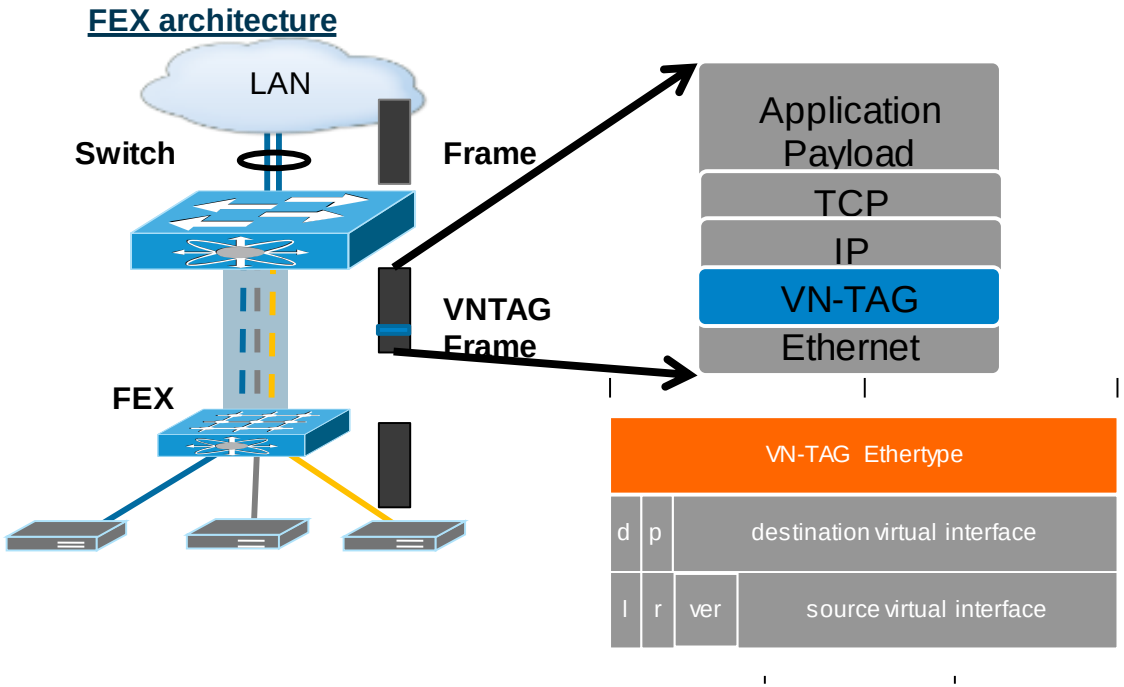
Switch port extended over Fabric Extender

Fabric Extender Evolution

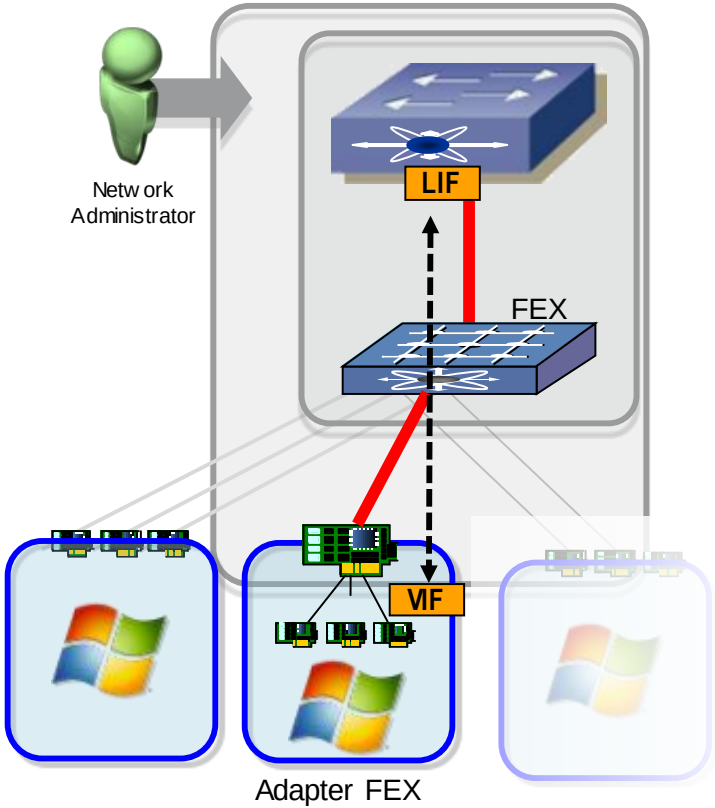


- Cisco VN-TAG is the pre-standard to IEEE 802.1BR Port Extension
- The 802.1BR Architecture provides the ability to extend the bridge (switch) interface to downstream devices

VN-TAG



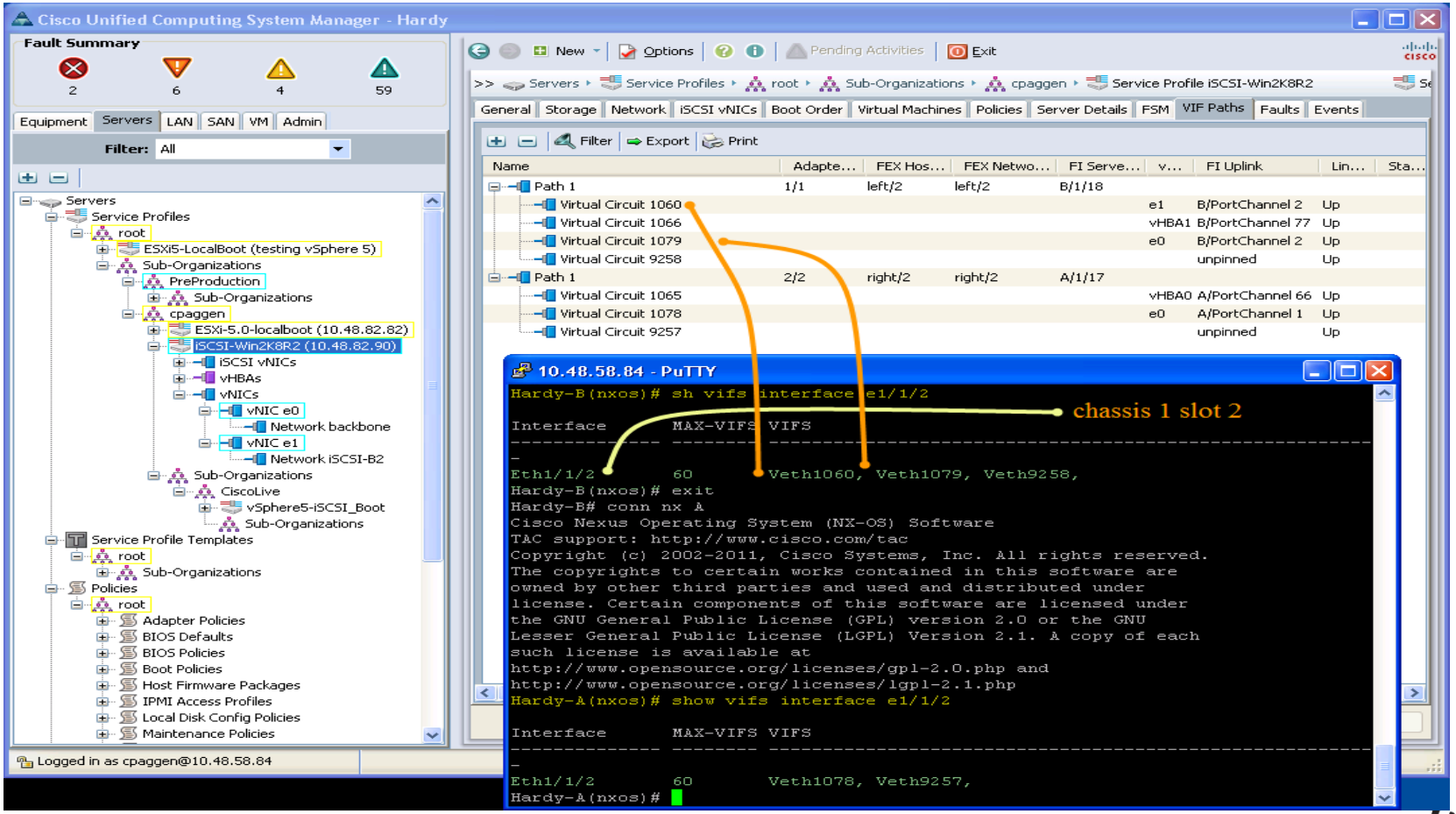
Fabric Extender Evolution – Virtual Interfaces



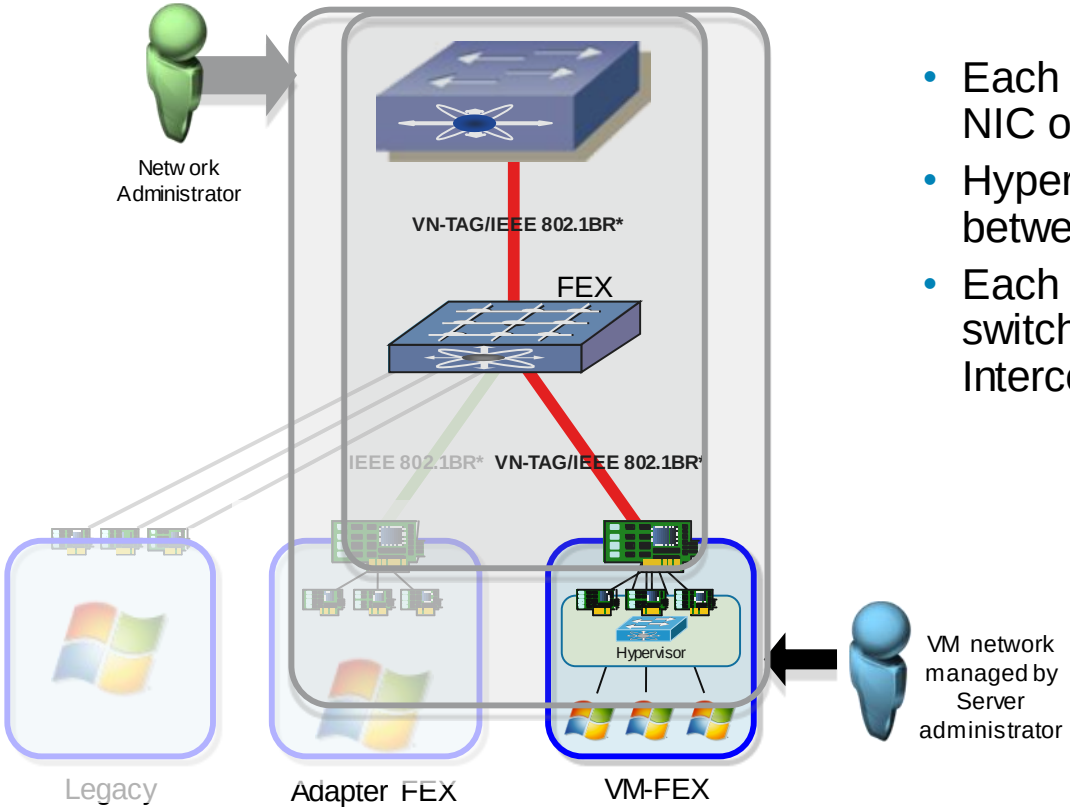
- 802.1BR associates the Logical Interface (LIF) to a Virtual Interface (VIF)

VN-Tag at the Adapter (Mezz Card) Level

```
Hardy-the-new-A# connect adapter 1/1/1
adapter 1/1/1 # connect
adapter 1/1/1 (top):1# attach-mcp
adapter 1/1/1 (mcp):1# vnic
vnic id       : internal id of vnic, use for other vnic cmds
vnic name/mac : ucsm provisioned name (-n) or mac address (-m)
vnic type     : enet=ethernet, enet_pt=dynamic ethernet, fc=fcoe
vnic bb:dd.f  : host pci bus/device/function id
vnic state    : state of vnic
lif           : internal logical if id, use for other lif/vif cmds
lif state     : state of lif
vif uif       : bound uplink 0 or 1, =:primary, -:secondary, >:current
vif ucsm      : ucsm id for this vif
vif idx       : switch id for this vif
vif vlan      : default vlan for traffic
vif state     : state of vif
------------------------------------ --------- --------------------------
               v n i c                  l i f            v i f
id  name           type    bb:dd.f state lif state uif  ucsm   idx vlan state
--- -------------- ------- ------- ----- --- ----- --- ----- ----- ---- -----
  5 vnic_1         enet    08:00.0 UP      2 UP    - 0   970   860    1 UP
                                                   =>1   969   873    1 UP
  6 vnic_2         enet    09:00.0 UP      3 UP    - 0   972   861    1 UP
                                                   =>1   971   874    1 UP
  7 vnic_3         enet    0a:00.0 UP      4 UP    - 0   974   862    1 UP
                                                   =>1   973   875    1 UP

adapter 1/1/1 (mcp):2#
```



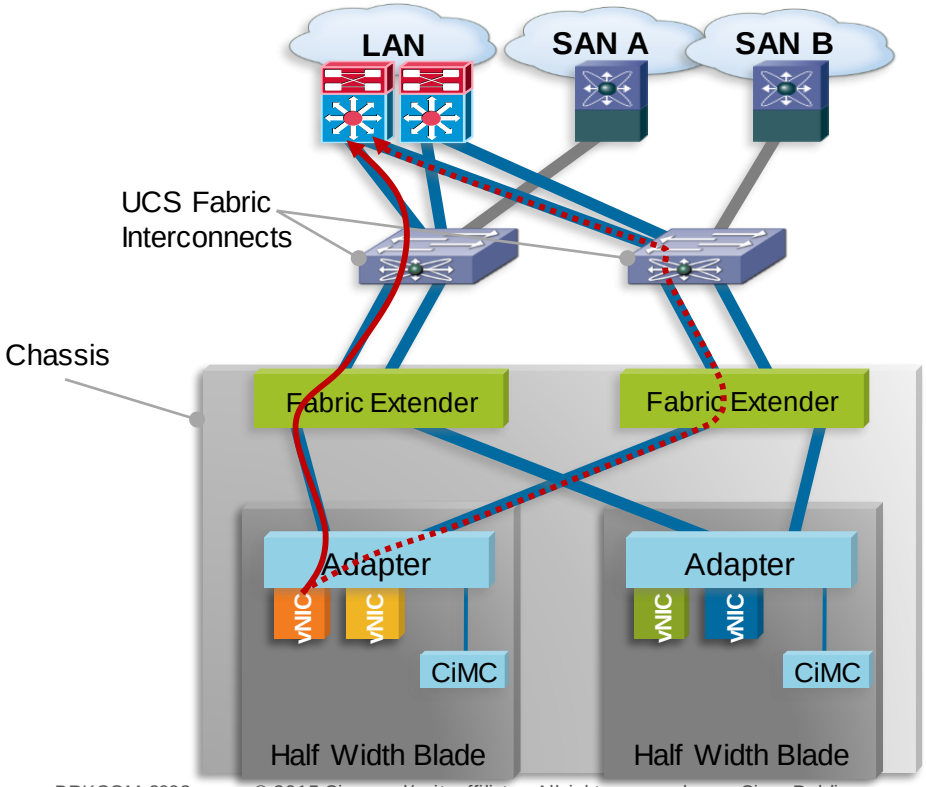
Fabric Extender Evolution (VM-FEX)



- Each VM assigned dedicated NIC on ESXi hosts PCIE bus
- Hypervisor arbitrates mapping between VM and PCI vNIC
- Each VM gets a dedicated switch port on the Fabric Interconnect

Fabric Failover

End Host Mode (only)



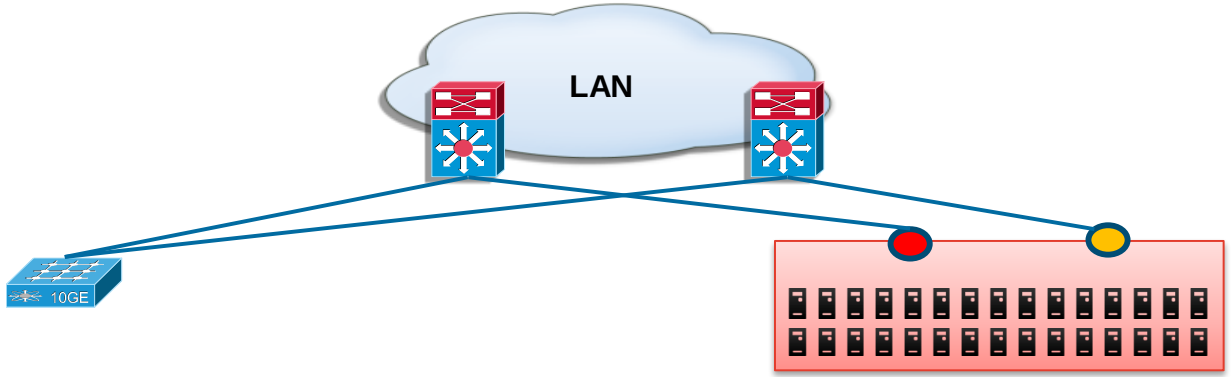
- Fabric provides NIC failover capabilities chosen when defining a service profile
- Traditionally done using NIC bonding driver in the OS
- Provides failover for both unicast and multicast traffic
- Works for any OS on bare metal and hypervisors

```
UCS-6200-A /chassis/server/adapter/host-eth-if # show vif

VIF:
     ID          Fabric ID Transport Tag    Status       Oper State
     ----------- --------- --------- ----- ------------ ----------
           1201 A         Ether           0 Allocated    Active
           1202 B         Ether           0 Allocated    Passive
```

Fabric Forwarding - Ethernet

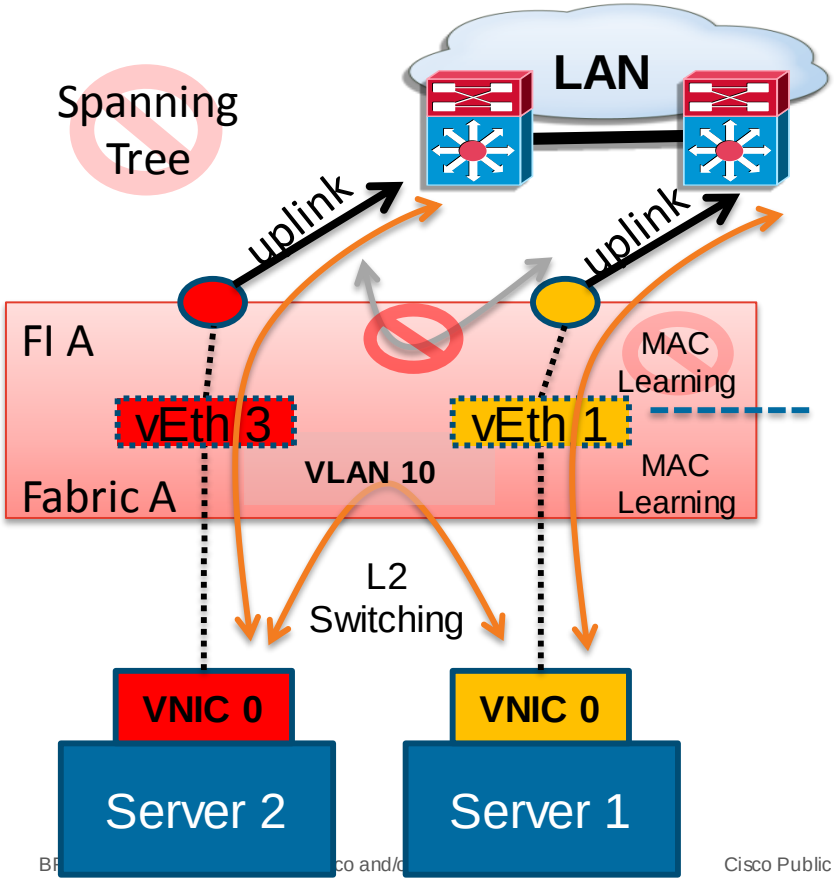
Cisco live!

Ethernet Fabric Forwarding Mode of Operations



- Switch mode: User configurable
  - Fabric Interconnects behave like regular ethernet switches
  - STP parameters are lock
  - VLAN/Mac based forwarding
- End-host mode (EHM): Default mode
  - No spanning-tree protocol (STP)
    - Active/Active for all links and VLANs
  - Port definitions
  - Policy based forwarding
    - No unknown unicast forwarding

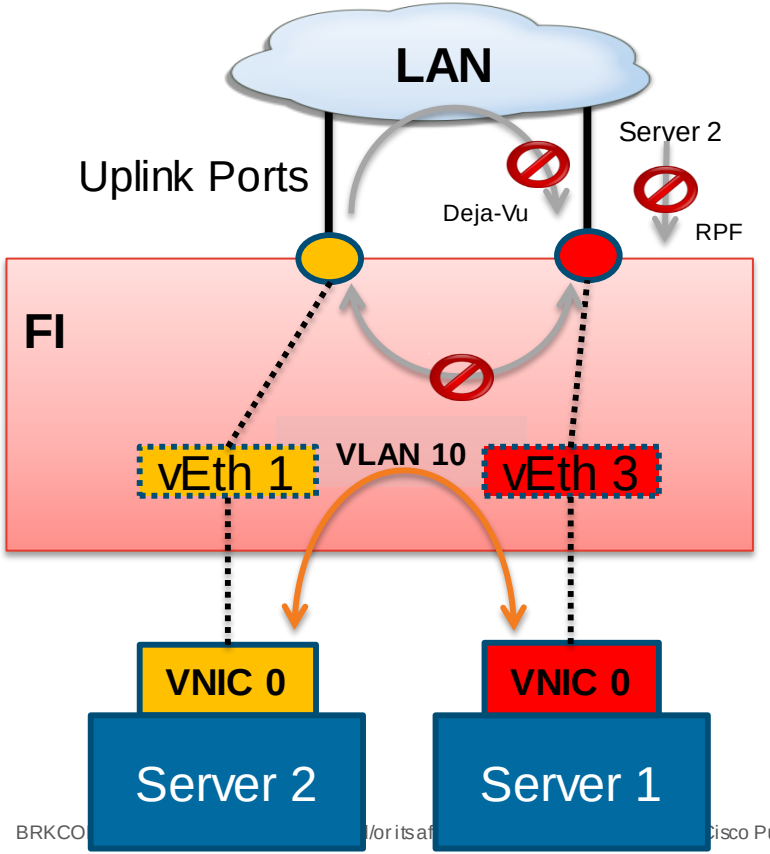
End Host Mode



- Completely transparent to the network
  - Presents itself as a bunch of hosts to the network
- No STP – simplifies upstream connectivity
- All uplinks ports are forwarding – never blocked

End Host Mode

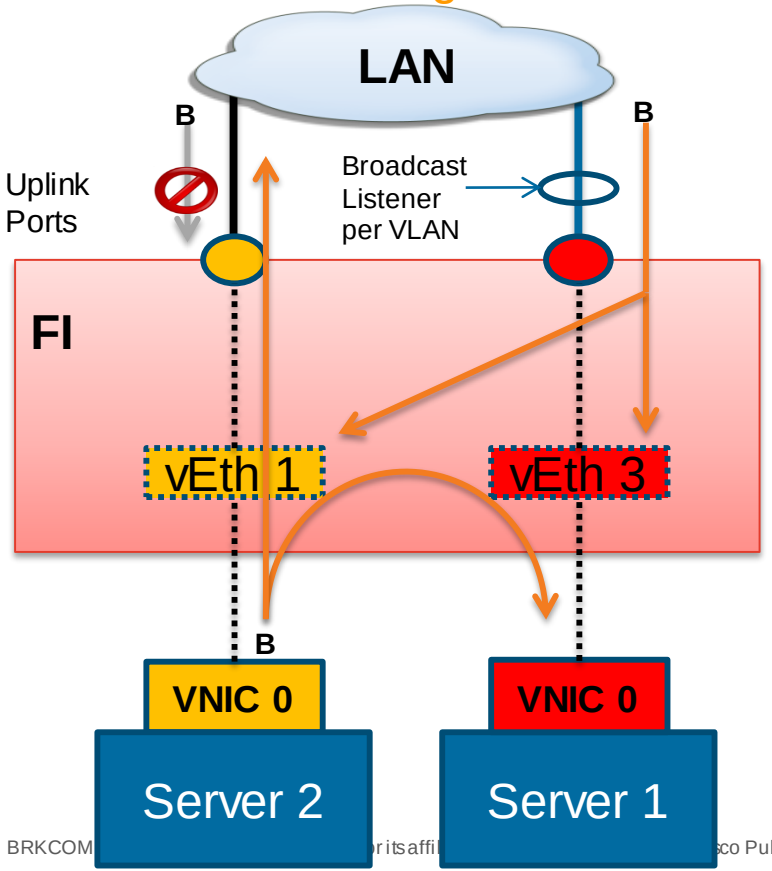
Unicast Forwarding



- MAC/VLAN plus policy based forwarding
  - Server pinned to uplink ports
- Policies to prevent packet looping
  - déjà vu check
  - RPF
  - No uplink to uplink forwarding
- No unknown unicast

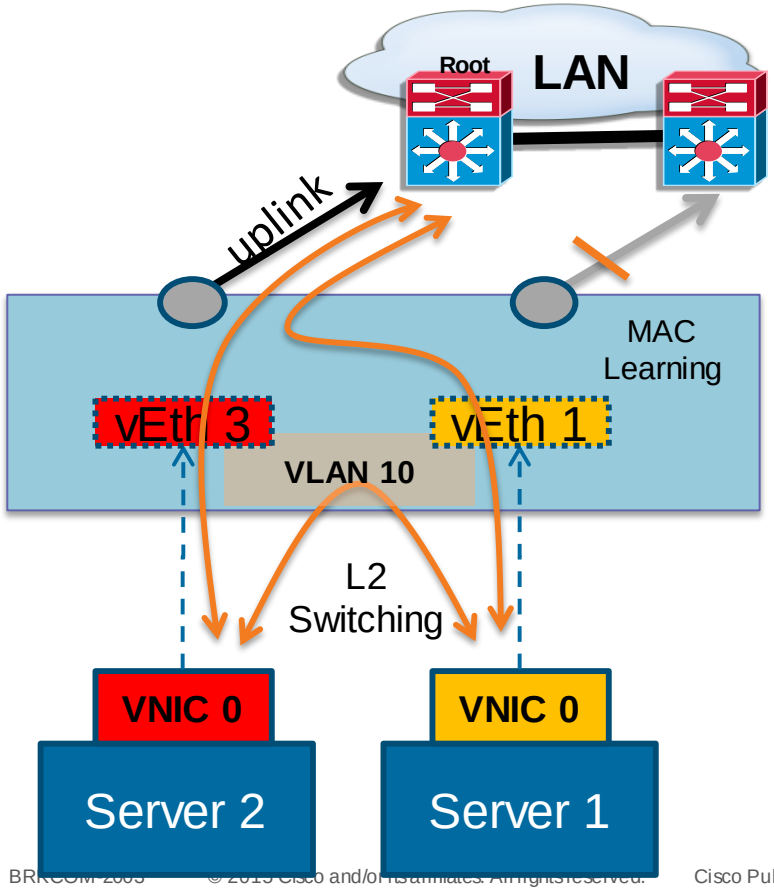
End Host Mode

Multicast Forwarding



- Broadcast traffic for a VLAN is pinned on exactly one uplink port (or port-channel) i.e., it is dropped when received on other uplinks
- Server to server multicast traffic is locally switched
- RPF and déjà vu check also applies for multicast traffic

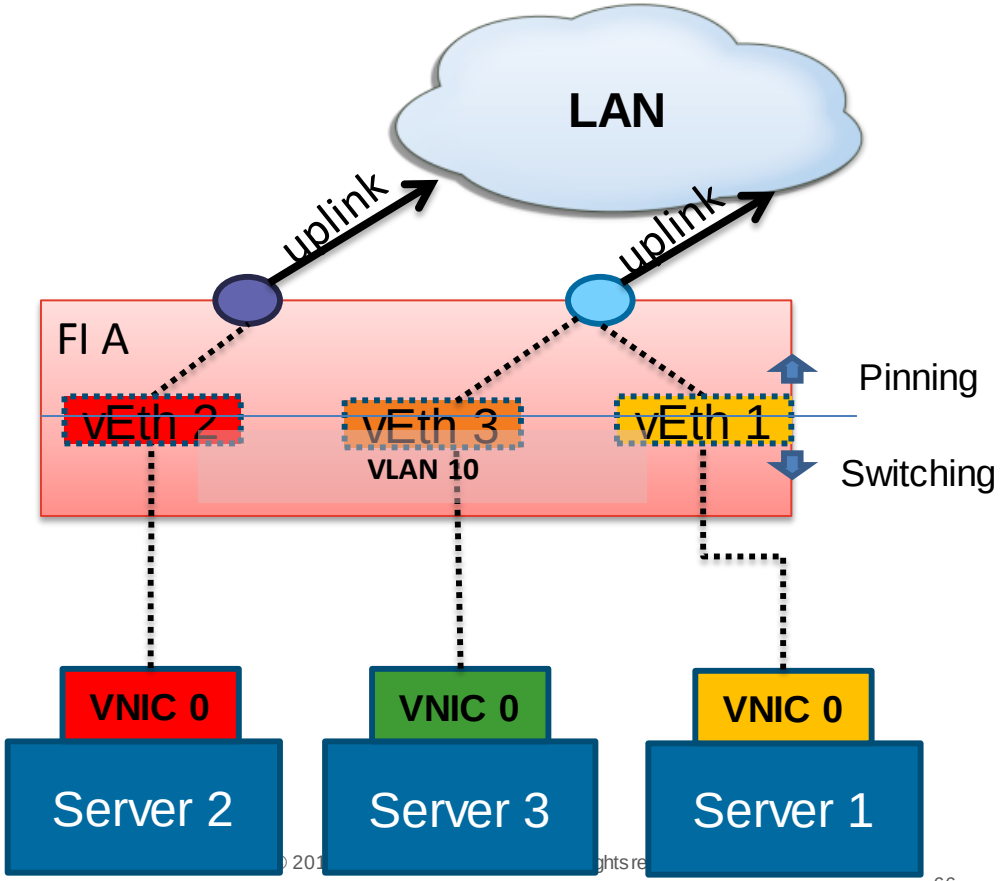
Switch Mode



- Fabric Interconnect behaves like a normal L2 switch
- Rapid-STP+ to prevent loops
  - STP parameters are not configurable
- Server vNIC traffic follows STP forwarding states
  - Use VPC to get around blocked ports
- VTP is not supported
- MAC address learning on both uplinks and server links


Uplink Pinning

End Host Mode - Dynamic Pinning

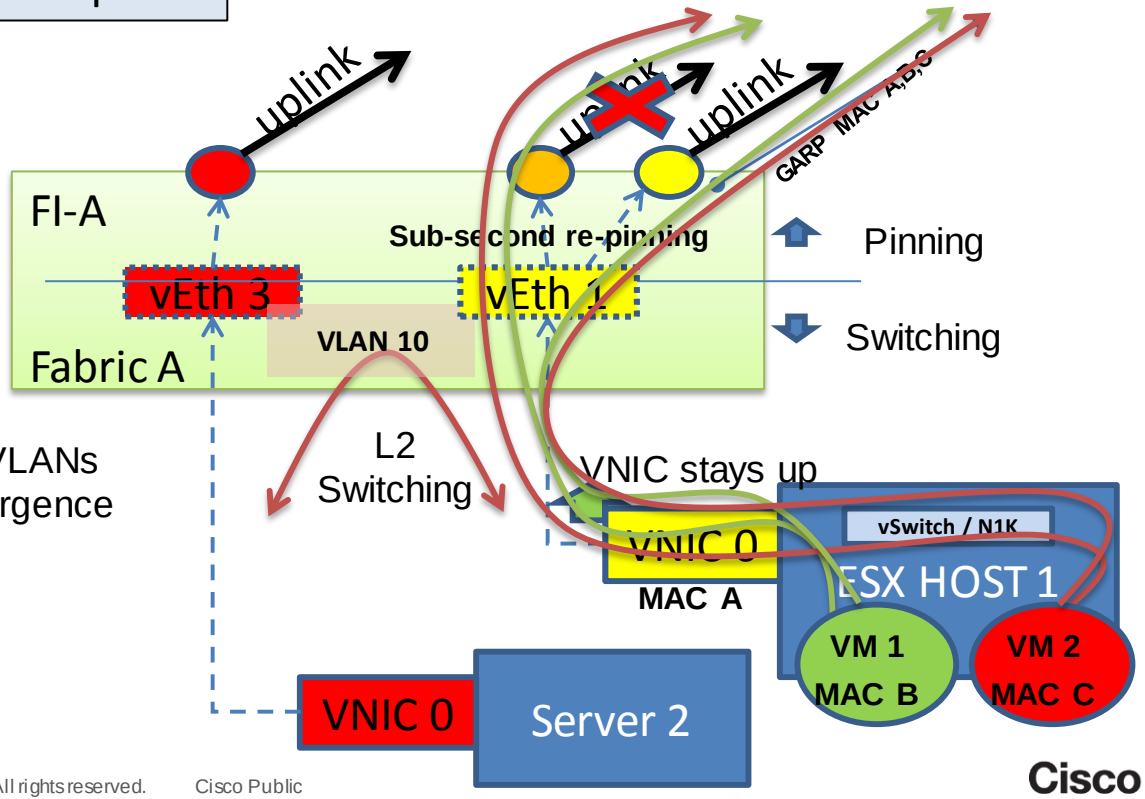


- UCSM manages the vEth pinning to the uplink
- UCSM will periodically vEth distribution and redistribute the vEths across the uplinks

Cisco live!

End Host Mode – Individual Uplinks

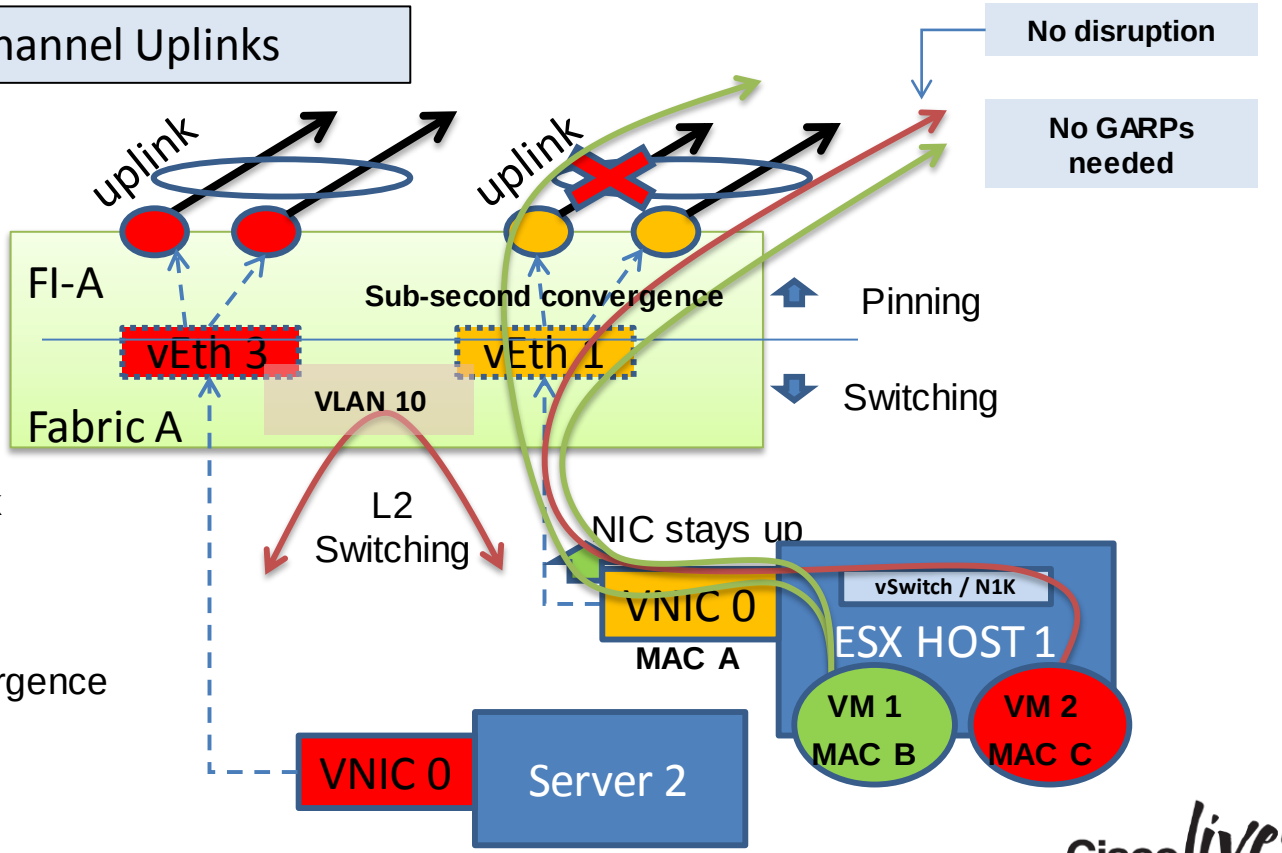
Dynamic Re-pinning of failed uplinks



- ✓All uplinks forwarding for all VLANs
- ✓GARP aided upstream convergence
- ✓No STP
- ✓Sub-second re-pinning
- ✓No server NIC disruption

End Host Mode – Port Channel Uplinks

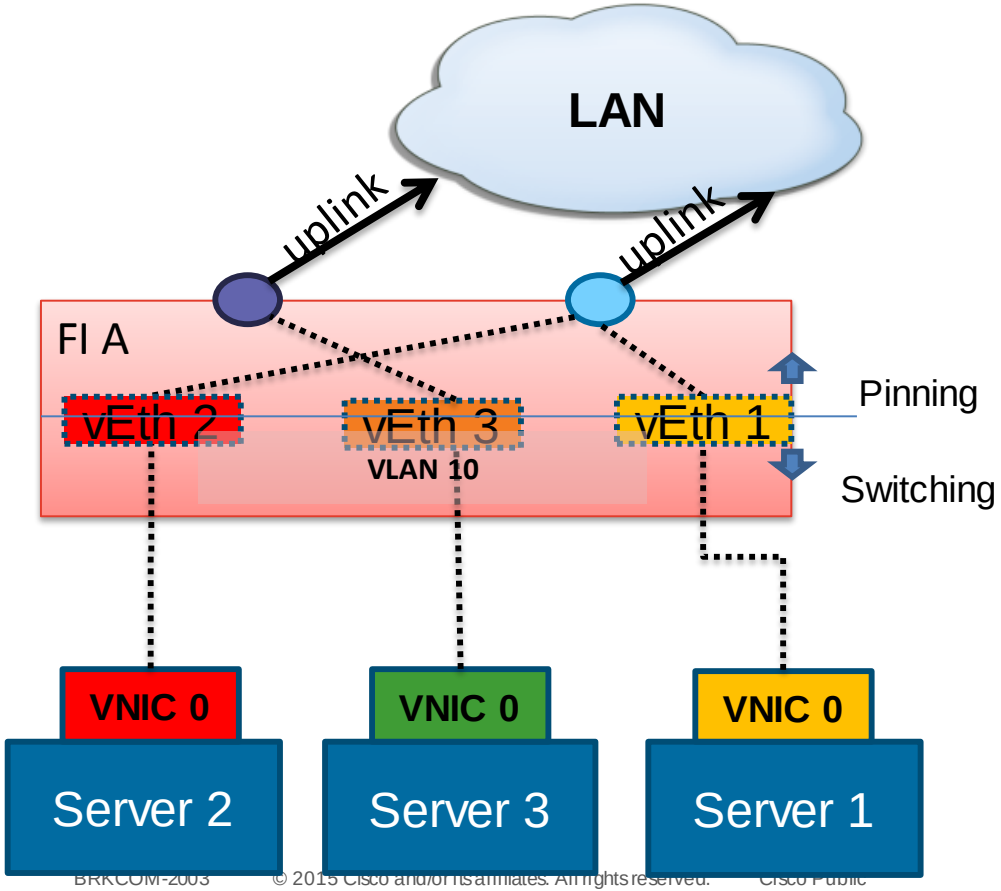
Recommended: Port Channel Uplinks



- ✓More Bandwidth per Uplink
- ✓Per flow uplink diversity
- ✓No Server NIC disruption
- ✓Fewer GARPs needed
- ✓Faster bi-directional convergence
- ✓**Fewer moving parts**

RECOMMENDED

End Host Mode – Static Pinning



Administrator Pinning Definition

vEth Interfaces	Uplink
vEth 1	Blue
vEth 2	Blue
vEth 3	Purple

- Administer controls the vEth pinning
- Deterministic traffic flow
- Pinning configuration is done under the LAN tab -> LAN Pin groups and assigned under the vNIC
- No re-pinning with in the same FI
- Static and dynamic pinning can co-exist

Which uplink is the servers' vEth pinned to?

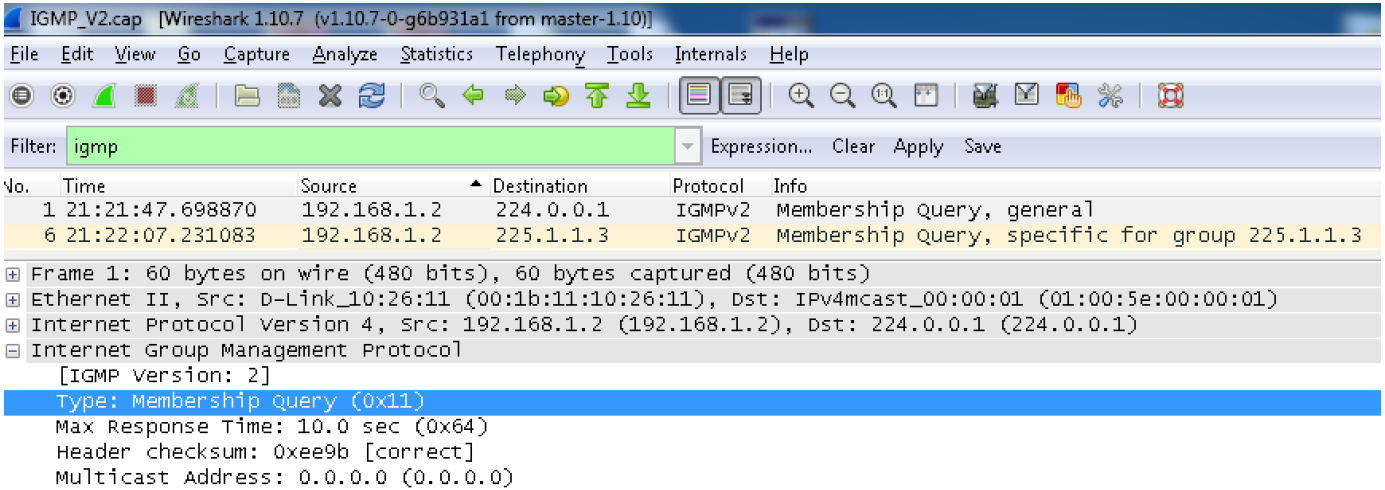
```
TME-UCS6100-A(nxos)# sh pinning border-interfaces

--------------------+---------+--------------------------------------
Border Interface    Status    SIFs
--------------------+---------+--------------------------------------
Po1                 Active    Veth1093 Veth1094 Veth1099
Po2                 Active    sup-eth2 Veth1103
Eth1/6              Down
Eth1/7              Down
Eth1/8              Down

Total Interfaces : 5
```

Fabric Forwarding - Multicast

Cisco live!

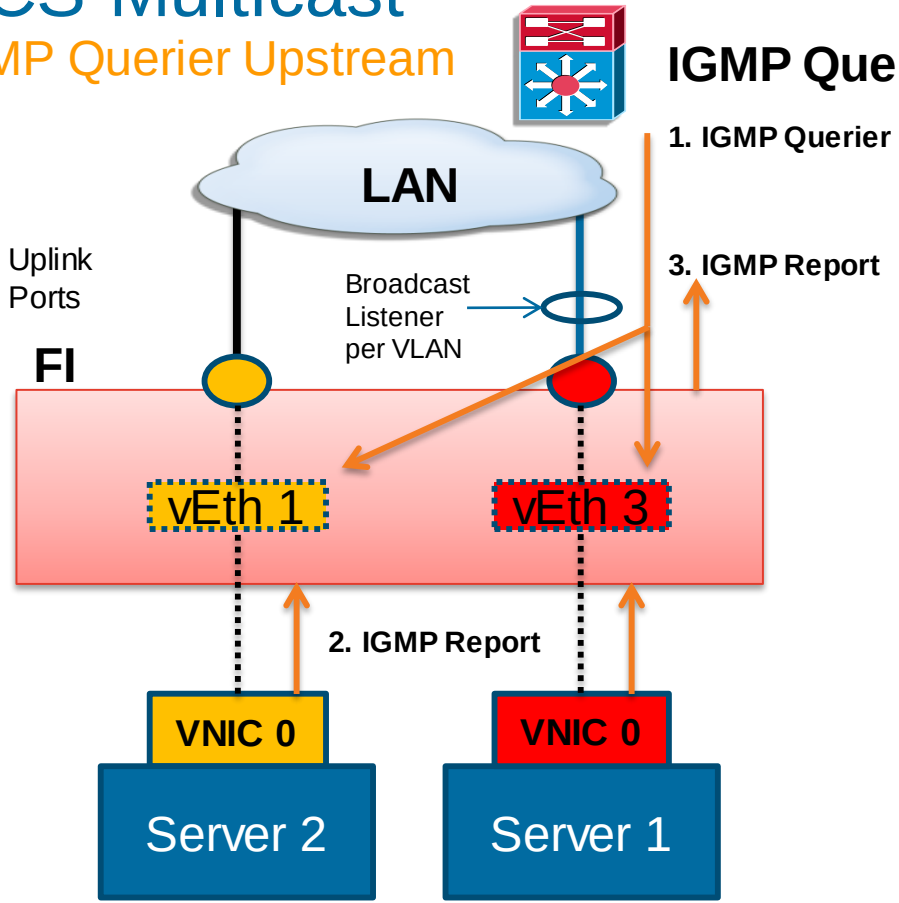
IGMP Querier?

- Three Options:
  1. Upstream IGMP Querier / PIM Router
  2. Fabric Interconnect IGMP Querier
  3. IGMP Snooping disabled



UCS Multicast

IGMP Querier Upstream



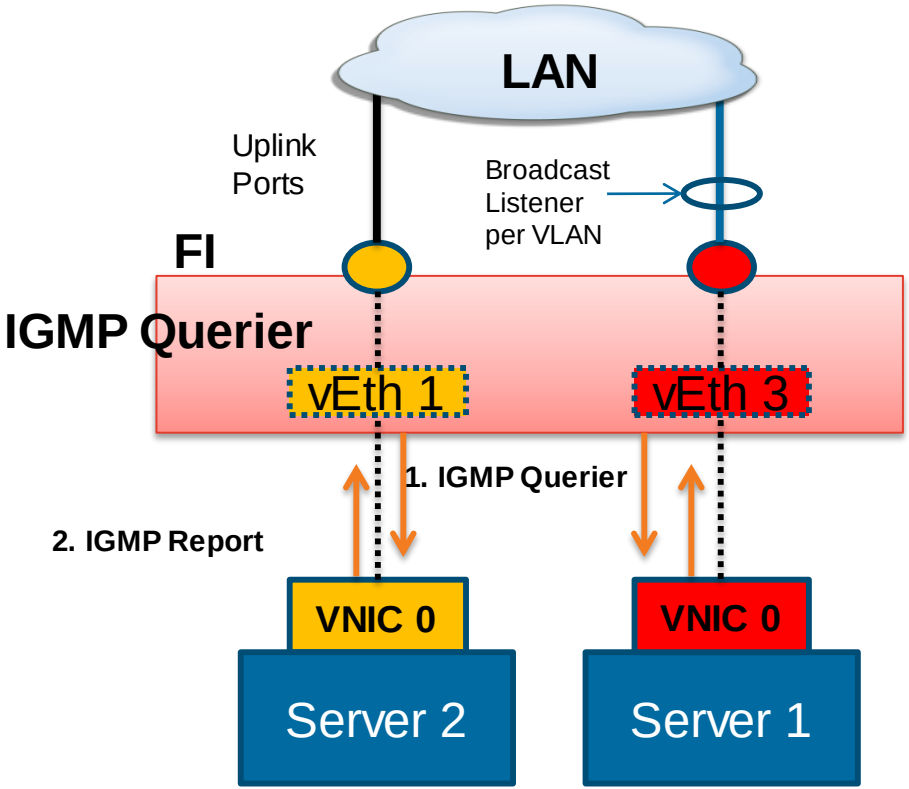

```
UCSB-2-A(nxos)# show ip igmp snooping vlan 511
IGMP Snooping information for vlan 511
  IGMP snooping enabled
  Optimised Multicast Flood (OMF) disabled
  IGMP querier present, address: 10.67.87.2, version: 3, i/f Eth1/2
  Querier interval: 125 secs
  Querier last member query interval: 1 secs
  Querier robustness: 2
  Switch-querier disabled
  IGMPv3 Explicit tracking enabled
  IGMPv2 Fast leave disabled
  IGMPv1/v2 Report suppression enabled
  IGMPv3 Report suppression disabled
  Link Local Groups suppression enabled
  Router port detection using PIM Hellos, IGMP Queries
  Number of router-ports: 1
  Number of groups: 0
  VLAN vPC function disabled
  Active ports:
    Eth1/1      Eth1/2  Veth32769       Veth32771
    Veth32772   Veth32773       Veth32774       Veth1315
```

G-pinned?

```
UCSB-2-A(nxos)# show ip igmp internal info global | grep "Vlan 511"
  Vlan 511: G-pinned if - curr: Eth1/2; prev: Eth1/1
  Vlan 511: Flood-pinned if - curr: Eth1/2; prev: Eth1/1
UCSB-2-A(nxos)#
```

UCS Multicast

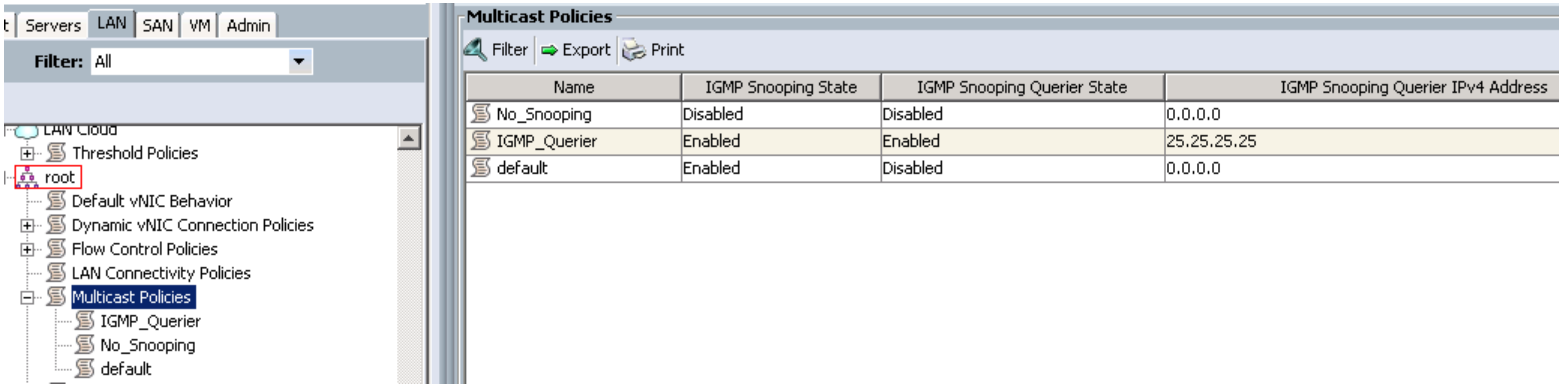
Internal Querier

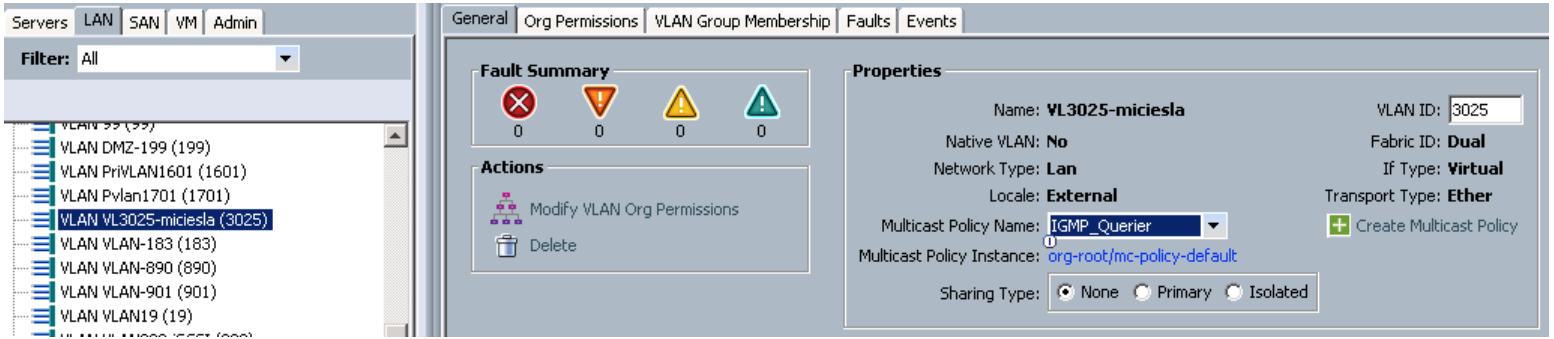



```
UCSB-2-A(nxos)# show ip igmp snooping vlan 511
IGMP Snooping information for vlan 511
  IGMP snooping enabled
  Optimised Multicast Flood (OMF) disabled
  IGMP querier present, address: 1.1.1.1, version: 3
  Querier interval: 125 secs
  Querier last member query interval: 1 secs
  Querier robustness: 2
  Switch-querier enabled, address 1.1.1.1, currently running
  IGMPv3 Explicit tracking enabled
  IGMPv2 Fast leave disabled
  IGMPv1/v2 Report suppression enabled
  IGMPv3 Report suppression disabled
  Link Local Groups suppression enabled
  Router port detection using PIM Hellos, IGMP Queries
  Number of router-ports: 1
  Number of groups: 0
  VLAN vPC function disabled
  Active ports:
    Eth1/1      Eth1/2  Veth32769       Veth32771
    Veth32772   Veth32773       Veth32774       Veth1315
```


Cisco live!

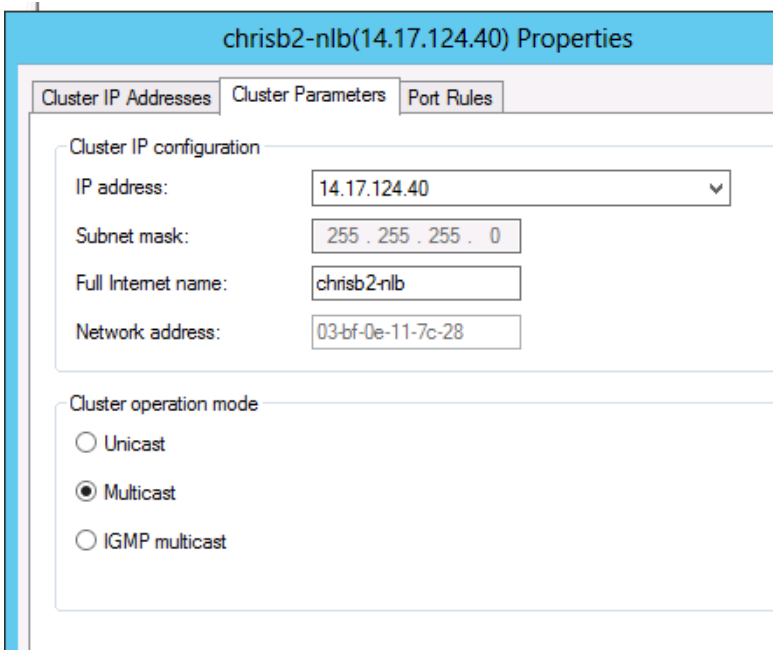
UCS Multicast Configuration (2.1+)





UCS & Microsoft Network Load Balancing (NLB)

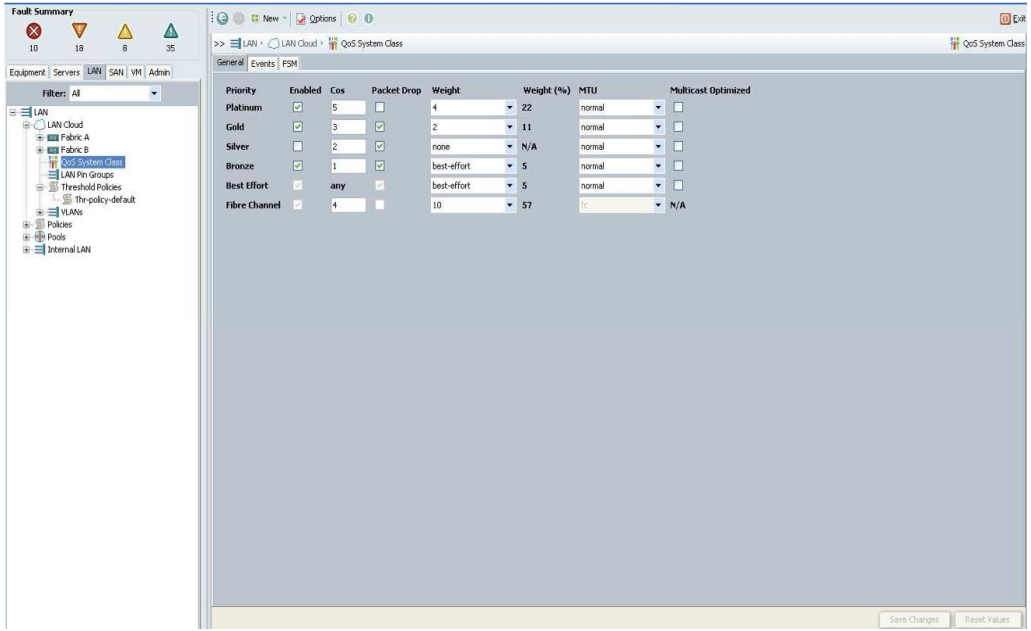
- Unicast Mode
  - Ethernet Switching mode only
  - Nexus 1000v (no mac auto-static-learn)
- Multicast Mode
- IGMP Multicast
  - Requires igmp querier



Fabric Forwarding - QoS

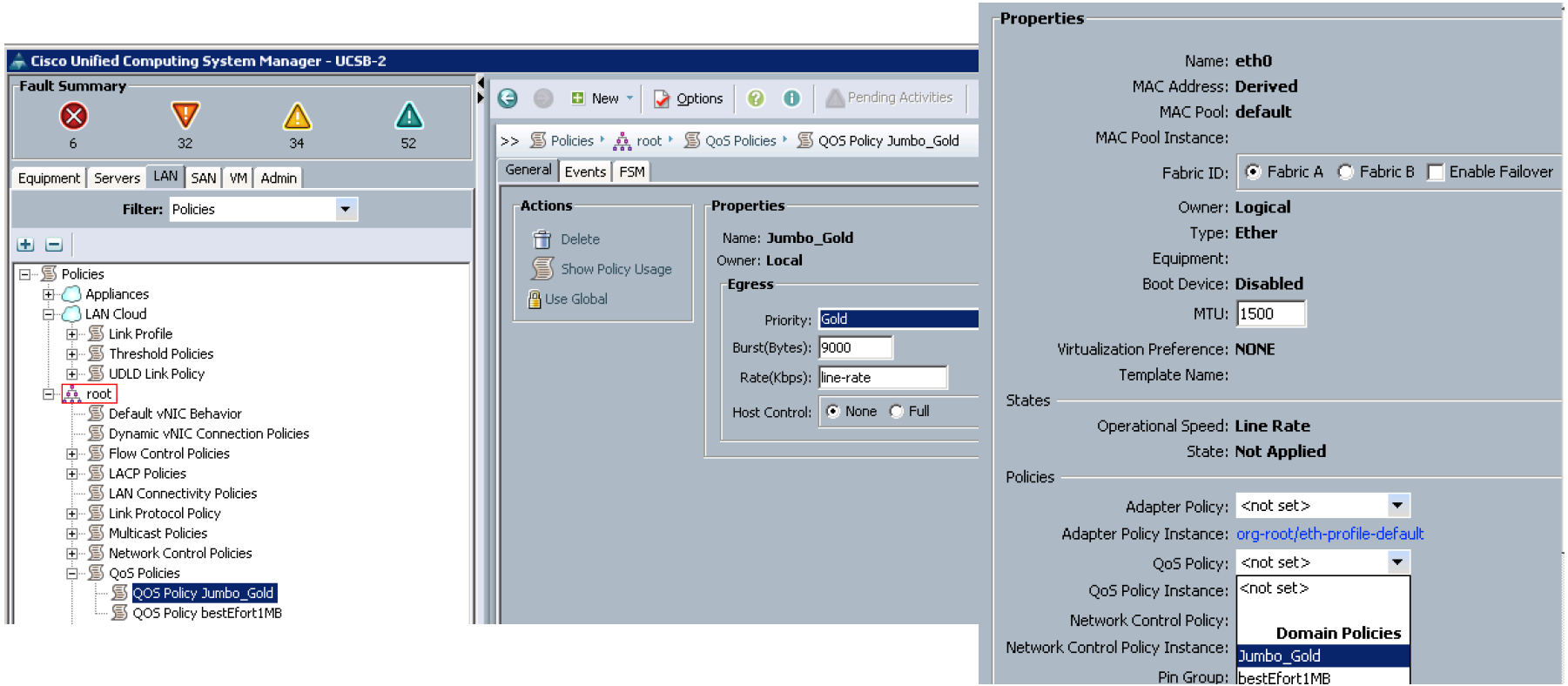
Cisco live!

UCS Congestion Management

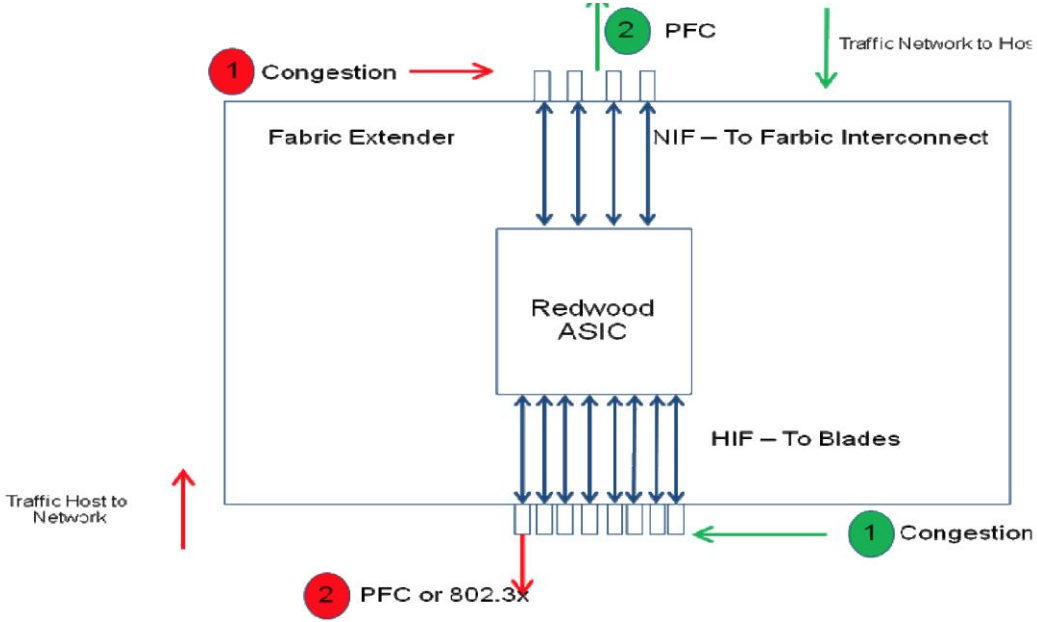


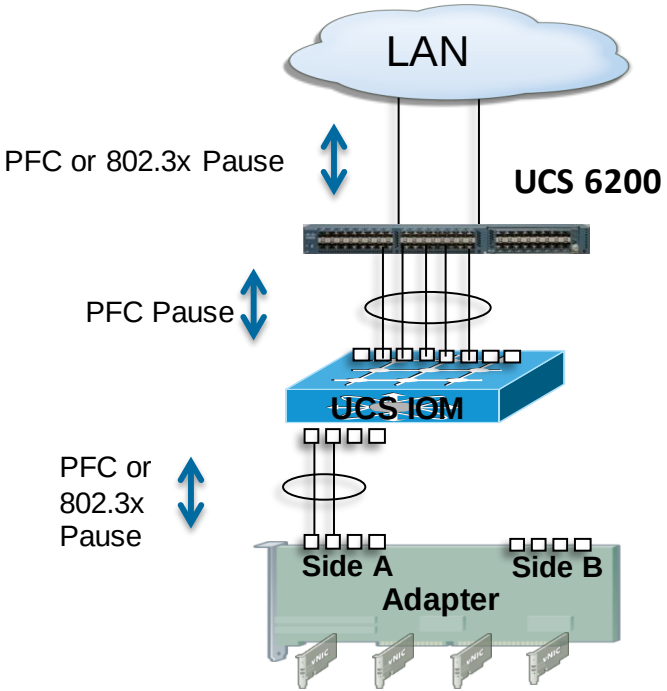
8 Classes: 1 FCoE, 1 best effort, 4 user-definable, 8 Classes: 2 reserved for control

UCS QoS – Marking / Classification



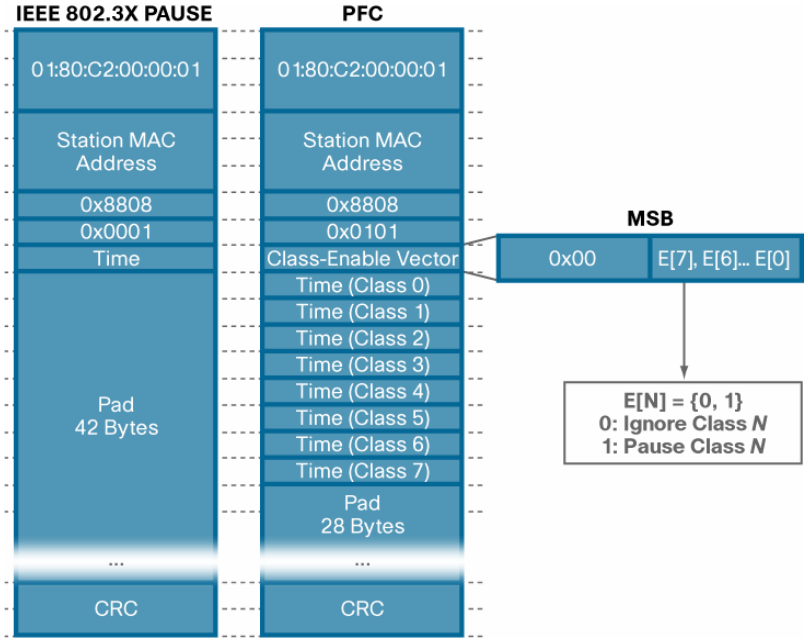
Pause!

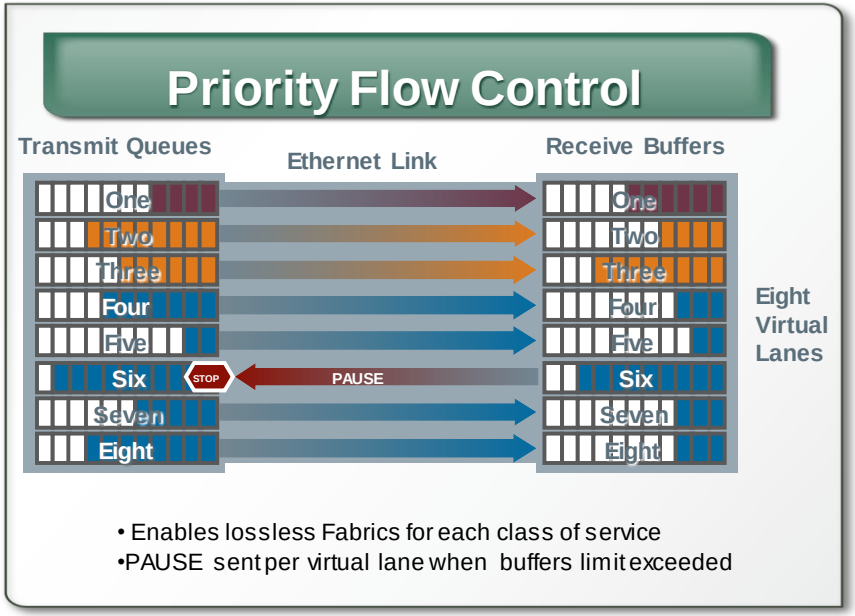




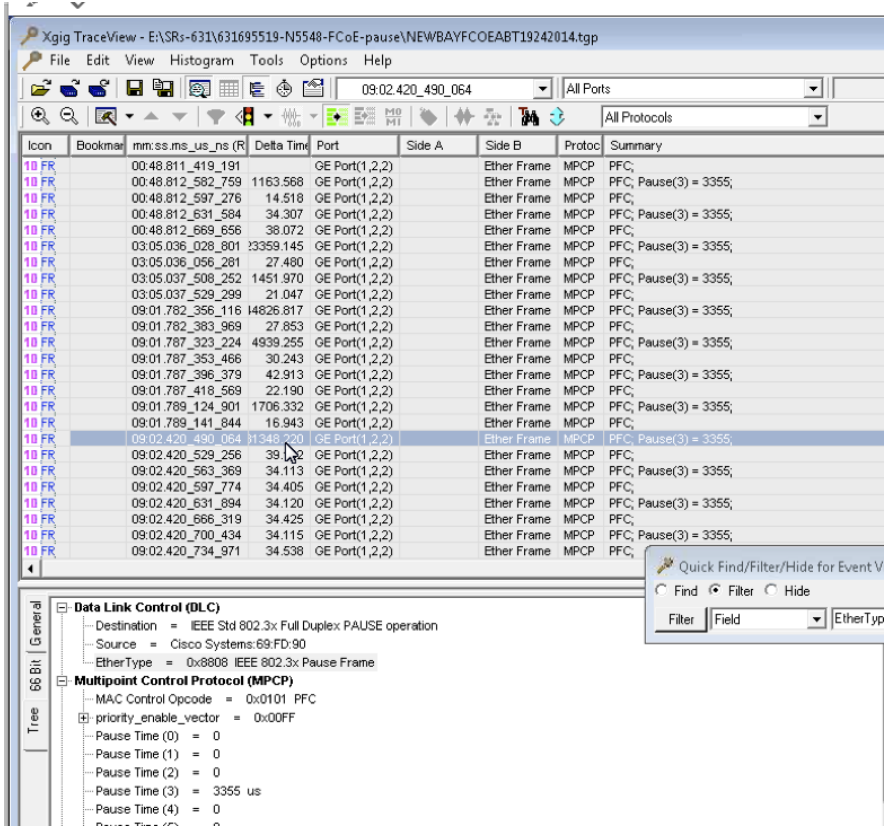
Priority Flow Control

Classical Ethernet Pause vs. Data Centre Ethernet PFC Pause





PFC Pause: What does it look like?



UCS QoS Identifying Congestion

```
UCSB-2-B(nxos)#  show  interface  priority-flow-control
=============================================================
Port                  Mode  Oper(VL  bmap)   RxPPP       TxPPP
=============================================================
Ethernet1/1           Auto  Off              0           0
Ethernet1/2           Auto  Off              0           0
Ethernet1/3           Auto  Off              39648       3597850
Ethernet1/4           Auto  Off              46784       4069738
Ethernet1/5           Auto  Off              0           0
```

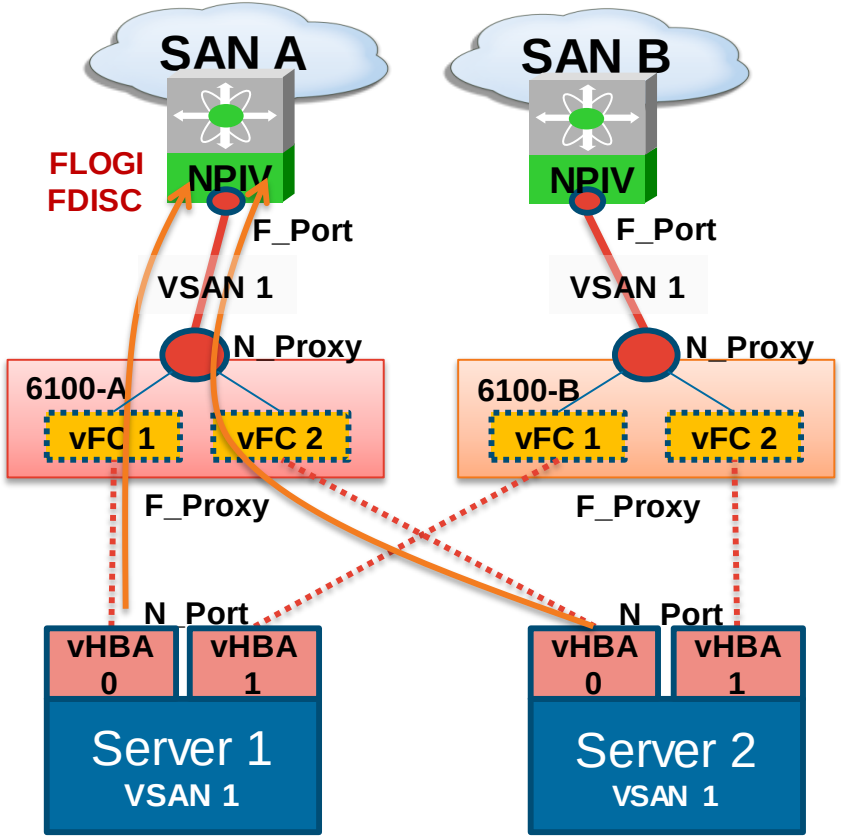
```
fex-1# show platform software woodside loss
+-------+-----------------------------------+-----------+-+--------------------------------+-------------------------------------+
|       |                                   |           | |                                |                                     |
|       |                                   |           | |                                | frm_to                              |
|       |                                   |Port Extra | |                                +-------------------------------------|
|       |          RMON                     |   Drop    |S|  SS Loss Counters              | COS                       | XOFF    |
|       +-----------+----------+------------+-----------|S|----------+----------+----------+-------------------------------------|
| Port  | Tx Pause  | Rx Pause | Errors     | Counters  |x| RX SS    | Tx SS    | SS Total |0  |1  |2  |3  |4  |5  |6  |7  |0  |1  |
+-------+-----------+----------+------------+-----------+-+----------+----------+----------+---+---+---+---+---+---+---+---+---+---+
| 0- NI1|          0|        32|           0|          0|0|         0|         0|         0|  0|  0|  0|  0|  0|  0|  0|  0|  0|  0|
+-------+-----------+----------+------------+-----------+-+----------+----------+----------+---+---+---+---+---+---+---+---+---+---+
| 0-HI19|         26|         0|           0|          0|0|         0|         0|         0|  0|  0|  0|  0|  0|  0|  0|  0|  0|  0|
+-------+-----------+----------+------------+-----------+-+----------+----------+----------+---+---+---+---+---+---+---+---+---+---+
fex-1#
```

Fabric Forwarding - Storage

Cisco live!

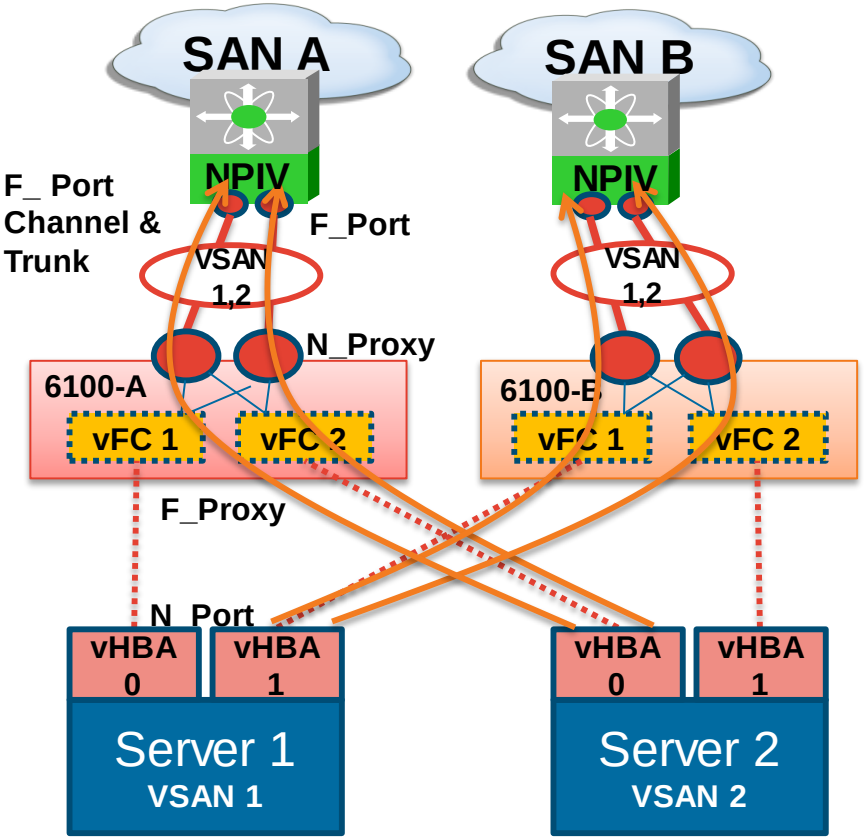
SAN “End Host” NPV Mode

N-Port Virtualisation Forwarding



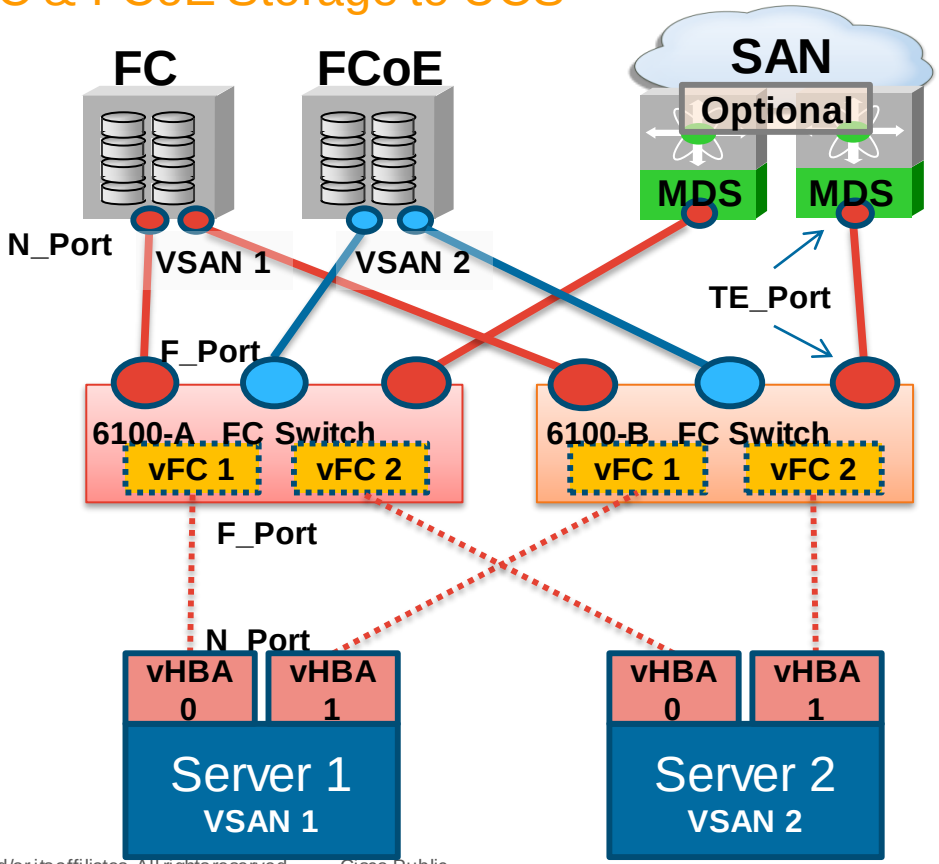
SAN “End Host” NPV Mode

N-Port Virtualisation Forwarding with MDS, Nexus 5000



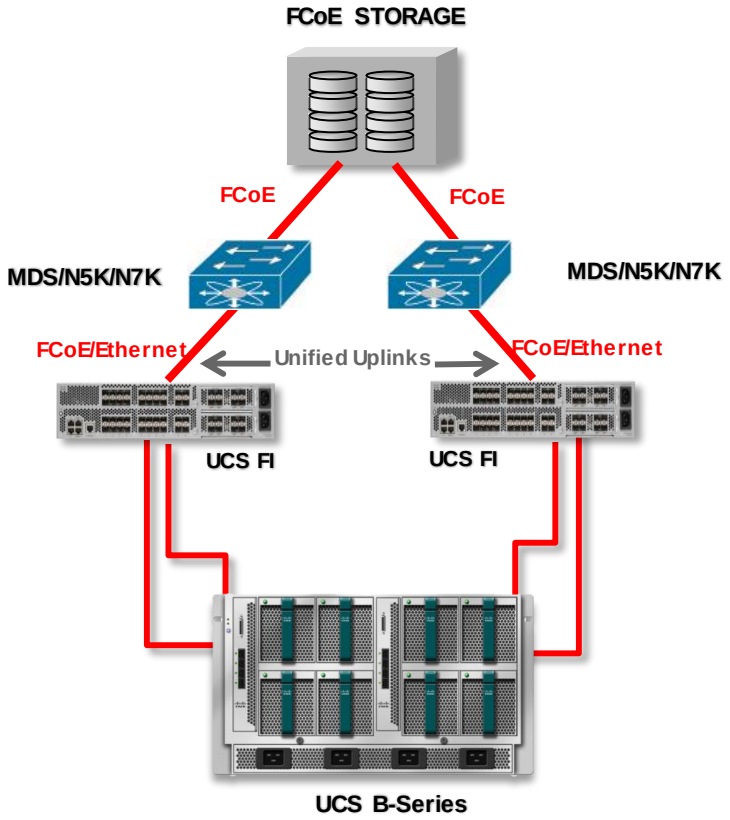
SAN FC Switch Mode

Direct Attach FC & FCoE Storage to UCS



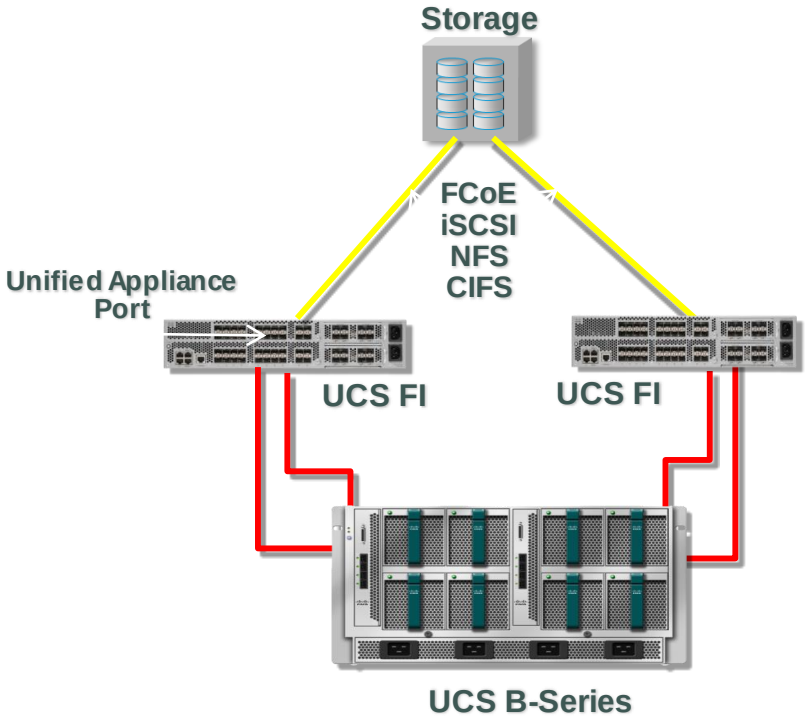
Multi-Hop FCoE

- Supports MDS, N5K & N7K
- “Unified Uplink” port type
- FI in ENM Switching Mode
  - VNP port type
- FI in FC Switching Mode
  - VE port type



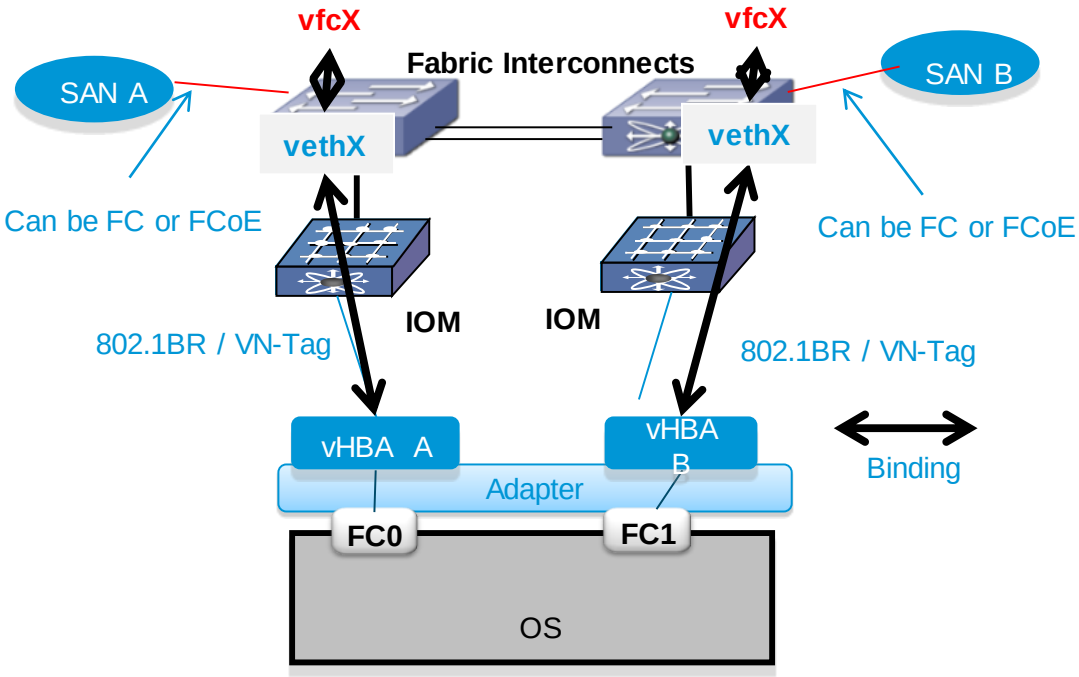
Unified Appliance Ports

Direct attach FCoE, iSCSI, NFS & CIFS storage



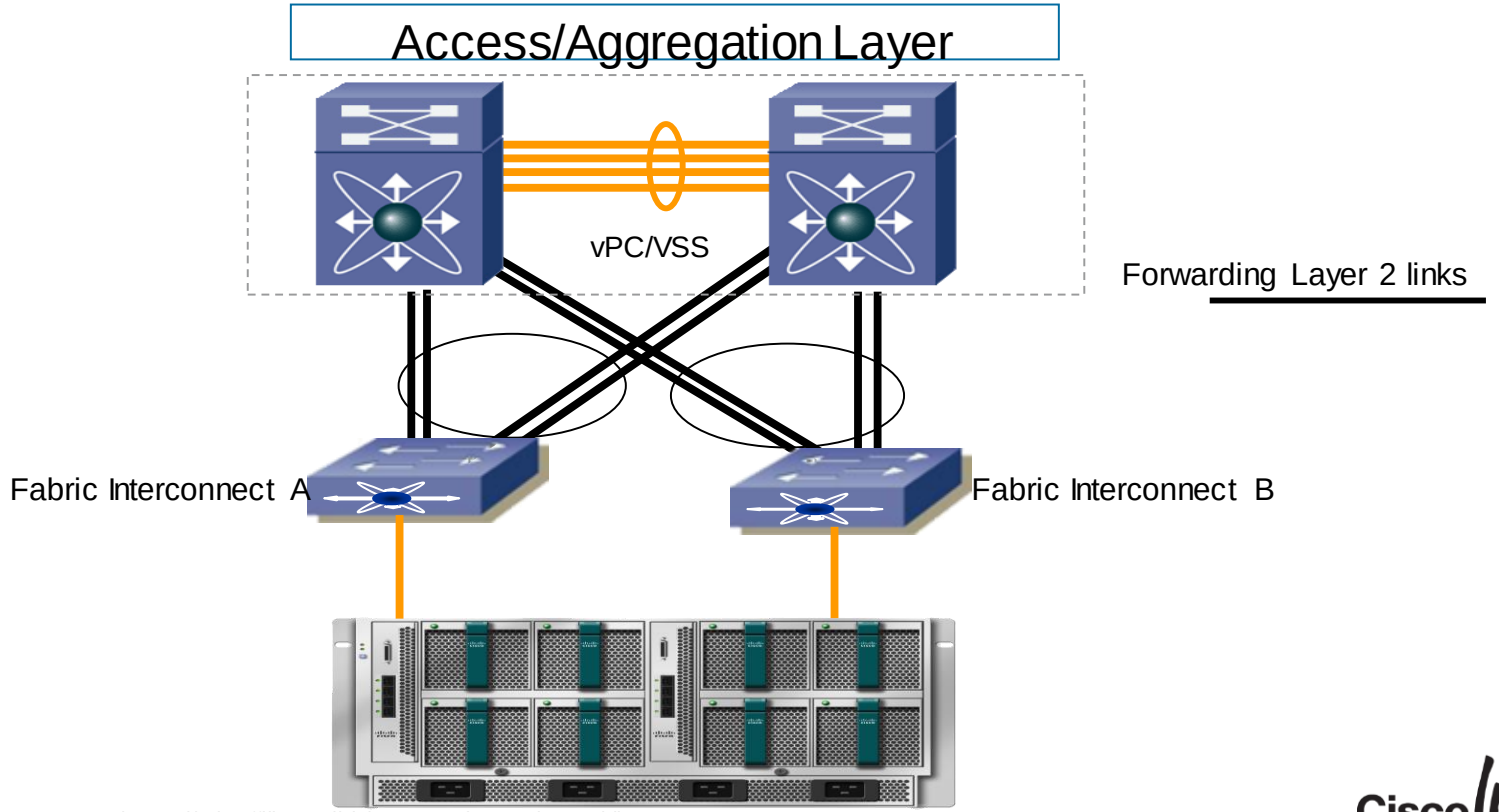
FCoE with Adapter FEX

VFC interface bound to 802.1BR / VN-Tag virtual ethernet interface

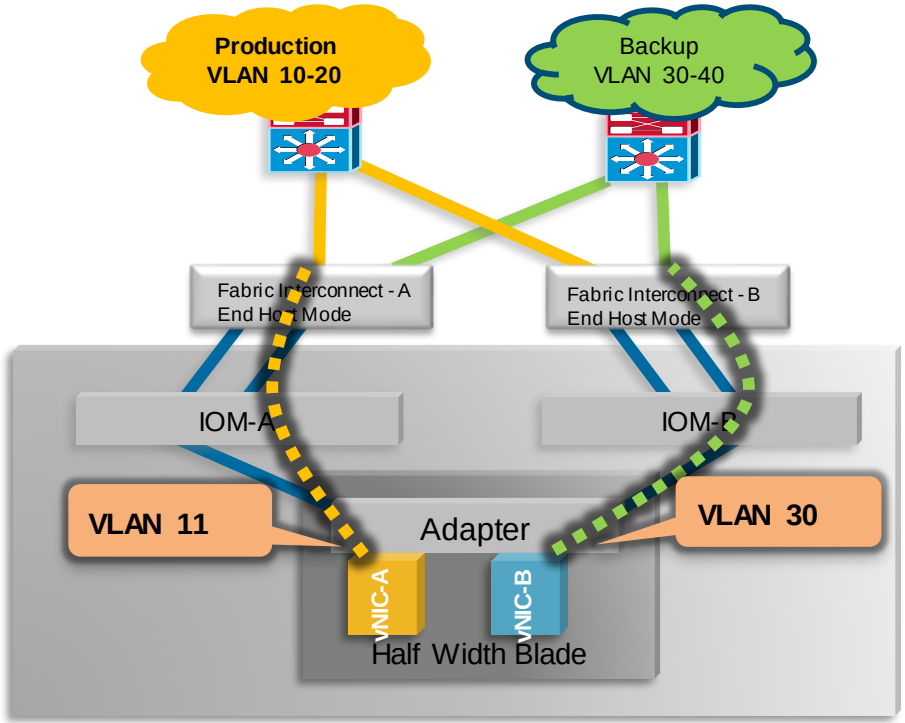



Topologies

Recommended Topology for Upstream Connectivity



Layer 2 Disjoint Topology



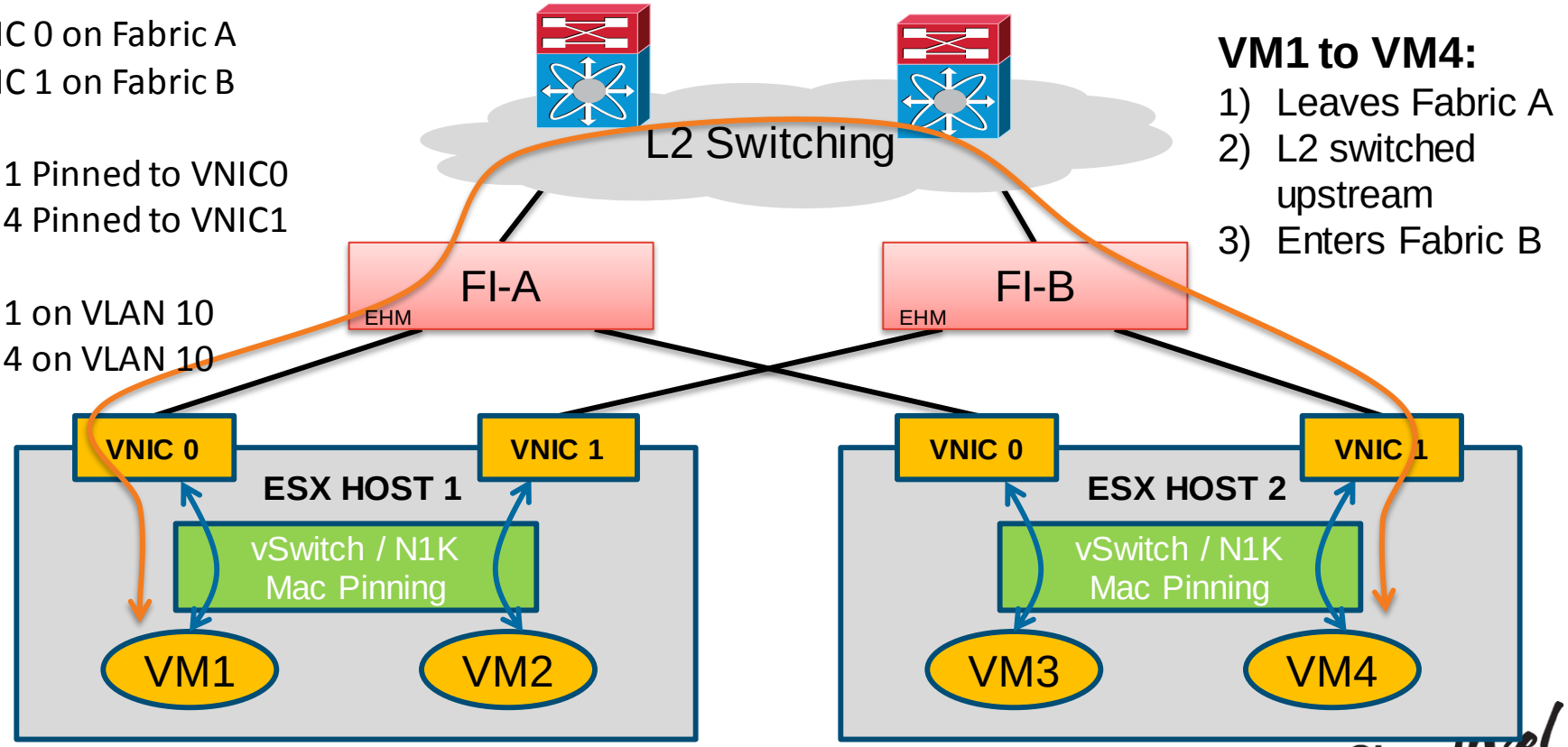
- A vNIC can only participate in one L2 network upstream
- Both dynamic and static pinning methods are supported

Inter-Fabric Traffic Example

- VNIC 0 on Fabric A
- VNIC 1 on Fabric B

- VM1 Pinned to VNIC0
- VM4 Pinned to VNIC1

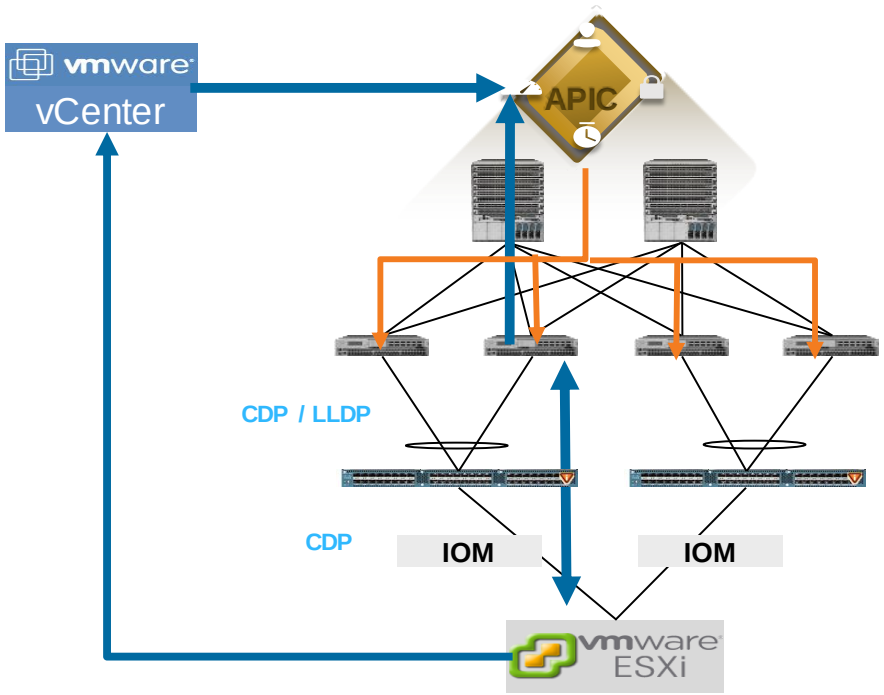
- VM1 on VLAN 10
- VM4 on VLAN 10



VM1 to VM4:

1) Leaves Fabric A
2) L2 switched upstream
3) Enters Fabric B

UCS & ACI Fabric



1. FI sends CDP/LLDP to leaf, CDP to ESXi blade.
2. ESXi blade & leaf send discovery data to vCenter & APIC
3. APIC receives discovery data from vCenter
4. APIC downloads policy on all leafs providing path to the ESXi blade

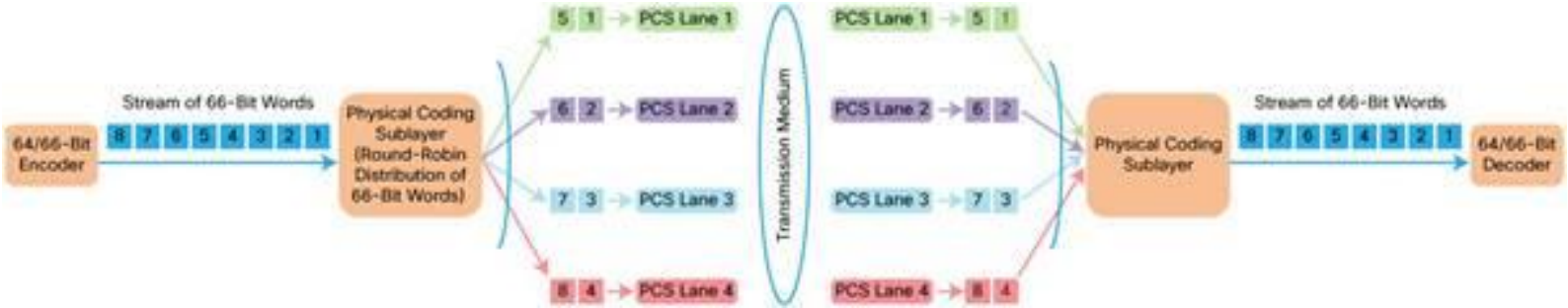
VIC 1300 New Features

Cisco live!

Native 40 Gigabit Ethernet

10 Gb Encoding

40 Gb Encoding = 4 x 10 Gb

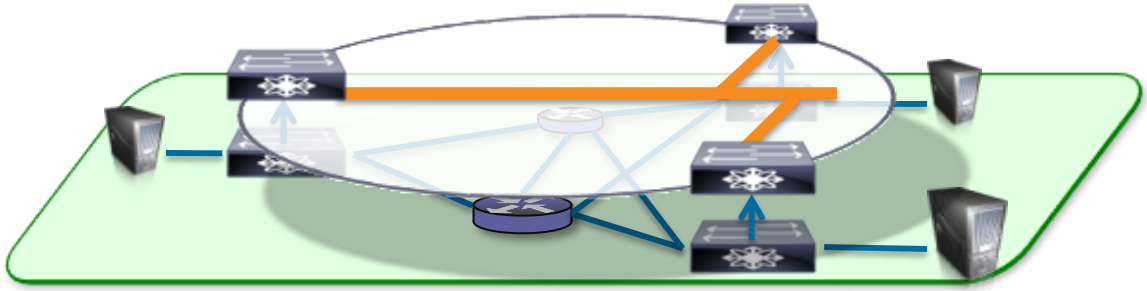


VIC 1300 VXLAN & NVGRE Offload

Hardware based TCP segmentation & checksum verification

- Overlay encap/decap performed by hypervisors:
  - Increases CPU utilisation
  - Decrease network throughput

- Inner & outer packet hardware processing:
  - TCP segmentation
  - TCP/UDP checksum
  - IP checksum

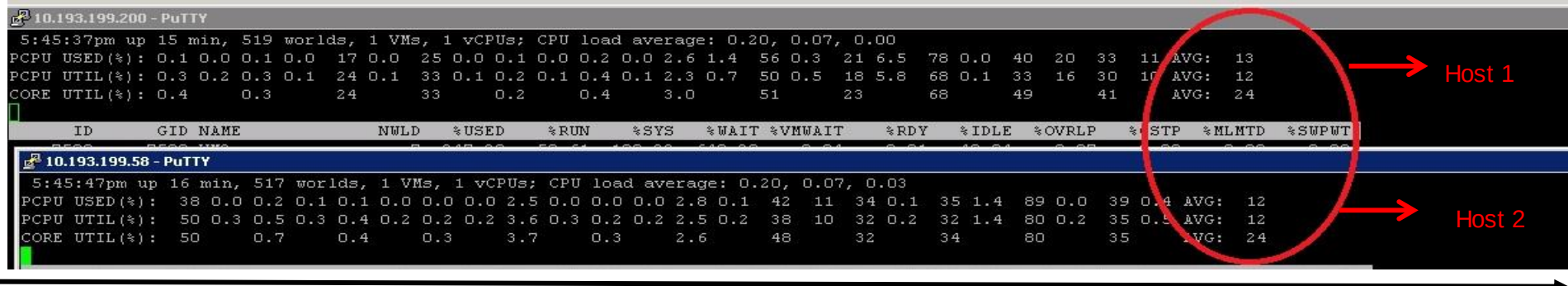


VXLAN Performance Test

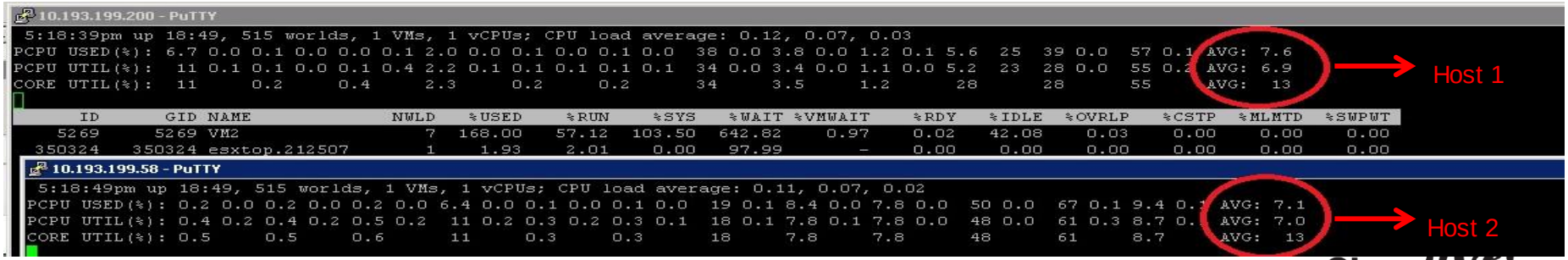
- 2 x ESXi 5.5 GA, VMware DVS
- UCS B200 M3, VIC 1380
- FI 6248, IOM 2208
- VM guest OS: RHEL6.4 64 bit, 1GB RAM, 1vCPU
- Unicast iPerf test between two VMs
  - iperf –s, iperf –c 17.1.1.1, TCP window size 23.2 KB
- Locally switch path on FI

VXLAN Offload Performance - CPU

VXLAN stateless offloads disabled



VXLAN stateless offloads enabled



VXLAN Offload Performance - Throughput

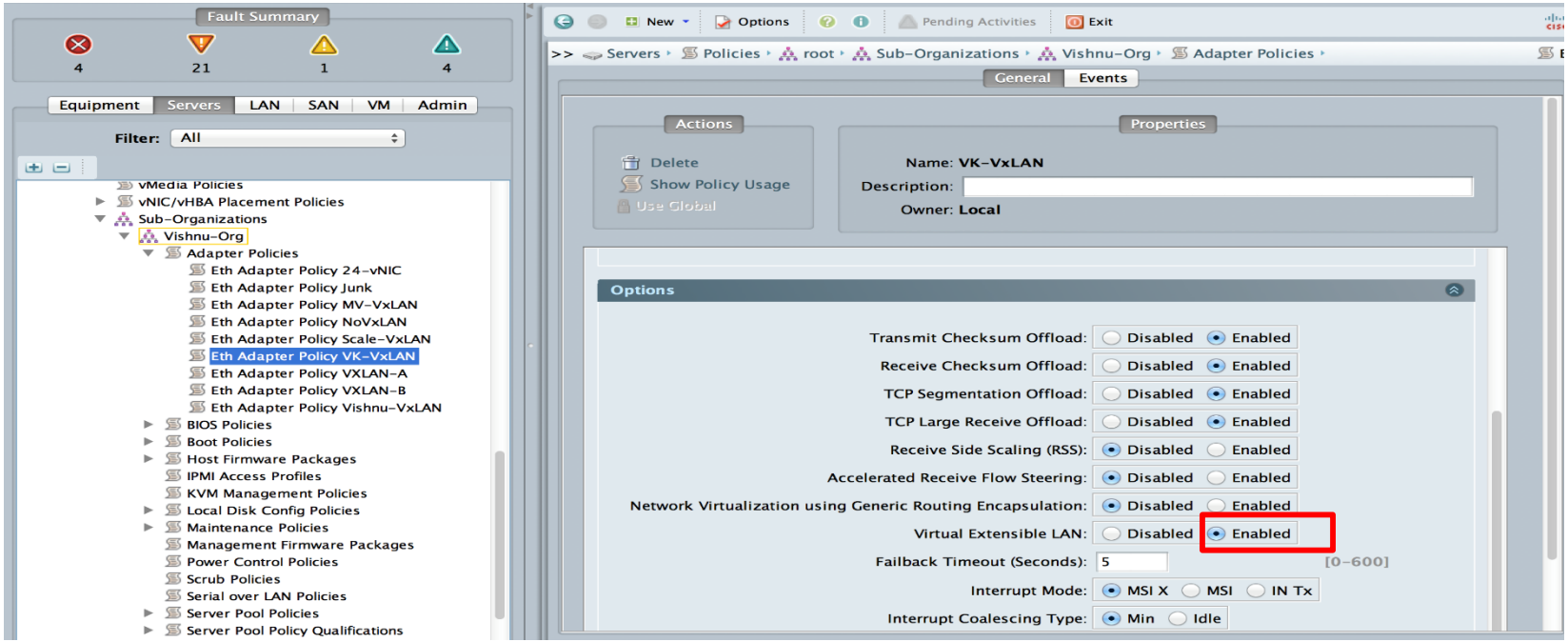
VXLAN stateless offloads disabled

```
[admin@netflow-traffic Desktop]$ iperf -c 17.1.1.1 -i 5
------------------------------------------------------------
Client connecting to 17.1.1.1, TCP port 5001
TCP window size: 23.2 KByte (default)
------------------------------------------------------------
[  3] local 17.1.1.2 port 56913 connected with 17.1.1.1 port 5001
[ ID] Interval       Transfer     Bandwidth
[  3]  0.0- 5.0 sec  2.40 GBytes  4.12 Gbits/sec
[  3]  5.0-10.0 sec  2.42 GBytes  4.16 Gbits/sec
[  3]  0.0-10.0 sec  4.82 GBytes  4.14 Gbits/sec
```

VXLAN stateless offloads enabled

```
[admin@netflow-traffic Desktop]$ iperf -c 17.1.1.1 -i 5
------------------------------------------------------------
Client connecting to 17.1.1.1, TCP port 5001
TCP window size: 23.2 KByte (default)
------------------------------------------------------------
[  3] local 17.1.1.2 port 56652 connected with 17.1.1.1 port 5001
[ ID] Interval       Transfer     Bandwidth
[  3]  0.0- 5.0 sec  3.54 GBytes  6.09 Gbits/sec
[  3]  5.0-10.0 sec  3.59 GBytes  6.17 Gbits/sec
[  3]  0.0-10.0 sec  7.14 GBytes  6.13 Gbits/sec
```

VXLAN Offload Configuration



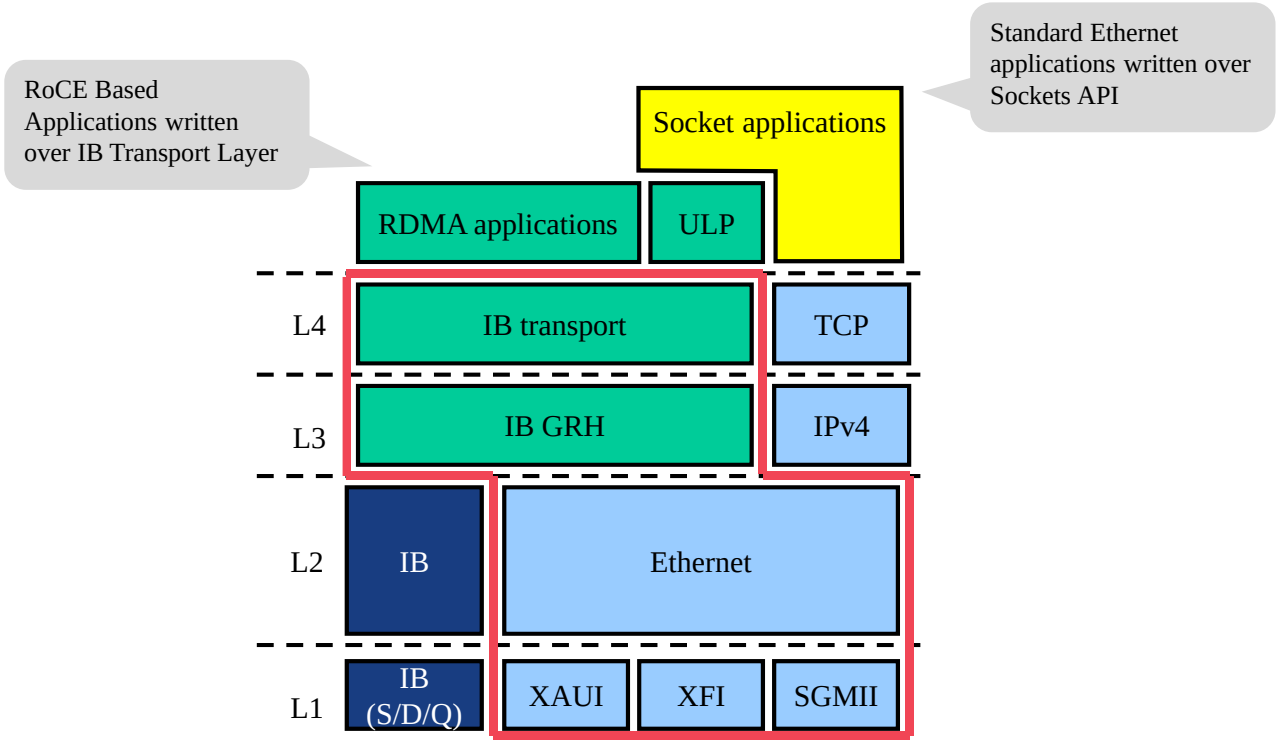
VIC 1300 RoCE Support

Remote Direct Memory Access over Converged Ethernet (RoCE)

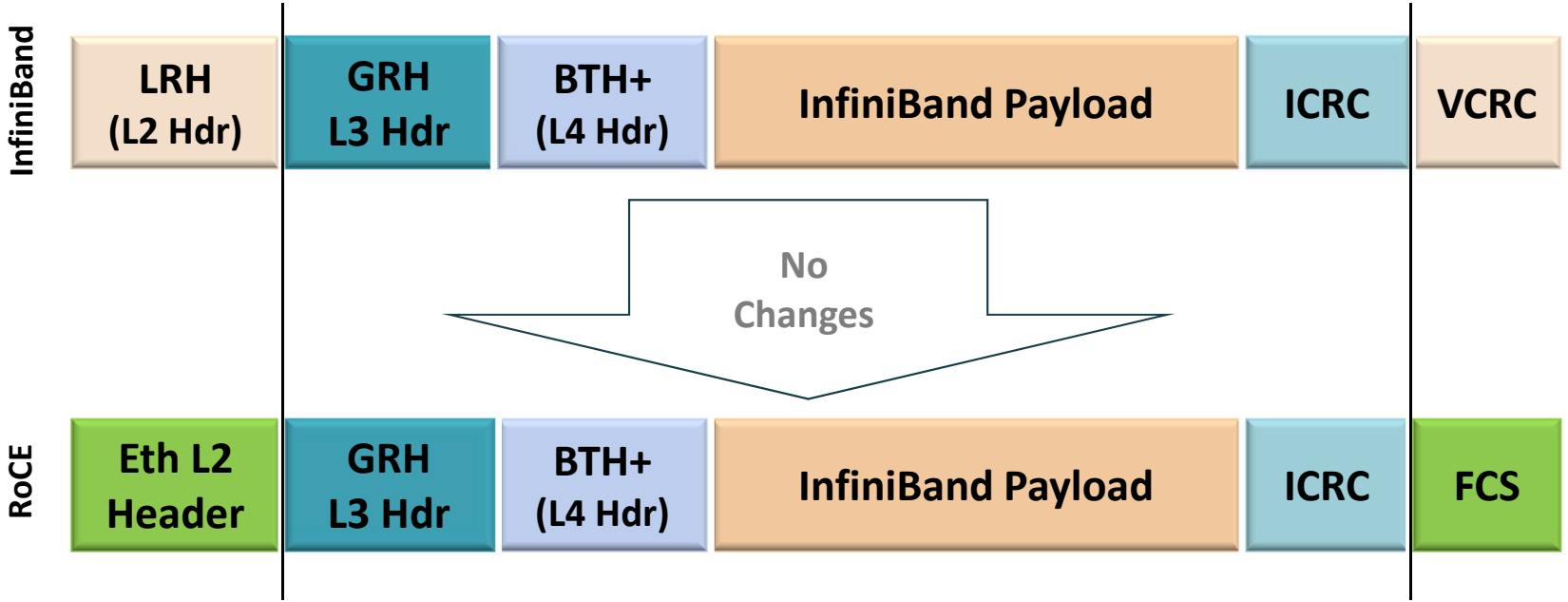
- Access remote node's memory w/o CPU interruption
- Lower latency, better CPU use
- "RoCE does for InfiniBand what FCoE did for Fibre Channel"*

* Scott Lowe - http://blog.scottlowe.org/2010/04/20/am-i-understanding-roce-correctly

RoCE vs InfiniBand Protocol Stack

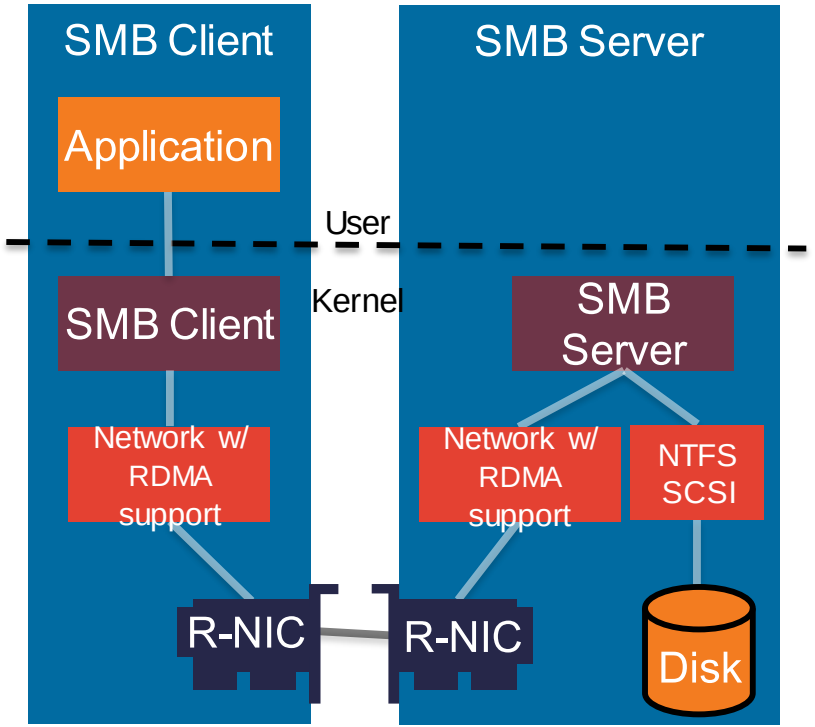


RoCE vs InfiniBand Packet Format



Microsoft SMB Direct (SMB over RMDA)

- Initial use case for RoCE is Microsoft Windows 2012 SMB Direct 3.0

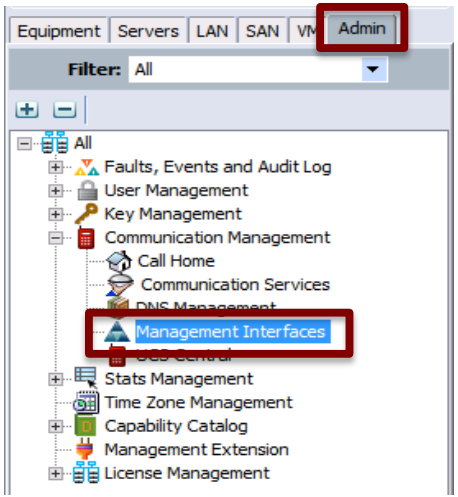


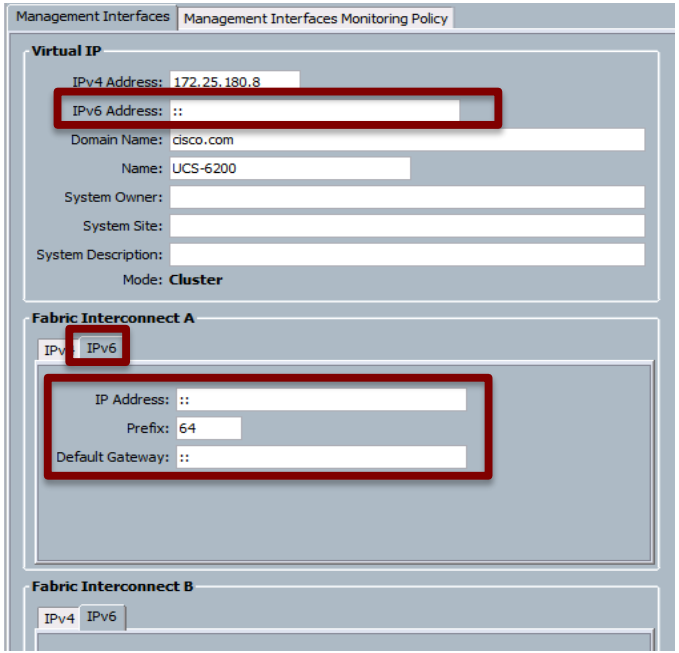
Jose Barreto, Tech Ed, 2013, http://channel9.msdn.com/Events/TechEd/NorthAmerica/2013/MDC-B335#fbid=

New Features in 2.2

IPv6 Management Support

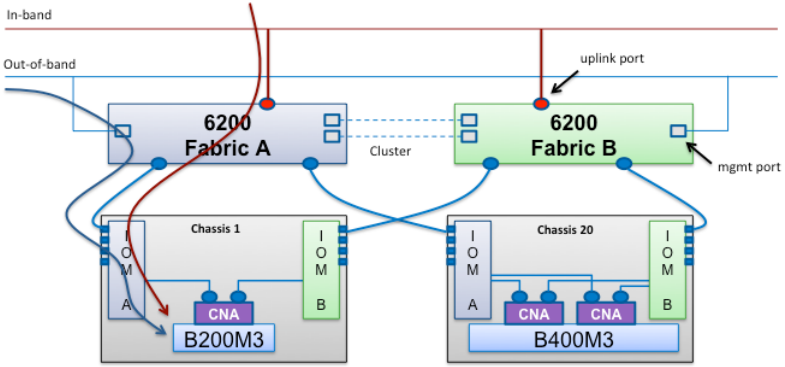
- UCS FI management can be configured with IPv6 address
  - IPv6 capable external services (i.e., NTP, SSH, TACACs, HTTP/HTTPs, etc)

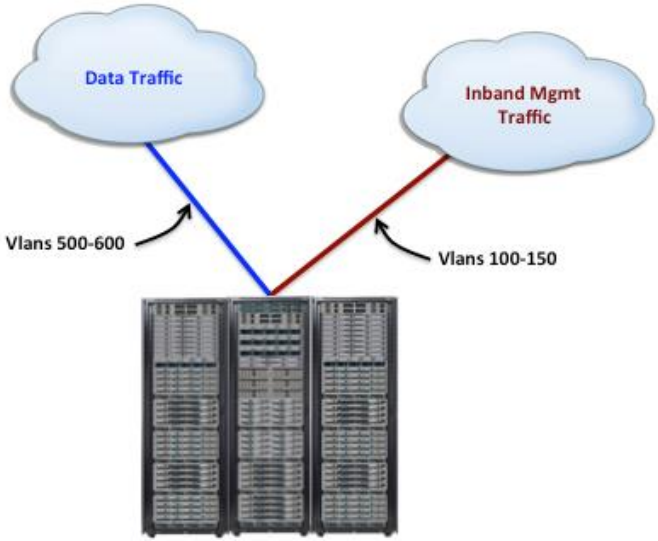




Inband Management for CIMC

- Separate server management (CIMC) traffic from UCSM
  - Designating servers' CIMC into different groups
- Higher bandwidth
  - 10G vs 1G



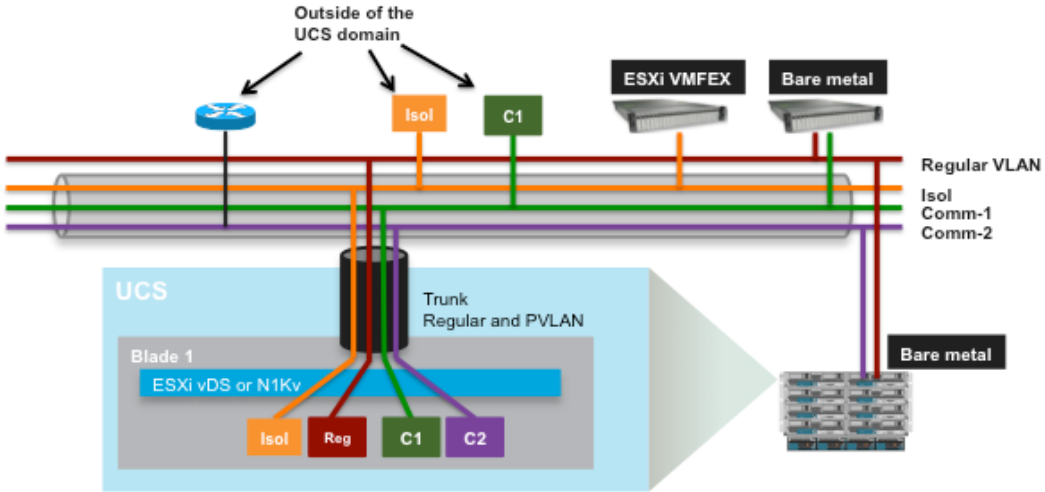


Faster and Better: Link Layer Enhancements

- Uplink ports only
- Faster link failure detection with UDLD
  - Layer 2 protocol that runs on top of the physical layer to help detect mis-wiring and uni-directional communication
  - UDLD Interval 7-90 seconds, with a detection time of 3x the interval
- LACP (pertain to uplinks only)
  - Choices between different timers, slow (30 sec) or fast (1 sec). Default is slow. Provides very fast failure detection
  - Enable suspend-individual link. Default is disable
    - Match settings of the LAN Switch

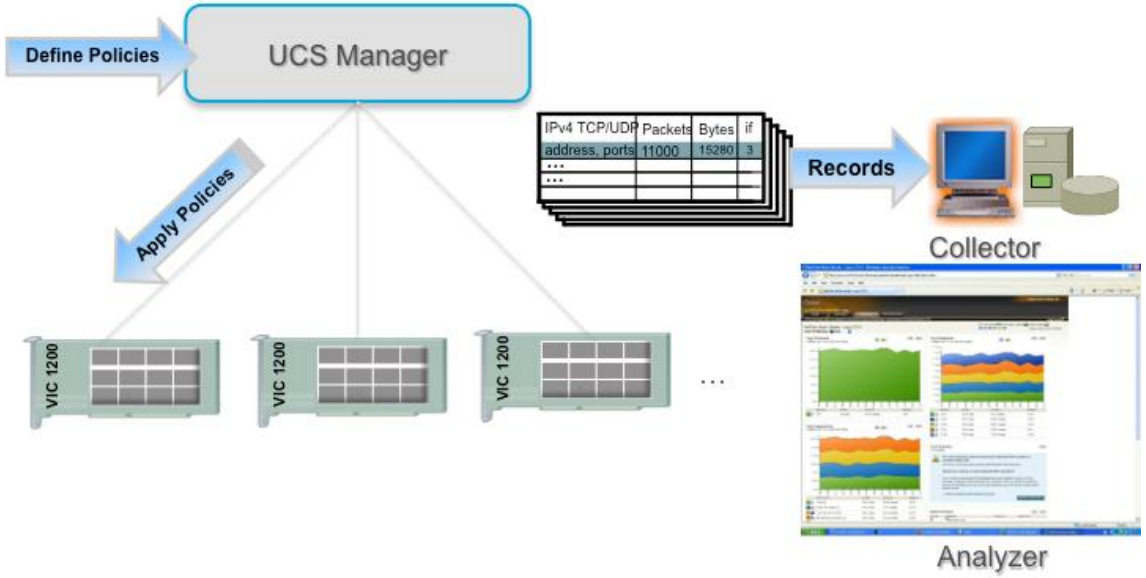
PVLAN Enhancements

- Promiscuous on Appliance Port
- Community Support
- PVLAN trunking on the vNIC (extend PVLAN to the virtual switches)



Netflow+

- Improve workload visibility
  - Capacity planning
  - Security
  - Troubleshooting



Other Sessions

- BRKCOM-1001 - UCS Fundamentals
- BRKCOM-2015 - Cisco UCS Network Performance Optimisation and Best Practices for VMware
- BRKCOM-2017 - UCS Systems Management Deep Dive with UCS Foundational Software
- BRKCOM-2601 - Hyper-Converged Computing
- BRKCOM-2602 - Next Generation Computing Architectures for Cloud Scale Applications
- BRKCOM-2640 - UCS C-Series Deployment Options, Best Practice and UCSM Integration
- BRKCOM-3002 - UCS Performance Troubleshooting
- BRKVIR-2044 - Multi-Hypervisor Networking - Compare and Contrast


Q & A
Cisco live!

Complete Your Online Session Evaluation

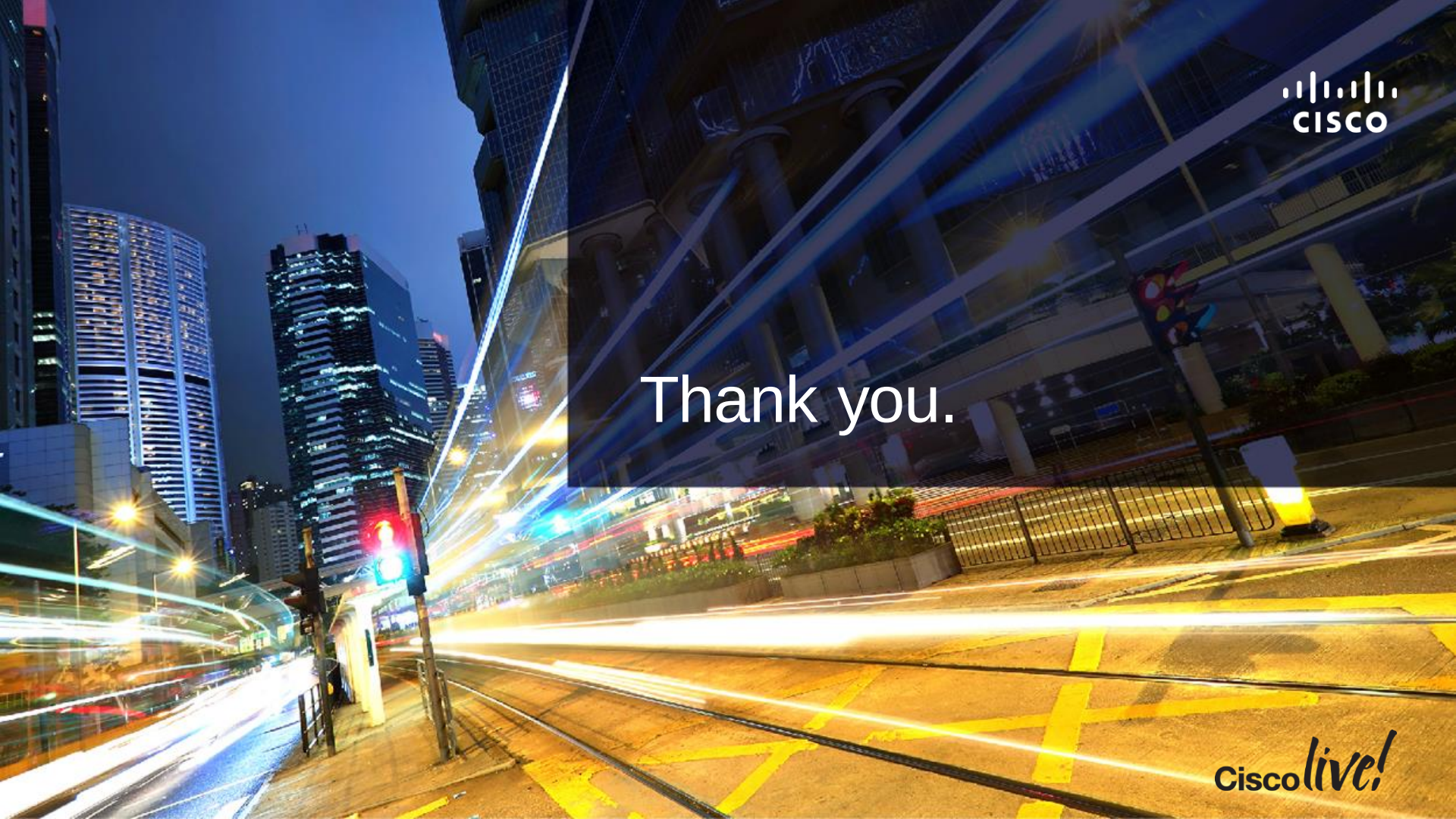
Give us your feedback and receive a Cisco Live 2015 T-Shirt!

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 http://showcase.genie-connect.com/clmelbourne2015
- Visit any Cisco Live Internet Station located throughout the venue

T-Shirts can be collected in the World of Solutions on Friday 20 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


Thank you.
CISCO
Cisco live!


CISCO