



# Virtualization and Data Center 3.0



**Višnja Milovanović, Sistem Inženjer**  
**[vmilovan@cisco.com](mailto:vmilovan@cisco.com)**

# Scalability: The server evolution

## Scale Up

*Monolithic servers*  
Large number of CPUs  
Proprietary platform  
Proprietary OS  
Many apps per server

## Scale Out

*Commoditized servers*  
Few CPUs  
X86 platform  
Commoditized OS  
1 App / server

## Scale In

*Bladed servers*  
Multi-core CPUs  
X86 platform  
Commoditized OS  
1 App / virtual machine

**The 90's**

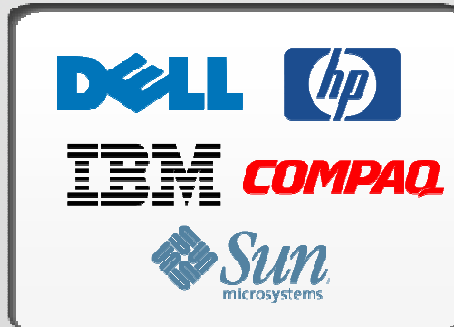
**Early 2000's**

**Now**

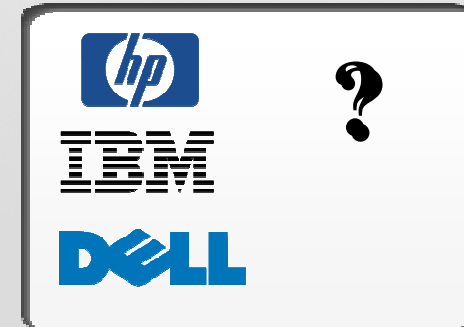
**High cost**  
**Large failure domain**



**Servers under-utilized**  
**Power & cooling**



**Management complexity**  
**Limited scale**



# Server Virtualization - Key DC Trend

## Today

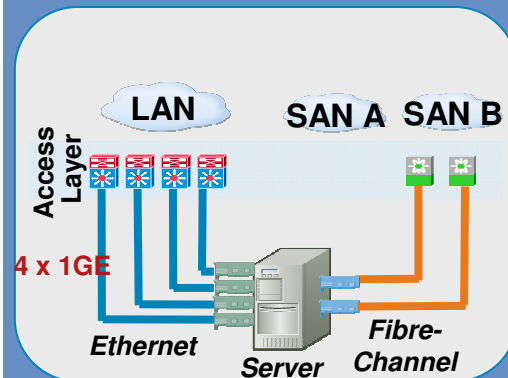


- Many under utilized servers
- Cable sprawl
- High power, cooling costs
- High CAPEX
- For every \$1 spent on server capex ~\$5 spent on opex

## Virtualization Step1

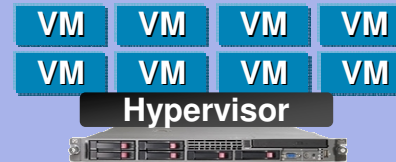


GE

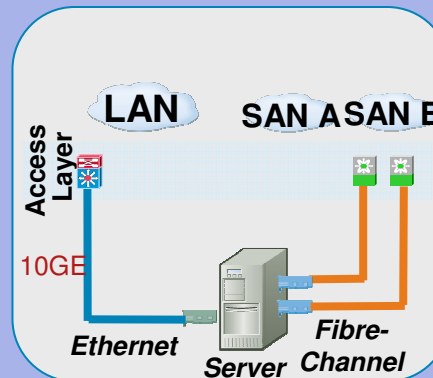


- Cable sprawl
- power, cooling costs
- Less number of access layer Ethernet ports

## Virtualization Step2

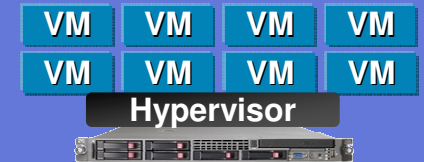


10 GE

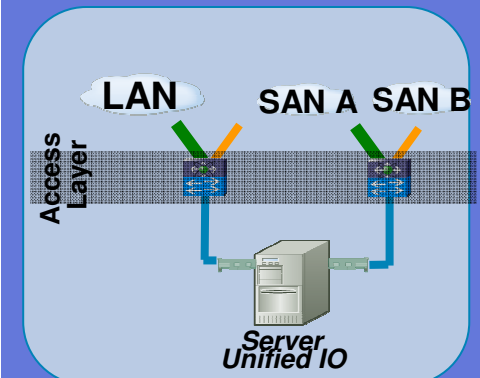


- GE to 10GE in access layer
- Less interfaces – reduced Cable sprawl
- Savings from power and cooling

## Virtualization Step3



10 GE/FCOE



- Unified I/O - LAN & SAN consolidation
- Reduce NICs, HBAs,
- Reduce cabling
- More Savings from power and cooling
- Lower capex

# Data Center Architecture

## Evolving Ethernet Capabilities - DCE



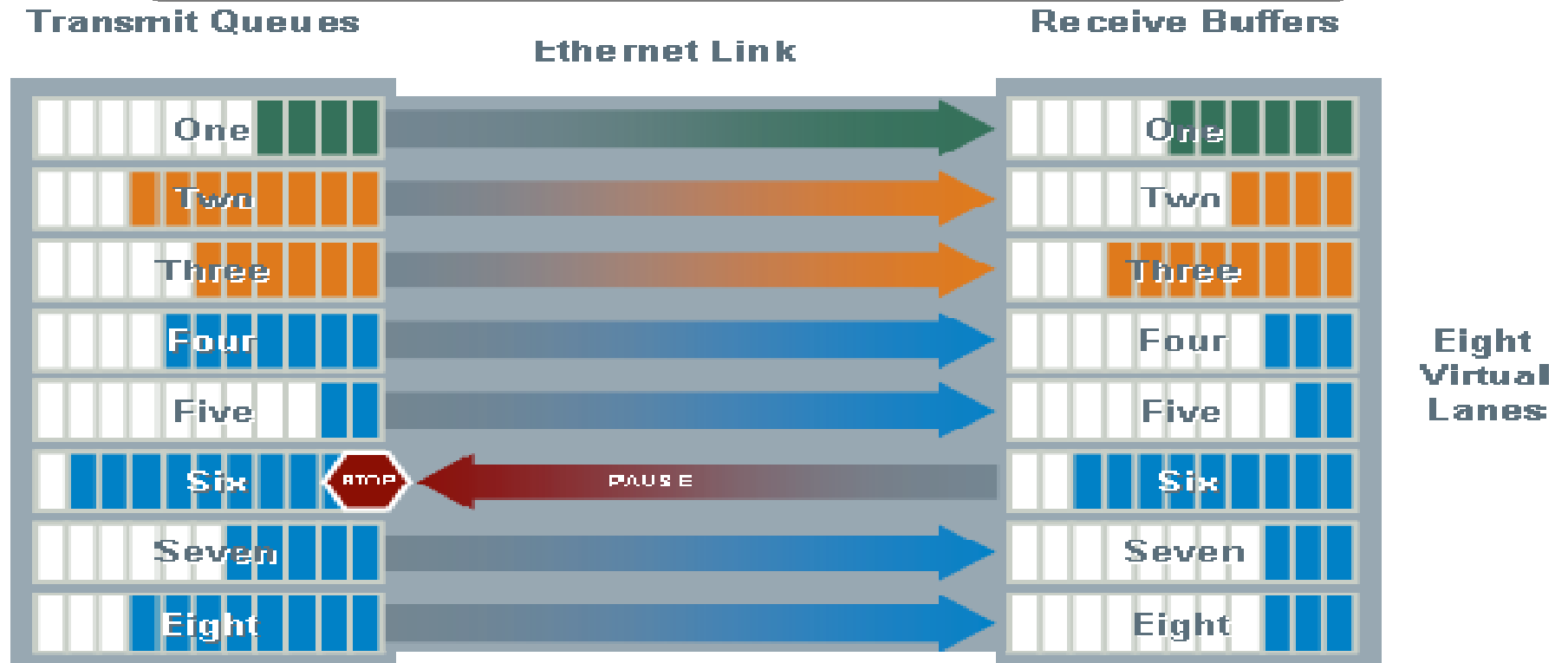
Feature / Standard	Benefit
Priority Flow Control (PFC) <b>IEEE 802.1Qbb</b>	Enable multiple traffic types to share a common Ethernet link without interfering with each other
Bandwidth Management <b>IEEE 802.1Qaz</b>	Enable consistent management of QoS at the network level by providing consistent scheduling
Congestion Management <b>IEEE 802.1Qau</b>	End-to-end congestion management for L2 network (future)
Data Center Bridging Exchange Protocol ( <b>DCBX</b> )	Management protocol for enhanced Ethernet capabilities
L2 Multipath for Unicast and Multicast	Increase bandwidth, multiple active paths. No spanning tree (future)

***Enabling Differentiated Services in an Ethernet Fabric***

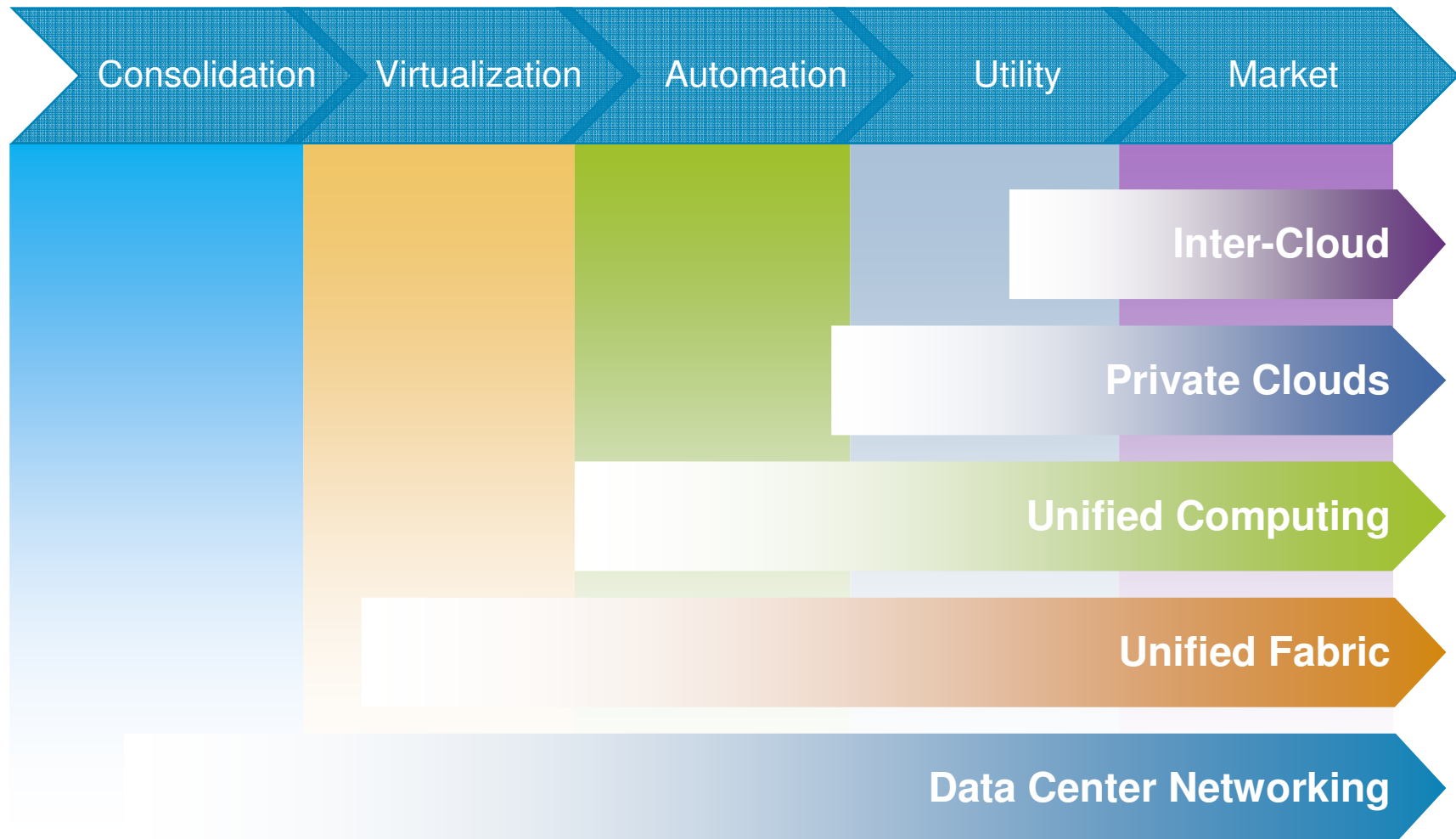


# Priority Flow Control

## *Priority based Flow Control*

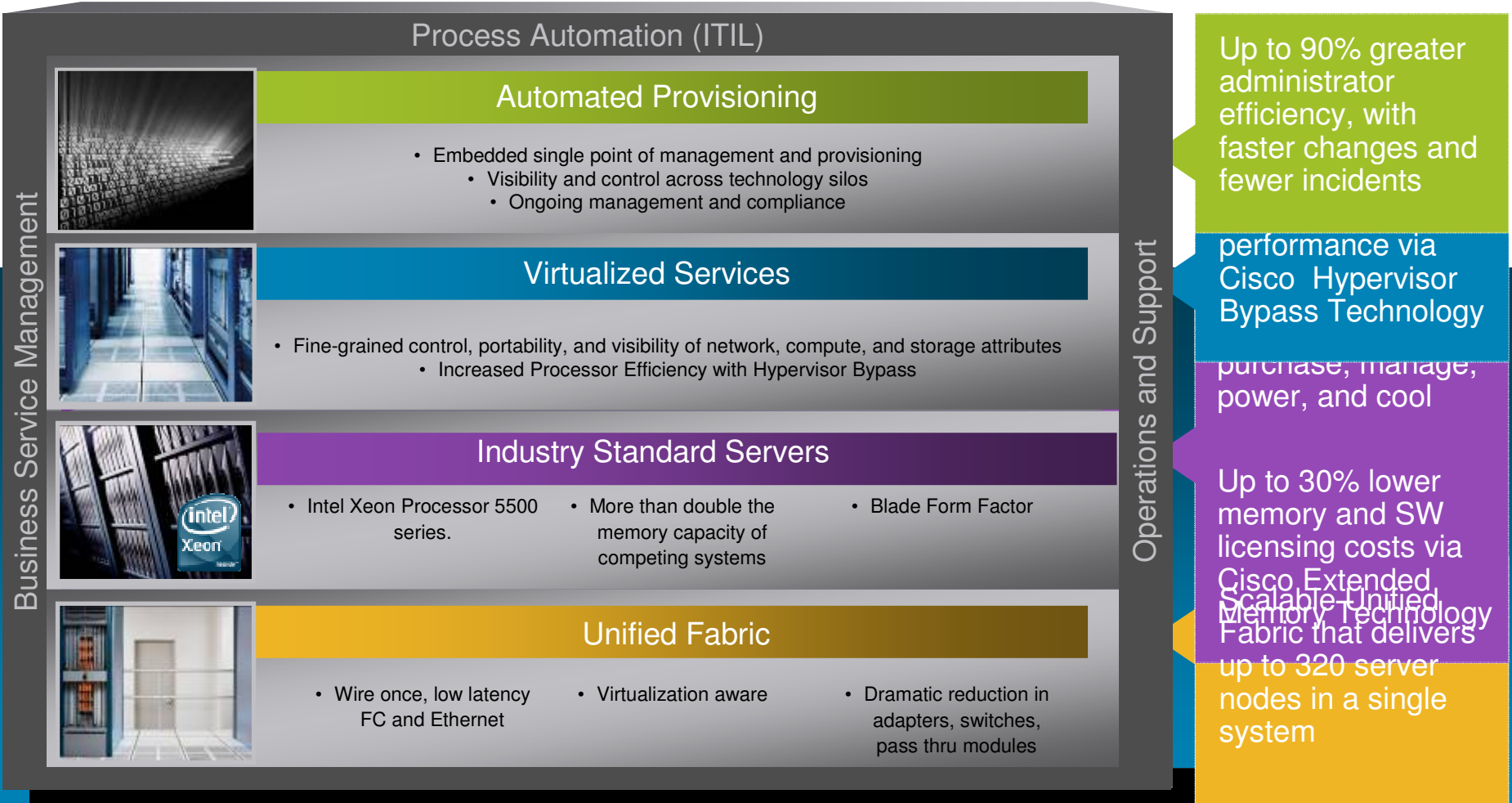


# Unified Computing Continues Data Center 3.0 Strategy

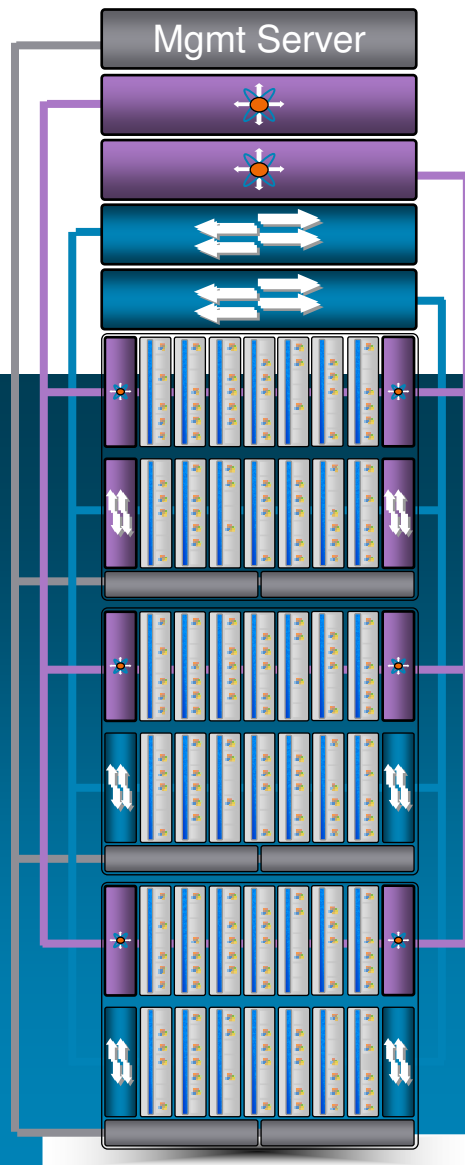


# Cisco Unified Computing System

The Cisco Unified Computing System is designed to dramatically reduce datacenter total cost of ownership while simultaneously increasing IT agility and responsiveness.

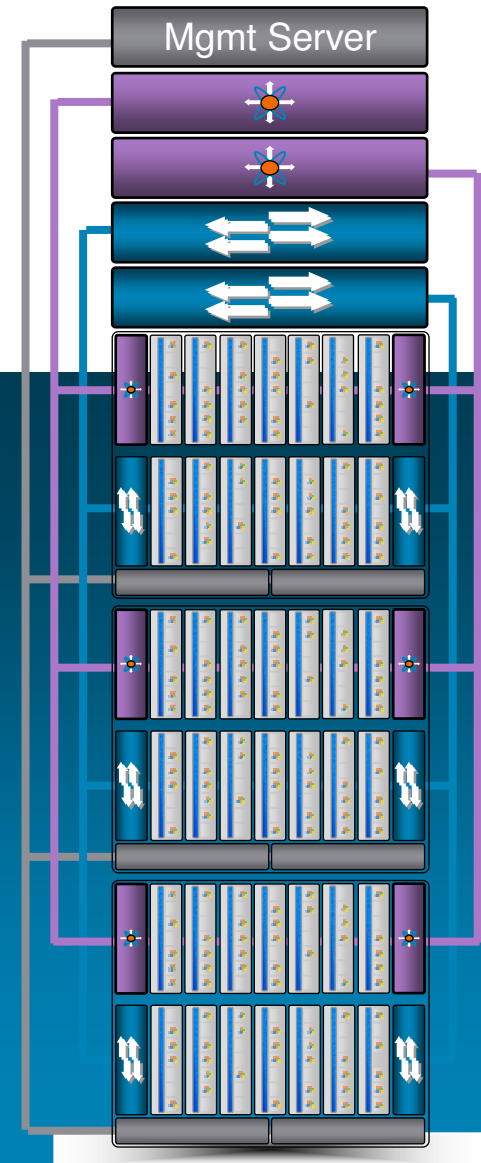
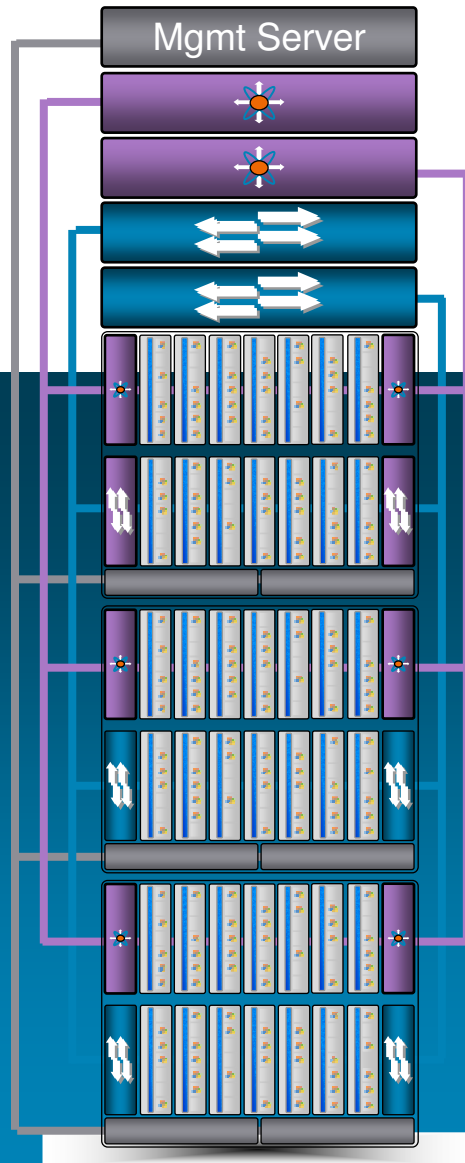


# Simplifying the Data Center



# Simplifying the Data Center

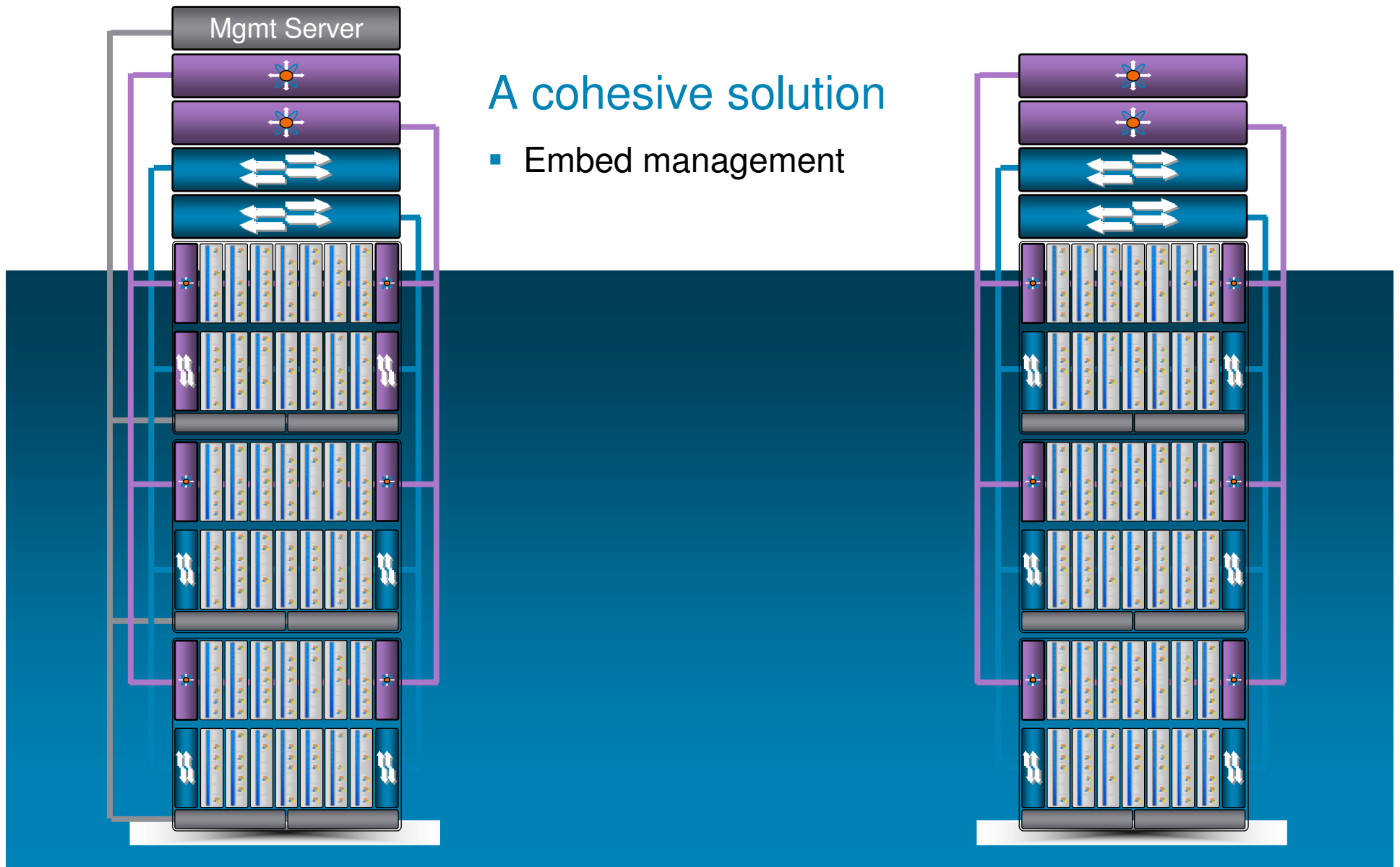
## A cohesive solution



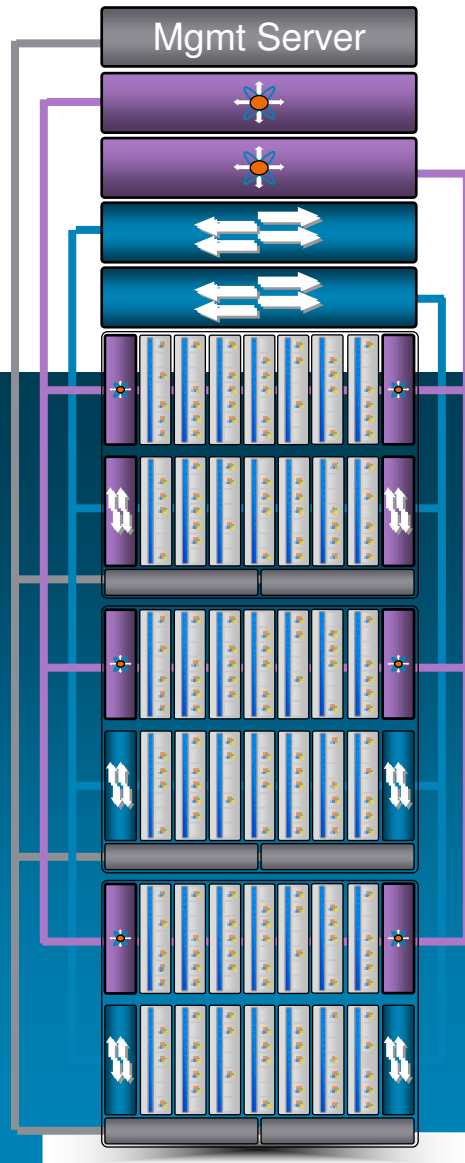
# Simplifying the Data Center

A cohesive solution

- Embed management

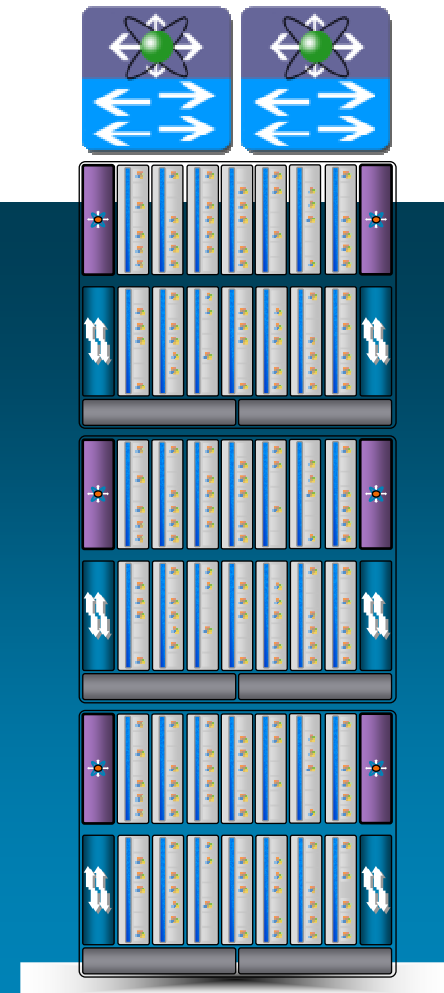


# Simplifying the Data Center



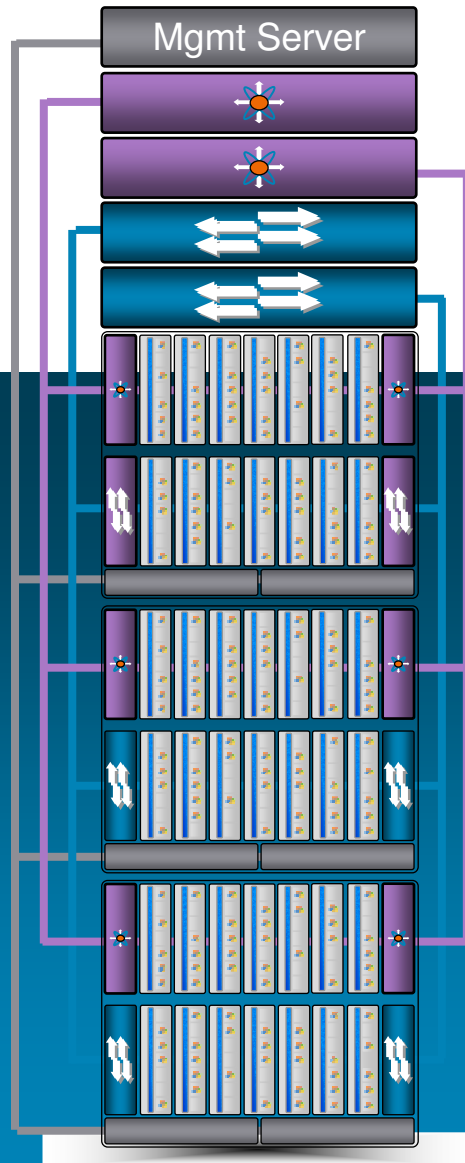
## A cohesive solution

- Embed management
- Unify fabrics



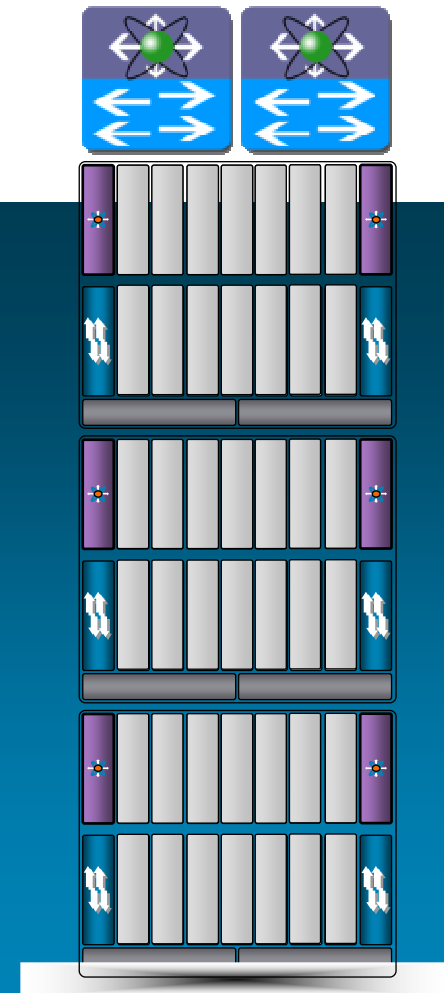


# Simplifying the Data Center

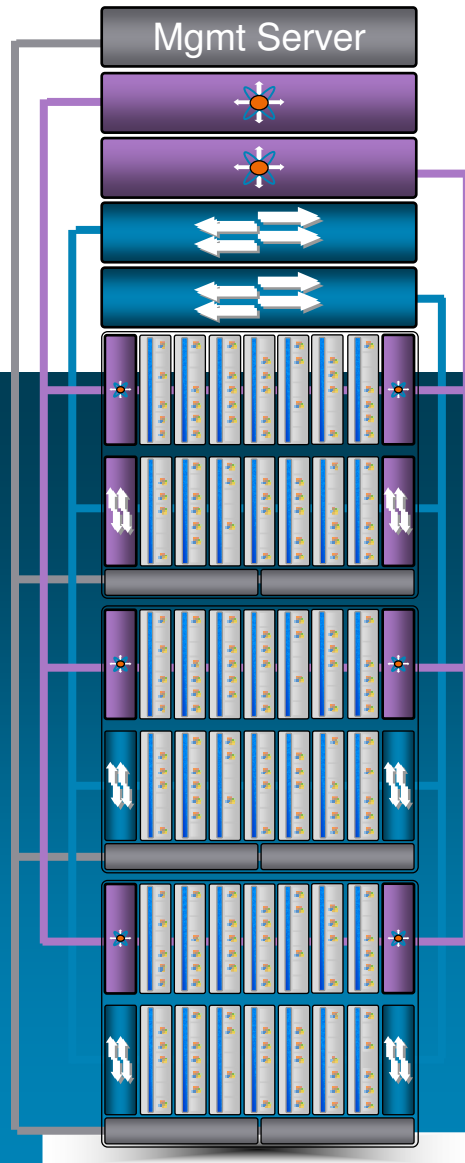


## A cohesive solution

- Embed management
- Unify fabrics
- Optimize virtualization

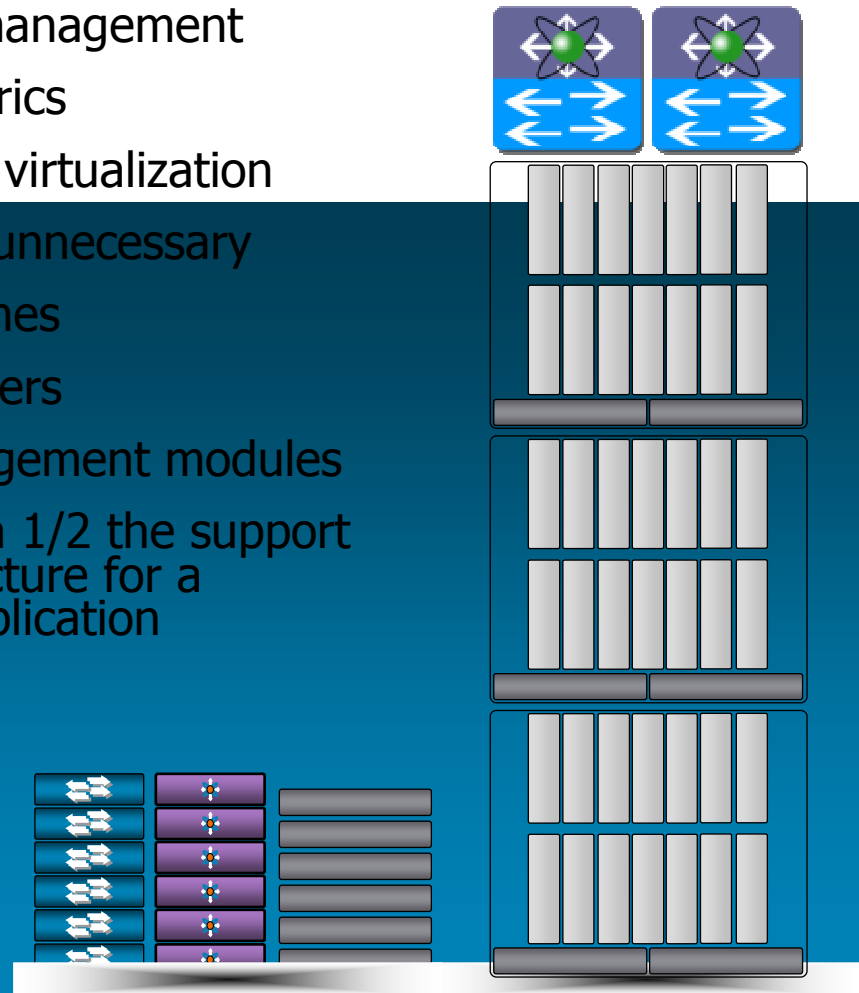


# Simplifying the Data Center



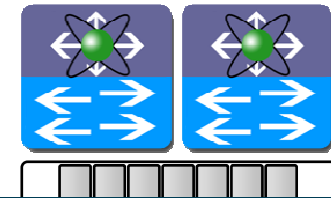
## A cohesive solution

- Embed management
- Unify fabrics
- Optimize virtualization
- Remove unnecessary  
Switches  
Adapters  
Management modules
- Less than 1/2 the support infrastructure for a given application



# Cisco Unified Computing System

- UCS
  - Mgmt Server
  - Scalable compute platform
  - Integrated virtualization
  - Natural aggregation point: Network
- Unified embedded management



# UCS Building Blocks

## **UCS Manager**

Embedded– manages entire system

## **UCS 6100 Series Fabric Interconnect**

20 Port 10Gb FCoE – UCS-6120

40 Port 10Gb FCoE – UCS-6140

## **UCS Fabric Extender – UCS 2100 Series**

Remote line card

## **UCS 5100 Series Blade Server Chassis**

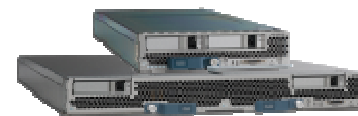
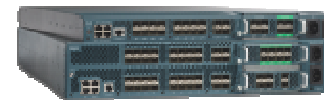
Flexible bay configurations

## **UCS Blade Server**

Industry-standard architecture

## **UCS Adapters**

Choice of multiple **C**onverged **N**etwork **A**dapters  
and Virtual Adapters



# Blade Chassis

- 6RU Chassis

Blades and Power supplies plug-in from front

- Blades

Power & cooling budget allows leading edge processor performance and memory capacity

Combinations of half slot and full slot blades

Up to 8 Half slot blades

Up to 4 Full slot blades

- Power Supplies

4x 2,500W hot-plug Power Supplies

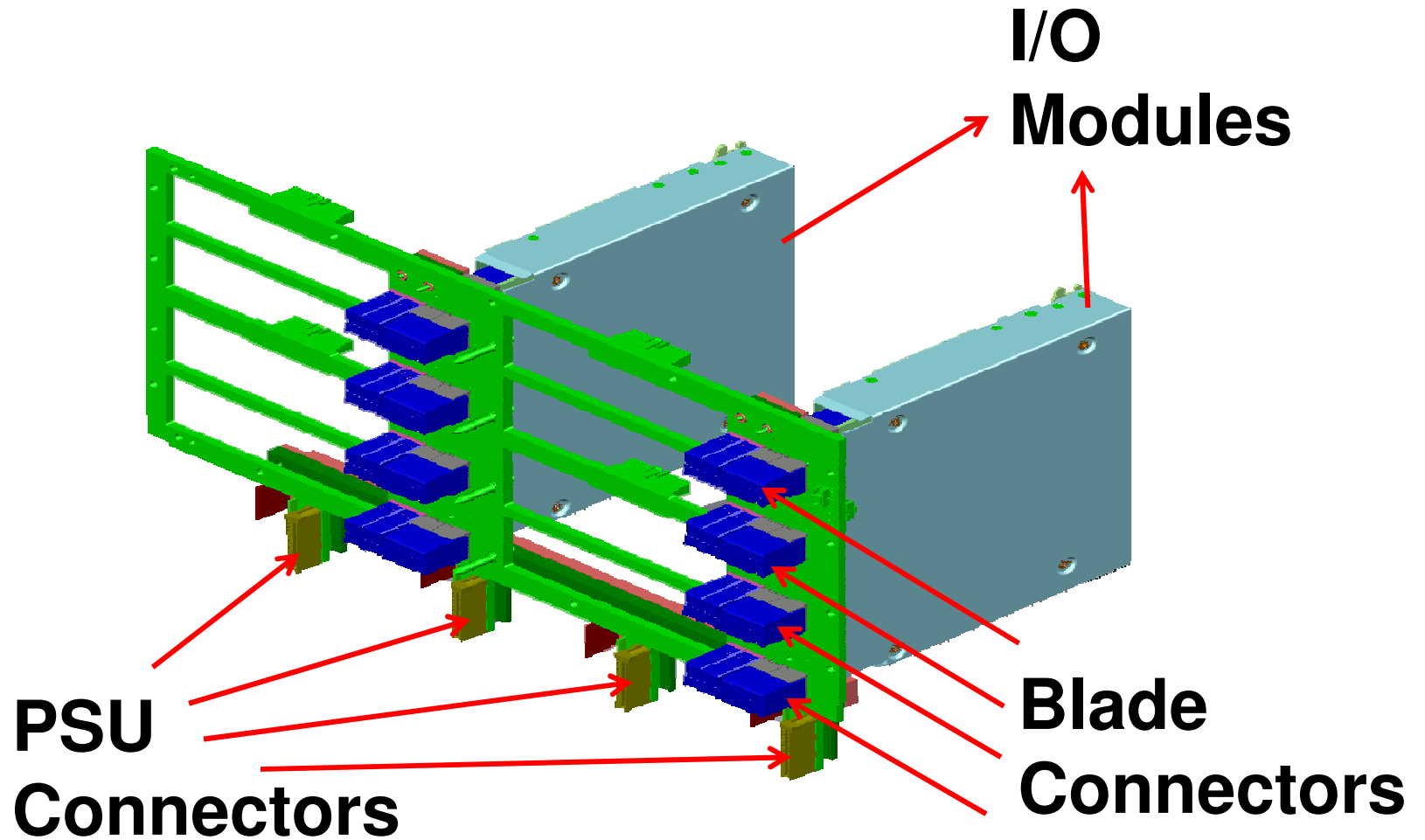
90% efficient

N+N redundancy (grid redundant)

4 single phase 220V, IEC320-C19

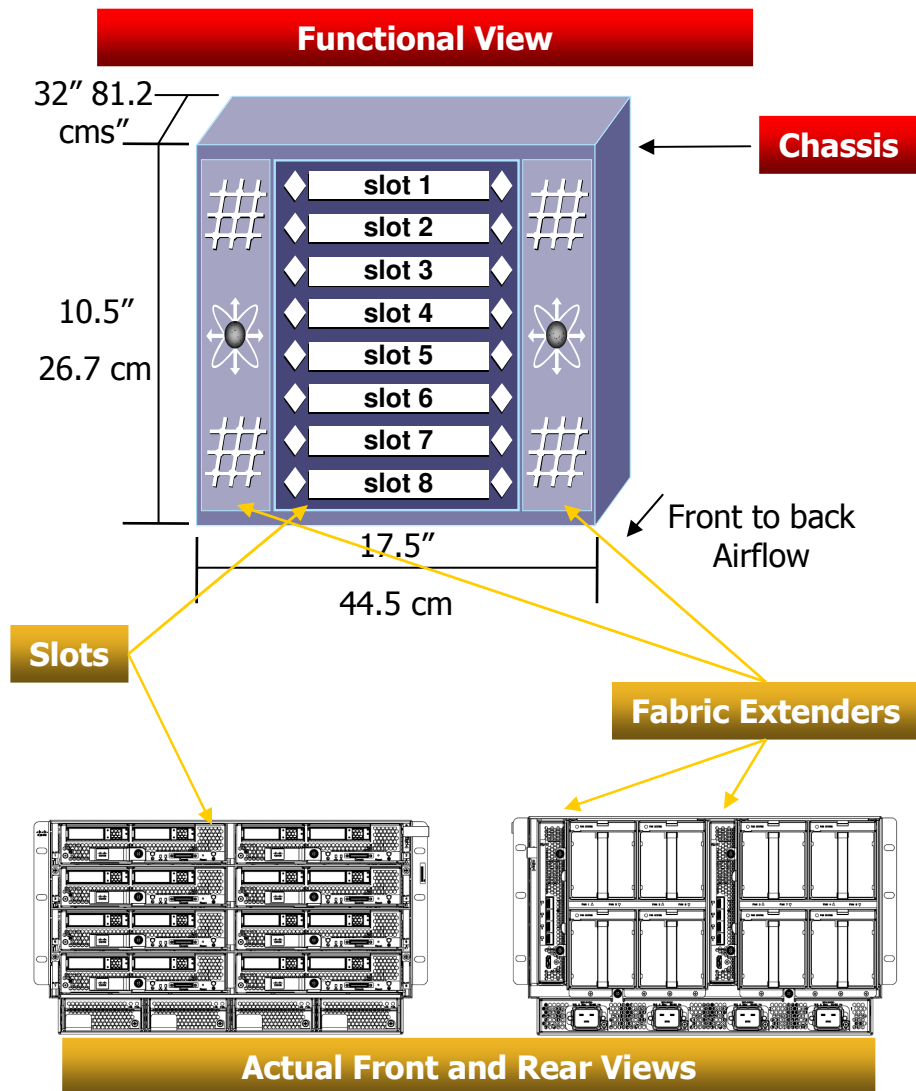


# Backplane



**Redundant data and management paths**

# Blade Chassis – UCS 5108 Details



## UCS 5108 Chassis Characteristics

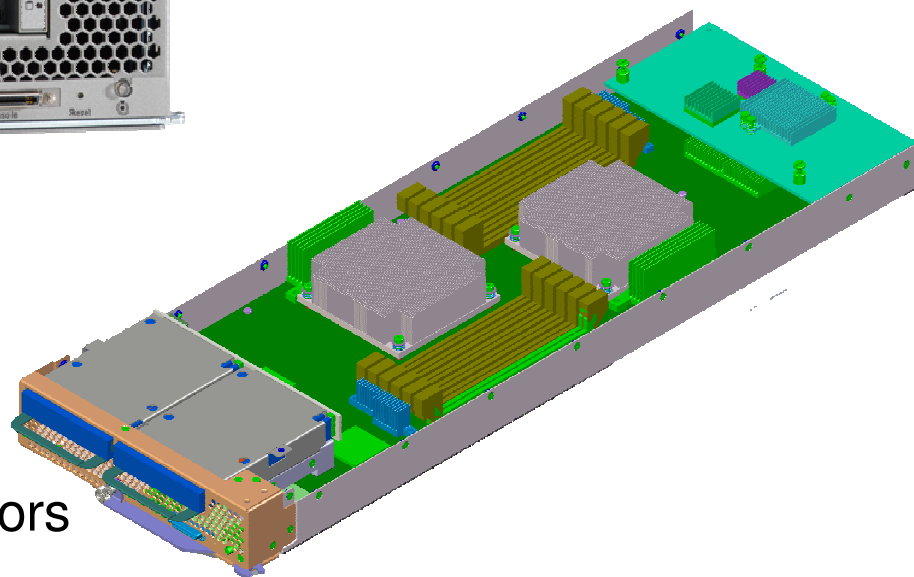
- ❖ 8 Blade Slots
  - 8 half-width servers or;
  - 4 full-width servers
- ❖ Up to two Fabric Extenders
  - Both concurrently active
- ❖ Redundant and Hot Swappable
  - Power Supplies
  - Fan Modules

## Additional Chassis Details:

- ❖ Size: 10.5" (6U) x 18.5" x 32"
- ❖ Power Consumption (nominal estimates)
  - Half-width Servers: 1.5 – 3 kW
  - Full-width Servers: 1.5 – 3 kW
- ❖ Cabling: SFP+ Connectors
  - For FEX to Fabric Interconnect
  - CX1: 3 & 5 meters
  - USR: 100 meters
  - SR: 300 meters
- ❖ Airflow: front to back



# UCS B200 M1 Blade

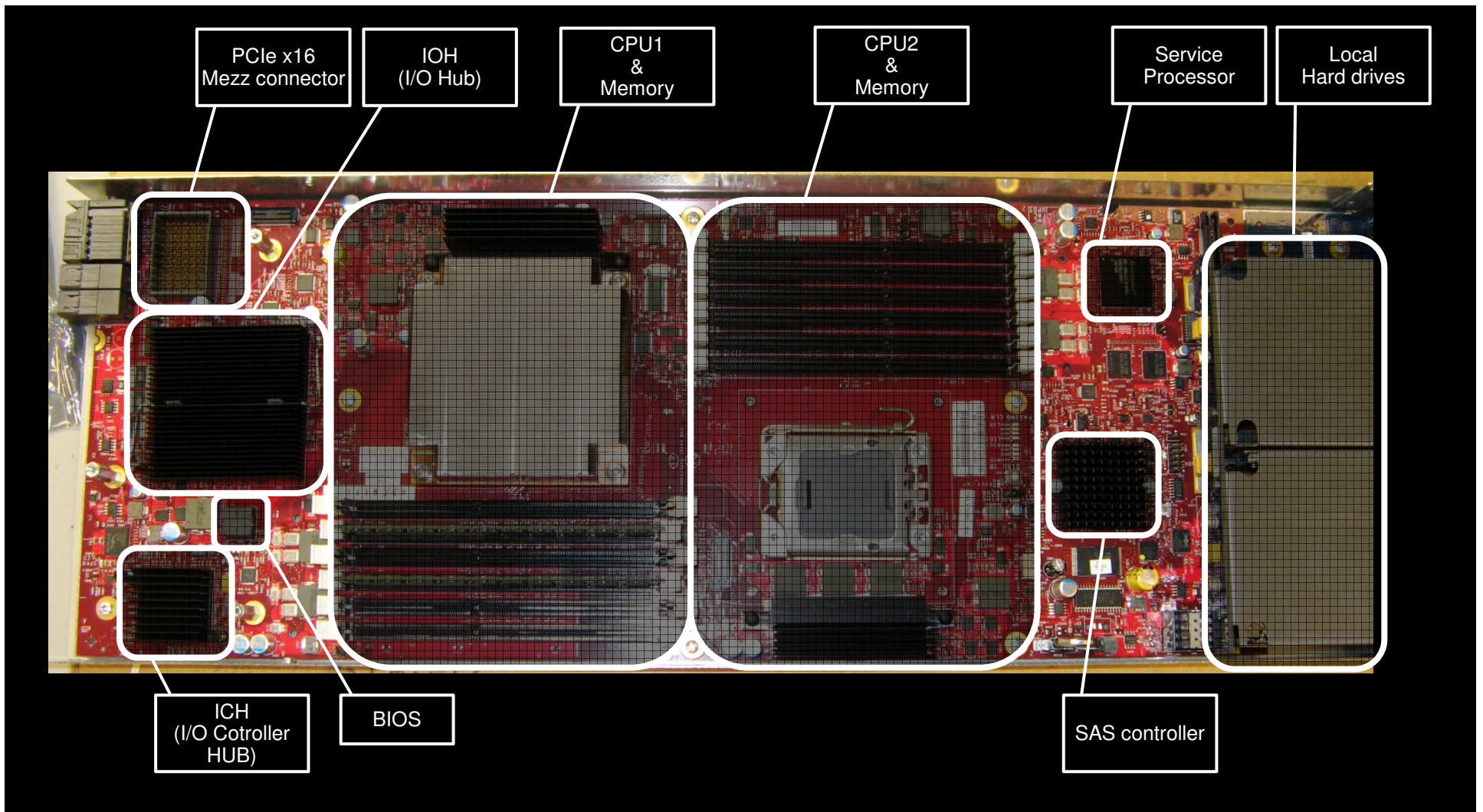


## **Blade Attributes:**

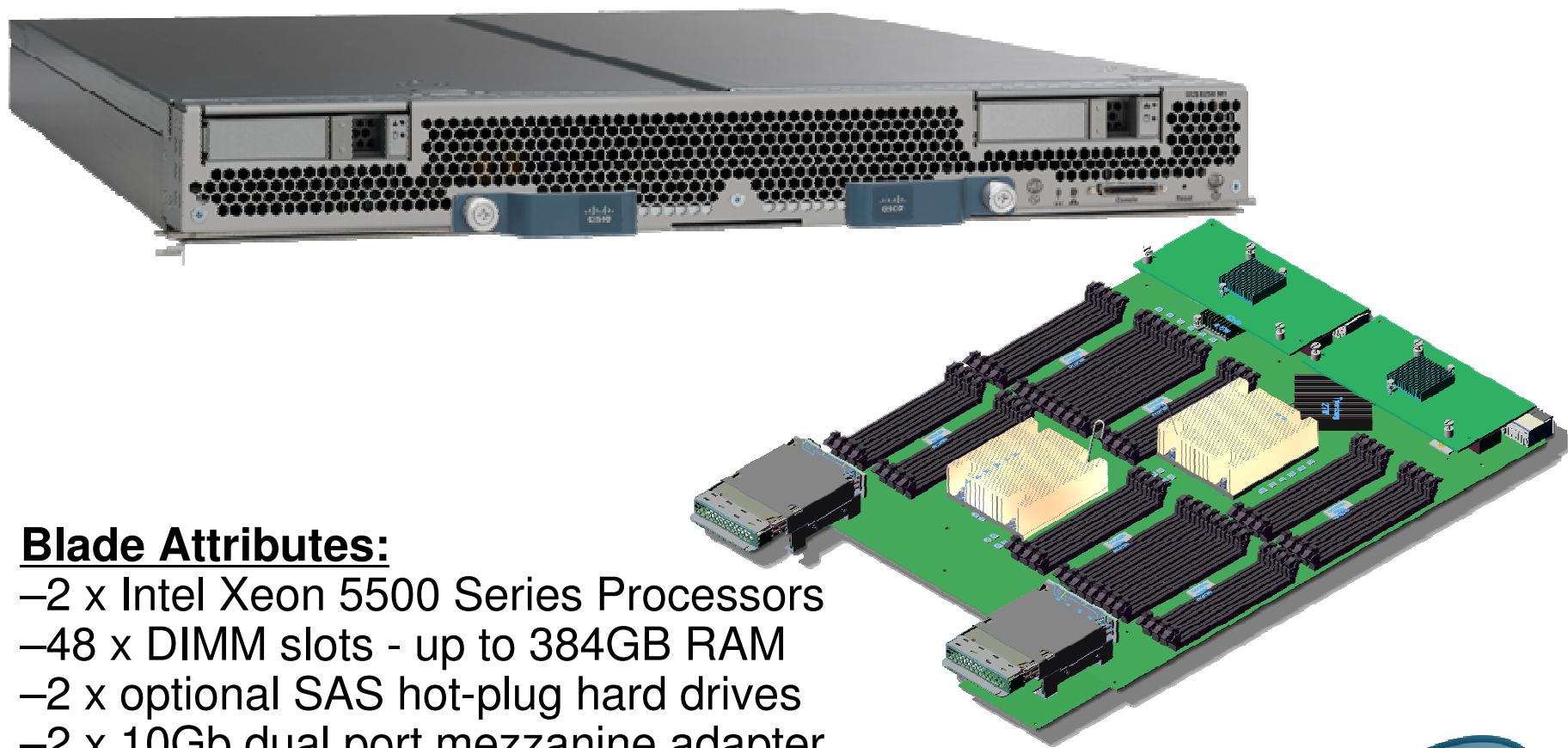
- 2 x Intel Xeon 5500 Series Processors
- 12 x DIMM slots - up to 96GB RAM
- 2 x optional SAS hot-plug hard drives
- 1 x 10Gb dual port mezzanine adapter
- Remote and local access to keyboard, video, mouse, serial
- Integrated with UCS Manager



# UCS B200 M1 overview



# UCS B250 M1 Blade

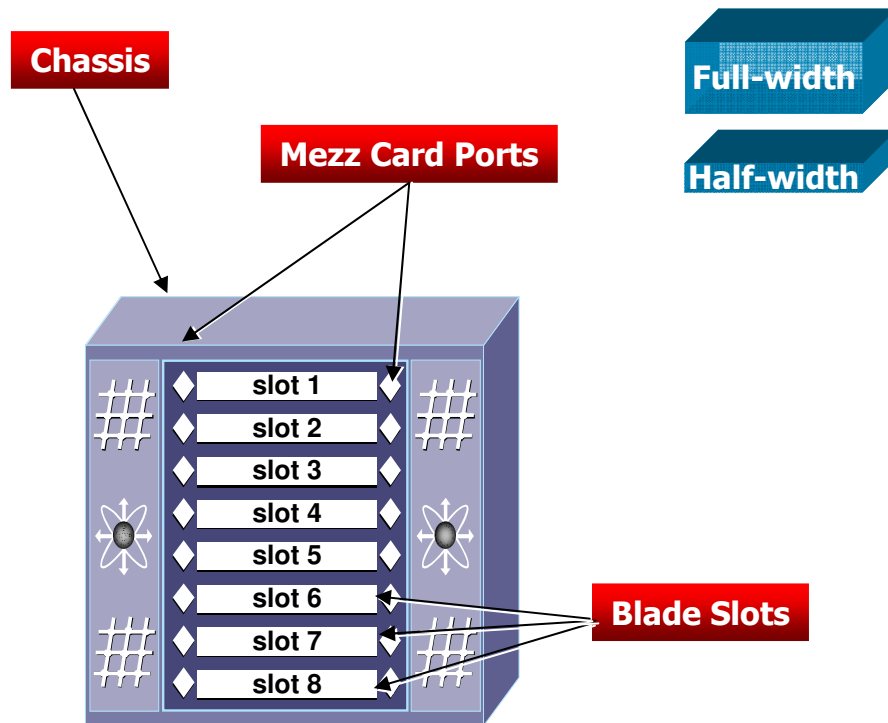


## **Blade Attributes:**

- 2 x Intel Xeon 5500 Series Processors
- 48 x DIMM slots - up to 384GB RAM
- 2 x optional SAS hot-plug hard drives
- 2 x 10Gb dual port mezzanine adapter
- Remote and local access to keyboard, video, mouse, serial
- Integrated with UCS Manager



# Blades, Slots and Mezz Cards



## B-Series Blades

- ❖ Half-width Blade – B200-M1
  - each blade uses one slot
- ❖ Full-width – B250-M1
  - each blade uses two slots
  - Always 1,2 or 3,4, or 5,6, or 7,8

## Slots and Mezz Cards

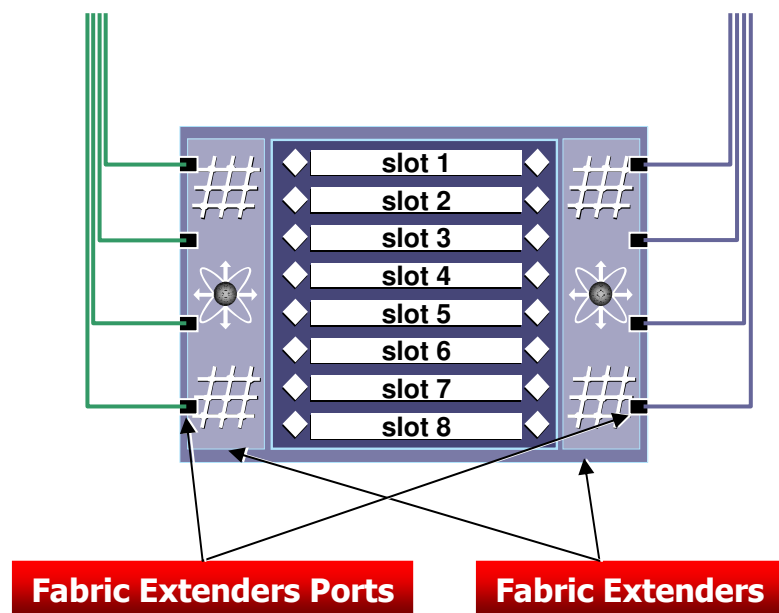
- ❖ Each Slot Takes a Single Mezz Card
- ❖ Each Mezz Card Has 2 Ports
  - each port connect to one FEX
- ❖ Total of 8 Mezz Cards Per Chassis

## Blades, Slots and Mezz Cards

- ❖ Half-width Blade – B200-M1
  - One mezz card two ports per server
  - Each Port to different FEX
- ❖ Full-width Blade – B250-M1
  - Two mezz cards four ports per server
  - Two: from each mezz card to each FEX



# UCS 2100 Series Fabric Extenders Details



## UCS 2104 XP Fabric Extender Port Usage

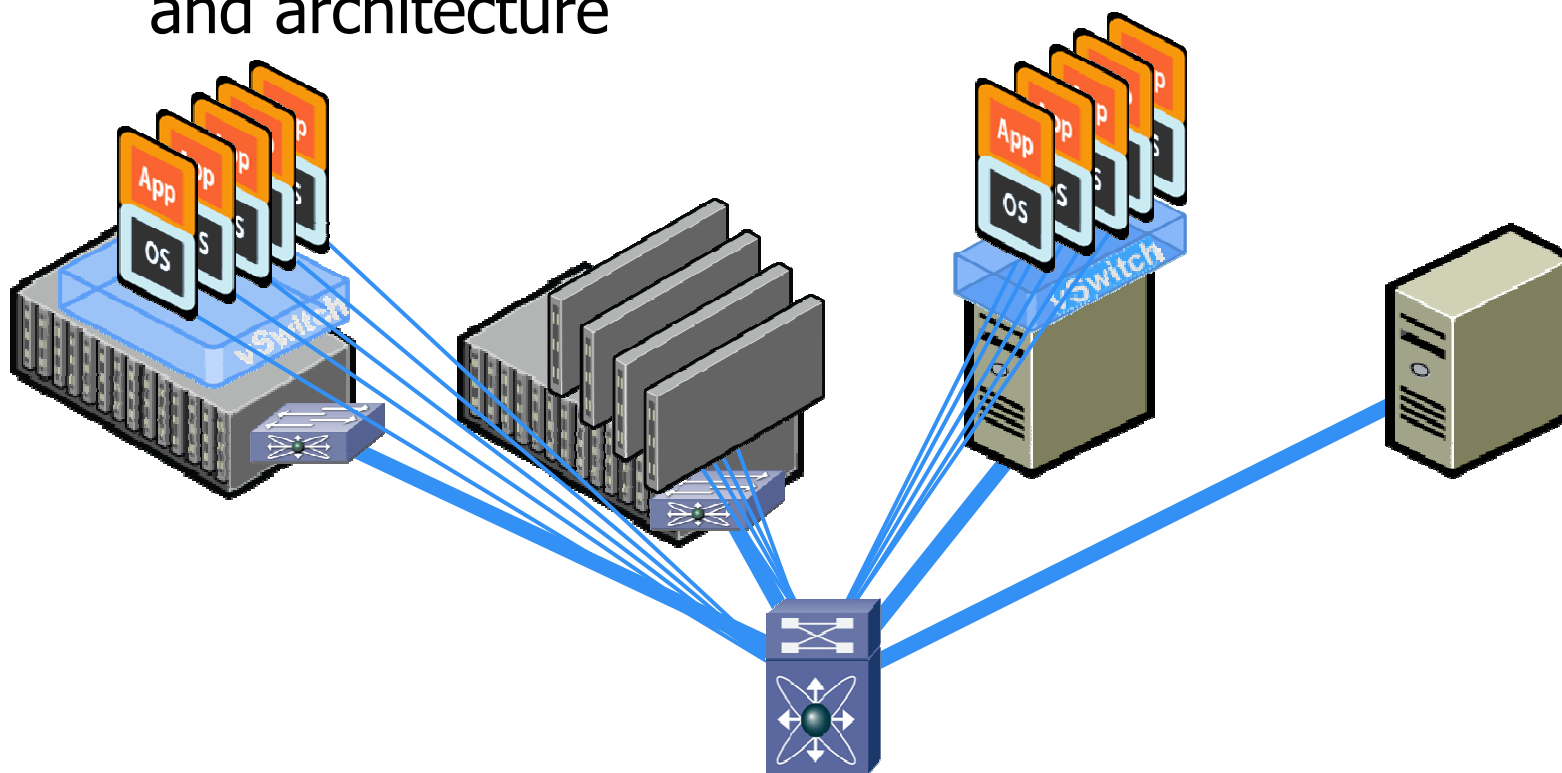
- ❖ Fabric Extenders per Chassis: 2
- ❖ Uplinks per Extender: 4x10GE
- ❖ Port Usage: 1, 2 or 4 uplinks
- ❖ Single FEX Uplink Connectivity:
  - All uplinks only one fabric interconnect
- ❖ Port Combination Across Fabric Extenders
  - Port count must match across FEXs
  - Any port on FEX could be utilized

## Fabric Extender Capabilities

- ❖ Managed as part of UC System
- ❖ 802.1q Trunking
- ❖ FCoE Capabilities
- ❖ Uplink Traffic Distribution
  - Selected when blades are inserted
  - Slot assignment occurs at power up
  - All slot traffic is assigned to a single uplink
  - Any port could be utilized
- ❖ Other Logic
  - Monitor and control of environmentals
  - Blade insertion/removal events

# Introducing Network Interface Virtualization (NIV)

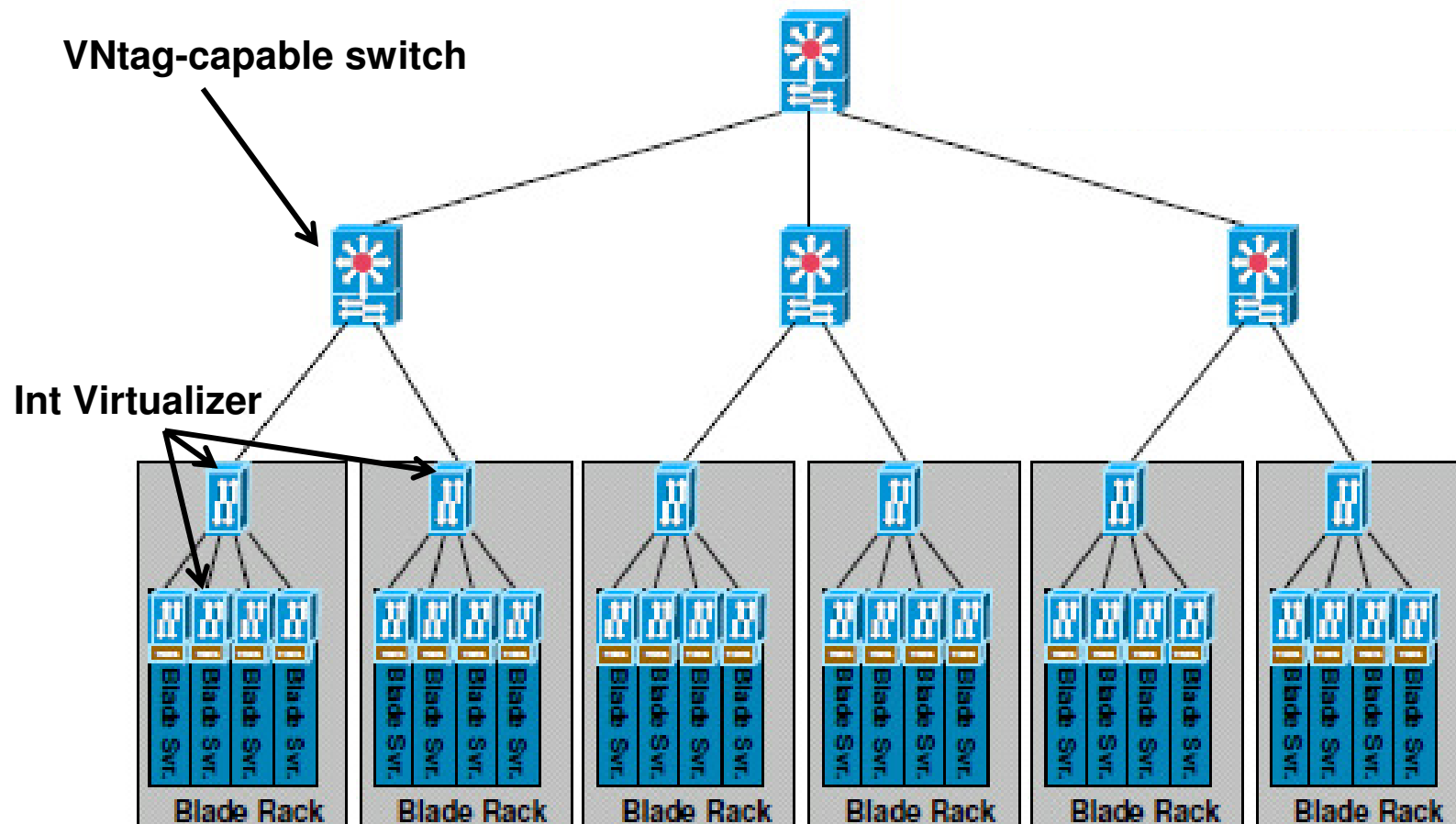
- Logically collapse access layer to simplify management and architecture



Proposal to

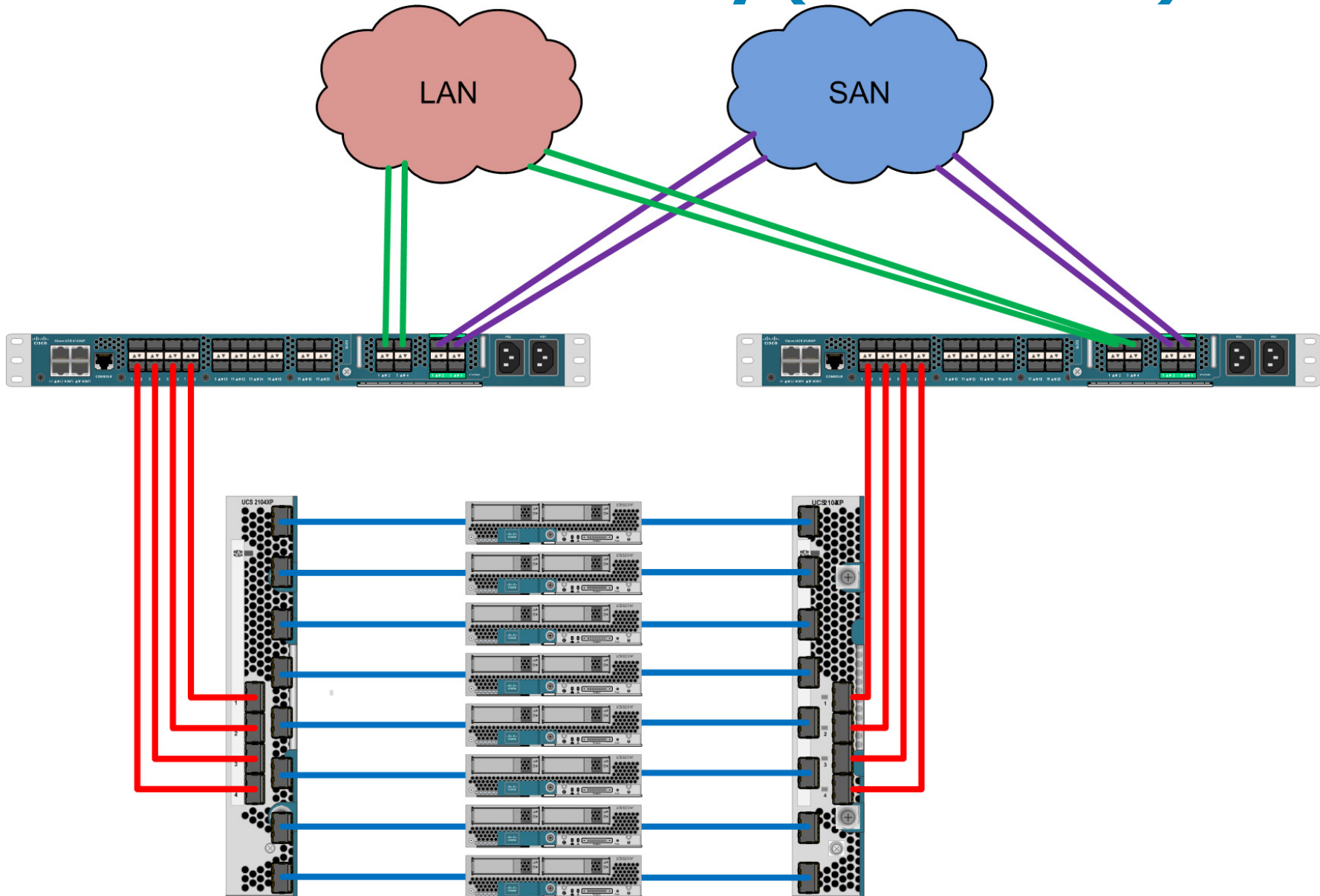


# VNtag: Bridges and Interface Virtualizers

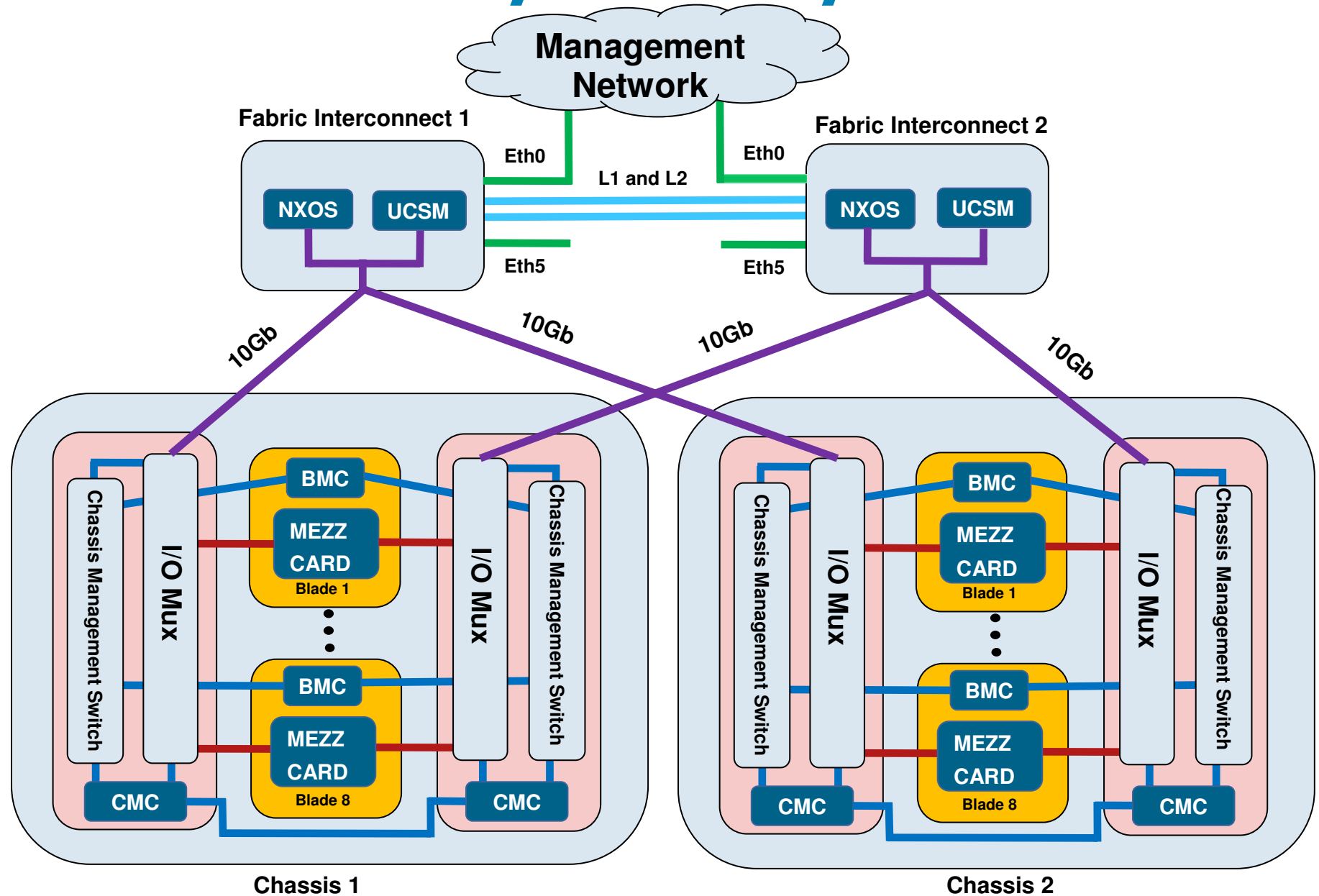




# UCS Data Connectivity (Data Plane)

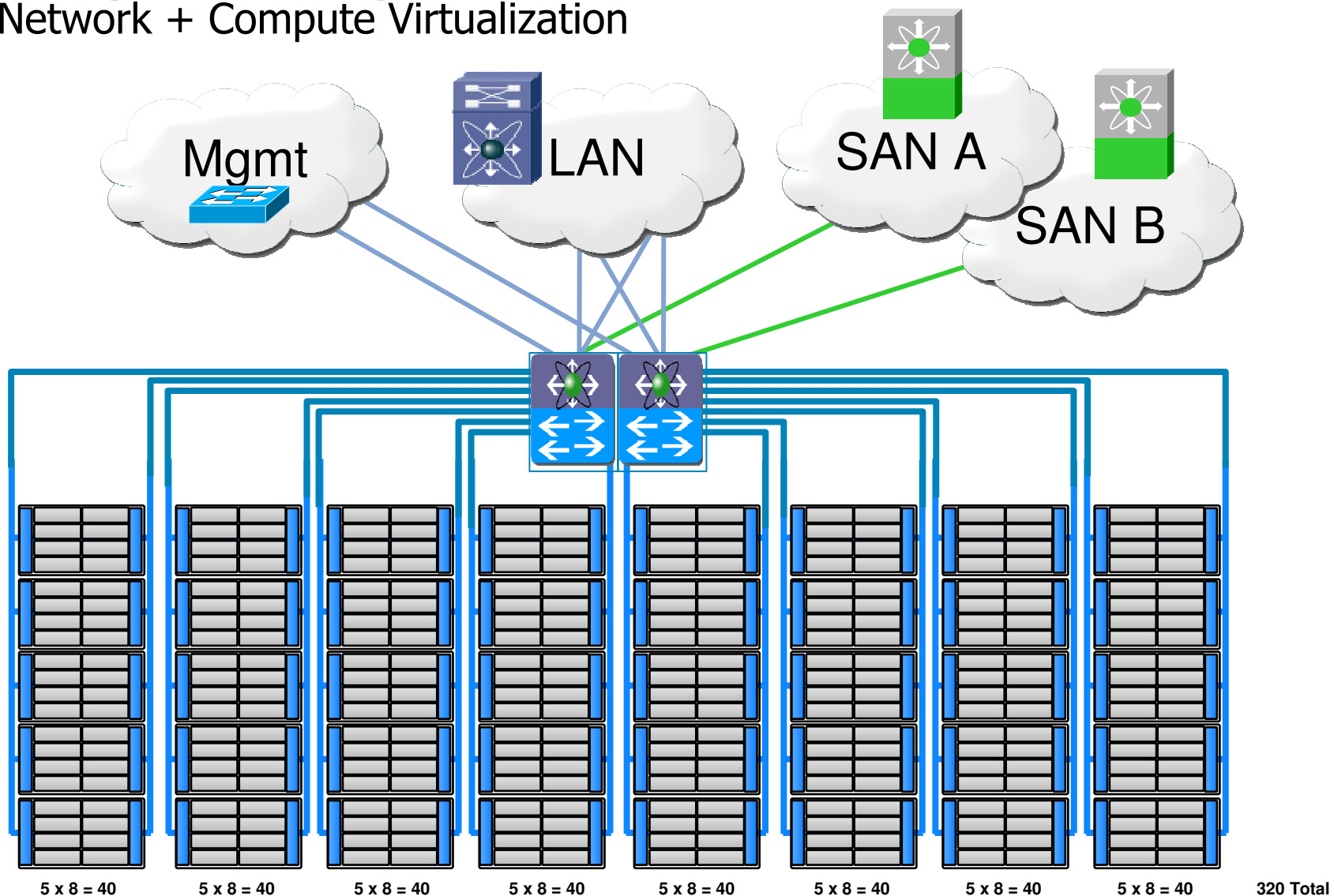


# UCS Connectivity Summary

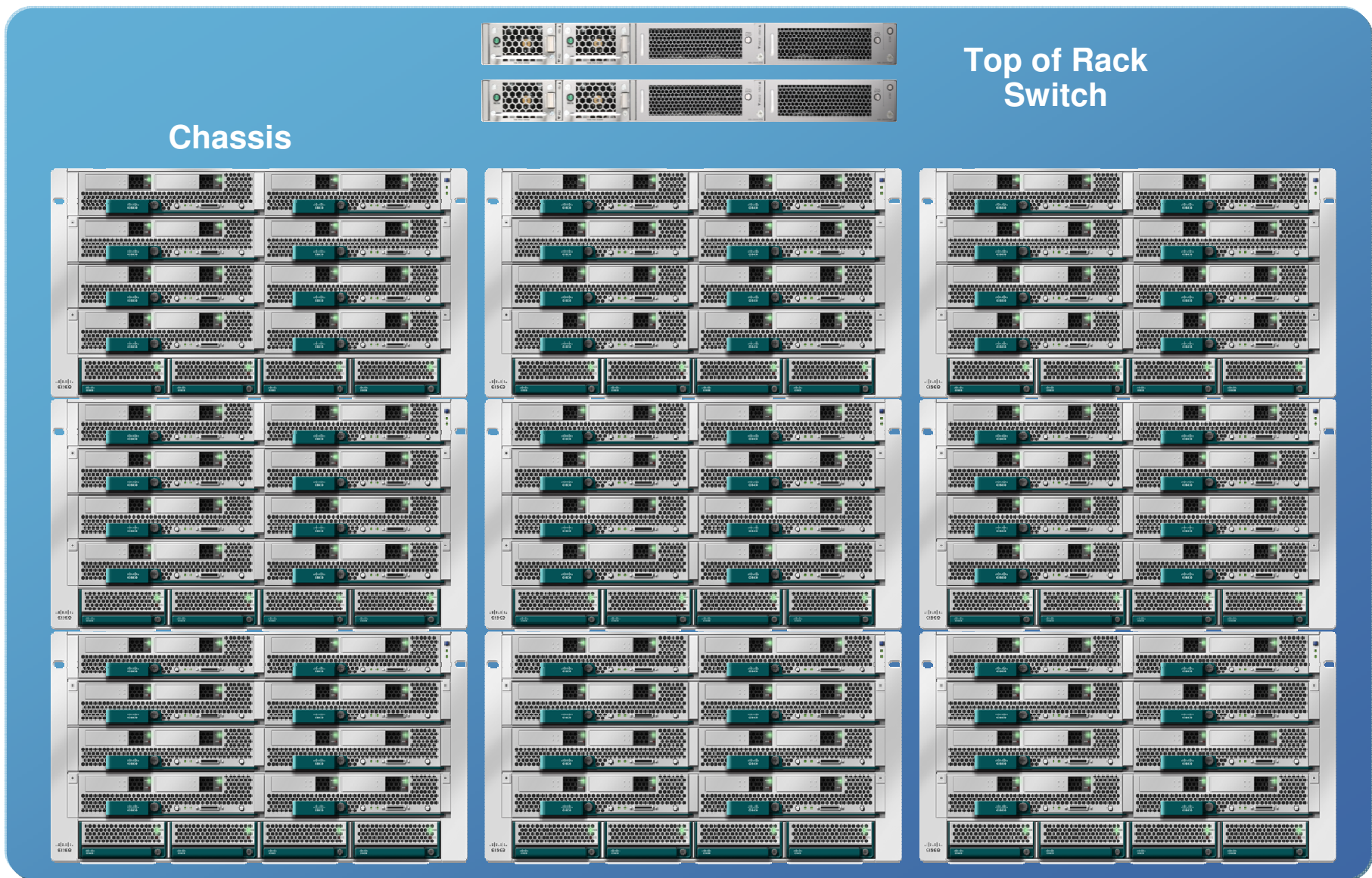


# Single Integrated System

- Network + Compute Virtualization

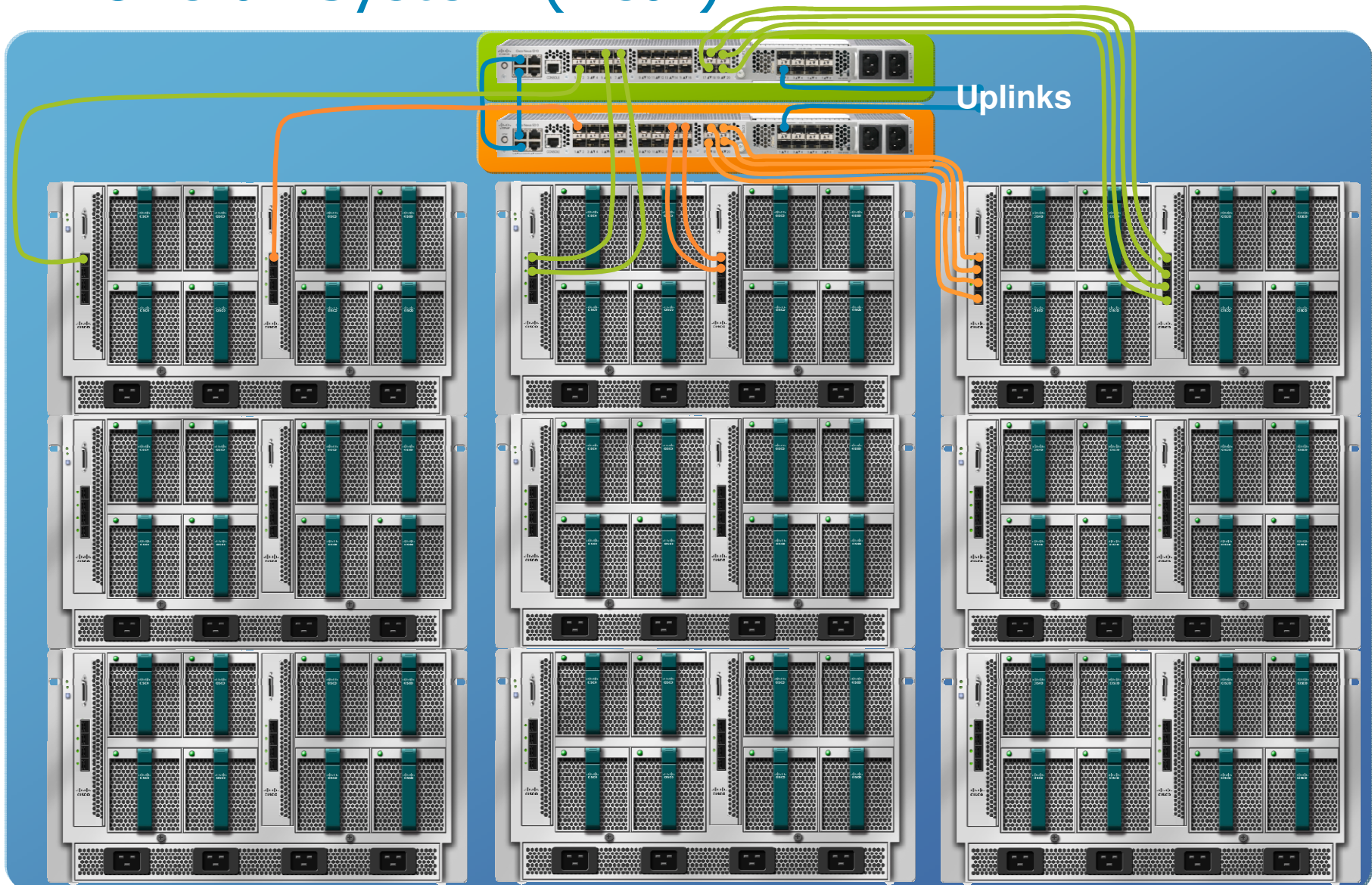


# Overall System (Front)





# Overall System (Rear)





# Memory Extension

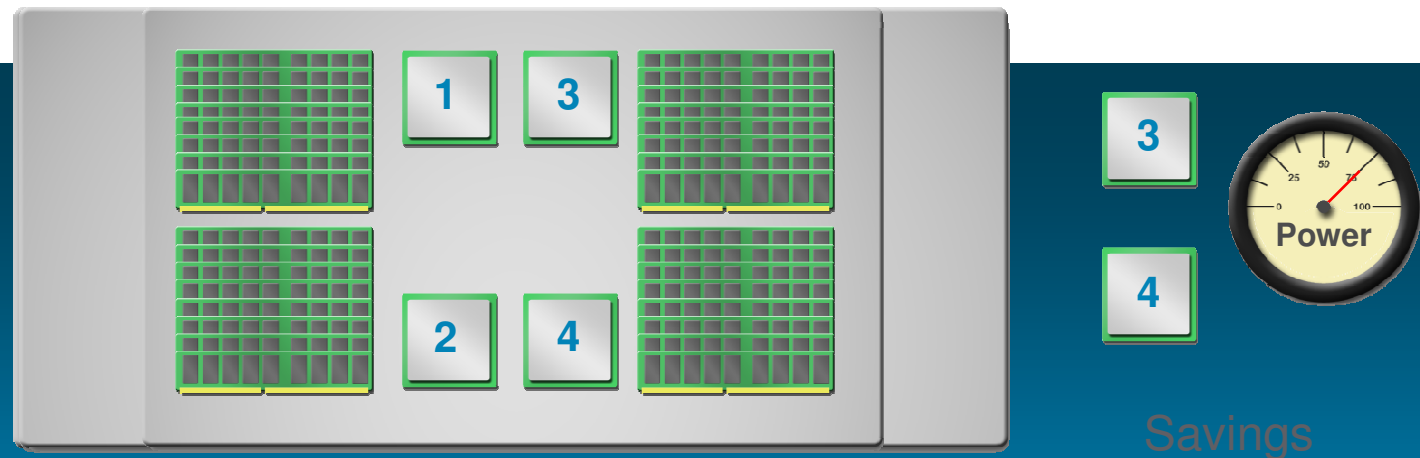


# Memory Expansion

Higher server consolidation & larger VM density

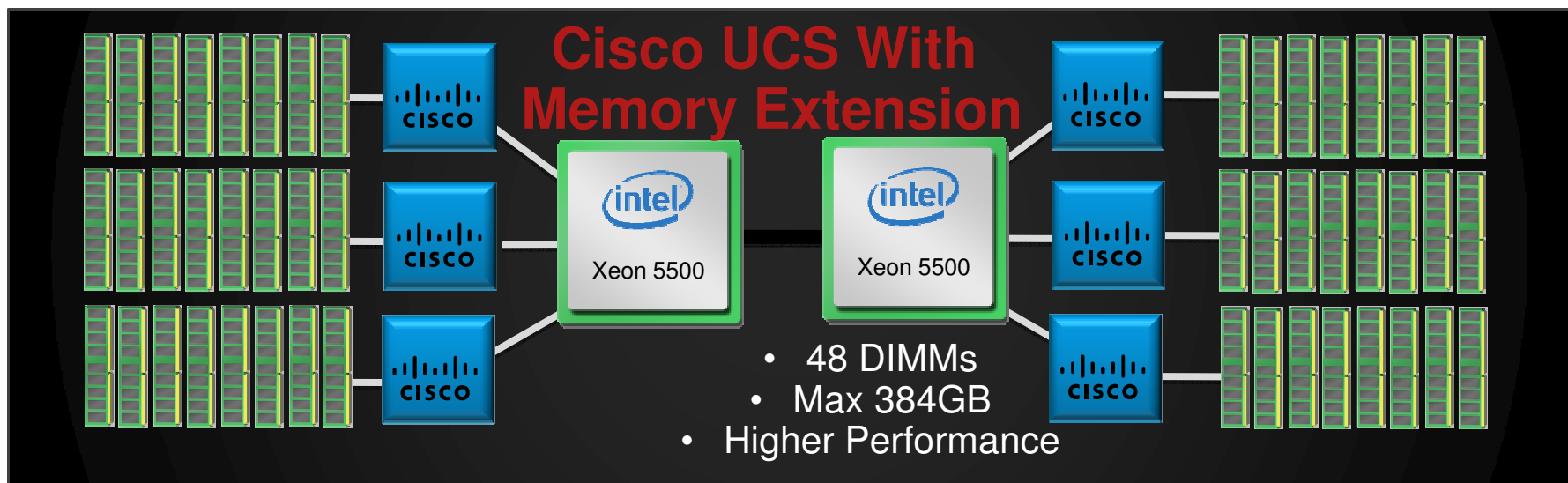
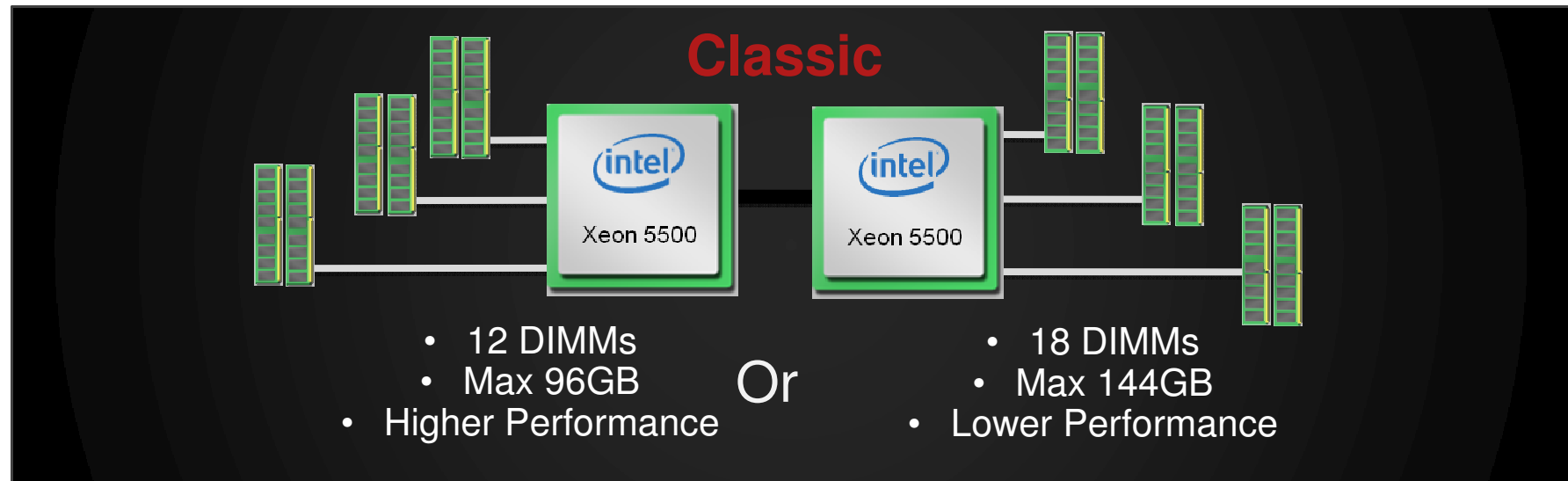
Reduces CPU, power/cooling, and SW licensing costs

Competition - max 18 & high density with proprietary tech

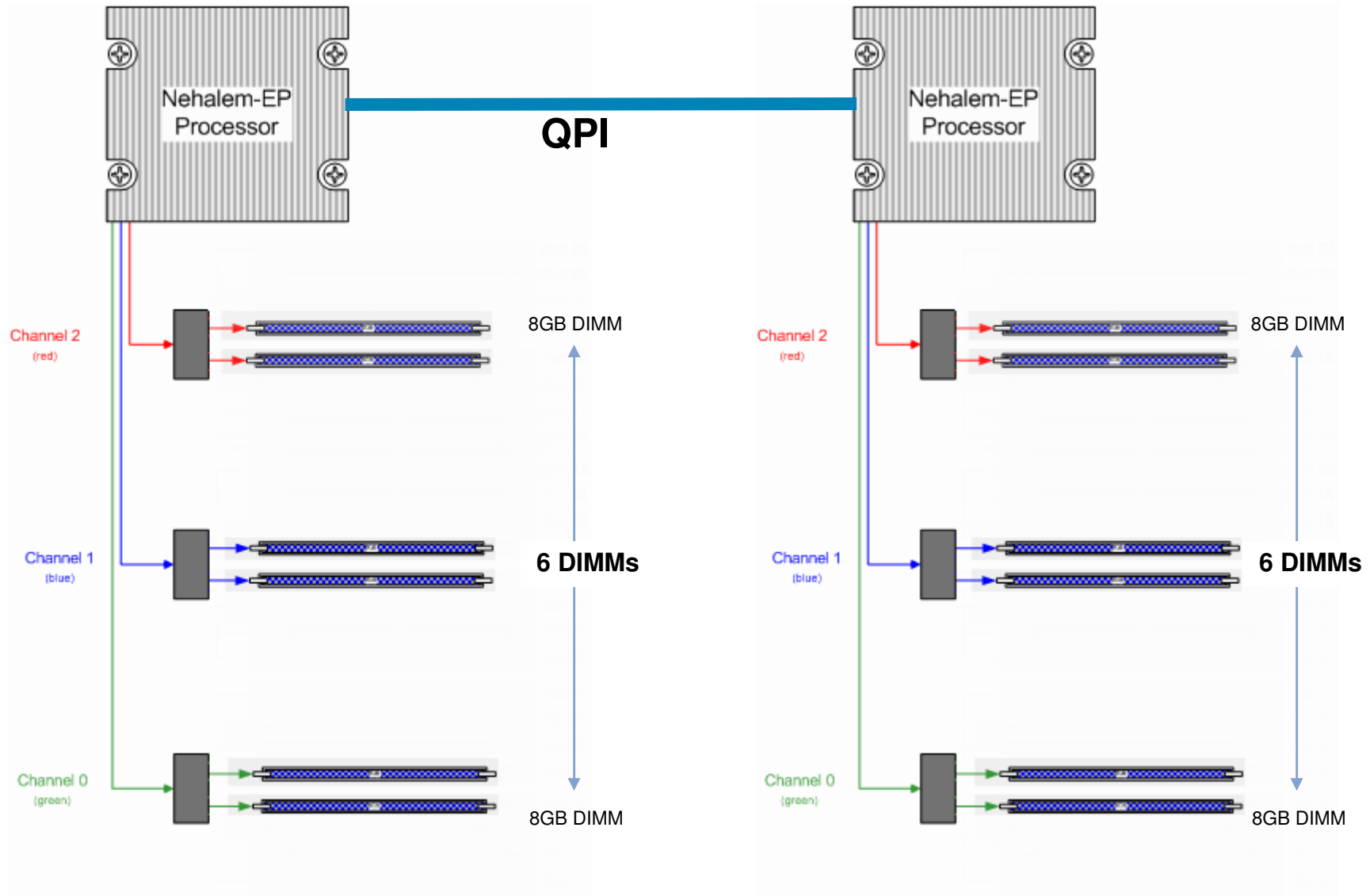




# Optimizing Memory with the Xeon 5500

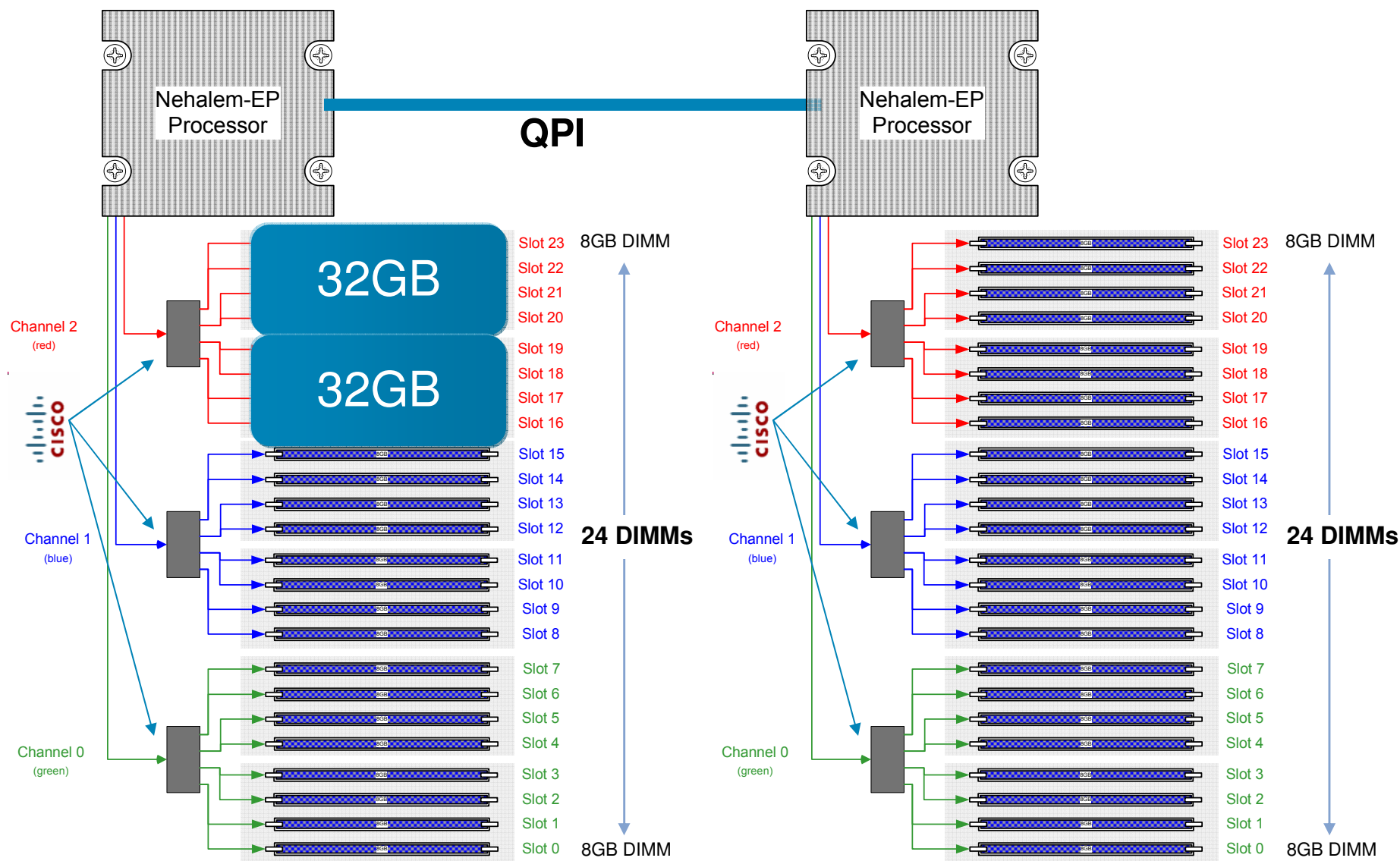


# Nehalem Memory Architecture



**96 GB max with 8 GB DIMMs**

# Expanded Memory Blade



**384 GB max with 8 GB DIMMs**

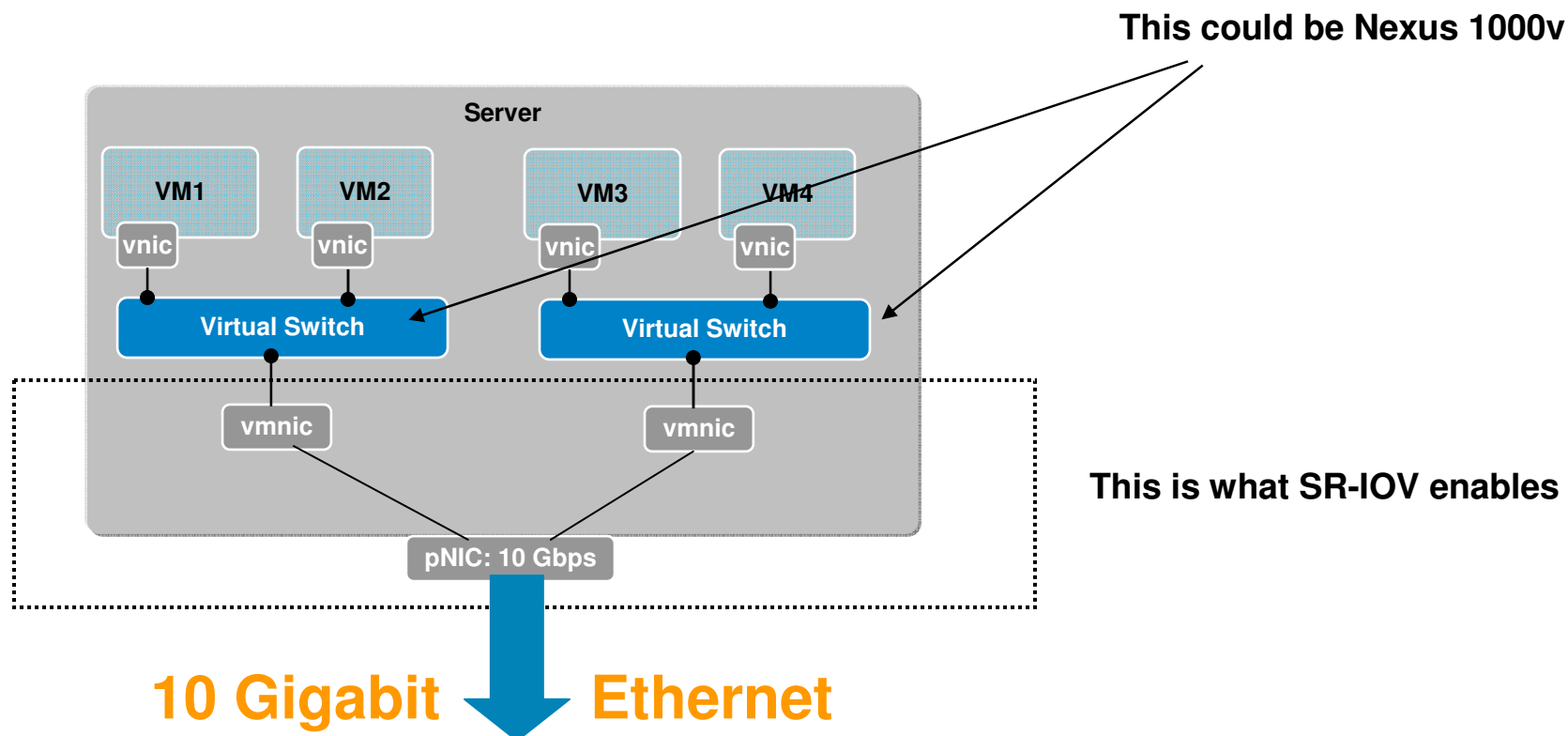


# Virtualized Adapter



# What is SR-IOV about?

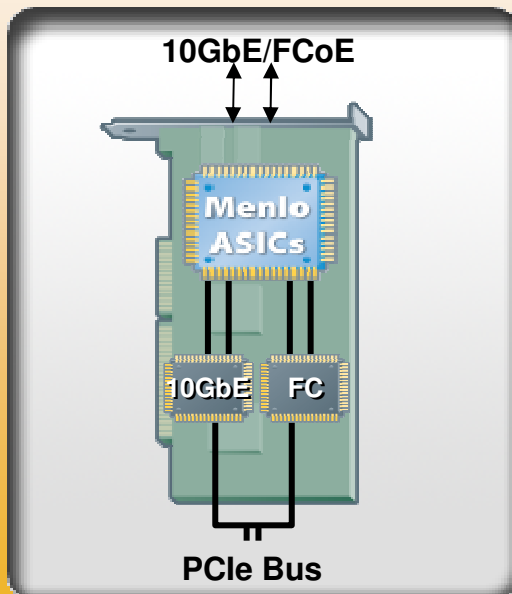
- Single Root IO Virtualization (SR-IOV) allows “virtualizing” the 10 GigE link (via the PCI-Express bus) into multiple “virtual links”.
- SR-IOV is a PCI-Sig standard
- In other words you can create multiple “vmnics” each with its own bandwidth allocation



# Adapter Offerings

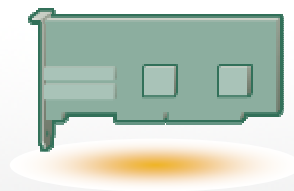
## Compatibility

Existing Driver Stacks



## Cost

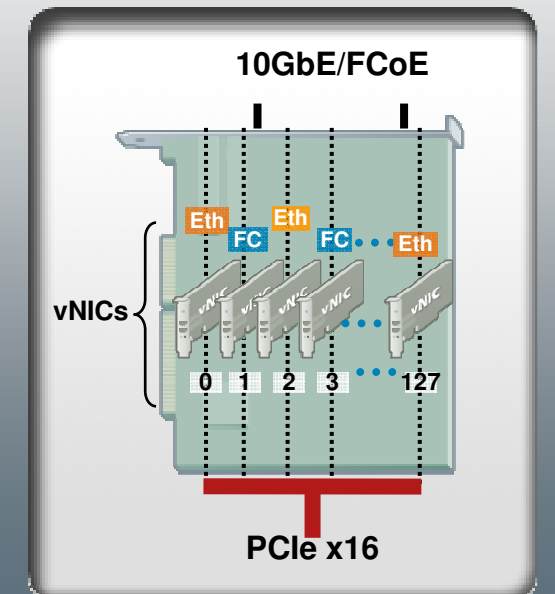
“Free” SAN Access  
for  
any Ethernet  
equipped host



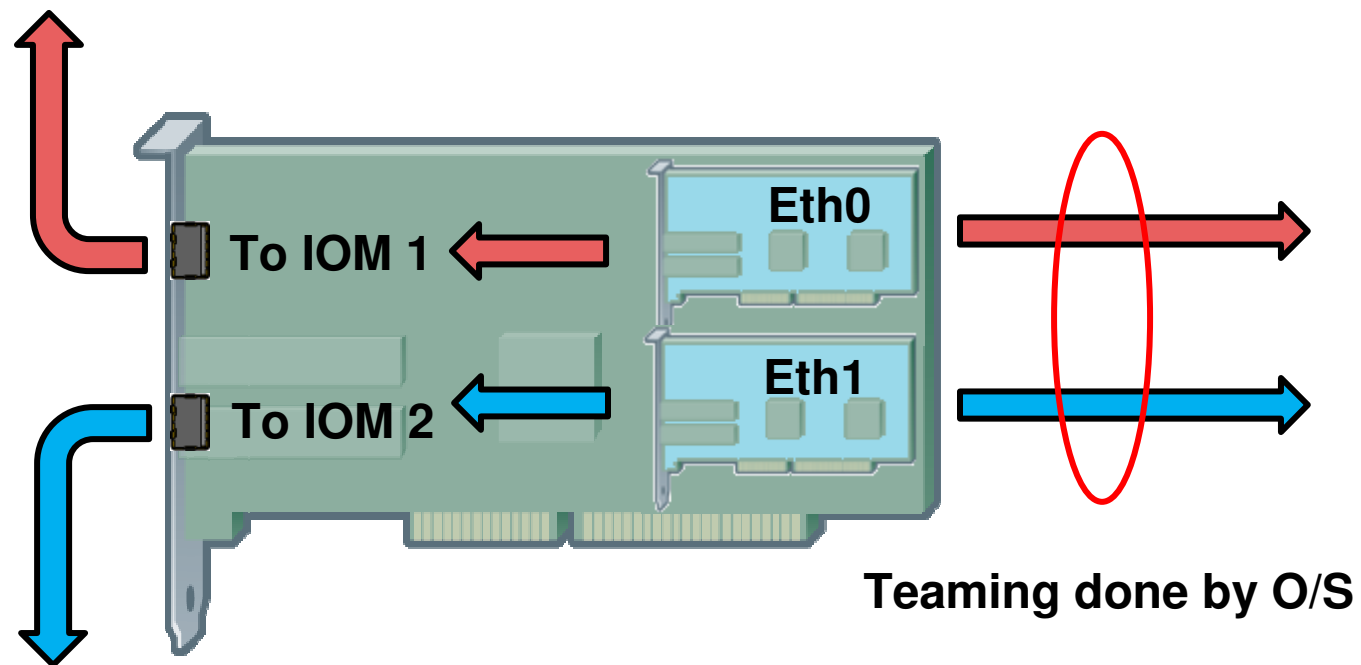
Software FCoE

## Virtualization

VM I/O Virtualization  
and Consolidation

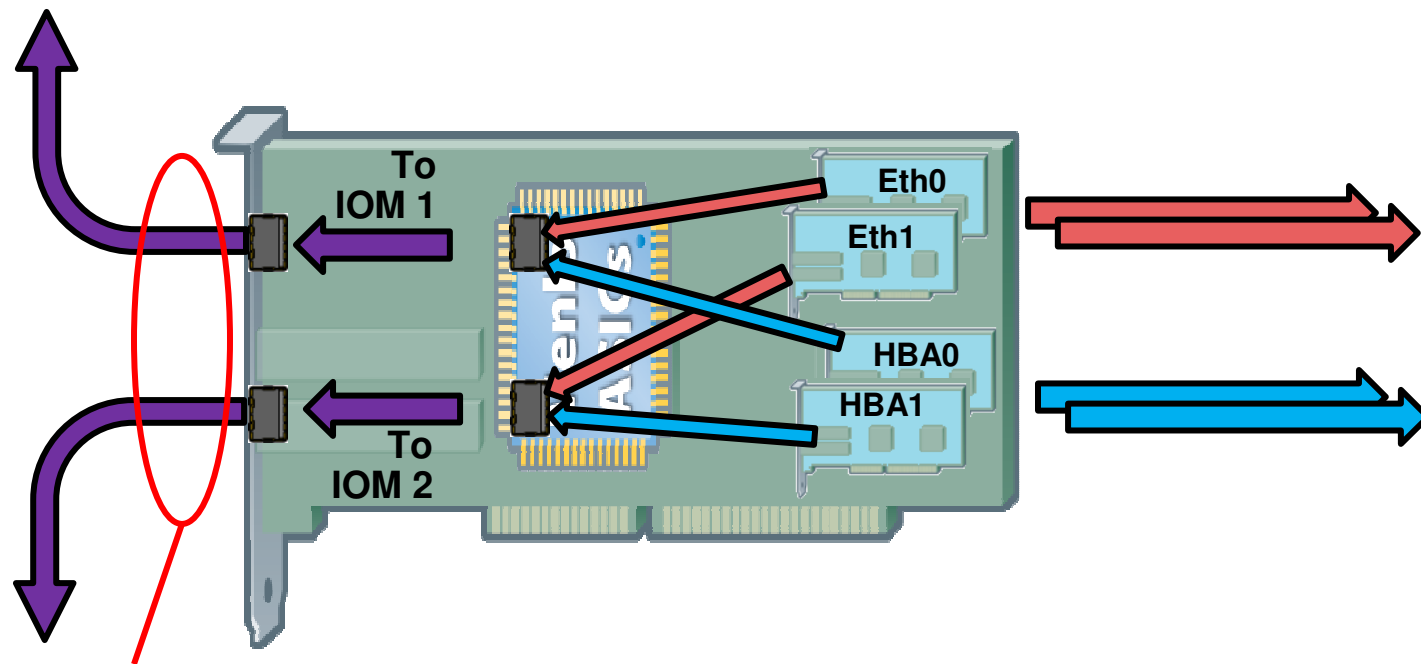


# Adapter Connectivity Details: Oplin



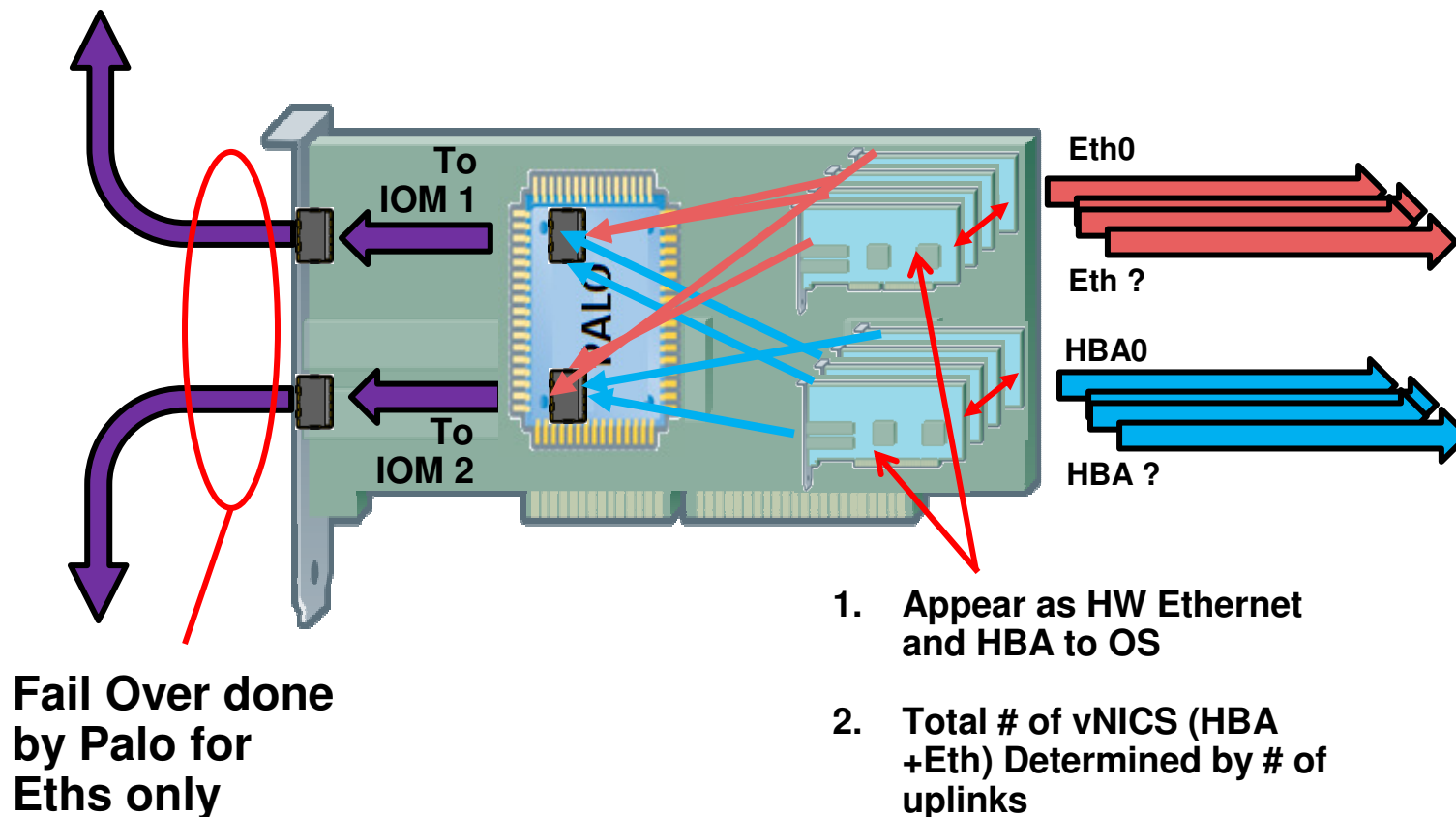


# Adapter Connectivity Details: CNA (Menlo)



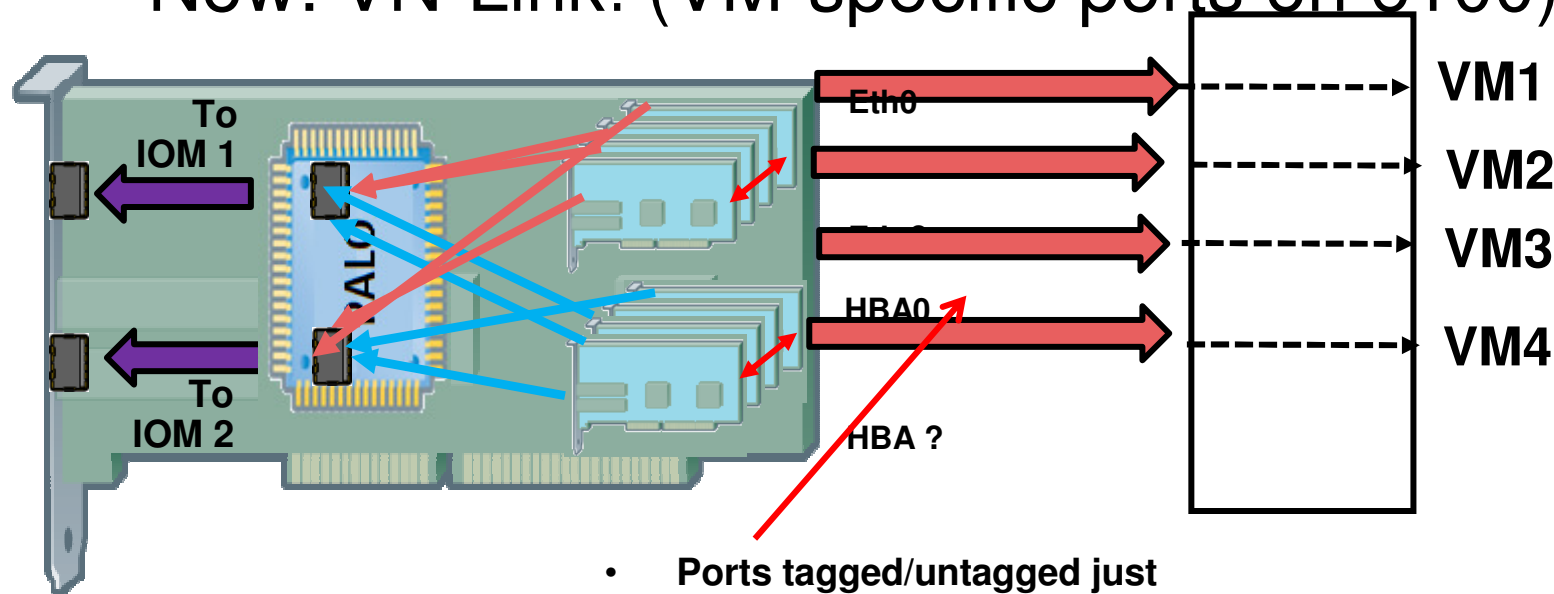
**Failover done  
by Menlo for  
Eths only**

# Adapter Connectivity Details: Palo



# What does it all have to do with Virtual Machines?

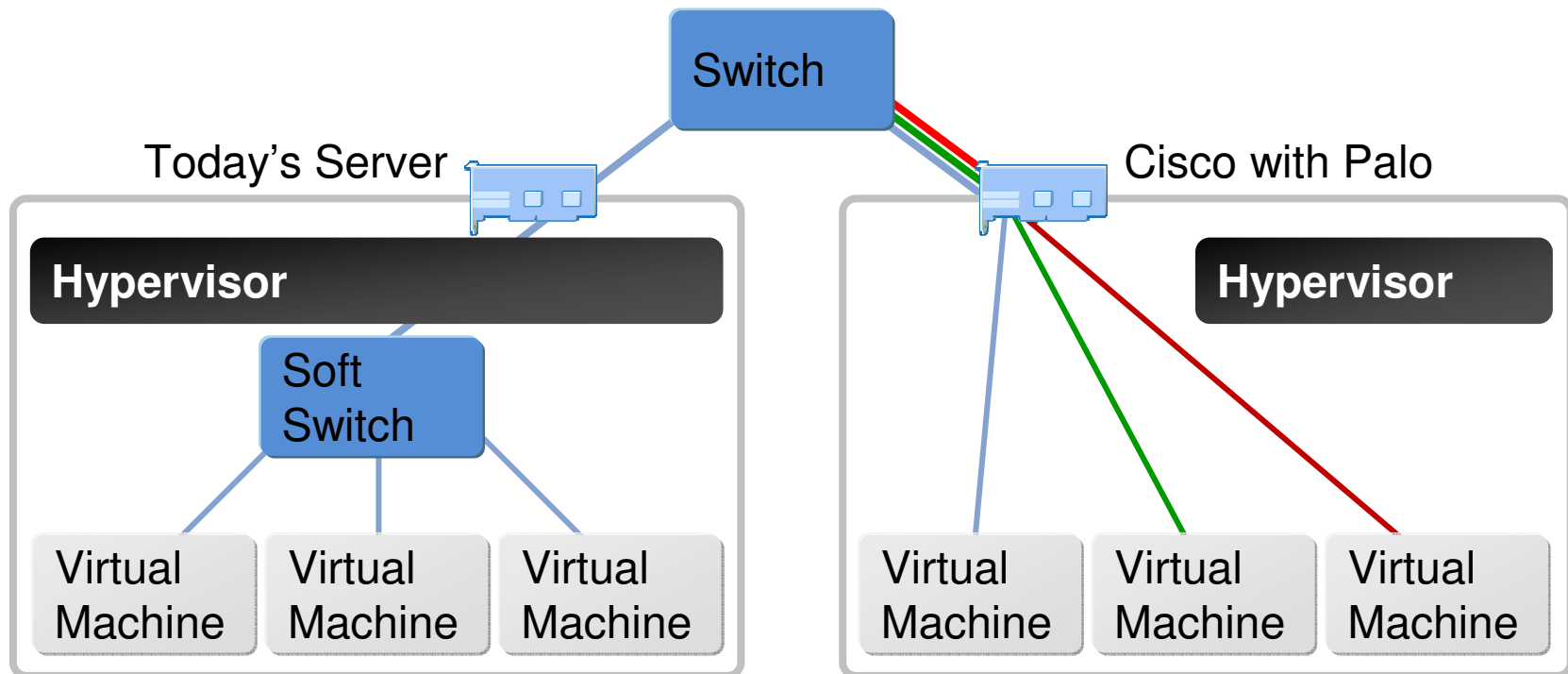
- Vntag part: nothing
- Now connect each Palo Port to Virt Machine
- Now: VN-Link! (VM-specific ports on 6100)



- Ports tagged/untagged just like fex ports
- Appear as virtual ports on top-level bridge (6100)

# Virtualization Adapter

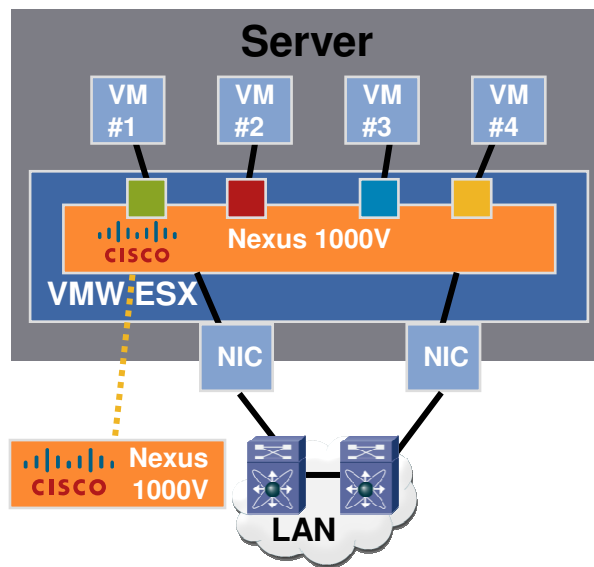
True wire once architecture – highly dynamic  
Network policy and visibility brought to VMs  
Hypervisor bypass support – increases performance  
Reduce NIC and mezz card infrastructure



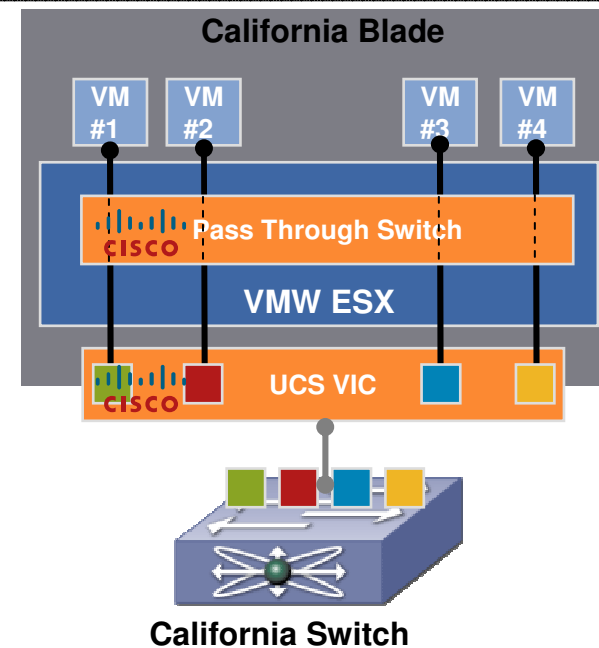
# VN-Link

## Complimentary Models for Evolving Requirements

### Cisco Nexus 1000V (Software Based)



### UCS VIC (Hardware Based)



Policy-Based  
VM Connectivity

Mobility of Network  
and Security Properties

Non-Disruptive  
Operational Model

# Cisco VN-Link

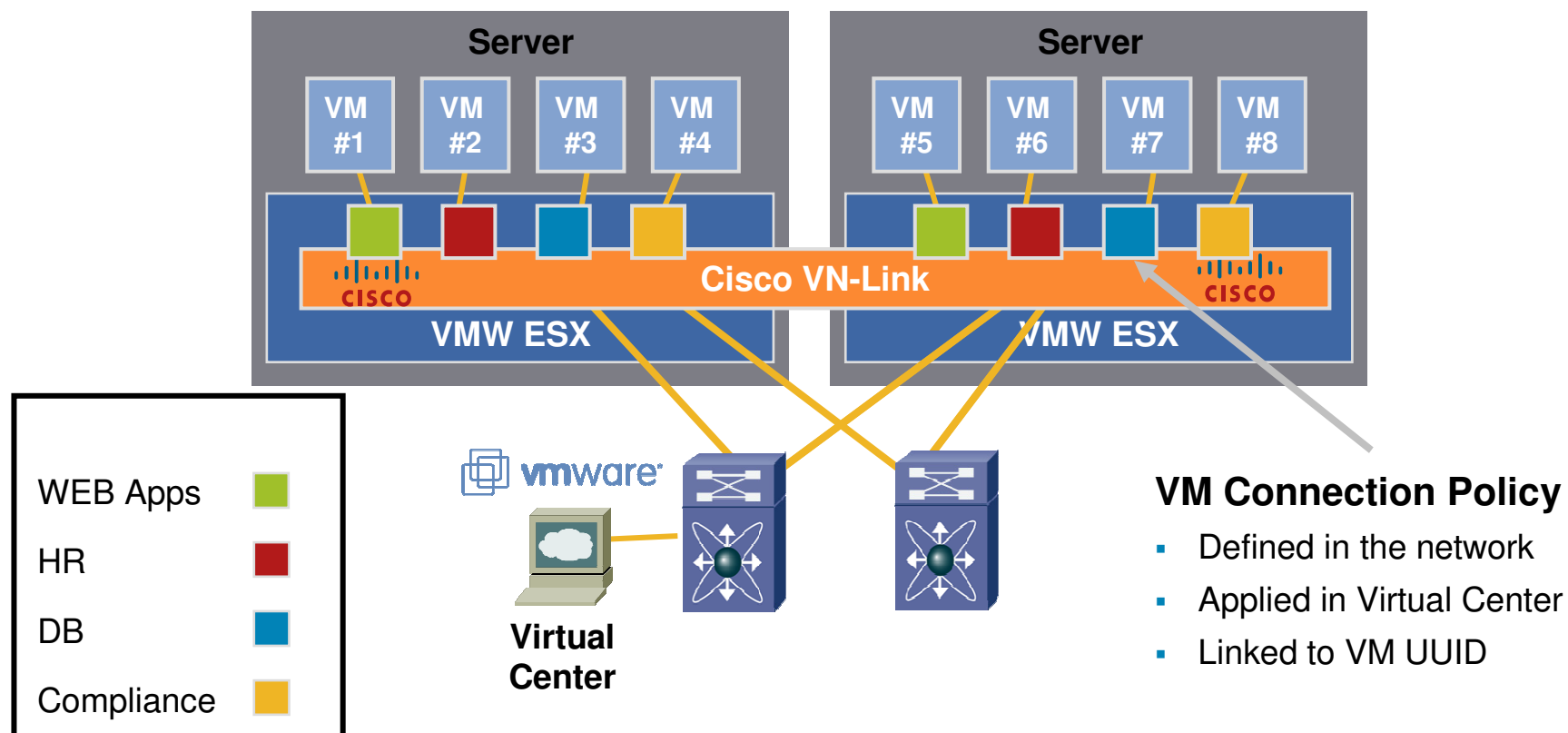
## Faster VM Deployment

### Cisco VN-Link—Virtual Network Link

**Policy-Based  
VM Connectivity**

**Mobility of Network  
& Security Properties**

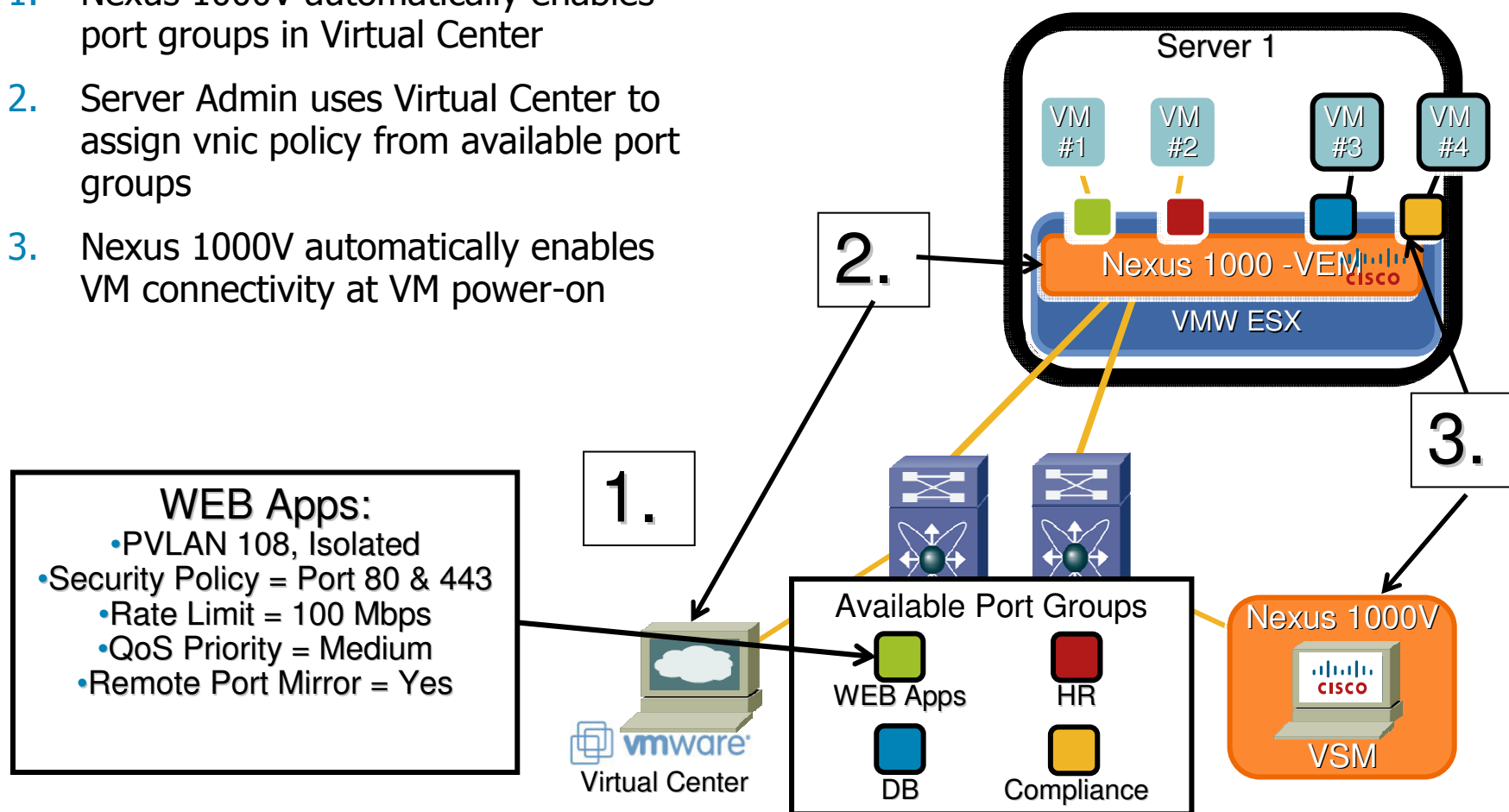
**Non-Disruptive  
Operational Model**



# Policy Based VM Connectivity

## *Enabling Policy*

1. Nexus 1000V automatically enables port groups in Virtual Center
2. Server Admin uses Virtual Center to assign vnic policy from available port groups
3. Nexus 1000V automatically enables VM connectivity at VM power-on



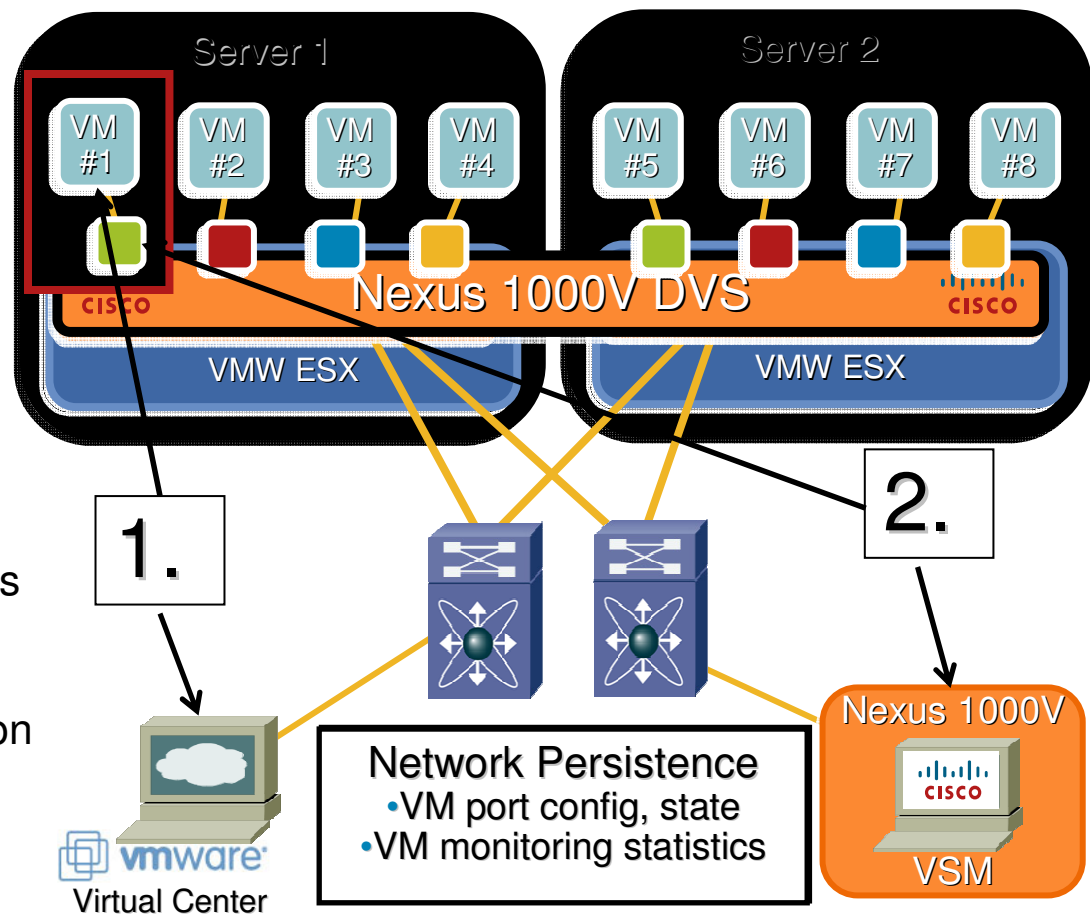
# Mobility of Security & Network Properties

*Following your VMs around*

1. Virtual Center kicks off a Vmotion (manual/DRS) & notifies Nexus 1000V
2. During VM replication, Nexus 1000V copies VM port state to new host

Mobile Properties Include:

- Port Policy
- Interface State & Counters
  - Flow Statistics
- Remote Port Mirror Session

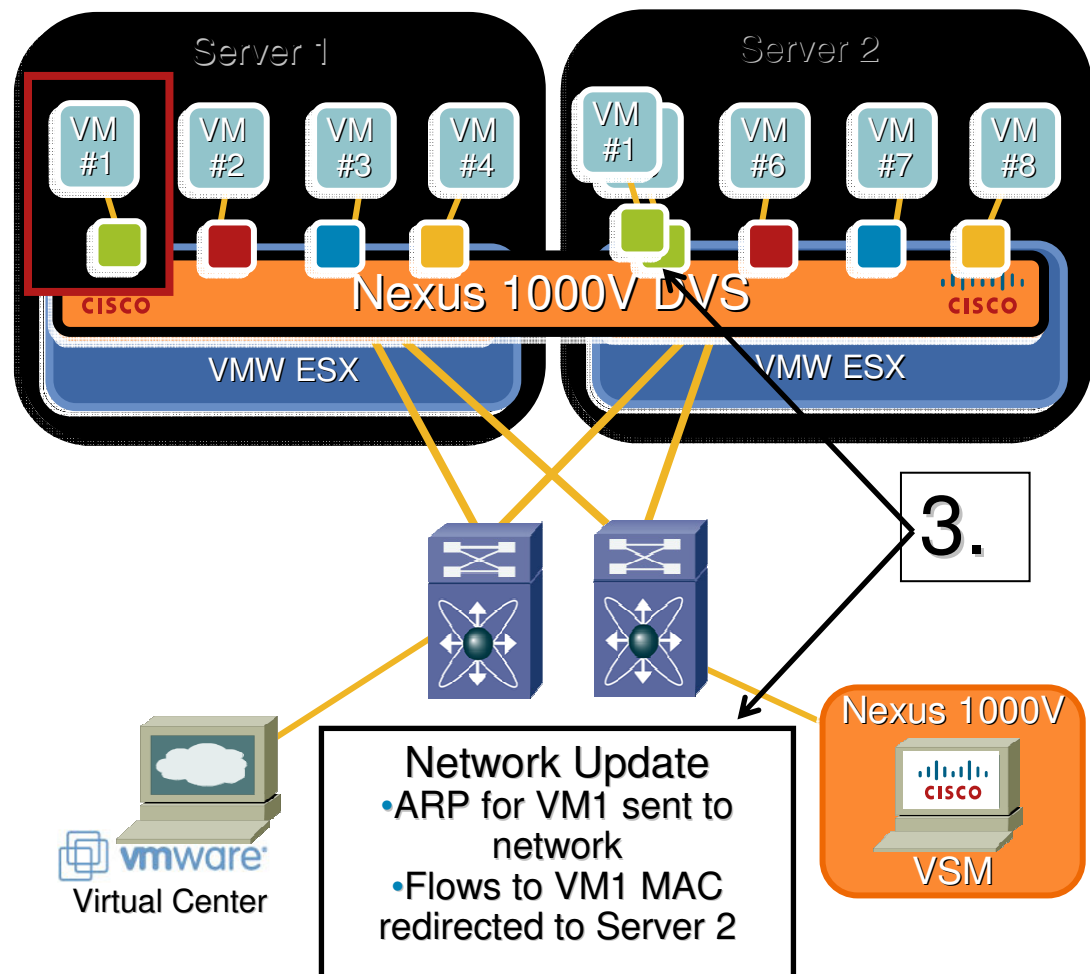




# Mobility of Security & Network Properties

*Following your VMs around*

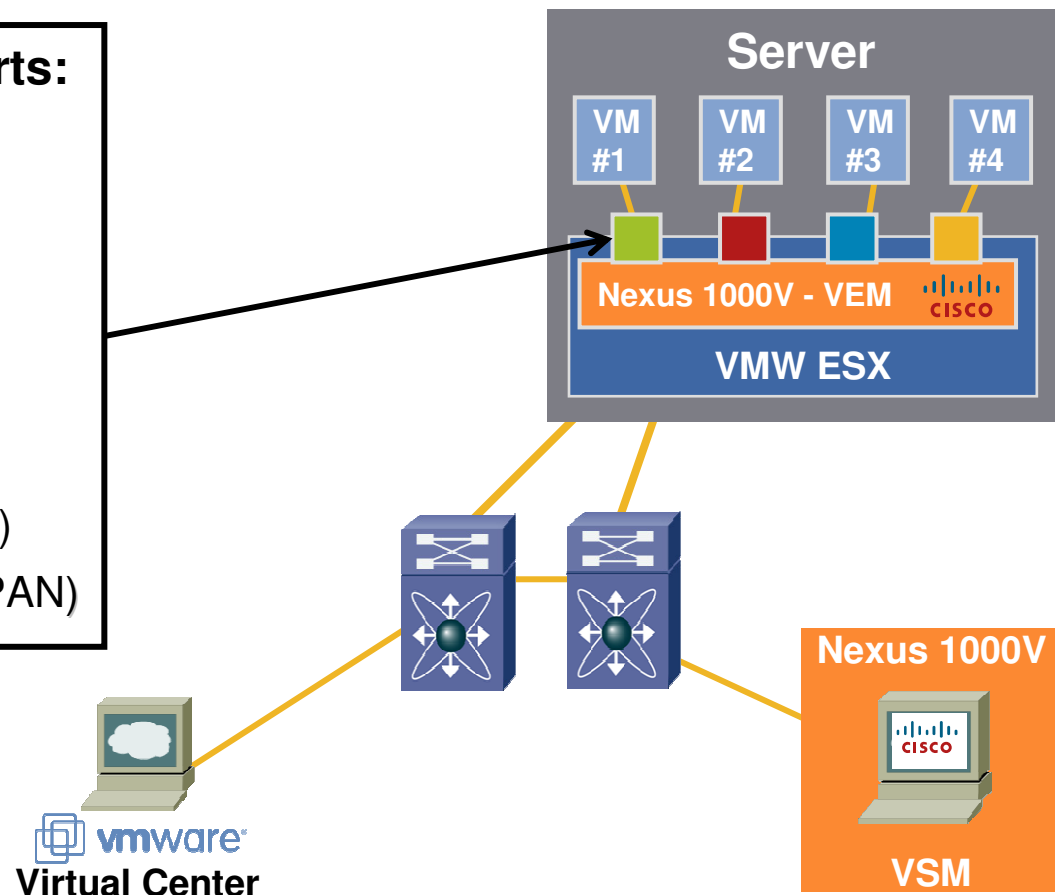
1. Virtual Center kicks off a Vmotion (manual/DRS) & notifies Nexus 1000V
2. During VM replication, Nexus 1000V copies VM port state to new host
3. Once VMotion completes, port on new ESX host is brought up & VM's MAC address is announced to the network



# What Can A Profile Contain?

## Policy definition supports:

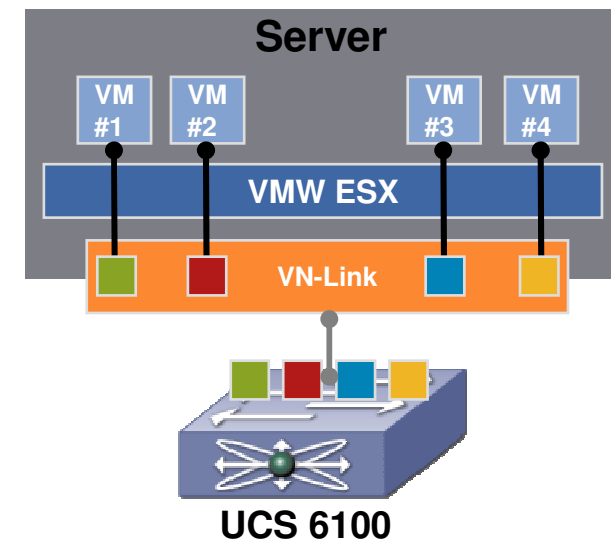
- VLAN, PVLAN settings
- ACL, Port Security, ACL Redirect
- Cisco TrustSec (SGT)
- NetFlow Collection
- Rate Limiting
- QoS Marking (COS/DSCP)
- Remote Port Mirror (ERSPAN)



# VN-Link with the UCS Fabric Interconnect

## Nexus Switch with VN-Link Hardware Based

- Allows scalable hardware-based implementations through hardware switches
- Standards-based initiative: Cisco & VMware proposal in IEEE 802 to specify “Network Interface Virtualization”
- Combines VM and physical network operations into one managed node

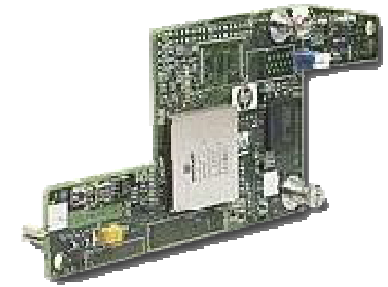


**Policy-Based  
VM Connectivity**

**Mobility of Network  
and Security Properties**

**Non-Disruptive  
Operational Model**

# California Palo Adapter



- Virtualized adapter designed for both single-OS and VM-based deployments
- Provides mobility, isolation, and management from the network

Secure

Transparent to hosts

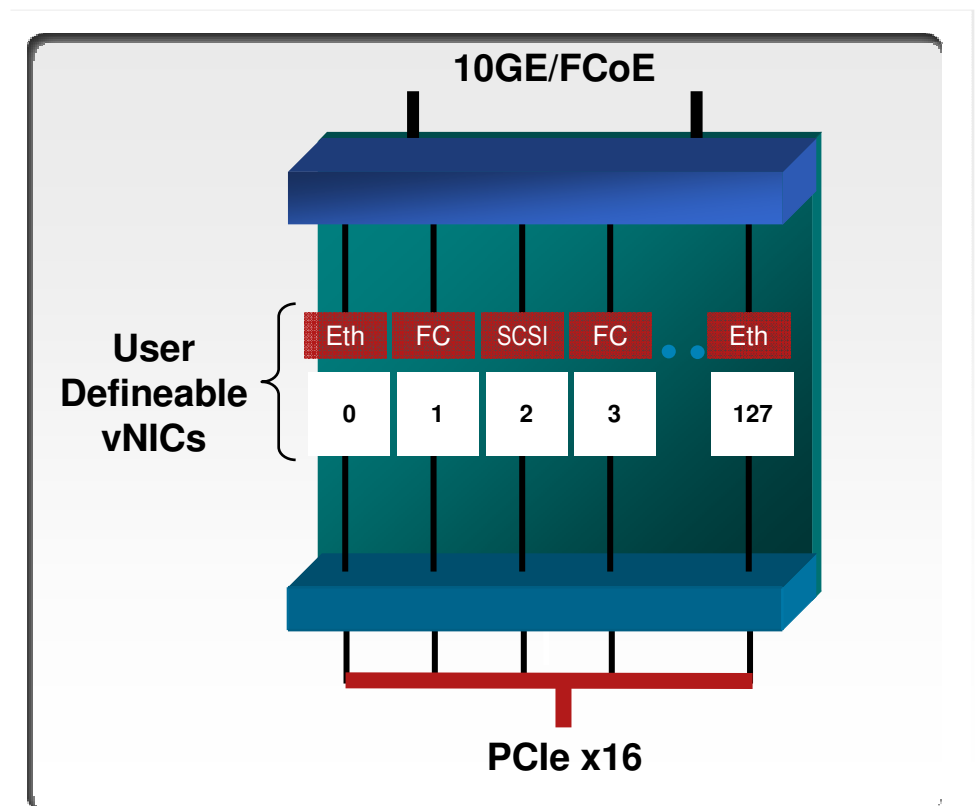
- Cut-through architecture
- High Performance

2x 10Gb

Low latency

High BW IPC support

- **128 vNICs**
  - Ethernet, FC or SCSI
  - 500K IOPS
  - Initiator and Target mode

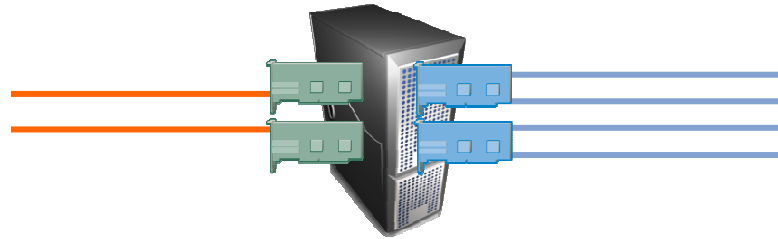




# Service Profiles



# Configuration Points



■ Server

# Service Profiles

Blade identities can be duplicated, automatically moved and deployed, and failed-over to another blade

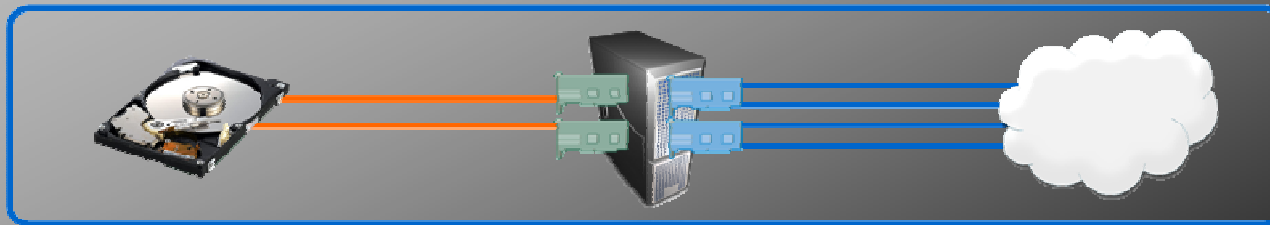
Firmware and bios included – competition does not do

“Stateless” environment

Significant process/labor savings

## Service Profile

- Encapsulation of HW state – MAC, WWN, Firmware, BIOS



# Service Profiles

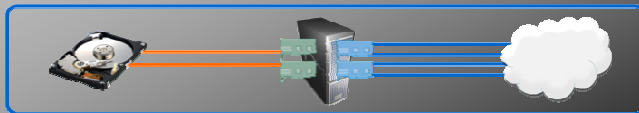
Blade identities can be duplicated, automatically moved and deployed, and failed-over to another blade

Firmware and bios included – competition does not do

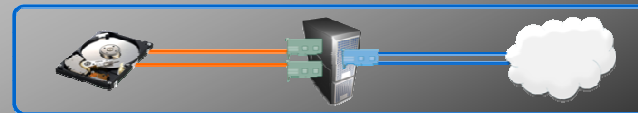
“Stateless” environment

Significant process/labor savings

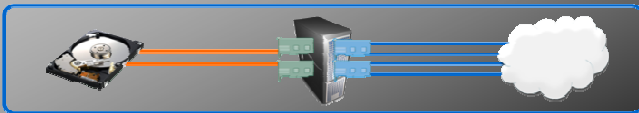
Service Profile



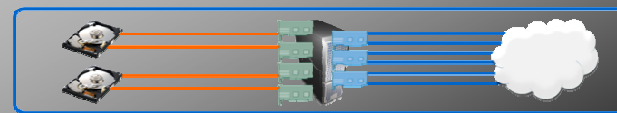
Service Profile



Service Profile



Service Profile





# Integrated Stateless Computing

- Attributes no longer tied to physical hardware

Not just identity

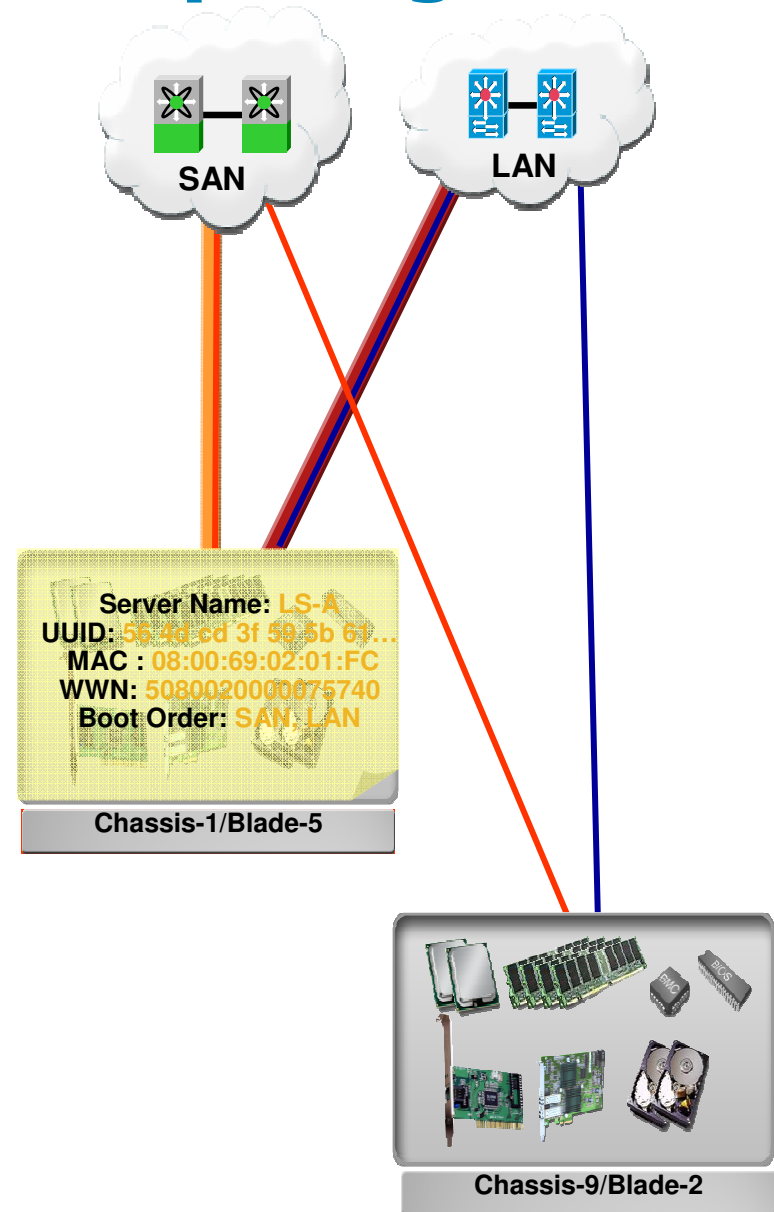
Seamless server mobility

Within interconnect domain

- Dynamic Provisioning

Complete infrastructure repurposing

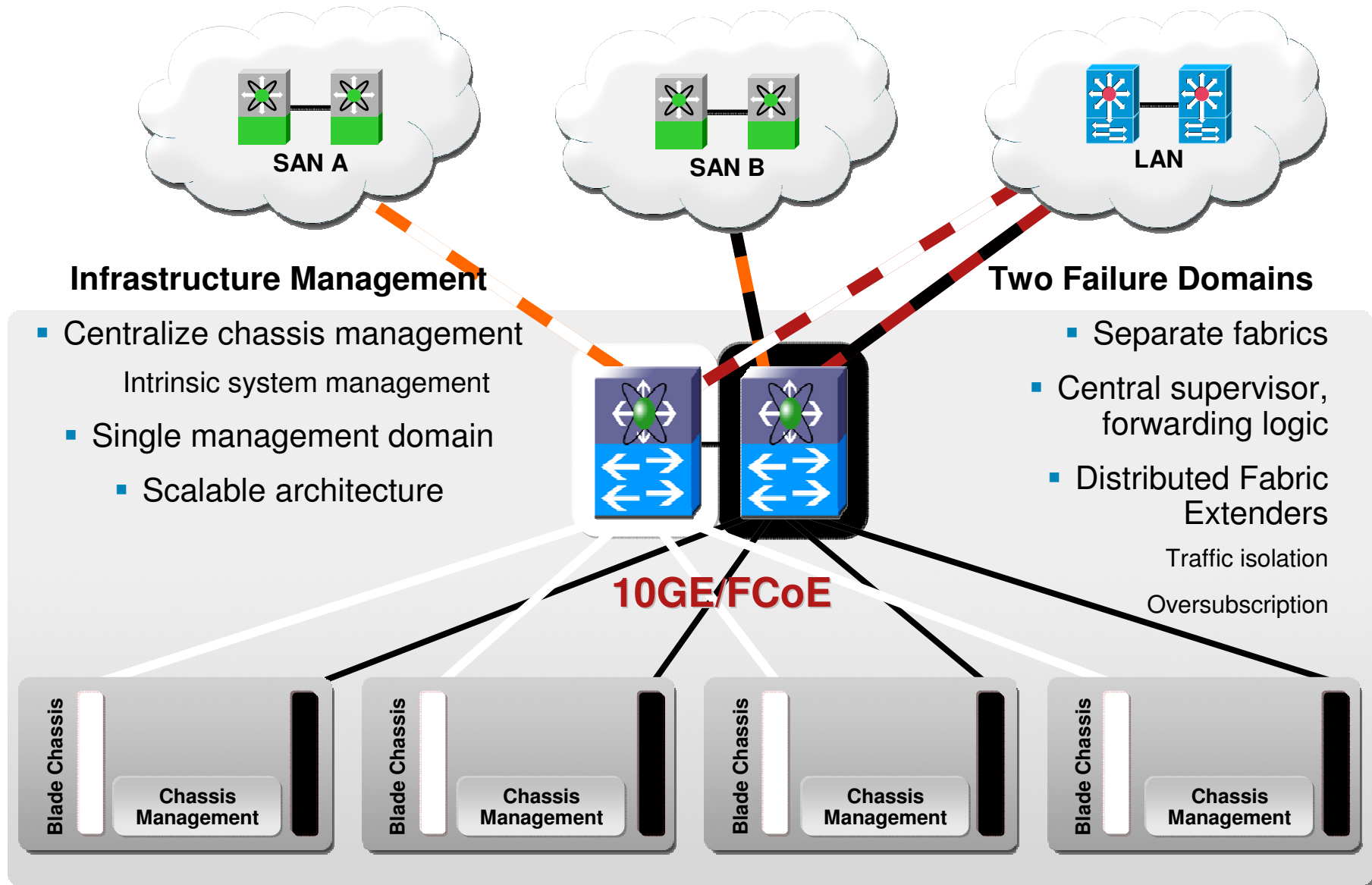
Integrated with 3<sup>rd</sup> part tools



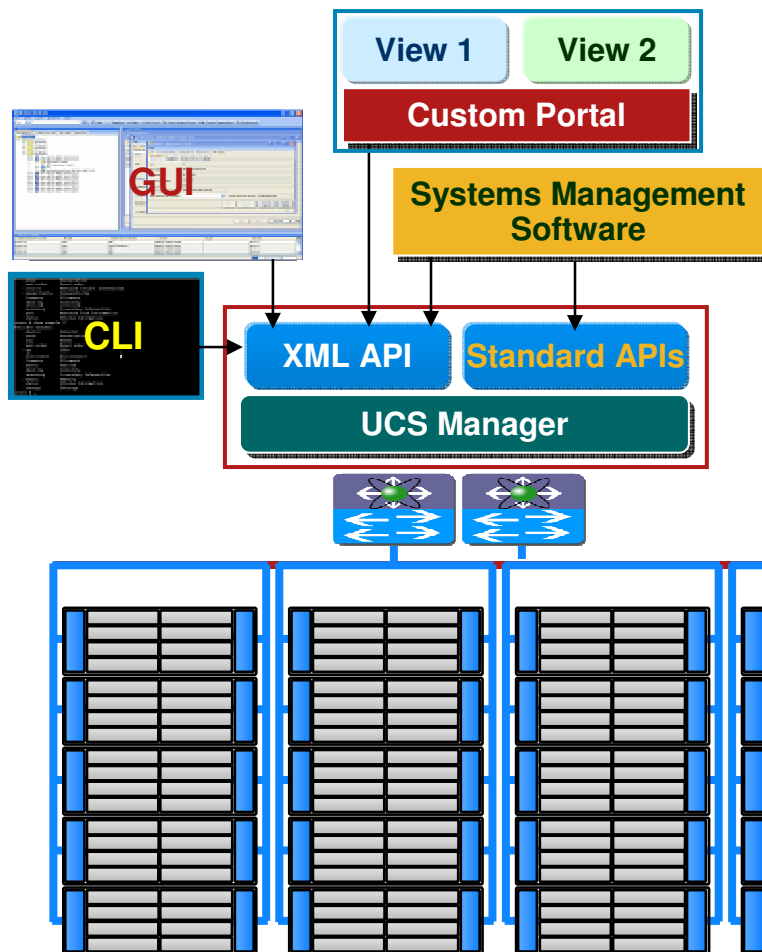
# Identity and Server Pools

- Server Pools
- UUID Pools
- WW Node Name Pools
- WW Port Name Pools
- Mac Pools

# Unified Management



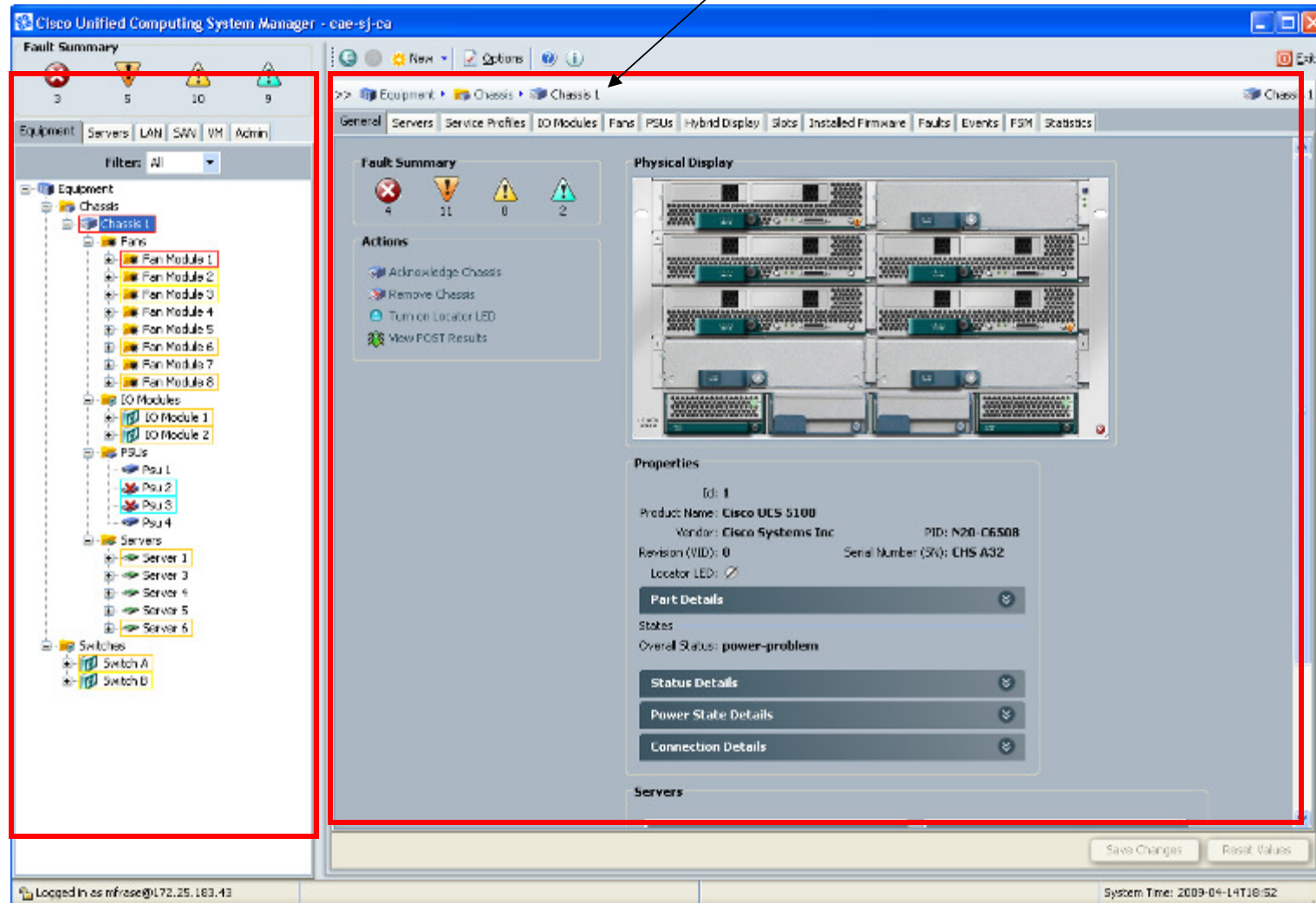
# Unified Management



- Single point of device management
  - Adapters, blades, chassis, LAN & SAN connectivity
  - Embedded manager
  - GUI & CLI
- Standard APIs for systems management
  - XML, SMASH-CLP, WSMAN, IPMI, SNMP
  - SDK for commercial & custom implementations
- Designed for multi-tenancy
  - RBAC, organizations, pools & policies

# UCS Graphical interface

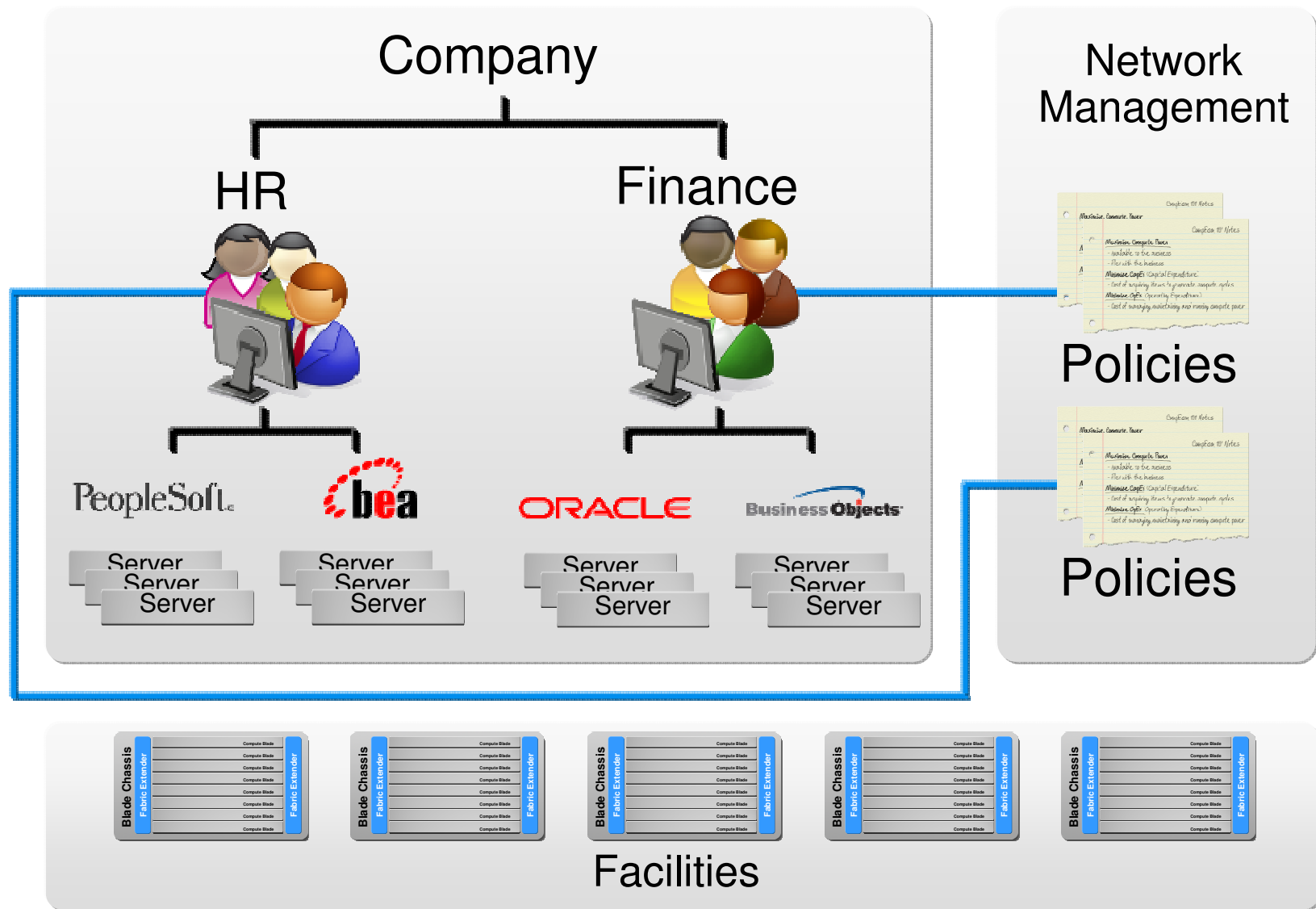
Top directory map tells you where you are in tree



NAVIGATION PANE

CONTENT PANE

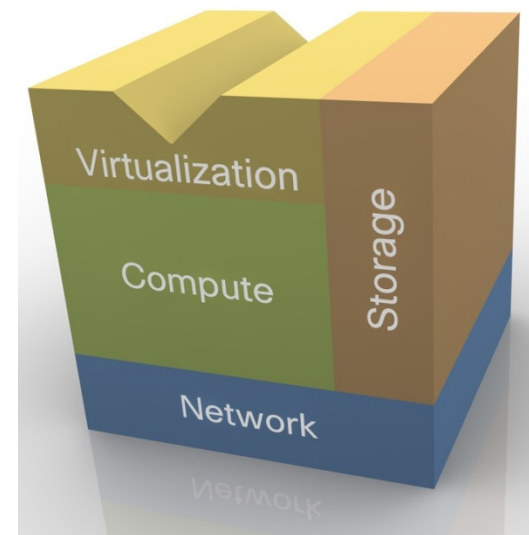
# Multi-Tenancy Model





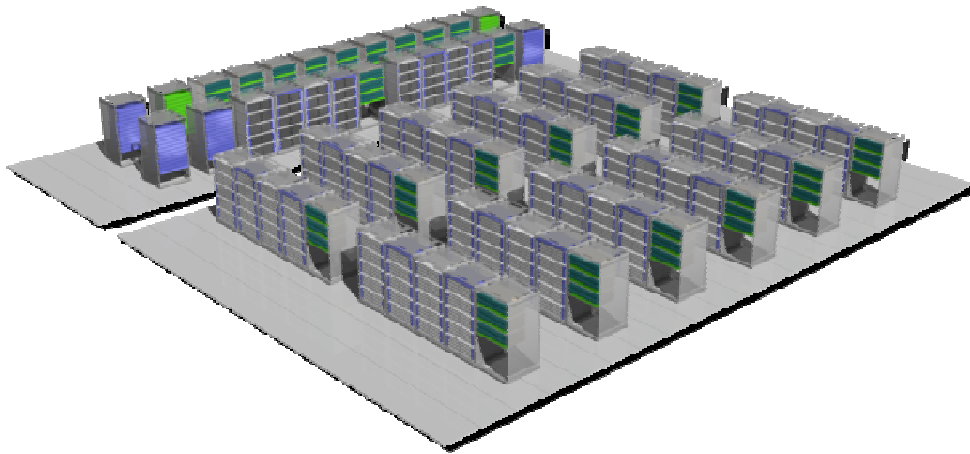
Efficiency. Control. Choice.

## Vblock Infrastructure Packages: Accelerating Deployment of the Private Cloud





# Vblock Design principles



**A Data Center is a collection of pooled 'Vblocks' aggregated in 'Zones'.**

- *A unit of assembly* that provides a set of services, at a known level, to target consumers
- Self contained, but it may also use external shared services
- Optimized for the classes of services it is designed to provide
- Can be clustered to provide availability - or aggregated for scalability, but each Vblock is still viable on its own
- Fault and service isolation - the failure of a Vblock will not impact the operation of other Vblocks (Service Level degradation may occur unless availability or continuity services are present)



# Vblock Infrastructure Packages

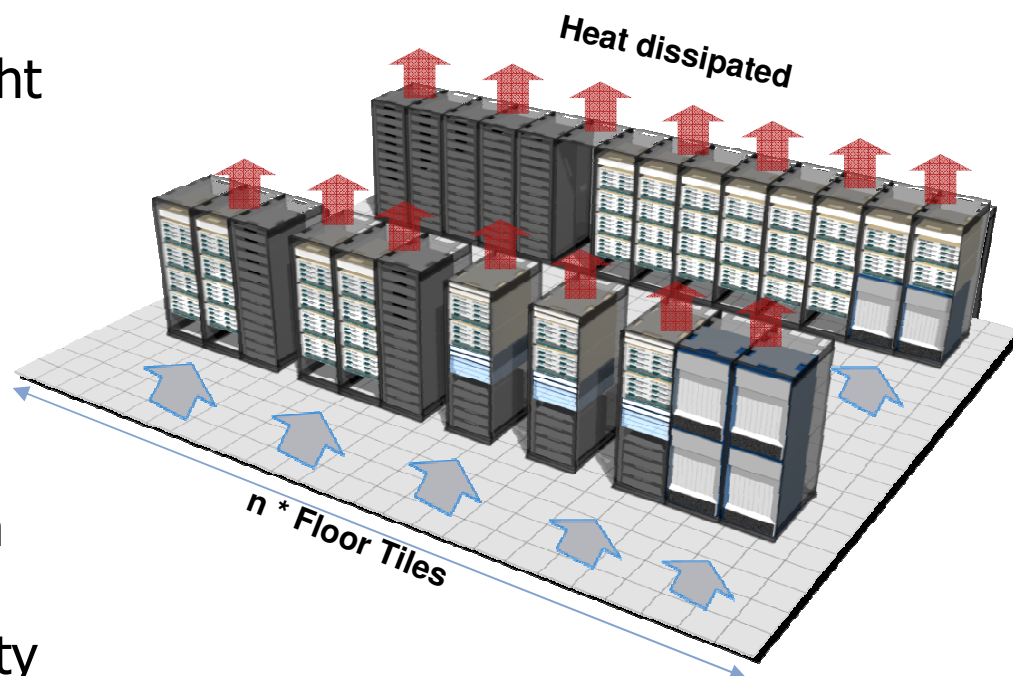
## Scalable Platform for Building Solutions

- Vblock 2 (3000 – 6000+ VMs)
  - A high-end configuration - extensible to meet the most demanding IT needs
  - Typical use case:
    - Business critical ERP, CRM systems
- Vblock 1 (800 – 3000+ VMs)
  - A mid-sized configuration - broad range of IT capabilities for organizations of all sizes
  - Typical use case:
    - Shared services – Email, File and Print, Virtual Desktops, etc.
- Vblock 0 (300 – 800+ VMs) ~1H 2010
  - An entry-level configuration addresses small datacenters or organizations
  - Test/development platform for Partners and customers



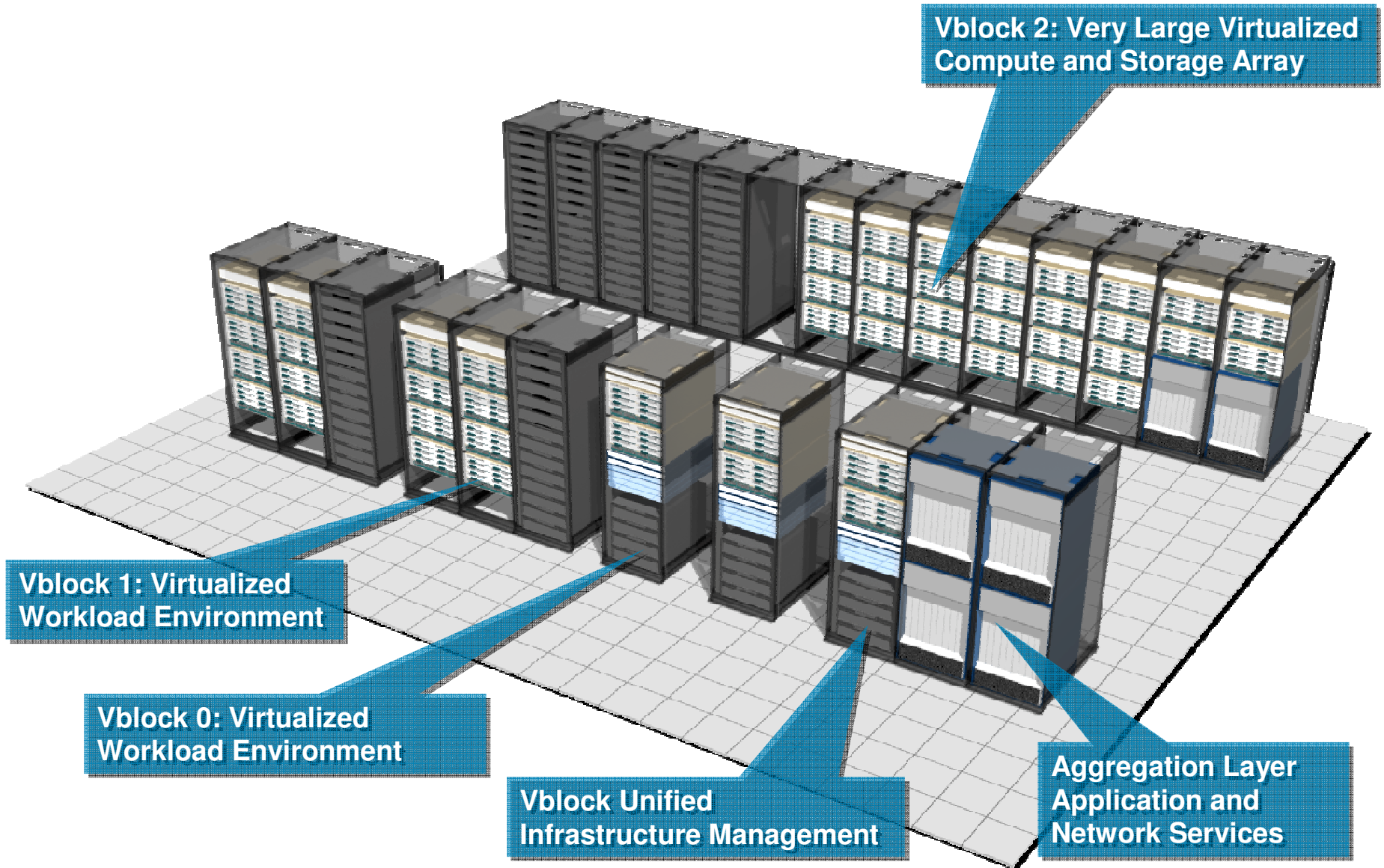
# Deterministic Performance, Predictable Architecture

- Predictable SLA:  
Granular SLA measurement and assurance
- Deterministic space and weight  
floor tiles become unit of capacity planning
- Power and Cooling:  
consistent power and cooling (KWh/BTUs) per unit
- Pre-determined capacity and scalability:  
Uniform workload distribution and mobility
- Deterministic fault and security isolation



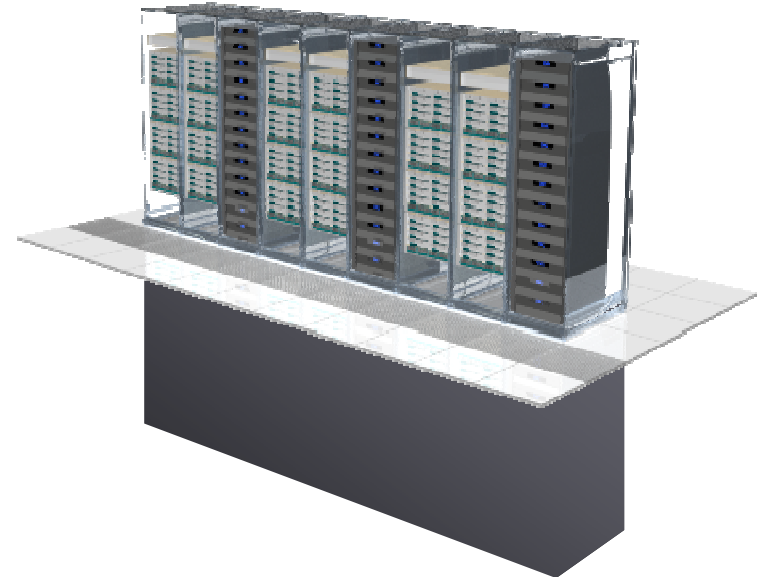
# Vblock Infrastructure Packages

## Scalable IT capability and performance



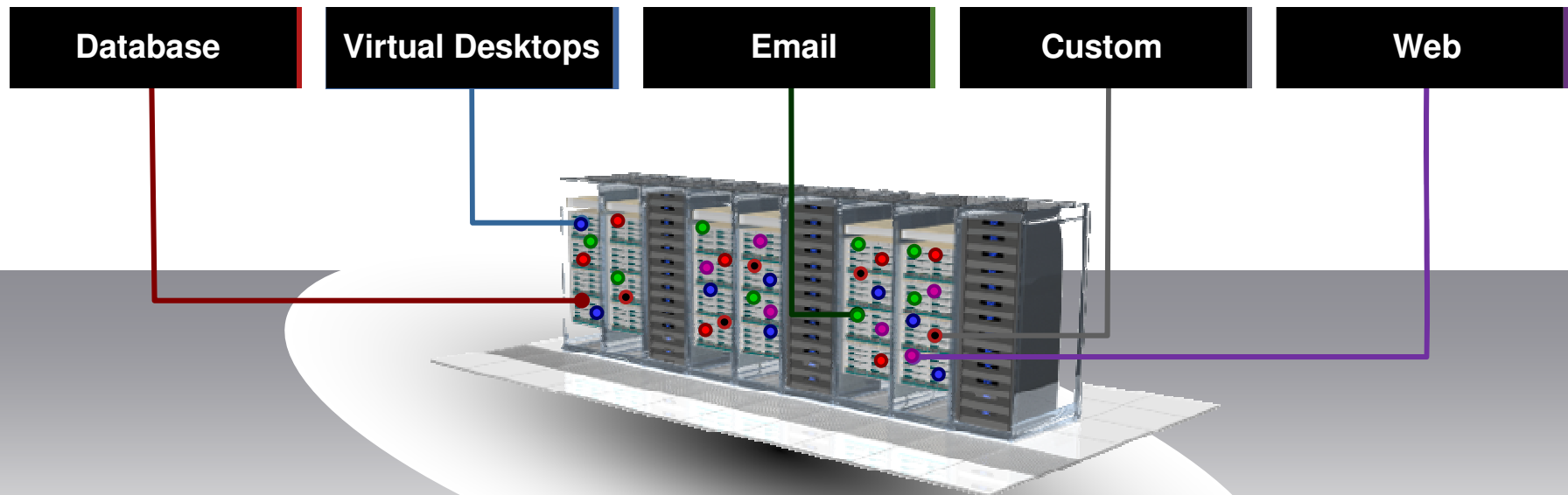
# Vblock 1 Components

- Compute
  - Cisco UCS B-series
- Network
  - Cisco Nexus 1000V
  - Cisco MDS 9506
- Storage
  - EMC CLARiiON CX4
- Hypervisor
  - VMware vSphere 4
- Management
  - EMC Ionix Unified Infrastructure Manager
  - VMware vCenter
  - EMC NaviSphere
  - EMC PowerPath
  - Cisco UCS Manager
  - Cisco Fabric Manager



# Accelerating Virtualization...

Accelerate IT Standardization and Simplification



**Enable Virtualization at Scale—Simplify IT**

