NEW

Cisco Nexus 7000

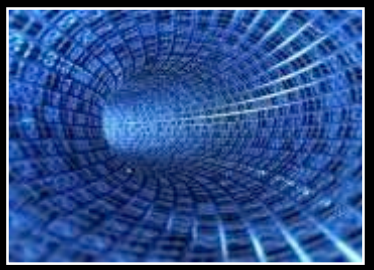
Therdtoon Theerasasana

ttheera@cisco.com

What Is Data Center Class?

Continuous Operations

Virtualized Resources



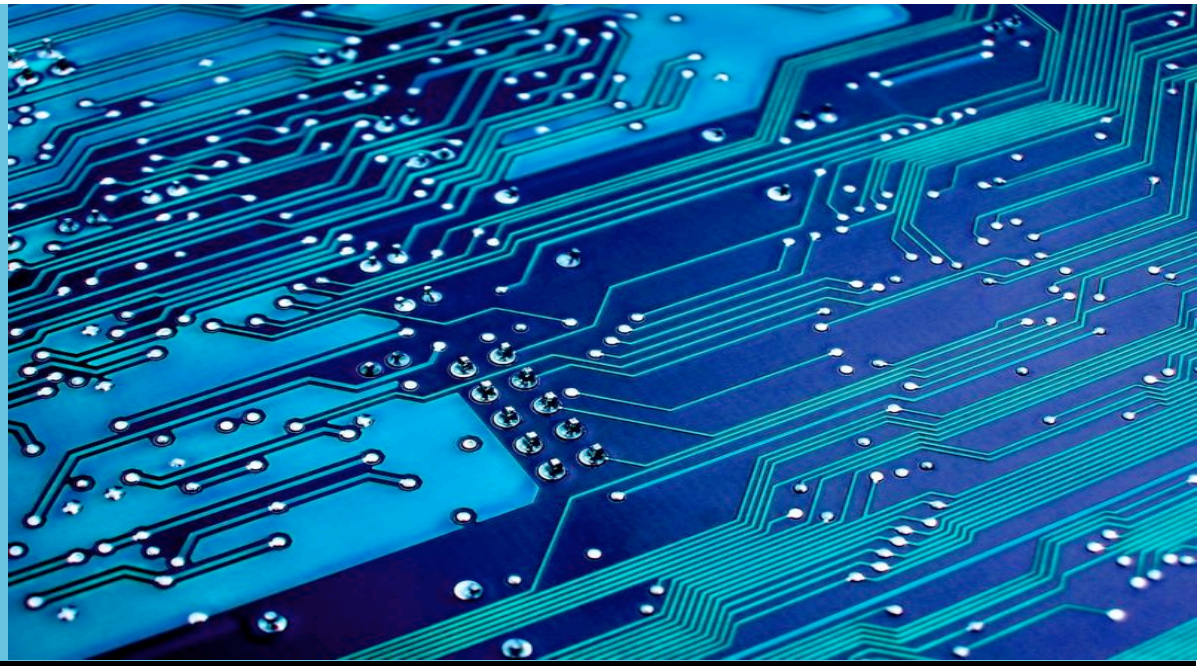
Operational Efficiency

Power and Cooling Efficiency

Cisco Nexus 7000 Series Switches

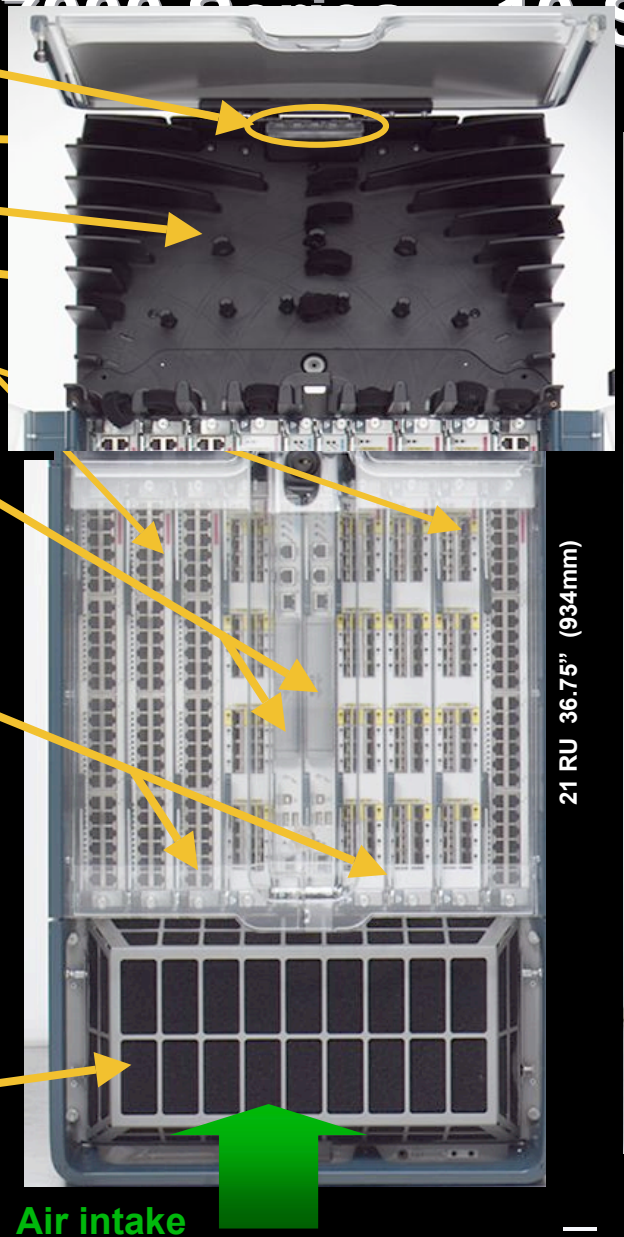
- Data-Center Class Network Platform
  - New family of data-center switching devices
- Hardware:
  - Scalable to over 15+Tbps
  - 10/100/1000 and 10GbE I/O modules at release
  - Optimized for 10GbE and future 40G/100G
  - Unified fabric capable (Ethernet, Future - FCoE)
- Software:
  - Modern, modular, scaleable operating system
  - Highly available and resilient
  - Virtualizes control plane and data plane
  - Stateful process restarts and supervisor switchovers
- Management:
  - New data-center optimized management platform
  - Open APIs and published SDK

Cisco Nexus 7000 Series Hardware



Chassis & Power Supply
Supervisor Engine
Crossbar Switch fabrics
Virtual Output Queuing
Central Arbitration

Nexus 7000 Series 10 Slot Switch



Front and Rear of 10 Slot Chassis

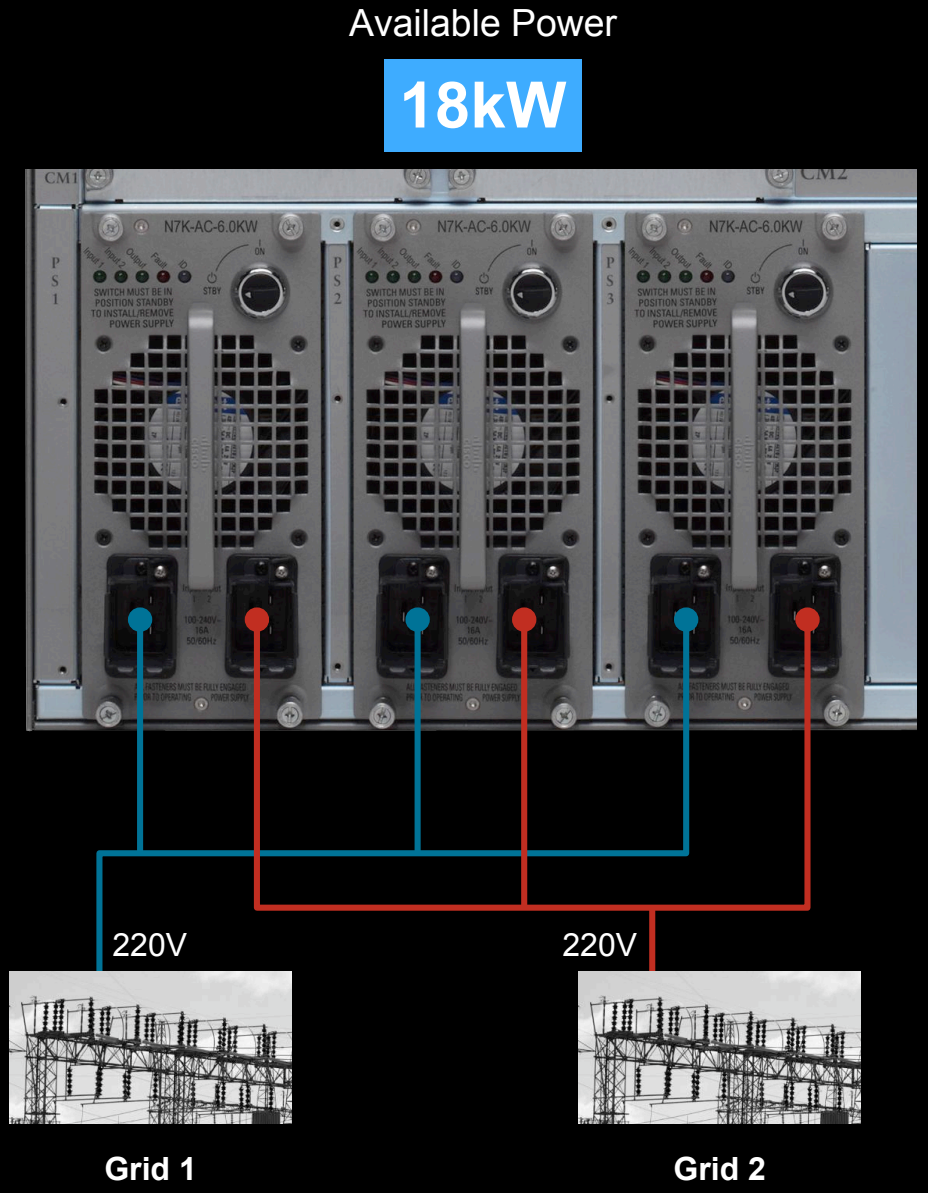
System Power

- 6000W power supply for Nexus 7000 Series chassis
- Supports dual inputs at 220/240V or 110/120V
- Hot swappable
- Blue beacon LED for easy location
- Power redundancy modes:
  - Full redundancy
  - N+1 redundancy (Default)
  - Grid/input source redundancy

AC Input		Full	N+1	Grid	Combined
1 input	110V 220V	n/a n/a	2400 6000	n/a n/a	3600 9000
2 inputs	110V 220V	3600 9000	4800 12000	3600 9000	7200 18000
2 inputs	110V + 220V	3600	8400	3600	12600

Power Redundancy

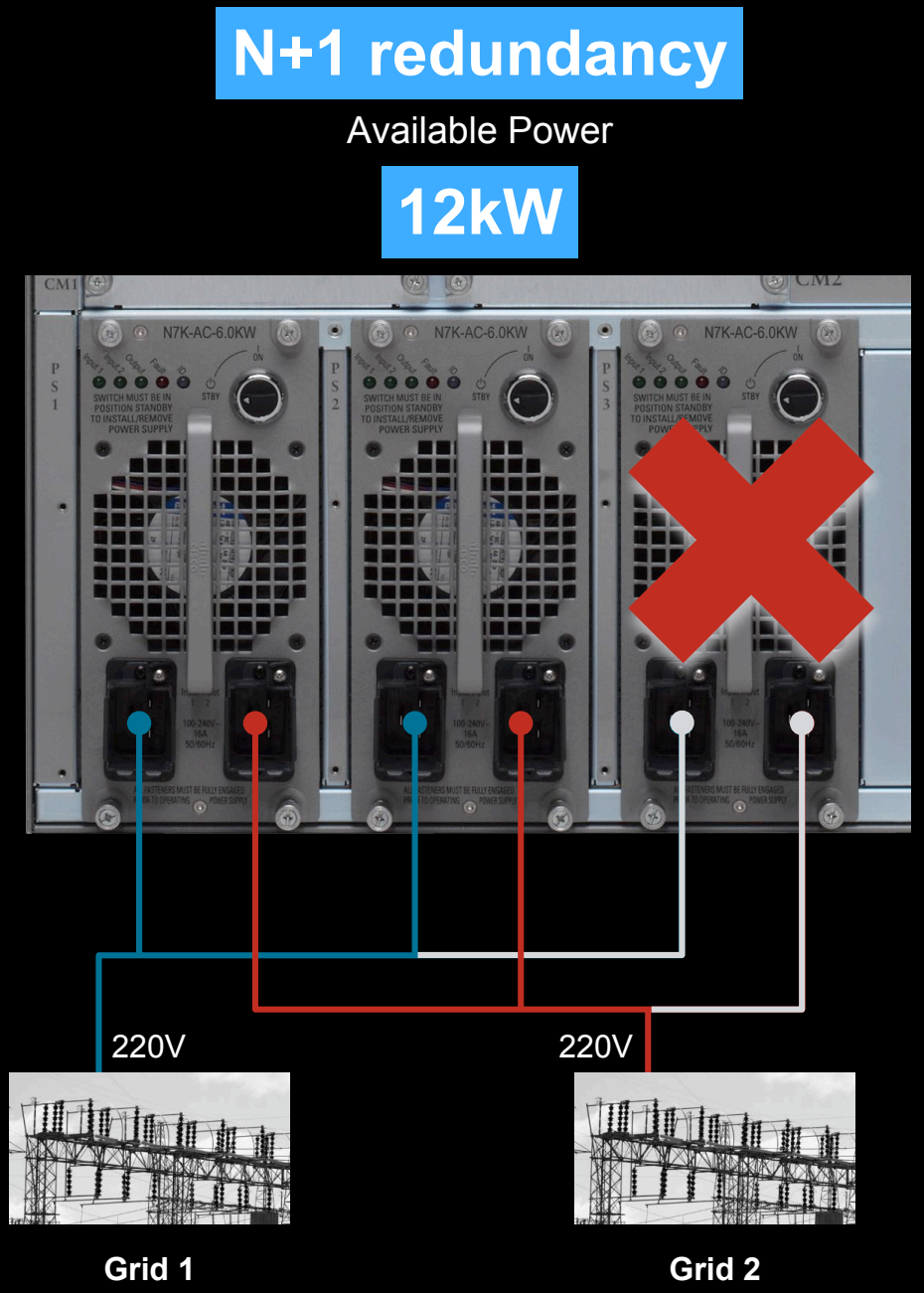
Power redundancy mode dictates how system budgets power



Power Redundancy

Power redundancy mode dictates how system budgets power

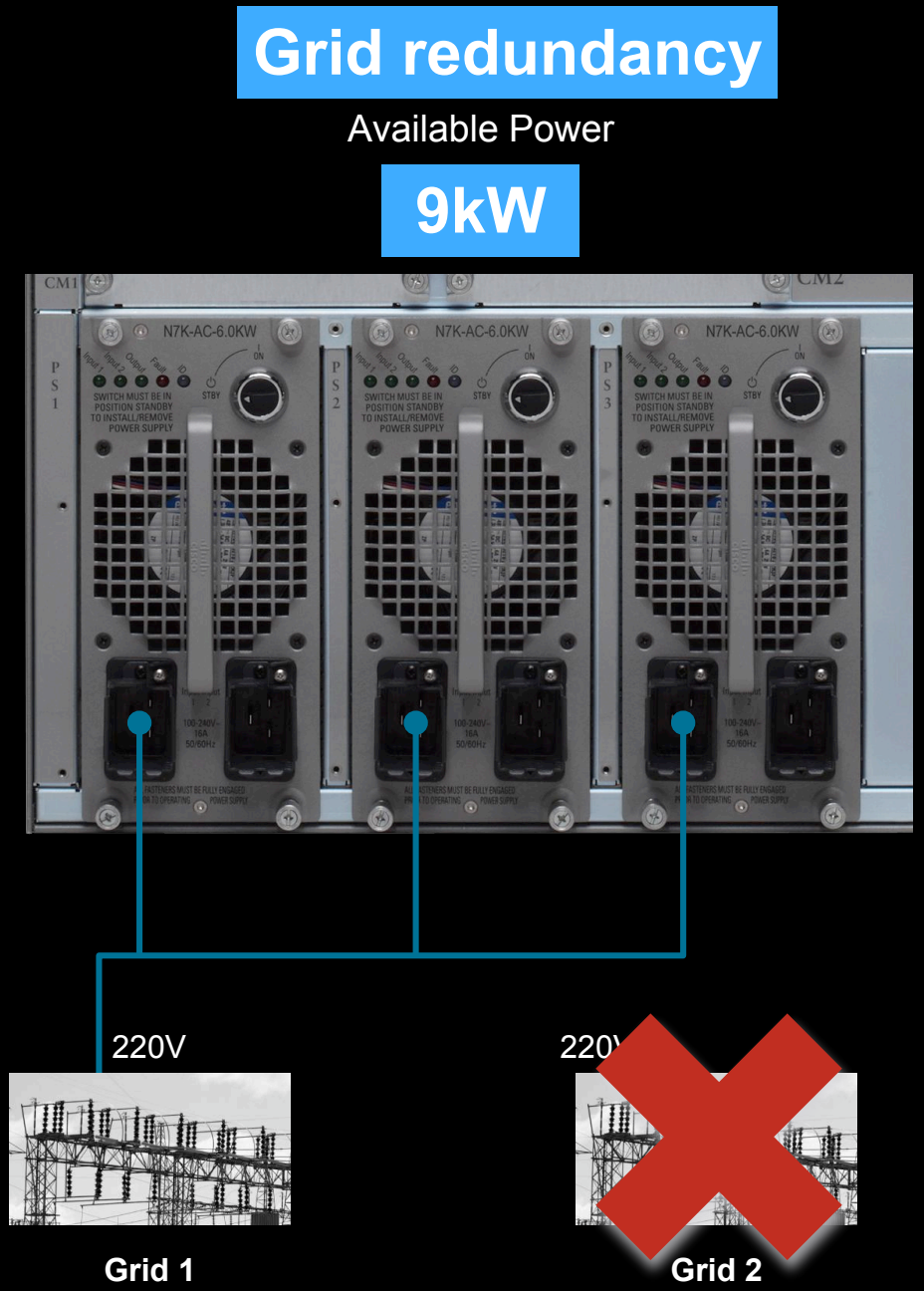
- N+1 redundancy – Reserves capacity equal to sum of lowest two power supplies



Power Redundancy

Power redundancy mode dictates how system budgets power

- N+1 redundancy – Reserves capacity equal to sum of lowest two power supplies
- Grid/input source redundancy – Reserves capacity equal to sum of half capacity of each power supply



Power Redundancy

Power redundancy mode dictates how system budgets power

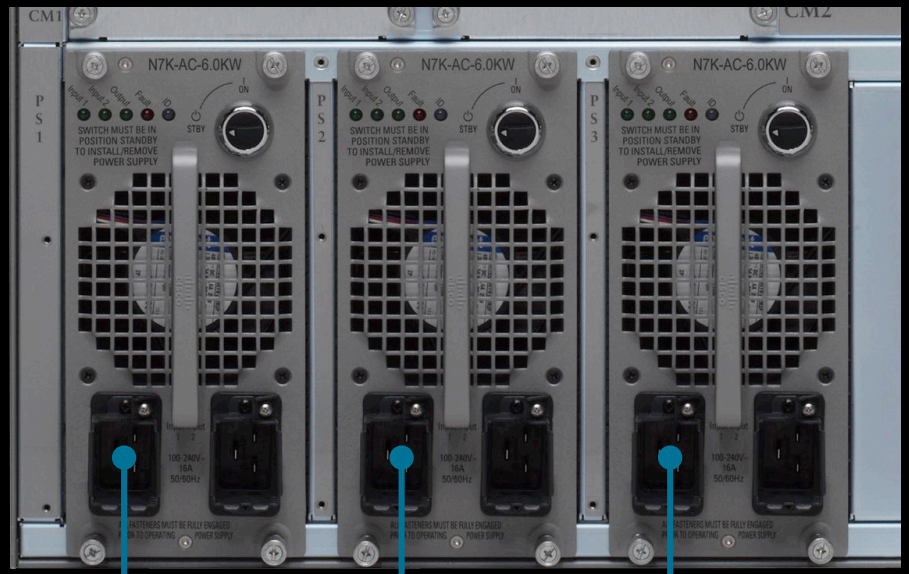
- **N+1 redundancy** – Reserves capacity equal to sum of lowest two power supplies
- **Grid/input source redundancy** – Reserves capacity equal to sum of half capacity of each power supply
- **Full redundancy** (default) – Provides capacity equal to the lesser of N+1 and grid redundancy capacity

System can provide full power redundancy for loaded 10G system (two supervisors, eight 10G I/O modules)

Full redundancy

Available Power

9kW



Grid 1

Grid 2

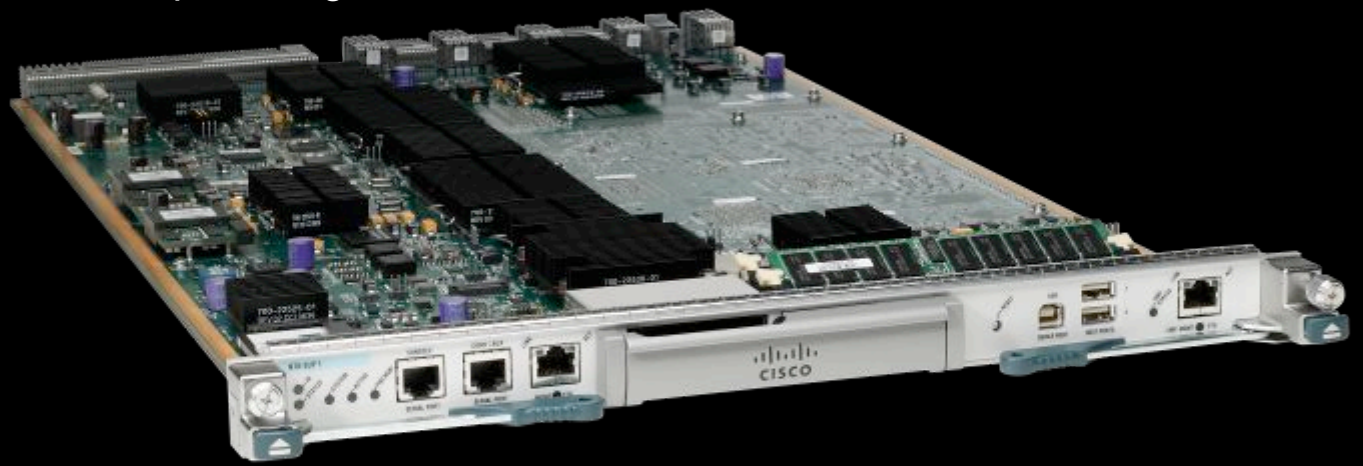
Cable Management

- Integrated cable management tray with straps
- Cable grooming to right, left, or split
- Can route up to 384 Cat6A cables to one side of chassis – worst-case scenario

er and lockable front
accidental

Supervisor Engine 1

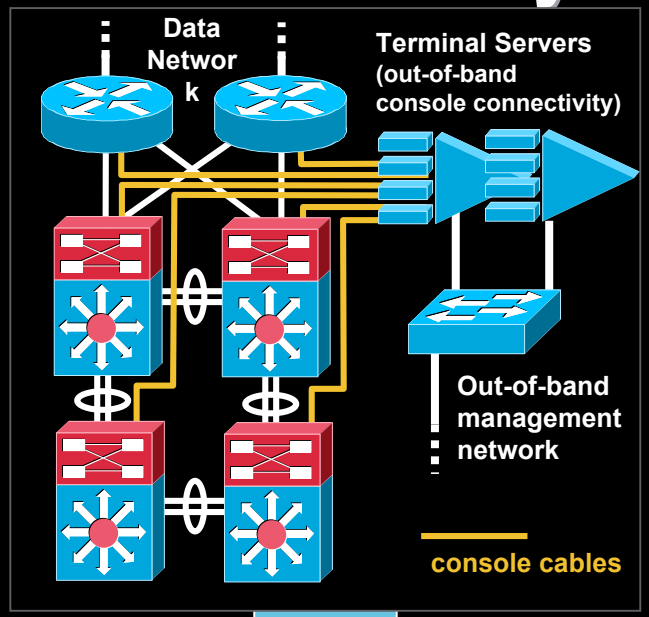
- High-performance dual-core 1.86GHz Intel Xeon processor
- Connectivity Management Processor (CMP) for lights-out management via dedicated 10/100/1000 Ethernet
- Memory – 4GB DRAM, 2GB internal bootdisk, 2GB External Expansion Flash (Optional), 8GB External Log flash
- 10/100/1000 management port with 802.1ae LinkSec
- Console & Auxiliary serial ports
- One USB device port, two USB host ports
- Blue beacon LED for easy location
- Resilient switched GE out-of-band channels to I/O modules
- Central arbiter providing I/O modules access to switch fabric

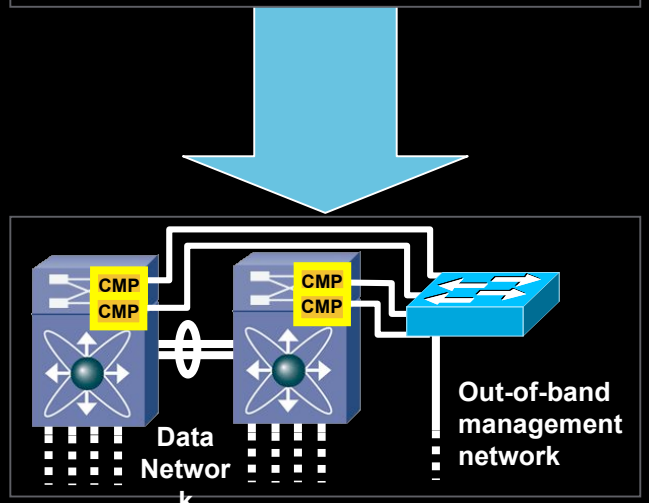


Management Ethernet Interface

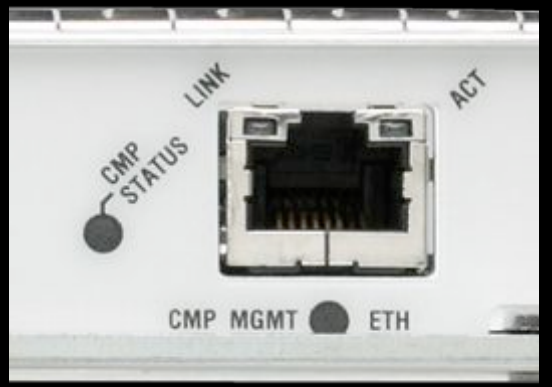
- 10/100/1000 interface used exclusively for system management
- Belongs to dedicated "management" VRF
  - Prevents data plane traffic from entering/exiting from mgmt0 interface
  - Cannot move mgmt0 interface to another VRF
  - Cannot assign other system ports to management VRF
- Supports IEEE 802.3ae LinkSec encryption

Connectivity Management Processor (CMP)





- Standalone, always-on microprocessor on supervisor engine
- Provides 'lights out' remote management and disaster recovery via 10/100/1000 interface
  - Removes need for terminal servers
- Monitor supervisor and modules, access log files, power cycle supervisor, etc.
  - Runs lightweight Linux kernel and network stack
  - Completely independent of NX-OS on main CPU



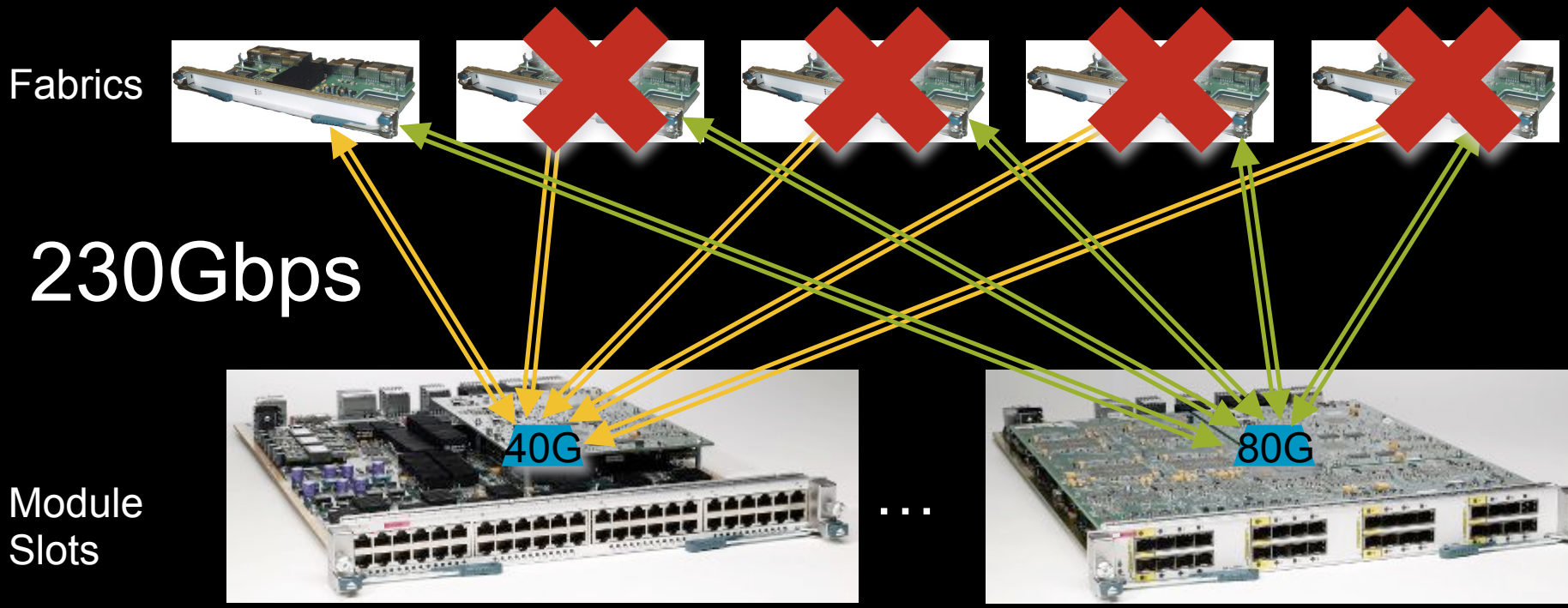
Fabric Module 1

- Provides 46Gbps per I/O module slot
- Up to 230Gbps per slot with 5 fabric modules

- Load-sharing across all fabric modules in chassis
- Multilevel redundancy with graceful performance degredation
- Non-disruptive OIR
- Blue beacon LED for easy location

Fabric Capacity and Redundancy

- Per-slot bandwidth capacity increases with each fabric module
- 1G module requires 2 fabrics for N+1 redundancy
- 10G module requires 3 fabrics for N+1 redundancy
- 4th and 5th fabric modules provide additional level of redundancy
- Future modules will leverage additional fabric bandwidth
- Fabric failure results in reduction of overall system bandwidth



Fabric Bandwidth Capacity

- Initial shipping fabric bandwidth: 230 Gbps per payload slot, 115 Gbps per supervisor slot
  - Forwarding Engine on initial shipping modules cannot leverage all fabric bandwidth
  - Future modules will be capable of leveraging full bandwidth
- In Nexus 7010 chassis:
  - **I/O Modules:** (230 Gbps/slot) x (8 payload slots) = **1840 Gbps**
  - **Sup Slots:** (115 Gbps/slot) x (2 supervisor slots) = **230 Gbps**
  - **Overall:** (1840 + 230 = 2070 Gbps) x (2 for full duplex operation) = **4140 Gbps** = **4.1 Tbps system bandwidth**

4.1 Terabits per second fabric bandwidth capacity

Future Vision for Platform Series

- Future goal to more than double the fabric bandwidth
  - 500+ Gbps bandwidth per slot
  - Requires future fabric module
- 10 slot chassis will scale to 9+ Tbps system bandwidth
- 18 slot chassis will scale to 15+ Tbps system bandwidth

15+ Terabits per second platform bandwidth capacity

What Is Virtual Output Queuing?

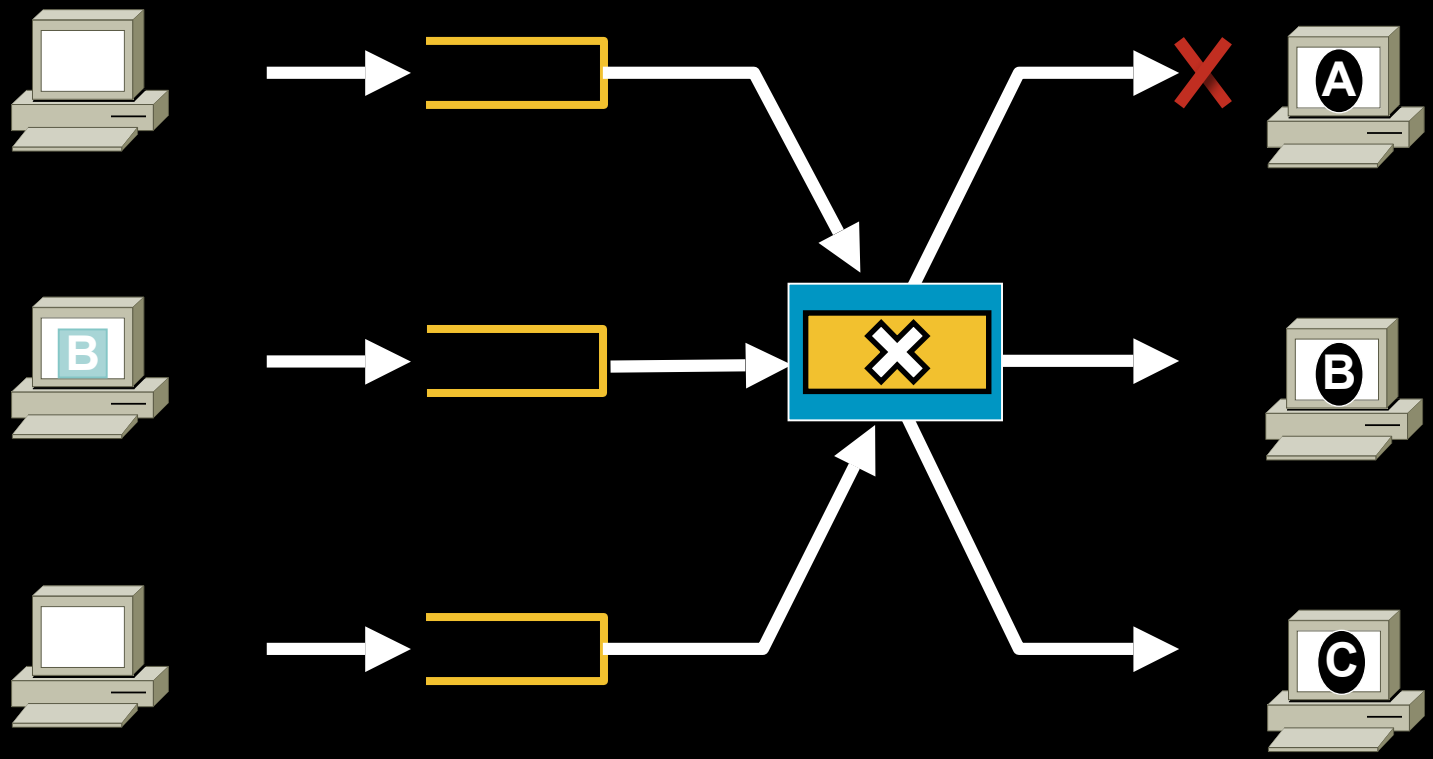
- **VOQs on ingress modules represent buffer availability on egress modules**
- Guaranteed delivery to egress module for arbitrated packets entering fabric
  - **If VOQ available on ingress, buffer available on egress**
- VOQ is NOT equivalent to ingress or egress port buffer or queues
  - Relates ONLY to ASICs at ingress and egress to fabric
- VOQ is "virtual" because it represents EGRESS buffer but resides on INGRESS module

AL e s

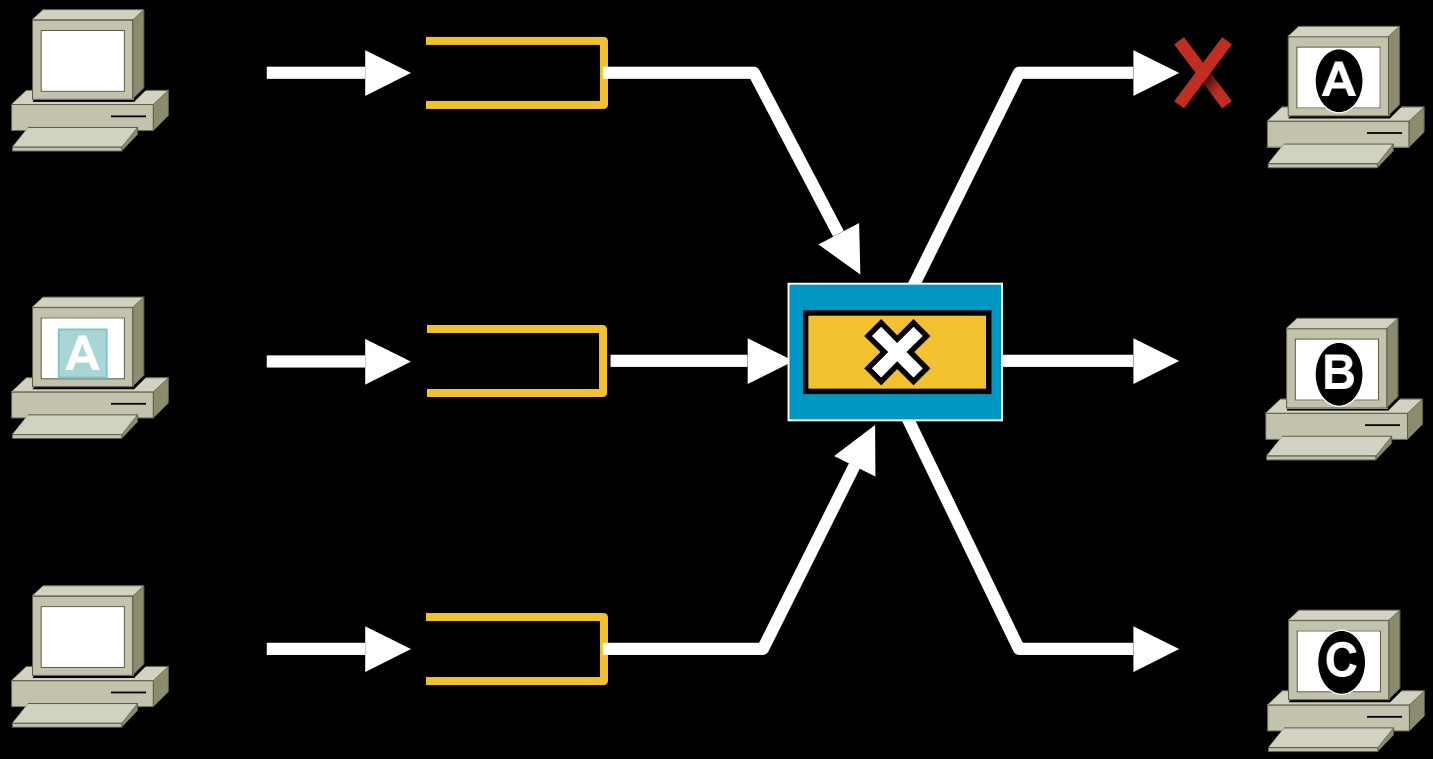
Centralized Fabric Arbitration

- Access to fabric bandwidth controlled by central arbiter on supervisor
  - In other words, access to the VOQ for the destination across the fabric
- Arbitration works on credit request/grant basis
  - Modules communicate egress fabric buffer availability to central arbiter
  - Modules request credits from supervisor to place packets in VOQ for transmission to destination over fabric
  - Supervisor grants credits based on egress fabric buffer availability for that destination
- Arbiter discriminates among four classes of service
  - Priority traffic takes precedence over best-effort traffic across fabric

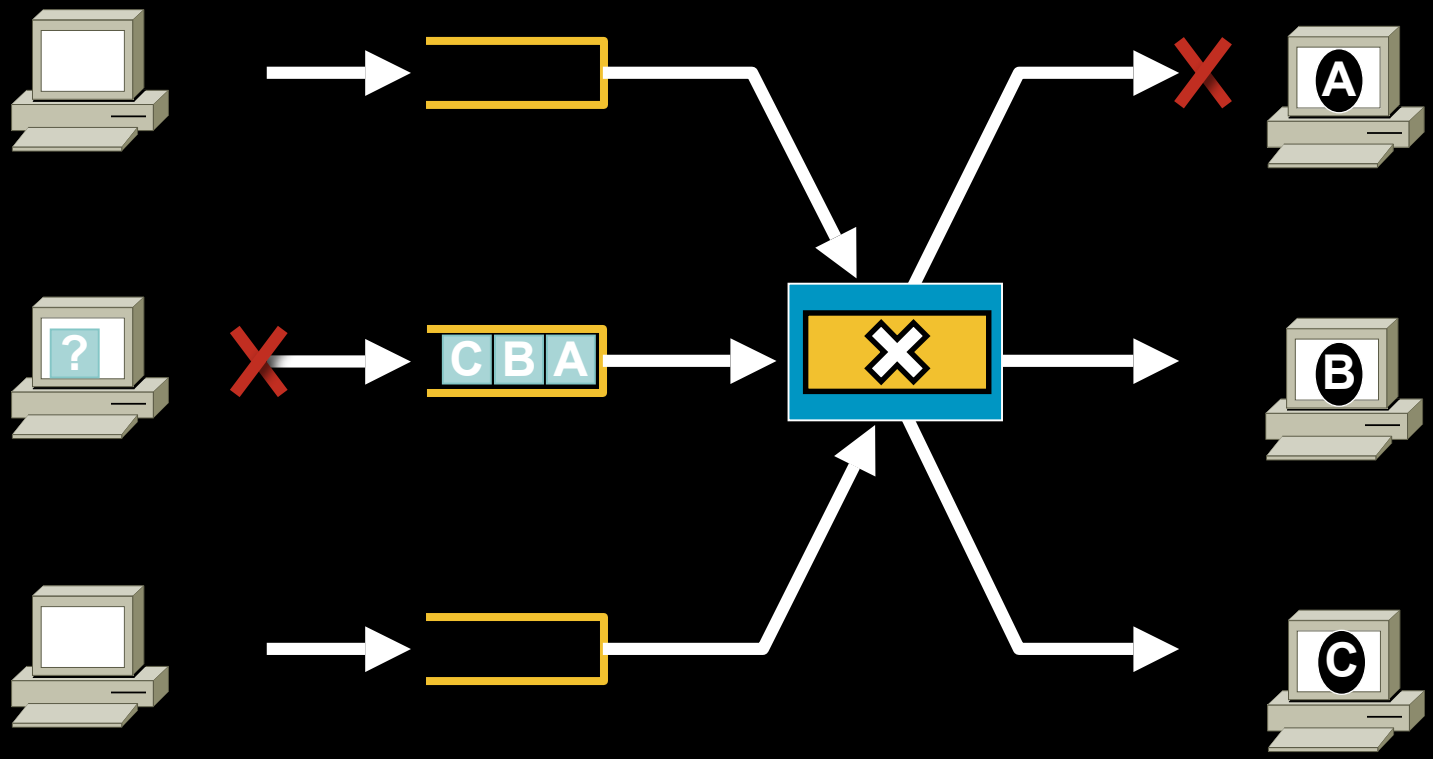
Head of Line Blocking



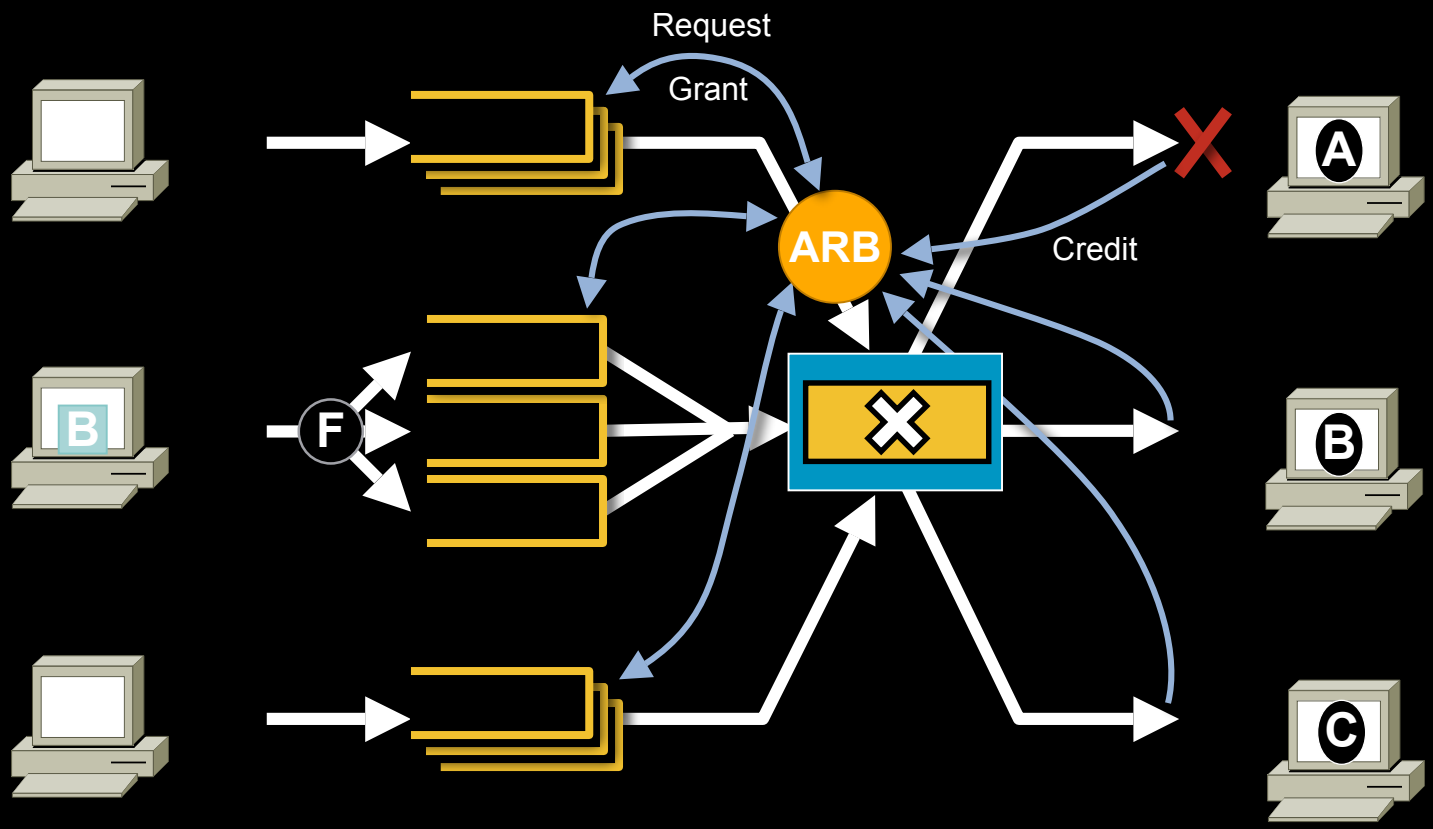
Head of Line Blocking



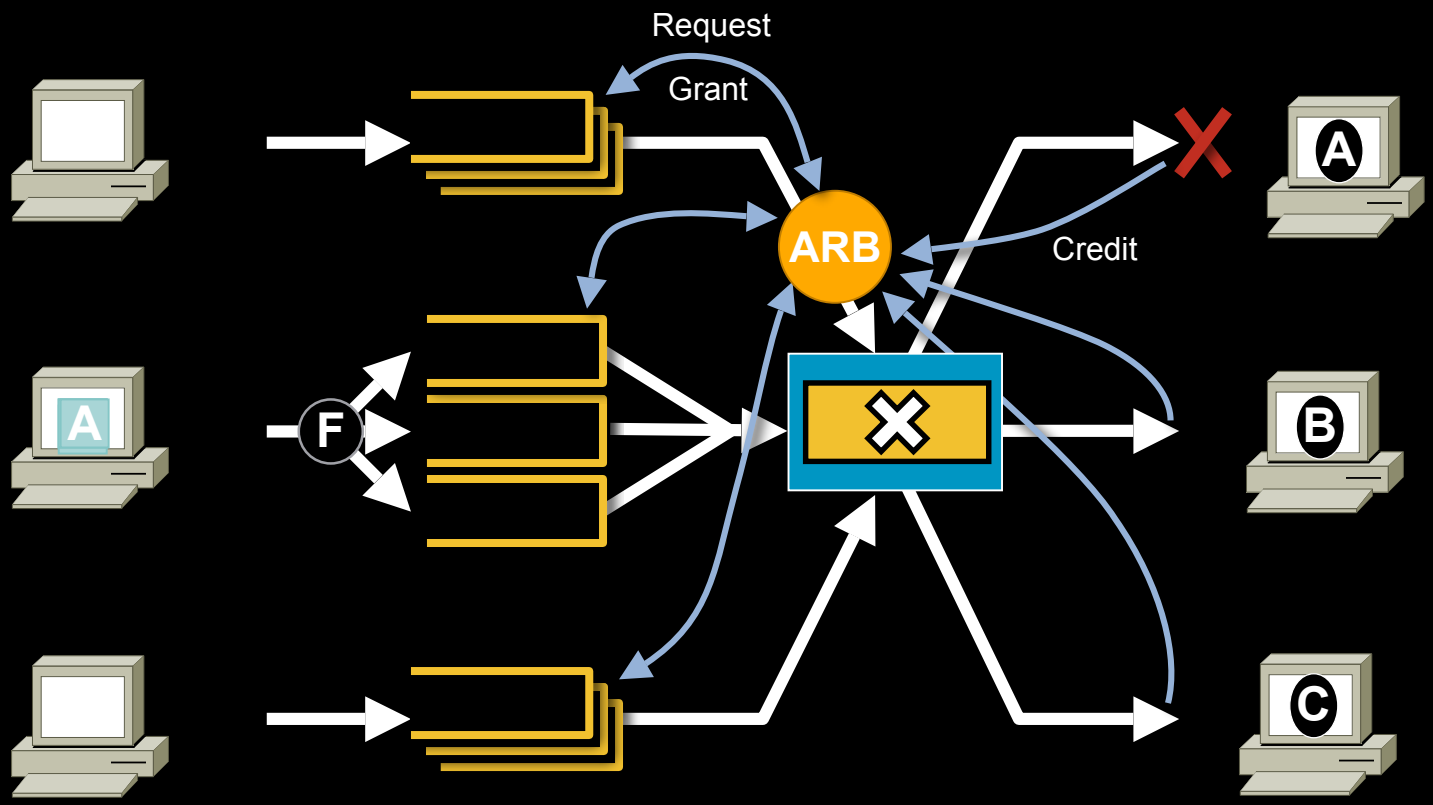
Head of Line Blocking



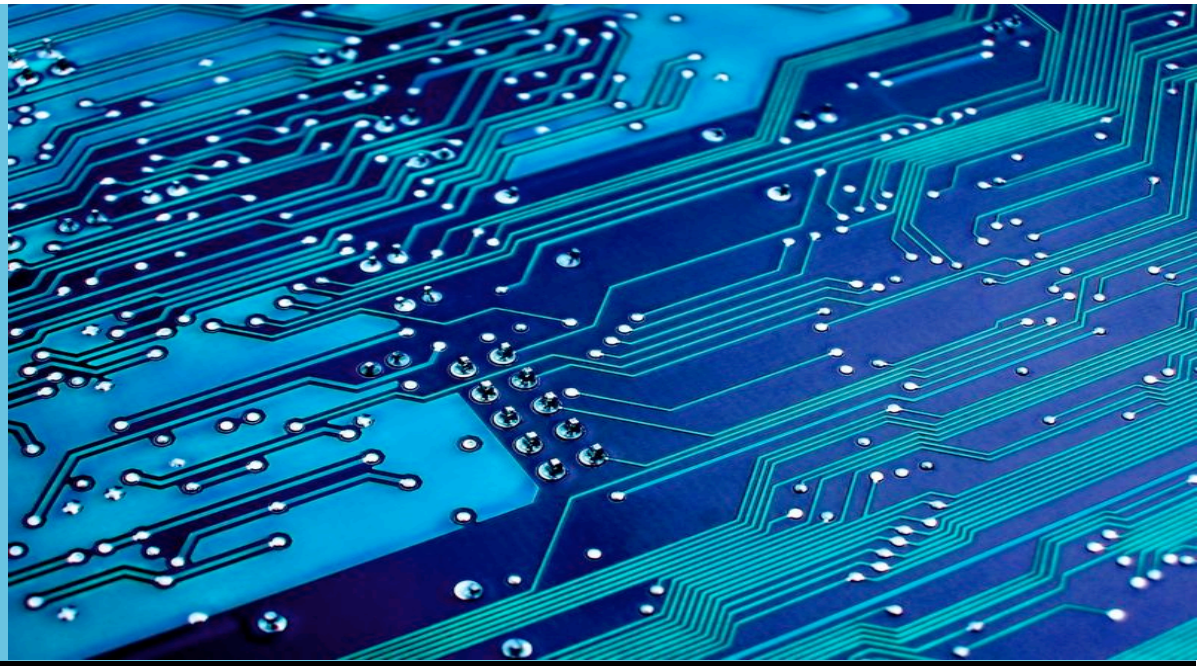
Virtual Output Queues and Arbitration



Virtual Output Queues and Arbitration



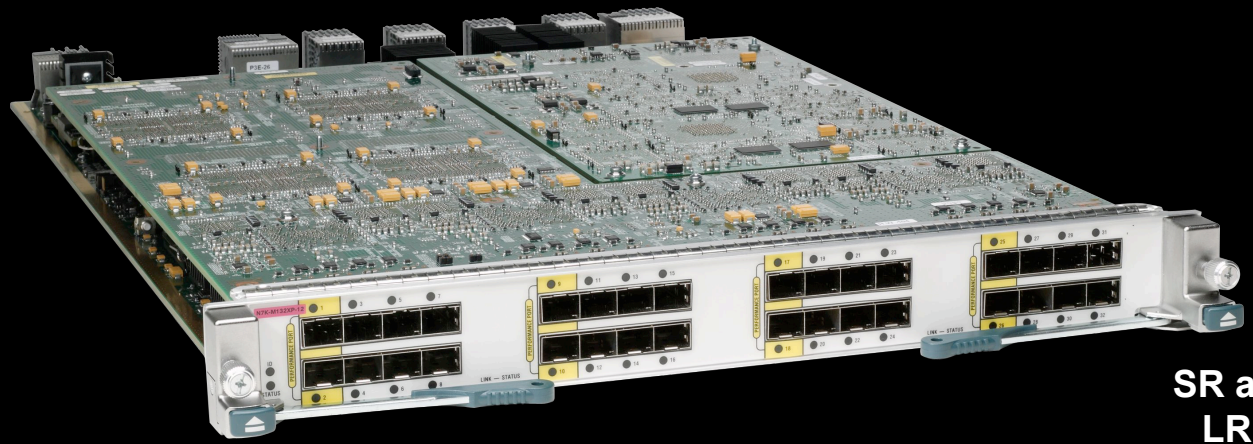
Cisco Nexus 7000 Series Hardware



I/O Modules (Linecards)
Integrated Packet Forwarding Engine (EARL8)

32-Port 10GE I/O Module, 80G Fabric

- 32 10GE ports with SFP+ transceivers
- 80G full duplex fabric connectivity
- Integrated 60Mpps forwarding engine for fully distributed forwarding
- 4:1 oversubscription at front panel
- Two operational modes, enabled on per-port-group basis
  - Dedicated mode: up to 8 line-rate 10G interfaces per module
  - Shared mode: up to 32 4:1 oversubscribed interfaces per module
- Virtual output queueing (VOQ) ensuring fair access to fabric bandwidth
- 802.1AE LinkSec on every port
- Buffering:
  - Dedicated mode: 96MB ingress, 80MB egress
  - Shared mode: 1MB ingress, 80MB egress (4 ports)
- Queues: 8q2t ingress, 1p7q4t egress
- Blue beacon LED for easy location

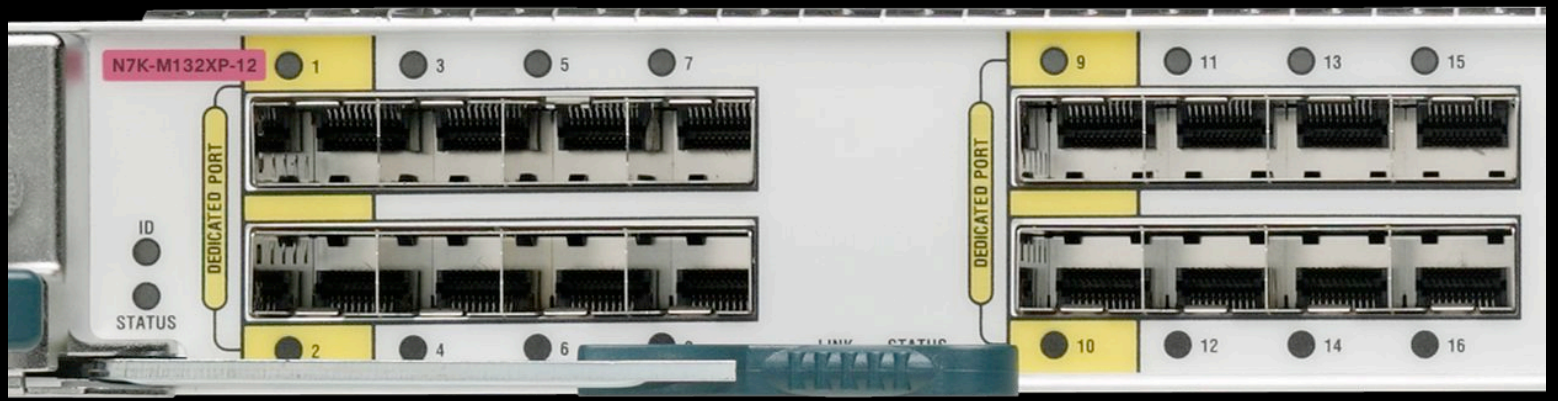


SFP+

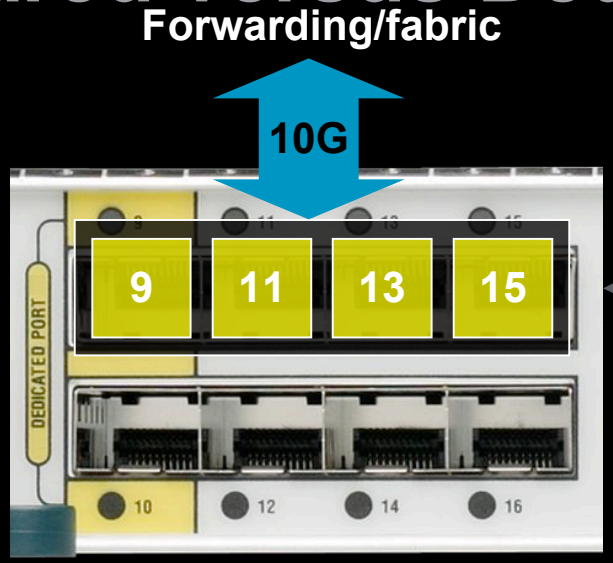
SR at initial release – 300m over MMF
LR post-release – 10km over SMF

10G Interface Rate Modes

- 10G interfaces support two rate modes
  - Dedicated mode
  - Shared mode
- Mode configurable on per-port-group basis
  - Port groups consist of four contiguous odd or even interfaces (e.g., {1,3,5,7}, {2,4,6,8}, {9,11,13,15}, etc.)
  - Dedicated mode disables three interfaces in port group
  - Dedicated mode interfaces: 1, 2, 9, 10, 17, 18, 25, 26

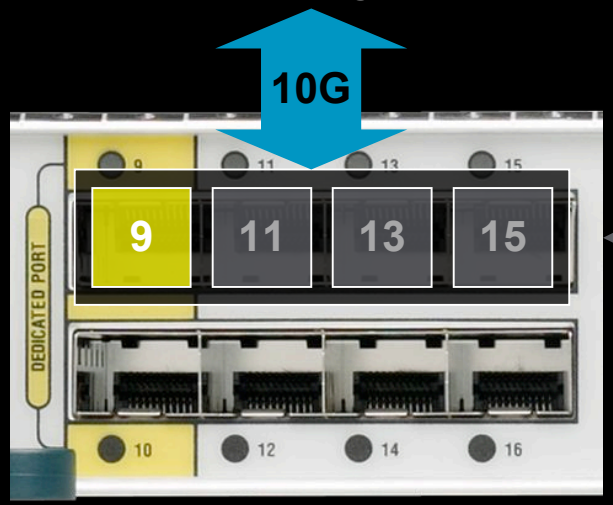


Shared versus Dedicated Mode



Shared mode

- Four interfaces share 10G bandwidth

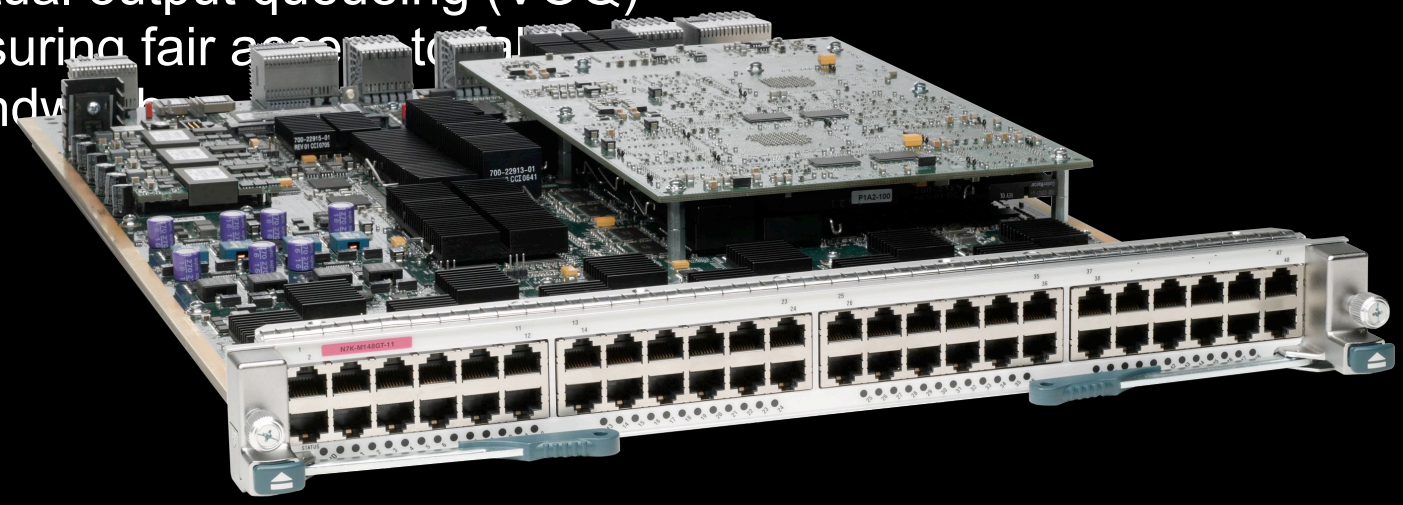


Dedicated mode

- One interface gets 10G bandwidth
- Three interfaces disabled

48-Port 1GE I/O Module

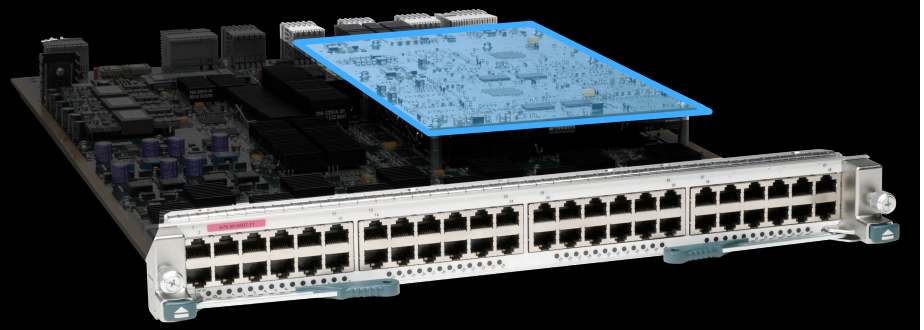
- 48 1GE 10/100/1000 RJ-45 ports
- 40G full duplex fabric connectivity
- Integrated 60Mpps forwarding engine for fully distributed forwarding
- Virtual output queueing (VOQ) ensuring fair a[illegible] t[illegible] bandw[illegible]
- 802.1ae LinkSec on every port
- Buffer: 7.5MB ingress, 6.2MB egress
- Queues: 2q4t ingress, 1p3q4t egress
- Blue beacon LED for easy location



Integrated Forwarding Engine Hardware

Advanced hardware forwarding engine **integrated on every I/O module**

- 60Mpps Layer 2 bridging with hardware MAC learning
- 60Mpps IPv4 and 30Mpps IPv6 unicast
- IPv4 and IPv6 multicast (SM, SSM, bidir)
- IPv4 and IPv6 security ACLs
- Cisco TrustSec security group tag support
- Unicast RPF check and IP source guard
- QoS remarking and policing policies
- Ingress and egress NetFlow (full and sampled)
- GRE tunnels



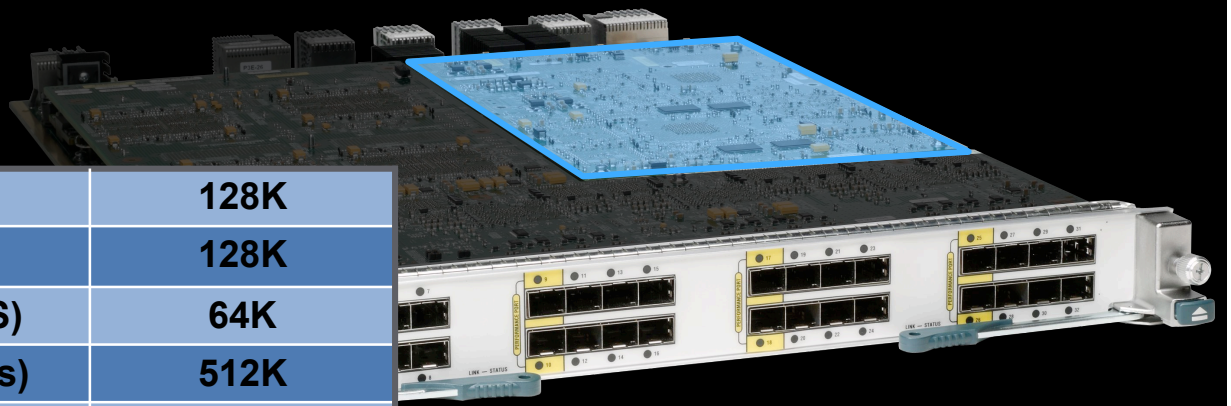


Table sizes optimized for Data Center

FIB TCAM	128K
MAC table	128K
Classification TCAM (ACL and QoS)	64K
NetFlow TCAM (Ingress and Egress)	512K
Policers	16K

EARL8 Forwarding Engine Details

- Forwarding engine chipset consists of two ASICs:
- Layer 2 Engine
  - 120Mpps (60Mpps effective) Layer 2 lookups
  - Performs ingress and egress SMAC/DMAC lookups
  - Hardware MAC learning
  - True IP-based Layer 2 multicast constraint
  - Performs lookups on ingress I/O module, and egress I/O module for bridged packets
- Layer 3 Engine
  - 60Mpps IPv4 and 30Mpps IPv6 Layer 3/Layer 4 lookups
  - Performs all FIB, ACL, QoS, NetFlow processing
  - Linear, pipelined architecture – every packet processed in ingress and egress pipe
  - Performs lookups on ingress I/O module, and egress I/O module for multicast replicated packets

Nexus 7000 Software: NX-OS

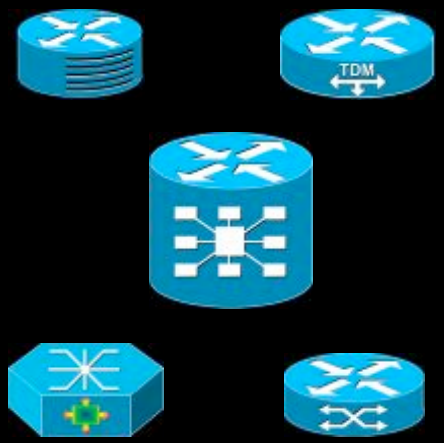
Cisco Network Operating Systems

The Right Tools in the Right Places

Cisco IOS® XR Software

Core WAN-Focused

Service provider-grade services
State-of-the-art resiliency
Large-scale networks
WAN-core link-layer types



Cisco IOS Software

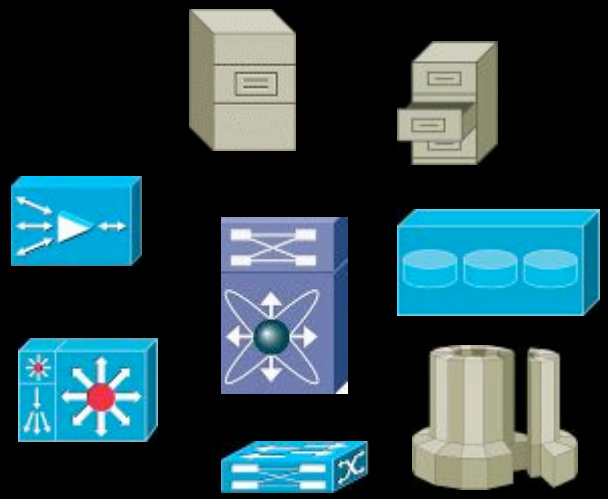
Ubiquitous

Powerful feature set
Powerful flexibility
Robust resiliency
LAN/WAN link-layer types

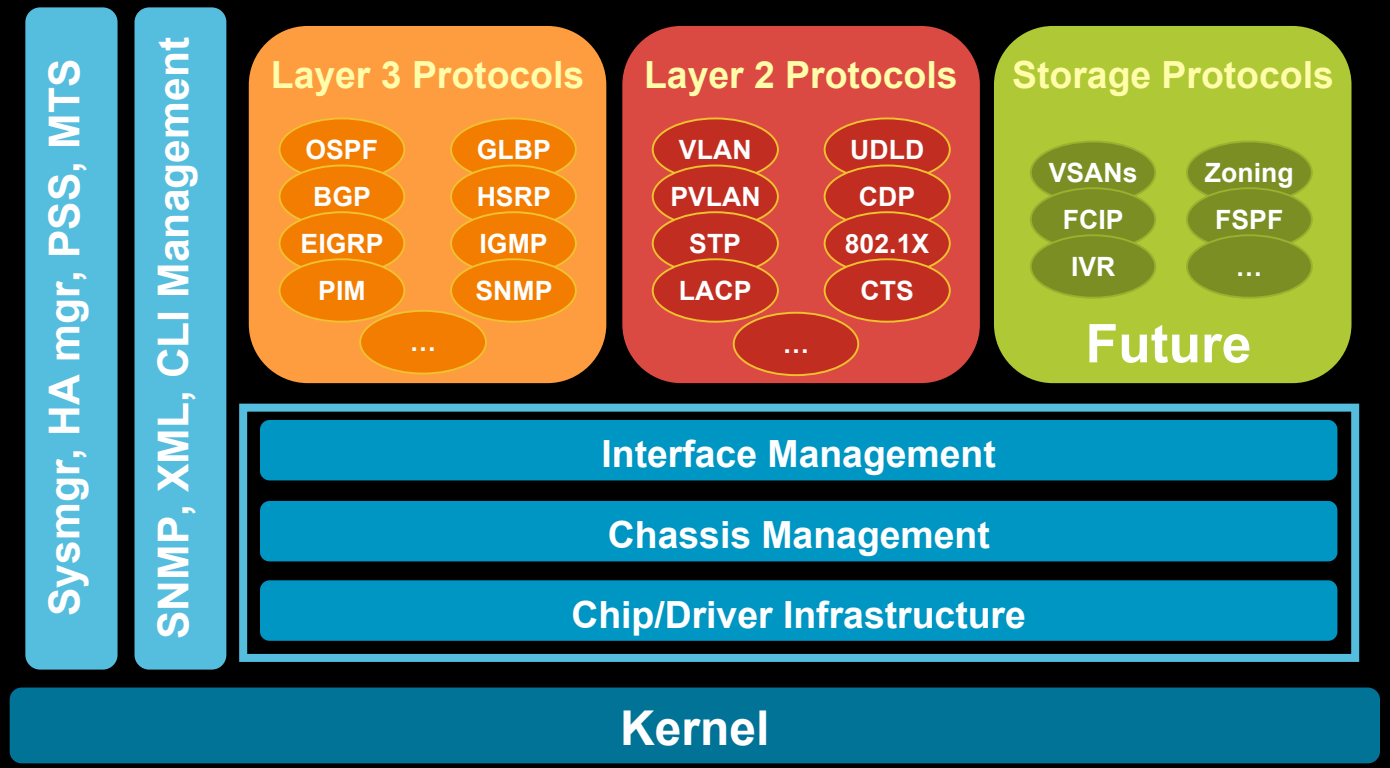
Cisco® NX-OS

Data Center-Focused Feature Set

Mission-critical environments
24-hour continuous operations
High density and performance Ethernet
Data center-specific link-layer types



NX-OS



- Core infrastructure leveraged from SAN-OS
  - Linux kernel plus system management, reliable messaging, state dbase, HA manager
- Layer 3 protocols and infrastructure
  - Modern, streamlined, scaleable routing implementation
- Layer 2 protocols and infrastructure
  - Data-center focused, standards-based Layer 2 feature set
- Storage protocols taken directly from SAN-OS
  - In future, when FCoE modules are delivered , storage functionality “plugs in” to OS

Feature Rich - r4.0 Overview

Resilient Architecture

ISSU+
Rapid Switchover
Highly Modular

Stateful Process Restart
NSF/GR
Plug-in/Patch Support

Connectivity Management Processor
Reliable Process Communication
Independent/Isolated Processes

Virtualized Services

Layer 2

802.1s/w - Spanning-tree
16k VLANs
802.1Q - Trunking
802.3ad - Link Agg.
IGMP snooping
Private VLANs.
UDLD

Layer 3

OSPF, EIGRP, ISIS, BGP
PIM – ASM, BiDir, SSM
BSR, Auto-RP, Static
MSDP
IGMPv1/2/3 MLDv2
HSRP, GLBP, VRRP

Security

Cisco Trusted Security
Cisco Integrated Security
CoPP
NAC
ACLs (L2 - L4)
Port Security
802.1x

Operations

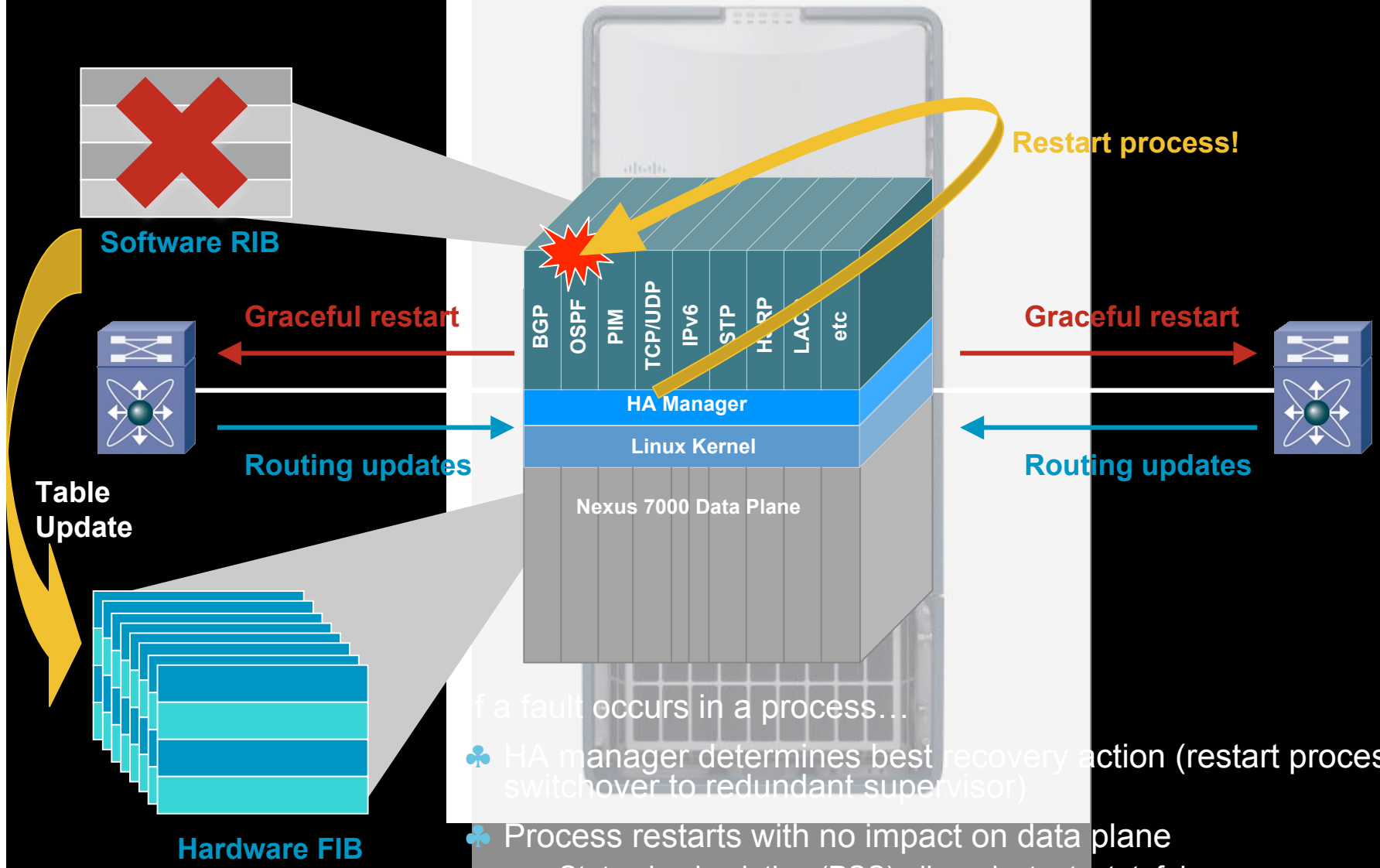
EEM, GOLD
Smart CallHome
XML API
RBAC
SPAN
QoS
NetFlow

Comprehensive Feature Set - End-to-end Data Center

Nexus 7000 Series High Availability

- Hardware provides redundancy at every component level:
  - Supervisors
  - Power
  - Fabrics
  - Fans
- Software offers multi-layered, multi-faceted resiliency:
  - Stateful process restarts
  - Graceful restart for routing protocols
  - Stateful supervisor engine switchovers
  - True in-service software upgrades
- Hardware and software combine to deliver data-center class high availability – zero service disruption

Stateful Fault Recovery

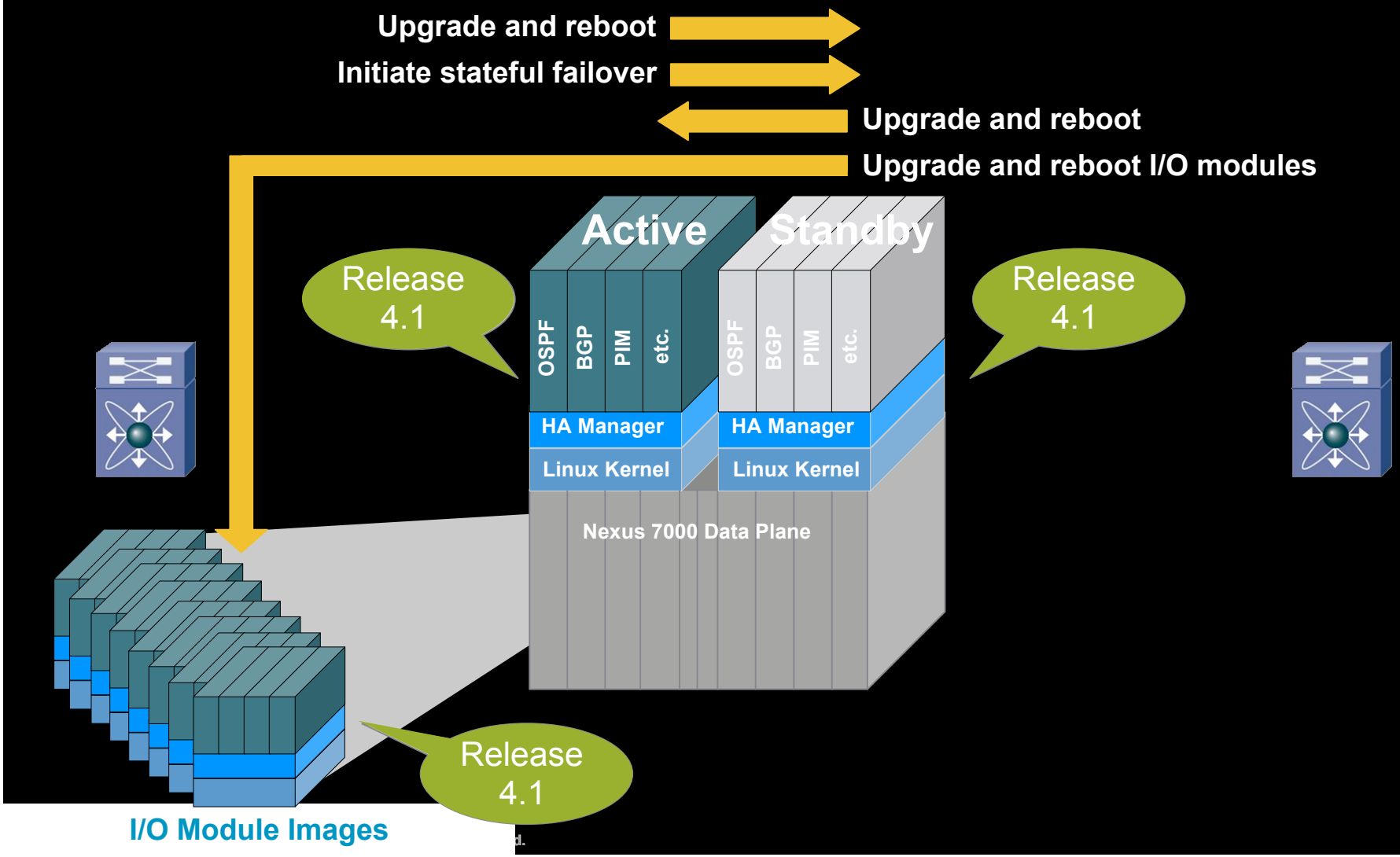


If a fault occurs in a process…

- HA manager determines best recovery action (restart process, switchover to redundant supervisor)
- Process restarts with no impact on data plane
  - State checkpointing (PSS) allows instant, stateful process recovery
  - Software utilizes Graceful Restart where appropriate

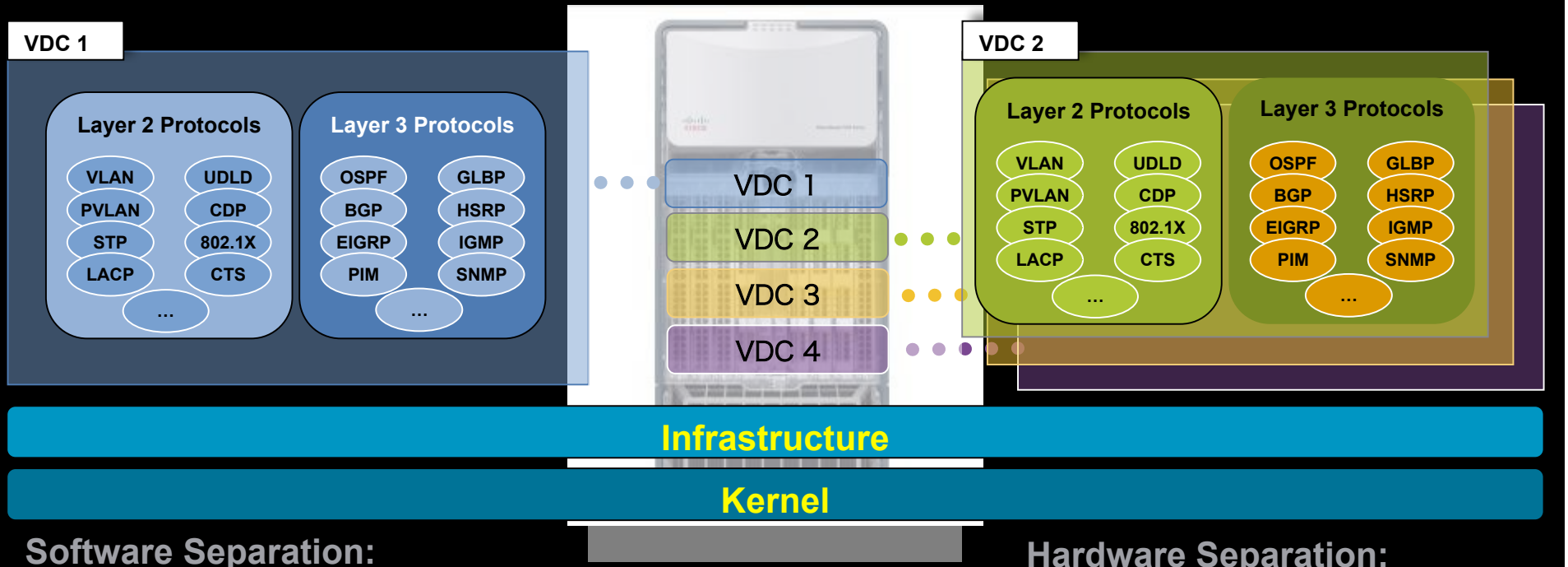
In-Service Software Upgrade

```
N7K# install all kickstart bootdisk:4.1-kickstart system bootdisk:4.1-system
N7K#
```



Virtualization with VDCs

Virtualization is key to maximizing resource utilization while providing strong security and software fault-isolation



Software Separation:

- Software fault isolation domains
- Addressing domains
- Service differentiation domains
- Management domains
- Resource allocation
- Security domains

Shared resources:

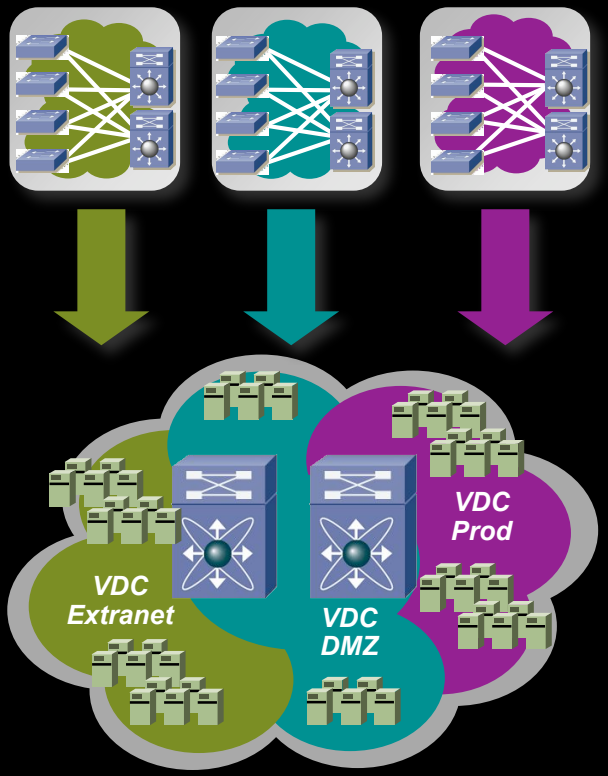
- Software Infrastructure
- Kernel
- Power Supplies
- Fans
- Chassis

Hardware Separation:

- Individual Physical Ports
  - Layer 2
  - Layer 3
- Port Channels
- Entire Linecards

VDC Use Cases

- Enables collapsing of multiple logical networks into single physical infrastructure while maintaining strong security, administration, and fault isolation
- Appropriate for typical silo/stovepipe designs such as:
  - Production, Development, Test
  - Intranet, Internet, DMZ, Extranet
  - Organization A, Organization B, Organization C
  - Application A, Application B, Application C
  - Customer A, Customer B, Customer C
  - Cluster A, Cluster B, Cluster C etc.



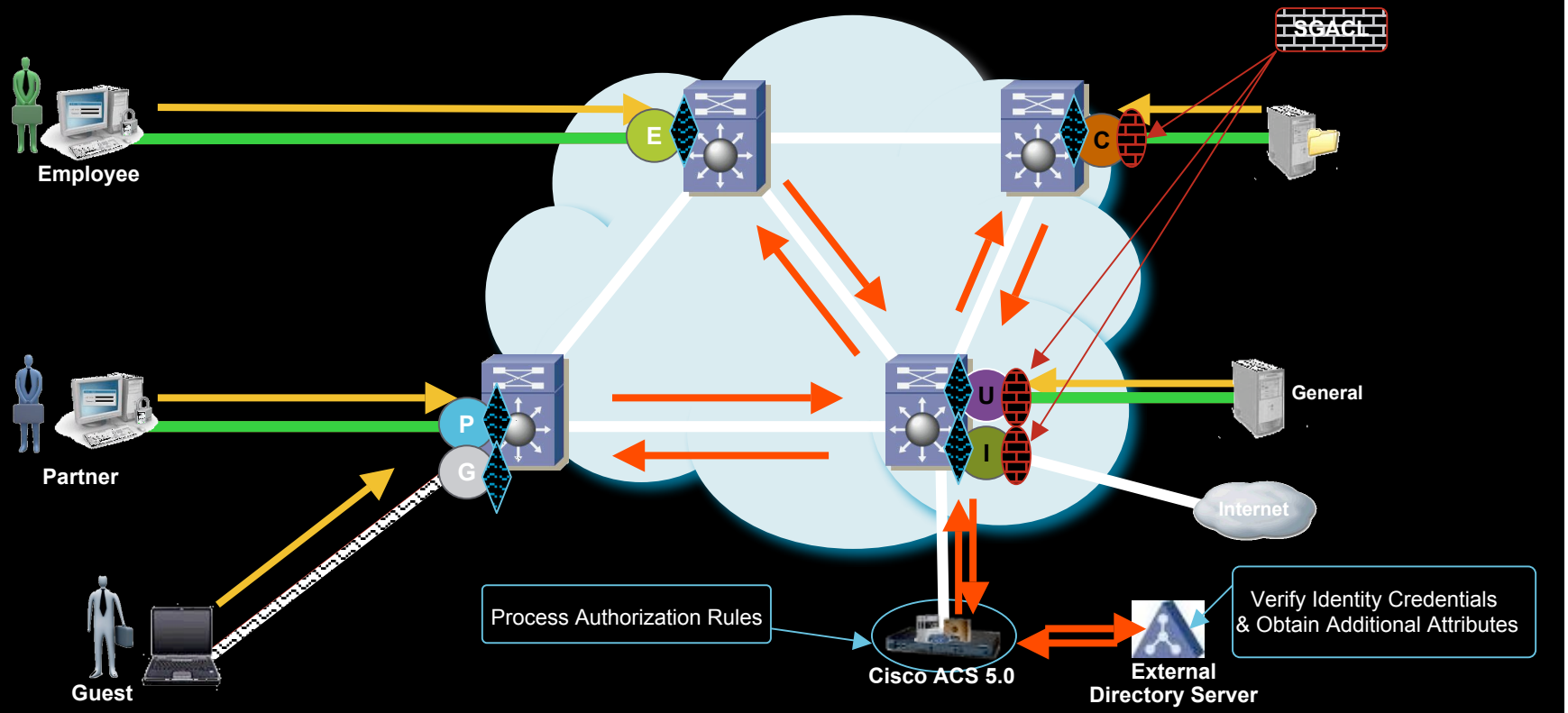
Different network islands *virtualized* onto common data center networking infrastructure

System Capacity…	…with 1 VDC on 1 I/O module	…with 1 VDC on 8 I/O modules	…with 8 VDCs, 1 VDC per I/O module
FIB TCAM	128K	128K	1M (8 x 128K)
MAC table	128K	128K	1M (8 x 128K)
Classification TCAM	64K	64K	512K (8 x 64K)
Policers	16K	16K	128K (8 x 16K)

Cisco TrustSec

- First Cisco product to deliver Cisco TrustSec, a cross-platform, multi-phase policy-based admission- and access-control solution
- Provides network-based confidential communications (IEEE 802.3AE MACSec) while preserving existing network services (such as Netflow, QoS, load-balancing, L4-L7 Services)
- Simplifies Access Control by tagging packets with a Security Group Tag.
  - Access Control doesn't depend on IP addresses anymore

CTS – Network Admission Control



Legend

Link/Port Status

- Unauthenticated
- Failed Auth
- Authenticated
- Shutdown

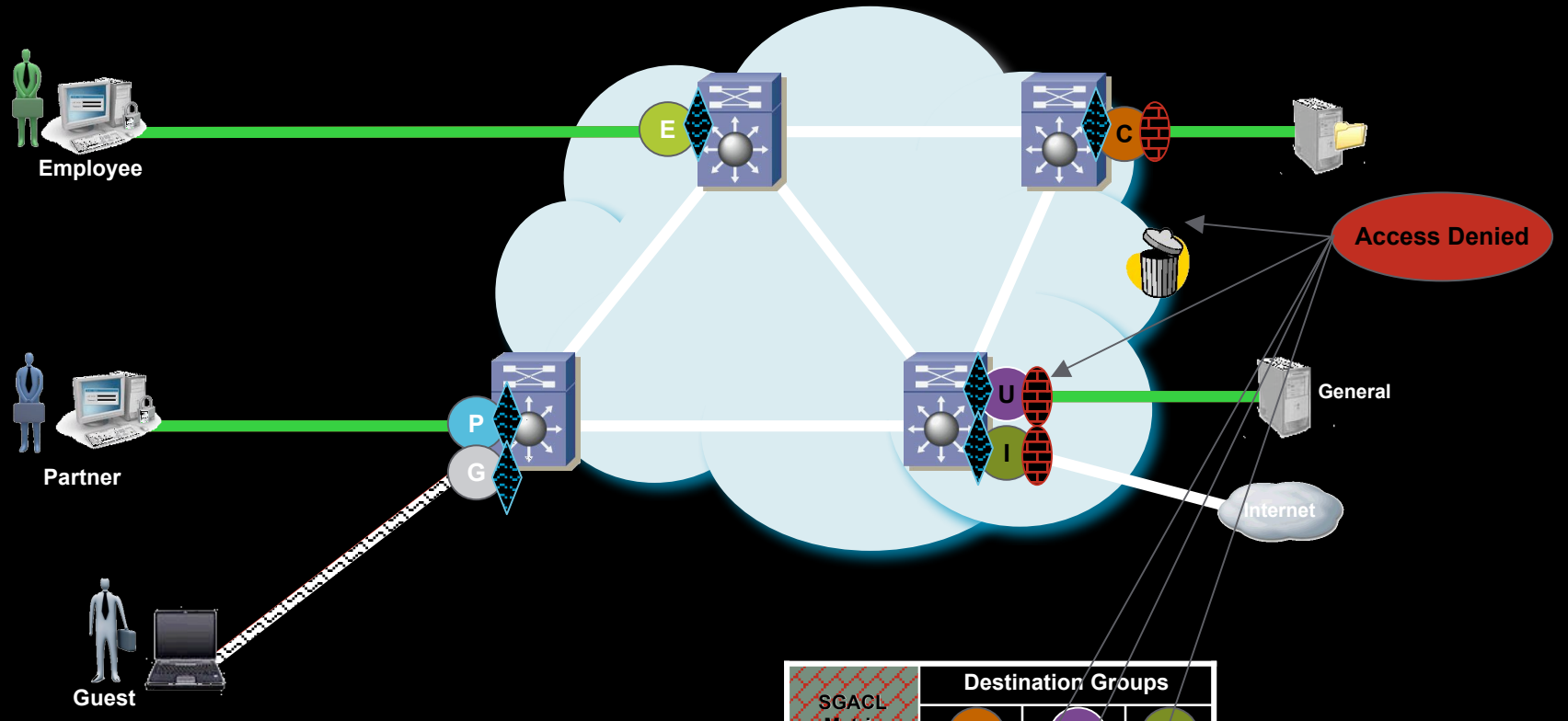
- Ingress Tagging
- Egress Filtering

Security Group Classifications

- E Employee Group
- P Partner Group
- G Guest Group
- C Confidential Group
- U Unrestricted Group
- I Internet Group

1. Authentication Request
2. Radius & AD Authc/Authz
3. SGT Dynamically Assigned
4. SGACL Dynamically Applied
5. Links Up

CTS – Topology Independent Access Control



SGACL Matrix	Destination Groups: C	U	I
Source Groups: E	✓	✓	✓
P	⊘	✓	✓
G	⊘	⊘	✓

Legend

Link/Port Status
- Unauthenticated
- Failed Auth
- Authenticated
- Shutdown

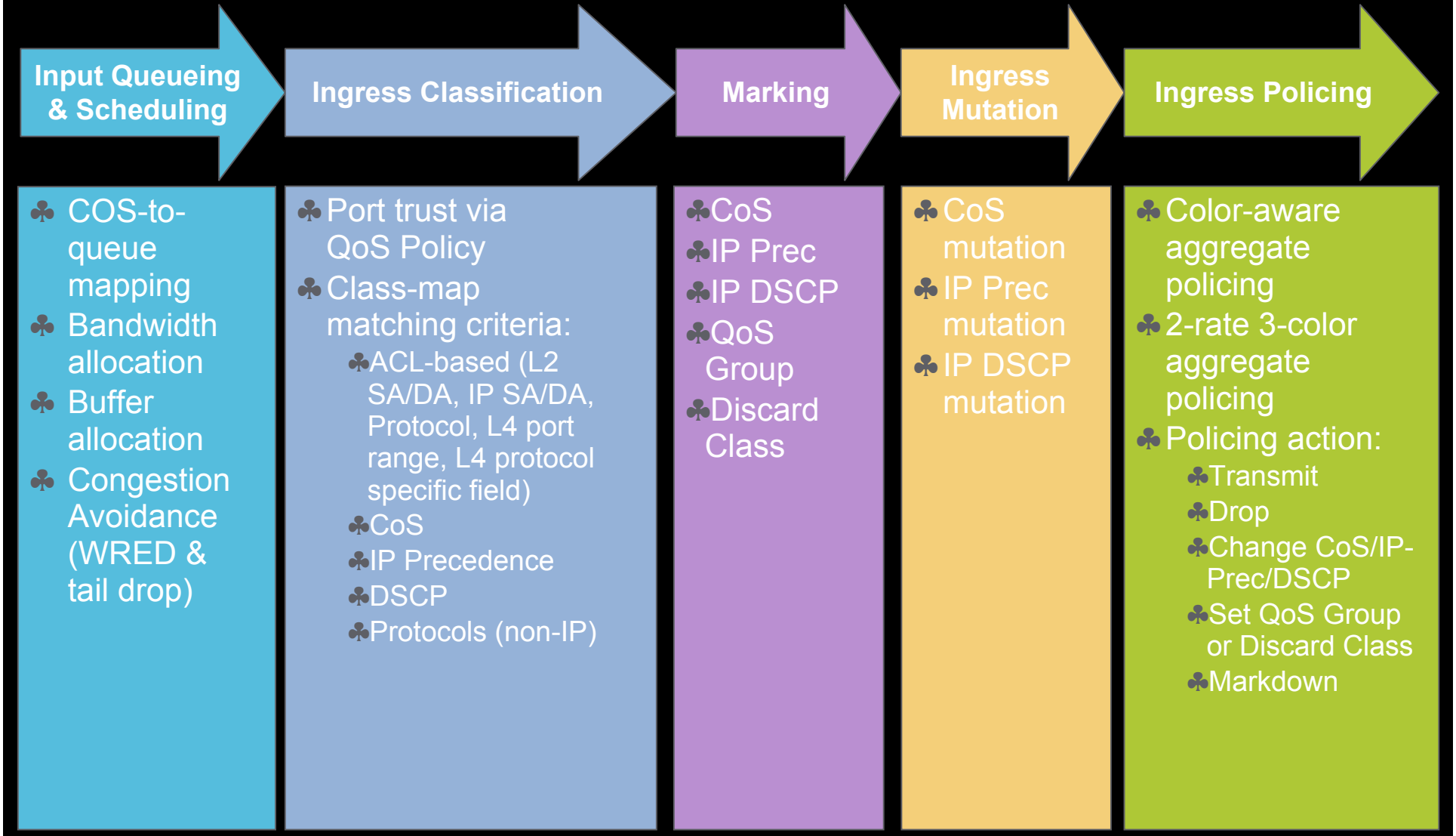
- Ingress Tagging
- Egress Filtering

Security Group Classifications
- E — Employee Group
- P — Partner Group
- G — Guest Group
- C — Confidential Group
- U — Unrestricted Group
- I — Internet Group

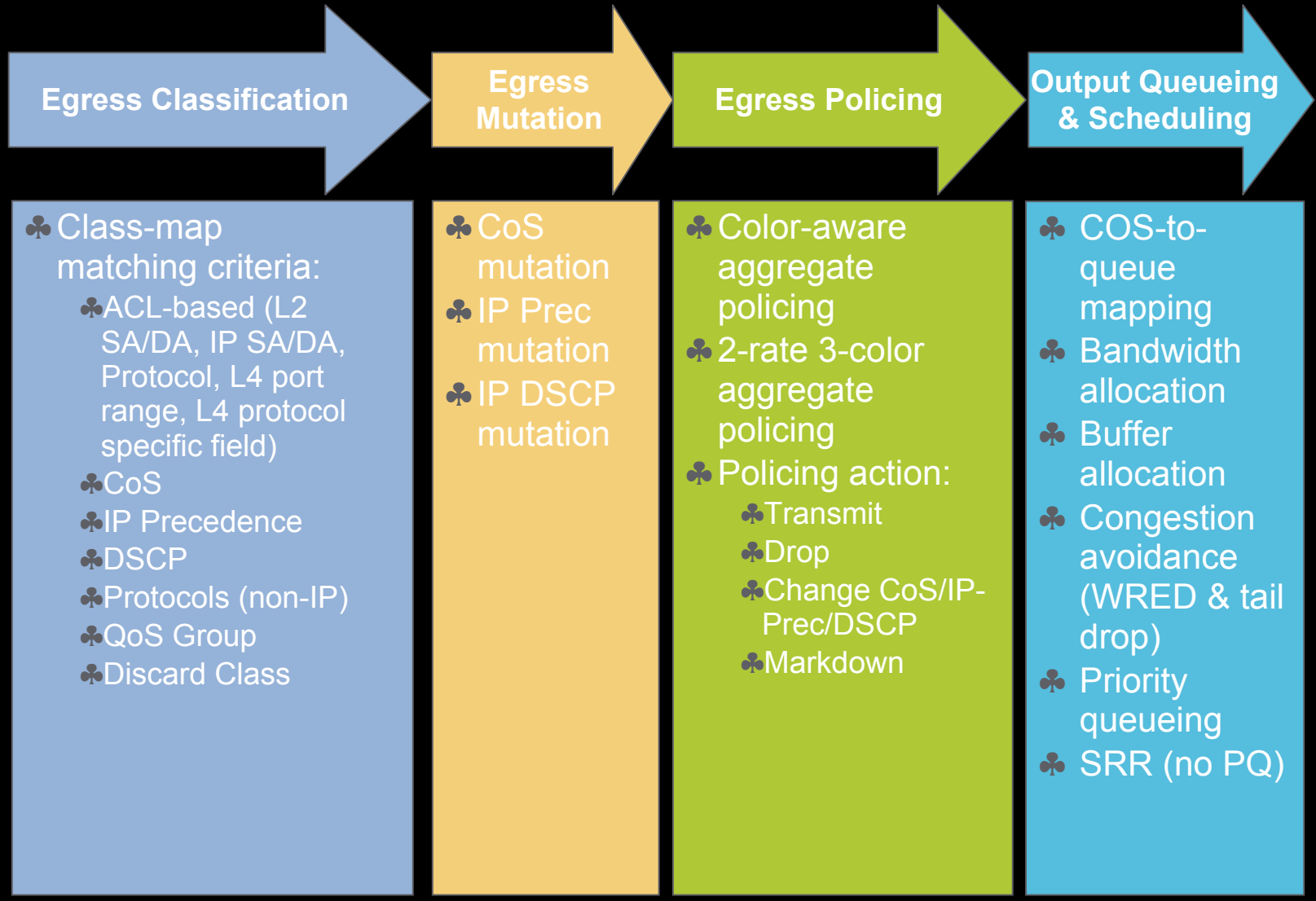
Quality of Service

- All configuration via Modular QoS CLI (MQC)
- Queueing policies leverage port ASIC capabilities to map traffic to queues and schedule packet delivery
- QoS marking and policing policies mark or police traffic, leveraging CL TCAM to match packets

QoS Ingress Feature Set



QoS Egress Feature Set



NX-OS Licensing

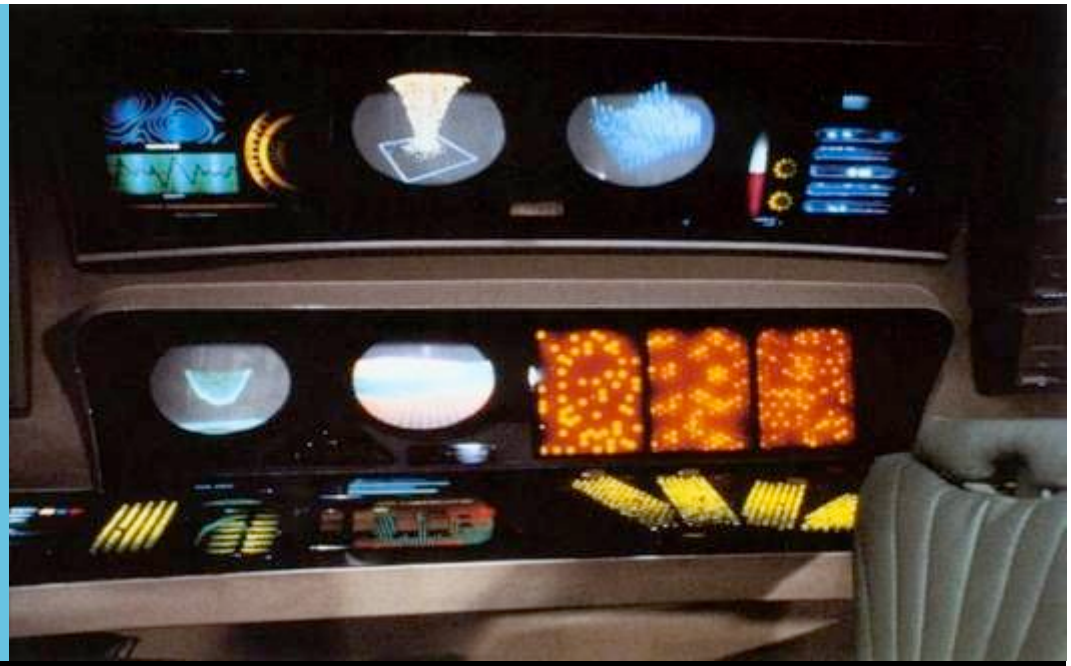
Simple, flexible licensing model

- Three levels of enforced licensing
- Base, Enterprise Services, and Advanced Services
- Grace periods enable feature testing/trials without buying a license (e.g. 120 days), with some restrictions
- Time-bound licenses allow feature use during sales cycle etc. (after expiration date, feature will continue to run if grace period has not been exhausted)

Base	ISSU	PVRST+	MSTP+	802.1Q	LACP	PVLANs	CDP	SPAN	QoS
	RIP/RIPng	IGMP snooping	DHCP helper	uRPF check	Port Security	SSHv2	RBAC	SNMP	RADIUS
	HSRP	GLBP	VRRP	VRF lite	CoPP	DHCP snooping	DAI	IPSG	802.1x
	Jumbo Frames	UDLD	Storm control	EEM	GOLDs	Call home	NAC	TACACS+	ACLs
Enterprise Services	OSPF	EIGRP	IS-IS	BGP	Graceful Restart	PIM-SM	PIM-Bidir	PIM-SSM	IGMP
	MSDP	PBR	GRE	NetFlow					
Advanced Services	VDCs	CTS							

Note: Enterprise Services is NOT included with Advanced Services license

Nexus 7000 Management: DCNM



Data Center Network Manager (DCNM)

Comprehensive data-center class administration architecture:

- Multi-protocol aware – consolidates and automates Ethernet, IP and Fiber Channel management
- Manages specific data-center network features
- Offers FCAPS coverage for full network service life cycle administration, with emphasis on provisioning, performance and accounting
- Provides management solution for all layers of Cisco-based data center networks

DCNM FCAPS Highlights

Fault, Configuration, Accounting, Performance, Security

Network Discovery & Inventory

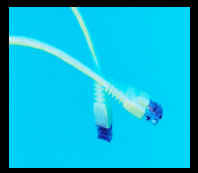
- Automated discovery
- Continuous resynchronization
- Resource management and inventory
- Physical network views

Network Fault Isolation & Diagnostics

- Event collection
- Event filtering
- Diagnostics (w /GOLDS/EEM)
- L2 and IP route topology maps

Performance Monitoring

- Real-time traffic statistics
- Historical trending reports
- Visuals and exporting



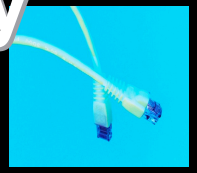
Provisioning

- Model-based configuration creation
- Workflow oriented provisioning wizard
- Configuration validation
- Time-based change request activation
- Network auditing

System & Infrastructure

- Role Based Access Control
- API: Web Services & Java
- Backup & Restore tool
- Image management / ISSU
- Embedded/commercial RDBMS

Ethernet, IP and Network Security Provisioning and Monitoring

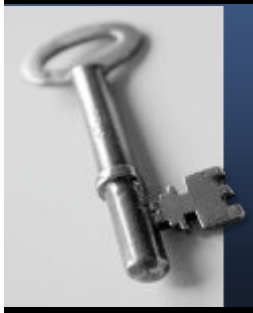


General

- Port
- PortChannel
- SPAN & ERSPAN
- Deep Packet Inspection
- HSRP, VRRP, GLBP
- MCEC
- Hardware Resource Utilization
- Power/Heat Analytics

Layer 2/3

- Port
- VLAN/ Private VLAN
- Spanning Tree
- UDLD
- PortChannel
- LACP
- Multicast (PIM, IGMP, IGMP snooping, MSDP, MLD)
- MRIB/URIB
- VRF
- Route-Map/FilterList
- BGP
- OSPF
- EIGRP
- VRF
- IS-IS



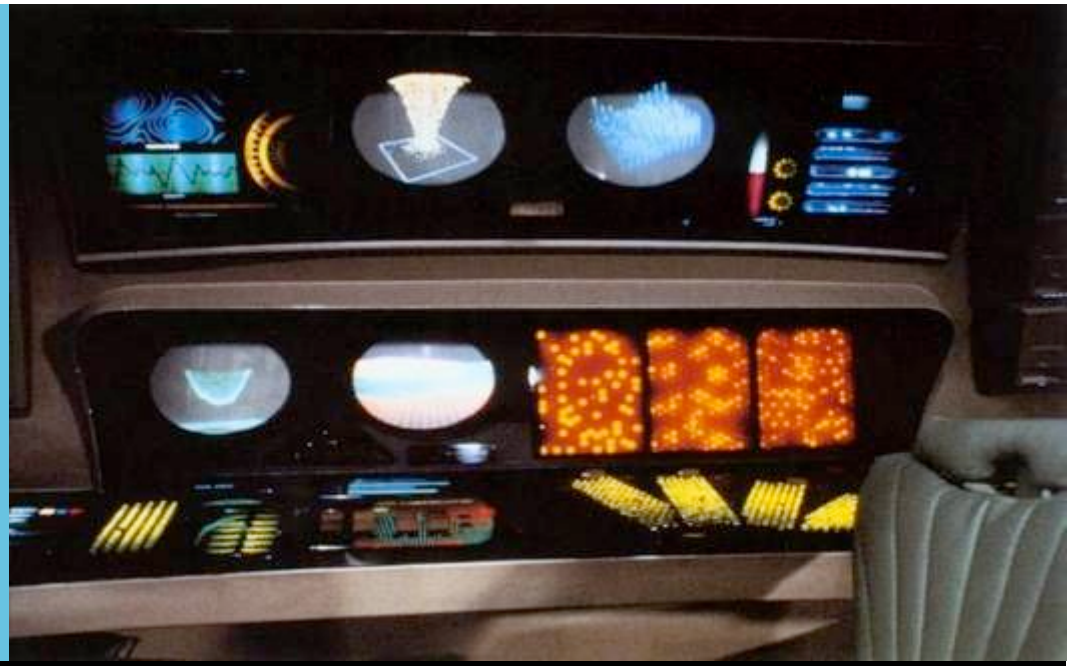
Security

- Port Security
- AAA (Tacacs – Radius)
- Virtualization: Virtual Device Context
- ACL
- Cisco Trusted Security
- 802.1X
- IP Source Guard
- DHCP Snooping
- Dynamic ARP Inspection
- GRE Tunnel
- Unicast / Broadcast / Multicast Suppression
- Network Admission Control

Additional Resources

- More Information
  - http://www.cisco.com/go/datacenter
  - http://www.cisco.com/go/nexus
  - http://www.cisco.com/go/nxos
  - http://www.cisco.com/go/dcnm

Nexus 7000: Product Positioning



Catalyst and Nexus: Complementary Focus

Cisco® Nexus 7000

15 Terabit Scalability
Unified Fabric

100GbE

40GbE

Transport Flexibility

Operational Continuity

15T

7.5T

3.7T

10GbE

1GbE

720G

2T

Cisco Catalyst® 6500

2 Terabit Scalability
Integrated Services

Nexus Launch News Coverage Reflects Cisco Messaging

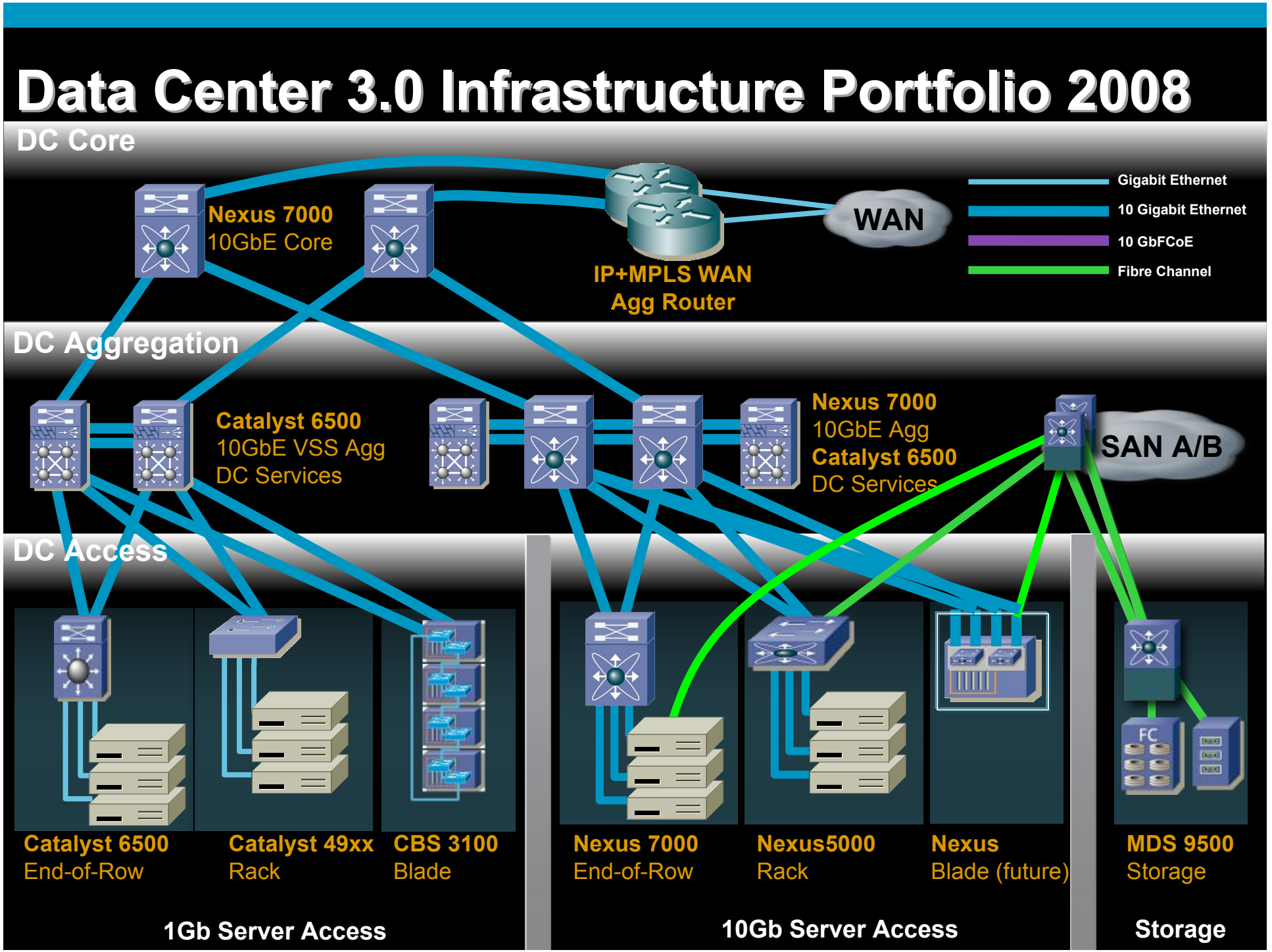
Message #4: Nexus does not replace Catalyst, is not disruptive, and fits into existing data center infrastructure. Cisco is committed to the Catalyst and announced Catalyst enhancements also.

"Nexus's arrival doesn't mean Catalyst and the storage-switching MDS line are doomed. Cisco says it has plans for all three in "Data Center 3.0," a phrase the company has been tossing around since August."-- **Byte & Switch, January 28, 2008**

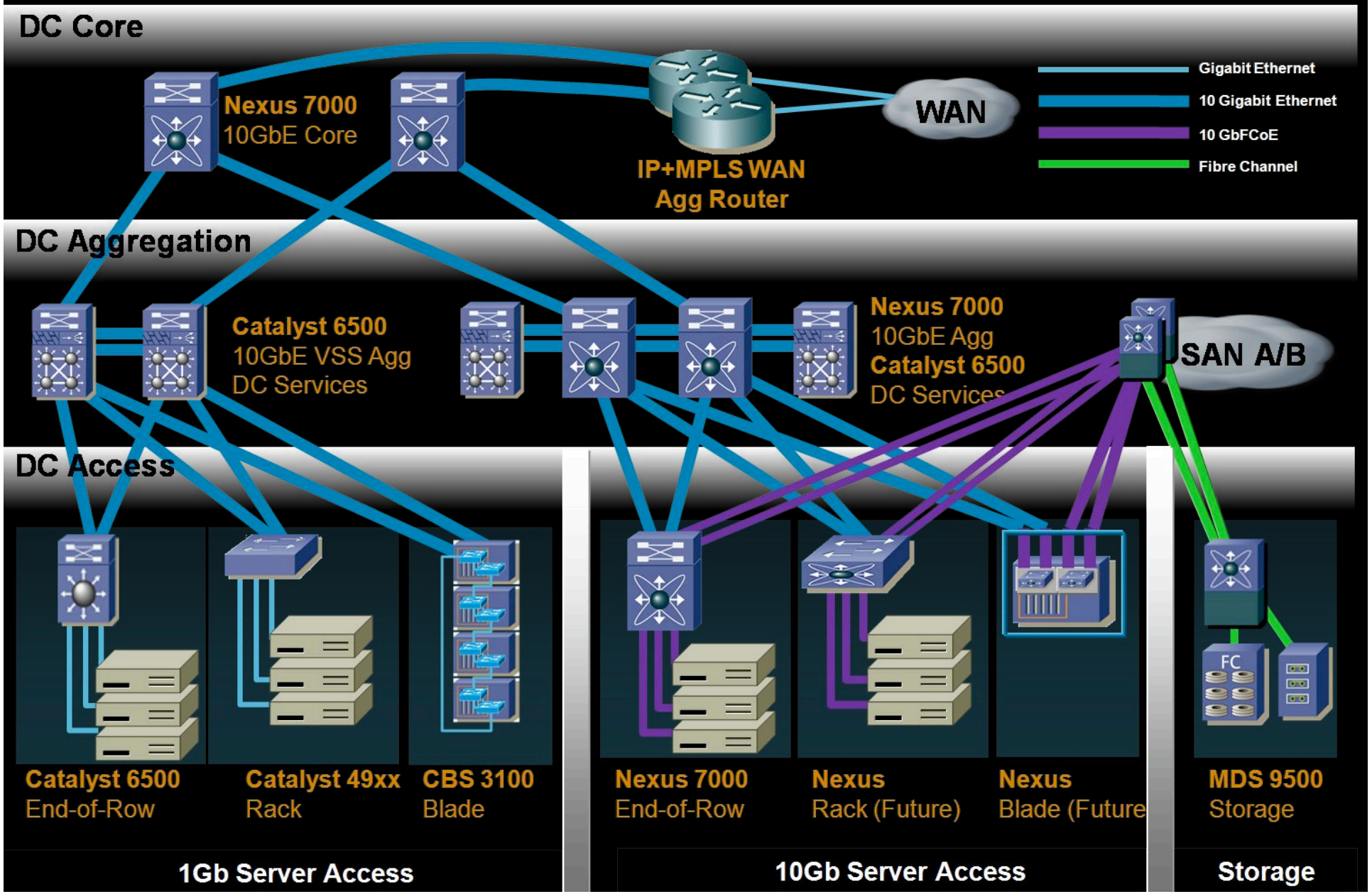
"Meantime, Cisco isn't abandoning the Catalyst 6500, though it's now describing it as a service switch. The company last week unveiled upgrades to two of the Catalyst 6500 blades, boosting Power-over-Ethernet wattage so the wireless blade can support 802.11n access points, and giving its Wide Area Application Services a software client for accelerating WAN traffic to mobile devices. Cisco plans to keep adding service blades to the Catalyst, though it probably won't go beyond 10 Gbps. The higher speeds will be reserved for the Nexus at the data center core."
-- *InformationWeek*, January 28, 2008

"Nexus's arrival doesn't mean Catalyst and the storage-switching MDS line are doomed. Cisco says it has plans for all three in "Data Center 3.0," a phrase the company has been tossing around since August…….Catalyst keeps kicking. Amid the hoopla about Nexus, Cisco is being adamant that the box doesn't replace Catalyst or MDS…..It's easy to believe the company for now, because Cisco's plans don't call for Nexus to ever support intelligent applications such as firewalls or load balancing. Those have been left to Catalyst.…Moreover, Nexus isn't intended for every type of data center. 'You don't need Nexus functionality everywhere. Catalyst makes a great edge switch or smaller data center,' Kerravala says."
--Light Reading, January 28, 2008

Data Center 3.0 Infrastructure Portfolio 2008



Data Center 3.0 Infrastructure Portfolio



Nexus 7000 and Catalyst 6500
Technology Linkages

L2 Scalability	Pervasive Security	Services Integration	Unified Fabric	System Scalability	Operational Continuity
MultiChassis Etherchannel	CTS – Secure Group ACL	Embedded Deep Packet Inspection (DPI) Services	Unified Fabric and SAN/LAN Interconnect	Upto 15 Terabit Switching w/ 40G/100G interfaces	Modular OS, Multi-layered resiliency, VDCs
L2 MultiPathing	Cisco TrustSec	Service Integration	Unified Fabric	40G/100G	Software
Virtual Switching	CTS – MultiHop SXP	Integrated Services cards; PISA (DPI & NBAR)	Uplinks w/ Multipathing	2 Terabit Switching w/ 40G interfaces	Modular OS; SSO across two switches