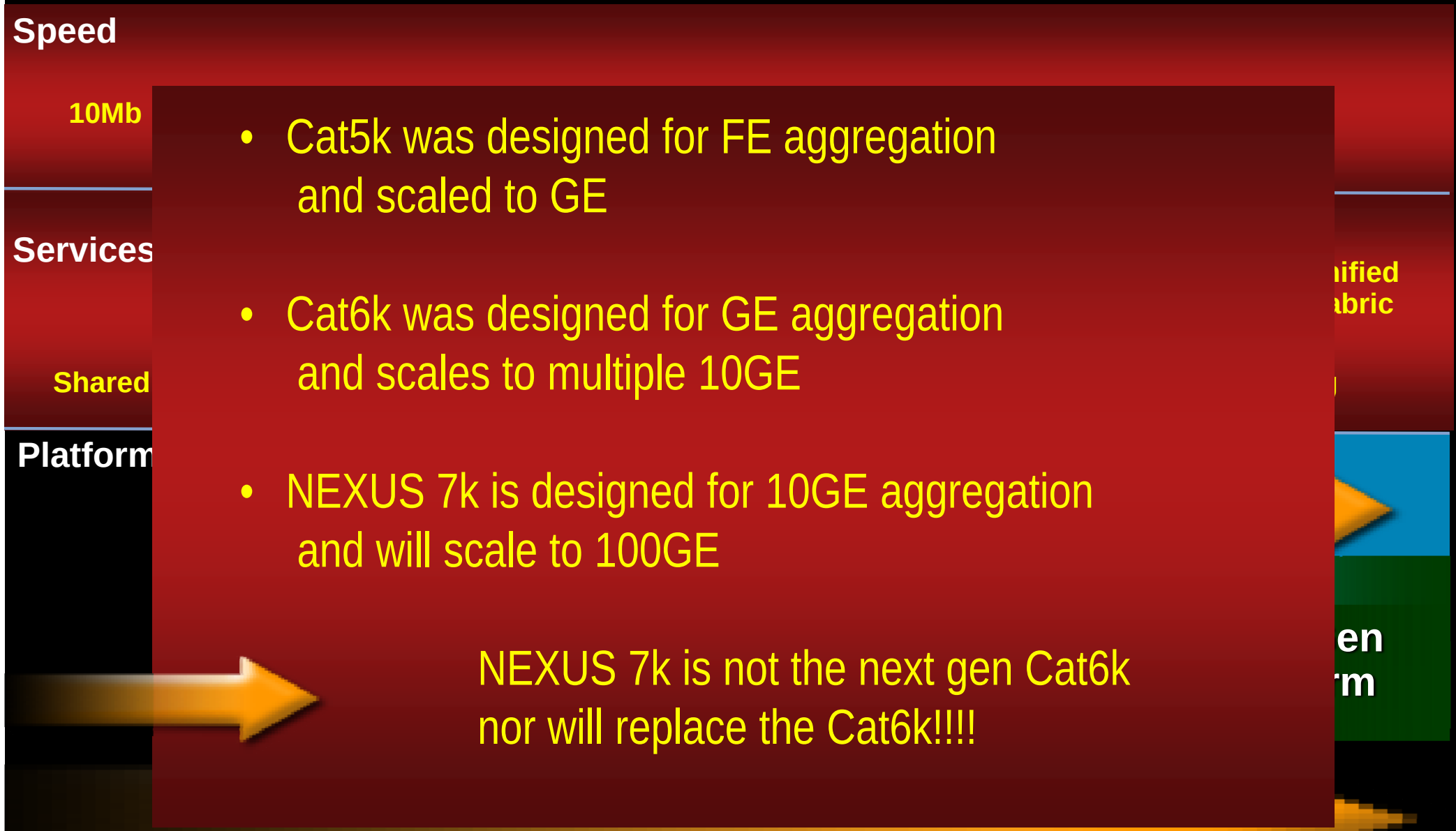
Cisco Expo 2008

Data Center:

NEXUS 7000 Switching-Platform für zukünftige Technologien

Wolfgang Riedel
Consulting System Engineer
Technical Marketing - Data Center
CCIE #13804
wr@cisco.com

Markets Transition To Meet New Needs



- Cat5k was designed for FE aggregation and scaled to GE
- Cat6k was designed for GE aggregation and scales to multiple 10GE
- NEXUS 7k is designed for 10GE aggregation and will scale to 100GE

NEXUS 7k is not the next gen Cat6k nor will replace the Cat6k!!!!

NEXUS and Catalyst Strategic Roadmap

Shared Innovation

L2 Scalability	Pervasive Security	L4-7 Services Integration	Data Center Ethernet (DCE)	Scalability
MultiChassis Etherchannel	CTS – Secure Group ACL	Embedded Deep Packet Inspection Services	Core Unified Fabric and SAN/LAN Interconnect	Upto 15 Terabit Switching w/ 40G/100G I/Fs
MultiChassis Etherchannel	Cisco TrustSec	Service Insertion Arch.	Data Center Ethernet	40GE/100GE
Virtual Switching	CTS – MultiHop SXP	L4-7 Integrated Services	DCE light Uplinks w/Multipathing	2 Terabit Switching w/ 40GE Uplinks

CY08 CY09 CY10+

Nexus 7000: First in Class

$1B in Data Center R&D
100+ Patents

15+ Tbps Lossless Fabric
500+ Gbps slot capacity

Designed for continuous system operations

Engineered for manageability and serviceability

Transport Flexibility: 1/10/40/100G Ethernet and Unified I/O Capable

Nexus 7000 Series Switches
Future Proofed Hardware

High Availability Design
- Hot swappable components
- Redundant common equipment

Physical Characteristics
- 8 Payload and 2 Supervisor Slots
- 21RU per chassis (2 per 42RU rack)
- Three 6KW PS - dual 20A inputs

Environmental Characteristics
- Front-to-Back airflow
- 9KW Maximum System Power

Operational Manageability & Serviceability
- Integrated Cable Management
- System & Module Locator LEDs
- Lockable Doors
- Easy access for common equipment

Highest Density for scalable designs
- 32 Port 10GE I/O Module
- 48 10/100/1000 I/O Module

NX-OS

The NeXt step in operating systems, designed to meet the operational needs of the Data Center, based on industry proven SAN-OS

Data Center Focused, Feature Rich Operating System

Intelligent IOS-like CLI

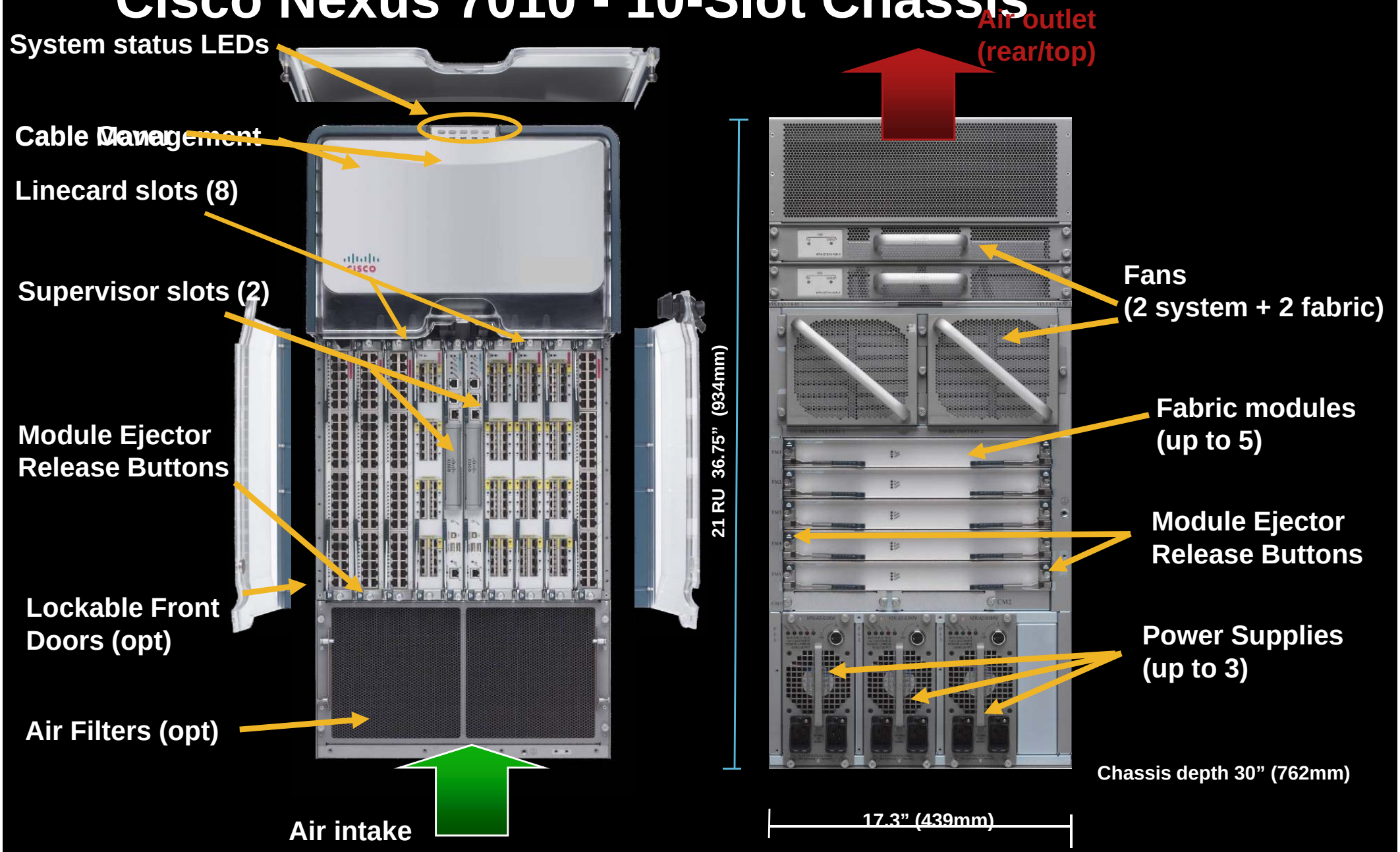
Innovative
- Tool to I/O Unification
- Truly Virtualized OS
- Cisco TrustSec

Resilient
- Mission critical environments
- Industry Leading ISSU
- Self Healing

Manageable
- Synergistic Management Tools
- Extensive Built-in Diagnostics

Nexus 7000 Hardware

Cisco Nexus 7010 - 10-Slot Chassis

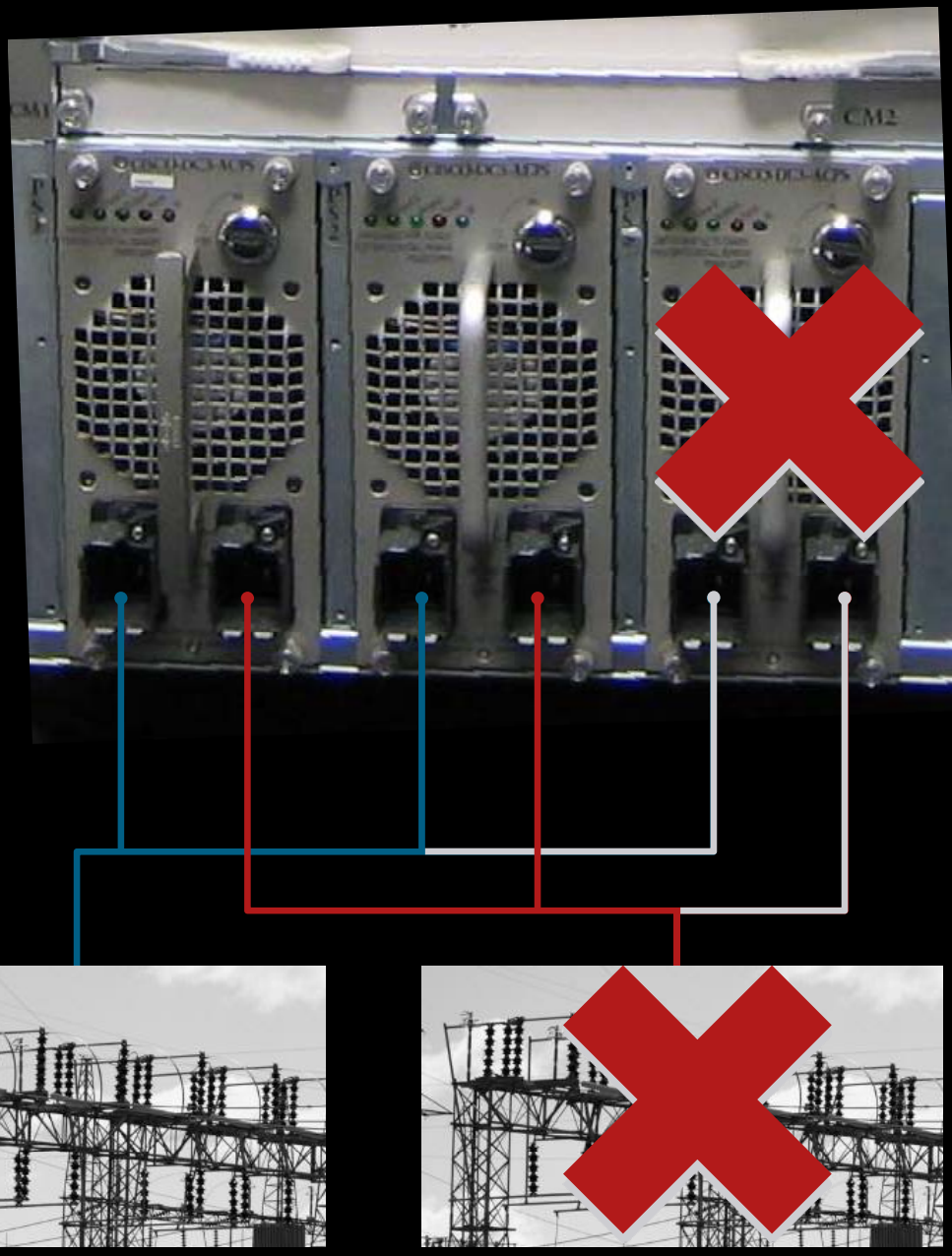


Front and Rear of 10 Slot Chassis

Power Redundancy

Three power redundancy modes:

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



N+1 redundancy

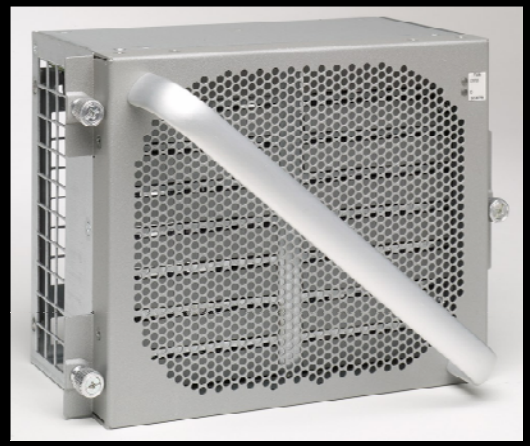
Grid redundancy

System Cooling

- Variable speed redundant fans provide complete system cooling
- Fans removed from chassis rear – no disruption of cabling
- Blue Indicator ID LED for easy location

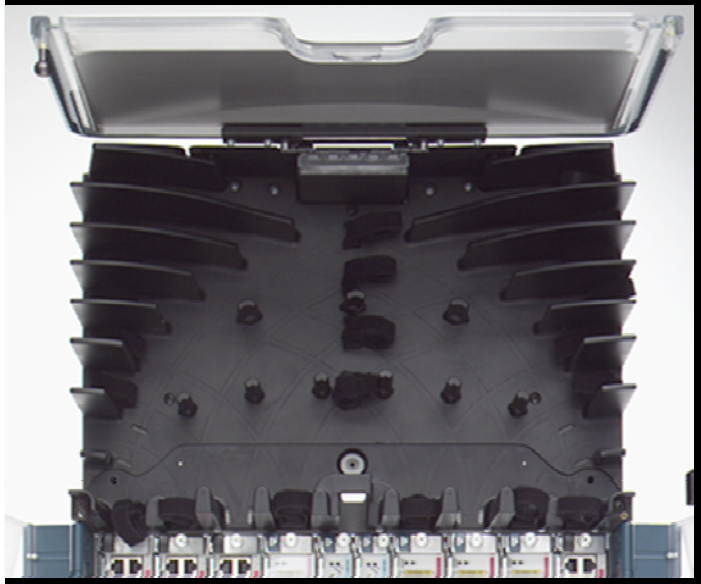
- Redundant system fan trays provide cooling of I/O modules and supervisor engines
- 6 fans per tray
- Hot swappable

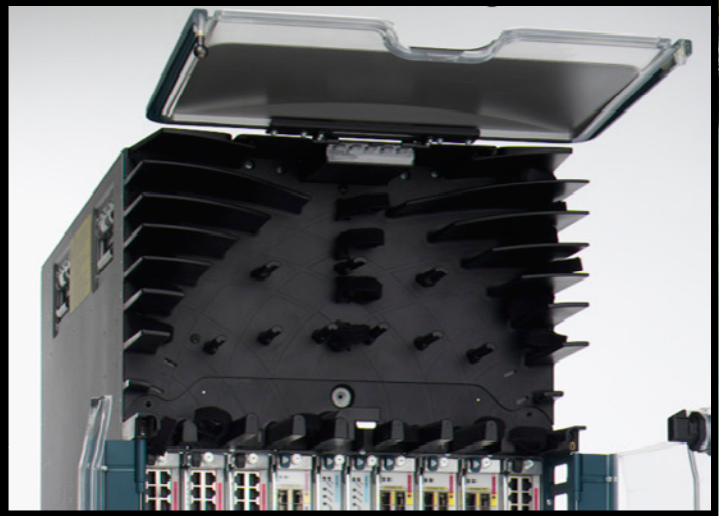
- Redundant fabric fan trays provide cooling of crossbar fabric modules
- One fan per tray
- Hot swappable



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
- Cable tray cover and lockable front doors prevent accidental interference





Other Hardware Features

Locking ejector levers ensure proper module seating and prevent accidental disengagement

Depressing both ejection buttons will start a module power down

System LEDs provide summary of system status

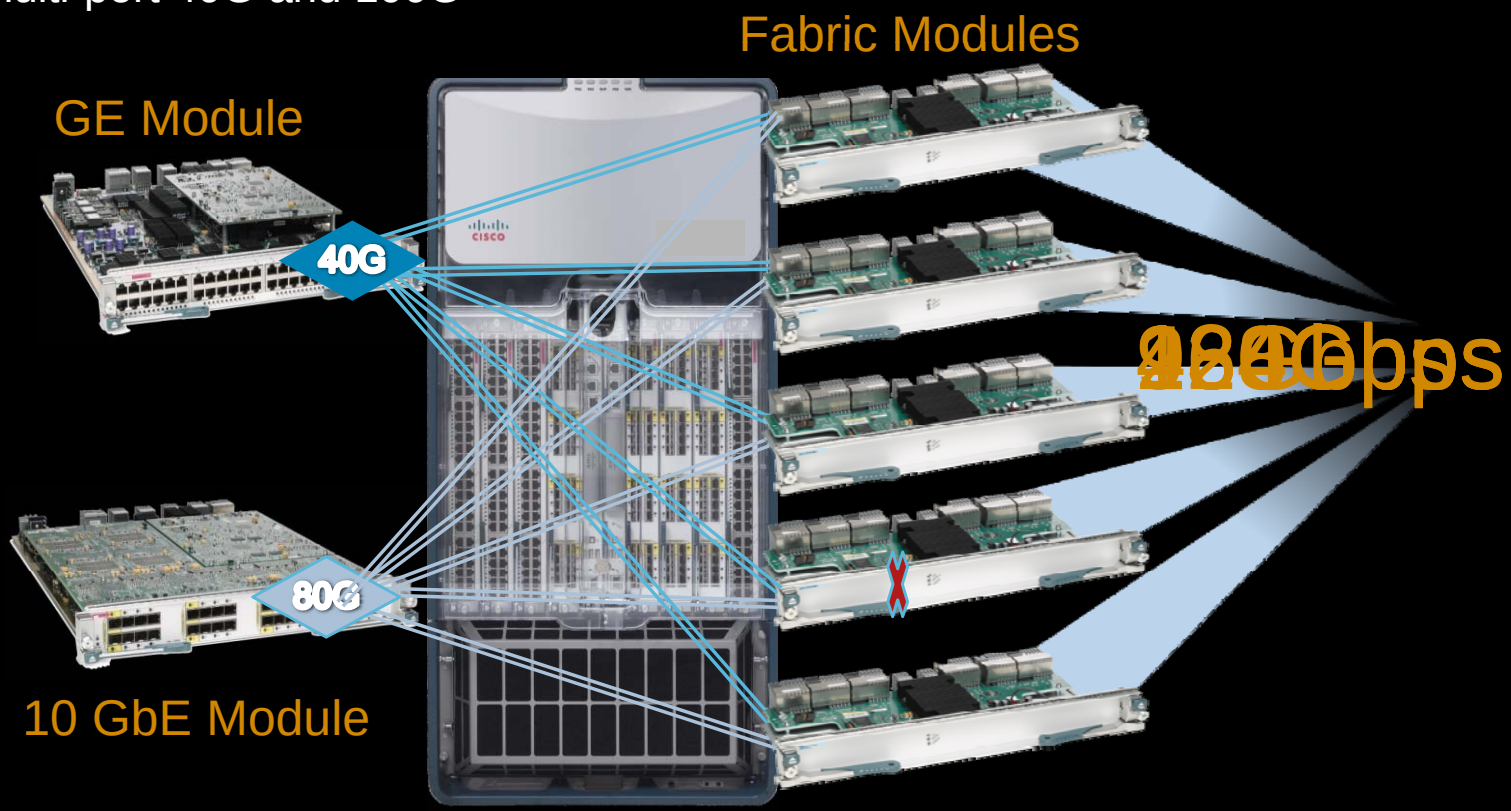
- Supervisor engines
- I/O modules
- Fabric modules
- Power supplies
- Fan trays

Indicator ID on I/O Modules, Supervisors, Power Supplies, Fabric Fan Trays and System Fan Trays controlled from software to allow remote operators to indicate to maintenance staff the chassis and part

Fabric Design and Capacity
Designed to handle tomorrow's bandwidth needs

- Per-slot bandwidth capacity increases with each fabric module
- Graceful Fabric failure/removal
- Leverages Virtual Output queues and central arbitration scheme for 'lossless' characteristics
- Day one Fabric Modules provide investment protection for future higher speed modules – scaling to multi-port 40G and 100G

Fabric module failure handled with controlled predicable reduction of capacity



Supervisor Engine 1

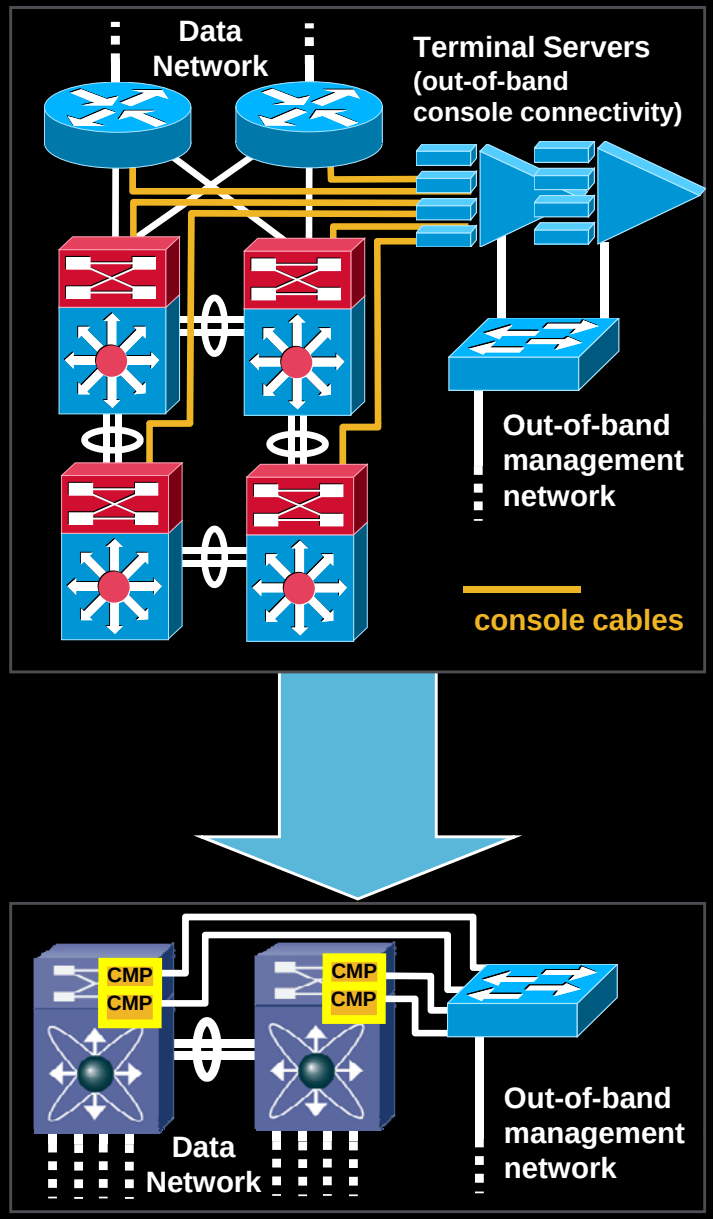
Simple Design

- No uplinks, no forwarding, no fabric.
- 100% mgmt & control plane
- Built for the future

Key Features

- High-performance dual-core 1.66GHz Intel Xeon processor
- Connectivity Management Processor (CMP) for lights-out management
- Separate 10/100/1000 management port with 802.1ae LinkSec
- Blue beacon LED for easy location

Connectivity Management Processor (CMP)



- Standalone, “always on” management processor on supervisor engine
- Provides ‘lights out’ remote management connectivity via 10/100/1000 Ethernet interface
- Provides ability to monitor supervisor and module console ports, access log files, power cycle supervisor engine, etc.
  - Runs lightweight Linux kernel and network stack
  - Completely independent of DC-OS on main CPU
- Removes need for separate terminal servers for out-of-band management
  - Direct Ethernet connection from management network switch to CMP management interfaces

Nexus 7000 I/O Modules (Classical Ethernet)

32 Port 10G (80G)
SFP+ Optics

48 Port 10/100/1000 (40G)
RJ-45

Overview

- **Integrated** 60Mpps forwarding engine for fully distributed forwarding
- Virtual output queueing (VOQ) ensuring fair access to fabric bandwidth
- 802.1ae LinkSec on every port
- Blue Beacon LED for easy location

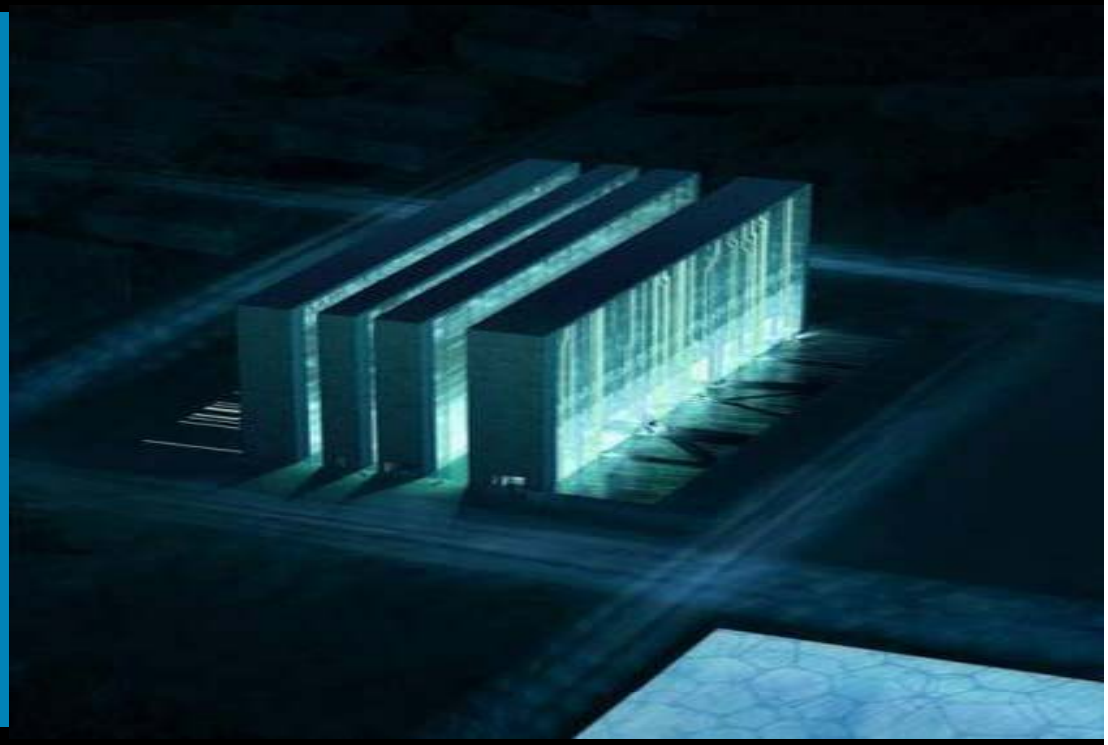
Table Sizes Optimized for the DC.

- 128K FIB, 128K MAC, 512K Netflow, & 64K ACLs

Features Highlights

- 16 Way ECMP for large scale designs
- Secure Group ACLs
- Security Enhancements (for CoPP, uRPF, IDS Checks)
- Netflow Sampling & TCP Flag Export

NX-OS Software



Cisco Network Operating Systems

The right tools in the right places

Cisco IOS XR

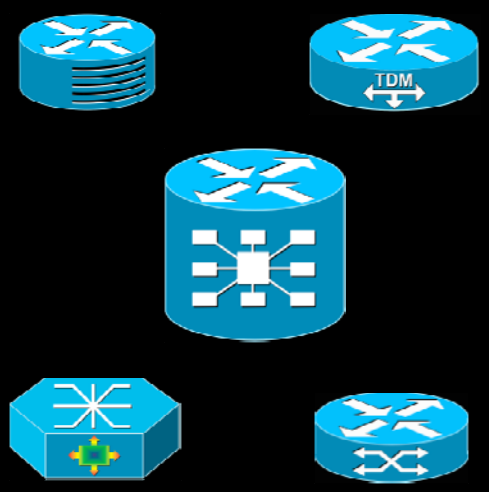
Core WAN focused

SP grade services

State of the art resiliency

Large scale networks

WAN-core link-layer types



Cisco IOS

Ubiquitous

Unmatched feature set

Unprecedented flexibility

Robust resiliency

LAN/WAN link-layer types

Cisco NX-OS

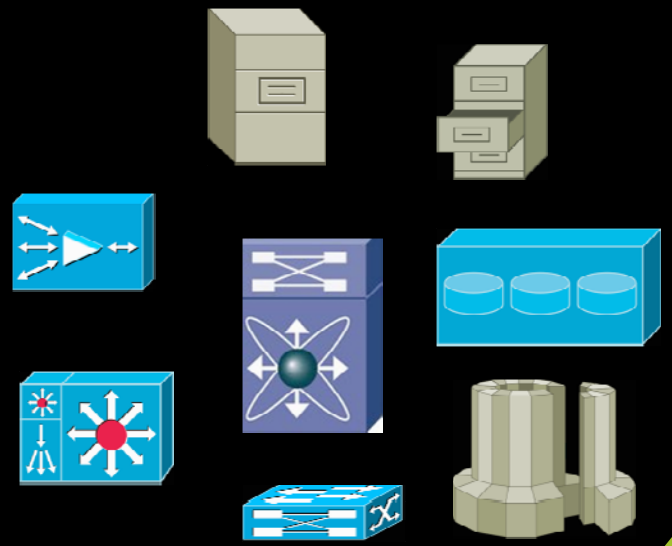
Data Center focused feature set

Mission critical environments

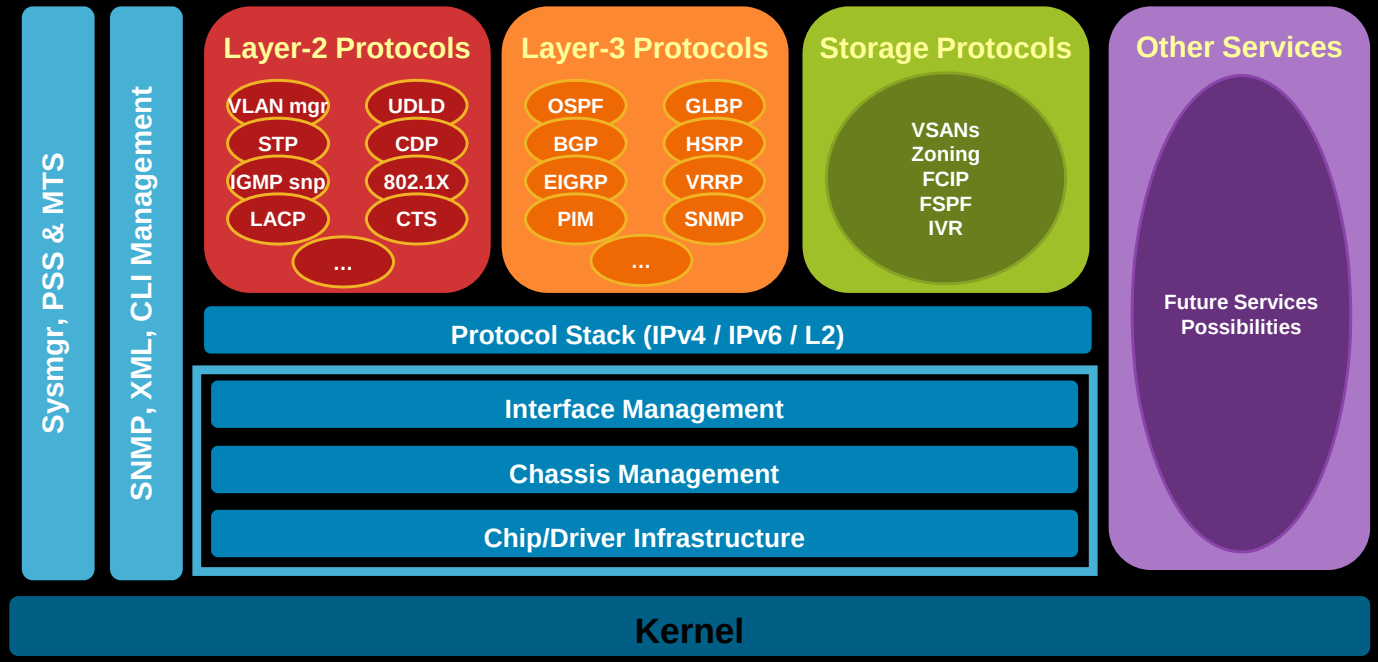
24x7 Continuous operations

High density / performance Ethernet

DC specific link-layer types



NX-OS
Software Architecture

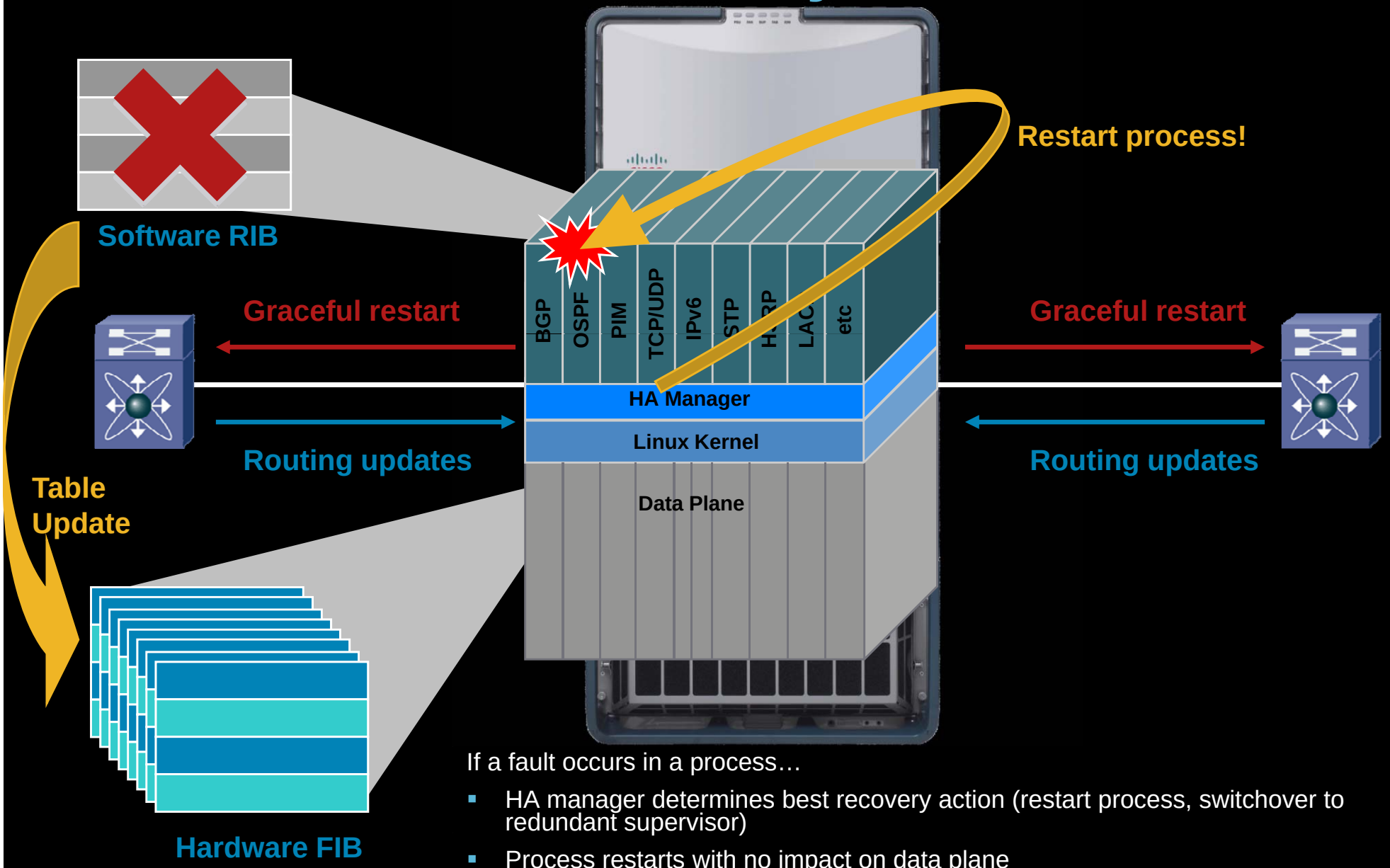


- **'Next-generation' operating system that brings 3 fundamental technologies into a single platform:**
  - **Layer-2** classical (now) and **unified I/O** switching (future)
  - **Layer-3** multi-protocol routing (now)
  - **Storage protocols** and **SAN switching** (future)
- **Separation of control-plane & data-plane**
  - **Complete separation** of control-plane and data-plane
  - **Multi-threaded & modular**
  - **The majority of computationally-intensive hardware/table programming tasks are offloaded** to linecard-local control-plane CPUs
  - **Modular code with real-time preemptive scheduling**
- **Modularity**
  - **Protocols, table managers and different subsystems all run as distinct memory-protected restart able processes**
  - **Non-Stop Forwarding for all of Storage/L2/L3**
  - **HA infrastructure built day-1**
  - **Granular modularity from day-1**

Nexus 7000 / NX-OS High Availability

- Hardware provides redundancy at every component level:
  - Supervisors
  - Fabrics
  - Power
  - Fans
- Software offers multi-layered, multi-faceted resiliency:
  - Stateful process restarts (PSS)
  - 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

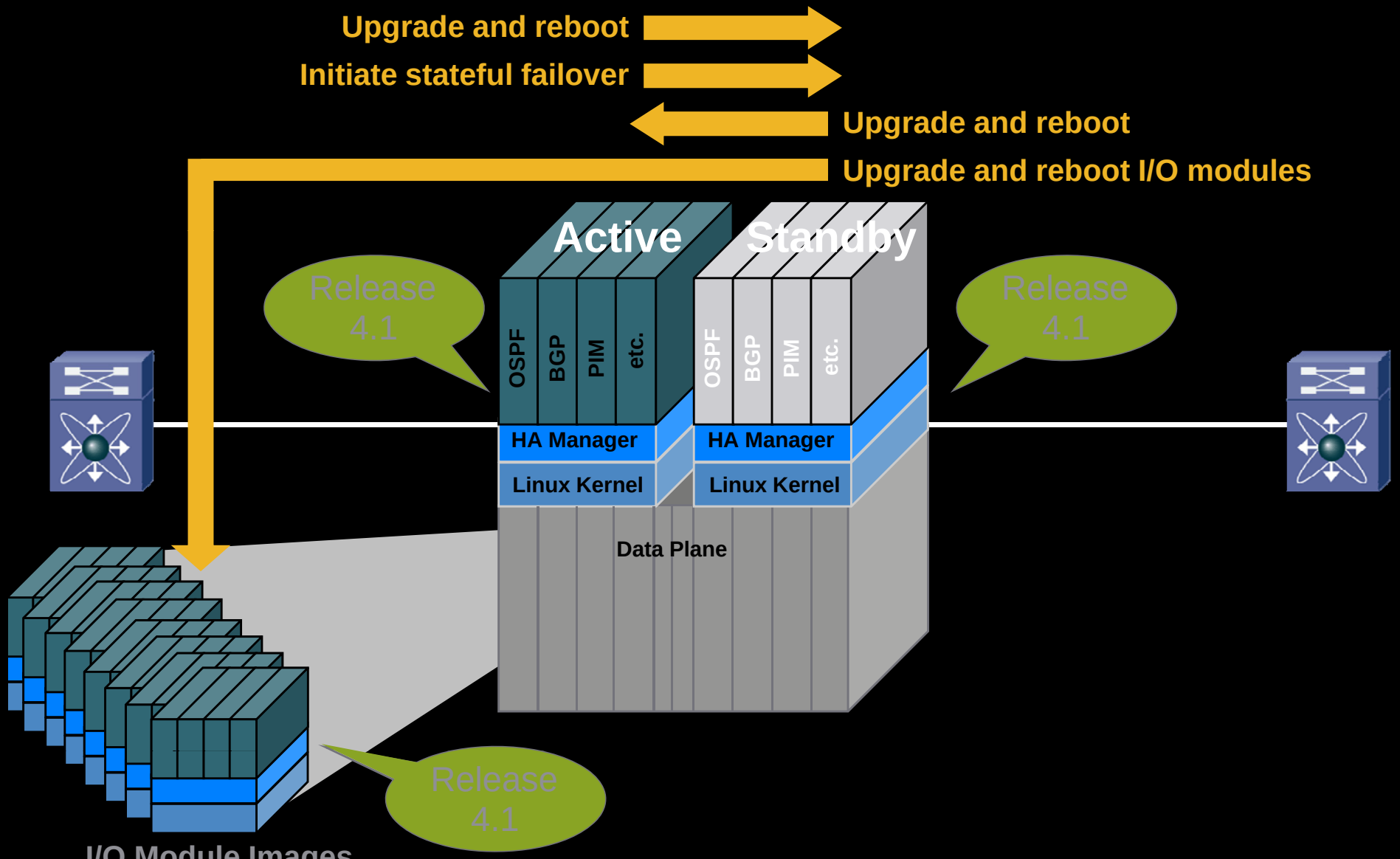


If a fault occurs in a process…

- 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

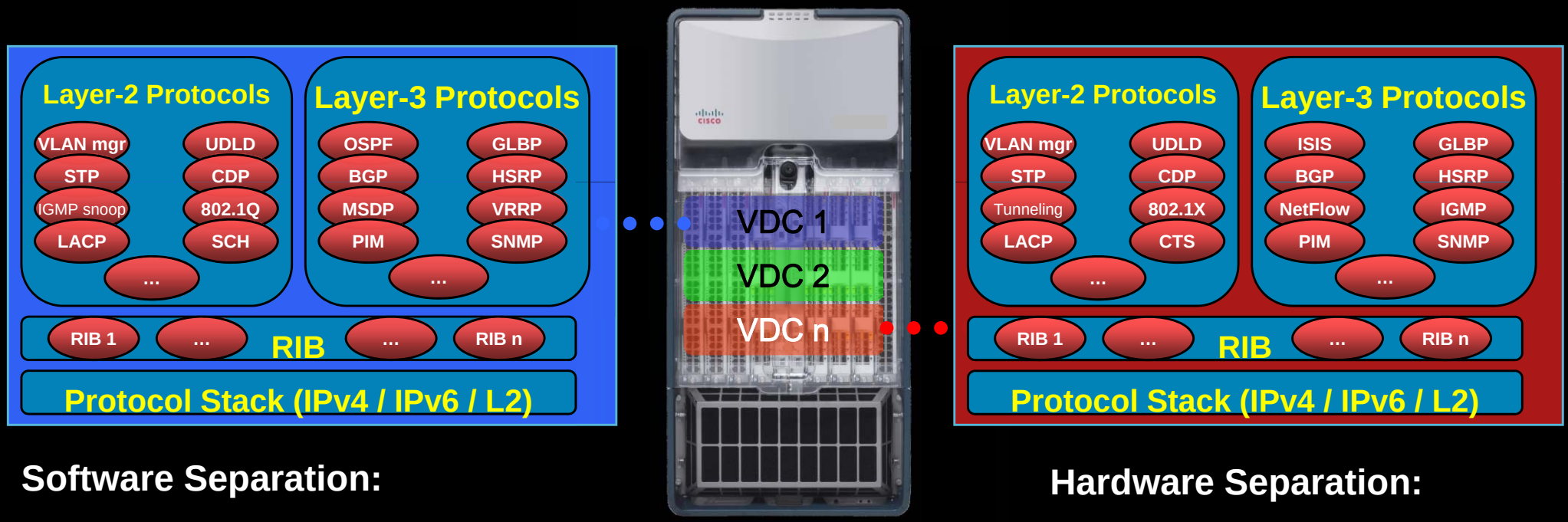
In-Service Software Upgrade

```
dc3# install all kickstart bootdisk:4.1-kickstart system bootdisk:4.1-system
dc3#
```



Virtualize the network
Virtual Device Context (VDC)

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:

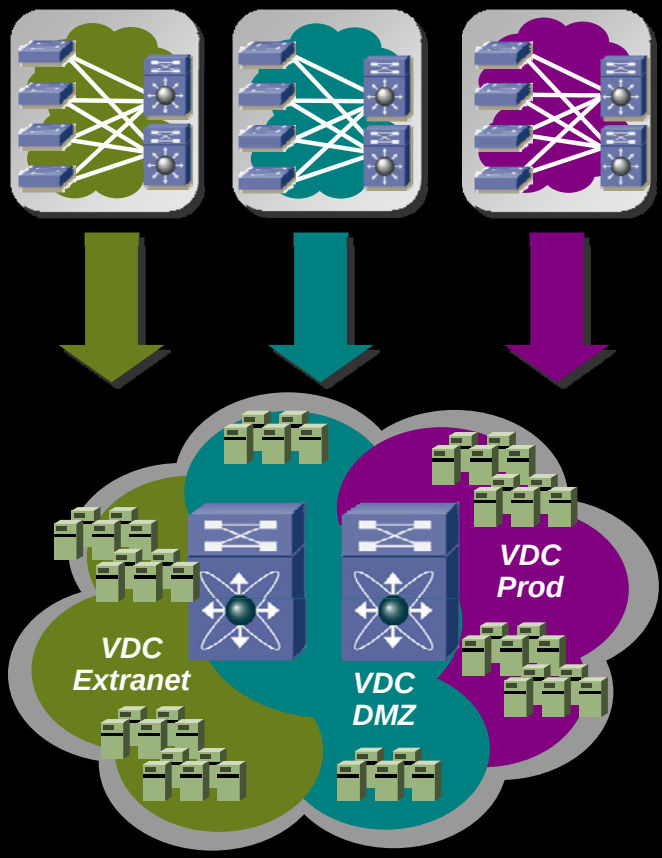
- Software Infrastructure
- Kernel
- Power Supplies
- Fans
- Chassis

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.
- Helps scale physical resources of device:

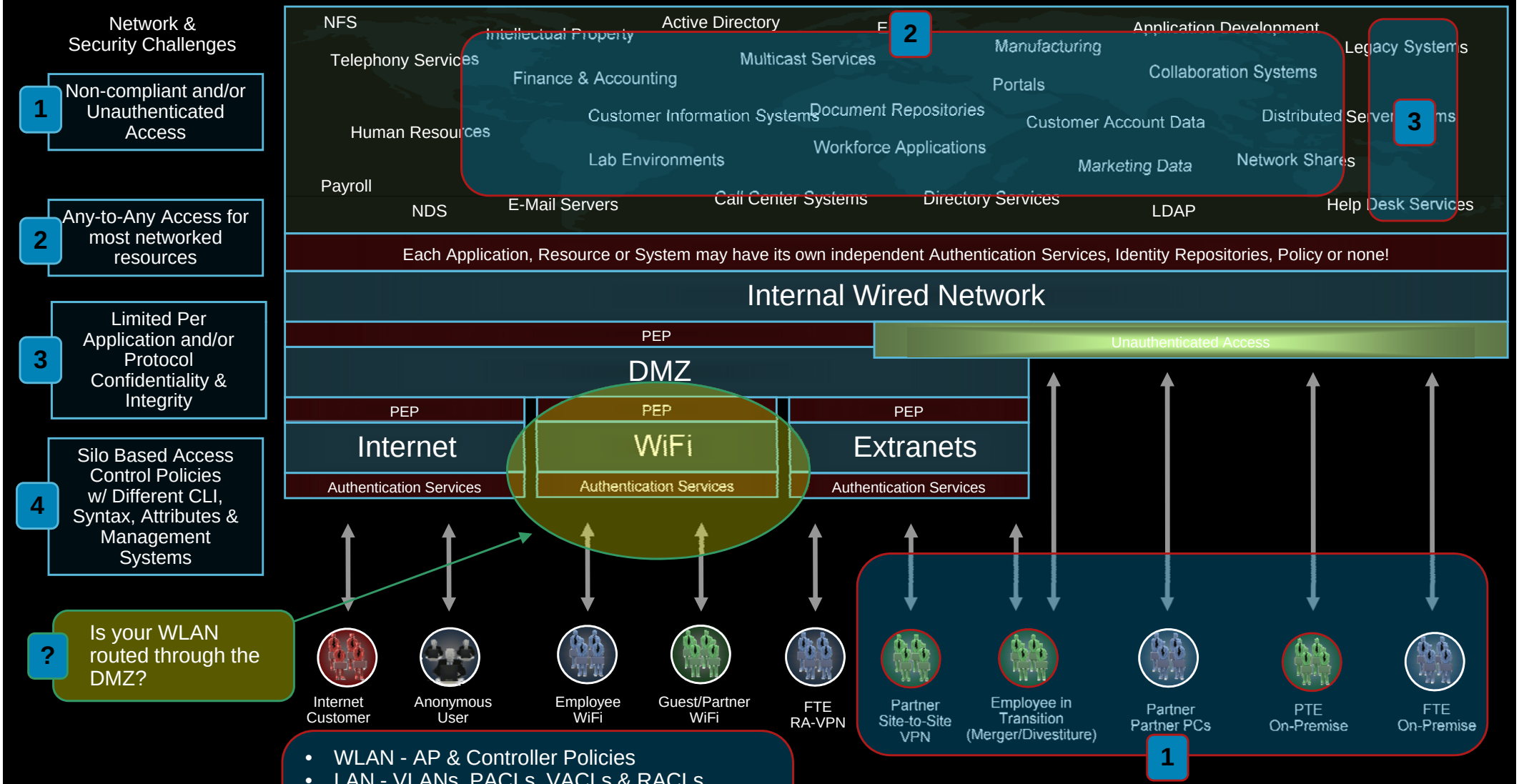


Different network islands *virtualized* onto common data center networking infrastructure

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 * 128K)
MAC table	128K	128K	1M (8 * 128K)
Classification TCAM	64K	64K	512K (8 * 64K)
Policers	16K	16K	128K (8 * 16K)

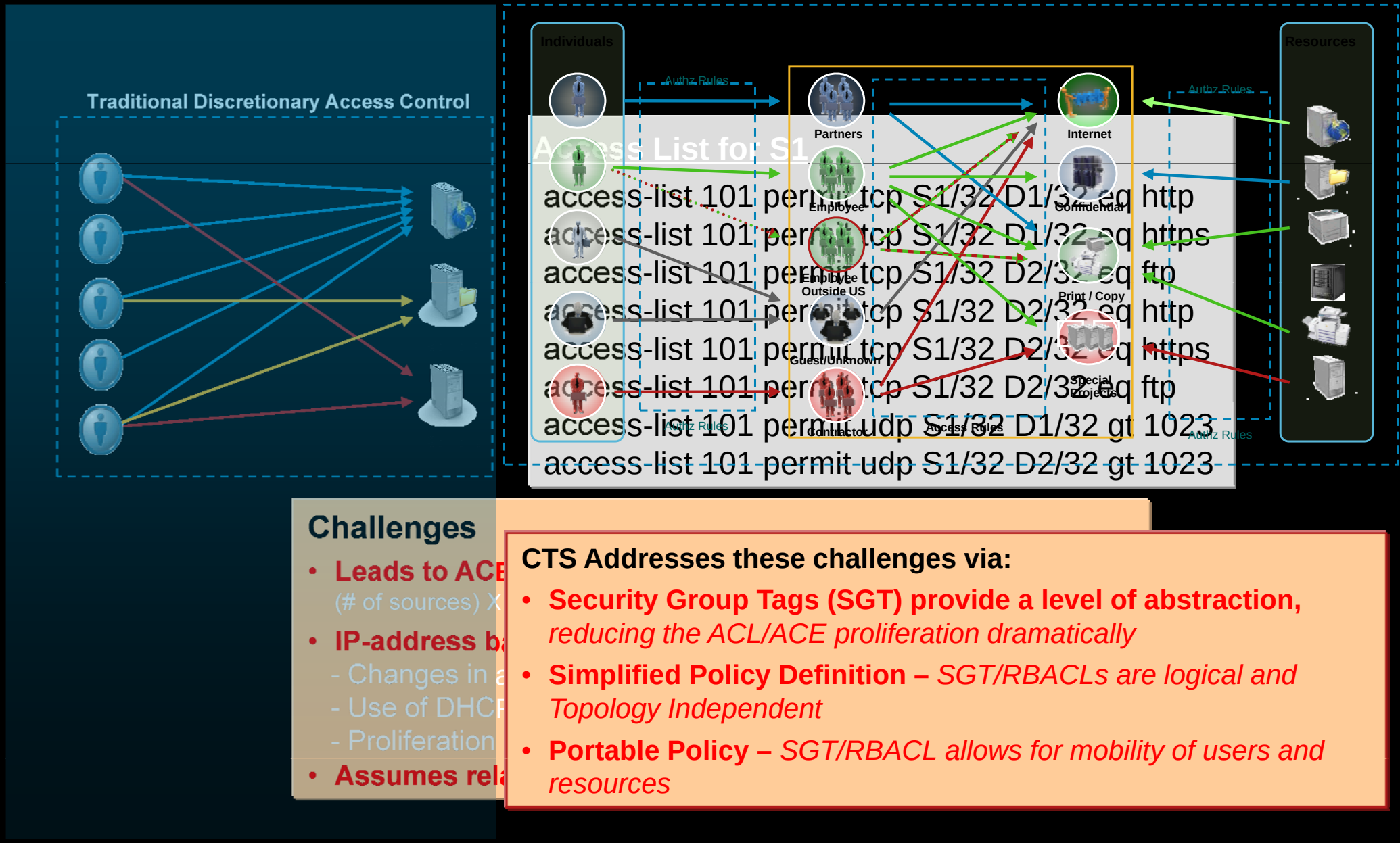
Cisco TrustSec on Nexus 7000

Why TrustSec?



Why Security Group Tags

Traditional ACLs vs. CTS Security Group Based Access Control



Challenges

- **Leads to ACE**
  (# of sources) X
- **IP-address ba**
  - Changes in a
  - Use of DHCP
  - Proliferation
- **Assumes rela**

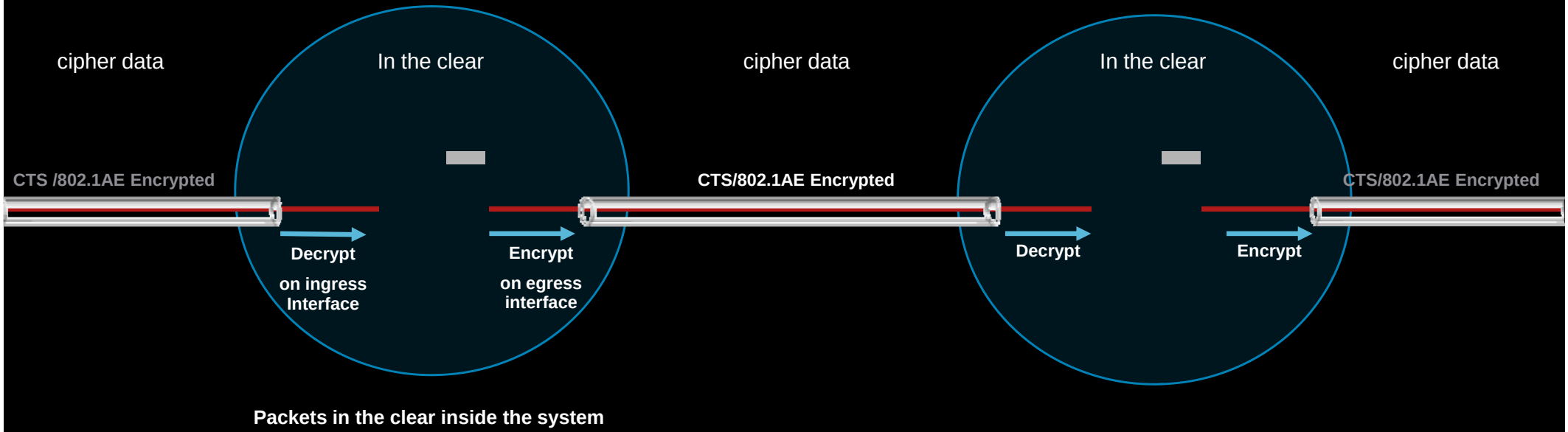
CTS Addresses these challenges via:

- **Security Group Tags (SGT) provide a level of abstraction,** *reducing the ACL/ACE proliferation dramatically*
- **Simplified Policy Definition –** *SGT/RBACLs are logical and Topology Independent*
- **Portable Policy –** *SGT/RBACL allows for mobility of users and resources*

Cisco TrustSec Link-layer encryption

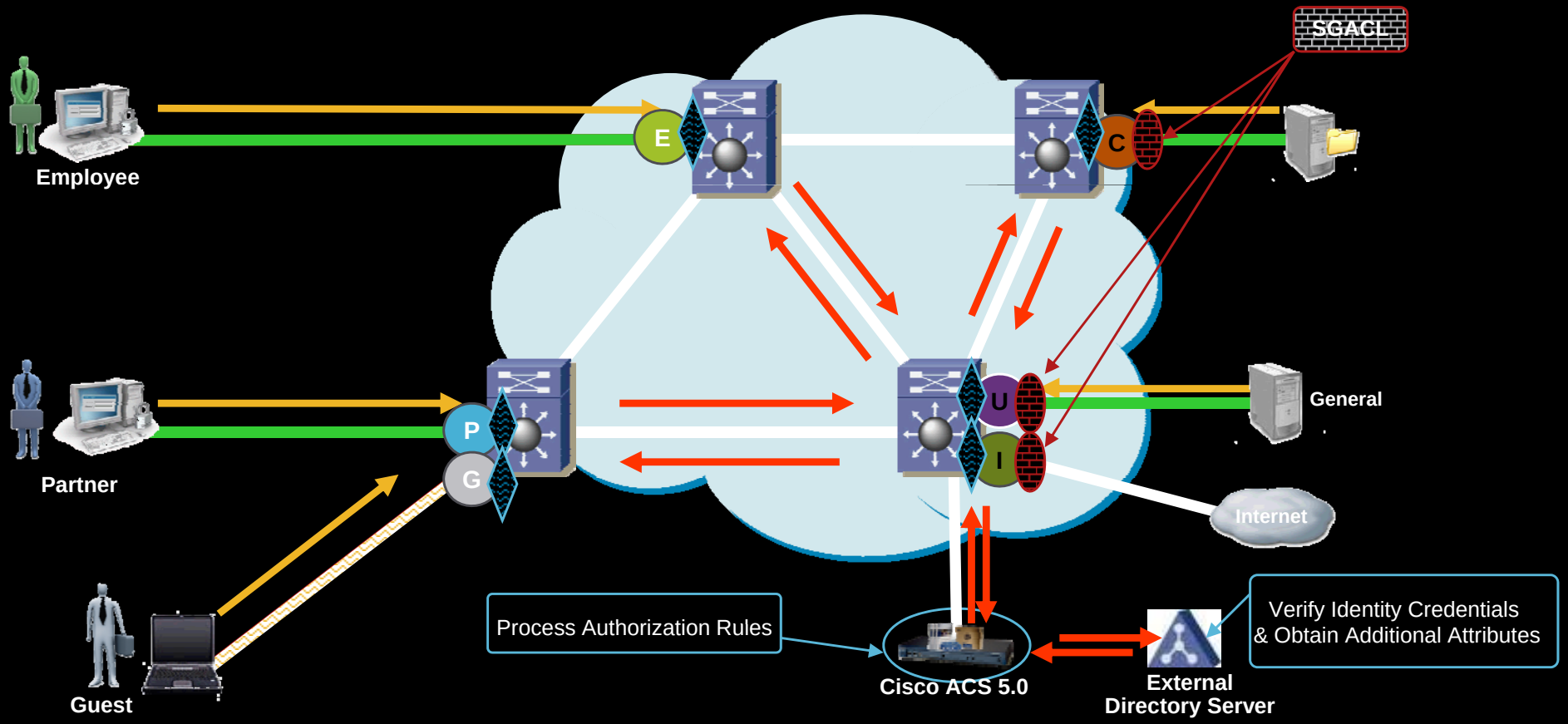
Hop-by-Hop Packet Confidentiality and Integrity via IEEE 802.1AE

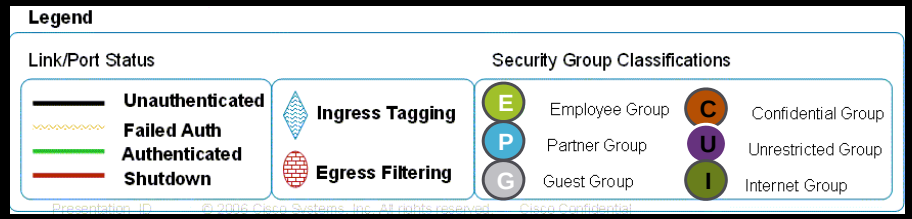
- "Bump-in-the-wire" model
  - Packets are encrypted on egress
  - Packets are decrypted on ingress
  - Packets are in the clear in the device
- Allows the network to continue to perform all the packet inspection features currently used
- Can be incrementally deployed depending on link vulnerability
- **DC3: Wire-rate link-layer encryption on every 10/100/1000/10GbE port**



wr@cisco.com

CTS – Network Admission Control





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

Control Plane Policing (CoPP)

- Prioritizes important control plane traffic and protects supervisor from DoS attacks
- Follows MQC model, with service-policy applied to “control-plane” interface
- Provides granular classification, marking, and rate control for control-plane bound packets

Receive packets	Broadcast MAC + non-IP packets
Multicast packets	Broadcast MAC + IP packets
Exception packets	Mcast MAC + IP packets
Redirect packets	Router MAC + non-IP
ARP packets	

NetFlow Key Points

- System Scalability. Up to **~500K** (with 95% utilization efficiency) cached flows for Forwarding Engine
- Sampled NetFlow. Effective hardware-based sampling to improve and preserve NetFlow table utilization
- Egress NetFlow and Bridged NetFlow
- TCP Flags are now exported as part of the flow information
- Export version 5 (the most used) and export version 9 (the most flexible) are both supported
- VRF aware export
- Hitless ISSU and process restartability
- Flexible NetFlow CLI look & feel

Policy Based Routing

- PBR forwards packets to next-hop device based on administrative policy rather than best routing metric
- Leverages CL TCAM to classify traffic for policy routing
- Supported on any Layer 3 interface type (including tunnels, subinterfaces, SVIs)
- Supports load-sharing among PBR next hops
- Matches based on:
  - IPv4 or IPv6 access-list
  - Packet length
- Set defines:
  - IPv4/IPv6 next-hop or default next-hop, with load-sharing
  - VRF

EtherChannel

- 1G and 10G EtherChannel
- Layer 2 (access or trunk) or Layer 3 (routed or with subinterfaces)
- Up to 8 active member ports per EtherChannel
- Up to 256 EtherChannel interfaces per system
- LACP for channel negotiation
- Variety of load-sharing options:
  - Source, destination, source + destination MAC
  - Source, destination, source + destination IP (default)
  - Source, destination, source + destination TCP/UDP ports
- Per-module load-sharing configuration

SPAN

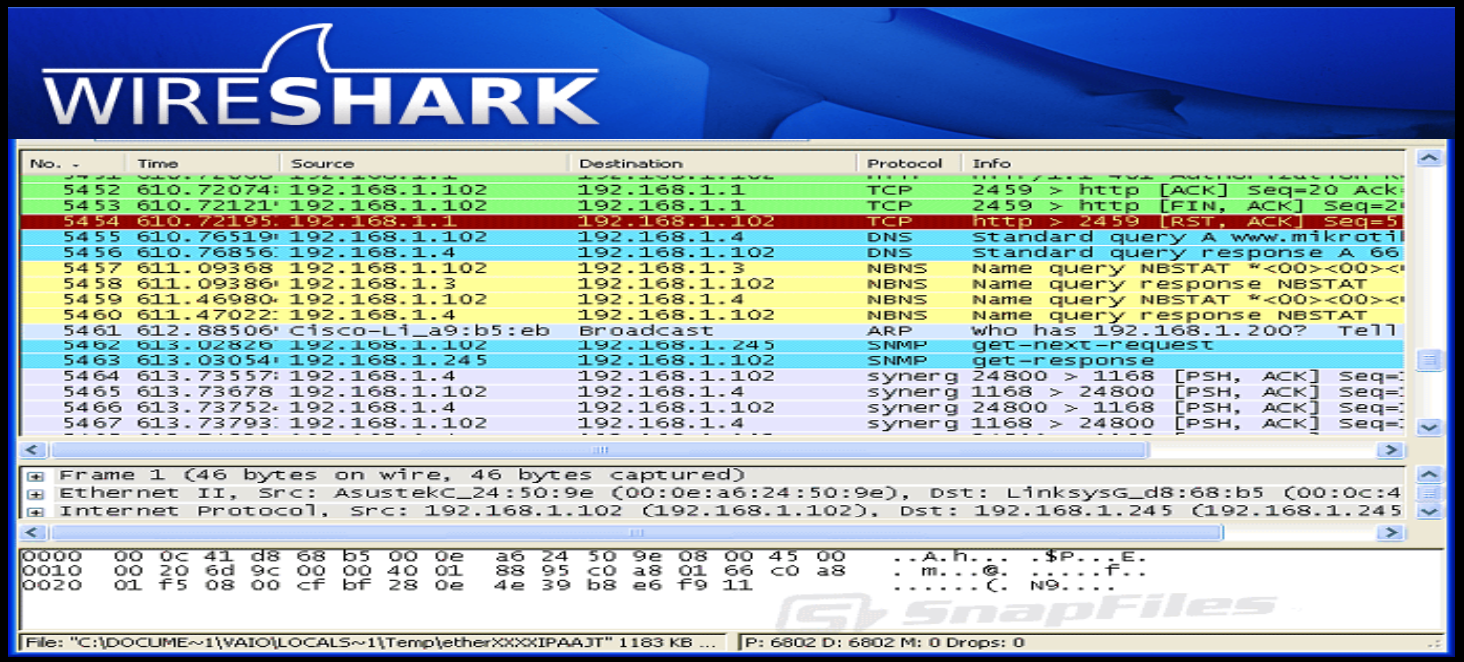
- 2 bidirectional SPAN sessions
- Configuration through submode

```
n7k(config)# mon ses 1
n7k(config-monitor)# source int e2/1
n7k(config-monitor)# destination int e2/24
n7k(config-monitor)#
```

- Support for L2, L3, subinterfaces, SVIs, VLANs
- Mix any source types in a single session
- Support for virtual SPAN using trunks with allowed VLAN lists

Wireshark

- DCOS offers an integrated packet capture tool for control packets built on top of Wireshark called Ethanalyzer
- Ethanalyzer can capture packets sent/received by the Supervisor
- Ethanalyzer cannot capture data plane traffic getting forwarded in hardware
- Ethanalyzer can open and/or save packet data captured



wr@cisco.com

Configuration Sessions

- Separate configuration mode allowing “dry-run” verification of hardware resource availability
- Initial release support security ACLs and QoS service-policies only
- Enter desired configuration, then “verify”
- System goes through motions of applying configuration, but does not actually apply it in the hardware
- System returns success or failure result
- If success returned, “commit” configuration to hardware

Configuration Rollback

- Provides checkpointing and rollback facility to return configuration to any previous state
- Options to name checkpoints, view contents of checkpointed configuration, diff checkpoints versus each other or running/startup configuration, etc.

```
tstevens-dc3-10# sh checkpoint
----------------------------------------------------------------------
Checkpoint_id   Label          UserName                TimeStamp
----------------------------------------------------------------------

16777476        10-8           tstevens                Mon Oct  8 21:55:45 2007

tstevens-dc3-10# rollback destination label 10-8
Note: Processing the Request... Please Wait
Note: Generating the Rollbackpatch... Please Wait
Note: Executing the patch... Please Wait
`conf t`
`interface Ethernet1/1`
`no service-policy type qos input foo stats-enable`
`no ip access-group test in`
tstevens-dc3-10#
```


IPv4 and IPv6 Unicast Forwarding

- Fully distributed Layer 3 IPv4 and IPv6 hardware switching

- IP services
  - uRPF check (strict and loose)
  - PBR
  - DHCP helper
  - Proxy ARP (off by default)

- First Hop Redundancy Protocols
  - IPv4 only
  - HSRP
  - GLBP
  - VRRP

Routing Protocols

Protocol	IPv4?	IPv6?	VRF-Aware?	GR?
OSPF	Yes	Yes	Yes	Yes
EIGRP	Yes	No	Yes	Yes
IS-IS	Yes	Yes	Yes	Yes
RIP	Yes	Yes	Yes	No
BGP	Yes	Yes	Yes	Yes

FIB TCAM

Protocol	Protocol Entries	TCAM Entries
IPv4 Unicast	56K	56K
IPv4 Multicast and IPv6 Unicast	32K	64K
IPv6 Multicast	2K	8K

IP Multicast Forwarding

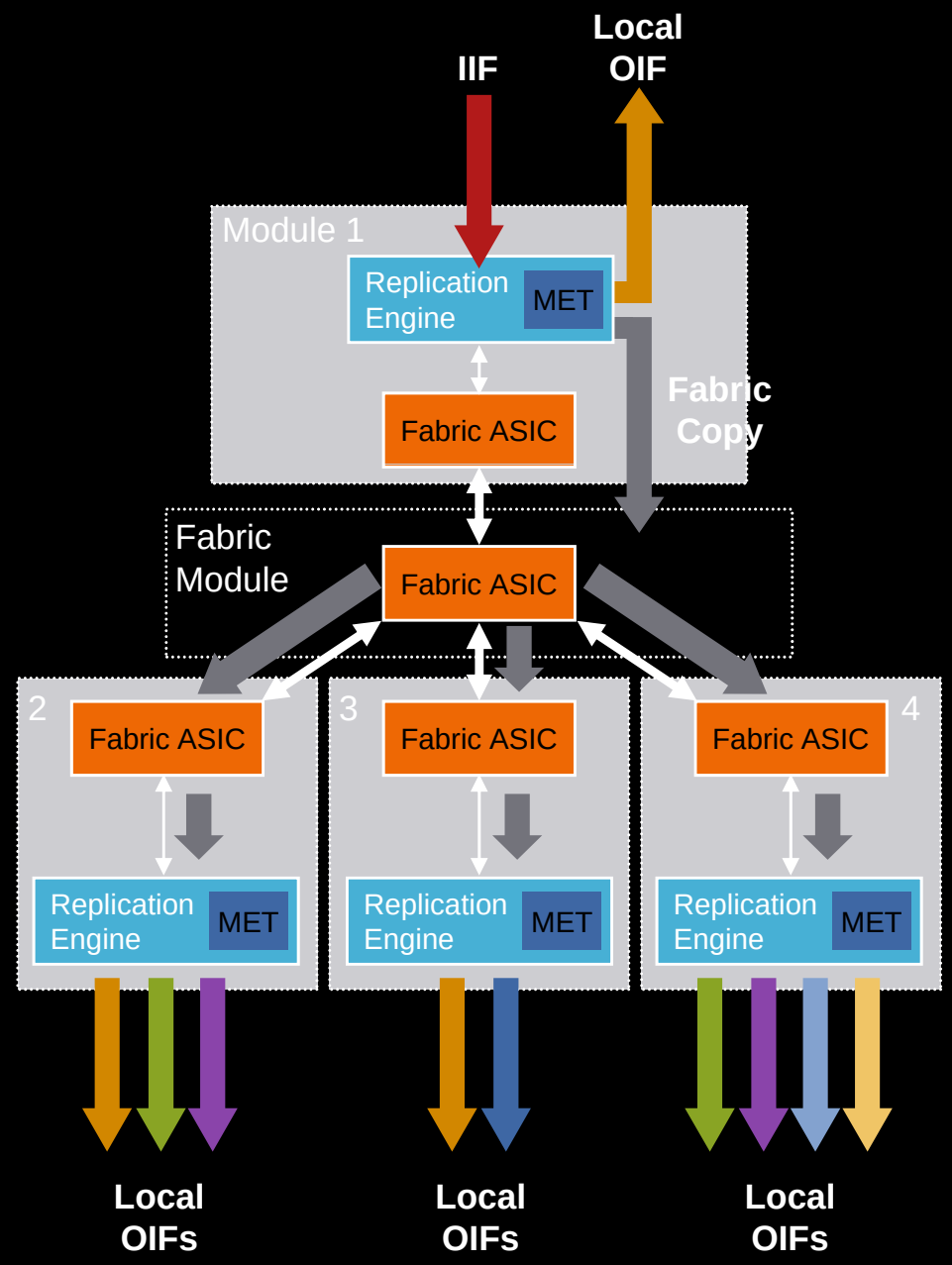
- Fully distributed Layer 2 multicast hardware switching
- Fully distributed Layer 3 IPv4 and IPv6 multicast hardware switching
- IPv4 and IPv6 (S,G), (*,G), and (*,G/m) mroute forwarding in hardware
- Distributed multicast packet replication using egress replication
- Up to 8 Bidir RPs per VRF
- IPv4 and IPv6 PIM-SM, PIM-SSM, and PIM-Bidir
- MBGP and MSDP
- Auto-RP, BSR, Anycast-RP (MSDP and PIM), and static RP
- IGMPv2/IGMPv3 and MLDv1/v2
- True IP-based IGMP snooping
- IGMP snooping querier
- VRF-lite for multicast
- SSM translation
- State limits for PIM, IGMP, MSDP, PIM6, and MLD
- Static mroutes and OIFs
- Multicast ping and traceroute

Layer 3 Multicast Replication

- “Multicast replication” typically refers to Layer 3 replication
- Creates copy of original packet for each Layer 3 OIF
- Multicast Expansion Table (MET) in replication engines contains OIFs
- Nexus 7000 system supports egress replication for Layer 3

Egress Replication

- Distributes multicast replication load among replication engines of all I/O modules with OIFs
- Input packets get lookup on ingress FE, replicated packets get lookups on ingress FE, egress FE, or both
- For OIFs on ingress module, ingress replication engine performs the replication
- For OIFs on other modules, ingress replication engine replicates a single copy of packet over fabric to all egress modules
- Replication engine on egress module performs replication for local OIFs
- MET tables on different replication engines are asymmetric wherever possible
  - Scales replication bandwidth and OIFs



wr@cisco.com

Layer 2 Forwarding

- Fully distributed Layer 2 hardware switching
- Hardware MAC learning with software synchronization and aging
- Per-VLAN Rapid Spanning Tree, Multi-Instance Spanning Tree
  - Stateful STP process restart, supervisor switchover, and ISSU
  - Backward compatibility with 802.1D STP using fallback mode
- STP guards: BPDU Guard, Root Guard, Loop Guard, BPDU Filter
- 16K VLANs – Up to 4K per VDC
- 802.1Q – VLANs and Trunking
- Private VLANs – Promiscuous/Isolated/Community PVLANs
- UDLD –Standard and aggressive mode

Data Center Network Manager

Centralized management throughout the data center network

- Fiber Channel, Ethernet, IP routing and Network Security domain awareness

Enables error-free provisioning

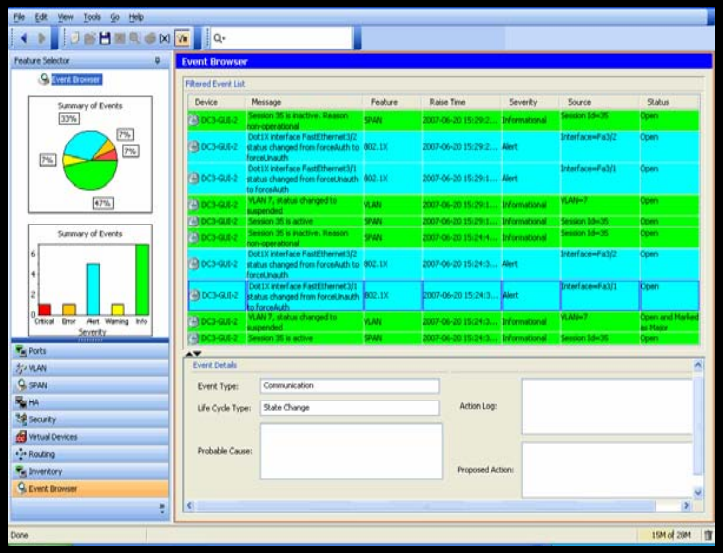
- Configuration validation via syntax and semantics checks

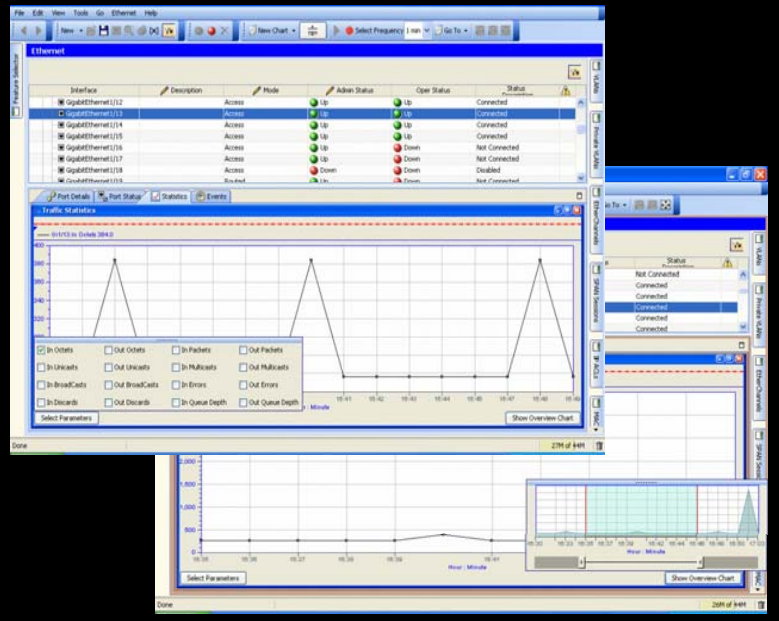
Health monitoring

- Real-time alarms and key traffic performance indicators

Facilitates the insertion of innovative network features

- Network virtualization transparently supported day 1





Intelligent Information Middleware

A middelware with a network model

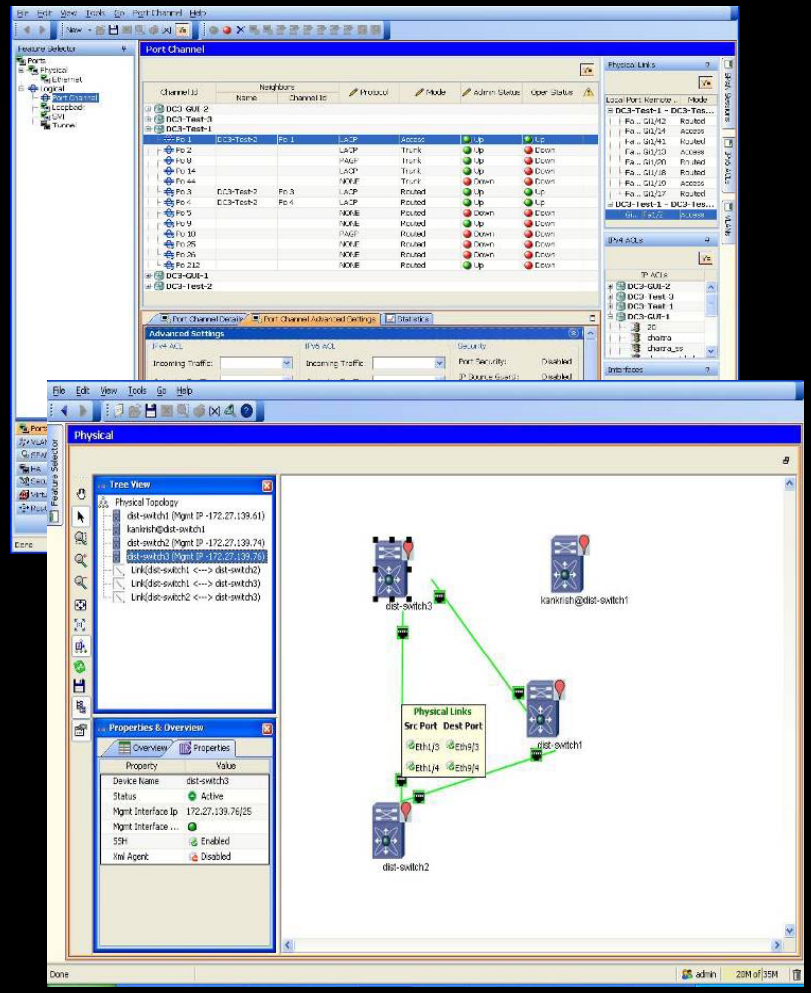
- Device mediation and network abstraction

SOA extensible framework

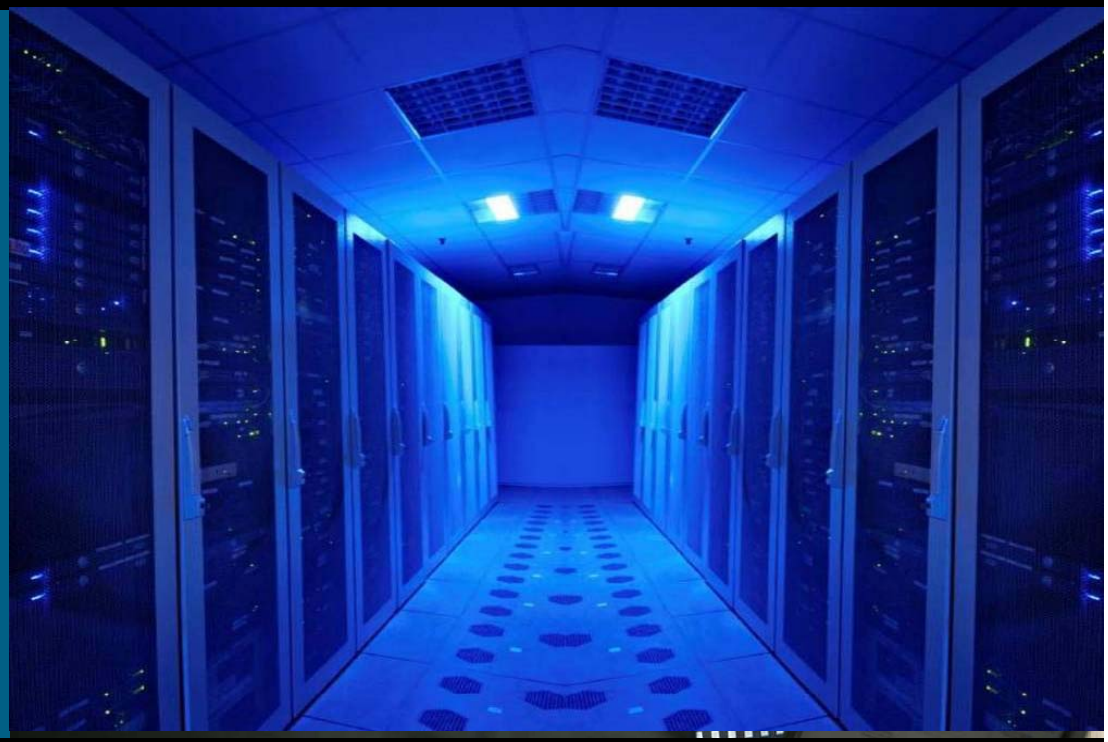
- Reuse of software components for rapid support of forthcoming NX-OS platforms

Industry-standard SOAP/XML API

- Stateful network information enabling network-aware applications



Data Center 3.0 Portfolio and Technology



NEXUS Product Line: Many Form Factors

Next Generation Data Center 3.0

Data Center Class OS

Unified Fabric

Performance & Density

Pervasive Security

MGMT Architecture

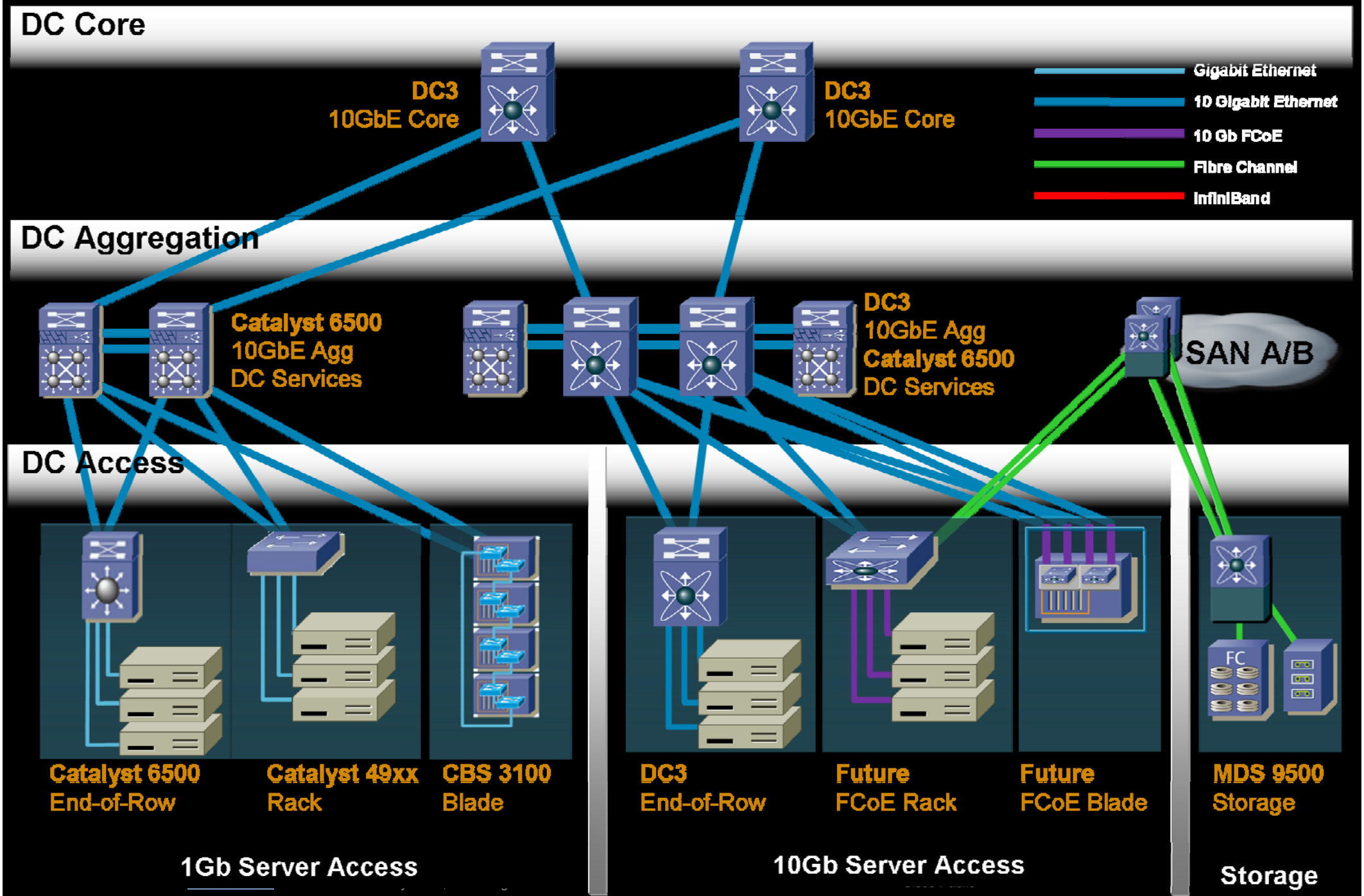
Virtualization Switching

Modular Switching

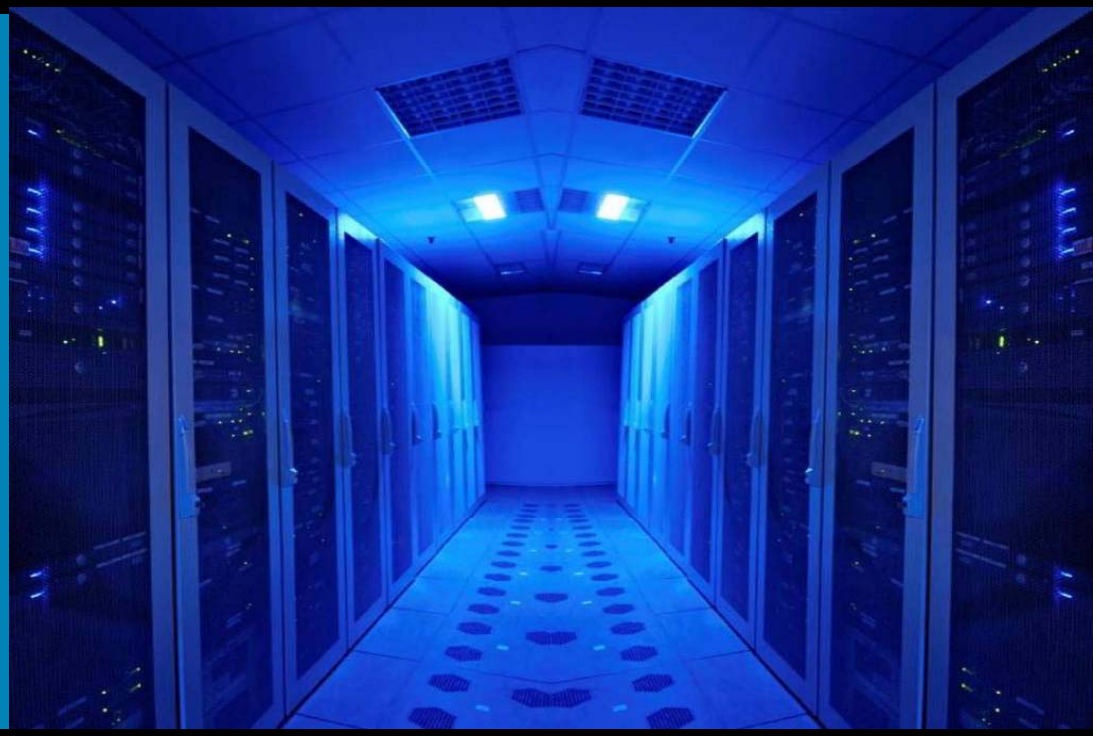
Blade Switching

ToR Switching

Data Center 3.0 Infrastructure Portfolio 2008



Summary

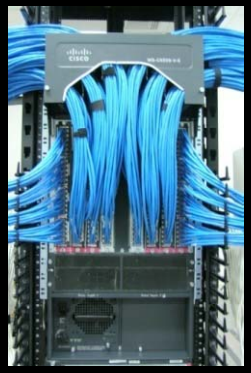


Cisco High End Switching Portfolio
Nexus 7000 and Catalyst 6500

Nexus 7000 Series

230 GbE / Slot
10G optimized

15T Switching; 500G+ / slot
10G/40G/100G optimized
Unified Fabric



Catalyst 6500

720G switching; 40G /slot
1G/10G optimized
Service modules

2T Switching; 80G+ / slot
1G/10G/40G optimized
Service Modules

Q and A

Complete Your Online Session Evaluation

Please complete the online evaluation under

www.cisco.at/expo2008/feedback

The first 100 to complete the survey will receive a copy of Don Tapscott´s book “Wikinomics”.

We very much appreciate and value your feedback, many thanks!