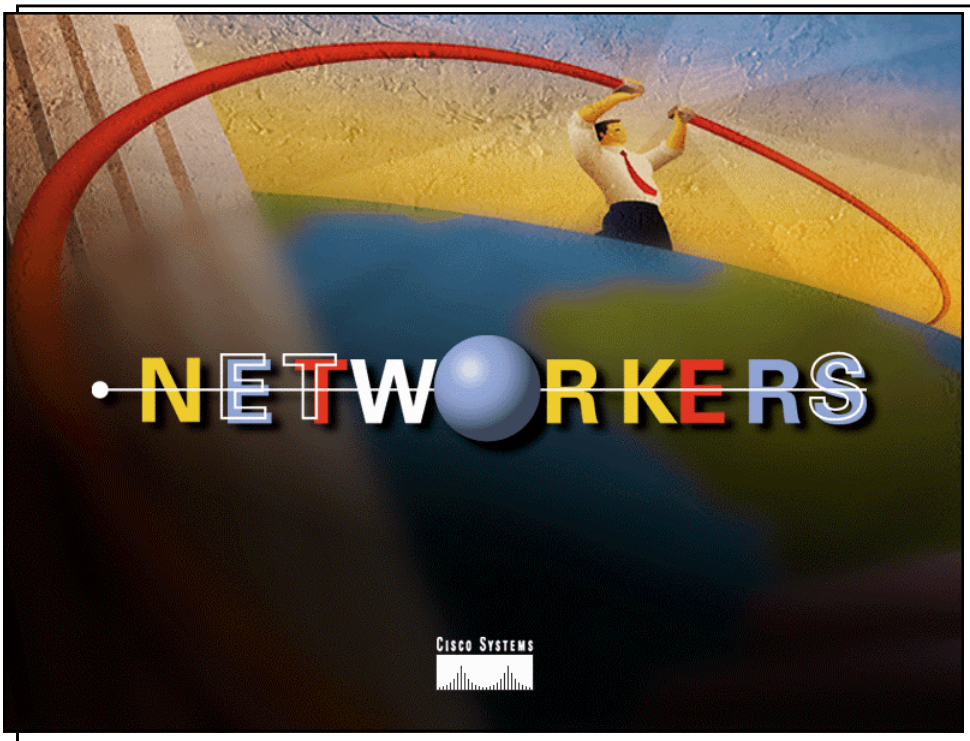
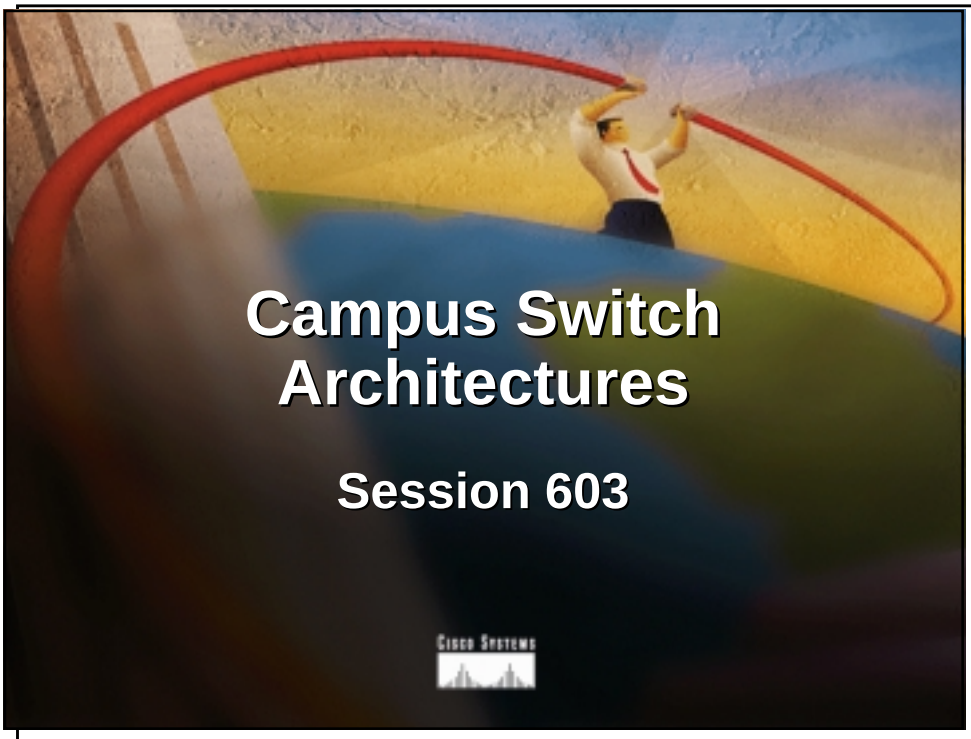




Agenda

Architecture Responsibilities

- Queuing Models
- Switching Implementations
- Switching Fabrics
- Example—Catalyst 4000 Series
- Example—Catalyst 8500 Series
- Example—Catalyst 5000 Family
- Example—Catalyst 6000 Family

www.cisco.com

Typical Responsibilities of the Hardware

- Switching packets
- Packet lookups (assuming cache is primed)
- Quality of Service
- Access Control Lists
- Address Learning (in some implementations)

www.cisco.com

Typical Responsibilities of the Software

- Anything not handled in hardware (i.e. Appletalk, DecNet, etc.)
- Server load balancing
- Security (IPSec, etc.)
- Intrusion detection systems

www.cisco.com

Typical Responsibilities of the Control Plane

- Spanning Tree
- Routing Protocol updates
- ICMP packets
- RMON and SNMP statistics
- The Routing Table
- Flow setup frames (ie. the first packet of a flow)
- Address Learning (if no hardware mechanism is enabled)

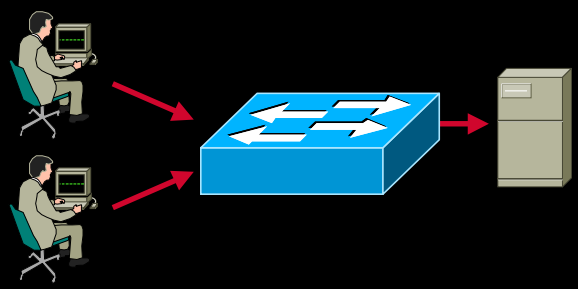
www.cisco.com

Agenda

- Architecture Responsibilities
- Queuing Models
- Switching Implementations
- Switching Fabrics
- Example—Catalyst 4000 Series
- Example—Catalyst 8500 Series
- Example—Catalyst 5000 Family
- Example—Catalyst 6000 Family

www.cisco.com

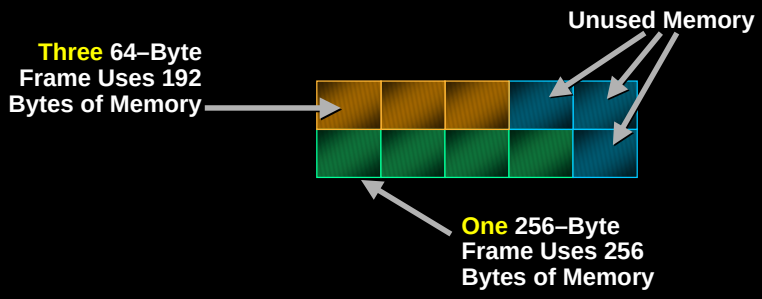
Congestion Management



- Required when multiple ports are contending for the same port
- Normalize bursts from the network to an outbound port

www.cisco.com

Dynamic Buffer Queuing



- Each buffer fixed in small increments (for example, 64 bytes each)
- Allows for efficient use of buffers

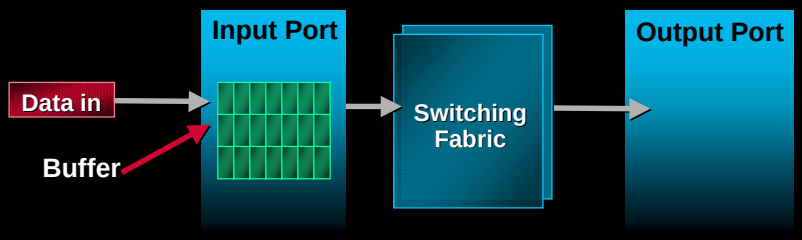
www.cisco.com

Fixed Buffer Queuing

- Buffer length fixed in size (often to MTU)
- Less expensive than custom controllers
- Inefficient use of buffers

www.cisco.com

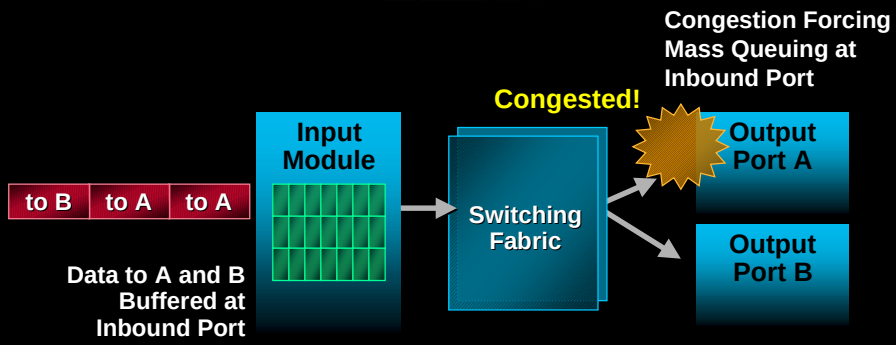
Input Queuing



- Packets buffered at the inbound port
- Results in head of line blocking
- Reduces throughput to 60% maximum

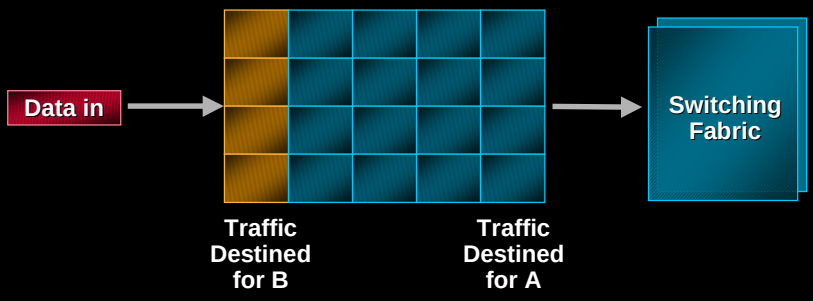
www.cisco.com

Head of Line Blocking



www.cisco.com

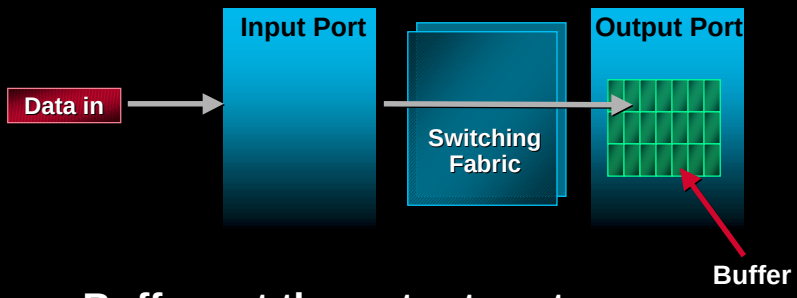
Head of Line Blocking with Input Queuing



- Traffic for B cannot enter switching fabric due to data for “A” being ahead of it

www.cisco.com

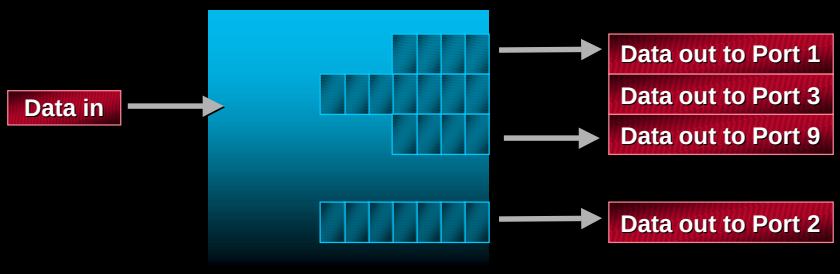
Output Queuing



- Buffers at the output port
- No head of line blocking
- Can overflow buffers at peak bursts

www.cisco.com

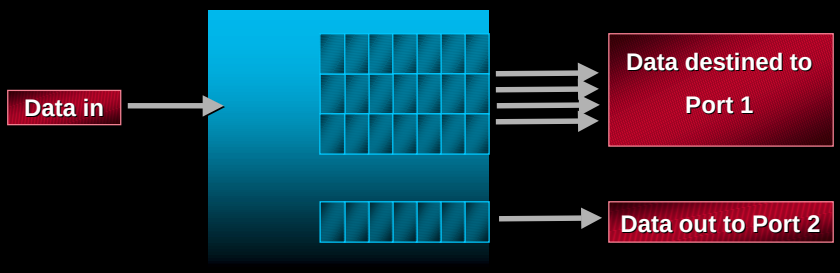
Output Queuing/Shared Buffer



- Central pool of buffers shared between all ports
- Maximum throughput with fewest buffers
- No head of line blocking

www.cisco.com

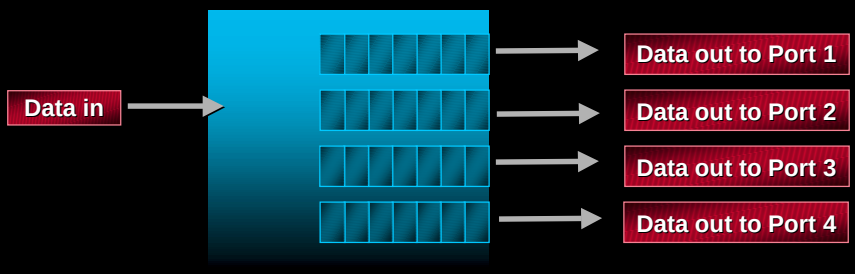
Rules of Shared Buffer Architectures



- Amount of buffer can be dynamically allocated per port ONLY IF there are no QoS requirements

www.cisco.com

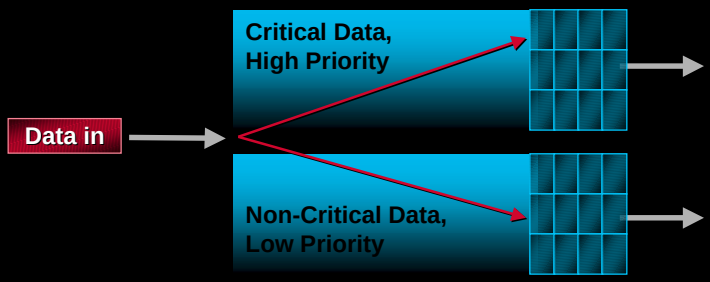
Rules of Shared Buffer Architectures



- If QoS is a requirement, buffers are typically fixed per port and per queue

www.cisco.com

Multiple Queues per Port

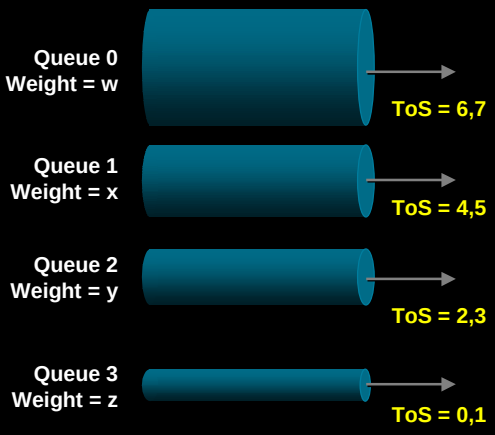


- Can be implemented in either output queuing or shared memory models
- Support multiple service levels for QoS
- Scheduling and/or congestion avoidance algorithm required

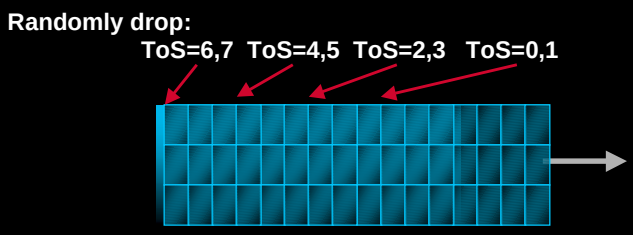
www.cisco.com

Weighted Round Robin

- Outbound scheduling algorithm
- Service assigned via user-defined weights on outbound port
- Scheduler enforces bandwidth requirements per queue



Weighted Random Early Detection



- Pre-emptive drops as buffer begins to fill
- Without WRED: tail drop of packets throttles multiple TCP sessions
- With WRED: limited numbers of TCP sessions throttle back due to random discard

Agenda

- Architecture Responsibilities
- Queuing Models
- Switching Implementations
- Switching Fabrics
- Example—Catalyst 4000 Series
- Example—Catalyst 8500 Series
- Example—Catalyst 5000 Family
- Example—Catalyst 6000 Family

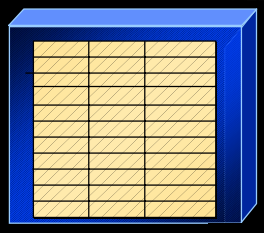
www.cisco.com

Switching Implementations

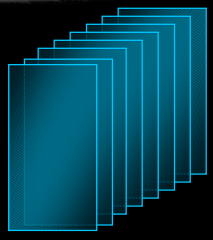
- Describe where and how a switching decision is made
  - Where: Locally or centrally
  - How: Longest match, exact match

www.cisco.com

Serial vs. Hashed Lookup



- Sequentially look up entries in CAM
- **Benefit**: more entries
- **Drawback**: decrease in performance as entries increase

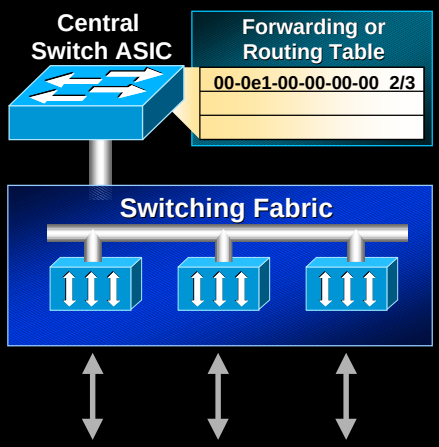


- Hash values together to obtain a value in memory
- **Benefit**: faster lookup
- **Drawback**: fewer entries due to hash collisions

www.cisco.com

Centralized Switching

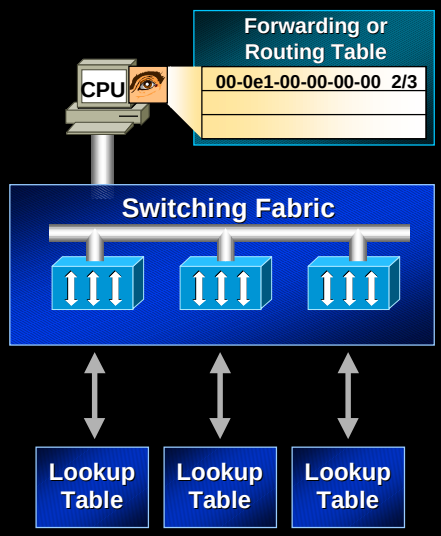
- Central forwarding table utilized
- Provides centralized control for switching and learning
- Lookup done is ASICs for fast processing
- Can perform a Layer 2 or Layer 3 lookup



www.cisco.com

Distributed Switching

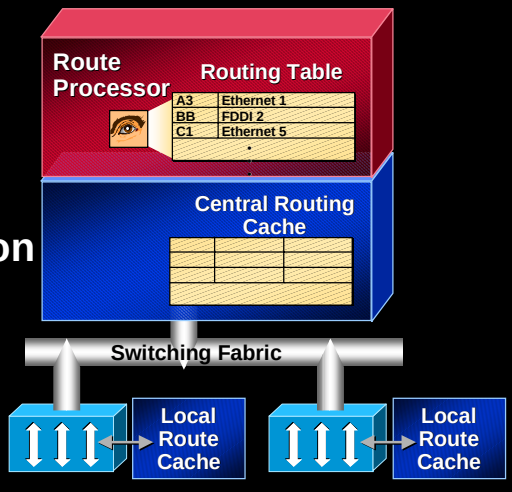
- Switching decision made locally by port or module
- L2 and L3 tables must be synchronized to account for adds, moves or changes
- Routing table not necessarily distributed on line cards



www.cisco.com

Route Caching (Traffic-Based Switching)

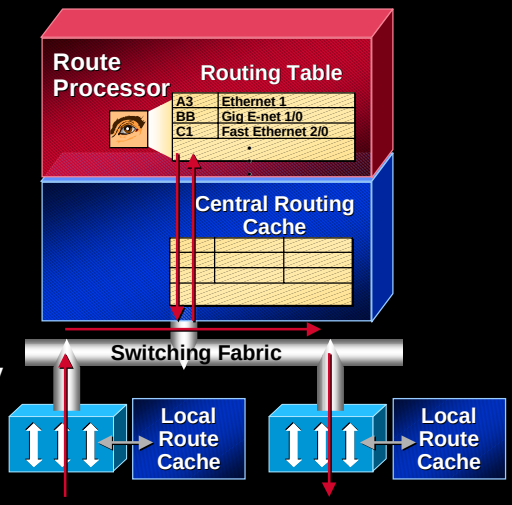
- Demand-based switching
- Entry containing source, destination and/or Layer 4 information
- Flows are unidirectional



www.cisco.com

Route Caching (Traffic-Based Switching)

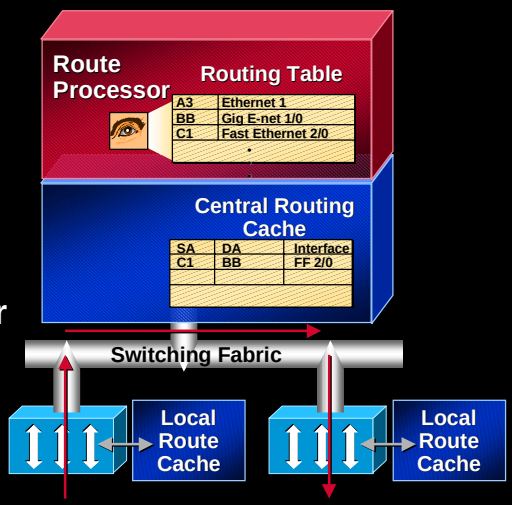
- First packet processed switched by processor
- Flow cache enabled (centrally and/or locally)



www.cisco.com

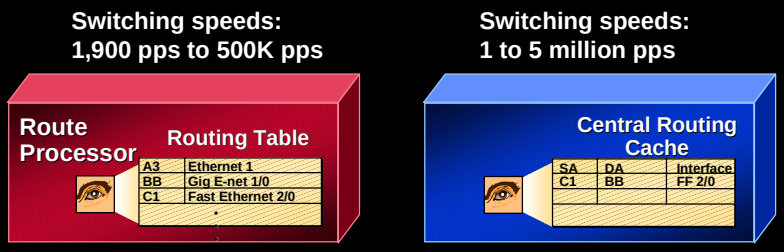
Route Caching (Traffic-Based Switching)

- Subsequent packets in that flow switched via flow cache
- Flow cache decision made centrally and/or locally
- All packets switched at Layer 3



www.cisco.com

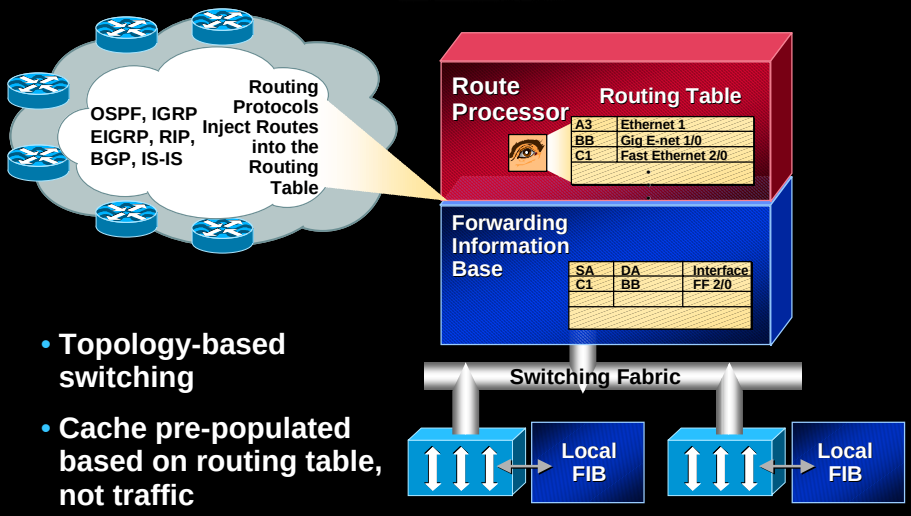
The Slow Path



- First packet in a traffic based switching implementation must go to Route Processor
- Speed of switch dependant on performance of this engine
- Slow path functions: cache entry, cache flush in route flap

www.cisco.com

Cisco Express Forwarding (Topology-Based Switching)

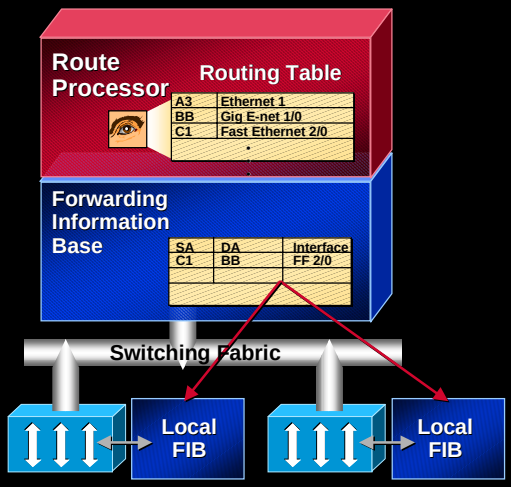


- Topology-based switching
- Cache pre-populated based on routing table, not traffic

www.cisco.com

Cisco Express Forwarding (Topology-Based Switching)

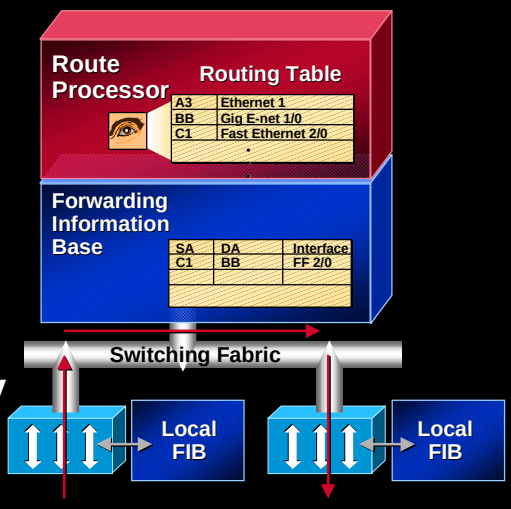
- FIB calculated based on routing table entries, not traffic flows
- FIB can be kept central or distributed



www.cisco.com

Cisco Express Forwarding (Topology-Based Switching)

- Packet enters router
- No process switching necessary
- Decision made locally or centrally irregardless of switching fabric



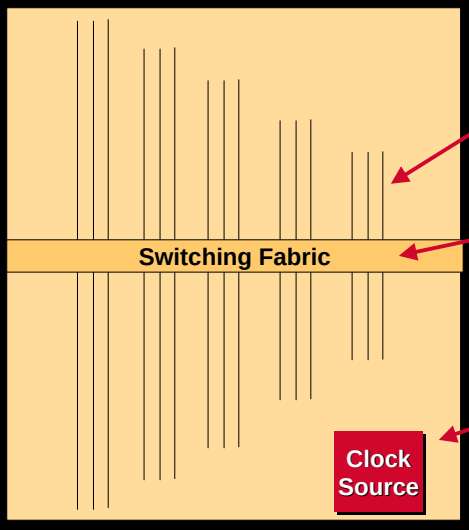
www.cisco.com

Agenda

- Architecture Responsibilities
- Queuing Models
- Switching Implementations
- Switching Fabrics
- Example—Catalyst 4000 Series
- Example—Catalyst 8500 Series
- Example—Catalyst 5000 Family
- Example—Catalyst 6000 Family

www.cisco.com

Backplane



Electrical traces (wires) run signals

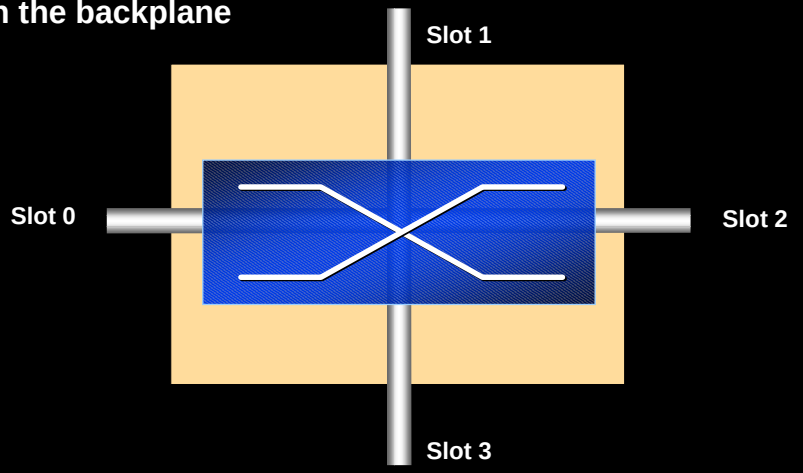
Switching Fabric switches signals between wires. Fabric can be located on backplane or line card

Wires clocked at certain speed (measured in MHz) via a clock source

www.cisco.com

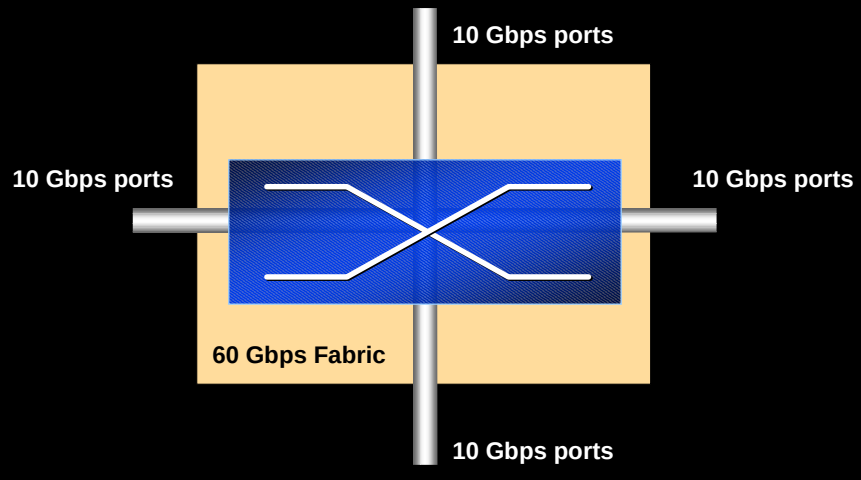
Switching Fabric

Switches between the “wires” on the backplane



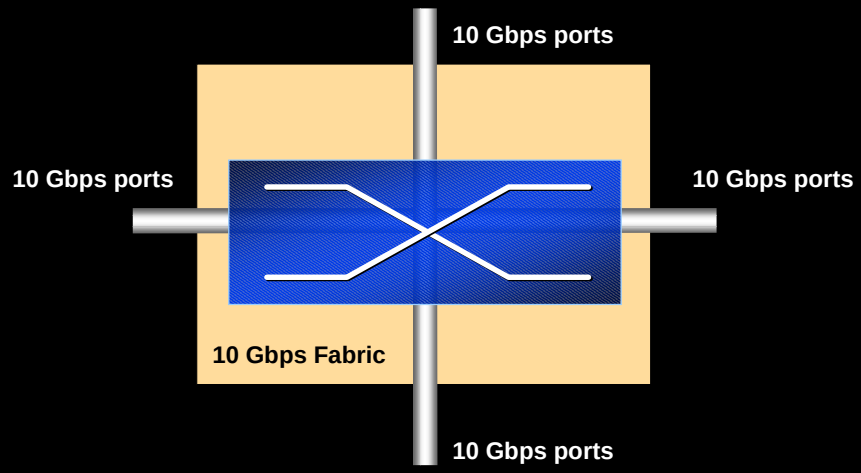
www.cisco.com

Non-Blocking Switching Fabric



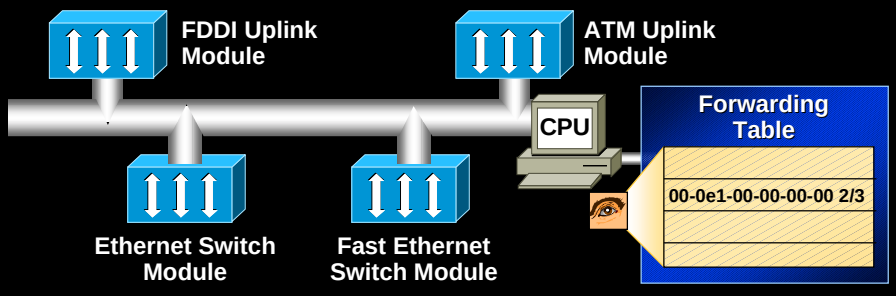
www.cisco.com

Over-subscribed Switching Fabric



www.cisco.com

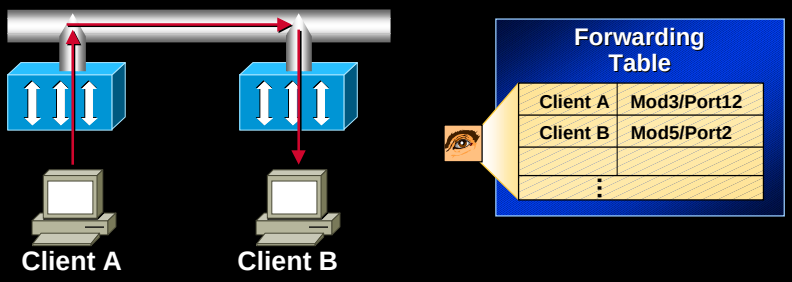
Single Bus



- One central fabric element
- Each port must arbitrate for access
- Broadcast and multicast easy
- Oversubscription normal

www.cisco.com

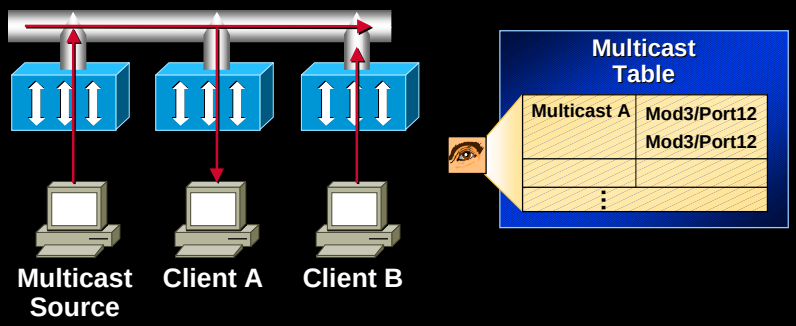
Unicast Switching over A Single Bus



- Data enters switched port
- Bus arbitration takes place
- Lookup in forwarding table performed
- Packet sent to destination port

www.cisco.com

Multicast/Broadcast over A Single Bus

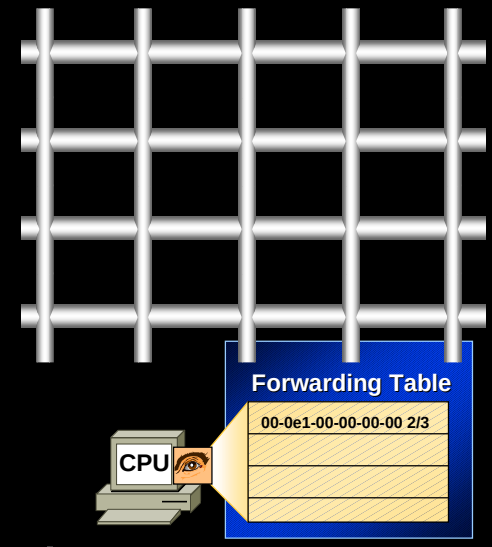


- Flooded data decreases end-station performance
- Destination must be only those ports who need that traffic
- Multicast or VLAN mechanism must limit traffic to certain ports

www.cisco.com

Crossbar Fabrics

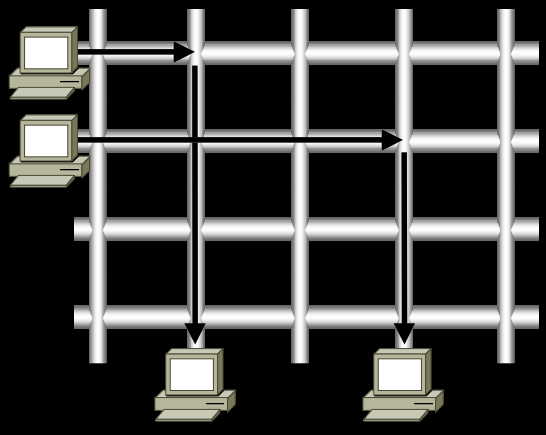
- Multiple input buses allows for meshed fabric
- Typically (but not always) non-blocking
- Broadcast/multicast complex
- More complex forwarding table lookup



www.cisco.com

Unicast Switching over Crossbar Fabrics

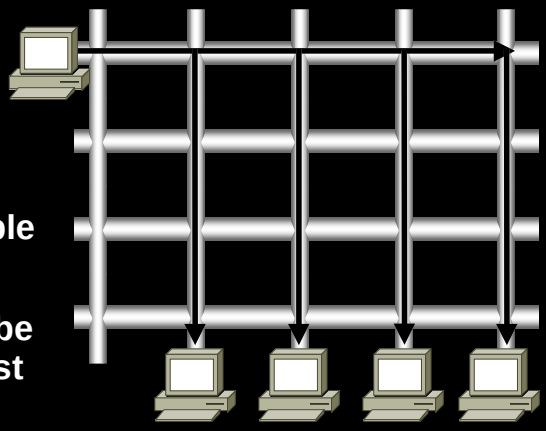
- Balanced traffic
- Ports/modules have access to part of the crossbar
- Multiple traffic streams simultaneously in fabric



www.cisco.com

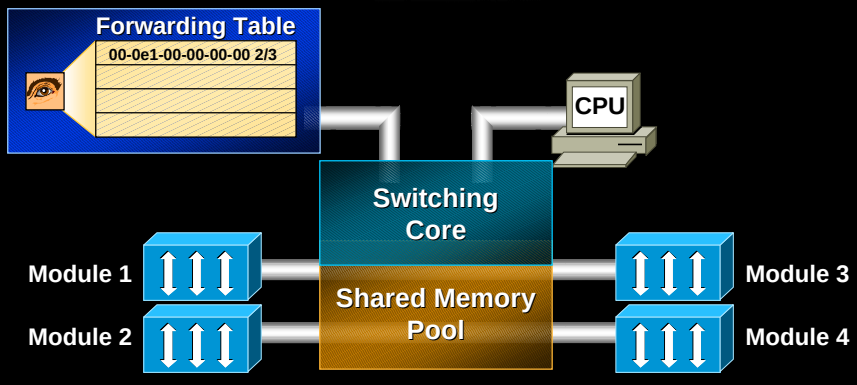
Multicast/Broadcast over a Crossbar

- Weakness in a crossbar architecture
- Packet replication must occur, possible performance loss
- Entire fabric must be “quiet” for multicast to occur



www.cisco.com

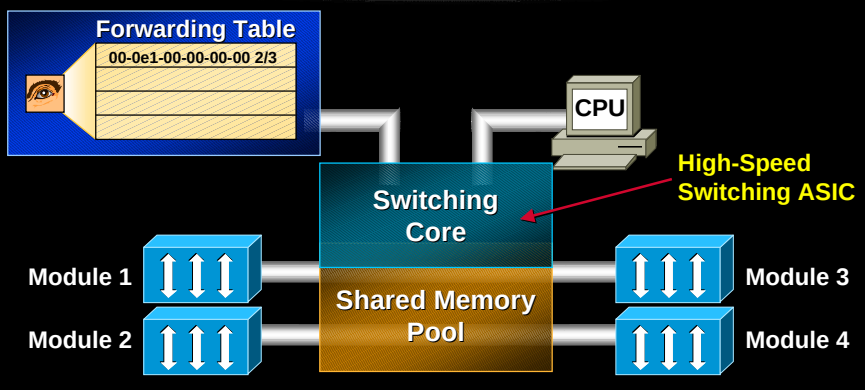
Shared Memory Architecture



- Buffering internal to the switching fabric
- Switching inputs to memory governed by ASICs
- Uses high-speed memory and switching fabric

www.cisco.com

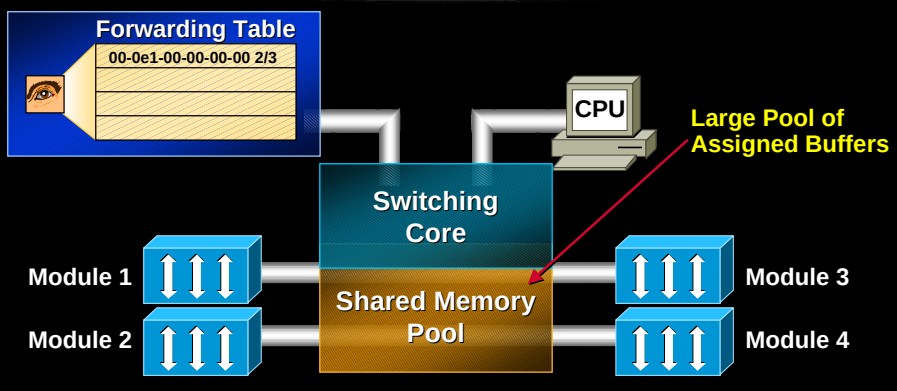
Shared Memory Architecture



- Switching core performs lookup, resolves destination to pointers in memory, switches the packet

www.cisco.com

Shared Memory Architecture

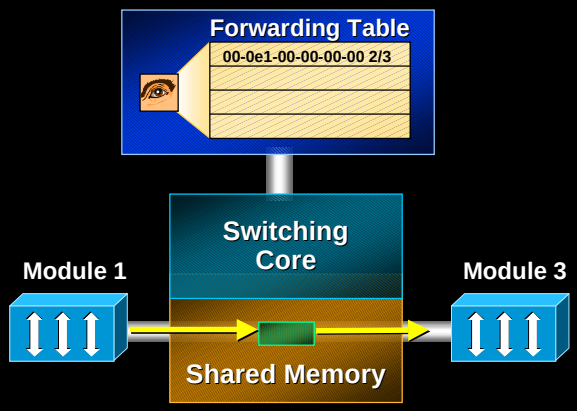


- Buffers either fixed or dynamic, less buffers needed if architecture is non-blocking

www.cisco.com

Unicast Switching over A Shared Memory Fabric

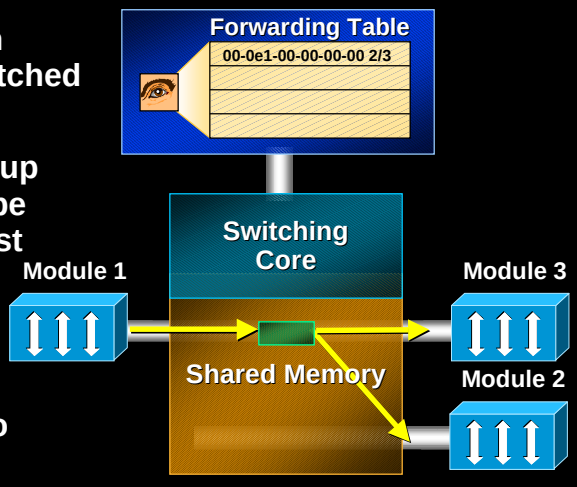
- Data enters switch module
- Switched into memory
- Switching core evaluates and resolves destination
- Packet switched to outgoing module



www.cisco.com

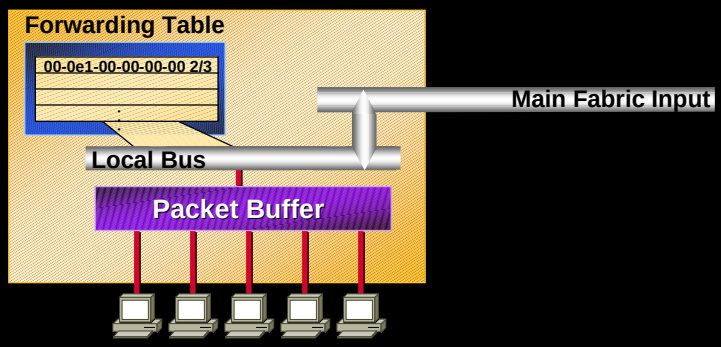
Multicast/Broadcast Switching over a Shared Memory Fabric

- Data enters switch module and is switched into memory
- MAC address lookup reveals packet to be broadcast/multicast
- Switching core identifies outgoing ports
- Packet switched to outgoing multiple outgoing ports



www.cisco.com

Local Switching



- Local switching fabric keeps local traffic off main fabric; increases switch bandwidth
- Local architecture can be bus, crossbar or shared memory

Agenda

- Architecture Responsibilities
- Queuing Models
- Switching Implementations
- Switching Fabrics
- Example—Catalyst 4000 Series
- Example—Catalyst 8500 Series
- Example—Catalyst 5000 Family
- Example—Catalyst 6000 Family

The Catalyst 4000 Series

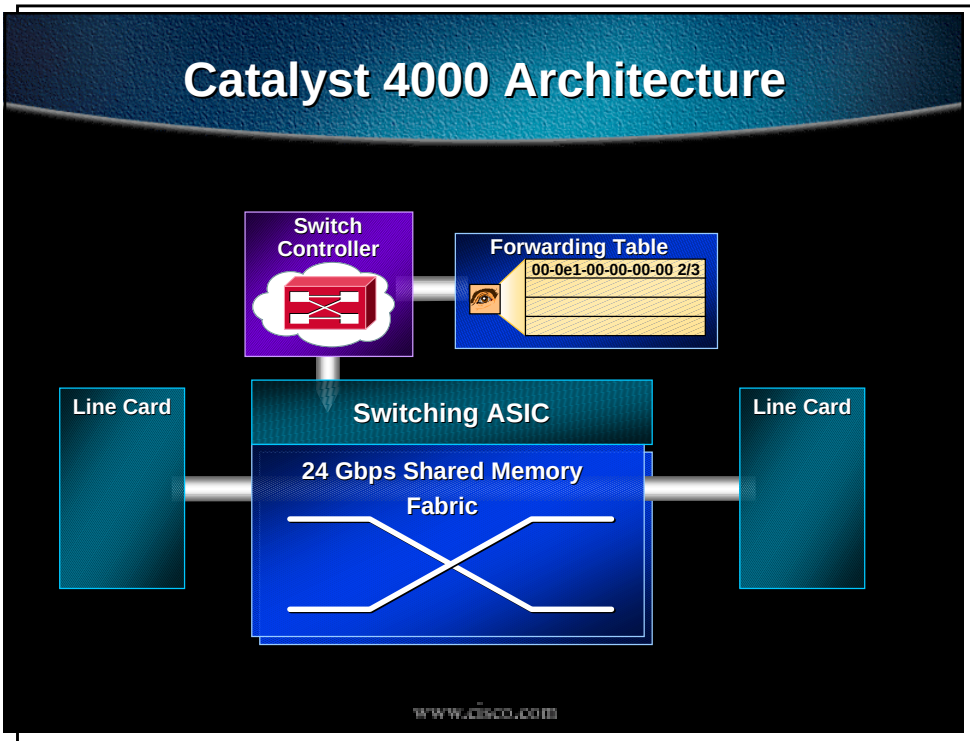
- Extension of the Catalyst 5000 Family for mid-range wiring closets
- Flexible modular configurations
- Cost-effective, high-performance 10/100/1000 Ethernet switching
- Enterprise software functionality

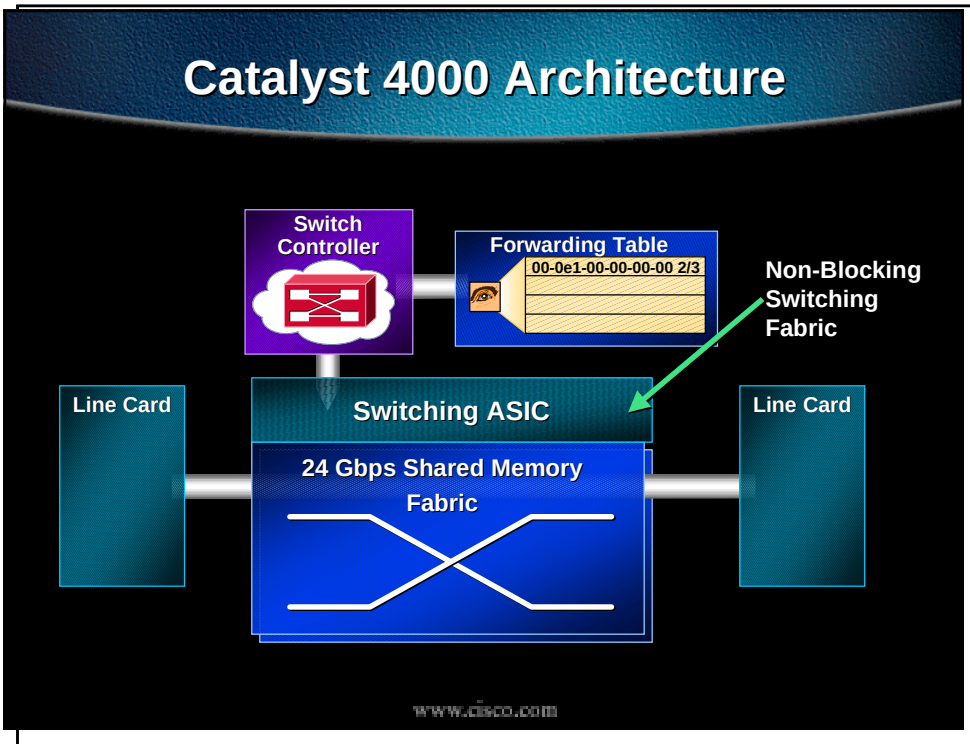
www.cisco.com

Catalyst 4912 and 2948G

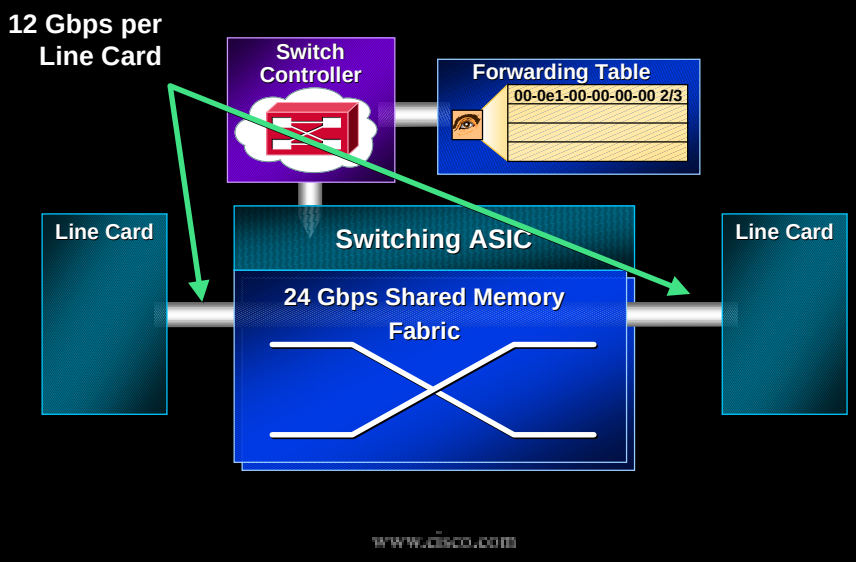
- Same architecture as the Catalyst 4000
- All ports access 24 Gbps of shared memory
- Catalyst 2948G—48 10/100 ports with two Gigabit Ethernet ports
- Catalyst 4912—12 1000BaseX ports

www.cisco.com



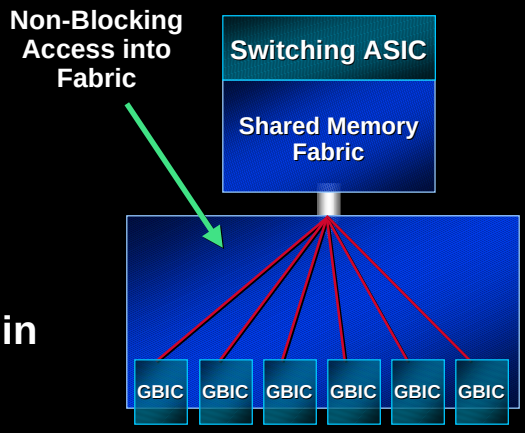


Catalyst 4000 Architecture



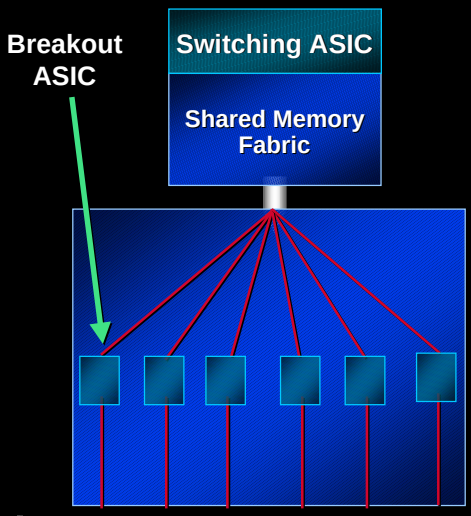
Catalyst 4000 Port Interface—Gigabit Ethernet

- 12 Gbps per line card
- No buffering or switching on line cards
- All intelligence in switching core ASIC



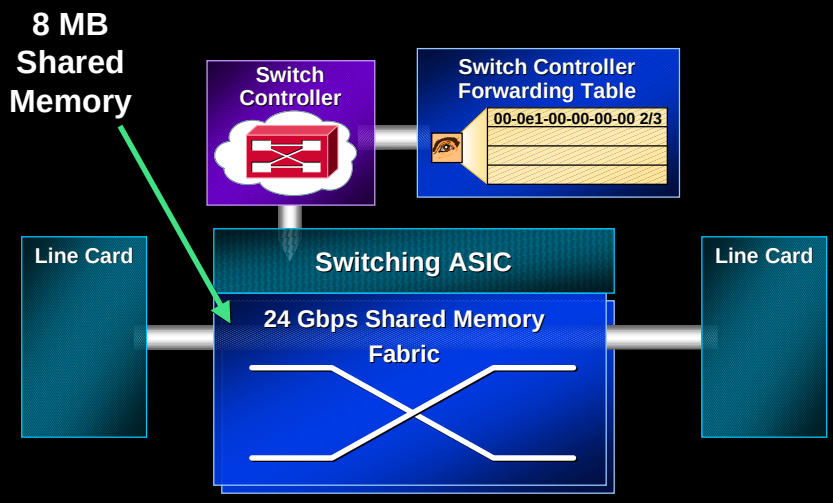
Catalyst 4000 Port Interface—Fast Ethernet

- Breakout ASIC connects 8 10/100 ports to single 1 Gbps fabric port
- Non-blocking 10/100
- All switching on supervisor engine

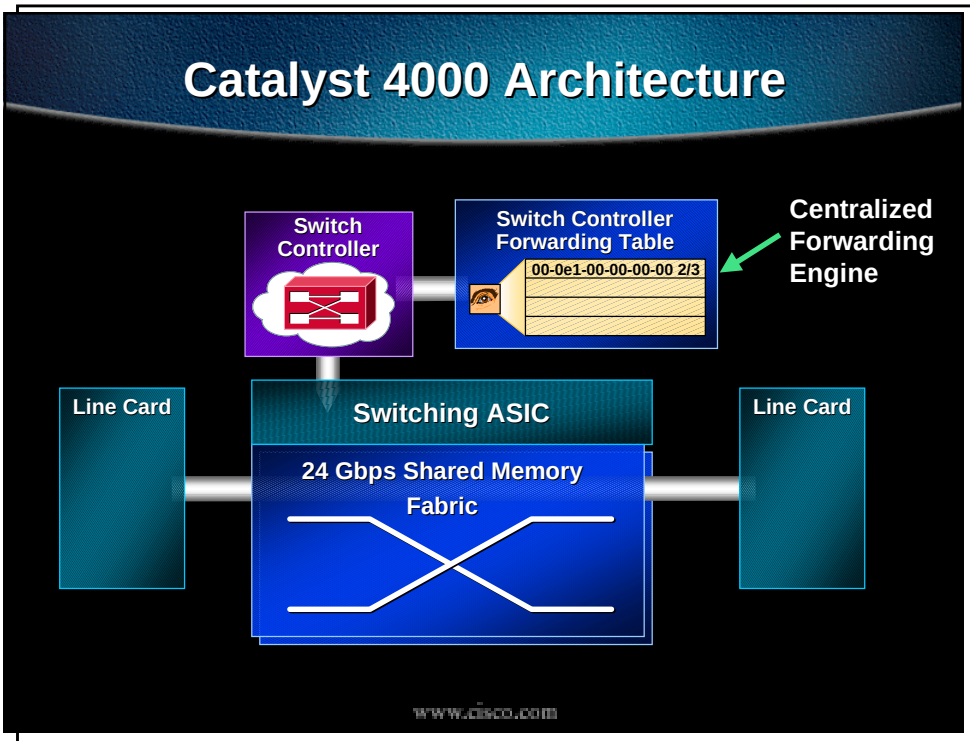


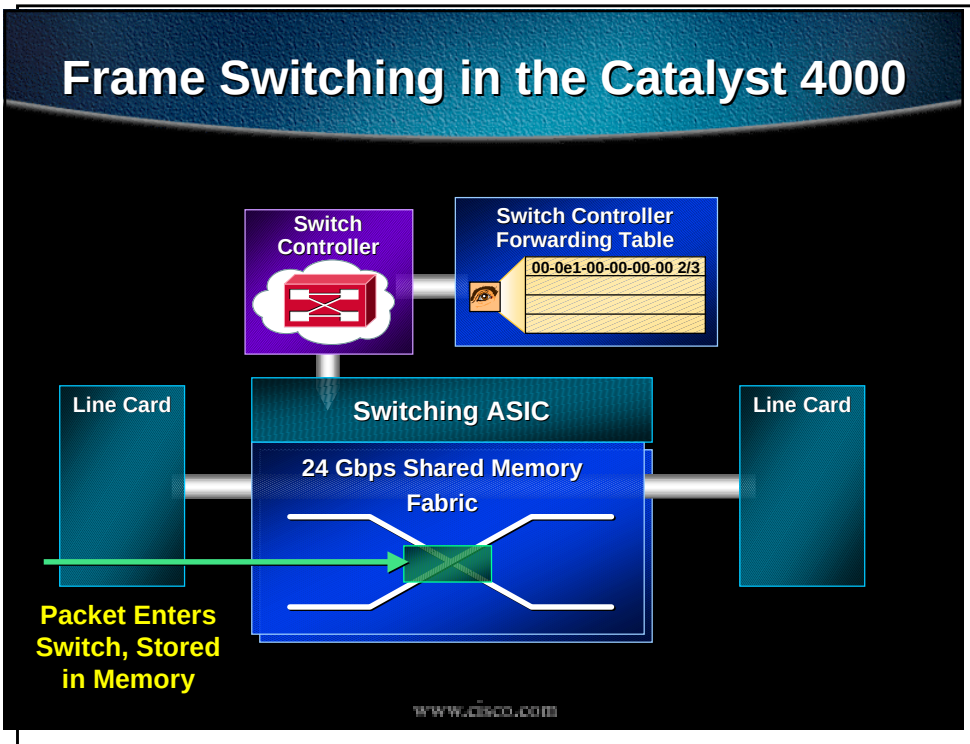
www.cisco.com

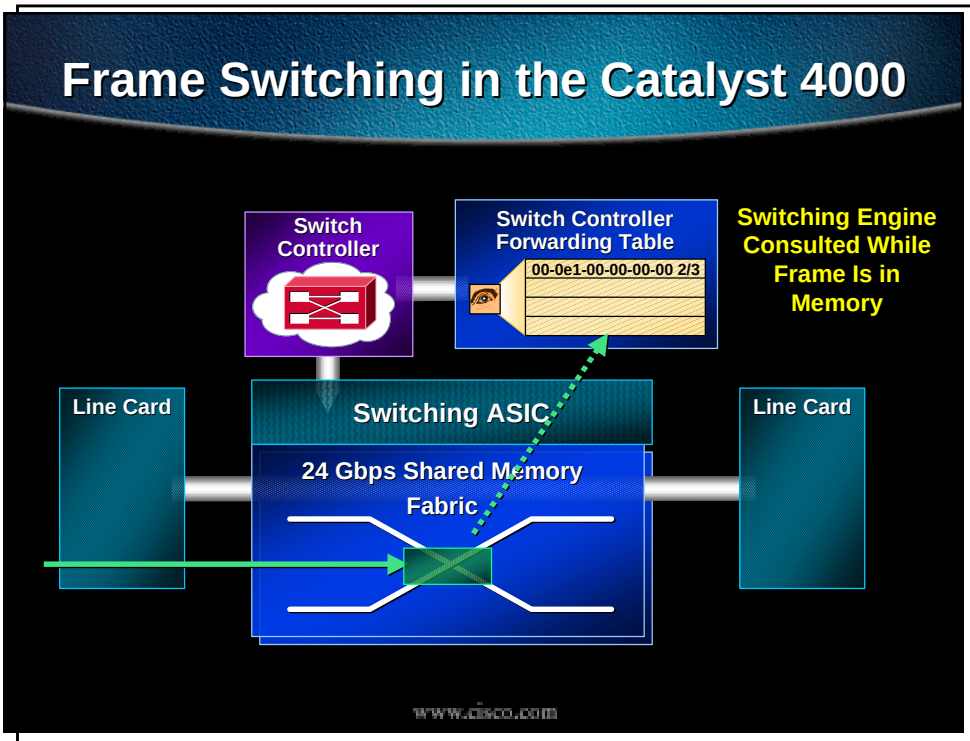
Catalyst 4000 Shared Memory

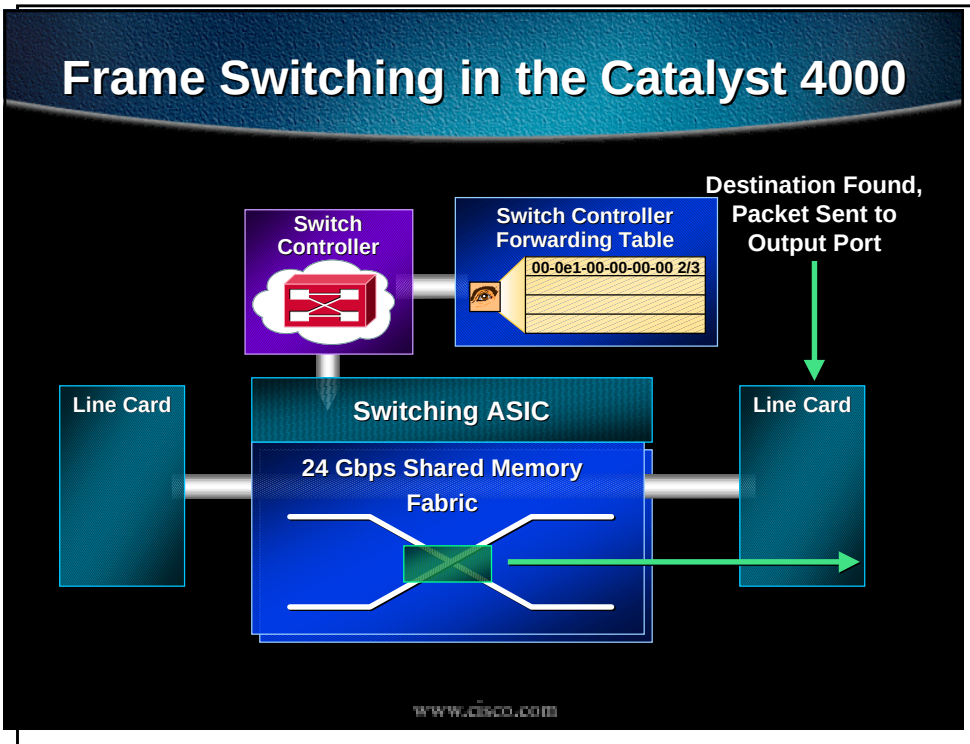


www.cisco.com









Catalyst 4003/4912 Performance

- 10/100 for 96 ports
  - 14,285,760 packets per second (64 bytes)
- 1000BaseX for 12 ports
  - 17,857,200 packets per second (64 bytes)

www.cisco.com

Catalyst 2948G Performance

- 10/100 for 48 ports
  - 7,142,880 packets per second (64 bytes)
- 1000BaseX for 2 ports
  - 2,961,319 packets per second (64 bytes)

www.cisco.com

Agenda

- Architecture Responsibilities
- Queuing Models
- Switching Implementations
- Switching Fabrics
- Example—Catalyst 4000 Series
- Example—Catalyst 8500 Series
- Example—Catalyst 5000 Family
- Example—Catalyst 6000 Family

www.cisco.com

Catalyst 8500 Series

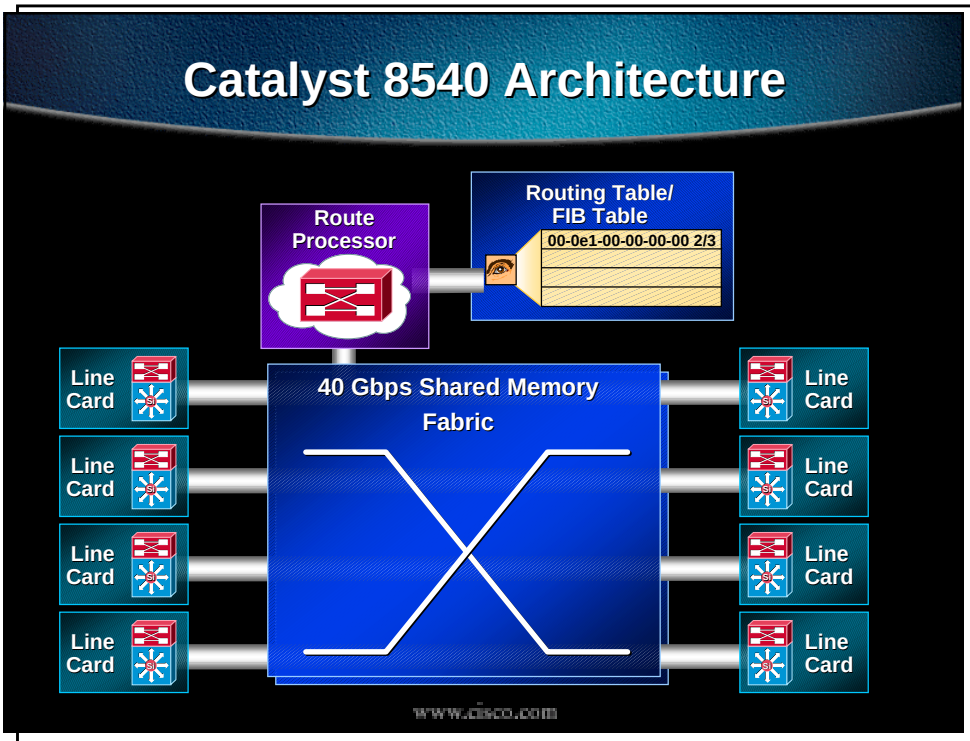
Cisco IOS®

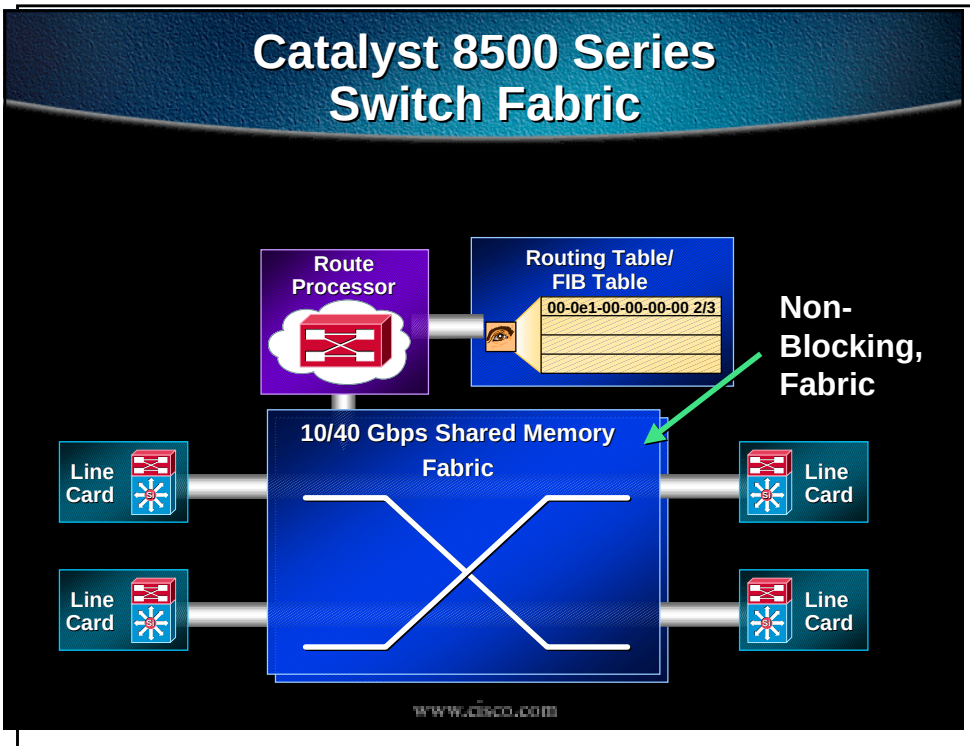
Cisco Express Forwarding	Wire-Speed IP, IPX Switching
Cisco IOS Routing Protocols	Wire-Speed IP Multicast
Extensive QoS Capabilities	Multiservice Integration

Catalyst 8510
5 Slot

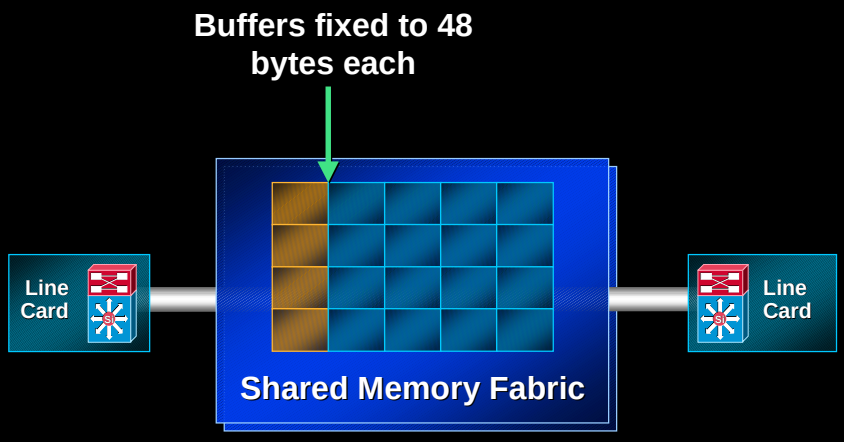
Catalyst 8540
13 Slot

www.cisco.com

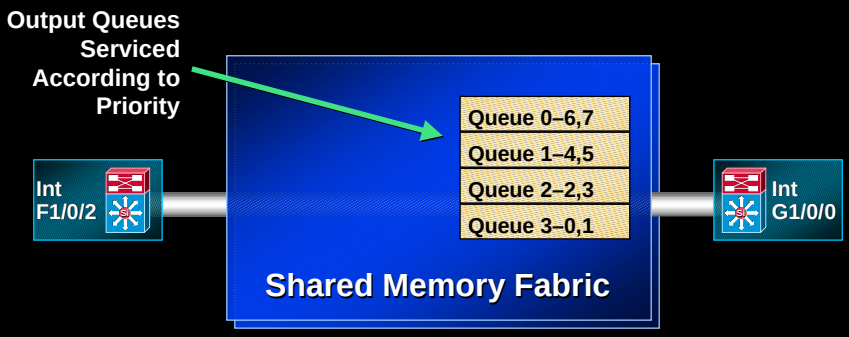




Dynamic Buffering Architecture

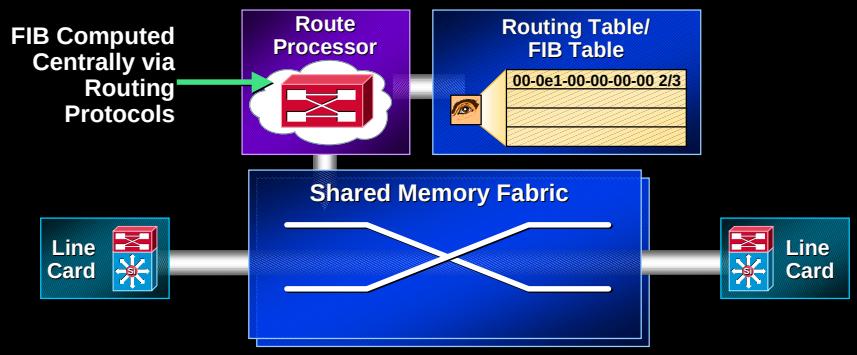


Multiple Queues per Port



- Four queues per port
- Scheduling algorithm allows average bandwidth per queue

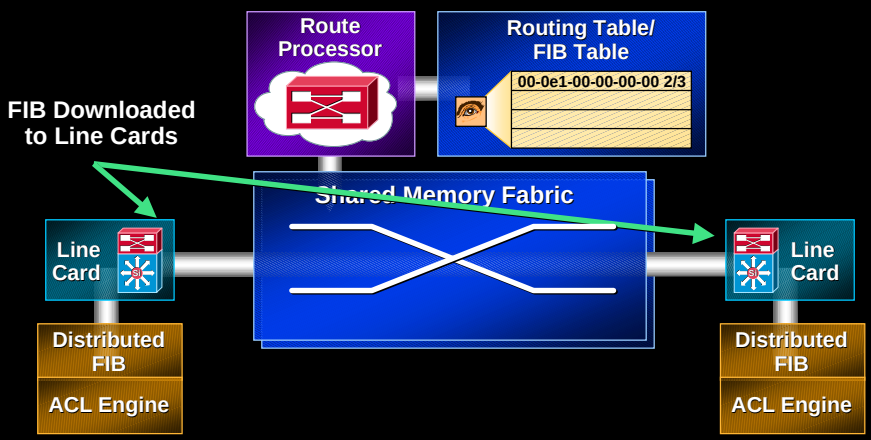
Catalyst 8500 Forwarding



- Cisco IOS processor maintains routing table
- Computes forwarding information base

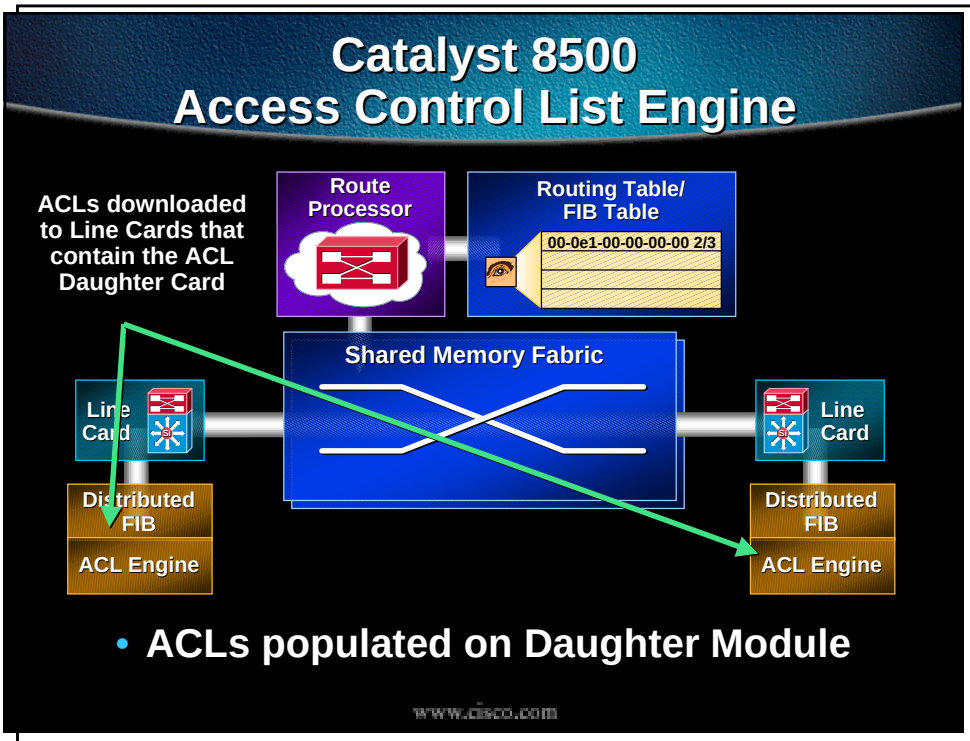
www.cisco.com

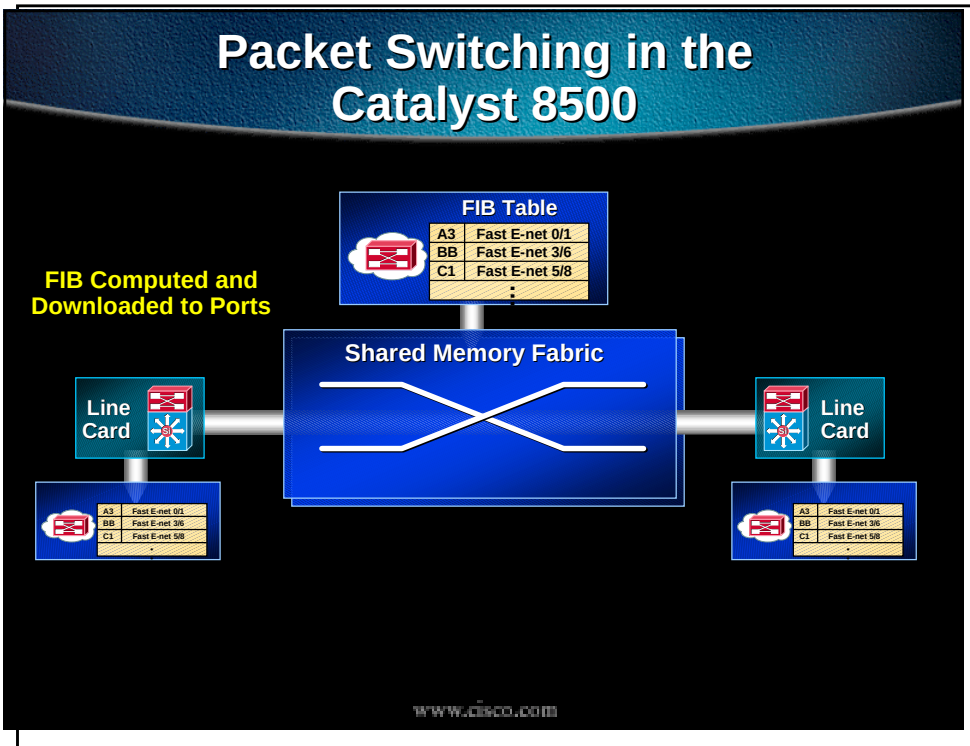
Catalyst 8500 Distributed FIB

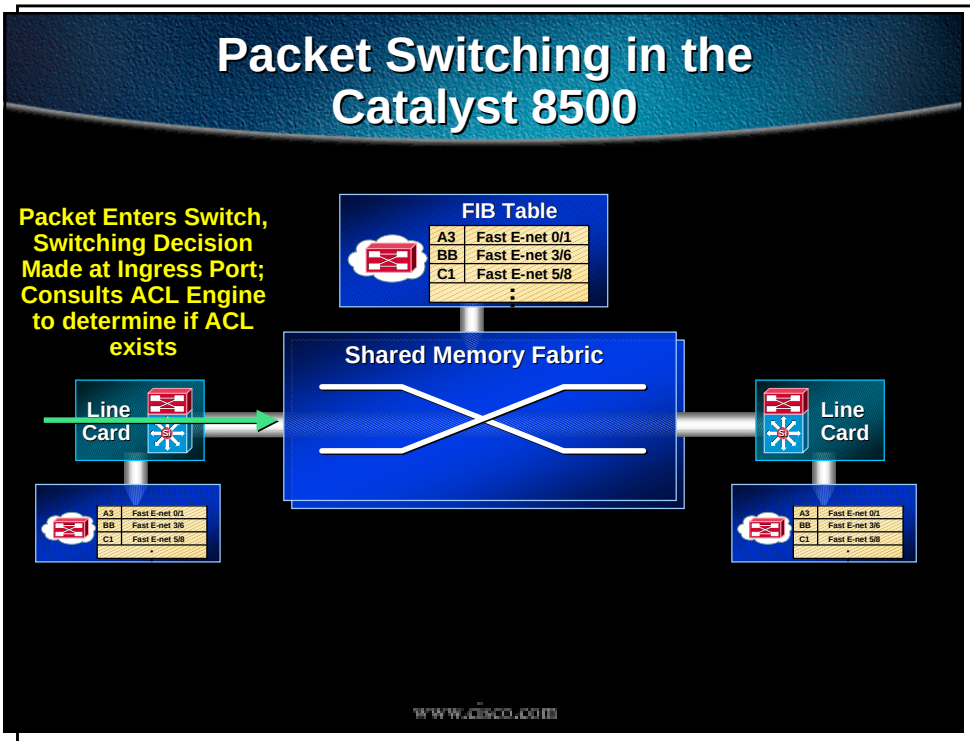


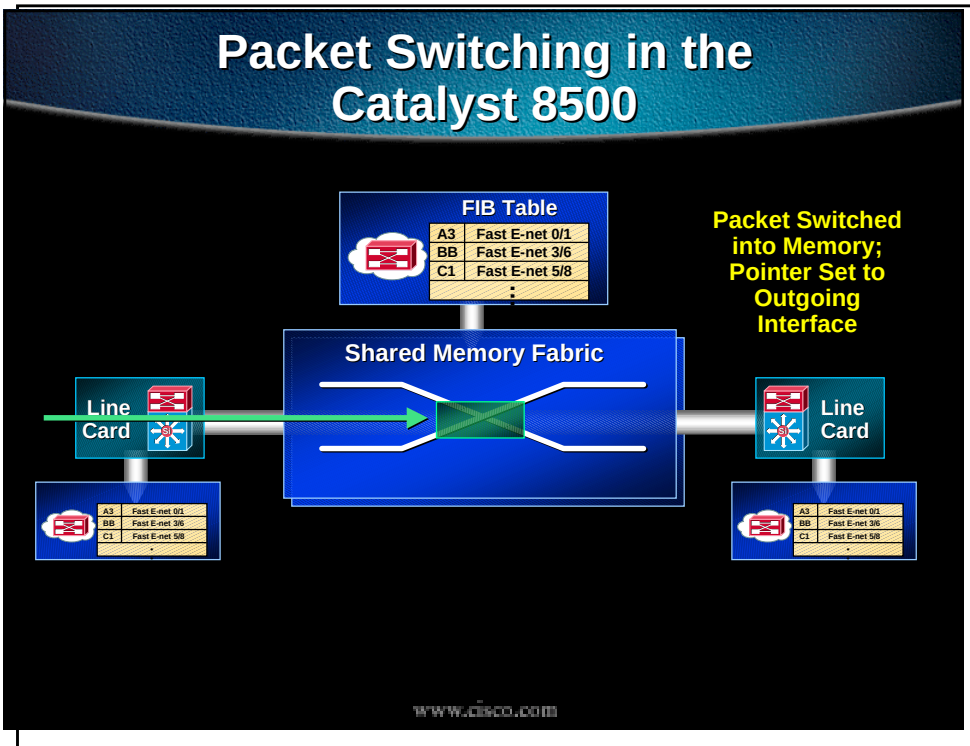
- FIB performs longest match network address lookup in hardware

www.cisco.com

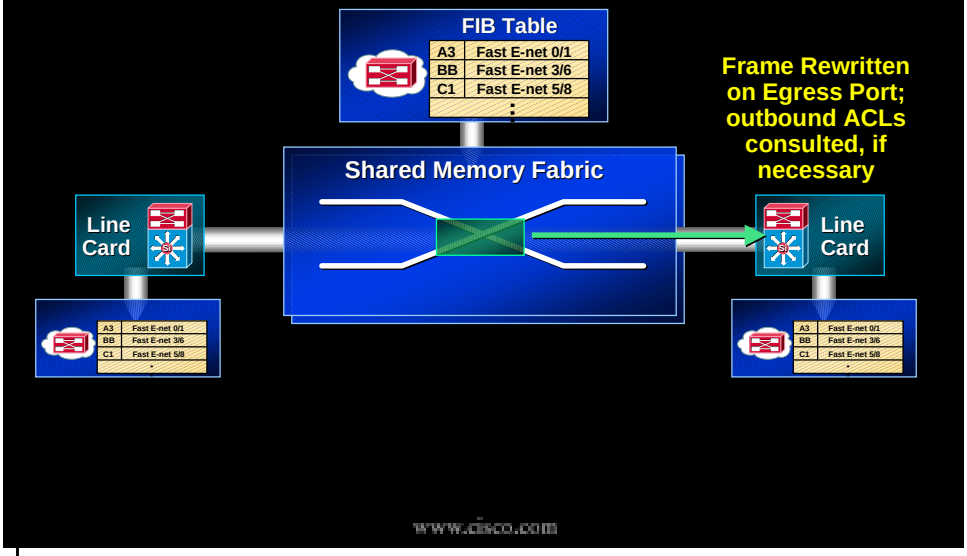








Packet Switching in the Catalyst 8500



Catalyst 8500 Series Performance

- **Fast Ethernet forwarding rate**
  - Catalyst 8510—4.76 million pps
  - Catalyst 8540—19 million pps
- **Gigabit Ethernet forwarding rate**
  - Catalyst 8510—5.77 million pps
  - Catalyst 8540—23 million pps

www.cisco.com

Agenda

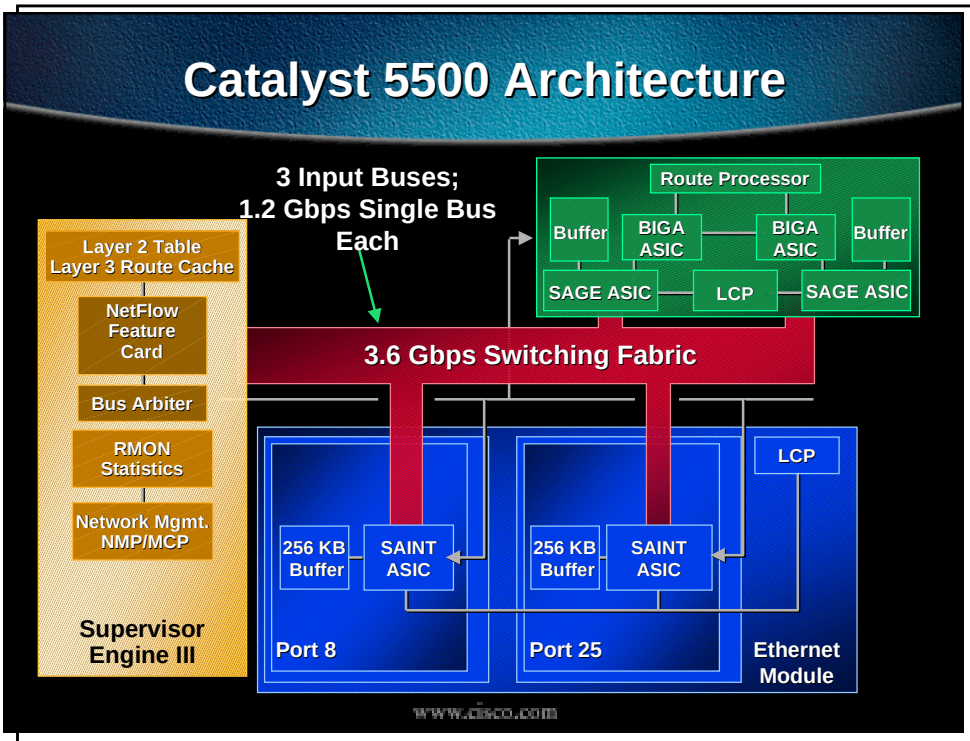
- Architecture Responsibilities
- Queuing Models
- Switching Implementations
- Switching Fabrics
- Example—Catalyst 4000 Series
- Example—Catalyst 8500 Series
- Example—Catalyst 5000 Family
- Example—Catalyst 6000 Family

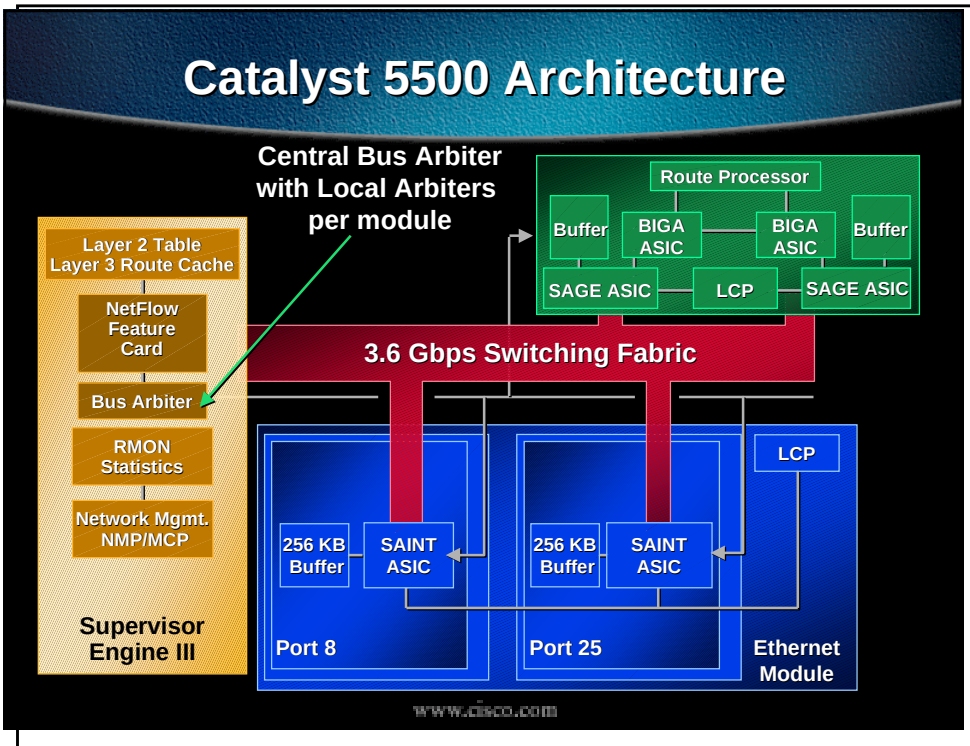
www.cisco.com

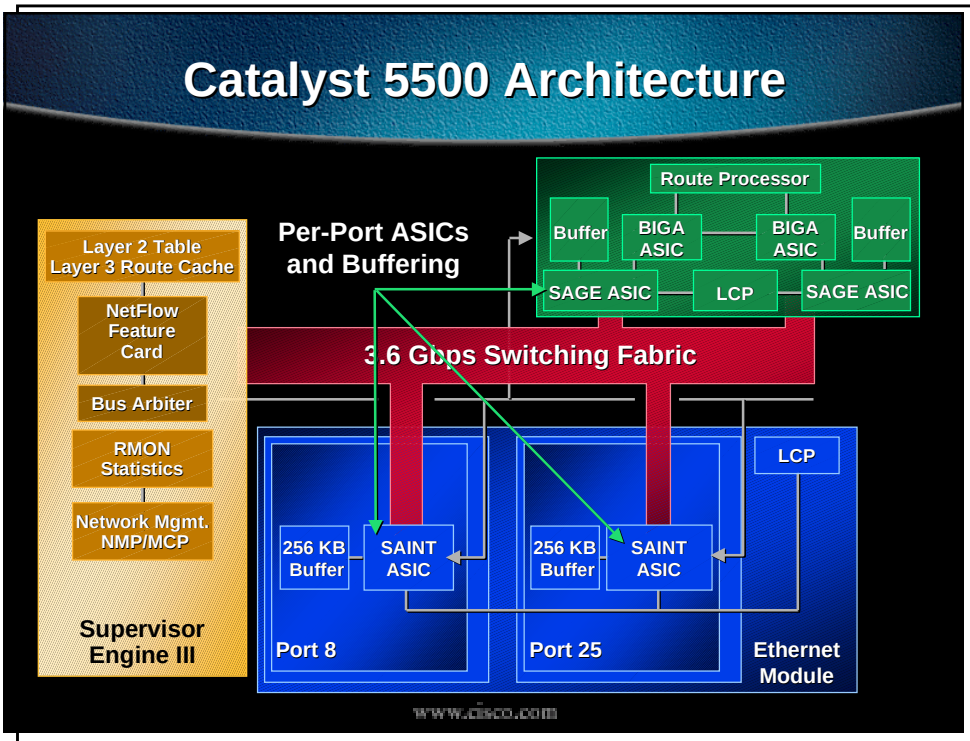
The Catalyst 5000 Family

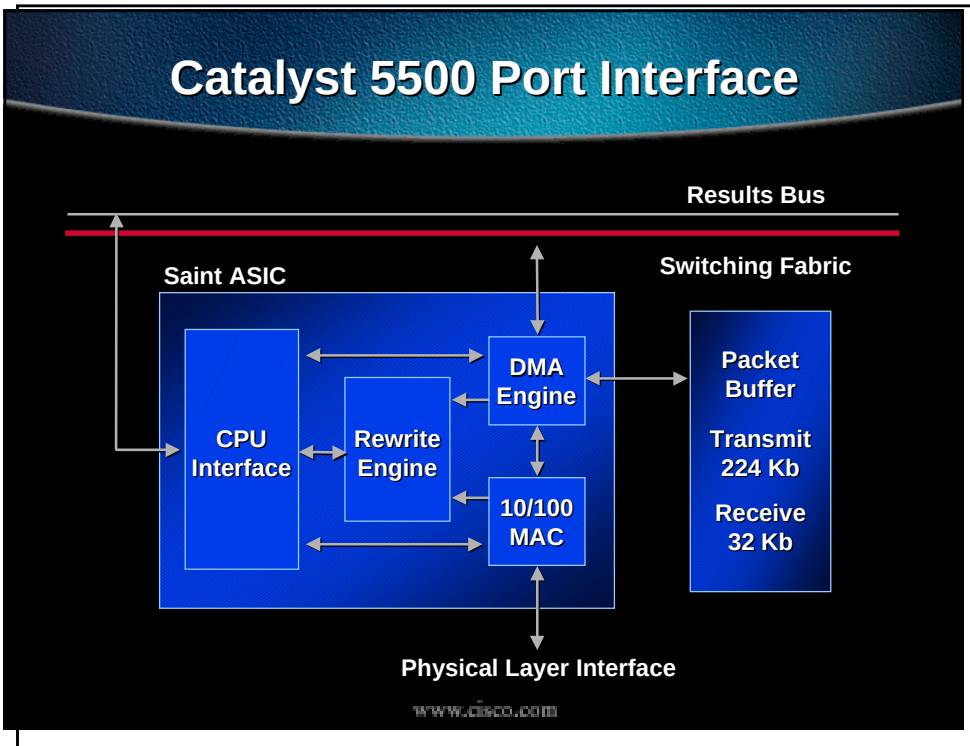
- Powerful switching solutions:
  - Multilayer, multiprotocol switching
  - Complete Cisco IOS network services support
  - Integration of gigabit/ATM/Layer 3 on a common platform
  - High-density LAN aggregation

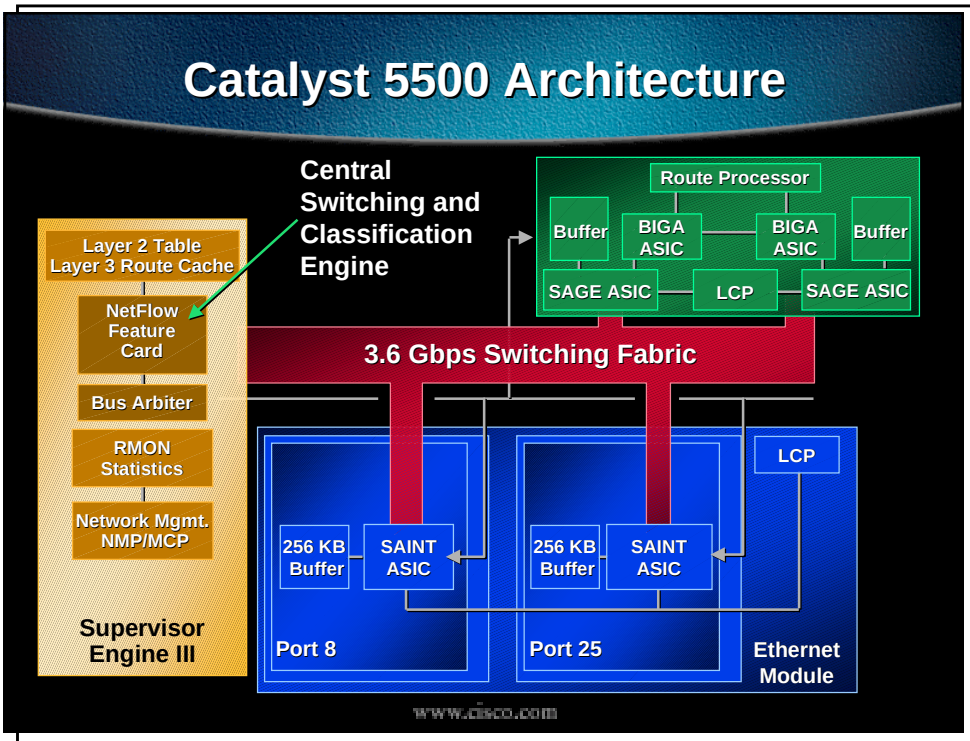
www.cisco.com

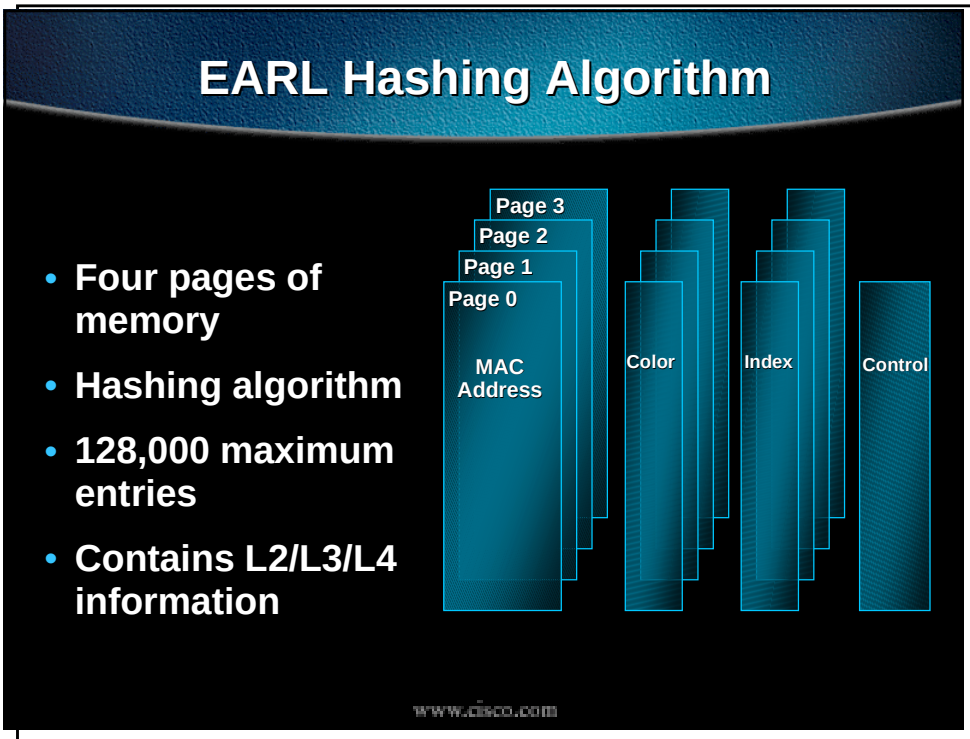












Components of Catalyst 5500 Multilayer Switching (MLS)

MLS-SE—Switch Engine

MLS-RP—
Route Processor

or

Cisco
7500 Series
7200 Series
4000 Series
3600

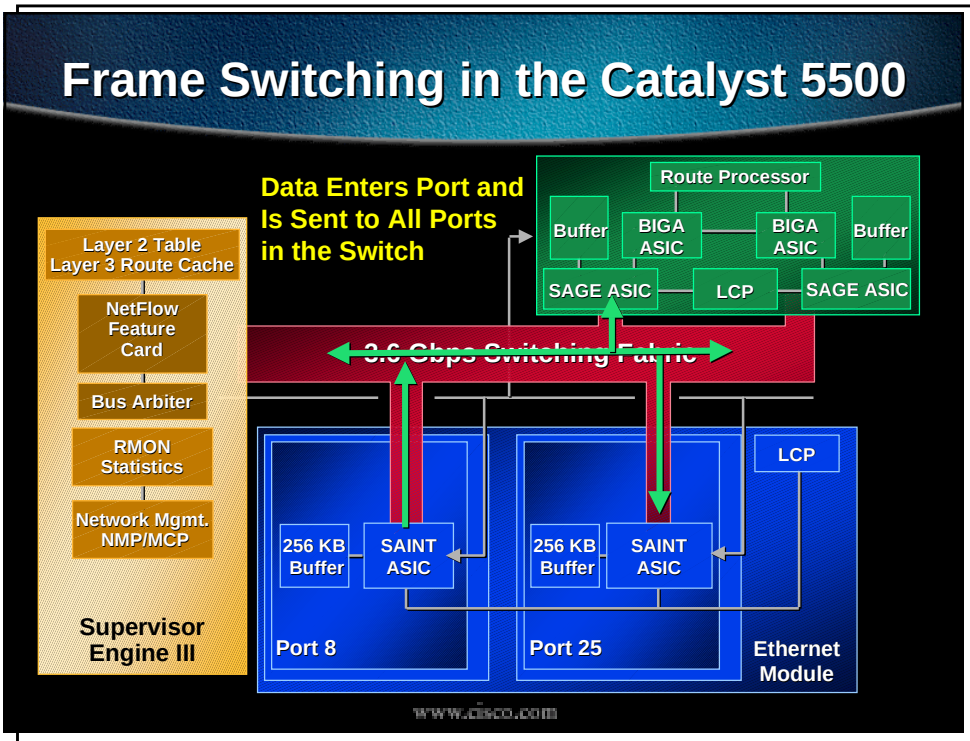
MLSP—Multilayer Switching Protocol
for Router Registration with MLS-SE

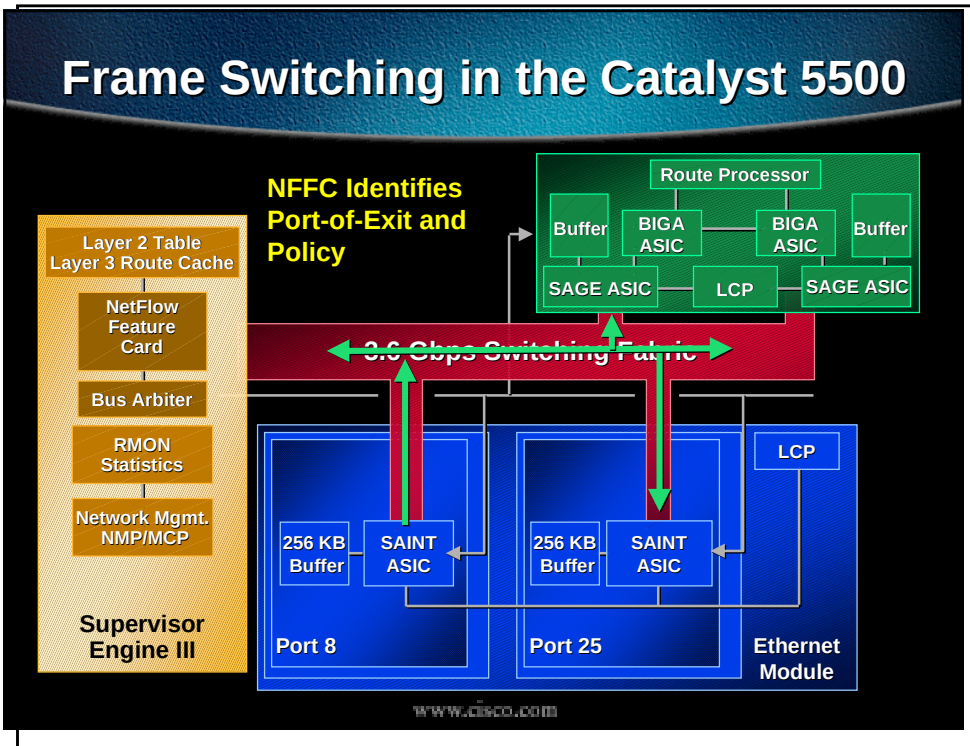
www.cisco.com

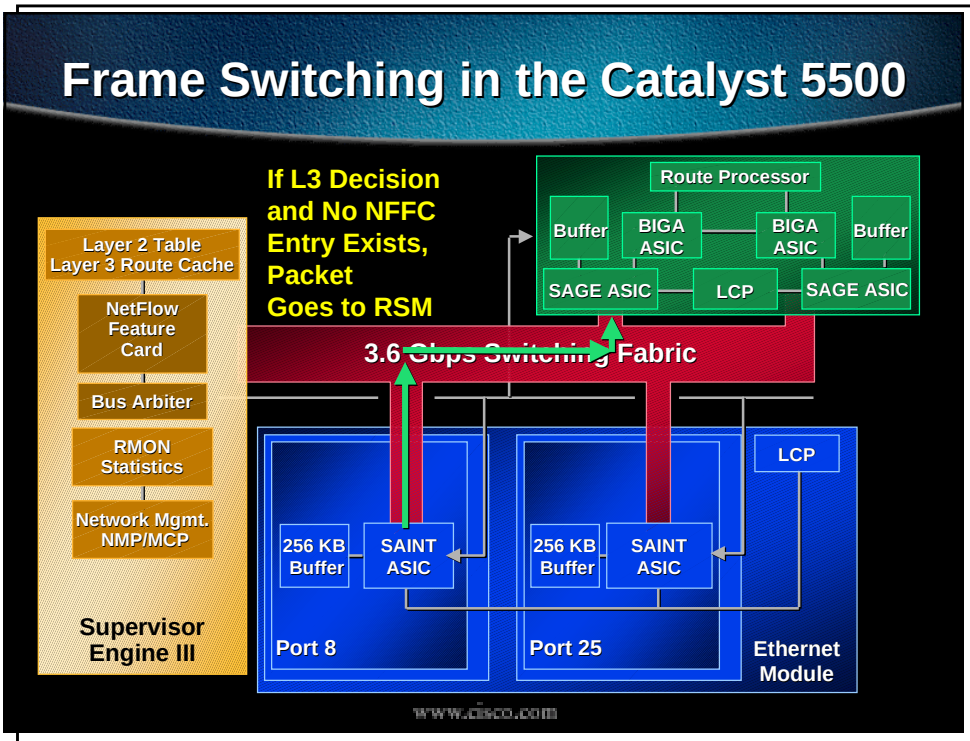
Multilayer Switching Protocol (MLSP)

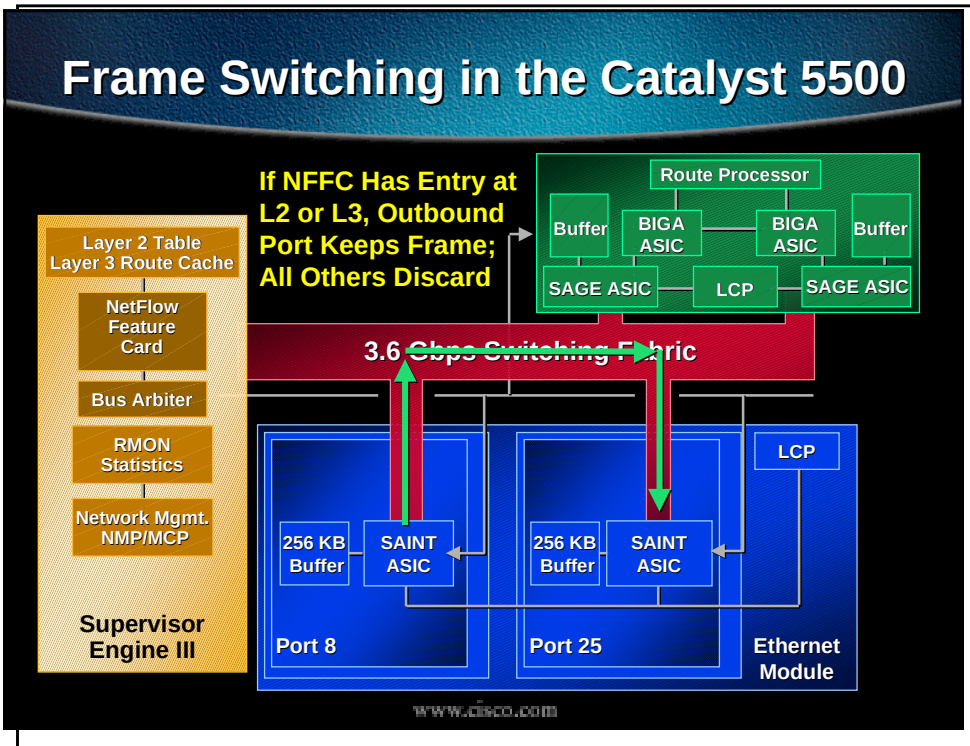
- Allows RSM, Cisco 7500, Cisco 4000, Cisco 3600 and Cisco 7200 series routers to provide route processing functionality
- Used for router registration and services (i.e., access list) updates for NFFC
- Not used in Layer 3 switching itself

www.cisco.com

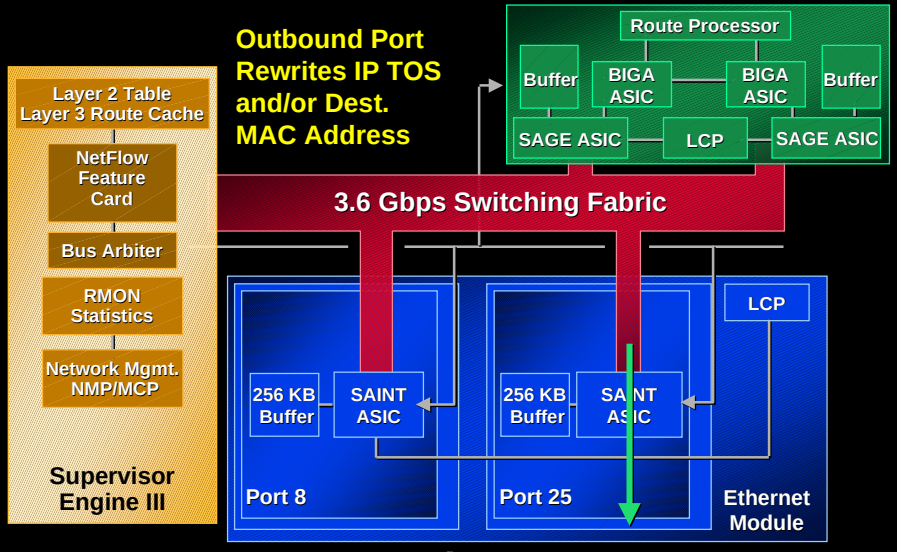








Frame Switching in the Catalyst 5500



Catalyst 5000 Family Layer 2 Performance

- Forwarding rate: 10 Mbps
  - 2,200,000 pps aggregate
- Forwarding rate: 100 Mbps
  - 2,200,000 pps aggregate
- Forwarding rate: 1000 Mbps
  - 35.7 million pps (with local switching)
- Multicast forwarding rate: 100 Mbps
  - 39.2 million pps

www.cisco.com

Catalyst 5000 Series Layer 3 Forwarding Rates

- Route Switch Module (RSM)—175,000 pps (based on RSP-2)
- Route Switch Feature Card (RSFC)—170,000 pps (based on NPE-200)
- NetFlow Feature Card (NFFC-II)—2.0 million pps

Agenda

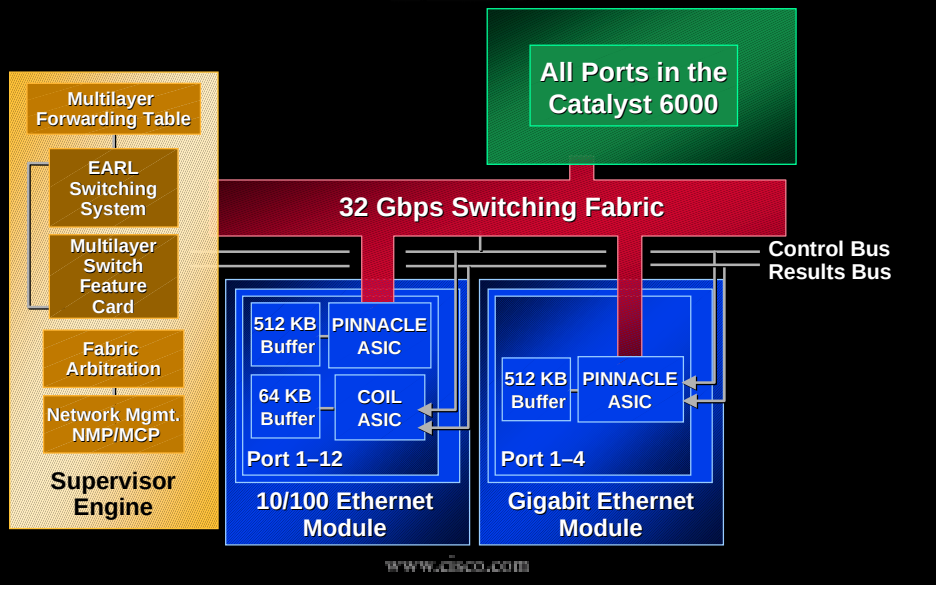
- Architecture Responsibilities
- Queuing Models
- Switching Implementations
- Switching Fabrics
- Example—Catalyst 4000 Series
- Example—Catalyst 8500 Series
- Example—Catalyst 5000 Family
- Example—Catalyst 6000 Family

The Catalyst 6000 Family

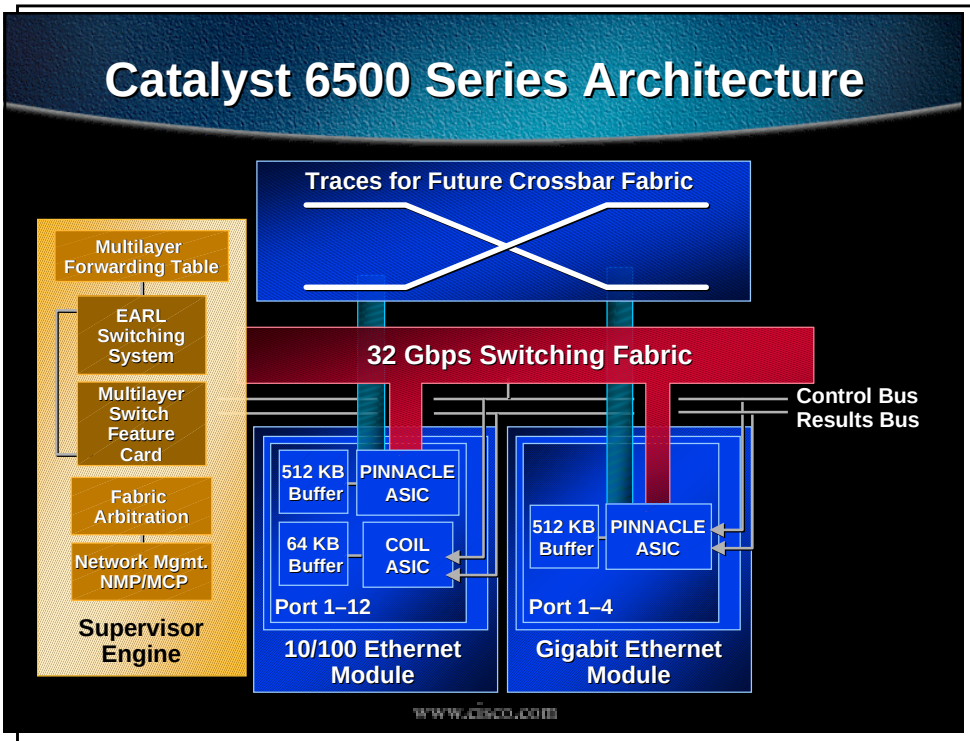
- Next Generation switching platform
- Provides support for
  - High Speed Switching
  - Cisco IOS features and functionality
  - Web Hosting features such as Server Load Balancing, WCCP, Security
  - Voice Support via powered line cards, analog ports, DSP functionality.

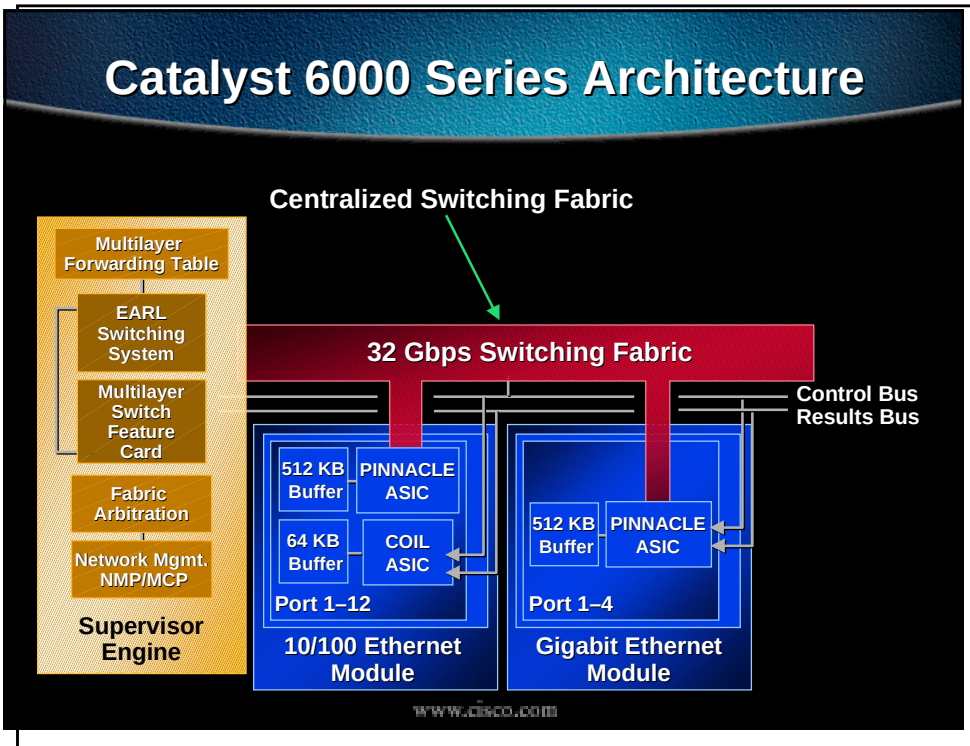
www.cisco.com

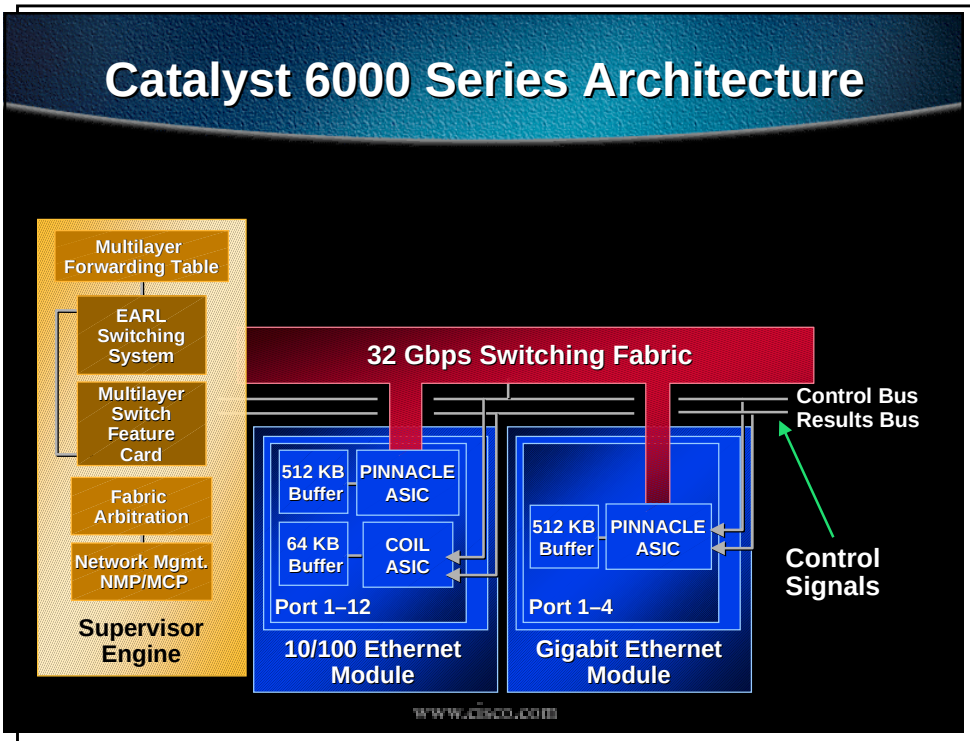
Catalyst 6000 Series Architecture

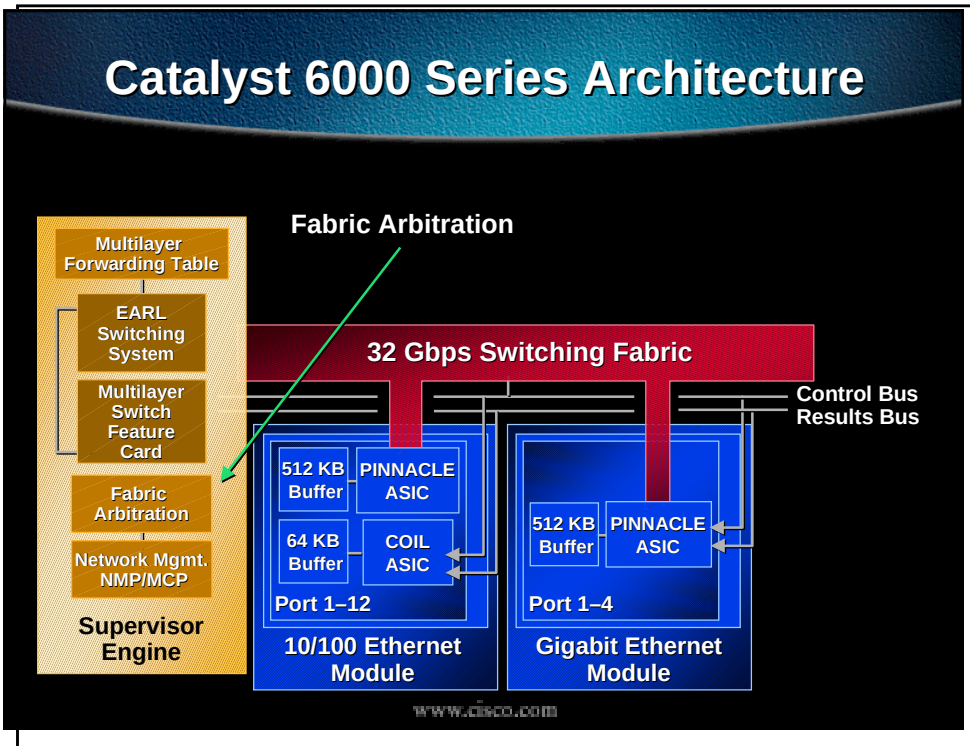


www.cisco.com

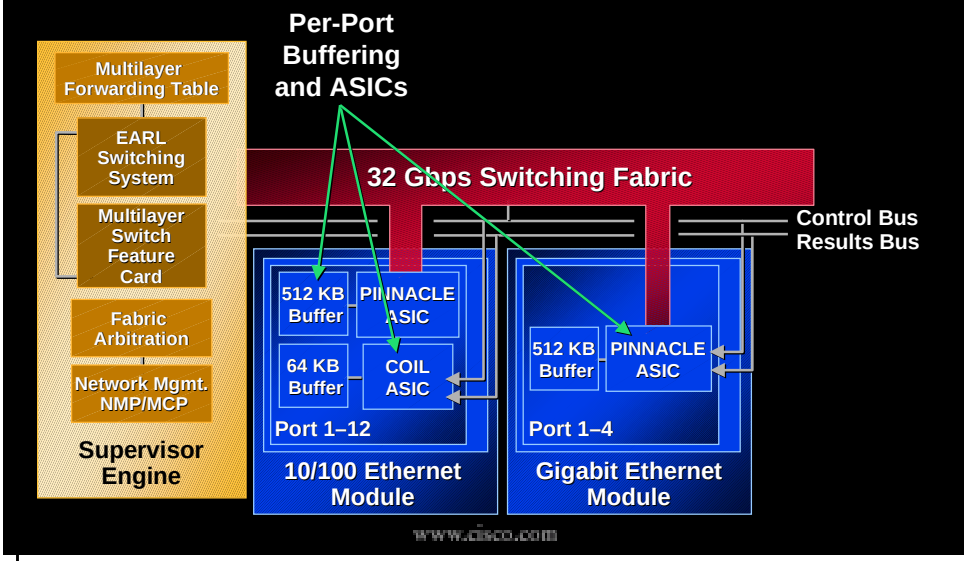








Catalyst 6000 Series Architecture

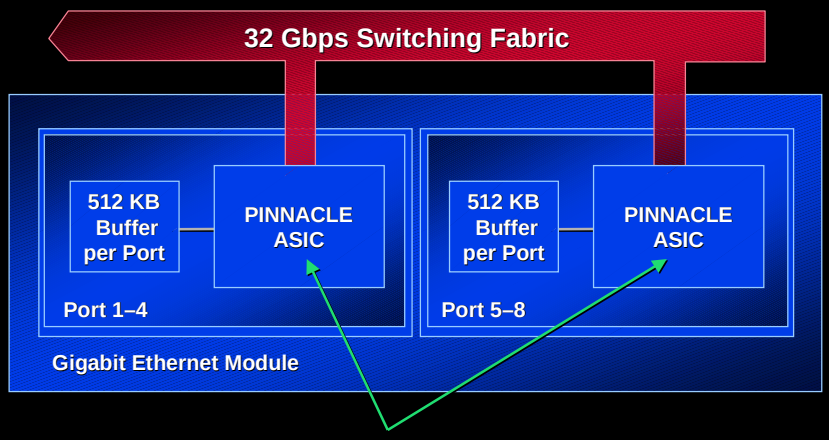


Port ASIC Built-in Features

- InterSwitch Link (ISL) and 802.1Q
- In-line rewrite capability for Layer 3 and QoS reclassification
- Fast and Gigabit EtherChannel
- Four groups of RMON per port
- Two queues per port
- Four drop thresholds per port

www.cisco.com

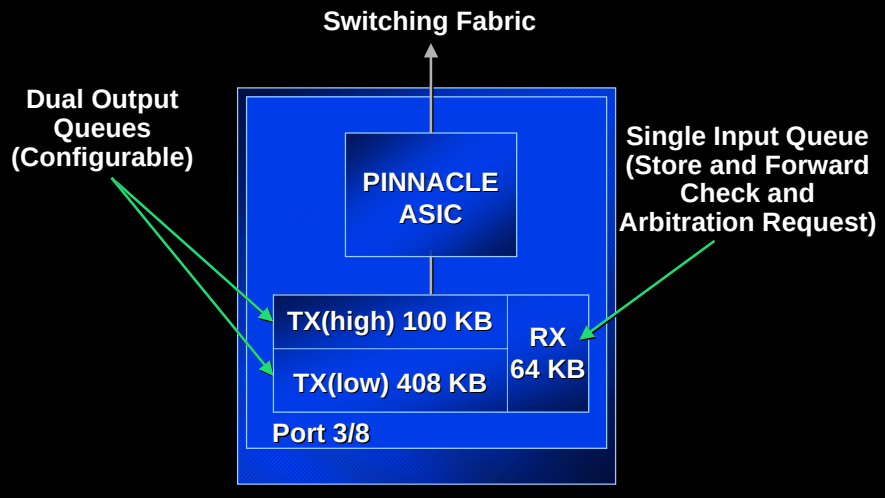
Catalyst 6000 Port ASICs—Gigabit Ethernet



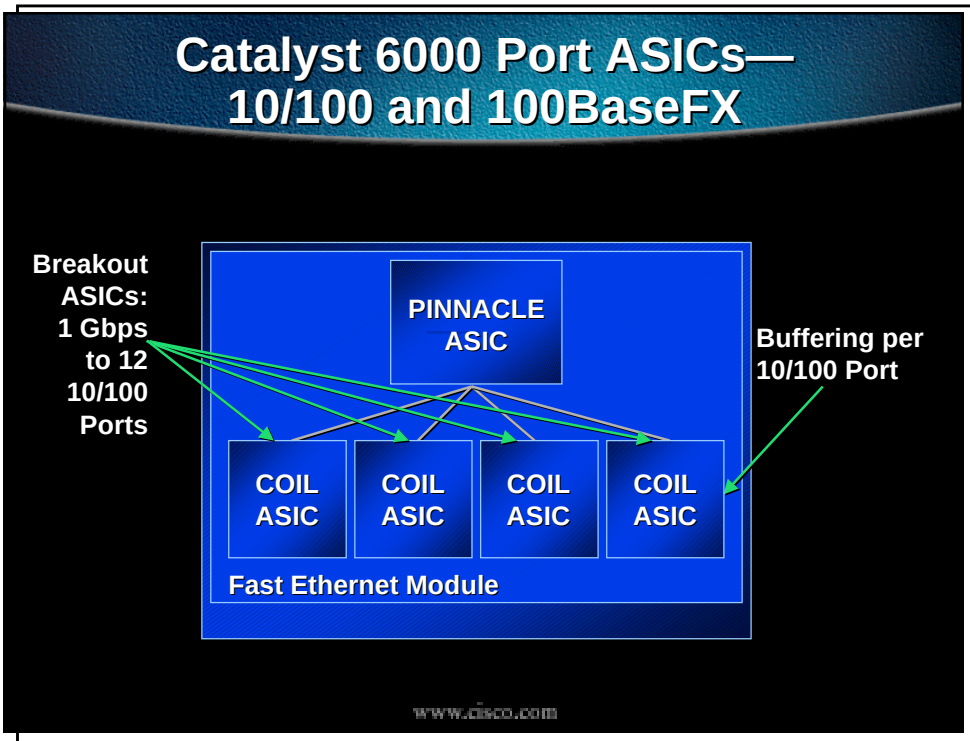
PINNACLE Controls DMA Access into Per-Port Memory and Performs Arbitration Requests

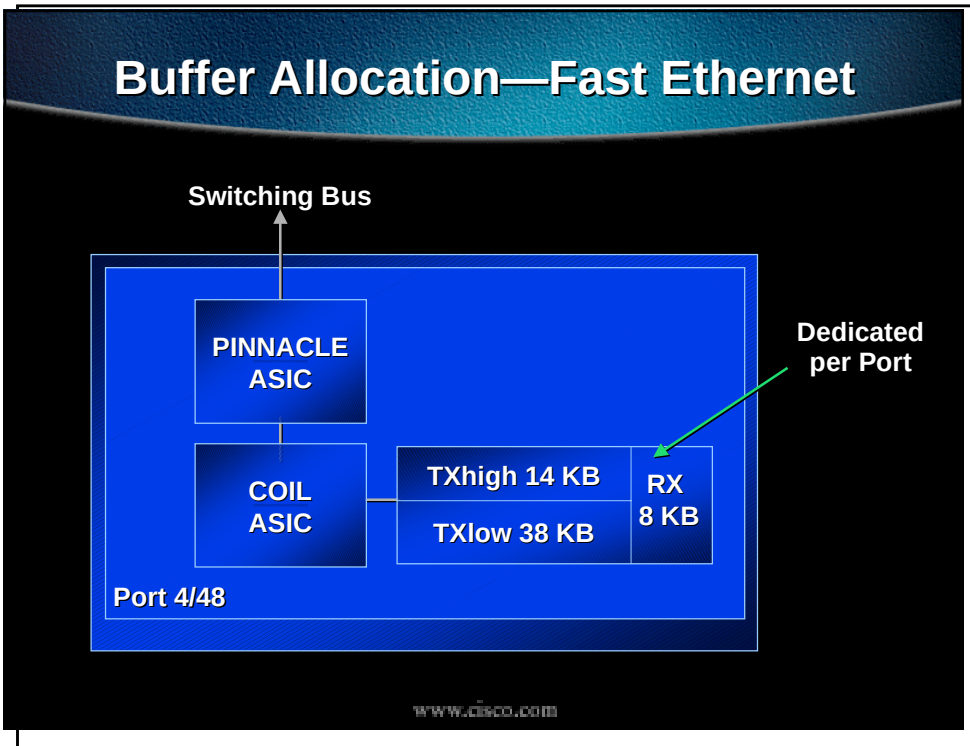
www.cisco.com

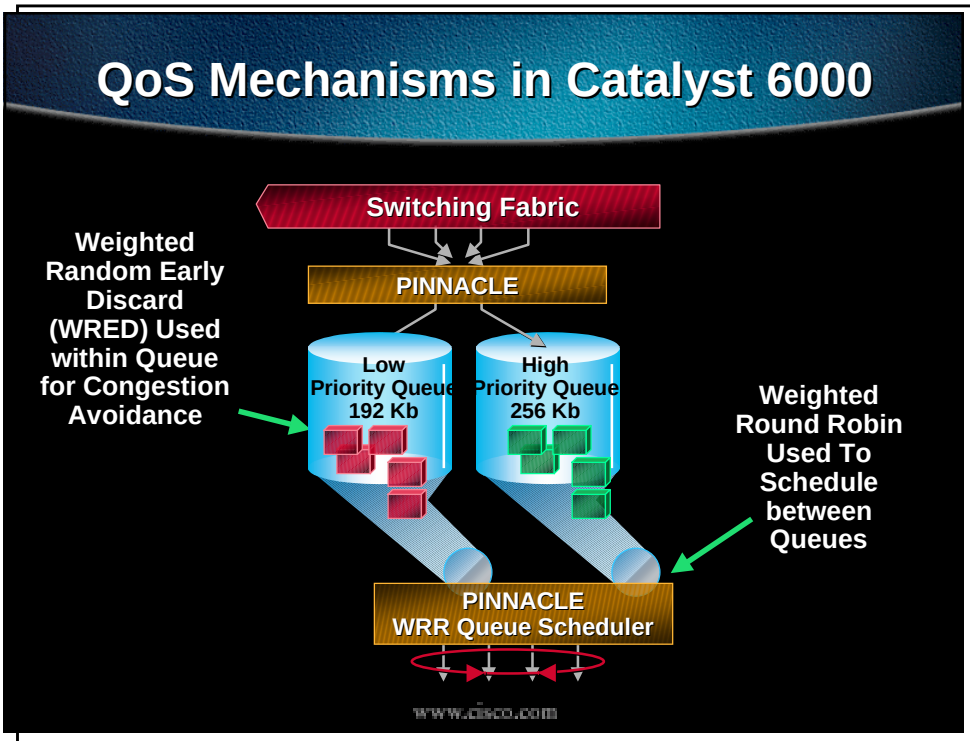
Buffer Allocation—Gigabit Ethernet

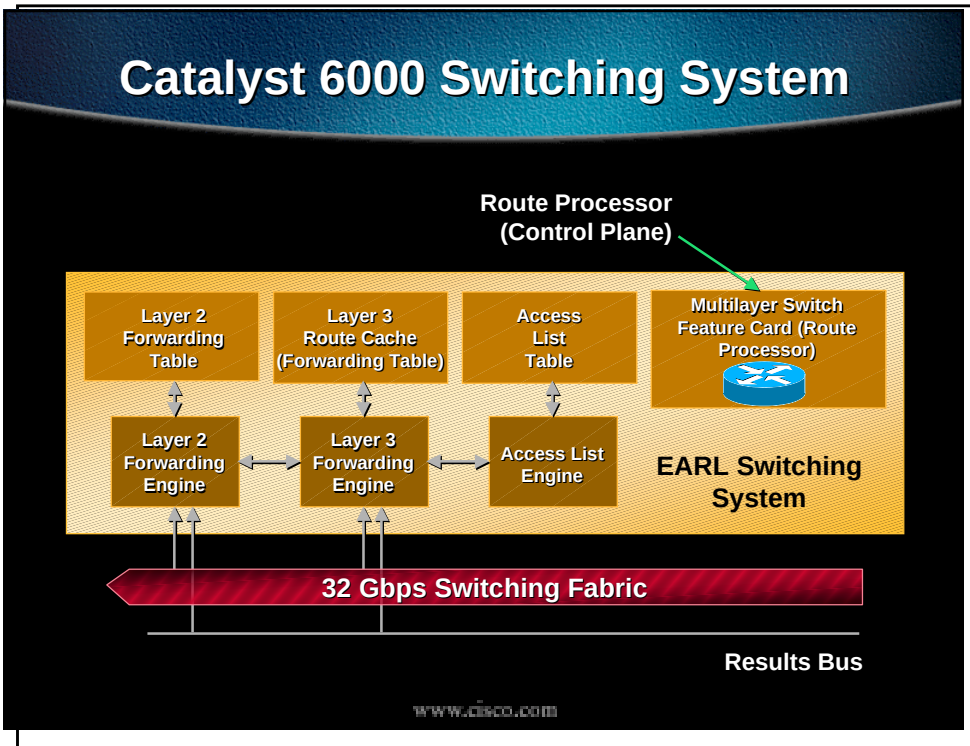


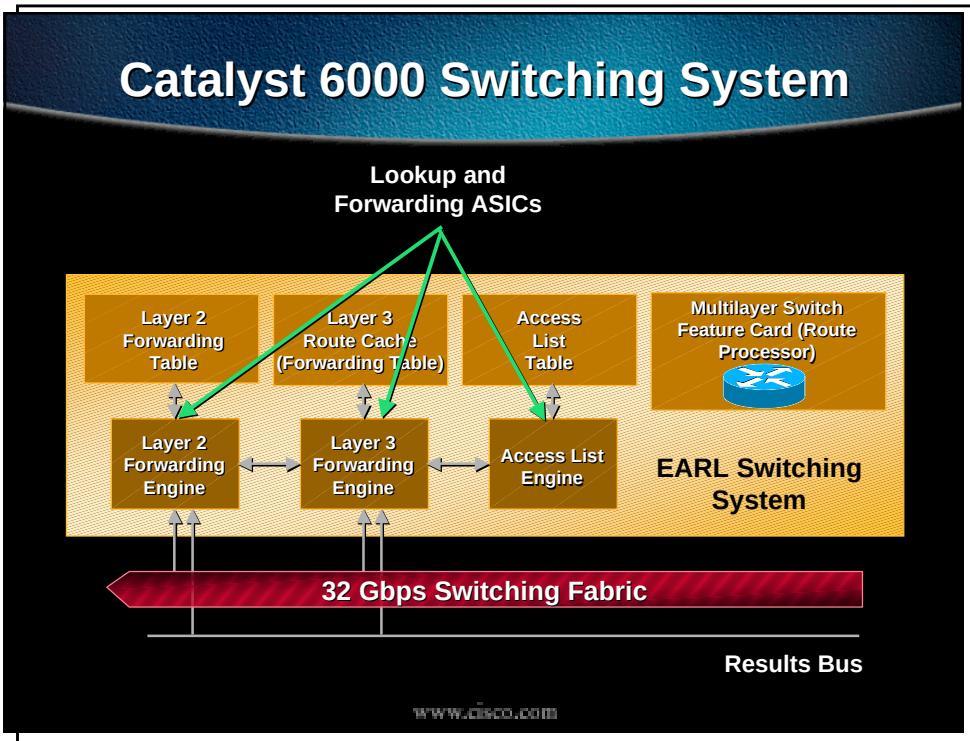
www.cisco.com

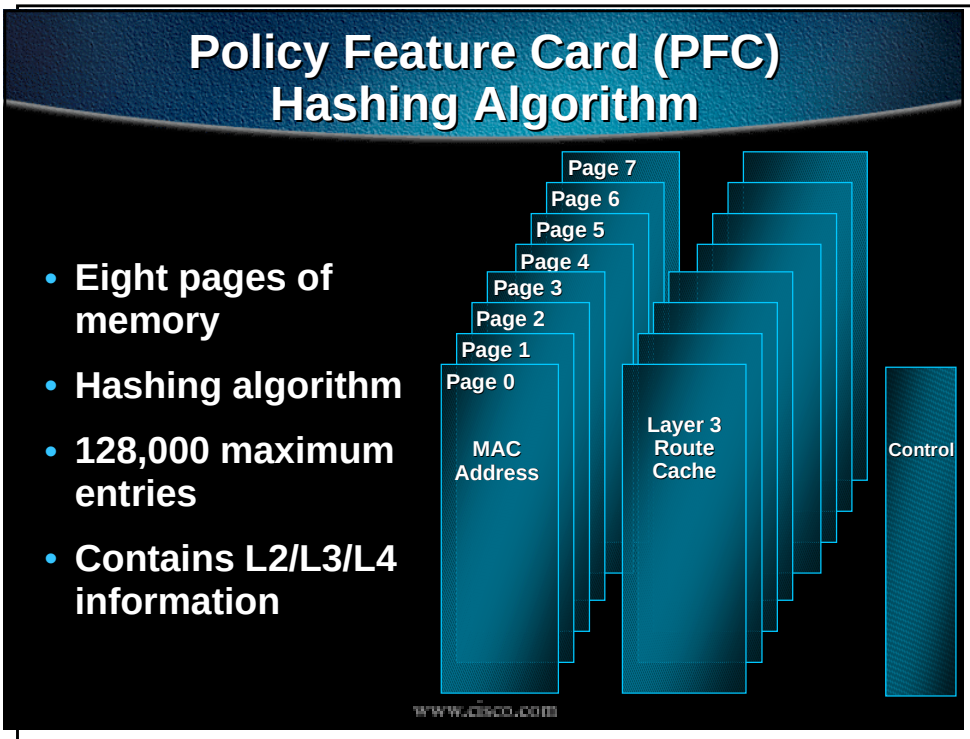


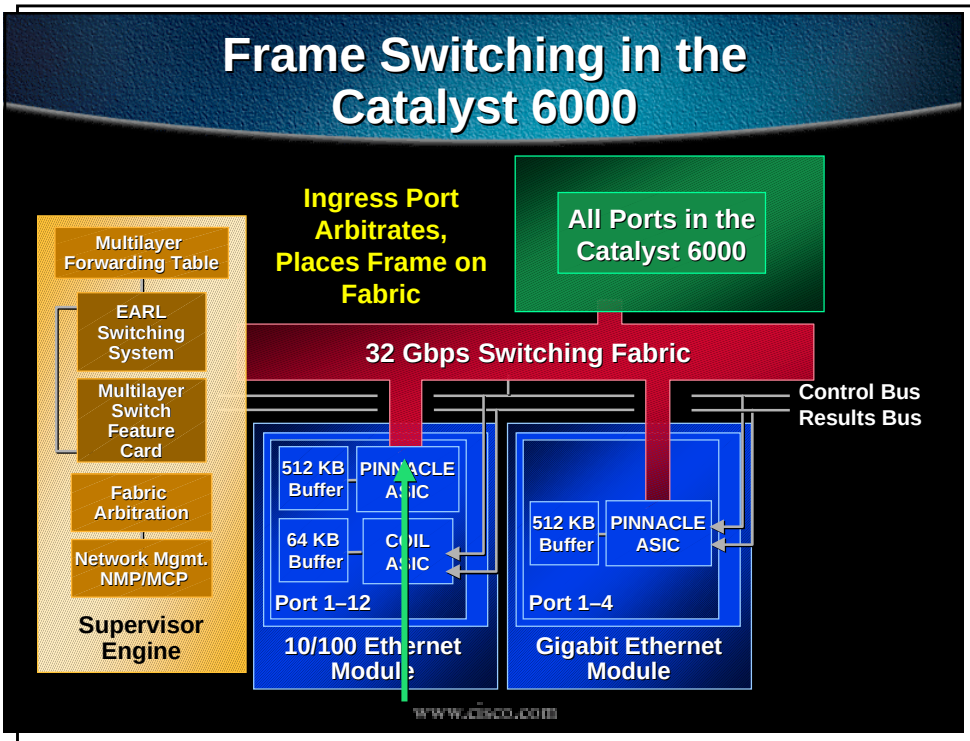


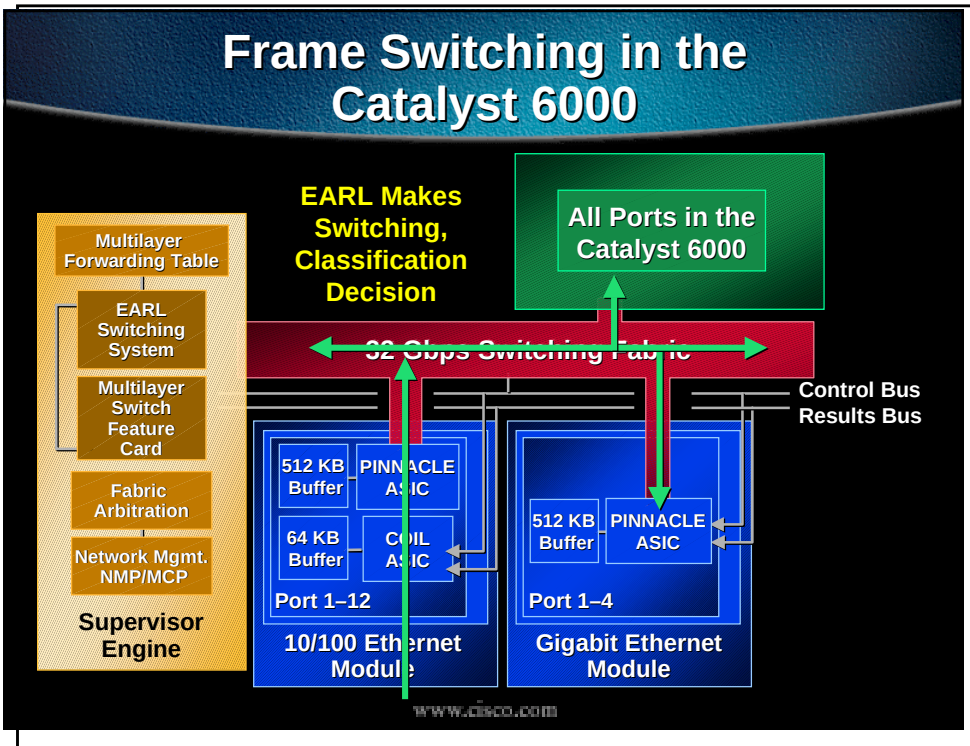




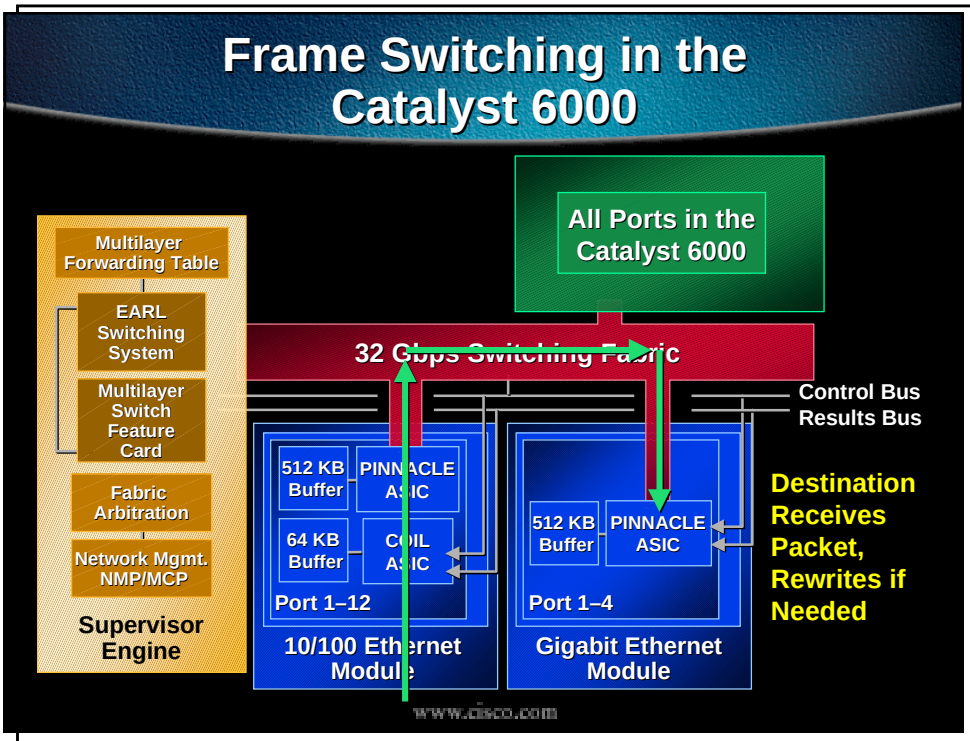








Frame Switching in the Catalyst 6000



Catalyst 6000 Performance

- Forwarding rate: 100 and 1000 Mbps
  - Layer 2: 15 million packets per second
  - Layer 3 (MSFC): 15 million packets per second

www.cisco.com