Networkers at Cisco live!

July 22–26, 2007 Anaheim, CA

CISCO

Catalyst 3750/ 3750-E and 3560/3560-E Architecture

BRKRST-3437

Questions We Will Answer Today

- The Differences between 3560/3560-E/3750/3750-E
- What is a stack ring?
- How is the stack ring controlled?
- How does the hardware work?
- How are stack processes controlled?
- What happens when I mix different switch types?
- How does QoS work?

3560-E

3750-E

3560-E

3750-E

Agenda

Differences Between Models

- 3560 and 3560-E standalone


3560


3560-E

- 3750 and 3750-E stackable


3750


3750-E

- Stackable means that it has Cisco's proprietary stacking cables running the StackWise and/or StackWise Plus Protocols
- Other than stackable features the 3750 and 3560 are identical
- Other than stackable features the 3750-E and 3560-E are identical
- E series and non-E-Series have some feature differences. These are outlined on the following slides

Main E-Series Feature Differentiators

- Dual 10GE Line Rate Uplinks, with Twin Gig SFP adapters, 2 per 10G slot
- StackWise Plus increases the effective stacking throughput to Nx64Gbps using spatial reuse, and also supports StackWise Mode
- Hardware-based per port power monitoring/policing
- Field replaceable power supplies, supports up to 48 ports POE at 15.4 watts
- On-Board Failure Logging (OBFL)
- Jumbo frame L3 routing
- IPv6 Multicast Routing
- Additional 10/100 management interface
- Universal Software Image
- *Detailed descriptions provided in the appendix


3750-E


3560-E

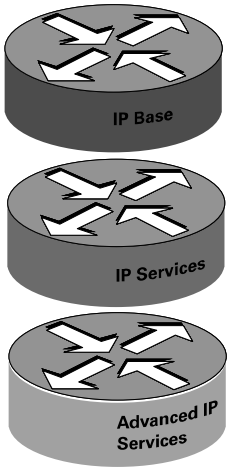

3750


3560

Images for Non-E-Series

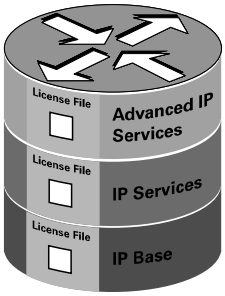
- There are 3 Images available for non E-series switches.
  - IP-Base (L2,Stub routing, IP ACLs)
  - IP Services (Full L3 Routing and Multicast routing)
  - Advanced IP services (IPv6 Routing)

3 Distinct Images



One Universal Image For E-Series

- A “Universal” IOS image contains all IOS features
- Licensing enables a specific level of IOS functionality
- Customers only upgrade their license to upgrade functionality
- A Universal IOS image is loaded in manufacturing



Universal IOS Image

License Installation (Upgrade)

- **Obtain license file through CCO or CLM**
  - Unique Device Identifier (UDI) is required along with a token Product Activation Key (PAK)
  - UDI codes the serial number and the product ID.
  - PAK is the proof of purchase
- **TFTP license file into flash**
- **Install the license using the "license install" command**

```
Switch# copy tftp flash:
Address or name of remote host [ ]? 172.20.244.138
Source filename [ ]? r1fs-ips
Destination filename [r1fs-ips]?
Accessing tftp://172.20.244.138/r1fs-ips...
Loading r1fs-ips from 172.20.244.138 (via GigabitEthernet1/0/1): !
[OK - 1161 bytes]
1161 bytes copied in 0.059 secs (19678 bytes/sec)

Switch# license install flash: r1fs-ips
    Installing licenses from "flash:r1fs-ips"
    Installing...Feature:ipservices...Successful:Supported
    1/1 licenses were successfully installed
    0/1 licenses were existing licenses
    0/1 licenses were failed to install
```

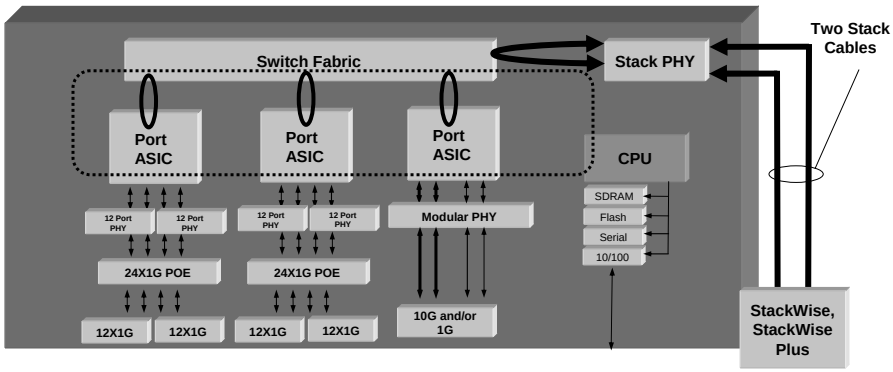

Show Commands

- Following administrative commands will be supported to administer software licensing:
  - Displaying the file
  - Show and detailed show of license type
  - Showing the Unique Device Identifier
  - Enabling debug mode

```
show license file [switch  <switch_id>]

show license status [switch <switch_id>]

show license detail <feature_name> [switch
<switch_id>]

show license udi

debug license <events | all | errors>
```

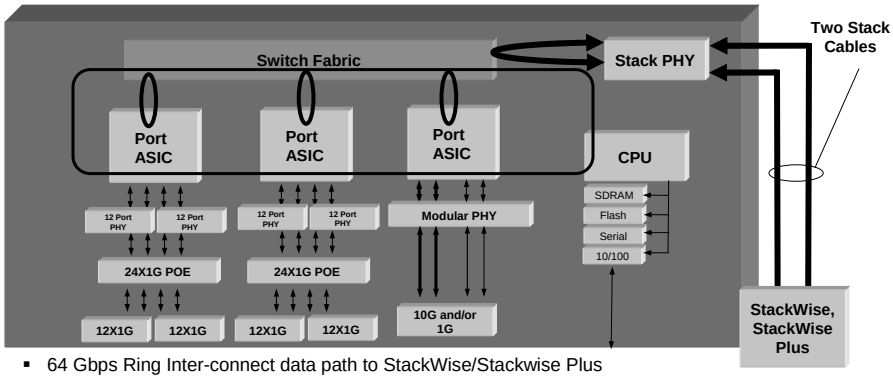

Agenda

Catalyst 3750-E Architecture Overview: Processor



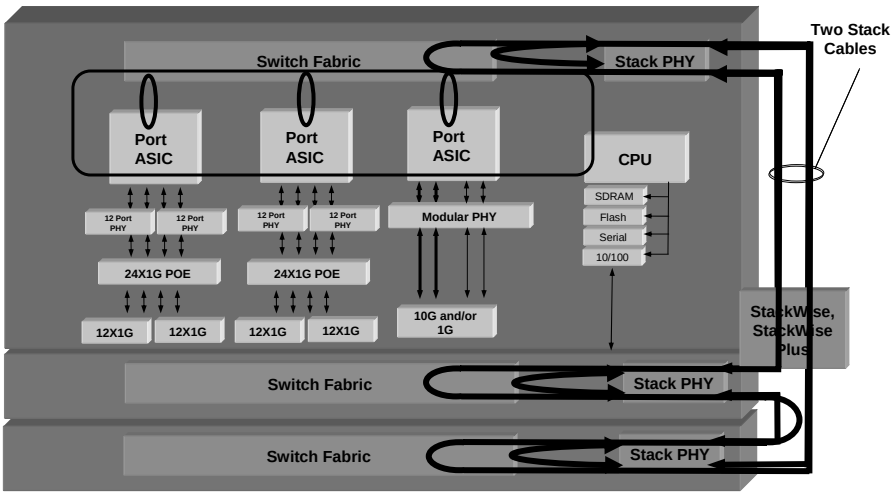
- Switch-to-Switch communication and synchronization
- Updates the MAC and Routing caches attached to each port ASIC
- Performs CPU-based slow-path forwarding when the TCAM is over its limits for MACs, Routes, ACL entries etc.
- The CPU communicates with the Port ASICs via a dedicated management 1G ring (the yellow dotted ring in the diagram)

Catalyst 3750-E Architecture Overview: Switch Fabric



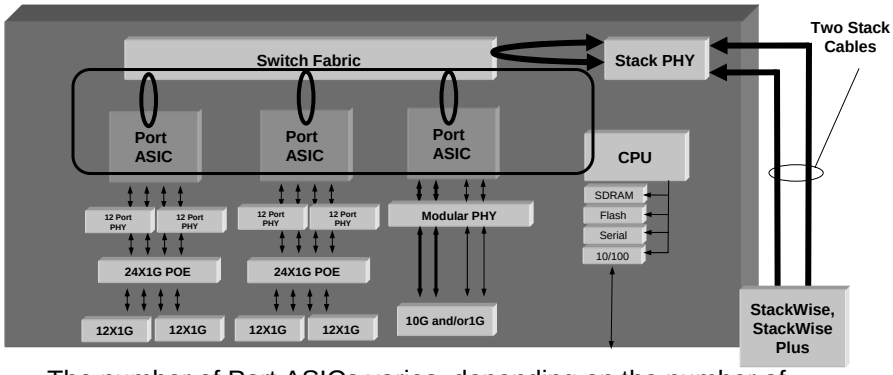
- 64 Gbps Ring Inter-connect data path to StackWise/Stackwise Plus
- 1 Gbps Ring Inter-connect control path to the Port ASICs to the CPU
- 1 P2P, 24 Gbps ring connecting each Port ASIC
- Provides line rate local switching within a switch and line rate stack connectivity
  - Local:48XG + 2X10G traffic can be local switched (68G)
  - Stack Ring: 64 Gbps
  - CPU Ring: 1 Gbps
- Performs local switching between Port ASICs that are connected to it without using Stackwise Plus resources

Catalyst 3750-E Ring View



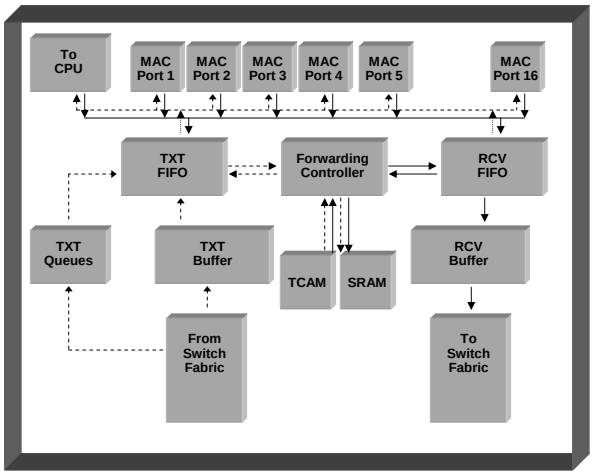
- Physically, the Ring Is a Series of Switch Fabrics Strung Together by Stack Cables
- The Switch Fabric performs token generation and ring control

Catalyst 3750-E Architecture Overview: Port ASIC



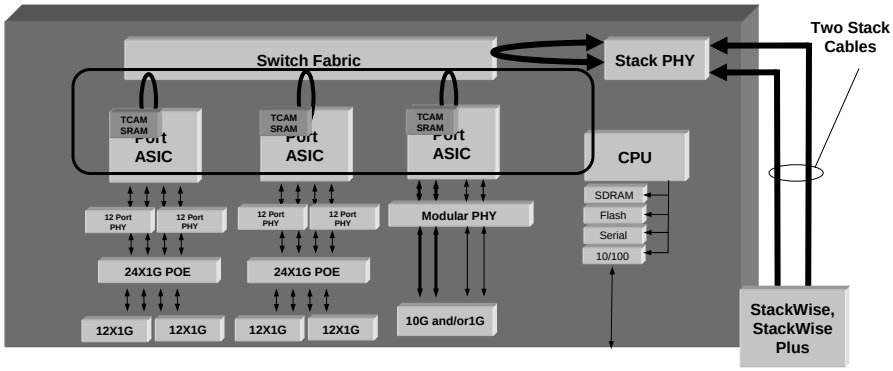
- The number of Port ASICs varies, depending on the number of ports
- The Port ASIC performs:
  - Traffic forwarding
  - QoS
  - ACL lookup

Port ASIC Architecture Exposed



Greater Detail Listed In the Appendix

Catalyst 3750-E Architecture Overview: TCAM/SRAM

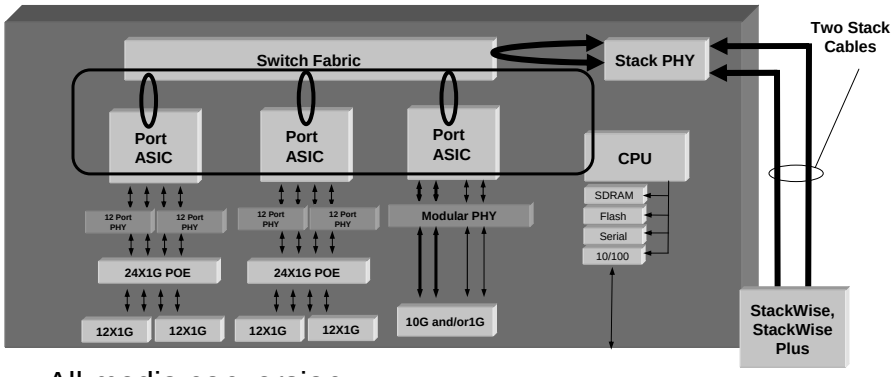


- The TCAM stores vital information including IPv4, IPv6 and MAC addresses
- With the 3750-E one can now perform a simultaneous IP and MAC lookup with one ACE
- With the 3750-E it is now easier to configure the full 2K ACEs

TCAM Templates

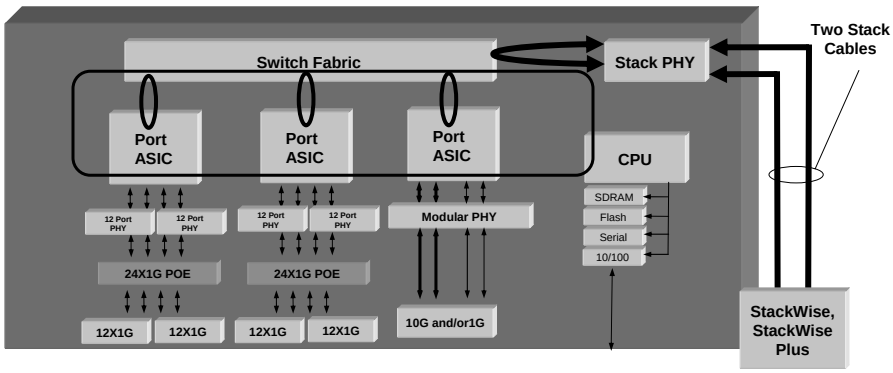
```
Switch# show sdm prefer routing

"aggregate routing" template:
The selected template optimizes the resources in
the switch to support this level of features for
8 routed interfaces and 1024 VLANs.
number of unicast mac addresses: 6K
number of igmp groups + multicast routes: 1K
number of unicast routes: 20K
number of directly connected hosts: 6K
number of indirect routes: 14K
number of policy based routing aces: 512
number of qos aces: 512
number of security aces: 1K
```


Catalyst 3750-E Architecture Overview: PHY



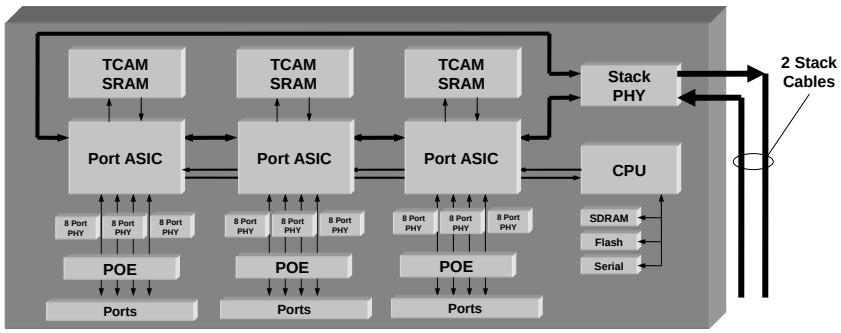
- All media conversion
- 10/100 Mbps
- 10/100/1000 Mbps
- 10 Gbps

Catalyst 3750-E Architecture Overview: POE



- 24 X 1G ports per POE per chip
- Terminates all power to/from the PHY
- Performs per port auto-sensing and controls all POE

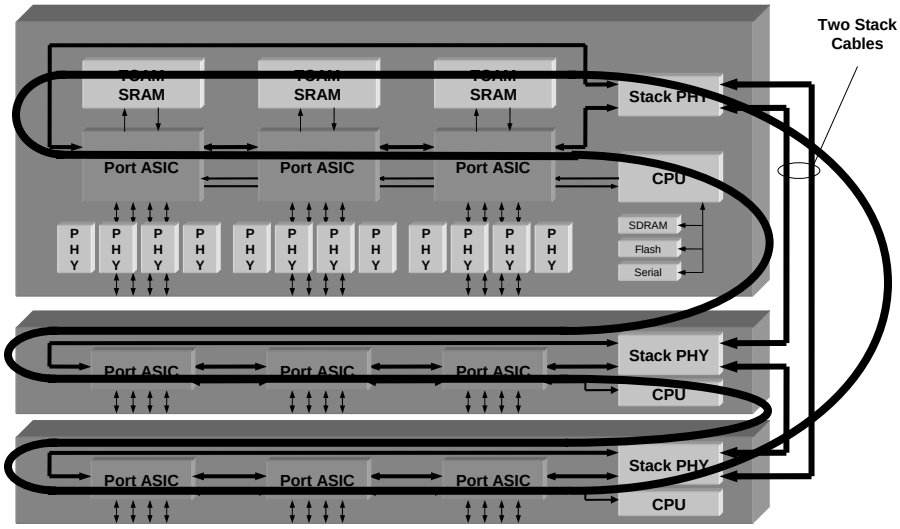
Catalyst 3750 Hardware Differences:
Block Diagram—48port POE



3750 and 3750-E Main Architectural Differences:

- 3750 Does not have a second tier switch fabric like the 3750-E and can not locally switch without sending packets on the ring
- 3750 has external TCAMs
- The 3750 does not have an Ethernet Management Port
- 3750 only runs in StackWise mode

Catalyst 3750 Ring View



- Physically, the Ring Is a Series of Port ASICs Strung Together by Stack Cables

Agenda

What Is the Stack?

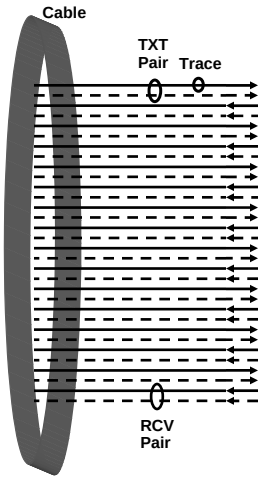
- The Cisco Catalyst 3750/3750-E switch is a switch that when stacked together forms a seamless single device
- This is made possible by Cisco StackWise and StackWise Plus
- The term “stack ring” is used because the stacking configuration is a true ring
- The stack interfaces form a hardware-based ring
- A hardware ring is beneficial because:
  - Non-ring stacks must block, just like spanning tree, or loops will occur and melt down the stack
  - A software ring would require the CPU to forward and this would result in dismal performance
  - Makes sure only one copy of a multicast packet is on the stack cables
- There are statistics and a MIB for stack ring functions

Stack MIB

MIB Object Type	Object Description
SET	• Enable stack notification
GET	• Max number of switches in the stack • Highest switch priority that can be configured • Indicates if the stackports are connected such that ring redundancy is available • List of switches in the stack • Current switch number and next switch number after next reload • Switch role in the stack • Switch priority • Switch state (for example, waiting, progressing, added, and so on) • Switch MAC • Switch image • Switch stackport info • Switch stackport neighbor • Switch stackport status
TRAP	• Switch stackport state change • New master elected • Stack mismatch for a new member joining • Stack ring redundancy change • New member added • Member removed

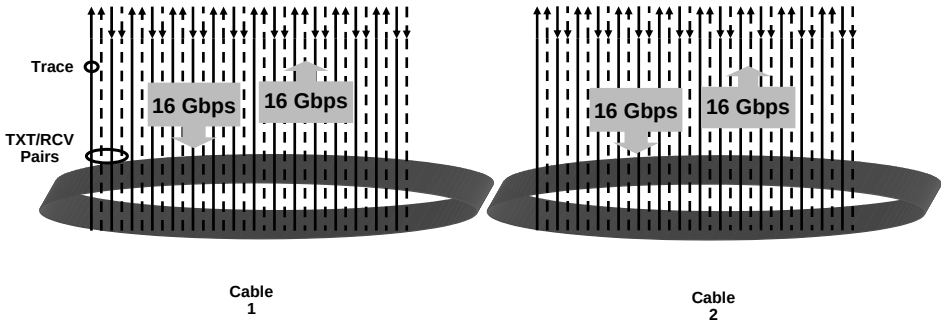
Understanding the Stack Cable

- Eight TXT/RCV pairs, that is 16 total pairs
- Each TXT/RCV pair has two traces that use differential signaling. That is 32 traces in total.
- Each TXT/RCV pair runs at 2.5 Gbps
- 8B/10B encoding is used. That is, for every ten bits sent, eight bits are user data and two bits are overhead



Understanding the Stack Ring Speed

- Two Cable X 16 Pairs/Cable X 2.5 Gbps/Pair X 8B/10B = 64 Gbps total
- Or 32 Gbps send and 32 Gbps receive total per cable
- Or 16 Gbps per cable bidirectionally

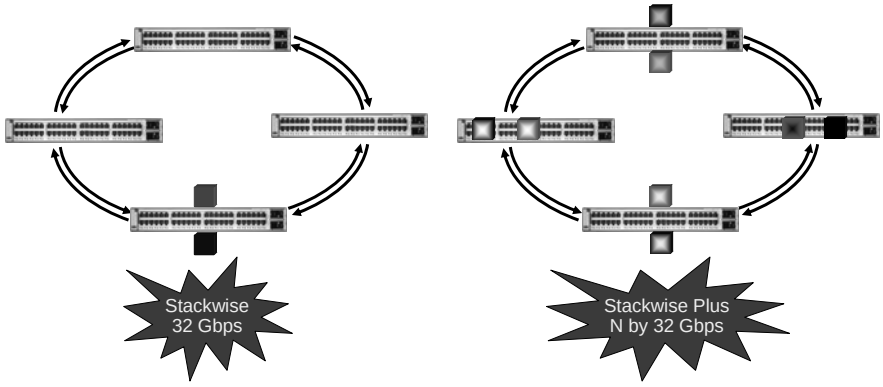


Physical Line Rate Only

Spatial Reuse
Stackwise Plus

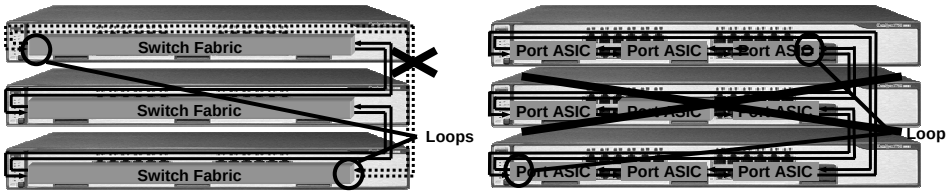
No Spatial Reuse 3750 StackWise
Only 2 Flows
Access-based tokens

Spatial Reuse 3750-StackWise Plus
Up to N by 2 Simultaneous Flows
Credit-based Tokens



Note: These are packets not tokens. There is are only 1 token per direction, 2 in total

Ring Healing

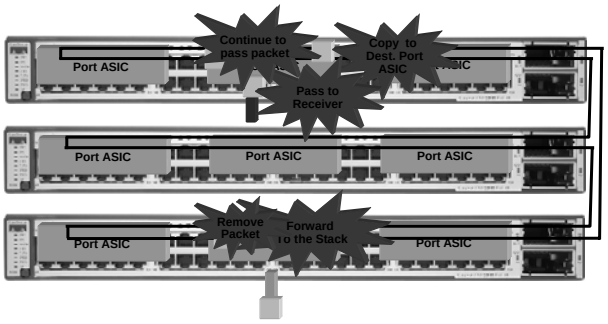


3750-E 3750

- The Switch Fabric or Port ASIC closest to cable detects link down
  - Criteria is coding violations in a period of time
  - Loss of at most one packet that was being transmitted when ring broke
  - Just microseconds for hardware to detect failure
- Each switch signals a bad link to stack its partner
- Both ends of the cable loop back on themselves

Agenda

3750 Packet Walk—All



- All types of packets are passed all the way around the ring, copied at the destination(s) and returned to the sender for stripping
- All packets are sent to the stack ring, the Port ASICs can not locally switch traffic

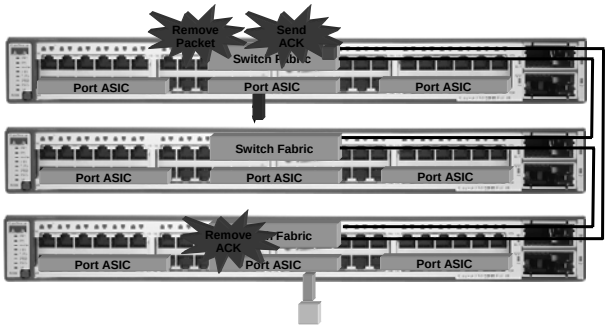
Source
Destination
Data

3750-E Unicast Packet Walk
Locally Switched

- The packet is sent to the switch Fabric and locally switched to the destination Port ASIC
- Simple switching with, no ACK necessary

Source
Destination
Data

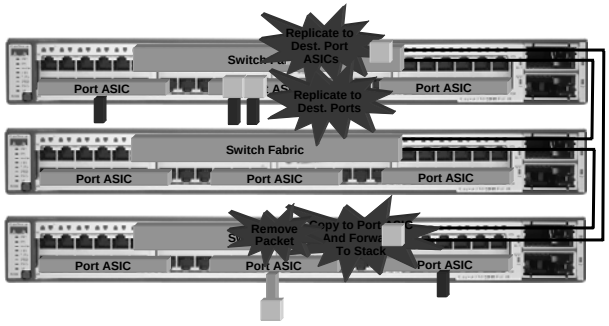
3750-E Unicast Packet Walk—Remote Destination



- The Source Port ASIC sends the packet to the Source Switch Fabric and it is switched to the Destination Switch Fabric
- The Destination Switch Fabric removes the packet and sends a 8 bit ACK
- The Originating Switch Fabric receives and removes the ACK

Source
Destination
Data
ACK

3750-E Multicast Packet Walk



- The packet is passed all the way around the ring
- The Switch Fabrics with multicast ports in that group copy the packet
- The originating Switch Fabric removes the packet from the ring
- Note: There is only one packet on the ring per multicast flow, replication only occurs at the local level
- Note: if the sender and all of the receivers are on the same switch no packets are sent to the ring

Source
Destination
Data

Agenda

- Switch Differences
- Hardware Overview
- StackWise Overview
- Packet Walks
- Stack Functions
- Configuration Management
- QOS Model
- Summary

Stack Master and Members

- A stack is created by connecting switches using Cisco proprietary Stacking Cable
- During the formation of stack, a stack master is elected
- All switches have the ability to be stack master—no special hardware/software required
- The stack master can be selected by assigning a user-configurable priority 1 through 15, 15 being the highest
- An LED indicates stack master
- The master controls all centralized functions
- On stack master failure, another switch in the stack takes over
- 1:N master redundancy
- All non-master switches are called members

Stack Master Election Criteria

- When adding switches or merging stacks, the master will be chosen based on the rules below, in the order specified

1. The stack (or switch) whose master has the higher user configurable mastership priority 1-15
2. The stack (or switch) whose master is not using the default configuration
3. The stack (or switch) whose master has the higher software priority
   - Cryptographic advanced IP services (IPv6)
   - Noncryptographic advanced IP services (IPv6)
   - Cryptographic IP services
   - Noncryptographic IP services
   - Cryptographic IP based
   - Noncryptographic IP based
4. The stack (or switch) whose master has the lowest MAC address

Switch Priority

```
Switch (config)# switch 3 priority 10

Switch (config)# exit

Switch# show switch

Switch# Role     Mac Address    Priority        State
-------------------------------------------------------------
1    Member 000a.fdfd.0100       5       Ready
2    Member 000a.fdab.0100       5       Ready
3    Master 000a.fd22.0100       10      Ready
4    Member 0003.fd63.9c00       5       Ready
```


Becoming A Stack Master

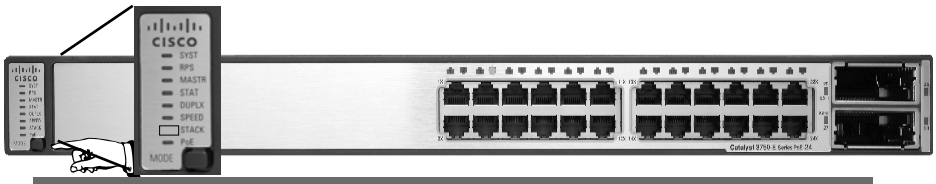
- A stack master can change if:
  - The stack master fails
  - The stack master is removed from the switch stack
  - The stack master is power cycled or powered off.
  - Stack membership is increased by adding one or more powered-on switches with a higher priority than the current master

Functions of the Stack Master

- The Stack Master:
  - Builds and propagates the L3 FIB
  - Propagates the configuration to the stack
  - Controls the console
  - Controls the CDP neighbor table
  - Controls the single VLAN database

Switch Numbers

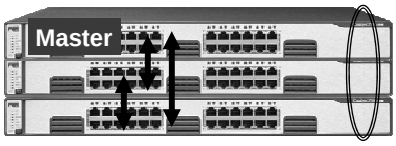
- Member switches, in a stack, are assigned switch numbers
- Valid switch numbers are 1 through 9
  - Numbering does not reflect physical location of the stack members
- Switch numbers are “sticky”, i.e. they switch will keep the same switch number after reboot
- The user has the ability to renumber the switch through the CLI
- The switch number can be shown by using the “STACK” LED




```
Switch(config)# set switch number 4
Switch(config)# exit
Switch # write mem
```

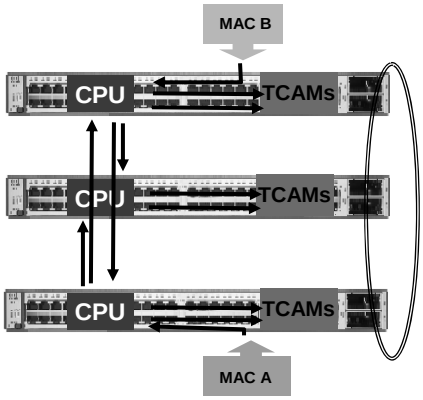

Centralized and Distributed Functions

- Centralized functions
  - Those that are reside on the master node
  - Those that are forwarded to the master node
  - Those that are controlled or synchronized by the master node
- Distributed functions
  - Those that are performed locally by each node
  - These functions are synchronized or updated between the nodes



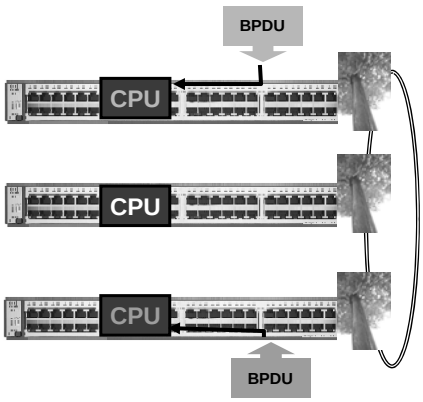
Distributed: MAC Address Management

- MAC address tables are synchronized across the stack
- How it is distributed:
  - A switch learns an address and sends a message to other switches in the stack
  - Learning an address that was previously learned on a different port (either same or different switch) is considered as move



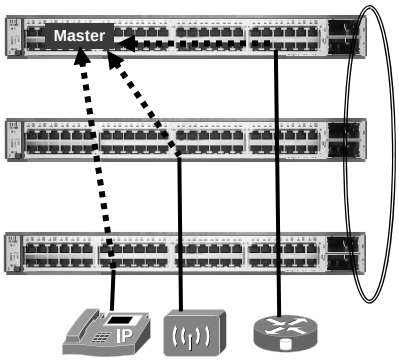
Distributed: STP

- Each switch in the stack runs its own spanning tree instance per VLAN
- Each switches will use the same bridge-id
- Each switch process its own BPDUs
- Show commands show spanning tree as a single entity
- Stacking ports are never blocked
- All packets on the ring have the internal ring header. Therefore, even broadcast packets are source stripped and do not continuously recirculate.
- Supports Cisco enhancements, like Uplink-fast, Backbone-fast, Port-fast, Root-guard, BPDU-guard, etc. are supported with no impact.
- There is support for 128 instances of STP per node/stack



Centralized: CDP

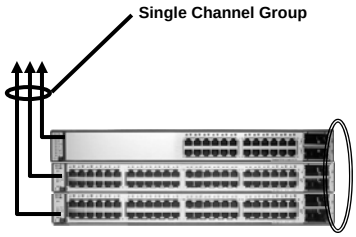
- CDP is implemented using centralized model
- The master will maintain CDP neighbor table and the neighbor tables will be empty on member nodes
- Upon a master switchover, a new master will build the CDP neighbor table



Centralized: Cross Stack Etherchannel/LACP

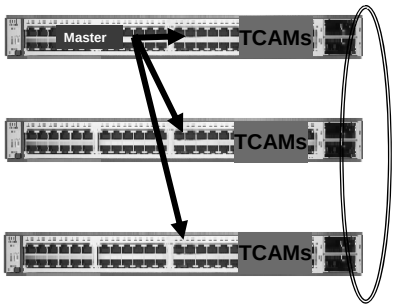
- An LACP-based Etherchannel can be formed with member ports from one or more switches in the stack
- Etherchannel control, not forwarding, is performed by the master node
- Benefits:

  In addition to port aggregation, load-balance and link redundancy and switch-level redundancy is provided



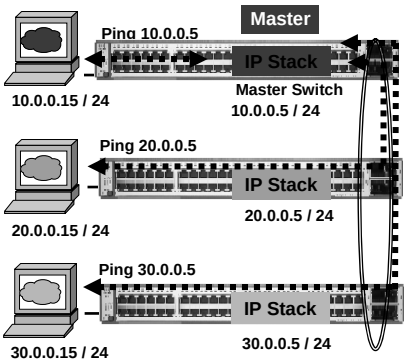
Centralized: VLAN Database

- All switches in the stack build from same VLAN database
- Members download VLAN database from master during initialization
- They are synchronized over the stack ports
- The stack supports all 3 VLAN Trunking Protocol (VTP) modes: server, client and transparent modes
- 1024 VLANs; 4K VLAN IDs are supported



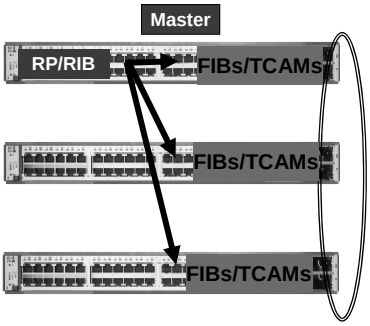
Centralized: Cross Stack IP Host

- The IP stack is active only on stack master
- All IP applications like ICMP, TFTP, FTP, HTTP, SNMP, etc. are handled on the stack master irrespective of, which switch the L3 interface is connected to



Centralized: L3 Routing Overview

- The route processor and Routing Information Base (RIB) lives on the Master
- All Switches have an identical copy of the Forwarding Information Base (FIB) a.k.a. Forwarding table
- Routing protocols include Static, RIPv1, RIPv2, OSPF, IGRP, EIGRP, BGP, PIM-SM/DM, DVMRP, HSRP
- The Cisco Catalyst 3750 uses cross stack equal cost routing
- The Cisco Catalyst 3750 Stack appears as a single router to the world
- No HSRP peering (and no need to do so) among stack members, stack and external router are peers
- Policy Based Routing (PBR), IPv4 and IPv6 routing are supported in hardware
- Distributed Stateful Switchover (SSO) Enabled
- Non Stop Forwarding (NSF) Aware and NSF Capable

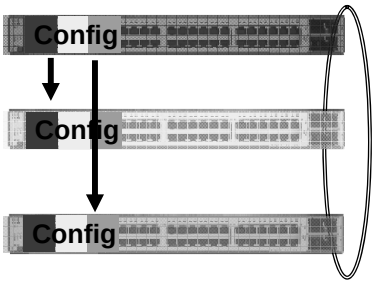



```
Switch(config)# router nsf
Switch(config)# stack-mac persistent timer 0
```


Agenda

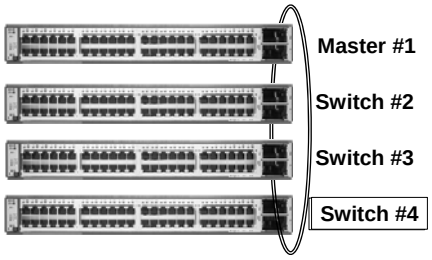
Configuration Management

- Master:
  - Copies of the startup and running config files are kept on all members in the stack
  - The current running-config is synched from the master to all members
  - On a switchover, the new master re-applies the running-config so that all switches are in sync
- Member:
  - Keeps a copy of startup and running config at all times
  - On boot-up waits for config file from master and parses it



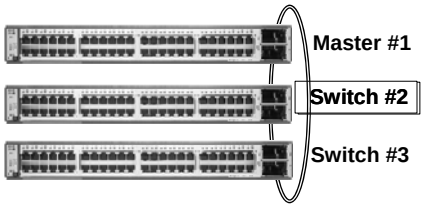
Switch Addition

- The stack has three members—with numbers 1, 2, 3
- A new switch with an existing #3 is added to the stack
- The new switch detects a conflict, and loses, based on the rules used for stack Master determination.
- It is assigned the #4 and reloads switch #4
- All configuration commands in the config file which apply to interfaces 4/0/* apply to the new switch



Switch Removal

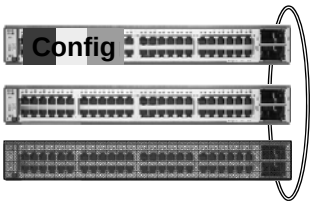
- The stack has three members—1, 2, 3
- Switch #3 is removed or powered down
  - Neighbor loss is detected by Switch #1 and Switch #2
  - Layer 2 and Layer 3 convergence may need to happen
  - Now there is a stack of two switches—Switch #1 and Switch #2
  - Switch#1 is still the master
- Switch #1 is removed or powered down
  - Switch #2 takes over as master
  - Layer 2 and Layer 3 convergence may need to happen
  - Now there is a stack of one switch—#2 which is the master



Replacing a Switch

Replacing a Failed Switch:

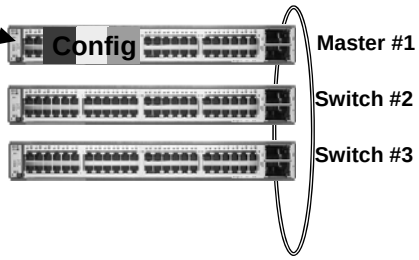
- For example, the failed switch is a Cisco Catalyst WS-C3750E-48TD
- If replaced by another Cisco Catalyst WS-C3750E-48TD , the new switch will receive the port-level configuration of the original unit
- If replaced by a different switch, the original configuration is lost and the new switch receives all stack global configuration



Switch Preprovisioning

Create a provision Switch #4 (Shadow).

Enter the port configuration of the New Switch.



Set the Switch Number (#4)

Switch #4

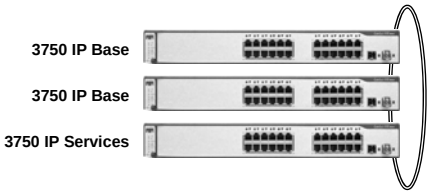
Preprovisioning a Switch

```
Switch(config)# switch 4 provision WS-C3750G-12S

Switch(config)# exit

Switch# write mem

Switch# show running-config | include switch 4
!
interface GigabitEthernet4/0/1
!
interface GigabitEthernet4/0/2
!
interface GigabitEthernet4/0/3
<output truncated>
```


Stack Mismatches

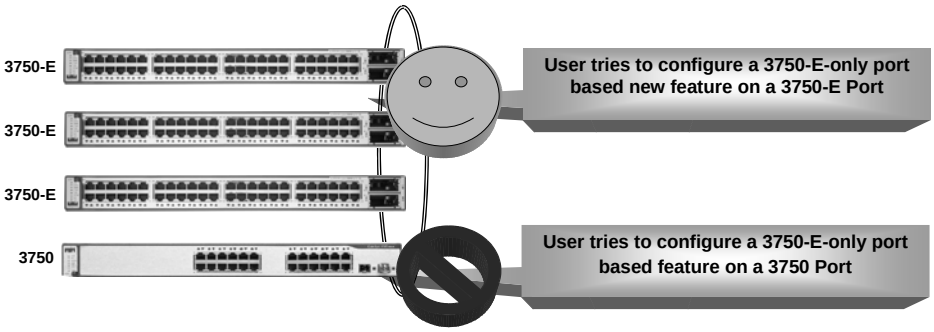
- Hardware Mixed Stack:
  - Feature Mismatch
    - Hardware features (POE Policing, Jumbo frame routing)

- Software Mixed Stack:
  - Software Mismatch
    - IPbase, IPservices



Mixed Hardware Stack:
Incompatible Port Level and Interdependent Features

- New 3750-E port level features are only allowed to be configured on the 3750-E



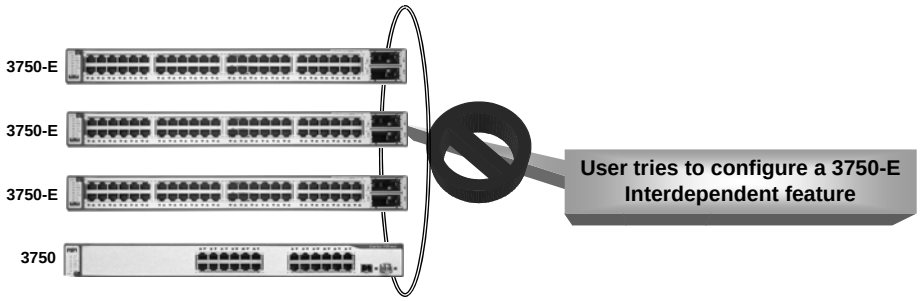
E-Series Port Level Features

Feature	Description
MAC ACLs on IP packets, configued on a port	Both MAC ACLS and IP ACLs on the IP packets
Port + IP ACL on a port	Apply port and VLAN Based ACL at the same time for the same packet
MAC+ IP ACL on a port	ACLs based on a combination of MAC+ IP fields in the same ACE
10G policing	Policing is supported up to 10G
ACL Timestamp support*	Per-entry timestamp/Dynamic ACLs
Per port per vlan/per vlan per port classification	Classification on Port+vlan and policing is per port
IPV6 keyword support*	IPV6 address prefix from /0 to /128 are supported. Matching on presence of routing header and flowlabel
Flow label Support for IPV6	QoS classification based on flowlabel
Egress Shaping	Shaping can go over 50% with improved granularity
Bandwidth limit	Bandwidth limit is in increments of 1%
MAC based QoS classification and policing for IP packets	Police Ip frames based on MAC ACLs in a policy map
Statistics support for Unicast Routing	Support either byte or frame counters
L2 Forwarding of Multicast Frame	Support programmable .1q other than 800
Unknown Unicast Storm Control	Unknown unicast traffic can be blocked at the ingress

* On a VLAN it is Interdependent, otherwise it is port level

Mixed Hardware Stack:
Incompatible Interdependent Feature Configuration

- New 3750-E Interdependent, or system-based, features can not be configured on any switch in a mixed stack.

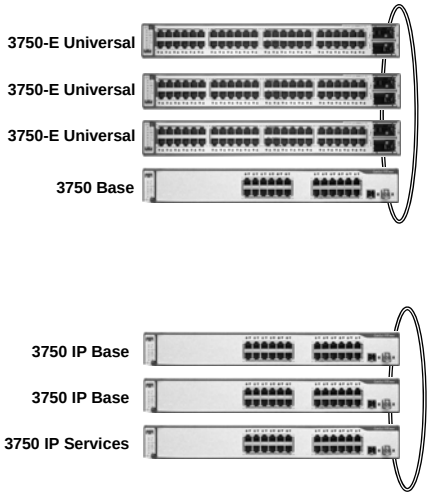


E-Series Switch Interdependent Level Features

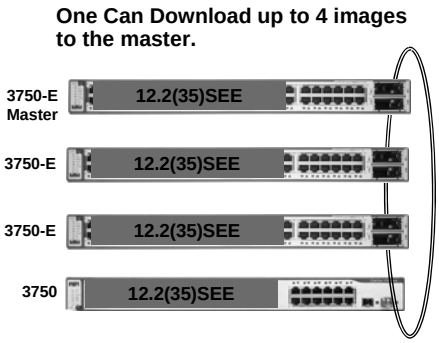
Feature	Description
MAC ACLs on IP packets, configured on Vlan	Both MAC ACLs and IP ACLs on the IP packets
MAC+IP ACL on a VLAN	ACLs based on a combination of MaC+IP fields in the same ACE
ACL Statistics Support	Statistics based on either byte or frame countess
Address learning for ACL for denied/redirected frames	L2 MAC addresses out of frames that are denied or redirected
Uncompressed IPv6	Allows for better utilization of the TCAM space
Unicast RPF	Discards IP addresses that do not have a verifiable IP source
QinQ Inner Tag	Look into inner tag while parsing
Jumbo frame routing	Routing of 9K+ frames

Software Mismatch

- Individually upgrade each Switch
- Use the multiple file download option
- Use the TFTP assistance option



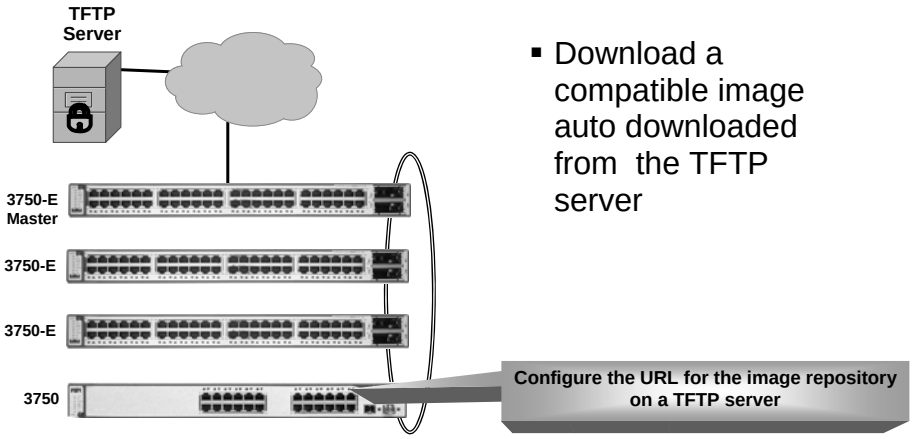
Upgrading with Homemade Image Bundle



- A Catalyst 3750 image is auto-extracted from the initial 3750-E bundle, to the new Catalyst 3750 switch
- The new switch the reloads and joins the stack seamlessly

```
Switch(config)# archive download-sw /allow-feature-upgrade [/directory]
/overwrite /reload <file 1.tar> <file 2.tar> <file 3.tar> <file 4.tar>
```


Upgrading A Mixed Hardware And/Or Software Stack with TFTP Assistance



- Download a compatible image auto downloaded from the TFTP server

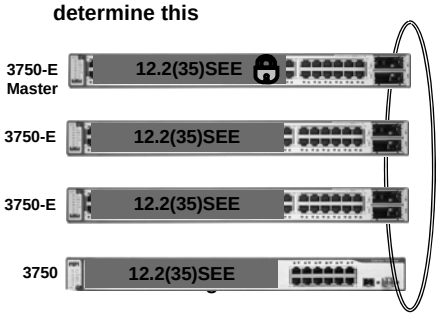
```
Switch(config)# boot auto-download-sw <URL>
```


Creating a Stack by Adding 3750 to a Stack of Catalyst 3750-E Switches

The Catalyst 3750-Es are shipped with a bundled image that contains E and non-E series images

This only works of all switches are running the same version of the stack ring protocol

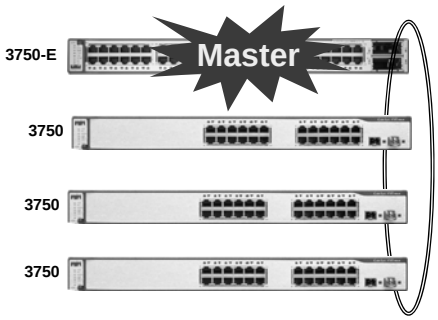
Use the "show switch" command to determine this



- A Catalyst 3750 image is auto-extracted from the initial 3750-E bundle, to the new Catalyst 3750 switch
- The new switch the reloads and joins the stack seamlessly

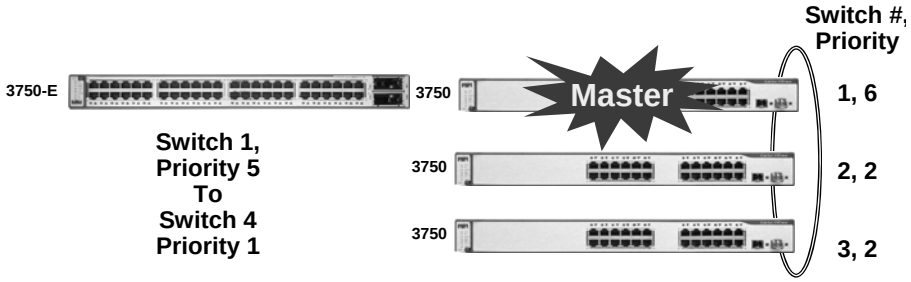
Mixed Stack Upgrading: Make the 3750-E the Master

- Making the a 3750-E the master switch gives the user more options for upgrading in a mixed stack scenario
- In a mixed stack the 3750-E will run in StackWise mode, not Stackwise Plus



Making the 3750-E The Master Super Safe Way Step 1 of 5

- Set the priority of the 3750-E to be lower than that of the switches in the 3750 stack.



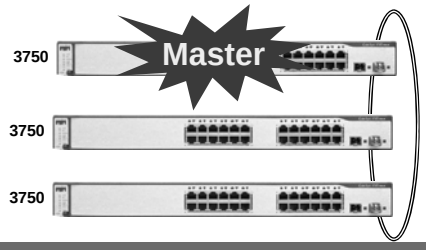

```
Switch_3750_E(config)# switch 1 renumber 4

Switch_3750_E(config)# exit

Switch_3750_E# write mem

Switch_3750_E# reload

Switch_3750_E# switch 4 priority 1
```

```
Switch# show switch

Switch# Role Mac Address Priority State
-----------------------------------------
1    Member 000a.fdfd.0100    6   Ready
2    Member 000a.fdab.0100    2   Ready
3    Master 000a.fd22.0100    2   Ready
```

Making the 3750-E The Master Super Safe Way Step 2 of 5

- Upgrade the 3750 stack to be equal to the code level of the 3750-E.

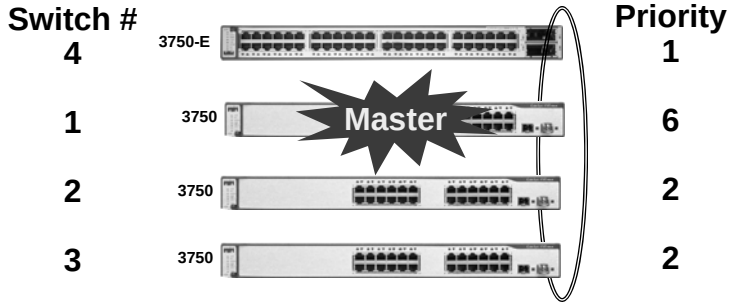
3750-E code level = 3750 code level




```
Switch_3750# copy tftp flash:10.1.1.1 <src_file> <dst_file>
Switch_3750# boot system flash: <new_image>
Switch_3750# reload
```

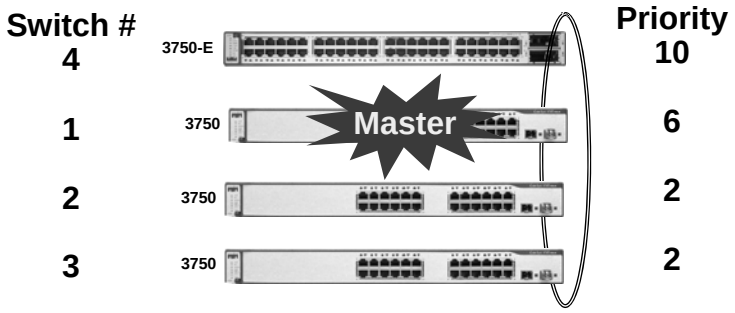
Making the 3750-E The Master Super Safe Way Step 3 of 5

- Power Down the 3750-E
- Physically add the 3750-E to the stack
- Power up the 3750-E (it ill now receive the switch config from the master switch)



Making the 3750-E The Master Super Safe Way Step 4 of 5

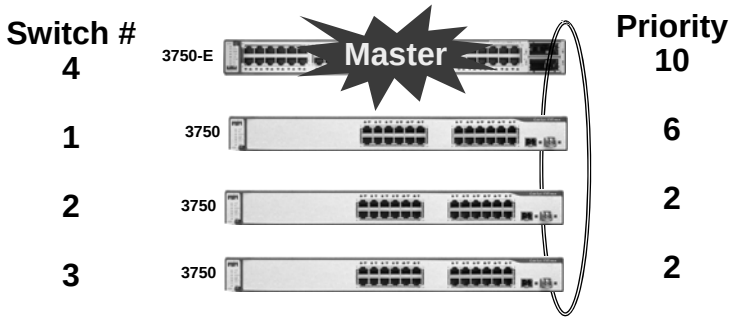
- Set the switch priorities so that when one reboots the stack master, the 3750-E will be come the master.




```
Switch_3750(config)# switch 4 priority 10
```


Making the 3750-E The Master Super Safe Way Step 5 of 5

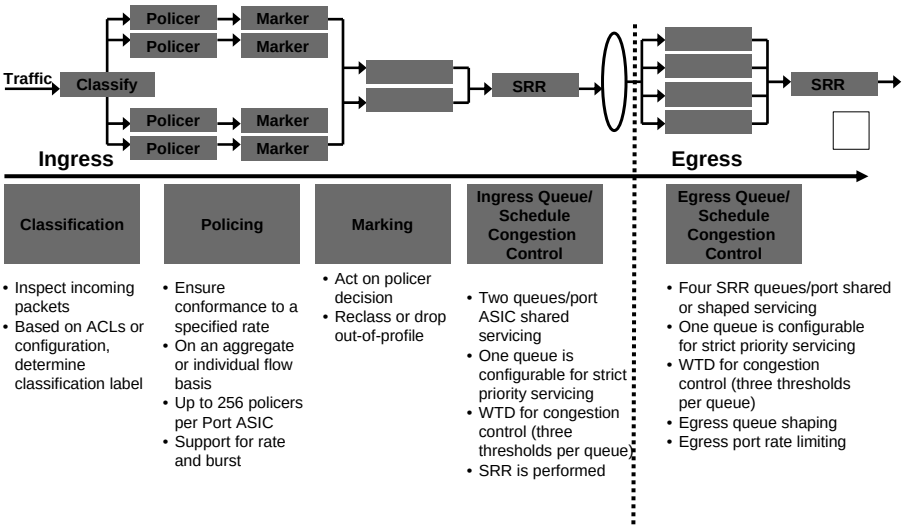
- Reload or powercycle the current stack master.




```
Switch_3750# reload slot 1

/* Note slot 1 is the means Switch ID 1 */
```


Agenda

Cisco Catalyst 3750 and 3750-E QoS Model



- Inspect incoming packets
- Based on ACLs or configuration, determine classification label

- Ensure conformance to a specified rate
- On an aggregate or individual flow basis
- Up to 256 policers per Port ASIC
- Support for rate and burst

- Act on policer decision
- Reclass or drop out-of-profile

- Two queues/port ASIC shared servicing
- One queue is configurable for strict priority servicing
- WTD for congestion control (three thresholds per queue)
- SRR is performed

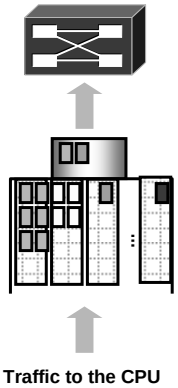
- Four SRR queues/port shared or shaped servicing
- One queue is configurable for strict priority servicing
- WTD for congestion control (three thresholds per queue)
- Egress queue shaping
- Egress port rate limiting

Catalyst 3750 Control Plane Protection 16 Processor Hardware Queues

- DoS protection via 16 CPU queues.
- The workload is distributed to processors on each switch of the stack.
- The stack ring reserves bandwidth for priority traffic

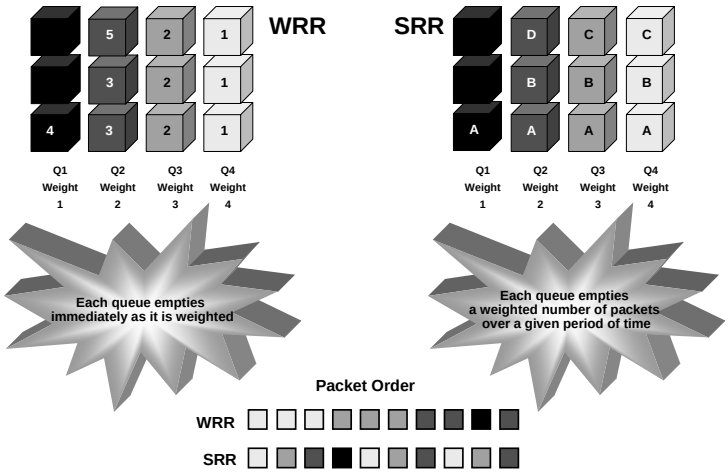
  Bandwidth reservations on the ring ensure the CPU communication is not affected by data traffic.
- These 16 processor queues are not configurable.

  STP, OSPF & inter-CPU packets on separate Queues

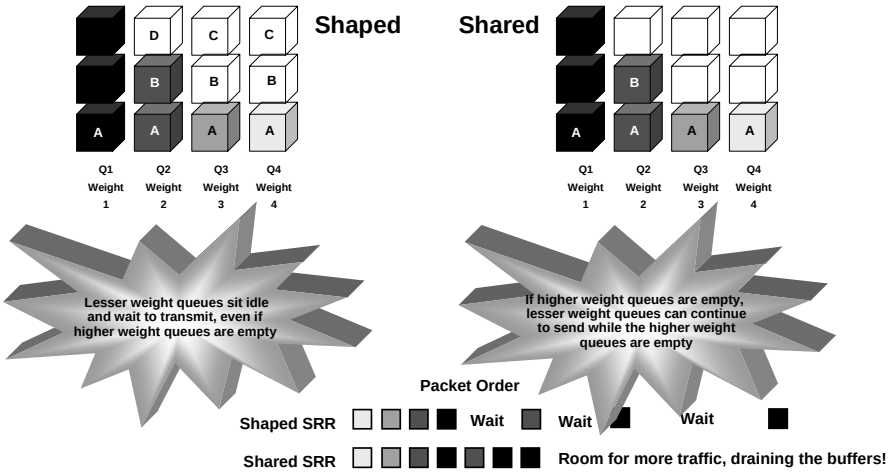


WRR vs. SRR

SRR Is an Evolution of WRR that Protects Against Overwhelming Buffers with Huge Bursts of Traffic by Using a Smoother Round-robin Mechanism



Shaped SRR vs. Shared SRR



Shared Queuing Drains Queues More Efficiently!

Shaped SRR vs. Shared SRR And Traffic Shaping

- Neither Shaped SRR or Shared SRR is better
- Shared SRR is used when one wants to get the maximum efficiency out of a queuing system, because unused time slots can be reused by queues with excess traffic. This is not possible in a standard WRR.
- Shaped SRR is used when one wants to shape a queue or set a hard limit on how much bandwidth a queue can use
- When one uses Shaped SRR one can shape queues within a ports overall shaped rate, and map traffic types to individual queues for shaping

Configuring a Priority Queue (Ingress)

- This example shows how to assign the ingress bandwidths to the queues, one of which is set to a priority queue
- Queue 2, the priority queue, is set with a 10% bandwidth guarantee
- Equal bandwidth weights are allocated to queues 1 and 2, 4/(4+4), for the remaining bandwidth.
- That is, SRR equally shares the remaining 90% of the bandwidth between queues 1 and 2 by allocating half, 45%, to each queue
- Weights range from 0 to 40 for the first command and 0 to 100 for the second command

```
Switch(config)# mls qos srr-queue input priority-queue 2 bandwidth 10
Switch(config)# mls qos srr-queue input bandwidth 4 4
```


Configuring Shaped Queues (Egress)

- This example shows how to configure bandwidth shaping on queue 1
- The bandwidth is weighted 1/10 or 10%
- The final 0 0 0 in the last field indicates that the remaining 3 queues do not operate in shaped mode, but in shared mode.
- Weights range from 0 to 63555

```
Switch(config)# interface gigabitethernet2/0/1

Switch(config-if)# srr-queue bandwidth shape 10 0 0 0
```

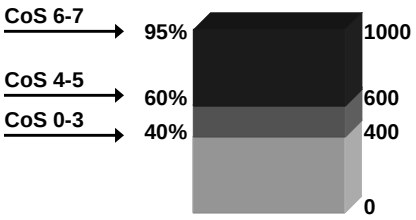

Configuring Shared Queues (Egress)

- This example shows how to configure the weight ratio of the SRR scheduler running on an egress port
- Four queues are used, and the bandwidth ratio allocated for each queue in shared mode is 1/(1+2+3+4), 2/(1+2+3+4), 3/(1+2+3+4), and 4/(1+2+3+4), which is 10 percent, 20 percent, 30 percent, and 40 percent for queues 1, 2, 3, and 4
- This means that queue 4 has four times the bandwidth of queue 1, twice the bandwidth of queue 2, and one-and-a-third times the bandwidth of queue 3
- Weights range from 1 to 255

```
Switch(config-if)# srr-queue bandwidth share 1 2 3 4
```


Cisco Catalyst 3750 Weighted Tail Drop

- WTD is a congestion-avoidance mechanism for managing the queue lengths and providing drop precedences for different traffic classifications
- WTD can be performed at either the Ingress Ring queues or the Egress queues
- User configurable thresholds determine when to drop certain types of packets
- As a queue fills up, lower priority packets are dropped first
- In this example, when the queue is 60% full, arriving packets marked with CoS 0-5 are dropped



One Is Displayed. All Queues Can Be Configured Independently

Configuring Weighted Tail Drop

- This example shows how to map DSCP values 0 to 6 to ingress queue 1
- It also maps DSCP values 10 to 16 to ingress queue 1 and threshold 2
- Last it maps DSCP values 20 to 26 to ingress queue1 and threshold 3
- The final command sets the drop thresholds of levels 1, 2 and 3 to 50%, 70% and 90% respectively

```
Switch(config)# mls qos srr-queue input dscp-map queue 1 threshold 1 0 1 2 3 4 5 6
Switch(config)# mls qos srr-queue input dscp-map queue 1 threshold 2 10 11 12 13 14 15 16
Switch(config)# mls qos srr-queue input dscp-map queue 1 threshold 3 20 21 22 23 24 25 26
Switch(config)# mls qos srr-queue input threshold 1 50 70 90
```


Configuring Ethernet Port Rate Limiting

- This example shows how to limit the bandwidth on a port to 80%
- Percentages can be set in increments of 1%, from 10% to 90%

```
Switch(config)# interface gigabitethernet2/0/1
Switch(config-if)# srr-queue bandwidth limit 80
```


Agenda

Did We Answer?

- The Differences between 3560/3560-E/3750/3750-E switches?
- What is a stack ring?
- How is the stack ring controlled?
- How does the hardware work?
- How are stack processes controlled?
- What happens when I mix different switch types?
- How does QoS work?

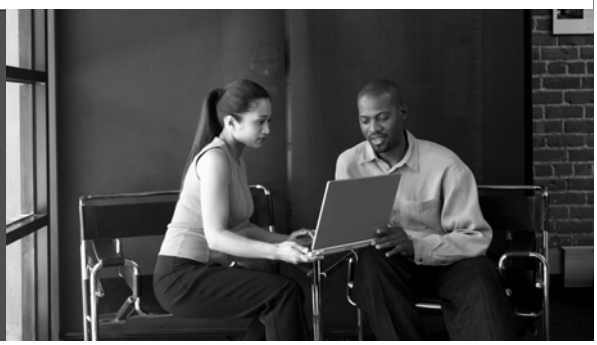
3560-E

3750-E

3560-E

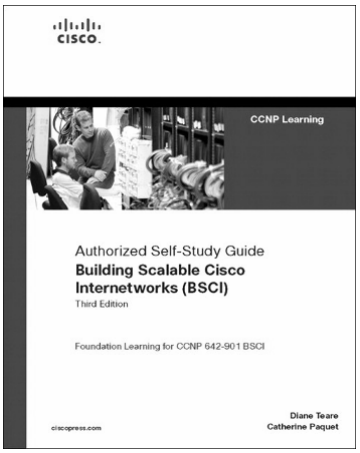
3750-E

Q and A



Recommended Reading

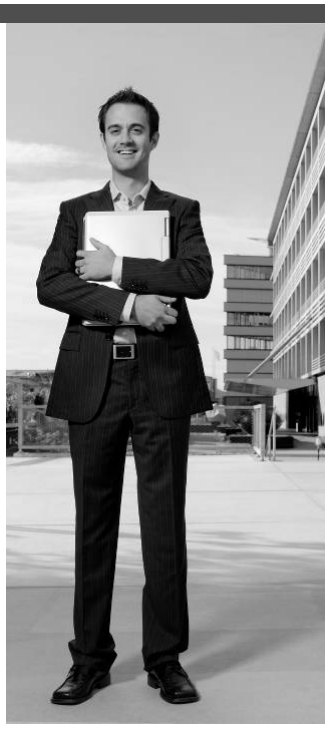
- Continue your Networkers at Cisco Live learning experience with further reading from Cisco Press
- Check the Recommended Reading flyer for suggested books



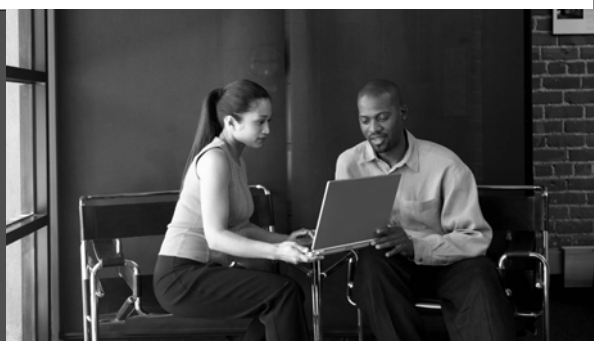
Available Onsite at the Cisco Company Store

Complete Your Online Session Evaluation

- Win fabulous prizes; give us your feedback
- Receive ten Passport Points for each session evaluation you complete
- Go to the Internet stations located throughout the Convention Center to complete your session evaluation
- Winners will be announced daily at the Internet stations



Appendix



Line Rate 10G Uplinks and Converter Module

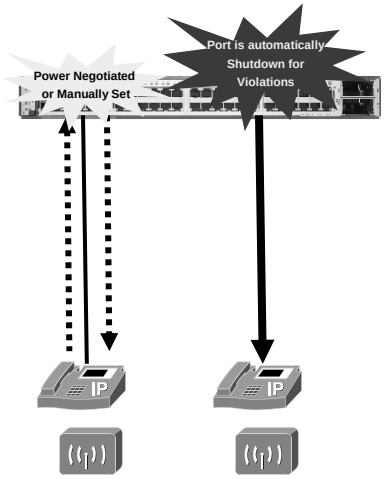
- Dual, line Rate 10GE uplink (X2) modules
- Converts X2 10GE interface into dual SFP interfaces
  - 1000BASE-SX
  - 1000BASE -LX/LH
  - 1000BASE-ZX SFP
  - 1000BASE-T SFP
  - CWDM 1470-1610 NM
  - Other SFPs will be evaluated

StackWise Plus

- StackWise Plus increases the effective stacking throughput to Nx64Gbps using spatial reuse
- E Series switches are backwards compatible, using StackWise, with non E Series switches
- Local switching, without placing packets on a StackWise or StackWise Plus ring

Hardware-Based Power Policing

- Each port negotiates or is manually configured for a power level.
- If a port overdraws (due to a misconfiguration, hardware issue or software bug) the power is turned off on that port.
- This protects the switch and the power being drawn via the other ports.
- Notifies the admin via SNMP



On-Board Failure Logging (OBFL)

- Provides “flight recorder” capability, enabled by default
- Collects operational data about the:
  - Switch
  - Field replaceable power supplies
  - Redundant power systems
  - Pluggable optics modules
- Stores the data as a circular buffer on the flash (2Mbytes)

- The Collected data can be retrieved by TAC and repair personnel to troubleshoot switches including:
  - CLI commands
  - Environment data
  - Message
  - Power over Ethernet (PoE)
  - Temperature Uptime data
  - Voltage
- Each switch on the stack records its own OBFL data
- Collected data can be copied to storage device
- Command : show log onboard

Field-Replaceable Power Supplies

- 1150W AC for full 48 x 15.4 W ports of PoE in a 1 RU switch
- 750W AC, providing 370W of PoE
- 265W AC for non-PoE SKUs
- 265W DC for non-PoE SKUs
- The 1150 W and 750W supplies can be used with the new Redundant Power System

Redundant Power System 2300

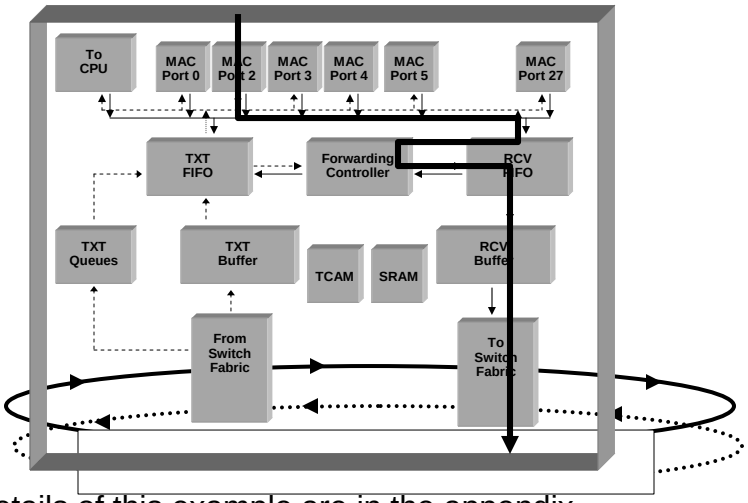
- Makes PS failure transparent to end users
  - Seamless failover to RPS when switch PS fails
  - Automatic back-off when internal power supply of switch resumes
- RPS can have a different AC source than switch(es)
- Programmable failover policy
- Backward compatible: Provides RPS675-compatible support for all Catalyst 3K and 2K switches as well as 2800 and 3800 ISRs
- Up to two switches actively backed up (up to six connected)
- The RPS can be managed via E-Series Switches

Other Enhancements

- All models have the ability to route Jumbo Frames up to 9216 byte sizes
- All models will have two management ports
  - RS-232 serial console port
  - 10/100BASE-TX Ethernet port for out-of-band management

- IPv6 Multicast routing

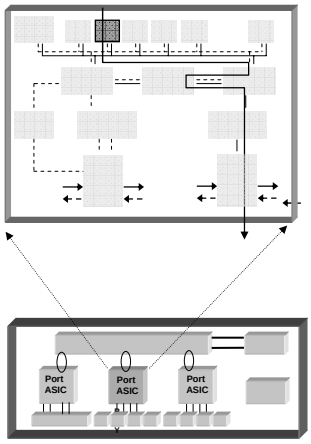
Port ASIC Ingress Flow



Details of this example are in the appendix

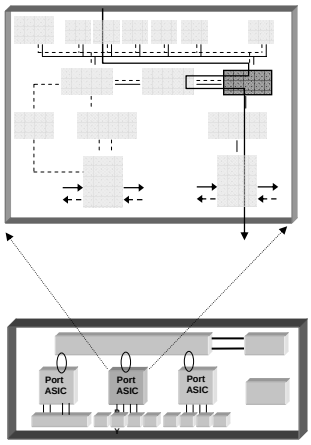
Ingress Flow: MAC Port

- All physical layer functionality is terminated prior to entering this port ASIC function
  - Encoding
  - Power over Ethernet
  - Etc.
- The MAC port's main function is to implement Ethernet Media Access Control
- The MAC port function also adds the 24B internal header, which may be modified later
- This header is used to guide the packet to its destination
- The packet is then passed to the RCV FIFO



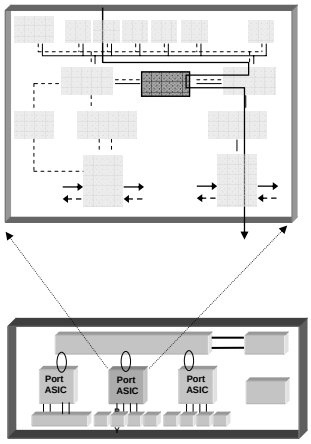
Ingress Flow: RCV FIFO

- The packet enters the RCV FIFO from the MAC port
- There is one physical memory divided into multiple logical RCV FIFOs to serve all of the MACS on the Port ASIC
- One FIFO per port
- The RVC FIFO absorbs time so the forwarding controller to do its job



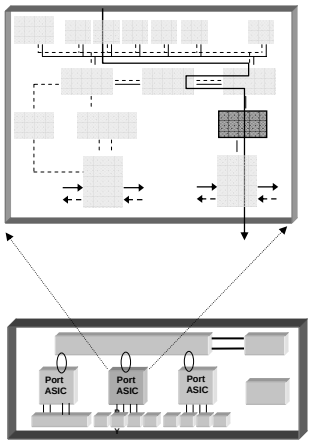
Ingress Flow: Forwarding Controller

- The forwarding controller reads the 24 Byte header and up to 200 Bytes of the packet and performs
  - Forwarding lookups
  - QoS labeling
  - Marking (packet dropping is not performed at this point)
  - ACL lookup
- After the header is updated to the RCV FIFO, the packet is passed to the RCV buffer



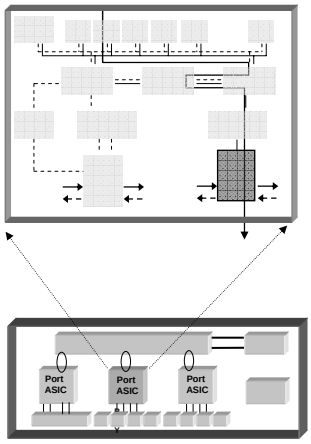
Ingress Flow: RCV Buffer

- The packet enters the RCV buffer while it waits for internal ring access
- This is where the two manageable ingress queues can be configured and packets can be dropped
- SRR is performed on these queues
- WTD can be/is also performed here
- Each buffer:
  - Is shared (common) between all flows
  - Minimum buffer space can be configured to makes sure ports are not buffer starved

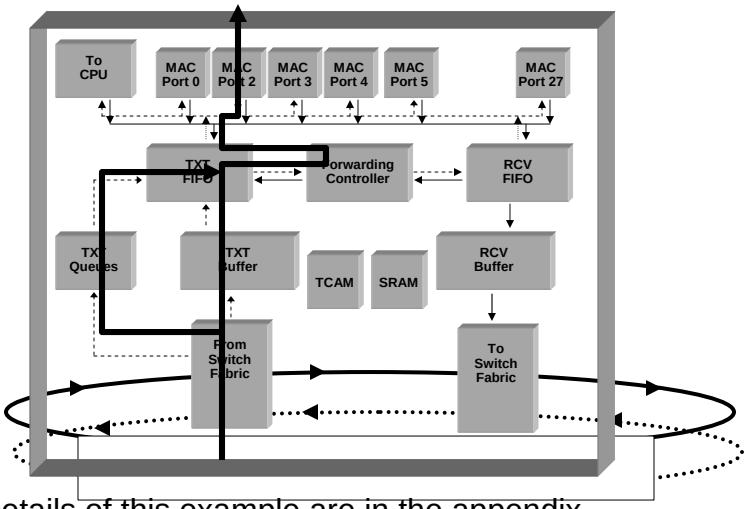


Ingress Flow: Ring Insert

- At this point the port ASIC sends the packet to the Switch Fabric via a point-to-point local ring connection.
- DLAP-PP is used by the Port ASIC
- The packets will be sent or received only on a local ring for the corresponding ports. If a packet arrives on the other ring it is ignored
- The three local ports connected to a Port ASIC are configured to be in DLAP-PP mode
  - Can transmit whenever required. No tokens
  - All packets are received
  - All frames stripped



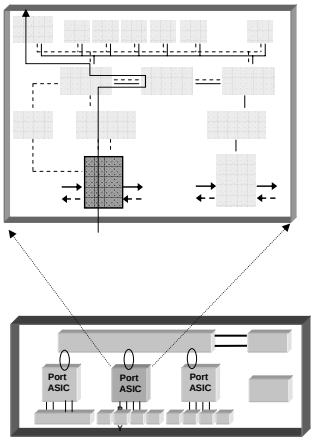
Port ASIC Egress Flow



Details of this example are in the appendix

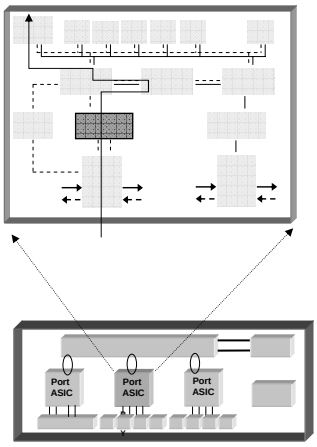
Egress Flow: Ring Copy

- At this point the packet enters the Port ASIC from the point-to-point ring that connects the port ASIC to the Switch Fabric



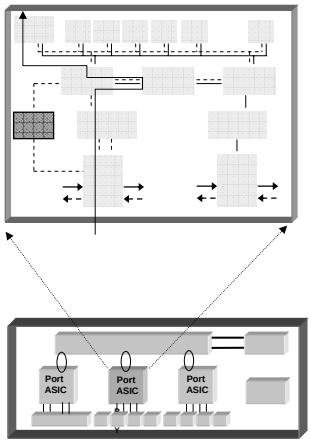
Egress Flow: TXT Buffer

- At this point the TXT queues control what happens to the packets in the TXT buffer
- The TXT buffer performs packet drops



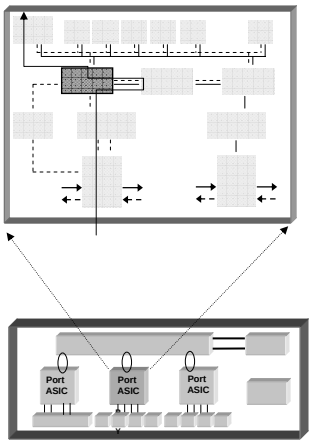
Egress Flow: TXT Queues

- There are four queues per MAC port
- Each queue is highly programmable
- The queues are scheduled with SRR and are susceptible to WTD
- Each buffer:
  - Is shared (common) between all flows
  - Minimum buffer space can be configured to makes sure ports are not buffer starved
- There also are 16 queues for the CPU. Each queue is statically allocated and dedicated to a different protocol



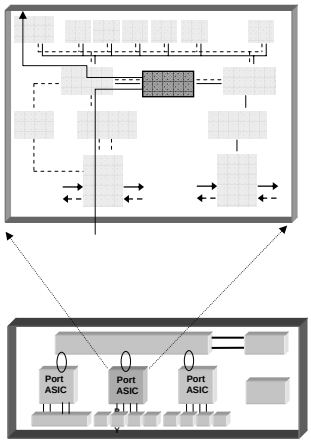
Egress Flow: TXT FIFO

- The packet enters the TXT FIFO from the TXT buffer
- There is one physical memory divided into multiple logical TXT FIFOs to serve all of the MACS on the Port ASIC
- One FIFO per port
- The TXT FIFO absorbs time so the forwarding controller to do its job



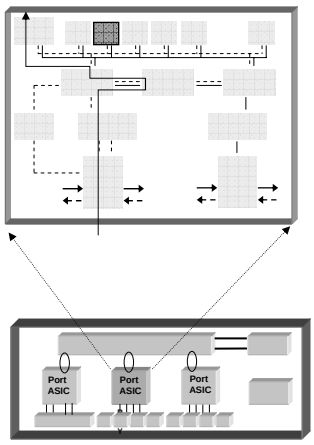
Egress Flow: Forwarding Controller

- The forwarding controller reads the 24B header + the first 200 B of the frame
- The controller performs:
  - Rewrites for the MAC header
  - Time To Live (TTL) decrements
  - Checksum calculation
  - SPAN coordination



Egress Flow: MAC Port

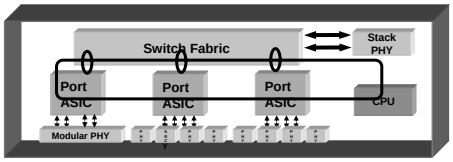
- The packet is received from the TXT FIFO
- The MAC port function performs all Ethernet Media Access Control
- The MAC port function strips the 24B internal header
- All physical layer functionality is performed after leaving the port ASIC function
  - Encoding
  - Power over Ethernet
  - Etc.



CPU Flows

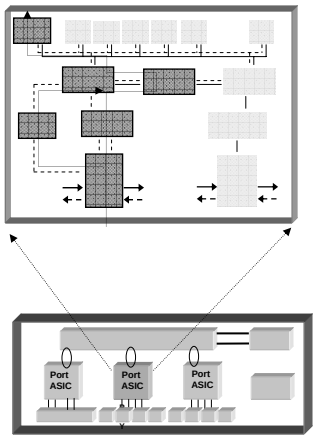
Flows Eligible for CP Forwarding Are:

- Control plane traffic
- Management traffic
- TCAM overflow traffic
  - ACL overflow
  - MAC entry overflow
  - Routing table overflow
- Special protocol flows, these are typically low volume and unofficially supported



CPU Flows: To the CPU

- To hit the CPU the packet must first enter the system
- The packet follows the typical egress path, because the CPU is treated like any other port
  - From Switch Fabric
  - TXT buffer
  - TXT queues
  - TXT FIFO
  - Forwarding controller
  - Off of the Port ASIC to the CPU



CPU Flows: Reentry

- The packet returns to the Port ASIC from the CPU and then follows the typical ingress path
  - RCV FIFO
  - Forwarding controller
  - TXT buffer
  - Switch Fabric
- After this it follows the transmit path to its destination port