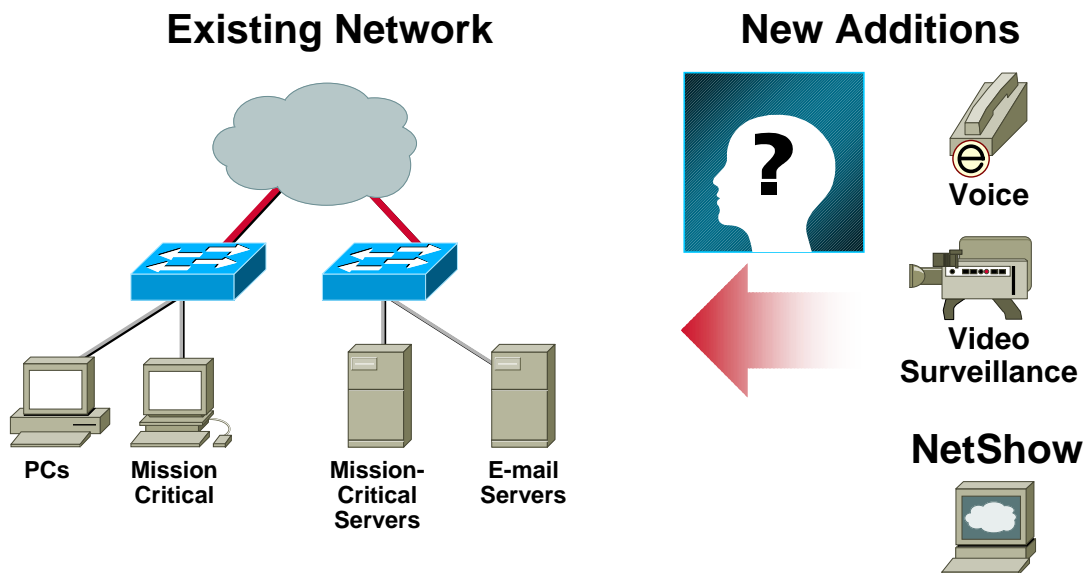


Agenda

- **QoS: What, Why, Where?**
- **QoS Features**
- **Platform Specifics**
- **Configuration**
- **Troubleshooting**

Evolving Campus Network



Protect Mission-Critical Traffic!

Not All Traffic Is Created Equal

	Voice	Video	Data (Best-Effort)	Mission-Critical Data
Bandwidth	Low to Moderate	Moderate to High	Moderate to High	Low to Moderate
Random Drop Sensitivity	Low	Low	High	Moderate to High
Delay Sensitivity	High	High	Low	Moderate to High
Jitter Sensitivity	High	High	Low	Low to Moderate

QoS in Campus Networks?

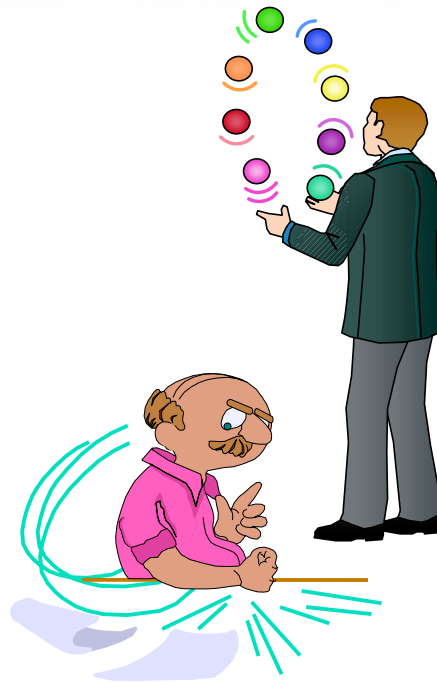


**Why Not Just Increase
the Link Speed?
Throw More Bandwidth
at the Problem!**



Maybe You Can, Maybe You Can't

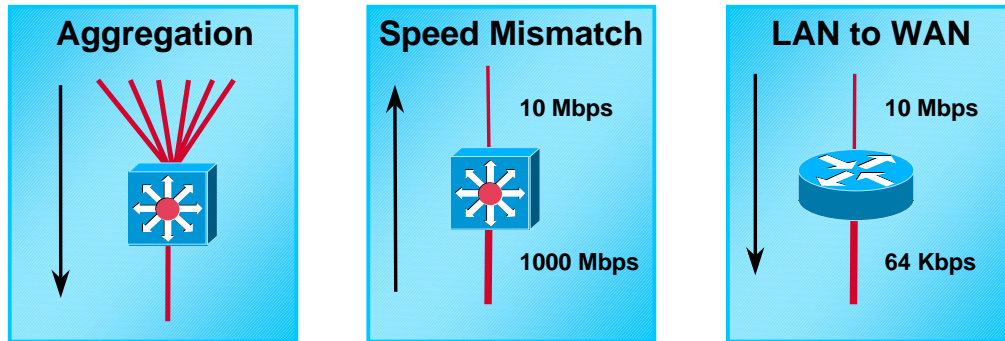
- Output buffers have tendency to be operating at 100% in the campus
- QoS in the campus is primarily about buffer management
- Buffer management is required to control delay or drops
- TCP will eventually retransmit dropped packets
- Do you care about delay or drops?



Buffers

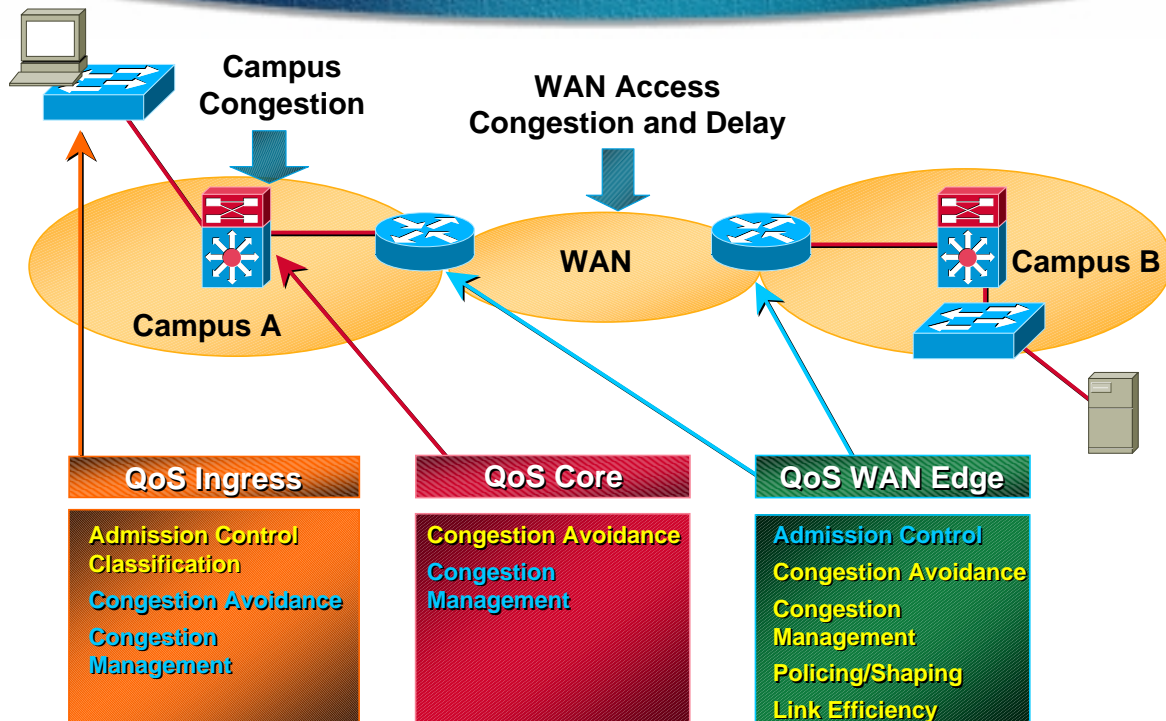
- Buffers can congest in LANs
- QoS required when there is congestion in buffers
- Buffer management can help reduce loss
- Buffering reduces loss but delay sensitive application could be negatively impacted

Where Does Congestion Exist? That's Where QoS Applies!

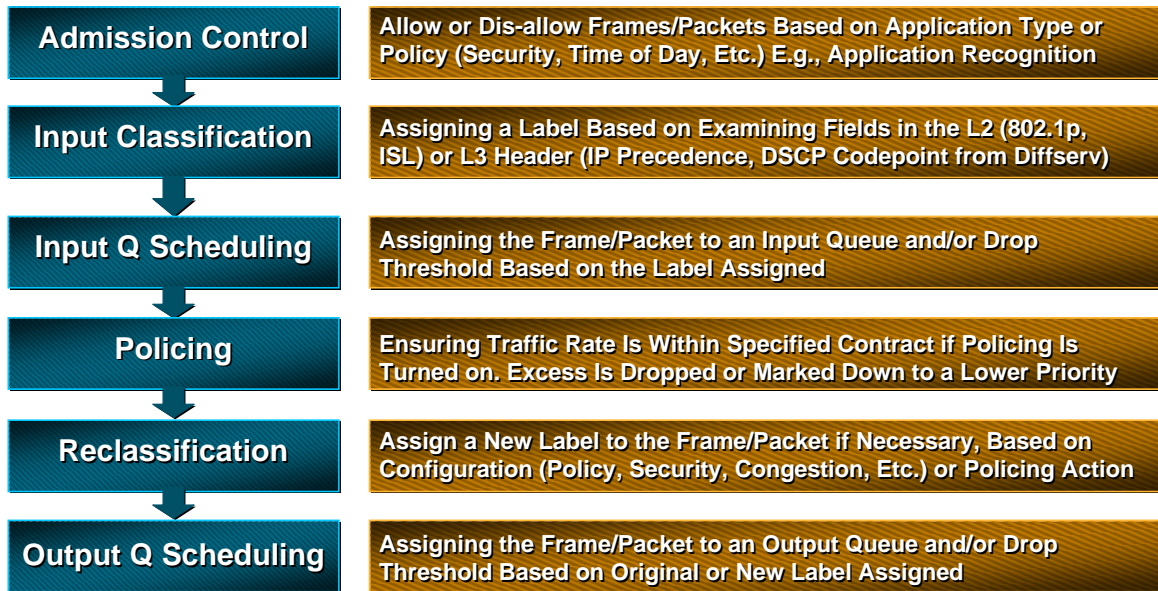


- QoS required where there is congestion in links and buffers
- Points of substantial speed mismatch and points of aggregation
- Transmit buffers have the tendency to fill (TCP windowing)
- Buffering reduce loss, introduce delay

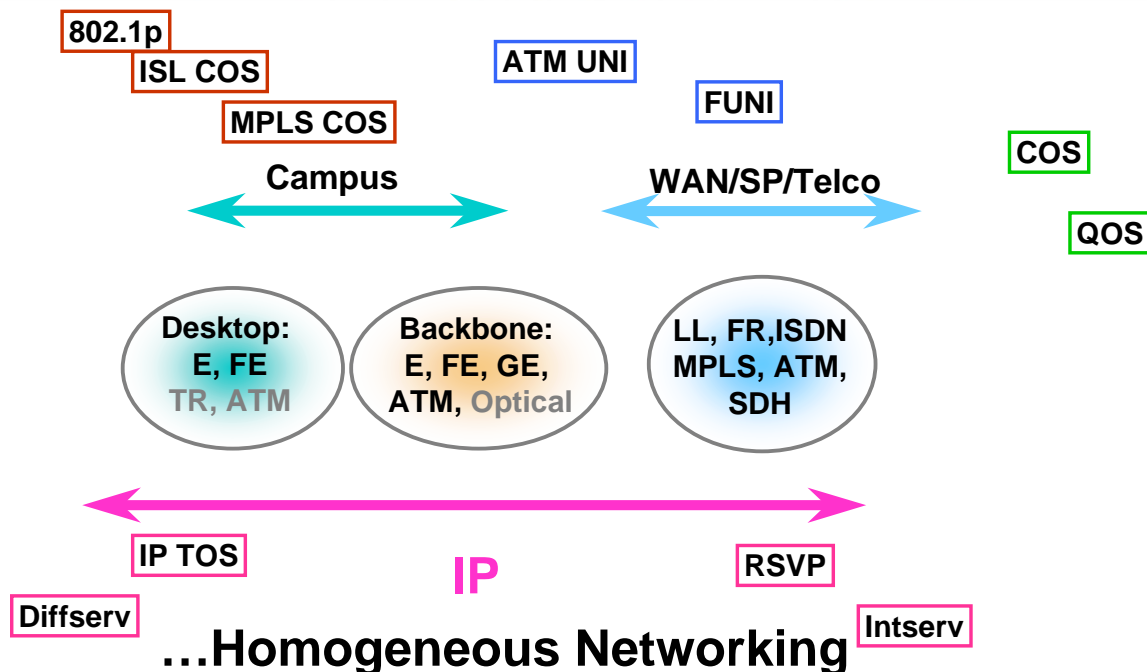
Where Are QoS Features Applied?



QoS Operational Model



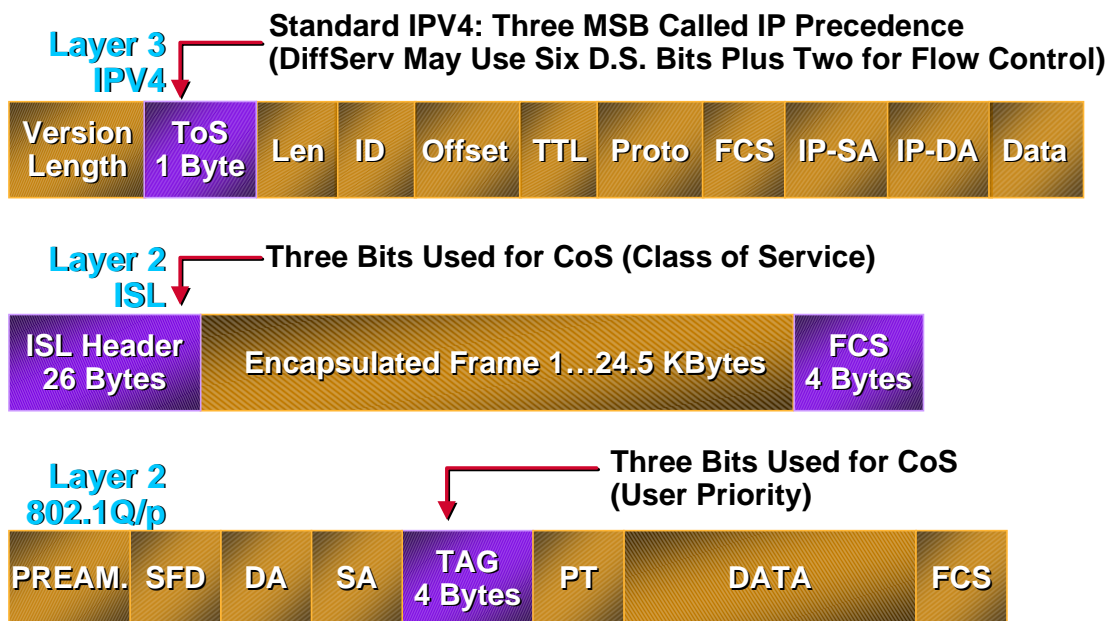
Heterogeneous Technologies...



Agenda

- QoS: What, Why, Where?
- **QoS Features**
- Platform Specifics
- Configuration
- Troubleshooting

Classify at Layer 3 and Layer 2

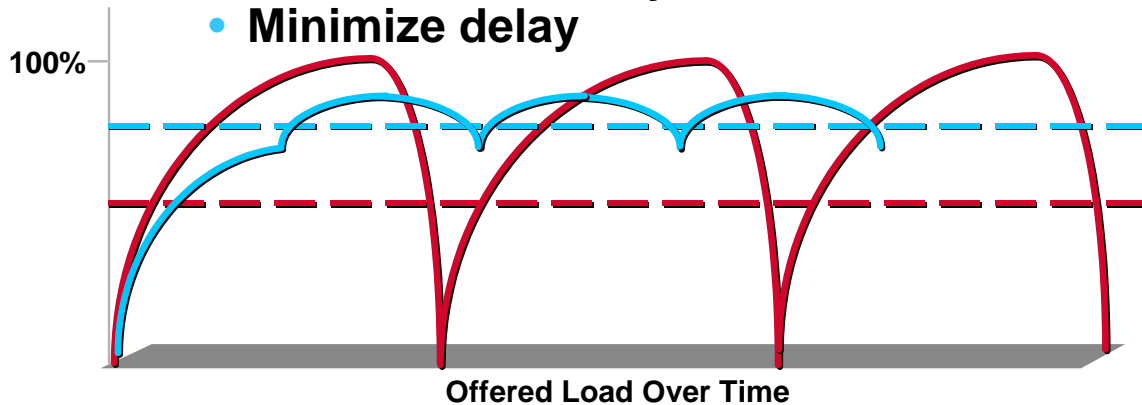


RED for Congestion Avoidance

Uncontrolled Congestion

Managed Congestion

- Maximize goodput
- Fairness to small users
- Reduces Global Synchronization
- Minimize delay



2808
1331_06_2000_c2 © 2000, Cisco Systems, Inc.

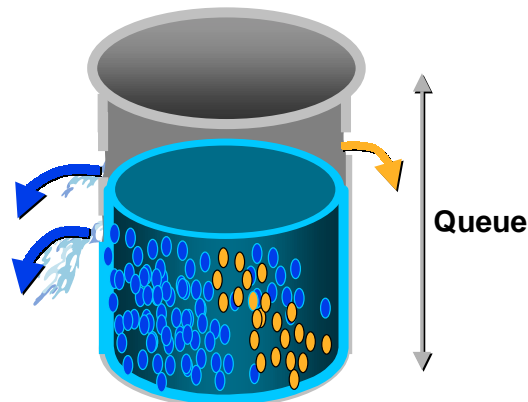
cisco.com

15

Weighted RED

- WRED addresses:

In the event packets need to be dropped, what class of packets should be dropped



Packets Classified as Blue Start Dropping at a 50% Queue Depth. Drop Rate Is Increased as Queue Depth Is Increased

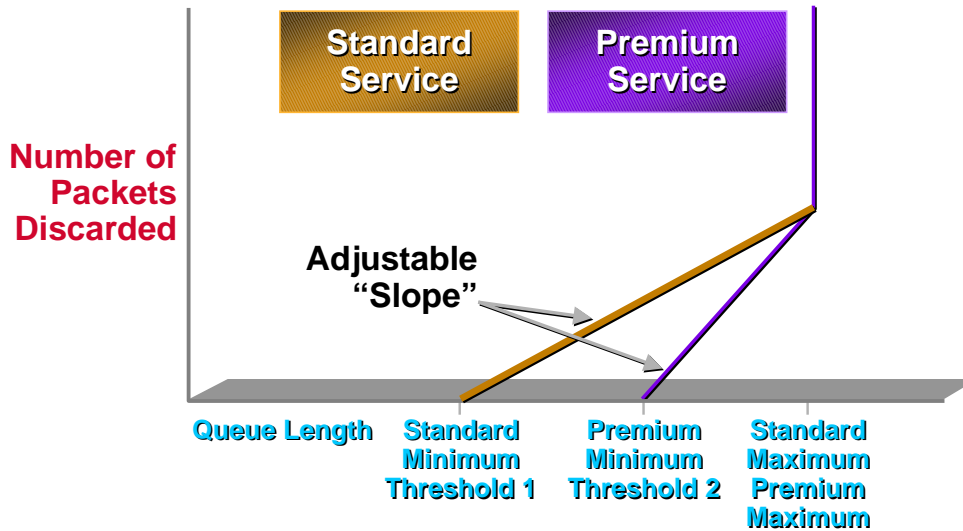
Packets Classified as Gold Are Dropped at 90% Queue Depth

2808
1331_06_2000_c2 © 2000, Cisco Systems, Inc.

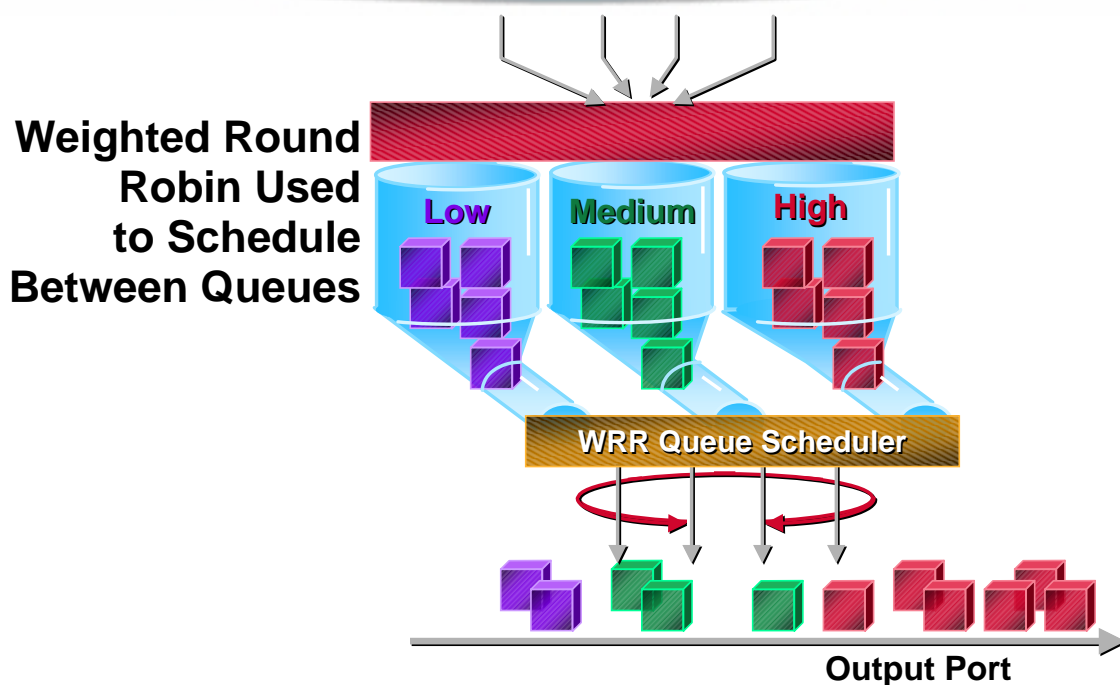
cisco.com

16

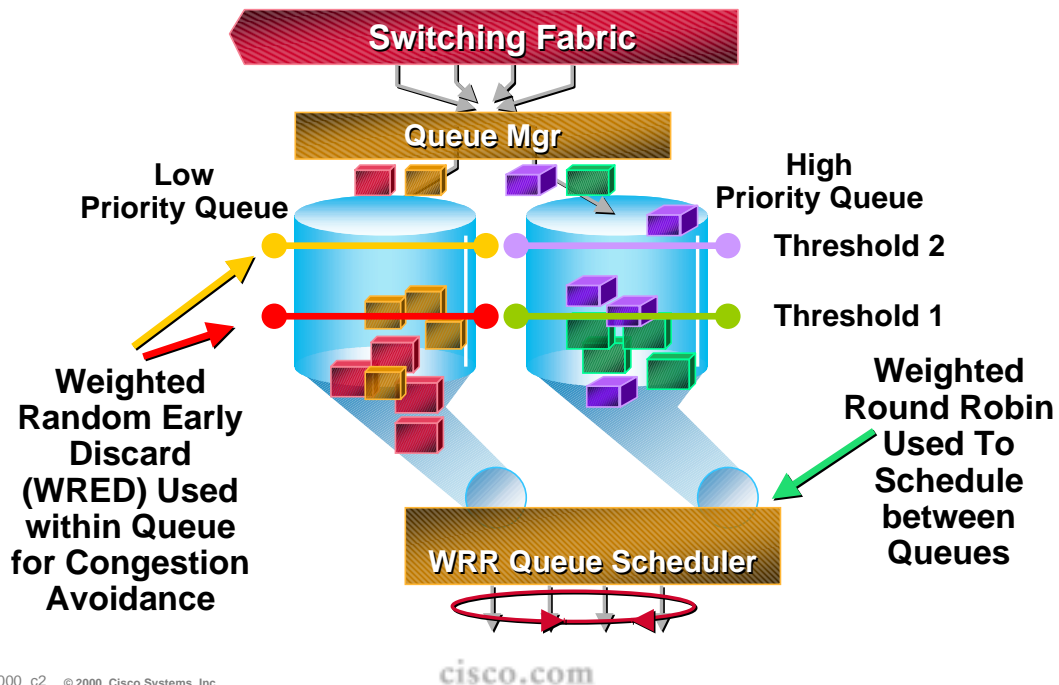
WRED Congestion Avoidance— Two Drop Thresholds



Output Scheduling—WRR



Dual Queues—Dual Thresholds



2808
1331_06_2000_c2 © 2000, Cisco Systems, Inc.

19

CoS Assignment

- On trusted ports, ingress port assigns CoS only to untagged packets
- On untrusted ports, the ingress port assigns a default CoS to all packets, tagged or not
- CoS assigned by an ingress port is independent of the VLAN

2808
1331_06_2000_c2 © 2000, Cisco Systems, Inc.

cisco.com

20

Trusted vs. Untrusted

- Cisco's QoS model assumes that the CoS carried in a packet can be both "trusted" and "untrusted"
- Ports are "untrusted" by default
- Servers may be reliable enough to correctly tag the packets
- End stations generally are not trusted to tag a packet's priority correctly

Agenda

- QoS: What, Why, Where?
- QoS Features
- **Platform Specifics**
- Configuration
- Troubleshooting

QoS in Catalyst Switches

- Deals with the (re-) classification, policing and input/output scheduling
- Hardware determines features
- Classification, policing and reclassification depends on Layer 3 capabilities of the hardware

QoS in Catalyst Switches

- Scheduling capabilities depends on hardware:

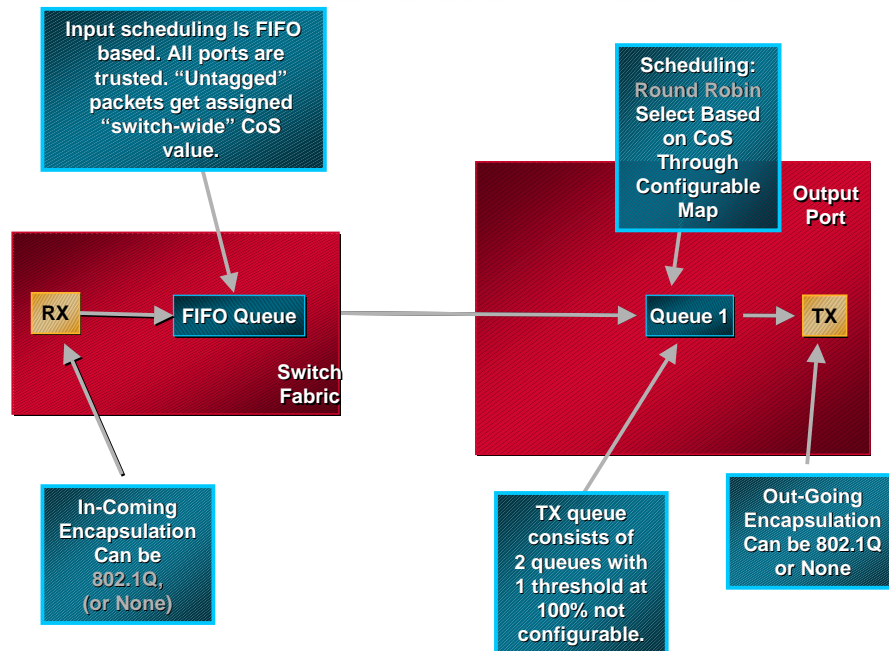
Catalyst 4000: 2Q1T (2 queues, 1 threshold, tx)

Catalyst 5000: 1Q4T (1 queue, 4 thresholds, tx)

Catalyst 6000: 1Q4T (1 queue, 4 thresholds, rx)
1P1Q4T (1 strict priority, 1 queue,
4 thresholds, rx)

2Q2T (2 queues, 2 thresholds, tx)
1P2Q2T (1 strict priority, 2 queues,
2 thresholds, tx)

QoS in Catalyst 4000/2948G



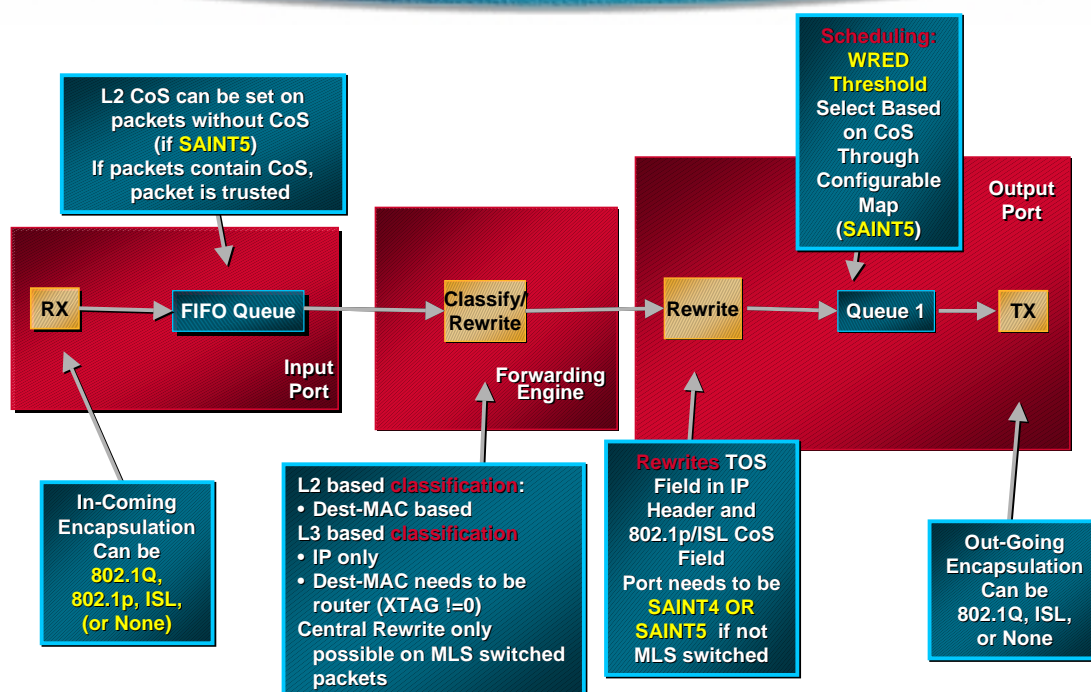
QoS in Catalyst 4000/2948G

- **Input classification based on incoming CoS label (802.1q)—ONLY if frames are tagged with a VLANid**
- **If no CoS, packet gets assigned a CoS value which is "switch-wide"**
- **All ports are considered "trusted"**

QoS in Catalyst 4000/2948G

- The output ports have a 2Q1T capability
- CoS values mapped to two output queues
- Queues are serviced in a round-robin fashion

QoS in Catalyst 5xxx



QoS in Catalyst 5xxx Switches

- **Classification done in three steps:**

Packet is classified based on ingress port. If packet has CoS value, this value is used. Else we use the port's CoS value

If CoS has been explicitly configured for the packets destination mac address, replace the previously assigned CoS

QoS in Catalyst 5xxx Switches

- 3. If packet is IP packet going to/from a router, check if packet matches ACE. If yes, replace CoS with CoS of that ACE**

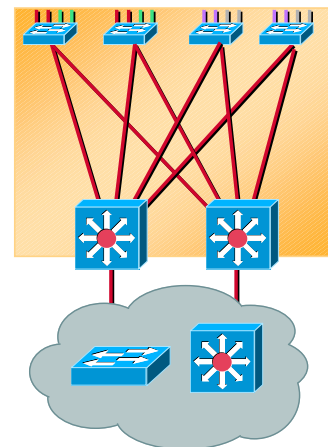
Note: if source and destination are not on the same switch, you have to configure the router's mac-address (qos) and exclude him as MLS-RP

QoS in Catalyst 5xxx Switches

- **Classification can occur based on:**
 - Ingress port CoS (saint5 needed)
 - Destination mac-address (any linecard)
 - IP packets matching a specific ACE
- **An ACE includes:**
 - Source/destination IP address
 - Protocol
 - Source/destination Layer 4 port (UDP/TCP)

QoS in Catalyst 5xxx Switches

- **There needs to be an entry in the Layer 3 table (full flow) in order to rewrite IP precedence**
- **In most cases enabler packet will not pass via the same switch, hence need for software initiated shortcuts (via “set QoS router-mac” command)**
- **These packets cannot be centrally rewritten, inline rewrite needed**



QoS in Catalyst 5xxx Switches

- **Result of the classification is:**
 - Rewrite of CoS (not matching an ACE)**
 - Rewrite of BOTH IP Precedence and CoS (ACE match)**
- **CoS might get lost on egress**
 - Although it is used for scheduling**

QoS in Catalyst 5xxx Switches

- **Output scheduling is only supported on the SAINT5 based linecards**
- **The Catalyst 5xxx has support for QoS output scheduling using 1Q4T ports**
- **WRED is used to randomly drop packets with four wred thresholds**
 - Note: MLS flowmask will change to full-flow when enabling QoS**

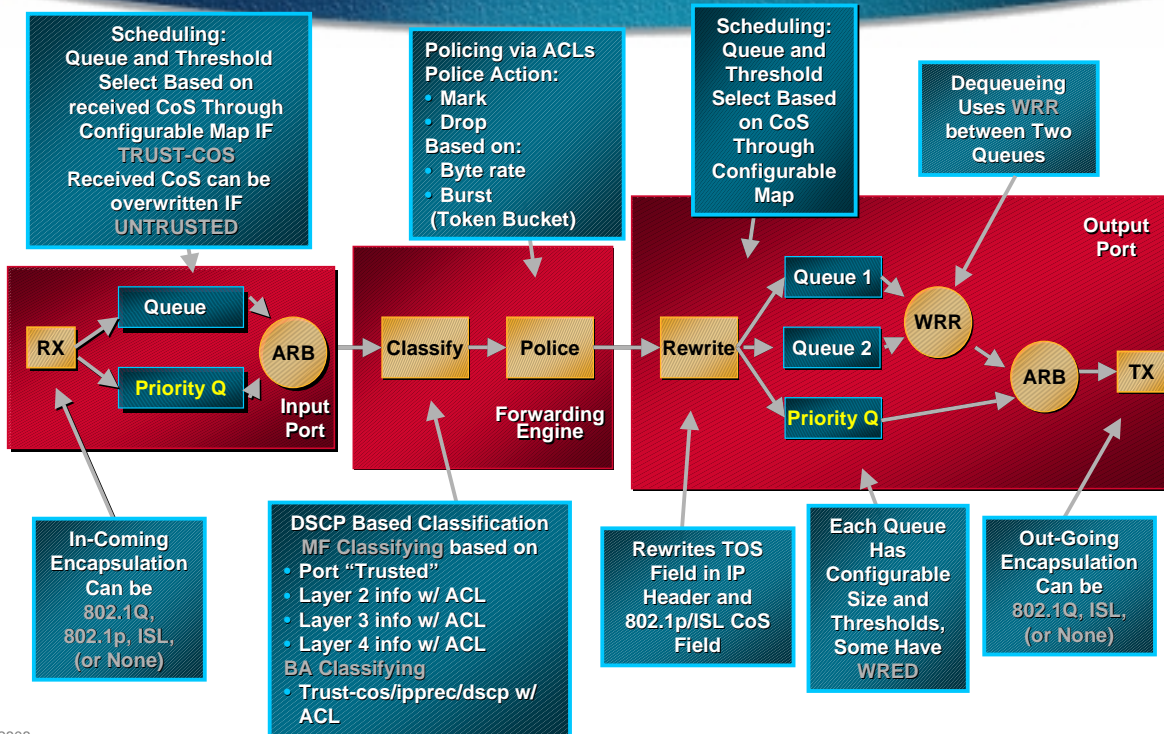
QoS in Catalyst 5xxx Switches

- **Saint5 based:**
 - 24-port 10/100BaseTX Fast Ethernet RJ-45 (WS-X5234-RJ45)
 - 24-port 100BaseFX Fast Ethernet MT-RJ (WS-X5236-FX-MT)
 - 36-port 10/100BaseTX Fast Ethernet Telco (WS-X5239-RJ21)
 - 24 Port 100 BaseFX SMF MT-RJ Switching Module (WS-X5237-FX-MT) (5.4 release)
- **Saint4 based:**
 - 24-port 10/100Base TX Fast EtherChannel RJ-45 (WS-X5225R-RJ45)
 - 12-port 100BaseFX MMF Fast EtherChannel (WS-X5201R-FEFX-MMF)
 - 4-port 10/100BaseTX Fast EtherChannel uplink module (WS-U5537-FETX)
 - 4-port 100BaseFX MMF MT-RJ Fast EtherChannel uplink module (WS-U5538-FEFX-MMF)

SAINT4 or SAINT5?

- **Port based CoS**
 - Requires SAINT5 ingress
- **For dest-mac based CoS**
 - No specific module for ingress/egress
- **For TOS rewrite**
 - Requires SAINT5 egress if not
MLS switched

QoS In Catalyst 6000



2808
1331_06_2000_c2 © 2000, Cisco Systems, Inc.

cisco.com

37

QoS in Catalyst 6xxx Switches

- Classification is handled by ACLs
- Each ACL contains various ACEs
- Three types of ACEs:
 - IP ACE
 - IPX ACE
 - MAC ACE

2808
1331_06_2000_c2 © 2000, Cisco Systems, Inc.

cisco.com

38

QoS in Catalyst 6xxx Switches

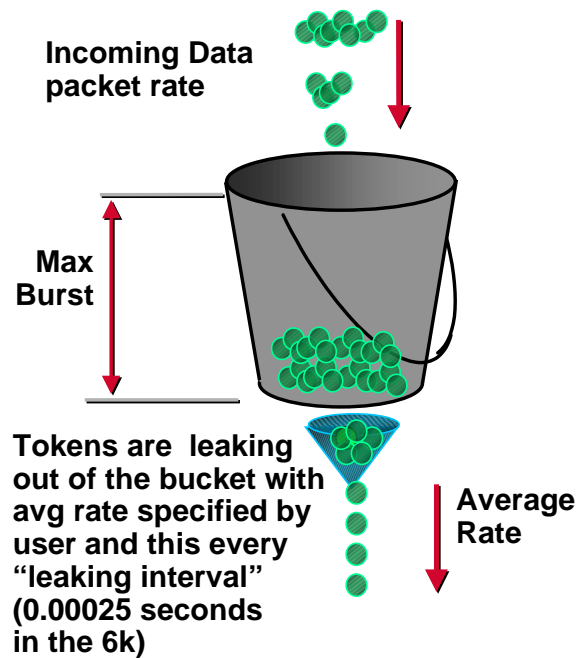
- **Classification can be done inter- and intra-vlan**
- **Classification can be port based or VLAN based**
- **No restriction on the use of MLS and QoS concurrently**
- **MLS flowmask can often stay destination-IP (except when microflow policing)**

QoS in Catalyst 6xxx Switches

- **Policing uses a token bucket scheme**
- **Out-of-profile packets:**
 - Dropped**
 - Mark down the DSCP (via a map)**
- **Two policers are applied to each packet:**
 - Individual microflow policer (each mls entry can have its own policer, up to 63)**
 - Aggregate flow policer (up to 1023)**

Token Bucket for Policing

- Start with a bucket without tokens
- Tokens can be added at a bursty rate
- Tokens are leaked at a specified constant rate
- If the bucket overflows, packet gets policed



Token Bucket for Policing(2)

- Example:

Rate: 16000 kbps, burstsize: X kbits

Means: each leaking interval we will leak $16000 \text{ kbps} * 0.00025 \text{ s} = 4000 \text{ bits}$

If $X=1$ and the input rate is 16 Mbps, the output rate will be 4 Mbps only: this is because between leaking intervals, we attempt to put $16000 \text{ kbps} * 0.00025 \text{ s} = 4000 \text{ bits}$ into the bucket. Since the bucket is already full after 1000 bits, 3/4 of the traffic will get policed

To make sure the bucket can fill at 16 Mbps, the bucketsize needs to be exactly 4000 bits, $X=4$. Right before it will overflow it is drained back to 0

To allow a burst above avg rate, increase burst, $x=5$ allows an extra 1000 bits coming into the bucket before it overflows, meaning a burst rate of up to 20000 kbps (for one second).

QoS in Catalyst 6xxx Switches

- **Output scheduling consists of:**
 - Assigning to queues based on CoS**
 - Dropping based on thresholds**
 - Configuring threshold levels**
 - Assigning weights for WRR**

Agenda

- **QoS: What, Why, Where?**
- **QoS Features**
- **Platform Specifics**
- **Configuration**
- **Troubleshooting**

Catalyst 4000/2948G/2900/3500-XL Example

- Configuration on Catalyst 4000:

```
c4k#set qos enable
c4k#set qos defaultqos 0
```

optional:

```
c4k#set qos map 2q1t 2 1 cos 4-7
c4k#show qos info config
```

QoS setting in NVRAM:

QoS is enabled

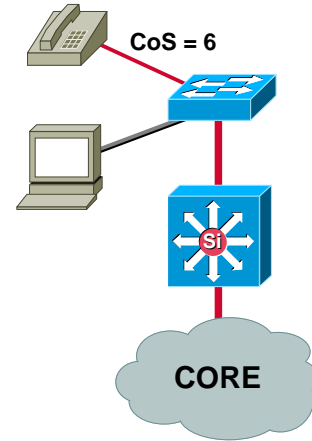
All ports have 2 transmit queues with 1 drop thresholds (2q1t).

Default CoS = 0

Queue and threshold mapping:

Queue threshold CoS

Queue	Threshold	CoS
1	1	0 1 2 3
2	1	4 5 6 7



QoS in Catalyst 5xxx Switches Access Layer Example

- Relevant commands:

```
c5k#set qos enable
c5k#clear mls include 2.2.2.254
c5k#set qos router-mac 00-00-0c-11-22-33
c5k#set qos ip 4 tcp any 0 host 1.1.1.1 2001
```

optional:

```
c5k#set qos map 1q4t 1 4 cos 6,7
c5k#set qos map 1q4t 1 3 cos 4,5
c5k#sh qos info config 1q4t tx
```

QoS setting in NVRAM for 1q4t transmit:

QoS is enabled

Queue and Threshold Mapping:

Queue Threshold CoS

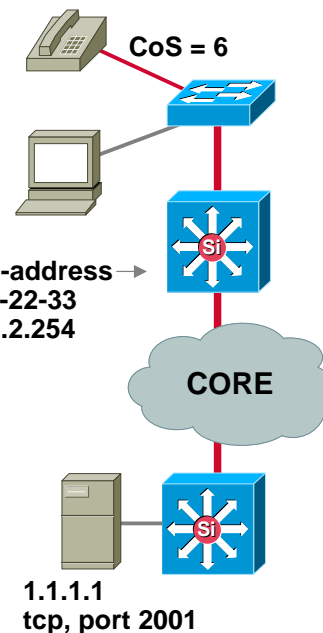
Queue	Threshold	CoS
1	1	0 1
1	2	2 3
1	3	4 5
1	4	6 7

Tx WRED thresholds

Queue # Thresholds in percentage (in abs values)

Queue	10%	20%	40%	100%
1	10%	20%	40%	100%

Router mac-address →
00-00-0c-11-22-33
MLS-RP 2.2.2.254



QoS in Catalyst 6xxx Switches Examples

- **Example1 : input scheduling through 1p1q4t ports. VOIP has CoS=5**

```
c6k#set qos enable
c6k#set qos port 1/1 trust trust-cos
c6k#set qos drop-threshold 1p1q4t rx queue 1 20 40 75 100
c6k#set qos map 1p1q4t rx 2 1 cos 5
c6k#set qos map 1p1q4t rx 1 1 cos 0,1
c6k#set qos map 1p1q4t rx 1 2 cos 2,3
c6k#set qos map 1p1q4t rx 1 3 cos 4
c6k#set qos map 1p1q4t rx 1 4 cos 6,7
```

Note: there is a default IP/IPX/MAC ACL that matches all traffic and has the 'dscp 0' keyword. This ACL is used to mark frames with dscp= 0 on all ports. There is a default mapping between CoS and DCSP and IP precedence and DSCP. On output IP-precedence and/or QoS will get derived from internal DSCP.

```
c6k#set qos acl default-action ip dscp
```

Note: an ACL can always override the per-port classification

QoS in Catalyst 6xxx Switches Examples

- **Example 2: output scheduling using 1p2q2t ports. VOIP has CoS=5, Mission critical has CoS 4**

```
c6k#set qos enable
c6k#set qos map 1p2q2t tx 3 1 cos 5
c6k#set qos map 1p2q2t tx 1 1 cos 0,1
c6k#set qos map 1p2q2t tx 1 2 cos 2,3
c6k#set qos map 1p2q2t tx 2 1 cos 4
c6k#set qos map 1p2q2t tx 2 2 cos 6,7
c6k#set qos wred 1p2q2t tx queue 1 50 100
c6k#set qos wred 1p2q2t tx queue 2 80 100
c6k#set qos wrr 1p2q2t 30 70
c6k#set qos txq-ratio 1p2q2t 70 15 15
```

QoS in Catalyst 6xxx Switches Examples

- **Example 3: Classify frames using ACLs.**
All traffic from vlan 200 to 1.1.1.1 gets CoS 4 (internal default DSCP 32). IP phones are on module 3 and get a DSCP of 40 (Cos=5)

```
c6k#set qos enable
c6k#set port qos 2/1-12 vlan-based
c6k#set qos acl ip mission_critical dscp 32 ip
any host 1.1.1.1
c6k#commit qos acl mission_critical
c6k#set qos acl map mission_critical 200
c6k#set port qos 3/1-48 port-based
c6k#set port qos 3/1-48 trust untrusted
c6k#set qos acl ip voip dscp 40 ip any any
c6k#commit qos acl voip
c6k#set qos acl map voip 3/1-48
```

QoS in Catalyst 6xxx Switches Examples

- **Example 4: policing of all http traffic in general and to web server 2.2.2.2 in particular**

```
c6k#set qos policer aggregate Policer_http rate
10000 burst 3 drop
c6k#set qos policer microflow Policer_2.2.2.2
rate 1000 burst 1 drop
c6k#set qos acl ip police_acl trust-dscp
microflow Policer_2.2.2.2 tcp any host 2.2.2.2
eq port 80
c6k#set qos acl ip police_acl trust-dscp
aggregate Policer_http tcp any any eq port 80
```

Agenda

- QoS: What, Why, Where?
- QoS Features
- Platform Specifics
- Configuration
- **Troubleshooting**

QoS Restrictions/Minimum Requirements—Catalyst 4000

- **Catalyst 4xxx/2948G do not support input scheduling, policing, Layer 3 TOS rewrite, or destination-based QoS**
- **Queue thresholds are hard-set to 100% and are not tunable**
- **Minimum software 5.3**
- **No CoS on “Native” VLAN**

QoS Restrictions/Minimum Requirements—Catalyst 5000

- **QoS on Catalyst 5000 requires one of: Sup III with NFFC2, or Sup II G, or Sup III G**
- **Must run 5.x or higher software**
- **SAINT5 based linecards required to support output scheduling**

QoS Restrictions/Minimum Requirements—Catalyst 6000

- **Catalyst 6xxx requires Policy Feature Card (PFC) to implement Layer 3 QoS features via Ternary CAM (TCAM)**
- **No PFC means no Layer 3 QoS on Catalyst 6xxx; Layer 2 only**

QoS In Catalyst 6xxx Switches Gotchas

- Preservation of CoS/ToS is function of port trust-state
- Software routed packets loose CoS
- 10/100 ports are **always** untrusted. Use ACL trust-cos/ipprec/dscp keyword in ACL to trust flows arriving on these ports. Note that input scheduling works as expected!
- Size of strict priority queue and high priority queue have to be the same
- No flow-control and QoS together

Classification of Multicast

- IGMP type fields in ACL are V1 definitions, not V2 report
- To avoid out-of-order packet delivery, map marked-down DSCP to same queue as in-profile DSCP. (Affects unicast also)

Show Port QoS

```
Cat6509> (enable) show port qos 4/1
```

QoS is disabled for the switch.

Configured settings are not used.

QoS policy source for the switch set to local.

Port	Interface	Type	Interface	Type	Policy	Source	Policy	Source
	config	runtime	config	runtime				

4/1	port-based	vlan-based			COPS		local	
-----	------------	------------	--	--	------	--	-------	--

Port	TxPort	Type	RxPort	Type	Trust	Type	Trust	Type	Def	CoS	Def	CoS
			config	runtime	config	runtime	config	runtime				

4/1	1p2q2t		1p1q4t	untrusted	trust-dscp		0	0				
-----	--------	--	--------	-----------	------------	--	---	---	--	--	--	--

(*)Trust type set to untrusted.

Config:

Port	ACL name	Type
------	----------	------

No ACL is mapped to port 4/1.

Show Port Capabilities

```
Cat6509> (enable) show port capabilities 4/1
```

Model	WS-X6416-GE-MT
Port	4/1
Type	1000BaseSX
Speed	1000
Duplex	full
Trunk encap type	802.1Q,ISL
Trunk mode	on,off,desirable,auto,nonegotiate
Channel	yes
Broadcast suppression	percentage(0-100)
Flow control	receive-(off,on,desired),send-(off,on,desired)
Security	yes
Membership	static,dynamic
Fast start	yes
QOS scheduling	rx-(1p1q4t),tx-(1p2q2t)
CoS rewrite	yes
ToS rewrite	DSCP
UDLD	yes
SPAN	source,destination

Show QoS Information

```
Cat6509> (enable) sh qos info runtime 4/1
Run time setting of QoS:
QoS is enabled
Policy Source of port 4/1: Local
Tx port type of port 4/1 : 1p2q2t
Rx port type of port 4/1 : 1p1q4t
Interface type: port-based
ACL attached:
The qos trust type is set to untrusted.
Default CoS = 0
Queue and Threshold Mapping for 1p2q2t (tx):
Queue Threshold CoS
-----
1 1 0 1
1 2 2 3
2 1 4 6
2 2 7
3 1 5
Queue and Threshold Mapping for 1p1q4t (rx):
All packets are mapped to a single queue.
Rx drop thresholds:
Rx drop thresholds are disabled.
Tx drop thresholds:
Tx drop-thresholds feature is not supported for this port type.
```

Show QoS ACL Map

```
Cat6509> (enable) show qos acl map all
ACL name      Vlan #      Ports
-----      -
---
ACLdeny_http  1,4-7
ACLpermit_isp1 2            1/1

Cat6509> (enable) show qos acl map 1
Vlan          ACL name
----          -
1            ACLdeny_http

Cat6509> (enable) show qos acl map ACLpermit_isp1

ACL name      Vlan #      Ports
-----      -
---
ACLpermit_isp1 2            1/1
```

Show QoS ACL Map Runtime All

```

Cat6509> (enable) show qos acl map runtime all
ACL name          Type      Vlans
-----
aciHTTP_LIMIT     IP        1
ACL name          Type      Ports
-----
aciHTTP_LIMIT     IP
ACL name          Type      Vlans
-----
aciMB1            MAC
ACL name          Type      Ports
-----
aciMB1            MAC      4/1
Cat6509 (enable)
  
```

Show QoS Policer Runtime All

```

Cat6509> (enable) show qos policer runtime all
QoS microflow policers:
Microflow name      Avg. rate (kbps)  Burst size (kb)  Exceed action
-----
mpMB1              992                32                drop
ACL attached
-----
aciHTTP_LIMIT

QoS aggregate policers:
Aggregate name      Avg. rate (kbps)  Burst size (kb)  Exceed action
-----
apMB1              992                32                drop
ACL attached
-----
aciMB1
  
```

QoS ACL Statistics

```
Cat6509> (enable) show qos stat l3stats
Packets dropped due to policing:      189003
IP packets with ToS changed:         29843
IP packets with CoS changed:         4532
Non-IP packets with CoS changed:     17254
```

```
Cat6509> (enable) show qos acl resource-usage
ACL resource usage:
ACL storage (mask/value): 0.29%/0.7%
ACL to switch interface mapping table: 0.39%
ACL layer 4 port operators: 0.0%
```

Parting Thoughts

- **Plan QoS around application requirements**
- **Ensure that your hardware supports intended configuration**
- **Use meaningful names for ACEs, ACLs, and policers**
- **Check and test thoroughly**

Related Sessions

- **Session 2806—Campus Switch Architectures**
- **Session 2802—Performance Metrics for Multilayer Switches**
- **Session 2810—Campus Modular Switching Product Update**



Troubleshooting Catalyst Switches (Part 2)

Session 2808



Please Complete Your Evaluation Form

Session 2808

2808
1331_06_2000_c2 © 2000, Cisco Systems, Inc.

cisco.com

67

CISCO SYSTEMS



EMPOWERING THE INTERNET GENERATIONSM

2808
1331_06_2000_c2 © 2000, Cisco Systems, Inc.

www.cisco.com

68