

Cisco *live!*

February 15 - 19, 2016 • Berlin, Germany

We're ready. Are you?

End-to-End Qos Implementation and Operation with Nexus

Rohan Grover, Technical Marketing Manager

BRKDCT-3346

Session Objectives

- Provide a refresh of QoS and Queuing Basics
- Understand the basic hardware architecture for the Nexus Platforms
- Provide a detailed understanding of QoS on Nexus platforms
- Examine real-world deployment examples



Session Non-Objectives

- Nexus hardware architecture deep-dive
- Nexus 3000/9000 QOS (in Backup slides)
- Application Centric Infrastructure (ACI) QOS



Related Sessions

Session Id	Session Name
BRKARC-3470	Cisco Nexus 7000/7700 Switch Architecture
BRKARC-3452	Cisco Nexus 5600 and 6000 Architecture
BRKRST-2509	Mastering Data Center QoS

Agenda

- Introduction
- QoS and Queuing Basics
- QoS Implementation on Nexus
- Nexus 7000/7700 QoS
- Nexus 5x00/6000 QoS
- Nexus 2000 QoS
- Real World Deployment
- Conclusion

Introduction

“Quality of service (QoS) is the overall performance of a telephony or computer network, particularly the performance seen by the users of the network.”

Wikipedia.org

“On the Internet and in other networks, QoS (Quality of Service) is the idea that transmission rates, error rates, and other characteristics can be measured, improved, and, to some extent, guaranteed in advance.”

Margaret Rouse
Whatis.com

Congestion Happens Everyday!

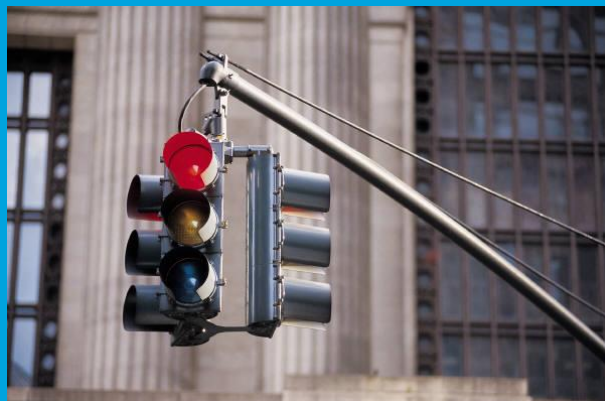


Why QoS in the Data Center?

**Assign
Color to Traffic**



**Manage
Congestion**



**Maximize
Throughput**



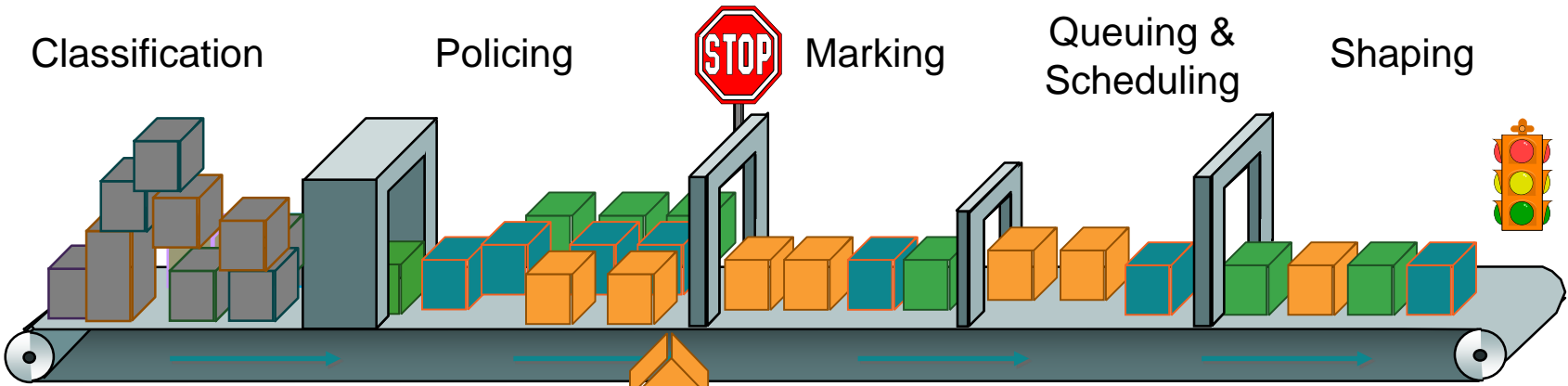
Maximize Throughput and Manage Congestion!

Agenda

- Introduction
- **QoS and Queuing Basics**
- QoS Implementation on Nexus
- Nexus 7000/7700 QoS
- Nexus 5x00/6000 QoS
- Nexus 2000 QoS
- Real World Deployment
- Conclusion

QoS and Queuing Basics

The QoS Toolset



Identify and Split Traffic into Different Classes

Discard Misbehaving Traffic to Maintain Network Integrity

Mark Traffic According to Behavior and Business Policies

Prioritize, Protect and Isolate Traffic Based on Markings

Control Bursts and Conform Traffic



Traffic Management Tools

Classification

- Traffic Categorization

Marking

- Traffic Re-categorization

Policing

- Limit misbehaving flows

Cisco *live!*



Classification and Marking – Two sides of a coin

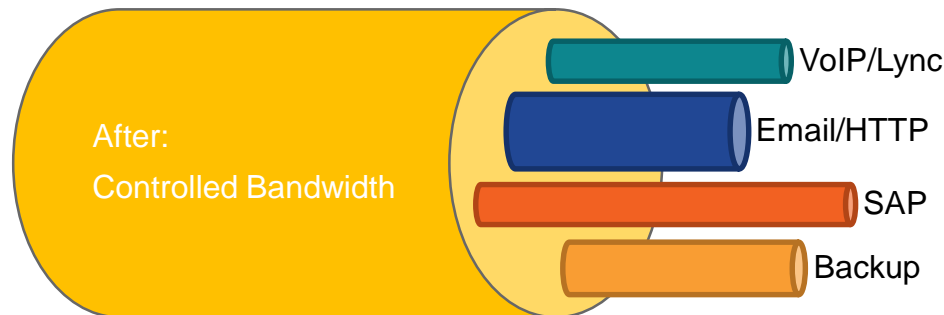
Identify traffic

- DSCP
- CoS
- ACLs
- IP PREC



Remap Traffic

- Like to Like (i.e CoS to CoS)
- Like to Unlike (i.e DSCP to COS)
 - Needs mapping tables
 - Also called Mutation



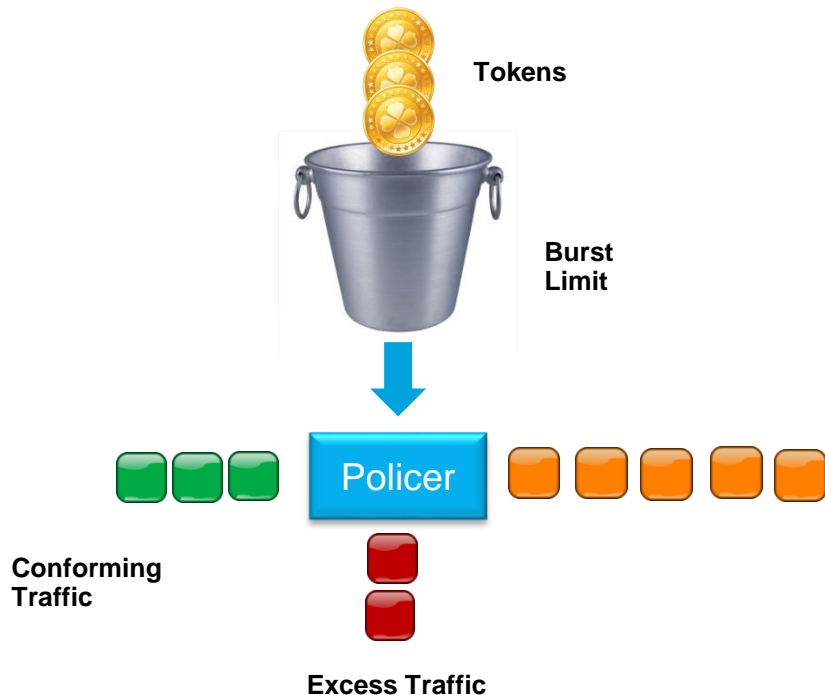
Policing – limit misbehaving traffic

Single rate Two Color Policer

- Conform Action (permit)
- Exceed Action (drop)

Two rate Three Color Policer

- Conform Action (permit)
- Exceed Action (markdown)
- Violate Action (drop)



Congestion Management Tools

Buffering

- Storing packets in memory

Queuing

- Buffering packets according to traffic class

Scheduling

- Order of transmission of buffered packets

Shaping

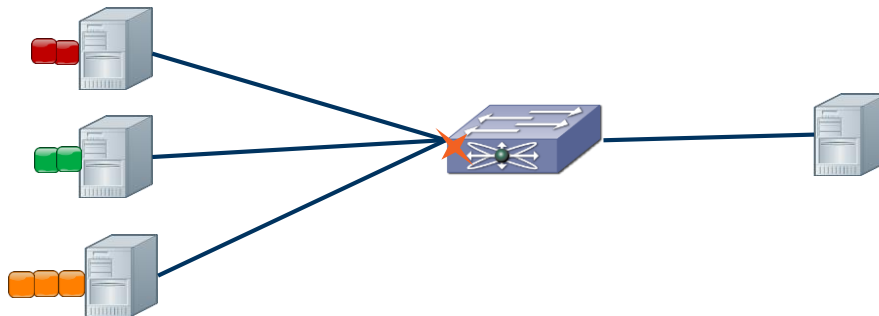
- Smooth bursty traffic



Buffering – Why do we need it?

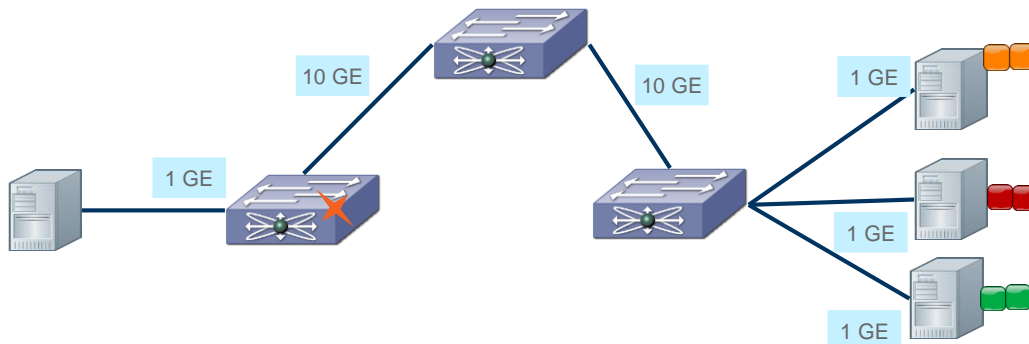
Many to One Conversations

- Client to Server
- Server to Storage
- Aggregation Points



Speed Mismatch

- Client to WAN to Server



Scheduling – Who goes first?

- Defines **Order of transmission**
- The **Priority-Queue** always serviced first
- **Normal Queues** serviced only after Priority Queue empty
- Different Scheduling algorithms for normal queues

Common Scheduling Algorithms

Round Robin (RR)

- Simple and **Easy to implement**
- Starvation-free

Weighted Round Robin (WRR)

- Serves n packets per non-empty queue
- Assumes a **mean packet size**

Deficit Weighted Round Robin

- **Variable sized** packets
- Uses a deficit counter

Shaped Round Robin

- More **even distributed ordering**
- Weighted interleaving of flows

4 Class Queuing Model example

Class	CoS	Queues
Priority	5-7	PQ
No-Drop	3, 4	Q2
Better than Best-Effort	2	Q1
Best-Effort	0, 1	Default-Q

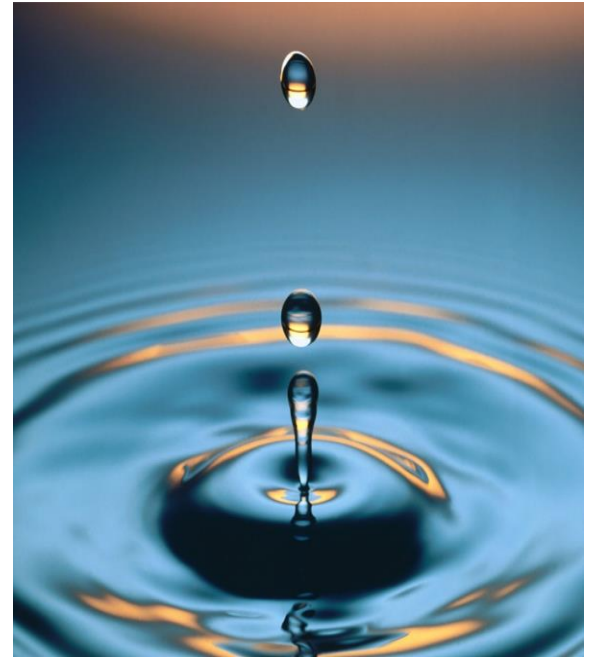
Congestion Avoidance Tools

Tail Drop (TD)

- Drop packets at **tail of the queue**
- **Single threshold** per queue

Weighted Random Early Drop (WRED)

- One or more thresholds per queue
- Threshold associated with **DSCP or COS**



Putting it all together!

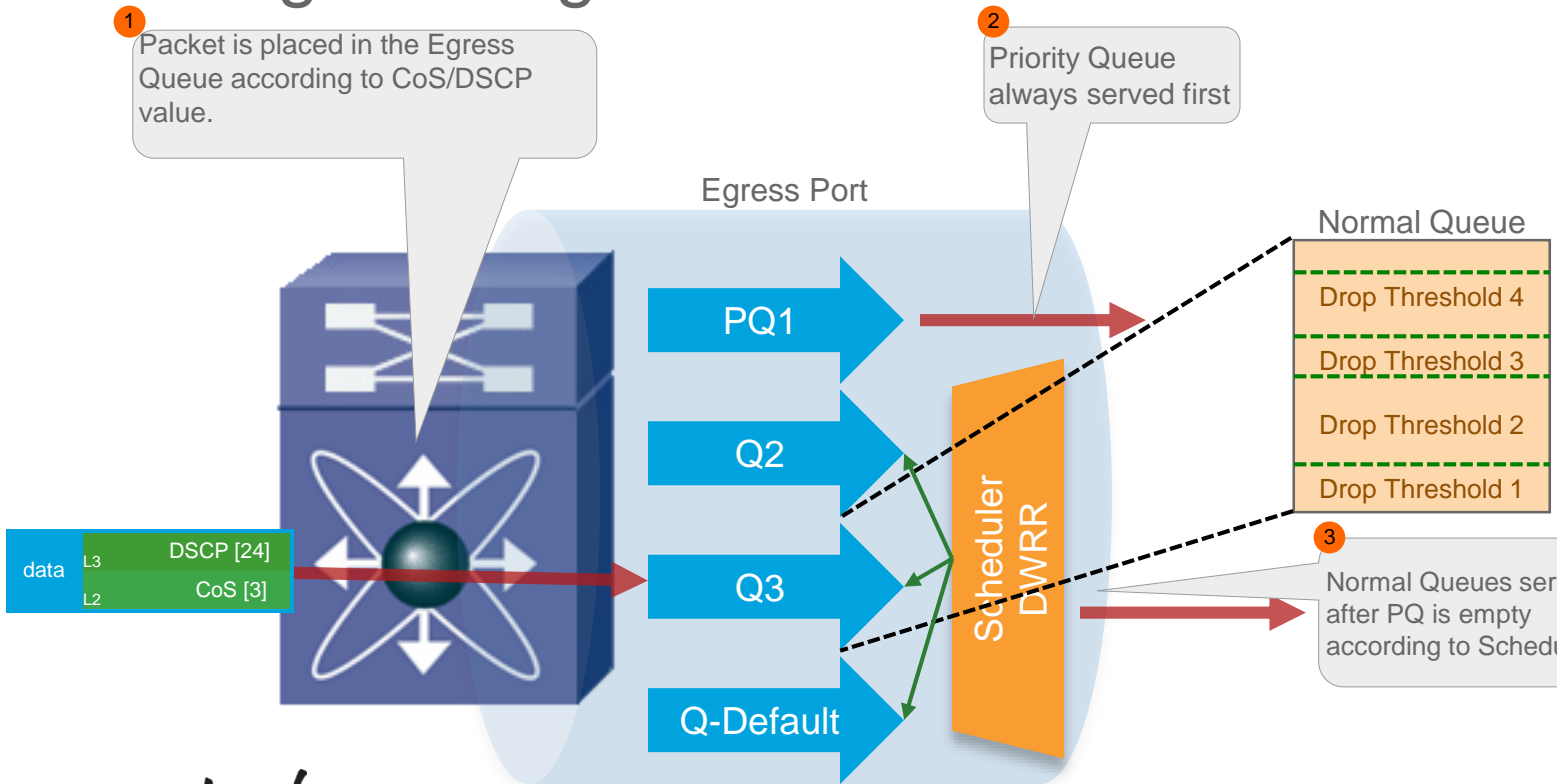
1 Packet is placed in the Egress Queue according to CoS/DSCP value.

2 Priority Queue always served first

4 Normal Queues have Threshold and drop packet accordingly

3 Normal Queues served after PQ is empty according to Scheduler

COS	Queue
5, 6, 7	PQ1
3, 4	Q3
2	Q2
0, 1	Q-Default



Congestion could result in?

Head

of

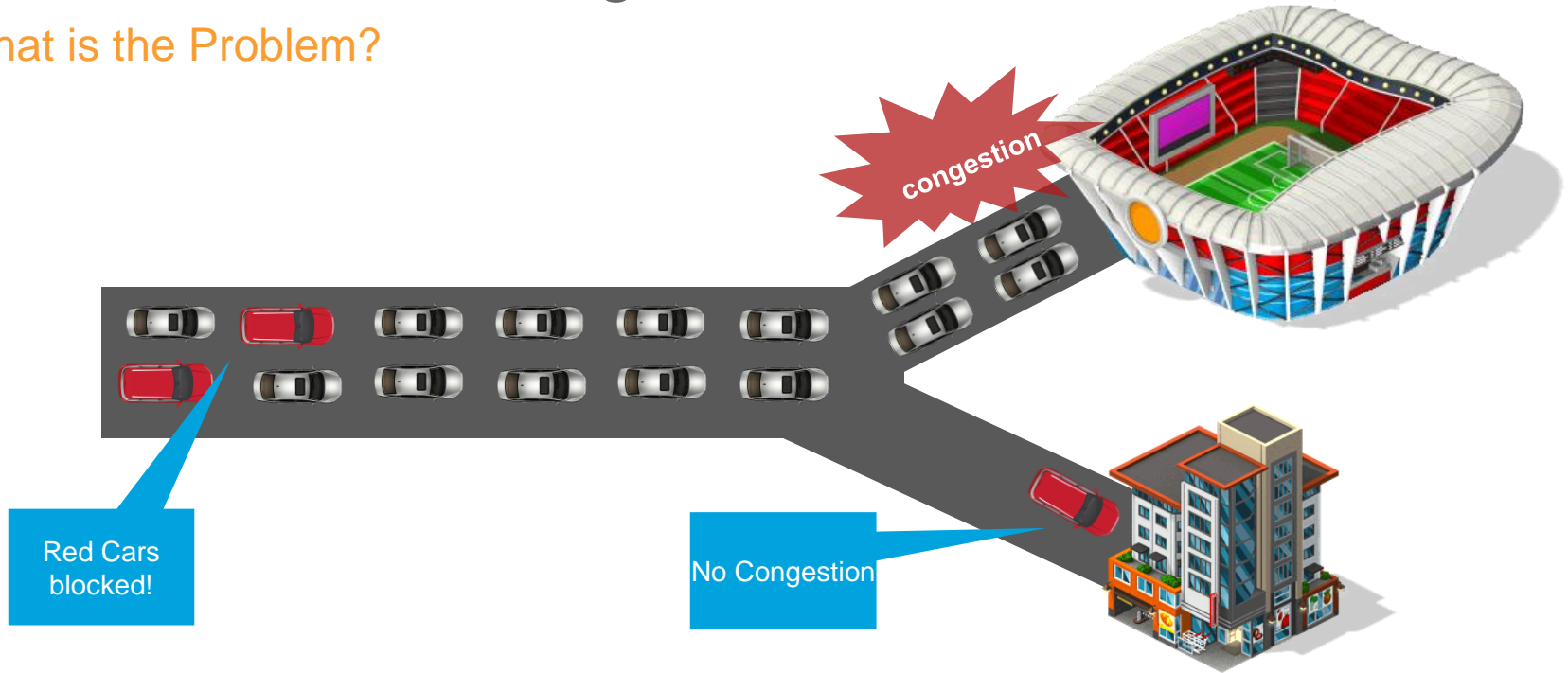
Line

Blocking



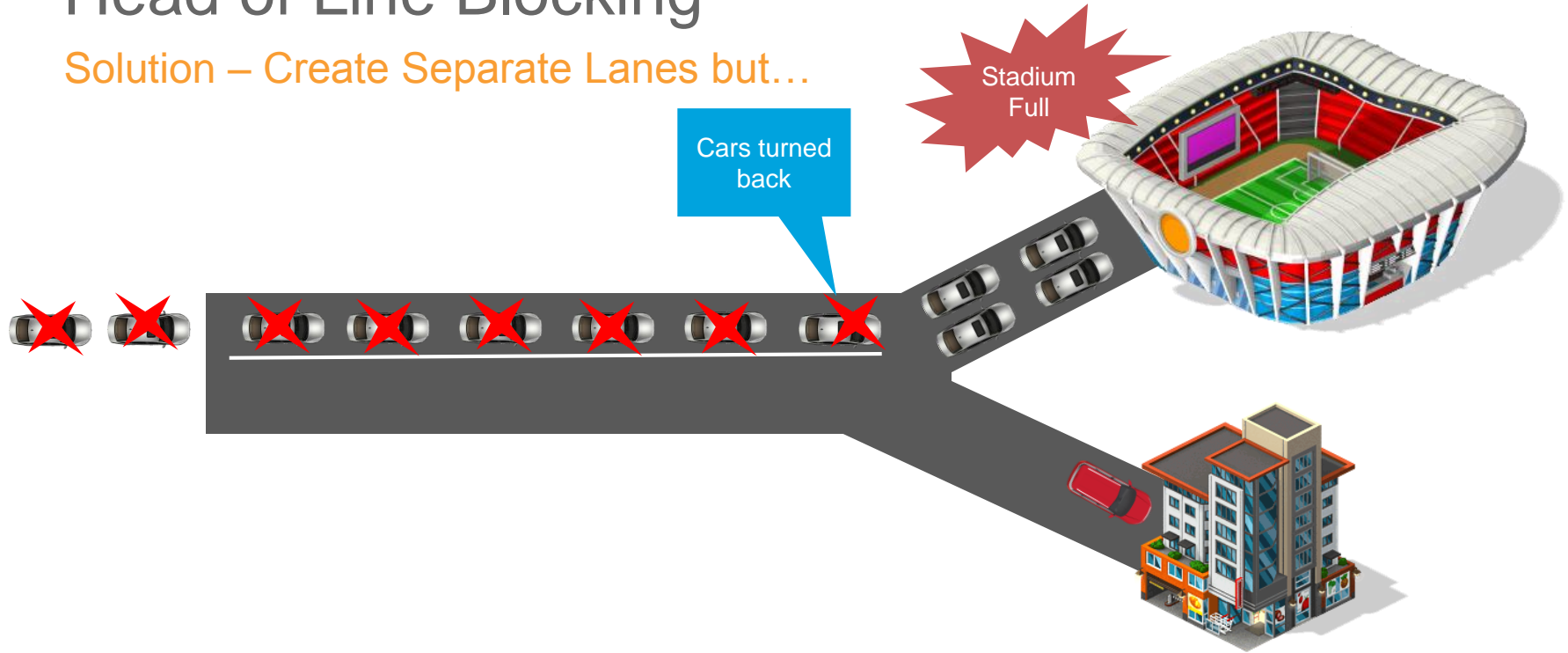
Head of Line Blocking

What is the Problem?



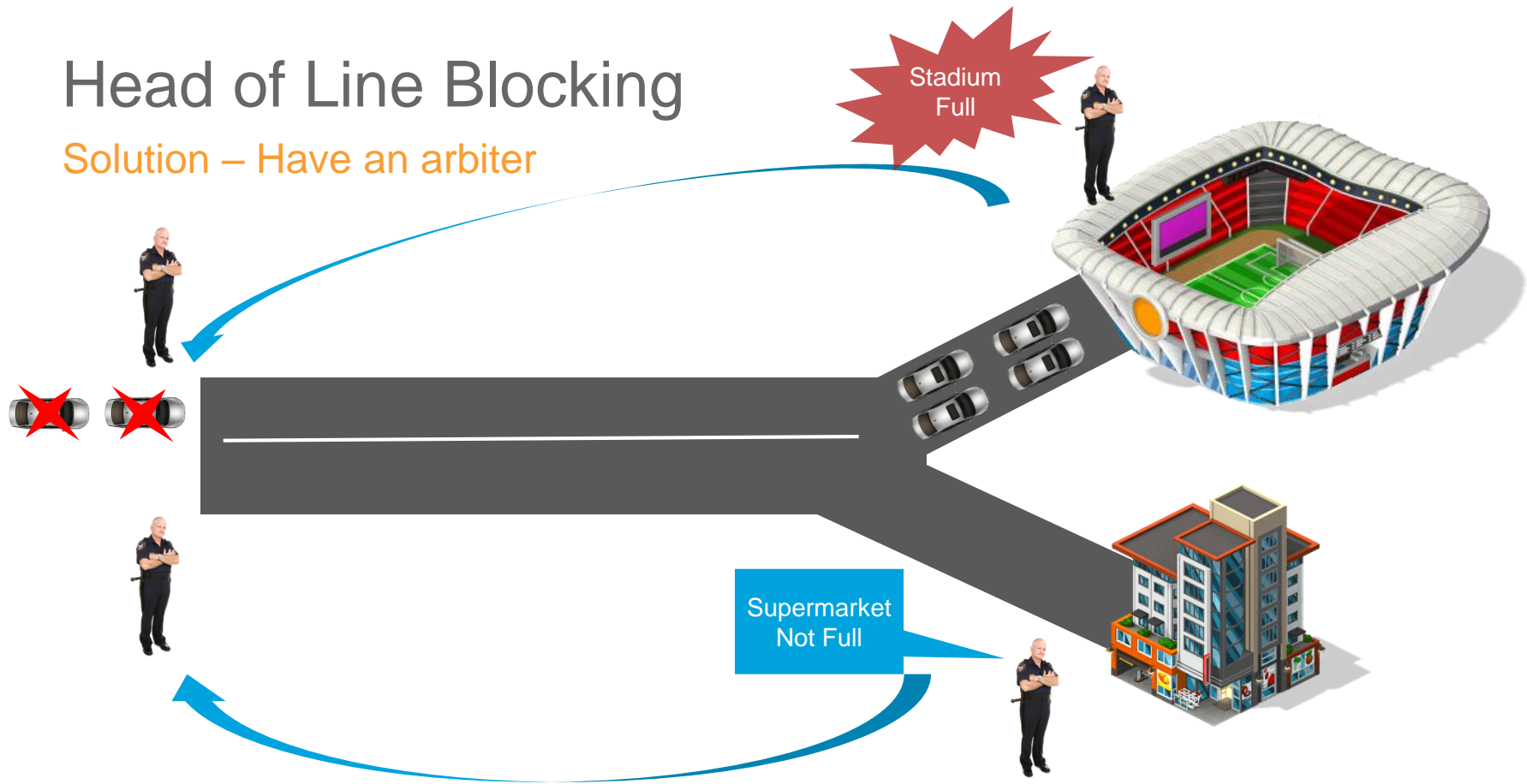
Head of Line Blocking

Solution – Create Separate Lanes but...



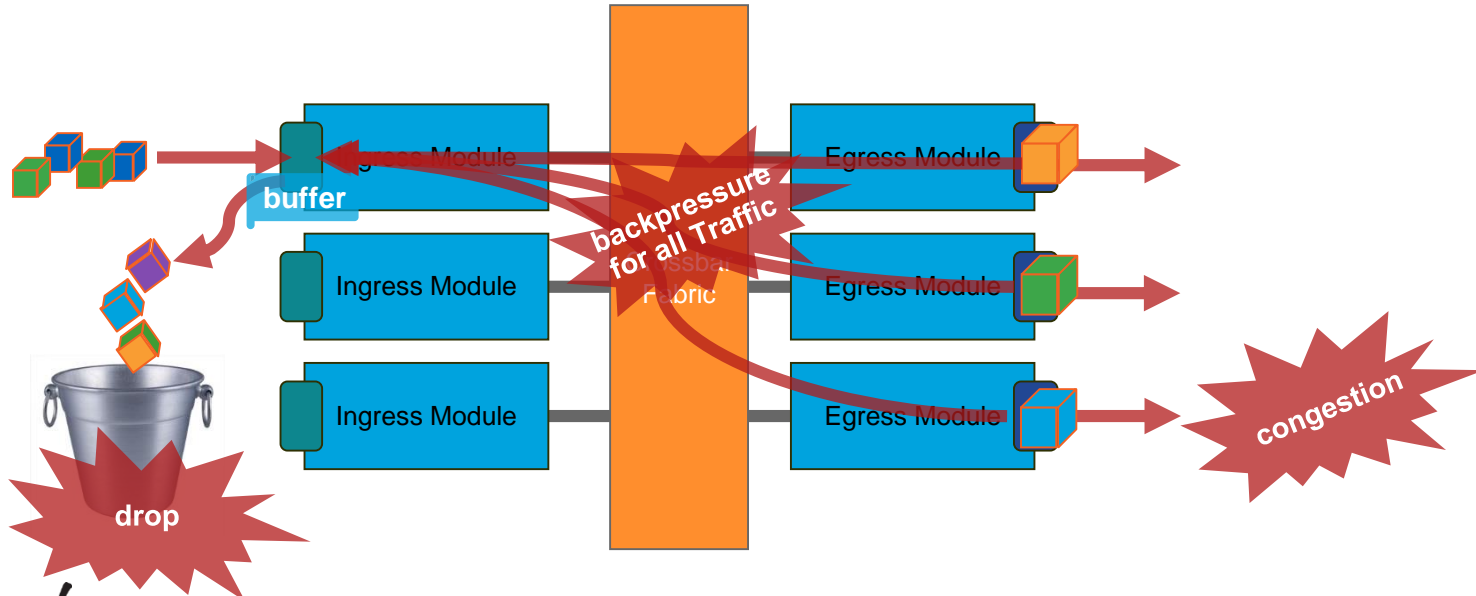
Head of Line Blocking

Solution – Have an arbiter



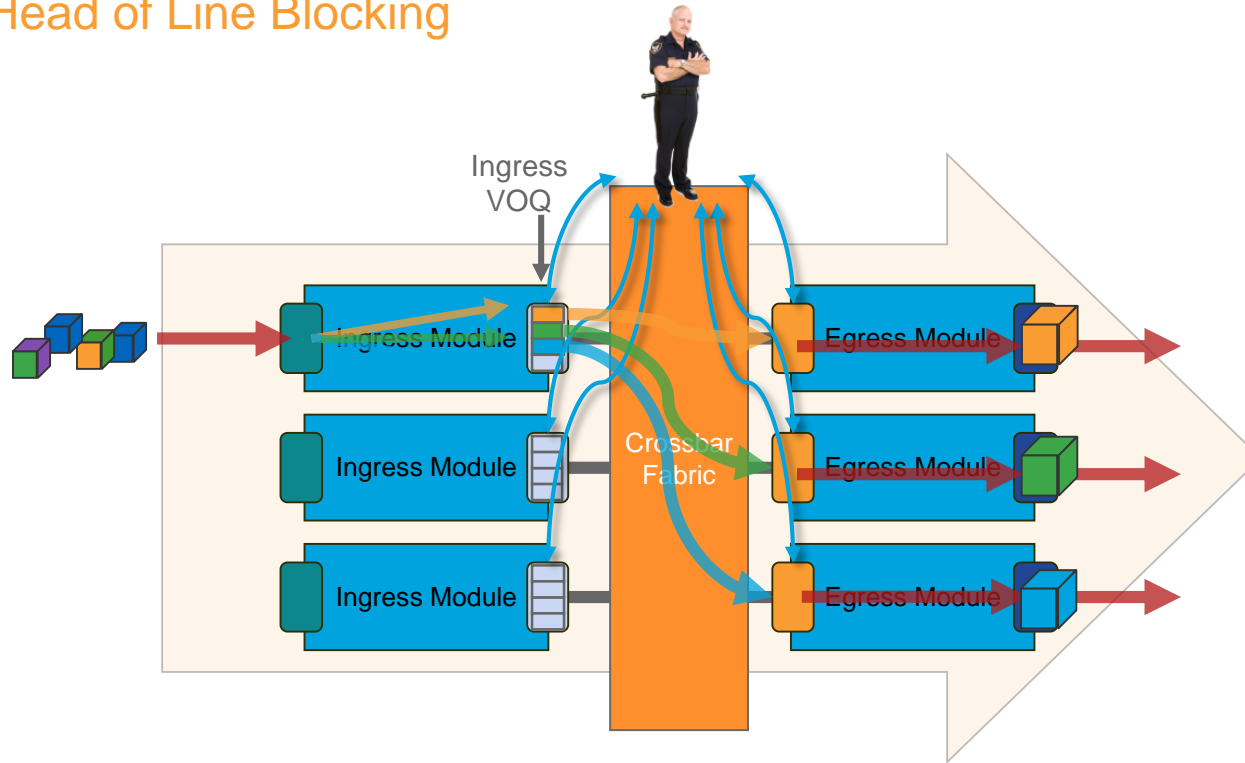
Head of Line Blocking

What is the Problem?



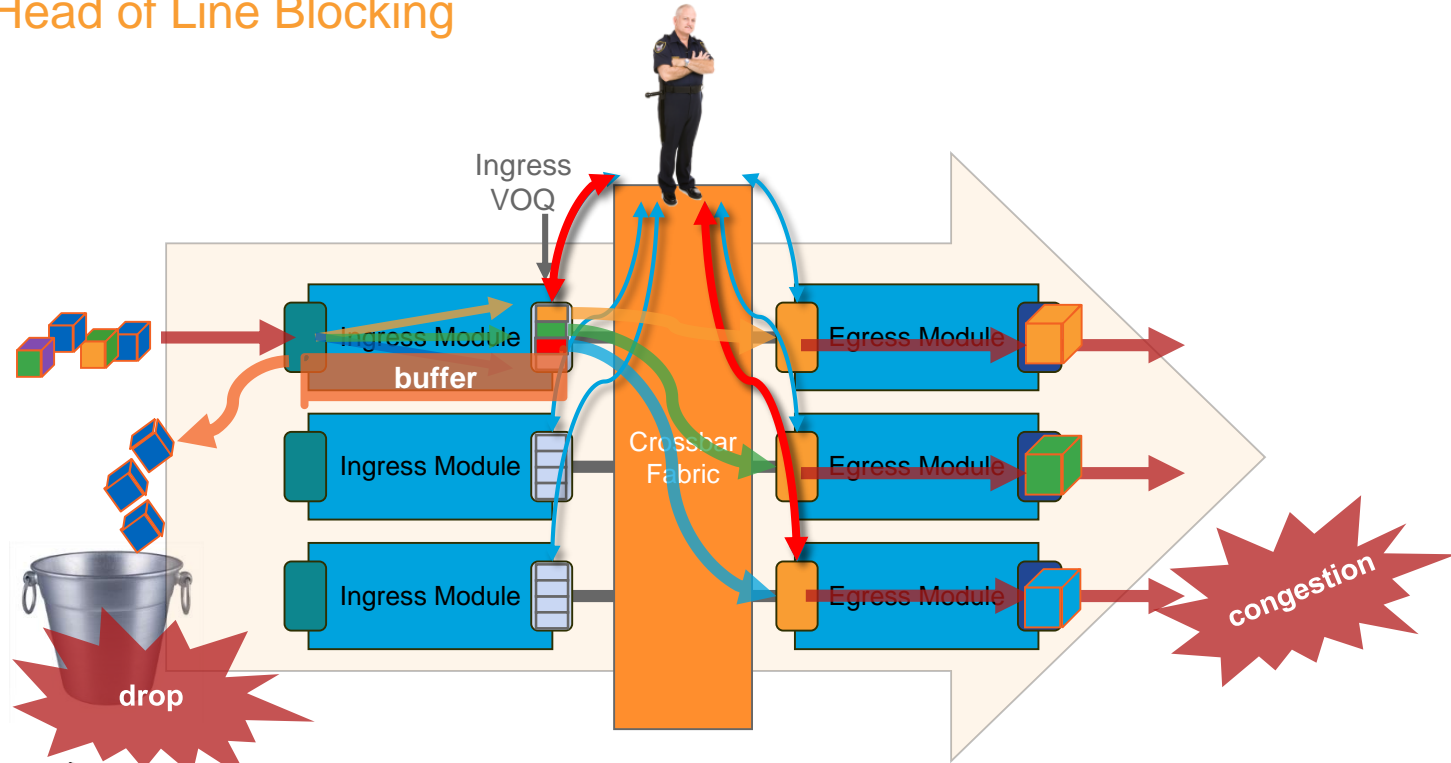
Virtual Output Queues

Avoid Head of Line Blocking



Virtual Output Queues

Avoid Head of Line Blocking



Agenda

- Introduction
- QoS and Queuing Basics
- QoS Implementation on Nexus
- Nexus 7000/7700 QoS
- Nexus 5x00/6000 QoS
- Nexus 2000 QoS
- Real World Deployment
- Conclusion

QoS Implementation on Nexus

Nexus uses Modular QOS CLI (MQC)

3 Block Construct

Class-Map

What Traffic do we care about?

- DSCP
- CoS
- IPPREC
- ACLs

Policy-Map

What action do I take on these classes?

- Policing
- Marking
- Scheduling

Service-Policy

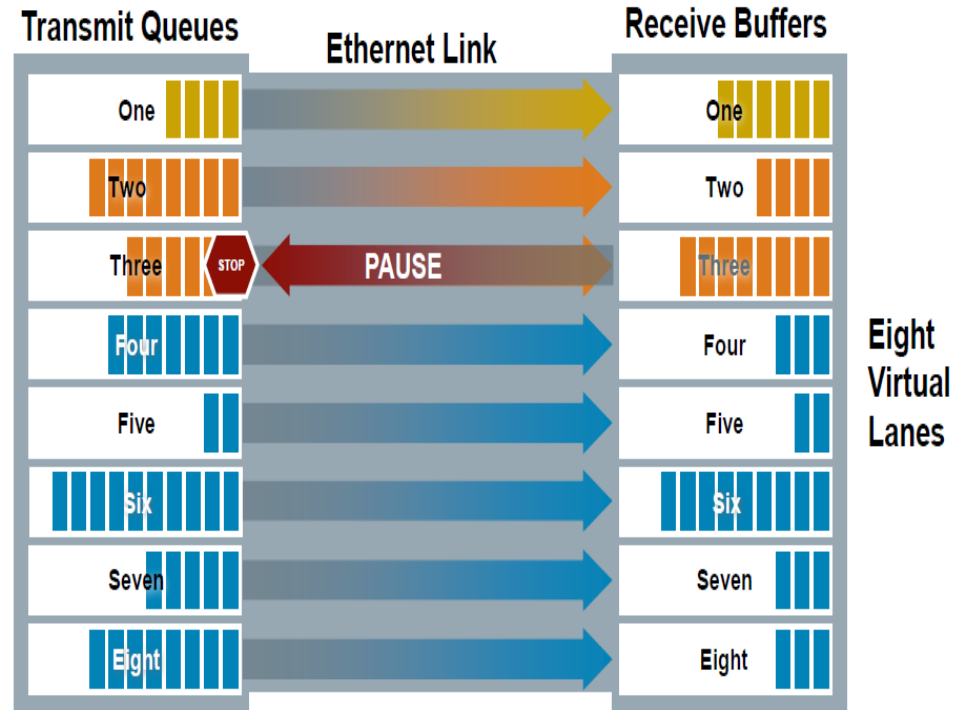
Where do I apply this policy?

- System Wide
- Vlan
- Interface (in/out)
- Port-channels

New Data Center QoS Capabilities

Priority Flow Control (802.1Qbb)

- Enables **Lossless Ethernet** using per COS pause
- During congestion, **no-drop COS** is paused
- No effect on other COS values



New Data Center QoS Capabilities

DCBXP (802.1Qaz)

- LLDP with new TLV Values
- **Negotiates capabilities** (like PFC) with other devices



ECN (Explicit Congestion Notification)

- Congestion Notification without dropping packets
- Uses **two LSB bits in DiffServ field** in IP header

DiffServ Field Values

0x00 – Non ECN Capable

0x10 – ECN Capable Transport (0)

0x01 – ECN Capable Transport (1)

0x11- Congestion Encountered

Data Center Applications

FCoE (Fiber Channel over Ethernet)

- Replaces the FC0 and FC1 of the Fiber Channel stack with Ethernet
- Reduces cabling and applicable to server virtualization applications

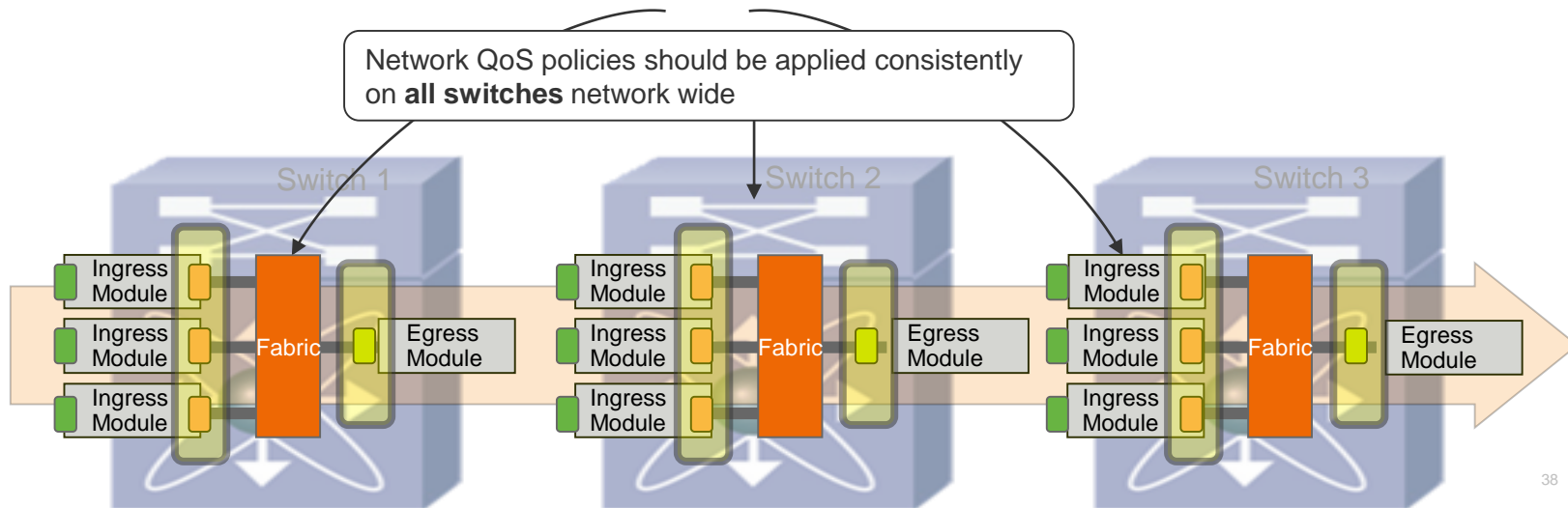
RoCE (RDMA over Converged Ethernet)

- Remote Direct Memory Access allows network adapters to do server-to-server data transfer between application memory
- RoCE is a IBTA (InfiniBand Trade Association) standard that brings RDMA capabilities over Converged Ethernet



Network-QoS Policy

- Define global queuing and scheduling parameters for all interfaces in switch
 - Identify drop/no-drop classes, MTU and WRED/TD, etc.
- One network-QoS policy per system, applies to all ports
- Assumption is network-QoS policy defined/applied consistently network-wide



Three Different Types

Class-map

QoS

- CoS
- DSCP
- PREC
- ACLs

Queuing

- CoS
- DSCP

Network-QoS

- CoS
- Protocol (fcoe)

Policy-map

QoS

- Marking
- Policing
- Mutation

Queuing

- Buffering
- Queuing
- Scheduling

Network-QoS

- Congestion-Control
- Pause / MTU per VL

Service-policy

QoS

- Interfaces
- Vlans
- Port-channel
- System-qos

Queuing

- Interfaces
- Port-channels
- System-qos

Network-QoS

- System-qos

To Trust or Not To Trust?

- Data Centre architecture provides a new set of **trust boundaries**
- Virtual Switch extends the **trust boundary into the Hypervisor**
- Hence Nexus Switches **always trust CoS and DSCP**



Agenda

- Introduction
- QoS and Queuing Basics
- QoS Implementation on Nexus
- **Nexus 7000/7700 QoS**
- Nexus 5x00/6000 QoS
- Nexus 2000 QoS
- Real World Deployment
- Conclusion

Nexus 7000/7700 QoS

Nexus 7000 / 7700 I/O Module Families

NEXUS 7000



M1 1G and 10G

M2 10G / 40G / 100G

M Series Modules

L2/L3/L4 with large forwarding tables and rich feature set

NEXUS 7000



F1 10G

F2 10G

F2E 10G

F2E 10G

F3 40G

F Series Modules

High performance, low latency with streamlined feature set

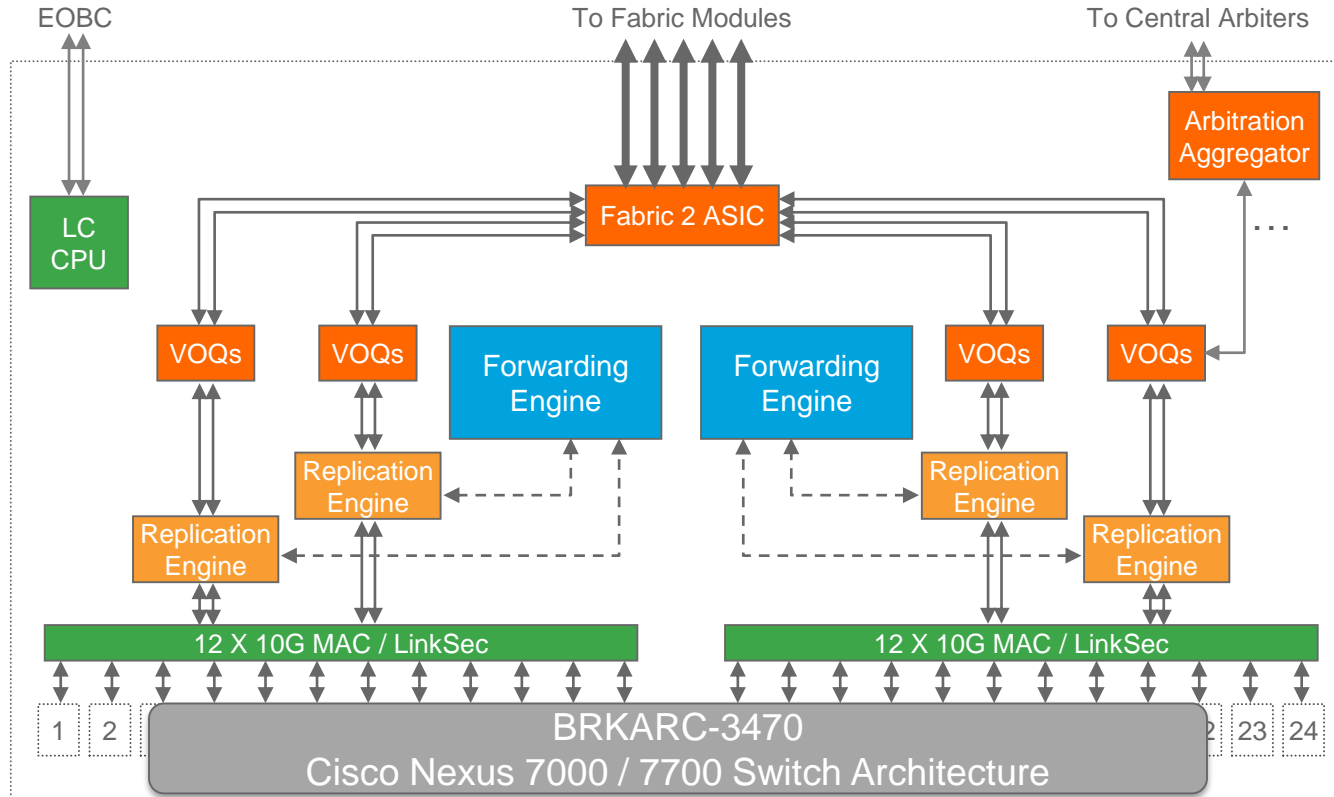
F3 10G / 40G / 100G

NEXUS 7700

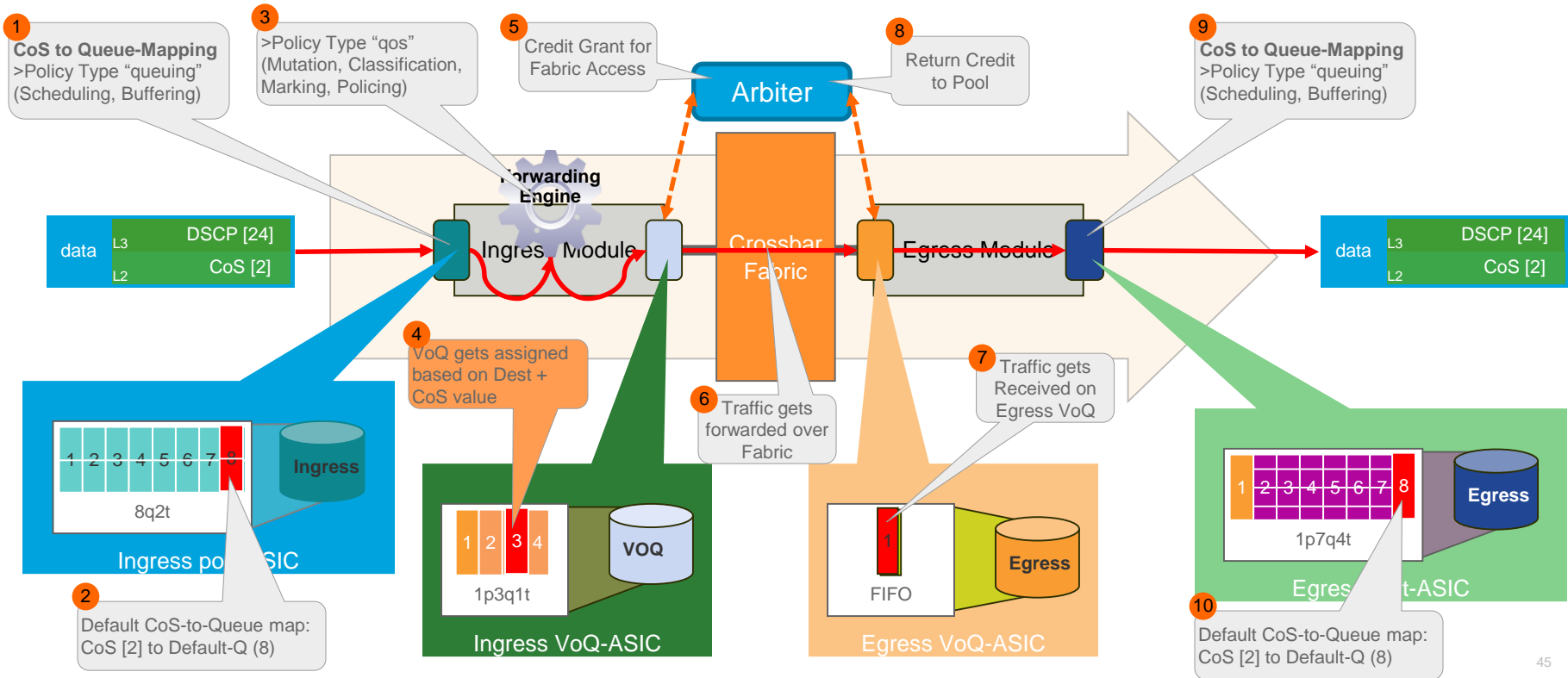


Cisco *live!*

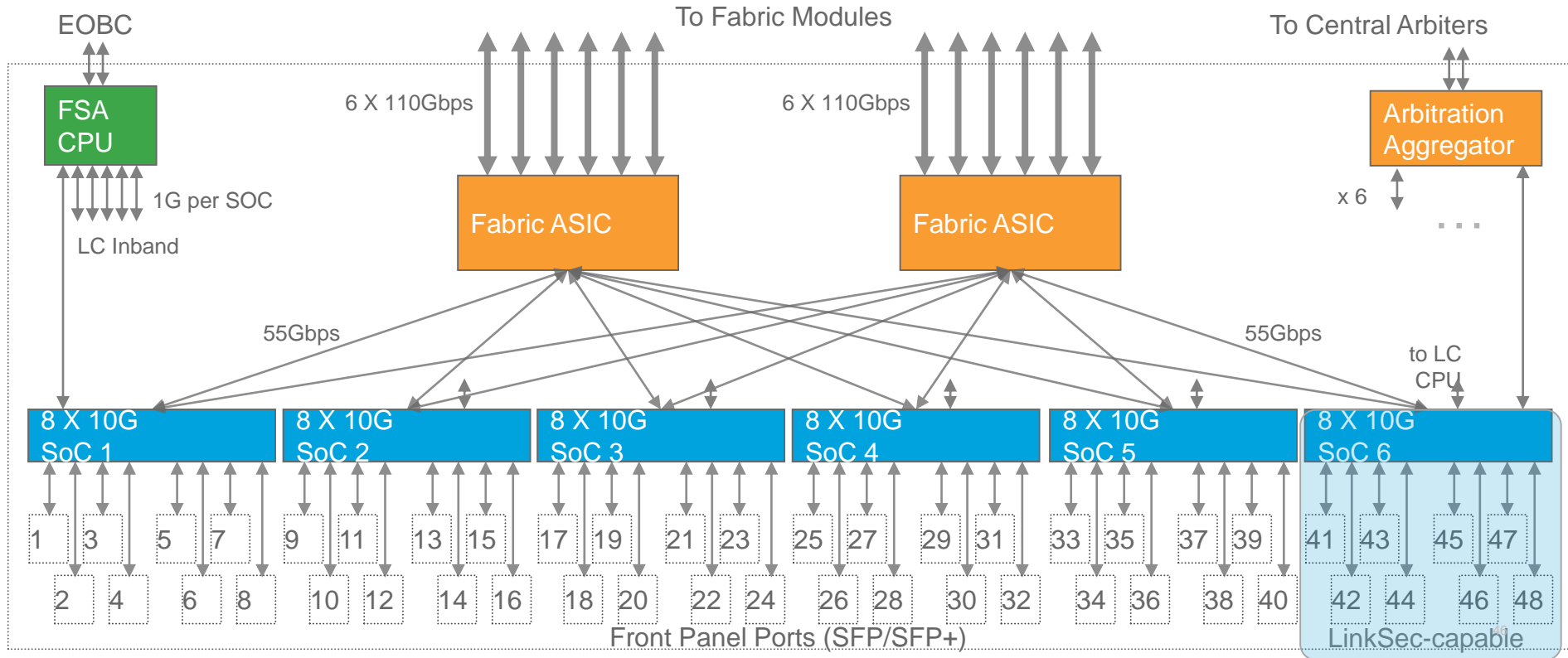
M2 I/O Module Architecture



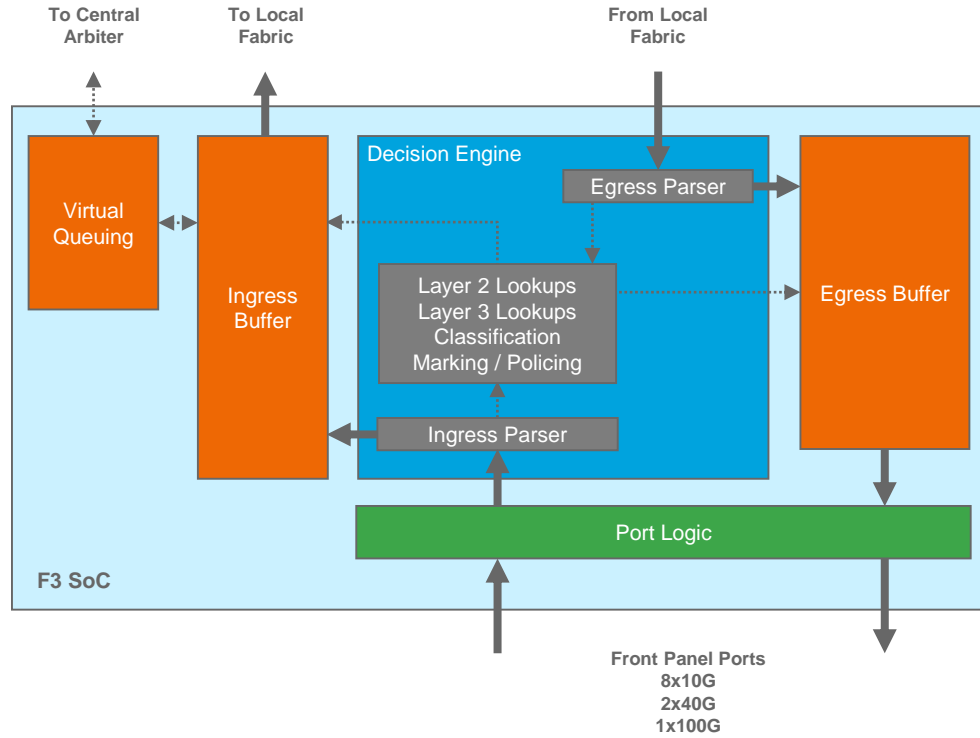
M2 I/O QoS Packet-Flow (10G)



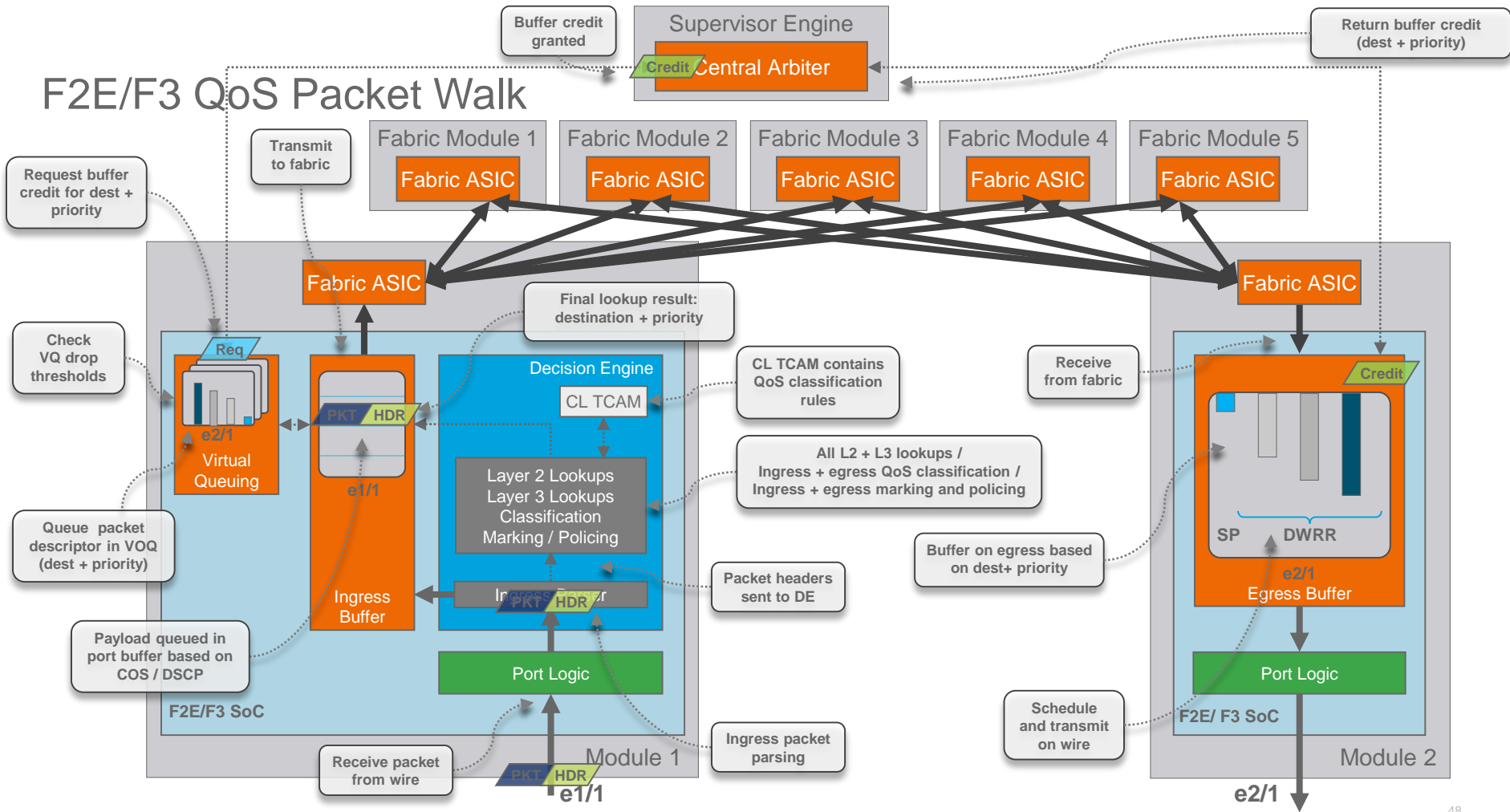
Nexus 7700 F3 48-Port 1G/10G Module Architecture



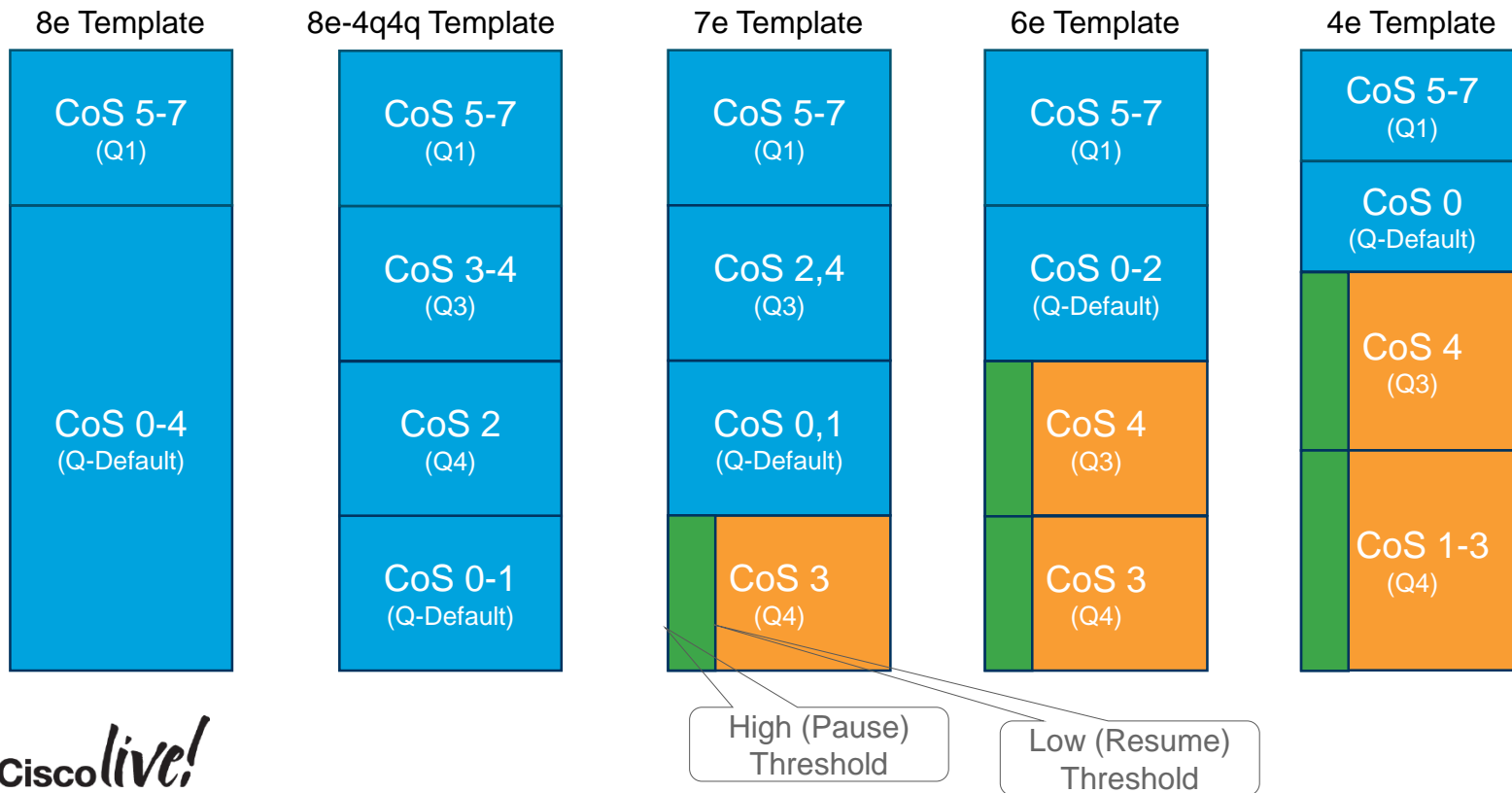
F3 SoC Architecture



F2E/F3 QoS Packet Walk



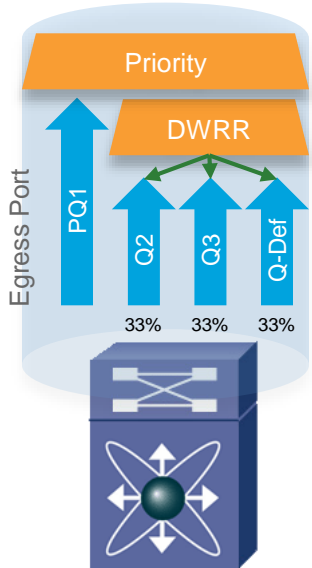
Ingress Queuing – Logical View



Egress Queuing – Logical View

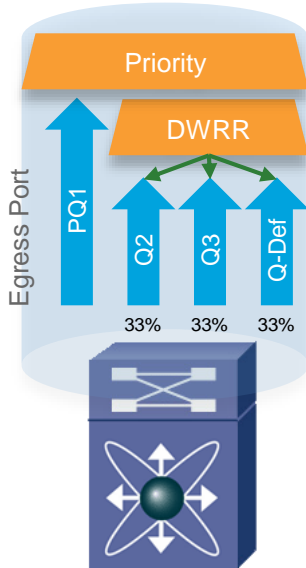
8e Template

PQ1 **Q2** **Q3** **Q-Def.**
(5,6,7) (3,4) (2) (0,1)



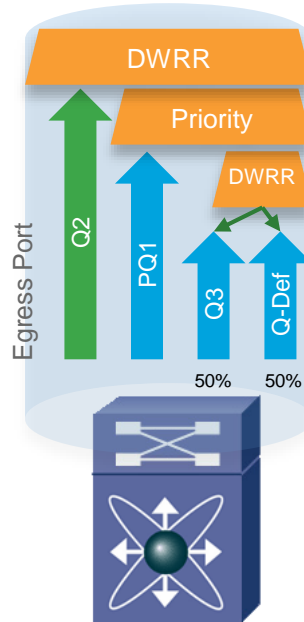
8e-4q4q Template

PQ1 **Q2** **Q3** **Q-Def.**
(5,6,7) (3,4) (2) (0,1)



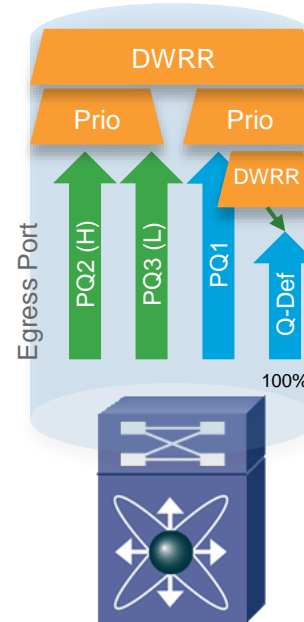
7e Template

Q2 **PQ1** **Q3** **Q-Def.**
(3) (5,6,7) (2,4) (0,1)



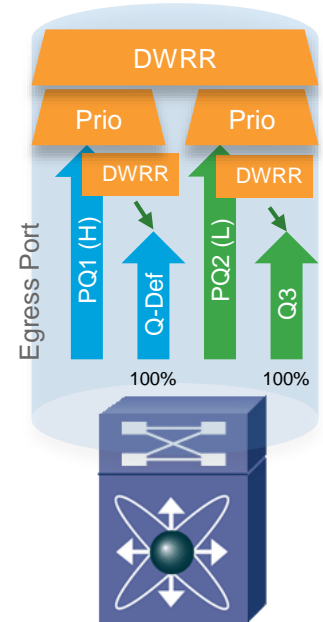
6e Template

PQ2. **PQ3** **PQ1** **Q-Def.**
(4) (3) (5,6,7) (0-2)



4e Template

PQ1 **Q-Def.** **PQ2** **Q3**
(5,6,7) (0) (4) (1,2,3)

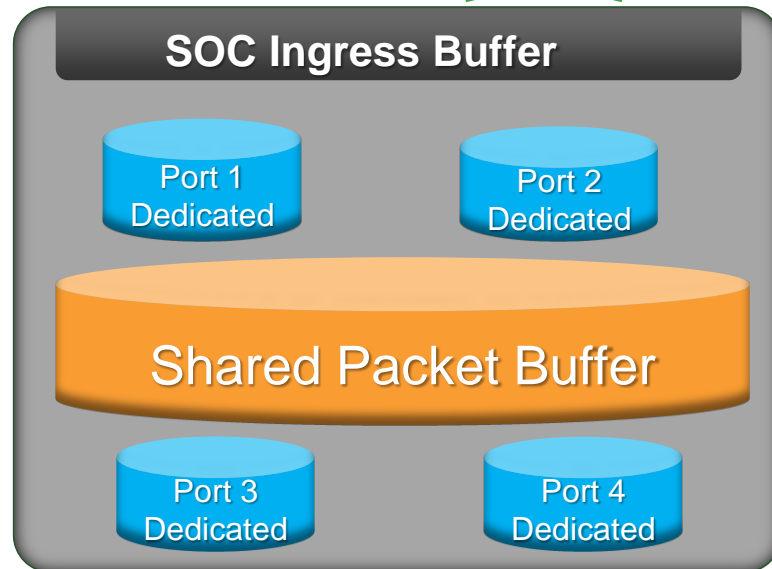


red indicates no-drop

Shared Buffer Queuing on F3

6.2.10
onwards

- Disabled by Default and enabled on per module-basis
- Default ratio is 50:50 for dedicated/shared
- Ratio can be changed per port-group (SOC) basis



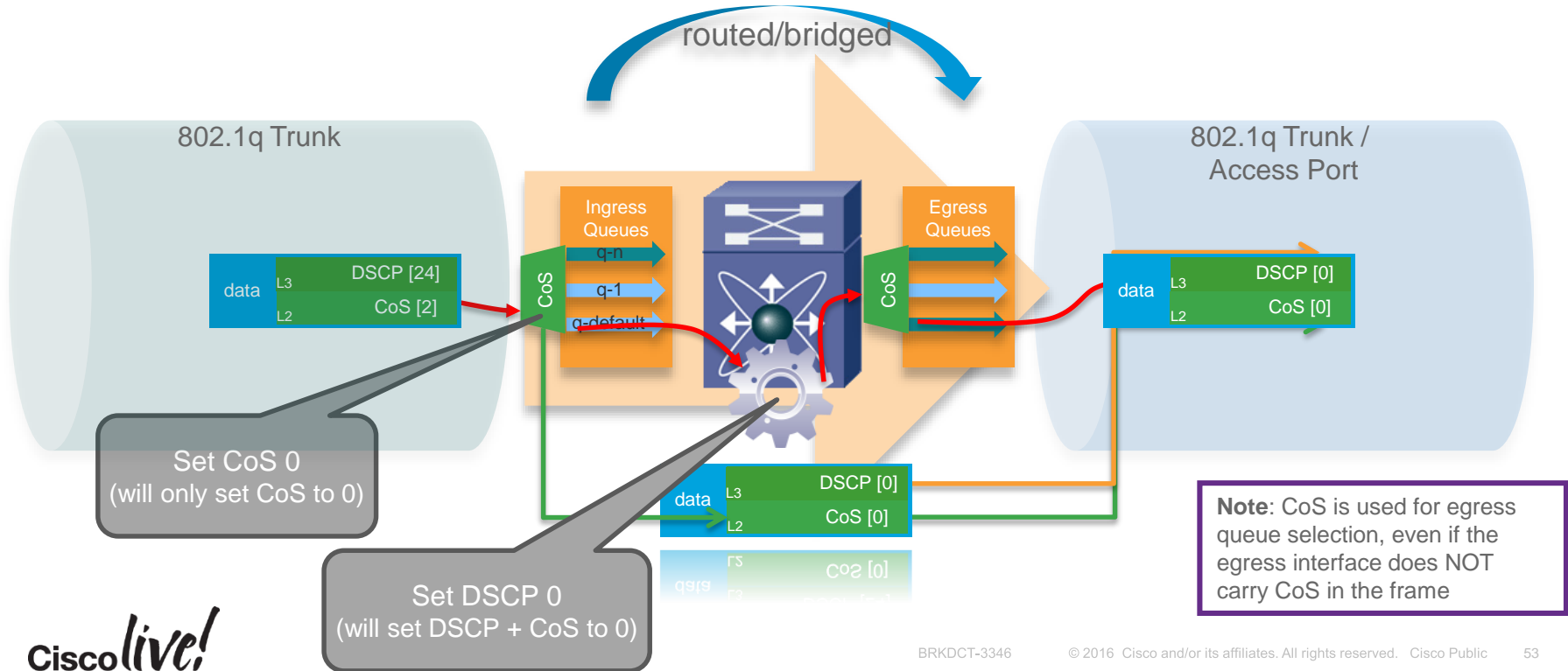
```
Switch(config)# hardware qos shared-buffer module module-number
switch(config)# hardware module module-number port-group port-group-number
switch(config-port-group)# qos shared-buffer queue-limit percent
```

CoS or DSCP to Queue Mapping

- **Default CoS to Queue Mapping** for **Nexus 7000** (F- and M-Series I/O Module)
 - Ingress: CoS to Queue
 - Egress: CoS to Queue
- **Default DSCP to Queue Mapping** for **Nexus 7700**
 - Ingress: DSCP to Queue
 - **Egress: CoS to Queue**
- Global Configuration (Admin/Default VDC) required to enable DSCP to Queue Mapping:

```
N7k(config)# hardware qos dscp-to-queue ingress module type {all | f-series | m-series}
```


Changing the Default Trust



Nexus 7000 QoS Golden Rules

- QoS is **enabled by default** and cannot be disabled
- CoS and DSCP are **TRUSTED by default**
- Default Queuing and QoS policies are applied to all physical interfaces across all VDCs
- For bridged traffic, CoS is preserved, DSCP is unmodified
- For routed traffic, DSCP is copied to CoS (first 3 bits)
 - Ex: DSCP 40 (b101000) becomes CoS 5 (b101)

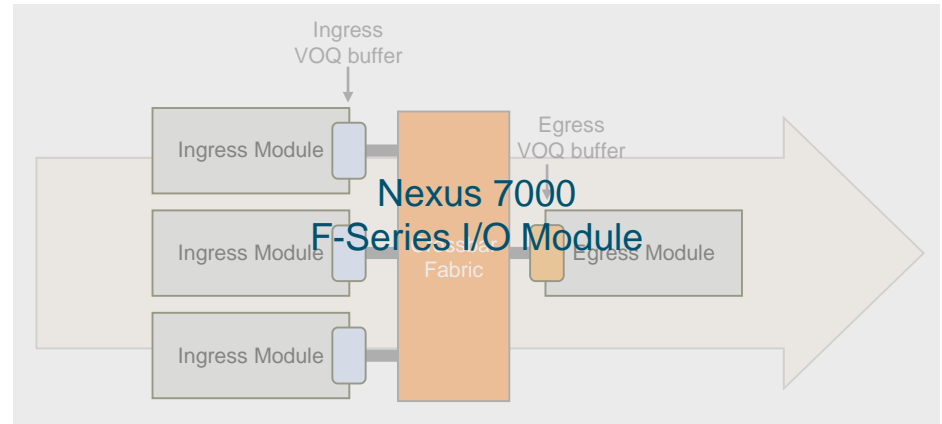
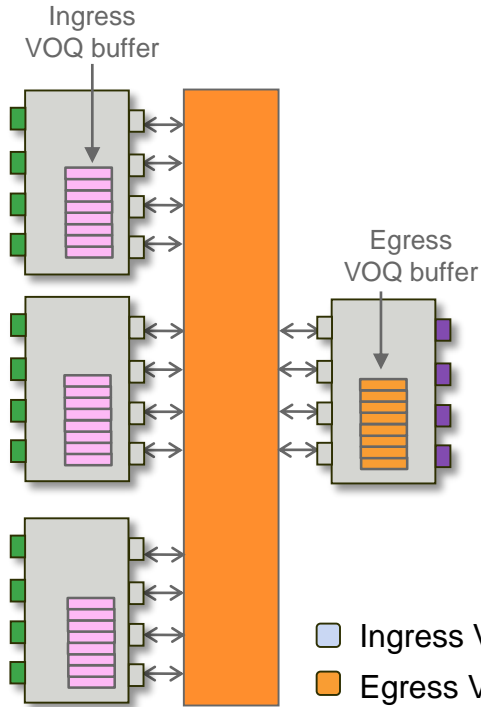
Agenda

- Introduction
- QoS and Queuing Basics
- QoS Implementation on Nexus
- Nexus 7000/7700 QoS
- **Nexus 5x00/6000 QoS**
- Nexus 2000 QoS
- Real World Deployment
- Conclusion

Nexus 5000/6000 QoS

Ingress Buffering Model

Nexus 5500/5600/6000 compared to Nexus 7000 F-Series I/O Module



- Ingress VOQ buffer – Manages congestion toward egress destinations
- Egress VOQ buffer – Receives frames from fabric; also buffers multi-destination frames

Key Concepts – Common Points

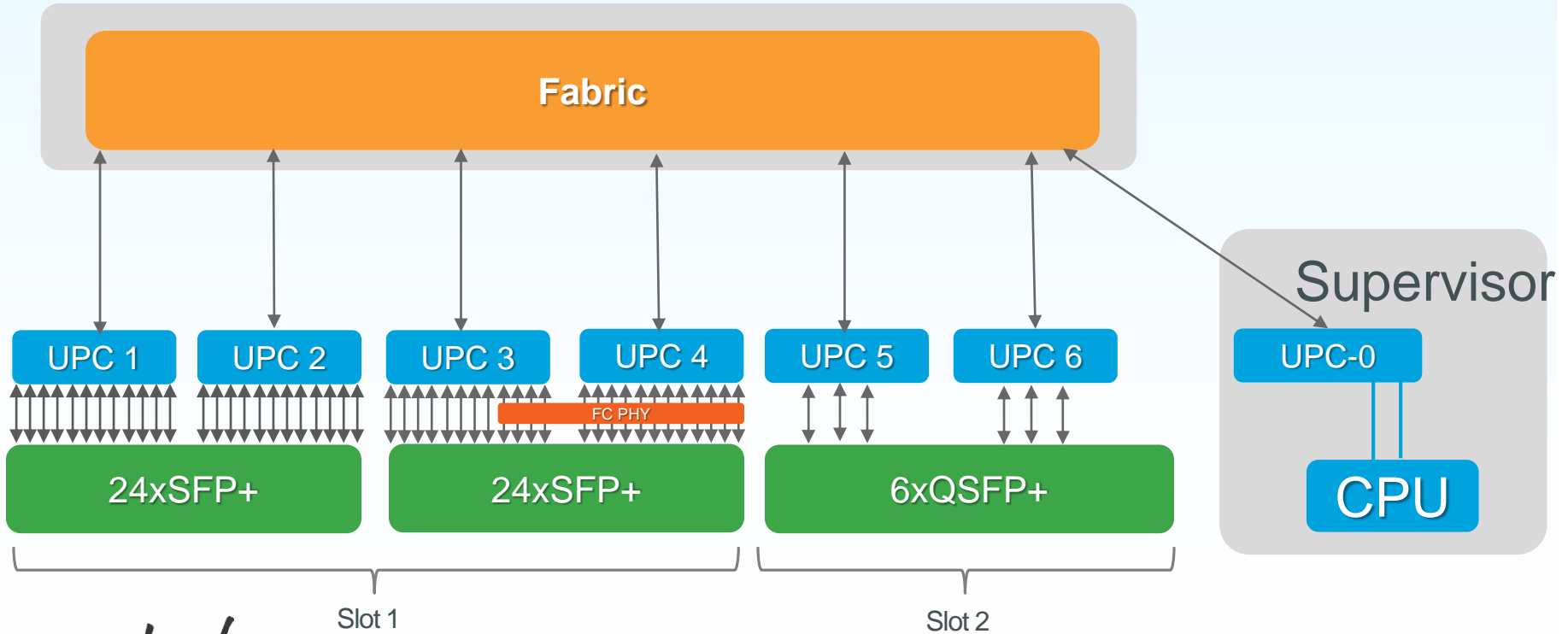
Nexus 7000 compared to Nexus 5000/6000 QoS

- Nexus 5000/6000 & Nexus 7000 F-Series I/O Modules share the **Ingress Buffer Model**
- Ingress buffering and queuing occur at **VOQ of each ingress port**
- Egress scheduling enforced by egress port
- No Egress QOS Policies

Cisco Nexus 5600/6000 QoS Features

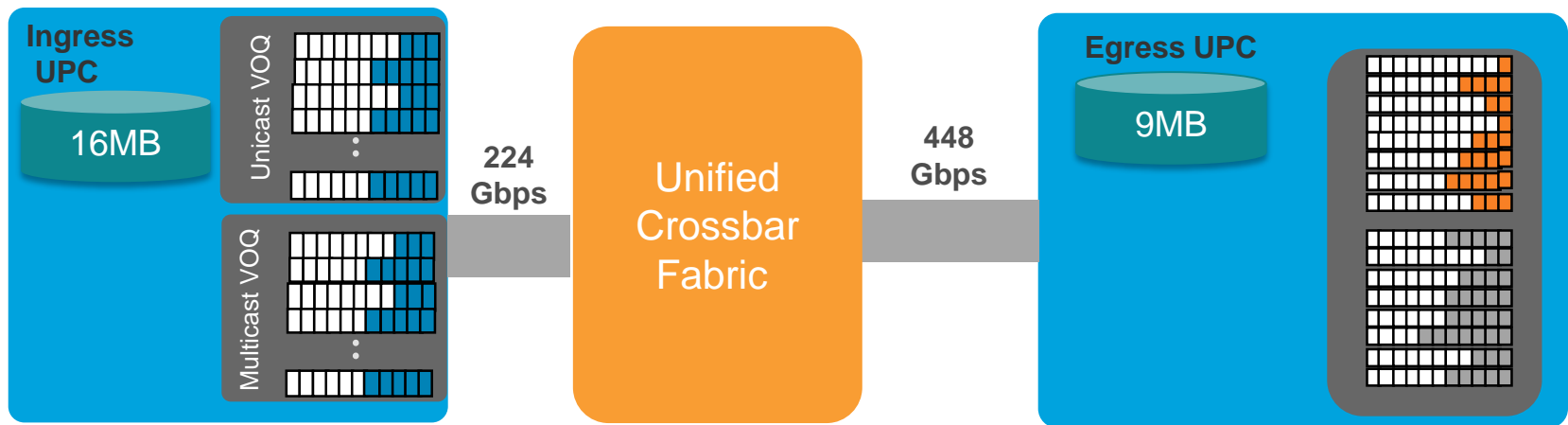
- **Traffic classification**
 - DSCP, CoS, IP Precedence and ACL
- **Packet marking**
 - DSCP, CoS, and ECN
- **Strict Priority Queuing and DWRR**
 - Priority Flow Control
 - DCBX 802.1Qaz
- **Ingress policing (No egress policing)**
 - 4096 policers per ASIC
- **Flexible buffer management**

Cisco Nexus 5672UP Internal Architecture



Packet Buffering

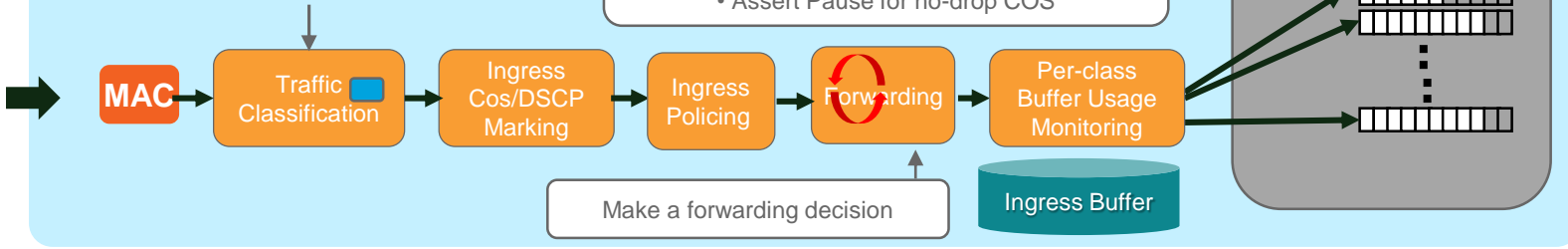
- 25MB packet buffer is shared by every three 40 GE ports or twelve 10 GE ports.
- Buffer is 16MB at ingress and 9MB at egress.
- Unicast packet can be buffered at both ingress and egress.
- Multicast Buffered at egress only



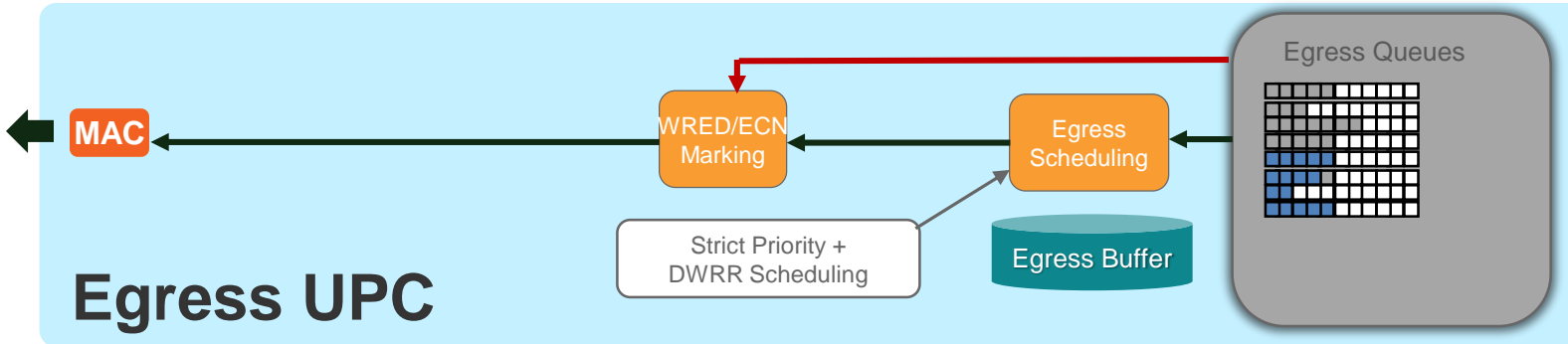
Nexus 5600 QoS Processing

Ingress UPC

Trust CoS/DSCP and Match on COS/DSCP/ACL



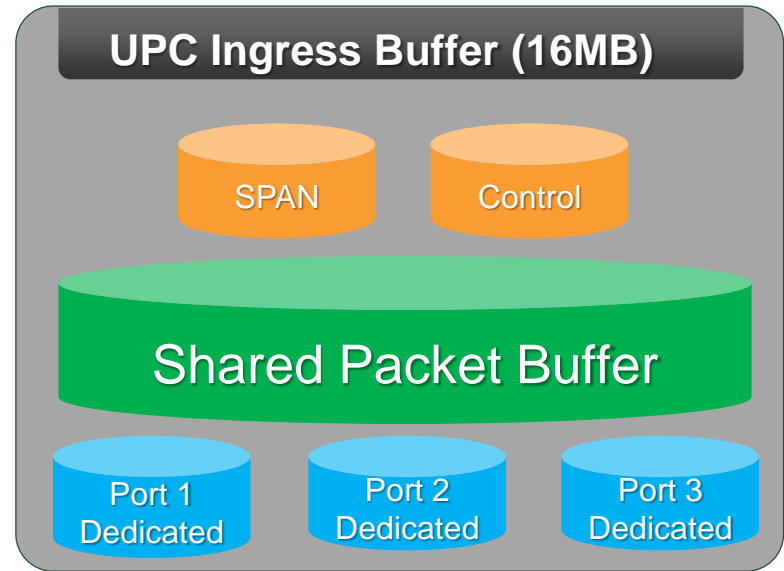
Egress UPC



Flexible Buffer Management

Ingress Buffer

- Shared buffer is good for burst absorption.
- Dedicated buffer is good for predictable performance for each port.
- On by default, no configuration needed
- Long-distance FCoE, video editing (i.e., AVID), Big Data, and distributed storage



Tune Buffer Allocation at Ingress

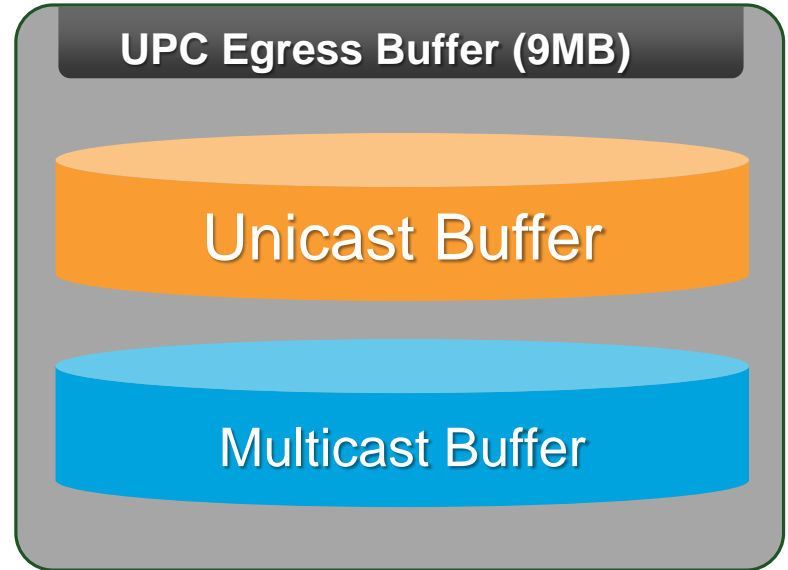
- “queue-limit” under “network-qos” policy specifies the dedicated buffer for each port and each class. The dedicated buffer can be used by the port for only that class of service.
- Without “queue-limit” each class of service will get 100 KB of dedicated buffer.
- The size of dedicated buffer can be different for different classes of service. The policy applies to all ports in the chassis.
- Total ingress buffer minus the dedicated buffer and buffer for control and SPAN will be in the shared buffer pool.
- The following example sets the dedicated buffer for “class-default” to be 400 KB for all ports.

```
switch(config)# policy-map type network-qos Policy-buffer
switch(config-pmap-nq)# class type network-qos class-default
switch(config-pmap-nq-c)# queue-limit 400000 bytes
switch(config-pmap-nq-c)# system qos
switch(config-sys-qos)# service-policy type network-qos Policy-buffer
```

Flexible Buffer Management

Egress Buffer

- 9-MB packet buffer is shared among three 40 GE or twelve 10 GE.
- CLI is provided to allocate buffer between unicast and multicast (future).
- Unicast traffic can be buffered at egress and ingress.
- Multicast is buffered at egress in case of interface oversubscription.



WRED/ ECN Configuration

- ECN parameters are configurable only at system level.
- ECN is enabled by default along with WRED
- Packet Threshold below minimum – **Transmit**
- Packet Threshold between minimum and maximum – **Mark ECN bits**
- Packet Threshold above maximum – **Drop**

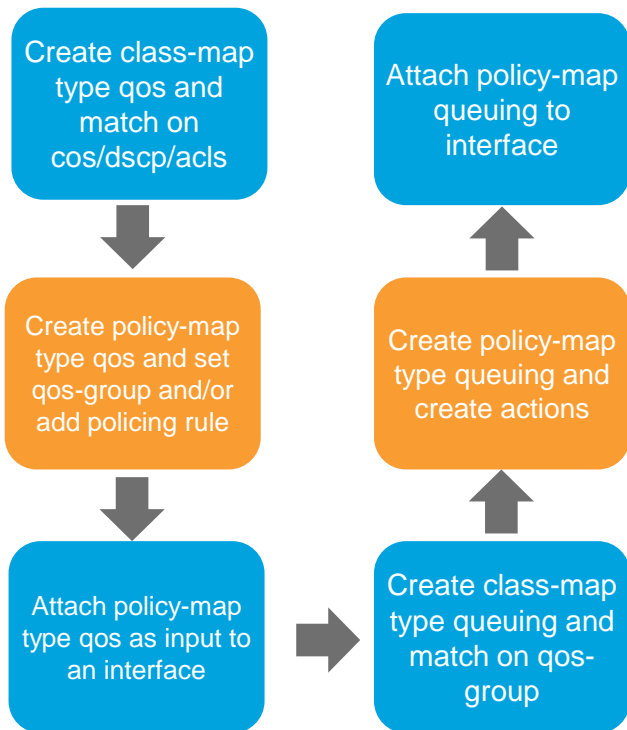
```
switch(config)# hardware random-detect min-thresh 10g 10g-min-threshold 40g 40g-min-  
threshold max-thresh 10g 10g-max-threshold 40g 40g-max-threshold ecn qos-group qos-  
group-number
```

Nexus 5600/6000 QoS Configuration Model

- Uses QoS-Groups to tie together QoS, Queuing and Network-QoS policies
- QoS-Group has no direct relation with CoS value
- QoS-Groups defined (set) in `policy-map type qos`.
- QoS-groups referenced (match) in `policy type queuing` and `policy-map type network-qos`



Putting it all together



```
class-map type qos class_foo
  match cos 3-4

policy-map type qos pm1
  class type qos class_foo
    set qos-group 1
  police cir 20 mbytes conform transmit violate drop
  class type qos class-default
    set qos-group 0

interface ethernet 1/1
  service-policy type qos input pm1

class-map type queuing class-foo
  match qos-group 1

policy-map type queuing policy-foo
  class type queuing class-foo
    bandwidth percent 20
  class type queuing class-default
    bandwidth percent 80

interface ethernet 1/3
  service-policy type queuing input policy-foo
```


Nexus 5600/6000 QoS Golden Rules

- ECN/WRED is enabled by default and cannot be disabled
- CoS and DSCP are TRUSTED by default
- Use QoS-Groups to tie policies together
- No Egress QOS policies

Agenda

- Introduction
- QoS and Queuing Basics
- QoS Implementation on Nexus
- Nexus 7000/7700 QoS
- Nexus 5x00/6000 QoS
- **Nexus 2000 QoS**
- Real World Deployment
- Conclusion

Nexus 2000 QoS

FEX Overview

- Scalable and Extensible Fabric
- Single point of management
- Homogeneous and consistent policies

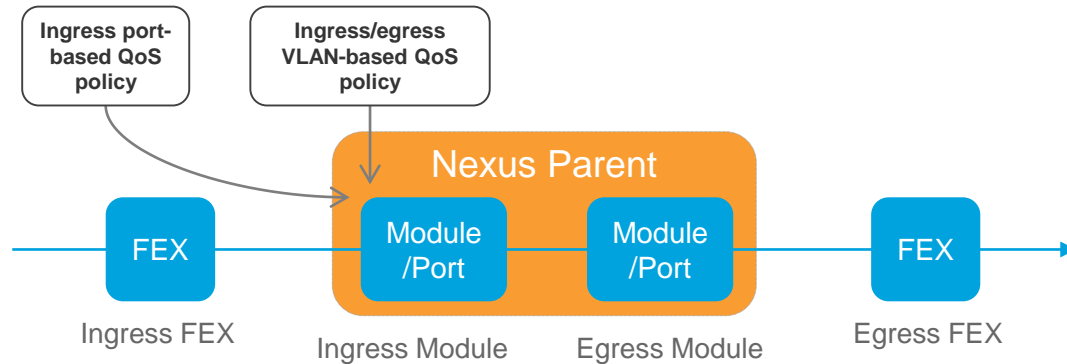


Cisco Nexus 2000 QoS Features

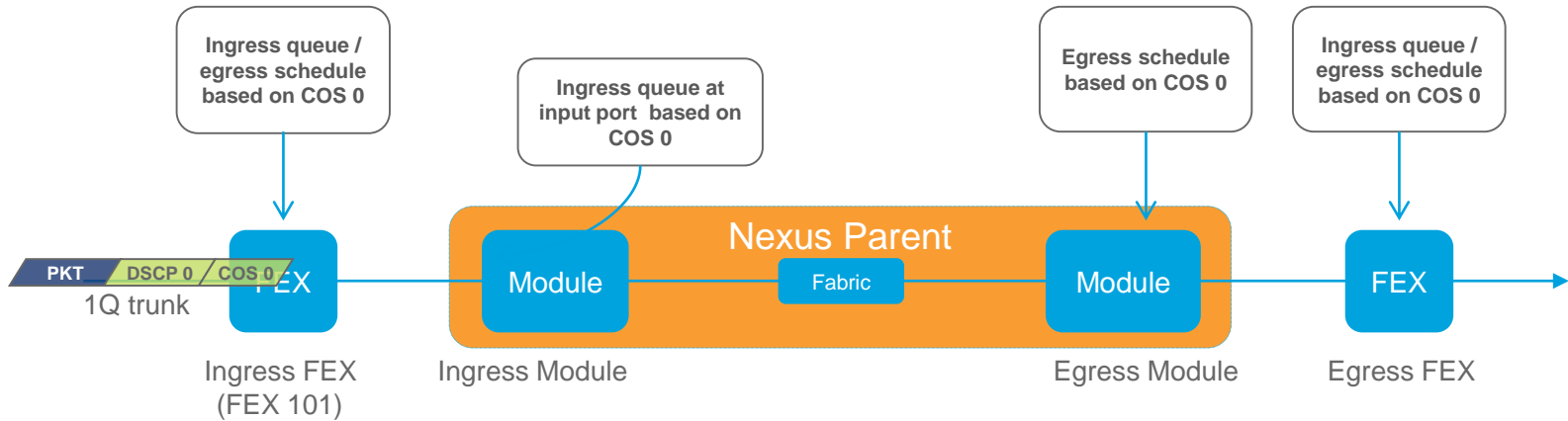
- **Traffic classification**
 - DSCP, CoS
 - ACL classification (FEX offload) –only on 5600/6000
- **Strict Priority Queuing and DWRR**
 - Priority Flow Control
- **Queue-limit Carving**

FEX QoS Policies

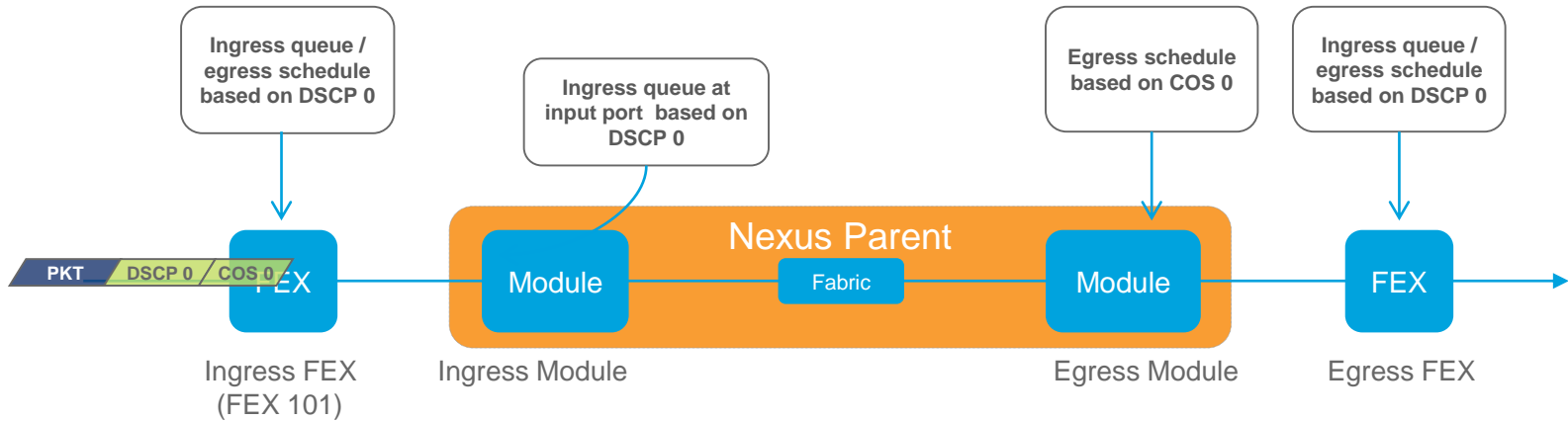
- Support for ingress port-based QoS policies on FEX HIF ports
- Support for ingress/egress VLAN-based QoS policies on FEX VLANs
- FEX QoS policies applied at ingress module of parent switch
 - No support for remarking, policing policies



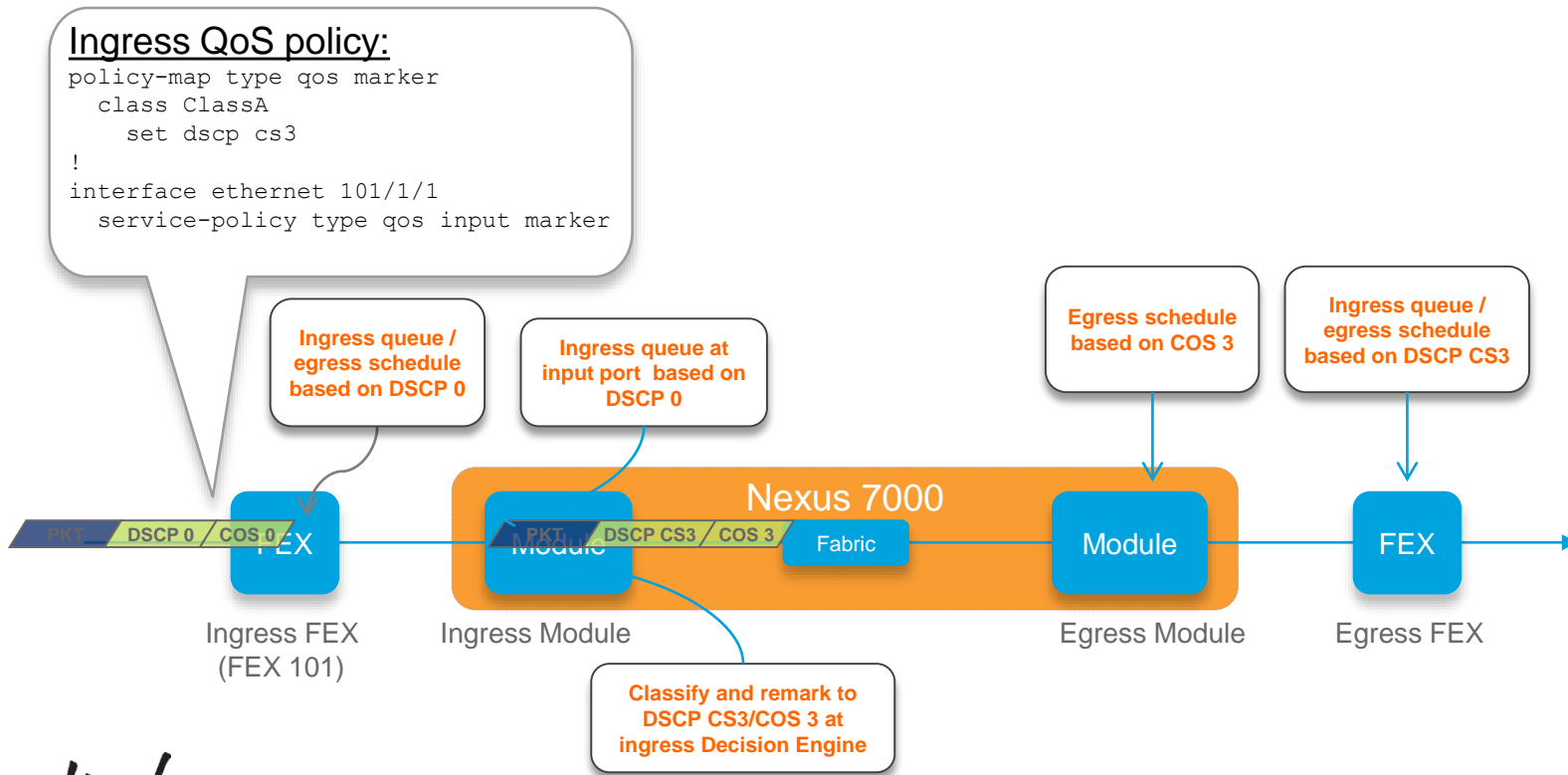
FEX QoS Packet Flow Example (CoS2Q)



FEX QoS Packet Flow Example (DSCP2Q)



FEX QoS Packet Flow Example (With Ingress Marking Policy and DSCP-to-Queue)



FEX Policy Offload (Nexus 5600/6000 only)

- TCAM resources on a FEX to perform ACL-based classification
- The feature is disabled by default
- By default, a FEX classifies packets on CoS value
- Both system level and interface level policies are offloaded to the FEX

```
switch# configure terminal
switch(config)# fex chassis_ID
switch(config)# hardware card-type qos-policy-offload
```

FEX Queuing Policies

- FEX queuing driven implicitly by parent switch queuing configuration
- Network QoS template drives:
 - Number of queues
 - MTU
- Ingress queuing class-maps drive:
 - **Both ingress and egress** COS/DSCP-to-queue mapping
- On Nexus 7000 with FEX + M-Series parent modules, network-qos and F-series ingress queuing class-maps still drive FEX queuing configuration

DSCP-to-Queue on FEX

- Enabling DSCP-to-queue on parent switch enables DSCP-to-queue on FEX
- Currently active DSCP mappings pushed to FEX when enabled
- DSCP-to-queue only active in the HIF→NIF direction
 - NIF→HIF direction always uses COS-to-queue mapping, based on COS transmitted by parent switch to FEX

FEX Queue-Limit

- Provides FEX queue-limit configuration option
- Manages buffer thresholds on FEX based on platform capabilities
- **Default has queue-limit enabled**, disabling not recommended
- Configuration applied **per-VDC** (on Nexus 7000/7700)
- Different FEX models have different capabilities

Nexus 2000 QoS Golden Rules

- FEX QoS classification on COS or DSCP unless FEX offload enabled
- FEX queuing driven implicitly by parent switch queuing configuration
- No support for per-queue shaping, policing or marking
- Drop thresholds are tail-drop only, no WRED support

Agenda

- Introduction
- QoS and Queuing Basics
- QoS Implementation on Nexus
- Nexus 7000/7700 QoS
- Nexus 5x00/6000 QoS
- Nexus 2000 QoS
- Real World Deployment
- Conclusion

Real World Deployment

What do we want to achieve?

Company XYZ's Business Intent

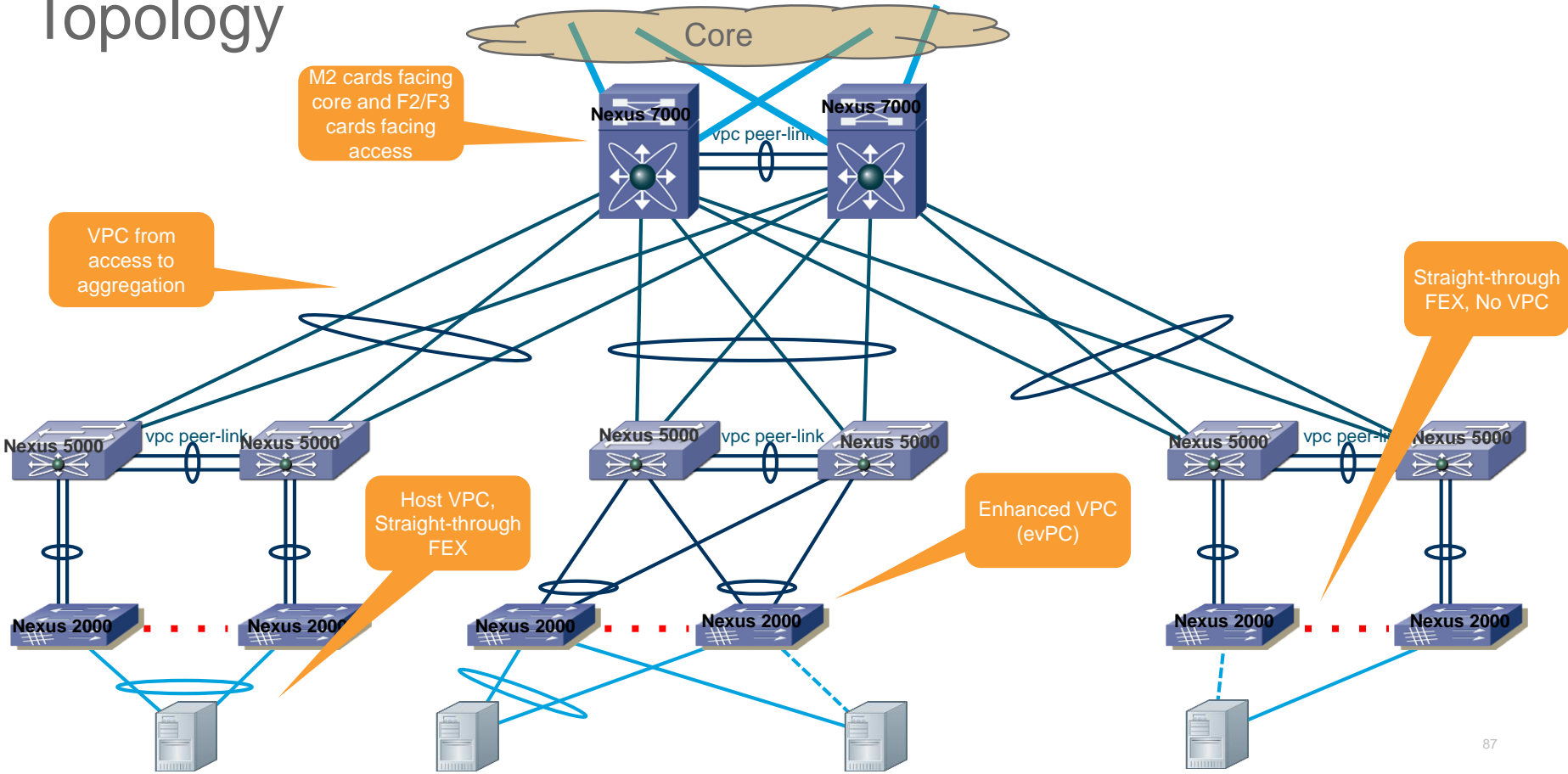
- Make sure no disruption in network services
 - *Put control traffic in priority queue*
- Video/voice hosting also an business objective
 - *Put voice traffic in priority queue*
 - *Dedicated bandwidth to video traffic*
- Flexibility in moving applications across servers
 - *Dedicated bandwidth to vmotion/mobility*
 - *Everything else best-effort*



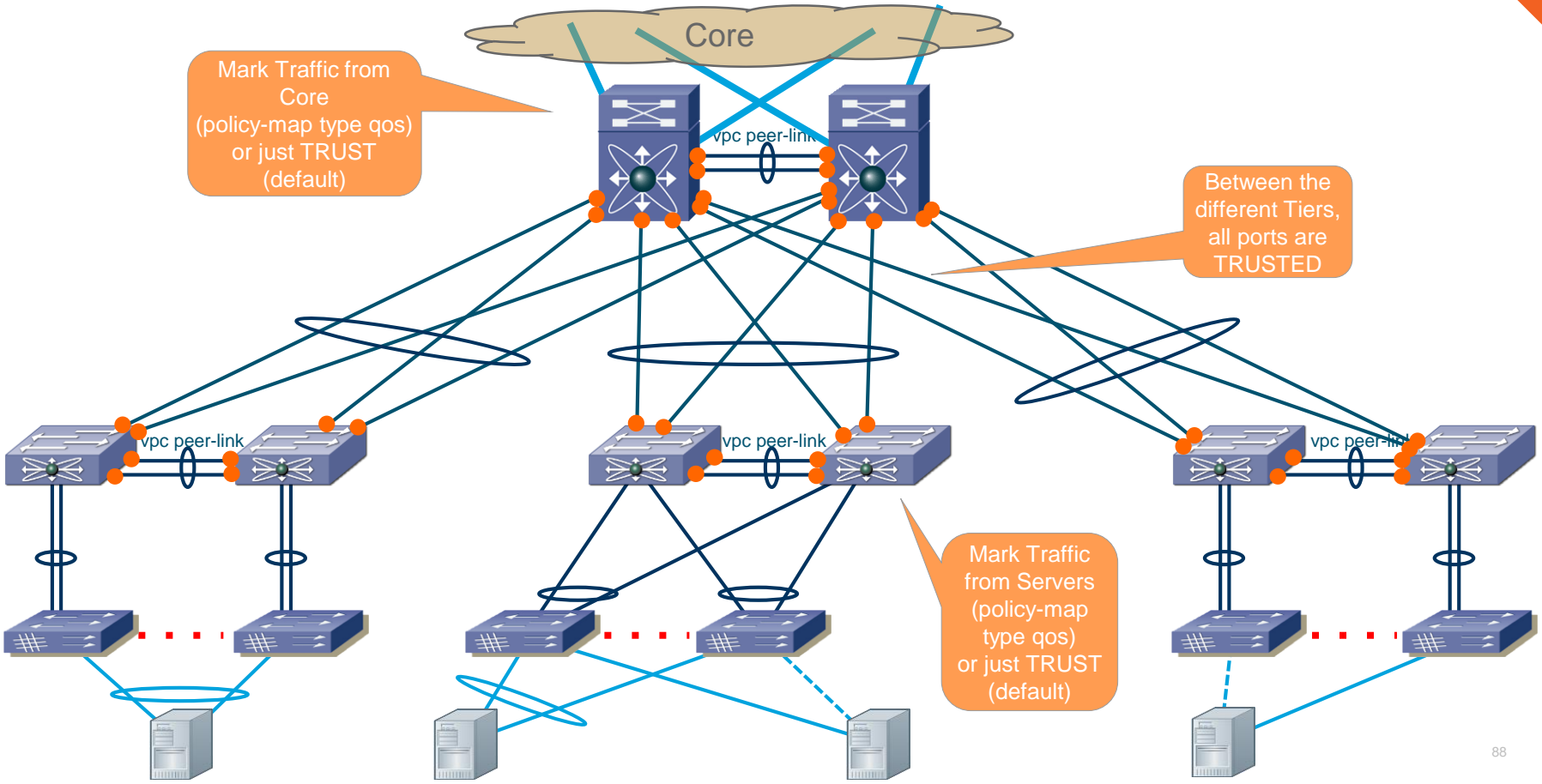
Translating to the language of QoS

Application	CoS	Queuing (Scheduling)	Queue-Limit (Buffer)	Character
Best Effort	0, 1	BW remaining 50%	60%	High Volume / Less Important
vMotion / Live Migration	2	BW remaining 20%	10%	Medium Volume / Important
Multimedia	3, 4	BW remaining 30%	20%	Medium Volume Very Important
Strict Priority	5	Priority Queue	10%	Low Volume / Important / Delay Sensitive
Network Control	6,7			Low Volume / Very important

Topology



Classification, Marking & Trust on Nexus 5000/7000



Classification & Marking: Nexus 7000

```
ip access-list ACL_QOS_LOWPRIO
 10 permit ...
ip access-list ACL_QOS_VMOTION
 10 permit ...
ip access-list ACL_QOS_MULTIMEDIA
 10 permit ...
ip access-list ACL_QOS_SCRIPTPRIO
 10 permit ...
!
class-map type qos match-any CM_QOS_LOWPRIO_COS1
 match access-group name ACL_QOS_LOWPRIO
!
class-map type qos match-any CM_QOS_VMOTION_COS2
 match access-group name ACL_QOS_VMOTION
!
class-map type qos match-any CM_QOS_MULTIMEDIA_COS4
 match access-group name ACL_QOS_MULTIMEDIA
!
class-map type qos match-any CM_QOS_STRICTPRIO_COS5
 match access-group name ACL_QOS_STRICTPRIO
```

```
policy-map type qos PM_QOS_MARK_COS_IN
 class CM_QOS_STRICTPRIO_COS5
   set cos 5
 class CM_QOS_MULTIMEDIA_COS4
   set cos 4
 class CM_QOS_VMOTION_COS2
   set cos 2
 class CM_QOS_LOWPRIO_COS1
   set cos 1
!
interface Ethernet1/1
 service-policy type qos input PM_QOS_MARK_COS_IN
!
vlan configuration 100
 service-policy input PM_QOS_MARK_COS_IN
```

Classification & Marking: Nexus 5600/6000 (1)

```
ip access-list ACL_QOS_LOWPRIO
 10 permit ...
ip access-list ACL_QOS_VMOTION
 10 permit ...
ip access-list ACL_QOS_MULTIMEDIA
 10 permit ...
!
class-map type qos match-any CM_QOS_LOWPRIO_COS1
 match access-group name ACL_QOS_LOWPRIO
!
class-map type qos match-any CM_QOS_VMOTION_COS2
 match access-group name ACL_QOS_VMOTION
!
class-map type qos match-any CM_QOS_MULTIMEDIA_COS4
 match access-group name ACL_QOS_MULTIMEDIA
!
class-map type qos match-any CM_QOS_STRICTPRIO_COS5
 match cos 5
```

```
policy-map type qos PM_QOS_MARK_COS_IN
 class CM_QOS_STRICTPRIO_COS5
  set qos-group 5
 class CM_QOS_MULTIMEDIA_COS4
  set qos-group 4
 class CM_QOS_VMOTION_COS2
  set qos-group 3
 class CM_QOS_LOWPRIO_COS1
  set qos-group 2
!
system qos
 service-policy type qos input PM_QOS_MARK_COS_IN
```

QoS-Group # is mapping between Slide 1 & Slide 2

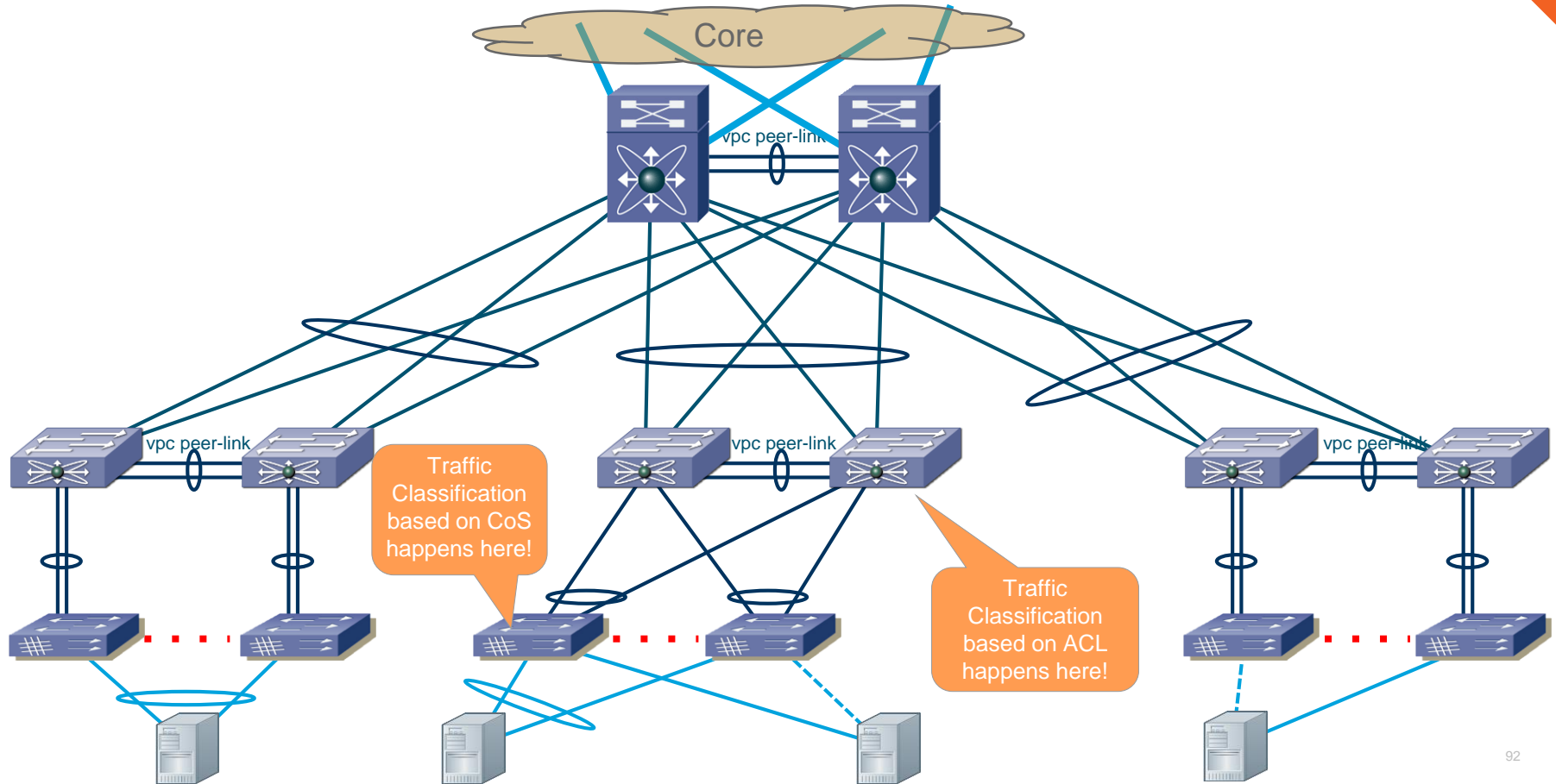
Classification & Marking: Nexus 5500/6000 (2)

```
class-map type network-qos CM_N-QOS_MATCH_QG2_COS1
  match qos-group 2
class-map type network-qos CM_N-QOS_MATCH_QG3_COS2
  match qos-group 3
class-map type network-qos CM_N-QOS_MATCH_QG4_COS4
  match qos-group 4
class-map type network-qos CM_N-QOS_MATCH_QG5_COS5
  match qos-group 5
```

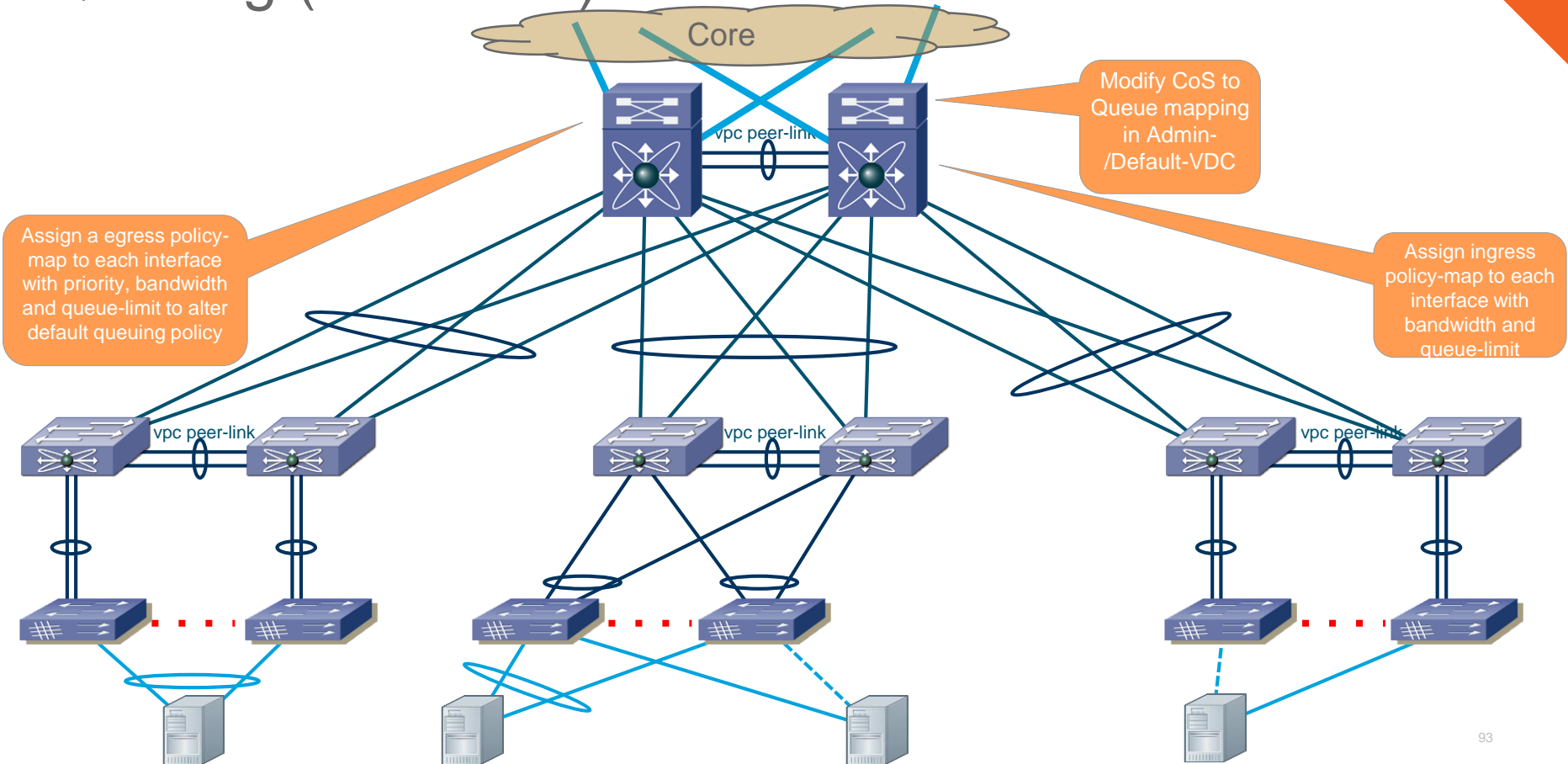
```
policy-map type network-qos PM_N-QOS_SYSTEM
  class type network-qos CM_N-QOS_MATCH_QG2_COS1
    set cos 1
  class type network-qos CM_N-QOS_MATCH_QG3_COS2
    set cos 2
  class type network-qos CM_N-QOS_MATCH_QG4_COS4
    set cos 4
  class type network-qos CM_N-QOS_MATCH_QG5_COS5
    set cos 5
    queue-limit 20480 bytes
!
system qos
  service-policy type network-qos PM_N-QOS_SYSTEM
```

QoS-Group # is mapping between Slide 1 & Slide 2

Classification & Marking (Nexus 2000)



Queuing (M2-cards)



CoS to Queue Mapping – M2 I/O Module

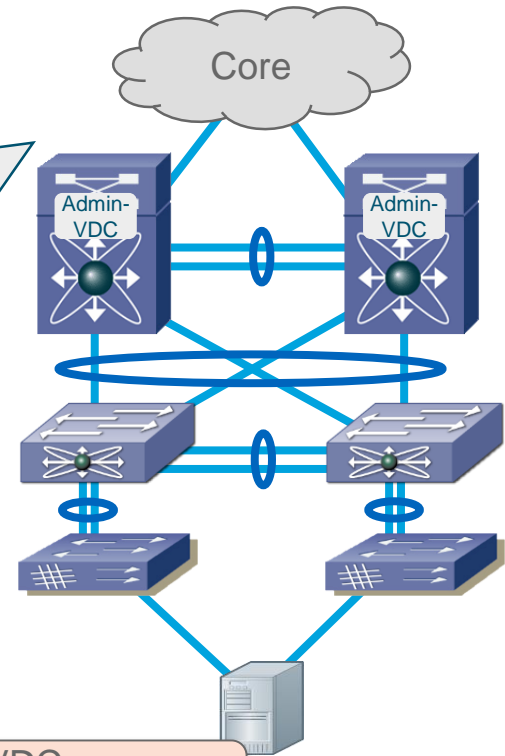
Example

Application	CoS	Queuing (Scheduling)	Queue-Limit (Buffer)	Queue (8q2t / 1p7q4t)	Character
Best Effort	0, 1	BW remaining 50%	60%	8q2t-in-q-default / 1p7q4t-out-q-default	High Volume / Less Important
vMotion / Live Migration	2	BW remaining 20%	10%	8q2t-in-q6 / 1p7q4t-out-q6	Medium Volume / Important
Multimedia	3, 4	BW remaining 30%	20%	8q2t-in-q2 / 1p7q4t-out-q2	Medium Volume Very Important
Strict Priority	5	Priority Queue	10%	8q2t-in-q1 / 1p7q4t-out-pq1	Low Volume / Important / Delay Sensitive
Network Control	6,7				Low Volume / Very important

CoS to Queue Mapping (on M2 cards)

Exmple (Admin- / Default-VDC)

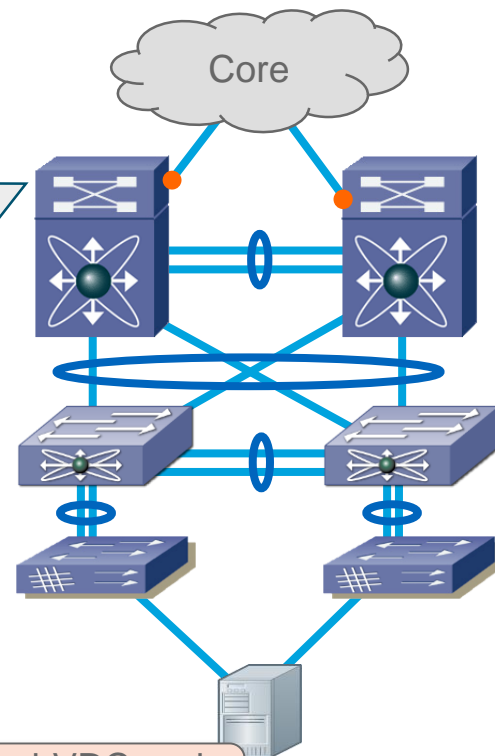
```
class-map type queuing match-any 8q2t-in-q1
  match cos 5-7
class-map type queuing match-any 8q2t-in-q2
  match cos 3-4
class-map type queuing match-any 8q2t-in-q6
  match cos 2
class-map type queuing match-any 8q2t-in-q-default
  match cos 0-1
!
class-map type queuing match-any 1p7q4t-out-pq1
  match cos 5-7
class-map type queuing match-any 1p7q4t-out-q2
  match cos 3-4
class-map type queuing match-any 1p7q4t-out-q6
  match cos 2
class-map type queuing match-any 1p7q4t-out-q-default
  match cos 0-1
```



Ingress Queuing Configuration on M2 cards

Example (Payload-VDC)

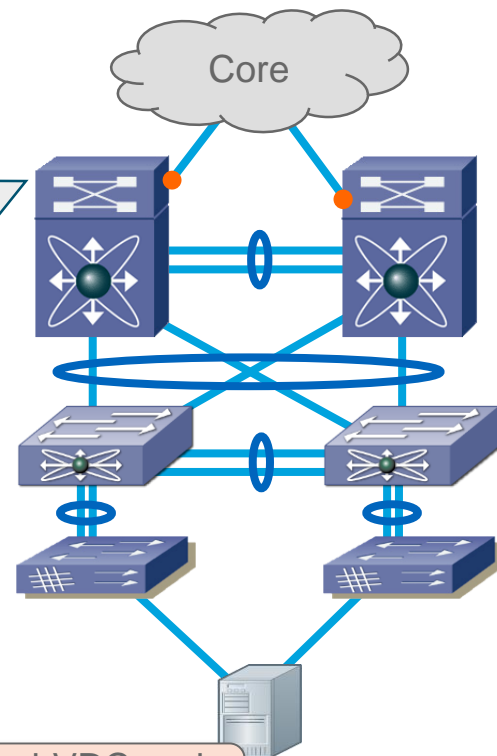
```
policy-map type queuing PM_QUEUE_10G-40G-100G_IN
  class type queuing 8q2t-in-q1
    queue-limit 10
    bandwidth remaining percent 10
  class type queuing 8q2t-in-q2
    queue-limit 20
    bandwidth remaining percent 30
  class type queuing 8q2t-in-q6
    queue-limit 10
    bandwidth remaining percent 10
  class type queuing 8q2t-in-q-default
    queue-limit percent 50
    bandwidth percent 50
!
interface Ethernet1/1
  service-policy type queuing input PM_QUEUE_10G-40G-100G_IN
```



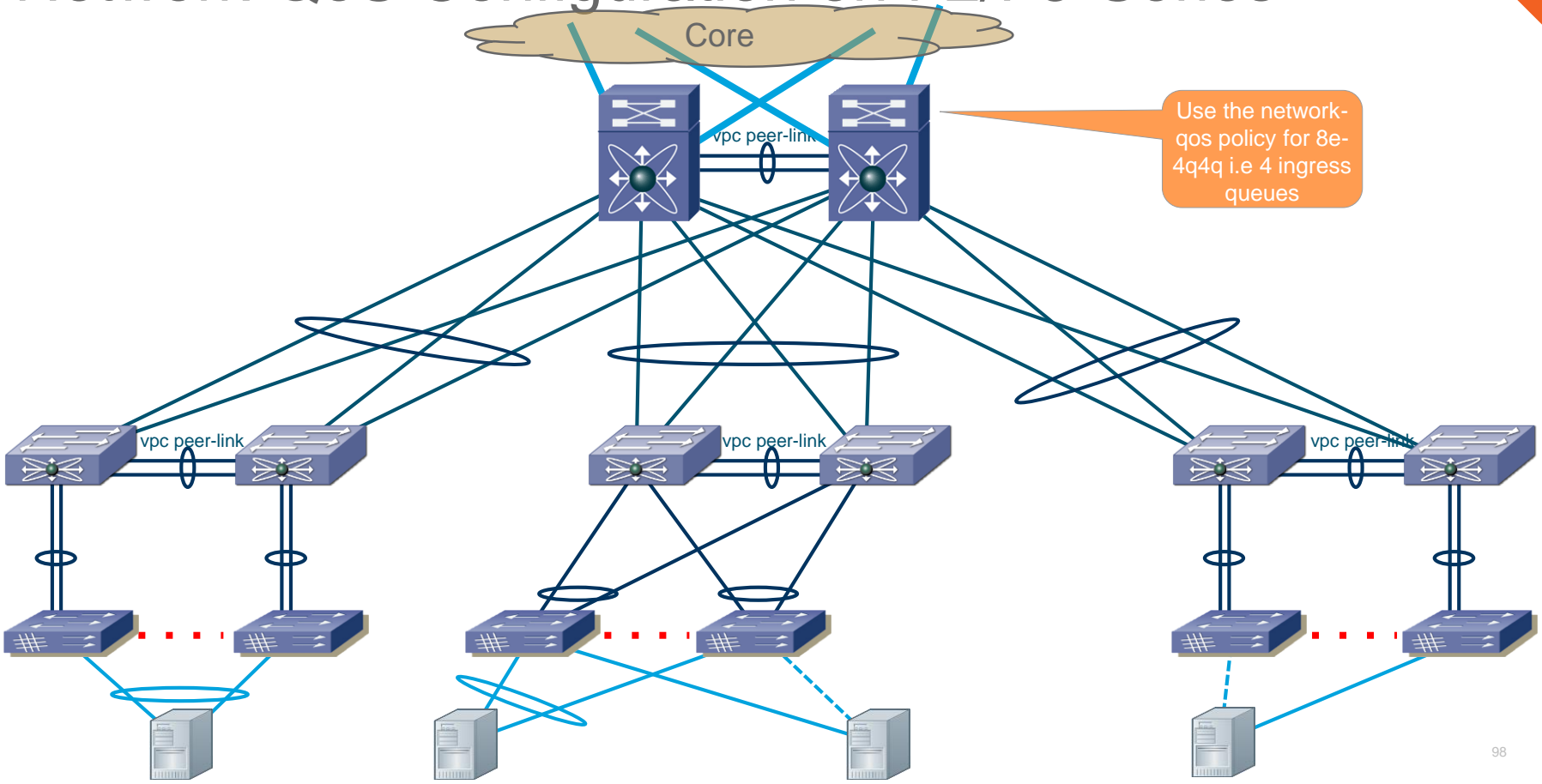
Egress Queuing Configuration on M2 cards

Example (Payload-VDC)

```
policy-map type queuing PM_QUEUE_10G-40G-100G_OUT
  class type queuing lp7q4t-out-pq1
    priority level 1
    queue-limit percent 10
  class type queuing lp7q4t-out-q2
    queue-limit percent 20
    bandwidth remaining percent 30
  class type queuing lp7q4t-out-q6
    queue-limit percent 10
    bandwidth remaining percent 20
  class type queuing lp7q4t-out-q-default
    queue-limit percent 50
    bandwidth remaining percent 40
!
interface Ethernet1/1
  service-policy type queuing output PM_QUEUE_10G-40G-100G_OUT
```



Network-QoS Configuration on F2/F3-Series

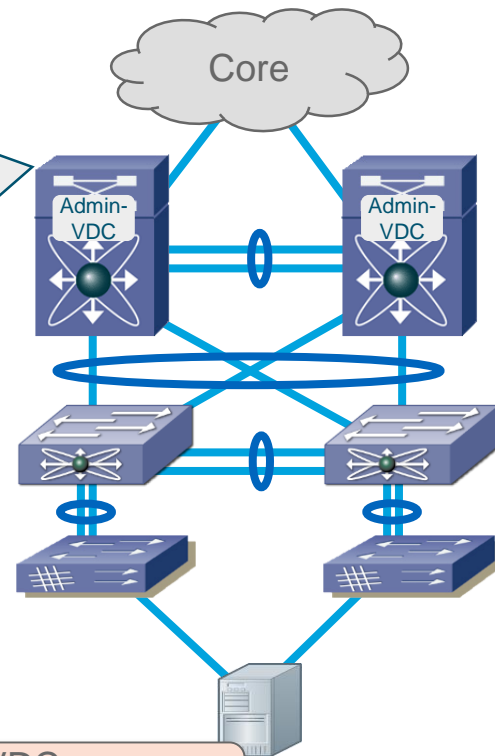


Network-QoS Configuration –F2/F3 cards

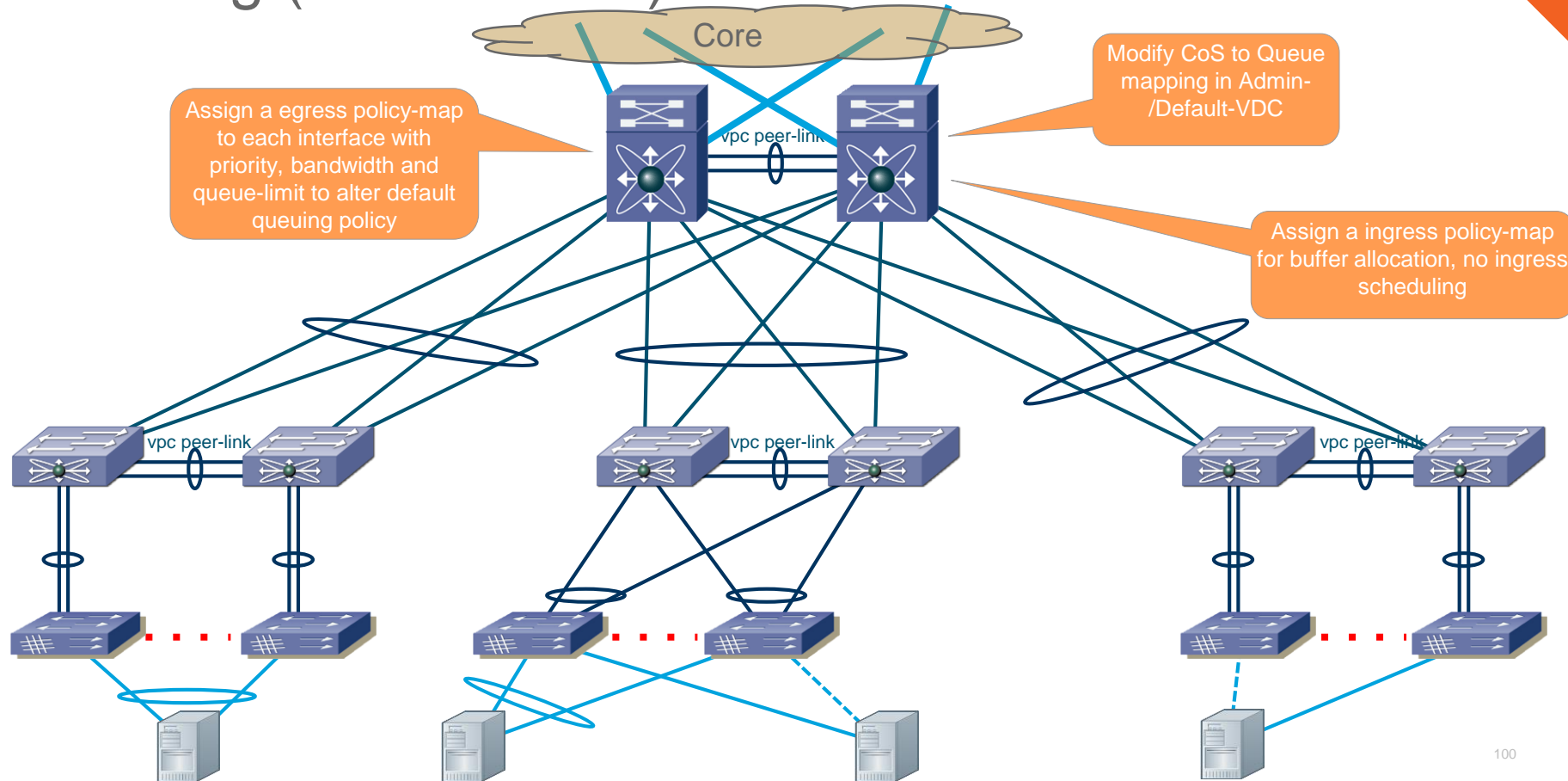
Example (Admin- / Default-VDC)

```
system qos
  service-policy type network-qos default-nq-8e-4q4q-policy

policy-map type network-qos default-nq-8e-4q4q-policy template 8e-4q4q
  class type network-qos c-nq-8e-4q4q
    match cos 0-7
    congestion-control tail-drop
    mtu 1500
```



Queuing (F2/F3 cards)



CoS to Queue Mapping – F2/F3 I/O Module

Example

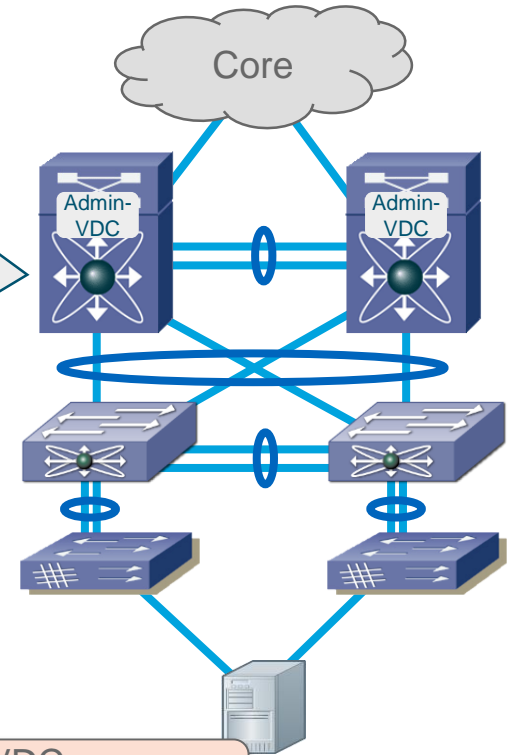
Application	CoS	Queuing (Scheduling)-egress	Queue-Limit (Buffer)-ingress	Queue (Ingress/Egress)	Character
Best Effort	0,1	BW remaining 50%	50%	4q1t-8e-4q4q-in-q-default / 1p3q1t-8e-4q4q-out-q-default	High Volume / Less Important
vMotion / Live Migration	2	BW remaining 20%	10%	4q1t-8e-4q4q-in-q4 / 1p3q1t-8e-4q4q-out-q3	Medium Volume / Important
Multimedia	3, 4	BW remaining 30%	30%	4q1t-8e-4q4q-in-q3 / 1p3q1t-8e-4q4q-out-q2	Medium Volume Very Important
Strict Priority	5	Priority Queue	10%	4q1t-8e-4q4q-in-q1 / 1p3q1t-8e-4q4q-out-pq1	Low Volume / Important / Delay Sensitive
Network Control	6/7				Low Volume / Very important

CoS to Queue Configuration –F2/F3 slides

Example (Admin- / Default-VDC)

```
class-map type queuing match-any 4q1t-8e-4q4q-in-q1
  match cos 5-7
class-map type queuing match-any 4q1t-8e-4q4q-in-q-default
  match cos 0-1
class-map type queuing match-any 4q1t-8e-4q4q-in-q3
  match cos 3-4
class-map type queuing match-any 4q1t-8e-4q4q-in-q4
  match cos 2

class-map type queuing match-any 1p3q1t-8e-4q4q-out-pq1
  match cos 5-7
class-map type queuing match-any 1p3q1t-8e-4q4q-out-q2
  match cos 3-4
class-map type queuing match-any 1p3q1t-8e-4q4q-out-q3
  match cos 2
class-map type queuing match-any 1p3q1t-8e-4q4q-out-q-default
  match cos 0-1
```



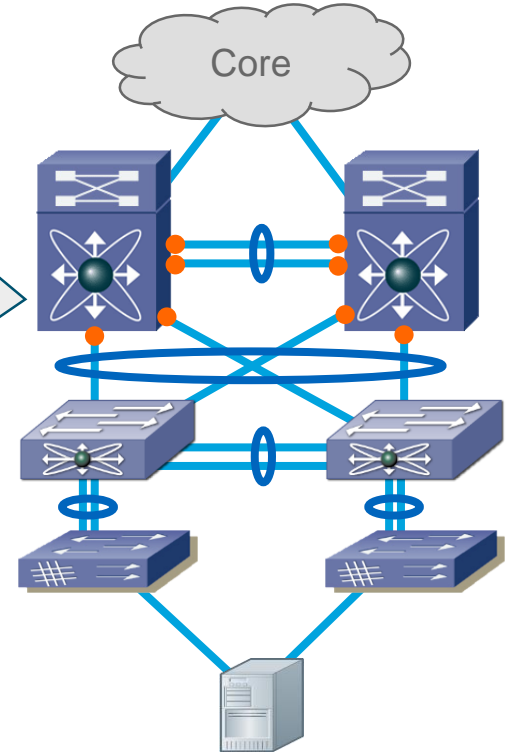
Ingress Queuing Configuration for F2/F3 cards

Example (Payload-VDC)

```
qos copy policy-map type queuing default-8e-4q4q-in-policy prefix  
Custom-
```

```
policy-map type queuing Custom-8e-4q4q-in  
  class type queuing 4q1t-8e-4q4q-in-q1  
    queue-limit percent 10  
    bandwidth percent 25  
  class type queuing 4q1t-8e-4q4q-in-q-default  
    queue-limit percent 50  
    bandwidth percent 25  
  class type queuing 4q1t-8e-4q4q-in-q3  
    queue-limit percent 30  
    bandwidth percent 25  
  class type queuing 4q1t-8e-4q4q-in-q4  
    queue-limit percent 10  
    bandwidth percent 25
```

```
interface Ethernet1/1  
  service-policy type queuing input Custom-8e-4q4q-in
```



All Policy-Map and Service-Policy are done in relevant Payload-VDC and only affect the interface to which they get applied

Egress Queuing Configuration for F2/F3 cards

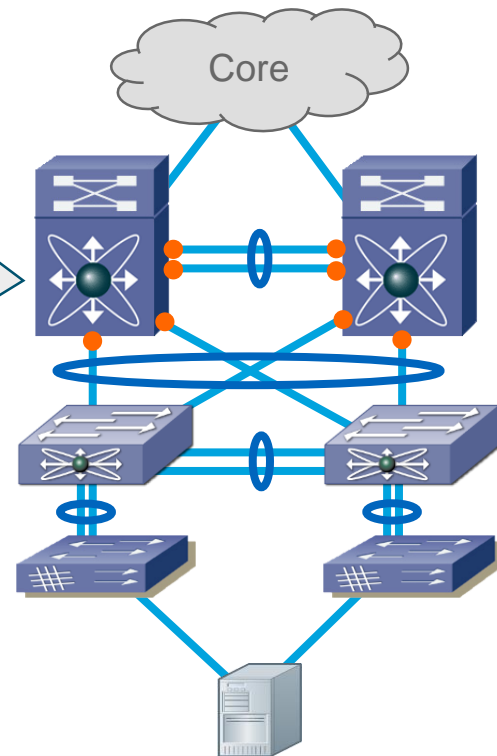
Example (Payload-VDC)

```
qos copy policy-map type queuing default-8e-4q4q-out-policy prefix  
Custom-
```

```
policy-map type queuing Custom-8e-4q4q-out  
  class type queuing lp3q1t-8e-4q4q-out-pq1  
    priority level 1  
  class type queuing lp3q1t-8e-4q4q-out-q2  
    bandwidth remaining percent 30  
  class type queuing lp3q1t-8e-4q4q-out-q3  
    bandwidth remaining percent 20  
  class type queuing lp3q1t-8e-4q4q-out-q-default  
    bandwidth remaining percent 50
```

```
!
```

```
interface Ethernet1/1  
  service-policy type queuing output Custom-8e-4q4q-out
```



All Policy-Map and Service-Policy are done in relevant Payload-VDC and only affect the interface to which they get applied

CoS to Queue Mapping - Nexus 5600/6000

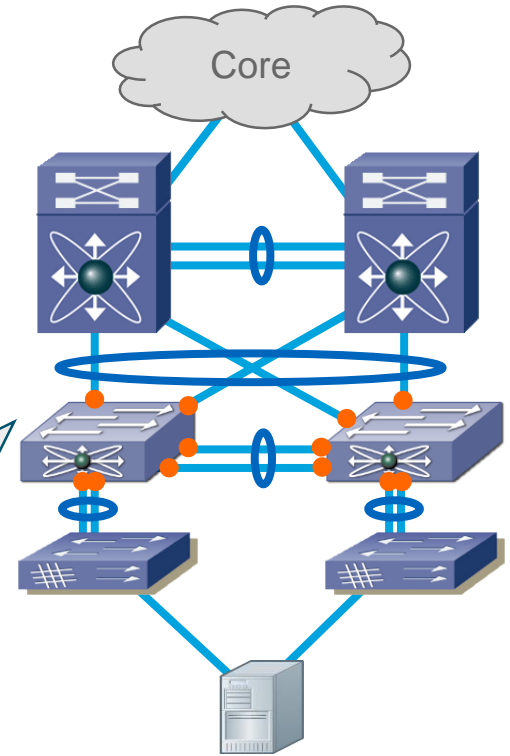
Example

Application	CoS	Queuing (Scheduling)	Queue-Limit (Buffer)	Queue (6q1t / 1p6q0t)	Character
Best Effort	0,1	BW percent 40%	remaining (226kByte)	qos-group 0 (default)	High Volume / Less Important
vMotion / Live Migration	2,3	BW percent 20%	41kByte	qos-group 3	Medium Volume / Important
Multimedia	4	BW percent 30%	41kByte	qos-group 4	Medium Volume Very Important
Strict Priority	5	BW percent 10%	20kByte (min)	qos-group5 / priority	Low Volume / Important / Delay Sensitive
Network Control	6,7				Low Volume / Very important

Egress Queuing Configuration: Nexus5600

Example

```
class-map type queuing CM_Q_MATCH_QG3_COS2
  match qos-group 3
class-map type queuing CM_Q_MATCH_QG4_COS4
  match qos-group 4
class-map type queuing CM_Q_MATCH_QG5_COS5
  match qos-group 5
!
policy-map type queuing PM_QUEUING_SYSTEM_OUT
class type queuing CM_Q_MATCH_QG3_COS2
  bandwidth percent 20
class type queuing CM_Q_MATCH_QG4_COS4
  bandwidth percent 30
class type queuing CM_Q_MATCH_QG5_COS5
  priority
  bandwidth percent 10
class type queuing class-default
  bandwidth percent 40
```

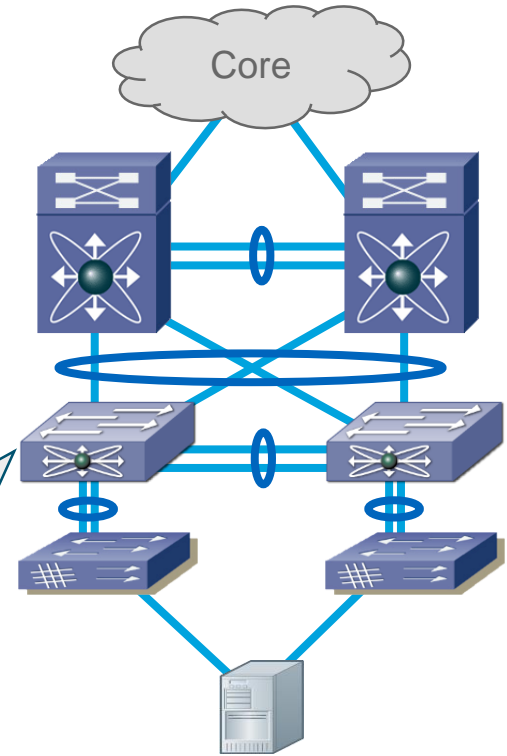


Queue-Limit (Buffer) Configuration

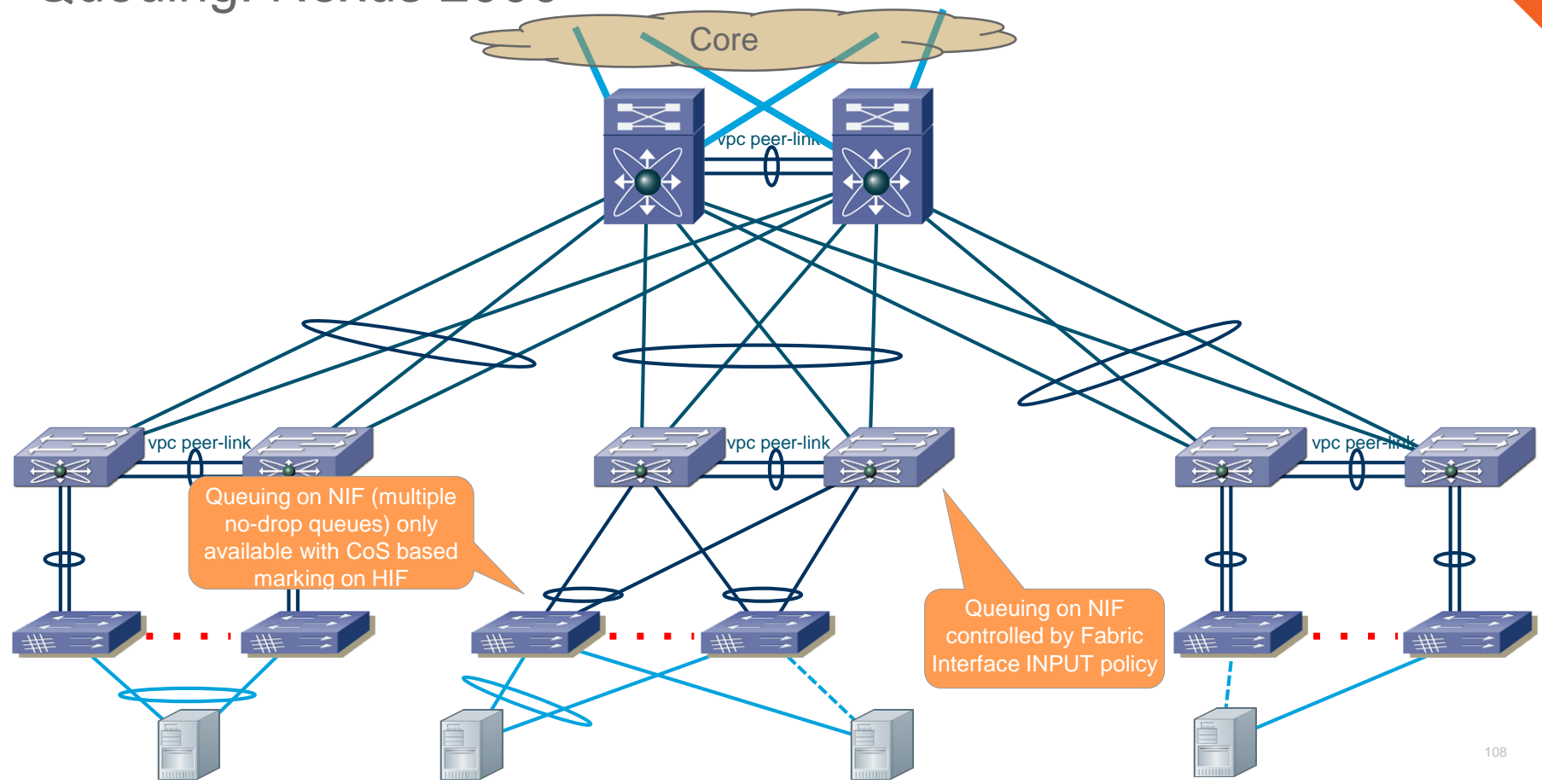
Example

```
policy-map type network-qos PM_N-QOS_SYSTEM
  class type network-qos CM_N-QOS_MATCH_QG3_COS2
    set cos 2
    queue-limit 40960 bytes
  class type network-qos CM_N-QOS_MATCH_QG4_COS4
    set cos 2
    queue-limit 40960 bytes
  class type network-qos CM_N-QOS_MATCH_QG5_COS5
    set cos 5
    queue-limit 20480 bytes
  class type network-qos class-default
```

```
system qos
  service-policy type network-qos PM_N-QOS_SYSTEM
```



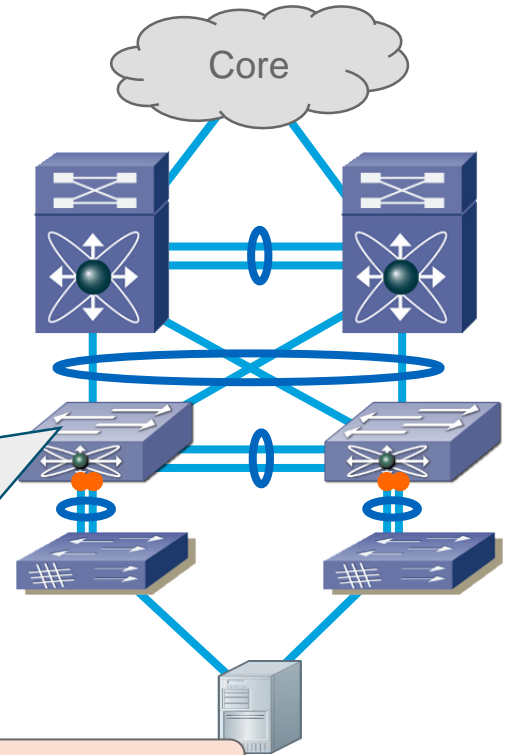
Queuing: Nexus 2000



Queuing Configuration (Nexus 2000)

Example

```
class-map type queuing CM_Q_MATCH_QG3_COS2
  match qos-group 3
class-map type queuing CM_Q_MATCH_QG4_COS4
  match qos-group 4
class-map type queuing CM_Q_MATCH_QG5_COS5
  match qos-group 5
!
policy-map type queuing PM_QUEUING_SYSTEM_N2K
class type queuing CM_Q_MATCH_QG3_COS2
  bandwidth percent 20
class type queuing CM_Q_MATCH_QG4_COS4
  bandwidth percent 30
class type queuing CM_Q_MATCH_QG5_COS5
  priority
  bandwidth percent 10
class type queuing class-default
  bandwidth percent 40
```



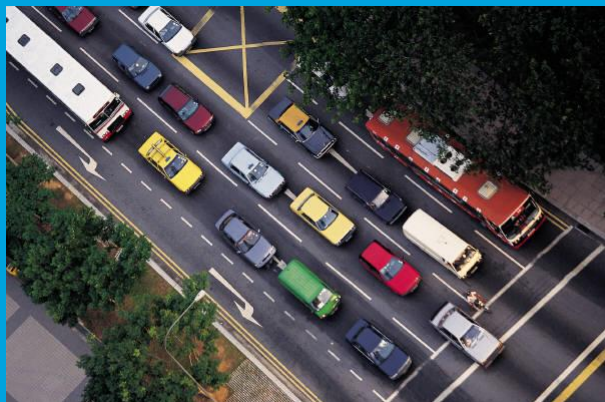
Agenda

- Introduction
- QoS and Queuing Basics
- QoS Implementation on Nexus
- Nexus 7000/7700 QoS
- Nexus 5x00/6000 QoS
- Nexus 2000 QoS
- Real World Deployment
- **Conclusion**

Conclusion

Why QoS in the Data Center?

**Assign
Color to Traffic**



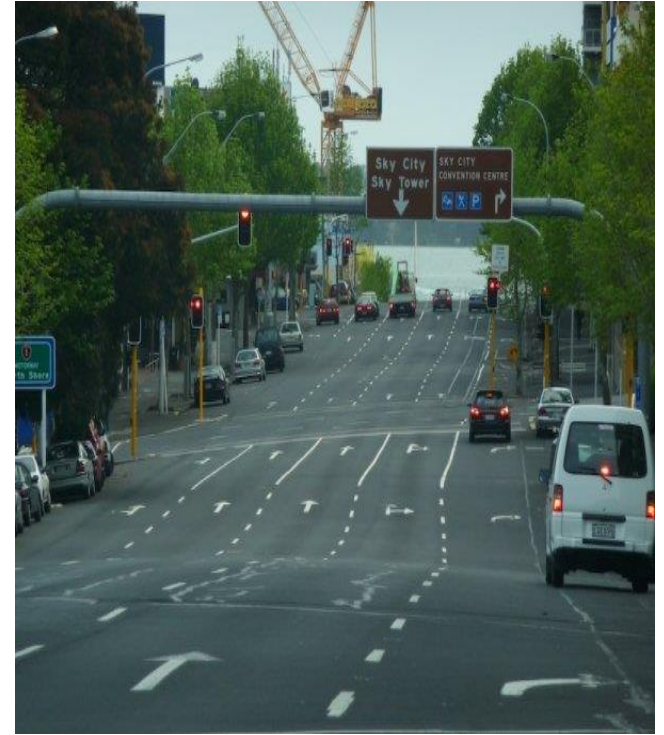
**Manage
Congestion**



**Maximize
Throughput**



Maximize Throughput and Manage Congestion!

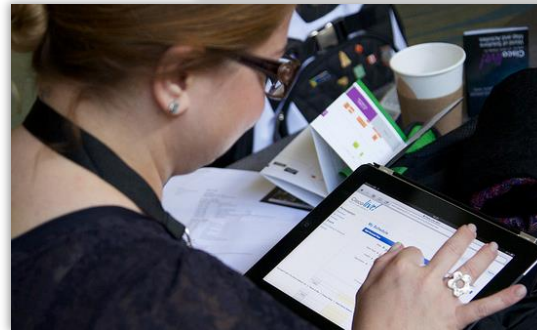


Call to Action

- Visit the World of Solutions for
 - Cisco Campus
 - Walk in Labs
 - Technical Solution Clinics
- Meet the Engineer
- Lunch and Learn Topics
- DevNet zone related sessions

Complete Your Online Session Evaluation

- Please complete your online session evaluations after each session. Complete 4 session evaluations & the Overall Conference Evaluation (available from Thursday) to receive your Cisco Live T-shirt.
- All surveys can be completed via the Cisco Live Mobile App or the Communication Stations



Thank you



We're ready. Are you?

Bonus Slides

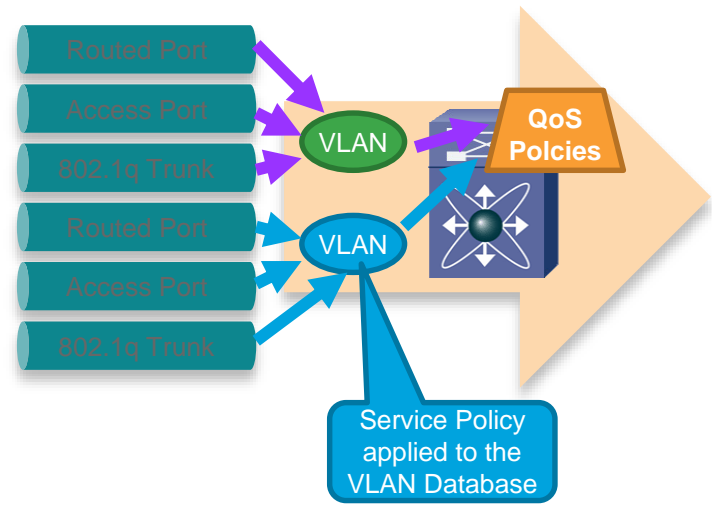
8 Class Queuing Model

Class	DSCP	Queues
Priority	CS6 (CS7)	PQ
Platinum	EF	
Gold	AF41	Q7
Silver	CS4	Q6
No-Drop	CoS3	Q5
Bronze	AF21	Q4
Management	CS2	Q3
Scavenger	AF11	Q2
Bulk Data	CS1	Q1
Best-Effort	0	Default-Q

- Matches a Campus QoS concept (RFC 4594)
- No-Drop still with CoS3 (**DSCP 24-30 are “unusable”**)
- Valid but **most complex**

VLAN based QoS Policy attachment

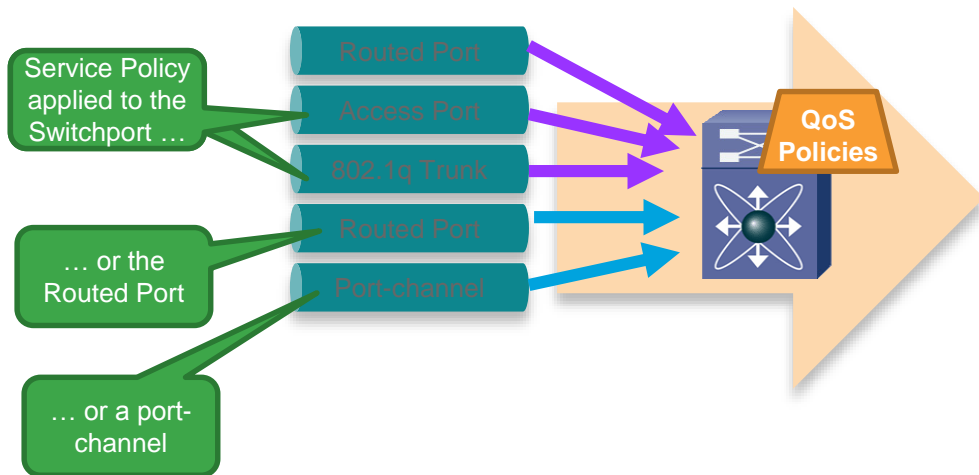
- VLAN based QoS Policy is configured in VLAN Database
- No SVI (aka L3 VLAN Interface) required



```
Nexus(config)# vlan configuration <vlan-id>  
Nexus(config-vlan)# service-policy input myPolicy
```

Interface based QoS Policy attachment

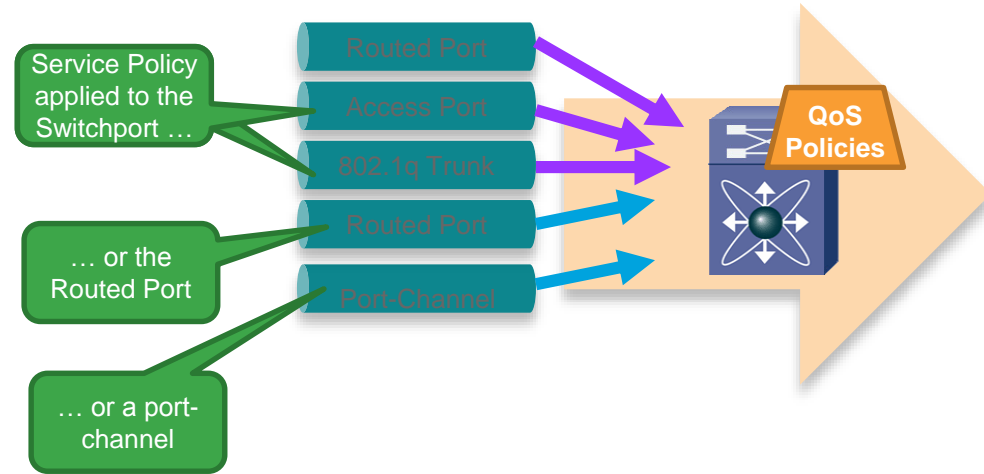
- Interface based QoS Policy takes precedence over VLAN
- Can also be attached to port-channel and applies to all member-ports
- No Egress QoS policies on L2 ports!



```
Nexus(config)# interface ethernet 1/1
Nexus(config-if)# service-policy input myPolicy
```

Interface based Queuing Policy attachment

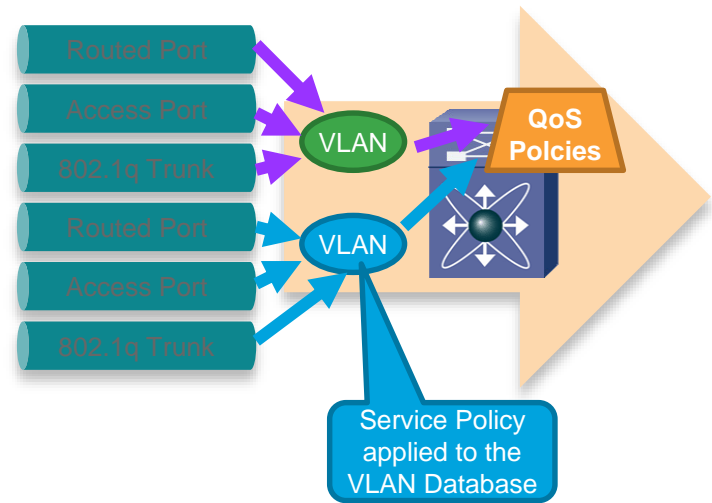
- Interface based QoS Policy takes precedence over VLAN
- Interface based QoS Policy is configured under the respective Interface
- Queuing Policy can be attached to port-channel also



```
Nexus(config)# interface ethernet 1/1
Nexus(config-if)# service-policy input myPolicy
```

VLAN based QoS Policy attachment

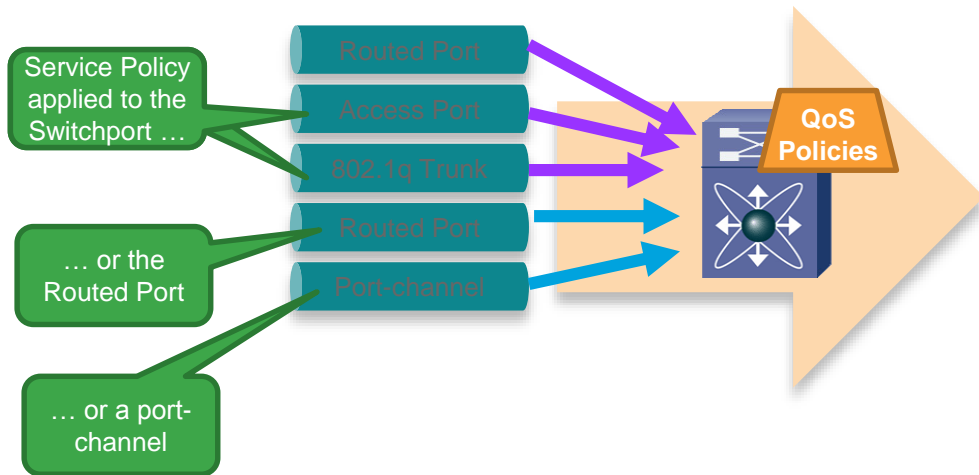
- VLAN based QoS Policy is configured in VLAN Database
- No SVI (aka L3 VLAN Interface) required



```
Nexus(config)# vlan configuration <vlan-id>  
Nexus(config-vlan)# service-policy input myPolicy
```

Interface based QoS Policy attachment

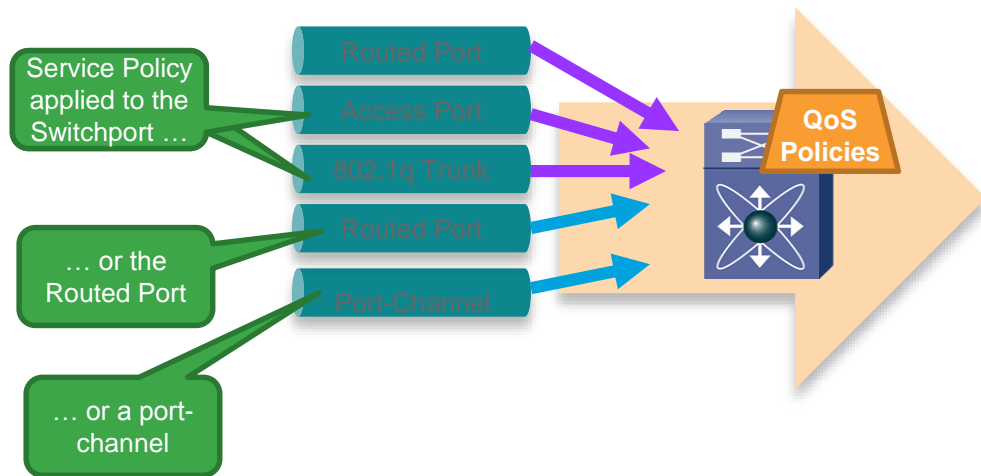
- Interface based QoS Policy takes precedence over VLAN
- Can also be attached to port-channel and applies to all member-ports
- No Egress QoS policies on L2 ports!



```
Nexus(config)# interface ethernet 1/1
Nexus(config-if)# service-policy input myPolicy
```


Interface based Queuing Policy attachment

- Interface based QoS Policy takes precedence over VLAN
- Interface based QoS Policy is configured under the respective Interface
- Queuing Policy can be attached to port-channel also



```
Nexus(config)# interface ethernet 1/1
Nexus(config-if)# service-policy input myPolicy
```

M2 I/O Module Buffering Capacity

Ingress

M2 Module	Ingress Queue Structure	Ingress Port Buffer	Ingress VOQ Buffer
24-port 10G N7K-M224XP-23L	8q2t	5.2MB / port	4.5MB / port
6-port 40G N7K-M206FQ-23L	8q2t	20.8MB / port	18MB / port
2-port 100G N7K-M202CF-22L	8q2t	62.8MB / port	54MB / port

Egress

M2 Module	VOQ Structure	Egress VOQ Buffer (Credited)	Egress VOQ Buffer (Uncredited)	Egress Queue Structure	Egress Port Buffer
24-port 10G N7K-M224XP-23L	1p4q	295KB / port	512KB / 6 ports	1p7q4t	5MB / port
6-port 40G N7K-M206FQ-23L	1p4q	1.2MB / port	1MB / 3 ports	1p7q4t	20.7MB / port
2-port 100G N7K-M202CF-22L	1p4q	3MB / port	1MB / port	1p7q4t	30.2MB / port

FYI

F I/O Module Buffering Capacity

Ingress

Module	Total VOQ Buffer Per Module	Ingress Queue Structure	Ingress VOQ Buffer
48-port 10G F2E N77-F248XP-23E	72MB	4q1t	1.5MB / port
F3 48-port 10G N77-F348XP-23	72MB	4q1t	1.5MB / port
F3 24-port 40G N77-F324FQ-25	144MB	4q1t	6MB / port
F3 12-port 100G N77-F312CK-26	144MB	4q1t	12MB / port

Egress

Module	Egress VOQ Structure	Egress VOQ Buffer (Credited)	Egress VOQ Buffer (Uncredited)
48-port 10G F2E N77-F248XP-23E	1p7q1t	295KB / port	256KB / 4 ports
F3 48-port 10G N77-F348XP-23	1p7q1t	295KB / port	512KB / 8 ports
F3 24-port 40G N77-F324FQ-25	1p7q1t	1.1MB / port	512KB / 2 ports
F3 12-port 100G N77-F312CK-26	1p7q1t	2.2MB / port	512KB / 1 port

Nexus 5600 Default Ingress Buffer Allocation

- Each cell is 320 bytes.
- Total number of cells for ingress buffer is 48,840.

Buffer Pool	10 GE Port	40 GE Port
Control traffic (per port)	64 KB	67.2 KB
SPAN (per port)	38.4 KB	153.6 KB
Class default (per port)	100 KB	100 KB
Shared buffer	13.2 MB	14.7 MB

Nexus 5600 Default Egress Buffer Allocation

- By default the majority of egress buffer is allocated for multicast traffic.
- Future software will provide CLI to tune the egress buffer allocation.
- At egress, unicast buffer is allocated on a per-port basis. For multicast, the egress buffer is shared among all ports.

Buffer Pool	10 GE Port	40 GE Port
Unicast (per port)	363 KB	650KB with 10G fabric mode 635KB with 40G fabric mode
Multicast (per ASIC)	4.3 MB	6.6 MB

FYI

Nexus 5600 Buffering Capacity

Ingress

Traffic Type	Ingress Queue Structure	10 GE Port	40 GE Port
Control traffic (per port)	6q1t	64 KB	67 KB
Span Traffic (per Port)	6q1t	38.4 KB	154 KB
Class Default (per Port)	6q1t	100 KB	100 KB
Shared Buffer	6q1t	13.2 MB	14.7 MB

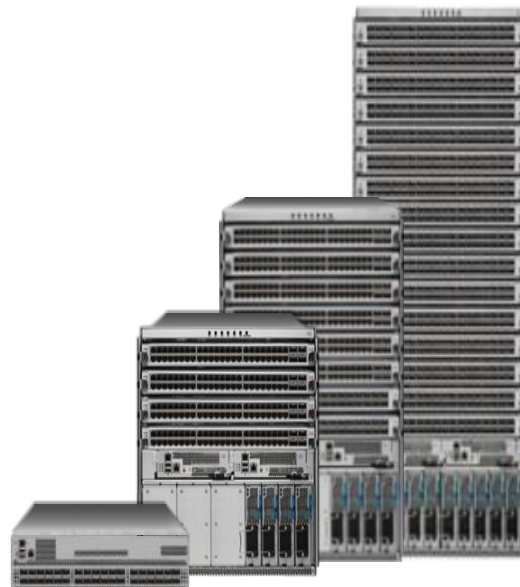
Egress

Traffic Type	Egress Queue Structure	10 GE Port	40 GE Port
Unicast	1p5q0t	363 KB	650 KB with 10GB Fabric Mode 635 KB with 40GB Fabric Mode
Multicast	1p5q0t	4.3MB	6.6 MB

Nexus 9000 QoS

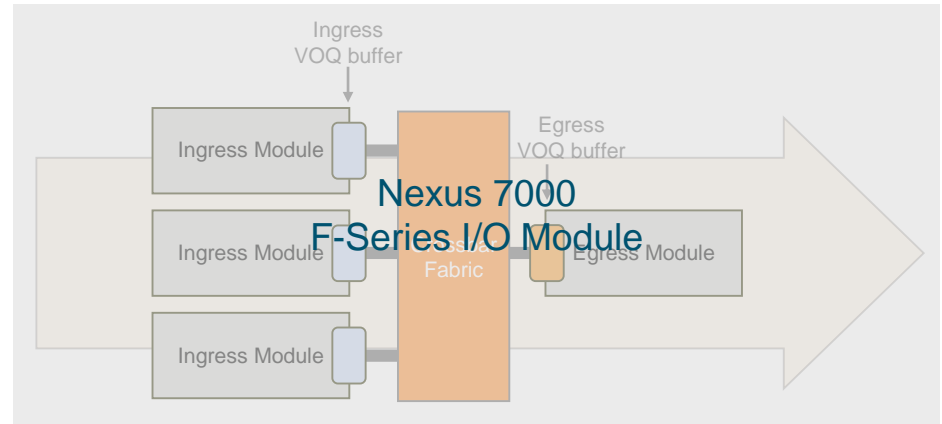
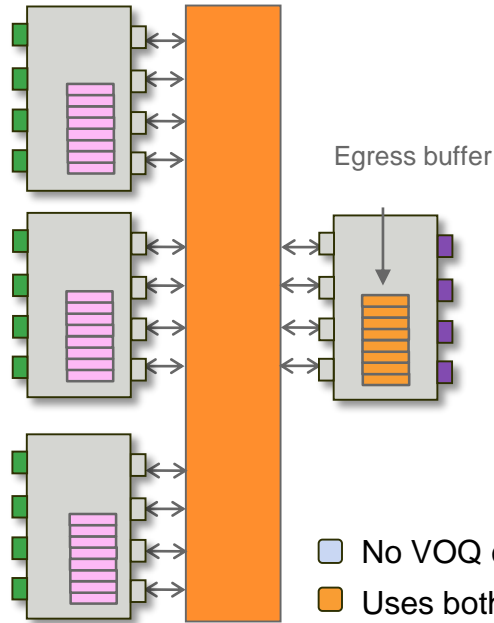
Nexus 9000 Overview

- Modular and fixed chassis
- Optimized for high density 40G/100G
- Standalone and ACI Mode
- Merchant+ Strategy
 - Mix of Merchant and Custom Silicon



Buffering Model

Nexus 9000 compared to Nexus 7000/Nexus 5000

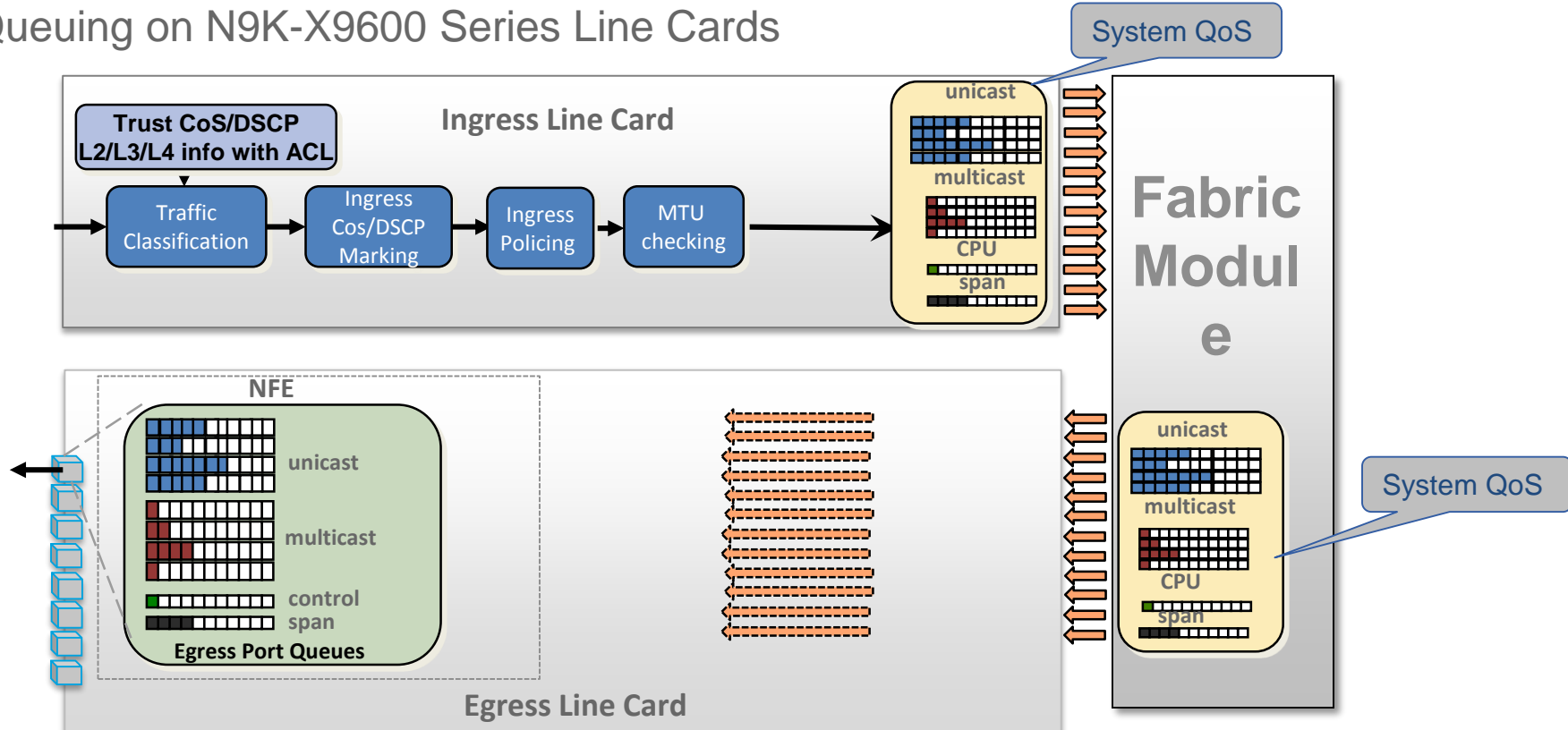


Cisco Nexus 9000 QoS Features

- **Traffic classification**
 - DSCP, CoS, and ACL
- **Packet marking**
 - DSCP, CoS, and ECN
- **Strict Priority Queuing and DWRR**
 - DCBX 802.1Qaz
- **Ingress policing only**
- **WRED, TD and ECN**
- **Shared buffer capability**
- **Buffer boost**

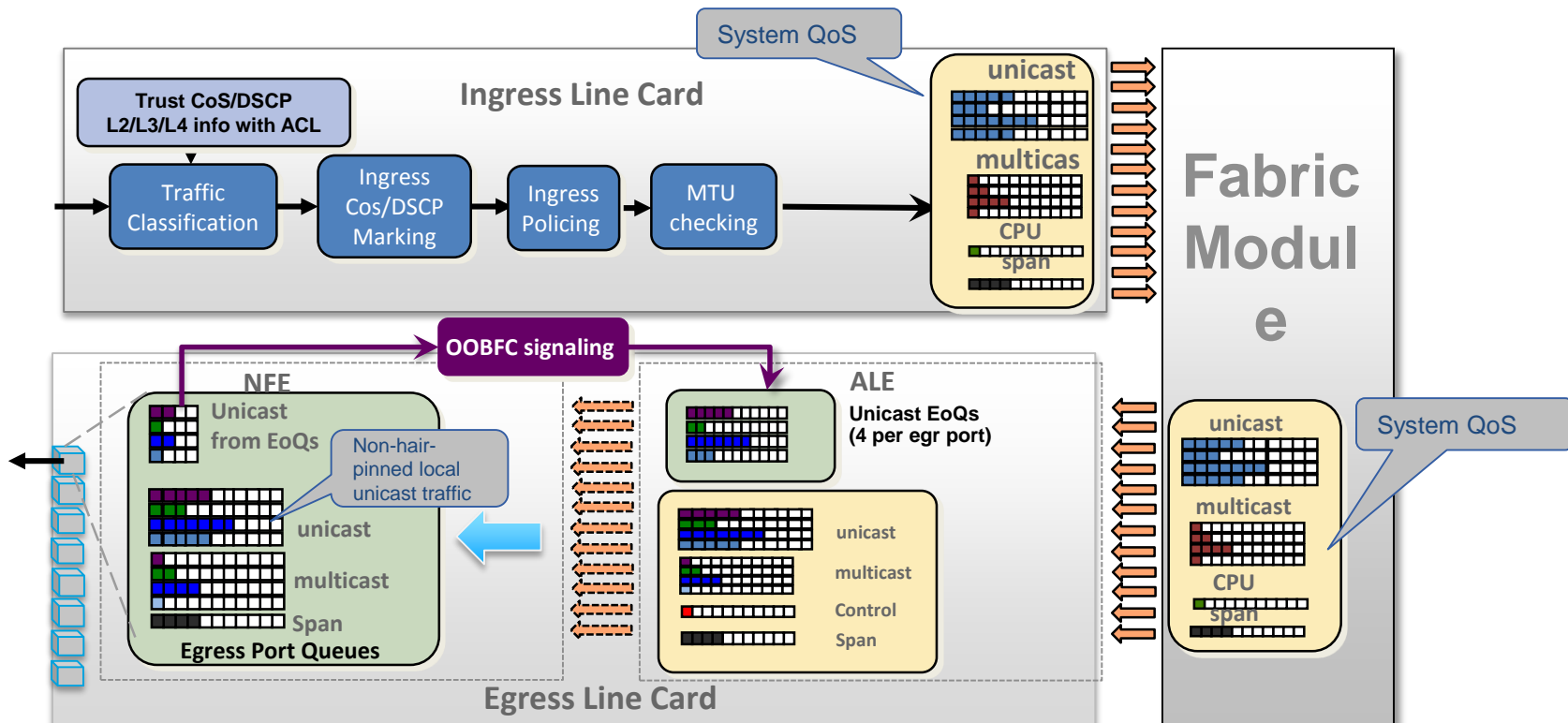
Queuing on Nexus 9500 Switch Line Cards

Queuing on N9K-X9600 Series Line Cards



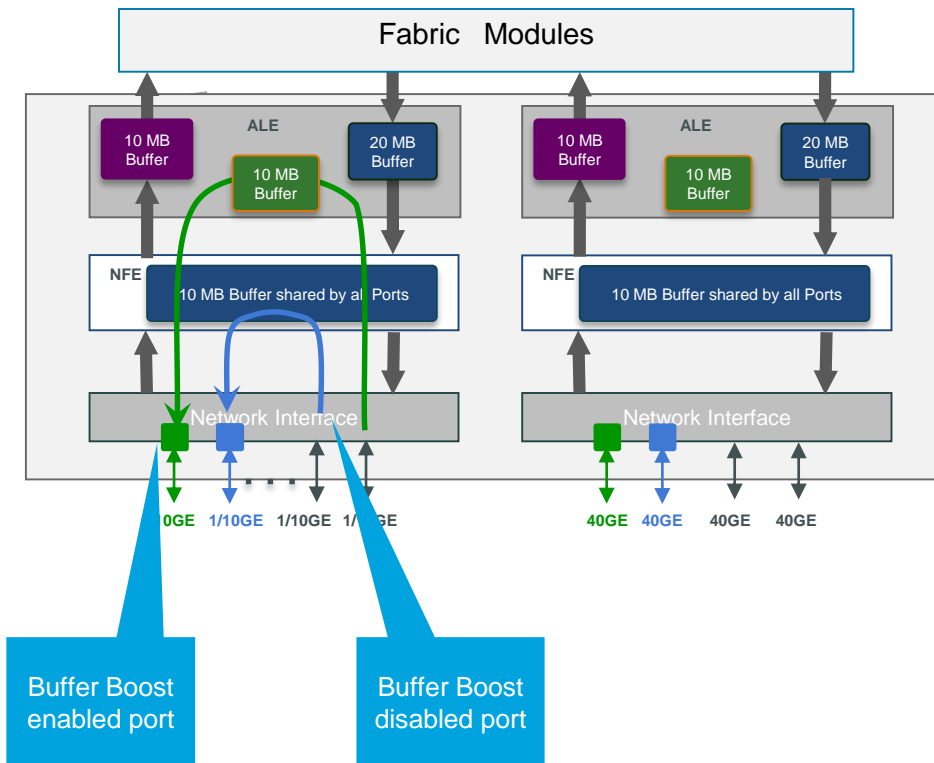
Queuing & Scheduling on Nexus 9500 Line Cards

Queuing on N9K-X9500 Series Line Cards



Buffer Boost on x9500 cards

- Line cards with ALEs can leverage the additional buffer on ALE for NFE local traffic.
- When buffer boost is enabled, NFE local traffic on the port is sent to ALE for additional buffer space.
- When buffer boost is disabled, NFE local traffic to the port remains local
- Buffer boost is enabled by default and can be disabled on a per-port basis.



FEX QoS Configuration Examples



Fex QoS Policy Configuration Example

```
policy-map type qos fex-qos
  class fex-qos-class-1
    set dscp 10
  class fex-qos-class-2
    set dscp 18
  class fex-qos-class-3
    set dscp 26
```

} Marking policy

!

```
interface Ethernet101/1/1
  service-policy type qos input fex-qos
```

} Policy applied on ingress of FEX HIF



Nexus 7000 Network-QoS Configuration Example #1

- Applying 8e-4q4q template to enable 4 ingress/egress queues on FEX with COS to queue mapping (also enables 4 ingress queues on F-series modules, if present)

```
system qos
service-policy type network-qos default-nq-8e-4q4q-policy
```

} Default 8e-4q4q template applied to "system qos" target

- FEX output ("show queuing interface"):

Queuing:

queue	qos-group	cos	priority	bandwidth	mtu
ctrl-hi	n/a	7	PRI	0	2400
ctrl-lo	n/a	7	PRI	0	2400
2	0	0 1	WRR	30	1600
3	1	2	WRR	30	1600
4	2	5 6	WRR	10	1600
5	3	3 4	WRR	30	1600

} 8e4q4q configuration (4 data traffic queues)





Nexus 7000 Network-QoS Configuration Example #2

- Applying custom 8e-4q4q-based template with new MTU

```

policy-map type network-qos custom-nq-8e-4q4q template 8e-4q4q
  class type network-qos c-nq-8e-4q4q
    congestion-control tail-drop
    mtu 9216
system qos
  service-policy type network-qos custom-nq-8e-4q4q

```

Custom network-qos policy with new MTU

Custom template applied to "system qos" target

- FEX output ("show queuing interface") after MTU change:

Queuing:

queue	qos-group	cos	priority	bandwidth	mtu
ctrl-hi	n/a	7	PRI	0	2400
ctrl-lo	n/a	7	PRI	0	2400
2	0	0 1	WRR	30	9280
3	1	2	WRR	30	9280
4	2	5 6	WRR	10	9280
5	3	3 4	WRR	30	9280

MTU increased on data traffic queues



Modifying CoS- or DSCP-to-Queue Mappings

- Changing CoS- or DSCP-to-queue mappings in parent switch F-type ingress queuing class-maps modifies mappings on FEX
- Queuing class-maps modified only in default/admin VDC (apply to entire system)

```
class-map type queuing match-any 4q1t-8e-4q4q-in-q1
  match cos 1-3
  match dscp 8-31
class-map type queuing match-any 4q1t-8e-4q4q-in-q-default
  match cos 0
  match dscp 0-7
class-map type queuing match-any 4q1t-8e-4q4q-in-q3
  match cos 4-5
  match dscp 32-47
class-map type queuing match-any 4q1t-8e-4q4q-in-q4
  match cos 6-7
  match dscp 48-63
```

Non-default F-series
ingress queuing
class-maps (COS and
DSCP match statements
modified)

FEX queue mappings
reflect changes

Queuing:

queue	qos-group	cos	priority	bandwidth	mtu
ctrl-hi	n/a	7	PRI	0	2400
ctrl-lo	n/a	7	PRI	0	2400
2	0	0	WRR	30	1600
3	1	6	WRR	30	1600
4	2	1 2 3	WRR	10	1600
5	3	4 5	WRR	30	1600

<...>

queue	DSCPs
02	0-7,
04	8-31,
03	48-63,
05	32-47,

EYE

Enabling FEX Queue Limits

- Example #1 – N2K-C2248TP-1GE

```
fex 101
```

```
hardware N2248T queue-limit 50000
```

- Example #2 – N2K-C2232TM-E-10GE

```
fex 102
```

```
hardware N2232TM-E queue-limit 50000
```

- FEX output (“show queuing interface”) **before**:

```
Queue limit: Disabled
```

- FEX output (“show queuing interface”) **after** (configured queue-limit rounded to nearest hardware supported value):

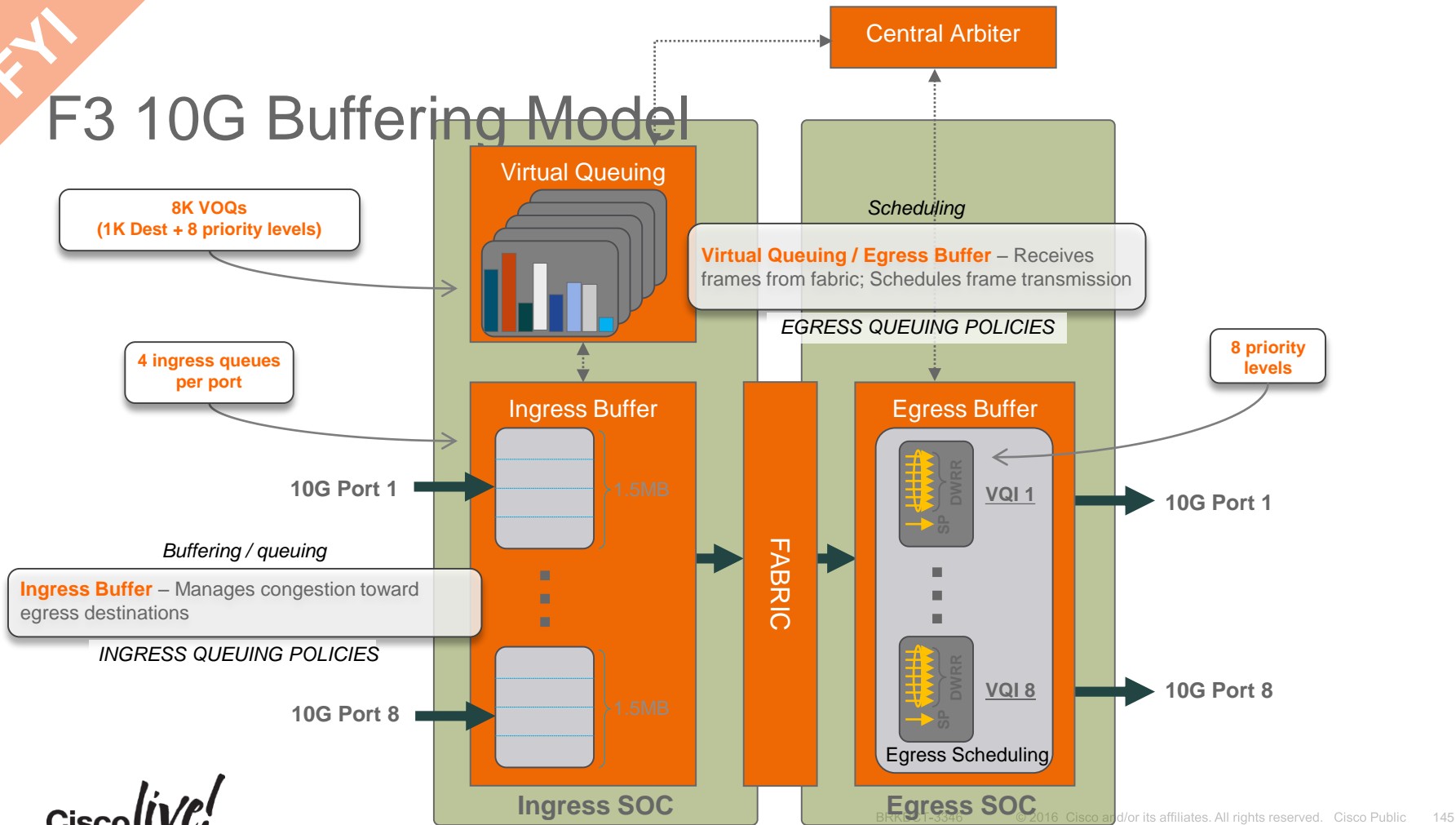
```
Queue limit: 51200 bytes
```



F3 Queuing Configuration Examples

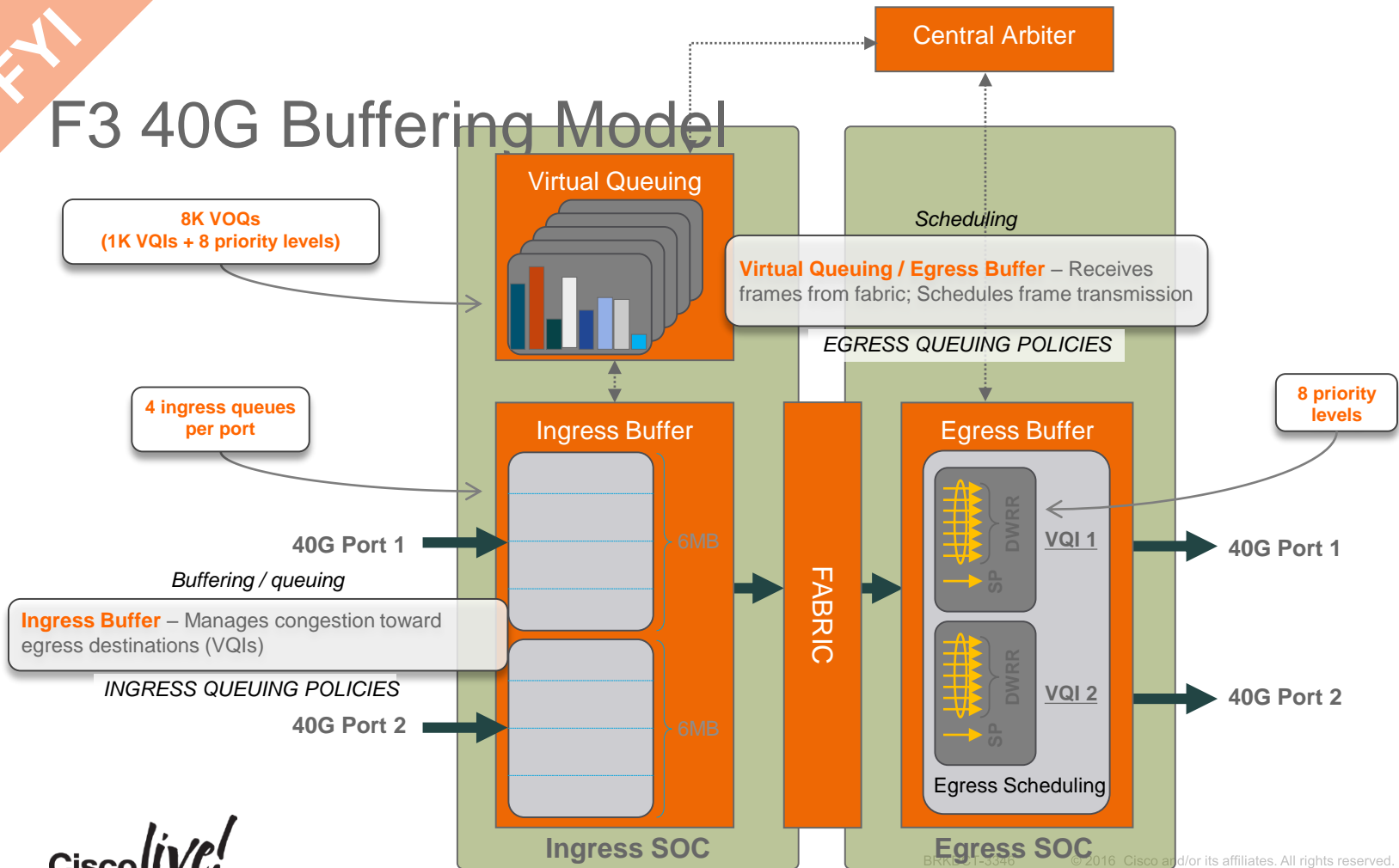
FYI

F3 10G Buffering Model



FYI

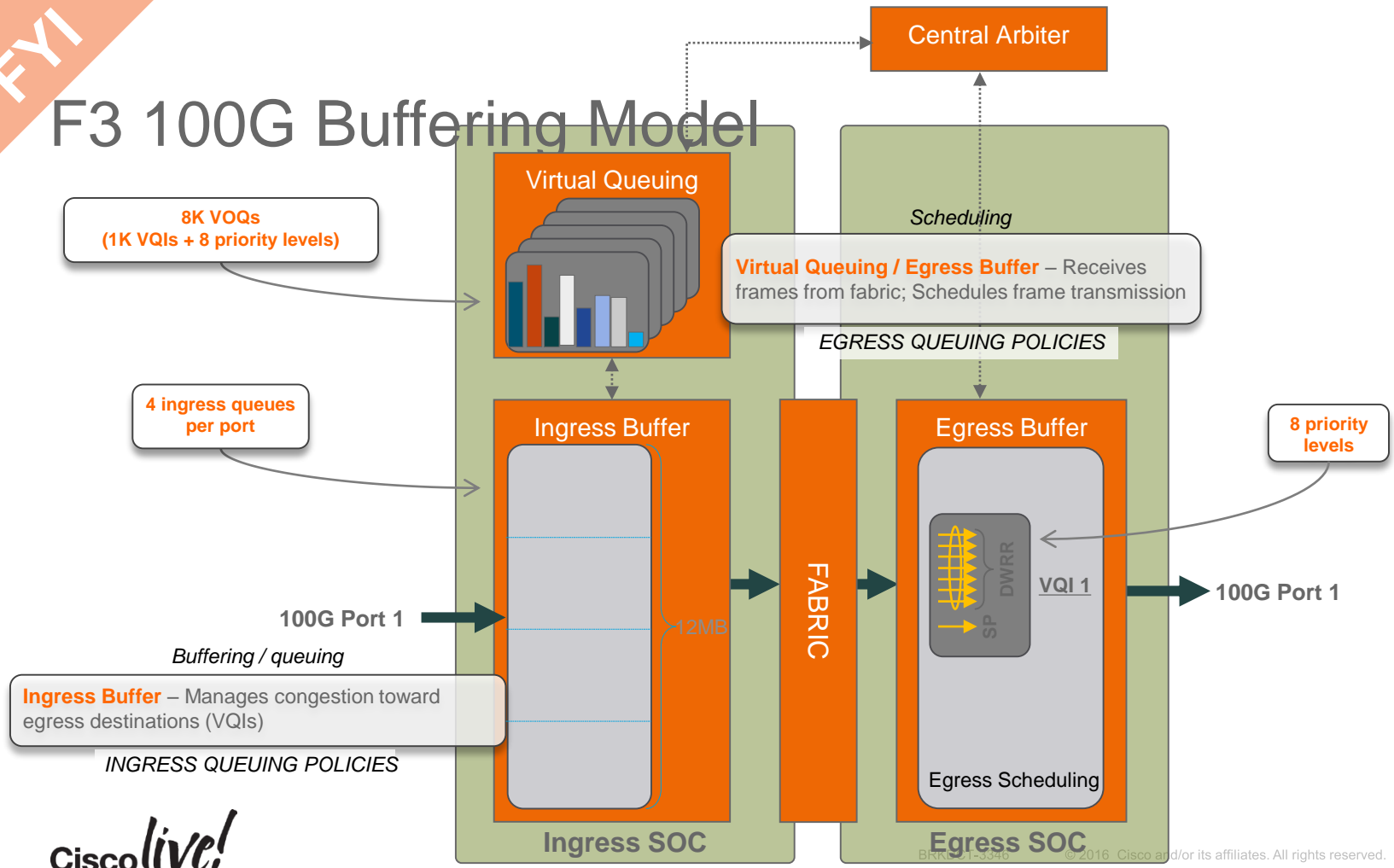
F3 40G Buffering Model



Cisco live!

FYI

F3 100G Buffering Model



Network-QoS and Default Queuing (global)

- default-nq-8e-policy is default `network-qos` policy and attached to `system qos` in Admin-/Default-VDC
- The system queuing policy applied by default can be overridden on a per port basis.

```
N7k# show policy-map type queuing | beg default 4q-8e
```

```
policy-map type queuing default-4q-8e-in-policy  
  class type queuing 2q4t-8e-in-q1  
    queue-limit percent 10  
    bandwidth percent 50  
  class type queuing 2q4t-8e-in-q-default  
    queue-limit percent 90  
    bandwidth percent 50
```

```
policy-map type queuing default-4q-8e-out-policy  
  class type queuing 1p3qlt-8e-out-pq1  
    priority level 1  
  class type queuing 1p3qlt-8e-out-q2  
    bandwidth remaining percent 33  
  class type queuing 1p3qlt-8e-out-q3  
    bandwidth remaining percent 33  
  class type queuing 1p3qlt-8e-out-q-default  
    bandwidth remaining percent 33
```

Note: show policy-map system does display similar output

2 ingress queues
with buffer ratio
1:9 and DWRR
weights 1:1

4 egress queues
with one priority
queue and DWRR
weights 1:1:1

Modifying Queuing and Scheduling Behaviour on F3 Modules

I want to...	Steps to follow
...remap COS/DSCP values from one queue to another queue without activating additional queues	<ol style="list-style-type: none">1. Modify the type queuing class-map(s) for the desired queue(s)
...change queuing behavior without changing COS-or DSCP-to-queue mapping	<ol style="list-style-type: none">1. Define new type queuing policy-map (you cannot modify the default policies)2. Modify class-map parameters3. Apply new policy-map to interfaces
...activate additional queues and remap COS/DSCP values	<ol style="list-style-type: none">1. Define new type queuing policy-map2. Modify COS-to-queue mapping for target port type3. Apply new policy-map to interfaces
...shape the SP queue	<ol style="list-style-type: none">1. (Optional) Clone the default egress queuing policy2. Shape the SP queue in the new (cloned) policy3. Apply the new queuing policy to the target interfaces

FYI Modifying Queuing Behavior

Remap Some COS/DSCP Values from One Queue to Another Queue without Activating Additional Queues

Modify “type queuing” class-map(s) for desired queue(s)

Remap COS- or DSCP-to-queue mapping for given queue(s)

Important: changing COS- or DSCP-to-queue mapping **takes effect immediately** and is **disruptive** to all ports

EY Modifying Queuing Behavior

Remap Some COS/DSCP Values from One Queue to Another Queue without Activating Additional Queues

Example: remap COS 4 and DSCP 32-39 to ingress queue "q1":

```
n77# show class-map type queuing 8q2t-in-q1
Type queuing class-maps
=====
class-map type queuing match-any 8q2t-in-q1
  Description: Classifier for ingress queue 1 of type 8q2t
  match cos 5-7
  match dscp 40-63
n77# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
n77(config)# ! Modify ingress queue q1
n77(config)# class-map type queuing match-any 8q2t-in-q1
n77(config-cmap-que)# ! Change COS- and DSCP-to-queue mapping for this queue
n77(config-cmap-que)# match cos 4
n77(config-cmap-que)# match dscp 32-39
n77(config-cmap-que)# show class-map type queuing 8q2t-in-q1
Type queuing class-maps
=====
class-map type queuing match-any 8q2t-in-q1
  Description: Classifier for ingress queue 1 of type 8q2t
  match cos 4-7
  match dscp 32-63
n77(config-cmap-que)#
```

Show current mapping

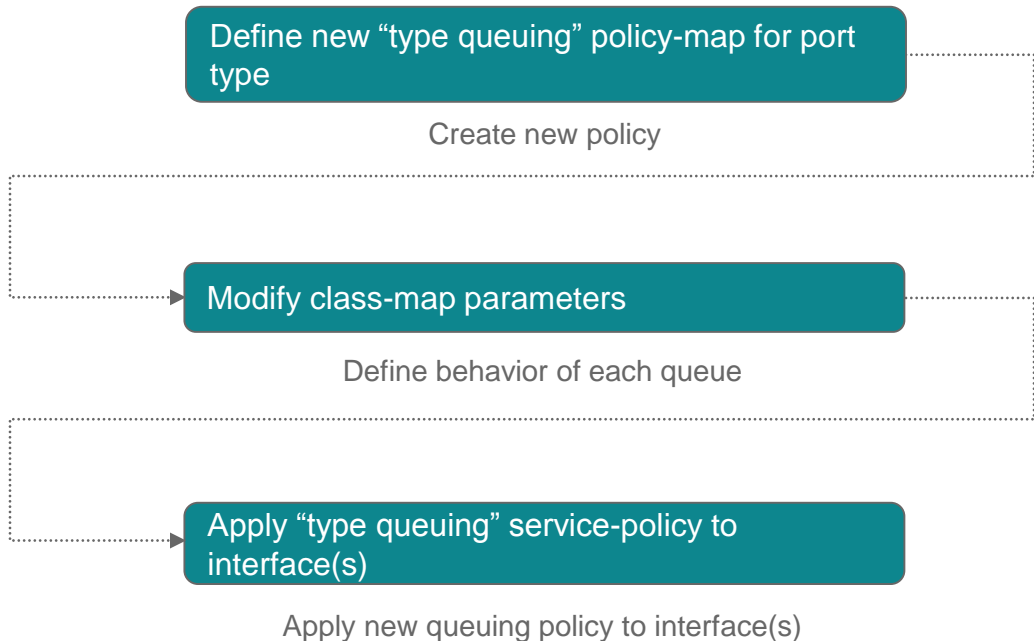
Configure new mapping

Show new mapping



FYI Modifying Queuing Behavior

Changing Default Queuing Behavior without Changing COS- or DSCP-to-Queue



Important: applying new queuing policy **takes effect immediately** and is **disruptive** to any ports to which the policy is applied

Modifying Queuing Behavior

Changing Default Queuing Behavior without Changing COS- or DSCP-to-Queue

Example: Resize ingress queues without modifying COS- or DSCP-to-queue mapping

```
n77# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
n77(config)# ! Define new "type queuing" policy
n77(config)# policy-map type queuing new-f3-ingress
n77(config-pmap-que)# ! Define behavior for F3 ingress q-default
n77(config-pmap-que)# class type queuing 8e-4q8q-in-q-default
n77(config-pmap-c-que)# ! Resize this queue
n77(config-pmap-c-que)# queue-limit percent 74
n77(config-pmap-c-que)# ! Define behavior for F3 ingress queue 1
n77(config-pmap-c-que)# class type queuing 8e-4q8q-in-q1
n77(config-pmap-c-que)# ! Resize this queue
n77(config-pmap-c-que)# queue-limit percent 24
n77(config-pmap-c-que)# ! Policy must include all queues (even inactive)
n77(config-pmap-c-que)# class type queuing 8e-4q8q-in-q3
n77(config-pmap-c-que)# ! Must give at least 1% to inactive queues
n77(config-pmap-c-que)# queue-limit percent 1
n77(config-pmap-c-que)# class type queuing 8e-4q8q-in-q4
n77(config-pmap-c-que)# queue-limit percent 1
n77(config-pmap-c-que)# interface e 2/1-48
n77(config-if-range)# ! Apply the new policy to F3 interfaces
n77(config-if-range)# service-policy type queuing input new-f3-ingress
n77(config-if-range)#
```

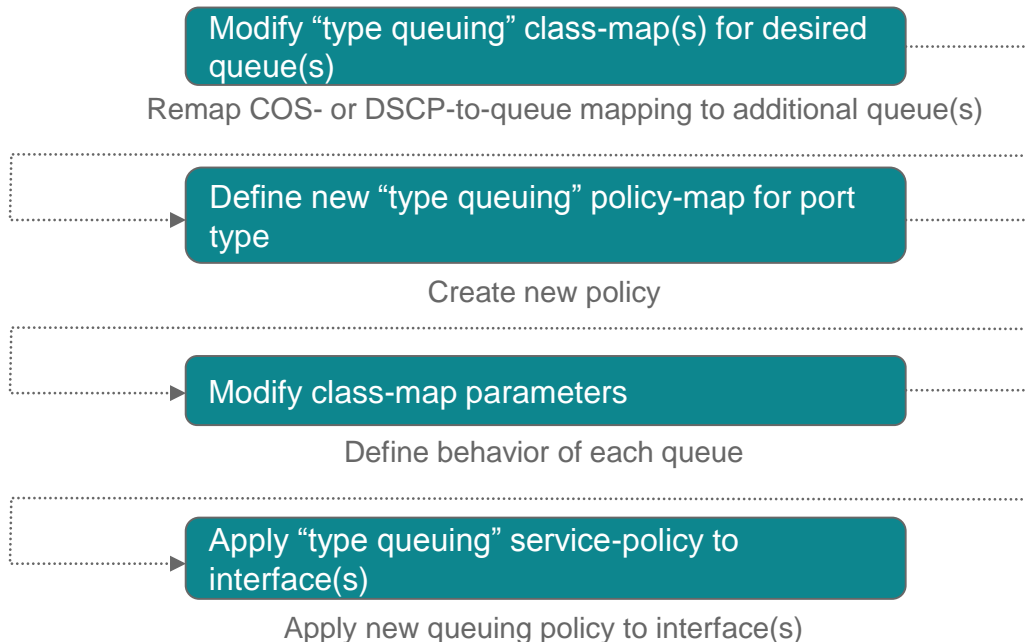
Create new queuing policy

Modify class-map parameters (resize queues)

Apply policy to interface(s)

Modifying Queuing Behavior

Activate Additional Queues and Remap COS/DSCP Values



Important: changing COS/DSCP-to-queue mapping and **takes effect immediately** and is **disruptive** to all ports; applying new queuing policy **takes effect immediately** and is **disruptive** to any ports to which the policy is applied

Modifying Queuing Behavior

Activate Additional Queues and Remap COS/DSCP Values

Example: Enable one additional ingress queue and map COS/DSCP values to all active queues

```
n77# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
n77(config)# ! Modify ingress queue q3
n77(config)# class-map type queuing match-any 8q2t-in-q3
n77(config-cmap-que)# ! Map COS and DSCP values to this queue
n77(config-cmap-que)# match cos 3-4
n77(config-cmap-que)# match dscp 24-39
n77(config-cmap-que)# ! Define new "type queuing" policy
n77(config-cmap-que)# policy-map type queuing new-f3-ingress
n77(config-pmap-que)# ! Define behavior for F3 ingress q-default
n77(config-pmap-que)# class type queuing 8e-4q8q-in-q-default
n77(config-pmap-c-que)# ! Resize this queue
n77(config-pmap-c-que)# queue-limit percent 60
n77(config-pmap-c-que)# ! Define behavior for F3 ingress queue 1
n77(config-pmap-c-que)# class type queuing 8e-4q8q-in-q1
n77(config-pmap-c-que)# ! Resize this queue
n77(config-pmap-c-que)# queue-limit percent 25
n77(config-pmap-c-que)# ! Define behavior for F3 ingress queue 3
n77(config-pmap-c-que)# class type queuing 8e-4q8q-in-q3
n77(config-pmap-c-que)# ! Resize this queue
n77(config-pmap-c-que)# queue-limit percent 14
n77(config-pmap-c-que)# ! Policy must include all queues (even inactive)
n77(config-pmap-c-que)# class type queuing 8e-4q8q-in-q4
n77(config-pmap-c-que)# queue-limit percent 1
n77(config-pmap-c-que)# int e 2/1-48
n77(config-if-range)# ! Apply the new policy to F3 interfaces
n77(config-if-range)# service-policy type queuing input new-f3-ingress
```

Map COS/DSCP values to inactive queue

Create new queuing policy

Modify class-map parameters (resize queues)

Apply policy to interface(s)

Important!

- If you change the COS- or DSCP-to-queue mapping for a port type, make sure **all** ports of that type in **all** VDCs have a queuing policy applied that defines behavior for **all** queues with COS/DSCP values mapped

- For example, if you do THIS...

```
n77(config)# class-map type queuing match-any 8e-4q8q-in-q-default
n77(config-cmap-que)# match cos 0-1
n77(config-cmap-que)# match dscp 0-15
n77(config-cmap-que)# class-map type queuing match-any 8e-4q8q-in-q4
n77(config-cmap-que)# match cos 2-4
n77(config-cmap-que)# match dscp 16-39
n77(config-cmap-que)# class-map type queuing match-any 8e-4q8q-in-q3
n77(config-cmap-que)# match cos 6-7
n77(config-cmap-que)# match dscp 48-63
n77(config-cmap-que)# class-map type queuing match-any 8e-4q8q-in-q1
n77(config-cmap-que)# match cos 5
n77(config-cmap-que)# match dscp 40-47
```

Changes the default COS/
DSCP-to-queue mapping

FYI

Important!

- ... then make sure you do THIS...

```
n77(config)# policy-map type queuing new-f3-ingress
n77(config-pmap-que)# class type queuing 8e-4q8q-in-q-default
n77(config-pmap-c-que)# queue-limit percent 50
n77(config-pmap-c-que)# class type queuing 8e-4q8q-in-q1
n77(config-pmap-c-que)# queue-limit percent 20
n77(config-pmap-c-que)# class type queuing 8e-4q8q-in-q3
n77(config-pmap-c-que)# queue-limit percent 20
n77(config-pmap-c-que)# class type queuing 8e-4q8q-in-q4
n77(config-pmap-c-que)# queue-limit percent 10
n77(config-pmap-c-que)#
```

Defines a new queuing policy that defines behavior of all queues that COS/DSCP values have been mapped to

FYI

Important!

- ... and then do THIS:

```
n77(config)# int e 2/1-48
n77(config-if-range)# service-policy type queuing input new-f3-ingress
n77(config-if-range)#
```

Maps the new policy to **ALL** interfaces in the system (do this on **ALL** ports in EVERY VDC!!)

- If you DON'T, traffic arriving on ports with default policy (i.e., without all queues activated that have COS/DSCP values mapped) will suffer – packet drops, poor performance, etc.
- Of course, you can have *different* non-default policies on different sets of interfaces, but all interfaces in the system must use some policy that defines all activated queues!

Modifying Queuing Behavior

Shape the SP Queue

(Optional) Clone a default egress “type queuing” policy-map

Create a copy of a default egress queuing policy

Shape SP queue in new (cloned) “type queuing” policy

Limit SP queue bandwidth consumption

Apply new “type queuing” policy to target interface(s)

Apply new queuing policy to interfaces

Important: applying new queuing policy **takes effect immediately** and is **disruptive** to any ports to which the policy is applied

Modifying Queuing Behavior

Shape the SP Queue

Example: Shape the SP queue to 2Gbps on a 10G interface, using a queuing policy cloned from the default “8e4q4q” egress queuing policy

```
n77# ! Clone the 8E egress queuing policy
n77# qos copy policy-map type queuing default-8e-4q8q-out-policy prefix new-
n77# conf
Enter configuration commands, one per line. End with CNTL/Z.
n77(config)# ! Modify new queuing policy
n77(config)# policy-map type queuing new-8e-4q8q-out
n77(config-pmap-que)# ! Modify egress queue q1
n77(config-pmap-que)# class type queuing 8e-4q8q-out-q1
n77(config-pmap-c-que)# ! Make this queue strict priority
n77(config-pmap-c-que)# priority level 1
n77(config-pmap-c-que)# ! Shape the queue to 20% (2G on 10G port)
n77(config-pmap-c-que)# shape percent 20
n77(config-pmap-c-que)# int e 2/1-48
n77(config-if-range)# ! Apply the new policy to target interfaces
n77(config-if-range)# service-policy type queuing output new-8e-4q8q-out
n77(config-if-range)#
n77#
```

Clone the default egress queuing policy

Modify the cloned policy

Make q1 Strict Priority and shape to 20% (2G)

Apply new policy to target interfaces

FYI

Changing The Default Trust (M-Series I/O Module)

- You can make an interface untrusted (CoS & DSCP)
 - CoS for bridged traffic
 - DSCP for routed traffic
- You need two Policies
 - A "type queuing" policy to set the CoS to 0
 - A "type qos" policy to set the DSCP to 0
- Set DSCP will set the CoS value for Bridged traffic as well.

```
policy-map type queuing Reset-CoS
  class type queuing 8q2t-in-q-default
    set cos 0
    bandwidth percent 100
    queue-limit percent 100
!
policy-map type qos Reset-DSCP
  class class-default
    set dscp 0
!
! Tie to an interface:
interface Ethernet1/1
  service-policy type queuing input Reset-CoS
  service-policy type qos input Reset-DSCP
```

FYI

Changing The Default Trust (F-Series I/O Module)

```
qos copy policy-map type queuing default-4q-8e-in-policy prefix UNTRUSTED-
!
policy-map type queuing untrusted-4q-8e-in
  class type queuing 2q4t-8e-in-q1
    queue-limit percent 1
  class type queuing 2q4t-8e-in-q-default
    queue-limit percent 99
    set cos 0
!
policy-map type qos UNTRUSTED
  class class-default
    set dscp 0
!
! Tie to an interface:
interface Ethernet1/1
  service-policy type queuing input untrusted-4q-8e-in
  service-policy type qos input UNTRUSTED
```