Catalyst 9000 Switching QoS
Deep Dive

Ninad Diwakar – Technical Marketing Engineer
BRKENS-2096

The bridge to possible
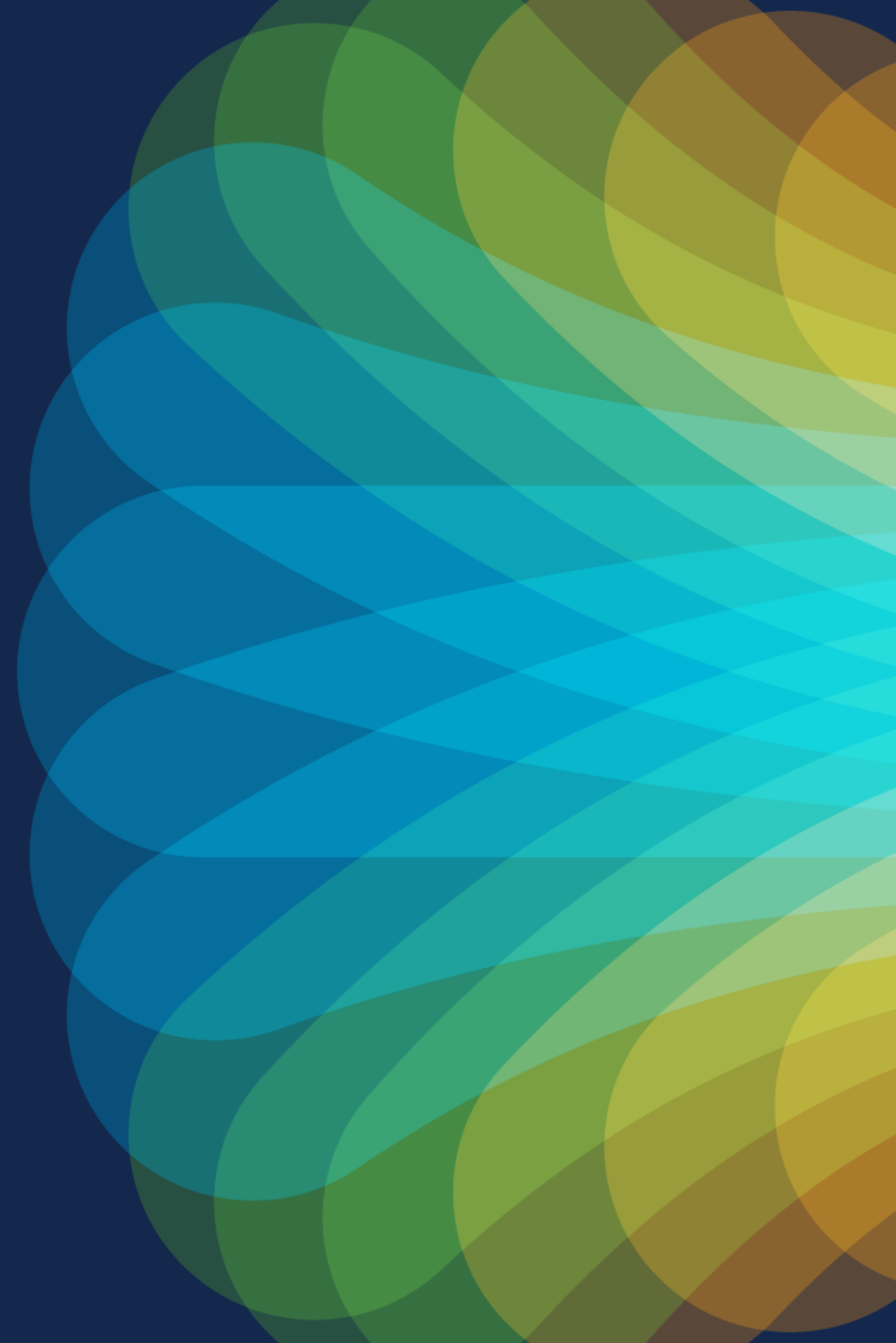
#CiscoLiveAPJC

Agenda

- Introduction

- QoS Overview

- UADP QoS

- Silicon One Q200 QoS

- Config Migration

- Conclusion

Introduction and Overview

Do we need QoS?

User Experience

Guaranteeing voice quality

Bandwidth Savvy Business Applications

protect network infrastructure to deal with abnormal events

Video Quality

de-prioritizing non-business applications

protecting the control planes

QoS helps define the latency priority for your traffic packets

What does QoS do?

QoS provides a set of tools to treat different traffic differently

Identify incoming traffic	Classification
Treat traffic differently	Scheduling and Queueing
Label traffic	Trust, Marking
Limit traffic	Policing and Shaping
What happens when queue is congested?	Congestion management (WTD and WRED)

Modular QoS CLI (MQC)

class-map
    What traffic do we care about?

policy-map
    What actions do we take on the classes?

service-policy
    Where do we apply the policy?

```
class-map match-any Voice
  match dscp ef
class-map match-any Video
  match dscp 34


Policy-map POLICY-QOS
  class Voice
    priority level 1
  class Video
    set dscp 10


interface x/y
  service-policy (input/output) POLICY-QOS
```
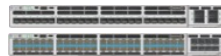
Catalyst 9000 family of Switches

UADP 2.0x

Catalyst® 9200CX Series

Catalyst 9300X models

Catalyst 9200 Series

Catalyst 9300 Series

Catalyst 9400 Series

Catalyst 9500 Series

UADP 3.0x
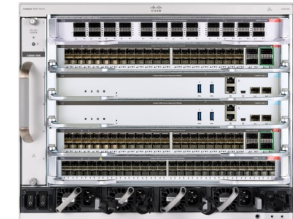
Catalyst 9600 Series with Sup-1

Catalyst 9400X models

Catalyst 9500 High performance Series

Silicon One Q200

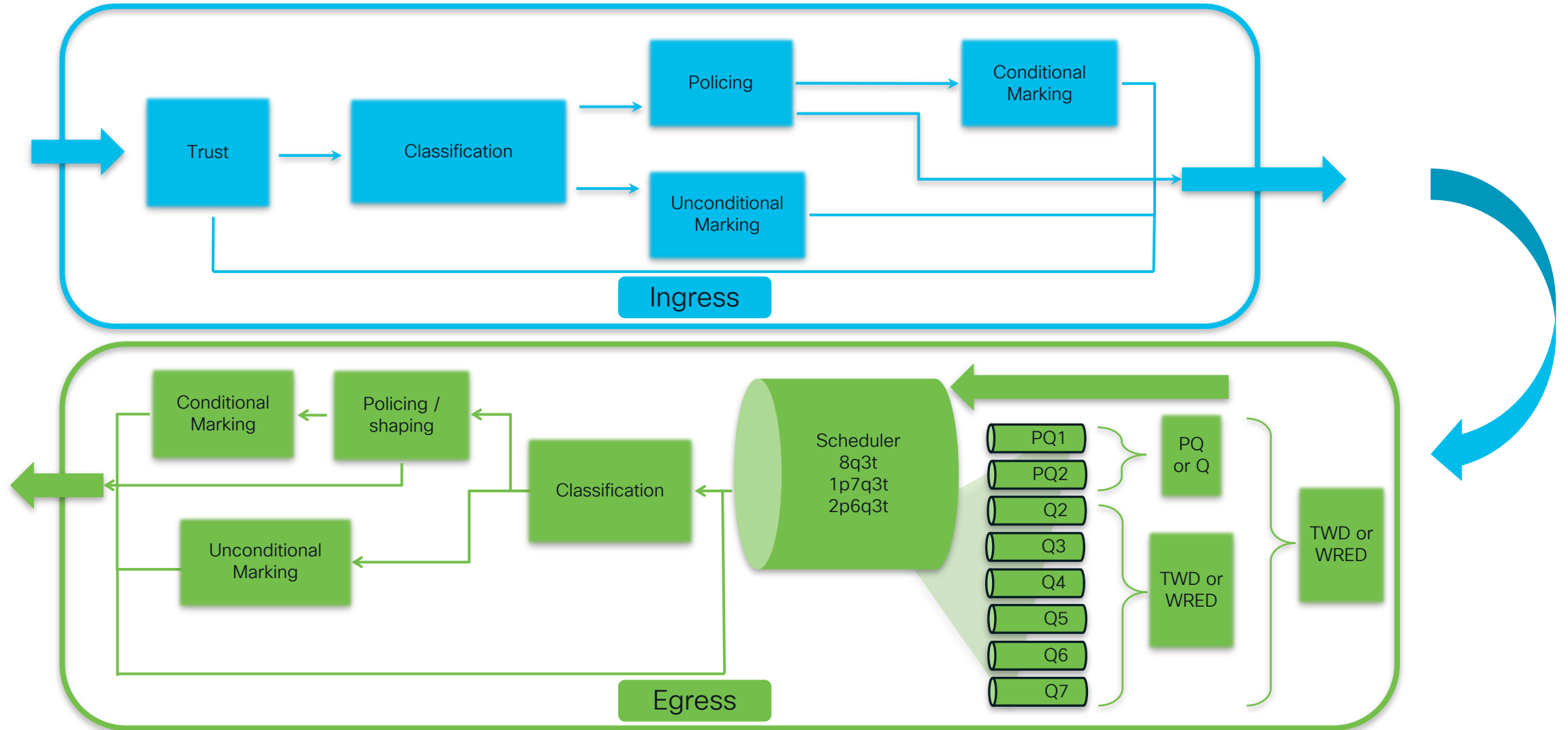Catalyst 9600 Series with Sup-2

Catalyst 9500X Series

Class-based Weighted Fair Queue (CBWFQ) QoS
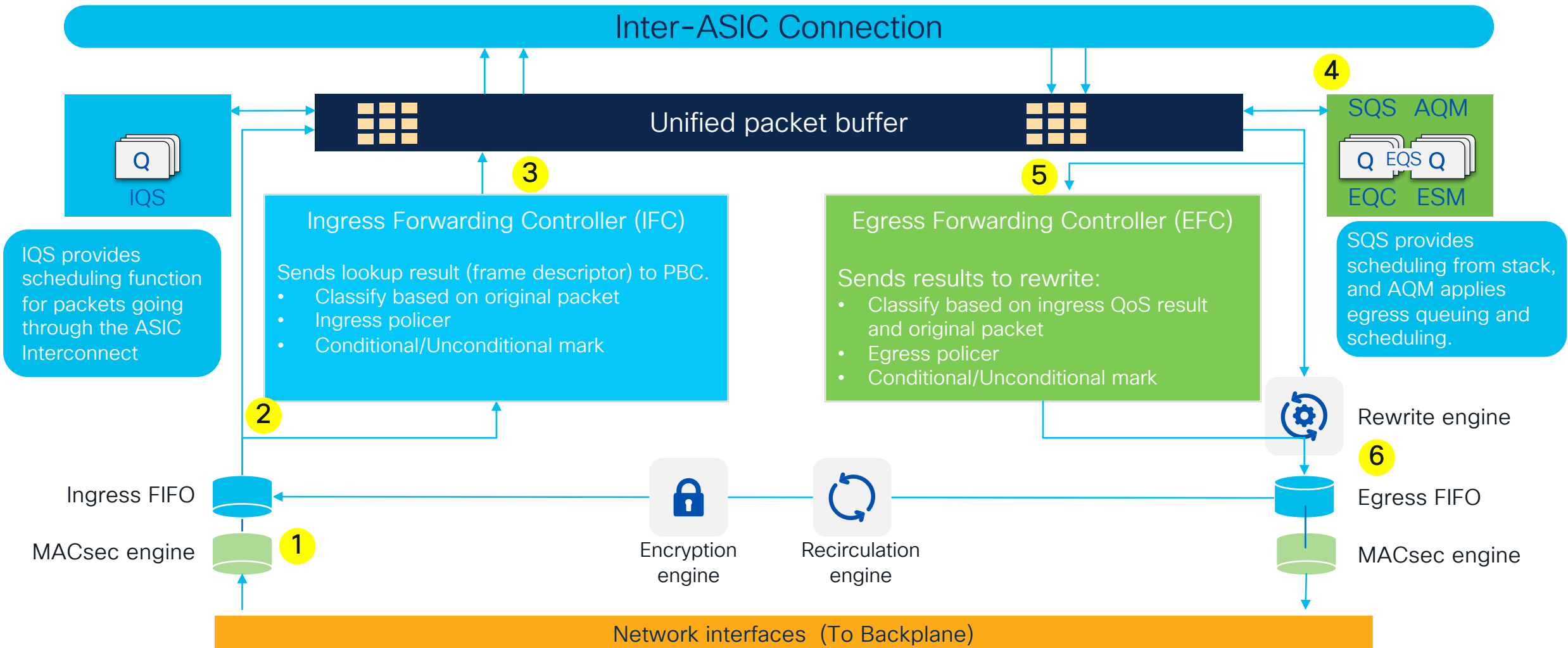
Virtual Output Queue (VoQ) QoS
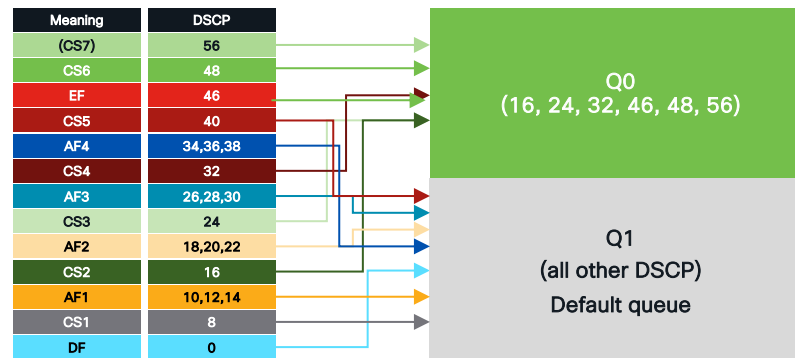
UADP QoS

QoS Fundamental Actions in UADP

WRED: up to 4 queues with UADP 2.0x; up to 8 queues with UADP 3.0x

UADP QoS forwarding
ingress and egress

Inter-ASIC Connection

Unified packet buffer

4

SQS AQM

Q EQS Q

EQC ESM

IQS

3

Ingress Forwarding Controller (IFC)

Sends lookup result (frame descriptor) to PBC.
- Classify based on original packet
- Ingress policer
- Conditional/Unconditional mark

5

Egress Forwarding Controller (EFC)

Sends results to rewrite:
- Classify based on ingress QoS result and original packet
- Egress policer
- Conditional/Unconditional mark

IQS provides scheduling function for packets going through the ASIC Interconnect

SQS provides scheduling from stack, and AQM applies egress queuing and scheduling.

2

Ingress FIFO

MACsec engine **1**

Encryption engine

Recirculation engine

Rewrite engine

6

Egress FIFO

MACsec engine

Network interfaces (To Backplane)

UADP QoS Default

- Catalyst 9000 Switches with UADP ASICs
  - QoS enabled
  - All ports trust at layer2 and layer3
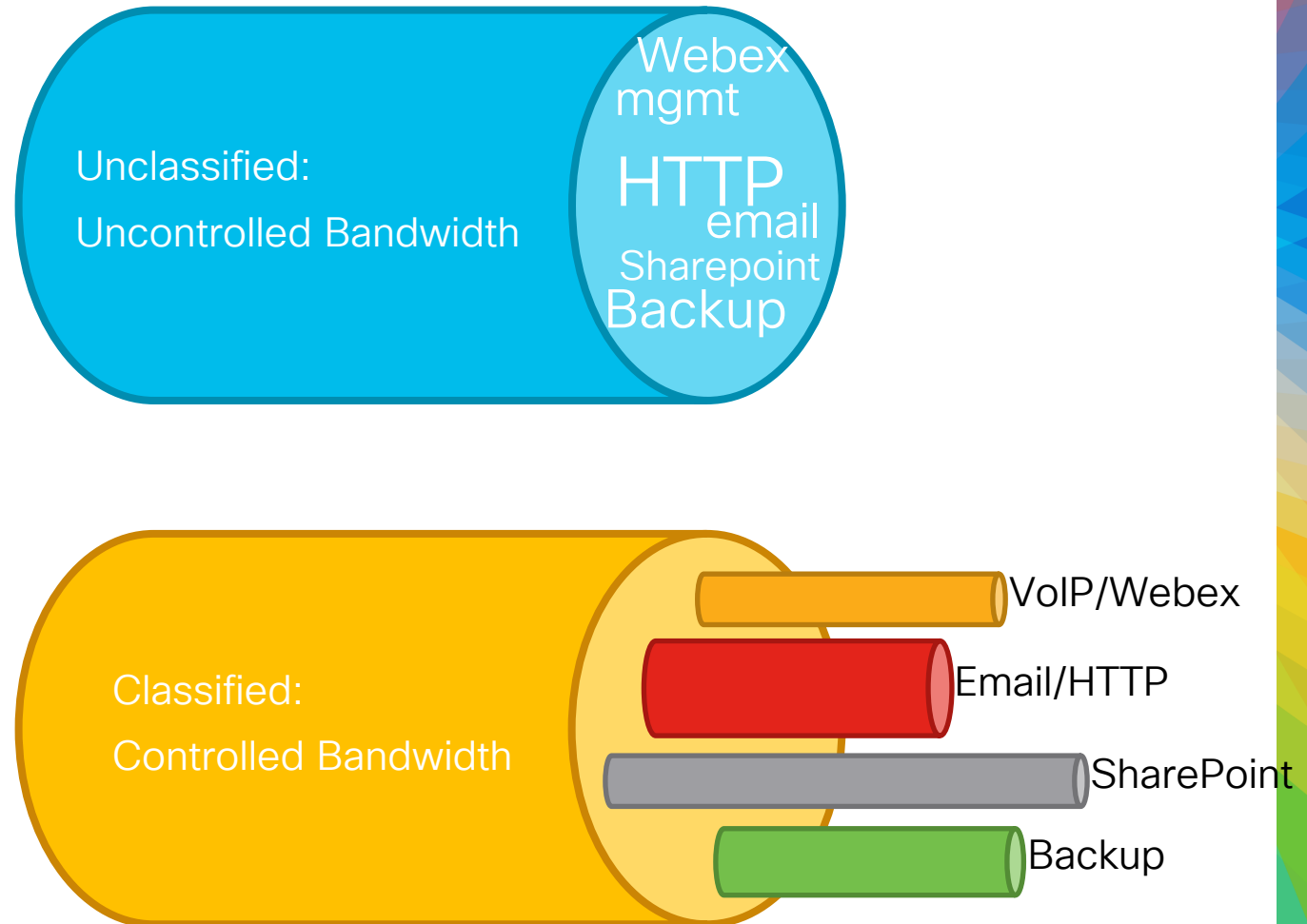 - Two queues (neither set as priority)

Meaning	DSCP
(CS7)	56
CS6	48
EF	46
CS5	40
AF4	34,36,38
CS4	32
AF3	26,28,30
CS3	24
AF2	18,20,22
CS2	16
AF1	10,12,14
CS1	8
DF	0

Q0
(16, 24, 32, 46, 48, 56)

Q1
(all other DSCP)
Default queue

Classification, Marking and Policing

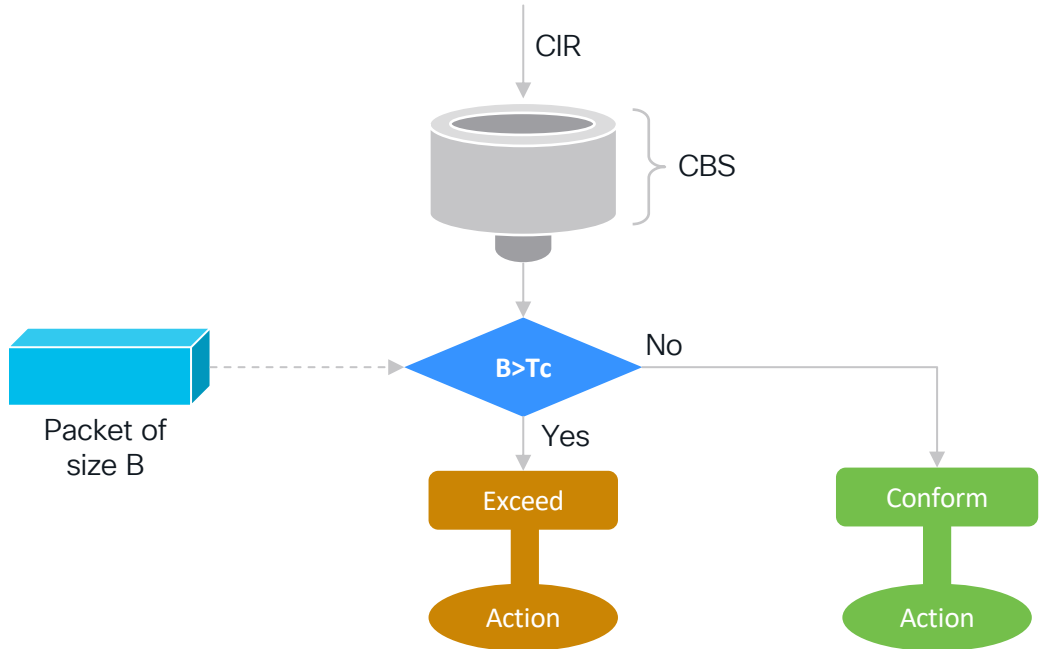# Classification and Marking

- Identify traffic
  - Access Control Lists (ACLs)
  - DSCP
  - IP precedence
  - CoS
  - QoS Group (local with the switch)
  - EXP (MPLS)
  - Network-Based Application Recognition (NBAR) protocols *
  - VLANs

- Marking
  - Conditional or unconditional
  - Table map (default-class)
  - QoS group (local within switch)

* Access platforms

Unclassified:
Uncontrolled Bandwidth

Webex
mgmt
HTTP
email
Sharepoint
Backup

Classified:
Controlled Bandwidth

VoIP/Webex
Email/HTTP
SharePoint
Backup

Policing – Limit the traffic

1 rate 2 color


```
police cir 1g bc 3125000
    conform-action set-dscp-transmit af41
    exceed-action drop
```

2 rate 3 color

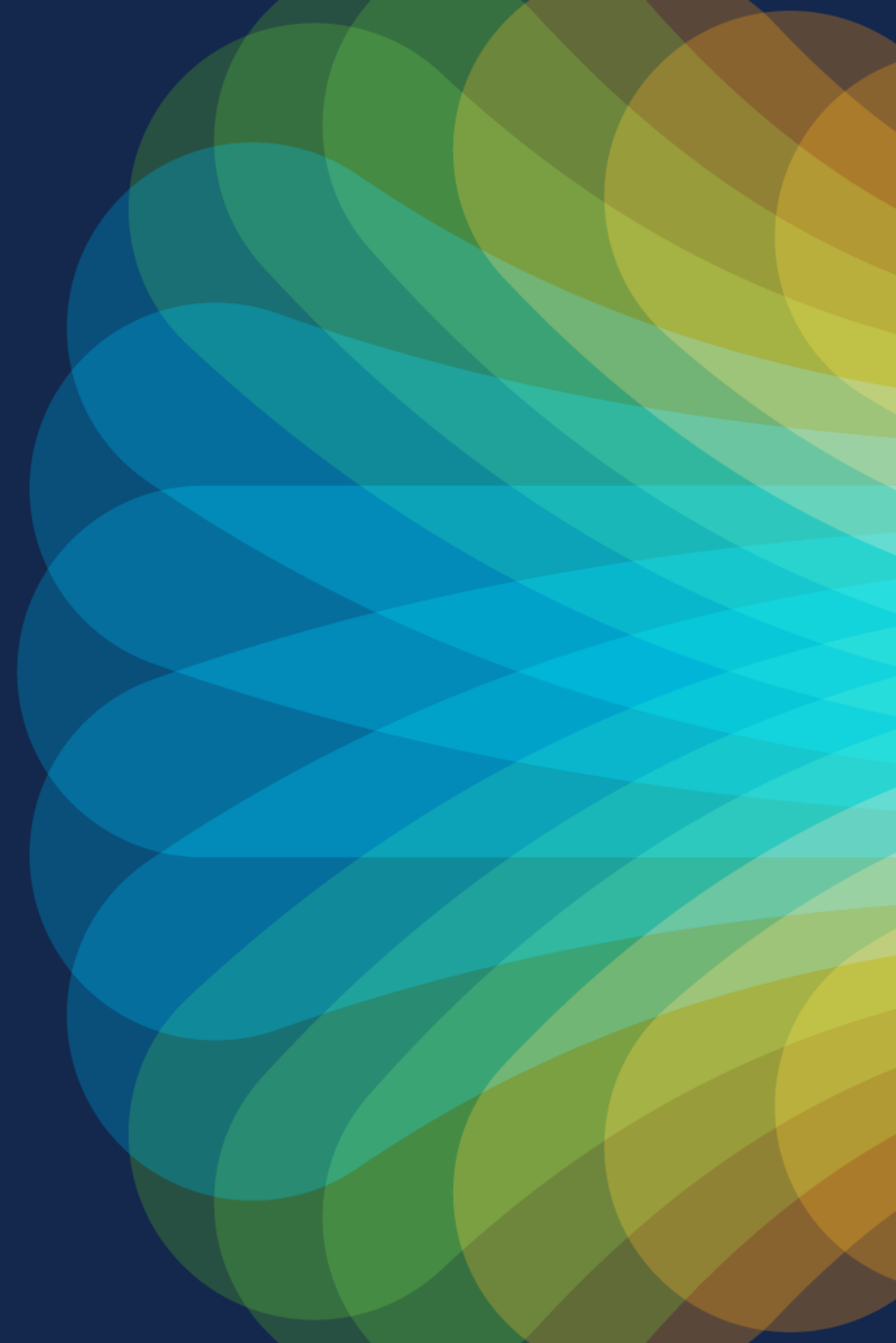

```
police cir percent 10 pir percent 50
    conform-action transmit
    exceed-action set-dscp-transmit dscp table MARKDOWN
    violate-action drop
```

CIR – Committed Information Rate   PBS- Peak Burst Size
PIR – Peak Information Rate   CBS – Committed Burst Size

Queueing, Scheduling and Shaping

CISCO *Live!*

Queueing

- Separate the traffic into the queues

- Traffic in different queue can be treated differently

- Up-to 8 queues per interface, 2 of which can be priority-queues (PQ).

- Both priority-queues are strict priority queues.

- Policer or a shaper on the priority queue will limit the traffic to the configured value regardless of the traffic level on other queues.
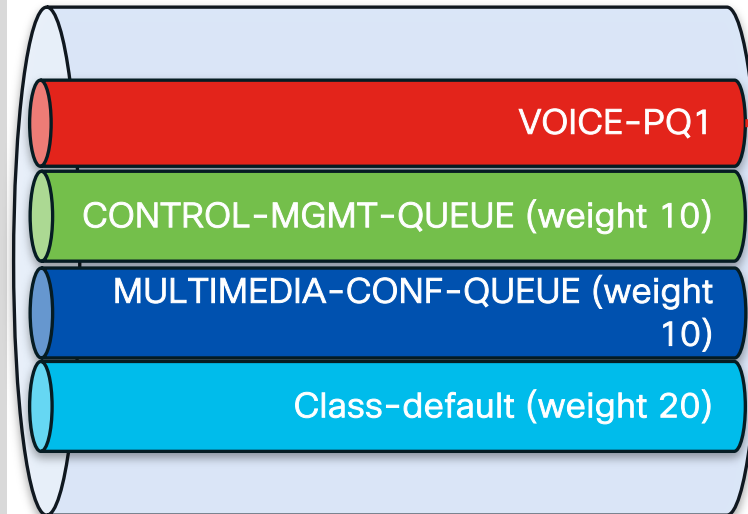

Buffer Memory
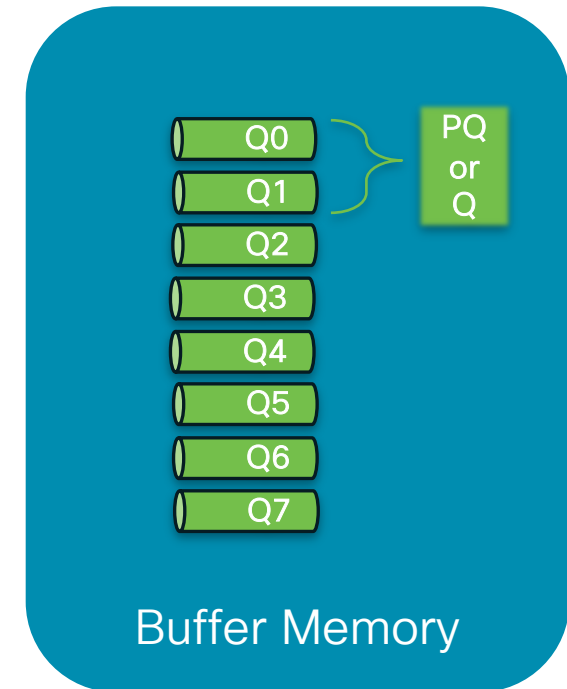
Queueing Terminology
1P3Q Example

Policy Map Configuration

```
policy-map 2P6Q3T
class VOICE-PQ1
  priority level 1
class CONTROL-MGMT-QUEUE
  bandwidth remaining percent 10
class MULTIMEDIA-CONFERENCING-QUEUE
  bandwidth remaining percent 10
class class-default
  bandwidth remaining percent 20
```

VOICE-PQ1

CONTROL-MGMT-QUEUE (weight 10)

MULTIMEDIA-CONF-QUEUE (weight 10)

Class-default (weight 20)

Strict Priority

Traffic in priority queue will always be sent out. Normal queues will be allowed to send traffic only when priority queue is empty.

Scheduler permits class-default to forward traffic twice as often as the other queues

WRR

Queueing Contd.

Policy Map Configuration

```
policy-map 2P6Q3T
class VOICE-PQ1
  priority level 1
class CONTROL-MGMT-QUEUE
  bandwidth remaining percent 10
class MULTIMEDIA-CONFERENCING-QUEUE
  bandwidth remaining percent 10
class class-default
  bandwidth remaining percent 20
```

VOICE-PQ1

CONTROL-MGMT-QUEUE (weight 10)

MULTIMEDIA-CONF-QUEUE (weight 10)

Class-default (weight 20)
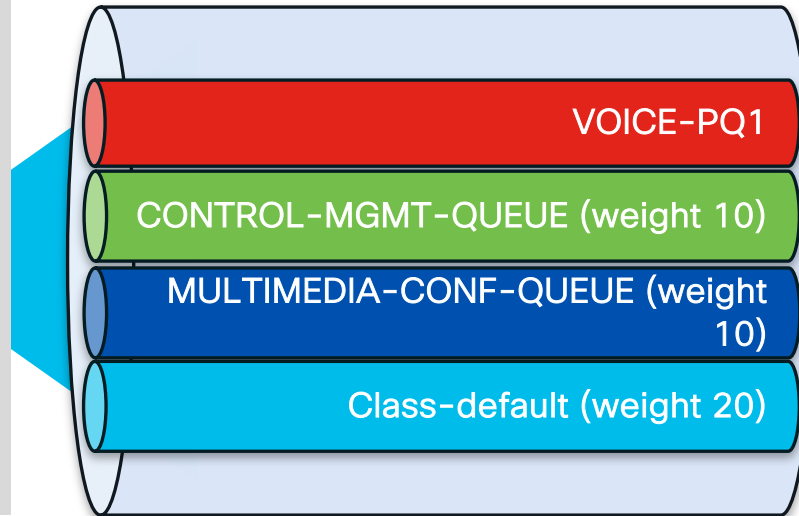
4 classes in the policy-map results in 4 queues created on the interface.

Different rules can be defined for each of the queues

- Queueing classes must match dscp/cos/Ip Prec
- Acl/VLAN matching on egress queueing is not supported
- Match Acl/VLAN on INGRESS and mark to qos-group. Match qos-group on EGRESS

Bandwidth is shared between queues so each queue can at right conditions forward traffic at 100% of link speed

Queueing
Priority queue

```
policy-map 2P6Q3T
  class VOICE-PQ1
    priority level 1
```

- Priority queue.
- As long as there is traffic in priority queue all other queues will not be serviced.

```
policy-map 2P6Q3T
  class VOICE-PQ1
    priority level 1
      police rate percent 10
```

- Priority queue.
- **Limit (police)** traffic to 10 percent of link speed regardless the utilization of other queues

```
policy-map 2P6Q3T
  class VOICE-PQ1
    priority level 1 percent 10
```

- Enable strict priority queue.
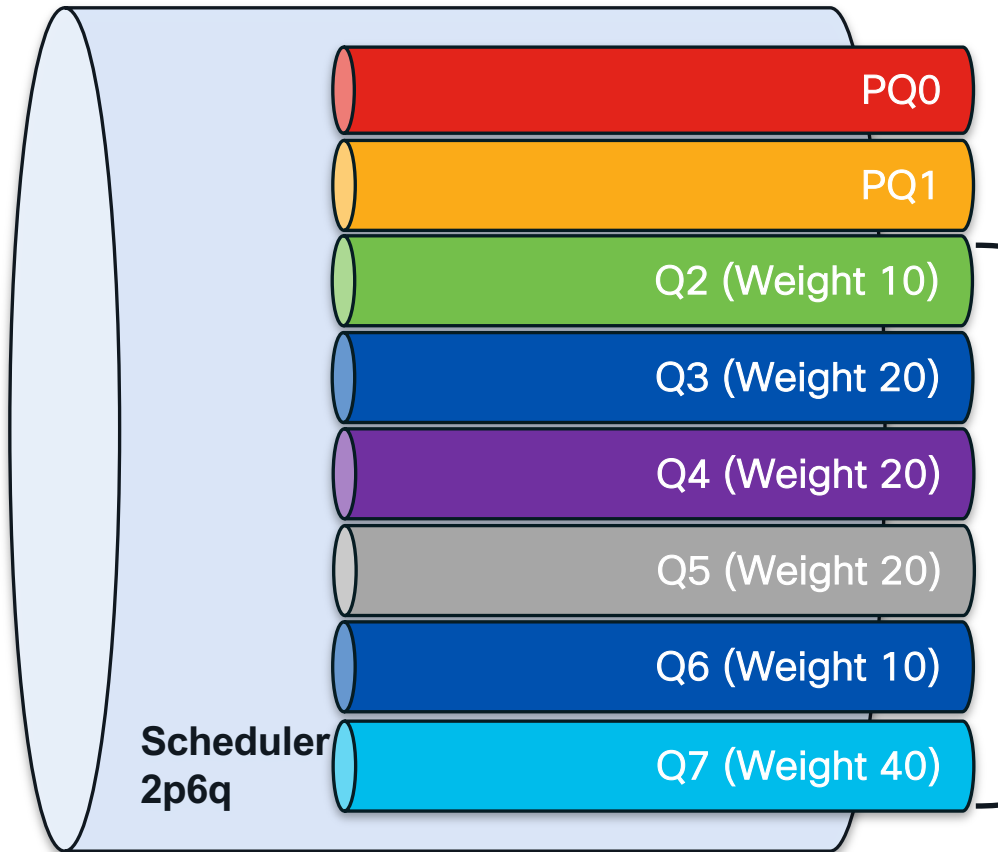- **Shape** traffic to 10 percent of link speed regardless the utilization of other queues

Scheduling - Example
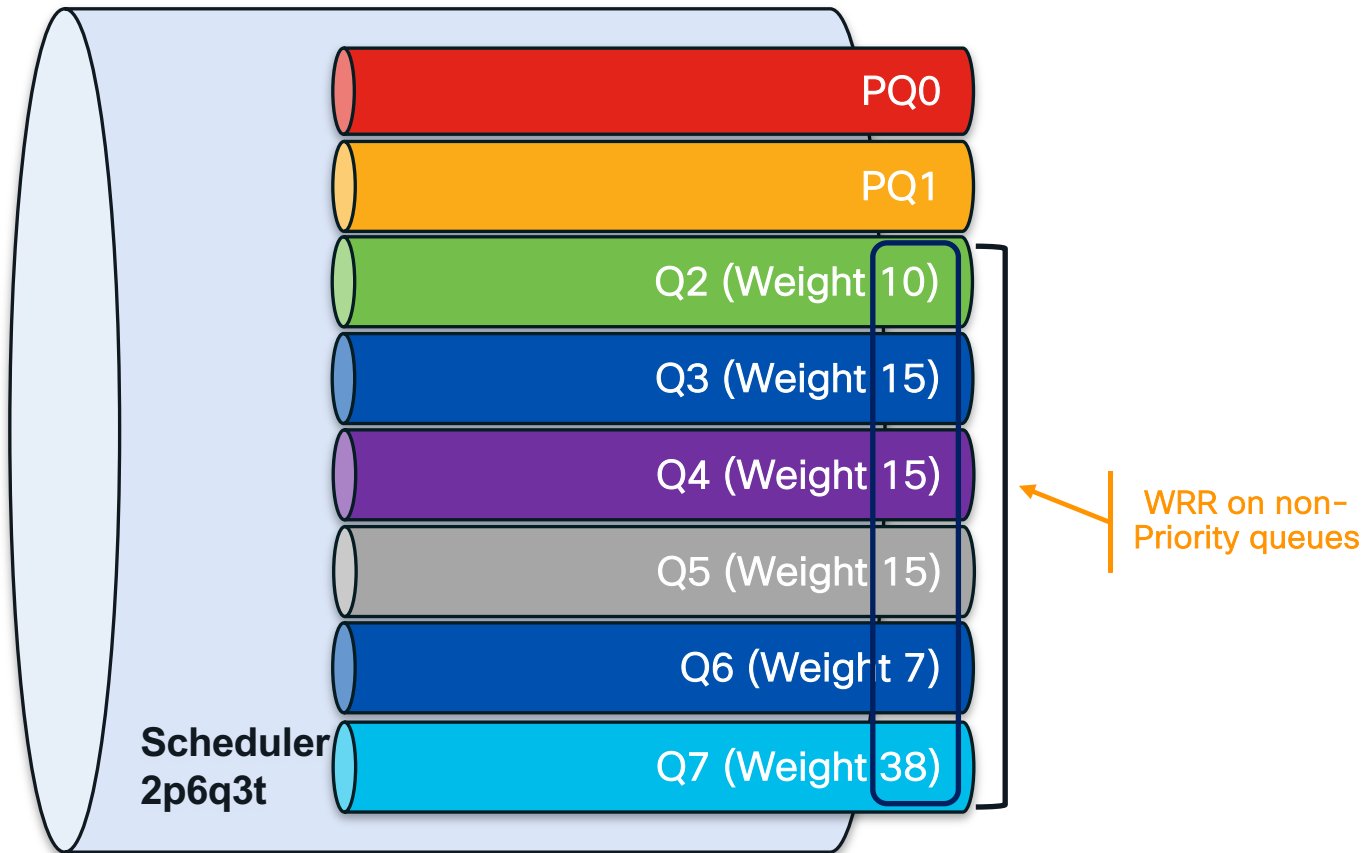
Policy Map Configuration

```
policy-map 2P6Q3T
  class VOICE-PQ1
    priority level 1
  class VIDEO-PQ2
    priority level 2
  class CONTROL-MGMT-QUEUE
    bandwidth remaining percent 10
  class MULTIMEDIA-CONFERENCING-QUEUE
    bandwidth remaining percent 15
  class MULTIMEDIA-STREAMING-QUEUE
    bandwidth remaining percent 15
  class TRANSACTIONAL-DATA-QUEUE
    bandwidth remaining percent 15
  class SCAVENGER-BULK-DATA-QUEUE
    bandwidth remaining percent 7
  class class-default
    bandwidth remaining percent 38
```

Scheduling - UADP

Scheduler 2p6q

Queues (top to bottom):
- PQ0
- PQ1
- Q2 (Weight 10)
- Q3 (Weight 20)
- Q4 (Weight 20)
- Q5 (Weight 20)
- Q6 (Weight 10)
- Q7 (Weight 40)

- Priority Queue is always processed first.
- Priority level 1 has absolute priority.
- Priority level 2 is processed as long as there is no traffic in priority level 1
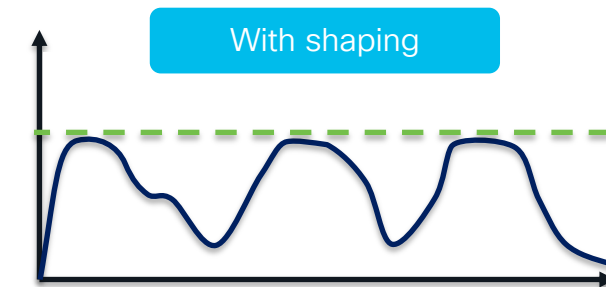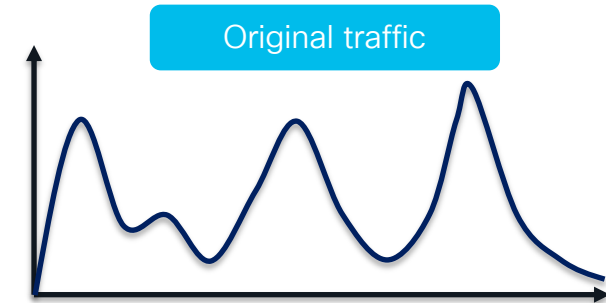
- Normal queues are processed as long as there is no traffic in either the priority queues.
- WRR algorithm instructs the scheduler how often to allow a queue to forward traffic.
- WRR uses weights relative to weights configured on other normal queues.
- In this example, considering same packet sizes, Q2 and Q6 are allowed to forward traffic once for every two time Q3, Q4 and Q5 can forward and every 4 times Q7 can forward traffic.
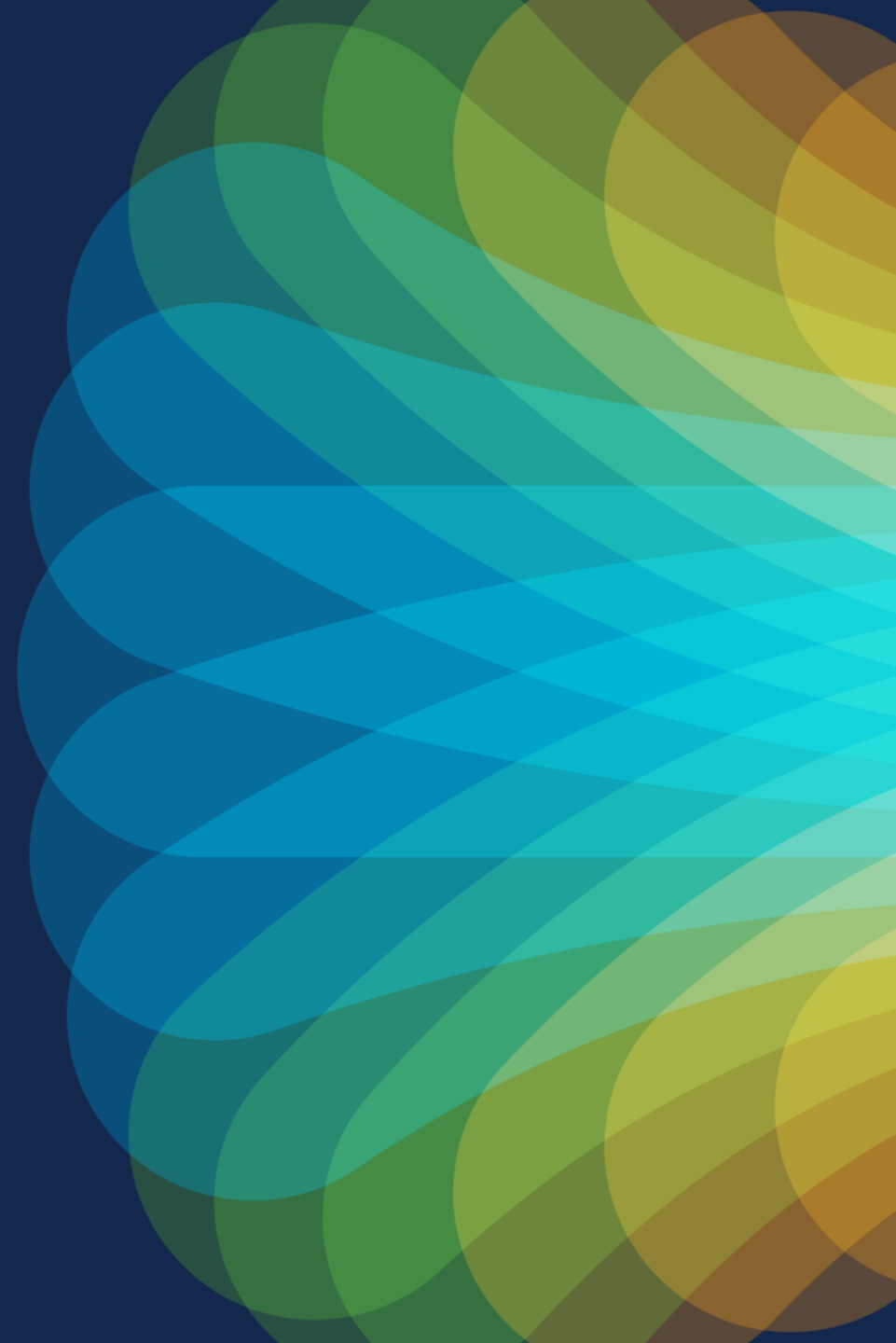
Shaping

- Smooth out traffic peaks, microburst, with preserving traffic

- Control traffic rate to the desired value with buffering.

- Usually in the egress direction

Original traffic

With shaping

Shaping Example

```
policy-map Shaper
    class Transactions
    shape average percent 30
```

UADP
Congestion
Management

Where do we need congestion management?
2P6Q3T Example

VOICE-PQ1

VIDEO-PQ2

CONTROL-MGMT-QUEUE

MULTIMEDIA-CONF-QUEUE

MULTIMEDIA-STREAM-QUEUE

TRANSACTIONAL-DATA

SCAVENGER-BULK-DATA

Class-default

WRR

Priority Queueing

DSCP af11
DSCP af12
DSCP af13

Congestion Management Algorithms

Allows for configuring different priorities for traffic hitting the same queue in the event of congestions
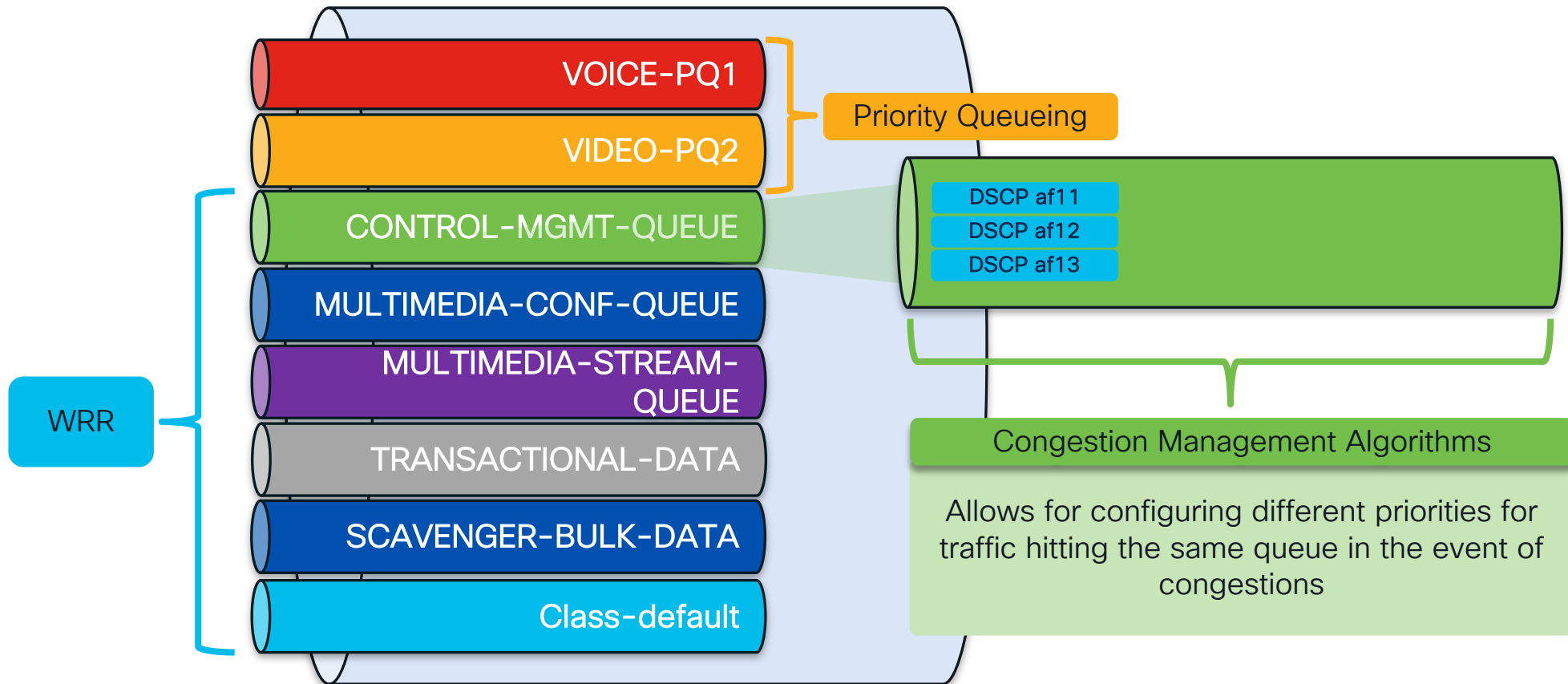
UADP - Congestion Management

Weighted Tail Drop (WTD)

- Default
- For non-priority queues
- Up to 3 thresholds per queue, one threshold per QoS tag
- Each queue need to use same QoS tag type

Weighted Random Early Detection (WRED)

- For non-priority queues
- Up to 4 queues with UADP 2.0X and up to 8 queues with UADP 3.0X
- Up to 3 threshold pairs per queue
- Each queue need to same QoS tag type

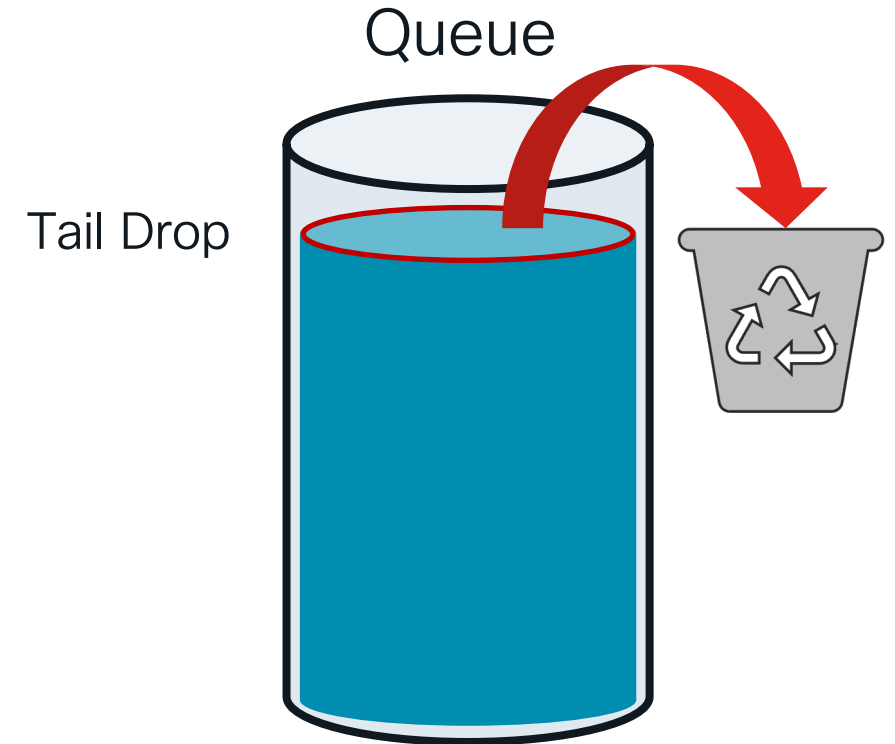**Weighted -> Multiple pair of thresholds**

Congestion Management Tools

- Tail Drop (TD)
  - Drop packets at tail of the queue
  - Single threshold per queue
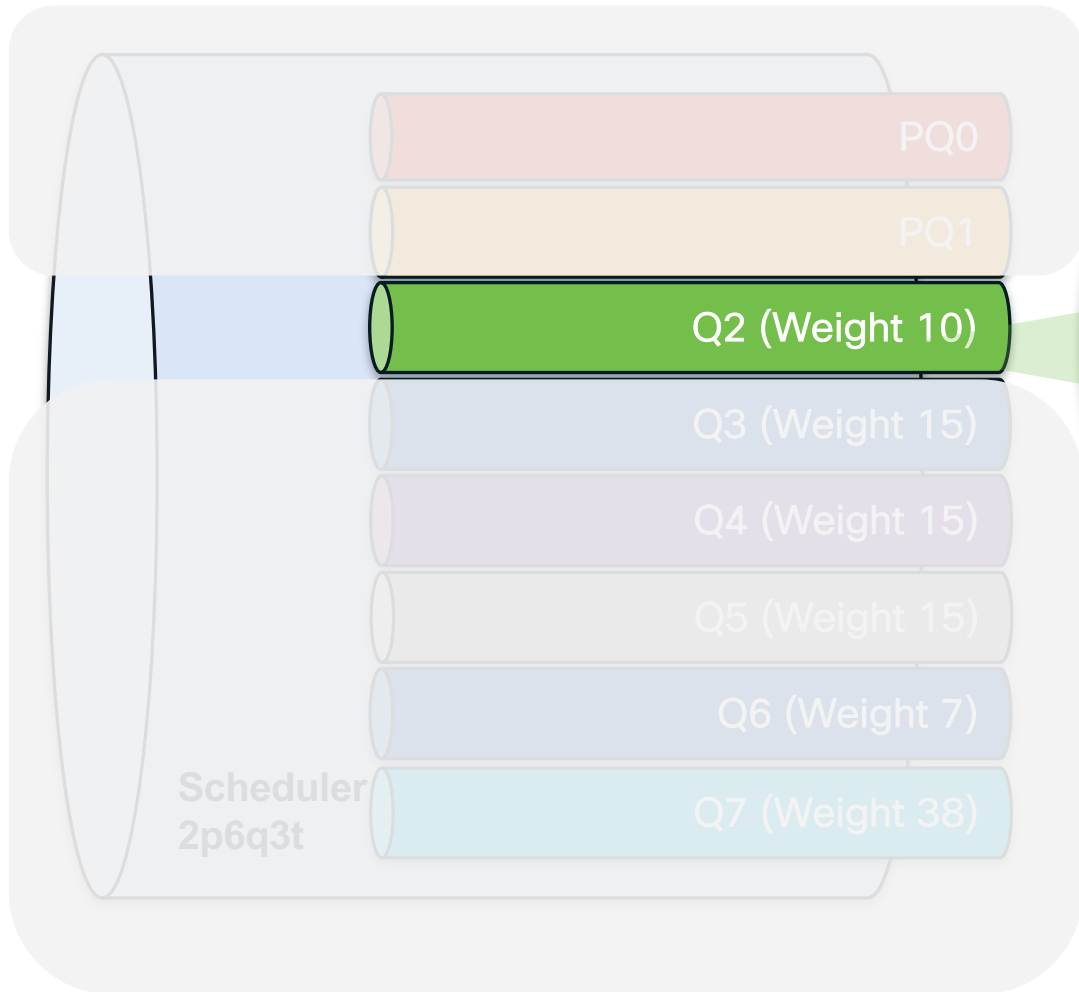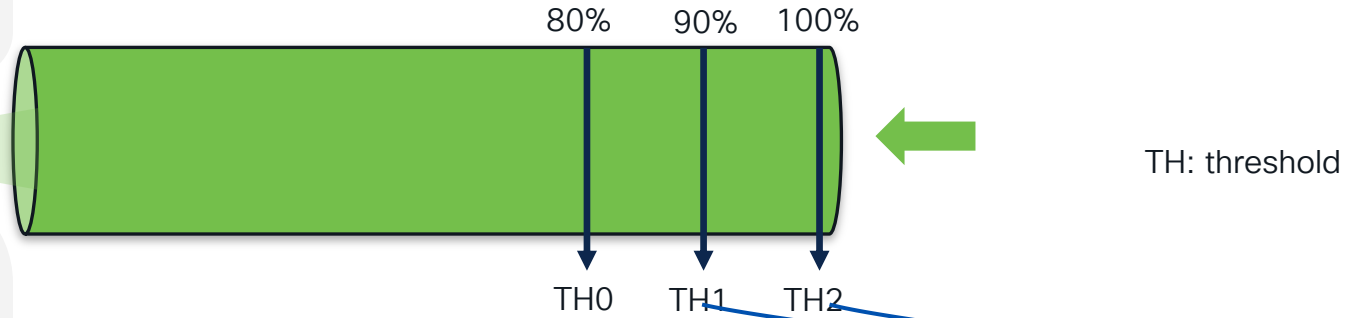
Queue

Tail Drop
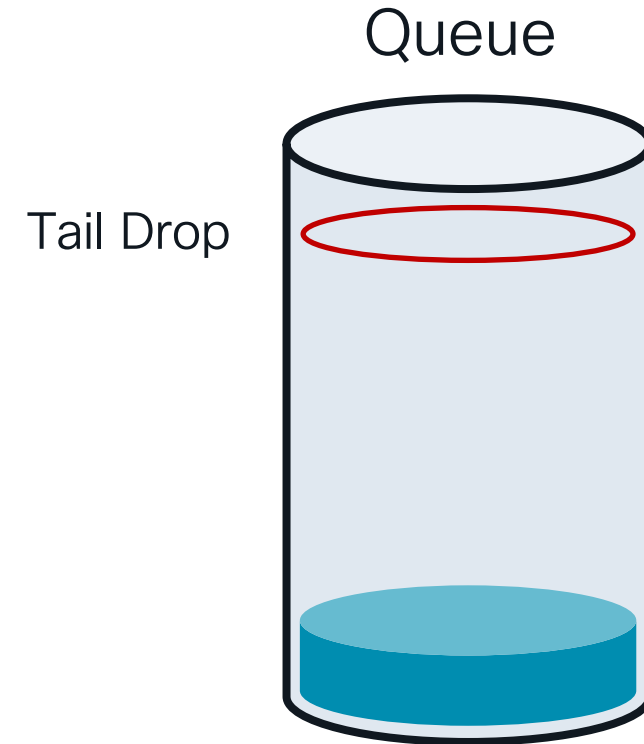
Congestion Management Tools

- Tail Drop (TD)
  - Drop packets at tail of the queue
  - Single threshold per queue

Queue

Tail Drop

WTD – UADP Example

Three thresholds to conditionally drop specific traffic in the event of congestion
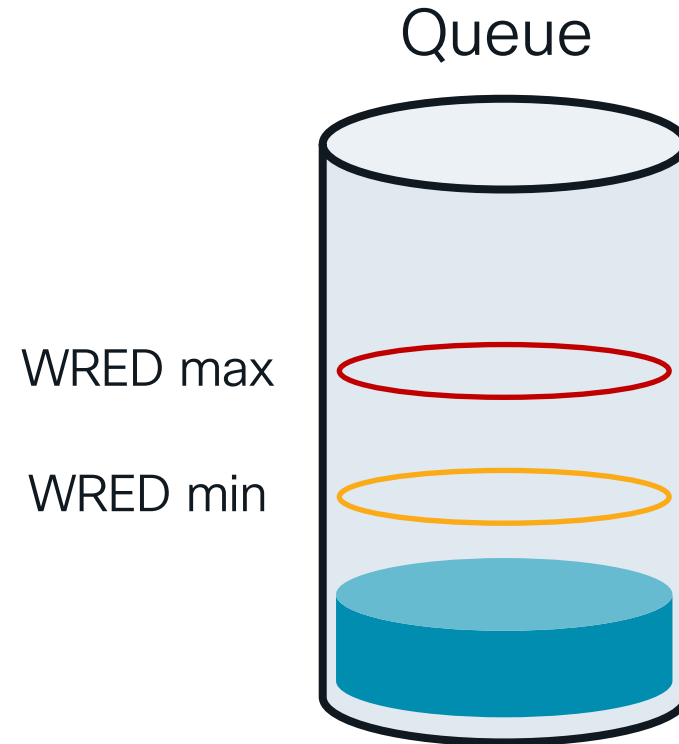
TH: threshold

Policy Map Configuration

```
policy-map 2P6Q3T
 class DATA-QUEUE
   queue-limit dscp values af13 percent 80
   queue-limit dscp values af12 percent 90
   queue-limit dscp values af11 percent 100
```

Congestion Management 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 priority
  - Buffer usage below threshold no affect

Queue

WRED max

WRED min

Congestion Management 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 priority
  - Buffer usage below threshold no affect
  - Buffer usage over min threshold = random drops

Queue

WRED max

WRED min

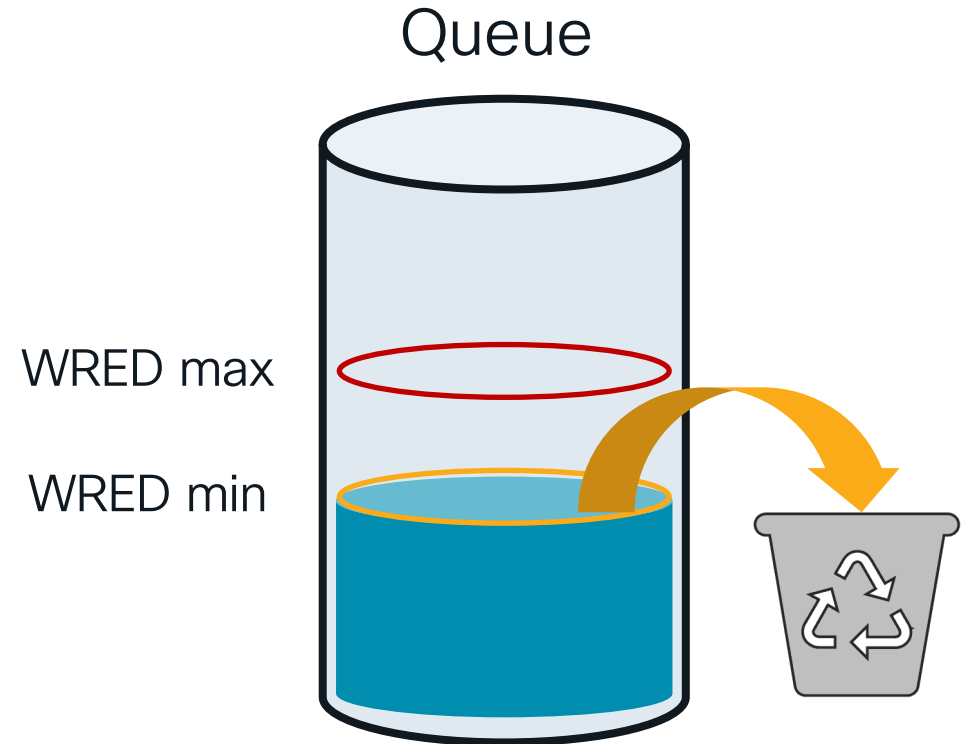# Congestion Management Tools

- 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 priority
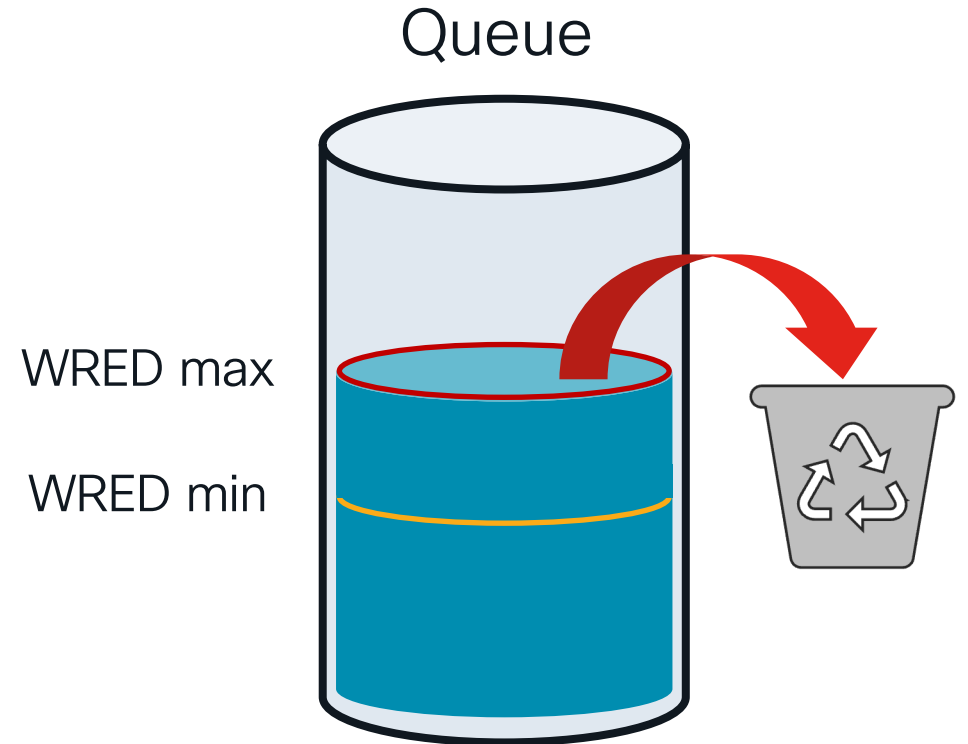 - Buffer usage below threshold no affect
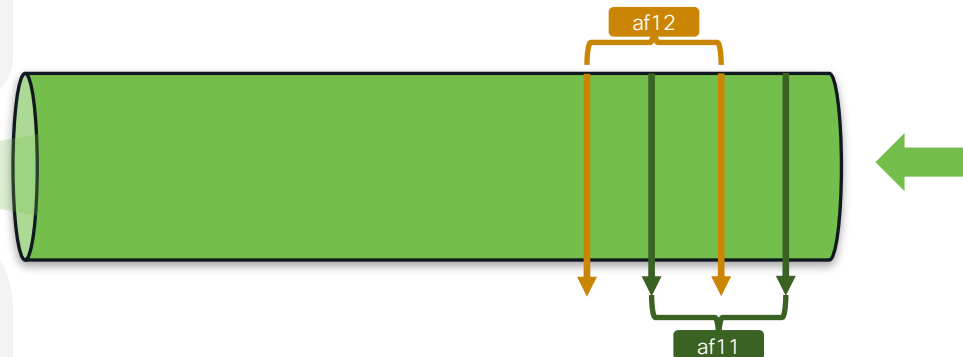 - Buffer usage over min threshold = random drops
  - Buffer usage over max threshold = all traffic drop

Queue

WRED max

WRED min

WRED – UADP Example

- Shown two pairs of WRED thresholds
- UADP supports up to 3 pairs of thresholds

PQ0

PQ1

Q2 (Weight 10)

Q3 (Weight 15)

Q4 (Weight 15)

Q5 (Weight 15)

Q6 (Weight 7)

Scheduler 2p6q3t

Q7 (Weight 38)
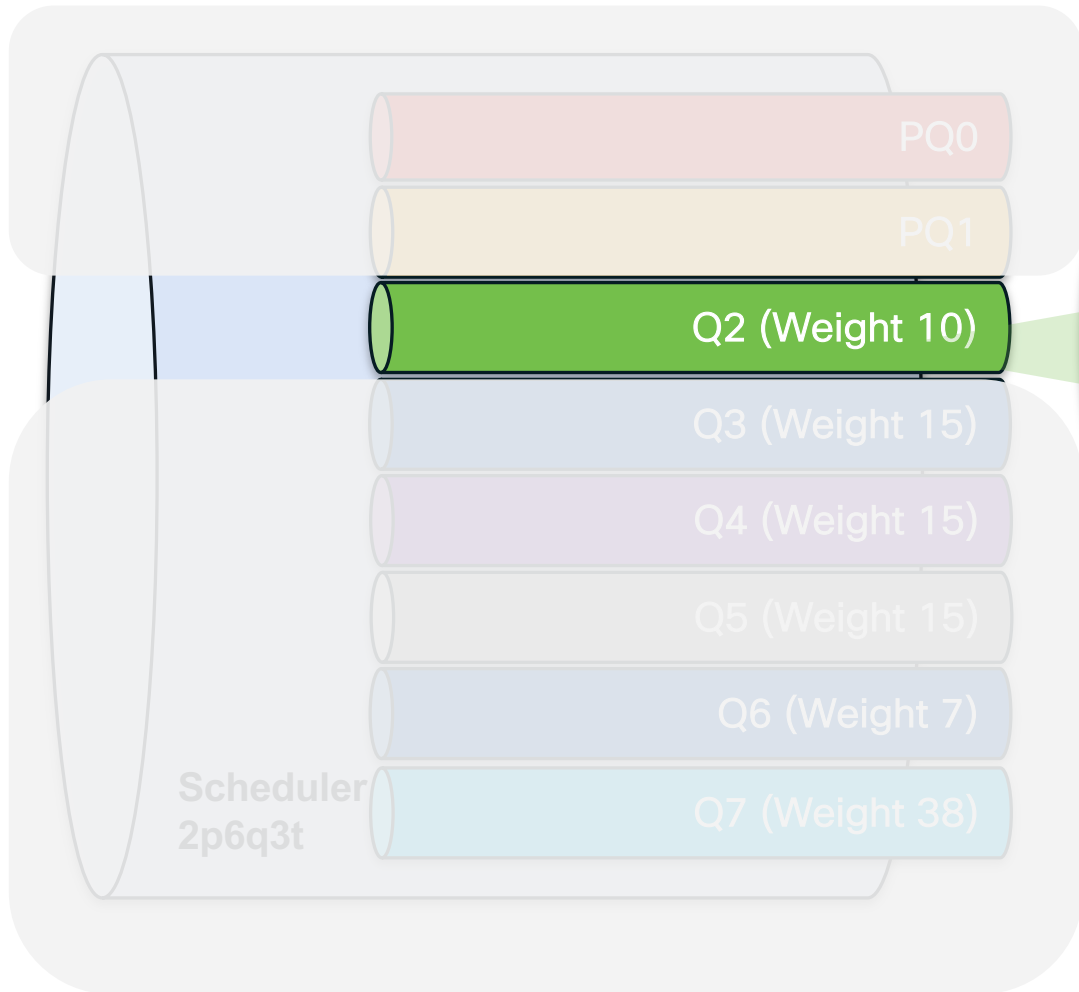
af12

af11

Policy Map Configuration

```
policy-map 2P6Q3T
 class DATA-QUEUE
    bandwidth remaining percent <number>
    queue-buffers ratio <number>
    random-detect dscp-based
    random-detect dscp af11 percent 70 90
    random-detect dscp af12 percent 60 80
```

Buffers

Buffers

- Allocation
  - Dedicated and shared: use dedicated first then shared
  - Dynamic Threshold Scale (DTS): Algorithm to managed the shared buffer

Dedicated

- Allocated to each port on boot.
- Cannot be dynamically changed/edited

Shared

- Dynamically assigned to ports for burst absorption.
- Returned to common pool when not in use.
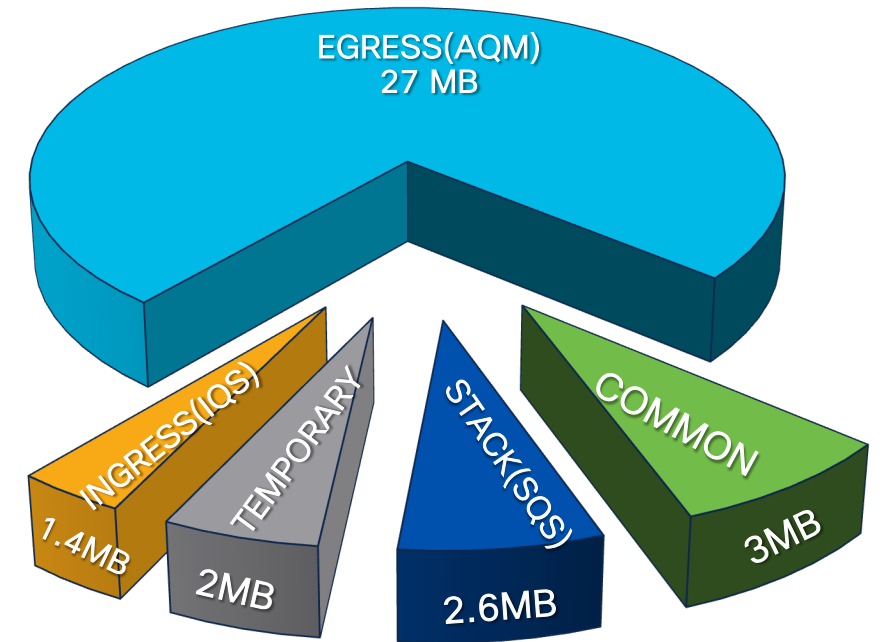
- UADP 3.0 specific
  - Buffer can be shared across two cores
  - "qos share-buffer" to enable the unified buffer

EGRESS(AQM)
27 MB

INGRESS(IQS)
1.4MB

TEMPORARY
2MB

STACK(SQS)
2.6MB

COMMON
3MB

UADP 3.0

Dynamic Threshold Scale (DTS)

Switch

Dynamic Shared Pool (DTS based)

Unused

Port 1

Port 2

Port N
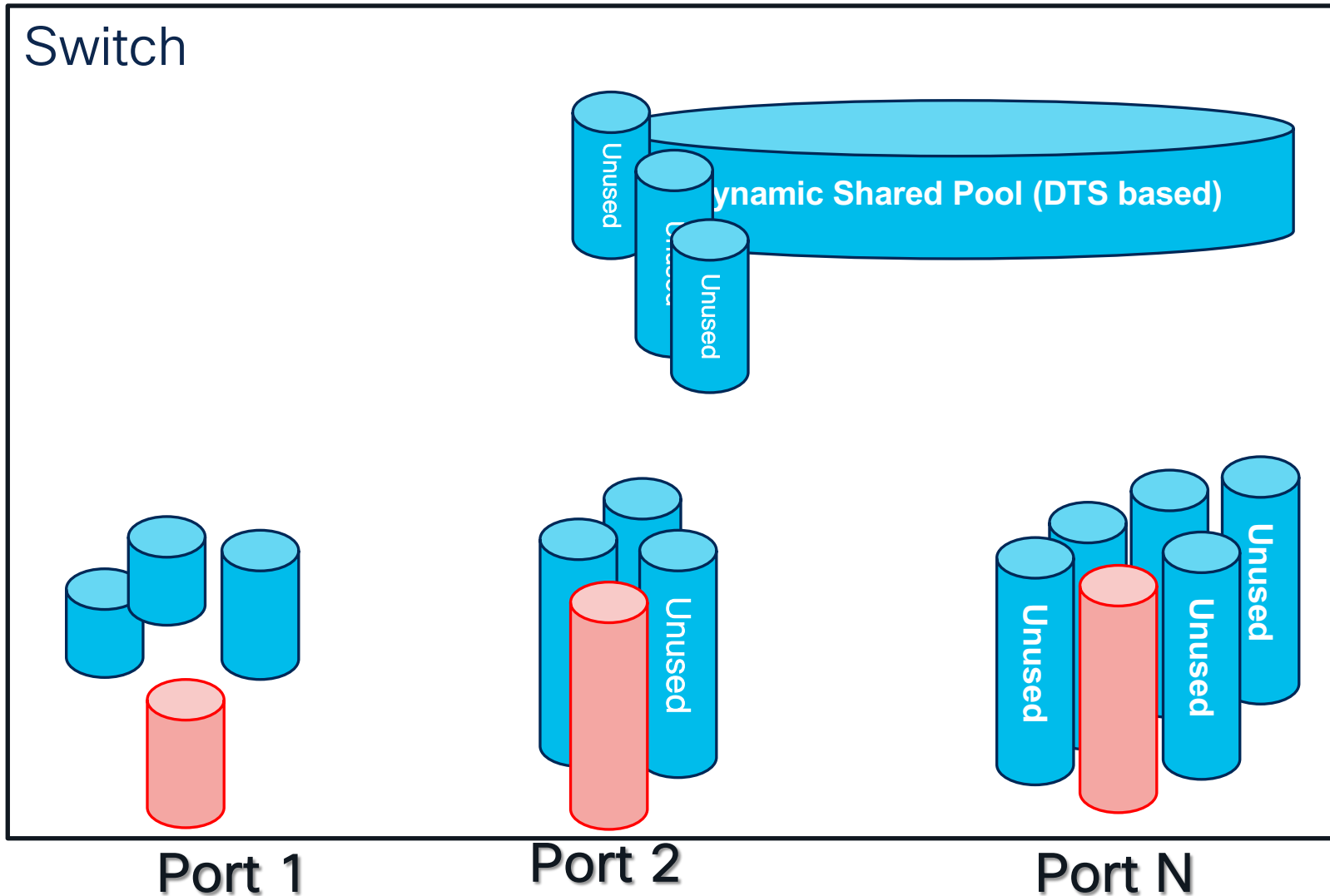
- Shared buffer is good for burst absorption.

- Dedicated buffer is good for predicated performance for each port.

- Buffer management is flexible: Dedicated plus shared.

- Configurable dedicated threshold per port/queue

- Configurable global maximum shared threshold

- Automatically adjusted depends on the available shared pool

The famous softmax multiplier

What does it do?

- `qos queue-softmax-multiplier #`

Switch

Dynamic Shared Pool (DTS based)

Unused

Determines maximum additional buffers each port can request from shared pool

Configuring the softmax multiplier

What value to use?
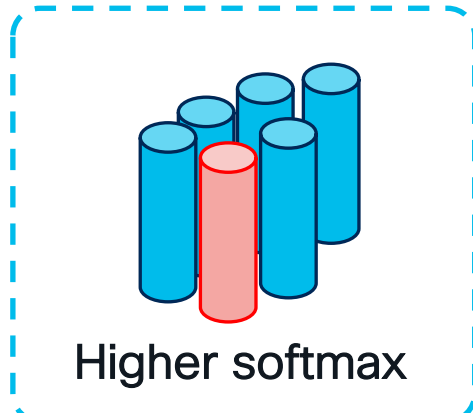
Higher softmax
- Higher burst absorption by individual ports
- Lower concurrent burst spread across ports
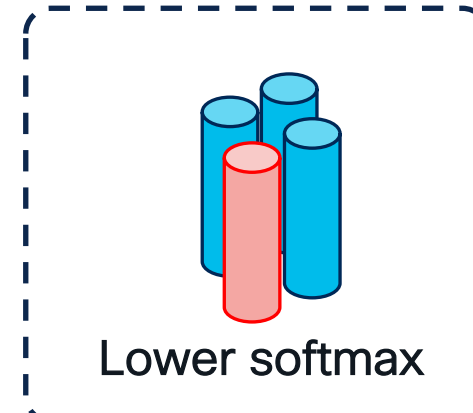
Lower softmax
- Lower burst absorption by individual ports
- Higher concurrent burst spread across ports

| UADP 2.0 | Max value of 1200 |
| UADP 3.0 | Max 1200 by default and up to 4800 with shared buffer |

Unused buffers are returned to the shared pool. No impact from maxing out the softmax multiplier

HQoS
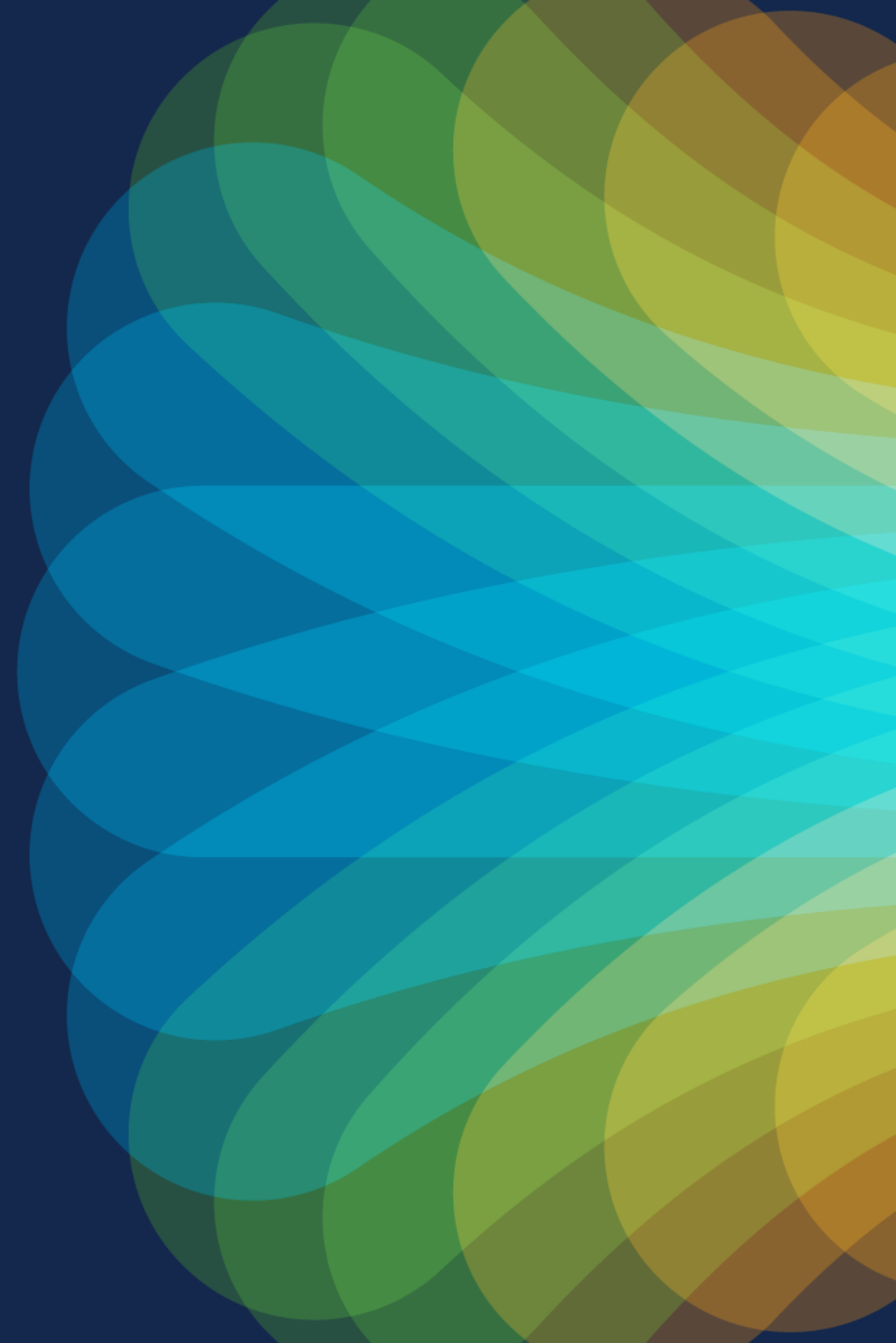
UADP Hierarchical QoS (HQoS)

HQoS (two-level hierarchy) allows you to perform the following functions:

- Classification
- Policing
- Shaping
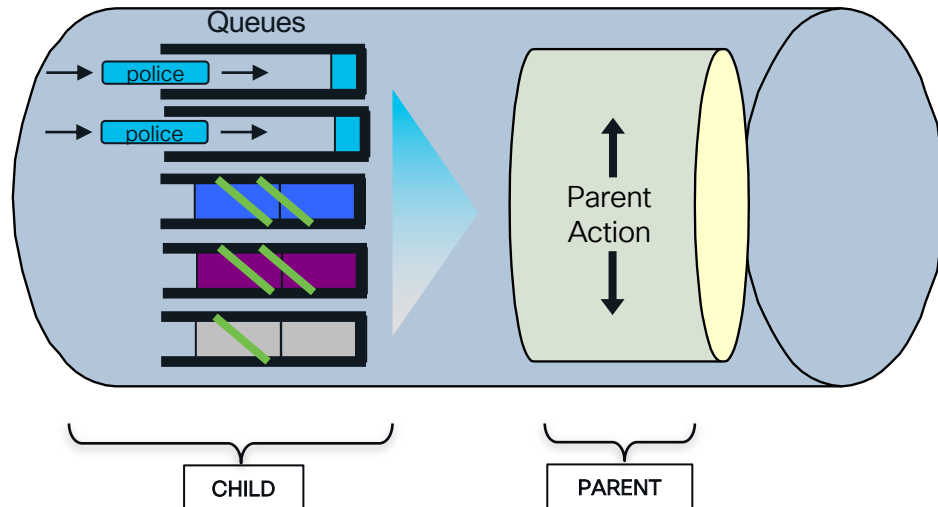
Child Policy	Parent Policy
Classification + Policing	Shaping
	Marking
Classification + Marking	Policing
	Shaping

Silicon One
Q200 QoS (VoQ)

VoQ and Head of Line Blocking

Buffer types – Silicon One vs UADP
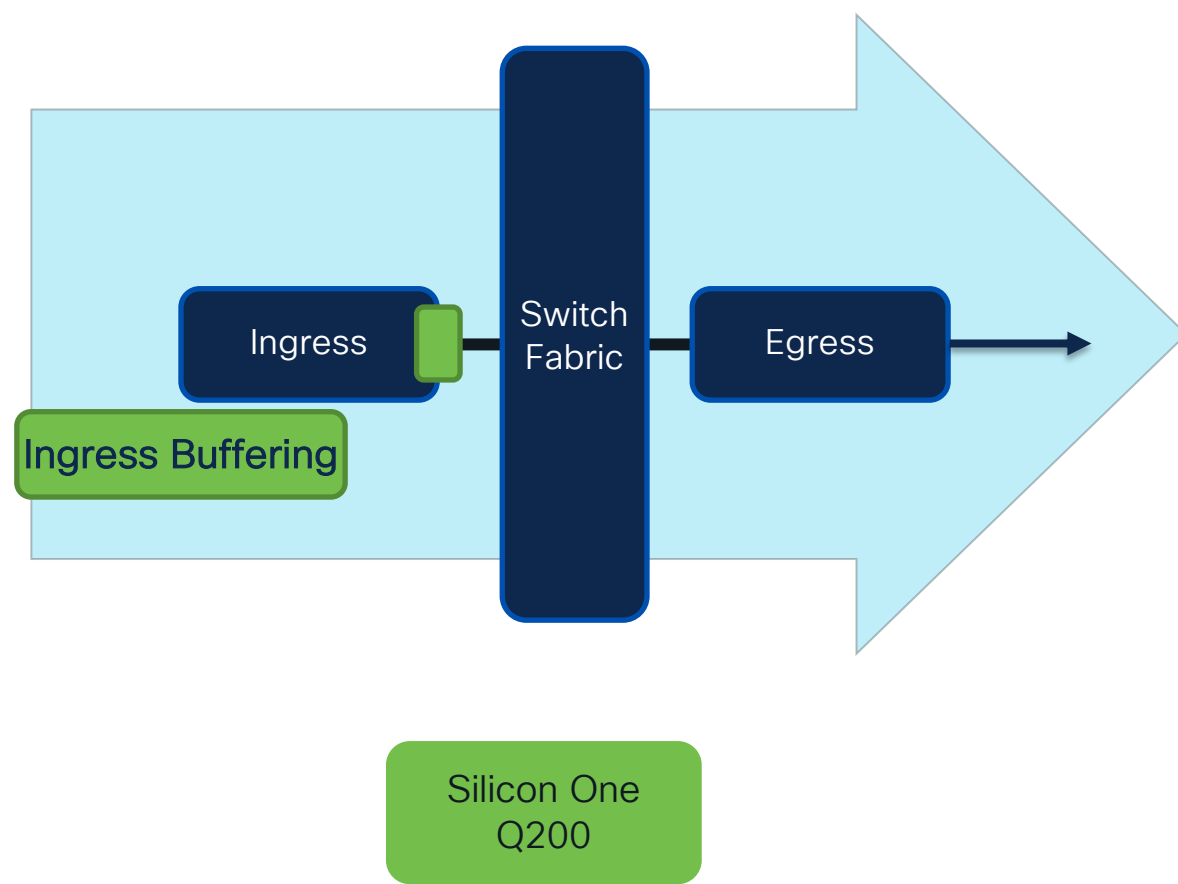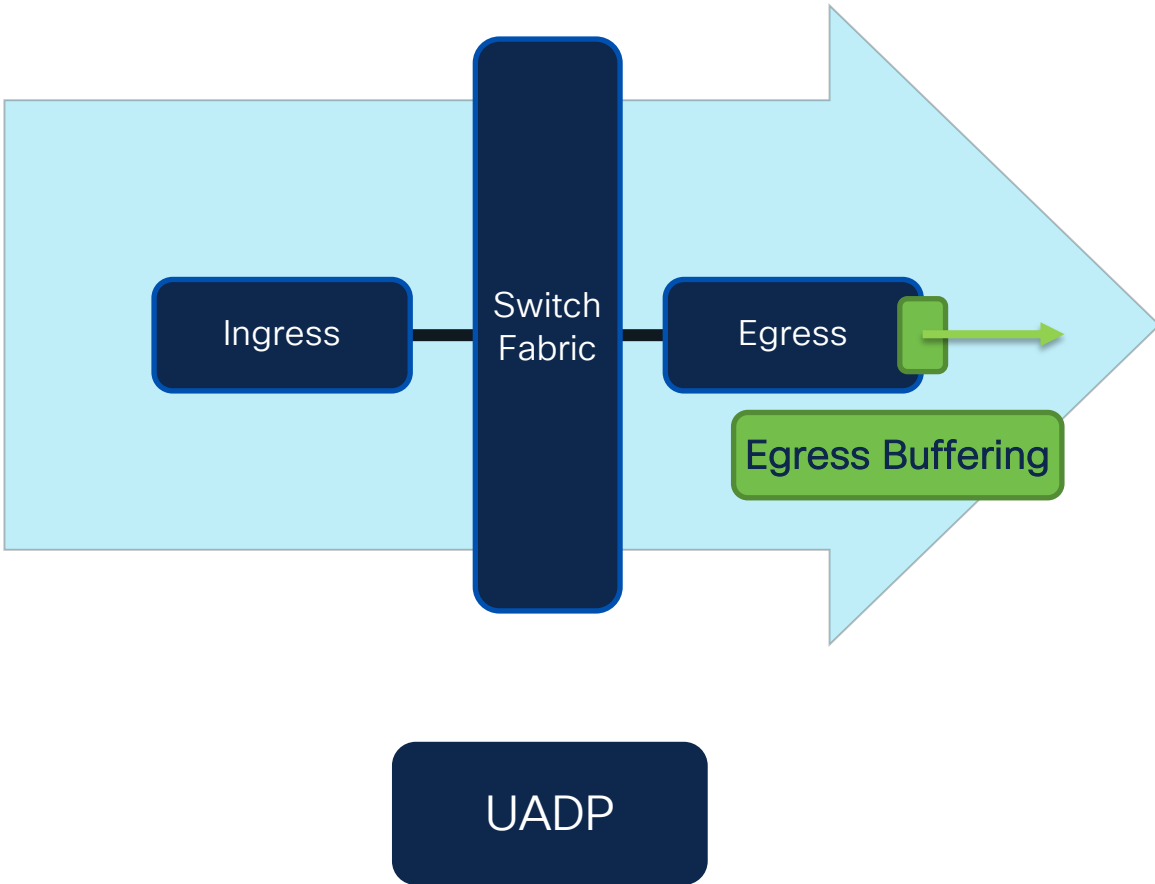
Ingress buffering – Head of Line Blocking
What is the Problem?
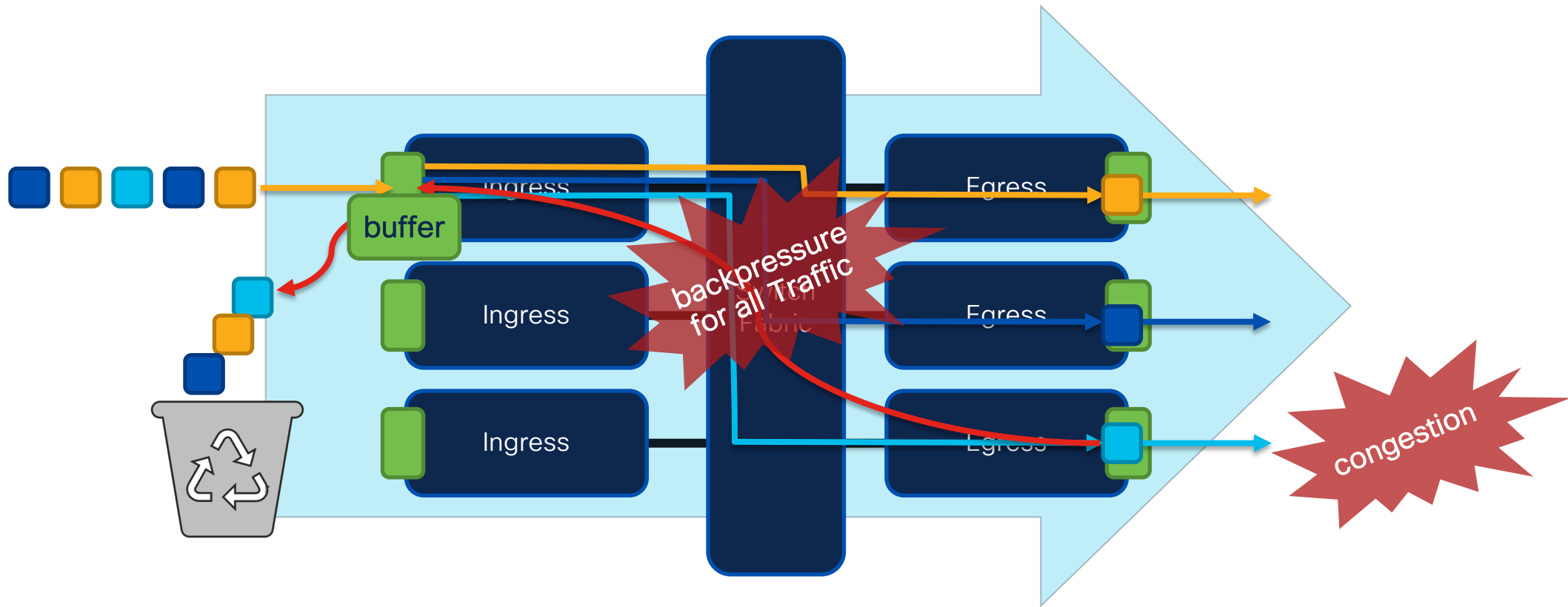
Egress buffering - UADP

No Head of Line Blocking
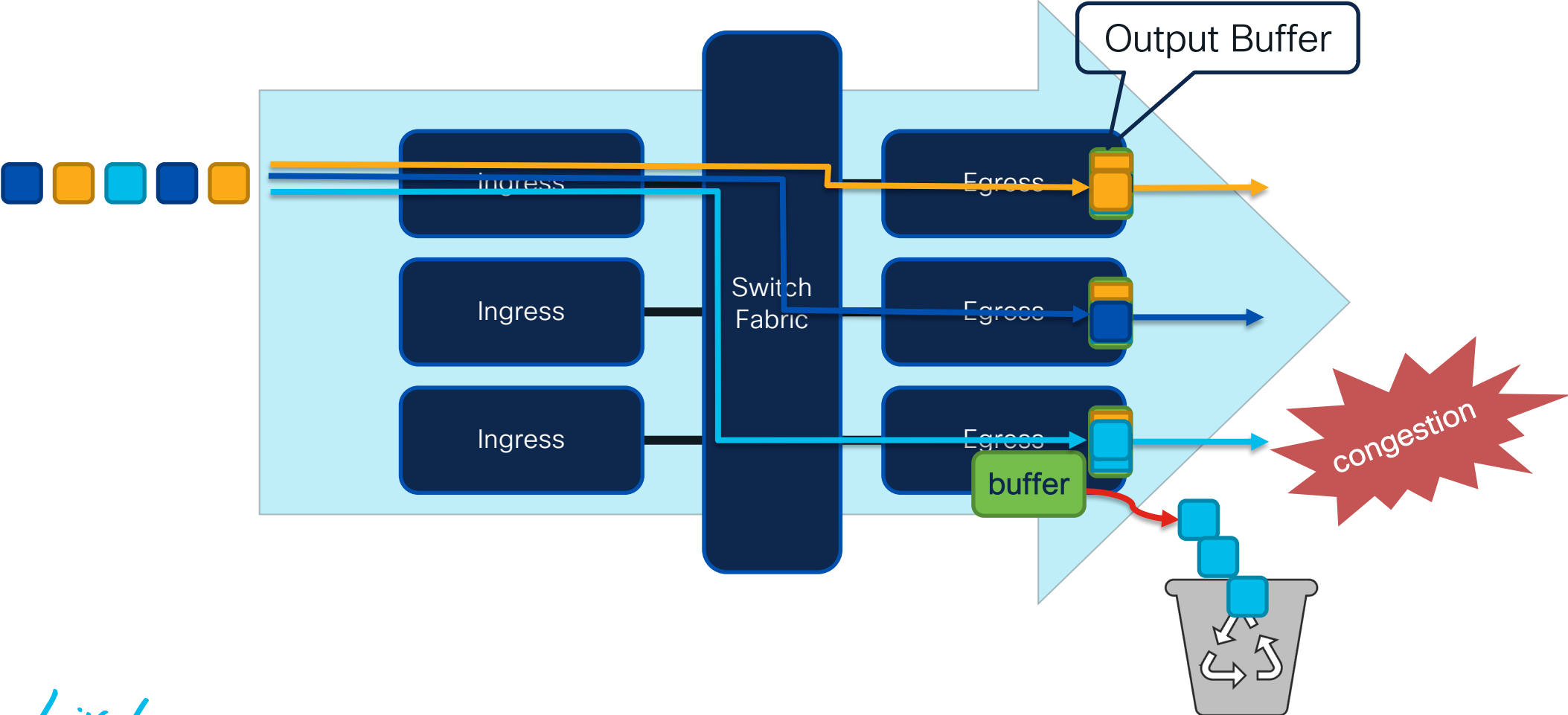
Virtual Output Queuing – Silicon One ASIC

Address Head of Line block

VoQ Buffer

Credit

buffer

Ingress

Ingress

Ingress

Switch Fabric

Egress

Egress

Egress

congestion
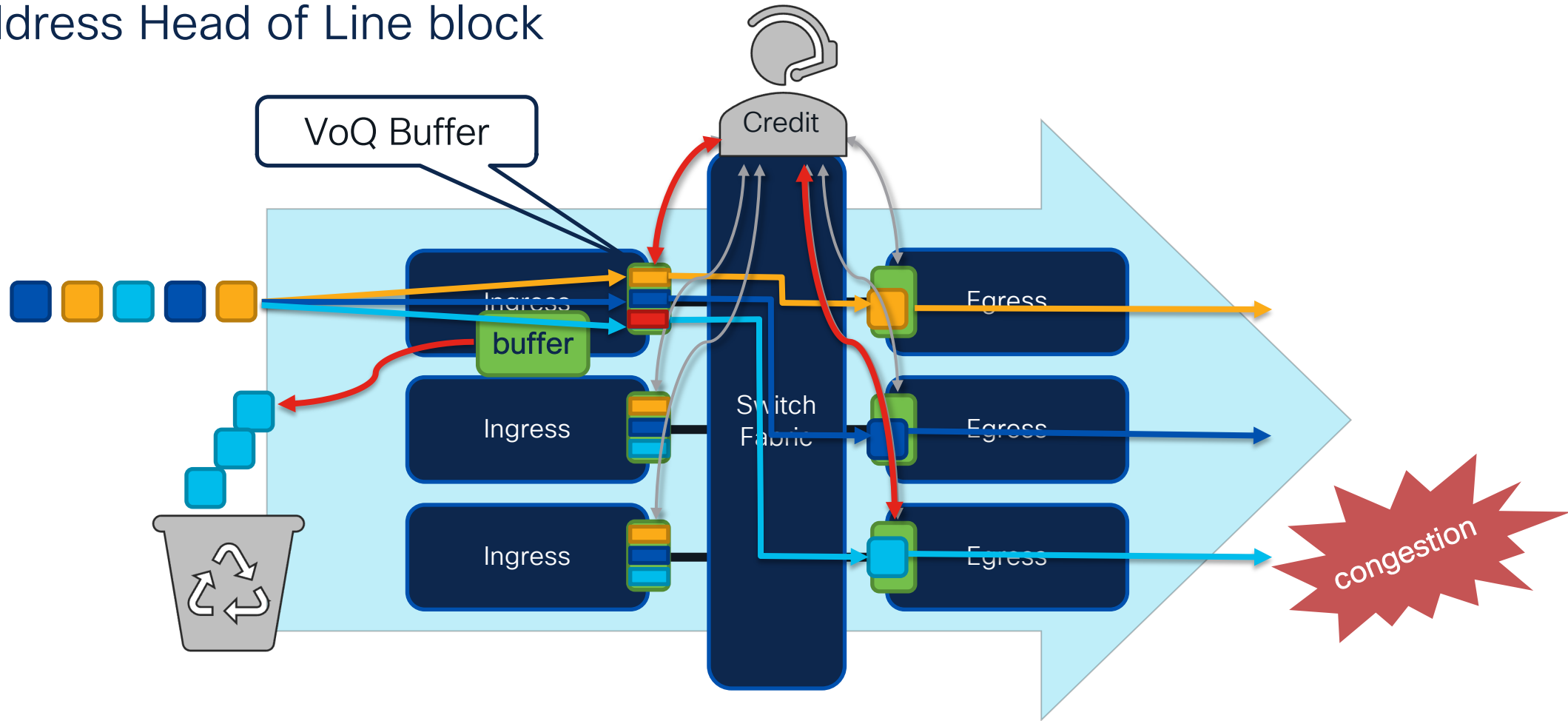
Silicon One Q200 QoS Overview

Silicon One
Features Mapping
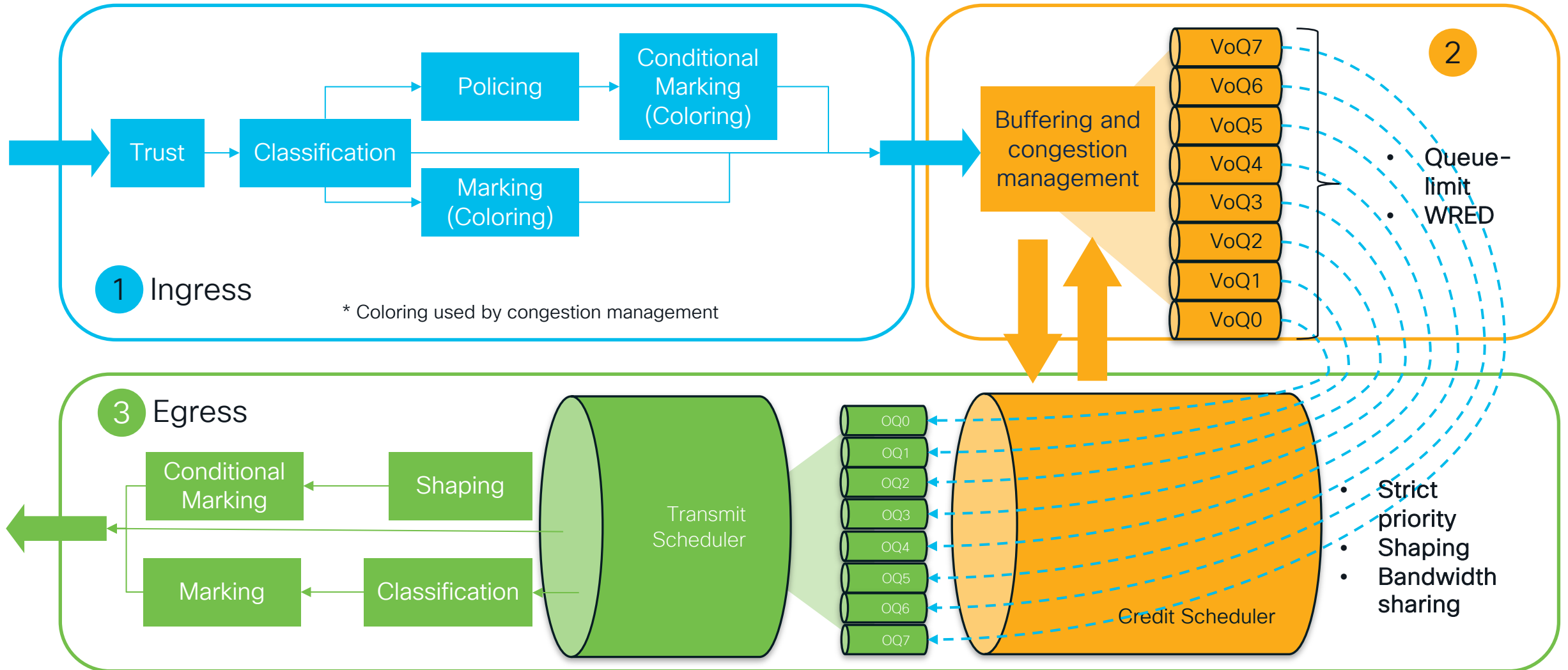
Traffic class

- S1 ASIC uses traffic classes to map traffic to different queues. "traffic-class" is local significant to the switch only

- 3-bit field => 8 values, traffic-class <0 - 7>

- Traffic-class 0 - lowest priority (maps to class-default); traffic-class 7 - highest priority (traffic-class 1 to 6 can be non-priority)

- **Ingress policies** classify packets to specific traffic classes

- Class-maps in **egress queuing policy** can **ONLY** match traffic-class

Priority Level	Traffic class	
1 (statically)	7	Highest priority
2 (optional)	6	
3 (optional)	5	
4 (optional)	4	
5 (optional)	3	
6 (optional)	2	
7 (optional)	1	
Class default	0	Least priority
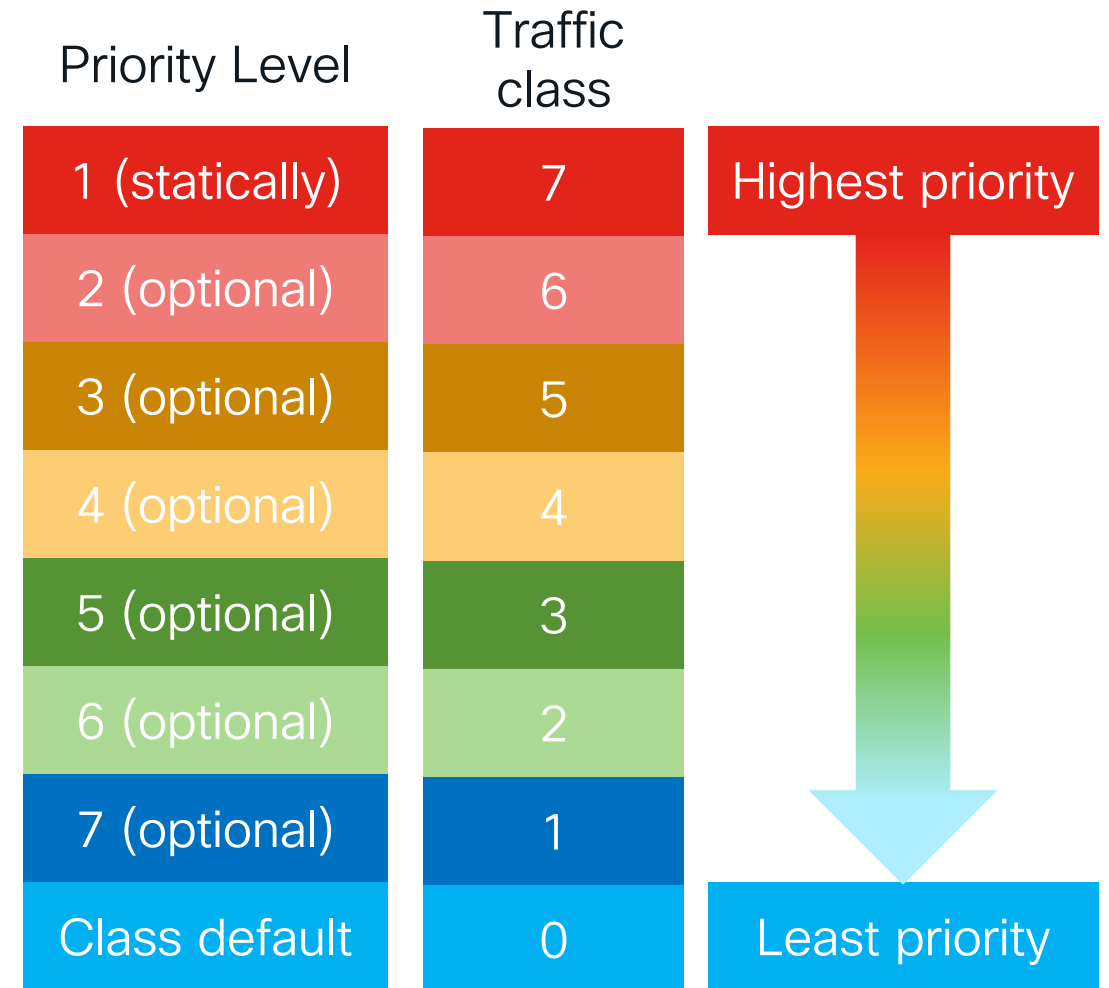
Traffic class to VoQ mapping

Ingress policy

```
policy-map INGRESS
 class REALTIME (matching dscp ef)
    set traffic-class 7
 class VIDEO (matching af41)
    set traffic-class 6
class class-default
```

Mandatory step to classify on ingress to map dscp/cos etc to traffic classes

VoQ	OQ	
VoQ7	OQ7	Always Priority level 1
VoQ6	OQ6	
VoQ5	OQ5	
VoQ4	OQ4	Disabled VoQ and OQ
VoQ3	OQ3	
VoQ2	OQ2	
VoQ1	OQ1	
VoQ0	OQ0	
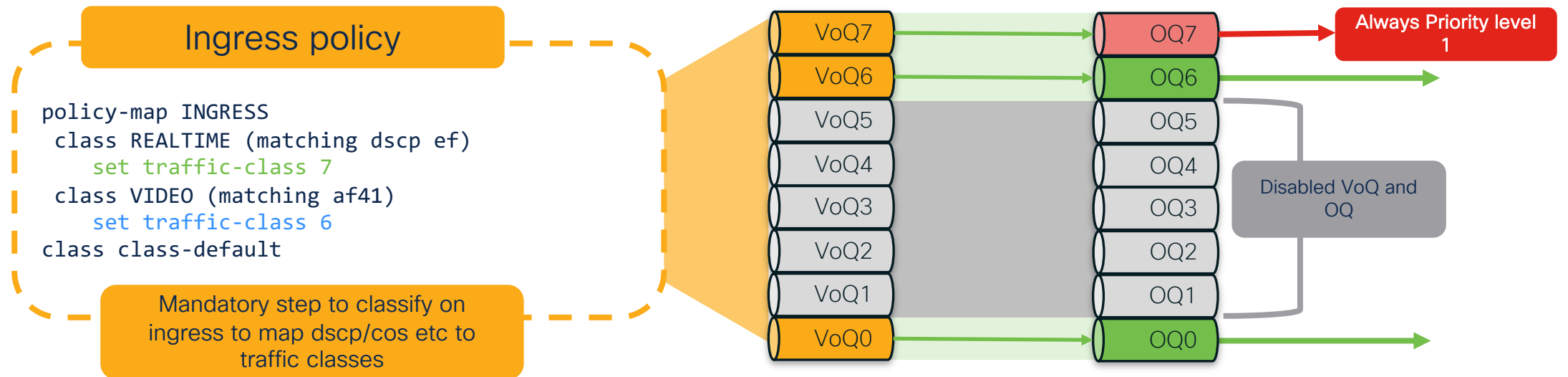
- Ingress policy determines how many VoQ are enabled.
- Each VoQ has a 1:1 mapping with corresponding OQ.
- Traffic class 7 is always a strict priority queue regardless if configured or not.

Silicon One ASIC mapping

1 Ingress Policy

2 Egress Queuing Policy

3 Egress Marking Policy

Ingress Classification, Policing, Marking

Queue Limit, WRED

Egress Shaping, Priority Queueing, Bandwidth/Sharing

Egress Classification, Marking

1 RXPP

2 VoQ

5 OQ

6 TXPP

RXPP

TM Port Transmit Scheduler

4 VSC / VSC / VSC

3 RR ... RR Credit Scheduler

RXPP

Other Egress Slices
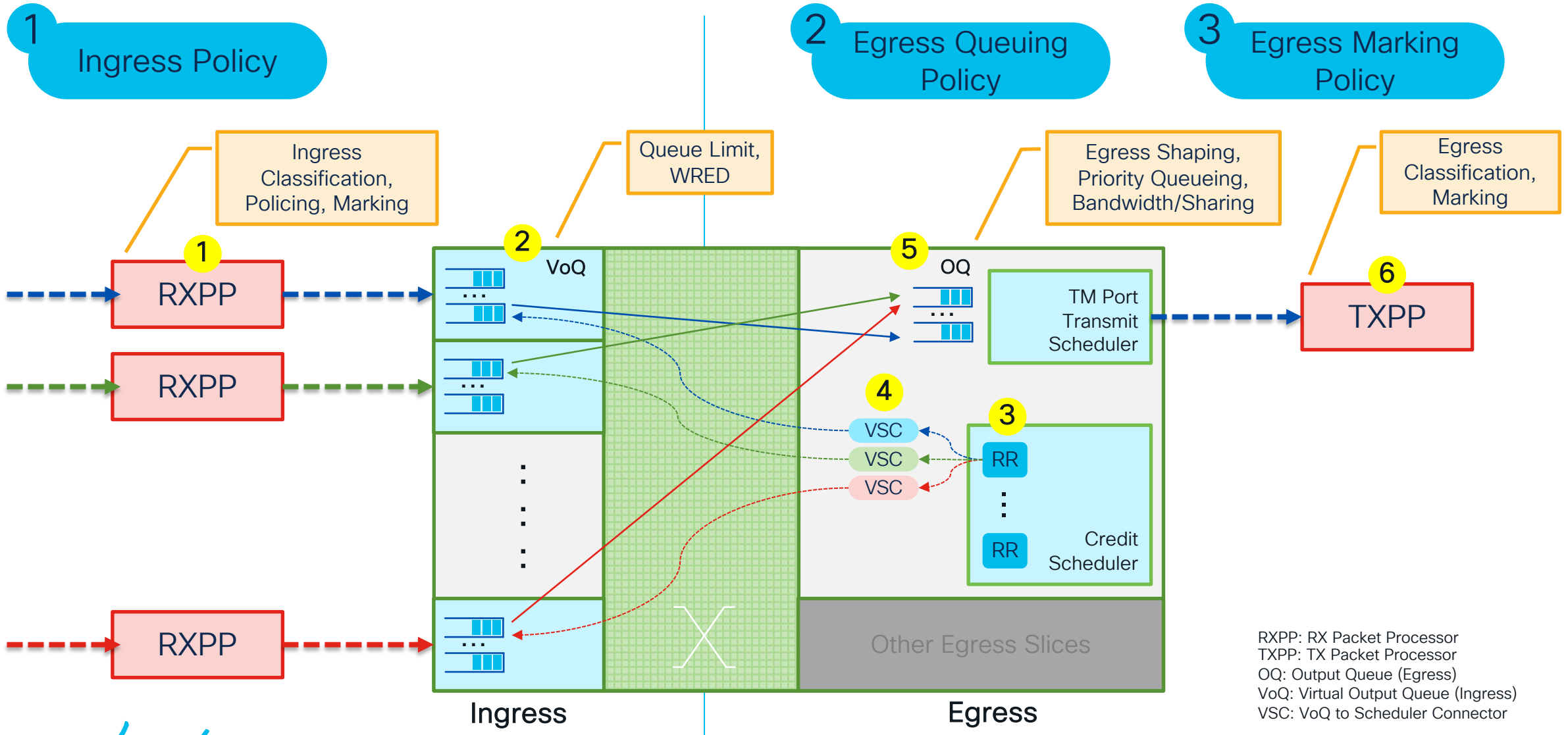
Ingress

Egress

RXPP: RX Packet Processor
TXPP: TX Packet Processor
OQ: Output Queue (Egress)
VoQ: Virtual Output Queue (Ingress)
VSC: VoQ to Scheduler Connector

Traffic color – discard-class

Ingress traffic

discard class

1

0

Traffic marked Yellow has lower priority and will be dropped first in event of congestion.

All traffic dropped green unless specified. Has higher priority then traffic marked Yellow.

Discard class is a 1-bit field
Two possible values – 0 (Green) and 1 (Yellow)
Silicon One Q200 congestion management algorithms <u>ONLY</u> use packet color
Mandatory step if using either WTD or WRED

Silicon One Q200 QoS Default

- QoS enabled
- All ports trust at layer2 and layer3
- Two queues (traffic-class 7 and traffic-class 0, traffic-class7 is priority level 1)
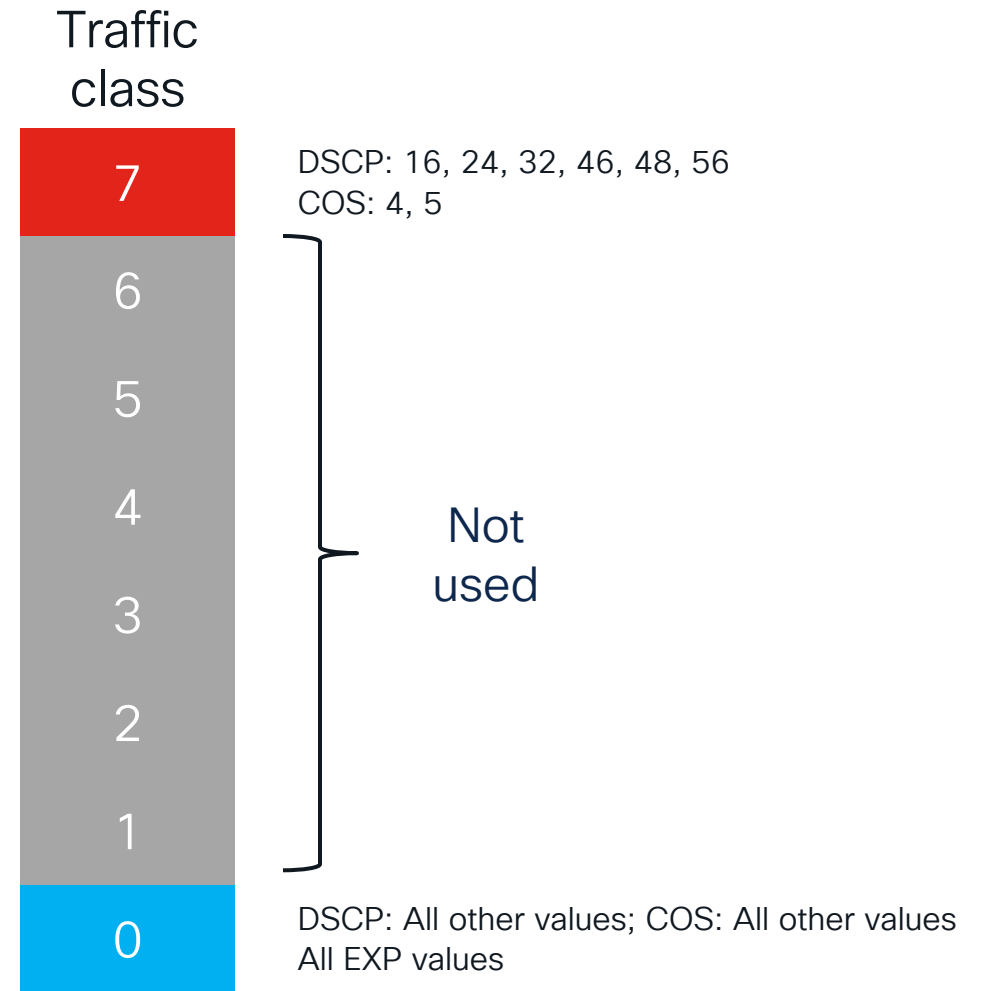
Traffic class

7	DSCP: 16, 24, 32, 46, 48, 56 COS: 4, 5
6	
5	
4	Not used
3	
2	
1	
0	DSCP: All other values; COS: All other values All EXP values

Classification, Marking and Policing

Classification and Marking

- Identify traffic
  - Access Control Lists (ACLs)
  - DSCP
  - IP precedence
  - CoS
  - QoS Group (local with the switch)
  - EXP (MPLS)
  - VLANs

- Marking (coloring)
  - Conditional or unconditional
  - Table map *
  - QoS group (local within switch)
  - Traffic-class (local within switch)
  - Traffic-color (local within switch)

* for policer markdown

Unclassified:
Uncontrolled Bandwidth

Webex
mgmt
HTTP
email
Sharepoint
Backup

Classified:
Controlled Bandwidth

VoIP/Webex
Email/HTTP
SharePoint
Backup

Policing – Limit the traffic

1 rate 2 color

CIR

CBS

Packet of size B

B>Tc — No → Conform → Action

Yes
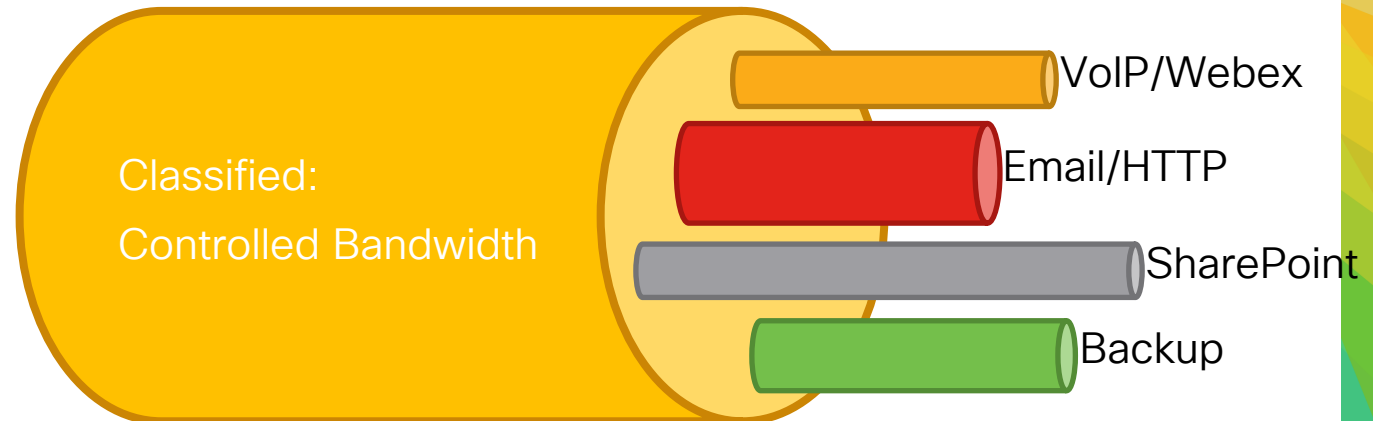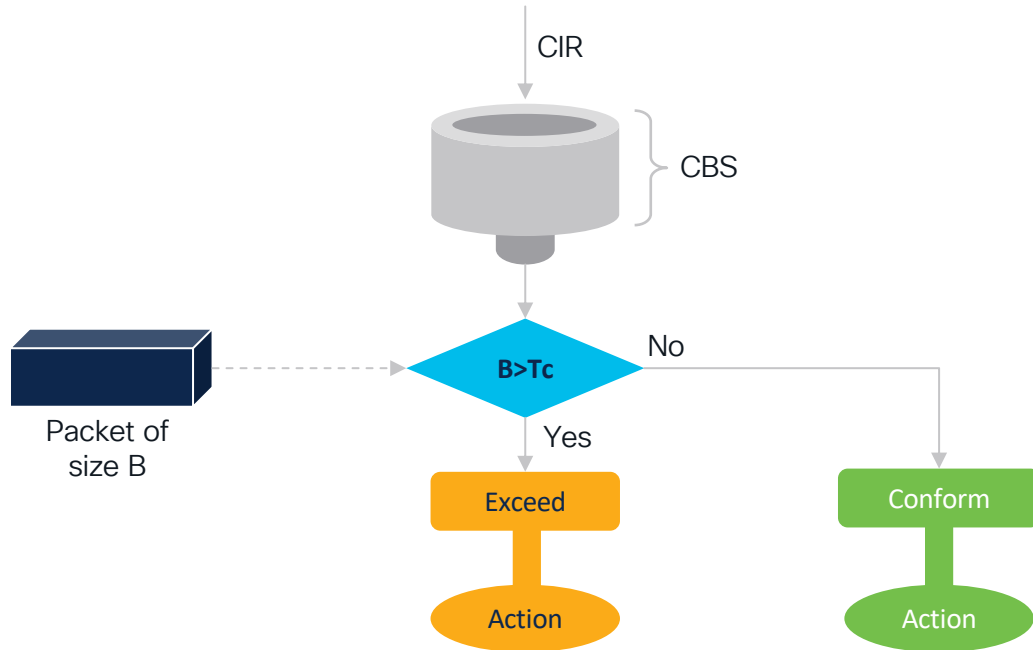
Exceed → Action

```
police cir 1g
    conform-action transmit
    exceed-action drop
```

2 rate 3 color

PIR

CIR

PBS

CBS

Packet of size B

B>Tp — No → B>Tc — No → Conform → Action

Yes   Yes

Violate → Action   Exceed → Action

```
police cir percent 10 pir percent 50
    conform-action transmit
    exceed-action set-discard-class-transmit 1
    violate-action drop
```

Traffic color

CIR – Committed Information Rate
PIR – Peak Information Rate

PBS- Peak Burst Size
CBS – Committed Burst Size

Policing and marking/coloring example

Unconditional Traffic Marking/Coloring

```
policy-map ingress-policy
 class class-5-green
  set traffic-class 5
 class class-5-yellow
  set traffic-class 5
  set discard-class 1
```

> Different class-map
> Same traffic-class

1R2C Policing:

```
policy-map test-police-1R2C
 class dscp1
  set traffic-class 3
  police rate 10g bps
     conform-action transmit
     exceed-action drop
  !
```

Conditional Traffic Marking/Coloring

```
policy-map ingress-policy
 class class-5
   set traffic-class 5
   police rate 5g bps peak-rate 10g bps
    exceed-action set-discard-class-transmit 1
```

2R3C Policing:

```
policy-map test-police-2R3C
 class dscp1
  set traffic-class 3
  police rate 10g bps peak-rate 20g bps
   conform-action transmit
   exceed-action set-discard-class-transmit 1
   violate-action drop
```
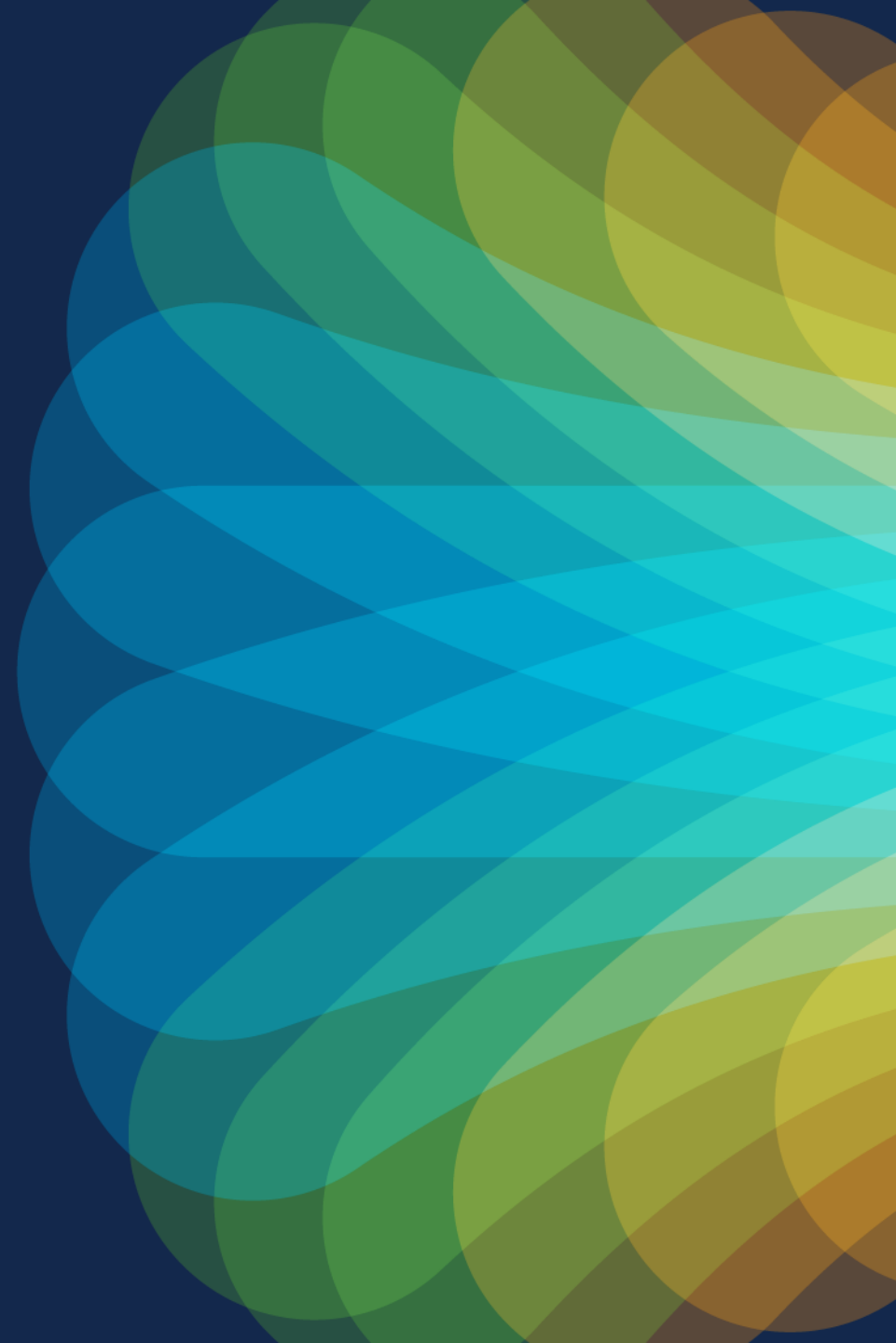
Queueing, Scheduling and Shaping

Queueing

Virtual output Queue (VoQ)
- 8 VoQ on each ingress slices for each interface
- Each traffic-class maps to a VoQ (multiple traffic-classes can map to same VoQ)

- VoQ maps to output Queue.

Output Queue
- 8 output queues (egress) for each interface
- Up to 7 strict priorities (level 1 – highest)
- Traffic-class 7 is always priority level 1, priority level is optional for other traffic-classes

Buffering and congestion management

VoQ7
VoQ6
VoQ5
VoQ4
VoQ3
VoQ2
VoQ1
VoQ0

2

- Queue-limit
- WRED

OQ0
OQ1
OQ2
OQ3
OQ4
OQ5
OQ6
OQ7

Credit Scheduler

- Strict priority
- Shaping
- Bandwidth sharing
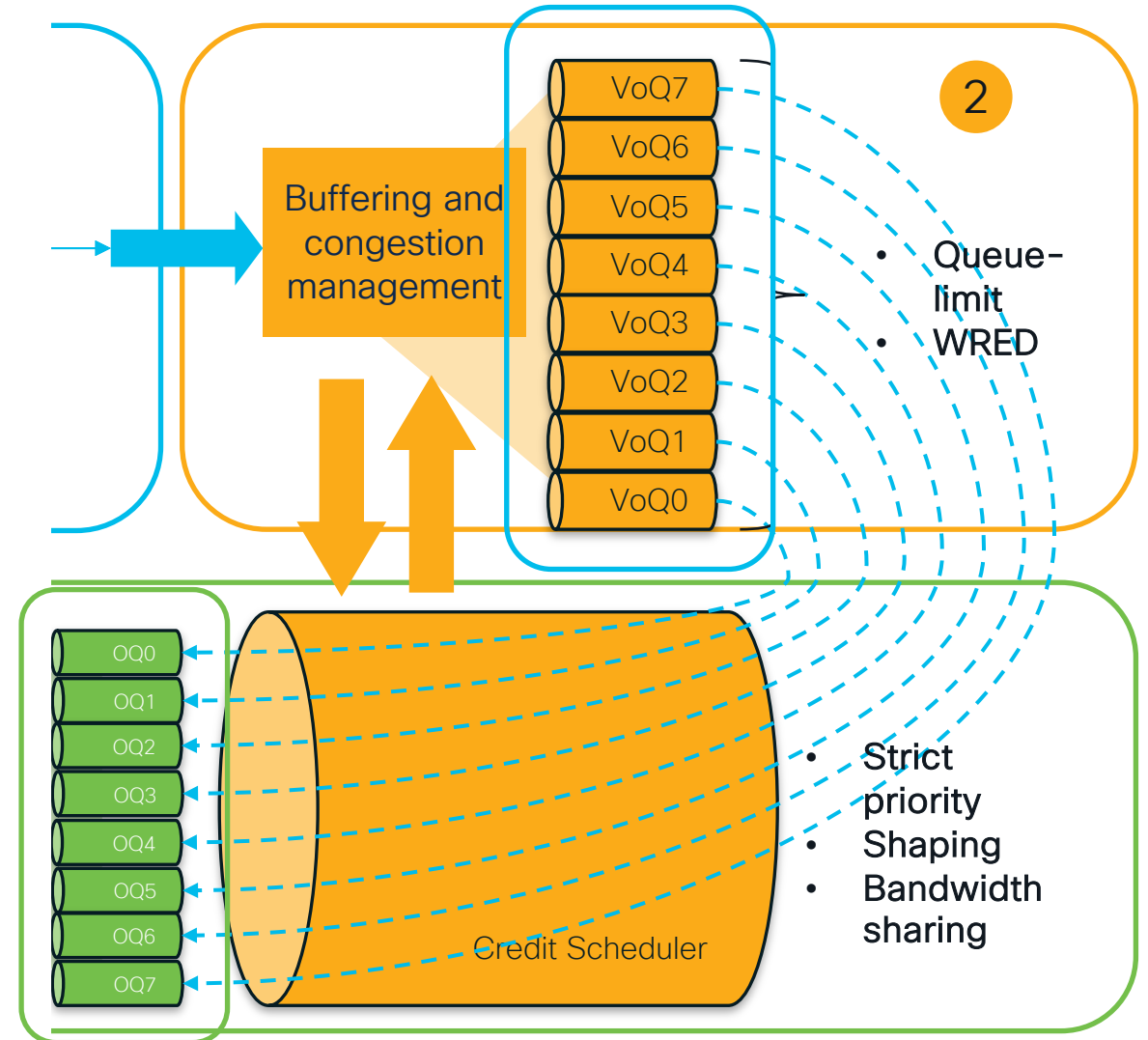
Scheduling

- Packet schedule from VoQ to OQ based on a credit scheduling system
- Packets are buffered at ingress (VoQ)
- Different type of queues are served differently
  - Strict priority queues
    - Always serviced first
    - Up to 7 PQs
  - Normal queues (without priority configured)
    - Served only after priority queues are empty
    - Use Weighted Round Robin (WRR) for scheduling

2

Buffering and congestion management

VoQ7
VoQ6
VoQ5
VoQ4
VoQ3
VoQ2
VoQ1
VoQ0

- Queue-limit
- WRED

OQ0
OQ1
OQ2
OQ3
OQ4
OQ5
OQ6
OQ7

Credit Scheduler

- Strict priority
- Shaping
- Bandwidth sharing
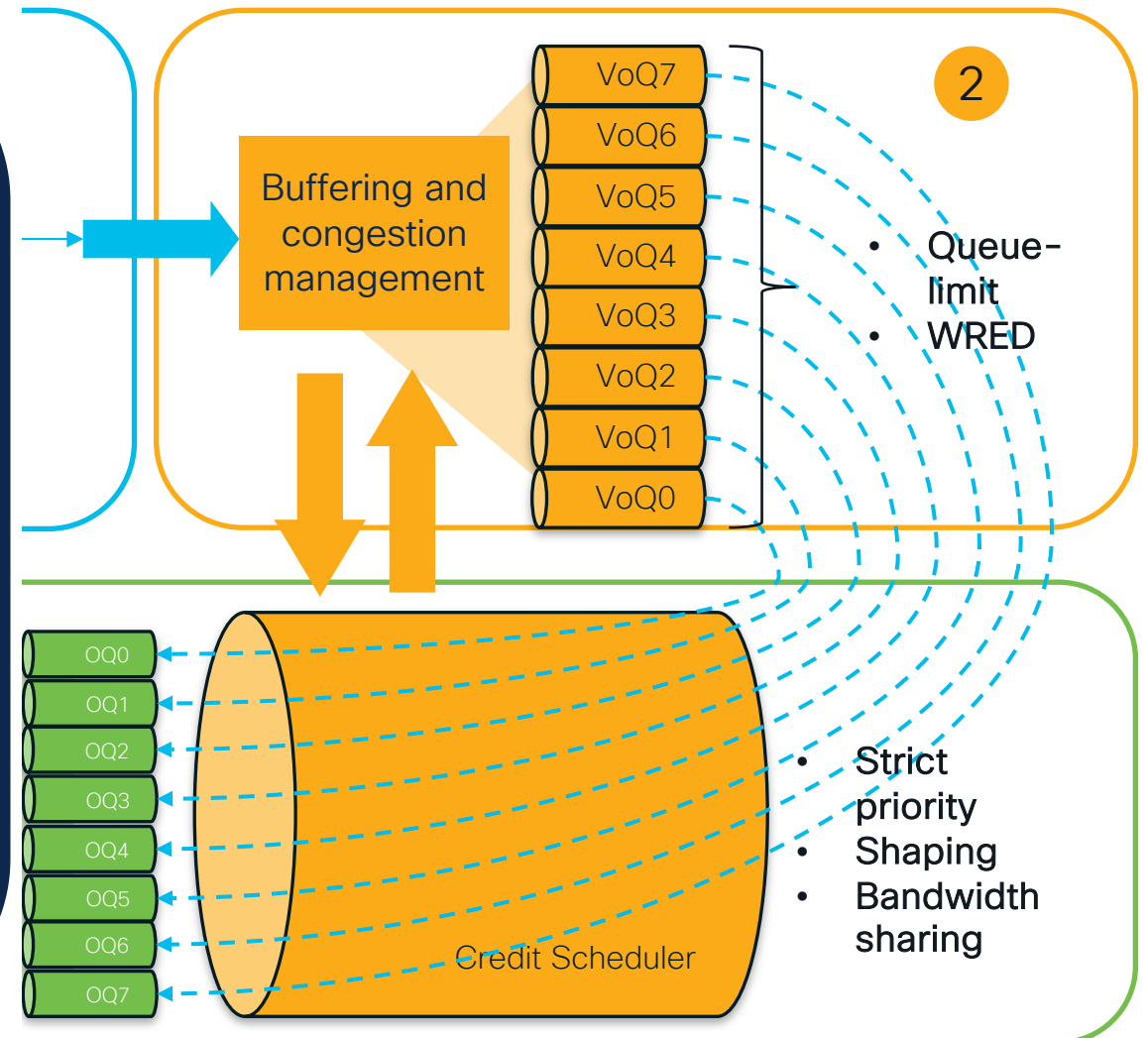
Scheduling - Example

```
class-map match-any tc-7
 match traffic-class 7
class-map match-any tc-6
 match traffic-class 6
…
class-map match-any tc-1
 match traffic-class 1
```

Map traffic to the queues

```
policy-map egress-policy
 class tc-7
  priority-level 1
 class tc-6
  Priority-level 2
 class tc-5
  bandwidth remaining ratio 1
 class t-4
  bandwidth remaining ratio 1
 class tc-3
  bandwidth remaining ratio 1
 class tc-2
  bandwidth remaining ratio 1
 class tc-1
  bandwidth remaining ratio 1
 class class-default
  bandwidth remaining ratio 4
```

- Two priority queues here
- Level 1 has the absolutely priority over level 2

- Use "bandwidth remaining ratio" to assign weight
- This example gives a higher weight to class-default and same weight to rest of the queues
- Served round robin around 6 queues if there isn't any traffic on the two PQs

Shaping

- Smooth out traffic peaks, microburst, with preserving traffic

- Control traffic rate to the desired value with buffering.

- Usually in the egress direction

- Can be applied on all classes, regardless of priority level.

Shaping Example:
```
policy-map type queueing egress-queueing
  class tc7
    priority level 1
    shape average 1g
  class tc6
    priority level 2
    shape average 5g
...

  class class-default
    shape average 5g
```

Original traffic

With shaping
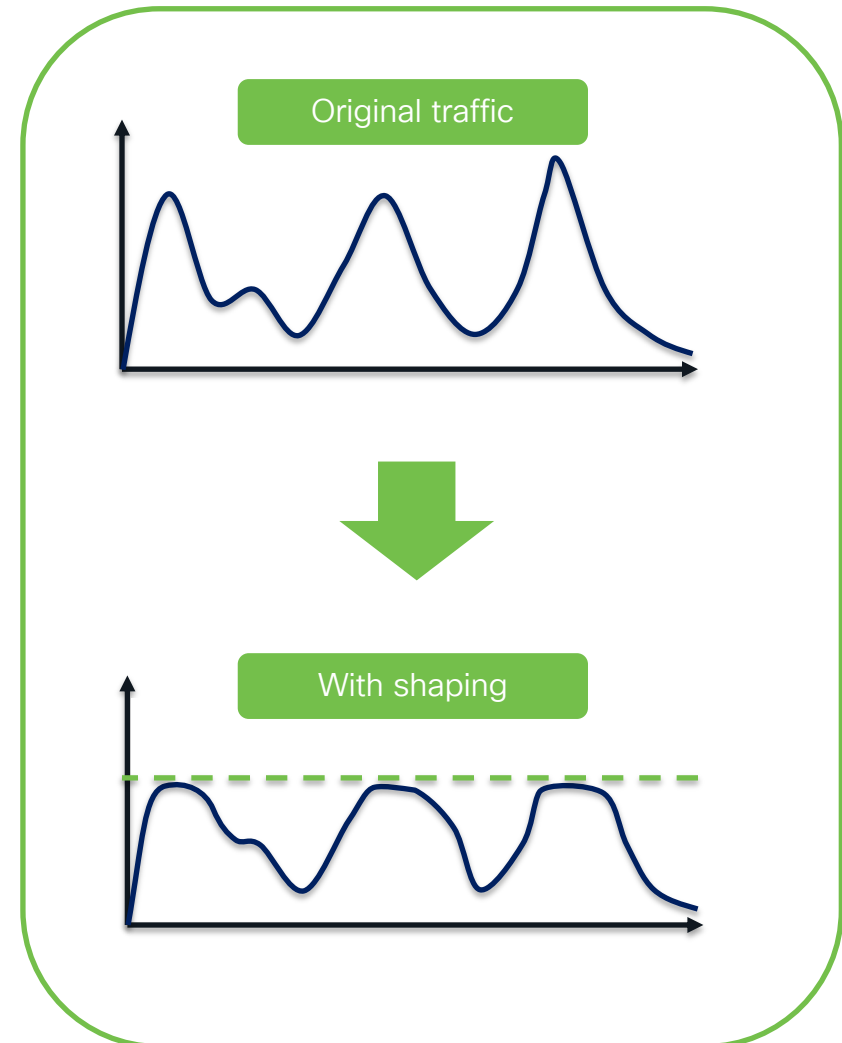
Egress Marking

Egress Marking

- Used to change packet tags of packets egressing the switch.

- A separate policy-map apart from the queueing policy-map.

- If both queueing and marking egress policy-maps are applied, marking happens after queueing actions.

- ACL matching in egress is no supported.

```
class-map match-any dscp-af41
 match dscp af41
!

policy-map egress-map
 class dscp-af41
  set dscp af31
  !

interface interface <#>
 service-policy output egress-map
```
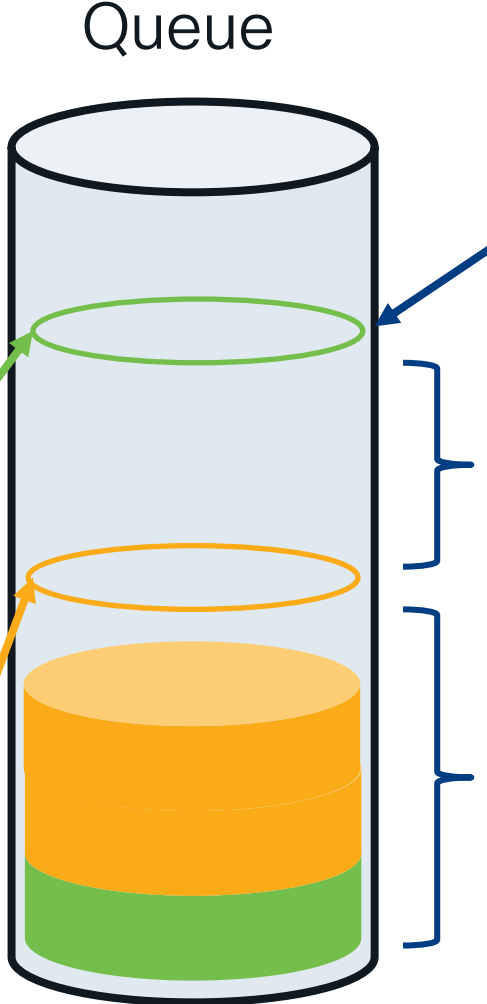
Silicon One Q200 Congestion Management

Weighted Tail Drop (WTD)

Queue

Ingress Policy-map

```
policy-map ingress-policy
        class class-5-green
          set traffic-class 5
        class class-5-yellow
          set traffic-class 5
          set discard-class 1
```

```
…
class-map match-any traffic-class-5
 match traffic-class 5
…
```

Egress Policy-map

```
policy-map type queueing queue-policy
 ...
 class traffic-class-5
  queue-limit 100000000 bytes
  queue-limit discard-class 1 50000000 bytes
```
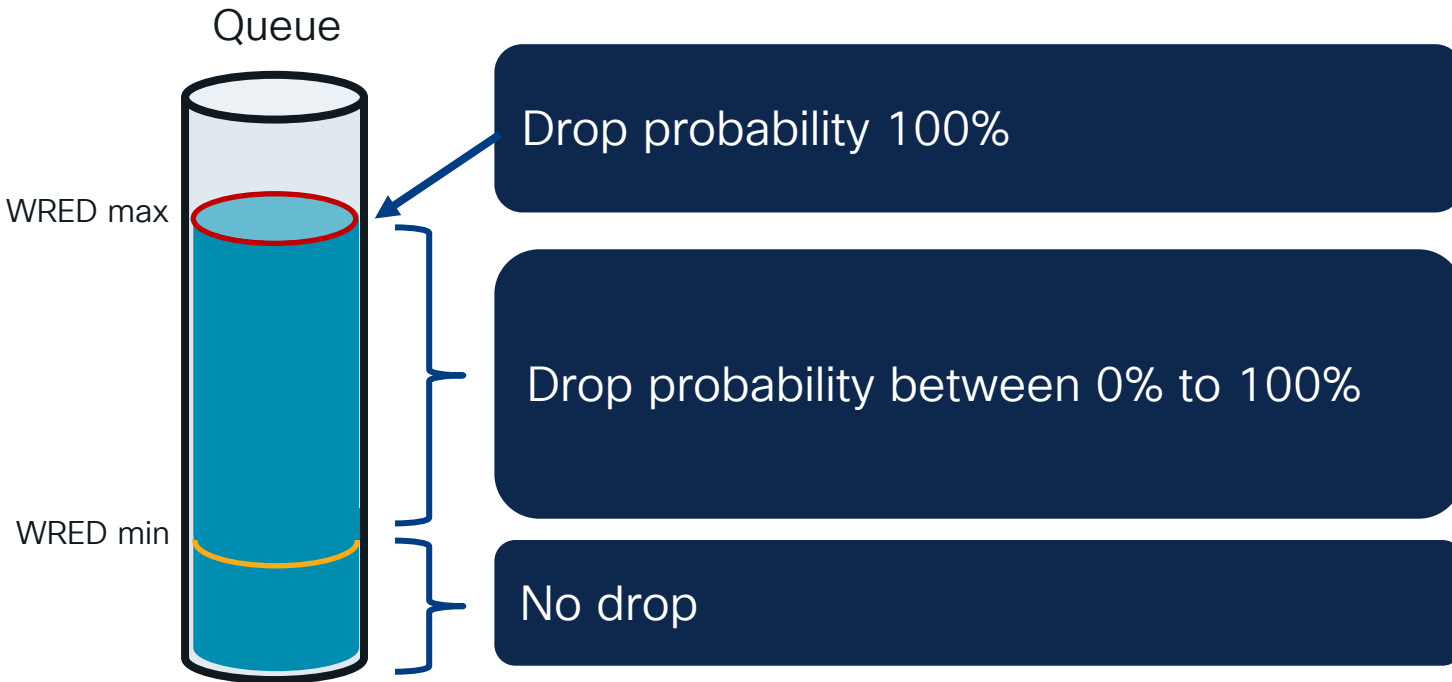
Green traffic will be dropped over this the threshold

Only traffic marked with green would be here
Yellow traffic would be dropped once it is over the yellow threshold

No drop for both Green and Yellow traffic here

Weighted Random Early Drop (WRED)

Queue

WRED max

WRED min

Drop probability 100%

Drop probability between 0% to 100%

No drop

WRED – Drop probability

Queue

CLI:
`random-detect discard-class <0-1> percent <Min> <Max> <K>`

WRED max

WRED min

Drop Probability

1

1/K

0

0   Min   1   2   3   4   5   6   7   8   9   10   11   12   13   14   15   Max

Queue Size

WRED min   WRED max
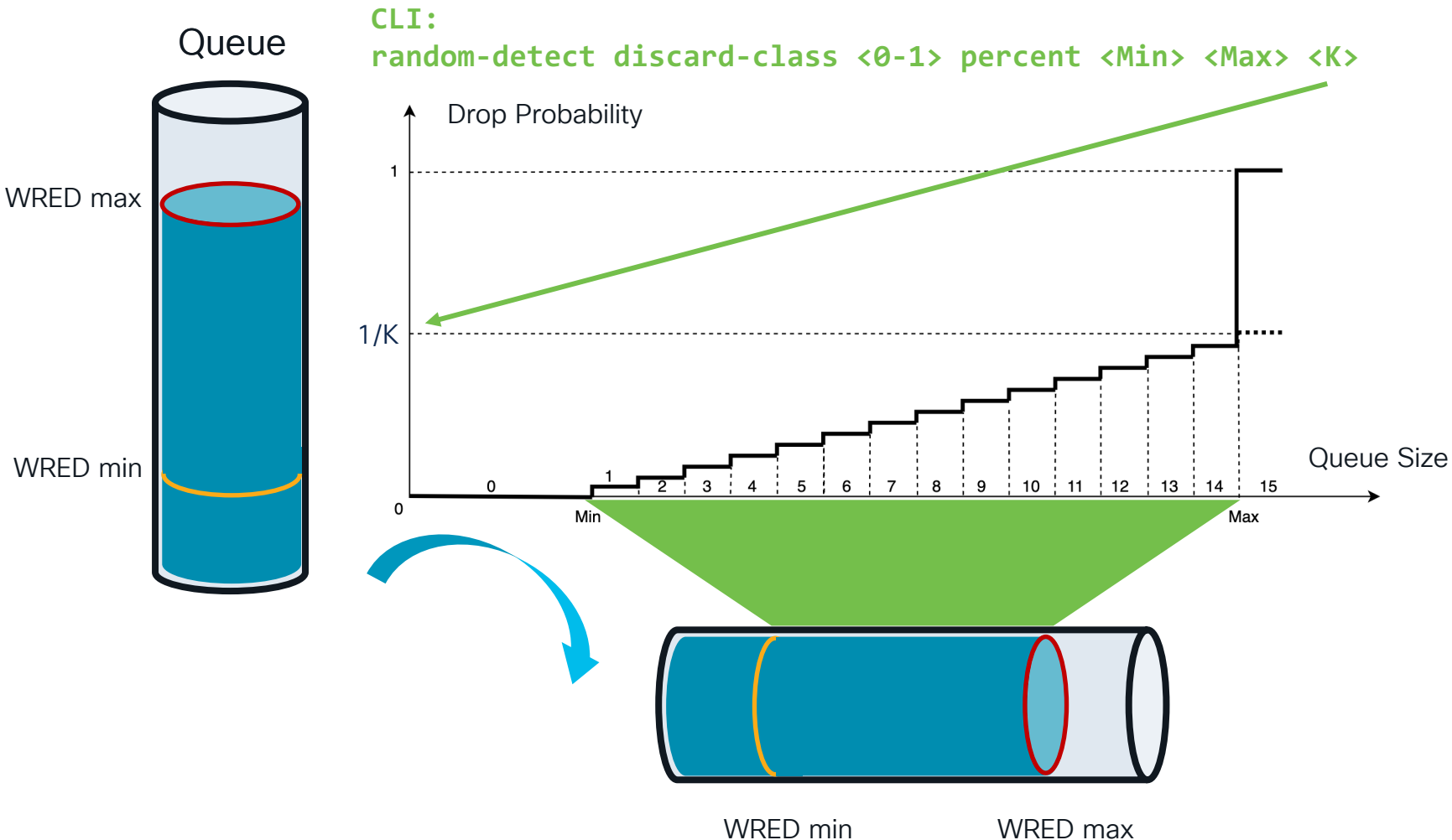
- Drop probability increases as the queue utilization increases

- Silicon One ASIC provides 16 regions (drop probabilities)

- Silicon One ASIC provides a knob to influence the drop probability

WRED – Example

Ingress Policy-map
```
policy-map ingress-policy
…
 class class-5-green
  set traffic-class 5
 class class-5-yellow
  set traffic-class 5
  set discard-class 1
…
```

← Ingress policy with marking/coloring of packets

class-map
```
class tc5
   match traffic-class 5
…
```

← Map the ingress class to one of the traffic-class

Egress Policy-map
```
policy-map type queueing queue-policy
...
 class tc5
  random-detect discard-class-based
  random-detect discard-class 0 percent 80 90 5
  random-detect discard-class 1 percent 40 70 2
...
```

←
- Green traffic has higher Min and Max threshold comparing to yellow traffic
- Green traffic also has higher forwarding probability (lower drop probability) comparing to yellow traffic
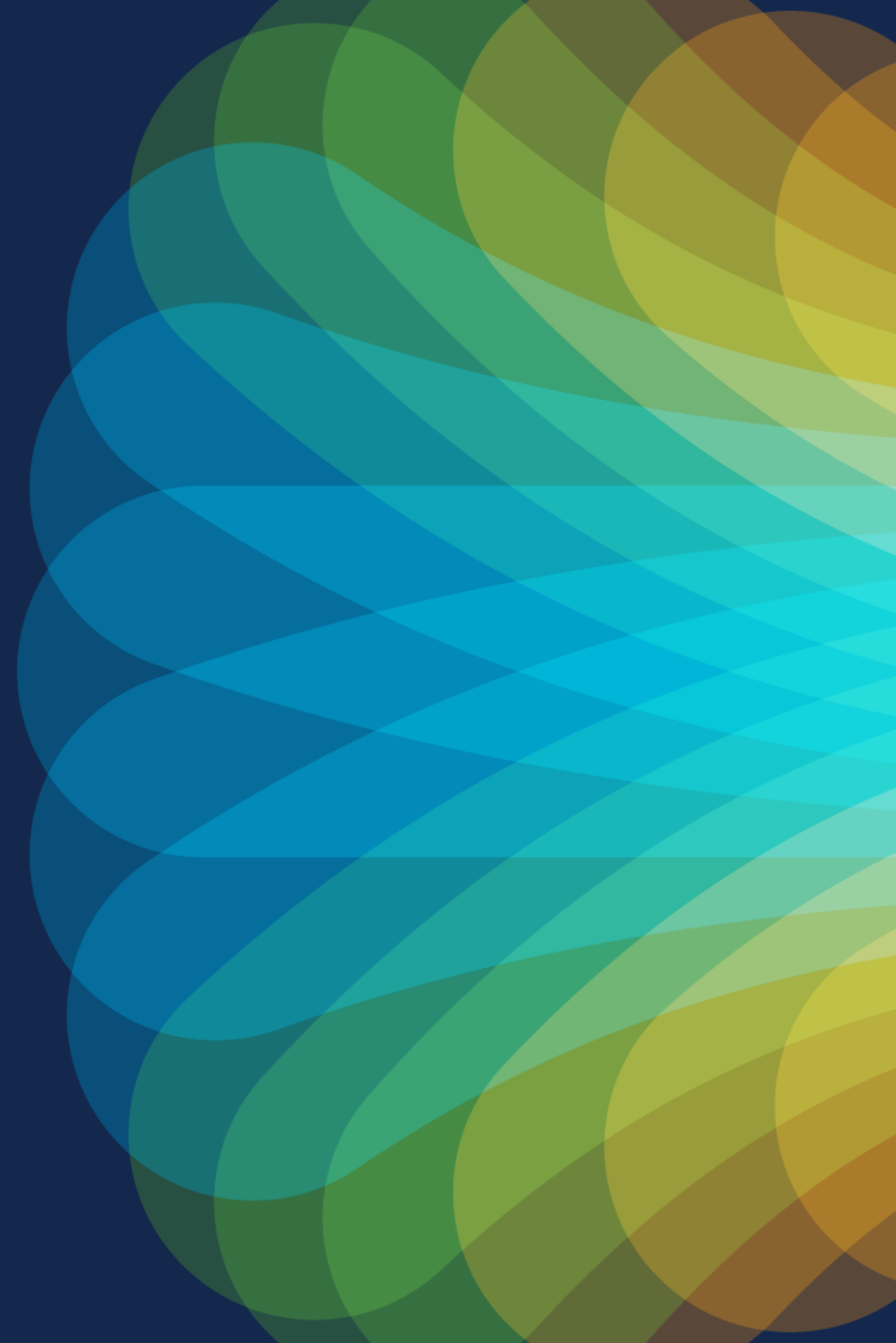
Buffers

Silicon One Buffers

- Two different buffers to address two different requirements.
  - 80 MB of Shared Memory Sub-system (SMS) buffers:
    - Low latency packet queueing (video/voice packets)
    - Shallow specialized pool of buffers for quick queueing.
  - 8 GB of High Bandwidth Memory (HBM) buffers:
    - Deep pool of on-demand buffers for guaranteed delivery.
    - Reserve to absorb occasional bursts or address speed over-subscription between ingress and egress.

Local vs. HBM Buffers

HBM: Guaranteed Delivery & Burst Absorption

8GB HBM

Video
Voice
Control
Transfer
Scavenger

80MB SMS

SMS: Low-Latency & Priority Queuing
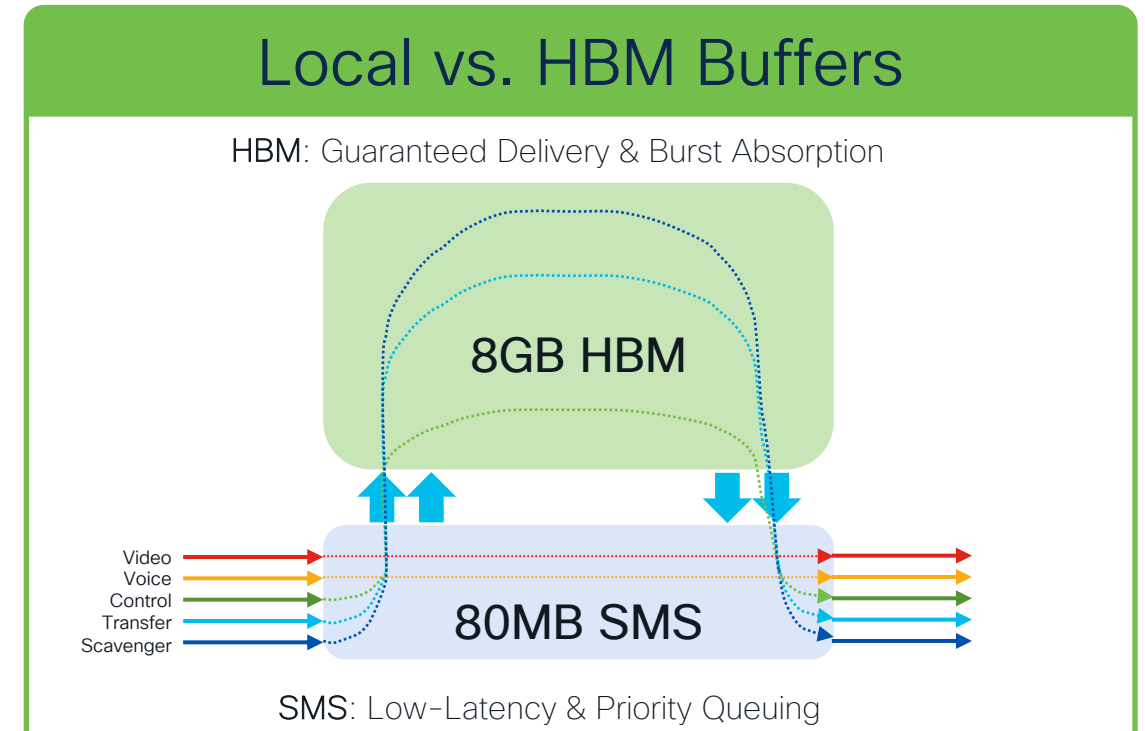
- Packet will always hit the SMS buffers first.
- SMS send the packet to HBM if additional buffers are needed.
- HBM **CANNOT** send the packet to the output queue, it has to be sent to the SMS again to be sent to the egress.

HQoS

Silicon One Hierarchical QoS (HQoS)

HQoS (two-level hierarchy) allows a parent and child policies on an interface for greater granularity. The Cisco Silicon One supports shaping as parent action.
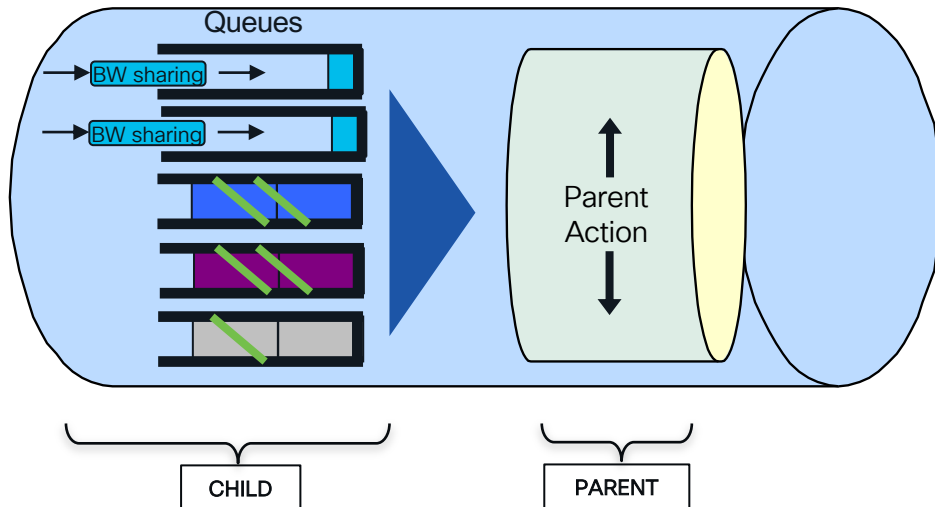
Child Action	Parent Action
Bandwidth sharing + Priority	Shaping

Silicon One vs UADP QoS

—

QoS tools on UADP and Silicon One ASICs 1/2

Features	UADP ASIC	Silicon One ASIC (Q200)
Trust	Trust all ports by default	Trust all ports by default
Classification	Based on Packet header and ACL for both ingress and egress	Based on packet header and ACL for ingress **Based on packet header for egress**
Marking	Header, Table-map, QoS-Group for ingress Header and table-map for egress	Header, Table-map, QoS-group, **traffic-class, discard-class** for ingress Header and table-map for egress
Policing	Both ingress and egress	Ingress only

QoS tools on UADP and Silicon One ASICs 2/2

Features	UADP ASIC	Silicon One ASIC (Q200)
Queueing	Based on header or QoS group Bandwidth and Bandwidth remaining	Based on **traffic-class** **Bandwidth remaining**
Buffering	Dedicated and shared buffer with DTS	SMS: Low-latency & priority queueing HBM: Guaranteed Delivery & Burst Absorption
Shaping	Egress	Egress
Congestion Management	WTD: three thresholds per class WRED: three thresholds' pairs per class	WTD: **two** threshold per class WRED: **two** thresholds' pairs per class

QoS Config Migrations

Config Migration steps

1 Define the problem/behavior addressed with QoS.

Simply copy-pasting existing configs between platform families will always throw errors due to differences in syntax and supported actions between platforms.

2 How many times do you want to split your traffic – Upto 8 queues possible with our ASIC

Its often not as much as you think you need. Broad generalized splits often are more efficient than granular splits

3 Do you want multiple strict priority classes?

Know what strict means. All traffic coming into it will be serviced at the expense of other classes.

4 Define traffic shaping/policing or sharing between queues.

Police/shape priority queues. Use weights to control bandwidth sharing with remaining queues

5 Do you want differential priority for packets within same class? – Use WTD or WRED

Advanced configuration options, not required for most use cases.

Config Migration from 6k to UADP
MLS configs

```
              Catalyst 6K MLS config

!
mls qos
!
interface TenGigabitEthernet2/14
  platform qos queue-mode mode-dscp
  wrr-queue bandwidth 20 1 14 10 15 2 3
  priority-queue queue-limit 20
  wrr-queue dscp-map 1 1 0 1 2 3 4 5 6 7
  wrr-queue dscp-map 2 1 8 14 32 35 36 37 38
  wrr-queue dscp-map 3 1 10 22 24 30 49 50 51 52
  wrr-queue dscp-map 4 1 16 18
  wrr-queue dscp-map 5 1 26
  wrr-queue dscp-map 6 1 48 56
  wrr-queue dscp-map 7 1 34
```

1 Enable QoS Globally

2 Use DSCP mapping to different classes/queues

3 1 priority queue

4 7 normal queues WRR and the weights

Config Migration from 6k to UADP
MLS configs

Catalyst 9K(UADP) Config

```
class-map match-any queue1
 match dscp 0 1 2 3 4 5 6 7
class-map match-any queue2
 match dscp 8 14 32 35 36 37 38
class-map match-any queue3
 match dscp 10 22 24 30 49 50 51 52
class-map match-any queue4
 match dscp 16 18
class-map match-any queue5
 match dscp 26
class-map match-any queue6
 match dscp 48 56
class-map match-any queue7
 match dscp 34
class-map match-any priority
 match dscp ef
policy-map egress-queue
 class priority
  priority level 1 percent 20
 class queue1
  bandwidth remaining percent 20
 class queue2
  bandwidth remaining percent 1
 class queue3
  bandwidth remaining percent 14
 class queue4
  bandwidth remaining percent 10
 class queue5
  bandwidth remaining percent 15
 class queue6
  bandwidth remaining percent 2
 class queue7
  bandwidth remaining percent 3
```

Catalyst 6K MLS config

```
!
mls qos
!
interface TenGigabitEthernet2/14
 platform qos queue-mode mode-dscp
 wrr-queue bandwidth 20 1 14 10 15 2 3
 priority-queue queue-limit 20
 wrr-queue dscp-map 1 1 0 1 2 3 4 5 6 7
 wrr-queue dscp-map 2 1 8 14 32 35 36 37 38
 wrr-queue dscp-map 3 1 10 22 24 30 49 50 51 52
 wrr-queue dscp-map 4 1 16 18
 wrr-queue dscp-map 5 1 26
 wrr-queue dscp-map 6 1 48 56
 wrr-queue dscp-map 7 1 34
```

default

1. Enable QoS Globally

2. Use DSCP mapping to different classes/queues

3. 1 priority queue

4. 7 normal queues WRR and the weights

* Consider weight as interface speed can be much higher now

Config Migration from 6k to UADP
MQC Configs

Catalyst 6K Configuration

```
class-map type lan-queuing match-all REALTIME
 match dscp ef
class-map type lan-queuing match-all NETWORK_CONTROL
 match dscp cs6  cs7
class-map type lan-queuing match-all VIDEO
 match dscp cs3  af31  af32  af33

policy-map type lan-queuing CAMPUS_EGRESS_6800_POLICY
 class type lan-queuing REALTIME
    priority level 1
 class type lan-queuing NETWORK_CONTROL
    bandwidth remaining percent 10
 class type lan-queuing VIDEO
    bandwidth remaining percent 20
 class class-default
    random-detect dscp-based
    random-detect dscp af11 percent 80 100

Interface gig1/0/1
 service-policy type lan-queueing output
CAMPUS_EGRESS_6800_POLICY
```

1 Use DSCP mapping to different classes/queues

2 1 priority queue

3 3 normal queues
WRR on non-default queue
WRED for class-deault

Config Migration from 6k to UADP
MQC Configs

Catalyst 6K Configuration

```
class-map type lan-queuing match-all REALTIME
 match dscp ef
class-map type lan-queuing match-all NETWORK_CONTROL
 match dscp cs6  cs7
class-map type lan-queuing match-all VIDEO
 match dscp cs3  af31  af32  af33

policy-map type lan-queuing CAMPUS_EGRESS_6800_POLICY
 class type lan-queuing REALTIME
    priority level 1
 class type lan-queuing NETWORK_CONTROL
    bandwidth remaining percent 10
 class type lan-queuing VIDEO
    bandwidth remaining percent 20
 class class-default
    random-detect dscp-based
    random-detect dscp af11 percent 80 100

Interface gig1/0/1
 service-policy type lan-queueing output
CAMPUS_EGRESS_6800_POLICY
```

1 Use DSCP mapping to different classes/queues

2 1 priority queue

3 3 normal queues
WRR on non-default queue
WRED for class-deault

Catalyst 9K Configuration

```
class-map match-any REALTIME
 match dscp ef
class-map match-any NETWORK_CONTROL
 match dscp cs6  cs7
class-map match-any VIDEO
 match dscp cs3  af31  af32  af33

policy-map CAMPUS_EGRESS_POLICY
 class type REALTIME
    priority level 1
 class type NETWORK_CONTROL
    bandwidth remaining percent 10
 class type VIDEO
    bandwidth remaining percent 20
 class class-default
    random-detect dscp-based
    random-detect dscp af11 percent 80 100

Interface gig1/0/1
 service-policy output CAMPUS_EGRESS_POLICY
```

Migration from Catalyst 6K to Silicon One Q200

Catalyst 6K Configuration

```
class-map type lan-queuing match-all REALTIME
 match dscp ef
class-map type lan-queuing match-all NETWORK_CONTROL
 match dscp cs6  cs7
class-map type lan-queuing match-all VIDEO
 match dscp cs3  af31  af32  af33
…
policy-map type lan-queuing CAMPUS_EGRESS_6800_POLICY
 class type lan-queuing REALTIME
    priority level 1
 class type lan-queuing NETWORK_CONTROL
    bandwidth remaining percent 10
 class type lan-queuing VIDEO
    bandwidth remaining percent 20
 class class-default
    random-detect dscp-based
    random-detect dscp af11 percent 80 100
```

1. Classified Based on DSCP value
2. 4 classes (3 defined + default)

3. 4 queues
4. 1 priority queue
5. Scheduling is WRR with "bandwidth remaining"
6. Congestion management is WRED with the default class

Config Migration from Catalyst 6K to Silicon One Q200

Catalyst 6K Configuration

```
class-map type lan-queuing match-all REALTIME
 match dscp ef
class-map type lan-queuing match-all NETWORK_CONTROL
 match dscp cs6  cs7
class-map type lan-queuing match-all VIDEO
 match dscp cs3  af31  af32  af33
…
policy-map type lan-queuing CAMPUS_EGRESS_6800_POLICY
 class type lan-queuing REALTIME
    priority level 1
 class type lan-queuing NETWORK_CONTROL
    bandwidth remaining percent 10
 class type lan-queuing VIDEO
    bandwidth remaining percent 20
 class class-default
    random-detect dscp-based
    random-detect dscp af11 percent 80 100
```

1. Classified Based on DSCP value
2. 4 classes (3 defined + default)
3. 4 queues (traffic-class), traffic-7 is priority level 1

Catalyst 9K(Q200) Configuration

```
class-map match-all REALTIME
 match dscp ef
class-map match-all NETWORK_CONTROL
 match dscp cs6  cs7
class-map match-all VIDEO
 match dscp cs3  af31  af32  af33
class-map match-all default-green
 match dscp af11


policy-map INGRESS
 class REALTIME
    set traffic-class 7
 class NETWORK_CONTROL
    set traffic-class 6
 class VIDEO
    set traffic-class 5
 class default-green
    set traffic-class 0
 class class-default
    set discard-class 1
```
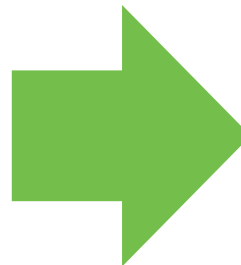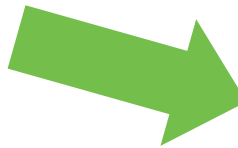
Note: class-default is always assigned with traffic-class 0

Config Migration from Catalyst 6K to Silicon One Q200

Apply policy on the egress interface

Map the traffic-class marking defined on the ingress
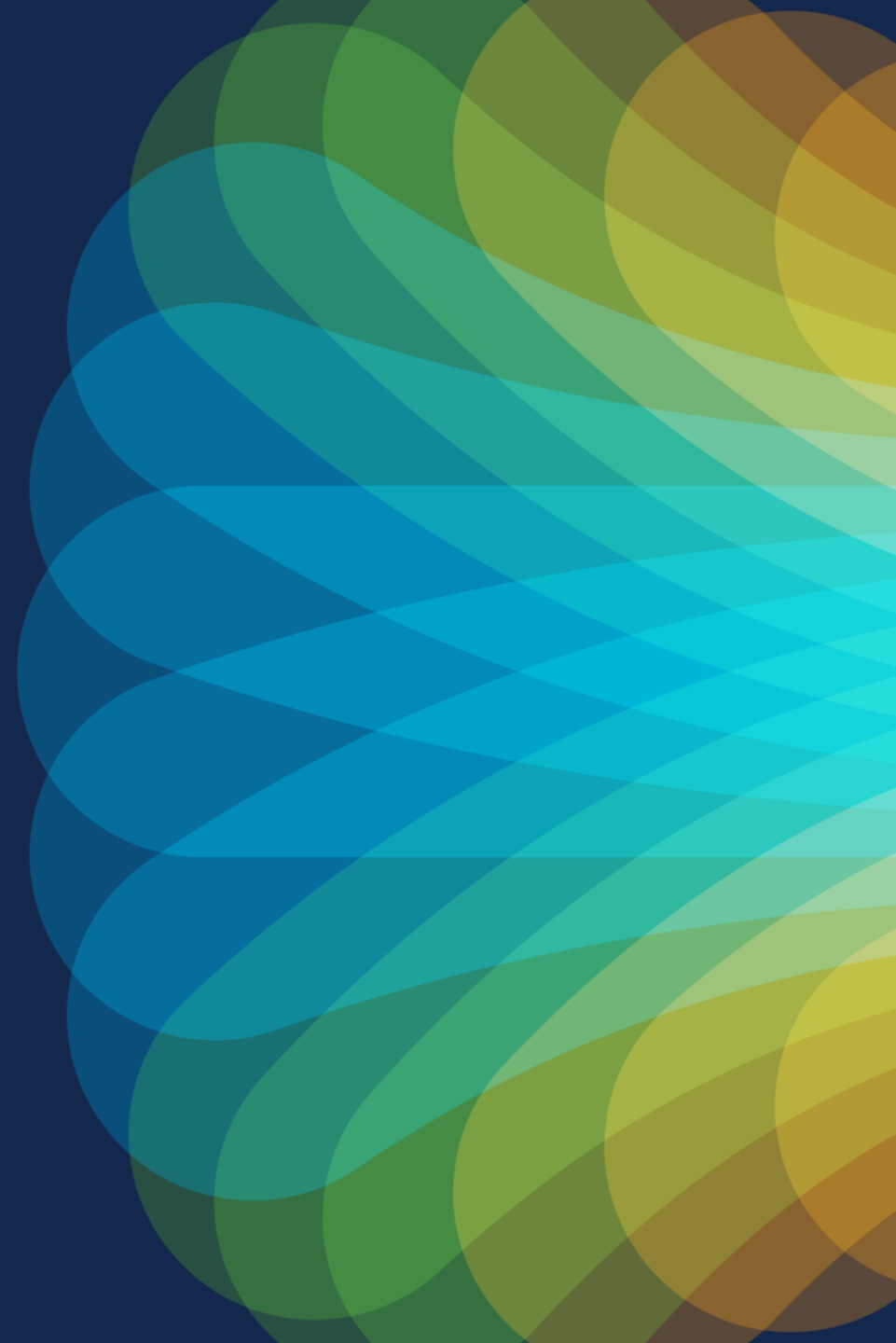
Catalyst 9K(Q200) Configuration

```
class-map tc7
 match traffic-class 7
class-map tc6
 match traffic-class 6
class-map tc5
 match traffic-class 5

policy-map type queuing EGRESS
class tc7
    priority level 1
class tc6
    bandwidth remaining ratio 1
class tc5
    bandwidth remaining ratio 2
class class-default
    random-detect discard-class-based
    random-detect discard-class 0 percent 80 100
    random-detect discard-class 1 percent 40 100
```

Catalyst 6K Configuration

```
policy-map type lan-queuing CAMPUS_EGRESS_6800_POLICY
class type lan-queuing REALTIME
    priority level 1
class type lan-queuing NETWORK_CONTROL
    bandwidth remaining percent 10
class type lan-queuing VIDEO
    bandwidth remaining percent 20
class class-default
    random-detect dscp-based
    random-detect dscp af11 percent 80 100
```

3. 1 priority queue
4. Scheduling is WRR with "bandwidth remaining"
5. Congestion management is WRED with the default class

Note: C6K WRED default min is 40, max is 100.

Summary

Do we need QoS?

User Experience

| Guaranteeing voice quality | Bandwidth Savvy Business Applications | protect network infrastructure to deal with abnormal events |
| Video Quality | de-prioritizing non-business applications | protecting the control planes |

QoS helps define the latency priority for your traffic packets

cisco Live!

Did you know?

You can have a
one-on-one session with
a technical expert!

Visit Meet the Expert in The HUB
to meet, greet, whiteboard & gain
insights about your unique questions
with the best of the best.

Meet the Expert Opening Hours:

Tuesday	3:00pm – 7:00pm
Wednesday	11:15am – 7:00pm
Thursday	9:30am – 4:00pm
Friday	10:30am – 1:30pm

Session Surveys

We would love to know your feedback on this session!

- Complete a minimum of four session surveys and the overall event surveys to claim a Cisco Live T-Shirt

Continue your education

- Visit the Cisco Showcase for related demos

- Book your one-on-one Meet the Expert meeting

- Attend the interactive education with DevNet, Capture the Flag, and Walk-in Labs

- Visit the On-Demand Library for more sessions at www.CiscoLive.com/on-demand

Thank you

CISCO

The bridge to possible

cisco Live!

#CiscoLiveAPJC