
CISCO
You make possible

Packet Forwarding & QoS Implementation on NCS 500

New Age Access Platforms

Paban Sarma, Technical Marketing Engineer
@pabanelb

BRKSPG-2012

Barcelona | January 27-31, 2020

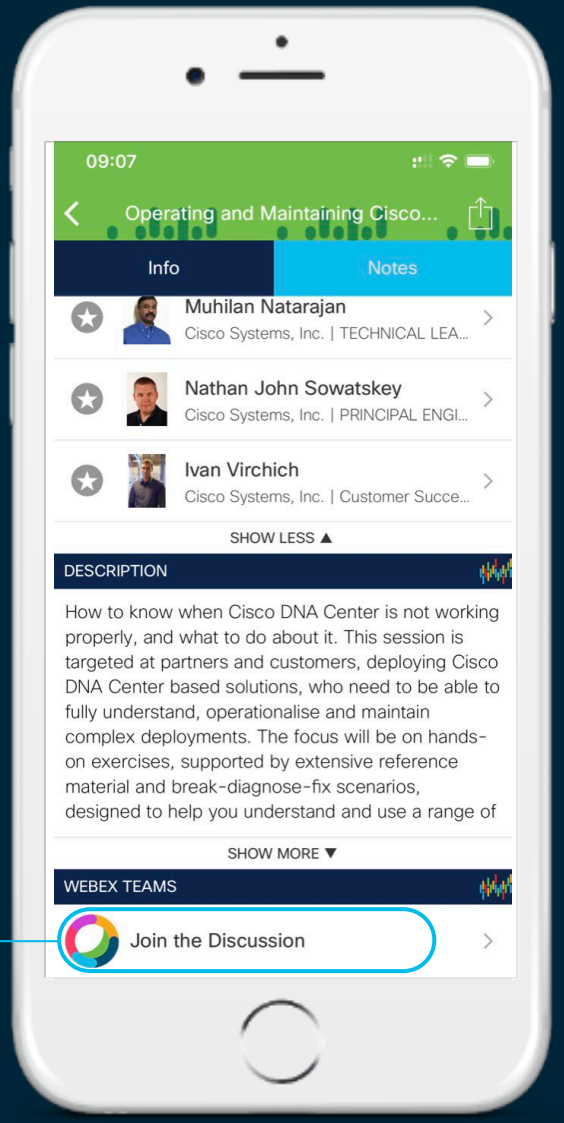
Cisco Webex Teams

Questions?

Use Cisco Webex Teams to chat with the speaker after the session

How

1. Find this session in the Cisco Events Mobile App
2. Click “Join the Discussion”
3. Install Webex Teams or go directly to the team space
4. Enter messages/questions in the team space



Session Takeaway

You can simply deliver QoS in your access network
with NCS 500 !!

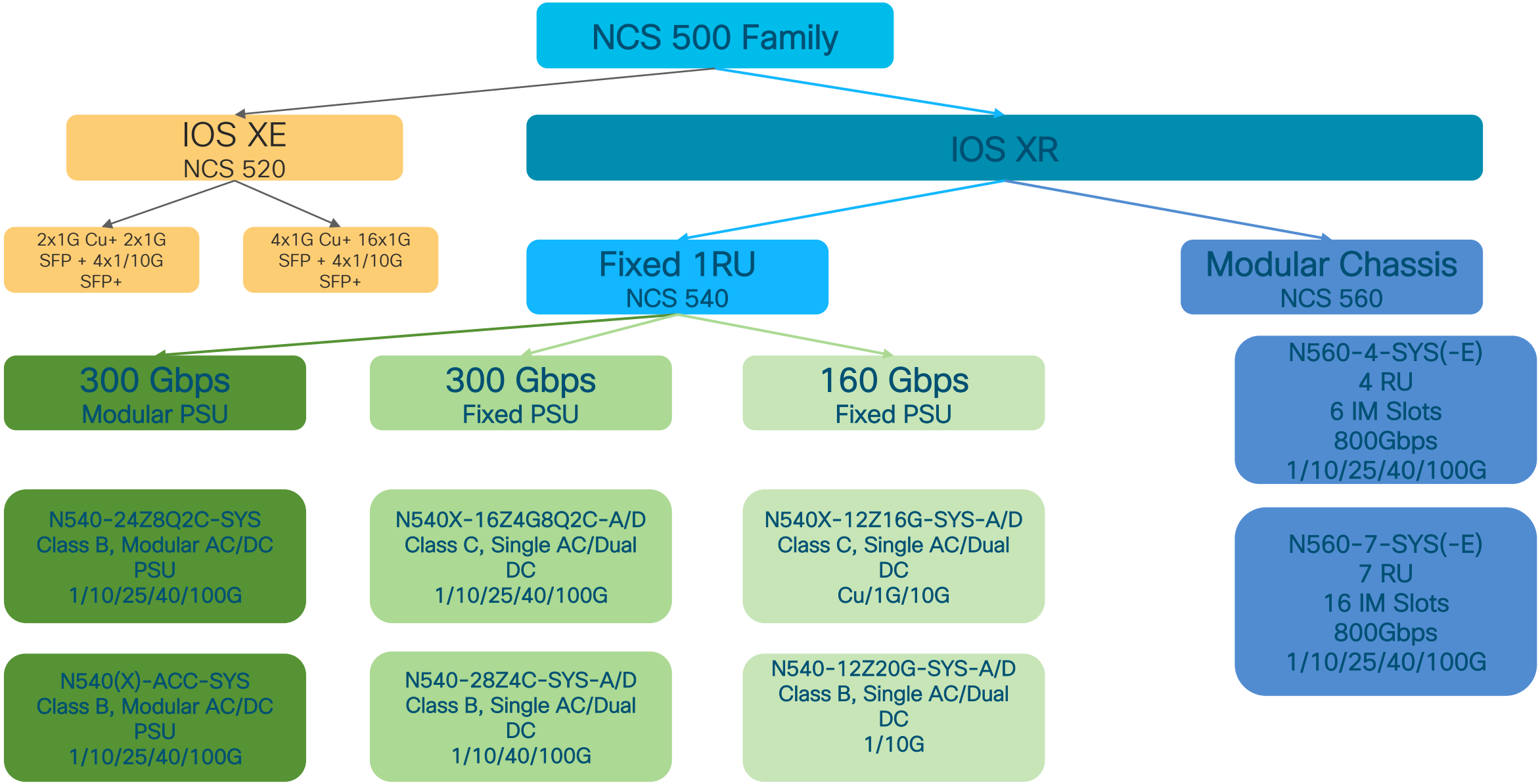
Agenda

- NCS 500 Family & Internal Architecture
- NCS 5xx QoS Overview
  - Ingress
  - Egress
- QoS Behaviors
  - H-QoS
  - Bundle QoS
  - Multicast QoS
  - QoS for DiffServ Tunneling Modes
- Behavioral Difference with ASR 9xx QoS
- Summary

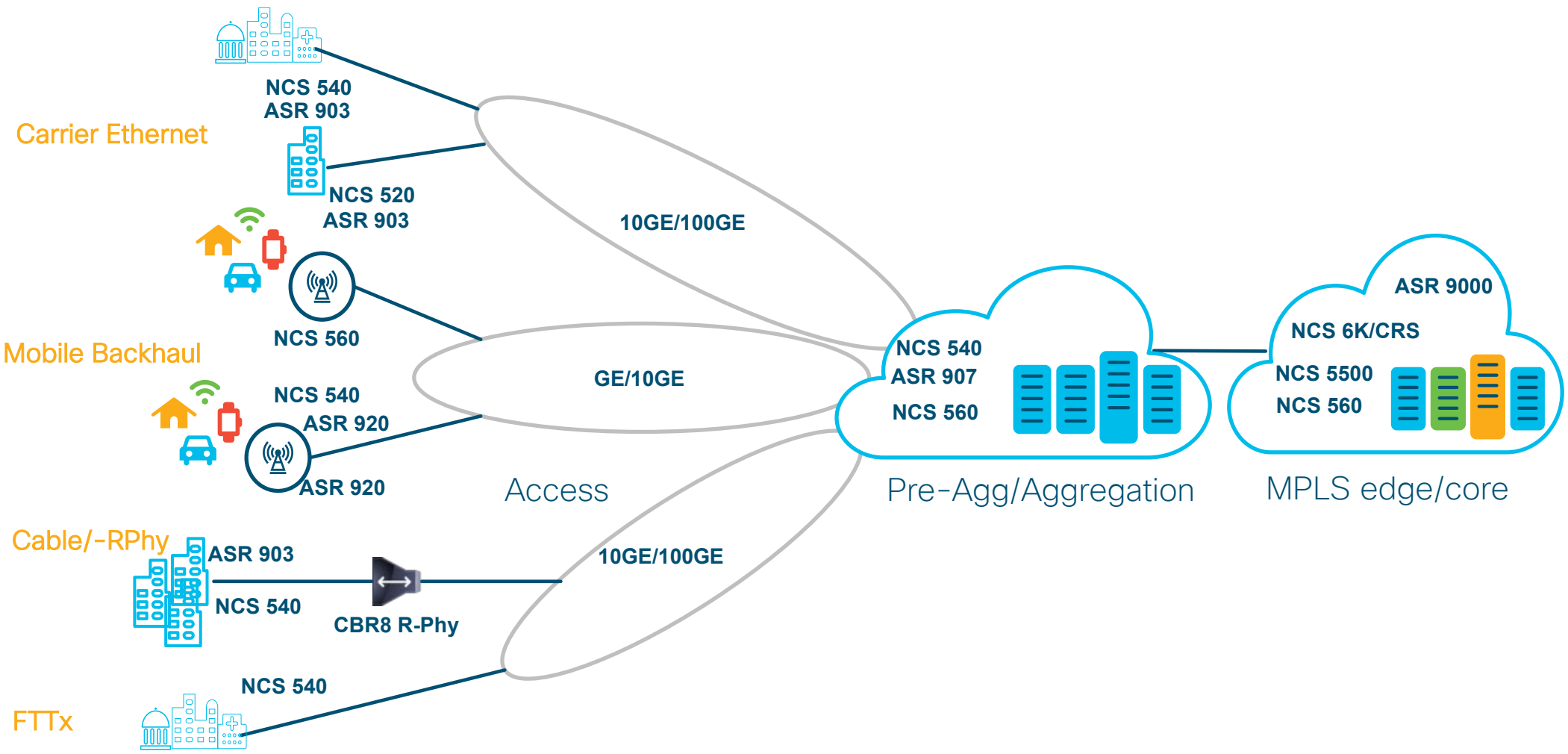
NCS 500 Family & Internal Architecture

cisco Live!

The NCS 500 Family



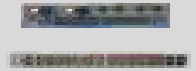
NCS 500 in the Network



Highlights

IOS XR

Dense fixed and modular XR portfolio



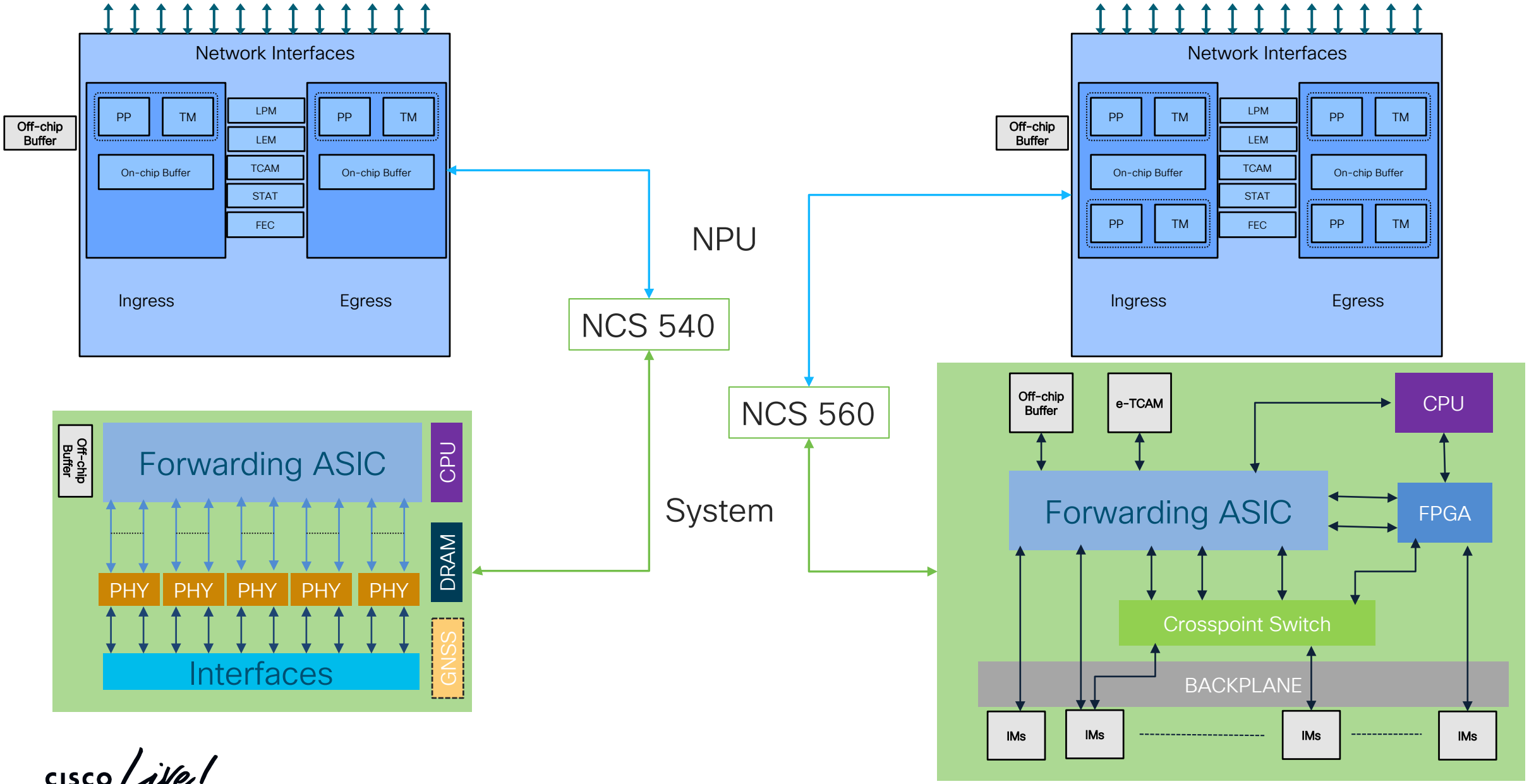
Small form factor with low power

Secure

Wide Range of BW

Indoor/outdoor deployments

NCS 500 Architecture



Refreshing The Basics

Why QoS !

Services

- ✓ Access is the starting point
- ✓ Connectivity starts here

Satisfaction

- ✓ Delivery with Guranted SLA
- ✓ Ease of Enhancement

- ✓ Connectivity with guarantee
- ✓ Prioritization
- ✓ Best Effort

SLA

Revenue

- ✓ Service pricing
- ✓ Add on Bandwidth
- ✓ Utilize Available infrastructure

Identify Traffic

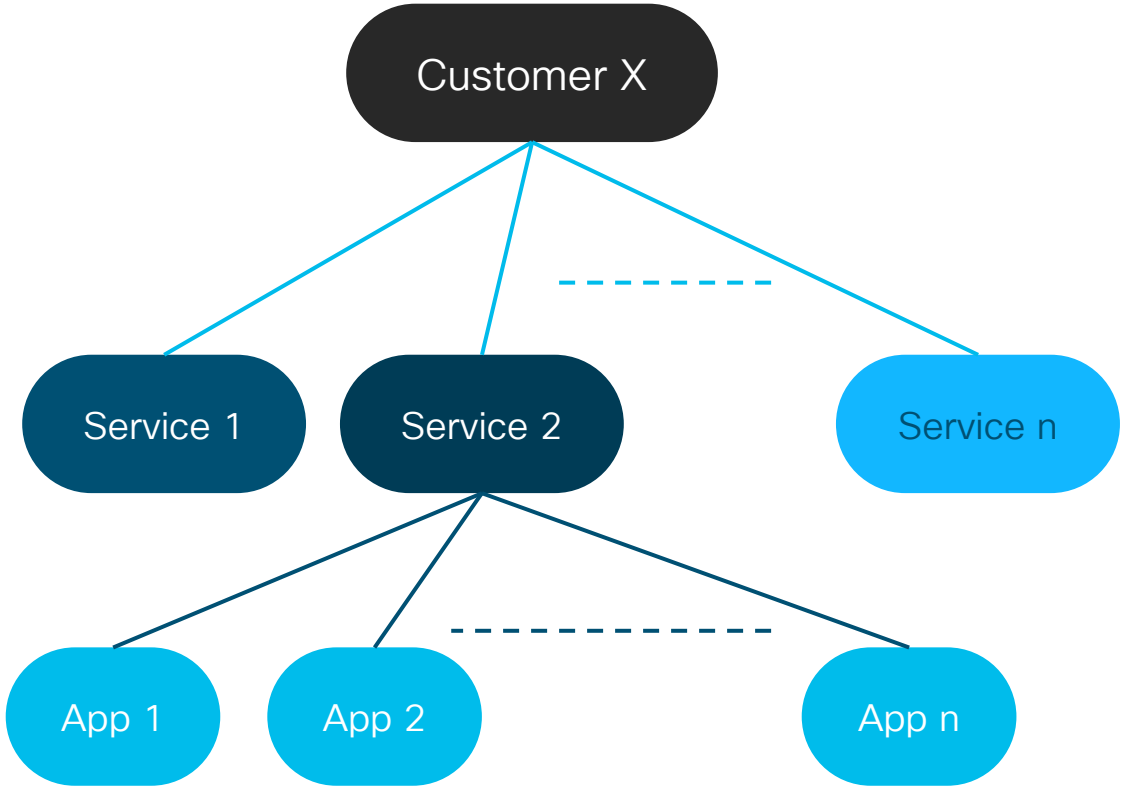
- Different Traffic Type
  - Latency sensitive/Real Time
  - Best Effort
  - Undesired
  - Control Plane
- Classification
  - Information present in headers
  - IP/PCP/DSCP/EXP etc.
- Marking
  - Simplify the Network QoS
  - Ensure E2E QoS

Manage SLA

- Policers
  - Cap users at the committed rate: 1R2C
  - Give something extra with Best Effort: 2R3C
- Shape/Queues
  - Cap Users with shape
  - Prioritize/schedule different traffic
  - Weight Between BE traffic

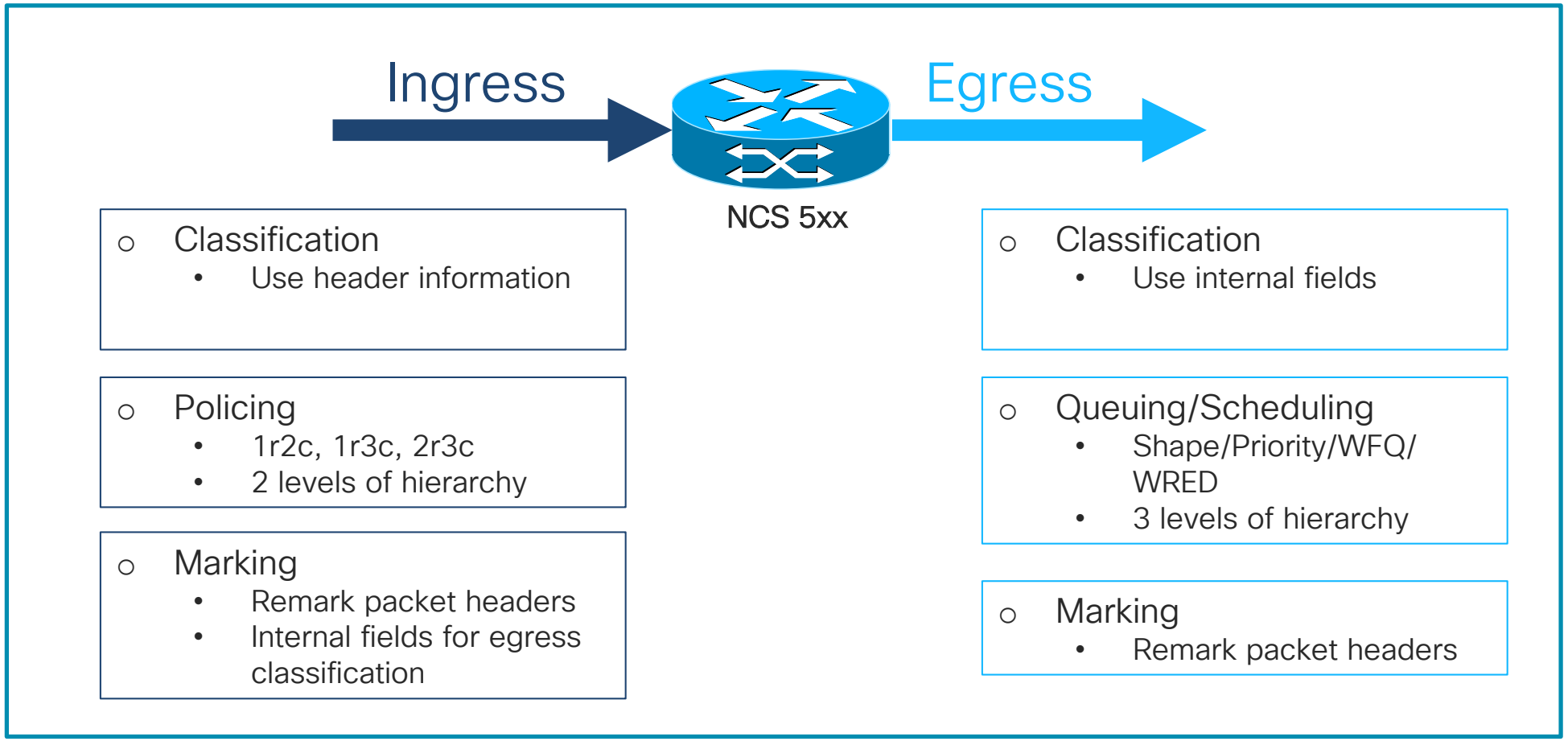
SLA Levels

- Differentiated QoS Per Application
- Aggregated QoS Per Service
- Differentiated QoS for Per Service
- Aggregated QoS Per Customer

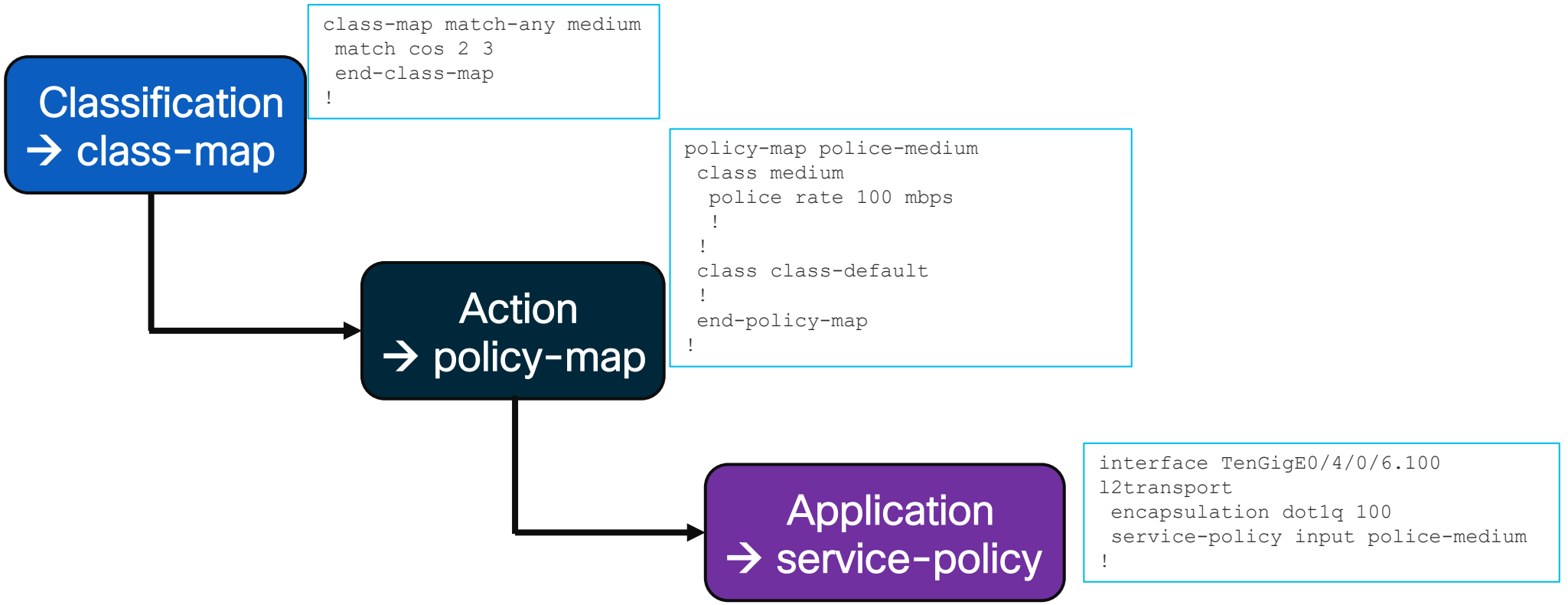


NCS 500 QoS Overview

NCS 500 QoS Behavior

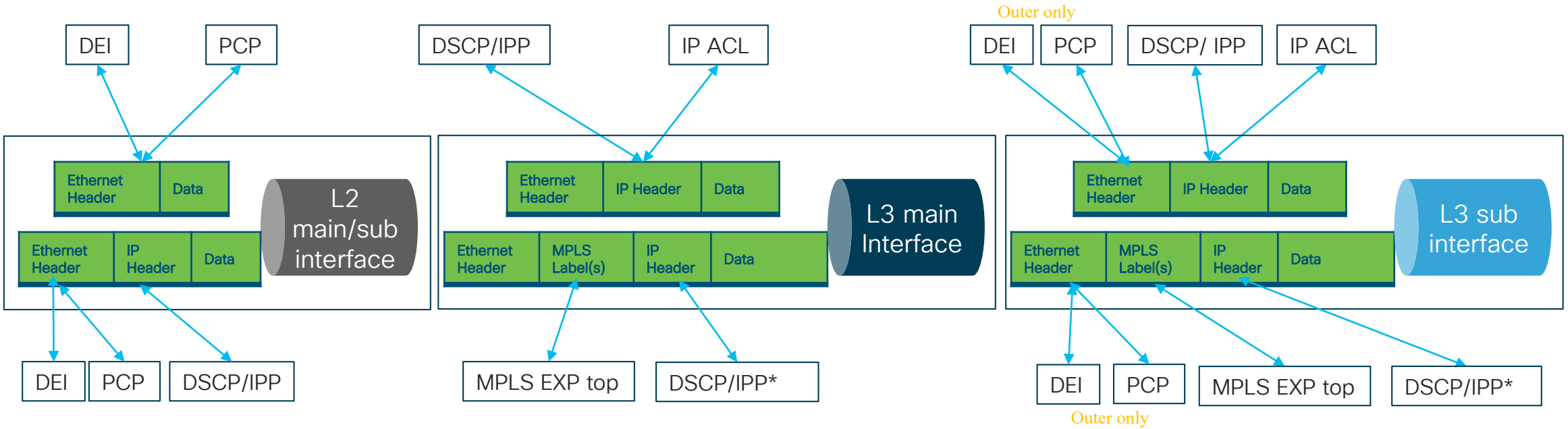


Modular QoS CLI (MQC)



Ingress QoS

Classification

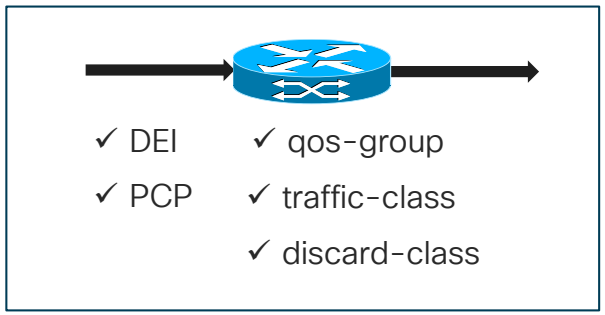


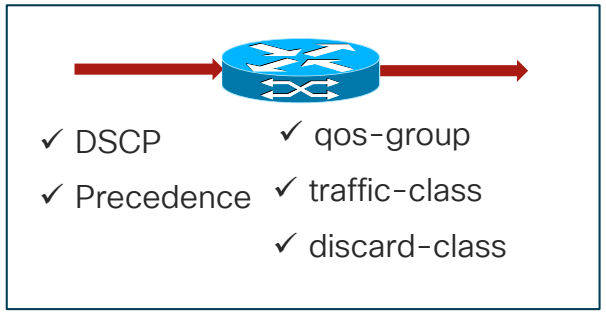
- Match on protocol supported for L3 interface
- Only Single Protocol
- Match-not is not supported for ACLs

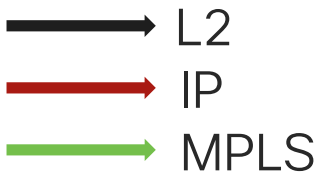
PCP : Priority Code Point (802.1p)
DEI : discard eligibility indicator
IPP : IP Precedence

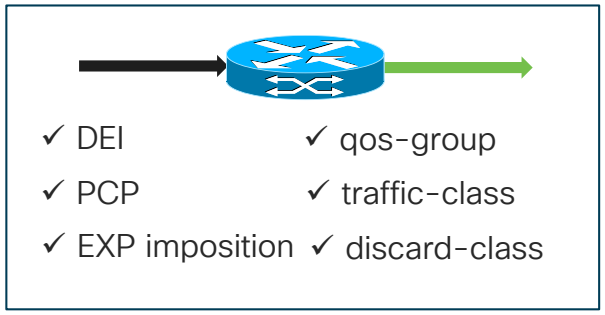
* With MPLS disposition- ingress short pipe mode

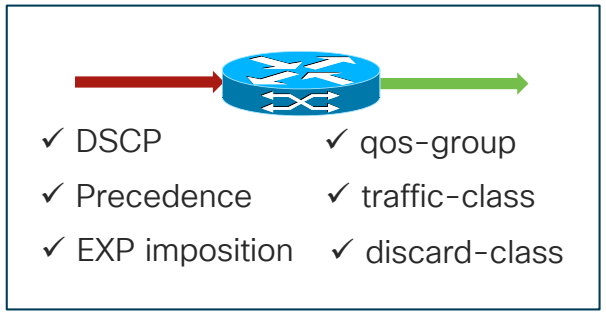
Supported Marking Matrix

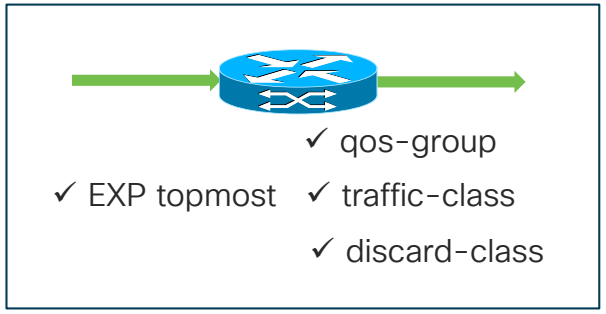


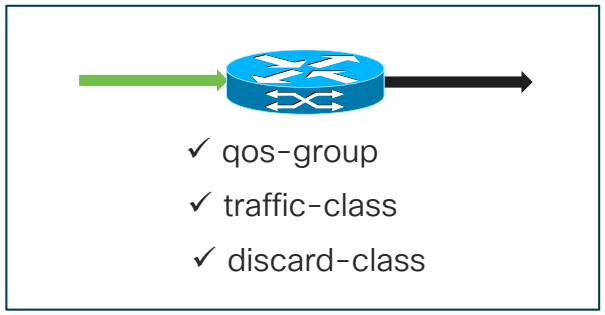


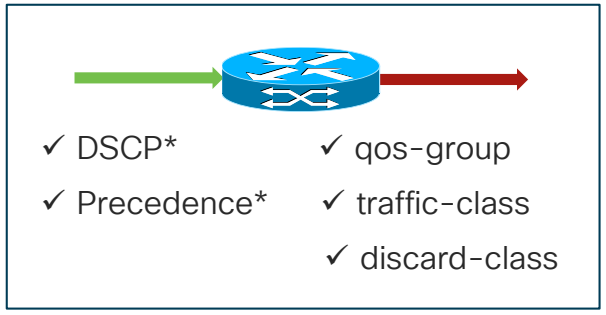












* With MPLS disposition- ingress short pipe mode

Policing Support

- Policer Types
  - 1R2C

    ```
    police rate 10 mbps burst 12 kbytes
    ```

  - 1R3C

    ```
    police rate 10 mbps burst 12 kbytes peak-burst 12 kbytes
    ```

  - 2R3C

    ```
    police rate 10 mbps burst 12 kbytes peak-rate 20 mbps
    peak-burst 12 kbytes
    ```

- Hierarchy
  - 2 Levels
  - Parent:
    - Class Default
    - Child conform-aware/non conform-aware
  - Child
    - Multiple class
    - Color blind

- Statistics
  - Normal Mode
    - Green and Red
  - Enhanced Mode*
    - `hw-module profile stats qos-enhanced`
    - Green, Yellow & Red

- MEF BWP Reference
- `police rate 10 mbps burst 12 kbytes peak-rate 20 mbps peak-burst 12 kbytes`

CIR CBS EIR** EBS

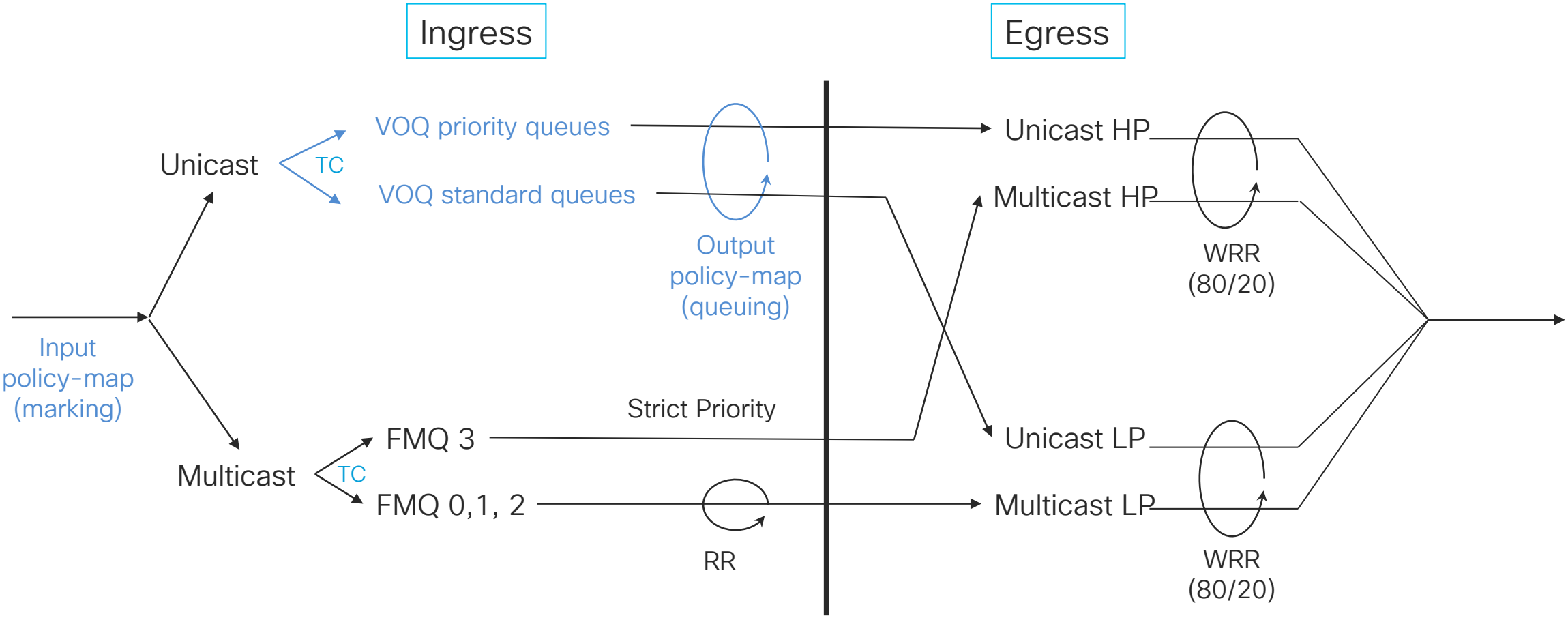
CM=blind CF=1

* Default on NCS 560

** peak-rate = CIR+EIR

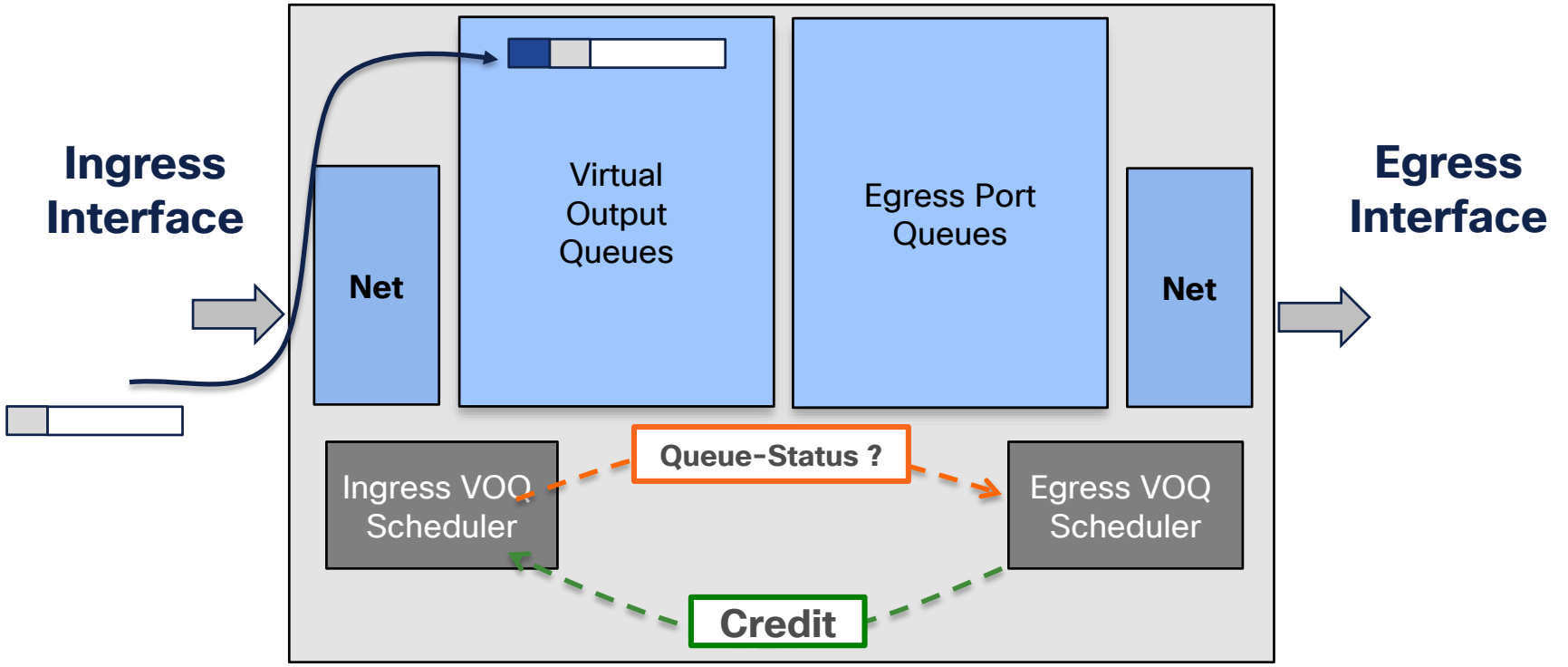
Egress QoS

Queues and Schedulers



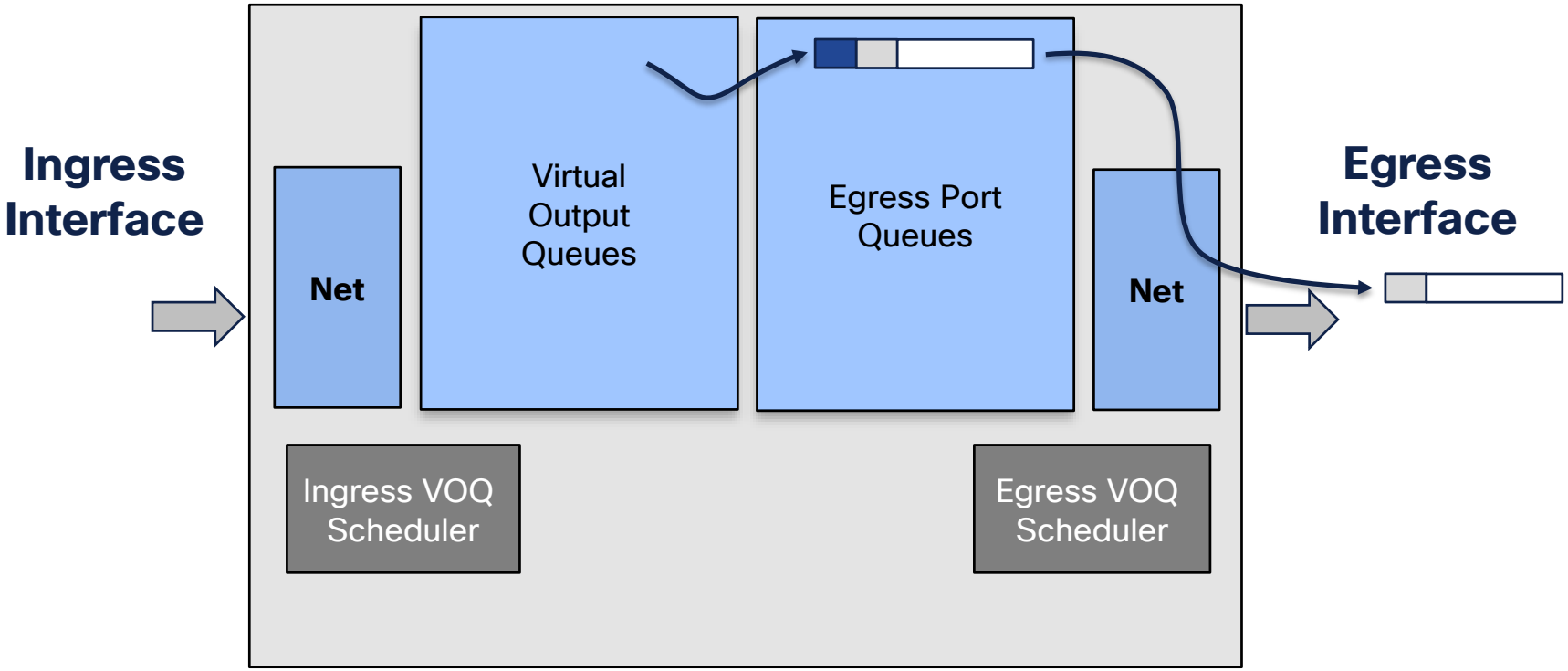
Blue items are done according to policy-maps. Black items are internal operations.

VOQ-Only Architecture



- Packet is received on ingress interface, classified, and stored in internal buffer
- Ingress VOQ scheduler polls Egress scheduler (maintaining a local VOQ DB)
- Egress answers with a credit-message

VOQ-Only Architecture contd.

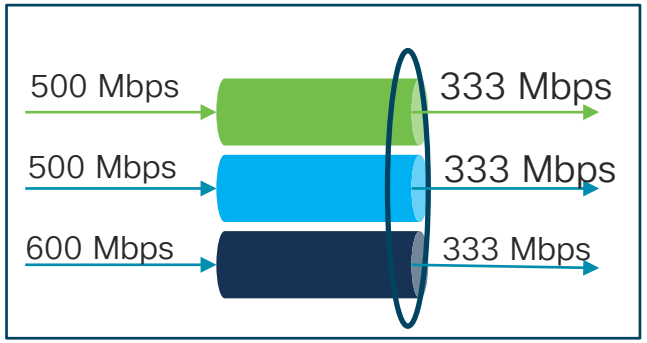


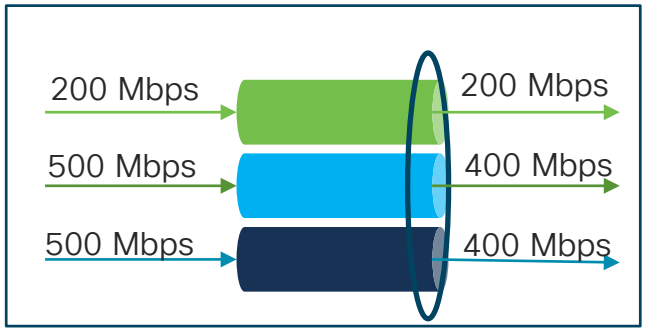
- Packet is stored in the port queue(HP/LP)
- Finally packet is transmitted through the egress interface

Scheduling Mechanism

Fair Scheduling

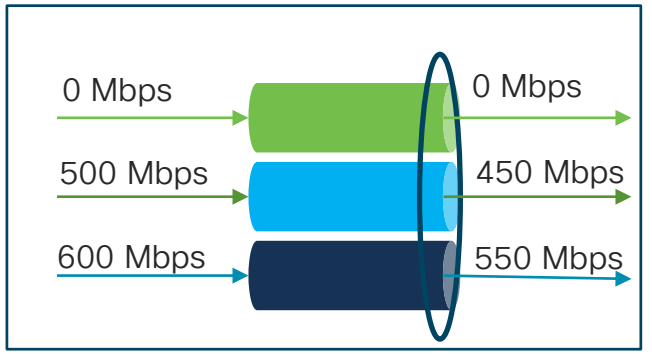
Profile 1





Profile 3

Profile 3



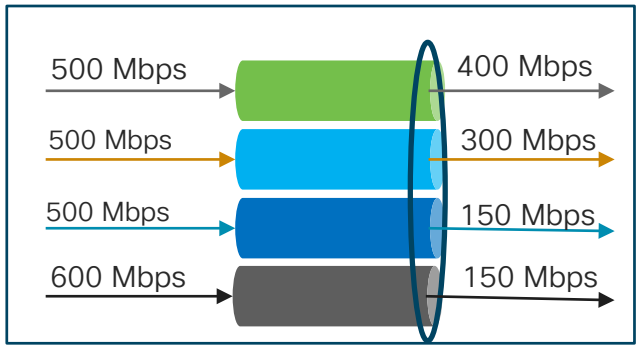

```
policy-map
fair_schedule
 class cos1
  shape average
percent 50
 !
 class cos2
  shape average
percent 45
 !
 class class-default
 !
 end-policy-map
!
```

- Shape average sets Maximum queue BW
- Traffic is fair among all classes
- Class without shape can consume all bandwidth available

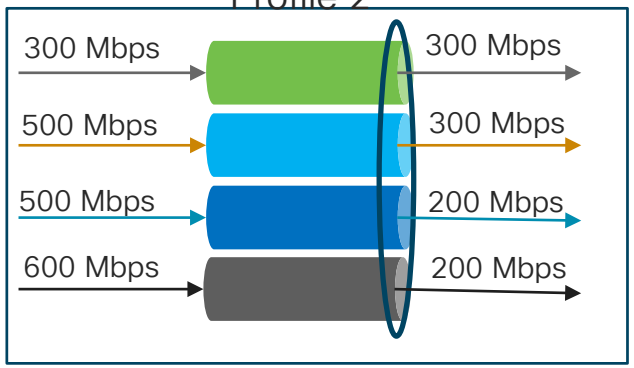
Note: all traffic profile explained assuming a 1Gbps reference BW

Priority Scheduling

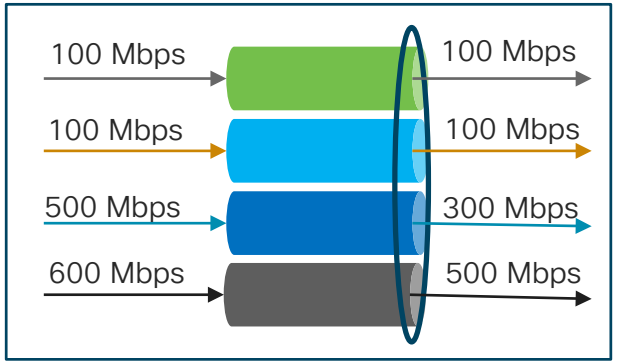
Profile 1



Profile 2



Profile 3



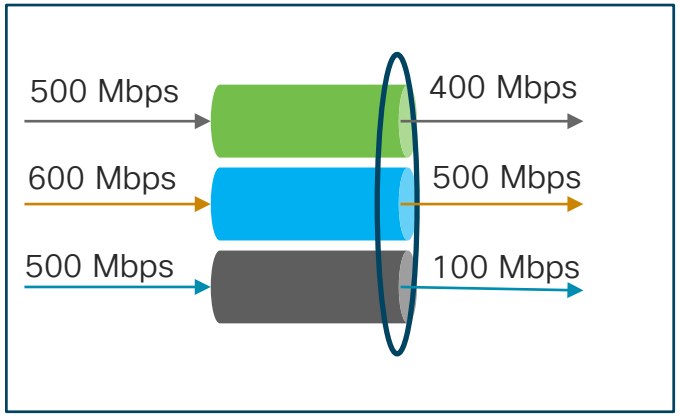

```
policy-map priority
 class cos1
  shape average percent 40
  priority level 1
  queue-limit 1 bytes
 !
 class cos2
  shape average percent 30
  priority level 2
  queue-limit 500 us
 !
 class cos3
  shape average percent 30
 !
 class class-default
 !
 end-policy-map
!
```

- Up to 8 priority levels in non h-qos mode
- Up to 4 priority levels in h-qos mode
- Minimum queue limit for latency sensitive priority traffic.
- Fair BW share for non priority classes as shape is configured

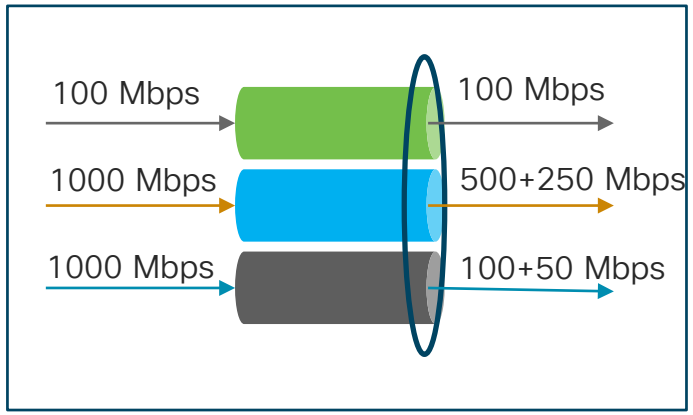
Note: all traffic profile explained assuming a 1Gbps reference BW

Guaranteed Service Rate

Profile 1



Profile 2



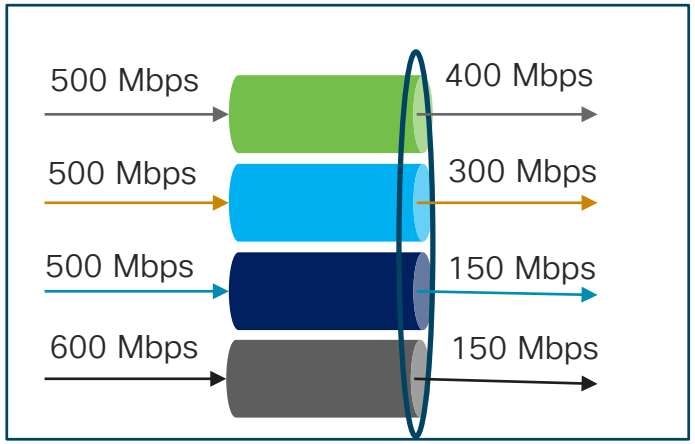

```
policy-map
guaranteed_sr
 class cos1
  shape average
percent 40
  priority level 1
  queue-limit 1 bytes
 !
 class cos2
  bandwidth percent 50
 !
 class class-default
  bandwidth percent 10
 !
 end-policy-map
!
```

- Bandwidth command sets queue minimum BW
- Inversely weighted bandwidth sharing
- Priority cannot be used in Bandwidth

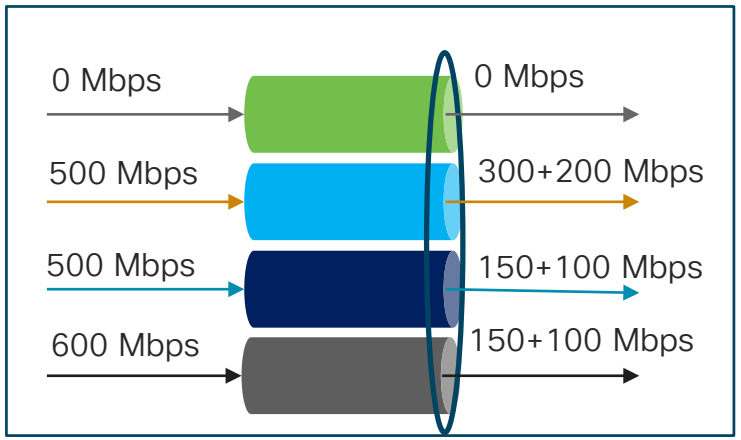
Note: all traffic profile explained assuming a 1Gbps reference BW

Weighted Queues with Bandwidth Remaining (BWR)

Profile 1



Profile 2



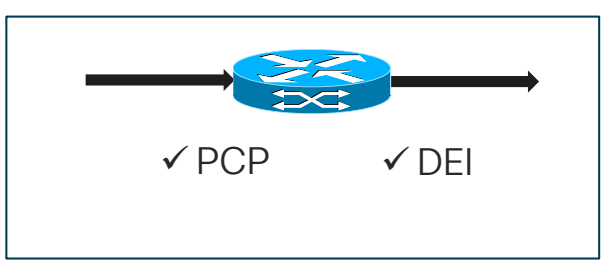

```
policy-map bwr_sr
 class cos1
  shape average percent 40
  queue-limit 1 bytes
  priority level 1
 !
 class cos2
  bandwidth remaining percent 50
 !
 class cos3
  bandwidth remaining percent 25
 !
 class class-default
 !
 end-policy-map
!
```

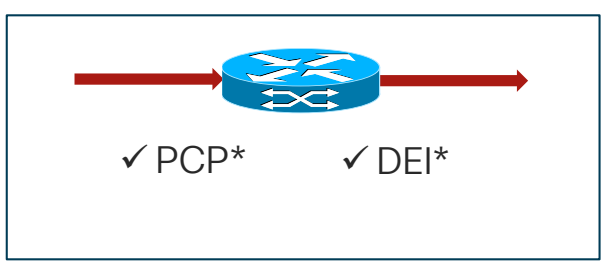
- Bandwidth remaining assigns weight for available bw sharing
- Both BW and BWR in a single policy map is not supported in HQoS mode

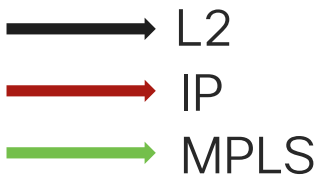
Note: all traffic profile explained assuming a 1Gbps reference BW

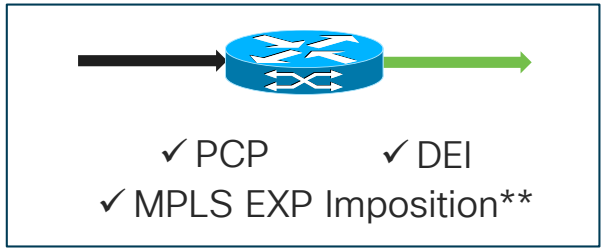
Egress Marking

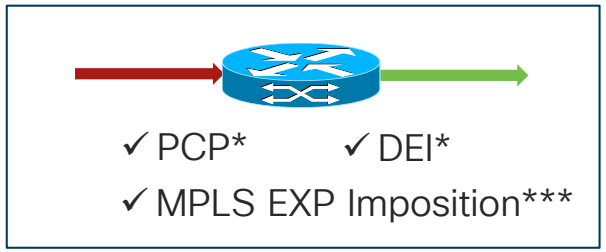
Marking Support

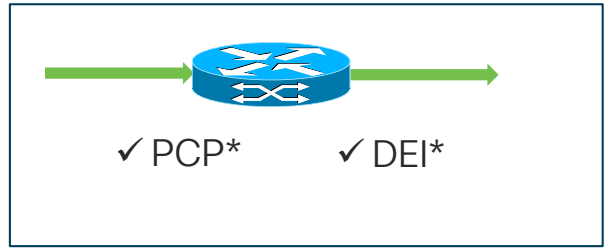


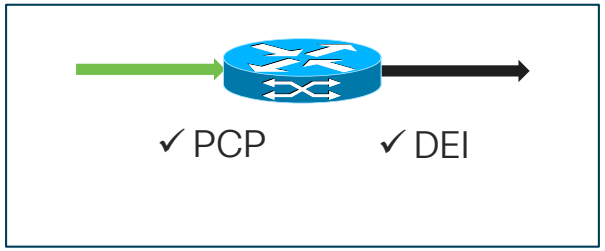


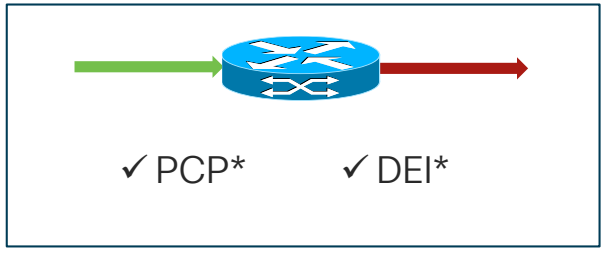










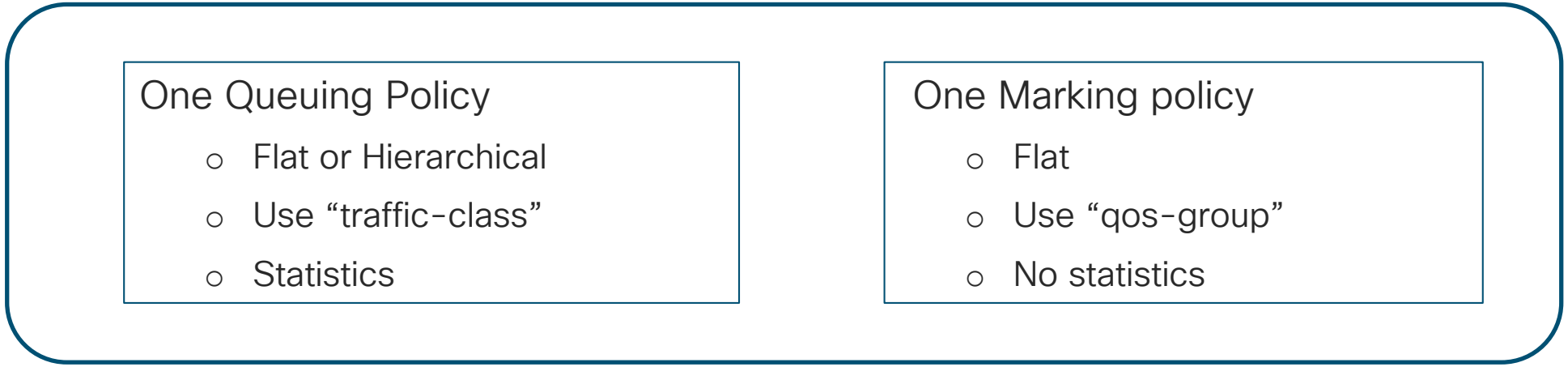


Note: Egress Remarking uses “qos-group” as classification criterion

* L3 sub-intf vlan priority fields
** Not for EVPN
*** 2 labels only, single BGP Label not supported

Dual Egress Policy
Queuing & Marking handled separately

- Can we Remark along with managing different queues ???



Dual Egress Policy: Example

```
class-map match-any cos1
   match cos 1
end-class-map
!
class-map match-any cos2
   match cos 2
end-class-map
!
policy-map in-classify
   class cos1
      set qos-group 1
      set traffic-class 3
  !
  class cos2
      set qos-group 2
      set traffic-class 5
  !
class class-default
!
end-policy-map
```

```
class-map match-any qos1
   match qos-group 1
end-class-map
!
class-map match-any qos2
    match qos-group 2
end-class-map
!
policy-map egress-marking
   class qos2
       set cos 2
   !
   class qos1
       set cos 1
       set dei 1
   !
class class-default
       set cos 7
   !
end-policy-map
```

```
class-map match-any tc3
    match traffic-class 3
end-class-map
!
class-map match-any tc5
    match traffic-class 5
end-class-map
!
policy-map egress-queuing
   class tc5
     bandwidth percent 30
   !
   class tc3
       bandwidth percent 20
   !
   class class-default
   !
end-policy-map
!
interface TenGigE0/0/1/0/0
 service-policy output egress-marking
 service-policy output egress-queuing
!
```

Use Cases

Session Takeaway

You can simply deliver QoS in your access network
with NCS 500 !!

Hierarchical QoS

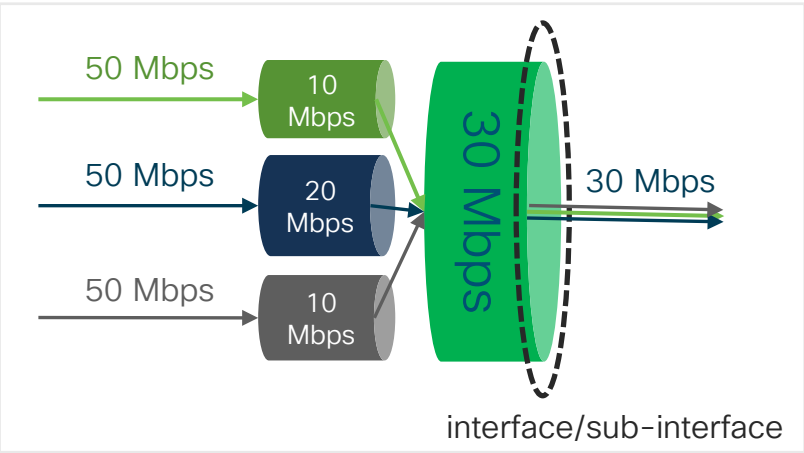
H-QoS Support Overview

- Ingress (Policing only)
  - 2 levels
  - Parent level:
    - class-default only
    - Child conform-aware Or non conform-aware
  - Child level:
    - Multiple Classes of Service
    - Color blind 2r3c

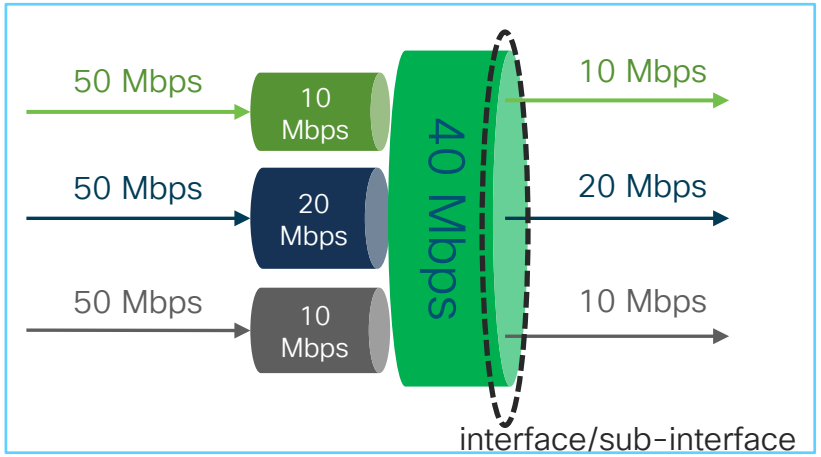
- Egress (Queueing Only)
  - 3 (1+2) levels starting XR 6.6.25
  - 1 level Grand Parent
    - class default shaper at port
  - 2 level nested at sub-interface
    - Parent:
      - Class Default only
    - Child:
      - Multiple Classes of Service (Upto 8)
      - Priority levels
  - "`hw-module profile qos hqos-enable`" config is necessary

H-QoS for Ingress Policing

- 2 levels of hierarchy
- Parent level supports class-default
- Parent policer is conform-aware/non conform-aware
- To achieve expected result: ∑child CIR <= Parent CIR



Profile with ∑child CIR > Parent CIR
Parent non-conform aware



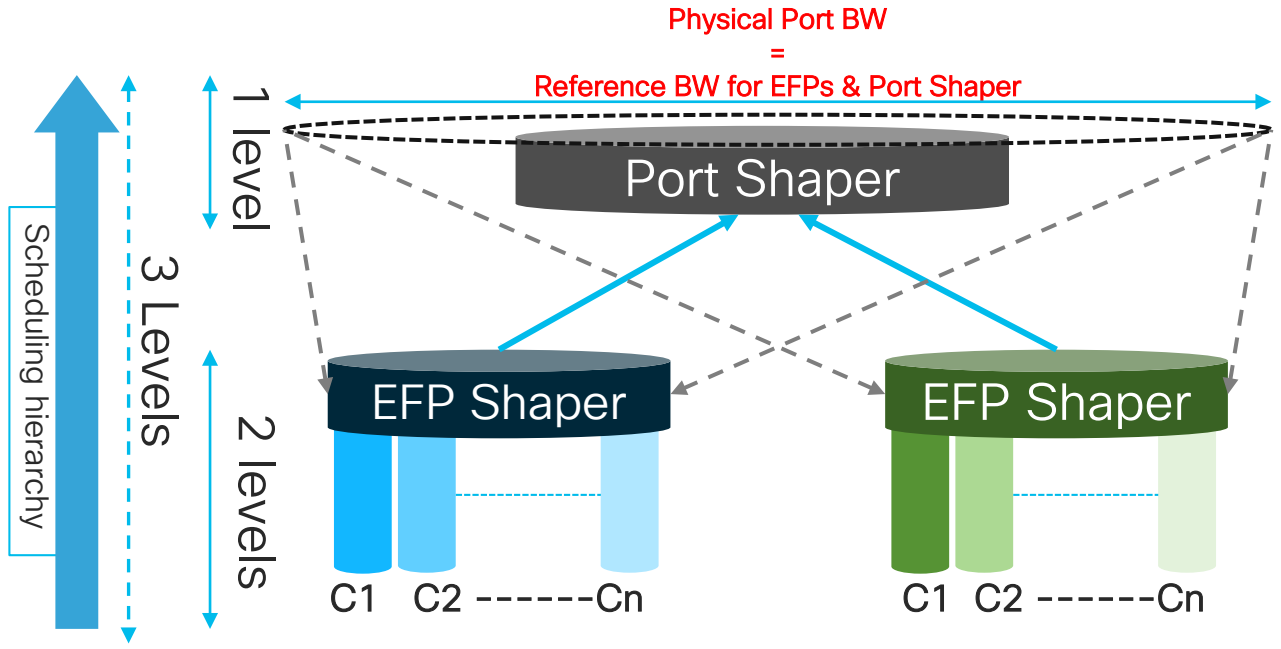
Profile with ∑child CIR < Parent CIR

```
policy-map parent-policer
 class class-default
  service-policy child-policer
  police rate 40 mbps
  !
 !
 end-policy-map
!

policy-map child-policer
 class cos3
  police rate 10 mbps
  !
 !
 class cos4
  police rate 20 mbps
  !
 !
 class class-default
  police rate 10 mbps
  !
 !
 end-policy-map
!
```

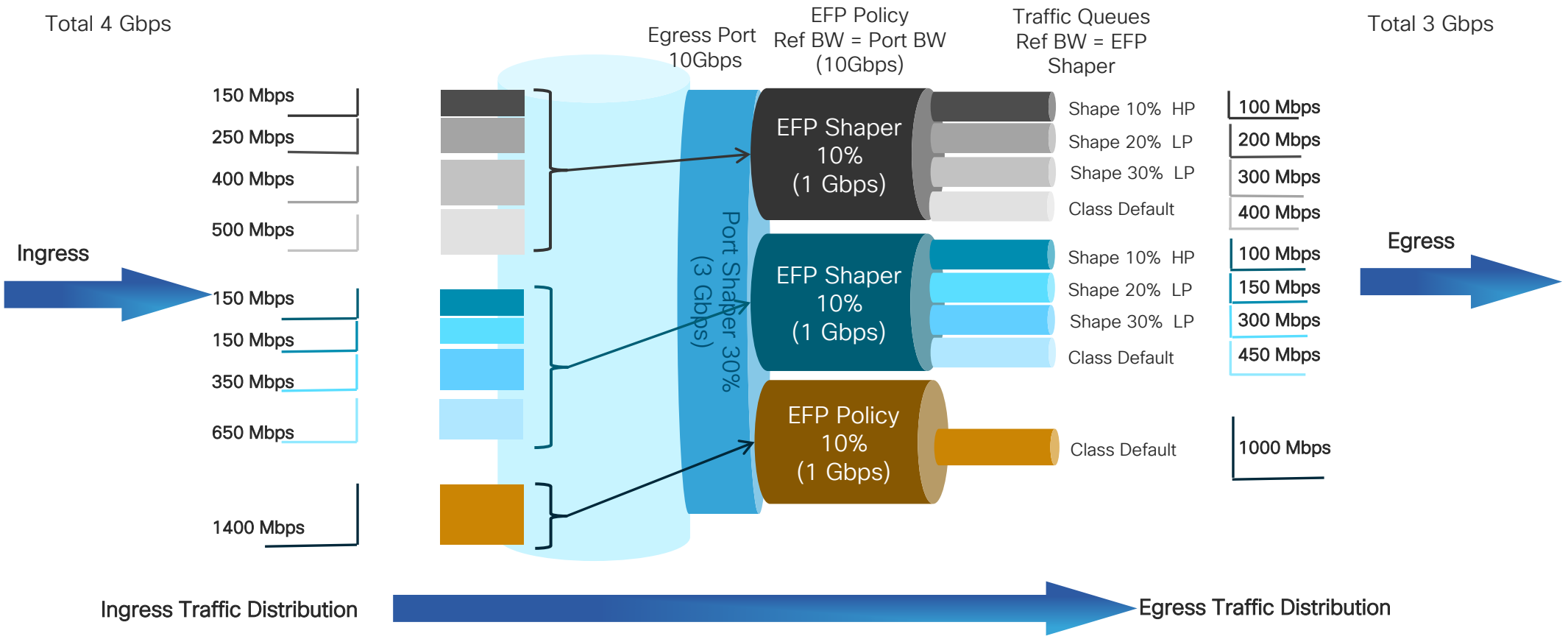
3-level H-QoS for Egress Queueing

- Implemented in an 1+2 Model
- Port shaper at Grand Parent Level (1 level class default only)
- EFP shaper with 2 levels
  - Parent “class default” only
  - Child multiple traffic classes
- EFP shaper reference BW is same as Port BW
  - Allows to oversubscribe Port/Port Shaper



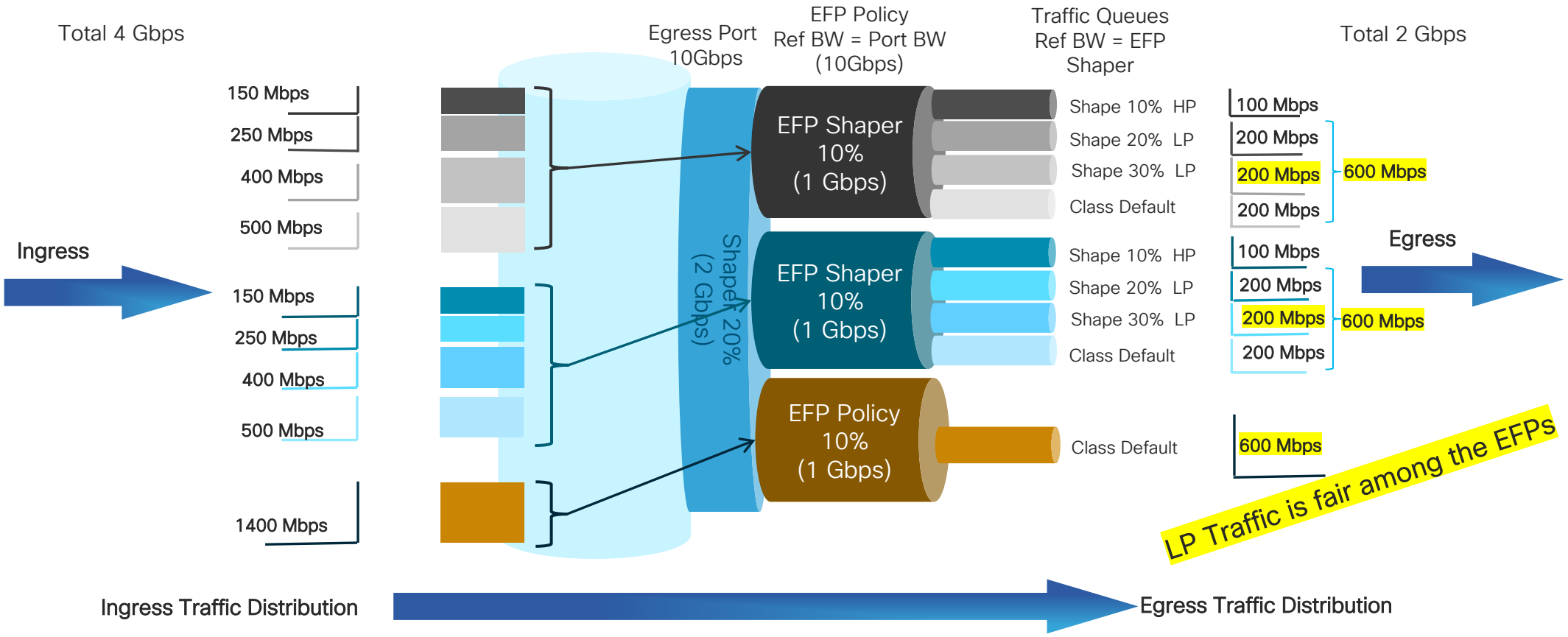
3- Level H-QoS Behavior

Scenario1: ∑(Child shape rate) < Port Shaper



3-Level H-QoS Behavior

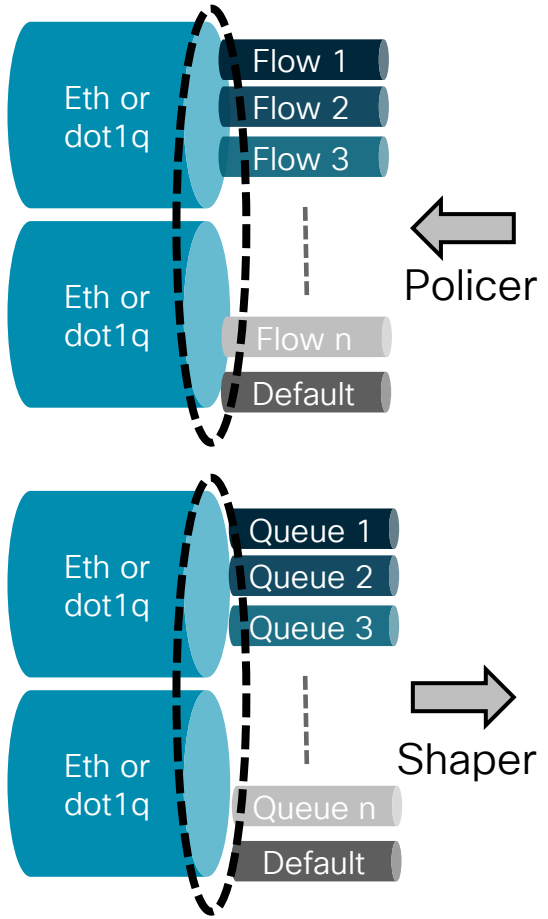
Scenario2: ∑(Child shape rate) > Port Shaper



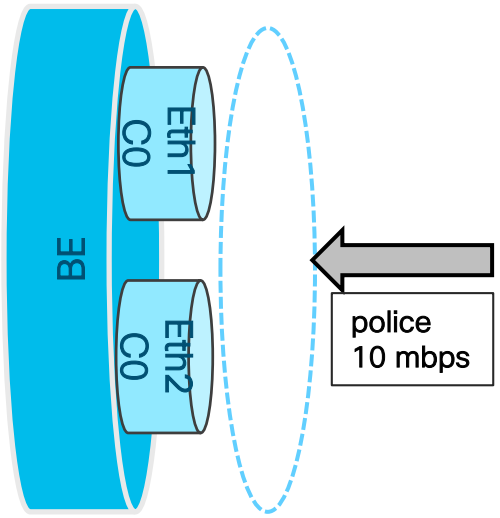
QoS on Bundles

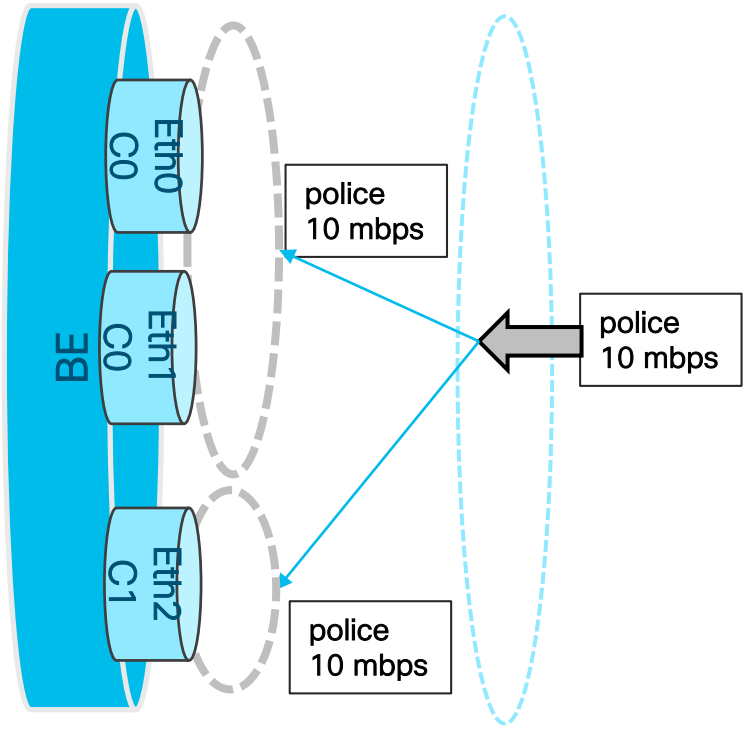
Bundle QoS

- Ingress Support
  - Policing only
  - Aggregate Behavior per core
- Egress Support
  - Queuing /Scheduling only
  - Distributed Behavior
  - Same policy replicated for all members
- QoS policy is configured under LAG bundle interface
  - The actual configured QOS policy is programmed onto each link bundle member (port)



Bundle QoS: Ingress Policing





NPU Core#	NCS 560-7 IM Slots#	NCS 560-4 IM Slots#
0	0-7	0,2,4
1	8-15	1,3,5

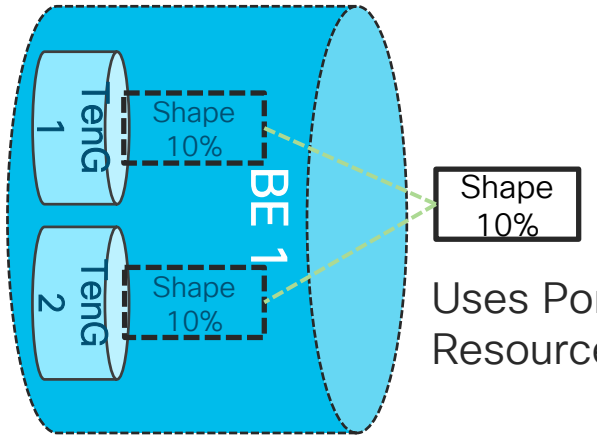
CLI to identify NPU core distribution:

```
show controllers npu voq-usage interface all
instance all location <>
```

- Aggregate policer Per Core (absolute rate allowed)
- Members on the same NPU core
  - aggregate behavior
  - Consume single Policer entry
- Members across different NPU core
  - No Aggregate Behavior
  - Policer consumption is doubled

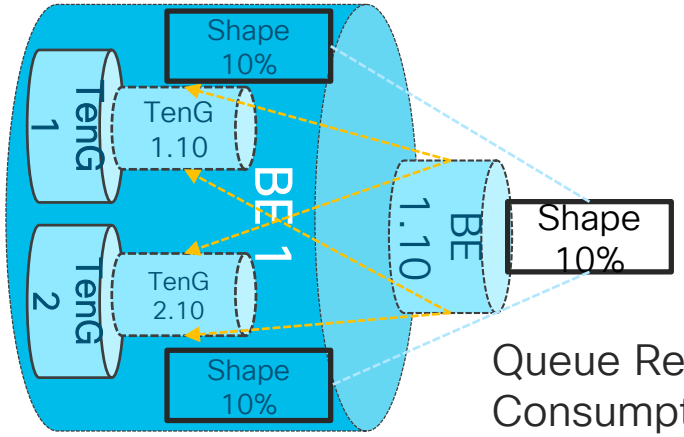
Bundle QoS: Egress Shaping

- Aggregate Rates not allowed/Recommended
  - Percent/per thousand shape rates allowed
  - Policy replicated into each member
  - Aggregate Behavior achieved of traffic is load balanced evenly



Uses Port Default Queue Resources

Shaping
bundle main interface

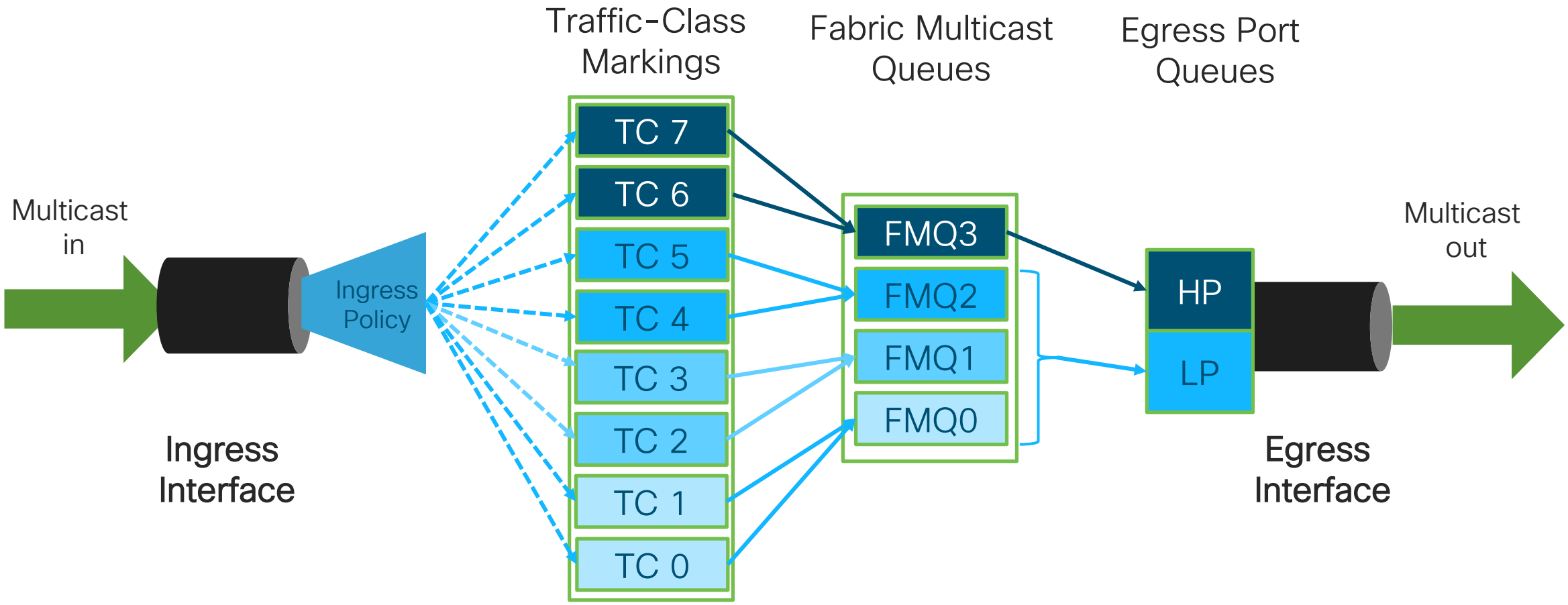


Queue Resources Consumption is multiplied

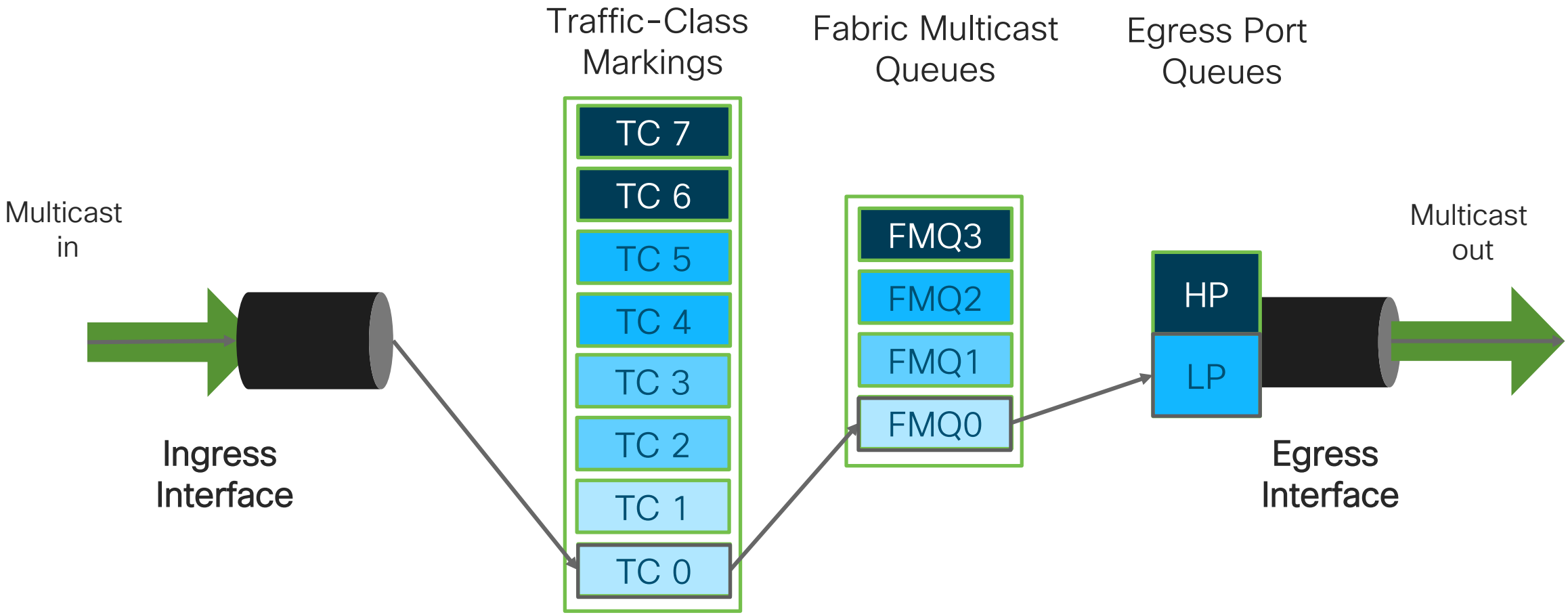
Shaping
bundle sub-interface

Managing Multicast QoS

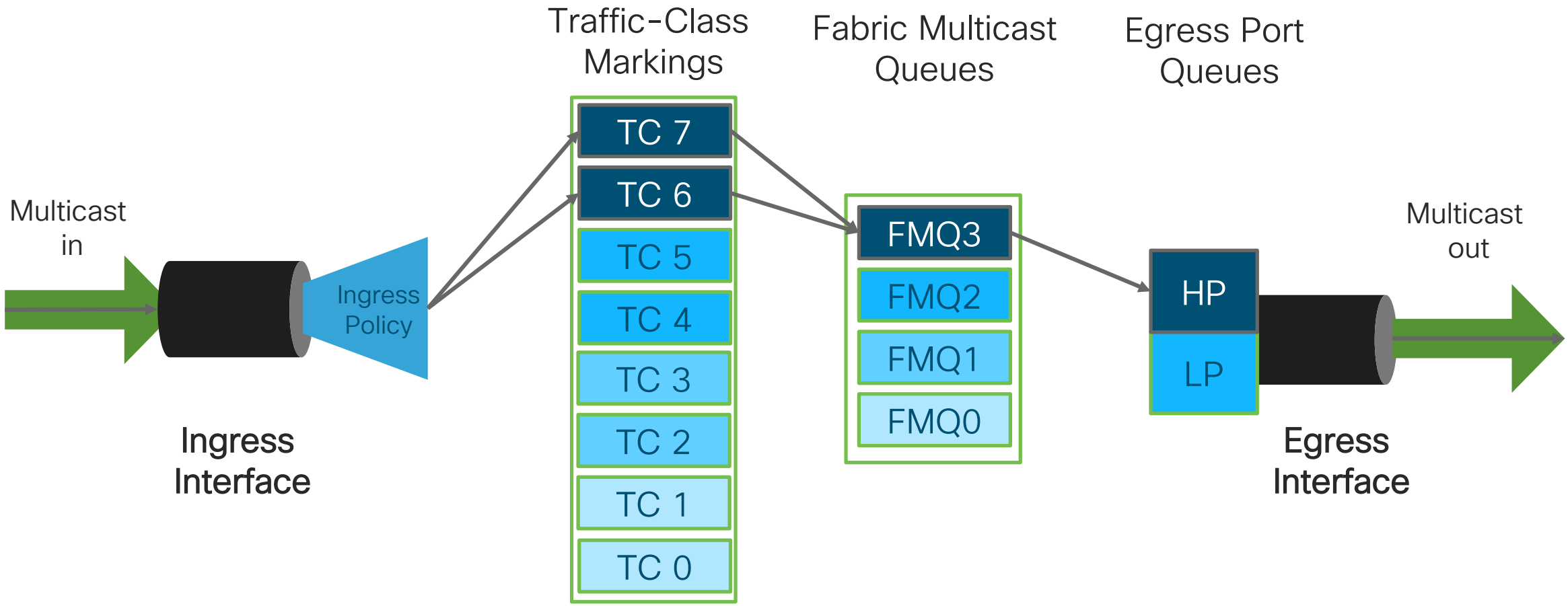
Fabric Multicast Queues (FMQ)
Multicast Packet Path



Default Behavior
No Ingress Policy



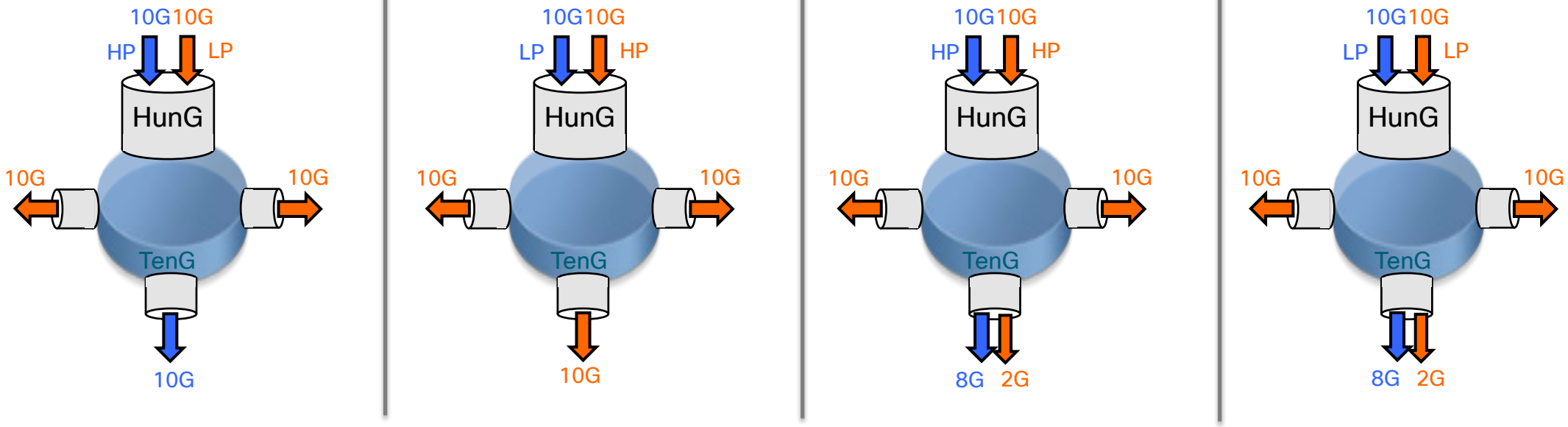
Fabric Multicast Queues (FMQ)
Managing Priority Multicast



Unicast vs Multicast

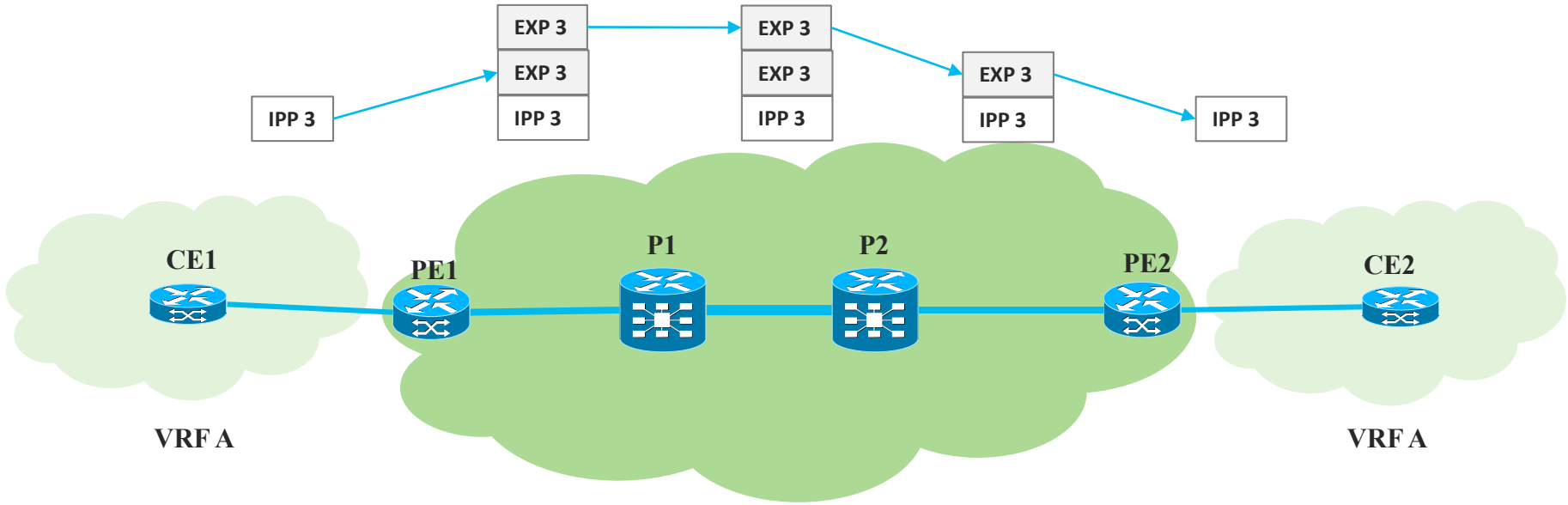
Unicast
Multicast

- HP Traffic is always preferred
- When in competition for same priority level 80:20 ratio maintained



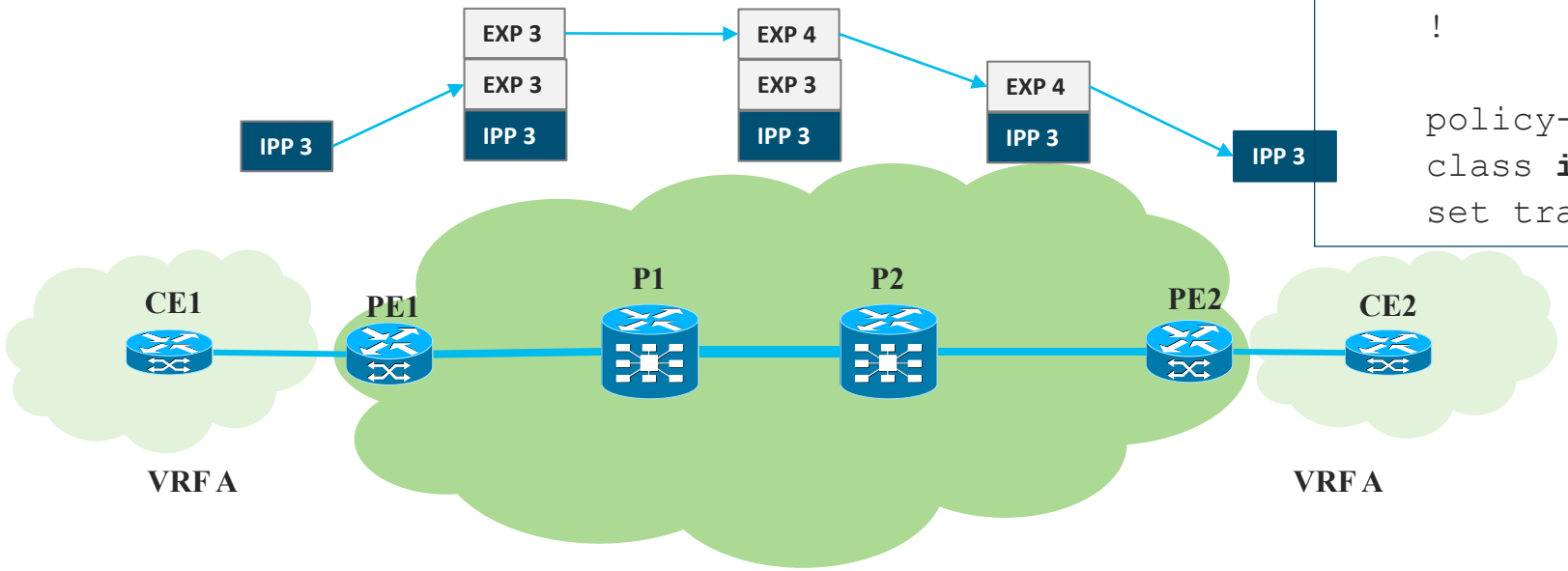
Diffserv Tunneling Modes

Uniform Mode



- IP Precedence and MPLS EXP is same across the Network.
- When a label is popped, PHB value is copied to new exposed label/DSCP
- Default Behavior on NCS 5xx

Pipe Mode

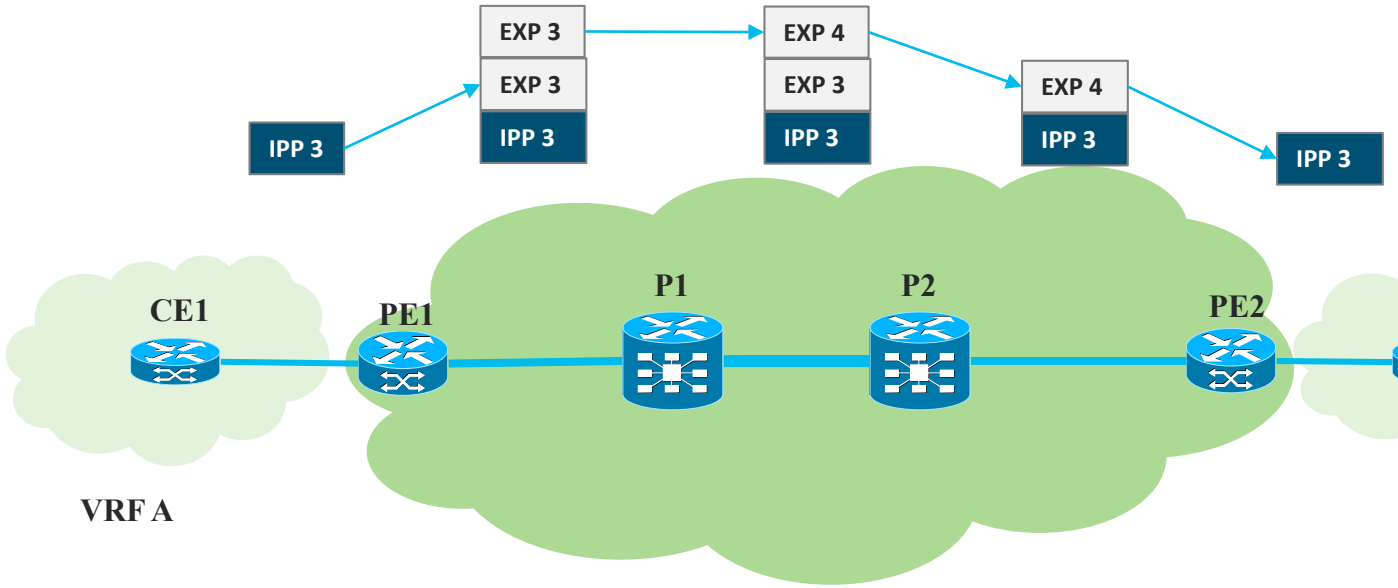



```
class-map match-any in_pipe
 match mpls exp topmost 4
 end-class-map
!

policy-map ingress-classifier
class in_pipe
set traffic-class 5
```

- DSCP & Precedence Preserved from CE to CE
- At egress PE , EXP-Precedence copy must be disabled using "`mpls-ip-ttl-propagate disable`"
- EXP=map (EXP) at P nodes
- EXP value is used for queuing decisions at Egress PE

Short Pipe Mode




```
class-map match-any in_pipe
 match mpls disposition class-
map child_pipe
 end-class-map
!
class-map match-any child_pipe
 match precedence 3
end-class-map
!

policy-map ingress-classifier
class in_pipe
set traffic-class 5
```

- DSCP/Precedence Preserved from CE to CE
- At egress PE , EXP-Precedence copy must be disabled using "`mpls-ip-ttl-propagate disable`"
- EXP=map (EXP) at P nodes
- DSCP/Precedence value is used for queuing/scheduling decisions at Egress PE
- DSCP/Precedence match at egress using "`match mpls disposition`"

QOS behavioral differences between ASR9xx and NCS5xx

Key differences –Summary

ASR 920/RSP2

- Classifications:
  - Allows egress classification based on packet header information
  - Allows match EFP (vlan)
- Policing:
  - Policing supported with priority at Egress
- Priority Queuing:
  - 2 levels only (P1&P2)

RSP3

- Classifications:
  - Egress classification based on “qos-group”
  - Allows match EFP (vlan)
- Policing:
  - Only on ingress
- Priority Queuing:
  - 2 levels only(P1&P2)

NCS 5xx

- Classifications:
  - Egress classification based on “qos-group”, “traffic-class”, “discard-class”
- Policing:
  - Only on ingress
- Priority Queuing:
  - 7 levels (P1-P7)

Policer Implementation (2r3c)

- In ASR9xx Policer default action can be changed by configuration
  - Default action :
    - Conform: transmit Exceed: drop Violate: drop
- In NCS 5xx Policer default action cannot be changed
  - Default action :
    - Conform: transmit (discard-class 0) Exceed: transmit (discard-class 1) Violate: drop

ASR 9xx

```
policy-map 2r3c
 description 2R3C policer on ASR9xx
 class class-default
  police cir 3000000 pir 6000000
   conform-action transmit
   exceed-action transmit
   violate-action drop
!
end
```

NCS 5xx

```
policy-map 2r3c
 description 2R3C policer on ncs5xx
 class class-default
  police rate 3 mbps peak-rate 6 mbps
  !
 !
 end-policy-map
!
```

Implementing conditional marking

- ASR 9xx uses single policy in ingress for Conditional marking
- NCS 5xx uses ingress and egress policy simultaneously to achieve same results
- NCS 5xx uses qos-group and discard-class together to mark traffic based on policing

ASR 9xx

```
policy-map conditional-mark
 description mark PCP based on policing results ASR9xx
 class class-default
  police cir 3000000 pir 6000000
   conform-action set-cos-transmit 5
   exceed-action set-cos-transmit 4
   violate-action drop
!
end
```

NCS 5xx

```
policy-map conditional-mark
 description policer marks qos group on ingress-NCS5xx
 class class-default
  police rate 3 mbps peak-rate 6 mbps
  !
  set qos-group 5
 !
 end-policy-map
!
```

```
class-map match-all exceed-qos
 match qos-group 5
 match discard-class 1
 end-class-map
!
class-map match-all conform-qos
 match qos-group 5
 match discard-class 0
 end-class-map
!
policy-map conditional-mark
 class conform-qos
  set cos 5
 !
 class exceed-qos
  set cos 4
 !
 class class-default
 !
 end-policy-map
!
```

Managing priority traffic-class

- Objective is to have a low latency priority traffic and restrict it to 25% of the link bw
- Priority traffic is identified by DSCP value EF

ASR 920/RSP2

```
class-map match-all dscp
 match dscp ef
!


policy-map egress_priority
 class dscp
  priority level 1 percent 25
  queue-limit <value>
!


policy-map priority-policing
 class dscp
  police cir percent 25
  queue-limit <value>
  priority
!
end
```

RSP3

```
class-map match-any ingress_dscp
 match dscp ef
 end-class-map
!
policy-map set_qg_from_dscp
 class ingress_dscp
  set qos-group 5
 !
 class class-default
 !
 end-policy-map
!
class-map match-any egress_qg
 match qos-group 5
 end-class-map
!
policy-map egress_priority
 class egress_qg
  priority level 1 percent 25
  queue-limit <value>
 class class-default
!
end
```

NCS 5xx

```
class-map match-any ingress_dscp
 match dscp ef
 end-class-map
!
policy-map set_tc_from_dscp
 class ingress_dscp
  set traffic-class 5
 !
 class class-default
 !
 end-policy-map
!
class-map match-any egress_tc
 match traffic-class 5
 end-class-map
!
policy-map egress_priority
 class egress_tc
  shape average percent 25
  priority level 1
  queue-limit <value>
 !
 class class-default
 !
 end-policy-map
!
end
```

Implementing hierarchical QoS

- ASR 9xx implements 3 level HQoS using nested policies
- NCS 5xx implements 3 level HQoS using 1+2 model

ASR920/RSP2

- Grand-Parent policy
  - Class-default only
  - Applied to Port
- Parent Policy
  - match-efp
  - referred in grand-parent policy
- Child Policy
  - Match dscp/cos/qos-group
  - Referred in parent policy

RSP3

- Grand-Parent policy
  - Class-default only
  - Applied to Port
- Parent Policy
  - match-efp
  - referred in grand-parent policy
- Child Policy
  - Match qos-group
  - Referred in parent policy

NCS 5xx

- Grand-Parent policy
  - Class-default only
  - Applied to Port
- Parent Policy
  - Class-default only
  - Applied to sub-interface (EFP)
- Child Policy
  - Match traffic-class
  - Referred in parent policy

Examples in next two slides

Example : Implementing hierarchical QoS: RSP3 (XE)

- ‘qos-group’ is set using ingress-policy

```
class-map match-any qgroup1
 match qos-group 1

class-map match-any efp2
 match service instance ethernet  2
class-map match-any efp1
 match service instance ethernet  1

policy-map egress-child
 class qgroup1
  priority percent 30
!


policy-map egress-efp
 class efp1
  shape average 200m
   service-policy egress-child
 class efp2
  shape average 300m
   service-policy egress-child
!
```

```
policy-map 3-hqos
 class class-default
  shape average 800m
   service-policy egress-efp
!


interface GigabitEthernet0/1/2
 no ip address
 negotiation auto
 service-policy output 3-hqos
 service instance 1 ethernet
  encapsulation dot1q 1
  bridge-domain 1
 !
 service instance 2 ethernet
  encapsulation dot1q 2
  bridge-domain 2
 !
```

Example: Implementing hierarchical QoS: NCS 500 (XR)

- ‘traffic-class’ is set using ingress-policy

```
class-map match-any tc1
 match traffic-class 1

policy-map egress-child
 class tc1
  shape average percent 30
  priority level 1
 !
 class class-default
 !
 end-policy-map
!

policy-map egress-efp1
 class class-default
  service-policy egress-child
  shape average 200 mbps
 !
 end-policy-map
!

policy-map egress-efp2
 class class-default
  service-policy egress-child
  shape average 300 mbps
 !
 end-policy-map
!
policy-map grand-parent
 class class-default
  shape average 800 mbps
 !
 end-policy-map
!

interface GigabitEthernet0/0/0/2
 service-policy output grand-parent
 negotiation auto
!
interface GigabitEthernet0/0/0/2.1 l2transport
 encapsulation dot1q 1
 service-policy output egress-efp1
!
interface GigabitEthernet0/0/0/2.2 l2transport
 encapsulation dot1q 2
 service-policy output egress-efp2
!
```

Implementing QoS Short-Pipe: ASR 920/RSP2 (XE)

- Egress policy can match dscp / precedence directly
- Egress queuing is based on dscp / precedence

```
class-map match-any dscp_ef
 match dscp ef

policy-map egress-child
 class dscp_ef
  shape average percent 30
  queue-limit 20000 us
 !
 class class-default
 !
 end-policy-map
!
```

Implementing QoS Short-Pipe: NCS 500 (XR)

- Ingress policy uses **mpls-disposition** to match dscp /precedence
- ‘traffic-class’ is set using ingress-policy
- Egress queuing is based on traffic-class match (tc=map[dscp])

```
class-map match-any child_pipe
 match dscp ef
end-class-map
!

class-map match-any in_pipe
 match mpls disposition class-map
child_pipe
 end-class-map
!

policy-map ingress-classifier
class in_pipe
set traffic-class 5
```

```
class-map match-any tc5
 match traffic-class 5
 end-class-map
!

policy-map egress-child
 class tc5
  shape average percent 30
  queue-limit 20ms
 !
 class class-default
 !
 end-policy-map
!
```

Summary

Summary

- QoS Requirements for Access
  - Police/shape/schedule
  - Hierarchy
- How NCS 500 Delivers QoS
  - Police & Remark on ingress
  - Shape/Schedule and remark on egress
  - Multiple levels of hierarchy
- Compared to ASR900
  - QoS delivery is intact
  - Change in configuration/implementation due to ASIC/OS difference

Session Takeaway

You can simply deliver QoS in your access network
with NCS 500 !!

Complete your online session survey

- Please complete your session survey after each session. Your feedback is very important.
- Complete a minimum of 4 session surveys and the Overall Conference survey (starting on Thursday) to receive your Cisco Live t-shirt.
- All surveys can be taken in the Cisco Events Mobile App or by logging in to the Content Catalog on ciscolive.com/emea.

Cisco Live sessions will be available for viewing on demand after the event at ciscolive.com.

cisco Live!

Continue your education

Demos in the Cisco Showcase

Walk-In Labs

Meet the Engineer 1:1 meetings

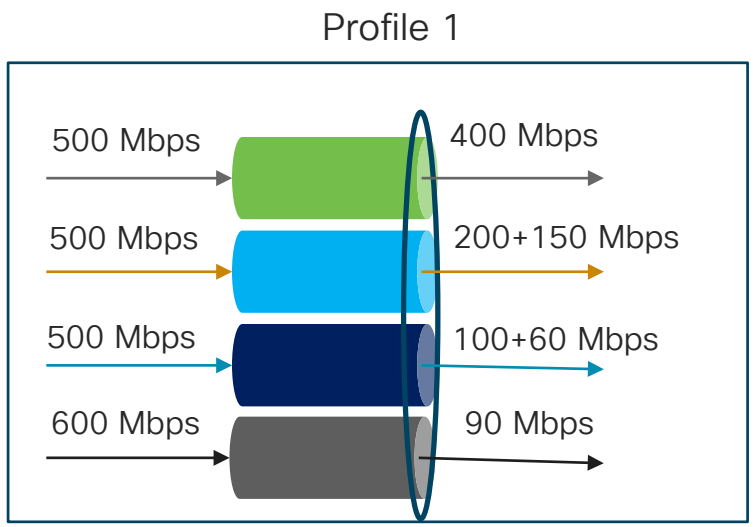
Related sessions

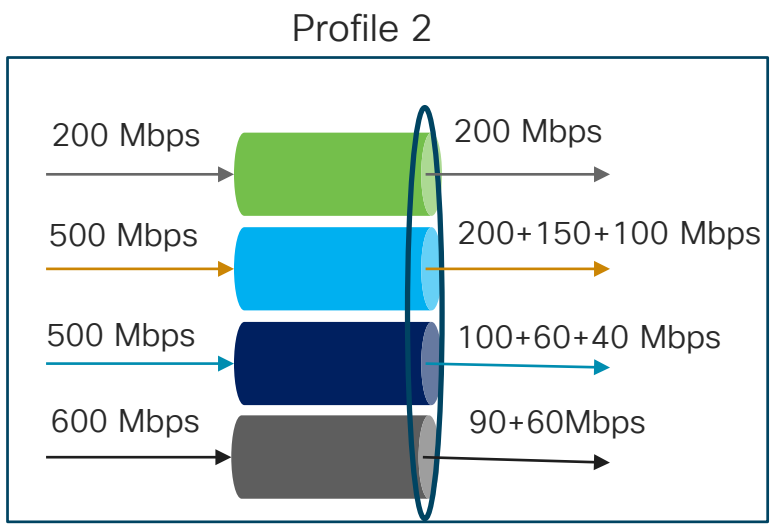

Thank you
cisco Live!

You make **possible**

Appendix

Guaranteed Service Rate with Weight





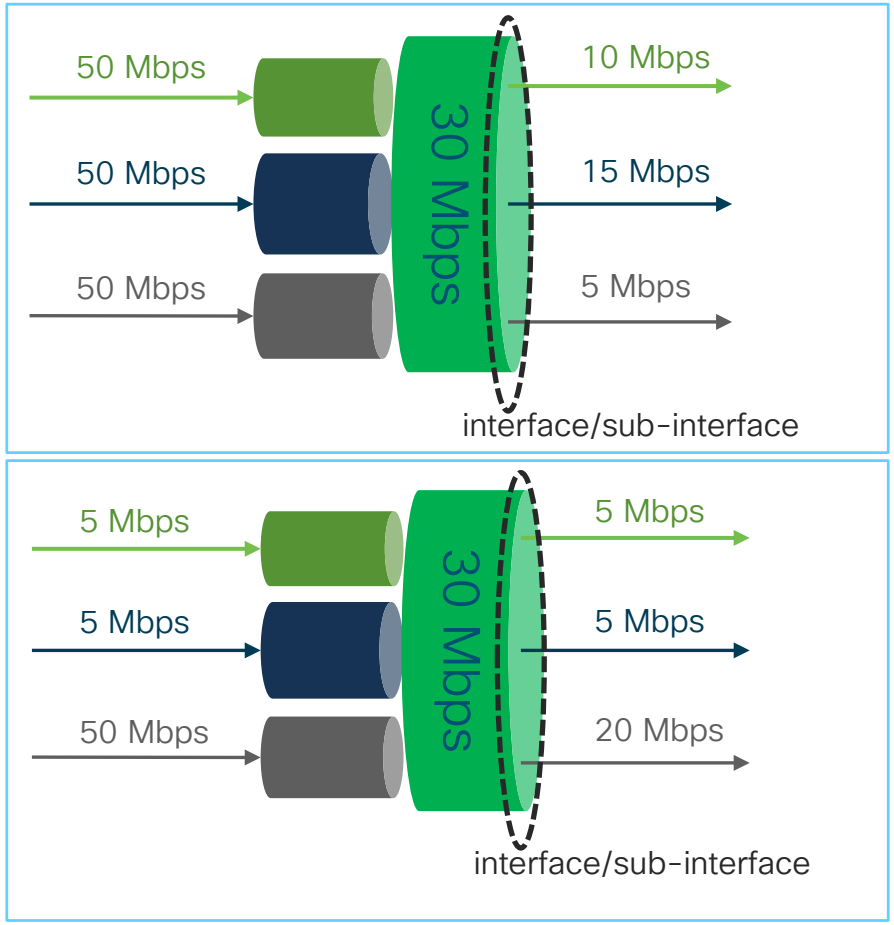

```
policy-map gsr_bwr
 class cos1
  shape average percent 40
  queue-limit 1 bytes
  priority level 1
 !
 class cos2
  bandwidth remaining percent 50
  bandwidth percent 20
 !
 class cos3
  bandwidth remaining percent 20
  bandwidth percent 10
 !
 class class-default
 !
 end-policy-map
!
```

- Bandwidth remaining assigns weight for available bw sharing
- Both BW and BWR in a single policy map is not supported in HQoS mode

Note: all traffic profile explained assuming a 1Gbps reference BW

Parent Conform Aware-Ingress Policing

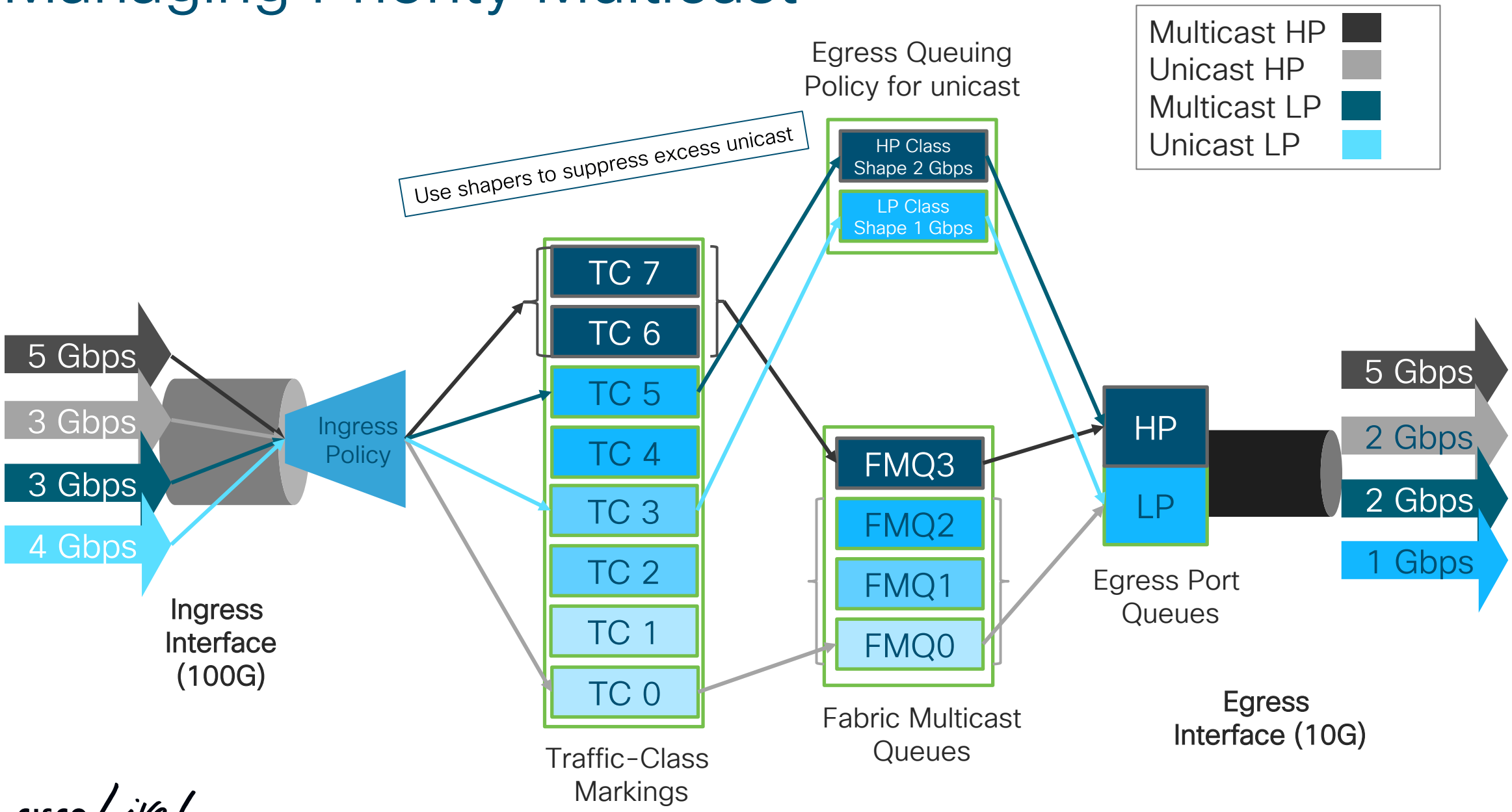
- ∑child CIR <= Parent CIR
- ∑child CIR+ ∑child EIR > Parent CIR




```
policy-map parent-policer
 class class-default
  service-policy child-policer
  police rate 30 mbps
  !
 !
 end-policy-map
!
policy-map child-policer
 class cos3
  police rate 10 mbps peak-rate 30 mbps
  !
 !
 class cos4
  police rate 15 mbps peak-rate 30 mbps
  !
 !
 class class-default
  police rate 5 mbps peak-rate 30 mbps
  !
 !
 end-policy-map
!

hw-module profile qos conform-aware-policer
```

Managing Priority Multicast



You make **possible**