Quality of Service (QoS)

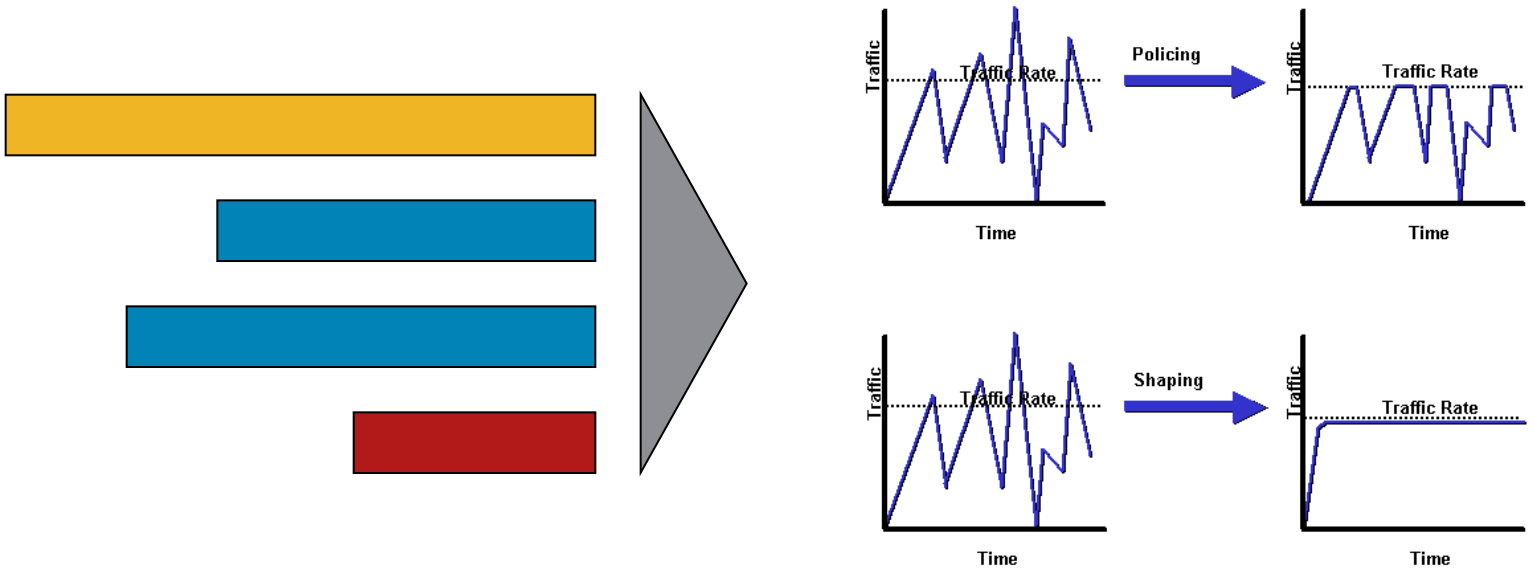
Ing. Pedro Escudero

Telf: 0994667184
Mail: pedro.escudero@unach.edu.ec
Web: https://www.researchgate.net/profile/Pedro-Escudero-3/research

Mayo – 2025

What is Quality of Service

- Quality of service is the key term to represent the set of requirements which a user (person or software component) wants to impose to his traffic.

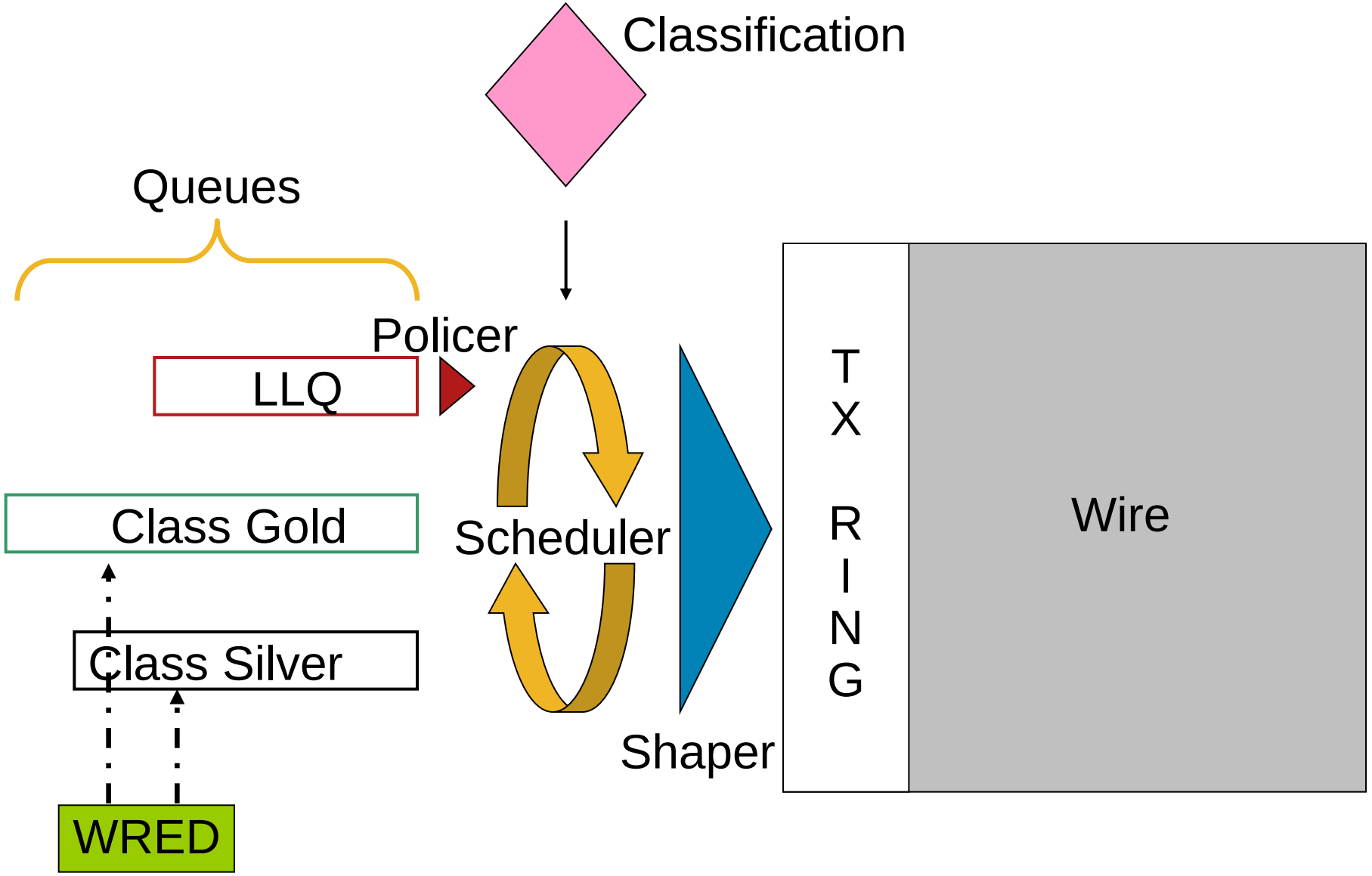


- Multiple mechanisms are available to fulfill these requirements.

Why do we need them ?

- Together, they provide tools that will allow us to define which action needs to be taken in case of congestion.

- Understanding where they fit, what they exactly do, and how they work together do is necessary to:
  - Correctly design and deploy QoS
  - Monitor its activity
  - Troubleshoot potential issues


Classification
Queues
Policer
LLQ
Class Gold
Class Silver
WRED
Scheduler
Shaper
TX RING
Wire

```
show policy-map interface
 Ethernet0/0

  Service-policy output: parent

    Class-map: class-default (match-any)
      4983 packets, 240814 bytes
      5 minute offered rate 0 bps, drop rate 0 bps
      Match: any
      Traffic Shaping
           Target/Average   Byte   Sustain   Excess     Interval  Increment
             Rate           Limit  bits/int  bits/int   (ms)      (bytes)
          4000000/4000000   25000  100000    100000     25        12500

        Adapt  Queue     Packets   Bytes     Packets   Bytes     Shaping
        Active Depth                        Delayed   Delayed   Active
        -      0         4         240       0         0         no

      Service-policy : child

        Class-map: one (match-all)
          575 packets, 10435 bytes
          5 minute offered rate 0 bps, drop rate 0 bps
          Match: ip precedence 5
          Match:  dscp default (0)
          Queueing
            Strict Priority
            Output Queue: Conversation 136
            Bandwidth 256 (kbps) Burst 6400 (Bytes)
            (pkts matched/bytes matched) 0/0
            (total drops/bytes drops) 0/0
```

```
  Class-map: two (match-all)
    254 packets, 16743 bytes
    5 minute offered rate 2000 bps, drop rate 0 bps
    Match: any
    Match:  dscp 1
    Queueing
      Output Queue: Conversation 137
      Bandwidth 512 (kbps) Max Threshold 64 (packets)
       (pkts matched/bytes matched) 2/0
   (depth/total drops/no-buffer drops) 0/0/0

   Class-map: class-default (match-any)
    1 packets, 60 bytes
    5 minute offered rate 0 bps, drop rate 0 bps
    Match: any
    Queueing
      Flow Based Fair Queueing
      Maximum Number of Hashed Queues 128
   (total queued/total drops/no-buffer drops) 0/0/0
       exponential weight: 9

class     Transmitted       Random drop       Tail drop     Minimum Maximum  Mark
          pkts/bytes        pkts/bytes        pkts/bytes     thresh  thresh  prob
    0         3/452             0/0               0/0          20      40  1/10
    1           0/0             0/0               0/0          22      40  1/10
    2           0/0             0/0               0/0          24      40  1/10
    3           0/0             0/0               0/0          26      40  1/10
    4           0/0             0/0               0/0          28      40  1/10
    5           0/0             0/0               0/0          30      40  1/10
    6           0/0             0/0               0/0          32      40  1/10
    7           0/0             0/0               0/0          34      40  1/10
 rsvp           0/0             0/0               0/0          36      40  1/10
```

QoS Models

QoS Service Models

- These are global, high level framework describing how QoS can be applied in a network.

- Three services models:

  Best Effort

  Integrated Services

  Differentiated Services

QoS Model #1: Best Effort

- First come, first served basis

- Network's behavior:

  Treats all traffic the same and on a first come, first served basis.

- Drawbacks

  Delivers data if it can, with no assurances of reliability, delay bounds, or throughput. So basically no QoS ;)

QoS Model #2: Integrated Services

- Dynamic allocation of resources
- Network's behavior:

  Applications requests a specific level of service before starting to send data.
- Drawbacks

  Requires explicit signaling through protocol (RSVP)

  Overhead in network services, scalability issues.

QoS Model #3: Differentiated Services

- Flows are aggregated in the network
- Network's behavior:

  Small numbers of aggregated flow follow the behavior implemented on each hop ('Per Hop Behavior').
- Drawbacks

  Needs standardized policies to ensure end-to-end services

DiffServ Architecture

- Network Boundaries: Traffic Conditioner Block

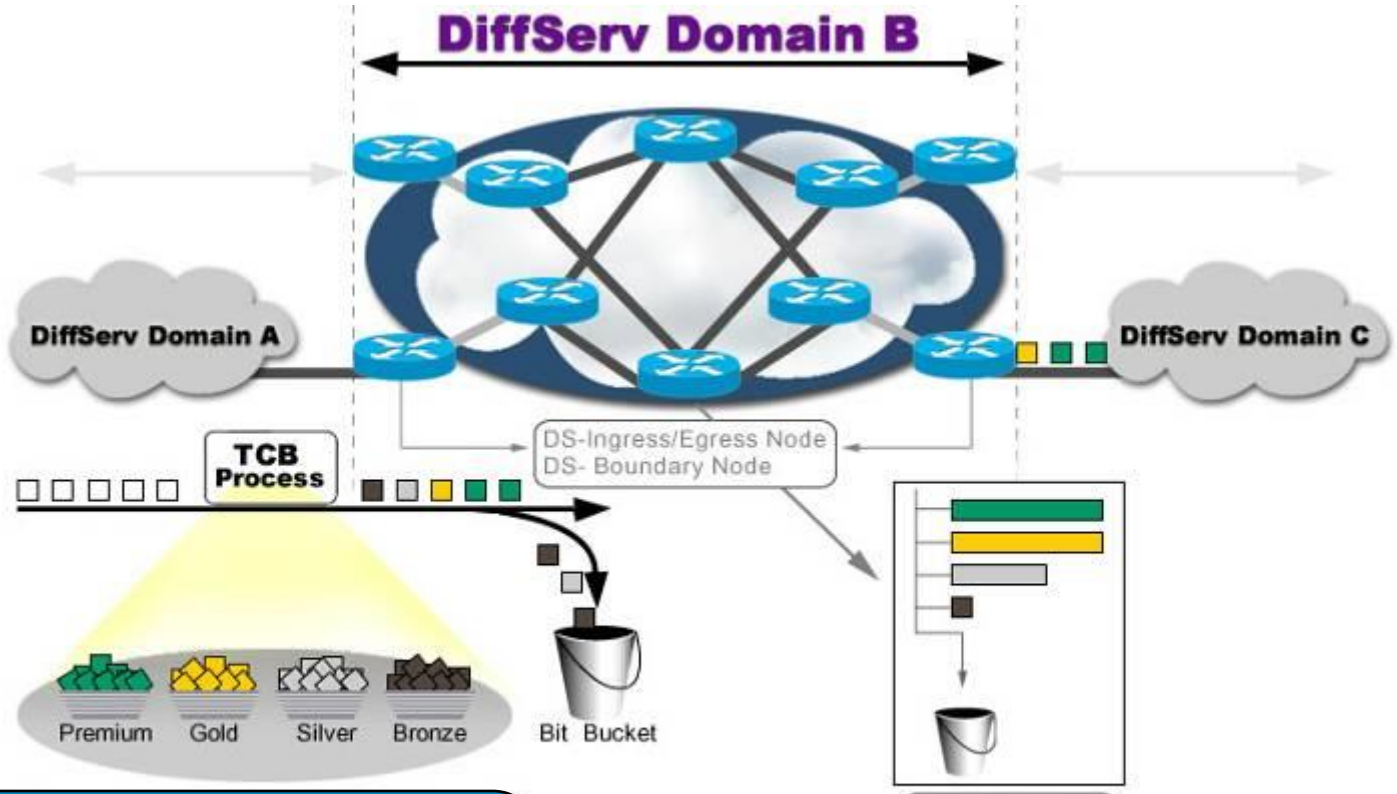
  Incoming traffic is classified and can be conditioned (metered, delayed, dropped)

  Is assigned to an aggregate flow matching a behavior. This is done by marking it with a DiffServ Code Point (DSCP).

- Network Core: Per Hop Behavior

  Traffic is forwarded/dropped according to the Per Hop Behavior corresponding to its DiffServ Code Point.

DiffServ Architecture



TCB in Cisco IOS

Class-Based Policing and Shaping

MQC-Based Classification and Marking

Packet "color" in **DSCP**

PHB in Cisco IOS

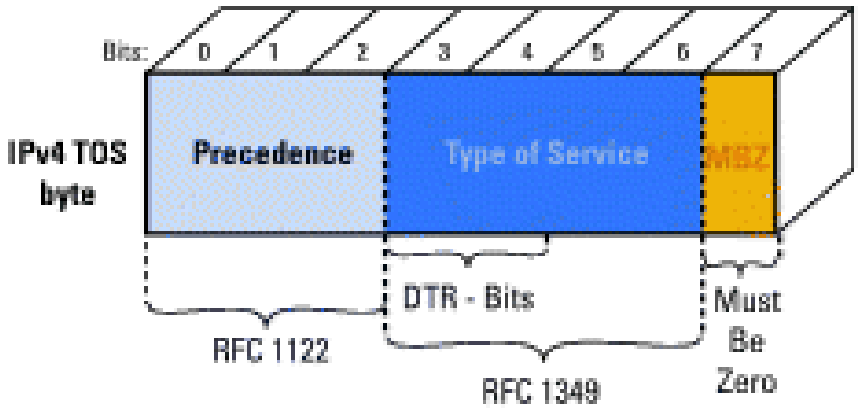
LLQ & WRED

Per Hop Behavior

- Defines the “Externally observable forwarding behavior” of a DiffServ node.
- The DiffServ model associates the standard behavior of a participating node to the DSCP of the packets.
- Some convention are used to ensure consistent usage of DSCP values across networks.

IP Precedence Field

- “An independent measure of the importance of this datagram.”
- 8 values, meaning is now merely historical.
- Defined in RFC 791 (Internet Protocol)

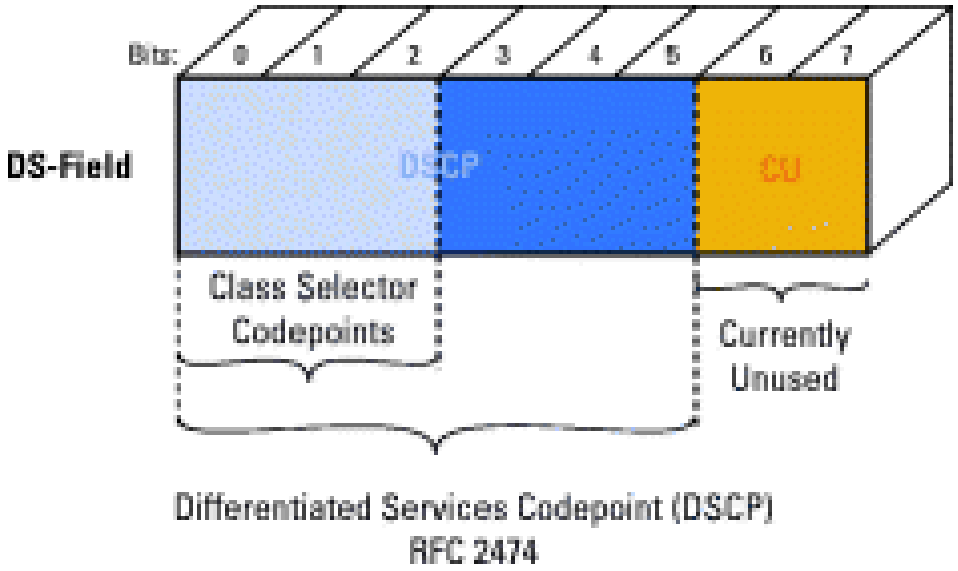


Bits (0-2): IP-Precedence Defined
111 - Network Control
110 - Internetwork Control
101 - CRITIC/ECP
100 - Flash Override
011 - Flash
101 - Immediate
001 - Priority
000 - Routine

Bits (3-6): The Type of Service Defined
0000 (all normal)
1000 (minimize delay)
0100 (maximize throughput)
0010 (maximize reliability)
0001 (minimize monetary cost)

DiffServ Code Points (DSCP) values

- RFC 2474 redefines the TOS byte into DSCP values
- 6 bits = 64 values



- DS-compliant nodes map a codepoint to a particular forwarding treatment or PHB.

Per Hop Behavior

- Can be split in 4 types:

1. Default PHB: 0
2. Class Selector PHB: IP Precedence
3. Assured Forwarding PHB: AF
4. Expedite Forwarding PHB: EF

Per Hop Behavior

1. Default PHB:

   DSCP Value 000000, maps to IP Precedence 0

2. Class Selector:

   DSCP Value xxx000, maps to IP Precedence dec(xxx)

   Values of 110000 and 111000 should always have preferential treatment to preserve common values of routing traffic (precedence 6 and 7)

Per Hop Behavior

- Assured Forwarding PHB:

  Guaranteed Bandwidth + Extra if available

  4 classes (af1, af2, af3, af4)

  3 drop probability values per class

  DSCP Value aaadd0, maps to 4 classes (aaa) and 3 drop probability (dd)

- Expedite Forwarding PHB:

  Minimum departure rate (minimum delay)

  Guaranteed Bandwidth + Drop if excess (Policed)

  DSCP Value 101110

Review Questions

Review Questions

- Name the three QoS Models.
- Which one of them involves the use of RSVP ? Why ?
- What does DSCP stand for ?
- To which type of PHB do those belong:

  101000

  101110

  010110
- Which PHB class would you use for voice traffic ?

References

- Cisco IOS – The Differentiated Services Model

  http://www.cisco.com/warp/public/cc/pd/iosw/prodlit/difsf_ds.htm

- DiffServ - The Scalable End-to-End QoS Model

  http://www.cisco.com/en/US/partner/tech/tk543/tk766/technologies_white_paper09186a00800a3e2f.shtml

- RFC 791 – Internet Protocol
- RFC 2474 - Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers
- RFC 2475 - An Architecture for Differentiated Services
- RFC 2597 - Assured Forwarding PHB Group
- RFC 2346 - An Expedited Forwarding PHB