Enhancing Network Operations With QoS

ine.com

Keith Bogart

CCIE #4923

kbogart@ine.com
@keithbogart1
linkedin.com/in/keith-bogart-2a75042

CCIE Routing & Switching

Course Objectives

- To explain why we need QoS in today’s networks
- To help you become familiar with common QoS terminology
- To give you exposure to some common QoS techniques

Course Prerequisites

- A basic understanding of how routers and switches forward IP packets
- Familiarity with the IPv4 packet header

INE

Introduction To QoS

ine.com

Topic Overview

+ What Problems Are Solved By QoS?
+ How Does QoS Control Traffic?
+ Day In The Life Of A Packet
+ The Differences Between Buffers & Queues

QoS Overview

- QoS = Quality of Service
- What problem does it solve?
  - Provides predictable management of network resources during times of congestion
  - Assists in maximizing the end-user experience of critical sessions
  - Provides differentiated services to packets based upon pre-defined user criteria

How Does QoS Control Network Traffic?

- Many different QoS features
- Some designed to accomplish only a single task
- Others designed to accomplish multiple tasks
- In general, the tasks that a QoS feature can accomplish can be categorized as:
  - Classification of data
  - Queue Management
    - Size, Placement of packets, Scheduling Order, Transmission Rate
  - Pre-Emptive Drops
  - Marking of packets

Memory Buffers

- Buffers
  - Physical memory used to store packets before and after a forwarding decision is made
  - On routers, this same memory can be allocated to interfaces as ingress/egress queues
  - Shared memory (of which, part is allocated as buffers) is also used by lots of other CPU processes

Memory Queues

- Queues
  - On routers, a queue is a logical part of the shared memory buffers
  - On switches, individual interfaces (or linecards) have their own memory which is used as interface queues

Routers: Day In The Life Of A Packet

1. Packet arrives on ingress interface (Rx-Ring)
2. Packet queued in memory buffer
3. Forwarding decision is made
4. Packet placed on hardware transmit ring
5. Packet transmitted onto egress media

Switches (Shared Memory): Packet Forwarding

1. Packet arrives on ingress interface
2. Interface/Module ASIC immediately forwards packet into a common, shared memory pool
3. Forwarding decision is made by forwarding ASIC(s)
4. Memory ownership of packet buffer transferred to egress interface
5. Packet transmitted onto egress media

Switches (Distributed Memory): Packet Forwarding

1. Packet arrives on ingress interface
2. Interface/Module ASIC places packet into a queue (buffer)
3. Forwarding decision is made by forwarding ASIC(s)
4. Packet transmitted (along with lookup result) onto shared ring/bus to all egress interfaces
5. Appropriate egress interface queues and then schedules the packet

Queuing & Congestion

- DMA = Direct Memory Access
- When egress traffic cannot immediately be transmitted, it is placed in an egress queue
- Without QoS = The queue is one large piece of memory
  - FIFO
- With QoS = One can control characteristics of the queue

What Is Affected By QoS?

- Bandwidth
- Delay
- Jitter
- Loss (or Drops)

Thanks for Watching!

Classification & Marking

ine.com

Topic Overview

+ Definition Of Classification
+ Layer-2 Classification
+ Layer-3 Classification
+ Overview Of NBAR
+ Trust Boundaries

Classification Defined

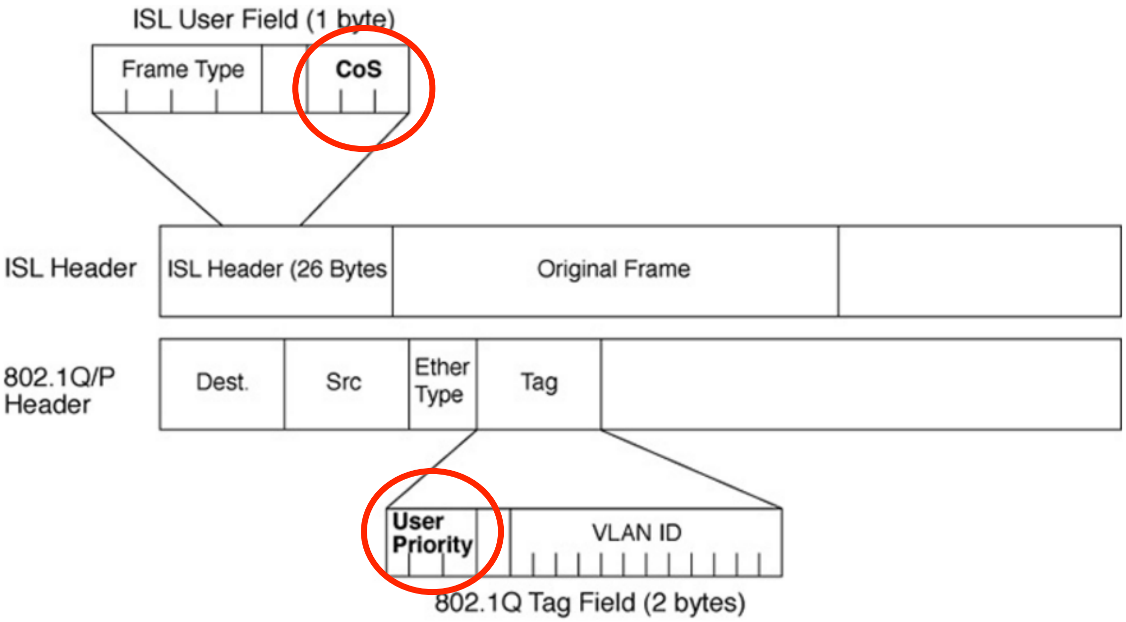
- You know what traffic is important to you…but how does the ROUTER know that?
- Traffic must first be divided into "classes"
  - A "class" of traffic will receive the same type of QoS treatment
  - Analyze packets to differentiate flows
- Classification = features that identify traffic based distinctive differences

Classification Defined

- Packets belonging to same class typically marked on ingress to allow for easier classification by upstream devices
- Most common ways of classifying traffic:
  - Markings
  - Addressing
  - Application Signatures

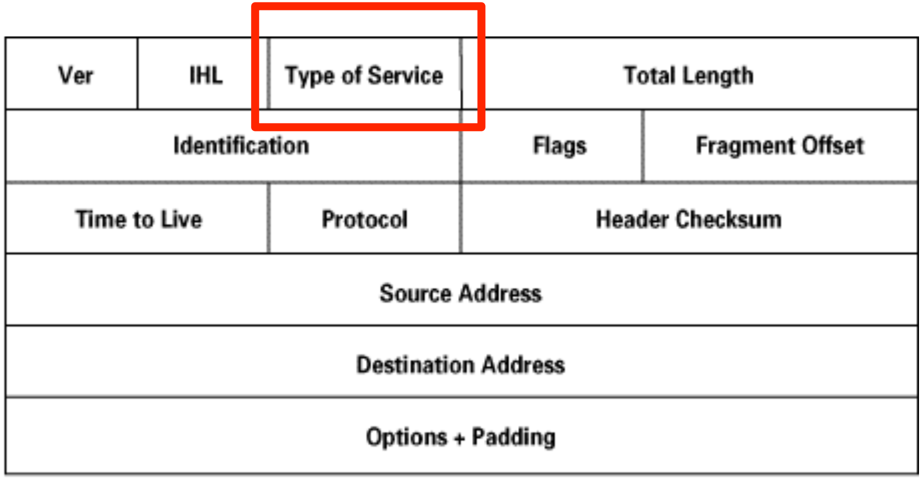
Layer-2 Classification

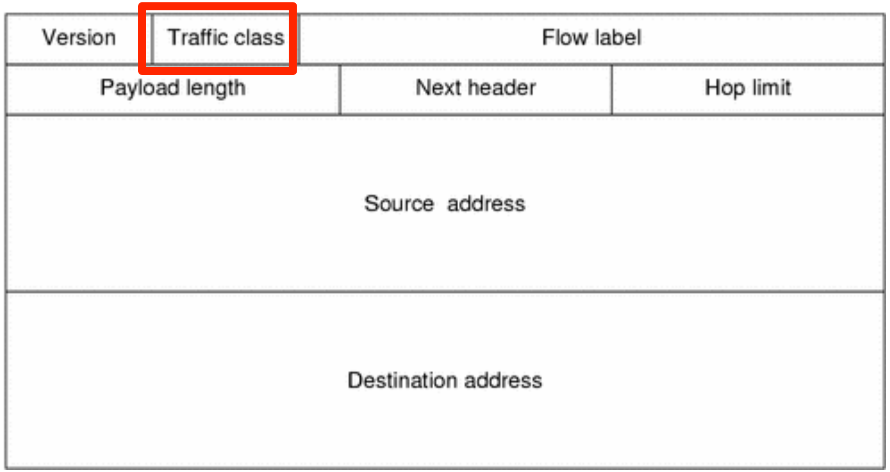
- Ethernet frames contain no distinctive "priority" field UNLESS carried by 802.1q or ISL trunks



Layer-3 Classification (ToS byte)

+ Both IPv4 and IPv6 contain a byte used for indicating relative priority of a packet





NBAR

- Network-Based Application Recognition
- Most protocols can be identified by matching on their well-known L3 or L4 numbers
  - Some protocols negotiate dynamic numbers and can't be matched this way
  - NBAR examines the data payload
  - More CPU intensive than other classification features

NBAR

- NBAR supports recognition of a large quantity of protocols
- Example: NBAR can be used to match on a full URL name, or a word or phrase within a URL
- Implemented by CPU of the device…so most (not all) Cisco switches don't support NBAR because their CPUs never see the traffic

Trust Boundaries

- Some host devices may mark traffic upon creation
- Do you “trust” these devices?
- QoS Trust Boundaries
  - Logical point in network beyond which, received QoS markings are not trusted
  - Typically access-layer ports
  - Default when QoS enabled = Untrusted

Thanks for Watching!

IP Precedence & DSCP

ine.com

Topic Overview

- Explaining The IPv4 ToS Byte
- Understanding DSCP
- Common DSCP Per-Hop Behavior Classes

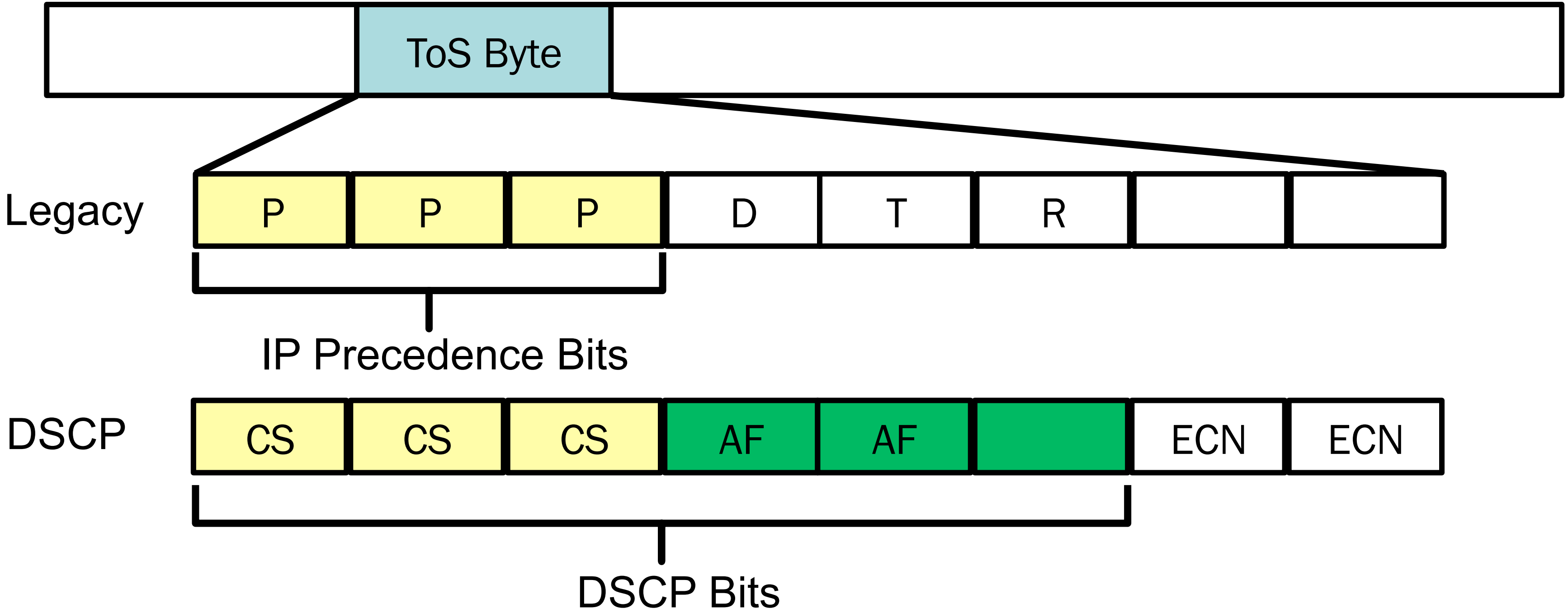
IPv4 & ToS

- Original meanings of ToS byte per RFC 791

```
Bits 0-2:  Precedence.
Bit    3:  0 = Normal Delay,       1 = Low Delay.
Bits   4:  0 = Normal Throughput, 1 = High Throughput.
Bits   5:  0 = Normal Relibility, 1 = High Relibility.
Bit  6-7:  Reserved for Future Use.

   0     1     2     3     4     5     6     7
+-----+-----+-----+-----+-----+-----+-----+-----+
|                 |     |     |     |     |     |
|   PRECEDENCE    |  D  |  T  |  R  |  0  |  0  |
|                 |     |     |     |     |     |
+-----+-----+-----+-----+-----+-----+-----+-----+

  Precedence

    111 - Network Control
    110 - Internetwork Control
    101 - CRITIC/ECP
    100 - Flash Override
    011 - Flash
    010 - Immediate
    001 - Priority
    000 - Routine
```

DSCP

- DSCP = Differentiated Services Code Point
- Utilizes six bits within ToS byte for QoS Prioritization



Defined DSCP Per-Hop Behavior Classes

- There are four high-level traffic classes defined for DSCP per-hop behaviors
  - Default Forwarding (DH) PHB
    000000xx (DCSP value 0)
  - Expedited Forwarding (EF) PHB
    101110xx (DSCP value 46)
  - Assured Forwarding (AF) PHB
    - AF11 through AF43
  - Class Selector (CS) PHB
    - Backwards compatible with older systems implementing IP Precedence

Thanks for Watching!

Congestion Avoidance With Policing & Shaping

Topic Overview

- Overview Of Congestion Avoidance
- Defining Policing, Shaping & Markdown
- Policing & Shaping Compared

Congestion Avoidance - Overview

- Term used to define a set of features that attempt to prevent queues from becoming congested
- Can be done in three places (depending on hardware platform)
  - Ingress interface queue, prior to lookup by forwarding engine
  - At the forwarding engine (policing)
  - Within the egress queue (queuing and shaping)

Policing, Shaping & Markdown

- Between ISP and Customer there is a pre-defined, contracted rate (called CIR)
- ISP will police ingress traffic
  - Traffic that is non-conforming is caught by policer and:
    - Dropped
    - Marked-down
- Customer typically doesn't want any traffic dropped (delay is better than drops)
  - Shaping done on egress interface leading to ISP

Sample Policing Configuration

```
policy-map INE
  class Prec3
    police cir <rate in bps> pir <rate in bps> conform-action transmit
exceed-action set-prec-transmit 0 violate-action drop
```

Policing & Shaping Compared

- On routers:
  - Policers can be applied on ingress or egress interfaces...but usually done on ingress
  - Typically ISPs will enforce contracts with Policers
  - Shapers usually done on egress connection TO the ISP
- Most Cisco Switches do not support Traffic Shaping

Sample Traffic Shaping Configuration

```
R2(config)#policy-map INE
R2(config-pmap)# class Prec0
R2(config-pmap-c)#shape ?
  adaptive         Enable Traffic Shaping adaptation to BECN
  average          configure token bucket: CIR (bps) [Bc (bits) [Be (bits)]],
                   send out Bc only per interval
  fecn-adapt       Enable Traffic Shaping reflection of FECN as BECN
  fr-voice-adapt   Enable rate adjustment depending on voice presence
  peak             configure token bucket: CIR (bps) [Bc (bits) [Be (bits)]],
                   send out Bc+Be per interval

R2(config-pmap-c)#shape av
R2(config-pmap-c)#shape average ?
  <1000-154400000>  Target Bit Rate (bits/sec). (postfix k, m, g optional;
                    decimal point allowed)
  percent           % of interface bandwidth for Committed information rate
```

Thanks for Watching!

Congestion Avoidance With Pre-Emptive Queue Drops

ine.com

Topic Overview

- What Is Queuing-Based Congestion Avoidance?
- Types Of Congestion Avoidance
- Overview Of WRED & WTD

Queuing-Based Congestion Avoidance

- Set of features to pre-emptively drop traffic within queues
- The goal: Prevent queues from becoming saturated with low-priority traffic by randomly dropping that traffic…thus leaving room in the queue for future, high-priority traffic

Congestion Avoidance At The Queue

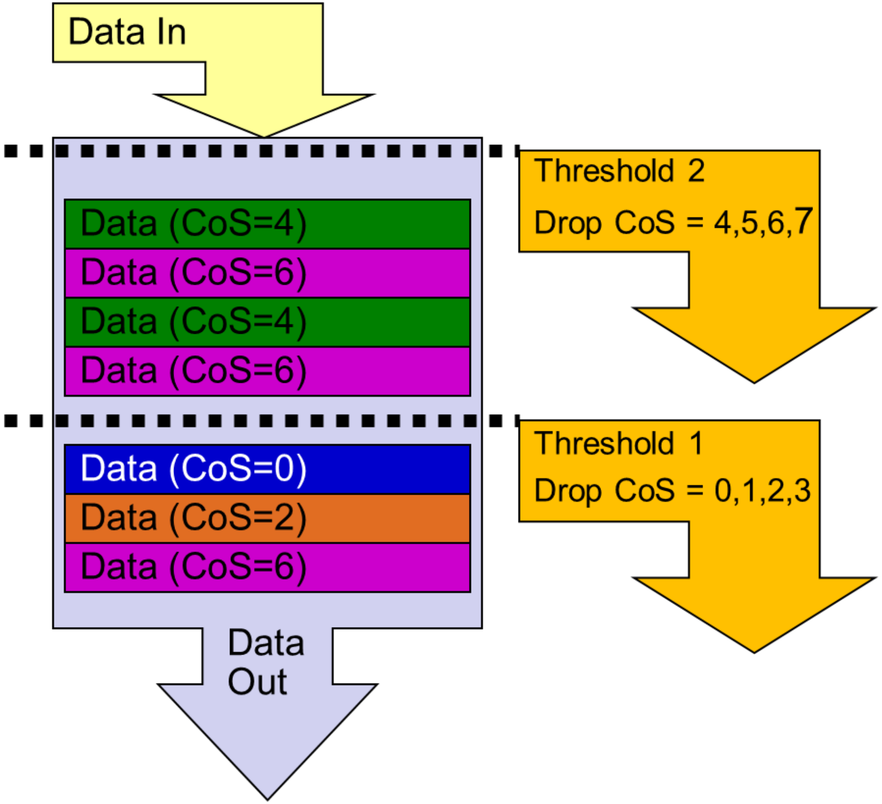
- Congestion Avoidance within queues on Switches
  - WTD (Weighted Tail Drop)
  - WRED (Weighted Random Early Discard)
  - DBL (Dynamic Buffer Limiting)
- Congestion Avoidance within queues on Routers
  - WRED (Weighted Random Early Discard)

WRED & WTD Terminology

- Drop Thresholds
  - Minimums
  - Maximums
- Minimum Threshold = When drops begin
- Maximum Threshold = Point at which 100% of matched traffic is dropped

Weighted Tail Drop

- Mechanism used on most switching hardware
- Configurable thresholds and DSCP-to-Threshold Mappings



WRED Operation

- Random packet drops start at the min-threshold
- Increase in a linear format until max-threshold is reached
- After max-threshold is reached, WRED drops 100% of all subsequent packets received

Thanks for Watching!

Congestion Management Control With Queuing & Scheduling

ine.com

Topic Overview

+ Queuing Defined
+ Scheduling Defined
+ What Is Congestion Management & Why Do We Need It?
+ Queuing & Scheduling Features

Queuing

- A single egress interface may have multiple, associated egress queues differentiated by priority
- QoS features designed for queuing provide control over which classified traffic is placed into each of these queues
- Can also pre-emptively drop traffic from within queues to make room for higher-priority traffic

Scheduling

- What is the “Scheduler”?
- On routers, QoS queuing features (such as WFQ) typically affect queuing and scheduling behaviors
- On switches, queuing and scheduling can be separate functions/features
- Traffic Shaping is a function of Scheduling

What Is Congestion Management?

- Congestion management features allow you to control congestion by determining the order in which packets are sent out an interface based on priorities assigned to those packets
  - Creation of queues
  - Assignment of packets to those queues based on the classification of the packet
  - Selectively dropping packets from within queues when those queues reach pre-defined thresholds
  - Scheduling of the packets in a queue for transmission

Why Do We Need Congestion Management?

- The fundamental reason we need Congestion Management is because:
  - By default, queues are configured for FIFO (First-In First-Out)
  - FIFO provides no control over the order of packet transmission
  - Incoming bursts can cause congestion of queues
  - Congestion management techniques provide some control of the order-of-transmission

Queuing & Scheduling Features

- Queuing features
  - FIFO (no congestion management)
  - Weighted Fair Queuing
  - LLQ
  - CBWFQ
- Scheduling Features
  - Round Robin
  - Weighted Round Robin
  - LLQ
  - CBWFQ

Thanks for Watching!