KHOA CÔNG NGHỆ THÔNG TIN-ĐHKHTN
CSC11004 - MẠNG MÁY TÍNH NÂNG CAO

QUEUE MANAGEMENT & QUALITY OF SERVICES

Lê Ngọc Sơn

TPHCM, 9-2021

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

fit@hcmus

Agenda

- Queue Management
- Drop Policy
- Scheduling Discipline
- Quality of Service
- IntServ and DiffServ

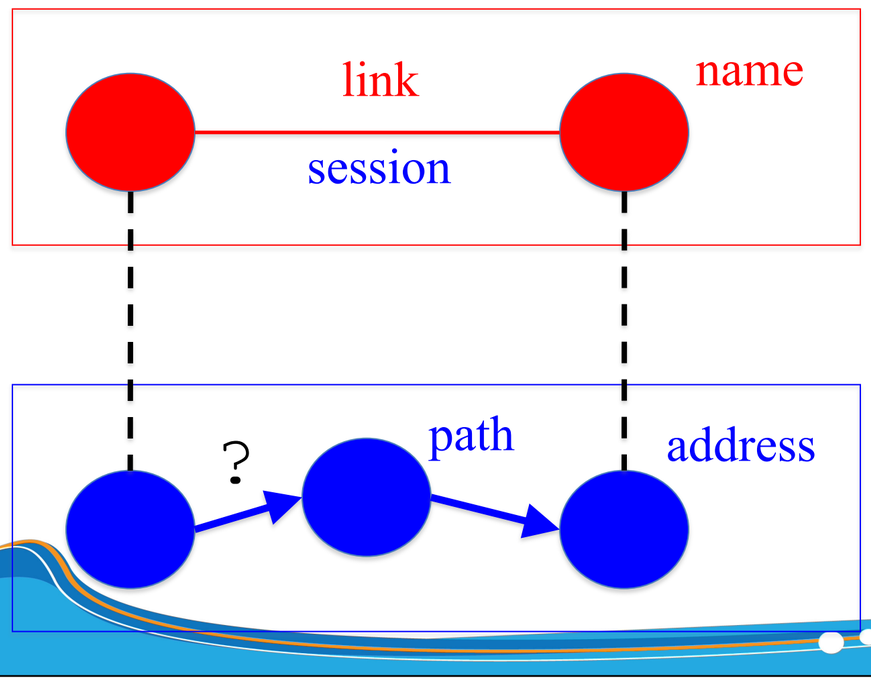
Queue Management

Congestion Control

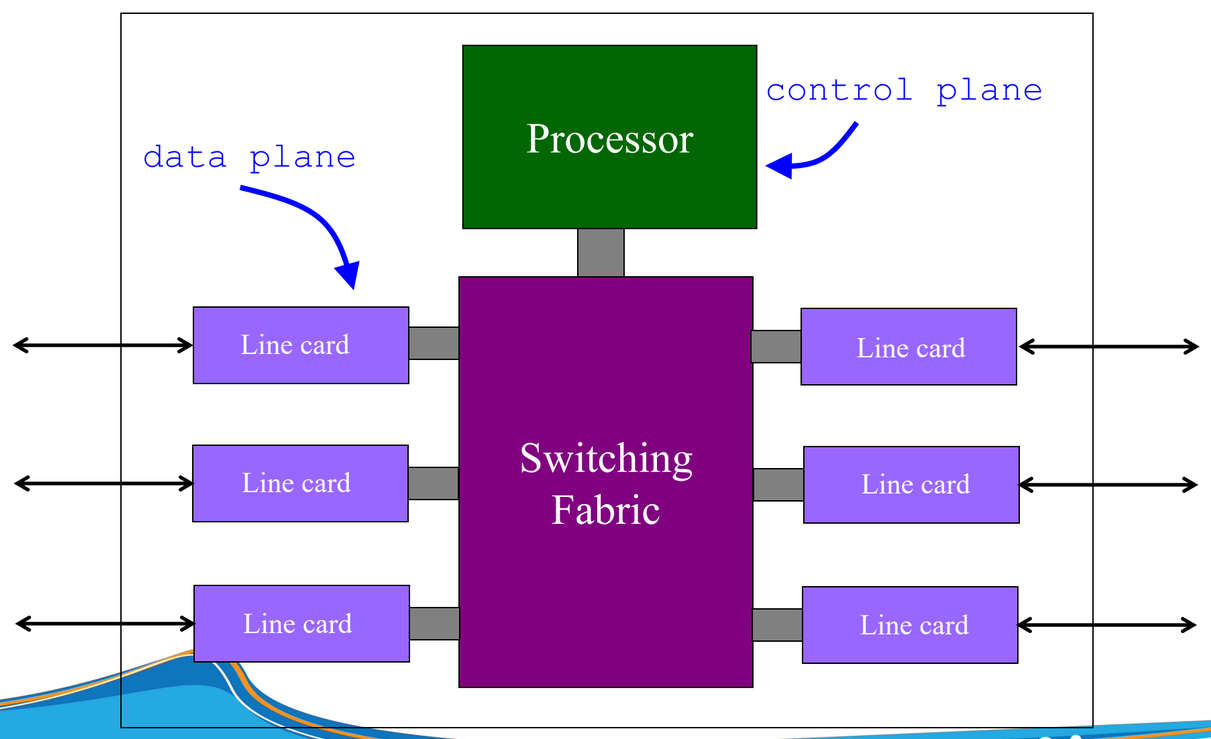
What can the *end-points* do to collectively to make good use of shared underlying resources?

Queue Management

What can the individual *links* do to make good use of shared underlying resources?

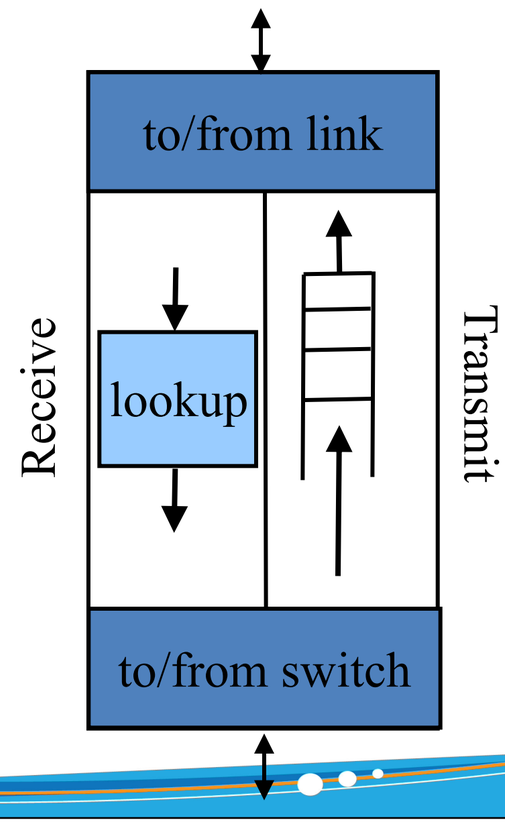


Router

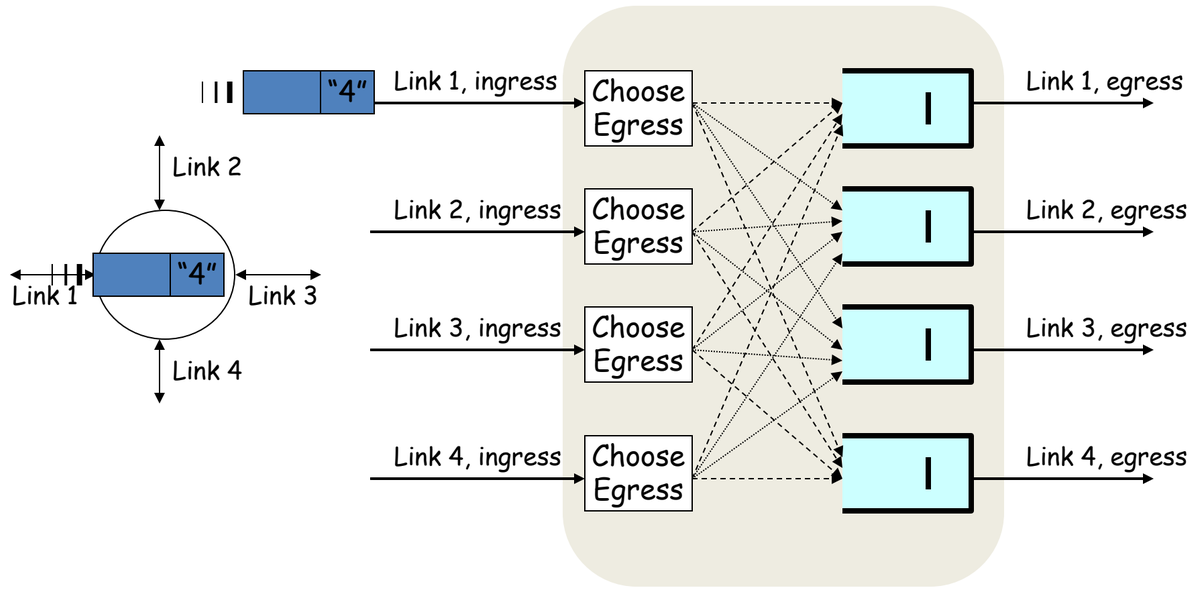


Line Cards (Interface Cards, Adaptors)

- Packet handling
  - Packet forwarding
  - Buffer management
  - Link scheduling
  - Packet filtering
  - Rate limiting
  - Packet marking
  - Measurement



Packet Switching and Forwarding



Buffer Size

- Why not use infinite buffers?
  - no packet drops!
- Small buffers:
  - often drop packets due to bursts
  - but have small delays
- Large buffers:
  - reduce number of packet drops (due to bursts)
  - but increase delays
- Can we have the best of both worlds?

Queue Management Issues

- Drop policy
  - When should you discard a packet?
  - Which packet to discard?
- Scheduling discipline
  - Which packet to send?
  - Some notion of fairness? Priority?
- Goal: balance throughput and delay
  - Huge buffers minimize drops, but add to queuing delay (thus higher RTT, longer slow start, …)

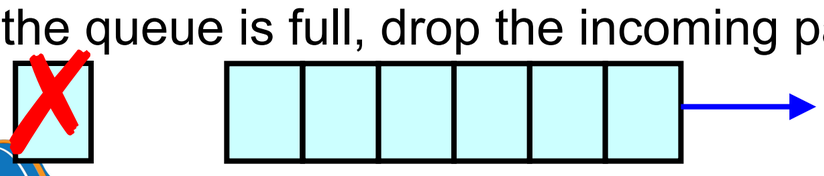
Drop Policies

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

FIFO Scheduling and Drop-Tail

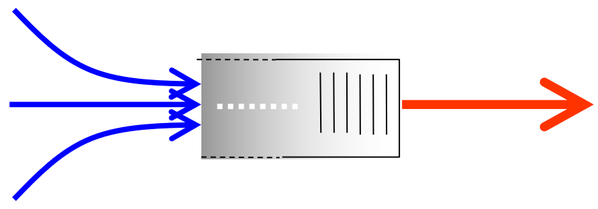
- Access to the bandwidth: first-in first-out queue
  - Packets only differentiated when they arrive

- Access to the buffer space: drop-tail queuing
  - If the queue is full, drop the incoming packet



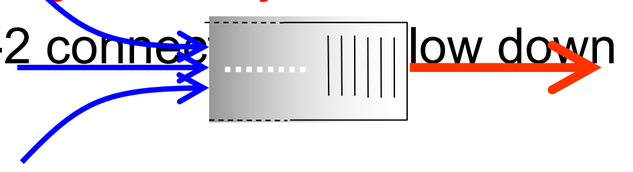
Bursty Loss From Drop-Tail Queuing

- TCP depends on packet loss
  - Packet loss is indication of congestion
  - TCP additive increase drives network into loss
- Drop-tail leads to *bursty* loss
  - Congested link: many packets encounter full queue
  - Synchronization: many connections lose packets at once



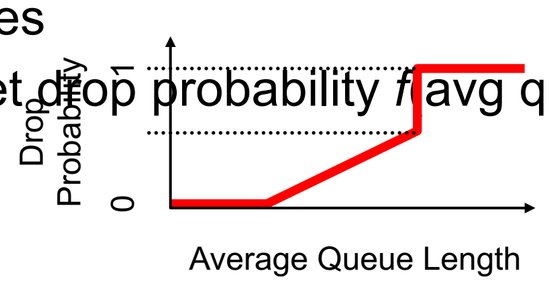
Slow Feedback from Drop Tail

- Feedback comes when buffer is completely full
  - … even though the buffer has been filling for a while
- Plus, the filling buffer is increasing RTT
  - … making detection even slower
- Better to give early feedback
  - Get 1-2 connections to slow down before it's too late!



Random Early Detection (RED)

- Router notices that queue is getting full
  - … and randomly drops packets to signal congestion
- Packet drop probability
  - Drop probability increases as queue length increases
  - Else, set drop probability *f*(avg queue length)



Properties of RED

- Drops packets before queue is full
  - In the hope of reducing the rates of some flows
- Drops packet in proportion to each flow's rate
  - High-rate flows selected more often
- Drops are spaced out in time
  - Helps desynchronize the TCP senders
- Tolerant of burstiness in the traffic
  - By basing the decisions on average queue length

Problems With RED

- Hard to get tunable parameters just right
  - How early to start dropping packets?
  - What slope for increase in drop probability?
  - What time scale for averaging queue length?
- RED has mixed adoption in practice
  - If parameters aren't set right, RED doesn't help
- Many other variations in research community
  - Names like "Blue" (self-tuning), "FRED"…

Feedback: From loss to notification

- Early dropping of packets
  - Good: gives early feedback
  - Bad: has to drop the packet to give the feedback
- Explicit Congestion Notification
  - Router marks the packet with an ECN bit
  - Sending host interprets as a sign of congestion

Explicit Congestion Notification

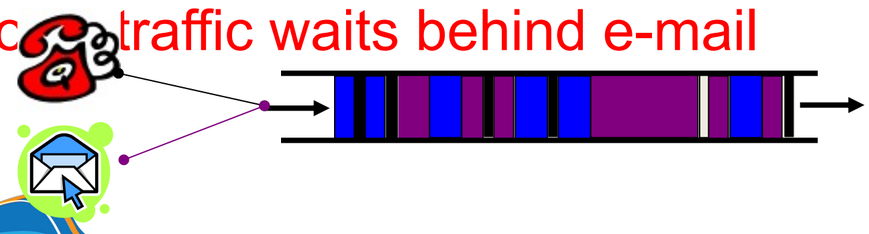
- Must be supported by router, sender, AND receiver
  - End-hosts determine if ECN-capable during TCP handshake
- ECN involves all three parties (and 4 header bits)
  1. Sender marks "ECN-capable" when sending
  2. If router sees "ECN-capable" and experiencing congestion, router marks packet as "ECN congestion experienced"
  3. If receiver sees "congestion experienced", marks "ECN echo" flag in responses until congestion ACK'd
  4. If sender sees "ECN echo", reduces cwnd and marks "congestion window reduced" flag in next TCP packet

Scheduling Discipline

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

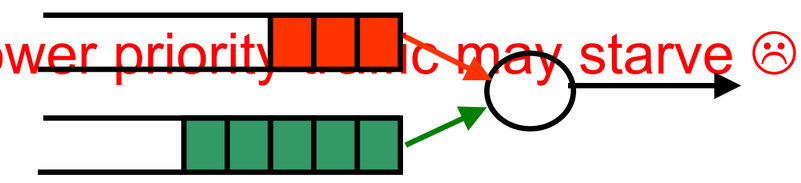
First-In First-Out Scheduling

- First-in first-out scheduling
  - Simple, but restrictive
- Example: two kinds of traffic
  - Voice over IP needs low delay
  - E-mail is not that sensitive about delay
- Voice traffic waits behind e-mail



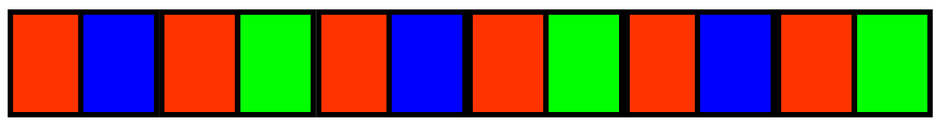
Strict Priority

- Multiple levels of priority
  - Always transmit high-priority traffic, when present
- Isolation for the high-priority traffic
  - Almost like it has a dedicated link
  - Except for (small) delay for packet transmission
- But, lower priority traffic may starve ☹



Weighted Fair Scheduling

- Weighted fair scheduling
  - Assign each queue a fraction of the link bandwidth
  - Rotate across queues on a small time scale



50% red, 25% blue, 25% green

- Work-conserving
  - Send extra traffic from one queue if others are idle

Implementation Trade-Offs

- FIFO
  - One queue, trivial scheduler
- Strict priority
  - One queue per priority level, simple scheduler
- Weighted fair scheduling
  - One queue per class, and more complex scheduler

Quality of Service

KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

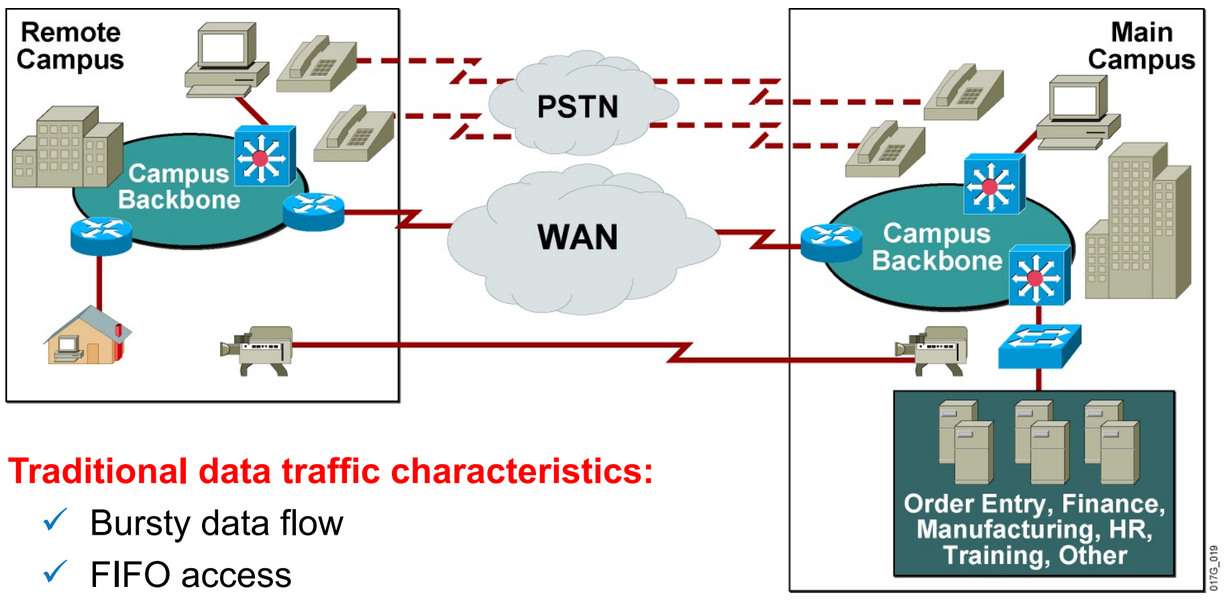
fit@hcmus

What is Quality of Service (Qos)?

"The measurable end-to-end performance properties of a network service, which can be guaranteed in advance by a Service Level Agreement between a user and a service provider, so as to satisfy specific customer application requirements. Note: These properties may include throughput (bandwidth), transit delay (latency), error rates, priority, security, packet loss, packet jitter, etc"

----------NIST ---

Traditional Nonconverged Network



Traditional data traffic characteristics:

- ✓ Bursty data flow
- ✓ FIFO access
- ✓ Not overly time-sensitive; delays OK
- ✓ Brief outages are survivable

What Is Quality of Service?

- The user perspective

  Users perceive that their applications are performing properly

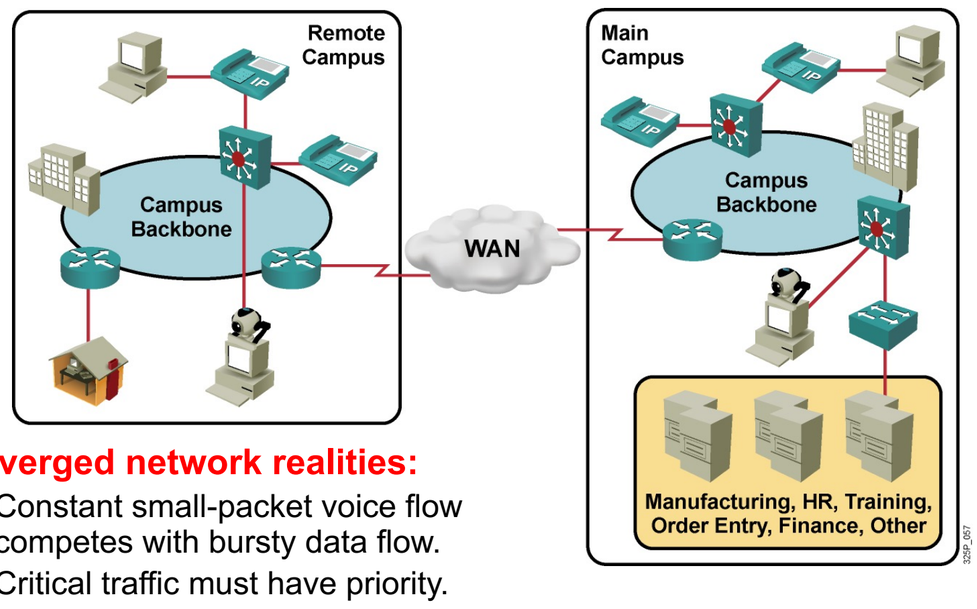
  Voice, video, and data

- The network manager perspective

  Need to manage bandwidth allocations to deliver the desired application performance

  Control delay, jitter, and packet loss

Converged Network Realities



Converged network realities:

- ✓ Constant small-packet voice flow competes with bursty data flow.
- ✓ Critical traffic must have priority.
- ✓ Voice and video are time-sensitive.
- ✓ Brief outages are not acceptable.

Different Types of Traffic Have Different Needs

Real-time applications especially sensitive to QoS

- ✓ Interactive voice
- ✓ Videoconferencing

Causes of degraded performance

- ✓ Congestion losses
- ✓ Variable queuing delays

The QoS challenge

- ✓ Manage bandwidth allocations to deliver the desired application performance
- ✓ Control delay, jitter, and packet loss

Application Examples	Sensitivity to QoS Metrics		
	Delay	Jitter	Packet Loss
Interactive Voice and Video	Y	Y	Y
Streaming Video	N	Y	Y
Transactional/ Interactive	Y	N	N
Bulk Data Email File Transfer	N	N	N

Need to manage bandwidth allocations

Converged Network Quality Issues

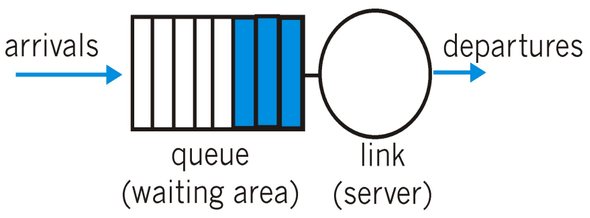
- Lack of bandwidth: Multiple flows compete for a limited amount of bandwidth.
- End-to-end delay (fixed and variable): Packets have to traverse many network devices and links; this travel adds up to the overall delay.
- Variation of delay (jitter): Sometimes there is a lot of other traffic, which results in varied and increased delay.
- Packet loss: Packets may have to be dropped when a link is congested.

Building blocks

- Scheduling
  - Active Buffer Management
- Traffic Shaping
  - Leaky Bucket
  - Token Bucket

Scheduling: How Can Routers Help

- Scheduling: choosing the next packet for transmission
  - FIFO/Priority Queue
  - Round Robin/ DRR
  - Weighted Fair Queuing
- Packet dropping:
  - not drop-tail
  - not only when buffer is full
    - Active Queue Management
- Congestion signaling
  - Explicit Congestion Notification (ECN)



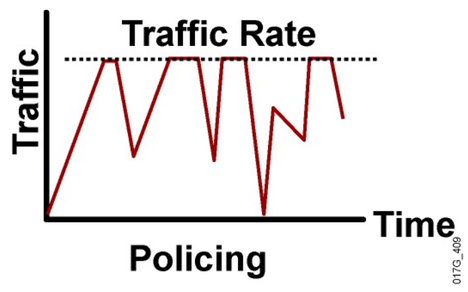
Discussion

Traffic Shaping & Traffic Policing Comparison?

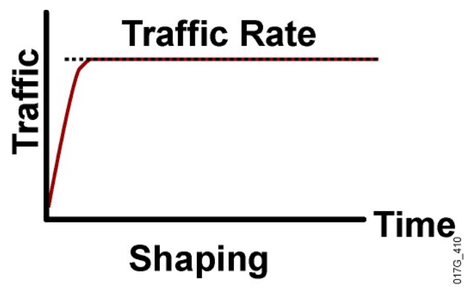
Traffic Conditioners

- **Policing**
  - ✓ Limits bandwidth by discarding traffic.
  - ✓ Can re-mark excess traffic and attempt to send.
  - ✓ Should be used on higher-speed interfaces.
  - ✓ Can be applied inbound or outbound.
- **Shaping**
  - ✓ Limits excess traffic by buffering.
  - ✓ Buffering can lead to a delay.
  - ✓ Recommended for slower-speed interfaces.
  - ✓ Cannot re-mark traffic.
  - ✓ Can only be applied in the outbound direction.

Policing Versus Shaping

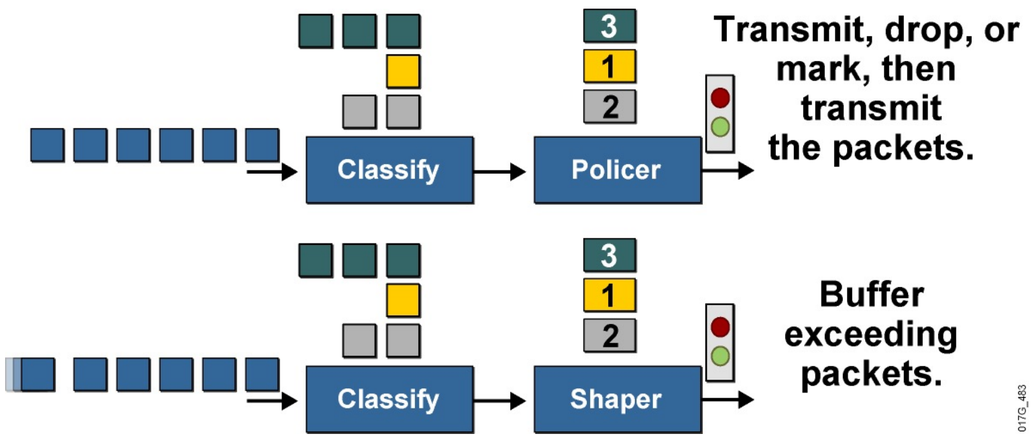


- ✓ Incoming and outgoing directions.
- ✓ Out-of-profile packets are dropped.
- ✓ Dropping causes TCP retransmits.
- ✓ Policing supports packet marking or re-marking.



- ✓ Outgoing direction only.
- ✓ Out-of-profile packets are queued until a buffer gets full.
- ✓ Buffering minimizes TCP retransmits.
- ✓ Marking or re-marking not supported.
- ✓ Shaping supports interaction with Frame Relay congestion indication.

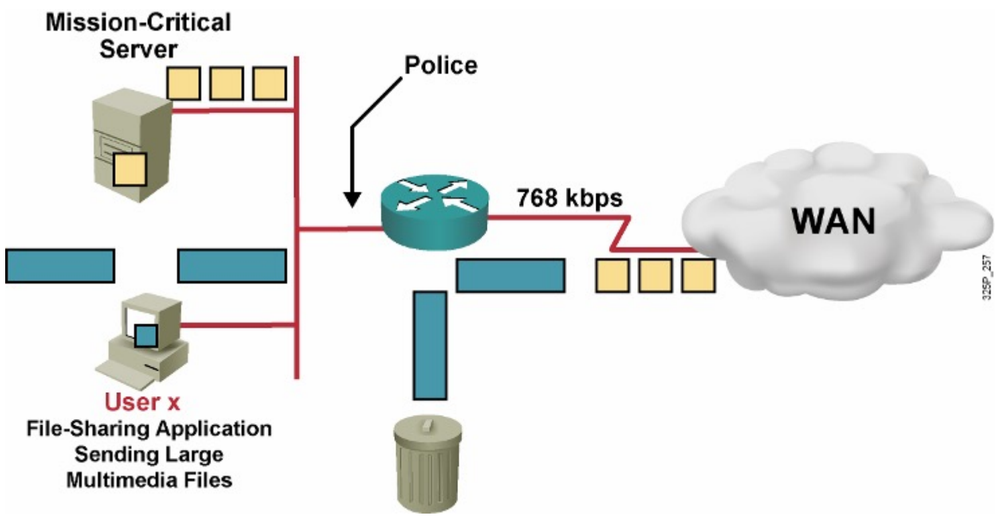
Traffic Policing and Shaping



- These mechanisms must classify packets before policing or shaping the traffic rate.
- Traffic policing typically drops or marks excess traffic to stay within a traffic rate limit.
- Traffic shaping queues excess packets to stay within the desired traffic rate.

Traffic Shaping Algorithms

Traffic Policing Example

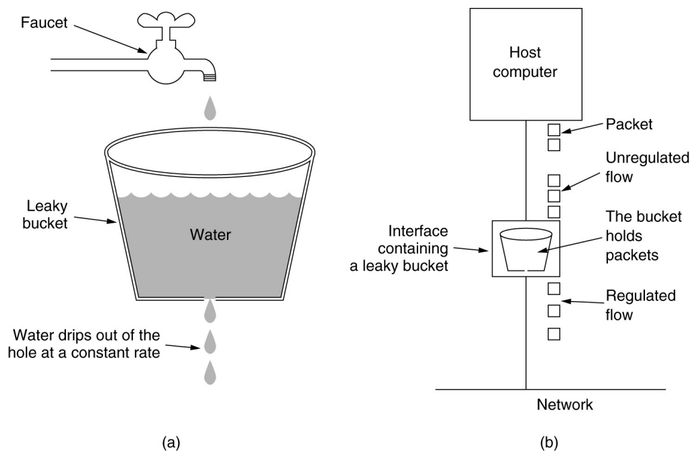


- Do not rate-limit traffic from mission-critical server.
- Rate-limit file-sharing application traffic to 56 kbps.

The Leaky Bucket

- The **Leaky Bucket Algorithm**
  - used to control rate in a network.
  - It is implemented as a single-server queue
    - with constant service time.
  - If the bucket (buffer) overflows then packets are discarded.
- Leaky Bucket (parameters r and B):
  - Every r time units: send a packet.
  - For an arriving packet
    - If queue not full then enqueue
- Note that the output is a “perfect” constant rate.

The Leaky Bucket Algorithm



(a) A leaky bucket with water. (b) a leaky bucket with packets.

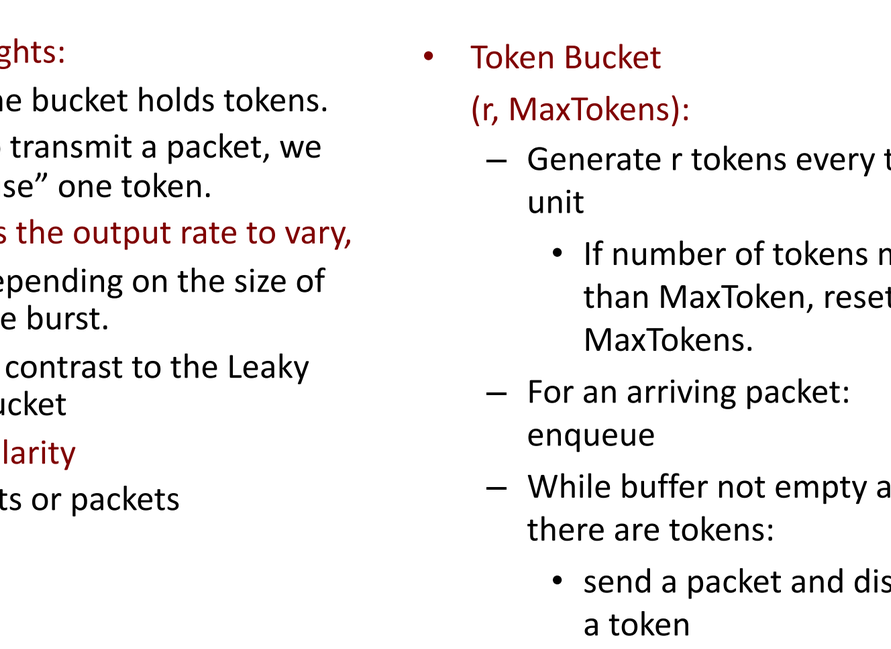
Discussion

Drawbacks of Leaky Bucket ?

Token Bucket Algorithm

- Highlights:
  - The bucket holds tokens.
  - To transmit a packet, we "use" one token.
- Allows the output rate to vary,
  - depending on the size of the burst.
  - In contrast to the Leaky Bucket
- Granularity
  - Bits or packets
- Token Bucket (r, MaxTokens):
  - Generate r tokens every time unit
    - If number of tokens more than MaxToken, reset to MaxTokens.
  - For an arriving packet: enqueue
  - While buffer not empty and there are tokens:
    - send a packet and discard a token

The Token Bucket Algorithm



(a) Before. (b) After.

Leaky Bucket vs Token Bucket

Leaky Bucket

- Discard:
  - Packets
- Rate:
  - fixed rate (perfect)
- Arriving Burst:
  - Waits in bucket

Token Bucket

- Discard:
  - Tokens
  - Packet management separate
- Rate:
  - Average rate
  - Burst allowed
- Arriving Burst:
  - Can be sent immediately

QoS Models

Three QoS Models

Model	Characteristics
Best effort	No QoS is applied to packets. If it is not important when or how packets arrive, the best-effort model is appropriate.
Integrated Services (IntServ)	Applications signal to the network that the applications require certain QoS parameters.
Differentiated Services (DiffServ)	The network recognizes classes that require QoS.

Best-Effort Model

- Internet was initially based on a best-effort packet delivery service.
- Best-effort is the default mode for all traffic.
- There is no differentiation among types of traffic.
- Best-effort model is similar to using standard mail—"The mail will arrive when the mail arrives."
- Benefits:
  - ✓ Highly scalable
  - ✓ No special mechanisms required
- Drawbacks:
  - ✓ No service guarantees
  - ✓ No service differentiation

IntServ and DiffServ

Integrated Services

- Network wide control
- Admission Control
- Absolute guarantees
- Traffic Shaping
- Reservations
  - RSVP

Differentiated Services

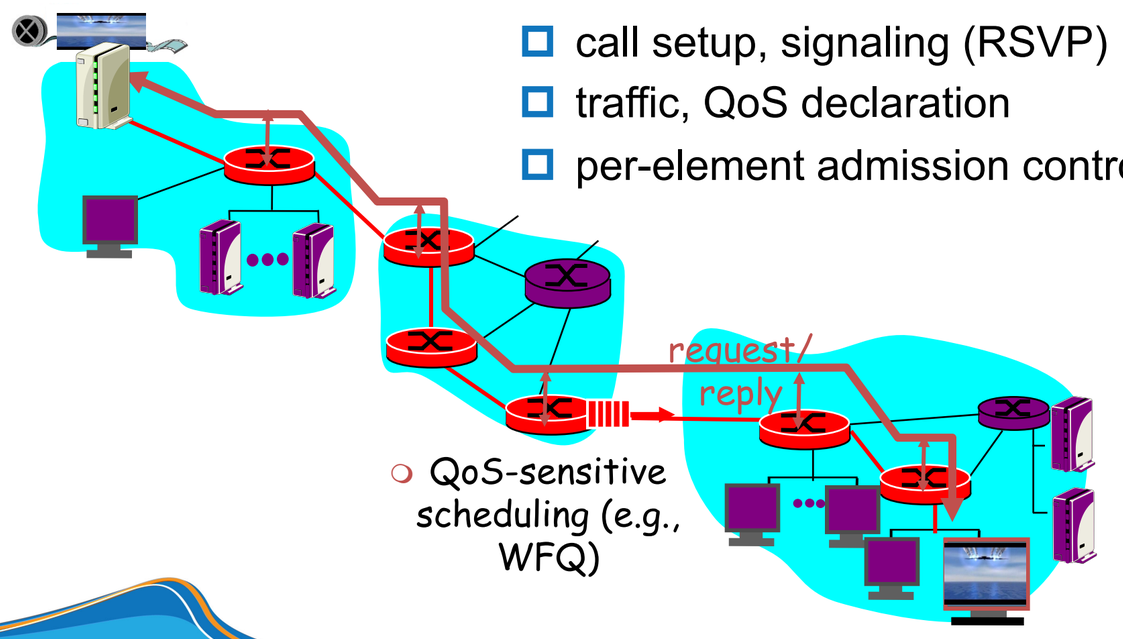
- Router based control
  - Per hop behavior
- Resolves contentions
  - Hot spots
- Relative guarantees
- Traffic policing
  - At entry to network

IETF Integrated Services

- architecture for providing QOS guarantees in IP networks for individual application sessions
- resource reservation: routers maintain state info (a la VC) of allocated resources, QoS req's
- admit/deny new call setup requests

Intserv: QoS guarantee scenario

- Resource reservation
  - call setup, signaling (RSVP)
  - traffic, QoS declaration
  - per-element admission control



- QoS-sensitive scheduling (e.g., WFQ)

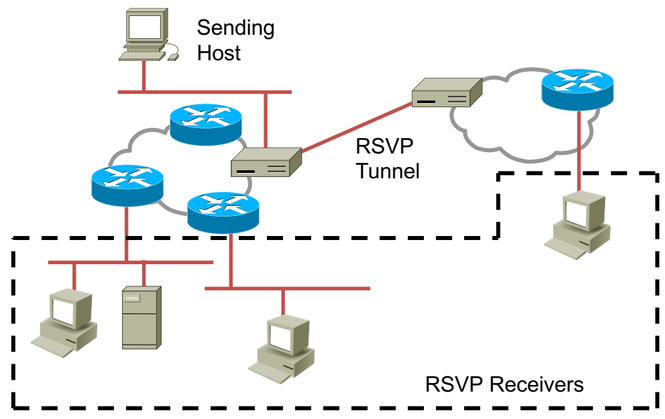
Call Admission

Arriving session must :

- declare its QoS requirement
  - R-spec: defines the QoS being requested
- characterize traffic it will send into network
  - T-spec: defines traffic characteristics
- signaling protocol: needed to carry R-spec and T-spec to routers (where reservation is required)
  - RSVP

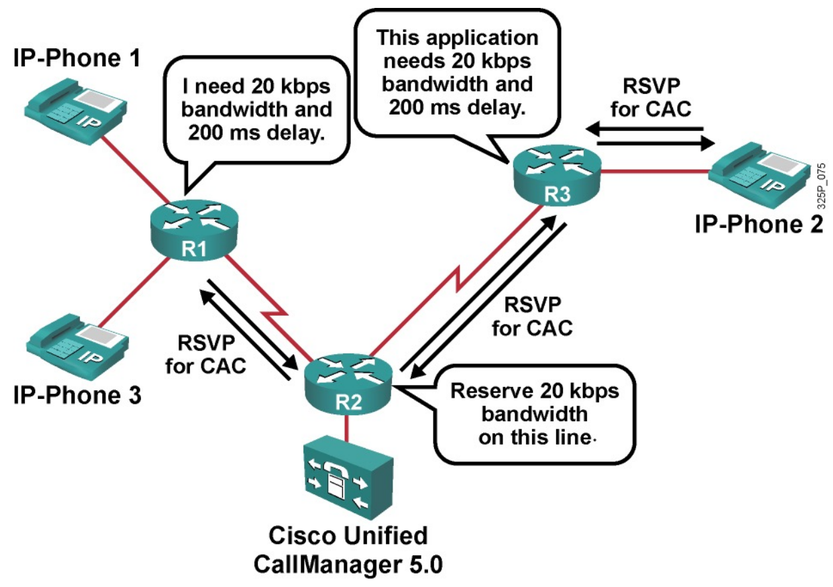
Resource Reservation Protocol (RSVP)

- Is carried in IP—protocol ID 46
- Can use both TCP and UDP port 3455
- Is a signaling protocol and works with existing routing protocols
- Requests QoS parameters from all devices between the source and destination



- Provides divergent performance requirements for multimedia applications:
  - ✓ Rate-sensitive traffic
  - ✓ Delay-sensitive traffic

RSVP in Action



- RSVP sets up a path through the network with the requested QoS.
- RSVP is used for CAC in Cisco Unified CallManager 5.0.

Benefits and Drawbacks

- Benefits:
  - Explicit resource admission control (end to end)
  - Per-request policy admission control (authorization object, policy object)
- Drawbacks:
  - Continuous signaling because of stateful architecture
  - Flow-based approach not scalable to large implementations, such as the public Internet

IETF Differentiated Services

Concerns with Intserv:

- Scalability: signaling, maintaining per-flow router state difficult with large number of flows
- Flexible Service Models: Intserv has only two classes. Also want "qualitative" service classes

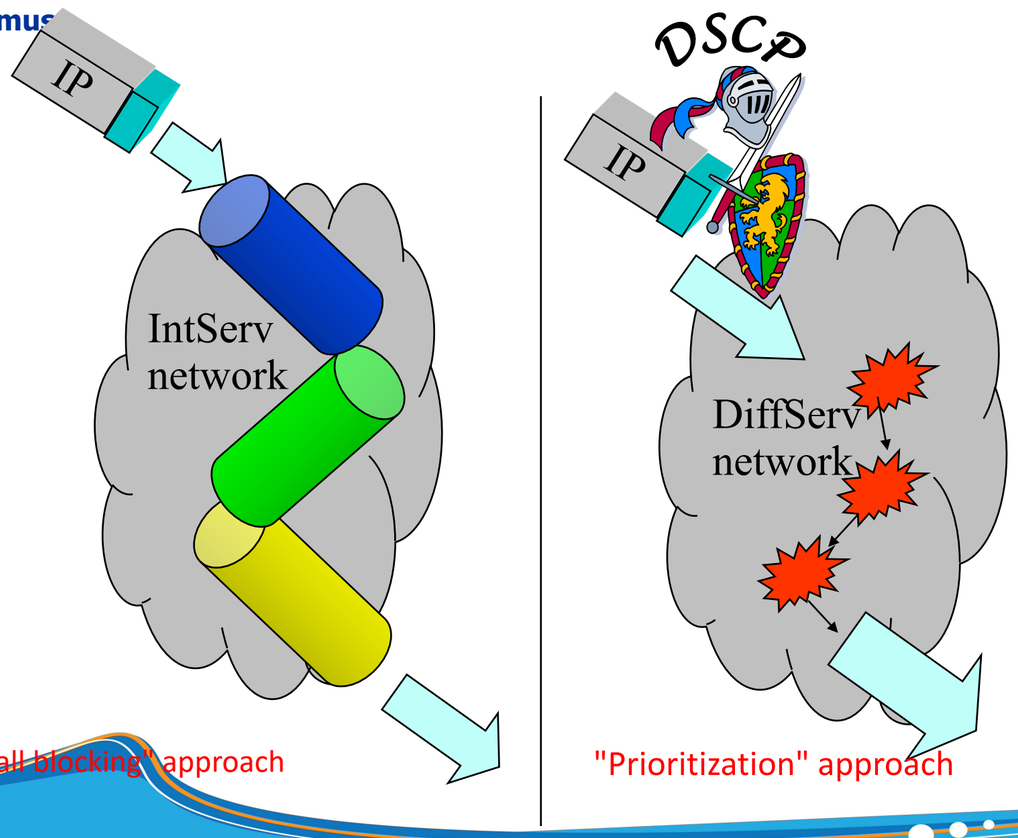
Diffserv approach:

- Simple functions in network core, relatively complex functions at edge routers (or hosts)
- Don't define define service classes, provide functional components to build service classes

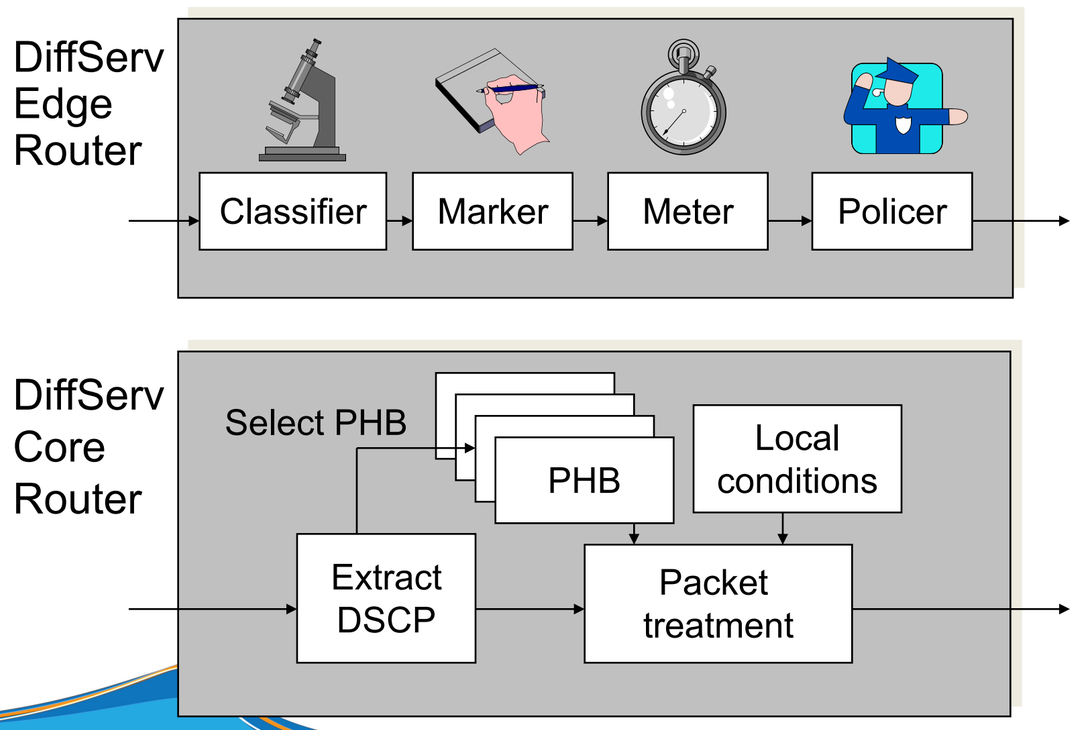
DiffServ Model

- Describes services associated with traffic classes, rather than traffic flows.
- Complex traffic classification and conditioning is performed at the network edge.
- No per-flow state in the core.
- The goal of the DiffServ model is scalability.
- Interoperability with non-DiffServ-compliant nodes

IntServ vs. DiffServ



DiffServ Routers



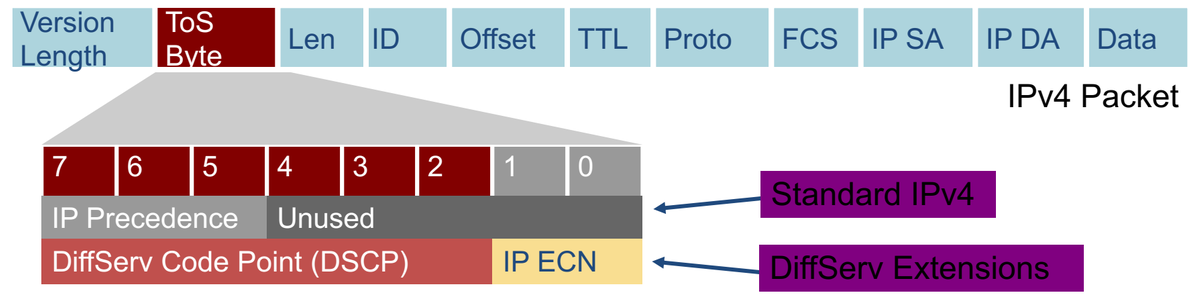
Classification

- Classification is the process of identifying and categorizing traffic into classes, typically based upon:
  - ✓ Incoming interface
  - ✓ IP precedence
  - ✓ DSCP
  - ✓ Source or destination address
  - ✓ Application
- Without classification, all packets are treated the same.
- Classification should take place as close to the source as possible.

Marking

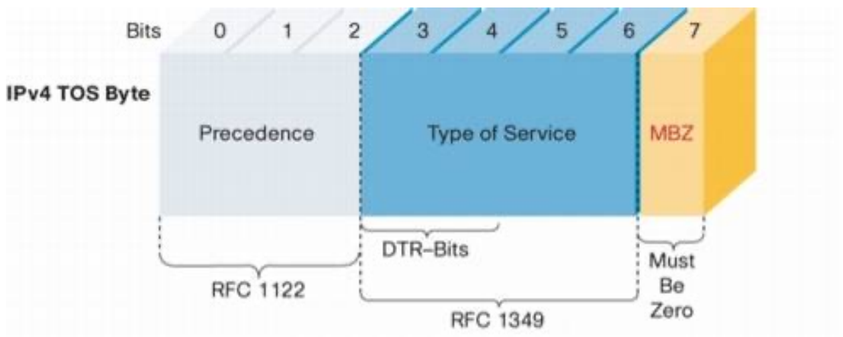
- Marking is the QoS feature component that “colors” a packet (frame) so it can be identified and distinguished from other packets (frames) in QoS treatment.
- Commonly used markers:
  - ✓ Link layer:
    - CoS (ISL, 802.1p)
    - MPLS EXP bits
    - Frame Relay
  - ✓ Network layer:
    - DSCP
    - IP precedence

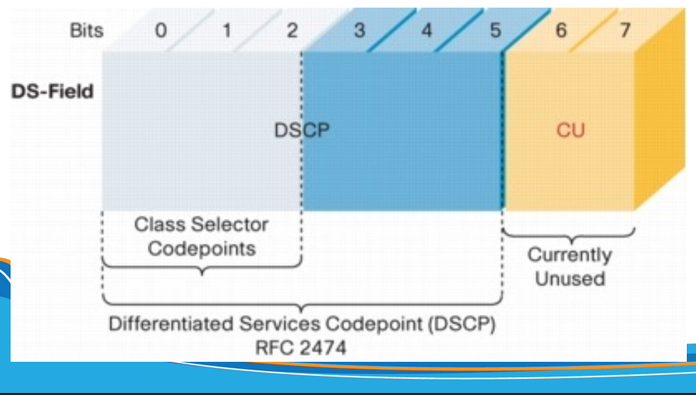
Classification Tools
IP Precedence and DiffServ Code Points



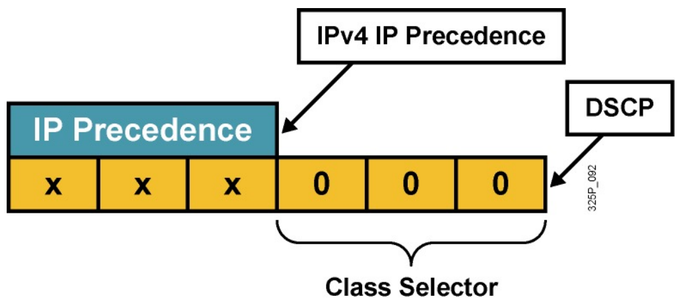
- IPv4: three most significant bits of ToS byte are called IP Precedence (IPP)—other bits unused
- DiffServ: six most significant bits of ToS byte are called DiffServ Code Point (DSCP)—remaining two bits used for flow control
- DSCP is backward-compatible with IP precedence

IP ToS Byte and DS Field Inside the IP Header





IP Precedence and DSCP Compatibility



- Compatibility with current IP precedence usage (RFC 1812)
- Differentiates probability of timely forwarding:

  (xyz000) >= (abc000) if xyz > abc
- That is, if a packet has DSCP value of 011000, it has a greater probability of timely forwarding than a packet with DSCP value of 001000.

Behavior Aggregator, Per-Hop Behaviors

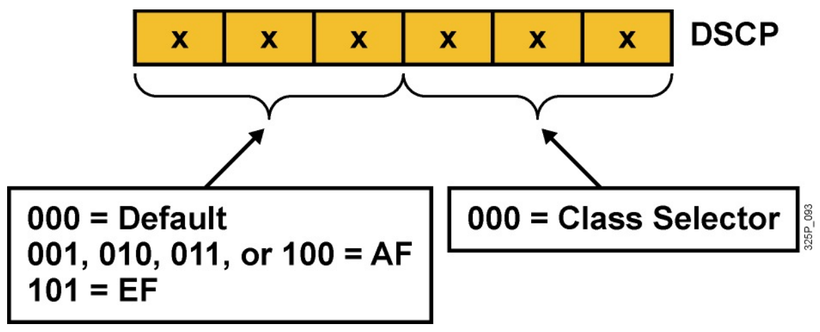
- Behavior Aggregate (BA): the collection of packets that have the same DSCP value (also called a codepoint) and crossing in a particular direction.
- Per Hop Behavior (PHB): the externally observable forwarding behavior applied at a DS-compliant node to a DS BA.

→ *PHB refers to the packet scheduling, queuing, policing, or shaping behavior of a node on any given packet belonging to a BA, and as configured by a Service Level Agreement (SLA) or policy.*

Per-Hop Behaviors

- *Default Forwarding* (DF) PHB — which is typically best-effort traffic
- *Expedited Forwarding* (EF) PHB — dedicated to low-loss, low-latency traffic
- *Assured Forwarding* (AF) PHB — gives assurance of delivery under prescribed conditions
- *Class Selector* PHBs — which maintain backward compatibility with the IP precedence field.

Per-Hop Behaviors



DSCP selects PHB throughout the network:

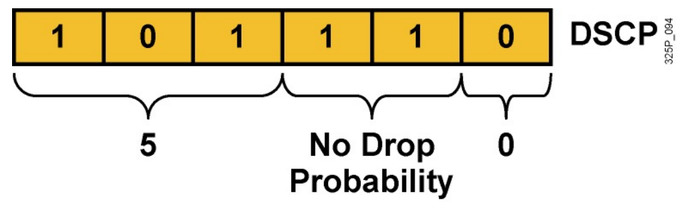
- ✓ Default PHB (FIFO, tail drop)
- ✓ Class-selector PHB (IP precedence)
- ✓ EF PHB
- ✓ AF PHB

Standard PHB Groups

PHB				DSCP			Maps to IP Precedence
Default (Best Effort)				000000 (0)			0
Scavenger (Less-than-Best-Effort)				001000 (8)			1
Assured Forwarding	Low Drop Pref.	Med Drop Pref.	High Drop Pref.				
Class 1	AF11	AF12	AF13	001010 (10)	001100 (12)	001110 (14)	1
Class 2	AF21	AF22	AF23	010010 (18)	010100 (20)	010110 (22)	2
Class 3	AF31	AF32	AF33	011010 (26)	011100 (28)	011110 (30)	3
Class 4	AF41	AF42	AF43	100010 (34)	100100 (36)	100110 (38)	4
Expedited Forwarding	EF			101110 (46)			5

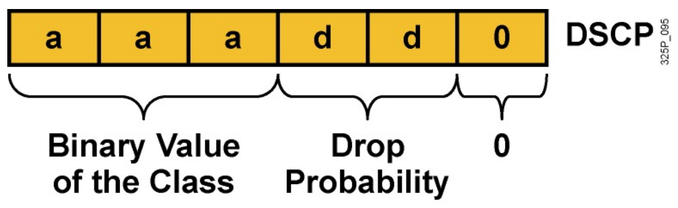
325P_098

Expedited Forwarding (EF) PHB



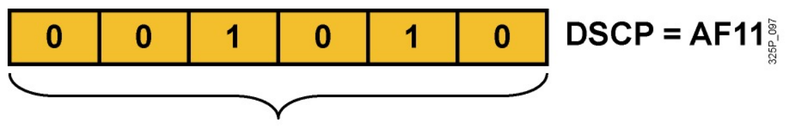
- EF PHB:
  - ✓ Ensures a minimum departure rate
  - ✓ Guarantees bandwidth—class guaranteed an amount of bandwidth with prioritized forwarding
  - ✓ Polices bandwidth—class not allowed to exceed the guaranteed amount (excess traffic is dropped)
- DSCP value of 101110: Looks like IP precedence 5 to non-DiffServ-compliant devices:
  - ✓ Bits 5 to 7: 101 = 5 (same 3 bits are used for IP precedence)
  - ✓ Bits 3 and 4: 11 = No drop probability
  - ✓ Bit 2: Just 0

Assured Forwarding (AF) PHB



- AF PHB:
  - ✓ Guarantees bandwidth
  - ✓ Allows access to extra bandwidth, if available
- Four standard classes: AF1, AF2, AF3, and AF4
- DSCP value range of aaadd0:
  - ✓ aaa is a binary value of the class
  - ✓ dd is drop probability

AF PHB Values

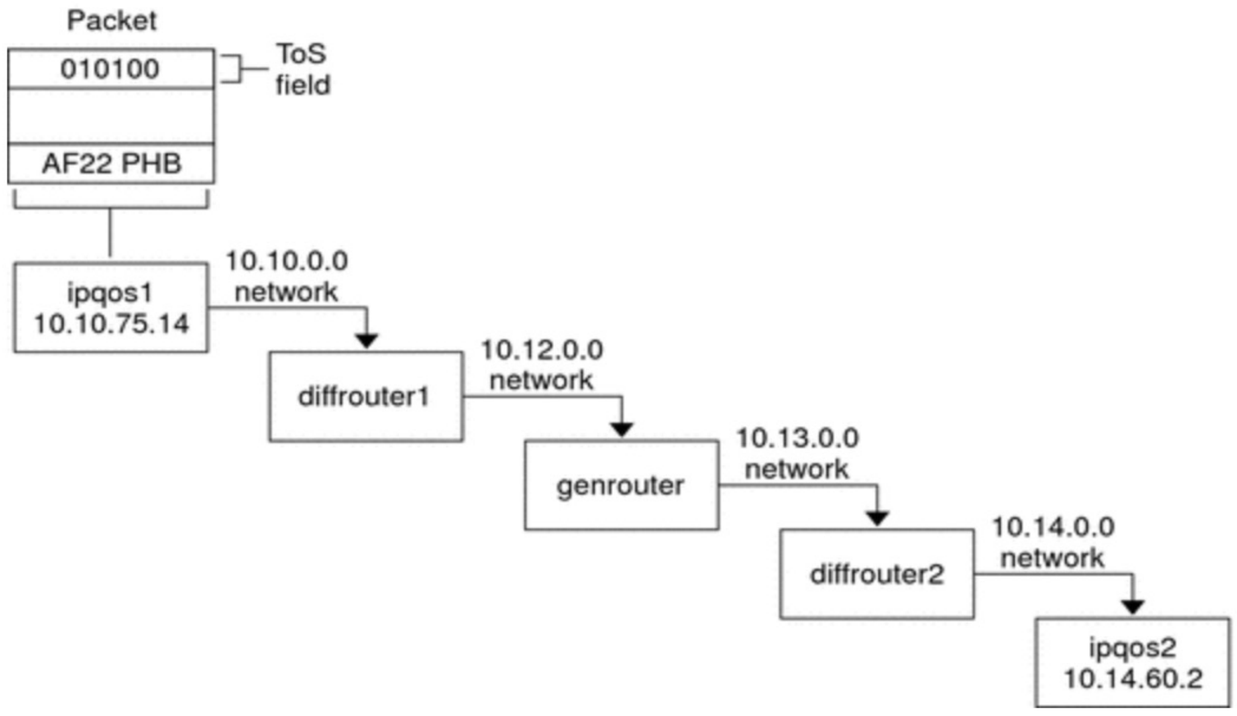


Class	Value		
AF1	001	dd	
AF2	010	dd	0
AF3	011	dd	0
AF4	100	dd	0

Drop Probability (dd)	Value	AF Value
Low	01	AF11
Medium	10	AF12
High	11	AF13

- Each AF class uses three DSCP values.
- Each AF class is independently forwarded with its guaranteed bandwidth.
- Congestion avoidance is used within each class to prevent congestion within the class.

Example



Example

1. The user on ipqos1 runs the ftp command to access host ipqos2, which is three hops away.
2. ipqos1 applies its QoS policy to the resulting packet flow. ipqos1 then successfully classifies the ftp traffic.
3. The system administrator has created a class for all outgoing ftp traffic that originates on the local network 10.10.0.0. Traffic for the ftp class is assigned the AF22 per-hop behavior: class two, medium-drop precedence. A traffic flow rate of 2Mb/sec is configured for the ftp class.
4. ipqos-1 meters the ftp flow to determine if the flow exceeds the committed rate of 2 Mbit/sec.
5. The marker on ipqos1 marks the DS fields in the outgoing ftp packets with the 010100 DSCP, corresponding to the AF22 PHB.
6. The router diffrouter1 receives the ftp packets. diffrouter1 then checks the DSCP. If diffrouter1 is congested, packets that are marked with AF22 are dropped.
7. ftp traffic is forwarded to the next hop in agreement with the per-hop behavior that is configured for AF22 in diffrouter1's files.
8. The ftp traffic traverses network 10.12.0.0 to genrouter, which is not Diffserv aware. As a result, the traffic receives “best-effort” forwarding behavior.
9. genrouter passes the ftp traffic to network 10.13.0.0, where the traffic is received by diffrouter2.
10. diffrouter2 is Diffserv aware. Therefore, the router forwards the ftp packets to the network in agreement with the PHB that is defined in the router policy for AF22

Questions ?