

Internet QoS: Differentiated Services

TELCOM2321 – CS2520

Wide Area Networks

Dr. Walter Cerroni

University of Bologna – Italy

Visiting Assistant Professor at SIS, Telecom Program

Slides based on Dr. Znati's material

Reading

- Textbook, Section 17.4

Differentiated Services

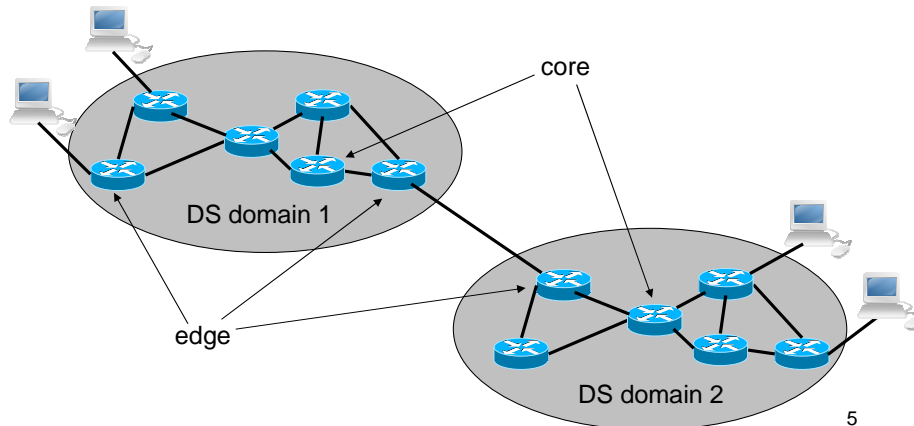
- The IETF DiffServ working group focused on the definition of an architecture with coarse service differentiation and related router functions
- RFC 2475 describes the DiffServ architecture
- RFC 2474 defines the DS field in the IP header
- RFC 2597 specifies the Assured Forwarding PHB
- RFC 2598 specifies the Expedited Forwarding PHB
- RFC 3246 updates the Expedited Forwarding PHB definition

DiffServ goals

- Create a relatively simple and coarse methods of providing differentiate classes of services for Internet traffic
 - the model must support various types of applications
 - service to be provided based on a Service Level Agreement (SLA) with the customer
- Keep the forwarding path simple to allow easy and early deployment
 - push complexity to network edge
- Avoid strong assumption about traffic types
 - assume the dominant internet traffic will be the best-effort traffic in a massive amount
 - priority treatment of packets are only for small portion of total traffic carried by the network

DiffServ architecture

- Built around the concept of a *domain* :
Contiguous region of a network using the same DS configuration

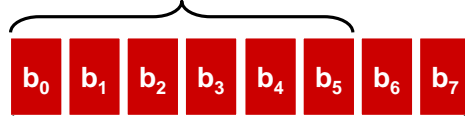


DiffServ architecture

- Distinction between *edge* and *core* routers
 - edge routers (ingress)
 - classify each incoming packet and associate it to the relevant Behavior Aggregate (BA)
 - mark packets with the corresponding service class identifier using the DS field in the IP header
 - perform per-aggregate shaping or policing (conditioning)
 - core routers
 - process packets and apply the appropriate Per-Hop Behavior (PHB) based on the DS field
- Scalable service differentiation achieved through the combination of
 - traffic conditioning and packet marking at the edges
 - local per-class forwarding behavior at the core
 - adequate network provisioning

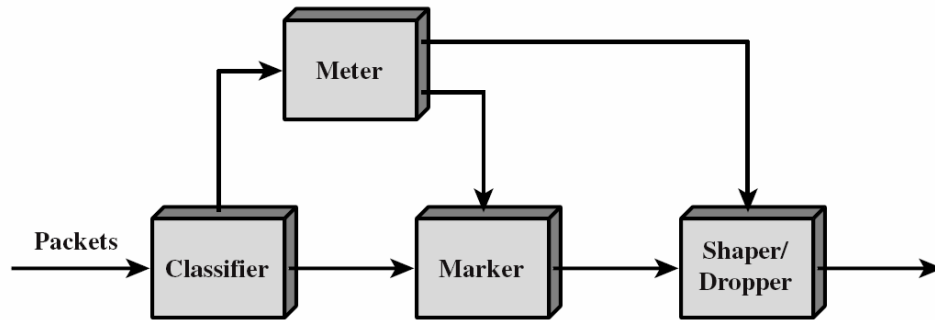
DS field in IPv4 header

DS Code Point (DSCP)



Version	IHL	Type of Service	Total Length			
Identification			R	D	M	Fragment Offset
Time to Live		Protocol	Header Checksum			
Source Address						
Destination Address						
Options			Padding			
Data						

DiffServ ingress router: Traffic conditioner



Packet classification

- In general a packet classifier selects packets based on the content of some fields in the packet header
- DiffServ defines two types of packet classifiers
 - Behavior Aggregate (BA) classifier
 - selects packets based on the DSCP value
 - Multi-Field (MF) classifier
 - uses a combination of the values of one or more header fields
 - e.g. source and destination addresses, DS field, protocol ID, source port, destination port, incoming interface, etc.
 - the result of the classification is written in the DS field
- Classification used by core routers to service packets

Traffic conditioning

- Traffic conditioning may consist of various elements that perform metering, marking, shaping, and dropping
- A *meter* measures the incoming traffic to determine whether a particular packet is in-profile or out-of-profile, according to the SLA
- A *marker* writes or re-writes the DS field of a packet header to a specific DSCP to reflect a particular DS behavior aggregate
- A *shaper* may delay traffic to bring it into conformance to the agreed profile
- A *dropper* may discard packets to bring traffic stream into conformance to the agreed profile (policing)

DiffServ core router: Per-Hop Behavior

- PHB describes the externally observable forwarding behavior applied to a particular behavior aggregate
- PHB is the means by which core routers allocate resources to different behavior aggregates
- A specific PHB can be defined in terms of
 - amount of allocated resources, e.g. buffer size and link bandwidth
 - the relative priority of the PHB compared to other PHBs
 - the observable traffic characteristics, in terms of delay, jitter and loss
- PHBs are not defined in terms of specific implementation mechanisms
 - this provides flexibility to use different mechanisms to implement a specific PHB

DiffServ PHBs

- Expedited Forwarding service (EF)
- Assured Forwarding service (AF)
- Best Effort service (BE)
 - default PHB: packets for which no agreements are in place are classified as BE service
 - DSCP = 000000
 - network will deliver as many of these packets as possible, as soon as possible, but there are no guarantees
- It is expected that new DSCPs may be assigned by a provider to its own local use, or by future standard definitions

Expedited Forwarding

- EF PHB, also called Premium Service, is designed to provide low loss, low delay, low jitter, assured bandwidth, end-to-end service
 - EF PHB simulates virtual leased lines to support highly reliable voice or video applications
 - recommended DSCP code is 101110
- Network guarantees that EF packets encounter extremely short or empty queues to ensure minimal delay and jitter
 - can only be achieved if the EF packet service rate on a given output port exceeds the typical EF packet arrival rate at that port
 - independently of the load on other non-EF PHBs
- Users are not allowed to send more traffic than the size of the pipe
 - when excess occurs, traffic is delayed and dropped when buffer overflows

Enforcing EF PHB

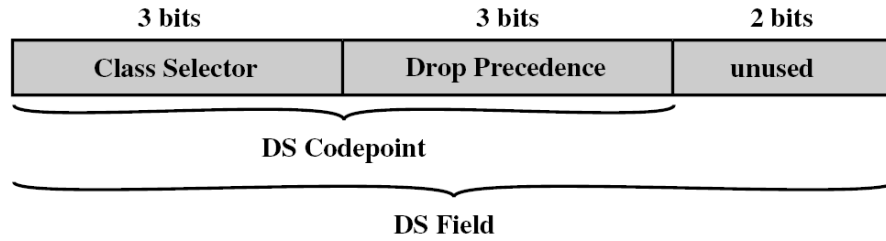
EF PHB can be enforced in different ways

1. By policing EF micro-flows to prescribed values at the edge of the DS Domain, ensuring that service rate for EF packets exceeds their arrival rate in the core of the network
2. By ensuring adequate bandwidth provisioning across the core network
3. By placing EF packets in the highest strict-priority queue and ensuring that the minimum output rate is at least equal to the maximum input rate
4. By rate-limiting the EF aggregate load in the core of the network to prevent inadequate bandwidth for other service classes

Assured Forwarding

- AF PHBs are designed to ensure that packets are forwarded with a high probability of delivery, as long as the aggregate traffic in a forwarding class does not exceed the subscribed information rate
 - if ingress traffic exceeds its subscribed information rate, then out-of-profile traffic is not delivered with as high a probability as traffic that is in-profile
- The AF PHB group includes four traffic classes
 - packets within each AF class can be marked with one of three possible drop-precedence values
 - the AF PHB group can be used to implement gold, silver, and bronze services
 - further differentiation can be achieved using low, medium or high drop precedence within a service class

Recommended DSCP for AF PHBs



Class Selector		Drop Precedence	
100	Class 4 - best service	010	Low - most important
011	Class 3	100	Medium
010	Class 2	110	High - least important
001	Class 1		

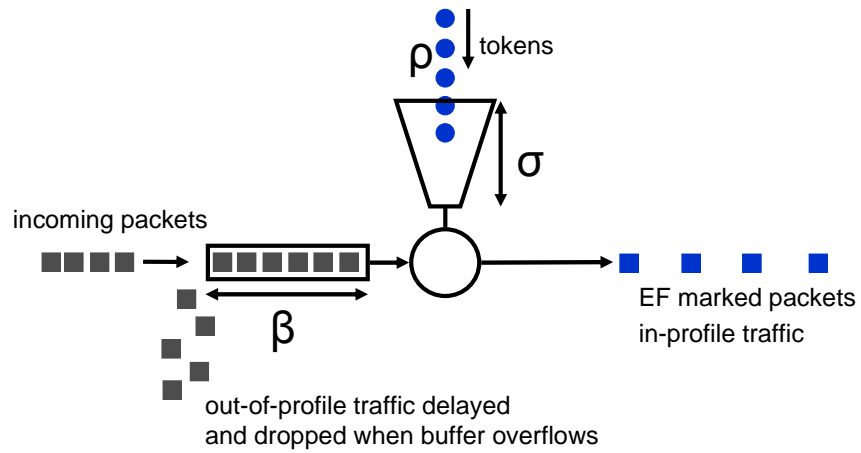
Enforcing AF PHBs

AF PHBs can be enforced in different ways

1. By policing AF micro-flows to prescribed values at the edge of the DS Domain
2. By ensuring adequate bandwidth provisioning across the core network
3. By placing each AF service class in a separate queue
4. By selecting the suitable queue scheduling discipline to allocate appropriate buffer space and bandwidth to each AF class
5. By configuring RED to honor the drop precedence bits, determining how aggressively packets of each class are dropped in case of congestion

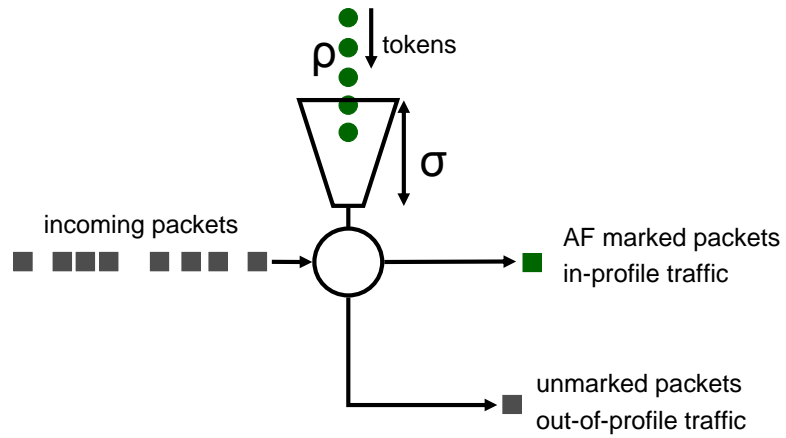
EF traffic conditioning

- Profile defined as average rate and maximum burst size
- Traffic conditioner implemented with a token bucket *shaping* scheme

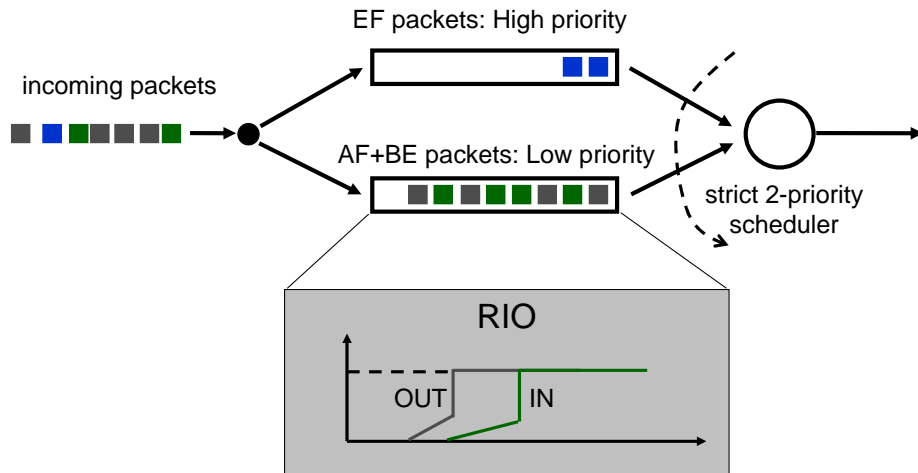


AF traffic conditioning: 2-colors marker

- Profile defined as average rate and maximum burst size
- Traffic conditioner implemented with a token bucket *policing* scheme



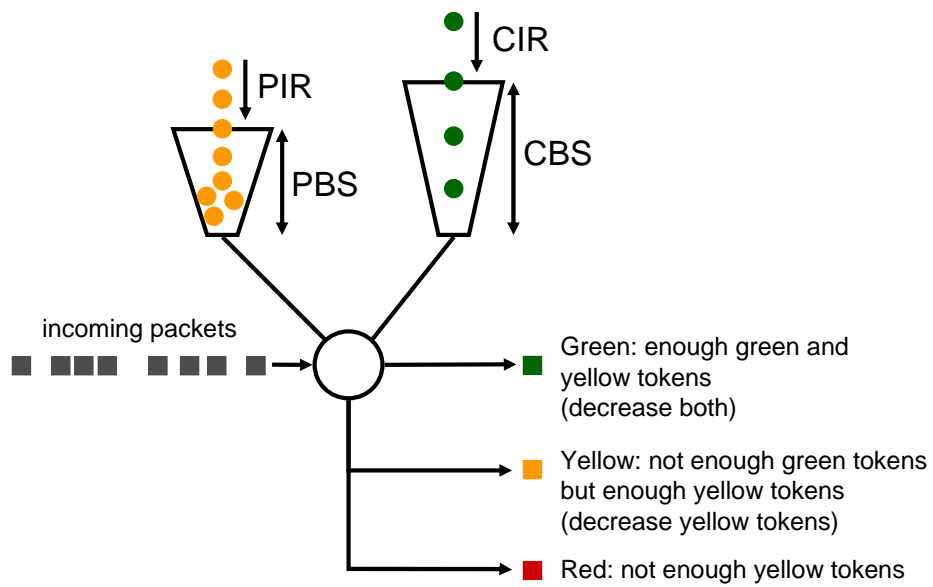
Queuing and scheduling at core routers



AF conditioning: 2-rates 3-colors marker

- 2-rates 3-colors marker (trTCM) exploits the three AF dropping precedence values
- Useful when a peak rate needs to be enforced separately from a committed rate
- Committed Information Rate (CIR) and Committed Burst Size (CBS)
- Peak Information Rate (PIR) and Peak Burst Size (PBS)
- $PIR \geq CIR$, PBS and $CBS > 0$
- Uses committed and peak values to define two token buckets
- Assigns dropping precedence based on both token bucket levels

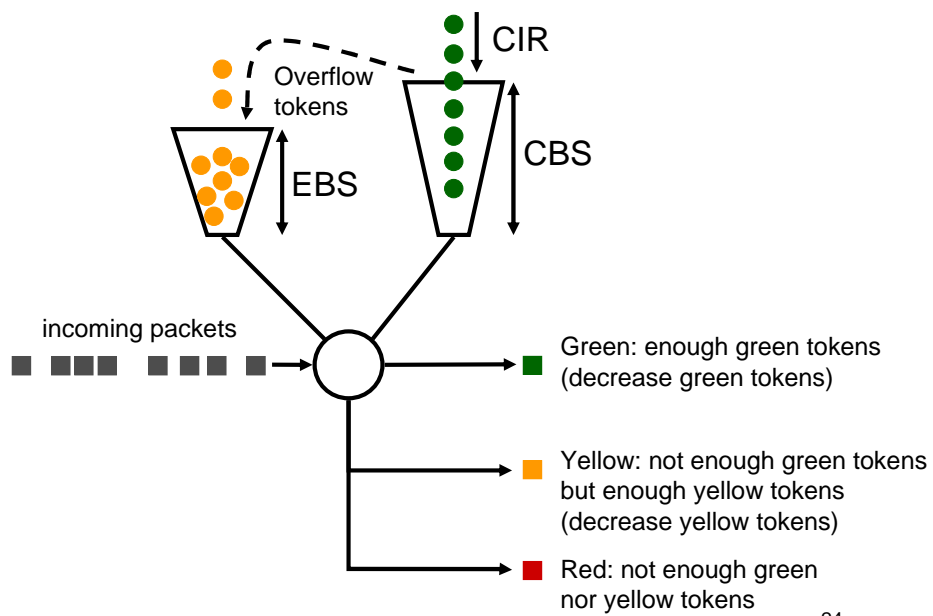
AF conditioning: 2-rates 3-colors marker



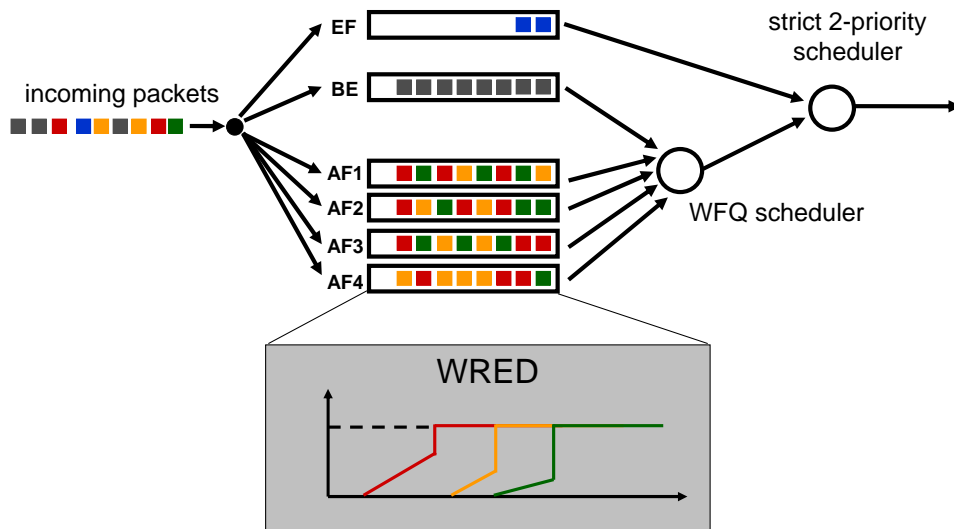
AF conditioning: Single-rate 3-colors marker

- Single-rate 3-colors marker (srTCM) also exploits the three AF dropping precedence values
- Useful when only the burst length determines service eligibility
- Committed Information Rate (CIR)
- Committed Burst Size (CBS)
- Excess Burst Size (EBS)
- $CIR > 0$, CBS or EBS or both > 0
- Uses CBS and EBS to define two token buckets sharing the same token rate CIR
- Assigns dropping precedence based on both token bucket levels

AF conditioning: Single-rate 3-colors marker

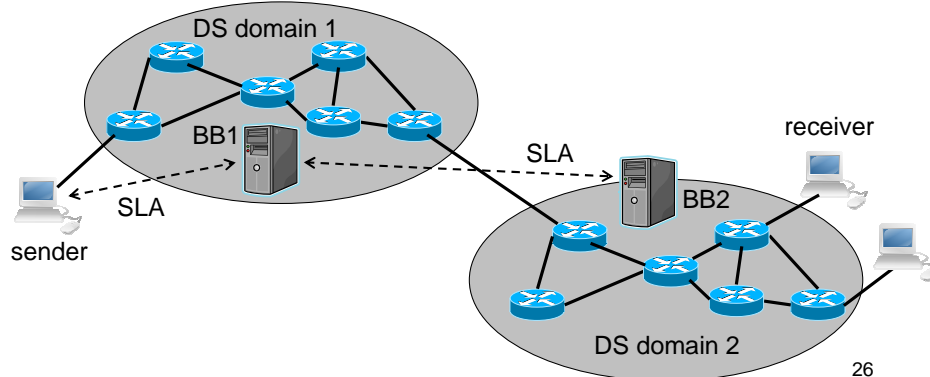


Queuing and scheduling at core routers



Admission control: Bandwidth broker

- Each domain is assigned a Bandwidth Broker (BB)
 - usually, used to perform ingress-egress bandwidth allocation
- BB is responsible to perform admission control in the entire domain
- SLAs needed also between domains



DiffServ summary

- DiffServ is more scalable than IntServ
 - coarse-grain “Class of Service”
 - edge routers maintain per aggregate state
 - core routers maintain state only for a few traffic classes
 - simplified core operations by pushing the complexity to the edge
 - classification and conditioning only at edge
 - no hop-by-hop signaling in the core
 - uniform forwarding treatments to each class of packets at every node
 - stateless uniform forwarding (PHB)
- DiffServ provides weaker services than IntServ
 - per-aggregate bandwidth and delay guarantees (EF) vs. per-flow bandwidth and delay guarantees
 - hard guarantee may be difficult
- BB is not an entirely solved problem
 - single point of failure
 - handle only long term reservations (hours, days)

Internet QoS: Comparison

	Best Effort	IntServ	DiffServ
Service	Connectivity No isolation No guarantees	Per-flow isolation and guarantees	Per-aggregate isolation and guarantees
Service scope	End-to-end	End-to-end	Domain
Complexity	No setup	Per-flow setup	Long-term setup
Scalability	Highly scalable (nodes maintain routing state only)	Not scalable (nodes maintain per-flow state)	Scalable (edges maintain per- aggregate state, core nodes per-class state)