Quality of Service Concepts in IP Networks

Introduction

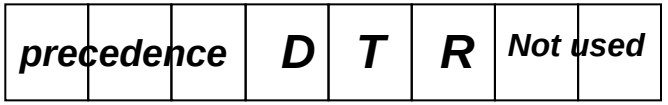
QoS in packet switching networks

- **WAN Technologies**
  - **FR: <Tc, CIR, EIR>**
  - **ATM**
    - **Traffic: PCR,SCR,MBS,MCR**
    - **QoS: CLR,CTD,CDV**
    - **Categories: CBR,VBR-RT,VBR-nRT,ABR,GFR,UBR**
    - **Traffic Control:**
      - CAC
      - UPC
      - Priority Control
      - Flow Control (EFCN, discard)
      - Congestion Control (preventive, reactive)
  - **IP**
  - **MPLS**

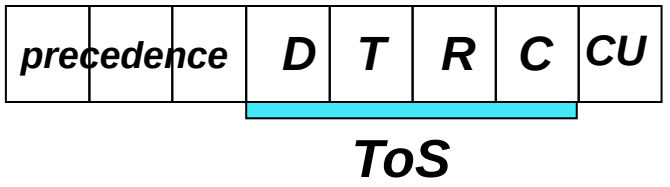
- **LAN Technologies**
  - **Ethernet IEEE 802.1p**
    - ✓ **Extension of 802.1q (tag VLAN added to the MAC layer). The VLAN tag has two parts: VLAN ID (12-bit) and Priority (3-bit), this last field specified in the 802.1p standard.**
    - ✓ **Priority Queuing**
    - ✓ **No reservations.**
  - **Extensions to other contexts (home, MAN, RAN)**
    - ✓ **TSN (Time Sensitive Networking) group**

QoS in the IP specification

- RFC 791
  - TOS (bits 9-16)
    - ✓ precedence: routine (ooo), priority, immediate, flash, flash override, Critical, Internetwork control, Network control (111)



- RFC 1349



- No consensus on its interpretation and implementation

Definitions

QoS (Quality of Service)

- A set of network performance characteristics such as delay, jitter, bit error rate, packet loss, etc affecting a flow. In a wide sense, it includes traffic parameters too. e.g. average bit rate, peak bit rate, maximum burst size available to the flow.
- Flow:
  - ✓ A stream of related packets from a sending application to a receiving application traversing a network that requires the same QoS
  - ✓ It is simplex

CoS (Class of Service)

- QoS for a set of packets from a number of flows classified as belonging to the same class. No per-flow treatment.
- Class (or Traffic Class):
  - ✓ Aggregate of flows with a common treatment by the router or network in terms QoS

Internet QoS models proposed in the IETF

- **IntServ (Integrated Services Internet) RFC1633**
  - **Flow-based processing**
  - **Dynamic soft-state reservations for individual flows**
  - **Signalling: RSVP (Resource reservation protocol).**
    - ✓ **Two basic messages:**
      - PATH : from sender to receiver
      - RESV : from receiver to sender
- **DiffServ (Differentiated Services Internet) RFC2474, RFC2475**
  - **Class-based processing**
  - **Policies pre-set across the network**
  - **No signalling required**
- **Best-effort and over-provisioning**

Necessary functions

- **Traffic Control**
  - **Packet classification**
    - Various queues with different treatments
  - **Queue management**
    - Usage policy of the capacity of each queue to provide a service
  - **Packet scheduling**
    - Link sharing algorithm to provide a service
  - **Admission Control**
    - Accept/reject reservation requests
  - **Shaping**
  - **Policing (Use control)**
    - To check compliance to contract of a flow or class
    - Separation between flows/classes
- **Signaling**
  - **Resource reservation protocol**

DiffServ Model

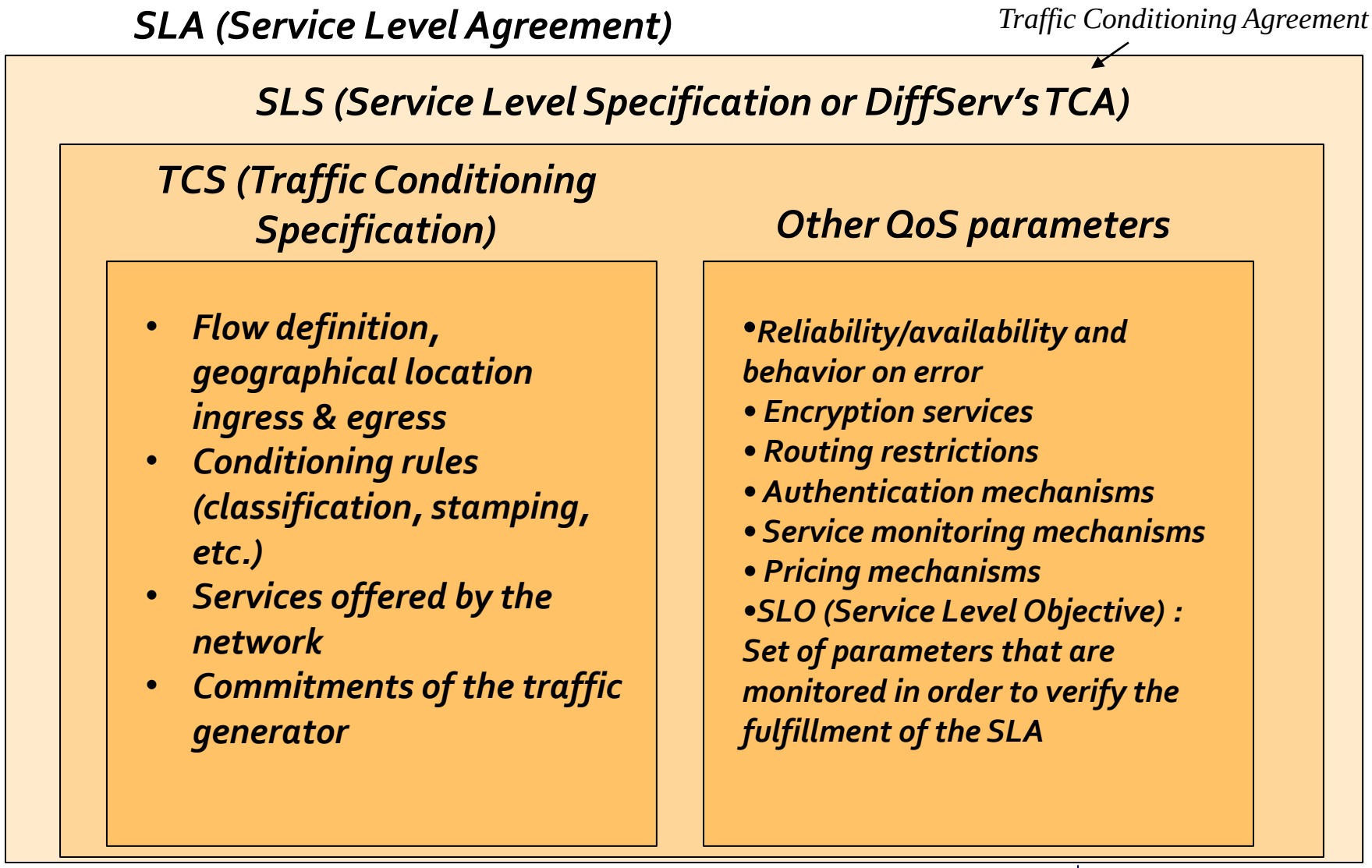
DiffServ and services

- **The DiffServ architecture standardizes PHBs, NOT services.**
  - **Nor does it standardize the PHB implementation**
  - **Services can be defined using the PHBs**
- **End-to-end service definition**
- **DS Domain: A network with a set of routers with the same policies.**
- **Types of service. Examples:**
  - **Qualitative**
    - **Transport of low latency traffic**
    - **Transport of low loss traffic**
      - Can only be checked by comparison
  - **Quantitative**
    - **90% of the traffic in a class will experience a delay of less than 50 ms.**
    - **The loss of traffic in a given sample will be less than 5%**
      - Can be verified, without needing comparisons
    - **The class makes use of at least 20% of the link capacity; the spare capacity of the link is shared in a fair manner.**

RFC2475, 3198 Terminology

Service Level Agreement (SLA)	*Legal service contract: General terms and conditions*
Service Level Specification (SLS)	*Technical terms and conditions on how the client traffic will be treated by the SP. DSCPs, PHBs, etc*
Per Domain Behaviour (PDB)	*Service classes and end to end aggregates in a DiffServ domain*
Per Hop Behaviour (PHB)	*Externally observable router behavior upon a given DSCP*
Scheduling (WFQ,…) *Packet discard (RED, …)*	*Implementation of PHB*

Service specification of DiffServ



Classification and Marking of packets

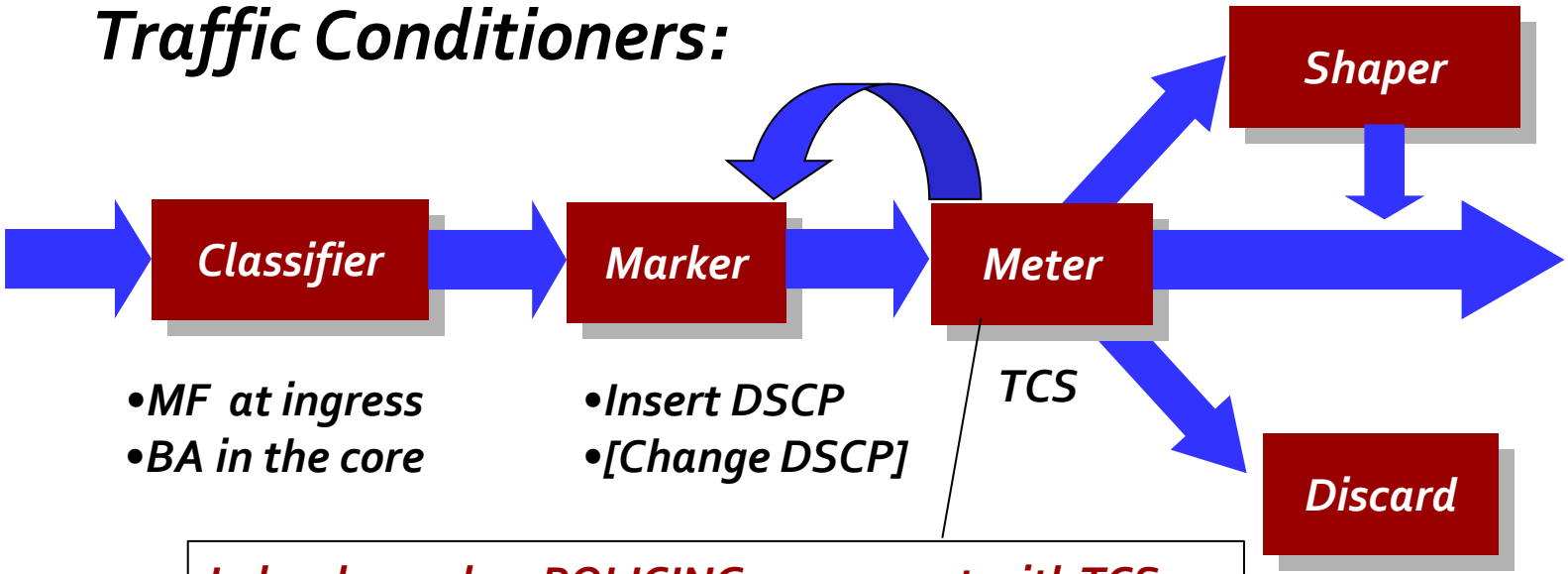
- DS Field (Differentiated Services)

  - *DSCP (Differentiated Services Code Point) determines the behaviour hop by hop: DSCP ↔ PHB*
    - *xxxxx0 standard action*
    - *xxxx11 experimental/local action*
    - *xxxx01 experimental/local action (subject to standardization)*
    - *000000 Best Effort forwarding*
    - *Xxx000 for IP precedence compatibility*
    - *CU not used (RFC 3168 uses it for ECN)*
  - *Note that it's at the network level: in the IPv4 or IPv6 header*
    - *IPv4 TOS (Type of Service)*
    - *IPv6 Traffic Class Field*

Classification

- **Identifiying a flow or a class of packets (how?)**
  - **Edge**
    - ✓ **MF (multi field) discriminator -- hardware support**
    - ✓ **<IPv6 source address, FlowID>**
    - ✓ **Application level: Bandwidth Managers**
      - ➢ Agents with knowledge of an organization's priorities
      - ➢ Manage admission control of entire flows (not packets!)
        - Edge routers mark or drop packets (not flows!)
  - **Core**
    - ✓ **DiffServ: DSCP (DiffServ Code Points) stamps**
    - ✓ **MPLS: label, TC (Traffic Class) field (3 bits, formerly known as EXP bit)**

A model of a DiffServ router



In border nodes: POLICING agreement with TCS

For example, token bucket metering

Traffic not obeying can (according to the SLS)

- Be declared as conforming
- Classified differently
- Discarded

Also used for accounting

A model of a DiffServ router

- *Classifier*:
  - The packet received by the DiffServ router is first classified by a classifier module. The classifier selects packets based on the values of one or more packet header fields.
- Two types of classification supported by Diffserv:
  - *Multifield (MF) classification*: Supports classification based on multiple fields (i.e. source and destination address, source and destination port, protocol identification). The ingress router of the Diffserv domain is the one that performs the MF classification.
  - *Behavior aggregate (BA) classification:* Sorts packets based on the ToS field that contains the DSCP. This classification is performed in the DSCP core routers and results in faster classification.

A model of a DiffServ router

- ***Marker***:
  - Inserts the appropriate DSCP value in the DS byte so that the packet receives appropriate service (PHB) in subsequent routers. Once the packet has been marked, all downstream routers need to perform only BA classification.
- ***Meter***:
  - A meter is used to compare the incoming flow with the negotiated traffic profile and pass the violating packets to the shaper and dropper or remark the packet with lower grade service using a different DSCP. The meter can be used for accounting management of the network.

A model of a DiffServ router

- *Shaper*:
  - A packet may be sent to the shaper module. This module may introduce some delay in order to bring the flow into compliance with its profile. The shapers usually have limited buffer, and packets that don't fit into the buffer may be discarded. The shaper buffers may accept a burst of traffic and then send it at an acceptable rate to the next hop.
- *Discard or Dropper*:
  - A dropper performs a policing function by simply dropping the packets that are out of profile. It is a special instance of a packet shaper with no buffer.

A model of a DiffServ router

- The meter, marker, shaper, and dropper are also known as traffic conditioners in the DiffServ world.
- Combination of these components facilitates building a scalable DiffServ network.
  - At the edge: MF classification combined with metering. Scalable, as the traffic volume is not very high (in comparison to the core).
  - At the core: Classification is performed based on BA (no need to maintain a per-flow state). QoS guarantees can be achieved by separating flows using different DSCP and by shaping and policing traffic.

Expedited Forwarding PHB - details

- Choice of service classes is up to each operator,
  - ..but packets are often forwarded between networks run by different operators,

→ The IETF defined two network-independent service classes: Expedited Forwarding and Assured Forwarding.

Expedited Forwarding (EF, RFC 3246)

- Low latency, jitter and minimum bandwidth guaranteed
- DSCP (Differentiated Services Code Point) recommended: 101110

Service goal in a domain:

- "Dedicated virtual line" or "Premium Service"
  - ✓ Abstraction of a link with a minimum link bandwidth guaranteed
  - ✓ The injected rate is measured over some interval of time (sliding window) equal or greater than the time of transmission of an MTU packet at the configured rate

$$T_{observation} = MTU/CIR$$

  - ✓ Application of interest: Voice over IP

Assured Forwarding PHB - details

Assured Forwarding (AF, RFC 2597)

- **Service goal in a DS domain:**
  - ✓ To offer different levels of forwarding guarantee to clients that have subscribed to the service
  - ✓ Example: gold, silver, bronze services.
  - ✓ Pre-marking by the provider or by the client.
- **4 independent AF classes**
  - ✓ For each class, a minimum bandwidth and memory is assigned
    - E.g. with WFQ . CAC is necessary.
  - ✓ Within each class, 3 packet discard priorities
    - Discard priorities are checked and applied at points of congestion
    - Discard priorities are implemented with WRED
    - A single queue per class, to avoid reordering

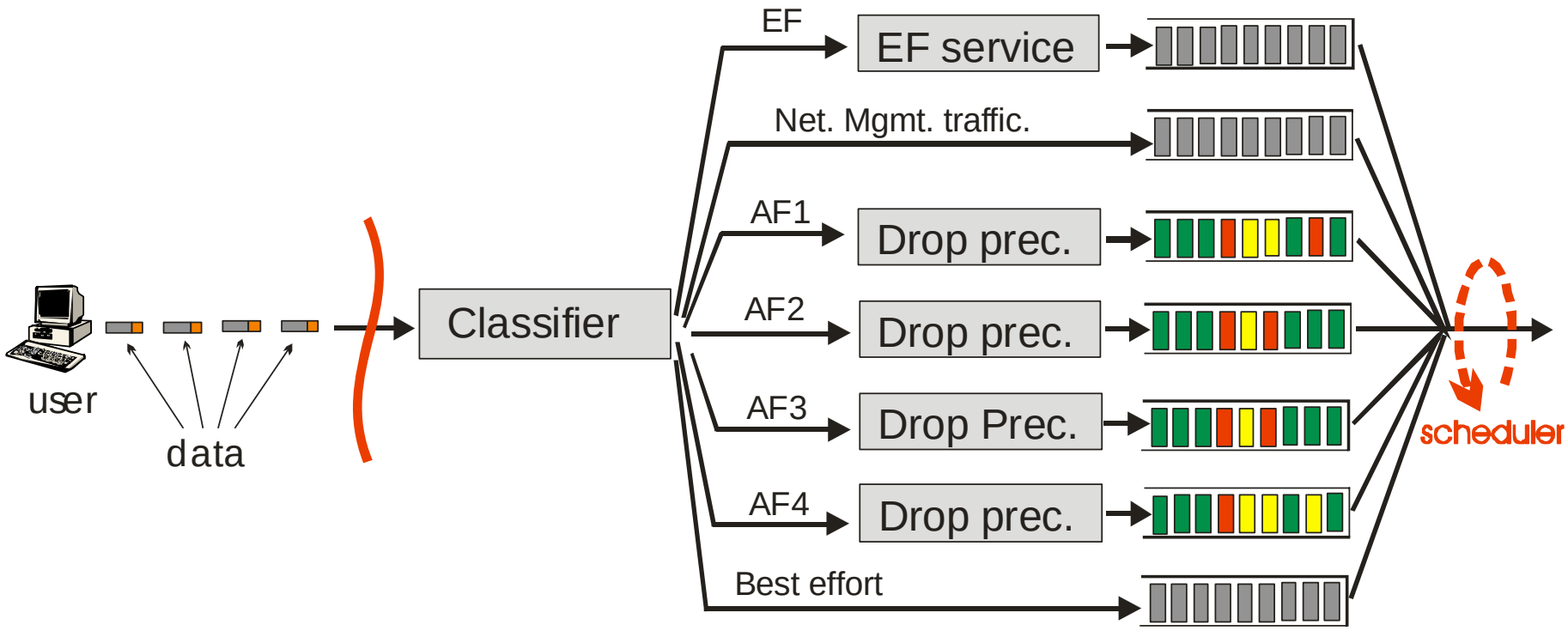
Assured Forwarding PHB - details

- DSCP

Drop precedence	Class AF1	Class AF2	Class AF3	Class AF4
Low 1	001010	010010	011010	100010
Medium 2	001100	010100	011100	100100
High 3	001110	010110	011110	100110

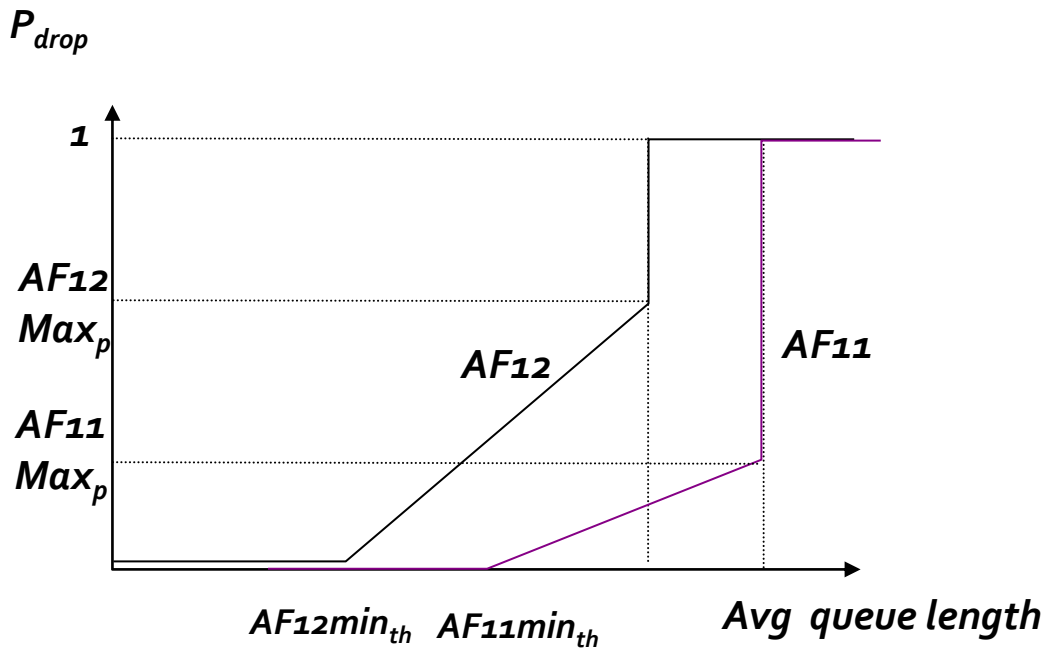
- *AFij, with i=class, j=drop precedence*
- *Example marking rule*
  - *Mark AF11: 1 Mbps : Any egress point: excess traffic is marked as AF43:*

DiffServ Router - example



WRED

- Allows to differentiate packet drop probability between classes



- Used in DiffServ

Service Types

RFC2210 specifies three end-to-end reservation service types:

- ***Guaranteed service*** **is intended to support inelastic applications with low-delay, low-jitter, low-loss, assured bandwidth requirements, such as VoIP and video. By comparison, such applications are typically supported with an EF PHB where DiffServ is deployed.**
- ***Controlled load service*** **is intended to support elastic applications with assured bandwidth requirements. By comparison, such applications are typically supported with an AF PHB where DiffServ is deployed. To quote RFC2211, the controlled load service provides: *"A QOS closely approximating the QOS that same flow would receive from an unloaded network element, but uses capacity (admission) control to assure that this service is received even when the network element is overloaded."***
- ***Best-effort service*** **is defined as the service, which flows receive that have neither had a successful GS or CL reservation established. The IntServ best-effort service is analogous to the service that would be supported with the default PHB where DiffServ is deployed.**

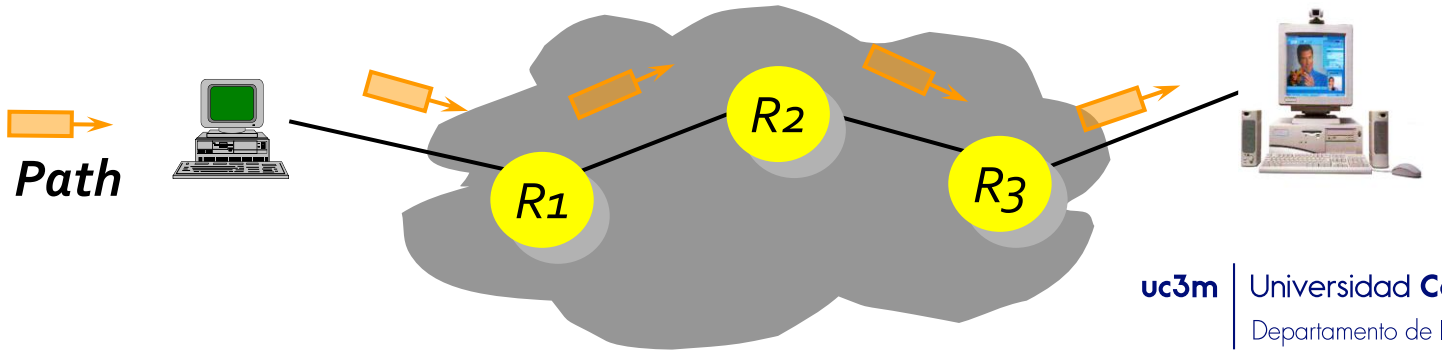
IntServ - RSVP (Resource reSerVation Protocol)

Introduction: RSVP objectives

- Node-to-node signaling
- Dynamic reservations
- Multicast-aware
  - Allows reservation aggregation
  - Protocol focused on the receiver
    - Heterogeneous receivers may make different reservations of the same multicast flow
- Routing Protocol independent
  - Separate signaling, but designed to follow the routing
  - Allows the possibility of topology changes
- Network protocol independent
  - IPv4, IPv6, ATM
- Load the network minimally

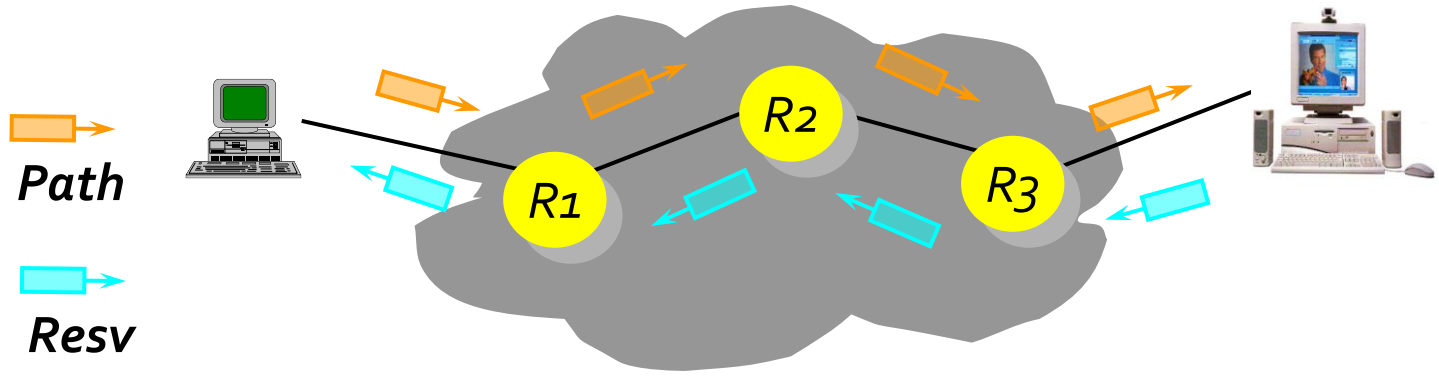
Reservations in RSVP

- How does the receiver know the path of the flow?
  - Source-destination path can be different from the reverse path
- Path messages: sourceIP->destinationIP
  - the same path as the flow
  - Identifies/characterises the flow and alert routers to wait for reservations coming upstream
  - The message carries the address of the previous router's output interface: r-1, r-2, r-3 (phop)

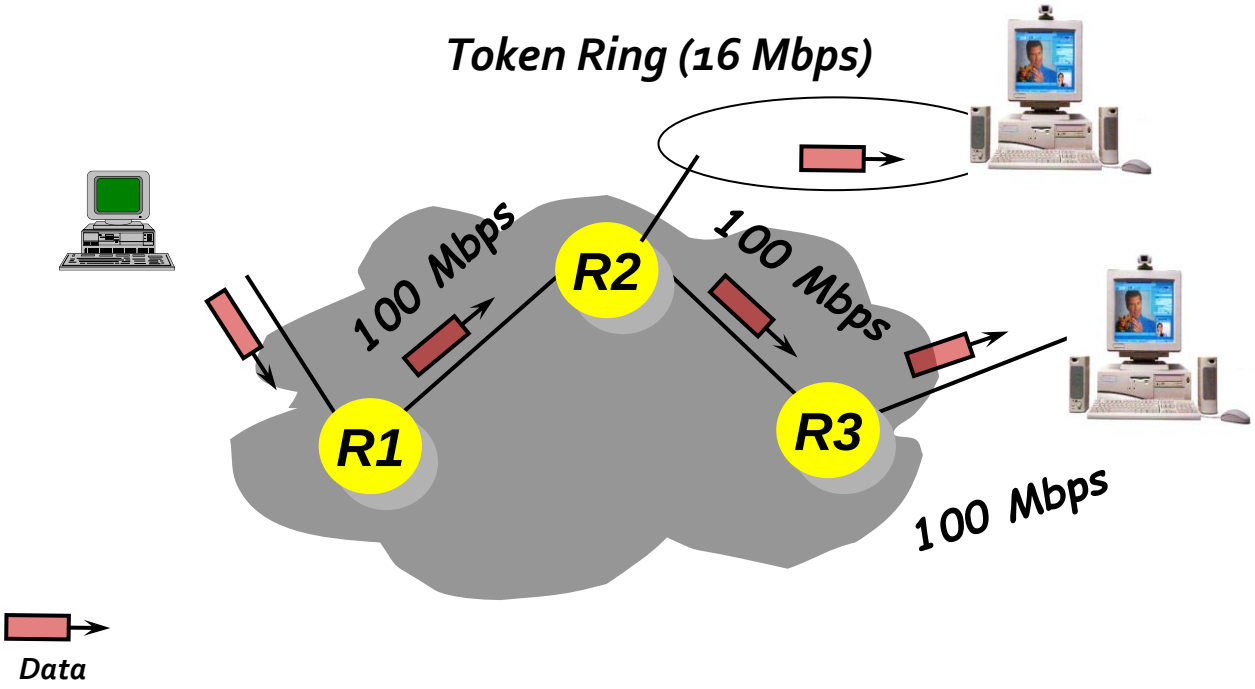


Reservations in RSVP

- **Resv message: sent by receiver**
  - Follows the same route as Path (backwards)
  - Describes the desired reservation
    - ✓ RSVP only transports it



Why is it started at the receiver?



Soft State

- IP networks are dynamic
  - Routes can change if a node does not work…
- Soft State philosophy
  - The reservation has a limited lifetime
  - You must refresh periodically the request, as if you were requesting new reservations
  - Source sends Path messages
    - ✓ In case the route changes
  - Destination sends Resv messages
    - ✓ To keep the reservation up
    - ✓ To update it (e.g. change rate)

RSVP packet format

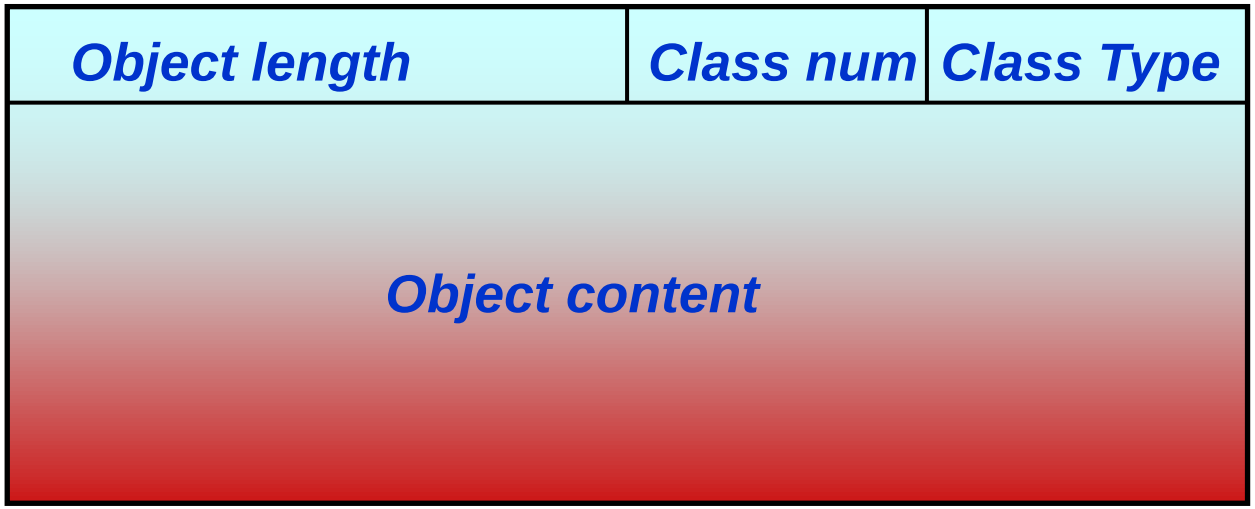
- RSVP works over IPv4, or directly over IPv6
  - Protocol identifier nº 46
  - If an element does not understand RSVP, it forwards the packet
- Common header to all the packets

Vers	*Flags*	*Msg Type*	*Checksum*
Send_TTL		*-------*	*RSVP Length (bytes)*

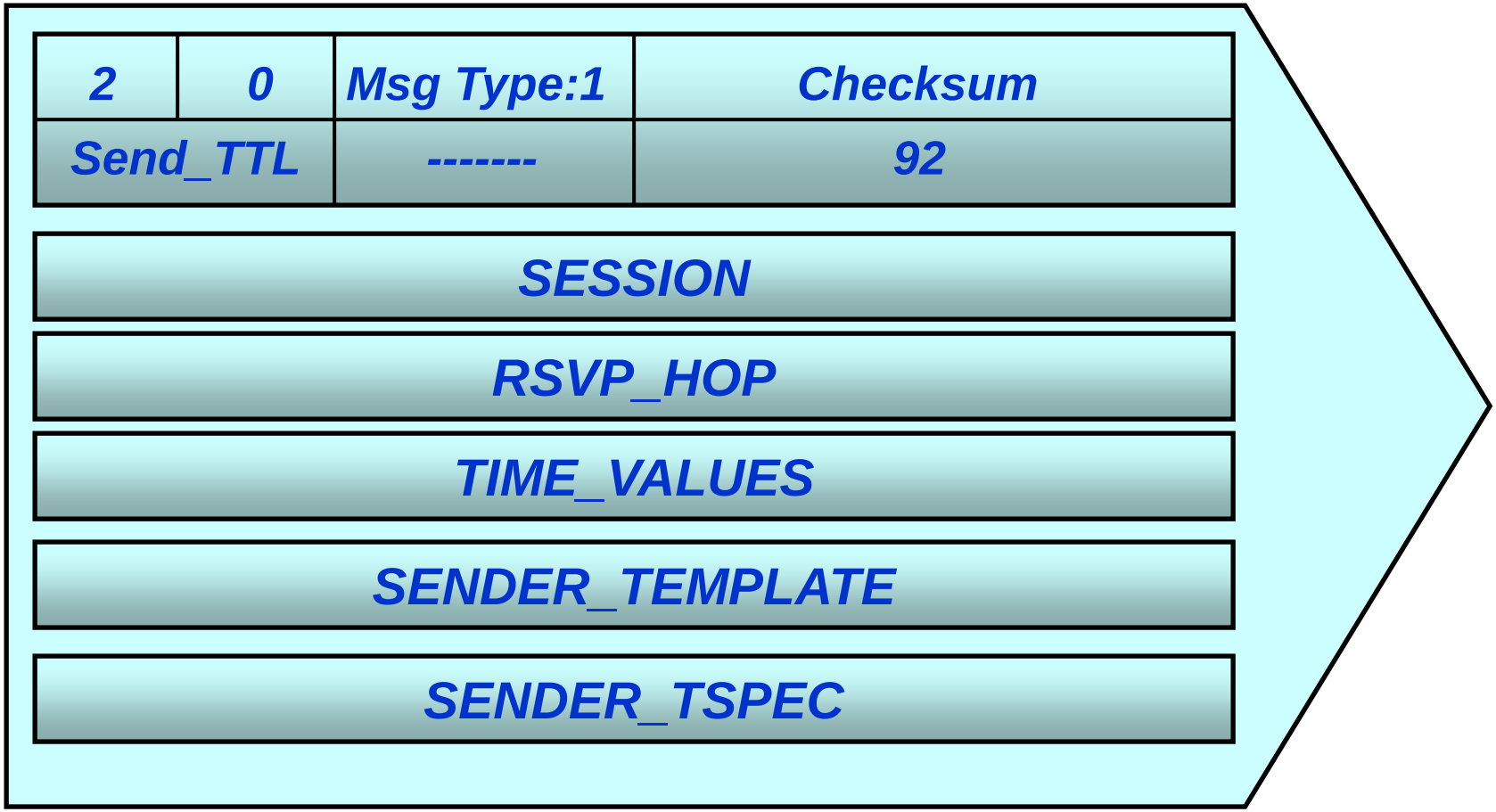
Types of messages	
1	Path
2	Resv
3	PathError
4	ResvError
5	PathTeardown
6	ResvTeardown
7	ResvConf

Format of MESSAGES

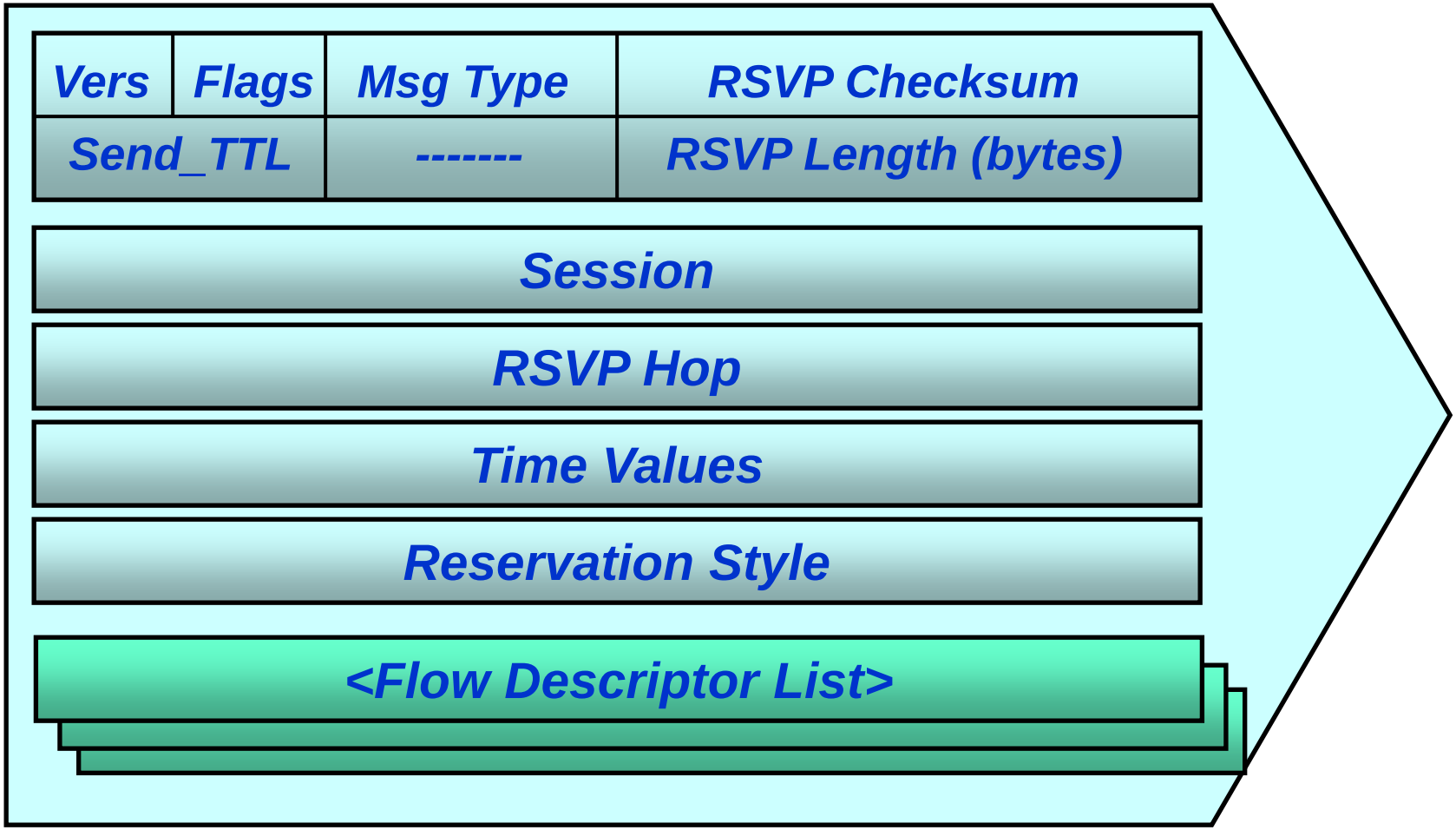
- Objects: blocks that can be inserted in the packets



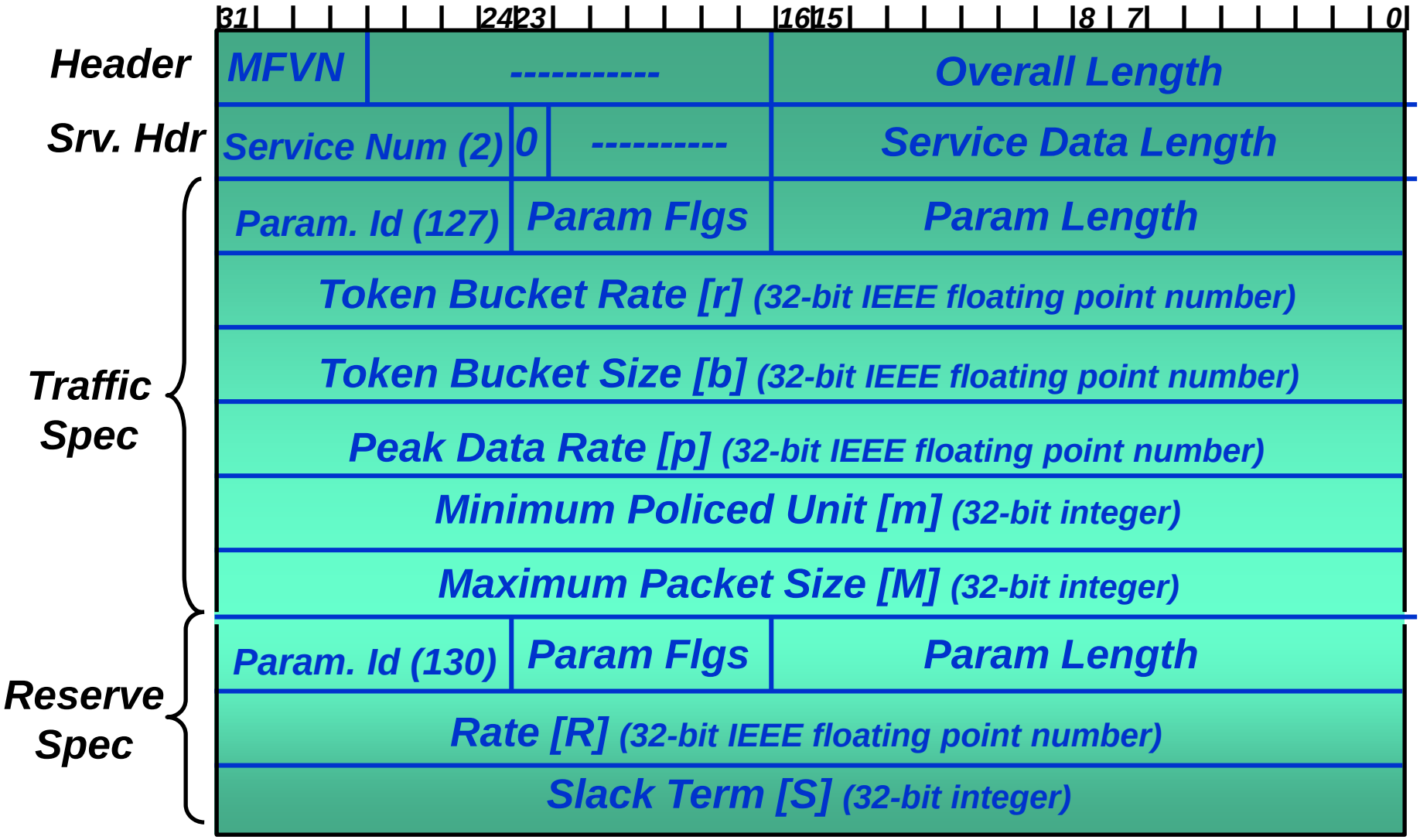
Path Message



Resv packet format



FLOW_SPEC object format



Main objective services of IntServ

- **Controlled load**
  - **best-effort emulation in a lightly loaded network**
    - ✓ **Very short delay, small losses because of queues full and small jitter**
    - ✓ **Behavior does not depend on the real load of the network**
  - **Implementation:**
    - ✓ **The network reserves enough bandwidth and buffer space**
    - ✓ **Scheduling mechanism with guarantee of minimum rate**
      - ➢ E.g. Deficit Round Robin
- **Guaranteed service**
  - **..**

Guaranteed service

- Provides applications with
  - Guaranteed bandwidth
  - Guaranteed specific maximum delay
  - No discards in queues
- It needs traffic contract and call admission control
  - Token bucket

$$delay_{queuing\_end2end} = \frac{(b-M)(p-R)}{R(p-r)} + \frac{(M+C_{tot})}{R} + D_{tot} \quad (p > R \geq r)$$

$$delay_{queueing\_end2end} = \frac{(M+C_{tot})}{R} + D_{tot} \quad (R \geq p \geq r)$$

GS invocation

- Up to know the information that travels is

Tspec
- p = peak rate
- b = bucket depth
- r = token bucket rate
- m = minimum policed unit
- M = maximum datagram size

Rspec
- R = bandwidth or service rate
- S = slack term

FLOW_SPEC

IntServ problems

- Complexity
- Scalability: "per-flow" reservation problems
  - Classify by flow
  - Resource use "by flow"
  - Signaling by flow
    - ✓ Soft-state: double-edged sword : dynamic but not so soft, a lot of signaling overhead
    - ✓ Compression (aggregate messages)
    - ✓ Why don't we remove all signaling?
- Constrained semantics for the service models
  - IntServ only "guaranteed" and "controlled load"
  - But perhaps… "Carry always class A before class B"
- Slow deployment
  - Poor coexistence with "no QoS" environments
    - ✓ In general it is desirable to avoid that situation where all routers must change in order for the scheme to work

General references

- Engineering Internet QoS. Sanjay Jha, Mahbub Hassan. Artech House 2002.
- Internet QoS. Architectures and mechanisms for Quality of Service, Zheng Wang, Morgan Kaufman, 2001.
- Computer Networking. A Top-Down Approach Featuring the Internet. James F. Kurose and Keith W. Ross. ISBN 0-201-61274-7, Addison-Wesley
- Internetworking Multimedia. Jon Crowcroft, Mark Handley, Ian Wakeman. UCL Press
- A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single-Node Case. Abhay K. Parekh, Robert G. Gallager. IEEE ACM Transactions on Networking, Vol. 1, Nº. 3, June 1993.
- A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Multiple Node Case. Abhay K. Parekh, Robert G. Gallager. lEEE ACM Transactions on Networking, Vol. 2, Nº. 2, April 1994.
- RSVP and Integrated Services in the Internet: A Tutorial. Paul P. White. IEEE Communications. May 1997.
- Inside the Internet's Resource Reservation Protocol. David Durham, Raj Yavatkar. John Wiley & sons. 1999
- RFC 2212. Specification of Guaranteed Quality of Service. S. Shenker, C. Partridge, R. Guerin. September 1997.

References on DiffServ

- DiffServ charter, http://www.ietf.org/html.charters/OLD/diffserv-charter.html
- "An Architecture for Differentiated Services", http://www.ietf.org/rfc/rfc2475.txt
- "New Terminology and Clarifications for DiffServ", http://www.ietf.org/rfc/rfc3260.txt
- "An Informal Management Model for Diffserv Routers", http://www.ietf.org/rfc/rfc3290.txt
- "An Expedited Forwarding PHB (Per-Hop Behavior)", http://www.ietf.org/rfc/rfc3246.txt
- "Assured Forwarding PHB Group", http://www.ietf.org/rfc/rfc2597.txt
- "Definition of Differentiated Services Per Domain Behaviors and Rules for their Specification", http://www.ietf.org/rfc/rfc3086.txt
- "Definition of the Differentiated Services Field (DS Field)in the IPv4 and IPv6 Headers", http://www.ietf.org/rfc/rfc2474.txt