Quality of Service in IP Networks

S. Hadjiefthymiades

University of Athens,

Department of Informatics & Telecommunications

Communication Networks Laboratory

Best Effort IP

Until recently, IP networks supported only one service class: best effort

The network would make its best attempt to deliver packets to their destination but with no guarantees and no special resources allocated for any of the packets.

Reasons for BE orientation

- The original TCP/IP suite was built on the idea of fair and equitable access to all and no special treatment of anyone.
- There weren’t any applications that needed QoS support.
- Early routers implemented simple queuing disciplines (FIFO).
- TCP tried to alleviate some of the problems of increased traffic and FIFO based routers.

Early mechanisms for Differentiation

First effort on IP QoS: written on 1992 by Clark, Shenker and Zhang.

Defined architecture, service classes and two mechanisms for QoS support: the Token bucket filter and Weighted Fair Queuing.

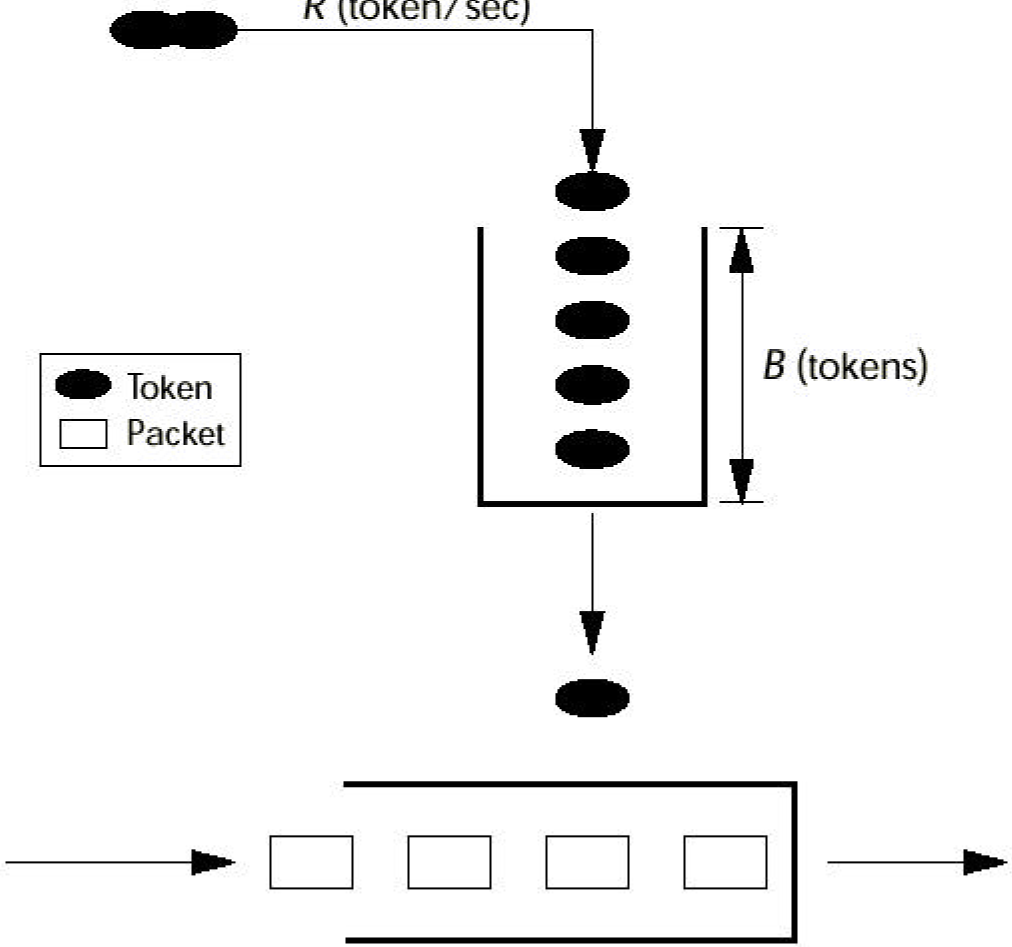
Token Bucket Filter

Characterisation of the application traffic load that receives a particular service.

Can be conceptualised as a bucket of tokens; depth B, rate of new token generation R.

Traffic source is said to conform to the parameters of a filter if it sends packets at a rate less than or equal to R.

Token Bucket Filter



Weighted Fair Queuing (WFQ)

WFQ algorithm used for scheduling packets for outbound transmission from routers.

Packets of a flow are time-stamped based on their arrival rate at the router, their scheduled departure time from the router and their length.

The departure queue of the WFQ scheduler is reordered every time a new packet arrives so the packets with smallest time stamps are txed first.

WFQ Characteristics

- **A particular flow is guaranteed its allocated share of the bandwidth irrespective of the behaviour of all other flows travelling through the same router.**

- **The WFQ algorithm is work conserving i.e. the link is never left idle.**

IETF IntServ Architecture

- IP architecture is extended to support both real-time and best effort traffic flows.
- Flow: a stream of packets with common source address, destination address and port number.
- The network maintains flow-specific state in order to provide the desired levels of service in terms of quantifiable bandwidth and delay.

Finite network resources

- Routers have finite buffers, CPU capacity and are linked to links with a certain maximum bandwidth.
- Hence, routers need to exercise control over what flows would be allocated what resources.
- For the 1st time, some router within the IP network would deny service provision.

Basic Components of IntServ

- Traffic Control: Admission Control + Packet Classifier + Packet Scheduler
- Traffic Classes: Guaranteed & Controlled
- Setup Protocol: RSVP

Traffic Control

- **Admission Control: checks to see if the resources in the router can support a particular service. Performs accept/reject decision.**

- **Packet Classifier: Examines the source/destination address and port fields in the processed packet.**

- **Packet Scheduler: Schedules the packet for outbound transmission. Applies algorithms like WFQ.**

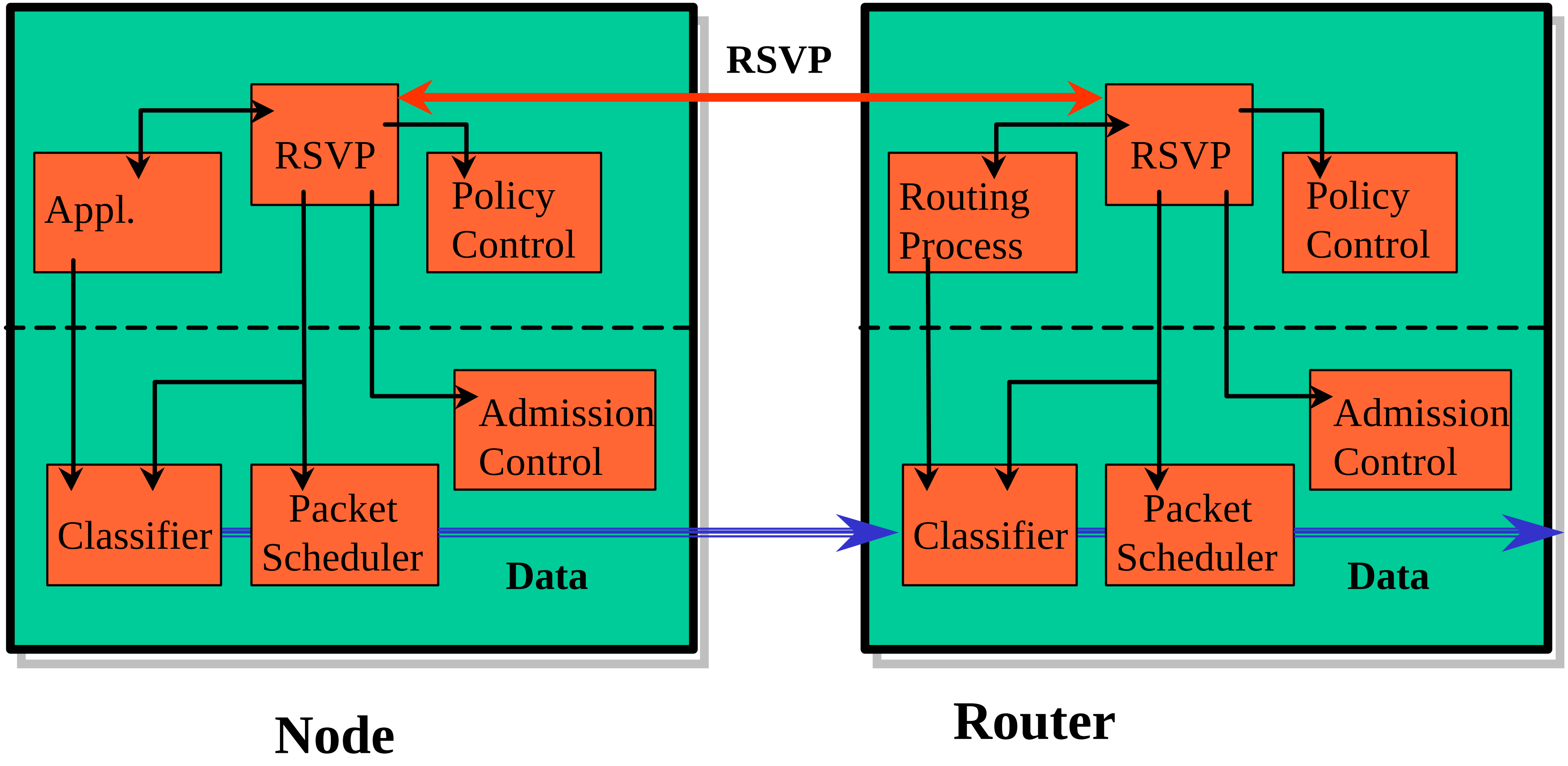
Traffic Classes

- Controlled Load: approximates a best effort service over an non-congested network.
- Guaranteed Load: supports real-time flows that require a quantifiable bound on delay.

IntServ FlowSpec

- **The flow specification is carried by RSVP messages into the network and defines and application’s QoS requirements as a series of objects (e.g., token bucket parameters).**

IntServ Architecture



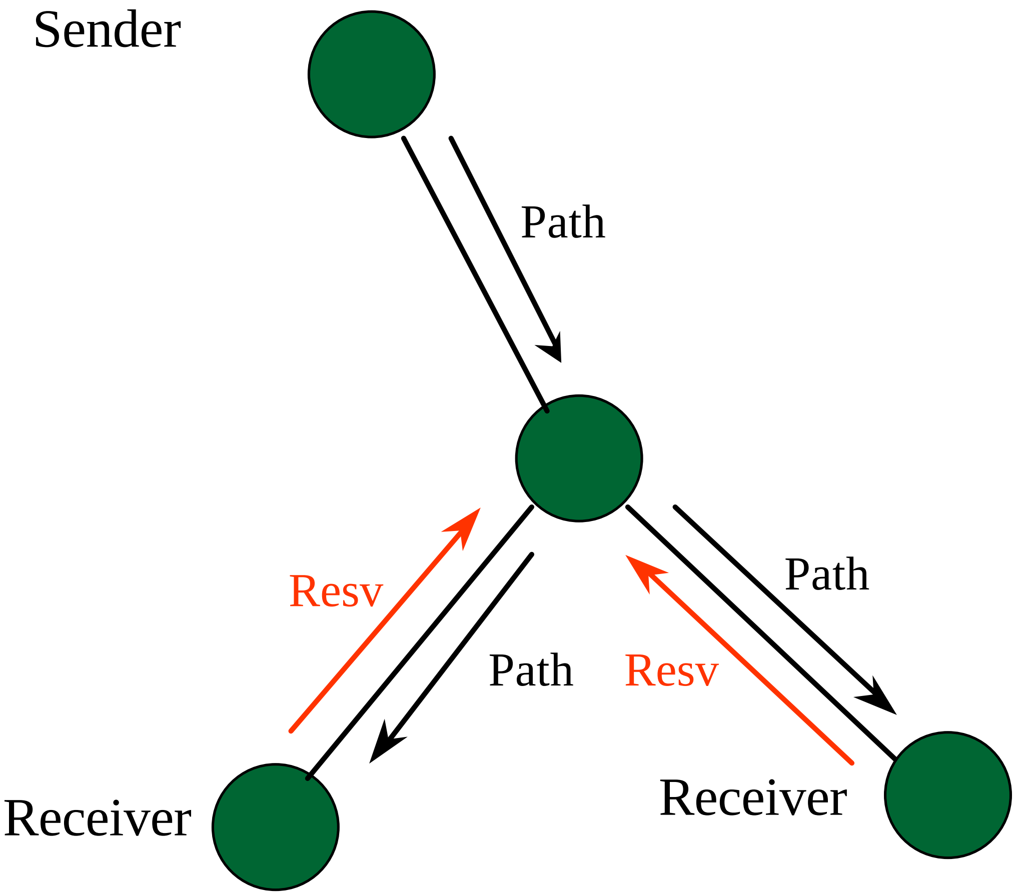
Resource Reservation Protocol

- The defacto setup protocol for IntServ is Resource Reservation Protocol (RSVP).
- The source transmits a PATH message along the routed path to the destination (unicast or multicast).
- PATH: marks the routed path across the network and collects information about the QoS viability of each router.

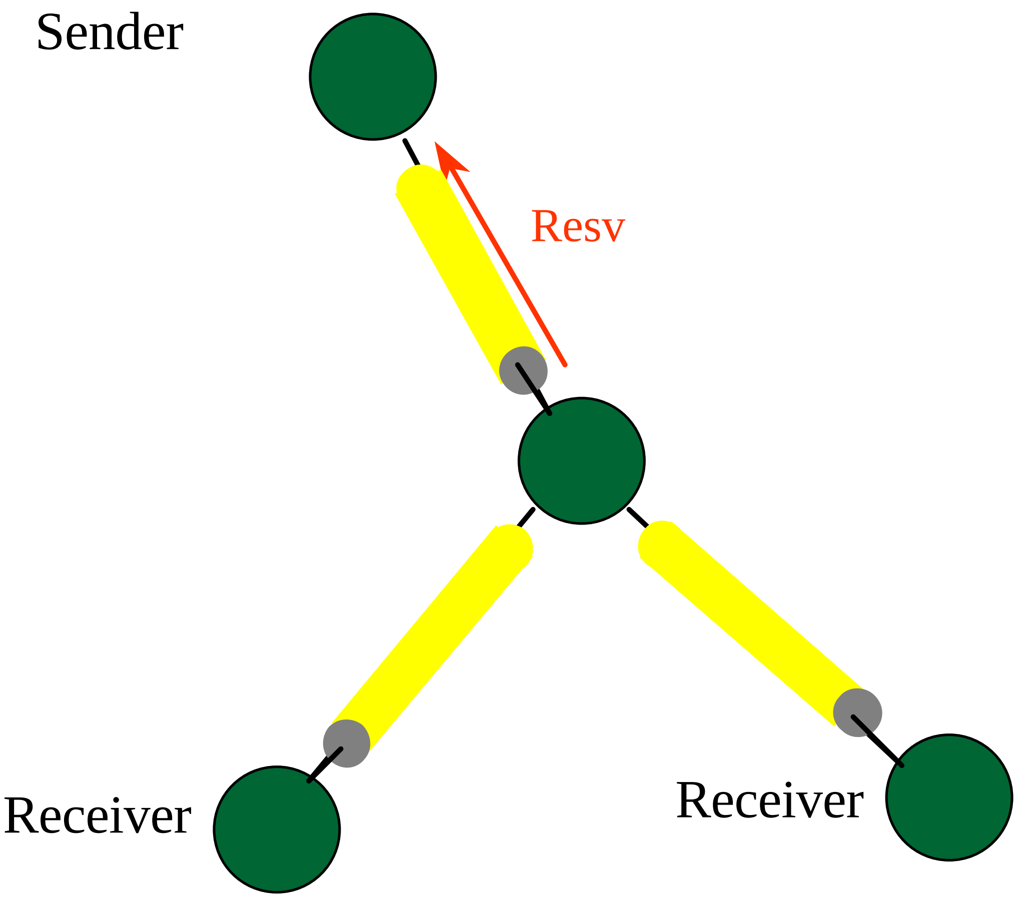
Resource Reservation Protocol

- **Protocol is unidirectional. Establishes reservation state in one direction.**
- **Receiver initiated reservations.**
- **Independent from current unicast/multicast routing protocols.**

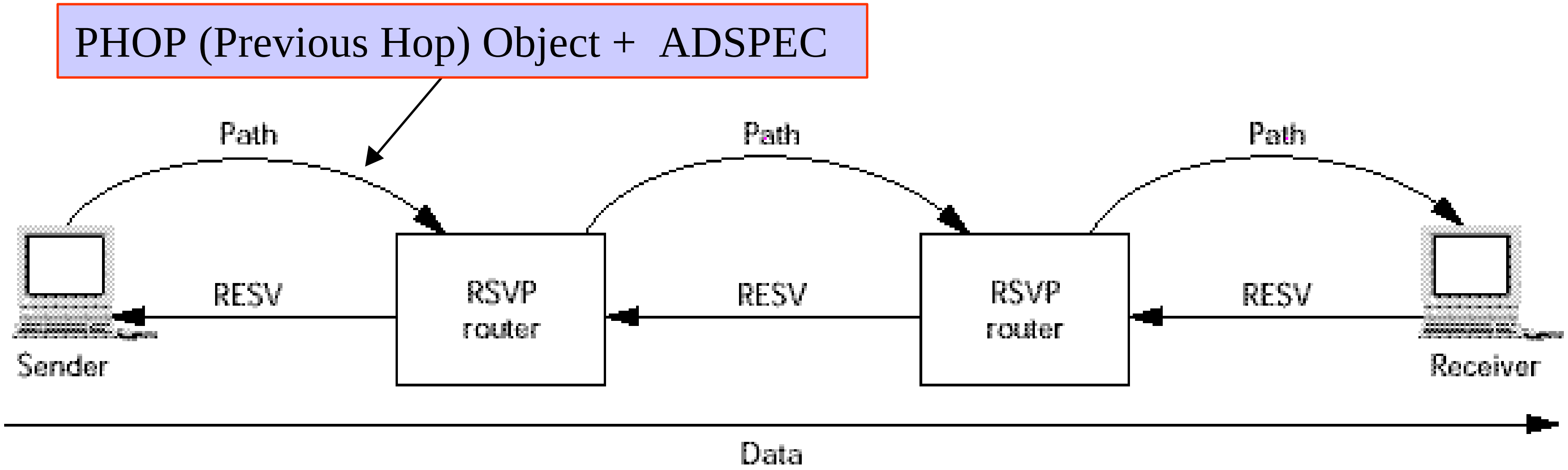
RSVP basic message flow



RSVP basic message flow



RSVP message flow



• PHOP allows the RESV messages to find their route upstream.

• ADSPEC: summarises the path’s characteristics and delivers this info to the receiver. <u>OPWA: One Pass With Advertisement</u>

RESV message

- Upon receiving the PATH message, the receiver can gauge what services the network can support and then generate an RESV message towards the sender.
- RESV contains traffic and QoS objects that are processed by the traffic control component of each router in the upstream direction.

RSVP Soft-state

- The per flow reservation state maintained in routers will be deleted unless RSVP PATH and RESV messages are periodically sent by the sender and receivers respectively.
- Soft-state orientation of the protocol.

Reservation Styles (1/2)

- To make best use of reserved state, RSVP supports different reservation styles.

- Styles specify whether a reservation should be dedicated to a particular sender or shared among multiple senders in the same session (RSVP session = same destination IP + port number).

Reservation Styles (2/2)

- Fixed Filter (FF) reservation: a single reservation will be allocated for a single sender.
- Shared Explicit (SE) reservation: multiple senders in the same session are allowed to share a single reservation.

Merging of reservations

- RSVP reservations are allowed to merge at branch points in a multicast tree.
- The router will install the largest of all received reservation requests for a sender or set of senders and then pass that merged reservation value to the next, upstream router.

IntServ/RSVP Open Issues

- The applicability and scalability of RSVP in large networks is limited.
- Example: core router in ISP supporting 10.000 VoIP flows established using RSVP. The router needs to maintain state for 20.000 flows (10.000 per direction) while processing refresh messages.
- RSVP lacks adequate security mechanisms.

RSVP Compliant API

- **Winsock2 is a Windows based networking API that supports features such as multicast and QoS.**
- **Winsock2 defines data structures and calls that enable applications to signal their requirements through a variety of techniques including RSVP and ATM.**
- **Generic QoS API by Microsoft.**

IETF RSVP Survey (1997)

? atas?e?ast??	**? ??t???**
3COM	CoreBuilder (?at?stas?: alpha prototype) ?a? NetBuilder II (?at?stas?: development)
Bay Networks	Backbone Node, Access Stack Node, Advanced Remote Node, Access Node (?at?stas?: beta prototype ??a ??a ta µ??t??a)
IBM	QoS-Switch-Router (QSR) (?at?stas?: Beta testing. Ta ???s?µ?p????e?? ? e?e???t??? p??t?t?p?).
CISCO	IOS 11.2. ? p?st?????eta?se ??e? t?? p?atf ??µe? 1xxx, 25xx, 4xxx, and 7xxx. ? ????? ap? t?? ? ????st? 1996.
FORE	RSVP Prototype (?at?stas?: alpha prototype)
Furukawa Electric	INFONET3740/3780/3791 (?at?stas?: beta prototype)
IBM	IBM Multi-protocol Router Family (Nways Router), p?? pe???aµß??e?ta µ??t??a 2210, 2216, and 8210 (? at?stas?: under testing).

Differentiated Services

- New IETF approach for supporting IP QoS.
- Individual micro-flows are classified at the edge of the network into one of several unique service classes and then a per-class service is applied in the middle of the network.

Packet Classification

- Packet classification is performed at the ingress of the network based on the analysis of one or more fields in the packet.
- Packet is then marked as belonging to a particular service and, then, injected into the network.

Reasons behind DS

- Scalability; individual host-2-host microflows are aggregated into a single aggregate flow and then this flow receives special treatment.
- Solution should be applicable to all applications. No new control protocol and APIs should be needed.
- New router technologies with constantly increasing speeds.

DS Architecture Components

- DS-field is a bit pattern contained in the header of each packet denoting the service (PHB) the packet should receive at each hop. The TOS/IPv4 and the TrafficClass/IPv6 fields have been redefined as the DS-fields.
- Per Hop Behaviour: the service that packets receive at each hop. PHB can be declared in a relative or absolute way.

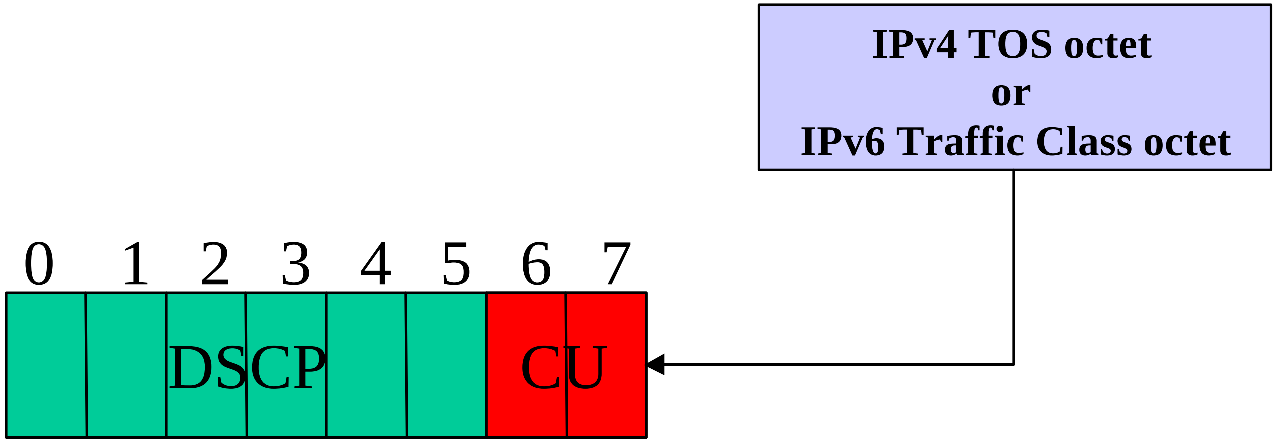
DS Architecture Components

- Behaviour Aggregate (BA): a group of packets with the same DS Code Point (DSCP).

- A boundary router is positioned at the edge of a DiffServ capable network. This device performs packet classification, metering, packet marking and possibly policing or shaping.

- Interior nodes: provide the PHB based on the DSCP bits. Use algorithms such as WFQ or RED for queue management and scheduling.

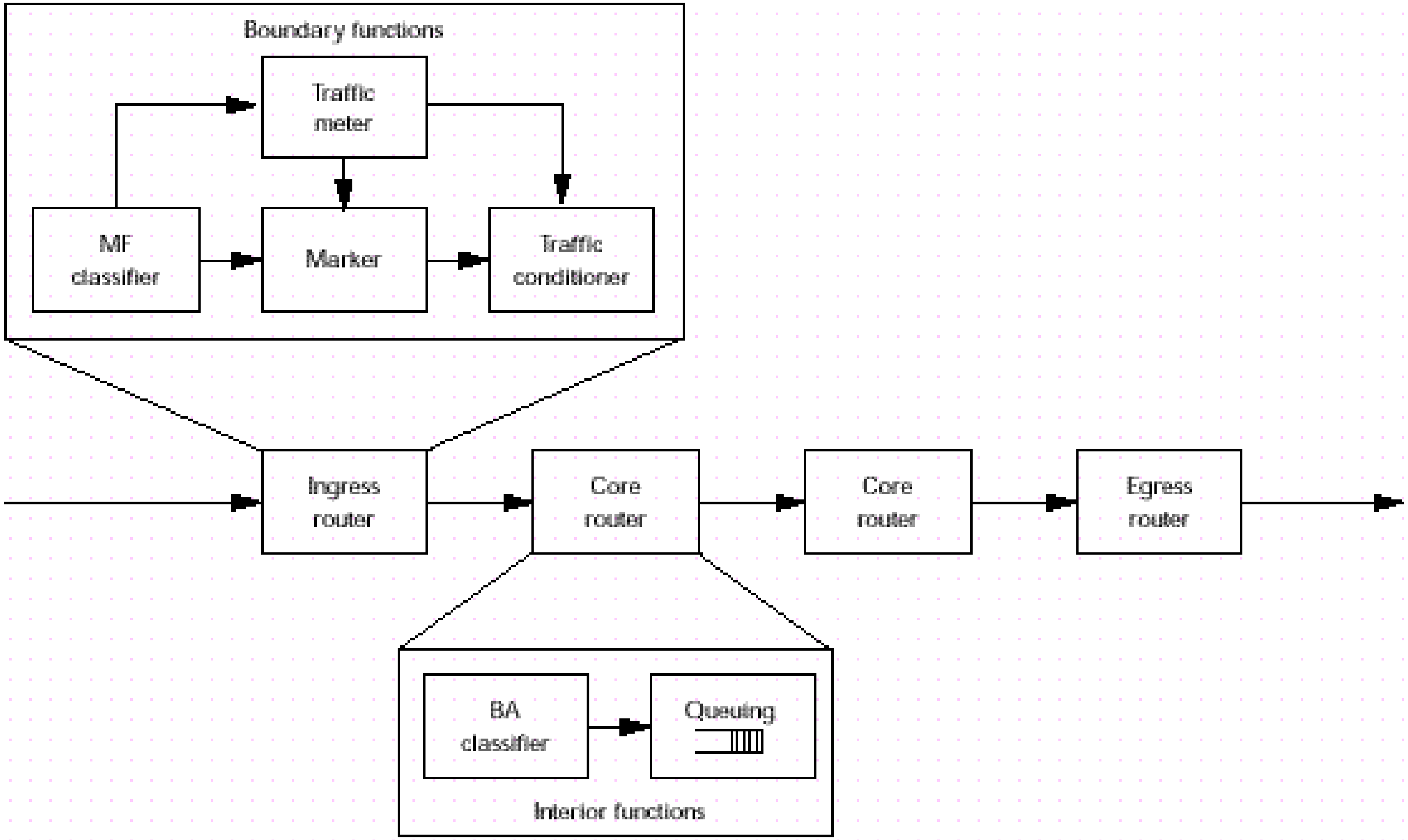
DS Field



DSCP:Differentiated Services CodePoint

CU:Currently Unused

DS Architecture Components



Packet processing in ingress router

- **Packet passes through a MultiField (MF) classifier which works with a traffic meter to measure packet’s conformance to with agreed traffic profiles. In-profile and out-of-profile packets may be treated differently.**
- **The DSCP bits are then marked.**
- **The packet may then be conditioned (shaped or dropped) before entering the network.**

Per-Hop Behaviours

- Expedited Forwarding (EF): low loss, low delay and low jitter connections. It appears as a point-2-point Virtual Leased Line (VLL) with a peak bandwidth.

- To minimise jitter and delay, packets must spent little or no time in router queues. Therefore the EF PHB requires that the traffic be conditioned to conform to the peak rate at the ingress router.

- A single DSCP bit is used to indicate EF PHB.

Per-Hop Behaviours

- **Assured Forwarding (AF): defines four relative classes of service with each service supporting 3 levels of drop precedence. 12 DSCP bit combinations define the AF classes and the drop precedence levels.**
- **When congestion occurs, packets with a higher drop precedence will be discarded ahead of those with a lower drop precedence.**