5. QoS Functions in Core and Backbone Networks

Dr. David Soldani

(david.soldani@nokia.com, tel. +358.50.3633527)

S-38.3215 Special Course on Networking Technology for Ph.D. students at TKK

Outline

- IP QoS
- Circuit-switched core QoS
- Packet-switched core QoS
- Backbone QoS

IETF: Definition (1/5)

- **Flow** - Sequence of packets that are sent from a particular source to a particular (unicast or multicast) destination and that are related in terms of their routing and any particular set of fields from the packet header used to identify the flow
- **Microflow** - A single instance of an application-to-application flow of packets which is identified by source address, source port, destination address, destination port and protocol id
- **Link** - A single link-level connection between two (or more) hosts; includes leased lines, ethernets, frame relay clouds, etc.

IETF: Definition (2/5)

- **Differentiated Services (DiffServ)** - IP header field, called the **DS-field**. In IPv4, it defines the layout of the ToS (**Type of Service**) octet; in IPv6, it is the **Traffic Class** octet. *Differentiated Services* is also an "area of use" for QoS policies (correspondence between code-points in the packet's DS-field and individual per-hop behaviours to achieve a specified per-domain behaviour)
- **Integrated Services (IntServ)** - The integrated services architecture assumes that some explicit setup mechanism is used to convey information to routers so that they can provide requested services to flows that require them

IETF: Definition (3/5)

- **DS behavior aggregate** - A collection of packets with the same DS code-point crossing a link in a particular direction
- **DS code-point (DSCP)** – A specific value of the DSCP portion of the DS field, used to select a PHB
- **Per-Hop-Behavior (PHB)** - The externally observable forwarding behavior applied at a DS-compliant node to a DS behavior aggregate

IETF: Definition (4/5)

- **Service Level Agreement (SLA)** - The documented result of a negotiation between a customer/consumer and a provider of a service, that specifies the levels of availability, serviceability, performance, operation or other attributes of the service
- **Service Level Objective (SLO)** - Partitions an SLA into individual metrics and operational information to enforce and/or monitor the SLA
- **Service Level Specification (SLS)** - An SLS is a specific SLA (a negotiated agreement) and its SLOs (the individual metrics and operational data to enforce) to guarantee quality of service for network traffic.
- **Service Provisioning Policy** - A policy that defines how traffic conditioners are configured on DS boundary nodes and how traffic streams are mapped to DS behavior aggregates to achieve a range of services

IETF: Definition (5/5)

- **Traffic Conditioning Agreement (TCA)** - An agreement specifying classifier rules and any corresponding traffic profiles and metering, marking, discarding and/or shaping rules which are to apply to the traffic streams selected by the classifier
- **Traffic Conditioning Blocks (TCBs)** - A generalized TCB might consist of the following stages: Classification stage; Metering stage, Action stage (involving Markers, Absolute Droppers, Counters, and Multiplexors); Queuing stage (involving Algorithmic Droppers, Queues, and Schedulers)

IETF: Definition of QoS

- Quality of Service refers to the “classification of packets for the purpose of **treating certain classes or flows of packets in a particular way compared to other packets**”
- Ideally, it makes the data delivery service of otherwise unpredictable best effort Internet Protocol (IP) networks, predictable
- **QoS protocols provide the mechanics to differentiate traffic, and policy defines how they are used**

IP QoS: Integrated Services (IntServ)

- Attempts to provide per-flow (defined by the 5-tuple) QoS assurances with dynamic resource reservation and RSVP signalling
- Policy control for individual flows, and regulate their ability to reserve network resources
- Extended service model targeted towards RT traffic
  - Guaranteed service
  - Predictive service (whose quality is sufficiently predictable)
  - ⇒ Resource reservation and CAC capability assumed

[RFC1633]

IP QoS: Differentiated Services (DiffServ)

- DiffServ are aimed at traffic aggregates that may not correspond to fine grained flows
- DiffServ relies on administrative control of bandwidth, delay or dropping preferences, rather than per flow signalling, to communicate service level information to network elements
- For such services IETF enables flexible definition of class-based packet handling behaviours and class based policy control

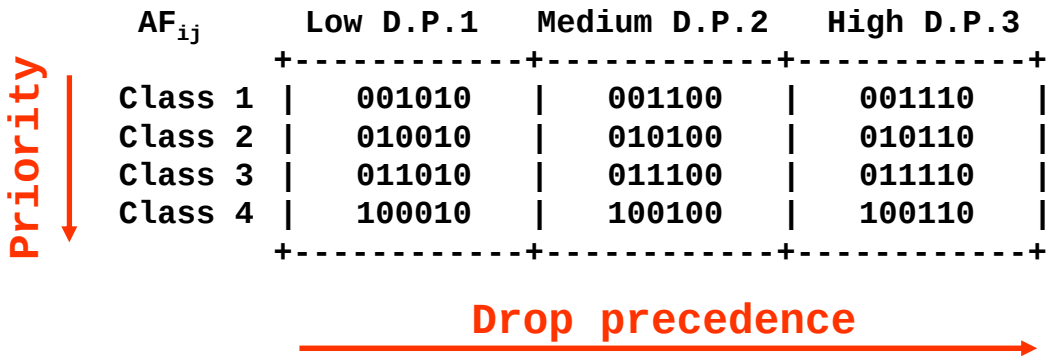
[RFC2475]

Expedited Forwarding [RFC3246]

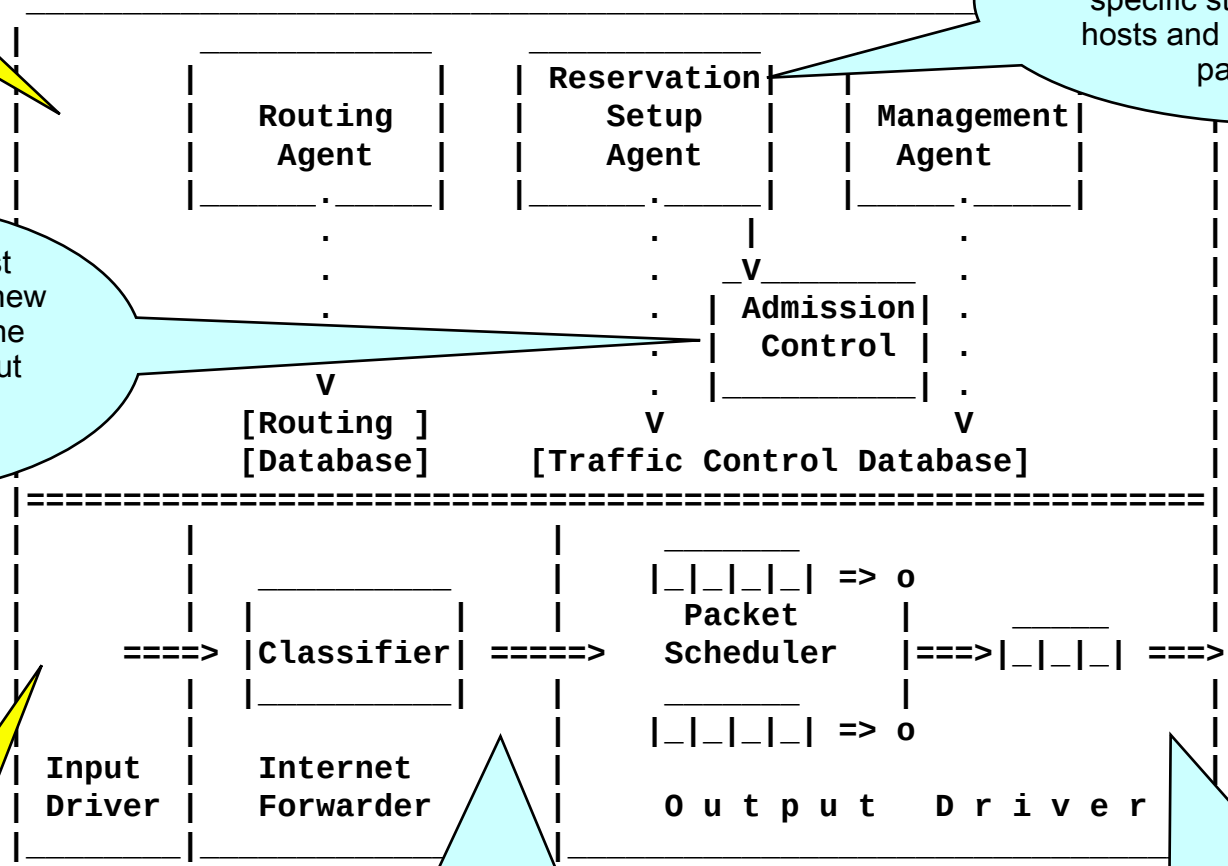
- The rate at which EF traffic is served at a given output interface should be at least the configured rate *R*, over a suitably defined interval, independent of the offered load of non-EF traffic to that interface
- The intent of the EF PHB is to provide a building block for low loss, low delay, and low jitter (variation between maximum and minimum delay) services
- **Note**: the EF PHB only defines the behavior of a single node, the of behavior of a collection of nodes may be provided by a Per-Domain Behavior (PDB) specification

Assured Forwarding [RFC2597]

- AF PHB group provides forwarding of IP packets in ***N*** independent AF classes, and within each AF class, an IP packet is assigned one of ***M*** different levels of Drop Precedence
- Currently, four classes (*N* = 4) with three levels of drop precedence in each class (*M* = 3) are defined for general use, more AF classes or levels of drop precedence MAY be defined for local use
- IP packets are assigned into one or more AF classes according to the services that the customer has subscribed to, and in case of congestion, the drop precedence of a packet determines the relative importance of the packet within the AF class



IntServ Architecture for Routers [RFC1633]



Background code

A reservation setup protocol creates and maintain flow-specific state in the endpoint hosts and in routers along the path of a flow

AC in a router or host determines whether a new flow can be granted the requested QoS without impacting earlier guarantees

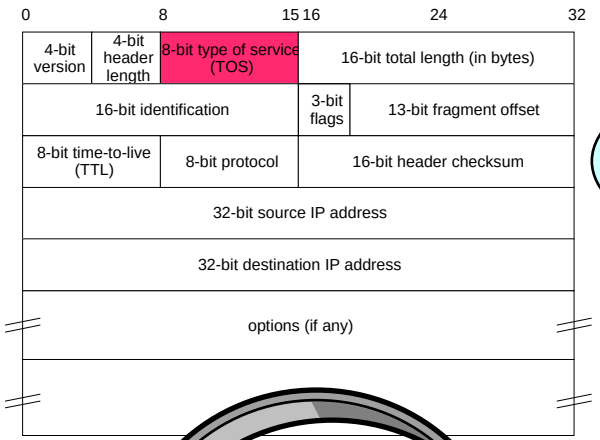
Note: an application must specify the desired QoS carried by the reservation setup protocol, and ultimately used to parameterise the packet scheduling mechanism

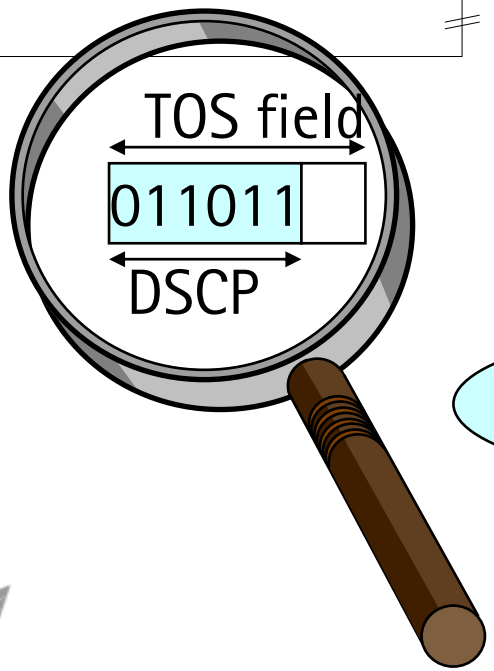
Forwarding path (for every packet)

Mapping of packets into some treatment class e.g. based upon contents of existing packet header(s)

The packet scheduler manages the forwarding of different packet streams using a set of queues and other mechanisms like timers

DiffServ Architecture for Routers [REC2475]

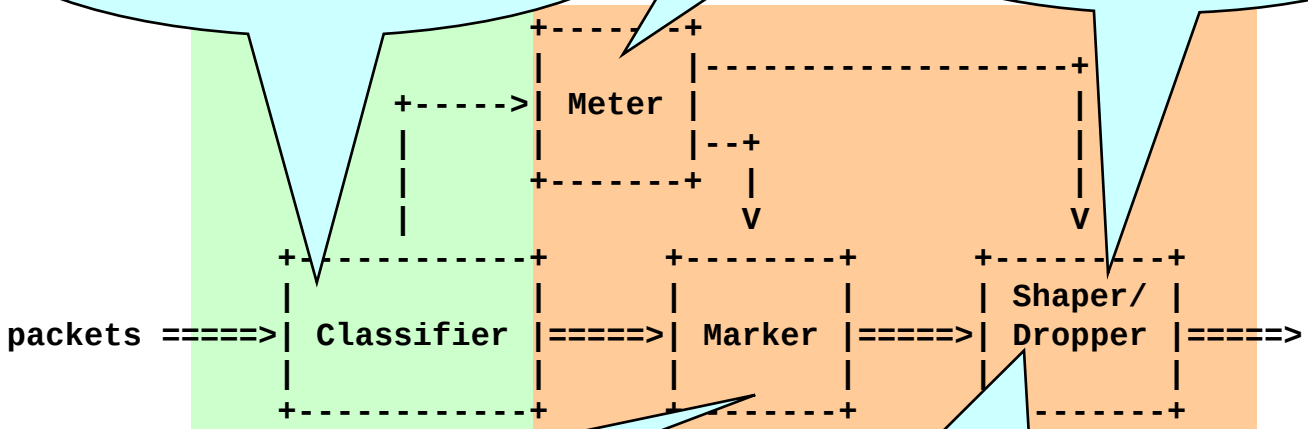




Measures the traffic stream against a traffic profile (e.g. all packets marked with DS codepoint *X* measured against a **token bucket meter** with rate *r* and burst size *b*) specified in a Traffic Conditioning Agreement (TCA), and affects marking, dropping, or shaping actions

Behavior Aggregate (BA): classification based on the DSCP only
Multi-Field (MF): Based on values of header fields, e.g. destination address, DS field, protocol ID, source port and destination port numbers

Shapers delay some or all of the packets in a traffic stream in order to bring the stream into compliance with a traffic profile



Sets the DS field of a packet to a particular codepoint, adding the marked packet to a particular DS behaviour aggregate

Discards some or all of the packets in a traffic stream in order to bring the stream into compliance with a traffic profile

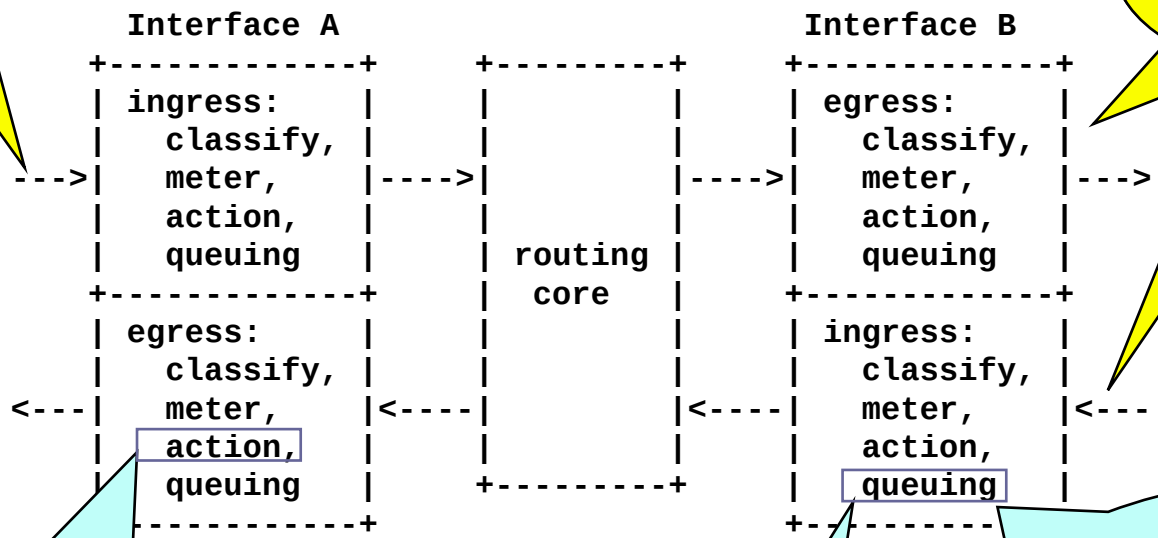
Classifier

Traffic conditioner*
* May not not necessarily contain all four elements

DiffServ Architecture for Routers [RFC3290]

Diffserv Functions at Ingress and Egress interfaces

Traffic Conditioning and Queuing Elements

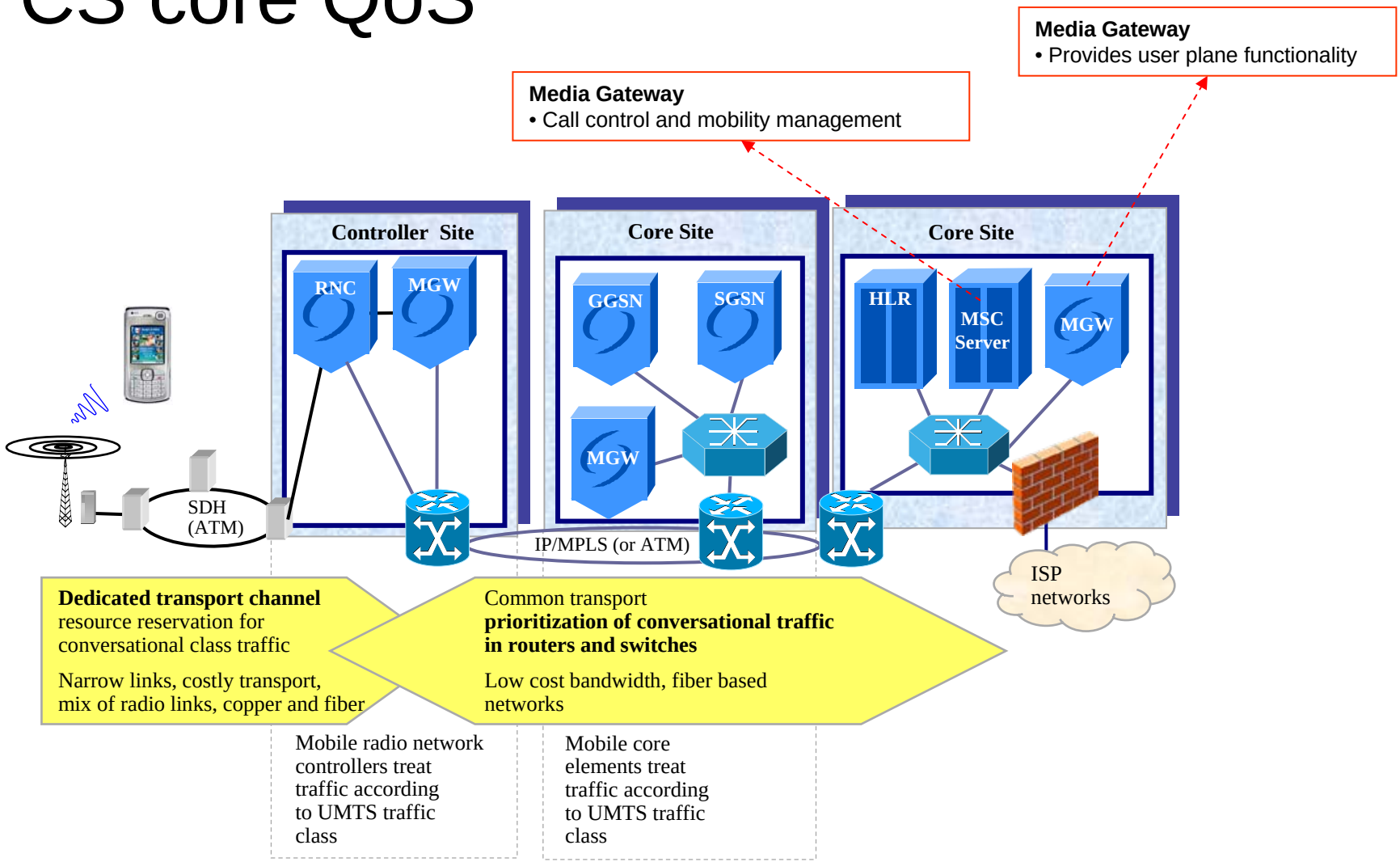


Action component may consist of **shaping**, **dropping** (policing, e.g. **Weighted Random Early Detection** (single queue), **DiffServ Urgency/Importance model** (multiple queues)), **multiplexing** and **marking** or **remarking** of packets

Queues for: **E**xpedited **F**orwarding, four **A**ssured **F**orwarding classes, and **B**est **E**ffort PHB

- Two possibilities: **Multi-stage and/or Single-stage scheduler**
- Possible scheduling methods: **Strict Priority (SP) queuing**, **Class-Based Weighted Fair Queuing (CB-WFQ)**, **Best-Effort queuing**.
- The number of queues needed are typically 3-8 per interface

R4: CS core QoS



PS: Session Management (SM)

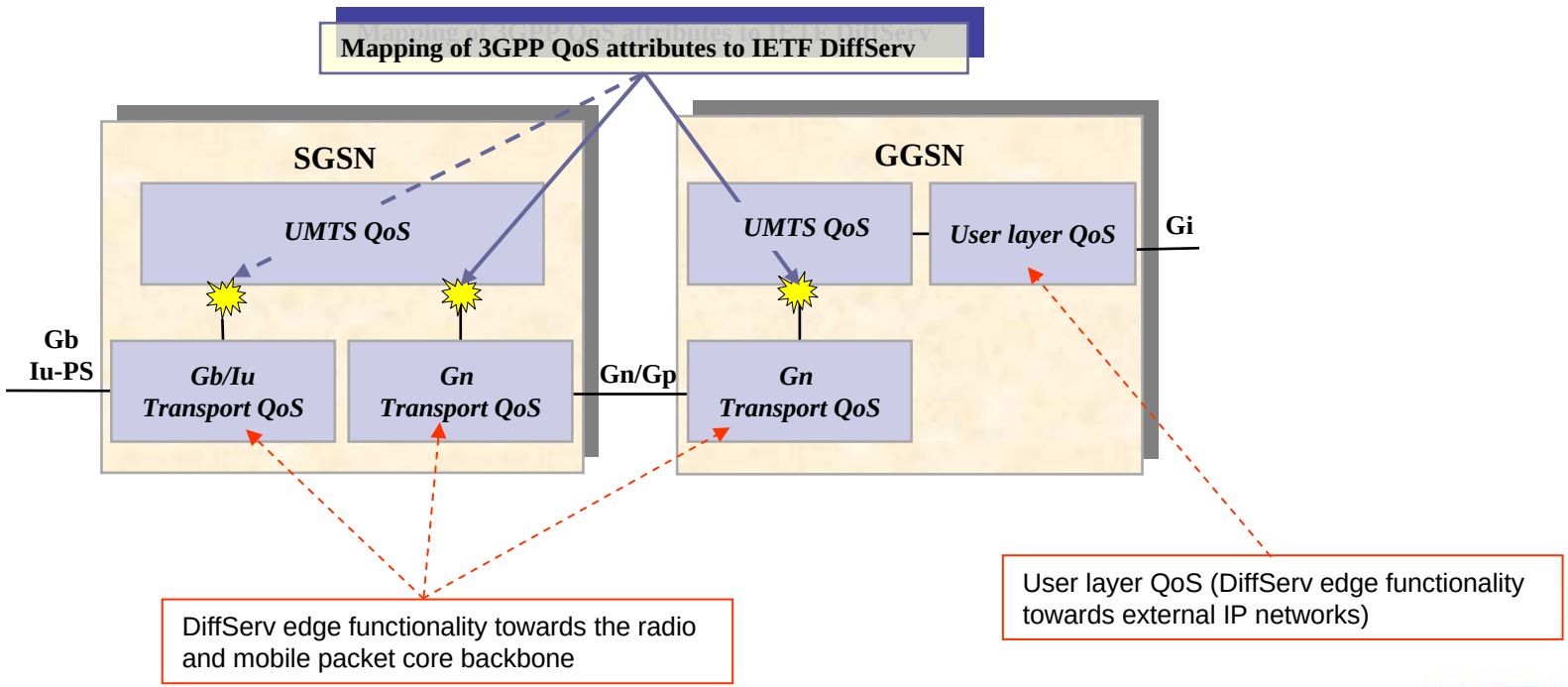
- SM: SGSN, HLR and GGSN
- SGSN: PDP context parameter selection
- Mapping of R97/98 QoS onto R99 and vice versa
- RT PDP context Admission Control (AC)
  - Configurable bandwidth for all RT PDP contexts in all NEs
  - AC based on NE utilization, e.g. central processing unit (CPU) load percentage

PS: Intelligent Edge concept

- Improve QoS and charging control based on actual services being used
- GGSN identifies which of the subscriber's services is in use by looking inside the **IP flow using a Layer 4/7 lookup mechanism**, and selects adequate QoS profile accordingly
- If there are several active IP flows associated with one PDP context, the **QoS profile suitable for the most demanding flow should be selected**
- In 3GPP R6, it is possible to modify the **PDP context depending** on access network capabilities based **on the RAT field** in the PDP context activation and PDP context update messages between the SGSN and GGSN

PS: Traffic Management

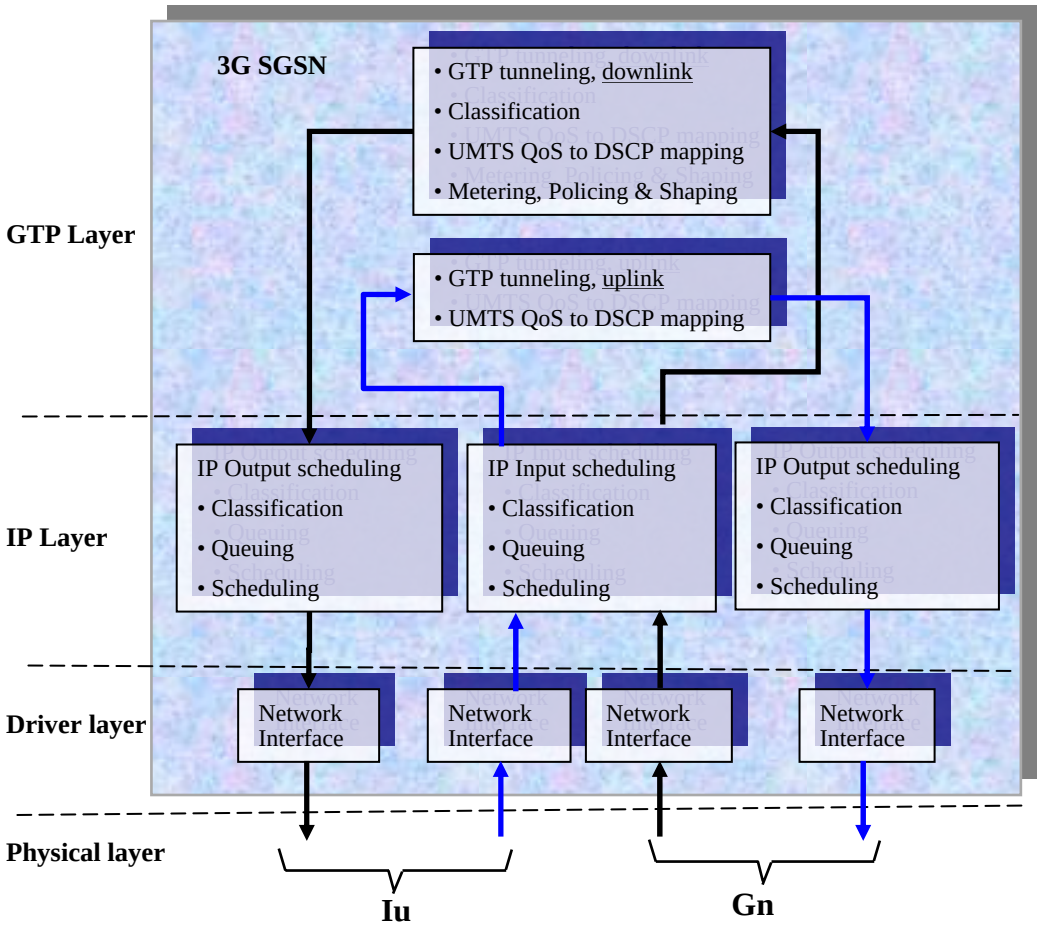
- Packet **classification** and **marking**, **queuing**, **scheduling** and **congestion avoidance** mechanisms
- SGSN and GGSN mark Diffserv code point (DSCP) field of the transport IP header according to the PDP context type



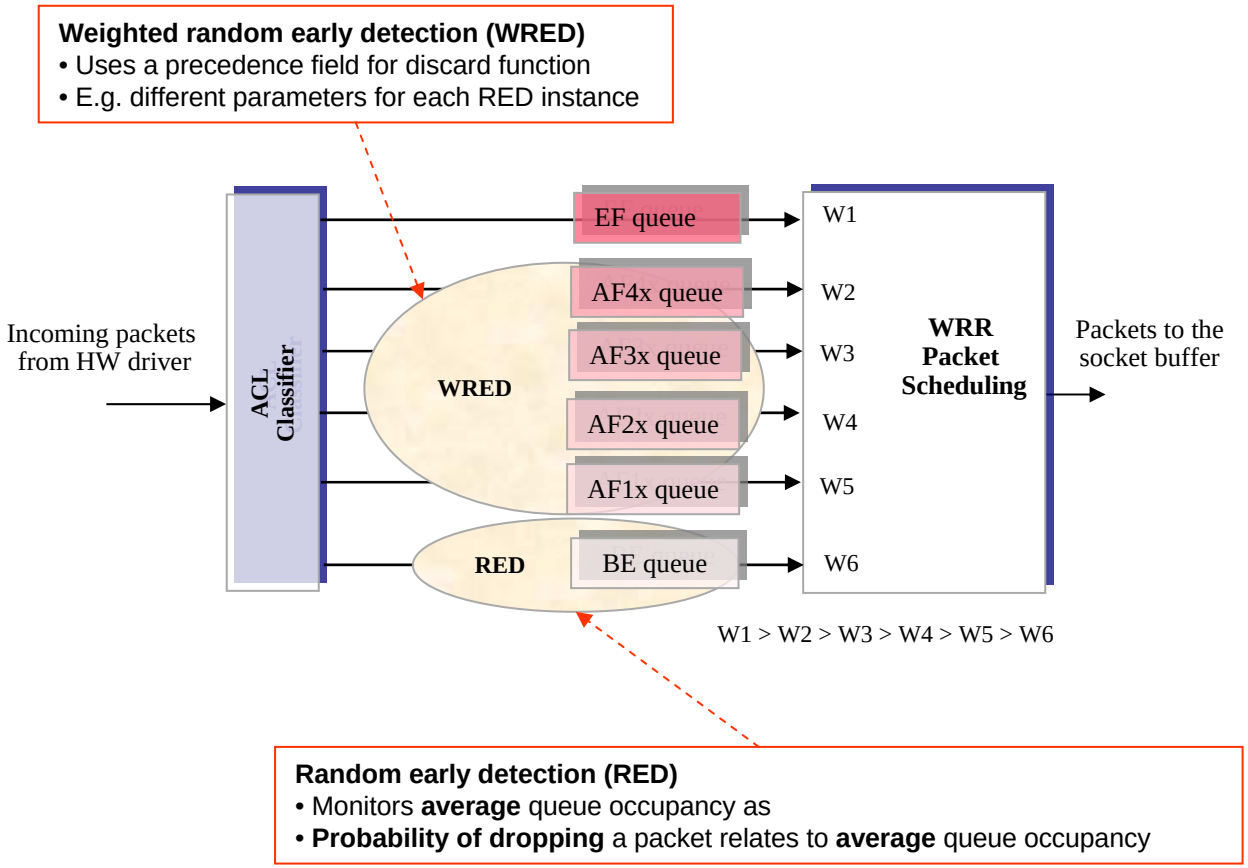
Mapping between 3GPP QoS and DSCP field

Classifier			Action	
Traffic Class	**THP**	**ARP**	**PHB**	**DSCP**
Conversational	-	ARP1	EF	101110
Conversational	-	ARP2	EF	101110
Conversational	-	ARP3	EF	101110
Streaming	-	ARP1	AF41	100010
Streaming	-	ARP2	AF42	100100
Streaming	-	ARP3	AF43	100110
Interactive	THP1	ARP1	AF31	011010
Interactive	THP1	ARP2	AF32	011100
Interactive	THP1	ARP3	AF33	011110
Interactive	THP2	ARP1	AF21	010010
Interactive	THP2	ARP2	AF22	010100
Interactive	THP2	ARP3	AF23	010110
Interactive	THP3	ARP1	AF11	001010
Interactive	THP3	ARP2	AF12	001100
Interactive	THP3	ARP3	AF13	001110
Background	-	ARP1	BE	000000
Background	-	ARP2	BE	000000
Background	-	ARP3	BE	000000

3G SGSN traffic management



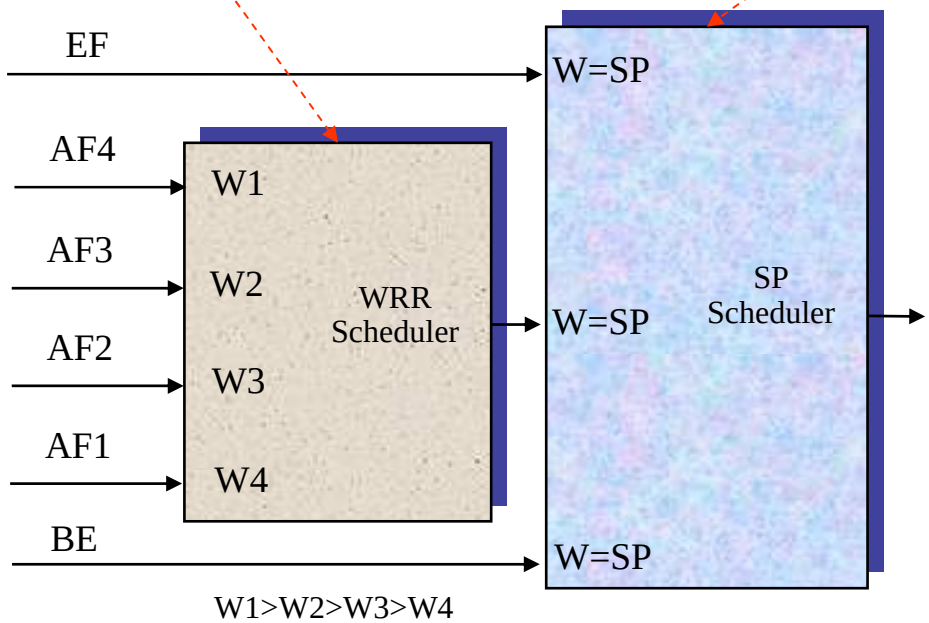
IP Congestion Avoidance mechanisms



IP Packet Scheduling mechanisms

Round Robin (RR) scheduler
Avoids local queue starvation by cycling through the queues one after the other, transmitting one packet before moving on to the next queue

Strict Priority (SP) scheduler
Orders queues by descending priority and serves a queue of a given priority level only if all higher priority queues are empty

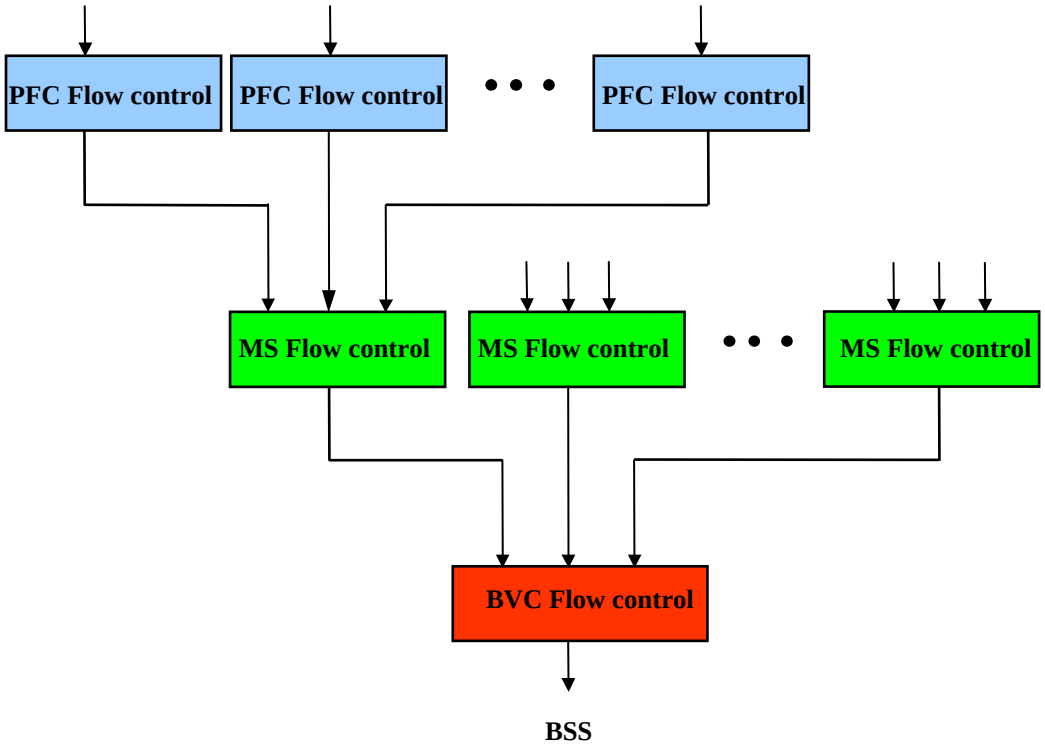


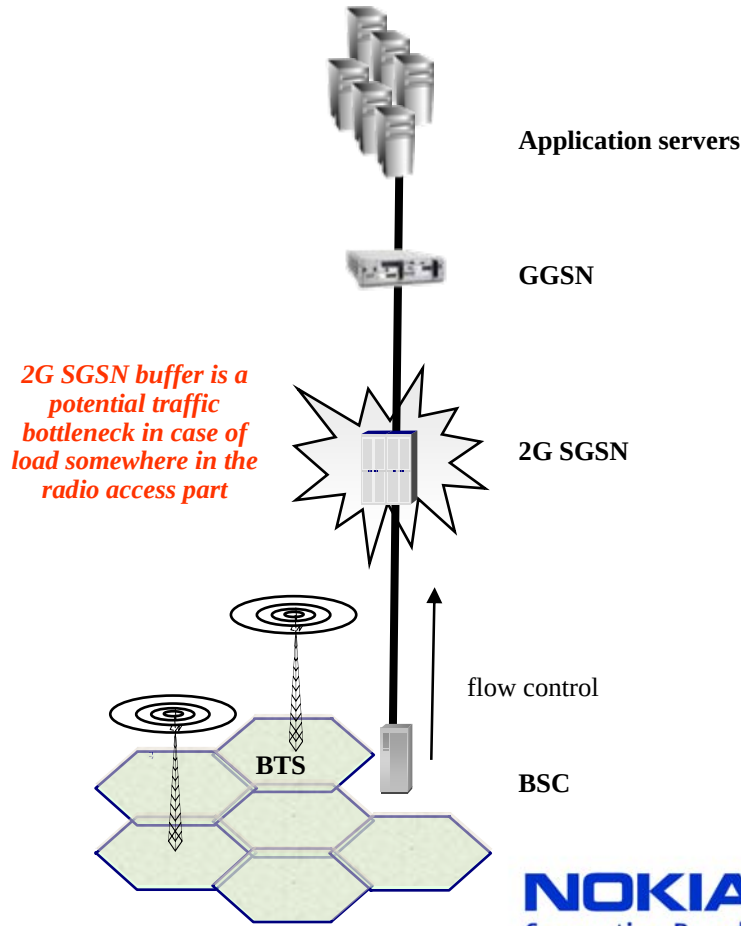
GGSN traffic management

- Scheduling, queuing and prioritization of IP traffic are typically done in a similar way to 3G SGSN
- **Metering and policing** functionalities for downlink
  - **Metering function:** ensures that downlink traffic conforms to the negotiated bit rate at the PDP context level
  - **Traffic conditioner (shaper/dropper) function:** provides conformance of downlink user data traffic
- **Marking IP header DSCP field** according to the PDP context QoS profile, **also for uplink traffic** the DSCP field can be marked in order to enable consistent traffic differentiation behind the Gi interface

2G SGSN: flow control levels

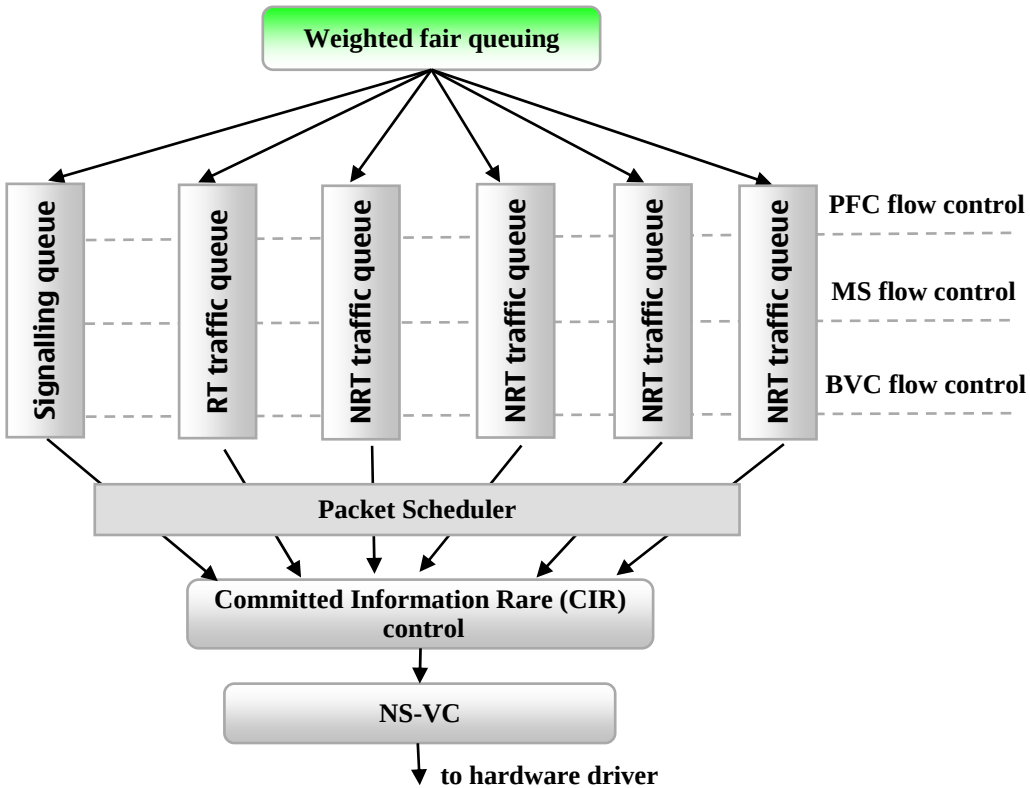
- Performed on each LLC-PDU first by the PFC flow control mechanism (if applicable and negotiated), then by the MS flow control mechanism and last by the BVC (cell) flow control mechanism





2G SGSN: traffic prioritization

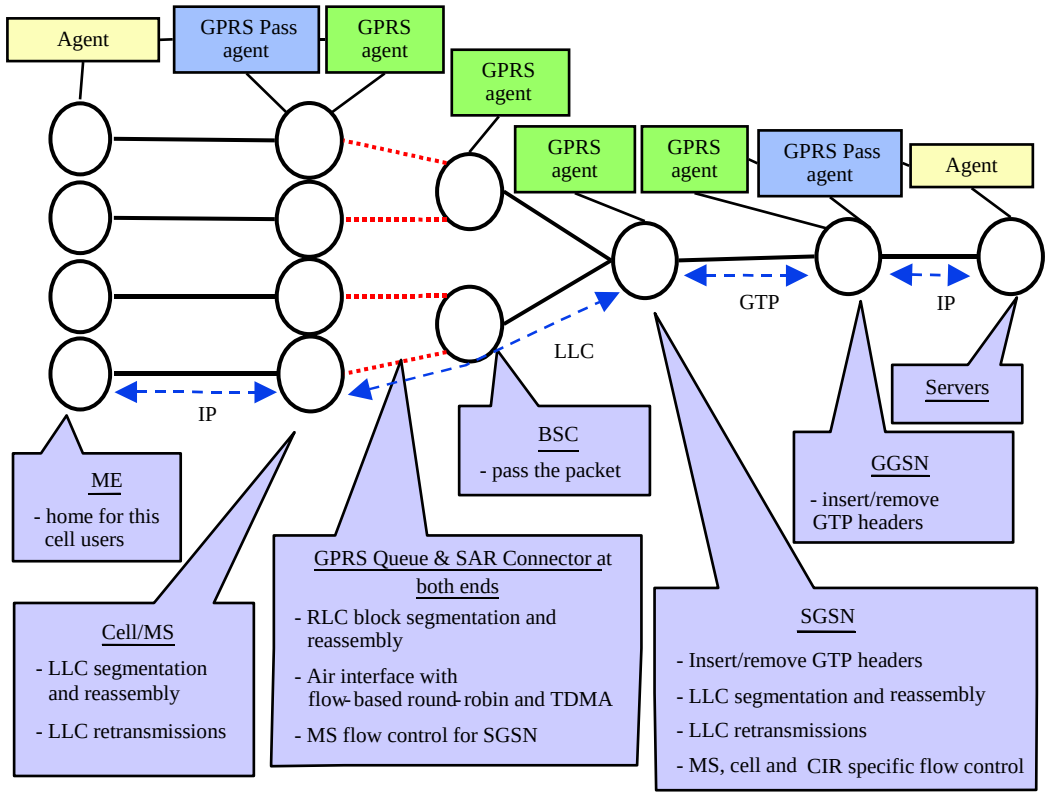
- LLC packets belonging to different TCs may be handled in separate buffers



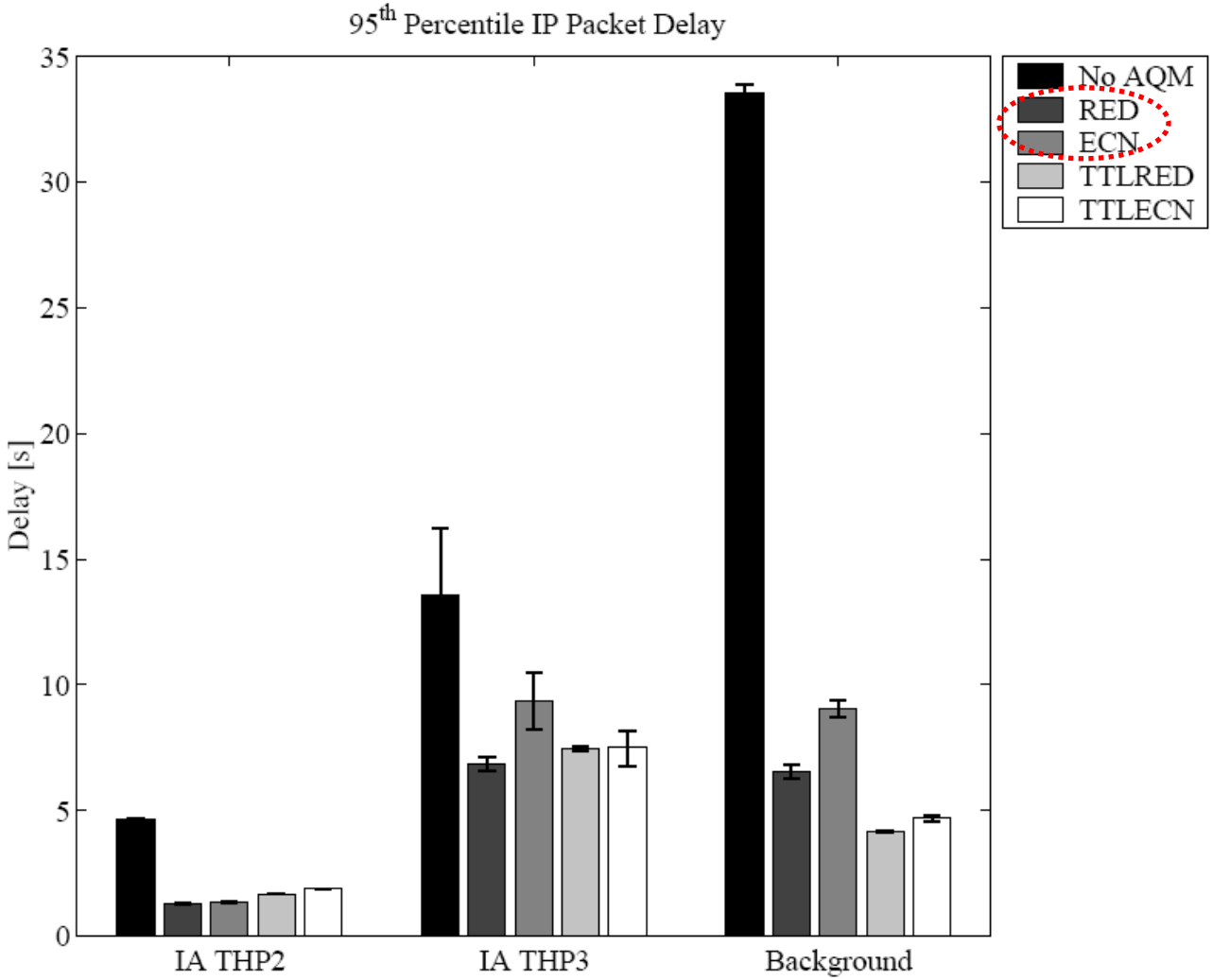
2G SGSN: buffer delay control

- **Pre-defined lifetime for LLC frames**: after having spent a certain pre-defined time in the 2G-SGSN and/or BSC buffers, the LLC frame is discarded
- **Limit the buffer size** (extracting a maximum buffer delay out of the 2G-SGSN buffer size is not an easy task)
- **Adopting RED** algorithm or **'explicit congestion notification' (ECN)**, which allows a TCP receiver to inform the sender of congestion in the network upon receiving an **IP packet marked with congestion experienced (CE) bit(s)**; TCP sender will then reduce its congestion window
- **TTL-based RED/ECN approach** since it is not straightforward to relate 2G-SGSN buffer occupancy and buffer delay
- **Window pacing**: **decrease the TCP-advertised window value** in uplink TCP acknowledgements if the defined buffer filling level threshold for a specific TC is reached

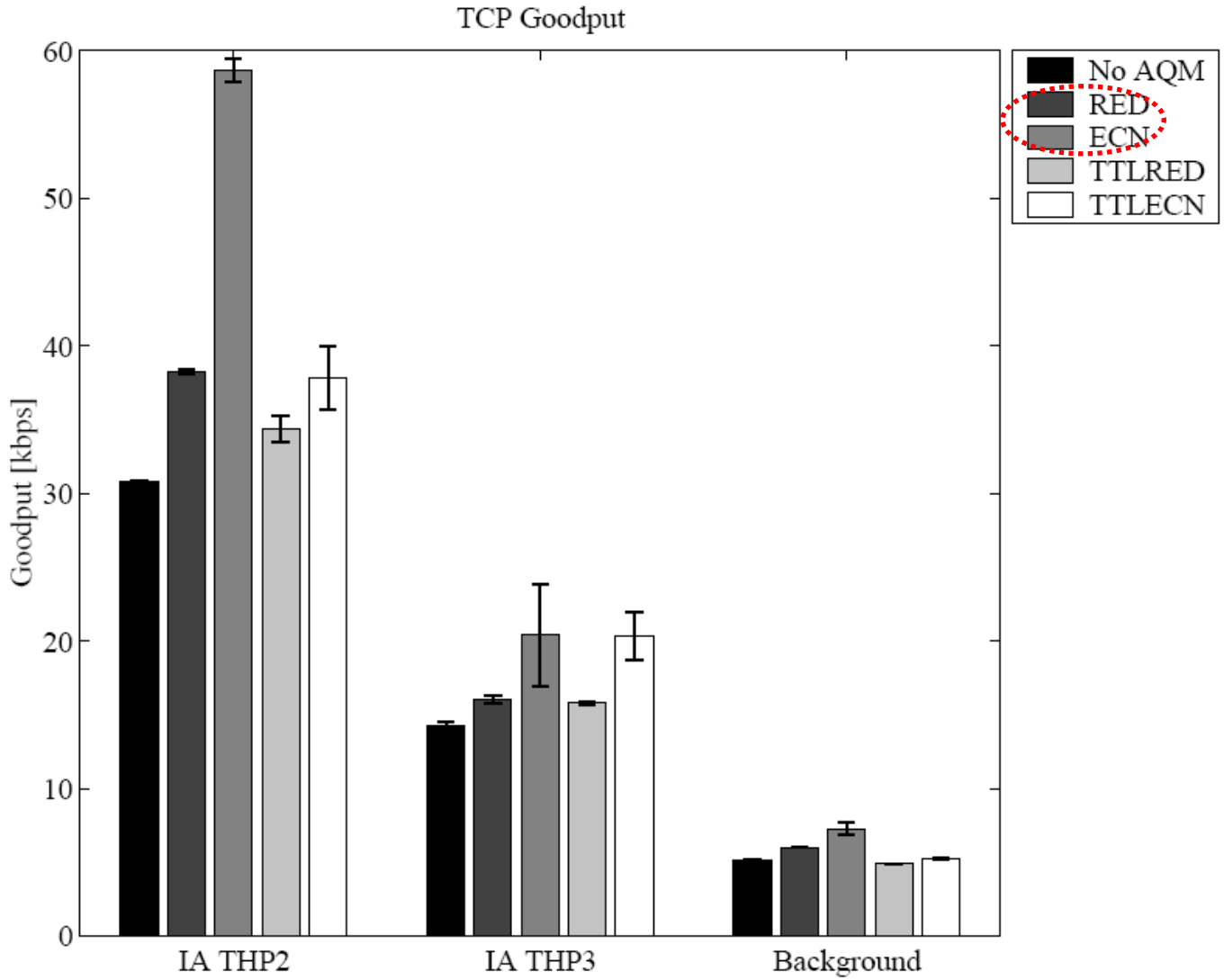
End-to-end (E)GPRS simulator using ns-2



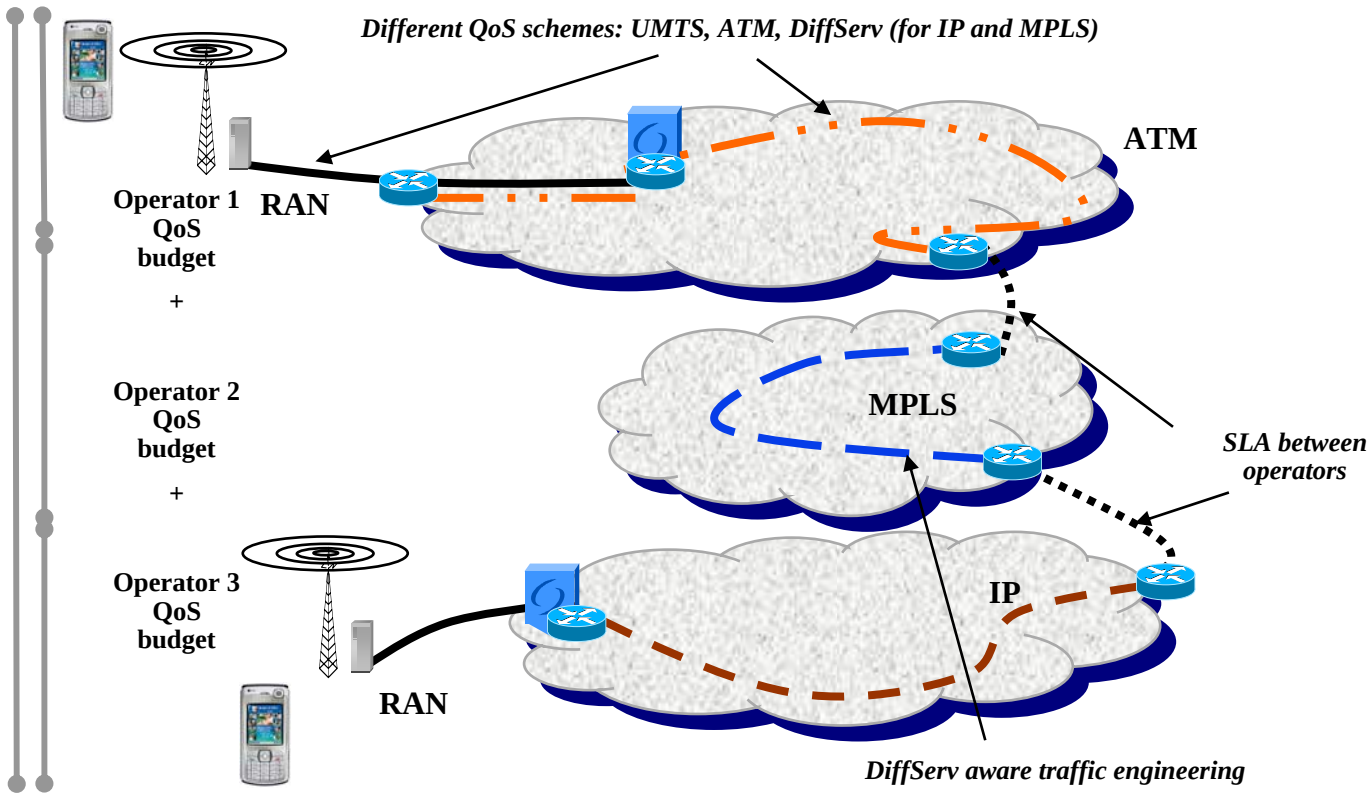
End-to-end delay (95th percentile)



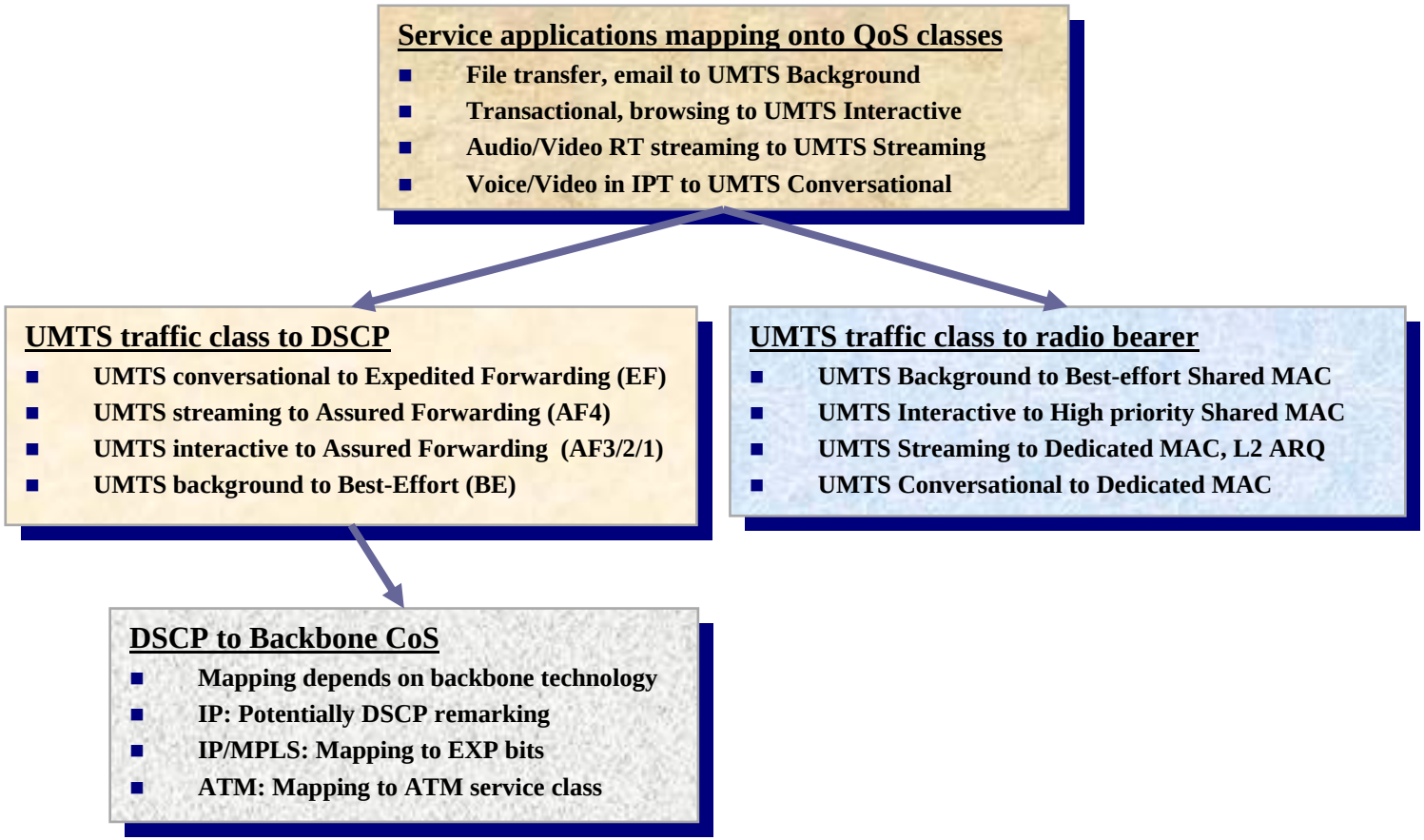
Mean end-user TPC throughput



Backbone QoS



QoS Mapping: examples



References

- D. Soldani, M. Li and R. Cuny (eds.), **QoS and QoE Management in UMTS Cellular Systems**, John Wiley and Sons, June, 2006, 460 pp.
  - http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0470016396.html
  - http://www.connecting.nokia.com/NOKIA/nns.nsf/a/78786C61AB5A7C5AC225718F0026BAA3

    (contact Mr. Geoff Farrell @ Wiley gfarrell@wiley.co.uk)

See also:

  - http://lib.tkk.fi/Diss/2005/isbn9512278340/