universidade de aveiro
theoria poiesis praxis

TCP/IP
Quality of Service

Qualidade de Serviço
2023/2024

Mestrado Engenharia Computadores e Telemática
DETI-UA

1

TODAY

- We will see mechanisms to add quality of service inside the network, providing "other" guarantees than the basic "TCP connection pipe" assurances

2

2

Multiservice Networks

Emerging services – heterogeneous requirements

QoS over IP networks

3

3

Current services

- Internet has many services beyond the basic network services
- Services
  - Interactive games
  - Audio/video
  - High definition moving image
  - Data base
  - Information storage
  - Communication networks

That require large transport systems

- Data networks can be very complex!

4

4

Services requirements

- Packet loss
  - Some applications (e.g., real-time audio/video) support losses
    - Voice supports more losses than video
    - TCP and its retransmissions
  - Other applications (e.g., file transfer, telnet) require 100% of success in transmission
    - However, they use TCP
- Bandwidth
  - Some applications (e.g., multimedia) require a minimum bandwidth
    - Buffer gets full
    - Large delays and some losses
  - Other applications ("elastic applications", e.g., email, file transfer) use the bandwidth they can get
- Timing: delay and jitter
  - Some applications (e.g., Internet telephony, multiplayer games) require low delays
  - Other applications (non-real-time) do not present strict limits on end-to-end delay
  - Some applications do not react well to delay variations (jitter)

5

5

Multimedia services

- Transmission of different types of information in the same service
  - Voice
  - Video
  - Data
- It is required synchronization between these types of information
  - Sending and reception in the terminal equipments
- It is required to handle different requirements of services in the same network
  - Interactivity vs non-interactivity
  - Bandwidth
  - Delay
  - Losses
- It is required to support interactivity in environments with variable delay

7

7

Multimedia services

- Transmission of different types of information in the same service
  - Voice
  - Video
  - Data
- It is required synchronization between these types of information
  - Sending and reception in the terminal equipments
- It is required to handle different requirements of services in the same network
  - Interactivity vs non-interactivity
  - Bandwidth
  - Delay
  - Losses
- It is required to support interactivity in environments with variable delay

Networks do not handle this usually: they are best-effort

8

8

Multimedia: quality assurance
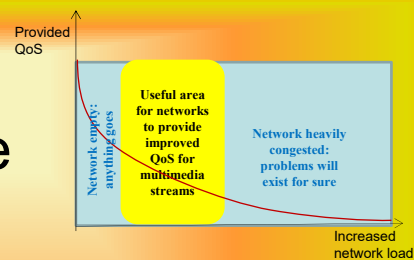
- **Multimedia applications assume best-effort networks**. As such there are interactions between the codecs and the network behaviour.
- Internet multimedia applications have several usual strategies:
  - Dynamically changing codecs, trading quality by bandwidth/resilience
  - Buffering, dynamically changing the size of the buffer
  - Progressive downloads, prefetching some seconds/minutes beforehand
- However, in networks medium loaded, some resource management is required, favouring multimedia flows over (best-effort) data flows, effectively creating an approach of weighted multiplexing gains.

9

9

What is needed to guarantee QoS?

- Example: 1Mbps IP audio/video stream and FTP transfer share a 1.5 Mbps connection.
  - FTP bursts can congest the router, causing loss in audio/video
  - It is intended to give priority to audio/video
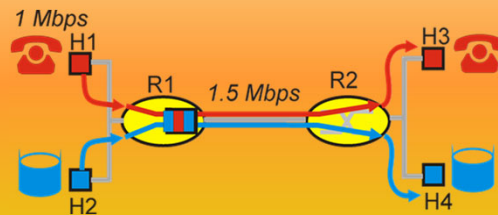
Principle 1

Packet marking is required so that the router can distinguish between the different traffic types; new policies are needed on the router for handling packets
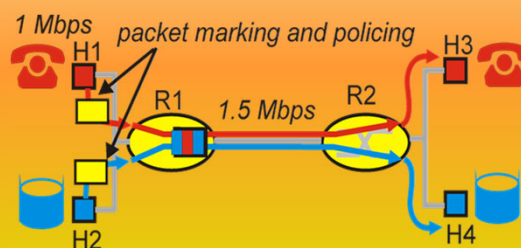
10

10

What is needed to guarantee QoS?

What if the applications "misbehave" (e.g. audio/video sends more than the declared bitrate)?
  - policing: forces the compliance of the sources to the agreed bandwidth
- marking and policing at the network entry

Principle 2

Provide **protection** (isolation) of one traffic class in relation to the other

11

11

5

Para garantir a Qualidade de Serviço (QoS), é necessário ter em consideração alguns princípios. Um deles é a atribuição de uma largura de banda fixa ao stream de áudio/vídeo. No entanto, esta abordagem pode ser ineficiente se o stream não utilizar toda a largura de banda atribuída.

O princípio 3 é que, ao fornecer isolamento (reservar recursos dedicados a determinados fluxos), é desejável utilizar os recursos tão eficientemente

What is needed to guarantee QoS?

- Can we assign a Fixed BW to the audio/video stream?
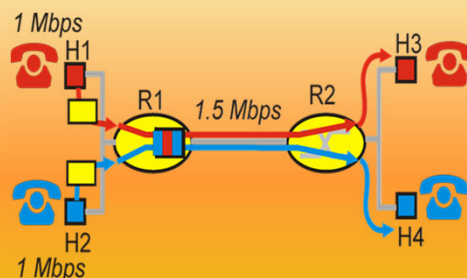 - Inefficient use of the bandwidth if the stream does not use the bandwidth that was assigned
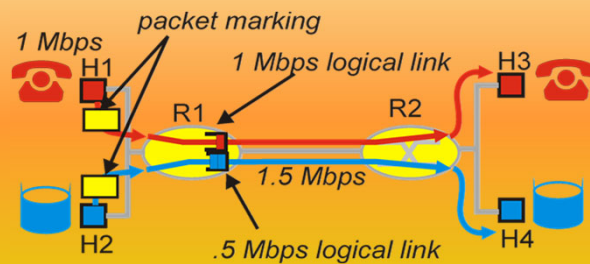
Principle 3

When providing isolation, it is desirable to use the resources **as efficiently as possible**

12

What is needed to guarantee QoS?

- It is not possible to support requests that exceed the connection capacity

Principle 4

Call admission: the stream declares its requirements, and the network can block the call (busy signal) if it can't support them

13

What is needed for QoS support

1. Some form of <u>signalling</u> between applications and network (and internally between parts of applications)
2. Signaling for <u>resource reservation</u>/management (typically RSVP)
3. Ability to <u>differentiate traffic</u> treatment inside network equipment (typically queueing strategies in routers, see last slides)
4. <u>Control and policing</u> of the network usage (see last slides).

Previous slides we saw the basic concept that allow 3) and 4) inside the network. This section discusses 1) and (mostly) 2, and how the concepts work together.

14

14

Main approaches in IP networks

Basic IP service:

– Packets suffer delays, losses, jitter and reordering.

Differentiated Services

– Classes of services

Integrated Services
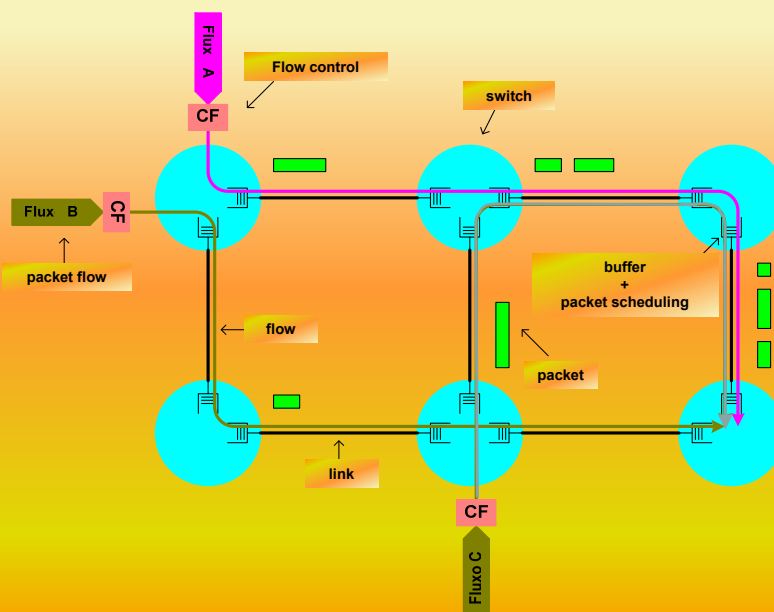
– Defined service levels

IETF

15

Main differences

- IntServ
  - ✓ Rely in flows, implementing two types of E2E services: GS (guaranteed service) and CL (Controlled Load).
  - ✗ Does not scale in the core!!!

- DiffServ
  - ✓ Works on simpler aggregates, with an approach effective for QoS on the core.
  - ✗ Does not provide E2E guarantees.

16

16

IP network with QoS support: summary

17

17

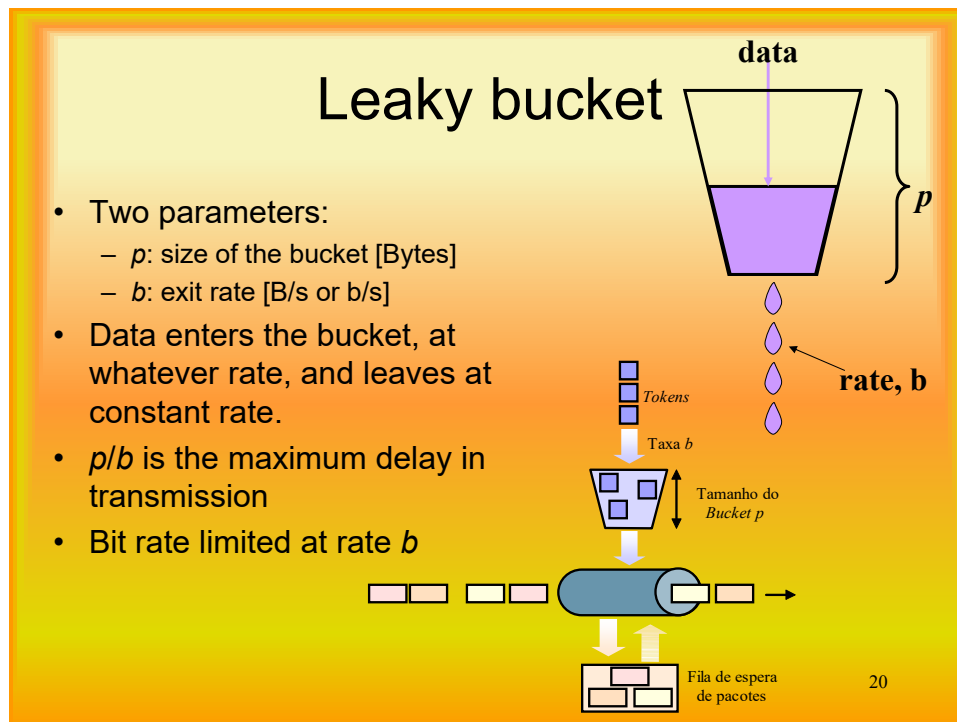# How to condition traffic?

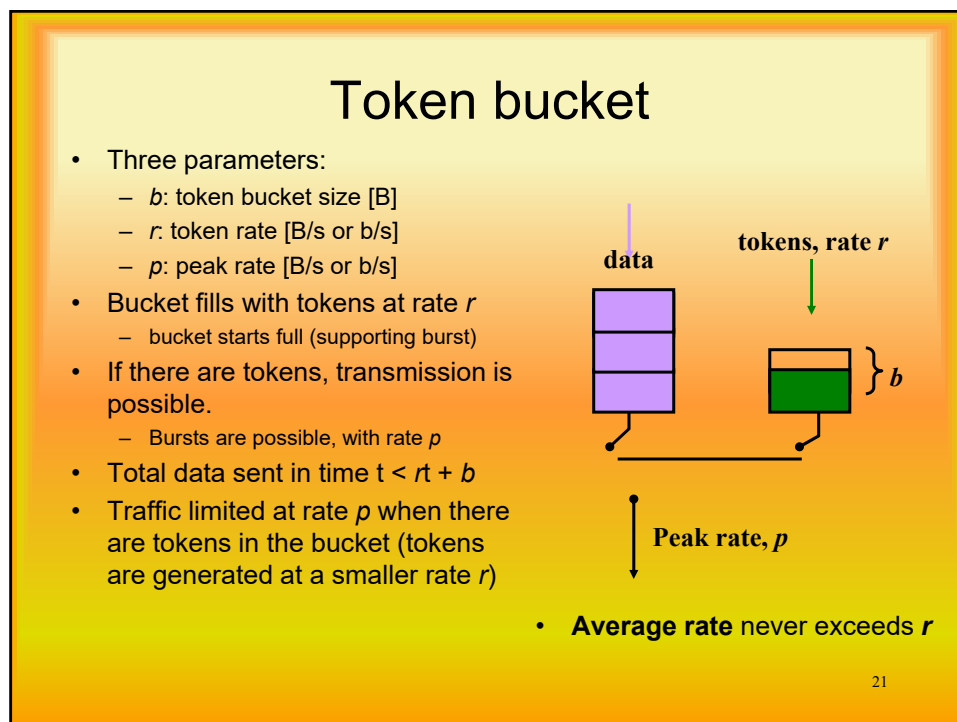## Basic concepts

18

18

Policing/Shaping

- Policing reduces the impact of excess traffic
  - Loss of excess packets
  - Tagging with lower qos
- Shaping stores traffic, smoothing bursts
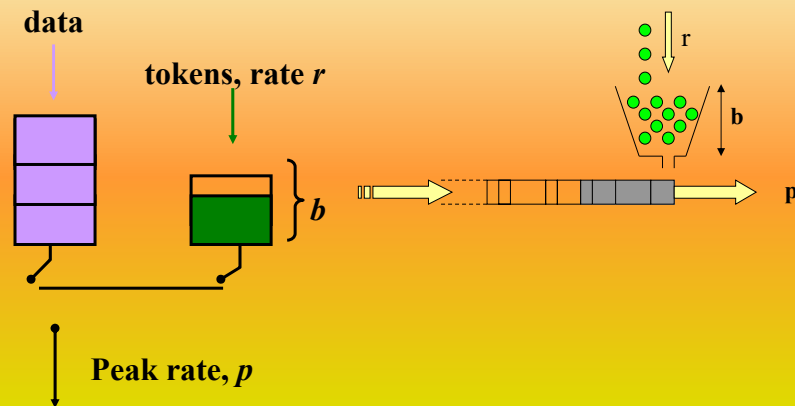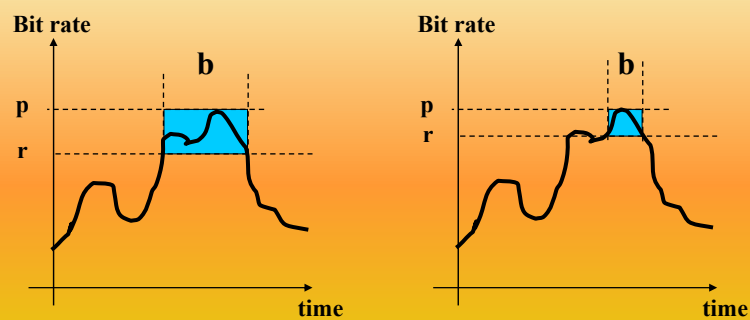 - Only allows traffic to be sent at A certain rate

19

19

Leaky bucket

data

- Two parameters:
  - *p*: size of the bucket [Bytes]
  - *b*: exit rate [B/s or b/s]
- Data enters the bucket, at whatever rate, and leaves at constant rate.
- *p/b* is the maximum delay in transmission
- Bit rate limited at rate *b*

p

rate, b

Tokens

Taxa *b*

Tamanho do *Bucket p*

Fila de espera de pacotes

20

20

Token bucket

- Three parameters:
  - *b*: token bucket size [B]
  - *r*: token rate [B/s or b/s]
  - *p*: peak rate [B/s or b/s]
- Bucket fills with tokens at rate *r*
  - bucket starts full (supporting burst)
- If there are tokens, transmission is possible.
  - Bursts are possible, with rate *p*
- Total data sent in time t < *r*t + *b*
- Traffic limited at rate *p* when there are tokens in the bucket (tokens are generated at a smaller rate *r*)

tokens, rate *r*

data

b

Peak rate, *p*

- **Average rate** never exceeds *r*

21

21

Token Bucket

data

tokens, rate *r*

b

r

b

p

Peak rate, *p*

22

22

Describing sender traffic

Bit rate

b

p

r

time

Bit rate

b

p

r

time

• In the time interval Δt the number of packets (bytes) accepted must be equal or less than (rΔt + b).

– Actions to take may be policing or shaping

23

23

Token(s)

- Basic tool for "measurement"
- Can be used to:
  - Describe traffic
  - Validate traffic compliance
  - Clarify terms we use regularly ("bandwidth", "burst",...)
  - Assist in mathematical analysis of networks

24

24

Example

- **2 tokens, size 100 bytes, added every second to the token bucket, with total capacity 500 bytes. Operations are handled in bytes**
  - **Average rate = 200 bytes/sec,**
  - **burst size = 500 bytes**
  - **Packets larger than 500 bytes are never sent**
- **Is it possible to get a peak rate above 200 bytes/sec ?**
  - **Yes, any rate is possible as long as you send in each second only 500 bytes**
  - Example: we can transmit 100 bytes in 1ms, meaning 100Kbps of peak transmission…

25

25

Admission control

26

Policing (drops)

- Dropping packets is one of the possible actions to ensure that the expected network performance is not exceeded
- For a some traffic types (e.g. voice) it makes no sense to miss only "a few" packets, minimum guarantees are required
- Admission control:
  - Reject/accept flows, with a well-defined traffic, making sure that a certain QoS will be guaranteed
  - admission control may however have actions such as shaping

  It is an action inherent in the operation of circuit switching networks (telephone networks) – all calls are previously "admitted" before being processed.

27

Quando ocorre uma sobrecarga nos buffers do router (filas de espera), é comum descartar pacotes para evitar a sobrecarga não só nos buffers, mas também na rede como um todo. No entanto, a questão de como descartar esses pacotes é importante.

Existem várias abordagens para descartar pacotes:

1. Último pacote a chegar: Nesta abordagem, o último pacote que chega é descartado. Isso pode ser útil quando se deseja dar prioridade aos pacotes mais
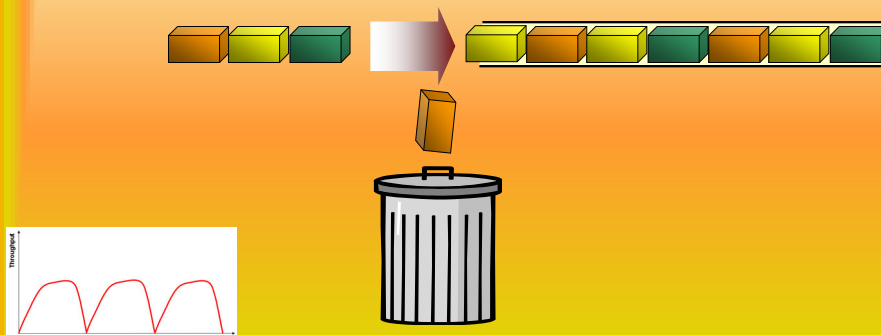
Packet drop techniques

- Dropping packets to stop overload of the router buffers (waiting queues)
  - And of the network.
- How to drop?
  - Last packet to arrive?
  - First packet to arrive?
  - Any random packet?
  - Differentiated packets per class?

28

28
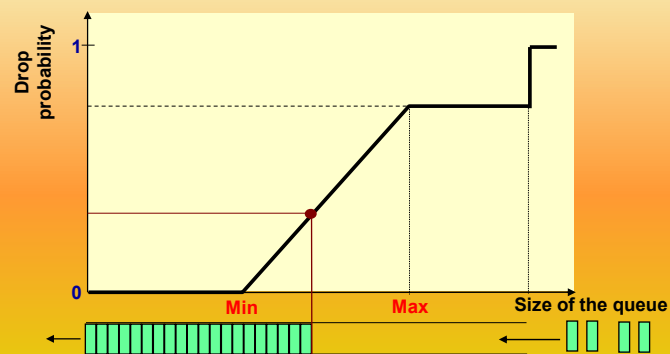
Congestion control (I) – *Tail Drop*

Tail drop: in each buffer, the packets received are dropped when the buffer is full

Problem: Interaction with TCP flow control mechanisms in the core routers ➔ **Global synchronization of the TCP traffic sources.**

29

Congestion control (II) –
RED (Random Early Detection)

30

RED: Random Early Detection

- Random Early Detection:
  - Handles congestion before it appears
  - packet loss → Congestion signal
    - Source slows down
    - Prevents actual congestion
- What packets to lose?
  - Probability of packet loss ∝ queue length
  - Monitoring of flows
    - Cost in processing vs overall network performance
  - Queue length – exponential average:
    - "smooths" reaction to traffic bursts
    - Limits constant heavy traffic, being good for Intserv (Controlled Load)
  - Packets may be lost or marked as "offending"
    - RED-aware routers will lose packets "offending", when needed
  - Source should adapt:
    - TCP: OK!
    - real-time traffic - UDP ?

31

31

RED

- **RED Parameters**
  - Minimum queue threshold (minQ)
  - Maximum queue threshold (maxQ)
  - Average Queue length (AvgQ).
    - Dynamic calculated
  - Maximum drop probability (maxP)
  - Drop probability (P)
    - Dynamic calculated
- **Algorithm**
- For each incoming packet
  - If AvgQ <= minQ
    - queue packet
  - If minQ <= AvgQ < maxQ
    - Mark packet with probability P
  - If maxQ <= AvgQ
    - Mark the packet

32

RED

- **Size of the waiting queing**
  - **AvgQ = (1 – weight) x AvgQ + weight x currQ**
    - 0 < Weight < 1
    - currQ é o tamanho actual da fila de espera

- **Drop probability**
  - **TempP = MaxP x (AvgQ – minQ)/(maxQ – minQ)**
  - **P = TempP/ (1 – count.TempP)**
    - Count is the number of new packets that reached the queue

33

# WRED (*Weighted Random Early Detection*)

- **3 drop levels – 3 drop levels for different sizes of que queue**
  - **The last to be discarded are those with more priority**
  - **The first to be discarded are those with lower priority**

34
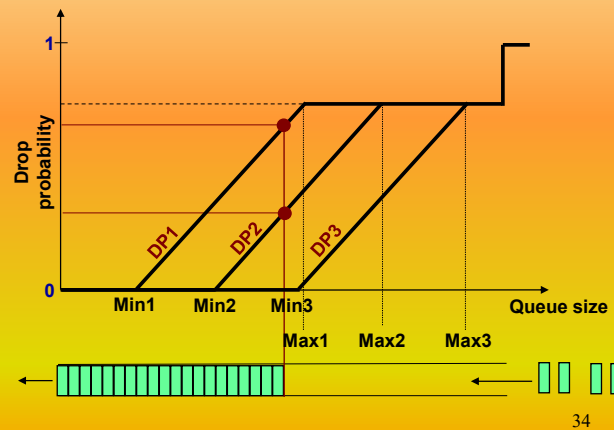
34

Congestion control (II) – RED impact in global synchronization

- **With RED we are able to decrease the TCP global synchronization in core routers,** created by all input flows reacting at the same time to the queue congestion in core routers

35
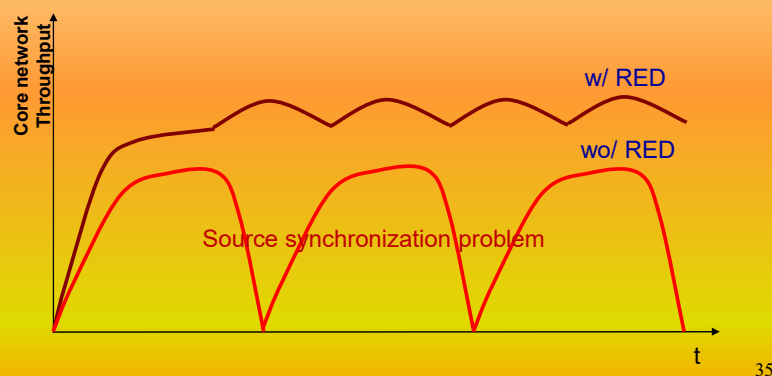
35

Scheduling algorithms

Scheduling algorithms: Decide the order in which packets belonging to different streams are served in a queue.

Work conserving scheduling algorithms ensure that the server is always busy if there is a packet waiting to be transmitted.
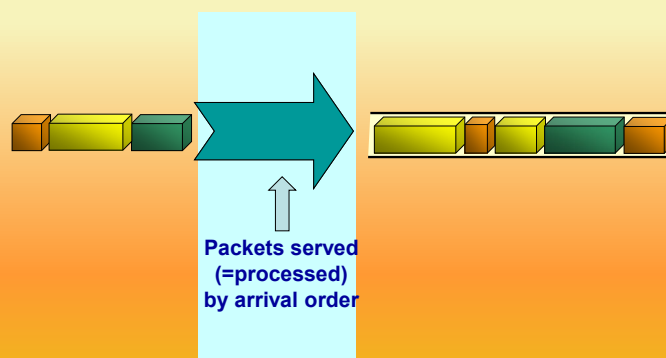
Examples of work conserving scheduling algorithms:

(1) FIFO,

(2) Strict priority,

(3) Fair Queuing,

(4) Weighted Fair Queuing.

36

36

First In First Out (FIFO)

Packets served (=processed) by arrival order
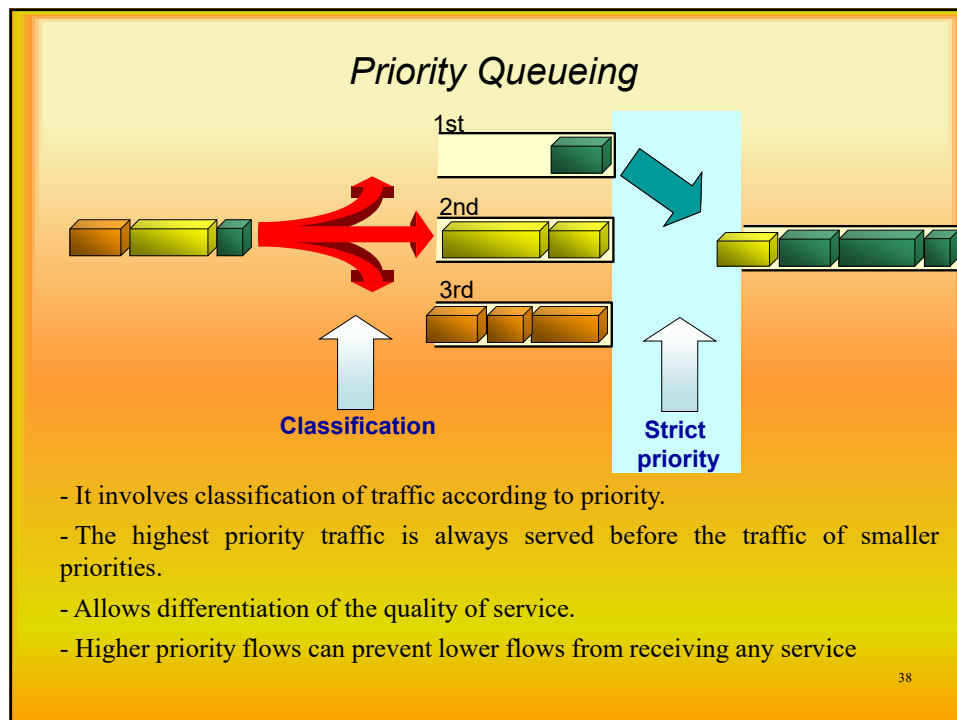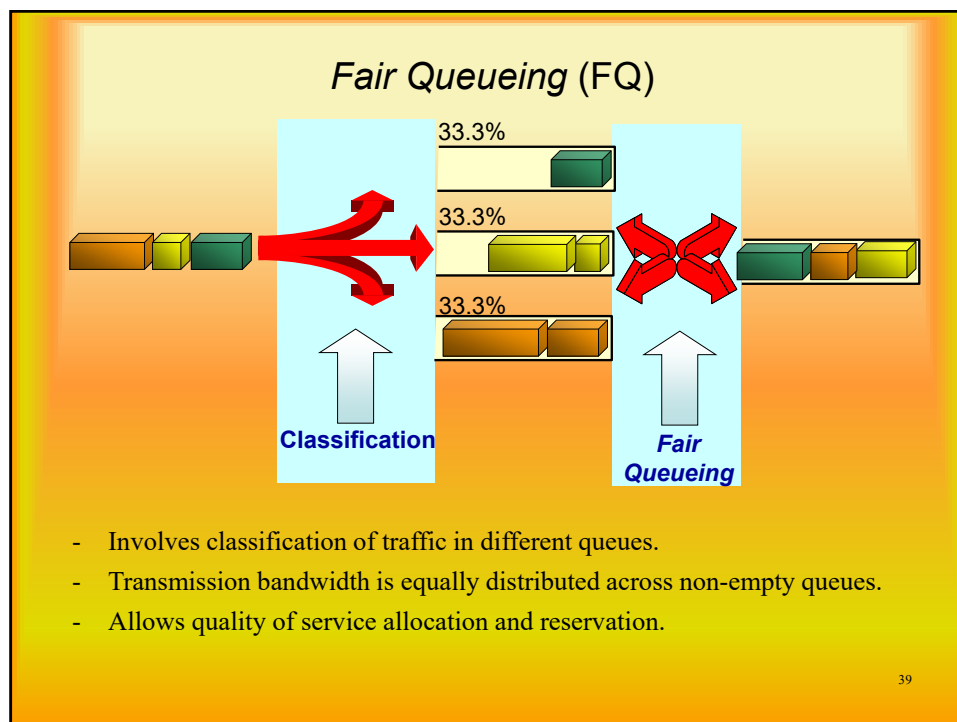
- It does not perform sort processing.
- It does not allow differentiation of quality of service.
- Flows with *n* times more traffic receive *n* times more service.
- In finite queues, streams with smaller packets, receive more service

37

37

Priority Queueing

- It involves classification of traffic according to priority.
- The highest priority traffic is always served before the traffic of smaller priorities.
- Allows differentiation of the quality of service.
- Higher priority flows can prevent lower flows from receiving any service

38

38

Fair Queueing (FQ)

- Involves classification of traffic in different queues.
- Transmission bandwidth is equally distributed across non-empty queues.
- Allows quality of service allocation and reservation.

39

39

Weighted Fair Queuing (WFQ)

This algorithm ensures that each queue achieves a percentage of the connection bandwidth <u>at least</u> equal to its weight divided by the sum of the weights of all queues

$$R_A = \frac{2}{2+3+4} LB$$

Queue A (weight = 2)

$$R_B = \frac{3}{2+3+4} LB$$

Queue B (weight = 3)

**Connection
(LB – bandwidth)**

$$R_C = \frac{4}{2+3+4} LB$$

Queue C (weight = 4)

40

40

Exercise

- Consider that in this network, the serial interfaces of the connections between routers 1, 2 and 3 only have FIFO mechanisms active. However, you can configure also Random Early Detection (RED) with 2 drop probabilities: high discard probability and low probability.
  - How does this mechanism work?
  - If you had to choose priorities for a video and file transfer service, which one would you choose? Justify.
  - If routers had Weighted Fair Queueing active with 3 different queues and with weights of 2 for the voice queue, 5 for video, and 3 for data, what bandwidth is available for each service?
  - In the same case, if there are no video packets, what bandwidth is available for voice and data?
- Considering the network with an active 50 Kb/sec video service and FIFO queues with unlimited capacity, determine the delay of the 1000-bit video service packets, considering that only the serial connections have a meaningful delay.
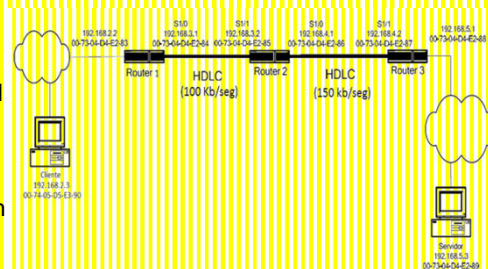
41

41

Quality of Service

Supporting network services

42

NOW

- We will see how we can signal and provide QoS in the network, using the previous mechanism.

43

43

Integrated Services
(IntServ)

44

Integrated Services:

- Controlled Load CL (RFC2211)
  - Assures service end-to-end (E2E) that is as load independent as possible (*good best-effort always)*
  - End nodes will receive most packets with minimal delays in routers
- Guaranteed Service GS (RFC2212)
  - Assures E2E service in terms of delay, for a given bandwidth.
- Best Effort BE

  Does not guarantee any quality of service,

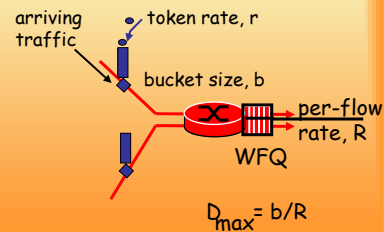 only the existence of connection.

  Other services are possible (RFC2215-6)

  Defined signaling (RFC2205,2210): RSVP

arriving traffic

token rate, r

bucket size, b

per-flow rate, R

WFQ

$D_{max} = b/R$
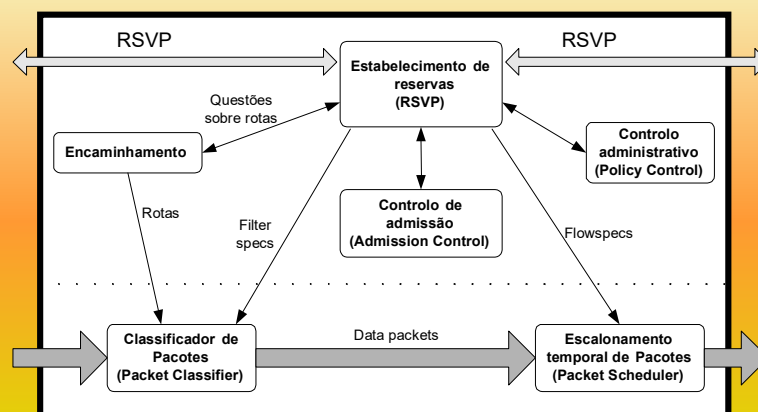
45
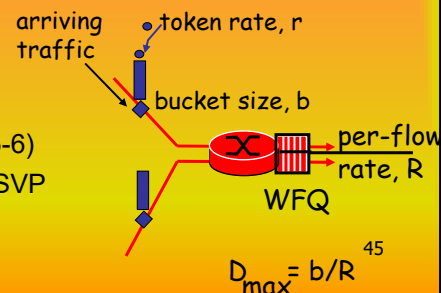
45

Control model for IntServ

- Flow specification
  - Intserv is used for flows!
- Routing
- Admission control
  - Sender should control the sending of packets using a token bucket model
- Policing
- Resource reservation
- Packet Scheduling

arriving traffic | token rate, r

bucket size, b

per-flow rate, R

WFQ

$D_{max} = b/R$

46

46

Router architecture

RSVP | RSVP

Estabelecimento de reservas (RSVP)

Questões sobre rotas

Encaminhamento

Controlo administrativo (Policy Control)

Rotas

Filter specs

Controlo de admissão (Admission Control)

Flowspecs

Classificador de Pacotes (Packet Classifier)

Data packets

Escalonamento temporal de Pacotes (Packet Scheduler)

47

47

Call admission process

The session that starts should:

- declare its QoS requirements
  - R-spec: defines the QoS that is being required
- characterize the traffic that will send to the network
  - T-spec: defines the characteristics of the traffic

A signaling protocol is required to carry the R-spec and T-spec towards routers (where the reservations are needed):

- RSVP [RFC 2205]

48

48

RSVP (*Resource Reservation Protocol*)

- RFC 2205
- Encapsulated in IP; protocol type = 46 (0x2E)
- Signalling is based on the PATH and RESV message exchange
  - PATH announces the traffic characteristics of the sender
  - RESV confirms the reservations, initiated by the receivers
  - If the reservation is not possible, the message RESV ERR is sent
- The states of routers must be refreshed periodically (process know as: soft states)

(1) *PATH*   (2) *PATH*   (3) *PATH*

Emissor   Nuvem RSVP   Receptor

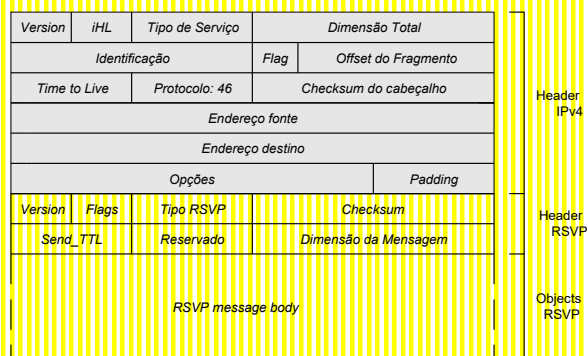(6) *RESV*   (5) *RESV*   (4) *RESV*
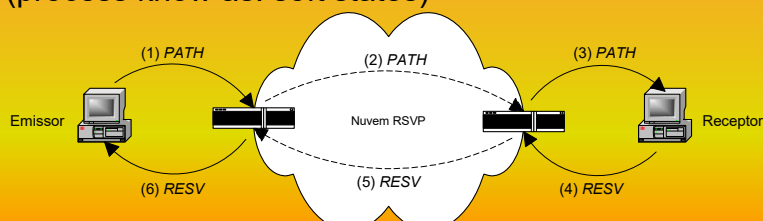
49

Service template: generic parameter format

- There are generic message parameters defined for IntServ operation
  - NON_IS_HOP (flag): node does not support Intserv
  - NUMBER_OF_IS_HOPS: counter of QoS-aware nodes
  - AVAILABLE_PATH_BANDWIDTH: available bandwidth (Adspec)
  - MINIMUM_PATH_LATENCY: path delay (Adspec)
  - PATH_MTU: MTU maximum transfer unit size possible to use.
  - TOKEN_BUCKET_TSPEC: traffic specifications as token bucket parameters
    - r (rate), b (bucket size), p (peak rate)
      m (minimum policed unit), M (maximum packet size)

50

50

RSVP Messages

Version	iHL	Tipo de Serviço	Dimensão Total		
Identificação			Flag	Offset do Fragmento	
Time to Live		Protocolo: 46	Checksum do cabeçalho		Header IPv4
Endereço fonte					
Endereço destino					
Opções				Padding	
Version	Flags	Tipo RSVP	Checksum		Header RSVP
Send_TTL		Reservado	Dimensão da Mensagem		
RSVP message body					Objects RSVP

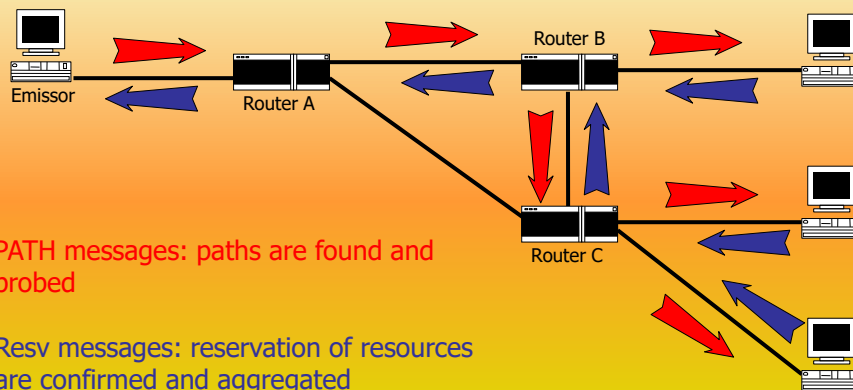|	Dimensão do Objecto	Nº Classe	Tipo de Classe
RSVP object format:	Conteúdo		

51

51

RSVP operation

- **Receiver joins a (multicast) group**
  - Operation outside RSVP
  - Senders do not need to join the group
- **Signalling sender-network**
  - *Message path:* makes the sender known to routers
  - Path erasure: removes from routers the path corresponding to a sender.
- **Signalling receiver-network**
  - *Message reservation:* reserves resources from the sender(s) to the receiver
  - Reservation cancelling: deletes the reservations made by receiver
- **Signalling network- end-system**
  - *path error*
  - *reservation error*

52

52

RSVP – Basic operation

Emissor | Router A | Router B | Router C

PATH messages: paths are found and probed

Resv messages: reservation of resources are confirmed and aggregated

53

53

Reservation styles supported in RSVP (STYLE)

- **"Fixed Filter"** (Style Option Vector = 0x00000A)
  - Receiver specifies a reservation value per each sender
- **"Wildcard Filter"** (Style Option Vector = 0x000011)
  - Receiver sets a single reservation value to receive info from any sender
- **"Explicit Filter"** (Style Option Vector = 0x000012)
  - Receiver specifies a list of senders from which to receive information, and a single reservation number to receive traffic from those identified senders.

In RSVP RESV:
  - The reservation style is set by the object STYLE
  - Senders are identified by the object FILTER_SPEC

54

54

Flows

- A flow is a set of packets related by some reasons
  - In RSVP a flow is a set of packets that cross a Network Element (NE), and that are covered by the same QoS request.
- A Packet Classifier sets which packets belong to each flow
  - IPv6: Flow label helps this classification
- In ISPs…
  - Microflow: TCP or similar connection…
  - Macroflow: Large set of packets between two NEs

Flowspec define the traffic parameters
  - LB, buffering needs, using token bucket specs

Filterspec identify the packets in the flow
  - Basic Filter: Source, Dest address/port pair
  - Advanced data filter: depends on packet content

55

55

Service models for Intserv

- There are service models that describe the semantics of service for the flow.
- Specify how the packets belonging to a flow are to be treated by the network elements.
- Parameters: general format:
  - \<service_name\>.\<parameter_name\>
    - Can have values between [1, 254]
- Services:
  - TSpec: specify the traffic pattern (CL+GS)
  - Rspec: specify the service request (GS)

56

56

IntServ General definitions

- Token bucket (rate, bucket-size):
  - Used to define data rate
- Admission control:
  - Verification before accepting a reservation
- Policing:
  - Verify if TSpec is fulfilled
  - The packet treatment may be changed if Tspec is violated (e.g. service degradation, packet discard, etc…)
- Parameters: locals and composed
  - Total path value is the combination of the local values with the path

57

57

Controlled Load

- Service that provides QoS similar to what exists in unloaded network "Best-effort in unloaded environments":
  - Statistical guarantees
  - No delay limitation
- Motivation
  - Support of applications sensitive to large delays
  - Keeping minimal functionalities
- Operation
  - Average delay in queues small or null
  - Small or null losses.
  - Analysis period: much larger than a burst period
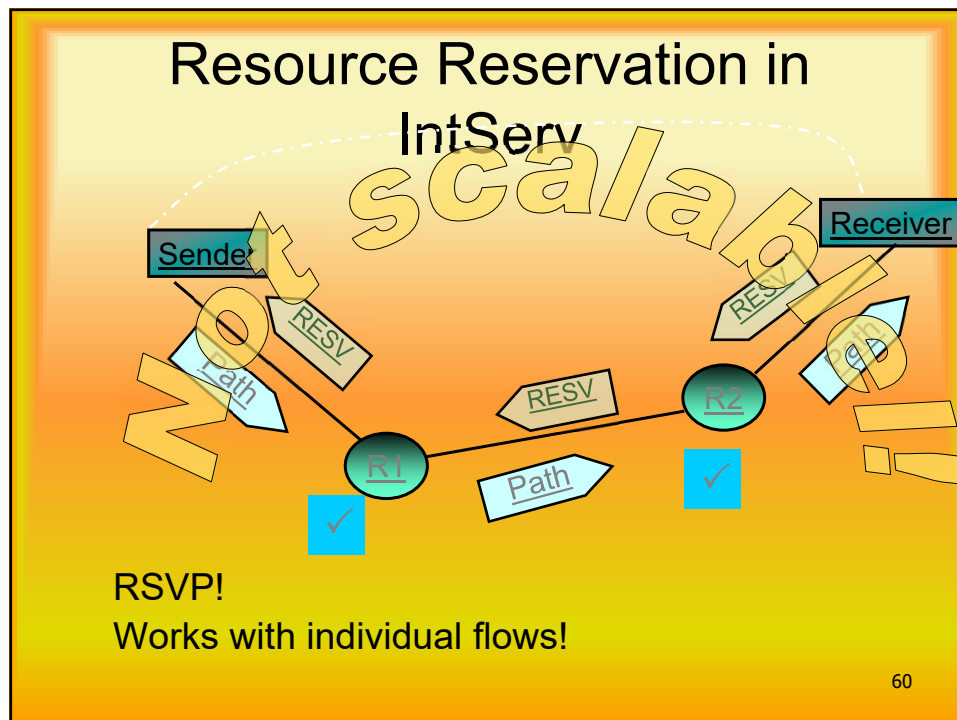
58

58

Guaranteed Service

Service providing na assured bitrate, with a limit on total delay
  - Deterministic guarantee
  - No jitter guarantee
- Parameters used:
  - TSpec: TOKEN_BUCKET_TSPEC
  - RSpec: R (rate), S (delay slack term, μs)
    - Larger R: smaller E2E
    - Larger S: larger delays, but better reservation possibilities
- Admission control:
  - Weighted Fair Queuing (WFQ)
- Policing:
  - drop, move to best-effort; reshape (delay)

59

59

Resource Reservation in IntServ

Not scalable!

Sender

Receiver

RESV

RESV

Path

RESV

R2

R1

Path

Path

✓

✓

RSVP!
Works with individual flows!

60

60

Example –
Resource Reserve in IP Networks

- RSVP reservations for voice and video services over IP
  - Low quality: Reservation of 64 kb/seg for voice and 1 Mb/seg for video
  - High quality:
    - IntServ Guaranteed Service for voice with 1Mb/seg bandwidth
    - IntServ Controlled Load for video with 5 Mb/seg bandwidth
    Or
   - IntServ Guranteed Service for voice and video with 6 Mb/seg

61

61

Differentiated Services (DiffServ)

62

Basic idea

> The real question is to choose which packets shall be dropped. The first definition of differential service is something like "not mine."
> -- Christian Huitema

- Some packets are more important than others
  - whoever "pays" for better service, should get it...
- Differentiated services should provide a mechanism to specify relative priority of packets
  - Implement simple routing operations on the network's core routers and leave complex operations to the network edge routers.

Nokia Sweden

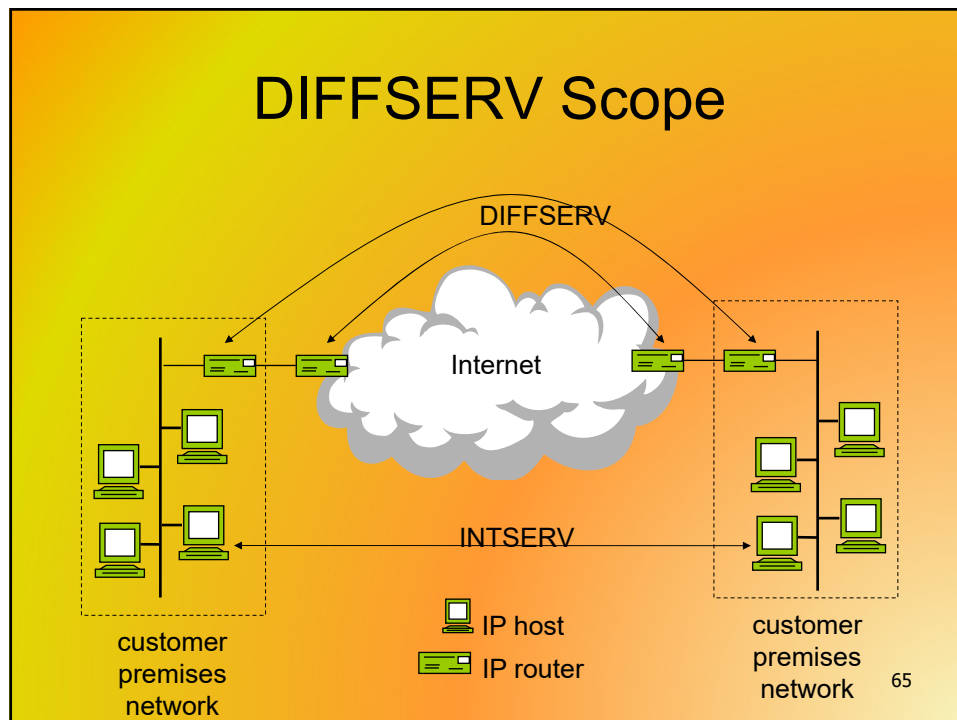Nokia Portugal

Limited Bandwidth

63
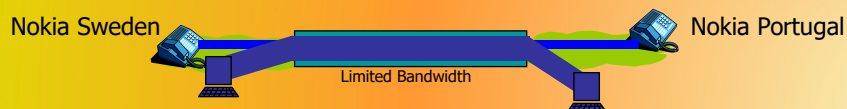
63

Objectives

- Ability to charge differently for different services
- Ability to discriminate services in a scalable way for the core, with low complexity
  - No per-flow state
  - No per-flow signalling
- Easy to evolve, with simple start-up implementation
  - Define only elements that may implement any class of service
- Simpler and more efficient than IntServ
  - With signalling separated by services
  - With "more-or-less" static user-services
  - With traffic aggregated in classes
- Without individual reservation per link

64

64

DIFFSERV Scope

DIFFSERV

Internet

INTSERV

💻 IP host
▭ IP router

customer premises network

customer premises network

65

65

DiffServ

- Oriented towards the core network (*core*)
- No direct E2E guarantees: service assurances are structured by layers
- No per-flow control
  - Packets marked by the network (not the app)
  - All existing applications can be supported.
- Simple control and marking tools (RFC2474)
- $\Rightarrow$ no E2E user service model (RFC2475)
- $\Rightarrow$ More proper to speak of CoS than of QoS
- $\Rightarrow$ Simpler to implement in core than IntServ
  - $\Rightarrow$ Can be deployed in current networks!

66

66

DiffServ approach

Based in three assumptions (implicitly assumed):

- The network is overprovisioned regarding the needs of QoS traffic ($\Longrightarrow$ small number of customers with DiffServ)
- Non-real-time traffic will be the largest load of all network traffic
  - No explicit reservation per link, and as such links in "popular" areas could have problems with congestion (when more people would like to use prioritize traffic)
- E2E services can be implemented over networks with different QoS features

67

67

Services and SLA

- Two types of services:
  - quantitative; require numerical metrics and information about entry and exit points
  - qualitative; require only entry information
- Not (necessarily) per-flux:
  - Packet treatment is performed over aggregates of a "source"
  - A service level specification (SLS) is defined
  - The service is provided by the network (and the host may not even perceive this)
- All services are unidirectional!

68

68

Architecture components

1. Define PHB (per-hop-behaviour) for the routers. These will operate over traffic aggregates. The PHB priotize traffic, in terms of delay and loss probability.
2. Traffic control and management is fundamentally performed at the borders, and then is aggregated.
3. Services are clearly separated from the network technical constrains.

69

69

Overview of a DiffServ network

All "QoS" definitions are local.
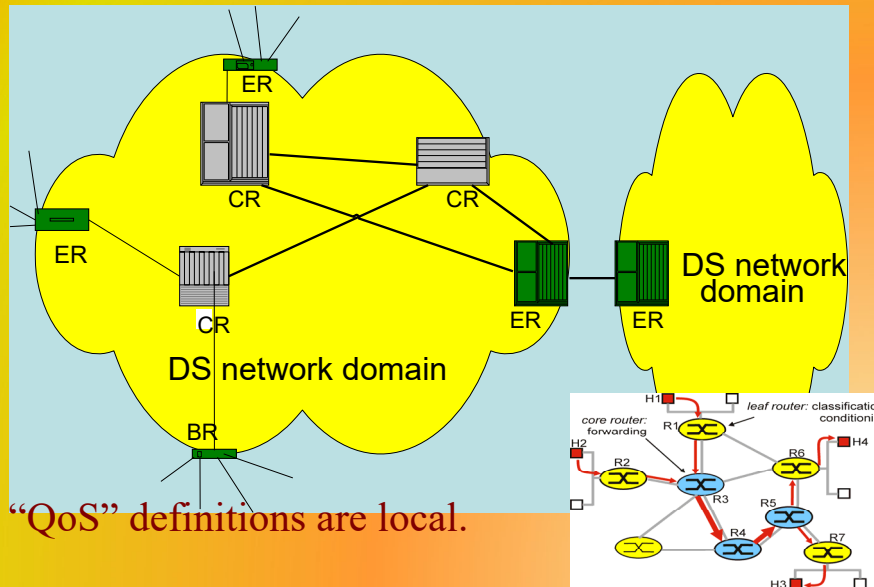
70

Functional elements

- *Edge (border) Routers:*
  - *Classify packets: Mark each packet in the Type of Service field of the IP header*
  - *Condition traffic: for example, they use a "Token Bucket" to verify that incoming traffic is contracted and*
    - *delay excess traffic or*
    - *drop non-conformant traffic*
- *Core (internal) Routers:*
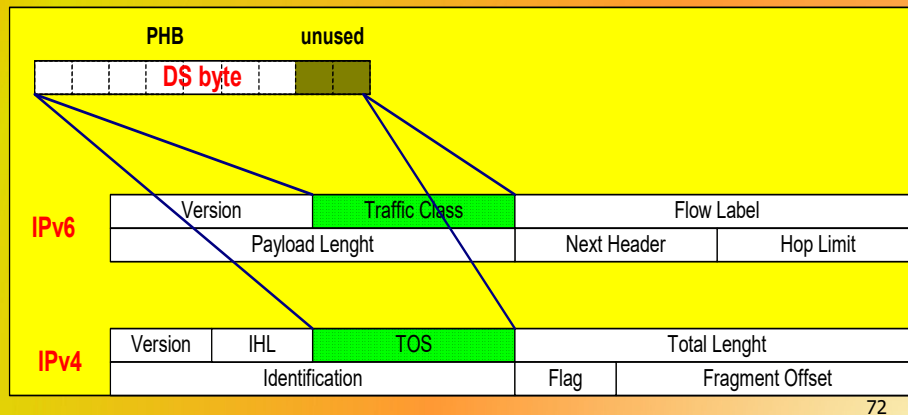 - *Identify the treatment to give to packets based on marking and according to a defined Per-Hop-Behavior (PHB)*

71

71

Field DS

Packets are marked in the Type of Service (TOS) field of the IPv4 header or Traffic Class of the IPv6:
- DSCP header - Differentiated Service Code Point, 6 bits
- rest - Currently Unused

PHB	unused
DS byte	

IPv6

Version	Traffic Class	Flow Label	
Payload Lenght		Next Header	Hop Limit

IPv4

Version	IHL	TOS	Total Lenght	
Identification			Flag	Fragment Offset

72

72

PHB

- Per-Hop-Behaviour is the forwarding behaviour that a DS node applies to a traffic aggregate. This is perceived in terms of:

  Delay and

  Packet loss probability

- The specification of transmission rate, losses and delay should be done in the SLS
- Any PHB is only defined inside an administrative domain

73

73

QoS Processing in
Core Routers

- Different Per-Hop-Behaviors (PHBs) result in different network performances that can be measurable
- PHBs do not specify which queuing mechanisms should be used
- Examples of PHBs:
  - Class A packets are assigned x% of the physical connection bandwidth during any time interval for a specified duration
  - Class A packets are always served first than Class B packets
  - Class A packets are served with twice the service bandwidth of Class B packets

74

74

DIFFSERV PHBs

- Two types of PHB already developed:
  - AF (Assured Forwarding) (RFC2597):
  - EF (Expedited Forwarding) (RFC2598):
    - virtual leased line (VLL) service

      +BE (best effort)

- Coupled to different types of services:
  - Premium (low delay) - EF
  - Assured (high transmission rates, low losses) - AF

75

75

DiffServ service classes

- *Default* (DE) → DSCP = 000000
  - *best-effort* service with a single queue, FIFO managed
- *Expedited Forwarding* (EF) → DSCP = 101110
  - "Virtual leased line" service
  - provides loss, delay, and delay variance control within a given maximum bandwidth
- *Assured Forwarding* (AF) → DSCP = aaadd0
  - provides a relative Quality of Service (AFi is served with more bandwidth than AFj for i<j)
  - in each class there are 3 precedence levels for packet deletion in case of congestion

AF Codepoints	AF1	AF2	AF3	AF4
Low drop precedence	001010	010010	011010	100010
Medium drop precedence	001100	010100	011100	100100
High drop precedence	001110	010110	011110	100110

76

Expedited Forwarding (EF)

- For critical traffic
  - Low delay, small jitter, no losses
- Nodes must forward these packets ASAP (PQ).
- Packets cannot be lost, or reorder
  - Resources must be reserved in a conservative way, by the maximum value.
  - Agreed bandwidth is assured
  - Packets out of profile are lost: stringent policing in the border
- EF can block all other network traffic.
- Defined for quantitative services, as it require entry and exit nodes well defined.
  - VLL: maximum LB defined, available when required

77

Assured Forwarding (AF)

- Defines four classes, and three levels of packet loss for each class.
  - AF11 - "best", AF13 - "worst"
- Relations between classes are not defined
  - Provisioning according with the expected usage
- Performance in each class should be degraded gradually (3 levels) in terms of packet loss, as traffic increases.
  - Packets inside profile will not be "usually" lost
  - Packets out of profile may be treated (almost) as BE, so higher bandwidths may be used if available
- Allows qualitative services, and only requires the knowledge of entry-node
  - Bandwidths are roughly respected

78

78

Borders: *Edge Routers*

- Control network access, policing and classifying entry traffic.
  - These is also required between DS networks, as diffserv definitions are local...
- Traffic may be in or out of profile.
- The network requires traffic conditioners at the borders (optional at *core*), that classify and act over traffic:
  - Meters – check the timing features of the flow, confronting with the SLA associated
  - Classifier – identifies the traffic class of the (for the) packet
  - Markers – set a DS codepoint to each packet (in/out profile)
    - Packets may be remarked
  - *Droppers* – remove packets out-of-profile
  - *Shapers* – delay packets out-of-profile, using buffers and smoothing methods
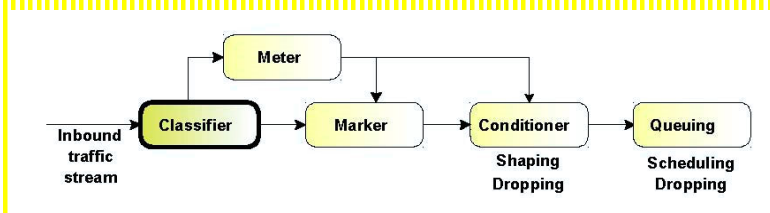
79

79

Border control model

Packet may be marked already
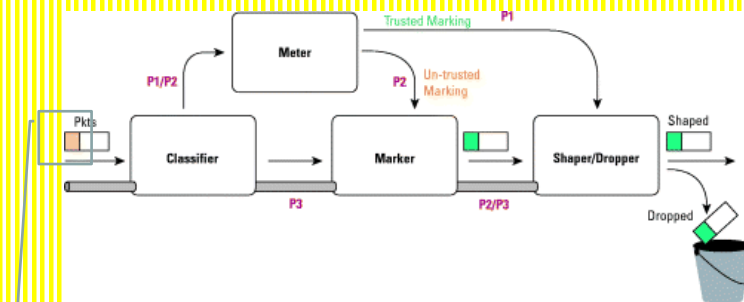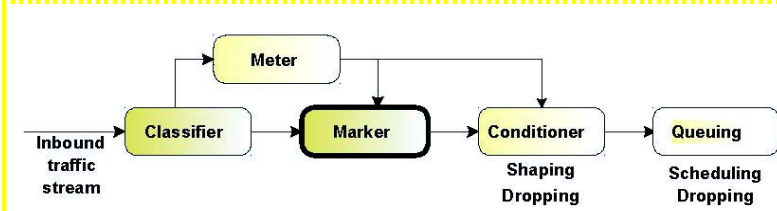
80

80

Classification

- Many traditional QoS mechanisms already include intrinsically classifiers
  - Committed Access Rate (CAR)
  - QoS policy propagations via BGP (QPPB)
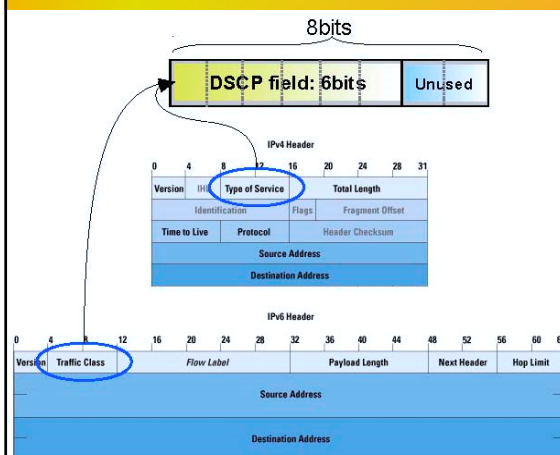 - Queuing Mechanisms
  - ...

81

81

Marking

- Marking is used to set:
  - IP precedence
  - DSCP
  - QoS group
  - MPLS experimental bits
  - IEEE 802.1Q or ISL CoS

- Marking mechanisms:
  - Comitted Access Rate (CAR)
  - QoS Policy Propagation
  - through BGP (QPPB)
  - Policy-based Routing (PBR)
  - Class-based Marking
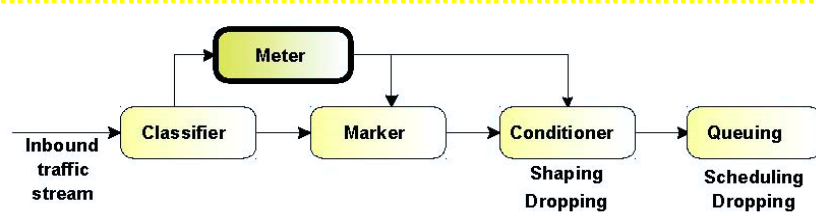
82

Edge Routers: packet marking

- Packets are marked in the field *Type of Service* (TOS) of IPv4 or *Traffic Class* in IPv6 DSCP – *Differentiated Service Code Point* CU – *Currently Unused*
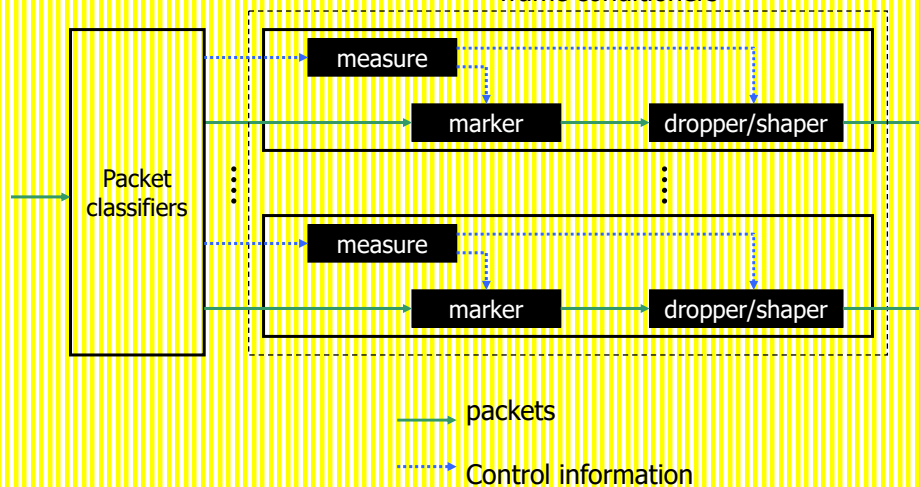
83

Metering

- Token Bucket models are used for metering
  - Committed Access Rate (CAR)
  - Generic Traffic Shaping (GTS)
  - Class-based Weighted Fair Queuing (CB-WFQ)
  - Class-based Policing
  - Class-based Shaping   84

84

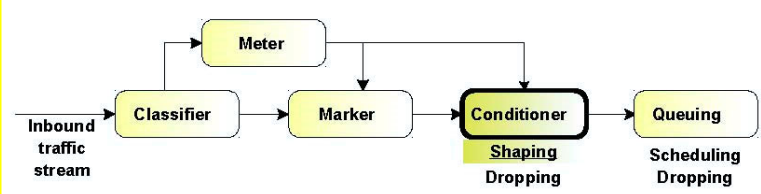# Traffic conditioning
shaping/dropping

Traffic conditioners

Packet classifiers

measure

marker

dropper/shaper

measure

marker

dropper/shaper

→ packets

⇢ Control information

85

85

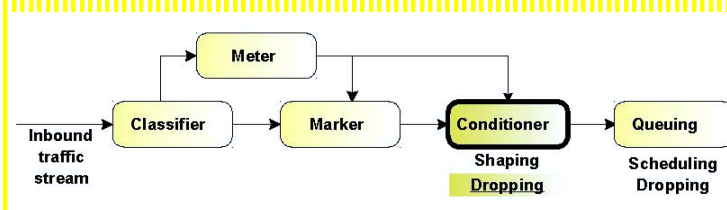# Conditioning traffic: shaping

- *Traffic Shaping mechanisms*:
  - Generic Traffic Shaping (GTS)
  - Class-based Shaping

86

86

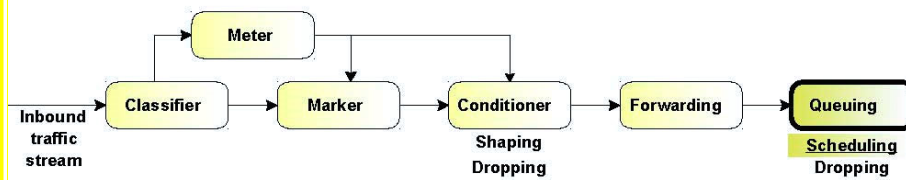# Conditioning traffic: dropping

- *Dropping* traffic:
  - Committed Access Rate (CAR) and Class-based Policing may drop packets that exceed agreed rate
  - Weighted Random Early Detection (WRED) may drop packets randomly when an interface starts to reach congestion
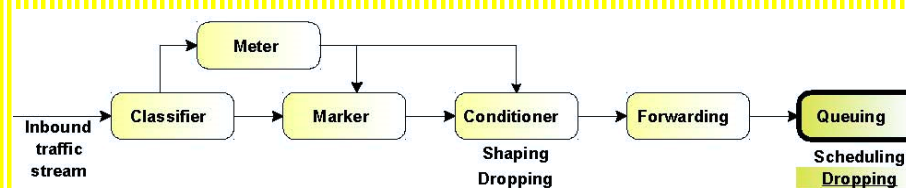
87

87

Final step: Queuing

- **Traditional queuing mechanisms**
  - FIFO, Priority Queuing (PQ), Custom Queuing (CQ)
- **Weighted Fair Queuing (WFQ) family**
  - WFQ, dWFQ, CoS-based dWFQ, QoS-group dWFQ
- **Advanced queueing mechanisms**
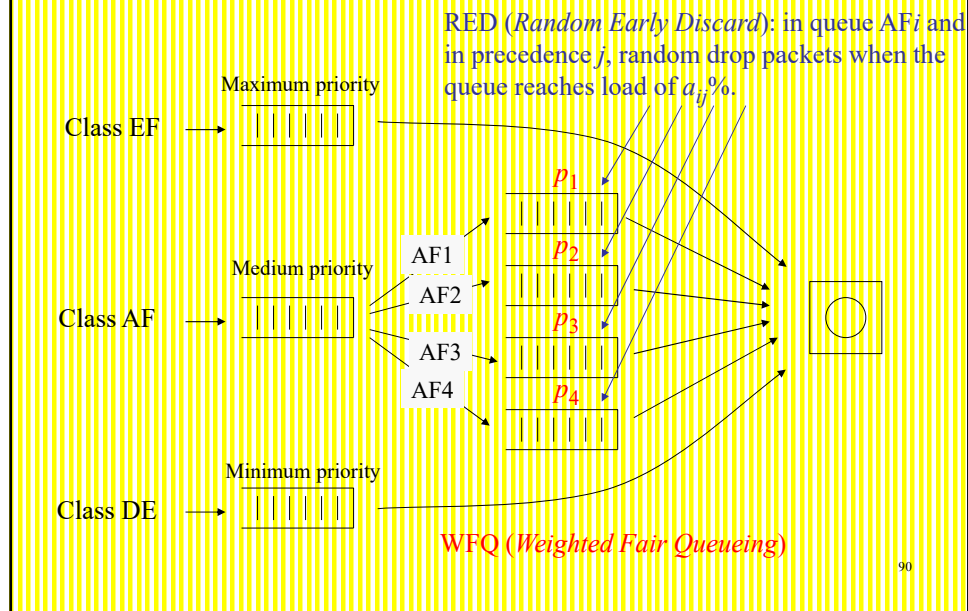 - Class-based WFQ, Class-based LLQ

88

88

Queuing

- *Dropping* mechanisms
  - *Tail drop* when there is congestion in the waiting queue
  - WFQ has an improved approach to *tail-drop*
  - Weighted Random Early Detection (WRED) may drop packets randomly when an interface starts to reach congestion

89

89

Implementation example

RED (*Random Early Discard*): in queue AFi and in precedence j, random drop packets when the queue reaches load of a_{ij}%.

Maximum priority

Class EF →

p_1

p_2

Medium priority

AF1
AF2

Class AF →

p_3

AF3
AF4

p_4

Minimum priority

Class DE →

WFQ (*Weighted Fair Queueing*)

90

90

DiffServ problems

- No standards for SLAs:
  - The same DS codepoint can be used by different services, between different ISPs
  - Different networks, using the same PHB, may provide different behaviors
  - No generalized edge-to-edge semantics (PDBs: per-domain-behavior)
- Lack of symmetry:
  - Protocols like TCP work best in nearly symmetric environments
- Multicast:
  - No support for multiparty, symmetric, communications
- Network configuration, for each PHB

91

91

Multiplexing effects

- "aggregates" means that we do not see simple flows
  - Thousands of different flows reach the same core router with the same PHB
  - Different delays $\Rightarrow$ variable traffic conditions
- $\Rightarrow$ QoS traffic percentage vs. Best Effort varies at each instant!!
- Total QoS bandwidth allocation must be carefully considered.
  - BW reservation for each QoS class must be assessed.
  - The buffer parameters for each queue has to be done separately.
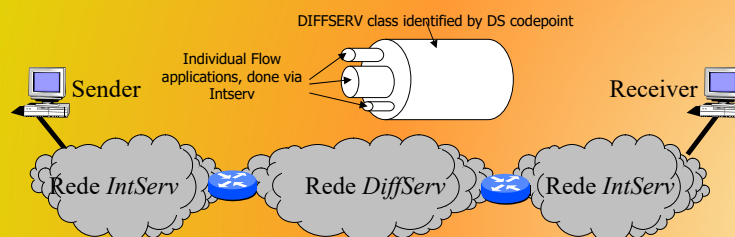 (RIO - RED with In and Out – is frequent in DS systems)

92

92

INTSERV vs DIFFSERV

93

- Complementary:
  - DIFFSERV: aggregate, per user/customer/groups/applications – oriented to the service provider
  - INTSERV: per flow – oriented to application
- One can integrate:
  - INTSERV reservations inside DIFFSERV "flows"
  - The border routers of the two types of network:
    - classify RSVP requests in the adequate DiffServ service classes.
    - If there are insufficient resources, refuse RSVP reservation requests

DIFFSERV class identified by DS codepoint

Individual Flow applications, done via Intserv

Sender

Receiver

Rede *IntServ*   Rede *DiffServ*   Rede *IntServ*

93

INTSERV and DIFFSERV

	INTSERV	DIFFSERV
signaling	By the application	Network management, application
granularity	Flow	flow, source, site (aggregation)
mechanism	Endereço de destino, protocolo e número de porto	Classe de pacotes (outros mecanismos possíveis)
scope	End-to-end	Between networks, E2E

94

94