IP Quality of Service

Jukka Manner, TKK

TELECOMMUNICATIONS SOFTWARE AND MULTIMEDIA LABORATORY

What is IP Quality of Service?

In this presentation, IP QoS means:

The technology, information and decisions used in routing IP packets in order to fulfill the expectations of clients on the quality of the IP connectivity.

TELECOMMUNICATIONS SOFTWARE AND MULTIMEDIA LABORATORY

Why do we need to talk about IP Quality of Service?

If the network had unlimited capacity, there would be no need to talk about "Quality of Service": everybody could send and receive as much as they want.

However in practice there is no such thing as unlimited capacity.

Contents

- Part I: The why and how of providing IP QoS, and what applications can do themselves
- Part II: Mechanisms for providing service to packets on a “case-by-case” basis (queuing mechanisms, filters, schedulers, …)
- Part III: (IETF) architectures that make use of these mechanisms to provide QoS to end-user applications

TELECOMMUNICATIONS SOFTWARE AND MULTIMEDIA LABORATORY

Part I: The why and how of IP QoS

Routing in the Internet

- The default service of the Internet is "best-effort"; no particular special effort is put in forwarding packets
- When a packet is received its destination is compared to the routing table and the packet is put to the queue of the output interface; only the destination address matters
- Each packet is served in the same way in each router
- This is an equal service between competing flows
- As the load varies, the network's service varies
- Any packet may be dropped when congestion occurs
- During and after congestion, senders are expected to adjust their packet sending rate

New Applications and Usage

- The load in the network is very heterogeneous
  - The main load is HTTP/FTP/email traffic based on TCP
  - Streaming multimedia applications are gaining popularity, e.g. network radio, streaming video, VoIP, etc.
  - Also, emergency calls, network control signalling, database queries for E-commerce, etc. are increasing
- Many of these need reliable and stable forwarding services, which means a need for steady transmission delay and low packet loss
- Requirements can be about bandwidths (10 Kbps up to Mbps), low packet loss, low delay, etc.
- The fundamental problem is that no dedicated circuit-switched connections exist in a packet switched network
- Thus, no dedicated resources can be allocated to single flows

How to support IP QoS

- Most simple form of IP QoS is over-provisioning
- Applications can try to adapt the data transfer to the network capacity, and congestion situation, e.g. TCP, DCCP, RTP (Note the difference between application and transport level)
- Applications can duplicate the payload, e.g.,use FEC, and hope that the receiver can rebuild the original data
- Admission control schemes can be used to prohibit new flows from entering a network:
  - Distributed algorithms base the decision on the knowledge of the load in the network
  - Explicit signaling protocols can carry requests of hosts to let the network know, e.g., the bandwidth needed by a new flow
- IP routers can forward incoming packets differently based on the sender/receiver, the type of flow, an explicit request, etc.

TELECOMMUNICATIONS SOFTWARE AND MULTIMEDIA LABORATORY

The Real-time Transport Protocol (RTP)

- With lack of QoS mechanisms in the Internet, how can audio-video applications deliver an adequate service?
- TCP is an earlier example of an adaptation mechanism:
  - When packets get lost, slow down the sending rate, since there ”must” be congestion in the network
- Ordinary multimedia flows use commonly a fixed codec and send a constant stream
- When packets get lost, the receiver notices a jerky and mismatched audio/video; delayed packet are as well “lost”
- No way for the sender to know about network congestion and change the transmission rate
- Thus, try to adapt yourself, e.g., lower the quality of the stream

The Real-time Transport Protocol (RTP)

- RTP is based on three components:
  - A family of codecs to code the multimedia flow, ranging from GSM-codecs to MP3, MPEG4 (over 40 payload formats, so far)
  - A transport protocol operating over UDP to deliver the audio/video information
  - A control protocol (RTCP) to control the media delivery and provide feedback to the sender: received/lost packets, timestamps, jitter, etc.
- RTP specifies payload formats, security, header compression, multiplexing of streams, etc.
- Codecs have different delay and bandwidth requirements
- The feedback provided to the sender allows it to change the stream properties and, thus, change its requirements from the connecting network

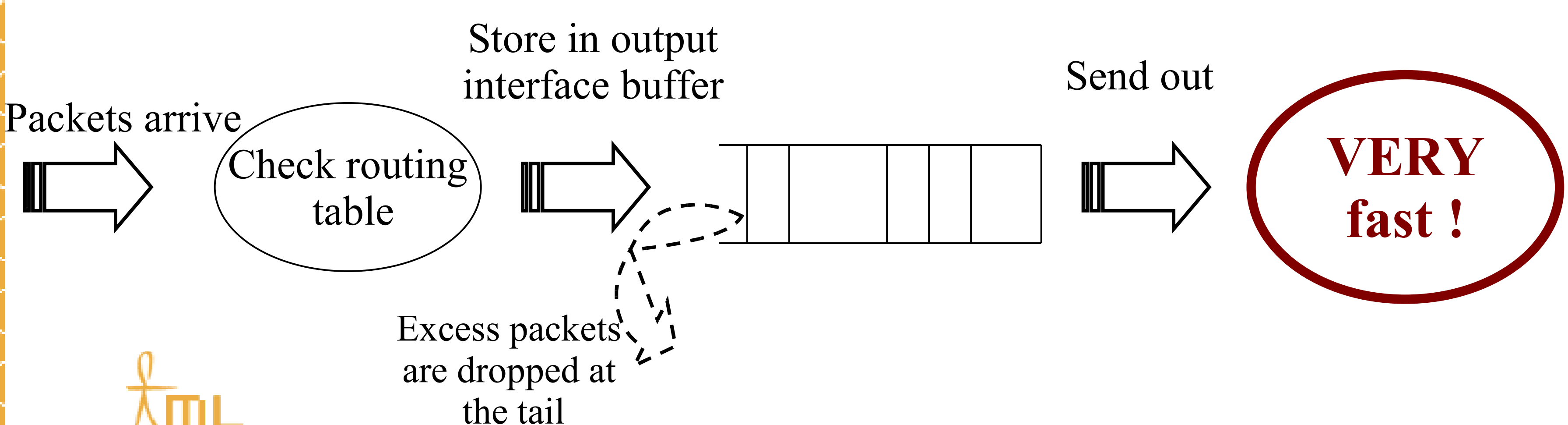
The Datagram Congestion Control Protocol

- Many applications need to send a constant stream of packets
- So far the only option has been to use UDP, which does not have flow and congestion control algorithms, or RTP/RTCP
- If the network becomes congested, TCP-based flows will react, slow down their transmission rate – UDP flows won't
- If applications want to support congestion control, they must implement their own messaging and algorithms
- The Datagram Congestion Control Protocol (DCCP) is an effort to define a "congestion-aware UDP"
    - Still an unreliable datagram transport protocol
    - Receiver sends feedback to the sender on how packets are received
    - Sender adjust the sending rate based on the feedback
    - Different congestion control algorithms can be used, currently two are defined: TCP-like and TCP-friendly algorithms
    - Still, does not really work welll with actual streaming applications

Part II: Mechanisms for IP QoS

The Simple Mechanisms for IP Routing

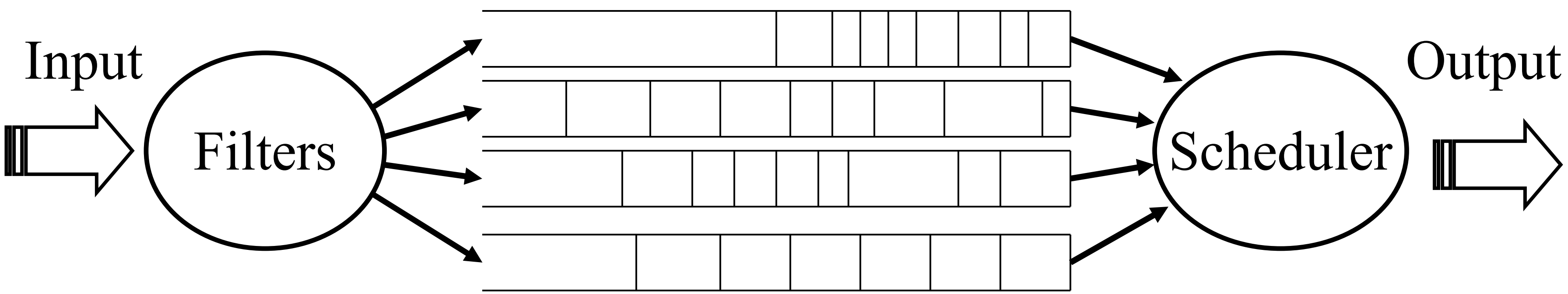
- Three components affect the packet handling in a router:
  - The routing table lookup,
  - The packet queuing mechanism, and
  - The packet scheduler at the output interface
- The most simple combination:
  - First-in-First-out scheduling and tail drop
  - Problem: cannot provide service distinction



TELECOMMUNICATIONS SOFTWARE AND MULTIMEDIA LABORATORY

Distinguishing Flows

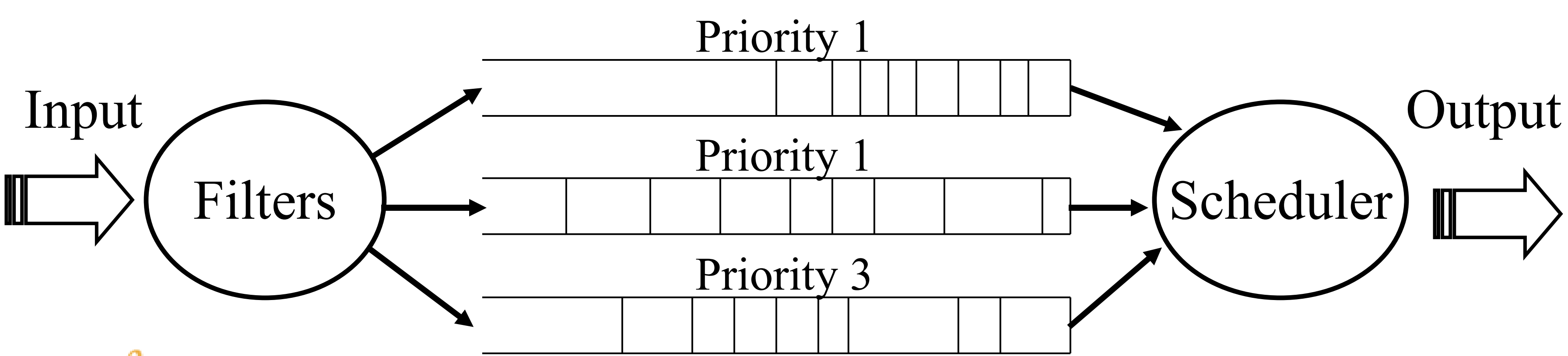
- A different service level can only be accomplished if
  - We can identify certain packet flows, and
  - Forward these separated from other flows
- In practice we need filters than can distinguish certain flows and more than one queue
- Filtering is most commonly based on protocol header fields
- Different scheduler principles can be used for the queues



TELECOMMUNICATIONS SOFTWARE AND MULTIMEDIA LABORATORY

Most Common Schedulers (PRIO)

- A straightforward scheduler uses priority queuing:
  - Packets are filtered and queued in several queues
  - Packets in highest priority queue (1) are all sent before any packets from the next highest queue (2), etc.
- Issues:
  - Very good service to the highest class of packets
  - Very unfair to lower priority queues
  - Easily leads to starvation in the lower priority queues



Most Common Schedulers (CBQ)

- A variation of priority queuing is Class Based Queuing:
  - Also called Weighted Round Robin
  - Queues are assigned relative priorities
  - During a time interval t each queue is emptied an amount relative to its priority, for example:
    - Queue 1: 8 pcks, 8/15 (53%) of link BW
    - Queue 2: 4 pcks, 4/15 (27%) of link BW
    - Queue 3: 2 pcks, 2/15 (13%) of link BW
    - Queue 4: 1 pck, 1/15 (7%) of link BW

But what if…
(60B audio x 8 = 480B)
(120B HQ audio x 4 = 480B)
(540B HQ video x 2 = 1080B)
(1500B FTP (TCP) = 1500B)

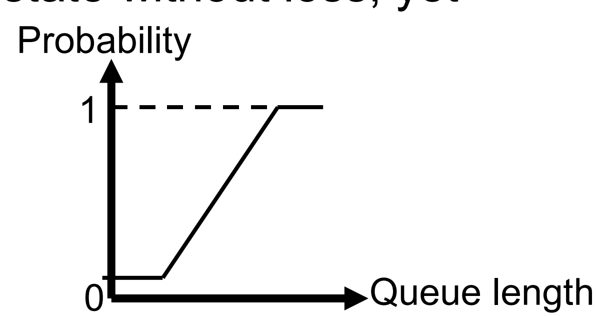
- Issues:
  - More fairness between different queues, avoids total starvation
  - Major concern is that the amount of data taken from a queue is based on packets
  - Also, setting up the queues requires many parameters (Linux: over 20)

Most Common Schedulers (WFQ)

- More complex scheduler: Weighted Fair Queuing
- Similar to CBQ but uses more information and calculations to decide which packets to forward next
- Estimates each packets time to be sent during a round and sends packets in that order
- Issues:
  - Accurate link sharing based on weights
  - Provides high level of fairness to different classes
  - Bounds delay (Parekh-Gallager) but not jitter
  - Consumes much more cpu cycles to calculate the next packet to handle
  - May lower the maximum throughput of the link
  - Should not be used if there are numerous queues

Queuing Strategies

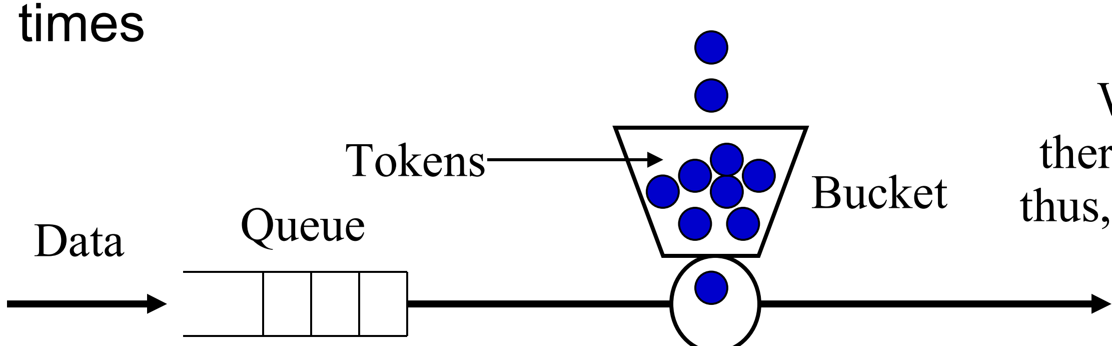
- Most typical is tail-drop: when congestion occurs, buffers are full and arriving packets are (all) dropped
- Should take precautions to not run out of buffer space
- Various active queue management mechanisms can be used to drop specific packets, not always the last arrived
- Random Early Detection (RED) is one solution:
  - Packet drop is based on queue length: larger queue, higher probability of an incoming packet to get dropped
  - RED is fair: probability of a flow’s packet to be lost is proportional to its share of link bandwidth
  - Can also mark packets instead of dropping them (IP ECN bits), allowing sources to detect network state without loss, yet
  - Numerous variations of RED



TELECOMMUNICATIONS SOFTWARE AND MULTIMEDIA LABORATORY

Additional Mechanisms

- Leaky bucket and token bucket filters
- Data packets ”leak” from a bucket of depth ***n*** with a rate ***R***
- Token bucket filter (TBF) adds “tokens” and a “bucket”:
  - Tokens are provided at a fixed rate ***R***
  - The bucket can accommodate some tokens for future use allowing bursts of packets
  - Still, the filter only allows a maximum rate ***R*** to flow through
- A number of derivate of WFQ and CBQ exist (Linux: HTB)
- Also pricing can be used to affect load in classes and at peak times



With leaky bucket,
there is no actual bucket,
thus, no bursts are possible

TELECOMMUNICATIONS SOFTWARE AND MULTIMEDIA LABORATORY

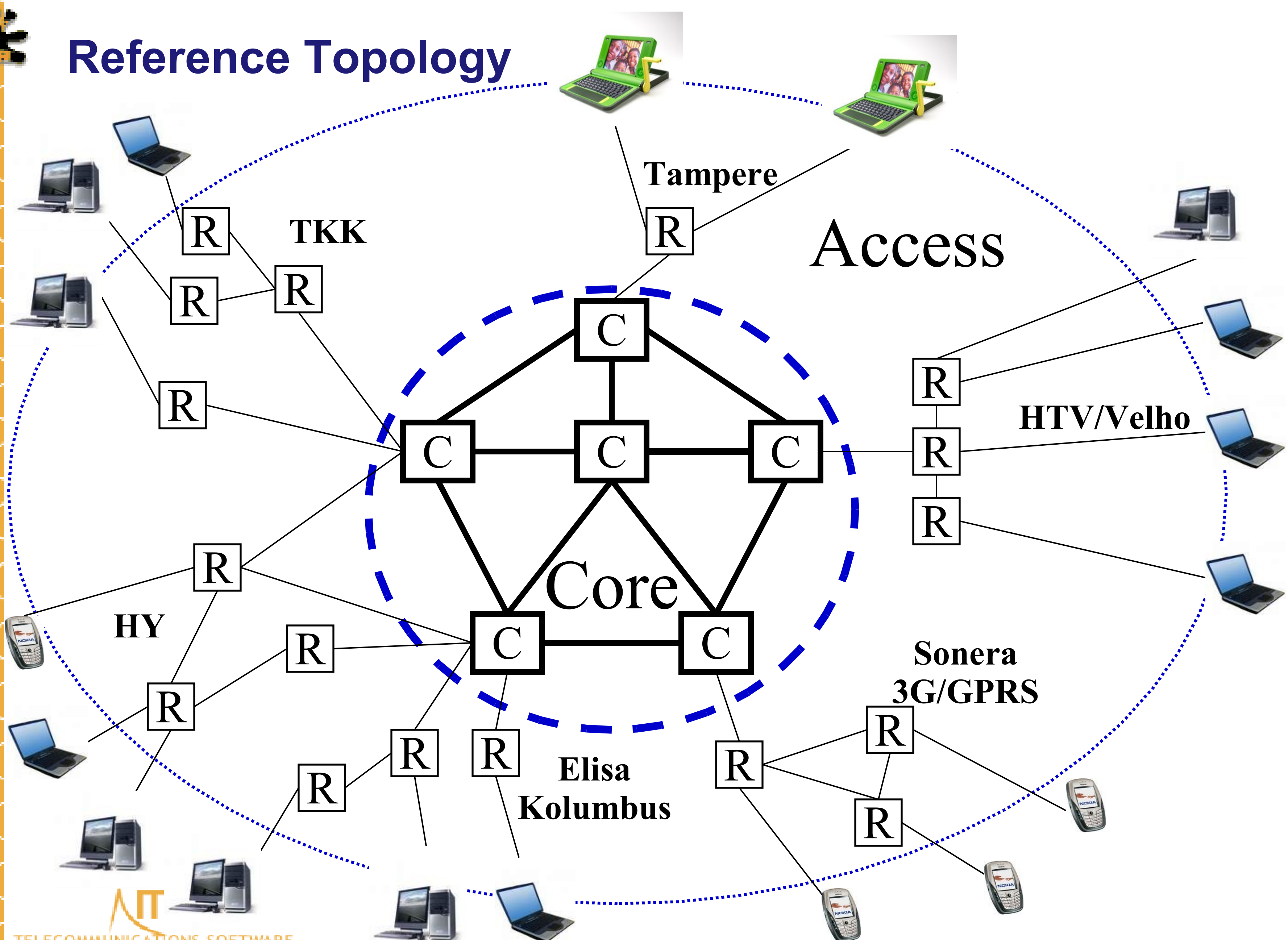
QoS Routing

- The default router forwarding does not take into account resource availability
- With a broader collection of links, we would need some way of finding various paths that would support the traffic
- Also in view of network efficiency, alternate QoS-enabled paths would be beneficial
- The routing protocol needs to include information about "resources", and
- The routing decision in a router needs to exploit this information to find the "best" next hop for each packet
- Still, must not often shift directions of individual flows: if the characteristics of the path change, transport protocols react

Where to Implement QoS?

- Link speeds faster towards the core; more traffic
- Access routers generally do not have to handle high packet switching rates and, thus, can do complex traffic filtering, classification and policing
- The overhead of implementing QoS in the core would affect a large amount of traffic and lower the throughput
- Thus, access networks could do per-flow packet handling, and
- Backbones forward packets in packet aggregates, a limited number of classes (or even FIFO)
- Note that the quality of the end-to-end connection is as good as the weakest link

Reference Topology



Part III: QoS Architectures

Objectives of QoS Architectures

- To control the network service response such that:
  - The response to a specific service is consistent and predictable
  - A client is provided with a level equal to or above a guaranteed minimum
- To allow a client to request in advance a service response
- To control the contention for network resources such that:
  - A client can be provided with a superior level of service
  - A client does not obtain unfair allocation of resources
- To allow for efficient total utilisation of network resources, while providing a range of different network services
- Quality can be bandwidth (min, max, average), delay, jitter, packet loss rate, etc.

IETF QoS Architectures

- Numerous proprietary architectures exist within the academic community and the telecommunications industry, e.g. YESSIR, INSIGNIA, RMD, Mobile RSVP, Localized RSVP
- IETF has two main, although very different, standards:
  - The Integrated Services architecture (IntServ)
  - The Differentiated Services architecture (DiffServ)
- QoS-related protocols:
  - RTP provides an adaptive multimedia transport protocol, that adjust to the properties of the underlying network
  - Multiprotocol Label Switching (MPLS) can be used to create dedicated links in a packet-switched network and provide a QoS-aware service

TELECOMMUNICATIONS SOFTWARE AND MULTIMEDIA LABORATORY

Two Architectures: IntServ

- The IntServ model includes two sorts of services targeted towards real-time traffic: guaranteed and predictive service
- Resource reservation and admission control are key issues
- The term "guarantee" is used in a broad sense:
  - Absolute or statistical,
  - Strict or approximate
- Source describes its desired flow rate and sends this information to the routers and the receiver
- Network admits requests and reserves resources
- Source must send at this rate (controlled by network)
- Provides a sort of “dedicated” connection within an IP packet-switched network
- Reservation of resources are usually done with the Resource Reservation Protocol (RSVP)

IntServ: Guaranteed Service (GS)

- Aims to control the maximal queuing delay
- A flow is described using a Traffic Specification (TSpec): average and burst rates, average and max packet size, etc.
- Routers compute various parameters describing how it will handle the flow's data
- By combining the parameters from the routers in a path, it is possible to compute the maximum delay a packet will experience when transmitted via that path

IntServ: Controlled-Load Service (CL)

- Goals of the service:
  - High percentage of packets successfully delivered to the receiving end-nodes
  - The transit delay experienced by a high percentage of the packets will not greatly exceed the minimum transmit delay experienced by any delivered packet
- Clients requesting controlled-load service provide a TSpec of the data traffic they will generate
- The controlled-load service does not use specific parameters such as delay or loss
- An accepted request for CL service implies a commitment by the network to provide service closely equivalent to *uncontrolled (best-effort) traffic under lightly loaded conditions*

TELECOMMUNICATIONS SOFTWARE AND MULTIMEDIA LABORATORY

Resource Reservation Protocol (RSVP)

- Used by a host to request specific qualities of service from the network routers for application packet flows
- Makes resource reservations for both unicast and multicast applications, adapts to group memberships and routes
- Receiver-oriented: receiver of a flow initiates and maintains the resource reservation used for that flow
- Sender of the data flow still informs the receiver about the traffic characteristics, to allow for a proper reservation from the network
- RSVP maintains soft state in routers and hosts; reservations can be modified and are not stored "forever" in routers

Resource Reservation Protocol (RSVP)

- RSVP makes reservations for unidirectional data flows
- It does not transport application data
- The traffic control and policy control parameters are opaque to RSVP
- RSVP provides several reservation models or "styles" to fit a variety of applications (shared and dedicated reservations)
- RSVP provides transparent operation through routers that do not support it: RSVP packets are just normal IP packets
- RSVP is not a routing protocol but depends upon present and future routing protocols
- RSVP supports both IPv4 and IPv6

IntServ & RSVP: Issues

- "All or nothing" -service
- Provides near absolute guarantees of QoS if the request is accepted (unless routers crash, mobile nodes move…)
- Need for state information, processing overhead and (sometimes heavy) signalling are the major concerns
- To be useful, all nodes on the path should support RSVP
- RSVP is also very complex because of the support for multicasting; today unicast would be mostly enough
- RSVP signalling in mobile networks is difficult
- RSVP, and other QoS protocols, have been analyzed by the IETF NSIS working group (RFC will be published very soon)

TELECOMMUNICATIONS SOFTWARE AND MULTIMEDIA LABORATORY

Two Architectures: DiffServ

- No signalling, no reservations
- No feedback about the resource availability
- Traffic entering a network is marked with a code to differentiate a certain packet/flow from others
- Uses the old 8-bit IP TOS-field, 6 bits (2 bits used by ECN)
- Different codes result in different service at routers
  - Premium service, priority-based service, best-effort
- Does not provide absolute guarantees to flows
- Scales well to large numbers of users
- “All, or something, maybe” -service

TELECOMMUNICATIONS SOFTWARE AND MULTIMEDIA LABORATORY

Two Architectures: DiffServ (cont.)

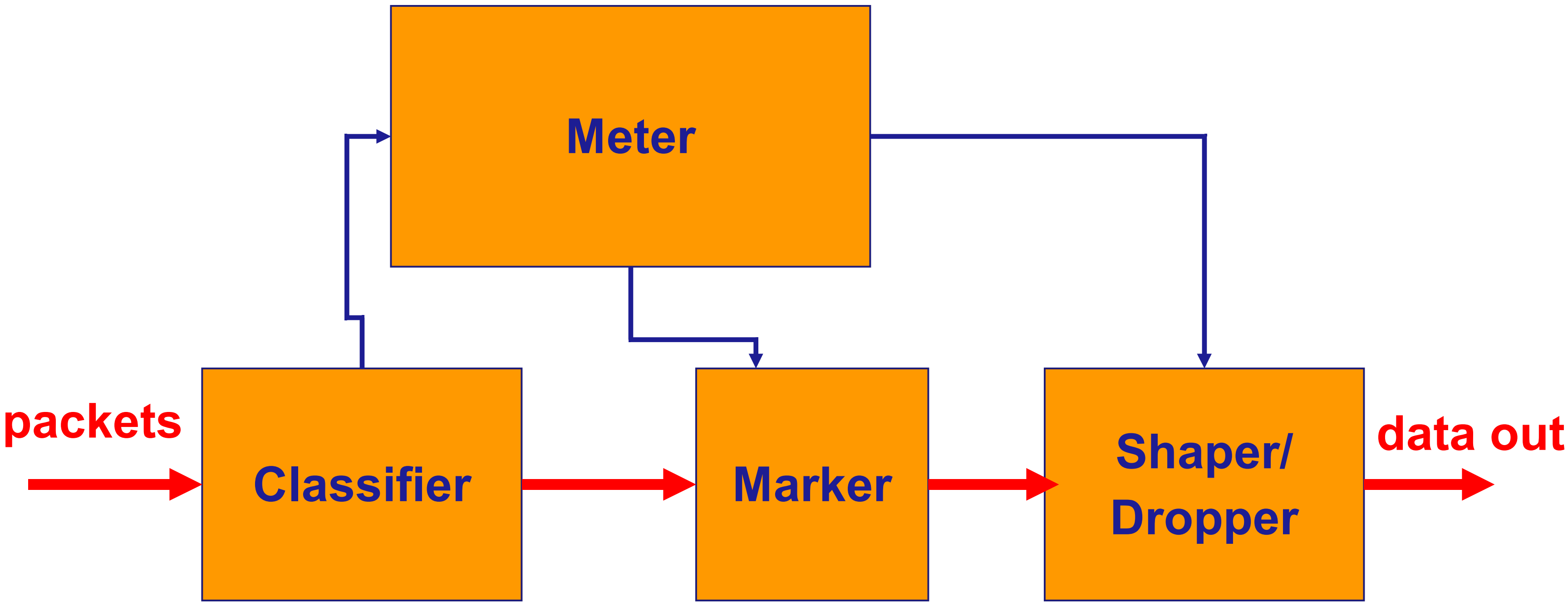
- Architecture for implementing service classes in the Internet, usually controlled by agreements between ISPs
- Architecture is composed
  - Per-hop forwarding behaviours (=classes),
  - Packet classification functions,
  - Traffic conditioning functions including metering, marking, shaping, policing, and
  - Service Level Agreements (SLA) between data senders and the forwarding network operator
- Per-Hop Behaviours (PHB) define the DS field handling (per link or “hop”) and service outcome of classes marked using the DS field in the IP header

Two Architectures: DiffServ (cont.)

- Complex classification and conditioning functions only at network boundary nodes (and possibly at sending host)
- Service differentiation separately per direction
- Two types of services standardized, operators may define other services and code points
- Does not necessarily provide end-to-end QoS:
  - Operators may have different meanings and implementations for classes and code points,
  - The code points can change, thus, may not remain the same on the whole end-to-end path.

TELECOMMUNICATIONS SOFTWARE AND MULTIMEDIA LABORATORY

Two Architectures: DiffServ (cont.)



TELECOMMUNICATIONS SOFTWARE
AND MULTIMEDIA LABORATORY

DiffServ: Expedited Forwarding (EF)

- ”A Premium service”
- To build a low loss, low latency, low jitter, assured bandwidth, end-to-end service through DiffServ domains
- Mimics "virtual leased line”
- The aggregate has a well-defined minimum departure rate
- The standard does not define the implementation, just the nature of service and how to achieve it

TELECOMMUNICATIONS SOFTWARE AND MULTIMEDIA LABORATORY

DiffServ: Assured Forwarding (AF)

- A “classic” priority-based service
- The AF PHB group provides four independent classes
- Each class has a certain share of the forwarding resources
- Within each AF class, an IP packet can be assigned one of three different levels of drop precedence
- A DiffServ node does not reorder IP packets of the same flow if they belong to the same AF class
- In case of congestion, the drop precedence of a packet determines the relative importance of the packet within the AF class

Stateless vs. Stateful Operation

- In IntServ the network must maintain the state of flows
- In DiffServ there is no need for states, the DSCP is enough
- Also the application does not give information of its traffic
- IntServ gives a fine level of granularity, while DiffServ is more approximate in its service outcome
- In IntServ, scalability and router performance are of concern
- DiffServ does not inform the client of the outcome of the service, applications need to monitor the service
- Aggregate RSVP, RSVP DCLASS Object, IntServ over DiffServ, and RSVP proxies are latest efforts to find better solutions

A Look into the Future

- The challenge is addressing the weaknesses of the two main architectures and integrate them (RFC2998):
  - IntServ/RSVP would provide the service signalling
  - DiffServ would be used for the core
  - RSVP allows applications to signal their needs to the network
  - RSVP requests are mapped to DSCPs for access admission and forwarding within the network
- The Next Steps in Signalling (NSIS) IETF WG is looking at new ways to do QoS signalling (ready soon)
- Another approach is to make applications and transport protocols behave better (e.g. DCCP)
- Lately, people have raised issues about TCP and "friendliness"
- A fundamental question remains:
  - What is the business model for QoS?
  - Who/when/why/how pays for a differentiated service?

Pointers

- IETF Web site: http://www.ietf.org/ (see working groups IntServ, DiffServ, RSVP, ISSLL, AVT (for RTP), NSIS, DCCP)
  - IntServ and RSVP: RFCs 2205, 2210, 2211, 2212, 2961
  - DiffServ: RFC 2474, 2475, 2597, 3246, 3247, 3248
  - IntServ and DiffServ together: 2990, 2998
  - NSIS: 4080, 4094,
  - RTP: RFC 3550
  - QoS Routing: RFC 2386
- RFCs at http://www.ietf.org/rfc/rfcXXXX.txt
- Also for RTP: http://www.cs.columbia.edu/~hgs/rtp/
- Very good book is (a little outdated, though):
  Geoff Huston: Internet Performance Survival Guide, QoS Strategies for Multiservice Networks, Wiley, 2000.

TELECOMMUNICATIONS SOFTWARE AND MULTIMEDIA LABORATORY