Quality of service in Internet

Lecture 9

Quality of service in networking

- We say that a network provides Quality of Service (QoS) if the users are guaranteed that, in certain conditions, their quality of service requirements will be satisfied by the network.
- QoS is expressed by measurable parameters like:
  - Data delay (latency)
  - Data delay variation, called *jitter*
    - Jitter is important in streaming applications (audio and video streaming)
  - Network throughput
  - Reliability (expressed by the percentage of data packets lost, received with errors, duplicated, received out of sequence, etc).
- The guarantees can be absolute (seldom), or, more often, statistical:
  - e.g. 95% of data packets will have a delay < … seconds

Quality of service in networking

- If there are no QoS guarantees, then the network is called **best effort** (it does its best !)
- Besides QoS, important is also **QoS differentiation**, i.e.:
  - The network can give to different (categories of) users different QoS guarantees, according to, e.g., their subscription, or according to their application.
  - Or, the guarantees are given to a category of users compared with other category of users (e.g. premium users will have a latency smaller, or, approx. half of the latency of standard users)
- QoS classes (categories) are defined
  - Users will have a priority, or precedence, parameter
- The priority can be based on users' subscription and/or on the application
- QoS differentiation can be very important, e.g., in case of congestion, when the network cannot deliver all data, but it will try to deliver the data belonging to higher priority users and to drop data only from lower priority users
- Usually it is easier to provide QoS differentiation than (absolute) QoS guarantees

QoS in networking: real-time applications

- Usually, real-time (*rt*) applications have more demanding QoS requests, especially in what concerns the delay.
- Example of real-time applications:
  - Remote login or telnet sessions
  - On-line games
  - Banking or stock exchange applications
  - Video conferences
  - Audio and video streaming
  - Voice (VoIP : voice over IP), …
- Some rt applications can tolerate the loss of a small amount of data packets:
  - E.g., the VoIP application: the human ear cannot tolerate a delay bigger than a certain value, but it does tolerate the loss of some data packets
  - But, stock exchange applications do not tolerate data loss !

Non real-time applications

- Example of non real-time (*nrt*) applications:
  - e-mail
  - File transfer (FTP)
  - Web browsing (WWW)
- In general, the *nrt* applications have less demanding requirements concerning data delay, but higher demands concerning data reliability
- There are differences between *nrt* applications: (e.g. the delay for www is more important that for e-mail)

Historical perspective

- The Internet was developed as a best-effort data network:
  - First aplications were file transfer and e-mail between the scientists working (on military programs) at different universities across US
- The Internet is based on packet switching (PS), which is not very suited (efficient) for QoS *rt* applications.
- In time, the idea of transmitting *rt* applications like voice, tv, etc, on the same network has gained importance
- Traditionally, voice was transmitted over Circuit Switching (CS) networks (like fixed telephony)

ATM

- In the 1970's have been proposed the Asynchronous Transfer Mode (ATM) networks
- They have been the only networks designed for QoS !
- They are based on virtual circuit switching, which is a compromise between PS and CS:
  - The same like at CS, circuits are established between source and destination, but, in the circuits' nodes, the resources are shared between several connections (like in PS)
  - Data packets, named cells (not to be confused with the cells from cellular telephony), have a fixed length of 53 octets: 5 octets header + 48 de octets payload
- ATM networks aimed to transmit rt and nrt traffic, integrating voce, data, TV, etc
- **Aimed to replace the Internet (huge mistake !)**
- The ATM networks didn't replace the Internet, having now a much limited role.
- However, ideas from ATM have been applied to Internet, in order to obtain QoS over the Internet.

QoS in Internet: IntServ and DiffServ

- IETF (Internet Engineering Task Force), the Internet standardization body, proposed first the ISA standard (Integrated Services Architecture), also known as Integrated Services (IntServ)
- Because of its high complexity and because IntServ doesn't scale well when the number of data flows is very high, the DiffServ architecture (Differentiated Services) has been proposed
- **IntServ and DiffServ will be presented after William Stallings, Data and Computer Communications, , 8th edition, Pearson Prentice Hall,2007 [Sta07], ISBN 0-13-243310-9, in the sense that ideas, text, images, are taken from [Sta07], chapter 19.**

IntServ. Traffic types

- The traffic can be:
  - Elastic
  - Inelastic
- Elastic traffic:
  - Can easily adapt to large variations of delay and throughput, continuing to fulfill the requirements of the application
  - It is the traditional traffic type in Internet
  - The applications that generate such traffic typically use TCP or UDP at the transport level
    - UDP: the application will use as much network capacity as it is available, up to the maximum rate at which the application can generate data
    - TCP: the application will use as much network capacity as it is available, up to the maximum rate at which the receiver of the end-to-end connection can receive data.

Elastic traffic

- Elastic traffic includes:
  - E-mail (the SMTP protocol): not sensitive to delay changes
  - File transfer (FTP): sensitive to changes in throughput, since the user expects the duration of the transfer to be proportional to the files' lenght
  - Interactive traffic: remote login (TELNET) and Web access (HTTP) – delay sensitive
  - Network management (SNMP): delay sensitive only in case of congestion
- The QoS perceived by the users doesn't concern the delay of a data (IP) packet, but to the delay of an element of the application: a keystroke or a line for Telnet, a file for FTP, a Web page at HTTP (but a web page can be very small or very large, if it contains many images)
- For small elements, the delay is given by the delay on the internet, while for large elements, the total elapsed time is given by "the sliding-window performance of TCP" [Sta07]
- The importance of QoS is obvious even for the elastic traffic

Inelastic traffic

- Definition: “Inelastic traffic does not easily adapt, if at all, to changes in delay and throughput across an internet.” [Sta07]
- Typical example: real-time traffic (rt)
- The requirements for the inelastic traffic may be:
  - Throughput: needs a minimum throughput in order to work
  - Delay: example of delay-sensitive application: stock trading
  - Jitter. Jitter = delay variation – it is a critical parameter for some rt application (streaming, videoconferences)
    - Buffers are used in order to compensate the jitter (packets are buffered and then delivered at a constant rate to the software application), but the buffer size should be limited, depending on the application
  - Packet loss: many rt applications tolerate a certain packet loss rate
- Requirements introduced by the inelastic traffic :
  - To give preferential treatment (to allocate more resources) to the applications with higher requirements
  - The inelastic traffic does not reduce the transfer rate in case of congestion => needs to reserve resources for it

ISA Approach

- The purpose of ISA: to provide QoS in the IP-based networks
- The central idea: how to allocate the existing network capacity in case of congestion
- The mechanisms for congestion control used in routers before ISA:
  - Routing algorithms: most routing algorithms try to minimize the delay, ensuring in this way a load balancing
  - Packet discard: when the buffers of the router are full, the packets that arrived most recently are dropped => the corresponding TCP connections will reduce their data generation rate => the congestion is reduced
  - These mechanisms are not sufficient for the inelastic traffic

ISA approach

- Each packet can be associated with a flow
- Flow = “a distinguishable stream of related IP packets that results from a single user activity and requires the same QoS” (RFC 1633, from [Sta07])
- The differences between flow and TCP connection:
  - The flow is unidirectional
  - The flow can have multiple destinations (multicast)
- An IP packet is identified as a member of a flow based on its source and destination IP addresses

ISA functions for congestion management

- Admission control (AC):
  - a flow has to reserve resources
  - If the routers don't have sufficient resources in order to guarantee the requested QoS, (without degrading the QoS for the existing flows), then the flow is not admitted (it is rejected)
  - For reservations it is used the RSVP protocol (resource ReSerVation Protocol)
- Routing algorithms:
  - Routing can be done based on different QoS parameters, not only on delay
- Queueing disciplines:
  - Describe how the resources are allocated to queues (scheduling)
- Discard policy:
  - What packets are discarded (or dropped) from queues in case of congestion

ISA architecture [Sta07]

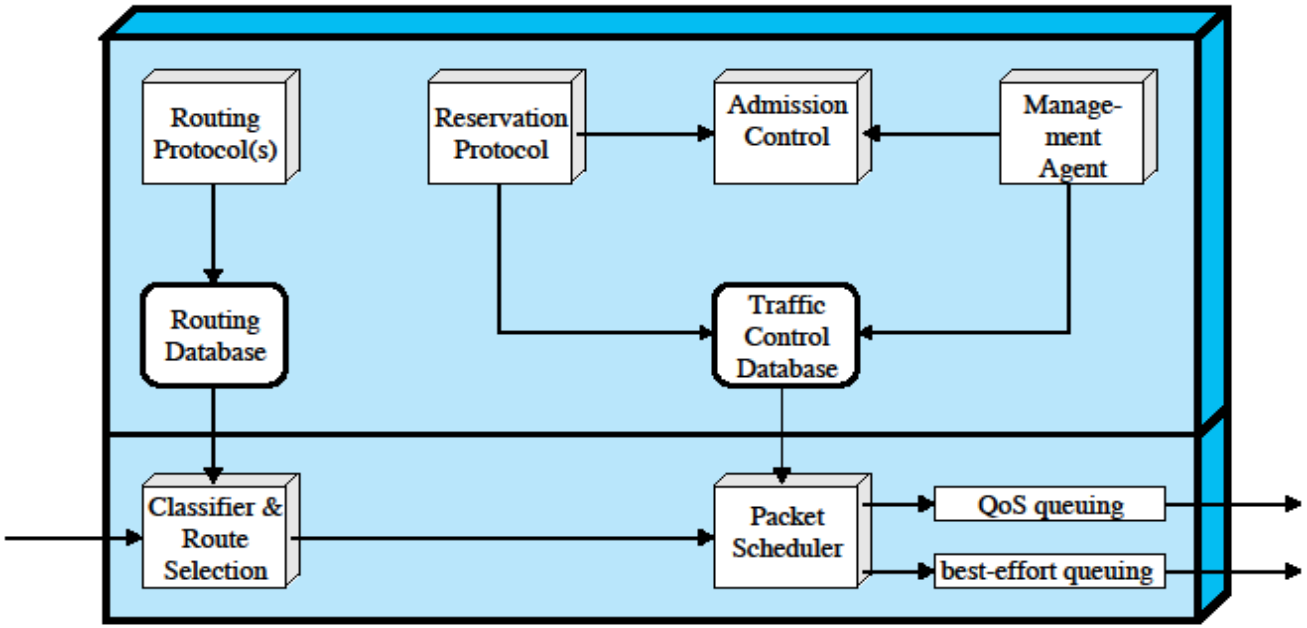


Figure 19.10 Integrated Services Architecture Implemented in Router

Components of ISA

- In figure 19.10, the functions from below the horizontal line are applied to every packet, hence, they must be optimized. They are called forwarding functions
- Functions above the line are called background functions:
  - They create data structures used by the forwarding functions
  - Reservation protocol: reserves resources for a flow and **maintain state information for each flow in routers and end systems (this generates scaling problems !)**
  - AC: when a new flow appears, the AC function is invocated. It determines if the existing resources are sufficient for the new flow, such that its QoS requirements are met, without negatively impacting the QoS of the existing flows
  - Management agent: it can modify the AC policies
  - Routing protocol: maintains a routing database, which gives the next hop for every destination address of each flow

Forwarding functions

- Classifier and route selection:
  - Packets are mapped to classes
  - A class can correspond to a single flow or to a set of flows that have the same QoS requirements(e.g: all flows that contain videostreaming, or all flows that belong to a certain organization)
  - Selection of a class is done based on some fields from the header of the IP packets
- Packet scheduler:
  - Does the management of one or several queues for the same output
  - Includes the policing function: determines if a flow exceeds the negotiated transmission rate and decides what measures to take in such a case

ISA services

- Are defined at two levels:
  1. General categories of services, each of them providing certain service guarantees.
  2. Within each category, the service for a certain flow is specified by the values of certain parameters. Together, these parameters form Tspec (traffic specification)
- When a flow makes a reservation, Tspec defines the exact amount of service requested.
  - If the reservation is accepted, the service commits to provide the requested QoS as long as the traffic flow is according to the description fromT spec.
  - The traffic can be specified using token bucket.

Token bucket [Sta07]

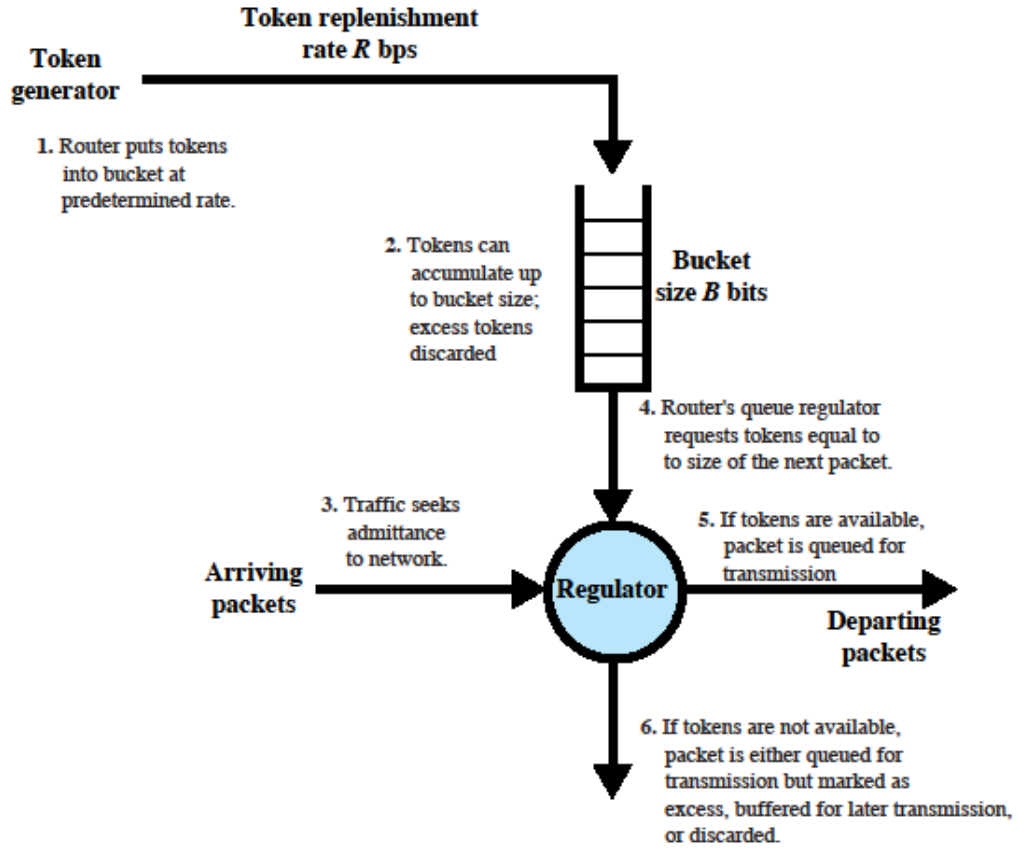


Figure 19.11 Token Bucket Scheme

Token bucket

- The token generator creates tokens with a rate R
- R = the continually sustainable data rate: it can be sustained a relatively long time for that flow.
- If tokens are not consumed, they are stored in the bucket up to the bucket capacity of B bits.
- B is the burstiness: the amount of data that can be generated in excess for a short period of time
- An IP packet is queued for later transmission only if the amount of tokens in the bucket is greater or equal than the packet size; in this case from the bucket there are removed the tokens corresponding to the packet size.
- If there are not enough tokens in the bucket, the packet is subject to a policing action: the packet can be either discarded, delayed, degraded to “best effort”, or marked as excess, in order to be discarded later, if necessary.
- In this way it is guaranteed that, during any period of time T, the maximum amount of data that can be transmitted by that flow cannot exceed the value $R \cdot T+B$.

Categories of ISA services

- Currently, three categories of services are available:
  1. Guaranteed
  2. Controlled load
  3. Best effort
- Guaranteed service:
  - It is specified an upper limit for the queueing delay, which has to be added to the propagation delay (or latency) in order to obtain the total delay of the packets
  - There are no queueing losses due to buffer overflow, but there can be packet losses due to other causes (the fall of some network elements or the change of the routing paths)
  - It is the most demanding service offered by ISA.

Categories of ISA services

- Controlled load:
  - Approximates the service received by the best effort applications under low network load.
  - No upper bound is specified for delay, however, the service guarantees that a very high percentage of the packets will not have delays greater than the transit time in the network (propagation delay + processing delay in routers)
  - Packet losses in queues will be near zero.
  - It is suitable for adaptive real-time applications.
- Best effort
  - The network does its best, but there are no QoS guarantees
- RSVP: the protocol for resource reservation in ISA.

Differentiated Services

- Problems of IntServ (ISA):
  - High complexity
  - Big amount of control signaling => the solution may not scale well for large volume of traffic
  - Routers maintain state information => problem for large volumes of traffic (don’t scale well)
- However, there is an immediate need to ensure different levels of QoS to different traffic flows =>
- Differentiated Services, or DiffServ (DS) architecture, RFC 2475:
  - Easy to implement, low cost, low-overhead

Characteristics of DiffServ

- Several key characteristics of DS contribute to its efficiency and ease of deployment:
  - • IP packets are labeled for differing QoS treatment using the existing IPv4 or IPv6 DS field.Thus, no change is required to IP.
  - • A service level agreement (SLA) is established between the service provider (internet domain) and the customer prior to the use of DS. This avoids the need to incorporate DS mechanisms in applications. Thus, existing applications need not be modified to use DS.
  - • DS provides a built-in aggregation mechanism. All traffic with the same DS octet is treated the same by the network service. For example, multiple voice connections are not handled individually but in the aggregate. This provides for good scaling to larger networks and traffic loads.
  - • DS is implemented in individual routers by queuing and forwarding packets based on the DS octet. Routers deal with each packet individually and do not have to save state information on packet flows.

DS domain

- DS domain = a contiguous portion of Internet over which a consistent set of DS policies are administrated.
- Services offered in a DS domain are defined by an Service Level Agreement (SLA), which is a contract between the customer and the service provider (the internet domain)
- SLA specifies the forwarding service that the client should receive for different classes of packets.
- The customer indicates the class of every IP packet through the DS field, marked in the header of the IP packet
- The service provider must configure in each router the forwarding policies needed in order to fulfill the contract and must measure the performance of each class of packets.
- If the destination of the packets is outside the DS domain, then the domain will try to send the packets to other domains, requesting from these domains services as close as possible to the service agreed with the customer in the SLA.
- The DS field is called DS codepoint (DSCP) and contains 6 bits.

DS field [Sta07]

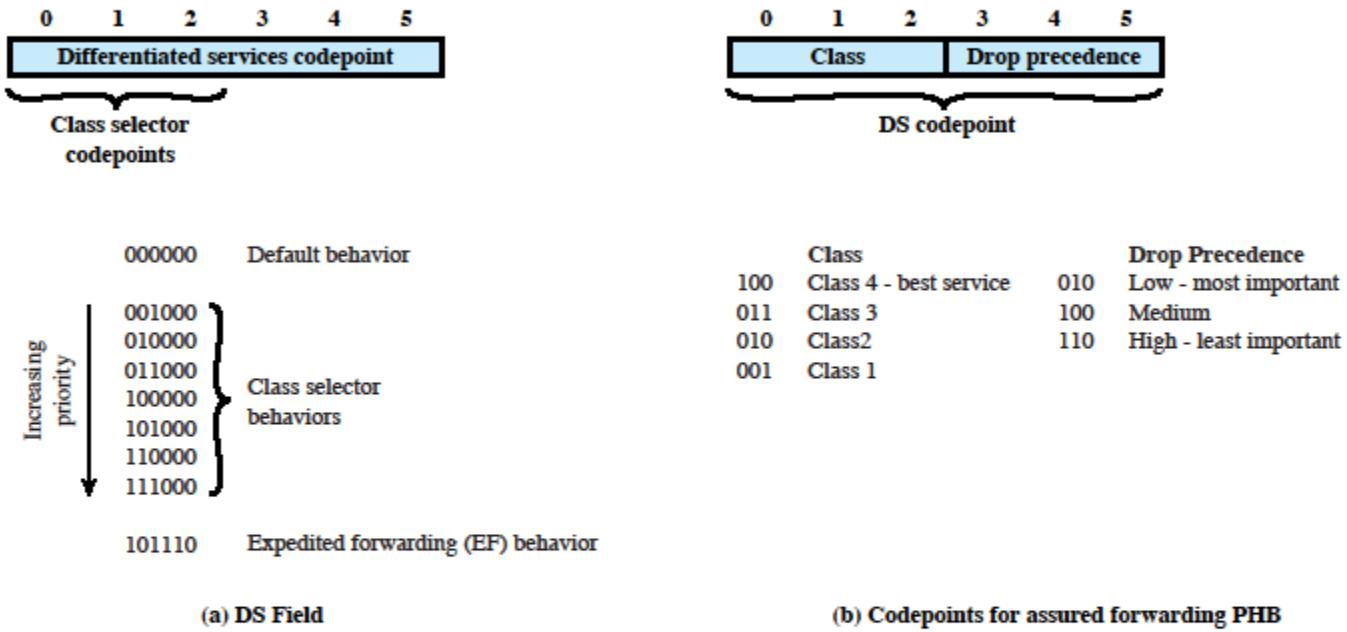


Figure 19.13 DS Field

DS configuration and operation [Sta07]

- A DS domain consist of a set of contiguous routers (it is possible to go from any router in the domain to any other router in the domain without going through routers from outside the domain).
- Routers in a DS domain may be:
  - either boundary nodes, or
  - interior nodes
- The forwarding treatment provided by a router is called Per Hop Behaviour (PHB).
- PHB must be available in all routers, but typically the interior routers offer only PHB
- Border nodes perform, besides PHB, other, more complex, functions : classifier, meter, marker, shaper and dropper.
- Classifier: separates the packets according to their class (based on DSCP)
- Meter: verifies if the packet is conform with the negotiated TSpec. If not, the packet is applied one of the functions: marking, delaying, shaping or dropping
- Marker: re-marks the packets, either because they exceed their profile, or at the boundary between two DS domains (e.g. highest priority can have the value 3 in one domain, and 7 in the other)

DS configuration and operation [Sta07]

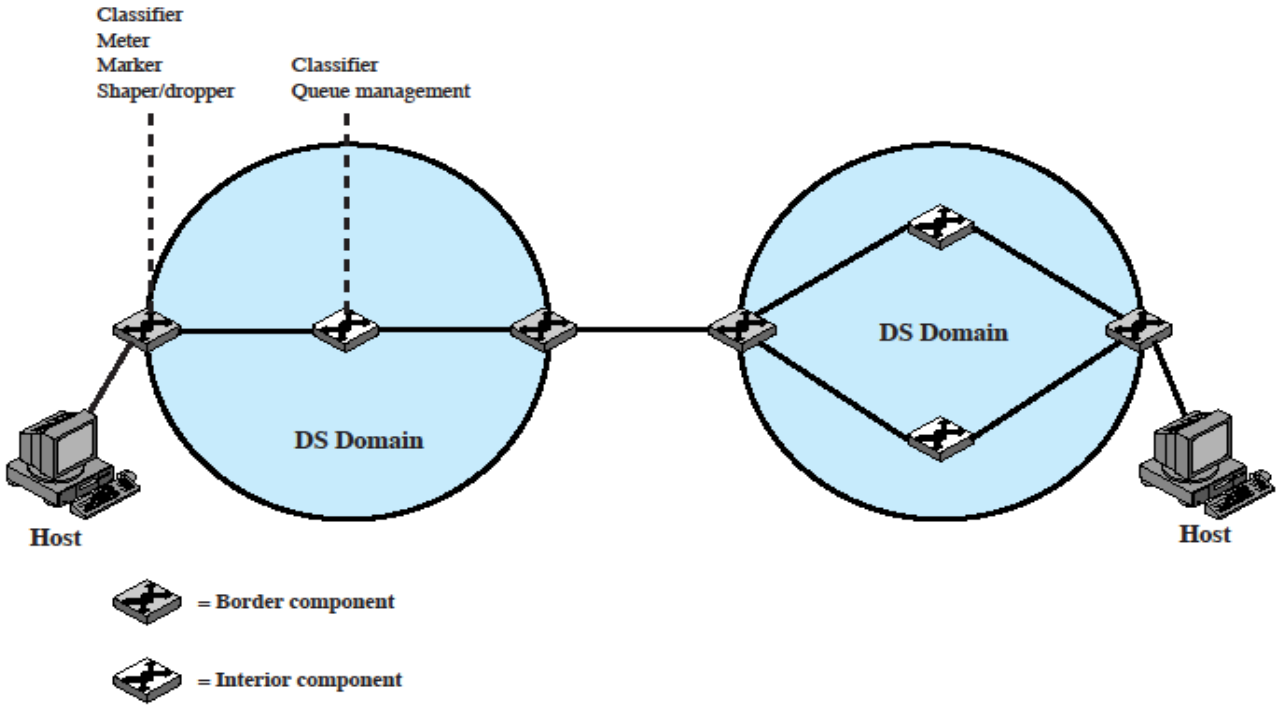


Figure 19.14 DS Domains

DS traffic conditioner [Sta07]

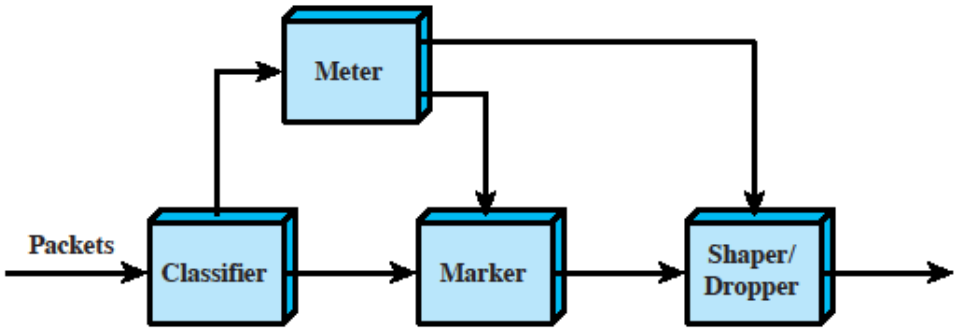


Figure 19.15 DS Traffic Conditioner

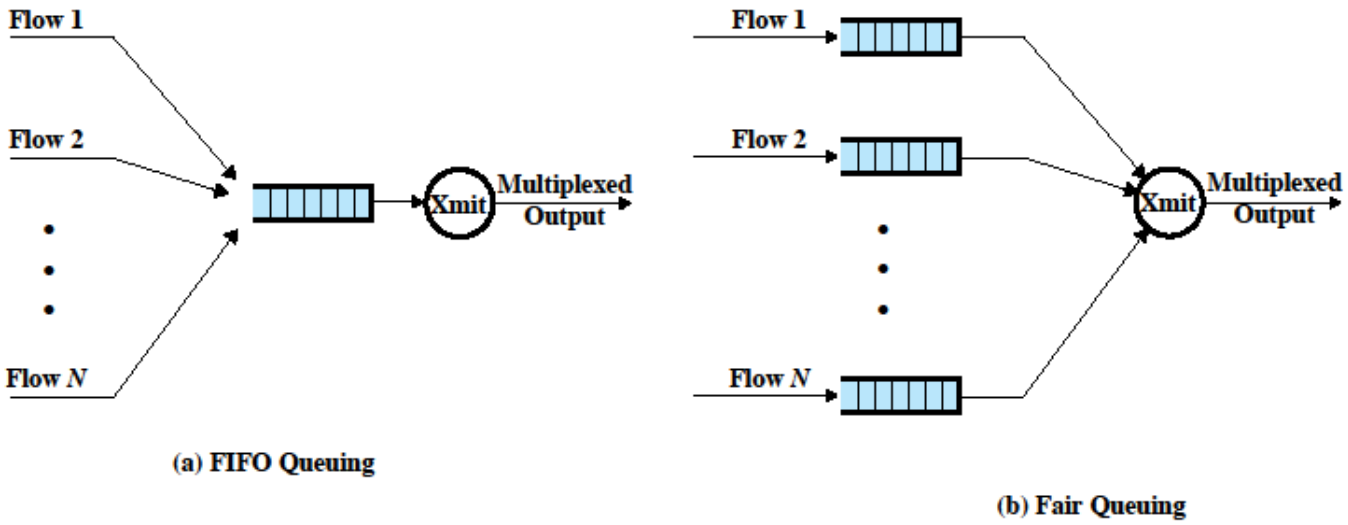
Per hop behaviour (PHB)

- Two PHB are defined: EF si AF:
- Expedited forwarding (EF) PHB
  - For low-loss, low-delay and low-jitter end-to-end services
  - It is expected that the packets from the EF class will meet only empty or very short queues
  - Border nodes will limit the rate and the burstiness of these packets to predefined values, and interior nodes will treat them such that no queueing effects will appear
  - Can be implemented using priority queueing, the EF traffic having higher priority than other queues (with other types of traffic).
- Assured forwarding (AF) PHB
  - Offers services better than best effort (BE), but without requiring to reserve resources and without requiring different treatment to flows from different users.
  - Four classes of AF are defined, each of them corresponding to a traffic profile
  - Within each class, 3 values of precedence are defined. In case of congestion, the packets will be eliminated according to their precedence (packets with lower precedence will be protected, and those with higher precedence will be dropped if necessary).

Queueing disciplines

- FIFO (first in first out):
  - Easy to implement, but:
  - Cannot provide QoS guarantees
  - cannot protect a well behaving flow against the ill behaving flows
  - If in a queue there are small packets queued after big packets, the small packets might experience very high delay
- EDF (earliest deadline first):
  - Every packet contains in its header the moment when it must depart from the queue
  - Packets are sorted in the queue in their departing order, which means that some packets will have to be inserted in the queue
  - Insertion of a packet in a queue is a very complex operation
  - EDF would have been the ideal algorithm, especially for rt traffic, but it is too complex for being implemented
- Fair queueing type algorithms: a solution between FIFO and EDF:
  - Unlike FIFO, they can provide QoS differentiation
  - Have lower complexity than EDF (do not sort packets in queues)
  - Algorithms: Weighted Fair Queueing and its derivates

Queueing disciplines [Sta07]



(a) FIFO Queuing

(b) Fair Queuing

Figure 19.12 FIFO and Fair Queuing

Weighted Fair Queueing (WFQ)

- WFQ algorithms approximate an ideal algorithm, named Fluid Fair Queueing (FFQ) or GPS (Generalized Processor Sharing), proposed by Parekh and Gallagher [PaGa93], that originates in the (ideal) algorithm bit by bit round robin (Keshav and Demers [DKS90]):
  - GPS assumes that every packet is indefinitely divisible, and that the packets from all the backlogged (i.e. non empty) flows are transmitted simultaneously
  - Every flow *i* is allocated a weight w_i, and, if the capacity of the network is C, every flow will receive C x W_i/(sum(W_j)) from the network capacity
    - The sum is for all backlogged flows
    - For each packet it is computed the start time (when it starts being transmitted) and the finish time (when its transmission from the queue is finished) in GPS (FFQ)
- The real algorithms approximate FFQ (GPS), transmitting the real packets in the order of their finish or/and start time of the packets in FFQ.
- Real algorithms approximate better or worse the ideal FFQ algorithm, and they have a higher or lower complexity
- In order to simplify the computations, the *virtual time* was introduced, and the packets are tagged with the virtual start time and virtual finish time.
- For the virtual time, the values of these times are not so important, but their order is important (start time and finish time are conventional values).
- Real algorithms: Packet by Packet GPS (PGPS) called also Weighted Fair Queueing (WFQ), Start time Fair Queueing (SFQ), Worst Case Weighted Fair Queueing WF2Q, etc…

Priority queueing

- It behaves like a limit case of WFQ, when the ratio between queues weights is infinite
- Data packets are put in queues, each queue having a certain priority
- A lower priority queue will be served only if all queues that have a higher priority are empty.
- Low priority queues can be easily starved.
- The algorithm is also known as static priority queueing (SPS) or simple priority queueing
- For DiffServ, it can be used a combination of SPS and WFQ: one queue with EF traffic, several queues with AF traffic and one queue for BE (best effort), SPS is applied between the categories EF, AF and BE, and WFQ is applied between the AF queues.

Bibliography

- [Sta07], William Stallings, Data and Computer Communications, , 8th edition, Pearson Prentice Hall,2007 ISBN 0-13-243310-9
- [PaGa93], Parekh, A, Gallagher, R., A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single-Node Case, IEE/ACM TRANSACTIONS ON NETWORKING, VOL. 1, NO. 3, JUNE 1993, pp. 344-357
- [DKS90] A, Demers, S. Keshav, and S. Shenkar, “Analysis and simulation of a fair queueing algorithm,” Internet. Res. and Eper., vol. 1, 1990