CHAPTER 20

INTERNETWORK QUALITY OF SERVICE

❖ **What is the difference between elastic and inelastic traffic?**

Ans:

Elastic Traffic Elastic traffic is that which can adjust, over wide ranges, to changes in delay and throughput across an internet and still meet the needs of its applications. Elastic include the common applications that operate over TCP or UDP, including file transfer (FTP), electronic mail (SMTP), remote login (TELNET), network management (SNMP), and Web access (HTTP).

Inelastic Traffic Inelastic traffic does not easily adapt, if at all, to changes in delay and throughput across an internet. The prime example is real-time traffic.

The **requirements for inelastic traffic** may include the following:

- **Throughput:** A minimum throughput value may be required.
- **Delay:** An example of a delay-sensitive application is stock trading; someone who consistently receives later service will consistently act later, and with greater disadvantage.
- **Jitter:** The magnitude of delay variation, called jitter, is a critical factor in real-time applications. The larger the allowable delay variation, the longer the real delay in delivering the data and the greater the size of the delay buffer required at receivers.
- **Packet loss:** Real-time applications vary in the amount of packet loss, if any, that they can sustain.

❖ **Which functions does ISA use to manage congestion and provide QoS transport?**

Ans:

ISA makes use of the following functions to manage congestion and provide QoS transport:

- **Admission control:** For QoS transport, ISA requires that a reservation be made for a new flow. If the routers collectively determine that there are insufficient resources to guarantee the requested QoS, then the flow is not admitted. The protocol RSVP is used to make reservations.

- **Routing algorithm:** The routing decision may be based on a variety of QoS parameters, not just minimum delay.
- **Queuing discipline:** A vital element of the ISA is an effective queuing policy that takes into account the differing requirements of different flows.
- **Discard policy:** A discard policy determines which packets to drop when a buffer is full and new packets arrive. A discard policy can be an important element in managing congestion and meeting QoS guarantees.

❖ **Explain about Integrated Services Architecture (ISA) components and services with figures.**

Ans:

Figure is a general depiction of the implementation architecture for ISA within a router. Below the thick horizontal line are the forwarding functions of the router; these are executed for each packet and therefore must be highly optimized. The remaining functions, above the line, are background functions that create data structures used by the forwarding functions.

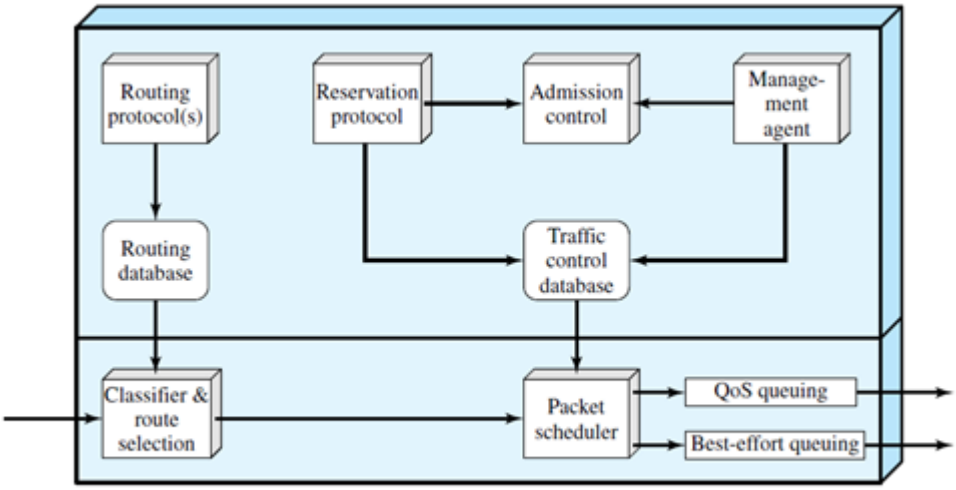


Figure Integrated Services Architecture Implemented in Router

❖ The principal **background functions** are as follows:

- **Reservation protocol:** This protocol is to reserve resources for a new flow at a given level of QoS. It is used among routers and between routers and end systems. The reservation protocol is responsible for maintaining flow-specific state information at the end systems and at the routers along the path of the flow. RSVP is used for this

purpose. The reservation protocol updates the traffic control database used by the packet scheduler to determine the service provided for packets of each flow.

- **Admission control:** When a new flow is requested, the reservation protocol invokes the admission control function. This function determines if sufficient resources are available for this flow at the requested QoS. This determination is based on the current level of commitment to other reservations and/or on the current load on the network.
- **Management agent:** A network management agent is able to modify the traffic control database and to direct the admission control module in order to set admission control policies.
- **Routing protocol:** The routing protocol is responsible for maintaining a routing database that gives the next hop to be taken for each destination address and each flow.

❖ The two principal functional areas that accomplish **forwarding** are the following:

- **Classifier and route selection:** For the purposes of forwarding and traffic control, incoming packets must be mapped into classes. A class may correspond to a single flow or to a set of flows with the same QoS requirements. The selection of class is based on fields in the IP header. Based on the packet's class and its destination IP address, this function determines the next-hop address for this packet.
- **Packet scheduler:** This function manages one or more queues for each output port. It determines the order in which queued packets are transmitted and the selection of packets for discard, if necessary. Decisions are made based on a packet's class, the contents of the traffic control database, and current and past activity on this outgoing port. Part of the packet scheduler's task is that of policing, which is the function of determining whether the packet traffic in a given flow exceeds the requested capacity and, if so, deciding how to treat the excess packets.

Token Bucket Scheme

The token bucket traffic specification, this is a way of characterizing traffic that has three advantages in the context of ISA:

1. Many traffic sources can be defined easily and accurately by a token bucket scheme.
2. The token bucket scheme provides a concise description of the load to be imposed by a

flow, enabling the service to determine easily the resource requirement.

3. The token bucket scheme provides the input parameters to a policing function.

A token bucket traffic specification consists of two parameters: a token replenishment rate R and a bucket size B. The token rate R specifies the continually sustainable data rate; that is, over a relatively long period of time, the average data rate to be supported for this flow is R. The bucket size B specifies the amount by which the data rate can exceed R for short periods of time. The exact condition is as follows: During any time period T, the amount of data sent cannot exceed RT + B.

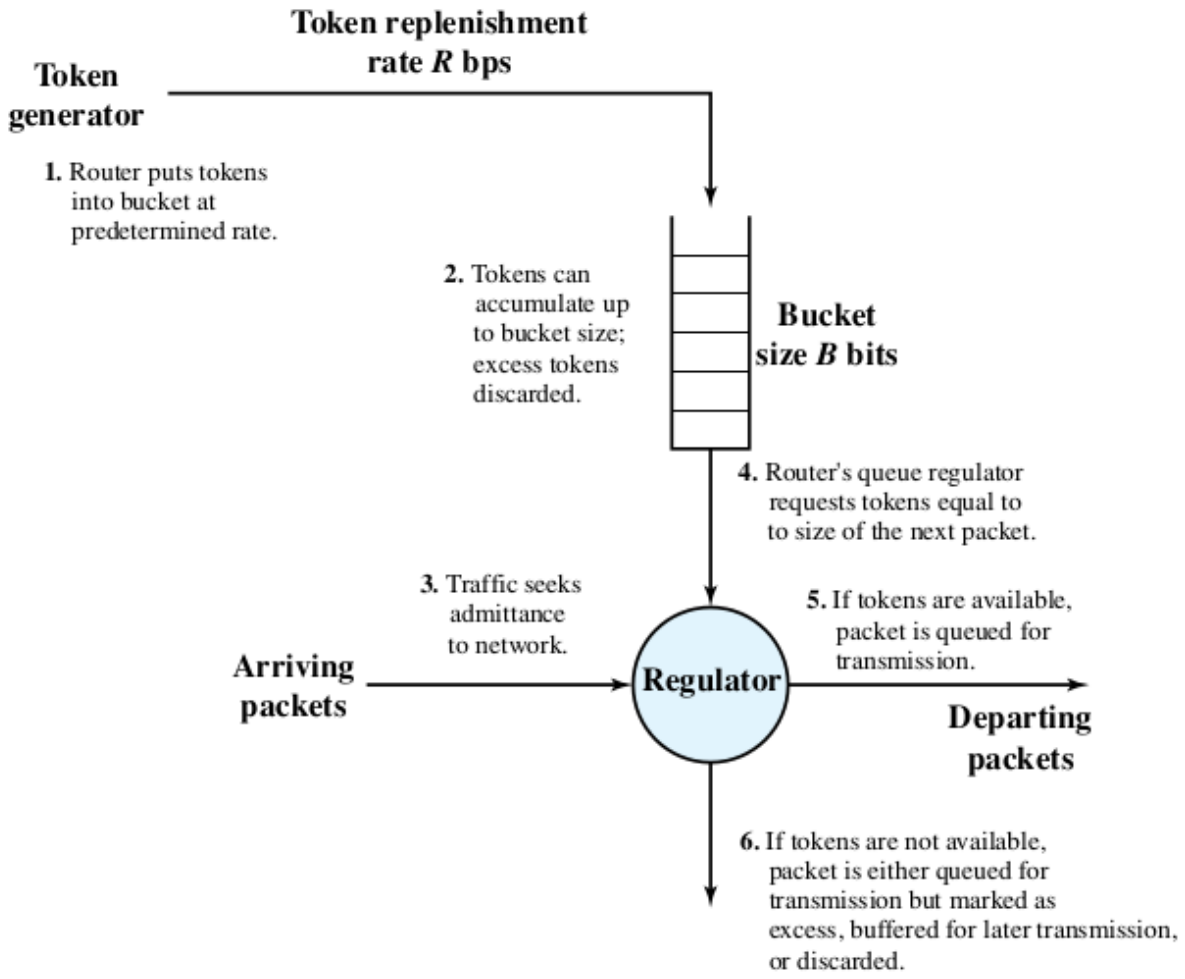


Figure: Token Bucket Scheme

❖ **Queuing Discipline**

FIFO Queue

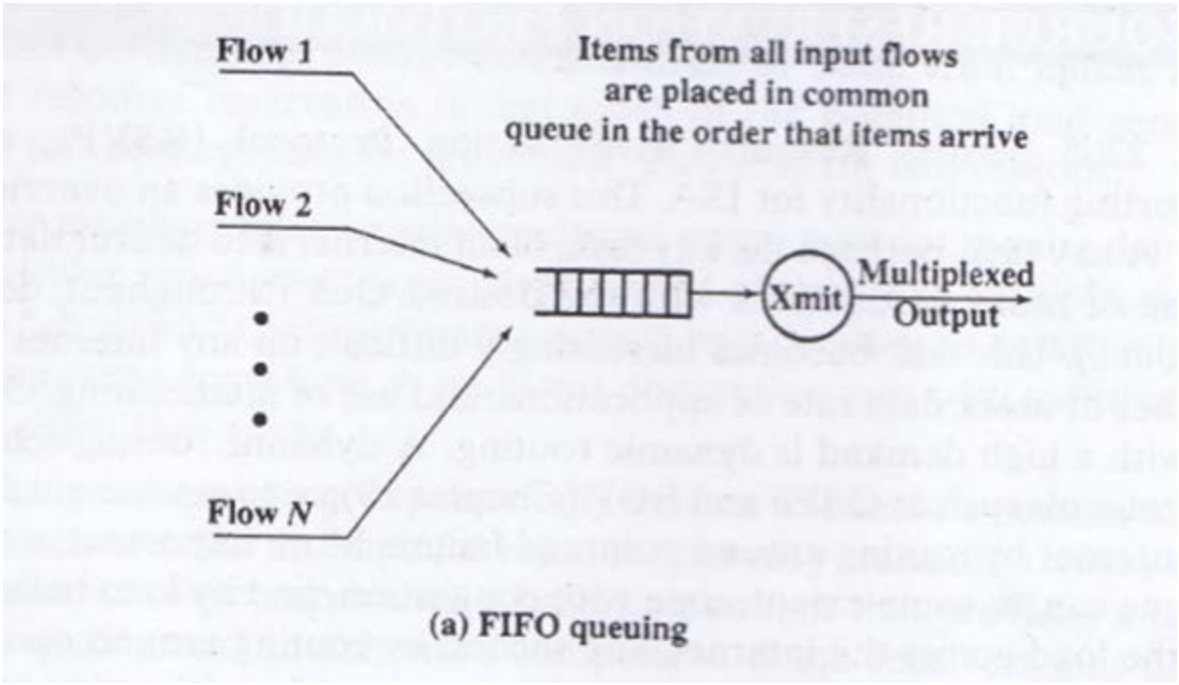
Router traditionally have a first-in-first-out (FIFO) queuing discipline at each output port. A single queue is maintained at each at each output port. When a new packet arrives and is routed to an output port, it is placed at the end of the queue. As long as the queue is not empty, the router transmits packets from queue, taking the oldest remaining packet next.

Fair Queue

A router maintains multiple queues at each output port. Each incoming packet is placed in the queue for its flow. The queues are serviced in round robin fashion, taking one packet from each non-empty queue in turn. Empty queues are skipped over. This scheme fair in that each busy flow gets to send exactly one packet per cycle. Further, this is a form of load balancing among the various flows. There is no advantage in being greedy. A greedy flow finds that its queues become long, increasing its delays, whereas other flows are unaffected by this behavior.

WFQ Queue

WFQ takes into account the amount of traffic through each queue and gives busier queues more capacity without completely shutting out less busy queues. In addition WFQ can take into account the amount of service requested by each traffic flow and adjust the queuing discipline accordingly.



(a) FIFO queuing

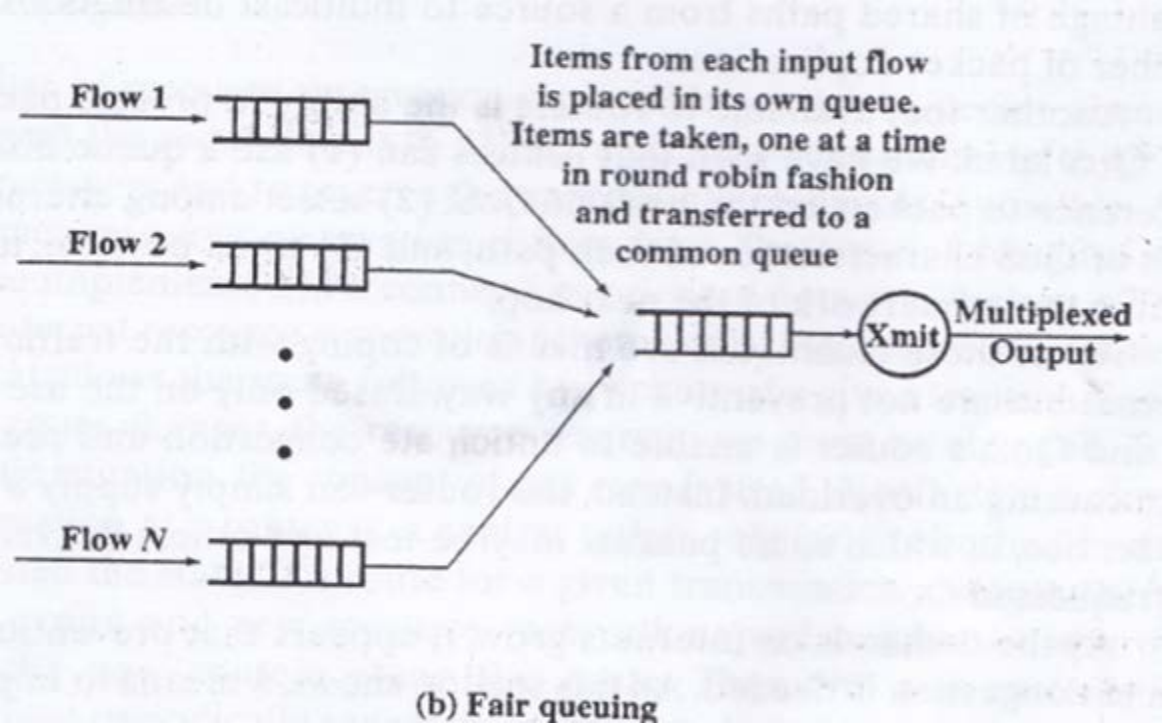


Figure: FIFO and Fair Queuing

- **What is the purpose of Integrated Service Architecture? List and briefly describe the three categories of service offered by ISA.**

Ans:

The purpose of Integrated Service Architecture is to provide QOS transport over IP based internets

The three categories of service offered by ISA (8 marks)

Guaranteed

Controlled load

Best effort

Guaranteed Service

The key elements of the guaranteed service are as follows:

- The service provides assured capacity, or data rate.
- There is a specified upper bound on the queuing delay through the network.
- There are no queuing losses. That is, no packets are lost due to buffer overflow; packets may be lost due to failures in the network or changes in routing paths.

The guaranteed service is the most demanding service provided by ISA.

Controlled Load

The key elements of the controlled load service are as follows:

- The service tightly approximates the behavior visible to applications receiving best-effort service under unloaded conditions.
- There is no specified upper bound on the queuing delay through the network. However, the service ensures that a very high percentage of the packets do not experience delays.
- A very high percentage of transmitted packets will be successfully delivered (i.e., almost no queuing loss).

The controlled service is useful for applications that have been referred to as adaptive real-time applications.

❖ **Explain about goals and characteristics of RSVP.**

Ans:

Goals and characteristics of RSVP

The specification lists the following characteristics of RSVP:

- **Unicast and multicast:** RSVP makes reservations for both unicast and multicast transmissions, adapting dynamically to changing group membership as well as to changing routes, and reserving resources based on the individual requirements of multicast members.
- **Simplex:** RSVP makes reservations for unidirectional data flow. Data exchanges between two end systems require separate reservations in the two directions.
- **Receiver-initiated reservation:** The receiver of a data flow initiates and maintains the resource reservation for that flow.
- **Maintaining soft state in the internet:** RSVP maintains a soft state at intermediate routers and leaves the responsibility for maintaining these reservation states to end users.
- **Providing different reservation styles:** These allow RSVP users to specify how reservations for the same multicast group should be aggregated at the intermediate switches. This feature enables a more efficient use of internet resources.
- **Transparent operation through non-RSVP routers:** Because reservations and RSVP are independent of routing protocol, there is no fundamental conflict in a mixed environment in which some routers do not employ RSVP. These routers will simply use a best-effort delivery technique.

Describe the several key characteristics of Differentiated Services (DS).

Ans:

- 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.
- 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 Field

Packets are labeled for service handling by means of the 6-bit DS field in the IPv4 header or the IPv6 header. The value of the DS field, referred to as the DS code-point, is the label used to classify packets for differentiated services.

With a 6-bit codepoint, there are in principle 64 different classes of traffic that could be defined. These 64 codepoints are allocated across three pools of code-points, as follows:

- Codepoints of the form **xxxxx0**, where x is either 0 or 1, are reserved for assignment as standards.
- Codepoints of the form **xxxx11** are reserved for experimental or local use.
- Codepoints of the form **xxxx01** are also reserved for experimental or local use but may be allocated for future standards action as needed.

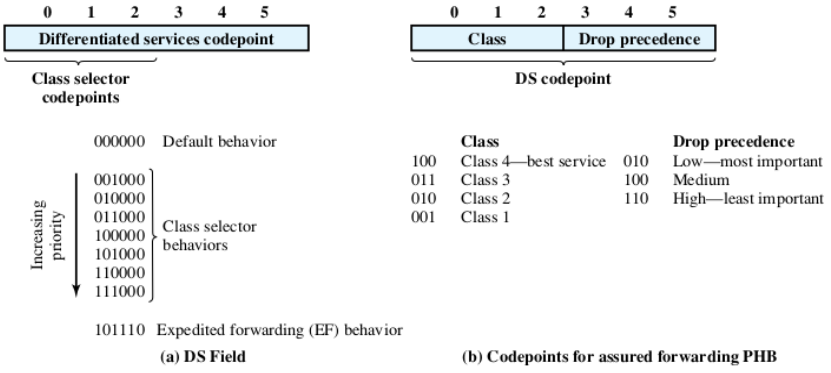
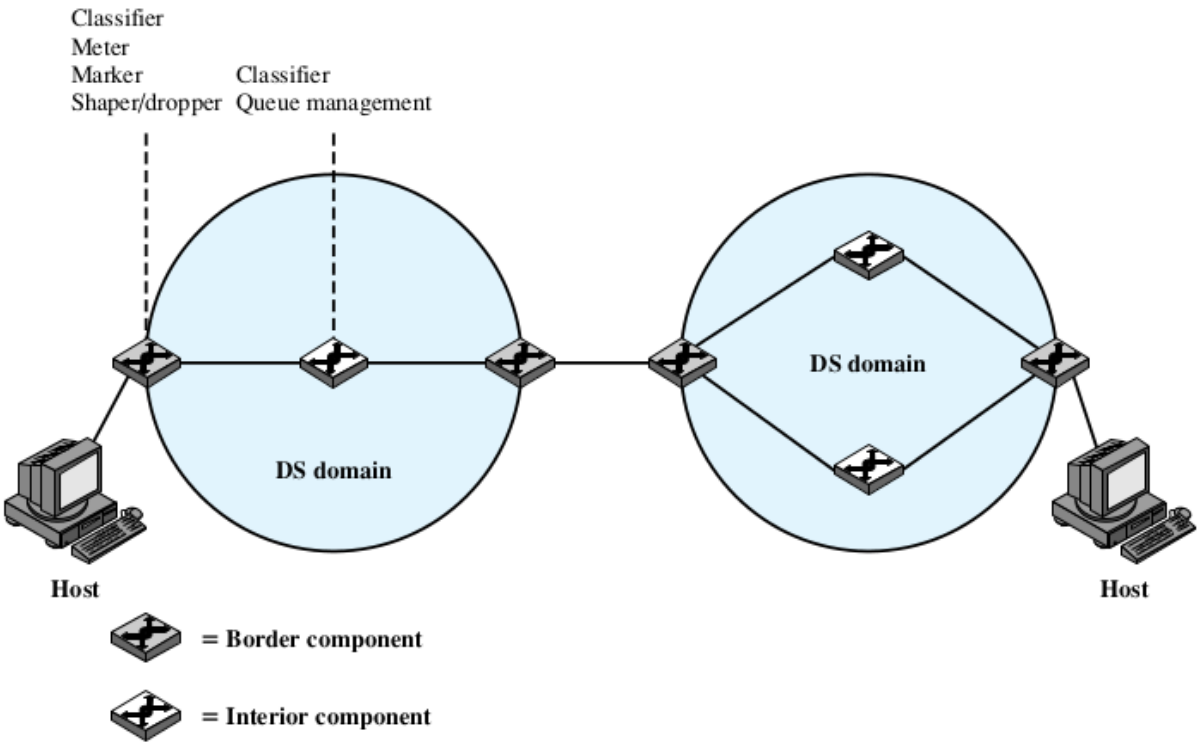


Figure: DS Field

DS Configuration and Operation



A DS domain consists of a set of contiguous routers; that is, it is possible to get from any router in the domain to any other router in the domain by a path that does not include routers outside the domain.

Within a domain, the interpretation of DS code-points is uniform, so that a uniform, consistent service is provided. Routers in a DS domain are either boundary nodes or interior nodes.

Typically, the interior nodes implement simple mechanisms for handling packets based on their DS codepoint values. This includes queuing discipline to give preferential treatment depending on codepoint value, and packet-dropping rules to dictate which packets should be dropped first in the event of buffer saturation. The DS specifications refer to the forwarding treatment provided at a router as per-hop behavior (PHB).

The boundary nodes include PHB mechanisms but more sophisticated traffic conditioning mechanisms are also required to provide the desired service. Thus, interior routers have minimal functionality and minimal overhead in providing the DS service, while most of the complexity is in the boundary nodes. The boundary node function can also be provided by a host system attached to the domain, on behalf of the applications at that host system.

DS Traffic Conditioner

Classifier: Separates submitted packets into different classes. A classifier may separate traffic only on the basis of the DS codepoint or based on multifield classifier.

Meter: The meter determines whether a given packet stream class is within or exceeds the service level guaranteed for that class.

Marker: Re-marks packets with a different codepoint as needed. If any packets in that class that exceed the throughput in some defined time interval may be re-marked for best effort handling.

Shaper: Delays packets as necessary so that the packet stream in a given class does not exceed the traffic rate specified in the profile for that class.

Dropper: Drops packets when the rate of packets of a given class exceeds that specified in the profile for that class.

DS Traffic Conditioner

The metering function measures the volume of packets over a particular time interval to determine a flow's compliance with the traffic agreement.

If the host is bursty, a simple data rate or packet rate may not be sufficient to capture the desired traffic characteristics.

If a traffic flow exceeds some profile, several approaches can be taken. Individual packets in excess of the profile may be re-marked for lower-quality handling and allowed to pass into the DS domain.

A traffic shaper may absorb a burst of packets in a buffer and pace the packets over a longer period of time. A dropper may drop packets if the buffer used for pacing becomes saturated.

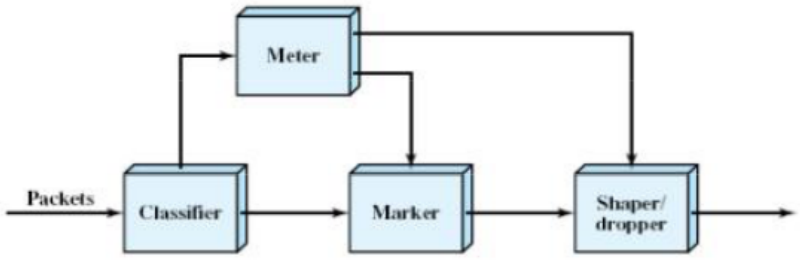


Figure : DS Traffic Conditioner

SERVICE LEVEL AGREEMENTS

SLA typically includes the following information:

- A description of the nature of service to be provided: A basic service would be IP-based network connectivity of enterprise locations plus access to the Internet. The service may include additional functions such as Web hosting, maintenance of domain name servers, and operation and maintenance tasks.
- The expected performance level of the service: The SLA defines a number of metrics, such as delay, reliability, and availability, with numerical thresholds.
- The process for monitoring and reporting the service level: This describes how performance levels are measured and reported.

Fill in the blanks.

No	Questions	Answers
1.	The purpose of ---------- is to enable the provision of QoS support over IP-based internets.	ISA
2.	Traffic on a network or internet can be divided into two broad categories: elastic and----------.	inelastic
3.	---------- is that which can adjust, over wide ranges, to changes in delay and throughput across an internet and still meet the needs of its applications.	Elastic traffic
4.	---------- traffic does not easily adapt, if at all, to changes in delay and throughput across an internet.	Inelastic
5.	The magnitude of delay variation is called ----------.	jitter
6.	Three categories of service are defined as Guaranteed, ----------and Best effort.	Controlled load
7.	In ----------, a router maintains a single queue at each output port.	FIFO Queuing
8.	In----------, a router maintains multiple queues at each output port.	Fair Queuing
9.	----------takes into account the amount of traffic through each queue.	Weighted Fair Queuing
10.	----------provides supporting functionality for ISA.	RSVP
11.	Reservation requests are sent by----------.	receivers
12.	RSVP stands for----------.	Resource Reservation Protocol
13.	Wildcard-filter (WF) style is represented in the form----------.	WF (*{Q})
14.	Fixed-filter (FF) style is represented in the form ----------.	FF(S1{Q1},
15.	Shared-explicit (SE) style is represented in the form ----------.	S2{Q2}, S3{Q4},… SE(S1,S2,S3, …{Q}).
16.	RSVP uses two basic message types:----------.	Resvand Path

17.	----------is a contiguous (connected) set of nodes, capable of implementing differentiated services,	DS Domain
18.	----------is a service contract between a customer and a service provider that specifies the forwarding service a customer should receive.	SLA
19.	When a router's buffer overflows, it discards----------.	packets
20.	---------- protocol is to reserve resources for a new flow at a given level of QoS.	Reservation
21.	To overcome the drawbacks of FIFO queuing, some sort of ---------- queuing scheme is used.	fair
22.	The ---------- services architecture is designed to support a range of	differentiated
23.	network services that are differentiated on the basis of performance. ---------- separates submitted packets into different classes.	Classifier
24.	The ---------- determines whether a given packet stream class is within or exceeds the service level guaranteed for that class.	meter
25.	---------- drops packets when the rate of packets of a given class exceeds that specified in the profile for that class.	Dropper