
# 第19章 服务质量

南京大学计算机系 黄皓教授

2007年11月30 日 星期二

# Contents

1. Integrated Services Architecture
2. Resource Reservation
3. Differentiated Services


# 1. Integrated Services Architecture

# Integrated Services Architecture

(1) Elastic Traffic

(2) Inelastic Traffic

(3) ISA Approach

(4) ISA Components

# (1) Type of Internet Traffic – Elastic

- Can adjust to **changes in delay and throughput**
- e.g. common TCP and UDP application
  - FTP: user expect delay proportional to file size, sensitive to changes in throughput
  - SNMP: delay not a problem, except when caused by congestion
  - Web, Telnet: sensitive to delay
  - EMail: insensitive to delay and throughput
- Not per packet delay – **total elapsed time concerned**
  - e.g. web page loading time, ftp file transfer time
  - For small items, **delay** across internet dominates
  - For large items it is **throughput** over connection
- QOS based internet service could be benefit.
  - Without QoS, routers are dealing evenhandedly with arriving IP packets, with no concern for the type of application and whether this packet is part of a large transfer element or a small one. It is unlikely that the resources will be allocated fairly.

# (2) Inelastic Traffic

- Not easily adapt to changes in delay and throughput
  - e.g. real-time traffic
- **Throughput**
  - A given minimum throughput required
- **Delay**
  - Delay sensitive application, e.g. stock trading
- **Jitter**
  - Magnitude of delay variation, a maximum limit need to be defined
  - More jitter requires a bigger buffer
- **Packet loss**
  - Some are sensitive, e.g. stock trading; some are not, e.g. online video

# (2) Inelastic Traffic

- Traffic Requirements of Internet Apps

| Application | Data Loss | Bandwidth | Time Sensitive |
|---|---|---|---|
| File transfer | no loss | elastic | no |
| Email | no loss | elastic | no |
| Web documents | no loss | elastic | no |
| Real-time audio/video | loss-tolerant | audio: 5k~1Mbps<br>video: 10k~5Mbps | 100's msec |
| Stored audio/video | loss-tolerant | same as above | few secs |
| Interactive games | loss-tolerant | few kpbs up | 100's msec |
| Instant messaging | no loss | elastic | yes and no |

# (2) Inelastic Traffic

## Requirements for Inelastic Traffic

- Difficult to meet requirements on network with **variable queuing delays and congestion**
  - Require preferential treatment
- Applications need to be able to **state their requirements**
  - Ahead of time, using some sort of resource reservation functions (preferred)
  - Or on the fly, using fields in IP header
- Require elastic traffic to be supported as well
  - Inelastic application **do not back off and reduce the demand** in face of congestion
  - Elastic traffic may be crowded off the internet
  - Deny service requests that leave too few resources

# (2) Inelastic Traffic

- Inelastic applications typically do not back off and reduce demand in the face of congestion, in contrast to TCP-based applications.
- In times of congestion , inelastic traffic will continue to supply a high load, and elastic traffic will be crowded off the internet.
- A reservation protocol can help control this situation by denying service requests that would leave too few resources available to handle current elastic traffic.

# (3) ISA Approach (1)

- The purpose
  - Providing QOS support over IP-based internets
  - One central issue is to share the available capability in **times of congestion**
  - An overall architecture, a number of enhancements to the traditional **best-effort** mechanisms
- Traditional congestion control
  - **Routing algorithms**: minimize delay
  - **Packet discard**: buffer overflows on router
  - Inadequate for traffic now on Internet

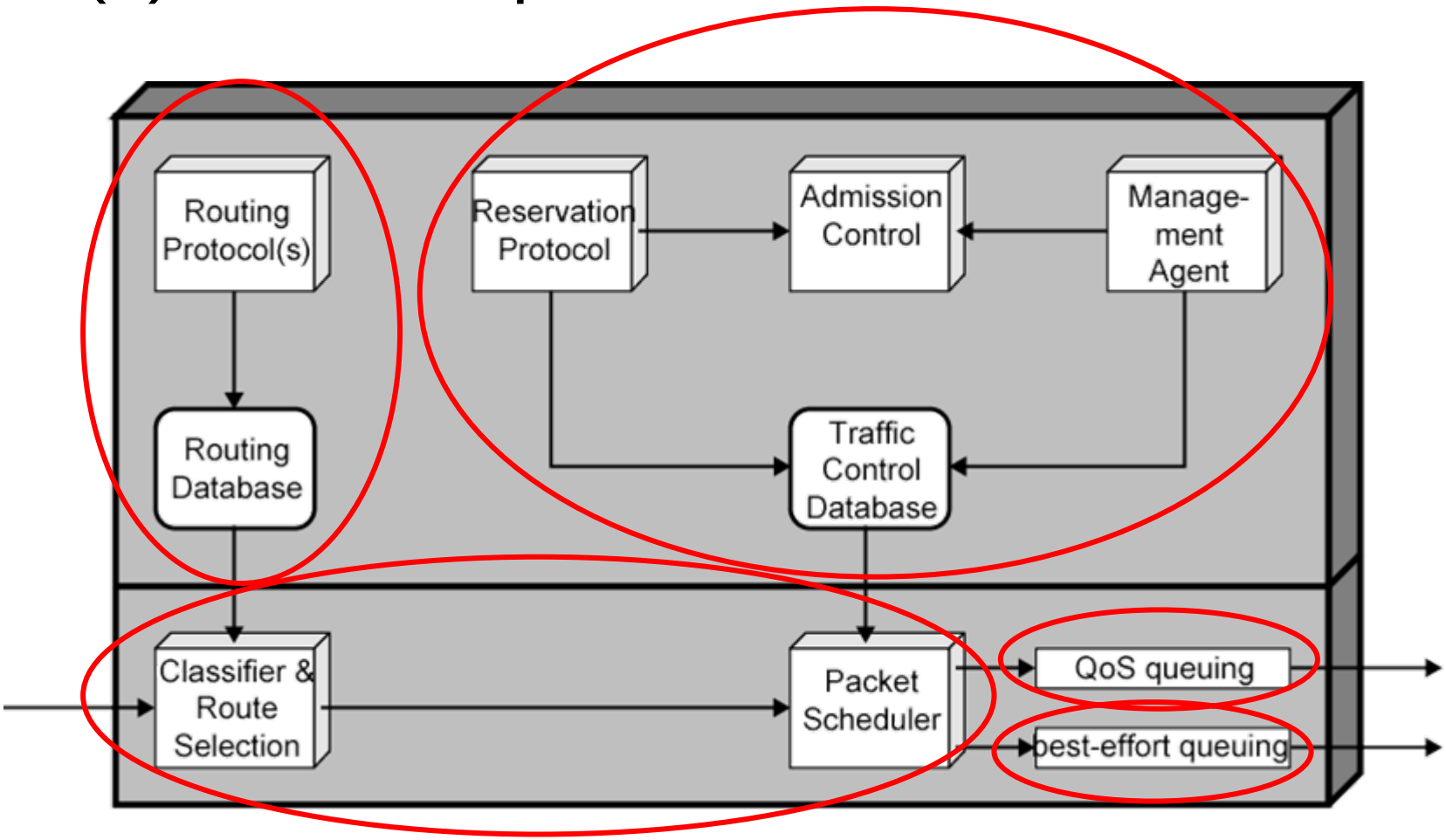
# (3) ISA Approach (2)

- ISA is a overall architecture
  - A number of enhancements to the traditional best-effort mechanisms are being developed.
- Flow
  - **Distinguishable stream of related IP packets that resulted from a single user activity and requires the same QoS.**
  - A flow might consist of one transport connection or one video stream.
  - A flow differs from a TCP connection
    - A flow is unidirectional;
    - Can be multicast
  - An IP packet is identified as a member of flow on the basis of source and destination IP addresses and port number and protocol type.

# (3) ISA Approach (3) — ISA Functions

- **Admission Control**
  - ISA requires that a reservation be made for a new flow.
  - Routers collectively determine whether the resources is sufficient.
  - RSVP is used to make reservations.
- **Routing Algorithm**
  - The routing decision may be based on a variety of QOS parameters, not just delay.
- **Queuing discipline**
  - A vital element of the ISA：effective queuing policy that take account of different flow requirements.
- **Discard policy**
  - Selective discard instead of just new comings.

# (4) ISA Components — ISA in a Router



# (4) ISA Components – Background Functions

- **Reservation protocol**
  - Reserves resource for new flow at a given level of QOS
- **Admission control**
  - Determines if sufficient resources are available for the flow at the requested QOS
- **Management agent**
  - Modifies the traffic control database
  - Directs the admission control module to set policies

**Traffic control database**

# (4) ISA Components – Background Functions

- Routing protocol
  - Maintaining a **routing database**
  - For each destination address and each flow.

# (4) ISA Components – Forwarding

- Classifier and route selection
  - Incoming packets mapped to **classes**
  - A class maybe a single flow or set of flows with same QOS requirements.
  - e.g. all video flows or all flows attributable to a particular organization may be treated identically.
  - Determines next hop based on packet's clsss and destination.
- Packet scheduler
  - Manages one or more queues for each output
  - Order queued packets sent, based on class, traffic control database etc.
  - Policing for packet priority, discard etc.

# (5) ISA Services

- **Traffic specification** (TSpec) defined as service for flow of packets
- Defined on 2 levels
  - **General categories** of service
    - Guaranteed
    - Controlled load
    - Best effort (default)
  - **Particular flow** within each category
    - Specified by the values of certain parameters
- TSpec is part of the contract between the data flow and the service after **reservation** accepted

# (5) ISA Services

- **Traffic specification** (TSpec) defined as service for flow of packets
- Defined on 2 levels
  - **General categories** of service
    - Guaranteed
    - Controlled load
    - Best effort (default)
  - **Particular flow** within each category
    - Specified by the values of certain parameters
- TSpec is part of the contract between the data flow and the service after **reservation** accepted

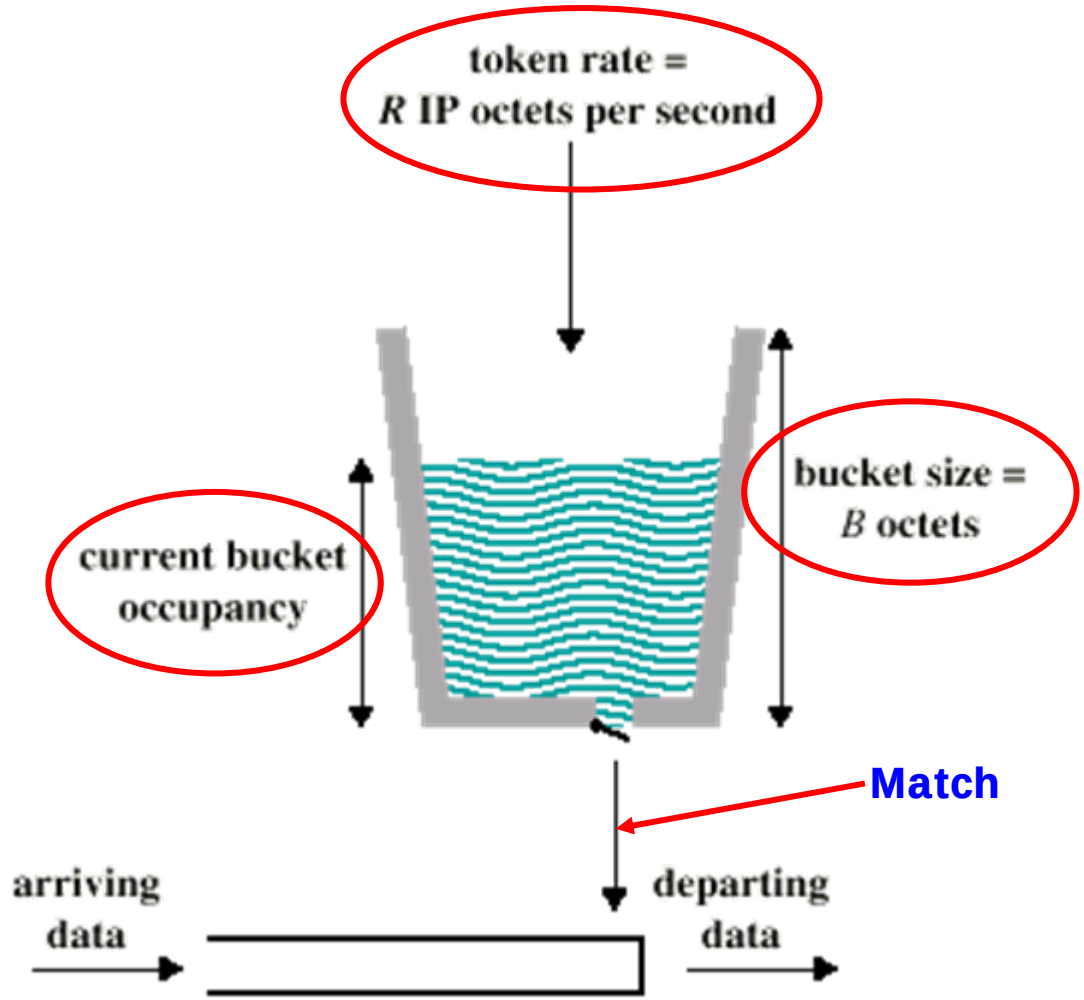
# (5) ISA Services — Token Bucket

- A way of charcaterizing traffic
- Three advantages in the context of ISA
  - Many traffic sources can be defined by token bucket scheme
  - Provides concise description of load imposed by a flow
  - Easy to determine resource requirements
  - Provides input parameters to policing function

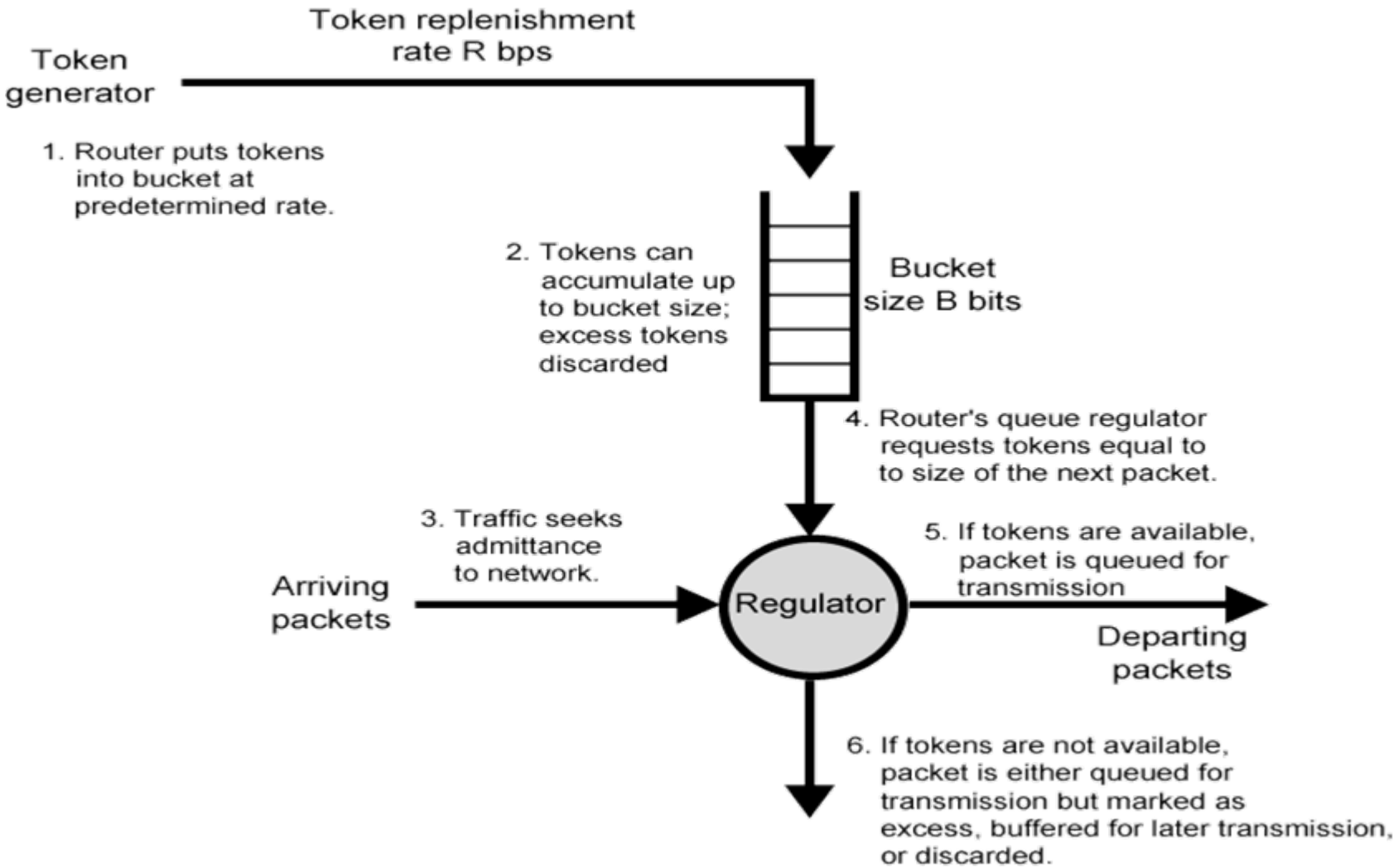
# Token Bucket Traffic Specification

- 2 parameters defined
- Token replenishment rate R
  - Continually **sustainable data rate**
- Bucket size B
  - Amount that data rate can **exceed R for short period**
- During time period T amount of data sent can not exceed "**RT + B**"

# (5) ISA Services — Token Bucket Diagram



# (5) ISA Services — Token Bucket Scheme



# (5) ISA Services — Guaranteed Service

## Key elements

- Assured capacity level or data rate
- Specifies upper bound on queuing delay through network
  - Must be added to propagation delay or latency to get total delay
  - Set high to accommodate rare long queue delays
- No queuing losses
  - i.e. no buffer overflow

## Application

- Real-time play back
  - Uses delay buffer and retains time pattern for incoming message which will not tolerate packet loss

# (5) ISA Services — Controlled Load

## Key elements

- Tightly approximates to best efforts under unloaded network
- No upper bound on queuing delay
  - High percentage of packets do not experience delay over minimum transit delay
  - i.e. propagation plus router processing with no queuing delay
- Very high percentage delivered
  - Almost no queuing loss

## Application

- Adaptive real-time applications
  - Receiver measures jitter and sets playback point
  - Video can drop a frame or delay output slightly
  - Voice can adjust silence periods

# (5) ISA Services — Queuing Discipline

## Traditional

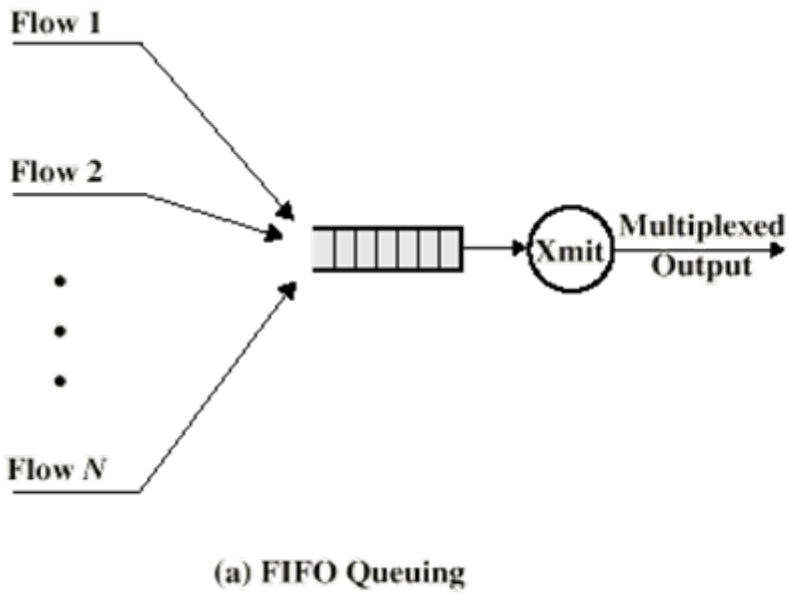
- First in first out (*FIFO*) or first come first served (*FCFS*) at each router port
- No special treatment to high priority packets or flows
- **Small packets held up** by large packets ahead of them in queue
  - Larger average delay for smaller packets
  - Flows of larger packets get better service
- **Greedy *TCP* connection** can crowd out **altruistic** connections
  - If one connection does not back off, others may back off more

# (5) ISA Services — Queuing Discipline – Fair

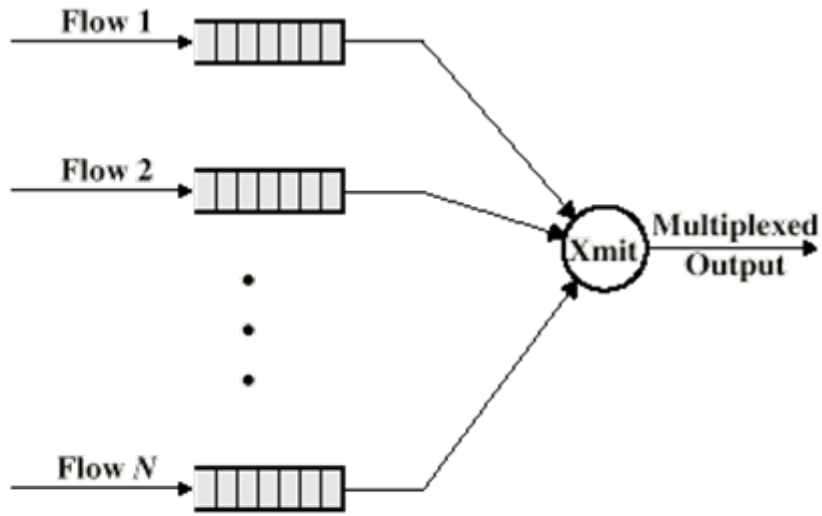
## Fair

- **One for each source or flow**
  - Packet placed in queue for its flow
- **Round robin** servicing
  - Each busy queue (flow) gets exactly one packet per cycle
  - Short packets penalized as each queue sends one packet per cycle
  - Skip empty queues
  - Load balancing among flows
- No advantage to being greedy
  - Your queue gets longer, increasing your delay, whereas other flows are unaffected by his greedy behavior.
- Can have **weighted fair queuing**

# (5) ISA Services — FIFO and Fair Queue




(a) FIFO Queuing




(b) Fair Queuing


# 2. Resource Reservation

# Resource Reservation: RSVP

- RFC 2205, allows **receiver to request** a special end to end *QOS* for its data flows
  - An internet control protocol, designed to operate with current and future unicast and multicast routing protocols
- Unicast applications can reserve resources in **routers** to meet *QOS*
  - If router can not meet request, application informed
- **Multicast is** more demanding, and maybe **reduced**
  - Some group members may not require delivery from particular source over given time
  - Some group members may only be able to handle a portion of the transmission

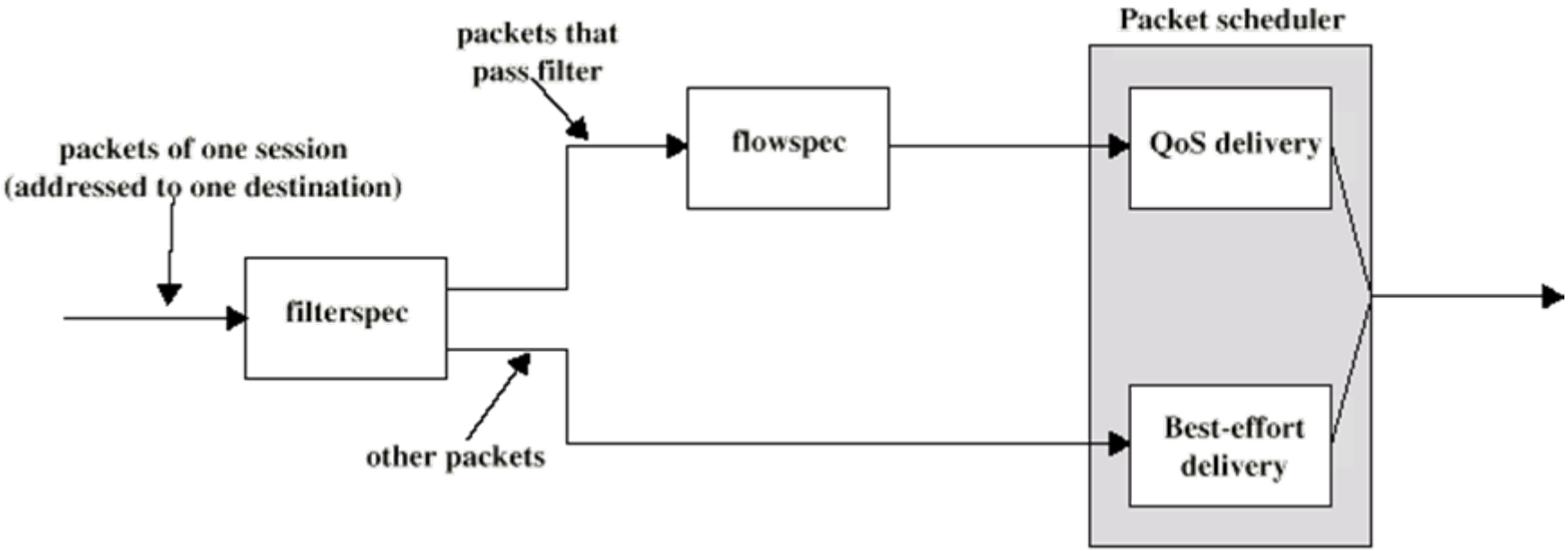
# RSVP

- **Goals**
  - Ability for receivers to make reservations
  - Deal with changes in multicast group membership
  - Enable receivers to select one source
  - Deal with changes in routes
  - Independent of routing protocol
- **Features**
  - Maintain soft state (periodically renewed) on each router
  - Provide different reservation styles
  - Transparent operation through non-RSVP routers
  - Support for IPv4 and IPv6

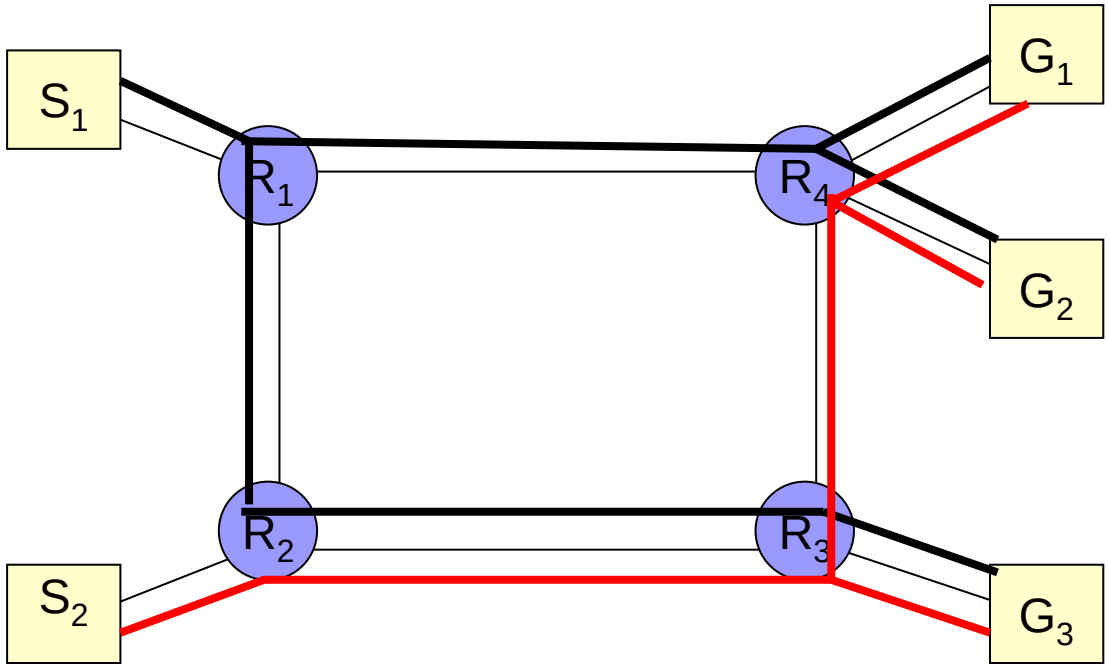
# Data Flows

- **Session**
  - A data flow with a particular destination and transport layer protocol
  - Defined by the triple: <DestAddress, Protocol ID, DstPort>
- **Reservation Model**
  - An RSVP request consists of a flow descriptor: <flowspec, filterspec>
  - Flowspec: specifies a desired QOS, used to set parameters in a node's packet scheduler
    - Flowspec contains the following elements
      - Service class, Rspec (R for reservation) parameter defines the desired QoS, and Tspec(T for traffic) parameter describes the data flow.
  - Filterspec: defines the set of data packets under QOS, used to set the packet classifier.
    - the filter spec together with the session define the set of packets, or flow, that are to receive the desired QoS.
    - Filterspec consisting of the elements of source address and source port

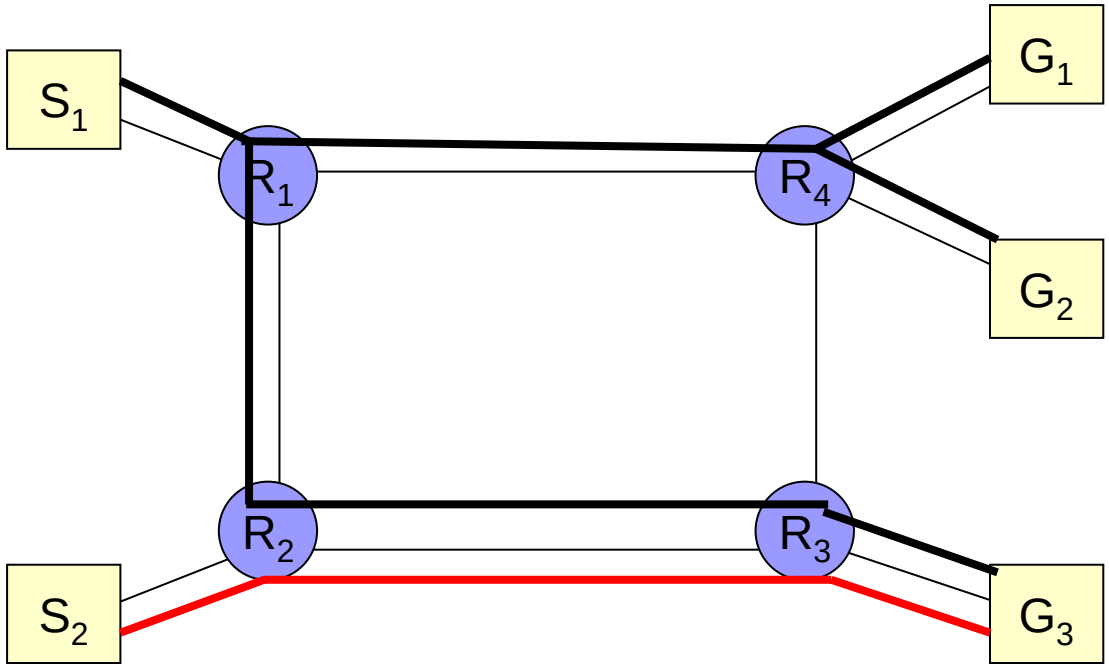
# Illustration of Data Flows



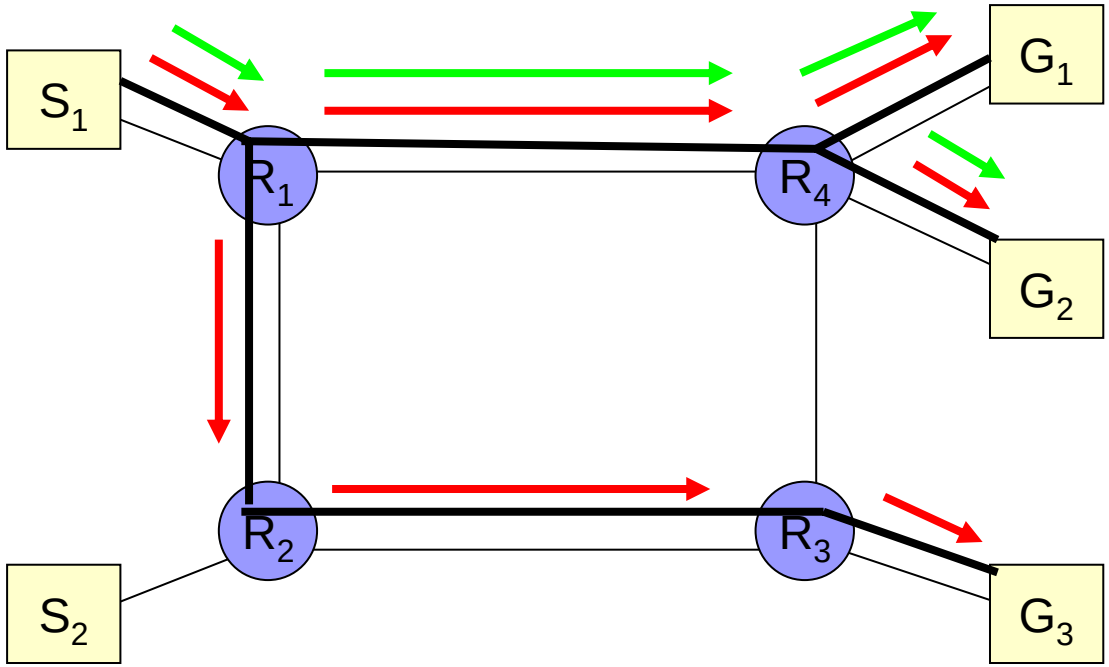
# Data distribute to multicast group



# Filtering by Source



# Filtering by Substream




William Stalling, Computer Networking with Internet Protocols and Technology.

# Reservation Process

- **Receivers**
  - Make RSVP msgs carrying reservation requests
  - Pass the msgs upstream towards the senders
- At each **intermediate node**
  - The RSVP module passes the request to Admission and Policy control and the check is executed
  - Reservation request is propagated
- **Scope** of the request
  - The set of sender hosts to which a reservation request is propagated

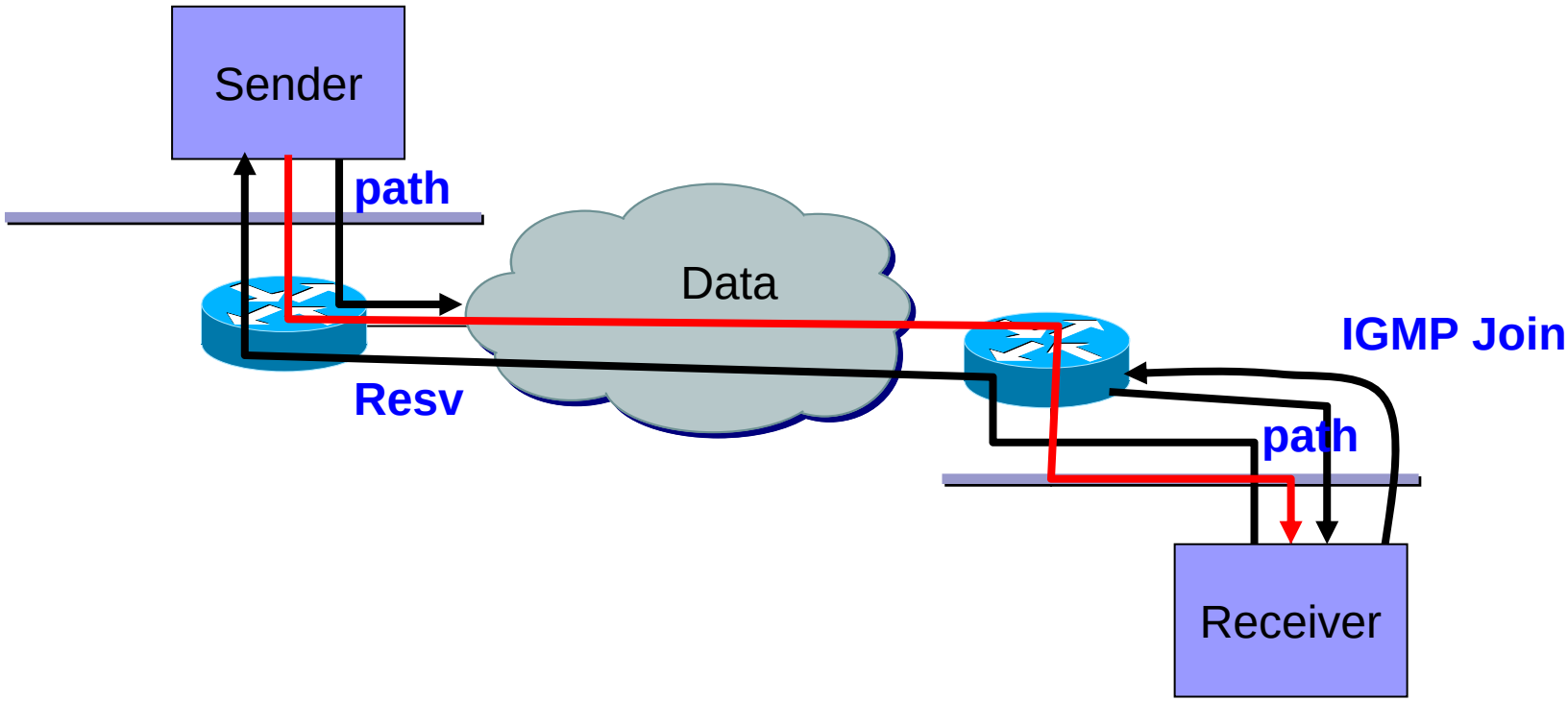
# RSVP Mechanisms (1)

- 2 fundamental RSVP messages
  - **Path**: from sender, pass downstream
  - **Resv**: from receiver, pass upstream
- Path message
  - Each RSVP sender transmits a Path message downstream, store the **path state** in each router along the way
  - The path state includes the IP address of the previous hop which is used for **reverse directing**
  - May gather information that can be used to predict the end to end QOS

# RSVP Mechanisms (2)

- Resv message
  - Each receiver sends a **Resv** message upstream towards the senders
  - Follows the exact reverse path the data will use (**path state**)
  - Creates and maintains the reservation state in each node along the path

# RSVP host model



# Soft State

- Manage the reservation state and/or path state in routers and hosts
- A **soft state** is created and periodically refreshed by **Path** and **Resv** messages
- Must interact with dynamic routing strategy of Internet
- When the route changes, the resource reservation must be changed
- Applications must periodically renew requests during transmission, or the state info will expired
- Teardown msgs can be used to remove **path or reservation state** immediately

# RSVP Operation for Multicast (2)

- Receiver joins a multicast group (IGMP)
- Potential sender issues Path message
- Receiver gets the message identifying sender
- Receiver has reverse path info and may start sending Resv messages
- Resv messages propagate through distribution tree and is delivered to sender
- Sender starts transmitting data packets
- Receiver starts receiving data packets


# 3. Differentiated Services

# Differentiated Services

- ISA and RSVP complex to deploy now
- May not scale well for large volumes of traffic
  - Amount of control signals
  - Maintenance of state information at routers
- DS architecture designed
  - Provide simple, easy to implement, low overhead tool
  - Support range of network services differentiated on basis of performance
- RFC 2475

# Characteristics of DS

- Use IPv4 header Type of Service or IPv6 Traffic Class field
  - No change to IP
- **S**ervice **L**evel **A**greement (*SLA*)
  - Established between provider (internet domain) and customer prior to use of DS
  - DS mechanisms not needed in applications
- Build in aggregation
  - All traffic with same DS field treated same
  - e.g. multiple voice connections
- DS implemented in individual routers by queuing and forwarding based on DS field
  - State information on flows need not saved by routers

# DS Services

- Defined within DS domain
  - Contiguous portion of internet over which consistent set of DS policies administered
  - Typically under control of one organization
- Defined in SLA
  - Customer may be user organization or other DS domain
  - Packet class marked in DS field
- Service provider configures forwarding policies routers
  - Ongoing measure of performance provided for each class
  - Destination within the DS domain, provides the agreed service internally
  - If destination in another domain, attempts to forward packets matching the most appropriate service

# SLA Parameters

- Detailed service performance parameters
  - Expected throughput
  - Drop probability
  - Latency
- Constraints on ingress and egress points
  - Indicate **scope of service**
- Traffic profiles to be adhered to
  - e.g. token bucket parameters
- Disposition of traffic in excess of profile

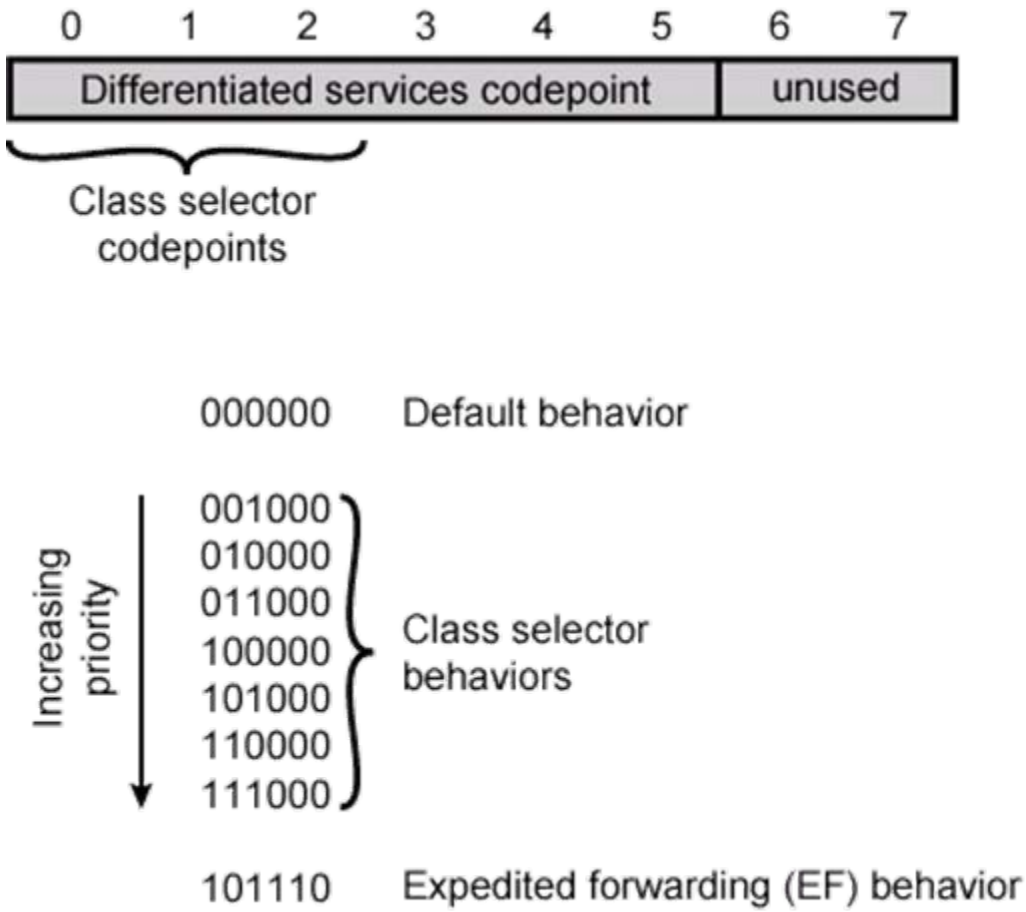
# Example Services

- **Qualitative**
  - Level A – low latency
  - Level B – low loss
- **Quantitative**
  - Level C – 90% of traffic < 50ms latency
  - Level D – 95% in profile traffic delivered
- **Mixed**
  - Level E – allotted twice bandwidth of level F traffic
  - Level F – traffic with drop precedence X higher probability of delivery than that with Y

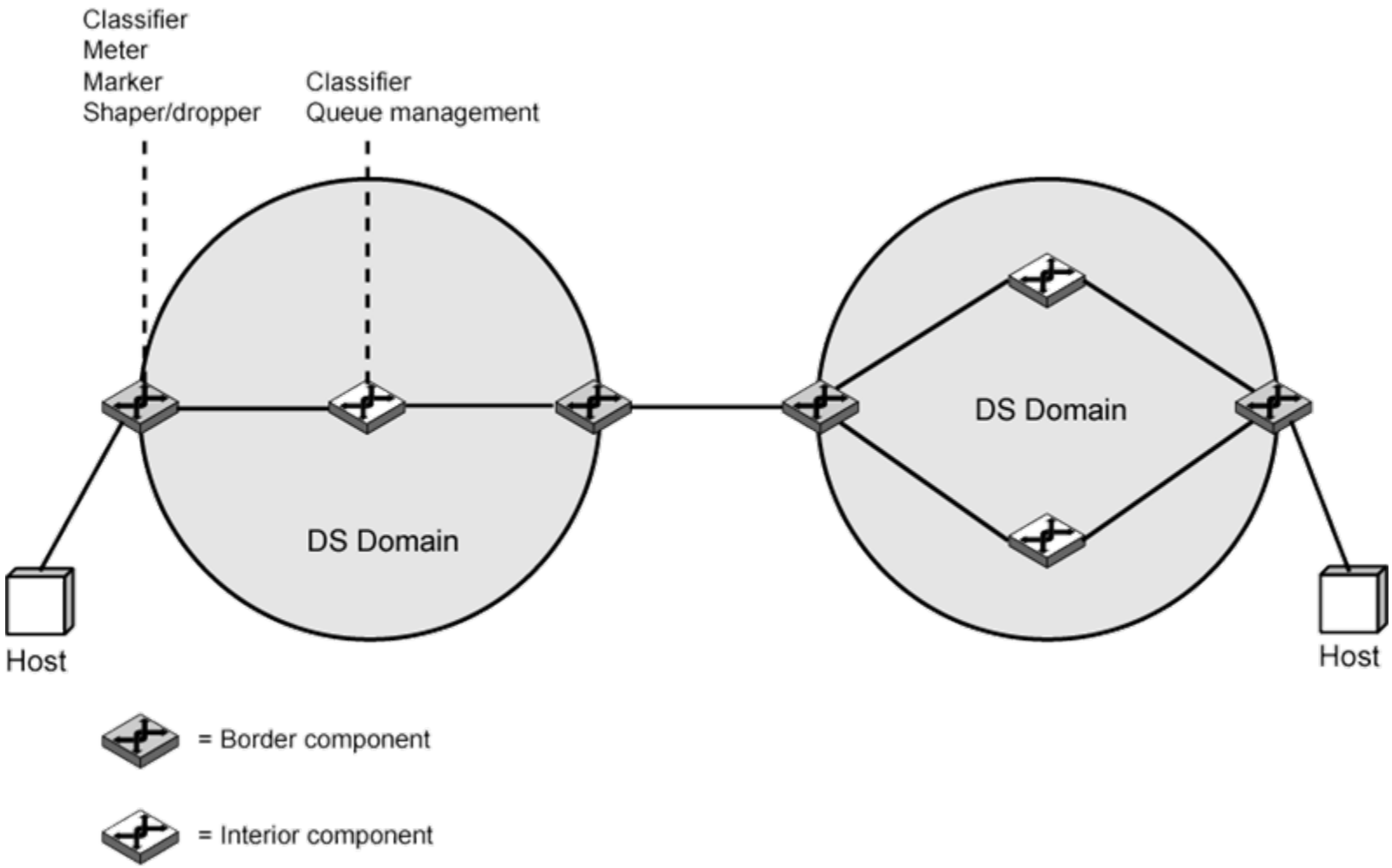
# DS Octet in Detail

- Leftmost 6 bits are DS code-point
  - 64 different classes available
- 3 pools defined
  - xxxxx0 : reserved for standards
    - 000000 : default packet class
    - xxx000 : backwards compatibility with IPv4 TOS
  - xxxx11 : reserved for experimental or local use
  - xxxx01 : experimental or local but may be allocated for standards in future
- Rightmost 2 bits unused

# Illustration of DS Fields



# DS Configuration Diagram



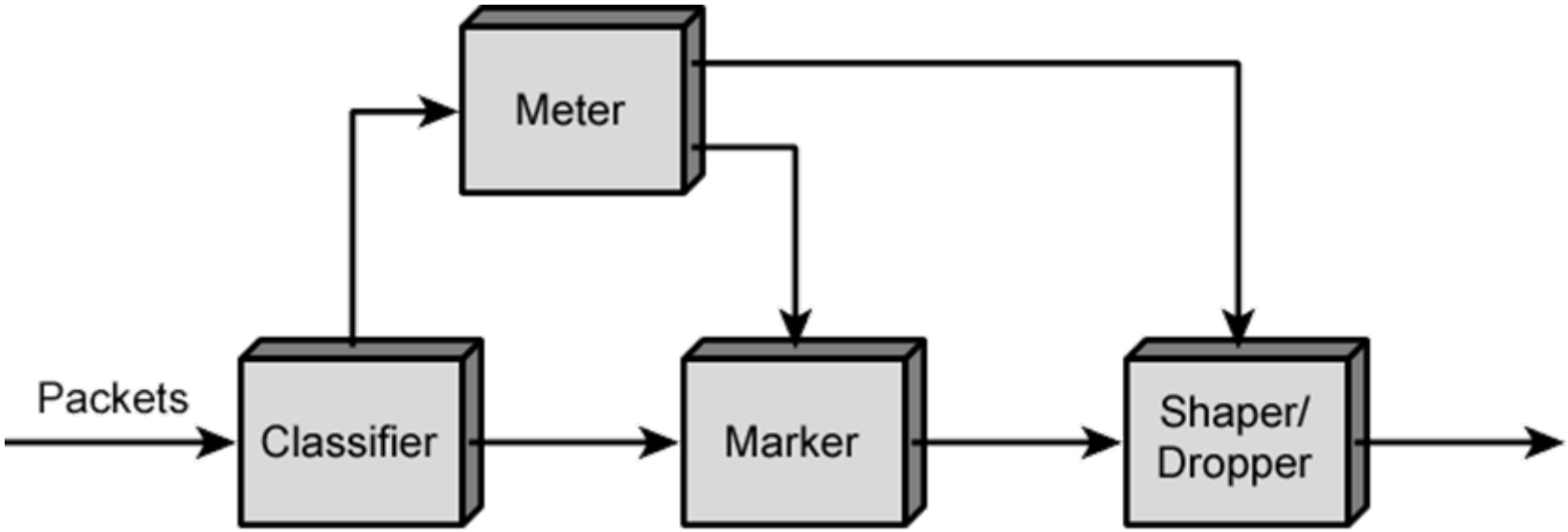
# Interior Routers

- Consistent interpretation of DS code-points within domain
- Simple mechanisms to handle packets based on code-points
  - Classifier and Queue management
- Classifier
  - Differentiate packets based on DS code-point, src & dst addresses, high-level protocol, etc.
- Queue Management
  - Queuing gives preferential treatment depending on code-point
    - Per Hop Behaviour (PHB)
    - Must be available to all routers
  - Packet dropping rule dictates which to drop when buffer saturated

# Boundary Routers

- Traffic conditioning to provide desired service
  - Classifier
    - Separate packets into classes
  - Meter
    - Measure traffic for conformance to profile
  - Marker
    - Policing by remarking code-points if required
  - Shaper
    - Policing by delaying packets
  - Dropper
    - Drops packets if packets rate exceeds those specified in the profile of the class
- Implement PHB rules

# DS Traffic Conditioner



# PHB – Expedited Forwarding

- Specific PHBs defined
  - Associated with specific differentiated services (DSs)
- RFC 3246 defines expedited forwarding (EF) PHB
  - Support for premium service
  - Low-loss, low-delay, low-jitter; assured bandwidth, end-to-end service through DS domains
  - Appears to endpoints as point-to-point connection or leased line
- Difficult in internet or packet-switching network
  - Queues (buffers) at each node (router) will result in loss, delays, and jitter
  - Configure nodes so traffic aggregate has well defined minimum departure rate
  - Condition aggregate so arrival rate at any node is always less than minimum departure rate

# Handle Expedited Forwarding

- Config minimum departure rate in each node by EF PHB
- Use network boundary conditioners (policing and shaping) to assure arrival rate < minimum departure rate
  - Border nodes control traffic aggregate
  - Limit characteristics (data rate, burstiness) to predefined level
- Interior nodes treat traffic so no queuing effects
  - Aggregate's maximum arrival rate must be less than Aggregate's departure rate.
- No specific queuing policy at interior nodes defined in RFC 3246
  - Use simple priority scheme, EF traffic given absolute priority, but
  - EF traffic must not overwhelm interior node, or packet flows for other PHB traffic disrupted

# PHB – Assured Forwarding

- Provides service superior to best-effort, RFC 2597
  - Does not require reservation of resources
  - Does not require detailed discrimination among flows from different users
- Based on explicit allocation
  - Uses offered choice of classes of service
    - Each class describes different traffic profile
    - Aggregate data rate and burstiness
  - Traffic monitored at boundary node
    - Each packet marked in or out of profile
  - Inside network, no separation of traffic from different users
    - Only distinction being whether packet marked in or out
  - When congested, out packets are dropped before in packets

# Classes of Assured Forwarding

- 4 classes defined
  - 4 distinct traffic profiles
  - Select one or more to meet requirements
- Within interior DS node, traffic from different classes treated separately
  - Different amounts of resources (buffer space, data rate)
- Within a class, packets marked by customer or provider with one of 3 drop precedence values
  - Used to determine importance when dropping packets as result of congestion
- Different users will see different levels of service
  - Have different quantities of in packets in service queues

# Assured Forwarding Characteristics

- Level of forwarding assurance depends on
  - How much forwarding resources allocated to AF class that the packet belongs to
  - Current load of class
  - If congested within the class, drop precedence of packet
- Simplicity
  - Very little work required by internal nodes
  - Marking of traffic at boundary nodes based on traffic profiles provides different levels of service to different classes
- Interior nodes use RED algorithm to manage AF traffic

# Problems

- **19.6**
- **19.7**