Computer Networks

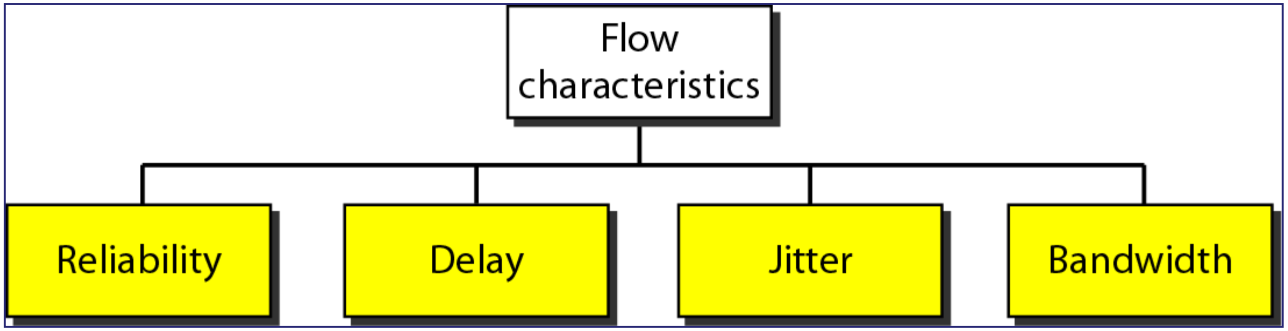
Wenzhong Li
Nanjing University

Chapter 6.
Congestion Control and QoS

- Network Congestion
- Congestion Control in FR
- Traffic Management in ATM
- Internet QoS
- Resource Allocation and RSVP
- Differentiated Services

Internet QoS

- New additions to Internet increasing traffic
  - High volume client/server application
  - Web with large amount of graphics
  - Real time voice and video
- Must support Quality of Service (QOS) within TCP/IP
  - In place of "best-effort"
  - Add traffic control to routers
  - Provide means of requesting QOS



Traffic Requirements of Internet Apps

Application	Data Loss (Reliability)	Throughput (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 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	nearly

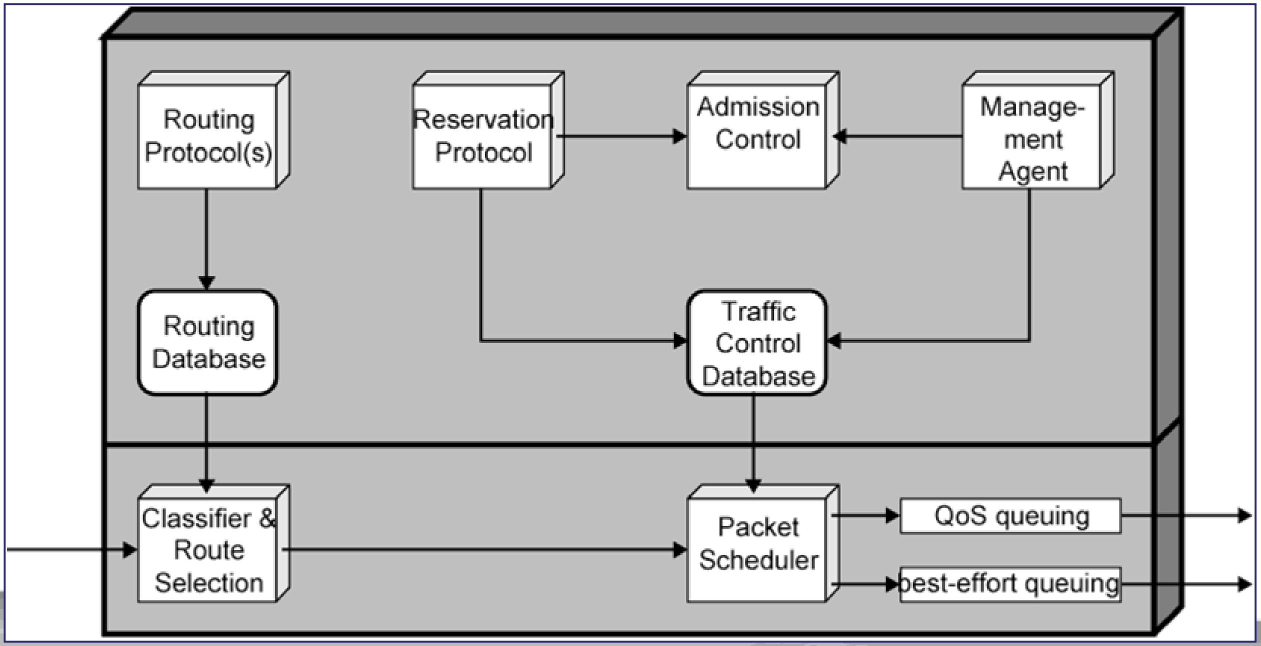
Requirements for Inelastic Traffic

- Difficult to meet requirements on IP network
  - Require preferential treatment to handle variable queuing delays and congestion
- Applications need to be able to state their requirements
  - Ahead of time, using resource reservation functions – ISA (Integrated Services Architecture)
  - Or on the fly, using fields in IP header – DS (Differentiated services)
- Require elastic traffic to be supported as well
  - Inelastic application do not back off and reduce the demand in face of congestion
  - Deny (traffic) service requests that leave too few resources

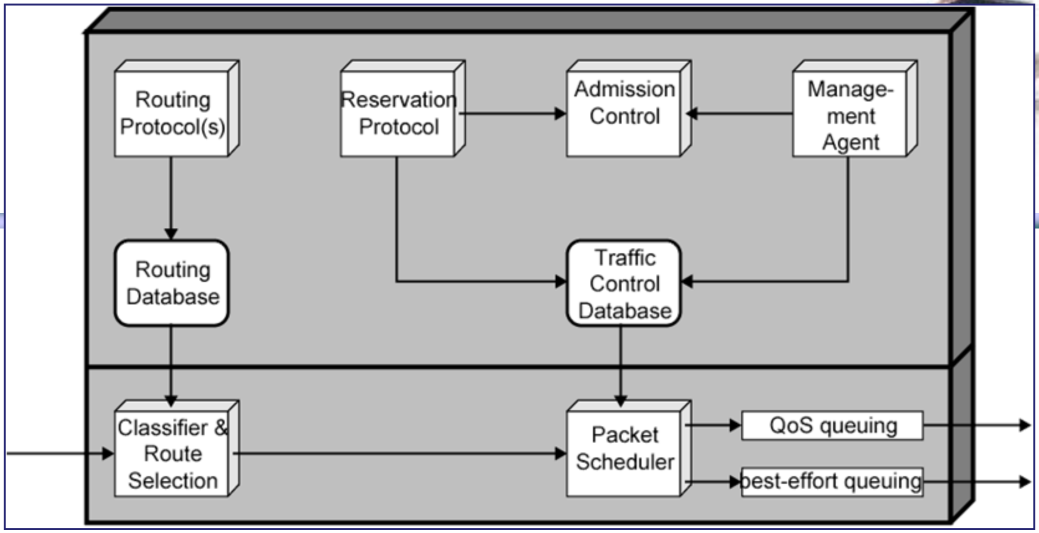
Integrated Services Architecture

- Associate a distinguishable stream of IP packets with a flow
  - With the same QOS parameters
  - Identified by source and destination IP address, port numbers, protocol type (TCP or UDP)
  - Unidirectional, Can be multicast

ISA Functions on a router

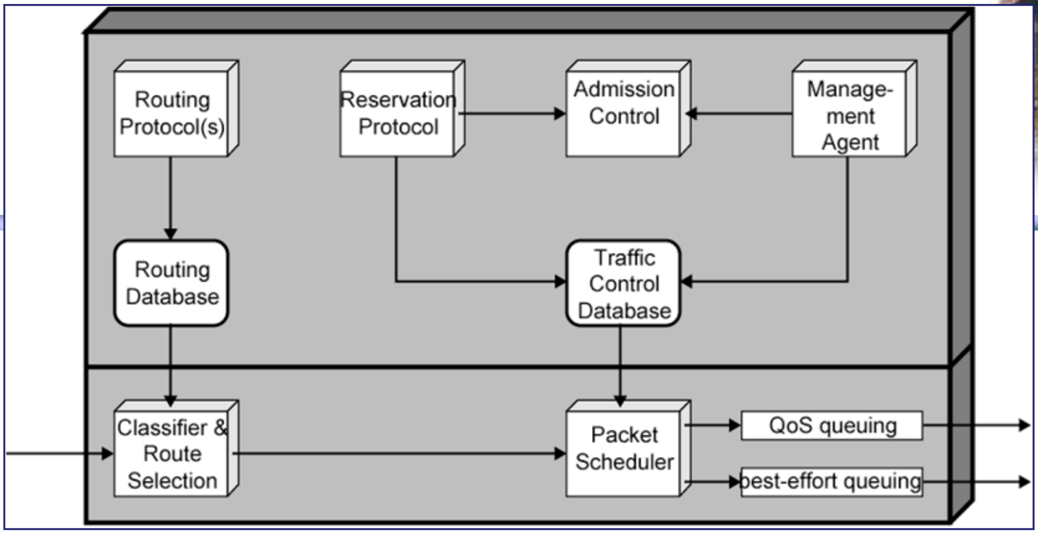


ISA Functions (1)



- Routing Algorithm
  - Link cost based on a variety of QOS parameters, not just delay
  - Routing / forwarding based on classes of flows with similar QoS
- Queuing discipline
  - Priority queuing
  - Multiple queues instead of one, taking account of different flow requirements
- Discard policy
  - Selective discard instead of just new comings

ISA Functions (2)



- Reservation protocol
  - RSVP, reserve 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
- Traffic control database
  - Parameters of traffic control
- Management agent
  - Modifies the traffic control database
  - Directs the admission control module to set policies

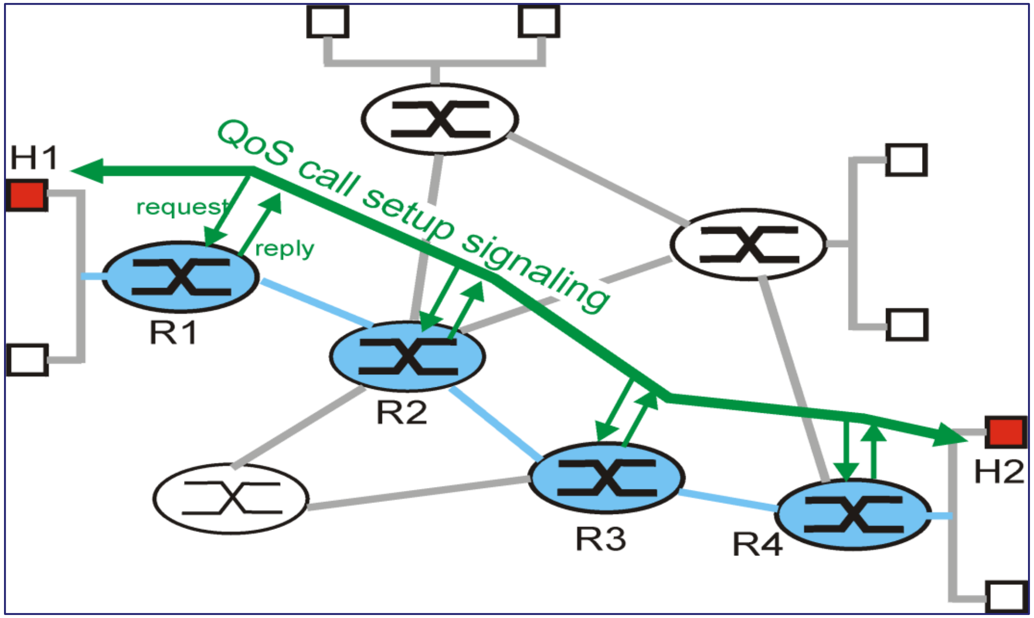
ISA Services

Defined on 2 levels

- General categories of service
  - Guaranteed service
    - Set upper bound on queuing delay through network
    - No queuing losses
    - Assume leased line
  - Controlled load service
    - No upper bound on queuing delay, but has priority scheduling on router
    - Very high percentage delivered
    - For Internet video and voice apps
  - Best effort (default)
- Particular flow within each category
  - Specified by the values of QoS parameters
  - Traffic specification (TSpec) defined for flow of packets
- Traffic that exceeds TSpec is given the best-effort service

Resource Allocation and RSVP

- Providing QOS guarantees in IP networks for individual application sessions
- Resource reservation: routers maintain state info of allocated resources for each session
  - A virtual circuit like mechanism is needed
- Admit or deny call setup requests for a new session



Data Flow in a Session

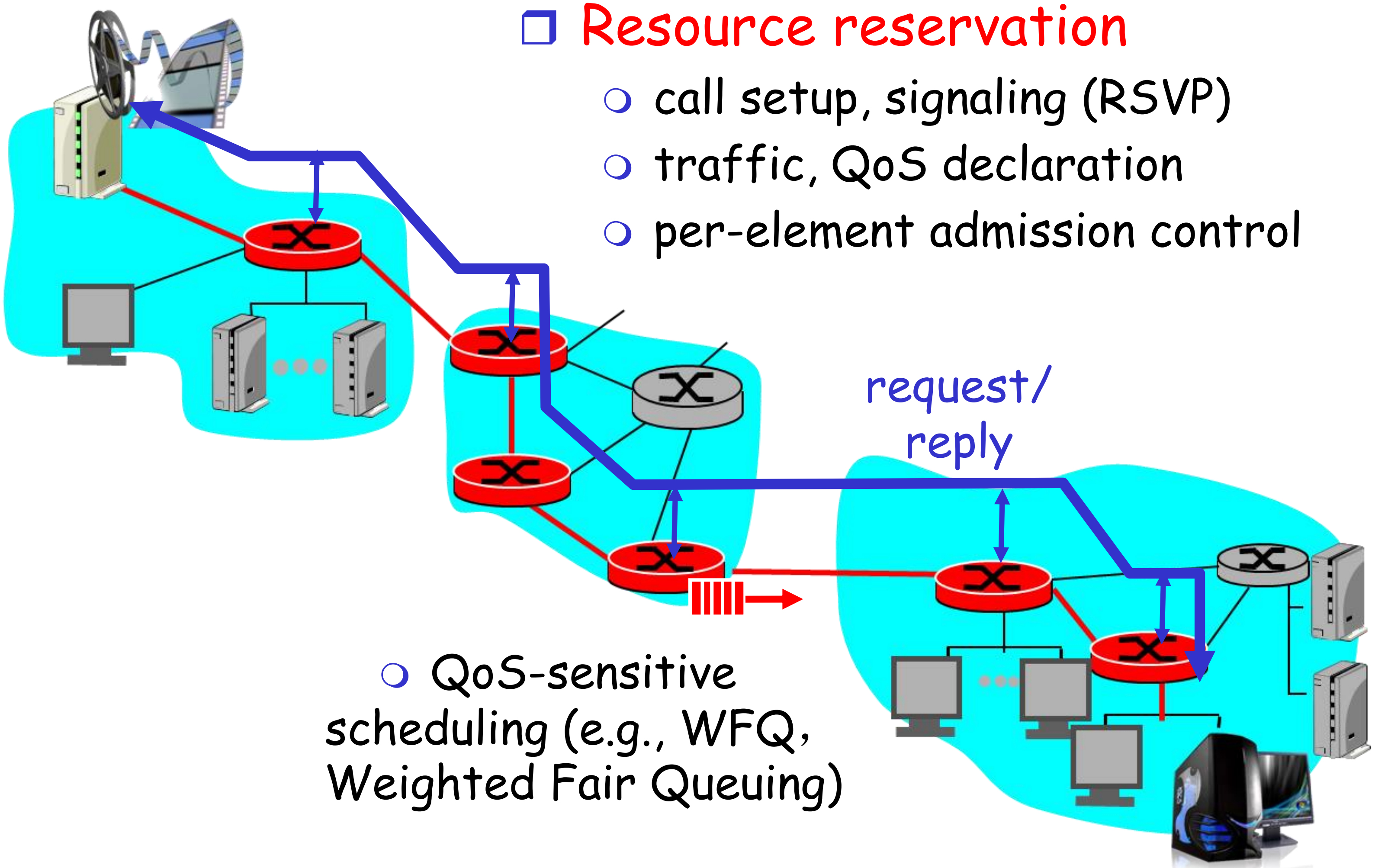
- Session
  - Unidirectional, acquired by a receiver
  - A data flow with a particular destination and transport layer protocol
  - Defined by the triple: <DestAddress, Protocol ID, DstPort>
- Reservation Model
  - An RSVP request consists of a flow descriptor: <R-Spec, T-Spec>
  - R-Spec: specifies a desired QOS, used to set parameters in the packet scheduler
  - T-Spec: defines traffic characteristics under QOS, used to set the packet classifier

Resource Reservation Protocol

IP信令协议

- RFC 2205
  - IP signaling protocol for multimedia applications
  - Reserve resources along end-to-end path (routers) for QoS support
  - Designed to operate with current and future unicast and multicast routing protocols
- Features
  - Ability for receivers to make reservations, transparent operation through non-RSVP routers
  - Support for IPv4 and IPv6, independent of routing protocol
  - Deal with changes in dynamic routes
  - Deal with changes in multicast group membership

RSVP Scenario



- Resource reservation
  - call setup, signaling (RSVP)
  - traffic, QoS declaration
  - per-element admission control
  - QoS-sensitive scheduling (e.g., WFQ, Weighted Fair Queuing)

RSVP Reservation Process

- Receivers
  - Make RSVP msgs carrying reservation requests
  - Pass the msgs upstream towards the senders
- Scope of the request
  - The set of sender hosts to which a reservation request is propagated
- At each intermediate router
  - The RSVP module passes the request to Admission and Policy control and the check is executed
  - Maintain soft state (periodically renewed) for each session
  - Reservation request is propagated upward

RSVP Mechanisms

- 2 fundamental RSVP messages
  - Path: from sender, pass downstream
  - Resv: from receiver, pass upstream
- Path message
  - 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
- 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

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

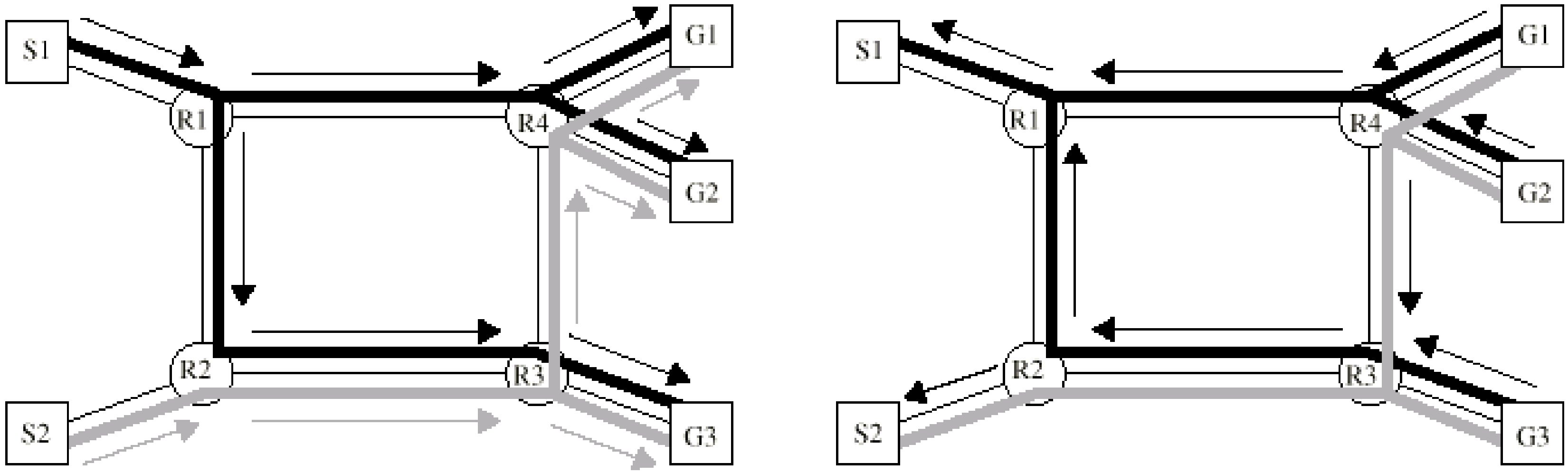
- Must interact with dynamic routing strategy of Internet
- When the route changes, the resource reservation must be changed

- Apps 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

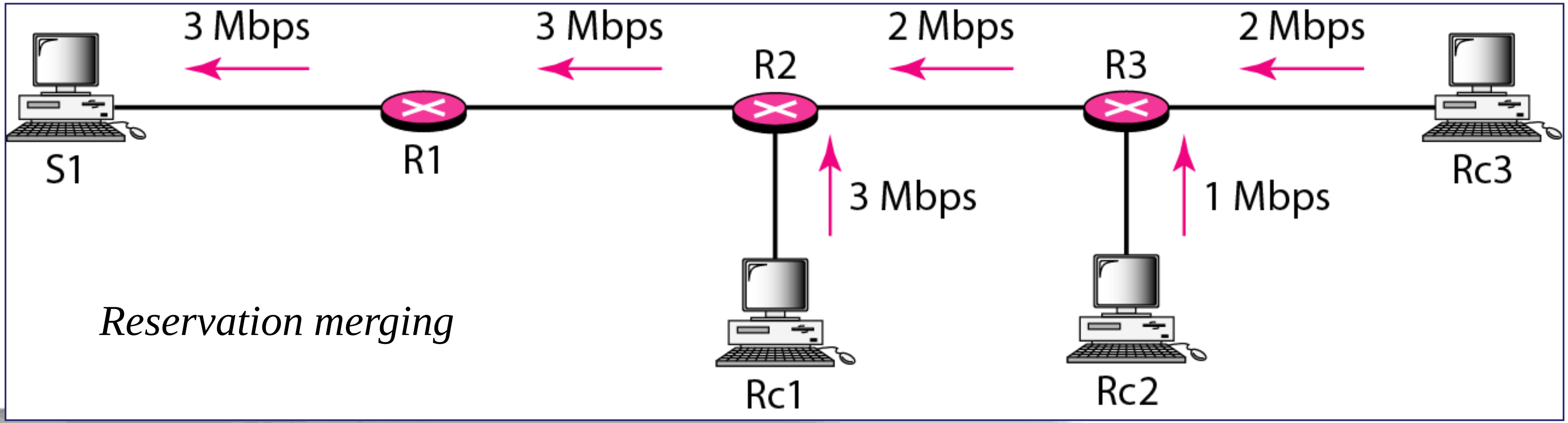
- Receiver joins a multicast group (IGMP)
- Sender-to-network signaling
  - Potential sender issues Path message
  - Receiver gets the message identifying sender
  - RSVP router records reverse path info along the mcast tree
- Receiver-to-network signaling
  - Receiver start sending Resv messages
  - Resv messages propagate through mcast tree and is delivered to sender
- Path / Reservation teardown
  - Remove sender's path state and receiver reservations

RSVP Operation for Multicast



(a) Data distrubution to a multicast group

(b) Merged Resv Messages



Disadvantage of ISA

- High cost: it need to maintain soft states for each flow, not scale well for large volumes of flows
- Complexity: the architecture is complicate, and it requires dedicate RSVP routers.
  - If there is a router in the path does not support RSVP, it will become best effort service
- Only very few service levels are defined, which is not flexible

Differentiated Services

- RFC 2475
  - Provide simple, easy to implement, low overhead mechanism
  - Support range of network services differentiated on basis of performance
  - Simple functions in network core, relatively complex functions at edge routers

Characteristics of DS

- Use IPv4 header Type of Service or IPv6 Traffic Class field
- Define Service level agreement (SLA)
  - Established between provider (ISP) and customer prior to use of DS
  - Apps only need to select appropriate DS, and all traffic with same DS field treated same
  - e.g. multiple voice connections
- Classification and Conditioning
  - Per-Hop-Behavior (PHB, queuing and forwarding) determined based on DS field
  - Edge router classifies and shapes the non-conforming traffic

Service Level Agreement

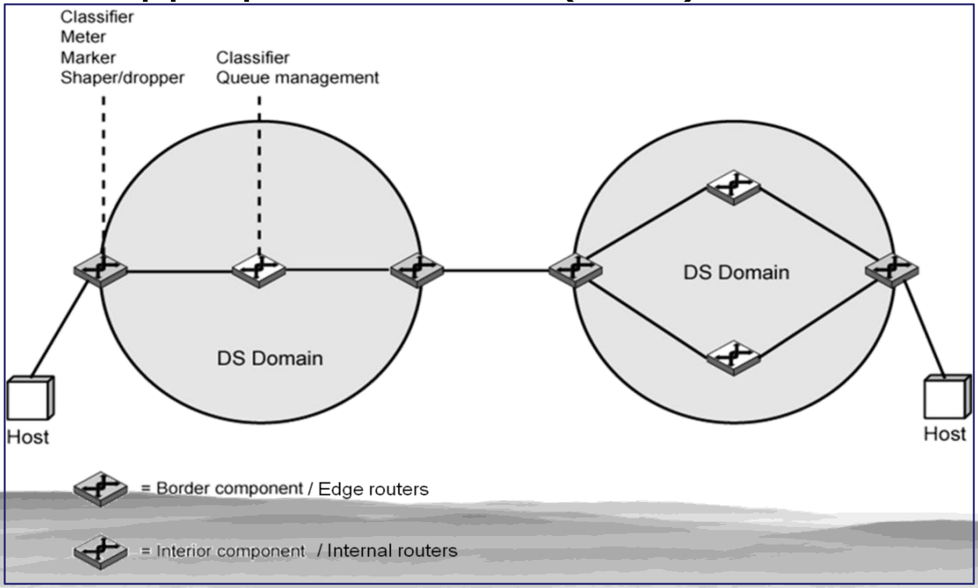
- Define service performance parameters
  - Expected throughput, latency
  - Drop probability
  - Jitters
- Define Traffic profiles to be adhered to
  - e.g. token bucket parameters
  - pre-negotiated rate r, bucket size B
- Disposition of traffic in excess of profile
- Constraints on ingress and egress points
  - Indicate scope of service

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 Domain

- Provided by singular ISP or group of ISPs
  - Contiguous portion of internet over which consistent set of DS policies administered
  - i.e. Similar explanation and handling of SLA parameters
- Service provider configures domain edge routers
  - Customer may be hosts or edge routers in other domain
  - Ongoing measure of performance provided for each class
  - Match the most appropriate service (class) for traffic from other domain



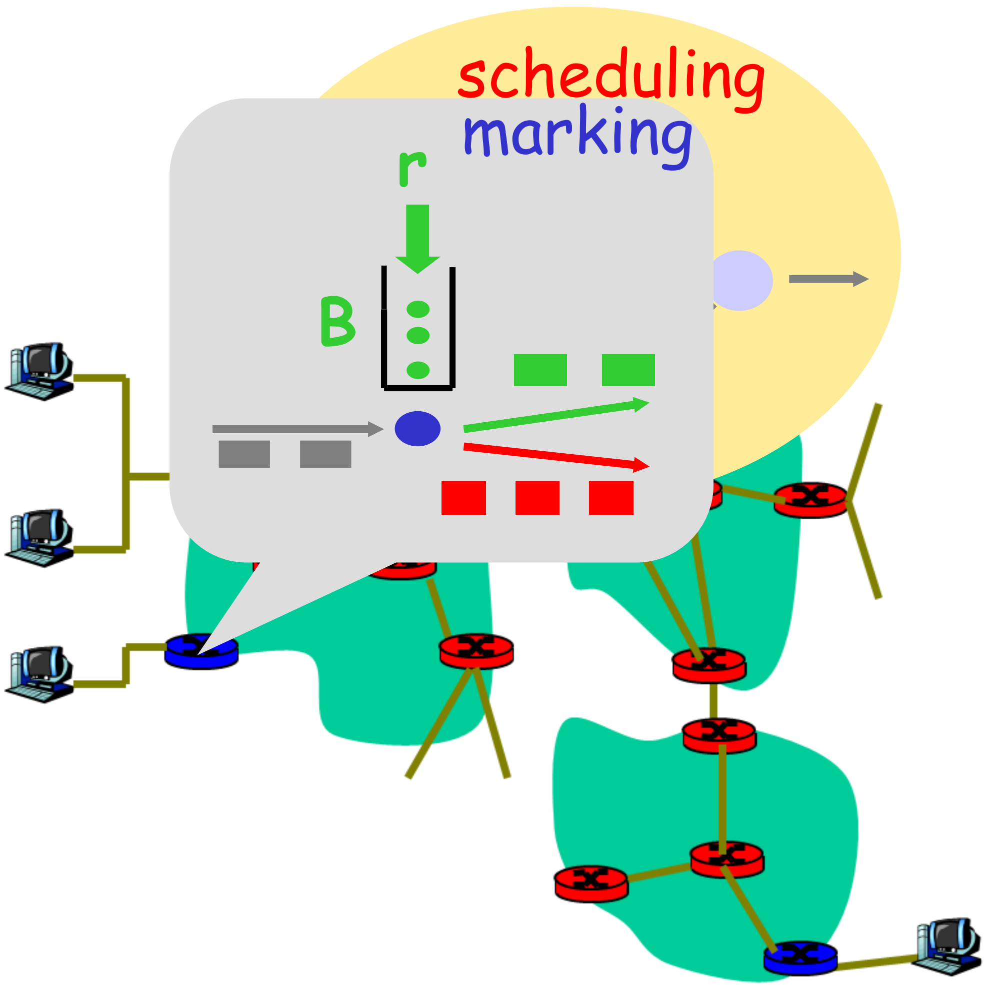
DS Architecture

Edge router

- Per-flow traffic management
- Marks packets as in-profile and out-profile

Internal router

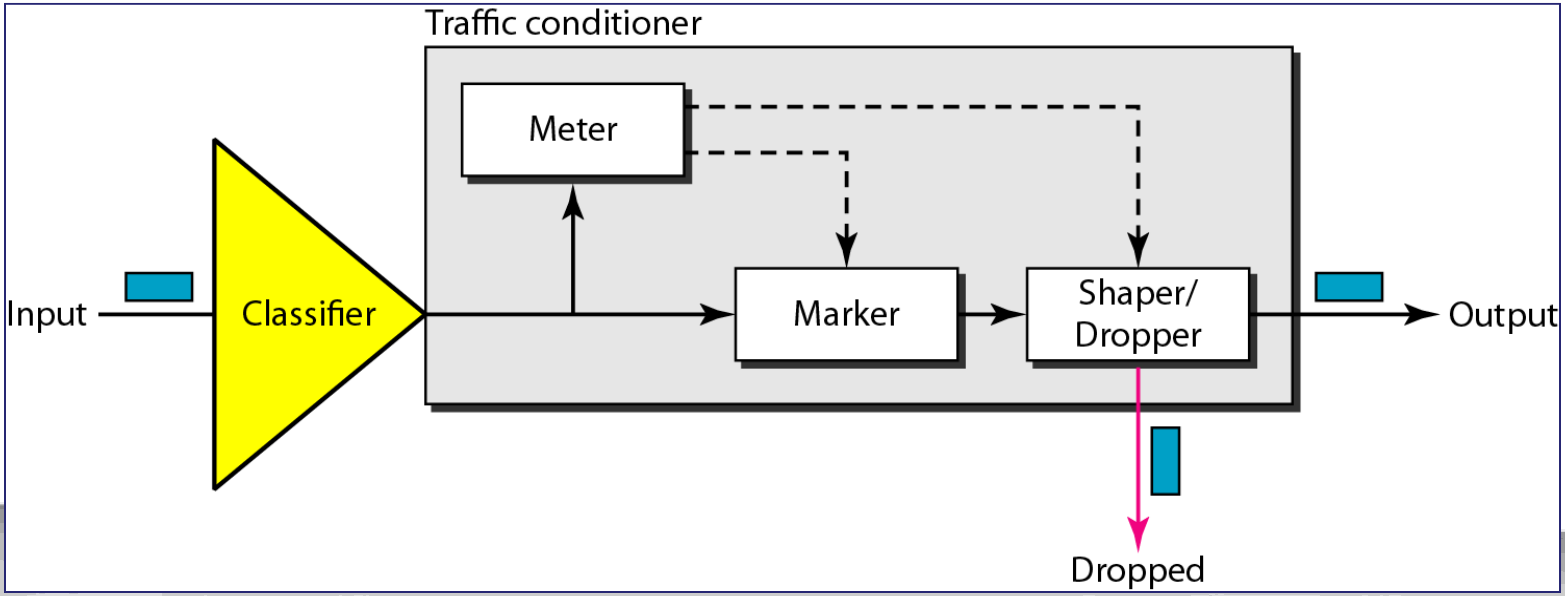
- Per class traffic management
- Buffering and scheduling based on Marking at edge
- Preference given to in-profile packets



Functions of Edge Routers

Traffic classification and conditioning per flow

- Classifier: separate packet flows into classes
- Meter: measure flow traffic for conformance to profile
- Marker: policing by remarking code-points if required
- Shaper: shaping packet flow using token bucket
- Dropper: drops packets if flow rate exceeds too much those specified in the class profile



Functions of Internal Routers

- Consistent interpretation of DS code-points within domain
  - Simple mechanisms to handle packets based on code-points (Class)
- Classifier
  - Differentiate packets based on DS code-point, src & dst addresses, high-level protocol, etc.
- Queue Management
  - Per Hop Behavior (PHB): queuing gives preferential treatment depending on code-point
  - Packet dropping rule dictates which to drop when buffer saturated

Per Hop Behavior

- For each hop (router) in a DS domain
  - Defines the policy and priority applied to a packet with specific DS
  - Results in a different observable (measurable) forwarding performance

- Examples:
  - Class A gets x% of outgoing link bandwidth over time intervals of a specified length
  - Class A packets leave first before packets from class B

- 2 PHBs under consideration
  - Expedited Forwarding
  - Assured Forwarding

PHB – Expedited Forwarding

加速型转发

- RFC 3246
  - 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 current Internet
  - Queues at each router will result in loss, delays, and jitter
  - Define the minimum departure (guaranteed) rate
  - Condition aggregate so arrival rate at internal routers is always less than minimum departure rate
  - In-profile EF traffic is given absolute queuing priority

PHB – Assured Forwarding

确保型转发

- RFC 2597, provides ranked services superior to best-effort
  - Based on explicit allocation on routers, do not require reservation of resources
- Performance of forwarding depends on
  - How much forwarding resources allocated to each class that the packet belongs to
  - Current load of the class
  - If congested within the class, drop precedence of packet
- Simplicity
  - Very little work required by internal routers
  - Marking of traffic at edge routers provides different levels of service to different classes
- Interior routers use RED algorithm to manage DS traffic

Summary

- Integrated Services Architecture
  - Architecture
  - Resource Allocation (RSVP)
    - Receivers make reservations
    - Independent of routing protocol
    - Deal with changes in dynamic routes
    - Deal with changes in multicast group membership
- Differentiated Services
  - Edge router
    - Per-flow traffic management
  - Internal router
    - Per class traffic management
    - Buffering and scheduling based on Marking at edge