Improving QOS in IP Networks

Thus far: “making the best of best effort”
Future: next generation Internet with QoS guarantees
  o RSVP: signaling for resource reservations
  o Differentiated Services: differential guarantees
  o Integrated Services: firm guarantees

simple model for sharing and congestion studies:

Principles for QOS Guarantees

Example: 1Mbps IP phone, FTP share 1.5 Mbps link.
  o bursts of FTP can congest router, cause audio loss
  o want to give priority to audio over FTP

Principle 1
packet marking needed for router to distinguish between different classes; and new router policy to treat packets accordingly
Principles for QoS Guarantees (more)

- what if applications misbehave (audio sends higher than declared rate)
  - policing: force source adherence to bandwidth allocations
- marking and policing at network edge:
  - similar to ATM UNI (User Network Interface)

Principle 2
provide protection (isolation) for one class from others

Principles for QoS Guarantees (more)

- Allocating fixed (non-sharable) bandwidth to flow: inefficient use of bandwidth if flows doesn't use its allocation

Principle 3
While providing isolation, it is desirable to use resources as efficiently as possible
Principles for QOS Guarantees (more)

- Basic fact of life: cannot support traffic demands beyond link capacity

Principle 4

Call Admission: flow declares its needs, network may block call (e.g., busy signal) if it cannot meet needs

Summary of QoS Principles

QoS for networked applications

Let's next look at mechanisms for achieving this ....
Scheduling And Policing Mechanisms

- **scheduling**: choose next packet to send on link
- **FIFO (first in first out) scheduling**: send in order of arrival to queue
  - real-world example?
  - **discard policy**: if packet arrives to full queue: who to discard?
    - Tail drop: drop arriving packet
    - priority: drop/remove on priority basis
    - random: drop/remove randomly

![Diagram of scheduling and queuing](image)

Scheduling Policies: more

**Priority scheduling**: transmit highest priority queued packet

- **multiple classes**, with different priorities
  - class may depend on marking or other header info, e.g. IP source/dest, port numbers, etc..
  - Real world example?

![Diagram of priority scheduling](image)
Scheduling Policies: still more

round robin scheduling:
- multiple classes
- cyclically scan class queues, serving one from each class (if available)
- real world example?

![Round Robin Scheduling Diagram]

Scheduling Policies: still more

Weighted Fair Queuing:
- generalized Round Robin
- each class gets weighted amount of service in each cycle
- real-world example?

![Weighted Fair Queuing Diagram]
**Policing Mechanisms**

**Goal:** limit traffic to not exceed declared parameters

Three common-used criteria:

- **(Long term) Average Rate:** how many pkts can be sent per unit time (in the long run)
  - crucial question: what is the interval length: 100 packets per sec or 6000 packets per min have same average!

- **Peak Rate:** e.g., 6000 pkts per min. (ppm) avg.; 1500 ppm peak rate

- **(Max.) Burst Size:** max. number of pkts sent consecutively (with no intervening idle)

---

**Policing Mechanisms**

**Token Bucket:** limit input to specified Burst Size and Average Rate.

- bucket can hold b tokens
- tokens generated at rate \( r \) token/sec unless bucket full
- over interval of length t: number of packets admitted less than or equal to \( r t + b \).
Policing Mechanisms (more)

- token bucket, WFQ combine to provide guaranteed upper bound on delay, i.e., QoS guarantee!

\[ D_{\text{max}} = \frac{b}{R} \]

IETF Integrated Services

- architecture for providing QoS guarantees in IP networks for individual application sessions
- resource reservation: routers maintain state info (a la VC) of allocated resources, QoS req’s
- admit/deny new call setup requests:

**Question:** can newly arriving flow be admitted with performance guarantees while not violated QoS guarantees made to already admitted flows?
**Intserv: QoS guarantee scenario**

- **Resource reservation**
  - call setup, signaling (RSVP)
  - traffic, QoS declaration
  - per-element admission control

**QoS-sensitive scheduling (e.g., WFQ)**

**Reservation Protocol**

- Consistent with robustness of today's connectionless model
- Uses soft state (refresh periodically)
- Designed to support multicast
- Receiver-oriented
- Two messages: PATH and RESV
- Source transmits PATH messages every 30 seconds
- Destination responds with RESV message
- Merge requirements in case of multicast
RSVP Example

Call Admission

Arriving session must:
- declare its QOS requirement
  - R-spec: defines the QOS being requested
- characterize traffic it will send into network
  - T-spec: defines traffic characteristics
- signaling protocol: needed to carry R-spec and T-spec to routers (where reservation is required)
  - RSVP
**Intserv QoS: Service models** [rfc2211, rfc 2212]

**Guaranteed service:**
- worst case traffic arrival: leaky-bucket-policed source
- simple (mathematically provable) bound on delay
  [Parekh 1992, Cruz 1988]

**Controlled load service:**
- "a quality of service closely approximating the QoS that same flow would receive from an unloaded network element."

---

**IETF Differentiated Services**

**Concerns with Intserv:**
- **Scalability:** signaling, maintaining per-flow router state difficult with large number of flows
- **Inflexible Service Models:** Intserv has only two classes. Also want "qualitative" service classes
  - "behaves like a wire"
  - relative service distinction: Platinum, Gold, Silver

**Diffserv approach:**
- simple functions in network core, relatively complex functions at edge routers (or hosts)
- Do't define define service classes, provide functional components to build service classes
**Diffserv Architecture**

**Edge router:**
- per-flow traffic management
- marks packets as in-profile and out-profile

**Core router:**
- per class traffic management
- buffering and scheduling based on marking at edge
- preference given to in-profile packets
- Assured Forwarding

**Edge-router Packet Marking**
- **profile:** pre-negotiated rate $A$, bucket size $B$
- packet marking at edge based on per-flow profile

Possible usage of marking:
- class-based marking: packets of different classes marked differently
- intra-class marking: conforming portion of flow marked differently than non-conforming one
Classification and Conditioning

- Packet is marked in the Type of Service (TOS) in IPv4, and Traffic Class in IPv6
- 6 bits used for Differentiated Service Code Point (DSCP) and determine PHB that the packet will receive
- 2 bits are currently unused

may be desirable to limit traffic injection rate of some class:
- user declares traffic profile (e.g., rate, burst size)
- traffic metered, shaped if non-conforming
Forwarding (PHB)

- PHB result in a different observable (measurable) forwarding performance behavior
- PHB does not specify what mechanisms to use to ensure required PHB performance behavior
- 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

Forwarding (PHB)

PHBs being developed:

- Expedited Forwarding: pkt departure rate of a class equals or exceeds specified rate
  - logical link with a minimum guaranteed rate
- Assured Forwarding: 4 classes of traffic
  - each guaranteed minimum amount of bandwidth
  - each with three drop preference partitions