Integrated and Differentiated Services

Christos Papadopoulos
(Remixed by Lorenzo De Carli)
CSU CS557, Fall 2017
Preliminary concepts: token buffer
Characterizing Traffic: Token Bucket Filter

- Parsimonious model to characterize traffic
- Described by 2 parameters:
  - token rate $r$: rate of tokens placed in the bucket
  - bucket depth $B$: capacity of the bucket
- Operation:
  - tokens are placed in bucket at rate $r$
  - if bucket fills, tokens are discarded
  - sending a packet of size $P$ uses $P$ tokens
  - if bucket has $P$ tokens, packet sent at max rate, else must wait for tokens to accumulate
Token Bucket Operation

- **tokens**
- **overflow**
- **Packet**
- **tokens**
- **tokens**

- Enough tokens packet goes through, tokens removed
- Not enough tokens - wait for tokens to accumulate
Token Bucket Characteristics

- In the long run, rate is limited to $r$
- In the short run, a burst of size $B$ can be sent
- Amount of traffic entering at interval $T$ is bounded by:
  \[ \text{traffic} = B + r*T \]
- Information useful to admission algorithm
## Token Bucket Specs

<table>
<thead>
<tr>
<th>Time</th>
<th>Flow A</th>
<th>Flow B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

BW (MB) vs. Time

- Flow A: 1 MB constant from 1 to 3
- Flow B: 2 MB constant from 2 to 3
Token Bucket Specs

Flow A: $r = 1$ MBps, $B=1$ byte

 BW (MB)  

1 2

Flow B

Flow A

1 2 3

Time
Token Bucket Specs

Flow A: $r = 1 \text{ MBps}$, $B=1 \text{ byte}$
Flow B: $r = 1 \text{ MBps}$, $B=1 \text{ MB}$

Diagram:
- **Flow A**: Constant rate of 1 MBps from 1 to 3 units of time.
- **Flow B**: Constant rate of 2 MBps from 2 to 3 units of time.

**Axis Labels:**
- **BW (MB)**
- **Time**
Possible Token Bucket Uses

• Shaping, policing, marking
  – delay pkts from entering net (shaping)
  – drop pkts that arrive without tokens (policing)
  – let all pkts pass through, mark ones without tokens
    • Then, network drops pkts without tokens during congestion
Preliminary concepts: RED queuing

Random Early Detection (RED)

• Motivation:
  – TCP detects congestion from loss - after queues have built up and increase delay (not good if goal is to keep queue utilization low!) (full queue problem)

• Aim:
  – keep throughput high and delay low
  – accommodate bursts

• Approach:
  – Probabilistically drop packets before congestions occurs
  – No per-flow state
Solving the Full Queues Problem

• Drop packets before queue becomes full (early drop)

• Intuition: notify senders of incipient congestion
  – example: early random drop (ERD):
    • if qlen > drop level, drop each new packet with fixed probability $p$
    • does not control misbehaving users
RED Operation

Max thresh

Min thresh

Average queue length

always drop

probabilistic drop

do not drop

P(drop)

MaxP

minthresh

maxthresh

Avg length
Integrated services (IntServ)
**Integrated Services**

- Basic idea: let applications specify whatever delay and bandwidth they desire, and network tries to satisfy the application

- Components:
  - **Service interface** between applications and network
  - **Admission Control** – which flows get in?
  - **Reservation Protocol** (e.g., RSVP) - signaling
  - **Scheduling algorithms** (e.g. Weighted Fair Queuing)

- A hot research area many years ago
  - Work has essentially stopped
  - But old ideas sometimes come back..
State of Integrated Services

• Lots of work done in the area
• We understand many of the problems
  – But no commercial interest in the technology
  – Too complex?
    • we can probably build schedulers in hardware
    • Need per-flow state for scheduling
    • Need end-to-end signaling

• Can we do something simpler?
Differentiated Services (DiffServ)
Key Ideas

• Traffic classes instead of flows
• Forwarding behaviors instead of end-to-end service guarantees
  – Tune applications to network services rather than network services to applications
  – Discrete vs. continuous space
• No resource reservation
• Somewhere between Best Effort and IntServ
Service Differentiation

• Analogy:
  – airline service, first class, coach, various restrictions on coach as a function of payment

• Best-effort expected to make up bulk of traffic, but revenue from first class important to economic base (will pay for more plentiful bandwidth overall)

• **Not motivated by real-time** but by economics and assurances
Types of Service

• **Premium service**: (type P)
  – admitted based on peak rate
  – conservative, virtual wire services
  – unused premium goes to best effort (subsidy!)

• **Assured service**: (type A)
  – based on expected capacity usage profiles
  – traffic unlikely to be dropped if user maintains profile.
    Out-of-profile traffic marked

• **Best effort**
Differences With Integrated Services

• No need for reservations: just mark packets
• Packet marking done at administrative boundaries before injecting packets into network
• Significant savings in signaling, much simpler overall
Service vs. Forwarding Treatment

• Service: end-to-end
• Forwarding treatment: hop-by-hop (at each router)
  – Reasoning: various forwarding treatments can be used to construct same e2e service
  – Free to implement treatments locally in various ways (buffer management and scheduling)
  – Example: no-loss service implemented with priority queue (but needs admission control)
Service Level Agreements

- Mostly static or long-lived. Specification:
  - Traffic profile (e.g., token bucket per class)
  - Performance metrics (throughput, delay, drop priority)
  - Actions for non-conformant packets
  - Additional marking/shaping
Where Things Happen

- Company A
  - Host
  - First hop router
  - Internal router
  - Border router

- ISP
  - Border router

Marked packets

Unmarked packet flow

Classify, police and mark

Only scheduling
No police or mark

Classify, police and mark
A Two-bit Differentiated Services Architecture for the Internet

Nichols99a
Premium vs. Assured Forwarding Behaviors

• **Premium** packets receive virtual circuit type of treatment
  – Appropriate for intolerant (of loss) and rigid (in delay) applications

• **Assured** packets receive “better than best effort” type of treatment
  – Appropriate for adaptive applications
2-bit Differentiated Service

- Precedence field encodes $P$ & $A$ type packets
- $P$ packets are BW limited, shaped and queued at higher priority than ordinary best effort
- $A$ packets treated preferentially wrt dropping probability in the normal queue
- Leaf and border routers have input and output tasks - other routers just output
Leaf Router Input Functionality

Markers: service class, rate, permissible burst size
Marker Function in Routers

- Leaf routers have traffic profiles - they classify packets based on packet header
- If no profile present, pass as best effort
- If profile is for A:
  - mark in-profile packets with A, forward others unmarked
- If profile is for P:
  - delay out-of-profile packets to shape into profile
Markers to Implement Two Different Services

Packet input → Wait for token → Set P bit → Packet output

Drop on overflow
Markers to Implement Two Different Services

Packet input → Wait for token → Set P bit → Packet output

Packet input → Test if token

No token:

Token:

Packet output

Drop on overflow
Output Forwarding

- 2 queues: P packets on higher priority queue
- Lower priority queue implements RED “In or Out” scheme (RIO)
- At border routers profile meters test marked flows:
  - drop P packets out of profile
  - unmark A packets
Router Output Interface for Two-bit Architecture

P-bit set?

yes

High-priority Q

no

If A-bit set
incr A_cnt

Low-priority Q

RIO queue management

If A-bit set
decr A_cnt

Packets out

30
Red With In or Out (RIO)

- For Assured Services
- Similar to RED, but with two separate probability curves
- Has two classes, “In” and “Out” (of profile)
- “Out” class has lower Minthresh, so packets are dropped from this class first
- As avg queue length increases, “in” packets are dropped
RIO Drop Probabilities

$P(\text{drop})$

Max $P_{\text{out}}$

Max $P_{\text{in}}$

Min $P_{\text{out}}$

Min $P_{\text{in}}$

AvgLen

32
RIO Drop Probabilities

More drop probability curves (WRED)?
Border Router Input Interface Profile

Meters

- Arriving packet
- Is packet marked?
  - A set
  - Token available?
    - no
    - P set
      - Token available?
        - no
          - Drop packet
        - token
          - Forwarding engine
    - token
      - Clear A-bit
Signaling

- **Where?**
  - static (long-term):
    - done out-of-band
  - dynamic:
    - from leaf to *Bandwidth Broker*
    - and from BB in one domain to another BB

- **How?**
  - not clear, but maybe RSVP
Signaling: BBs
Diffserv V.S. Intserv Summary

- Resources to aggregated traffic, not flows
- Traffic policing at the edges, class forwarding in the core
- Define forwarding behaviors, not services
- Guarantees by provisioning and SLAs, not reservations
- Focus on single domain, not e2e (need BBs for e2e)