QoS for Real Time Applications over Next Generation Data Networks

Dr. Arjan Durresi
The Ohio State University
Columbus, OH 43210
durresi@cis.ohio-state.edu
These slides are available at
http://www.cis.ohio-state.edu/~durresi
Overview

- Task 4: Improving Explicit Congestion Notification with the Mark-Front Strategy

- Task 5: Multiplexing VBR over VBR

- Task 6: Achieving QoS for TCP traffic in Satellite Networks with Differentiated Services
Task 4: Buffer Management using ECN

- Explicit Congestion Notification
- Standardized in RFC 2481, expected to be widely implemented soon.
- Two bits in IP header: Congestion Experienced, ECN Capable Transport
- Two bits in TCP header: ECN Echo, Congestion Window Reduced
No standard action specified

One possibility is to randomly mark at lower congestion levels. Random Early Detection (RED) marking instead of dropping.
Simulation Model

TCP senders

TCP receivers

router

router
Threshold & Buffer Requirement

- In order to achieve full link utilization and zero packet loss, ECN routers should
  - have a buffer size of three times the bandwidth delay product;
  - set the threshold as 1/3 of the buffer size;
- Any smaller buffer size will result in packet loss.
- Any smaller threshold will result in link idling.
- Any larger threshold will result in unnecessary delay.
Setting the Threshold

![Graph showing link utilization vs threshold]

Threshold \( T = rd \)
Buffer Requirement for No Loss

ECN buffer size requirement

Threshold

Theoretical
Simulation
Problem with Mark Tail

- Current ECN marks packets at the tail of the queue:
  - congestion signals suffer the same delay as data;
  - when congestion is more severe, the ECN signal is less useful;
  - Flows arriving when buffers are empty may get more throughput $\Rightarrow$ unfairness.
Proposal: Mark Front

- Simulation Analysis has shown that mark-front strategy
  - reduces the buffer size requirement, from 3rd to 2nd
  - increases link efficiency,
  - avoids lock-in phenomenon and improves fairness,
  - alleviates TCP’s discrimination again large RTTs.

The Ohio State University
Arjan Durrezi
Our theoretical bound is tight when threshold $\geq \text{rd}$

- Mark Front requires less buffer
Mark-Front improves the link efficiency for the same threshold
Unfairness

Mark-Front improves fairness
Task 4: Summary of Results

- Mark-front strategy
  - reduces the buffer size requirement,
  - increases link efficiency,
  - avoids lock-in phenomenon and improves fairness
- This is specially important for long-bandwidth delay product networks
Task 5: Multiplexing

- Internet is becoming the new infrastructure for telecommunications.
- Everything over IP and IP over everything
- IP has to provide one key function: Decomposition of high-capacity channels into hierarchically ordered sub-channels
- Or how to build a multiplexing node?
Multiplexing (cont.)

- Real-time or high QoS traffics include voice over IP, virtual leased line, video over IP etc.

- Goal and Objectives:
  - Provide the needed QoS
  - High network resource utilization or multiplexing gain
  - “Simple” solutions to be deployed extensively in practice
Problems

- How to characterize IP traffic to be aggregated:
  - Different models and parameters
  - Models should be “practical” i.e. easy to be used by applications
- How to characterize the output traffic
- Provide multiplexing rules: balancing QoS for the aggregates and multiplexing gain: find an optimized solutions
- Select the appropriate scheduling mechanism
Optimized Solution

- Find the best rules how to multiplex different IP traffics
How to characterize IP traffic flows

- Envelope-Based methods to specify the traffic, both deterministically or stochastically
- Leaky Bucket Model
  - Using Leaky Bucket Model to do simulation
- D-BIND/H-BIND Model
Envelop Concepts

- $E(t) \geq A[s, s+t]$, for all $t > 0$, and $s > 0$
- Empirical Envelope is the tightest envelope for a given traffic
- Definition: Let $A(t)$ be the arriving traffic, then
  \[ E(t) = \max_{s>0} A[s, s+t], \text{ for all } t > 0 \]
  is the Empirical Envelope of $A(t)$. $A[t1, t2]$ represents the amount of traffic arrives during interval $[t1, t2]$
- Mathematical tool: Network Calculus to do envelope based derivation and analysis
Leaky Bucket Model

- Most research work uses this model to characterize the arriving traffic
- \( E(t) = b + r*t \), \( b \) is the bucket size, and \( r \) is the leaking rate
- \( E(t) \geq A[s, s+t] \), where \( A(t) \) is the arriving traffic

- Multiple Leaky Bucket Model
- \( E^*(t) = \min_{1\leq i \leq n} \{b_i + r_i*t\} \)
D-BIND/H-BIND Model

- Proposed by Edward Knightly and Hui Zhang
- With the D-BIND model, sources characterize their traffic to the network via multiple rate-interval pairs, \((R_k, I_k)\), where the rate \(R_k\), is a bounding or worst-case rate over every interval of length \(I_k\). With \(P\) rate-interval pairs, the model parameterizes a piece-wise linear constraint function with \(P\) linear segments given by

\[
E(t) = \frac{(R_k*I_k - R_{k-1}*I_{k-1})(t-I_k)/(I_k - I_{k-1}) + R_k*I_k, I_{k-1} < t <= I_k \text{ with } I_0 = 0}{I_k}
\]
D-BIND/H-BIND Model (Cont.)

- Simulations have showed that, P, the number of pairs needs not to be too large. P=4 is good.
- Better performance than Leaky Bucket Model
- Better describe the correlation structure and burstiness properties for a given traffic (Video, etc)
- Drawbacks:
  - Larger number of parameters to characterize the traffic
  - Unrealistic to let users to accurately specify such parameters, need some on-line traffic measurement
Multiple Leaky Bucket Model is a special case of D-BIND

H-BIND extends D-BIND

It uses D-BIND to characterize traffic, and achieves statistical multiplexing
How to get characteristics of aggregates of several IP flows

- Deterministic Approaches
- Stochastic/Statistical Approaches
Deterministic Approaches

- Consider worst case
- Provide 100% QoS guarantees
- Drawback:
  - waste network resource
  - bandwidth utilization is low (under 50% with D-BIND)
Stocastic/Statistical Approaches

- No 100% guarantee
- Probabilistic guarantee. For example, to guarantee packet loss rate is smaller than $10^{-6}$.
- Better network utilization.
Stochastic/Statistical Approaches (Cont.)

- Connection Admission Control Algorithms for statistical QoS.
  - Average/Peak Rate combinatorics
  - Additive Effective Bandwidths
  - Loss Curve
  - Maximum Variance Approaches
  - Refined Effective Bandwidths and Large Deviations
  - Measurement based algorithms.
  - Enforceable statistical service
  - Algorithms for special-purpose system (video on demand)
CBR vs. VBR

- IP CBR: Simple multiplexing rules, Provide guarantee for QoS, No multiplexing gain
- IP VBR, More complex multiplexing rules, Guarantees for QoS depend on multiplexing rules, More multiplexing gain:
  - With CBR overbooking is performed by burst absorption at the multiplexer.
  - With VBR it is possible to let burst go through the multiplexer and count on statistical multiplexing inside the network, where the number of connections and the trunk bit rates are larger
Effective Bandwidth:
The queue with constant rate $C$ guarantees a delay bound $D$ to a flow with arrival curve $\alpha$ if $C \geq e_D(\alpha)$ where:

$$e_D(\alpha) = \sup \alpha(s)/(s+D) \text{ for } s \geq 0$$

Example: For IETF traffic specification:

$$e_D = \max \{M/D, r, p, (1-(D-M/p)/(x+D))\}$$

where $x = (b-M)/(p-r)$

$$\sum e_D(\alpha_i) - e_D(\sum \alpha_i) \text{ is non statistical multiplexing gain}$$
Output - CBR (cont.)

- Equivalent capacity

Bound the buffer B:

\[ C \geq f_B(\alpha) = \sup (\alpha(s) - B)/s; \ s \geq 0 \]

- Again: \( f_B(\alpha) \leq \sum f_{B_i}(\alpha_i) \)

- Also given a predicted traffic we can find the optimal VBR parameters that can carry the traffic.
For $\sigma(t) = \min(Pt, St+B)$

P: Peak rate, S: sustainable rate, B: burst tolerance

\[
\begin{align*}
(1) \quad (s+D)P & \geq \alpha(s) \\
(2) \quad (s+D)S+B & \geq \alpha(s)
\end{align*}
\]

From (1) $\Rightarrow P \geq P_0 = e_D(\alpha)$

Using a VBR trunk rather than a CBR is all benefit since, by definition of effective bandwidth, the CBR has at least a rate $P_0$

We can also optimize S and B.
Given a predicted traffic we can find the optimal VBR parameters that can carry the traffic.

A number of flows, with an aggregated curve $\alpha$, is multiplexed into a VBR trunk.

The VBR trunk is viewed as a single stream by downstream nodes and is constrained by an arrival curve $\sigma$.

If the constraint at the multiplexor is to guarantee a maximum delay $D$:

$\sigma(s + D) \geq \alpha(s)$ for $s \geq 0$
Each source is leaky bucket regulated: token rate $\rho$, token size $\sigma$ and the peak rate $P$.

- We can calculate queue length.
- We can build the optimal Buffer/Bandwidth curve.
The impact of burstiness when heterogeneous sources are multiplexed $C=45Mbps$, $\rho_1 = \rho_2 = 0.15$, $P_1=1.5$, $P_2=6$, $T_{on1} >> T_{on2}$

Statistical service with even small loss probability increases the multiplexing gain.
Advantages of using VBR trunks:

- More statistical multiplexing gain than CBR
- With CBR overbooking is performed by burst absorption at the multiplexer.
- With VBR it is possible to let burst go through the multiplexer and count on statistical multiplexing inside the network, where the number of connections and the trunk bit rates are larger.
Loss prob. = $10^{-9}$, Peak = 115 Mbps, 
M (sustained) = Peak / 1.5, B (bust tolerance) = 5 Mb, 
X (buffer) = 10 Mb

Input VBR flows: p = 2 Mbps, b = 0.5 Mb 
m = 0 to p

Figure shows the advantage of statistical multiplexing.
Connection Grouping

Compare three types of multiplexing:

1) No connection grouping
2) Connections grouped as CBR trunks
3) Connections grouped as VBR trunks
Performance of VBR over VBR traffic aggregation can significantly exceed the performance of CBR aggregation.
VBR Multiplexor

N. conn.

Mean rate $m = 0.1$

Ratio P/M

The Ohio State University

Arjan Durresi
Task 5: Summary

- Why IP multiplexing?
- Models for input and output flows and multiplexing rules: find an optimized solution
- Leaky Bucket Model
- D-BIND/H-BIND Model
- Deterministic Approaches
- Stochastic/Statistical Approaches
- VBR over VBR better than CBR over CBR
- Simulations using Leaky Bucket Model
Task 6: Study of Assured Forwarding in Satellite Networks

- Key Variables
- Buffer Management Classification: Types of RED
- Traffic Types and Treatment
- Level of Reserved Traffic
- Two vs Three: Best Results
- Summary
Key Variables

First Class

Business Class

Coach Class
Differentiated Services

- DiffServ to standardize IPv4 ToS byte’s first six bits
- Packets get marked at network ingress
- Marking ⇒ treatment (behavior) in rest of the net
- Six bits ⇒ 64 different per-hop behaviors (PHB)

<table>
<thead>
<tr>
<th>Ver</th>
<th>Hdr Len</th>
<th>Type of Service (ToS)</th>
<th>Tot Len</th>
</tr>
</thead>
<tbody>
<tr>
<td>4b</td>
<td>4b</td>
<td>8b</td>
<td>16b</td>
</tr>
</tbody>
</table>
DiffServ (Cont)

- Per-hop behavior = % of link bandwidth, Priority
- Services: End-to-end. Voice, Video, ...
  - Transport: Delivery, Express Delivery,...
    Best effort, controlled load, guaranteed service
- DS group will not develop services
  They will standardize “Per-Hop Behaviors”
- Marking based on static “Service Level Agreements” (SLAs). Avoid signaling.
Expedited Forwarding

- Also known as “Premium Service”
- Virtual leased line
- Similar to CBR
- Guaranteed minimum service rate
- Policed: Arrival rate < Minimum Service Rate
- Not affected by other data PHBs
  \[ \Rightarrow \text{Highest data priority (if priority queueing)} \]
- Code point: 101 110
Assured Forwarding

- PHB Group
- Four Classes: No particular ordering
- Similar to nrt-VBR/ABR/GFR
Key Variables

- **Bandwidth Management**:  
  - Number of colors: One, Two, or Three  
  - Percentage of green (reserved) traffic: Low, high, oversubscribed

- **Buffer Management**:  
  - Tail drop or RED  
  - RED parameters, implementations

- **Traffic Types and their treatment**:  
  - Congestion Sensitivity: TCP vs UDP  
  - Excess TCP vs Excess UDP

- **Network Configuration**:  
  Our goal is to identify results that apply to all configs.
Buffer Management Classification

- Accounting (queued packets): Per-color, per-VC, per-flow, or Global Multiple or Single
- Threshold: Single or Multiple
- Four Types:
  - Single Accounting, Single threshold (SAST)
  - Single Accounting, Multiple threshold (SAMT)
  - Multiple Accounting, Single threshold (MAST)
  - Multiple Accounting, Multiple threshold (MAMT)
Types of RED

  - Used in present diffserv-unaware routers

- Single Accounting Multiple Threshold (SAMT): Color-Aware RED as implemented in some products
  - Used in this study
Types of RED (Cont)

- Multiple Accounting Single Threshold (MAST):
  - G, G+Y, G+Y+R Queue
  - Used in our previous study

- Multiple Accounting Multiple Threshold (MAMT):
  - R, Y, G Queue

Conclusion:
More Complexity
⇒ More Fairness
Traffic Types and Treatment

- Both TCP and UDP get their reserved (green) rates
- Excess TCP competes with excess UDP
- UDP is aggressive
  - UDP takes over all the excess bandwidth
  - Give excess TCP better treatment than excess UDP

The Ohio State University
Two Drop Precedences

- All packets up to CIR are marked Green
- Overflowed packets are marked Red

TCP/UDP

Committed Burst Size (CBS)
Committed Information Rate (CIR)

Green Red
Three Drop Precedences

- Tokens in Green, Yellow buckets are generated independently.
- Parameters: Token generation rate and Bucket Size for Green and Yellow buckets
- Color Aware ⇒ Excess packets overflow to next color
Level of Reserved Traffic

- Percentage of reserved (green) traffic is the most important parameter

- If the green traffic is high
  - ⇒ No or little excess capacity
  - ⇒ Two or three colors perform similarly

- If the green traffic is low
  - ⇒ Lots of excess capacity
  - ⇒ Behavior of TCP vs UDP impacts who gets excess
  - ⇒ Need 3 colors + Need to give excess TCP yellow
  + Need to give excess UDP red colors
Simulation Configuration

- TCP_1
- TCP_5
- UDP

Customer

1
2
9
10

Satellite

R_1
R_2

Snk_1
Snk_46

1 μs | 5 μs | 250 ms | 5 μs
10 Mbps | 1.5 Mbps | 1.5 Mbps | 1.5 Mbps
## Link Parameters

<table>
<thead>
<tr>
<th>Link B/W</th>
<th>Link Delay</th>
<th>Link Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between TCP/UDP &amp; Customer i</td>
<td>10 Mbps</td>
<td>1 µs</td>
</tr>
<tr>
<td>From Customer i to R1</td>
<td>1.5 Mbps</td>
<td>5 µs</td>
</tr>
<tr>
<td>From R1 to Customer i</td>
<td>1.5 Mbps</td>
<td>5 µs</td>
</tr>
<tr>
<td>From R1 to R2</td>
<td>1.5 Mbps</td>
<td>250 µs</td>
</tr>
<tr>
<td>From R2 to R1</td>
<td>1.5 Mbps</td>
<td>250 µs</td>
</tr>
<tr>
<td>Between R2 and sink i</td>
<td>1.5 Mbps</td>
<td>5 µs</td>
</tr>
</tbody>
</table>
Simulation Parameters

- Single Accounting Multiple Threshold RED
- RED Queue Weight for All Colors: \( w = 0.002 \)
  \[
  Q_{\text{avg}} = (1-w)Q_{\text{avg}} + w Q
  \]
- Maximum Queue Length (For All Queues): 60 packets
- TCP flavor: Reno
- TCP Maximum Window: 64 packets
- TCP Packet Size: 576 bytes
- UDP Packet Size: 576 bytes
- UDP Data Rate: 1.28Mbps
## Two Color Simulations

<table>
<thead>
<tr>
<th>Simulation Configuration</th>
<th>Green Token Generation Rate [kbps]</th>
<th>Green Token Bucket Size (in Packets)</th>
<th>Maximum Drop Probability {Green, Red}</th>
<th>Drop Thresholds {Green, Red}</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Through 1152</td>
<td>12.8, 25.6, 38.4, 76.8, 102.4, 128, 153.6, 179.2</td>
<td>1, 2, 4, 8, 16, 32</td>
<td>{0.1,0.1}, {0.1,0.5}, {0.5,0.5}, {0.1,1}, {0.5,0.5}, {1,1}</td>
<td>{40/60,0/10}, {40/60,0/20}, {40/60,0/5}, {40/60,20/40}</td>
</tr>
</tbody>
</table>
# Three Color Simulations

<table>
<thead>
<tr>
<th>Simulation Config.</th>
<th>Green Token Gener. Rate [kbps]</th>
<th>Green Token Bucket Size in Packets</th>
<th>Yellow Token Bucket Size in Packets</th>
<th>Max Drop Probability {Green, Yellow, Red}</th>
<th>Drop Threshold {Green, Yellow, Red}</th>
<th>Yellow Token Gener. Rate [kbps]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Through 2880</td>
<td>12.8, 25.6, 38.4, 76.8</td>
<td>1, 2, 4, 8, 16, 32</td>
<td>1, 2, 4, 8, 16, 32</td>
<td>{0.1,0.5,1} {0.1,1,1} {0.5,0.5,1} {0.5,1,1} {1,1,1}</td>
<td>{40/60,20/40, 0/10} {40/60,20/40, 0/20}</td>
<td>128, 12.8</td>
</tr>
</tbody>
</table>
Fairness Index

- Measured Throughput: \((T_1, T_2, ..., T_n)\)
- Use any criterion (e.g., max-min optimality) to find the Fair Throughput \((O_1, O_2, ..., O_n)\)
- Normalized Throughput: \(x_i = T_i/O_i\)

\[
\text{Fairness Index} = \frac{(\sum x_i)^2}{n\sum x_i^2}
\]

Example: 50/50, 30/10, 50/10 \(\Rightarrow\) 1, 3, 5

\[
\text{Fairness Index} = \frac{(1+3+5)^2}{3(1^2+3^2+5^2)} = \frac{9^2}{3(1+9+25)} = 0.81
\]
ANOVA

- Analysis of Variance (ANOVA) - Statistical tool
- Most Important Factors Affecting Fairness and Throughput:
  - What % of the Variation is explained by Green (Yellow) rate?
  - What % of the Variation is explained by Bucket Size?
  - What % of the Variation is explained the Interaction between Green (Yellow) Rate and Bucket Size
ANOVA For 2 Color Simulations

- Most Important Factors Affecting Fairness:
  - Green Rate (Explains 65.6% of the Variation)
  - Bucket Size (Explains 19.2% of the Variation)
  - Interaction between Green Rate and Bucket Size (Explains 14.8% of the Variation)
ANOVA For 3 Color Simulations

- Most Important Factors Affecting Fairness:
  - Yellow Rate (Explains 41.36% of the Variation)
  - Yellow Bucket Size (Explains 28.96% of the Variation)
  - Interaction Between Yellow Rate And Yellow Bucket Size (Explains 26.49% of the Variation)
Two vs Three: Best Results

Fairness

Three colors

Two colors

Reservation

The Ohio State University
1. The key performance parameter is the level of green (reserved) traffic
2. If reserved traffic level is high or if there is any overbooking, two and three colors give the same throughput and fairness
3. If the reserved traffic is low, three colors give better fairness than two colors
4. Classifiers have to distinguish TCP and UDP:
   Reserved TCP/UDP $\Rightarrow$ Green, Excess TCP $\Rightarrow$ Yellow, Excess UDP $\Rightarrow$ Red
5. RED parameters and implementations have significant impact.