QoS Amplification Research

Kihong Park
Network Systems Lab
Dept. of Computer Sciences
Purdue University
park@cs.purdue.edu

http://www.cs.purdue.edu/nsl
Goal  Achieve QoS amplification over imperfect network service substrate

→  end-to-end control & per-hop control

- End-to-end QoS amplification
  - Multiple time scale traffic control
  - Adaptive redundancy control
  - Adaptive label control
- Per-hop QoS amplification
  - Aggregate-flow label switching
  - Optimal classifiers
  - WAN experiments and collaborations
Multiple Time Scale Traffic Control

Self-similar Network Traffic

- Data traffic is fundamentally different from telephony traffic (Leland et al. ’93)
  - self-similar or long-range dependent

- Causality
- Performance Impact
- Control


Network Systems Lab
Multiple Time Scale (cont.)

- Causality
  - Single-source causality (e.g., MPEG video)
Multiple Time Scale (cont.)

- Structural causality

Network Systems Lab
Multiple Time Scale (cont.)

- Structural causality (cont.)

→ UNIX file system (G. Irlam)
Structural causality (cont.)

→ impervious to "details"
Multiple Time Scale (cont.)

- Structural causality (cont.)

- on/off traffic (0/1 reward renewal process)
- asymptotic second-order self-similarity

- Two principal traits
  - Invariant correlation structure across multiple time scales
  - Correlation at a distance (long-range dependence)
Multiple Time Scale (cont.)

- Detrimental performance impact: queueing

- polynomial (vs. exponential) queue length distribution
- infinite memory/asymptotic analysis

Network Systems Lab
Multiple Time Scale (cont.)

- Empirical validation with feedback control (e.g., TCP)
Multiple Time Scale (cont.)

- Impact of long-range structure can be curtailed
  - extreme: bufferless queueing
  - time horizon implied by finite memory
  - short-range correlation can dominate

- Small buffer/large bandwidth resource provisioning policy
  - statistical multiplexing
  - central limit theorem
Importance of second-order performance measures
→ e.g., jitter

- concentrated periods of over- and under-utilization
- bufferless queueing does not help
Traffic Control

- Premise: exploit long-range correlation for traffic control
  - correlation/predictability structure at large time scales

  ➔ relevant in broadband WANs with high delay-bandwidth product
Multiple Time Scale Traffic Control (cont.)

Large time scale predictability:

\[
\text{Traffic Volume } X_i^{(m)} \quad X_{i+1}^{(m)} \\
\text{Traffic Level } L_1 \quad L_2 \\
X_i \
\text{Time Block } i \quad \text{Time Block } i+1
\]

\[
\text{alpha}=1.95, T=2\text{sec} \\
\text{Relative Frequency} \\
\text{Traffic Level L2} \quad \text{Traffic Level L1}
\]

\[
\text{alpha}=1.95, T=2\text{sec} \\
\text{Relative Frequency} \\
\text{Traffic Level L2} \quad \text{Traffic Level L1}
\]
Multiple Time Scale Traffic Control (cont.)

Large time scale predictability (5 sec):

- For $\alpha = 1.05, T = 5.0\sec$
- For $\alpha = 1.95, T = 5.0\sec$

Network Systems Lab
Implications: mitigate reactive cost of feedback control

\begin{itemize}
  \item \textbf{LRD time scale} \( \rightarrow \) RTT
\end{itemize}
Multiple Time Scale Traffic Control (cont.)

Multiple time scale traffic control:

Diagram showing the components of a network system with labels for sender, SAC, feedback congestion control, and network.
Application Domains:

- Bulk data transport – congestion control
  → throughput maximization (TCP-MT)
- Real-time data transport – adaptive redundancy control
  → end-to-end QoS (AFEC-MT)
Multiple Time Scale Traffic Control (cont.)

Congestion control: TCP and rate-based

Idea:

- Modulate slope of linear increase phase in AIMD

Low Contention

- \( \lambda_H \)
- Increased Slope

High Contention

- \( \lambda_L \)
- Decreased Slope
Multiple time scale TCP (TCP-MT):

- Multiple time scale TCP (TCP-MT):
Multiple time scale rate-based congestion control:

ATM:

- Explicit Predl. E[L\,L\,L;\,L]=L
- Aggre. Sched.
  \[ \varepsilon(i) = \frac{A - a}{h-1} (i-1) + a \]
- Feedback Congestion Control
  \[ \frac{d\lambda}{dt} = \begin{cases} \varepsilon(i) & \text{if } d\gamma/d\lambda > 0 \\ -b\lambda & \text{if } d\gamma/d\lambda < 0 \end{cases} \]

Network Systems Lab
TCP-MT: performance gain as function of RTT
TCP-MT: performance gain as function of self-similarity
Multiple Time Scale Traffic Control (cont.)

- Principal performance effect:
  - impart proactivity above and beyond AFEC
  - proactivity of reactive control in broadband WANs
  - mitigate reactive cost

  predictability at time scales exceeding RTT imparts timeliness

  ➞ applications: broadband WAN, TCP-over-Satellite
Real-time Traffic Transport

- Achieve invariant end-to-end QoS
- User-specified QoS
- ARQ infeasible (RTT & timeliness)
- Packet-level FEC
  - proactive QoS protection
- Purely end-to-end (black box network)
- MPEG video/audio implementation (UDP)
Adaptive redundancy control (AFEC):

\[ 0 \leq \gamma \leq 1 \]
Adaptive Redundancy Control (cont.)

- Redundancy-recovery relation:

\[ \gamma = f(h) \]

\( h_g, h^*, h_h \)

\[ \gamma_g = \gamma^* \]

\[ \gamma_g = \gamma^* \]

→ stability & optimality
Adaptive Redundancy Control (cont.)

- AFEC structure:
Adaptive Redundancy Control (cont.)

- Experimental set-up:
  - UltraSparc 1 & 2, SGI, x86
  - Solaris UNIX, Windows NT
  - Optibase, Futuretel MPEG I & II compression boards
  - Sony DCR-VX 1000, Panasonic F250

Network Systems Lab
Adaptive Redundancy Control (cont.)

- Impact of redundancy: Static FEC
Adaptive Redundancy Control (cont.)

- Adaptive FEC vs. static FEC
Adaptive Redundancy Control (cont.)

- Stable target QoS: symmetric control
Adaptive Redundancy Control (cont.)

- Unstable target QoS: asymmetric control

![Graphs showing measured hit rate and target hit rate over frames, and redundancy h over frames.](image-url)
Adaptive Redundancy Control (cont.)

- Multiple time scale redundancy control

---

→ level control
Adaptive Redundancy Control (cont.)

- AFEC-MT structure:

```
Sender

Time Scale 2

Explicit Prediction

C^2_S

h_2

Time Scale 1

C^1_S

h_1, h_2

MPEG Encoder

FEC Encoder

Receiver

MPEG II Player

FEC Decoder

C_R

Feedback (Implicit Prediction)

Network Systems Lab
```
Adaptive Redundancy Control (cont.)

- AFEC-MT:

  static FEC

  hit trace:

  AFEC

  AFEC-MT

Network Systems Lab
Adaptive Label Control

Motivation:

→ diverse QoS requirements
→ shared network environment
Adaptive Label Control (cont.)

Differentiated services network:

\[ n \text{ users} \geq L \text{ labels (colors)} \geq m \text{ classes} \]
Adaptive Label Control (cont.)

Questions:

- What is a “good” (optimal) per-hop control?
  → optimal aggregate-flow per-hop behavior

- What is a “good” (optimal) edge control?
Adaptive Label Control (cont.)

- What is the loss of power due to aggregation?
  - $n \Rightarrow L \geq m$
  - loss of resolution vis-à-vis per-flow switching

- What is the impact of finite, discrete label set \{1, 2, \ldots, L\}?
  - $\eta \in \mathbb{Z}_+, \mathbb{R}_+, [0,1], \text{ or } \mathbb{R}_+^s$
Adaptive Label Control (cont.)

- What is the system dynamics when driven by selfish users?
  - end-to-end label control
  - stability (Nash equilibria) and efficiency (system optimality)

- What is the impact of selfish service provider (ISP)?

Network Systems Lab
Adaptive Label Control (cont.)

Theory
- optimal PHB
  - differentiation/shaping
  - efficiency
- adaptive label control
- selfish users
- selfish service provider
- performance analysis

Simulation
QSim: WAN QoS Simulator

Implementation
Cisco 7206 VXR IP-over-SONET QoS Testbed
Purdue Infobahn

Network Systems Lab
Adaptive Label Control (cont.)

Performance Results

→ QSim: ns based WAN QoS simulation environment
Adaptive Label Control (cont.)

- Structural: bottleneck BW, $L = 16$ ($m = 16$)
Adaptive Label Control (cont.)

- Structural: $L = 1, 4, 16, 32$

Network Systems Lab
Adaptive Label Control (cont.)

- **Structural:** $\log L = 0, 1, 2, 3, 4, 5$ (bits)
Adaptive Label Control (cont.)

- Structural: system optimal BW requirement
Dynamical: adaptive label control (end-to-end) → reachability
Adaptive Label Control (cont.)

- Dynamical: adaptive label control (cont.)
Optimal aggregate-flow per-hop control:

\[ n \gg L \geq m \]

\[ \rightarrow n \text{ users, } L \text{ labels, and } m \text{ service classes} \]
Adaptive Label Control (cont.)

- Of interest: \( n \geq L \geq m \)
- Special case: \( n = m \)
  \[ \rightarrow \text{per-flow per-hop control} \]
- Of special interest: \( L = m \)
  \[ \rightarrow \text{as many service classes as label values} \]

Optimality I: service differentiation/shaping
Per-flow Control ($n = m$):

- Label value $\eta$ viewed as “code” of user requirement
  - e.g., 1.5 Mbps, relative share of link bandwidth, etc.
- If infinite resources, then no interaction/coupling
  - e.g., INDEX
- In resource-bounded systems, $\exists$ coupling (externality)
Adaptive Label Control (cont.)

- Illustration of coupling in simple single switch case:
Adaptive Label Control (cont.)

- INDEX (Varaiya et al.)

<table>
<thead>
<tr>
<th>Service Class</th>
<th>BW</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platinum Service</td>
<td>$B_{W1}$</td>
<td>$P_{rice1}$</td>
</tr>
<tr>
<td>Gold Service</td>
<td>$B_{W2}$</td>
<td>$P_{rice2}$</td>
</tr>
<tr>
<td>Silver Service</td>
<td>$B_{W3}$</td>
<td>$P_{rice3}$</td>
</tr>
<tr>
<td>Bronze Service</td>
<td>$B_{W4}$</td>
<td>$P_{rice4}$</td>
</tr>
</tbody>
</table>

→ service class: volume insensitive
→ infinite resources
→ no externality
Adaptive Label Control (cont.)

- Assume label set is metric space (totally ordered)
  - e.g., Euclidean distance ($L_2$ norm)
  - e.g., $\eta = 1 < 2 < \ldots < L$

- Mean square measure of goodness:

  Given $\eta$, find resource configuration $\omega$ s.t.

  $$\min_{\omega} \sum_{i=1}^{n} (\eta_i - \omega_i)^2$$
Adaptive Label Control (cont.)

- **GPS**: \( \omega_i = \alpha_i / \lambda_i \)

\[ \eta_i \in \{1,2,...,L\}; \quad \xi : \{1,...,L\} \to \{1,...,m\} \]
Normalization:
\[
\frac{\eta_i - \eta_{\text{min}}}{\eta_{\text{max}} - \eta_{\text{min}}} \in [0,1]
\]

Solution:
\[
\alpha_i = (1 - \nu) \frac{\lambda^i \eta^i}{\sum_k \lambda^k \eta^k} + \nu \frac{\lambda^i}{\sum_k \lambda^k}
\]
Optimal aggregate-flow classifier:

Given $\eta$, find resource configuration $\omega$ s.t.

$$\min_{\omega} \sum_{i=1}^{n} (\eta_i - \omega_i)^2$$

Optimal solution:

Reduce to per-flow optimal solution

$\rightarrow$ optimal clustering problem
Properties (A1), (A2), and (B)

- (A1) If $\eta_i$ increases, then QoS of user $i$ improves
- (A2) If $\eta_i$ increases, then QoS of user $j$ deproves
- (B) If $\eta_i \geq \eta_j$ then QoS of user $i$ is better than QoS of user $j$

Optimal per-flow classifier satisfies (A1), (A2), (B)

Optimal aggregate-flow classifier with $L = m$ satisfies (A1), (A2), (B)
Adaptive Label Control (cont.)

Overall Architecture

→ three control planes

Network Systems Lab
End-to-end QoS control: label control

- open-loop
- closed-loop
  → adaptive label control
Adaptive Label Control (cont.)

- Integrated QoS control:
  - e.g., TCP over adaptive label control
Adaptive Label Control (cont.)

Benchmark Environment

- Purdue Infobahn QoS testbed: 4 Cisco 7206 VXR routers
  - IP-over-SONET backbone
  - custom classifier implementation in Cisco IOS (Fred Baker)
- NSF vBNS and Abilene connectivity (DS-3)
  - Purdue vBNS/Internet2 Advisory Committee
  - Internet2 collaboration
- Fore ATM, FastEthernet switches
Adaptive Label Control (cont.)

Purdue Infobahn

Network Systems Lab
Adaptive Label Control (cont.)

- Real-time MPEG I & II video/audio compression engines
  → Optibase, Futuretel (Windows NT)
- Video/audio capture equipment
- 35+ Sun/Intel/SGI workstations & PCs
- Prototype software systems: UNIX, Windows NT
Adaptive Label Control (cont.)

Performance Evaluation and Benchmarking

- Internet2 benchmarking of
  - Multiple time scale traffic control (TCP-MT, AFEC-MT)
  - Adaptive redundancy control (AFEC)
  - Adaptive label control (Diff-Serv router support)
    - vBNS/Abilene

- Commodity Internet benchmarking

- Evaluate effectiveness of end-to-end QoS amplification
  - model of future Internet (NGI)
Adaptive Label Control (cont.)

- Integration with Purdue Infobahn & QoS peering

Network Systems Lab

Multimedia DB & Network Security Apps

Indy NOC

Abilene

Network Systems Lab
Adaptive Label Control (cont.)

- Application Benchmarking:

Network Systems Lab
Collaborations

- **Academic:**
  - Boston Univ. (A. Bestavros)
  - Ohio State Univ. (J. Hou)
  - Santa Fe Institute (Fellow-at-Large)
  - Univ. of Wisconsin (P. Barford; WAWM)
  - Seoul National Univ. (S. Bahk)

- **Industry/Research Labs:**
  - AT&T Research (W. Willinger)
  - Cisco (F. Baker)
  - Sprint (K. Metzger)
Acknowledgments & More Info

- Supported by:
  - NSF ANI-9714707, ANI-9875789 (CAREER), ESS-9806741, EIA-9972883; ANI-9729721 (vBNS)
  - Purdue Research Foundation
  - Santa Fe Institute
  - Sprint
  - CERIAS, SERC

- Research assistants & postdocs:
  - RAs: A. Balakrishnan, S. Chen, J. Cruz, G. Nalawade, H. Ren, M. Tripunitara, T. Tuan, W. Wang
  - Postdocs/visting scientists: S. Bahk, H. Lee, J. Park

- Network Systems Lab
  - [http://www.cs.purdue.edu/~nsl](http://www.cs.purdue.edu/~nsl)
Acknowledgments & More Info (cont.)

- Related publications: