Application Layer Switching: A Deployable Technique for Providing Quality of Service

Raheem Beyah
Communications Systems Center
School of Electrical and Computer Engineering
Georgia Institute of Technology

Globecom 2003  San Francisco, California
December 3, 2003

Communications Systems Center
Presentation Outline

- Motivation & Background
- QoS-based Routing Review
- Traffic Measurement & Characterization
- Generalized Application-Layer Overlay
- Application Layer Switching
Presentation Outline

- Motivation & Background
- QoS-based Routing Review
- Traffic Measurement & Characterization
- Generalized Application-Layer Overlay
- Application Layer Switching
Quality of Service (QoS) Background

- Internet was initially used for simple data transfer
- Now provides a multitude of services
- The current Internet is a best-effort, delay and jitter prone Network of networks
- Users and applications now push for a overall better and more predictable experiences on the Internet
- To accomplish this some level of QoS must be established
Presentation Outline

- Motivation & Background
- QoS-based Routing Review
- Traffic Measurement & Characterization
- Generalized Application-Layer Overlay
- Application Layer Switching
QoS-based Routing

- Current Internet routes packets using OSPF, taking into account only link cost
- Common minimum cost paths leads to over utilized links – “hot spots”
- Results in improper load balancing on the Internet
- QoS-based routing considers multiple metrics (distance, delay, loss, jitter, bandwidth)
Presentation Outline

- Motivation & Background
- QoS-based Routing Review
- Traffic Measurement & Characterization
- Generalized Application-Layer Overlay
- Application Layer Switching
In this experiment we generated:
- Single TCP flows, multiple TCP flows, single UDP flows

We discussed:
- Link asymmetry
- Effects of file size and transport protocol on throughput
- Overall throughput when using multiple TCP flows
- UDP and TCP behaviors
Experimental Setup

- Four end nodes (GT, NCAT, UCR, UCLA)
- Nodes are mesh connected - each node takes a different route to every other node
- Total of 12 paths
- Modified *sock* program – (Stevens vol. 1)
- Perl scripts used to run instances of program
- Cron scripts for Synchronization

- Scripts executed in two twelve-hour blocks
- Experiment Duration
  → Nov. 15, 2001 – Feb. 1, 2002
TCP Single Flows (Per Link Throughput)

<table>
<thead>
<tr>
<th>Link #</th>
<th>Link Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Link1</td>
<td>UCR→NCAT</td>
</tr>
<tr>
<td>Link2</td>
<td>GT→NCAT</td>
</tr>
<tr>
<td>Link3</td>
<td>UCLA→NCAT</td>
</tr>
<tr>
<td>Link4</td>
<td>UCR→GT</td>
</tr>
<tr>
<td>Link5</td>
<td>NCAT→GT</td>
</tr>
<tr>
<td>Link6</td>
<td>UCLA→GT</td>
</tr>
<tr>
<td>Link7</td>
<td>NCAT→UCR</td>
</tr>
<tr>
<td>Link8</td>
<td>GT→UCR</td>
</tr>
<tr>
<td>Link9</td>
<td>UCLA→UCR</td>
</tr>
<tr>
<td>Link10</td>
<td>UCR→UCLA</td>
</tr>
<tr>
<td>Link11</td>
<td>GT→UCLA</td>
</tr>
<tr>
<td>Link12</td>
<td>NCAT→UCLA</td>
</tr>
</tbody>
</table>

Fig. 1. Single TCP flows – All file sizes
Presentation Outline

- Motivation & Background
- QoS-based Routing Review
- Traffic Measurement & Characterization
- Generalized Application-Layer Overlay
- Application Layer Switching
Generalized Application-Layer Overlay (GALO)

- Overlay Architecture: current Infrastructure remains in place and a virtual network is run atop it.
- Previous Relevant Overlay Architectures
  - Detour
  - MIT RoN
  - Mbone
- Strictly uses end nodes to generate path quality updates
- Deployable solution: No infrastructure modifications (i.e. routers, etc.)
**Generalized Application-Layer Overlay (GALO)**

- GALO is used to gather path quality along the logical links of the collaborating nodes.
- Three components:
  - Distributed Client Engine (DCE)
  - Forwarding Engine (FE)
  - QoS Routing Engine (QRE)
- The Application Layer Communications Protocol (ALCP) is the scheme in place to communicate between components.
## Generalized Application-Layer Overlay (GALO)

<table>
<thead>
<tr>
<th>DCE</th>
<th>FE</th>
<th>QRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resides on remote node</td>
<td>Resides on remote node</td>
<td>Controlling unit of the system</td>
</tr>
<tr>
<td>Primary purpose is to calculate and transmit path quality updates (throughput) to QRE</td>
<td>Only invoked at transit nodes</td>
<td>Maintains an accurate picture of the network</td>
</tr>
<tr>
<td>May be used to implement loss detection or perform packet reordering</td>
<td>Switching engine for passing traffic</td>
<td>Receives UPDATE messages from DCEs</td>
</tr>
<tr>
<td></td>
<td>When appropriate signaling is received, incoming and outgoing ports are invoked to perform switching</td>
<td>Maintains reachability information</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Table uses a modified version of Dijkstra’s algorithm</td>
</tr>
</tbody>
</table>
Generalized Application-Layer Overlay (GALO)

Figure 11. GALO Logical Architecture.
Generalized Application-Layer Overlay (GALO)

- Emulate regular communications
- Background traffic - sock
- Passive sampling

<table>
<thead>
<tr>
<th>Path</th>
<th>Thpt</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-&gt;B</td>
<td>12</td>
</tr>
<tr>
<td>A-&gt;C</td>
<td>18</td>
</tr>
<tr>
<td>A-&gt;D</td>
<td>4</td>
</tr>
<tr>
<td>A-&gt;E</td>
<td>19</td>
</tr>
<tr>
<td>B-&gt;E</td>
<td>8</td>
</tr>
</tbody>
</table>

Internet Cloud
Summary of Contributions & Conclusion

- Deployable overlay architecture, GALO
- Obtain the state of the monitored network via invasive or non-invasive sampling/probing
- GALO’s modular design allows easy extensions
Presentation Outline

- Motivation & Background
- QoS-based Routing Review
- Traffic Measurement & Characterization
- Generalized Application-Layer Overlay
- Application Layer Switching
Application Layer Switching (ALSW)

- Motivated by improper traffic balance on Internet
- Related Work: Savage, et. al. [6]
  - Performed a measurement-based study of path quality (used traceroute)
  - Primary metrics considered were: rtt and loss, with minimal emphasis on thpt
  - After synthetically constructing alternate paths, found that 80% had a better path quality than default path
Application Layer Switching (ALSW)

- ALSW – Introduced to provide a better than traditional, best-effort QoS
- Has the ability to utilize alternate paths when routing
- Routing decisions can be made based on multiple metrics
- Hence, helping to balance the load on the Internet
Application Layer Switching (ALSW)

- Uses end nodes to route packets, therefore no modification of current infrastructure
- ALSW module is an extension of GALO
- GALO has the current state of the network, this is passed to the ALSW module to perform QoS-based routing
- If a better path for a flow is known, the ALSW module signals using the ALCP to the path nodes to allow forwarding along the new alternate path
- The path nodes provision appropriately and expect data to be routed through them
Figure 15. GALO Architecture extended to support Application Layer Switching.
Application Layer Switching (ALSW)

Path | Alternate Path | Min(Thpt) |
-----|----------------|-----------|
B->A | DEFAULT        | 12        |
B->A | B->D->A        | 18        |
B->A | B->E->A        | 4         |
B->A | B->D->C->A     | 19        |
B->A | B->E->C->A     | 8         |

Communications Systems Center
Experimental Setup (ALSW)

- 4 nodes running Linux, 2 nodes running Solaris
- All workstations are general purpose shared machines
- Code executed in user space
- Default file size - 100KB
- Sampling block – traffic generated to sample – Default: 130 seconds
- Reroute block – period used to reroute selected flow – Default: 130 seconds
Performance Analysis of ALSW

Figure 19. CDF of the total number of rerouted flow’s average throughput compared to the overall average throughput of the corresponding direct path (ALL). SAMP1.

- 70% of flows achieve higher throughput than default path
- 30% perform worse than default path
  - Overhead: going to application layer to route traffic
  - Node variable processing load
  - Temporal traffic fluctuations
  - Increased number of hops
  - Probability of loss increases
Sample Rate Variations

Figure 25. CDF of the total number of rerouted flows’ average throughput compared to the overall average throughput of the corresponding direct path. (ALL) SAMP1 vs. SAMP2.

- Sampling rate and the number of benefited flows are directly proportional
Performance Analysis of ALSW (continued)  

The Effect of Hop Count - Flows

- 90% of single hop rerouted flows benefit from ALSW
- Maximum gain – 120%

Only 40% of multiple hop rerouted flows benefit from ALSW
- Maximum gain – 20%
- Limitation in our scheme as number of hops increase, the effect of overhead increases

Figure 21. CDF of the total number of rerouted flows’ average throughput compared to the overall average throughput of the corresponding direct path (SINGLE HOP). SAMP1.

Figure 22. CDF of the total number of rerouted flows’ average throughput compared to the overall average throughput of the corresponding direct path (MULTIPLE HOP). SAMP1.
Summary of Contributions & Conclusion

- Deployable approach to improving QoS, by using the GALO overlay architecture along with the ALSW
- 70% of the reroute flows achieved a better average throughput than that of the default path
- Limitation of current implementation as paths get more complex
- Current scheme independent of file size variations
Application Layer Switching: A Deployable Technique for Providing Quality of Service

Raheem Beyah
Communications Systems Center
School of Electrical and Computer Engineering
Georgia Institute of Technology

Globecom 2003   San Francisco, California
December 3, 2003
Future Work

~Scalability and Increased Throughput~

- **Distributed QRE**
  - Placed in each autonomous system or domain
  - DCE would report to the nearest QRE
  - QREs will communicate on a regular basis, updating accessibility information for every other domain

- **Hierarchical Routing with Domain Speakers**
  - Best when you have several densely populated domains
  - Nodes clustered in groups, and each group would have a dedicated group speaker
  - The group speaker of each domain would report (to the QRE) its accessibility to the group speaker of every other domain
Application Layer Switching (ALSW)

- Related Work (continued):
  - Detour project [57]
    - Discussed routing and transport inefficiencies as a motivation for rerouting traffic around bottleneck and lossy links
    - Describe an architecture and provide preliminary results obtained using a testbed in a lab with traffic generators
  - MIT RoN [61] goals were to:
    - Fast failure detection and recovery
    - Tighter integration with applications
    - Fine-grained expressive routing policies
  - Primary emphasis is rerouting around network outages. They also reroute traffic during congestion periods and produce some basic empirical results considering only single-hop alternate paths