Large Data Transfers on Shared Wide-Area Networks

Hamidreza Anvari
hanvari@ualberta.ca
Department of Computing Science, University of Alberta
Edmonton, Canada
June 22, 2016
Acknowledgement

This presentation is mainly based on a submission for BigDAW’16

Co-author: Paul Lu, paullu@ualberta.ca


- Paul Lu, Professor
- Nooshin Eghbal, PhD Candidate
- Hamidreza Anvari, PhD Candidate
- Soham Sinha, MSc Student
Agenda

Introduction
High-profile Use-cases
Background: Networking
Data Transfer Tools
Measurements: Performance in Shared Networks
Conclusion & Future Works

Take-away message:
• TCP-based (E.g. GridFTP) for,
  • Dedicated network
  • Shared with TCP-based traffic
• UDP-based tools (E.g. UDT) for,
  • Shared with UDP-based traffic
Introduction

Large-scale data processing projects
- Sharing/transferring data across Wide-Area Networks (WANs)
- Large datasets → substantial transfer time
- Affecting the overall pipeline
Introduction

Large-scale data processing projects
- Sharing/transferring data across Wide-Area Networks (WANs)
- Large datasets → substantial transfer time
- Affecting the overall pipeline
Introduction

Large-scale data processing projects
- Sharing/transferring data across Wide-Area Networks (WANs)
- Large datasets $\rightarrow$ substantial transfer time

Characteristics of wide-area networks
- High BDP (Bandwidth Delay Product)
- A.k.a. LFN (Long Fat Networks)
Introduction

Large-scale data processing projects
- Sharing/transferring data across Wide-Area Networks (WANs)
- Large datasets → substantial transfer time

Characteristics of wide-area networks
- **High BDP** (Bandwidth Delay Product)
- A.k.a. **LFN** (Long Fat Networks)

~ 3,700 KMs

Map from: d-maps.com
Motivation

Performance issues on High-BDP networks

Emulated Testbed, RTT: 128ms, Bandwidth: 1000Mbps

* Numbers adopted from experiments done by Nooshin Eghbal, 2014.
Motivation

Performance issues on High-BDP networks

Emulated Testbed, RTT: 128ms, Bandwidth: 1000Mbps

* Numbers adopted from experiments done by Nooshin Eghbal, 2014.
Dedicated vs. shared networks

Networks are not always private, In many cases shared

- Dynamic workload
- Problematic performance on shared WANs

Emulated Testbed, RTT: 128ms, Bandwidth: 500Mbps
Background – Network Protocols

Reliability as a service on top of unreliable network
- Tracking delivery acknowledgement

TCP (reliable) vs. UDP (unreliable)

Congestion Control
- effective, network-aware use of available bandwidth
Background - Network Properties

Round Trip Time (RTT)
- 2 * End-to-end delay
- A function of Geo-distance, middleboxes, and network load
  - Dynamically increased by network load

![Diagram showing two points A and B with 35ms delay in both directions]
Background - Network Properties

Round Trip Time (RTT)
- 2 * End-to-end delay
- A function of Geo-distance, middleboxes, and network load
  - Dynamically increased by network load

Loss rate
- How often a packet is dropped in path
  - Overflowed buffers, physical noise, ...
- Very dynamic

![Diagram showing physical bandwidth and data transfer over time]
Background - Network Properties

Round Trip Time (RTT)
- 2 * End-to-end delay
- A function of Geo-distance, middleboxes, and network load
  - Dynamically increased by network load

Loss rate
- How often a packet is dropped in path
  - Overflowed buffers, physical noise, ...
  - Very dynamic

![Diagram showing network properties]
High-Performance Data Transfer Tools

Establishing Parallel Connections
- GridFTP
- Decrease impact of single packet loss on overall performance
High-Performance Data Transfer Tools

Establishing Parallel Connections
- GridFTP
- Decrease impact of single packet loss on overall performance

Tools based on unreliable UDP protocol
- UDT
- Reliability at user-level
- Efficient bandwidth management
Experimental Methodology

Emulated Testbed
- Dumbbell topology

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAN bandwidth</td>
<td>1 Gbps</td>
</tr>
<tr>
<td>WAN bandwidth</td>
<td>500 Mbps</td>
</tr>
<tr>
<td>RTT</td>
<td>128 ms</td>
</tr>
<tr>
<td>Router buffer size</td>
<td>6 MB</td>
</tr>
</tbody>
</table>
Experimental Methodology

Emulated Testbed

- Dumbbell topology

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>LAN bandwidth</td>
<td>1 Gbps</td>
</tr>
<tr>
<td>WAN bandwidth</td>
<td>500 Mbps</td>
</tr>
<tr>
<td>RTT</td>
<td>128ms</td>
</tr>
<tr>
<td>Router buffer size</td>
<td>6 MB</td>
</tr>
</tbody>
</table>
Experimental Methodology

Synthetic Traffic Pattern

<table>
<thead>
<tr>
<th>Background Traffic Type</th>
<th>x (seconds)</th>
<th>y (seconds)</th>
<th>Traffic Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform TCP</td>
<td>10</td>
<td>0</td>
<td>TCP</td>
</tr>
<tr>
<td>SQ UDP 1</td>
<td>10</td>
<td>10</td>
<td>UDP</td>
</tr>
<tr>
<td>SQ UDP 2</td>
<td>15~20</td>
<td>10</td>
<td>UDP</td>
</tr>
<tr>
<td>SQ TCP</td>
<td>10</td>
<td>10</td>
<td>TCP</td>
</tr>
<tr>
<td>Variable Bursty UDP</td>
<td>parameter</td>
<td>parameter</td>
<td>UDP</td>
</tr>
</tbody>
</table>

![Graph showing synthetic traffic pattern]
Performance in Shared Networks

Throughput Inefficiency

Emulated Testbed, RTT: 128ms, Bandwidth: 500Mbps, UDP1: 1Gbps, UDP2: 500Mbps
Performance in Shared Networks

Fairness: Effect of protocols on cross traffic
- Single TCP connection throughput

Emulated Testbed, RTT: 128ms, Bandwidth: 500Mbps
Performance in Shared Networks

Fairness: Effect of protocols on cross traffic

- Single TCP connection throughput

![Graph showing throughput over time for different types of streams]

<table>
<thead>
<tr>
<th>Type</th>
<th>Stream</th>
<th>no_bg</th>
<th>bg_iperf</th>
<th>bg_gftp16</th>
<th>bg_udt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dedicated</td>
<td>Background</td>
<td>-</td>
<td>478</td>
<td>473</td>
<td>481</td>
</tr>
<tr>
<td></td>
<td>Foreground (iperf)</td>
<td>474</td>
<td>474</td>
<td>474</td>
<td>474</td>
</tr>
<tr>
<td>Shared</td>
<td>Background</td>
<td>-</td>
<td>238</td>
<td>415</td>
<td>398</td>
</tr>
<tr>
<td></td>
<td>Foreground (iperf)</td>
<td>474</td>
<td>253</td>
<td>75.6</td>
<td>85.3</td>
</tr>
<tr>
<td></td>
<td>Jain fairness</td>
<td>-</td>
<td>1</td>
<td>0.68</td>
<td>0.7</td>
</tr>
</tbody>
</table>

**HAMIDREZA ANVARI – LARGE DATA TRANSFERS ON SHARED WIDE-AREA NETWORKS**

2016-06-22
Future Works

Hybrid solution for different conditions of network
- Combination of different protocols
- Automated switching between protocols
- Periodic parameter tuning for efficiency
Summary

Large-scale data processing projects
- Large data transfers
- Wide-Area Networks

Take-away messages
- TCP-based (E.g. GridFTP) tools work well if,
  - Dedicated network
  - Shared network with TCP-based traffic
- Consider using UDP-based tools (E.g. UDT) when,
  - Shared network with UDP-based traffic
Thank you

Large Data Transfers on Shared WANs

Hamidreza anvari
hanvari@ualberta.ca
Department of Computing Science
University of Alberta
Edmonton, Canada