Cause Analysis of Packet Loss in Underutilized Enterprise Network Links

2005.12.21
Thesis Defense

Deepali Agrawal
deepali@postech.ac.kr

Distributed Processing & Network Management Lab.
Department of Computer Science & Engineering
POSTECH
Contents

- Introduction
- Related Work
- Traffic Monitoring and Loss Detection
- Cause Analysis Method
  - Traffic Data Collection
  - Analysis Tools
- Cause Analysis of Packet Loss
- Conclusion and Future Work
Introduction

- ISPs Employ Overprovisioning
  - Increasing number of users and applications

- Performance Problems
  - Packet loss, delay and jitter
  - Network applications: VoIP, multimedia, games and P2P

- Goal
  - Study traffic and packet loss characteristics
  - Determine the root cause of packet loss phenomenon

- Approach
  - Developed a root cause analysis method
  - Developed set of tools to execute the methodology
Related Work

- Most of the studies focus on packet delay analysis
  - Packet loss and delay are closely related

- Network Performance Monitoring at Small Time Scales
  - Papagiannaki et. al. (IMC 2003)
  - Attributed high delays to congestion in the routers

- Origins of Microcongestion in an Access Router
  - Papagiannaki et. al. (PAM 2004)
  - Identified and discussed causes of microcongestion: link bandwidth, multiplexing and traffic burstiness

- Detection and Analysis of Packet Loss on Underutilized Enterprise Networks
  - Chung et al. (E2EMON 2005)
  - Indicated that only bursty packets affect the packet loss
Traffic Monitoring and Loss Detection

Linux Traffic Monitor System

<table>
<thead>
<tr>
<th>Module Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNMP Polling Module</td>
<td>This module polls Cisco standard MIB II and private MIB variables</td>
</tr>
<tr>
<td>Packet Capture Module</td>
<td>This module captures packet trace using DAG API</td>
</tr>
</tbody>
</table>
SNMP Polling

- Monitor traffic and switch status
  - By polling standard MIB II variables at 1 second granularity

<table>
<thead>
<tr>
<th>Object</th>
<th>OID</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ifInUcastPkts</td>
<td>1.3.6.1.2.1.2.2.1.11</td>
<td>The number of subnetwork-unicast packet delivered to a higher-layer protocol</td>
</tr>
<tr>
<td>ifOutUcastPkts</td>
<td>1.3.6.1.2.1.2.2.1.17</td>
<td>The total number of packets that higher-level protocols requested be transmitted to a subnetwork-unicast address, including those that were discarded or not sent</td>
</tr>
<tr>
<td>ifInOctets</td>
<td>1.3.6.1.2.1.2.2.1.10</td>
<td>The total number of octets receive on the interface, including framing characters</td>
</tr>
<tr>
<td>ifOutOctets</td>
<td>1.3.6.1.2.1.2.2.1.16</td>
<td>The total number of octets transmitted out of the interface, including framing characters</td>
</tr>
</tbody>
</table>
Packet Loss Detection Using SNMP

- Packet loss information
  - Using Cisco enterprise MIB variables

- Each interface owns input queue and output queue

- Packet Loss = Input Queue Drops + Output Queue Drops

<table>
<thead>
<tr>
<th>Object</th>
<th>OID</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>cpuLoad</td>
<td>1.3.6.1.4.1.9.2.1.56</td>
<td>CPU Utilization (5 sec avg.)</td>
</tr>
<tr>
<td>locIfInputQueueDrops</td>
<td>1.3.6.1.4.1.9.2.2.1.1.26</td>
<td>The number of packets dropped because the input queue was full</td>
</tr>
<tr>
<td>locIfOutputQueueDrops</td>
<td>1.3.6.1.4.1.9.2.2.1.1.27</td>
<td>The number of packets dropped because the output queue was full</td>
</tr>
</tbody>
</table>
Packet Capture Using Network TAP

- **SNMP**
  - Inaccuracies due to high response loss rate
  - Time granularity not satisfactory enough
    - Need finer time analysis for cause detection

- **Optical TAP**
  - Packet trace capture using DAG card
  - Guaranteed lossless performance in gigabit link
  - High precision time stamps

- **Packet Capture Module**
  - Implemented using C API of DAG
  - Provides highest performance
Traffic Data Collection

Overall switch status and traffic flow

SNMP Tools

Protocol Based Analysis

IP Traffic

Inter-domain Traffic

TCP

UDP

Other

Intra-domain Traffic

TCP

UDP

Other

Non-IP Traffic

Broadcast Packet Analysis

Loss Correlation Analysis

SNMP Tools

TCP Flow Analysis

Distinct Destinations

Count

Sizes

Lifetime

New and Terminated Flows

Large TCP Flow

Throughput

Sequence Number

Loss Rate
Traffic Data Collection

- POSTECH campus network
  - Composed of two IP routers, two core backbone switches, dozes of gigabit switches and hundreds of 100mbps switches

- Monitored Link
  - Always underutilized
  - Convey traffic composed of many internet application

- Link located in campus dormitory network
Traffic Data Collection

INTERNET

Cisco 7513
Router 1

100Mbit/Sec
1Gbit/Sec

Cisco 7401
Router 2

Catalyst 6513
Core1 Switch

Catalyst 6513
Core2 Switch

Catalyst 5500
Dorm. Building Switches

Dorm. Backbone Switch

Catalyst 2950
Nakwon APT Switch

Backbone Switch

Catalyst 3550
PIRL

Catalyst 3508

Catalyst 2924
DPNM Lab. Switch

PC

PC

Campus Dormitory Network
POSTECH Dormitory Network Overview

- Monitored link connecting Nakwon Apt to dormitory backbone switch
- Data collection over 4 months, interleaved with analysis
## Analysis Tools
### SNMP Logs Processing Tools

<table>
<thead>
<tr>
<th>Tool Name (Perl Script)</th>
<th>Input</th>
<th>Output</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interface Count</td>
<td>SNMP Log file</td>
<td>Text file</td>
<td>Count of interfaces those experience losses simultaneously</td>
</tr>
<tr>
<td>Interface Index</td>
<td>SNMP Log file</td>
<td>Text file</td>
<td>Index of interfaces those experience losses simultaneously</td>
</tr>
<tr>
<td>Response Loss Rate</td>
<td>SNMP Log file</td>
<td>Text file</td>
<td>Time in Unix seconds when the SNMP response is not received</td>
</tr>
<tr>
<td>Average</td>
<td>SNMP Log file</td>
<td>Text file</td>
<td>Averaged values for specified seconds</td>
</tr>
<tr>
<td>Total</td>
<td>SNMP Log file</td>
<td>Text file</td>
<td>Total count across all ports</td>
</tr>
</tbody>
</table>
## Analysis Tools

### DAG Logs Processing Tools

<table>
<thead>
<tr>
<th>Tool Name (C programs)</th>
<th>Input</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utilization measure</td>
<td>DAG Logs</td>
<td>Inter domain and intra domain bit counts (at 1s, 1ms and 1us scale)</td>
</tr>
<tr>
<td>Protocol Analysis</td>
<td>DAG Logs</td>
<td>Inter &amp; intra domain Bit Counts : TCP, UDP and Other (1s, 1ms &amp; 1us)</td>
</tr>
<tr>
<td>Flow generator</td>
<td>Binary and ascii flow file</td>
<td>Generate 4 (src/dst IP, src/dst port) tuple based TCP flows</td>
</tr>
<tr>
<td>Flow count</td>
<td>Binary Flow file</td>
<td>Inter/intra domain flow counts per second</td>
</tr>
<tr>
<td>Flow lifetime and size</td>
<td>Binary Flow file</td>
<td>Inter/intra domain flow lifetime and sizes</td>
</tr>
<tr>
<td>New and Exit Flow</td>
<td>Binary Flow file</td>
<td>New and terminated flow counts and sum of their sizes per second</td>
</tr>
<tr>
<td>Distinct destination</td>
<td>Binary Flow file</td>
<td>Count of distinct destinations to which each source IP connects</td>
</tr>
<tr>
<td>Top_n_flow</td>
<td>Binary Flow file</td>
<td>flows of size in the specified range</td>
</tr>
<tr>
<td>Data Rate</td>
<td>Binary Flow file &amp; DAG logs</td>
<td>pps and BPS of the selected flow (at 1s, 1ms and 1us scales)</td>
</tr>
<tr>
<td>Run length and loss rate</td>
<td>Binary Flow file &amp; DAG logs</td>
<td>run-length magnitudes and interval between two run-lengths in microsecond</td>
</tr>
<tr>
<td>Broadcast packets</td>
<td>DAG logs</td>
<td>Count of IP level broadcast packets (at 1s, 1ms and 1us scales)</td>
</tr>
<tr>
<td>Non-IP packets</td>
<td>DAG logs</td>
<td>Count of non-IP packets and ARP packets (at 1s, 1ms and 1us scales)</td>
</tr>
</tbody>
</table>
Analysis Tools

- SNMP Logs
- DAG Logs

Generating complete diagnosis document
- Diagnosis Script: Process data to get useful information
- Gnuplot Script: Generate plots
- Latex script: Arrange plots in document
### Dorm Backbone Switch
### Underutilized Links

<table>
<thead>
<tr>
<th>Interface Number</th>
<th>Mean Ingress Bits (Mbps)</th>
<th>Standard Deviation (Mbps)</th>
<th>Max (Mbps)</th>
<th>Avg. Utilization (AVG/BW) %</th>
<th>Max Utilization (Max/BW)%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2 (Backbone Link)</td>
<td>265</td>
<td>399.4</td>
<td>829.61</td>
<td>26.5</td>
<td>82.9</td>
</tr>
<tr>
<td>3</td>
<td>25.85</td>
<td>38.2</td>
<td>87.10</td>
<td>12.9</td>
<td>43.5</td>
</tr>
<tr>
<td>4</td>
<td>12.49</td>
<td>18.3</td>
<td>33.24</td>
<td>6.2</td>
<td>16.6</td>
</tr>
<tr>
<td>5</td>
<td>65.86</td>
<td>94.2</td>
<td>196.07</td>
<td>32.9</td>
<td>98</td>
</tr>
<tr>
<td>6</td>
<td>30.60</td>
<td>44.5</td>
<td>96.46</td>
<td>15.3</td>
<td>48.2</td>
</tr>
<tr>
<td>7</td>
<td>54.07</td>
<td>10.1</td>
<td>37.77</td>
<td>27.0</td>
<td>18.8</td>
</tr>
<tr>
<td>8</td>
<td>18.84</td>
<td>29.9</td>
<td>58.31</td>
<td>9.4</td>
<td>29.1</td>
</tr>
<tr>
<td>9</td>
<td>4.13</td>
<td>13.0</td>
<td>48.51</td>
<td>2.0</td>
<td>24.2</td>
</tr>
<tr>
<td>10</td>
<td>11.69</td>
<td>16.7</td>
<td>28.58</td>
<td>5.8</td>
<td>14.2</td>
</tr>
<tr>
<td>11(Nakwon Link)</td>
<td>17.44</td>
<td>27.4</td>
<td>61.82</td>
<td>1.7</td>
<td>6.1</td>
</tr>
</tbody>
</table>
**Dorm Backbone Switch**

**Backbone VS Non-backbone Traffic**

- Switch have single backbone link
- Total ingress traffic across all non-backbone link is equivalent to backbone egress traffic
Protocol Based Analysis

**IP Traffic**

<table>
<thead>
<tr>
<th>Nak-won link Traffic</th>
<th>TCP</th>
<th>UDP</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra Domain</td>
<td>2.8%</td>
<td>0.02%</td>
<td>0.003%</td>
</tr>
<tr>
<td>Inter Domain</td>
<td>96.7%</td>
<td>0.39%</td>
<td>0.01%</td>
</tr>
</tbody>
</table>
Protocol Based Analysis

Non-IP Traffic
Broadcast Packet Analysis

- IP level broadcast packets
- Tie up system resources
- Broadcast packets and packet loss correlation coefficient: -0.009
Loss Correlation

- Packet Loss
  - 1 sec scale
  - Over 30 mins

Losses are highly correlated
Flow Analysis - Distinct Destinations

- 4 Tuple-based TCP Flow (Src/Dst IP Address, Src/Dst Port)
- Max number of distinct destinations
  - Loss period: 492
  - No loss period: 460
- Number of distinct sources are 63
- No IP spoofing
**Flow Analysis - Flow Count**

- More number of flows during loss period

<table>
<thead>
<tr>
<th>Packet Loss</th>
<th>Mean</th>
<th>Max</th>
<th>Min</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heavy</td>
<td>300.11</td>
<td>393</td>
<td>197</td>
<td>26.277</td>
</tr>
<tr>
<td>Rare</td>
<td>240.490</td>
<td>316</td>
<td>111</td>
<td>26.669</td>
</tr>
</tbody>
</table>
Flow Analysis – Flow Size

- $\Pr [x = a] \sim c x^{\{-\alpha\}}$

<table>
<thead>
<tr>
<th>Packet Loss</th>
<th>Mean</th>
<th>Min (Bytes)</th>
<th>Max</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heavy</td>
<td>0.24</td>
<td>40</td>
<td>436.61</td>
<td>6.06</td>
</tr>
<tr>
<td>Rare</td>
<td>0.233</td>
<td>40</td>
<td>241.228</td>
<td>4.583</td>
</tr>
</tbody>
</table>
Flow Analysis – Flow Lifetime

- Flows are longer during loss period

<table>
<thead>
<tr>
<th>Loss</th>
<th>Mean</th>
<th>Max</th>
<th>Min</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heavy</td>
<td>42.25034</td>
<td>2000</td>
<td>0</td>
<td>196.6926</td>
</tr>
<tr>
<td>Rare</td>
<td>36.51392</td>
<td>2000</td>
<td>0</td>
<td>177.6445</td>
</tr>
</tbody>
</table>
Flow Analysis – New & Terminated Flows

Number of flows initiated and terminated every second are comparable.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Max</th>
<th>Min</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loss - New Flow</td>
<td>7.660</td>
<td>116</td>
<td>0</td>
<td>6.039</td>
</tr>
<tr>
<td>Loss - Terminated Flow</td>
<td>7.660</td>
<td>73</td>
<td>0</td>
<td>5.624</td>
</tr>
<tr>
<td>New flow</td>
<td>7.081</td>
<td>83</td>
<td>0</td>
<td>5.753</td>
</tr>
<tr>
<td>Terminated Flow</td>
<td>7.081</td>
<td>63</td>
<td>0</td>
<td>5.733</td>
</tr>
</tbody>
</table>
Flow Analysis - Large TCP Flows

- Sum of top 50 flows size: 83% of total traffic
Flow Analysis - Large TCP Flows

Throughput

- Flow sizes: 179Mbytes and 140Mbytes
Flow Analysis - Large TCP Flows

Sequence Number

- SNMP Loss Rate: 40 packets/second
- Flow Loss Rate: 37 packets/second
Conclusion

- Study traffic and packet Loss characteristics
  - To determine root cause

- Cause Analysis
  - Collected data from a underutilized link (2%)
  - Defined a complete methodology
  - Developed tools to execute methodology
  - Analyzed data using our tools

- Results
  - TCP protocol is present in highest percentage
  - No malicious or spurious traffic is present in our link
  - Broadcast packets do not affect packet loss
  - Losses are strongly correlated
  - Number of flows are higher during loss period
  - TCP large flows are suspected as one of the possible reasons
    - Complicated to verify due to SNMP inaccuracies

- Contribution
  - Developed a cause analysis method
  - Developed set of tools to execute the analysis
Future Work

- Detail study of TCP characterization by monitoring and analyzing both ingress/egress directions and backbone link traffic
- Collecting data sets from various networks including IPv6
- Studying packet loss characteristics at core switch or router
- Checking if losses occur only on the links that convey large TCP flows
- If large TCP flows are responsible for losses
  - Need to determine remedy to avoid these losses
Questions?