KNOM Tutorial 2003

Internet Traffic Matrix Measurement and Analysis

Sue Bok Moon
Dept. of Computer Science

KAIST
Overview

- Definition of Traffic Matrix
  - Traffic demand, delay, loss
- Applications of Traffic Matrix
  - Engineering, research, SLAs
- Challenges in Obtaining Traffic Matrix
  - Limitation of NetFlow and active probes
  - Challenges in measurement and modeling
- Summary & Future Work
Definition of Traffic Matrix

- What is a traffic matrix?
  - A matrix built on metric of interest
  - Traffic demand matrix
    - How much traffic flows from point A to point B
    - Granularity: PoP, router, link, prefix
  - Delay matrix
    - How much delay from point A to point B
    - Granularity: PoP, router, link, end hosts
  - Loss matrix
    - How many packets are dropped from point A to point B
    - Granularity: PoP, router, end hosts
Example: AT&T Latency Matrix

Current Average: 35 msec

Latency in milliseconds
Traffic Demand Matrix

- Not part of SLAs
  - Hard to obtain
  - Few available publicly
Delay Matrix

- Usually a matrix of average delay of *pings* between routers of random selection per PoP
  - Average of all PoP-to-PoP delays => SLA

- At end hosts
  - Easy to get using *pings* between hosts of interest
Loss Matrix

- Usually a matrix of average loss rate of *pings* between routers of random selection per PoP
  - Average of all PoP-to-PoP loss rates => SLA

- At end hosts
  - Easy to get using *pings* between hosts of interest
Applications of Traffic Matrix

- Marketing/Sales
  - How much traffic does customer A send from point #1 to point #2?
    - Where should customer A buy more capacity from us?
  - Is most traffic originating in Korea stay within Korea?
    - What is the trend in international traffic growth?
  - What is the performance that customer A sees?
    - Do we have an edge over our competitors?
Applications of Traffic Matrix

- Network Operators
  - Capacity Planning
    - How much traffic do we have from point A to point B?
    - How much capacity should we add?
    - When should we add more capacity?
  - Network Engineering
    - Where is the hot spot? – From SNMP
    - What if a link fails from point A to point B?
    - What if we move traffic from point A to point B?
Applications of Traffic Matrix

- Customers: SLAs
  - What quality of service am I getting?
    - How much delay do I get from ISP A?
    - How much loss do I experience from ISP A?
    - Can I get delay under X ms from ISP A?
    - What is the most popular destination of my traffic?
Applications of Traffic Matrix

- Researchers
  - Traffic modeling
    - How does TM evolve over time?
    - What is the fanout factor of traffic?
    - How much more capacity do we expect between point A and point B?
  - Example: IP over WDM
    - Given physical topology of routers and optical nodes, what is the “best” virtual topology?
    - Based on traffic demand matrix
Challenges in Obtaining Traffic Matrix

- Traffic Demand Matrix
  - resource requirements in routers
    - # of concurrently active flows
  - resource requirements in measurement infrastructure
    - production rate of flow statistics
  - traffic characterization
    - packet/byte rate of original traffic
    - rate of occurrence of original flows
    - average packet/bytes per original flow
Resource Requirements

- Server
- NetFlow data
- Analysis
- Traffic reports
- Router
- Memory
- Fast link
- Resource in Router
- Resource in Measurement Infrastructure
- Network
- Network Operation Center
Most Popular Tools of Choice?

- NetFlow for traffic demand matrix
- ping for delay and loss matrix
NetFlow

- Cisco’s “proprietary” tool
  - Not an IETF standard

- Basic idea
  - Based on (src ip, src port, dst ip, dst port, proto)
  - Records byte/packet/duration per flow
  - Cannot keep up with high speed links
  - Can sample every N\textsuperscript{th} packet
NetFlow Sampling

Original Packets

Sampled Packets (every 1/N, N=3)
Limitation of NetFlow

- Scalability
  - Historically NetFlow had a “performance issue”
  - Never deployed at the core
  - Number of flows in case of DDoS attacks beyond capacity
    - Network melt down
Number of Active Flows on a OC-48 Link
Limitation of NetFlow

- Representativeness
  - Can we estimate # of total flows from # of sampled flows accurately?
  - Can we estimate # of total WWW flows from # of sampled WWW flows accurately?
- Metrics of interest:
  - # of flows, flow rate,
- Packet sampling
  - reduce effective packet rate
  - save cost: slower memory sufficient (DRAM vs SRAM)
NetFlow Sampling

Original Packets

Sampled Packets (every 1/N, N=3)

Flow Splitting
Comparison of sparse and non-sparse applications

- Flow definition
  - 5-tuple = (src ip, src port, dst ip, dst port, proto)
  - interflow timeout = T
- Increase timeout T
  - potentially less splitting
  - fewer measured flows, more active flows
- Sparse vs. non-sparse flows
  - napster vs. www
  - # of mean active flows change differently over T
  - No simple model of rate and # active flows based on aggregate traffic rates
  - Model sparse and non-sparse flows separately
    [Duffield03]
Challenges in Delay Monitoring

- Not much is known about delay within ISP
  - People think they know delay, but ...
  - Cisco SAA implementation on GSR did not consider clock synchronization, and outputs meaningless numbers

- Too many paths to cover
  - hop-by-hop addition not yet possible
Limitation of Active Probes

- Representativeness [Choi04]
  - Average? Median?
Suitable Statistic: Percentile!

- Mode detection is hard
  - Difficult to distinguish small from big
  - Don’t know how many ahead of them

- High-percentile
  - represents upper bound for “most” delay
  - requires a very small number of probes to estimate
Sampling for Demand Matrix

- Periodic sampling does not answer:
  - What are the top 10 flows?
  - What is the most dominant application and who is the heaviest user?
  - What is the total # of packet for every flow?
Hash Function

- Mapping from a very large space to a smaller space
  - $h: X \rightarrow Y$ where $|X| >> |Y|$  
  - IP address to 10-bit hashed key  
  - 5-tuple address to 30-bit hashed key
- Load factor = collision probability
What are the top 10 flows?
Sampling for Elephants [Estan02]

All packets

Every n-th packet

Update entry or create a new one

Large flow memory

Has entry?

Pass with $p \sim \text{size}$

Small flow memory

no

create new entry
Sampling for Elephants [Estan02]

Flow Memory

h1

h2

h3

All Large?
What is the most dominant application and who is the heaviest user?
Who is using my link? [Estan03]
Looking at the traffic

Too much data for a human
Do something smarter!
Looking at traffic aggregates

- Aggregating on individual packet header fields gives useful results:
  - Traffic reports are not always at the right granularity (e.g. individual IP address, subnet, etc.)
  - Cannot show aggregates defined over multiple fields (e.g. which network uses which application)
- The traffic analysis tool should automatically find aggregates over the right fields at the right granularity.

Traffic Destination IP

<table>
<thead>
<tr>
<th>Source IP</th>
<th>Destination IP</th>
<th>Traffic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ran</td>
<td>2.83%</td>
<td>risc.cs.bigU.edu</td>
</tr>
<tr>
<td>3.12%</td>
<td>Tracy.dorm.bigU.edu</td>
<td></td>
</tr>
<tr>
<td>11.9%</td>
<td>jeff.dorm.bigU.edu</td>
<td></td>
</tr>
</tbody>
</table>

Most traffic goes to the dorms…

Traffic Destination Network

<table>
<thead>
<tr>
<th>Source IP</th>
<th>Destination Network</th>
<th>Traffic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ran</td>
<td>17.8%</td>
<td>dorm.bigU.edu</td>
</tr>
<tr>
<td>18.1%</td>
<td>cs.bigU.edu</td>
<td></td>
</tr>
<tr>
<td>27.5%</td>
<td>library.bigU.edu</td>
<td></td>
</tr>
</tbody>
</table>

What apps are used?

- Which network uses the web and which one kazaa?
- Most traffic goes to the dorms…
This paper is about giving the network administrator insightful traffic reports.

**Ideal traffic report**

<table>
<thead>
<tr>
<th>Traffic aggregate</th>
<th>Traffic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web traffic</td>
<td>42.1%</td>
</tr>
<tr>
<td>Web traffic to library.bigU.edu</td>
<td>26.7%</td>
</tr>
<tr>
<td>Web traffic from <a href="http://www.schwarzenegger.com">www.schwarzenegger.com</a></td>
<td>13.4%</td>
</tr>
<tr>
<td>ICMP traffic from sloppynet.badU.edu to jeff.dorm.bigU.edu</td>
<td>11.9%</td>
</tr>
</tbody>
</table>
Traffic Clusters and Reports

- Traffic clusters are multidimensional aggregates.
- Traffic reports give volume of chosen clusters.
- Only those over threshold are reported.
- To avoid redundant data, compress inferrable data (up to error H).
- Highlight non-obvious aggregates with unexpectedness label.
Structure of regular traffic mix

- Backups from CAIDA to tape server
  - Semi-regular time pattern
- FTP from SLAC Stanford
- Scripps web traffic
- Web & Squid servers
- Large ssh traffic
- Steady ICMP probing from CAIDA
What is the total # of packet of every flow?
Space-Code Bloom Filter

- Bloom filter answers set-membership.
- Space-code bloom filter answers multiset-membership
- Use a number of "virtual Bloom-filters, spread multiplicity information over space.
- Write-only
- At OC768, it can work at 5ns SRAM
- What about storage space at the router?
Future Work

- One traffic matrix to rule?
  - Can we answer all questions with one matrix?
- Continuous monitoring
  - data export in real-time
  - query over streaming data
- Availability/survivability
  - Implications in SLAs?
Failures are part of everyday operations
Time between Failures (network-wide)

43% : <1 min

81% : <20 min
Sources of failures

- Duration can provide hints, e.g.,
  - long (>1 hour): fiber cuts, severe failures
  - medium (>10 min): router/line card failures
  - short (>1 min): line card resets
  - very short (<1 min): software problems, optical equipment glitches

- Other hints
  - shared equipment (routers, optical)
  - router logs (e.g., SONET alarms), etc.
Network-wide Failure Duration

![Graph showing failure duration distribution](image)

- 40% in 1-60sec
- 40% in 1-15min
- 10% in 15-60min
- 10% >1h
References


Acknowledgements

- C. Estan’s SIGCOMM 2002 talk.
- S. Bhattacharyya and G. Iannaccone’s ICNP 2003 Tutorial.