Jiří Navrátil

SLAC CT measurement tools

Remarks to BW estimation and

My experience, results and
My interest is concerned on

- Full BW utilization on the path
- Tools for BW estimation
- Improving of CT detection
The main goal of our daily effort is Maximum BW Utilization (Inspiration) – Feng Wu-Chun simulations
- Jin Guojun netest-2 recommendation
- Les Cottrell IEP metrics
Maximun BW utilization (Inspiration)
Area of saturation = area of FULL utilization

- Combination of both methods
- TCP windows size (not sufficient for HSL)
- Parallel streams
Parallel TCP via Iperf

 SLAC-LANL

#N parallel streams

<table>
<thead>
<tr>
<th>248</th>
<th>162</th>
<th>243</th>
<th>128</th>
</tr>
</thead>
</table>

Total speed [Mbps]

Saturation point

# parallel streams

Parallel TCP via Iperf

(SLAC-LANL)
<table>
<thead>
<tr>
<th>Mbytes</th>
<th>Median</th>
<th>Average</th>
<th>Sdev</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>113.4</td>
<td>23.65</td>
<td>0.057735</td>
</tr>
<tr>
<td>0</td>
<td>52.7</td>
<td>22</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>26.4</td>
<td>22</td>
<td>1</td>
</tr>
</tbody>
</table>
Parallel TCP via Iperf

( SLAC-LANL )

Stream speed [Mbps]

Inflexion point

Parallel TCP via Iperf (SLAC-LANL)

#N parallel stream

0 5 10 15 20 25

2 4 8 16 24 32 64 96 128
Parallel TCP - Bandwidth allocation (1)

Total speed as SUM of partial streams speed [Mbps]
Why such distribution of throughput of individual streams?
Problem on slow and high loaded lines is

1. All parallel streams share same "virtual queue"
2. All "my traffic" share same queue with outside CT

visible in time reports (In iperf) visible in statistics for "streams speed"

10 second in Iperf means 10 second of sending ! Not receiving
Metric with 8 parallel streams doesn’t fit to the BW estimation
Parallel TCP Bandwidth

(iperf from SLAC-CERN)

#N parallel streams

Total speed [Mbps]

Stream speed

sample-1
sample-2
sample-3
sample-4
sample-5

Parallel TCP Bandwidth

(iperf from SLAC-CERN)
BW allocation

SLAC-CERN(32 streams)

Mbps per stream

#N parallel streams

0 8 16 24
BW allocation (SLAC-CERN)

#N parallel streams

N=16
N=64
N=32
N=24

Mbps per stream

0
0.5
1
1.5
2
2.5
3
3.5
Parallel TCP via Iperf (32 streams SLAC-IN2P3)
BW allocation (SLAC LANL)

#N parallel streams

Mbps per stream

N=128
N=64
N=32
N=24
N=16

Bw allocation
Used tools for BW estimations

• Pathrate
• Incite BW
• Nettest-2
• Iperf
• Pathload
• Pathtrace
Short characteristic of methods

**Patharate**
- First version 2.1 not very reliable (reject more than 60%)
- Long run time
- Limited range of operation (\( > 155 \text{ Mbps} \))
- Accurate, fast, light
- Sensitive to heavily loaded lines

**Pathload**
- Accurate, fast, light
- Limited range of operation (\( < 155 \text{ Mbps} \))
- Long run time
- Report of results (too many text info)
Short characteristic of methods

**Iperf**
- Different timing scheme, problematic accuracy
- Must be configured to use full BW
- Reliable, gives recommendations of WS

**Netest-2**
- Creates heavy load on lines during operation
- Must be configured to use full BW and post processed
- Reliable
Small comparisons
Iperf vers. Netest-2

SLAC - CERN (pdaqga)
Iperf vers. Netest-2

# N parallel streams

2 4 8 16 24 32 64

Aggregated Speed

Mbps

2 4 6 8 10 12 24 32 48 64
INCITE: Edge-based Traffic Processing and Service Inference for High-Performance Networks

Richard Baraniuk, Rice University; Les Cottrell, SLAC; Wu-chun Feng, LANL
\[ dTS(20\ \text{ms}) \]

\[ BWe \approx f(VQ) = f(dTR) \]

\[ <dTR > 0, RTT \]

\[ dtp1 \]

\[ dtp2 \]

\[ dTR \]

\[ RTT \]
What is pkt_length?

Open Problem

\[ \text{BW} = \text{Mean(freq)} \]

\[ \text{dt} \sim F(VQ) \sim F(\text{Internet-path}) \]

\[ \text{data volume} \sim \text{pkt length}^8 \]

\[ \text{freq} = \frac{dv}{dt} \]
All "my test traffic" share same queue

The average packet length ~ 1000 bytes

packet distribution on particular path.

The accuracy is dependent on the knowledge of

Virtual queue

It means that delay caused to my pkts by VQ

is not dependent only on my pkt-lengths !
CEP is connected via 155 Mbit to US

(Dominant bottleneck)
Sunstream sitting on 100 Mbit LAN

3 peaks = 3 bottlenecks

BW $= 9.1768 \times 10^7$
Reaction to the network problems
\( CT = f(\text{dTR}) \)
MF-CT Features and benefits

- No need access to routers!
  - Current monitoring systems for load of traffic are based on SNMP or Flows (needs access to routers)

- Low cost:
  - Allows permanent monitoring (20 pkts/sec ~ overhead 10 Kbytes/sec)
  - Can be used as data provider for ABW prediction (ABW = BW - CT)
  - We have 2 data points of CT per second!!

- Weak point for common use
  - No need access to routers!
Probing