Towards a Measurement Based Networking approach for Internet QoS improvement

Philippe Owezarski, Nicolas Larrieu

LAAS-CNRS
Toulouse, France
{owe, nlarrieu}@laas.fr
Outline

- Internet traffic characterization and QoS issue analysis
- TFRC: a step in the right direction for QoS
- Comparative impact of TCP vs TFRC on traffic characteristics
- MBN: a new measurement based networking approach
- MBCC: a case study of MBN
- MBCC evaluation
- Conclusion and future projects
Introduction

- New applications with various and changing requirements in terms of QoS appear
- New P2P applications make traffic characteristics change
Internet traffic evolution (May 2000)

Main TCP applications throughputs (SPRINT)

Time

Throughput (kbits/s)

- Other
- Telnet
- RealAudio
- MediaPlayer
- Quake
- NNTP
- SMTP
- FTP
- HTTPS
- HTTP
Internet traffic evolution (August 2000)

Main TCP applications throughputs (SPRINT)
Internet traffic evolution (May 2003)

Main TCP applications throughputs (Renater)

<table>
<thead>
<tr>
<th>Time</th>
<th>Throughput (kbits/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>May20 21:30</td>
<td></td>
</tr>
<tr>
<td>May20 21:53</td>
<td></td>
</tr>
<tr>
<td>May20 22:15</td>
<td></td>
</tr>
<tr>
<td>May20 22:30</td>
<td></td>
</tr>
<tr>
<td>May20 23:00</td>
<td></td>
</tr>
<tr>
<td>May20 23:23</td>
<td></td>
</tr>
<tr>
<td>May20 23:45</td>
<td></td>
</tr>
</tbody>
</table>

Legend:
- ip.tcp.ftp
- mac.ip.udp
- tcp.ftp
- tcp.udp
- tcp.http
- tcp.http.azza
- tcp.udp.dns
- tcp.udp.edonkey
- tcp.udp.dns
- tcp.udp.ssh
- tcp.udp.https
- tcp.udp.sftp
- tcp.udp.tls
Impact of P2P on traffic

- Thousands of mice
- A large number of elephants

⇒ Change flow size distribution
Flow size distribution

%  100  90  80  70  60  50  40  30  20  10

- Exponential
- August 2000
- May 2003

Nb of packets / flow

Séminaire OLC, 26 avril 2004
Traffic oscillation issues

Integrated Throughput on:

- 0.01 s
- 0.1 s
- 1 s

IP traffic

Simulated Poisson traffic
Illustration: LRD and losses
Illustration: LRD and losses
Illustration: LRD and losses
Illustration: LRD and losses
Illustration: LRD and losses
Wavelett analysis of the traffic
Oscillations persistence characterization

Logscale Diagram, $N=5$  \[
(j_1,j_2)=(9,15), \quad \alpha\text{-est} = 0.873, \quad Q = 1.6199e-007 \]

$H = 0.641$
Multiple causes for Internet oscillations

1. TCP like congestion control mechanisms
   (Slow Start and Congestion Avoidance mechanisms / Closed control loop)

2. Increase of transmitted file size

3. Increase of network capacities (and over-provisioning)

⇒ Increase of oscillations

- Amplitude
- Range

traffic oscillations limit network performance
« High variability » paper of Willinger (IEEE ToN 96)
Disturbances are mainly due to elephants

- We need to increase elephant flows regularity

⇒ We are going to use the TFRC mechanism to transmit elephant flows
TFRC principles

- TFRC is a new congestion control mechanism dedicated for stream oriented applications.

- TFRC proposes a smooth sending rate with very soft increases and decreases:
  - Computed once by RTT by receiver.
  - According to the loss event rate (LER).

LER = a loss event is considered if at least one loss appears in a RTT.
Experiment description

Objective: comparative evaluation of the global traffic characteristics if elephants use TCP or TFRC as the transmission protocol

Start points:
- Traffic profile based on microscopic monitoring traces
- NS-2 simulations based on replaying actual traffic traces

Simulation principles:
- Elephant flows are transmitted using TFRC
- Others flows use TCP New Reno
Traffic parameters

- Classical traffic ones: throughput mean and standard deviation

- One related to traffic variability:

  Stability Coefficient (SC) = \frac{\text{exchanged average traffic}}{\text{exchanged traffic standard deviation} (\sigma)}

- QoS statistical one: LRD
  \Rightarrow \text{Estimation of traffic oscillating range}
TFRC impact on flow QoS: throughput analysis

Table 1. Throughput evolution during time for TCP and TFRC protocols

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Average throughput (kB)</th>
<th>Throughput $\sigma$ (kB)</th>
<th>SC</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCP New Reno (NR): real case</td>
<td>82.335</td>
<td>157.959</td>
<td>0.521</td>
</tr>
<tr>
<td>TCP NR &amp; TFRC: simulated case</td>
<td>77.707</td>
<td>102.176</td>
<td>0.761</td>
</tr>
</tbody>
</table>
TFRC impact on flow QoS: LRD analysis

Real traffic

Simulated traffic

Logscale Diagram, N=5 \[ (\alpha_{L2})= (9,15), \ \alpha_{est} = 0.873, \ Q= 1.6199\times10^{-07} \]

Logscale Diagram, N=3 \[ (\alpha_{L2})= (5,12), \ \alpha_{est} = 0.127, \ Q= 1.2244\times10^{-05} \]
Partial conclusion

1. Traffic oscillations highlighted:
   - Causes (TCP + elephants + network capacity)
   - Illustration of the bad impact of LRD on QoS

2. TFRC which generates smoother traffic than TCP
   - Helps to optimize performances
   - Smoothing traffic is essential for being able to guarantee stable QoS
   - Validate the use of LRD to characterize oscillations
But...

- TFRC limitation problem: it cannot generate more traffic than TCP (equation based...)
- It cannot benefit from the traffic characteristics improvements
Additional problematics (1)

New traffic analysis exhibited that:

- Traffic characteristics are different on different links, at different times...
- Traffic is not stationary
- Many ruptures arises
  - On daily, weekly, monthly yearly basis
  - Random unexpected ruptures
    - Failures, Byzanthin behaviours
    - DoS attacks
    - Legitimate traffic
Additional problematics (2)

Topological issues for end to end QoS

- The Internet is split into AS and domains
- Each domain / AS is designed and managed without regard of other domains / AS
- Few cooperations between carriers and ISP
  → they are competing to attract clients
Measurement Based Networking

Principle: Extend preceding approach (MBNE) with mechanisms reacting in real time to measurements performed in a large number of points of the network

Points to address:

- RT measurement system (passive and active)
- Measurements signaling
- Mechanisms to reacting to measurements (routers or end hosts)
Measurement Based Architecture

Router with active and passive measurement system

Signaling of measurements: intra- (& inter-)domain(s)

End host able to take into account measurement information

Intra-domain measurement

Intra & inter-domains measurement

AS 1

AS 2

AS 3

Séminaire OLC, 26 avril 2004
RT measurement system

▪ What to measure?
  ▪ Throughput (passive in intra-domain, active in inter-domains) and available capacities
  ▪ Ruptures in the traffic (attacks, events driven, ...)
  ▪ Traffic matrices (passive)
  ▪ Oscillations (passive)
  ▪ Losses (passive)
  ▪ Delays (active)
  ▪ ...

▪ RT ⇒ what granularity?
Signaling system and protocol

- What parameter to signal?
- How to signal these parameters?
  - COPS, SIP, ...
  - Mcast/P2P/... ?, Push/pull?
Reaction to measurements

- How to react to measurements?
- Are measurements trustable?
  - Especially for inter-domain on a market where competition is the standard
- What to do if measurements are missing?
  - Signaling issue if the network is congested
  - Differentiated QoS services (PQ, ...)
  - Game theory: dead reckoning
MBN case study

- Let us assume:
  - RT measurement system is available
  - Signaling system is available
MBCC objectives

- Smooth sending rate
- Optimize resources and in particular bandwidth utilization
- Provide fairness between competing flows
- Continue working efficiently even if measurements are missing

Measurement based extension of TFRC
MBCC equations

- Only applied to elephants
  - Mice traffic is a white Gaussian noise
  - Elephant flow arrival is a Poisson process

\[
\begin{align*}
\text{if } (LER = 0), \quad X_{MBCC} &= X_{TFRC} + \frac{\text{total\_available\_BW}}{N} \\
\text{if } (LER \neq 0), \quad X_{MBCC} &= \min(X_{TFRC}, \text{flow\_consumed\_BW})
\end{align*}
\]

\(N \rightarrow 2N\) in case of low BW network (wireless)
MBCC test topology

TCP SACK & MBCC Senders (Elephant Flows)

TCP New Reno Senders (Mice Flows) NS-2 replay module

100 Mbps 20 ms

2 Mbps 75 ms

Passive Probe

TCP SACK & MBCC Receivers

TCP New Reno Receivers

100 Mbps 20 ms
MBCC simulations scenarios

- Scenario 1:
  - Mice sent using TCP New Reno
  - Elephants sent using MBCC

- Scenario 2:
  - Mice sent using TCP New Reno
  - Elephants sent using TCP SACK

- Scenarios 3:
  - Both mice and elephants sent using MBCC
### MBCC vs. TCP results

<table>
<thead>
<tr>
<th></th>
<th>Mice Flows</th>
<th>Elephants Flows</th>
<th>Mice &amp; Elephants Flows</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TCP New Reno</td>
<td>MBCC</td>
<td>TCP SACK</td>
</tr>
<tr>
<td>Average Throughput (B/s)</td>
<td>10348,7</td>
<td>243738,2</td>
<td>242014,2</td>
</tr>
<tr>
<td>Throughput Standard Deviation (B/s)</td>
<td>6898,3</td>
<td>16350,5</td>
<td>19585,5</td>
</tr>
<tr>
<td>Stability Coefficient</td>
<td>1,500</td>
<td>14,907</td>
<td>12,357</td>
</tr>
</tbody>
</table>
MBCC vs. TCP losses

Eléphants (MBCC) + Souris (TCP New Reno)
Eléphants (TCP SACK) + Souris (TCP New Reno)
MBCC seul
MBCC vs. TCP LRD

- Scenario 1
  
  \[ H = 0.67 \]

- Scenario 2
  
  \[ H = 0.88 \]
MBCC in a multi-domain case
### Results

<table>
<thead>
<tr>
<th>Scenario 1</th>
<th>Scenario 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MBCC</strong></td>
<td><strong>TCP New Reno Bottleneck n°1</strong></td>
</tr>
<tr>
<td>Average Throughput (B /s)</td>
<td>109434.9</td>
</tr>
<tr>
<td>Throughput Standard Deviation (σ) (B /s)</td>
<td>31840.4</td>
</tr>
<tr>
<td>Stability Coefficient (SC)</td>
<td>3.437</td>
</tr>
</tbody>
</table>
MBCC vs. TCP losses

![Graph showing cumulative loss number per second over time for Scenario 1 and Scenario 2.](image-url)

- **Scenario 1**
- **Scenario 2**

---

Séminaire OLC, 26 avril 2004
MBCC vs. TCP LRD

Fig. 8. Traffic LRD estimation (simulation 2, Sc = scenario)
Conclusion

- MBNE (Network Engineering) → MBN
- MBN/MBA proved to work well on congestion control
- MBA and MBN principle can be used in other areas than congestion control
- Still to be done
  - RT measurement system
  - MBN measurement signaling
- Proposed for EuQoS
Other related future projects

- GdX + Grid’5000 → MBA/MBN experimentations
  - STM: Emulation methodology and traffic sources
- MetroSec: improving network robustness to ruptures