

Measurement Based Approach of Congestion Control for enforcing a robust QoS in the Internet

Philippe Owezarski, Nicolas Larrieu

LAAS-CNRS

7, avenue du Colonel ROCHE

31077 TOULOUSE Cedex 4

FRANCE

Email: {owe, nlarrieu}@laas.fr

Abstract—How to provide quality of service (QoS), has been a major issue for the Internet for the past decade. But, recent monitoring projects showed that Internet traffic exhibited large variations, and non-stationary traffic, making difficult to guarantee a stable and robust QoS. The objective of this paper is then to guaranty a robust QoS which means providing the requested QoS under all circumstances, including the most difficult ones. Among the most difficult circumstances are enforcing QoS even in the presence of denial of service (DoS) attacks. This paper then proposes to use a measurement based architecture (MBA) suited for coping with actual non-stationary traffic, as well as traffic disruptions or anomalies. The idea of our measurement based networking (MBN) approach relies on a real time analysis of traffic characteristics and QoS evolution, and on the design of mechanisms able to adapt their reactions accordingly. In particular, we designed a new congestion control mechanism, called MBCC (Measurement Based Congestion Control), able to cope with the large variability in traffic throughput. MBCC proved to optimize the use of resources and to improve QoS. In this paper, we show that using MBCC instead of TCP makes the Internet more robust to DoS attacks, i.e. the QoS provided by the attacked network using MBCC, is better than using TCP. These preliminary results are shown on different NS-2 simulations.

Keywords. Traffic characterization, Measurement Based Networking, congestion control, QoS, Internet robustness

Acknowledgment — This work is achieved in the framework of the METROSEC project, a French national project granted and funded by the French ministry of research, CNRS, INRIA and DGA It has also been partly supported by the European Union under the E-Next Project FP6-506869.

I. INTRODUCTION

How to provide quality of service (QoS), has been a major issue for the Internet for the past decade. Though many proposals have been put forward, such as IntServ, DiffServ, and others, until now none have met the needs of users or operators (Internet service providers, carriers, etc.).

Guaranteeing QoS means providing the requested QoS under all circumstances, including the most difficult ones. Among the most difficult circumstances are denial of service (DoS) attacks. Because of this, protection against DoS is a defining characteristic for guaranteed QoS mechanisms (especially for high level QoS classes).

But, the line between a DoS attack and legitimate traffic is blurred. That is why, in our work, attacks are included in a wider family of events we call “disruptions” in the traffic. Traffic disruptions include all events that provoke a large change in network traffic characteristics, and that can badly impact the QoS provided by the network, be they DoS or legitimate marked variations as flash crowds. Indeed, Internet traffic is very far from being regular, and presents large variations in its throughput at all scales [14]. These projects have shown that Internet traffic exhibits characteristics such as self-similarity [15], (multi-) fractality [6], and long-range dependence (LRD) [5], which is to say in all cases that traffic can vary significantly.

Such significant variations are very damageable for enforcing a stable QoS. Our previous work [12] analyzed the impact of traffic high variability on QoS. In particular, it showed that such variability is mainly due to the congestion control mechanisms of the TCP transport protocol (the main protocol used in the Internet) that proved to be inefficient for transmitting large flows on high speed networks. And recent evolution of network usages are due to the arrival of P2P applications for exchanging files, in particular music tracks or movies which range from few Mbytes to several Gbytes. Facing such large flows (called elephants), traffic generated by TCP is highly variable, made of very large traffic bursts, which induce “disruptions”. We have also shown [12] that such oscillations are related to the LRD notion. The impact of LRD (and then oscillations) on QoS has been analyzed, showing that the greater the LRD, the lower the QoS. In addition, we have already illustrated this point with DoS attacks [13], showing how attacks impact the LRD and then decrease the network QoS and performance.

Given these previous traffic analysis results, this paper then proposes to use monitoring techniques and to take advantage, in real time, of monitoring results to adapt to frequent changes in the traffic and especially traffic anomalies or disruptions. Monitoring and measurement techniques are the basis of this new approach for managing communications, traffic and QoS in the Internet. This paper then proposes to use our measurement based architecture (MBA), we introduced recently

[9], and the related MBCC congestion mechanism. We have already demonstrated that MBCC, when facing current self-similar and LRD traffic, limits LRD, induces less variability, and a higher and stable QoS [9]. In this paper, we will show that MBCC is also suited for coping with actual non-stationary traffic, including traffic disruptions or anomalies, and improves network robustness (i.e. contributes to continue providing a good QoS even in the case of DoS attacks).

This paper shortly summarizes our measurement based approach, called MBN for Measurement Based Networking (Section II), the MBA principles (section II-A), and the related MBCC congestion control mechanism (section II-C). The main objective of this paper is to show that, in addition of solving QoS issues for normal self-similar and LRD traffic, MBCC is very robust to any kinds of traffic disruptions / anomalies, and especially DoS attacks (which appears as the most difficult case for the network to enforce a guaranteed QoS). MBCC robustness has been evaluated thanks to NS-2 simulations. In particular, section III shows that DoS attacks have a very limited impact on network QoS when MBCC is used instead of TCP. The comparative results between TCP and MBCC in the presence of attacks are also proposed. Finally, section IV concludes this paper.

II. MEASUREMENT BASED NETWORKING PRINCIPLES

A. Measurement Based Architecture design

Given all the issues related to current traffic as un-stability, non-stationary nature, huge oscillating nature, correlation, dependence, and a huge versatility of traffic types during time, it is easy to understand that it is impossible to find a static solution optimal for all connections in the Internet. This statement leads us proposing MBN in order to react in real time and locally to some events on the network.

The first requirement of MBN is then to be aware of the network and traffic changes. It is then necessary to measure traffic and QoS parameters locally, as well as on long distances when the connection crosses several domains or autonomous systems (AS), and to exchange measurement and monitoring results between all concerned components in the network. Figure 1 depicts how measurement tools can be deployed in the Internet for this purpose. This figure more specifically depicts the case of a MBCC connection from a source to a destination crossing two Internet AS, and all the core and edge routers having measurement capabilities and running MRP (Measurement Reporting Protocol), the protocol we designed for signaling measurement results to concerned network components. In this architecture, core routers provide intra-domain measurements, while edge routers provide inter-domains measurements / estimations¹.

All these measurements performed in real time and reported to all concerned network components (traffic sources for

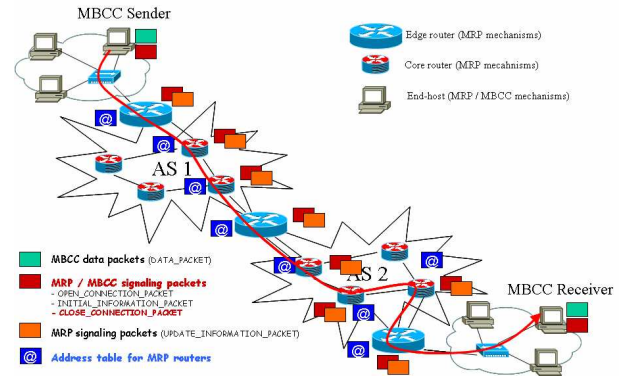


Fig. 1. Network entities needed in the network to deploy MBN: case of the MBCC congestion control

instance), can give an accurate knowledge of network and traffic state, and allow them to perfectly adapt their sending rate (for instance) to available resources. Note however that one important aspect of MBN deals with the design of the protocol for reporting measurement information.

B. Measurement Reporting Protocol principles

The measurement reporting protocol is then a key component of the MBA architecture. The challenges for MRP are first to be efficient, i.e. to provide traffic information very rapidly, in order to make network components react according to very recent information on network state. But MRP must also be scalable. This means that MRP must generate a limited amount of traffic, not to impact the way the network is working. For scalability purpose, MRP components also have to store a limited amount of data: their link and connection table has to be as small as possible, with a limited number of entries, in order to reduce to the maximum the time required for searching any information in it.

Figure 1 depicts the way MRP is working (more details are given in [9]). First of all, MRP is closely related to connections, i.e. the signaling path has to be the same as the one of the considered connection. For this, we are then using the principle of RSVP [3] with a first packet that finds a path from the source to the destination, and then, a reverse packet that goes back to the source. The differences are that the back packet that is a reservation packet in RSVP is in MRP a signaling packet, transporting measurement information. As well, the signaling packets are sent any times it is required, whereas in RSVP they are just sent once at the connection opening. In fact, MRP only uses the RSVP principle of being able to find a path and to go back along this path. Proceeding like this allows MRP to perfectly identify which are the concerned network components (routers) on the path, and to limit the number of sources and destinations for the measurement messages.

For addressing the scalability issue we chose:

- To consider only elephant flows. In fact, mice do not create troubles in the traffic, and all issues that appear in

¹For more information on measurements and monitoring tools and techniques used in MBA, interested readers can refer to this detailed description of MBN approach [9].

the traffic are related to the sending of elephants. Then MRP routers just keep information on elephants passing through them. This is a way to limit the amount of entries in the connection table, as elephants just represent a very small proportion of the amount of flows;

- To send measurement information only when some disruptions arise in the traffic, in order to generate traffic only when the network conditions are changing. It then should limit the amount of signaling traffic, and provide to senders or routers important information very rapidly.

Proceeding like this, we solve the scalability problem that has been previously encountered in the Internet while addressing the QoS management and enforcement issues².

C. Measurement Based Congestion Control principles

The main MBCC objective is then to bring more stability to elephant flows. To increase elephant flows regularity (i.e. to suppress observable oscillating behaviors at all scales), the new TCP Friendly Rate Control (TFRC) congestion control mechanism seems to be able to provide a great contribution. TFRC has been designed to provide a service suited for stream oriented applications requiring smooth throughputs. TFRC, then, tries as much as possible to avoid brutal throughput variations that occur with TCP because of loss recovery. The sending rate of each TFRC source is made thanks to a receiver oriented computation, that calculates, once by RTT, the sending rate X_{TFRC} according to the loss event rate p measured by the receiver [7] according to equation 1:

$$X_{TFRC} = \frac{s}{R * \sqrt{2 * b * \frac{p}{3}} + (t_{RTO} * (3 * \sqrt{3 * b * \frac{p}{8}}) * p * (1 + 32 * p^2))} \quad (1)$$

where:

- X is the transmit rate in bytes/second,
- s is the packet size in byte,
- R is the round trip time in second,
- p is the loss event rate (between 0 and 1.0), of the number of loss events as a fraction of the number of packets transmitted,
- t_{RTO} is the TCP retransmission timeout value in second and is normally equal to $4 * R$,
- b is the number of packets acknowledged by a single TCP acknowledgement.

The benefits of using TFRC instead of TCP have been already demonstrated [8]. However, if TFRC is able to cope with TCP oscillations, it is not able to perfectly adapt to more brutal disruptions in the traffic due to DoS attacks. The measurement based approach is proposed as a solution to cope with such traffic disruptions / anomalies. However, in order to take advantage of all the TFRC benefits, MBCC has been designed as an extension of TFRC including some capabilities for using measurement results coming from monitoring equipments in the network.

The main principle of MBCC then consists in using the TFRC algorithm for computing the nominal sending rate of each elephant connection, and to correct it thanks to the

knowledge of the available and consumed bandwidth in the network. Then, if some bandwidth is available, sources should generate more traffic than indicated by TFRC throughput equation 1 (X_{TFRC} corresponding to a TCP flow throughput) without creating any loss or congestion in the network. Thus, the network congestion level should be significantly decreased by having “proactive” sources able to adapt in real time their sending rate according to available resources. As well, such mechanism will help to improve fairness between flows, as the correction on the sending rate will not depend on the RTT value, but on the real bandwidth available, equally shared between competing flows.

MBCC can only be used for elephants which are the flows inducing the more disturbances in the network [8]. At the opposite, as “mice” traffic represents a white Gaussian noise [2] and then does not induce large transfer issues, it is not necessary to modify their transport protocol.

Then, for a normal period (when monitoring information is correctly received and when there is no congestion, i.e. there is some available bandwidth), each elephant flow can get an additional fraction of available resources. This fraction is computed by dividing the total available bandwidth by the average number of elephant flows in the network at this time. It makes sense to divide the available bandwidth by the average number of active flows (N) crossing this link, as it has been demonstrated that elephant flow arrivals nearly follow a Poisson process [2]. Indeed, in a Poisson process, as mean equals variance, the average number is significant because the process values will never be far from this average. In addition, as the traffic considered has some elastic properties, the number of losses will be quite reduced. Indeed, we already checked [9] that the bandwidth usage optimization overruns the wastes due to retransmissions.

At last, for a congested period, MBCC senders have to reduce their sending rate for resolving congestion, and this trying to be as fair as possible. Thus, MBCC sources send the minimum value between the possible TFRC throughput and the effective throughput got by the flow at this time in the bottleneck of the network (this information being given by monitoring tools met all along the path).

So, the equations of this algorithm can be summed up as follows:

- For a period without congestion ($p = 0$):
 $X_{MBCC} = X_{TFRC} + aBW_{flow}$;
- For a congested period ($p \neq 0$):
 $X_{MBCC} = \min(X_{TFRC}; cBW_{flow})$;

Where:

- aBW_{flow} is the Available BandWidth per flow in the bottleneck link(s) of the path. It is computed with the ratio $\frac{\text{total available bandwidth}}{N}$, this information is provided by MRP routers met on the path;
- cBW_{flow} is the Consumed BandWidth per flow in the bottleneck link(s) on the path, this information is provided by MBCC receiver with other end-to-end information such as RTT and loss ratio for instance; RTT estimation as described in equation 1.

²MRP performances have been accurately evaluated previously [9].

III. EXPERIMENTAL VALIDATION OF MBN CONTRIBUTION TO NETWORK ROBUSTNESS

MBN proved to greatly improve network performance and QoS when facing normal self-similar and LRD traffic [9], but we expect it to make network less sensitive to DoS attacks. By avoiding traffic bursts in the presence of attacks, we expect to limit congestions, losses, and unsuited TCP responses leading to traffic instability and QoS decrease. Therefore, in this section we present experiment results which illustrate MBCC capabilities to improve QoS robustness. In particular, we compare network QoS when MBCC or TCP are facing traffic containing DoS attacks.

A. Simulation principles

These new congestion control (MBCC) and reporting (MRP) mechanisms have been implemented and evaluated using NS-2. It has then been needed to develop a set of tools for monitoring available and consumed bandwidth in the simulated network and to exchange the measurement results between routers and traffic sources.

1) *Simulation topology*: The topology used is described in figure 2. In these simulations, we have created one bottleneck link to improve concurrent behaviors between the different flows. Then, elephant flows, using either MBCC or TCP SACK, and cross traffic using TCP New Reno are transmitted in order to make them compete in the bottleneck³. The goal is then to study how they behave the ones relatively to the others, and to compare the impact of DoS attacks on network QoS when MBCC or TCP are used for sending elephants.

The core link represents then the most “congested” link on the considered path, i.e. the one that will mainly influence the MBCC and TCP sending rates. Every simulation is based on real traffic traces collected on the Renater⁴ network. These traces are replayed in NS-2 with a special methodology [11] whose goal is to make simulation realistic, i.e. replay in simulations traffic samples in order to reproduce all characteristics of real traffic, with all its variability and LRD characteristics for example (interested readers can refer to an evaluation of this monitoring-based replay methodology already published [11]).

In simulation, short and long flows are differentiated. First ones (mice) do not induce any transfer problem in the network. Thus, they will be transmitted using TCP and more precisely TCP New Reno that is the most frequent version of TCP in the Internet. At the opposite, elephant flows create in the network long range oscillations which induce congestions. This is the reason why MBCC has been designed for suitably transmitting such elephants flows. Thus, simulations compare the case in which elephants are transmitted using our new MBCC congestion control mechanism, and the one in which they are transmitted using TCP SACK⁵.

³Cross traffic potentially contains DoS attacks.

⁴Renater is the French National Network for Education and Research.

⁵TCP SACK serves as the TCP reference as it is the best performing version of TCP.

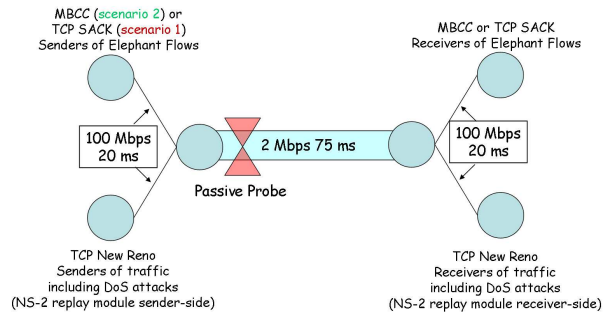


Fig. 2. Network topology used in NS-2 simulations

For evaluating QoS robustness when MBCC vs. TCP are used, we have replayed a traffic trace including DoS attacks. This trace has been captured on LAAS’ access network to the Internet one day we launched a UDP flooding attack toward our lab. The trace we used lasts 40 minutes. This capture was realized with DAG equipments [4]. Its throughput characteristics are represented in figure 3. The sixteen first minutes (cf. slot 1 on figure 3) represent standard Internet traffic containing mainly classical applications: web, mail ou ftp, etc. The rest of the trace (cf. slot 2) contains in addition to the same initial traffic multiple DoS attacks generated from one computer placed out of the LAAS’ network toward a specific computer placed in LAAS’ network. More precisely, the DoS attacks are UDP flooding consisting of 10,000 packets and whose parameters as attack strength and frequency change. More precisely:

- attack strength deals with the size of UDP packets sent (0, 20, 40, 100, 1000 or 1500 bytes);
- attack frequency deals with the time period between two consecutive UDP packets: 100 ns, 1,000 ns or 10,000 ns.

Consecutive attacks are separated by a 30 seconds idle time.

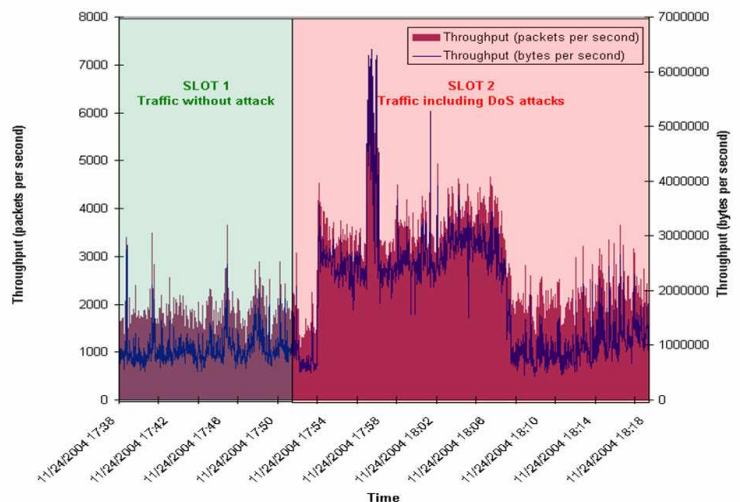


Fig. 3. DoS attack characteristics

2) *Analysis parameters*: The main goal of this study is to measure MBCC adaptation capabilities to network anomalies

as DoS attacks, and to compare the results with other congestion control mechanisms. For evaluating MBCC and its contribution to QoS robustness (good QoS being associated to a smooth traffic having a low LRD), several parameters have been evaluated in simulations:

- throughput evolution by traffic type (TCP or MBCC) to study the traffic variability: computing of average throughput (A), standard deviation (σ) and stability coefficient ($SC = \frac{A}{\sigma}$);
- loss process evolution in order to evaluate MBCC adaptation capabilities compared to TCP;
- traffic oscillation range by computing the Hurst factor.

B. MBCC contribution to QoS in the presence of attacks

This simulation includes two different scenarios: in scenario 1, elephant flows are transmitted using TCP SACK, while in scenario 2 elephants are transmitted using MBCC. First scenario is used as a reference experiment. In these two scenarios, the cross traffic (if we exclude DoS attacks) is a normal Internet traffic consisting of both mice and elephants. They both are sent using TCP New Reno.

Each simulation lasts 200 seconds. 100 elephants and about 4000 cross flows (consisting of both mice and elephants), and attack constituting packets have been replayed. Simulation results are the following: first, traffic throughput was computed for the slot period 2 (traffic with DoS attack). Table I shows result values for scenarios 1 and 2. This experiment then exhibited that MBCC performs better than TCP SACK, as throughput and resource usage are higher, and the traffic is also much smoother. Moreover, it seems that another interesting information deals with cross traffic in bottleneck when there is MBCC elephant traffic in the network (scenario 2). We can see that in the case where TCP SACK is used for transmitting elephants (scenario 1), cross traffic average throughput is lower and exhibits more variability than when MBCC is used in the network (cf. for instance $SC(\text{TCP New Reno}_{\text{Scenario 1}}) < SC(\text{TCP New Reno}_{\text{Scenario 2}})$). Thus, this demonstrated that MBCC helps to make the network more robust and traffic much smoother than TCP SACK in the presence of DoS attacks.

TABLE I
TRAFFIC VARIABILITY ANALYSIS

	SLOT 2: traffic including DoS attacks					
	Scenario 1 (TCP)			Scenario 2 (MBCC)		
	Global traffic	Cross traffic	Elephant traffic (TCP SACK)	Global traffic	Cross traffic	Elephant traffic (MBCC)
Average throughput (bytes/s)	243,872.13	221,827.32	22,044.81	248,004.64	227,691.13	22,313.51
Throughput standard deviation (bytes/s)	31,553.07	83,943.63	44,704.32	22,690.51	57,127.70	31,840.41
Throughput stability coefficient (SC)	7.73	2.64	0.49	10.93	3.99	0.70

This result is confirmed with the loss process analysis. Indeed, figure 4 depicts a more important loss level in the network when using TCP SACK than when using MBCC. This result has been analyzed both on elephant only traffic (cf. figure 4(a)) but also on global traffic exchanged (cf. figure 4(b)). Indeed, traffic variability with TCP is more important, making congestion more likely to occur in the network and the loss number higher.

Finally, as it is depicted on figure 5(a), MBCC impacts in a very positive way traffic LRD⁶. In fact, thanks to MBCC, the LRD is much reduced in scenario 2 (where $H = 0.69$) compared to the reference scenario with TCP SACK where LRD is very high ($H = 0.88$). Consequently, there are less oscillations (cf. related stability coefficient values of table I), this feature inducing more stability on traffic profile and then less congestion in the network.

As a conclusion, these results prove that MBCC makes the network QoS more robust than when using TCP. Indeed, the impact of an UDP flooding attack is much more limited with MBCC than with TCP: throughput is higher, loss and congestion ratio are lower, and the LRD is much lower. This implies a smoother traffic and a better QoS. In particular, the results show that the performance of MBCC with and without attacks are very close, demonstrating the insensitivity of MBCC to attacks, and then explaining the improved network QoS robustness.

IV. CONCLUSION AND FUTURE WORK

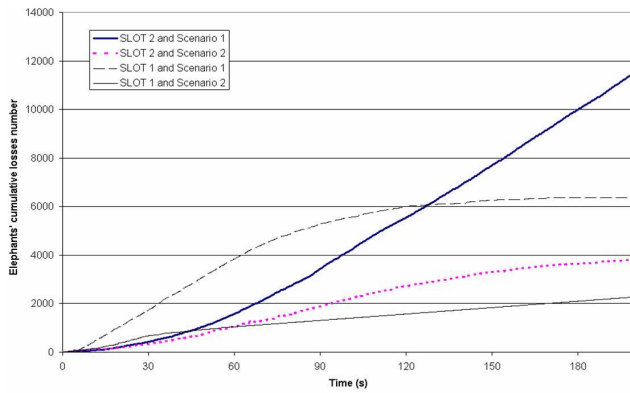
This paper has illustrated the contribution of monitoring and measurements for improving network robustness and for continuing providing high quality services even when they are attacked. We proposed a new measurement based networking approach and its related congestion control mechanism (MBCC), based on the use of on-line traffic analysis, for adapting in real time to network and traffic conditions. MBCC proved to reach all its objectives and improves significantly network QoS by providing a very regular traffic (reducing that way traffic LRD responsible of most of QoS and performance issues). But this paper has shown that MBCC improves also QoS robustness, and even when facing attacks, MBCC continues to provide to users (almost) the same QoS level. In addition, results showed that MBCC in the presence of attacks performs even better than TCP without attacks.

However, a lot of work remains to do. First, this paper only focuses on UDP flooding attacks, as they impact the network in a similar way as normal (TCP) traffic variations (except that they impact LRD at a single scale, whereas TCP variations impact LRD at all scales). Therefore future work includes the analysis of the impact of many other kinds of DoS attacks, and in particular the new family of "light attacks". Given the results of DoS attacks analysis, we will continue studying new MBN solutions for improving network QoS, and making it more robust. In particular, MBCC will be improved, but we also plan to adapt router mechanisms and routing protocols for fighting the impact of DoS attacks.

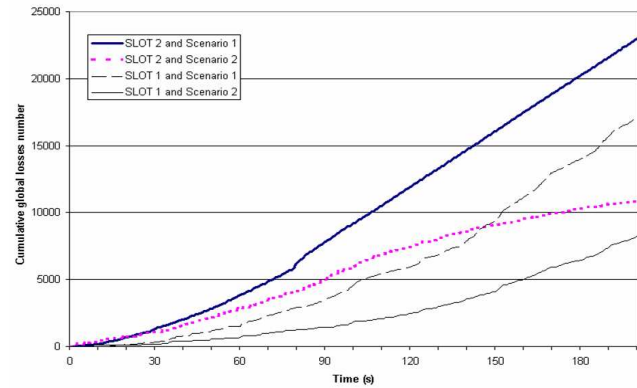
REFERENCES

- [1] P. Abry, D. Veitch, *Wavelet Analysis of Long Range Dependent Traffic*, Trans. Info. Theory, Vol.44, No.1 pp.2-15, Jan 1998.
- [2] N. Ben Azzoune and F. Guillemin, *Analysis of ADSL traffic on an IP backbone link*, In Proc. Globecom 2003, San Francisco, December 2003.

⁶LDR diagrams are produced thanks to the LDEstimate tool, developed by Abry and Veitch [1].



(a) Elephant traffic



(b) Global traffic

Fig. 4. Congestion level estimation

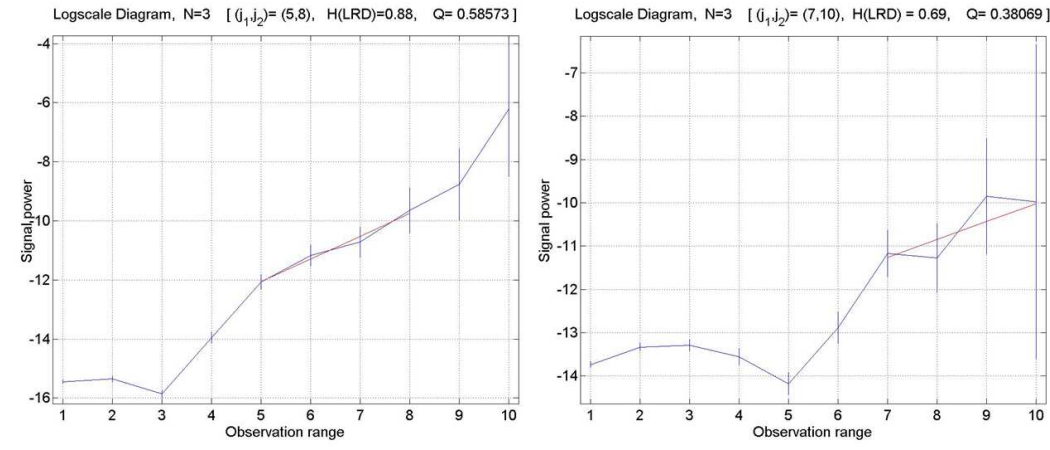


Fig. 5. Traffic LRD estimation with TCP (left figure) vs. MBCC (right figure)

- [3] R. Braden, L. Zhang, *Resource ReSerVation Protocol (RSVP) – Version 1 message processing rules*, RFC 2209, September 1997.
- [4] J. Cleary, S. Donnelly, I. Graham, A. McGregor and M. Pearson, *Design principles for accurate passive measurement*, PAM (Passive and Active Measurements) Workshop, Hamilton, New Zealand, April 2000.
- [5] A. Erramilli, O. Narayan, W. Willinger, *Experimental queuing analysis with long range dependent packet traffic*, IEEE/ACM Transactions on Networking, Vol. 4, No. 2, pp 209–223, 1996.
- [6] A. Feldmann, A. Gilbert, and W. Willinger, *Data networks as cascades: Investigating the multifractal nature of Internet WAN traffic*, Proc. of ACM SIGCOMM'98, Vancouver, Canada, 1998.
- [7] S. Floyd, M. Handley, J. Padhye, J. Widmer *Equation-Based Congestion Control for Unicast Applications*, SIGCOMM 2000, August 2000.
- [8] N. Larrieu, P. Owezarski, *TFRC contribution to Internet QoS improvement*, in Proc. of the fourth COST 263 international workshop on Quality of Future Internet Services (QoFIS'2003), Stockholm, Sweden, 1-3, October 2003
- [9] N. Larrieu, P. Owezarski, *Measurement based networking approach applied to congestion control in the multi-domain internet*, to be published in the proceedings of the 9th IEEE / IFIP International Symposium on Integrated Network Management (IM'2005), Nice, France, May 15th-19th, 2005.
- [10] M. Mathis, J. Mahdavi, S. Floyd, A. Romanow, *TCP Selective Acknowledgment Options*, Request for Comments 218, October 1996.
- [11] P. Owezarski, N. Larrieu, *A trace based method for realistic simulations*, IEEE International Conference on Communications (ICC'2004), 20-24 June, Paris (France).
- [12] P. Owezarski, N. Larrieu, *Internet traffic characterization – An analysis of traffic oscillations*, 7th IEEE International Conference on High Speed Networks and Multimedia Communications (HSNMC'2004), Toulouse, France, July 2004.
- [13] P. Owezarski, *On the impact of DoS attacks on Internet traffic characteristics and QoS*, 14th IEEE International Conference and Computer Communications and Networks (ICCCN 2005), San Diego, CA, USA, 17-19 October 2005.
- [14] K. Park, G. Kim, M. Crovella, *On the relationship between file sizes, transport protocols, and self-similar network traffic*, IEEE ICNP, 1996.
- [15] K. Park and W. Willinger, *Self-similar network traffic: an overview*, In Self-similar network traffic and performance evaluation, J.Wiley & Sons, 2000.