Chapter 7

Metrology of Internet networks[113]

7.1. Introduction

Internet is changing in terms of its uses. Monoservice network to transport binary or textual files twenty years ago, Internet must today be a multi-service network for the transport of diverse and various data such as audio and video (movies, video on demand, telephony, etc.). Indeed, there is a need to operate a technological mutation of the network to make it capable of transporting different types of information offered by all the applications using Internet, with adequate and multiple Quality of Service (QoS). However, all attempts to ensure the Quality of Service of Internet have failed, especially because of a complete ignorance of Internet traffic and the reasons of this complexity. In the end, as it exists today, nobody has controlled it, not even a complete knowledge of the network, which goes against the implementation of multiple services of communication of guaranteed quality.

The metrology of Internet networks – literally « the science of measurements » applied to Internet and its traffic – must provide an answer to these problems. First, to provide services with predetermined qualities, it must be able to measure these qualities. Secondly, metrology must answer the questions about the model(s) of Internet traffic that are now missing. If recent science it is (it appeared in the early

---

Chapter written by Nicolas LARRIEU & Philippe OWEZARSKI.

2000s), it changes the whole process of research and engineering of Internet networks and becomes the cornerstone.

The metrology of Internet is divided into two distinct tasks: the first consists of measuring physical parameters of Quality of Service offered by the network or on the traffic. In a network of the size and complexity of Internet, it is already – we will see – a complex task. However, this activity of measurement and observation allows only the highlighting of visible phenomena. But, in the case of networks, it is even more important to deduce the causes, that is to determine the components and/or mechanisms of protocol which generate them. We find ourselves in fact facing the same problem as Plato in the allegory of the cave [PAL 02]. In this cave where a wood fire was crackling, he saw only the shadow of the men wandering in the cave. And the shadows projected on the walls of the cave were huge, nearly giving the impression they were those of giants. Metrology – measurement of the QoS or simple analysis of traffic – confront us with Platonic problem. It only shows us the effects of all the mechanics of networks, while what intrigues us is the second part of metrology: by highlighting the causes of the shortcomings of Internet, we pave the way of research to change the mechanisms, architectures and protocols of Internet.

This chapter introduces therefore the problems associated with the implementation of measurements globally in computer networks such as Internet, which will be the target of all the works described below (Section 7.2). It shows the different active and passive measurement techniques, their needs, their qualities and faults (Section 7.3). Then, from measurements or observations on the traffic, this chapter highlights the causes of the current limitations of traffic growth, thus showing the importance of metrology for research and networks engineering 9Section 7.4). Section 7.5 concludes this chapter by stating a number of networks domains that derive large profits from using metrology.

7.2. Problems associated with Internet measurement

When a specialist of computer networks talks about measurement of Internet, this can, depending on the cases, be related to three types of metrics to measure:

– the size of Internet, which is expressed in number of machines installed in the Internet;

– the traffic carried on the links (wired or wireless) of Internet;

– the measurement of Internet QoS.

However, the term QoS can have different connotations depending on who uses it. To simplify the problem, we can consider that there are two main viewpoints:

– *point of view of applications* (or of users): each application has needs in terms of « delay, throughput, loss », the famous triptych for the new generation computer networks. These needs are obviously different from one application to another, and each application would wish to have a communication service specifically tailored to its needs;

– *point of view of operators*, whose objectives are to optimise the utilisation of the resources of the communication infrastructure (and, thus, maximise their earnings), to limit losses and delays, and to charge fairly and consistently the services provided to users.

Another characteristic of the measurement of physical parameters of QoS (also called simple) is related to the size of the Internet, consisting of a multitude of interconnected networks, themselves consisting of numerous nodes and points of presence in the different cities served around the world. Thus, it is important to break the measurements into two new classes:

– end-to-end measurements, concerning the triptych « delay, throughout, loss » between two user terminals. These are measurements that join the point of view of applications aforementioned;

– step-by-step measurements, concerning the triptych « delay, throughput, loss » between intermediary network nodes (routers, switches, gateways, etc.). These are measurements that join the point of view of operators.

7.2.1. *Geographical and administrative dimension*

Clearly, the major inherent difficulties in the realisation of traffic or QoS measurements are related to the size of the network and its geographical expanse. Thus, it is essential to have many measurement points in the network. To watch the traffic at a single point at a given time is finally quite insignificant because Internet communications are often made over very long distances. However, Internet is an interconnection of networks, and, as such, multiple operators are owners and in charge of the management and operation of one of these networks that make up the Internet. Also, it is today impossible to position measurement tools at certain points of the network against the will of the operator who owns and manages this point. To carry out the metrology of the network, we should be able to place measurement points at all network nodes, but the construction of Internet prohibits it by default. Furthermore, « to monitor » all Internet nodes is a titanic task. To circumvent this difficulty, many works have been underway for three to four years on sampling problems in both space and time. The idea would be to find techniques which, from some measurement points that would work only during short periods of time, would accurately estimate the traffic on the entire network over 24 hours. But, today, despite the efforts made, the results on sampling are at a standstill and, indeed, there are no significant positive results in this field. Furthermore, the global scope of Internet leads to daily variation in activity. It thus makes little sense to watch only one point of the network because on one continent where night has fallen, for example, an intercontinental connection can be faced with a potentially more reduced traffic (which will therefore pose few problems concerning compliance to QoS constraints), then crossing a few milliseconds later a continent during the peak of the day, and thus come up to an important traffic and to possible congestion phenomena very problematic for the communication QoS. In such a context, it is thus essential to have measurement points on these different continents because the local traffic can have significant consequences on the intercontinental connection which we were considering. We find the problem mentioned above regarding end-to-end or step-by-step measurements. We clearly see from this intercontinental connection example that end-to-end measurement only gives an overall result of the actual QoS for this connection, while step-by-step information help locate the problem, causing service limitations.

7.2.2. *Distribution problems*

Another set of problems to perform measurements on Internet are related to distributed systems – problems even more difficult to solve as Internet is vast and, more importantly, uncontrolled in its entirety by a single entity. Thus, one of the main issues to be resolved is about the possibility of using a unique time reference. Indeed, if the clocks of « probe machines » involved in delay measurement, for example, are out of synchronisation, the result will have no value and will not be exploitable. There are many network clock synchronisation protocols, the most known and used is NTP [MIL 96], but performance evaluations of this protocol have shown that even if it succeeded in giving fairly satisfactory results for local networks, it is totally inadequate on expanded long distance networks such as Internet. In the absence of clock synchronisation protocols on long distance networks better than NTP, it is a whole piece of research on large scale and geographically expanded distributed systems that must be invented. Today, many engineers and researchers in networks use GPS which, in addition of giving the current position on the earth surface, also carry the ticks of reference atomic clocks.

Another problem encountered in distributed systems – and even more pronounced in the Internet because of its size – is related to the location of the measurements that are not necessarily made at the point where they will be exploited. These measurements should therefore be repatriated to their place of exploitation. This repatriation generates at least two other problems:

– the first is related to the amount of traffic that this can generate on the network. Indeed, even if sometimes, for example, for a delay measurement, only a scalar value is transmitted by the measuring probes and therefore does not generate an excess of significant traffic, it is possible that the information to be repatriated is a complete trace of packets which can represent several megabytes, or even gigabytes. In such a case, measurement information signalling is far from being transparent to the network and its traffic;

– the second is when we wish to exploit the measurements performed in real time. Obviously, the repatriation of data to their place of exploitation takes more or less time depending on the case, but, however, enough time to question the time validity of this measurement. In a high-speed network – which is the case for Internet today – and whose traffic varies a lot and very quickly, it is legitimate to wonder if the value of the received measurement is still valid. It is in fact a problem encountered in distributed networks and systems from the fact that control or signalling messages in the network use the same transmission medium as the communications data of users… In the case of Internet, signalling messages and user data are carried in packets that follow the same links, go through the same routers, etc. By making an analogy with road traffic, it would mean for example that to determine the time that it would take on a Saturday to go from Paris to Marseille, a vehicle is sent on the same journey on the night before the day of departure. Obviously, the value obtained on Friday night may not be very different from what it will be on Saturday. But, if the Saturday in question is the first Saturday of August with the infamous coming and going of summer holidays, it is clear that the value obtained on Friday night is not connected with what the vehicle will take on Saturday. The sudden increase on car traffic between Paris and Marseille on the first Saturday of August, with the creation of bottlenecks (similar to network congestions), makes the measurement performed on the night before is worthless. However, in network, there is no parallel medium to convey urgent control or signalling messages, which will make the performances of network management mechanisms or measurement signalling sensitive to the data traffic existing on the network. This is also a major problem that must be fixed to be able to implement and deploy measurement systems in the Internet, and effectively exploit the results.

7.2.3. *Measurements, estimations and analysis*

This problem leads naturally to the approach concerning the dependence of measurement techniques or their « calibration » to traffic nature. Thus, it seems clear that we would not be able at all to use the same solutions to measure 2 Mbps ADSL traffic and traffic of an operator at the heart of a network at tens of Gbps. In the same way, the granularity of observation may not be the same in both cases. This granularity will also need to be adapted to other traffic characteristics: for example, if the average size of transmitted traffic flows varies on different links, it will certainly be wise to adapt the granularity of observation to the size of the carried traffic flows.

Finally, and this is certainly the main problem of networks measurement and metrology, it is almost impossible to work on the multitude of mechanisms and protocols that is included in the Internet architecture called TCP/IP. This general problem discussed in the introduction of the chapter makes it such that measurement alone does not provide information on the network and its behaviour. Seeing only the effects of mechanisms and protocols is indeed far from being satisfactory. Measurement should therefore be followed by a phase during which metrology methods and techniques, for example based on traffic characterisation or analysis and/or QoS measurements will help explain the causes of the observed phenomena. To do so, we will show below different techniques that have been designed and developed in recent years to measure the basic physical parameters in networks or collect traffic traces.

7.3. Measurement techniques

7.3.1. *Active measurements*

The principle of active measurements is to generate traffic in the network to be studied and observe the effects of components and protocols – networks and transport – on the traffic: loss rate, delay, RTT[114], etc. This first approach has the advantage of a user-oriented positioning. Active measurements remain the only for a user to measure the parameters of the service which he may receive. One of the

---

1. This chapter is taken essentially from the article of the same authors: « Metrology techniques and tools for Internet and its traffic », *Techniques of the engineer*, Treatise « Mesurements and Control », reference R 1090, 2006, with the friendly authorisation of the editors.

2. RTT (*Round Trip Time*): it is the time taken by a data packet transmitted from a source to a receiver to come back to the source. It is a key performance parameter in a computer network.

major drawbacks for the network with active measurements is the disturbance introduced by traffic measurement, which can change the network state and, thus, distort the measurement. Indeed, the measurement result provides information on the state of the network carrying both normal data of users and of signalling of the network control plan, but also, all the « probes packets ». However, we would wish to have information that matches normal traffic only, without « probes packets», which inevitably have an impact on networks performances. We should therefore be able to estimate the impact of these packets on networks performances or be sure that they would have a minimal impact, if possible almost zero. It is this last proposition, *a priori* simpler, which requires the most research efforts. We talk about non-intrusive measurement traffic. Thus, many current works address this problem by trying to find profiles of measurement traffic that minimise the effects of additional traffic on network state [ALM 99a, ALM 99b, ALM 99c, PAX 98].

Another question raised by such measurements is related to the convergence speed of measurements to a result whose reliability is good. Indeed, in order to asses some parameters, it is sometimes necessary to implement a whole complex process to approximate the solution. For example, to measure the rate available on a path between a source and a destination, some tools transmit « probes packets » rates which increase at each attempt until losses appear, these losses being treated as congestion phenomena. The value of the generated rate is thus the value retuned by the measurement tool as available rate on the path. However, the process can be lengthy and, where the transported traffic on the path is very variable, the result is unreliable. Sometimes even, it is possible that the tool does not converge to a result. Converging rapidly is thus essential to know exactly the traffic changes on the path.

Finally, accuracy is crucial. If the measurements have significant confidence intervals, the results are of no interest for researches, engineers or network administrators. For example, in the case of the measurement tool of rate available on a link, the accuracy is closely related to the convergence speed of the measurement process and to the step increase of the « probe traffic » at each iteration change. But, in other cases, like delay measurement, accuracy can only be related to the quality of the time synchronisation between the source and destination of the « probes packets ».

7.3.2. *Passive measurements*

Passive measurements projects appeared much more recently than those of active measurements because they require systems for capture or analysis of traffic in transit relatively advanced and recently developed – even if some simple software, but with limited capacity, already existed, as TSTAT, NTOP, LIBCAP, TCPdump, TCPtrace, etc. They nevertheless highlighted that supervision tools, working with a a passive approach, were likely to solve many problems about the design, engineering, operation and management of Internet networks.

More powerful hardware are in fact the basis of the current boom in the field of passive metrology, and have particularly paved the way to microscopic passive metrology (defined below). The principle of passive measurements is to watch the traffic and study its properties in one or several points of the network. The advantage of passive measurements is that they are not intrusive at all and do not change anything to the network state when using dedicated hardware solutions (for example on the basis of *Dag* maps [DAG 01]). They allow very advanced analysis. However, it is very difficult to determine the service that may be offered to a customer based on information obtained by passive metrology. It is better, in this case, to use active techniques.

Passive metrology system scan also be differentiated according to the method of analysis of traces. Thus, the system can do an analysis online or offline. For an online analysis, the whole analysis must be undertaken in the time frame corresponding to the passage of the packet in the measurement probe. Such a « real time » approach performs analysis over very long periods and produces significant statistics. However, the maximum complexity of these analyses remains very limited because of the low computation time allowed (even lower than the capacity of the network is important). In contrast, an offline analysis requires the probe to save a traffic record for future examination. Such an approach requires thus considerable resources, which represents a limitation for traces of very long duration. However an offline analysis allows extremely complete and difficult analyses, capable of studying non trivial traffic properties. Moreover, as the traces are saved, it is possible to perform several analyses on the traces, then correlate the different results obtained on the trace or obtained on different traces, for a better understanding of the complex network mechanisms.

The main constraint for installing measurement probes is from the fact that the network whose traffic we wish to analyse is almost always an operational network (with the exception of a few experimentation networks in research laboratories), and that, despite the presence of probes, this network should continue to function without any service degradation:

– the primary need for the measurement system to set up is thus a total transparency for the network and its traffic. This means to be non intrusive, this equipment should not cause failures, transmission errors and not introduce delays in order not to change the traffic profile and network performances;

– the second need in the selection of passive measurement probes is its accuracy and validity of traces it will produce. Thus, it is essential not to « miss » packets transiting on a network, and to have precise information on the transition of these packets, especially in time. The system should therefore be well designed and provide a precise clock that does not deviate;

– the third need focuses on the possibility to correlate the events of several traces, for example to track a packet at several points of the network, or to analyse in an overlapped way the passage of packets and their acknowledgements, etc. In order to finely analyse such events occurring at geographically distant points and at distinct instants but slightly apart in time, it is necessary to have a common and universal time basis for all the probes.

Finally, there are also other needs which may have secondary importance when designing a measurement tool, but which should not be overlooked. The first concerns the problem of metrology of a complete network and the repatriation of data collected from different metrology points. It is important to find a way to bring these data back to the analysis machines efficiently and without affecting the network and its load. The second concerns the nature of the information to be collected, and especially the size of recordings made on each packet. Indeed, collecting the all the data of the packet, with thus all the « applicative » information, is *a priori* contrary to the rules laid down by the CNIL on the right to privacy. We should therefore consider carefully the data to be collected on each packet in relation to the analysis we wish to do.

7.4. Characteristics of Internet traffic

Let us begin this section by describing the evolution of traffic distribution per application measured in the Internet in recent years. Figure 7.1 illustrates this distribution measured in May 2000 on the Sprint network. The large proportion shown by http traffic (more than 75% of Internet traffic) is remarkable. We also note that the main standard applications are represented: web, secure web, email, ftp or news. However, new emerging applications (at that time) are present: traffic flow of multimedia streaming (such as MediaPlayer ot RealAudio) or network distributed games (like Quake). Nevertheless, the most important feature of this traffic remains its elasticity and its QoS time constraints which are not important (for the vast majority of its applications).

Three months later, in August 2000, the distribution was almost the same with the exception of a new application, which has become, in a few weeks, one of the major applications of the Internet. It was Napster, decked of the term « killer application » because, within three months, it represented between 20 and 30% of the trafic. This type of P2P application has known, over time, a very important success among users, thus representing an increasingly important part of the traffic in the Internet. Although Napster had some problems with the American justice, it has paved the way to a whole set of P2P applications such as Gnutella, E-donkey, Morpheus and others. Indeed, three years later, P2P traffic has steadily increased and, at present, on some links of the Renater network, it can represent the same proportion as http traffic (see Figure 7.2). Of course, Napster has been replaced by Kazaa or E-donkey. Such an increase of P2P traffic has fatally affected the characteristics of global traffic in particular because of the nature the exchanged files (mostly music or films), which are comparatively much longer than the web traffic flow, the majority traffic few years ago.

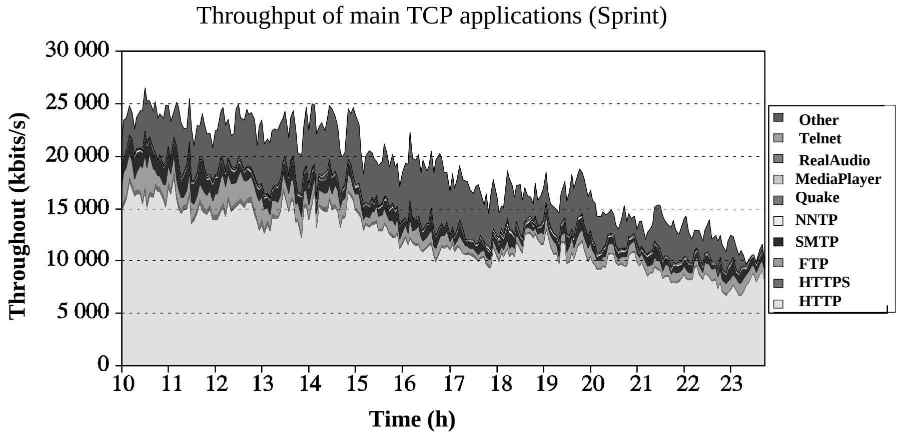

Throughput of main TCP applications (Sprint)

Figure 7.1. *Distribution of traffic on the Sprint network (May 2000)*[115]

Throughputs of main Internet applications (*ReNatER* network)

---

3. The applications are classified in the same order in the legend and on the graph.

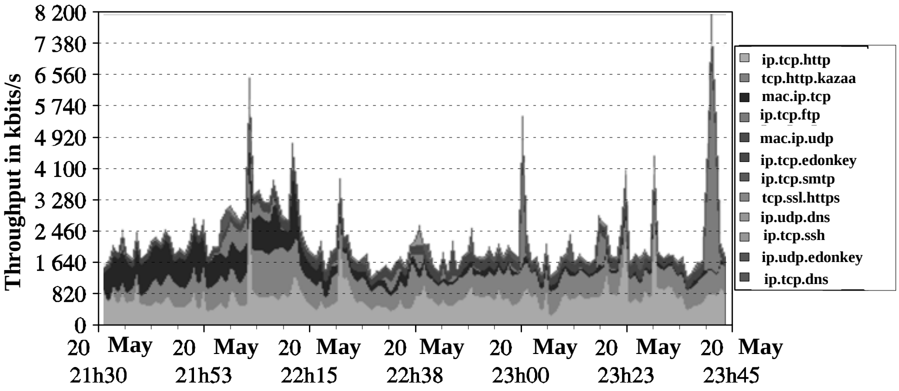


Figure 7.2. *Traffic distribution on the* ReNatER *network (May 2003)*[116]

In fact, the increase in P2P traffic coupled with the presence of traditional traffic induces the following characteristics:

– Internet traffic is always composed of thousands of small flows called mouse (mainly due to web traffic as well as P2P control traffic);

– a number of increasing elephant flows.

So much so that distribution of the size of traffic flows in the Internet changes significantly. This phenomenon has been analysed since the early 2000s and the results are shown in Figure 7.3. The exponential distribution (the one with the least cumbersome tail) serves as a reference because it is close to the Poisson model[117]. We can see on this figure that the proportion of very long flows has increased dramatically since 2000. If in 2000, this distribution was not very far from an exponential, it is no longer the case at present. On the contrary, it has a heavy tail and very far from the exponential distribution. This is a major achievement of the research in metrology in recent years; until then, we considered that the Poisson model or slightly more complex Markovian models – those who had been validated for telephone networks traffic – applied to Internet traffic. It is clear today that this is not true and that Internet traffic is incredibly more complex and diverse.

---

4. The applications are classified in reverse order in the legend and on the graph.

5. This model is used as reference in most cases when it comes to perform simulations or evaluations of network performance. It was considered – still very recently – as a model correctly representing Internet traffic.

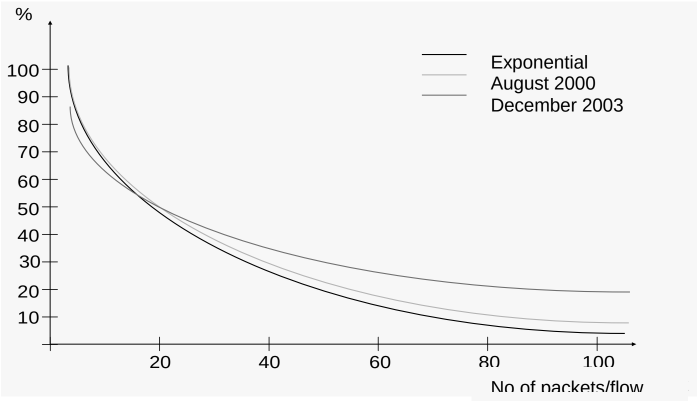


Figure 7.3. *Evolution of the distribution of traffic flow sizes in the Internet between 2000 and 2003*

Returning to this major change in Internet traffic, which consists of an increasing number of long traffic flows, Figure 7.4 illustrates the changes that we can observe. To do so, it compares the current Internet traffic with a traffic following the Poisson model. These two traffic are observed at different granularities (0,01, 0,1 and 1 second), and it is easy to notice that Internet traffic does not smooth out as fast as the Poisson traffic. The analysis showed that this result is totally due to elephants present in the Internet traffic. Indeed, elephant transmission creates in the traffic the arrival of a great wave of data that has the particularity to last a relatively longer time (more than one second[118]). That is why we observe this difference between the two traffic types: elephant transmission induces persistent oscillations in the current traffic.

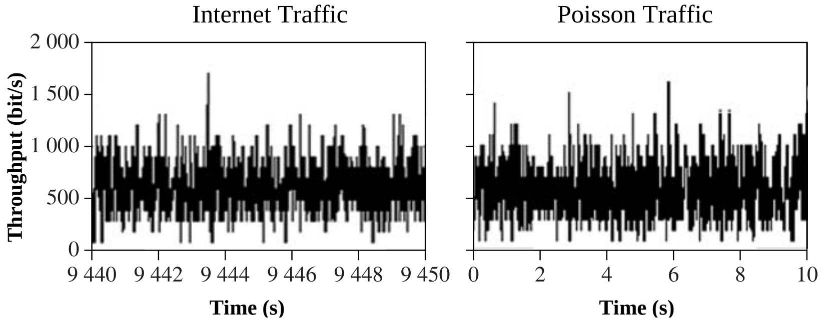
In addition, the TCP connections used to transmit larger elephant flows last longer and the dependence existing between the packets of the same connection thus spreads on longer time scales. This phenomenon is traditionally called LRD. It is attributed several causes, the main one due to congestion control mechanisms of TCP (the dominant protocol of Internet). Among all TCP mechanisms, it is clear that the one based on a closed loop control introduces dependence in the short term, since the acknowledgements result from the arrival of a packet and that the emission of all the following the packets of the connection is determined by this acknowledgement. In the same way, the two TCP mechanisms (slow-start and congestion avoidance) introduce dependence on a longer term between packets of different congestion windows. Thus, by generalising these observations, it is clear

---

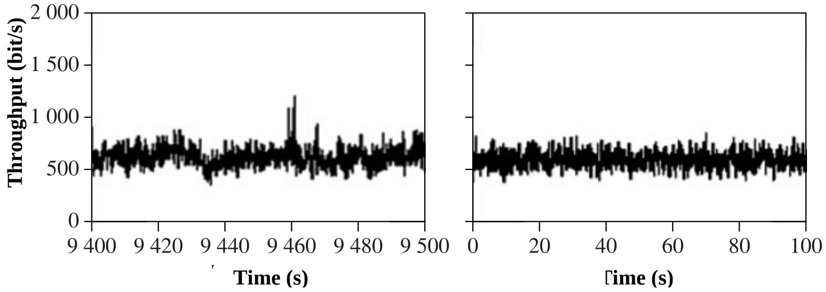
6. Web traffic flows are traditionally transmitted in less than one second in current Internet.

that the TCP packets of a connection are dependent on each other. In addition, the increase in Internet links capacity, by allowing the transmission of increasingly long flows, increases the LRD phenomenon. This is why we observe on Figure 7.4 the persistence of an oscillatory behaviour in Internet traffic, even with a significant granularity of observation (1 second).

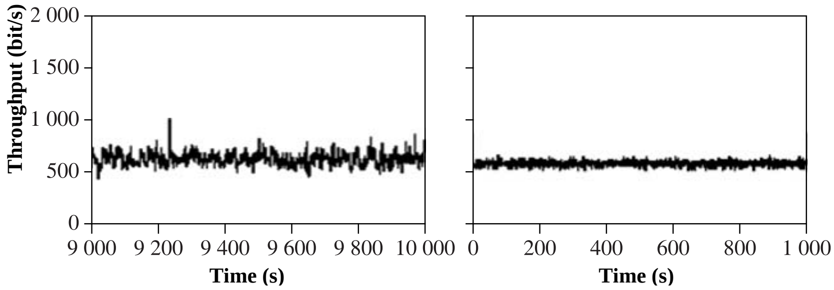
Since the dependence phenomenon of TCP spreads in the traffic through the flows (that is the TCP connections) [VER 00], the increase in flow sizes induces an increase in the extent of the dependence, which can reach very significant scales. Thus, an oscillation at time *t* then induces other oscillations at other moments that can be potentially very far from *t*. It is clear, moreover, that the elephants, because of their important lifetime in the network and the large capacity of the latter (most of the time the links being oversized), have time to reach large values for their congestion control window. Thus, a loss induces for the traffic flow that experiences it a significant decrease, followed by a significant increase in its throughput. The increase in size of flows favours therefore oscillations with high amplitude and a phenomenon of long-term dependency. Of course, the oscillations are very harmful for an optimal use of the global resources of the network, given that the capacity released by a flow undergoing loss can not be immediately used by another (due to the slow-start phase in particular). This means there is a waste of resources and implies a decrease of the global QoS of the network. In fact, the more the traffic oscillates, the less significant the performances are [PAR 97].



a) Combined throughput: 0,01s



b) Combined throughput: 0,1s



c) Combined throughput: 1s

Figure 7.4. *Comparison between the observed oscillations in Internet traffic and Poisson traffic* [119]

7.5. Conclusion

With the advent of the Internet network, and the problems posed today by its phenomenal growth, metrology becomes the cornerstone of many activities, both

---

7. This study, conducted in the context of *Metropolis*, is based on traffic on an ADSL plate of *France Telecom*.

in network research and its engineering. In particular, Internet and its mechanisms, the change in demands of users, etc. give to the structure of the IP traffic great complexity, one of whose manifestations is for example the phenomenon of long-term dependence or self-similarity. Its knowledge requires thus to be sought (if not controlled) more or less continuously in terms of its upgradeability. The need for traffic measurements and analyses is justified especially by the following needs:

– assess the demand of user traffic, particularly in the context of the supply to come of differentiated service classes;

– size network resources: processing capacity of routers; transmission rate of links; size of « buffers » at interfaces. Adapt the operational management of these resources to the time scalability of traffic demand. This includes in particular all aspects that are related to routing in network;

– control the QoS offered by the network: rate of packet loss, transfer delay and jitter of end-to-end packets for applications with « real-time » constraints, useful transport throughput of data traffic flows;

– test the adequacy of performance models developed using analytical calculations or simulations, both in terms of the validation of considered hypotheses and the relevance of results.

Finally, metrology is used for many functions in networks, a non-exhaustive list of which might start with:

– measurement, implementation and management of QoS,

– management and administration of networks,

– pricing and billing,

– congestion control and admission control,

– routing,

– security (intrusion detection, network protection through tools such as firewalls, etc.),

– etc.

It follows from all this that metrology is not an easy science. It thus concerns only computer networks professionals: engineers, researchers, operators, administrators, etc. Although this science is very young in the field of networks, its capabilities have made it an obligatory passage for all these professionals and, ransom of success, metrology has even started being taught in tertiary education (universities, schools of engineers). However, much remains to be done so that the metrology tool is able to answer all the questions asked about Internet and its traffic. One of the major difficulties comes from its multidisciplinary side, requiring technical expertise in the field of computers, of course, but also statistics or signal processing for the analysis of the captured traffic traces. Now, metrology is also of interest for human sciences, but also the media, advertising, polling organisations, etc.

7.6. Bibliography

[ALM 99a] ALMES G., KALIDINDI S., ZEKAUSKAS M., A One-way Delay Metric for IPPM, RFC 2 679, September 1999.

[ALM 99b] G. ALMES, S. KALIDINDI, M. ZEKAUSKAS, A One-way Packet Loss Metric for IPPM, RFC 2 680, September 1999.

[ALM 99c] ALMES G., KALIDINDI S., ZEKAUSKAS M., A Round-trip Delay Metric for IPPM, RFC 2 681, September 1999.

[DAG 01] DAG 4 SONET NETWORK INTERFACE, http://dag.cs.waikato.ac.nz/

[MIL 96] MILLS D., Simple Network Time Protocol (SNTP) Version 4 for IPv4, IPv6 and OSI, Request for Comments 2030, October 1996.

[PAR 97] PARK K., KIM G., CROVELLA M., « On the Effect of Traffic Self-similarity on Network Performance », *SPIE International Conference on Performance and Control of Network Systems*, Dallas, United States, November 1997.

[PAX 98] PAXSON V., ALMES G., MAHDAVI J., MATHIS M., Framework for IP Performance Metrics, RFC 2 330, May 1998.

[PLA 02] PLATON, *Allegory of the cave*, The Republic, Book VII, Fourth century B.C., Flammarion, Paris, 2002.

[VER 00] VERES A., KENESI Z., MOLNAR S., VATTAY G., « On the propagation of long-range dependence in the Internet », *SIGCOMM'2000*, Stockholm, September 2000.