On the Use of Passive Network Measurements for Modeling the Internet

Klaus Mochalski and Klaus Irmscher

University of Leipzig, Computer Science Department
{mochalski|irmscher}@informatik.uni-leipzig.de

Abstract. Many important discoveries in Internet research have been made on the basis of traces collected by passive measurement systems. However, the various difficulties involved in capturing a network traffic trace have resulted in a poor availability of high-quality traces. Furthermore, with the rapidly increasing data rates of today's Internet links, capturing a trace long enough to yield meaningful analysis results becomes more and more difficult. This paper emphasises the usefulness of passive measurements and outlines the methodology used to conduct them. We present three trace data sets to describe our analyses, among them a new trace captured at the University of Leipzig. We also show how packet traces can be used to support and evaluate Internet simulations.

1 Introduction

Passive measurements are a powerful tool for modeling Internet traffic. They produce a trace of the actual traffic on the measured link at a certain time. Such a trace can be seen as a snapshot of an Internet link. The achieved accuracy of a trace depends on the quality of a measurement. What exactly accuracy means and how to achieve it is discussed in this paper. Once a high-quality trace has been captured, it can be used for further analyses. This includes for instance verification of models of certain traffic classes, support and verification of simulations, and also very practical analyses like detection of denial of service attacks or misbehaving hosts and applications. Classical examples of discoveries made by trace analysis are, for instance, [10] and [11]. These works have greatly influenced subsequent research. Passive measurements will be especially useful to support current efforts toward Internet QoS support. They enable precise analyses of packet burst and delay behaviour—two important issues linked to QoS.

Capturing packet traces becomes more and more challenging with the increasing bandwidths of today's Internet. Different scalability issues have to be addressed including capturing speed during the actual measurement session, storage capabilities for longer traces, and processing power for offline analysis of trace data. In this paper we present the methodology we used to capture and analyse trace data sets at various locations—among them the Auckland-IV and VI [5] and the Leipzig-I [6] data sets.

The availability of specialised hardware to support Internet packet capturing turned out to be of crucial importance. Dag measurement cards [7] developed by the University of Waikato, Hamilton, New Zealand, have been used to collect these packet header traces. They provide high precision timestamps with clock synchronisation.

The rest of this paper is organised as follows: Section 2 gives a brief comparison of passive and active measurement techniques. Section 3 discusses various aspects of passive measurement methodology which have been applied to the Auckland and Leipzig measurements described in Sect. 4. Section 5 presents various analyses based on these data sets. We show how traces can be used to improve and evaluate network simulations. Section 6 discusses the results and concludes the paper.

2 Active vs. Passive Network Measurements

There are two forms of network measurements—active and passive.

Active measurements create and inject artificial packets into the network under observation. Later, these packets are intercepted and metrics based on their behaviour are calculated. The idea behind this technique is to use a well-defined sample to draw conclusions about the overall behaviour of a certain part of the network.

Passive measurements capture packets transmitted by applications running on network-attached devices over a network link. Usually, the arrival of each packet is earmarked with a timestamp. Storing all captured packets along with their timestamps in a trace file provides an accurate representation of network traffic.

The achievable measurement accuracy strongly depends on the accuracy of the timestamps supplied by the measurement system. While this applies to both active and passive measurements, it is a more important issue for the passive case. Since legacy hardware does not provide high-precision clocks and clock synchronisation to support accurate timestamps, specialised hardware has to be deployed in many cases. Section 3 provides further details about timestamping of network packets.

Active and passive measurements both have their specific advantages and disadvantages making them suitable for different purposes. One of the major drawbacks of active measurements is the potential interference of injected packets with normal network traffic. Depending on the network load and the amount of data transmitted by an active measurement platform, this could not only lead to a distortion of the very effects to be measured but also actually create an overload situation. This can pose a serious limitation as network measurements are especially interesting during periods of high load.

The passive approach does not have such a limitation. There is no interference of the measurement with network traffic. However, the above mentioned accuracy of passive measurements comes at a price. Each and every packet needs to be captured to gain a complete picture of a link's traffic behaviour. This imposes a serious scalability problem to passive measurements. With the Internet link capacities growing faster than other computer technologies such as CPU, memory, disk, and tape performance, it is just a matter of time until full network packet traces—even for short periods of time—become all but unfeasible. A thorough study of bandwidth requirements for passive measurements can be found in [1]. In this respect, active measurements scale much better because they often work with a data sample of negligible size in comparison to the overall traffic on a measured link. Section 3 discusses possible improvements of the poor scalability properties of passive measurements.

Safety and privacy are very important issues of any network measurement. Neither network operation nor user privacy should be adversely affected. The first aspect applies to active measurements as discussed in the last two paragraphs whereas user privacy is more of a concern for passive measurements. Active measurements generate their own data. Only these data are used for analyses, and user data remain untouched.

The situation is somewhat different for passive measurements. User data are intentionally captured and often stored for analysis purposes. This is one of the major sources of difficulties involved in conducting a passive measurement in an operational network. Most organisations running a network usually have strong objections to this kind of data collection. These privacy concerns have to be addressed by dropping any unnecessary data (e.g. any packet payload) and by anonymising IP addresses to prevent end user identification from the trace data. Section 3 describes how this is achieved for the Auckland and Leipzig data sets.

3 Passive Measurement Methodology

Timestamping is one of the central functions of a passive measurement system. To understand its importance, one has to be aware of the timescale at which events happen on a network link. With increasing link rates, requirements to timestamping precision rise proportionally. Table 1 gives an idea of timescales of different link technologies.

Table 1. Packet serialisation times for selected packet sizes and network technologies[2]

Packet size (bytes)	Link	Serialisation (in ns)
64	10BaseT	51,200
512	10BaseT	410,000
1500	10BaseT	1,200,000
64	100BaseTX	5,200
64	100BaseTX	520
53	OC3c	2,720
53	OC12c	675
53	OC48c	168

The simplest and cheapest approach to timestamping network packets is to use a standard PC equipped with a network interface card and to rely on the PC's clock facility along with a synchronisation package like NTP (Network Time Protocol). The disadvantage of this approach is the relatively low accuracy. NTP provides an accuracy within some milliseconds. Table 1 shows that this is by no means sufficient for today's high speed links.

The Dag network measurement cards developed by the University of Waikato [7] provide a solution to this problem. These cards are able to do packet capturing at full link rate for a wide variety of link layer technologies including Ethernet, ATM and Packet over Sonet (PoS). They are equipped with a high-quality clock which can be synchronised to UTC by feeding a PPS (pulse per second) signal into the cards using GPS (Global Positioning System) or CDMA (Code Division Multiple Access). The achievable clock accuracy is in the range of some 100 ns and is the basis for the timestamping process. More information about precision timestamping can be found in [2].

There are several approaches beside these two just mentioned which range in their timestamp accuracy somewhere in between. RIPE NCC, for instance, use GPS-conditioned PC clocks for their measurements [8]. Another approach is the utilisation of CPU cycles to improve the accuracy of the PC clock [9].

In Sect. 2, the scalability problem of passive measurements has been posed. There are several ways to improve the scalability properties of passive measurements. The most obvious one is to collect only a subset of the actual traffic on a link. This approach has two dimensions: first, instead of capturing complete packets, it is often sufficient to store only the packet header. The header contains most of the important information (e.g. IP addresses, port numbers, protocol information) needed for analyses. Second, filters can be applied to capture only certain classes–or flows–of packets.

All data sets presented in this paper are such packet header traces. They are stored in native Dag format [7] consisting of 64-byte records with a 64-bit timestamp at the beginning of each record and the rest used for the packet header. Depending on the link technology, each record stores at least the first 40 bytes of an IP packet thus including 20 bytes of IP header and 20 bytes of a possibly present TCP header (assuming no IP options).

A useful side effect of dropping packet payload is improved privacy. No user data are collected. However, there is still concern for the IP addresses which could be used to track an individual user's behaviour. This issue is addressed by an anonymisation scheme applied to a trace shortly after it has been captured. It basically performs an irreversible mapping of real IP addresses into an anonymous address range. The disadvantage of this anonymisation procedure is that it circumvents any analysis involving routing properties as these are lost with the change of IP addresses.

Packet filtering can be done based on different criteria. These criteria should be flexible enough to allow for different levels of flow granularity (e.g. TCP connections, application or host conversations). The specific filter properties depend on the intended analysis. However, once such a filtering decision has been made for a capturing session, any following analyses are limited to the collected sample. This may prove disadvantageous in case some unexpected results occur in the course of an analysis. The choice with the highest flexibility always is a full packet trace.

The challenge is to reduce the amount of collected data without compromising the purpose of a measurement. It is of central importance—both to the usefulness and the collected data volume of a measurement—where it is conducted. Network capacities in today's Internet are being deliberately over-provisioned. The Internet is a mesh of interconnected autonomous systems. The operator of each autonomous systems will design network resources in a way that there be virtually no data loss even during periods of extreme network utilisation. Problems are much more likely to arise at peering points between such autonomous systems. These interconnections are carefully rate-limited and their capacity is regulated by service level agreements and peering arrangements. This commercially motivated bandwidth reduction renders these points most likely for congestion-caused high packet delays and losses to occur. Our analysis presented in Sect. 5 will provide proof that this is indeed true.

These considerations make such points especially interesting for passive measurements. Moving a measurement point away from the Internet's core toward its edges means reducing the bandwidth that has to be handled. Furthermore, a peering point with a bandwidth reduction from OC48 speed to OC12 will likely display a similar behaviour as a point with a reduction from OC12 to OC3. With the scalability concerns of passive measurements in mind, it appears sensible to conduct a measurement at the lower speed link. It should then be possible to extrapolate the results to higher speed links.

In addition to the lossy data reduction described above, standard compression utilities (e.g. gzip) can be used to further reduce the amount of collected data. Tests with various traces have shown that gzip delivers the best trade-off between cost and efficiency. Depending on the data rate of the measured link and the performance of the measurement system, compression can either be done in real-time during the capturing session or afterwards. This has to be individually assessed for each measurement session.

Despite all these data reduction efforts, the amount of data collected in network traces is still huge. This calls for a means of automatic trace data processing. We have developed a set of tools [7] as part of the Dag software package to support a partially automated analysis of large data sets. This tool set has been used throughout the analyses of the data sets presented here.

4 The Auckland and Leipzig Data Sets

We have captured three different data sets at two locations—the University of Auckland, New Zealand (Fig. 1), and the University of Leipzig, Germany (Fig. 2). In both cases, the university's Internet access link has been instrumented on either side of the main Internet router, though only Auckland-VI is a multi-point measurement with taps on both sides being used to capture data. Auckland-IV and Leipzig-I were captured at the outer side of the router.

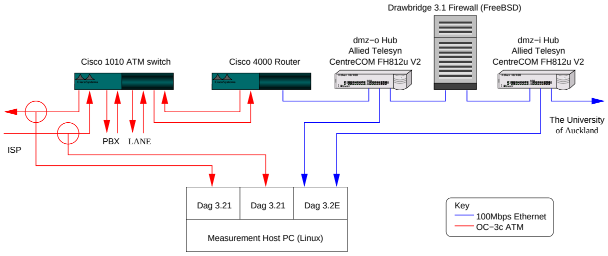


Fig. 1. Measurement configuration for the Auckland-IV and VI data sets

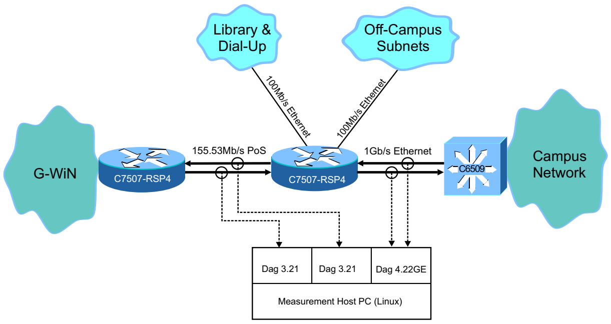


Fig. 2. Measurement configuration for the Leipzig-I data set

The ISP connection in Auckland is an OC3c ATM link with a 4.048 Mbit/s PVC. The University of Leipzig is connected to the German Research Network (G-WiN) via an OC3c PoS (Packet over Sonet) link running at 155.53 Mbit/s. Both campus networks are hooked to the router via an Ethernet connection running at 100 Mbit/s in Auckland and at 1 Gbit/s in Leipzig.

Instrumenting a university Internet access link means to exactly follow the principles for selecting an interesting monitoring point as proposed in Sect. 3. The disparity of link capacities on either side of the router induces heavy queuing of outgoing traffic with its adverse impact on the traffic behaviour. The analysis presented in Sect. 5 provides proof that this assumption is indeed true. It will be shown that the queuing delay added by the router renders any realtime application running over such a link without additional QoS functionality all but impossible.

Table 2 lists some statistical information about the traces. The Auckland data sets are available to the public [5].

Table 2. Trace data statistics

	Duration	Number of	File size (GB)	
Data set	(days)	Packets (million)	Uncompressed	Compressed
Auckland-IV	45	3157	188	65
Auckland-VI	$4\frac{1}{2}$	$3 * 312$	50	18
Leipzig-I	$5\frac{1}{2}$	3800	226	113

The Auckland measurement system was equipped with three Dag measurement cards—two Dag 3.2 with OC3c/OC12c interface, one for each direction of the OC3c link, and one Dag 3.2E, a dual-port 10/100 Mbit/s Ethernet card, which was connected to the DMZ hubs (see Fig. 1). In Leipzig we used two Dag 3.21 with OC3c/OC12c interface to tap into the ISP link. Additionally, the measurement system is equipped with a Dag 4.22 with 1000Base-SX interface which hasn't been used for any measurements yet. However, we plan to capture a multi-point trace similar to Auckland-VI in Leipzig.

Clock synchronisation was provided by a Trimble Palisade GPS antenna in Auckland and a Trimble Acutime 2000 in Leipzig. The antennas were connected to the Dag cards ensuring that timestamping precision is within 600 ns to UTC at any time during the measurement session. Log files of the timekeeping process have been recorded for verification. In addition to these log files, the analyses did not reveal any timestamp inconsistencies (e.g. outliers) further reassuring the confidence in this data set.

5 Analysing Passive Traces

Analysing trace data is a time-consuming process. Table 2 gives an idea about what amounts of data have to be handled. We have developed a set of tools [7] written in C which take traces in Dag format as input and generate various statistics as text or graphics. Among them are programs to generate statistics about packet rate, bandwidth, application mix, and flow behaviour.

We have used these tools to generate comparable graphics about all three trace data sets. An examples is given in Fig. 3. A complete set of graphs can be found at the web sites [5] and [6].

We have also developed an algorithm to calculate one-way packet delays for multi-point measurements. It generates a 32-bit CRC over the whole packet header excluding IP TTL and checksum fields which will change when a packet passes a router. The CRC values of the trace captured at one point are stored in a hash table and matched against those of the trace taken at the second point. Multiple matches of a single packet are considered ambiguous and discarded from the analysis. Unmatched packets may occur for several reasons:

- cross traffic from untapped router interfaces
- packets sent to or originating from one of the router's interface IP addresses

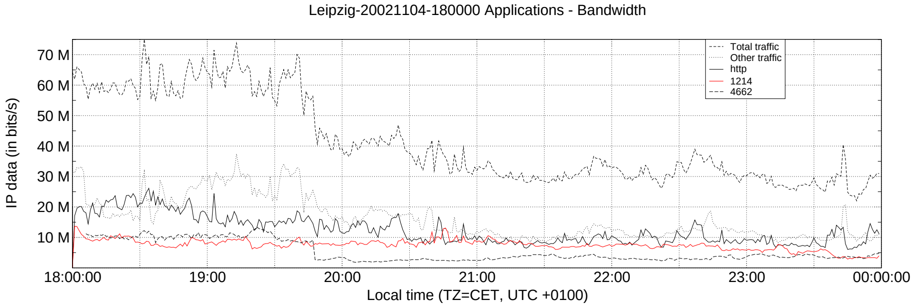


Fig. 3. Average bandwidth of incoming and outgoing traffic for a 6-hour period of Leipzig-I

- packets discarded due to filtering rules (e.g. firewall as for Auckland-VI, see Fig. 1)
- true packet loss due to link overload

Mismatches of the first two categories can easily be filtered out if the subnet addresses connected to the router and its interface addresses are known. However, this is only possible if the traces have not been anonymised as described in Sect. 3. Recognising missing packets from the second category is possible only if the firewall rules are known.

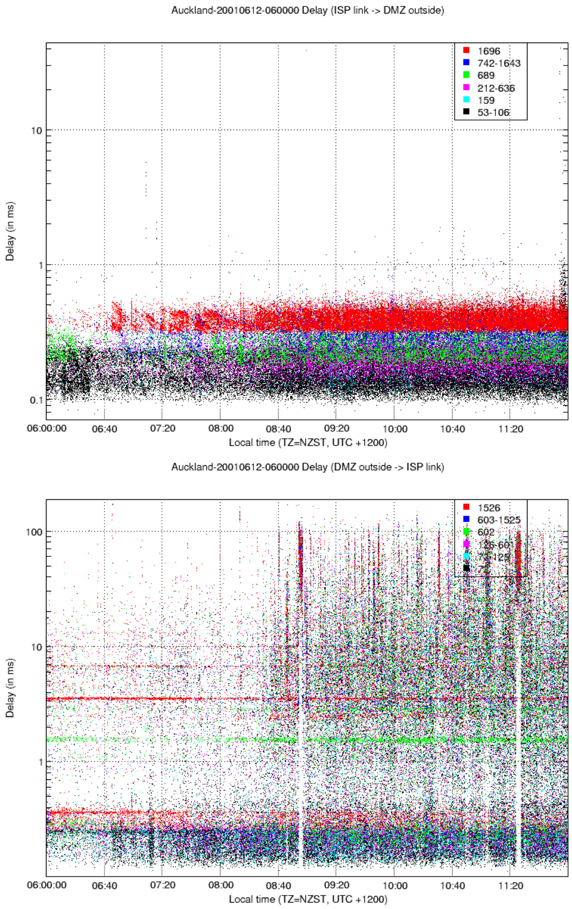
5.1 Analysing a Router's Behaviour

A router is basically a packet forwarding device with queues buffering packets. Watching all incoming and outgoing packets provides a complete view of a router's behaviour. That was done for the Auckland-VI data set. There are only two basic things which can happen to a packet passing a router—it gets forwarded with a certain delay, or it gets lost due to buffer congestion. A measurement configuration as deployed in Auckland and Leipzig (Fig. 1 and Fig. 2) can detect both.

The following paragraphs will outline the results of our delay analysis of the Auckland-VI data set to demonstrate the usefulness of such measurements. For a more complete discussion of this subject refer to [4].

During our analyses of the Auckland-VI data set we found scatter plots of the delay value distribution to be adequate visualisations of the router's behaviour. We applied a colouring scheme to reveal packet size dependencies of certain prevalent delay values. Five colour groups which represent packets of a certain size have been defined. The groups are different for ATM and Ethernet because the graphs show link layer size (i.e. multiple of 53 bytes for ATM, or Ethernet frame size). For ATM the groups are 53-106, 689, 1696, 159, 212-636, and 742-1643. For Ethernet they are 72, 602, 1526, 73-125, 126-601, 603-1525. Note that the first three groups respectively have been chosen to represent the typical IP packet sizes of 40, 576, and 1500 bytes. An analysis of packet size distribution of the trace revealed a strong prevalence of these packet sizes.

Figure 4 shows such scatter plots of the delay distribution with various clearly separated bands of different colours (i.e. different packet sizes). All bands for outbound packets and the bottom bands for inbound packets (below 0.4 ms) can be attributed to different serialisation times depending on packet size. The upper bands for outgoing traffic represent bursts of packets of the same size. Outgoing packet bursts arrive with up to 100 Mbit/s but can only be transmitted with 4.048 Mbit/s and thus have to be buffered. This only happens for outgoing packets which is the reason for the huge disparity of delay values for both direction. 99% of all inbound packets experience a delay of less than 0.53 ms, whereas for outbound traffic it is 82 ms—a difference of two orders of magnitude.


Auckland-20010612-060000 Delay (ISP link -> DMZ outside)

Auckland-20010612-060000 Delay (DMZ outside -> ISP link)

Fig. 4. Delay distribution of incoming (top) and outgoing packets at the University of Auckland on Tuesday, June 12, 6am-12pm

Apart from the horizontal banding, the plot of outbound packet delays reveals a number or periods during which the delay distribution seems distorted compared to general behaviour. Figure 4 shows two major incidents—one at about 8:55am, the second at about 11:25am—during which virtually no delay value is below about 30 ms. Considering that during normal operation a large proportion of delay values lies below 0.4 ms, this has to be taken seriously.

Our analysis revealed a plateau of SMTP traffic perfectly coinciding with the high-delay incidents. Interestingly, similarly high spikes of HTTP traffic do not have such a devastating impact on the overall delay. Checking the IP addresses has shown that a single host is responsible for this malicious behaviour. More about these anomalies can be found in [4].

5.2 Support and Evaluation of Network Simulations

The ultimate simulation is one based on real data. The main obstacle of many network simulation is that they use especially generated traffic based on a—possibly wrong—model of real traffic. Several influencing papers (e.g. [12]) have described the difficulties involved in conducting proper simulations of Internet traffic behaviour. Trace data collected by passive measurement system can provide a solution for at least a subset of these problems.

In [4] we presented a simple router simulation based on a leaky bucket model. This simulation takes a Dag trace file as its input and calculates a delay value for each packet based on buffer occupancy, which in turn depends on the burst behaviour of the input traffic.

The delay distribution calculated by such a simulation is visualised in Fig. 5. Its most obvious feature is the striking visual similarity to the real trace data (Fig. 4) which promises a good match of simulation and real world. The simulation is able to reproduce both, the horizontal banding due to packet queuing as well as the vertical patterns representing high-delay incidents as described in Sect. 5.1.

By using a trace file of the Auckland-VI data set, we can not only run a simulation based on real data but also evaluate its quality by checking simulated against observed delay values. This reveals some weaknesses of this simulation which are not obvious from the looks of Fig. 5. Only about 58% of all simulated values have an error of less than 20%. If we set the error threshold to 10%, the proportion of valid simulation values amounts to a mere 45%. Furthermore, the simulation predicts about 2,300 lost packets during a certain 6-hour interval. This value gets invalidated by the observed number of about 50,000 lost packets.

Although the router simulation largely fails, the comparison against the observed data reveals some useful hints at possible sources of its weaknesses which can be used to improve the simulation accuracy. By using real data for the simulation we can rule out an adverse influence by poor models used to generate test traffic. This allows to focus on the actual simulation instead of having to deal with proper test traffic generation. We plan to conduct a similar simulation based on traffic captured in Leipzig.

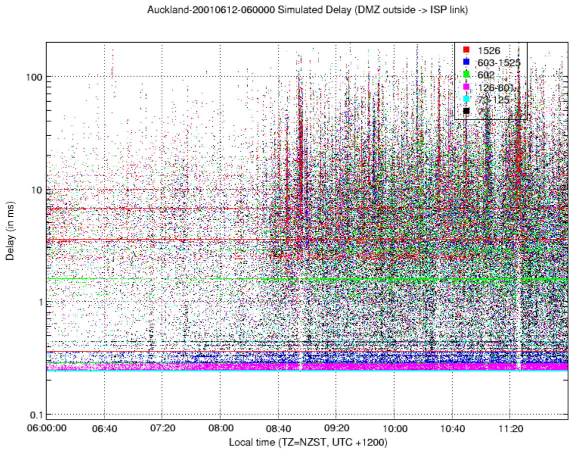


Fig. 5. Simulated delay distribution

5.3 Realtime Capabilities – An Outlook

All analyses described so far have one common disadvantage: they consume a significant amount of time in the order of some minutes up to hours for larger traces depending on the available computing resources. This fact renders such approaches unusable for network management purposes were immediate availability of information about network status is indispensable.

On the other hand, a passive measurement system equipped with Dag cards provides an undistorted view of the network traffic that few other solutions can match. That raises the question about possible realtime capabilities of such a system. We are currently planning on conducting a study about the kind and granularity of information that can be provided using such a passive measurement system.

6 Conclusion

This paper advocates the use of passive measurements as a modeling tool for today's still rapidly growing Internet. The specific advantages of passive over active measurements have been described. The key feature of passive measurements is the ability to provide a high-fidelity snapshot of real-world Internet behaviour. The importance of a thoughtful deployment of measurement systems has been discussed. Access and peering points have been identified as the most interesting monitoring points. Extending the measurement methodology by monitoring several points which maintain a traffic relation enables the analysis of packet delay and loss—key metrics when it comes to Internet quality of service. Such multi-point measurements can be conducted by instrumenting a single device (e.g. a router) or an end-to-end path.

We have presented several trace data sets to prove the usefulness of passive measurements. The Auckland-VI data set instrumented a router and a firewall at an Internet access link. This measurement confirms the presumption that a significant bandwidth reduction implies an adverse impact on the delay and jitter behaviour of real network traffic. Delay values ranging from below 0.2 ms up to about 200 ms have been observed. These results are of special interest if one seeks to implement QoS features to support real-time applications. The measurement methodology can easily be adopted to support QoS-specific analyses by taking into account different traffic classes.

We are planning to conduct subsequent measurements at the University of Leipzig. There will be a router instrumentation similar to Auckland-VI. This will put the results of Auckland-VI analyses into perspective and improve their value by allowing or ruling out generalisations. Furthermore, it is planned to conduct an end-to-end path measurement to focus on the user's perception of service quality. We will also explore the realtime capabilities of passive measurements.

References

1. Jörg Micheel, Hans-Werner Braun and Ian Graham: Storage and bandwidth requirements for passive Internet header traces, Workshop on Network-Related Data Management, Santa Barbara, California, USA, May 25th 2001
2. Jörg Micheel, Ian Graham and Stephen Donnelly: Precision Timestamping of Network Packets, Proceedings of the ACM SIGCOMM Internet Measurement Workshop, San Francisco, California, USA, November 1st/2nd 2001
3. Stephen Donnelly: High Precision Timing in Passive Measurements of Data Networks, PhD thesis, University of Waikato, June 12, 2002
4. Klaus Mochalski, Jörg Micheel and Stephen Donnelly: Packet Delay and Loss at the Auckland Internet Access Path, PAM2002 Passive and Active Measurement Workshop, Fort Collins, Colorado, USA, March, 25-26th, 2002
5. Waikato Internet Traffic Storage: http://wand.cs.waikato.ac.nz/wand/wits/
6. Leipzig Trace Archive: http://rnvs.informatik.uni-leipzig.de/ traces/
7. Web sites of Dag development at the University of Waikato and Endace Measurement Systems: http://dag.cs.waikato.ac.nz, http://www.endace.com
8. Maximo Alves, Luigi Corsello, Daniel Karrenberg, Cagdas Ögüt, Mark Santcroos, Reinhard Sojka, Henk Uijterwaal and René Wilhelm New: Measurements with the RIPE NCC Test Traffic Measurements Setup, PAM2002 Passive and Active Measurement Workshop, Fort Collins, Colorado, USA, March, 25-26th, 2002
9. Attila Pasztor, Darryl Veitch: PC Based Precision Timing Without GPS, ACM CLIOMETRICS 2002, June 15-19, 2002, Marina Del Rey, California, USA
10. Will E. Leland, Murad S. Taqqu, Walter Willinger, and Daniel V. Wilson: On the Self-Similar Nature of Ethernet Traffic, Proceedings ACM SIGCOMM 93, 13-17 September 1993
11. V. Paxson and S. Floyd: Wide-Area Traffic: The Failure of Poisson Modeling, IEEE/ACM Transactions on Networking, Vol. 3 No. 3, pp. 226-244, June 1995
12. S. Floyd and V. Paxson: Difficulties in Simulating the Internet, IEEE/ACM Transactions on Networking, Vol.9, No.4, pp. 392-403, August, 2001.