

The Auckland data set: an access link observed

Jörg Micheel, Ian Graham; The University of Waikato, Hamilton, New Zealand
joerg@cs.waikato.ac.nz, ian@cs.waikato.ac.nz,

Nevil Brownlee, The University of Auckland, New Zealand
n.brownlee@auckland.ac.nz

Abstract

In 1999 the WAND network research group of the University of Waikato set up a bi-directional measurement system on the OC3c ATM access link that connects the University of Auckland to the public Internet. The system uses Dag cards, synchronised by a GPS time receiver. A total volume of 118 Gigabytes of compressed IP header traces (more than 2 billion IP headers) have now been taken and archived, and we describe various research projects that have used these data.

1. Introduction

Internet measurements have been taken as a research effort since the existence of the Internet. The purpose of early measurements was for debugging and development of Internet protocols and applications. Examples for active measurements include the `ping` and `traceroute` tools; for passive measurements the `tcpdump` network analysis tool [`Tcpdump`] is commonly used.

Since the commercialisation of the Internet, and the advent of the World Wide Web the study of Internet traffic patterns and the behaviour of applications has become a true research area of its own. For example, the efforts at Lawrence Berkeley National Laboratory by Van Jacobsen, Vern Paxson *et al* have lead to a better understanding of the effects of different TCP implementations on the global Internet [Paxson97]. Archives on captured traces are also kept [ITA], but the volume is sparse and contributions are made on the basis of availability.

Since 1995 the Measurement and Analysis Team at the National Laboratory for Applied Network Research [MOAT] has been taking regular measurements of Internet traffic at high-speed interconnection points of the vBNS network (and more recently the Internet2/Abilene). These measurements are targeted at studies of Internet load profiles (applications), routing stability and other information that can be derived from Internet packet header traces. Traces are taken as samples of 90 seconds eight times per day. The archive maintained at NLANR MOAT allows for historical studies of Internet behaviour for a period of more than 5 years.

When looking for good sources of Internet traffic to study models of statistical behaviour our research group quickly discovered that the support for accurate time was either poor or non-existent. What was missing was information on relative time of packet inter-arrivals at a single link and packet arrivals at one point in the network relative to another point in the network. A feature of the first generation of passive Internet monitors [OC3MON] was the use of off-the-shelf network interface cards with no support for accurate, high-resolution packet timestamping. As a result, timestamp precision and accuracy were poor and frequent timestamp wraps would create ambiguities. In general, those errors cannot be fixed by post-processing.

For many of the current monitors on bi-directional fibre links (ATM and POS) timestamps for the inbound direction captured with one card bear almost no relationship to the timestamps gathered for the opposite direction with a second card. This effect is due to clock drift and is particularly a problem for long-term studies, where the gain in time by one crystal can produce offsets in the order of milliseconds, even seconds, within a single day.

Classic methods for clock synchronisation across networked computer systems — NTP — are too crude to be useful for the task, partly because their utilisation involves communication via the network and thus suffers from the very effects we are trying to measure: network latency, network delay jitter, asymmetric routing, route instability. The polling interval and the resulting clock adjustments of NTP also severely affect the medium term stability of timekeeping. It was to meet the requirements of accurate time measurement in a low-cost system that the Dag project was initiated in the mid 90's at the University of Waikato, New Zealand.

2. Dag passive network measurement cards

At the end of 1998 the availability of the Dag2 PCI network measurement card made it possible to accurately study network behaviour on a larger scale. The Dag2 has now been replaced by the Dag3 series, which includes cards for OC12c and OC3c ATM/POS interfaces and 10/100BaseT Ethernet. The new Dag4 has an OC48c ATM/POS interface and a Gigabit Ethernet card is in development. [Dag] [Graham98]

A major feature on the advanced Dag cards is a conditioned clock, which uses an external synchronising pulse as a time reference. For historical reasons the clock system is known as the Dag Universal Clock Kit, or DUCK [DUCK]. The DUCK's synchronising pulse may be generated by another Dag card, thus allowing the synchronisation of two or more cards at one site, or may be derived from a GPS time receiver. The present Dag3 series has a clock resolution of 60 nanoseconds, and the arrival time of a cell or packet can be measured with an absolute accuracy of better than one microsecond.

The accuracy of timing has three major advantages for network measurement. Firstly, two network measurement cards attached to the same link can now be synchronised with each other and timestamps can be correlated for long measurement periods (hours, even days). In this paper we refer to these as one-point or single-point measurements. Secondly, with the use of an external GPS time reference, the short and the long-term stability and accuracy of time support can now be guaranteed. Thirdly, because of the accuracy of the GPS time reference, measurements can be taken at geographically distant points in the network and the timestamps be correlated. This allows for time-of-flight (unidirectional delay and jitter) measurements [Donnelly98][Graham98a]. In this paper we refer to these as two-point, three-point, or, in general, as multi-point measurements.

3. The Auckland measurement point

In July 1999 the Waikato Applied Network Dynamics (WAND) research group decided to launch a passive measurement effort on the University of Auckland's Internet access link. The major reason for our interest was that the link is the only connection of the university to the public Internet and so carries the aggregated application traffic of some 30,000 users. The measurements will not be disturbed by dynamic route changes and a complete view of application flows is possible. Yet another advantage is the placement of the monitor at the boundary between a local area and a wide area network. From the measurement point it is possible to observe behaviour of packet delay in both environments (see the study on delay by H. Stele Martin et al below).

Technically, the link is a virtual circuit on a SONET OC3c ATM link, which was the physical layer supported by Dag2 cards at the time. The clock synchronisation is provided by a Trimble Palisade GPS system with ± 100 nanoseconds accuracy to UTC [Trimble], resulting in a packet arrival timestamp precision of better than 1 microsecond to UTC.

Dag cards currently deliver a fixed 64-byte record per each received packet. The record contains a 64-bit timestamp reflecting the arrival of the packet or ATM cell at the measurement point, plus link specific fields, such as ATM cell headers, AAL5 CRC or Ethernet MAC headers. In the case of the Dag2 cards the timestamp is a free running counter clocked at 12.5 MHz. Dag3 cards and beyond deploy the DUCK technology described above. For the ATM cards as deployed at the Auckland measurement point packet headers have to be filtered from the ATM cell stream. With a limited number of onboard processor cycles available per ATM cell inter-arrival time (for OC3c around 2.7 microseconds) the filtering scheme on the Dag2 cards only passes cells with the characteristic LLC/SNAP / EtherType IP payload pattern to the monitoring host. Dag3 cards and beyond use the PT bit in the ATM header to synchronise with the AAL5 framing structure, and discard all cells except the first in each packet.

The trace records contain the first 40 bytes of the IP packet. This is sufficient to capture all the information present in IP/ICMP, IP/TCP and IP/UDP headers in the case where no protocol options are present. This approach is identical to other monitors [OC3MON]. We are not interested in or permitted to collect user data. For UDP packets the first 12 bytes of the user payload is also present in the trace record. Those must be appropriately sanitised in a post-processing stage (see security and privacy considerations below).

For data capture on bi-directional fibre links, such as ATM and POS, a measurement card needs to be installed for each direction. Consequently, the trace taken will consist of two files. Those files must be kept separate in case an analysis step requires information about the direction of a packet flow. For other evaluations only the combined packet stream is of interest. The dagmerge tool as supplied with the dagtools software package [DagSoft] allows for merging of the two directions for a bi-directional trace. The individual packets are ordered by increasing time, timestamp synchronisation of the two Dag cards is thus implicitly assumed.

	Auckland I	Auckland II	Auckland III	Auckland IV
Dates	July 1999	November 1999 – June 2000	August 2000	Scheduled for end of 2000
Link	ATM Virtual Circuit Connection on STM-1c (OC3c equivalent)			
Bandwidth	2 MBits/sec packet peak rate per each direction			
System	Celeron 333A, 128KB cache, Asustek P2B (BX chipset)			
Memory	32MB PC100	32MB PC100	32MB PC100	96MB PC100
Disk	6.3 GB ATA	6.3 GB ATA	50 GB SCSI	50 GB SCSI
Capture cards	Dag 2.1	Dag 2.1	Dag 3.21	Dag 3.21
Synchronisation	None	Palisade GPS	Palisade GPS	Palisade GPS
Timing	12.5 MHz	12.5 MHz soft	16 MHz DUCK	16 MHz DUCK
Archive	Internet	DDS-2 Tape	SCSI disk	SCSI disk

Table 1 – Auckland Dag monitor configuration

4. The Auckland-I data set

A one-week trial at the beginning of July 1999 revealed a number of interesting results. The system was able to keep up with the traffic load. A single day produced approximately 1.5 to 2 Gigabytes of trace data. Uninterrupted 24-hour trace runs are thus possible. A brief study with conventional file compression tools (GNU `gzip`, `bzip2`, UNIX `compress` and `pack`) revealed that the GNU `gzip` program produces the best results. It achieves superior or equal compression ratio, compression speed is fair, and the performance of the `gunzip` decompression tool is superior to any other tools studied. Since it is expected that the captured traces will be studied frequently, our choice of `gzip` was clear. At the time, `gzip` was also found to be memory and CPU hungry. For that reason we did not decide to do on-the-fly (real-time) compression of trace files, but used a post-processing stage. As a result, subsequent traces see a minimum gap of 90 to 120 minutes from one another: the time it takes to compress the last day's trace files. A typical compression ratio using `gzip` level 9 is 55%.

The parameters of the archive are as follows:

Duration	1 week
Frequency	7x24 hours
Volume per day (uncompressed/compressed)	0.8 – 1.8 / 0.4 – 0.8 GByte
Volume total (uncompressed/compressed)	10.8 / 4.5 GByte
IP headers	170 million

Table 2 – Auckland-I data set

Very simple analysis was done with the traces on the data server; we only printed packet and bandwidth statistics for the entire week. The need to develop tools for post-processing and analysis was recognized for the first time.

As a result of the study several changes to the measurement set-up were made.

Packet loss monitoring was implemented with a change to the Dag firmware. Loss can occur as a result of temporary shortage of PCI bus bandwidth due to activity of other devices, such as the disk drive or the second Dag card. Software support for GPS time stamping was added. For every GPS pulse-per-second (PPS) interrupt a timestamp of the free running 64-bit counter was recorded and merged as an artificial trace record into the capture data stream. Post-processing software would pick up those records and produce timestamps compatible with the modern Dag DUCK [DUCK] format. This timestamp records time as two 32-bit fields: the number of seconds since January 1st 1970 (compatible with a UNIX `time_t`) and 32-bit fractions of a second. The resolution of the format is approximately 0.25 nanoseconds, which we consider adequate for future high-bandwidth networks. The actual resolution may differ from card to card, but the format stays the same and thus allows for hard- and software upward compatibility.

A benefit of the artificial trace records was that the health of the PPS timekeeping support could be monitored for the duration of the measurement. Any two PPS timestamps should never be further apart than about 12.5 million ticks. As it turns out, the Palisade GPS antenna does have a firmware bug, which lets it miss a pulse every 30 to 45 minutes. This does not actually affect the timekeeping accuracy and the software was instructed to deal with the situation. Log files of these outages are kept along with the trace files. Trimble has fixed this bug in a recent firmware update.

We also considered loss monitoring for the physical layer. The network measurements might be skewed when the connection of the monitor to the link is impaired. The current architecture of the card does not allow for monitoring of the physical layer interface chip during the capturing session itself. However, SONET/SDH error counters are cumulative and can be latched and read before and after the session. For the time of our measurements we could not find any reason for concern that the connection to the link was impaired.

The test also highlighted the problem of storage and transport of the traces. A one-week trace consumes more disk space than is available on the local machine. Some slack space must be provided for the trace compression process. The main data server of the WAND group is located in Hamilton, about 130 km south of the city of Auckland. During the Auckland-I measurement we had to offload completed trace files during the later days of the week. This resulted in skews of the traffic pattern on the access link as measured by our monitoring system. Another concern was costs. For some of the New Zealand universities there is a charge for Internet traffic on a per-megabyte basis. The costs for transferring the huge trace files via the Internet turned out to be enormous. For the next set of measurements we changed our archival scheme to magnetic tape. A used DDS-2 tape drive was deployed and the data shipped via courier to Hamilton.

In general, the test run encouraged a systematic long-term study at the university, which, after about 5 month of preparatory work, resulted in the Auckland-II data set.

5. The Auckland-II data set

All of the above improvements were made for the Auckland-II data set. Individual trace runs were started whenever possible, and trace length was targeted at 24 hours. The first data set was, in fact, taken for as long as disk space lasted about 38 hours. Due to a bug introduced into the Dag2 firmware, a number of the traces ended prematurely. Since there is not much point in keeping uneven legs for each of the directions, traces have been trimmed in a post processing stage to start and end at exactly the same time. We considered chasing and fixing the bug, but development of the Dag3 cards was close to completion and so we decided to replace the Dag2 cards with Dag3 for the next data set to be captured.

The parameters of the archive are as follows:

Measurement period	7 months
Frequency	irregular intervals
Duration per trace	2.5 – 38.5 hours
Volume per trace (uncompr/compr)	0.2 – 3.5 GByte / 0.1 – 1.6 GByte
Volume total (uncompr/compr)	59 GByte / 26 GByte
IP headers total	985 million
Number of traces	42
Active trace duration total	24 days 3 hours

Table 3 – Auckland-II data set

Some basic data integrity checking was done on the measurement host. Timestamps were verified to be increasing at all times. The bulk of the post processing was deferred until traces had been loaded onto the main data server, coined the Waikato Internet Traffic Storage [WITS].

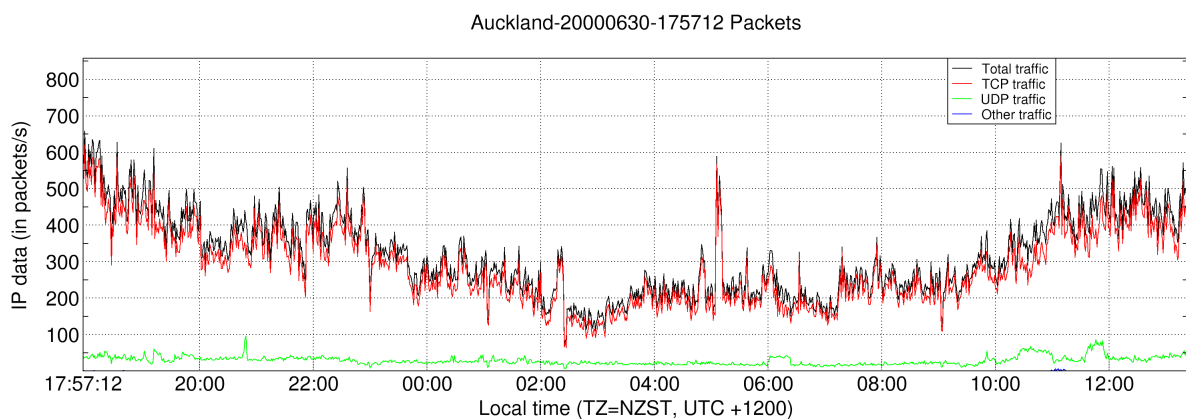
6. Post-processing and publishing of Auckland-II

It is difficult to assure data integrity for long term measurements which result in voluminous data sets. Some inspection *must* be done, since the trace data is going to be used by a larger community of researchers, who rely on the fact that the data truly reflects the actual behaviour of Internet traffic on the link observed. Limited redundancy is available in the trace content which can be used to develop tools for automatic checking. For instance, TCP sessions can be inspected to assure that events happen in the order they are supposed to occur according to protocol specifications.

For our work on the Dag trace data we decided to produce a set of simple analysis tools that demonstrate the use of the trace data to researchers for development of their own tools. In addition, the toolset is targeted to provide simple visualization of the trace content. We call this approach an “illustrated trace archive”. No in-depth analysis is done, instead, the graphs displayed are meant to allow for inspection and selection of trace files for detailed analysis.

We considered using available packages for network trace analysis, such as CoralReef [Coral]. An advantage in using such packages is that development time is low and most of the errors have already been taken care of. The main reasons not to work with CoralReef were that it is extremely resource consuming (CPU and memory) and that it is not meant to deal with data sets of the volume we are generating. As a result, simple pipelining tools and a supporting Makefile system have been written that perform orders of magnitude better and achieve fair results. For example, the three graphs shown below have been generated in about 10 minutes time on our main data server. Passes on the entire Auckland-II data to regenerate the full set of graphs and HTML pages take just a few hours and can be completed during the idle hours of the night.

Below we show the standard graphs currently produced for the Auckland data set: packets per second, bytes per second and new connections per second. Binning of the data is done over a period of one second; the graphs typically display one-minute averages because of the resolution.



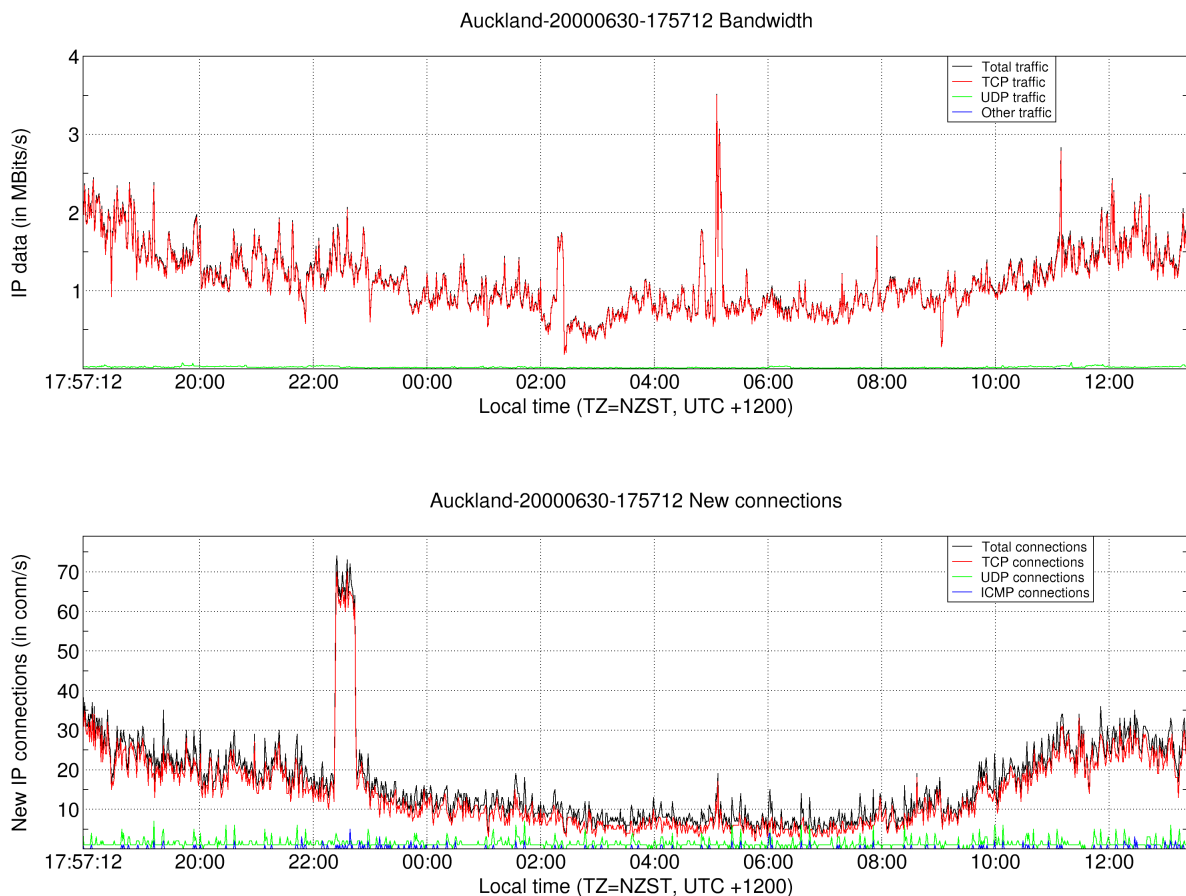


Figure 1 – Standard data set analyses

Once a trace has been selected, detailed inspection is possible. We demonstrate typical analysis work based on the data set displayed above. Most noticeable is the spike in the new connections graph between 22:20 and around 23:00, which is not reflected in any of the graphs displaying packet and byte statistics. We use several tools from the `dagtools` [DagSoft] package in combination. First we select the given time interval from the trace using `dagcut`. This shrinks the amount of data to be processed from 1.4 GB to 65 MB. The `dagsess` tool displays active TCP and UDP connections in terms of their start time, duration, end time, packets and bytes sent in both directions. We use the standard UNIX `sort` command to order the output based on start time. We find a number of single-packet sessions originating at one particular host, which have different destination IP addresses, all located at the University of Auckland. Looking at the same trace file with the `tcpdump` network analysis tool (the `dagbpf` tool provides for conversion) we find that indeed this is an active scan of the university's IP address space for DNS servers, followed by packets with suspicious combinations of TCP flags and header content, leaving little doubt that we are observers of an attack. For network administrators the procedures to be followed from here onwards are very clear, we omit the details.

In a different study we analysed packet size distributions based on the IP header total length field. For the Auckland access link more than 50% of all packets are smaller than 100 bytes. About 70% are smaller or equal to 576 bytes. Ten per cent of all traffic is 1500 byte packets. As disturbing as this sounds; apart from packet counts one also has to consider traffic volume based on the data bytes carried. Here, only 20% of the data is carried in packets smaller than or equal to 576 bytes. An overwhelming 50% of the total data volume is carried in packets of 1500 bytes. These results

vary slightly with diurnal changes, but the overall picture is consistent. We therefore omit the graphs. It might be interesting to compare the data for inbound and outbound packets separately.

The data is then prepared for use by the research public. Since IP addresses can in some circumstances reveal the identity and behaviour of a particular user, these privacy concerns have to be addressed. We post-process all traces by a procedure coined *munging*, which maps real IP addresses into virtual IP addresses selected from network address space 10 (10.X.Y.Z). This process is non-reversible. A downside of address munging is that some of the structure associated with IP addresses can currently not be preserved (see below for future enhancements).

Once sanitized for remaining user payload and IP addresses the trace data can then be shipped via magnetic tape or the Internet to the individual trace users. With the current Internet charging scheme in New Zealand in mind we have chosen to publish a subset of the Auckland-II trace data on the NLANR MOAT server [Kiwitraces].

7. The Auckland-III data set

The major change from Auckland-III onwards is use of the Dag3 network capture cards. Since the Dag3 produces trace records with proper timestamps by default, artificial trace records can no longer be used to assure health of the GPS/PPS support. Instead, additional health monitoring software has been developed, which keeps track of the various parameters of the DUCK. From those statistics, very detailed statements about the accuracy of the timekeeping system can be derived.

With the Dag3 cards available there are no technical restrictions, which limit traces to certain time intervals. Contiguous captures can now be made which last for as long as storage space is available. We felt that the previous sparse trace taking with Auckland-II may have introduced some bias, since the times at which traces started were in fact not arbitrary, but depended on the working hours of the human monitor administrator. For that reason, we have upgraded the Auckland monitor with a 50 GB SCSI disk. We expect that the storage capacity is sufficient to allow for at least one month of uninterrupted capturing. For this project, on-the-fly compression of trace files became a necessity. We therefore had to re-evaluate the performance assessment for the `gzip` compression tool as discussed above. We picked a few of the newer traces at random and did a study on the compression ratio and compression speed for various levels of `gzip` compression. Traces were found to behave identically.

Gzip compression level	Compression ratio	Throughput (Mbytes/sec input data)	Storage space occupied relative to level 1
1	54.8%	1.88	1.00
2	56.2%	1.72	0.97
3	56.7%	1.32	0.96
4	58.4%	1.28	0.92
5	59.4%	1.04	0.90
6	60.2%	0.77	0.88
7	60.6%	0.59	0.87
8	61.0%	0.31	0.86
9	61.0%	0.22	0.86

Table 4 – Gzip compression performance on trace data

The study suggests that there is not much of a difference between the compression that can be gained with level 9 as compared to level 1: the disk space saved is a total of about 14 per cent. However, compression speed (throughput) drops dramatically with increasing compression level. There is a factor of almost 10 in between level 1 and level 9 compression speeds. For our purposes it was sufficient to study whether on-the-fly compression is feasible at all. The Auckland link sees a total of about 0.5 Mbytes/sec peak data in both directions. Theoretically, it is possible that there is an inflation of captured data compared to link load since a single ATM cell occupies 53 bytes whereas a Dag trace records consist of 64 bytes of data. For practical purposes some headroom must be reserved since it is possible that the configured packet peak rates are not accurate. Also, `gzip` compression speed is certainly not uniform for the duration of the trace, but varies based on the time for dictionary lookup. Assuming 3:1 compression between packets and trace records for typical Internet traffic mix we can assume that `gzip` level 1 compression is good for links of up to 10 Mbits/sec and leave enough headroom for short bursts of small packets. Therefore, there is no harm in using `gzip` level 1 to take traces in Auckland. Finally, a tool called `dagsplit` has been developed which combines the functions of separating trace files based on packet time with `gzip` compression, utilizing the `zlib` compression library [Zlib].

During our preparations for Auckland-III another measurement point at the New Zealand Internet Exchange [NZIX] became available. NZIX is to be dismantled by the end of 2000. For that reason, the Auckland-III/NZIX-III data set has become the first two-point measurement in our Internet trace collection. The Auckland monitor was instructed to capture contiguously as planned, whereas at NZIX we filtered packets originating from or destined for Auckland. A post-processing step will filter all the Auckland trace data to the packets supposed to appear at NZIX. At the time of this writing initial analysis of the trace data is in progress.

Duration	3 weeks
Specifics	2-point measurement: NZIX-III
Volume (total)	1.8 Gbytes
IP headers	29 million

Table 5 – Auckland-III data set

As a result of the restrictive rules for capturing at NZIX and the necessary IP address munging of the Auckland data set we will not be able to publish the Auckland-III data set as a single-point trace for the research public, which will make it necessary to schedule another capturing session: see Auckland-IV below.

Before the actual start of the trace we also introduced active measurement [AMP] data into the data stream to be captured by both monitors. We intend to do calibration studies between active and passive measurement techniques with the Auckland-III/NZIX-III data set [AMPNZ].

8. The Auckland-IV data set

As a result of the experience with Auckland-III a memory upgrade from 32 MB to 96 MB has been performed so as to not run short of main memory and impair the performance of the two `dagsplit` (`gzip`) processes running in parallel, which may result in loss of measurement data.

Auckland-IV is scheduled as a contiguous 1-month single-point trace for the end of 2000, which is as soon as the Auckland-III data currently stored at the monitor has been processed.

9. Studies on the Auckland data set

More than a dozen researchers at different organizations have been using the Auckland-II data set for analysis on Internet traffic dynamics. We list a few of them here, which we find representative for the wide scope of different analysis and research that can be done with the data.

9.1. H Stele Martin et al – Analysis of Internet Delay Times, PAM2000 workshop, Hamilton

For this study [Martin00] time synchronisation between the two measurement cards as well as stable long-term clock support on the measurement system were required.

In their work the authors study the round-trip times for traffic between the measurement site and a collection of different hosts, some on the local network, some within New Zealand, some overseas. The focus of the study is on Web servers. An attempt to break down delay times into 3 different components is made: physical network latency, variable network latency and server processing delay.

The analysis work done highlights the importance for efficient algorithms and library support to handle the voluminous data sets with certain speed and a given time frame. The authors wish that data sets would include more information about network topology, which could be achieved by combining active and passive network measurement approaches, for instance doing active traceroute's during passive trace taking.

9.2. Sarah K Joyce: Traffic on the Internet – A study of Internet games traffic –Honours project report

The aim of this project [Joyce00] was to produce a model to simulate Internet games traffic. The accuracy of the simulation depends on the accuracy of the measurement data. For this study it was also important that both directions of the UDP session could be observed (see above: properties of the Auckland University access link configuration). The length of the trace files (several hours) was crucial to gain an unbiased view of entire sessions. For comparison, the 90-second snapshot traces as captured by NLANR MOAT would not support such a study.

Internet games typically use UDP as the underlying transport protocol. Default port numbers are used to identify particular Internet games. For instance, UDP port 7777 identifies a server for Unreal Tournament; Quake World servers use UDP port 27500. This way traffic sent by the server and the client can also be distinguished from one another. The Auckland-II data set was pre-filtered for packets with these characteristics. A second step would identify individual sessions based on the 5-tuple: IP protocol UDP / IP source address / IP destination address / UDP source port / UDP destination port. Game sessions were found to typically last for about 20 minutes; few lasted for more than 90 minutes.

For analysis, plots of different game sessions where produced. For the study, packet size vs time was of particular interest, however inter-arrival times where also investigated.

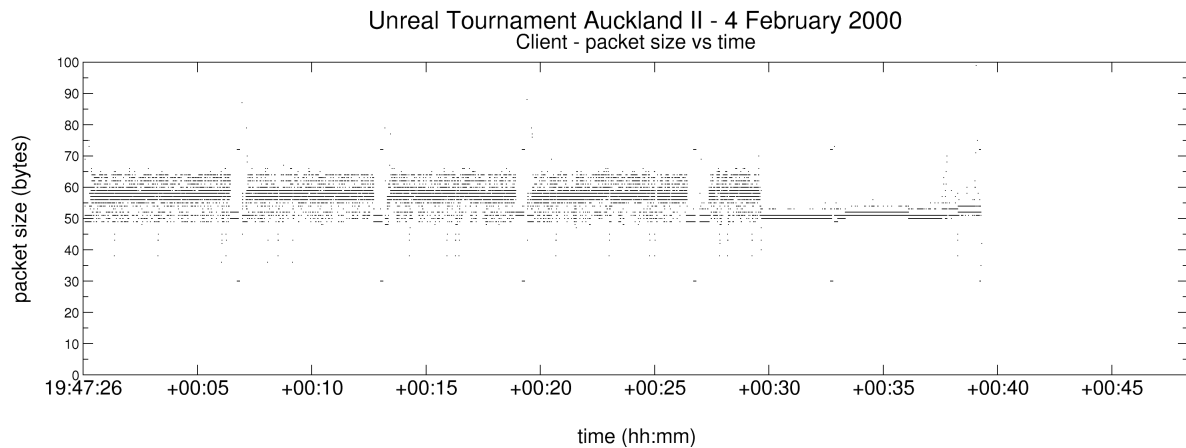


Figure 2 – Internet games traffic – client side

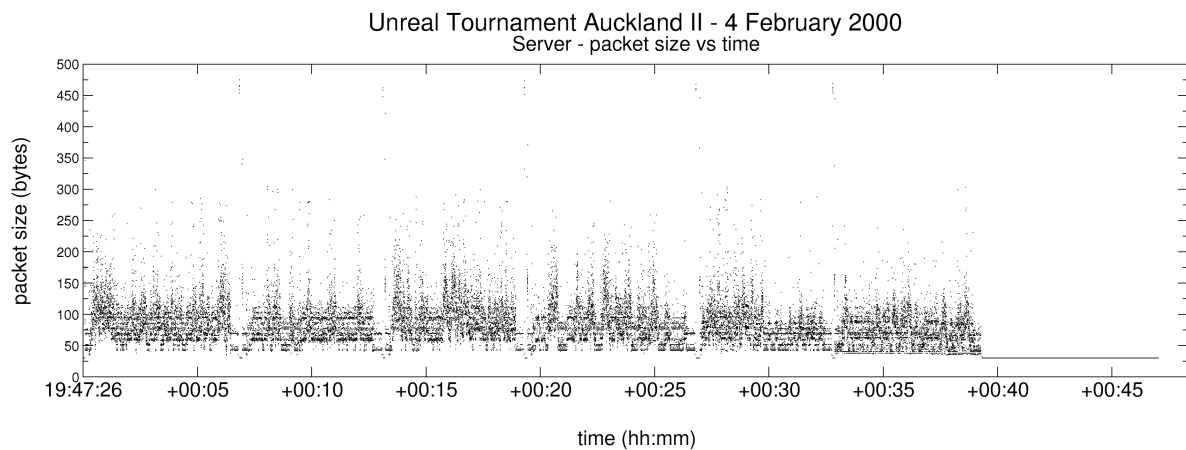


Figure 3 – Internet games traffic – server side

Graphs courtesy Sarah K Joyce, WAND research group, University of Waikato.

The server graph reveals a bug in the Unreal Tournament server implementation, the server continues to send packets for about 8 minutes after the client has disconnected. This bug has since been fixed in the server. In comparison, the client sends fairly small packets. Traffic originating from the server varies a lot more.

The game sessions as extracted from the raw traces were fed into a network simulator along with models for HTTP (WWW) traffic [Pearson99]. The simulation revealed that UDP traffic pushes TCP traffic out of the way. UDP traffic is considered aggressive to network friendly applications deploying adaptive congestion control.

9.3. Darryl Veitch et al – Studies of Long Range Dependencies and Self-Similarity in Internet traffic patterns using Wavelet analysis [Veitch]

The plots in figure 4 show the number of TCP connections active on the Auckland link within 10 ms intervals over a three-hour period. They were produced using a single pass through the data set, a technique which could be used for on-line analysis.

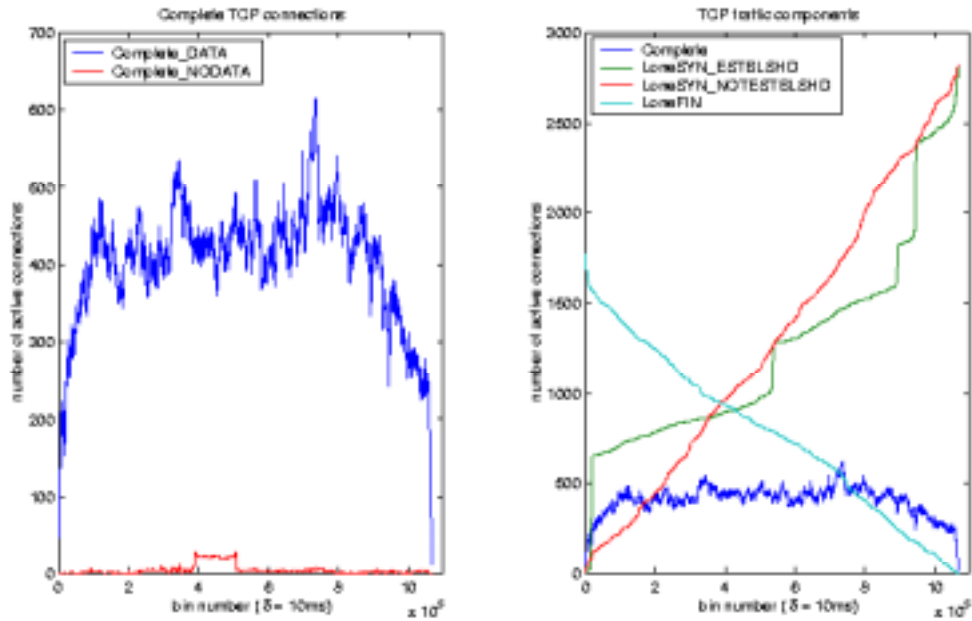


Figure 4 – TCP connection analysis

The left plot of figure 4 shows the number of active 'complete' TCP connections. Note the 'edge' effects – early in the trace there are many flows which have already started, near its end there are many flows, which we cannot see terminating. Connections, which don't transmit any data, are plotted separately; there can be a non-negligible number of them.

The right plot of figure 4 shows the number of active TCP connections, which begin (SYN was seen) but don't end (FIN wasn't seen). These are subdivided according to whether connection was successfully established or not, giving insight into the amount of traffic coming from unsuccessful versus broken connections.

The 'edge' effects shown in these plots will be less significant with longer analysis periods. The Auckland traces lend themselves to this, since they provide real traffic data for periods of days.

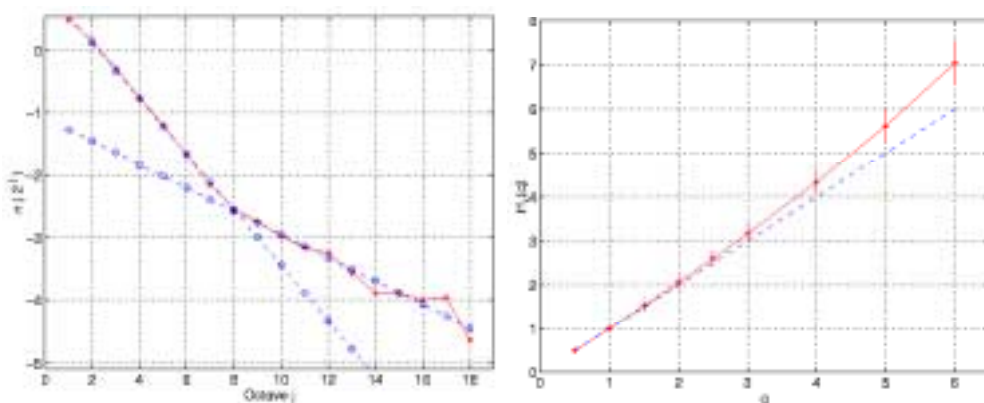


Figure 5 – Hurst parameter as a function of q

Scaling, or “fractal” properties of traffic are now well documented. In the Auckland-II (and Auckland-I) data set both long range dependence – a long memory property involving a single scaling exponent, and multifractal behaviour – a description of high variability at small scales involving an entire spectrum of scaling exponents, have been found. Ongoing work involves the

more general approach of “indefinitely divisible cascades” (IDC). These models can be written in terms of the wavelet coefficients $d(j,k)$ of the data: $E |D(j,k)|^q = c_q \exp(n(a)H(q))$ where the separability between scale $a=2^j$ and moment order q is a direct consequence of the defining assumptions of the model, which is hereby characterized by the functions $n(a)$ and $H(q)$. Multifractal modelling corresponds to the special case of $n(a) = \ln(a)$, which reduces a general evolution along scales to the special case of power laws. In figure 5 these two functions are given for a time series corresponding to the number of new arrivals of (complete) TCP connections per 10-millisecond bin, extracted from an Auckland-II data set of 5 hours. We observe that $H(q)$ (left plot) departs from a trivial linear behaviour, indicating non-trivial scaling behaviour, and giving insights into the structure of the cascades, that is the cascading relationship from one scale to the next, underlying the IDC model. In figure 5 we see also that $n(a)$ is not $\ln(a)$ but is however approximately “piecewise log” (note linear fits). The departure from the simple $n(a) = \ln(a)$ carries statistical information on the way the cascades evolve with scale, and explains why multifractal modelling does not apply at all scales uniformly, whereas an IDC model does, integrating the observed scaling behaviour at both small and large scales. Ongoing work involves explaining the origin of such behaviour, both statistically and in terms of TCP dynamics.

Text and graphs courtesy Darryl Veitch and Lidong Huang of EMUlab, Department of Electronics and Electrical Engineering, University of Melbourne, Australia, and Patrice Abry, Ecole Normale Supérieure de Lyon, Laboratoire de Physique, Lyon, France.

9.4. Vinay Ribero et al - Multifractal Cross-Traffic Estimation, 13th ITC seminar

In their paper [Ribero00] Ribero et al develop a new model-based technique for estimation of cross-traffic for and end-to-end path through the Internet. The presentation at the ITC seminar [RiberoT00] showed the usage of the Auckland-II data set as used for the verification of their simulations.

9.5. Guoqiang Mao - Loss performance analysis for heterogeneous ON-OFF sources with application to connection admission control

This paper is to appear [Mao00].

9.6. Ilze Ziedins - On the output process from a finite buffer with long range dependent input

This paper is to appear [Ziedins00].

10. Conclusions

Auckland data sets are currently the most precise, publicly available, long term Internet access link traces. Contiguous traces allow for unbiased studies of Internet application flows (TCP and UDP connections). The traces captured have been used by more than a dozen researchers working on a broad variety of topics on Internet traffic dynamics. The Auckland data sets are currently single-point measurements, however, publication of a first two-point measurement is in progress. We strongly believe that the research community can benefit from a set of well-documented high-precision Internet traces. We also think that researchers should start sharing analysis results for the same data sets with each other, to provide better insights into the nature of Internet traffic. We have plans to create an Internet trace user community.

11. Future

We currently see much room for improvements on simple post-processing for traces and we intend to enhance and enrich the `dagtools` in this direction and to make more graphs per data set available on the data set web pages. We also seek cooperation with other researchers for this work. For multi-point measurements it becomes even more important to preserve the relationship of IP addresses to CIDR blocks or Autonomous Systems. We are looking into ways to preserve such IP address information via structured munging without compromising security and privacy.

A number of improvements are directly or indirectly related to the Dag network capture cards. For some studies it is necessary to have access to the full header information. For example, many of today's TCP implementations use high-bandwidth or transaction extensions, which result in TCP options that are not currently being captured by the cards. The current size of the trace record "64 Bytes" is a major limitation in a number of ways: host PCI and memory bandwidth, disk bandwidth, network transport capacity, disk and tape storage. We intend to make the compression scheme more efficient.

We see a tremendous potential for GPS-synchronised multi-point measurements. We intent to produce different series of trace files providing insight into the behaviour of the Internet, such as router performance instrumentations, campus studies, backbone studies, transatlantic and transpacific studies.

Acknowledgements

The New Zealand Public Good and Science Fund (PGSF) funded part of this work. Much collaboration exists with US research institutions and funding agencies and it is impossible to list all of them here. We would like to acknowledge the long lasting and fruitful work with our colleagues of NLANR MOAT and CAIDA at SDSC/UCSD. For much of the work we relied on the support of everyone in the WAND research group.

References

- [AMP] NLANR MOAT Active Measurement Program: <http://watt.nlanr.net>.
- [AMPNZ] Measurements of the NZ AMP monitor at Waikato for Auckland: <http://erg.cs.waikato.ac.nz/active/cgi-bin/weekly.cgi?amp-kiwi/hosts/data/130.216.1.239>.
- [Coral] Cooperative Association for Internet Data Analysis (CAIDA): CoralReef. <http://www.caida.org/tools/measurement/coralreef/>.
- [Dag] <http://dag.cs.waikato.ac.nz/>.
- [DagSoft] Dag Software Web Page: <http://dag.cs.waikato.ac.nz/dag/dag-soft.html>.
- [Donnelly98] Stephen Donnelly: Internet Time of Flight measurements, University of Waikato, 1998. <http://atm.cs.waikato.ac.nz/wand/delay/>.
- [DUCK] Dag synchronization and time stamping: http://dag.cs.waikato.ac.nz/dag/docs/dagduck_v2.1.pdf

- [Graham98] Ian Graham, Murray Pearson, Jed Martens, Stephen Donnelly, John Cleary: A remote ATM network monitoring system, University of Waikato, Hamilton, New Zealand, 1998. <http://wand.cs.waikato.ac.nz/wand/publications/>.
- [Graham98a] Ian D. Graham, Stephen Donnelly, Stele Martin, Jed Martens and John Cleary: Non-intrusive and Accurate Measurement of Unidirectional Delay and Delay Variation on the Internet. Proceedings of the INET'98 Conference, July 1998. http://www.isoc.org/inet98/proceedings/6g/6g_2.htm
- [ITA] Internet Traffic Archive, <http://www.acm.org/sigcomm/ITA/>.
- [Joyce00] Sarah K Joyce: Traffic on the Internet – A study of Internet Games Traffic. BCMS 420 Honours project report, University of Waikato, Hamilton, New Zealand, 2000. <http://wand.cs.waikato.ac.nz/wand/publications/>.
- [Kiwitraces] NLANR MOAT: Auckland-II data set. <http://moat.nlanr.net/Traces/Kiwitraces/>.
- [Mao00] Guoqiang Mao, Daryous Habibi: Loss performance analysis for heterogeneous ON-OFF sources with application to connection admission control, submitted paper for IEEE/ACM transactions on networking, network traffic analysis, October 2000. http://www-soem.ecu.edu.au/~gmao/Loss_performance_analysis.ps.gz.
- [Martin00] H. Stele Martin, Anthony J McGregor, John G. Cleary: Analysis of Internet Delay Times, Proceedings of the First Passive and Active Measurement Workshop, PAM2000, April 3-4 2000, Hamilton, New Zealand, p.141ff. http://www.cs.waikato.ac.nz/pam2000/pdf_papers/P033.pdf.
- [MOAT] NLANR MOAT, San Diego Supercomputing Center, University of California at San Diego, <http://moat.nlanr.net/>.
- [NZIX] New Zealand Internet Exchange, a major neutral peering point. See <http://wand.cs.waikato.ac.nz/wand/wits/nzix/2/>.
- [OC3MON] Joel Apisdorf, K Claffy, Kevin Thompson, Rick Wilder: OC3MON: flexible, affordable, high performance statistics collection, NLANR MOAT and MCI/vBNS, 13 Sept 1996. <http://www.nlanr.net/NA/Oc3mon/>.
- [Paxson97] Vern Paxson: Measurements and Analysis of End-to-End Internet Dynamics, PhD dissertation, Lawrence Berkeley National Laboratory, University of California at Berkeley, California, USA, 1997. <ftp://ftp.ee.lbl.gov/papers/vp-thesis/dis.ps.gz>
- [Pearson99] Murray Pearson, Anthony McGregor: A simulation study of network architectures to support HTTP Traffic on Symmetric high-bandwidth*delay circuits. Proceedings of the Asia Pacific Web Conference 1999. <http://wand.cs.waikato.ac.nz/wand/publications>.

- [Ribero00] Vinay Ribero, Mark Coates, Rudolf Riedi, Shiram Sarvotham, Brent Hendricks and Richard Baraniuk: Multifractal Cross-Traffic Estimation. Proceedings of the 13th ITC Specialist Seminar on IP Traffic Measurement, Modelling and Management, September 18th-20th, 2000, Monterey, California, USA. Pages 15-1ff. http://www.dsp.rice.edu/publications/pub/itc00_cross_traffic.ps.gz.
- [RiberoT00] Vinay Ribero et al: Multifractal Cross-Traffic Estimation. Talk at 13th ITC Monterey http://www.dsp.rice.edu/~vinay/talks/itc00_talk.pdf.gz.
- [Tcpdump] <http://www.tcpdump.org/>.
- [Trimble] Palisade NTP Synchronisation Kit, Trimble Navigation, Inc., <http://www.trimble.com/products/catalog/timing/ntp.htm>
- [Veitch] Darryl Veitch's home page: <http://www.emulab.ee.mu.oz.au/~darryl/>.
- [WITS] Waikato Internet Traffic Storage. <http://wand.cs.waikato.ac.nz/wand/wits/>.
- [Ziedins00] Ilze Ziedins: On the output process from a finite buffer with long range dependent input submitted to IEEE Transactions on Networking, August 2000. <http://wand.cs.waikato.ac.nz/wand/publications>.
- [Zlib] Jean-loup Gailly and Mark Adler: zlib compression library. <http://www.info-zip.org/pub/infozip/zlib/>.

About the authors

Jörg Micheel grew up on the east side of the Berlin wall. He went overland studying computer science in Siberia and received his diploma engineer degree in 1992 from the Novosibirsk Institute for Electrical Engineering, Russia. He returned to unified Germany and worked for 5 years at GMD FOKUS in Berlin before being attracted by foreign countries again. For 18 month he was network researcher with SingAREN in Singapore. Since then he has found a new home with the University of Waikato, where he is working on Dag software and network measurements. Recently, Jörg has also become involved in the work of the NLANR MOAT team at SDSC/UCSD where he is expected to take the lead for the passive measurements and analysis group. Jörg is desperate to add more countries and different continents to his carrier path. If you are from a place not listed here and you wish to work with him on passive measurements, please send email to joerg@cs.waikato.ac.nz.

Ian Graham received his PhD from the University of Cambridge, UK, in Radio Astronomy. He spent most of his academic life developing hardware of various kinds. Before getting into network research and taking the lead on the Dag project he was obsessed by transputers. These days there are very few people who dare to distract him from his research work and make him attend to his obligations as the Dean of the School of Computing and Mathematical Sciences. Ian loves travelling, which is the only reason why we got ourselves into the trouble of writing this paper in the first place.

Nevil Brownlee received his PhD in Atmospheric Physics from the University of Auckland in 1975. He currently serves as Director for Technology Development at the University of Auckland's ITSS computer center. Nevil has been staring at IP packet headers for more than 10 years. As one of the pioneers of the Internet in New Zealand he has become famous for his NeTraMet Internet traffic meter. He is the author and co-author of several Internet RFC's and chair of the IETF RTFM working group. During the year 2000 Nevil spent much of his time on sabbatical with CAIDA and he continues to develop NeTraMet as a platform to gain insights into the nature of Internet traffic dynamics.