One-Way-Delay Measurements with CM Toolset[1]

Ulrich Hofmann
FH Salzburg
Schillerstr.30
A-5020 Salzburg / Austria

T:+43.662.4665.620
F:+43.662.4665.609

Thomas Pfeiffenberger
F&E Techno-Z
Jakob-Harringerstr 5/3
A-5020 Salzburg / Austria

T:+43.662.454888.725
F:+43.662.456174

Bernhard Hechenleitner
FH Salzburg
Schillerstr.30
A-5020 Salzburg / Austria

T:+43.662.4665.631
F:+43.662.4665.609

{ulrich.hofmann, bernhard.hechenleitner}@fh-sbg.ac.at, thomas.pfeiffenberger@newmedia.at

Abstract

Internet Quality of Service (QoS) is an actual research topic. The need for more or less guaranteed transmission rates (e.g. for bulk-data transfer applications), upper bounds for transmission delay and jitter (e.g. for real time applications like IP telephony, video-conferencing) stimulated the development of QoS transmission technologies like ATM, IP-RSVP and the actual discussion about Diff-Serv mechanisms. This paper describes the implementation and the first measurement trials with the One-Way-Delay measurement components of CM-Toolset.

The overall tool architecture consists of a distributed system of load generators and receivers (agents) and a measurement server, which stores the measurement results in a data-base. To measure the exact One-Way-Delay (OWD) the synchronised GPS clock system was integrated. The measurements were realised in a 4 hop heterogeneous IP and ATM network. The clock drift during the loop measurements was about 10ms, linear and deterministic. So the drift could be eliminated by statistical computations. The measurements were made for UDP flows under different load situations (1 UDP ....5 UDP, TSDU size, interarrival time), and for n*UDP + 1 TCP.

The measurements show very clearly the interaction and dependencies between the number of flows, their load parameters and the QoS parameters like packet loss and one way delay. The measurements of UDP+TCP multiplexed flows explains, why the concept of flow separation is important for QoS guarantees. CM Toolset can be used by protocol engineers for the development of new QoS-based network technologies by collecting measurement data for modelling the network behaviour. Network managers can tune the operational network to improve the quality e.g. of IP-Telephon.

Keywords

Quality of Service, Communication Measurement Toolset (CM Toolset), One-Way-Delay (OWD), Internet Protocol, TCP, UDP

1 Motivation

Quality of Service (QoS) becomes important in today's networks. These networks provide different QoS mechanisms to the user. Generally it is not possible for the user to control these different QoS mechanisms offered by the Internet Service Providers (ISP). For some IP based real-time services, like VoIP, a stochastic packet transfer delay can correspond with a degradation of the QoS, even if these variations do not generate packet losses.

For that reason the Communication Measurement Toolset (CM Toolset) was redesigned and implemented with new features like the integration of a database and the GPS clock.

With the new CM Toolset measurements were done in a testbed to test the new features and to evaluate the One-Way-Delay (OWD) [AKZ99] and the One-Way-Delay-Variation (OWDV) [D98].

[1] This project was sponsored by Telekom Austria

2 One-Way-Delay Testbed

The One-Way-Delay-Testbed includes two IP network clouds connected via an Asynchronous Transfer Mode (ATM) Wide Area Network (WAN) connection, going from Salzburg via Linz to Vienna and back to Salzburg.
For the connection a CISCO 7204 Router with a PA-A3 Enhanced ATM Port Adapter, which provides a traffic shaping algorithm, was used. The connection to the ATM network was established with a Fore Runner ASX 200 WG ATM switch with an OC3 Singlemode Interface.

Figure 2-1 shows the measurement testbed.

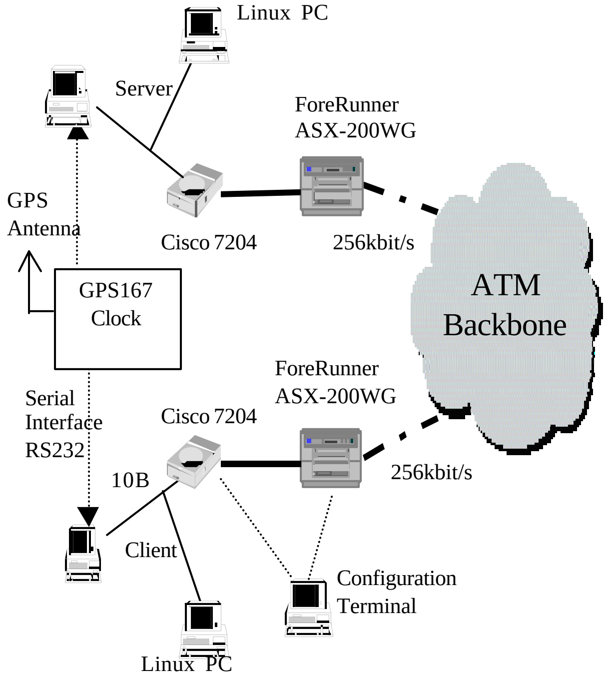

Figure 2-1: Infrastructure of the testloop Salzburg-Wien

3 Components and Parameters:

Fore ASX 200 WG ATM Switch:
For the connectivity to the ATM Backbone a 155Mbit/s OC-3c/STM-1 3MM/1SM Module (Series C) with four SONET/SDH ports (155Mbit/s) was used.
Cisco 7204 Router:
For the connection to the ForeRunner ASX 200WG ATM Switch a PA-A3 Enhanced ATM Port Adapter is used. This ATM Port Adapter provides a traffic shaping algorithm. Older PA-A3 ATM Port Adapters caused problems when a connection with a bandwidth lower than 10Mbits was established, because the older ATM Adapters did not provide traffic shaping, so too much ATM cells were sent. As a result, the next ATM switch discarded ATM cells on these flows and the maximum throughput went down to about 8-10 kbits on a 256 kbits link.

End-system:
The operating system *Linux* was used in the delay measurement testbed. The synchronisation with the GPS Clock is one background process. The resolution of the system-clock is 1 µs.
The socket programming code of the CM Toolset was written in UNIX C programming language.

4 Implementation of the CM Toolset

The CM Toolset offers measurement features for an evaluation of data communication protocols and networks. It provides a simple management platform to handle the measurement scenarios and to measure the IP performance parameters for different transport protocols including their multiplexing. The different protocol parameters (e.g. TSDU size, parameters of the load generators) can be controlled. The result and the parameters of the measurement is stored in a database. The impact of different network configuration and protocol parameters on QoS parameters can be analysed.

Components:
The CM Toolset consists of the following components:

- Communication Platform, Graphical User Interface (GUI)
- Database
- CMCaller
- CMDaemon
- Generator

Figure 4-1 shows the architecture of CM Toolset.

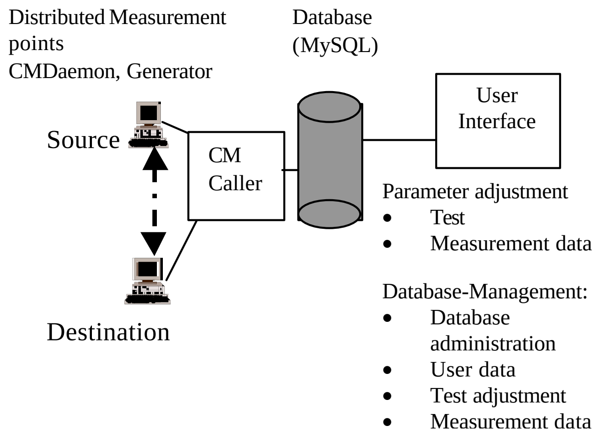

Figure 4-1: Architecture of CM Toolset

Graphical User Interface GUI:
The Graphical User Interface enables the user to specify the parameters for the measurement between the measurement points. Furthermore a test-scheduling for every hour or every day or other scheduling can be defined.

Parameters which can be adjusted:

- Source address (node where the traffic is generated).
- Destination address (the sink of the traffic).
- Amount of data (the number of Bytes or the number of packets)
- Ipv4-Sockets
- Ipv6-Sockets
- Nativ ATM (SEAN)
- Protocol to use (TCP,UDP)
- Window size (only for TCP), this is the number of bytes that the receiver is willing to accept. It is a 16-bit field, limiting the window to 65535 bytes.
- Inter-arrival-time of the sender (inter-transmittal-time of outgoing packets)
- Distribution of the inter-arrival-time (constant, exponential, bivalue, uniform, normal, lognormal, see below)
- Transport Service Data Unit (TSDU) size, this is the number of bytes which will be generated and sent to the network as one packet.
- Distribution of the TSDU size(constant, exponential, bivalue, uniform, linear, normal, lognormal, see below)
- Bundle of measurements, which means how often the test will be started at the same time and with the same parameters.
- Bundle of applications, which means how often the measurement will be started at the same time and with different parameters.
- Point to Multipoint, which means to start the measurement between one source and many destinations with the same parameters.
- Increment of Bundles, which means to start the measurement for example with one Bundle and raise the number of Bundles during the measurement.
- Number of tests, this is the number of tests started one after the other.

Database:
The database manages the test-parameters and the measurement results for each test and user. The database makes it possible to handle the tests automatically. Each test has its start time, and so it is possible to make a time schedule to repeat tests daily, weekly, and so on. This enables to run the tests automatically, triggered by the CMCaller.

CMCaller:
The CMCaller is the management module of the CM Toolset. It is responsible for the test scheduling and the handling of the measurement data. The CMCaller is the link between the DB and the remote measurement points. The information about the nodes which are concerned by a test are handed over from the DB to the CMCaller. With this information, which represents the testbed, each measurement point will be controlled from the

CMCaller.
The CMCaller polls the DB and searches for the next test. If there is more than one test in the DB with the same start-up time, a DB flag indicates which test has already been started or finished. So it is possible to start another test in parallel with different test-parameters nearly at the same time. The delay between the start of two or more tests, which appears because of the restriction to start only one test after the other, is very short regarding to the length of the test, and should not affect the measurement data of simultaneous running tests.

CMDaemon:
On each single measurement point the CMDaemon controls the start of a measurement. This means that the CMDaemon starts the Generator with the

specified parameters. The CMDaemon receives the test-parameters from the CMCaller and sends the measurement results back to the CMCaller.

Generator:
The Generator generates different kinds of traffic between two or more measurement points. The traffic is generated according to the parameters specified by the user. The parameters are shown above.
The generator is started by the CMDaemon on each measurement point in the network and receives the test parameters from the CMDaemon. On ending the test, the measurement data is sent to the CMDaemon and forwarded to the database via the CMCaller.
The following load and time distributions can be set by the generator:
Linear -, Exponential -, Uniform -, Normal -, Log Normal distribution.

5 Integration of the GPS clock for One-Way-Delay measurement

The time-synchronisation of two PCs or Workstations in the testbed was done by using a Meinberg GPS167 receiver [AKZ99]. With the installation of a GPS system, the internal PC software clock can be synchronised.
Figure 2-1 shows the test-loop Salzburg-Linz-Wien-Salzburg with the GPS receiver. Two end-systems are synchronised with the same GPS receiver. This caused the problem of a periodical drift between the two end-systems. Because of the internal control of the two RS232 interfaces of the GPS 167 receiver, a periodical linear drift of the two clocks occurs. This can be seen in the measurements of the next chapter.

Clocksync Application:
The synchronisation of the sender and receiver clock between two measurement points is very important for a OWD measurement. By using the serial interface of the GPS receiver, the GPS Time datagram from the GPS clock is sent. The time datagram is converted from an ASCII string to the number of seconds since 1.1.1970. This is the UNIX internal time format. At each remote measurement point the system clock is synchronised with this GPS time datagram.

6 Measurements

Clock Drift:

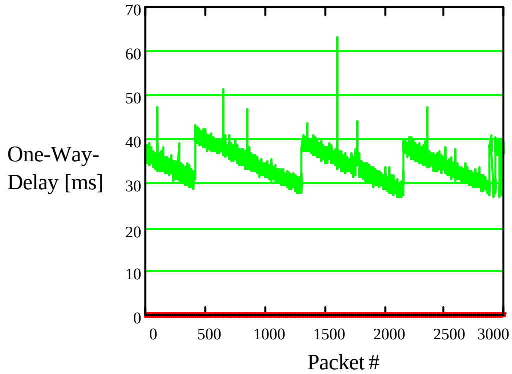

Figure 6-1: OWD of UDP 160 Byte, 40 ms sending interval (160/40) with clock drift

Figure 6-1 shows the clock drift during the measurement of 1 UDP flow with 160 Byte payload and a sending interval of 40 ms (160/40). The variation of the clock is about 10 ms. For some QoS measurements this phenomenon can be ignored - e.g. the variation for the OWD of intercontinental audio lies in the range of 200 ... 300 ms (see ITU delay-recommendations for audio and fax). Because of the deterministic growth of the inaccuracy, the drift can be eliminated by statistical computation.

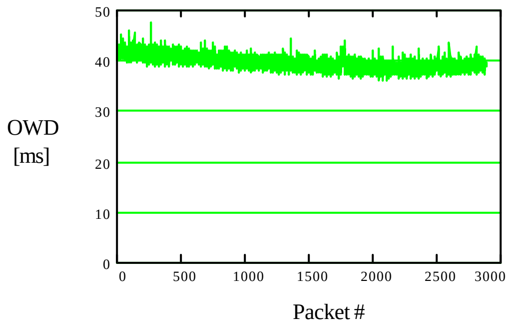

Figure 6-2: OWD of UDP 160/40 without clock drift

Figure 6-2 shows the result of the computational correction of the clock drift. The small variations of the OWD are the results of the stochastic behaviour of the operating systems of the end-systems and the routers.

OWD for Multiplexing:

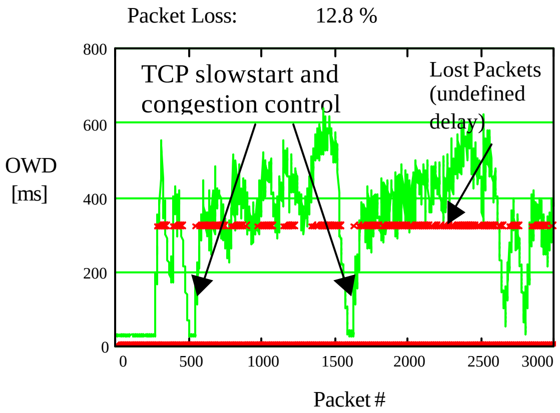

Figure 6-3: OWD of 3 UDP 80/20 streams multiplexed with 1 TCP flow

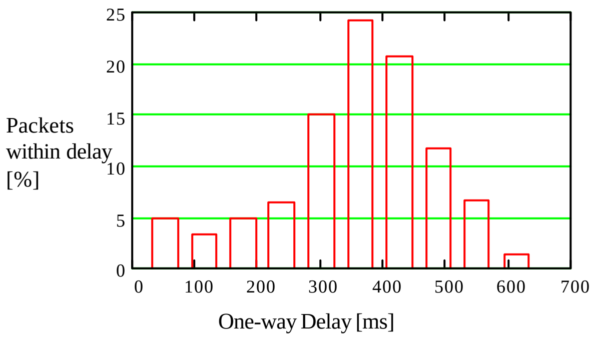

Figure 6-4: Histogram of the OWD of UDP 80 Byte, 20 ms sending interval

Figure 6-3 shows the statistics for the OWD of one of three UDP packet streams (80/20) which were multiplexed with a TCP flow. The system was heavily loaded which resulted in 12.8 % UDP packet loss. The fluctuations in Figure 6-3 illustrate the typical interaction of responsive and unresponsive flows. The IETF concept of resource reservation (Int.-Serv., Diff.-Serv., [BBC98], [W97], [SPG97], [BBB99]) offers algorithms to isolate these different types of packet flows. For best-effort traffic the responsive flows can be isolated from the unresponsive flows by the use of statistical algorithms implemented in the routers or/and edge devices [ZFH99].

Influence of the TCP-slow-start and congestion control algorithms:

There is no packet loss during the start of the measurements (TCP slow-start). After about 300 packets TCP causes packet loss and heavy load with high one-way-delays. Then the TCP-congestion control algorithm gets active. This is the reason for better delay and loss values for UDP in the 400-packet-region. After some time (packet # >900) TCP converges to a sending rate with low TCP loss. The consequence is that the delay and loss values of UDP converge to 350 ms and 12% packet loss.

Figure 6-4 shows the statistical aggregation of the packet series of Figure 6-3. The 5% bar around the OWD of 40ms represents the repeating low load situations in Figure 6-3 during the TCP slow start and congestion control intervals.

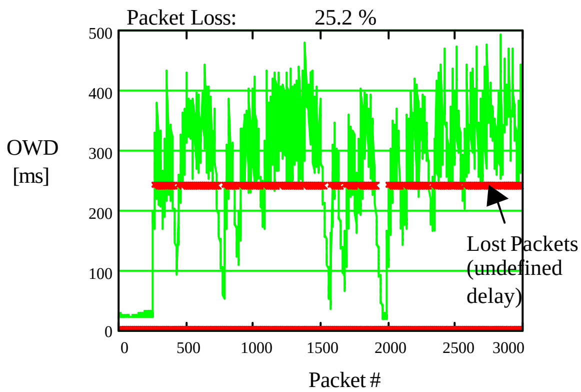

Figure 6-5: OWD of 5 UDP 80/20 streams multiplexed with 1 TCP flow

Figure 6-5 shows the same measurement scenario with 5 UDP flows. The higher load results in smaller intervals without packet loss and in a larger OWD variation.

Correlation of OWD for Multiplexing:

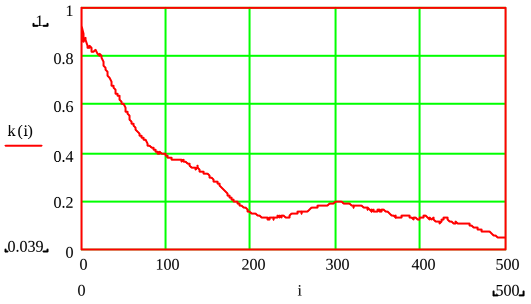

Figure 6-6: Correlation of OWD of UDP stream regarding to figure 6-3

Figure 6-6 shows the correlation of the OWD of the UDP stream (80/20) described in figure 6-3. This UDP stream was multiplexed with two other UDP streams (80/20) and one TCP flow. As expected it shows that successive UDP packets have similar OWD values. The value of correlation decreases strongly within a correlation distance of 200 packets and then converges slowly against 0 when increasing the correlation distance. This indicates a noticeable dependence of the OWD of successive packets. At a correlation distance of 100 packets the value of correlation still is 0.4.

There is a small correlation peak at i=300. This correlation is caused by the repeating OWD-peaks in Figure 6-4 . There we have about 10 peaks (max or min) in the interval of 3000 packets. This results in a peak-period of about 300.

As can be seen in figure 6-3, several packets of the measured UDP stream have been lost (the packet loss rate was 25.2 %). This has to be considered in the calculation of the correlation values.

Imagine the following sequence of transmitted packets (‘x’ indicates a lost packet):

1 2 3 4:x 5 6:x 7 8 .....

When calculating the correlation value with a correlation distance of e.g. k := 2, then the following pairs have to be considered (‘x’ indicates that this pair must not be added in the calculation):

(1,3) (2,4):x (3,5) (4,6):x (5,7) (6,8):x ...

7 Analysis of the results

The integration of the GPS clock in the CM Toolset allows measurements with an accuracy of about 10 ms. The deterministic inaccuracy can be eliminated by a correction calculation. It was shown that the measured OWD of the unloaded operational network ,of about 40 ms, satisfies the demands of isochronous UDP based applications, like IPTel.

For multiplexed traffic of responsive (TCP) and unresponsive (UDP) flows the one-way-delay increases from 40 ms (no load on the network) to peak-values of more than 400 ms (peak-factor > 10). This is caused by the waiting times in the routers and the overload situation on the links.

Therefore, the routers should drop a higher rate of the UDP-flow to guarantee a delay lower than 200 ms for all packets that are not dropped and to avoid unnecessary network load. An end-to-end QoS-control using CM Toolset in a real time measurement scenario could use these results to decrease the sending rate. The correlation analysis shows the short- and long-term correlation and an OWD-period at a distance of about 300 packets. These results can be used for further research in protocols and router-mechanisms. The OWD-values represent the stochastic behaviour of oscillating router buffer-occupation and the interaction between multiplexed responsive and non-responsive flows.

CM Toolset gives the possibilities to analyse the OWD and loss behaviour of UDP and TCP flows. The network engineers can use CM Toolset for an optimum tuning of the network components, like increasing bandwidth and separating UDP and TCP flows by InServ or DiffServ mechanism.

CM Toolset can be used by protocol engineers for the development of new QoS-based network technologies by collecting measurement data for modelling the network behaviour.

References:

[AKZ99] Almes.G , S. Kalidindi, M. Zekauskas, “A One-way Delay Metric for IPPM”, <draft-ietf-ippm-delay-0.txt>, November 1999

[D98] Demichelis,C., CSELT, “ Instantaneous Packet Delay Variation Metric for IPPM”, <draft-ietf-ippm-ipdv-02.txt>, November 1998

[ZFH99] Th.Ziegler, S.Fdida, U.Hofmann, “Distribution of unfriendly Best Effort Flows in the Internet“, IEEE Conf. on Broadband Communication, Hongkong 99

[HML99] U.Hofmann, I.Milouchewa, L.Lamont, “Global Measurement Infrastructur of Application Classes and their Bundling over ATM Wide Area Network“, ECMAST99, Madrid

[BBC98] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss, “An Architecture for Differentiated Services” , RFC2475, December 1998

[W97] J. Wroclawski, “Specification of the Controlled-Load Network Element Service”, RFC2211, September 1997

[SPG97] S. Shenker, C. Partridge, R. Guerin, “Specification of Guaranteed Quality of Service”, RFC2212, September 1997

[BBB99] Y.Bernet, James Binder, S.Blake, M.Carlson, B.E.Carpenter, S.Keshav, E.Davies, B.Ohlman, D.Verma, Z.Wang, W.Weiss, “A Framework for Differentiated Services”, <draft-ietf-diffserv-framework-02.txt>, February 1999