High Speed TCP/IP Experiment over International ATM Test Bed

Toru Hasegawa, Mitsuru Yamada*, Kanji Hokamura*,*
Kohei Yoshiizumi, Teruyuki Hasegawa*, Toshihiko Kato*,*
*Linda Galasso**, and Hiroyuki Fujii****
** KDD, ** AT&T, *** NTT*
** KDD (Kokusai Denshin Denwa Co., Ltd.) R&D Labs.,*
2-1-15, Ohara, Kamifukuoka-shi, Saitama, Japan.
phone: +81-492-78-7368 fax: +81-492-78-7510
e-mail: hasegawa@hsc.lab.kdd.co.jp

Abstract

AT&T, KDD and NTT conducted a joint ATM trial called Multimedia Application Project (MAP) for the purpose of testing and validating network-based broadband multimedia applications and services between US and Japan. Although the Internet based network is promising for broadband multimedia communications, there is a performance problem that TCP is the throughput bottleneck in an international ATM network with large propagation delay. In MAP, the high speed TCP/IP experiment was jointly performed by the three companies with the leadership of KDD, for the purpose of validating the high speed TCP/IP infrastructure using TCP gateway proposed by KDD. In this infrastructure, a pair of TCP gateways are introduced in international ATM networks, and they improve the TCP throughput by introducing a link-by-link flow control and a sufficient window size between gateways, without any modifications to end terminals.

In the high speed TCP/IP experiment, the throughput of TCP communications with and without TCP gateways were measured under various conditions. This paper describes the outline and the results of this experiment. The experiment was successful and the TCP gateways greatly improved the throughput of TCP based applications. For example, the response time of WWW access retrieving a 600 Kbyte image file was improved from 13.5 seconds to 2.6 seconds, and the ftp throughput was improved from 1.5 Mbps to 22 Mbps.

Keywords

ATM, TCP/IP and International ATM Networks

Broadband Communications P. Kühn & R. Ulrich (Eds.)
© 1998 IFIP. Published by Chapman & Hall

1 INTRODUCTION

Recent technological advances in the fields such as ATM (asynchronous transfer mode) and SDH (synchronous digital hierarchy) have realized a high speed digital transport network. Many communication providers have started ATM network services for domestic and international environments, for the purpose of allowing broadband multimedia applications and services to be deployed in wide area networks. However, further investigations are required to understand the behaviors of multimedia applications over ATM networks and to identify advanced ATM network functionality supporting multimedia services.

From these backgrounds, AT&T, KDD and NTT conducted a joint ATM trial called *Multimedia Application Project* (MAP), between July 1996 and February 1997. This trial created an international ATM test bed for validating network-based broadband multimedia applications and services between US and Japan, and performed four kinds of broadband multimedia experiments[1]. The high speed TCP/IP experiment is one of the experiments and was jointly performed by the three companies with the leadership of KDD.

The Internet based network is one of the most promising network configurations, because it allows the use of the huge legacy of existing application software. However, there is a performance problem that TCP is the throughput bottleneck in a network with large *bandwidth-delay products*, such as an international ATM network (Stevens, 1994). That is, the throughput is limited by the window based flow control because the maximum value of TCP window size is limited by 64 Kbyte and most applications use smaller values. Against this problem, RFC 1323 proposed the TCP extensions including the TCP window scale option which uses the window sizes larger than 64 K bytes (Jacobson, 1992). However, there are still some problems such that RFC 1323 is supported by only limited computers, and that the performance is also degraded due to the TCP congestion control (Kato, 1997). Therefore, it will be an effective approach for an wide area ATM network to provide the functionality to attain a high speed TCP/IP communication.

KDD has been studying the realization of a *high speed TCP/IP infrastructure* over an international ATM network, and has proposed a *TCP gateway approach*[2]. Here, a pair of gateways are introduced in an international ATM network, and each of them works as an intermediate system along a TCP connection and supports both the IP function and the transport protocol handling (Kato, 1997 and Hasegawa, 1995). As for the transport protocol, the gateway uses TCP to communicate with an end terminal, and a KDD proprietary high speed transport protocol with a sufficient window size to communicate with the other TCP gateway. Using these two transport protocols, the acknowledgment and retransmission of data segments are performed on a link-by-link basis, i.e. between

---

[1] *MAP focused on broadband PC-VMS (Virtual Meeting Service) (Beaken, 1995) led by AT&T, high speed TCP/IP (Kato, 1997 and Hasegawa, 1995) led by KDD, Virtual LAN (Hariu, 1997) led by NTT and broadband imaging (Nakabayashi, 1995) led by NTT.*

[2] *TCP gateway is KDD proprietary.*

the gateway and an end terminal or between the gateways. As a result, it is possible to increase the end-to-end throughput without giving any modification to the end terminals and, furthermore, without making the end terminals aware of the existence of the TCP gateways.

In the high speed TCP/IP experiment, the three companies collaboratively evaluated a high speed TCP/IP infrastructure using TCP gateway over an international ATM test bed. In this experiment, the throughput of TCP/IP communications with and without TCP gateway was measured using various applications under various conditions. This paper describes the results of the high speed TCP/IP experiment conducted in MAP. In section 2, we describe the problem of TCP and the structure of high speed TCP/IP infrastructure. In section 3, we describe the network configuration of international ATM test bed used in this experiment. In section 4 , we describe the experiment results.

2 HIGH SPEED TCP/IP INFRASTRUCTURE USING TCP GATEWAY

2.1 Problems of TCP over Wide Area ATM Networks

TCP performs the flow control based on the sliding window mechanism between end terminals on an end-to-end basis. In order to obtain a high throughput corresponding to the network bandwidth, it is required that the TCP window size is large enough to transmit data segments continuously during one RTT (round trip time). As shown in Figure 1, if the window size is insufficient, the throughput will be degraded because the sender waits for ACKs (acknowledgments) after it finishes sending all data segments within the window. In this situation, the throughput Th [bps] is given by Equation (1). In this equation, W [byte], Rtt [sec], M [byte] and C [bps] indicate the TCP window size, RTT, the maximum transmission unit size and the bandwidth respectively.

$$Th = W * 8 / (Rtt + M * 8 / C) \qquad (1)$$

Many applications over TCP use window size such as 8 K bytes. In the case of W = 8 K bytes, M = 9188 bytes and C = 6 Mbps, Equation (1) gives the result that Th is roughly 310 Kbps and 110 Kbps when RTT is 200 ms and 600 ms, respectively. This throughput does not increase no matter how large bandwidth the wide area ATM networks provide.

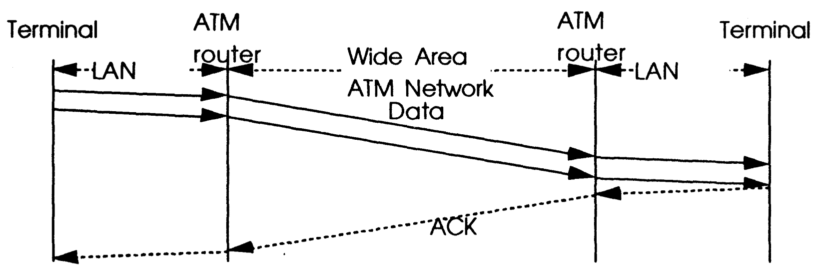

Figure 1 Sequence of End-to-End Flow Control with Insufficient Window Size.

2.2 Structure of High Speed TCP/IP Infrastructure Using TCP Gateway

Figure 2 shows a configuration of the high speed TCP/IP infrastructure over an international ATM network. As described in this figure, a TCP gateway works as an adjunct equipment of an international ATM network and a pair of gateways collaborate with each other to provide a high speed TCP/IP for customers. The functions of TCP gateway are summarized as follows.

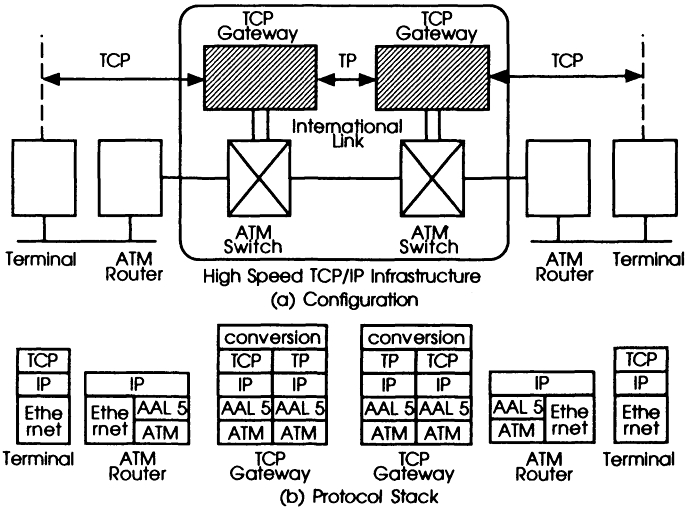

Figure 2 High Speed TCP/IP Infrastructure.

(1) TCP gateway provides both IP functions such as IP routing and the gateway function at the transport layer. It applies IP and gateway functions to TCP segments.
(2) TCP gateway introduces a KDD proprietary high speed transport protocol, called *TP*. It is designed so as to be suitable for wide area ATM networks.
(3) TCP gateway uses one VC (virtual channel) with a customer, and another VC with the other TCP gateway. When more than one customers are connected to a TCP gateway, the data from the VCs between the gateway and the customers are aggregated into the VC between the gateways. TCP gateway communicates with end terminals using TCP, and communicates with the other gateway using TP.
(4) TCP gateway performs the protocol conversion between TCP and TP in the following way.

- The establishment and release of TCP connections and TP connections are performed together on an end-to-end basis.
- The data transfer function such as the flow control and the error recovery is performed on a link-by-link basis. For the link between two TCP gateways, TCP gateway allocates sufficient window sizes for the long propagation delay.

- In order to relay data segments, TCP gateway transfers them according to the flow control individually for the customer side and the TCP gateway side, buffering them if required by the flow control, and sends ACKs individually for the customer side and the TCP gateway side.

A communication sequence during the data transfer phase is illustrated in Figure 3.

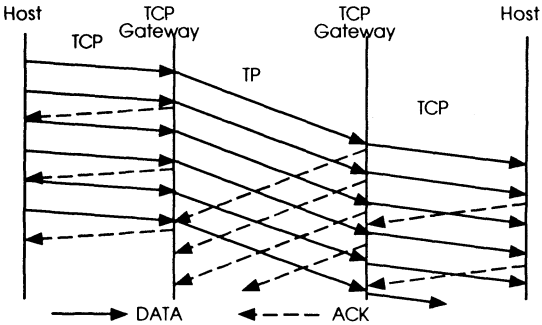

Figure 3 Communication Sequence During Data Transfer Phase.

(5) TP introduces two levels of flow control, the flow control for an individual TP connection, and that for the aggregate TP connections over VC between the TCP gateways. The purpose of the latter flow control is used for the aggregated TP traffic not to excess the bandwidth of the VC between the gateways.

3 INTERNATIONAL ATM TESTBED

In MAP, an international ATM test bed was established among AT&T (Holmdel in US), KDD (Shinjuku in Japan) and NTT (Musashino and Yokosuka in Japan) through the Pacific Ocean. Figure 4 illustrates the network configuration for the high speed TCP/IP experiment, which is a part of the whole test bed. The features are as follows:

(1) Network Facility

- A 45 Mbps transmission facility (DS-3) was used for an international link. The traffic was routed on DS-3 facility in TPC-4, an optical fiber submarine cable between US and Japan.
- A 45 Mbps transmission facility (DS-3) was used for a US domestic network.
- A 155 Mbps transmission facility (STM-1) was used for a Japanese domestic network.
- RTT between AT&T and KDD, and RTT between KDD and NTT Yokousuka, measured by ping commands, were 175 ms and 2 ms, respectively.

(2) VP (Virtual Path) Characteristics

A virtual path network was established among the four ATM VP cross-connects at AT&T Holmdel, KDD Shinjuku, NTT Musashino and NTT Yokousuka. VP connections were established between every pair of the cross-connects.

AT&T and KDD networks (International network and US domestic network) use VBR (variable bit rate), and NTT networks use CBR (constant bit rate).

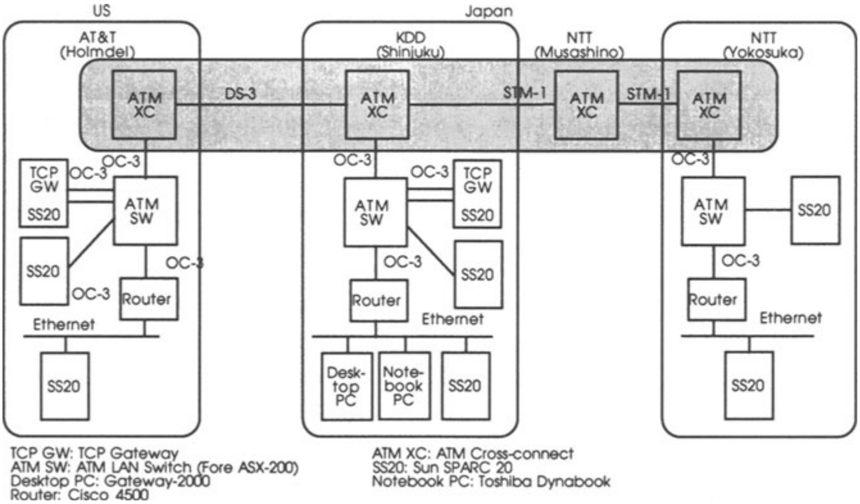

Figure 4 International ATM Test Bed.

(3) TCP Gateways
The TCP gateways were installed at AT&T Holmdel and KDD Shinjuku. A VC connection was established between two TCP gateways.
(4) LANs and Terminals
At AT&T, KDD and NTT Yokosuka, both ATM LAN and Ethernet LAN were installed as customer premises networks. Ethernet LANs were accommodated to ATM network using ATM routers.

TCP terminals were installed at AT&T Holmdel, KDD Shinjuku and NTT Yokosuka. TCP terminals at AT&T Holmdel and KDD Shinjuku communicate with the TCP gateway at the same location. TCP terminals at NTT Yokosuka communicate with the TCP gateway at KDD Shinjuku.

4 EXPERIMENT RESULTS

4.1 Methodology

In order to validate the performance of the high speed TCP/IP infrastructure using TCP gateway, the following approaches are taken in the experiment.
(1) Both the application level throughput of widely used applications, such as WWW (world wide web) and ftp, and the TCP level throughput are measured. The application level throughput includes the processing overhead of application protocols and the overhead of accessing disks and displaying. The TCP level throughput is measured by memory to memory copy through TCP protocol, and indicates the pure protocol processing overhead. A free software called ttcp is used for measuring the TCP level throughput.
(2) As TCP terminals, both workstations and personal computers are used in the experiment. By using different kinds of terminals, it is possible to test different

cases for processing capability of terminals, and to test different TCP implementations.
(3) As a customer premises network, both ATM LAN and Ethernet are used in the experiment. This can evaluate different transmission speed of terminals, and different values of TCP parameters such as MSS (maximum segment size) and window size.
(4) The number of TCP connections between TCP terminals is changed from one to eight. This can evaluate the performance of TCP gateway for multiple connections, and the fairness of TCP gateway for individual connections.
(5) As described in section 3, TCP gateway is installed at only KDD Shinjuku in Japan side. The performance is measured for the case that TCP terminals are co-located in KDD Shinjuku and the case that they are located in NTT Yokosuka. This can evaluate how the domestic propagation delay in Japan affects the throughput.

4.2 Test Conditions

(1) SS-20 SPARC station with 60 MHz SuperSPARC and SunOS 4.1.3 is used for TCP gateway. It has 64 M byte main memory, and two Fore SBA-200E ATM NICs (Network Interface Cards).
(2) The following workstations and personal computers are used as TCP terminals.

- Workstation: SS-20 SPARC station with 150MHz HyperSPARC, 32 M byte main memory, a Fore SBA-200E ATM NIC, and Solaris 2.5.1.
- Notebook Personal Computer: Toshiba Dynabook with 90MHz Pentium, 40 M byte main memory, and Windows 95.
- Desktop Personal Computer: Gateway-2000 with 200MHz Pentium Pro, 64 M byte main memory, and Windows 95.

(3) The window sizes of high speed transport protocol between the TCP gateways are 1M bytes and 256 K bytes, when ATM LAN and Ethernet are used as customer premises networks, respectively. These values of window size are large enough for 45 Mbps and 10 Mbps, the values corresponding to the bandwidth of bottleneck networks, respectively.

Table 1 Combination of Computers and Customer Premises Network.

	WWW	ftp	ttcp
SS-20 -> SS-20 (ATM LAN)	done	done	done
SS-20 -> SS-20 (Ethernet)	done	done	done
SS-20 -> Notebook PC (Ethernet)		done	
SS-20 -> Desktop PC (Ethernet)		done	
Notebook PC -> SS-20 (Ethernet)	done	done	
Desktop PC -> SS020 (Ethernet)	done	done	

(4) All the computers connected to ATM networks perform traffic shaping with 30.0 Mbps as a peak cell rate. This corresponds to 27.1 Mbps as an ATM payload transmission rate.

(5) In the experiment, throughput is measured for WWW, ftp and ttcp communication. Table 1 indicates the combinations of computers and customer premises networks used for individual throughput measurements. It should be noted that the arrow between computers indicates the direction from client to server.

4.3 Results of WWW and FTP Throughput Measurements

4.3.1 WWW Throughput Measurements

A WWW server is installed in the SPARC station at AT&T Holmdel, and it stores twelve image files formatted in GIF (graphics interchange format) whose sizes are from 321,051 bytes to 572,101 bytes. A WWW client using Netscape Navigator 3.0 is installed in KDD Shinjuku, and it retrieves image files and displays the received data. As an application level throughput of WWW communication, the response time is measured, which is the time between when the retrieval request is invoked and the time when the retrieved data is fully displayed. The response time is measured for the cases with and without TCP gateway using the combinations of computers and customer premises networks described in Table 1. For all the measurements, TCP gateway improves the response time.

Figure 5 shows the measurement results of response time using a SPARC station as a WWW client and ATM LANs as customer premises networks in both client and server sides. For example, while the response time for 572,101 byte file without TCP gateway is about 13.5 seconds, that with the gateway is just 2.6 seconds. For the cases that a personal computer is used as a WWW client and that Ethernet is used as a customer premises network, the measurements give similar results for response time. As the results of this measurement indicate, it can be said that the high speed TCP/IP infrastructure using TCP gateway increases an application level WWW performance more than five times.

4.3.2 FTP Throughput Measurements

The throughput is also measured between an ftp server installed at AT&T Holmdel and an ftp client at KDD Shinjuku. In the measurement, an ftp client gets a binary

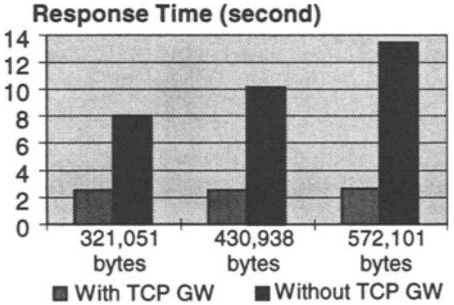

Figure 5 Response Times.

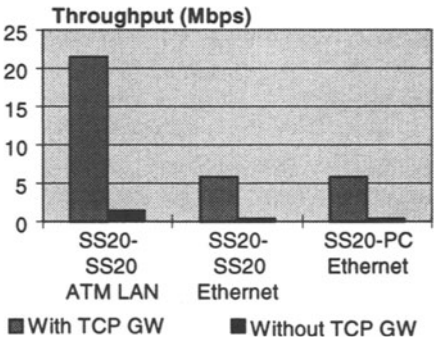

Figure 6 Ftp Throughput.

file stored in a disk whose size is 100 Mega bytes. The throughput is measured for the cases with and without TCP gateway using the combinations of computers and customer premises networks described in Table 1. As a result, TCP gateway improves the throughput for all the measurements.

Figure 6 shows the results for some cases of measurements. For the case that ATM LAN is used as customer premises network, the high speed TCP/IP infrastructure using TCP gateway improves the throughput more than ten times, and for the case of Ethernet, it improves the throughput more than eighteen times.

4.4 Results of Ttcp Throughput Measurements

In order to evaluate the TCP level performance of the high speed TCP/IP infrastructure, the throughput measurement using ttcp software is performed using SPARC stations. The improvement of the TCP level throughput by TCP gateway is evaluated for the cases that one TCP connection is established and that more than one TCP connections are established simultaneously. The measurement also focuses on the case that there is some propagation delay between the TCP terminal and the TCP gateway in Japan.

The TCP parameters used in this measurement are summarized as follows.

- Nagle algorithm is used.
- Ttcp writes 8192 byte data at a time.
- The values of MSS are 9148 bytes and 1460 bytes when the customer premises network is ATM LAN and Ethernet, respectively. It should be noted that, in this measurement, SPARC stations running Solaris 2.5.1 sent data segments whose length is 8192 bytes for the ATM LAN customer premises network.

The send and receive socket buffer sizes are default values used in Solaris 2.5.1. They are 36,592 bytes and 8760 bytes for the ATM LAN and the Ethernet customer premises networks, respectively.

4.4.1 Throughput Measurement for One TCP Connection

Figures 7 and 8 show the TCP level throughput with and without TCP gateway, when the number of TCP connections are changed for the ATM LAN customer premises network and Ethernet customer premises network, respectively.

For the case of one TCP connection, the following results and discussions are obtained.

(1) When TCP gateway is not used, the throughput of one TCP communication is 1.42 Mbps and 0.32 Mbps for ATM LAN and Ethernet, respectively. On the contrary, when TCP gateway is used, the throughput is improved more than ten times faster than the original TCP throughput. The improved throughput is 21.2 Mbps and 5.69 Mbps for ATM LAN and Ethernet, respectively.

(2) The reason why the throughput without TCP gateway is different for ATM LAN and Ethernet is that the socket buffer sizes, which correspond to the window sizes, are different. In the case of ATM LAN customer premises network, four 8192 byte segments are sent continuously corresponding to one window size. The

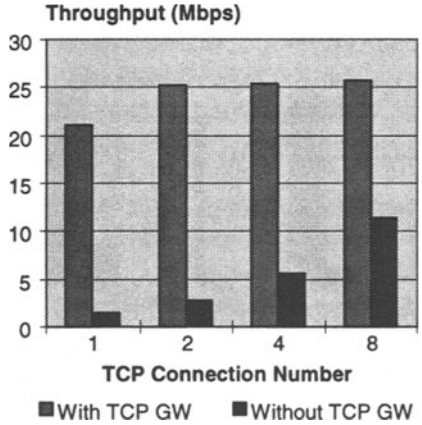

Figure 7 TCP Level Throughput for ATM LAN Customer Premises Network.

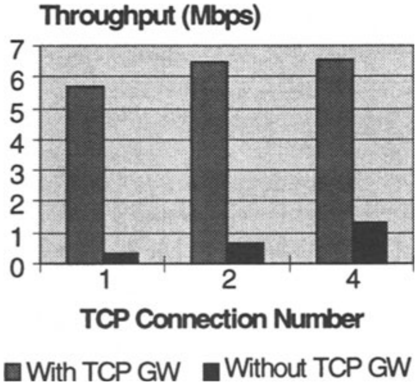

Figure 8 TCP Level Throughput for Ethernet LAN Customer Premises Network.

throughput without TCP gateway, which can be calculated by equation (1), is 1.47 Mbps. In the case of Ethernet customer premises network, five 1460 byte segments are sent for one window size, and the throughput without TCP gateway is 0.33 Mbps. The obtained results match these calculated values.

4.4.2 Throughput Measurement for Multiple TCP Connections

Figures 7 and 8 also give the TCP level throughput with and without TCP gateways, when TCP connections are used between multiple pairs of TCP terminals. For multiple TCP connections, the following results and discussions are obtained.

Table 2 TCP Throughput of Individual TCP Connection.

Connection / Num. of Connection	1	2	3	4	5	6	7	8
2	12.64	12.64	-	-	-	-	-	-
4	6.36	6.34	6.34	6.34	-	-	-	-
8	3.25	3.22	3.21	3.20	3.20	3.20	3.19	3.19

Throughput : Mbps

(1) When TCP gateway is not used, the total throughput of multiple TCP connections increases in proportion to the number of connections. However, the total throughput is still much lower than the throughput when TCP gateway is used.

(2) When TCP gateway is used, the total throughput of multiple TCP connections is more than 25 Mbps for ATM LAN customer premises network, and is more than 6.5 Mbps for Ethernet customer premises network. Since the ATM payload

transmission rate is 27.1 Mbps as described in section 4.2, the total throughput for ATM LAN is considered to be limited by this rate.
(3) The throughput of individual TCP connections is also measured for ATM LAN customer premises network. The results are listed in Table 2, and the throughput is the same for individual connections. This means that TCP gateway handles multiple TCP connections fairly.

4.4.3 Throughput Measurements for Domestic Propagation Delay in Japan

Figure 9 shows the TCP level throughput with TCP gateway and ATM LAN customer premises network for the case that TCP gateway and TCP terminals are co-located in KDD Shinjuku (without domestic delay), and the case that TCP gateway in KDD Shinjuku and TCP terminals in NTT Yokosuka (with domestic delay). Figure 10 shows the corresponding throughput when Ethernet is used as a customer premises network. As for one TCP connection, the TCP throughput with domestic delay is a little worse than that without the delay. The throughput degradation by the local access line delay is more significant for ATM LAN than for Ethernet. On the contrary, when multiple TCP connections are used, the total throughput is almost the same for both cases, as shown in Figs. 9 and 10.

5 CONCLUSION

In this paper, we described the result of the high speed TCP/IP experiment jointly performed in Multimedia Application Project conducted by AT&T, KDD and NTT. The experiment focused on validating the high speed TCP/IP infrastructure using TCP gateway over an international ATM network. In the experiment, the

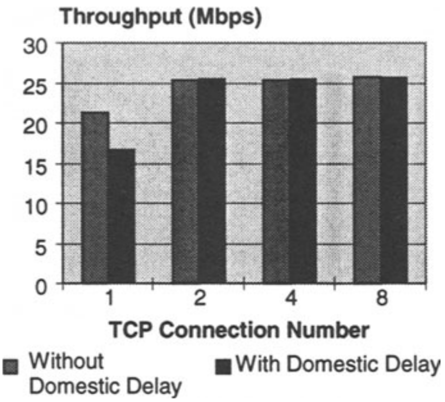

Figure 9 TCP Level Throughput with Domestic Delay for ATM LAN Customer Premises Network.

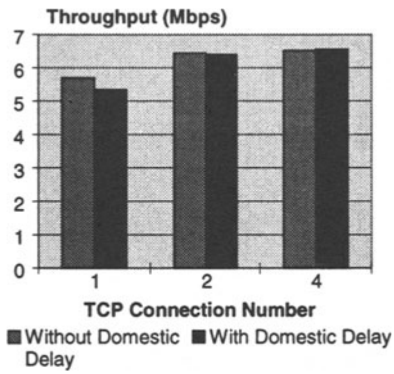

Figure 10 TCP Level Throughput with Domestic Delay for Ethernet Customer Premises Network.

throughput of TCP communications with and without TCP gateway were measured under various conditions. The experiment results are summarized as follows.

- TCP gateway greatly improved the throughput of TCP based applications. For example, the response time of WWW access retrieving a 600 Kbyte image file was improved from 13.5 seconds to 2.6 seconds, and the ftp throughput was improved from 1.5 Mbps to 22 Mbps.
- The throughput improvement by TCP gateway was confirmed for different TCP terminals and customer premises networks, and for different number of TCP connections.
- When multiple TCP connections were established simultaneously, TCP gateway improved the throughput of each connection fairly.
- If TCP gateway and terminal were located in different locations and if one TCP connection was used, the TCP throughput was a little worse than if they were co-located. On the contrary, when multiple connections were used, the delay did not degrade the total TCP throughput at all.

As described above, the high speed TCP/IP experiment was successful and has proved that TCP gateway can realize a high throughput infrastructure across an international ATM network for Internet based broadband applications and services. multimedia

6 REFERENCES

Stevens, W. R. (1994), TCP/IP Illustrated, Volume 1 [The Protocols]. Addison-Wesley.

Jacobson, V., Braden, R. and Borman, D. (1992), TCP Extensions for High Performance. RFC 1323.

M. Beaken, M. (1995), The Provision of Intelligent Agent-Based Enhanced Multimedia Network Services. AT&T TECHNICAL JOURNAL, pp.68-77.

Kato, T., Hokamura, K., Yamada, M., Hasegawa, T., Hasegawa, T. and Sawada, K. (1997) TCP Gateway Improving Throughput of TCP/IP over Wide Area ATM Networks. ISS'97 (International Switching Symposium), pp. 35-42.

Hasegawa, T., Hasegawa, T., Kato T. and Suzuki, K. (1995) Implementation and Performance Evaluation of TCP Gateway for LAN Interconnection through Wide Area ATM Network (in Japanese). IEICE Trans. Comm., vol.J79-B-I, no.5, pp.262-270.

Hariu, T. and Tanimoto S. (1997), 'Logical Office' Service: A Mobile VLAN Service for a Mobile Computing Environment. APSITT'97 (Asia-Pacific Symposium on Information and Telecommunication Technologies).

Nakabayashi, K. (1995) An Intelligent Tutoring System on World Wide Web: Towards an Integrated Learning Environment on a Distributed Hypermedia. Proc. of ED-MEDIA95, June 1995.

Acknowledgments

We wish to thank all the members of KDD, AT&T and NTT who participated in Multimedia Application Project. We would like to express special thanks to Dr. Y. Hatori, KDD, Dr. D. Rajala and Mr. R. Ramamurthy, AT&T, and Dr. S. Hatano and Mr. Y. Harada, NTT for their valuable advice.