Delay Characteristics of High-Speed Internet Access Network
: One Case Study

Kug Chang Kang, Ki-Dong Nam, Duk Joong Jeon, Seong Youn Kim

Network Strategy Dept., Network Technology Labs., ETRI,
161, Gajeong-Dong, Yuseong-Gu, Daejeon, 305-350, Korea
TEL: +82-42-860-6149 FAX: +82-42-860-6858
{kckang, kdnam, jooniii, kimsy}@etri.re.kr
ETRI

Abstract

In Korea, the number of high-speed Internet users grows so explosively that it seems to have touched its saturation point already. For NSPs to survive in this market situation, 'competition-of-volume' has no more meanings but 'competition-of-quality' matters. Network performance management to provisioning QoS-guaranteed network is one of urgent problems for Korean NSPs currently.

Until quite recently, NSPs in Korea have had the 'war-of-speed' but the speed is no more critical performance parameter in the age of quality-competition. It is getting more and more important to characterize network performance with regard to packet delay, packet loss, delay variation, and so on.

In this research, we examined the delay characteristics of ADSL service which is the most prevailing high-speed internet access service in Korea. Under end-to-end and end-to-server environment, we measured traffic delays path by path across the whole network intervals to see on which interval the bottleneck of the delay is. Additionally, we plotted delay distribution from the measurement to check the stability of delay performance. Further, we studied the relationship between the delay and packet size, and between the delay and network utilization.

Keywords : network performance, end-to-end delay, delay distribution

Introduction

- The growth in the number of high-speed Internet subscribers is being saturated in Korea
  - The Management of network performance is crucial in the age of 'Competition of Quality'
- Questions to answer;
  - What is the network performance?
  - How to measure and how to manage performance parameters?
- *What are the delay characteristics of the Internet?*
  - *How much is the delay?*
  - *Is it stable?*
  - *What factors are affecting the delay?*
  - *What is the critical bottleneck across the whole network path?*

ETRI 2

1. Introduction

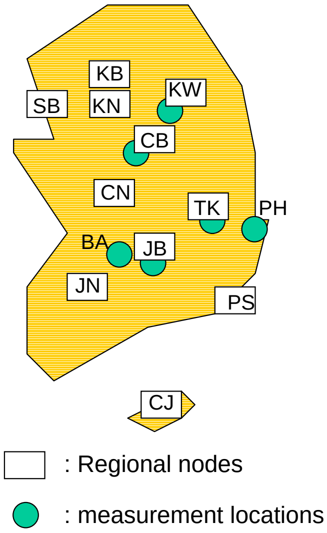
Due to the rapid advance of the Internet services market in Korea, the subscriber growth seems to have touched it's saturation point so that the 'competition-of-volume' between service providers has no more meanings in this market situation. The subscribers are demanding more refined QoS, more diversified features of services and more enhanced performance. That is, the era of 'competition-of-quality' has already come. Consequently, network performance management to provisioning QoS-guaranteed network is one of critical issues for network providers to survive in the Internet services market.

This study has focused on ADSL performance since it is the most prevailing high-speed Internet access service in Korea. Main research topic is about the delay characteristics of ADSL services; How much is the delay? Is it stable? What is the critical bottleneck across the whole network path? What factors are affecting the delay?

We examined two typical types of connection on the Internet. One is end-to-end connection for Peer-to-Peer communications, the other, end-to-server connection for Client-Server communications. In the early stage of the Internet the Client-Server model prevailed but, the use of P2P application is being increased nowadays. So we examined both end-to-end traffic delay and end-to-server traffic delay. To grab the real characteristics of delay, we measured the delays on a commercial Internet in Korea.

Measurement environment

- Target network is KORNET
  - 11 regional nodes
  - DQMEs are used under ADSL-Lite environment
  - 82 DQMEs are installed across the country
- Measurements are done at 6 locations
- RTTs are measured using ping command



ETRI 3

2 The Measurements

2.1 Measurement environment

The target network of this study is the KT Internet, called KORNET, and KT is the largest network operator in Korea. There are 11 regional nodes in KORNET. To examine the delay characteristics of the KORNET, measurements were performed at 6 regional offices, some of which are at regional nodes others are at branch nodes. At the same time with delay measurement, packet loss was also monitored. Observations were made for tens of hours consecutively to include both peak period and off-peak period of a day.

To measure end-to-end delay, we utilized measurement equipment called DQME(Distributed Quality Measurement Equipment). 82 DQMEs are installed on KORNET and they are geographically distributed over regional offices or branch offices of KT. They are placed under the environment of ADSL-Lite subscribers to measure network performances from end-users' perspective; Each DQME is located 2Km apart from DSLAM to emulate real situation of ADSL subscribers. The line speed of the ADSL-Lite is 2Mbps.

We wrote a small program to build a tool measuring RTT between source and destination based on ping command. We adapted various packet sizes of ping, from 32 bytes up to 1380 bytes, to check the relationship between delay and packet size.

measurement architecture

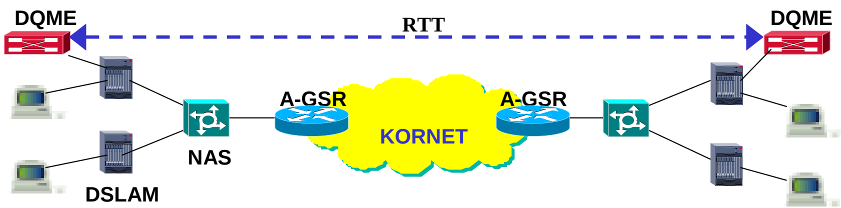


Figure 1. Configuration of end-to-end delay measurement

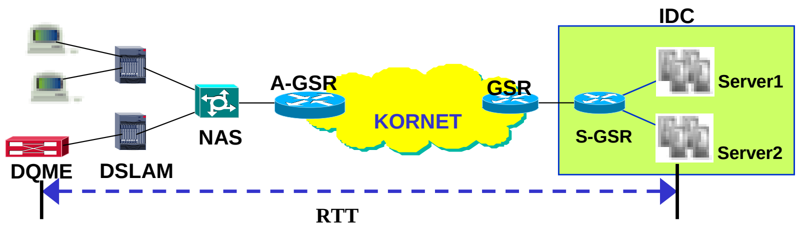


Figure 2. Configuration of end-to-server delay measurement

ETRI 4

2.2 End-to-end delay measurement

The configuration of end-to-end delay measurement is shown at Figure 1. At each measurement location, we observed end-to-end delay between DQME of the office and every available DQMEs across the KORNET. So, the traffic traverses 9 or 10 hops including AGSRs(Access GSRs) and backbone GSRs. The routing path of end-to-end interval is; DQME-DSLAM-NAS-AGSR-GSRs-AGSR-NAS-DSLAM-DQME. The hop count of each DQME pair varies according as how many GSRs the packet traverses.

Packet sizes of ping are 32, 64, 128, 256, 512, 1024 and 1380 bytes. Observations were taken for a few days at each measurement location, and more than 450,000 times in total.

2.3 End-to-server delay measurement

To examine end-to-server delay, we measured RTT from DQME to the main server of KORNET, www.kornet.net. The KORNET servers are located at the CDC(central data center) of KT and we picked one of them as a target server.

As shown at Figure 2, the routing path of whole interval was 7 or 8 hops including SGSR(server GSR); DQME-DSLAM-NAS-AGSR-GSRs-SGSR-Server. Packet sizes of ping are same as in the case of end-to-end delay measurement, and observations were taken for a few days at each location, more than 60,000 times in total.

End-to-end delay of KORNET-ADSL

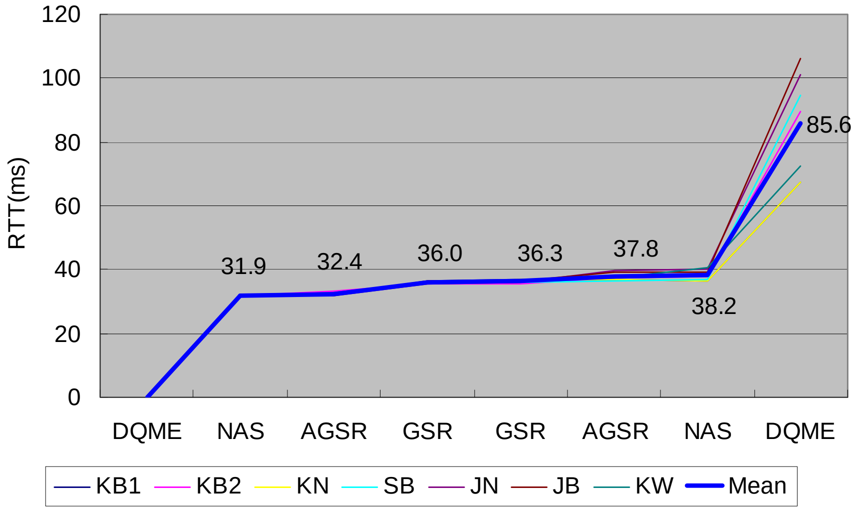


Figure 3. End-to-end delay across the routing path(TK;1380bytes)

ETRI 5

3 Results and Analysis

3.1 End-to-end delay

3.11 End-to-end delay of the KORNET ADSL

Figure 3. shows the result of end-to-end delay of the KORNET ADSL, measured at TK office, the packet size of ping is 1380 bytes. Through the 9 hops, average delay was 85.6ms while average packet loss ratio was 0.22%.

79.7ms, around 93% of total delay, was due to the edge part(between DQME and AGSR; last mile and access interval) and only 5.4ms was due to the core part(from AGSR to AGSR). Summarizing the results of all the measurement offices, we have found that 93.5% of total delay was due to edge part. This implies that, for the network performance of end-to-end interval, the most critical and significant bottleneck is at the edge part in case of current KORNET.

It can be recommended that, for network operators to enhance their network performance, it should be examined more carefully that what factors are affecting delay performance in the last mile and access network. Since routing processors handle ping packets, i.e., ICMP packets, in different ways from that of normal user data packets, additional measurement schemes should be designed for in-depth analysis to grab the real characteristics of delay over network edge interval.

The collected data of delay measurement of this study has quite small variances; in most cases, the CoVs(coefficient of variation) of the measurement fall on the interval of 0.04 to 0.15

End-to-end delay by packet size

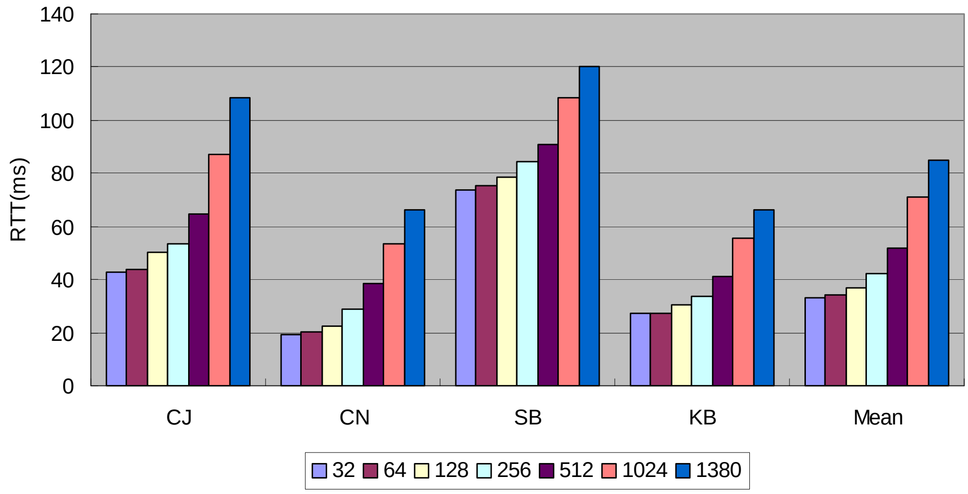


Figure 4. End-to-end delay by ping packet sizes(CB)

ETRI

6

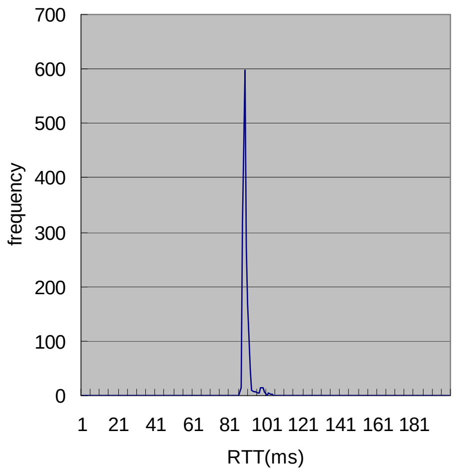
3.12 End-to-end delay by ping sizes

To examine the effect of packet size on delay performance, we varied the size of ping packets. From 32 bytes to 1380, we chose 7 sizes randomly. We found that end-to-end delay varies according to the packet size of ping. The larger the packet size is, the bigger the delay. For the paths from CB-DQME to other DQMEs, the end-to-end delay of 32 byte packet size was 33.4ms on average while 1380 bytes, 85.0ms. That is, the delay of 1380 byte ping packets was found to be 2.5 times greater than that of 32 byte ping packets. Even if we consider packetizing delay, the difference is greater than expected. With small ping size, 32 and 64 bytes, the end-to-end delay grows slightly but when it comes to over 512 bytes the delay increases rapidly.

We infer that this variance of delay performance by packet size can be caused by the way the packets are handled at each node they traverse; routing policy or processing power of node systems can affect the delay.

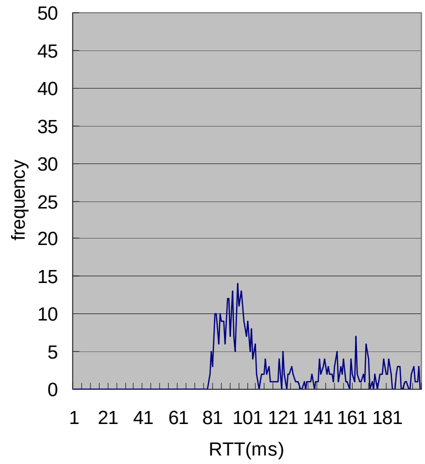
Since ping packet is ICMP packet, we don't know if the same rule can be applied to the real situation of general data packets. If the relation is systematically uncovered, NSPs or ISPs could optimize their network or service with regard to packet size.

End-to-end delay distribution



Packet loss ratio = 0.27%

Figure 5. Gamma-like shape with heavy tail
(TK to KB;1380bytes)



Packet loss ratio = 0.25%

Figure 6. Many peak
(CB to SB;1380bytes)

ETRI 7

3.13 End-to-end delay distributions

It is well known that the most typical delay distribution of the Internet is Gamma-like shape with heavy tail[2,5]. The result of this study shows that 91.6%(77 out of 84 cases) of end-to-end delay of the KORNET ADSL services follow the typical distribution. An example is shown at Figure 5.

As seen at Figure 6, some intervals appear very unstable. From the RIPE project, it was argued that this kind of instability may cause high ratio of packet loss[2]. The result of this study, however, shows no significant difference in packet loss ratio between stable cases and unstable cases; the packet loss ratio of Figure 6 case was 0.27% while that of Figure 7, 0.25%.

The unstable delay distribution could be resulted from other factors such as performance of node systems, network status, and so on. Especially, traffic is prone to be congested at network edge part to over-burden node systems of access interval. We infer that, for the case of instability seen at Figure 6, destination-NAS has to process so many in/out packets that downlink at destination part was unstable. Bovy et al.(2002) suspected that an overloaded router caused many peaks distribution[2].

End-to-server delay of KORNET-ADSL

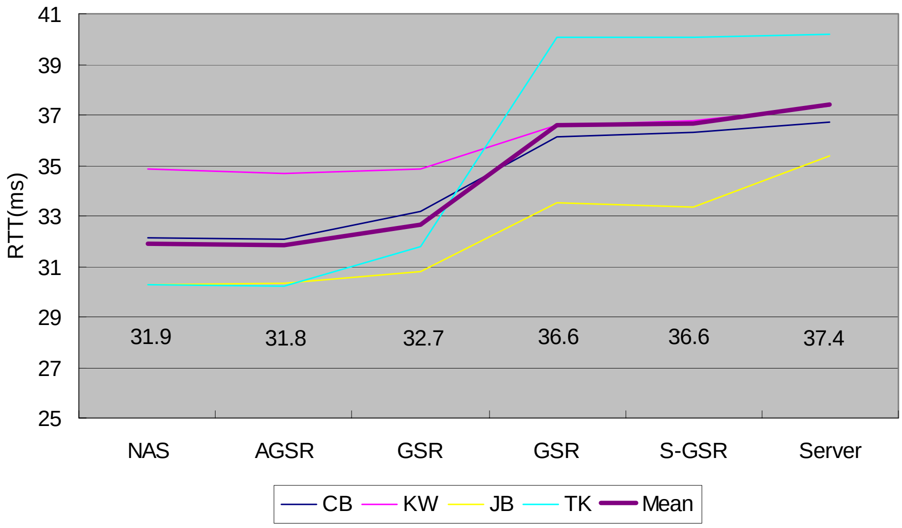


Figure 7. End-to-server delay across the routing path

ETRI 8

3.2 End-to-server delay

3.21 End-to-server delay of the KORNET ADSL

Average end-to-server delay of the KORNET ADSL was lower than 40ms while average hop count was 7.5 and average packet loss ratio was 0.12%.

Figure 7 shows the end-to-server delay at several locations where hop count is 7 and ping size is 1380 bytes. Similar to the result of the end-to-end delay, most of the delay is due to edge part(between DQME and AGSR). On average, the delay across the whole path was 37.4ms while the delay of edge part was 31.8, which is about 85% of total delay.

It should be noted that the delay in the server interval, from GSR to server, was only 0.8ms. So, it can be said that the main server of KORNET, which is located at the central data center, shows reasonable delay performance. Of course, the server reaction time for subscribers will be far much longer than the ACK time for ping packets.

End-to-server delay by packet size

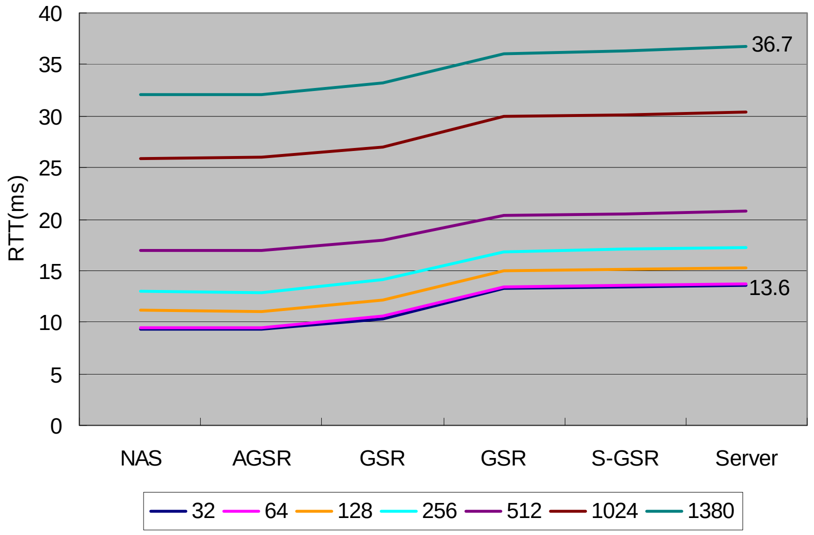


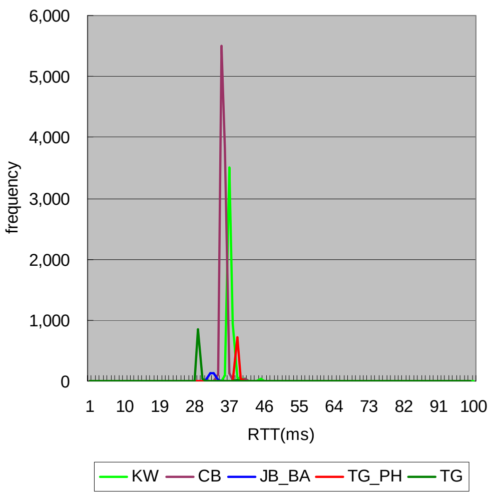
Figure 8. End-to-server delay by ping packet sizes(CB)

ETRI 9

3.22 End-to-server delay by packet sizes

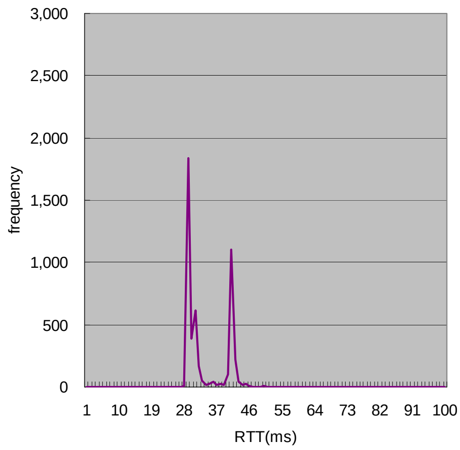
End-to-server delay performance by ping packet sizes brought in similar results as well. On average, in the case of CB, the end-to-server delay with 32 byte ping packets was 13.6 while 1380 bytes, 36.7ms. Exact delay time for data packets can be changed by the way the server handles that packets. Anyway, we can say that the end-to-server delay performance depends on the unit packet size of the traffic the server transfers.

End-to-server delay distribution



Packet loss ratio = 0.07%

Figure 9. Gamma-like shape with heavy tail (1380bytes)



Packet loss ratio = 0.01%

Figure 10. 2 Gamma-like shape (KW;1024bytes)

ETRI

10

3.23 End-to-server delay distributions

In case of end-to-server delay distributions, 17 cases out of 20, which is about 85.0%, appeared to be Gamma-like shape with heavy tail. Some examples of the cases are shown at Figure 9. Other cases showed two Gamma-like shape as seen at Figure 10. For this case, we found that alternative routing paths to the server result in the overlapping of two Gamma distributions.

Delay distribution of many peak shape was not shown in the end-to-server delay. We can regard this result as a desirable one since it means the delay characteristics of the network to the server is quite stable.

End-to-end delay vs. network utilization

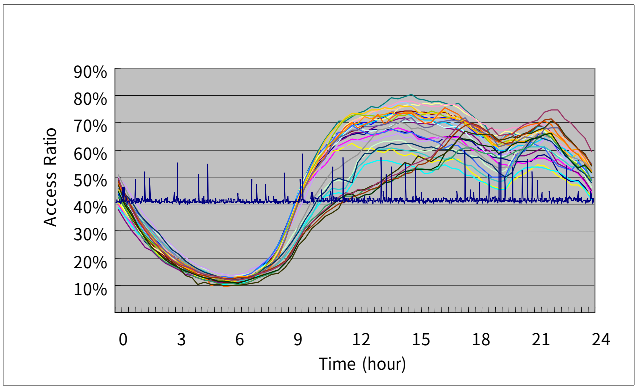


Figure 11. Relationship between packet delay and network utilization

ETRI 11

3.3 Relationship between packet delay and network utilization

Figure 11 shows an example of relationship between packet delay and network utilization at the JB_BA office. Owing to data availability, we used ‘access ratio’ as a network utilization factor. Access ratio refers to how many IP addresses are assigned from NAS IP pool at a certain time. The access ratios at the graph are daily averages of one month of measurement.

Applying several statistical tests on the result, we have found no case which has significant relationship between two factors. It can be argued that access ratio has not much effect on delay performance. However, since 90% of access ratio does not mean 90% of network utilization in the access interval, further study is required. Mochalski et al.(2002) reported that shortage of bandwidth is a major reason for increased delay in access path[7].

Summary & Future works

- The bottleneck is at last mile and access network
- Packet delay can be varied severely by packet size
- Unstable delay distribution can result from NAS or access router problems
- Experiment path should be extended
  - Domestic and foreign ISPs
- More extensive metrics are required
  - IPDV
- Protocol specific delay should be measured
  - Application specific performance measurement

ETRI

12

4 Summary and future works

We examined the delay characteristics of KORNET ADSL services. The main results of our research and their implication could be summarized as followings.

First, the large portion of packet delay occurred at the edge part of network, especially at the end-to-NAS interval. So, the key to enhancing delay performance is in the edge part of network, not in the core part; last mile and access network should be managed more carefully.

Secondly, packet delay is affected by packet size. If the relationship between delay performance and packet size is well uncovered for real data packets, delay-sensitive applications can be designed so that resulting delay does not exceed required level.

Thirdly, delay distributions followed typical Gamma-like shape with heavy tail at most of end-to-end intervals. At only a few intervals the distributions showed unstable many peak shape. The main reason of unstable distribution can be the status of NAS or access router. By putting monitoring function into QoS management system, NSPs can make their network more stable by checking the delay distributions of each interval. If it is alarmed that a certain interval shows unstable delay distribution, the status of the node systems in that interval should be checked.

Finally, access ratio has less influence on delay performance. To find out the relationship between network utilization and packet delay, the real data for bandwidth usage are required.

The measurements of this study was done under ADSL user environment. But to our sorry, the packets traversed only one network. Experiment path should be extended to the subscribers or servers of other ISPs including foreign ISPs to characterize inter-domain delay and international delay. Also, it is desirable to measure more extensive performance metrics like IPDV since the Internet users demand more and more real-time applications.

There are so many types of applications on the Internet and moreover, each application has different performance objectives. To examine application specific network performance, protocol specific performance should be studied in future works as well.

References

[1] M. Alves et al., “New Measurements with the RIPE NCC Test Traffic Measurements Setup,” Proc. PAM 2002 (2002)

[2] C.J. Bovy et al., “ Analysis of End-to-end Delay Measurements in Internet, ” Proc. PAM 2002 (2002)

[3] k. claffy et al., “The nature of the beast: Recent traffic measurements from an Internet backbone,” Proc. INET '98 (1998)

[4] A. Fei et al., “Measurements on Delay and Hop-Count of the Internet,” Proc. IEEE GLOBECOM'98 (1998)

[5] G. Hooghiemstra, P. Van Mieghem, “Delay distributions on fixed Internet paths,” Report TU-Delft.

[6] S. McCreary and k. claffy, “Trends in wide area IP traffic patterns - A view from Ames Internet Exchange,” in ITC Specialist Seminar (2000)

[7] K. Mochalski et al., “Packet Delay and Loss at the Auckland Internet Access Path, ” Proc. PAM 2002 (2002)

[8] K. van der WAL et al., “Delay Performance Analysis of the New Internet Services with Guaranteed QoS,” Proc. IEEE 85 (1997)

[9] V. Paxson, “End-to-End Internet Packet Dynamics”, IEEE/ACM Trans. on Networking, Vol.7, No.3 (1999)

[10] ITU-T SG13, “Network Performance Objectives for IP-Based Services,” Y.1541