Available online at www.sciencedirect.com
ScienceDirect
Computer Communications 29 (2006) 3113–3121
computer communications
www.elsevier.com/locate/comcom

Evaluating IPv6 on a large-scale network

Wen-Lung Shiau [a,*], Yu-Feng Li [b], Han-Chieh Chao [c], Ping-Yu Hsu [d]

[a] *Department of Information Management, Ming Chuan University, No. 5, Teh-Ming Rd., Gwei-Shan District, Taoyuan, County 333, Taiwan, R.O.C*
[b] *Computer & IT Center at National Dong Hwa University, Hualien, Taiwan, R.O.C*
[c] *Department of Electrical Engineering, National Dong Hwa University, Hualien, Taiwan, R.O.C*
[d] *Department of Business Administration, National Central University, Chung-Li, Taiwan, R.O.C*

Available online 9 January 2006

Abstract

We evaluate an ideal model and a real large-scale network environment using available end-to-end measurement techniques that focuses on a large-scale IPv6 backbone and made performance comparisons between the current Internet (IPv4) and next generation Internet (IPv6). In this paper, we compiled the performance statistics of each network in terms of TCP and UDP throughput, delay jitters, packet loss rate, and round trip time. Our conclusions show that, in a real large-scale network environment, a minor degradation in the throughput of the TCP, a slightly higher throughput of the UDP, a somewhat emerging frequency of the delay jitter, a lower packet loss rate, and a slightly longer round trip time happens when we compare the IPv6 network to the IPv4 network.

Keywords: IPv4; IPv6; Performance measurement; End-to-end performance

1. Introduction

More and more enterprise clients replace their mission-critical applications on the Internet in favor of cheaper and more ubiquitous Internet networks. For these applications, business enterprises are evaluating Internet networks by their ability to meet much more rigorous standards of availability and performance [18]. Today network backbones offering bandwidths in excess of 1000 Mbps with very low transmission error rates are becoming widespread. The capacities of IP backbones are growing quickly to meet customer requirements [18]. But sometimes users still have to wait for information longer than the desired time and endure unsatisfactory performance.

Performance is a key factor in the development and implementation of modern computer systems and networks [3]. From general users to advanced researchers, each user seeks minimum service time. Since service time is composed of: (a) local hardware sending requests, (b) network transmitting time, and (c) remote hardware receiving requests, end-to-end performance evaluation becomes a very crucial part in network assessment.

In order to understand the Internet performance of users, more and more researchers evaluate end-to-end performance from different viewpoints. Krishnamurthy and Wills [1] analyzed factors that influence end-to-end measurement of Web performance. They examined the components of delay or the effectiveness of the recent changes on the HTTP protocol. They found *pipelining* as an improvement over existing practice, but concluded that servers serving a small number of objects or closing a persistent connection without explicit notification can decrease or eliminate any performance improvement. ElAarag and Bassiouni [4] measured end-to-end performance of TCP connections in an ideal and non-ideal network environment. In their ideal model, they provided an upper bound limit for the throughput and a lower bound limit for the transfer time of the TCP connection. For the non-ideal environment, the simulation results show the relative performance of four standard TCP implementations. Ferrari [16] measured end-to-end performance analysis with traffic aggregation. The effect aggregation is fundamental in evaluating the capability of a *diffserv* network to preserve

* Corresponding author. Tel.: +886 3 350 7001; fax: +886 3 359 3875.
E-mail address: mac@mcu.edu.tw (W.-L. Shiau).

0140-3664/$ - see front matter
doi:10.1016/j.comcom.2005.11.011

the original stream profile, in particular, the delay and jitter-sensitive traffic. The author developed a study of priority queuing in detail and compared the performance with weighted fair queuing. Zhang and Zheng [9] measured end-to-end performance for the IPv6 traffic model with multiple qualities of service (QoS) classes in virtual private networks (VPN). They presented the performance trade-off between the delay sensitive traffic and delay insensitive traffic in terms of traffic throughput, packet loss probability and end-to-end delay in VPN networks.

The next generation Internet protocol version (IPv6) was designed to replace the current Internet protocol (IPv4) and cope up with the world's needs in the future. Today the next generation Internet (NGI) has already been deployed worldwide, but there have been no publicly available end-to-end performance measurement that focuses on a large-scale IPv6 backbone. The reason is because the chance is scare and the time period for performing the measurement is very limited. The best time to measure the performance is in the duration when the network is completed and no traffic has traveled on the network yet. Traffic on the network may skew the measurement. The paper publishes a result measuring the performance of TWAREN in the precious time period.

IPv6 end-to-end performance tests are needed when IPv6 is moving towards commercial usage with high quality service requirements. The purpose of this paper is to evaluate a large-scale IPv6 backbone and make a performance comparison between current Internet (IPv4) and next generation Internet (IPv6). The rest of the paper is organized as follows: Section 2 reviews related works. Section 3 presents our test bed and measurement procedures. Section 4 then follows with the results and discussions of this study. Finally, on Section 5, we have our final conclusions.

2. Related works

Performance measurement is valuable because it provides historical data on how the network is fulfilling its objectives as well as revealing performance or configuration problems proactively [11]. Today the relocation from the IPv4 network to the IPv6 network is happening quite gradually. This probably means that dual stack transition mechanism will exist for a while. Performance evaluation for both the IPv4 and IPv6 networks are necessary because we want to know how each network will execute under the dual stack transition mechanism. The protocol stack is expected to have a definite impact on end-to-end performance of the final system. More and more researchers have published performance comparisons between IPv4 and IPv6 protocol stacks at the end-system.

Draves and Zill [12] presented a performance evaluation on IPv6 network in comparison with IPv4 results on Windows NT using a Fast Ethernet adapter. The results show that throughput is lower by about 2% in IPv6 cases compared to IPv4. The author conducted tests only on throughput in a prototype IPv6 stack and did not perform detailed parameter checks such as packet size testing on TCP and UDP for a real large-scale IPv6 network. Anand [10] measured network TCP stream throughput performance on an IPv6 network and compared them with IPv4 results on Linux using a gigabit Ethernet adapter. The author used different message sizes to measure the TCP stream throughput on both the server and the client. The results show throughput is higher in IPv4 cases compared to IPv6. That is, IPv6 is not driving the system hard enough to achieve greater throughput. According to the profiling analysis, the author found out that the IPv6 network is doing a check-sum of the data in the software even though the data is already check-summed by the hardware. He, however, conducted only TCP tests by simulation and did not compare the UDP of IPv4/IPv6 stacks for real IPv6 networks. Zeadally and Raicu [14], evaluating IPv6 on Windows and Solaris, compared IPv4 and IPv6 networks using the above-mentioned operating systems. They measured throughput, round-trip time, CPU utilization, and other connection characteristics. The experimental results show that IPv6 protocol stacks for Solaris outperform IPv6 stacks of the Windows operating system, while the IPv4 protocol stacks outperform IPv6 stacks on both TCP and UDP. Ariga et al. [13] assessed the performance of large data transmissions and applications such as digital videos with various security protocols over both IPv4 and IPv6 networks. The authors used ordinary personal computers (PC) to simulate routers and end-hosts with *FreeBSD2.2.8* and a *KAME* [7] IPv6 protocol stack. The results show that the typical PC can handle digital video transmissions with *IPSec* over the IPv6 network. The authors conducted only TCP and UDP streams and did not perform detailed testing for a real large-scale IPv6 network. Karuppiah [2] analyzed the performance of IPv6 and IPv4 networks using the *Ping* utility and *FTP* applications. The author used the *Ping* utility to find the latency and *FTP* applications to find out the throughput rates over the IPv6 and IPv4 networks. He used a PC to simulate the routers and end-hosts with *FreeBSD* and a *KAME* [7] IPv6 protocol stack. The results show the IPv6 network has an inferior performance compared to IPv4 for the files transferred. He did not experiment with detailed parameters such as packet size on TCP and UDP, connection time or protocol type (since they could not perform any UDP tests due to the nature of *FTP*).

Zeadally et al. [15] appraised IPv6 performance on Windows, Solaris, and Linux. The authors measured throughput of TCP and UDP, latency, CPU utilization, and other web-based performance characteristics. Their investigation shows that the IPv6 protocol stacks for Linux outperforms the IPv6 stacks of the other operating systems. Loiacono et al. [6] measured a worldwide network performance of NASA's Earth Observation System. They developed a Web-based network-monitoring tool *ENSIGHT* to detect and troubleshoot performance problems. The results show a comprehensive view of relevant performance parameters. Although the authors conducted a very large network performance measurement, they only tested the IPv4 network and did not perform any IPv6 performance measurements.

There are three relevant evaluation techniques: analytical modeling, simulation, and measurement experiment [8]. Analytical modeling involves the creation of a mathematical model, which describes the basic characteristics of a network system. Simulation is a flexible technique that is used to replicate almost any desired system behavior. The results, however, only show the behavior in controlled environments. Measurement experiment provides a collection of raw data, which is then used in analyzing the characteristics of the targeted system. For end-to-end performance evaluation, we did not focus on the mathematical portion of analytical modeling. Most researchers actually prefer simulation techniques because with it, it is easier to control the environment as compared to measurement experiments. Based on the study of previous researches, we made some comparisons between previous studies conducted along this theme as Table 1.

This work differs from the previous efforts of other researchers in that we performed a performance evaluation on a real, large-scale network backbone and made a comparison between the current Internet (IPv4) and next generation Internet (IPv6).

3. Measurement procedures and test bed

For the purpose of this study, we followed the lead of ElArag and Bassiouni [4]. In the first stage, we measured end-to-end performance in an ideal network environment to provide upper bound limits for throughputs and lower bound limits for latency. Two identical workstations were connected using a point-to-point link of gigabit Ethernet as projected in Fig. 1.

Both workstations were each equipped with an Intel Pentium 2.8 GHz processor, 1000 MB of DDR RAM, 36 Gb SEAGATE ALTRA 320 SCSI 10,000 RPM hard drives, and 1000 Mbit/s PCI64 Ethernet network adapters. The workstations were each loaded with the Linux Fedora Core II (kernel version 2.6.5-1) operating system.

In the second stage, we measured end-to-end performance in a real large-scale network including IPv6 and IPv4 backbones to get real end-to-end throughput and other characteristics. Two identical workstations were each connected to Cisco 3750 gigabit switches connected to

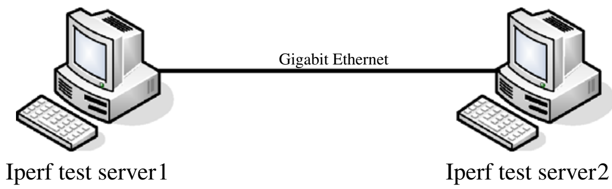

Fig. 1. An ideal network environment.

Cisco 7609 routers as shown in Fig. 2. Taiwan Advanced Research and Education Network (TWAREN) is a next generation research and education network. TWAREN has the bandwidth of 10 Gbps and is based on the technology of Dense Wavelength Division Multiplexer (DWDM) technologies. There are eleven regional centers serve to get all major institutions of higher education and research connected to the backbone. Every institution will connect to the backbone at 1 Gbps [17]. In the future, total of more than one hundred and fifty institutions will connect to TWAREN. The workstations have similar infrastructure as assembled in the ideal network environment. Each workstation were equipped with an Intel Pentium 2.8 GHz processor, 1000 Mb of DDR RAM, 36 Gb SEAGATE ALTRA 320 SCSI 10000 RPM hard drives, and 1000 Mbit/s PCI64 Ethernet network adapters. The workstations were each installed with Linux Fedora Core II (kernel version 2.6.5-1) operating systems.

The measurement tool which we used is *Iperf* [5] (available from the National Lab for Advanced Network Research (NLANR)) because of its rich set of features. An older program *tcpwatch* has been mostly phase-out already [6]. Another famous tool *ttcp* is very old and possesses confusing options [5]. *Iperf* was developed as a modern tool to measure maximum TCP bandwidth, allow the tuning of various parameters such as UDP characteristics, report bandwidth, delay jitter, and datagram loss. We also used *Ping* to measure round trip time. After we did the performance evaluation, we show the results in the succeeding section and make comparisons between the current Internet (IPv4) and next generation Internet (IPv6).

4. Results and discussions

In this section, we present and discuss the results obtained from our tests for an ideal model and a real, large-scale network.

Table 1
Comparisons between previous studies and this research

Researches	Simulation or measurement experiment	Scale	IPv4 TCP /UDP throughput	IPv6 TCP/UDP throughput	IPv4 latency	IPv6 latency
Draves and Zill [12]	Simulation	Small	Only TCP	Only TCP	Yes	Yes
Anand [10]	Simulation	Small	Only TCP	Only TCP	No	No
Zeadally and Raicu [14]	Simulation	Small	Both	Both	Yes	Yes
Ariga et al. [13]	Simulation	Small	Both	Both	No	No
Karuppiah [2]	Simulation	Small	Only TCP	Only TCP	Yes	Yes
Zeadally et al. [15]	Simulation	Small	Both	Both	Yes	Yes
Loiacono et al. [6]	Measurement experiment	Large	Both	No	Yes	No
This research	Measurement experiment	Large	Both	Both	Yes	Yes