Regular Paper

Evaluations of Web Server Performance with Heavy-tailedness

Takuo Nakashima[†]

Providing quality of service (QoS) in a business environment requires accurate estimation of Internet traffic, especially HTTP traffic. HTTP traffic, mainly consisting of file transmission traffic, depends on the sizes of files with a heavy-tailed property. Analyzing the performance of specific Web servers and comparing server performance is important in QoS provisioning for user applications. In this paper, we present the experimental analysis of Web server performance using our active measurement. We capture activity data from 1984 diverse Web servers by sending measurement packets 20 times in order to evaluate the spatial properties, and we select 14 Web servers to send packets 2,000 times in 5-second intervals to evaluate the temporal properties. The main contribution of our study is to provide methods of evaluating the temporal and spatial properties on any Web server by measuring from a remote observation host, and to illustrate the current activity of temporal and spatial properties in a series of figures. Crovella's well known work merely describes the self-similar properties for the total transmission time generated by a heavy-tailed file size distribution. We divided the total transmission time into network and server system dependable elements, of which the heavy-tailed properties are captured. We found that temporal properties consist of two factors: stability and activity in the Poisson process. Robustness of heavy-tailedness in terms of constructing elements of total transmission time was evident from the analysis of spatial properties. Finally, we observed the upper boundary and classified groups in a mean-variance plot of primitive elements of the Web server.

1. Introduction

Providing quality of service (QoS) for end users requires accurate estimation of the amount of Internet traffic as well as analysis of traffic elements. Currently, the trend in Internet traffic is shifting to peer-to-peer traffic; however, web-based technology will continue to increase, especially in the business environment. Since the seminal study of Leland, et al.[1], self-similarity of network traffic has been widely adopted in the modeling and analysis of network performance. Self-similar traffic satisfies the following two properties: firstly, the correlation structure retains its status with respect to time aggregation; secondly, the autocovariance function decays hyperbolically. This means that long-range dependence is led by heavy-tailed traffic. The relationship between file size and self-similar traffic was explored by Park, et al.[2]. Crovella, et al.[3] surmised that the file size of Web servers had heavy-tailed distribution that might cause self-similarity in Web traffic, i.e., total data transmission. This manifestation of the application layer causes self-similarity at multiplexing points in the network layer in the form of self-similar traffic[4]. In conjunction with this evidence, properties of

heavy-tailedness appear in computer systems and Internet traffic. File size distribution in UNIX is heavy-tailed[5], and UNIX processes have been observed to possess heavy-tailed lifetimes[6]. Paxson, et al.[7] found that the upper tail of the distribution of data bytes in FTP bursts was well fitted to a Pareto distribution with $0.9 \leq \alpha \leq 1.1$. Park, et al.[4] also found through simulation that the presence of self-similarity at the link and network layer depends on whether reliable and flow-controlled communication is employed at the transport layer.

Our purpose in this research is to determine the characteristics of Web servers in order to evaluate their systems working with the heavy-tailed property described above. Each Web server is accessible from all over the world, and operates differently in various time series. Our proposal in both temporal and spatial property analyses covers evaluations of time series for specific Web servers as well as performance evaluations for a large number of Web servers. Valuable tools exist to measure these performances. For example, Httperf[8] has been widely used to evaluate Web servers from a remote site. However, it is necessary to customize this tool in order to probe time data that is related to the TCP connection and the initial processing time. On the contrary, SPECweb[9] focuses on testing high performance servers using

† Kyushu Tokai University

a number of benchmarks. Performance evaluation functions, however, are limited to special purpose Web servers. To analyze temporal and spatial properties of Web servers, we have implemented performance evaluation programs using the active measurement method. Our motivation is to clarify the following points: first, what kind of distribution are the temporal properties dependent on; and second, how does each element of Web performance contribute to heavy-tailed distribution?

In this paper, we present the experimental analysis of Web server performance with heavy-tailedness using our active measurement method. This method collects sending and receiving timestamps on the measuring host to estimate the remote processing and transmitting time on Web servers. We did extensive measurements of Web servers and analyzed timing data based on temporal and spatial property evaluation procedures. We found that the two key factors of the temporal properties are stability and their behavior in the Poisson process. Robustness of heavy-tailed distribution was evident from the analysis of spatial properties in respect to constructing elements of the total transmission time. Finally, we observed the upper boundary and classified groups on a performance evaluation plot of primitive elements independent of network distance and file size.

This paper is organized as follows. First, the definition of heavy-tailed distribution is described in Section 2. In Section 3, validation of the measurement method used in the experiments is given. Using this method, we probed a large number of Web server performances. Section 4 is the discussion of temporal property and spatial property with heavy-tailedness; then the performance evaluations of Web servers are explained in Section 5. Section 6 is the summary and discussion of future work.

2. Heavy-tailed Distribution

The basic concepts related to heavy-tailed distribution will be introduced shortly; a more detailed treatment can be found in Park, et al. [4].

First, a definition of heavy-tailed distribution is given, followed by a comparison of the distribution drawn from a Poisson process.

A random variable X has a heavy-tailed distribution if

$$P[X > x] \sim cx^{-\alpha}, 0 < \alpha < 2, \quad (1)$$

for some positive constant c, where $a(x) \sim b(x)$ means $\lim_{x \to \infty} a(x)/b(x) = 1$, and α is called the tail index or shape parameter. Regardless of the behavior of the distribution for small values of the random variable, if the asymptotic shape of the distribution is a power law, it is heavy-tailed. If $P[X > x]$ is heavy-tailed, then X shows very high variability. Also, X has infinite variance, and, if $\alpha \leq 1$, X has infinite mean.

The simplest heavy-tailed distribution is the Pareto distribution. The Pareto distribution is a power law over its entire range; its probability density function is

$$p(x) = \alpha k^\alpha x^{-\alpha-1}, \alpha, k > 0, x \geq k, \quad (2)$$

and its cumulative distribution function is given as

$$F(x) = P[X \leq x] = 1 - (k/x)^\alpha. \quad (3)$$

The parameter k represents the smallest possible value of the random variable and is called the location parameter.

To assess the property of heavy-tailedness, we employ $\log - \log$ plots of the complementary cumulative distribution given as

$$\overline{F}(x) = 1 - F(x) = P[X > x] = (k/x)^\alpha, \quad (4)$$

for random variable X. Plotted in this way, heavy-tailed distributions have the property that

$$\frac{d \log \overline{F}(x)}{d \log x} = -\alpha, \quad x > \theta, \quad (5)$$

for some θ. To check the property of heavy-tailedness, we form $\log - \log$ plots and look for approximately linear behavior with slope $-\alpha$ in the tail. The heavy-tailed property causes the long-range dependence that is one of the properties of self-similarity by the restriction of the Hurst parameter (H) with $1/2 < H < 1$. Park [10] stated that heavy tails lead to predictability, and by relation, they lead to long-range dependence in network traffic; and furthermore, the Hurst parameter is related to the tail index by $H = (3 - \alpha)/2$.

On the other hand, if X is exponentially distributed, its cumulative distribution function is given by

$$\overline{F}(x) = P[X > x] = e^{-\lambda x}, \quad (6)$$

and has the property that

$$\frac{d \log_e \overline{F}(x)}{dx} = -\lambda. \quad (7)$$

When we employ log plots for the distribution function, the slope becomes linear.

3. Measurement Method of Experiments

3.1 Measurement Method

Currently, the tendency in Internet traffic is shifting to peer-to-peer traffic. However, Internet traffic is for personal and business use, and is also used to find on security holes, and we are interested in business-use traffic, especially Web server performance that should be discussed separately from multiplexing behavior in Internet traffic. In addition, we think Web-based technology will continue to increase in order to provide protection from security problems. All HTTP traffic is generated by people accessing diverse Web servers spread all over the world, and this consists of Internet traffic. In this research, we classify methods to evaluate Web server performance into two types: temporal evaluation and spatial evaluation. Temporal evaluation is based on data that is observed in a time sequence for a specific Web server, and it is used to evaluate the performance of a specific server. In contrast, spatial evaluation is based on multiplexed data for Web servers at a specific time. It assumes that accumulation of each traffic flow for Web servers at different times produces Internet traffic at a specific time. We examined both types of evaluation methods.

The measurement methods for Web servers are classified into two types, local and remote, based on the network distance between a measuring host and a measured host. Local measurement detects the performance of a local server without network delay, and evaluates local activity, e.g., available memory ratio and hard disk access ratio. This method is relevant when the efficiency of server applications on a specific OS and Web server are being addressed, and when performance metrics on different kinds of OSs and Web servers are being compared. Remote measurement, on the other hand, observes the performances of a remote server with a long network distance, and evaluates response activities, e.g., the time taken to establish a TCP connection, and the total file transfer time. Remote measurement can also measure a huge number of remote, diverse Web servers. As Web applications are currently put into practice with such different technologies as PHP (PHP Hypertext Preprocessor), ASP (Ac-
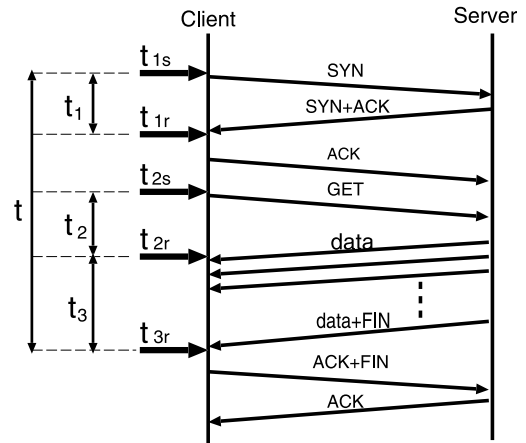

Fig. 1 Packet sequence.

tive Server Pages), JSP (Java Server Pages) and Servlet, remote measurement provides a uniform comparison method for remote users. In addition, metrics observed by remote measurement are close to the real activities of clients. These reasons have led us to adopt the remote measurement method.

3.2 Proposed Measurement Method

In these experiments, we implemented an original measuring program to collect the processing time at remote Web servers.

Figure 1 illustrates the packet sequence of the measuring program. First, the measuring host or client stamps the time (t_{1s}) before the SYN packet of TCP at the connection establishment phase; this is followed by the time (t_{1r}) that is stamped after receiving the SYN+ACK packet. Just prior to sending data, i.e., a GET request on HTTP, the measuring host stamps the time (t_{2s}), and then this host stamps the time again (t_{2r}) after receiving the first data from the Web server. Finally, the time (t_{3r}) is stamped after receiving the final data, eventually enabling us to obtain five different timestamps.

If we regard $t_1 = t_{1r} - t_{1s}$ as the round-trip time (RTT), and RTT does not change over a short period, we will consider $t_2 = t_{2r} - t_{2s}$ as the initial transmission time including RTT, and $t_2' = t_2 - t_1$ as the initial processing time. Here, t_2' is the proper metrics of the Web server and is independent of network distance. In addition, we regard $t_3 = t_{3r} - t_{2r}$ as the data processing time on the Web server, and $t = t_{3r} - t_{1s}$ as the total transmission time. Under the assumption that one packet constitutes 1,460 bytes, we propose the packet unit processing

time (t_3'). Both t_2' and t_3' are packet unit processing times, and exhibit proper system performance independent of network distance and file size.

When we precisely identify the meaning of the value of t_2' and t_3', the focus of consideration should be on the effect of protocol stacks and system architectures. The value t_2' includes the processing time of the GET request on the HTTP application layer, the response time for data packets on the application layer, and the generating time that is taken to compose the first packet in response. On the contrary, t_3' does not include the processing time of HTTP as a major factor, but rather the time to devide a data file on the TCP layer into packets, and to deliver it to the IP layer. The value t_2' mainly represents the performance of Web server applications, and t_3' represents the performance of the TCP/IP layer, operating system, and hardware.

3.3 Experimental Setup

We implemented a measurement program in C language on our client host on a FreeBSD 5.X platform with a Pentium 4.2-GHz processor on which an NTP (Network Time Protocol) stratum-2 server daemon keeps accurate time. First, we randomly selected 1,984 Web servers including domestic and foreign portal nodes as well as newspaper and broadcasting company nodes that seemed to have a high concentration of user access. This measurement program gathers text data from the top page of selected Web servers, and gets the five important HTTP timestamps over the TCP on our client host. Because our host belongs to the leaf domain of SINET, our measuring packets go through the SINET backbone and finally reach each Web server. Distribution of RTT, i.e., t_1, of these Web servers is shown in **Fig. 2**. In this figure, the x-axis indicates RTT partitioned into 20 millisecond non-overlapping blocks, whereas the frequency of Web servers is illustrated on the y-axis. These servers with RTTs less than 200 milliseconds are mainly categorized in domestic nodes, leaving the others as foreign nodes.

We conducted two experiments to characterize both temporal and spatial properties. For the spatial properties, we captured activity data from 1,984 diverse Web servers by sending measuring packets 20 times, and to evaluate the temporal properties we selected 14 Web servers to send packets 2,000 times in 5-second time
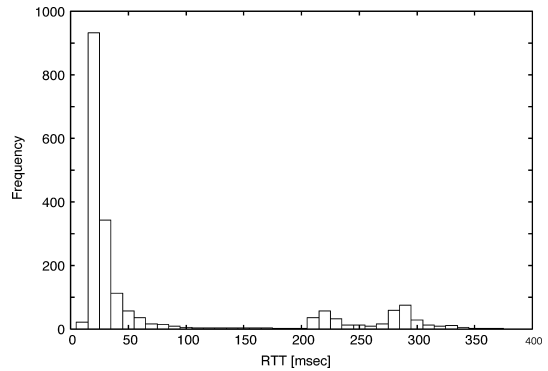

Fig. 2 RTT distribution of Web servers.

intervals.

The Web server performance on the specific Web server software, OS, and hardware specification was just a sample performance. Due to the diversity of Web servers, the server performance can hardly be formalized as generalized Web server performance by choosing one sample performance under limited circumstances. This type of performance should be analyzed in a different component, such as Web server software or OS, and then compared to other types of performance. In this paper, we focus on the total transmission time, which includes the Web server software performance and TCP/IP performance as a whole. In addition, we present a detailed discussion of each element of the total transmission time to extract the generalized properties of Web servers.

4. Properties of Web Servers

In this section, we focus on temporal and spatial properties, and discuss them in relation to heavy-tailed distribution.

4.1 Temporal Property

We will discuss how each Web server performs in a long sequential time series in this subsection. If we assume that the sample Web servers are identified as having average behavior, then Internet traffic would be similar to the accumulation of traffic in such sample Web servers.

Figure 3 illustrates complementary cumulative distribution of the total transmission time in four different servers that delegate different curvatures. The vertical axis in Fig. 3 is the $\log_{10}$ scaled value, whereas, the horizontal axis is a linear scale. Every curvature reveals two features: a straight drop and a linear decrease. The lines that drop straight down indicate that each server operates constantly to transfer the top page. We measured ten other servers sepa-
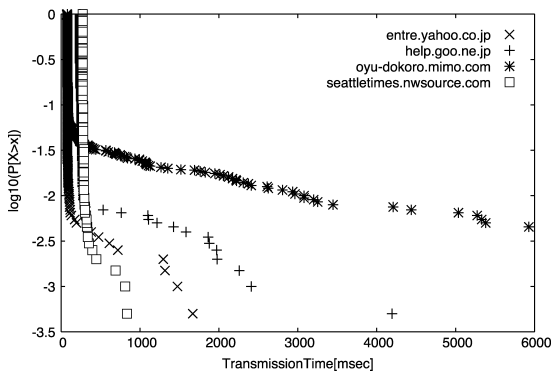
Fig. 3 Temporal property of delegated Web servers.

(a) Unix file system.

(b) Web server.

Fig. 4 File size distribution.

rately, which are not shown here, that operate similarly. The linear decreases appearing after the straight lines reveal that each total transmission time obeys an exponential distribution.
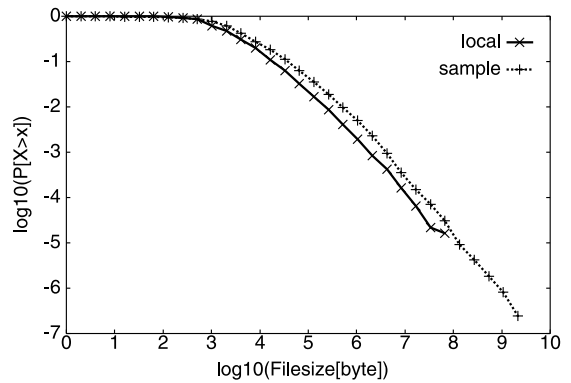
4.2 Spatial Property

We will also indicate how a huge number of Web servers perform individually in a short time period. If traffic in such diverse Web servers is generated at the same time, then Internet traffic would be multiplexed by these Web servers.

Firstly, we examine the file size distribution of UNIX systems and the top pages of Web servers, then discuss the distribution of the file transmission time in detail.
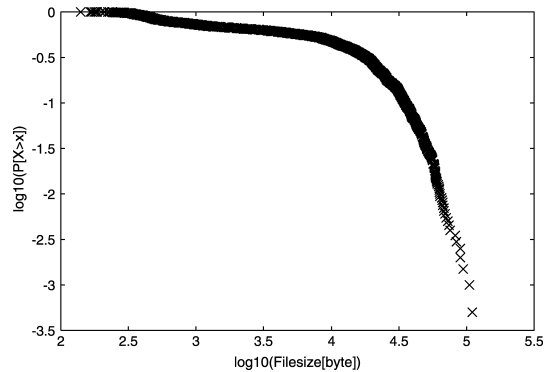
4.2.1 File Size Distribution

Figure 4 (a) shows the property of the UNIX file size as the log-log complementary cumulative distribution. The data labeled as "local" in this figure indicates the property of kernel data pertaining to FreeBSD 5.2.1 with a 40-Gigabyte hard disk. "Sample" data is from Irlam's work [5]. These two different UNIX file size distributions seem to be reasonably well modeled by a Pareto distribution with $\alpha = 0.91$. Considering the curvature in this figure, this distribution of file sizes is consistent with a hyperbolic tail for file sizes greater than 1,000 bytes. Both sets of data have no great divergence; they have not changed significantly for 10 years. This indicates that UNIX system programmers unconsciously tend to divide system files to be shaped in the heavy-tailed distribution.

We illustrate the complementary cumulative distribution of text file gathering by Web servers in our experiments depicted in Fig. 4 (b). Compared with the distribution in Fig. 4 (a), the file sizes were a concentration of values of those greater than 10 Kbytes and

smaller than 100 Kbytes, in which $\alpha = 3.00$. The distribution curve of this text file is similar to the curve of the study by Crovella, et al. [3], but the current distribution of text file size shows an earlier decline for small file sizes. This is because current trends are for top pages to have more visual content and links: these top pages mainly include small sized text data. Fujita, et al. [11] discussed performance modeling and evaluation of Web systems, and pointed out that the document size is proportional to the response time at the part of the document that is large, i.e., at the tail part. This means that a heavy-tailed property of the file size leads to a heavy-tailed property of the response time.

4.2.2 Robustness for Heavy-tailed Distribution

In this subsection, we focus on the tail part of distribution in terms of the total transmission time, and discuss how robustly heavy-tailedness is dominated by other factors in the total transmission time. We describe the robustness to reveal how other factors in the total transmission time contribute to making the distribution
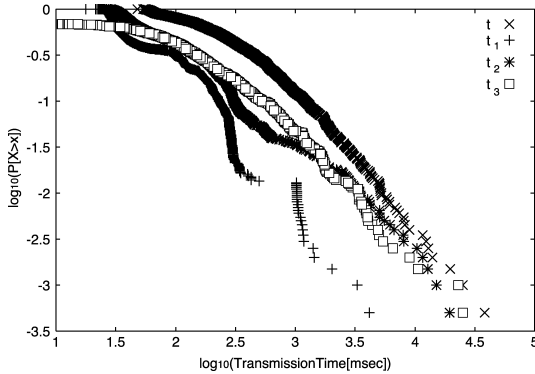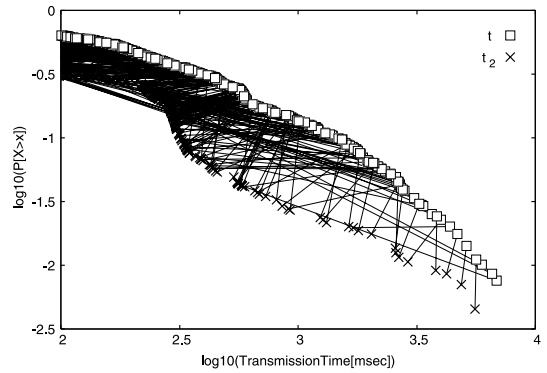
Fig. 5 Distribution of transmission times.

(a) Relationship between t and t_2.

(b) Relationship between t and t'_2.

Fig. 6 Relationship of initial processing time.

heavy-tailed.

Figure 5 plots the complementary cumulative distribution of t, t_1, t_2, and t_3. The figure implies that this distribution of t, t_2, and t_3 have the heavy-tailed property. The study by Crovella, et al.[3] obtained an estimate of $\alpha = 1.29$ for total transmission time, while the current distribution has a slope of $\alpha = 1.34$. Both slopes have a similar curvature, but current distribution declines one digit less than that of the old text files. This means that current Web servers respond more quickly to requests from Web clients.
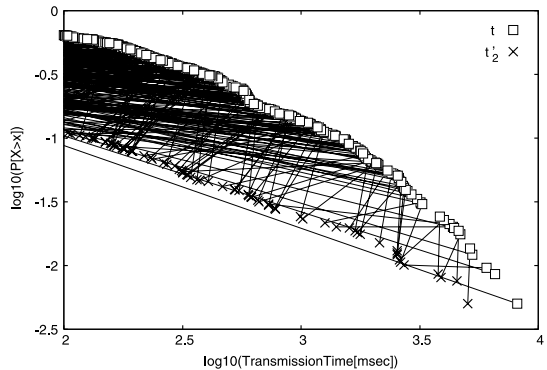
Compared to the study by Crovella, et al.[3], the file size distribution of top pages declines rapidly, whereas the distribution of t remains heavy-tailed, as Crovella, et al.[3] indicated. Apart from t, both t_2 and t_3 generate heavy-tailed distribution. This indicates that robustness exists in different processing times even though t_2 and t_3 do not have an explicit relationship. The distribution of t_1, i.e., RTT, depends on the network distance; thus, such distribution is integrated into two parts: domestic and foreign servers.

Using the proposed measuring program, we analyze sets of timestamps in detail to discover the effects of each element of total transmission time. In this analysis, we pay particular attention to the tail part of the total transmission time to exhibit the relationship for t_2, t'_2, and t_3 against t on the same Web server.

We have connected respective plots shown in boxes and crosses with straight lines for the data sets observed on the same Web server in Figs. 6 and 7. If the slope of a straight line is almost vertical, total transmission time shown by the box is mainly occupied by the other connected element.
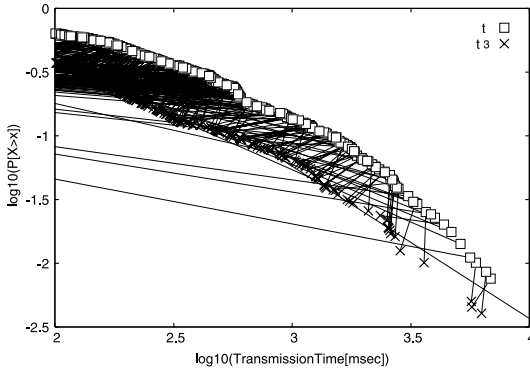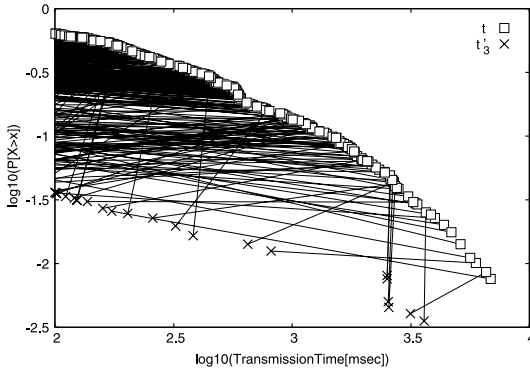
Figure 6 (a) displays the relationship between t and t_2. Lines connected to the tail part of t_2, shown by the cross marks, are almost vertical in this figure. These lines indicate that the heavy load of t_2 contributes to shaping the heavy-tailed distribution of t. Some lines connected to the tail part of t curve horizontally. This indicates that t is partially dominated by another element. Figure 6 (b) illustrates the relationship between t and t'_2 in eliminating the effect of RTT. Most of the tail part of t'_2 is connected vertically to the t on the same Web server. We also conclude that some t'_2 dominate t regardless of its relation with the file size or network distance. Both results indicate that heavy-tailed distribution of t is robust, since the tail part of t remains heavy regardless of whether t is dominated by the file size or initial processing time.

The other major element dominating total transmission time is the data processing time (t_3). **Figure 7** (a) plots the relationship between t and t_3 on the same Web server. The vertical connection line indicating the domination of t mainly with t_3 reflects the common
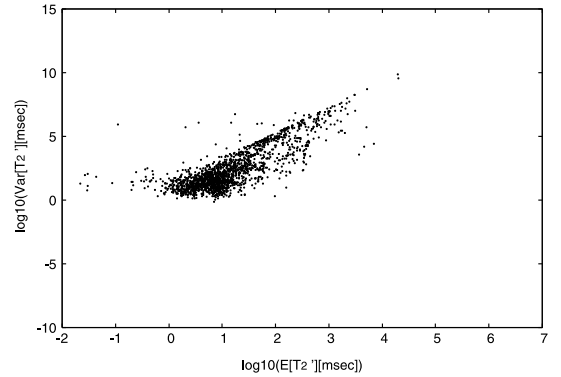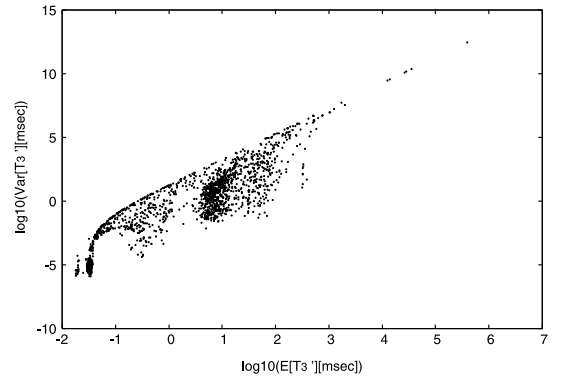
(a) Relationship between t and t_3.

(b) Relationship between t and t_3'.

Fig. 7   Relationship of data processing time.

(a) Distribution of t_2'.

(b) Distribution of t_3'.

Fig. 8   Mean-variance plot of each element.

acknowledgement that t depends on file size. Figure 7 (b) indicates that there is no clear relationship between the two metrics of t and packet unit processing time (t_3'). A number of flat slopes appear, especially in the tail part of t in Figure 7. Most flat slopes in Fig. 7 (a) indicate that t has been mainly established by t_2.

As a result of this evidence we can state firstly that the total transmission time obeys a heavy-tailed distribution; secondly, this causal mechanism is robust in respect that the initial setting time or the file transmission time can contribute to the heavy-tailed property. Each Web server has the variation that t is dominated by t_2' or t_3. In addition, the heavy-tailed property is maintained irrespective of different dominant elements of Web performance.

5. Performance Evaluations of Web Servers

In Section 4.1, we illustrated the temporal properties of typical Web servers and focused on the tail part of the heavy-tailed property as a spatial property; we also mentioned that heavy-tailedness of the total transmission time

was dominated by t_2' and t_3 in Section 4.2. In this section, we focus on the internal activity in Web servers, and illustrate both temporal and spatial properties in the same figure and discuss the relationship between two proper metrics. First, to eliminate the effects of file size, we selected the measuring metrics of two packet unit processing times, t_2' and t_3'. Second, we used a mean-variance plot to illustrate the current activity of a particular Web server as a temporal property on a map of the activity of multiple Web servers as a spatial property.

Figure 8 shows the mean-variance plot for each packet processing time, and both axes are scaled in log order. One plot illustrates the activity of one Web server for a short period. The mean values indicate how rapidly on average a Web server responded to a client request, and the variance values indicate how steadily a server responded to a client request. The upper boundary for variance appears in Figure 8 (a) and (b), indicating that any Web server converges to some level of stability. These values of t_2' converge to large values compared to those
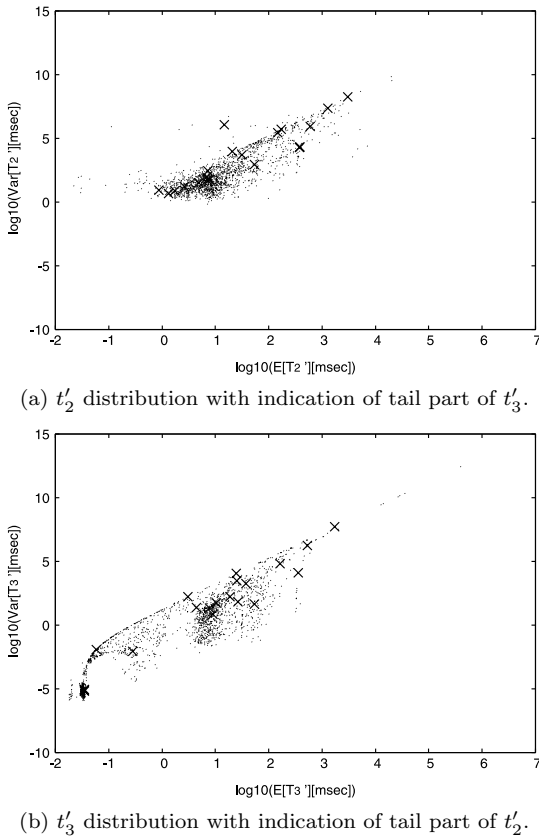
(a) t_2' distribution with indication of tail part of t_3'.

(b) t_3' distribution with indication of tail part of t_2'.

Fig. 9 Overlapping of tail part.

of t_3', which means that the initial processing consumes more time, while values of t_3' are classified in just a few groups.

We mentioned previously that t_2' mainly represents the performance of Web server applications, and t_3' represents the performance of TCP/IP. The results mean that Web server applications require more time for loading data, and the fluctuations in the TCP/IP performance generates a few groups of activity.

Figure 8 shows the spatial distribution of metrics that have a heavy-tailed property as a short-time temporal property. It indicates that a host with heavy-tailedness will be plotted at the tail part of the mean value in Fig. 8 (a) and (b). We introduce a short discussion of the relationship of two proper metrics with heavy-tailedness; that is, we show how some Web servers with a heavy-tailed property in one metric are activated in another metric in **Fig. 9**.

Figure 9 (a) is overlapped with 20 cross marks selected from the tail part of Fig. 8 (b). The mean value of these cross marks exceeds 1 second, and 20 Web servers are selected from

those with maximum variance in Fig. 8 (b). Figure 9 (b) is configured from the tail part of Fig. 8 (a) as well. In Fig. 9 (a), we focus on how a specified host that obeys a heavy-tailed property for t_3' distributes and activates in the entire t_2' distribution, and how that for t_2' activates in the entire t_3' distribution in Fig. 9 (b). The relationship between the mean-variance plot of one metric and that of another with heavy-tailedness is shown in Fig. 9. These cross marks are scattered widely in Fig. 9 (a), while the cross marks in Fig. 9 (b) tend to be located at the upper boundary. These results indicate that if the Web server application operates slowly, then the system software also tends to work slowly, but not vice versa.

6. Conclusion

In this paper, we presented the experimental analysis of Web server performance using our method of active measurement. Our observed data consist of the activity of 1,984 diverse Web servers sending measuring packets 20 times for spatial property evaluation, and 14 Web servers sending packets 2,000 times in 5-second intervals for temporal property evaluation. The main contribution of our study is that we have provided a method to evaluate temporal and spatial properties on any Web server by measuring from a remote observation host; we have also included figures that illustrate the current activity of temporal and spatial property. These illustrations make it possible to determine the current status in a temporal property and to compare it with an average status in a spatial property. The study by Crovella, et al.[3] merely describes the self-similar property for the total transmission time generated by the heavy-tailed file size distribution. We divided the total transmission time into time used by the network and time used by elements that depend on server systems, and we captured the heavy-tailed property of such elements. We found that temporal property consists of two key factors: stability and activity during the Poisson process. Robustness of heavy-tailedness in terms of constructing the elements of total transmission time appeared in the analysis of spatial property. Finally, we observed the upper boundary and classified groups in a mean-variance plot of primitive elements of the Web server.

As the analysis in this paper is confined to text data on Web servers, analyses of many

other data types are required for further study to estimate accurate Web server performance and HTTP traffic. The improvement of our program will also contribute to understanding the features of multiplexed transmission of different data types.

References

1) Leland, W., Taqqu, M., Willinger, W. and Wilson, D.: On the Self-Similar Nature of Ethernet Traffic (Extended Version), *IEEE/ACM Trans. Networking*, Vol.2, No.1, pp.1–15 (Feb. 1994).

2) Park, K., Kim, G. and Crovella, M.: On the relationship between file sizes, transport protocols, and self-similar network traffic, *Proc. IEEE International Conference on Network Protocols*, pp.171–180 (1996).

3) Crovella, M and Bestavros, A: Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes, *IEEE/ACM Trans. on Networking*, Vol.5, No.6, pp.835–845 (1997).

4) Park, K., Kim, G. and Crovella, M.E.: The Protocol Stack and Its Modulating Effect on Self-similar Traffic, *Self-Similar Network Traffic and Performance Evaluation*, Park, K. and Willinger, W. (Eds.), pp.349–366, Wiley-Interscience, New York (2000).

5) Irlam, G. (Sep. 1994). Unix file size survey-1993, http://www.base.com/gordoni/ufs93.html

6) Harchol-Balter, M. and Downey, A.: Exploiting process lifetime distributions for dynamic load balancing. *Proc. SIGMETRICS '96*, pp.13–24 (1996).

7) Paxson, V. and Floyd, S.: Wide-Area traffic: the failure of Poisson modeling, *IEEE/ACM Trans. Networking*, Vol.3, No.3, pp.226–244 (1995).

8) Mosberger, D. and Jin, T.: httperf: A Tool for Measuring Web Server Performance, *First Workshop on Internet Server Performance, ACM*, pp.59–67 (1998).

9) S.P.E. Corporation (2000). http://www.spec.org/osg/web99/

10) Park, K. and Willinger, W.: Self-similar Network Traffic: An Overview *Self-Similar Network Traffic and Performance Evaluation*, Park, K. and Willinger, W. (Eds.), pp.1–38, Wiley-Interscience, New York (2000).

11) Fujita, Y., Murata, M. and Miyahara, H.: Performance Modeling and Evaluation of Web Server Systems, *The Transactions of IEICE (Japanese Edition)*, Vol.J82-B, No.3, pp.347–357 (1999).

Takuo Nakashima was born in 1956. He received his M.E. degree from Kumamoto University in 1986. After a period of working at Fujitsu Ltd., he completed coursework and further research at the Graduate School of Science and Technology at Kumamoto University in 1991 and started his career as a researcher the same year. He has been in his current position at Kyushu Tokai University since 2000. His current research interests are network analysis and active measurements. He received the IPSJ Yamashita SIG research award in 2006. He is a member of IPSJ, IEEE-CS, and ACM.