Computer Communication Review

acm sigcomm

Special Interest Group on Data Communication

The SIGCOMM Quarterly Publication

Volume 23
Number 4

October, 1993

SIGCOMM'93

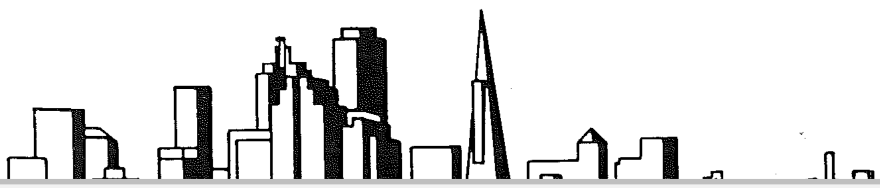
CONFERENCE PROCEEDINGS

Communications Architectures, Protocols and Applications

September 13-17, 1993

San Francisco, California, USA

acm PRESS



The Association for Computing Machinery
1515 Broadway
New York, N.Y. 10036

ISBN: 0-89791-619-0

Additional copies may be ordered prepaid from:

ACM Order Department
P.O. Box 12114
Church Street Station
New York, N.Y. 10257

Phone: 1-800-342-6626
(U.S.A. and Canada)
1-212-626-0500
(All other countries)
Fax: 1-212-944-1318
E-mail: acmpubs@acm.org

ACM Order Number: 533930

Printed in the U.S.A.

End-to-End Packet Delay and Loss Behavior in the Internet

Jean-Chrysostome Bolot
INRIA
B. P. 93
06902 Sophia-Antipolis Cedex
France
bolot@sophia.inria.fr

Abstract

We use the measured round trip delays of small UDP probe packets sent at regular time intervals to analyze the end-to-end packet delay and loss behavior in the Internet. By varying the interval between probe packets, it is possible to study the structure of the Internet load over different time scales. In this paper, the time scales of interest range from a few milliseconds to a few minutes. Our observations agree with results obtained by others using simulation and experimental approaches. For example, our estimates of Internet workload are consistent with the hypothesis of a mix of bulk traffic with larger packet size, and interactive traffic with smaller packet size. We observe compression (or clustering) of the probe packets, rapid fluctuations of queueing delays over small intervals, etc. Our results also show interesting and less expected behavior. For example, we find that the losses of probe packets are essentially random unless the probe traffic uses a large fraction of the available bandwidth. We discuss the implications of these results on the design of control mechanisms for the Internet.

1 Introduction

Current data networks typically use packet switching as a means of dynamically allocating network resources on a demand basis. Packet-switching has been widely used because it facilitates the interconnection of networks with different architectures, and it provides flexible resource allocation and good reliability against node and link failure [7]. However, packet switching provides little control over the packet delay at the switches.

SIGCOMM'93 - Ithaca, N.Y., USA /9/93

Therefore, one fundamental characteristic of a packet-switched network is the delay required to deliver a packet from a source to a destination[1]. Each packet generated by a source is routed to the destination via a sequence of intermediate nodes. The end-to-end delay is thus the sum of the delays experienced at each hop on the way to the destination. Each such delay in turn consists of two components, a fixed component which includes the transmission delay at a node and the propagation delay on the link to the next node, and a variable component which includes the processing and queueing delays at the node. Packets may be rejected at the intermediate nodes because of buffer overflow. Hence, another important characteristic of a packet-switched network is its packet loss rate.

Our objective is to understand the packet delay and loss behavior in the Internet. Understanding this behavior is important for the proper design of network algorithms such as routing and flow control algorithms (e.g. [12]), for the dimensioning of buffers and link capacity (e.g. [14]), and for choosing parameters in simulation and analytic studies (e.g. [4]). It is also essential for designing the emerging audio and video applications [8, 24, 27]. For example, the shape of the delay distribution is crucial for the proper sizing of playback buffers [24].

Many studies of packet delay and loss in various network environments have been reported in the literature. Next we describe analytic, simulation, and experimental approaches.

The obvious analytic approach is to use queueing network models to analyze packet delay in computer networks. The analysis of queueing models is relatively

[1] Throughout the paper, we refer to this delay as the end-to-end delay or simply the packet delay.

simple if certain independence assumptions hold. In this case, the analysis gives rise to so-called product-form solutions, i.e. the queue size distribution for an entire network of queues is equal to the product of the queue size distribution of the individual queues [14]. Using this result, a number of parameters including the queue size and delay distributions can be easily calculated. However, product-form networks cannot incorporate features of real-life networks such as the correlations introduced when traffic streams merge and split, the regulation of traffic by the routing and flow control mechanisms, or the packet losses due to buffer overflow. Although progress in these areas has been recently reported (e.g. [5]), it should be pointed out that due to the lack of analytic solutions, many studies of packet delay and loss behavior have been conducted with simulation and experimental approaches.

Regarding simulation approaches, recent work has examined the impact of routing and flow control mechanisms on end-to-end delay. For example, reference [25] concludes that both link state and distance vector routing yield similar average packet delay statistics in a NSFNET-like network. References [28, 29] investigate the dynamic behavior of TCP connections. In realistic situations (i.e. for connections with so-called two-way traffic), it is found that the interactions between data and acknowledgement packets generate a clustering of the acknowledgement packets which in turn gives rise to rapid fluctuations in queue lengths. These results emphasize the importance of studying the dynamics, i.e. the time-dependent behavior, of computer networks.

Regarding experimental approaches, systematic measurements of packet delay and loss were carried out on the ARPANET as early as 1971 [14, ch. 6]. They examined the variations of packet delay for different paths, different times of day and days of the week, etc. Other measurements were taken to determine how delays across the ARPANET were influenced by packet length. The results were used to assess whether TCP performance could be improved by including a dependence on packet length in the retransmission timeout algorithm [15]. Several other studies have addressed timeout adjustment in TCP, and they have proposed improvements to take into account packet losses, packet retransmissions, and the variance of packet round trip delays [12, 13].

The NSFNET replaced the ARPANET in 1990. Recent studies have measured the delay and loss behavior in the NSFNET, and more generally in the Internet. They have examined this behavior over different time scales.

Merit Network Inc. publishes monthly statistics of packet delay between the nodes of the NSFNET. These statistics are obtained from measurements performed at 15 minute intervals. They are used in [6] to examine the distribution of median delay between nodes of the NSFNET. Unfortunately, the Merit statistics are based on measurements performed between the exterior interfaces of the backbone nodes. Thus, they might not accurately characterize end-to-end delay over paths which span a combination of backbone and regional or international networks.

The behavior of end-to-end round trip delays over somewhat shorter time scales is examined in [19]. There, groups of 10 ICMP echo packets [26] are sent periodically from a source node to a destination node and echoed back to the source node, with a 1 minute interval between successive groups. Packets within a group are sent at regular 1 second intervals. Round trip delays are measured for each packet, and then averaged over a group. Various paths, i.e. source destination pairs, are considered. The results indicate that the delay distribution for all paths is best modeled by a constant plus gamma distribution, where the parameters of the gamma distribution depend on the path (e.g. a path over a regional network vs. a path over the NSFNET backbone) and the time of the day. A spectral analysis of the average delays shows a clear diurnal cycle, suggesting the presence of a base congestion level which changes slowly with time. Furthermore, packet losses and reorderings are positively correlated with various statistics of delay.

The behavior of end-to-end round trip delays over even shorter time scales is examined in [21, 22]. There, small UDP packets are sent every 39.06 ms from a source node to a destination node, and echoed back to the source node. The authors show how their measurements can be used to detect problems in the Internet. For example, they observed in May 1992 that round trip delays would increase dramatically every 90 seconds. They identified the problem as being caused by a 'debug' option in some gateway software. They identified other problems caused by synchronized routing updates, by faulty Ethernet interfaces, etc. [22]. Their measurements were also used to observe the dynamics of the Internet, e.g. the changes in round trip delays caused by route changes [21].

Despite all the efforts and results described above, the end-to-end performance of the Internet remains an area which deserves more research attention. For example, there is no clear consensus yet on how "well" the Internet performs, or on how to characterize its performance.

In this paper, we use measurements of end-to-end delay and loss to characterize the behavior of the Internet. We obtain these measurements with the UDP echo tool used in [21, 22], which provides the round trip delays of UDP packets at regular time intervals. By varying the interval between successive packets, we can examine the delay and loss behavior of the Internet over different time scales.

Our observations agree with results obtained by others using simulation and experimental approaches. For example, our estimates of Internet traffic are compatible with the hypothesis of a mix of bulk traffic using large packet size, and interactive traffic characterized by smaller packet size. We observe compression (or clustering) of the probe packets and rapid fluctuations of queueing delays over small intervals. Our estimates of Internet traffic are compatible with the hypothesis of a mix of bulk traffic using larger packet size, and interactive traffic characterized by smaller packet size. Our results also show interesting and less expected behavior. For example, we find that the losses of probe packets are essentially random unless the probe traffic uses a large fraction of the available bandwidth.

The rest of the paper is organized as follows. In Section 2, we describe the data collection process, i.e. how the measurements of packet delay and loss are obtained. In Section 3, we outline our strategy for analyzing the measurements. In Section 4, we analyze the characteristics of the measured packet delays. In Section 5, we analyze the characteristics of the measured packet losses. Section 6 concludes the paper.

2 Data collection

Recent measurements indicate that the number of hosts in the Internet is fast approaching the 1 million mark. Clearly, it is impossible to study the delay and loss characteristics for all possible connections, i.e. for all source-destination pairs. In this paper, we examine one specific connection in detail. This connection links INRIA in France to the University of Maryland (UMd) in the United States. The routes taken by the packets sent over the connection can be obtained either with the route record option of `ping`, or with `traceroute` [26]. Table 1 shows the route between INRIA and UMd as obtained with traceroute in July 1992. Nodes 5 and 6 are distinct nodes in the Ithaca Nodal Switching System. Nodes 4 and 5 are the endpoints of the transatlantic link between France and the United States. At the time the experiments were carried out (July 1992), the transatlantic link was the the bottleneck link with with a bandwidth equal to 128kb/s.

1	tom.inria.fr
2	t8-gw.inria.fr
3	sophia-gw.atlantic.fr
4	icm-sophia.icp.net
5	Ithaca.NY.NSS.NSF.NET
6	Ithaca1.NY.NSS.NSF.NET
7	nss-SURA-eth.sura.net
8	sura8-umd-c1.sura.net
9	csc2hub-gw.umd.edu
10	avwhub-gw.umd.edu

Table 1: Route between INRIA and the University of Maryland in July 1992

Packet delays and losses on the INRIA-UMd connection are obtained using *NetDyn*, a measurement tool developed by Dheeraj Sanghi [22]. This tool sends UDP packets at regular intervals from a source host to a destination host via an intermediate host. Throughout the rest of the paper, we refer to these packets as probe packets, or simply probes. Upon receipt of a probe packet from the source, the intermediate host immediately echoes the packet to the destination host. The user can specify the number of probe packets to be sent, the size of the packets, and the interval between successive packets sent by the source. In our experiments, we send probe packets of 32 bytes each. The interval between successive packets ranges over the following values: 8, 20, 50, 100, 200, and 500 ms. Each experiment lasts 10 minutes.

A packet includes three 6-byte timestamp fields. The source timestamp is written when the packet is sent by the source host. The echo timestamp is written when the packet is received by the intermediate host. The destination timestamp is written when the packet is received by the destination host. Furthermore, each packet has a unique packet number in order to detect packet losses.

If the source, intermediate, and destination hosts are geographically distant, then their local clocks may not be synchronized and hence the timestamps in the UDP probe packets would be difficult to interpret. To avoid this problem, we let the source host be the same as the destination host. Furthermore, we measure only the difference between the source timestamp and the destination timestamp, i.e. we measure only roundtrip delays. In our experiments, we use a DECstation 5000 as a source host. Its clock resolution is 3.906 ms.

We have taken measurements of end-to-end packet delay and loss on connections other than the INRIA-UMd connection, e.g. connections between UMd and MIT, between UMd and the University of Pittsburgh, between INRIA and universities in Europe, etc. Even

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.