What TCP/IP Protocol Headers Can Tell Us About the Web*

F. Donelson Smith Félix Hernández Campos Kevin Jeffay David Ott
University of North Carolina at Chapel Hill
Department of Computer Science
Chapel Hill, NC 27599-3175, USA
http://www.cs.unc.edu/Research/dirt
{smithfd,fhernand,jeffay,ott}@cs.unc.edu

October 2000

Abstract: We report the results of a large-scale empirical study of web traffic. Our study is based on over 500 GB of TCP/IP protocol-header traces collected in 1999 and 2000 (approximately one year apart) from the high-speed link connecting a large university to its Internet service provider. We also use a set of smaller traces from the NLANR repository taken at approximately the same times for comparison. The principal results from this study are: (1) empirical data suitable for constructing traffic generating models of contemporary web traffic, (2) new characterizations of TCP connection usage showing the effects of HTTP protocol improvement, notably persistent connections (*e.g.*, about 50% of web objects are now transferred on persistent connections), and (3) new characterizations of web usage and content structure that reflect the influences of "banner ads," server load balancing, and content distribution. A novel aspect of this study is to demonstrate that a relatively light-weight methodology based on passive tracing of only TCP/IP headers and off-line analysis tools can provide timely, high quality data about web traffic. We hope this will encourage more researchers to undertake ongoing data collection programs and provide the research community with data about the rapidly evolving characteristics of web traffic.

Keywords: WWW, measurement and monitoring, workload characterization, case study.

1. Introduction and Background

By any measured quantity — bytes, packets, or flows — web traffic has become the single largest consumer of Internet resources [9, 25, 33]. Prior to the explosive growth of the web, applications such as FTP, mail (SMTP), and news (NNTP) dominated the Internet. For those applications, user-oriented performance metrics (*e.g.*, response times) were much less important than they are today for web-based applications. This has profound implications for how network infrastructure and protocols are designed and their performance evaluated. Good characterizations of how web traffic "looks" in the network are essential for networking experiments (either simulations or live testing in laboratory networks) investigating end-to-end performance issues in the web. Usually this involves constructing a model of web traffic and using the model to introduce synthetically generated web traffic into the simulation or laboratory network. For example, a critical element of networking research involving the effect of TCP's end-to-end congestion-control mechanism, or router-based mechanisms such as RED, on web traffic is the generation syn-

* This work supported in parts by grant from the National Science Foundation. (grants CDA-9624662, ITR-0082870, and ITR-0082866), the Intel Corporation., the IBM Corporation, Cabletron and Aprisma Corporations, and the North Carolina Networking Initiative.

thetic web traffic for experiments [4, 8, 21, 29]. Paxson and Floyd [32] have presented compelling arguments for the importance of using application-dependent but network-independent traffic sources layered over (real or simulated) TCP implementations in such experiments. Constructing a traffic generator for TCP applications depends ultimately on the availability of high quality measurement data that represents application characteristics. In this paper we address this requirement for web traffic. We present the results of a measurement study that collected over 500 GB of TCP/IP headers in 1999 and 2000 from a high-speed link connecting a major university campus to its Internet service provider.

While the primary motivation for our measurements was to provide the modeling foundation for generating synthetic web traffic, a secondary motivation was to provide new results on the effects recent versions of the HTTP protocols are having on the characteristics of web traffic in the Internet. For example, measurements of TCP connection usage for early versions of the HTTP protocols pointed to clear inefficiencies in design, notably the creation of a different TCP connection for each web object reference [26]. Recent revisions to the HTTP protocol, notably version 1.1 [28], have introduced the concepts of *persistent connections* and *pipelining*. Persistent connections are provided to enable the reuse of a single TCP connection for multiple object references at the same IP address (typically embedded components of a web page). Pipelining allows the client to make a series of requests on a persistent connection without waiting for a response between each request (the server must return responses in the same order as the requests are sent). As browsers and servers have migrated to support the 1.1 version of HTTP, there has been very little data collected from production networks to show how these protocol modifications have changed the usage and behavior of TCP connections. Our results show that HTTP 1.1 is already producing very significant effects. We also show that other rapidly evolving developments such as the presence of "banner ads," server load balancing, and content distribution networks are influencing several observable characteristics of the web.

A final motivation is related to methodology. Our approach emphasizes simplicity and a passive non-invasive method of measuring — we used widely available packet-capture tools to gather traces consisting only of TCP/IP headers, each timestamped with its arrival time. Using only the information in the TCP/IP headers and knowledge of the TCP and HTTP protocol behaviors, we created trace-processing tools to analyze individual TCP connections and reconstruct properties of the higher-level protocol (HTTP) from the TCP segment headers. We explicitly decided not to capture HTTP protocol headers mostly because of privacy concerns about tracing users' data falling beyond the TCP/IP headers[1] and also to reduce the volume of data in traces covering extended periods of time. We show that the use of relatively straightforward methodology can produce timely, high quality data and we hope to encourage more researchers to undertake ongoing data collection programs. If this happens, it will help provide the research community with more data about the rapidly evolving characteristics of web traffic.

From the analysis of our trace data, as well as data acquired by NLANR during the same periods, we make a number of observations about the evolving nature of web traffic, the use of the web, the structure of web pages, the organization of web servers that provide their content, and the use of packet tracing as a method for understanding the dynamics of web traffic. Specifically:

---

[1] IP addresses are made anonymous in our traces to remove privacy concerns related to identifying individual users.

- From a methodology standpoint, we conclude that a substantial and detailed analysis of HTTP request/response exchanges is possible given only the TCP/IP headers of packets sent from servers to clients. Given only a unidirectional trace, one can discern transport protocol phenomena including the effects of higher-level protocols (e.g., the use of persistent HTTP connections), as well as application-level phenomena (e.g., the number of embedded objects per web page and the number of servers delivering this content). Moreover, these data can be quickly and cheaply obtained. The most recent traces were acquired with commodity PCs and publicly available packet capturing software. We have also developed a rich set of tools for processing unidirectional packet traces that enabled us to present much of the data in this paper less than a month after the raw traces were acquired.[2]
- We conclude that traces of web traffic must be quite long (on the order of hours) in order to fully capture the tail of the distribution for measures such as the size of responses from servers to clients. We have observed that a very small number of responses from servers are multiple hundreds of megabytes (some close to a gigabyte) and that the TCP connections carrying these responses can be active for upwards of an hour.
- From a protocol usage standpoint, we observe that web traffic has declined recently as a percentage of the total number of TCP connections recorded. For web traffic, 15% of all the TCP connections carrying HTTP request/response exchanges are persistent in the sense that they actually carry more than one request/response exchange. These persistent connections deliver 40-50% of all the web objects requested and these objects account for approximately 40% of all the bytes transferred. Thus, although the fraction of connections that are persistent is small, their use represents a 50% reduction in the total number of TCP connections required to deliver web content (compared to a similar environment in which persistent connections are not used).
- From an analysis of HTTP responses, we see that 65% of all web pages are constructed entirely by responses from a single server while 35% of the pages receive content from two or more servers. Close to 70% of the consecutive top-level page references go to an IP address that is different from the address used for the previous top-level page reference. We conjecture that this reflects the increasing trend of large organizations to manage their web sites with a "server farm" for load balancing.
- We also see an increase in the number of embedded objects per web page and an increase in the frequency of smaller objects. This is possibly due to the pervasive use of "banner adds" and icons to decorate pages. Overall the number of bytes transferred to deliver embedded objects is increasing.
- Overall we see a slight increase in the frequency of small response objects but a marked increase in the size of the largest objects transferred. We find that the top 15% of object sizes account for 80% of the bytes sent by servers (objects greater than 1 MB in size account for 25% of the bytes).
- From an analysis of HTTP requests we see novel uses of web requests to implement applications such as "web email." These uses result in a shift in the distribution of request sizes because files are being transferred to servers (*e.g.*, email attachments) in the context of an HTTP request.

The remainder of this paper is organized as follows. Section 2 reviews the literature in previous measurement studies of network traffic in general and web traffic in particular. We also review the methods used to gather and process network traces. Section 3 gives an overview of our measurement methodology and presents summary statistics for the traces comprising our study. Section 4 presents our methodology for reconstructing and analyzing HTTP connections given only the information contained in TCP/IP headers. Section 5 presents the analysis of user browsing metrics such as think time. Section 6 gives our analysis of the usage of TCP connections by HTTP including the effects of persistent HTTP connections. Section 7 analyzes HTTP request and response sizes and Section 8 analyzes the structure of web pages in

[2] It is our intent to distribute both our tools and the data we have acquired (and are continuing to acquire).

terms of objects, including distributions of top-level and embedded objects, and the number of servers contacted per page. Finally, Section 9 comments of the use of these data in the generation of web traffic for network simulation and testing.

2. Related Work

Two important measurement efforts that focused on application-specific traffic models, but which preceded the growth of the web, were conducted by Danzig *et al.*, [7, 13, 14], and by Paxson and Floyd [30, 31]. More recently, measurements to characterize web usage have become a very active area for research. Because caching and content delivery are widely considered vital to the long-term viability of the web, most of the high quality data currently available is focused on providing inputs to cache or content delivery evaluations, *e.g.*, [15, 16, 17, 19, 22, 35, 36]. For these studies, the critical data are traces or logs of URL references, typically collected at proxies or servers. There is much less data available that is focused on how web browsing behaviors by users result in the creation of network traffic. For networking studies, the critical data are related to characterizing the TCP connections between web browsers and servers in terms of connection establishment rates and the sizes and timing of exchanges of request and response data.

Web traffic generators in use today are usually based on data from the two pioneering measurement projects that focused on capturing web-browsing behaviors: the Mah [23], and Crovella, *et al*., [3, 5, 11, 12] studies. Traffic generators based on both of these sources have been built into the widely used *ns* network simulator [6] that has been used in a number of studies related to web-like traffic, *e.g.*, [21, 29]. These models have also been used to generate web-like traffic in laboratory networks [4, 8]. For both sets of measurements, the populations of users were highly distinctive and the sizes of the traces gathered were relatively small. Mah captured data reflecting a user population of graduate students in the Computer Science Department at UC Berkeley. His results were based on analysis of approximately 1.7 million TCP segments carrying HTTP protocols. The measurement programs by Crovella and colleagues reflected a user population consisting primarily of undergraduate students in the Computer Science Department at Boston University and in aggregate represented around 1 million references to web objects. In addition, both sets of data are now relatively old (especially as measured in “Internet time”) The Mah data were collected in 1995 and the Crovella, *et al*., data in 1995 and 1998. It is especially important to note that these studies were conducted before significant deployment of HTTP version 1.1 protocol implementations. For comparison, our study involved traces consisting of over 800 million TCP segments generated by a user population of approximately 35,000 and representing the transfer of some 55 million web objects. Moreover, we have developed a capability for nearly real-time analysis of our data and, because of this, half of the results we report here are based on data less than one month old.

We are aware of at least five projects involving the analysis of large-scale packet-level traces containing web traffic. None of these projects have reported on any analysis of the traces to extract information complete enough to create traffic generating models or characterize user browsing behavior. Gribble and Brewer [22] in late 1996 collected 45 consecutive days of continuous traces from the dial-in Home IP service for the UC Berkeley campus. Their traces comprise 24 million references to web objects from over 8,000 unique clients. By using an on-line analysis tool they examined the HTTP request and re-

sponse headers as well as TCP/IP header information. While the primary focus of their study was on factors that would be of use to web cache designers (locality, cache-control headers, etc.) they also reported some data on request interarrival times and mean sizes for HTML, GIF, and JPEG object types. Balakrishnan, *et al.* [2], collected packet traces at the official web site for the 1996 Olympic Summer Games. Their traces consisted of over 1.5 billion packets from 720,000 unique client IP addresses. Their emphasis was on improving the performance of TCP protocols and algorithms when used to carry web traffic. Their analysis of the traces, therefore, was concerned with issues such as TCP loss recovery, ACK compression, receiver bottlenecks, and congestion control for multiple parallel connections. Cleveland, *et al*. [10] used traces of 23 million TCP connections for web browsing captured between November 1998 and July 1999 on a link connecting a Bell Labs internal network to the Internet. These traces were used to develop a statistical model for generating TCP connection start times from web clients using a notion of connection-rate superposition but they did not model other TCP connection characteristics or web browsing behavior.

Researchers at the University of Washington have been collecting continuous traces on the switches that are transited by all the university's Internet traffic. They use custom trace software to identify TCP connections carrying web traffic, extract HTTP headers from TCP segments, and log information parsed from the headers. These data were used in two studies of web proxy caching [35, 36] both of which were based on seven days of traces in 1999 that included over 82 million references to web objects from about 23,000 clients. Feldman [18] summarizes the results of over three years of large-scale trace-gathering projects using PacketScope monitors (with special on-line analysis software to process HTTP headers) at several locations in the AT&T WorldNet IP network. She reviews the many challenges faced in reconstructing TCP connections from individual segments and reconstructing HTTP protocol characteristics from TCP connections and HTTP headers (many of which we encountered with our own analysis tools). She also presents a survey of several projects in web cache design and content compression that successfully used these traces [15, 19, 20, 27].

Our approach is to use off-the-shelf hardware and publicly available packet capture tools. Privacy considerations limit us from capturing more than the TCP/IP header, however, as we show below, with careful analysis, significant and substantial data on protocol usage and the nature, structure, and distribution of web content can be gleaned from just the TCP/IP headers. One drawback to our approach is that all processing of the traces is done off-line. This means that the length of traces is fundamentally limited by the amount of disk space available, the packet arrival rate, and the bytes traced per packet. As a practical matter this limited individual traces of TCP/IP headers to about one hour (we expect to soon add enough disk capacity to trace over eight hour intervals). More elaborate hardware/software instrumentation that supports continuous capture and real-time analysis has been developed elsewhere [18, 22, 24, 35]. For our intended uses of the data, however, we believe our simpler approach represents a viable alternative to instrumentation embedded directly in browser software or to specialized tracing hardware and software that analyzes entire packet contents (including user data) and logs the results. Both of these approaches present significant barriers to widespread use. Given the rapidly falling prices and increasing sizes of PC hard disks (*e.g.*, under $4.00 per GB for 60 GB disks), it is quite feasible to use very large pools of storage for intermediate processing of TCP/IP header traces.

3. The Trace Data

The data used in our study are from two sources. The two largest trace collections were obtained by placing network monitors on the high-speed link connecting our university campus network to the Internet via our Internet service provider (ISP). This campus is home to a large public research-I university. All units of the university including administration, academic departments, research institutions, and a medical complex (including a hospital that is the center of a regional health-care network) all use a single ISP link for Internet connectivity. The user population is large (over 35,000) and diverse in their interests and how they use the web — including, for example, student "surfing" (and music downloading), access to research publications and data, business-to-consumer shopping, and business-to-business purchases by the university. In addition to the thousands of on-campus computers (the large majority of which are Intel architecture PCs running some variant of Microsoft Windows), several units of the university operate dial-in modem pools (total of about 250 ports) that are used by some students, faculty, and staff for access from home. In effect, the university is a local ISP for this population, forwarding all Internet traffic to its upstream ISP. There are only a handful of small proxy servers on campus so almost all the web traffic is generated directly by browser actions. It is important, however, to remember that all web traffic we observed represents only requests that could not be satisfied from local browser caches.

We used network monitors to capture traces of TCP/IP headers from all packets entering and leaving the campus network. The traces were collected during six one-hour sampling periods each day (periodic sampling was necessary because of storage and processing capacity limitations given the high volume of packets traversing this link). The one hour sampling periods were 8:30-9:30AM, 11:00-12:00 noon, 1:30-2:30PM, 4:00-5:00PM, 7:30-8:30PM, and 10:00-11:00PM. These periods were chosen somewhat arbitrarily to produce four traces during the normal business day and two during non-business hours when traffic volumes were still reasonably high. One set of traces consists of all the TCP/IP headers collected during these sampling intervals over a seven-day period, in late September 1999. This seven-day period provided a set of six traces from each of the seven weekdays for a total of 42 one-hour traces. This set of traces will be referred to in this paper as "university-99." The second set of 42 traces also consists of traces taken at the same hours over a seven-day period, in late September 2000 (referred to as "university-00"). This allows us to compare results from traces gathered approximately one year apart.

When the "university-99" traces were gathered, our campus was connected to the ISP by an OC-3 (155 Mbps) full-duplex, ATM link. This link carried all network traffic between the campus and the "public" Internet (traffic between the campus and Internet 2 sites was routed over a separate OC-3 link). We placed the monitor on the OC-3 link to the "public" Internet. The specific monitor used was the OC3mon developed initially at MCI for vBNS [1] and now distributed by CAIDA (CoralReef) [37]. This monitor was passively inserted in the link using a fiber splitter to divert some light from the optical signal. The signal is input to the receive port of an ATM interface card hosted in an Intel-architecture PC equipped with large, high-performance hard disks. Because the link was full-duplex (two fibers) the monitor required two ATM interface cards, each receiving the signal for one of the directions of transmission (*i.e.*, inbound or outbound with respect to the campus). The OC3mon software ran on FreeBSD (version 2.2.8) and produced traces of timestamped entries giving the contents of ATM cells carrying the TCP/IP headers. The trace entries for each interface were filtered into separate traces for the inbound and outbound traffic. We

also converted the OC3mon trace into the format used by *tcpdump* using a locally-modified version of a tool, *mon2dump*, originally developed at MCI. As a result of this conversion, the resolution of the time-stamps was reduced to one microsecond.

A year later when the "university-00" traces were taken, the ISP link had been replaced by an OC-12 (622 Mpbs) path based on Cisco-proprietary DPT technology instead of ATM. Thus we could not use the same monitoring tool we used for the 1999 traces. Fortunately, all the traffic between the campus and the Internet traversed a single full-duplex gigabit Ethernet link from the campus aggregation switch to the edge router with the DPT interface. In this configuration, both "public" Internet and Internet 2 traffic are co-mingled on the one Ethernet link (the only traffic on this link is traffic to and from the ISP edge router). We placed a monitor on this gigabit Ethernet by passively inserting a fiber splitter to divert some light to the receive port of a gigabit Ethernet network interface card (NIC) set in "promiscuous" mode. Because the link is full-duplex (two fibers) the monitor required two NICs, and each received the signal for one of the directions of transmission (*i.e.*, inbound or outbound with respect to the campus). The NICs were hosted in an Intel-architecture PC equipped with large, high-performance hard disks. The operating system for this monitor was FreeBSD (version 3.2) and an instance of the *tcpdump* program was run on each of the interfaces to collect a trace of TCP/IP packet headers. Buffer space of 3 MB was allocated to the *bpf* devices used by *tcpdump* in order to buffer transient overloads in packet arrivals. The *tcpdump* program reports statistics on the number of packets dropped. We found that in many traces no packets were dropped and that the maximum number of drops in any trace was less than 0.02% (average of 0.004% drops over all traces)[3].

In this paper we use data only from the traces of packets flowing into the campus network from the ISP ("inbound" packets). Summary statistics for the inbound packets in the two sets of university traces is given in Table 1. The volume of Internet data increased significantly between 1999 and 2000. Some part of this can be attributed to the fact that the 2000 data includes both "public" Internet and Internet 2 traffic. However, most of it is simply growth of Internet usage by the university population. While the number of HTTP bytes flowing into the campus doubled between 1999 and 2000, web traffic actually declined as a percentage of all TCP traffic. This is attributed to the sudden popularity of the *Napster* application for downloading audio files [39].

As a "sanity check" of our data, two considerably smaller sets of traces from the repository of traces at NLANR were used for comparison [38]. This allowed us to both debug our measurement methodology and verify that other university campuses see similar patterns of web traffic. (in Section 7 we comment on some of the limitations the NLANR traces for our analysis.) The NLANR traces were gathered using the OC3mon/Coral software described above. We selected from the NLANR repository two sets of traces collected at approximately the same time periods in 1999 and 2000 as our university traces. Within these time periods, we selected traces from sites that appeared to have relatively high volumes of traffic during the sampling times used by NLANR. The set of traces we refer to as "NLANR-99" is composed of eight traces from each of two sites (Merit-U. of Michigan, and the San Diego Supercomputer Center "com-

[3] The OC3mon does not explicitly report dropped packets but halts with an error if the trace overruns the disk. This did not occur in any of our traces.

Table 1: Summary data for the University and NLANR traces.

	University-99	University-00	NLANR-99	NLANR-00
TCP Packets	525,258,293	1,872,964,615	16,919,364	18,656,630
% TCP Packets	85.08%	90.77%	84.66%	90.13%
UDP Packets	89,759,578	180,482,572	2,681,268	1,596,875
% UDP Packets	14.54%	8.75%	13.42%	7.71%
HTTP Packets	232,245,911	602,183,175	9,263,705	7,617,906
% HTTP Packets	37.62%	29.18%	46.35%	36.80%
Total Packets	617,333,998	2,063,351,175	19,985,154	20,699,713
TCP Bytes	211,610,632,288	721,866,693,794	8,374,506,562	9,744,043,417
% TCP Bytes	86.15%	89.84%	93.08%	95.85%
UDP Bytes	33,760,062,473	80,921,395,625	480,644,915	385,298,496
% UDP Bytes	13.74%	10.07%	5.34%	3.79%
HTTP Bytes	138,050,696,899	278,484,679,075	4,198,123,465	3,351,340,432
%HTTP Bytes	56.20%	34.66%	46.66%	32.97%
Total Bytes	245,636,674,456	803,493,236,871	8,996,799,288	10,165,498,893

modity connection") all taken on September 19, 1999. Our "NLANR-00" set also consists of eight traces from the same sites. These traces were all taken on September 28, 2000. Summary statistics for the NLANR traces are also given in Table 1.

4. Analysis of TCP Connections

Because we have only traces of TCP/IP headers, all the statistics we report here have been derived from analysis of these headers knowing their formats and the dynamic behaviors of the TCP and HTTP protocols. The primary information used from the TCP/IP headers was the IP source and destination addresses, the protocol number (to identify TCP segments), the source and destination ports, the TCP flags (to detect SYN, FIN, and Reset), the data sequence number, and the acknowledgement (ACK) number.

The OC3mon tool used to gather the university-99 and all the NLANR traces puts an arrival timestamp on each trace entry using a clock local to the ATM interface card that receives the data. Because the trace entries from each direction of data flow on the full-duplex OC-3 link are timestamped with a different clock, it is difficult to create a merged bi-directional trace in the correct order. Even though the Coral software attempts to initialize the clocks on both interface cards to the same value, there is usually a random initial offset between the clocks and the clocks may drift at different rates. Our attempts to produce merged bi-directional traces by sorting on timestamps resulted in a large number of TCP connections with obviously incorrect orderings (*e.g.*, SYN+ACK before SYN).[4] Fortunately, it is not necessary to merge (or even use) both directions of data flow in order to infer important elements of the internal dynamics within a TCP connection and reconstruct parts of the HTTP protocol.

Consider a trace consisting only of TCP/IP packet headers captured on the ATM interface receiving IP packets arriving on the inbound path of the OC-3 link (to the university from its ISP). We first filter this trace to pull out only those IP packets where the protocol field in the IP header designates TCP and the

[4] Later versions of the Coral software have improved the bounds on clock synchronization.

source port field in the TCP header contains the value 80 (the normal HTTP server port). This produces a trace of TCP/IP packet headers that were sent from Web servers somewhere in the Internet to Web clients (browsers) located at the university. This filtered trace is then sorted on three keys in the following order: source IP address and port, destination IP address and port, and timestamp. This produces a time ordered trace of TCP segments within each TCP connection (actually within unique TCP connection address 4-tuples; port reuse occasionally produces multiple TCP connections using the same 4-tuple within a trace). The university-00 traces taken with *tcpdump* are handled the same way.

The HTTP protocol is asymmetric — the client always initiates the connection (sends the TCP initial SYN segment). The server normally continues the connection establishment protocol by responding with a SYN+ACK segment. This SYN+ACK segment should appear in our trace and its timestamp is used for the beginning time of a TCP connection. Similarly the connection is considered to end when a FIN or Reset segment from the server is found in the trace. For the HTTP protocol, we are interested in the exchanges of data between the browser and server that occur within the TCP connection. Specifically, we want to identify the first and last bytes of browser requests and the first and last byte of server responses. For those cases where the TCP connection is used for more than one request/response pair (when both the browser and server support persistent connections), we need to identify the beginning and end of requests and responses for multiple exchanges between the browser and server. Pipelining complicates the identification of request/response exchanges in persistent connections and is discussed further below. Fortunately the TCP protocol allows us to infer this information from examination of only the TCP segments flowing from the server to the browser.

Consider the common case where the first TCP segments that flow on an established TCP connection are the HTTP-request protocol elements sent by the browser. As the TCP protocol stack on the server receives the segment(s) comprising the request, it will send TCP acknowledgment sequence numbers (ACKs) indicating the in-order byte sequence it has received. These may be sent in segments containing only an ACK or in segments containing an ACK and HTTP-response protocol elements plus the requested object. The ACK may be sent immediately, delayed by up to 200 milliseconds, or be sent on the next outbound data segment on that connection. The important observations are that the ACK value will advance by an amount equal to the size of the request protocol elements and that all of the request message will be ACKed *no later than* the first segment carrying any data for the corresponding response. The size of the response is indicated by the amount the data sequence number advances in segments from the server. In the case of persistent connections with more than one request/response exchange, we will see an alternating pattern of advancing ACK values followed by advancing data sequence numbers from the server. Some examples from our traces should aid in understanding these observations. Shown in Figures 1 and 2 below are examples from *tcpdump* of the by-far most common cases, one that is not a persistent TCP connection and one that is.

In Figure 1, the first segment from the server following the SYN+ACK segment acknowledges 334 bytes of request and contains the first 1,460 bytes of the response. In the following segments, the data sequence numbers advance to 4,818 (the size of the response) with no further changes in the ACK values. In Figure 2, the first segment from the server following the SYN+ACK segment acknowledges 304 bytes of request. In the next segment, the ACK number does not advance but the data sequence number does indi-

```
924200314.823839 xxx.yyy.128.6.80 > aaa.bbb.38.4.1080: S 998480228:998480228(0) ack 3850385 win 8760 1460>
924200315.117611 xxx.yyy.128.6.80 > aaa.bbb.38.4.1080: P 1:1461(1460) ack 335 win 8760
924200315.812598 xxx.yyy.128.6.80 > aaa.bbb.38.4.1080: . 1461:2921(1460) ack 335 win 8760
924200315.812721 xxx.yyy.128.6.80 > aaa.bbb.38.4.1080: P 2921:4381(1460) ack 335 win 8760
924200316.360468 xxx.yyy.128.6.80 > aaa.bbb.38.4.1080: P 4381:4818(437) ack 335 win 8760
924200331.938448 xxx.yyy.128.6.80 > aaa.bbb.38.4.1080: F 4818:4818(0) ack 335 win 8760
```

Figure 1: *tcpdump* TCP connection trace for a non-persistent HTTP connection with one request/response exchange.

```
924200313.814433 uuu.vvv.46.83.80 > xxx.yyy.130.45.1038: S 1343529051:1343529051(0) ack 2062207851 win 32736
924200313.997334 uuu.vvv.46.83.80 > xxx.yyy.130.45.1038: . ack 305 win 32736
924200314.099386 uuu.vvv.46.83.80 > xxx.yyy.130.45.1038: P 1:1461(1460) ack 305 win 32736
924200314.180664 uuu.vvv.46.83.80 > xxx.yyy.130.45.1038: . 1461:2921(1460) ack 305 win 32736
924200314.611830 uuu.vvv.46.83.80 > xxx.yyy.130.45.1038: . 2921:4381(1460) ack 305 win 32736
924200314.845951 uuu.vvv.46.83.80 > xxx.yyy.130.45.1038: . 4381:5841(1460) ack 305 win 32736
924200314.953118 uuu.vvv.46.83.80 > xxx.yyy.130.45.1038: . 5841:7301(1460) ack 305 win 32736
924200315.250695 uuu.vvv.46.83.80 > xxx.yyy.130.45.1038: . 7301:8761(1460) ack 305 win 32736
924200315.718471 uuu.vvv.46.83.80 > xxx.yyy.130.45.1038: P 8761:9502(741) ack 305 win 32736
924200330.989115 uuu.vvv.46.83.80 > xxx.yyy.130.45.1038: . ack 609 win 32736
924200331.291394 uuu.vvv.46.83.80 > xxx.yyy.130.45.1038: P 9502:10962(1460) ack 609 win 32736
924200331.811170 uuu.vvv.46.83.80 > xxx.yyy.130.45.1038: . 10962:12422(1460) ack 609 win 32736
924200332.216693 uuu.vvv.46.83.80 > xxx.yyy.130.45.1038: . 12422:13882(1460) ack 609 win 32736
924200332.663373 uuu.vvv.46.83.80 > xxx.yyy.130.45.1038: P 13882:15161(1279) ack 609 win 32736
924200347.307346 uuu.vvv.46.83.80 > xxx.yyy.130.45.1038: F 15161:15161(0) ack 609 win 32736
```

Figure 2: *tcpdump* TCP connection trace for a persistent HTTP connection with two request/response exchanges.

cating that the initial request was 304 bytes in size. The data sequence numbers advance in all later segments until time 924200330.989115 when we find an ACK-only segment that advances the ACK sequence number and indicates the first response (which has a length of 9,502 bytes) has ended and a new request has begun. In the next segment, the ACK number does not advance but the data sequence number does indicating that the second request was also 304 bytes (609 – 305) in size. In the following segments, the data sequence numbers advance with no further changes in the ACK values. The size of the second response is 5,659 bytes (15,161 – 9,502).

Fundamental to all of this is the observation that a server should not be sending data in response to a request unless that data is accompanied or preceded by an advance in the ACK sequence covering receipt of the request segments. Similarly, any new data segment sent by a server that follows or is accompanied by an advance in the ACK sequence number is assumed to be a response to the request that caused the ACK sequence to advance. Put another way, response segments (a sequence number advance) mark the end of a request and ACK advances mark the end of a response. Of course other events such as FIN or Reset can mark ends also. A request's start time is the timestamp on the trace entry containing the first advance in the ACK field following the connection establishment or a sequence of response segments. A response's start time is the timestamp on the trace entry containing the first advance in the sequence number field following a sequence of request segments. A response's ending time is the timestamp on the trace entry

that last advanced the data sequence number before the response ended (a new request starts, a FIN is sent, *etc*.). Similarly, a request is considered to end at the timestamp on the last trace entry of the request.

All of this would be quite straightforward if it were not for all the ways real TCP connections deviate from such well-behaved traces.[5] Retransmissions and segment reordering in the network disturb this use of advancing (ACK or data) sequence numbers to mark requests and responses. In traces of TCP segments, the data sequence numbers may not be monotonically increasing. In some cases, such as reorderings or retransmissions of data-only segments in a response, this presents no problem since only the highest sequence number seen is used. The length of a response in a persistent connection can be computed as the difference between the (largest) sequence at the end of one response and the (largest) sequence at the end of the subsequent response.

Reordering of ACKs (especially in data-carrying segments) presents problems since boundaries between requests and responses may be missed which can result in overstating or understating request and response sizes. ACKs should be monotonically increasing so the length of a request in a persistent connection can be computed as the difference between the ACK value marking the end of one request and the ACK value marking the end of the next request. Unfortunately, out-of-order segments can cause ACKs to appear to "go backward." For segments without data (ACK only), simply ignoring the "backward" ACK is fine. For persistent connections, out-of-order ACKs in segments carrying response data can lead to incorrect results. Each case of suspected ACK ordering that might lead to erroneous request or response lengths was recorded for off-line investigation. In the university traces, we found about 4,900 instances of connections with out-of-order segments ordered such that the analysis tool could not make an unambiguous decision. We examined a random sample of several hundred of these cases and found that less than half were analyzed incorrectly. Thus we believe that the very small percentage (less than 0.005 %) of incorrect values included in the data has no appreciable effect on the results.

Pipelining introduces the possibility of errors in determining the lengths of requests and responses. As long as the server TCP stack receives all the request segments generated in a pipeline before it sends the first response segment, the entire pipeline of requests will be treated as one (larger) request. Similarly, as long as the server sends all the response segments for a pipelined response before receiving segments for a new client request, the entire pipeline of responses will be treated as one (larger) response. This inflates the sizes of requests and responses and deflates the number of exchanges per persistent connection. If a pipeline of requests did overlap with (one or more) responses or vice versa, the analysis tool may calculate the sizes of individual requests and responses incorrectly. Our detailed examination of a substantial number of traces from persistent connections found no obvious indications of overlapped pipelines, an observation that correlates well with the findings in [34] that pipelining is not widely implemented.

A final complication arises from the fact that our traces cover specific intervals of time. This means that at the beginnings and ends of the traces we find incomplete TCP connections (5-7% of the total connec-

[5] We describe here only the "expected" exceptional behaviors related to loss, retransmissions, and reordering; the truly bizarre sequences that our analysis tools unearthed are fortunately so rare that completely discarding those connections does not effect the results.

tions in the university traces[6]). We have excluded from our analysis any data that was made ambiguous by missing the beginning or end of the TCP connection.

5. User and Web Content Characterizations

Our traces do not include any part of the user data carried in TCP segments. Because we do not have access to any of the HTTP protocol headers, we must use heuristics to infer characteristics of user and browser behavior from the analysis of TCP connections. The first step in this process is to aggregate TCP connection traces by unique client IP addresses (the unique IP destination addresses found in our traces filtered for source port 80). We then create a time-sorted summary of the TCP connection activity between each individual client and the server(s) that client used. We assume that in the vast majority of cases a client IP address identifies a single human user running one or more browser instances on a personal computer or workstation. Although we know that there are times when multiple users concurrently run browsers on a shared compute server (single client IP address), we believe this to be rare on our campus where there is basically one computer for each Internet user. The time-sorted summary of TCP connections used by a client contains the connection start times, the server IP address for each connection, the beginning time and size in bytes of each request in all connections, the beginning and ending times of each response along with its size, and the ending time of the connection. We then use this time-ordered information to infer certain characteristics of activity by the user or the browser software.

Using a heuristic approach similar to those developed originally by Mah [23] and Barford and Crovella [3], we attempted to identify points in each client's activity that is likely to mark a request for a new (or refreshed) page. We use the term "page" as a convenient label for a web object referenced in a "top-level" sense, *i.e.*, it is not referenced through interpreting references found internal to some other object (*e.g.*, embedded references in HTML). We also use the term "object" synonymously with a response from a web server. Server responses that are error reports (*e.g.*, "404 – Not found") are counted as objects (or pages) in this discussion. We assume that page references normally occur after some period of idle or "think" time at the client, *e.g.*, the time a user spends digesting the contents of one browser display and selecting a link to (or entering) a new page reference. This same model of a page request following an idle period also captures the behavior of periodically refreshed pages.

We define an idle period heuristically by examining the time-ordered set of TCP connections used by a client. We identify periods in which the client either has no established TCP connections or where no established connection has an active request/response exchange in progress. We consider a request/response exchange to be active from time the request begins until the corresponding response ends. If any such period persists for longer than a time threshold, it is classified as an idle period. We found empirically that a threshold of 1 second works well for distinguishing idle periods (as did Mah and Barford and Crovella). It is important to note that this approach works only on traces for which we can be reasonably certain that *all* the TCP connections for a given browser appear in the traces. Since the NLANR traces have no information about where clients are located relative to the monitoring point, we perform this analysis only on

[6] We found about 20% of the connections in the NLANR traces to be incomplete because of the short (90 second) trace durations.

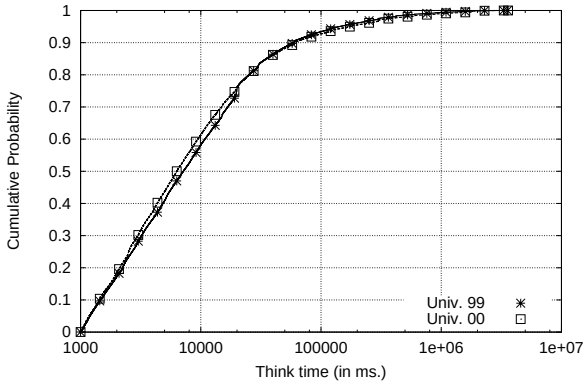


Figure 3: Cumulative idle ("think") time distribution.

the university traces where we know the clients are located on (or dialed into) the campus network and the servers are located somewhere in the Internet. Figure 3 shows the distribution of idle periods greater than one second observed in the university data. There are no appreciable differences between the 1999 and 2000 results; 60% of idle periods are between 1 and 10 seconds and approximately 90% of idle periods are less than 10 minutes.

We consider the initial request/response exchange following an idle period to be for the "top-level" page object (typically HTML) and all the subsequent request/response exchanges before the next idle period to be for the "embedded" object references (if any) within the initial page object. The server IP address involved in the request/response exchange for the top-level page object is considered to be the *primary* server for the page. All server IP addresses not equal to the primary IP address involved in subsequent request/response exchanges for objects related to that page are considered to be *non-primary* servers. Embedded objects may come from either the primary server or from non-primary servers.

6. TCP connection usage

We have implemented a suite of tools for processing *tcpdump* formatted traces that produce statistical data to characterize TCP connection behaviors as observed in the 1999 and 2000 traces from the university and NLANR. We first present results concerning the usage of persistent and non-persistent connections. Our classification of a TCP connection used in HTTP as persistent reflects actual usage characteristics, not whether the browser and server using the connection have used the HTTP protocol to establish a persistent connection. Our definition is thus of *effective persistence*, that is, whether or not a TCP connection is actually used for multiple request/response exchanges. A TCP connection is considered persistent if it is actually used for two or more request/response exchanges. All TCP connections used for one request/response exchange are considered non-persistent. This means that TCP connections in which the browser and server have enabled a persistent connection but make only one request/response exchange are considered non-persistent. Partial connections (normally at the beginning and end of traces as discussed above) and those terminated without any exchanges are not counted.

Table 2 gives summary information about how TCP connections are used in the web. While 15% or fewer of all TCP connections are effectively persistent, they are now used for 40-50% of all object references representing about 40% of all bytes transferred. This means that persistent connections now have a sig-

Table 2: Summary data for TCP connections used in the web.

	University-99	University-00	NLANR-99	NLANR-00
Users Browsing	140,552	244,025	36,245	27,598
Top-level Objects	3,722,038	6,586,547	N/A	N/A
Embedded Objects	14,799,560	30,836,389	N/A	N/A
% Embedded Objects	79.90%	82.40%	N/A	N/A
Non-Persistent Connections	9,620,971	17,672,258	261,185	187159
Persistent Connections	1,551,642	3,152,326	34,282	29,895
% Persistent Connections	13.89%	15.14%	11.60%	13.77%
Unclassified Connections	577,517	1,513,872	58,440	64,083
% Unclassified	4.91%	6.78%	16.51%	22.79%
Objects on Non-Persistent	9,620,971	17,672,258	261,185	187,159
Objects on Persistent	8,280,849	17,497,714	153,179	140,192
% Objects on Persistent	46.26%	49.75%	36.97%	42.83%
Bytes on Non-Persistent	66,522,598,025	124,665,042,556	1,490,813,422	991,967,768
Bytes on Persistent	45,471,488,454	84,205,017,825	680,123,628	550,485,402
% Bytes on Persistent	40.60%	40.41%	31.33%	35.69%

nificant influence on the dynamics of TCP connections for the web. Put another way, the number of TCP connections required for web traffic is now approximately 50% lower than it would have been with the original HTTP protocol. Figures 4 and 5 show the distributions of the number of request/response exchanges for persistent connections. Over 60% of persistent connections are used for three or more request/response exchanges and 10% carry more than ten. As Figure 5 shows, some TCP connections carry more than 100 request/response exchanges.

Figures 6 and 7 show the usage of TCP connections for all the elements of a web page (both top-level and embedded objects). Figure 6 is the distribution of unique TCP connections used in requesting all the objects for a page. Around 55% of all pages are fetched using two or more unique TCP connections (which may be any mix of persistent and non-persistent connections depending on the capabilities of the server and browser at the endpoints of the connection). Around 50% of all pages required 2-10 unique TCP connections. The number of unique TCP connections used for a page is a result of complex factors including the number of objects in the page, the number of servers holding the objects for the page, the number of concurrent TCP connections the browser opens to each server, whether the client and server support per-

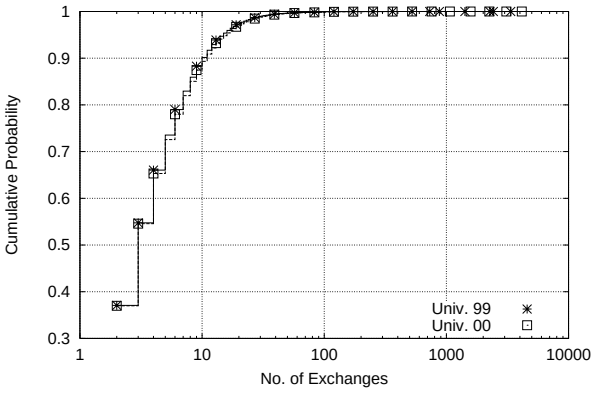


Figure 4: Cumulative distribution of request/ response exchanges per persistent connection.

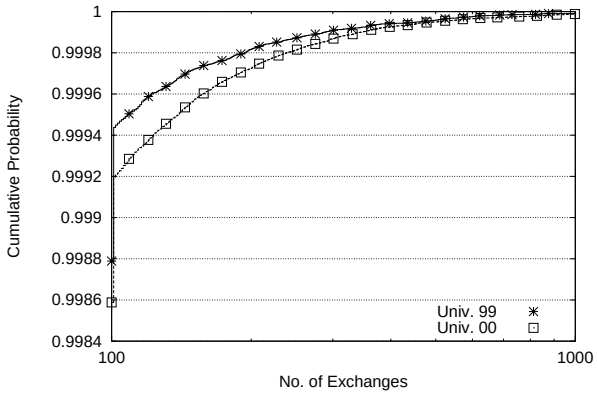


Figure 5: Distribution of request/response exchanges per persistent connection >100.

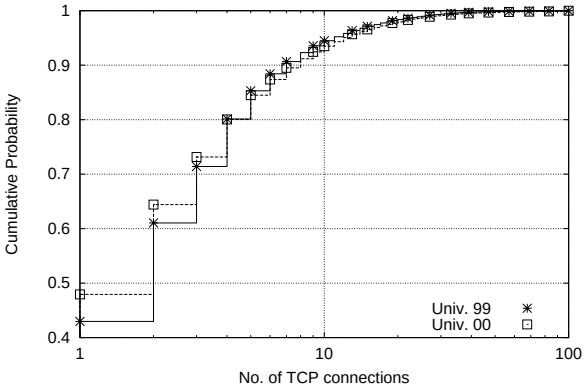


Figure 6: Cumulative distribution of unique TCP connections per page.

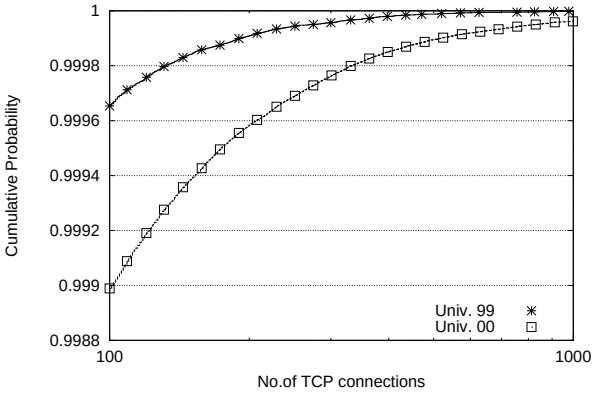


Figure 7: Cumulative distribution of unique TCP connections per page >100.

sistent connections, and how aggressive they are about keeping persistent connections active over multiple pages. Figure 7 shows that in some cases several hundred connections are used for a single page. Figure 8 gives the distribution of unique server IP addresses per page. While about 65% of all pages can be obtained from a single server IP address, about 35% require connections to 2-10 different IP addresses (and rarely as many as 100 IP addresses). We believe these results are a reflection of the ways page content is obtained dynamically from a number of sources including banner ads from agency sites and content that has been explicitly distributed to content servers (*e.g.*, Akamai).

7. Request and Response Data Sizes

We now consider the sizes of individual request and response elements of the HTTP protocol as identified by our analysis tools. Figures 9, and 10 show the distribution of request sizes. Request sizes are defined in our analysis as the number of bytes sent by the browser in a single request/response exchange as explained above. As expected, over 90% of the requests are between 100 and 1,000 bytes in size. The most surprising feature of this distribution, however, is the tail (see Figure 10) which indicates the presence of some very large requests, especially in the 2000 traces. Table 3 shows the sizes of the 10 largest requests in each of the trace collections. Sizes of the very largest requests are significantly bigger in the 2000 traces. We have looked more closely at how these very large requests arise (which are larger than one would expect from submitting forms with web browsers). An interesting example we found was the use

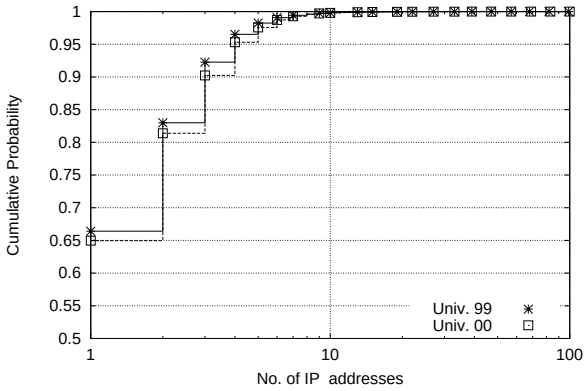


Figure 8: Cumulative distribution of unique server IP addresses visited per page.

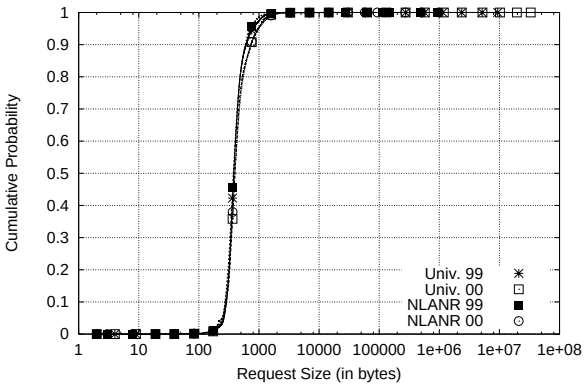


Figure 9: Cumulative distribution of request data sizes.

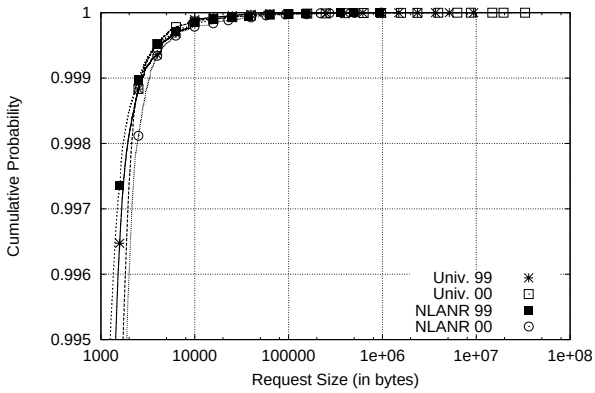


Figure 10: Cumulative distribution of request data sizes >1,000 bytes.

Table 3: Sizes in bytes of the ten largest requests in each set of traces.

University-99	University-00	NLANR-99	NLANR-00
3,212,320	8,610,634	94,266	141,924
3,252,718	8,752,959	96,056	155,241
3,315,157	8,758,838	99,500	173,031
3,555,055	10,216,308	112,345	214,756
3,573,932	11,055,604	143,560	216,984
3,674,698	11,187,714	149,608	218,724
4,302,188	14,393,587	352,256	271,408
5,166,961	14,815,754	358,928	450,927
8,025,483	19,526,090	487,359	504,412
9,389,131	33,120,413	926,994	974,561

of web-enabled email, specifically users of Yahoo email, that send email messages with large attachments. In one trace, a user sent email with an approximately 500K attachment producing a request of that size but eliciting a response from the Yahoo server of only 2.2K bytes. We speculate that as browsers become more widely used as interfaces for web-enabled applications such as email, large request elements will become more significant.

The requests over 1,000 bytes in size already represent a non-trivial contribution to the total bytes transmitted as requests. Figure 11 shows the cumulative percentage of total bytes in requests as a function of request size. Figures 12 and 13 zoom in on request sizes between 100 and 1,000 bytes and above 1,000 bytes, respectively. These plots clearly show the growing contribution of large requests to the total volume of request bytes in the university and NLANR traces (*e.g.*, over 20% of the request bytes come from requests larger than 1,000 bytes in the university-00 traces). There has been a noticeable change from 1999 to 2000 (5% of the bytes were in requests larger than 10,000 bytes compared to 3% in 1999).

Response sizes are defined in our analysis as the number of bytes sent by the server in a single request/response exchange as explained above. Figure 14 gives the overall distribution of response sizes while Figures 15 and 16 zoom in on the distributions of sizes between 100 and 100,000 bytes and those larger than 50,000 bytes, respectively. Overall we find that about 85% of the responses observed were 10,000 bytes or less. These distributions also indicate some potentially interesting shifts between 1999 and 2000 in the proportions of responses in the range of 200 bytes to 10,000 bytes. In both the university

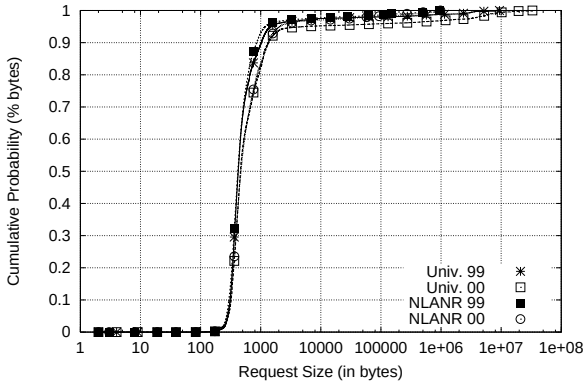


Figure 11: Cumulative distribution of request bytes transmitted weighted by request size.

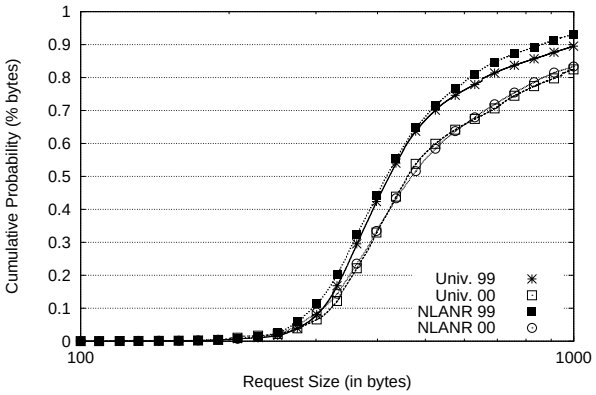


Figure 12: Cumulative distribution of request bytes transmitted weighted by request size <1,000 bytes.

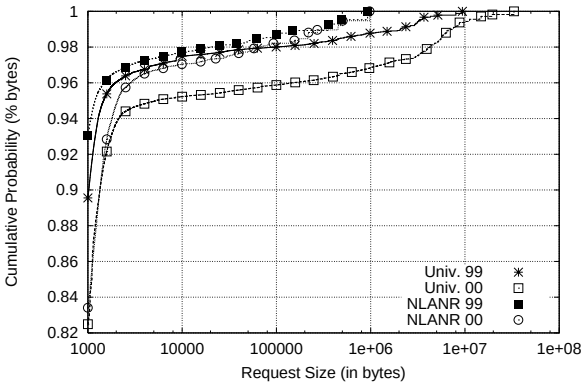


Figure 13: Cumulative distribution of request bytes transmitted weighted by request size >1,000 bytes.

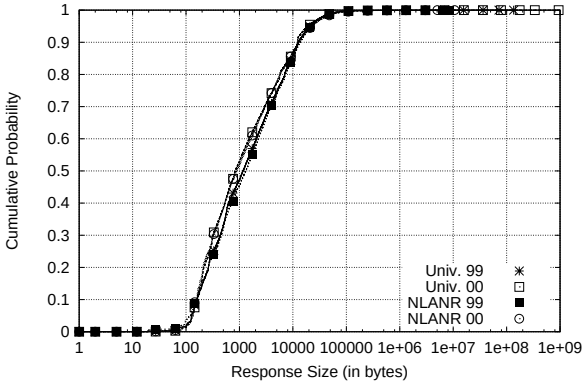


Figure 14: Cumulative distribution of response data sizes.

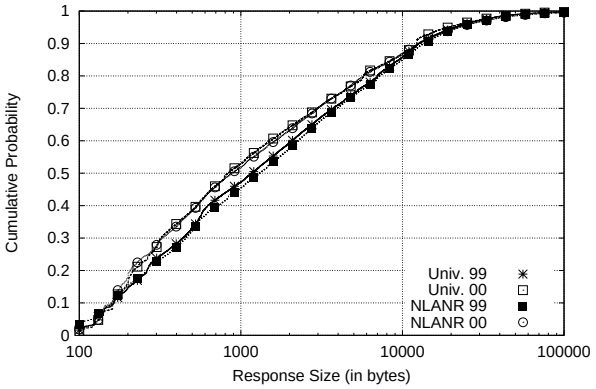


Figure 15: Cumulative distribution of response data sizes (100 – 100,000 bytes).

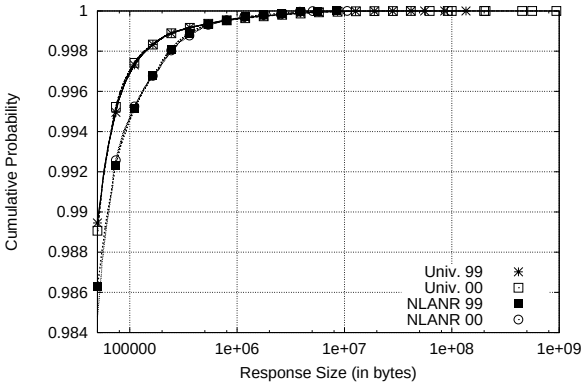


Figure 16: Cumulative distribution of response data sizes > 50,000 bytes.

Table 4: Sizes in bytes of the ten largest responses in each set of traces.

University-99	University-00	NLANR-99	NLANR-00
50,719,045	206,245,341	4,550,940	4,533,040
54,018,012	336,847,821	4,600,268	4,637,561
55,836,918	451,132,232	4,878,659	4,703,281
67,856,932	476,597,590	4,926,040	4,726,020
71,942,389	478,029,615	5,511,166	4,924,580
84,533,622	485,477,754	5,613,872	5,005,312
85,885,308	496,845,215	6,388,668	5,053,449
95,492,044	559,874,428	6,435,388	9,087,873
99,618,720	828,097,371	7,723,409	10,229,998
135,294,044	940,341,530	8,463,773	10,611,470

and NLANR data, we see some shift to a greater proportion of smaller response objects. For example, in the university-99 traces about 47% of responses were 1000 bytes or smaller while in the 2000 traces, about 53% of the responses were 1000 bytes or less. Data from future years will be needed before determining if there is a definite shift in the relative proportions of object sizes in the 200-10,000 byte range.

For response sizes greater than 50,000 bytes, we found that the very largest objects increased significantly in size. This probably reflects the increasing use of HTTP instead of FTP to distribute large files.[7] Table 4 shows the sizes of the 10 largest response elements in each of the trace collections. The sizes of the very largest responses in the university traces are bigger in 2000 by a factor of 4-7 while there appears to be no change for the NLANR traces. The sizes in the NLANR traces are smaller by almost two orders of magnitude. We believe that much of this difference in the NLANR data is not due to real changes in the sizes of responses but is an artifact of the tracing environment. This illustrates an important methodological point — the ability to collect data on very large response objects is influenced by the trace interval and link speeds. In the case of the NLANR traces, the nominal trace interval of 90 seconds is just too short to capture large responses entirely. For the university traces, it is likely that tracing for one hour on a gigabit-speed link results in a greater chance of observing larger responses than tracing for one hour on an OC-3 link. It will be interesting to see if the sizes of the very largest responses increase when we trace for eight hours or more on gigabit-speed links.

We found significant differences in the distributions of response sizes for top-level and embedded objects. Figure 17 shows the overall distributions of top-level and embedded object sizes from the university traces while Figures 18 and 19 zoom in on portions of the response size range. We see clear indications that top-level objects tend to be larger than embedded objects. For example, about 30% of top-level objects are larger than 10,000 bytes while only 10% of embedded objects are. About 50% of top-level objects are smaller than 2,000 bytes while 70% of embedded objects are smaller than 2,000 bytes. This observation is consistent with the number of complex web pages we all see that are composed with embedded objects for icons or advertisements. Figure 18 shows indications of a year-to-year trend toward larger proportions of smaller objects in the range of sizes between 100 and 10,000 bytes.

[7] One specific example is CD images for software distribution (e.g., Linux releases)

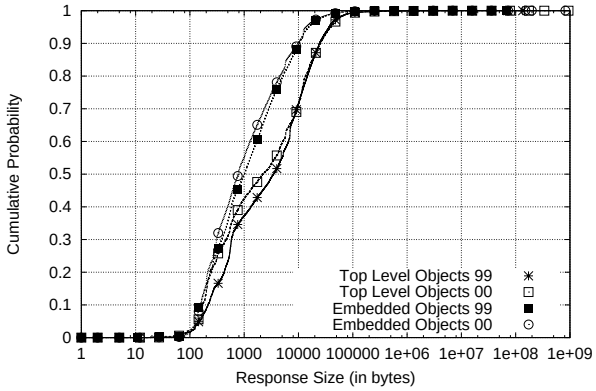


Figure 17: Cumulative distribution of response data sizes for top-level and embedded objects.

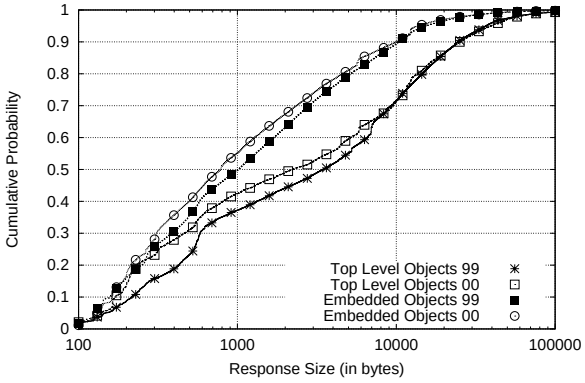


Figure 18: Cumulative distribution of response data sizes for top-level and embedded object sizes (100 – 100,000 bytes).

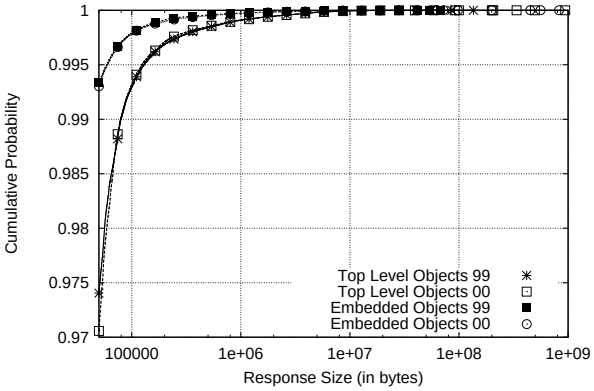
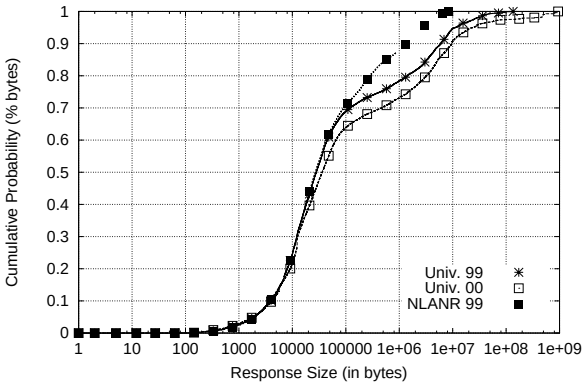


Figure 19: Cumulative distribution of response data sizes for top-level and embedded object sizes >50,000 bytes.

While we have found that 85% of all responses are 10,000 bytes or less, these responses account for only about 20% of the bytes actually transferred from servers to clients, while responses larger than 100,000 bytes represent over 35% of the bytes transferred from servers. Figures 20, 21, and 22 give different views of the cumulative percentage of total bytes in responses as a function of response size. Figure 20 gives the overall contribution of different response sizes to the total bytes returned by servers. The NLANR traces are quite similar to the university traces except above 100,000 bytes where the truncated



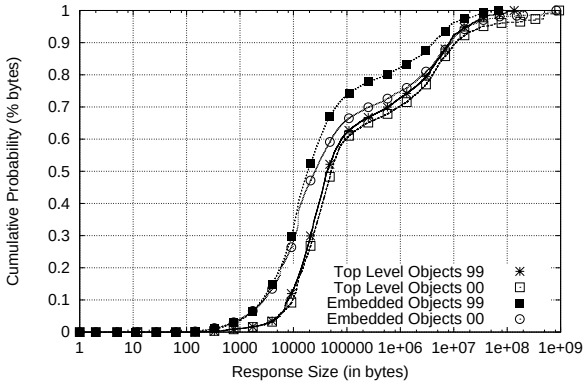


Figure 22: Cumulative distribution of response bytes transmitted weighted by response size for top-level and embedded objects

trace intervals limit the observation of large responses. There is a clear influence of the larger response sizes in the 2000 university traces when compared to 1999. In Figure 21 we see evidence that larger objects account for a higher proportion of the total bytes transmitted on non-persistent connections. Figure 22 confirms our observation that top-level objects are larger than embedded objects and also indicates embedded objects larger than 10,000 bytes may be increasing in size year-to-year. For example, only 10% of the total bytes from top-level objects are from objects smaller than 10,000 bytes while about 30% of the bytes from embedded objects are from those smaller than 10,000 bytes.

8. User Activity and Web Content Structure

We now report briefly on interesting results that provide some insight into how users access web objects and how web content is distributed among servers. In interpreting these data it is important to keep in mind that all the web traffic we observed represents only objects that could not be obtained from local browser caches**.** Figure 23 gives the distribution of top-level page requests per "user" observed in the traces. A "user" is defined as a unique client IP address and the same IP address in different traces is counted as a different user. We counted 140,522 users for the 1999 traces and 244,025 users in 2000. A slight majority of users were observed to make more than 10 top-level page requests in the one hour tracing intervals. There were some users, however, (about 5%) that made more than 100 page requests in one hour. We also observed a slight trend between 1999 and 2000 for users to make fewer requests. Figure 24 gives the distribution of consecutive top-level page requests by a given user to the same server IP address. About 68% of the consecutive top-level page references by the same user go to a different server IP address than the immediately prior reference. We believe these results reflect the basic organization of web servers for a web site into "server farms" for load balancing and content distribution.

Web page structure is reflected in Figure 25 that shows the distribution of objects (top-level plus embedded) per page. While the proportion of quite simple pages is large (70% of pages are 3 or fewer objects), there are significant numbers of pages with complex structure. It is again important to keep in mind that many objects in pages that a user views often may be cacheable and not be included in these traces.

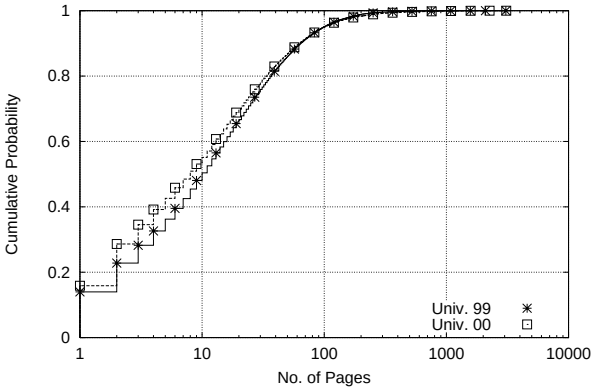


Figure 23: Cumulative distribution of page requests per user.

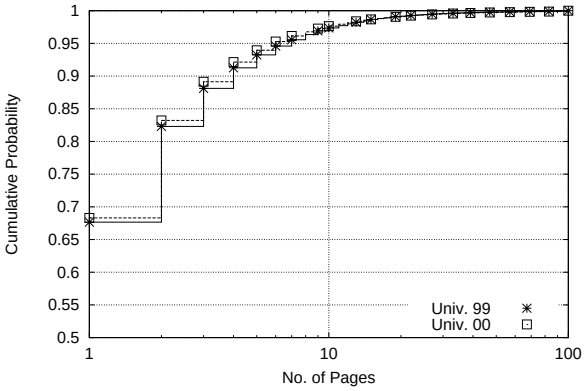


Figure 24: Cumulative distribution of consecutive page requests by one user to the same server.

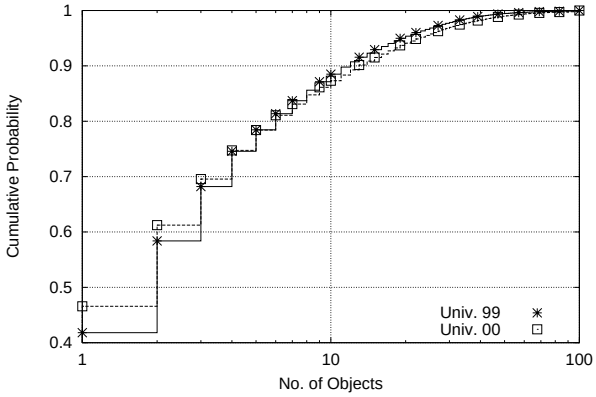


Figure 25: Cumulative distribution of objects per page.

9. Generating Synthetic Web Traffic for Network Simulations and Laboratory Experiments

Our goal here is to sketch a simple design for a Web traffic generator that uses the data presented here to create realistic web traffic. This model is based solely on the characteristics observed about TCP connections and does not depend on characterizing features that are more directly related to application data or usage such as pages and embedded objects. The traffic generating program based on our models simply initiates TCP connections at the desired arrival rate based on the work by Cleveland *et al.* (see [10] for the details of generating TCP connection start times). For each TCP connection, a critical influence on its behavior is the round-trip time (RTT) along the path(s) between the two end-points. We assume that the association of the connection with a path between two endpoints depends on the particular simulated or experimental environment and can be handled as an extension to the generator sketched here. Typical ways of selecting the path are based either on the desired distribution of path RTTs or on the desired distribution of traffic loads over paths in the network.

When the TCP connection is started, a random choice between persistent and non-persistent connection types is made using the data in Table 2. If the connection is to be non-persistent, a request size and a response size for the single exchange are chosen by sampling from the distributions plotted in Figures 9 and 14, respectively. The appropriate number of bytes is transferred in each direction by whatever means is used in the network (real or simulated) for injecting application data — typically by socket write calls. The connection is then closed. If the connection is to be persistent, the behavior is slightly more compli-

cated. A number of request/response exchanges is selected for the connection by sampling from the distribution plotted in Figure 4. To generate each request/response exchange, a request size and response size are chosen by sampling from the distributions plotted in Figures 9 and 14, respectively.[8] When all the exchanges selected for this persistent connection are completed, the connection is closed.

Using the data reported here, we can also construct a traffic generating program using Mah's modeling framework based on user document requests and document structure [23]. The necessary distributions in addition to those discussed above are plotted in Figure 3 (Mah's "think time"), Figure 24 (Mah's "consecutive document retrievals"), and Figure 25 (Mah's "document size"). Figures 3, 14, 24, 25 also represent distributions that are elements of the modeling framework used by Barford and Crovella [3, 4].

10. Summary and Conclusions

By any measure, web traffic is the single largest identifiable consumer of bandwidth on the Internet. A contemporaneous characterization of web traffic is therefore important for driving network simulations and live testing of network components such as congestion control mechanisms. While previous models of web traffic have been presented in the literature, it is important to continually re-populate these models with new data and update the models to account for the evolution of protocols and their use. We are advocating the use of a lightweight measurement methodology for gathering these data based on capturing TCP packet headers. The method balances the tradeoff between ensuring the privacy of users and gathering sufficient data to capture the HTTP protocol dynamics. We have developed a sufficient set of tools for processing multi-gigabyte *tcpdump* files to create distributions of the major structural elements of an HTTP connection. We have conducted two measurement studies of HTTP traffic arriving at the ingress router on the campus of a major research university. We compare the characteristics of traces of HTTP connections recorded during September and October in 1999 and 2000. We also compare our traces to set of similar (but smaller) traces acquired by NLANR during the same periods as our traces.

From the analysis of our trace data, and the NLANR data, we draw a number of conclusions about the evolving nature of web traffic, the use of the web, the structure of web pages and the organization of web servers that provide their content, and the use of packet tracing as a method of understanding the dynamics of web traffic.

From a methodology standpoint, we conclude that a substantial and detailed analysis of HTTP request/response exchanges is possible given only the TCP/IP headers of packets sent from servers to clients. Given only a unidirectional trace, one can discern transport protocol phenomena and the effects of higher-level protocol usage such as the use of persistent HTTP connections, as well as application-level phenomena such as the number of embedded objects per web page and the distribution of servers delivering this content. We believe that future traces of web traffic must be considerably longer (on the order of hours) in order to fully capture the tail of the distribution for measures such as the size of responses from servers to clients.

---

[8] Browser induced delays between request/response exchanges on a persistent connection can also be modeled, if needed, using distributions which have been derived from the traces but not presented here.

From a protocol usage standpoint, we observed that 15% of all the TCP connections carrying HTTP request/response exchanges are persistent in the sense that they actually carry more than one request/response exchange. These persistent connections deliver 40-50% of all the web objects requested and these objects account for approximately 40% of all the bytes transferred. Thus, although the fraction of connections that are persistent is small, their use represents a 50% reduction in the total number of TCP connections required to deliver web content (compared to a similar environment in which persistent connections are not used).

From an analysis of HTTP responses we see that 65% of all web pages are constructed entirely by responses from a single server while 35% of the pages requested receive content from two or more servers. Close to 70% of the consecutive top-level page references go to an IP address that is different from the address used for the previous top-level page reference possibly representing the impact of large organizations managing their web sites with a "server farm" for load balancing. We also see an increase in the number of embedded objects per web page but a decrease in frequently occurring sizes. This is possibly due to the pervasive use of "banner adds" and icons to decorate pages. Overall the number of bytes transferred to deliver embedded objects is increasing.

Overall we see a slight decrease in the frequently occurring sizes of response objects but a marked increase in the size of the largest objects transferred. The largest 15% of object sizes account for 80% of the bytes transferred from servers. From an analysis of HTTP requests we see novel uses of web requests to implement applications such as "web email" that result in a shift in the distribution of request sizes because files are being transferred (*e.g.*, email attachments) in the context of an HTTP request. This also gives rise to HTTP request/response exchanges in which the request is orders of magnitude larger than the corresponding response.

Researchers interested in repeating our study on their campus or enterprise network may find copies of our *tcpdump* processing and analysis tools at *http://www.xxx.yyy.zzz/*.[9] The empirical distributions reported here are also available in a form suitable for generating random variates in traffic generators. Sample traffic generating programs are also available at this site.

11. References

[1] J. Apisdorf, K Claffy, K. Thompson, and R. Wilder. *OC3mon: Flexible, Affordable, High-Performance Statistics Collection*, Proceedings of INET '97, June 1997. (*http://www.isoc.org/isoc/whatis/conferences/inet/97/proceedings/F1/F1_2.HTM*)

[2] H. Balakrishnan, M. Stemm, S. Seshan, V. Padmanabhan, R. H. Katz, *TCP Behavior of a Busy Internet Server: Analysis and Solutions*, Proceedings of IEEE INFOCOMM '98, March 1998, pp. 252-262.

[3] P. Barford and M. E. Crovella, *Generating Representative Web Workloads for Network and Server Performance Evaluation*, Proceedings of ACM SIGMETRICS '98, 1998, pp. 151-160.

[4] P. Barford and M. E. Crovella, *A Performance Evaluation of HyperText Transfer Protocols*, Proceedings of ACM SIGMETRICS '99, May 1999, pp. 188-197.

[9] SIGMETRICS submission guidelines on maintaining anonymity require hiding the actual URL until after the review process.

[5] P. Barford, A. Bestavros, A. Bradley, and M. E. Crovella, *Changes in Web Client Access Patterns: Characteristics and Caching Implications*, World Wide Web, Special Issue on Characterization and Performance Evaluation, Vol. 2, 1999, pp. 15-28.

[6] L. Breslau, D. Estrin, K. Fall, S. Floyd, J. Heidemann, A. Helmy, P. Huang, S. McCanne, K. Varadhan, Y. Xu, and H. Yu, *Advances in Network Simulation*, IEEE Computer, Vol. 33 No. 5, May 2000, pp. 59-67.

[7] R. Caceres, P. Danzig, S. Jamin, and D. Mitzel, *Characteristics of Wide-Area TCP/IP Conversations*, Proceedings of ACM SIGCOMM '91, pp. 101-112.

[8] M. Christiansen, K. Jeffay, D. Ott, and F. D. Smith, *Tuning RED for Web Traffic*, Proceedings of ACM SIGCOMM 2000, September 2000, pp. 139-150.

[9] K Claffy, G. Miller, and K. Thompson. *The nature of the beast: recent traffic measurements from an Internet backbone*, Proccedings of INET '98, (*http://www.isoc.org/inet98/proceedings/6g/6g_3.htm*).

[10] W. S. Cleveland, D. Lin, D. X. Sun, *IP Packet Generation: Statistical Models for TCP Start Times Based on Connection-Rate Superposition*, Proceedings of ACM SIGMETRICS 2000, June 2000, pp.166-177.

[11] Crovella, M. and A. Bestavros, *Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes*, IEEE/ACM Transactions on Networking, vol. 5, no. 6, December 1997, pp. 835-846.

[12] C. R. Cunha, A. Bestavros, and M. E. Crovella, *Characteristics of WWW Client-based Traces*, Technical Report TR-95-010, Boston University Computer Science Department, June 1995.

[13] P. Danzig and S. Jamin, *tcplib: A Library of TCP Internetwork Traffic Characteristics*, USC Technical Report USC-CS-91-495, 1991.

[14] P. Danzig, S. Jamin, R. Caceres, D. Mitzel, and D. Estrin, *An Empirical Workload Model for Driving Wide-Area TCP/IP Network Simulations*, Internetworking: Research and Experience, vol. 3, no. 1, 1992, pp. 1-26.

[15] F. Douglis, A. Feldmann, B. Krishnamurthy, and J. Mogul, *Rate of Change and other Metrics: a Live Study of the World Wide Web*, Proceedings of the USENIX Symposium on Internet Technologies and Systems, December 1997.

[16] B. M. Duska, D. Marwood, and M. J. Feeley, *The Measured Access Characteristics of World-Wide-Web Client Proxy Caches*, Proceedings of the USENIX Symposium on Internet Technologies and Systems, December 1997, pp. 23-36.

[17] L. Fan, P. Cao, J. Almeida and A. Broder, *Summary Cache: A Scalable Wide-Area Web Cache Sharing Protocol*, Proceedings of ACM SIGCOMM '98, pp. 254-265.

[18] A.Feldmann, *BLT: Bi-Layer Tracing of HTTP and TCP/IP*, Proceedings of WWW-9, May 2000. (*http://www.research.att.com/~anja/feldmann/blt_httptrace.abs.html*)

[19] A. Feldmann, R. Caceres, F. Douglis, G. Glass, M. Rabinovich, *Performance of Web Proxy Caching in Heterogeneous Bandwidth Environments*, Proceedings of IEEE INFOCOM '99, 1999, pp. 107-116.

[20] A. Feldmann, J. Rexford, and R. Caceres, *Efficient policies for carrying Web traffic over flow-switched networks*, IEEE/ACM Transactions on Networking, Dec. 1998, pp. 673-685.

[21] W. Feng, D. Kandlur, D. Saha, K. Shin, *Blue: A New Class of Active Queue Management Algorithms*, University of Michigan Technical Report CSE-TR-387-99, April 1999.

[22] S.D. Gribble and E.A. Brewer, *System Design Issues for Internet Middleware Services*, Proceedings of the USENIX Symposium on Internet Technologies and Systems, December 1997.

[23] B. Mah. *An Empirical Model of HTTP Network Traffic*, Proceedings of IEEE INFOCOM '97, April 1997. (an extended version is at http://www.ca.sandia.gov/~bmah/Software/HttpModel/)

[24] G. R. Malan and F. Jahanian, *An Extensible Probe Architecture for Network Protocol Performance Measurement*, Proceedings of ACM SIGCOMM '98, September 1998, pp. 215-227.

[25] S. McCreary, K. C. Claffy, *Trends in Wide Area IP Traffic Patterns: A View from Ames Internet Exchange*, CAIDA Technical Report (*http://www.caida.org/outreach/papers/AIX0005/*).

[26] J. Mogul, *The Case for Persistent-Connection HTTP*, Proceedings of ACM SIGCOMM '95, pp. 299-313.

[27] J. Mogul, F. Douglis, A. Feldmann, and B. Krishnamurthy, *Potential benefits of delta-encoding and data compression for HTTP*, Proceedings of ACM SIGCOMM '97, September 1997, pp. 181-194.

[28] H. Nielsen, J. Gettys, A. Baird-Smith, E. Prud'hommeaux, H. Lie, and C. Lilley, *Network Performance Effects of HTTP/1.1, CSS1, and PNG*, Proceedings of ACM SIGCOMM '97, September 1997, pp. 155-166.

[29] T. Ott, T. Lakshman, and L. Wong, *SRED: Stabilized RED*, Proceedings IEEE INFOCOM '99, March 1999, pp. 1346-1355.

[30] V. Paxson, *Empirically Derived Analytic Models of Wide-Area TCP Connections*, IEEE/ACM Transactions on Networking, vol. 2, no. 4, August 1994, pp. 316-336.

[31] V. Paxson, and S. Floyd, *Wide-Area Traffic: The Failure of Poisson Modeling*, IEEE/ACM Transactions on Networking, vol. 3, no. 3, June 1995, pp. 226-244.

[32] V. Paxson, and S. Floyd, *Why We Don't Know How To Simulate The Internet*, Proceedings of the 1997 Winter Simulation Conference, December 1997.

[33] K. Thompson, G. Miller, and R. Wilder, *Wide-Area Internet Traffic Patterns and Characteristics*, IEEE Network, vol. 11 no. 6, November/December 1997.

[34] Z. Wang and P. Cao, *Persistent Connection Behavior of Popular Browsers*, December 1998, (*http://www.cs.wisc.edu/~cao/papers/persistent-connection.html*)

[35] A. Wolman, *et al*., *On the Scale and Performance of Cooperative Web Proxy Caching*, Proceedings of ACM SOSP '99, December 1999, pp. 16-31.

[36] A. Wolman, G. Voelker, N. Sharma, N. Cardwell, M. Brown, T. Landray, D. Pinnel, A. Karlin, and H. Levy, *Organization-Based Analysis of Web-Object Sharing and Caching*, Proceedings of the 2nd USENIX Conference on Internet Technologies and Systems, October 1999.

[37] *http://www.caida.org/tools/measurement/coralreef/*

[38] *http://moat.nlanr.net/Traces/*

[39] *http://www.napster.com*