

LBL-40319
UCB//CSD-97-945

Measurements and Analysis of End-to-End Internet Dynamics

Vern Paxson
Ph.D. Thesis

Computer Science Division
University of California, Berkeley

and

Information and Computing Sciences Division
Lawrence Berkeley National Laboratory
University of California
Berkeley, CA 94720

April, 1997

This work was supported by the Director, Office of Energy Research, Scientific Computing Staff, of the United States Department of Energy under Contract No. DE-AC03-76SF00098.

Measurements and Analysis of End-to-End Internet Dynamics

by

Vern Edward Paxson

B.S. (Stanford University) 1985

M.S. (University of California, Berkeley) 1991

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Computer Science

in the

GRADUATE DIVISION

of the

UNIVERSITY of CALIFORNIA at BERKELEY

Committee in charge:

Prof. Domenico Ferrari, Chair

Prof. Michael Luby

Prof. John Rice

1997

Measurements and Analysis of End-to-End Internet Dynamics

Copyright 1997

by

Vern Edward Paxson

The U.S. Department of Energy has the right to use this document
for any purpose whatsoever including the right to reproduce
all or any part thereof

Abstract

Measurements and Analysis of End-to-End Internet Dynamics

by

Vern Edward Paxson

Doctor of Philosophy in Computer Science

University of California at Berkeley

Prof. Domenico Ferrari, Chair

Accurately characterizing end-to-end Internet dynamics—the performance that a user actually obtains from the lengthy series of network links that comprise a path through the Internet—is exceptionally difficult, due to the network's immense heterogeneity. It can be impossible to gauge the generality of findings based on measurements of a handful of paths, yet logistically it has proven very difficult to obtain end-to-end measurements on larger scales.

At the heart of our work is a “measurement framework” we devised in which a number of sites around the Internet host a specialized measurement service. By coordinating “probes” between pairs of these sites we can measure end-to-end behavior along $O(N^2)$ paths for a framework consisting of N sites. Consequently, we obtain a superlinear scaling that allows us to measure a rich cross-section of Internet behavior without requiring huge numbers of observation points. 37 sites participated in our study, allowing us to measure more than 1,000 distinct Internet paths.

The first part of our work looks at the behavior of end-to-end routing: the series of routers over which a connection's packets travel. Based on 40,000 measurements made using our framework, we analyze: routing “pathologies” such as loops, outages, and flutter; the stability of routes over time; and the symmetry of routing along the two directions of an end-to-end path. We find that pathologies increased significantly over the course of 1995, indicating that, by one metric, routing *degraded* over the year; that Internet paths are heavily dominated by a single route, but that routing lifetimes range from seconds to many days, with most lasting for days; and that, at the end of 1995, about half of all Internet paths included a major routing asymmetry.

The second part of our work studies end-to-end Internet packet dynamics. We analyze 20,000 TCP transfers of 100 Kbyte each to investigate the performance of both the TCP endpoints and the Internet paths. The measurements used for this part of our study are much richer than those for the first part, but require a great degree of attention to issues of *calibration*, which we address by applying *self-consistency checks* to the measurements whenever possible. We find that packet filters are capable of a wide range of measurement errors, some of which, if undetected, can significantly taint subsequent analysis. We further find that network clocks exhibit adjustments and skews relative to other clocks frequently enough that a failure to detect and remove these effects will likewise pollute subsequent packet timing analysis.

Using TCP transfers for our network path “measurement probes” gains a number of advantages, the chief of which is the ability to probe fine time scales without unduly loading the network. However, using TCP also requires us to accurately distinguish between connection dy-

namics due to the behavior of the TCP endpoints, and dynamics due to the behavior of the network path between them. To address this problem, we develop an analysis program, `tcpanaly`, that has coded into it knowledge of how the different TCP implementations in our study function. In the process of developing `tcpanaly`, we thus in tandem develop detailed descriptions of the performance and congestion-avoidance behavior of the different implementations. We find that some of the implementations suffer from gross problems, the most serious of which would devastate overall Internet performance, were the implementations ubiquitously deployed.

With the measurements calibrated and the TCP behavior understood, we then can turn to analyzing the dynamics of Internet paths. We first need to determine a path's *bottleneck bandwidth*, meaning the fastest transfer rate the path can sustain. Knowing the bottleneck bandwidth then lets us determine which packets a sender transmits must necessarily *queue* behind their predecessors, due to the load the sender imposes on the path. This in turn allows us to determine which of our probes are perforce *correlated*. We identify several problems with the existing bottleneck estimation technique, “packet pair,” and devise a robust estimation algorithm, PBM (“packet bunch modes”), that addresses these difficulties. We calibrate PBM by gauging the degree to which the bottleneck rates it estimates accord with known link speeds, and find good agreement. We then characterize the scope of Internet path bottleneck rates, finding wide variation, not infrequent asymmetries, but considerable stability over time.

We next turn to an analysis of packet loss along Internet paths. To do so, we distinguish between losses of “loaded” data packets, meaning those which necessarily queued behind a predecessor at the bottleneck; “unloaded” data packets, which did not do so; and the small “acknowledgement” packets returned to a TCP sender by the TCP receiver. We find that network paths are well characterized by two general states, “quiescent,” in which no loss occurs, and “busy,” in which one or more packets of a connection are lost. The prevalence of quiescent connections remained about 50% in both our datasets, but for busy connections, packet loss rates increased significantly over the course of 1995. We further find that loss rates vary dramatically between different regions of the network, with European and especially trans-Atlantic connections faring much worse than those confined to the United States.

We also characterize: loss symmetry, finding that loss rates along the two directions of an Internet path are nearly uncorrelated; loss “outages,” finding that outage durations exhibit clear Pareto distributions, indicating they span a large range of time scales; the degree to which a connection's loss patterns predict those of future connections, finding that observing quiescence is a good predictor of observing quiescence in the future, and likewise for observing a busy network path, but that the proportion of lost packets does not well predict the future proportion; and the efficacy of TCP implementations in dealing with loss efficiently, by retransmitting only when necessary. We find that most TCPs retransmit fairly efficiently, and that deploying the proposed “selective acknowledgement” option would eliminate almost all of their remaining unnecessary retransmissions. However, some TCPs incorrectly determine how long to wait before retransmitting, and these can suffer large numbers of unnecessary retransmissions.

We finish our study with a look at variations in packet transit delays. We find great “peak-to-peak” variation, meaning that maximum delays far exceed minimum delays. Delay variations along the two directions of an Internet path are only lightly correlated, but correlate well with loss rates observed in the same direction along the path. We identify three types of “timing compression,” in which packets arrive at their receiver spaced more closely together than when originally

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.