Addressing the Challenges of Web Data Transport

by

Venkata Narayana Padmanabhan

B.Tech. (Indian Institute of Technology, Delhi) 1993
M.S. (University of California at Berkeley) 1995

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:

Professor Randy H. Katz, Chair
Professor Steven R. McCanne
Professor J. George Shanthikumar

1998

Addressing the Challenges of Web Data Transport

by

Venkata Narayana Padmanabhan

Abstract

Addressing the Challenges of Web Data Transport

by

Venkata Narayana Padmanabhan

Doctor of Philosophy in Computer Science

University of California at Berkeley

Professor Randy H. Katz, Chair

In just a few years since its inception, the World Wide Web has grown to be the most dominant application in the Internet. In large measure, this rapid growth is due to the Web's convenient point-and-click interface and its appealing graphical content. Since Web browsing is an interactive activity, minimizing user-perceived latency is an important goal. However, layering Web data transport on top of the TCP protocol poses several challenges to achieving this goal.

First, the transmission of a Web page from a server to a client involves the transfer of multiple distinct components, each in itself of some value to the user. To minimize user-perceived latency, it is desirable to transfer the components concurrently. TCP provides an ordered byte-stream abstraction with no mechanism to demarcate sub-streams. If a separate TCP connection is used for each component, as with HTTP/1.0, uncoordinated competition among the connections could exacerbate congestion, packet loss, unfairness, and latency.

Second, Web data transfers happen in relatively short bursts, with intervening idle periods. It is difficult to utilize bandwidth effectively during a short burst because discovering how much bandwidth is available requires time. Latency suffers as a consequence.

To address these problems, we first developed a new connection abstraction for HTTP, called *persistent-connection HTTP* (*P-HTTP*). The key ideas are to share a persistent TCP connection for multiple Web page components and to pipeline the transfers of these components to reduce latency. These ideas, developed by us in 1994, have been adopted by the HTTP/1.1 protocol. The main drawback of P-HTTP, though, is that the persistent TCP connection imposes a linear ordering on the Web page components, which are inherently independent.

This drawback of P-HTTP led us to develop a comprehensive solution, which has two components. The first component, *TCP session*, decouples TCP's ordered byte-stream service abstraction from its congestion control and loss recovery mechanisms. It *integrates* the latter mechanisms across the set of concurrent connections between a pair of hosts, thereby combining the flexibility of separate connections with the performance efficiency of a shared connection. This integration decreases download time by up to a factor of ten compared to HTTP/1.0 layered on standard TCP. TCP session does not alter TCP's messaging semantics, so deployment only involves local changes at the sender.

The second component of our solution, *TCP fast start*, improves bandwidth utilization for short transfers by reusing information about the network conditions cached in the recent past. To avoid adverse effects in case the cached information is stale, TCP fast start exploits priority dropping at routers, and augments TCP's loss recovery mechanisms to quickly detect and abort a failed fast start attempt.

In addition to the two challenges we have discussed, a third challenge arises from the increasing deployment of *asymmetric access networks*. Although Web browsing has asymmetric bandwidth requirements, bandwidth asymmetry could adversely impact Web download performance due to a disruption in the flow of acknowledgement feedback that is critical to sustaining good TCP throughput. To avoid performance degradation, we have developed end-host and router-based techniques, that both reduce disruption in the feedback and reduce TCP's dependence on such feedback. In certain situations, these techniques help decrease download time by a factor of fifteen.

This thesis includes mathematical and trace-based analysis, simulation, implementation and performance evaluation. In addition to the algorithms that we have designed and software that we have developed, our contributions include a set of paradigms for advancing the state-of-the-art in Internet transport protocols. These paradigms include the use of shared state and/or persistent state, and the exploitation of differentiated services mechanisms in routers.

Professor Randy H. Katz, Chair

To my parents and my sisters, Priya and Sumi.

Vidya dadaathi vinayam

vinayaath yaathi paathrathaam

paathrathwaath dhanam apnothi

dhanaath dharmam thatha sukham

(ancient Sanskrit saying)

English Translation

Education gives modesty;

through modesty one becomes worthy;

through worthiness one gets wealth;

through wealth one can be righteous and happy.

Table of Contents

List of Figures

List of Tables

Acknowledgements

As I stand at the threshold of earning my doctorate, I am overwhelmed when I recall all the people who have helped me get this far.

First and foremost, I would like to thank my Ph.D. advisor, Professor Randy Katz, for his constant support, guidance, and inspiration. Randy is a truly remarkable advisor who grants students a lot of freedom to explore new ideas, but at the same time interacts closely with them. I have greatly benefited both from his excellent technical advice and from his role-model as a successful team leader and researcher. I look forward to continuing my association with him in the future.

I am indebted to Jeff Mogul for introducing me to the problem of improving HTTP performance, which eventually led to my dissertation. Jeff has been a great mentor and collaborator. I would like to thank him and the Compaq (formerly Digital) Western Research Laboratory for providing me generous support, beginning with a summer internship, and continuing with access to their computing facilities as an honorary researcher.

I would like to thank Professor Steve McCanne for being on my qualifying exam and thesis committees. Steve's insightful comments and constructive criticisms have been extremely useful to me. I am also grateful to Professors George Shanthikumar and Bob Brodersen for serving on my committees and providing valuable feedback on the ideas presented in this thesis.

I have had very productive collaborations with Hari Balakrishnan, Jeff Mogul, Srini Seshan and Mark Stemm, which have contributed significantly to this thesis. I have also had valuable discussions on various aspects of the thesis with B. R. Badrinath, Sally Floyd, Tom Henderson, T. V. Lakshman, Giao Nguyen, Vern Paxson, and K. K. Ramakrishnan. My graduate student career has also been enriched by interactions with several researchers including Mary Baker, Eric Brewer, Ramón Cáceres, Domenico Ferrari, Akthar Jameel, Barry Leiner, Reiner Ludwig, Kevin Mills, Rob Ruth, Subir Varma, and Yongguang Zhang. I would like to thank them all.

This work was supported primarily by DARPA under contract DAAB07-95-C-D154. It was also supported by grants from Hughes Aircraft Corporation, Hybrid Networks Inc., Metricom Inc. and the California MICRO program. Son Dao and Yongguang Zhang of Hughes Research Labs, Ed Moura of Hybrid Inc., and Mike Ritter of Metricom Inc. helped set up the network testbed for my experimental work. Brian Shiratsuki, Keith Sklower and Ken Lutz were always willing to help straighten out any kinks that developed in the testbed. Keith Sklower was largely responsible for getting our connection to the Hughes DirecPC network off the ground. I am grateful to them all.

Kathryn Crabtree was ever helpful in guiding me through the administrative labyrinth in Berkeley. I am convinced that every research organization should have someone as adept as her in keeping bureaucratic red tape at bay! Terry Lessard-Smith and Bob Miller were great project administrators. I doubt if large systems projects in Berkeley would run half as smoothly as they do without their tireless efforts, or be half as much fun as they are without the excellent project retreats they organize.

My graduate student colleagues have made my stay at Berkeley a truly enjoyable one. Their collective intellectual energy has been a great motivator for me. I would like to thank Elan Amir, Hari Balakrishnan, Yatin Chawathe, Armando Fox, Jeff Gilbert, Steve Gribble, Tom Henderson, Todd Hodes, Dan Jiang, Bruce Mah, Giao Nguyen, Mark Stemm, Helen Wang, and Tao Ye.

The "443 Soda gang" of Elan, Hari, Todd, Mark, and members-at-large, Tom, Giao, and Srini was always an enthusiastic (and leak-proof!) audience for almost any topic under the sun, including many politically-incorrect ones. I certainly hope the gang will continue to flourish in cyberspace (with the protection of PGP encryption, of course!) as the ranks of the members-at-large swells.

I would like to thank all my teachers, from elementary school through graduate school, for providing me with an excellent education. I am particularly indebted to my professors at the Indian Institute of Technology for giving me a solid foundation in Computer Science and Engineering.

Most important of all, I would like to express my gratitude to my family for being an unstinting source of support and encouragement. My grandparents have inspired me through their courage in overcoming the challenges of life. I am saddened that they did not live to see my graduation. My parents have taught me the value of education and have worked hard to provide me the very best of it. They have always been there when I have needed them. My sisters, Priya and Sumi, have been wonderful sources of support. The tinge of jealousy I feel in them both having beaten me to earning the title "Doctor" is effectively neutralised by their abiding affection! Last but not the least, my brothers-in-law, Sekhar and Umesh, have been great friends and have given me excellent advice drawing on their own careers as doctoral students, and subsequently, researchers.