

DESIGNING A PREDICTABLE BACKBONE NETWORK USING
VALIANT LOAD-BALANCING

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL
ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Rui Zhang-Shen

June 2007

© Copyright by Rui Zhang-Shen 2007
All Rights Reserved

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Nick McKeown) Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Balaji Prabhakar)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Ramesh Johari)

Approved for the University Committee on Graduate Studies.

*To my parents,
who taught me the most important lessons.*

Abstract

The backbones form the core of the Internet and carry large volumes of data across long distances. Designing a backbone network with performance guarantees is important but difficult due to the uncertainty in the traffic matrix, the need to accommodate component failures, and the desire to avoid congestion. Backbone networks today are highly over-provisioned but cannot guarantee to support all traffic matrices.

In this thesis we propose using Valiant Load-Balancing (VLB, named after L.G. Valiant) to design backbone networks that can support all traffic matrices efficiently, even under a number of failures.

Chapter 2 shows that VLB is efficient for any network. A simple scheme, the Gravity Full Mesh, requires a total capacity no more than twice the theoretical lower bound, and another scheme, Minimum Network Fanout, requires a total capacity at most 20% more than the theoretical lower bound.

The path diversity in VLB enables the network to tolerate failures efficiently. Chapter 3 shows that in order to tolerate k *arbitrary* failures, the fraction of extra capacity required is only approximately $\frac{k}{N}$. In addition, VLB provides fast rerouting after failure.

In Chapter 4 we first argue that most applications will not be affected by the extra propagation delay caused by VLB. Then we propose adaptive VLB, which adjusts the amount of load-balancing according to the traffic condition in order to minimize packet delay.

In Chapter 5 we propose using VLB to route traffic between two networks. VLB can efficiently utilize the peering links so that there is no congestion unless the total

peering traffic rate exceeds the total peering capacity. It can further ensure that peering packets receive the same quality of service as non-peering traffic. We derive the optimal load-balancing parameters for routing local traffic and peering traffic.

We finally apply VLB to designing circuit-switched networks with performance guarantees. Current circuit-switched networks use heuristics to minimize average blocking probability. The heuristics require parameters that depend on the traffic metrics, are hard to analyze, and cannot give performance guarantees. Valiant Load-Balancing, on the other hand, can give theoretical bounds on blocking probabilities for all flows under worst case traffic.

Acknowledgments

It has been a great privilege to work with my adviser Nick McKeown. Nick is a teacher, a mentor, and a role model. On research he is visionary; on everything else he is extremely understanding. Nick is always there to support, to guide, and to help, with his calming influence and amazing wisdom. I have learned so much from you, and thank you, Nick.

My undergraduate adviser Doug Osheroff is another amazing person I met at Stanford. Thank you, Doug, for inviting me to join Stanford and to join your lab, for guiding me, and for supporting me.

I am lucky to have worked with such great professors as Balaji Prabhakar and Ramesh Johari. Thank you for what you taught me and for being the readers of this thesis. I am also lucky to have worked with some top researchers and engineers outside of Stanford, especially John Chuang, Nasser El-Aawar, Sandy Fraser, Albert Greenberg, Frank Kelly, Murali Kodialam, TV Lakshman, Darren Loher, Jennifer Rexford, and Jennifer Yates. I thank them for what I learned from them.

My special thanks go to Nandita Dukkupati, my office mate and best friend. My graduate school years would not have been the same without you and your support. I will always remember our passionate discussions, work related or not. I also thank the other members of my research group, especially Clay, Da, David, Gireesh, Glenn, Guido, Isaac, Jad, Jianying, John, Justin, Martin, Masa, Neda, Pablo, Pankaj, Sundar, and Yashar, and members of Balaji's group, especially Abtin, Chandra, Damon, Devavrat, Elif, Kostas, Lu Yi, Mei, Paolo, and Rong. I enjoy all the meetings and discussions we had.

I am grateful to the Lillie Family Stanford Graduate Fellowship, the National

Science Foundation, and the Stanford Childbirth Accommodation Fund for their financial support.

I thank Marcia Keating and Judy Polenta for making my life at Stanford a much easier one than it could be, and Dr. Evelin Sullivan for editing part of this thesis.

My gratitude also goes to my earlier teachers, especially Yousheng Shu, Xun Yang, Zengcai Liu, and Xiaobin Li, who not only taught me but also encouraged me to pursue my dreams.

Last but not least, my deepest gratitude goes to my family: my grandparents, my parents, my husband, and my son. You all taught me a lot and your love and support define who I am. I am proud of you all as you are of me.

Contents

Abstract	v
Acknowledgments	vii
1 Introduction	1
1.1 Background	1
1.2 Motivations for Designing a Predictable Backbone	3
1.2.1 Uncertainty in the traffic matrix	4
1.2.2 Difficulty in supporting all traffic matrices	6
1.2.3 The need to accommodate interruptions	10
1.2.4 Peering	10
1.3 Over Provisioning	11
1.4 Valiant Load-Balancing: The Basics	12
1.5 Related Work	15
1.6 Organization of the Thesis	16
2 VLB in Heterogeneous Networks	17
2.1 Introduction	17
2.1.1 Notations	18
2.1.2 Oblivious Load-Balancing	20
2.2 Lower Bound and Optimality of VLB	22
2.2.1 Minimum Total Capacity	22
2.2.2 Optimality of VLB	26
2.3 Designing a Heterogeneous VLB Network	27

2.3.1	Gravity Full Mesh	28
2.3.2	Network Fanout	29
2.3.3	Minimizing Network Fanout	32
2.3.4	Oblivious Minimum Network Fanout and the Lower Bound	35
2.4	Summary and Possible Extensions	38
3	Designing for Failures in a VLB Network	40
3.1	Introduction	40
3.1.1	Background	40
3.1.2	Fault Tolerance in VLB	44
3.2	Node Failures	46
3.3	Link Failures	47
3.3.1	Adversarial link failures	47
3.3.2	Random link failures	49
3.4	Combination of Node and Link Failures	50
3.5	Designing for Any Failure Scenarios	51
3.6	Discussions	53
3.7	Summary	56
4	VLB and Network Latency	58
4.1	Introduction	58
4.2	Adaptive Valiant Load-Balancing	63
4.2.1	Adaptive Load-Balancing with Local Traffic Information	64
4.2.2	Adaptive Load-Balancing with Global Traffic Information	66
5	Using Valiant Load-Balancing in Peering	68
5.1	Introduction	68
5.1.1	Notation	71
5.2	Provisioning Internal Links for Peering Traffic	72
5.3	Finding the Optimal Peering Parameters	76
5.3.1	Finding the Optimal Peering Parameters in a Single VLB Network	76
5.3.2	Finding the Optimal Peering Parameters in Two VLB Networks	78

6	VLB in Circuit-Switched Networks	80
6.1	Introduction	80
6.2	Problem Statement	82
6.3	Oblivious Valiant Load-Balancing	84
6.3.1	Linear Programming with Approximate Blocking Probabilities	85
6.4	Generalized Load-Balancing	91
6.5	Algorithms for Comparison	96
6.6	Simulation Study	98
6.7	Conclusion	104
7	Concluding Remarks	105
A	Arbitrary Link Failures	108
	Bibliography	113

List of Tables

1.1	Percentage of traffic matrices supported	8
6.1	Experimental Comparison of Routing Schemes	101
6.2	Experimental Comparison in the Scaled Down Network	102
6.3	Experimental Comparison in a 20-Node Network	103

List of Figures

1.1	Calculating the percentage of traffic matrices supported by a network	8
1.2	A hierarchical network with N backbone nodes	12
1.3	Valiant Load-Balancing in a network of N identical nodes	13
2.1	Valiant Load-Balancing in a heterogeneous N -node network.	19
2.2	The star topology	25
2.3	The two-star topology	27
2.4	A comparison of the network fanout given by different schemes	36
3.1	The required link capacity vs. the number of link failures	46
3.2	The Abilene network.	54
3.3	Node aggregate traffic rates of the Abilene network.	55
5.1	An example illustrating hot-potato routing	69
5.2	Two VLB networks peer with each other.	71
5.3	Delay sensitive load-balanced peering	73
6.1	The flowchart for Dynamic Alternative Routing.	97
6.2	The 7-node network used in simulations.	99
A.1	The scenario in which $k + 1$ paths are destroyed by k link failures.	110