

# CIRCUIT SWITCHING IN THE INTERNET

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

Pablo Molinero Fernández

June 2003

© Copyright by Pablo Molinero Fernández 2003  
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.

---

Nicholas Bambos

Approved for the University Committee on Graduate Studies:



# Abstract

The motivation for this thesis is our desire to build faster routers and switches to accommodate for the traffic growth in the Internet. For the past few years, Internet traffic has been doubling every year, and nothing indicates that this growth rate will slow down in the near future. The Internet forwards information through packet switching, which has so far proven to scale from the early slow phone modems to the current fast link rates. However, it is unclear whether it will continue scaling to match future optical link rates.

Fiber optics and optical switching elements have demonstrated a capacity to forward information that today looks unattainable by electronic switching elements. As a consequence, one possible way of increasing network capacity is to build all-optical packet switches. However, these switches are not possible today because packet switching requires the buffering and processing of packets, and we do not (yet) know how to perform them in optics. On the other hand, optical circuit switches do not have these constraints, and thus they are already in use. The simplicity of the forwarding path in a circuit switch makes it faster than an equivalent router, even when implemented in electronics. In this thesis, I argue that we would greatly benefit from the use of circuit switching in the core of the network, in either electronic or optical form.

Circuit switching is already used in the Internet. Since the beginning of the Internet it is widely used in the core of the network; when early Internet service providers wanted to interconnect remote sites, the only option was to lease a circuit from the long-distance telephone carrier. Chapters 2 and 3 of this thesis analyze what type of network we would build were we to start with a clean slate. After analysis,

modeling and simulation, I conclude that we would be better off with a hybrid network similar to the current one.

A problem with the current circuits in the core is that they are completely decoupled of packets in the edges. Rather than following traffic patterns in real time, circuits are usually provisioned manually, and thus they change very slowly. IP considers circuits to be static, point-to-point, layer-2 links between routers. Chapters 4 and 5 propose two evolutionary ways of integrating circuit and packet switching, so that circuits are automatically controlled by the traffic carried by IP. The first approach uses lightweight, fine circuits to carry single user flows, whereas the second multiplexes several flows onto heavyweight, coarse circuits.

... un primer axioma para establecer cualquier sistema educativo: Es objetivo primordial e irrenunciable mantener el sentido universal de la Ciencia y no sólo en un aspecto informativo, sino en el creativo de la investigación.

D. Luciano Fernández Penedo en “Momentos estelares de la enseñanza Española”

# Acknowledgements

I would like to thank my advisor Nick McKeown, for the guidance he has provided me through my Ph.D., as well as the members of my reading committee, Balaji Prabhakar and Nick Bambos, and my former advisor, Fouad Tobagi. I am also very grateful to those that helped me improve this thesis by reading it in its early stages, Sundar Iyer, Nandita Dukkupati, Greg Watson and Mary McDevitt. I also would like to thank Hui Zhang for his help in Chapter 2, Byung-Gon Chun for his implementation of a TCP Switch, and Mor Harchol-Balter for her suggestions for the analysis of the CS-SJF discipline in Chapter 3. I want to thank NLANR, Sprint Labs (Chuck Fraleigh and Brian Lyles), CAIDA and Ciena for providing part of the information that have analyzed in this thesis.

I will not forget everybody with whom I been in Gates 342: Youngmi, Pankaj, Amr, Guido, Paul, Yashar, Giulio and Gireesh, and those other members of the research group, Isaac, Rui, Adisak, Da, Steve, Kersten and Mina. Certainly, my life at Stanford has not always been centered around work, and I also would like to remember John, Oskar, Katya, Krishna, Brad, Waël, Athina, Mansour, Victor, Kristin, Lorenz, Chuck, Kostas, Charlie, and Kevin. I also would like to thank the members of the Spanish communities at Stanford, Iberia, and the Bay Area, AESV, specially Juanjo, Alberto, Carlos y Carlos, Victor y Natacha, Victor y Esperanza, José Manuel, Mario, César y Teresa, David y María, Leo, Diego, and Cintya.

Last but not least, I would like to thank those who really made it possible for me to do my Ph.D. at Stanford because of their encouragement and emotional support: my parents, my grandfather, my twin sisters, my brother, my elder sister and my godmother. My love and gratitude to all of you.



# Contents

<b>Abstract</b>	<b>v</b>
<b>Acknowledgements</b>	<b>viii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	2
1.2 Technology trends in routers and switches . . . . .	2
1.2.1 Technology trends . . . . .	5
1.2.2 Optical switching technology . . . . .	9
1.3 Circuit and packet switching . . . . .	11
1.3.1 Virtual circuits . . . . .	13
1.4 Performance metrics for core IP routers . . . . .	14
1.5 Understanding Internet traffic and failures . . . . .	15
1.6 Organization of the Thesis . . . . .	16
<b>2 Circuit and Packet Switching</b>	<b>20</b>
2.1 Introduction . . . . .	20
2.1.1 Organization of the chapter . . . . .	22
2.2 Background and previous work . . . . .	22
2.2.1 Circuit switching . . . . .	23
2.2.2 Packet switching . . . . .	24
2.3 IP Folklore . . . . .	25
2.3.1 IP already dominates global communications . . . . .	26
2.3.2 IP is more efficient . . . . .	28

2.3.3	IP is robust . . . . .	32
2.3.4	IP is simpler . . . . .	34
2.3.5	Cost of ownership of IP is small . . . . .	38
2.3.6	Support of telephony and other real-time applications . . . . .	40
2.4	Discussion . . . . .	42
2.4.1	Dependability of IP networks . . . . .	42
2.4.2	Interaction of IP and circuits . . . . .	43
2.4.3	What if we started with a clean slate? . . . . .	44
2.5	Conclusions and summary of contributions . . . . .	46
<b>3</b>	<b>Response Time of Circuit and Packet Switching</b>	<b>48</b>
3.1	Introduction . . . . .	48
3.1.1	Organization of the chapter . . . . .	50
3.2	Background and previous work . . . . .	50
3.3	LANs and shared access networks . . . . .	51
3.3.1	Example 1: LANs with fixed-size flows . . . . .	52
3.3.2	Example 2: LANs with heavy-tailed flow sizes . . . . .	52
3.3.3	Model for LANs and access networks . . . . .	53
3.4	Core of the Internet . . . . .	57
3.4.1	Example 3: An overprovisioned core of the network . . . . .	57
3.4.2	Example 4: An oversubscribed core of the network . . . . .	59
3.4.3	Model for the core of the Internet . . . . .	60
3.5	Simulation of a real network . . . . .	64
3.6	Discussion . . . . .	68
3.7	Conclusions and summary of contributions . . . . .	69
<b>4</b>	<b>TCP Switching</b>	<b>70</b>
4.1	Introduction . . . . .	70
4.1.1	Organization of the chapter . . . . .	73
4.2	Advantages and pitfalls of circuit switching . . . . .	73
4.2.1	Pitfalls of circuit switching . . . . .	73
4.2.2	State maintenance . . . . .	74

4.2.3	Signaling overhead and latency . . . . .	74
4.2.4	Wasted capacity . . . . .	74
4.2.5	Blocking under congestion . . . . .	75
4.3	TCP Switching . . . . .	75
4.3.1	Typical Internet flows . . . . .	79
4.3.2	Design options . . . . .	81
4.3.3	Design choices . . . . .	83
4.3.4	Experimentation with TCP-Switching networks and nodes . . . . .	86
4.4	Discussion . . . . .	87
4.4.1	Single-packet flows . . . . .	88
4.4.2	Bandwidth inefficiencies . . . . .	88
4.4.3	Denial of service . . . . .	90
4.5	Conclusions and summary of contributions . . . . .	91
<b>5</b>	<b>Coarse circuit switching in the core</b>	<b>92</b>
5.1	Introduction . . . . .	92
5.1.1	Organization of the chapter . . . . .	94
5.2	Background and previous work . . . . .	94
5.3	Monitoring user flows . . . . .	97
5.4	Modeling traffic to help identify the safeguard band . . . . .	101
5.5	Discussion . . . . .	105
5.6	Conclusions and summary of contributions . . . . .	107
<b>6</b>	<b>Related work</b>	<b>109</b>
6.1	Introduction . . . . .	109
6.1.1	Organization of the chapter . . . . .	109
6.2	Circuit switching in the Internet . . . . .	109
6.2.1	Generalized Multi-Protocol Label Switching (GMPLS) . . . . .	110
6.2.2	ASTN: Automatic Switched Transport Network . . . . .	114
6.2.3	OIF: Optical Internetworking Forum . . . . .	115
6.2.4	ODSI: Optical Domain Service Interconnect . . . . .	116
6.2.5	Grid computing and <i>CA*Net 4</i> . . . . .	116

6.2.6	Proposal by Veeraraghavan et al. . . . .	117
6.2.7	IP Switching . . . . .	118
6.3	Packet switching in the optical domain . . . . .	119
6.3.1	Optical Packet Switching (OPS) . . . . .	120
6.3.2	Optical Burst Switching (OBS) . . . . .	121
6.3.3	Performance of OPS/OBS . . . . .	123
6.4	Flow Measurement . . . . .	126
6.4.1	RFC 2722 and NetFlow . . . . .	127
6.4.2	Proposal by Estan and Varghese . . . . .	127
6.5	Conclusions . . . . .	128
<b>7</b>	<b>Conclusions</b>	<b>129</b>
7.1	Future directions . . . . .	131
7.2	Final words . . . . .	132
	<b>Glossary</b>	<b>133</b>
	<b>Bibliography</b>	<b>136</b>

# List of Tables

1.1	Switching capacities of commercial switches . . . . .	10
1.2	Concerns of carriers for network equipment . . . . .	14
1.3	New features required by carriers . . . . .	15
2.1	World telecommunications infrastructure market in 2001 . . . . .	27
2.2	Frequency of failures in an ISP . . . . .	34
2.3	Cost structure for an Internet carrier . . . . .	39
3.1	Average and maximum response times in Example 3.3.1 . . . . .	52
3.2	Average and maximum response times in Example 3.3.2 . . . . .	53
3.3	Average and maximum response times in Example 3.4.1 . . . . .	59
3.4	Average and maximum response times in Example 3.4.2 . . . . .	60
4.1	Typical TCP flows in the Internet . . . . .	80

# List of Figures

1.1	Functionality of a packet switch . . . . .	3
1.2	Functionality of a circuit switch . . . . .	5
1.3	Trends of traffic demand and underlying technologies in the Internet .	8
1.4	Simple architecture of the public Internet . . . . .	11
1.5	Architecture of the public Internet in the real world . . . . .	12
1.6	Heavy-tailed traffic . . . . .	17
2.1	Architecture of the public Internet . . . . .	28
3.1	Network scenario for motivating examples 3.3.1 and 3.3.2 . . . . .	51
3.2	Queueing model used for circuit and packet switching. . . . .	54
3.3	Average response time of CS-FCFS and CS-SJF vs. PS-PrSh for a single bimodal server . . . . .	56
3.4	Average response time of CS-FCFS vs. PS-PrSh for a single Pareto server . . . . .	58
3.5	Network scenario for motivating examples 3.4.1 and 3.4.2 . . . . .	59
3.6	Average response time of CS-FCFS vs. PS-PrSh for N bimodal servers	61
3.7	Average response time of CS-FCFS vs. PS-PrSh for N Pareto servers	62
3.8	Time diagram of three M/Pareto/N/CS-FCFS systems . . . . .	63
3.9	Topology used in the ns-2 simulation . . . . .	65
3.10	Average goodput as a function of the size of the transferred file . . .	66
3.11	Average relative response time vs. the size of the transferred file . . .	67
3.12	Hybrid network architecture recommended in this thesis . . . . .	68

4.1	An example of a TCP-Switching network. . . . .	72
4.2	Time diagram of a TCP connection using TCP Switching . . . . .	76
4.3	Functional block of a TCP-Switching boundary router . . . . .	77
4.4	Functional block of a TCP-Switching core circuit switch . . . . .	78
4.5	cumulative histogram of flow bandwidths for TCP and non-TCP flows	81
4.6	Correlation of lengths and durations for TCP and non-TCP flows . .	82
4.7	Bandwidth inefficiencies in TCP Switching. . . . .	89
5.1	Network topology considered in this chapter. . . . .	93
5.2	Daily, weekly and monthly average traffic . . . . .	95
5.3	Time diagram of the instantaneous link bandwidth and the average flow rate. . . . .	96
5.4	Time diagram of the calculation of the safeguard band . . . . .	99
5.5	Safeguard band vs. overflow probabilities and circuit-creation latencies.	101
5.6	Histogram of the peak-bandwidth envelope. . . . .	102
5.7	Histogram of flow interarrivals. . . . .	103
5.8	Histograms of flow durations and bandwidths . . . . .	104
5.9	Joint histogram of flow durations and average bandwidths. . . . .	105
5.10	Safeguard band according to real traces and a simple model. . . . .	106
6.1	Hierarchy of label-switched paths in GMPLS. . . . .	113
6.2	Network architecture of the Automatic Switched Transport Network (ASTN). . . . .	115
6.3	Time diagram of Optical Burst Switching. . . . .	122
6.4	Topology used to simulate the effect of Optical Packet and Burst Switching on TCP. . . . .	124
6.5	Response time in Optical Packet and Burst Switching using TCP. . .	126

