Cheng-Shang Chang and Duan-Shin Lee

Principles, Architectures and Mathematical Theories of High Performance Packet Switches

Feb. 2006

Preface

Due to the recent advances in optical transmission technologies, the transmission speed of optical links is much faster than the switching speed of current electronic Internet routers. The challenge is then to find switch architectures that scale with the transmission speed of fiber optics. There are two approaches for this: (i) the electronic approach and (ii) the optical approach. The electronic approach is to use parallel electronic devices to acquire the needed speedup for fiber optics. On the other hand, the optical approach is to explore the possibility of building intelligent logic control directly with optical devices. It is the intent of this book to introduce recent advances in switch architectures from both the electronic approach and the optical approach.

In this book, we will first give a detailed review of existing switch architectures in Chapter 2, including both the output-buffered switches and input-buffered switches that are currently used in most Internet routers. Output-buffered switches, though ideal in the mathematical sense, suffer from the memory accessing speed problem. As such, its speed is in general limited by the memory accessing speed of the state-of-the-art memory technology. To acquire the needed speedup, parallel buffers are needed. Input-buffered switches are built for that purpose. However, as there are parallel input buffers, coordination of of parallel buffers results in a problem of finding "good" matchings. It seems that the matching problem was solved by two great mathematicians, G. Birkhoff [17] and J. von Neumann [163], long before it became a problem in input-buffered switches. The Birkhoff-von Neumann switch, an input-buffered switch that implements the algorithm developed by Birkhoff and von Neumann, achieves the ideal 100% throughput for all admissible traffic.

The load-balanced Birkhoff-von Neumann switches in Chapter 3 eliminate the need for finding "good" matchings in input-buffered switches. They are, therefore, much more scalable than input-buffered switches. The idea of the load-balanced Birkhoff-von Neumann switches is quite simple. In such a switch, parallel buffers are placed between two switch fabrics. The first switch fabric performs load balancing so that the traffic coming to the parallel buffers is uniform, and thus can be easily switched by the second switch fabric. As there are multiple routing paths through the load-balanced Birkhoff-von Neumann switches, packets in such switches may be delivered out of sequence. In Chapter 3, we shall introduce several variants in the literature that address the out-of-sequence problem in such switches.

Chapter 4 is a short chapter. There we introduce the concept of quasi-circuit switching. Traditionally, circuit switching is used for quality of service, while packet switching is used for bandwidth sharing. Quasi-circuit switching is a concept that falls in between packet switching and circuit switching, and thus can be viewed as a performance compromise between packet switching and circuit switching. The advantage of quasi-circuit switching is that quasi-circuit switches can be built with less complexity by using the load-balanced Birkhoff-von Neumann switches.

In Chapter 5, we introduce the optical approach. The key problem with optical packet switches is the lack of inexpensive memory. We start from an optical memory cell that is capable of storing one fixed-size packet. Then we use that as a basic building block to construct various optical queues, including time slot interchanges, First-In-First-Out (FIFO) multiplexers, FIFO queues, linear compressors, non-overtaking delay lines, and priority queues. The most interesting part of such a development is its connection to classical switching theory. For instance, linear compressors are connected to banyan networks, and non-overtaking delay lines and FIFO queues are connected to Benes networks.

This book is the result of courses developed in packet switch architectures at National Tsing Hua University. The material in this book can serve as a basis for a semester-long graduate level course that covers all the chapters in this book. Readers are recommended to take an undergraduate course in *computer networks* as a prerequisite. Also, for some of the mathematical theories in the book, it might be helpful to have some knowledge of *linear algebra (matrices)*, *differential equations*, *discrete math (graphs)* and *elementary probability.*

Several chapters of this book were rewritten from papers jointly coauthored with our colleagues and students in the last seven years.

For this, we gratefully thank Wen-Jyh Chen, Yi-Ting Chen, Jay Cheng, Ching-Te Chiu, Hsien-Chen Chiu, Chih-Chieh Chou, Hsiang-Yi Huang, Yi-Shean Jou, Issac Keslassy, Ching-Ming Lien, Nick McKeown, Ying-Ju Shih, Chih-Ying Tu, Chao-Kai Tu, Chao-Lin Yu, and Chi-Yao Yue. We give special thanks to Jay Cheng and Anne Bouillard for carefully reviewing an earlier draft and Yu-Hao Hsu for setting up the Web site

http://gibbs.ee.nthu.edu.tw/COM5353.htm

that contains slides and solutions for several sets of problems in the book. We are also grateful to the National Science Council and the Ministry of Education, Taiwan, R.O.C., for support of much of the work under the Program for Promoting Academic Excellence of Universities NSC 94-2752-E-007-002-PAE, and the National Innovative Communication Education Program.

Hsinchu, Feb. 2006 *Cheng-Shang Chang* and *Duan-Shin Lee*

Table of Contents