ABSTRACT OF THE DISSERTATION

Buffering and Flow Control in Communication Switches
for Scalable Multicomputers

by

Gregory Lee Frazier
Doctor of Philosophy in Computer Science
University of California, Los Angeles, 1995
Professor Yuval Tamir, Chair

Small $n \times n$ switches are key components of the communication networks of multicomputers. For fine-grained distributed tasks to efficiently execute on large multicomputers, high throughput, low latency communication must be supported. The architecture of these switches, particularly their internal buffers, is critical for achieving high performance communication with cost-effective implementations.

The focus of the work in this dissertation is the design of a new buffer, the *dynamically allocated, multi-queue* (DAMQ) buffer. This buffer maintains multiple queues of packets and allows the queue heads to be accessed in any order rather than restricting access to first in, first out (FIFO) order for the entire buffer. This reduces the effect of output port contention, giving a network of switches with DAMQ buffers (DAMQ switches) a significantly higher bandwidth and a lower average latency at given throughputs than a network of FIFO switches.

The DAMQ buffer is easily enhanced to support low latency delivery of high-priority packets even under heavy network load. This is important for distributed

real-time systems where a subset of the packet transmissions must happen within a bounded time.

The implementation complexity and performance characteristics of several non-FIFO buffer structures are evaluated. Extensive simulation and analysis are used in the evaluation of the DAMQ buffer and alternative buffers. A VLSI implementation of the DAMQ buffer as a component for communication switches is presented. Detailed circuit simulations are used to demonstrate the clear superiority of the DAMQ buffer over conventional FIFO buffers.

A methodology for evaluating flow control mechanisms is presented. A *congestion benchmark suite* is used to ‘‘stress’’ selected flow control mechanisms with a variety of congestion patterns. It is shown that hop-level flow control schemes in conjunction with multi-queue buffers can be used to optimize network performance under uniform and non-uniform traffic conditions.