Dynamically Reconfigurable Reduced Crossbar: 
A Novel Approach to Large Scale Switching

Tri Caohuu
Department of Electrical Engineering
San Jose State University
San Jose, California 95192-0084
ciaohuut@email.sjsu.edu

Thuy Trong Le
High Performance Computing Group
Fujitsu America, Inc.
San Jose, California 95134
thuy@fujitsu.com

Manfred Glesner, Jürgen Becker
Institute of Microelectronic System
Darmstadt University of Technology
64283 Darmstadt, Germany
glesner@mes.tu-darmstadt.de, becker@mes.tu-darmstadt.de

Abstract: The cross-bar is the fastest switching architecture for multiprocessor system, yet the most expensive in terms of hardware cost. The hardware complexity is of order $O(n^2 w)$ where $n$ is the number of processors and $w$ is width of the data path. In this paper we present a novel reconfigurable architecture where the hardware cost is reduced to $O(w(n/k)^2)$ while maintaining the same operating speed most of the time, $k$ is the reduction factor. The approach bases on a cache-like connectivity table. This table controls the reconfigurable data path of the cross bar. A hit in the connectivity table results in a direct connection of unit delay while a miss result in a miss penalty for updating the connectivity table. We assume that the local and spatial locality principles are applicable for this class of switching networks.

1. Introduction

The interconnection networks allow high-end multiprocessor system to fully exploit parallelism for higher performance. The ideal interconnection network connects instantaneously any output port of a processor to the input port of any other processor in the system. In practise, the physical location, the connection topology, and the contention of resources prevent this idea situation from happening. Typically, the interconnection can be a bus-based system, a multistage network, or a cross-bar and the design trade of is speed vs cost.

The bus-based network allows only one processor to communicate at anytime puts a significant constraint on the bus bandwidth particularly when a large number of processors is involved. This method is however the cheapest in term of hardware cost. The multistage network is a more popular choice for larger number of proces-

1 Currently a Visiting Professor at TU Darmstadt

© Springer-Verlag Berlin Heidelberg 1999
sors which allow several transactions to take place simultaneously. However the packets must travel through \( \log_p(n) \) where \( p \) is the switching factor and \( n \) is the number of processors. A major drawback of this scheme is that the latency increases as packets encounter delays due to contention for common paths. The multistage network therefore represents a middle ground between the performance speed and the hardware cost. The cross-bar switch connects each processor to any other processor in a system in one time unit delay and allows several transfers to occur simultaneously. Provided that no two processors trying to access the same destination at the same time, the cross-bar facilitates the connection of \( n! \) permutations without blocking. The cost of the hardware is however inhibiting for large size systems since it is of the order \( O(n^2 w) \) where \( n \) is the number of processor and \( w \) is the bus width [1,2,3].

In reality, the applications of “one-to-all broadcast”, “single-node accumulation”, “all-to-all broadcast”, etc... are very limited to the cases such as dense matrix-vector operations and database related operations. As a matter of fact, most of parallel scientific applications strongly have spatial and temporal communication localities [7]. Scientific computing problems generally are represented by systems of algebraic equations that are discretized in forms of sparse matrix and vector operations with some specific patterns. For example, the most popular sparse matrix patterns are the block-tridiagonal and the banded unstructure which are mostly originated by finite difference and finite element discretization classes of system of partial differential equations, respectively [8,9]. Parallel implementation of these methods will lead to the communications of only neighbouring processors. The number of neighbouring processors in fact depends on the decomposition of the physical domain and mostly is fixed for each parallel implementation.

In this paper we investigate the implementation of new type of reconfigurable cross-bar, coined reduced-crossbar, which offers the performance speed approximating that of a cross-bar at a fraction of the hardware cost. The novel approach is based on cache-based principle, which allows unit delay access for a hit and a longer penalty in case of a miss, will be presents in the next section. As in case of cache, if the hit rate is high, the reduced cross-bar is virtually a cross-bar at a fraction of the hardware cost.

### 2. Approach

In the past two decades, cache memory is a very important concept that had been successfully implemented to improve the processing speed of a computer system. Loosely speaking, if the data is the cache memory, the processor can access to it at a much faster speed that that of the main memory (hit). If the data is not in the cache (miss) then addition delay will be incurred to read in new line from memory. For a hit rate of 95% or higher, the processor is mostly operate from the cache with much faster access time.

Similar to the cache concept, the reduced-crossbar design is based on the assumption that the spatial and temporal locality properties also exist in large scale switching