Title	Hierarchical Interconnection Networks for Massively Parallel Computers
Author(s)	M., M., Hafizur Rahman
Citation	
Issue Date	2006-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/981
Rights	
Description	Supervisor:日比野　靖，情報科学研究科，博士

Hierarchical Interconnection Networks for Massively Parallel Computers

M. M. Hafizur Rahman

School of Information Science

Japan Advanced Institute of Science and Technology (JAIST)

January 10, 2006

Abstract

The most critical component in determining the ultimate performance potential of a multicomputer is its interconnection network. In this dissertation, we propose a new hierarchical interconnection network, called the Hierarchical Torus Network (HTN) for large scale 3D multicomputers. It consists of multiple basic modules (BMs) which are 3D-tori ($m \times m \times m$) and are hierarchically interconnected by 2D-tori ($n \times n$). Both the BMs and the interconnection at higher levels are toroidally connected, hence the name Hierarchical Torus Network (HTN). To reduce the vertical links between silicon planes, we consider higher-level networks as 2D-torus instead of 3D-torus network. We have explored various aspects such as network diameter, cost, average distance, bisection width, peak number of vertical links, and VLSI layout area of the HTN and compared them with those for other networks. It is shown that the HTN possesses several attractive features including constant node degree, small diameter, small average distance, high arc-connectivity, better bisection width, small number of wires, a particularly small number of vertical links, and an economic layout area.

We have used wormhole routing for switching because it has low buffering requirements, and more importantly, it makes latency independent of the message distance. Deadlock-free routing is the most critical issue in wormhole-routed networks and is achieved by using virtual channels (VCs). Since the hardware cost increases as the number of VCs increases, the unconstrained use of VCs is cost-prohibitive. In this dissertation, we present a deadlock-free routing algorithm for the HTN with a minimum number of VCs. By using the dimension-order routing and various traffic patterns, we have evaluated the dynamic communication performance of the HTN as well as other networks. HTN yields low latency and high throughput, which are indispensable for high performance massively parallel computers. It is also shown that the impact of non-uniform traffic patterns on the HTN is less than on the other networks. We have also described a suite of low-cost adaptive routers, LS, CS, and LS+CS with dimension order routing, analyzed their cost, and evaluated the dynamic communication performance for the HTN. The hardware cost for the LS, CS, and LS+CS algorithms is exactly equal to dimension order routing. The only overhead imposed is router delay for header selection. The dynamic communication performance using LS+CS algorithm is better than when the other algorithms are used. Therefore, an HTN with the LS+CS algorithm is a good choice for future massively parallel computers.

A fault tolerant network is very essential for the reliability of massively parallel computer systems. We have presented a hierarchical redundancy approach to reconfigure a faulty node by redundant node for the HTN. With a 25% redundancy, the system yield at the BM and second level are satisfactory. To show the suitability of the HTN, we have discussed mapping of some commonly used advanced applications. It is shown that the number of communication steps for applications mapping on the HTN is lower than for conventional and other hierarchical interconnection networks.

Pruning technique reduces the wiring complexity. We have explored the 3D-WSI implementation aspect of pruned-HTN. It is shown that the peak number of vertical links and layout area of pruned HTN in 3D-WSI is less than that of non-pruned HTN. To show the versatility of torus-torus combination for hierarchical networks, we have modified two other hierarchical networks (H3D-torus and TESH) using torus-torus networks.