A new Interconnection Network that achieves High Performance for Many-Core Processors

Faiz Al Faisal (1310059)

School of Information Science,
Japan Advanced Institute of Science and Technology

January 30, 2015

Keywords: Interconnection Network, 3D-TESH, Static Network Performance, Estimation of Power Consumption, Dynamic Communication Performance.

1 Introduction

High performance computer (HPC) is the increased demand for next generation of computers. Sequential computers fails to meet this demand and already reached to the saturation point due to the scaling difficulties of uniprocessor architectures. Hence the need for massively parallel computers (MPC) is increasing day by day. To facilitate the millions of node for MPC systems, interconnection networks are the key elements. Interconnection network acts as a path between one node to another. Considering millions of nodes, the large diameter of conventional interconnects is completely infeasible. On the other hand, to jump into the next level there is few key constraints that limit network performance, cost-performance ratio, power consumption, throughput and latency. One of the main constraints for MPC systems is the suitable interconnection network that could be scaled up to millions of nodes with small diameters. Next is the cost-performance issue, increased outgoing link from each node increases the total cost of the network. On the other hand, interconnection links consume the most of the power; shorter link consumes smaller power where

as longer links consumes more power. Even according to the advancement of current technology to build a one exa-flop system, it needs close to 1000MW of electrical power, which indeed a major constraint for next generation computers. And also it is always desirable that an interconnection network with low latency and high throughput for high network performance.

2   Introduction of 3D-TESH Network

Hierarchical interconnection networks (HINs) are a cost-effective way to interconnect a large number of nodes. In this research plan, we have introduce a multi-dimensional HIN network (3D-TESH). The 3D-TESH network consists of basic modules (BMs) that are hierarchically interconnected for higher level of networks.

Definition: A 3D-TESH(m, L, q) network, by definition is built using 2^m number of 2D-TESH $(2^m \times 2^m)$ basic modules, which has L levels of hierarchy, q is used for the inter-level connectivity and m is a positive integer.

In 3D-TESH, a BM is similar to 3D-mesh network consists of $(2^m \times 2^m \times 2^m)$ connected processing elements (PEs) having $(2^m \times 2^m)$ rows and $(2^m \times 2^m)$ columns. Higher level of 3D-TESH network is built by the recursive interconnection of the immediate lower level of sub-networks. A Level-2 3D-TESH network can be formed by $(2^{2 \times 2})$ 16BMs (16 level-1 3D-TESH network). Similarly, a Level-3 3D-TESH network can be formed by interconnecting 16 Level-2 sub-networks. The maximum level for 3D-TESH can be built by a $(2^m \times 2^m \times 2^m)$ BM is $L_{max} = 2^{m-q} + 1$. If inter-level connectivity, q = 0 & m = 2, L_{max}= 5; Level-5 is the highest possible level. The maximum number of nodes in each level of network can be defined, $N = (2^{2mL} \times 2^m)$. If m = 2 & L = 2, N = 1024. Similarly, a level-3 3D-TESH network will be consists of $N = 2^{2 \times 2 \times 3} \times 4 = 16384$ nodes, which is equal to 16 Level-2 3D-TESH networks.

3 Static Network Performance Evaluation

The topology of interconnection network affects the performance metrics. Performance metrics can be used to evaluate and compare different network topologies. It is being expected from the interconnection network with low cost, low degree, low congestion, high connectivity and high fault tolerance than the other networks.

3.1 Diameter Performance

Increase of path length increases the communication delay. Shortest path is desirable. The diameter of a network is the maximum inter-node distance i.e., the maximum number of links that must be traversed to send a message to any node along the shortest path. If the diameter is preferable it will take less time to route a packet. It has been observed from this research that the 3D-TESH requires less diameter than 2D-mesh, 2D-torus, 3D-mesh, 3D-torus, TESH networks for any hierarchical level of networks.

3.2 Average Distance

It is not always preferable to compare the network performance against only the diameter because a node has to communicate with others; hence on an average, shorter path than the lower diameter is being expected. The average distance is the mean distance between all distinct pairs of nodes in a network. Small average distance is preferable which allows small communication latency. It has been found from this research that the 3D-TESH requires less average distance than 2D-mesh, 2D-torus, 3D-mesh, 3D-torus, TESH networks for any hierarchical level of networks.

4 Dynamic Communication Performance Evaluation

Bad performance of the communicational network will severely limit the speed and efficiency of the entire MPC system. The dynamic communication performance (DCP) of an interconnection network is characterized by latency and throughput.

4.1 Estimation of Power Consumption

Power dissipation is a major concern for the next generation of supercomputers. The k-computer with Tofu interconnect requires about 9.89MW of electrical power to run the complete system. The power model for 3D-TESH can be defined by 2-ways; one is on-chip power model and other is the off-chip power model. The on-chip power model is based on the Orion energy model using 65nm fabrication process, considering the static and dynamic power dissipation within the routers and in the inter-router interconnects. We have used the garnet on-chip network model for using the Orion energy model, considering the clock frequency is 1GHz, supply voltage 1.0V, using 128bits message size and uniform traffic pattern with 2mm link length. And comparing the power analysis at the on-chip network, we can confirmed that 3D-TESH requires less power than the 3D-TORUS network and almost similar to 3D-mesh network.

Now, using the off-chip interconnect model we can also simulate the total required power for various levels of 3D-TESH network. To find the total required power at the level-2 3D-TESH network, we have assumed the off-chip wire length is 100mm, $V_{NF} = 8.8mV$, 1 virtual channels and similarly, to find the total required power at the level-3 3D-TESH network, we have assumed the off-chip wire length is 1m, $V_{NF} = 8.8mV$ & 1 virtual channels. Comparing the power consumptions for various networks using the off-chip interconnect model explains that 3D-torus network will require much more power than 3D-TESH network with the increase of the network size due to the increased number of outgoing links at various levels.

4.2 Simulation Environment for Dynamic Performance Evaluation

To evaluate dynamic communication performance, we use the TOPAZ interconnection network simulator. Here, we have showed the dynamic performance of 3D-TESH with regards to the 64 nodes and compare the result with the other networks having the same common parameters. In all the simulations, we use 6 virtual channels (VCs) for per physical link and 5 flits of packet size, having 16 bytes for each flit. We have used the variable

load and flits are transmitted at 20,000 simulation cycles using the Virtual Cut-Through (VCT) flow control and 2 cycle wire delay for level-1 links.

Uniform Traffic: We have observed the dynamic performance of 3D-TESH under uniform traffic pattern with the variable load. According to the analysis, the normalized supply throughput of the 3D-TESH network is higher than the Hierarchical Hypercube network(5-HHC) having 32 nodes, takes less buffer message latency and average transfer time than the 3D-mesh. But requires more buffer message latency and average transfer time than the 3D-torus network. Hence, 3D-TESH network achieves better dynamic communication performance than the Hierarchical Hypercube alike networks; and worse performance than that of 3D torus network.

Tornado: Here, 3D-TESH shows better performance than the 3D-mesh and 5-HHC network having 32 processing nodes but worse than the 3D-torus network in terms of both total message latency and buffer message latency. In tornado traffic pattern, 3D-TESH networks requires less zero load latency than the 3D-mesh network but slightly higher than 3D-torus network. As the comparing nodes for TTN(2,1,0) is only 16, TTN should be outperform other networks.

5   Conclusion

In this research plan, our main objective was to find a new interconnection network, which achieves high performance for many-core processors. And also we like to introduce a new interconnection network that could reduce the existing problems of interconnection networks such as- high power consumption, longer wire length, high cost-performance ratio, high throughput and high latency. Hence we have introduced a new multi-dimensional 3D-TESH HIN network. From our analysis we have found that the static network performance of 3D-TESH network is better than the conventional topologies of 2D-mesh, 2D-torus, 3D-mesh, 3D-torus and even better or equal than the 2D-TESH network in terms of diameter, average distance, node degree and bisection bandwidth. But the cost performance of 2D-TESH network shows better than the other networks due

to the limited outgoing link from the nodes. In summary, we have found that 3D-TESH network achieved about 52.08% better diameter performance and near about 45.71% better average distance performance than the 3D-torus network with 262,144 nodes. In case of dynamic network performance, we have able to find the buffer message latency and the total message latency along with the power consumption analysis of 3D-TESH network against the other networks. In case of power estimation, we could able to find the required power consumption for 3D-TESH network at various levels of networks, which shows that 3D-TESH network requires near about 14.81% less power than the 3D-torus networks with 16384 nodes. Now for the dynamic communication performance, 3D-TESH network is able to show better dynamic performance than 5-HHC (about 30% better performance for uniform traffic pattern when load = 0.2 flits/cycle/node) & 3D-mesh (about 2% better performance for uniform traffic pattern when load = 1.2 flits/cycle/node) in terms of uniform, tornado, perfect shuffle and bit-reversal traffic patterns even at the level-1 network with 64 nodes. As we have compared the 3D-TESH(2, 1, 0) topology at the level-1 network, it is obvious that we could get more better performance at higher levels of hierarchy for 3D-TESH network than the other networks due to hierarchical structure of 3D-TESH network.

Hence, in summary we can prefer that 3D-TESH would be a good choice for next generation supercomputers.