

Communication Locality Analysis of Triplet-Based Hierarchical Interconnection Network in Chip Multiprocessor

Shahnawaz Talpur^{1,2}, Feng Shi¹, and Yizhuo Wang¹

¹ Beijing Institute of Technology, Beijing P.R. China

² Department of Computer Systems Engineering,

Mehran University of Engineering & Technology Jamshoro, Sindh, Pakistan

talpur@bit.edu.cn, frankwyz@126.com

Abstract. Interconnection topology inside chip multiprocessor acts as fundamental role in communication locality. Considering compiler optimization data locality has been an inmost hypothesis in the high performance computing. Conversely, data locality sphere has several troubles when its degree of dimension is two or higher. In mesh network of two dimensions, each core is connected with its four neighbors. The data locality can potentially be exploited in two dimensions considering the specified processor's perspective. A Triplet-Based Hierarchical Interconnection Network (TBHIN) has straightforward topology and fractal attribute for chip multiprocessor. In this paper, a static (no contention) performance analysis of TBHIN and 2-D mesh is presented, based on the premise of locality in communication. The dynamic (contention) software simulation of TBHIN shows that the stronger the locality in communication, the lower the delay of the communication.

Keywords: Chip multiprocessor, communication locality, interconnection network, mesh.

1 Introduction

In computing, performance can be achieved by reducing data access latency, which is dependent on interconnect network on-chip. The issue of the highest performance on significant applications is the main effort of the microprocessor architecture, which also meets the limitation in the design time, power consumption and the area [1]. CMP is much more promising because it is partitioned into individual processing cores and has simple design scalability with great performance/power ratios and the potential to provide higher peak throughput [2].

In parallel with interconnection network, Cache organization has also the main impact on data latency which can affect the performance of multicore. Association of memory hierarchy, the operational work load and interconnection on-chip are the factors on which data access latency is built-up. Rapid path or routes are assembled for reducing latency in several Network-on-chip designs, utilizing communication

locality [3]. Latency on data stored in cache on-chip can be suffered by interconnection network, while the distance in blocks of data stored on cache can be affected by cache association. This affects the overall communication model against NoC.

Multicore has one decisive issue, the locality, which is alienated in temporal and spatial. When two or more cores access shared data in short time, it is known as temporal locality, whereas; in flat time, one core's data is accessed constantly by other core(s), known as spatial locality. Data latency will be even more critical issue as the number of cores increases in the system.

Packet interconnection network is an extensive approach in recent research, for controlling on chip communication. Wires deliver the suitable electrical constraint in structured network, which allows high recital circuits to raise bandwidth and dipping latency [4][5]. Researchers conclude that the power efficiency should be balanced with wire utilization through topologies.

The mature topologies, explored by parallel computing, such as complete connection, torus and cube are all too complex to implement on a single chip. Recently, 2-D mesh is much focused for implementation as interconnection network in CMP [6]. Although it is not expected that it will benefit at all.

The rest of the paper is prepared as follows: Section 2 presents TBHIN topology and node addressing. The static analysis and dynamic analysis are given in Section 3. We compare TBHIN with 2-D mesh on the condition of existence of communication locality. Section 4 consists of related work and Section 5 concludes this paper.

2 Triplet-Based Hierarchical Interconnection Network

Fig 1 illustrates the basic TBHIN topology and it can be seen that its structure is similar to Sierpinski Gasket, which is a fractal object. Iterartor Function System (IFS) produces Sierpinski Gasket according the theory of fractal geometry [7].

Single node in a scalable and hierarchical interconnection network described as a lowest level 0 TBHIN. On next step level 1 three cores shaping one triangle. The level 2 consists of three triangular shape communication links [8], as shown in Fig.1.

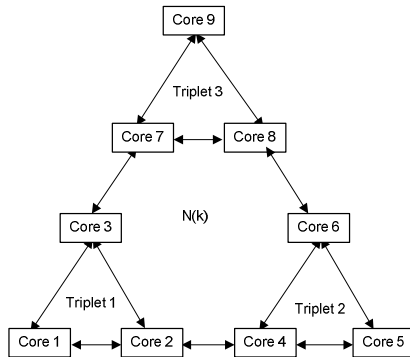


Fig. 1. Triplet-Based Hierarchical Interconnection Network Level 2(TBHIN)

Using Iterated Function System (IFS), higher levels of TBHIN can be generated from level 1. Assume that IFS for this design is IFS {F1, F2, F3}. Level 1 TBHIN $N(1)$ is the outcome after one iteration of IFS and $N(k)$ represents level k , after k th iteration that is the result of IFS. Then the procedure of TBHIN structure can be represented by

$$N_{k+1} = \bigcup_{i=1}^3 F_i(N_k) \quad (1)$$

Based on level 1 each core builds an upper level TBHIN. For the convenience three sub-networks of $N(k)$ are described here as Triplet 1, Triplet 2 and Triplet 3 [9]. However; in order to understand the organization and binary representation of basic addressing scheme of TBHIN, we show only the 2 levels of TBHIN in the following Table 1.

Table 1. TBHIN Addressing

| Triplet 1(01) | | | Triplet 2(10) | | | Triplet 3(11) | | |
|---------------|--------|--------|---------------|--------|--------|---------------|--------|--------|
| Core 1 | Core 2 | Core 3 | Core 4 | Core 5 | Core 6 | Core 7 | Core 8 | Core 9 |
| 0101 | 0110 | 0111 | 1001 | 1010 | 1011 | 1101 | 1110 | 1111 |

The ratio of the result of sum of minimal path lengths between two nodes to the total number of paths in the networks is known as mean minimal distance (MMD) of network [10]. The estimated MMD of TBHIN can be calculated as [11]:

$$P_{avg} = \frac{D_k}{3^k \times 3^k} = \frac{3^k \times \left(3 + \frac{16 \times (6^{k-1} - 1)}{5} - 3^{k-1} \right)}{3^k \times 3^k} \quad (2)$$

$$= \frac{1}{3^{k-1}} + \frac{16 \times (6^{k-1} - 1)}{5 \times 3^k} - \frac{1}{3}$$

Where D_k is sum of all shortest distance and, k is level of TBHIN.

3 Performance Analysis of Network

This section analyzes the performance of TBHIN. Through this analysis, it is showing the advantages of this interconnection network in comparison to 2-D Mesh, most commonly used network on chip. TBHIN uses less number of links than 2-D mesh, but it provides larger MMD. Section 4.1 presents a performance analysis assuming that there is no contention. Section 4.2 presents contention analysis and the simulation experiments that are on the queuing analysis.

3.1 Analysis with No Contention

Here, we consign all assumptions and performance comparison standards proposed in [12]. In this paper, we consider α as the probability at both the source and destination nodes of a message in the same sub-network, whereas; in [12] α is the probability at both the source and destination nodes of a message in the same cluster. Therefore $(1-\alpha)$ represents the probability of different sub-networks communication. Communication locality is stronger if the value of α is larger. Denoting q_m the probability of both source and destination nodes of a message being in N (m) and q'_m is the probability of a message being in N (M), but not in the same sub-network N (m-1). For N (k), we have:

$$\begin{aligned} q_1 &= \alpha = 1 - (1 - \alpha) \\ q_2 &= q_1 + q'_2 = 1 - (1 - \alpha)^2 \\ &\dots \\ q_{k-1} &= q_{k-2} + q'_{k-1} = 1 - (1 - \alpha)^{k-1} \\ q_k &= q_{k-1} + q'_k = 1 \end{aligned} \tag{3}$$

$$\begin{aligned} q'_2 &= \alpha (1 - q_1) = \alpha (1 - \alpha) \\ q'_3 &= \alpha (1 - q_2) = \alpha (1 - \alpha)^2 \\ &\dots \\ q'_{k-1} &= \alpha (1 - q_{k-2}) = \alpha (1 - \alpha)^{k-2} \\ q'_k &= 1 - q_{k-1} = (1 - \alpha)^{k-1} \end{aligned} \tag{4}$$

For larger value of α , the probability of inter-sub-networks communication will be smaller. As in [12], the cost and complexity of network N(k) is represented with L_k , and P_k , that is used to present the efficiency of the network N(k). Therefore, a useful measure is the product of L_k , and P_k ($L_k \times P_k$), which gives a cost-benefit part of static performance and should be minimized. For N(k), the total number of links can be presented as:

$$\begin{cases} L_1 = 3 \\ L_k = 3L_{k-1} + 3 \end{cases} \tag{5}$$

From Eq. (11) we can know that:

$$L_k = 3 \times (3^k - 1) / 2 \tag{6}$$

Denote P'_k as the source and destination nodes, which reside within the same network of N(k), but not within the same network of Triplet.

$$\begin{aligned}
P_k &= P_1 q_1 + P_2 q_2 + P_3 q_3 + \dots + P_k q_k \\
&= \alpha P_1 + \alpha(1-\alpha) P_2 + \alpha(1-\alpha)^2 P_3 + \dots + \alpha(1-\alpha)^{k-2} P_{k-1} + (1-\alpha)^{k-1} P_k \quad (7)
\end{aligned}$$

According to [11] suppose v_i and v_j are two nodes and they belongs to two different Triplets.

$$\begin{aligned}
P'_k &= \frac{\sum_{i=1}^{3^{k-1}} \sum_{j=1}^{3^{k-1}} d(v_i, v_j)}{3^{k-1} \times 3^{k-1}} = \frac{\sum_{i=1}^{3^{k-1}} \sum_{j=1}^{3^{k-1}} (d(v_i, v_{12})) + 3^{k-1} \times 3^{k-1} \times 1 + \sum_{i=1}^{3^{k-1}} \sum_{j=1}^{3^{k-1}} (d(v_{21}, v_j))}{3^{k-1} \times 3^{k-1}} \\
&= \frac{3^{k-1} \times \sum_{i=1}^{3^{k-1}} d(v_i, v_{12}) + 3^{k-1} \times 3^{k-1} \times 1 + 3^{k-1} \times \sum_{j=1}^{3^{k-1}} d(v_{21}, v_j)}{3^{k-1} \times 3^{k-1}} \quad (8) \\
&= \frac{2 \times 3^{k-1} \times Q_{k-1} + 3^{k-1} \times 3^{k-1} \times 1}{3^{k-1} \times 3^{k-1}}
\end{aligned}$$

From Eq. (8) and already calculated sum of shortest distance from any node with Triplet N(K)[11]

$$P'_k = \frac{2 \times Q_{k-1} \times 3^{k-1} + 3^{k-1} \times 3^{k-1} \times 1}{3^{k-1} \times 3^{k-1}} = \frac{2^{k+1} - 1}{3} \quad (9)$$

And because

$$P_1 = \frac{D_1}{3 \times 3} = \frac{6}{3 \times 3} = \frac{2}{3} \quad (10)$$

Thus from Eq. (7), Eq. (9) and Eq. (10) we can know:

$$P_k = \alpha - \frac{1}{3} + \frac{8 \times \alpha(1-\alpha) - 2^{k+1}(1-\alpha)^k}{3 \times (2\alpha - 1)} \quad (11)$$

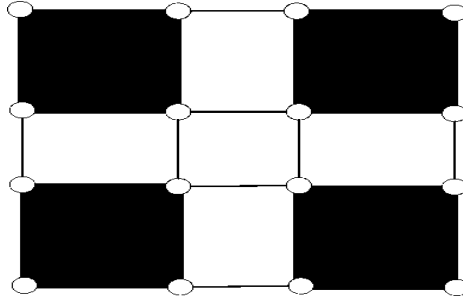


Fig. 2. The group scheme for 2-D mesh

To compare with TBHIN, every four nodes of 2-D mesh are grouped together as a sub-network, as shown in Fig. 2. After that four groups constitute a bigger group. Analogically, the 2-D mesh with $N = 4^k$ nodes can be considered as a k hierarchy network. The number of links and MMD of 2-D mesh are listed below:

$$L_k = 2 \times (N - \sqrt{N}) = 2 \times (4^k - 2^k) \tag{12}$$

$$P_{avg,k} = \frac{2(N-1)}{3\sqrt{N}} \approx \frac{2}{3} \sqrt{N} = \frac{2}{3} \times 2^k \tag{13}$$

Like TBHIN, suppose the probability of communication locality is α , yielding:

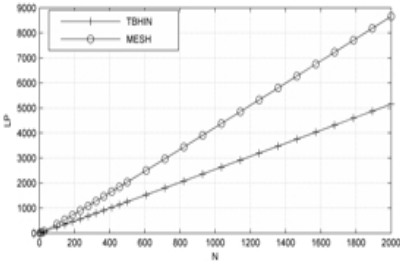
$$\begin{cases} P_1 = \frac{16}{4 \times 4} = 1 \\ P_k = \alpha P_1 + \alpha(1-\alpha) P_2' + \alpha(1-\alpha)^2 P_3' + \dots + \alpha(1-\alpha)^{k-2} P_{k-1}' + (1-\alpha)^{k-1} P_k' \end{cases} \tag{14}$$

P_k' is given by:

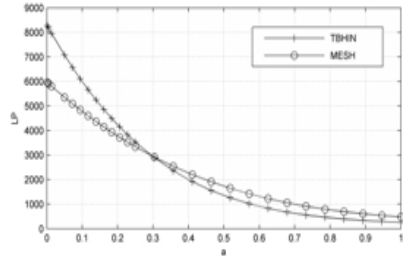
$$P_k' = \frac{P_{arg,k} \times 4^k \times 4^k - 4 \times P_{arg,k-1} \times 4^{k-1} \times 4^{k-1}}{4^k \times 4^k - 4 \times 4^{k-1} \times 4^{k-1}} = \frac{7 \times 2^k}{9} \tag{15}$$

Eq. (14) and Eq. (15) then give:

$$P_k = \alpha + \frac{28\alpha(1-\alpha) - 7 \times 2^k (1-\alpha)^k}{9(2\alpha-1)} \tag{16}$$



(A) $\alpha = 0.7$



(B) $N=256$

Fig. 3. The comparison of LP with $\alpha=0.7$ and $N=256$

Fig. 3(A) gives LP comparison of TBHIN and 2-D Mesh, when α is 0.7 as function of the all nodes in the network; it could be observed that TBHIN has lower LP than 2-D Mesh. In Fig. 3(B), N is fixed at 256 and the value of α is increased, to increase the probability of communication locality. The graph of Fig. 3(B) indicates that TBHIN has lower LP value when α is larger than 0.3, which proves better cost benefit ration than 2-D mesh on the condition of $\alpha > 0.3$. This condenses; TBHIN is applicable in the fields having higher communication locality.

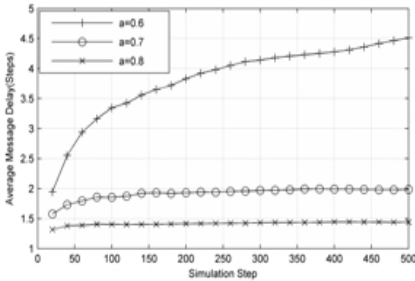
3.2 Analysis with Contention

The analysis with contention becomes more complex, since we have to consider the competition for communication resources among messages. In order to simulate the transition of messages over network, we assume that every link is full-duplex in the network. Each full-duplex link in our study is conceptually sub divided in two half-duplex links, each of which is modeled as a queuing center [12]. Between the source and destination nodes, all messages are routed over the shortest path. We also suppose that all the messages have the same size and the capacity of link buffers is unlimited.

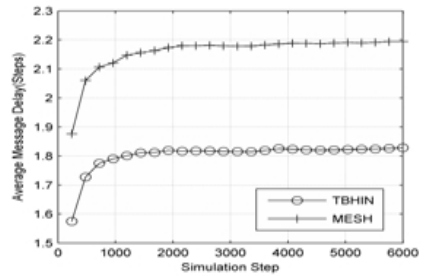
The simulation process is composed of consecutive steps; following tasks should be finished within each step:

1. Each node generates λ pieces of messages, whose source and destination nodes are randomly obtained according to the distribution of probability. The shortest path from source node to destination nodes is calculated using the algorithm proposed in section IV. Meanwhile, the message is sent to the waiting queue of the first link. Additionally, the shortest path is attached to the corresponding message.
2. Each link picks up μ pieces of messages from the waiting queue, and sends the message to the next link according to the attached routing information. When message reaches the destination node, calculate the total delay and delete the message.
3. Calculate the average delay by dividing the total number of messages into total delay of all messages.

Fig. 4(A) gives the average delay of messages along with the simulation steps. As we can see, the higher the communication locality, the shorter will be average delay. What's more, the network will be stable much more quickly. In reverse, the average delay becomes longer. And what is worse, the network may paralysis due to the heavy workload. In fig. 4(B), the average delay of TBHIN is around 25% lower than 2-D mesh when $\alpha = 0.6$, $N = 81$, $\lambda = 1$, $\mu = 2$. Therefore, Triplet-Based Hierarchical Interconnection Network dynamically performs much better than 2-D mesh.



(A) $N=81$, $\lambda=1$, $\mu=1$



(B) $\alpha=0.6$, $\lambda=1$, $\mu=2$ and $N=243$ for TBHIN, $N=256$ for Mesh

Fig. 4. The comparison of message delay

4 Related Work

Interconnection network on-chip emphasized more intensely as a prominent factor in Communication locality. In [3], interconnection for cache scheme and design circuit is classified periodically for the improvement in communication latency. Study essential diverse approach focuses on hindrance in locality and partitioning internal memory between cores in CMP shared memory system for reducing locality obstruction [13]. When data is shared among processors, allowing diverse application, optimal data locality scheme has been expressed with inimitable characteristic [14]. Parallel program and locality is achieved by enabling direct manipulation of tiles expressed in [15]. Optimal behavior of cache memory is stated and using sampling techniques elucidate the predictable scheme for CMEs solution in [16]. In literature survey of [17][18], latency improvement is explored in hop count, topology mapping, different mesh dimensions and interconnects graph etc. Data in groups is accessed collectively with reference to the association in memory as expressed in [19]. Despite the fact that, many of the researcher's proposed simple 2D packet-switched mesh interconnects for the cache. The novelty of our analysis demonstrates that Triplet based hierarchical internetwork can be replaced as improved one. Due to different attributes of each methodology for data accessing, cache/memory hierarchy can't fully benefit from 2D interconnections.

5 Conclusion

In this paper, we investigate communication locality between Triplet-Based Hierarchical Interconnection Network and 2-D mesh. It has been found that TBHIN is easily scalable and has better cost benefit ratio than 2-D mesh in case of higher communication locality. It is also observed that TBHIN dynamically performs much better than 2-D Mesh because of having lower average delay. Due to the facility of hardware implementation, it can be used as the interconnection network of CMP.

References

1. Codrescu, L., Wills, D.S.: Architecture of the Atlas chip-multiprocessor: dynamically parallelizing irregular applications. In: International Conference on Computer Design, ICCD 1999 (2002), 10.1109/ICCD.1999.808577
2. Kumar, R., Zyuban, V., Tullsen, D.M.: Interconnections in Multi-Core Architectures: Understanding Mechanisms, Overheads and Scaling. In: International Symposium on Computer Architecture, ISCA, pp. 408–419 (2005), doi:10.1109/ISCA.2005.34
3. Ahmed, A., Jones, A.K., Rami, M.: Codesign of NoC and Cache Organization for Reducing Access Latency in Chip Multiprocessors. IEEE Transactions on Parallel and Distributed Systems 23 (2012)
4. Dally, W.J., Towles, B.: Route Packets, Not Wires: On-Chip Interconnection Networks. In: Proceedings of the 38th Design Automation Conference, pp. 681–689 (2001)

5. Ching, D., Schaumont, P., Verbauwheide, I.: Integrated Modeling and Generation of a Configurable network-on-chip. In: Proceedings 18th International Parallel and Distributed Processing Symposium, p. 139 (2004)
6. Kumar, S., Jantsch, A., Soininen, J.-P., Forsell, M., Millberg, M., Oberg, J., Tiensyrja, K., Hemani, A.: A network on chip architecture and design methodology. In: Proc. Symposium on VLSI, pp. 117–124 (April 2002)
7. Feng, S., Weixing, J., Baojun, Q., Bin, L., Haroon, R.: A Triplet Based Computer Architecture Supporting Parallel Object Computing. In: Proceedings of the Eighteenth International Conference on ASAP, pp. 192–197 (July 2007)
8. Baojun, Q., Feng, S., Weixing, J.: A New Routing Algorithm in Triple-based Hierarchical Interconnection Network. In: The First International Conference on Innovative Computing, Information and Control, pp. 725–728. IEEE, Beijing (2006)
9. Bin, L., Zhi-Chen, T., Yu-Jin, G.: Triplet-based architecture and its process migration mechanism. In: 2009 International Conference on Machine Learning and Cybernetics, vol. 5 (2009), 10.1109/ICMLC.2009.5212644
10. Dong, Y., Wang, D., Zheng, W.: Exact Computation Of The Mean Minimal Path Length of N-MESH and N-TORUS. *Journal of Software* 20(4) (1997)
11. Weixing, J., Feng, S., Baojun, Q., Liu, B.: Study on an interconnection network for complex embedded systems. *Chinese High Technology Letters* 17(9), 886–890 (2007)
12. Dandamudi, S., Eager, D.: Hierarchical Interconnection Networks for Multicomputer Systems. *IEEE Transactions on Computers* 39(6), 786–797 (1990)
13. Min Kyu, J., Doe Hyun, Y., Dam, S., Mike, S., Ikhwan, L., Mattan, E.: Balancing DRAM Locality and Parallelism in Shared Memory CMP Systems. In: 2012 IEEE 18th International Symposium on High Performance Computer Architecture, HPCA (2012)
14. Kandemir, M.: Data locality enhancement for CMPs. In: Proc. ICCAD (2007)
15. Bikshandi, G., Jia, G., Daniel, H.: Programming for parallelism and locality with hierarchically tiled arrays. In: Proc. PPOPP (2006)
16. Xavier, V., Nerina, B., Josep, L., Antonio, G.: A fast and accurate framework to analyze and optimize cache memory behavior. In: TOPLAS 2004 (2004)
17. Abhinav, B., Laxmikant, V.K.: Benefits of Topology Aware Mapping For Mesh Interconnects. *Parallel Processing Letters* (April 16, 2010)
18. Abhinav, B., Eric, B., Laxmikant, V.K.: Optimizing communication for Charm++ applications by reducing network contention. *Concurrency and Computation: Practice and Experience* (2009); 00-17 Prepared using cpeauth.cls (Version: 2002/09/19 v2.02)
19. Zhang, C., Yutao, Z., Youfeng, W.: A hierarchical model of data locality. In: Proc. POPL (2006)