Wire Length of Midimew-Connected Mesh Network

Md Rabiul Awal[1,⋆], M. M. Hafizur Rahman[1], Rizal Mohd Nor[1], Tengku Mohd Bin Tengku Sembok[2], Yasuyuki Miura[3], and Yasushi Inoguchi[4]

[1] Department of Computer Science, KICT, IIUM, Jalan Gombak, Kuala Lumpur, 50728, Malaysia
[2] Cyber Security Center, National Defense University Malaysia, Kuala Lumpur, 57000,Malaysia
[3] Graduate School of Technology, Shonan Institute of Technology, 1-1-25, Tsujido Nishikaigan, Fujisawa, Kanagawa, Japan
[4] Research Center for Advanced Computing Infrastructure, JAIST, Nomi-Shi, Ishikawa 923-1292, Japan
{rabiulawal1,tmtsembok}@gmail.com, {hafizur,rizalmohdnor}@iium.edu.my, miu@info.shonan-it.ac.jp, inoguchi@jaist.ac.jp

Abstract. Midimew connected Mesh Network (MMN) is a Minimal DIstance MEsh with Wrap-around links (midimew) network. In this paper, we present the architecture of MMN and evaluate the total wire length of MMN, TESH, mesh, and torus networks. It is shown that the proposed MMN possesses simple structure and moderate wire length. The total wire length of MMN is slightly higher than that of mesh network and lower than that of 2-D torus network. Overall performance suggests that, MMN is an optimal network among these networks.

Keywords: Massively Parallel Computers, Interconnection Network, MMN, and Total Wire Length.

1 Introduction

Current trend [1] suggests that, the demand for computation power is increasing rapidly and found as constant over the last half century. Massively parallel computer (MPC) is introduced to meet this demand. Nevertheless, the scaling of MPC is increasing as well. In nearby future, MPC will contain 10 to 100 millions of nodes [2] in a single system with computing capability at the tens of petaflops or exaflops level. In MPC, interconnection network dominates the system performance [3, 4]. In relation, hierarchical interconnection network (HIN) is a plausible alternative way to interconnect the future generation MPC [5] systems. Nevertheless, the performance of proposed HIN does not yield any obvious choice of a network for MPC. Among a lot of HINs, several k-ary n-cube based HIN have been proposed [6–9] for better performance.

⋆ Corresponding author.

C.-H. Hsu et al. (Eds.): NPC 2014, LNCS 8707, pp. 132–143, 2014.

The scaling of the processor is an arising concern with the attributes of high performance. Application driven technology trends pressing the geometry of silicon fabrication technology. This advancements make the transistor very small and allow greater transistor densities. Eventually, MPC with more than million nodes is feasible with current and future technology. Hence, the functionality becomes more complex of a MPC system with the shrinking geometry. As a matter of fact, interconnection network becomes the steering point, in the context of power dissipation and cost. In an MPC system, more than 50% of total power dissipated by the interconnection network. Also, the cost of MPC is related to the communication links of the network. In other words, interconnection network is composed of nodes and wires. Hence, the network is wire limited [10] on a VLSI surface. Wire length determines the communication delay [11–13] of the network. It also indicate the network size on a VLSI surface. Total wire length of network indicates the average locality of links of the network. It also explores the ease of Network-on-Chip implementation of the network. Therefore wire length is an influential factor for the network [14].

Midimew-connected Mesh Network (MMN) [15] was proposed to improve performance of fixed degree network while keeping the diameter short which is still desirable [7, 16]. Basic module of MMN is 2-D mesh and higher level network are midimew [17] network. Hence, MMN offers simple and hierarchical structure and this translate to the ease of VLSI implementation. The focus point of this paper is to explore the feasibility of VLSI implementation of MMN in terms of wire length. We compare the total wire length with several fixed degree network. For fair comparison we consider degree 4 networks only.

The remainder of the paper is organized as follows. In Section 2, we present the basic architecture of the MMN. Wire length evaluation is discussed in Section 3. Finally, in Section 4, we conclude this paper.

2 Architecture of the MMN

Midimew connected Mesh Network (MMN) is a hierarchical interconnection network. Multiple basic modules (BM) are hierarchically interconnected to form a higher level network of MMN. Architecturally the MMN consists of two major parts, the basic module (BM) and higher level networks. The BMs act as the basic building blocks of MMN whereas higher level networks determines the construction of MMN from BMs.

Basic Module is the basic building blocks of MMN. BM of MMN is a 2D-mesh network of size $(2^m \times 2^m)$. BM consists of 2^{2m} processing elements (PE). PEs are arranged in 2^m rows and 2^m columns, where m is a positive integer. Considering $m = 2$, a BM of size (4×4) is portrayed in Figure 1. Each BM has $2^{(m+2)}$ free ports at the contours for higher level interconnection. These free ports are used as communication links for higher levels and denoted by q. All Intra-BM links are done by free ports of the interior nodes. All free ports of the exterior nodes, either one or two, are used for inter-BM links to form higher level networks. In this paper, BM refers to a Level-1 network.

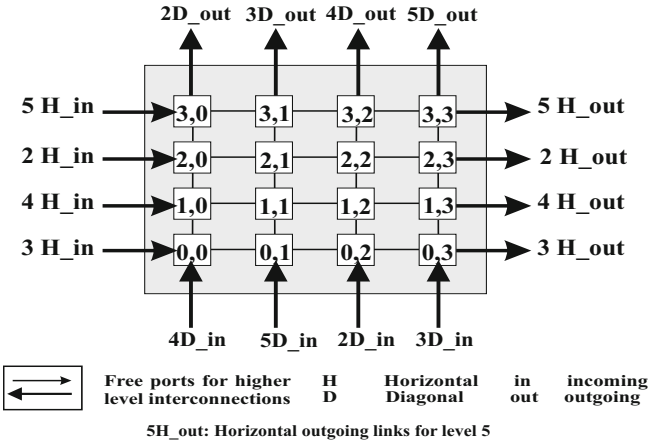

Fig. 1. Basic Module of MMN

Successive higher level networks are built by recursively interconnecting 2^{2m} immediate lower level subnetworks in a $(2^m \times 2^m)$ midimew network. In a midimew network, one direction (either horizontal or vertical) is symmetric tori connected and other direction is diagonally wrap-around connected. We have assigned the vertical free links of the BM for symmetric tori connection and horizontal free links are used for diagonal wrap-around links.

As portrayed in Figure 2, considering (m = 2) a Level-2 MMN can be formed by interconnecting $2^{(2\times 2)} = 16$ BMs. Similarly, a Level-3 network can be formed by interconnecting 16 Level-2 sub-networks, and so on. Each BM is connected to its logically adjacent BMs. It is useful to note that for each higher level interconnection, a BM uses $4 \times (2^q) = 2^{q+2}$ of its free links, $2^{(2q)}$ free links for diagonal interconnections and $2^{(2q)}$ free links for horizontal interconnections. Here, $q \in \{0, 1, ..., m\}$,, is the inter-level connectivity. $q = 0$ leads to minimal interlevel connectivity, while $q = m$ leads to maximum interlevel connectivity.

A MMN(m, L, q) is constructed using $(2^m \times 2^m)$ BMs, has L levels of hierarchy with inter-level connectivity q. In principle, m could be any positive integer value. However, if $m = 1$, then the network degenerates to a hypercube network and if $m \geq 3$, the granularity of the family of networks is coarse. If $m = 2$, then it is considered the most interesting case, because it has better granularity than the large BMs. In the rest of this paper we consider m = 2, therefore, we focus on a class of MMN(2,L,q) networks.

The highest level network which can be built from a $(2^m \times 2^m)$ BM is $L_{max} = 2^{m-q} + 1$ with $q = 0$ and $m = 2$, $L_{max} = 5$, Level-5 is the highest possible level. The total number of nodes in a network having $(2^m \times 2^m))$ BMs is $N = 2^{2mL}$. If the maximum hierarchy is applied then number of total nodes which could be connected by MMN(m,L,q) is $N = 2^{2m(2^{m-q}+1)}$. For the case of (4×4) BM with $q = 0$, a MMN network consists of over 1 million nodes.

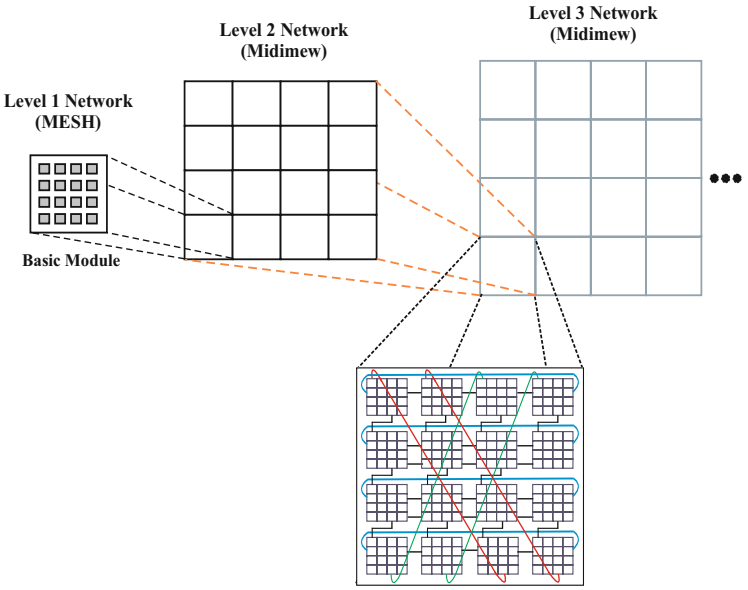

Fig. 2. Higher Level Networks of MMN

3 Wire Length Evaluation

Efficient use of wires is important to accomplish required performance from an interconnection network. Here we assume that all networks have a 2D-planar implementation and each node is implemented in one tile area. The width and height of a node depends upon its underlying CMOS technology. As the network consists of tiles (nodes), eventually the wire length depends on the size of tile. Let us consider, the tile height is x and tile width is y, hence tile area is xy. All the tiles are interconnected by wires to construct an interconnection network. For simplicity, we can consider total wires of the network are of two direction, Vertical and Horizontal direction. For Vertical and Horizontal direction, the wire length depends on the tile height and width respectively. Wire length between two particular nodes is the number of tiles needs to pass to interconnect the nodes. Consequently, total wire required to connect all the nodes of a network is the number of total tiles needs to be passed and can be expressed as,

$$Wire\ Length = Tile\ distance_X + Tile\ distance_Y$$

$$Tile\ distance = \#\ of\ tiles \times \#\ of\ groups$$

Here $\#\ of\ groups$ indicate the total number of same patterned communication links. For example, in a (4×4) 2D-mesh network, there are 4 columns and 4 rows with 16 nodes. So, in this network, $\#\ of\ groups$ for both vertical and horizontal directions is 4.

It is convenient to point out that, each node used in MMN have a router. These routers are used to interconnect all nodes, either directly or indirectly by the communication links. This interconnection network is implemented by direct wires. To evaluate the total wire length, we have calculated the length of the wires used to connect all the routers.

The communication links used to interconnect nodes are considered as bidirectional links. That is, each link is used for data in and data out by sharing the time. Therefore, each link contains just one wire to transfer data. For the intercommunication among nodes, unidirectional links can considered as well which contain multi wire for each link. Now, the point of using this type of links is, unidirectional links are faster than bidirectional links as they contain more channels for data passing. Thus, use of unidirectional links can improve the performance by saving time. Nevertheless, unidirectional links also require at least double wire than that of bidirectional links. As a result, unidirectional links increase the wire length, wiring complexity and cost of the network. On the other hand, use of bidirectional links save additional expenses and implement area of the network. Hence, for low cost, high performance network, bidirectional links are more appropriate.

We evaluated the total wire length of 2D-mesh, 2D-torus, TESH (2,2,0) and MMN (2,2,0). We have considered 45 nm technology to define the nodes size. According to 45nm technology [18], the tile height is 5.2 mm and tile width is 3.6 mm. Thus the tile size is 18.72 mm^2. So the node height, width and size are same as tiles. The wire length between two nodes suggests the number of tiles to be passed. So wire length between two nodes in horizontal direction is the product of number of tiles needs to be passed and tile width. Similarly for vertical direction, wire length is the product of number of tiles needs to be passed and tile height.

Hence, with the considered tile size, wire length between two neighbor nodes in horizontal direction is 3.6 mm and for vertical direction is 5.2 mm. For a 2D-mesh network, wire length depends on the grid size in vertical and horizontal direction. For example, in a $(N_x \times N_y)$ 2D-mesh network, wire length in one group in horizontal direction is $(N_x - 1) \times tilewidth$. Again for total N_x groups wire length is $N_x \times ((N_x - 1) \times tilewidth)$. Correspondingly, for total N_y groups $N_y \times ((N_y - 1) \times tileheight)$ is the wire length in vertical direction. Thus, for a (4×4) 2D-mesh network, wire length in vertical direction is 10.8 mm. With 4 groups and space for system interface, input/output (I/O), message class (MC), wire length in vertical direction is $(10.8 \times 4) + 4.9 = 48.1$ mm. Similarly for horizontal direction, wire length is 63 mm and total wire length is 111.1 mm. Figure 3(a) shows the wire length of a 4×4 mesh network. The wire length of a 16×16 (256 nodes) mesh is evaluated in same pattern and it is 211.75 cm.

In a 2D-torus network, with the mesh links, additionally there are wrap around torus links for both vertical and horizontal direction. Like mesh, 2D-torus is also dimension size dependent. For a $(N_X \times N_Y)$ 2D-torus network, two neighboring nodes in horizontal direction has the wire length of 3.6 mm, the width of a tile and in vertical direction wire length is 5.2 mm, the height of a tile. wire length of

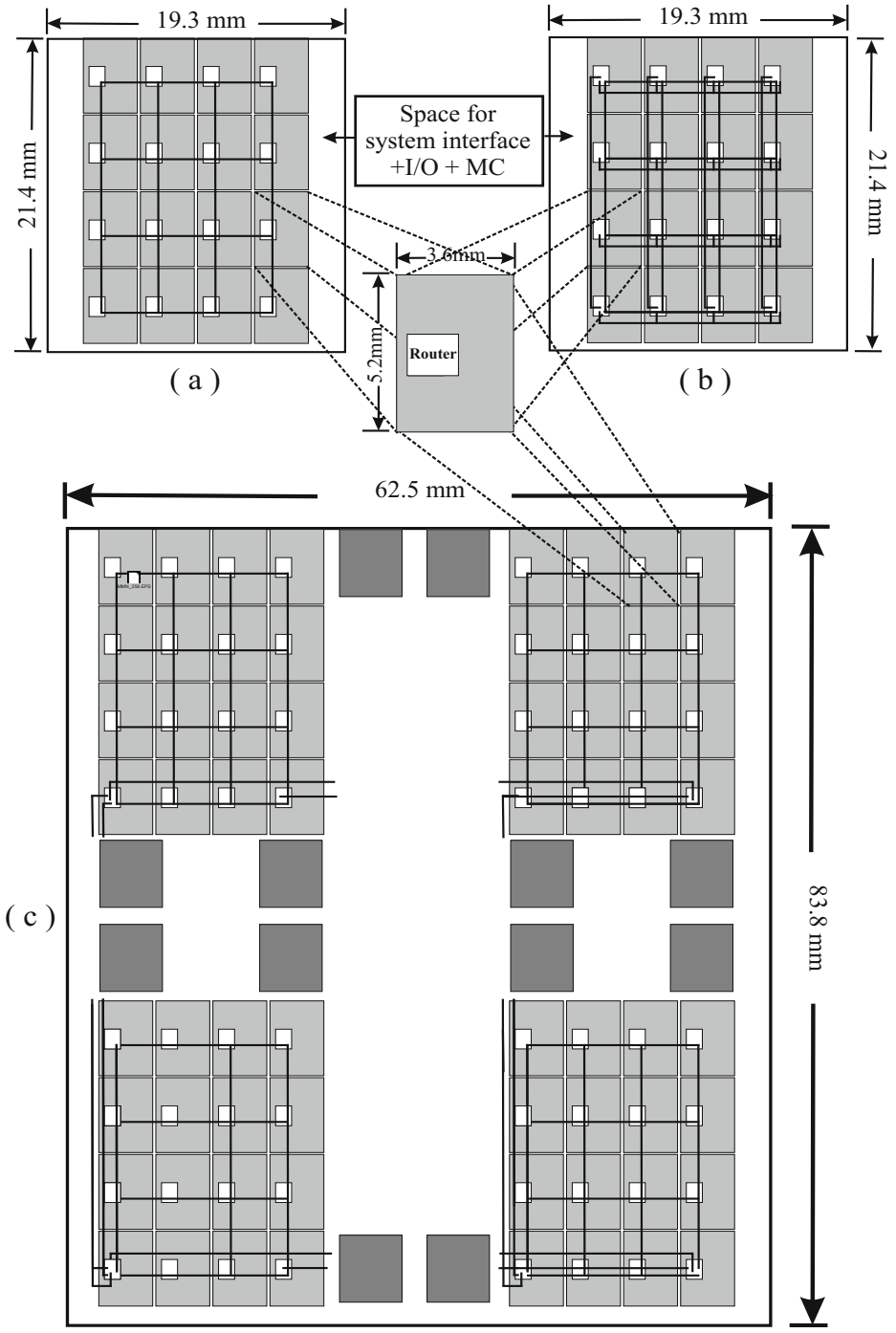

Fig. 3. Illustration of (a) 4×4 mesh, (b) 4×4 torus and (c) 256 nodes TESH Network

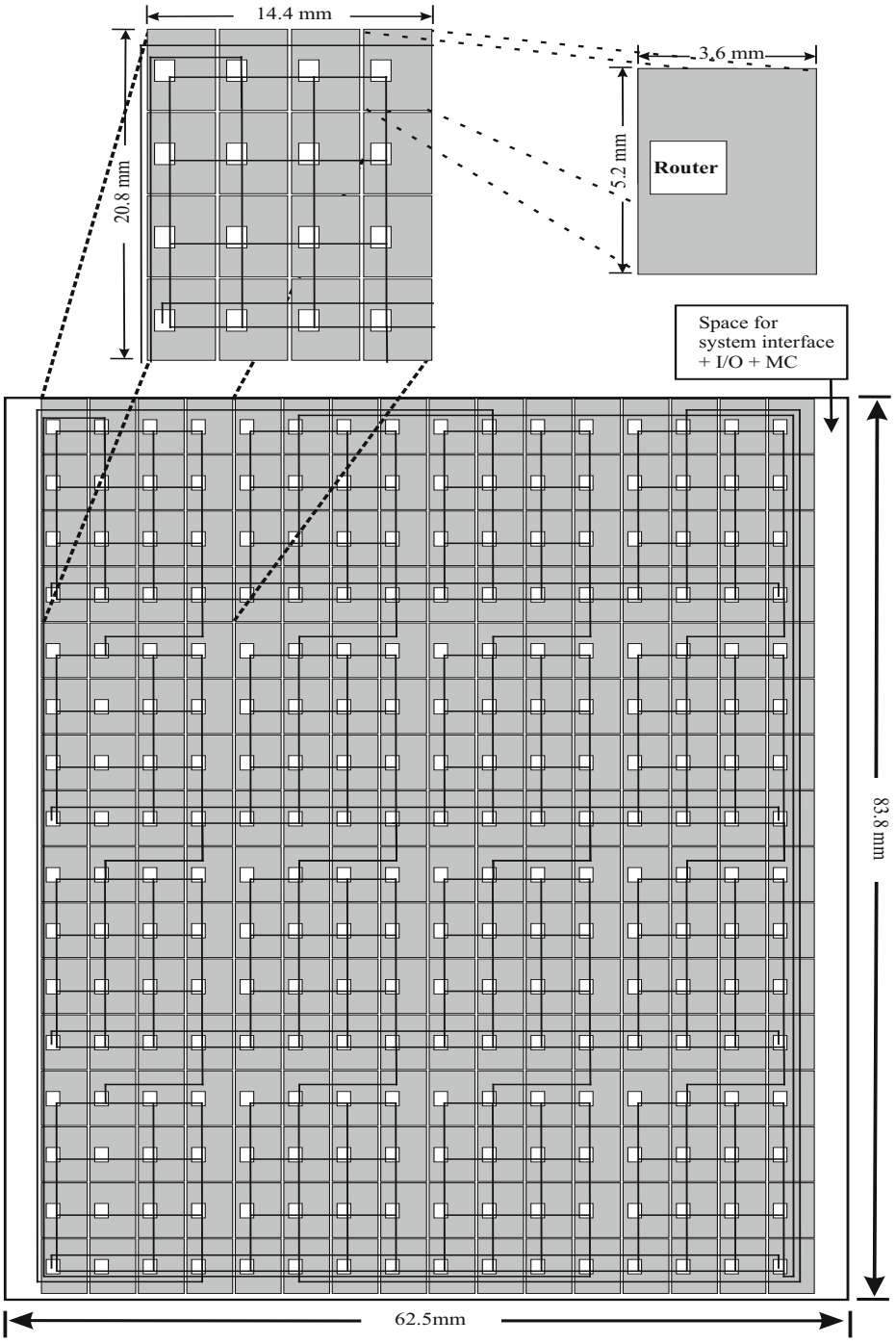

Fig. 4. 256 nodes MMN

the wrap around links in horizontal direction is N_{X-1}. Again in vertical direction, the wire length is N_{Y-1}. Number of groups are equal to N_X and N_Y. Figure 3(b) is showing a 4×4 2-D torus network. In horizontal direction wire length is 21.6 mm. # of groups is 4, therefore 86.4 mm is the wire length in horizontal direction. Similarly in vertical diction wire length of each group is 31.2 mm, for 4 groups, it is 124.8 mm. Finally with the reserved space for system interface, input/output (I/O), message class (MC) total wire length is 216.7 mm for a 4×4 2-D torus network. Wire length of a 16×16 (256 nodes) 2-D torus is calculated in a same manner and it is 422.95 cm in total. In Figure 3(b) a 4×4 2-D torus network is depicted.

TESH network is a hierarchical interconnection network with multiple basic modules. The basic module is a 2-D mesh network. Hence it is convenient to find out the wire length in basic modules, then calculate the connecting links for higher levels. The wire length of a 4×4 basic module is 105.6 mm which gives 1689.6 mm for 16 basic modules. In higher level links to horizontal direction wire length is 345.6 mm for 4 groups and 499.2 mm wire length for vertical direction. So in total the wire length of a 256 nodes TESH network is 253.99 cm. Figure 3(c) demonstrates the wire length of a 256 nodes TESH network.

Like TESH, Wire length of MMN evaluated by calculating the wires in BMs and in higher levels. Wire length in BMs of MMN is determined by the number of links in horizontal direction and in vertical direction and exactly equal to TESH, 1689.6 mm for 16 4×4 basic modules. Nevertheless, the links of higher levels is different from TESH. Higher level links of MMN are composed of wrap around and diagonal links. Each wraparound links has length of 54 mm which compute 216 mm wire length for 4 wrap around links. 4 diagonal links have length of 120 mm, 120 mm, 109.6 mm, and 109.6 mm. The rest inter BM links gives 480 mm. Hence 263.43 cm is the total wire length of a 256 nodes MMN. Figure 4 illustrates the wire length of a 256 nodes MMN.

The wire length dominates the initial system cost of networks. Networks with much wire eventually results a high installation cost and a large VLSI area which responsible for poor performance. In correlation, diameter indicates the worst case scenario of a network and has direct influence on the overall static network performance. Hence, the product of total wire length and diameter is a good criteria to get the static operating cost of the network. We can express the static operating cost as follows,

$$C_{static} = L \times D$$

Here, C_{static} represents the static operating cost, L for total wire length and D stands for diameter. It is already mentioned that, we have considered the links as bidirectional links. Hence, calculation of the links are the wire length. This calculation is valid for bidirectional links only.

Cost is one of the important parameter for evaluating an interconnection network. Though the actual cost of a system depends on the implemented hardware and the physical network in total but total wire length and diameter effect the performance metrics of the network including message traffic density, fault tolerance and average distance. Low diameter impose low cost, small space and better

performance, while high diameter requires high cost, large space, and poor performance. Therefore, the static operating cost is a good criterion to indicate the relationship between cost and performance of a network. Hence, it can give a pre-idea about the network before installation. The evaluation of Total Wire Length and static operating cost of various networks are tabulated in Table 1,

Table 1. Comparison of Total Wire Length of Various Networks

Network	Wiring Complexity	Total Wire Length (*cm*)	Static Operating Cost
2D-Mesh	480	211.75	6652.5
2D-Torus	512	422.95	6767.2
TESH(2,2,0)	416	253.99	5333.79
MMN(2,2,0)	416	263.43	4478.31

From Table 1, it is clear that 2D-mesh network can be constructed with minimum amount of wires among the networks, 211.75 *cm* in total. On the other hand, 2D-torus network contains maximum 422.95 *cm* of wires to interconnect all nodes of it. Total wire length for TESH (2,2,0) is 253.99 *cm* and in case of MMN (2,2,0), it is slightly higher than that of TESH network, 263.43 *cm*. The static network performance of different networks are shown in Table 2[15]. For fair comparison we consider degree 4 networks only.

2D-mesh is a very simple network. It is very easy to construct and does not contain any wrap around links. So 2D-mesh results least amount of wire length in total, but the performance of mesh keeps the network under the table. The diameter of mesh network is 30. Hence, this large diameter results the network ending with relatively high static operating cost 6652.5. It is higher than TESH, MMN and less than 2-D torus network.

2-D torus network presents better performance than that of others. 2D-torus network consists of 2D-mesh and warp around torus links. The wrap around links are equal to the sum of column and row numbers. It has $N_x + N_y$ wrap around links where the network size is $N_x \times N_y$. These long wrap around links results a significant amount of wires. Therefore, 2-D torus network possess maximum wire length among the networks. Though the diameter of 2-D torus is better than other networks which is 16, there is no wonder that it includes the maximum static operating cost among the networks which is 6767.2.

TESH (2,2,0) is a hierarchical network. It has optimized architecture combined of 2-D mesh and 2-D torus network. Hence, the wiring complexity and total wire length both are optimized for this network. As a result, TESH network has smaller diameter than mesh network and slightly higher than MMN and 2-D torus network. But, TESH require less wire than 2-D torus network. Eventually, TESH has less static operating cost than mesh and 2-D torus network and higher than MMN and it is 5333.79.

Like TESH (2,2,0), MMN (2,2,0) is also a hierarchical network. MMN has the combination of mesh and midimew networks for the architecture. Now, mesh is

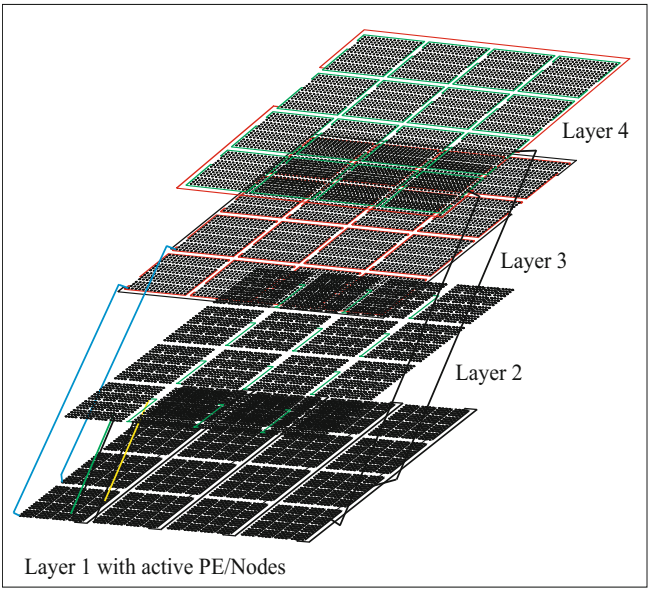

Fig. 5. Metal layers to implement MMN

the simplest network among all grid networks and midimew network has minimum diameter among all degree 4 networks. Hence, MMN is a perfect example of optimized network. So, naturally MMN has optimized diameter and total wire length. It possess less amount of communication links. Also the communication links increase with higher levels only not with the grid size. Therefore, MMN has better static operating cost than mesh, 2-D torus and TESH network and that is 4478.31.

Table 2. Comparison of Static Network Performance of Various Networks [15]

Network	Node Degree	Diameter	Average Distance	Ark Connectivity	Bisection Width
2D-Mesh	4	30	10.67	2	16
2D-Torus	4	16	8	4	32
TESH(2,2,0)	4	21	10.47	2	8
MMN(2,2,0)	4	17	9.07	2	8

The MMN can be implemented either on a 2-D or 3-D metal plane. For the case of 2-D plane, the network can be implemented on one metal layer only. In this case the implemented network will have a lot of jump crossing junction among the links. Jump crossing links cause serious affects. Therefore, multiple layers is a solution to avoid jump crossing of links. We have implemented

the MMN on 4 layer 2-D plane [19]. Figure 5 depicts the metal layers used to implement the MMN.

Despite the fact that MMN consumes more wires to be constructed than mesh and TESH network, MMN has smaller diameter than mesh and TESH [15]. As a result, MMN possesses better static operating cost. Hence, among these networks, MMN is the optimal network in the context of static network performance and required wire to implement the network.

4 Conclusion

The architecture and wire length of the MMN have been discussed in detail. In addition the wire length evaluation of 2-D mesh, 2-D torus and TESH are also explored and compared with MMN as well. It is shown that the MMN possess a simple architecture, composed of 2-D mesh and midimew network. From the wire length evaluation, it is clear that, the MMN presents moderate wire length in total with fixed degree nodes. The total wire length of MMN is slightly higher than that of 2-D mesh and TESH network. However total wire length of MMN is far lower in comparison with 2-D torus. This paper focused on the architectural structure and wire length evaluation. Issues for future work include wire length evaluation of MMN in a 3-D environment.

Acknowledgments. This work is partly supported by FRGS13-065-0306, Ministry of Education, Government of Malaysia. The authors would like to thank the anonymous reviewers for their constructive comments and suggestions on the paper which have helped to improve the quality of the paper.

References

1. Koomey, J.G., Berard, S., Sanchez, M., Wong, H.: Assessing trends in the electrical efficiency of computation over time. In: IEEE Annals of the History of Computing (2009)
2. Beckman, P.: Looking toward exascale computing. In: 9th International Conference on Parallel and Distributed Computing, Applications and Technologies, p. 3 (2008)
3. Yang, Y., Funahashi, A., Jouraku, A., Nishi, H., Amano, H., Sueyoshi, T.: Recursive diagonal torus: an interconnection network for massively parallel computers. IEEE Transactions on Parallel and Distributed Systems 12, 701–715 (2001)
4. Rahman, M.M., Hafizur., J.X., Masud, M.A., Horiguchi, S.: Network performance of pruned hierarchical torus network. In: 6th IFIP International Conference on Network and Parallel Computing, pp. 9–15 (2009)
5. Abd-El-Barr, M., Al-Somani, T.F.: Topological properties of hierarchical interconnection networks: a review and comparison. J. Elec. and Comp. Engineering 1 (2011)
6. Lai, P.L., Hsu, H.C., Tsai, C.H., Stewart, I.A.: A class of hierarchical graphs as topologies for interconnection networks. J. Theoretical Computer Science 411, 2912–2924 (2010)

7. Liu, Y., Li, C., Han, J.: RTTM: a new hierarchical interconnection network for massively parallel computing. In: Zhang, W., Chen, Z., Douglas, C.C., Tong, W. (eds.) HPCA 2009. LNCS, vol. 5938, pp. 264–271. Springer, Heidelberg (2010)
8. Rahman, M.M.H., Horiguchi, S.: HTN: a new hierarchical interconnection network for massively parallel computers. IEICE Transactions on Information and Systems 86(9), 1479–1486 (2003)
9. Jain, V.K., Ghirmai, T., Horiguchi, S.: TESH: A new hierarchical interconnection network for massively parallel computing. IEICE Transactions on Information and Systems 80, 837–846 (1997)
10. Dally, W.J.: Performance Analysis of k-ary n-cube Interconnection Networks. IEEE Trans. on Computers 39(6), 775–785 (1990)
11. Chi-Hsiang, Y., Parhami, B., Emmanouel, A., Varvarigos, E.A., Hua Lee, H.: VLSI layout and packaging of butterfly networks. In: Proceedings of the Twelfth Annual ACM Symposium on Parallel Algorithms and Architectures, pp. 196–205 (2000)
12. Dally, W.J., Towles, B.: Route packets, not wires: On-chip interconnection networks. In: Proceedings of Design Automation Conference, pp. 684–689 (2001)
13. Parhami, B.: Introduction to parallel processing: algorithms and architectures, vol. 1. Springer (1999)
14. Parhami, B., Kwai, D.M.: Challenges in Interconnection Network Design In the Era of Multiprocessor and Massively Parallel Microchips. In: Proc. Int'l Conf. Communications in Computing, pp. 241–246 (2000)
15. Awal, M.R., Rahman, M.H., Akhand, M.A.H.: A New Hierarchical Interconnection Network for Future Generation Parallel Computer. In: Proceedings of 16th International Conference on Computers and Information Technology, pp. 314–319 (2013)
16. Camarero, C., Martinez, C., Beivide, R.: L-networks: A topological model for regular two-dimensional interconnection networks. IEEE Transactions on Computers 62, 1362–1375 (2012)
17. Puente, V., Izu, C., Gregorio, J.A., Beivide, R., Prellezo, J., Vallejo, F.: Improving parallel system performance by changing the arrangement of the network links. In: Proceedings of the 14th International Conference on Supercomputing, pp. 44–53 (2000)
18. Howard, J., Dighe, S., Vangal, S.R., Ruhl, G., Borkar, N., Jain, S., Erraguntla, V., Konow, M., Riepen, M., Gries, M., Droege, G., Larsen, T.L., Steibl, S., Borkar, S., De, V.K., Wijngaart, R.V.D.: A 48-core IA-32 processor in 45 nm CMOS using on-die message-passing and DVFS for performance and power scaling. IEEE Journal of Solid-State Circuits 46(1), 173–183 (2011)
19. Awal, M.R., Rahman, M.H.: Network-on-Chip Implementation of Midimew-Connected Mesh Network. In: Proceedings of 14th International Conference on Parallel and Distributed Computing, Applications and Technology, pp. 265–271 (2013)