Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 179 (2021) 590–597

Procedia Computer Science

www.elsevier.com/locate/procedia

5th International Conference on Computer Science and Computational Intelligence 2020

Midimew Connected Torus Network for Next Generation Massively Parallel Computer System

Maryam Manaa Al-Shammari[a], Asrar Haque[a], M.M. Hafizur Rahman[b]*

[a] *Dept. of Computer Science, CCSIT, King Faisal University, Al Hassa 31982, Kingdom of Saudi Arabia*
[b] *Dept. of Computer Networks & Communications, CCSIT, King Faisal University, Al Hassa 31982, Kingdom of Saudi Arabia*

Abstract

Important scientific and engineering applications need high performance computing. This can be provided by massively parallel computer (MPC) systems. A sensitive step of maintaining such systems is the interconnection network used to interconnect the computing nodes. The topology used effects the network costs and performance significantly. Hierarchal interconnection networks (HIN) were introduced having several attractive features including low latency, low cost, and high fault tolerance. This paper proposes a new HIN called Midimew connect Torus Network (MTN) that provides constant node degree, high arc connectivity, high fault tolerance and a reasonable bisection width. Static network performance evaluation for the proposed MTN has been conducted and compared with other networks. The comparison included conventional topologies such as 2D Mesh and 2D Torus and hierarchal ones which are TESH and TTN. The comparison of result shows that with the cost of extra communication links MTN attains higher fault tolerance than that of TESH and mesh networks, and equal to that of TTN, and less than that of torus network. Also, the hierarchical networks such as MTN, TTN, and TESH yield moderate bisection width; and the bisection width of MTN is lower than that of mesh and torus networks. This paper shows the superiority of the proposed MTN by comparative study of static network performance.

Peer-review under responsibility of the scientific committee of the 5th International Conference on Computer Science and Computational Intelligence 2020

Keywords: Messively Parallel Computers; Hierachical Interconnection Network; Static Network Performance;

* Corresponding author.
E-mail address: mhrahman@kfu.edu.sa

1877-0509
Peer-review under responsibility of the scientific committee of the 5th International Conference on Computer Science and Computational Intelligence 2020
10.1016/j.procs.2021.01.044

1. Introduction

Here MPCs play an important role in our daily lives through solving computationally intensive problems. They are used to solve problems or meet specific goals [1] with exaflops level of computational speed [2]. They help in solving the grand challenge problems such as brain initiative, reducing solar energy costs, dealing with asteroid threats and affordability of electric vehicles [3]. In addition, MPCs are important in the development and testing of powerful and sophisticated security equipment including nuclear systems. In fact, almost all important areas make use of such systems [4] including education and scientific research, energy coordination, weather forecasting, neural networks and medicine developments [5]. Thus, MPCs has drawn interest significant number of researchers proposing new topologies to build networks for these systems. The necessity of proposing such networks was due to the inability to use conventional networks because their large diameter makes it infeasible to implement networks consisting millions of nodes (next generation of MPC) [6]. For example, in a 2D mesh even with only 256 nodes could be interconnected ending up having diameter as high as 30 hop-distance. Thus, latency will increase significantly making message traveling inefficient with respect to time and resource consumption [7].

MPC system also plays an important role in improving health care and developing new medicines. To maintain the national health system database consisting of medical images and to process this massive amount of database for critical decision making regarding national health system, MPC system is indispensable. To reduce community transmission and forecast nature of spreading of a disease like the COVID-19 pandemic, we need to identify and track affected people and closely monitor their surrounding community [8]. To do this vital task in a timely manner, enormous computing power of MPC system is being utilized [9]. To discover drugs for any new diseases in a shorter period of time, simulation considering many factors helps scientists from medicinal chemistry, biochemistry, and virology to follow the complex process of target identification, compound screening, and lead identification for potential drugs. This simulation followed by preclinical test gives better results in clinical trials [10]. To develop a safe and effective vaccine against COVID-19 or any other virus, the genome sequence of that virus needs to be analyzed. It is noteworthy that the analysis of COVID-19 virus is quite challenging because of its genetic mutation in new environment and geographical location. Indeed, MPC system can play a vital role in the genome sequence analysis of COVID-19 [11].

HINs allow connecting millions of nodes with a reasonable diameter resulting in cost-effective implementation [12]. Since topology used to build MPC plays a sensitive role that impacts the overall performance of the system, a new HIN for next generation of MPC [7] is proposed in this paper. It is important to note that even many MPCs still use direct networks for their attractive features including large bandwidth and low dimeter. However, this is only possible for small sized systems [13]. Consequently, next generation where up to million nodes are needed, HINs could provide more attractive features where efficient and fast message delivery could be attained [14].

The proposed topology, MTN, is built from interconnecting of 2D torus as a basic module and midimew to form higher levels. The rationale to have 2D torus network as the basic module is that in comparison to other similar conventional networks it can provide higher fault tolerance and lower latency. On the other hand, a midimew is an optimal topology with respect to diameter and average distance for a network of degree four. The combination of 2D torus and midimew provides lower network latency and bandwidth leading to lower contention. MTN also addresses other factors such as bisection width, arc connectivity and wire complexity. A static evaluation has been carried out to illustrates the features of MTN comparing results with the other HIN.

Many HINs topologies were introduced in the literatures such as Tori connected mESH (TESH) [15,16] and Tori connected Torus Network (TTN) [12]. These networks provide low diameter and constant node degree with variations of some network parameters. In analogy, MTN provides constant node degree that remains same with the increase of the network size making it easily scalable. However, MTN provides more attractive features including low cost, reasonable bisection width, variations of high connectivity and a suitable fault tolerance ability, as described in section 3. In addition, MTN is expected to reduce the message traveling time with respect to diameter.

The rest of the paper is organized as following: Section 2 includes a brief description of MTN architecture. Section 3 provides a discussion of a comparative static network evaluation of MTN. In addition, some generalizations with respect to network architecture were introduced in Section 4. Finally, concluding results are summarized in Section 5.

2. Network Architecture of the MTN

Midimew connected Torus Network (MTN) is a HIN that consists of 2D Torus as a basic module and interconnected by Midimew topology to form the higher levels of the network. The basic module (BM) is a 2D torus that represents level-1 which is the lowest level. The next level, level-2, is formed by the hierarchal interconnection of $2^m \times 2^m$ of level-1 where *m* is a positive integer. The direction for the connections between the BMs inside each level are bidirectional due to small distances and absence of intermediate buffers. However, the connections forming the higher level are unidirectional as intermediate buffers are used because of long wires. MTN architecture can be defined as the following:

Definition 1: A $2^m \times 2^m$ BM is a 2D torus network consisting of 2^{2m} processing elements formed by 2^m rows and 2^m columns, where m is a positive integer.

Definition 2: An MTN (*m, L, q*) is built by $2^m \times 2^m$ BM having L level of hierarchy and q for inter-level connectivity. The connections to form the higher levels makes each BM use 4×2^q of its free ports where $2(2^q)$ for horizontal and $2(2^q)$ for diagonal connections where $q \in \{0, 1, \ldots, m\}$, $q = 0$ provides the minimal level of inter-connectivity and maximum one provided by $q = m$. For better building of network m is considered in this paper to be 2 as it is shown to provide better graininess.

Definition 3: The highest level that can be formed from MTN *(m, L, q)* is $L_{max} = 2^{m-q} + 1$ level.

Definition 4: The wire complexity, W_L, in MTN is defined as the total number of wires used to form the overall network of level L. Let W_{BM} be number of links in a BM. Then W_L can be found using equation 1.

$$W_L = W_{BM} \times 2^{m \times 2(L-1)} + \sum_{x=2}^{L} 2(2^q) \times 2^{m \times 2(L-1)} \quad (1)$$

To support better understanding of the above definitions, an MTN (2,3,0) is represented in Fig. 1. In addition, Table 1 is used to illustrate virous levels of MTN to provide clear understanding of higher levels connections.

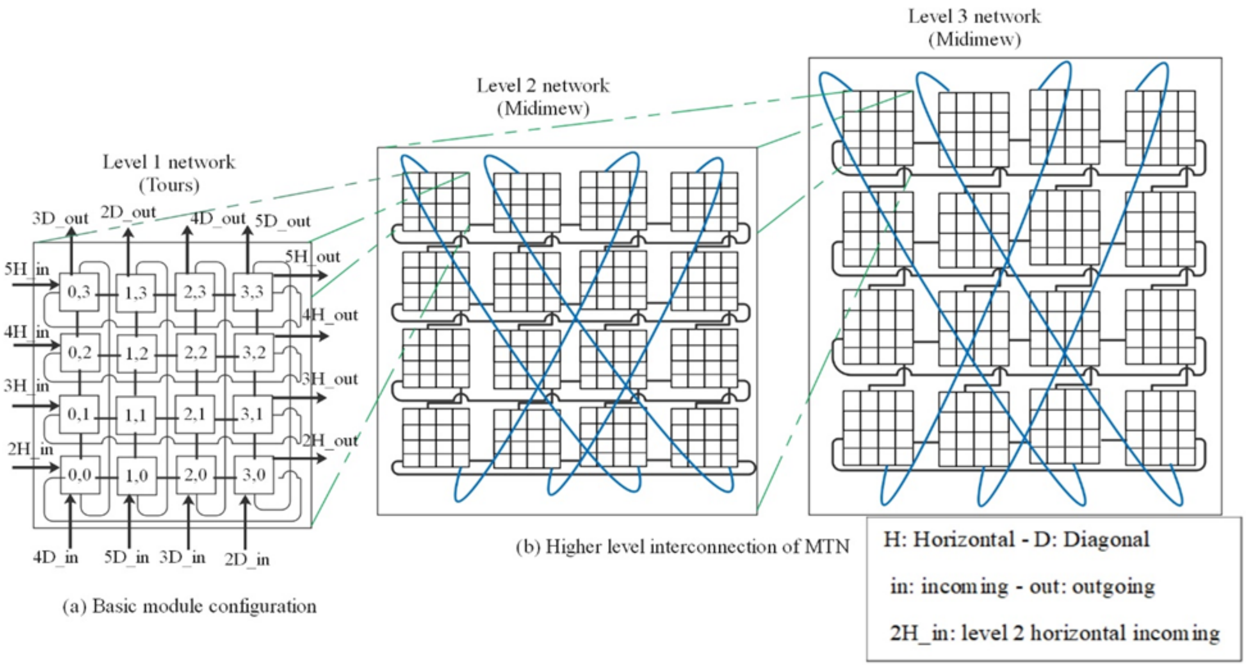

Fig. 1. Hierarchy of MTN (2,3,0)

Table 1. Various levels of MTN.

	Consists of	Total nodes	Free ports used	Wire complexity
Level-2	16 of 4 x 4 torus	256	4	544
Level-3	16 of 4 x 4 level 2	4,096	8	8,736
Level-4	16 of 4 x 4 level 3	65,536	12	139,808
Level-5	16 of 4 x 4 level 4	1,048,576	16	2,236,960

3. Comparative Static Network Performance Evaluation

The topology of MTN described in the previous section can provide performance parameters in the light of graph theory. These parameters show some attractive features of the network such as low costs, high connectivity and scalability. However, real implementation will be impacted by physical issues and technological properties. The following section describes four static network parameters that effects the overall performance of a system [14]. MTN was evaluated with respect to other networks: Mesh, Torus, TESH and TTN. The evaluation was done according to the number of total nodes that could be provided using different levels and inter-connectivity links for HINs as shown in Table 2. However, for Mesh and Torus corresponding 2D architectures with respect to node number are used.

Table 2. Static network performance parameter of several networks.

	Node Degree	Bisection Width	Arc Connectivity	Wire Complexity	Total # Nodes
16 × 16 Mesh	4	16	2	480	256
16 × 16 Torus	4	32	4	512	
TESH (2, 2,0)	4	8	2	416	
TESH (2,2,1)	4	16	2	448	
TTN (2,2,0)	6	8	4	544	
TTN (2,2,1)	6	16	4	576	
MTN (2,2,0)	6	8	4	544	
MTN (2,2,1)	6	16	4	576	
64 × 64 Mesh	4	64	2	8,064	4,096
64 × 64 Torus	4	128	4	8,192	
TESH (2,3,0)	4	8	2	6,688	
TESH (2,3,1)	4	16	2	7,232	
TTN (2,3,0)	4	8	4	8,736	
TTN (2,3,1)	6	16	4	9280	
MTN (2,3,0)	6	8	4	8,736	
MTN (2,3,1)	6	16	4	9,280	

3.1. Node Degree

Node degree indicates the highest number of connections between a node and its neighboring nodes [17]. It is directly related to the network cost as higher node degree requires expensive routers. In addition, constant node degree is important to allow the network scaling without changing the network interface configuration [18]. MTN and the other networks shown at Table 2 have constant node degree with respect to different network sizes. However, even though MTN has 6 as a node degree it is important to note that this could be compromised with lower diameter and average distance [19]. Also, this highest node degree is considered since only some of the external nodes have 6 connections. For example, in MTN (2,2,0) only 20 nodes out of 256 nodes are connected to 6 neighboring nodes. Thus, MTN provides constant node degree that supports scaling capabilities with indications of lower diameter as Midimew interconnection to form higher levels.

3.2. Bisection Width

When disconnecting the network into two halves of equal sizes, bisection width can be defined as the lowest amount of connections that must be removed to allow this division. Its importance becomes evident when applying divide and conquer technique where the problem input is divided into halves and then further more. The results are then merged together until reaching the original two halves of the network. If the network bisection width is low, this indicates minimal bandwidth between the network halves resulting in slower merging process. However, even though large bisection width means higher bandwidth, it comes with higher costs due to complexity and extra wires [20]. These costs become serious issues for VLSI where the design principle makes use of divide and conquer to allow fast and large amount of functionalities to take place. Bisection bandwidth could give a good indication of the true bandwidth of the entire system [21]. MTN provides a reasonable bisection width with respect to total number of nodes [21]. For example, MTN allows implementing more than one million nodes with only 16 connections between the two equal halves. In comparison to other networks, bisection width of MTN, TESH and TTN are similar that are lower than mesh and torus.

3.3. Arc Connectivity

Connectivity of a network refers to how much multiple paths are available between neighboring nodes. Higher connectivity is desirable to have lower contention. One of the important connectivity measures is arc connectivity that indicates least number of connections that when removed will divide the network into two disconnected parts. Even though routing properties depend on real implementation of networks and technology used, arc connectivity can be used as a reliable measure of static network robustness [12]. It represents multiple connections which the packets can use, thus, it is useful in avoiding network congestion [22]. Also, coordination of these multiple links will take place in case of faulty operations [23]. The ratio between arc connectivity and node degree indicates the static fault tolerance performance of an interconnection network. Arc connectivity cannot be higher than node degree. Thus, it is useful to note that equal arc connectivity to node degree provides maximum fault tolerance as found in torus networks [24]. The proposed MTN results higher fault tolerance than that of mesh and TESH networks and equal fault tolerance to that of TTN as represented in Table 2.

3.4. Wire Complexity

Wire complexity refers to all the connections used to layout the overall network. It is directly proportional to the node degree of the network since the wires are connecting the node according to the links going out from it to its neighbours. Higher wire complexity means higher costs of the network since it requires higher amount of connecting wire [25]. It is noteworthy that the number of static wire complexity is not the actual costs of the network. However, it can represent a good indication to which extent the network will cost in terms of number of physical links. Even though HINs provide higher wire complexity than conventional networks, they enable different levels where different number of nodes and wires are served [26]. MTN provides similar number of wires as TTN which is higher than other networks as shown at Table 2. However, this comes with better bisection width and dimeter with respect to same number of nodes.

4. Some Generalizations

Massively parallel computers having exaflops level of computational speed are used in solving many vital scientific and engineering applications. A crucial design issue of MPC is the performance of the underlying interconnection network connecting millions of nodes. Designing an interconnection network having the essential features like low latency, low cost, and high fault tolerance is a challenge. The proposed MTN is a promising HIN showing to possess most of these features. The basic module (BM) of MTN, referred as level 1, is a 2D Torus and the next level i.e. level 2 is formed by the hierarchal interconnection of $2^m \times 2^m$ of level 1. Throughout this paper, we have considered $m = 2$ for the performance evaluation. It is evident that when $m = 2$ provides better network as far as granularity is concerned. Considering $m = 2$, the size of the basic module is 4×4 and the level of hierarchy corresponding to total number of nodes required to connect the respective level of MTN is shown in Table 1. Considering $m = 2$, the highest level of hierarchy is 5 and as many as 1 million nodes can be connected. However, right now the fastest computer Fugaku [2] needs more than 7 million cores. Thus, the future generation massively parallel computers will require tens of millions of cores. MTN is also suitable to support this huge number of cores by increasing the value of m.

In case $m = 3$, the size of the network BM is 8×8 resulting 4,096 nodes in level 2 and level 3 will contain up to 262,144 nodes, thus, it is a coarse granularity. This case might be desirable due to huge number of nodes for each level. Both the maximum level of hierarchy and number of nodes to be interconnected to form those respective levels of hierarchy is depicted in the Table 3. However, it will be complex to reconfigure and maintain redundancy of the network elements for the fault tolerance. It is clear from the above discussion the subnetwork size is increasing with increasing the m. For example, moving further to $m = 4$ the BM size is 16×16 resulting in 65536 nodes even in level-2.

Another parameter that is needed to be considered is the inter-level connectivity, (q). In case $q = 0$ it results in minimal inter-level connectivity where a single channel is used for diagonal incoming, diagonal outgoing, horizontal incoming and horizontal outgoing. This case is illustrated in the Figure 1 and tabulated in the Table 1 in section 2. However, if we could consider the network to have 2 or 3 levels with respect to nodes needed, then different inter-level connectivity could be used. These increase of q are expected to reduce diameter and average distance which in turn reduces the congestion and latency of message transfer. On the other hand, the increase of q increases the bisection width which in turn results a bit complexity in the NoC realization. For example, when a maximum level is needed up to two, we could have eight inter-level connections for diagonal and eight for horizontal. These eight will be divided to four for incoming and four for outgoing connections horizontally and diagonally. The variations of q and their impacts on the illustrated static network parameter in section 3 are illustrated in Table 3.

In this early stage of work on the MTN, we have studied details about the architecture and some static network performance mainly related to the router cost (node degree), wire cost (wiring complexity), and NoC/VLSI realization cost (bisection width). However, our main objective is to evaluate and analyze the network throughput and message latency [27]. These two dynamic communication performance parameters directly depend upon the diameter and average distance. In our immediate next step of the research work, we will evaluate the diameter and average distance. It is believed that the diagonal midimew connection in the higher level of the MTN and torus connection in the basic module will result less hop distance, i.e., diameter and average distance. These torus connection in the basic module yielded many short distance communication links as depicted in the Table 2 and 3.

Table 3. Static Network Performance of the MTN with various inter-level connectivity (q)

	Node Degree	Bisection Width	Arc Connectivity	Wire Complexity	Total Nodes
MTN (2,2,2)	6	32	4	1042	256
MTN (3,2,0)	6	16	4	8,320	4,096
MTN (3,2,1)	6	32	4	8,976	
MTN (3,2,2)	6	64	4	8,704	
MTN (3,2,3)	6	64	4	9,216	
MTN (3,3,0)	6	16	4	5352,608	262,144
MTN (3,3,1)	6	32	4	574,736	
MTN (3,3,2)	6	64	4	557,568	
MTN (3,4,0)	6	16	4	34,087,040	16,777,216
MTN (3,4,1)	6	32	4	36,783,376	
MTN (3,5,0)	6	16	4	2,181,570,688	1,037,741,824
MTN (3,5,1)	6	32	4	2,354,136,336	

5. Conclusion

This paper introduces a new HIN called Midimew connected Torus Network (MTN) for the next generation of Massively Parallel Computer Systems. The paper discussed the network architecture and introduced some generalizations that might be utilized for better features interconnecting very large size interconnection network. In addition, the static network performance evaluation has conducted for MTN, TTN, TESH, 2D Mesh, and 2D Torus networks. It has resulted in the following: constant node degree for all networks, moderate bisection width (neither too high nor too low) for MTN, TESH and TTN, better fault tolerance for torus and a reasonable one for MTN. MTN yields a good static network performance making it a promising hierarchical interconnection network topology. Thus, MTN can be used for high performance MPC systems where significantly huge amount of processing elements can be interconnected together.

This research work is in the early stages of work on MTN. The scope of the paper included discussion of the structure of MTN and its static network performance evaluation. Future work related to MTN will be carried in the light of (1) assessing hop distance performance in terms of diameter and average distance through simulation [24] (2) evaluation of the cost effectiveness analysis [25,26] (3) evaluation of the dynamic network performance [16,27].

References

1. Mishra , Shankarprasad B, Dehuri S. Parallel Computing Environments: A Review. IETE Technical Review. 2011; 28(3): p. 240-247.
2. Erich S, Jack D, Horst S. TOP500.org. [Online].; 2020 [cited 2020 June 23. Available from: https://www.top500.org/lists/top500/2020/06/.
3. The White House. [Online].; 2013 [cited 2020 June 27. Available from: https://obamawhitehouse.archives.gov/administration/eop/ostp/grand-challenges.
4. William J. Dally , Brian T. Introduction to Interconnection Networks. In Penrose DEM, editor. Principles and Practices of Interconnection Networks. San Francisc: Elsevier Morgan Kaufmann; 2010. p. 1-2.
5. Haohuan F, Liao J, Yang L, Wang Z, Song X, Huang C, et al. The Sunway TaihuLight Supercomputer. Science China Information Science. 2016; 59.
6. Dongarra JJ, Gottlieb S, Kramer WTC. Race to Exascale. Computing in Science and Engineering. 2019; 21(1): p. 4-5.
7. William J. Dally , Brian T. Introduction to Interconnection Networks. In Penrose DEM, editor. Principles and Practices of Interconnection Networks. San Francisco: Elsevier Morgan Kaufmann; 2010. p. 45-60.
8. Munayco C, Tariq A, Rothenberg R, SotoCabezas G, Reyes M, Valle A, et al. Early transmission dynamics of COVID-19 in a southern hemisphere setting: Lima-Peru: February 29the-March 30th,2020. Infectious Disease Modelling. 2020 May; 5: p. 338-345.
9. Car Z, Šegota B, Anđelić N, Lorencin I, Mrzljak V. Modeling the Spread of COVID-19 Infection Using a Multilayer Perceptron. Hindawi Computational and Mathematical Methods in Medicine. 2020 May; 2020: p. 1-10.
10. Robson B, Li J, Dettinger R, Peters A, Boyer S. Drug discovery using very large numbers of patents. General strategy with extensive use of match and edit operations. Springer. 2011 May; 25: p. 427–441.
11. King Abdullah University of Science and Technology. [Online].; 2020 [cited 2020 June 27. Available from: https://www.kaust.edu.sa/en/news/shaheen-ii-open-to-serve-covid-19-research-around-the-kingdom.
12. Rahman MMH, Inoguchi Y, Sato Y, Horiguchi S. TTN: A High Performance Hierarchical Interconnection Network for Massively Parallel Computers. IEICE Transactions on Information and Systems. 2009; 92(5): p. 1062-1078.
13. Lai PL, Hsu HC, Tsai CH, Stewart. A class of hierarchical graphs as topologies for interconnection networks. Theoretical Computer Science. 2010; 411(31-32): p. 2912-2924.
14. Abd-El-Barr M, Turki FAS. Topological Properties of Hierarchical Interconnection Networks: A Review and Comparison. Journal of Electrical and Computer Engineering. 2011; 2011.
15. Jain VK, Ghirmai T, Horiguchi S. TESH: A new hierarchical interconnection network for massively parallel computing. IEICE Trans. on Inf. & Syst. 1997; E80-D(9): p. 837-846.
16. Rahman MMH, Inoguchi Y, Sato Y, Miura Y, Horiguchi S. Dynamic communication performance of a TESH network under the nonuniform traffic patterns. In 11th International Conference on Computer and Information Technology; 2008; Khulna. p. 941-951.
17. Ali MNM, Rahman MMH, Nor RM, Behera DK, Sembok TMT, Y. Miura YI. SCCN: A Time Effective Hierarchical Interconnection Network for Network-on-Chip. Mobile Networks and Applications. 2019; 24(4): p. 1-10.
18. Abhinav B, Nikhil J, William D, Gropp , Kale LV. Avoiding hot-spots on two-level direct networks. In Proceedings of International Conference for High Performance Computing, Networking, Storage and Analysis; 2011; Seatle. p. 1-11.
19. Phing NY, Mohd Warip MN, Ehkan P, Zulkefli FW, Ahmad RB. Topology Design of Extended Torus and Ring for Low Latency Network-on-Chip Architecture. TELKOMNIKA (Telecommunication Computing Electronics and Control). 2017; 15(2): p. 869-876.
20. Awal R, Rahman MMH, Akhand MAH. A New Hierarchical Interconnection Network for Future Generation Parallel Computer. In 16th Int'l

Conf. Computer and Information Technology; 2014; Bangladesh. p. 314-319.

21. Solihin Y. Network Topology. In Simon H, editor. Fundamentals of Parallel Multicore Architectur. Oakville: CRC Press, Taylor & Francis Group; 2016. p. 377-380.
22. Singh W, Deb S. Energy Efficient and Congestion-Aware Router Design for Future NoCs. In 29th International Conference on VLSI Design and 2016 15th International Conference on Embedded Systems (VLSID); 2016; Kolkata, India. p. 81-85.
23. Nasirian S, Farhad F. Crystal: A Scalable and Fault-Tolerant Archimedean-Based Server-Centric Cloud Data Center Network Architecture. Computer Communications. 2019; 147: p. 159-179.
24. Rahman MMH, Yukinori S, Yasushi I. High and Stable Performance Under Adverse Traffic Patterns of Tori-Connected Torus Network. Computers and Electrical Engineering, Elsevier. 2013 March; 39(3): p. 973-983.
25. Rahman MMH. HMMN: a cost-effective derivative of midimew-connected mesh network. Int'l Journal of Computers and Applications. 2019 April; 24(4): p. 1-6.
26. Wang T, Wang L, Hamdi M. A cost-effective low-latency overlaid torus-based data center network architecture. Computer Communications. 2018; 129: p. 89-100.
27. Moudi M, Othman M. On the relation between network throughput and delay curves. Automatika. 2020; 61(3): p. 415-424.