3D-TTN: a power efficient cost effective high performance hierarchical interconnection network for next generation green supercomputer

Faiz Al Faisal[1] · M. M. Hafizur Rahman[2] · Yasushi Inoguchi[3]

Abstract

Green computing is an important factor to ensure the eco-friendly use of computers and their resources. Electric power used in a computer converts into heat and thus, the system takes fewer watts ensuring less cooling. This lower energy consumption allows to be less costly to run as well as reduces the environmental impact of powering the computer. One of the most challenging problems for the modern green supercomputers is the reduction of current power consumptions. Consequently, regular conventional interconnection networks also show poor cost performance. On the other hand, hierarchical interconnection networks (like-3D-TTN) can be a possible solution to those issues. The main focus for this paper is the estimation of power usage at the on-chip level for 3D-TTN with the various other networks along with the analysis of static network performance. In our analysis, 3D-TTN requires about 32.48% less router power usage at the on-chip level and can also achieve near about 21% better diameter performance as well as 12% better average distance performance than the 5D-Torus network. Similarly, it also requires only about 14.43% higher router power usage; however, can achieve 23.21% better diameter performance and 26.3% better average distance than recent hierarchical interconnection network- 3D-TESH. The most attractive feature of this paper is the static hop distance parameter and per watt analysis (power-performance). According to our power-performance results, 3D-TTN can also show better result than the 3D-Mesh, 2D-Mesh, 2D-Torus and 3D-TESH network even at the lowest network level. Moreover, this paper is also featured with the static effectiveness analysis, which ensures cost and time efficiency of 3D-TTN.

Keywords 3D-TTN · Estimation of power consumption · Diameter · Average distance · Performance per watt · Cost effectiveness factor · Time cost effectiveness factor

1 Introduction

Green computing reflects the designing, manufacturing/ engineering, using and disposing of computing equipment in a way to reduce their environmental impact. And, reduction of the electric power is the key to achieve the green computing. In 2016, supercomputer PEZY-SCnp at RIKEN

✉ Faiz Al Faisal
faisal@cub.edu.bd

[1] Department of Computer Science and Engineering, Canadian University of Bangladesh (CUB), Bir Uttam Rafiqul Islam Avenue (Progati Soroni), Dhaka 1212, Bangladesh

[2] Department of Computer Networks & Communications, CCSIT, King Faisal University, Al Ahsa 31982, Saudi Arabia

[3] Research Center for Advanced Computing Infrastructure, Japan Advanced Institute of Science and Technology (JAIST), Nomi, Ishikawa, Japan

(Japan) achieved 6673.8 MFLOPS/watt and ranked top in the Green500 list, while beating the Sunway TaihuLight with 6051.3 MFLOPS/watt. On the other hand, the requirement of exa-scale performance is enormous. In fact, today's molecular research in health (specially for the analysis on COVID-19)(For example-world's fastest supercomputer, Fugaku is being used to chose dozens of possible COVID-19 remedies through analysing more than 2000 drugs [1]), nuclear analysis and organic simulation highly depends on the parallel computers. Fugaku supercomputer used Tofu interconnect with 7,299,072 cores and can achieve about 415PFLOPS requiring about 28,335kW power usage [2]. Another very interesting topic could be the use of supercomputers in smart cities. Supercomputers certainly could be very useful for food safety inspections, optimize energy generation, response to emergency situations and even for traffic congestions [3, 4]. In addition, modern MPC systems like-K-computer has already achieved 10.51 petaflops

performance with more than 80,000 computing nodes and also requires about 12.6 MW electrical power using the Tofu interconnect. Therefore, low energy consumption is the most desired choice for the next generation supercomputers with continuing the other constraints like- low network performance, low scalability, low throughput and latency.

The overall performances as well as the power consumption of MPC systems are heavily affected by the interconnection networks and its processing nodes. Interconnection network acts as a communicating path for processing nodes as well as for the memory units [5]. Consequently, every MPC system requires interconnection network as an obvious choice. Vastly used interconnect for the MPC systems is Flat-Tree network [6], which has a big concern in case of network performance. In MPC systems, the total number of outgoing links like-on-chip as well as off-chip links is a big concern, due to power usages as well as high latency [7]. Hence, the interconnect pattern for network topology is a vital issue. Moreover, network topology at the off-chip level should maintain less number of physical outgoing links to reduce the power usage. For example, in modern supercomputers off-chip links for intra-rack requires 0.0101 W and inter-rack requires 0.035 W [8].

The later part of this paper describes about the related research analysis, then the architectural structure of 3D-TTN, reviews the routing algorithm for 3D-TTN in Sect. "Routing algorithm for 3D-TTN", estimates the on-chip power consumption for the 3D-TTN, after that Sect. "Power-performance analysis" shows the analysis of performance verses power and finally Sect. "Cost effectiveness analysis" shows the cost and time-cost effectiveness factor of various networks, while Sect. "Conclusion" concludes the overall outcome of the paper.

2 Related works

Exa-scale performance is the prime goal for next generation supercomputers and most likely next generation high performance computing is solely depends on the massively parallel computers. In contrast, sequential computers are not feasible for meeting the increased computational demand due to its small processing limit for example- according to the Geekbench, Intel Core i9-9980XE can achieve about 1360.0 GFLOPS through its 18 cores [9]. Flat networks like- torus networks show better performance than the mesh networks [10]. However, static electric power for torus network consumes more than the mesh networks due the extra wrap-around connection. Eventually, one of the probable solutions to reduce power consumption as well as to maintain the stable network performance is to use Hierarchical Interconnection Networks (HIN) [11] or undirected interconnection networks [12] or multistage interconnection networks

[13, 14]. However, many HIN networks have already been introduced like- TESH [15], 3D-TESH [5], TTN [16], which are unable to show good performance in comparison to 5D-Torus network and this paper focuses only on the hierarchical networks. Even the networks like- 3D-TESH has proved the power efficiency, but has fallen behind the 5D-Torus in case of performance efficiency due to its mesh connection at the on-chip level. Furthermore, torus networks are more performance efficient over the mesh networks, which is our key motivation for a new network. Hence, in this paper, we like show the detailed analysis of a hierarchical interconnection network- 3D-TTN (Three Dimensional Tori-connected Torus network) was first introduced in 2016 [17], was focused on the power usage and topological analysis rather than the network performance and power-performance comparisons.

3 Network architecture of 3D-TTN

Hierarchical interconnection networks are one of the probable solutions for obtaining the low powered network as well as maintaining suitable network performance with high degree of network scalability. 3D-TTN is a HIN network, which contains multiple basic modules (BMs) that are hierarchically interconnected for higher levels [17].

Definition A BM of 3D-TTN(m, L, q) network is similar to 3D-torus network, which consists of 2^{3m} connected processing elements (PEs) having 2^m rows and $2^m \times 2^m$ columns, where m is a positive integer, L is defined for levels of hierarchy and q is used for inter-level connectivity.

3.1 Basic module

The construction of the lowest level network for 3D-TTN is called as the Basic Module (BM). A $(2^m \times 2^m \times 2^m)$ BM of 3D-TTN has 2^{2m+2} free ports for the higher level interconnected hierarchy. Each BM uses $2^m \times 4 \times (2^q) = 2^{m+q+2}$ of its free links for the upper level networks, where $2(2^{m+q})$ free links are used for the vertical connections and similarly $2(2^{m+q})$ free links for the horizontal connections. Here, q defined as the inter-level connectivity (q ϵ 0,1,..., m). Consequently, according to Fig. 1, a $(4 \times 4 \times 4)$ BM has $2^{2\times2+2} =$ 64 free ports. Moreover, in Fig. 1, node(0,0,0) has two off-chip connectivity of level-2 vertical in and vertical out connections. Similarly, node(0,0,1), node(0,0,2) and node(0,0,3) has both the level-2 vertical in and vertical out connections. And, even node(1,0,1), node(1,0,2) and node(1,0,3) has similar off-chip connection like node(1,0,0) as the level-4 vertical in connection. Figure 1, has skipped some of those links in order to reduce the figure complexity. In this paper, we are particularly analyzing the network class with 3D-TTN(2, L, 0).
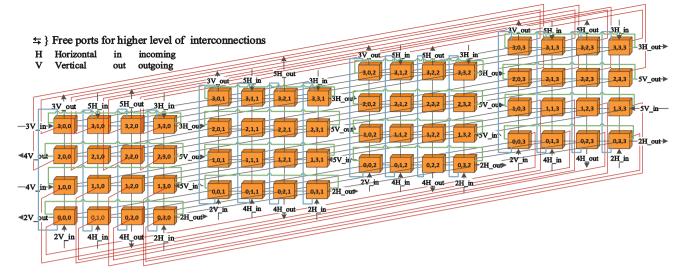
Fig. 1 A Basic Module of 3D-TTN(2, L, 0)

3.2 Higher level connectivity for 3D-TTN

Higher level of 3D-TTN follows the recursive structural pattern of the immediate lowest level of sub-networks (where, 3D-torus is for the on-chip level). Therefore, constructing a level-2 network level-1 network will be used for the 3D-TTN. Figure 2 illustrates the high-level interconnection of 3D-TTN. For example, a level-2 network can be built by ($2^{2 \times 2}$) 16 BMs (16 level-1 3D-TTN). Based on the Table 1, the total number of nodes at level-2 3D-TTN(2, 2, 0) network is N = ($2^8 \times 2^2$) = 1024. In considering the highest level for 3D-TTN network is based upon the ($2^m \times 2^m \times 2^m$) BM by $L_{max} = 2^{m-q} + 1$ (with m = 2 and q = 0, $L_{max} = 5$ and this case total number of nodes will be as N = ($2^{2 \times 2 \times 5} \times 2^2$) = 4,194,304). Table 1 generalizes the various parameters for the 3D-TTN.
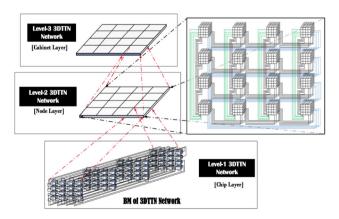
Table 1 Generalization for 3D-TTN

Basic module	Max levels	Number of nodes
($2^m \times 2^m \times 2^m$)	$L_{max} = 2^{m-q} + 1$	$N_L = (2^{2mL} \times 2^m)$

3.3 Addressing of 3D-TTN processing nodes

Node addressing for 3D-TTN requires 3 digit combination for the BM level and 2 digit combination at the higher levels. At the BM level 3D-TTN can be presented by Y-index as the first digit, then for X-index and finally for the Z-index. On the other hand, in case of higher levels; the first digit represents the Y-index and then the X-index for higher levels. In general, a Level-L 3D-TTN can be represented by:

$$A^L = \begin{cases} (a_{yL}, a_{xL}, a_{zL}) & \text{if } L = 1 \\ \\ (a_{yL}, a_{xL}) & \text{if } L_{max} \geq L \geq 2 \end{cases}$$

More generally, in 3D-TTN(m, L, q) the node address is represented by-

$$\begin{aligned} A &= A^L A^{L-1} A^{L-2} \ldots \ldots A^2 A^1 \\ &= (a_{2L}, a_{2L-1}) \ldots \ldots (a_4, a_3)(a_2, a_1, a_0) \end{aligned} \quad (1)$$

Here, the node address (a_2, a_1, a_0) has been defined for the lowest level 3D-TTN, where a_0 has been treated for z-axis of the level-1 network, a_1 is used for x-axis nodes and a_2 has been considered for the Y-directional nodes. On the other hand, the upper levels from level-2 to level-5 networks are contained with two-dimensional structures. Now, a connection path from the source node $n^1 = [(s_{2L}, s_{2L-1})..$

Fig. 2 Higher-level interconnections for 3D-TTN(2, L, 0)

..$(s_4, s_3) (s_2, s_1, s_0)]$ to destination node n^2 included in BM_2 is represented as $n^2 = [(d_{2L}, d_{2L-1})..\,..(d_4, d_3) (d_2, d_1, d_0)]$, are connected if the following connections are satisfied for n^2(when m = 2, q = 0)-

Link for BM-
- $[(s_{2L}, s_{2L-1})..\,..(s_4, s_3) (s_2, s_1, s_0)]$ to $[(s_{2L}, s_{2L-1})..\,..(s_4, s_3) (s_2, s_1, s_0 \pm 1 \mod 2^m)]$ where $2^m > s_0 \geq 0$
- $[(s_{2L}, s_{2L-1})..\,..(s_4, s_3) (s_2, s_1, s_0)]$ to $[(s_{2L}, s_{2L-1})..\,..(s_4, s_3) (s_2, s_1 \pm 1 \mod 2^m, s_0)]$ where $2^m > s_1 \geq 0$
- $[(s_{2L}, s_{2L-1})..\,..(s_4, s_3) (s_2, s_1, s_0)]$ to $[(s_{2L}, s_{2L-1})..\,..(s_4, s_3) (s_2 \pm 1 \mod 2^m, s_1, s_0)]$ where $2^m > s_2 \geq 0$

Link for L2_Vertical-
$[(s_{2L}, s_{2L-1})..\,..(s_4, s_3) (0, 0, s_0)]$ to $[(s_{2L}, s_{2L-1})..\,..(s_4 \pm 1 \mod 2^m, s_3) (0,0,s_0)]$

Link for L2_Horizontal-
$[(s_{2L}, s_{2L-1})..\,..(0, 2^m - 1, s_0)]$ to $[(s_{2L}, s_{2L-1})..\,..(s_4, s_3 \pm 1 \mod 2^m) (0, 2^m - 1, s_0)]$

Link for L3_Vertical-
$[(s_{2L}, s_{2L-1})..\,..(s_6, s_5)..(2^m - 1, 0, s_0)]$ to $[(s_{2L}, s_{2L-1})..\,..(s_6 \pm 1 \mod 2^m, s_5)..(2^m - 1, 0, s_0)]$

Link for L3_Horizontal-
$[(s_{2L}, s_{2L-1})..\,..(s_4, s_3) (2^m - 1, 2^m - 1, s_0)]$ to $[(s_{2L}, s_{2L-1})..\,..(s_6, s_5 \pm 1 \mod 2^m)..(2^m - 1, 2^m - 1, s_0)]$

Here, addressing for 3D-TTN has been defined for level-1 to level-3 network. However, similarly we can also define the upper level interconnections as well as increased interconnectivity with q = 1 or q = 2.

4 Routing algorithm for 3D-TTN

A simple deterministic dimension-order routing (DOR) algorithm has been considered for 3D-TTN [17]. In dimension-order routing a packet starts its routing from the source node to the destination though checking the same
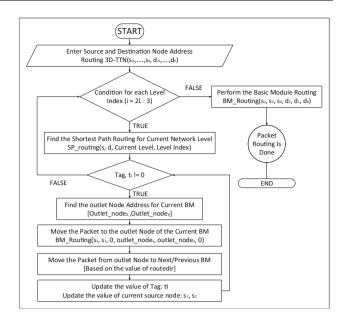

Fig. 3 Packet routing for 3D-TTN

BM first. However, if it is destined for another BM, then the source node will send the packet to the outlet_node which connects the outer BM at which the routing will be performed. The function SP_routing will help to find the shortest route for the higher levels. Now, if we consider a source node as s = $[(s_{2L}, s_{2L-1})..\,..(s_4, s_3) (s_2, s_1, s_0)]$ and destination node d = $[(d_{2L}, d_{2L-1})..\,..(d_4, d_3) (d_2, d_1, d_0)]$, then routing tag can be defined as t = $[(t_{2L}, t_{2L-1})..\,..(t_4, t_3) (t_2, t_1, t_0)]$. Algorithm 4.1 shows the routing algorithm for 3D-TTN. Moreover, in order to simplify the routing algorithm, we have showed the overall packet routing through Fig. 3 packet routing flowchart. Figure 3 shows the step by step packet routing for each corresponding level. Routing algorithm requires the source and destination node address and tag value is based s_i and d_i. If the tag value t_i is not zero, then upper level routing is required and packet will be moved to next BM based on the value of routedir. Next, if any packet moves to a new source node, tag value as well as the current source address will be updated. This process continues for all the value of level index, i. Finally, routing will be done at the basic module level.

Algorithm 4.1 Routing Algorithm for 3D-TTN [17]

```
Routing 3D-TTN(s_{2L}, s_{2L-1}, ..., s_1, s_0, d_{2L}, d_{2L-1}, ..., d_1, d_0);
tag: t_{2L}, t_{2L-1}, t_{2L-2}, ...., t_1, t_0;
        for i = 2L : 3;
                routedir = SP_routing(s, d, ⌊(i − 1)/2 + 1⌋, i);
                if (routedir = positive) then t_i = ((d_i - s_i + 2^m) mod 2^m);
                else t_i = (2^m - (d_i - s_i + 2^m) mod 2^m); endif;
                while (t_i ≠ 0) do
                        if (i mod 2) = 1, then
                                outlet_node_x  =  outlet_x(s, d, ⌊(i − 1)/2 +
1⌋, H, routedir);
                                outlet_node_y  =  outlet_y(s, d, ⌊(i − 1)/2 +
1⌋, H, routedir);
                        else    outlet_node_x  =  outlet_x(s, d, ⌊(i − 1)/2 +
1⌋, V, routedir);
                                outlet_node_y  =  outlet_y(s, d, ⌊(i − 1)/2 +
1⌋, V, routedir);
                        endif;
                        BM_routing(s_2, s_1, 0, outlet_node_y, outlet_node_x, 0);
                        if (routedir = positive) then send the packet to the next BM;
                        else move the packet to previous BM; endif;
                        if (t_i > 0) then t_i = t_i - 1; endif;
                        if (t_i < 0) then t_i = t_i + 1; endif;
                        if (i mod 2) = 1, then
                                s_1 = receiving_node_x(s, d, ⌊(i−1)/2+1⌋, H, routedir);
                                s_2 = receiving_node_y(s, d, ⌊(i−1)/2+1⌋, H, routedir);
                        else  s_1 = receiving_node_x(s, d, ⌊(i−1)/2+1⌋, V, routedir);
                                s_2 = receiving_node_y(s, d, ⌊(i−1)/2+1⌋, V, routedir);
                        endif;
                endwhile;
        endfor;
        BM_routing(s_2, s_1, s_0, d_2, d_1, d_0);
end
BM_routing(s_2, s_1, s_0, d_2, d_1, d_0);
source: s_2, s_1, s_0; destination: d_2, d_1, d_0;
BM_tag: t_2, t_1, t_0 = destination address(d_2, d_1, d_0) - source address(s_2, s_1, s_0);
        for i = 0 : 2;
                if (t_i > 0), movedir = positive; endif;
                if (t_i < 0), movedir = negative; endif;
                if (movedir = positive and t_i > 2^m/2) then t_i = t_i - 2^m; endif;
                if (movedir = negative and t_i < -2^m/2) then t_i = t_i + 2^m; endif;
        endfor;
        while(t_0 ≠ 0) do
                if (t_0 > 0) then move packet to +z node; t_0 = t_0 - 1; endif;
                if (t_0 < 0) then move packet to -z node; t_0 = t_0 + 1; endif;
        endwhile;
        while(t_1 ≠ 0) do
                if (t_1 > 0) then move packet to +x node; t_1 = t_1 - 1; endif;
                if (t_1 < 0) then move packet to -x node; t_1 = t_1 + 1; endif;
        endwhile;
        while(t_2 ≠ 0) do
                if (t_2 > 0) then move packet to +y node; t_2 = t_2 - 1; endif;
                if (t_2 < 0) then move packet to -y node; t_2 = t_2 + 1; endif;
        endwhile;
end
SP_routing(s, d, Level, i);
        if ((d_i - s_i + 2^m) mod 2^m) > 2^m/2, then routedir = negative;
        elseif ((d_i - s_i + 2^m) mod 2^m) = 2^m/2, {
                if (Level mod 2) = 0, then
                        if (i mod 2) = 0, then routedir = positive;
                        else routedir = negative; endif;
                else
                        if (i mod 2) = 0, then routedir = negative;
                        else routedir = positive; endif;
                endif; }
        else routedir = positive; endif;
end
```

5 Estimation of power consumption

Reduction of the power usages is the most desirable target for the supercomputers. Sunway TaihuLight System has achieved about 93 PFLOPS (requiring about 15.3M W electrical power) performance with about 10.65M cores considering the 2D mesh network for the interconnectivity of its cores [18]. This section considers only the on-chip electrical power analysis.

5.1 Assumptions for power model

Power requirement at the on-chip network level can be up to 50% of total chip power usage [19]. Hence, this paper focuses only the on-chip power estimation for 3D-TTN. The power consumption for on-chip network has been estimated by the leakage and dynamic power consumption model for both the links and routers using an on-chip power model simulator. Therefore, in the on-chip level, we have considered all the interior links for 3D-TTN at the BM level. One of the interesting features of this paper is the power estimation of 5D-Torus network (used in Blue Gene/Q supercomputer) [6]. Apart from this, another attractive feature of our paper is to show the performance per watt for 3D-TTN, which has also been evaluated by the on-chip power usage.

5.2 On-chip power model

On-chip power model for this paper considered the Orion energy model [20] using 65nm fabrication process for the 3D-TTN and others. To simulate the on-chip model, we have used the GARNET network simulator [21] along with Orion energy model. In order to integrate the network simulation along with the Orion energy model evaluation, we have considered the GEM5 [22], which is a full system simulator specially designed for computer architecture research. Gem5 requires Unix platform in order to build and run simulations [22]. The dynamic power and leakage power are the main source of power consumption. Hence, both the router and link power dissipation are entirely responsible for total power consumption. Router total energy depends on the read and write operations in buffers, energy consumption by the total activity at the local and global arbiters and then for the total number of crossbar traversals. Equation 2 shows the total energy consumption inside the router [21]. On the other hand, the dynamic energy is defined by $E = 0.5\alpha C V^2$, where α is the switching activity, C is the capacitance and V is the supply voltage [20]. Dynamic power of physical links is evaluated through the charging and discharging of capacitive loads. In a CMOS circuits link power formulated as $P = E f_{clk}$, where f_{clk} is the clock frequency. Hence, the link dynamic power is defined as, $P_{link} = \alpha C_1 V_{dd}^2 f_{clk}$, where C_1 is the load capacitance. However, static power of physical links is due to the inserted repeaters.

$$E_{router} = E_{buffer\_write} + E_{buffer\_read} + E_{vc\_arb} + E_{sw\_arb} + E_{xb} \quad (2)$$

Now, considering the clock frequency with 1 GHz, supply voltage 1.0 V, 128 bits message size and uniform traffic pattern with 2 mm per link length, Table 2 shows the

Table 2 Simulation condition for power analysis

Parameter	Value	Units
Fabrication process	65 nm	–
Link length	2.00	[mm]
Operating freq.	1×10^9	Hz
Transistor type	NVT	–
Supply voltage	1.00	V
Traffic pattern	uniform traffic	–
Message inject rate	0.01	flits/cycle/node
Message size	128	bits
Simulation Cycle	20,000	–

Fig. 5 Router power analysis of various networks
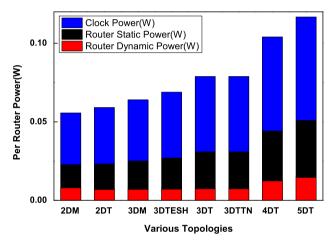
Fig. 6 Per node power analysis of various networks

simulation condition for the on-chip 3D-TTN. On the other hand, Fig. 4 shows the power dissipation for total number of links on various networks and Fig. 5 shows the power dissipation for total number of routers for various networks with only one virtual channel. However, in both the figures the total number of nodes/routers or links is much higher than the 3D-TTN (64 nodes (2DM, 2DT, 3DM, 3DTESH, 3DT have also 64 nodes)) comparing with 4D-Torus (256 nodes) and 5D-Torus (512 nodes) network due to the minimum node requirement for 4D-Torus and 5D-Torus at level-1 network. This is the prime reason why Figs. 4 and 5 is showing high power usage for 4D-Torus and 5D-Torus. In order to compare with 4D-Torus and 5D-Torus, Fig. 6 shows the per router power consumption and has been compared with respect to static and dynamic power. Moreover, it is expected to link power will have also have a large impact in comparing with 3D-TTN (router radix is 8) as the router radix for 5D-Torus is 10. Figure 6 explains that considering per router consumption of 4D-Torus requires about 24.21% and for 5D-Torus
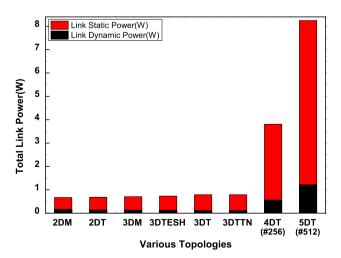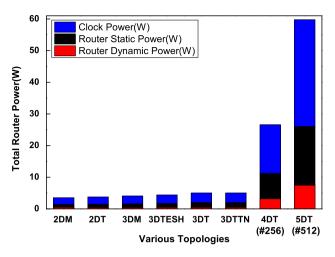
network is about 32.48% higher electric power than the 3D-TTN.

6 Power-performance analysis

Performance per watt can be one of the most attractive features for the supercomputers. As the modern MPCs are highly affected by the electrical power consumptions, Performance per watt can able to trace the system performance with respect to power, which is a relatively new feature in the field of interconnection network. The choice of this parameter has been taken from the observation of scenario where network with little poor performance, but having much better power efficiency, had always been rejected. Similarly, another motivation for this parameter came from the Green500 list, where listed supercomputers are required to ensure the energy-efficiency through gigaflops/watt parameter [23]. By definition, performance per

Fig. 4 Link power analysis of various networks

watt can be defined by the ratio between the network performances against the total power consumption. Regarding the definition of performance, it can be treated as the static or dynamic network performance of the corresponding network. In this paper, we have considered the only the static network performance for analyzing the performance per watt.

$$\text{Performance per watt} = \frac{\text{Achieved Performance}}{\text{Total Power Usage}} \quad (3)$$

6.1 Diameter power-performance analysis

In this section, the diameter performance has been used against the total router power consumption. In massage passing, source node should communicate by the curtain route to transmit the data from source to destination, which may not be directly connected. Shortest routed path is expected for interconnection networks as the increased routing path also increases the communication delay. Diameter is the maximum inter-node distance between all distinct pairs of nodes along the shortest path. Interconnection network with smaller diameter is preferable [24]. Now, to evaluate the calculated value for diameter performance of 3D-TTN(2, L, 0) Eq. 4 can be used-

$$Dia. = max(D_z + D_s + (\sum_{i=2}^{L} (D_{si} + D_i)) + D_d) \quad (4)$$

Here, D_z is considered the total path required at the Z-directions, D_s is the value to move to the outgoing node of highest level, then D_{si} used as the value to go to next level of routing, D_i is for corresponding level routing and finally D_d is for level-2 receiving node to the destination node. Table 3 shows the calculated analysis of equation 4.

Figure 7 shows the diameter for the 3D-TTN, which explains that it is much better than the 2D or 3D mesh and torus networks. It has outperformed the 3D-TESH [5] and about 21% better performance than the most considerable recent interconnection network 5D-Torus [25] at the maximum level as per our analysis. However, equation 5 defines the consideration for diameter performance per watt considering the router power. Total router power usage depends on the total router leakage and dynamic power usage with also the clock power obtained from the Fig. 5. On the other hand, equation 6 defines the diameter
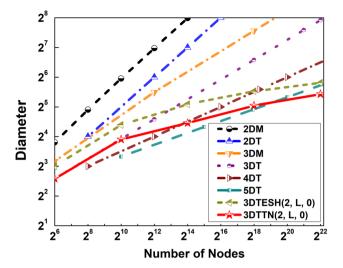

Fig. 7 Diameter of various networks

performance per watt considering the link power (considering the same simulation condition as Table 2).

$$\text{Dia. Perf. Per Watt(Router)} = \frac{\text{Diameter Performance}}{\text{Total Router Power}} \quad (5)$$

$$\text{Dia. Perf. Per Watt(Link)} = \frac{\text{Diameter Performance}}{\text{Link (Static + Dynamic) Power}} \quad (6)$$

Fig. 8 (considering router power usage) and Fig. 9 (considering link power usage) shows those result analysis, where 3D-TTN(#64) shows much better performance per watt than the 2D-Mesh(#64), 2D-Torus(#64), 3D-Mesh(#64), 3D-TESH (#64) network. Here, in Figs. 8 and 9, we have showed the same analysis for 4D-Torus with 256 nodes and 5D-Torus with 512 nodes (according to their lowest network level) and it is obvious that having high node number 4DT and 5DT will outperform others. As the small diameter is preferable hence the network with low performance per watt for diameter is more desirable. Moreover, 2D-Mesh network (used in Sunway TaihuLight System having 10,649,600 cores and achieving about 6 Gflops/W with 8 × 8 mesh network [26]]) shows the worst diameter performance per watt than the 3D-TTN.

Table 3 Calculated formulation of diameter for 3D-TTN

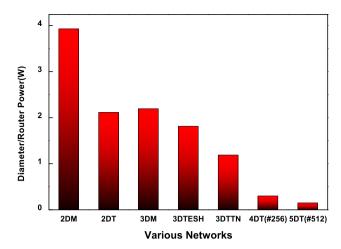| Network	D_z	D_s	D_{si} and D_i	D_d	Diameter
Level-1	2	4	$D_{si} = 0$, $D_i = 0$	0	6
Level-2	2	4	for i = 2: $D_{si}=0$, $D_i=5$	4	15
Level-3	2	4	for i = 2: $D_{si} = 2$, $D_i = 5$; for i = 3: $D_{si} = 0$, $D_i=5$	4	22

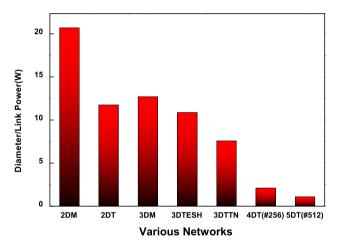**Fig. 8** Diameter performance per watt (Router)

Fig. 9 Diameter performance per watt (Link)

6.2 Average distance power-performance analysis

In case of interconnection networks, low average distance is more preferable over the diameter due to the communication patterns, where every node requires to communicate with the every other [27]. The average distance can be treated as the mean distance between all distinct pairs of nodes in a network. Hence, it is expected for the networks to have small average distance. The average distance for 3D-TTN(2, L, 0) is shown in the Fig. 10, which confirms that the average distance performance of our network is far better than conventional 2D or 3D networks. Considering our network with the 5D-Torus network, it also shows about 12% better performance at the maximum level. Now, average distance performance per watt for the MPC systems can be defined by the achieved average distance over the total power usage. Equation 7 shows the average distance performance per watt with regards to total router power usage and equation 8 shows average distance
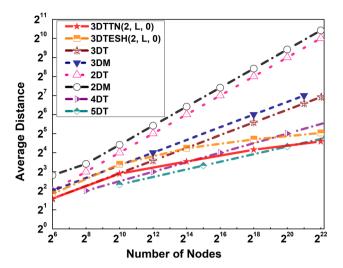

Fig. 10 Average distance of various networks

performance per watt with regards to total link power usage (considering the same simulation condition as Table 2).

Fig. 11 (considering 64 routers power consumption) and Fig. 12 (considering total link power usage) show the average distance performance per watt for various networks, which illustrates that 3D-TTN(#64) can show better performance than 2D-Mesh(#64), 2D-Torus(#64), 3D-Mesh(#64) and 3D-TESH(#64) network with respect to total router and link power usage. Here, also (like Figs. 8 and 9), in Figs. 11 and 12, we have showed same analysis for 4D-Torus with 256 nodes. As smaller average distance is preferable, 3D-TTN has obviously outperformed 2D-Mesh, 2D-Torus and 3D-TESH network.

$$\mathrm{Av. Dist. Perf. Per Watt (Router)} = \frac{\mathrm{Average\,Dist.\,Performance}}{\mathrm{Total\,Router\,Power}}$$
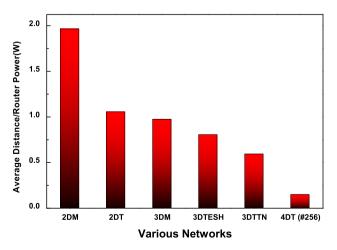(7)

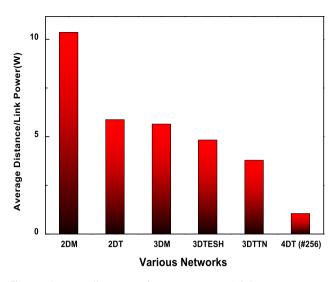**Fig. 11** Average distance performance per watt(Router)

Fig. 12 Average distance performance per watt(Link)

$$\text{Av. Dist. Perf. Per Watt(Link)} = \frac{\text{Average Dist. Performance}}{\text{Link (Static + Dyna.) Power}} \tag{8}$$

7 Cost effectiveness analysis

Cost effectiveness analysis ensures a suitable profit over the faster solution for curtain number of processors. Considering the total number of communication links in the system, this analysis shows two measures- cost-effectiveness and time-cost effectiveness.

7.1 Cost-effectiveness factor

Speedup and efficient computing are the common parameters that have been used for the performance evaluation for MPC systems. However, number of communication links is a big concern for the MPC systems due to the total system cost. System cost are not only depends on the number of processors, but also through the communication links [28]. Hence, cost-effectiveness factor can be handy for the MPCs. This parameter considers system cost through the communication links. The cost-effectiveness factor (CEF) for 3D-TTN has been defined by the equation 9.

$$\text{CEF(N)} = \frac{1}{1 + \rho \times \dfrac{\text{Total communication links}}{\text{Total number of nodes}}} \tag{9}$$

Now, suppose the cost for a single processor including its processing unit, control unit and memory unit is defined by C_p and C_L is the cost for single communication link, then ρ is defined by the ratio of C_L against C_p. Fig. 13 shows the CEF for various networks (for $\rho = 0.1$), which explains that
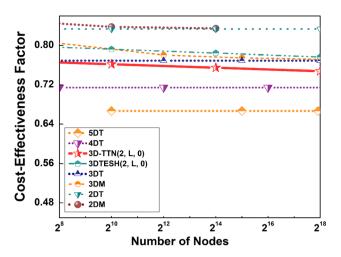

Fig. 13 CEF of various networks ($\rho = 0.1$)

the cost-effectiveness factor for 3D-TTN is better than the 2D as well as 3D mesh and torus networks and little worst than 4DT and 5DT networks with obvious high wiring complexity. It has also outperformed 3D-TESH network with a big margin. On the other hand, Fig. 14 shows the CEF for various networks with variable ρ, which explains that the cost-effectiveness factor for 3D-TTN is better than the 2D as well as 3D mesh and torus networks. However, little poor than the 4DT and 5DT network due to the low wiring complexity of 3D-TTN.

7.2 Time-cost-effectiveness factor

The requirement for a MPC system to show the time efficiency for any kind of program can be obtained from the time-cost-effectiveness factor (TCEF) [28]. A faster solution is more desirable than the low cost effectiveness for the MPC systems. Hence, TCEF can be a useful parameter in order to characterize any interconnection network. The
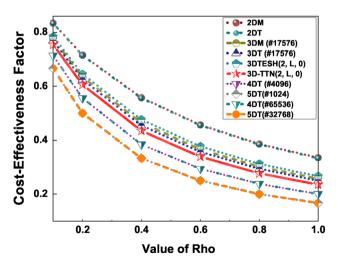

Fig. 14 CEF of various networks

TCEF of 3D-TTN has been shown in equation 10, where ρ is defined by the ratio of C_L against C_p. T_1 has been used for the time to solve a single problem by a single processor & α is a linear time penalty in T_p. T_p is the time which is required by p processing nodes to solve a single problem.

$$G(p) = \frac{\text{Total number of communication links}}{\text{Total number of nodes}}$$

$$\text{TCEF}(p, T_p) = \frac{1 + \sigma T_1^{\alpha-1}}{1 + \rho\, G(p) + \frac{\sigma}{p} T_p^{\alpha-1}} \quad (10)$$

Fig. 15 shows the TCEF for 3D-TTN, which is better than any 2D or 3D networks of mesh and torus connections and slightly worst than 4D or 5D networks. Even it has outperformed the other hierarchical interconnection networks like- 3D-TESH(2, L, 0). As TCEF considers time for the solution of a problem, 3D-TTN can produce a faster solution together with increased profit. Similar to Fig. 14, Fig. 16 shows the TCEF analysis with variable ρ values, which depicts that 3D-TTN is an obvious choice over other 2D and 3D networks and highly comparable with 4D torus network even having much low wiring complexity.

8 Conclusion

In conclusion, in this paper there are three contributions; power analysis, static hop distance parameter and performance per watt analysis, and static cost effectiveness analysis. However, our main objective was to find an interconnection network, which achieves high performance as well as reducing the current power usages for MPC systems and also to introduce parameter (performance per watt) in the field of interconnection network that could show the analysis of system performance against power usage.
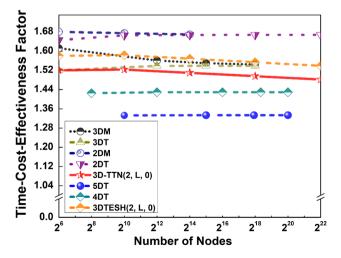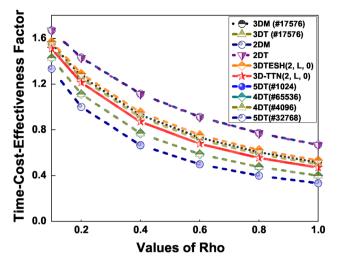

Fig. 16 TCEF of various networks ($\sigma = 1$)

Power efficiency for 3D-TTN has been compared with the various networks, which shows that 3D-TTN requires 24.21% less router power usage than the 4D-Torus network and also 32.48% less router power usage than the 5D-Torus network (explained in Sect. "Estimation of power consumption"). In contrast, it requires about only 14.43% higher router power usage than the 3D-TESH network. However, if we consider power-performance analysis for diameter and average distance in comparing with 3D-TTN, 2D-Mesh (69.74% worst router diameter power-performance), 2D-Torus (43.72% worse router diameter power-performance) and even 3D-TESH (34.46% worse router diameter power-performance) shows the worse result than the 3D-TTN (explained in Sect. "Power-performance analysis"). On the other hand, 4D-Torus and 5D-Torus networks with 256 nodes and 512 nodes (lowest network level nodes), in comparing with the only 64 nodes of 3D-TTN (lowest network level nodes), definitely will show better power-performance for their high power usage. Moreover, this research focuses only the on-chip power usages. As 3D-TTN shows better performance at the higher level, it is expected that it will show much better performance per watt at the upper level network. Concerning the analysis on diameter and average distance, 3D-TTN has outperformed the latest HIN network 3D-TESH with a big margin in case of diameter (23.21%) and average distance (26.3%). It has achieved near about 53% better diameter and 47% better average distance than the 4D-Torus network over 4 millions of nodes (explained in Sect. "Power-performance analysis"). Comparing with the 5D-Torus network, it has also outperformed 5D-Torus by near about 21% diameter and 12% average distance performance at over 4 millions of nodes. Now, considering cost-effectiveness (CEF) and time-cost-effectiveness (TCEF) parameters, it is obviously better choice than the 2D and

Fig. 15 TCEF of various networks ($\sigma = 1$, $\rho = 0.1$)

3D mesh and torus, and also 3D-TESH network (explained in Sect. "Cost effectiveness analysis"). In addition, having higher communication links than 3D-TESH network, 3D-TTN also shows better CEF and TCEF results than those other interconnection networks of 2D and 3D mesh and torus networks along with 3D-TESH for variable ρ values. On the other hand, with 16,384 nodes 3D-TTN have the wiring complexity of 53,248 links whereas 4D-Torus network requires 58,564 links for its only 14,641 nodes, which is a 9.08% higher interconnected links with having 1743 less nodes. This exemplifies the main cause for the slight better performance of 4DT and 5DT in CEF and TCEF over 3D-TTN. Issues for future work include the following: (1) evaluation of dynamic network performance, (2) fault-tolerant analysis and (3) assessment of the performance improvement for 3D-TTN with an adaptive routing algorithm.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Japan's Fugaku Supercomputer, World's Fastest. https://www.japantimes.co.jp/news/2020/07/04/national/science-health/japan-fugaku-supercomputer-coronavirus-drugs/. Accessed 25 June 2020
2. Fugaku Supercomputer, Tofu Interconnect, Fujitsu. https://en.wikipedia.org/wiki/Fugaku_(supercomputer). Accessed 25 June 2020
3. Liu, Y., Feng, T., Peng, M., Guan, J., Wang, Y.: DREAM: online control mechanisms for data aggregation error minimization in privacy-preserving crowdsensing. IEEE Trans. Depend. Secure Comput. (2020). https://doi.org/10.1109/TDSC.2020.3011679
4. Liu, Y., Wang, H., Peng, M., Guan, J., Wang, Y.: An incentive mechanism for privacy-preserving crowdsensing via deep reinforcement learning. IEEE Internet Things J. (2020). https://doi.org/10.1109/JIOT.2020.3047105
5. Al Faisal, F., Rahman, M.H., Inoguchi, Y.: A new power efficient high performance interconnection network for many-core processors. J. Parallel Distrib. Comput. **101**, 92–102 (2017)
6. Minkenberg, C.: IBM Research. http://www.systems.ethz.ch/
7. Pavlidis, V.F., Friedman, E.G.: 3-D topologies for networks-on-chip. IEEE Trans. VLSI syst. **15**, 1081–1090 (2007)
8. Bergman, K.: Silicon photonics for extreme scale computing. https://press3.mcs.anl.gov/
9. Geekbench 4 benchmarks platform, Processor performance. https://gadgetversus.com/processor/intel-core-i9-9900ks-gflops-performance/. Accessed 15 June 2020
10. Miura, Y., Shimozono, K., Watanabe, S., Matoyama, K.: An adaptive routing of the 2-D torus network based on turn model. In: Proceedings of the International Symposium on Computing and Networking. pp. 587–591 (2013)
11. Hafizur Rahman, M.M., Horiguchi, S.: HTN: a new hierarchical interconnection network for massively parallel computers. IEICE Trans. Inf. Syst. **E86**, 1479–1486 (2003)
12. Anjali, N., Mathew, S.: A generalization of block structures in undirected interconnection networks. J. Interconnect. Netw. **18**, 1850008 (2018)
13. Panda, D.K., Dash, R.K.: Network reliability evaluation and analysis of multistage interconnection networks. Int. J. Eng. Adv. Technol. (IJEAT) **8**(3), 1729–1737 (2019)
14. Multistage Interconnection Networks (MINs). https://en.wikipedia.org/wiki/Multistage_interconnection_networks . Accessed 15 June 2020
15. Jain, V.K., Ghirmai, T., Horiguchi, S.: TESH: a new hierarchical interconnection network for massively parallel computing. IEICE Trans. Inf .Syst. **E80**, 837–846 (1997)
16. Hafizur Rahman, M.M., Sato, Y., Inoguchi, Y.: High and stable performance under adverse traffic patterns of tori-connected torus network. Comput. Electr. Eng. **39**, 973–983 (2013)
17. Al Faisal, F., Rahman, M.H., Inoguchi, Y.: Power analysis with variable traffic loads for next generation interconnection networks. In: Proceedings of the IEEE 18th International Conference on High Performance Computing and Communications. Sydney, NSW, Australia (2016)
18. Dongarra, J.: Report on the Sunway TaihuLight System, University of Tennessee, Oak Ridge National Laboratory (2016)
19. John, E.: Juan Rubio. CRC Press, Unique chips and systems (2007)
20. Kahng, A.B., Li, B., Peh, L.S., Samadi, K.: ORION 2.0: a power-area simulator for interconnection networks. IEEE Trans. VLSI Syst. **20**, 191 (2011)
21. Agarwal, N., Krishna, T., Peh, L.-S., Jha, N.K.: GARNET: a detailed on-chip network model inside a full-system simulator. In : Proceedings of the IEEE International Symposium on Performance Analysis of Systems and Software. p. 33–42 (2009)
22. The gem5 simulator, Computer-system architecture platform. https://www.gem5.org/. Accessed 10 Mar 2021
23. The Green 500, Energy-efficient computers. http://www.green500.org/. Accessed 15 June 2020
24. Rahman, M.H., Inoguchi, Y., Al-Faisal, F., Kundu, M.K.: Symmetric and folded tori connected torus network. J. Netw. **6**, 26–35 (2011)
25. Chen, D., Eisley, N.A., Heidelberger, P.: The IBM Blue Gene/Q interconnection network and message unit. In: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC). pp. 1–10 (2011)
26. Fu, H., Liao, J., Yang, J., Wang, L., Song, Z., Huang, X., Yang, C., Xue, W., Liu, F., Qiao, F., Zhao, W.: The sunway TaihuLight supercomputer: system and applications. Sci. China Inf. Sci. **59**, 1–16 (2016)
27. Abd-El-Barr, M., Al-Somani, T.F.: Topological properties of hierarchical interconnection networks: a review and comparison. J Electr. Compt. Eng. (2011). https://doi.org/10.1155/2011/189434
28. Sarkar, D.: Cost and time-cost effectiveness of multiprocessing. IEEE Trans. Parallel Dist. Syst. **4**(6), 704–712 (1993)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Faiz Al Faisal received his B.Sc. degree inComputer Science and Engineering from KhulnaUniversity of Engineering and Technology (KUET),-Khulna, Bangladesh, in 2009. He received his M.Sc.degree in Information Science from JapanAdvanced Institute of Science and Technology(JAIST), Japan, in 2015. He has also persuaded thedoctoral degree in Information Science at JapanAdvanced Institute of Science and Technology(JAIST) in 2018. Currently, Dr. Faisal is working as an assistantprofessor in the Dept. of Computer Science and Engineering (CSE) atCanadian University of Bangladesh (CUB), Dhaka, Bangladesh. Beforejoining CUB, he worked as an assistant professor in North SouthUniversity (NSU), Dhaka, Bangladesh. Prior to his M.Sc. graduation, heserved as an IN-VAS associate Engineer at Orascom TelecomBangladesh Limited in 2011–2013. He was also provisioned as a softwareengineer at Leads Corporation Limited, Bangladesh, in 2010. His currentresearch interest has been centered on parallel computer systems,interconnection networks and machine learning algorithms.

M. M. Hafizur Rahman received his B.Sc.degree in EEE from KUET, Khulna, Bangladesh, in1996. He received his M.Sc. and Ph.D. degrees inInformation Science from the JAIST in 2003 and2006, respectively. Dr. Rahman is currentlyworking as an assistant professor in the Dept. ofCN, CCSIT, KFU, Saudi Arabia. Prior to join in theKFU, he was assistant professor in the XiamenUniversity, Malaysia & IIUM, Malaysia andassociate professor in the CSE, KUET, Khulna, Bangladesh. He was alsoa visiting researcher in the School of Information Science at JAIST and aJSPS postdoctoral research fellow at Graduate School of InformationScience (GSIS), Tohoku University, Japan & Center for InformationScience, JAIST, Japan in 2008 and 2009 & 2010-2011, respectively. Hiscurrent research interests include hierarchical interconnection networksand optical switching network, software defined network.

Yasushi Inoguchi received his B.E. degree fromDepartment of Mechanical Engineering, TohokuUniversity in 1991, and received MS degree and Ph.Dfrom JAIST (Japan Advanced Institute of Science andTechnology) in 1994 and 1997, respectively. He is aProfessor of Research Center for Advanced ComputingInfrastructure, JAIST. He was a research fellow of theJapan Society for the Promotion of Science from 1994to 1997. From 2002 to 2006 he was a researcher ofPRESTO program of Japan Science and Technology Agency. From 2008to 2009 he was a courtesy senior research scholar at University of SouthFlorida. His research interest has been mainly concerned with parallelcomputer architecture, interconnection networks, GRID/Cloudarchitecture, and high performance computing on parallel machines. Dr.Inoguchi is a member of IEEE, IEICE & IPS of Japan.