http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 13 (4): 795-800, 2014
ISSN 1812-5638 / DOI: 10.3923/itj.2014.795.800

Research of Network Properties and Power Consumption on THIN

[1]Baojun Qiao, [1]Shengbin Liang and [2]Wei Wei
[1]College of Software, Henan University, Kai Feng, 475008, China
[2]School of Computer Science and Engineering, Xi'an University of Technology, Xi'an, Shaanxi, 710048, China

Abstract: A new chip design paradigm called Network-on-Chip (NoC) offers a promising architectural choice for future System-on-Chip (SoC). Communication latency and power efficiency are the most important concerns in NoC architecture design. Triple-based Hierarchical Interconnection Network (THIN) was proposed that aims to decrease the node degree, reduce the links and shorten the diameter. In this study, the zero-load latency and energy consumption are thoroughly studied and compared with 2-D mesh. The compare results show that THIN is a better candidate for constructing the NoC than 2-D mesh, when there are not too many nodes.

Key words: Triple-based hierarchical interconnection network, communication latency, power consumption

INTRODUCTION

With the advance of the VLSI technology, future System-on-Chip (SoC) will integrate from several dozens to hundreds of cores in a single billion-transistor chip and the on-chip communication is soon becoming the bottleneck. Bus-based architecture suffers from the clear bottleneck of the share media used for the transmission. The bus allows only one communication at a time, all the cores in the system share its bandwidth and its operating frequency decreases with the system growth.

Network on Chip (NoC), a new chip design paradigm concurrently proposed by many research groups (Benini and De Micheli, 2002; Kumar *et al*., 2002), is expected to be an important architectural choice for future SoCs. Using a network to replace global wiring has advantages of structure, performance and modularity. Performance and power efficiency are the important concerns in NoC architecture design. Consider a 10×10 tile-based NoC, assuming a regular mesh topology and 32 bit link width in 0.18 um technology and minimal spacing, under 100 Mbit sec^{-1} pair-wise communication demands, interconnects will dissipate 290W of power (Hu *et al*., 2005). Thus, reducing the power consumption on global interconnects is a key factor to the success of NoC designs.

A new interconnection architecture named THIN (Triple-based Hierarchical Interconnection Network) was proposed by Qiao *et al*. (2007) and Shi *et al*. (2006). THIN is a particular case of WK-recursive topology whose basic modules are 3-node complete graph. THIN offers a high degree of regularity, scalability and symmetry which very well conform to a modular design and implementation of NoC. In this study, we thoroughly studied the network properties and power consumption of THIN and compared them with 2-D mesh from a theoretical perspective.

NETWORK TOPOLOGY PROPERTIES OF THIN

THIN is a hierarchical and scalable interconnection network and it emphasizes particularly on decreasing the node degree, reducing the links and shortening the diameter. Figure 1 shows the topology of THIN. As shown in Fig. 1a, we define a single node as a level 0 THIN. A level 1 THIN can be constructed by connecting three nodes with three communication channels and then forming a triangle, as shown in Fig. 1b. The level 1 THIN is the base component to form any level THIN.

THIN is easily scalable and the constructing process is: replacing the node of level 1 THIN with lower level THIN to structure a higher one, reiterating this process, we can get any level THIN, as illustrated in Fig. 2.

The topology of THIN is very simple and the node degree is very low. THIN has obviously hierarchical, symmetric and scalable characteristic. The nodes in the lowest hierarchy of THIN are fully connected, whilst other hierarchies have relatively less number of links and thus the complexity of network is reduced and the silicon costs is decreased.

As interconnection network can be mainly characterized by two factors: Number of links and

Corresponding Author: Baojun Qiao, College of Software, Henan University, Kai Feng, 475008, China Tel: 86-371-15136108958

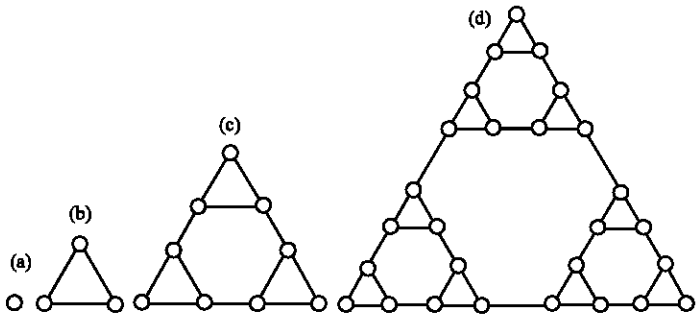

Fig. 1: Topology of THIN

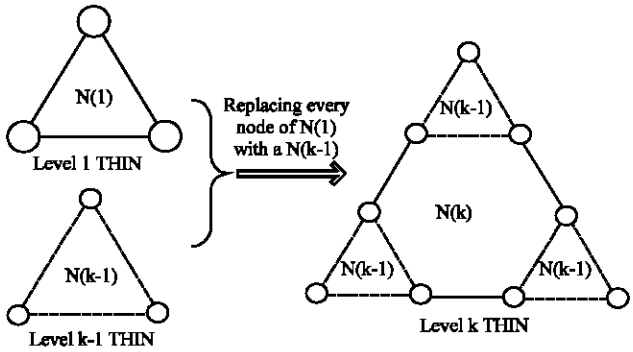

Fig. 2: Constructing process of level k THIN

diameter. This section addresses these principle properties of THIN and compares it with 2-D mesh. In the following, we first introduce the definitions and notations of these properties.

Definition 1: We use L_k to denote the number of links in a level k THIN.

According to the constructing process of THIN, the number of links in a level k THIN can be represented by Eq. 1:

$$\begin{cases} L_1 = 3 \\ L_k = 3L_{k-1} + 3 \end{cases} \quad (1)$$

From Eq. 1 we can know:

$$L_k = \frac{3\cdot(3^k - 1)}{2} = \frac{3\cdot(N-1)}{2} \quad (2)$$

where, $N = 3^k$, represents the number of nodes in level k THIN.

Definition 2: $P_{i,j}$ is used to represent the path from vertices i and j of a graph G, we call the distance between i and j the length of the shortest $P_{i,j}$ and denote it by $D_{i,j}$.

Definition 3: The diameter of a graph G, denoted by D_G, is the maximum of the distance $D_{i,j}$ over all pairs of vertices of G. By the definition:

$$D_G = \max(D_{i,j}) \quad (3)$$

Let $D_{THIN(k)}$ denotes the diameter of a level k THIN. Following the constructing process of THIN, $D_{THIN(k)}$ can be represented as Eq. 4:

$$\begin{cases} D_{THIN(1)} = 1 \\ D_{THIN(k)} = 2D_{THIN(k-1)} + 1 \end{cases} \quad (4)$$

From Eq. 4, we can know that:

$$D_{THIN(k)} = 2^k - 1 = 2^{\log_3^N} - 1 \quad (5)$$

Table 1 compares the network properties of THIN with 2-D mesh, where N denotes the number of nodes in network.

The number of links is used to represent the cost and complexity of a network. When the nodes of a network increase, the links should increase in the linear model in order to minimize the connect cost. The links of THIN are

Table 1: Comparison of topoloty properties of THIN and 2-D mesh

Type of network	Parameter	Value
THIN	$\frac{3\cdot(N-1)}{2}$	$2^{\log_3^N}-1$
2-D Mesh	$2\cdot(N-\sqrt{N})$	$2\cdot(\sqrt{N}-1)$

the fewest when they have the same network size. This is very important for constructing NOC, because the fewer the number of links is, the less the chip resource will be cost.

The diameter is one of important parameters for interconnection network and it impacts the communication delay between nodes. In a packet switching network, the diameter is always required to be as short as possible. The diameter of THIN is shorter than 2-D mesh when the network size is not very large. The comparison results show that when there are not too many cores, THIN is a better candidate for constructing NOC, taking into account the number of links and diameter.

LATENCY OF THIN

To evaluate an interconnection network, the network latency must be taken into account. In this section, we mainly research the zero-load latency of THIN and compare it and 2-D mesh.

In the study, Duato *et al.* (1997), a zero-load latency model is presented for wormhole switching networks. Suppose the message contains L-bit data. The phit size and flit size are assumed to be equivalent and equal to the physical data channel width of W bits. The routing header is assumed to be 1 flit; thus the message size is L+W bits. The latency to transfer the message in the network is:

$$T = D_{avg}\bullet(t_r + t_s + t_w) + \max(t_s, t_w)\bullet\left\lceil\frac{L}{W}\right\rceil \quad (6)$$

where, t_r is the time spent by the router to make a routing decision; t_s is the intra-router or switching delay and t_w is the inter-router delay (the propagation delay across the wires of an external channel). L/W is the packet payload and when addresses and data must be transmitted.

The first expression in Eq. 6 computes the latency to transfer the packet header, while the second one determines the time spent by the packet payload to reach the destination node following the header in a pipelined fashion. In this study, D_{avg} is taken as the average distance of the interconnection network.

Definition 4: The average distance of a interconnection network is the result of the sum of all the minimal path length between any two nodes in the network dividing by the total number of paths (Dong *et al.*, 1997) and we denote it by D_{avg}. D_{avg} can be calculated as the following equation:

$$D_{avg} = \sum_{i=0}^{D_G} i\times\rho(i) \quad (7)$$

where, D_G represents the diameter and $\rho(i)$ denotes the probability of the message which transmission distance is i over all the messages in the network.

The average distance of a level k THIN, denoted by D_{avg_THIN}, is given in:

$$D_{avg\_THIN} = \frac{1}{3^{k-1}} + \frac{16\times(6^{k-1}-1)}{5\times 3^k} - \frac{1}{3} \quad (8)$$

The study Dong *et al.* (1997) gives the equation to calculate the average distance of 2-D mesh:

$$D_{2-D\ mesh} = \frac{2(N-1)}{3\sqrt{N}} \quad (9)$$

Increasing network degree can reduce the average distance of an interconnection network. So, it is very difficult to accurately evaluate the latency of interconnection networks with different degree, if only using the average distance without taking into account the network degree. In this study, we use the normalized average distance (Peng, 2004) when analyzing the latency.

Definition 5: The normalized average distance of an interconnection network, denoted by μ, is the result of the average distance D_{avg} multiplied by the degree d:

$$\mu = d.D_{avg} \quad (10)$$

The normalized average distance of a level k THIN, denoted by μ_{THIN}, can be obtained by Eq. 11:

$$\begin{aligned}\mu_{THIN} &= 3\times\left(\frac{1}{3^{k-1}} + \frac{16\times(6^{k-1}-1)}{5\times 3^k} - \frac{1}{3}\right)\\ &= 3\times\left(\frac{3}{N} + \frac{16\times(6^{\log_3^N-1}-1)}{5\times N} - \frac{1}{3}\right)\end{aligned} \quad (11)$$

where, N represents the number of nodes in the THIN.

We use $\mu_{2\text{-}D\ Mesh}$ to denote the normalized average distance of a 2-D mesh and it can be given by:

$$\mu_{2-D\ mesh} = \frac{8\times(N-1)}{3\sqrt{N}} \quad (12)$$

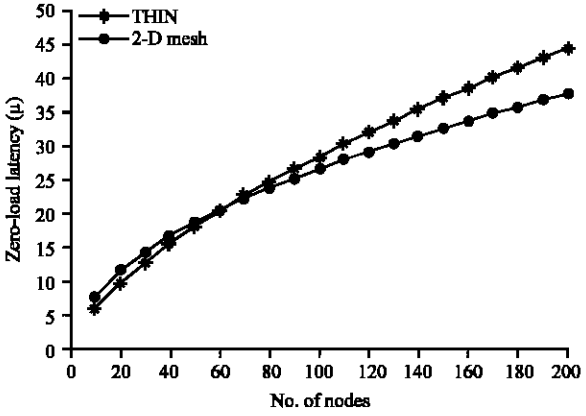

Fig. 3: Comparison of zero-load latency between THIN and 2-D mesh

In this study, when comparing the zero-load latency of different interconnection networks, we use the normalized average distance μ to take place the average distance D_{avg}. Based on Eq. 6 and 11, the zero-load latency of level k THIN, denoted by T_{THIN}, is given by:

$$T_{THIN} = 3\times\left(\frac{3}{N}+\frac{16\times(6^{\log_6^N-1}-1)}{5\times N}-\frac{1}{3}\right)\times (t_r + t_s + t_w) + \max(t_s, t_w)\left\lceil\frac{L}{W}\right\rceil \quad (13)$$

The zero-load latency of a 2-D mesh is given by:

$$T_{2-D\ mesh} = \frac{8\times(N-1)}{3\sqrt{N}}(t_r + t_s + t_w) + \max(t_s, t_w)\left\lceil\frac{L}{W}\right\rceil \quad (14)$$

Figure 3 compares the zero-load latency generated by THIN and 2-D mesh, respectively. The same routing decision, network switching and communication bandwidth are used by both interconnection networks. Suppose the routing decision time (t_r), switching delay (t_s) and channel delay (t_w) all are fixed constant and we use the normalized average distance to measure the zero-load latency. Figure 3 indicates that the zero-load latency of THIN is lower than 2-D mesh when the scale of network is not very large.

THIN AND 2-D MESH POWER MODEL AND COMPARISON

The power consumption of SoC especially from the interconnection network becomes more and more important for the whole power optimization. With the high speed and great amount of data exchange among the chip components, power problem must be solved, because consumers demand longer battery life in addition to lower cost in computers, battery-operated systems and many consumer products. In this section, we analyze and model the power consumption of THIN and compare it with 2-D mesh. The model computes power dissipation for a packet crossing the interconnect network in low traffic mode that there is no packet contention.

As defined in the study Ye *et al.* (2002), the average energy consumed for transferring a packet is dissipated on three components: (1) The internal node switches, located on the intermediate nodes between ingress and egress ports, (2) The internal buffers, used to temporarily store the packers and (3) The interconnect wires for packet transfer. E_s is defined as the average switching energy dissipated in a node for packet transfer, E_b is the buffering energy and the average wiring energy dissipated between two nodes is defined as E_w. We use a single parameter E_{sb} to denote both the buffering and switching energy dissipated in a node. We call E_{sb} router energy. Defining E_{packet} as the average power dissipated for packet transfer, we can have:

$$\begin{aligned} E_{sb} &= E_s + E_b \\ E_{packet} &= D_{avg}\bullet E_w + (D_{avg}+1)\bullet(E_s + E_b) \\ &= D_{avg}\bullet(E_w + E_{sb}) + E_{sb} \end{aligned} \quad (15)$$

where, D_{avg} represents the average distance of a interconnection network. In order to simplify the calculations, we define:

$$\beta = E_w/E_{sb} \quad (16)$$

The parameter β shows the relation of wiring and router energy dissipation for a packet transfer.

Taking into account the network degree, we use the normalized average distance described in section 3 when model the power consumption of interconnection network. Based on Eq. 11 and 15, the average energy dissipated for a packet transfer in THIN, denoted by E_{THIN}, is given by:

$$\begin{aligned} E_{THIN} &= \mu_{THIN}\bullet(E_w + E_{sb}) + E_{sb} \\ &= 3\times\left(\frac{3}{N}+\frac{16\times(6^{\log_6^N-1}-1)}{5\times N}-\frac{1}{3}\right)\bullet(E_w + E_{sb}) + E_{sb} \end{aligned} \quad (17)$$

The average power dissipated for a packet transfer in 2-D mesh is given by:

$$\begin{aligned} E_{2-D\ mesh} &= \mu_{2-D\ mesh}\bullet(E_w + E_{sb}) + E_{sb} \\ &= \frac{8\times(N-1)}{3\sqrt{N}}\bullet(E_w + E_{sb}) + E_{sb} \end{aligned} \quad (18)$$

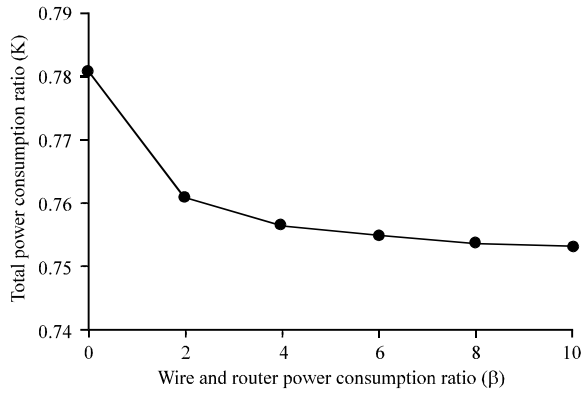

Fig. 4: Ratio of packet transfer power dissipation for a level 2 THIN and mesh (3×3)

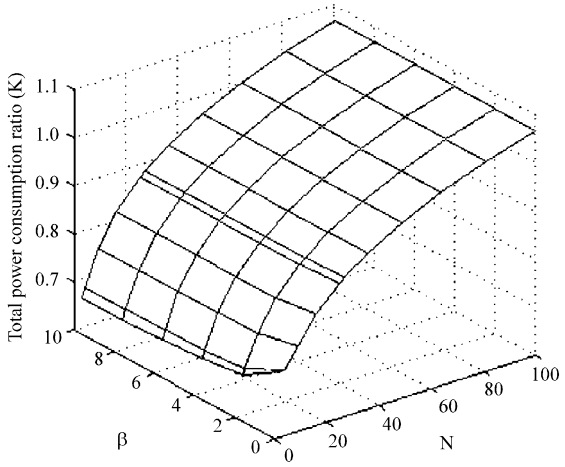

Fig. 5: Ratio of packet transfer power dissipation in THIN and 2-D mesh for different values of β and N

In this study, we use the method mentioned by Rahmati *et al.* (2006) to compare the energy consumption of THIN and 2-D mesh. Therefore, we define K as follows:

$$K = \frac{E_{THIN}}{E_{2-D\ mesh}} = \frac{(\beta+1)\bullet\mu_{THIN}+1}{(\beta+1)\bullet\mu_{2-D\ mesh}+1} \quad (19)$$

where, K shows the ration of the energy dissipated for 2-D mesh and THIN for a packer transfer. Figure 4 shows the ratio of packet transfer power dissipation for a level 2 THIN and 2-D mesh (3×3) as function of β. Depending on all values of β, the power consumption ratio may vary from 0.75-0.78. This means that the THIN consumes lower amount of power than its mesh counterpart.

Figure 5 shows the K as a function of β and N (number of nodes). As can be seen in the figure, with the network scale enlarged, THIN will consumes more power than 2-D mesh. Because with the number of nodes increasing, the normalized average distance of THIN will be longer than 2-D mesh.

CONCLUSION

In this study, we study the communication latency and power consumption of THIN and compare them with 2-D mesh. THIN is preferable to construct interconnection network for system-on-chip when the network size is not very large. Our future research should be on modeling the energy consumption for high traffic loads in THIN.

ACKNOWLEDGMENTS

The authors wish to acknowledge the support of the National Natural Science Foundation of China (No. 61100173, 51305021, U1334211, No. 11226173 and 61102163) and fund of ShaanXi Province Education Department (No. 2013JK1139) and Fundamental Research Funds for the Central Universities (2012JBZ014). This project is also supported by China Postdoctoral Science Foundation (No. 2013M542370) and supported by the Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant No. 20136118120010). This work is also supported by Scientific Research Program Funded by Xi'an University of Science and Technology (Program No. 201139). This job is also supported by Science and Technology Project of Xi'an (CXY1139 (7) and CX1262⑨).

REFERENCES

Benini, L. and G. De Micheli, 2002. Networks on chips: A new SoC paradigm. IEEE Comput., 35: 70-78.

Dong, Y.F., D.X. Wang and W.M. Zheng, 1997. Exact computation of the mean minimal path length of N-Mesh and N-Torus. Chinese J. Comput., 20: 376-380.

Duato, J., S. Yalamanchili and L.M. Ni, 1997. Interconnection Networks: An Engineering Approach. IEEE Computer Society Press, Los Alamitos, CA., USA., ISBN: 978-0818678004, Pages: 515.

Hu, Y.F., H.Y. Chen, Y. Zhu, A. Chien and C.K. Cheng, 2005. Physical synthesis of energy-efficient networks-on-chip through topology exploration and wire style optimization. Proceedings of the IEEE International Conference on Computer Design: VLSI in Computers and Processors, October 2-5, 2005, USA., pp: 111-118.

Kumar, S., A. Jantsch, M. Millberg, J. Oberg and J.P. Soininen *et al*., 2002. A network on chip architecture and design methodology. Proceedings of the IEEE Computer Society Annual Symposium on VLSI, April 26-26, 2002, Pittsburgh, PA., USA., pp: 105-112.

Peng, G., 2004. Research on interconnection networks of parallel signal processor. Ph.D. Thesis, Harbin Engineering University, Harbin, China.

Qiao, B.J., F. Shi and W.X. Ji, 2007. THIN: A new hierarchical interconnection network-on-chip for SOC. Proceedings of the 7th International Conference on Algorithms and Architectures for Parallel Processing, June 11-14, 2007, Hangzhou, China, pp: 446-457.

Rahmati, D., A.E. Kiasari, S. Hessabi and H. Sarbazi-Azad, 2006. A performance and power analysis of WK-recursive and mesh networks for network-on-chips. Proceedings of the International Conference on Computer Design, October 1-4, 2006, San Jose, CA., USA., pp: 142-147.

Shi, F., W.X. Ji and B.J. Qiao, 2006. A new non von Neumann architecture TriBA. Trans. Beijing Inst. Technol., 26: 847-849.

Ye, T.T., L. Benini and G. de Micheli, 2002. Analysis of power consumption on switch fabrics in network routers. Proceedings of the 39th Design Automation Conference, June 10-14, 2002, New Orleans, LA., USA., pp: 524-529.