Performance Comparison of 3D NoC Topologies using Network Calculus

Nallasamy Viswanathan [1], Kuppusamy Paramasivam [2], Kanagasabapathi Somasundaram [3]

[1.] Department of ECE, Mahendra Engineering College, Namakkal-637 503, INDIA.
e-mail: vissivn@gmail.com
[2.] Department of ECE, Bannari Amman Institute of Technology, Sathyamangalam-638 401, INDIA.
e-mail: kp_sivam@yahoo.com
[3.] Department of Mathematics,Amrita Vishwa Vidyapeetham, Coimbatore-641 112, INDIA.
e-mail: s_sundaram@cb.amrita.edu

Abstract: Nowadays, System-on-Chips (SoCs) designers are forced to integrate tens to hundreds of functional and storage blocks in a single die to implement emerging complex computation, multimedia and network services. The integration of huge degree of the blocks in a single die poses new challenge in designing the interconnect architecture of the blocks in SoCs. The traditional bus based interconnect infrastructure is ineffective as the number of the blocks increases more than ten. The packet based Network- on- Chip (NoC) is an obvious interconnect design alternative to the bus based on-chip communication architecture in SoCs. The advent of 3D NoC architecture attracts added interest as it offers improved performance and shorter global interconnect. Evolving an efficient 3D network topology and developing 3D routing scheme play a crucial role in determining the performance of 3D NoC interconnect architecture. In this paper, two 3D NoC topologies, namely 3D Recursive Network Topology (3D RNT) and 3D Modified Mesh Topology (3D MMT) are presented. End-to-end delay, switch buffer size and the influence of the buffer size in determining area overhead requirement of the two topologies are evaluated using an analytical model of network calculus and the evaluation results of the two topologies are compared. It is shown that the 3D RNT outperforms the 3D MMT even though 50% of the vertical links are trimmed down in the former topology. Further, 20 % reduction in average switch buffer size and 16% reduction in area overhead requirement are achieved in the 3D RNT. The results of the analysis are of use to evaluate and optimize 3D NoC interconnect architecture as far as the design space is concerned.
[Viswanathan. N, Paramasivam. K, Somasundaram. K. **Performance Comparison of 3D NoC Topologies using Network Calculus**. *Life Sci J* 2013;10(1):3379-3385] (ISSN: 1097-8135). http://www.lifesciencesite.com. 429

Keywords: 3D NoC topology; 3D routing; vertical links; network calculus; performance evaluation

1. Introduction

Recently, on-chip transistor density has been increased to many folds owing to the advances in the semiconductor technology and hence it is possible to integrate tens to hundreds of the functional and storage blocks together in a System-on-Chips (SoCs). Nowadays, a lot of commercial chips are available that contain a huge degree of the blocks to implement emerging complex computation, multimedia and network services. Optimization of the interconnect architecture in SoCs is a challenging task due to the escalating electromagnetic interference and crosstalk effects. The traditional bus based interconnect infrastructure is ineffective as the number of the blocks increases more than ten due to the poor scalability, parallelism integration, high latency and power dissipation [Owens, 2007].

The packet based Network- on- Chip (NoC) has been identified recently as an obvious interconnect design alternative to the bus based on-chip interconnect architecture in SoCs as NoC scales better than the bus based on-chip interconnections and has better performance and power consumption characteristics[Dally, 2001]. On the other hand, longer global interconnects and more area overhead requirement are becoming the major performance bottleneck in NoC as the number of the blocks over a 2D plane increases to meet the ever growing applications. The advent of three-dimensional (3D) stacked technologies provides a new perspective for on-chip interconnect design. 3D ICs have higher packaging density which reduces power consumption due to shorter global interconnect wires and circuitry is more immune to noise, have multiple layers where the active functional and storage blocks are stacked in each layer and the layers are vertically interconnected using through-silicon vias (TSVs) [Feero, 2009],[Topol, 2006], [Davis, 2005]. Even though both 3D integrated circuits and NoCs are proposed to reduce the global interconnect wires, area overhead requirement and improve the on chip communication performance, there are several challenges to design three-dimensional NoCs such as non availability of EDA tools, high peak temperatures, large area footprints of vertical interconnects and lack of optimized network topology, etc.

In this paper, we have presented two 3D NoC topologies, namely 3D Recursive Network Topology

(3D RNT) and 3D Modified Mesh Topology (3MMT) and 3D routing scheme to transverse flits among the blocks interconnected using the topologies. Traffic related parameters, i.e. end-to-end delay, switch buffer size and the influence of the switch buffer size in reducing the area overhead are computed using an analytical model of network calculus and the performance of the two topologies are compared.

In the rest of the section, related works are summarized in section 2, section 3 presents two 3D NoC topologies and 3D routing schemes, section 4 provides evaluation method, section 5 and 6 describes the evaluation results. Discussion and Conclusion are given in the last section.

2 Related Works

Many 3D NoC topologies are proposed in [Viswanathan, 2011], [Viswanathan, 2012], [Rohit, 2009]. Among the topologies, fully and partially connected regular 3D mesh topologies are focused in many current literatures [Feero, 2009]. The 3D RNT has been identified as an appropriate topology for 3D NoC as it offers a high degree of scalability, and well suited for modular approach and to implement heterogeneous blocks. However, for the best of our knowledge, finding a suitable topology for 3D NoC remains an open problem.

A near optimal oblivious routing algorithm for conventional 3D mesh networks is called as randomized partially minimal (RPM) routing. The RPM works in both vertical and horizontal links using either a Z-XY-Z or Z-YX-Z dimensional path [Ramanujam, 2008].

On-chip interconnect macro architectures proposed for SoCs are evaluated by various trade-offs with regard to end-to-end delay, buffer size, through put, energy consumption and silicon area requirements. Performance evaluation of NoC is generally performed using simulation which is commonly slow for larger systems and provides trivial insight on how the different traffic related parameters influence the actual NoC performance [Pande,2005], [Bakhouya, 2009].

Several works have illustrated that there is a crucial need for analytical model to evaluate the performance of NoC interconnect architectures as analytical models can allow a fast evaluation of the performance metrics of larger systems in early design process and give more insights on how the traffic related parameters affect the performance of NoC interconnect architecture. [Suboh, 2008], [Bakhouya, 2009 & 2010].

Network calculus is an analytical model that computes the input and output arrival curves, end-to-end delay bounds, buffer size of the switch and other performance metrics of 2D mesh on-chip interconnect design and the results of the model are compared against simulation results in the papers [Suboh, 2010], [Suboh, 2005], [Moadeli, 2007], [Bakhouya , 2009].

[Bakhouya, 2010] illustrated the practical approach of network calculus to evaluate the performance of 2D regular Mesh, Spidergon and WK recursive on-chip interconnect architectures under different traffic patterns. The results obtained from the analytical model showed many insights on the interconnect architectures performance and trade-off between the traffic related parameters.

[Suboh, 2010] presented a network calculus based model to evaluate the performance metrics like latency, throughput and communication load and the cost metrics, i.e. energy and area requirement of 2D WK recursive on-chip interconnect architecture. The results of the model are validated with the simulation and it is observed that the changes of the two methodologies are in the same order of the magnitude.

3 3D NoC Topologies and Routing Scheme

In this section, we have presented the 3D Recursive Network Topology (3D RNT) and 3D Modified Mesh Topology (3D MMT) as shown in Figure 1 and Figure 2. In each topology, only 3 layers are considered for experimental analysis at ease even though *n* layers can be accommodated in the topologies.

The 3D RNT is derived from WK recursive network topology of W (4, 2) where four additional flipping links are used in each layer except in the top and bottom layer of the topology and the topology has many favorable properties such as Hamiltonian connectedness, topology expandability without changing existing vertex degree, scalability, fault tolerance and symmetry [Suboh, 2008], [Jung-Sheng, 2005]. Only 25% of the blocks in each layer are interconnected with the neighboring layers using the vertical links.

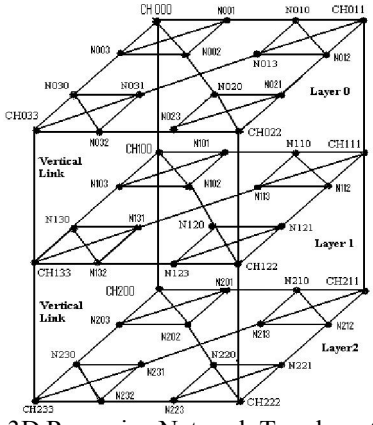

Figure 1. 3D Recursive Network Topology (3D RNT)

The 3D MMT is a 4×4×3 partially connected 3D mesh topology where 50% of the blocks in each layer are interconnected with their counterpart blocks of the neighboring layers. In the given two topologies, each node comprises of a switch and either a functional or storage block. And each four nodes are grouped together to create four clusters in each layer thus total number of nodes in a layer is sixteen.

In each cluster, a node is identified as cluster head (CH) and each CH and other nodes are identified with an ID of three digits xyz, first digit x represents a layer, second digit y represents a cluster and the digit z is either a CH or a node in the cluster y and layer z. The nodes are connected to the switches, in turn the switches are interconnected by using bidirectional horizontal and vertical interconnect links and the vertical links are utilized to establish inter layer communications [Pavlidis, 2007].

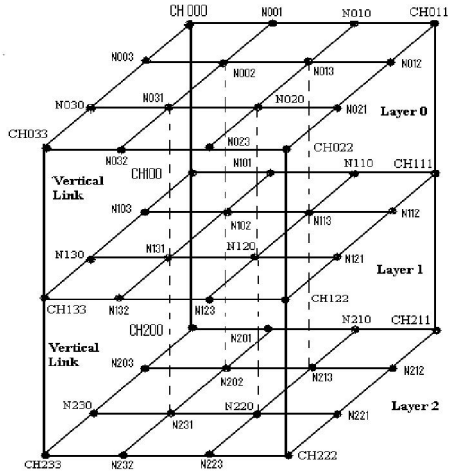

Figure 2. 3D Modified Mesh Topology (3D MMT)

In the routing, we have used Z-dimensional routing [Viswanathan, 2012] to route the flits through vertical links to reach destination node if source and destination nodes are located in different layers and X, Y- dimensional routing is used to find destination node in the destination layer. The 3D RNT routing scheme follows 3 hierarchical steps. Step 1: finding destination layer, step 2:finding destination cluster in the destination layer and step 3: finding destination node in the destination cluster. In the 3D MMT, routing scheme compares the current node ID (x_c , y_c , Z_c) to the destination node ID (X_d, Y_d, Z_d) stored in the header flit. Step 1:Flits are moved to the layer x_d where the destination node is located. Step 2: If $Y_c > Y_d$, the flits are routed to the west port of the switch and if $Y_c < Y_d$, the flits are routed to the east port and $Y_c = Y_d$, the flits are in the destination cluster Y_d. Step 3: If $Z_c > Z_d$, the flits are routed to the north port and if $Z_c < Z_d$, the flits are moved to the south port and $Z_c = Z_d$, the flits are in destination node.

4 Evaluation Method

Data flows f_i among the nodes are realized by the transfer of flits hop by hop from source node to destination. Each switch has an input buffer which has a FIFO queue management system [Coenen, 2006]. Network calculus is a mathematical model that allows us to specify the NoC traffic related parameters into the model.

Incoming traffic of a switch can be defined by the arrival curve α(t) of the incoming data flows. The incoming data flows consist of cumulative flits arriving from a block interfaced with the switch and neighboring switches. If a switch S_{xyz} has input data flows from n neighboring switches and its own block, then the arrival curve of the switch S_{xyz} can be defined at time t as $\alpha_{xyz}(t) = \alpha(t) + \sum_{j=1}^{n} \alpha_j(t)$, where α(t) is the arrival curve of the input data flow of its own IP block and $\sum_{j=1}^{n} \alpha_j(t)$ is the arrival curve of the input data flows arriving from n neighboring switches [Bakhouya, 2008]. Consider a straight line is $y\text{-}y_1 = m(x\text{-}x_1)$, where m is slope of the line and (x_1, y_1) is a point on the line. The maximum input data flows to a switch is constrained by the arrival curve $\alpha_{xyz}(t) = rx + b$, where b is the maximum burst size of the data, r is the average rate of the data flow or average service rate. In other words, the arrival curve of the switch S_{xyz} is [Suboh, 2010]

$$\alpha_{xyz}(t) = rt + b \qquad (1)$$

Similarly, the service curve of the switch is constrained by the equation β(t) = R (x - T),where R is the guaranteed service rate and T is the maximum latency caused by a switch. In other words, when t >T, the service curve of the switch is

$$\beta(t) = R(t\text{-}T) \qquad (2)$$

where T= k /R, k is the flit size in bytes. The output data flows of a switch is constrained by the following output curve

$$\alpha^*_{xyz}(t) = \alpha_{xyz}(t) + rT \qquad (3)$$

Input buffer size B of a switch is computed as follows. At time t, number of bytes available in the buffer is B= b + rt. If maximum number of the bytes residing in the buffer at t = T, then the maximum input buffer size B of a switch can be

$$B = b + rT \qquad (4)$$

Similarly, the delay bound can be calculated as follows. Let x_1=c, y_1=0, equation of the straight line becomes y/m = x- c, where y = b, m = R, x = t and c =T, then the delay bound D of a switch becomes

$$D= b / R + T \quad (5)$$

The data flows considered here as case study are generated by six different source-sink pairs that impersonate particular application data traffic. The date flows f_1, f_2, f_3, f_4, f_5 and f_6 generated by the source nodes whose IDs are 123,022,023,000,003 and 100 are consumed by the sinks whose IDs are 122,021,122,202,220 and 223. The sources are selected randomly and the sinks are selected according to the communication locality principle in which 33 % of the traffic takes place between neighboring nodes with distance d =1 and the rest of the traffic is uniformly distributed among the other blocks. Routing path for the six traffic flows is established according to the routing scheme given in the previous section.

In the 3D MMT, there are two input data flows f_2 and f_3 to the switch S_{122} and the following can be computed using the equations (1) and (3).

The arrival curve of the switch S_{122} is

$$\alpha_{122}(t) = 2rt+2b+5/2\ rT \quad (6)$$

From the equation (6), b = 2b+ (5/2) rT and r = 2r then the maximum input buffer size B_{122} of the switch S_{122} becomes

$$B_{122} = 2b+ (9/2)\ rT \quad (7)$$

Similarly, the delay bound D_{122} of the switch S_{122} becomes

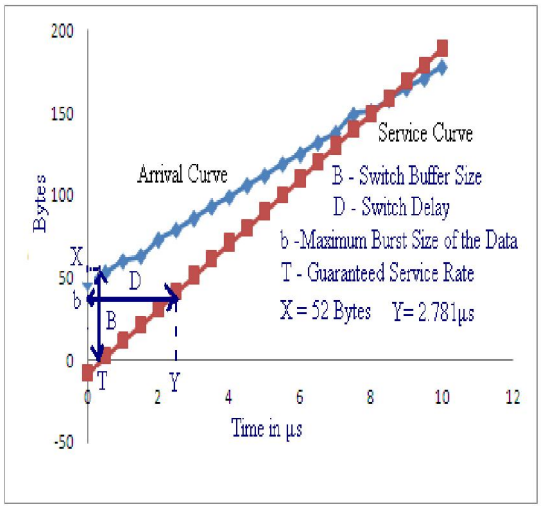

Figure 3. Arrival and service curve of the switch S_{122} in the 3D MMT

$$D_{122} = 2b/R+ (5r/2R +1)\ T \quad (8)$$

Let us consider a traffic pattern where source injection rate r = 50 Mbps, R = 150 Mbps, maximum burst size of the data b= 20 bytes and flit size k=8 bytes then the maximum latency caused by the switch is T= k/R = 0.407 µs. The arrival curve and service curve of the switch S_{122} is drawn as shown in Figure 3 and it is observed from the Figure 3 that the maximum input buffer size B_{122} (X) of the switch S_{122} is 52 bytes and the delay bound D_{122} (Y) of the switch is 2.781 µs.

5 Evaluated Result I - End-to-end latency

The following assumptions are made to compute the end-to-end latency

- A switch can function as either source or sink or intermediate node
- Sources injecting the data have the same injection rate and burst size in order to make analysis at ease
- Links used to interconnect the nodes have same length
- Switches have same speed in processing the flits.

The data flows arrived to the switch S_{122} is constrained by the arrival curve which is computed and given in the equation (6). Similarly, arrival curve of the other switches participating in the data flows are computed and given in Table 1.Having obtained the arrival curves, the data transfer delay bound D_{xyz} of each switch S_{xyz} is computed using the equation (5) and given in Table 2.

Table 1. Arrival curve of the switches participating in the data flows

3D MMT	3D RNT
$\alpha_{023}(t)= \alpha_{123}(t)= \alpha_{0003}(t)=b+rT$	$\alpha_{023}(t)=\alpha_{123}(t)=$
$\alpha_{022}(t)=2rt+2b+rT$	$\alpha_{0003}(t)=b+rT$
$\alpha_{021}(t)=rt+b+3/2rT$	$\alpha_{022}(t)=2rt+2b+rT$
$\alpha_{122}(t)=2rt+2b+5/2rT$	$\alpha_{021}(t)=rt+b+3/2rT$
$\alpha_{000}(t)=2rt+2b+rT$	$\alpha_{122}(t)=2rt+2b+5/2rT$
$\alpha_{100}(t)=3rt+3b+3rT$	$\alpha_{000}(t)=2rt+2b+rT$
$\alpha_{200}(t)=3rt+3b+6rT$	$\alpha_{100}(t)=3rt+3b+3rT$
$\alpha_{201}(t)=3rt+3b+9rT$	$\alpha_{200}(t)=3rt+3b+6rT$
$\alpha_{202}(t)=rt+b+4rT$	$\alpha_{202}(t)=3rt+3b+9rT$
$\alpha_{210}(t)=2rt+2b+8rT$	$\alpha_{220}(t)=2rt+2b+8rT$
$\alpha_{213}(t)=2rt+2b+10rT$	$\alpha_{223}(t)=rt+b+5rT$
$\alpha_{220}(t)=2rt+2b+12rT$	
$\alpha_{223}(t)=rt+b+7rT$	

The delay bound D_{xyz} of each switch participating in the data flows is computed at b=20 bytes, R=150 bytes, T=0.407µs and r is varied between 0 and 50Mbps. Having computed the delay bound of each switch, the end-to-end delay bound of each data flow is computed by summing up the individual delay bound of the switches participating in each data flow as shown in Figure 4. It is observed from the Figure 4 that the delay increases as the

injection rate is increased. When increasing the injection rate, more flits are waiting in the buffer before being serviced by a switch causes the increase in the delay bound.

Table 2. Delay bound of the switches participating in the data flows

3D MMT	3D RNT
$D_{023}=D_{123}=D_{0003}=b/R+T$	$D_{023}=D123=D\ 0003=\ b/R+T$
$D_{022}=2b/R+(r/R+1)T$	$D_{022}=\ 2b/R+(r/R+1)T$
$D_{021}=\ b/R+(3r/2R+1)T$	$D_{021}=b/R+(3r/2R+1)T$
$D_{122}=\ 2b/R+(5r/2R+1)T$	$D_{122}=2b/R+(5r/2R+1)T$
$D_{000}=\ 2b/R+(r/R+1)T$	$D_{000}=\ 2b/R+(r/R+1)T$
$D_{100}=\ 3b/R+(3r/R+1)T$	$D_{100}=\ 3b/R+(3r/R+1)T$
$D_{200}=\ 3b/R+(6r/R+1)T$	$D_{200}=\ 3b/R+(6r/R+1)T$
$D_{201}=\ 3b/R+(9r/R+1)T$	$D_{202}=\ 3b/R+(9r/R+1)T$
$D_{202}=b/R+(4r/R+1)T$	$D_{220}=\ 2b/R+(8r/R+1)T$
$D_{210}=\ 2b/R+(8r/R+1)T$	$D_{223}=\ b/R+(5r/R+1)T$
$D_{213}=\ 2b/R+(10r/R+1)T$	
$D_{220}=\ 2b/R+(12r/R+1)T$	
$D_{223}=\ b/R+(7r/R+1)T$	

On comparing the end-to-end latency of each data flow of the two topologies, it is obvious that the 3D RNT outperforms the 3D MMT as the number of hops between the source and sink node pairs is reduced duo to the interconnect design of the 3D RNT.

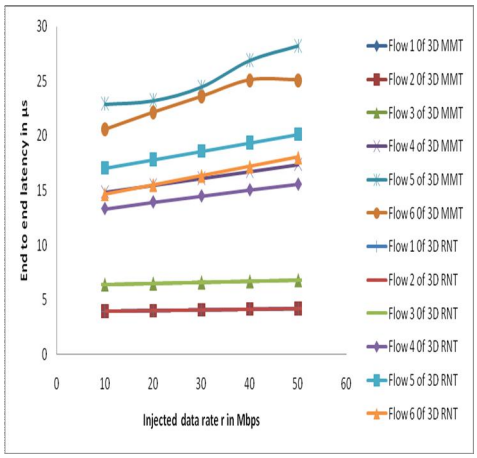

Figure 4. End-to-end delay of the data flows

6 Evaluation Result II - Switch buffer size and its influence in the area overhead requirement

Each switch has input buffer in which the incoming flits are temporarily stored before the flits being serviced by the switch. Dominant part of the area occupied in a switch is buffer and hence an accurate modeling of the switch input buffer size plays vital role in reducing the chip area [Suboh, 2010]. The input buffer size B_{122} of the switch S_{122} is computed by using the equation (7). Likewise, the input buffer size of the other switches participating in each data flow is computed using the equation (4) and given in Table 3. Average buffer size of the switches participating in each data flow is computed and shown in Figure 5 and 6.

Table 3. Buffer size of the switches participating in the data flows

3D MMT	3D RNT
$B_{023}=B_{123}=B_{003}=b+rT$	$B_{023}=B_{123}=B_{003}=b+rT$
$B_{022}=2b+2rT$	$B_{022}=2b+2rT$
$B_{021}=b+5/2rT$	$B_{021}=b+5/2rT$
$B_{122}=2b+9/2rT$	$B_{122}=2b+9/2rT$
$B_{000}=2b+3rT$	$B_{000}=2b+3rT$
$B_{100}=3b+6rT$	$B_{100}=3b+6rT$
$B_{200}=3b+9rT$	$B_{200}=3b+9rT$
$B_{201}=3b+12rT$	$B_{202}=3b+12rT$
$B_{202}=b+5rT$	$B_{220}=2b+10rT$
$B_{210}=2b+10rT$	$B_{223}=b+6rT$
$B_{213}=2b+12rT$	
$B_{220}=2b+14rT$	
$B_{223}=b+8rT$	

It is observed from the evaluation that the average buffer size of a switch is $(n \times 20)$ bytes where n is the number of data flows arriving to the switch and if $n=3$, the buffer size of the switch becomes round 60 bytes. On comparing the

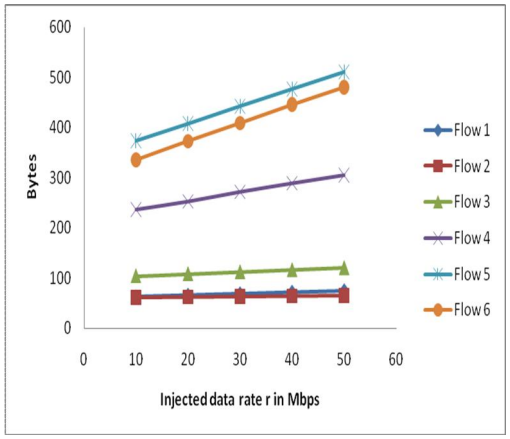

Figure 5. Average buffer size of the switches participating in each data flow of 3D MMT average buffer size of the switches of the 3D MMT with 3D RNT, 20% average buffer size reduction is achieved in the latter topology.

In 3D NoC, there are three major sources of area overhead which are switches, blocks and links. The average area overhead requirement can be computed by using the following equation [Suboh, 2010]

$$A=N_s+(R_s+a_s d_g S_f B_s)+N_b A_b+a_\ell N_\ell L_\ell \qquad (9)$$

where N_s is the number of switches, R_s is the switch silicon area required for routing table and logic to implement the routing algorithm, a_s is the area required for one byte, d_g is the average number of buffers inside the switch, S_f is the flits size in bytes, B_s is the average buffer size in bytes, N_b is number of

blocks, A_b is the area requirement for a block, a_ℓ is the area required for a link with link length L_ℓ and N_ℓ is the number of bidirectional links.

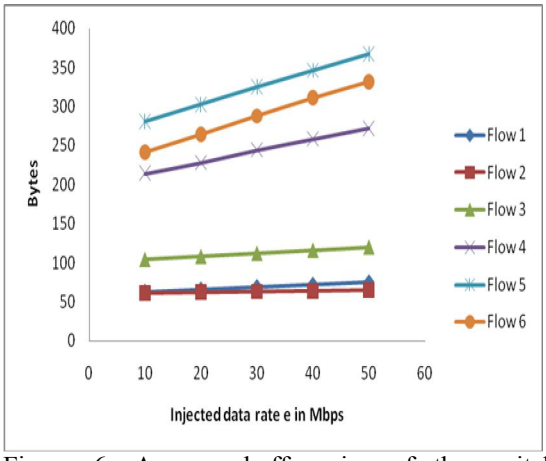

Figure 6. Average buffer size of the switches participating in each data flow of 3D RNT

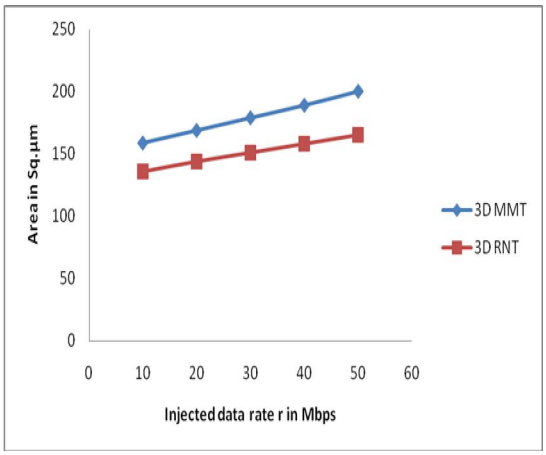

Figure 7. Average area overhead requirement

In the equation (9), the parameters other than the average buffer size B_s and N_ℓ are assumed as follows: N_s=16, R_s=1μm^2, a_s=0.001 μm^2, d_g=5, N_b=16, A_b =1 μm^2, a_ℓ = 0.01 μm^2, L_ℓ =5μm and S_f = 8 bytes. In the 3D MMT, N_ℓ = 24 and B_s=b+rT, equation (9) becomes

$$A= 33.20+0.64\ (b+rT) \quad (10)$$

In the case of 3D RNT, N_ℓ = 30 and B_s= b+rT, then the equation (9) becomes

$$A= 33.50+0.64\ (b+rT) \quad (11)$$

Average area overhead requirement is computed by substituting the average buffer size of the switches of each flow in the above equations (10) and (11). Figure 7 shows the average area overhead requirement of the two topologies and it is observed that 16% area overhead reduction is achieved in the 3D RNT when comparing with the 3D MMT.

Discussion and Conclusion

In this paper, an analytical model based on network calculus methodology was used to evaluate the performance and cost metrics i.e., end-to-end delay, buffer size and the influence of the buffer size in determining the area overhead requirement of the two 3D NoC topologies. On comparing the 3D RNT with3D MMT, the 3D RNT outperforms the latter topology with regard to end-to-end latency. Further, 20 % reduction in average buffer size and 16% reduction in area overhead requirement are achieved in the 3D RNT even though 50% of the vertical links are trimmed down in the 3D RNT as against the 3D MMT. This approach can provide the designers with intricate insights on the influence of the traffic related parameters in determining the performance and cost metrics of the on chip interconnect architecture. It is concluded that the network calculus based analysis is useful to evaluate and optimize 3D NoC interconnect architecture as far as the design space is concerned.

Corresponding Author
N. Viswanathan,
Faculty Member,
Department of ECE,
Mahendra Engineering College,
Mallasamudram – 637 503
Namakkal District, Tamil Nadu, India.
E-mail: vissivn@gmail.com

References

1. Owens.J.D, Dally.W.J, Ho. R, Jayasimha.D.N, Keckler. S.W, Peh. L.S. Research Challenges for On-Chip Interconnection Networks. IEEE Micro, 2007;27(5):96-108.
2. Dally .W. J, Towles .B. Route Packets, Not Wires: On-Chip Interconnection Networks. In Proc. of Design Automation Conference. 2001; 684–689.
3. Feero. B. S., Pande. P. P. Networks-on-Chip in a Three-Dimensional Environment: A Performance Evaluation. IEEE Transactions on Computers. 2009; 32-45.
4. Topol. A.W. et al. Three-Dimensional Integrated Circuits. IBM J. Research and Development. 2006; 50(4/5): 491.
5. Davis. W. R. et al. Demystifying 3D ICS: The Pros and Cons of Going Vertical. IEEE Design and Test of Computers. 2005;22(6): 498-510.
6. Ramanujam R. S., Lin. B. Near-optimal oblivious routing on three dimensional mesh networks. In Proc. IEEE Int. Conf. Comp. Design. Lake Tahoe, CA.2008.

7. Viswanathan.N,Paramasivam.K,Somasundaram.K. Performance Analysis of Cluster based 3D Routing Algorithms for NoC. In IEEE Proc. of Recent Advances in Intelligent Computational Systems –RAICS. 2011;157-162.
8. Viswanathan.N,Paramasivam.K,Somasundaram.K. Exploring Optimal Topology and Routing Algorithm for 3D Network on Chip. American Journal of Applied Sciences.2012; 9(3): 300-308.
9. Rohit Sunkam Ramanujam and Bill Lin . A layer-multiplexed 3D On- Chip Network Architecture. IEEE Embedded Systems Letters.2009;1(2).
10. Pande. P. P., Grecu. C., Jones. M., Ivanov. A., Saleh .R. Performance evaluation and design tradeoffs for network-on chip interconnect architectures. IEEE Trans. on Computer. 2005; 54(8): 1025–1040.
11. Suboh. S., Bakhouya. M., Lopez-Buedo. S., El-Ghazawi T. Simulation-based approach for evaluating on-chip interconnect architectures. SPL Proc., 2008; 75–80.
12. Bakhouya. M., Suboh. S., Gaber. J, El-Ghazawi. T. Analytical modeling and evaluation of on-chip interconnects using network calculus. NoCS Proc., 2009; 74–79.
13. Bakhouya. M., Suboh. S., Gaber. J, El-Ghazawi. T. Analytical performance comparison of 2D Mesh, WK-recursive, and Spidergon NoCs. In Proc. of IEEE International Symposium on Parallel & Distributed Processing, Workshops and Phd Forum (IPDPSW, 2010; 1-6.
14. Suboh Suboh, Mohamed Bakhouya, Jaafar Gaber, Tarek El-Ghazawi. Analytical modeling and evaluation of network-on-chip architectures. In Proc. of the 2010 International Conference on High Performance Computing and Simulation. HPCS. 2010;615-622.
15. Suboh. S., Bakhouya. M., El-Ghazawi. T. Simulation and evaluation of on-chip interconnect architectures: 2D Mesh, Spidergon, and WK-Recursive Networks. NoCS Proc., 2005–2006.
16. Moadeli. M., Shahrabi. A., Vanderbauwhede. W., Ould-Khaoua. M. An analytical performance model for the spidergon noc. 21st AINA Proc.2007; 1014–1021.
17. Suboh. S., Bakhouya.M, Gaber. J, El-Ghazawi. T. An interconnection architecture for network-on-chip systems. Telecommunication Systems. 2008;37:137–144.
18. Jung-Sheng Fu. Hamiltonicity of the WK-Recursive Network with and without Faulty Nodes. IEEE Transactions on Parallel and Distributed Systems. 2005; 16(9):853-865.
19. Pavlidis.V. F., Friedman. E. G. 3-D topologies for networks- on chip. IEEE Transactions on Very Large Scale Integration (VLSI) Systems. 2007;15(10):1081-1090.
20. Coenen. M, Murali. S, Ruadulescu. A., Goossens. K, Micheli. G. D. A buffer-sizing algorithm for networks on chip using TDMA and credit-based end-to-end flow control. CODES+ISSS Proc., 2006.

3/5/2013