International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 11 (2015) pp. 28155-28160

http://www.ripublication.com

Congestion Based Delay Model For Asynchronous NOC

[1.]Mrs. K. Ilamathi, [2.]Dr.P.Rangarajan
[1.]ilamathivelu@gmail.com,Assistant Professor,R.M.D Engineering College,India
[2.]Professor,R.M.D Engineering College,India

Abstract

The prediction of delay is of importance while considering the performance of NOC. Hence, an analytical performance model for asynchronous NOC is proposed based on queuing theory taking into account congestion in the network and is defined for heterogeneous topology. The mathematical model derived is based on queuing theory, taking congestion in links into account. The model is verified for synthetic traffic using a NOC simulator .It is found that the percentage of error between analytical value and simulation is 32.36 % on an average.

Keywords: Network on Chip (NOC), multiprocessor system-on-chips (MPSoCs)

Introduction

The computing in embedded systems have a continuously increasing computation and reliability, which is seen in multi-media and mobile communication. This is being fulfilled by multiprocessor system-on-chips (MPSoCs) soon scale to many-core SoCs with thousands of processors on a single chip[1].

With large number of cores, the interconnection becomes a serious challenge which is solved by network on chip (NoC).It has evolved as a flexible and high-performance solution for the interconnection problem during the last decade [2].

Reducing packet delay in NoC interconnect for many-core SoCs is a very challenging task, which decides the choice of routing and switching methods, selecting topology, application mapping, etc.

Therefore, fast and accurate delay model of NoC will be required that give an insight into the system and enable us to reduce the design space already in early design stages.

Cycle-accurate simulation based approaches are too slow for this purpose.

Simple high-level system models (e.g. only considering the propagation latency and ignoring queuing delays), on the other hand, are able to provide results in very short time, however these models loose some accuracy. Hence , Analytical models are used.

In this paper, we propose an analytic NoC model based on queueing theory [3]

The remainder of this paper is structured as follows. In Section 2,related work is discussed. Section 3 shows the Proposed Queuing Theory Model for delay of router, while the list of symbols are dealt on section 3.1.And Section 4 deals with performance evaluation. Finally section 5 concludes the work.

Related Work

Much research has gone into finding traffic models for the analysis of off-chip and (later) on-chip networks. The development of analytic tools for investigating latency and throughput in networks was done by Dally[4].A M/G/1 queueing model for wormhole switched two dimensional (2D) torus NoC topologies, assuming deterministic routing was presented in [5]. A fast and flexible analytic approach was proposed in [6] for the mean value performance analysis of virtual channel first-come first-serve (FCFS) input buffered routers .

Proposed Queuing Theory Model For Delay of Router

A generic router model for NoC performance analysis along with flow –control feedback probability has been considered in [7].However, the model does not discuss the link congestion and its delay. In our model we consider the link congestion delay along with feedback delay.

An insight into the effects of link congestion on delay experienced by the packets of wireless sensor networks is dealt in [8].

A queuing-theory-based analytical model for 2D mesh networks, for determining latency and power at the granularity of individual router sub-modules is discussed in [9].

However, it does not consider delay outside router and restricted to homogeneous topology. Our model is different from [7] and [9] , in that we consider delay in link transmission time due to congestion in links and defined for heterogeneous topology.

Average waiting time and propagation time of a packet transmitted from router (i, j) to router $(i + 1, k)$, is given by

$$TP_{i,j,k} = E(T)N_j + \frac{1}{2}\lambda_{i,j}E(T^2) + E(T)\sum_{q=1,q\neq j}^{w}\sum_{k=1}^{w} f_{i,j,k}f_{i,q,k}\left(1 - P_{i+1,k}\right)^2 N_q + \left(\frac{1}{1-PB}\right) * \frac{1}{\mu C_i} \quad (1)$$

$$NR = \frac{PB}{1-PB} \quad (2)$$

$$NT = \frac{PB}{1-PB} + 1 = \frac{1}{1-PB} \quad (3)$$

$NR -$ Total number of retransmissions

$NT - \text{Total number of transmissions}$

$$NR = PB + PB^2 + PB^3 + \cdots = PB \sum_{i=0}^{\infty} PB^i = \frac{PB}{1-PB} \quad (4)$$

$$c_{i,j,q} = \sum_{k=1}^{w} f_{i,j,k} f_{i,q,k} \left(1 - P_{i+1,k}\right)^2 \quad (5)$$

$$f_{i,j,k} = \frac{\lambda_{i,j,k}}{\sum_{l=1}^{w} \lambda_{i,l,k}} \quad (6)$$

$(i,j), (i+1,k), (i,q)$

$$\frac{dp^K(t)}{dt} = \lambda(t) p^{k-1}(t) - \mu C p^K(t) \quad (7)$$

$$PB = p^K(t) \quad (8)$$

The first term of equation (1) gives the service time of packets waiting in the same queue. The second term is the residual service time seen by an incoming packet. The third term is the blocking delay caused by competition and flow control and the fourth term is the congestion delay of the link through which the packet traverses so as to reach the next router.

Using Little's Theorem

The average number of customers in a stable system $N_{j,k}$, is obtained using Little's theorem

$$N_{j,k} = \lambda_{i,j} TP_{i,j,k} = \lambda_{i,j} E(T) N_j + \frac{1}{2} \lambda_{i,j}{}^2 E(T^2) + \lambda_{i,j} E(T) \sum_{q=1,q \neq j}^{w} \sum_{k=1}^{w} f_{i,j,k} f_{i,q,k} \left(1 - P_{i+1,k}\right)^2 N_q + \lambda_{i,j} \left(\frac{1}{1-PB}\right) * \frac{1}{\mu C_i} \quad (9)$$

List of Symbols

Symbol	Explanation
$E(T)$	mean service time
N_q	average number of packets waiting in (i,q)
$c_{i,j,q}$	Competition probability of the header flits in (i,j) and (i,q) transmitting to the same input port of router (i+1).
$\lambda_{i,j,k}$	traffic rate from (i,j) to $(i+1,k)$
$f_{i,j,k}$	Competition probability of header flit
(i,j)	Port(j) of router (i),where $0 \leq j < w$
$TP_{i,j,k}$	Average waiting time and propagation time of a packet transmitted from router (i,j) to router $(i+1,k)$
μC_i	Mean service rate of the link
C_i	Link capacity

$N_{j,k}$	The average number of customers in a stable system
L	Average packet latency
$X_{s,d}$	Traffic rate from s to d

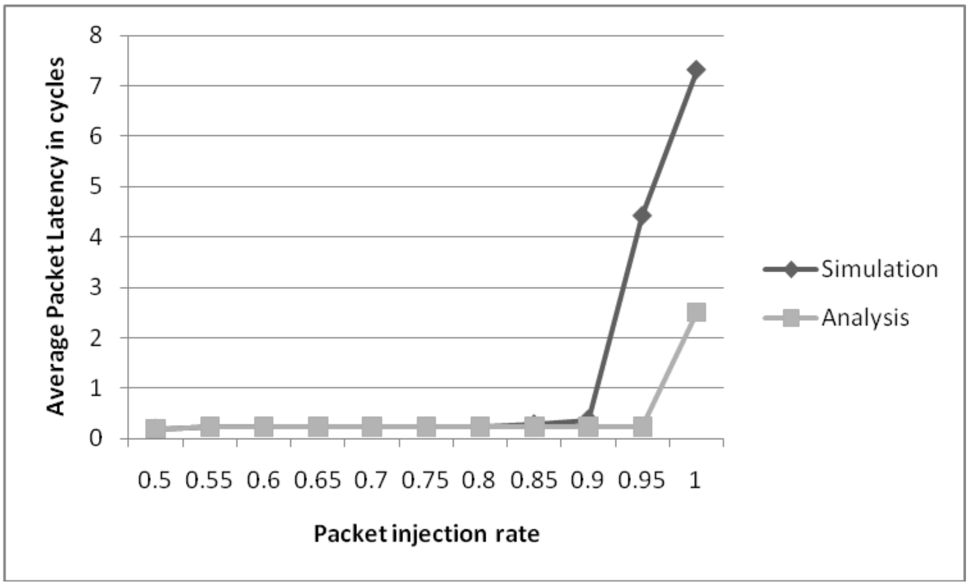
Buffer utilization ratio is the ratio between N and B (the queue capacity) , where N is the average number of packets waiting in the queue.

The average packet latency,

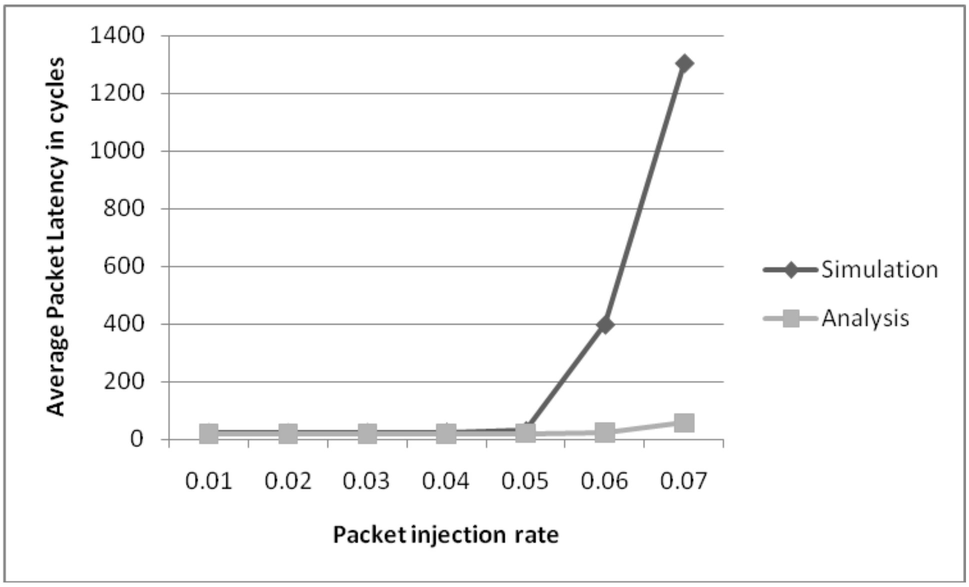
$$L = \frac{1}{\sum_{\forall s,d} x_{s,d}} \sum_{\forall s,d} \sum_{(i,j)\in\pi_{s,d}} x_{s,d}(TP_{i,j,k} + E(T)) \quad (10)$$

Performance Evaluation

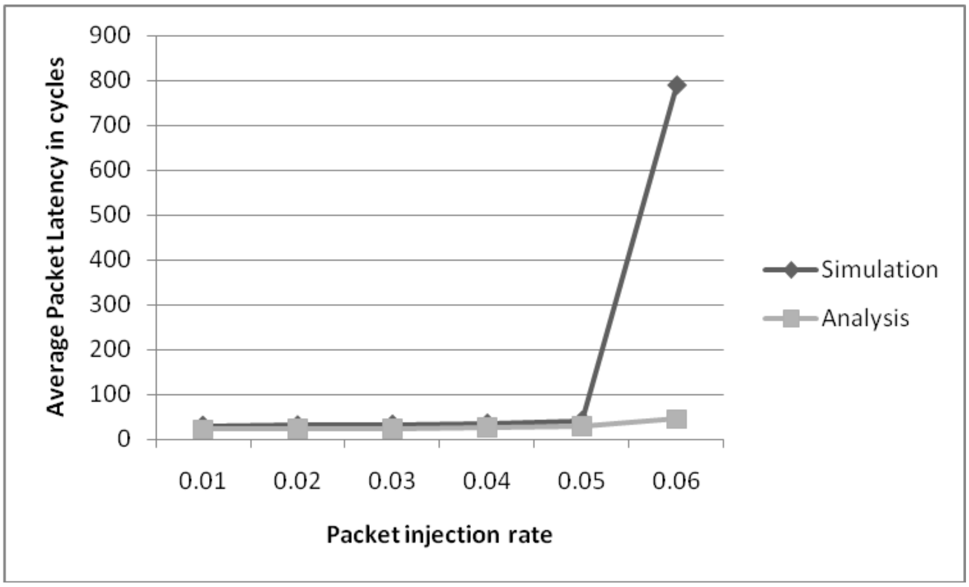
We show the accuracy of the proposed NoC model by comparing it against cycle-accurate NoC simulation using HNOC[10].It is OMNeT++ based a modular open-source simulator is used. It supports heterogeneous NoCs with variable link capacities and number of VCs per each unidirectional port. HNOCS allows researching and exploring new paradigms and phenomena of heterogeneous NoCs. The simulations are carried out for Uniform , Transpose and Tornado traffics as shown in Fig.1. In this experiment, we adopt the XY deterministic routing and a 5x5 2D-mesh network. The percentage of error between analytical value and simulation is 15.89% for *Uniform Traffic*, whereas it is 43.09% for *Transpose* and 38.11% for *Tornado traffic*. The percentage of error is more in *Transpose* and *Tornado* due to nature of the traffic.



(a)



(b)



(c)

Figure 1: The latency curve for different packet injection rates for .(a) Uniform Traffic.(b)Transpose Traffic.(c)Tornado Traffic.

Conclusion

A analytical model for determining the delay in asynchronous NOC was proposed based on queuing theory ,taking congestion in the network into account. It was found that on average the percentage of error between analytical values and simulation results is 32.36%.

References

[1] S. Borkar, “Thousand core chipsa technology perspective,” in Proc. of DAC, 2007.

[2] L. Benini and G. De Micheli, “Networks on chips: a new soc paradigm,” Computer, vol. 35, no. 1, pp. 70 –78, Jan 2002.

[3] L. Kleinrock, Queueing systems - 1 : Theory. New York: Wiley, 1975.

[4] W. Dally, “Performance analysis of k-ary n-cube interconnection networks,” computers, IEEE Transactions on, vol. 39,no. 6, pp. 775 –785, Jun 1990.

[5] A. Kiasari, D. Rahmati, H. Sarbazi-Azad, and S. Hessabi, “A markovian performance model for networks-on-chip,” in Proc. of Euromicro PDP, 2008.

[6] U. Ogras, P. Bogdan, and R. Marculescu, “An analytical approach for network-on-chip performance analysis,” Computer Aided Design of Integrated Circuits and Systems, IEEE Trans actions on, vol. 29, no. 12, pp. 2001 –2013, Dec. 2010.

[7] Youhui Zhang et-al, “Model of network -on-chip routers and performance analysis”,IEICE electronic express,Vol.8,No.13July-2011

[8] David Tipper et-al, “An Analysis of the Congestion Effects of Link Failures in Wide Area Networks”.IEEE journal on selected areas in communications,vol.12,No.1,Jan 1994

[9] J. Kim, D. Park, C. A. Nicopoulos, N. Vijaykrishnan, and C. R. Das, “Design and Analysis of an NoC Architecture from Performance, Reliability and Energy Perspective”, in Proc. the 2005 Symposium on Architectures for Networking and Communications Systems (ANCS 2005), pp. 173-182, October 2005, Princeton, NJ.

[10] Y. Ben-Itzhak, I. Cidon, and A. Kolodny, “Delay analysis of wormhole based heterogeneous NoC,” in Proceedings of the fifth ACM/IEEE international symposium on Networks-on-Chip (NOCS 2011), 2011.