

MODELING ROUTER HOTSPOTS ON NETWORK-ON-CHIP

Siti Aisah Mat Junos@Yunus¹, Muhammad Nadzir Marsono²,
Izzeldin Ibrahim²

¹Faculty of Electronics and Computer Engineering,
Universiti Teknikal Malaysia Melaka (UTeM), Malaysia.

²Faculty of Electrical Engineering,
Universiti Teknologi Malaysia (UTM), Johor, Malaysia

aisah@utem.edu.my, nadzir@fke.utm.my, izzeldin@fke.utm.my

Abstract

A Network-on-Chip (NoC) is a new paradigm in complex System-on-Chip (SoC) designs that provides efficient on-chip communication architecture. It offers scalable communication to SoC and allows decoupling of communication and computation. In NoC, design space exploration is critical due to trade-offs among latency, area, and power consumption. Hence, analytical modeling is an important step for early NoC design. This paper presents a novel top-down approach router model, and utilizes this model for analysis mesh NoC performance measured in terms of throughput, average of queue size, efficiency, and loss and wait time. As case study, the proposed model is used to map a MPEG4 video core to a 4x4 mesh NoC with deterministic routing to measure the overall NoC quality of service. The model is used also to present how much occupancy of average queue size for each router that reduces resources (hardware) area and cost. The accuracy of this approach and its practical use is illustrated through extensive simulation results.

Keywords: Markov chain, Network-on-Chip, Queue, Router, System-on-Chip.

I. INTRODUCTION

Network-on-chip (NoC) has been proposed [1] to replace system bus as the primary on-chip communication method. Due to the separation of computation and communication, NoC could be designed separately from the computational entities (termed intellectual properties, IPs) [2]. Hence, analysis and optimization of NoC performance in terms of delay, latency, and loss are required. Quality-of-Service (QoS)

for NoC defines the level of commitment for packet delivery among IPs. Such a commitment can be the correctness and completion of the transaction, and bounds on the performance [3].

Researchers addressed the NoC router modeling from different perspectives [9]–[12]. The work in [12] proposed a delay model for a variable pipelined wormhole router with fixed time cycle to address Lopez's model problem. However, these models cannot be applied for router designs that use both clock edges and furthermore they did not study the impact of changing router design parameters on its delay. The router queue modeling also still need to be addressed because it is very important to get an estimation of the optimum queue size that matches the target traffic characteristics at higher levels of abstraction.

Different NoC-based SoC applications require different NoC design aspects in terms of the topology, router type, queue size, and switching technique. The challenges for NoC research have thus been on assisting early design space exploration for NoC-based SoC. Good traffic and NoC performance estimation could significantly shed some light to probable NoC design aspects. This work looks at router model that estimates an NoC router performance. This paper presents a performance analysis of mesh NoC in terms of throughput, average queue size, efficiency, loss, and waits time. A discrete time Markov model of

mesh NoC topology is obtained, which helps decision-making in terms of switching techniques, buffer sizes, and router types. This also helps to identify router and link hotspots for better packet routing.

II. RELATED WORKS

System bus is a circuit-switching, connection-oriented on chip communication backbone. In contrast, NoC uses packet-switching, which segments the message into a sequence of packets [4] sent to a shared network. On-chip networks share the same characteristics in topology, switching, routing, and flow control with local area network [13]. Furthermore, NoC has to provide high and predictable performance [5] with small area overhead and low power consumption.

An important problem in NoC is the router design, because it significantly affects the network performance and power consumption. An efficient router design is determined by its switching technique (packet switching, wormhole, etc.), flow control type (hand shaking, credit based, etc.), queue size, arbiter design (round robin, rate-proportional servers, etc.), routing scheme (adaptive, deterministic). Researchers tried to address the design and modeling of NoC routers from different perspectives. In this subsection, we highlight the work done up to date in this area through representative research work.

Researchers addressed the NoC router modeling from different perspectives [9]–[12]. Chien et al. [14] proposed a delay model for wormhole and virtual channel routers. However, this model was designed for 0.8-micron CMOS and also cannot be applied to pipelined architectures. Lopez et al. [16] proposed an extension to Chiens model for pipelined routers. But Lopez's model assumes that the time duration of the clock cycle depends on the router latency,

which is not a practical assumption. Peh et al. [12] proposed a delay model for a variable pipelined wormhole router with fixed time cycle to address Lopez's model problem. However, these models cannot be applied for router designs that use both clock edges and furthermore they did not study the impact of changing router design parameters on its delay.

The router queue modeling still needs to be addressed because it is very important to get an estimation of the optimum queue size that matches the target traffic characteristics at higher levels of abstraction. Due to limited buffers [6] and link bandwidth, packets may be blocked due to contention [7]. Buffer sizing has a direct association with bounds in bandwidth, delay and jitter [8].

III. NOC STRUCTURE

This chapter provides background on NoC architectural issues. These issues include network topology, router structure, switching techniques and routing algorithms. The trade-off analysis in NoC modeling can be performed by optimally considering NoC architecture parameters based on a specified application.

A. Topology

This architecture is based on an $m \times n$ mesh network where every router, except those at the edges, is connected to four neighbouring routers and one computation resource, (IP) through communication channels [15]. This topology allows integration of large number of IP cores in a regular shape structure. A channel consists of two unidirectional links between two routers or between a router and a resource. Fig. 1 shows a 3×3 mesh NoC with nine functional IP blocks.

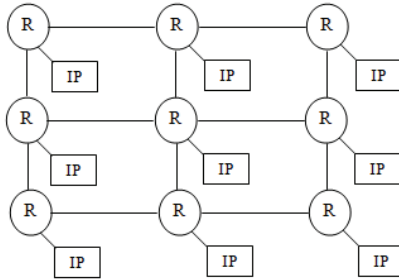


Fig. 1. Mesh Topology

The 2D-torus architecture is basically similar as a regular mesh except that routers at the edges are connected to the routers at the opposite edge through wrap-around channels [17]. Every router has five ports, one connected to the local resource and the others connected to the closest neighboring routers. The long end-around connections can yield excessive delays. An octagon NoC consisting of 8 nodes and 12 bidirectional links. In an octagon topology, exchange message between any pair of nodes takes at most two hops. To design a system consisting of more than eight nodes, the octagon can be extended to multidimensional space, however with a significantly increased wiring complexity [16].

B. Router Structure

Router operates at the network layers similar to the one for computer network. A router uses packet headers and a forwarding table to determine the best way a packet should go between the networks. An NoC router has three main architectural components, input/output ports, queues, and switch fabric (SF). The switch fabric establishes the required paths between pairs of input and output ports according to a certain routing mechanism such as round-robin scheduler, weighted round-robin scheduler, and max-min fairness scheduling [18].

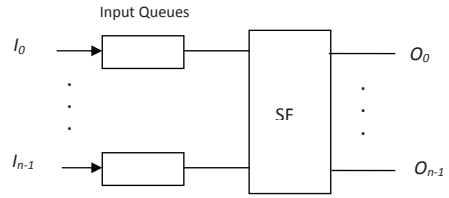


Fig. 2. Input-queuing router

Fig. 2 shows an input-queuing router. Each input port has a dedicated first-in first-out (FIFO) queue for storing incoming packets. In one time step, an input queue must be able to support one write and one read operations. Assuming an $n \times n$ router, the switch fabric must connect n input ports to n output ports [18]. The main advantage of an input queuing router is the low memory speed requirement, distributed traffic management at each input port, and also distributed table lookup at each input port. It supports packets broadcast and multicast without the need to duplicate the packet. The main disadvantage is the head of line (HOL) problem when the packet at the head of the queue is blocked from accessing the desired output port [18]. There are three potential causes for packet loss, fully populated input queue, internal blocking due to blocked the switch fabric, and when switch fabric is busy serving another packet.

C. Switching Techniques

As an alternative to circuit switching, a message can be partitioned and transmitted as fixed-length packets by packet switching. Packets are individually routed from source to destination. A packet is stored at each intermediate node then forwarded to the next node. Packet switching is good for short and frequent messages [19]. However, unlike in circuit switching where a physical path is reserved for the whole message, each packet of a message has to be routed at each intermediate node. Moreover, splitting a message into packets also increases overhead.

Traditional designs borrowed from local area networks (LANs) result in limiting performance bottleneck. Some new switching techniques, such as virtual cut-through (VCT) and wormhole switching techniques have been proposed to improve NoC performance [19]. In packet switching, a packet must be received, in whole, at an intermediate node before a routing and forwarding decision to the destination. However, the header of a packet usually arrives to an intermediate node earlier than the tail of a packet by several clocks. To construct small router that resides in an on-chip component, wormhole switching is usually used [19]. This work assumes wormhole switching because this switching requires less queue capacity and allows low-latency communication.

D. Routing

Routing algorithms are used to specify the path from source to destination for each message. They can be implemented in two ways which are either deterministic or adaptive [19].

Deterministic routing protocol chooses the path for a message only by its source and destination. All packets with the same source and destination pair will follow one single path. The packet will be delayed if any channel along this path is loaded with heavy traffic, and if a channel along this path is faulty, the packet cannot be delivered. Thus, the deterministic routing protocols suffer from poor use of bandwidth, and blocking even when alternative paths are available.

A common deterministic routing algorithm is dimension order routing, in which the packet is routed in one dimension at a time, arriving at the proper coordinate in each dimension before proceeding to the next dimension. Deterministic routing has been widely used in multi-computers due to its simplicity for router implementation [19]. It is because in the deterministic routing, messages with the same source and destination always traverse the same path.

Adaptive routing protocols are proposed to make more efficient use of bandwidth and to improve fault tolerance of interconnection network. In order to achieve this, adaptive routing protocols provide alternative paths for communicating nodes. Thus, it could overcome the congested areas in the network. Several adaptive routing algorithms have been proposed, showing that message blocking can be considerably reduced, thus strongly improving throughput [16].

IV. MARKOV CHAIN APPROACH FOR NOC MODELING

There are several approaches to modeling NoC. Several works [1]–[4] focus on stochastic models. This project could be conceptualized from top-level system design. It starts with the highest level of NoC view, and works its way down to every single component in NoC block diagram.

A. Modeling Abstractions

An NoC-based SoC system is composed from an NoC topology and IP blocks. The NoC provides decoupling computation (IP) and communication parts. This allows for IPs and interconnects to be designed independently. At a level below is the router abstraction. Routers are pivotal modules in NoC based design.

- 1) *NoC-level Abstraction*: Fig. 3 shows a SoC system is composed from NoC and IP blocks. The NoC provides decoupling computation (IP) and communication parts. This allows for IPs and interconnects to be designed independently.
- 2) *Topology-level Abstraction*: Fig. 1 shows the top level view of a 3 mesh topology for NoC modeling. Two elements on NoC are router and network interface (NI). The NI is used as interfaces between IP blocks and NoC. Function of the router is to transport data from one

network interface to another. This work considered analysis on NoC router only.

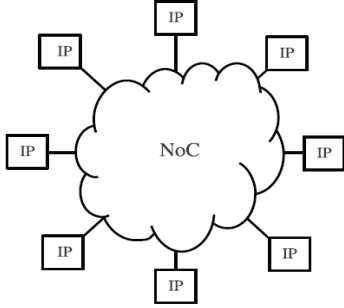


Fig. 3. NoC-level Model Abstraction

- 3) *Router-level Abstraction:* Mesh topology is used with each router has the maximum 5 input-output ports. Four ports are connected with others routers and one port to the IP. Fig. 2 shows an input-queuing router internal structure. Each input port has a first-in first-out (FIFO) queue for storing incoming packets. In mesh topology, the top queue is fed by the link connected to the IP associated with that router and the other four bottom queues are fed by the inter-router links.

B. Performance Metrics

The NoC performance is analyzed in term of throughput, average queue size, and packet delay.

- 1) Throughput in units of packets per time step which demonstrate how many end-to-end packet/flit transfer.
- 2) Latency in terms of time step, where time step is define as the time to transfer a packet on a local link (between two routers or between an NI and a router).
- 3) Average lost traffic is measured in units of packets per time step.

- 4) Queue Occupancy Queue size is measured in units of packets.

C. Queue Modeling

This section presents an analytical model for input-queuing router. Each queue is considered as a first-in-first-out (FIFO) queue. The model has simple close-form calculations and produces the performance of the queue.

A simple M/M/1/B queue is used in this model. This model provides a discrete-time Markov chain [18] analysis of queue where the time step is taken equal to the time required to transmit a packet. Poisson distribution traffic arrival process and the exponential distributed service time is assumed. For each queue model, one server queue with B finite buffer size are assumed.

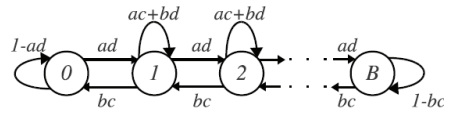


Fig. 4. State transition diagram for an M/M/1/B queue

Fig. 4 shows, the state transition diagram for the discrete time Markov chain M/M/1/B queue. A homogeneous Markov Chain is considered since packet arrivals and departures are independent of the time index value [18]. With an assumption that a packet cannot arrive and be served in the same time slot. Each state represents the number of e-mails in the queue. In Fig. 4, $b = (1 - a)$ and $d = (1 - c)$. From state transition diagram in Fig. 4, the state transition matrix P [18] is defined as

$$\mathbf{P} = \begin{bmatrix} 1 - a & bc & \dots & 0 & 0 \\ a & f & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & f & bc \\ 0 & 0 & \dots & ad & 1 - bc \end{bmatrix} \quad (1)$$

The difference equations for the state probability vector [18] can be expressed as

$$s = [s_0 \ s_1 \ \dots \ s_{B-1} \ s_B]t \quad (2)$$

where s_i is the probability that the queue contain i packets, and is given as

$$s_i = \rho^i d^{i-1} s_0 \quad \text{for } 1 \leq i \leq B \quad (3)$$

with ρ is defined as

$$\rho = \frac{a}{bc} \quad (4)$$

Satisfying the condition $\sum_{i=0}^B s_i = 1$ from [18] gives:

$$s_i = \frac{(1-\rho d)\rho^i d^{\max(0,i-1)}}{1+\rho(c-\rho^B d^B)} \quad \text{for } 0 \leq i \leq B \quad (5)$$

From Little's result [18], the average queuing delay Φ is given by:

$$\phi = \frac{Q}{T} \quad (6)$$

where Q is the average queue occupancy and T is the average queue throughput. The throughput, T , is defined as the probability that the queue is served while it is not empty, and is given by:

$$Th = c(1-s_0) \quad (7)$$

The average queue occupancy Q is given by:

$$Q = \sum_{i=0}^B i s_i = \frac{1-\rho d}{1+\rho(c-\rho^B d^B)} \sum_{i=1}^B i \rho^i d^{i-1} \quad (8)$$

A packet is said to be lost when the queue is full when the a packet arrives and none (in the queue) is serviced. Thus, the loss probability, L , is given by:

$$L = s_B a d \quad (9)$$

The queue size, B , gives a trade-off between queuing delay and loss probability. Bigger B increases the average queuing delay Φ due to higher Q and at the same time decreases the loss probability, and vise versa.

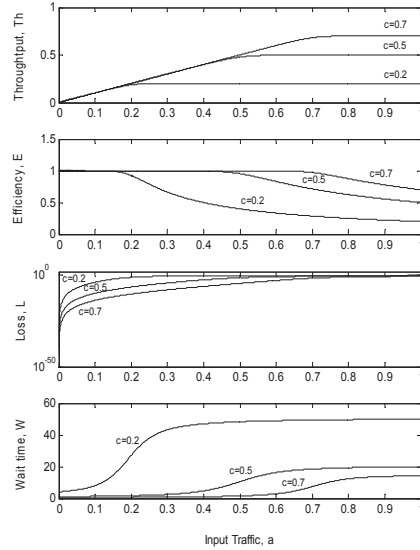


Fig. 5. Performance of an M/M/1/B queue. Throughput, efficiency, loss and wait time versus input traffic.

D. Router and NoC Modeling

This section discusses the placement of a router in mesh NoC can generate different value of output traffic c at each that router output port. It is also shows generation of general router performance. Router loading performance for each router is identified that determinates possible router hotspots in the NoC topology.

Table I. General router and NoC performance equations. q_j refers to j -Th queue in router r_i .

Variables	Router r_i	NoC
Throughput Th	$\frac{1}{ N } \sum_{j=1}^N Th_{q_j}$	$\frac{1}{ r_{total} } \sum_{i=1}^N Th_{r_i}$
Queue occupancy Qa	$\frac{1}{ N } \sum_{j=1}^N Qa_{q_j}$	$\frac{1}{ r_{total} } \sum_{i=1}^N Qa_{r_i}$
Loss probability L	$\frac{1}{ N } \sum_{j=1}^N L_{q_j}$	$\frac{1}{ r_{total} } \sum_{i=1}^N L_{r_i}$
Waiting time W	$\frac{1}{ N } \sum_{j=1}^N W_{q_j}$	$\frac{1}{ r_{total} } \sum_{i=1}^N W_{r_i}$

The output traffic of a router is important value that significantly impact router performance in the network. Congestion condition occurs as soon as input traffic exceeds the maximum output traffic c . All routers have different number of input/output port, n , and their placement in the NoC topology. Therefore, probability c for each router may be different. Assume that packet arriving at a certain input switching fabric router is destined to other $n - 1$ router output with equal probability c that depends on n . The general equation of output traffic c is:

$$c = \frac{\sum_{m=0}^{n-1} \frac{n!}{n-m}}{n \times n!} \quad \text{where } \frac{1}{n} \leq c \leq 1 \quad (10)$$

General router equation performance is generated from average of all queues in a router. There are the equations of throughput, efficiency, average queue size, loss, and wait time for a router ri . The overall performance in term of throughput, efficiency, average queue size, wait time, and loss are obtained by averaging all 16 routers in a 4×4 mesh NoC. The general NoC performance is given in Table I.

V. ANALYSIS OF MPEG-4 CORES

Fig. 6 shows a proposed methodology in a case study for video application (MPEG4 core) to analyse performance of mesh NoC in term of throughput, efficiency, average queue size, loss, and latency.

A. MPEG4 SoC Traffic Distribution

As shown in Fig. 6, a typical traffic distribution graph (TDG) for the video applications (MPEG4 core) discussed in [20] is considered the main design input. The numbers written on the arrows are the average number of packets transmitted and the numbers written on the circles represent the IPs number.

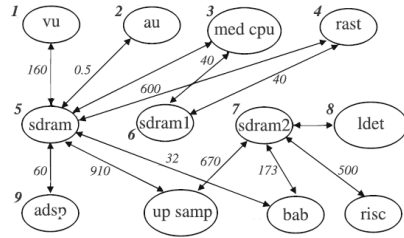


Fig. 6. MPEG4 core [20]

We assume that loads between two IPs are equal for the incoming and outgoing packets. The generating traffic distribution matrix (λ) from a given TDG, which represents the initial mapping of the MPEG4 core. The generated matrix is organized such that λ_{ij} represents the number of packets transmitted from a node IP_j to IP_i .

$$\lambda = \begin{pmatrix} 0 & 0 & 0 & 0 & 190 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 60 & 40 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 600 & 40 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 190 & 0.5 & 60 & 600 & 0 & 0 & 0 & 0 & 0 & 0.5 & 910 & 32 & 0 & 0 & 0 \\ 0 & 0 & 40 & 40 & 190 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 250 & 0 & 670 & 173 & 500 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 250 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 910 & 0 & 670 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 32 & 0 & 173 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 500 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (11)$$

B. Mapping Routes to Connectivity Matrix

IPs mapping and routing for the MPEG4 cores through the routers in 4×4 Mesh NoC with deterministic routing is shown in Fig. 7. All communication between IP blocks with same source and destination always go through same path of router through shortest path. Connectivity matrix is formed through IPs routing. A packet with same source and destination go through the specific routing path had been determined. Total no of hops used is 35 as shown in Table II.

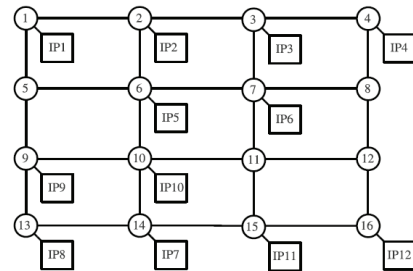


Fig. 7. MPEG4 core IPs is mapped in 4×4 mesh NoC.

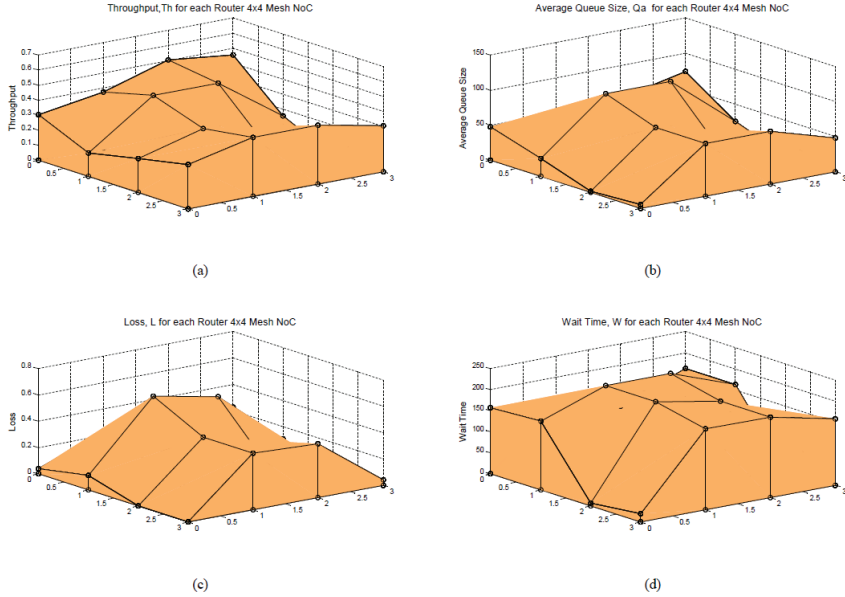


Fig. 8. The average performance for all 16 routers in the 4x4 mesh NoC (a) Throughput, (b) Average queue size, (c) Loss, and (d) Wait time.

Table II. MPEG4 connectivity matrix.

Communicating IPs	Routing path	No of hops
$IP_1 \leftrightarrow IP_5$	$r_1 \leftrightarrow r_5 \leftrightarrow r_6$	3
$IP_2 \leftrightarrow IP_6$	$r_2 \leftrightarrow r_6$	2
$IP_3 \leftrightarrow IP_5$	$r_3 \leftrightarrow r_7 \leftrightarrow r_6$	3
$IP_3 \leftrightarrow IP_6$	$r_3 \leftrightarrow r_7$	2
$IP_4 \leftrightarrow IP_5$	$r_4 \leftrightarrow r_3 \leftrightarrow r_2 \leftrightarrow r_6$	4
$IP_4 \leftrightarrow IP_6$	$r_4 \leftrightarrow r_8 \leftrightarrow r_7$	3
$IP_5 \leftrightarrow IP_6$	$r_6 \leftrightarrow r_{10} \leftrightarrow r_9$	3
$IP_5 \leftrightarrow IP_{10}$	$r_6 \leftrightarrow r_{10}$	2
$IP_5 \leftrightarrow IP_{11}$	$r_6 \leftrightarrow r_7 \leftrightarrow r_{11} \leftrightarrow r_{15}$	4
$IP_7 \leftrightarrow IP_8$	$r_{14} \leftrightarrow r_{13}$	2
$IP_7 \leftrightarrow IP_{10}$	$r_{14} \leftrightarrow r_{10}$	2
$IP_7 \leftrightarrow IP_{11}$	$r_{14} \leftrightarrow r_{15}$	2
$IP_7 \leftrightarrow IP_{12}$	$r_{14} \leftrightarrow r_{15} \leftrightarrow r_{16}$	3

C. Routing Matrix

From the connectivity matrix, the input port is identified for communicating with output ports of each router. Then, remove those set of input to output connections that are not used for routing paths. It is reduce resources area in the NoC. Equation (12) shows the example of routing table and routing matrix for router 10 in the 4×4 mesh NoC.

$$rm = \begin{vmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{vmatrix} \quad (12)$$

D. Overall NoC Performance and Router Hotspots

To verify performance analysis, overall mesh NoC performance in term of throughput, average queue size, efficiency, loss, and waiting time are performed. Initial mapping of 12-IPs MPEG4 core in 4×4 mesh NoC modeling is made through the shortest path deterministic routing given in Table I and Fig. 7. Simulation was performed and the router loading performance analysis for all 16 routers is shown in Fig. 8.

Fig. 8 shows that all 16 routers give different performance in mesh NoC due to the router loading and IPs traffic. Router six (R6) is determined as a router hotspot that gives the worst performance, only 39% efficiency, 100% wait time and 61% loss. R6 is connected to IP5 has a highest traffic rate. Therefore, all five queues of R6 have been mostly occupied with packets. The efficiency of centre routers are R7, R10, and

also R6 less than 50%, while 100% occupancy of queue size, almost 100%

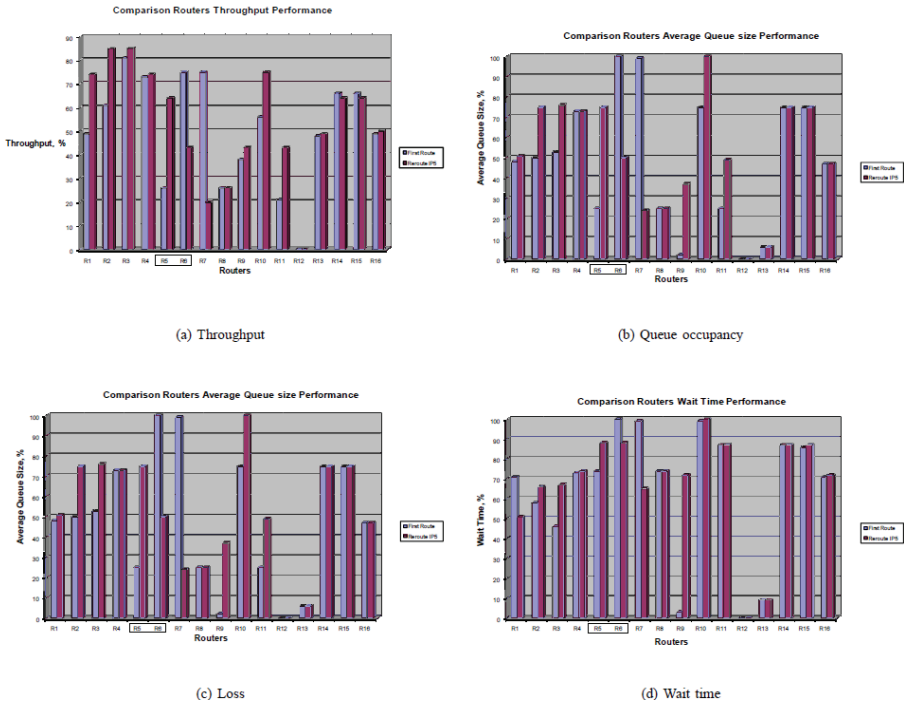


Fig. 9. Comparison efficiency between first routes and reroute IP5 in terms of (a) Throughput, (b) Average queue size, (c) Loss, and (d) Wait time.

wait time, and more than 50% loss. The R9 that is connected to IP9 gives a best performance with closely 100% efficiency, no loss, and only 8% wait time. Congestion could be characterized by decreased efficiency, increased loss, and increased wait time. Throughput only depends on the probability output traffic c , for each router. The average of throughput for this mesh NoC is 54%.

E. Hotspots Analysis

The main reason to identify router hotspots in an NoC is to improve the overall performance NoC. Thus, IP of router hotspot is rerouted possibly until the best performance could be achieved. From this router loading analysis, an early idea of mapping IPs that have different traffic rate could be set up for the each router. The occupancy of average queue size also could be obtained with specific routes. Therefore, the resource hardware area and cost could be reduced in the NoC topology.

Fig. 9 shows the percentage of loss between two routes among all 16 routers in mesh NoC. It can be seen that R6 is known as a router hotspot was decreased 34% of loss after rerouting IP5. In contrast, for the loss percentage of R5 increased 42% due to congestion occurred in all queue ports of R5. The only router of R7 offered the most decreasing of loss after rerouting IP5. It was happened because the only north, east, and IP6 queue ports have been used to traverse the path. In general, apparently only 3% differences of performance between two routes due to the number of hops is used. The first route used 35 hops and second route used 38 hops. It is shows that the first route provided a better overall performance than second route. Table III also reveals that the average queue size for both routes was reduced around 50% NoC area complexity.

Table III. Overall comparison before and after reroute.

Variables	First Route	After re-route
Throughput Th (packet/timestep)	54	57
Queue occupancy Qa	51%	55%
Loss probability L	23%	25%
Waiting time W (timestep)	69	72

VI. CONCLUSION

This paper presented a Markov chain model for identifying router loading and hotspots. An analytical model for 4-4 mesh NoC with a video application MPEG4 cores is presented. NoC performance metrics such as throughput, waiting time, queue size, efficiency, and loss could be easily identified from the model output.

As with most other research works, the models described in this work cannot yet claim that they are finished. There are many interesting possibilities for future research here and the most important of these are, extending this modeling approach with other NoC topologies, router types, and queue models. Another challenging problem is to improve the model to automatically place each IP to the most optimum NoC tile. Finally is to prototype the NoC.

ACKNOWLEDGMENT

The authors would like to thank those who contribute directly and indirectly towards completion of this project.

REFERENCES

[1] T. Ahonen, S. Virtanen, J. Kylliainen, D. Truscan, T. Kasanko, D. Siguenza-Tortosa, T. Ristimäki, J. Paakkulainen, T. Nurmi, I. Saastamoinen, H. Isannainen, J. Lilius, J. Nurmi, and J. Isoaho "A brunch from the coffee table - case study in NoC platform design," in Interconnect-Centric Design for Advanced SoC and NoC, J. Nurmi, H. Tenhunen, J. Isoaho, and A. Jantsch, Eds. Kluwer Academic Publishers, 2004, pp.425-453.

[2] W. J. Bainbridge and S. B. Furber, "CHAIN: A delay insensitive chip area interconnect," IEEE Micro special Issue on Design and Test of System on Chip, vol. 142, No.4., pp. 16-23, Sep. 2002.

[3] L. T. Smit, G. J. M. Smit, P. J. M. Havinga, J. A. Huisken, K. G. W. Goossens, and J. T. M. H. Dielissen, "Towards a model for making a trade-off between QoS and costs," in Proceedings of the CTIT workshop. Mobile Communications in perspective, pp. 105-109, Feb. 2001.

[4] M. Liu, "Improving the performance of a wormhole router and wormhole flow control," Master's thesis, School for Information and Communication Technology, Royal Institute of Technology, Stockholm, Sweden, Dec.2005. [Online]. Available: <http://www.imit.kth.se/axel/papers/2005/MScming-liu.pdf>

[5] L. T. Smit, G. J. M. Smit, P. J. M. Havinga, J. A. Huisken, K. G. W. Goossens, and J. T. M. H. Dielissen, "Towards A model for making A trade-off between QoS and costs," in Proceedings of the CTIT workshop. Mobile Communications in perspective, Feb. 2001.

[6] J. Hu and R. Marculescu, "Application-specific buffer space allocation for networks-on-chip router design," in Proceedings of the IEEE/ACM International conference on Computer-aided design, San Jose, CA, Nov. 6-10, 2004, pp. 354-361.

[7] P. Avasare, V. Nolle, J.-Y. Mignolet, D. Verkest, and H. Corporaal, "Centralized end-to-end flow control in a best-effort network-on-chip," in Proceedings of the 5th ACM international conference on Embedded software (EMSOFT '05). New York, NY, USA: ACM Press, 2005, pp.17-20.

[8] P. Vellanki, N. Banerjee, and K. Chatha, "Quality-of-Service and Error Control Techniques for Network-on-Chip Architectures," in Proceedings of the Great Lakes Symposium on VLSI, 2004.

[9] J. Chan and S. Parameswaran, "NoCGEN: A template based reuse methodology for networks on chip architecture," in Proceedings of 17th International Conference on VLSI Design, Mumbai, India, Jan. 5-9, 2004,

- pp. 717-720.
- [10] K. Goossens, J. Dielissen, and A. Radulescu, "Æthereal network on chip: concepts, architectures, and implementations," *IEEE Design and Test of Computers*, vol. 22, no. 5, pp. 21-31, Sept. 2005.
- [11] D. Ching, P. Schaumont, and I. Verbauwhede, "Integrated modeling and generation of a reconfigurable network-on-chip," in *18th International Parallel and Distributed Processing Symposium*, Santa Fe, NM, Apr. 26-30, 2004, pp. 139-146.
- [12] L.-S. Peh and W. J. Dally, "A delay model for router microarchitectures," *IEEE Micro*, vol. 21, no. 1, pp. 26-34, Jan. 2001.
- [13] D. S. Tortosa and J. Nurmi, Topology design for global link optimization for application specific network-on-chip, in *Proc. International Symposium on System-on-Chip SoC2004*, Tampere, Finland, 2004, pp. 135-138.
- [14] A. Chien, A cost and speed model for k-ary n-cube wormhole routers, *IEEE Transactions on Parallel and Distributed Systems*, vol. 9, no. 2, pp. 293-306, Feb. 1998.
- [15] S.Kumar, A.Jantsch, J.-P.Soininen, M.Forsell, M.Millberg, J.Oberg, K.Tiensuu and A.Hemani. A Network-on-Chip Architecture and Design Methodology. In *proceedings of the IEEE Computer Society Annual Symposium on VLSI*. 2002.
- [16] E. Baydal, P. Lopez and J. Duato. Increasing the Adaptivity of Routing Algorithms for k-ary n-cubes. In *Proc. 10th Euromicro Workshop on Distributed and Network-based Processing*. pp. 455-462. Jan. 2002.
- [17] W.J. Dally and B. Towles. Route Packets, Not Wires: On-Chip Interconnection Networks. *Proc. Design Automation Conf. (DAC)*. pp. 683-689. 2001.
- [18] Fayeze Gebali. *Computer Communication Networks Analysis and Design*. Springer. 2008.
- [19] J.Duato, S. Yalmanchili and L. Ni. *Interconnection Networks*. IEEE Computer Society. 1997.
- [20] D.Bertozzi and A.Jalabert. NoC Synthesis Flow for Customized Domain Specific Multiprocessor System-on-Chip. *IEEE Transaction on Parallel and Distributed System*. vol. 16, no. 2, pp. 113-129. Feb. 2005.

