Dynamic Migratory Selection Strategy for Adaptive Routing In Mesh NoCs

John Jose
Indian Institute of Technology
Guwahati, India
johnjose@iitg.ernet.in

Joe Augustine
Rajagiri School of Engineering
and Technology, Kochi, India
joe00thayil@gmail.com

Sijin Sebastian
Rajagiri School of Engineering
and Technology, Kochi, India
sijinsebastian@gmail.com

***Abstract*—NoC architectures are the most commonly used communication framework for multicore processors. Few factors that affect the performance of an on-chip network include the efficiency of the routing algorithm and the effectiveness of the output selection strategy used. All popular selection strategies use a static technique that behaves uniformly across various traffic patterns. This paper proposes a cost effective adaptive model of Regional Congestion Awareness (RCA) selection strategy. The traffic analyser incorporated in the proposed model learns the flit-flow pattern at each router and makes RCA behaves like the local best selection strategy under local traffic. Under non-local traffic, the normal RCA selection strategy works as it is. This switching (migration) between two selection strategies is done by proper controlling of aggregation and propagation mechanisms of RCA. As only in-router information is used for this switching, the design has no additional communication overhead. This dynamic switching decreases average packet latency and effectively optimises the network resources depending on traffic pattern, thereby, reducing power consumption. Our experiments on 8×8 mesh NoC with various synthetic and real traffic patterns show promising improvements compared to the existing baseline adaptive selection strategies.**

Keywords*—*flit-flow pattern, switching, traffic analysis.

I. INTRODUCTION

Traditionally, System on Chip (SoC) has been designed with dedicated point-to-point bus based interconnections. For large designs, this has several limitations from a physical design view point. Buses are not only non-scalable, but also consume significant area and power on the chip. Network on Chip (NoC) replaces traditional bus based communication system with an on-chip packet-switched interconnection network. NoC is the most popular communication framework used in modern multicore designs [11].

Figure 1 shows a 9-core multiprocessor SoC that uses a two dimensional 3x3 mesh NoC topology for on-chip communication. From the figure, we can see that each core is connected to a router. Routers are interconnected using bi-directional links. NoC based system uses packet based switching for inter-core communication. Packets are divided into flow control units called flits. Buffers in the routers and handshaking signals between routers enable flow control and smooth movement of packets from the source router to the destination router.

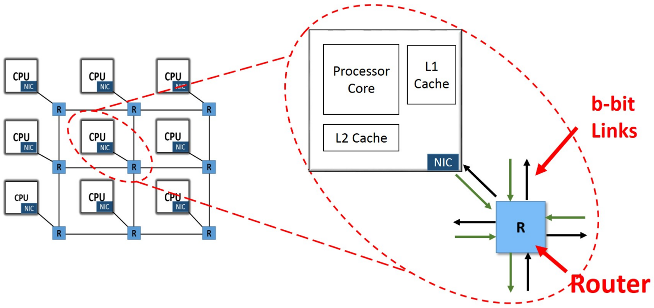

Fig. 1. Core-router interaction in a 3x3 mesh topology.

Each core consists of an out-of-order superscalar processor, a private L1 cache and a shared distributed L2 cache. Whenever there is a cache miss, a miss request is generated from the source core in the form of a packet. Packet is forwarded through various intermediate routers to reach the destination core. Routing algorithm employed in the router determines the next outgoing link for every incoming packet. Similarly the reply packet also traverses through the on-chip network. Typically the cache miss request is a single flit packet whereas the reply packet can have multiple flits based on cache block size and inter-router link width.

The miss penalty of a cache miss depends on the round-trip latency of the cache miss request and the cache miss reply packet. Since the latency of cache miss request packet is a very critical performance parameter, the underlying NoC which carries the cache miss request and reply packet has to deliver them in the minimum time. In this context, the performance of an NoC framework is highly critical in determining the throughput of applications running on the cores.

The performance of an NoC depends mainly on the routing algorithm and selection strategy used in the routers [2]. The routing algorithm implemented on the router computes the output link for every incoming packet. If the routing algorithm returns more than one output link, a selection strategy is used to choose the most suitable link. By developing cost effective output channel selection techniques, the average packet latency and power consumption can be reduced.

The basic and the most simple selection strategy employed is the local best selection strategy [11]. Here, each router receives the number of free buffers in each of its downstream neighbor through a dedicated 4-bit wide control channel. It then chooses the router with higher number of free buffers. This can take greedy decisions at each hop, without considering the congestion scenario beyond neighbor nodes [6]. A better and more efficient selection strategy is implemented using the Regional Congestion Awareness (RCA) algorithm [3]. In this strategy, congestion information from all the downstream neighbors (local as well as non-local) of a router is aggregated and propagated to its upstream neighbors. RCA uses 9-bit wide control channel for transferring this aggregated congestion information. This causes extra overhead in the network, even though it delivers good performance.

We identify few limitations of RCA selection strategy and propose a cost effective dynamic model of RCA that can improve the performance of an NoC under varying traffic patterns. Conventional RCA performs best under non-local traffic patterns. We modify RCA such that it delivers improved performance under local traffic also. We incorporate a traffic analyser on each router to analyse the run time traffic patterns. Based on the traffic, a switching technique chooses the normal RCA under non-local traffic load and cut down the resource over-head of RCA when the traffic is local. This migration (switching) makes RCA behave as local best selection strategy in routers where traffic is mostly towards local (nearby) destinations.

The rest of the paper is organised as follows. We describe the related works in Section II. In Section III, the motivation for the proposed work is explained. The architectural details of the proposed model is given in Section IV. Experimental methodology and result analysis are covered in Section V followed by the conclusion of our work in Section VI.

II. Related Work

The packets generated at the source core need to be directed towards the destination core through the network. Routing algorithms are used to determine the sequence of channels (inter-router links) a packet traverses from the source to the destination. Minimal Odd-Even (MOE) Routing [4] is a commonly used adaptive deadlock free routing algorithm. It restricts the location where certain turns that a packet can take while moving to the downstream routers. Once a packet reaches a router, the MOE routing may return more than one admissible output ports. To enhance the performance of a routing function, output selection functions are employed on top of MOE routing. Selection function captures the congestion metric of the reachable downstream neighbors and chooses that neighbor which is less congested, thus reducing the delay in movement of packets towards their respective destinations. Several parameters of the network as well as routers are used as congestion metric. Neighbors-on-Path [5] checks the Free Virtual Channels (FVCs) of reachable neighbors of adjacent downstream routers. TRACKER [6] uses the history of flow of flits through all the output ports of reachable downstream routers. In BOFAR [7] , the cycles spend by a flit in a buffer is taken as the congestion metric. Global Congestion Awareness (GCA) [8] is yet another technique that uses local as well as non-local status information for computing congestion metric.

RCA is an effective path selection technique that improves the load balance in the network. It aggregates and propagates congestion information about a region of the network beyond the adjacent routers to the upstream routers. This helps the upstream routers in estimating the best path with minimal congestion. The RCA unit of a router basically consists of two modules. The aggregation module combines the local as well as the non-local congestion metric by assigning appropriate weights to each of them. The propagation module sends the aggregated congestion metric to respective neighbors through dedicated control network.

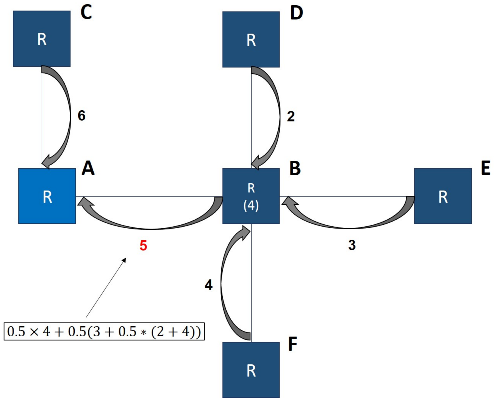

Fig. 2. Illustration of aggregation and propagation in RCA

The working of RCA is illustrated in Figure 2. The figure shows the aggregation and propagation of the congestion metric values in RCA-FanIn [3] selection strategy through a link. Here A, B, C, D, E and F are routers in an 8x8 mesh network. The curved edges connecting the routers are control channels. Assume that a flit in router A has two permissible output neighbors (B or C) after MOE routing. In RCA, the propagated congestion information from the downstream neighbors B and C is used for choosing one of the neighbor. The congestion metric from B is computed as follows:

B adds the local congestion metric (i.e., 4) and the average congestion metric from its neighbors (i.e., 2, 3, and 4 from D, E and F, respectively) in the ratio 0.5:0.25:0.5:0.25. This calculation is illustrated in the bottom left side of the figure. This gives a higher weightage for local information over non-local information. Also, the congestion metric given by the east neighbor of B (i.e., value 3 passed by E to B) is given more weightage than north and south directions. This is because both north-east and south-east destinations include east channel. The congestion metric given to B by routers D, E and F are also an aggregated value of a similar computation done in those routers. Thus A is given the value 5 from B as shown in the figure. The congestion value is a combination of the number of free buffers and crossbar demand of routers. Similarly, an aggregated value is obtained from C also. Based on these values, A chooses the next router to send out the flit.

Local best selection strategy uses the number of FVCs [11] in the adjacent downstream routers. If local best selection strategy was used in the above case, the router B and C transmit their local free virtual channel count to A. This information alone will be used for the output port selection. Even though it has a simpler circuitry and smaller control channel, it takes greedy decision without considering the congestion status beyond downstream routers.

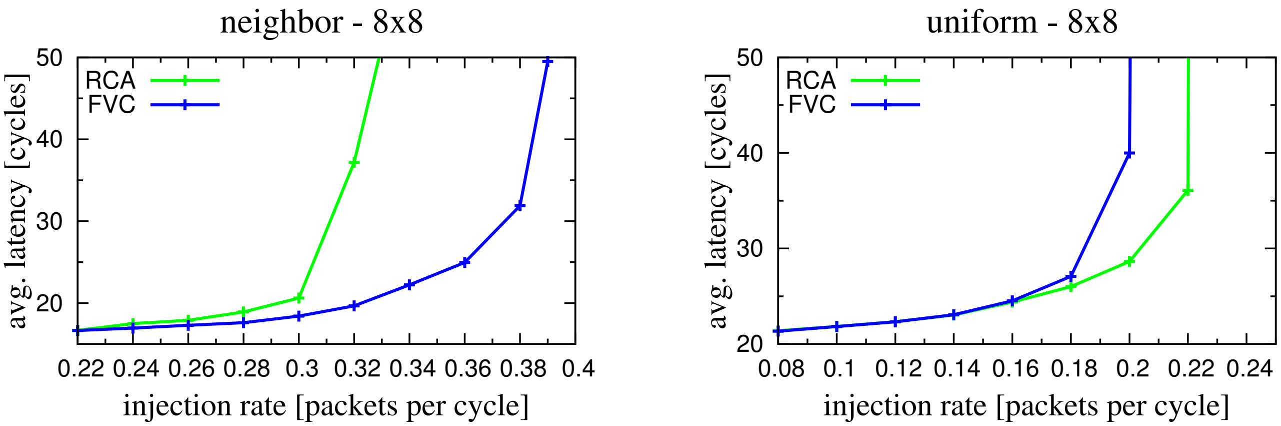

Fig. 3. Comparative analysis of average packet latency versus injection rate for neighbor and uniform traffic patterns in 8×8 mesh network.

III. MOTIVATION

Booksim [9] is a cycle accurate simulator specialized for NoCs with a highly precise underlying network model. Using *Booksim*, we model an 8x8 mesh network that employs MOE routing. We studied the impact of the FVC selection strategy under uniform and neighbor traffic patterns. Synthetic traffic patterns are abstract models of message passing in NoCs. Simple synthetic traffic patterns like uniform, tornado, bit-complement and neighbor traffics allow a network to be stressed with a regular, predictable pattern which aid NoC designers in acquiring new insights [1], [12]. In uniform traffic, each node sends messages to other nodes with an equal probability (i.e., destination nodes are chosen randomly using a uniform probability distribution function). In neighbor traffic, the source and destination are nodes at 2 hop distance that differ in one cordinate both row-wise and column-wise. Similarly, *Booksim* is modified to model the RCA-FanIn architecture as mentioned in [3]. We obtained the average packet latency for varying injection rates from zero to saturation using uniform and neighbor traffic. The load vs. latency graph is shown in Figure 3.

From the figure, we can observe that under neighbor traffic, FVC technique outperforms RCA by a significant margin. In neighbor traffic, every packet's source core and destination core are at a two hop distance (except for packets originating from edge and corner routers). i.e., a packet generated into the network will travel only through two intermediate routers before reaching its destination. So, a selection strategy like RCA that aggregates non-local congestion metric (congestion metric of routers beyond two hops) is meaningless. In our experiments using neighbor traffic, in many cases we observed that RCA selection strategy selects output channels that are not in favour for a packet whose destination is within two hops. Local best selection strategy is a simple technique which requires only 4 bit-lines for communicating the free buffer count of neighbors. It performs well in local traffic loads compared to other techniques.

But RCA was significantly outperforming FVC technique under uniform traffic. In uniform traffic the average hop-length of a packet (in an 8x8 mesh network) is more than six, which indicates that uniform traffic is an example of a non-local traffic. So as expected RCA outperformed FVC technique. These two contradicting observations emphasizes the fact that selection strategy as such cannot improve performance. Certain selection strategy is meaningful and productive under particular traffic patterns only. For a fabricated chip containing a collection of NoC routers, we cannot change the selection strategy on individual routers based on run time traffic pattern.

RCA requires a 9-bit lines in the case of an 8×8 mesh for propagating the congestion feedbacks [3]. These lines are always used irrespective of the traffic load and pattern. When the traffic is mostly towards neighboring nodes (local traffic), congestion information about an entire region is irrelevant. Thus RCA uses equal amount of power and area in the feedback network even when the signals through them are not productively used.

IV. THE PROPOSED WORK

We propose that the performance of an NoC can be improved irrespective of the traffic pattern, if selection strategy is decided based on run-time traffic. We put forward a technique that can take dynamic decision based on the traffic pattern in the network. We implement a runtime traffic pattern analyser for each router that checks whether on-chip traffic is local or non-local. Based on the nature of the traffic, the appropriate selection function is used. If the traffic is non-local, the RCA technique is used and it is made to behave as FVC selection strategy when the traffic is local by proper weight adjustment in the RCA aggregation module.

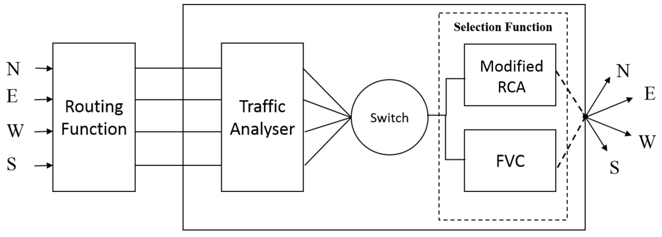

Fig. 4. Proposed Router Architecture

Figure 4 shows the router architecture of our proposed system. We use MOE routing algorithm to obtain the admissible output ports for the incoming packets. The traffic analyser decides the nature of the traffic at each router. The switching circuit enables a suitable selection strategy that can yield favourable results.

A. *Traffic Analyser*

The traffic analyser, extracts the destination address of each packet passing through a router and analyses this information for every T clock cycles. Two 5-bit saturating counters L (local) and N (Non-local) are used to represent the traffic pattern through each router.

If a packet's destination is more than 2-hops away from the current router, it is considered as a non-local packet, otherwise a local packet. The analyser calculates the number of hops to the destination and updates the L and N counters. This information is delivered to the switch for deciding the selection strategy. The counters are cleared at the end of the T cycles.

B. *Switch*

We implement a switching technique that enables either the RCA or the local best selection strategy at a given time in a router. The switching condition is checked at every T clock cycles. If T value is high, the network's response to change in traffic pattern will be slow. If it is low, frequent switching can affect throughput. We study the performance of the system for various switching intervals at T = 8, 16, 32, 64, 128 and 256 clock cycles. Based on these experiments, we fix T at 32 clock cycles, since it give better results compared others.

At the end of 32 clock cycles, the traffic pattern is decided using the analyser output. If the traffic trend is prominent towards non-local destinations, the customised RCA-FanIn selection strategy is enabled. Otherwise, aggregation portion of RCA circuitry is shutdown to make RCA behave as local best selection strategy. A router is evaluated to be local or non-local based on the ratio of the traffic through it.

For a non-edge router, ratio of number of local nodes to non-local nodes in an 8×8 mesh network is 12:51. i.e., the influence of the non-local packets will be more prominent in deciding the overall performance of the network. So, in routers, having a fixed minimum ratio of non-local traffic through it, the RCA selection strategy (which has higher visibility) was to be used. Thus a decision parameter value (i.e., a ratio of non-local packets to local packet through that particular router) must be slightly in-favour of choosing the RCA over FVC. We carried out a study by varying this value (x) in the range 0.25 to 0.45 under various traffic patterns. After detailed analysis of the results, we decided to fix this value as 0.40 to obtain a consistent performance. That is, if at least 40% of the router traffic is towards non-local destinations, then the router should work in RCA. Algorithm 1 explains the switching logic.

In an extended study, we apply different switching points (decision value) for different regions of the mesh network. Better results were obtained when switching is done at x=0.35 for the two outermost layer routers and x=0.45 for the inner layer routers of an 8×8 mesh NoC (48 out of 64 routers are positioned in the outer 2 layers).

The switching logic is implemented on every router, so that, independent switching of each routers rather than a collective switching of the entire network takes place. Thus different routers in the NoC operate in different selection strategies at a given time. The analyser and switch takes only data within each router, so no additional network communication overhead is added for implementation of this migration.

```
Data: L=Local value; N=Non-Local value
Result: Chooses the appropriate selection strategy
for every T clock cycles do
    Compute $x$= N/(L+N)
    if $x > 0.4$ then
        Switch to modified RCA;
    else
        Switch to FVC;
    end
end
```

Algorithm 1: Algorithm for switching

C. *Cost effective RCA module*

The original RCA transfers the congestion metric periodically to the upstream routers. In our design, all the routers need not be operating in RCA. This can affect the network performance if the current version of RCA is used, as some routers will not respond to a status request. So we customize RCA, so that, the latest status information is transferred to the neighbors before the router shutdown few RCA feedback lines. This helps in propagating a fair congestion information across the network than fixing a static value for the RCA-dormant routers.

The local best selection strategy is implemented using the resource subset of RCA itself. The design is such that, turning off a certain portion of RCA circuit is in fact the local best selection strategy itself. This is done by weight adjustment in the aggregation module of RCA. RCA gives 50-50 weightage for local and non-local information. When FVC is to be used in our design, rather than using a different circuitry, 100% weightage is given to local congestion metric and zero weightage to non-local congestion metric. This ensures that the use of two selection strategies does not increase the router hardware cost.

V. EXPERIMENTAL ANALYSIS

We customize the simulator as per our proposed architecture and analysed for 8x8 mesh network using various standard synthetic traffic patterns. We implement our system in view of combining the advantages of both FVC and RCA techniques. So we perform a comparative study with respect to these two selection strategies.

A. *Analysis of Average Packet Latency*

Latency of a packet is defined as the number of cycles needed for the packet to travel from its source to destination. It is a crucial factor for evaluating the performance of an NoC based system. Lower the average latency of the packets, faster the cache miss will be serviced for the application running on the source core. Hence for better performance, the average latency should be as low as possible.

The average latency of RCA, FVC and our selection strategy at different injection rates for an 8x8 mesh network is obtained. Figure 5 shows the injection rate vs average

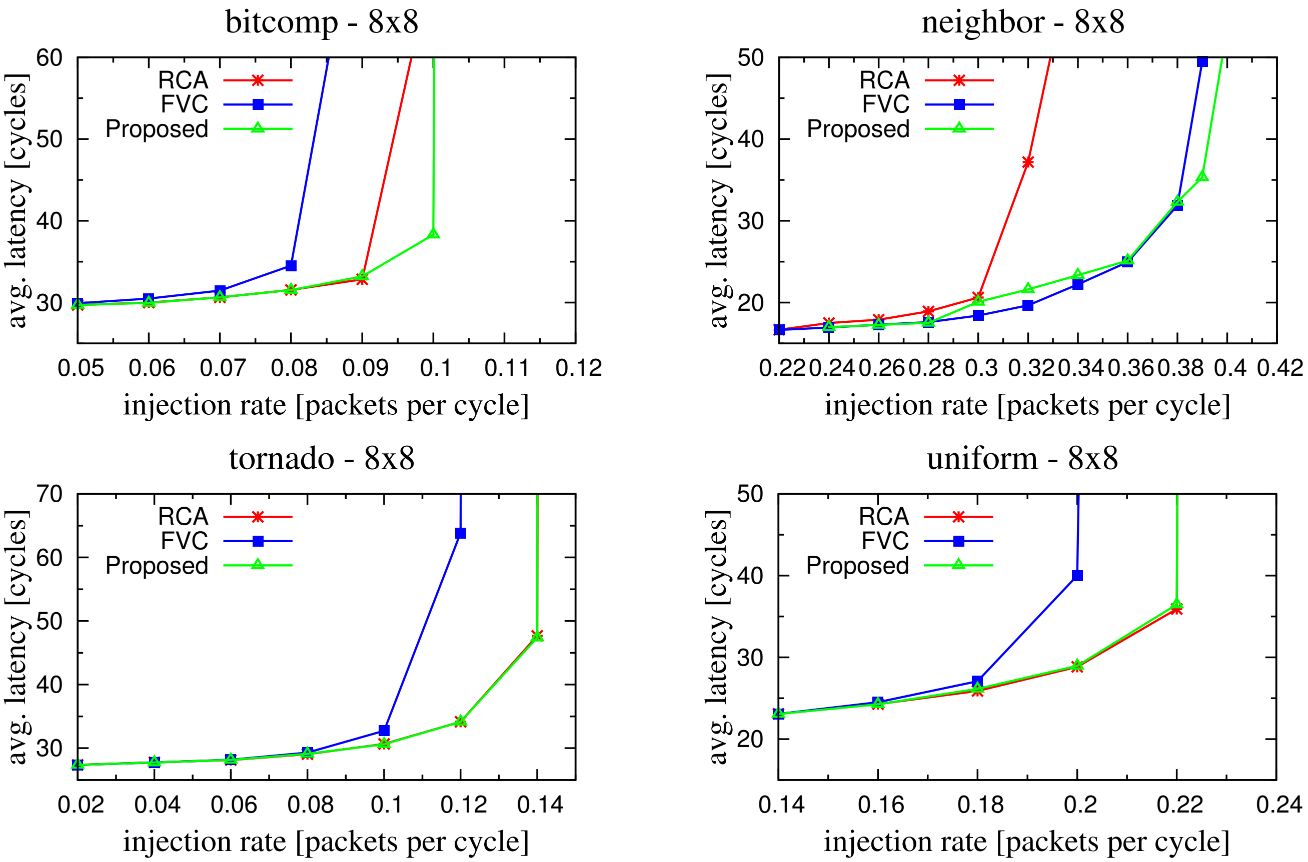

Fig. 5. Comparison of average packet latency in 8×8 mesh network using various synthetic traffic patterns.

packet latency graphs for bit-complement, neighbor, tornado and uniform traffics. From the graphs it is clear that across all synthetic traffic patterns our proposed technique shows lower latency values than FVC and conventional RCA.

In neighbor traffic, our technique shows values closer to FVC. In tornado, bit-complement and uniform traffics the RCA performs better than FVC, as the traffic pattern in most of the routers are non-local as evident from Figure 6. So, the latency curve of our system is close to that of RCA. This validates the claim that, if for a traffic RCA outperforms FVC, then our technique has latency values close to that of RCA. Otherwise, if FVC is better than RCA, our system gives latency near to that of FVC. i.e., based on the traffic pattern it adapts to the best performing strategy to deliver the least average latency. The migration across selection strategies helps our system to achieve a robust performance across traffics.

B. Control Network Design

The additional network resource needed for any selection strategy is the communication channels required for transferring the status information across the network. RCA-FanIn requires 9 bit control channel, whereas, FVC needs only 4 bit-lines for an 8×8 mesh with 16 virtual channels per router port.

Our design proposes to hardwire the resource requirements of RCA, but use them only when required. That is when the router has to operate in conventional RCA, it uses the entire 9-bit control channel, but as the traffic is either very low or local, only the 4 bit-lines are needed to operate as FVC. So, in the worst case (all the routers handling non-local heavy load) the network resource requirement of our system is equal to that of RCA.

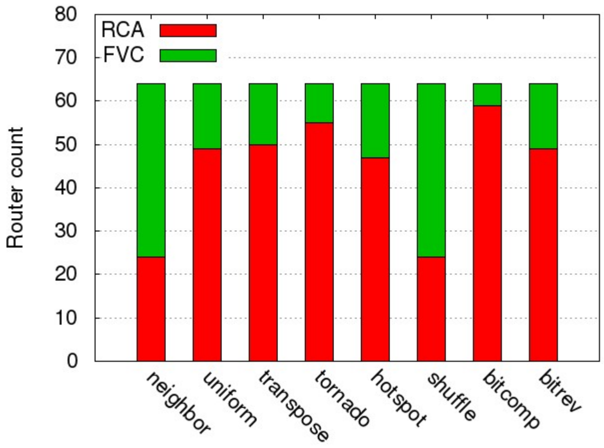

Fig. 6. Count of routers working in RCA and FVC

Figure 6 shows the number of routers that are operating in RCA and FVC for various traffics in an 8x8 mesh network at saturation load. This is also a representation of the average communication channel utilisation. We see that many of the routers are working in FVC only, hence, the additional RCA bit lines associated with it will be shutdown. Thus the average network resource requirement is considerably less compared to that of RCA. This in-turn reduces the dynamic power consumption of the system.

For example, consider the uniform traffic. When RCA alone is used the average bit-lines per port in each router is 9. From the graph it is clear that 49 routers are in RCA (9 bit-lines) and the rest is in FVC (4 bit-lines). So the average bit-line utilisation can be calculated as follows: $(49/64\times9+15/64\times4)/9=0.8697$. That is 13% of the entire communication channels are turned off saving a fraction of dynamic power dissipation.

C. Analysis of Router Complexity

It may seem that use of 2 selection strategy along with a traffic analyser and switching technique makes the router architecture complex. In our system, the implementation is done so that, the FVC is obtained by turning off a part of RCA, rather than as a separate selection logic. Hence, the architecture complexity of selection strategy is comparable to that of RCA. The analysis and switching techniques requires only 2 counters and a comparator in addition to an existing RCA router design. This adds a negligible overhead. Thus the routers do not significantly increase the overall system requirements. Verilog synthesis of the proposed router using Synopsys Design Compiler at 65nm shows an area overhead of 3.2% and static power overhead of 2.1% w.r.t RCA design. This overhead is due to the additional hardware units.

D. Analysis using Real Workloads

Apart from synthetic traffic, we evaluate the performance of our proposed system using traces of multi-programmed workloads also. We use Multi2sim [10] simulator to model a 64-core CMP set up with CPU cores, cache hierarchy, and coherence protocols in detail and accuracy. Each core consists of an out-of-order x86 processing unit with a 64KB, 4-way set-associative, 32 byte block, private L1 cache and a 512KB, 16-way set associative, 64 byte block, shared distributed L2 cache. Each core is assigned with a SPEC 2006 CPU benchmark application for running on it.

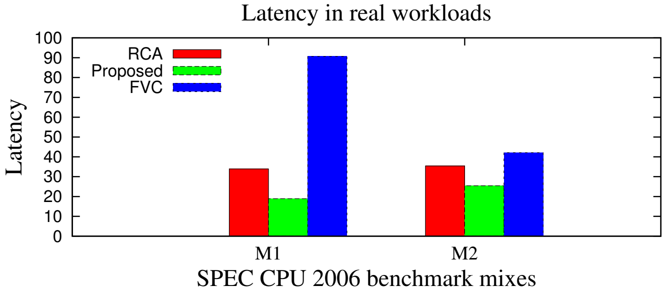

Fig. 7. Comparison of Average packet latency using real workloads traffic traces.

We prepared 2 mixes of 64 core multiprogram workloads, M1 consists of 16 instances of medium *misses per kilo instructions* (MPKI) applications like *bwaves*, *bzip2*, *gamess*, and *gcc* and M2 consists of 16 instances of low MPKI applications like *calculix*, *gobmk* and 16 instances of high MPKI applications like *mcf*, and *leslie3d*. After sufficient fast forwarding, we capture the L1 cache misses that generate network traffic and feed it to the modified Booksim model to simulate the network operations.

Figure 7 shows the performance comparison graph of the proposed selection strategy with RCA and FVC. In both mixes the proposed technique has slightly lower latency than RCA, and much lower than FVC. This establishes the fact that a performance equivalent to that of the better performing strategy will be delivered by our technique in real workloads also.

VI. CONCLUSION

A refined NoC router architecture with a dynamic traffic analyser and a run-time switching is implemented to make the best use of RCA and FVC techniques. Routers uses RCA as such or a cost reduced version of it (effectively behaving as FVC), based on the real-time traffic patterns. Experiments on 8×8 mesh NoCs showed that the proposed design has less latency values consistently across various traffics compared to RCA and FVC. The overall energy utilisation is also reduced compared to RCA as many routers will in FVC. Hence the on-chip network which uses our selection logic can minimize network resource utilisation without affecting performance. As the FVC was designed to be obtained by a weight adjustment on RCA, it has only a very small hardware overhead. Thus using RCA and local best selection strategy in an equilibrium can bring down the power consumption and deliver stable performance. Hence, we conclude that our proposed design will be a good design alternative to future NoCs.

REFERENCES

[1] W. Dally and B. Towles *Principles and Practices of Interconnection Networks* USA: Morgan Kaufmann Publishers Inc., 2003.

[2] W. Dally and B. Towles, *Route packets, not wires: on-chip interconnection networks* in DAC, pp. 684-689, 2001.

[3] P.Gratz, B.Grot, and S.W. Keckler, *Regional Congestion Awareness for Load Balance in Network-on-Chip* in HPCA, pp. 203-214,February 2008.

[4] G. M. Chiu., *The odd-even turn model for adaptive routing* in IEEE TPDS,11 (7):729-738, July 2000.

[5] Giuseppe Ascia, Vincenzo Catania, Maurizio Palesi and Davide Patti, *Implementation and analysis of a new selection strategy for adaptive routing in NoC* in IEEE TOC, 57 (6), pp. 809-820, 2008.

[6] John Jose, K.V. Mahathi, J. Shiva Shankar and Madhu Mutyam, *TRACKER: A low overhead adaptive NoC router with load balancing selection strategy* in ICCAD, pp. 564-568, 2012.

[7] John Jose, J. Shiva Shankar, K.V. Mahathi, Damarla Kranthi Kumar and Madhu Mutyam, *BOFAR: Buffer occupancy factor based adaptive router for mesh NoC* in NoCArc, pp. 23-28, 2011.

[8] Mukund Ramakrishna, Paul V. Gratz and Alexander Sprintson, *GCA: Global congestion awareness for load balance in Networks-on-Chip* in NOCS, pp. 21-24, 2013.

[9] Nan Jiang, Daniel U. Becker, George Michelogiannakis, James Balfour, Brian Towles, John Kim and William J. Dally, *A detailed and flexible cycle-accurate Network-on-Chip simulator* in ISPASS, pp. 86 - 96, 2013.

[10] R Ubal et al., *Multi2sim: A simulation framework to evaluate multicore-multithreaded processors* in SBAC-PAD, 2007, pp. 62-68.

[11] J. Kim, et al., *A low latency router supporting adaptivity for on-chip interconnects* in DAC, pp. 559-564, 2005.

[12] K. Lahiri et al., *Evaluation of the traffic-performance characteristics of system-on-chip communication architectures*, Proceedings of the International Conference on VLSI Design (2000), pp. 2935.