An Energy Efficient Load Balancing Selection Strategy for Adaptive NoC Routers

John Jose, Bivil M Jacob, Hashim P Kamal
Department of Computer Science and Engineering
Rajagiri School of Engineering and Technology, Kochi, India
johnj@rajagiritech.ac.in, bivilj@gmail.com, hashimpkamal@gmail.com

ABSTRACT

Modern chip multi core systems are using Network on Chip (NoC) as the communication infrastructure. Effective output channel selection techniques are used in adaptive routers, which form the back bone of NoC systems to reduce the average packet latency of inter-core communications in multicore systems. We propose a selection strategy, *Cool Centers*, for output port selection that ensures load balancing on a mesh NoC system. Cool centers reduces the possibility of traffic hot-spot formation in the network and can be applied on any minimal adaptive routing algorithm for improving the system performance. The proposed system equally distributes the traffic load among the available minimal paths without any significant architectural overhead. This reduces the rate of non-uniform wear and tear of routers and links, and prevent early aging of chips.

Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design**]: [Network communications]

Keywords

traffic distribution, hot-spot elimination, aging

1. INTRODUCTION

Network on Chip (NoC) is an emerging communication architecture for multi-core processors. It replaces the traditional bus based communication framework between IP cores in a System on Chip by a regular, well-structured interconnection network. The basic idea of NoC is inherited from off-chip multi-processors and distributed computing networks. The scalable and modular nature of NoCs and their support for efficient on-chip communication increase its popularity.

In majority of the recent homogeneous multi-core designs, a single core consists of an out-of-order superscalar processor with a private L1 cache and a shared distributed L2 cache as shown in Figure 1. Cores communicate each other through a well structured on-chip network that routes packets between them. Each core interfaces with the network over an input port for inserting packets into the network and an output port through which the network delivers packets to the core [1]. Inter-core communication is needed during cache misses, coherence and synchronization transactions. When a core wants to communicate with another, it creates a packet and forwards it to its router via the Network Interface Controller (NIC).

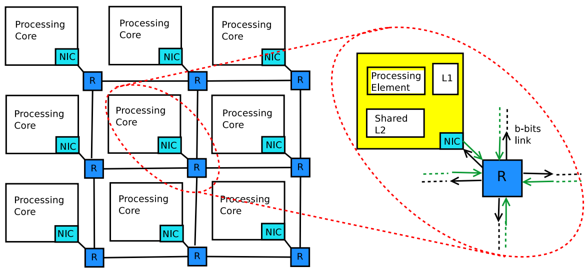

Figure 1: Interaction of a core and an NoC router in a 2D mesh topology.

Traditional NoC routers have buffers associated with each input port and are generally called as input buffered virtual channel (VC) routers [2]. The buffers in the VCs act as the storage units for the flits from the time they arrive at the router till they leave the router. Three tasks are performed on the packets when they reach a router. Route computation identifies the next output port for the packet. VC allocation reserves a VC in the downstream router for every routed packet (packet on which route computation is done). When multiple packets are competing for the same downstream router, switch allocation performs the arbitration to decide the winner.

Based on the routing algorithms implemented in routers, the packets move from source core to destination core in a hop-by-hop manner. Buffering within the routers and two way handshaking between routers enable the smooth flow of the packets [16], [17]. Routing algorithms are broadly classified into static and adaptive. Static algorithm uses only one path between a pair of source and destination. Adaptive routing algorithms provide multiple paths from a source to its destination. Here we make use of a selection strategy that chooses the best output port from the set of available ports in every intermediate router. The selection strategy exploits the possibility of routing packets efficiently based

NoCArc '14, December 13 - 14 2014, Cambridge, United Kingdom
Copyright 2014 ACM 978-1-4503-3064-0/14/12 ...$15.00
http://dx.doi.org/10.1145/2685342.2685350.

on the dynamic congestion status of the network. The decision is made based on various parameters like the intensity of congestion and number of free buffers in possible downstream routers.

In this paper, we propose an energy efficient load balancing selection strategy for adaptive routers that routes packets through paths that are more closer to the edges of the network so that traffic through the centrally located routers can be minimized. The organization of the rest of the paper is as follows. We describe the related works in Section 2. The motivation behind the proposed idea is explained in Section 3. Section 4 presents our proposed selection strategy. Section 5 covers the experimental analysis and we conclude the paper in Section 6.

2. RELATED WORKS

One of the simple and the most commonly used deadlock-free adaptive routing algorithm in mesh NoCs is the Minimal Odd-Even (MOE) routing [5]. MOE routing restricts the locations where certain turns can be taken so that deadlock is avoided. In MOE routing, EN turn and ES turn are prohibited at routers in the even columns. Similarly NW turn and SW turn are prohibited at routers in odd columns. The naive MOE routing does not have any output port selection strategy, but it makes a random selection from the available ports. For enhancing the performance, output port selection strategies are employed on top of MOE routing.

In the Free Buffer Priority (FBP) [7] based selection strategy the count of free buffers in the input port of the adjacent downstream router is used as the selection metric. The free buffer status of reachable neighbors of adjacent routers of current node is used in the Neighbors-on-Path (NOP) [8] strategy. In Buffer Occupancy Factor based Adaptive Router (BOFAR) [9] selection strategy is done based on a congestion metric computed from the history of buffer occupancy time of flits. TRACKER [10] is a load balancing selection strategy that make use of history of flit flow estimates for output port selection.

Regional Congestion Awareness (RCA) [11] and Global Congestion Awareness (GCA) [12] present a comprehensive evaluation and usage of non-local congestion information for improving the dynamic load balancing properties of fully adaptive minimally routed networks. Path-Based Randomized Oblivious Minimal Routing [13] proposes a load balancing routing scheme by random channel selection. Congestion Aware Deterministic Routing [14] estimates the congestion level in the network based on past flow pattern and computes optimized routing paths for all traffic flows deterministically. This is suited only for systems which generate repetitive traffic patterns.

All the above selection strategies except FBP approach handle congestion by making use of additional control lines between routers for fetching congestion information from neighbors and use additional circuits for computing the selection metric. Even though some of them [10], [11] improve performance, they incur energy and hardware overhead on implementation. The cost overhead of some of these approaches in terms of area, power and control network is discussed in [6]. The uneven distribution of traffic on the chip is a major performance issue and must be dealt with at most care to ensure long life for routers by reducing uneven wear and tear. An aging-aware adaptive routing algorithm [4] that routes the packets along the paths which are both

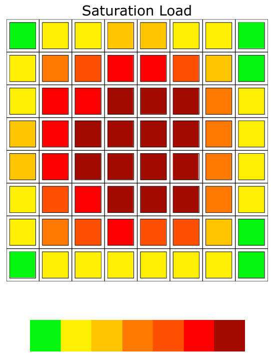

Figure 2: TDCG for MOE routing with Free Buffer Priority (FBP) selection strategy in 8x8 mesh network using uniform traffic at saturation load.

least congested and experience minimum aging stress.

The proposed selection strategy *cool centers* is an alternative for these mechanisms without imposing any hardware overhead. It ensures load balancing to a great extend in comparison with the existing strategies and guarantees extended chip life by reducing non-uniform wear and tear.

3. MOTIVATION

Majority of the existing adaptive routers employing MOE routing use free buffer count in the downstream node (FBP) as the priority scheme for output channel selection decisions. FBP mechanism gives higher priority to that direction which has the highest number of available free buffers in the input port of the downstream router. This sounds to be a simple and meaningful approach as congestion is measured by the count of free buffers in the down stream routers. But it can lead to load imbalance in centrally located routers.

To understand the traffic behavior of FBP selection strategy we modify the cycle accurate *BookSim* [3] simulator and implement MOE routing with FBP output channel selection on top of it for an 8x8 2D mesh network using uniform traffic. At the end of the network simulation we captured the count of the packets passed through each router. From this data, we plot the Traffic Density Color Graph (TDCG) that represents the packet flow density in each router.

TDCG for an 8x8 mesh NoC consists of 64 squares organized as 8 rows and 8 columns. Each square represents a router in the 8x8 mesh network. Each square is assigned a color from among the range given in the rectangular chart at bottom of Figure 2. Colors towards green end represent low traffic flow and colors towards brick red end represent high traffic flow. Based on the traffic density value obtained from the simulation we assign the appropriate color to each of the router squares. The TDCG for an 8x8 mesh network at saturation load using MOE routing with with FBP selection metric is given in Figure 2. TDCG of TRACKER and RCA also have an almost similar pattern. Saturation load is defined as the minimum injection rate (packets injected/node/cycle) at which linear increase in load results in exponential increase in packet latency. Upon analyzing the TDCGs in Figure 2 we make the following observations:

- The number of packets moving through the centrally located routers is very high in comparison with the routers at the edges and corners of the network.

- This results in non-uniform usage of network resources leading to uneven wear and tear of routers and links.
- Non-uniform wear and tear of routers and its associated channels causes early aging of the chip.
- This differential traffic density may lead the centrally located routers to loose its electrical and semiconductor properties faster than the rest of the routers, thereby reducing the chip life.

From the above observations we explore the possibility of a new selection strategy that can eliminate the accumulation of traffic density on centrally located routers of the chip and spread a fraction of the load to edge and corner routers.

4. COOL CENTERS STRATEGY

We propose a new selection strategy, called Cool Centers Priority (CCP) for adaptive routers employing minimal routing algorithms in mesh NoCs. Since we use MOE routing, this selection strategy will not increase the number of hops the packet has to take to reach the destination. The heart of the selection strategy is to choose an output port from the existing set of output ports so as to reduce the traffic flow through the center of the mesh. This is done by giving a higher priority to that port which routes the packet away from the center of the mesh. Due to load spreading towards edge and corner routers, it ensures load balancing and reduces the occurance of hot-spot formation at the center of the mesh.

Algorithm 1 Cool Centers Priority (CCP) selection strategy algorithm for an nxn mesh network.

Input: *Output port P, Current router C.*
Output: *Hot_ spot value of the output port P.*
Find Router R, the neighbor of (C) connected through the output port P.
if R is the destination router for the packet **then**
  return -1 {no need for channel selection}
else
  col = column number (R).
  row = row number (R).
  x = Minimum of (col, n-1-col).
  y = Minimum of (row, n-1-row).
  hotspot_ value = x + y.
  return (hotspot_ value).
end if

In CCP selection strategy, once the adaptive route function returns the set of possible neighbors, the relative distance of the identified neighbor from the edge of the mesh is computed. The neighbor with minimum distance (in hops) from the edge (in both X and Y dimensions) is given high priority in output channel selection. The working of the CCP selection strategy is explained in Algorithm 1. For each of the output port P, identified by the adaptive routing function, the algorithm compute the hot-spot value of the respective port. The port with lower hot-spot value is chosen for forwarding the packet. We name the proposed algorithm as *cool centers* as it reduces the traffic density at the centrally located routers of the chip. Load reduction in central routers reduces the dynamic power dissipation thereby making them much cooler than the conventional designs.

Table 1: Percentage of different network injection intensity applications in various benchmark mixes.

Benchmark Mix	M1	M2	M3	M4
% of Low	100	0	0	31
% of Medium	0	100	0	31
% of High	0	0	100	38

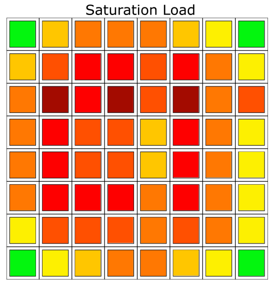

Figure 3: TDCG for MOE routing with Cool Centers Priority (CCP) selection strategy in 8x8 mesh network using uniform traffic at saturation load.

5. EXPERIMENTAL ANALYSIS

We customize the *Booksim* [3] simulator to model the MOE routing with CCP selection strategy for 8x8 mesh network. We use 128-bit wide flit channel and 4-bit wide credit channel. We capture the average latency and traffic density values for various traffic patterns from zero load to saturation. We evaluate the performance of our proposed system using traces of multiprogrammed workloads. We use Multi2sim [15] simulator to model a 64-core CMP set up with CPU cores, cache hierarchy, and coherence protocols in sufficient 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. Based on the *misses per kilo instructions* (MPKI) values calculated on a 64KB L1 cache, we classify the benchmarks into Low (i.e., MPKI less than 5), Medium (i.e., MPKI between 5 and 25) and High (i.e., MPKI greater than 25).

In our experiments, we use *calculix*, *gobmk*, *gromacs*, and *h264ref* in the Low MPKI group, *bwaves*, *bzip2*, *gamess*, and *gcc* in the Medium MPKI group, and *hmmer.nph3*, *lbm*, *mcf*, and *leslie3d* for the High MPKI group. We construct 40 multiprogrammed workloads, each with 64 single threaded benchmark instances. We categorize these workloads into 4 mixes (M1 to M4) based on the proportion of the network injection intensity *(Low / Medium / High)* of the constituent benchmarks. Details of the mixes are given in Table 1. For example M3 consist of High MPKI applications in all the 64 cores. After sufficient fast forwarding, we capture the L1 cache misses that generate network traffic and feed to *Booksim* to simulate the network operations.

Even though other techniques [10], [11] achieve better performance than our proposed method they incur significant hardware overheads. so we compare CCP with FBP only because of the comparable level of overheads.We conduct a comparative analysis to study the performance of CCP with respect to FBP selection strategy.

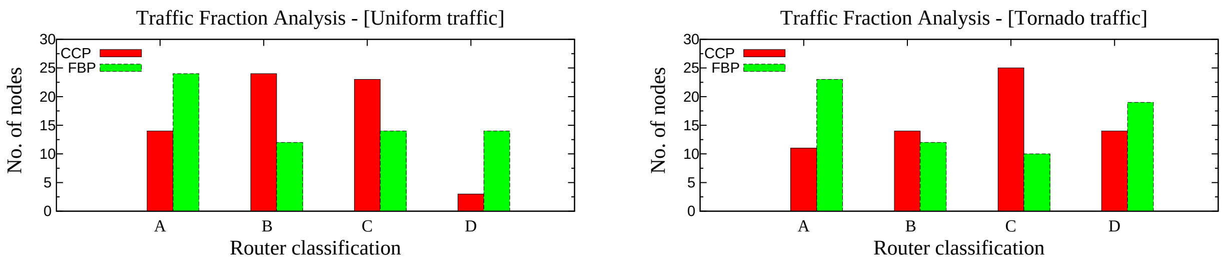

Figure 4: Traffic fraction analysis using synthetic traffic at saturation load on 8x8 mesh network.

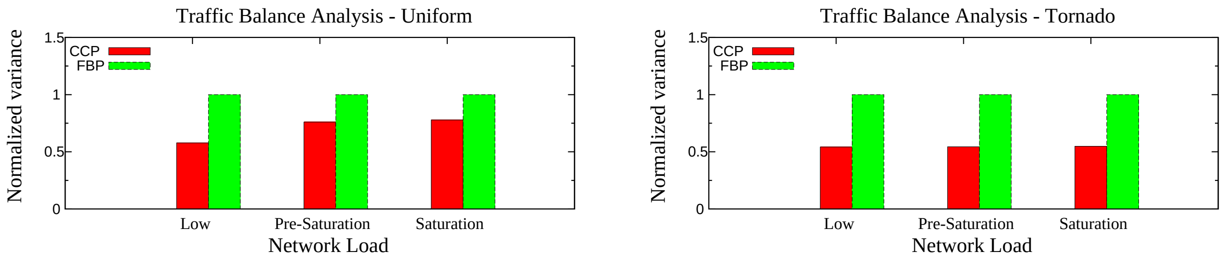

Figure 5: Traffic variance analysis in 8x8 mesh network using synthetic traffic patterns.

5.1 Effect on Load Balancing

The load on a router is defined as the number of packets that pass through the router in a given time period. We capture the load on each router for a period of 30,000 cycles. Figure 3 shows the TDCG for MOE routing with cool centers selection strategy using uniform traffic in 8x8 mesh network. We can clearly see that when compared to FBP (Figure 2) CCP selection strategy reduces the number of hot-spots (routers with very high traffic density) as well as cool-spots (routers with very low traffic density). This clearly shows that CCP selection strategy is able to balance the load in the network evenly across available routers.

From TDCG of CCP selection strategy, we observe that the traffic is highly distributed, thereby eliminating the formation of traffic hot-spots at the centrally located routers. The reduction in the number of very low traffic routers and very high traffic routers and the increase in the number of routers with average traffic ensures the high degree of load balancing in our proposed cool centers selection strategy.

To get a better understanding of this load distribution, we perform a load fraction analysis of the network in a time frame of 30,000 cycles. For this, we classify the routers into four category; Type A (routers that handle very low traffic of less than 60,000 packets), Type B (routers that handle below average traffic of between 60,000 to 80,000 packets), Type C (routers that handle above average traffic of between 80,000 to 1,00,000 packets), and Type D (routers that handle very high traffic of more than 1,00,000 packets). We plot the count of routers that belong to these classifications (Type A, Type B, Type C, and Type D) at saturation load for uniform and tornado traffic patterns in Figure 4.

From Figure 4 we can see that the count of Type A and Type D routers are very high in FBP selection strategy whereas using CCP selection strategy the count of Type A and Type D routers have significantly reduced. This means that cool centers selection strategy significantly reduces hot-spots (routers with very high traffic - Type D routers) and cool-spots (routers with very low traffic - Type A routers). Another observation is the significant increase of Type B and Type C routers (routers with moderate traffic) using CCP selection strategy. Cool centers significantly distribute the load evenly across available routers.

5.2 Effect on Traffic Variance

We analyse the traffic flow variance across various routers. From the count of packets passing through each router, we calculate the Mean (M) of the total number of packets per router as per the following equation.

$$M = \frac{\Sigma_{i=1}^{N} count(i)}{N} \tag{1}$$

where N is the total number of routers in the network and count (i) is the total number of packets passing through router i.

From the mean M, we define the traffic variance (V) as per the following equation.

$$V = \frac{\Sigma_{i=1}^{N} abs(M - count(i))}{N} \tag{2}$$

Variance is an indication of evenness in packet flow. It can also be interpreted as the measure of traffic imbalance in the system. Lower the traffic variance, better will be the load balance. Figure 5 shows the plot of normalized traffic variance with respect to MOE routing with FBP selection strategy at low load, pre-saturation load and saturation load. Low load is the load at which the packet latency is almost constant even if injection rate increases. Pre-saturation load is the load at which the packet latency increases linearly with injection rate. Saturation load is the load at which the packet latency increases exponentially with increase in packet injection rate. Across all these three loads, we can see that CCP has the lowest traffic variance than other two techniques.

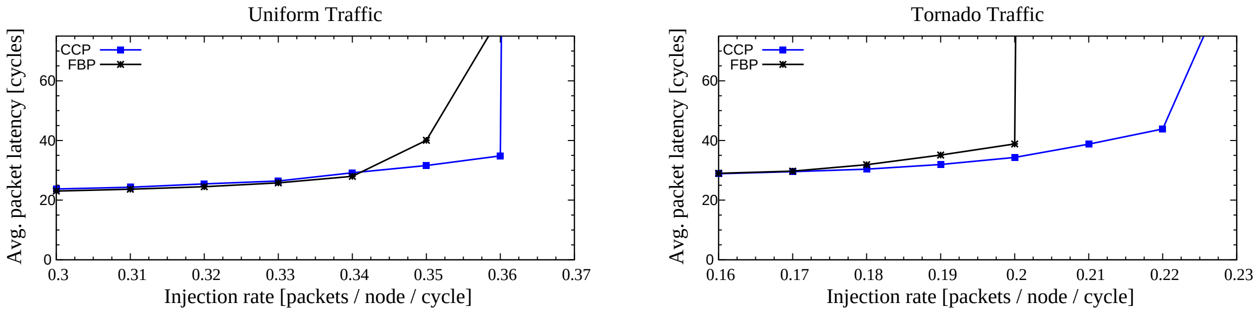

Figure 6: Comparison of average packet latency in 8x8 mesh network using synthetic traffic patterns.

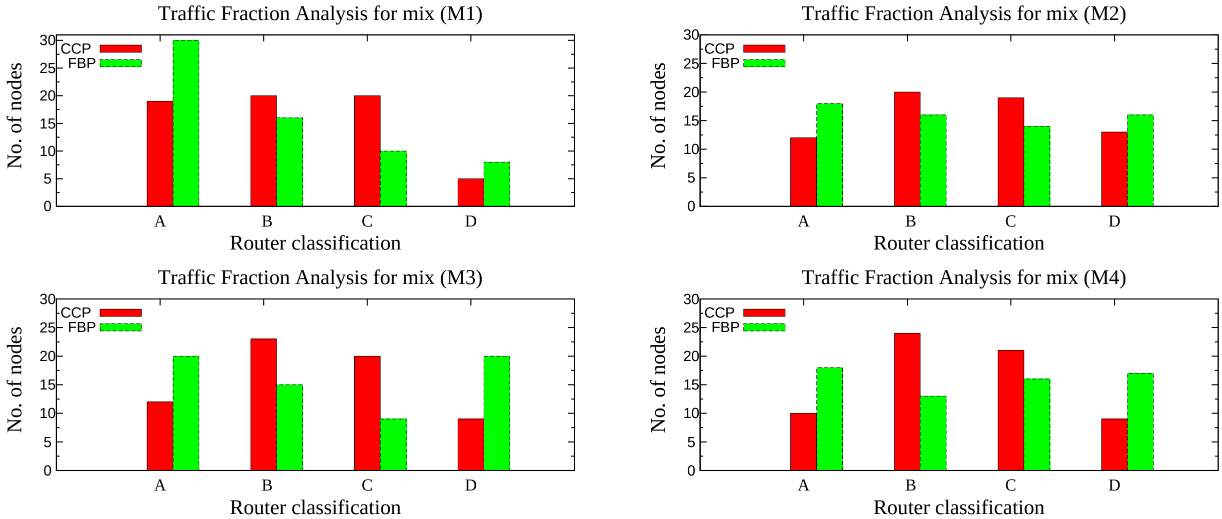

Figure 7: Traffic fraction analysis using SPEC 2006 CPU benchmark mixes in 64-core CMP.

5.3 Effect on Average Packet Latency

Average packet latency is defined as the average time (cycles) taken by the packets to reach the destination from the source. Latency is a crucial factor for evaluating the performance of an NoC system as it represents a significant fraction of the cache miss penalty for cache miss packets. Lower the average latency for the packets, smaller is the stall time for the application running on the respective source core. Hence for better application level performance, the average latency should be as low as possible. Average latency T_{avg} is given by the following equation.

$$T_{avg} = \frac{\Sigma_{i=1}^{i=P} t_i}{P} \tag{3}$$

where t_i is the time taken by the i^{th} packet to reach from source core to its destination core, P is the total number of packets generated in the system during the evaluation phase.

Figure 6 shows injection rate vs average packet latency plot for FBP and CCP selection strategy. Since the CCP selection strategy forwards the packets through routers that are as close as possible through the edges of the mesh, they experience less congestion and as a result packets move to the respective destination without much contention. This is visible in the latency reduction using the CCP selection strategy at the saturation load. We can also observe that CCP selection strategy extends the saturation load of the network than the FBP selection strategy. The reduction in average packet latency and extension of saturation load is more predominant in the tornado traffic pattern as the source destination pair in tornado pattern is diagonally oriented in the mesh network. This diagonal orientation of the source and destination leads to more number of output channel selection decisions. Thus we show that CCP technique is a good design choice for high injection rate applications.

5.4 Analysis Using Real Workloads

We also analyze the effect of our proposed technique using traffic traces generated by execution of SPEC 2006 CPU benchmark applications on a 64 core Chip Multi-core Processor (CMP). The traffic fraction analysis is done for all the mixes M1, M2, M3 and M4 (details of the mixes are given in Table 1) and the results are shown in Figure 7 . We can see that across all the workload mixes CCP selection strategy reduces the number of Type-A and Type-D routers and increases the Type-B and Type-C routers thereby ascertaining its load balancing capacity even under real traffic.

We also compute the percentage reduction in average packet latency by using the same mixes (M1-M4) and plot them in Figure 8. The Figure shows the percentage reduction in average packet latency using MOE routing with FBP and CCP selection strategies with respect to NP (No Priority, i.e., random selection) selection strategy for various mixes taken during a window of five lakh clock cycles.

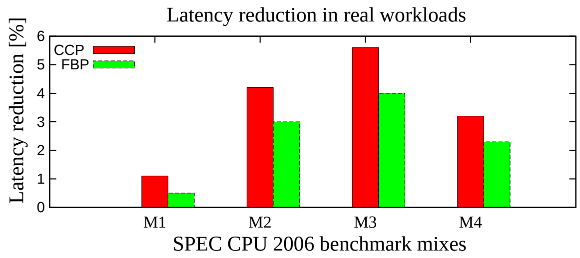

Figure 8: Percentage reduction in average packet latency with respect to MOE routing with random selection strategy using real workloads in 64-core CMP.

In mix M1, where the number of packets flowing through the network is relatively very low due to low MPKI applications running on the cores (hence few cache miss packets in the network), we observe that the latency reduction is marginal. But in mixes M2 and M4 we are able to observe more reduction in latency. We can see that the reduction is even more in mix M3 where there is heavy packet injection. These observations in real work load show that CCP selection strategy is able to handle load imbalance efficiently for high injection rate applications. We observe that the packet latency reduction is a byproduct of a good load balancing selection strategy.

6. CONCLUSION

The proposed CCP selection strategy based NoC router architecture eliminates traffic hot-spots and guarantees balanced load on every router on the network. The non-uniform wear and tear experienced in existing systems is reduced in the proposed method. We also showed that there is no concentrated accumulation of load anywhere in the system. This improves the chip life. Experimental analysis on real and synthetic traffic patterns shows that cool centers can be used as an energy efficient load balancing selection strategy for adaptive routers in future NoC designs.

7. REFERENCES

[1] W. Dally and B. Towles, Route packets, not wires: on-chip interconnection networks, In Proceedings of Design Automation Conference, Pages. 684-689, 2001.

[2] W. Dally and B. Towles, Principles and practices of interconnection networks, San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2003.

[3] 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 Proceedings of IEEE International Symposium on Performance Analysis of Systems and Software, Pages 86-96, 2013.

[4] Kshitij Bhardwaj, Koushik Chakraborty, Sanghamitra Roy, Towards graceful aging degradation in NoCs through an adaptive routing algorithm, In Proceedings of the Design Automation Conference, Pages 382-391, 2012

[5] G.M. Chiu, The odd-even turn model for adaptive routing, IEEE Transactions on Parallel and Distributed Systems, Vol. 11, Pages 729-738, July 2000.

[6] John Jose, Madhu Mutyam, Implementation and Analysis of History-Based Output Channel Selection Strategies for Adaptive Routers in Mesh NoCs, ACM Transactions on Design Automation of Electronic Systems, Vol.9, Number.4, Article 35, August 2014.

[7] W. Dally, Virtual-channel flow control, IEEE Transactions on Parallel and Distributed Systems, Vol. 3, Number.2, Pages 194-205, March 1992.

[8] Giuseppe Ascia, Vincenzo Catania, Maurizio Palesi and Davide Patti, Implementation and analysis of a new selection strategy for adaptive routing in NoC, IEEE Transations On Computers, Vol. 57, Number. 6, Pages 809-820, June 2008.

[9] 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 Proceedings of International workshop on Network on Chip Architectures Pages 23-28, 2011.

[10] John Jose, K.V. Mahathi, J. Shiva Shankar and Madhu Mutyam, TRACKER: A low overhead adaptive NoC router with load balancing selection strategy, In Proceedings of the International Conference on Computer-Aided Design, Pages 564-568, 2012.

[11] Paul Gratz, Boris Grot and Stephen W. Keckler, Regional congestion awareness for load balance in Networks-on-Chip, In proceedings of the International Symposium on High-Performance Computer Architecture, Pages 203-215, 2008.

[12] Mukund Ramakrishna, Paul V. Gratz and Alexander Sprintson, GCA: Global congestion awareness for load balance in Networks-on-Chip, In Proceedings of the International Symposium Networks on Chip, Pages 21-24, 2013.

[13] Myong Hyon Cho, Mieszko Lis, Keun Sup Shim and Srinivas Devadas, Path-based, randomized, oblivious, minimal routing, In the Proceedings of the International Workshop on Network on Chip Architectures, Pages 23-28, 2009

[14] Abbas Eslami Kiasari, Axel Jantsch, and Zhonghai Lu, A framework for designing congestion-aware deterministic routing, In Proceedings of the International Workshop on Network on Chip Architectures, Pages 45-50, 2010.

[15] R Ubal, J Sahuquillo, S Petit and P Lopez, Multi2sim: A simulation framework to evaluate multicore multithreaded processors, In Proceedings of the International Symposium on Computer Architecture and High Performance Computing, Pages 62-68, 2007.

[16] Yaming Yin and Shuming Chen, An application-specific buffer allocation algorithm for network-on-chip, In proceedings of the International Conference on ASICON Pages 439-442, 2009.

[17] Yaming Yin, Shuming Chen, Xiao Hu, Input buffer planning for network-on-chip router design, In Proceedings of International Conference on Computer Application and System Modeling, Pages 201-204, 2010.