iDEAL: Inter-Router Dual-function Energy and Area-efficient Links for Network-on-Chip (NoC) Architectures

Avinash Karanth Kodi[†], Ashwini Sarathy[‡] and Ahmed Louri[‡]
[†] Dept. of Electrical Engineering and Computer Science, Ohio University, Athens, OH 45701
[‡] Dept. of Electrical and Computer Engineering, University of Arizona, Tucson, AZ 85721
[†]kodi@ohio.edu, [‡]{sarathya, louri}@ece.arizona.edu

ABSTRACT

Network-on-Chip (NoC) architectures have been adopted by a growing number of multi-core designs as a flexible and scalable solution to the increasing wire delay constraints in the deep sub-micron regime. However, the shrinking feature size limits the performance of NoCs due to power and area constraints. Research into the optimization of NoCs has shown that a reduction in the number of buffers in the NoC routers reduces the power and area overhead but degrades the network performance. In this paper, we propose **iDEAL**, a low-power area-efficient NoC architecture by reducing the number of buffers within the router. To overcome the performance degradation caused by the reduced buffer size, we propose to use adaptive dual-function links capable of data transmission as well as data storage when required. Simulation results for the proposed architecture show that reducing the router buffer size in half and using the adaptive dual-function links achieves nearly 40% savings in buffer power, 30% savings in overall network power and about 41% savings in the router area, with only a marginal 1-3% drop in performance. Moreover, the performance in iDEAL can be further improved by aggressive and speculative flow control techniques.

Categories and Subject Descriptors

B.4.3 [**Hardware**]: Input/Output and Data Communications—*Interconnections (Subsystems)*; C.5.4 [**Computer Systems Organization**]: Computer System Implementation—*VLSI Systems*

General Terms

Design, Performance

Keywords

Network-on-Chip, Low-Power architecture, Interconnects

1. INTRODUCTION

Technology scaling in the deep sub-micron regime has exacerbated the interconnect design issues such as the global wire delays that do not scale as fast as the gate delays [1], thereby restricting the use of ad-hoc and global shared wiring in the current MultiProcessor System-on-Chip (MPSoC) paradigm [2, 3] and leading to the emergence of modular and scalable packet-switched Network-on-Chip (NoC) architectures [3, 4, 5, 6, 7, 8, 9, 10]. One of the major research challenges currently faced by NoC designers is that of power dissipation, as concluded by the recent NSF sponsored workshop on on-chip networks [11] - "The most important technology constraint for on-chip networks is power consumption" [9]. Power is dissipated by the NoCs in communicating data across the *links* as well as in the storage and switching functions within the *routers* [9]. Researchers have shown that almost 46% of the router power was consumed by the input buffers and 54% of the router area was dominated by the crossbar [11]. With the increasing need for low-power architectures, these power consumption and chip area trends for NoCs have initiated several research efforts into optimizing the buffer design [6, 8], minimizing the crossbar power [10, 12], incorporating topological [12, 13, 14, 15] and routing optimizations [16], and improving network performance [17, 18, 19, 20].

It is well known that the input buffers account for significant router power budget and chip area[8, 11]. Reducing the number of input buffers to reduce the power consumption and area overhead degrades the network performance as the performance and flow control are primarily characterized by the input buffers [21]. Wormhole switching [22] allowed the flits (the basic unit of flow control, a packet consists of several flits) of the same packet to be in several routers and alleviated the need for large buffers. Virtual Channel (VC) flow control [23] decoupled the channel state from the channel bandwidth and organized the input buffer as several independent flit buffers allocated to different packets, alleviating the head-of-line (HoL) blocking effects. Credit-based flow control ensures that the *downstream* router has sufficient buffer resources to accept the flit from the *upstream* router, thereby preventing the packet/flit from being dropped. For power and area constrained NoC design, reducing the size of the input buffer leads to a reduction in either the number of VCs or the buffer depth, both of which are very critical for overall network performance.

Current high speed VLSI designs require repeater insertion along the wires in order to meet the stringent timing

1063-6897/08 $25.00 © 2008 IEEE
DOI 10.1109/ISCA.2008.14

requirements and overcome the quadratic increase in delay with the wire length [1, 5, 24]. Research into the optimization of these repeaters has shown that the repeaters can also be designed to sample and maintain data line voltage levels when required [25]. Therefore, with repeaters as potential storage elements, we can use them as buffers along the links by triggering a control signal at high network loads when there are no more buffers in the router.

In this paper, we propose **iDEAL**, inter-router dual-function, energy and area-efficient links for NoCs that achieve a reduction in power consumption and area overhead without significant loss in performance, by employing circuit and architectural techniques at the inter-router links and the router buffers respectively. At the links, we deploy circuit level enhancements to the existing repeaters so that they double as buffers when required. We propose two control techniques that enable the repeaters to adaptively function as buffers during congestion. While the first technique using a switched capacitor is suitable for designs with low implementation overhead, the second circuit employs a self-correcting double-sampling technique to guarantee error-free operation at high frequencies at a marginally higher power consumption. Both the proposed control techniques achieve reliable operation at variable clock speeds and consume significantly lower power compared to a conventional repeater-inserted control line, as the proposed control blocks can be disabled in the absence of congestion.

At the router buffer, we deploy architectural techniques such as static and dynamic buffer allocation to prevent performance degradation, while sustaining or improving the performance of a generic router. The static buffer management scheme decreases the network performance due to either insufficient buffers or unused buffer slots [8]. Unlike static allocation, dynamic buffer management allocates the incoming flit to any free buffer slot, leading to higher network throughput, although at the cost of higher control management. As a congestion control circuit exists from the downstream to the upstream router, flits can be transmitted aggressively without waiting for credits to return, thereby overcoming credit-loop turn-around latency and further improving the throughput.

This combination of circuit and architectural techniques in iDEAL using the adaptive dual-function links and the dynamic router buffer management allows us the flexibility of reducing the router buffer size without significantly degrading the network throughput and latency. Unlike other NoC designs where performance is improved at the cost of major changes to the router design, the changes in the proposed architecture pertain only to the input buffer and the allocation of the input buffer space to the incoming flits. Synthesized designs using the Synposys Power Compiler in the 90nm technology at 500 Mhz and 1.0 V, show a power reduction of 40% and an area reduction of 41% when half the router buffers are removed. Cycle accurate network simulation on 8×8 mesh and folded torus network topologies shows only a marginal 1-3% loss in throughput. In addition with aggressive transmission of flits without credit turn-around, the throughput can be further improved by 10%.

2. RELATED WORK

As buffer management is critical to the overall network performance, research efforts have explored efficient buffer management techniques such as application-specific buffer allocation [6] and repositioning of buffers at the output or between the input and output of the crossbar [26]. Dynamically Allocated Multi-Queue (DAMQ) [27] buffers make use of link lists by fixing the number of VCs for each input port. This eases the operation for a given input port, but leads to an unacceptable three-cycle delay for every flit arrival/departure as the pointer logic has to be updated by the link lists to maintain the free list. Fully Connected Circular Buffers (FC-CBs) [28] avoid the link list approach and use registers to selectively shift some flits within the buffer. However, being fully connected, it requires $P^2 \times P$ crossbar instead of the regular $P \times P$ crossbar. Moreover, shifting the existing flit at a new flit arrival adds considerable power and area overhead.

In a non-shifting approach such as ViChaR [8], the number of VCs and the depth of each VC are dynamically adjusted based on the traffic load. As there can be as many VCs as there are flit buffers, control logic becomes complicated. Instead of v:1 (v is the number of virtual channels), the first stage arbitration logic increases to vk:1 (k is the buffer depth). While increasing the number of VCs arbitrarily can increase network throughput, it also increases latency due to higher interleaving of packets [21]. Moreover, it has been shown in [29], that increasing the number of VCs is beneficial for uniform traffic, while increasing the depth is beneficial for non-uniform traffic. Therefore, in the proposed work, we adopt a dynamic VC table based approach with fixed number of VCs, thereby achieving the flexibility of dynamic buffer allocation without excessive control overhead.

3. DESIGN OF ADAPTIVE DUAL-FUNCTION LINKS

3.1 Dual-function Link Implementation

In this section, we detail the implementation of the dual-function links and the associated control logic. Figure 1 shows the proposed repeater-inserted interconnect, with the conventional repeaters replaced by three-state repeaters. A single stage of the three-state repeaters comprises of a three-state repeater inserted segment along all the wires in the link. When the control input to a repeater stage is low, the three-state repeaters in that stage function like the conventional repeaters transmitting data. When the control input to the repeater stage is high, the repeaters in that stage are tri-stated and hold the data bit in position. Once congestion is alleviated, the control logic is disabled and the three-state repeaters return to the conventional mode of operation. The adaptive dual-function links hence enable a decrease in the number of buffers within the router and save appreciable power and area.

3.2 Control Block Implementation

The control block enables the three-state repeater inserted link to function as a dual-function link during congestion. A single control block is sufficient to control the functionality of all the repeaters in one stage. Thus the overhead of the control circuitry is negligible compared to the savings in power and area obtained by reducing the router buffer size. We provide two possible implementations of the control circuit considering different design requirements such as implementation complexity and reliable operation at varying frequencies. Figure 2 shows the implementation of the proposed control block using a switched-capacitor design. The

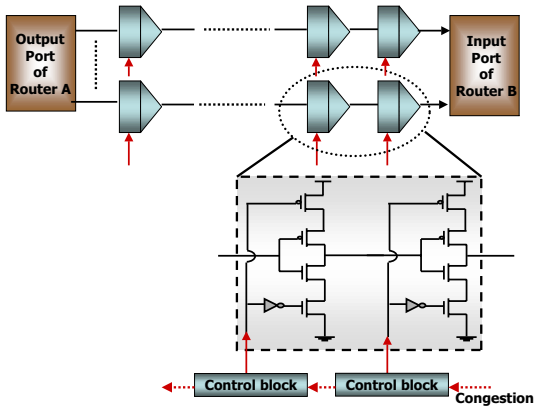

Figure 1: The proposed inter-router link using three-state repeaters that function as channel buffers during congestion.

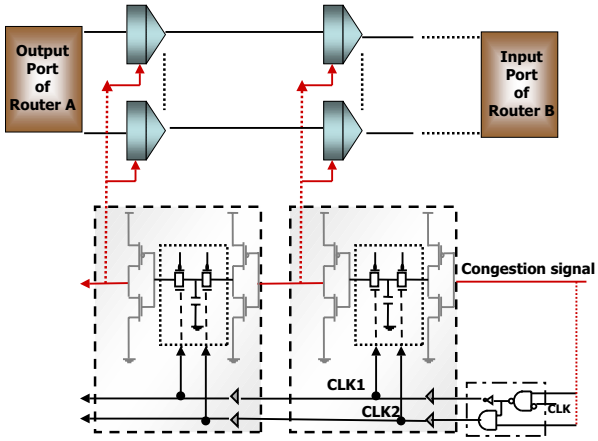

Figure 2: Proposed switched-capacitor based control block.

capacitor charges and discharges through the pass transistors that are controlled by the clock. The switched capacitor circuit delays the incoming congestion signal by one clock cycle. Though this circuit offers a low implementation overhead, it does not have an error-recovery mechanism in case of timing errors at high frequencies. In Figure 3, the control block is implemented using a self-checking double-sampling technique that enables reliable operation at high frequencies. The incoming congestion signal is sampled by two flip-flops operating at the same clock speed. The supplement flip-flop shown in Figure 3 receives a slightly offset clock with respect to the prime flip-flop, such that the data is ensured to be correctly sampled at the offset clock edge in spite of any timing errors on the data signal. The multiplexer (MUX) selects the data from the supplement flip-flop, in the event of an error. This circuit consumes a slightly greater area and power than the circuit in Figure 2, but offers a reliable error-free operation under varying frequencies.

The proposed control block implementations in Figures 2 and 3 provide the following advantages : (1) The control circuit behaves as a delay module as well as a repeater for

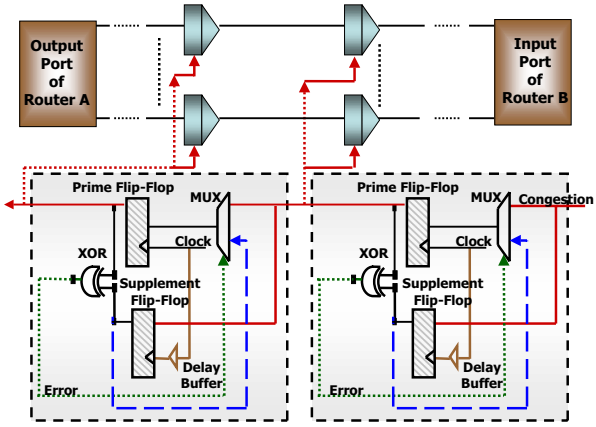

Figure 3: Proposed control block using a self-checking double-sampling technique.

the congestion signal. In addition, the control circuit shown in Figure 3 operates accurately at variable clock speeds and enables error-recovery in case of timing errors. (2) The control block can be turned OFF by the clocking circuitry when there is no congestion, thus reducing the power consumption along the congestion control line.

Figure 4 illustrates the data-flow control along the link using four repeater stages and corresponding control blocks. During cycle 1, the incoming congestion signal causes the data bit to be held by the zeroth repeater stage, while the remaining stages function as conventional repeaters. After a one clock-cycle delay in the control block, the congestion signal travels to the next stage in cycle 2 and causes it to hold the data bit in position. The remaining two stages still continue to function as conventional repeaters. Cycle 3 shows the congestion-release signal arriving at the zeroth stage. This causes the data in that stage to be output while the congestion signal travels to the second stage and causes it to hold the data. Thus the three-state repeaters are successively switched to function as link buffers during congestion, and then successively released to continue as repeaters once congestion is alleviated.

4. DESIGN OF ROUTER BUFFER

4.1 NoC Router Architecture

In packet-switched NoCs, every processing element (PE) is connected to a NoC component (router), with most NoCs commonly adopting network topologies such as mesh, or folded torus for regularity and modularity[16, 21, 26, 28] as shown in Figure 5(a). In wormhole switching, each packet that arrives on the input port progresses through router pipeline stages (routing computation(RC), virtual channel allocation (VA), switch allocation (SA), switch traversal (ST)) before it is delivered to the appropriate output port [21]. At each intermediate router, only the header flit of every packet is responsible for the first two pipeline stages of RC and VA, where as individual flits arbitrate for the SA stage. Each router pipeline stage requires a single clock cycle for every operation. After ST, the flit is transferred on the channel between the routers in the Link Traversal (LT) stage.

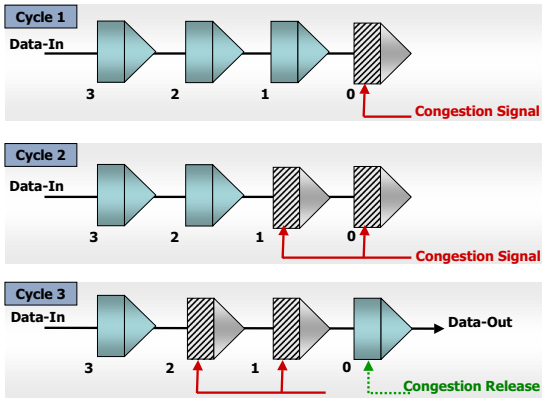

Figure 4: Data-flow control in the repeater stages during congestion.

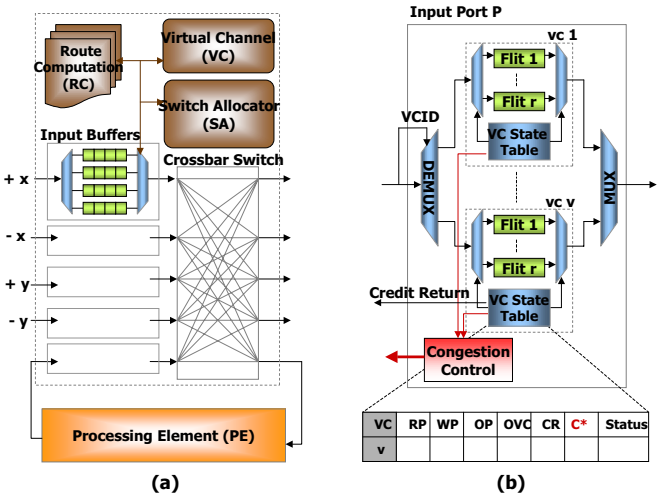

Figure 5: (a) A generic 5×5 NoC router architecture (b) The proposed static buffer allocation with congestion control.

4.2 Static Allocated Router Buffers

The proposed statically allocated router buffer design with congestion control is shown in Figure 5(b). Router buffers can be implemented as either SRAMs (Static Random Access Memory) or as FIFO (First-In-First-Out) shift registers. FIFO registers are better suited for power-constrained area-efficient NoC architectures [30] as SRAMs require additional area for the address decoding logic and involve higher switching activity during memory accesses. Hence a parallel FIFO implementation is used in iDEAL architecture as shown in Figure 5(b).

For a router architecture with P ports, v VCs/port and r flit buffers/VC the total number of buffers/port is $z = vr$. Each input VC is associated with a VC state table [8, 21]. It maintains the state for each incoming packet and ensures that the body flits are routed to the correct output port. The VCID (VC Identifier) of the incoming flit allows the DEMUX to switch to the correct input VC. The RP (read pointer) and the WP (write pointer) are used to read the flit into the buffer and write the flit out to the crossbar. The RP points to the next flit to be transmitted and WP points to a null pointer, indicating an empty flit to write the incoming data. OP (output port) is provided by the RC stage, OVC (output VC) is provided by the VA stage. CR (credits) indicates the total amount of storage available at the downstream router. Given that each VC has r credits, for every flit transmitted to the downstream router, a credit is consumed. Status field at the end indicates the current status of the VC - idle, waiting, routing, VA, SA, ST, and others. When the RP reads a flit out of the buffer, a credit is returned to the upstream router to indicate that it can send another flit.

In the generic NoCs design, the total number of input buffers is vr per input port. With the wires doubling as buffers, we have additional c buffers in the channel. Therefore, the total storage now available becomes $vr + c$. The number of credits available at each VC is $(vr + c)/z$. This allows routers to send additional flits into the network, even if the storage is in the channel, instead of the router buffer. Other than the congestion control unit, all other functionalities are identical to the generic router architecture. Every VC state table maintains another field C^* which indicates congestion. As the buffer implemented is a FIFO buffer, if the WP does not point to a null buffer, and WP = RP, then the C^* field is set. This causes the congestion control to be activated which in turn holds the data in the network channel itself. When a flit is read from the buffer, RP moves to the next buffer, clears congestion C^* field, which in turn allows data flits to enter into the router.

From the perspective of implementation, this nominal change does not impact the design of the network router architecture. Moreover, significant power savings and area gain can be obtained. However, from the perspective of performance, this design leads to head-of-line (HoL) blocking in the channel buffers at high network load. When the congestion field C^* is set for a particular VC, the corresponding flits are held in the network channel. These flits block the flits headed towards other VCs, although the other VCs may have their C^* field cleared. Therefore unavailability of buffers in any one of the VCs causes flits headed to all other VCs to be blocked. A more attractive alternative is dynamic allocated router buffers as explained in the next section.

4.3 Dynamic Allocated Router Buffers

In designing dynamically allocated router buffers, our goal is to maximize the throughput of the network without increasing the router latency. Link list [27] and circular buffers [28] have either the latency penalty or the crossbar scaling issue. As ViChaR's [8] table based approach had solved issues pertaining to latency, we have adopted a similar idea but limited the number of VCs to prevent excess control overhead.

Figure 6 explains the dynamically allocated router buffer proposed for iDEAL. We adopt the unified buffer architecture and augment the architecture with a 'Unified VC State Table' (UVST). In this case, there are v VCs/port, z buffer slots/port and c channel buffers, with r approximately z/v. This state table is simply an extension of the VC state table of the generic case. This unified state table is comparable in size to the generic case. Given the resource-constrained environment for NoCs, the size of this table is minimal and

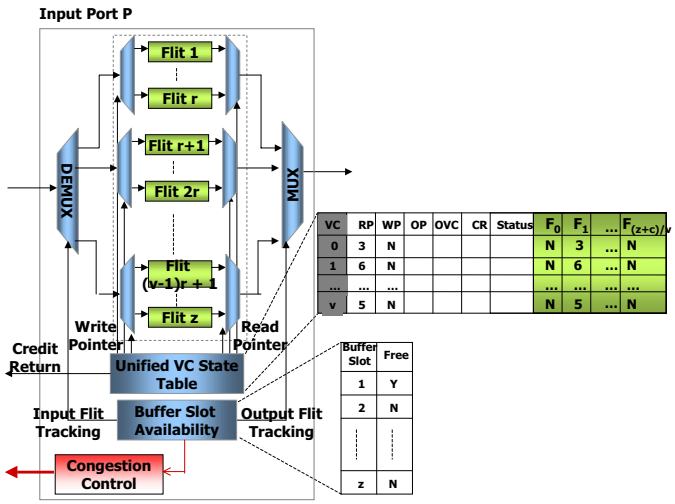

Figure 6: The proposed dynamic buffer allocation with congestion control.

does not grow with the number of VCs. The maximum size of the unified VC table is $O(v)$ as compared to the ViChaR which is $O(vr)$. When a new flit arrives, its VCID cannot be used to switch as all buffer slots are unified. For that purpose, we use the 'Buffer Slot Availability' (BSA) tracking system. BSA allocates/deallocates arriving/departing flits with buffer slots. Therefore, the DEMUX switches to the buffer slot provided by the BSA at the input flit tracking. BSA keeps track of all buffer slots currently available and allocates the first buffer slot found to be free. If the buffer slot number points to NULL, then such a slot can be selected for the newly arriving flit. After allocating the buffer slot to the incoming flit, BSA then searches for the next free slot to be allocated. Similarly, for a departing flit, BSA will de-allocate the buffer slot using the output flit tracking and add the free slot to the list of free slots maintained in the table (shown in the inset of Figure 6).

Once the flit is associated with the input flit tracking number identifying which flit buffer it is destined to, the flit now arrives at the second DEMUX. Here, the WP logic writes the flit to the buffer slot allocated by the BSA. In the same cycle, UVST identifies the VCID of the newly arriving flit and accordingly updates the UVST. If the newly arriving flit is the header flit, then it will undergo the usual stages of RC, VA, SA, and ST. The arbitration logic (v:1 at the input and Pv:1 at the output) is similar to the generic case as there is no increase in the number of VCs. The table contains buffer slots F_0, F_1, ... $F_{(z+c)/v}$ in addition to the regular fields of RP, WP, OP, OVC, CR and Status fields. The total number of credits is limited to $(z+c)/v$ per VC slot when used without speculation. The buffer slots are used to identify the location of the flit assigned to the particular VC. The number of buffer slots available depends on the maximum number of credits available for a particular VC. For fairness purposes, the number of credits is equally divided between all the different VCs. The responsibility for congestion detection rests with the BSA. When BSA finds only a single non-null pointer in its base table, it will trigger the congestion signal. To determine whether the input buffers are full, a small counter that counts the number of free slots is maintained and when this counter reaches one, we trigger the congestion signal. Similarly a departing flit will create a free buffer slot releasing the congestion signal. A single buffer slot combined with a dynamic spare VC for every output port can be maintained to ensure deadlock recovery[8].

In the proposed aggressively speculated iDEAL architecture, the number of credits is doubled to $2(z+c)/v$ per VC slot. In the generic case, this aggressive speculation will trigger dropped flits due to lack of buffer storage. However, in iDEAL, when there is no buffer storage available, the upstream router will feel the backpressure from the downstream router due to the congestion signal. This will prevent the upstream router from transmitting any further flits. Therefore, we can overcome the credit turn-around time without having to sacrifice throughput in iDEAL. Further increasing the number of credits arbitrarily can lead to deadlocks as packets from a single VC may then occupy the channel buffers, thereby blocking other flits/packets.

In iDEAL, HoL blocking in the channel buffers due to static buffer allocation can be overcome by dynamic buffer allocation based on a table approach as described above. These channel buffers can be viewed as serial FIFO buffers as opposed to the parallel FIFO buffers used within the routers. Therefore, eliminating the HoL blocking is critical in iDEAL. Moreover, the throughput of the network can be further increased by aggressive speculation as explained before. Static allocation of buffer slots simplifies the overall design as it requires minimum extension over a generic NoC router architecture. Dynamic allocation of buffer slots significantly reduces the HoL blocking. This achieves much higher throughput while saving in chip area and reducing the power consumption.

5. PERFORMANCE EVALUATION

In this section, we evaluate the router buffers and the proposed dual-function links in terms of power dissipation, area overhead and overall network performance. We consider 8 × 8 mesh and folded torus topologies with 4-stage pipelined router design. Each router has $P = 5$ input ports (4 for each direction and 1 for the PE). The baseline design considered has 4 VCs per input port, with each VC having 4 flit buffers in the router, for a total of 80 flit buffers ($= 5 \times 4 \times 4$). Each packet consists of 4 flits and each flit is 128 bits long. For the design with the adaptive dual-function links, we consider 5 different cases where some or all of the repeaters along the link are replaced by the link buffers. The notation followed for the different cases is of $vn_V - rn_R - cn_C$, where n_V is the number of VCs per input port, n_R is the number of router buffers per VC and n_C is the number of link buffers. For example, the baseline is denoted as $v4 - r4 - c0$, implying 4 VCs per input port, 4 router buffers per VC and 0 link buffers. For a fair comparison with the baseline, the number of buffers eliminated from the router is added to the set of link buffers. In each case, the design is implemented in Verilog and synthesized using the Synopsys Design Compiler tool and the TSMC 90 nm technology library. The power dissipation and area overhead in the links and the router are obtained for each case at a supply voltage of 1 V and an operating frequency of 500 MHz.

5.1 Power and Area Estimation for the Inter-router Links

The power per segment of the repeater-inserted link is given by

$$P_{segment} = P_{dynamic} + P_{leakage} + P_{short-ckt} \quad (1)$$

where $P_{dynamic}$ is the switching power, $P_{leakage}$ is the power due to the subthreshold leakage current and $P_{short-ckt}$ is the power due to the short-circuit current. The power per segment is multiplied by the number of segments and the link width to obtain the total link power dissipation for a flit traversal. When a conventional repeater is replaced by a three-state repeater, there is an additional capacitance due to the added transistors, as shown in Figure 1. The increase in the switching capacitance increases the total power consumed by the links. Power is also dissipated in the control blocks controlling the repeater stages, when they are enabled during congestion.

In calculating the power values, the inter-router links are assumed to be 2 mm long for the mesh network. The average channel length doubles in case of the folded torus network [12] and hence the inter-router links are 4 mm long. In the baseline design, there are 8 optimally spaced conventional repeaters along each wire of the 128-bit wide links. The total power consumed by the link per flit traversal is 2.45 mW for the 8×8 mesh and 3.94 mW for the 8×8 folded torus. When all the 8 conventional repeaters are replaced by channel buffers, the total power consumed in the link for every flit traversal is found to be 3.55 mW for the mesh and 5.04 mW for the folded torus. In the presence of congestion, the power dissipated by the switched-capacitor control block is found to be 2.089 μW and by the control block with double sampling technique is found to be 6.1 μW. The additional control logic thus consumes only a small fraction of the total power dissipated in the inter-router links. The repeaters and the wires utilize different metal layers and their area overheads are independent of each other [31]. The area consumed by the repeater stages along a 128-bit link is found to be 32 μm^2 in case of the baseline and 80 μm^2 when all the 8 conventional repeaters along the link are replaced by three-state repeaters.

5.2 Power and Area Estimation for the Router

This section summarizes the power estimation for the buffers, the crossbar and the arbiter in the router. The router buffers are implemented as FIFO registers with the associated control logic. The control logic maintains the read/write pointers that select the appropriate signals from an input demultiplexer and an output multiplexer. When the number of VCs or the buffer depth per VC is changed, the size and number of components within the buffer changes, altering the power consumption and area. Considering both the write and read operations in the buffer, the total power (including the dynamic and the leakage power) consumed for a 128-bit flit in the buffer is estimated to be 19.54 mW, for the baseline design with 16 buffer slots. Decreasing the buffer size by 4 buffer slots (25%) leads to a power savings of 25.74% compared to the baseline. Power reduces by 40.78% when the buffer size is reduced to 50% of the baseline.

A two-stage matrix arbiter design [32] is considered with the first stage selecting one output from the v VCs of a port and the second stage arbitrating among the Pv inputs from each of the P ports. In case of 4 VCs, the two-stage arbiter consumes a power of 0.15 mW, for a single arbitration task. When the number of VCs is decreased to 3, the power consumed by the arbiter reduces to 0.09 mW per arbitration. The switch in the router consumes 0.31 mW per flit traversal, in case of the design with 4 VCs per port and 0.27 mW per flit traversal in the case of 3 VCs per port. The area of the router buffers, arbiter and the switch are obtained from the synthesized designs using the Synopsys Design Compiler tool and the TSMC 90 nm technology library. In case of the baseline design with 16 buffer slots in the router, the buffer area is 81,407 μm^2. A 50% decrease in the buffer size leads to a 40.95% reduction in the buffer area.

5.3 Comparison of the different cases

Table 1 shows a comparison of the power estimations for various link and router buffer configurations. The first configuration shown is the baseline case and uses no link buffers. Change in power in each of the other cases is expressed as a percentage increase (+) or a percentage decrease (−) with respect to the baseline. As the number of router buffers is decreased, the power consumed by the buffer per flit reduces significantly. A maximum power savings of 31.15% is achieved for the third case that uses only half the number of router buffers compared to the baseline. The last case shown replaces only a single router buffer with a buffer along the link and does not achieve significant power savings compared to the baseline.

5.4 Simulation Methodology

A cycle-accurate on-chip network simulator was used to conduct a detailed evaluation of the proposed link and router buffer design in both a 8×8 mesh and a 8×8 folded torus networks. The test configurations are represented in the results as $vn_V - rn_R - cn_C$, where n_V is the number of VCs per input port, n_R is the number of router buffers per VC and n_C is the number of link buffers. For simplicity, they will be referred to as $n_V - n_R - n_C$ in the following discussion. The test configurations evaluated were $n_V - n_R - n_C$ = 4-4-0 (baseline), 4-3-4, 4-2-8, 3-4-4, 3-3-7 and 5-3-1. For synthetic traffic patterns, packets were injected according to Bernoulli process based on the network load for a given simulation run. The network load is varied from $0.1 - 0.9$ of the network capacity. The simulator was warmed up under load without taking measurements until steady state was reached. Then a sample of injected packets were labelled during a measurement interval. The simulation was allowed to run until all the labelled packets reached their destinations. For SPLASH-2 suite benchmarks [33], the network traces with precise timing information were gathered by running the benchmarks on RSIM [34] for 64 nodes and then simulated on our proposed cycle accurate network simulator.

We tested our hypothesis of using static and dynamic buffer allocation schemes on several traffic patterns such as: (1) Uniform Random, where each node randomly selects its destinations with equal probability and (2) Permutation Patterns, where each node selects a fixed destination based on the permutations. We evaluated the performance on the following permutation patterns: Bit-Reversal, Butterfly, Matrix Transpose, Complement, Perfect Shuffle, Neighbor and (3) SPLASH-2 suite benchmarks covering a spectrum of memory sharing and access patterns[33]. These include FFT with input data set $64K$ points; LU with 256×256,16×16

Table 1: Power Estimation for Various Link and Router Buffer Configurations in a 8 × 8 mesh and 8 × 8 folded torus interconnection networks. Power values are for one flit traversal. n_V is the number of VCs per input port, n_R is the number of router flit-buffers per VC and n_C is the number of link buffers.

vn_V- rn_R- cn_C	Buffer Power (mW)	Mesh Link+ Control Power (mW)	Folded Torus Link+ Control Power (mW)	Mesh Total Buffer + Link Power (mW)	% Change	Folded Torus Total Buffer + Link Power (mW)	% Change
v4-r4-c0	19.54	2.45 + 0	3.94 + 0	21.99	–	23.48	–
v4-r3-c4	14.51	2.90 + 0.012	4.39 + 0.012	17.42	-20.78	18.91	-19.46
v4-r2-c8	11.57	3.55 + 0.020	5.04 + 0.020	15.14	-31.15	16.63	-29.17
v3-r4-c4	15.09	2.90 + 0.012	4.39 + 0.012	18.00	-18.14	19.49	-16.99
v3-r3-c7	12.56	3.49 + 0.018	4.98 + 0.018	16.06	-26.96	17.55	-25.25
v5-r3-c1	19.29	2.81 + 0.005	4.28 + 0.005	22.10	+0.50	23.57	+0.03

block; mp3d with 48000 molecules; Radix with $1M$ integers, 1024 radix and Water-nsquared with 512 molecules.

5.5 Simulation Results and Discussion

We evaluate the proposed iDEAL architecture in terms of the input buffer power consumed, saturation throughput achieved, average latency and the overall power consumed by the network. The following discussion presents the simulation results for the individual cases as well as a comparison of the throughput and buffer power for all the cases considered.

Input Buffer Power: Figures 7(a) and 7(b) show the total power (both the dynamic and the leakage power) dissipated in the input buffers for the uniform traffic pattern in the 8 × 8 mesh and the 8 × 8 folded torus networks respectively, at a network load of 0.5. For the mesh topology, the power savings in the 4-3-4 configuration using dynamic buffer allocation is nearly 24% as shown in Figure 7(a). The power savings for the 4-2-8 configuration (reducing the buffer depth from 4 to 2 per VC) is about 40%. The 3-4-4 configuration shows 22% savings in buffer power alone. Similar results are observed for the folded torus topology with the 4-2-8 configuration achieving a power savings of almost 39% as shown in Figure 7(b). Therefore significant power savings is obtained in all the cases by reducing the buffer size.

Throughput, Latency and Power: Figure 8 shows the saturation throughput, average latency and overall network power for uniform traffic at varying network load, for the 8 × 8 mesh and the 8 × 8 folded torus networks. From Figure 8(a), for the mesh topology, the saturation throughput shows almost similar performance for 4-4-0, 3-4-4 and 4-3-4. The decrease in the number of VCs for the 3-4-4 or the buffer depth for the 4-3-4 do not significantly affect the throughput. The more interesting point is 4-2-8 which shows only about 4% drop in performance. This result is significant as we can save almost 42% of the buffer size and yet achieve similar performance as the baseline configuration by dynamically allocating the buffer resources to flits. The additional buffers along the link ensure that the flow of data flits is not hampered at high network loads even though there are fewer buffers in the router. For the folded torus topology, Figure 8(d) shows that the saturation throughput is almost unaffected for all cases.

The average latency plots shown in Figures 8(b) and 8(e) indicate that the network saturates at about 0.3 for the mesh topology and at about 0.35 for the folded torus topology. The total power consumed in the network, including the buffer, arbiter, switch, links and control blocks, is shown in Figures 8(c) and 8(f) for a network workload of 0.5. The plots indicate that the buffers and the links account for a significant fraction of the total power. The power dissipated in the control blocks is negligible and is not visible at the scale considered in the plots. For both the mesh and the folded torus topologies, the 4-2-8 case shows about 30% decrease in the total network power. The 4-3-4 and 3-4-4 configurations show a reduction of almost 20% of the network power while the 3-3-7 configuration shows the network power reducing by almost 27%. All the configurations achieve a reduction in power compared to the baseline, by reducing the router buffer size.

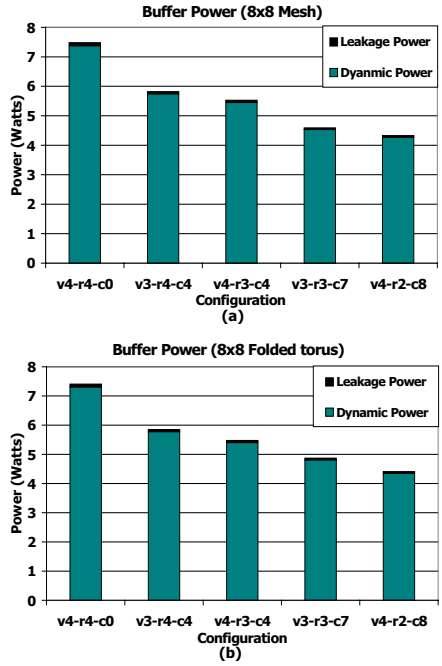

Figure 7: Buffer power for 8 × 8 mesh and folded torus networks under Uniform traffic, at network load = 0.5. In the notation $vn_V - rn_R - cn_C$, n_V is the number of VCs per input port, n_R is the number of router buffers per VC and n_C is the number of link buffers.

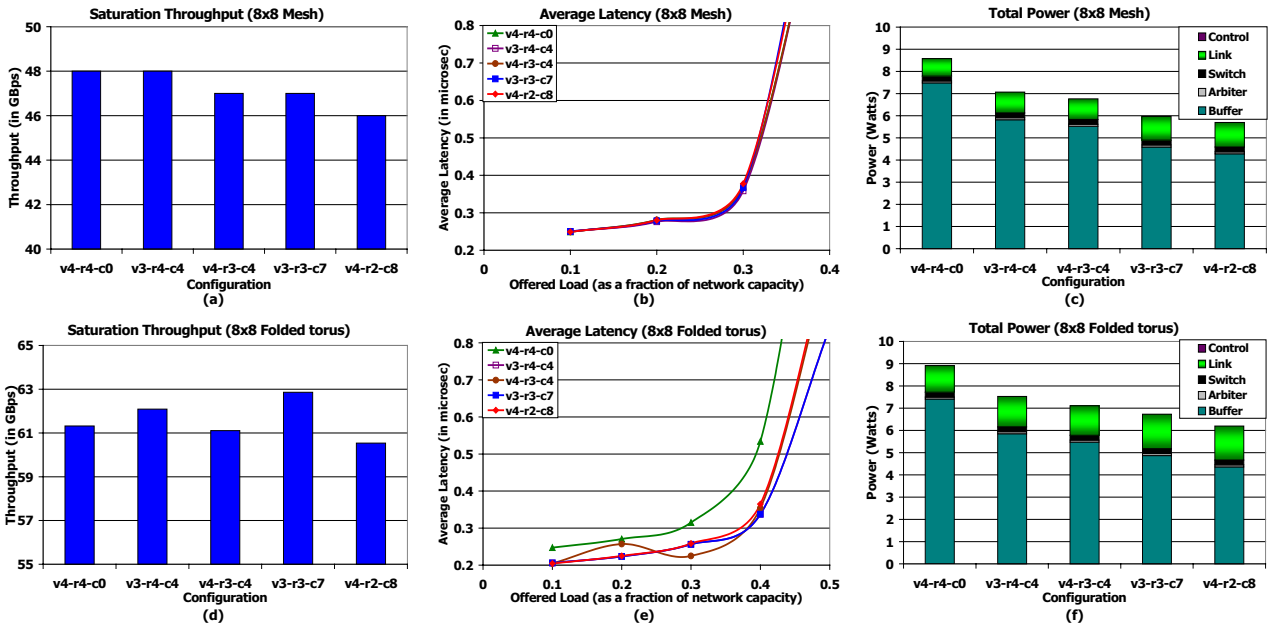

Figure 8: Saturation Throughput, Average Latency and Overall network power (at network load = 0.5) under Uniform Traffic for 8 × 8 mesh and folded torus networks.

Comparison of Throughput: Figure 9(a) shows a comparison of 4-3-4 and the 4-2-8 schemes of the iDEAL architecture, the Fully-Connected Circular Buffer (FC-CB) technique [28] and the Dynamically Allocated Multiqueue (DAMQ) buffers [27]. The FC-CB shows a similar performance as the dynamically allocated 4-4-0 configuration. The 4-3-4 and the 4-2-8 cases achieve about 4% improvement in throughput over the FC-CBs, as the FC-CB design involves a significant latency in shifting flits through the circular buffer at every new flit arrival. Compared to the DAMQ architecture, the 4-3-4 achieves about 12.5% improvement in saturation throughput. The DAMQ requires a three-cycle delay to update the pointers at every flit arrival. Therefore the iDEAL architecture achieves significant improvement in performance over both the FC-CB and the DAMQ designs.

Throughput and Power using Aggressive Speculation: Figures 9(b) and 9(c) show the saturation throughput and overall network power using an aggressive speculation technique, for the 8 × 8 folded torus network under uniform traffic at varying network load. The number of credits available to the upstream router is speculatively increased to 8 as the congestion control circuit enables an aggressive flit transmission without waiting for the credits from the downstream router. Figure 9(b) shows that the saturation throughput for the 4-2-8 improves by about 10% compared to the baseline. Therefore the speculative flow control technique improves performance improvement for the iDEAL architecture along with significant power and area savings.

Throughput and Power for All Synthetic Traffic Patterns and SPLASH-2 suite benchmarks: Figure 10(a) shows the power consumed at the input buffers and Figure 10(c) shows the throughput achieved at a network load of 0.5 for the 8 × 8 mesh network, with static and dynamic buffer allocation for all traffic patterns including Uniform (UN), Complement (CO), Perfect Shuffle (PS), Butterfly (BU), Bit Reversal (BR), Matrix Transpose (MT), Neighbor (NE) and Tornado (TO) for 3 configurations, namely 4-4-0, 4-3-4 and 4-2-8. Power savings is obtained for both the 4-3-4 and the 4-2-8 cases under all the traffic patterns. For the Complement traffic pattern, static buffer allocation provides 57% savings in buffer power for the 4-2-8 configuration as compared to the baseline, where as with dynamic buffer allocation, the savings decreases to 40%. From Figure 10(c), there is no appreciable decrease in throughput for the dynamic case for all traffic patterns. Dynamic buffer allocation provides the flexibility for the flits to be allocated to any available buffer slot and is not as restrictive as the static allocation.

Figures 10(b) and 10(d) show the normalized power and normalized execution time respectively, for the selected SPLASH-2 suite benchmarks for 4-4-0, 4-3-4 and 4-2-8 configurations with dynamic buffer allocation. The normalization is carried out with respect to the baseline 4-4-0 configuration. From Figure 10(b), the power savings from the 4-3-4 and 4-2-8 configurations are 20% and 30% respectively. From Figure 10(d), the 4-3-4 and 4-2-8 configurations do not show significant drop in performance, in fact the drop is less than 1%. Therefore, dynamic allocation with link buffers does not degrade performance and provides significant power savings for all SPLASH-2 benchmarks.

6. CONCLUSION

As recent research has shown, the major issue facing on-chip networks is the ever increasing power consumption. iDEAL proposes to reduce the number of buffers within the router, thereby achieving a significant savings in power and area. As this impacts performance, we provide dual-function links which can be used for storage when required. Simulation results show that by reducing the router buffer size in half, iDEAL achieves nearly 40% reduction in buffer power

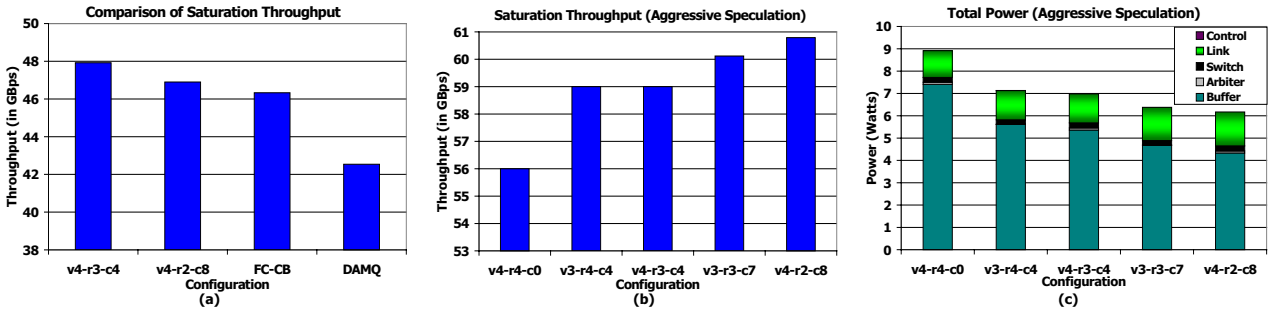

Figure 9: (a) Throughput Comparison for the 4-3-4 and 4-2-8 iDEAL configurations, Fully Connected Circular Buffers (FC-CB) and Dynamically Allocated Multiqueue (DAMQ) Buffers, under Uniform traffic for 8 × 8 mesh network. (b) Throughput and (c) Overall network power (at network load = 0.5) for Aggressive Speculation using 8 credits, under Uniform traffic for 8 × 8 folded torus network.

alone and more than 30% savings in the overall network power. The dynamically assigned buffers with aggressive speculative flow control show up to 10% improvement in performance, dynamically assigned buffers without speculation show a marginal 1-3% drop in performance and statically assigned router buffers show a 10 - 20% drop in the network performance. This paper shows that eliminating some of the buffers in the router and using adaptive link buffers saves an appreciable amount of power and area, without significant degradation in the throughput or latency.

Acknowledgements: We thank Dr. Dongsheng (Brian) Ma and Minkyu Song for their assistance with the switched capacitor control block. We also acknowledge the anonymous reviewers for their insightful comments. This research was partially supported by NSF grants CCR-0538945 and ECCS-0725765.

7. REFERENCES

[1] R. Ho, K. W. Mai, and M. A. Horowitz, "The Future of Wires," *Proceedings of the IEEE*, vol. 89, pp. 490–504, April 2001.

[2] L. Benini and G. D. Micheli, "Networks on Chips: A New SOC Paradigm," *IEEE Computer*, vol. 35, pp. 70–78, 2002.

[3] W. J. Dally and B. Towles, "Route packets, not wires : On-Chip Interconnection Networks," in *Proceedings of the Design Automation Conference (DAC)*, Las Vegas, NV, USA, June 18-22 2001.

[4] R. Kumar, V. Zyuban, and D. Tullsen, "Interconnections in Multi-core Architectures: Understanding Mechanisms, Overheads and Scaling," in *Proceedings of the 32nd Annual International Symposium on Computer Architecture (ISCA)*, Madison, Wisconsin, USA, June 4-8 2005, pp. 408–419.

[5] S. Heo and K. Asanovic, "Replacing Global Wires With an On-chip network: A Power Analysis," in *Proceedings of the International Symposium on Low Power Electronics and Design (ISLPED)*, San Diego, CA, USA, August 8-10 2005, pp. 369–374.

[6] J. Hu and R. Marculescu, "Application-specific Buffer Space Allocation for Network-on-chip Router Design," in *Proceedings of the IEEE/ACM International Conference on Computer Aided Design (ICCAD)*, San Jose, CA, USA, November 7-11 2004, pp. 354–361.

[7] T. Bjerregaard and S. Mahadevan, "A Survey of Research and Practices of Network-on-chip," *ACM Computing Surveys*, vol. 38, no. 1, 2006.

[8] C. A. Nicopoulos, D. Park, J. Kim, N. Vijaykrishnan, M. S. Yousif, and C. R. Das, "ViChaR: A Dynamic Virtual Channel Regulator for Network-on-chip Routers," in *Proceedings of the 39th Annual International Symposium on Microarchitecture (MICRO)*, Orlando, FL, USA, December 9-13 2006, pp. 333–344.

[9] J. D. Owens, W. J. Dally, R. Ho, D. N. Jayasimha, S. W. Keckler, and L. S. Peh, "Research Challenges for On-chip Interconnection Networks," *IEEE Micro*, vol. 27, no. 5, pp. 96–108, September-October 2007.

[10] H. S. Wang, L. S. Peh, and S. Malik, "Power-driven Design of Router Microarchitectures in On-chip Networks," in *Proceedings of the 36th Annual ACM/IEEE International Symposium on Microarchitecture*, Washington DC, USA, December 03-05 2003, pp. 105–116.

[11] P. Kundu, "On-die Interconnects for Next Generation CMPs," in *2006 Workshop on On- and Off-Chip Interconnection Networks for Multicore Systems*, Stanford, CA, USA, December 6-7 2006.

[12] J. Balfour and W. J. Dally, "Design Tradeoffs for Tiled CMP On-chip Networks," in *Proceedings of the 20th ACM International Conference on Supercomputing (ICS)*, Cairns, Australia, June 28-30 2006, pp. 187–198.

[13] J. Kim, W. J. Dally, and D. Abts, "Flattened Butterfly : A Cost-efficient Topology for High-radix Networks," in *Proceedings of the 34th Annual International Symposium on Computer Architecture (ISCA)*, San Diego, CA, USA, June 9-13 2007, pp. 126–137.

[14] D. Park, R. Das, C. Nicopoulos, J. Kim, N. Vijaykrishnan, R. Iyer, and C. Das, "Design of a Dynamic Priority-based Fast path Architecture for On-chip Interconnects," in *15th IEEE Symposium on High Performance Interconnects (HOTI 2007)*, Stanford, CA, USA, August 6-8 2007, pp. 15–20.

[15] P. P. Pande, C. Grecu, A. Ivanov, and R. Saleh, "Performance Evaluation and Design Trade-offs for Network-on-chip Interconnect Architectures," *IEEE Transactions on Computers*, vol. 54, no. 8, August 2005.

[16] J. Hu and R. Marculescu, "DyAD Ű Smart Routing for Networks-on-chip," in *Proceedings of the 41st IEEE/ACM Design Automation Conference*, San Diego, CA, USA, June 7-11 2004.

[17] J. Kim, C. A. Nicopoulos, D. Park, N. Vijaykrishnan, M. S. Yousif, and C. R. Das, "A Gracefully Degrading and Energy-efficient Modular Router Architecture for On-chip Networks," in *Proceedings of the 33rd Annual International Symposium on Computer Architecture (ISCA)*, Boston, MA, USA, June 17-21 2006, pp. 4–15.

[18] P. Abad, V. Puente, J. A. Gregorio, and P. Prieto, "Rotary Router: An Efficient Architecture for CMP Interconnection Networks," in *Proceedings of the 34th Annual International Symposium on Computer Architecture (ISCA)*, San Diego, CA, USA, June 9-13 2007, pp. 116–125.

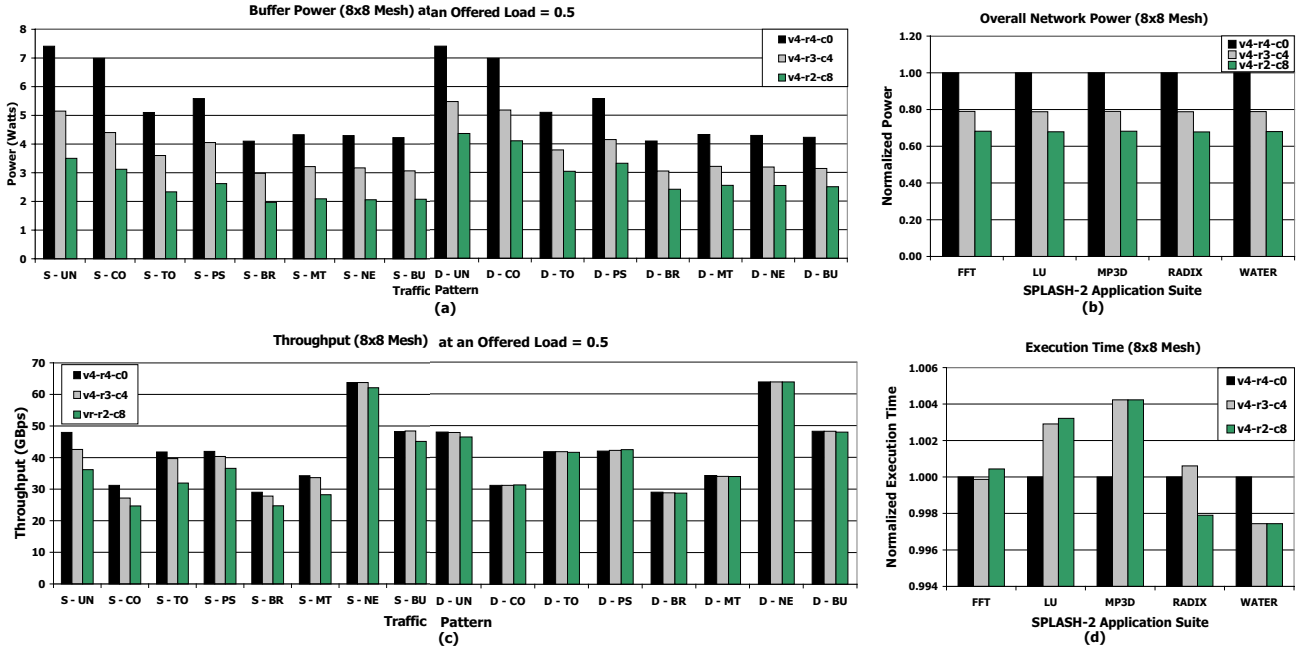

Figure 10: (a) Buffer power for synthetic traffic (b) Normalized Overall network power for SPLASH-2 benchmarks (c) Throughput for synthetic traffic (d) Normalized Execution time for SPLASH-2 benchmarks. Synthetic traffic patterns (Uniform (UN), Complement (CO), Tornado (TO), Perfect Shuffle (PS), Bit-Reversal (BR), Matrix Transpose (MT), Neighbor (NE) and Butterfly (BU)) are considered under both static (S) and dynamic (D) buffer allocation schemes.

[19] L. S. Peh and W.J. Dally, "A Delay Model and Speculative Architecture for Pipelined Routers," in *Proceedings of the 7th International Symposium on High-Performance Computer Architecture (HPCA)*, Nuevo Leone, Mexico, January 2001, pp. 255–266.

[20] R. Mullins, A. West, and S. Moore, "Low-latency Virtual Channel Routers for On-chip Networks," in *Proceedings of International Symposium on Computer Architecture (ISCA)*, Munchen, Germany, June 19-23 2004, pp. 188–197.

[21] W. J. Dally and B. Towles, *Principles and Practices of Interconnection Networks*, Morgan Kaufmann, San Fransisco, USA, 2004.

[22] W. J. Dally, "Performance Analysis of k-ary n-cube Interconnection Networks," *IEEE Transactions on Computers*, vol. 39, no. 6, pp. 775–785, June 1990.

[23] W. J. Dally, "Virtual-channel Flow Control," in *Proceedings of the 17th Annual International Symposium on Computer Architecture (ISCA)*, Seattle, WA, USA, June 1990, pp. 60–68.

[24] K. Banerjee and A. Mehrotra, "A Power-optimal Repeater Insertion Methodology for Global Interconnects in Nanometer Designs," *IEEE Transactions on Electron Devices*, vol. 49, no. 11, pp. 2001–2007, November 2002.

[25] M. Mizuno, W. J. Dally, and H. Onishi, "Elastic Interconnects: Repeater-inserted Long Wiring Capable of Compressing and Decompressing Data," in *Proceedings of the IEEE International Solid-State Circuits Conference*, San Fransisco, CA, USA, February 5-7 2001, pp. 346–347.

[26] Y. M. Boura and C. R. Das, "Performance Analysis of Buffering Schemes in Wormhole Routers," *IEEE Transactions on Computers*, vol. 46, pp. 687–694, 1997.

[27] Y. Tamir and G. L. Frazier, "High-performance Multiqueue Buffers for VLSI Communication Switches," in *Proceedings of the 15th Annual International Symposium on Computer Architecture (ISCA)*, Honolulu, Hawaii, USA, May-June 1988, pp. 343–354.

[28] N. Ni, M. Pirvu, and L. Bhuyan, "Circular Buffered Switch Design with Wormhole Routing and Virtual Channels," in *Proceedings of the International Conference on Computer Design (ICCD)*, Austin, TX, USA, October 1998, pp. 466–473.

[29] M. Rezazad and H. Sarbazi-azad, "The Effect of Virtual Channel Organization on the Performance of Interconnection Networks," in *Proceedings of the 19th International Parallel and Distributed Processing Symposium*, Denver, CO, USA, April 3-8 2005.

[30] J. Hu and R. Marculescu, "Energy-aware Mapping for Tile-based NoC Architectures Under Performance Constraints," in *Proceedings of the 2003 Conference on Asia South Pacific Design Automation*, Kitakyushu, Japan, January 21-24 2003, pp. 233–239.

[31] M. A. El-Moursy and E. G. Friedman, "Optimum Wire Sizing of RLC Interconnect with Repeaters," *Integration, the VLSI Journal*, vol. 38, no. 2, pp. 205–225, December 2004.

[32] H. S. Wang, X. Zhu, L. S. Peh, and S. Malik, "Orion: A Power-performance Simulator for Interconnection Networks," in *Proceedings of the 35th Annual ACM/IEEE International Symposium on Microarchitecture*, Istanbul, Turkey, November 18-22 2002, pp. 294–305.

[33] C. S. Woo, M. Ohara, E. Torrie, J. P. Singh, and A. Gupta, "The SPLASH-2 programs: Characterization and Methodological Considerations," in *Proceedings of the 22nd Annual International Symposium on Computer Architecture (ISCA)*, Santa Margherita Ligure, Italy, June 22-24 1995, pp. 24–37.

[34] V. Pai, P. Ranganathan, and S.V. Adve, "RSIM Reference Manual version 1.0," *Dept. of Electrical and Computer Engineering, Rice University*, July 1997.