RapidLink: A Network-on-Chip Architecture with Double-Data-Rate Links

Anastasios Psarras*, Savvas Moisidis*, Chrysostomos Nicopoulos† and Giorgos Dimitrakopoulos*
*Electrical and Computer Engineering, Democritus University of Thrace, Xanthi, Greece
†Electrical and Computer Engineering, University of Cyprus, Nicosia, Cyprus

***Abstract*—Various state-of-the-art Network-on-Chip (NoC) architectures, employing either low- or high-radix topologies, have exploited the speed provided by on-chip wires – after appropriate wire engineering – to transfer flits over longer distances in a single clock cycle. In this work, motivated by the same principle of fast link traversal, we propose the *RapidLink* NoC architecture, which exploits said speed to rapidly transfer flits between adjacent routers (connected with links of reasonable short-to-medium length, up to a few millimeters) in *half a clock cycle*, thereby enabling Double-Data-Rate (DDR) link traversal. RapidLink can markedly increase network performance, without increasing the area/power cost of the NoC. Extensive cycle-accurate network simulations and hardware implementation results demonstrate the efficiency of RapidLink and its potential as a scalable NoC architecture.**

I. Introduction

Networks-on-Chip (NoC) have been established as the dominant communication backbone in multi-core environments, primarily due to their innate scalability attributes. On-chip network scalability is demanded not only in terms of network performance (latency and throughput), but also with respect to physical implementation traits (energy/power consumption, area cost, and place-and-route amenability).

These two facets of NoC scalability (performance and area/power of the design) are often contradictory, thereby forcing designers to typically pursue only one avenue at a time. In this paper, we exploit an inherent idiosyncrasy of NoC architectures to expose a new design opportunity that can embrace both scalability flavors. Specifically, we harness the intrinsic asymmetry between the *intra-* and *inter-*router delays encountered in modern multi-core systems. For reasonable inter-router distances of up to a few millimeters, the Link Traversal (LT) stage of a generic NoC router pipeline is substantially shorter than the delay encountered by flits within the router [1]. Researchers have capitalized on this well-documented phenomenon in a three-fold manner: (1) to allow for the traversal of multiple network hops in a single clock cycle [2], [3]; (2) to enable single-cycle crossing of multi-millimeter links employed by high-radix routers in low-diameter network topologies [4], [5]; (3) to fuse the LT stage with the switch traversal pipeline stage of the router [6].

This work presents – for the first time – an additional exploitation prospect for the short inter-router link delay. Rather than using the fast link traversal to cover longer distances in a given cycle, the proposed *RapidLink* architecture exploits said speed to *rapidly transfer flits* between adjacent routers in *half a clock cycle* and *utilizes both edges of the clock* during the sending and receiving of flits. A similar approach has been used to tackle low power operation [7]. RapidLink enables two separate data transfers to be made between two routers per cycle. As a result, each upstream/downstream router pair benefits from Double-Data-Rate (DDR) transfers, whereby two flits can be sent/received per clock cycle. The proposed approach can markedly decrease latency and increase throughput significantly, without incurring any additional hardware cost. In RapidLink, the original clock frequency of the NoC is unaffected. The only constraint is that the LT delay cannot exceed one half of the delay of the router, which is feasible for small/medium wire lengths, and after appropriate wire engineering (Section II).

The new architecture is investigated in detail and quantitatively explored to highlight its potential. Extensive cycle-accurate simulations using a variety of traffic patterns validate the efficiency of RapidLink. Furthermore, detailed hardware analysis using placed-and-routed designs in a 45 nm standard-cell library reveal that network performance is increased without increasing the hardware cost of the NoC.

II. Motivation and Key Concepts

The delay experienced during link traversal in a NoC is determined by the capacitance and the resistance of the wires of the link, and by the manner that the wires are driven. Reducing link delay can be achieved using several approaches, such as (a) promoting NoC links to upper metal layers, (b) increasing the wire spacing, (c) using wire shielding, and (d) using across-wire repeaters.

In most systems, the NoC links can be routed only in intermediate metal layers with $2\times$ and $4\times$ lower resistance than the resistance of local routing layers. Metal layers reserved for local routing are used by the processing cores and their caches, while top metal layers (with almost $8\times$ lower resistance) are primarily occupied by power and clock routing [8], [9]. Therefore, the NoC links are accommodated within certain intermediate layers that are neither too resistive, nor too dense, thus allowing for low-delay link traversal. As measured in [8], and used in a real prototype at 32 nm, the wire delay of this group of metal layers ranges from 60-300 ps/mm, depending on repeater placement and wire spacing.

Similar results have been shown by a variety of real prototypes. For example, IBM has shown that, with appropriate wire spacing and metal layer selection, wires can cross distances of up to 2.7 mm in 210 ps at 45 nm [10], while, at the same technology node, Intel drives a wire of 5.4 mm in 270 ps [11] with appropriate wire engineering. Recently, SMART [1], [12] was demonstrated to traverse 16 mm of wire (16 hops of 1 mm each) at 1 GHz, by utilizing 1-mm-spaced repeaters and $3\times$ larger wire spacing than the minimum allowed. This translates to crossing 4 mm in less than 250 ps. Similarly, NoCs designed recently with high-radix routers assume repeated wire delays

of 66 ps/mm, which are used to cross long links of 5.4 mm in a single cycle [5].

On the contrary, the delay of the routers, which involves a large range of parameters, such as the router's radix N, the number of virtual channels per port V, and the flit width W, is increased relative to link traversal. After a large set of experiments[1] using a 45 nm standard-cell library at 0.8 V, we verify that the router's delay spans 650–1100 ps, assuming fast single-cycle routers employing a combined-allocation policy [6], [13]. Hence, based on the delay profile of routers and the range of wire delays, *the delay of a router is always* $2\times$ *larger than the wire delay per mm*. Equivalently, the delay of a router corresponds roughly to wire traversals of 3–6 mm, assuming a conservative wire delay in the middle of the possible wire delay range.

Previous work tried to exploit this asymmetry between the intra- and inter-router delays (link traversal) in several ways, with the main goal being the transfer of flits over a larger distance in a single cycle. Such solutions led to either high-radix networks [5], [8] with long connecting links, or to networks that allowed flits to traverse multiple network hops of shorter wires in a single clock cycle (through router bypassing) [2], [12].

Motivated by the same asymmetric delay property (between router and link traversal), we follow a different design direction, which allows for substantial performance optimizations to the NoC architecture. Rather than allowing flits to traverse longer distances in each clock cycle, we, instead, choose *to send one flit in each half-cycle using double-edge-triggered interfaces* and targeting medium-length wires of up to 2–3 mm, which sufficiently cover low-to-medium-radix topologies. For instance, at the 32 nm technology node, the longest side of a typical Chip Multi-Processor (CMP) tile (i.e., CPU core + 32 kB L1 instruction and data caches + a 512 kB L2 cache slice) is 3.27 mm, as demonstrated in [8]. For longer inter-router links, dual-edge-triggered pipeline registers can be added, in order to pipeline the links. These link pipeline registers can maintain the required (half-cycle) link traversal delay profile for each link segment, which, in turn, allows for double-data-rate transfers on each link segment.

III. RapidLink NoCs with DDR Links

In RapidLink, the sender and the receiver on each link are able to send and receive flits at both the positive and negative edges of the clock. This effectively *doubles the frequency of data transmission* on the links. In this scenario, the NoC routers maintain both (a) their original link bit-width, and (b) their original operating frequency, without experiencing tighter delay constraints, but they should provide two separate send/receive paths to each inter-router link. These two separate paths each carry a separate data stream. Both streams are then simultaneously (in a time-shared manner) transferred in DDR mode across the same inter-router link; one stream "rides" the positive phase of the clock, while the other stream "rides" the negative phase.

A scalable and cost-effective design option allowing for two distinct send/receive paths to each inter-router link involves

[1]Evaluations were performed for the following set of NoC router design parameters: N={3,5,8}, W={16,32,64}, and V={2,4,6}. For each {N,W,V} triplet, an independent delay-constrained optimization was performed.

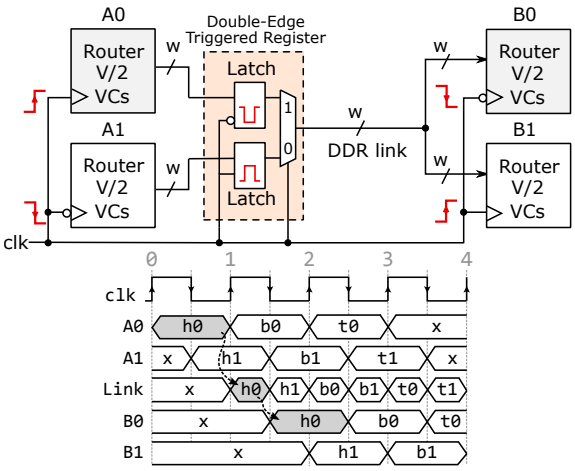


Fig. 1. A RapidLink NoC employing VC partitioning and DDR links. The two sub-routers (each hosting half of the VCs of the original router) time-share a W-bit inter-router link using both edges of the clock.

router/network partitioning. Under this approach, illustrated in Figure 1, the RapidLink NoC is partitioned into two sub-networks of W-bit widths, with each sub-network hosting half of the Virtual Channels (VC) of the baseline design. The two sub-routers time-share a W-bit inter-router link using both edges of the clock. For half a clock period (e.g., during the positive phase of the clock), the link is traversed by flits of sub-network 0, while during the other half of the clock period (e.g., the negative phase), flits of sub-network 1 use the link. Setting the appropriate clock edge for each router must now satisfy two conditions: (1) the pattern of alternating clock edges must be applied across routers of the same sub-network to guarantee half-cycle link traversal; (2) in order to assign a different half-period to each sub-network, the two sub-routers located at each node should operate under different clock edges.

As it will be demonstrated in Section IV, the total area consumed by the two sub-routers is almost identical to the area of the original router (less than 3% difference), while the two sub-routers are faster than the original router, due to the simplification of the allocation and multiplexing logic. Therefore, network performance is gained without increasing the NoC area, or the width of the links.

The output registers of each sub-router in Figure 1 are removed and the output data is fed into a shared Double-Edge Triggered (DET) register. Each DET register consists of two latches placed in parallel, which are enabled on opposite phases of the clock, and an output multiplexer driven by the clock signal [14]. Thus, *DET registers incur no additional overhead*, as compared to generic single-edge-triggered registers, which are also built using two latches (master and slave latches placed in series). The clock signal driving the multiplexer of the DET register is appropriately gated when no new valid flits arrive from any of the sub-routers, thus preventing unnecessary switching activity on the DDR link.

Figure 1 also illustrates the activity on the link, as flits flow from routers A0/A1 to routers B0/B1. On the positive edge of cycle 0, flit 'h0' enters router A0 and spends a full cycle inside the router. As the clock transitions to the negative level, the 'h1' flit is written in the input buffer of router B0, and all of the following occur simultaneously: (a) the A0 output latch becomes transparent, (b) the A1 output latch becomes opaque,

latching the data at its input, and (c) the multiplexer selects A1 as the link output. As the positive edge of cycle 1 arrives, the A0 latch stores flit 'h0,' which has completed a whole cycle inside the router, and the multiplexer switches to forward its flit on the link. Half a cycle later, the flit reaches router B0 and is captured on the positive clock edge. At the same time, 'h1' appears on the link, after the B0 latch becomes opaque, and the multiplexer's select signal switches again, while the same pattern continues in the following cycles.

Since each sub-network serves a particular group of VCs (half the VCs of the original router), moving from one VC of the first sub-network to a VC in the second sub-network is impossible, due to the physical separation of the clock-edge-interleaved sub-networks. When the NoC serves only 2 VCs (V=2), then the two sub-networks would not employ VC-based routers, since they would only need to serve 1 VC each. Instead, each sub-network would employ simpler wormhole routers, which save both area and reduce the delay.

Even though RapidLink affects the timing of the NoC links, it does not affect the full-cycle (single-edge) operation of the Network Interfaces (NIs). Each NI can safely assume an inject/eject throughput of at most 1 flit/full-cycle/NI, as in any baseline NoC. However, since RapidLink's DDR links can effectively receive and send 2 flits/node/cycle (due to the DDR transfers), re-time buffers and multiplexing should be added to the NIs to determine which flit should be sourced, or sinked. This modification adds at most 1.5 cycles of latency to each flit. However, this is amortized by the lower latency of RapidLink, which saves – by construction – half a cycle per hop. Sustained throughput is left unaffected, since (in most cases) throughput is limited by the utilization of the inter-router NoC links, not of the NIs. Once RapidLink's DDR links double the NoC link throughput (without increasing the link width), the overall NoC throughput increases significantly, as compared to full-cycle, single-data-rate NoCs.

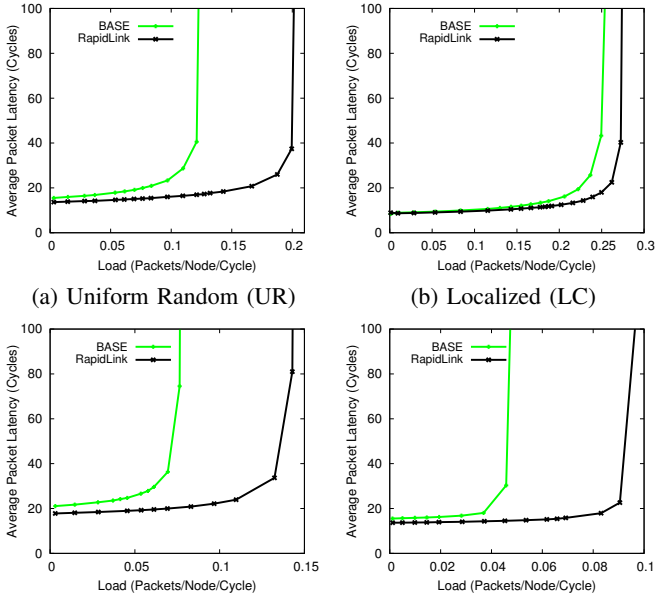
It should be noted that there is another design alternative to enabling DDR NoC operation. Instead of partitioning the NoC into two W-bit-wide sub-networks (as previously described), one may build routers with *double internal* data width (i.e., $2W$-bit-wide), and (de)serialize the flits at the input and output of the routers using the two edges of the clock. However, the extra (de)serialization steps introduce additional latency at each hop, while the routers experience almost double the area/power budgets, due to the width duplication of their buffering and switching components. Hence, the employed design solution with two separate W-bit-wide sub-networks is much more efficient.

IV. Experimental Evaluation

In this section, we evaluate the performance of RapidLink and compare it, in terms of network performance and hardware complexity, with baseline NoC architectures that assume full-cycle and single-data-rate link traversal.

Network performance comparisons were performed using a cycle-accurate SystemC network simulator that models all micro-architectural components of a NoC. We employ an 8×8 2D mesh network with XY dimension-ordered routing. Each router implements combined allocation [6] using also a MARX-based organization [15], and supports 4 VCs per input port with 3 flits/VC, using ElastiStore buffers [16], as needed to cover the credit Round-Trip Time (RTT) in single-cycle routers with full- or half-cycle links.

The performance evaluation involves four synthetic traffic patterns: Uniform Random (UR), non-uniform Localized (LC) traffic, and two versions of permutation traffic: Bit-Complement (BC) and Transpose (TS) traffic patterns. Under UR traffic, every node sends its packets to all other nodes of the network with equal probability. For LC traffic, we assume that 75% of the overall traffic is local (i.e., the destination is one hop away from the source), while the remaining 25% of the overall traffic is uniform-randomly distributed to the non-local nodes. The injected traffic consists of two types of packets to mimic realistic system scenarios: 1-flit short packets (just like request packets in a CMP), and longer 5-flit packets (just like response packets carrying a cache line). For the latency-throughput analysis, we assume a bimodal distribution of packets with 50% of the packets being short, 1-flit packets, and the rest being long, 5-flit packets.



(a) Uniform Random (UR) (b) Localized (LC)

(c) Bit-Complement (BC) (d) Transpose (TS)

Fig. 2. Latency vs. load curves for single-stage, 4-VC routers, under UR, LC, BC, and TS traffic patterns for baseline NoCs with full-cycle links and RapidLink NoCs with DDR links.

RapidLink with DDR links employs router partitioning across VCs, in order to create two independent paths that can be interleaved on the link using opposite edges of the clock. This allows for DDR operation on the links without increasing the clock frequency of the partitioned NoC routers. The DDR link operation is expected to offer up to 2× the saturation throughput of the baseline full-cycle and single-data-rate NoC, while still reducing zero-load latency.

This behavior is verified by the network performance results shown in Figure 2, which include baseline and RapidLink DDR NoCs. Full-width DDR NoCs achieve a 1.7–2× throughput increase under UR, BC, and TS traffic. In the case of LC traffic, saturation throughput is not limited by the NoC inter-router links and, thus, DDR links offer a marginal 10% improvement.

The NoC utilization increase under UR traffic is also

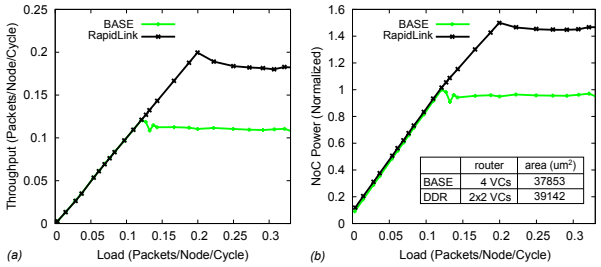


Fig. 3. (a) The throughput, and (b) the normalized power consumption of RapidLink with W-bit DDR links and two sub-networks (each one serving $V/2$ VCs), versus a baseline NoC with W-bit single-data-rate links. The NoC node area for each case (i.e., one router in the baseline case; two sub-routers in the RapidLink case) is also included for W=64 bits.

highlighted in Figure 3(a), together with the power of both designs with increasing traffic load, which is depicted in Figure 3(b). Up to the point where the baseline NoC saturates, both architectures consume the same amount of power. RapidLink utilizes – in parallel – two sub-networks of simpler routers (with $V/2$ VCs), with each one consuming almost half the power of a router supporting all V VCs. Hence, the overall power is almost equivalent to that of the baseline NoC. Beyond saturation, the baseline NoC cannot increase its utilization, thereby consuming a constant amount of power. On the contrary, RapidLink continues to higher utilization rates, which, inevitably, results in higher power consumption. However, the extra power is the result of increased utilization that cannot be achieved by the single-data-rate baseline NoC with full-cycle LT.

The significant throughput increase offered by RapidLink is achieved without dedicating more resources to the NoC than the baseline NoC with full-cycle and single-data-rate links (neither within the routers, nor on the links).

The table in Figure 3(b) reports the layout area occupied by a RapidLink router consisting of 2 parallel sub-routers (each one supporting 2 VCs) and 64-bit input-output ports, and the area of a 4-VC baseline router with a 64-bit datapath. Both designs are sized to operate at 1 GHz. The outputs of the two sub-networks of RapidLink are multiplexed on a 64-bit link (i.e., same width as in the baseline NoC) using the two edges of the clock, as depicted in Figure 1. The area of both designs is almost the same with less than 3% difference. All area-delay characteristics were obtained by following the Cadence back-end flow (synthesis and place-and-route), driven by a commercial low-power 45 nm 0.8 V standard-cell library, under worst-case conditions (0.8 V, 125 °C).

Note that the reduced latency and increased throughput provided by RapidLink's DDR link operation allows for a reduction in the NoC's clock frequency. This would reduce the power consumption of both the clock tree and the routers, while still delivering the performance of the baseline NoC.

V. Conclusions

The asymmetry between the NoC intra- and inter-router delays has been exploited in many forms in the past, primarily aiming to allow flits to traverse longer distances within a single clock cycle. NoCs with high-radix routers constitute such an example; they allow flits to reach their destinations using fewer hops, albeit through the use of longer links and fairly complex routers. The increased number of ports complicates allocation and switching logic, and requires custom design to achieve acceptable operating frequencies. Additionally, the design of high-radix networks typically leads to complicated layouts and wire-routing congestion, which also necessitate custom design effort. Finally, high-radix NoCs incur higher latencies and power consumption when handling local traffic (e.g., near-neighbor), because of unnecessary data movement over longer distances. Another alternative is to employ single-cycle multi-hop link traversal. This approach relies on complicated flow control and router bypassing to cross multiple hops "asynchronously" in one cycle. Once again, this philosophy does not offer a true benefit under localized traffic.

On the other hand, the proposed RapidLink NoC architecture complements previous state-of-the-art proposals by following a distinct and more scalable design path, which improves network performance without increasing design cost. RapidLink is minimally intrusive to both the router's micro-architecture and the flow-control policies, and it can be applied to any low- or medium-radix topology.

References

[1] T. Krishna, C.-H. O. Chen, S. Park, W.-C. Kwon, S. Subramanian, A. P. Chandrakasan, and L.-S. Peh, "Single-cycle multihop asynchronous repeated traversal: A smart future for reconfigurable on-chip networks," *IEEE Computer*, 2013.

[2] T. Krishna and L.-S. Peh, "Single-cycle collective communication over a shared network fabric," in *International Symposium on Networks-on-Chip (NOCS)*, 2014.

[3] B. Daya, L.-S. Peh, and A. Chandrakasan, "Towards high-performance bufferless nocs with scepter," *IEEE Computer Architecture Letters*, 2015.

[4] R. Manevich and et al., "Designing single-cycle long links in hierarchical nocs," *Microprocessors and Microsystems*, no. 8, pp. 814–825, 2014.

[5] N. Abeyratne, R. Das, Q. Li, K. Sewell, B. Giridhar, R. Dreslinski, D. Blaauw, and T. Mudge, "Scaling toward kilo-core processors with asymmetric high-radix topologies," in *IEEE International Symposium on High Performance Computer Architecture (HPCA)*, 2013.

[6] J. Postman, T. Krishna, C. Edmonds, L.-S. Peh, and P. Chiang, "Swift: A low-power network-on-chip implementing the token flow control router architecture with swing-reduced interconnects," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Aug. 2012.

[7] A. Psarras, J. Lee, P. Mattheakis, C. Nicopoulos, and G. Dimitrakopoulos, "A low-power network-on-chip architecture for tile-based chip multi-processors," in *ACM Great Lakes Symp-VLSI*, 2016, pp. 335–340.

[8] K. Sewell, R. Dreslinski, T. Manville, S. Satpathy, N. Pinckney, G. Blake, M. Cieslak, R. Das, T. Wenisch, D. Sylvester, D. Blaauw, and T. Mudge, "Swizzle-switch networks for many-core systems," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 2, pp. 278–294, 2012 2012.

[9] W. J. Dally, C. Malachowsky, and S. W. Keckler, "21st century digital design tools," in *Proceedings of the 50th Annual Design Automation Conference*, ser. DAC '13, 2013, pp. 94:1–94:6.

[10] A. Golander, N. Levison, O. Heymann, A. Briskman, M. J. Wolski, and E. F. Robinson, "A cost-efficient L1–L2 multicore interconnect: Performance, power, and area considerations," *IEEE Transactions on Circuits and Systems-I: Regural Papers*, vol. 58, no. 3, pp. 529–538, March 2011.

[11] P. Salihundam et al., "A 2Tb/s 6x4 Mesh Network with DVFS and 2.3Tb/s/W router in 45nm CMOS," in *VLSI Circuits*, 2010.

[12] T. Krishna, C.-H. O. Chen, W. C. Kwon, and L.-S. Peh, "Smart: Single-cycle multi-hop traversals over a shared network-on-chip," *IEEE Micro*, May/June 2014.

[13] Y. Lu, C. Chen, J. V. McCanny, and S. Sezer, "Design of interlock-free combined allocators for networks-on-chip," in *EEE 25th International SOC Conference (SoCC)*, 2012, pp. 358–363.

[14] Q. Wu, M. Pedram, and X. Wu, "A new design of double edge triggered flip-flops," in *Asia and South Pacific Design Automation Conf.*, 1998, pp. 417–421.

[15] G. Dimitrakopoulos, E. Kalligeros, and K. Galanopoulos, "Merged switch allocation and traversal in network-on-chip switches," *IEEE Transactions on Computers*, vol. 62, no. 10, pp. 2001–2012, 2013.

[16] I. Seitanidis, A. Psarras, G. Dimitrakopoulos, and C. Nicopoulos, "Elastistore: An elastic buffer architecture for network-on-chip routers," in *Proc. of Design Automation and Test in Europe*, Mar. 2014.