

# A 4-PAM Interconnect in Network-on-Chip for High-Throughput and Latency-Sensitive Applications

Ahmad Mansour<sup>1</sup>, Ahmed Elnaggar<sup>1</sup>, Bassma Alabassy<sup>2</sup>, Mostafa Khamis<sup>2</sup>, Ahmed Shalaby<sup>3</sup>

<sup>1</sup>Electrical Engineering Department, Alexandria University, Egypt

<sup>2</sup>Mentor Graphics Corporation, Cairo, Egypt

<sup>3</sup>Department of Computer Science, Faculty of Computers and Informatics, Benha University, Egypt

**Abstract**— In this paper, a network-on-chip four-level pulse amplitude modulation (4-PAM) scheme is proposed to be used for communication within the network itself in MPSoCs. A current-mode based 4-PAM transmitter is used to encode data transactions between neighboring routers. Decoding data streams is done by a flash-ADC based receiver using clocked latched type comparators. Additionally, this scheme is implemented on networks utilizing high-radix routers with a local concentration factor of 2 IPs per node to encode data streams injected into the network at the network interface and decode them at the input port of the router. We also discuss the required modifications to the router architecture in the input port buffers and introduce a two-stage allocation method to resolve conflicts of output port requests which is essential to maintain system stability after saturation by utilizing a fair flow control methodology. This results in a reduction in wiring load for each router which is an added value that facilitates the routing stage. The evaluation is extended to reflect the overall network performance supporting the use of multi-valued logic and estimate the overhead of implementation on area and power budgets.

**Keywords**—*System-on-Chip (SoC), Network-on-Chip (NoC), Pulse Amplitude Modulation (PAM), Multi-Valued Logic(MVL).*

## I. INTRODUCTION

As a result of the on-going advancement in technology of fabrication, the feature size of CMOS devices is scaling down. This in-turn allows the integration of multiple cores on a single die leading to multi-processor systems-on-chip (MPSoCs). Recently, these systems are used in applications that demand high performance in spite of the strict requirements imposed on resources and power consumption. Accordingly, using network-on-chip was proposed as a feasible solution to improve chip performance. For instance, a neural network NoC was implemented for Augmented Reality applications to monitor the workload history of all processing cores and control the inter-communication and data transactions between cores in distributed tasks resulting in a 29.1% improvement in throughput [1]. Another major contribution of NoCs can be viewed in heterogeneous multi-core architectures where the high-volume data exchange between CPUs and accelerated GPUs forms a bottleneck for conventional platforms. This will overcome the need for expensive power-hungry off-chip connections that degrade performance in terms of data-transfer latency. Implementations of such hybrid NoCs resulted in 25% improvement in system energy delay product when used for deep learning kernels [2]. Any NoC can approach its optimal

performance if the design is fitted according to the application. In the presence of high communication-bandwidth between two separate cores or IP blocks, a modification in the topology by adding a direct link can be of major impact on the system performance. This approach is extended further by utilizing high-radix routers to allow application-driven configuration at runtime by controlled activation of port links [3]. However, the need for topology reconfiguration will not be beneficial if the application is known beforehand. Thus, a proper mapping algorithm will simply place cores with high data-exchange rate on the same high-radix router (router with a local concentration factor of more than 1) to minimize the average number of hops throughout the application run. However, the emphasis in NoC architecture development proposals is mainly targeting the router design itself; optimizing in routing algorithms, arbitration and buffer allocation whereas the interconnection between routers within the NoC itself is rarely discussed. In this paper, we propose using multi-valued logic (4-PAM) to transmit packets on port channels, this will contribute to several aspects:

1. Improve bandwidth utilization of all interconnects within the network itself offering higher throughput and lower latency for all in-flight packets.
2. Resolve conflicts for the arbitration stage by granting multiple-packet to traverse the physical link when multiple redundant requests are present.
3. Implementing this scheme in high-radix routers where each two local ports can inject and collect using the same port-links. Thus, reducing the wiring area overhead of an entire port as it forms a major bottleneck in the routing stage.

The following section will showcase the required background and related work. Section III will discuss in-depth the NoC router architecture implemented and the 4-PAM-signaling design. On follows, a thorough evaluation of the proposed scheme in section IV. Finally, section V concludes the paper.

## II. BACKGROUND AND RELATED WORK

### A. Multiple Valued Logic Approach

Instead of limiting the domain of numerical representation of logic levels to two-level logic  $D = \{0,1\}$  as commonly used in

digital circuits. Additional flexibility can be introduced by increasing the domain of representation to  $D = \{0, 1, 2, \dots, L\}$  using multi-valued logic (MVL). The idea of using MVL mainly originates from the Lukasiewicz logic [6]. MVL offers the ability to transmit more information with fewer interconnects than possible by using traditional binary logic. Accordingly, it appealed as a potential approach that was utilized in many applications; MVL memories, reversible computing and even some applicability in quantum computing [7].

### B. Link Optimization Techniques

The link bandwidth is mainly affected by two parameters; the link width and the operating frequency. Accordingly, it is a key-contributor to the area overhead and power budget. Increasing either in an attempt to increase the bandwidth will have side-effects that must be taken into consideration.

1. Increasing the link-width: Results in increased wiring resources, routing problems and wider buffers that in turn add on power and area.
2. Increasing the operating frequency: Results in a pump in power consumption in all parts of the system.

Thus, the bandwidth overhead must be minimized according to the system requirements and constraints. Some schemes were previously proposed like using both phases of the clock to transmit two different flits over bidirectional links to double the data rate. A link module is inserted between each two neighboring routers, mainly composed of two D flip-flops, to store the data received from both sides in the first half cycle and forward it in the second half cycle [6]. Another approach, as suggested for multiple unicast and multicast communication, is to use network coding to forward two encoded packets on common links towards their destinations. This requires additional encoding, decoding and replication mechanisms to be integrated in the router architecture [7,8]. In proposals targeting link area, a NoC SerDes transceiver was proposed in [9] by serializing 16 data streams onto one data line. Another, proposed a 4-level PAM transceiver for transmitting and receiving two data streams on the same channel but only limited the implementation to the transceiver itself with no evaluation or integration in an actual NoC [10].

## III. PROPOSED ARCHITECTURE AND IMPLEMENTATION

Any IP block in the network is connected to a local port via a network interface. The task of the interface is to encapsulate raw data into the supported NoC packet format. These packets then propagate through the network from the source to the target destination indicated by the routing information present in the packet header. Within each hop, it undergoes several stages; storage, routing computation, switch allocation and link traversal.

### A. NoC Router Architecture

As soon as a packet arrives, it is stored in the input port buffer. An input port controller is responsible for controlling the flow of the buffer storage and slot de-allocation when a packet departs. The next stage is to perform routing computation. In this

work, an XY deterministic routing is used as it is deadlock-free and reduces the overhead of computation. In the routing stage, the destination router address is compared to the router ID on each hop. In an 2D mesh for instance, if the destination X dimension is greater than the source X dimension, the east port is requested. If it is smaller, the west port is requested. The same happens with the Y dimension after the X dimensions of both addresses match. On follows, the allocation stage where all requests on each router port are forwarded to an output port controller that uses Round Robin (RR) arbiters to shift priority and achieve a fair flow control. The arbiter grants only one input port request to allocate the channel bandwidth between the two neighboring routers.

### B. Modifications to Support MVL in the Behavioral Domain

For each interconnection link between two routers, two packets can be forwarded. Accordingly, in the presence of conflicts where more than one request is present on the same output port, this will reduce the blocking rate of the allocation stage. A modification to the allocation logic is needed to maintain the system stability and flow control fairness. A two-stage allocation method as described in Fig. 1 is used where the first stage of arbitration grants the first agent request and the second stage uses this grant to mask the request of the second stage and grants a different agent. Both stages use RR priority rotation. In case there is only one buffer slot available in the upstream router, the second stage is bypassed.

### C. Modifications to Support MVL at the Circuit Level

In the architectural domain, the modification can be viewed as a redefinition of what a logic wire is. Fig. 2 shows a conceptual impression of a 4-PAM logic wire where each pair of input or output binary logic wires are perceived as a single entity by the 4-PAM structure in the same sense as singleton bits are perceived as single entities for purely binary structures. The 4-PAM wire comprises a 4-PAM transmitter, a regular wire and a flash-ADC-based receiver.

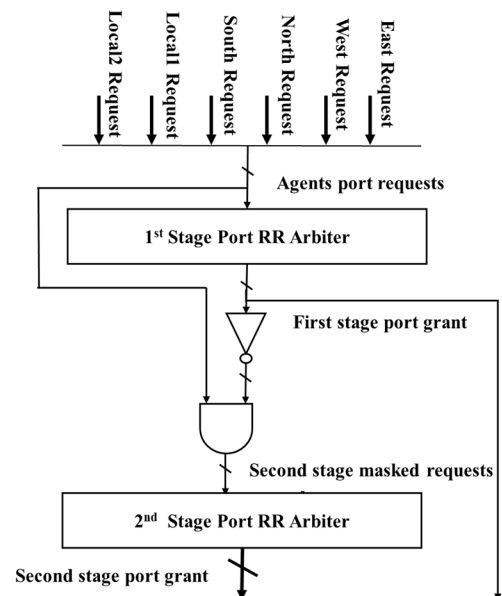


Fig. 1. The proposed two-stage allocation scheme of output port controller.

The 4-PAM transmitter is a current-mode-based operator where the pair of input bits;  $\{S_1, S_0\}$ , modulates a current through a resistor to generate the corresponding 4-PAM representation. For all purposes, the input data pushes through within a single clock cycle. Input encoding correspondence is given in Table I which depicts a modified thermometer code with a Gray-encoded-input arrangement for less error backlash.

The transmitter is shown in Fig. 3. A principal current mirror mirrors a reference current into a reference branch P1B-P2B-R. This branch is then roughly replicated to form the generating network where the gate of P2B serves as the gate bias for the current sources P2T2 and P2T1. Under perfect matching conditions, each branch should replicate the reference current. Transistors P1T2 and P1T1 are switching transistors that activate or deactivate their respective branch depending on the control signals T1-T2. With these transistors in place, the always-on P1A-P1B transistors are added to the mirror to keep the branches as symmetric as possible while, simultaneously, allowing minimization of the mirror's systematic error through the negative-feedback-connected op-amp's gain. A compensation capacitor might be required between the op-amp's output and VDD but is found not necessary for the given case and compensation is achieved indirectly by the loading devices. Transistors PT0-NT0 form a push-pull pair for the case of  $\{S_1, S_0\}$  corresponding to VDD or GND which is entirely resolved by T0 and should be more power economic for this case. Following the above description, it can be seen that the output will swing by an  $R \cdot I_{ref}$  increment per input-pair increment. Choosing  $R \cdot I_{ref} = V_{DD}/3$  will thus correspond to an evenly distributed 4-PAM representation.

The 4-PAM receiver is a flash-ADC based type as shown in Fig. 4. The reference levels are chosen to be midway between the transmitted levels for maximum noise margins. The reference generator can be as simple as a resistive ladder or a dedicated multiple-output switched capacitor DC-DC converter for low power budgets. The comparator (Fig. 5) is a clocked-latch-type comparator reported in [11] and operates as following; when CLK is high, MP0 and NM0 are off whereas MPC1 and MPC2 are on. Consequently, the collective equivalent parasitic capacitances at the gates of NR1 and NR2 are pre-charged to VDD which in return set both outputs of the latch to GND. When CLK turns low, MP0 and NM0 turn on, MPC1 and MPC2 turn off and the differential pair ND+/ND- is activated.

If  $V_+ > V_-$ , the charge at the gate of NR1 gets steered to GND, NR1 loses its grip on OUT and OUT is latched high. The inverse is true for the case of  $V_+ < V_-$ .

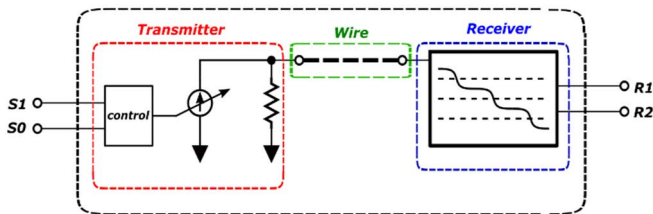


Fig. 2. 4-PAM logic wire.

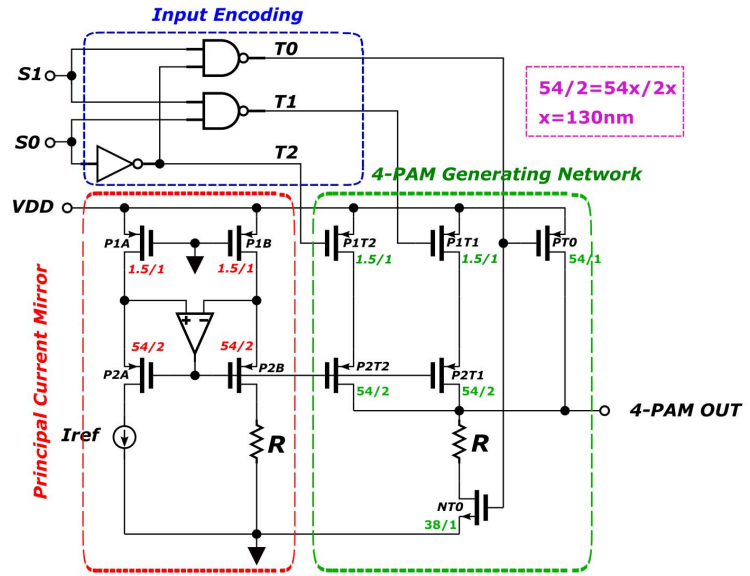


Fig. 3. 4-PAM transmitter

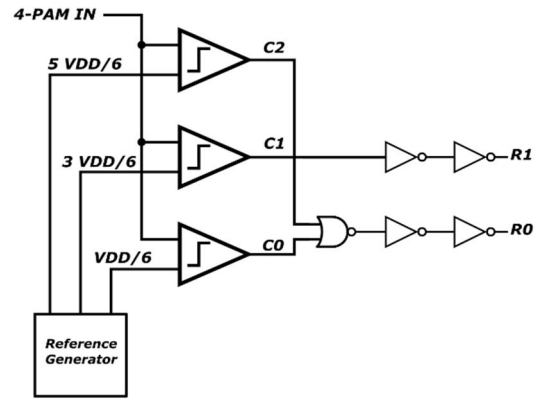


Fig. 4. Flash ADC architecture

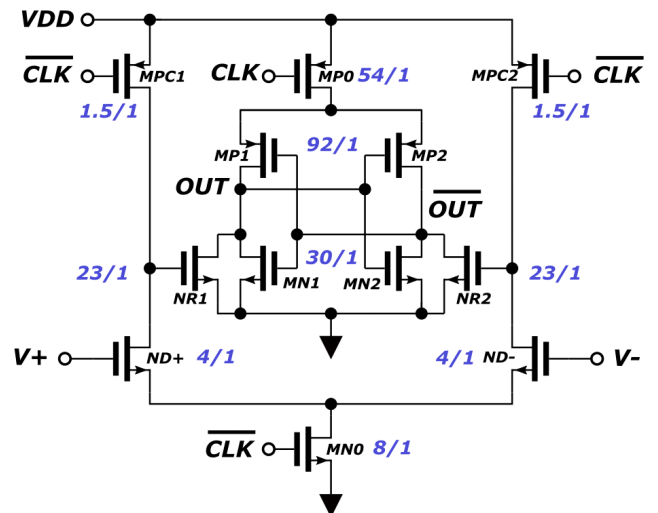


Fig. 5. Latch-type comparator

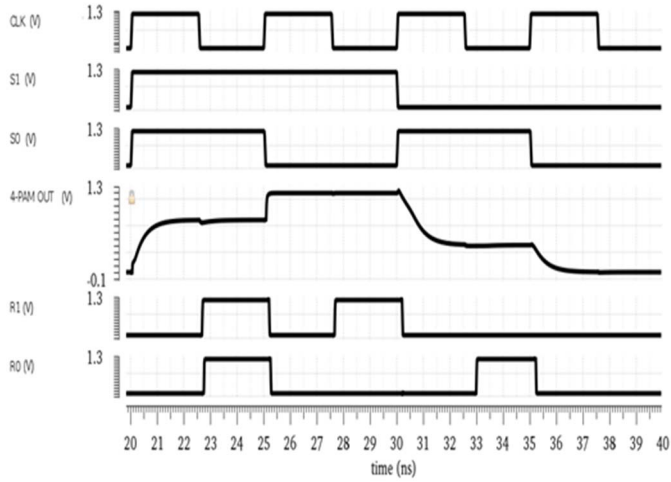


Fig. 6. Back-to-Back transmitter-receiver connection. The signals in order (from top to bottom) are CLK,  $S_1$ ,  $S_0$ , 4-PAM OUT,  $R_1$  and  $R_0$

Fig. 6 shows the result of connecting the transmitter and the receiver in a back-to-back configuration for the typical corner condition at 27 °C in a 130 nm CMOS technology. The data is clocked in at a 200 MHz frequency,  $R=20$  k $\Omega$ ,  $I_{ref} = 20$   $\mu$ A and  $V_{DD}= 1.2$  V which results in a 0/400/800/1200-mV-4-PAM system. The ADC references are therefore located at 200 mV, 600 mV and 1 V respectively. The ADC decoding correspondence is given in Table II.

#### IV. MEASUREMENTS AND EVALUATION

To achieve proper evaluation of any digital design architecture, there are 3 primary physical characteristics that need to be taken into consideration; speed, area and power. An all-inclusive evaluation of the first should consider the following 3 metrics:

1. **Throughput:** In the context of network performance, it defines the maximum accepted traffic by the network and can vary depending on its resources configuration and implementation methods. It can be expressed by the number of packets received every cycle normalized to the total number of cores or IP blocks connected in the network.
2. **Latency:** The timing between injection of a packet from a source node to receiving it at the destination node expressed in number of cycles.
3. **Timing:** Mainly conveyed by the delay of the combinational logic present on the critical path of the NoC router and it defines its maximum operating frequency.

##### A. Measurement Setup

In order to test each of the aforementioned criteria, the measurement setup shown in Fig. 7 is implemented as a means to simulate the network traffic. Architectures supporting a local concentration factor will generally yield optimum results in traffic patterns where the communication bandwidth between the two local cores of the unit router represents a large portion.

The input and output timing and counting units act as monitors collecting performance information that is processed at the end of the simulation to compute the latency and throughput of the system at different injection rates. The sweep on the injection rate is achieved by manipulating the average delay in the cycles between each two consecutive packets generated.

TABLE I  
INPUT ENCODING CORRESPONDENCE

$S_1$	$S_0$	$T_2$	$T_1$	$T_0$	4-PAM
0	0	1	1	1	0
0	1	0	1	1	400 mV
1	1	0	0	1	800 mV
1	0	1	1	0	1.2 V

TABLE II  
ADC DECODING CORRESPONDENCE

4-PAM	$C_2$	$C_1$	$C_0$	$R_1$	$R_0$
0	0	0	0	0	0
400 mV	0	0	1	0	1
800 mV	0	1	1	1	1
1.2 V	1	1	1	1	0

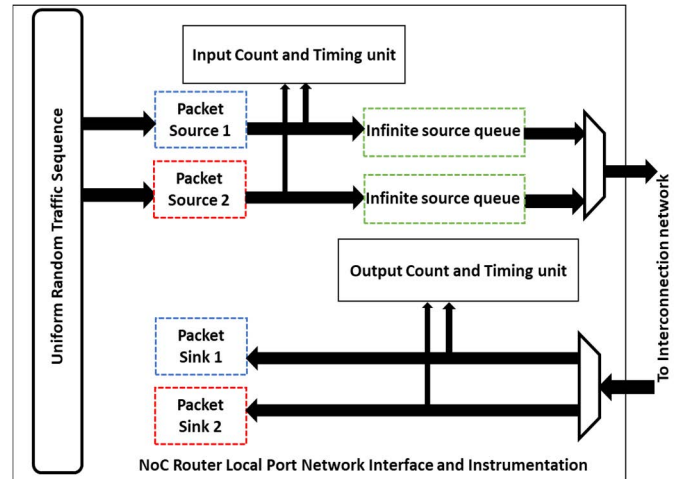


Fig. 7. Measurement setup

##### B. Performance Evaluation

Comparing our proposed modified architecture against its counterpart of the exact same configuration for buffering resources and bandwidth, as shown in Fig. 8(a), there is a 6.67% improvement in the throughput saturation level yielding a maximum of 0.08 packets/cycle/PE instead of 0.075 packets/cycle/PE. On the other hand, as indicated in Fig. 8(b) latency is improved by 8% dislocating the latency saturation point to be at an injection rate of 0.054 packets/cycle/PE. Additionally, a 9.3% enhancement in zero-load-latency is also achieved reducing the average latency by 3 cycles/packet to be 29 cycles/packet instead of 32 cycles/packet.

### C. Conflicts Rate Analysis

The main contribution of the proposed architecture can be attributed to resolving conflicts of allocation and arbitration. The allocation problem is due to having a set of agents requesting on a set of resources and the goal is to achieve the optimal matching between the available resources to grant the requesting agents. The allocation problem is broken down into smaller fragments forming the arbitration problem. That is the set of agents requesting on one physical resource and the availability of a single grant for one agent.

In our architecture, the packets in-flight at the top of each input port queue represent the agents and the channel's bandwidth of the router output port represents the resource. Conflicts occur when more than one packet requests on the same output port. Accordingly, analysis of the effect of the proposed scheme in resolving conflicts is a key-parameter in our evaluation. Fig. 9 shows the Resolved Conflicts Ratio (RCR) for different injection rates where:

$$RCR = \frac{\sum_{i=1}^L N_i}{\sum_{i=1}^L F_i}$$

Where  $N_i$  is the total number of resolved conflicts associated with the  $i^{\text{th}}$  router,  $F_i$  is the total number of inflights associated with the  $i^{\text{th}}$  router and  $L$  is the total number of routers.

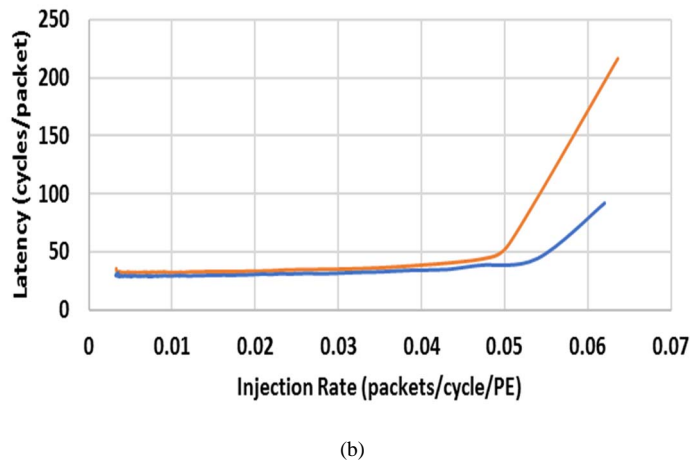
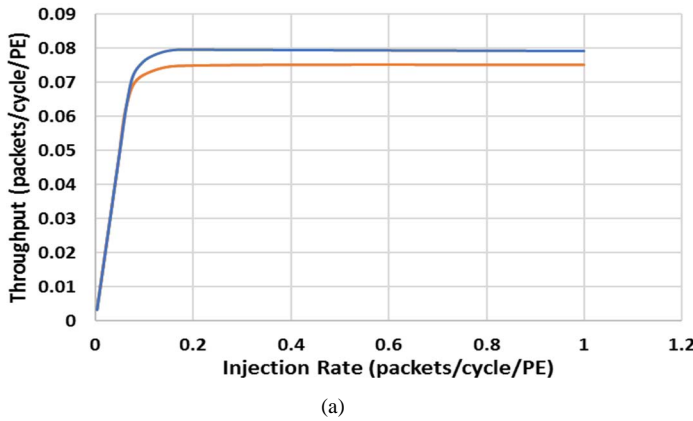


Fig. 8. Performance evaluation of the proposed architecture (in blue): (a) throughput vs. injection rate (b) latency vs. injection rate

As the results show, 18% of link traversals that occurred in the network utilized the 4-PAM technique to resolve the arbitration conflict. This percentage remains roughly constant over a wide range of injection rates. However, a noticeable drop occurs after the saturation point at an injection rate of 0.054 packets/cycle/PE. This drop can be attributed to the fact that at large injection rates, buffering resources are fully occupied and are blocking the traffic flow. Thus, the probability of presence of two empty buffer slots in the input port buffer of the upstream router is reduced significantly. Accordingly, it will not be able to utilize the bandwidth as efficiently as possible due to buffering limitations.

The results above can somehow be misleading to the assumption of steady utilization of PAM over the entire network. However, further analysis of the network status will uncover a variation in bandwidth utilization using PAM depending on the location of the router in question and the even the port direction.

### D. Variations in Contribution to Conflicts

Any core connected to a NoC router will inject into the network to communicate with another one either on the same router or a different one. Table III discusses the possible combinations of conflicts that may occur for the channel bandwidth of each output port and can be resolved by using PAM. The combinations are noted by the name of the two input ports requesting the channel bandwidth. We have only three constraints to consider:

1. No core will forward a packet to itself.
2. A packet in a port's input buffer will never request on it as an output port.
3. Packets coming from north and south ports will never request on east or west ports as we are using XY routing. Those cases are the invalid ones (marked as yellow in the table).

However, depending on the router position in the network these combinations will contribute differently to the number of conflicts within the router and to the overall network traffic. For instance, in a 4x4 mesh as shown in Fig. 10, the router 00 is in a corner case. It has only two active links (East and South) and the two local injectors. Accordingly, it will contribute the least to the network traffic and to the overall number of conflicts. To further illustrate, our measurements of the contribution percentage of each router to the overall network traffic along with its contribution to the total sum of resolved conflicts is shown in Fig.11.

To get an estimate, all cases are considered in Table IV. As expected the minimum percentage is in corner routers R00, R30, R03 and R33 with an average of 5.77%. Edge routers contribute by an average of 6.26% and surrounded routers contribute by the maximum average percentage of a 6.7%. Within each router, the conflicts resulting from requests on specific ports (East and West) form the major contributions and the rest contributes slightly. This is due to the nature of the routing computation algorithm used.

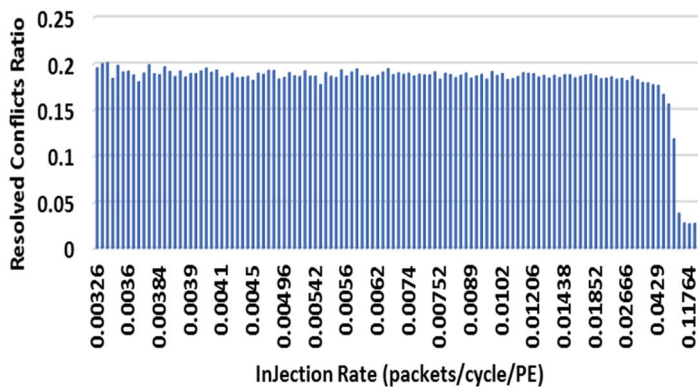


Fig. 9. Resolved conflicts ratio vs injection rate

TABLE III  
POSSIBLE COMBINATIONS OF CONFLICTS

Requested Port	Possible combinations of conflict									
East (E)	WN	WS	WL1	WL2	NS	NL1	NL2	SL1	SL2	L1L2
West(W)	EN	ES	EL1	EL2	NS	NL1	NL2	SL1	SL2	L1L2
North(N)	EW	ES	EL1	EL2	WS	WL1	WL2	SL1	SL2	L1L2
South(S)	EW	EN	EL1	EL2	WN	WL1	WL2	NL1	NL2	L1L2
Local1(L1)	EW	EN	ES	EL2	WN	WS	WL2	NS	NL2	SL2
Local2(L2)	EW	EN	ES	EL1	WN	WS	WL1	NS	NL1	SL1

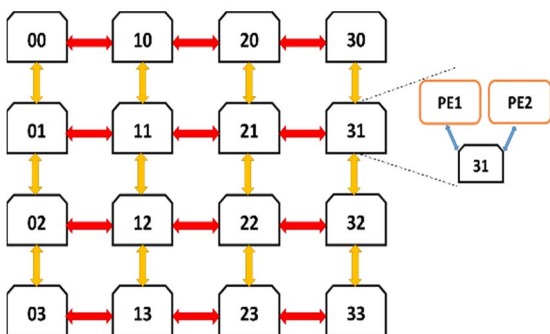


Fig. 10. A 4x4 mesh of interconnected routers

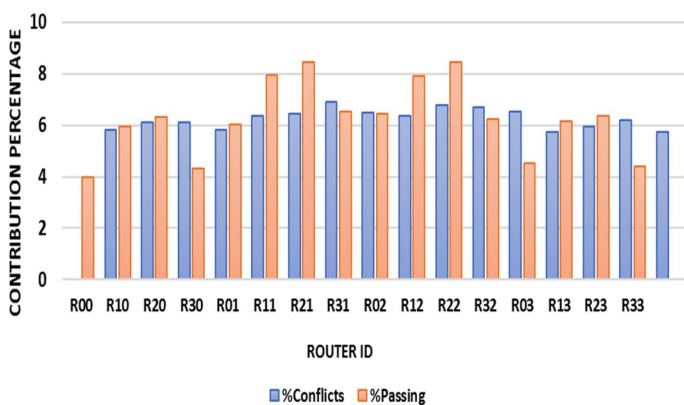


Fig. 11. Resolved Conflicts Percentage and all in-flights vs Router ID

Using deterministic XY routing in generic  $m \times n$  mesh, any packet in router  $(X_{src}, Y_{src})$  headed to router  $(X_{dest}, Y_{dest})$  follows the following probabilities described by Eq. 1 to Eq. 5 under uniform traffic pattern:

$$P(East) = P(X_{dest} > X_{src}) = ((m - 1) - X_{src}) * n * p \quad (1)$$

$$P(West) = P(X_{dest} < X_{src}) = X_{src} * n * p \quad (2)$$

$$P(North) = P(X_{dest} = X_{src}) * P(Y_{dest} < Y_{src}) = Y_{src} * p \quad (3)$$

$$P(South) = P(X_{dest} = X_{src}) * P(Y_{dest} > Y_{src}) = ((n - 1) - Y_{src}) * p \quad (4)$$

$$P(Local) = P(X_{dest} = X_{src}) * P(Y_{dest} = Y_{src}) = p \quad (5)$$

Where  $p$  is the probability of each destination router and that is the sum of probabilities of the two destination IP blocks connected to it. This is confirmed from the results shown in Fig. 12 where the maximum number of conflicts in any router is a result of requesting on the east port with an average of 37.6 % and the west port with an average of 36.2 %. Request on north and south port with lower probability where conflicts in requests occur with an average of 12.75% on the north port and an average of 13.34% on the south port. Finally, the local port has the minimum contribution where conflicts on it occur with an average of 1.04% only. These probabilities approach the ideal values when increasing the number of simulation cycles used for measurement.

#### E. PPA Estimations

All the HDL implementations of the NoC router are synthesized on Synopsys DC to estimate the area and power overhead from modifications in input port buffers and adding a second stage to the allocation unit which is expected from the cost-performance tradeoff. In our evaluation, we will consider the following parameters:

1. Maximum operating frequency: It is calculated by the synthesis tools to overcome the deficiency in network parameters evaluation. A low-latency system in terms of average number of cycles with a very high critical path delay will still end up resulting in a high actual delay in terms of seconds. A router will generally operate in a frequency range of 100MHz up to 1GHz as a maximum. Timing analysis shows that the router can operate on frequencies up to 900MHz although such high-range speeds are used less frequently.
2. Power and area overhead: The added functionality comes at the expense of 31.8% increase in power consumption of the unit router from 2.2mW to reach 2.9mW. The area overhead is mainly due to the second arbitration stage that doubled the allocation unit in the output-port controller. However, the entire output port controller still contributes by only 4.6% of the router area. Additionally, modifications in the input buffers to support two write operations affected the area to occupy 87% of the input port area instead of 85% using a single write buffer.

TABLE IV  
RESOLVED CONFLICTS CONTRIBUTION PERCENTAGE OF  
EACH ROUTER

Router ID	R00	R10	R20	R30	R01	R11	R21	R31
% Conflicts	5.808	6.118	6.105	5.808	6.365	6.451	6.882	6.491
Router ID	R02	R12	R22	R32	R03	R13	R23	R33
% Conflicts	6.379	6.774	6.689	6.509	5.754	5.956	6.19	5.713

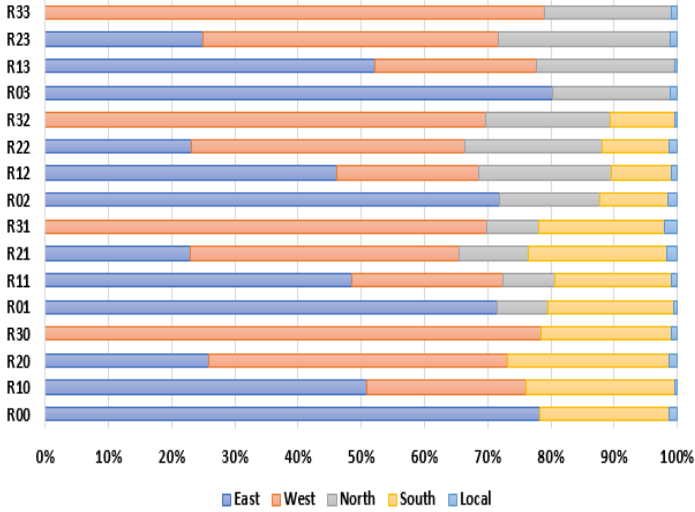


Fig. 12. Percentage of Resolved conflicts per port vs router position

## V. CONCLUSION AND FUTURE WORK

In this work, we discussed the usage of MVL for efficient bandwidth utilization by using a 4-PAM link designed on 130 nm CMOS technology at 1.2V supply for interconnections between routers and local port transactions in high-radix routers. We also reflected the required overhead in implementation of the NoC router itself. Our proposal resulted in a 6.67 % improvement in throughput, 9.3% in zero-load-latency and 8% in latency saturation point at the cost of 2% increase in buffer overhead on the input port and 31.8% on

power budget. This performance supports applications demanding high-throughput and low-latency communication.

Our future work may include exploring the integration of MVL memories and the use of recursive approaches for resource sharing in allocation to minimize the overhead.

## REFERENCES

- [1] Kim, G., Lee, K., Kim, Y., Park, S., Hong, I., Bong, K. and Yoo, H. (2015). A 1.22 TOPS and 1.52 mW/MHz Augmented Reality Multicore Processor With Neural Network NoC for HMD Applications. *IEEE Journal of Solid-State Circuits*, 50(1), pp.113-124.
- [2] Choi, W., Duraisamy, K., Kim, R., Doppa, J., Pande, P., Marculescu, R. and Marculescu, D. (2017). Hybrid network-on-chip architectures for accelerating deep learning kernels on heterogeneous manycore platforms.
- [3] Jain, A., Parikh, R. and Bertacco, V. (2014). High-radix on-chip networks with low-radix routers. *2014 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*.
- [4] J. Lukaszewicz, "On three valued-logic," in L. Borkowski, SelectWorks, Amsterdam, North-Holland, 1920, pp. 169-171.
- [5] Gaudet, V. (2016). A survey and tutorial on contemporary aspects of multiple-valued logic and its application to microelectronic circuits. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 6(1), pp.5-12.
- [6] Zhu, J., Qian, Z. and Tsui, C. (2016). BiLink: A high performance NoC router architecture using bi-directional link with double data rate. *Integration, the VLSI Journal*, 55, pp.30-42.
- [7] Duong-Ba, T., Nguyen, T. and Chiang, P. (2011). Network coding in multicore processors. *30th IEEE International Performance Computing and Communications Conference*.
- [8] Shalaby, A., Goulart, V. and Ragab, M. (2013). Study of Application of Network Coding on NoCs for Multicast Communications. *2013 IEEE 7th International Symposium on Embedded Multicore Socs*.
- [9] Hoppner, S., Walter, D., Hocker, T., Henker, S., Hanzsche, S., Sausner, D., Ellguth, G., Schlusser, J., Eisenreich, H. and Schuffny, R. (2015). An Energy Efficient Multi-Gbit/s NoC Transceiver Architecture With Combined AC/DC Drivers and Stoppable Clocking in 65 nm and 28 nm CMOS. *IEEE Journal of Solid-State Circuits*, 50(3), pp.749-762.
- [10] Byun, G. and Navidi, M. (2015). A Low-Power 4-PAM Transceiver Using a Dual-Sampling Technique for Heterogeneous Latency-Sensitive Network-on-Chip. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 62(6), pp.613-617.
- [11] Schinkel, D., Mensink, E., Klumperink, E., van Tuijl, E. and Nauta, B. (2007). A Double-Tail Latch-Type Voltage Sense Amplifier with 18ps Setup+Hold Time. *2007 IEEE International Solid-State Circuits Conference. Digest of Technical Papers*.
- [12] W. Dally and B. Towles, *Principles and practices of interconnection networks*. Amsterdam: Elsevier Morgan Kaufmann, 2010, pp. 449-458.