ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

High Data Transmission with Shared Buffer Routers Using RoShaQ Architecture

Chilukala Mohan Kumar Goud
M.Tech,
Santhiram Engineering College,
Nandyal, Kurnool(Dt), Andhra Pradesh.

C.V.Subhaskar Reddy, M.Tech, (Ph.D)
Associate Professor,
Santhiram Engineering College,
Nandyal, Kurnool(Dt), Andhra Pradesh.

Abstract:

Router architecture plays a central role in the performance of a Network on Chip (NoC). Router consists of buffers which arededicated to their input or output ports for temporarily storing packets during the times of congestion. Unfortunately, significantportion of the router area and power is consumed by the buffers alone. While running some tested traffic patterns, however, notall input ports of routers have incoming packets needed to be transferred simultaneously. Therefore, a large number of bufferqueues in the network are empty whereas the other queues are mostly busy.

This observation has led to the design of routerarchitecture with shared queues (RoShaQ) which maximizes the buffer utilization by sharing the multiple buffer queues amonginput ports. Buffers become more efficient by sharing queues, hence the router is able to achieve higher throughput when thenetwork load becomes heavy. On the other side, at light traffic load, this router is designed to achieve low latency by allowingpackets to effectively bypass these shared queues.A new adaptive weighted XY routing algorithm for eight port routerArchitecture is proposed in order to decrease the latency of the network on chip router.

Keywords: NoC, Router, RoShaQ, Buffer utilization.

Introduction:

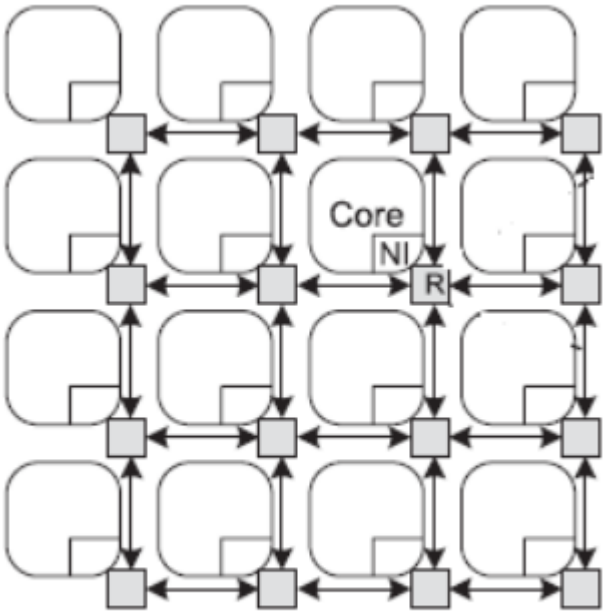
Network on chip or network on a chip (NoC or NOC) is a communication subsystem on an integrated circuit (commonly called a "chip"), typically between intellectual property (IP) cores in a system on a chip (SoC).

NoCs can span synchronous and asynchronous clock domains or use unclocked asynchronous logic. NoC technology applies networking theory and methods to on-chip communication and brings notable improvements over conventional bus and crossbar interconnections. NoC improves the scalability of SoCs, and the power efficiency of complex SoCs compared to other designs. The wires in the links of the NoC are shared by many signals. A high level of parallelism is achieved, because all links in the NoC can operate simultaneously on different data packets. Therefore, as the complexity of integrated systems keeps growing, a NoC provides enhanced performance (such as throughput) and scalability in comparison with previous communication architectures (e.g., dedicated point-to-point signal wires, shared buses, or segmented buses with bridges). Of course, the algorithms must be designed in such a way that they offer large parallelism and can hence utilize the potential of NoC.

Network on chip (NoC) is a communication subsystem on an integrated circuit. In the design of NoCs, high throughput and low latency are both important design parameters and the router micro architecture plays a vital role in achieving these performance goals. In a typical router, each input port has an input buffer for temporarily storing the packets in case that output channel is busy. This buffer can be a single queue as in a wormhole (WH) router or multiple queues in parallel as in Virtual Channel (VC) routers. High throughput routers allow an NoC to satisfy the communication needs of multicore and many core applications, or the higher achievable throughput can be traded off for power savings with fewer resources being used to

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

attain a target bandwidth. Further, achieving high throughput is also critical from a delay perspective for applications with heavy communication workloads because queueing delays grow rapidly as the network approaches saturation. Another approach is by sharing buffer queues that allows utilizing idle buffers or emulating an output buffer router to obtain higher throughput. Our work differs from those router designs by allowing input packets at input ports to bypass shared queues hence, it achieves lower zero load latency. In addition, the proposed router architecture has simple control circuitry making it dissipate less packet energy than VC routers and achieving higher throughput by letting queues share workloads when the network load becomes heavy. The proposed routing algorithm, which is transparent with respect to the router implementation, are presented and discussed, and assessed by means of simulation on synthetic and real traffic scenarios. The analysis takes into account several aspects and metrics of the design.



Multi Cores connected by a network of Routers
NI: Network Interface R: Router

Benefits of adopting NoCs

Traditionally, ICs have been designed with dedicated point-to-point connections, with one wire dedicated to each signal. For large designs, in particular, this has several limitations from a physical design viewpoint. The wires occupy much of the area of the chip, and in nanometer CMOS technology, interconnects dominate both performance and dynamic power dissipation, as signal propagation in wires across the chip requires multiple clock cycles. (See Rent's rule for a discussion of wiring requirements for point-to-point connections).

Related Work & research on NoCs

Although NoCs can borrow concepts and techniques from the well-established domain of computer networking, it is impractical to blindly reuse features of "classical" computer networks and symmetric multiprocessors[citation needed]. In particular, NoC switches should be small, energy-efficient, and fast[citation needed]. Neglecting these aspects along with proper, quantitative comparison was typical for early NoC research but nowadays they are considered in more detail[citation needed]. The routing algorithms should be implemented by simple logic, and the number of data buffers should be minimal[citation needed]. Network topology and execution properties may be application-specific on MPSoCs.

Some researchers think that NoCs need to support quality of service (QoS), namely achieve the various requirements in terms of throughput, end-to-end delays, fairness,[1] and deadlines[citation needed]. Real-time computation, including audio and video playback, is one reason for providing QoS support. However, current system implementations like VxWorks, RTLinux or QNX are able to achieve sub-millisecond real-time computing without special hardware. This may indicate that for many real-time applications the service quality of existing on-chip interconnect infrastructure is sufficient, and dedicated hardware logic would be necessary to achieve microsecond precision, a degree that is rarely needed in practice for end users (sound or video jitter need only tenth of milliseconds latency guarantee). Another motivation for NoC-level quality-of-service is to support multiple concurrent users sharing resources of a single chip multiprocessor in a public cloud

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

computing infrastructure. In such instances, hardware QOS logic enables the service provider to make contractual guarantees on the level of service that a user receives, a feature that may be deemed desirable by some corporate or government clients Many challenging research problems remain to be solved at all levels, from the physical link level through the network level, and all the way up to the system architecture and application software. The first dedicated research symposium on networks on chip was held at Princeton University, in May 2007. The second IEEE International Symposium on Networks-on-Chip was held in April 2008 at Newcastle University.Research has been done on integrated optical waveguides and devices comprising an optical network on a chip (ONoC).

Peh et al. [19] and Mullins et al. [16] proposed speculative techniques for VC routers allowing a packet to simultaneously arbitrate for both VCA and SA giving a higher priority for non speculative packets to win SA; therefore reducing zero load latency in which the probability of failed speculation is small. This low latency, however, comes with the high complexity of SA circuitry and also wastes more power each time the speculation fails. Sophisticated extensions to IBR micro architectures have been proposed for improving throughput, latency, and power. For throughput, techniques like flit-reservation flow control, variable allocation of VCs, and express VCs [13] have been proposed.

As these designs are input-buffered, they are only able to multiplex arriving packets from their input ports across the crossbar switch, unlike our proposed router architecture which can shuffle incoming packet flows from all input ports and then onto the crossbar switch. Recently, Passas et al. [18] designed a 128 × 128 crossbar allowing connecting 128 tiles while occupying only 6% of their total area. This fact encourages us to build RoShaQ that has two crossbars while sharing cost expensive buffer queues.

Nicopoulos et al. [17] proposed ViChar, a router architecture which allows packets to share flit slots inside buffer queue so that can achieve higher throughput. Our paper manages buffers at coarser grain that is at queue-level rather than at flit-level, hence allows reusing existing generic queue design which makes buffer and router design much simpler and straightforward. Ramanujam et al. [21] recently proposed a router architecture with shared-queues named DSB which emulates an output-buffered router. The majority of state of the art on-chip router designs utilize input queuing buffers; we, however, can find in the literature a few output queuing router architectures [9]. If looking into the whole network picture, buffers at an output router port should act the same as input buffers of its downstream router. Depending on network load, RoShaQ can dynamically adapt to use the bypass paths or the shared queues. Only the initial, introductory paragraph has a drop cap.

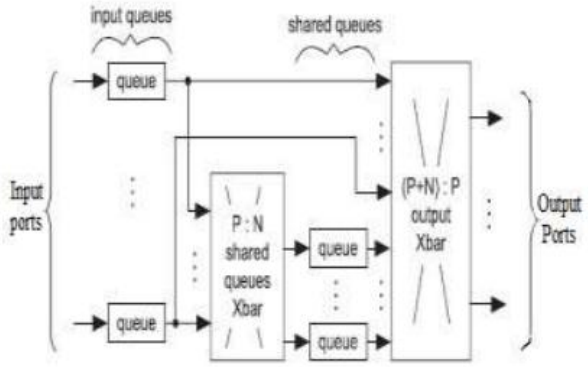
Concept of sharing Queues

For maximizing queue utilization, a router is designed to share all the queues [1],[2]. With this architecture, incoming packets from an input port has to be written to shared queue before reaching the output port even if the data is facing congestion or not. Also because this architecture has no buffer at the input port, when a packet needs to be forwarded, it has to send request to the destination router and wait for grant before sending data. Therefore the shared queue arbiter for each router is highly complicated because it must handle many external requests from multiple shared queues of all neighbouring routers and also the intra router requests.

To alleviate this latency, each input port is dedicated with one buffer queue and share all remaining queues, also providing it with a facility to bypass the shared queues as shown in the Fig.2. With this design, a packet from an input queues simultaneously arbitrates for both shared queues and an output port. If it wins the output port, it would be forwarded to the downstream router in the next cycle. Otherwise, that means having congestion at the corresponding output port., it can be buffered to the shared queues in the

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

same cycle. At low load, the network would have low latency because packets can bypass shared queues. While at heavy load, shared queues are used to temporarily store packets hence reducing their stall times at input ports that would improve the network throughput.



Concept of Sharing Queues.
P:number of router ports N: number of shared queues

RoShaQ Architecture RoshaQ, a router architecture with shared queues based on the idea of Fig.2 is shown in Fig.3. When an input port receives a packet, it calcualtes its output port for the next router. At the same time it arbitrates for both its decided output port and shared queues. If it receives a grant from the Output Port Allocator (OPA), it will advance to its output port in the next cycle. Otherwise, if it receives a grant to a shared queue, it will be written to that shared queue at the next cycle. Incase that it receives both grants, it will prioritize to advance to the output port.

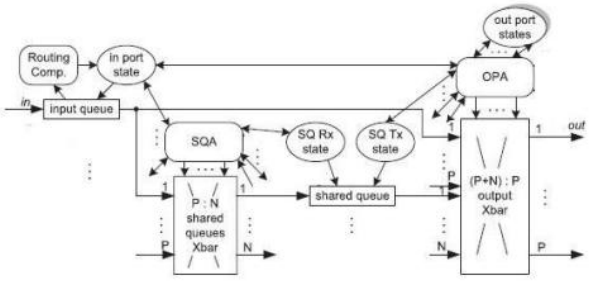


Fig: Generalized architecture of RoShaQ router. SQA:Shared Queue Allocator OPA:Output Port Allocator SQRx State:Shared Queue receiving/writing state SQTx State: Shared Queue Transmitting/reading State. P: Number of Router Ports N:Number of Shared Queues

RoShaQ's Properties

1) A network of RoShaQ routers is deadlock-free.

At light load, packets normally bypass shared queues, so RoShaQ acts as a WH router hence the network is deadlock-free [10]. At heavy load, if a packet cannot win the output port, it is allowed to write only to a shared queue which is empty or contains packets having the same output port. Clearly, in these cases RoShaQ is deadlock-free [11].

2) A network of RoShaQ routers is livelock-free.

In both OPA and SQA, each packet always has a chance to advance to the next router closer to its destination hence the network is free from livelock.

3) RoShaQ supports any adaptive routing algorithm.

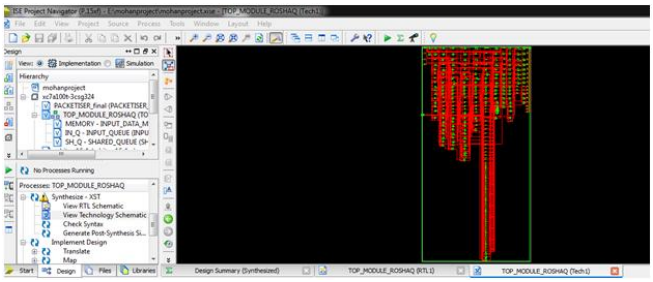
The output port for each packet is only computed at its input queue, not at shared queues. Therefore, any adaptive routing algorithm which works for WH routers also works for RoShaQ.

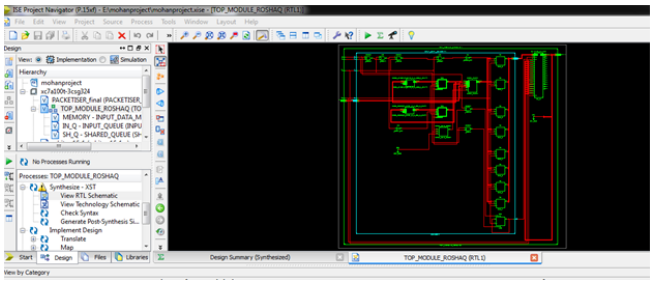
4) RoShaQ can be used for any network topology.

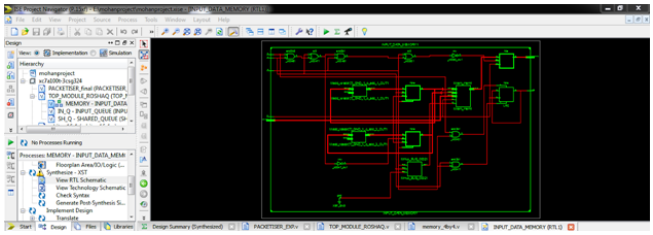
If all design details are hidden inside RoShaQ then it only has one buffer queue at each input port similar to a WH router. Therefore the number of RoShaQ's I/O ports can be changed to make it compatible with any network topology along with an appropriate routing algorithm.

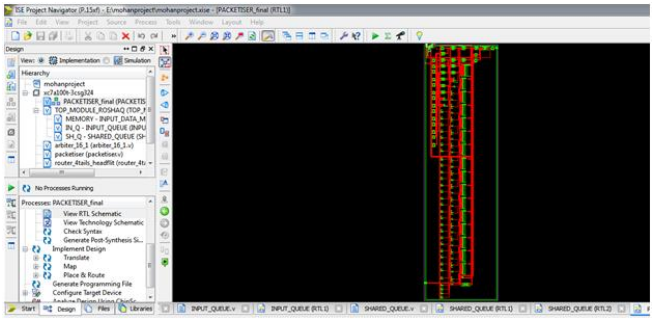
Experimental Results:

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal









Conclusion:

The various challenges faced by researchers in SoC design forced them to look for new alternatives which paved the way for Network-on-chip technology. The NoC is a vast and emerging research area that is still in its initial stages. The NoC area has a significant influence in the design of next generation SoC or multicore architectures. Router is a significant component of the NoC.It is a novel router architecture that allows sharing of multiple buffer queues for improving network performance. Sharing queues, makes using buffers more efficient hence is able to achieve higher throughput when the network load becomes heavy. On the other side, at light traffic load, this router achieves low latency by allowing packets to effectively bypass these shared queues.

References:

[1] Anh T. Tran and Bevan M. Bass, "Achieving High Performance On-Chip Networks With Shared Buffer Routers", IEEE Transactions on VLSI Systems,Vol 22, No.6,June 2014

[2] K.Priyanka, D.Gopinath & G. Rajaiah, Design and Implementation of Network Topology for Digitally Enhanced Networks, IJMETMR, Volume No: 2 (2015), Issue No: 7 (July) ,

http://www.ijmetmr.com/oljuly2015/KPriyanka-DGopinath-GRajaiah-91.pdf

[3] William J Dally and B.Towels, "Route Packets, not Wires: On-Chip interconnection Networks," DAC2001

[4] William J Dally, "Virtual channel Flow Control", IEEE Transactions on Parallel Distribution. Systems., Vol 3, No 2, Mar 1992

[5] T.Moscibroda and O.Mutlu, "A case for bufferless routing in Onchip Networks" ISCA, June 2009

[6]A.T. Tran, D.N. Truong and B.M.Baas, "A GALS many-core heterogenous DSP platform with source-synchronous on-chip interconnection network," in Proc. ACM/IEEE Int. NOCS, May 2009

[7] E.Begine, "An asynchronous power aware and adaptive NoC based circuit", IEEE J. Solid State Circuits, vol 44, Apr. 2009

[8] R.S.Ramanujam, V.Soteriou, B.Lin and L.S.Peh, "Extending the effective throughput of NoC's with Distributed Shared Buffer Routers", IEEETransactions on Computer-Aided Design Integrated Circuits Systems., Vol 30, no. 4, Apr 2011.

[9] H. Matsutani, Y. Hirata, M. Koibuchi, K. Usami, H. Nakamura, and H. Amano, "A multi-Vdd dynamic variablepipeline on-chip router for CMPs," in Proc. 17th ASP DAC, Feb. 2012, pp. 407–412.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

[10]J. L. Hennessy and D. A. Patterson, Computer Architecture: A QuantitativeApproach. San Francisco, CA, USA: Morgan Kaufmann, 2007.

[11] K. Latif, T. Seceleanu, and H. Tenhunen, "Power and area efficient design of network-on-chip router through utilization of idle buffers," in Proc. 17th IEEE Int. Conf. Workshops ECBS, Mar. 2010,pp. 131–138.

[12] L. Shang, L.-S. Peh, and N. K. Jha, "Dynamic voltage scaling with links for power optimization of interconnection networks," in Proc. Int. Symp. HPCA, Feb. 2003, pp. 91–102.