Broadcast- and Power-aware Wireless NoC for Barrier Synchronization in Parallel Computing

Hemanta Kumar Mondal[1], Rodrigo Cadore Cataldo[1], Cesar Augusto Missio Marcon[2], Kevin Martin[1], Sujay Deb[3], Jean-Philippe Diguet[1]
[1]*University of Southern Brittany (UBS), Lab-STICC, Lorient, France,* [2]*Pontifical University of Rio Grande do Sul, Brazil,* [3]*IIIT Delhi, India*
Email: [1]hemantam@iiitd.ac.in, [1]rodrigo.cataldo@acad.pucrs.br, [2]cesar.marcon@pucrs.br,
[1]{kevin.martin, jean-philippe.diguet}@univ-ubs.fr, [3]sdeb@iiitd.ac.in

***Abstract*—Efficient synchronization is one of the basic requirements of effective parallel computing. A key operation of the POSIX Thread standard (PThread) is barrier synchronization, where multiple threads block on a user-specified point of execution until all of them have reached it. Conventional architectures for broadcast operations limit the achievable performance benefits as synchronization is significantly affected due to critical path communications. This increases the network latency and degrades the performance dramatically. A Wireless Network-on-Chip (WiNoC) offers a promising solution to reduce the long distance/critical path communication bottlenecks of conventional architectures by augmenting them with single hop, long-range wireless links. In this paper, we propose a power-aware broadcast enabled WiNoC architecture to reduce the cost of broadcast operations for barrier-based applications. The proposed architecture reduces the barrier synchronization cost up to 43.97% regarding network latency under the PARSEC benchmarks. It also saves up to 80.49% idle-state power consumption in WIs for a 64-core system compared with the conventional WiNoC architecture without incurring significant overhead.**

Keywords—Broadcast operation, wireless network-on-chip, parallel computing, barrier synchronization, low power

I. INTRODUCTION

Parallel computing is currently being explored for High-Performance Computing (HPC) platforms for scientific research, automobile, cloud computing and data center applications. Network-on-Chip (NoC) architecture can be employed as a communication infrastructure in parallel applications, and it improves the performance of the system significantly [1]. One of the primary components of parallel applications is synchronization, a necessity for sharing information. Out of all the synchronization primitives, barrier synchronization faces significant challenges during broadcast operations due to critical path communications. Conventional NoC architectures support broadcast operations in the form of multiple unicast transmissions, which results in significant system performance penalties concerning network latency and energy consumption overhead.

The hardware-level barrier synchronization for NoC-based systems has already been explored for sending broadcast messages to all involved cores. A transmission-line based hardware barrier implementation is proposed in [2] to allow a single chip-spanning transmission line network to support many barriers simultaneously. However, this kind of architecture adds significant routing overhead throughout the network. Apart from this, there is also a fanout constraint. It increases with system size, which is not a feasible solution for scalable barrier synchronization. Many authors [3][4][5][6][7] have proposed a hybrid tree-based all-to-all barrier for NoC-based system, which improves the performance by avoiding the off-centered barrier core when compared to a master-slave and tree barrier. The number of processing elements in a system is increasing rapidly, and this trend will undoubtedly continue into the future. Therefore, the distance between core-to-core and length of the critical path is increasing rapidly. One of the promising and CMOS-compatible solutions for heavy broadcast traffic is the single hop, long-range wireless links based NoC architecture. WiNoC with omnidirectional setup plays a vital role in providing efficient broadcasting capabilities for parallel computing on HPC platforms.

Most existing WiNoC architectures use omnidirectional antennas along with token passing protocol to access wireless medium, where only a single wireless pair can communicate at a time. The primary component of WiNoC is the wireless interface (WI) responsible for handling the wireless communication. The major components of a WI are Mixers (modulator and demodulator), Low Noise Amplifier (LNA) and Power Amplifier (PA). The power consumption by these components is shown in Fig. 1 [8]. Recently, the number of broadcast and unicast messages in a cache coherence protocol are explored in [9]. The percentage of broadcast messages found is very small in case of cache coherence protocol (~5%) [9]. Similarly, for barrier synchronization, large portion of communications is based on unicast. Therefore, it is essential to reduce the static/idle-state power consumption of the WIs. As the number of WIs increases with system size, the idle-state power consumption is a major factor in highly parallel architectures.

To provide efficient barrier synchronization for effective parallel computing, we propose a power-aware broadcast

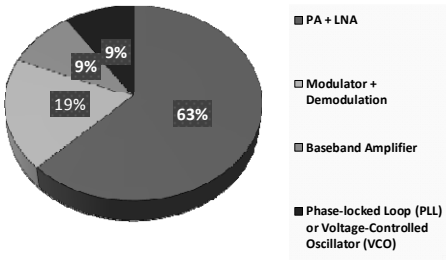

Fig. 1. Power budget of a WI. Specifications: Distance- 20mm, Data rate- 16Gbps, Modulation- OOK, Technology node- 65nm, Total power- 32mW [8].

978-1-5386-1491-4/18/$31.00©2018 IEEE

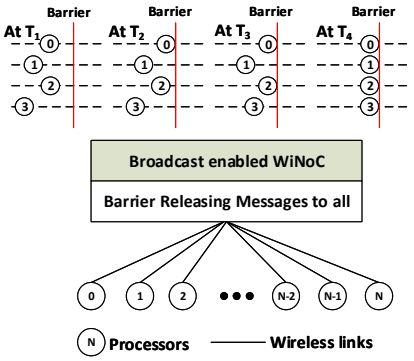

Fig. 2. Representation of broadcast enabled parallel computing with barrier synchronization.

enabled WiNoC architecture to reduce the cost of broadcast operations significantly. Fig. 2 shows that, at time T_4, messages releasing threads on a barrier are broadcasted to all involved cores using wireless links to improve the network latency. To make a more power-efficient wireless interface, we also explore the power gating technique with WIs to reduce idle-state power. Although our architecture must be efficient for barrier synchronization, these events occur far between, as they demand milliseconds of application computation. From the hardware perspective, delays of milliseconds can result in millions of inactive cycles, which can be timely exploited by power gating the WIs. We demonstrate the proposed work under real applications and show the benefits over existing.

The major contributions of this work are as follows:

1. The proposal of a power-aware broadcast enabled WiNoC architecture to improve the performance and power saving of WIs for parallel applications.
2. The implementation and discussion of the scalability of a partial power gated WI. To the best of our knowledge, none of the existing works explore power gated WI for parallel computing running wireless/ wired based NoC.
3. The validation of our architecture under the PARSEC benchmarks [10] and comparison with existing architectures. To evaluate the wireless interconnect-based proposed architecture, we have modified the existing Noxim simulator [11] to handle broadcast traffic.

The remainder of this paper is organized into the following sections. Section II briefly describes related works. The proposed architecture including synchronization, broadcast and power-aware WI is discussed in section III. Section IV presents the results of performance evaluation and section V concludes this work.

II. Related Works

Many research works have investigated the barrier synchronization in parallel computing using NoC-based architecture. Software-based barrier synchronization implementations increase the latency significantly with system size due to serialization of the barrier operation. To overcome this, a hardware-based barrier synchronization is implemented to support the multiple thread groups, where each has its barrier. The proposed method reduces the latency for serialization compared with the traditional planner wired. However, transmission-line based approach faces several challenges in manycore system [2]. To communicate with every core, transmission line needs to spread the entire chip area and require excessive branching points to connecting cores. It is not an efficient solution due to cross-talk, inter-channel interference for long transmission lines problems, large fanout and then power consumption for large-size systems. A hardware-based barrier synchronization is implemented using G-line based network to allow for efficient signaling of barrier arrival and departure [3]. A hybrid tree-based all-to-all barrier for NoC-based manycore system processors is explored in [5] to improve the performance by avoiding the off-centered barrier core. However, they are all based on multiple unicast packets, which is not efficient in parallel computing. The latency is also destination-dependent, so the message delivery time is unbalanced. The WiNoC architecture is one of the potential solutions to handle efficiently long-distance packets. The WiNoC architecture has been explored to incorporate multicast support for cache coherency protocols [12]. Recently, multiple WiNoC architectures have been explored for efficient on-chip communications. In directional wireless NoC, multiple wireless links can operate at the same time without interference [13]. A hierarchical NoC architecture with zigzag antennas and millimeter wave transceivers are proposed to design an mm-wave wireless NoC in [14]. The transceiver is an integral part of WIs for WiNoC architectures, and power optimized designs are required for achieving energy-efficient systems. Power gating is an effective technique to save idle-state power and has been explored for NoC infrastructures. Different distributed controller implementations for power gating operating at various granularity levels are introduced to save leakage/static power [15][16].

In this work, we propose a power-efficient WiNoC architecture to reduce the cost of broadcast operations for barrier synchronization. We present a detailed performance evaluation of the proposed WiNoC architecture and explore the performance overheads and associated trade-offs for realizing the proposed WiNoC architecture.

III. Proposed Architecture

The WiNoC consists of a large number of tiles. The

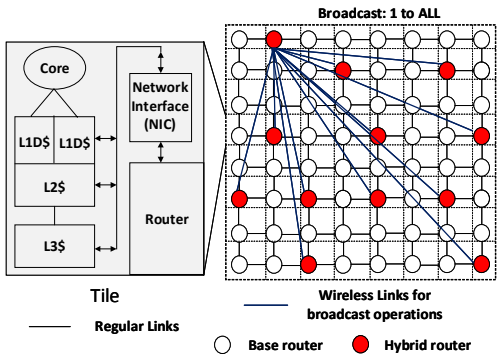

Fig. 3. Broadcast enabled WiNoC architecture.

contents of the tile are a router, a core, caches (L1, L2, and L3), a network interface controller and links (hops). A conventional NoC architecture consists of routers attached to message source/sink components; with all routers interconnected by wires in a specific topology. WiNoC architectures rely on strategically and optimally placed WIs at some routers, providing long-distance wireless communication to improve NoC performance [14]. A broadcast enabled WiNoC architecture is shown in Fig. 3. The WiNoC topology, with Base Routers (BRs) and Hybrid Routers (HRs), is shown in Fig. 3. The hybrid routers are combinations of BRs and WI. The number of WIs can be varied with system size. In our topology, the optimal number of WIs are placed based on simulated annealing algorithm, which is discussed in [14]. In a conventional NoC, a dedicated synchronization controller is usually integrated with each core to control barrier messages [5]. The barrier control registers in a synchronization controller handle barrier arriving and releasing messages between involved cores. They record the number of cores that have arrived at the barrier and activate the barrier releasing flag to all involved cores when all the cores have reached the barrier. Multiple types of message delivery (e.g., all-to-all, master-slave, butterfly and tree) can be implemented. However, the performance of the network is dramatically affected using these conventional approaches. It is due firstly to longest core-to core path/critical path, and secondly to the unicast implementation of broadcast messages. To minimize these, we explore the WiNoC architecture for efficient broadcast communications.

A. Barrier Synchronization

The PThread standard is one of the most well-known interfaces used for parallel computing. For developers, it offers multiple procedures to synchronize data between application threads or the threads themselves. For this work, we chose to improve the barrier synchronization procedure because it has the most to benefit for our architecture, as it releases threads by sending the same message to multiple, up to all, destinations. Barriers are responsible for synchronizing threads to a user-specified location in the application. When all participating threads reach the specified location, they can continue to execute; otherwise, they are blocked. Therefore, we have changed the releasing procedure to generate a single broadcast message instead of multiple unicast ones.

TABLE I. NUMBER OF SYNCHRONIZATION PROCEDURE CALLS DURING THE EXECUTION OF BODYTRACK AND STREAMCLUSTER APPLICATIONS FOR A NOC-BASED CMPS

Application	Type	Events per number of Threads		
		16-core System	*32-core system*	*64-core System*
Bodytrack	Barrier	2,112	4,288	17,788
	Condition	447	750	4,264
	Mutex	9,000	10,472	37,818
Streamcluster	Barrier	208,064	364,480	728,960
	Condition	381	802	1,274
	Mutex	510	1,054	2,142

The use of synchronization procedures provided by PThread is contingent on the design of the application. The *PARSEC* benchmark [10] is a collection of applications intended for next-generation shared-memory architectures that employ the PThread standard. For the twelve applications available, we analyze two of them as a representative of the PARSEC workload: *Bodytrack,* for a small number of broadcast messages, and *Streamcluster,* for a large number of broadcast messages. *Bodytrack* is a computer-vision application that tracks a 3D pose of a mark-less body. *Streamcluster* is a data-mining application that solves the online clustering problem for a stream of input points. Both of these applications use multiple synchronization procedures; however, as can be seen from Table I, the barrier procedure has the most requests. Therefore, we are able to exploit the broadcast enabled network for a 64-core system. When considering all data traffic generated by the application, *Streamcluster* requires more than 5% of the overall traffic for broadcast messages, as shown by *Karkar* et al. [9]. Performance is affected by these messages dramatically. It increases the congestion and provides poor quality of service. It also increases the power consumption due to the retransmission of the same packet. CMOS compatible wireless emerging interconnect offers many significant advantages to overcome these drawbacks of conventional NoCs. Hence, we consider the WiNoC architecture to reduce the number of hops required for communication and to implement broadcast by means of simultaneous receptions. Once the packet arrives at the WI, the packet is then transmitted to neighboring cores by tree-based load-balanced paths using the protocol described in *Section III.B.c*. Experimentally, we found that the proposed architecture reduces the network latency significantly during broadcasting for barrier synchronization. On the other hand, to reduce power consumption, we employ our power gating scheme with unused WIs.

B. Power-aware WI

In this section, we discuss the partially power gated WI to reduce the idle-state power consumption. This section consists of the communication infrastructure and protocol, the controller and the power gating scheme used on WIs. The WiNoC architecture is explored for manycore systems [13][14].

a) Communication Infrastructure

The primary components of the communication infrastructure are on-chip antenna and transceiver. A metal zigzag antenna is adopted from [17]. In this work, we employ a

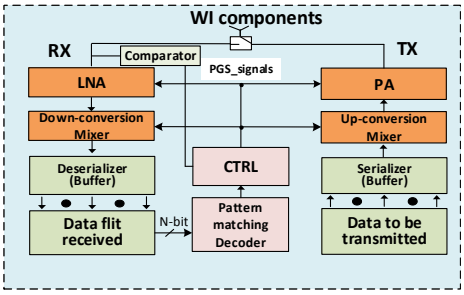

Fig. 4. Power gated LNA, PA and Mixers with controller.

non-coherent On-Off keying (OOK) modulation scheme based transceiver for WiNoC from [14]. The wireless channel is shared between all WIs and token passing mechanism with round robin arbitration is used to provide the access to the shared wireless medium. Hence, only a single broadcast operation is possible at the same time. As illustrated in Fig. 4, the transmitter (TX) circuitry consists of an up-conversion mixer and a power amplifier (PA). On the receiver (RX) side, direct-conversion topology is adopted, consisting of a low noise amplifier (LNA), a down-conversion mixer and a baseband amplifier. The OOK transceiver design keeps the additional overhead added by WIs in WiNoC architecture to a minimum. However, of all the components in WI, Mixers, PA and LNA consume more than 80% of total WI power. Power gating is employed with each WI to reduce the idle-state/static power consumption.

b) Controller and Power Gating Operation

When WIs are not in used, the controller sends them into sleep mode to reduce the power consumption using power gating switches as illustrated in Fig. 4. These switches are controlled by a distributed light-weight controller (CTRL) based on valid signals from a comparator [18]. The associated comparator at all WI detects a valid received signal. Then pattern matching decoder process the WI address bits. For broadcast, the header flits employ a unique address pattern. The controller controls the broadcast and unicast operations based on communication traces from cores.

The proposed method works when WIs are not active. As seen in the previous section, the percentage of broadcast messages (~5%) is very small in case of cache coherence protocol [9]. Similarly, for barrier synchronization, the percentage of broadcast operations is also small. Hence, the significant amount of (~95%) communications is based on unicast. During unicast, only a single frequency channel can communicate based on the token passing protocol. Therefore, there is a huge scope to reduce the idle-state power consumption at WIs. Applying this proposed method saves significant amount of power consumption. This WiNoC architecture can be used in manycore architectures for scalable and efficient on-chip communications.

c) Communication Protocol

For deadlock-free seamless communications in WiNoC architecture, we have adopted the routing algorithm from [19], which is a combination of South-Last routing [19] and XY routing to achieve an efficient deadlock-free routing for wireless links. In [19], the South-Last routing algorithm has been proposed for long-range single hop wired links

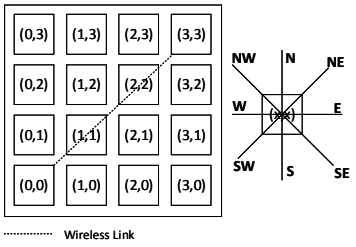

Fig. 5. Example of a 4×4 mesh NoC system augmented with wireless link with NE-SW direction.

communication in the network. The concept of wireless links is similar to that of long-range links, which is used to replace the multi-hop communications in large-scale networks. The South-Last/North-Last algorithm is a turn model routing strategy that avoids cyclic dependencies as discussed in [19]. In a conventional NoC, we have implemented a broadcast mechanism in the form of tree-based load-balanced paths inspired by the Whirl method [20]. The route for a packet is decided by the source node and destination wireless interface. This scheme balances the link loads for broadcast and ensures non-duplicate packet reception.

The traditional implementation of a barrier scheme selects a master node for collecting barrier arriving messages from all other cores and for broadcasting barrier releasing messages to them. This master node transmits broadcast messages to the nearest WI and its neighbors using South-Last routing. Otherwise, XY routing is used as a default for conventional routing. The nearest WI broadcasts barrier releasing messages to all nearest destination WIs using an omnidirectional setup. Fig. 5 demonstrates how our proposed routing algorithm avoids the deadlock and livelock situations. Fig. 5 also presents an example of a 4×4 mesh NoC system augmented with wireless link, which is connected between two routers with the NE-SW direction. The routing strategy follows traditional XY routing at all basic routers without a wireless link. XY routing avoids S-to-E and S-to-W turns to ensure that routing is deadlock free. However, the wireless link may take a turn from one of the middle directions NE, NW, SE, or SW to one of the main directions N, S, E, and W and hence may introduce 180° turns. Hence, we need to forbid extra turns to prevent deadlock cycles caused by the wireless links. In our routing, we forbid all turns from the south direction along with the ones forbidden by XY routing. To avoid such cycles, 180° turns from the south to west directions (180° turns from negative directions) are also prohibited, thereby providing deadlock-free routing with wireless links.

IV. Performance Evaluation

In this section, we explore the performance benefits of parallel applications that employ barrier synchronization. We discuss the implementation of a power-aware WiNoC architecture along with power gated WIs in detail.

TABLE II. Simulation Setup

Setup	Components	Configuration
Details of system architecture	CPU	ALPHA ISA cores, Out-of-order cores, 2.5GHz
	L1 Cache	64KB, 4-way, LRU policy, 64B line, 1 cycle latency
	L2 Cache	256KB, 8-way, LRU policy, 64B line, 10 cycle latency
	Cache Coherence Protocol	MESI
Details of network architecture	Topology	8×8 Mesh NoC, 8×8 Mesh WiNoC
	Routing	XY for baseline, South-Last for hybrid router node
	Pipeline	4 stages
	Flit size and packet size	32 bits and 64 flits
	Workload	PARSEC: *Bodytrack* and *Streamcluster*

A. Simulation Setup

To bring out characteristics of the proposed WiNoC architecture in the presence of real workloads, we have considered the *PARSEC* Benchmarks [10]. The workloads are simulated using Gem5 [21] and our internal simulation tool for parallel applications for collecting traces for broadcast based communication in the former and fast trace-based simulation in the latter. The traces are executed on a modified version of the Noxim simulator [11]. We consider a system size of 64 cores for our experiments, which is representative of current manycore technology trends. The width of all wired links is the same as the flit size (i.e., 32 bits). The NoC switches are driven with a clock of frequency 2.5GHz. The power gated components are implemented using Cadence tools at 65nm technology node. The controller is synthesized from RTL level design with Synopsys Design Compiler using 28nm technology node. To obtain area, power, and delay of sleep transistors, we use Cadence tools. The summary of simulation setup is presented in Table II.

B. Network Latency Reduction

Fig. 6 summarizes performance regarding network latency under two *PARSEC* benchmarks: *Bodytrack* and Streamcluster [10]. In this work, we have considered these two applications to validate our proposed architecture, as *Bodytrack* uses a smaller percentage of broadcast messages, and *Streamcluster* uses a higher percentage of broadcast messages over the total communication messages (*Streamcluster* has the highest barrier usage on the PARSEC benchmarks [10]). In case of *Bodytrack*, communication traces are extracted for two different inputs such as medium and large. It employs 4 barriers for the application implementation, so 4 sources, and uses 3 types of threads to sort inputs of threads. *Streamcluster* employs a single barrier for the application implementation, hence, only a single source. For both applications, only when 63 cores have reached the barrier it can be released. The release procedure is done with broadcast. Hence, every 64th barrier event is a broadcast. Therefore, the total number of broadcast messages is 210, 267 and 11390 for *Bodytrack_medium*, *Bodytrack_large* and *Streamcluster_medium*, respectively. From the figure, it can be observed that WiNoCs have lower network latency as compared with conventional mesh wired NoC architectures. It can be seen that latency for *Bodytrack* and *Streamcluster* is high for mesh architectures for 64-core system. The proposed architecture reduces the network latency

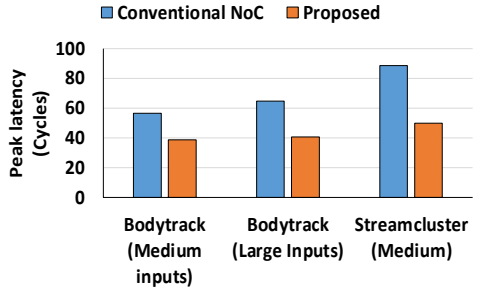

Fig. 6. Peak packet latency over real benchmarks of NoC architectures.

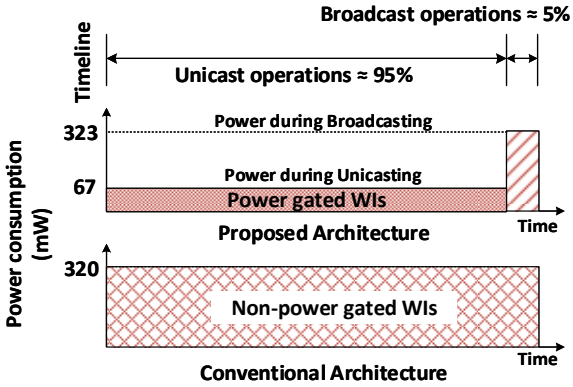

Fig. 7. Power gated WIs vs. non-power gated WIs.

up to 31.28% for *Bodytrack* with medium inputs, 36.92% for *Bodytrack* with large inputs and 43.97% for *Streamcluster* benchmarks. From these results, it is clear that WiNoC architecture achieves significant performance improvements over conventional mesh wired NoC.

C. Implementation Cost of Power-aware WI

Power saving is an essential feature of our implementation; thus, we discuss its overheads in WiNoC architecture. WI uses a power gating switch to keep idle-state power to a minimum. Each power gating switch consumes 3.16μW of power. The power consumption of additional components (i.e., comparator, controller and power gating switch) is 0.30mW. In case of a hybrid router, mixers, LNA, and PA consume 26mW of the total 32mW of the transceiver power [8]. The total power consumption by the transceiver components along with additional components during active mode is 32.30mW. During sleep mode, power consumption in WI comes down to 6.30mW. The proposed scheme saves up to 80.49% of the power for each WI in sleep mode. Usually, total power consumption by conventional WiNoC architecture or non-power gated WIs is 320mW during active mode as shown in Fig. 7. However, in case of barrier synchronization, the percentage of broadcast messages is ~5% (utilized all the WIs), and the rest ~95% is based on unicast (i.e., only a single WI pair can communicate), which is discussed in *Subsection III.A*. The proposed architecture consumes around 67mW of power including controller overheads for ~95% of the total simulation cycles, whereas non-power gated WIs consume 320mW of power for throughout the application. In case of our proposed architecture, all WIs utilize only ~5% for the total simulation cycles and consume 323mW power including the controller overheads. As can be seen in Fig. 7, the proposed architecture

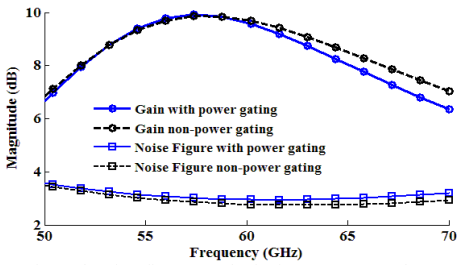

Fig. 8. Gain and noise figure analysis for power gated LNA and non-power gated LNA.

TABLE III. SUMMARY OF PROPOSED AND EXISTING WORKS

Ref.	Approaches	Saving	Penalty
[2]	Transmission-line based broadcast network	Worst-case latency: 4ns to 10ns	0.07% of total metal area overhead
[5]	Tree-based all-to-all barrier synchronization	Saved 20% during off-centered	0.2% area and 1.3% power overheads
[7]	OrthoNoC/ WiSync	Latency improvement: 30%	Less than 5% area overhead
This	Broadcast enabled WiNoC	Latency reduction: 43.97%; Idle-state power per WI: 80.49%	Less than 1% area overhead; only 10 WIs for 64-core system

saves a significant amount of power consumption in WIs as compared with the conventional WiNoC. In addition, the controller unit occupies $1.72\times10^{-3}mm^2$. The area overhead of the comparator circuit is $0.46\times10^{-3}mm^2$.

In the last paragraph, we discussed the implementation of power gated LNA only. We follow the exact same steps for the power gated PA and mixer. All components are designed to operate at a frequency of 58 GHz. The gain and noise figure performance of power gated LNA and non-power gated LNA are illustrated in Fig. 8. The peak gain of power gated LNA remains the same. Performance deviations of all the power gated components are dependent on the size of the PMOS switch. For this experiment, the size of the PMOS transistor is 10 times larger than NMOS' size for LNA/PA/Mixers design. With proper sizing, performance deviations are minimized with the cost of area penalty.

D. Summary of Proposed and Existing Works

In this section, we present a summary of proposed and recently proposed barrier synchronization for NoC-based systems in Table III. Our proposed method saves the cost of broadcast operation in barrier synchronization up to 43.97% using wireless interconnects, reducing significantly the latency as compared with the conventional interconnect architectures. Our proposed architecture reduces idle-state power consumption in WIs up to 80.49%, which is compared with the results in [18].

V. CONCLUSION

In this paper, we propose a power-aware WiNoC architecture to maintain the parallelization gains in parallel applications by reducing barrier synchronization latency significantly. As can be observed from the results, the small percentage of broadcast messages has great impact on system performance regarding network latency. In addition, we also employed the power gating method with WIs to reduce the idle-state power consumption. The proposed architecture is validated under the PARSEC benchmarks. We observed that the proposed WiNoC architecture reduces network latency up to 43.97% over conventional architectures. The proposed method also saves the power consumption in WIs up to 80.49% per WI compared with the conventional WiNoCs.

REFERENCES

[1] Abadal, S., et al., "Broadcast-enabled massive multicore architectures: A wireless RF approach." IEEE micro, 35(5), 52-61, 2015.

[2] Oh, J., et al., "TLSync: support for multiple fast barriers using on-chip transmission lines." In 38th Annual International Symposium on Computer Architecture (ISCA), 105-115, 2011.

[3] Abellán, J.L., et al., "Efficient hardware barrier synchronization in many-core cmps." IEEE Transactions on Parallel and Distributed Systems, 23(8), 1453-1466, 2012.

[4] Villa, O., et al., "Efficiency and scalability of barrier synchronization on noc based many-core architectures." In conference on Compilers, architectures and synthesis for embedded systems, 81-90, 2008.

[5] Wei, Z., et al., TAB barrier: Hybrid barrier synchronization for NoC-based processors. In IEEE International Symposium on Circuits and Systems, 409-412, 2015.

[6] Yang,J.-S., et al. "Designing tree-based barrier synchronization on 2D mesh networks. *IEEE TPDS*, vol. 9, no. 6, 526-534, 1998.

[7] Abadal, S., et al. "OrthoNoC: A Broadcast-Oriented Dual-Plane Wireless Network-on-Chip Architecture." IEEE Transactions on Parallel and Distributed Systems, 2017.

[8] Yu, X., et al., "Architecture and design of multichannel millimeter-wave wireless NoC." IEEE Design & Test, 31(6), pp.19-28, 2014.

[9] Karkar, A., et al. "A survey of emerging interconnects for on-chip efficient multicast and broadcast in many-cores. IEEE Circuits and Systems Magazine, vol. 16, 58–72, 2016.

[10] Bienia, C. et al. "The PARSEC Benchmark Suite: Characterization and Architectural Implications". PACT, 72-81, 2008.

[11] Catania, V., et al., "Cycle-Accurate Network on Chip Simulation with Noxim." ACM Transactions on Modeling and Computer Simulation (TOMACS), 27(1), p.4, 2016.

[12] Duraisamy, K., et al. "Multicast-aware high-performance wireless network-on-chip architectures." IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 25(3), pp.1126-1139, 2017

[13] Mondal, H.K., et al., "Interference-Aware Wireless Network-on-Chip Architecture Using Directional Antennas," in *IEEE Transactions on Multi-Scale Computing Systems*, vol. 3, no. 3, pp. 193-205, 2017.

[14] Deb, S., at el., "Design of an energy efficient CMOS compatible NoC architecture with millimeter-wave wireless interconnects." IEEE Transactions on Computers, 99(1), p.1, 2012.

[15] Mondal, H. K. et al., "Adaptive multi-voltage scaling in wireless NoC for high performance low power applications," *Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pp. 1315-1320, 2016.

[16] Mondal, H.K. et al., "An energy efficient wireless Network-on-Chip using power-gated transceivers," *2014 27th IEEE International System-on-Chip Conference (SOCC)*, pp. 243-248, 2014

[17] Chen, L., D. et al., "Power punch: Towards non-blocking power-gating of NoC routers," IEEE International Symposium on High-Performance Computer Architecture, 378,389, 2015.

[18] Parikh, R., et al. Power-aware NoCs through routing and topology reconfiguration," ACM/EDAC/IEEE Design Automation Conference (DAC), 1-6, 2014.

[19] Floyd, B.A. et al.., "Intra-chip wireless interconnect for clock distribution implemented with integrated antennas, receivers, and transmitters." IEEE Journal of Solid-State Circuits, 543-552, 2002.

[20] Mondal, H.K., et al., "Adaptive Multi-Voltage Scaling with Utilization Prediction for Energy-efficient Wireless NoC." in IEEE Transactions on Sustainable Computing (TSUSC), vol. 2, 382-395, 2017.

[21] Ogras, U.Y. et al., "It's a small world after all": NoC performance optimization via long-range link insertion. IEEE Transactions on very large scale integration (VLSI) systems, 14(7), 693-706, 2006.

[22] Krishna, T., et al. "Towards the ideal on-chip fabric for 1-to-many and many-to-1 communication." In 44th Annual IEEE/ACM International Symposium on Microarchitecture, 71-82, 2011.

[23] Nathan B., et al. "The Gem5 Simulator" ACM SIGARCH Computer Architecture News, 2011.