

Escalator Network for a 3D Chip Stack with Inductive Coupling ThruChip Interface

Akio Nomura, Yusuke Matsushita, Junichiro Kadomoto,
Hiroki Matsutani, Tadahiro Kuroda, Hideharu Amano
Graduate School of Science and Technology
Keio University
Yokohama-shi 223-8522 Japan

Received: February 14, 2017
Revised: May 2, 2017
Revised: July 6, 2017
Accepted: July 21, 2017
Communicated by Koji Nakano

Abstract

A wireless inductive coupling ThruChip Interface (TCI) is a flexible system-in-package (SiP) technique which enables to build a powerful interconnection network between stacked chips. For easy use of TCI, we have developed intellectual properties (IPs), and proposed an interconnection network which can make the use of IPs. We also developed a real chip embedded the IP, and evaluated the performance. By stacking multiple chips with the proposed IP, an inter-chip network with link-to-link flow control by piggyback control is established. The new proposed escalator network which uses piggyback of the credit packets outperforms the ring network used in the first prototype by 28%-59% in terms of throughput. The performance overhead by the piggyback control was less than 3%-4% of that without control messages.

1 Introduction

Mobile devices require wide range of performance, various functions, and low energy consumption because of their development and widespread applications. On the other hand, because of growing non-recurring engineering cost of recent advanced process technology, developing a single System-on-Chip (SoC) for each target application has become costly. Instead, combining small dedicated chips; microprocessors, memory modules and accelerators with a technology of System-in-Package (SiP) is believed to be an alternative way to easily integrate a huge number of transistors with low-cost.

Among various kinds of SiP structure, 3D chip stacking is advantageous for such mobile devices because of its high implementation density. For building 3D chip stacking systems, interconnections of chips are classified into two categories: wired and wireless. Wired interconnections (e.g. wire bonding, micro-bump bonding[1], and through-silicon-via (TSV)[2]) have been mature techniques. Especially, TSV is advantageous in terms of high bandwidth and small footprint, and has been utilized in memory systems including HMC and HMB. 2.5D implementation which combines the TSV and micro-bump bonding techniques has been used for building a large scale FPGAs[3]. However, such wired techniques require special manufacturing processes, and tend to be expensive. Furthermore, the replacement, addition, and deletion of the stacked chips are not allowed once they are connected with TSV.

On the other hand using wireless interconnections the structure of the stacked chips is easily changed since connections between these chips are physically contact-less. Wireless interconnections

are classified into capacitive coupling [4] and inductive coupling [5, 6, 7]. Since the use of capacitive coupling is limited to face-to-face stacking of two chips, inductive coupling which can interconnect more than two chips has been mainstream of wireless chip stacking. To cover a wide range of performance requirements with low-cost, it is important to make the system more scalable and flexible. We have thus focused on the inductive coupling interconnections in our research. By using this technique, various heterogeneous systems can be built just by stacking various types of chips; microprocessors, memory modules and accelerators which provide a standard wireless inductive coupling interface. We call such systems building block computational systems.

We developed the first prototype of building block computational system called Cube-1 [8]. It consists of a low-power microprocessor called Geysers [9] and coarse grained accelerators called CMA-Cube [10] connected with the wireless inductive coupling Thru-Chip Interface (TCI). In Cube-1, a uni-directional ring network with bubble-flow control [11] is formed just by stacking chips. It can guarantee deadlock freedom without virtual channels (VCs), and the link-to-link flow control is not needed by slotting the injection. It was planned that the required trade-off between performance and cost could be achieved by changing the number of accelerator chips for the target application.

However, we found several problems on the interconnections which prevented Cube-1 to work as the initial plan. First, the TCI was not well designed as the form of Intellectual Property (IP), combining it into the digital design flow made the chip development difficult. For this reason, other types of accelerators could not be developed. Second, although two-chip stack worked, more than three-chip stacking system did not stably work because of the problem on inductor location for building a ring network. Also, VCs are required by the operating system not for resolving deadlock but for controlling different classes of messages.

By making use of this experiences, the next prototype called CUBE-SOTB is under development. It consists of a host microprocessor Geysers CUBE-SOTB and several accelerators which provide the same TCI interface. All chips use novel silicon on insulator (SOI) technology called silicon-on-thin-buried oxide [12][13] that can work with low supply voltage. To cope with problems on the uni-directional ring network, a bi-directional escalator network is proposed for CUBE-SOTB. A credit based flow control is introduced to implement a number of virtual channels, and the data link is used for credit packets as well as data packets. By introducing the bi-directional network, the whole part of the network can be packaged into an intellectual property (IP) including a router, a link controller and the TCI physical layer.

The main contribution of this paper is summarized as follows:

- The IP including a router, a link controller and the TCI physical layer is developed.
- An escalator network which is based on the IP is proposed and its performance is evaluated both with network simulation and instruction level simulation.

Compared an earlier stage of our research [14], this paper includes the integrated IP structure, comprehensive evaluation of the escalator network with application programs, and the most recent real chip evaluation results.

The rest of the paper is organized as follows. Section 2 introduces inductive coupling TCI. The ring and escalator network is described in Section 2.2 and 3 with an integrated IP of TCI. Then, the evaluation results by simulation are shown in Section 5. Finally, the chip implementation and evaluation results are shown in Section 6, and the last section concludes the paper.

2 Inductive Coupling TCI and Cube-1

2.1 Inductive Coupling Channels

Inductive coupling ThruChip Interface (TCI) uses square coils implemented with common metal layers. An inductive coupling channel is formed between two chips by stacking a transceiver coil on the receiver coil of a different chip, as shown in Fig. 1. Paired coils, one for the clock and the other for the data, are implemented as a channel. A high frequency clock (1 - 8 GHz) is generated by a ring oscillator, and the serialized data are transferred in synchronization with the high frequency

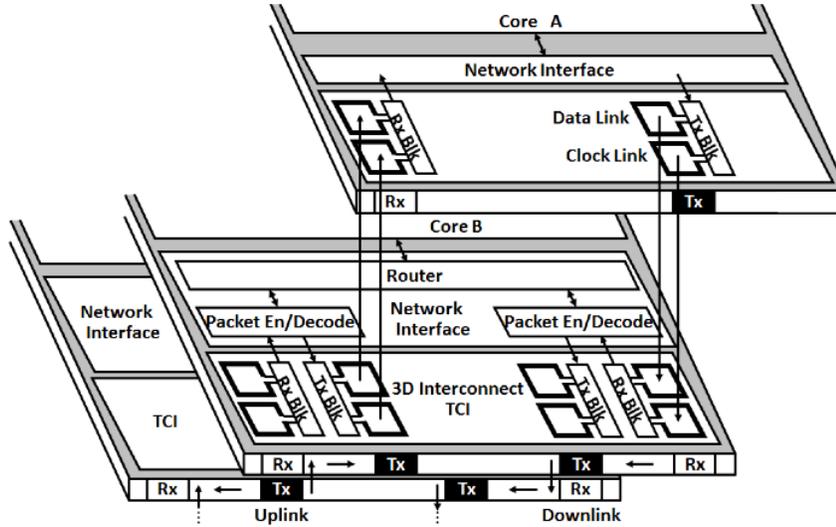


Figure 1: TCI with transmitter and receiver [8]

clock directly through the driver. The driver and coil pair for sending data is called the Tx channel or TX, while the receiver and coil pair for receiving is called the Rx channel or RX. In figures which show cross-cutting view of a chip, they appear as boxes marked TX or RX.

Although TCI requires a certain amount of logic to form a link between two chips, it has the following benefits.

- 8 Gbps of data at maximum can be transferred with low energy dissipation (0.14 pJ / bit) and low bit-error rate ($BER < 10^{-12}$) [7]
- The number of chips which can be stacked is not limited by TCI but a physical environment.
- Since testing is done before chips are stacked.
- Since TCI is electrically contact-less, electro-static-discharge (ESD) protection device is unnecessary.
- Since the coil is implemented with the common wire layers of a CMOS process, no extra process is needed. Although a coil occupies a large footprint, standard cells used for digital circuits can be implemented inside the coil.

2.2 Uni-directional ring for Cube-1

The problem of the TCI is a large footprint for coils, drivers and receivers. It means that a link should be used for high speed data transfer rather than handshake signal for the flow control. In order to form a high speed network without using link-to-link flow control, we adopted a uni-directional ring network with a bubble-flow control, as shown in Fig. 2, in our first prototype of building block computational system, Cube-1.

Each chip has two routers each of which has a TX and a RX, and one of the routers is connected with a CPU core or a memory core of an accelerator on each chip. By placing the TX channel of an upper chip on the RX channel of a lower chip (the RX channel of an upper chip on the TX channel of a lower chip), an uplink (a downlink) is formed to build a uni-directional ring network. The ring network can be theoretically formed in any number of chips, by giving each chip chip-ID and whether the chip is stacked on the top/bottom or not. In order to place the TX coil on the RX coil, each chip must be shifted a certain distance and stacked as shown in Fig. 2. This space is useful to provide bonding wires required for power supply and debugging signals. In Cube-1, 35bit data

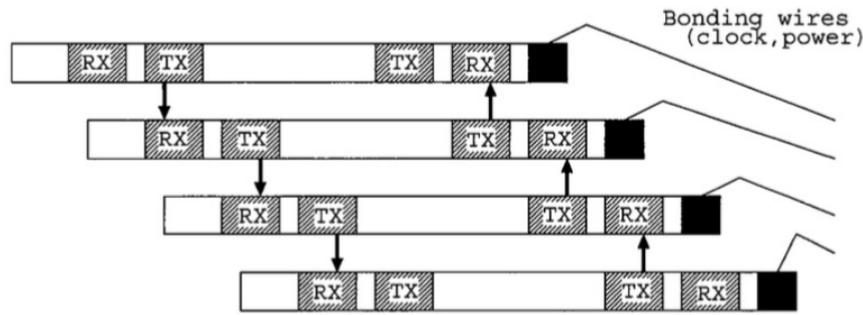


Figure 2: Ring network in Cube-1

can be transferred with 50MHz clock cycle on a link. As shown in Fig. 1, the 35bit data are stored in the serializer, and transferred serially synchronized with a 2GHz clock generated with cyclic buffers through the transceiver. The receiving data by the receiver are stored into the de-serializer, and sent to the digital part with a clock delay.

The bubble flow control[15] is a deadlock avoidance technique for rings with virtual cut-through (VCT) switching. It does not require any VCs but requires a single buffer with a capacity of at least two packets for each input port. By limiting the packet injection so as not to consume all the buffer resources in a ring, packets on the ring continuously move without any deadlocks. By keeping simple rules to save an empty space in the buffer, packets can continuously move along the ring. As long as the packets are continuously moving on the ring, no deadlocks can occur even though multiple message classes coexist in the same virtual network. For detail about the bubble flow control, the reader can refer to [15]. Also, the detail implementation in Cube-1 is shown in [16].

2.3 Problems of the network in Cube-1

Although the uni-directional ring has such benefits for chip stacking, the disadvantages listed below were found through development of Cube-1.

- Since two TCI links are needed to be provided at separate position in a chip, the position of coils tends to interference other coils as shown in Fig. 2. Although the TCI does not suffer the interference from neighboring coils, a large interference is observed from vertically overlapping coils. In the case of Cube-1, we assumed that the interference between the coils of the top chip and the bottom chip is not so severe, since the vertical distance is large. However, it was main reason why the communication of Cube-1 is not stable when more than three chips are stacked. We must keep the distance between coils but considering the size of the chip, it was difficult.
- In order to keep enough bubble on the network without the acknowledging mechanism, the router can insert packets only with a certain interval. Also, because of the large hop counts of uni-directional ring network, the latency between chips tend to be large. It increased the overhead for reading/writing the data from/to the accelerator chips.
- Bubble flow control does not require any VCs for deadlock free communication. However, VCs are required by the operating system to handle multiple message classes.

3 Escalator Network

3.1 Structure

Instead of the ring network, an escalator network shown in Fig. 3 is introduced for the second prototype CUBE-SOTB. Each chip has a three input/output ports router, and it connects with a

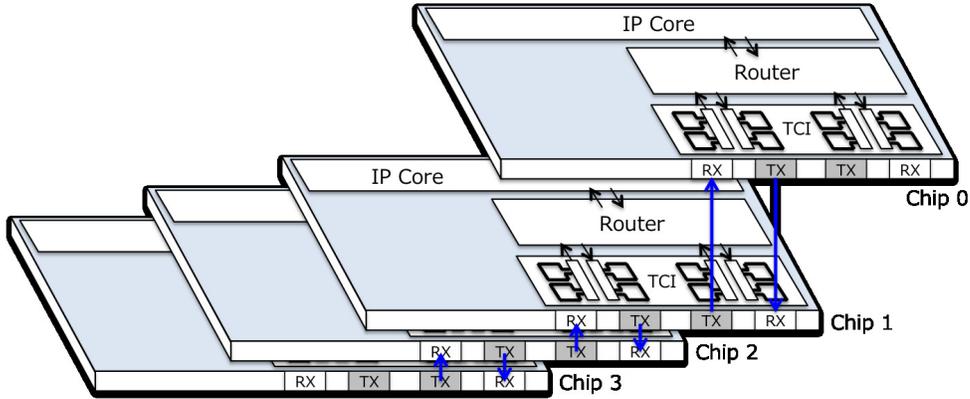


Figure 3: Escalator network in CUBE-SOTB

TX and a RX to upward/downward and the core on the chip. Unlike the ring network, full-duplex bi-directional interconnection is formed between the neighboring stacked chip. By using the opposite direction link, the sending router can detect the buffer status in the receiving router. Packets are routed to their destination along up-direction or down-direction link based on the chip-IDs of current and destination chips. If the target chip-ID is the same as the current one, the packet is transferred to the core. If it is smaller than the number of the current chip, the uplink is used, otherwise the downlink is used. This simple linear network is called the escalator network.

Here, the input port of each router has eight virtual channels (VCs), although the escalator network is deadlock free without any VCs. They are implemented not for increasing throughput but in order that the packets can use different VCs depending on their message type. Thus, the VC allocation is done by the operating system of the host processor during the initialization using the default VC number. A credit based flow control is implemented and the packet in the router can be sent through the TCI link only when the credit for free VC is given. Although the packet has to stop when the credit has not been given, the deadlock-free is guaranteed and the risk of overwriting has been removed in this network.

3.2 Piggyback Control

Since TCI is the interconnection using a coil, its footprint for a link is much larger than other techniques like TSV as shown in the previous section. The dedicated coil only for the information for flow control should not be provided considering such a large footprint. Thus, a uni-directional ring network in Cube-1 uses a bubble flow control. On the other hand, by making the use of bi-directional communication, the information for flow control for one direction is mixed with data signals for opposite direction in the escalator network.

Assuming that the router-X is a sender of data packet and the router-Y is a receiver, the flow of sending the credit packet through the opposite link is shown in Fig. 4. When some buffers of VCs in the router-Y have become free since the last credit was sent, the credit packet can be sent if the opposite link is available. If it is unavailable due to the data packet transfer from router-Y to X, the credit packet waits to be sent until the link becomes available. Then, the contents of the packet is updated if the state of VCs has changed while waiting.

This piggybacking control mechanism introduces a potential overhead to the performance with the following two aspects. First, the network throughput is degraded by sharing the opposite link by the normal data and credit packets. Second, the arrival of the credit packet and thereby the notification of the VC status is delayed particularly when a large data packet is being transferred along the opposite direction. In order to examine this problem, we evaluate the overhead of the piggyback control besides the comparison between the uni-directional ring and the escalator network later.

The linear network itself is a fundamental topology and piggyback of control data is a common

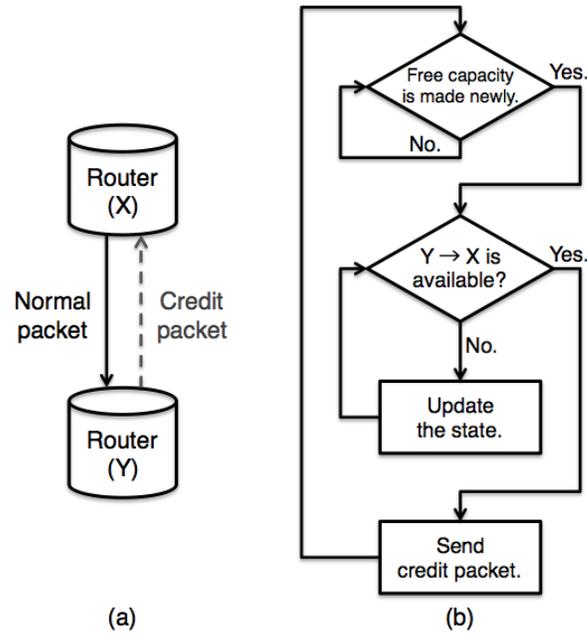


Figure 4: The flow of sending credit packet: (a) The direction focused on (b) Flowchart

technique in vehicular ad-hoc networks[17] or wireless sensor networks[18] which does not have enough bandwidth. However, they have not been applied to the inter-chip networks, since traditional inter-chip communication with TSVs can easily provide the dedicated wires for flow control. The following benefits of the proposed escalator network are mostly for using TCI links.

- TCI link only for flow control signals is not needed. It saves the relatively large area for TCI.
- Since the escalator network is consisting of bi-directional links between neighboring chips, the upper layers including the flow control with multiple VCs can be packaged into a TCI IP. By providing such IPs, the network can be formed just by stacking multiple chips.
- Although the overhead of piggyback is required, bubbles in the uni-directional ring adopted in Cube-1 are not needed. The performance can be comparable to the one with a dedicated handshake lines as shown in later.

4 IP (Intellectual Property) for TCI

For building bi-directional networks like the escalator network easily, we developed an integrated IP for TCI network. As shown in Fig. 5, the IP consists of the router layer, link control layer and TCI physical layer. Here, four physical layer IPs DR(Down Receiver), DT(Down Transmitter), UR(Up Receiver) and UT(Up Transmitter) are used as described later. Each layer can be used separately.

4.1 TCI physical layer

TCI physical layer is consisting of coils, transmitter, receiver and SERDES (Serializer/De-serializer) as shown in Fig. 6. As shown in the layout in Fig. 7, the receiver coil and transmitter coil are double wound up. The inner coil with four square wires is for the receiver, while the outer coil with two square wires is for the sender. That is, a channel of the IP is half-duplex. The direction can be switched within a few-clock interval. A pair of coils for clock and data are provided. From the user, the IP can be treated as a simple 35-bit uni-directional registered channel. Fig. 8 shows the waveform of data transfer. The transmitter accepts 35bits digital data with asserting TXWrite signal. The

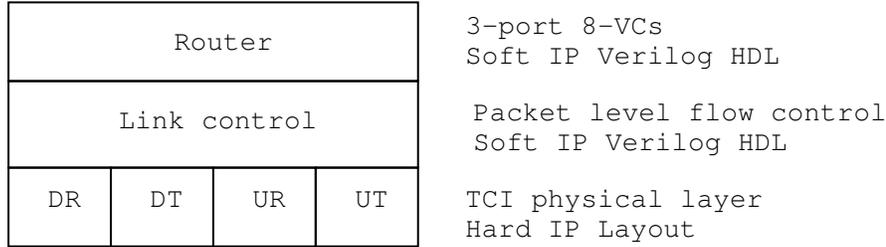


Figure 5: TCI IP layers

data are transferred to the next chip in the serial manner synchronized with a high speed (2.5GHz) clock generated by the ring oscillator in the IP. SERDES in the transceiver serializes the TX data. In order to save the power, high speed clock is only generated during data transfer. The start bits, parity, and end bits are automatically attached in the serializer. When end bits are received at the de-serializer of the receiver SERDES, RXReady is asserted. After reading the data, the receiver digital part (the link layer) asserts RXRead to allow the IP to receive the next data. In order to be used in various systems, the interface of the IP is asynchronous, and its performance is scaled by the power supply voltage and body bias voltage. The standard operational frequency is 50MHz because of the conservative design of the IP.

Three error bits are provided. PERROR indicates a parity error in the transferred data, OSEDI-TECT is asserted when an error is detected in the start bits and end bits, and RXOW is asserted when the next data are transferred before RxRead is asserted by the receiver digital part. In the physical layer, the sender provides no mechanism to recognize transmission error detected in the receiver. Thus, the error recovery must be provided in the upper layer.

Also, the flow control mechanism itself is not provided in the IP. During the transmission, TXBUSY is asserted to avoid overwriting the transmitted data. However, the sender router cannot know whether the receiver reads the data without switching the direction. If the next flit is transferred before the receiver router reads the previous flit, in the receiver side RXOW flag is asserted to indicate an overwriting at the receiver, but it is not automatically transferred to the sender. In our integrated IP, the flow-control is treated in the link layer.

The IP shown in Fig. 7 uses two $160\mu\text{m} \times 160\mu\text{m}$ coils for clock and data. They are marked in the figure as MSACOIL_D160_Data and MSACOIL_D160_Clock, respectively. Note that MSA is a name of the IP, and D160 shows the size of the coil. They are used when the chip is thinner than $60\mu\text{m}$. Unfortunately, the size of IP is relatively large $510\mu\text{m} \times 410.8\mu\text{m}$, because of the conservative implementation to keep the area inside and between coils empty. Now, the serializer and transceiver for the TX channel is implemented in MSATXBLK, while the de-serializer and receiver are implemented in MSZRXBLK shown in Fig.7.

The size can be reduced by implementing the SERDES inside or between the coil. Note that TCI physical IP can be used any type of networks including uni-directional ring used in Cube-1.

4.2 TCI4 layer

In order to build the escalator network, four TCI physical IPs: down transmitter (DT), down receiver (DR), up transmitter (UT) and up receiver (UR) are required as shown in Fig. 9. We developed a hard IP called TCI4 consisting of the 2x2 array of TCI physical IPs. As mentioned in the previous section, the TCI physical IP itself can change the direction dynamically. However, in this layer, the direction is fixed. That is, the DT and UT are for transmitters while DR and UR are for receivers. A bi-directional channel is formed just by stacking two chips with TCI4 by a gap corresponding to the distance of two IPs. By the stacking between IPs shown in Fig. 9, the escalator network shown in Fig. 3 is physically implemented.

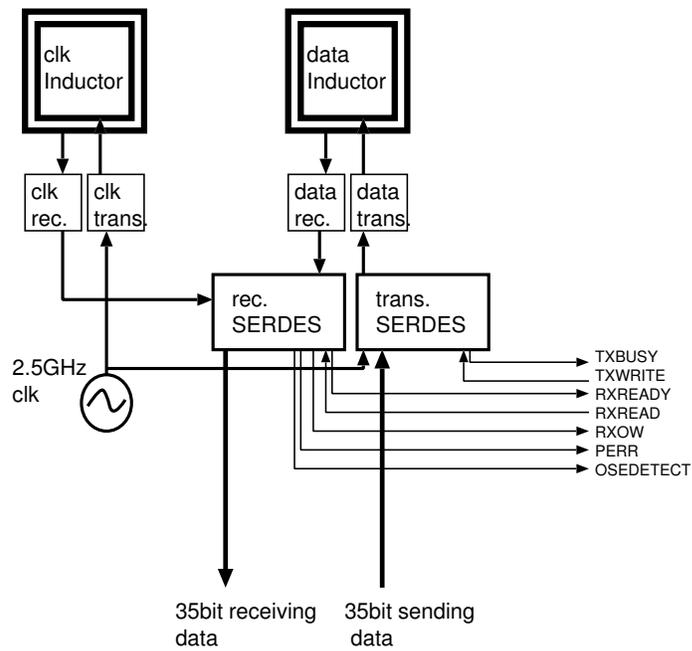


Figure 6: Structure of TCI Physical layer

4.3 Link control layer

Through the TCI physical layer, the sender chip can transfer a 35-bit data flit to the receiver chip. The link control layer treats the data as a form of packet, and provides a link-level flow control. In the IP, eight VCs each of which has 24-flit buffers are provided for up/down link. Here, the virtual cut-through (VCT) flow control is adopted. As shown in Fig. 10, a flit is classified into three types: HEAD/HEADTAIL, TAIL/DATA and Credit according to the flit type. Credit packets include ACK0-ACK7 fields each of which is corresponding to each VC. Unlike an acknowledge signal on the dedicated wire, the acknowledge flit is piggy backed in the data packet and so the instant acknowledgment is not possible. Thus, five bits are required to show the state of 24-flit buffer for the flow control. The VC number and the valid signal are embedded in the HEAD/HEADTAIL flit since they are in the same direction with the main data.

On the other hand, the credit signals for the status of VCs are sent through the opposite direction links. Then, STAT0 and STAT1 in Fig. 10 show the sending chip how much the capacity of VCs in the receiving router has become free since they were sent in the last time. Since each link has eight VCs and the capacity of the VC is 24 flits i.e. it needs 5 bits for one VC to indicate the amount of the free capacity in the VC. It is impossible to gather the status of all of the VCs into one packet since the data size of one flit is 32 bit. Here, the STAT0 stores the state of VCs 0-3, while the STAT1 stores No.4-7. The generation of credit packets and flow control shown in Fig. 4 are done in the link control layer. The link control layer ensures the link-to-link control between two neighboring chips. Since bi-directional link is required, it is useful in the escalator network, but cannot be used in the uni-directional ring in Cube-1.

4.4 Router layer

Router layer is consisting of two ports for other chips, a switch, an arbiter, and a port to the host. Each port for other chip provides a link layer. It is a parametrized 3-stage pipelined router with virtual cut-through packet transfer control. Although it is designed for escalator network which requires 3 ports, the number of ports can be easily extended for other network topologies.

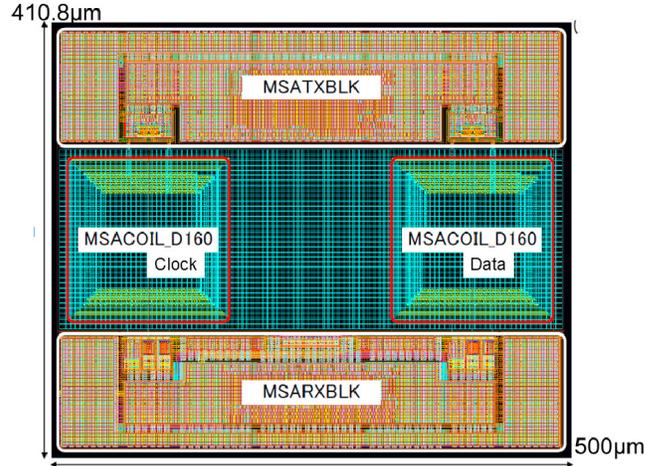


Figure 7: The layout of TCI IP

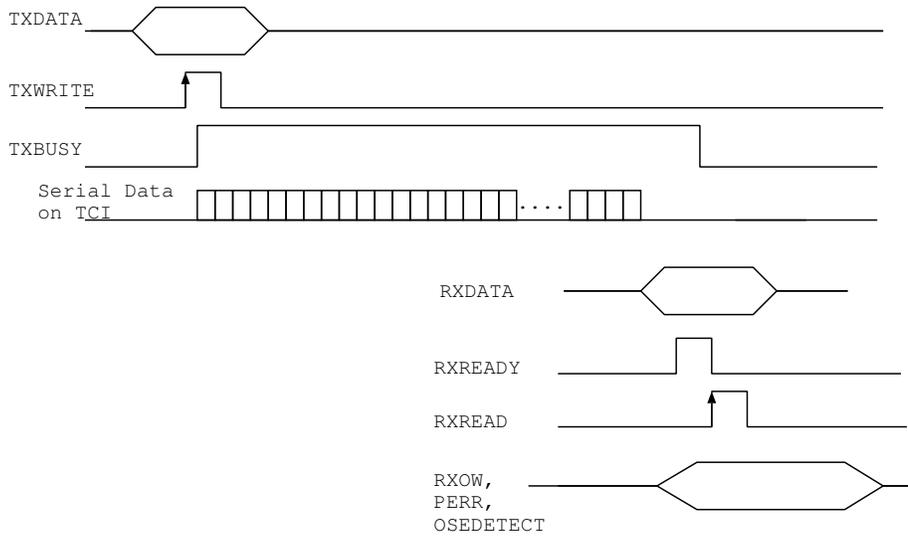


Figure 8: Data transfer between two chips

5 Evaluation

5.1 Experimental Setup

We evaluated the packet transmission performance of the networks used in Cube-1 and CUBE-SOTB. Each network in a chip stack using four chips is implemented at RTL by using Verilog HDL with the integrated TCI IP. All parameters are based on the real chip implementation of Cube-1 and CUBE-SOTB. In the escalator network, each router does not perform the VC allocation i.e. it is determined which VC is used when the packet is injected, and the same VC ID is used in the subsequent routers.

Parameters in this evaluation are listed in Table. 1. We executed the RTL simulation of the packet transmission under uniform and bit-reverse traffic pattern using Cadence NC-Verilog. Although packet size in the escalator network is a maximum of 17 flits, the size is set to 5 flits in order to compare it with the ring network that works with 5-flit packets.

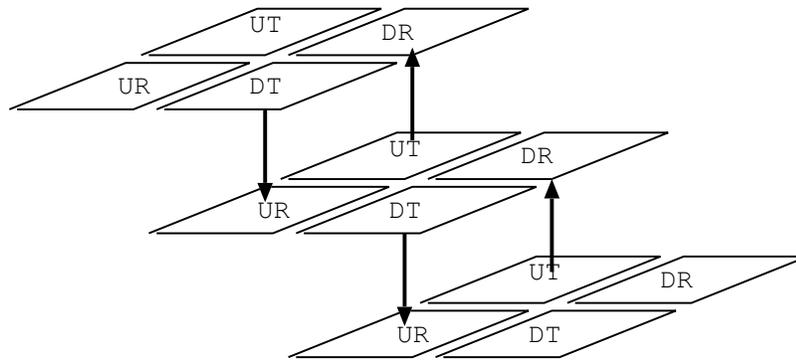


Figure 9: Chip stacking with TCI4

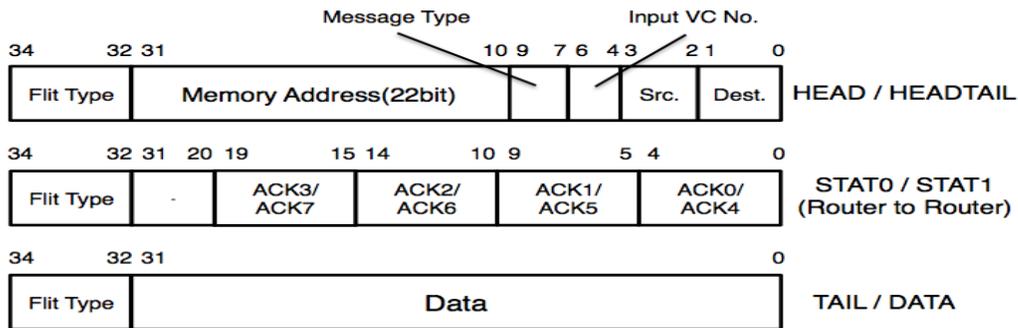


Figure 10: Packets treated in Link control layer

Table 1: Simulation parameter

Router switching	Virtual Cut Through
Switching latency of router	3 cycle
Latency of TCI	1 cycle
Packet size	5 flit
Input buffer size	24 flit
Flit size	35 bit
Num. of VCs (for escalator network)	8 VCs
Synthesis traffic pattern	Uniform, Bit-reverse
Num. of injected packets	2000

5.2 Performance Comparison

The results of comparison between the ring and escalator network are shown in Fig. 11. The escalator network without VCs was also evaluated since the uni-directional ring network does not have any VCs. In the escalator network with VCs, each packet chooses its VC in turn.

From the result under the uniform traffic, the throughput of the escalator network without and with VCs are 26% and 59% higher than that of the ring network, respectively. As a result, even without VCs, the escalator network is superior to the ring network in terms of network throughput under the uniform traffic. This shows that the input buffers are utilized more efficiently in the credit based flow control by piggybacking than the bubble-flow control that requires a certain amount of free capacity in the buffers.

On the other hand, from the result under bit-reverse traffic, the throughput of the escalator network without VCs is 7% lower than that of the ring network. The traffic load is given more to

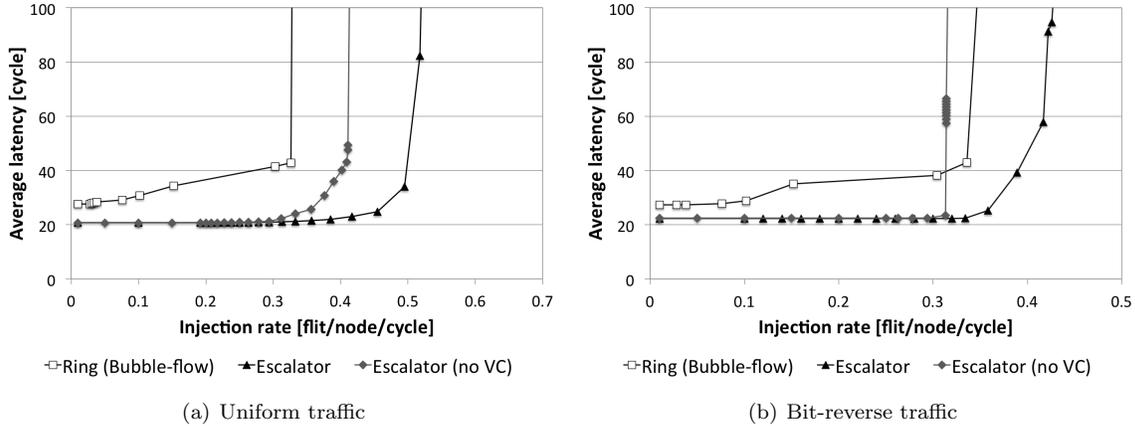


Figure 11: Performance comparison by the synthesis traffics

three input/output ports routers formed in the escalator network than to one or two input/output ports routers in the ring network. In addition, in this traffic, packets from the top and bottom of the stacked chips are always transferred to the bottom and top, respectively. Nevertheless, the throughput of the escalator network with VCs is still 28% higher than that of the ring network.

In terms of the zero-load latency, the escalator network was superior to the ring under both traffics, as shown in Fig 11. This is because the number of hops in the escalator network is smaller than that in the ring network that has two routers for each chip. As a result, the zero-load latency of the escalator network even without using VCs is 25% in uniform traffic and 18% in bit-reverse lower than that of the ring network.

5.3 Overhead of the Piggyback Control

In order to examine the overhead of the piggyback control, the network that performs the credit based flow control without the piggyback transmission of the credit packets was implemented for comparison as an ideal performance. The ideal network is the same structure as the escalator network except that additional TCI links for transferring the credit signals are provided.

From Fig. 12 showing the difference of the two, a certain amount of influence on the throughput is recognized under both traffic patterns. However, it can be considered as the negligible difference because the throughput of the piggybacking network is 4% in uniform traffic and 3% in bit-reverse lower than that of the ideal one. Moreover, almost no difference of the transmission latency between the two is confirmed when the offered load is relatively low. The latency after the network is saturated was larger than ideal case especially under bit-reverse traffic, since a large number of credit control packets are needed in this region. However, this region is not supposed to be used for common applications.

5.4 Performance of benchmark programs

Here, the performance of the ring and escalator network is evaluated when benchmark programs are executed. In this evaluation, four chips; two for GCSOTB CPUs and others for L2 cache are assumed to form a microprocessor which executes benchmark programs in parallel. Two CPUs are assigned on the top two chips on the stack, and bottom two was for the L2 cache. Tasks are evenly distributed into two CPUs and the L2 cache is shared by them. Eight programs described in OpenMP are selected from NAS Parallel Benchmarks[19] with a full-system simulator GEM5[20] and the captured traffic trace was used in the RTL simulation. The code of GEM5 was slightly modified to extract the packets for the RTL simulation. Note that the execution time of benchmark programs is not influenced by the network. So, we show the average latency of a packet transfer during execution of benchmark programs in Fig. 13. It shows that average latency is not so changed in all application

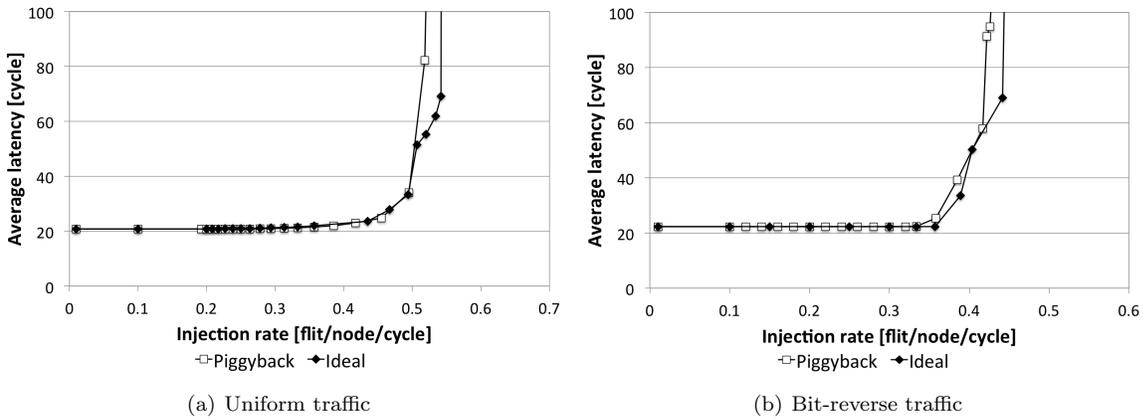


Figure 12: Overhead of piggyback

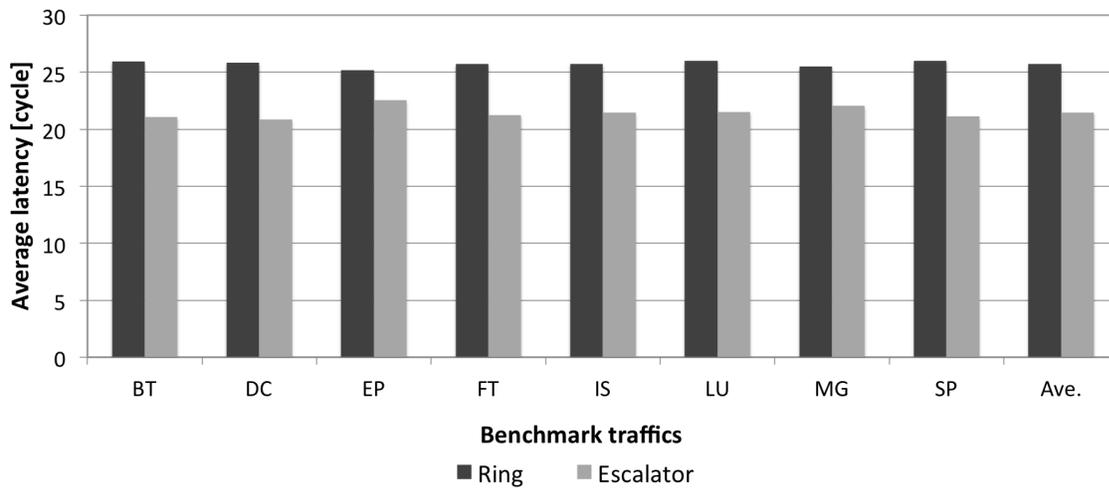


Figure 13: Average latency on executing benchmark programs

programs, that is, the traffic congestion is not severe. The latency with escalator network is 10% - 20% better than that of ring network. This shows that the low latency communication was done in real application programs.

6 Real Chip Implementation

The CUBE-SOTB which uses the escalator network and integrated IPs evaluated in the previous section was implemented, and now a prototype stack is available. Two chips; a host processor GCSOTB and coarse grained reconfigurable accelerator CCSOTB were implemented using Renesas SOTB 65nm process. The specification of CUBE-SOTB compared with Cube-1 is summarized in Table 2. The link-layer and router layer in the IP consume 47215 cells, almost the same as the case in Cube-1. It can be layouted with the other logic for reducing the size. Although the core of each chip works well, because of the bug in the power network layout, the TCI of GCSOTB does not work. Thus, we developed a chip stack only with CCSOTB chips.

As shown in Fig. 14, TCI4 IP is located at the upper left of the chip. It is consisting of four TCI physical IPs in Fig. 7 each of which is marked with a white frame. Two rectangle shapes are for the microcontroller and the PE array of the reconfigurable accelerator which is out of focus in the paper. The chip stack was implemented on the TEG daughter board, and connected with a small

Table 2: Spec. of Cube-1 and CUBE-SOTB

	Cube-1	CUBE-SOTB
Process	e-shuttle 65nm CMOS (12-Metal)	SOTB 65nm CMOS (7-Metal)
Area	2.1mm × 4.2mm	3mm × 6mm
TCI	3Gb/s 240 μ m	2.5Gb/s 240 μ m
CPU ISA I/D Cache TLB	Geyser-Cube MIPS R3000 4KW 2way 16-Entry Shared	GCSOTB MIPS R3000 4KW 2way 16-Entry Shared
CMA PE Array	CMA-Cube 8×8	CCSOTB 12×8
Supply Voltage	0.5-1.2V	0.3-1.2V
Target Freq.	50-100MHz	50-100MHz
Chip Thickness	40-80 μ m	80 μ m

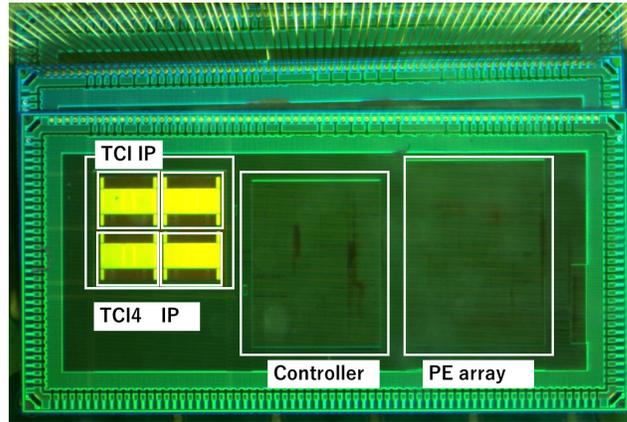


Figure 14: Three-chip-stack of CCSOTBs

FPGA card on the mother board shown in Fig.15. The test data generated in the FPGA board are sent to a chip in the stack, and used for data transfer with the escalator network. The receiving data are transferred to the FPGA and examined. All error signals are also monitored by the FPGA. Up to now, no errors were found.

Now, a three-chip-stack shown in Fig. 14 is stably operational. Table 3 shows the operational conditions and evaluation results from the real system. We will further optimize operational conditions including body bias control to be used with higher clock frequency.

7 Conclusion

We proposed the escalator network and an IP for a building block system with wireless inductive TCIs. By stacking multiple chips with the same IP, an inter-chip network with link-to-link flow control is established. In order to investigate the influence of piggybacking credit packets for the credit based flow control, the performance is evaluated with network simulations and RTL simulations, and compared with the uni-directional ring network used in the previous prototype.

By efficiently utilizing the input buffers and reducing the hop counts, the throughput of escalator network is 28%-59% higher than that of the ring network, and the proposed network outperforms the



Figure 15: Testing environment of CUBE-SOTB

Table 3: Current Operational Conditions

Clock for TCI	1.8GHz
Operational frequency	30MHz
TX voltage	1.3V
TX current	58.7mA
RX current	28.9mA
Digital part voltage	0.8V

ring network by 18%-25% in terms of zero-load latency. In addition, the performance overhead by the piggyback control could be lowered by 3%-4%. As a result, problems of the ring network used in Cube-1 have been solved. The real chip implementation and evaluation demonstrated the proposed IP is operational on three-chip stack. Evaluation of the inter-chip network in the real system is our future work.

Acknowledgment

This work is partially supported by JSPS KAKENHI S grant number 25220002. The authors are grateful for useful comments from Prof. Michihiro Koibuchi with National Institute of Technology.

References

- [1] Kouichi Kumagai, Changqi Yang, Satoshi Goto, Takeshi Ikenaga, Yoshihiro Mabuchi, and Kenji Yoshida. System-in-Silicon Architecture and its application to an H.264/AVC motion estimation for 1080HDTV. In *Proceedings of the International Solid-State Circuits Conference (ISSCC'06)*, pages 430–431, February 2006.
- [2] J. Burns, L. McIlrath, C. Keast, C. Lewis, A. Loomis, K. Warner, and P. Wyatt. Three-Dimensional Integrated Circuits for Low-Power High-Bandwidth Systems on a Chip. In *Proceedings of the International Solid-State Circuits Conference (ISSCC'01)*, pages 268–269, February 2001.

- [3] D.P.Seemuth, A.Davoodi, and K.Morrow. Flexible interconnect in 2.5D ICs to minimize the interposer's metal layers. In *22nd Asia and South Pacific Design Automation Conference (ASPDAC)*, pages 372–377, 2017.
- [4] Kouichi Kanda, Danardono Dwi Antono, Koichi Ishida, Hiroshi Kawaguchi, Tadahiro Kuroda, and Takayasu Sakurai. 1.27-Gbps/pin, 3mW/pin Wireless Superconnect (WSC) Interface Scheme. In *Proceedings of the International Solid-State Circuits Conference (ISSCC'03)*, pages 186–187, February 2003.
- [5] William Rhett Davis, John Wilson, Stephen Mick, Jian Xu, Hao Hua, Christopher Mineo, Ambarish M. Sule, Michael Steer, and Paul D. Franzon. Demystifying 3D ICs: The Pros and Cons of Going Vertical. *IEEE Design and Test of Computers*, 22(6):498–510, November 2005.
- [6] Noriyuki Miura, Daisuke Mizoguchi, Mari Inoue, Kiichi Niitsu, Yoshihiro Nakagawa, Masamoto Tago, Muneo Fukaiishi, Takayasu Sakurai, and Tadahiro Kuroda. A 1Tb/s 3W Inductive-Coupling Transceiver for Inter-Chip Clock and Data Link. In *Proceedings of the International Solid-State Circuits Conference (ISSCC'06)*, pages 424–425, February 2006.
- [7] Noriyuki Miura, Hiroki Ishikuro, Takayasu Sakurai, and Tadahiro Kuroda. A 0.14pJ/b Inductive-Coupling Inter-Chip Data Transceiver with Digitally-Controlled Precise Pulse Shaping. In *Proceedings of the International Solid-State Circuits Conference (ISSCC'07)*, pages 358–359, February 2007.
- [8] N.Miura, Y.Koizumi, Y.Take, H.Matsutani, T.Kuroda, H.Amano, R.Sakamoto, M.Namiki, K.Usami, M.Kondo, H.Nakamura. A Scalable 3D Heterogeneous Multicore with an Inductive ThruChip Interface. *IEEE Micro*, 33(6):6–15, Nov 2013.
- [9] L. Zhao, D. Ikebuchi, Y. Saito, M. Kamata, N. Seki, Y. Kojima, H. Amano, S. Koyama, T. Hashida, Y. Umahashi, D. Masuda, K. Usami, K. Kimura, M. Namiki, S. Takeda, H. Nakamura, and M. Kondo. Geyser-2: The second prototype CPU with fine-grained run-time power gating. *Proceedings of the Asia and South Pacific Design Automation Conference, ASP-DAC*, pages 87–88, 2011.
- [10] N. Ozaki et al. Cool Mega Arrays: Ultra-Low-Power Reconfigurable Accelerator Chips. *IEEE Micro*, vol.31, No.6, pages 6–18, 2011.
- [11] Hiroki Matsutani, Paul Bogdan, Radu Marculescu, Yasuhiro Take, Daisuke Sasaki, Hao Zhang, Michihiro Koibuchi, Tadahiro Kuroda, and Hideharu Amano. A Case for Wireless 3D NoCs for CMPs. In *Proceedings of the Asia and South Pacific Design Automation Conference (ASP-DAC'13)*, pages 23–28, January 2013.
- [12] R. Tsuchiya, et al. Silicon on Thin BOX : A New Paradigm of the CMOSFET for Low-Power and High-Performance Application Featuring Wide-Range Back-Bias Control. *Tech. Dig. Int, Electron Devices Meet., 0-7803-8684-1, San Francisco*, pages 631–634, 2004.
- [13] Takashi Ishigaki, et al. Ultralow-power LSI Technology with Silicon on Thin Buried Oxide (SOTB) CMOSFET. *Solid State Circuits Technologies, Jacobus W. Swart (Ed.), ISBN: 978-953-307-045-2, InTech*, pages 146–156, 2010.
- [14] Akio Nomura, Hiroki Matsutani, Junichiro Kadomoto, Tadahiro Kuroda, Yusuke Matsushita, and Hideharu Amano. Vertical Packet Switching Elevator Network Using Inductive Coupling ThruChip Interface. In *International Symposium on Computing and Networking (CANDAR'16)*, pages 195–201, Nov 2016.
- [15] Valentin Puente, R. Beivide, J. A. Gregorio, J. M. Prellezo, Jose Duato, and Cruz Izu. Adaptive Bubble Router: A Design to Improve Performance in Torus Networks. In *Proceedings of the International Conference on Parallel Processing (ICPP'99)*, pages 58–67, September 1999.

- [16] H.Matsutani, Y.Take, D.Sasaki, M.Kimura, Y.Ono, Y.Nishiyama, M.Koibuchi, T.Kuroda and H.Amano. A Vertical Bubble Flow Network using Inductive-Coupling for 3-D CMPs. In *Proceedings of the International Symposium on Networks on Chip (NoCS'11)*, pages 49–56, May 2011.
- [17] H.Zhang and X.Zhang. An Adaptive Control Structure based Fast Broadcast Protocol for Vehicular Ad hoc Networks. *IEEE Communication Letters*, 21, 2017.
- [18] M.Arifuzzaman and O.A.Dobre and M.H.Ahmed and T.M.N.Ngatched. Joint ROuting and MAC Layer QoS-Aware Protocol for Wireless Sensor Networks. *IEEE Global Communications Conference (GLOBECOM)*, 2016.
- [19] H. Jin, M. Frumkin, and J. Yan. The OpenMP Implementation of NAS Parallel Benchmarks and Its Performane. In *NAS Technical Report NAS-99-011*, October 1999.
- [20] Milo M. K. Martin, Daniel J. Sorin, Bradford M. Beckmann, Michael R. Marty, Min Xu, Alaa R. Alameldeen, Kevin E. Moore, Mark D. Hill, and David A. Wood. Multifacet General Execution-driven Multiprocessor Simulator (GEMS) Toolset. *ACM SIGARCH Computer Architecture News (CAN'05)*, 33(4):92–99, November 2005.