

A Vertical Bubble Flow Network using Inductive-Coupling for 3-D CMPs *

Hiroki Matsutani¹, Yasuhiro Take², Daisuke Sasaki², Masayuki Kimura², Yuki Ono², Yukinori Nishiyama², Michihiro Koibuchi³, Tadahiro Kuroda², and Hideharu Amano²

¹The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan
matutani@hal.ipc.i.u-tokyo.ac.jp

²Keio University
3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Japan
cube@am.ics.keio.ac.jp

³National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, Japan
koibuchi@nii.ac.jp

ABSTRACT

A wireless 3-D NoC architecture for CMPs, in which the number of processor and cache chips stacked in a package can be changed after the chip fabrication, is proposed by using the inductive coupling technology that can connect more than two known-good-dies without wire connections. Each chip has data transceivers for uplink and downlink in order to communicate with its neighboring chips in the package. These chips form a single vertical ring network so as to fully exploit the flexibility of the wireless approach that enables us to add, remove, and swap the chips in the ring. To avoid protocol and structural deadlocks in the ring network, we use the bubble flow control which is more flexible and efficient compared to the conventional VC-based deadlock avoidance. We implemented a real 3-D chip that has on-chip routers and inductive-coupling data transceivers using a 65nm process in order to show the feasibility of our proposal. The vertical bubble flow control is compared with the conventional VC-based approach and vertical bus in terms of the throughput, hardware amount, and application performance using a full system CMP simulator. The results show that the proposed vertical bubble flow network outperforms the VC-based approach by 7.9%-12.5% with a 33.5% smaller router area.

*This research was performed by the authors for STARC as part of the Japanese Ministry of Economy, Trade and Industry sponsored “Next-Generation Circuit Architecture Technical Development” program. The authors thank to VLSI Design and Education Center (VDEC) and Japan Science and Technology Agency (JST) CREST for their support. This work was also supported by Grant-in-Aid for JSPS Fellows.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Copyright 20XX ACM X-XXXXX-XX-X/XX/XX ...\$10.00.

Categories and Subject Descriptors

C.1.2 [Computer Systems Organization]: Multiprocessors—*Interconnection architectures*

General Terms

Design, Performance

Keywords

On-chip networks, many-core, 3-D ICs, inductive-coupling

1. INTRODUCTION

The three-dimensional Network-on-Chip (3-D NoC) [26] is an emerging research topic exploring the network architecture of 3-D ICs that stack several smaller wafers or dies in order to reduce the wire length and wire delay. Regarding the 3-D NoC architecture, the network topology [17, 22], router architecture [11, 13, 21], and routing strategy [24] have already been extensively studied.

Various interconnection technologies have been developed to be used as a medium for the 3-D NoCs: wire-bonding, micro-bump [4, 12], wireless (e.g., capacitive or inductive coupling) [7, 9, 19, 20] between stacked dies, and through-silicon via (TSV) [5, 7] between stacked wafers. The comparisons of these 3-D IC technologies are discussed in [7]. Many recent studies on 3-D NoC architecture focus on the TSV that offers the largest interconnect density. Unlike them, we focus on the inductive coupling that can connect more than two known-good-dies without wire connections, because it offers a large degree of flexibility to build the desired 3-D chip multiprocessors (3-D CMPs).

In this paper, we propose a wireless NoC architecture for 3-D CMPs, in which the number of processor and cache chips stacked in a package can be changed after the chip fabrication by using the inductive coupling. Each chip has data transceivers for uplink and downlink in order to communicate with its neighboring chips in the package. These chips form a single vertical ring network so as to fully exploit the flexibility of the wireless stacking that enables us to add, remove, and swap the chips in the ring. To avoid deadlocks in the ring network, we use the bubble flow control [1, 23] that does not require any virtual channels (VCs), because

the conventional VC-based deadlock avoidance techniques limit the flexibility of the wireless 3-D CMPs, depending on the number of VCs available.

We implemented a real 3-D chip that has on-chip routers and inductive-coupling data transceivers using a 65nm process. The vertical bubble flow control is compared with the conventional VC-based approach and vertical bus in terms of the throughput, hardware amount, and application performance using a full system CMP simulator.

The rest of this paper is organized as follows. Section 2 surveys existing 3-D IC technologies. Section 3 proposes the wireless 3-D NoC architecture and Section 4 illustrates the test chip implementation. Section 5 shows the evaluation results and Section 6 concludes this paper.

2. 3-D IC TECHNOLOGY

Interconnection technologies for 3-D ICs are classified into the wired and wireless approaches. In this section, we first survey both approaches, and then advantages of the inductive coupling for our purpose are shown.

2.1 Wired 3-D Interconnection

- **Wire-bonding** is a die-to-die interconnection using bonding wires. It is a popularly utilized technique for the current System-in-Package. However, since only edges of a chip can be used for the bonding, the number of wires and their density are limited. Also, the long bonding wires cause a considerable delay for communication.
- **Micro-bump** [4, 12] is a die-to-die interconnection through solder balls. This approach is limited to the face-to-face connection of only two dies.
- **Through-silicon via (TSV)** [5, 7] is a wafer-level interconnection that uses via-holes formed through multiple wafers. The footprint of the TSV is small, and high-density implementation is possible. However, the fabrication cost is increased due to the extra process steps for forming the TSVs. Although more than two wafers can be connected, the number and order of the wafers cannot be changed after the fabrication, since it is a wired approach.

Although these wired approaches have been utilized in real products, they are not enough flexible for customizing the chips in a package in response to application requirements.

2.2 Wireless 3-D Interconnection

- **Capacitive coupling** [9] connects the known-good-dies without wire connections. However, only the face-to-face connection is allowed; so the number of stacked dies is limited to only two.
- **Inductive coupling** [7, 19, 20, 25] also supports a wireless interconnection between the known-good-dies, and more than two dies can be stacked.

As an interconnection technology for custom 3-D CMPs, the inductive coupling is hopeful, since it can stack a number of known-good-dies wirelessly. That is, addition, removal, and swapping of processor or cache chips are possible after the chip fabrication in response to the application. Recently, the inductive coupling techniques have been improved, and the contact-less interface without an electrostatic-discharge (ESD) protection device achieves high speed more than 1GHz with a low energy dissipation (0.14pJ per bit) and a low bit-error rate ($BER < 10^{-12}$) [19].

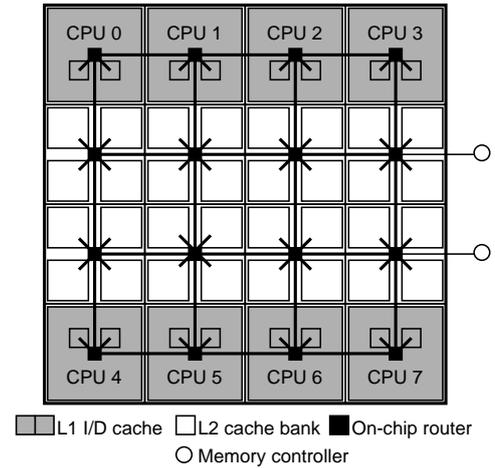


Figure 1: Baseline 2-D CMP

In this approach, an inductor is implemented as a square coil with metal in common CMOS layout (see Figure 6 for the layout). The data modulated by a driver are transferred between two coils placed at exactly the same position of two stacked dies, and it is received at the other chip by the receiver. Here, a set of a driver and a coil for sending data is called TX channel, while one with a receiver and a coil is RX channel. If a TX channel is placed at the same location of multiple RX channels in different chips, multicast of the data can be done. On the other hand, stacked multiple TX channels at the same location cannot send the data simultaneously in order to avoid the interference.

The footprint of the channel ranges from $30\mu\text{m} \times 30\mu\text{m}$ [19, 20] to $150\mu\text{m} \times 150\mu\text{m}$ [7], depending on the process technology and communication distance (i.e., chip thickness). In addition, a number of TX and RX channels can be implemented for the parallel data transfer, depending on the required vertical bandwidth. A 1-Tbit/sec inductive-coupling inter-chip clock and data link has been developed by using 1024 transceivers arranged with a pitch of $30\mu\text{m}$ [20]. The inductive coupling has been applied to various purposes, such as multi-core processors and dynamically reconfigurable processors [25].

The next section proposes the wireless 3-D NoC architecture for CMPs, in which processor and cache chips are stacked wirelessly using the inductive coupling links.

3. WIRELESS 3-D CMP ARCHITECTURE

We apply the inductive coupling to build the wireless 3-D CMP, in which the numbers of processors and cache banks can be customized in response to the application performance and cost requirements. In this section, first the baseline 2-D CMP architecture is illustrated. Then, the wireless 3-D CMP architecture is proposed.

3.1 Baseline 2-D CMP Architecture

Figure 1 illustrates an example of CMP inspired by [3], in which eight processors (or CPUs) and 32 shared L2 cache banks are interconnected by sixteen on-chip routers. These L2 cache banks are shared by all processors; so the cache architecture is SNUCA [10] and a cache coherence protocol is running on it. We extend this CMP to 3-D.

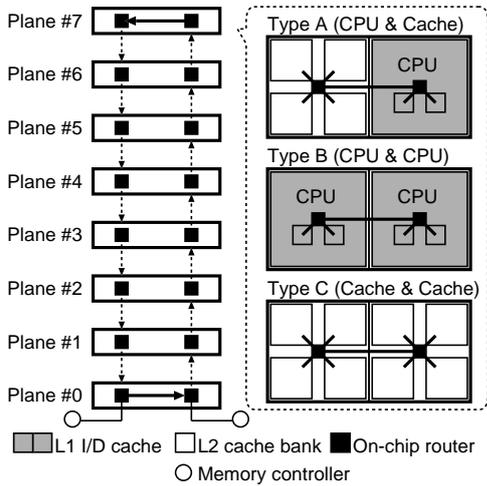


Figure 2: Proposed wireless 3-D CMP. Type B for computation-bound applications while Type C for memory-bound ones.

3.2 Wireless 3-D CMP Overview

Figure 2 illustrates the proposed wireless 3-D CMP architecture. The baseline 2-D CMP is divided into a number of planes, each of which has on-chip routers, processors, and/or cache banks. These planes are stacked vertically and connected by using the inductive coupling links. Memory controllers and external I/Os are connected to the bottom chip. We assume this simple 3-D SNUCA as a baseline for simplicity, although various optimization techniques for 3-D SNUCA are discussed in [14].

Three types of planes are illustrated in this figure: Type A (CPU and cache), Type B (CPU and CPU), and Type C (cache and cache). Typically, application performance is limited by either memory bandwidth or computation power; thus the applications can be classified into memory-bound and computation-bound ones [18]. Depending on the set of target applications, the wireless 3-D CMP using the inductive coupling enables us to customize the number and types of planes stacked in a package after the chip fabrication, such as more Type B planes for computation-bound applications while more Type C planes for memory-bound ones. This flexibility is attractive, because designing a new mask pattern for each set of applications is too costly in recent advanced process technologies. This is the firm advantage of the wireless 3-D CMPs over the conventional ones.

3.3 Network Design

Here, we consider the intra-plane and inter-plane networks separately. Their requirements are summarized as follows.

- **Intra-plane network** connects processors and cache banks on a single chip by using on-chip wires. Various types of chips will be stacked depending on the application requirements in the case of wireless 3-D CMPs; thus we should not expect any pre-determined network topology on it as long as each chip has data transceivers for uplink and downlink in pre-specified locations. That is, some chips may have 2-D mesh based intra-plane network, while the others may not have any intra-plane network except for the data transceivers for vertical communication. Needless

to say, each intra-plane network itself must be deadlock-free if it exists, because adding “deadlocked chips” will kill the whole CMP.

- **Inter-plane network** connects the data transceivers of all intra-plane networks in a package wirelessly by using the inductive coupling and forms a single vertical network. It should be simple and highly tolerable for adding, removing, and swapping the chips in a package. Of course, it must be deadlock-free.

To meet the above mentioned requirements, in this paper, we employ a uni-directional ring network for the inter-plane network as shown in Figure 2, because it is simple and easy to add, remove, and swap the nodes in a ring network without updating any routing information.

Typically, the downside of uni-directional ring is poor scalability on the throughput and communication latency. However, the inductive coupling link can offer ample throughput (e.g., 1-Tbit/sec using 1024 transceivers [20]). Also, communication latency is still reasonable for up to eight chips in a package, or we can just duplicate another uni-directional ring to form a single bi-directional ring. We will confirm this using a full system simulator assuming latency-sensitive shared-memory CMPs in Section 5.

Note that a ring network inherently contains a cyclic dependency, which introduces structural deadlocks. The router and flow control design for deadlock-freedom will be discussed in Section 3.4.

To sum up the network design for the wireless 3-D CMP, we summarize the rules each chip must comply as follows.

- **Network design rule 1:** Each chip has a pair of data transceivers for uplink and downlink in pre-specified locations for enabling the vertical communication.
- **Network design rule 2:** All processors and cache banks are connected to the data transceivers on the same chip via the intra-plane deadlock-free network.
- **Network design rule 3:** Only the intra-plane networks for the top and bottom chips must have a horizontal link that connects the uplink data transceiver and downlink data transceiver.

Network design rule 3 is required to form a single uni-directional ring (see Planes 0 and 7 in Figure 2). Otherwise, no such connection is needed.

3.4 Router and Flow Control Design

This section discusses the router and flow control design for the deadlock-free inter-plane ring network.

Various deadlock-free strategies have been used for ring networks, and they are summarized as follows.

- **VC-based approach** is the most conventional deadlock avoidance technique for rings. Two VCs (e.g., VC-0 and VC-1) are at least required for each message class. Packets are firstly transferred with VC-0 while VC-1 is used after they go over a wrap-around channel or dateline in the ring. Thus, the cyclic dependency of a ring is cut at the dateline.
- **Bubble flow approach** [1, 23] is also a deadlock avoidance technique for rings with virtual cut-through (VCT) switching. It does not require any VCs but requires a single buffer with capacity of at least two packets for each input port. By limiting the packet injection not to consume all the buffer resources in a ring, packets on the ring continuously move without any deadlocks.

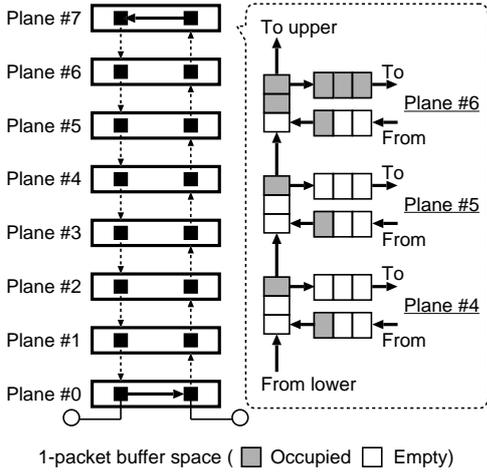


Figure 3: Vertical bubble flow network

- **Detection and recovery approach** [6] detects deadlocks when they really occur. Then, it recovers from the deadlock situation by discarding related packets or moving them to another network resource, such as escape VCs.

The detection and recovery approach requires a deadlock detection mechanism for routers. It also imposes a certain performance overhead for the recovery. Since a simple deadlock-free mechanism is preferable for on-chip purposes, we focus on the deadlock avoidance in this paper.

3.4.1 VC-Based Approach

Considering the avoidance approaches, we believe that the bubble flow approach is more reasonable compared to the VC-based one in shared-memory CMPs that rely on cache coherence protocols, in which packets with multiple message classes (e.g., cache request and reply) are transferred [27].

The downside of the conventional VC-based approach is the hardware complexity of additional VCs required for removing the dependencies between request and reply messages of a protocol. That is, it requires a separate virtual network for each message class to avoid the protocol deadlocks. Since it requires two VCs for each message class to avoid the structural deadlocks, for example, six VCs are required for MOESI directory-based protocol that defines three message classes, while ten VCs are required for MSI/MOSI directory-based protocol with five message classes. Router hardware complexity increases as the number of VCs increases.

However, the most serious problem of the VC-based approach is that the communication protocol to be used is limited, depending on the number of VCs available. For example, if the network equips only four VCs, it never uses the MOESI directory-based protocol that requires six VCs. Such limitation will harm the flexibility of the wireless 3-D CMPs that can customize their structure after the chip fabrication. On the other hand, the bubble flow approach stated below does not rely on any VCs for avoiding the protocol and structural deadlocks.

3.4.2 Vertical Bubble Flow

The bubble flow control [1, 23] is applied to the inter-plane vertical ring network of the wireless 3-D CMP.

Here, we illustrate its behavior using Figure 3. In this figure, each input port of the routers has a buffer with capacity of three packets. Each white box indicates an empty buffer with capacity of a single packet while gray box is an occupied one. The routers must follow the following flow control rules.

- **Flow control rule 1:** A packet on a ring can move to the next hop along the ring when the input buffer of the next hop has an empty space of at least one packet.
- **Flow control rule 2:** An intra-plane network can inject a packet to a ring when the vertical input buffer of the ingress router has an empty space of at least two packets. For example, Planes 4 and 5 can inject a packet to the ring, while Plane 6 cannot.
- **Flow control rule 3:** A packet can exit from a ring only when the horizontal output buffer of the egress router has an empty space of at least one packet. Otherwise, the packet must go around the ring again (i.e., miss routing). For example, packets destined to Planes 4 and 5 can exit from the ring, while that to Plane 6 cannot exit and must go around the ring until an empty space appears in the horizontal output buffer.

Based on the above mentioned rules, the packet injection is limited so as not to consume all the buffer resources in a ring, which guarantees that packets on the ring can continuously move. As long as the packets are continuously moving on the ring, no deadlocks occur even though multiple message classes coexist in the same virtual network. For more details about the bubble flow control, refer to [1] and [23].

Adding these rules to the conventional VCT router is simple. To confirm this, we implemented test chips to build the vertical bubble flow network using the inductive coupling links in Section 4.

Note that the miss routing, in which packets go around the ring again, significantly degrades the network performance; so we use routers with deep input buffers that can absorb three packets in order to suppress the miss routing. We compare the VC-based approach and vertical bubble flow control in terms of the application performance on CMPs in Section 5.

4. TEST CHIP IMPLEMENTATION

To show the feasibility of our proposal, we implemented a real test chip that has on-chip routers and inductive-coupling data transceivers using a 65nm process.

Figure 4 shows the top view of the test chip architecture, and Figure 5 shows the side view when four test chips are stacked. This test chip supports the following two communication schemes.

- Vertical bubble flow:** Each chip has a pair of data transceivers for uplink and downlink. The data transceiver for uplink receives data from a neighboring lower chip and transmits the data to the neighboring upper chip. The downlink is used for the opposite direction of the uplink.
- Vertical broadcast bus:** Each chip uses a single data transceiver that can switch its communication modes (i.e., TX and RX modes) in a system clock cycle. All chips share the same clock counter, and an 8-cycle time-slot is assigned for each chip periodically. A chip transmits data (if any) with TX mode when a time-slot is assigned to it. Otherwise, it must listen with RX mode.

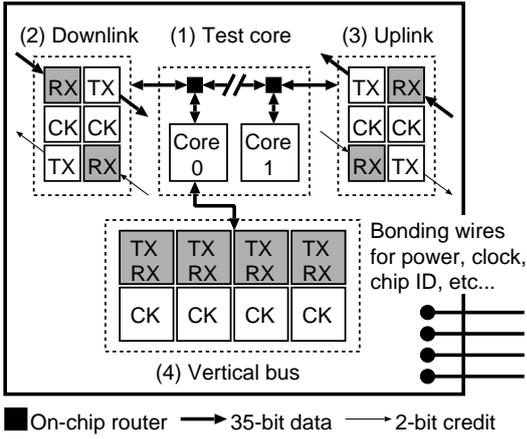


Figure 4: Test chip architecture (top view)

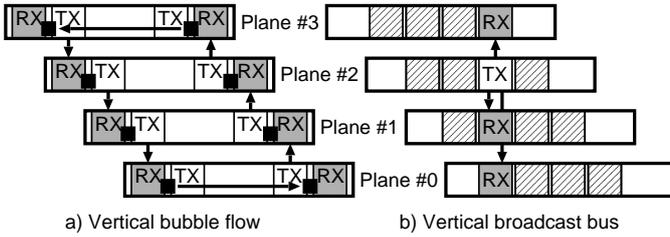


Figure 5: Test chip architecture (side view)

We chose the conservative design parameters for this chip. Table 1 lists the design parameters, and Figure 6 shows the layout of the test chip. We can duplicate the inductors for parallel transfer if a higher vertical bandwidth is required.

The test chip can be divided into the following four parts, as shown in Figures 4 and 6.

- 1) **Test core** consists of two cores and two routers. The core has a packet generator and a packet receiver with a 45-bit packet counter. One router is connected to the downlink data transceiver and another for the uplink. Although these two routers are connected via bi-directional on-chip wired link, only the top and bottom chips use one uni-directional link to form an inter-plane vertical ring.
- 2) **Downlink and 3) uplink inductors** consist of two pairs of TX and RX channels. One TX and RX pair is used for 35-bit flit transfer, while another pair for 2-bit credit back for flow control. The wireless data are transferred serially using the doubled communication rate of a 4GHz local clock shared by neighboring two chips. Thus, a TX channel can transmit a 35-bit flit in each 200MHz system clock.
- 4) **Vertical bus inductors** consist of four TX/RX channels and four clock channels. The TX/RX channel works as TX mode when a time-slot is assigned to the chip, while it is in RX mode otherwise.

In the vertical broadcast bus, only one TX/RX channel among four is used, depending on the chip ID, as shown in Figure 5 (b). For example, only the leftmost channel is used in Plane 0. Although this is inefficient for real use since the other three channels are not used, in this test chip, we chose this architecture in order to implement and test both

Process technology	Fujitsu CS202SZ 65nm
Chip size	2.1mm×2.1mm
System clock	200MHz
Router input buffer	16-flit FIFO
Flit size	32-bit data + 3-bit control
Packet size	5-flit
Inductor for bubble	150μm×150μm
Inductor for bus	250μm×250μm
Inductor bandwidth	35 [bit/cycle/channel]

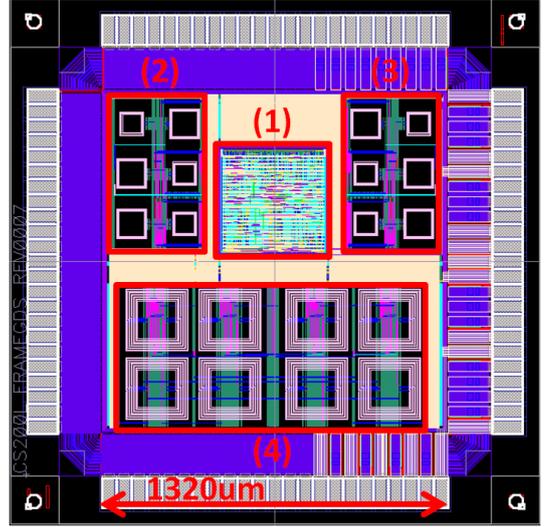


Figure 6: Layout of test chip

vertical bubble flow and vertical bus by using a single mask pattern.

The test chip is still under fabrication as of December 2010. Its test and evaluations are our future work.

5. EVALUATIONS

The vertical bubble flow control is compared with the conventional VC-based approach and vertical bus in terms of the throughput, application performance, and hardware amount.

5.1 Zero-Load Latency

Zero-load latency for the uni-directional ring, $T_{0,ring}$, is calculated as

$$T_{0,ring} = (H + 1)T_{router} + HT_{link} + L/BW, \quad (1)$$

where H is average hop count and L is packet length. T_{router} and T_{link} are latencies for transferring a header flit on a router and a link, respectively.

Zero-load latency for the vertical bus, $T_{0,bus}$, is calculated as

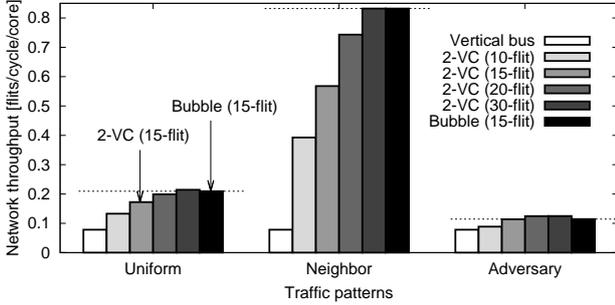
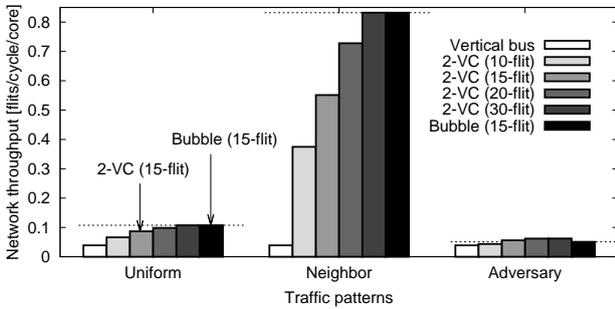
$$T_{0,bus} = T_{link} + L/BW + T_{slot}/N \sum_{i=0}^{N-1} i, \quad (2)$$

where N is number of chips stacked and T_{slot} is length of each time-slot. The rightmost term indicates the average waiting time to be assigned a time-slot.

Here, we assume the following three traffic patterns.

Table 2: Zero-load latency ($N = 4, 6, 8$) [cycle]

Traffic	Topology	$N = 4$	$N = 6$	$N = 8$
Uniform	Vertical ring	19	25	31
Neighbor	Vertical ring	10	10	10
Adversary	Vertical ring	28	40	52
Any	Vertical bus	18	26	34

**Figure 7: Network throughput ($N = 4$)****Figure 8: Network throughput ($N = 8$)**

- **Uniform traffic:** A source node sends packets to randomly selected destinations. Assuming each chip has two nodes, $H = N$.
- **Neighbor traffic:** A source node sends packets to the nearest destination. Thus, $H = 1$.
- **Adversary traffic:** A source node sends packets to the farthest destination. Thus, $H = 2N - 1$.

Table 2 shows the zero-load latencies for these traffic patterns, assuming $L = 5$, $T_{router} = 2$, $T_{link} = 1$, and $T_{slot} = 8$. Zero-load latency for the vertical bus is constant regardless of the traffic patterns applied. In the case of uniform traffic, zero-load latencies for the vertical ring and vertical bus are comparable.

5.2 Network Throughput

RTL simulations of the vertical bubble flow, conventional VC-based approach, and vertical bus are performed to measure their network throughputs for uniform, neighbor, and adversary traffics.

In the vertical bubble flow control, we implemented a 15-flit FIFO buffer for each input port. In the VC-based approach, two VCs are required for deadlock-freedom, assuming a single message class. Here, 2-VC (n -flit) means each input port has two VCs, each of which has an $(n/2)$ -flit FIFO buffer. The exception is 2-VC (15-flit). Because VCT

Table 3: Processor parameters ($N = 4, 8$)

Processor	UltraSPARC-III
L1 I/D cache size	64 KB (line:64B)
# of processors	4/8
L1 cache latency	1 cycle
L2 cache bank size	256 KB (assoc:4)
# of L2 cache banks	16/32
L2 cache latency	6 cycle
Memory size	4 GB
Memory latency	160 (± 2) cycle
# of memory controllers	2

Table 4: Network parameters ($N = 4, 8$)

Topology	Uni-directional ring
# of routers	8/16
Router pipeline	[RC/VSA][ST][LT]
Flit size	128 bit
Protocol	MOESI directory
# of message classes	3
Control packet size	1 flit
Data packet size	5 flit

switching is used for this router, 2-VC (15-flit) indicates the average throughput of the following two sub-configurations: 1) a 5-flit buffer for VC-0 and a 10-flit buffer for VC-1, and 2) a 10-flit buffer for VC-0 and a 5-flit buffer for VC-1. Notice that the buffer requirements of Bubble (15-flit) and 2-VC (15-flit) are the same.

Figures 7 and 8 show their network throughputs for 4-chip and 8-chip networks, respectively. As shown, Bubble (15-flit) outperforms 2-VC (15-flit) and is almost comparable to 2-VC (30-flit). Notice the throughput of Vertical bus is quite low compared to the ring-based networks and is constant regardless of the traffic patterns.

5.3 Application Performance

Full system simulations of the wireless 3-D CMPs that stack four and eight Type A chips (see Figure 2) are performed to measure the real application performance. As their communication schemes, the vertical bubble flow, conventional VC-based approach, and vertical bus are compared.

Tables 3 and 4 list the processor, memory, and network parameters. Each Type A chip has one processor, four shared L2 cache banks, and two on-chip routers. Two memory controllers are connected to the bottom chip, as shown in Figure 2. The cache architecture is SNUCA [10].

To simulate the wireless 3-D CMPs, we use a full-system multi-processor simulator: GEMS [16] and Wind River Simics [15]. We modified a detailed network model of GEMS, called Garnet [2], to accurately simulate the vertical bubble flow, conventional VC-based approach, and vertical bus.

A directory-based MOESI coherence protocol that defines three message classes is used. Thus, the VC-based approach requires six VCs for each input port, because each message class requires two VCs for avoiding structural deadlocks. We used default VC assignments of GEMS for this protocol.

Here, 6-VC (n -flit) means each input port has six VCs, each of which has an $(n/6)$ -flit FIFO buffer. We compare 6-VC (18-flit), 6-VC (30-flit), and Bubble (15-flit). Because

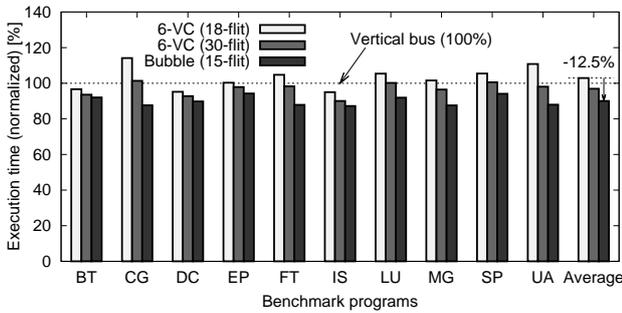


Figure 9: Application performance ($N = 4$)

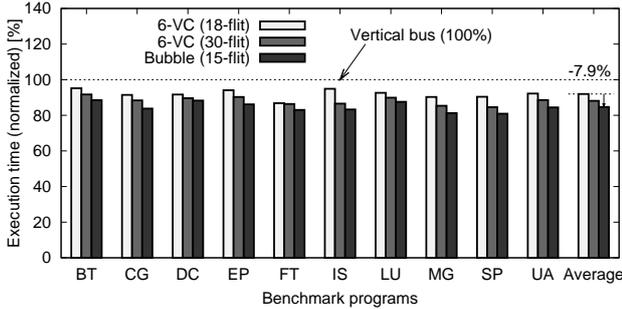


Figure 10: Application performance ($N = 8$)

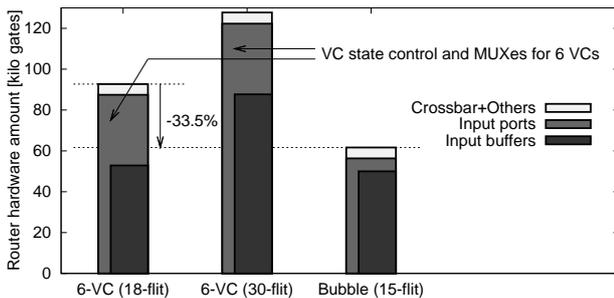


Figure 11: Router hardware amount (3 ports)

the packet length is up to 5-flit, 6-VC (30-flit) and Bubble (15-flit) use VCT switching, while 6-VC (18-flit) uses worm-hole. The buffer requirements of Bubble (15-flit) is less than that of 6-VC (18-flit).

To evaluate the application performance of these communication schemes on the wireless 3-D CMPs, we use ten parallel programs from the OpenMP implementation of NAS Parallel Benchmarks [8]. Sun Solaris 9 operating system is running on the 4-chip and 8-chip CMPs. These benchmark programs were compiled by Sun Studio 12 and are executed on Solaris 9. The number of threads was set to four or eight, depending on the number of Type A chips stacked.

Figures 9 and 10 show the application execution cycles of ten benchmark programs (BT, CG, DC, EP, FT, IS, LU, MG, SP, and UA) for 4-chip and 8-chip networks, respectively. The application execution time (Y-axis) is normalized so that the execution time using the vertical bus indicates 100%.

As shown, Bubble (15-flit) outperforms 6-VC (18-flit) by

7.9%-12.5%, because it uses a deep 15-flit FIFO buffer while 6-VC (18-flit) and 6-VC (30-flit) use shallow 3-flit and 5-flit FIFO buffers, respectively. Also, Vertical bus still outperforms 6-VC (18-flit) only for the 4-chip network. However, its performance is degraded as the number of chips stacked increases, such as the 8-chip network.

5.4 Router Hardware Amount

Using the RTL model of the test chip, 6-VC (18-flit), 6-VC (30-flit), and Bubble (15-flit) are compared in terms of the router hardware amount.

As shown in Figure 4, the router has three bi-directional ports: two bi-directional ports for a local core and another router on the same chip, and two uni-directional ports for routers on upper and lower chips, respectively. The flit size was set to 128-bit in this design, although the test chip conservatively employed 32-bit data width (Table 1). These routers were synthesized with Synopsys Design Compiler, and then placed and routed with Synopsys IC Compiler. We used a 65nm process for these designs.

Figure 11 shows their gate counts. As shown, input ports consume the most of the router area while the crossbar area is quite small, because the number of crossbar ports is only three in these routers. The input port area is divided into FIFO buffers and the other control circuits. Bubble (15-flit) and 6-VC (18-flit) have almost the same buffer areas. However, 6-VC (18-flit) requires more control circuits for supporting six VCs, such as the VC state controllers and VC multiplexers. As a result, Bubble (15-flit) requires a 33.5% smaller router area compared to 6-VC (18-flit).

6. SUMMARY AND FUTURE WORK

In this paper, we proposed a wireless 3-D NoC architecture for CMPs, in which the number and types of CMP chips stacked in a package can be changed after the chip fabrication, by using the inductive coupling. These chips form a single vertical ring network so as to fully exploit the flexibility of the wireless that enables us to add, remove, and swap the chips in a package without updating any routing information. As the communication schemes, we compared the vertical bubble flow control, conventional VC-based approach, and vertical bus, in terms of the latency, throughput, hardware amount, and application performance.

Below are the observations from the evaluation results.

- **Vertical bus** is simple and still low-latency (Table 2). However, its network throughput is quite low (Figures 7 and 8) and its application performance is also lower than the other approaches in the 8-chip network (Figure 10).
- **VC-based approach** is the most conventional deadlock avoidance technique for rings. However, the required number of VCs increases depending on the number of message classes on the network. Also, network protocols to be used are limited by the number of VCs already equipped. This will harm the flexibility of the wireless 3-D CMPs that can customize their structure for given purposes.
- **Vertical bubble flow** does not rely on any VCs; thus the communication protocols to be used are not limited by the number of VCs available. It outperforms the VC-based approach by 7.9%-12.5% (Figures 9 and 10) with a 33.5% smaller router area (Figure 11). However, the performance improvement decreases slightly as the number of chips increase, since the negative impacts of the

miss routing on performance increases as the ring size is enlarged.

As future work, we will focus on the following issues.

- More scalable wireless 3-D NoC architecture: We employed the ring network with bubble flow control, since the network sizes we assumed were eight chips at most in this paper. As more scalable approach, we are implementing a plug-and-play protocol that detects whole network structure at boot time and then configures the routing tables based on multiple spanning trees.
- Evaluations using the test 3-D chip: The test 3-D chip presented in Section 4 is still under fabrication as of December 2010. We will evaluate it as our future work. Also, we are planning to produce a more practical test chip, in which the packet generator/counter cores of the current chip will be replaced with MIPS R3000 processors.

7. REFERENCES

- [1] P. Abad, V. Puente, P. Prieto, and J. A. Gregorio. Rotary Router: An Efficient Architecture for CMP Interconnection Networks. In *Proceedings of the International Symposium on Computer Architecture (ISCA'07)*, pages 116–125, May 2007.
- [2] N. Agarwal, L.-S. Peh, and N. Jha. Garnet: A Detailed Interconnection Network Model inside a Full-system Simulation Framework. Technical Report CE-P08-001, Princeton University, 2008.
- [3] B. M. Beckmann and D. A. Wood. Managing Wire Delay in Large Chip-Multiprocessor Caches. In *Proceedings of the International Symposium on Microarchitecture (MICRO'04)*, pages 319–330, Dec. 2004.
- [4] B. Black, M. Annavaram, N. Brekelbaum, J. DeVale, L. Jiang, G. H. Loh, D. McCaule, P. Morrow, D. W. Nelson, D. Pantuso, P. Reed, J. Rupley, S. Shankar, J. P. Shen, and C. Webb. Die Stacking (3D) Microarchitecture. In *Proceedings of the International Symposium on Microarchitecture (MICRO'06)*, pages 469–479, Dec. 2006.
- [5] 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, Feb. 2001.
- [6] W. J. Dally and B. Towles. *Principles and Practices of Interconnection Networks*. Morgan Kaufmann, 2004.
- [7] W. R. Davis, J. Wilson, S. Mick, J. Xu, H. Hua, C. Mineo, A. M. Sule, M. Steer, and P. D. Franzon. Demystifying 3D ICs: The Pros and Cons of Going Vertical. *IEEE Design and Test of Computers*, 22(6):498–510, Nov. 2005.
- [8] H. Jin, M. Frumkin, and J. Yan. The OpenMP Implementation of NAS Parallel Benchmarks and Its Performane. In *NAS Technical Report NAS-99-011*, Oct. 1999.
- [9] K. Kanda, D. D. Antono, K. Ishida, H. Kawaguchi, T. Kuroda, and T. 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, Feb. 2003.
- [10] C. Kim, D. Burger, and S. W. Keckler. An Adaptive, Non-Uniform Cache Structure for Wire-Delay Dominated On-Chip Caches. In *Proceedings of the International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS'02)*, pages 211–222, Oct. 2002.
- [11] J. Kim, C. Nicopoulos, D. Park, R. Das, Y. Xie, N. Vijaykrishnan, M. Yousif, and C. Das. A Novel Dimensionally-Decomposed Router for On-Chip Communication in 3D Architectures. In *Proceedings of the International Symposium on Computer Architecture (ISCA'07)*, pages 138–149, 2007.
- [12] K. Kumagai, C. Yang, S. Goto, T. Ikenaga, Y. Mabuchi, and K. Yoshida. System-in-Silicon Architecture and its application to an H.264/AVC motion estimation fort 1080HDTV. In *Proceedings of the International Solid-State Circuits Conference (ISSCC'06)*, pages 430–431, Feb. 2006.
- [13] F. Li, C. Nicopoulos, T. Richardson, Y. Xie, V. Narayanan, and M. Kandemir. Design and Management of 3D Chip Multiprocessors Using Network-in-Memory. In *Proceedings of the International Symposium on Computer Architecture (ISCA'06)*, pages 130–141, June 2006.
- [14] N. Madan, L. Zhao, N. Muralimanohar, A. N. Udipi, R. Balasubramonian, R. Iyer, S. Makineni, and D. Newell. Optimizing Communication and Capacity in a 3D Stacked Reconfigurable Cache Hierarchy. In *Proceedings of the International Symposium on High-Performance Computer Architecture (HPCA'09)*, pages 262–274, Feb. 2009.
- [15] P. S. Magnusson et al. Simics: A Full System Simulation Platform. *IEEE Computer*, 35(2):50–58, Feb. 2002.
- [16] M. M. K. Martin, D. J. Sorin, B. M. Beckmann, M. R. Marty, M. Xu, A. R. Alameldeen, K. E. Moore, M. D. Hill, and D. A. Wood. Multifacet General Execution-driven Multiprocessor Simulator (GEMS) Toolset. *ACM SIGARCH Computer Architecture News (CAN'05)*, 33(4):92–99, Nov. 2005.
- [17] H. Matsutani, M. Koibuchi, Y. Yamada, D. F. Hsu, and H. Amano. Fat H-Tree: A Cost-Efficient Tree-Based On-Chip Network. *IEEE Transactions on Parallel and Distributed Systems*, 20(8):1126–1141, Aug. 2009.
- [18] A. Merkel, J. Stoess, and F. Bellosa. Resource-Conscious Scheduling for Energy Efficiency on Multicore Processors. In *Proceedings of the European Conference on Computer Systems (EuroSys'10)*, pages 153–166, Apr. 2010.
- [19] N. Miura, H. Ishikuro, T. Sakurai, and T. 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, Feb. 2007.
- [20] N. Miura, D. Mizoguchi, M. Inoue, K. Niitsu, Y. Nakagawa, M. Tago, M. Fukaishi, T. Sakurai, and T. 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, Feb. 2006.
- [21] D. Park, S. Eachempati, R. Das, A. K. Mishra, V. Narayanan, Y. Xie, and C. R. Das. MIRA: A Multi-layered On-Chip Interconnect Router Architecture. In *Proceedings of the International Symposium on Computer Architecture (ISCA'08)*, pages 251–261, 2008.
- [22] V. F. Pavlidis and E. G. Friedman. 3-D Topologies for Networks-on-Chip. *IEEE Transactions on Very Large Scale Integration Systems*, 15(10):1081–1090, Oct. 2007.
- [23] V. Puente, R. Beivide, J. A. Gregorio, J. M. Prellezo, J. Duato, and C. 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, Sept. 1999.
- [24] R. S. Ramanujam and B. Lin. Randomized Partially-Minimal Routing on Three-Dimensional Mesh Networks. *IEEE Computer Architecture Letters*, 7(2):37–40, July 2008.
- [25] S. Saito, Y. Kohama, Y. Sugimori, Y. Hasegawa, H. Matsutani, T. Sano, K. Kasuga, Y. Yoshida, K. Niitsu, N. Miura, T. Kuroda, and H. Amano. MuCCRA-Cube: a 3D Dynamically Reconfigurable Processor with Inductive-Coupling Link. In *Proceedings of the Field-Programmable Logic and Applications (FPL'09)*, pages 6–11, Sept. 2009.
- [26] A. Sheibanyrad, F. Petrot, and A. Janstch. *3D Integration for NoC-Based SoC Architectures*. Springer, 2010.
- [27] Y. Solihin. *Fundamentals of Parallel Computer Architecture*. Solihin Publishing & Consulting LLC, 2009.