

A thermal management system for building block computing systems

Yu Fujita¹, Kimiyoshi Usami², and Hideharu Amano¹

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

²Shibaura Institute of Technology
3-7-5 Toyosu, Kohtoh-ku,
Tokyo, Japan
usami@shibaura-it.ac.jp

Abstract

Cube-1 is a heterogeneous multiprocessor consisting of 3D stacked chips connecting with inductive coupling through chip interface (TCI). The most important problem of Cube-1 is the thermal management. Unlike TSV which can be used for heat dissipation, stacked chips are electrically contact-less in inductive coupling TCI.

First, by measuring the relationship between the chip temperature and leakage monitor, we examined that the leakage monitor can be used as a temperature sensor of the chip. Then, we measured the thermal characteristics of Cube-1 by leakage monitors. The chip temperature change due to the internal power was evaluated, and it appeared that the chip temperature was not changed with this level of power consumption even if the chip was sandwiched with other chips. The heat conductance through the stacked chip was also evaluated. Evaluation results show that the heat dissipation of the chip sandwiched with other chips is almost the same as that of the chip placed top on the stack.

Finally, we proposed the supply voltage control system of the stacked chip by making the best use of the chip temperature data from the leakage monitor. By using the proposed control, the energy efficiency can be improved by 5% at maximum.

I. Introduction

Recent mobile devices require various functions with wide range of performance and power consumption, and it is difficult to cope with them by a single System-on-Chip (SoC). Building block computational systems[1] can tailor a system for each requirement by stacking various types of chips with inductive coupling through-chip interface (TCI).

The first prototype Cube-1[2] a heterogeneous multicore system with R3000 compatible host and reconfigurable accelerators is now available.

The most important problem of the building block computational system is thermal management. Unlike TSV which can be used for heat dissipation, stacked chips are electrically contact-less in inductive coupling TCI. Although they are tightly stuck by a bonding agent, thermal dissipation is not so expected, and some management system is required. In this paper, we first evaluate the temperature of each chip by using leak monitors embedded in each chip, and temperature of the sandwiched chip is shown. The results show that at least Cube-1 which uses low power host CPU and accelerators, the risk of thermal run-away is quite low. However, for the case of emergency, the thermal management system might be required. Here, we proposed the thermal management system which can improve the energy efficiency based on the data from the leak monitor.

The rest of the paper is formed as follows. Section 2 describes our target building block computing system, Cube-1. Leakage monitor is introduced and the relationship between chip temperature is analyzed in Section 3. In Section 4, thermal characteristics of Cube-1 are evaluated. Section 5 proposes voltage control considering the chip temperature measured with the leakage monitor. Section 6 concludes this paper.

II. Related Work

Run-time thermal management for three dimensional chip-multiprocessors have been widely researched[3]. Most of them distributes the load of core so as not to concentrate heat generation vertically in the system or controlling hotspots[4]. The thermal aware mapping is also researched for FPGA[5] and Network-on-Chips (NoCs)[6].

Although these researches are based on accurate heat modeling for TSV connected 3D IC, temperature monitor and workload monitor, no experimental results were presented. A practical power management scheme for 3D MPSoC has been proposed[7], it also based on the theoretical model based on TSV.

This research is different from previous researches in the following two issues.

- The target is 3D multicore processors connected with wireless inductive coupling TCI, while previous researches focused on TSV connected 3D multicore systems.
- Real chip evaluation results are presented, while most of previous researches are based on theoretical model.

In other words, since theoretical thermal model for wireless inductive coupling TCI has not been well established, this research is just at the beginning stage. Nevertheless, this kind of study is needed to establish theoretical thermal model.

III. Cube-1

Cube-1 is the first prototype building block computational system[1] consisting of a R3000 compatible host CPU Geyser and reconfigurable accelerators CMAs. The specification and design environment are shown in Table I and Table II.

TABLE I. Cube-1 Chip specification

Proc-ess	Technology	65nm CMOS (12-Metal)
	Chip Area	2.1mm × 4.2mm
	Core Area	1.5mm × 3.6mm
Host CPU	CPU Core	MIPS R3000 Compatible
	Cache	4KB 2way Instruction Cache 4KB 2way Data Cache 16-Entry Shared TLB
	I/O	3Gb/s TCI 100Mb/s 32bit External I/O
	Supply Voltage	Core+TCI: 1.2V External I/O: 3.3V
	No. of Transistors	8k for TCI, 1249k for others
Accele-rator	PE Array	64 (8×8) PEs
	Micro-Controller	1Cycle Non-Pipelined
	Memory	25bit 2KB 2Bank Data Memory 14bit 128depth Instruction Memory
	I/O	3Gb/s TCI × 2Channels for Up/Down links
	Supply Voltage	PE Array: 0.5-1.2V DVS Core+TCI: 1.2V External I/O: 3.3V
	No. of Transistors	16k for TCI, 2008k for CMA
3D Proc-essor	System Clock	50-100MHz
	Chip Stack	Staircase Stacking Hot CPU+Accelerator × 1-3 Stack
	Chip Thickness	40-80μm(Bottom Chip:300μm)

TABLE II. Design Environment

Digital Design	Verilog HDL
Simulation	Cadence ncverilog
Synthesis	Synopsys Design Compiler 2009.06.SP5
Layout	Synopsys IC Compiler 2008.09.SP4
Verification	Mentor Calibre 2010.4
Analog Design	Vertuoso 5.1.41

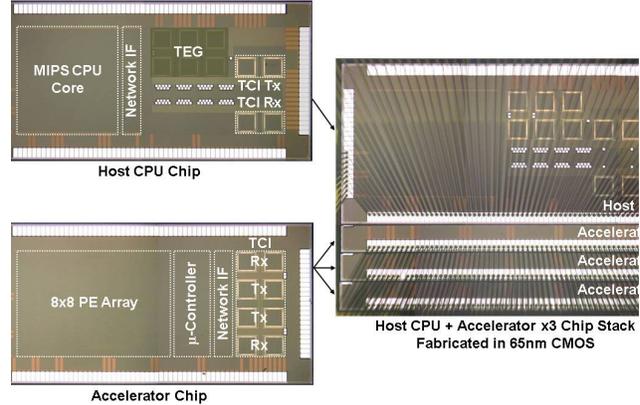


Fig. 1. Chip Stacking

A. TCI

The most important feature of Cube-1 is its inductive coupling through chip interface (TCI). The data is transferred between two coils stacked exactly on the same position of the chips. Since each coil is formed using a normal wires, any additional process is needed to a common CMOS process unlike TSV or other SiP techniques. The data transfer speed is more than 2Gb/sec at maximum with an extremely low error rate less than 10^{-11} . The power consumption is of each channel is less than 10mW. Although the footprint of each coil is much larger than that of TSV, common digital circuits can be implemented inside the coil. That is practical overhead of a coil is just area of wires used to form it.

When chips are stacked two communicating coils must be placed exactly the same place. For locating a sender coil over the receiver coil, we provide a certain length of gap between chips as shown in Figure 1. This gap can be also utilized for the bonding space to supply VDD, GND and digital clock signals.

B. NoC using TCI

A single directional ring network is formed by shifting stacked chips as shown in Figure 2. Each chip has two routers and they are connected with upper or lower channel as well as the memory of the CPU and accelerator. Just

by setting information that the chip is the top, middle or bottom, the ring network is formed automatically.

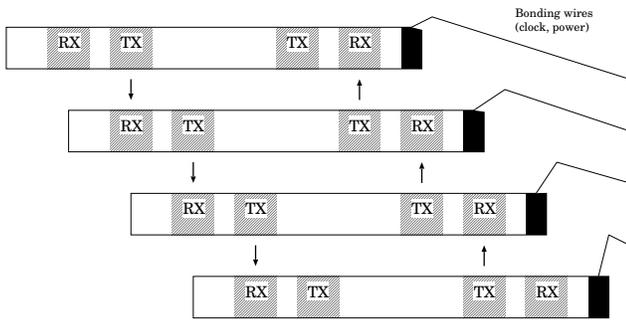


Fig. 2. Ring Network of Cube-1

Babble flow control [8] [9] is adopted in order to allow deadlock free packet transfer without caring the topology of inner chip network and the number of virtual channels. The drawback of the method is out-of-order delivery of packets caused by unfortunate packets circulating in the ring long time. However, since Cube-1 is shared memory system, and exceptional out-of-order delivery is not a serious problem.

C. Geyser-Cube

Geyser-Cube is an R3000 compatible low power embedded processor, and placed on the top of the chip stack for keeping a lot of I/O pins. It provides a CPU with standard 5-stage pipeline, 8KB instruction/data cache and 64-entry shared TLB. Linux OS and an embedded OS are running on this chip. As shown in left upper side of Figure 1, CPU itself is implemented on the left half, and others are TCI test circuits in which the rightmost two coils are used for forming the ring network. The most interesting property of Geyser-Cube is its fine-grained power gating but it is omitted in this paper.

D. CMA-Cube

CMA-Cube is a reconfigurable accelerator using Cool Mega Array (CMA) in its core. CMA is consisting of a large array of processing elements (PEs), a simple light weight microcontroller and data memory. In order to reduce the power for storing the intermediate results and clock distribution, the PE array is consisting of a pure combinatorial circuits. A data-flow graph of the target application program is directly mapped on the PE array statically without power hungry dynamic reconfiguration. Instead, the programmable microcontroller flexibly aligns the data at input of the PE array and stores the results in

the data memory. The computation of the PE array and data reading/writing from/to data memory are executed in the pipelined manner. When the delay of the PE array is smaller than the data manipulation time of the microcontroller, dynamic voltage scaling (DVS) is applied to the PE array. Since the PE array is consisting of pure combinatorial logic, the supply voltage can be lowered to 0.5V.

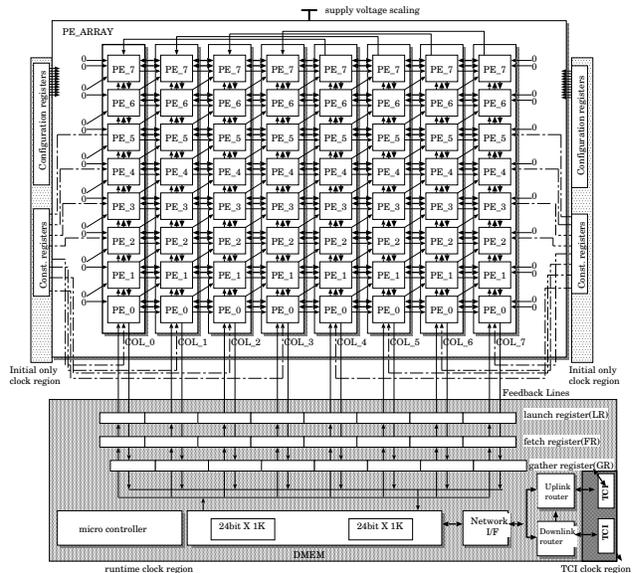


Fig. 3. Block diagram of CMA-Cube

As shown in Figure 3, CMA-Cube has 8×8 PE array like CMA-1, TCI is connected to the data memory through the interface module. CMA is implemented in the left side of the chip as shown in Figure 1 and TCI is located in the right side, The memory of Geyser-Cube and CMA-Cube chips are mapped into a single logical address space. The block data transfer is done using the cache control mechanism of Geyser-Cube. Also, DMA controller embedded in the interface manages data transfer between multiple CMA-Cubes directly.

IV. Leakage Monitor

Both Geyser-Cube and CMA-Cube provide a leakage monitor[10] for controlling the fine grained power. It collects the leakage current of the circuit and returns the time to reach a certain value. Since the leakage current increases as the chip temperature, the value from the leakage monitor can be used to evaluate chip temperature. It is difficult to stick thermometer on chips sandwiched with other chips, and both chips do not provide the embedded thermal diode[11] dedicated for evaluating temperature.

Here, we used the leakage monitor for evaluating the chip temperature, though it is not the best way.

A. Circuit of the leakage monitor

Figure 4 shows the circuit of the leakage monitor. When EN_{model} is asserted, N2 is initialized. The measurement starts when N2 turns off by negating EN_{model} . With the same signal, the counter starts to count. The leakage current going through P1 is collected in the virtual ground (VGND). The leakage current from P1 is larger, the time to increase the VGND level is shorter. When VGND level reaches the reference voltage V_{REF} , N2 turns off and the counter stops. From the value of the counter, the time to reach the V_{REF} is obtained.

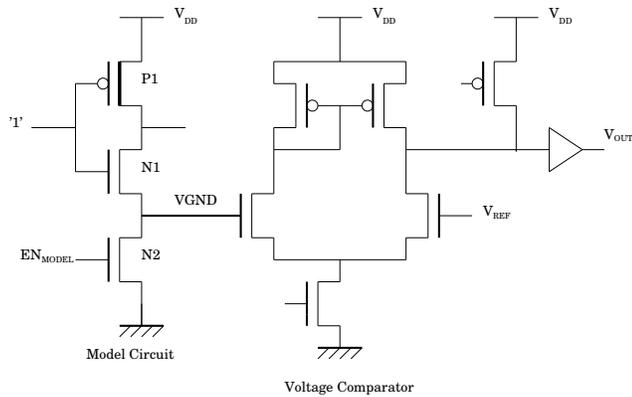


Fig. 4. Circuit of Leakage Monitor

B. Temperature versus the counter value

In order to evaluate the chip temperature by the counter value from the leakage monitor, we measured them with the environment shown in Figure 5. The crosscutting view is shown in Figure 6. For heating the chip, a compact thermal system (CTS) which can change the temperature of the heat metal from -20 to 150 centigrade. As shown in Figure 6, the heat metal of CTS is not directly touch on the chip but warms through a thermal conductive gum and the board. In order to keep the temperature around the chip stable, the chip is surrounded by the heat isolation material. In order to evaluate the relationship between temperature and counter value, we used Geysler-Cube and CMA-Cube without chip stacking. Each chip is solely implemented on the board and wires of a thermocoupled thermometer are stuck directly on the chip.

Figure 7 shows the relationship between temperature versus the counter value of the leakage monitor in CMA-Cube. Almost the same results were obtained in Geysler-

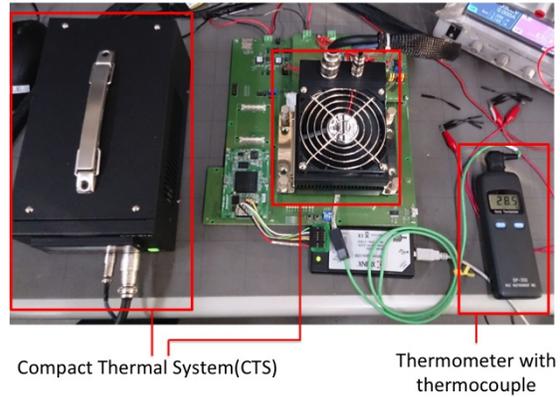


Fig. 5. Measurement Environment

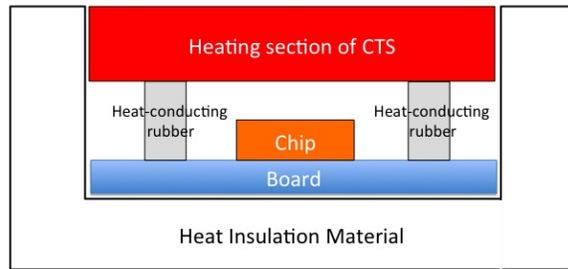


Fig. 6. Cross-cutting view

Cube, and by using the table reference, we can estimate the chip temperature by the value of the leakage monitor.

V. Thermal Characteristics

One of most important problems of Cube-1 is a possibility of thermal run-away inside the chip sandwiched with other chips. First, we examined the temperature change due to internal power consumption, and then evaluate the influence of the external heat.

A. The temperature charger due to the internal power

The chip on the top of the stack can dissipate heat directly to the air and the bottom chip can do it through the board. Thus, the thermal problem may occur in the middle layer. Here, we measured the temperature of the CMA-Cube located in the second layer as shown in Figure 8. In order to generate heat as possible, we gave maximum supply voltage (1.2V) to PE array and executes the heaviest

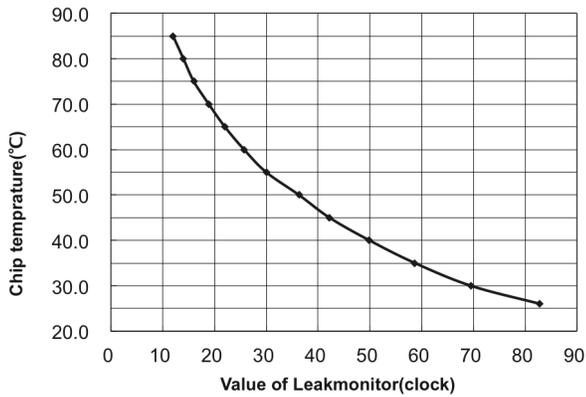


Fig. 7. Chip temperature vs. Counter value of the leak monitor

job. First, Geyser-Cube in the first layer reads the data from outside memory, then it executes it with the burst data transfer. Then, CMA-Cube executes an image filter application and returns the results to Geyser-Cube. Since both CMA-Cube and wireless TCI work almost all time, the CMA-Cube chip consumes about 100mW. After the computation of an image, the leakage monitor is accessed and temperature is recorded. We executed this loop more than 2 hours.

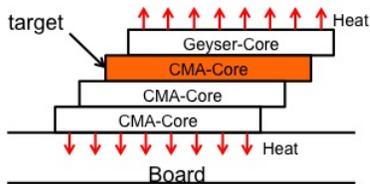


Fig. 8. Target of measurement

The measurement result is shown in the Figure 9. There was completely no change in the chip temperature. It appears that the chip temperature was not changed with this level of power consumption even if the chip was sandwiched with other chips. That is, the chip stack is not dangerous at least for embedded usage with under 100mW power consumption. For increasing the power consumption, we need to increase the operational clock frequency of CMA-Cube, but with the current configuration, it is difficult. We plan to embed special circuits to generate heat in the next trial.

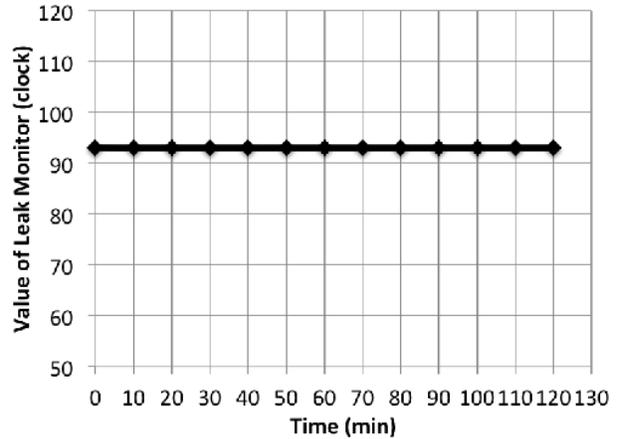


Fig. 9. Time versus the counter value

B. Heat conductance through the chip

The next concern is how the heat of the chip stacked in the middle layer is dissipated through other chips, that is heat conductance through the chip. In order to examine the change of the chip temperature when external heat is given, we used the measurement environment shown in Figure 5. Four-chip stack is used, and chip temperature of Geyser-Cube located the top and CMA-Cube is measured with the leak monitor embedded in each chip.

First, we kept the temperature of the CTS 100 centigrade, put on the top of the chip, and then starts the measurement. We measured the chip temperature every 30 seconds.

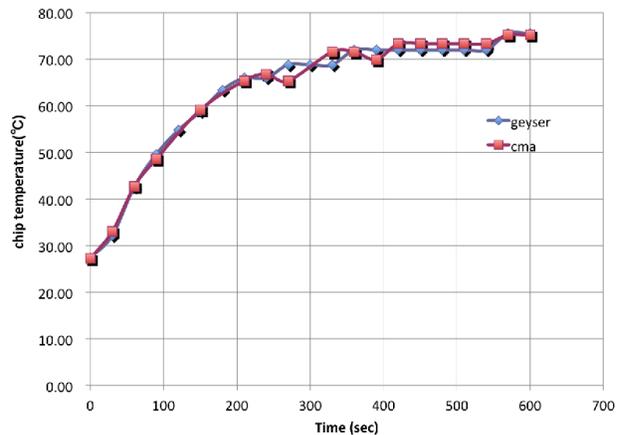


Fig. 10. Changes of the chip temperature at heating

As shown in Figure 10, chip temperature of the top Geyser-Cube rose slightly faster than that of CMA-Cube in

the second layer. However, the difference is very small and the temperature of the second layer chip quickly followed the first chip.

Next, we heated the whole chip stack 100 centigrade enough long time to all chips are well heated. Then we removed the CTS and measured chip temperature using leak monitor embedded in each chip every 60 seconds.

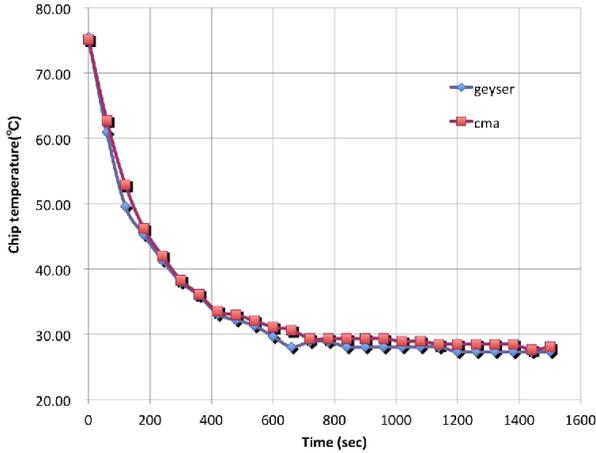


Fig. 11. Changes of the chip temperature at cooling

As shown in Figure 11, the temperature of the top chip, Geyser-Cube went down rapidly, but the second CMA-Chip followed quickly. The difference is also small.

From these measurement, we can say that the heat dissipation of the chip sandwiched with other chips is almost the same as that of the chip placed top on the stack. This comes from the fact that in Cube-1 which uses high speed wireless TCI, chips are tightly stacked with bonding agent. There is almost no room between multiple chips. This implementation enables to give enough heat conductance between chips.

VI. Voltage control considering the chip temperature

Although the heat dissipation of Cube-1 is enough for current configuration, for the emergency, the voltage management system should be required. The system controls the supply voltage according to the counter value from the CMA-cube. If the temperature becomes more than 80 centigrade, the process in CMA-cube is stopped and the supply voltage is lowered. If the temperature becomes enough low, the application program starts again. However, if it is provided only for the emergency state, the required hardware may be too heavy. So, we propose a voltage management system which can improve the

energy efficiency using the counter value from the leakage monitor.

A. Delay and temperature

The program of CMA-cube is designed with a certain time margin so that the correct results can be obtained even when the chip temperature is extremely high. If the chip temperature can be evaluated precisely by the leak monitor, we can reduce the supply voltage of the CMA-cube with the minimum time margin under the temperature.

In order to manage in such manner, we evaluated the relationship between the minimum supply voltage to the PE array versus chip temperature. Figure 12 shows the evaluation results using a real chip. Parameters are programmed delay value used in CMA-cube depending on the application program. The delay is depending on the chip temperature, although its variation is not so large as expected by the specification from datasheet. For example, from Figure 12 shows that the supply voltage that achieves 10nsec delay can be lowered from 1.6V to 1.52V when the temperature is 30 centigrade (the data from leakmonitor is about 70clock cycles). By lowering the supply voltage, we can save about 5% energy consumption.

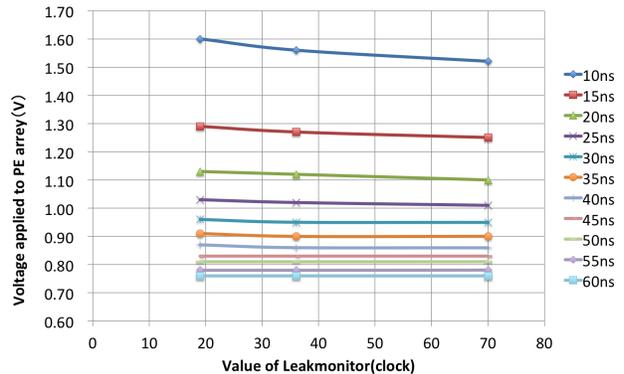


Fig. 12. Minimum voltage supply vs. chip temperature

B. Flowchart of voltage management

A flowchart of the voltage control considering the chip temperature can be drawn as shown in Fig. 13.

First, the value of the leakage monitor is checked. If the value is larger than a limit, the power supply of the CMA-Cube must be shut off in order to save the chip. Then, the value of the leakage monitor is checked again after a few minutes. If the value shows the normal temperature, the optimum supply voltage of each chip temperature is applied. This loop is iterated until the end of calculate.

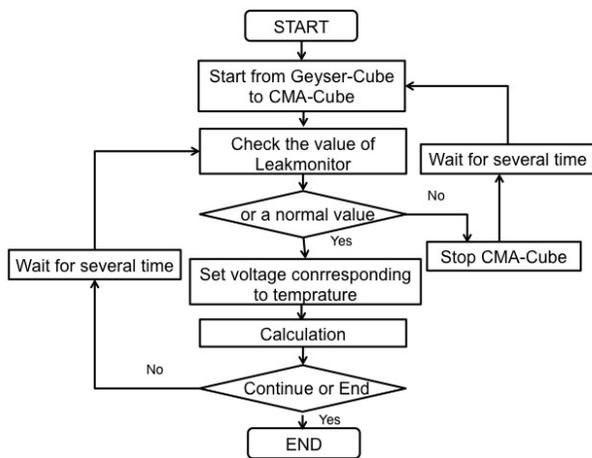


Fig. 13. Voltage control flowchart

VII. Conclusion

Heat dissipation problems for three-dimensional stacked multi-core processors Cube-1 were analyzed. By measuring the relationship between the chip temperature and leakage monitor, we examined that the leakage monitor can be used as a temperature sensor of the chip. Then, we measured the thermal characteristics of Cube-1 by the leakage monitor. First, the chip temperature change due to the internal power was evaluated. It appeared the chip temperature was not changed with this level of power consumption even if the chip was sandwiched with other chips. Second, heat conductance through the stacked chip was evaluated. Evaluation results show that the heat dissipation of the chip sandwiched with other chips is almost the same as that of the chip placed top on the stack.

Finally, we proposed the supply voltage control system by making the best use of the chip temperature data from the leakage monitor. By using the proposed control, the energy efficiency can be improved by 5% at maximum.

Since the current Cube-1 is designed as a low power system for embedded processing, the internal power from CMA-Cube is up to 100mW and the measured results here are only for low power systems in this class. Also the number of stacked chip is only four. The chip temperature and heat conductance problems will be more severe when 10's of chips each of which consumes more than 1W are stacked. For making the wireless TCI a practical implementation techniques, evaluation under such tough environment is necessary and it is our future work. Also, establishing thermal model which can be used for theoretical researches for wireless inductive TCI is also our future work.

References

- [1] N. Miura, et al, "A-Scalable 3D Heterogeneous Multicore with an Inductive ThruChip Interface," in *IEEE Micro*, Vol.33, No.6, 2013, pp. 6–15.
- [2] Y.Koizumi, H.Amano, H.Matsutani, N.Miura, T. Kuroda, R. Sakamoto, M. Namiki, K. Usami, M. Kondo, and H. Nakamura, "Dynamic power control with a heterogeneous multi-core system using a 3-D wireless inductive coupling interconnect," in *ICFPT*, Dec 2012, pp. 293–296.
- [3] Changyun Zhu, et al., "Three-Dimensional Chip-Multiprocessor RunTime Thermal Management," in *IEEE Transaction on Computer-Aided Design of Integrated Circuits and Systems*, Aug 2008, pp. 1479–1492.
- [4] M.Puttaswamy and G.H.Loh, "Thermal herding: Microarchitecture techniques for controlling hotspots in high performance 3D-integrated processors," in *Proc. of Int. Symp. on High Performance Computer Architecture*, Feb 2007, pp. 193–204.
- [5] A. Wold, D. Koch, and J. Torresen, "Thermal Aware Modle Placement for Heterogeneous 3D-IC Based FPGAs," in *Proceedings of the IEEE 27th International Symposium on Parallel and Distributed Processing Workshop (IPDPSW)*, 2013, pp. 281–286.
- [6] C.-H. Chao, K.-Y. Jheng, H.-Y. Wang, and J.-C. Wu, "Traffic- and Thermal Aware Run-Time Thermal Management Schime for 3D NoC Systems ," in *Proceedings of the IEEE 27th International Symposium on Networks-on-Chip(NOCS)*, 2010, pp. 223–230.
- [7] Arnica Aggarwal, Sumeet S.Kumar, Amir Zjajo, Rene van LeuKen, "Temperature Constrained Power Management Scheme for 3D MPSoC," in *Proceedings of IEEE 16th Workshop on Signal and Power integrity (SPI'12)*, May 2012, pp. 7–10.
- [8] 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)*, Sep. 1999, pp. 58–67.
- [9] 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 *Networks on Chip(NoCS)*, May 2011, pp. 49–56.
- [10] K.Usami, Y.Goto, S.Koyama, D.Ikebuchi, H.amano, H.nakamura, "On-Chip detection methodology for break-even time of power gated function units," in *proceedings of the International Symposium on Low Power Electronics and Design (ISLPED'10)*, Aug. 2011, pp. 241–246.
- [11] N.H.E.Wesgte, D.M.Harris, "CMOS VLSI Design: A Circuit and SYstems Perspective ," *Addison Wesley Pub. Co.*, 2011.