Thermal Management of Fujitsu's High-Performance Servers

● Jie Wei

(Manuscript received July 6, 2006)

With the continued increase in power dissipation and power density of high-performance microprocessors, as well as requirements for high-density packaging and lower device junction temperatures, improvements of electronic cooling technologies have been becoming strategically important in the challenge of advanced thermal solutions for achieving higher cooling efficiencies while meeting reliability, packaging, cost, and environmental requirements appropriate to various electronic equipments. This paper gives an overview of thermal design and cooling technology development for Fujitsu's high-performance servers, using the latest high-end UNIX server PRIMEPOWER 2500 as an example. The thermal management is outlined from viewpoints of the server cabinet, system board, and CPU package. It also discusses the challenges in cooling technology developments arising from thermal management of high-density and asymmetric CPU power dissipation, investigations of thermal interface and heat spreading materials, and enhancements of heatsink cooling capabilities.

1. Introduction

High-performance computer servers are widely used in leading-edge research, development, and service fields, where enormous resources, data processing abilities, and calculating abilities are needed. To achieve high-speed and large-scale transmission performance, servers must simultaneously have high reliability, effective cost-performance, and environmental features such as low power consumption, low noise, compact size, and use of green materials. Generally, technologies of system packaging, thermal management, and power supply are considered essential in server design and development.

The power dissipation of high-performance microprocessors has been increasing continuously in the past decade. **Figure 1** shows the trend of CPU power dissipation as predicted by the International Technology Roadmap for Semiconductors (ITRS),[1] together with those of some high-performance servers from major computer manufacturers. The power dissipation of high-performance CPU processors is predicted to exceed 150 W in the near future. The increased local power densities and air-cooling temperatures due to design complexities and high-density packaging, as well as the need for lower CPU junction temperatures to achieve higher reliability and less power leakage, are making it difficult for thermal management of high-performance servers, which strongly drives the cooling design and technology developments.

Forced air cooling is considered a cost-effective and relatively simple technology that has been playing a major role in thermal management of most electronic systems, including high-performance servers. However, constraints in cooling capability and packaging conditions have been pushing advantages of air cooling to

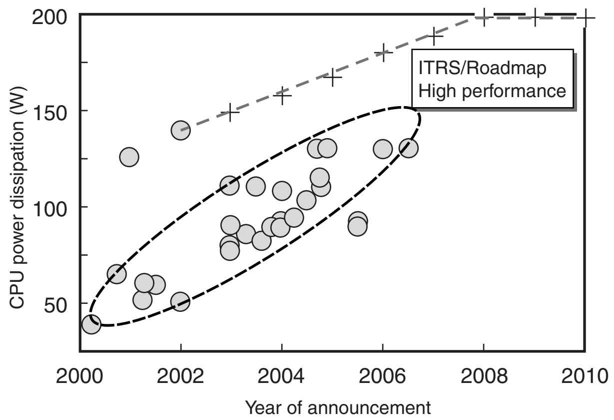

Figure 1
Roadmap of high-performance CPU power dissipation.

some limits, especially for servers having a high power dissipation and high-density packaging. It is becoming more challenging to extend the air cooling capability and increase the cooling efficiency of high-performance servers, and as a result an integrated approach from the chip, package, to overall system level is required.[2),3)]

To satisfy worldwide needs and expectations, Fujitsu is providing leading-edge high-performance servers such as the GS21, PRIMEPOWER, PRIMEQUEST, and PRIMERGY series. Advanced thermal solutions and state-of-the-art cooling technologies are widely used in these servers to meet the requirements of high performance, high reliability, and environmental compatibility.[4)-6)]

This paper gives an overview of thermal management for the latest high-end UNIX server PRIMEPOWER 2500 (hereafter called the PW2500), which was released in 2005. Perspectives are given on the cooling solution and design strategy at the server cabinet, system board, and CPU package levels. Furthermore, thermal challenges in dealing with a high-density and asymmetric power dissipation are also investigated. Especially, we focus on heat transportation and removal at the CPU package level, including the development of a new metallic thermal interface material (TIM), effects of thermal interface and heat spreading materials on cooling performance and thermal design, and investigations for extending heatsink cooling capabilities by applying high thermal conductivity devices such as heat pipes and vapor chambers.

2. Thermal management of Fujitsu's PW2500

In its maximum configuration, the PW2500 contains 128 high-performance SPARC64 V CPU processors and 512 GB of main memory on 16 system boards. The CPU processors use 90 nm logic and advanced copper interconnect technologies and operate at up to 2.08 GHz.

The server cabinet is 107 cm wide, 179 cm deep, and 180 cm high and weighs 1550 kg (**Figure 2**). It has a maximum power dissipation of about 40 kW, and forced air cooling is applied throughout the system. Thermal management can be generally simplified at the server cabinet, system board, and CPU package levels, while it is implemented precisely at inter-level interfaces.

2.1 Cooling at the server cabinet level

The system configuration of the PW2500 is approximately symmetrical and consists of the system board unit, power supply unit, and cooling fan unit (**Figure 3**). The system board unit contains 16 system boards and two Input/Output (I/O) boards installed vertically on two back-panel boards. The two back-panel boards are interconnected by a crossbar unit consisting of six crossbar boards installed horizontally.

The cooling fan unit, which contains forty-eight 200 mm-diameter fans, is installed between the system board unit and the power supply unit. It provides forced air cooling for the system board and power supply units, with air flowing in the server cabinet vertically from bottom to top. In addition, two fan trays containing six 140 mm-diameter fans are installed on one side of the crossbar unit to cool the crossbar boards with a horizontal airflow. The system and crossbar fan units both have full-redundancy features. The

Figure 2
Fujitsu's PW2500.

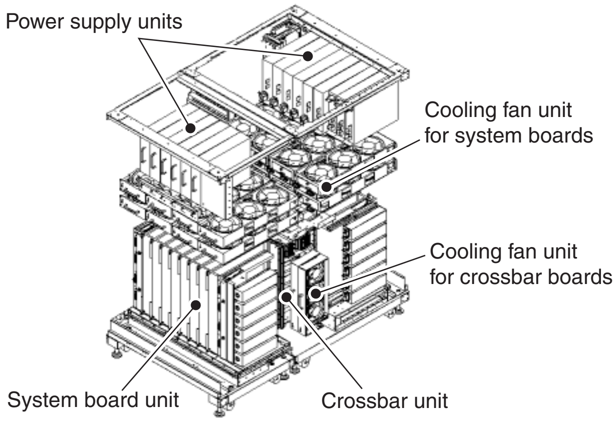

Figure 3
System configuration of PW2500.

server can operate normally at a 35°C ambient temperature at an altitude of 1500 m.

The system board unit is 79 cm wide, 124 cm deep, and 61 cm high (**Figure 4**). The system boards are mounted at a pitch of about 8 cm.

2.2 Cooling at the system board level

The system board is 58 cm wide and 47 cm long (**Figure 5**). It contains 8 CPU processors, 32 Dual In-line Memory Modules (DIMMs), 15 system controller processors, and associated DC-DC converters. The maximum power dissipation of the system board is about 1.6 kW.

The CPU packages on the system board are offset from each other along the airflow. This arrangement enhances the airflow for the downstream CPU packages, while achieving a balanced distribution across the system board for effective cooling of the board's components. Air ducts and baffles are used to direct and partition the airflow to improve the cooling efficiency and uniformity across the system board, as well as the overall system board unit. Evaluations have confirmed that the downstream CPU packages have about the same temperature characteristics as the upstream ones.

2.3 Cooling at the CPU package level

Figure 6 shows the CPU package and heatsink module. The CPU package consists of an integrated heat spreader (IHS) attached to the CPU chip with a high-performance TIM that thermally and mechanically couples the CPU chip to the IHS. The IHS effectively spreads the heat from the CPU chip across a wider area, thereby reducing the on-chip thermal gradient and improving the heatsink cooling efficiency. The heatsink module is mounted on the IHS with another TIM in between and dissipates the heat to the ambient environment.

The maximum power dissipation from the CPU processor is around 90 W, with an average power density of about 30 W/cm^2. However, due to an asymmetric power distribution, the highest power density on the local chip surface — called the hot-spot — was predicted to be about 90 W/cm^2. Thermal designs were conducted based on the proposed cooling capacity, airflow characteristics, packaging requirements, and asymmetric power distribution properties.

The heatsink was designed in accordance with required cooling performance and packaging constraints and optimized from associated thermal and airflow characteristics. We also investigated the relationships among the cooling performance, manufacturability, materials, cost, weight, and other factors. The heatsink consists of aluminum fins soldered to a copper base. The aluminum fins were used to reduce the cost and

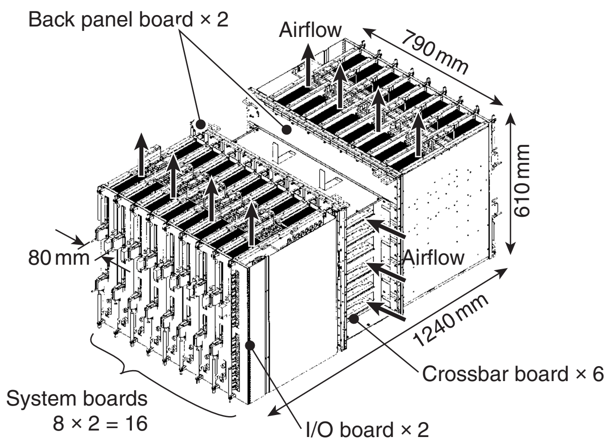

Figure 4
System board unit of PW2500.

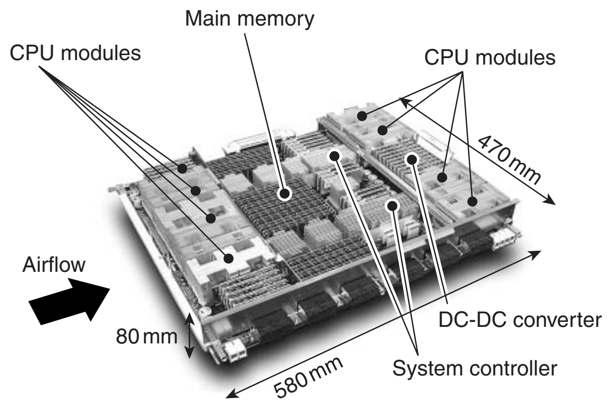

Figure 5
System board layout of PW2500.

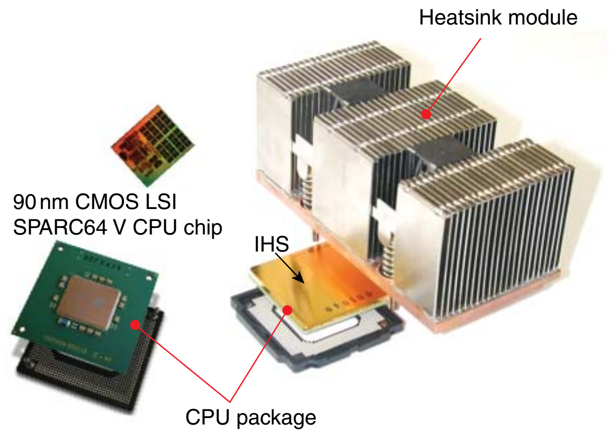

Figure 6
CPU package and heatsink module of PW2500.

weight, and the copper base was used to enhance heat spreading and therefore heat dissipation from the fins.

2.4 Simulations of thermal and airflow characteristics

A combination of extensive numerical simulations and experimental evaluations was applied throughout the thermal solution development to comprehensively investigate and understand the thermal and airflow behaviors related to the cooling and packaging requirements, scheme and component selections, system configuration and performance confirmations, and other details. Simulations were conducted from the CPU packages to the overall server cabinet, with boundary conditions applied to the inter-level interfaces. An example of one of these simulations is shown in **Figure 7**.

Furthermore, we performed automatic thermal modeling for system-level designs by using a Virtual Product Simulator (VPS)/simulation hub developed by Fujitsu. The VPS/simulation hub has a user-friendly interface and advanced features for converting 3D Computer Aided Design (3D-CAD) data directly to accurate numerical analysis models and automatically deleting unnecessary components.[7] **Figure 8** shows an example conversion by the VPS/simulation hub from a 3D-CAD system board model to a numerical analysis model. The 3D-CAD model contains more than 1500 components, and at least 1.5 million meshes would be needed for a traditional analysis model. However, by automatically deleting unnecessary components, the VPS/simulation hub produced a model with less than 1.0 million meshes, while achieving the same or possibly an even higher analytical accuracy.

3. Challenges in cooling technology developments

To cope with the continually increasing average and local power densities of high-performance microprocessors, improvements in

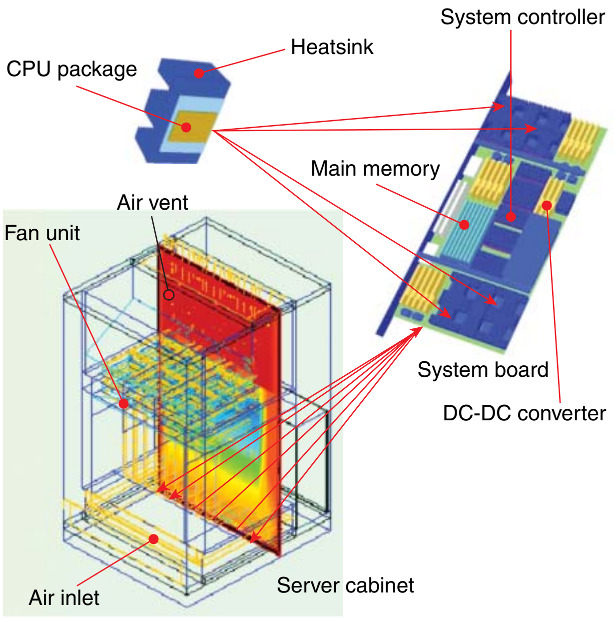

Figure 7
Thermal and flow simulations from CPU packages to server cabinet.

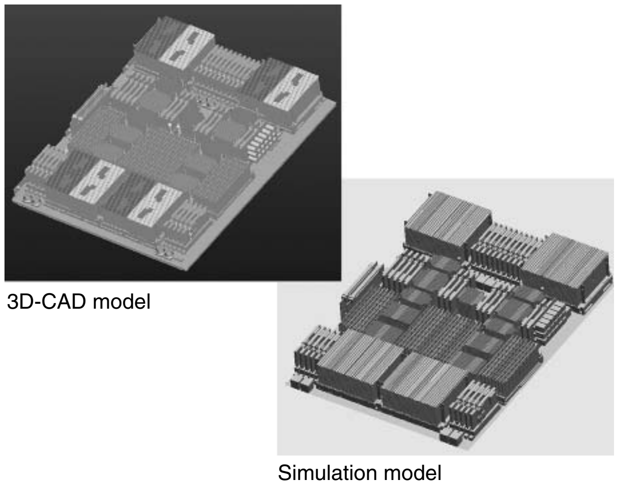

Figure 8
Example conversion of server system-board assembly.

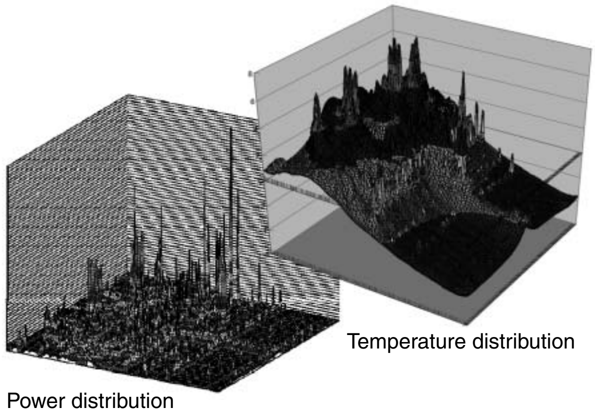

Figure 9
Power and temperature distributions at CPU junction.

cooling methodologies, techniques, and materials are required to achieve higher cooling efficiencies while meeting the packaging, cost, and environmental requirements.

3.1 High-density and asymmetric power dissipation

Figure 9 shows a simulation of the power and temperature distributions at the junction of an example CPU processor design.[8] The local hot-spots are at the CPU clock drivers, which comprise only a small portion of the chip area, and as the heat of these hot-spots moves towards the CPU chip's surface, it becomes less concentrated. The simulation shows that the hot-spots on the chip surface have about three to five times the average power density. These hot spots cause large temperature gradients across the chip that reduce the CPU performance and reliability. Cooling technologies of advanced thermal materials, structures, and devices as well as alternative novel techniques are required to deal with the high-density and asymmetric heat dissipation.

3.2 Advanced TIMs

TIMs are mostly used to facilitate thermal conduction from a chip to its heatsink via an IHS. Generally, they have two key functions in chip packages: to conduct heat from a chip and to absorb thermal stress resulting from material-stress mismatch between a chip, substrate, and IHS. The TIM between a chip and IHS is extremely important in package-level cooling technology developments for lowering the interface thermal resistance and matching thermal stresses.

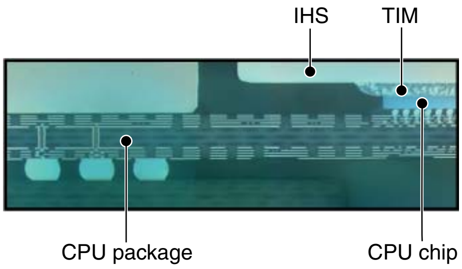

Figure 10
Structure of CPU package that uses In-Ag solder TIM.

Fujitsu has developed a new TIM made of an In-Ag composite solder for use between a chip and IHS[9] with the objective of achieving higher cooling performance while matching the thermal stresses and meeting environmental requirements. This TIM is an effective alternative to TIMs made from Sn-Pb, Sn-Ag, Su-Bi, In-Sn, and other metals.

The new In-Ag composition has a higher thermal conductivity than polymer-based TIMs and even most typical solders. Furthermore, it is relatively soft and has a low melting point, making it more tolerant of thermal stresses. Detailed simulations and optimizations were conducted to find the best bond line thickness (BLT) and soldering area for dispersing thermal stress. **Figure 10** shows the structure of a package that uses this In-Ag solder TIM.

In addition, the In-Ag composition has a wider temperature range between the liquidus and solidus than Sn-Pb, Sn-Ag, and certain other solder compositions, which makes it easier to eliminate air-voids inside the soldering interface by adjusting the TIM composition and reflow process.

3.3 Effects of heat-spreading material

An IHS in the form of a flat plate with good thermal conductivity is generally used to spread heat from a chip to a large-footprint heatsink and is an effective way of meeting the needs for sophisticated, high-density cooling options.

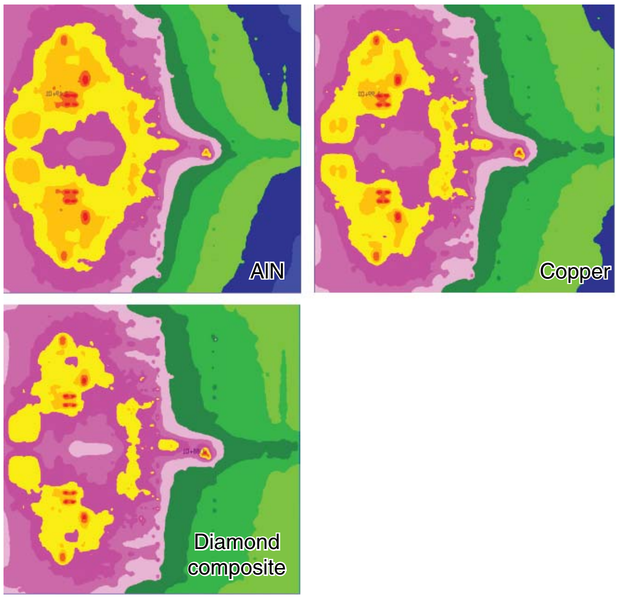

Figure 11
Effects of IHS on CPU junction temperature distribution.

We studied the effects of IHS thermal conductivity on CPU junction temperature distribution.[8] We found that, compared to 200 W/(m·K) of aluminum nitride (AlN) and 400 W/(m·K) of copper, a diamond composite IHS with a thermal conductivity of 600 W/(m·K) would result in a lower temperature gradient across the chip and lower temperature increments at hot-spots (**Figure 11**).

3.4 Improving heatsink cooling capability

Heatsinks are one of the most important cooling schemes, and much work has been done on designing and optimizing their structures. **Figure 12** shows the volume, weight, and thermal resistance of various heatsinks used in high-performance servers. It can be seen that the volume and weight of a heatsink for 150 W cooling, for example, would exceed 0.5 L and 1.0 kg. This indicates that to maintain sufficient air-cooling on high-performance devices, it is necessary to relax the packaging requirements or introduce

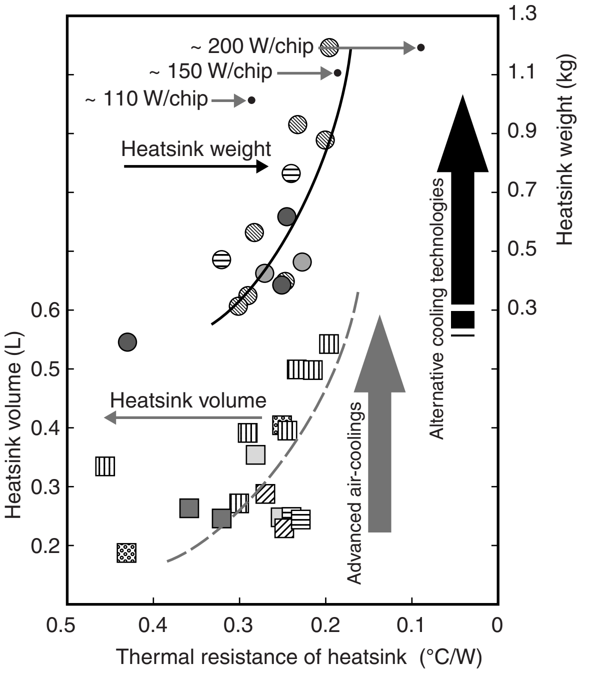

Figure 12
Characteristics of heatsinks for high-performance cooling.

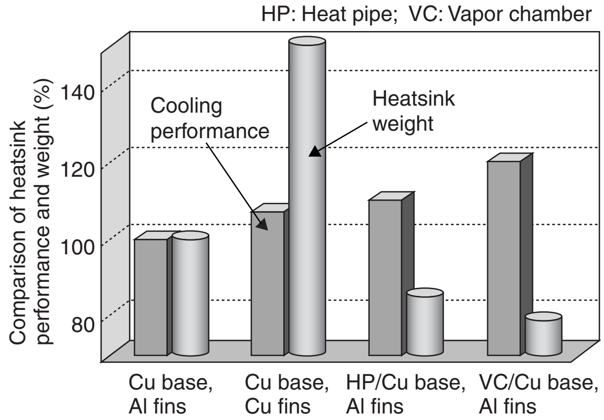

Figure 13
Performance and weight comparison of various heatsinks (Cu-base/Al-fins heatsink values set to 100%).

new technologies.

In addition to improving the fin structure and surface heat transfer coefficient, heatsink cooling capability can be further increased by improving the heat conduction inside the heatsink base, especially for large-footprint heatsinks. Some phase-change cooling devices, for example, heat pipes and vapor chambers (or planar heat pipes), are considered highly effective for enhancing heat spreading.[10),11)] These devices use an internal liquid-to-vapor phase-change mechanism to efficiently spread local high-density heat to an air-cooled fin structure. **Figure 13** compares the performance and weight of various heatsinks. The figure shows that, compared to a copper-base heatsink with aluminum fins, performance can be improved by 8% with a 50% weight increase using copper fins, by about 10% with a 15% weight reduction using heat-pipes, and by 20% with a 20% weight reduction using a vapor-chamber embedded heatsink.

4. Conclusion

In this paper, the thermal design and development methodology of Fujitsu's high-performance server PW2500 were described at the server cabinet, system board, and CPU package levels. We also discussed cooling technologies and challenges in dealing with high-performance microprocessors having a high-density and asymmetric power dissipation.

With the trend towards miniaturization, high performance, and high-density packaging, there will be continuing challenges in developing high-performance, high-efficiency thermal solutions. The industry faces a critical need to enhance conventional cooling technologies and develop more aggressive and cost-effective cooling technologies.

References

1) International Technology Roadmap for Semiconductors 2005 Edition Assembly and Packaging. *http://www.itrs.net/Links/2005ITRS/AP2005.pdf*
2) J. Wei: Challenges in Package-Level High Power Density Cooling. Proceedings of the 17th International Symposium on Transport Phenomena, Toyama, Japan, 2006.
3) I. Sauciuc et al.: Air-cooling Extension — Performance Limits for Processor Cooling Applications. Proceedings of 19th IEEE Semi — Therm Symposium, San Jose, USA, 2003, p.74-81.

4) J. Wei et al.: Thermal Management of Multiple MCMs with Low-Temperature Liquid Cooling. Proceedings of InterPACK'01, Hawaii, USA, 2001, p.1-10.
5) O. Hamada: High-Reliability Technology of Mission-Critical IA server PRIMEQUEST. *FUJITSU Sci. Tech. J.*, **41**, 3, p.284-290 (2005).
6) J. Wei et al.: Thermal Management of Fujitsu High-End UNIX Servers. Proceedings of InterPACK'05, San Francisco, USA, 2005, p.1-7.
7) K. Aoki et al.: Automatic Thermal Modeling for System Level Design of Electronic Equipment. Proceedings of InterPACK'03, Hawaii, USA, 2003, p.1-6.
8) J. Wei et al.: Effects of Asymmetric Power Distributions on Thermal Management of High Performance LSI Processors. Proceedings of IFHT2004, Kyoto, Japan, 2004, p.37-39.
9) M. Koide et al.: High-Performance Flip-Chip BGA Technology Based on Thin-Core and Coreless Package Substrates. Proceedings of 56th ECTC, San Diego, USA, 2006, p.1869-1873.
10) J. Wei et al.: Measurement of Vapor Chamber Performance. Proceedings of 19th IEEE Semi-Therm Symposium, San Jose, USA, 2003, p.191-194.
11) J. Y. Chang et al.: Thermal Performance of Vapor Chambers Under Hot Spot Heating Conditions. Proceedings of InterPACK'05, San Francisco, USA, 2005, p.1-7.

Jie Wei, *Fujitsu Ltd.*
Mr. Wei received the Dr. Eng. degree in Mechanical Aerospace Engineering from Tokyo Institute of Technology, Tokyo, Japan in 1997. He joined Fujitsu Limited in 1997, where he has been engaged in thermal designs and development of cooling technologies for Fujitsu high-end servers, as well as mid-range and personal computers.