Power Reduction Techniques Used in SPARC64 VIIIfx Processor for Fujitsu's Next-Generation Supercomputer

● Yukihito Kawabe   ● Ryuji Kan   ● Hideo Yamashita   ● Hiroshi Okano

The SPARC64 VIIIfx processor is part of the SPARC64 series developed by Fujitsu. Slated for use in Fujitsu's next-generation supercomputer, it features eight cores and an operating frequency of 2 GHz with arithmetic computational extensions for high performance computing such as single instruction multiple data (SIMD) instructions and register extensions as enhancements over the previous-generation SPARC64 VII processor used for UNIX servers. While achieving a high peak performance of 128 GFLOPS, this chip achieves a low power consumption of 58 W by reducing leakage power through the adoption of water cooling and reducing wasted power on the basis of power-analysis values obtained by a gate-level power analysis flow. The performance-per-watt factor achieved here is six times that of the SPARC64 VII processor. This paper introduces a power analysis flow for developing power-saving measures in the SPARC64 VIIIfx processor and presents specific techniques for reducing both leakage and dynamic power. It also describes the final power-analysis results for a core and the entire chip and presents the results of power measurements.

1. Introduction

Computer simulation has come to be used in manufacturing, meteorology, astronomy, and many other fields, increasing the importance of ultra-high-performance computers commonly referred to as "supercomputers." The performance of supercomputers has been increasing year by year, and supercomputers of the petaflops performance class have been around for several years. However, the power required for supporting calculations at this level has been increasing as well. As can be seen from the GREEN500 list,[1] the power consumption of supercomputer systems has reached the megawatt level. Thus, for petascale and future exascale computer systems, it is essential that the cost of power-supply equipment, electricity expenses, and other power-related costs be reduced and that power consumption itself be reduced to counter heat buildup caused by high-density device mounting. This means that it is becoming increasingly important that the power consumption of processors—the main source of power consumption in a supercomputer—be reduced simultaneously with gains in performance.

In this paper, we describe power-analysis and power-reduction techniques for the SPARC64 VIIIfx[2] processor that we developed for use in Fujitsu's next-generation supercomputer.

2. Overview of SPARC64 VIIIfx

The chip layout of the SPARC64 VIIIfx processor is shown in **Figure 1**, and the chip specifications are listed in **Table 1**. This chip features a variety of enhancements with respect to the SPARC64 VII processor for UNIX servers, which is the previous-generation chip in the SPARC64-series[3] design base. For example, it increases the number of cores from four to eight
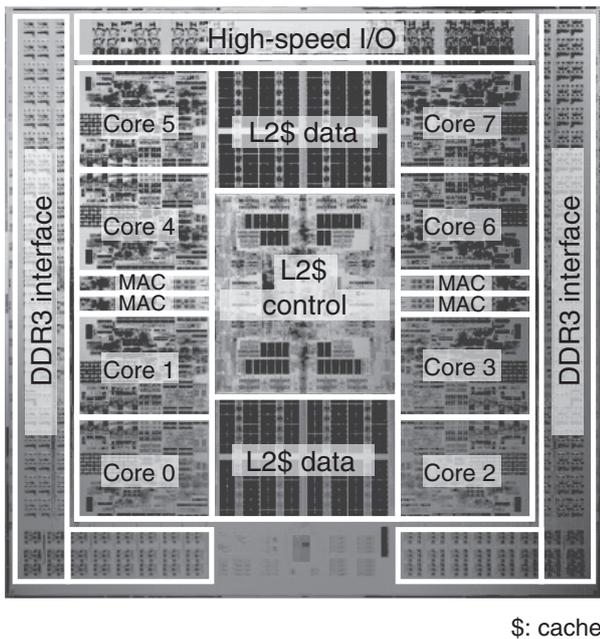
$: cache

Figure 1
SPARC64 VIIIfx chip layout.

Table 1
SPARC64 VIIIfx chip specifications.

Architecture	SPARC-V9/HPC-ACE
	  SIMD instructions
	  256 FPRs

	8 cores
	32 KB L1I$, 32 KB L1D$
	6 MB common L2$
	Clock frequency 2 GHz
Peak performance	Computational performance: 128 GFLOPS
	Memory throughput: 64 GB/s
Other	FSL 45-nm CMOS
	22.7 × 22.6 mm
	760 M transistors
	1271 signal pins

FSL: Fujitsu Semiconductor LTD.
HPC-ACE: High Performance Computing- Arithmetic
             Computational Extensions

and adds arithmetic computational extensions for high performance computing, including single instruction multiple data (SIMD) instructions, instructions for scientific computing, and a large register set featuring 256 floating point registers (FPRs) versus 32 in the previous chip.[2] Clock frequency is 2 GHz, and peak performance is 128 GFLOPS.

The above enhancements had to be achieved under the design constraint of a chip-wide power consumption of 58 W (typical process corner at 30°C) during peak operation. This was a tough power-consumption target as it represented a performance-per-watt factor six times that of the SPARC64 VII processor.

3. Power analysis flow

Power analysis in the design of the previous SPARC64 VII chip made use of average operating rate in units of function modules. This approach made it possible to determine power consumption for each function module but not for individual cells or macros. For the SPARC64 VIIIfx chip, we introduced a power analysis flow at the

gate level with the aim of achieving a drastic reduction in power consumption (**Figure 2**).[4] The power consumption information obtained from the results of this power analysis in units of cells and macros was used in the design process to isolate those locations where measures for reducing power consumption were needed and to then implement those measures and check their effects.

The operating rate information used in this power analysis was obtained in units of clock cycles using cycle-based logic emulation on a logic emulator. The operating information itself consisted of operating rates for all nets and cell pins inside the chip. Additionally, to keep the size of the operating-information file generated in this analysis and the time required for conducting the power analysis within realistic values, we selected—from the entire logic emulation period—several intervals of about 3000 cycles for which the usage rate of execution units was stable at the peak value and focused our analysis on those intervals to obtain detailed operating information.

As part of this power analysis flow, we created a power library, which is a file that defines information related to the power
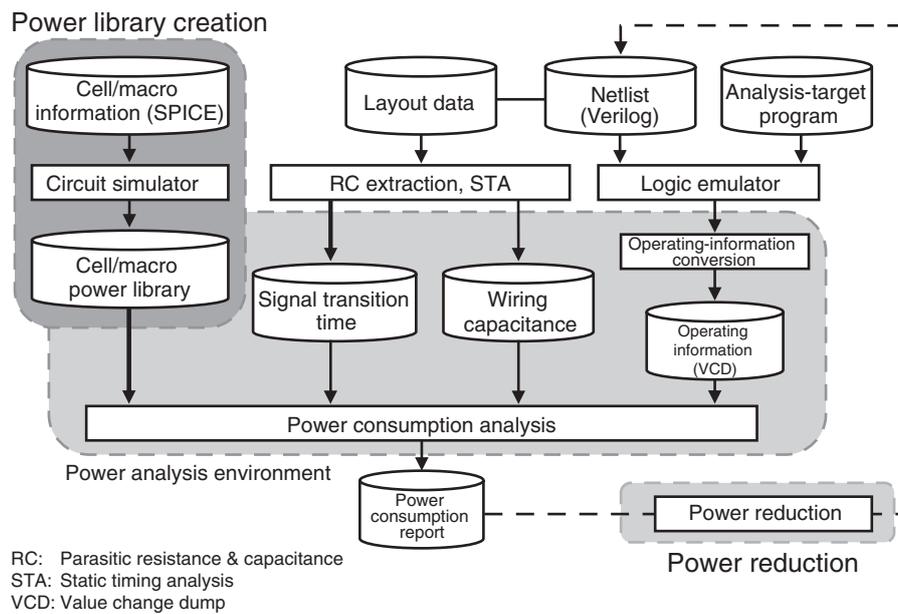
Figure 2
Power analysis flow.

consumption of individual circuit elements, i.e., cells and macros. Power analysis at the gate level can calculate the power consumption of cells and macros on the basis of the information defined in the power library. We created a power library based on the results of the circuit simulation of standard cells, I/O cells, and RAM macros from among the cells and macros that can be used for power analysis of this chip. However, as this chip was developed by gate-level design as opposed to register-transfer level (RTL) design, the number of custom macros for increasing speed and meeting other objectives was exceptionally high, which meant that it was unrealistic to create a highly detailed power library for all of those macros. We therefore created a simple library for custom macros that is based on power estimates from physical quantities. Specifically, we used leakage power and dynamic power estimated from total transistor width for each type of transistor. For dynamic power, we created a library by allocating it to macro pins as switching power, and, for macros of sequential circuits, we treated a fixed percentage of total transistor width as a latch and allocated that latch's worth

of power to the clock pins of that macro as power consumed during operation.

As for register file macros, we began to create a power library just as we did for custom macros, but as the power consumption per unit macro obtained as a result of power analysis was extremely large compared to that of custom macros, we decided to create a highly precise power library based on power values obtained from circuit simulation in the same manner as RAM macros.

As this chip is a product of gate-level design, there are many cells with relatively complicated structures having multistage-CMOS configurations such as multi-input-signal AND, multi-input-signal OR, and multi-input-signal AND-OR. We consider such a cell to be divided into power consumed by internal circuits connected to the input pins and that consumed by internal circuits connected to the output pin. We define the power consumed while a circuit on the input side is operating in terms of the input pins connected to that circuit and define the power consumed by a circuit on the output side in terms of the output pin connected to that

circuit and create a power library accordingly. In this way, power will not be underestimated even in the case of cell operation with no output transition.

During the design of this chip, we performed power analysis repeatedly across the entire chip. Moreover, using the results obtained after each analysis, we checked the degree to which the power target set for each functional block was being achieved and revised measures accordingly. In the design of a functional block, we reduced power consumption progressively on the order of µW to mW toward the established target value by making power-saving improvements from the micro-architecture level to the cell/macro level.

4. Power-reduction techniques

Power consumption in large scale integration (LSI) devices can be broadly divided into two types: leakage power caused by leakage current that continues to flow even when the circuit in question is not operating and dynamic power consumed by the circuit while operating. This section introduces techniques for reducing both leakage power and dynamic power.

4.1 Leakage power

For this chip, we decided on a water-cooling method to cool the chip with the aim of reducing leakage power. With this method, the operating temperature of the processor can be held to 30°C or less. On comparing power consumption between operating (junction) temperatures of 85 and 30°C, a power-reduction effect of 7 W across the entire chip can be seen (under the typical process corner).

Furthermore, as another measure for reducing leakage power, we used mostly transistors having a channel 15% longer than that of standard transistors. Such long-channel transistors have leakage power a fraction that of standard transistors, and for this reason we used them in 91.5% of all logic cells in the chip. As for the remaining logic cells, we applied transistors

of standard channel length to 0.1% of them used as high-speed cells and high-Vth transistors with low leakage power to 8.4% of them used as hold-countermeasure buffers and for other purposes.

In this way, we have been able to achieve a significant reduction in chip leakage power under conditions that minimize the impact on operating frequency. This accomplishment enabled us to focus our efforts on measures for reducing dynamic power.

To evaluate dynamic power in the SPARC64 VIIIfx chip, a maximum-power evaluation program is needed, and to this end, we used a program that emulates the main arithmetic computational sections of the LINPACK numerical-calculation benchmark program used for ranking the TOP500[5] supercomputer sites. During the design process, we implemented measures to reduce power consumption while running this program, which executes double-precision, floating-point multiply-and-add (FMA) instructions for nearly 100% of the time in all 32 FMA execution units on the chip. In the following subsection, we introduce several techniques for reducing unnecessary dynamic power consumed by the chip.

4.2 Dynamic power

For this chip, we incorporated a new single-loop entry buffer in the branch-prediction circuit. The information read from the branch-prediction table is stored in this buffer. The next branch address is read from this small-scale buffer instead of the branch-prediction table if "taken" branch instructions with the same address have been frequently executed—as in the case of a computational loop that is continuously executed. This scheme cuts down on RAM access to the branch-prediction table and consequently reduces power consumption when executing the maximum-power evaluation program by 890 mW across the entire chip.

A two-cycle-access pipelined RAM is used for the L1 cache. In previous versions of this RAM,

the clock could be suspended only when RAM had not been accessed for two consecutive clock cycles. However, by changing the specifications for the RAM macro, we added a clock-suspension function for each pipeline stage, enabling unnecessary clock operation to be suspended for any cycle not using RAM. As a result, power is now reduced for every 5 of 16 cycles, making for a power reduction of 540 mW across the entire chip.

We also eliminated unnecessary reading of FPRs to reduce power consumption. In the original circuit for supplying data to an execution unit to perform calculations, the FPR read address would change and the FPR unnecessarily operated even if computational data were supplied to the execution unit from other than a FPR, that is, even if the FPR was bypassed. In the new circuit, the clock for the latch governing FPR address supply is suppressed at the time of bypass, thereby suppressing unnecessary FPR operation. This results in a power-consumption reduction of 1.4 W.

In addition to the above, we reduced unnecessary operations by using clock gating and other techniques across the entire chip and used a cell-size adjustment tool to switch from large to small cells within the range that does not exceed delay constraints.

5. Results of power analysis

The results of initial and final power analyses for each functional unit in a core during design using the power analysis flow described above are shown in **Figure 3**. The power consumption for a core was reduced from about 6 W in the initial design stage to a final value of 4.4 W. Specifically, the power consumed by register files and RAM macros in each circuit element was reduced by using the power-reduction techniques presented in the previous section, and clock power and power used by logic sections were reduced by reducing unnecessary power consumption throughout the chip by using clock gating and other techniques.

The results of the final power analysis for the entire chip were within the design target of 58 W. Leakage power turned out to be about 10% of the total power consumed by the chip.
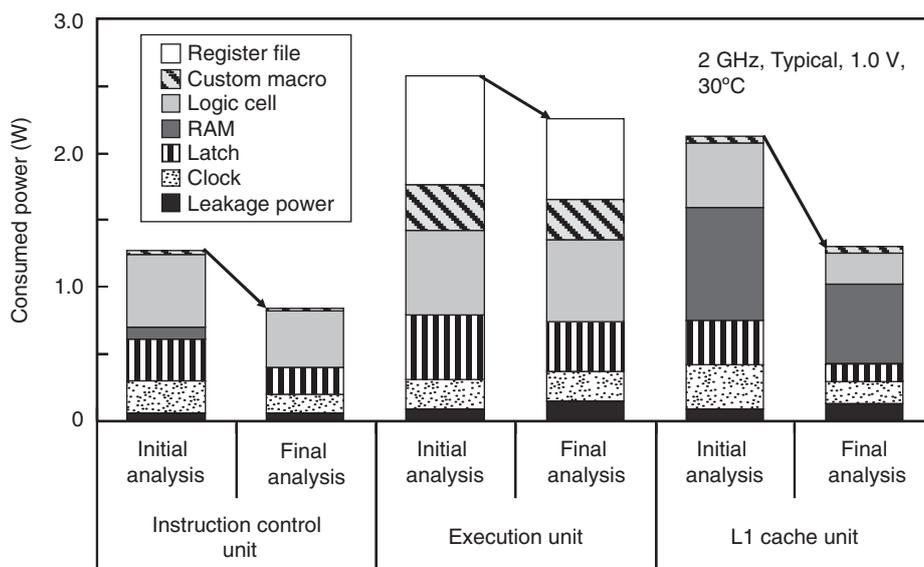
We note here that this power analysis was

Figure 3
Power analysis results for processor core.

performed under the typical process corner and that the results obtained do not necessarily apply to all chips given the variation in characteristics in manufactured chips. However, by applying the adaptive supply voltage (ASV)[6] method for adjusting the voltage applied to each chip, the average power consumption per chip across an entire supercomputer system can be made to meet the 58 W target.

6. Comparison with power measurements

Figure 4 compares the measured power consumption when various test programs were executed on a sample chip on a test board with the results of power analysis for the power-supply domain (VDD1 = 1.0 V) that feeds logic inside the chip as one type of power consumed by the processor. The values obtained by power

measurement were corrected for temperature and process so as to match the measurement conditions with the power-analysis conditions. As shown in the figure, the power consumed in the VDD1 power-supply domain by the maximum-power evaluation program was about 48 W, accounting for most of the maximum power (58 W) consumed by the chip. The bars on the far right shows the results for the maximum-power evaluation program when some of the power-suppression functions, such as clock gating, were disabled by making appropriate chip settings.

These results clearly demonstrate that the results of power analysis agree very well with the results of power measurement.

7. Conclusion

In this paper, we introduced power-analysis and power-reduction techniques for the

PVT: Process, voltage, temperature
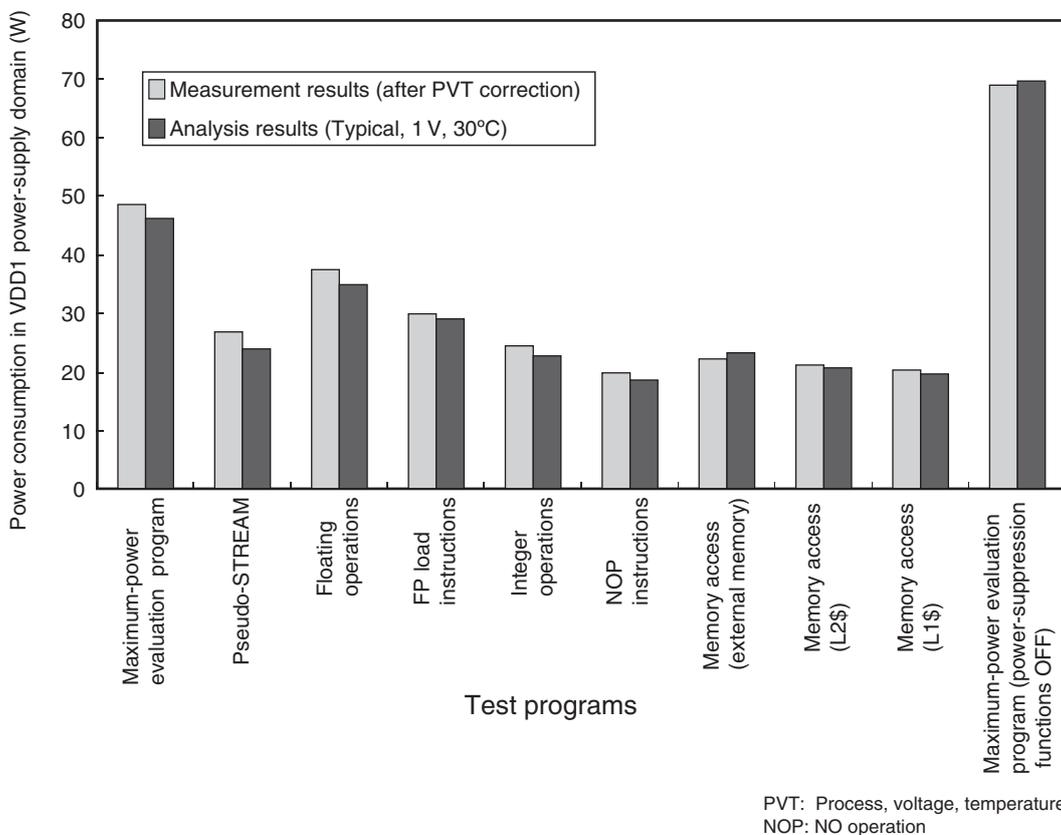NOP: NO operation

Figure 4
Comparison between power measurement and power analysis results.

SPARC64 VIIIfx processor to be used in Fujitsu's next-generation supercomputer. We showed that leakage power could be reduced by water cooling and the use of long-channel transistors and that other reductions in power consumption could be achieved by high-accuracy power analysis. As a result, the SPARC64 VIIIfx processor achieves low chip power consumption of 58 W when running a maximum power evaluation program while maintaining a high peak performance of 128 GFLOPS.

References

1) THE GREEN500: Ranking the World's Most ENERGY-EFFICIENT SUPERCOMPUTERS.
   http://www.green500.org/
2) T. Maruyama: SPARC64TM VIIIfx: Fujitsu's New Generation Octo Core Processor for PETA Scale Computing. HotChips21, August 25, 2009.
3) T. Maruyama et al.: Past, Present, and Future of SPARC64 Processors. *Fujitsu Sci. Tech. J.*, Vol. 47, No. 2, pp. 130–135 (2011).
4) H. Okano et al.: Fine Grained Power Analysis and Low-Power Techniques of a 128GFLOPS/58W SPARC64™ VIIIfx Processor for Peta-scale Computing. Symposium on VLSI Circuits, June 18, 2010.
5) TOP500 SUPERCOMPUTER SITES.
   http://www.top500.org/
6) H. Okano et al.: Supply Voltage Adjustment Technique for Low Power Consumption and its Application to SOCs with Multiple Threshold Voltage CMOS. Digest of Technical Papers. 2006 Symposium on VLSI Circuits, pp. 208–209 (2006).

Yukihito Kawabe
Fujitsu Laboratories Ltd.
Mr. Kawabe is engaged in the research of power-consumption estimation in the early stages of LSI design.

Hideo Yamashita
Fujitsu Ltd.
Mr. Yamashita is engaged in the development of processors.

Ryuji Kan
Fujitsu Ltd.
Mr. Kan is engaged in the design of CPU execution units.

Hiroshi Okano
Fujitsu Laboratories Ltd.
Mr. Okano is engaged in the research and development of low power consumption technology for integrated circuits.