SPARC64™ X: Fujitsu’s New Generation 16 Core Processor for the next generation UNIX servers

August 29, 2012
Takumi Maruyama
Processor Development Division
Enterprise Server Business Unit
Fujitsu Limited
Agenda

◆ Fujitsu Processor Development History

◆ SPARC64™ X
  ▪ Design concept
  ▪ SWoC (Software on Chip)
  ▪ Processor chip overview
  ▪ u-Architecture
  ▪ Performance

◆ Summary
SPARC64™ X Design Concept

◆ Combine UNIX and HPC FJ processor features to realize an extremely high throughput UNIX processor.
  - SPARC64 VII/VII+ (UNIX processor) feature
    - High CPU frequency (up-to 3GHz)
    - Multicore/Multithread
    - Scalability: up-to 64 sockets
  - SPARC64 VIIIfx (HPC processor) feature
    - HPC-ACE: Innovative ISA extensions to SPARC-V9
    - High Memory B/W: peak 64GB/s, Embedded Memory Controller

◆ Add new features vital to current and future UNIX servers
  - Virtual Machine Architecture
  - Software On Chip
  - Embedded IOC (PCI-GEN3 controller)
  - Direct CPU-CPU interconnect
Software on Chip 1/2

◆ HW for SW
Accelerates specific software function with HW

◆ The targets
- Decimal operation (IEEE754 decimal and NUMBER)
- Cypher operation (AES/DES)
- Database acceleration

◆ HW implementation
- The HW engines for SWoC are implemented in FPU
  - To fully utilize 128 FP registers & software pipelining
- Implemented as instructions rather than dedicated co-processor to maximize flexibility of SW.
- Avoid complication due to “CISC” type instructions
  - Various “RISC” type instructions are newly defined, instead.
  - 18 insts. for Decimal, and 10 insts. for Cypher operation
Decimal Instructions

- **Supported data type**
  - IEEE754 DPD (Densely Packed Decimal)
    - 8B fixed length
  - NUMBER
    - Variable length (max 21Byte)

- **Instructions**
  - Both DPD/NUMBER instructions are defined as 8B operation (add/sub/mul/div/cmp) on FP registers
    - To maximize performance with reasonable HW cost
    - When the data length is > 8byte, multiple such instructions will be used.
  - An instruction for special byte-shift on FP registers is newly added to support unaligned NUMBER
SPARC64™ X Chip Overview

- **Architecture Features**
  - 16 cores x 2 threads
  - SWoC (Software on Chip)
  - Shared 24 MB L2$
  - Embedded Memory and IO Controller

- **28nm CMOS**
  - 23.5mm x 25.0mm
  - 2,950M transistors
  - 1,500 signal pins
  - 3GHz

- **Performance (peak)**
  - 288GIPS/382GFlops
  - 102GB/s memory throughput
### SPARC64™ X Core spec

<table>
<thead>
<tr>
<th>Instruction Set Architecture</th>
<th>SPARC-V9/JPS HPC-ACE VM SWoC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Branch Prediction</td>
<td>4K BRHIS 16K PHT</td>
</tr>
<tr>
<td>Integer Execution Units</td>
<td>156 GPR x 2 + 64 GUB</td>
</tr>
<tr>
<td></td>
<td>ALU/SHIFT x2</td>
</tr>
<tr>
<td></td>
<td>ALU/AGEN x2</td>
</tr>
<tr>
<td></td>
<td>MULT/DIVIDE x1</td>
</tr>
<tr>
<td>FP Execution Units</td>
<td>128 FPR x 2 + 64 FUB</td>
</tr>
<tr>
<td></td>
<td>FMA x4, FDIV x2</td>
</tr>
<tr>
<td></td>
<td>IMA/Logic x4</td>
</tr>
<tr>
<td></td>
<td>Decimal x1 / Cypher x2</td>
</tr>
<tr>
<td>L1$</td>
<td>L1I$ 64KB/4way</td>
</tr>
<tr>
<td></td>
<td>L1D$ 64KB/4way</td>
</tr>
</tbody>
</table>
u-Architecture enhancements from SPARC64™ VII+

CPU Core
- Deeper pipeline to increase Frequency
- Better Branch Prediction Scheme
- Various Queue-size and #Floating point register increase
- Richer execution Units, including
  - 2EX + 2EAG → 2EX + 2EX/EAG
  - 2FMA → 4FMA to support 2way-SIMD
  - SWoC engine (Decimal and Cypher)
- More aggressive O-O-O-O execution of load and store
- Multi-banked 2port L1-Cache

System On Chip
- #core and L2$ size (4core/12MB→16core/24MB)
- Memory Controller, IO Controller, and CPU-CPU I/F are all embedded to increase performance and reduce cost.
SPARC64™ VII/VII+ Pipeline

Fetch (4 stages)  Issue (2 stages)  
<table>
<thead>
<tr>
<th>Fetch Port 16Entry</th>
<th>PC x2</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSE 64Entry</td>
<td></td>
</tr>
<tr>
<td>4-stage Process</td>
<td></td>
</tr>
<tr>
<td>2-stage Process</td>
<td></td>
</tr>
<tr>
<td>4-core</td>
<td></td>
</tr>
<tr>
<td>L1 I$ 64KB 2Way</td>
<td></td>
</tr>
<tr>
<td>L1 D$ 64KB 2Way</td>
<td></td>
</tr>
<tr>
<td>L2$ 6MB/12MB 12Way</td>
<td></td>
</tr>
<tr>
<td>System Bus Interface</td>
<td></td>
</tr>
</tbody>
</table>

Branch Target Address 8Kentry  

Decode & Issue  

RSF 8x2Entry  
RSE 8x2Entry  
L1 IS$ 64KB 2Way  

GPR 156Registers x2  
GUB 32Registers  
FPR 64Registers x2  
FUB 48Registers  

EAGA  
EAGB  
EXA  
EXB  
FLA  
FLB  

RSBR 10Entry  
RSF 8x2Entry  
RSE 8x2Entry  

Dispatch (4 stages)  

Reg.-Read (4 stages)  

Execute (L1$: 3 stages)  

Memory (L1$: 3 stages)  

Commit (2 stages)
**Execution units enhancements (Ex.)**

**Integer Execution Unit**
- 2EX + 2EAG → 2EX + 2EX/EAG
- 2 → 4W GPR
- → 4 integer instructions can be executed per cycle (sustained)

**Load Store Unit**
- Aggressive load/store O-O-O-O execution:
  - Execute load without waiting for preceding store address calculation.
- Multi-banked 2port L1-cache to execute 2 load or 1 load+1 store in parallel
- Doubled L1$ bus size
- Doubled L1$ associativity (2→4way)
- → Increase L1-cache throughput and hit-rate
SPARC64™ X interconnects

SPARC64™ VII/VII+ interconnects
(SPARC Enterprise M8000)

- 4 CPU require 8 additional LSIs to be connected with DIMM
- DIMM i/f: 4.35GB/s (STREAMtriad)

SPARC64™ X interconnects

- No additional LSIs to be connected with DIMM
- DIMM i/f: 65.6GB/s (STREAMtriad)
- CPU i/f: 14.5GB/s x 5ports (peak)
  - 3 ports: glueless 4way CPU interconnect
  - 2 ports: > 4way CPU
High Speed Transceivers (SerDes)

CPU-CPU glue-less communication links
- 14.5Gb/s x 8 lanes bi-directional serial interface, 5 ports
- Embedded equalizer circuit enables long distance signal transmission
- Embedded adaptive control logic optimizes equalizer parameters automatically depending on the various system configurations

PCI Express ports
- 8Gb/s x 8 lanes (Gen 3), 2 ports

Built-in SerDes provides peak 88.5GB/s x2 (up/down) total throughput
Reliability, Availability, Serviceability

<table>
<thead>
<tr>
<th>Units</th>
<th>Error detection and correction scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cache (Tag)</td>
<td>ECC Duplicate &amp; Parity</td>
</tr>
<tr>
<td>Cache (Data)</td>
<td>ECC Parity</td>
</tr>
<tr>
<td>Register</td>
<td>ECC (INT/FP) Parity(Others)</td>
</tr>
<tr>
<td>ALU</td>
<td>Parity/Residue</td>
</tr>
<tr>
<td>Cache dynamic degradation</td>
<td>Yes</td>
</tr>
<tr>
<td>HW Instruction Retry</td>
<td>Yes</td>
</tr>
<tr>
<td>History</td>
<td>Yes</td>
</tr>
</tbody>
</table>

◆ New RAS features from SPARC64™ VII/VII+
  - Floating-point registers are ECC protected
  - #checkers increased to ~53,000 to identify a failure point more precisely

→ Guarantees Data Integrity
Hardware Instruction Retry

1. Error
2. Flush
3. Single step execution
4. Update of SW visible resources
5. Back to normal execution after the re-executed Instruction gets committed without an error.

◆ When an error is detected, Hardware re-execute the instruction automatically to remove the transient error by itself.
SPARC64™ X realizes 7x INT/FP/JVM throughput and 15x memory throughput of SPARC64™ VII+

- The INT/FP/JVM result is with un-tuned Compiler/JVM.

SWoC of SPARC64™ X results in max 98x throughput.

- The NUMBER score is for scalar. Expect to be much better for vector data.
4 integer execution units and write port increase of GPR (integer register) improves overall performance.

Memory latency reduction, Large L2$, branch prediction, and L1$ improvement also contribute to the high performance dramatically.
Summary

◆ SPARC64™ X is Fujitsu’s 10th SPARC processor which has been designed to be used for Fujitsu’s next generation UNIX server.

◆ SPARC64™ X integrates 16 cores + 24MB L2 cache with over 100GB/s (peak) memory B/W.

◆ SPARC64™ X keeps strong RAS features.

◆ SPARC64™ X chip is up and running in the lab.
◆ It has shown 7 times throughput of SPARC64™ VII+ w/o compiler tuning.
◆ SWoC is effective to accelerate specific software functions

◆ Fujitsu will continue to develop SPARC64™ series.
Abbreviations

• **SPARC64™ X**
  – IB: Instruction Buffer
  – RSA: Reservation Station for Address generation
  – RSE: Reservation Station for Execution
  – RSF: Reservation Station for Floating-point
  – RSBR: Reservation Station for Branch
  – GUB: General Update Buffer
  – FUB: Floating point Update Buffer
  – GPR: General Purpose Register
  – FPR: Floating Point Register
  – CSE: Commit Stack Entry