Gemini: A Power-efficient Chip Multi-Threaded (CMT) UltraSPARC® Processor

Sanjiv Kapil
Gemini Architect
Sun Microsystems, Inc.
Design Goals

- Designed for compute-dense, transaction oriented systems (webservers, appservers ....)
  - Small form factors (1U, 2U, Blades)
  - Low power, low cost
  - Large physical memory (64-bit addressing)
  - Enterprise-class RAS
- Focus on throughput performance rather than single thread performance
- Support for 1- to 4-way SMP with simultaneous execution of 2-8 threads
- Maintains binary compatibility with SPARC® V9 code
UltraSPARC Processor Roadmap

Data Facing

s-Series
UltraSPARC III 1X
Today
UltraSPARC IV 2X

i-Series
UltraSPARC IIi 0.5X
UltraSPARC IIIi 1X 1.3X

Network Facing
h-Series
Gemini 3X
Niagara 15X

2002 2003 2004 2005 2006
Features

- Dual 4-issue superscalar cores
  - Based on 64-bit SPARC V9 processor architecture
  - 2 integer ops, 2 FPU/graphics (VIS 1.0) ops/cycle

- Dual integrated 0.5MB L2 caches (one per core)
  - 4-way set associative
  - Way-locking per core

- Pin-compatible with UltraSPARC IIIi processor
  - JBus system bus interface

- Integrated memory interface to standard DRAM
  - Shared 128-bit DDR1 controller
  - Up to 16 GBs total using 2-4 DIMMS
Features (con.)

- Power management modes (E*Star)
  - Transition to either $F_{cpu}/2$ or $F_{cpu}/32$

- Robust RAS support
  - Parity checking on L1 caches, JBus
  - ECC on L2 caches, DRAM

- CMT features
  - Software-controlled core disable
  - Dynamic cache-aware core parking
  - Adheres to UltraSPARC CMT Architectural Specification
Gemini-based System

4-Way SMP/8-thread System

CPU 0

CPU 1

CPU 2

CPU 3

ODR 1

DDR 1

DDR 1

DDR 1

1.5V DTL

JBUS 0

JBUS 1

Repeater

128

128

133 MHz

4.2 GB/s

200 MHz

3.2 GB/s

PCI A B

PCI A B

32/64 bit

33/66 MHz
Technology

- Dual-core CMT design
- Transistor count: 80 million
- Die size: 206 sq. mm.
- Package: 959 pin ceramic uPGA
- Maximum power consumption: 32 Watts @ 1.2 GHz, 1.3V

Technology:
- TI's advanced 130 nm CMOS process,
  300 mm wafers
- 53 nm $L_{eff}$
- 7LM Cu, low-k dielectric (OSG)
- dual $V_t$
L2 Caches

- Unified 512KB L2 cache per core: 4-way set associative, 64B line size, Pseudo-random replacement
- 2-2 mode access with 4-cycle latency
- 8 ECC bits for 64b data with physical bit scrambling, parity-protected tags
- L2 datapath and control logic integrated with SRAMs

L2Cache Area = 36.39 mm²
Power = 2.07 Watts

L2Cache Pipeline Diagram
### Memory Controller

<table>
<thead>
<tr>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Shared controller for two cores</td>
</tr>
<tr>
<td>- DDR1 compliant</td>
</tr>
<tr>
<td>- Max 4 DIMMs</td>
</tr>
<tr>
<td>(single-/double-sided)</td>
</tr>
<tr>
<td>- Total memory 256MB to 16GB</td>
</tr>
<tr>
<td>- SSTL Input/Output</td>
</tr>
<tr>
<td>- Memory I/F:</td>
</tr>
<tr>
<td>128-bit data + 9 ECC check bits</td>
</tr>
<tr>
<td>- E*Star support</td>
</tr>
<tr>
<td>(½ or 1/32 of Fcpu)</td>
</tr>
<tr>
<td>- PTB to track 16 open pages</td>
</tr>
<tr>
<td>- 4.2 GB/s peak B/W</td>
</tr>
<tr>
<td>- Measured local access latency: 96ns</td>
</tr>
</tbody>
</table>

#### DDR Unit

<table>
<thead>
<tr>
<th>DDR IO @ 266 mhz</th>
</tr>
</thead>
<tbody>
<tr>
<td>128+EC</td>
</tr>
<tr>
<td>PT</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>DP</td>
</tr>
<tr>
<td>AS</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>DP</td>
</tr>
<tr>
<td>128</td>
</tr>
<tr>
<td>QUEUE</td>
</tr>
<tr>
<td>Ct</td>
</tr>
<tr>
<td>I</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ADDR/Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>JBus Unit</td>
</tr>
</tbody>
</table>

| R |
| F |
Memory Subsystem Details

<table>
<thead>
<tr>
<th>Specifications</th>
<th>Main Features</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Caches:</strong></td>
<td></td>
</tr>
<tr>
<td>I: 16KB 2-way associative PIPT with 32B line size</td>
<td>I: Parity protection on tags/data s/w recoverable</td>
</tr>
<tr>
<td>D: 16KB direct VIPT with 32B line size</td>
<td>D: Parity protection on data and tag</td>
</tr>
<tr>
<td>L2: 512 KB, 4-way set associative 64B line size</td>
<td>L2: Parity protection on tags, ECC on data</td>
</tr>
<tr>
<td></td>
<td>2-2 mode access</td>
</tr>
<tr>
<td></td>
<td>3-ways locking support</td>
</tr>
<tr>
<td></td>
<td>Flush to memory support</td>
</tr>
<tr>
<td><strong>Memory:</strong></td>
<td></td>
</tr>
<tr>
<td>DDR1: 2.5 v (2.625v tolerant) SSTL</td>
<td>Shared for dual cores</td>
</tr>
<tr>
<td>128-bit data + 9-bit ECC</td>
<td>PTB tracks 16 open pages</td>
</tr>
<tr>
<td>Peak memory B/W: 4.2GB/@133Mhz</td>
<td>E*Star modes capable</td>
</tr>
<tr>
<td>128/256/512/1G bit 2-4 Dimm support</td>
<td>Separate PLL for DDR/E*Star support</td>
</tr>
<tr>
<td>64-bit virtual address space</td>
<td>4-Way SMP can address 64GB space</td>
</tr>
<tr>
<td><strong>CMP:</strong></td>
<td></td>
</tr>
<tr>
<td>Core disable/parking/error reporting</td>
<td>Non-core/Core-specific error reporting</td>
</tr>
<tr>
<td>JTAG/RAMTEST access to CSR's</td>
<td>Full 16 GB addressing by each core</td>
</tr>
<tr>
<td></td>
<td>Cache-aware parking (No instr. Execution)</td>
</tr>
</tbody>
</table>
JBus Unit

Specifications

IO: JBus : 1.5 V DTL I/O's
16 byte shared address/data
120 – 200 MHz clock
Impedence control on DTL I/O's

JBUS controller:
- Two level arbitration of JBus for dual cores.
- Snoop capability for dual L2 cache
- E*Star transitions
- 128-bit datapaths to dual L2 cache
- CMP features support:
  - Core disable
  - Core parking
  - DMA 16 GB addressing support
- Error detection for address/data

![Diagram of JBus Unit](image-url)
Area Distribution

Core Area = 28.6 mm$^2$

CPU Area = 206 mm$^2$
Power Distribution

Core Power = 6.7 W

CPU Power = 29.3 W
UltraSPARC Design Evolution

UltraSPARC I – 1995
- 5.2M transistors
- 500 nm CMOS
- 4 LM Al
- 3.3V, 30W @ 167 MHz
- SPEC2K int/rate: 76/0.9*

UltraSPARC II – 1997
- 5.4M transistors
- 350 nm CMOS
- 5 LM Al
- 2.5V, 25W @ 250 MHz
- SPEC2K int/rate: 120/1.4*

Gemini – 2003
- 80M transistors
- 130 nm CMOS
- 7 LM Cu
- 1.3V, 32W @ 1200 MHz
- SPEC2K int/rate: 511/11
- 4x perf, 8x throughput

*calculated from measured SPEC95 scores
Performance

- Peak 3.2 GB/s system I/O bandwidth
- Peak 4.26 GB/s memory bandwidth
- Average memory latency for local DRAM 96ns

<table>
<thead>
<tr>
<th>CPUs</th>
<th>SPECjbb2k</th>
<th>/Watt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>18700</td>
<td>584</td>
</tr>
<tr>
<td>2</td>
<td>34500</td>
<td>539</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CPUs</th>
<th>SPECweb99</th>
<th>/Watt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2100</td>
<td>66</td>
</tr>
<tr>
<td>2</td>
<td>3400</td>
<td>53</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CPUs</th>
<th>Int_Rate</th>
<th>/Watt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
<td>0.34</td>
</tr>
<tr>
<td>2</td>
<td>16.2</td>
<td>0.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CPUs</th>
<th>FP_Rate</th>
<th>/Watt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.2</td>
<td>0.38</td>
</tr>
<tr>
<td>2</td>
<td>18.0</td>
<td>0.28</td>
</tr>
</tbody>
</table>

All scores @ 1.2 GHz with 16 GB memory, 2 active cores/processor

All numbers preliminary based on measured data
Benchmark: SPECint/fp_rate

Benchmark: SPECweb, SPECjbb

**SPECweb/ watt**

- GM 1P: 66
- GM 2P: 53
- HM 1P: 1.8
- HM 2P: 37
- Cel 1P: 52
- Cel 2P: 28
- Xn 1P: 3.06
- Xn 2P: 2.4

**SPECjbb/ watt**

- GM 1P: 584
- GM 2P: 539
- HM 1P: 1.8
- HM 2P: 365
- Cel 1P: 1.8
- Cel 2P: 294
- Xn 1P: 3.06
- Xn 2P: 2.4


18
Summary

- First generation CMT design for high-density, low-cost, network-facing systems
- Optimized for throughput rather than single-thread performance
- Maintains binary compatibility with existing SPARC code base
- Highly leveraged design methodology