3.2: A 3.6 GHz 16-Core SPARC SoC Processor in 28nm

ORACLE
Santa Clara, CA

Outline

- T5 Processor and S3 core overview
- Physical design
- Power management
- Power distribution and clocking
- Frequency Locked Loop and Clock MUX
- L3 Cache overview
- SerDes Wide Band Amplifier
- Conclusion
T5 Processor

- Next generation SPARC processor with 16 cores and 128 threads
- Achieves core frequency of 3.6GHz
- Architectural improvements in the SoC
- Significant increases in I/O bandwidth
- Enables 8-socket glueless systems
- Advanced power management including Dynamic Voltage and Frequency Scaling
T5 Block Diagram

- DDR3 - 1066 MHz
- Memory Control
- Coherence Unit
- L3$ B0 1MB, 16-way
- L3$ B1 1MB, 16-way
- L3$ B2 1MB, 16-way
- L3$ B3 1MB, 16-way
- L3$ B4 1MB, 16-way
- L3$ B5 1MB, 16-way
- L3$ B6 1MB, 16-way
- L3$ B7 1MB, 16-way
- 8 x 9 Crossbar
- 8 threads per Core

2 x 8 PCIe 3.0 @ 8 Gbps
16 Gbps each direction

12.8 Gbps Links
T5 Micrograph

- 16 SPARC S3 cores
- 128 threads/chip
- Shared 8MB L3 cache, 8-bank, 16-way
- 8x9 Crossbar
- 8 DDR3 Schedulers providing 80 GB/s BW
- 7 Coherency @12.8Gb/s
- 2x8 PCIe 3.0 @8GT/s
- Advanced power management with DVFS
S3 Core Overview

- Dual-Issue, Out-of-Order execution
- 8 threads per core
- 16-stage integer pipeline
- Extensive branch prediction
- Dynamic threading
- 64-entry ITLB and 128-entry DTLB
- 16KB 4-way L1 instruction and data cache
- 128KB 8-way unified private L2 cache
S3 Core Block Diagram

- **Crypto**
- **Register Files (IRF/FRF/WRF)**
  - ALU1
  - ALU0
  - Unified Pick Queue
  - Rename
  - Instruction Decode
  - Thread Select
  - L1 I-Cache 16KB 4-Way
  - MMU
  - L1 D-Cache 16KB 4-Way
  - TLU
  - Private 128KB 8-Way L2 Cache

© 2013 IEEE
Physical Design

- Clusters are composed flat to allow better routing and design optimization
- Families of swappable static gates used for performance versus power trade-offs
- Emphasis on design for manufacturability and performance in cell layout and routing
- Rigid context requirements for cell and block abutment to enhance predictability
• Layout dependent effects are extracted in context
• Abutment cells are created to model worst case max and min timing effects
• Context is guaranteed in composition using abutment cells and rigorous cell boundary rules
Swapping Cells to Save Power

- Gate bias and HVT cells are used to reduce power
- LVT cells are used to increase performance
- High performance flops are used in critical paths
- Low power flops are used to reduce hold time and power
- Cell footprints compatible and swappable in context
T5 Power Breakdown

- Aggressive power management through:
  - Gate sizing and design
  - High VT and long channel gates
  - Wire route optimization
  - Clock gating
  - Core cycle skipping
  - DVFS

- SPARC Cores 39.1%
- L3 Caches 6.8%
- Crossbar 3.4%
- SoC 6.9%
- Top Level 15.0%
- IO 1.0%
- Misc. 4.4%
- Leakage 23.4%
T5 Power Reduction with DFS vs. DVFS

- Scaling both voltage and frequency is significantly more efficient than just scaling frequency
- Improves performance in a constrained power envelope
- Reduces idle power
Full Dense Mesh Power Plane

- Two power layers with bidirectional current
- Reduces inductive loop distance and lateral resistance
- Global clocks routed in zipper channels
- Bump structure creates high density of vias for low resistance
Final Flop Clock Distribution

Clock Header Schematic

Datapath Physical Layout

Control Block Physical Layout
Core Flop Timing
T5 vs. T4

- Significant reduction in final stage clock skew
- Better clock slew rates
- Reduced overall latency
- Much tighter latency and slew rate dispersion
- Distributed drivers reduce SigEM and IR issues
Frequency Lock Loop Usage

• Primary clock source during full frequency operation
• Monitors processor power supply for noise and adjusts frequency to compensate
• Reduces overall timing margin needed to account for power supply noise
• Reduces power by allowing the chip to run at the same frequency at a lower pin voltage
Frequency Lock Loop Schematic

- Vdd (local)
- Frequency Detector
- Loop filter
- Ref Clock
- Vdd (local)
- Encoder
- Sub DCO [3:0]
- DCO
- Control
- Select
- 4:1 MUX
- DCO Clock
- Vdd (global)
Clock MUX Usage

- Enables DVFS and FLL functionality
- Used to dynamically switch between clock sources without causing glitches
- Switches between two PLL's as well as the FLL and external pin driven test clocks
- Disables clocks during initialization and reset sequences
- Used during test modes for clock bypass
Glitch-Free Clock MUX Schematic
L3 Cache Overview

- 8MB, 8-bank, 16-way set associative inclusive cache
- MOESI Coherence States per 64B cache line
- Speeds up IO by allocating DMA buffers in the cache
- Supports coherent flushing and retirement of cache lines to avoid persistent errors
L3 Data Cache Repair

Sub-array Repair

16KB sub-array redundancy

Hit logic

2:1 mux select

From sub-arrays

Shift to right

Shift to left

256 rows
16 col IO

3:1 Redundancy MUX

IO Repair

Redundant IO

Failed IO

Regular IO
Wide Band Amplifier Input Buffer

- Wide band flat gain is required to limit SNR impact of crosstalk amplification at high frequency.
- AC Path (top) provides a wide band flat gain from 1/10 of Nyquist to Nyquist.
- DC Path (bottom) provides flat gain up to Nyquist.
- 1/10 of Nyquist is selected to complement TX FIR, which typically has little eq power below 1/10 of Nyquist.
Input Buffer Eye Plots

- Tail suppression beyond first tap
- No boosting on the main pulse
- Enhance link margin by 30-40mV
T5 Tester Shmoo Plot
Conclusion

- Introduced next generation SPARC T5 processor with 16 cores and 128 threads
- Enables 8-socket glueless systems
- Provides significant performance improvement over previous generations
- Significant advances in power management including DVFS, enables a new generation of power efficient, fully secure data centers
- The T5 system paper will be presented in Session 3.7