UltraSPARC(TM)-I: A 64-bit Superscalar Processor with Multimedia Support

Marc Tremblay

Sparc Technology Business

Sun Microsystems Inc.

SPARC Technology

A Sun Microsystems Inc. Business

UltraSPARC-I Vital Statistics

Architecture	**64-bit SPARC-V9 with multimedia instruction extensions**
Pipeline	**4-issue superscalar, 9 stage pipeline**
Clock Freq	**167 Mhz**
Performance @ 167 Mhz, 2MB E$	**240 SPECint92 350 SPECfp92**
Power Supply	**3.3v**
Power @ 3.3V, 167 Mhz	**28W* 20mW sleep mode**
Die Size	**17.7 mm * 17.8 mm = 315 mm^2**
Transistors	**5.2 million (3.4 million logic)**
Technology	**CMOS, 4-layer metal, 0.5 μm drawn**

7.1-03

Architectural Highlights

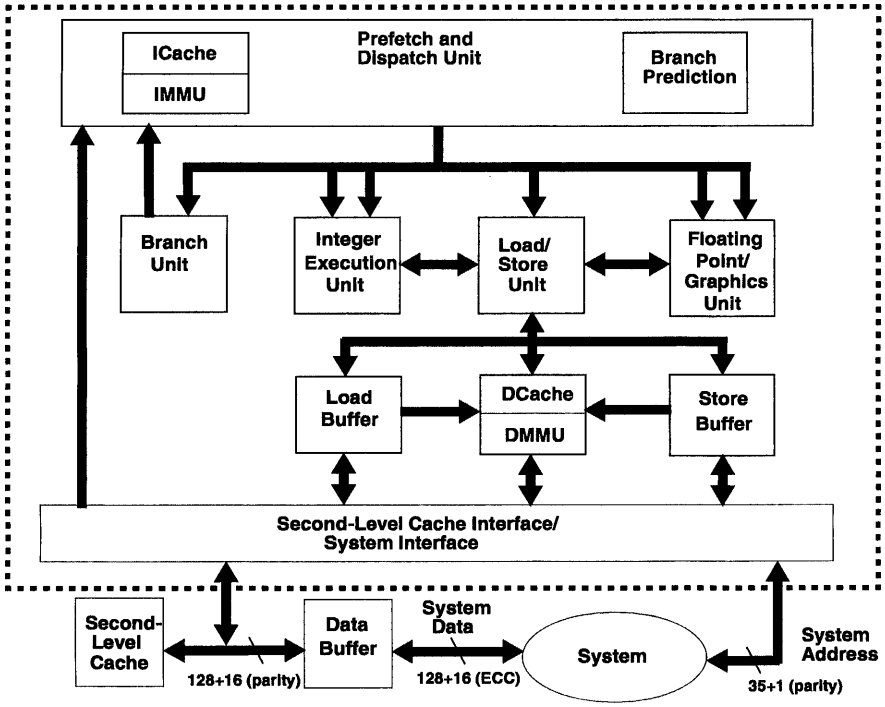
- **Goal: Sustain an execution rate of 4 instructions/cycle at a high clock rate even in the presence of:**
  - conditional branches
  - cache misses
- **Accomplished through:**
  - simple execution model
    - in-order dispatch, out-of-order completion
  - 9 functional units: 2 ALUs, 1 ld/st, 3 FPUs, 2 Graphics, 1 branch
  - single-cycle branch following mechanism
  - full throughput to the 2nd-level off-chip cache
    - non-blocking data cache
    - 9-entry load buffer
    - pipelined access to the SRAMs

7.1-04

Architectural Highlights (continued)

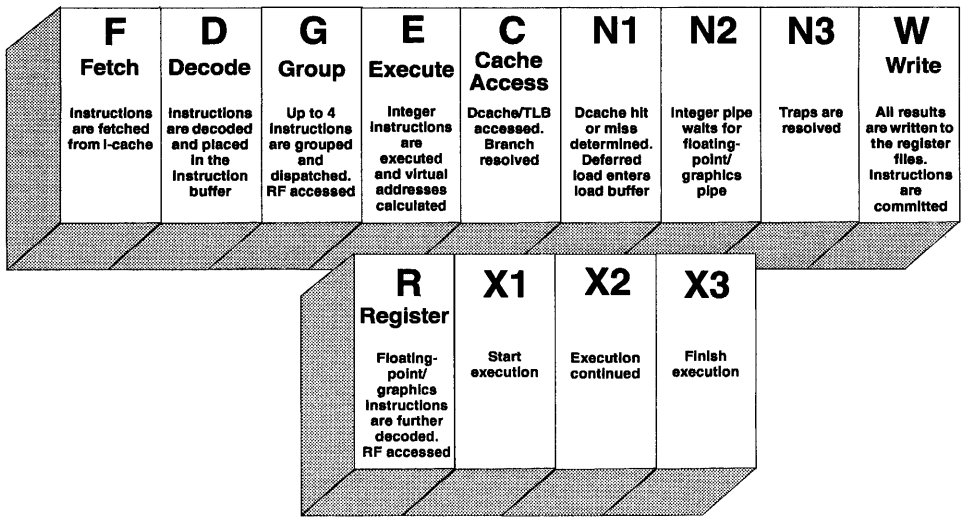
- **Provide on-chip support for system level functions**
  - Comprehensive multi-media support
    - VIS (Visual Instruction Set)
  - Microarchitecture support for networking
    - Block load, block store: transfer 600 MBytes/s without polluting caches
    - Parallel check-summing, encryption
  - Integrated 2nd-level cache controller
    - L2 cache is formed of synchronous SRAMs
  - Flexible MP support
    - Snooping and directory-based coherence protocols
    - Interconnect supports multiple address and data buses of different widths

Block Diagram



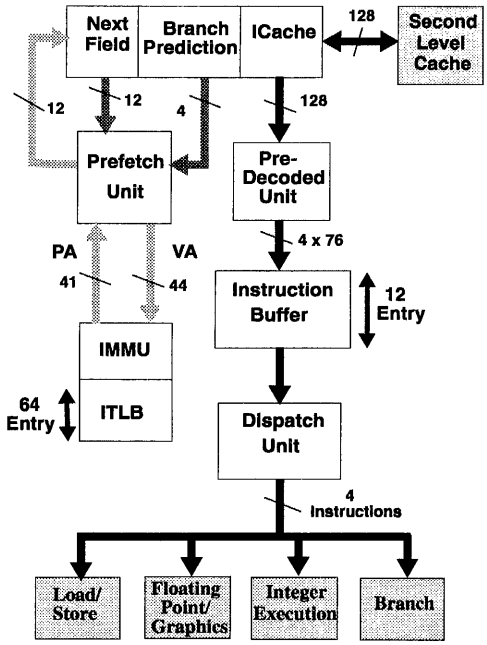
Pipeline Diagram

Integer Pipe



Floating-point/Graphics Pipe

Prefetch and Dispatch Unit

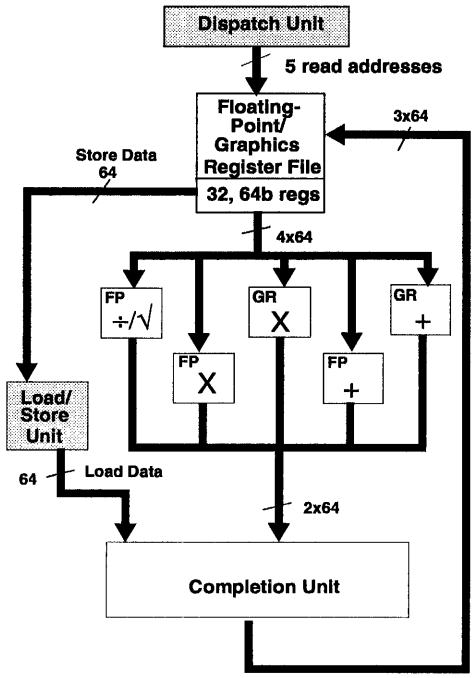


- 16kB ICache, 2-way set associative, 32B line size w/pre-decoded bits
- Next Field RAM which contains 1k branch target addresses and 2k dynamic branch predictions.

I0	I1	BP	I2	I3	BP	NFA

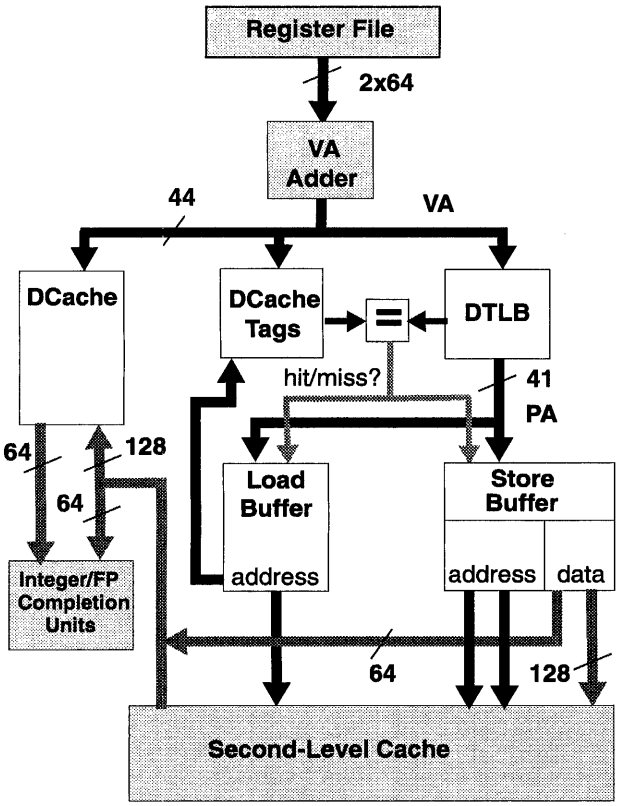
- 4-entry Return Address Stack
- Branch prediction accuracy: 88% SPECint92 / 94% SPECfp92
- 64-entry, fully associative ITLB backs up 1-entry μTLB
- 12-entry Instruction Buffer fed by ICache or second-level cache
- Single-cycle dispatch logic considers "top" 4 instructions

Floating-Point/Graphics Unit



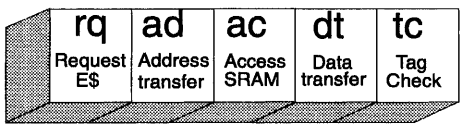
- Floating-point/Graphics Register File has 5 read ports/3 write ports
- 32 single-precision/32 double-precision registers
- 5 functional units, all fully pipelined except for Divide/Square Root unit
- High bandwidth: 2 FGops per cycle
- Short latencies
  - FP compares: *1 cycle*
  - FP add/subtract/convert: *3 cycles*
  - FP multiplies: *3 cycles*
  - FP divide/square root(sp/dp): *12/22 c.*
  - Partitioned add/subtract, align, merge, expand, logical: *1 cycle*
  - Partitioned multiply, pixel compare, pack, pixel distance: *3 cycles*

Load/Store Unit

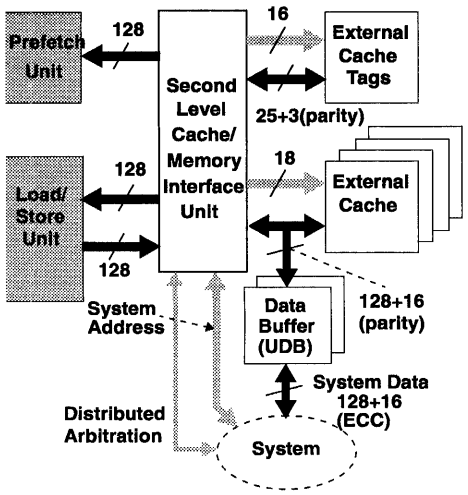


- 16 kB DCache (D$), direct-mapped, 32B line size w/16B sub-blocks
- 64-entry, fully associative DTLB supports multiple page sizes*
- D$ is non-blocking, supported by 9-entry Load Buffer
- D$ tags are dual-ported, allowing a tag check from the pipe and a line fill/snoop to occur in parallel
- Sustained throughput of 1 load per cycle from second-level cache (E$)
- Single-cycle stores to D$ and E$ via decoupled data and tag RAMs.
- Store compression dramatically reduces E$ bus utilization

External Cache Unit and System Interface



External Cache Pipeline



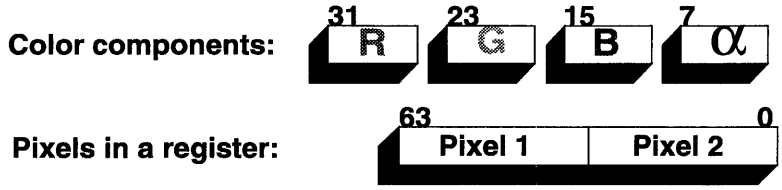
- E$ sizes from 0.5 MB to 4MB
- E$ is direct-mapped, physically indexed and tagged, w/64B line size
- 5 cache coherency states: MOESI
- Uses synchronous SRAMs with 1-entry internal write buffer
- 16B (+ parity) interface to E$ supports 2.6 GB/s sustained bandwidth.
- UDBs electrically isolate CPU/E$ from system data bus.
- Packet-based, 16B system interconnect with separate address and data busses
- Distributed arbitration w/low latency parking mode
- System is 1/2 or 1/3 CPU clock rate
- System address bus supports 4 CPU modules and system controller
- Supports snoopy and directory-based cache coherency protocols.

Multi-media Support

VIS - Data Format

- Tailored for graphics

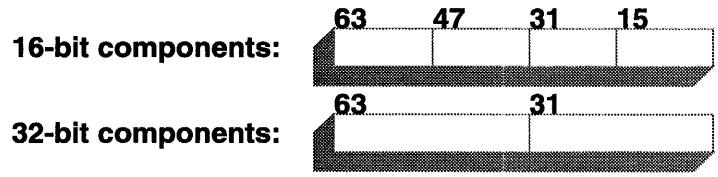
Pixel format:



- 8-bit Red, Green, Blue components
- Alpha value for transparency coefficient
- Band-interleaved and Band-sequential supported
- Conversion instructions

VIS - Data Format (cont.)

Fixed data:

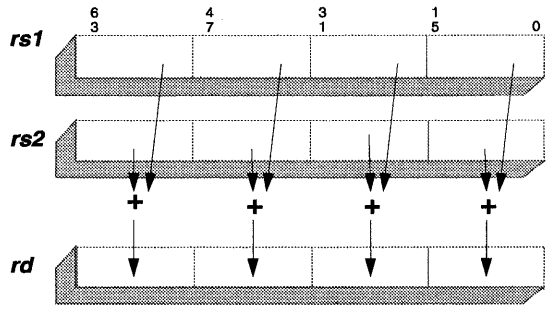


16/32-bit fixed point components

- **used for additional precision or larger dynamic range**
- **e.g. intermediate results during image processing**
- **very high quality imaging**
  - **medical processing**
  - **color pre-press imaging**

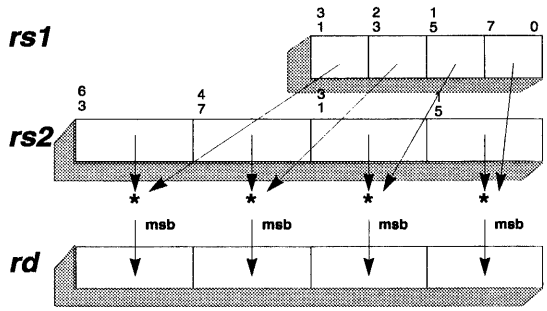
Arithmetic Instructions

- **Core of 2D, 3D graphics and image processing**
  - **filtering, smoothing, alpha blending -> pixel add/mul**
  - **Gouraud shading -> pixel interpolation**
  - **scaling/rotation, DCT, convolution, warping -> pixel add/mul**
- **Partitioned add and subtract:**
  - **4 16-bit add/sub or 2 32-bit add/sub**
  - **FPADD16 instruction:**



Arithmetic Instructions (cont.)

- **Partitioned multiply**
  - **4 8-bit X 16-bit multiplications**
  - **signed and unsigned**
  - **distributed and pair-wise**
  - **Mul8x16 instruction (pair-wise):**



Motion Estimation Instruction

- **dominant computation in real-time video compression**
- **minimal changes in the position of images from one frame to the next**
- **search for a motion value that minimizes estimation error**
  - Pdist operates on 8 pixels in parallel:

$$|a_1-b_1| + |a_2-b_2| + |a_3-b_3| + |a_4-b_4| + |a_5-b_5| + |a_6-b_6| + |a_7-b_7| + |a_8-b_8|$$

  - 8 subtracts, 8 absolute values, 7 adds, and 1 accumulate -> 24 operations
  - also replaces numerous loads and shifts required
    - replaces ~48 instructions
    - 16x16 blocks requires 32 pdist instructions
    - would typically require ~1500 conventional instructions

Code Example

- **(to be provided later)**
  - sofware pipelined loop
  - accesses L2 cache at full througput
  - uses VIS to accelerate pixel processing

Summary

- 64-bit SPARC V9 CPU - binary compatibility
- 4-issue, superscalar
- 167 Mhz
- High bandwidth memory system
- Flexible multiprocessing support
- Advanced on-chip multimedia features
- High performance machine for real world applications