Module 8: Memory Consistency Models and Case Studies of Multi-core

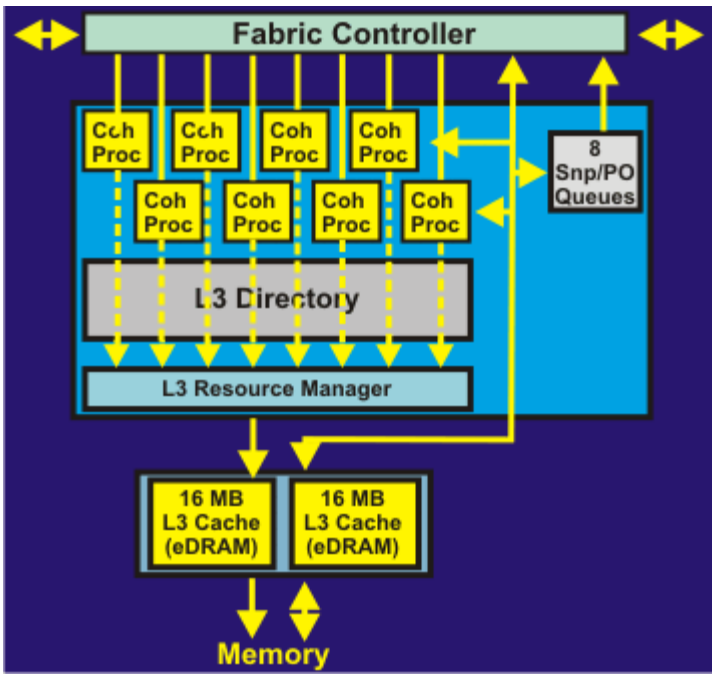
Lecture 16: Case Studies of Multi-core

The Lecture Contains:

Previous Next

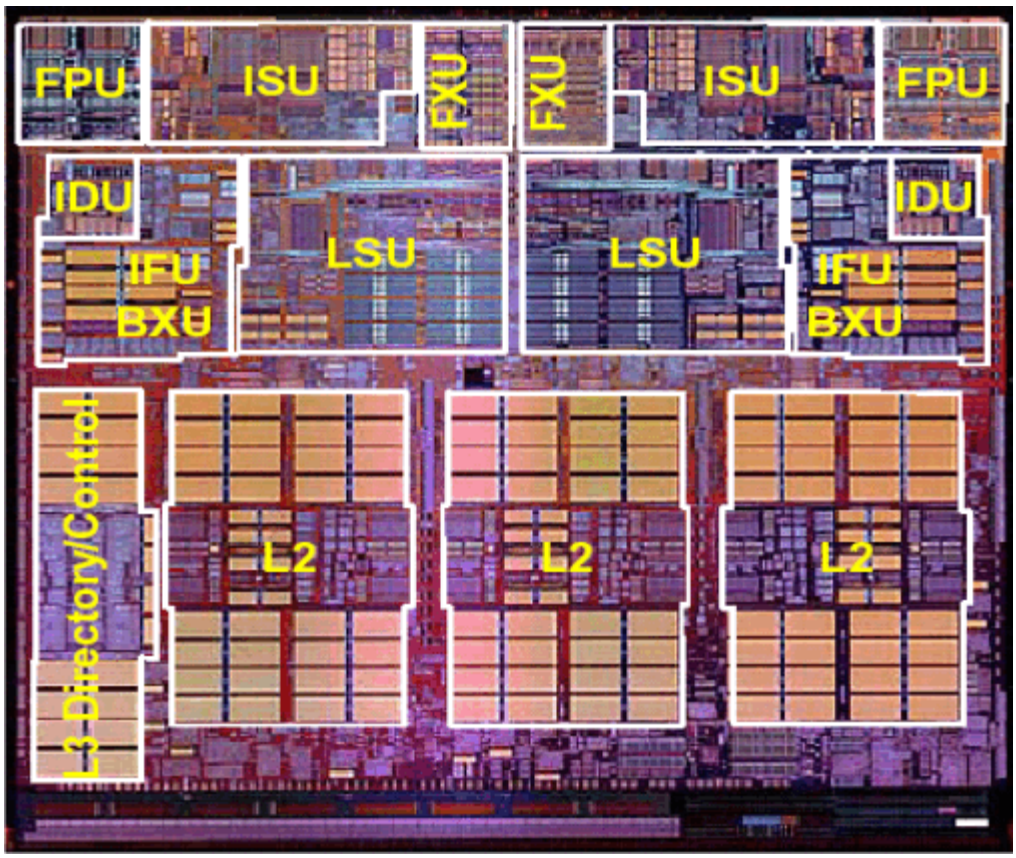
POWER4 L3 Cache

- On-chip tag (IBM calls it directory), off-chip data
  - 32 MB/8-way associative/512 bytes line
  - Contains eight coherence/snoop controllers
  - Does not maintain inclusion with L2: requires L3 to snoop fabric interconnect also
  - Maintains five coherence states
  - Putting the L3 cache on the other side of the fabric requires every L2 cache miss (even local miss) to cross the fabric: increases latency quite a bit



Previous Next

POWER4 Die Photo



IBM POWER5

- Carries on POWER4 to the next generation
  - Each core of the dual-core chip is 2-way SMT: 24% area growth per core
  - More than two threads not only add complexity, may not provide extra performance benefit; in fact, performance may degrade because of resource contention and cache thrashing unless all shared resources are scaled up accordingly (hits a complexity wall)
  - L3 cache is moved to the processor side so that L2 cache can directly talk to it: reduces bandwidth demand on the interconnect (L3 hits at least do not go on bus)
  - This change enabled POWER5 designers to scale to 64-processor systems (i.e. 32 chips with a total of 128 threads)
  - Bigger L2 and L3 caches: 1.875 MB L2, 36 MB L3
  - On-chip memory controller

Previous Next

IBM POWER5

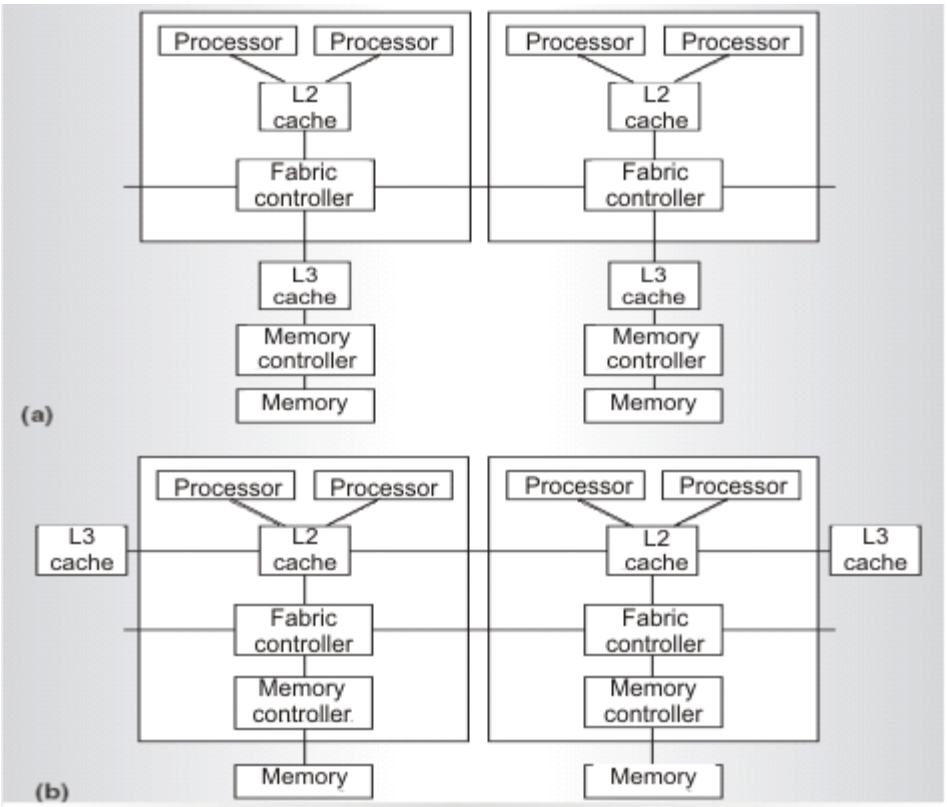


Figure1. Power4 (a) and Power (b) system structures.

Reproduced from IEEE Micro

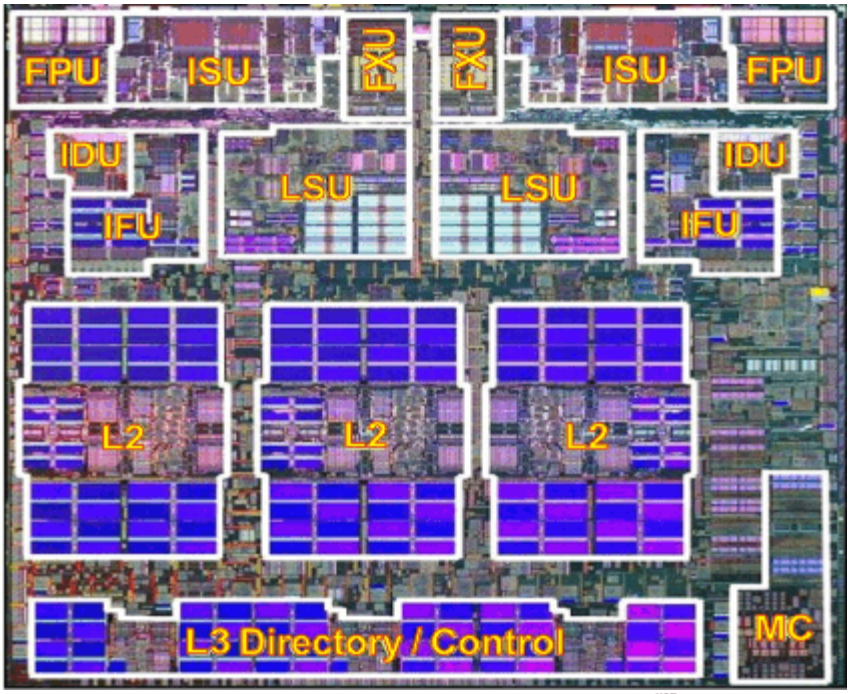
- Same pipeline structure as POWER4
  - Added SMT facility
  - Like Pentium 4, fetches from each thread in alternate cycles (8-instruction fetch per cycle just like POWER4)
  - Threads share ITLB and ICache
  - Increased size of register file compared to POWER4 to support two threads: 120 integer and floating-point registers (POWER4 has 80 integer and 72 floating-point registers): improves single-thread performance compared to POWER4; smaller technology (0.13 µ m) made it possible to access a bigger register file in same or shorter time leading to same pipeline as POWER4
  - Doubled associativity of L1 caches to reduce conflict misses: icache is 2-way and dcache is 4-way

Previous Next

IBM POWER5

- Dynamic power management
  - With SMT and CMP average number of switching per cycle increases leading to more power consumption
  - Need to reduce power consumption without losing performance: simple solution is to clock it at a slower frequency, but that hurts performance
  - POWER5 employs fine-grain clock-gating: in every cycle the power management logic decides if a certain latch will be used in the next cycle; if not, it disables or gates the clock for that latch so that it will not unnecessarily switch in the next cycle
  - Clock-gating and power management logic themselves should be very simple
  - If both threads are running at priority level 1, the processor switches to a low power mode where it dispatches instructions at a much slower pace

POWER5 Die Photo



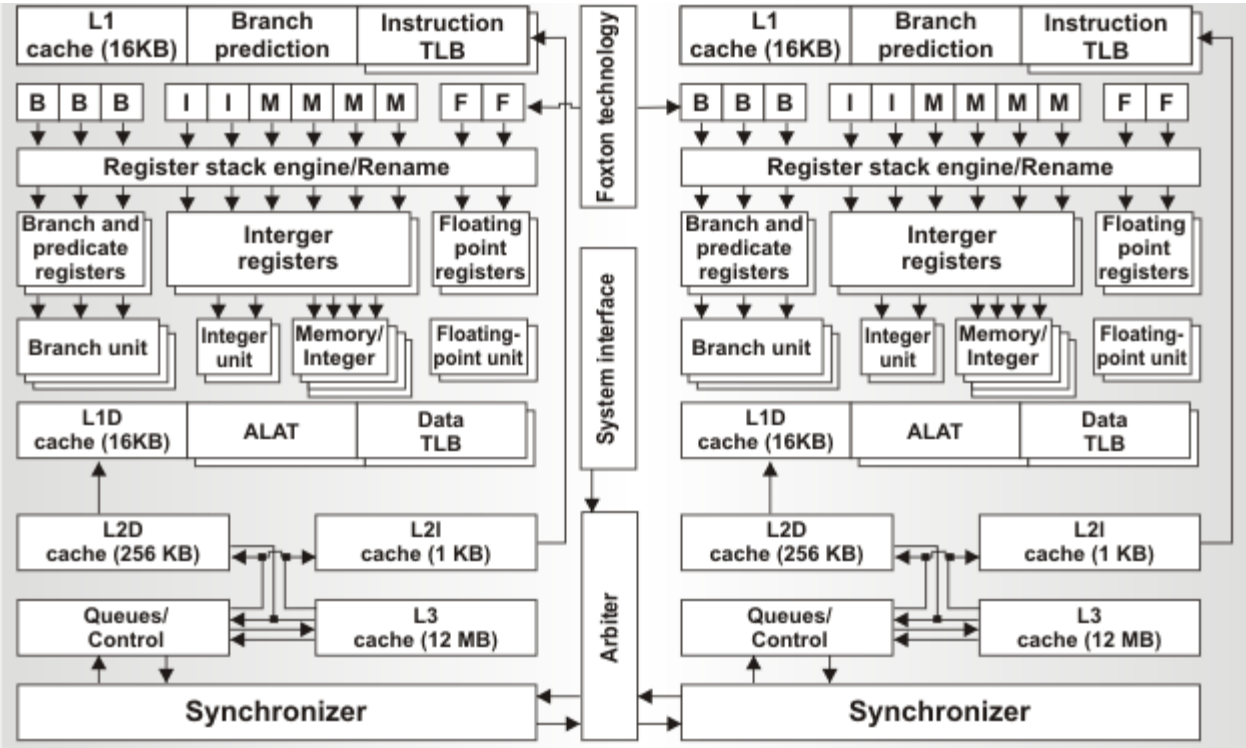
Previous Next

Intel Montecito

Features

- Dual core Itanium 2, each core dual threaded
- 1.7 billion transistors, 21.5 mm x 27.7 mm die
- 27 MB of on-chip three levels of cache
  - Not shared among cores
- 1.8+ GHz, 100 W
- Single-thread enhancements
  - Extra shifter improves performance of crypto codes by 100%
  - Improved branch prediction
  - Improved data and control speculation recovery
  - Separate L2 instruction and data caches buys 7% improvement over Itanium2; four times bigger L2I (1 MB)
  - Asynchronous 12 MB L3 cache

Overview



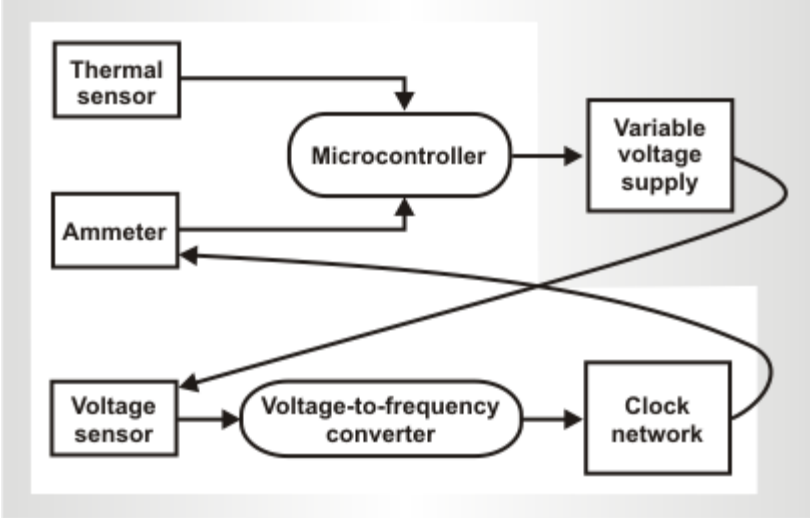
Reproduced from IEEE Micro

Previous Next

Power Efficiency

- Foxton technology
  - Blind replication of Itanium 2 cores at 90 nm would lead to roughly 300 W peak power consumption (Itanium 2 consumes 130 W peak at 130 nm)
  - In case of lower than the ceiling power consumption, the voltage is increased leading to higher frequency and performance
  - 10% boost for enterprise applications
  - Software or OS can also dictate a frequency change if power saving is required
  - 100 ms response time for the feedback loop
  - Frequency control is achieved by 24 voltage sensors distributed across the chip: the entire chip runs at a single frequency (other than asynchronous L3)
  - Clock gating found limited application in Montecito

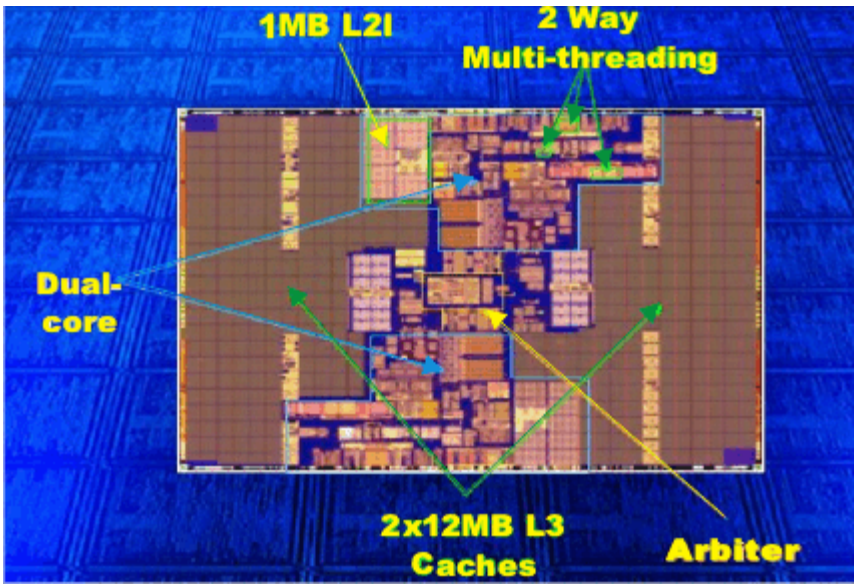
Foxton Technology



Reproduced from IEEE Micro

- Embedded microcontroller runs a real-time scheduler to execute various tasks

Die Photo



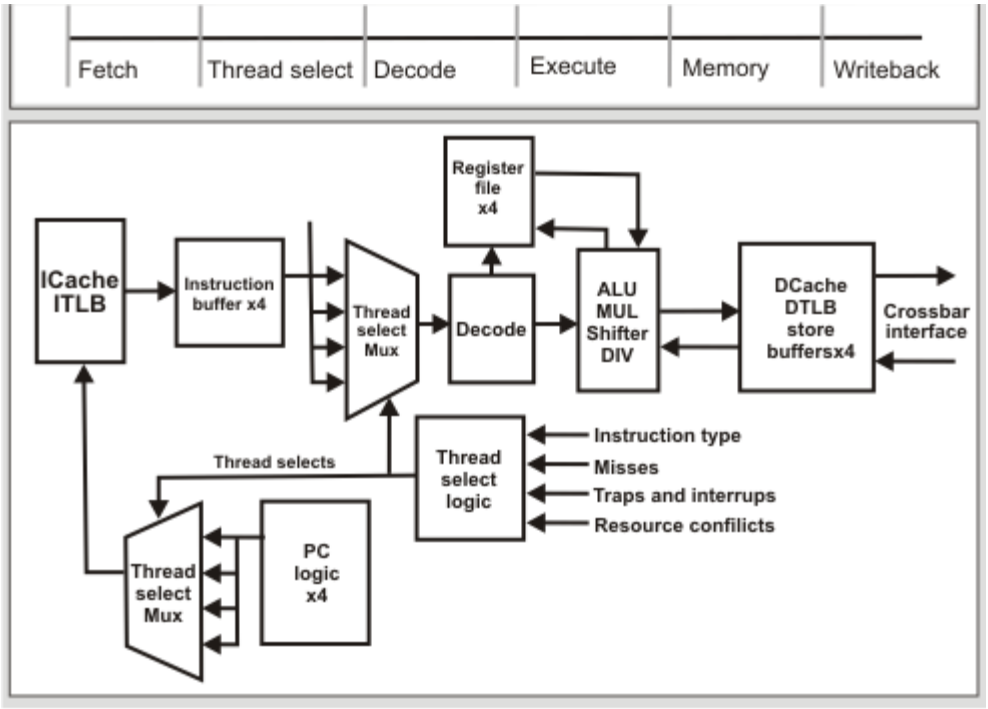
Sun Niagara
OR
Ultrasparc T1

Features

- Eight pipelines or cores, each shared by 4 threads
  - 32-way multithreading on a single chip
  - Starting frequency of 1.2 GHz, consumes 60 W
  - Shared 3 MB L2 cache, 4-way banked, 12-way set associative, 200 GB/s bandwidth
  - Single-issue six stage pipe
  - Target market is web service where ILP is limited, but TLP is huge (independent transactions)
- Throughput matters

Previous Next

Pipeline Details



Reproduced from IEEE Micro

Cache Hierarchy

- L1 instruction cache
  - 16 KB / 4-way / 32 bytes / random replacement
  - Fetches two instructions every cycle
  - If both instructions are useful, next cycle is free for icache refill
- L1 data cache
  - 8 KB / 4-way / 16 bytes/ write-through, no-allocate
  - On avearge 10% miss rate for target benchmarks
  - L2 cache extends the tag to maintain a directory for keeping the core L1s coherent
- L2 cache is writeback with silent clean eviction

Previous Next