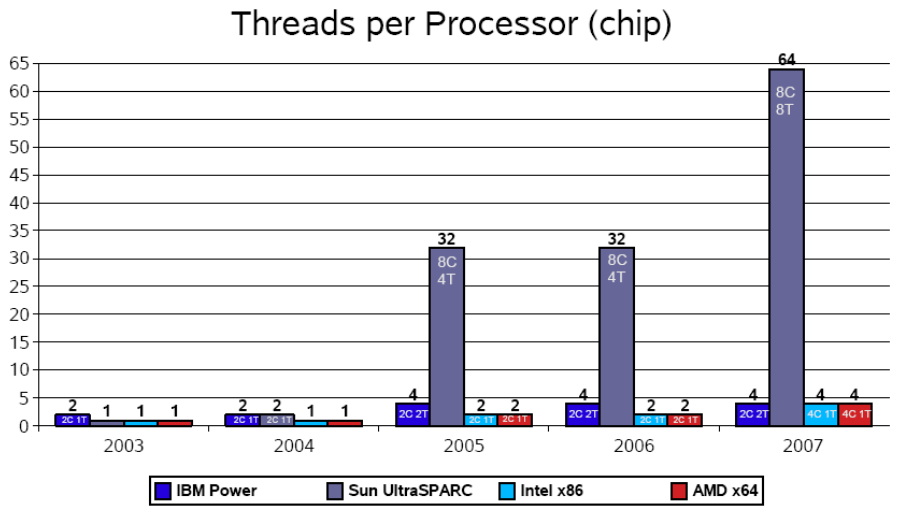
Multi-core Systems
What can we buy today?

Ian Watson & Mikel Lujan
Advanced Processor Technologies Group

A Bit of History

- AMD Opteron introduced in 2003
  - Hypertransport
  - On-chip memory controller
- In 2006
  - Sun Niagara I: 8 cores/32 thread & 4 on-chip memory controllers
  - Intel & AMD: dual core
  - IBM: dual core & Cell
  - Azul: 24 cores
- In 2007
  - Sun Niagara II: 8 cores/64 threads
  - Intel quad-core vs. AMD quad-core (Barcelona)
  - Azul: 48 cores
- In 2009
  - Sun Victoria Falls: 4 Niagara II+ (256 threads)
    - Server (web, DBs)
  - Intel Nehalem (core i7 & Xeon) vs. AMD quad-core (Shanghai) six-core (Istanbul)
    - Desktop/workstations/Server
  - ARM: Cortex A9 4 cores
    - low-power embedded systems (e.g. mobile phones)
  - Nvidia: Tesla (HPC), GTX

A bit of history (2)

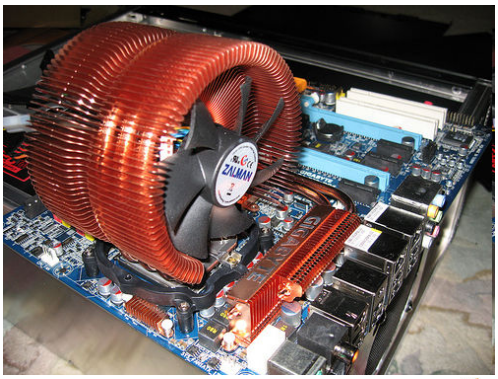


3

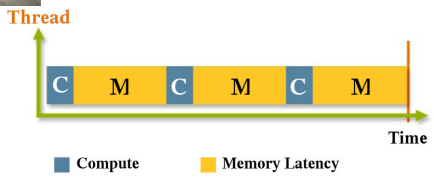
Is cooling a problem?

4

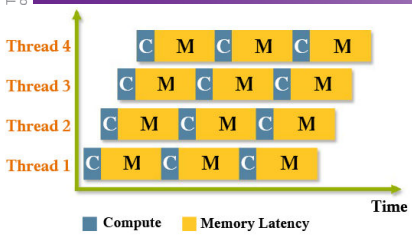
Problems: heat & memory wall

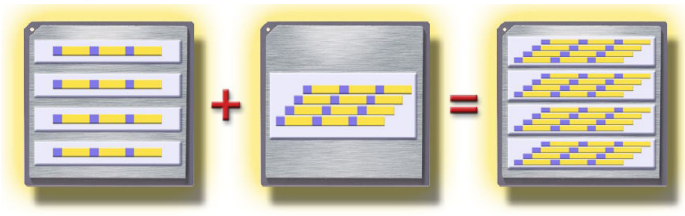


- Processor utilization (15%-25%)



Chip Multi-Threading – Towards Niagara



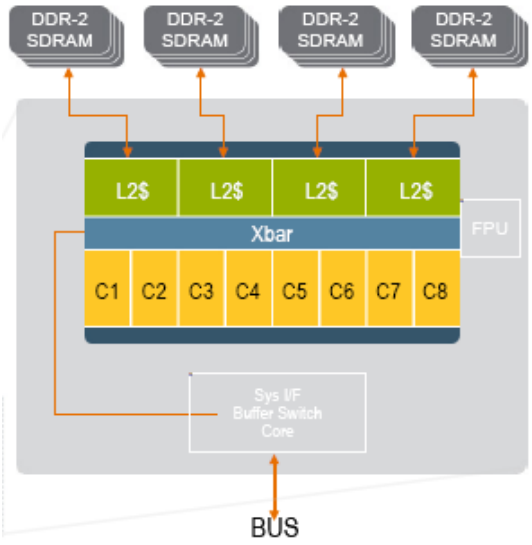


CMP
(Chip MultiProcessing, a.k.a. "multicore")
n cores per processor

HMT
(Hardware Multithreading)
m threads per core

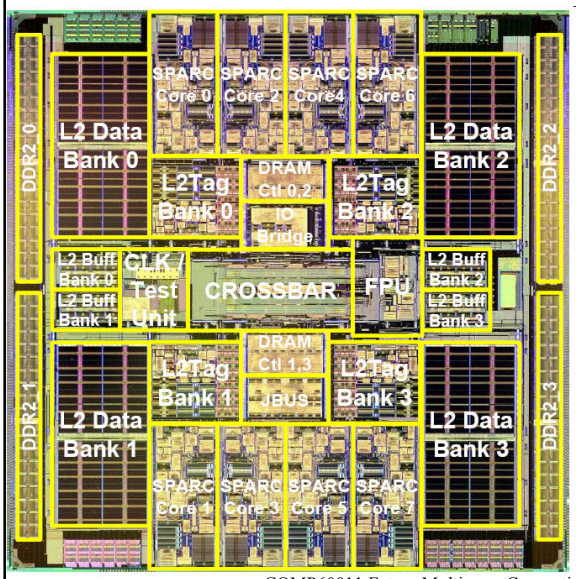
CMT
(Chip MultiThreading)
n x *m* threads per processor

Niagara



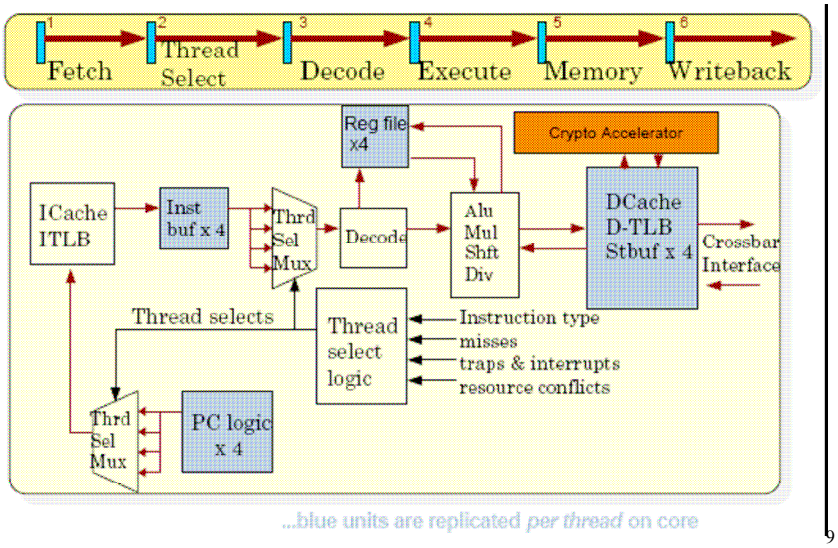
- SPARC architecture
- 8 cores (1GHz - 1.4GHz)
- 4 hardware threads per core
- Banked L2 cache
- 4 on-chip memory controllers
- 23GB/s off-chip bandwidth
- One FPU
- Crossbar

Niagara Floorplan



- 378 mm^2 die
- 90nm process
- ~300M transistors
- Single core 11 mm^2
- Multi-Threading 20% extra area

Niagara Core Pipeline



Niagara Memory System

- Instruction cache
  - 16kB, 4-way set associative, 32B line size
- Data cache
  - 8kB, 4-way set associative, 16B line size
  - Write-through cache, write-around on miss
- L2 cache
  - 3 MB, 12-way set associative, 64B line size
  - Write-through
  - 4-way banked by line
- Pseudo-random eviction policy
- Coherency
  - Data cache lines have 2 state: valid or invalid
  - Data cache is write through -> no modified state
  - Crossbar enforces ordering according to memory consistency
- Caches kept coherent by tracking lines in directories in L2
- Latencies
  - Load-to-use: 3 cycles
  - L2 latency: 22 cycles
  - Memory latency: ~90 ns

Thread Selection Policy

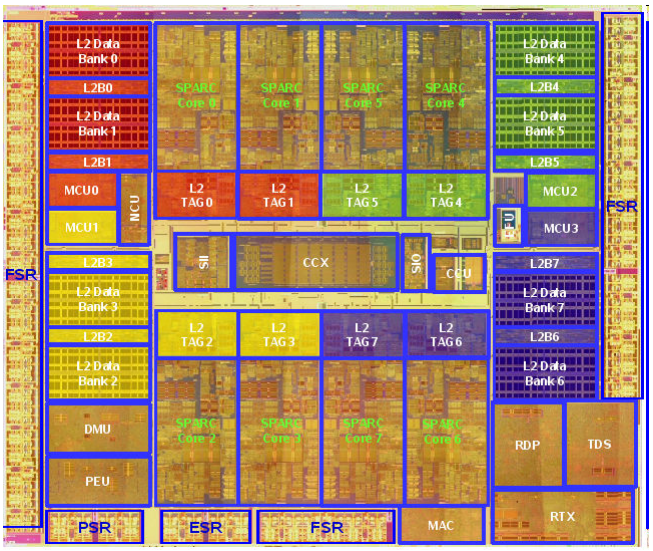
Towards Niagara 2

- Goal:
  - Double throughput versus Niagara 1 staying within the same power envelop

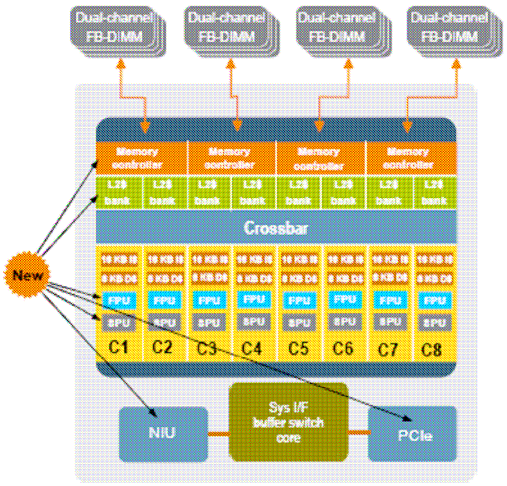
Modifications in Niagara 2

- Double number of threads from 4 to 8
  - Choose 2 threads out of 8 to execute each cycle
- Double execution units from 1 to 2 and add FPU
- Double set associativity of L1 instruction cache to 8-way
- Double size of fully associative DTLB from 64 to 128 entries
- Double L2 banks from 4 to 8
- Hardware tablewalk for ITLB and DTLB misses
- One extra stage in the pipeline (select 2 threads out of 8)

Niagara 2 Floorplan (65nm &342mm^2)



Niagara 2 – New Elements



- 8 threads per core
- Two integer pipes per core
- One fully-pipelined FPU per core
- 60+GB/s off-chip
- 4MB L2$ (8 banks) 16 way
- Security co-processor per core
- On-chip

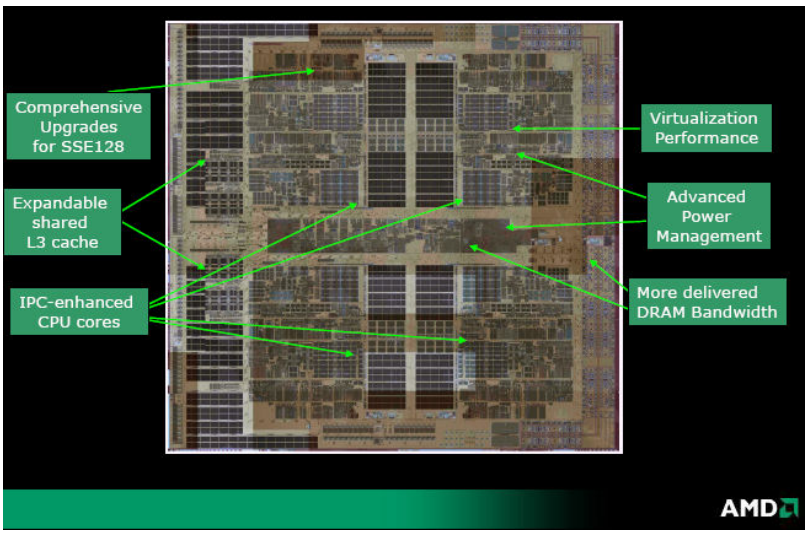
Sun Microsystems' Open Source CMP

- http://www.opensparc.net
- OpenSPARC T1 and T2
  - Verilog available (RTL)
  - Book available
  - Webcast of talks
  - More information on support for Virtualization, RAS

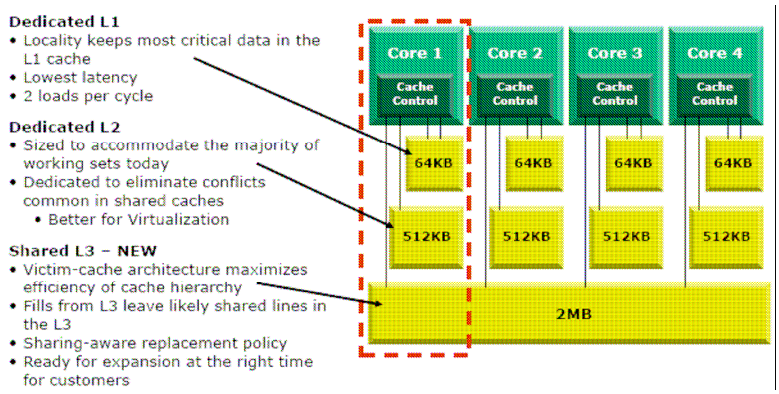
A Bit of History

- AMD Opteron introduced in 2003
  - Hypertransport
  - On-chip memory controller
- In 2006
  - Sun Niagara I: 8 cores/32 thread & 4 on-chip memory controllers
  - Intel & AMD: dual core
  - IBM: dual core & Cell
  - Azul: 24 cores
- In 2007
  - Sun Niagara II: 8 cores/64 threads
  - Intel quad-core vs. AMD quad-core (Barcelona)
  - Azul: 48 cores
- In 2009
  - Sun Victoria Falls: 4 Niagara II+ (256 threads)
    - Server (web, DBs)
  - Intel Nehalem (core i7 & Xeon) vs. AMD quad-core (Shanghai) six-core (Istanbul)
    - Desktop/workstations/Server
  - ARM: Cortex A9 4 cores
    - low-power embedded systems (e.g. mobile phones)
  - Nvidia: Tesla (HPC), GTX

AMD Barcelona (65nm, <3GHz)

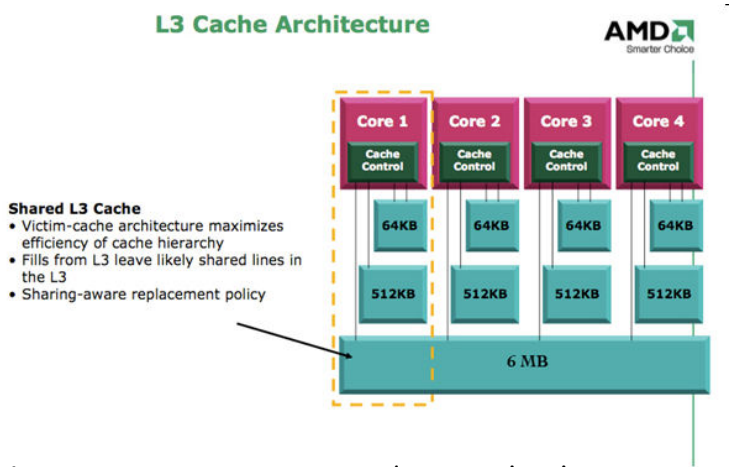


AMD Barcelona



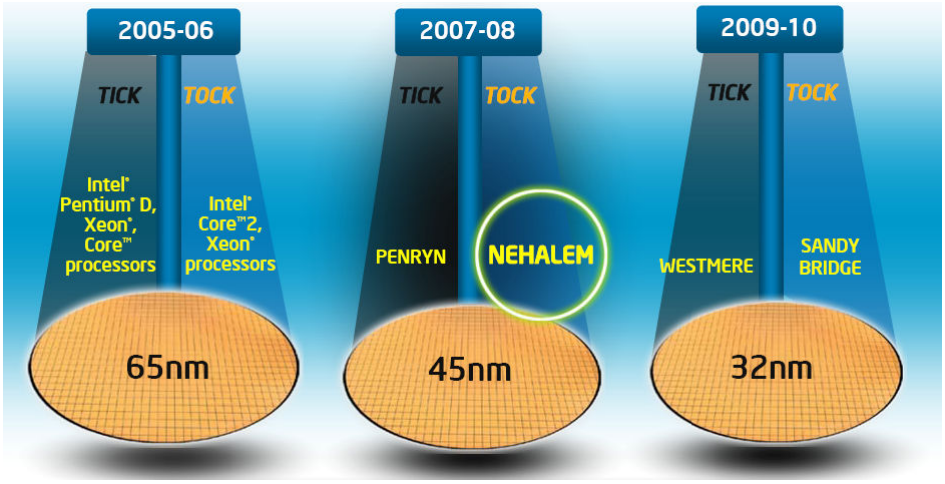
- L1: 2-way set associative, LRU replacement, block size 64 bytes, split I & D, write-back & write-allocate, 3 cycles latency
- L2: idem. except 9 cycles latency
- L3: idem. except evict block shared by fewest core and 34 cycles latency

AMD Shanghai (45nm, <2.6GHz)

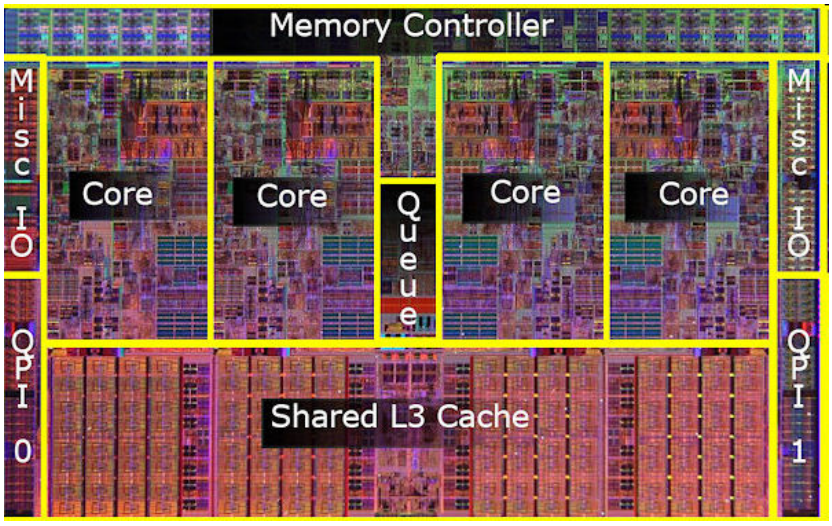


L3: 48-way set associative and 29 cycles latency

Intel Tick Tock

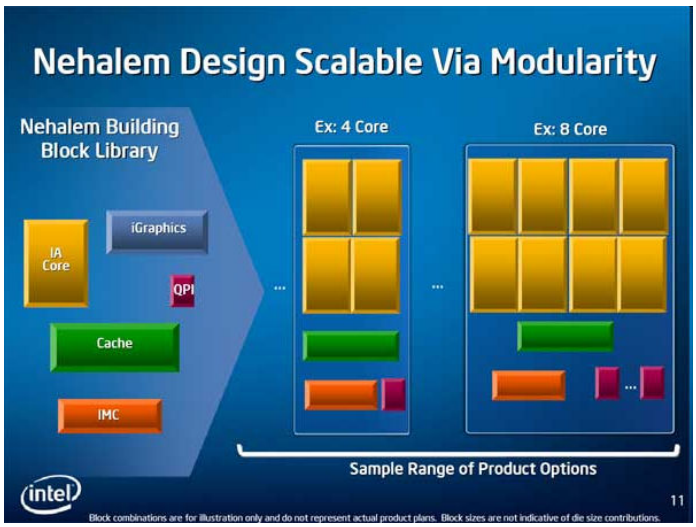


Core i7 (Nehalem, 2.6GHz -3.2GHz)

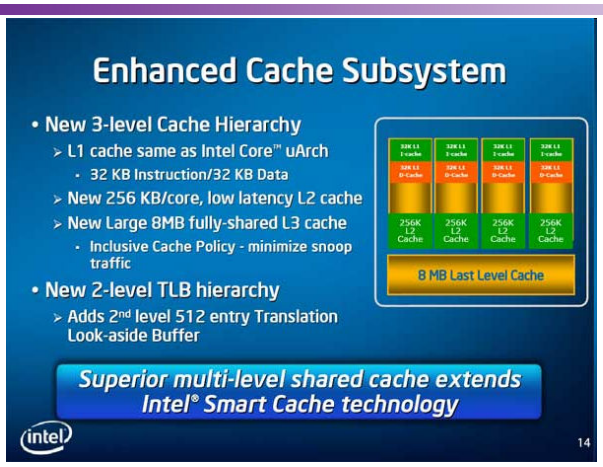


2 Simultaneous Multi-Threading per core

Nehalem



Nehalem Caches

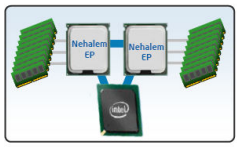


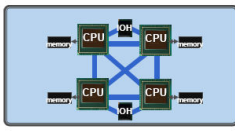
- L1: split D$ & I$, 32KB each, 4-way I$ & 8-way set associative, approx. LRU, block size 64 bytes, write-back & write-allocate
- L2: 8-way set associative, idem.
- L3: 16-way set associative, idem

NUMA

Intel® QuickPath Interconnect

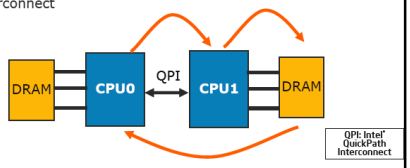
- Nehalem introduces new Intel® QuickPath Interconnect (QPI)
- ***High bandwidth, low latency*** point to point interconnect
- Up to 6.4 GT/sec initially
  - 6.4 GT/sec -> 12.8 GB/sec
  - Bi-directional link -> 25.6 GB/sec per link
  - Future implementations at even higher speeds
- Highly ***scalable*** for systems with varying # of sockets





Remote Memory Access

- CPU0 requests cache line X, not present in any CPU0 cache
  - CPU0 requests data from CPU1
  - Request sent over QPI to CPU1
  - CPU1's IMC makes request to its DRAM
  - CPU1 snoops internal caches
  - Data returned to CPU0 over QPI
- Remote memory latency a function of having a low latency interconnect



Other things not discussed Sun, AMD & Intel

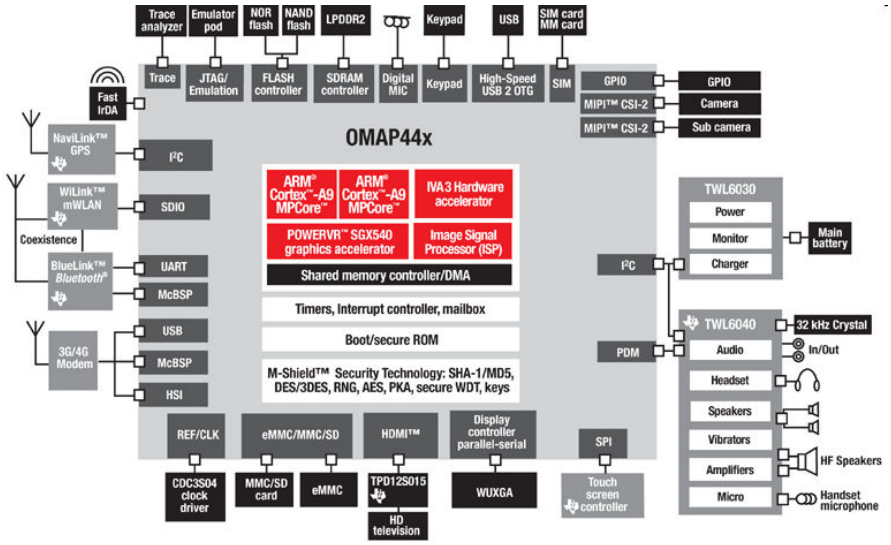
- Throttling to reduce power
- Support for VMM
- New instructions

A Bit of History

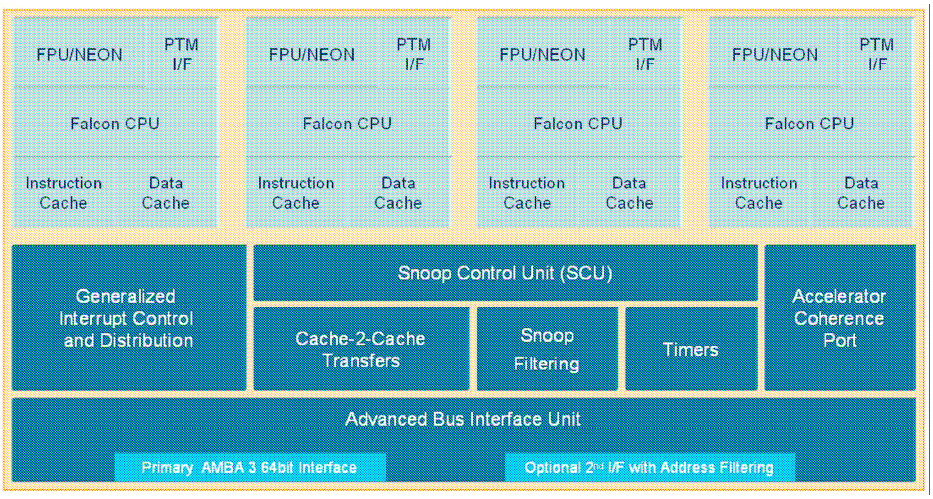
- AMD Opteron introduced in 2003
  - Hypertransport
  - On-chip memory controller
- In 2006
  - Sun Niagara I: 8 cores/32 thread & 4 on-chip memory controllers
  - Intel & AMD: dual core
  - IBM: dual core & Cell
  - Azul: 24 cores
- In 2007
  - Sun Niagara II: 8 cores/64 threads
  - Intel quad-core vs. AMD quad-core (Barcelona)
  - Azul: 48 cores
- In 2009
  - Sun Victoria Falls: 4 Niagara II+ (256 threads)
    - Server (web, DBs)
  - Intel Nehalem (core i7 & Xeon) vs. AMD quad-core (Shanghai) six-core (Istanbul)
    - Desktop/workstations/Server
  - ARM: Cortex A9 4 cores
    - low-power embedded systems (e.g. mobile phones)
  - Nvidia: Tesla (HPC), GTX

Mobile phone

Texas Instrument OMAP 44x MPSoC



ARM Cortex A9 MPCore



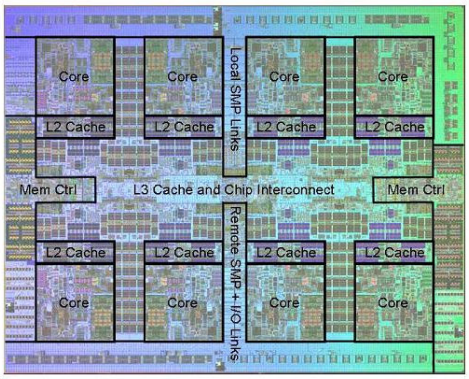
Atomic Instructions

- Compare-And-Swap
  - Niagara, Intel & AMD
- Load Linked & Store Conditional
  - ARM & IBM Power

2010 – New Systems

POWER7 Processor Chip

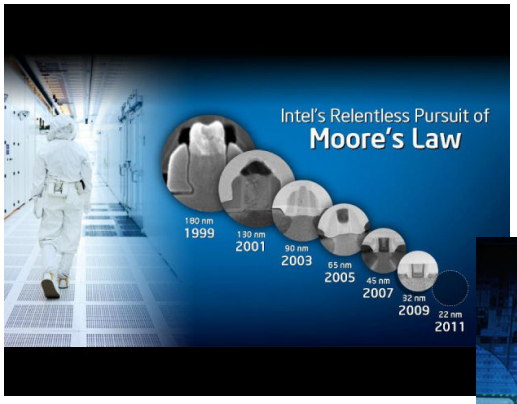
- 567mm² Technology: 45nm lithography, Cu, SOI, eDRAM
- 1.2B transistors
  - Equivalent function of 2.7B
  - eDRAM efficiency
- Eight processor cores
  - 12 execution units per core
  - 4 Way SMT per core
  - 32 Threads per chip
  - 256KB L2 per core
- 32MB on chip eDRAM shared L3
- Dual DDR3 Memory Controllers
  - 100GB/s Memory bandwidth per chip sustained
- Scalability up to 32 Sockets
  - 360GB/s SMP bandwidth/chip
  - 20,000 coherent operations in flight
- Advanced pre-fetching Data and Instruction
- Binary Compatibility with POWER6

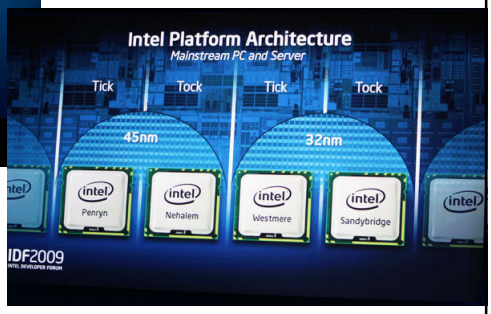


2010 – New Systems

- AMD 12-core Opteron
- Sun Microsystems - Niagara 3 (or RainbowFalls)
  - 16-cores (8 or 16 threads per core?)
- Nvidia Fermi – 512 Cuda cores
- ARM – Cortex A9 at 2GHz

Looking at the future





Papers for Next Week

- Simultaneous multithreading: maximizing on-chip parallelism. ISCA 1995.
  - http://doi.acm.org/10.1145/223982.224449
- The SGI Origin: a ccNUMA highly scalable server. ISCA 1997.
  - http://doi.acm.org/10.1145/264107.264206
- Java Tutorial: Concurrency
  - http://java.sun.com/docs/books/tutorial/essential/concurrency/
- Remember to email the slides before 14:00 13th November 2009