Lecture 4: Parallel Hardware: Examples

Mohamed Zahran (aka Z)
mzahran@cs.nyu.edu
http://www.mzahran.com
Processors We Will Look at

- SPARC T4
- IBM Power 7
- Intel Sandy Bridge
SPARC T4

- The next generation of Oracle multicore
- 855M transistors
- Supports up to 64 threads
  - 8 cores
  - 8 threads per core
    - Cannot be deactivated by software
- Private L1 and L2 and shared L3
- Shared L3
  - Shared among 8 cores
  - Banked
  - 4MB
  - 16-way set associative
  - Line size of 64 bytes
PEU: PCI-Express unit;    DMU: Data management unit;    NIU: Network interface unit;    NCU: Non-cacheable unit;   SIU: System interface unit;    BoB: Buffer on Board
The Cores in SPARC T4
The Cores in SPARC T4

• Supports up to 8 threads
• DL1 and IL1:
  – 16KB
  – 4-way set associative
  – 32 bytes cache line
  – Shared by all 8 threads
• IL1 has 3 line prefetch on-miss
• DL1 has stride-based and next-line prefetchers
What to Do About Prefetching?

• Use arrays as much as possible. Lists, trees, and graphs have complex traversals which can confuse the prefetcher.

• Avoid long strides. Prefetchers detect strides only in a certain range because detecting longer strides requires a lot more hardware storage.

• If you must use a linked data structure, pre-allocate contiguous memory blocks for its elements and serve future insertions from this pool.

• Can you re-use nodes from your linked-list?
Questions

• Suppose that you have 8 threads that are computation intensive and another 8 memory bound... how will you assign them to cores on T4?

• What if all threads are computation bound?

• What if they are all memory bound?

• T4 gives the software the ability to pause a thread for few cycles. When will you use this feature?
POWER7 Processors: The Beat Goes On

Approaching 20 Years of POWER Processors

- RS64i V Sstar
- RS64i II Pulsar
- RS64i II North Star
- RS64i Apache BiCMOS
- Muskie A35
- -Cobra A10 -64 bit
- .5um

- .22um
- .35um
- .72um

- RSC
- -603
- -601

- .6um
- .35um
- .25um

- 1.0um

- .35um

- POWER2™ P2SC

- POWER3™ -630

- POWER4™ -Dual Core

- POWER4™ -SMT

- POWER5™ -Ultra High Frequency

- POWER7™ -Multi-core

Major POWER® Innovation
- 1990 RISC Architecture
- 1994 SMP
- 1995 Out of Order Execution
- 1996 64 Bit Enterprise Architecture
- 1997 Hardware Multi-Threading
- 2001 Dual Core Processors
- 2001 Large System Scaling
- 2001 Shared Caches
- 2003 On Chip Memory Control
- 2003 SMT
- 2006 Ultra High Frequency
- 2006 Dual Scope Coherence Mgmt
- 2006 Decimal Float/VSX
- 2006 Processor Recovery/Sparing
- 2009 Balanced Multi-core Processor
- 2009 On Chip EDRAM

Source: Slides from Joel M. Tendler from IBM
IBM Power 7

- Supports global shared memory space for POWER7 clusters
  - So you can program a cluster as if it were a single system
- Design for power-efficiency, unlike Power 6
- ~1.2B Transistors
- Up to 8 cores and 4-way SMT
- TurboCore mode that can turn off half of the cores from an eight-core processor, but those 4 cores have access to all the memory controllers and L3 cache at increased clock speeds.
- 3.0 – 4.25 GHz clock speed
IBM Power 7: Cache Hierarchy

- 32KB DL1 and IL1 per core
- 256KB L2 per core
- eDRAM L3 4MB per core (total of 32MB)
  - Very flexible design for L3
Questions

• If you are writing the same programs for T4 and Power 7, will you change anything?

• As a programmer, how can you make use of the cache hierarchy?
Intel Sandy Bridge

- New microarchitecture
- 22nm
Intel Sandy Bridge

- New microarchitecture
- 22nm

Heterogeneous Multicore

Sandy Bridge, source: Intel
Intel Sandy Bridge

- Improvement over its predecessor Nehalem
- Targeting multimedia applications
  - Introduced Advanced Vector Extensions (AVX)
• The GPU can access the large L3 cache
• Intel’s team totally re-designed the GPU
How About Supercomputers?

http://www.top500.org/
IBM BlueGene/L

- Started December 1999
- Main goal: to build a petaflop/s scale supercomputer to attack science problems such as protein folding.
- Strategy: Massive collection of low-power CPUs instead of a moderate-sized collection of high-power CPUs.
- BlueGene is a family of supercomputers.
  - BlueGene/L is the first generation
  - BlueGene/P is the petaflop generation
  - BlueGene/Q is the third generation
IBM BlueGene/L

- A large number of nodes (65,536)
  - Low-power (20W) nodes for density
  - High floating-point performance
  - System-on-a-chip technology
- Nodes interconnected as 64x32x32 three-dimensional torus
  - Easy to build large systems, as each node connects only to six nearest neighbors – full routing in hardware
  - Bisection bandwidth per node is proportional to $n^2/n^3$
  - Auxiliary networks for I/O and global operations
- Applications consist of multiple processes with message passing
  - Strictly one process/node
  - Minimum OS involvement and overhead

- 65,536 dual-processor compute nodes
  - 700MHz IBM PowerPC 440 processors
  - 512 MB memory per compute node, 16 TB in entire system.
  - 800 TB of disk space
- 2,500 square feet
Supercomputer Peak Performance

Doubling time = 1.5 yr.
RAS (Reliability, Availability, Serviceability)

- System designed for RAS from top to bottom
  - System issues
    - Redundant bulk supplies, power converters, fans, DRAM bits, cable bits
    - Extensive data logging (voltage, temp, recoverable errors ... ) for failure forecasting
    - Nearly no single points of failure
  - Chip design
    - ECC on all SRAMs
    - All dataflow outside processors is protected by error-detection mechanisms
      - Access to all state via noninvasive back door
  - Low power, simple design leads to higher reliability
  - All interconnects have multiple error detections and correction coverage
Conclusions

• You need to know the big picture at least
  – number of cores and SMT capability
  – Interconnection
  – Memory hierarchy
  – What is available to software and what is not

• Actual performance of program can be a complicated function of the architecture
  – Slight changes in the architecture or program change the performance significantly

• The art of delegation
  – What to do at user level and what to leave for the compiler, OS, and runtime