Plan for Today

• Announcement
  – No lecture next Wednesday (Thanksgiving holiday)
  – Take Home Final Exam
    • Available Dec 7
    • Due via email to Patrick by Dec 14, 4:00PM

• Questions

• Today’s discussion
  – Project demo details
  – Memory system case studies
Project Demos

• Demo logistics
  – Create a demo screencast, no more than 20 mins long
  – Upload to YouTube (you may need two videos due to the 10 min limit)
  – Submit YouTube links to discussion site by 4pm Monday 11/23

• Demo content: show the following (if you can’t, explain why)
  1. Start a simulation. How do you load programs and data? How have you configured your caches?
  2. Execute one sort program, and show how you detect a hazard.
  3. Measure the performance of one instance of sort running alone. How many cycles did it take?
  4. Run 2 sort instances together or, if that’s not possible, run sort and stats together. How many cycles did they take to complete?
  5. Change your input data, your cache miss rate, and run sort alone again. How many cycles did it take?
  6. Explain the final division of labor in your group.
Memory System Case Studies

• Compaq’s Alpha 21264
• Intel’s IXP2800
• Sony/Toshiba/IBM Cell Processor
Alpha 21264

• Application: servers

• Out-of-order superscalar processor
  – Fetches 4 ints/cycle
  – Executes up to 6 insts/cycle
  – 48 bit virtual address and 44 bit physical address
    (or 43 and 41, resp.)
    • Lower half for memory addresses, upper half for I/O addresses
- Instruction cache
  - 64KB, 2-way, Virtually tagged and indexed
- Data cache
  - 64KB, 2-way, Physically tagged, virtually indexed
- 128 Entry TLBs
- Off-chip
  - 1MB-16MB, DM
Performance

- Vs. 21164
  - In-order
  - Smaller L1 & L2 (on-chip)
  - Similar clock rate
Intel IXP2800

• Application: networking
• Novel aspects
  – Chip-multiprocessor
    • XScale: ARM-compatible core
    • Microengines: 16 embedded RISC-like cores
      – Each supports 8 thread contexts in hardware
      – Each one has its own local data and instruction memory
  – Multiple, heterogeneous memory interfaces
  – Multiple, heterogeneous I/O interfaces
  – Great programmer flexibility and responsibility
IXP2800 Physical Organization

- XScale runs at 700 MHz
- Microengines run at 1.4 GHz
XScale Components

- Notable
  - No floating point
  - Non-blocking data cache
  - DSP extensions
  - Mini-data cache holds a “page”
Microengine Organization

- 4K instruction store
- 640 word local memory
- 16 entry CAM
- Registers
  - 2x128 general purpose
  - 2x128 for static memory access
  - 2x128 for dynamic memory access
  - 128 next-neighbor (pipelined applications)
- Multithreading
  - Thread scheduling is non-preemptive, round-robin
  - Threads do not have to wait for memory operations to complete before continuing
Consequences

• **Homogenous structure**
  – Eases programmer’s task
  – Makes performance unpredictable but good

• **Heterogeneous structure**
  – Complicates programmer’s task
  – Makes performance predictable

• **Choice is application-specific**
  – Although architects have historically preferred homogenous structures
Cell Processor

- Heterogeneous Multiprocessor architecture
- Developed by Sony, Toshiba and IBM
- Design and first implementation: 2001-2005
- First application: PlayStation 3 game console
  - Candidate platform for future applications having high parallelism and floating point operations
- Focus
  - Performance/watt
  - Peak throughput over programming simplicity
- Great programmer flexibility and responsibility
Cell Processor Basic

• 2 processor types:
  – Power Processor Element – **PPE** (1/cell)
    • General purpose Power Architecture core
      – Modest performance
      – Controller
  – Synergistic Processor Element – **SPE** (8/cell)
    • Coprocessing elements
      – Accelerate dedicated computations, such as multimedia and vector processing
      – Handle most of the computational workload

• **XDR DRAM** (extreme data rate)
  – High performance memory interface
Cell logical organization

Power Processor Element (PPE)
(64 bit PowerPC with VMX)

I/O Controller

I/O Controller

Memory Controller

Memory Controller

RAM

RAM

SPE 1

SPE 2

SPE 3

SPE 4

SPE 5

SPE 6

SPE 7

SPE 8

Dual "configurable"
High speed I/O
channels

(76.8 GBytes per
second in total)

Dual 12.8 GByte per
second memory busses
give Cell huge memory
bandwidth. (25.6 GBytes
per second in total)

EIB (Element Interconnect Bus)
is the internal communication system.

© Nicholas Blachford 2005
Power Processor Element (PPE)

- 64-bit Power Architecture processor
- Dual-issue
- Dual-threaded
- In-order
- Caches:
  - 32KB L1 I- and D-cache
  - 512KB L2 unified cache
- Support for VMX vector instructions (SIMD)
- Runs @ 3.2 GHz
Synergistic Processor Element (SPE)

- 128-bit vector unit
  - in-order execution
- 256 KB Local store (I and D)
  - Visible to PPE
- 128 x 128-bit registers
- In a clock cycle
  - 16 8-bit, 8 16-bit, 4 32-bit int, 4 32-bit fp ops
  - Memory operation
- @ 3.2GHz
  - 25.6 GFLOPS single precision
  - 14 GFLOPS double precision
- 64-bit system memory accessed via DMA ops
Element Interconnect Bus - EIB

• Communication bus interconnecting PPE, SPEs, I/O interface, memory controller

• Circular ring
  – 4 16B-wide unidirectional channels
  – Up to 3 parallel transaction/channel
  – One 16B read and one 16B write port/participant

• Runs @ ½ system frequency
  – Peak \textit{instantaneous} EIB bandwidth: 96B/clock $\rightarrow$ 307.2 GB/s @ 3.2GHz (system)
  – Limitations on I/O and memory controller: $\sim$200GB/s
Consequences

• Very powerful architecture
• Programming and deployment complexity
  – On SPE:
    • Statically parallelize code for SIMD execution
    • No cache
    • Plan DMAs to memory and I/O
  – System-wise:
    • Split PPE and SPEs tasks
    • SPEs logical operation model
      – All SPEs perform same operations on different data
      – Pipelining
      – Hybrid scheme
    • EIB, I/O and Memory interface utilization
Assignment

• Readings
  – For Monday
    • Project Demonstrations
  – For Wednesday (No lecture; Thanksgiving)
    • None
  – For Monday (11/30)
    • H&P: sections 6.1-6.2