EECS 252 Graduate Computer Architecture

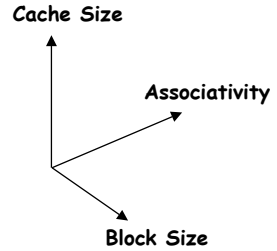
Lec 5 – Projects + Prerequisite Quiz

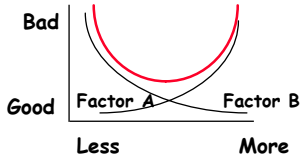
David Patterson
Electrical Engineering and Computer Sciences
University of California, Berkeley

http://www.eecs.berkeley.edu/~pattrsn
http://www-inst.eecs.berkeley.edu/~cs252

Review from last lecture #1/3: The Cache Design Space

- Several interacting dimensions
  - cache size
  - block size
  - associativity
  - replacement policy
  - write-through vs write-back
  - write allocation
- The optimal choice is a compromise
  - depends on access characteristics
    - » workload
    - » use (I-cache, D-cache, TLB)
  - depends on technology / cost
- Simplicity often wins





Review from last lecture #2/3: Caches

- The Principle of Locality:
  - Program access a relatively small portion of the address space at any instant of time.
    - » Temporal Locality: Locality in Time
    - » Spatial Locality: Locality in Space
- Three Major Categories of Cache Misses:
  - Compulsory Misses: sad facts of life. Example: cold start misses.
  - Capacity Misses: increase cache size
  - Conflict Misses: increase cache size and/or associativity. Nightmare Scenario: ping pong effect!
- Write Policy: Write Through vs. Write Back
- Today CPU time is a function of (ops, cache misses) vs. just f(ops): affects Compilers, Data structures, and Algorithms

Review from last lecture #3/3: TLB, Virtual Memory

- Page tables map virtual address to physical address
- TLBs are important for fast translation
- TLB misses are significant in processor performance
  - funny times, as most systems can't access all of 2nd level cache without TLB misses!
- Caches, TLBs, Virtual Memory all understood by examining how they deal with 4 questions:
  1) Where can block be placed?
  2) How is block found?
  3) What block is replaced on miss?
  4) How are writes handled?
- Today VM allows many processes to share single memory without having to swap all processes to disk; today VM protection is more important than memory hierarchy benefits, but computers insecure

Problems with Sea Change

1. Algorithms, Programming Languages, Compilers, Operating Systems, Architectures, Libraries, … not ready for 1000 CPUs / chip
2. Software people don't start working hard until hardware arrives
   - 3 months after HW arrives, SW people list everything that must be fixed, then we all wait 4 years for next iteration of HW/SW
3. How get 1000 CPU systems in hands of researchers to innovate in timely fashion on in algorithms, compilers, languages, OS, architectures, … ?
4. Skip the waiting years between HW/SW iterations?

Build Academic MPP from FPGAs

- As ~ 25 CPUs fit in Field Programmable Gate Array, 1000-CPU system from ~ 40 FPGAs?
  - 16 32-bit simple "soft core" RISC at 150MHz in 2004 (Virtex-II)
  - FPGA generations every 1.5 yrs; ~2X CPUs, ~1.2X clock rate
- HW research community does logic design ("gate shareware") to create out-of-the-box, MPP
  - E.g., 1000 processor, standard ISA binary-compatible, 64-bit, cache-coherent supercomputer @ 200 MHz/CPU in 2007
  - RAMPants: Arvind (MIT), Krste Asanovíc (MIT), Derek Chiou (Texas), James Hoe (CMU), Christos Kozyrakis (Stanford), Shih-Lien Lu (Intel), Mark Oskin (Washington), David Patterson (Berkeley, Co-PI), Jan Rabaey (Berkeley), and John Wawrzynek (Berkeley, PI)
- "Research Accelerator for Multiple Processors"

Characteristics of Ideal Academic CS Research Supercomputer?

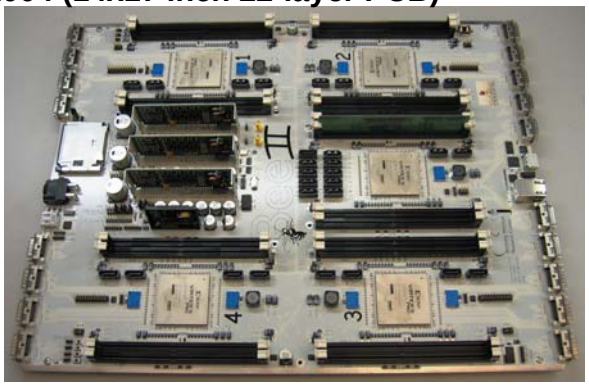
- Scale – Hard problems at 1000 CPUs
- Cheap – 2006 funding of academic research
- Cheap to operate, Small, Low Power – $ again
- Community – share SW, training, ideas, …
- Simplifies debugging – high SW churn rate
- Reconfigurable – test many parameters, imitate many ISAs, many organizations, …
- Credible – results translate to real computers
- Performance – run real OS and full apps, results overnight

Why RAMP Good for Research MPP?

	SMP	Cluster	Simulate	RAMP
Scalability (1k CPUs)	C	A	A	A
Cost (1k CPUs)	F ($40M)	C ($2-3M)	A+ ($0M)	A ($0.1-0.2M)
Cost of ownership	A	D	A	A
Power/Space (kilowatts, racks)	D (120 kw, 12 racks)	D (120 kw, 12 racks)	A+ (.1 kw, 0.1 racks)	A (1.5 kw, 0.3 racks)
Community	D	A	A	A
Observability	D	C	A+	A+
Reproducibility	B	D	A+	A+
Reconfigurability	D	C	A+	A+
Credibility	A+	A+	F	A
Perform. (clock)	A (2 GHz)	A (3 GHz)	F (0 GHz)	C (0.1-.2 GHz)
GPA	C	B-	B	A-

RAMP 1 Hardware

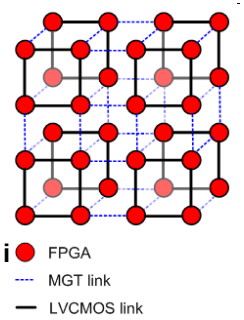
- **Completed Dec. 2004 (14x17 inch 22-layer PCB)**
- **Module:**
  - **5 Virtex II FPGAs, 18 banks DDR2-400 memory, 20 10GigE conn.**
  - **Administration/ maintenance ports:**
    - » **10/100 Enet**
    - » **HDMI/DVI**
    - » **USB**
  - **~$4K in Bill of Materials (w/o FPGAs or DRAM)**



BEE2: Berkeley Emulation Engine 2

By John Wawrzynek and Bob Brodersen with students Chen Chang and Pierre Droz

Multiple Module RAMP 1 Systems



- **8 compute modules (plus power supplies) i 8U rack mount chassis**
- **2U single module tray for developers**
- **Many topologies possible**
- **Disk storage: via disk emulator + Network Attached Storage**

Quick Sanity Check

- **BEE2 uses old FPGAs (Virtex II), 4 banks DDR2-400/cpu**
- **16 32-bit Microblazes per Virtex II FPGA, 0.75 MB memory for caches**
  - **32 KB direct mapped Icache, 16 KB direct mapped Dcache**
- **Assume 150 MHz, CPI is 1.5 (4-stage pipe)**
  - **I$ Miss rate is 0.5% for SPECint2000**
  - **D$ Miss rate is 2.8% for SPECint2000, 40% Loads/stores**
- **BW need/CPU = 150/1.5*4B*(0.5% + 40%*2.8%) = 6.4 MB/sec**
- **BW need/FPGA = 16*6.4 = 100 MB/s**
- **Memory BW/FPGA = 4*200 MHz*2*8B = 12,800 MB/s**
- **Plenty of room for tracing, …**

RAMP Development Plan

1. **Distribute systems internally for RAMP 1 development**
   - **Xilinx agreed to pay for production of a set of modules for initial contributing developers and first full RAMP system**
   - **Others could be available if can recover costs**
2. **Release publicly available out-of-the-box MPP emulator**
   - **Based on standard ISA (IBM Power, Sun SPARC, …) for binary compatibility**
   - **Complete OS/libraries**
   - Locally modify RAMP as desired
3. **Design next generation platform for RAMP 2**
   - **Base on 65nm FPGAs (2 generations later than Virtex-II)**
   - **Pending results from RAMP 1, Xilinx will cover hardware costs for initial set of RAMP 2 machines**
   - Find 3rd party to build and distribute systems (at *near-cost*), open source RAMP gateware and software
   - Hope RAMP 3, 4, … self-sustaining

- **NSF/CRI proposal pending to help support effort**
  - **2 full-time staff (one HW/gateware, one OS/software)**
  - **Look for grad student support at 6 RAMP universities from industrial donations**

RAMP Milestones

Name	Goal	Target	CPUs	Details
Red (S.U.)	Get Started	1Q06	8 PowerPC 32b hard cores	Transactional memory SMP
Blue (Cal)	Scale	3Q06	1024 32b soft (Microblaze)	Cluster, MPI
White	Features			Cache coherent, shared address, deterministic, debug/monitor, commercial ISA
1.0		2Q06?	64 hard PPC	
2.0		3Q06?	128? soft 32b	
3.0		4Q06?	64? soft 64b	
4.0		1Q07?	Multiple ISAs	
2.0	Sell	2H07?	4X '04 FPGA	New '06 FPGA, new board

the stone soup of architecture research platforms

Wawrzynek — Hardware

Chiou — Glue-support

Patterson — I/O

Hoe — Coherence

Kozyrakis — Monitoring

Asanovic — Cache

Oskin — Net Switch

Arvind — PPC

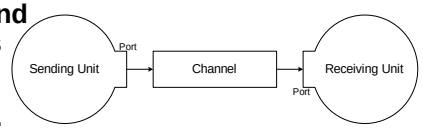
Lu — x86

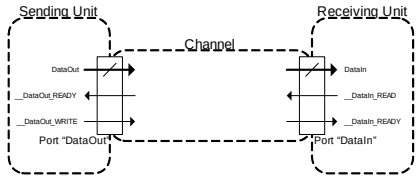
Gateware Design Framework

- Insight: almost every large building block fits inside FPGA today
  - what doesn't is between chips in real design
- Supports both cycle-accurate emulation of detailed parameterized machine models and rapid functional-only emulations
- Carefully counts for *Target Clock Cycles*
- Units in any hardware design language (will work with Verilog, VHDL, BlueSpec, C, ...)
- RAMP Design Language (RDL) to describe plumbing to connect units in

Gateware Design Framework

- Design composed of units that send messages over channels via ports
- Units (10,000 + gates)
  - CPU + L1 cache, DRAM controller….
- Channels (~ FIFO)
  - Lossless, point-to-point, unidirectional, in-order message delivery…





RAMP FAQ

- Q: How will FPGA clock rate improve?
- A1: 1.1X to 1.3X / 18 months
  - Note that clock rate now going up slowly on desktop
- A2: Goal for RAMP is system emulation, not to be the real system
  - Hence, value accurate accounting of target clock cycles, parameterized design (Memory BW, network BW, …), monitor, debug vs. clock rate
  - Goal is just fast enough to emulate OS, app in parallel

RAMP FAQ

- Q: What about power, cost, space in RAMP?
- A: Using very slow clock rate, very simple CPUs in a very large FPGA (RAMP blue)
  - 1.5 watts per computer
  - $100-$200 per computer
  - 5 cubic inches per computer

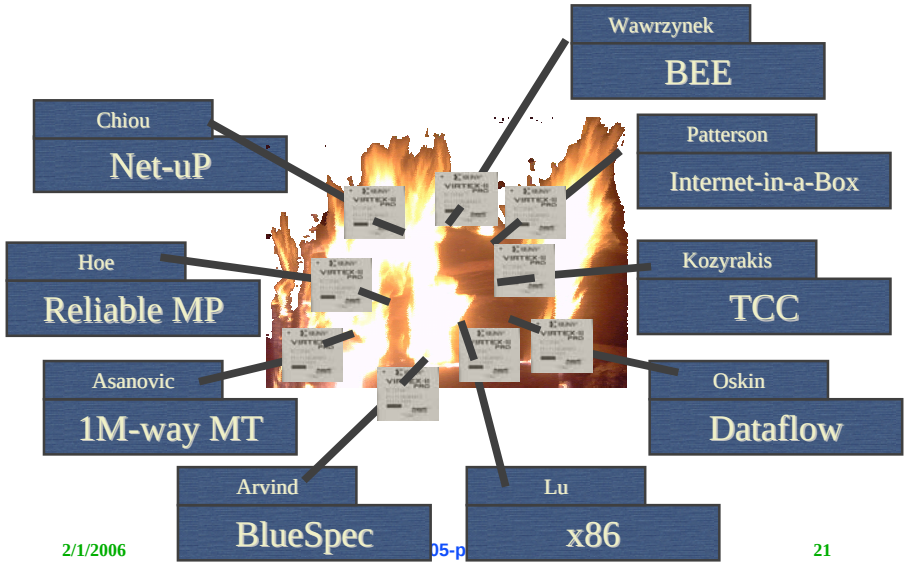
RAMP FAQ

- Q: But how can lots of researchers get RAMPs?
- A1: Official plan is RAMP 2.0 available for purchase at low margin from 3rd party vendor
- A2: Single board RAMP 2.0 still interesting + FPGA generation 2X CPUs/18 months
  - RAMP 2.0 two generations later than RAMP 1.0, so 256 simple CPUs per board?

RAMP Status

- ramp.eecs.berkeley.edu
- Sent NSF infrastruture proposal August 2005
- Biweekly teleconferences (since June 05)
- IBM, Sun donating commercial ISA, simple, industrial-strength, CPU + FPU
- Technical report, RAMP Design Language
- RAMP 1/RDL short course/board distribution in Berkeley for 40 people @ 6 schools Jan 06
- 1 Day RAMP retreat with 12 industry visitors
- Berkeley style retreats 6/06, 1/07, 6/07

RAMP uses (internal)



Multiprocessing Watering Hole

Parallel file system Dataflow language/computer Data center in a box
Thread scheduling Security enhancements Internet in a box
Multiprocessor switch design Router design Compile to FPGA
Fault insertion to check dependability Parallel languages

- **Killer app: All CS Research, Ind. Advanced Development**
- **RAMP attracts many communities to shared artifact**
  ⇒ Cross-disciplinary interactions
  ⇒ Accelerate innovation in multiprocessing
- **RAMP as next Standard Research Platform?**
  (e.g., VAX/BSD Unix in 1980s)

Supporters *(wrote letters to NSF)*

- **Gordon Bell (Microsoft)**
- **Ivo Bolsens (Xilinx CTO)**
- **Norm Jouppi (HP Labs)**
- **Bill Kramer (NERSC/LBL)**
- **Craig Mundie (MS CTO)**
- **G. Papadopoulos (Sun CTO)**
- **Justin Rattner (Intel CTO)**
- **Ivan Sutherland (Sun Fellow)**
- **Chuck Thacker (Microsoft)**
- **Kees Vissers (Xilinx)**
- **Doug Burger (Texas)**
- **Bill Dally (Stanford)**
- **Carl Ebeling (Washington)**
- **Susan Eggers (Washington)**
- **Steve Keckler (Texas)**
- **Greg Morrisett (Harvard)**
- **Scott Shenker (Berkeley)**
- **Ion Stoica (Berkeley)**
- **Kathy Yelick (Berkeley)**

RAMP Participants: Arvind (MIT), Krste Asanovíc (MIT), Derek Chiou (Texas), James Hoe (CMU), Christos Kozyrakis (Stanford), Shih-Lien Lu (Intel), Mark Oskin (Washington), David Patterson (Berkeley), Jan Rabaey (Berkeley), and John Wawrzynek (Berkeley)

RAMP Summary

- **RAMP accelerates HW/SW generations**
  - **Trace anything, Reproduce everything, Tape out every day**
  - **Emulate anything: Massive Multiprocessor, Distributed Computer,…**
  - **Clone to check results (as fast in Berkeley as in Boston?)**
- **Carpe Diem: Researchers need it ASAP**
  - **FPGA technology is ready today, and getting better every year**
  - **Stand on shoulders vs. toes: standardize on design framework, Berkeley effort on FPGA platforms (BEE, BEE2) by Wawrzynek et al**
  - **Architects get to immediately aid colleagues via gateware**
- **"Multiprocessor Research Watering Hole" ramp up research in multiprocessing via standard research platform ⇒ hasten sea change from sequential to parallel computing**

CS 252 Projects

- RAMP meetings Wednesdays 3:30-4:30
  - February 1st (today) and February 8 meetings will be held in Alcove 611 (sixth floor - Soda Hall)
  - February 15th - May 17th in 380 Soda Hall
  - Big cluster, DP fl. Pt., Software, workload generation, DOS generation, …
- Other projects from your own research?
- Other ideas:
  - How fast is Niagara (8 CPUs, each 4-way multithreaded); run unpublished benchmarks
  - How fast is Mac on x86 binary translation?

CS252: Administrivia

Instructor: Prof. David Patterson

Office: 635 Soda Hall, pattrsn@eecs, Office Hours: Tue 4-5

(or by appt. Contact Cecilia Pracher; cpracher@eecs)

T. A: Archana Ganapathi, archanag@eecs

Class: M/W, 11:00 - 12:30pm 203 McLaughlin (and online)

Text: *Computer Architecture: A Quantitative Approach, 4th Edition* (Oct, 2006), Beta, distributed free provided report errors

Wiki page : vlsi.cs.berkeley.edu/cs252-s06

Wed 2/1: Great ISA debate (4 papers) + 30 minute Prerequisite Quiz

1. Amdahl, Blaauw, and Brooks, "Architecture of the IBM System/360." *IBM Journal of Research and Development*, 8(2):87-101, April 1964.
2. Lonergan and King, "Design of the B 5000 system." *Datamation*, vol. 7, no. 5, pp. 28-32, May, 1961.
3. Patterson and Ditzel, "The case for the reduced instruction set computer." *Computer Architecture News*, October 1980.
4. Clark and Strecker, "Comments on 'the case for the reduced instruction set computer'," *Computer Architecture News*, October 1980.

4 Papers

- Read and Send your comments
  - email comments to archanag@cs AND pattrsn@cs by Friday 10PM; posted on Wiki Saturday
- Read, comment on wiki before class Monday
- Be sure to address:
- B5000 (1961) vs. IBM 360 (1964)
  - What key different architecture decisions did they make?
    - » E.g., data size, floating point size, instruction size, registers, …
  - Which largely survive to this day in current ISAs? In JVM?
- RISC vs. CISC (1980)
  - What arguments were made for and against RISC and CISC?
  - Which has history settled?

Computers in the News

- The *American Competitiveness Initiative* commits $5.9 billion in FY 2007, and more than $136 billion over 10 years, to increase investments in research and development (R&D), strengthen education, and encourage entrepreneurship and innovation.
- NY Times today: "In an echo of President Dwight D. Eisenhower's response after the United States was stunned by the launching of Sputnik in 1957, Mr. Bush called for initiatives to deal with a new threat: intensifying competition from countries like China and India. He proposed a substantial increase in financing for basic science research, called for training 70,000 new high school Advanced Placement teachers and recruiting 30,000 math and science professionals into the nation's classrooms."

SOTU Transcript

- "And to keep America competitive, one commitment is necessary above all. We must continue to lead the world in human talent and creativity. Our greatest advantage in the world has always been our educated, hard-working, ambitious people, and we are going to keep that edge. Tonight I announce the American Competitiveness Initiative, to encourage innovation throughout our economy and to give our nation's children a firm grounding in math and science. "
  - [American Competitiveness Initiative: www.whitehouse.gov/news/releases/2006/01/20060131-5.html]

SOTU Transcript

"First, I propose to double the federal commitment to the most critical basic research programs in the physical sciences over the next 10 years. This funding will support the work of America's most creative minds as they explore promising areas such as nanotechnology and supercomputing and alternative energy sources.

"Second, I propose to make permanent the research and development tax credit to encourage bolder private-sector initiative in technology. With more research in both the public and private sectors, we will improve our quality of life and ensure that America will lead the world in opportunity and innovation for decades to come."

SOTU Transcript

- "Third, we need to encourage children to take more math and science and to make sure those courses are rigorous enough to compete with other nations. We've made a good start in the early grades with the No Child Left Behind Act, which is raising standards and lifting test scores across our country. Tonight I propose to train 70,000 high school teachers to lead Advanced Placement courses in math and science, bring 30,000 math and science professionals to teach in classrooms and give early help to students who struggle with math, so they have a better chance at good high-wage jobs. If we ensure that America's children succeed in life, they will ensure that America succeeds in the world."

SOTU Transcript

"Preparing our nation to compete in the world is a goal that all of us can share. I urge you to support the American Competitiveness Initiative, and together we will show the world what the American people can achieve."