Microsoft® Research
Faculty Summit 2010
The Dinner Cruise will depart from the Kirkland Dock at 6:30PM.
Buses from here are going straight to the boat without stopping at the hotel.
Remember to take your coats and bags with you off of the bus and onto the boat; There is luggage check on the boat.
At the end of the cruise, buses will return to the Conference Center and the hotels.
Charles P. Thacker
ACM’s A.M. Turing Award Winner
Microsoft Research

Faculty Summit 2010

Chuck Thacker
Technical Fellow
RARE: Rethinking Architecture
Research and Education

Chuck Thacker (cthacker@microsoft.com)
Microsoft Research Silicon Valley
12 July 2010
Influential D. Patterson columns:

Seven Reason to Shave Your Head and Three Reasons Not to: The bald truth. Commun. ACM. 49, (4): 31-32 (April, 2006)

Alas, sometimes, Dave is wrong...
Points in Dave’s second CACM column:

• Use tools and libraries
  – “For many CS courses, a dramatic change would simply be if students first wrote a clear specification and then built software using modern tools and software components”.

• Embrace Parallelism.

• Join the OS movement.

• Build your own supercomputer.
  – Described RAMP, which led to my latest projects.
BEE3 (‘06 – ’09)
BEE3 Subsystems

User1
5VLXT
User2
5VLXT
User3
5VLXT
User4
5VLXT
DDR2 DIMM0
DDR2 DIMM1
DDR2 DIMM2
DDR2 DIMM3
DDR2 DIMM0
DDR2 DIMM1
DDR2 DIMM2
DDR2 DIMM3

QSH-DP-040

PCI-E 8X

QSH-DP-040

PCI-E 8X

QSH-DP-040

PCI-E 8X

QSH-DP-040

PCI-E 8X

QSH-DP-040

PCI-E 8X

QSH-DP-040

PCI-E 8X
BEE3 Program

- MSR, UCB did the detailed specifications
- MSR engaged Celestica for the implementation.
  - Better than burning out grad students
  - Pros can do the job better and faster
  - Resulting board worked the first time (unprecedented for me)
- MS licensed the design to a 3rd party company (BEECube) to build, sell, and support systems.
  - ~75 shipped to date
  - MSR supplied some basic IP (DRAM controller)
- Means that both academics and industrial customers can buy them.
  - Not the case if NSF/DARPA funded
Beehive (‘09 – present)

• An FPGA-based many-core system
  – 13 RISC cores (100 MHz)
  – 2 GB DDR2 DRAM controller
  – Display controller
  – 1 Gb Ethernet controller
  – ~6K lines of Verilog
  – Students can understand and modify it, using only the basic Xilinx tools (ISE, ChipScope). No high-powered CAD needed.

• A software tool chain
  – C compiler, assembler, linker

• A small but growing set of libraries for frequently needed things

• Licensed for academic research use
Beehive on Xilinx XUPV5

Academic price: $750
Core local IO subsystem

- AQ
- WQ
- Multiplier
  - Device 1
- RS232
  - Device 0
- DCache
  - Device 3
- Messenger
  - Device 4
- Lock Unit
  - Device 5

AQ[2:0], AQ[31]

RQ

To CPU
Beehive instructions

- 32-bit instructions, 32-bit registers
- Rw = Ra Function Rb Op Count
  - Function: add, sub, logic
  - Op: Shifts
- Variants for Jumps, Memory accesses
- Support for constants

<table>
<thead>
<tr>
<th>Ra</th>
<th>Rw</th>
<th>Count</th>
<th>Rb</th>
<th>Const</th>
<th>Function</th>
<th>??</th>
<th>Op</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>31</td>
<td>27</td>
<td>22</td>
<td>17</td>
<td>16</td>
<td></td>
</tr>
</tbody>
</table>

- Op
  - 6: No shift, CountConst, AQW :=
  - 7: No shift, CountConst, AQR :=
  - 4: ARSH (Count)
  - 5: A + B
  - 3: A & ~B
  - 2: A ^ B
  - 1: A | ~B
  - 2: A | B
  - 3: A ^ B
  - 4: LLI (Load Link Immediate)
  - 5: RSH (Count)
  - 6: LSH (Count)
  - 7: RCY (Count)
  - 8: LCY (Count)
  - 9: RwCountConst, jump if out < 0
  - 10: RwCountConst, jump if out = 0
  - 11: RwCountConst, jump if out > 0
  - 12: RwCountConst, jump Always
  - 13: RwCountConst, jump if out # 0
  - 14: RwCountConst, jump if no ALU carry

Const ??
Beehive Ring Interconnect

- All wires are local
- Passes through:
  - Each core
  - Display controller
  - Ethernet controller
  - DRAM controller
- “Train” contains token + contents
- Each node can modify/append to the train
Architectural Curiosities

- No coherent memory
- No byte addressing
  - We fudge this
- No protection
  - We may add this
- No VM
- No kernel mode
Beehive uses: Education

• Architecture lab courses
  – Boards are inexpensive, so every student can have one.
  – Verilog is simple enough for students to make changes, try new things.
    • Like Stanford’s NetFPGA
  – Tool chain and libraries are familiar
    • GCC, make, ...
  – Initial results are promising
Beehives at MIT

Two-week IAP course in January, full-semester course in Fall ’10.

Xilinx donated 20 boards. Students modified the Verilog and successfully tested their changes.

http://projects.csail.mit.edu/beehive
Beehive uses: Research

• Forget shared memory. Use message passing
• Transactional memory.
  – Allows apples-to-apples comparison with Monitors/CVs
  – Gets coherence where you need it.
• Do we really need...
  – Coherent shared memory?
  – Interrupts?
  – VM?
  – An OS?
Beehive Non-goals

• Emulate an existing ISA
  – Modern ISAs are *not* simple
  – Can’t do direct comparisons, only A/B experiments.

• Run Linux or other extant OSes
  – Small test programs, benchmarks
  – Barreelfish is the exception

• Have high performance
  – Can’t have this with FPGAs anyway.
  – Only needs to be fast enough to run programs much faster than a simulator.
Next steps

• Port (back) to BEE3
• Use in our own research
• TM (MSR SVL)
• Barreelfish (MSR Cambridge, ETH)
• Make it more widely available for academic use (email me)
Q&A
Microsoft® Research
Faculty Summit 2010
The Dinner Cruise will depart from the Kirkland Dock at 6:30PM.

Buses from here are going straight to the boat without stopping at the hotel.

Remember to take your coats and bags with you off the bus and onto the boat; There is no luggage check on the boat.

At the end of the cruise, buses will return to the Conference Center and the hotels.