Is the Multi-Core Roadmap going to Live Up to its Promises?

--A Sobering Point of View--

Michel Dubois

Department of Electrical Engineering

University of Southern California

Panel IPDPS 2007

USC
VITERBI
SCHOOL OF
ENGINEERING

DISCLAIMER

- I am not "anti-CMPs"
- Of course CMPs are useful and will help
  - For a while

How swiftly things change

- Currently Parallel architectures are simple extensions of uniprocessors, mostly used for throughput processing
- Most researchers in parallel architectures have moved into other fields
- Funding for parallel architectures has all but dried out
- Number of papers on multiprocessors in flagship architecture conferences has dwindled to a trickle
- Conferences dedicated to parallel processing (ICPP, IPPS…) have taken a second role/had to consolidate

USC
VITERBI
SCHOOL OF
ENGINEERING

How swiftly things change

- These bullets were part of a keynote I presented in May 2004 at IPDPS titled: "Are we entering the golden age of parallel processing?"

USC
VITERBI
SCHOOL OF
ENGINEERING

Fast Forward Today--2007

- From an academic standpoint point it is almost unthinkable today to submit an architecture paper or proposal that does not have the acronym CMP all over it.

- Industry is embracing CMPs currently, mostly because of power issues

- GHz of announced chips are barely inching up

USC
VITERBI
SCHOOL OF
ENGINEERING

Some Observations

- The demise of parallel architecture research back in the beginning of the 90's mainly came from the emergence of the "killer micro"

- We may be entering the era of the "killer CMP"

- The roadmap for the killer micro was a "slam dunk" because it used the traditional computing model, one that had been honed and widely accepted.

- By contrast, the CMP roadmap is drafted OUT of Technological NECESSITY
  - i.e., we don't know what else to do
  - So we tell the market: "take it and shove it"
  - Things could again change swiftly

Scary Times

- For academia, the emergence of CMPs is a bonanza, at least for the next few years
- For industry, which is pushing CMP technology, this must be very scary times
    - Past Experience with Parallel Programming and scalability
        - Market acceptance
        - Scalability???
    - The pyramid of sales
        - At the top, very expensive machines sell few units
        - The bulk of the market is at the bottom
        - Systems with large number of threads only benefit the machines at the top of the pyramid
    - Then, above all, there is the looming closing of the technological window for CMOS

Closing Technological Window

- Progressive degradation of the technological fabric
  - Discrete effects and uncertainty
    - Heisenberg principle
    - Quantum effects:
  - Transient faults
  - Manufacturing defects
    - The manufacture of circuits becomes less reliable

Closing Technological Window

- The net result is that logic gates won't behave deterministically any more

- Circuit designers are starting to think of gates as probabilistic devices.

- What will we do at the system level with such devices?

- Unless we are able to master the effects of uncertainty, miniaturization will stop being useful, at least for general purpose systems

Closing Technological Window

- To meet Moore's law as applied to computing by doubling the number of threads every other year machines will have to physically become larger and larger
- Not a good prospect for society

The Current Status

- It seems that researchers in technologies dealing with quantum effects live in a totally different world.
- It also seems that the architecture community is ignoring the inevitable outcome
- Is there a "Unified Theory"??
- We may hit a wall much taller than any other wall we have met so far
- This challenge cannot be met in an evolutionary fashion
    - Wait for technology to save us, as we always do
- Do we have the time for a revolution?

What Should We Do?

- Ride the current wave of CMP craze
  - How to design MPs
  - Rehash of an old problem, with different parameters
  - Industry can do it
  - Seems trivial compared to the looming challenges
- Attack hard problems in CMPs
  - How to speed up general purpose code
  - Using assist threads for example, or any other hardware support
  - PB is software support
  - Could there be anything new out there?
  - Do we have time??

What Should We Do?

- Attack very hard problems in CMPs
  - Software problem
    - Could there possibly be anything new out there????
    - Scalable parallelization of important software
  - How to use plentiful threads to tolerate (at least for a while) the increase in uncertainty??
  - New computing/execution paradigm
  - Do we have time??

- Attack extremely hard (impossible??) problems such as how to deal with uncertainty at the system level??
  - Is there a meeting point with quantum research??

Alternative Ideas

- Alternative to CMPs: Heterogeneous chips with various units which may possibly be configurable.
  - Could be hellish to program and orchestrate
  - Use of "macro-instructions"
  - Still limited by integration
- Accept the end of general-purpose computing
  - All environments are special purpose
  - Big, useful applications can deal with uncertainty

Implications

- We can work on short term issues and ride the current market-driven CMP wave or we can decide to work to solve BIG problems with big societal impact.
- The IT revolution has brought about huge rewards to society
- The societal impact of abruptly stalling its progress would be huge and could strike us much sooner than we think.
- CMPs won't solve the longer term problems
  - In fact in the grand scheme of things it could be that in retrospect, CMPs were a big distraction

USC VITERBI SCHOOL OF ENGINEERING