Workshops on Advancing Computer Architecture Research (ACAR)

Organizers:

Mark Oskin (University of Washington)
Josep Torrellas (University of Illinois)

Contributors: Mark Oskin, Josep Torrellas, Sarita Adve, George Almasi, Luis Ceze, Almadena Chtchelkanova, Chita Das, John Davis, Sandhya Dwarkadas, Lieven Eeckhout, Bill Feiereisen, William Harrod, Daniel Jimenez, Mark Hill, Jon Hiller, Sampath Kannan, Krishna Kant, Martha Kim, Christos Kozyrakis, James Larus, Margaret Martonosi, Richard Murphy, Onur Mutlu, Satish Narayanasamy, Kunle Olukotun, Yale Patt, Andrew Putnam, Tim Sherwood, Anand Sivasubramaniam, Kevin Skadron, James Smith, Karin Strauss, Steven Swanson, Dean Tullsen, David Wood, Craig Zilles

Sponsored by Computing Community Consortium (CCC) Computing Research Association (CRA)
Held on February 22-23, 2010 in San Diego and on September 20-21, 2010 in Seattle

http://www.cra.org/ccc/acar.php, http://iacoma.cs.uiuc.edu/acar1/

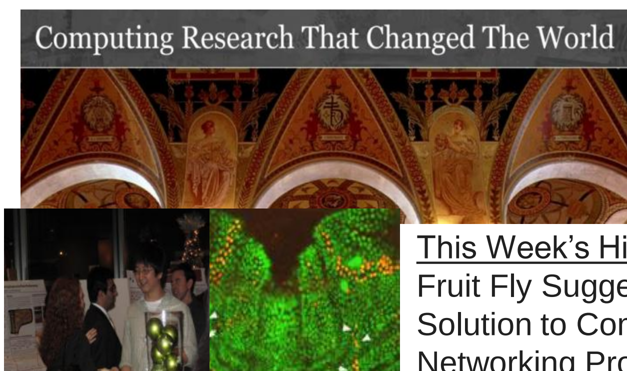
April 2011

Computing Community Consortium

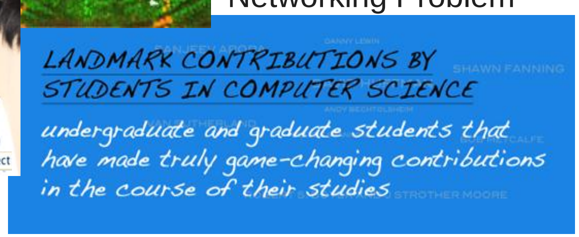
Focused on empowering the computing research community to pursue more audacious visions

A broad-based standing committee of 20 leading U.S. computer scientists

- Housed within the Computing Research Association (CRA), representing >200 U.S. & Canadian academic departments and industrial research labs
- Chair: Ed Lazowska, U-Washington
- Vice-Chair: Susan Graham, UC-Berkeley
- Director: Erwin Gianchandani, CRA [erwin@cra.org]



This Week's Highlight:
Fruit Fly Suggests New Solution to Computer Networking Problem



A multitude of activities:

- **Community-initiated visioning workshops –** bringing researchers together to generate "out-of-the-box" ideas
- **White papers for the White House & others –** short reports to inform policymakers
- **Public relations efforts –** Library of Congress symposiums, Research "Highlight of the Week," CCC Blog
- **Nurturing the next generation of leaders –** Computing Innovation Fellows, "Landmark Student Contributions"

Funded via a Cooperative Agreement with the National Science Foundation

Context of the Visit

- Two CCC-sponsored visioning workshops on “Advancing Computer Architecture Research”:
  - One on parallel computing (Feb 2010); other more general (Sep 2010)
  - Issued community-wide calls for position papers
  - Each workshop had 20-30 attendees
  - Invited funding agencies and industry
  - Generated 2 reports and summary slides

- Purpose of this visit:
  - Make you aware of what emerged from this visioning effort
    - Visions have backing of the broad comp. arch. research community
  - Provide the detailed reports and slides

Executive Summary

- Need for bold research investment in computer architecture/hardware
  - Computer architecture is at the core of the IT revolution
  - We are at a turning point for computer architecture
  - Industry’s emphasis on evolutionary technologies/methods won’t do
  - Danger of missing major societal benefits
  - Slowly losing leadership as a nation?

- Pressing/transformative issues:
  - Extreme scale computing
  - Architectures for programmability
  - Specialized architectures and heterogeneity
  - The end of the road for conventional instruction sets
  - Security and reliability from the hardware up
  - Exploiting emerging technologies

Why Investing in Architecture/Hardware is Essential

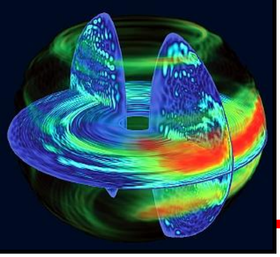
- Computer architecture is a foundation of IT
- Basic ideas laid 50 years ago with modest resources/aspirations
- We have been extending the same bungalow foundation and built skyscrapers on top of it
  - Delivered exponential growth in IT
- We are at a turning point: technology/applications/goals
- Bungalow foundation has reached its limits
  - Industry has turned to largely evolutionary multicore designs
  - Unlikely to go back to exponential capability growth
- Architecture advances have potential for sweeping societal impact
  - Foundation to build mobile to warehouse-scale computers

Turning Point for Computer Architecture

- Technology Drivers:
  - Progressive divergence between real and classical (Denard) scaling
  - Power limits and reduced semiconductor reliability
- Application Drivers:
  - Changing nature of applications and software
  - Growth: mobile/embedded computing and warehouse-scale data center
- Metrics and Goals:
  - Traditionally: exclusively performance-focused
  - Now: reliable, secure, and operating with dynamic power budgets

Example: Why Industry will not Do it

- Rethink architecture from the ground for 100-1000x energy efficiency
  - Circuits: low supply voltage while handling process variations at 8nm
  - Aggressive power and clock gating
  - Computing with 1Kcores/chip: Clustering, heterogeneity, simplicity
  - Architectures to minimize data movement: processing in memory
  - Hardware for fine-grain synchronization
  - Very low power networks
  - New resiliency problems
    - Voltage droops
    - Lightweight checkpointing
  - Stacked DRAM, photonics, new device types
- Do we want the innovation to occur in the US or in East Asia?



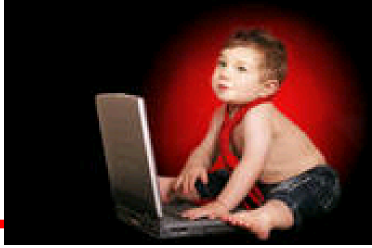
Extreme Scale Computing

- Energy and power consumption are the key limiters to progress
- They require complete rethinking of computer architecture
  - Cloud computing: Utility computing for billions of consumers
  - High-perf. computing: revolutionize science and engineering

- Goals:
  - Improve machines’ energy efficiency 1000x.
  - Minimize cost and power of datacenter infrastructure for typical user
- Research to do:
  - chip and node architecture, interconnect, memory and storage

- ***Transformative:*** Extreme scale computing can provide cheap utility computing to billions of citizens

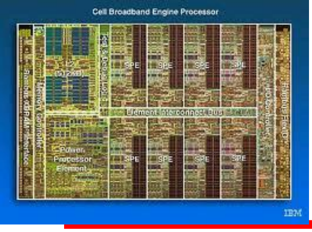
Architectures for Programmability



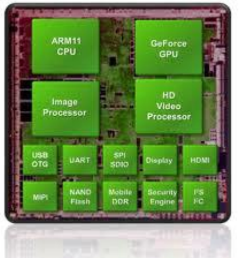
- Performance scaling requires extensive HW changes to exploit paralsm
- Need 1Kcore chip that is high-perf, energy efficient, and programmable

- Goals:
  – Programming for parallel arch. should be as easy as for sequential
  – Maintain Moore's law for performance
  – Eliminate concurrency bugs
- Research to do:
  – Scalable memory/communication fabric, correctness, introspection, resource management

- ***Transformative:*** Enabling programming for the masses of programmers



Specialized Arch. and Heterogeneity



- HW specialization eliminates inefficiencies/overheads of gen-purpose
- Need technologies to develop turn-key specialized computing systems quickly and economically

- Goals:
  - Fully-automated generation of appl-specific hardware cheapily
  - Manufacturing costs should be no greater than for conventional
- Research to do:
  - Heterogeneous designs, static and dynamic reconfigurability, identification of abstractions

- ***Transformative:*** Obtain orders of magnitude improvement in computing perf, perf per Watt, perf per dollar

End-of-Road for Conventional Instruc Sets

- Modern systems are skyscrapers build on bungalow foundation
- Hardware abstractions built 50 years ago:
  - Instruction grain forces processors view apps through narrow window
  - Obfuscates concurrency, intent, knowledge that compiler/user knew

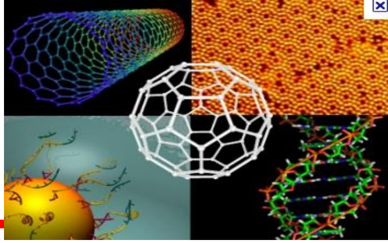
- Goals:
  - Revisit fundamental assumptions of computer architecture
  - Hardware has to know more about the software intent
- Research to do:
  - Specialized virtualization layers, QoS, dataflow operation

- ***Transformative:*** Unlock solutions for important challenges such as security, reliability, energy efficiency, QoS

Secure, Reliable, Predictable from HW Up

- Current architectures are fragile
  - Do not fail gracefully, have poor isolation mechanisms
  - Unsecure by default, impossible to keep a secret
- Goal: machines worthy of the trust we place in them:
  - Verifiably correct in the face of faults and errors
  - Hardware foundations for info containment, privacy, reliability
- Research to do:
  - Strong containment, assurances about info leak, predictability, deterministic hardware
- ***Transformative:*** Orders-magnitude reduction in cost of trustworthy systems will transform what things we trust computers to do.

Exploiting Emerging Technologies



- Emerging technologies offer orders of magnitude improvement
- Researching the architectural use-cases brings science to market faster

- Goal:
  - Shape the directions and provide solutions in emerging technologies
- Research to do:
  - Architectures for quantum computing, synthetic biology, resistive memories, 3D stacking, photonics, nanotubes

- ***Transformative:*** Architecture research enables new, revolutionary technologies to enter the market faster

Funding Recommendations

- Each of these areas should be the nucleus of research programs or be consciously incorporated in mission-oriented programs
- Fund ambitious, larger proposals. Eschew incrementalism
- Fund cross-cutting proposals with teams that are willing to change the architecture
- Coordinate NSF and DARPA/DOE efforts