Bridging the Computation Gap in a Future of Massive Data

Fred Chong
Director, Greenscale Center for Energy-Efficient Computing
Director, Computer Engineering
UC Santa Barbara
Computation Gap

- Optimistic technology scaling assumptions
- “Internet of Things”
- Greg Papadopoulos, keynote IGCC June 2012:
  
  “$1 trillion market for ubiquitous sensors.”
Outline

• Bridging the gap
• Environmental Costs
• Cultural change
Efficiency Gap

- IEE/Kavli Roundtable

- 40X efficiency gap in 13 years

- No single solution
  - Gains needed at many levels of the system

[IEEE Design and Test 2014]
Solutions

• Eliminate Waste
  – Turn stuff off
  – Avoid overprovisioning

• Change the Rules
  – New Technologies
  – Approximate Computing

• Will give several examples
  – About 20% savings each
Barely-Alive Servers

- Turn off microprocessors but allow other servers to use memory
- Decouple load variation from data variation
LogStore: Extending Low-Power Disk Modes

- Random disk writes are energy intensive (require higher speed to meet performance needs)
- *Sequentially logging writes can defer high-speed operation*

[FAST’12]
Targeted Thermoelectric Cooling

- Superlattice layer on microprocessor
- Acts as a Peltier heat spreader targeted at hot spots
- Avoids worst-case provisioning in datacenter-level cooling
Heterogeneous 3D Phase-Change Memory

- Different operating temperatures in 3D stack
- Tailor GST mixture to operating temperature
- 10% memory energy savings
Computational Sprinting

Power & Temperature Response

Max sprint: 3s @ 3.2GHz, 19s @ 1.6GHz

Wenisch, U Mich
Phase-Change Heat Sink

PCM Heat Sink Prototype

- Aluminum foam mesh **filled with Paraffin wax**
  - Relatively form-stable; melting point near \(55^\circ\) C
  - Working on a fully-sealed prototype w/ thermocouples
Deep Memory Hierarchies

- Motivation: Hierarchy is very non-energy-proportional
  - Existing technologies: faster flash, multi-speed & IDP disks
  - New byte-addressable technologies: PCM & STT-RAM
  - Deep hierarchy can improve energy-proportionality

- Recent Progress:
  - Predict data location instead of search
  - Simpler design allows compact table to be recalibrated periodically (22% energy savings)

[IPDPS’14 (Best Paper)]
Memory De-duplication

- 32% avg memory savings on MPI apps
  - 60% max

[IPDPS’11]
Approximation

• Approximate de-duplication
• Approximate computation
  – NPU 3.0X energy savings [Esmailzadeh 13]
• Guided approximation with information flow techniques
3D Beamforming in Datacenters

- Zheng and Zhao with Vahdat at Google
- 60 Ghz links with 2-6 Gbps
- *Flexible BW for burst loads*

---

**A Wireless Way Around Data Center Traffic Jams**

To battle information overload at data centers, researchers are testing a new option: high-frequency wireless links that can help move the data along during crunch times.

[Sigcomm’12]
Datacenter Placement

Datacenter Placement Example

Electricity Rates

Temperature

Cost Breakdowns

Cost of Green Datacenters

[Goiri et al, ICDCS’11]
Bridging the Gap

• 40X in 13 years
• Assuming 20% improvements can be compounded:
  – Need a new idea deployed every 7-8 months!
  – Probably much worse!
Part II: Environmental Costs to Bridging the Gap
Server = SUV

• More precisely:
  – 80 billion terawatt-hr / yr = 6 million SUVs in carbon production (10 mpg, 11K miles/yr)
Google at Columbia River Gorge => hydroelectric power

$30B annual energy bill worldwide

Energy starting to cost more than capital expenditures
Resource Use in Silicon Fabrication

1. 1.6 kilowatt-hrs / cm²
2. 20 liters water / cm²
3. 3.3 billion active cell phone subscriptions
   (212 Billion wireless devices by 2020)
4. ~20 cm² / phone
5. 106 billion kilowatt-hrs (recall that datacenters use 80 billion kwh annually)
Throughput

• 280 Million phones sold / quarter
• Average lifetime of a phone: 1.5-2 yrs
• Old phones sitting in drawers, but throughput of over 1 billion phones / yr
• 32 billion kilowatt-hrs / yr just for uproc
Other Impacts

• 400 billion liters of water
  – 160,000 olympic swimming pools
  – More than double annual global bottled water consumption

• 400 million kg of soil to remediate just the copper (more copper on surface than inside the earth!)
Biodegradable Materials

- Biodegradable plastics
  - Fire retardants are bad
- Organic LEDs / transistors
Microprocessor Reuse?

• Problem: obsolescence resulting from rapid improvements

• Solution: microprocessor food chain

[IEEE Computer ’07]
Example Applications

The BDTImark2000™ is a summary measure of signal processing speed. For more info and scores see www.BDTI.com
Lifetime Energy Savings

- Depends on die-size
  - Die sizes are getting smaller

- Depends on in-use energy consumption
  - Assume 3 hours of use per day
    - .5 W processors probably should be re-used
    - 20 W processor, upgrade!

Graphs showing energy savings over time for .5W and 20W processors with different replacement strategies.
Technical Challenges of Re-Use

• Form Factor
  – Can’t put a Pentium in the space of an 8051

• Battery Life
  – Is adequate power consumption good enough?
  – Voltage scaling

• ISA compatibility
  – Some ISA are more efficient on specific workloads
  – May require extra cycles
    • Erode the efficiency of our re-use strategy
Design for Reuse

- Design for several applications and lifetimes, not just one
- More severe wearout
- Added overhead to support different applications
- Design for easier reprogramming
- Design for easier reclamation and re-tasking
  - form factor, wireless or serial communication

Standard building blocks
Reclamation Costs

- < $7 cell phone
- Recycling surcharge + deposit

<table>
<thead>
<tr>
<th>Cost (US $)</th>
<th>Cell Phone**</th>
<th>Computer***</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collection*</td>
<td>6.00</td>
<td>23.50</td>
</tr>
<tr>
<td>Transportation</td>
<td>0.35</td>
<td>0.43</td>
</tr>
<tr>
<td>Sorting</td>
<td>-</td>
<td>3.50</td>
</tr>
<tr>
<td>Dismantling</td>
<td>0.03</td>
<td>2.75</td>
</tr>
<tr>
<td>Refining</td>
<td>0.32</td>
<td>7.87</td>
</tr>
<tr>
<td>Dispose of non-hazardous waste</td>
<td>0.01</td>
<td>0.83</td>
</tr>
<tr>
<td>Dispose of hazardous waste</td>
<td>0.03</td>
<td>5.00</td>
</tr>
</tbody>
</table>

*Average (cell phone:$4 to $8; computer:$13 to $34)

**Results from survey conducted with fifteen private US electronic recycling firms [Bhuie et al., 2004]

***[Boon et al., 2000]
Handset Reuse

- Refurbished phones
  - Only millions captured
  - Political issues

- PDAs
  - Learning tool / diary in elementary schools
  - Parking permit / navigator

- Location beacon

- Shipping container tracking

- Just park benches?
Reuse Summary

- Silicon fabrication and disposal are serious environmental concerns
- Reuse is a challenging goal, but we have to face the impact of our exponentially-growing computing demands
Part III: Cultural Change
Cultural Change

• Some sustainable technologies and practices exist, but managers and designers unaccustomed to the tradeoffs

• Need to develop frameworks and educate the next generation of technical leaders
Measuring Energy

- Coal-fired electric plants – 35% efficient
- Electrical transmission lines – 90% efficient
- Datacenter power distribution – also optimized for peak
- Other inefficiencies
  - Server power supplies
  - Battery charger / battery efficiency

www.epa.gov/cleanenergy/energy-resources/calculator.html
Life-Cycle Analysis

• Sustainable systems require a higher-level analysis
  – Energy and carbon metrics
  – Supply chains, end-of-life
  – Challenge: proprietary data
    – Study academic fabrication facilities
  – Make friends with your local industrial ecologist!
Reacting to Policy

• Standards and policies
  • Energy-star, SPEC power
  • Standards need knowledgeable participants
• Companies need to know how to respond to legislation and standards
  • WEEE, RoHS, Energy-star
Caveat: Jevons Paradox

- Efficiency in coal-fired machines led to greater demand for coal
Jevons Paradox

- Demand for computing could be elastic
- Need to measure productivity
Closing Remarks

• Computing for massive data poses significant sustainability challenges
• Good technical problems, but many are multidisciplinary
• We need to train the next generation of multidisciplinary engineers

energy.cs.ucsb.edu
Acknowledgements

• Luiz Barroso, Urs Hoezle, Bill Weihl (Google)
• Partha Raganan than (Google)
• Ricardo Bianchini (Rutgers)
• Roland Geyer (UCSB)
• Raj Amirtharajah, Venkatesh Akella (UC Davis)
• John Oliver (Calpoly)