Silicon Dreams: Inventing The Future

Dan Reed
reed@microsoft.com
www.hpcdan.org

Corporate Vice President
Microsoft Research (MSR)
A Question to Ponder ...

“What probability of successful return would you accept to be the first human to set foot on Mars?”
“The future is here, it is just not evenly distributed.”  

*William Gibson*

- Tools, culture and research
- Technology exponentials
- Scientific and cultural challenges
- Innovation and commitment
Looking back, in the public mind
- There were few or no experiences with …
  - web sites, email, spam, phishing, computer viruses
  - e-commerce, digital photography or telephony
- Cell phones were rare and expensive
- A Sony Walkman was still cool
- WiFi was almost unknown

Similarly, on the computing front
- The dot.com boom was still underway
- Robot vacuum cleaners were still science fiction
- A terabyte was a lot of data
- A peak teraflop was nation-scale infrastructure

The future depends on vision and context …
Imagine a Future Where …

- Your car drives and navigates for you
  ... and also parks the car (already a feature on some cars)
- Your sound system only plays music you love
  ... because it knows about every song you’ve ever heard
- Your phone only rings when you want to answer
  ... because it knows your emotional state
- All your family memories are recorded automatically
  ... via MEMS-based sensors and solid state storage
- Your body calls an ambulance when you’re ill
  ... via implanted, biologically powered diagnostic sensors
- Your DNA sample determines personalized treatment
  ... because genotype-phenotype models are specific
- Your office adjusts its behavior to your needs
  ... because it is smarter than you are
Imagine a Future Where ...

- Your every physical movement is tracked/logged
  - ... by embedded sensors on all human artifacts
- Your neighbors know all the {e-}books you read
  - ... because your electronic financial identity was stolen
- Your every call is monitored for content
  - ... by deep semantic analysis and logging
- Pollutant rationing determines lifestyle
  - ... via social network analysis and predictive characterization
- Your utilities fail due to a virus attack
  - ... because security was penetrated by a 10 year old
- Your DNA sample/lifestyle determine health cost
  - ... because you are targeted as a high risk genotype/lifestyle
- Cyberwar destroys U.S. financial institutions
  - ...because the U.S. lacks ability to construct IT infrastructure
Sapir–Whorf Hypothesis (SWH)
- Language influences the habitual thought of its speakers

Computing analog
- Available systems shape research agendas

Consider some past examples
- Cray-1 and vector computing
- VAX 11/780 and UNIX
- Workstations and Ethernet
- PCs and web
- Inexpensive clusters

Today’s examples
- multicore, sensors, clouds and services …

What lessons can we draw?
Rethink the nature of computing at extreme scale, from alternative, quantum computing models, through the transformative effects of manycore parallelism on programming systems and architectures, to massive cloud computing designs that each drive consumer, business and social applications and create value for Microsoft.

**extreme**: (ik-ˈstrēm)
- either of the two limits of a scale or range
- of a high or the highest degree or intensity
- a maximum or minimum value of a function
Some rules of thumb
- In the *near term*, we *overestimate* change
- In the *long term*, we *underestimate* change

Outside their field of expertise
- Experts are often better at predictions
- The contra-Delphi effect

*Inventing the future* is more successful
- Recognize exponentials
- Quantitative change brings qualitative change
- Recognize multidisciplinary coupling

Technological and social change
- Different rates with differing consequences
Computing History and Exponentials

- 1890-1945
  - Mechanical, relay
  - 7 year doubling
- 1945-1985
  - Tube, transistor,..
  - 2.3 year doubling
- 1985-2003
  - Microprocessor
  - 1 year doubling
- 2006-present
  - Multicore

Exponentials
- Chip transistor density: 2X in ~18 months
- Graphics: 100X in three years
- WAN bandwidth: 64X in two years
- Storage: 7X in two years

Source: Jim Gray
Storage: The Exponentials

Megabyte
- A small novel

Gigabyte
- A pickup truck filled with paper or a DVD

Terabyte: one thousand gigabytes – ~$100 today
- The text in one million books
- Entire U.S. Library of Congress is ~ten terabytes of text

Petabyte: one thousand terabytes
- 1-2 petabytes equals all academic research library holdings
- Coming soon to a pocket near you!
- Routinely generated annually by many scientific instruments

Exabyte: one thousand petabytes
- 5 exabytes of words spoken in the history of humanity

Source: Hal Varian, UC-Berkeley
Networks: The Exponentials
The Many Computing Eras

- Mainframe Era
- Pre-PC Era (1980)
- PC Era (1995)
- Internet Era (2000)
- Consumer Era (Today+)

**Implicit computing (21st century)**
- Natural interfaces and computing on behalf
- Embedded intelligence
- Adaptive, mobile context
- Number of cores/person → infinity
It seems reasonable to envision, for a time 10 or 15 years hence, a 'thinking center' that will incorporate the functions of present-day libraries together with anticipated advances in information storage and retrieval.

The picture readily enlarges itself into a network of such centers, connected to one another by wide-band communication lines and to individual users by leased-wire services. In such a system, the speed of the computers would be balanced, and the cost of the gigantic memories and the sophisticated programs would be divided by the number of users.

J.C.R. Licklider, 1960
December 8, 1993, C
Section (Front Page) John
Markoff, “A Free and Simple
Computer Link - NCSA’s
Mosaic Program”
Moore’s “Law” and Exponentials ...
The Brave New World

Free Lunch For Traditional Software

Operations per second for serial code

24 GHz 1 Core
12 GHz 1 Core
6 GHz 1 Core
3 GHz 1 Core
3 GHz 2 Cores
3 GHz 4 Cores
3 GHz 8 Cores

Additional operations per second if code can take advantage of concurrency

No Free Lunch For Traditional Software
(Without highly concurrent software, it won’t get any faster!)
Many challenges
Interfaces
Materials and physics
Display and storage
Various input scenarios

Many-core GPUs
Multitouch gestures
Multi-touch Physics
Paint rendering
3D display
Storage services

Brush and paint physics

XCG Project Lead: Bill Baxter

Project Gustav
Pastel Simulation Demo
Industry Giants Try to Break Computing’s Dead End

NY Times March 19, 2008 John Markoff

Intel and Microsoft said Tuesday that they planned to finance two groups of university researchers to start over and design a new generation of computing systems intended to break the industry out of a technological cul-de-sac that threatens to end decades of performance increases in computers.

$30M (U.S.) in funding (Microsoft/Intel/universities)
Can Multicore Supplant Clock Rates?

- Double the number of cores instead of speed
- No, at least without major innovation
  - Sequential code
  - Lack of parallel algorithms
  - Difficult programming
  - Few abstractions

If existing applications cannot use large parallelism
- New applications and systems will arise
- Software plus services is one “obvious” outcome …
Cloud Application Frameworks

Infrastructure as a Service

Applications as a Service

Software as a Service

Enables

Facilitates

Azure™ Services Platform

Live Services
.NET Services
SQL Services
SharePoint Services
Microsoft Dynamics CRM Services

Windows Azure
Each data center is 11.5 times the size of a football field.
Data Center "PacMan"

- Land - 2%
- Core and shell costs – 9%
- Architectural – 7%
- Mechanical/Electrical – 82%
  - 16% increase/year since 2004

Source: Christian Belady

Belady, C., “In the Data Center, Power and Cooling Costs More than IT Equipment it Supports”, *Electronics Cooling Magazine* (February 2007)
Consider These Services Challenges

- Environmental responsibility
  - Managing under a 100 MW envelope
  - Adaptive systems management
- Provisioning 100,000 servers
  - Hardware: at most one week after delivery
  - Software: at most a few hours
- Resilience during a blackout/disaster
  - Data center failure
  - Service rollover for 20M customers
- Programming the entire data center
  - Power, environmentals, provisioning
  - Component tracking, resilience, …
Philosophy: The data center is a computer that must be designed and programmed as an integrated system.
Optimizing Cloud Infrastructure …

- Watts alone are irrelevant
  - Turn off the equipment and declare victory
- The real metric is the following …
  
  **Effective Operations**
  
  **Watts \( \times \) Dollars**

- Many convoluted ideas
  - Application execution efficiency
  - Microarchitecture and system design
  - Power supply efficiency
  - Packaging and cooling overhead
  - Market costs for power and hardware
  - Cost of people and money

Bounding Boxes Matter!
## Step One: Be Efficient

<table>
<thead>
<tr>
<th>Class</th>
<th>Reliability</th>
<th>Maintainability</th>
<th>Mean PUE</th>
<th>Peak PUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>No UPS No Generator</td>
<td>No</td>
<td>1.08-1.14</td>
<td>1.08-1.14</td>
</tr>
<tr>
<td>B</td>
<td>UPS Optional Generator</td>
<td>Yes</td>
<td>1.14-1.19</td>
<td>1.14-1.20</td>
</tr>
<tr>
<td>C</td>
<td>UPS Generator + Maintenance Generator</td>
<td>Yes</td>
<td>1.18-1.23</td>
<td>1.22-1.28</td>
</tr>
<tr>
<td>D</td>
<td>UPS + Reserve UPS Generator + Reserve Generator</td>
<td>Yes</td>
<td>1.27-1.33</td>
<td>1.66-1.73</td>
</tr>
</tbody>
</table>

**GFS Gen4 Data Centers**

*Including capacities and utilizations...And we are not waiting for industry definitions*
Step Two: Rethink Assumptions ...

- Run Hot!
- Embrace Minimalism!
- Think Low Power!

![Graph showing Performance vs. Elapsed Time with Desired Lifetime and Utility Threshold]

Accept Failure!
All answers must be complete, backed by rigorous mathematical, scientific and business analyses. Each incorrect response will be penalized $5B (U.S.) and loss of turn.

Q1: Compare and contrast two possible designs for an international network of data centers for client plus cloud applications that optimize energy, performance, reliability, total cost of ownership and time to market.

Q2: Repeat Q1, but for a workload of (choose one) financial services, scientific and technical applications, or rich multimedia.
“In the last two decades advances in computing technology, from processing speed to network capacity and the Internet, have revolutionized the way scientists work.

From sequencing genomes to monitoring the Earth's climate, many recent scientific advances would not have been possible without a parallel increase in computing power - and with revolutionary technologies such as the quantum computer edging towards reality, what will the relationship between computing and science bring us over the next 15 years?”

http://research.microsoft.com/towards2020science
Black Hole Collision Problem

1963
Hahn and Lindquist
IBM 7090
One Processor
Each 0.2 MF
3 Hours

1977
Eppley and Smarr
CDC 7600
One Processor
Each 35 MF
5 Hours

1999
Seidel and Suen, *et al.*
NCSA SGI Origin
256 Processors
Each 500 MF
40 Hours

2001
Seidel *et al*
NCSA Pentium III
256 Processors
Each 1 GF
500,000 Hours total
plus 500,000 hours at NERSC

1,800,000,000X

300X

30,000X

~200X
Unanswered Physical Questions

- The Standard Model
  - Quantum theory
  - Electroweak and strong forces
- Gravity and relativity
  - No “Grand Unified Theory”
  - Gravity integration and rationale for mass
- Dark matter and dark energy
  - Most of the universe’s mass is “invisible”
  - Expansion is accelerating
- Wilkinson Microwave Anisotropy Probe
  - Universe is ~13.7 billion years old
- Computing insights
  - LSST, LHC and QCD
Large Hadron Collider (LHC)

- Probing the Standard Model
  - The Higgs boson and mass
- Online at CERN “soon”
- Six detectors (2 general purpose)
- International data hierarchy
- 10-20 PB/year
Lattice QCD Optimization

Iterative solution
Conjugant gradient solver
Memory intensive
Machines built for just this problem

LQCD Applications
- MILC
- Chroma
- FermiQCD
- CPS

Level 4
QCD Physics Toolbox
(Shared Alg, building blocks, Visualization, performance tools)
Workflow and Data analysis tools

Level 3
QOP
(Dirac Operators, inverters, forces)
Uniform User Env
(Runtime, accounting, grid)

Level 2
QDP/QDP++
(QCD Data Parallel)
QIO
(QCD I/O API)

Level 1
QLA
(QCD linear Algebra)
QMP
(QCD Message Passing)
QOP
(Dirac Operators, inverters, forces)
QMPL
(QCD Multicore Interface)

Lattice Dirac Operator

\[
[D\psi]^\alpha(x) = \frac{1}{2a} \sum_{\beta,\mu} [U^\alpha_\mu(x)\psi_\beta(x+\mu) - U^{*\beta\alpha}(x-\mu)\psi_\beta(x-\mu)] \quad \forall \alpha, x
\]
Large Synoptic Survey Telescope (LSST)

- Key project from the astronomy decadal survey
- Celestial cinematography
  - 3 gigapixel detector for wide field imaging
- Science
  - Beyond the standard model
  - Dark matter and dark energy
  - Observation targets
    - Near-Earth object survey
    - Weak lensing of wide fields
    - Supernovae measurements
- Features
  - >30 TB of data/night
  - Entire sky every 3 nights
  - Target first light 2013
Biochemical Physical Questions

Genomics
Proteomics
Cell biochemistry and structure
Cilia
Mucus
Airway/flow

Genomics
Proteomics
Cell biochemistry and structure
Cilia
Mucus
Airway/flow
Integrated, Predictive Biology

Computing, physics, engineering and biology
- Control theory, mathematical models, phase spaces
- From biological cartoons to predictive models
Genetics Gets Really Personal

The New York Times
Science

I.B.M. Joins Pursuit of $1,000 Personal Genome

Navigenics

There's DNA: And then there's what you do with it.

Our genetic testing services reduced to $999.

Success Stories

Find a physician

There's DNA.

Genome.

Cracking The Code!

Science

THE HUMAN GENOME

nature

THE HUMAN GENOME

Microsoft
The Future: One Possible Vision
“Consider a future device for individual use, which is a sort of mechanized private file and library. It needs a name, and to coin one at random, “memex” will do. A memex is a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory.”

Vannevar Bush

“As We May Think,” Atlantic Monthly, 1945
Dan TV: Microsoft SenseCam™

**Metrics**
- Light, XYZ acceleration
- Temperature, infrared

**Current use**
- Alzheimer’s research

Canberra, Australia
Long-Cycle Innovation

Technological, Scientific & Societal Trends

Responsive
Disruptive
Sustaining
Long-Cycle Innovation
Long-Cycle Innovation
Long-Cycle Innovation
Long-Cycle Innovation
Microsoft Research (MSR)

MSR Sites
- Redmond, Washington (Sept, 1991)
- San Francisco, California (June, 1995)
- Cambridge, United Kingdom (July, 1997)
- Beijing, China (Nov, 1998)
- Silicon Valley, California (July, 2001)
- Bangalore, India (Jan, 2005)
- Cambridge, Massachusetts (July, 2008)

Extreme Computing Group (XCG)
PITAC’s 1999 overall assessment

*Information Technology Research: Investing in Our Future*

During 2003-2005, focused PITAC assessments

- health care and IT, cybersecurity
- computational science

PCAST’s 2007 overall assessment

successor to 1999 review
A Question to Ponder ...

“What probability of successful return would you accept to be the first human to set foot on Mars?”