MURPA (Monash Undergraduate Research Projects Abroad)

Emerging Heterogeneous Technologies for High Performance Computing

Jack Dongarra
University of Tennessee
Oak Ridge National Lab
University of Manchester
Simulation: The Third Pillar of Science

- **Traditional scientific and engineering paradigm:**
  1) Do **theory** or paper design.
  2) Perform **experiments** or build system.

- **Limitations:**
  - Too difficult -- build large wind tunnels.
  - Too expensive -- build a throw-away passenger jet.
  - Too slow -- wait for climate or galactic evolution.
  - Too dangerous -- weapons, drug design, climate experimentation.

- **Computational science paradigm:**
  3) Use high performance computer systems to *simulate* the phenomenon
     » Base on known physical laws and efficient numerical methods.
Why Turn to Simulation?

- When the problem is too... 
  - Complex
  - Large / small
  - Expensive
  - Dangerous

- to do any other way.
Why Turn to Simulation?

- Climate / Weather Modeling
- Data intensive problems (data-mining, oil reservoir simulation)
- Problems with large length and time scales (cosmology)
Computational Science

Applications to Energy

Turbulence
Understanding the statistical geometry of turbulent dispersion of pollutants in the environment.

Energy Storage
Understanding the storage and flow of energy in next-generation nanostructured carbon tube supercapacitors.

Biofuels
A comprehensive simulation model of lignocellulosic biomass to understand the bottleneck to sustainable and economical ethanol production.

Nuclear Energy
High-fidelity predictive simulation tools for the design of next-generation nuclear reactors to safely increase operating margins.

Smart Truck
Aerodynamic forces account for ~53% of long haul truck fuel use. ORNL’s Jaguar predicted 12% drag reduction and yielded EPA-certified 6.9% increase in fuel efficiency.

Nano Science
Understanding the atomic and electronic properties of nanostructures in next-generation photovoltaic solar cell materials.

Source: Steven E. Koonin, DOE
High-Performance Computing Today

In the past decade, the world has experienced one of the most exciting periods in computer development.

Microprocessors have become smaller, denser, and more powerful.

The result is that microprocessor-based supercomputing is rapidly becoming the technology of preference in attacking some of the most important problems of science and engineering.
Technology Trends: Microprocessor Capacity

Gordon Moore (co-founder of Intel) *Electronics Magazine, 1965*

Number of devices/chip doubles every 18 months

2X transistors/Chip Every 1.5 years

Called “Moore’s Law”
Moore’s Law is Alive and Well

Data from Kunle Olukotun, Lance Hammond, Herb Sutter, Burton Smith, Chris Batten, and Krste Asanović
Slide from Kathy Yelick
Transistors into Cores
Today’s Multicores
99% of Top500 Systems Are Based on Multicore

- Sun Niagra2 (8 cores)
- Intel Xeon Phi (60 cores)
- IBM Power 7 (8 cores)
- Fujitsu Venus (16 cores)
- Nvidia Kepler (2688 Cuda cores)
- AMD Interlagos (16 cores)
- IBM BG/Q (18 cores)
- Intel Westmere (10 cores)
But Clock Frequency Scaling Replaced by Scaling Cores / Chip

15 Years of exponential growth ~2x year has ended

Data from Kunle Olukotun, Lance Hammond, Herb Sutter, Burton Smith, Chris Batten, and Krste Asanović
Slide from Kathy Yelick
Performance Has Also Slowed, Along with Power

Power is the root cause of all this

A hardware issue just became a software problem

Data from Kunle Olukotun, Lance Hammond, Herb Sutter, Burton Smith, Chris Batten, and Krste Asanović
Slide from Kathy Yelick
Power Cost of Frequency

- **Power** $\propto$ **Voltage**$^2$ x **Frequency** $\ (V^2F)$
- **Frequency** $\propto$ **Voltage**
- **Power** $\propto$ **Frequency**$^3$

<table>
<thead>
<tr>
<th></th>
<th>Cores</th>
<th>V</th>
<th>Freq</th>
<th>Perf</th>
<th>Power</th>
<th>PE (8ops/watt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superscalar</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>&quot;New&quot; Superscalar</td>
<td>1X</td>
<td>1.5X</td>
<td>1.5X</td>
<td>1.5X</td>
<td>3.3X</td>
<td>0.45X</td>
</tr>
</tbody>
</table>
# Power Cost of Frequency

- **Power** \(\propto\) **Voltage\(^2\)** *x* **Frequency** \((V^2F)\)

- **Frequency** \(\propto\) **Voltage**

- **Power** \(\propto\) **Frequency\(^3\)**

<table>
<thead>
<tr>
<th></th>
<th>Cores</th>
<th>V</th>
<th>Freq</th>
<th>Perf</th>
<th>Power</th>
<th>PE (Bops/watt)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Superscalar</strong></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>“New” Superscalar</strong></td>
<td>1X</td>
<td>1.5X</td>
<td>1.5X</td>
<td>1.5X</td>
<td>3.3X</td>
<td>0.45X</td>
</tr>
<tr>
<td><strong>Multicore</strong></td>
<td>2X</td>
<td>0.75X</td>
<td>0.75X</td>
<td>1.5X</td>
<td>0.8X</td>
<td>1.88X</td>
</tr>
</tbody>
</table>

(Bigger # is better)

50% more performance with 20% less power
Preferable to use multiple slower devices, than one superfast device
Look at the Fastest Computers

- **Strategic importance of supercomputing**
  - Essential for scientific discovery
  - Critical for national security
  - Fundamental contributor to the economy and competitiveness through use in engineering and manufacturing

- **Supercomputers are the tool for solving the most challenging problems through simulations**
Example of typical parallel machine
Example of typical parallel machine
Example of typical parallel machine

Shared memory programming between processes on a board and a combination of shared memory and distributed memory programming between nodes and cabinets.
Example of typical parallel machine

Combination of shared memory and distributed memory programming
What do you mean by performance?

- **What is a xflop/s?**
  - xflop/s is a rate of execution, some number of floating point operations per second.
    - Whenever this term is used it will refer to 64 bit floating point operations and the operations will be either addition or multiplication.

- **What is the theoretical peak performance?**
  - The theoretical peak is based not on an actual performance from a benchmark run, but on a paper computation to determine the theoretical peak rate of execution of floating point operations for the machine.
  - The theoretical peak performance is determined by counting the number of floating-point additions and multiplications (in full precision) that can be completed during a period of time, usually the cycle time of the machine.
  - For example, an Intel Xeon 5570 quad core at 2.93 GHz can complete 4 floating point operations per cycle or a theoretical peak performance of $11.72$ GFlop/s per core or $46.88$ Gflop/s for the socket.
- Listing of the 500 most powerful Computers in the World

- Yardstick: Rmax from LINPACK MPP

\[ Ax = b, \text{ dense problem} \]

- Updated twice a year

  SC‘xy in the States in November
  Meeting in Germany in June

- All data available from www.top500.org
Performance Development of HPC Over the Last 20 Years

- My Laptop (70 Gflop/s)
- My iPad2 & iPhone 4s (1.02 Gflop/s)

SUM

- N=1
- N=500

6-8 years

- 96.62 TFlop/s
- 1.17 TFlop/s
- 59.7 GFlop/s
- 400 MFlop/s
TOP500 Editions (41 so far, 20 years)
TOP500 Editions (41 so far, 20 years)
Processors Used in the Top500 Systems

Intel 80% (402)
AMD 10% (50)
IBM 8.4% (42)
Countries Share

Absolute Counts
US: 252
China: 66
Japan: 30
UK: 29
France: 23
Germany: 19
Australia: 5
## Top Systems in Australia

<table>
<thead>
<tr>
<th>Rank</th>
<th>Computer</th>
<th>Site</th>
<th>Manufacturer</th>
<th>Segment</th>
<th>Cores</th>
<th>Rmax</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>Fujitsu PRIMERGY CX250 S1, Xeon E5-2670 8C 2.600GHz, InfB FDR</td>
<td>National Computational Infrastructure, ANU</td>
<td>Fujitsu</td>
<td>Research</td>
<td>53504</td>
<td>978600</td>
<td>88</td>
</tr>
<tr>
<td>39</td>
<td>BlueGene/Q, Power BQC 16C 1.60GHz, Custom</td>
<td>Victorian Life Sciences Computation Initiative</td>
<td>IBM</td>
<td>Research</td>
<td>65536</td>
<td>715551</td>
<td>85</td>
</tr>
<tr>
<td>289</td>
<td>Nitro G16 3GPU, Xeon E5-2650 8C 2.000GHz, InfB FDR, NVIDIA 2050</td>
<td>CSIRO</td>
<td>Xenon Systems</td>
<td>Government</td>
<td>7308</td>
<td>133700</td>
<td>59</td>
</tr>
<tr>
<td>320</td>
<td>Sun Blade x6048, Xeon X5570 2.93 Ghz, InfB QDR</td>
<td>National Computational Infrastructure National Facility</td>
<td>Oracle</td>
<td>Academic</td>
<td>11936</td>
<td>126400</td>
<td>90</td>
</tr>
<tr>
<td>460</td>
<td>xSeries x3650M3 Cluster, Xeon X5650 6C 2.660GHz, InfB</td>
<td>Defence</td>
<td>IBM</td>
<td>Government</td>
<td>10776</td>
<td>101998</td>
<td>89</td>
</tr>
<tr>
<td>Name</td>
<td>Rmax Linpack# Pflops</td>
<td>Country</td>
<td>11</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>---------------------------</td>
<td>----------------------</td>
<td>--------------------</td>
<td>----</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Tianhe-2 (MilkyWay-2)</td>
<td>33.86</td>
<td>China NUDT: Hybrid Intel/Intel/Custom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Titan</td>
<td>17.59</td>
<td>United States Cray: Hybrid AMD/Nvidia/Custom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequoia</td>
<td>17.17</td>
<td>United States IBM: BG-Q/Custom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K Computer</td>
<td>10.51</td>
<td>Japan Fujitsu: Sparc/Custom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mira</td>
<td>8.59</td>
<td>United States IBM: BG-Q/Custom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stampede</td>
<td>5.17</td>
<td>United States Dell: Hybrid/Intel/Intel/IB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JUQUEEN</td>
<td>5.01</td>
<td>Germany IBM: BG-Q/Custom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vulcan</td>
<td>4.29</td>
<td>United States IBM: BG-Q/Custom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SuperMUC</td>
<td>2.90</td>
<td>Germany IBM: Intel/IB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tianhe-1A</td>
<td>2.57</td>
<td>China NUDT: Hybrid Intel/Nvidia/Custom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pangea</td>
<td>2.10</td>
<td>France Bull: Intel/IB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fermi</td>
<td>1.79</td>
<td>Italy IBM: BG-Q/Custom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DARPA Trial Subset</td>
<td>1.52</td>
<td>United States IBM: Intel/IB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spirit</td>
<td>1.42</td>
<td>United States SGI: Intel/IB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curie thin nodes</td>
<td>1.36</td>
<td>France Bull: Intel/IB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nebulae</td>
<td>1.27</td>
<td>China Dawning: Hybrid Intel/Nvidia/IB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yellowstone</td>
<td>1.26</td>
<td>United States IBM: BG-Q/Custom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blue Joule</td>
<td>1.25</td>
<td>United Kingdom IBM: BG-Q/Custom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pleiades</td>
<td>1.24</td>
<td>United States SGI Intel/IB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Helios</td>
<td>1.24</td>
<td>Japan Bull: Intel/IB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSUBAME 2.0</td>
<td>1.19</td>
<td>Japan NEC/HP: Hybrid Intel/Nvidia/IB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cielo</td>
<td>1.11</td>
<td>United States Cray: AMD/Custom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DiRAC</td>
<td>1.07</td>
<td>United Kingdom IBM: BG-Q/Custom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hopper</td>
<td>1.05</td>
<td>United States Cray: AMD/Custom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tera-100</td>
<td>1.05</td>
<td>France Bull: Intel/IB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oakleaf-FX</td>
<td>1.04</td>
<td>Japan Fujitsu: Sparc/Custom</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Performance of Countries

Total Performance [Tflop/s]


US

Graph showing the increase in total performance [Tflop/s] from 2000 to 2012 for the US.
Performance of Countries

Total Performance [Tflop/s]

- US
- EU

Year:
- 2000
- 2002
- 2004
- 2006
- 2008
- 2010
- 2012

Performance Scale:
- 0
- 1
- 10
- 100
- 1,000
- 10,000
- 100,000

Tflop/s
Performance of Countries

Total Performance [Tflop/s]

- US
- EU
- Japan

Year:
- 2000
- 2002
- 2004
- 2006
- 2008
- 2010
- 2012
Performance of Countries

Total Performance [Tflop/s]

0  1  10  100  1,000  10,000  100,000


US  EU  Japan  China
Performance of Countries

Total Performance [Tflop/s]

- US
- EU
- Japan
- China
- Others

## June 2013: The TOP10

<table>
<thead>
<tr>
<th>Rank</th>
<th>Site</th>
<th>Computer</th>
<th>Country</th>
<th>Cores</th>
<th>Rmax [Pflops]</th>
<th>% of Peak</th>
<th>Power [MW]</th>
<th>MFlops/Watt</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>National University of Defense Technology</td>
<td>Tianhe-2 NUDT, Xeon 12C 2.2GHz + IntelXeon Phi (57c) + Custom</td>
<td>China</td>
<td>3,120,000</td>
<td>33.9</td>
<td>70</td>
<td>17.8</td>
<td>1905</td>
</tr>
<tr>
<td>2</td>
<td>DOE / OS Oak Ridge Nat Lab</td>
<td>Titan, Cray XK7 (16C) + Nvidia Kepler GPU (14c) + Custom</td>
<td>USA</td>
<td>560,640</td>
<td>17.6</td>
<td>66</td>
<td>8.3</td>
<td>2120</td>
</tr>
<tr>
<td>3</td>
<td>DOE / NNSA L Livermore Nat Lab</td>
<td>Sequoia, BlueGene/Q (16c) + Custom</td>
<td>USA</td>
<td>1,572,864</td>
<td>16.3</td>
<td>81</td>
<td>7.9</td>
<td>2063</td>
</tr>
<tr>
<td>4</td>
<td>RIKEN Advanced Inst for Comp Sci</td>
<td>K computer Fujitsu SPARC64 VIIIfx (8c) + Custom</td>
<td>Japan</td>
<td>705,024</td>
<td>10.5</td>
<td>93</td>
<td>12.7</td>
<td>827</td>
</tr>
<tr>
<td>5</td>
<td>DOE / OS Argonne Nat Lab</td>
<td>Mira, BlueGene/Q (16c) + Custom</td>
<td>USA</td>
<td>786,432</td>
<td>8.16</td>
<td>81</td>
<td>3.95</td>
<td>2066</td>
</tr>
<tr>
<td>6</td>
<td>Texas Advanced Computing Center</td>
<td>Stampede, Dell Intel (8c) + Intel Xeon Phi (61c) + IB</td>
<td>USA</td>
<td>204,900</td>
<td>2.66</td>
<td>67</td>
<td>3.3</td>
<td>806</td>
</tr>
<tr>
<td>7</td>
<td>Forschungszentrum Juelich (FZJ)</td>
<td>JuQUEEN, BlueGene/Q, Power BQC 16C 1.6GHz+Custom</td>
<td>Germany</td>
<td>458,752</td>
<td>5.01</td>
<td>85</td>
<td>2.30</td>
<td>2178</td>
</tr>
<tr>
<td>8</td>
<td>DOE / NNSA L Livermore Nat Lab</td>
<td>Vulcan, BlueGene/Q, Power BQC 16C 1.6GHz+Custom</td>
<td>USA</td>
<td>393,216</td>
<td>4.29</td>
<td>85</td>
<td>1.97</td>
<td>2177</td>
</tr>
<tr>
<td>9</td>
<td>Leibniz Rechenzentrum</td>
<td>SuperMUC, Intel (8c) + IB</td>
<td>Germany</td>
<td>147,456</td>
<td>2.90</td>
<td>90*</td>
<td>3.42</td>
<td>848</td>
</tr>
<tr>
<td>10</td>
<td>Nat. SuperComputer Center in Tianjin</td>
<td>Tianhe-1A, NUDT Intel (6c) + Nvidia Fermi GPU (14c) + Custom</td>
<td>China</td>
<td>186,368</td>
<td>2.57</td>
<td>55</td>
<td>4.04</td>
<td>636</td>
</tr>
</tbody>
</table>

**500** US Navy DSRC Cray XT5 **USA** **12,720** **.096** **79**
Comparison

1 Gflop/s  
70 Gflop/s  
33,000,000 Gflop/s

~factor of 500,000
China, 2013: the 34 PetaFLOPS
Developed in cooperation between NUDT and Inspur for National Supercomputer Center in Guangzhou
Peak performance of 54.9 PFLOPS
- 16,000 nodes contain 32,000 Xeon Ivy Bridge processors and 48,000 Xeon Phi accelerators totaling 3,120,000 cores
- 162 cabinets in 720m² footprint
- Total 1.404 PB memory (88GB per node)
- Each Xeon Phi board utilizes 57 cores for aggregate 1.003 TFLOPS at 1.1GHz clock
- Proprietary TH Express-2 interconnect (fat tree with thirteen 576-port switches)
- 12.4 PB parallel storage system
- 17.6MW power consumption under load; 24MW including (water) cooling
- 4096 SPARC V9 based Galaxy FT-1500 processors in front-end system
ORNL’s “Titan” Hybrid System: Cray XK7 with AMD Opteron and NVIDIA Tesla processors

SYSTEM SPECIFICATIONS:
- Peak performance of 27 PF
  - 24.5 Pflop/s GPU + 2.6 Pflop/s AMD
- 18,688 Compute Nodes each with:
  - 16-Core AMD Opteron CPU
  - NVIDIA Tesla “K20x” GPU
  - 32 + 6 GB memory
- 512 Service and I/O nodes
- 200 Cabinets
- 710 TB total system memory
- Cray Gemini 3D Torus Interconnect
- 9 MW peak power

4,352 ft²
404 m²
Titan: Cray XK7 System

System:
- 200 Cabinets
- 18,688 Nodes
- 27 PF
- 710 TB

Cabinet:
- 24 Boards
- 96 Nodes
- 139 TF
- 3.6 TB

Board:
- 4 Compute Nodes
- 5.8 TF
- 152 GB

Compute Node:
- 1.45 TF
- 38 GB
**Sequoia**

- **USA, 2012: BlueGene strikes back**
- Built by IBM for NNSA and installed at LLNL
- 20,123.7 TFLOPS peak performance
  - Blue Gene/Q architecture
  - 1,572,864 total PowerPC A2 cores
  - 98,304 nodes in 96 racks occupy 280m²
  - 1,572,864 GB DDR3 memory
  - 5-D torus interconnect
  - 768 I/O nodes
  - 7890kW power, or 2.07 GFLOPS/W
  - Achieves 16,324.8 TFLOPS in HPL (#1 in June 2012), about 14 PFLOPS in HACC (cosmology simulation), and 12 PFLOPS in Cardioid code (electrophysiology)
Japanese K Computer

K computer Specifications

<table>
<thead>
<tr>
<th>CPU (SPARC64 VIIIfx)</th>
<th>Cores/Node</th>
<th>8 cores (@2GHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Performance</td>
<td>128GFlops</td>
</tr>
<tr>
<td></td>
<td>Architecture</td>
<td>SPARC V9 + HPC extension</td>
</tr>
<tr>
<td></td>
<td>Cache</td>
<td>L1(I/D) Cache: 32KB/32KB L2 Cache: 6MB</td>
</tr>
<tr>
<td></td>
<td>Power</td>
<td>58W (typ. 30 C)</td>
</tr>
<tr>
<td></td>
<td>Mem. bandwidth</td>
<td>64GB/s.</td>
</tr>
<tr>
<td>Node</td>
<td>Configuration</td>
<td>1 CPU / Node</td>
</tr>
<tr>
<td></td>
<td>Memory capacity</td>
<td>16GB (2GB/core)</td>
</tr>
<tr>
<td>System board(SB)</td>
<td>No. of nodes</td>
<td>4 nodes /SB</td>
</tr>
<tr>
<td></td>
<td>Rack</td>
<td>24 SBs/rack</td>
</tr>
<tr>
<td></td>
<td>System</td>
<td>&gt; 80,000</td>
</tr>
</tbody>
</table>

Interconnect

- Topology: 6D Mesh/Torus
- Performance: 5GB/s. for each link
- No. of link: 10 links/node
- Additional feature: H/W barrier, reduction
- Architecture: Routing chip structure (no outside switch box)

Cooling

- CPU, ICC*: Direct water cooling
- Other parts: Air cooling

System

- LINPACK 10 Pflops
- Over 1PB mem.
- 800 racks
- 80,000 CPUs
- 640,000 cores

Node

- 128 GFlops
- 16 GB Memory
- 64GB/s Memory bandwidth

System Board

- 512 GFlops
- 15TB memory

Linpack run with 705,024 cores at 10.51 Pflop/s (88,128 CPUs), 12.7 MW; 29.5 hours
Fujitsu to have a 100 Pflop/s system in 2014
First Chinese Supercomputer to use a Chinese Processor
- Sunway BlueLight MPP
- ShenWei SW1600 processor, 16 core, 65 nm, fabbed in China
- 125 Gflop/s peak
- #14 with 139,364 cores, .796 Pflop/s & 1.07 Pflop/s Peak
- Power Efficiency 741 Mflops/W

Coming soon, Loongson (Godson) processor
- 8-core, 65nm Loongson 3B processor runs at 1.05 GHz, with a peak performance of 128 Gflop/s
Industrial Use of Supercomputers

- Of the 500 Fastest Supercomputer
- Worldwide, Industrial Use is ~ 50%

- Aerospace
- Automotive
- Biology
- CFD
- Database
- Defense
- Digital Content Creation
- Digital Media
- Electronics
- Energy
- Environment
- Finance
- Gaming
- Geophysics
- Image Proc./Rendering
- Information Processing Service
- Information Service
- Life Science
- Media
- Medicine
- Pharmaceuticals
- Research
- Retail
- Semiconductor
- Telecomm
- Weather and Climate Research
- Weather Forecasting
Critical Issues at Peta & Exascale for Algorithm and Software Design

• **Synchronization-reducing algorithms**
  - Break Fork-Join model

• **Communication-reducing algorithms**
  - Use methods which have lower bound on communication

• **Mixed precision methods**
  - 2x speed of ops and 2x speed for data movement

• **Autotuning**
  - Today’s machines are too complicated, build “smarts” into software to adapt to the hardware

• **Fault resilient algorithms**
  - Implement algorithms that can recover from failures

• **Reproducibility of results**
  - Today we can’t guarantee this. We understand the issues, but some of our “colleagues” have a hard time with this.
Conclusions

• For the last decade or more, the research investment strategy has been overwhelmingly biased in favor of hardware.

• This strategy needs to be rebalanced - barriers to progress are increasingly on the software side.

• Moreover, the return on investment is more favorable to software.
  - Hardware has a half-life measured in years, while software has a half-life measured in decades.
  - High Performance Ecosystem out of balance
    - Hardware, OS, Compilers, Software, Algorithms, Applications
      • No Moore’s Law for software, algorithms and applications