Leading-edge Computers and the Extraordinary Research They Enable

Thom H. Dunning, Jr.

Department of Chemistry
University of Illinois at Urbana-Champaign

Northwest Institute for Advanced Computing
Pacific Northwest National Laboratory & University of Washington

Department of Chemistry
University of Washington
More & More FLOPS & Bytes

Computational scientists always seem to need more and more computing power and storage. What is the outcome of access to increasing amounts of flops & bytes?
More & More FLOPS & Bytes

Increasing Accuracy of Predictions

- **Bond Energies**
  - Critical for describing many chemical phenomena
  - Difficult to determine experimentally

- **Accuracy of Predictions**
  - Increased dramatically from 1970-2000

- **How?**
  - New theoretical approaches
  - New computational techniques
  - More computing power

Date: 3/29/2016
More & More FLOPS & Bytes

Increasing reach of Simulations

• In 1990
  ▪ Model systems, e.g., ether–alkali ion complexes

• In 2000
  ▪ Model separations agents, e.g., 18-crown-6–alkali ion complexes

• In 2010
  ▪ Real-world separations agents, e.g., Still’s crown ether–ion complexes
More & More FLOPS & Bytes

NWChem: Evolution

- Complete basic suite of tools
  - High accuracy
  - Density Functional Theory
  - Ab-initio MD
  - Classical MD
  - Basic NMR

**NWChem 4.X (2001-2006)**
- Expanding capability suite
  - Environmental models
  - Spectroscopic properties
  - TDDFT excited states
  - Relativistic effects
  - POLYRATE direct dynamics

**NWChem 5.X (2006-2008)**
- Fast time-to-solution
  - High accuracy reaches 1.3 Pflop/s
- Partners contribute
  - Q-HOP in MD
  - VENUS interface
  - Constrained DFT

- Open-source release
  - Education Community License 2.0
- Unique tools
  - Real-time TDDFT response
  - Optical rotation
- Scalability on new hardware

**New functionality**
- DFT with long range dispersion
Every field of computational science has a similar story to tell!

The purpose of computing is insight, not numbers.

Richard W. Hamming, 1962

The purpose of computing is numbers as well as insight.

with apologies to Dr. Hamming
Petaflops & Petabytes

Many areas of science and engineering require extraordinary computing power to solve the mathematical equations describing the phenomena of interest and enormous data handling capability to explore the massive data sets now becoming available.
Who Needs Petaflops?

- To calculate the energy content of Iso-octane
  - Iterative solution of 275 million coupled equations
  - Exchange of 2.5 petabytes of data between processors
  - Exchange of 15 terabytes of data between memory and disks
  - Execution of 30 quadrillion arithmetic operations

- Modeling Combustion of Fuels
Petaflops & Petabytes

Who Needs Petabytes?

Astronomy has become one of the first digital science, replacing photographs with digital images.

The Large Synoptic Survey Telescope (LSST) has a 3.2 gigapixel camera and will produce 15-20 terabytes of data per night and more than 100 petabytes over its first 10 years of operation.

With the genomic revolution, biology and biomedicine are rapidly becoming digital sciences. The opportunities for breakthroughs in these areas are just beginning to be explored as exemplified by the Genome 10K project.
Petaflops & Petabytes

Similar Needs Across Science & Engineering

Molecular Science

Weather & Climate

Astronomy

Geosciences

Health

Date: 3/29/2016
Blue Waters: An Exemplary Petascale Computer

Blue Waters and the National Petascale Computing Facility at the University of Illinois at Urbana-Champaign are truly extraordinary research resources for the nation.
Blue Waters
Petascale Computing Facility

- Modern Data Center
  - 90,000+ ft² total
  - 30,000 ft² raised floor
  - 20,000 ft² machine room gallery

- Energy Efficiency
  - LEED certified Gold
  - Power Utilization Efficiency
    = 1.1–1.2

Partners
EYP MCF/
Gensler
IBM
Yahoo!

Date: 3/29/2016
Blue Waters

Overview of Blue Waters Computing System

- Sonexion: 26 PBs
- Spectra Logic: 300 PBs

10/40/100 Gb Ethernet Switch

IB Switch

>1 TB/sec

120+ Gb/sec

~100 GB/sec

Date: 3/29/2016
# Blue Waters

## Specifications: Blue Waters & Titan

<table>
<thead>
<tr>
<th></th>
<th>Blue Waters</th>
<th>Titan</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vendor(s)</strong></td>
<td>Cray/AMD/NVIDIA</td>
<td>Cray/AMD/NVIDIA</td>
</tr>
<tr>
<td><strong>Processors</strong></td>
<td>Interlagos/Kepler</td>
<td>Interlagos/Kepler</td>
</tr>
<tr>
<td><strong>Peak Performance</strong></td>
<td>13.1 PF</td>
<td>27.1 PF</td>
</tr>
<tr>
<td><strong>CPU/GPU</strong></td>
<td>7.6 / 5.5</td>
<td>2.6 / 24.5</td>
</tr>
<tr>
<td><strong>Number of Chips (CPU/GPU)</strong></td>
<td>48,352 / 4,224</td>
<td>18,688 / 18,688</td>
</tr>
<tr>
<td><strong>Amount of Memory</strong></td>
<td>1.66 PB</td>
<td>0.71 PB</td>
</tr>
<tr>
<td><strong>Disk Storage, Capacity (usable)</strong></td>
<td>26 PB</td>
<td>&gt;10 TB</td>
</tr>
<tr>
<td><strong>Disk Storage, Bandwidth (sustained)</strong></td>
<td>1.2 TB/s</td>
<td>0.24 TB/s</td>
</tr>
<tr>
<td><strong>Archival Storage, Capacity (usable)</strong></td>
<td>300 PB</td>
<td>125 PB</td>
</tr>
<tr>
<td><strong>Archival Storage, Bandwidth (sustained)</strong></td>
<td>88 GB/s</td>
<td>18 GB/s</td>
</tr>
</tbody>
</table>

Date: 3/29/2016
Blue Waters
Exploring New Materials for Desalination

Nanoporous Molybdenum Disulfide (MoS$_2$)
N. Aluru, UIUC

Date: 3/29/2016
Blue Waters

Modeling the HIV-1 Capsid

J. Perilla, G. Zhao, A. Gronenborn, P. Zhang and Klaus Schulten
Blue Waters
Predicting the Impact of Earthquakes

Map of California’s San Andreas Fault.

Date: 3/29/2016
Blue Waters

One of Many Earthquake Scenarios

Courtesy of T. Jordan, SCEC
Spread of Contagious Diseases

No Intervention

**Intervention:** next-day treatment of 90% of cases with anti-virals, school closures, 50% household quarantine.

Video courtsey of N. Ferguson, Imperial College, London
What’s Next

Computing technology continues to advance, although advancement means new computing technologies—Dennard Scaling is over, Moore’s Law is approaching its end.
## What’s Next

### Argonne’s Mira & Aurora Systems

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor</td>
<td>PowerPC</td>
<td>Knights Hill</td>
</tr>
<tr>
<td>Peak Performance</td>
<td>10 PF</td>
<td>180 PF</td>
</tr>
<tr>
<td>Cores/Processor</td>
<td>16</td>
<td>&gt;72</td>
</tr>
<tr>
<td>Number of Nodes</td>
<td>49,152</td>
<td>&gt;50,000</td>
</tr>
<tr>
<td>Memory</td>
<td>786 TB</td>
<td>&gt;7 PB</td>
</tr>
<tr>
<td>File System Capacity</td>
<td>26 PB</td>
<td>&gt;150 PB</td>
</tr>
<tr>
<td>File System Bandwidth</td>
<td>300 GB/s</td>
<td>&gt;1 TB/s</td>
</tr>
<tr>
<td>Peak Power</td>
<td>4.8 MW</td>
<td>13 MW</td>
</tr>
</tbody>
</table>
What’s Next
Intel’s Many Integrated Core (MIC) Processors

Unveiling Details of Knights Landing
(Next Generation Intel® Xeon Phi™ Products)

Platform Memory: DDR4 Bandwidth and Capacity Comparable to Intel® Xeon® Processors

Compute: Energy-efficient IA cores
- Microarchitecture enhanced for HPC
- 3X Single Thread Performance vs Knights Corner
- Intel Xeon Processor Binary Compatible

On-Packagge Memory:
- up to 16GB at launch
- 1/3X the Space
- 5X Bandwidth vs DDR4
- 5X Power Efficiency

Jointly Developed with Micron Technology

All products, computer systems, dates and figures specified are preliminary based on current expectations, and are subject to change without notice. 1 Over 3 Terasflops of peak theoretical double-precision performance is preliminary and based on current expectations of cores, clock frequency and floating point operations per cycle. FLOPS = cores x clock frequency x floating-point operations per second per cycle. 2 Modified version of Intel® Silvermont microarchitecture currently found in Intel® Atom™ processors. 3 Modifications include AVX512 and 4 threads/core support. 4 Projected peak theoretical single-thread performance relative to 1st Generation Intel® Xeon Phi™ Coprocessor 7120P (formerly codenamed Knights Corner). 5 Binary compatible with Intel Xeon processors using Howell instruction Set (except TXT). 6 Projected results based on internal Intel analysis of Knights Landing memory vs Knights Corner (DDR3). 7 Projected result based on internal Intel analysis of STREAM benchmark using a Knights Landing processor with 16GB of ultra high-bandwidth versus DDR4 memory only with all channels populated.

Conceptual—Not Actual Package Layout
From Here to the Future

As computing technology continues to advance, and change, how do we ensure that computational science and engineering continues to advance?
From Here to the Future

Assertions About Computational Modeling

• **All Science & Engineering Impacted**
  – All scientists and engineers must understand computational simulation

• **Fidelity of Computational Models Will Continue to Advance**
  – Power of computational simulation will continue to grow

• **It’s the Software!**
  – Computational codes encapsulate current state of knowledge
  – Developing computational codes to describe natural or engineered systems requires knowledge of:
    • Foundational science
    • Applied mathematics
    • Computer science, including modern software engineering practices
  – As computing technology changes, software challenge increases
From Here to the Future

What is Needed to Continue to Advance

Computers of the future will be built on new technologies: many-core processors in the near term

• Need New, More Efficient Algorithms
  – Fully exploit available concurrency
  – Adapted to decreased memory per flop
  – Minimized data movement

• Need New Programming Models
  – Need better programming languages, domain specific languages

• Need Better Education
  – Undergraduate level use of HPC
  – Graduate level instruction in HPC
From Here to the Future
Educating a New Generation of Scientists

• Undergraduate Education
  – Use of simulation in undergraduate courses:
    • Used, but largely limited to simulations that run on laptop
  – Courses in Computational “X” are not widespread
    • Largely driven by faculty interest, not national requirements
    • Usually do not address basic computational issues – largely focused on teaching how to use existing applications
  – Existing mathematics and computer science courses are often not well suited for computational science and engineering—need specialized courses

• Graduate Education
  – Graduate programs in computational science and engineering are growing, but “best practices” have not been established
  – Existing mathematics and computer science courses are often not well suited for computational science and engineering—need specialized courses
From Here to the Future

Statement from American Chemical Society

Undergraduate Professional Education in Chemistry
ACS Guidelines and Evaluation Procedures for Bachelor’s Degree Programs

Computational Capabilities and Software. The ability to compute chemical properties and phenomena complements experimental work by providing understanding and predictive power. Students should use computing facilities and computational chemistry software in their course work and research.

Spring 2008
ACS Committee on Professional Training

But, there is a huge gap between this statement and the state of practice in universities, even major research universities.
From Here to the Future

Virtual School of Computational S&E

- ** Goals of Virtual School
  - Increase and enhance computing-related curricula available to graduate students in science & engineering
    - Prepare the current and next generation of scientists and engineers to utilize leading-edge computer systems

- ** Organization of Virtual School
  - Multi-institutional virtual organization, pooling expertise of faculty/staff in universities and national laboratories
    - One week intensive courses
      » Morning lectures with afternoon “hands-on” projects
    - Lecturer at central site, other sites linked by HD video
    - All sites required to be staffed (faculty member plus TAs)
Thank You!