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

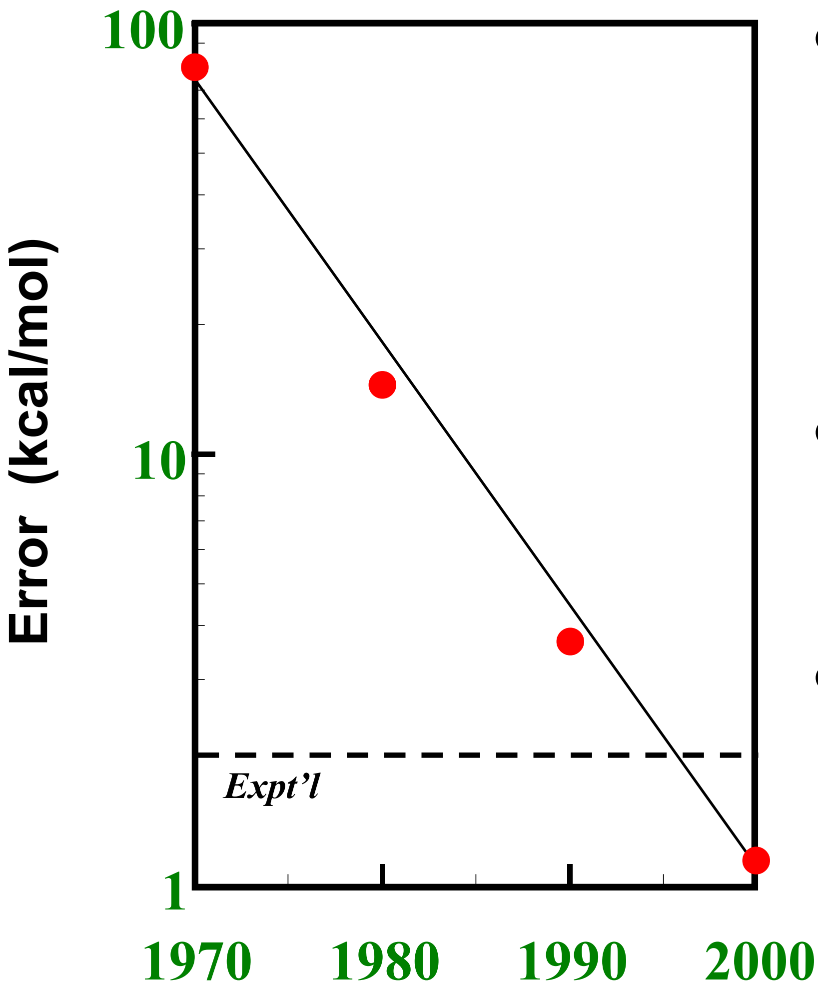
Department of Chemistry
University of Washington

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

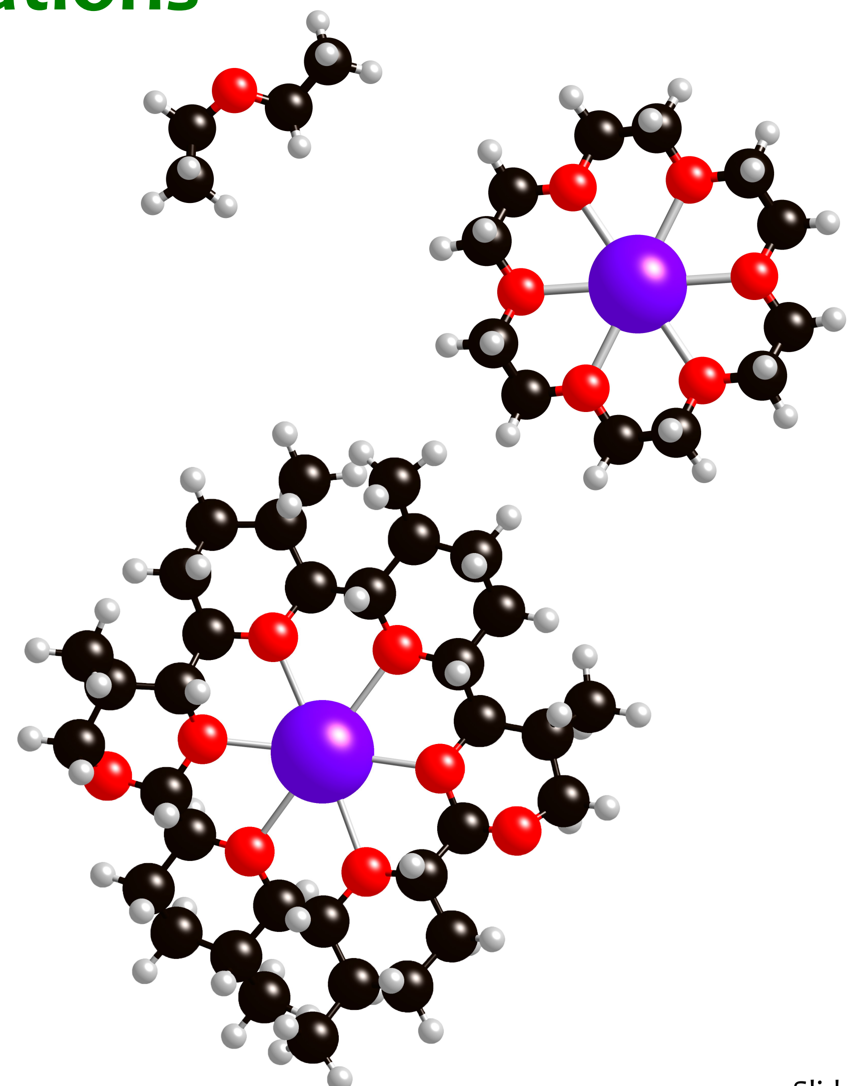


- **Bond Energies**
  - Critical for describing many chemical phenomena
  - Difficult to determine experimentally
- **Accuracy of Predictions**
  - Increased dramatically from 1970-2000
- **How?**
  - New theoretical approaches
  - New computational techniques
  - More computing power

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



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

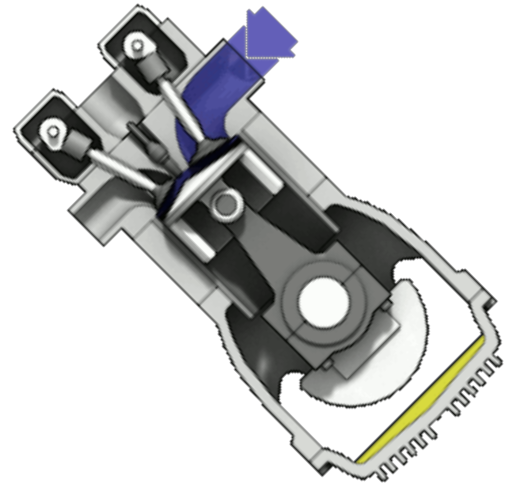
Petaflops & Petabytes

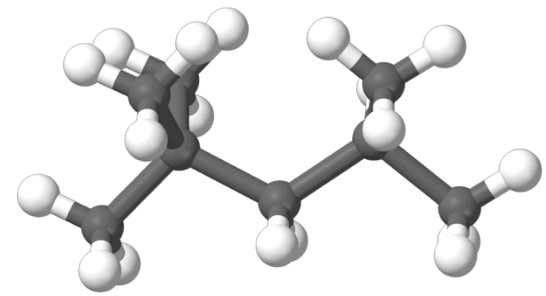
Many areas of science and engineering require extraordinary computing power to solve the mathematical equations describing the phenomena of interest and enormous data handling capability to explore the massive data sets now becoming available

Petaflops & Petabytes

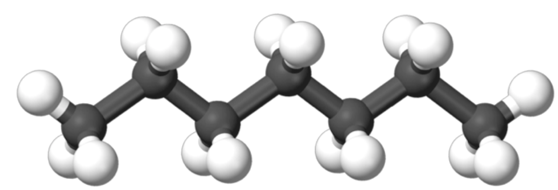
Who Needs Petaflops?

- **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**





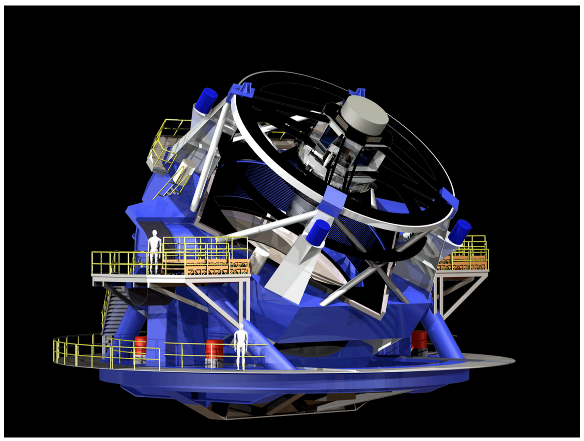
Iso-octane
(Octane Rating = 100)



***n*-heptane**
(Octane Rating = 0)

Petaflops & Petabytes

Who Needs Petabytes?

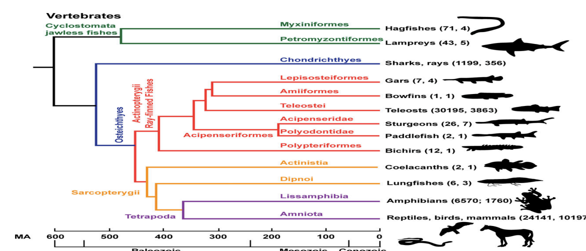


Astronomy has become one of the first digital science, replacing photographs with digital images.

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.

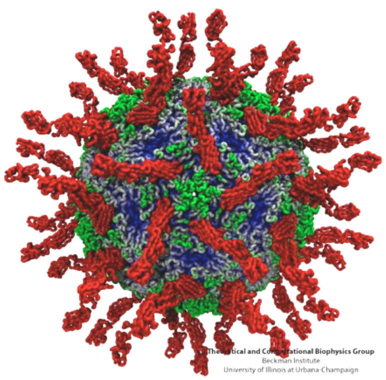
GENOME 10K.

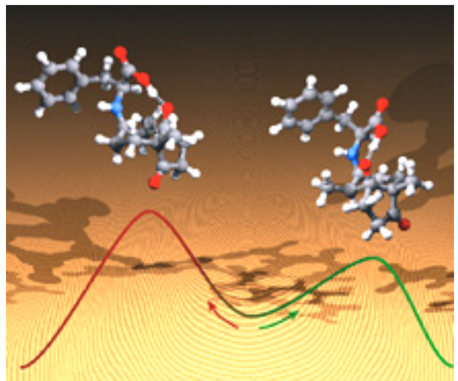


Petaflops & Petabytes

Similar Needs Across Science & Engineering

Molecular Science

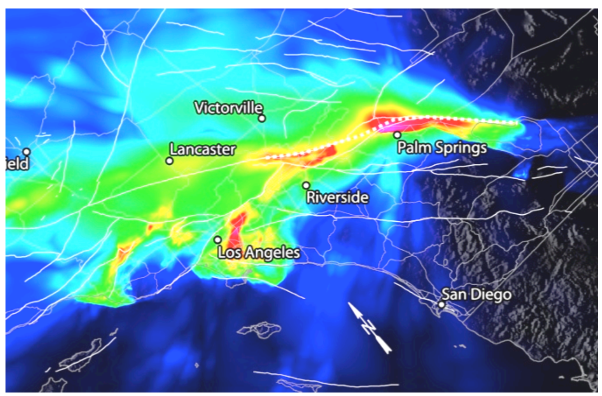




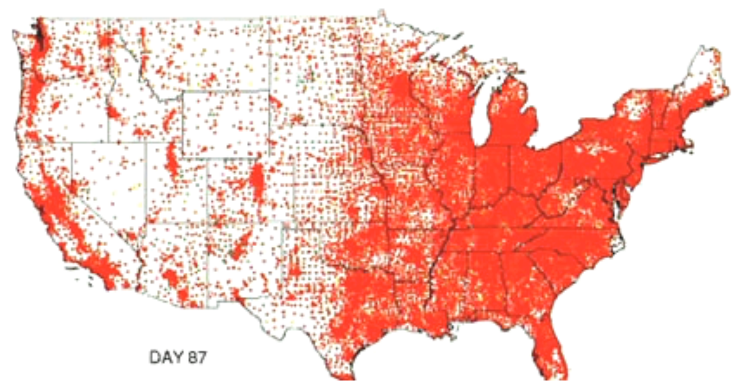
Weather & Climate

Astronomy

Geosciences



Health

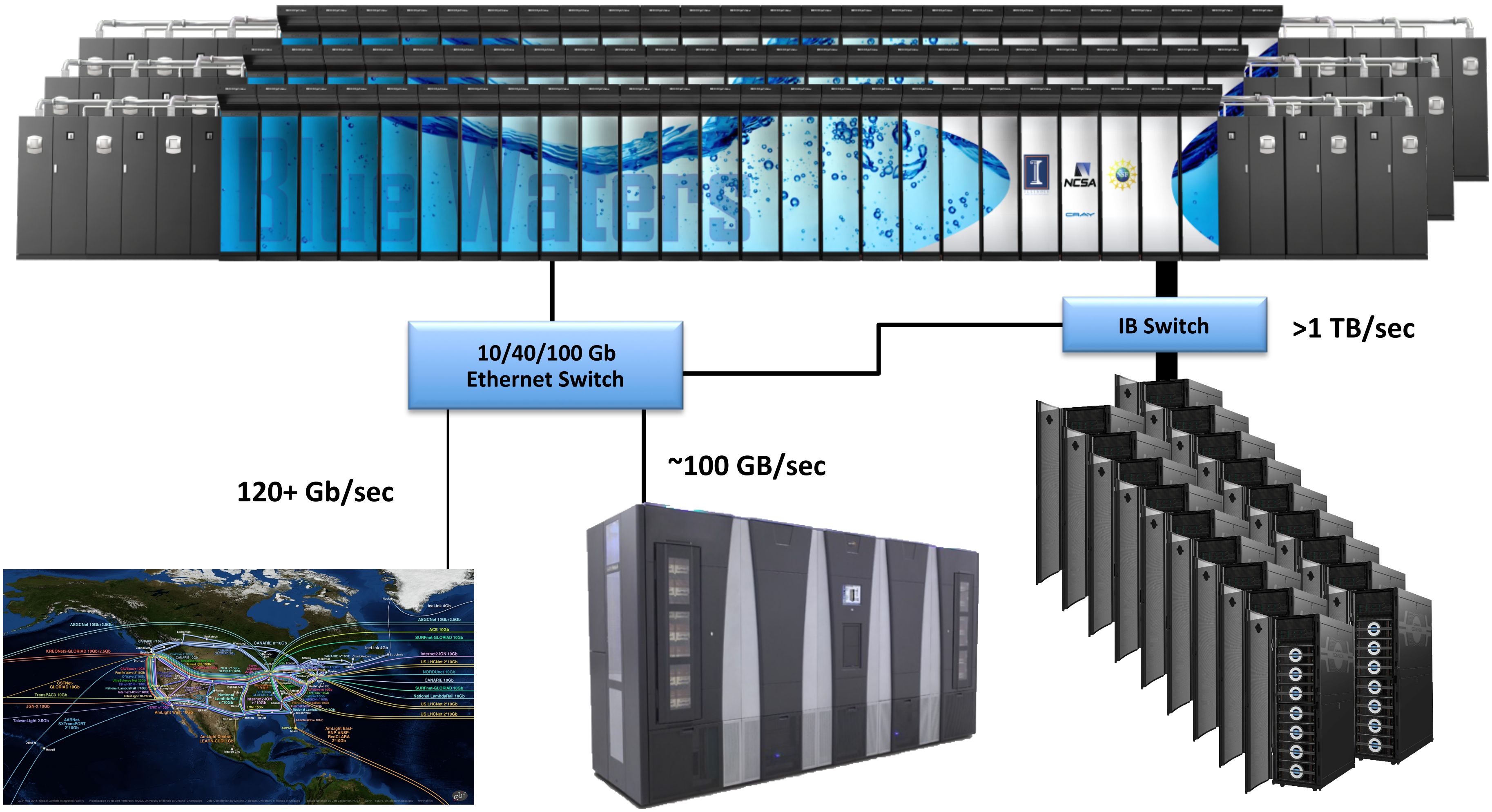


Current Leading-edge Computers: Blue Waters

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

Blue Waters Computing System



WAN

Spectra Logic: 300 PBs

Sonexion: 26 PBs

Blue Waters

Petascale Computing Facility

Partners

EYP MCF/
Gensler
IBM
Yahoo!

- **Modern Data Center**
  - 90,000+ ft^2 total
  - 30,000 ft^2 raised floor
    20,000 ft^2 machine room gallery
- **Energy Efficiency**
  - LEED certified Gold
  - Power Utilization Efficiency = 1.1–1.2

Blue Waters

Specifications: Blue Waters & Titan

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

Blue Waters

Specifications: Blue Waters & Titan

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

A CAUTIONARY NOTE:

SUSTAINED PERFORMANCE

BLUE WATERS: 1.37 PFs

TITAN: 0.64 PFs

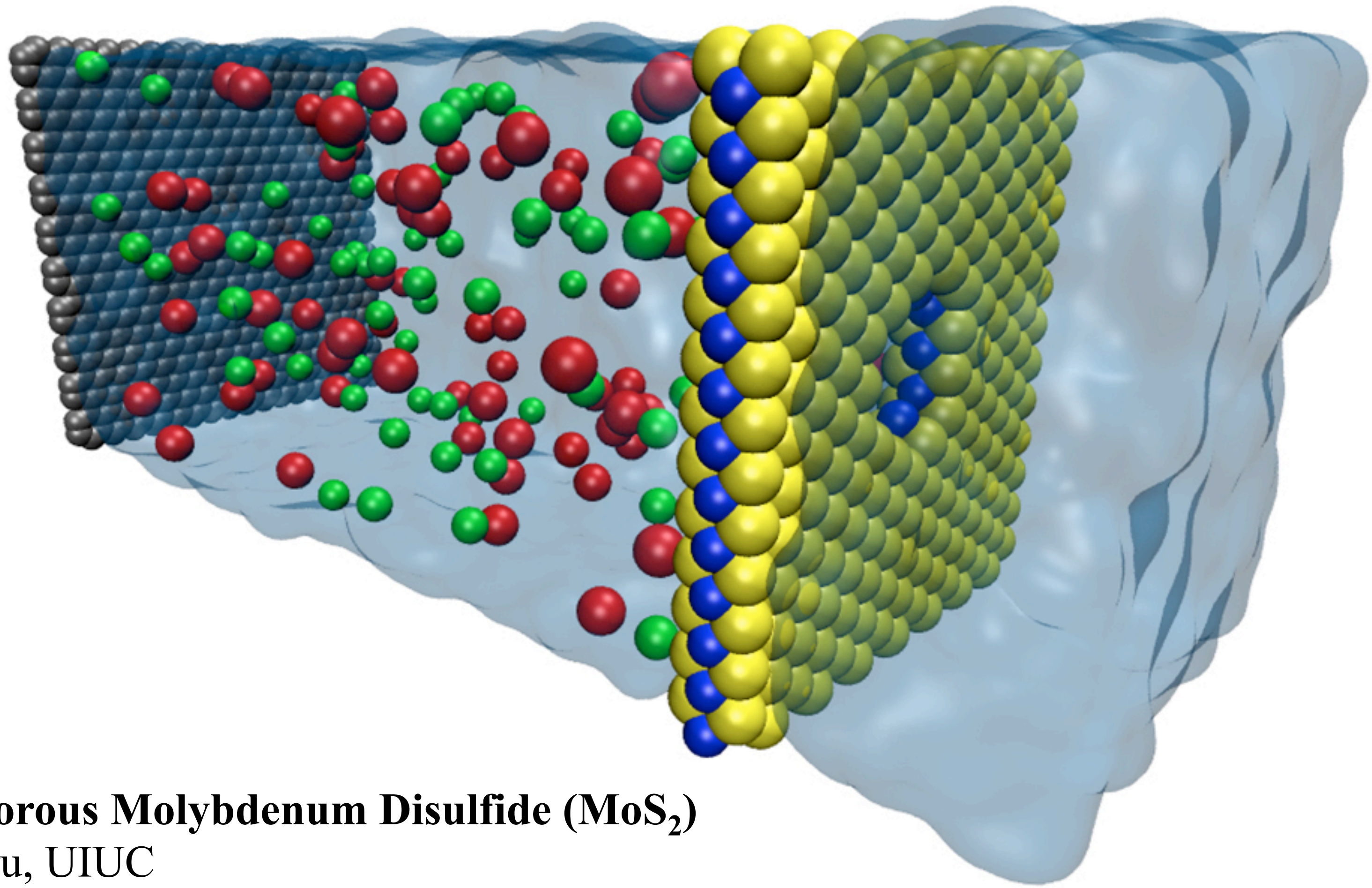
Blue Waters

Specifications: Blue Waters & Titan

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

Blue Waters

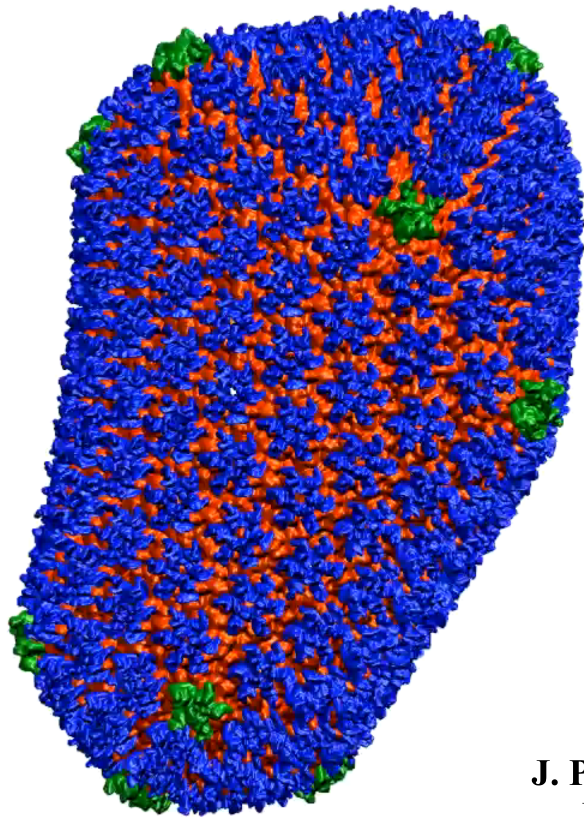
Exploring New Materials for Desalination



Nanoporous Molybdenum Disulfide (MoS_2)
N. Aluru, UIUC

Blue Waters

Modeling the HIV-1 Capsid



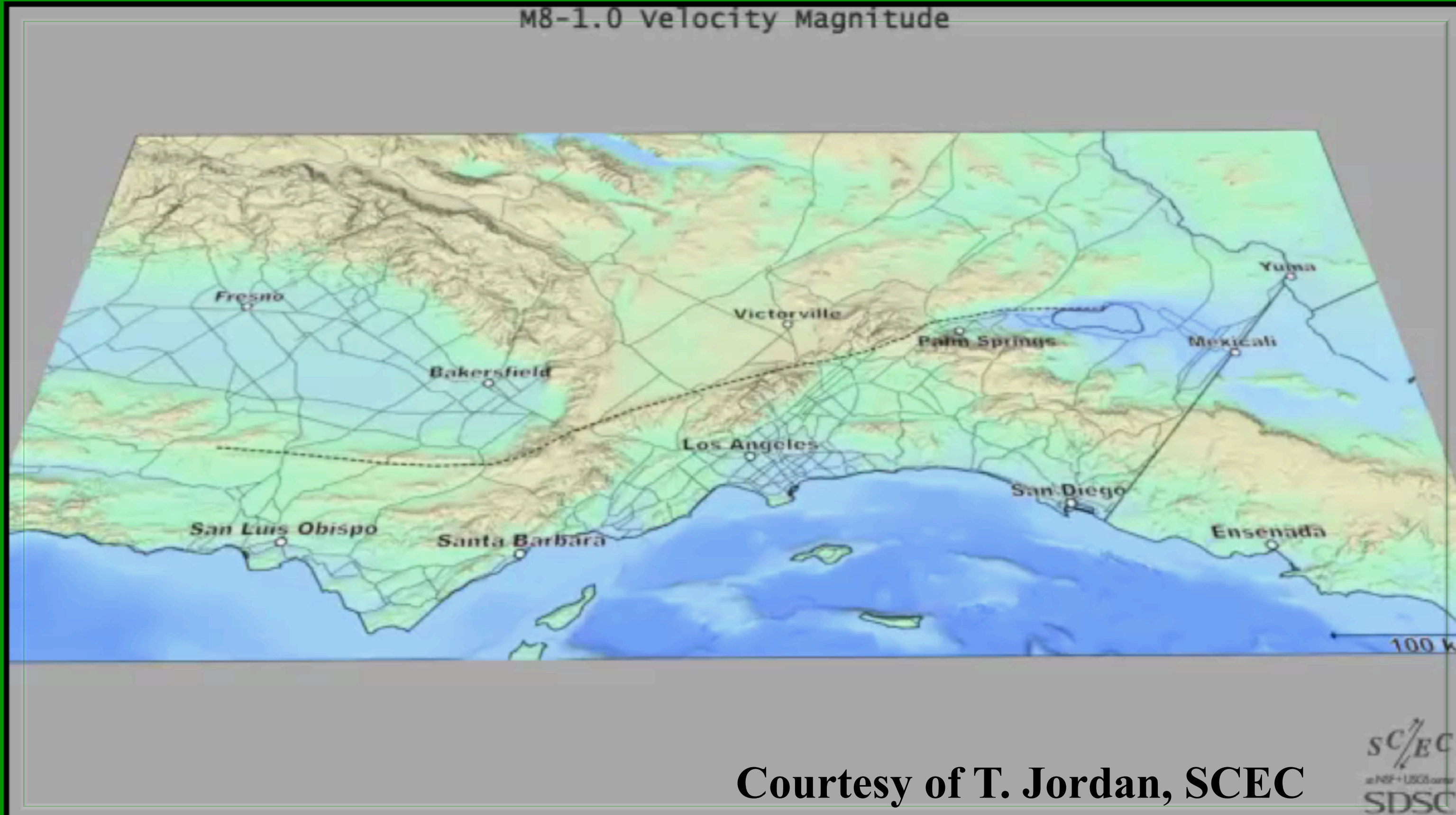
J. Perilla, G. Zhao, A. Gronenborn, P. Zhang and Klaus Schulten

Blue Waters

Predicting the Impact of Earthquakes

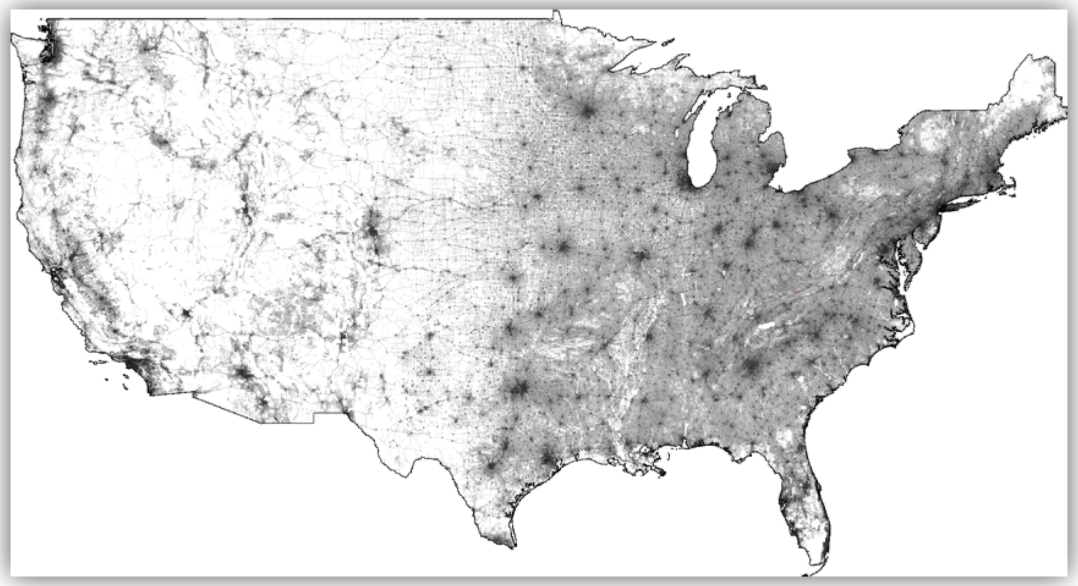
Blue Waters

One of Many Earthquake Scenaios



Courtesy of T. Jordan, SCEC

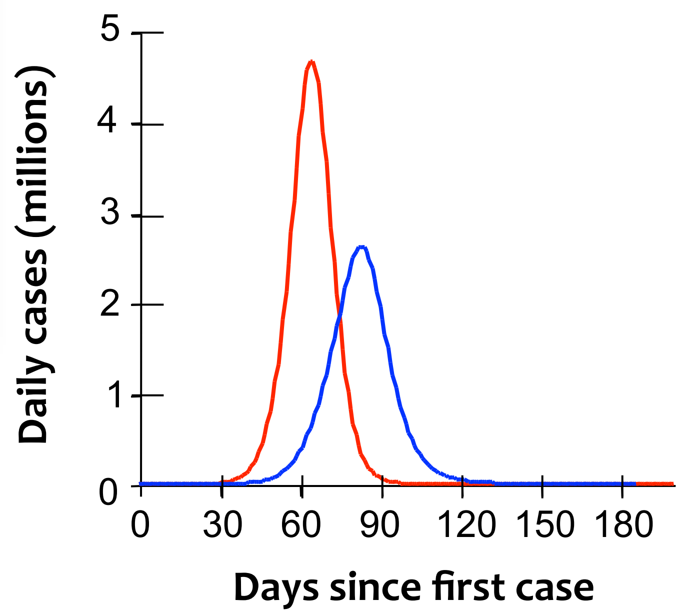
Spread of Contagious Diseases



Video courtsey of N. Ferguson, Imperial College, London

No Intervention

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



Future Leading-edge Computers: Summit & Aurora

Although computing technology continues to advance, the best strategy for architecting exascale computers is unclear—the U.S. Department of Energy has decided that two different architectures will be explored.

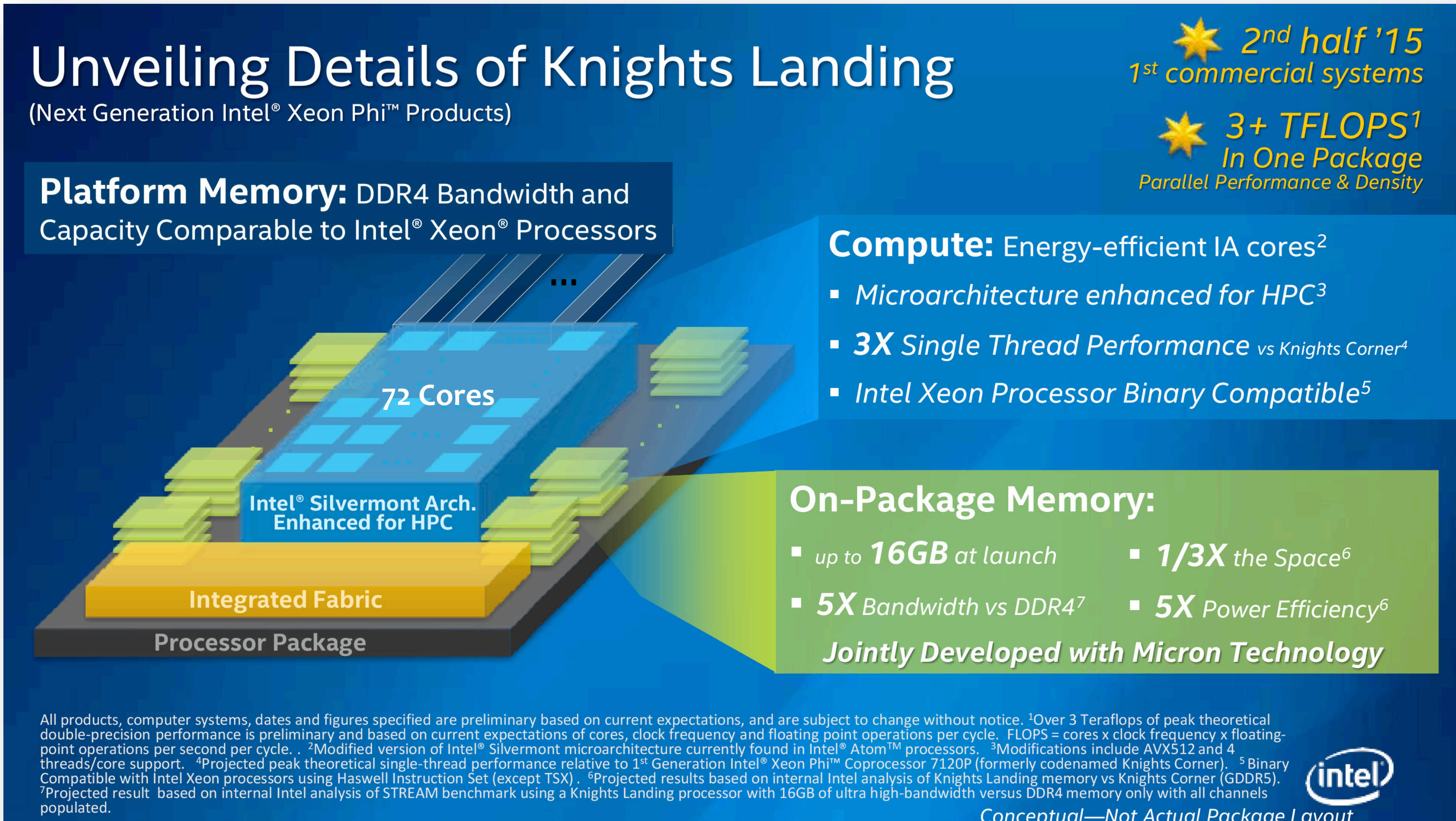
Future Leading-edge Supercomputers

Oak Ridge's Summit & Argonne's Aurora Systems

	Summit (2018)	Aurora (2018)
Processor	IBM Power9/NVIDIA Volta	Intel Knights Hill
Peak Performance	>150 PF	180 PF
Cores/Processor	Up to 24	>72
Number of Nodes	~3,400	>50,000
Memory	>1.7 PB	>7 PB
Interconnect BS Bandwidth	?	>500 TB/s
File System Capacity	~120 PB	>150 PB
File System Bandwidth	~1 TB/s	>1 TB/s
Peak Power	~ 10 MW	13 MW

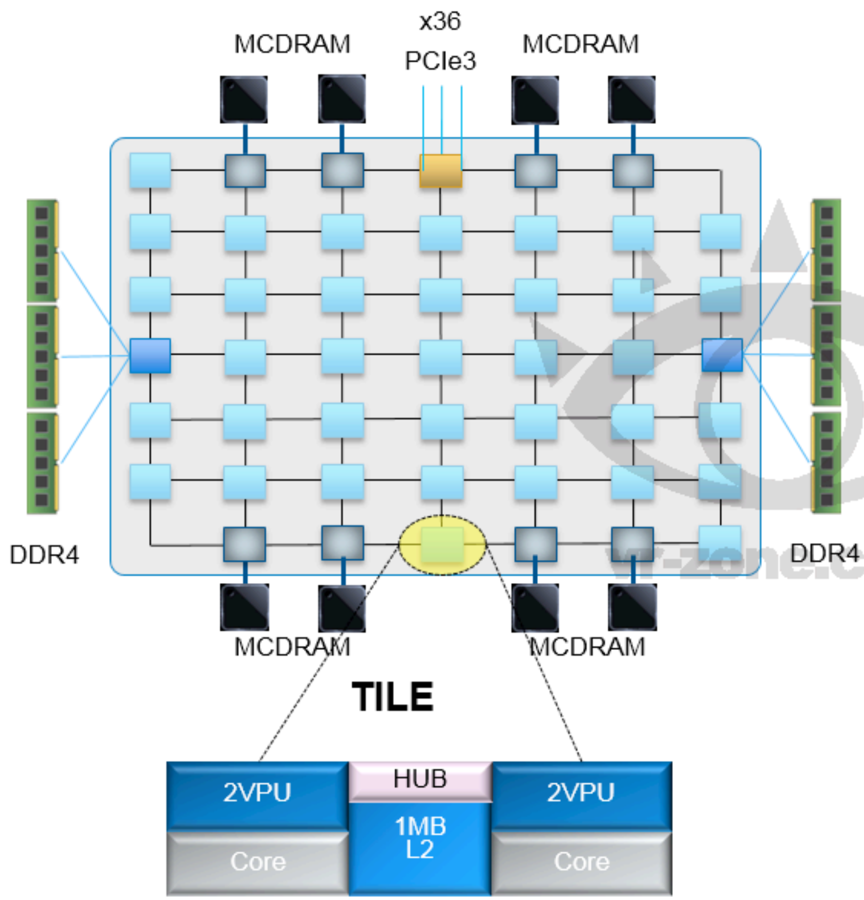
Future Leading-edge Supercomputers

Intel's Many Integrated Core (MIC) Processors



Future Leading-edge Supercomputers

Knights Landing Processor Architecture



Up to 72 Intel Architecture cores based on Silvermont (Intel® Atom processor)

- Four threads/core
- Two 512b vector units/core
- Up to 3x single thread performance improvement over KNC generation

Full Intel® Xeon processor ISA compatibility through AVX-512 (except TSX)

6 channels of DDR4 2400 MHz -up to 384GB

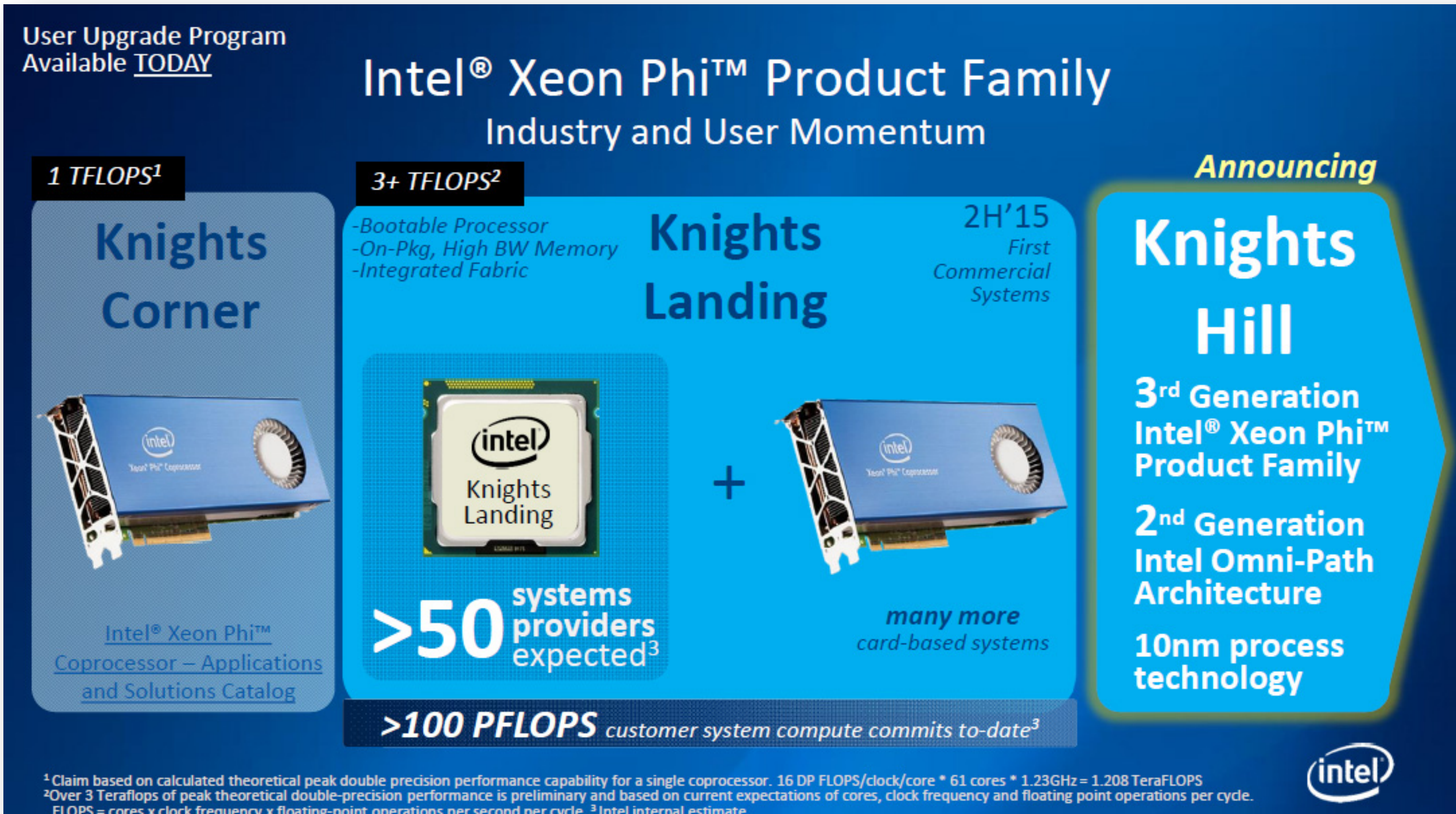
36 lanes PCI Express* Gen 3

8/16GB of high-bandwidth on-package MCDRAM memory >500GB/sec

200W TDP

Future Leading-edge Supercomputers

Knights Hill Processor



Future Leading-edge Supercomputers

International Efforts

- China
  - Dramatic push in supercomputing
    - Now has 1/3-rd of the supercomputers on Top500
  - Building domestic HPC ecosystem (hardware, software, systems)
  - Continuing series of #1 systems being built by China
    - 2011: Tianhe-1A (4.7 PFs); 2013: Tianhe-2 (55 PFs); 2016: Sunway (125 PFs); ...
- Japan
  - Flagship 2020 Project: Post-K computer development
  - RIKEN AICS with Fujuitsu + ARM (with HPC extensions)
  - Target: 50-100x K computer

Future Leading-edge Supercomputers

International Efforts

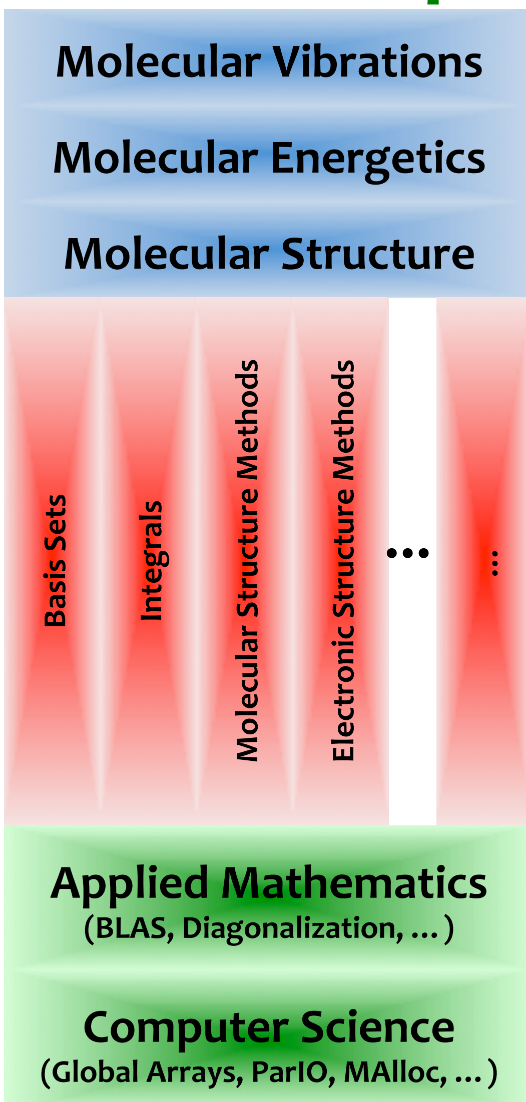
- Europe
  - PRACE
    - 25 members, 2 observers
    - Major efforts in Spain, France, Germany, Italy

Supercomputing Applications

The true value of any computing technology is measured by the applications that it enables. As we saw in comparing Blue Waters & Titan, enabling applications to take full advantage of new computing technologies is a challenging task.

Supercomputing Applications

NWChem: An Exemplary SC Application



Deep collaboration between computational chemists, applied mathematicians, and computer scientists

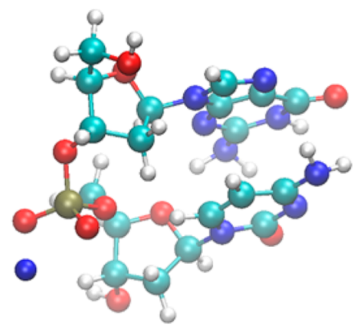
Performance: Another Cautionary Tale

NWChem can achieve impressive performance on petascale computers for the most flop-intensive calculations. For example, for CCSD(T) calculations, which is the current "gold" standard in electronic structure theory, this is the (T) algorithm:

Method	Time(s)*	GFLOP Count	PF/s
CCSD			
(T)	5024	5,948,249,197	1.18
CCSD(T)			

* On 20,000 XE6 nodes (Blue Waters)

V. M. Anisimov, G. H. Bauer, K. Chadalavada, R. M. Olson, J. Glenski, W. T. C. Kramer, E. Aprà, and K. Kowalski, *J. Chem. Theory Comput.* **10**, 4307-4316 (2014).



guanine− cytosine deoxydinucleotide monophosphate + Na^+

However, this is only part of the story. One needs the CCSD amplitudes for the (T) algorithm. The CCSD algorithm is far more complex with a much higher communication/compute ratio than the (T) algorithm:

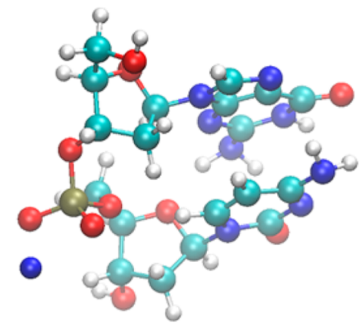
Supercomputing Applications

Performance: Another Cautionary Tale II

Method	Time(s)*	GFLOP Count	PF/s
CCSD	14,406	195,796,351	0.01
(T)	5024	5,948,249,197	1.18
CCSD(T)	19,430	6,144,045,548	0.32

* On 20,000 XE6 nodes (Blue Waters)

V. M. Anisimov, G. H. Bauer, K. Chadalavada, R. M. Olson, J. Glenski, W. T. C. Kramer, E. Aprà, and K. Kowalski, *J. Chem. Theory Comput.* **10**, 4307-4316 (2014).



guanine− cytosine deoxydinucleotide monophosphate + Na^+

So, the CCSD algorithm consumes ¾-th of the time. Further, the algorithm uses a substantial amount of memory, duplicating arrays to minimize communication costs, which limits the number of cores/node that can be used—just 1 of 16 cores on a Blue Waters node that has 64 GBs of memory on the node.

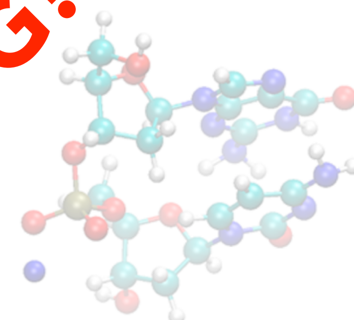
Supercomputing Applications

Performance of NWChem on Blue Waters II

Method	Time(s)*	GFLOP Count	PF/s
CCSD	14,406	195,796,351	0.01
(T)	5024	5,948,249,197	1.18
CCSD(T)	19,430	6,144,045,548	0.32

* On 20,000 XE6 nodes (Blue Waters)

V. M. Anisimov, G. H. Bauer, K. Chadalavada, R. M. Olson, J. Glenski, W. T. C. Kramer, E. Aprà, and K. Kowalski, *J. Chem. Theory Comput.* 10, 4307-4316 (2014).



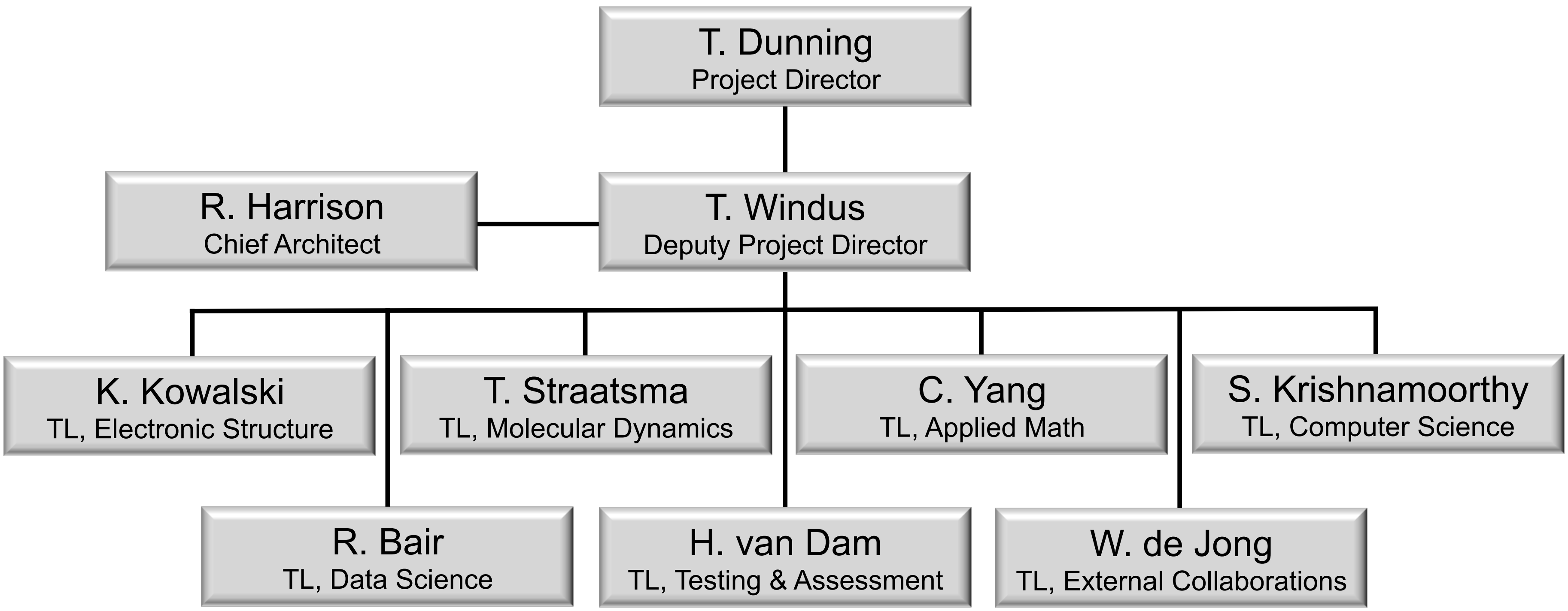
guanine− cytosine deoxydinucleotide monophosphate + Na^+

So, the CCSD algorithm consumes ¾-th of the time. Further, the algorithm uses a substantial amount of memory, duplicating arrays to minimize communication costs, which limits the number of cores/node that can be used—just 1 of 16 cores on a Blue Waters node that has 64 GBs of memory on the node.

THIS IS THE CHALLENGE OF EXASCALE COMPUTING!

Supercomputing Applications

NWChemEx ECP Project



- Redesign NWChem to create a more modular and library-oriented framework
- Incorporate new mathematical algorithms to reduce complexity and improve scalability
- Incorporate new computer science approaches and technologies to reduce/more effectively use memory, separate the details of the hardware from the software
- Identify, assess, and implement new developments in computational chemistry

Thoughts on What Needs to be Done

*As computing technology continues to advance—**and change**—how do we ensure that computational science and engineering continues to be able to take advantage of these advances?*

From Here to the Future

What is Needed to Continue to Advance

- **Better Understanding of Software-Hardware Interface**
  - Current benchmarks do not represent full range of applications
  - Current benchmarks provide little information on programmability
- **New Algorithms**
  - To fully exploit increasing concurrency
  - That are adapted to:
    - Decreased memory per flop
    - Decreased interconnect bandwidth per flop
    - Decreased I/O bandwidth per flop
- **New Programming Models**
  - Need better programming languages, domain specific languages
- **Better Education and Training**

Thank You!