Oak Ridge Leadership Computing Facility

20 Years of Excellence in Computational Science

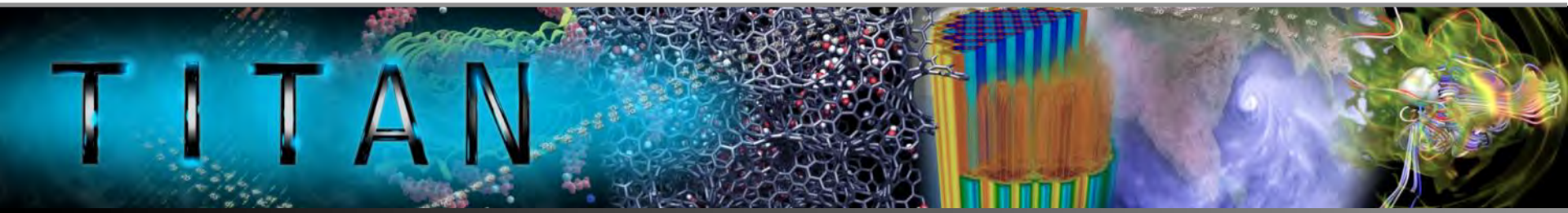
OLCF

Oak Ridge Leadership Computing Facility

1992–2012

Jack C. Wells
Director of Science
National Center for Computational Sciences
Oak Ridge National Laboratory

MAGIC Meeting, 2 April 2014



Outline:

- Introduction to Leadership Computing Program
- Overview of the Titan (OLCF-3) Project
- Support Infrastructure @ OLCF
- Introduction to the CORAL Procurement
- Integration of Compute and Data: SNS & OLCF

What is the Leadership Computing Facility (LCF)?

- Collaborative DOE Office of Science program at ORNL and ANL
- Mission: Provide the computational and data resources required to solve the most challenging problems.
- 2-centers/2-architectures to address diverse and growing computational needs of the scientific community
- Highly competitive user allocation programs (INCITE, ALCC).
- Projects receive 10x to 100x more resource than at other generally available centers.
- LCF centers partner with users to enable science & engineering breakthroughs (Liaisons, Catalysts).

OAK RIDGE National Laboratory
Argonne NATIONAL LABORATORY

The OLCF has delivered five systems and six upgrades to our users since 2004

- Increased our system capability by 10,000x
- Strong partnerships with computer designers and architects
- Worked with users to scale codes by 10,000x
- Science delivered through strong user partnerships to scale codes and algorithms

Phoenix X1
• Doubled size
• X1e

2004

Jaguar XT3
• Dual core upgrade

2005

Jaguar XT4
• Quad core upgrade

2007

Jaguar XT5
• 6 core upgrade

2008

Titan XK7
• GPU upgrade

2012

Science breakthroughs at the OLCF:

SELECTED science and engineering advances over the period 2003 - 2013

Researchers solved the 2D Hubbard model and presented evidence that it predicts HTSC behavior, ***Phys. Rev. Lett*** *(2005)*
105 citations, 3/2014

First-Principles Flame Simulation Provides Crucial Information to Guide Design of Fuel-Efficient Clean Engines, ***Proc. Combust. Insti.*** (2007)
78 citations, 3/2014

Largest simulation of a galaxy's worth of dark matter, showed for the first time the fractal-like appearance of dark matter substructures, ***Nature*** (2008)
326 citations, 3/2014

Astrophysicists discover supernova shock-wave instability, ***Astrophys. J.*** *(2003)*
254 citations, 3/2014

World's first continuous simulation of 21,000 years of Earth's climate history, ***Science*** *(2009)*
116 citations

Biomass as a viable, sustainable feedstock for hydrogen production for fuel cells, ***Nano Letters*** *(2011)*
J. Phys. Chem. Lett. *(2010)*
71 & 74 citations, respectively

MD simulations show selectivity filter of a trans-membrane ion channel is sterically locked open by hidden water molecules, ***Nature*** *(2013)*

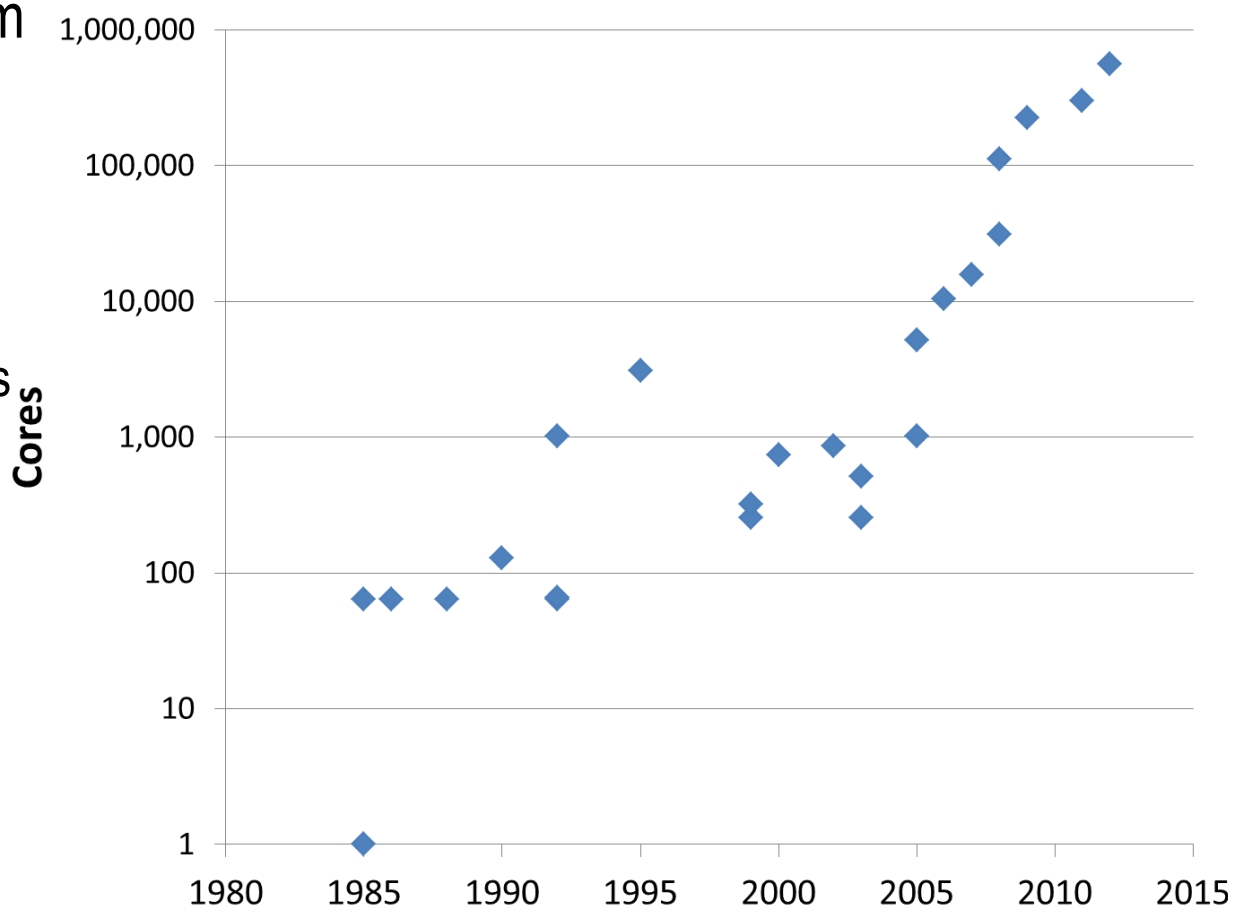
Calculation of the number of bound nuclei in nature, ***Nature*** *(2012), 36 citations, 3/2014 ,* 36 citations, 3/2014

Global Warming preceded by increasing carbon dioxide concentrations during the last deglaciation, ***Nature*** *(2012).*
64 citations, 3/2014

Demonstrated that three-body forces are necessary to describe the long lifetime of ^{14}C
Phys. Rev. Lett. (2011)
28 citations, 3/2014

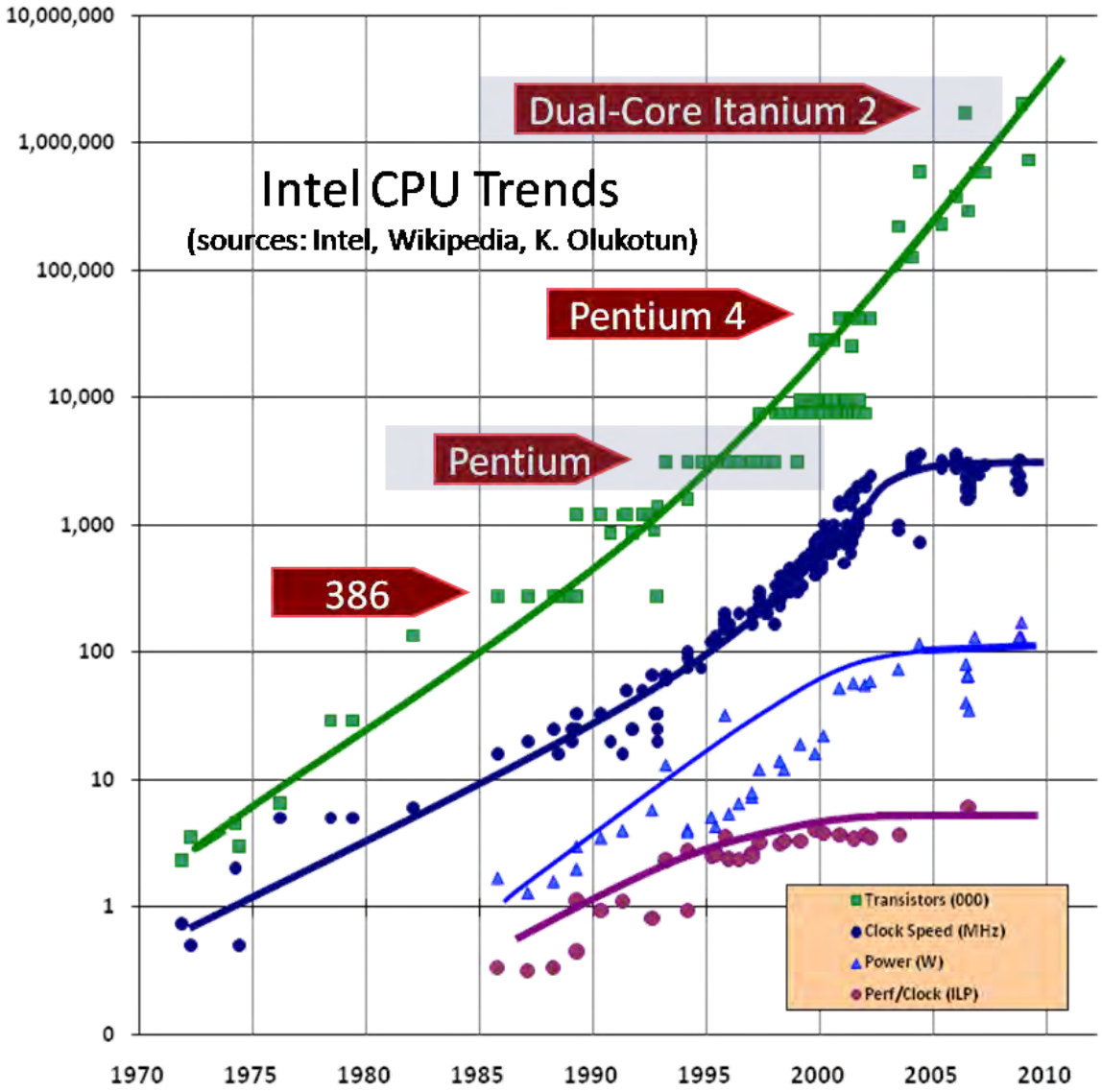
Hardware Trend of ORNL's Systems 1985 - 2013

- In the last 28 years, our systems have scaled from 64 cores to hundreds of thousands of cores and millions of simultaneous threads of execution
  - Multiple hierarchical levels of parallelism
  - Hybrid processors and systems
- The last 28 years of application development have been about finding ways to exploit that parallelism!



Architectural Trends – No more free lunch

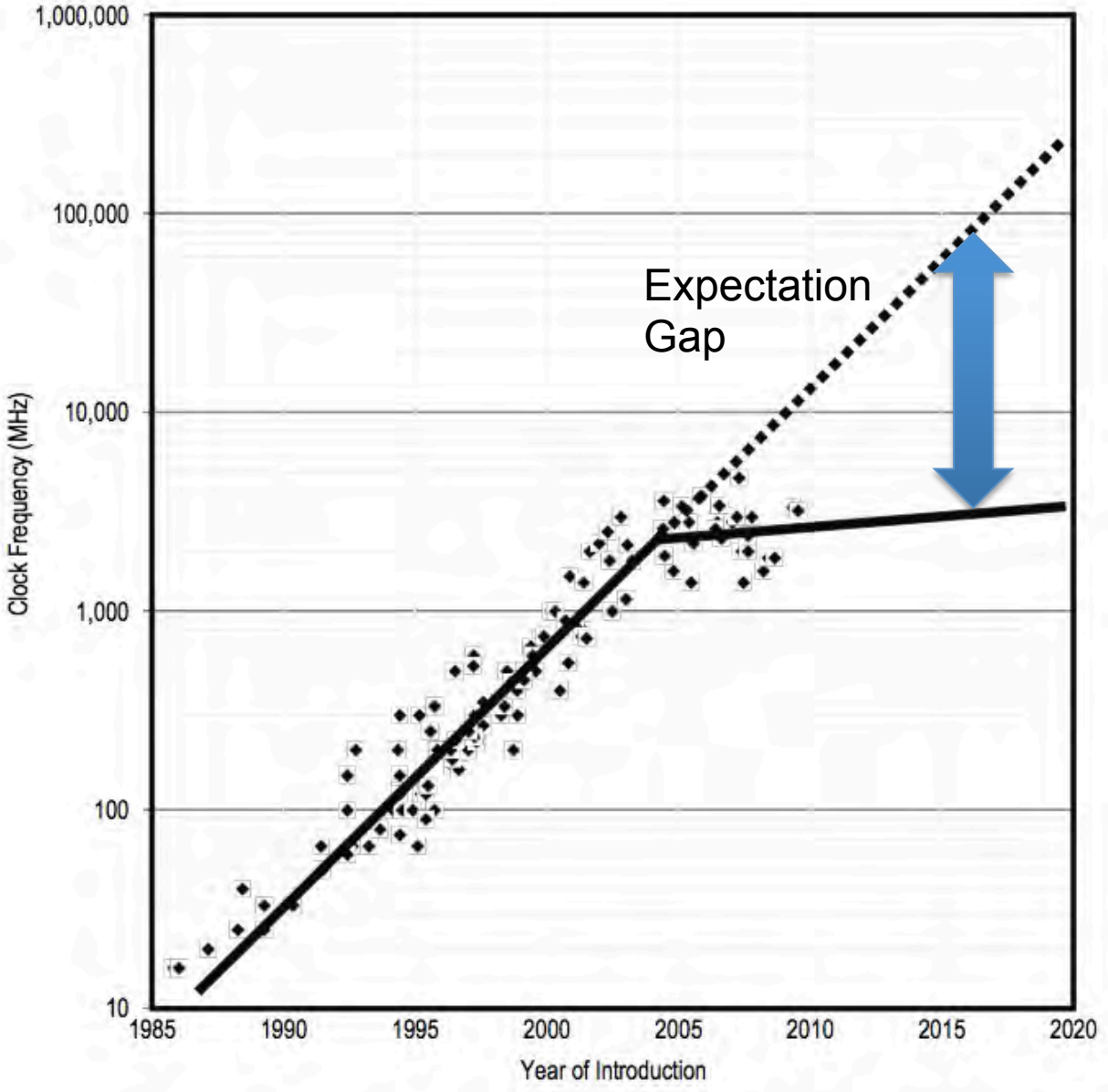
- Moore's Law continues (green line)
- But CPU clock rates stopped increasing in 2003 (dark blue line)
- Power (light blue line) is capped by heat dissipation and $$$
- Single-thread performance is growing slowly (magenta line)



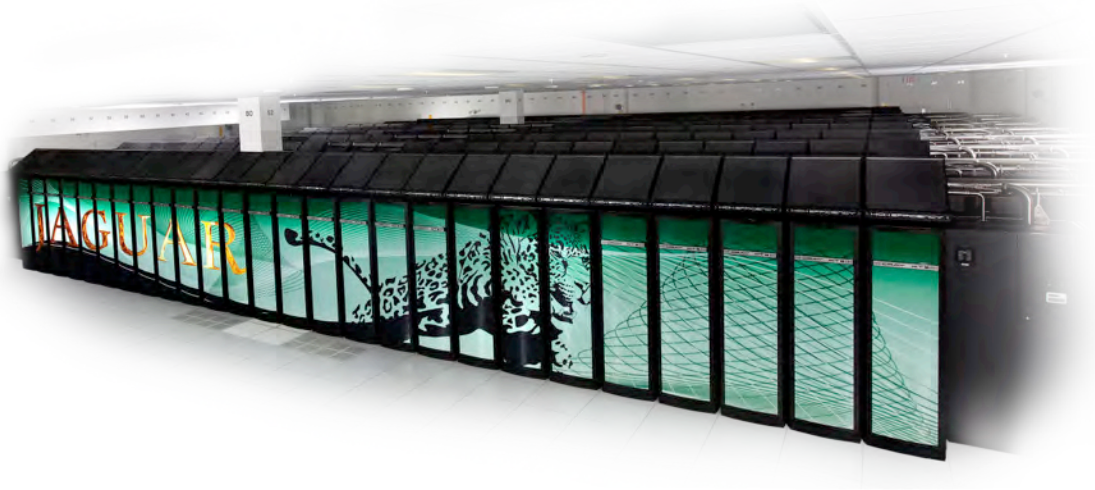
Herb Sutter: Dr. Dobb's Journal:
http://www.gotw.ca/publications/concurrency-ddj.htm

Microprocessor Performance "Expectation Gap"

National Research Council (NRC) – Computer Science and Telecommunications Board (2012)



Power is THE problem



Power consumption of 2.3 PF (Peak) Jaguar:
7 megawatts, equivalent to that of a small city (5,000 homes)

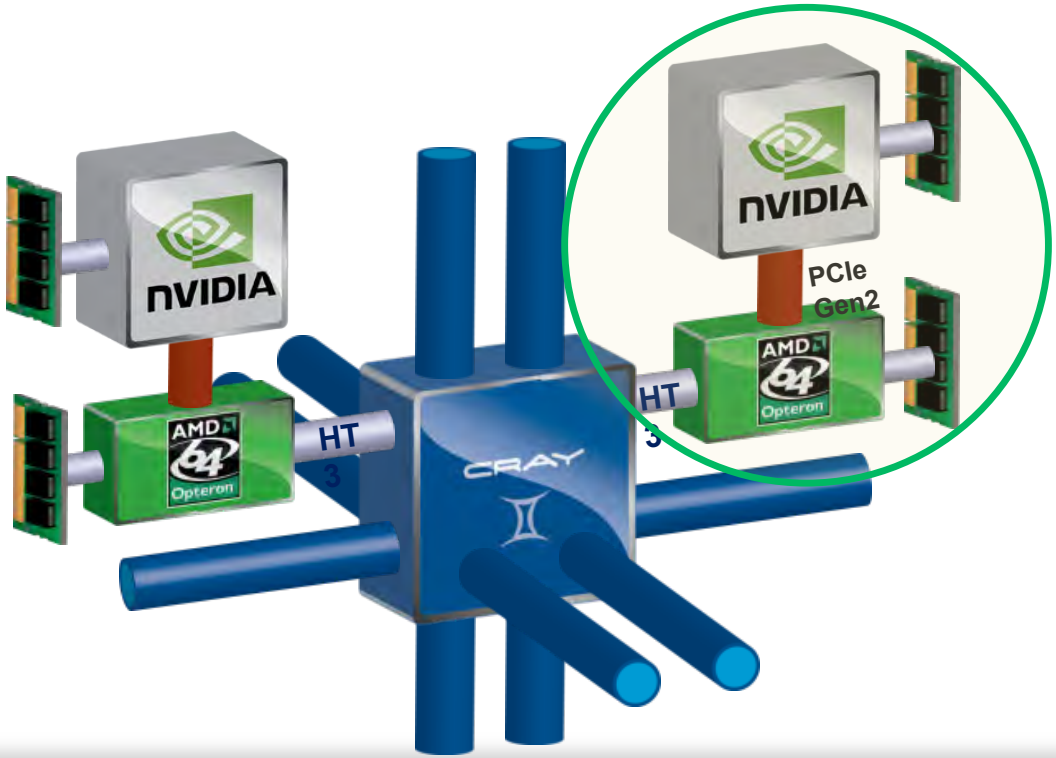
Peak FLOPS per Node was #1 Hardware Requirement in 2009 User Survey

System Attribute	Climate	Astrophysics	Fusion	Chemistry	Combustion	Accelerator physics	Biology	Materials science
Node peak flops								
MTTI								
WAN network bandwidth								
Node memory capacity								
Local storage capacity								
Archival storage capacity								
Memory latency								
Interconnect latency								
Disk latency								
Interconnect bandwidth								
Memory bandwidth								
Disk bandwidth								

(Priority: Red = high, pink = medium, grey = low.)

"Preparing for Exascale" OLCF Application Requirements and Strategy, December 2009

Titan Compute Nodes (Cray XK7)			
Node	AMD Opteron 6200 Interlagos (16 cores)	2.2 GHz	32 GB (DDR3)
Accelerator	Tesla K20x (2688 CUDA cores)	732 MHz	6 GB (DDR5)



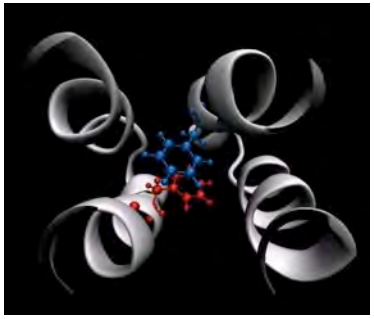
#2 TOP500 SUPERCOMPUTER SITES

Titan System (Cray XK7)			
Peak Performance	27.1 PF 18,688 compute nodes	24.5 PF GPU	2.6 PF CPU
LINPACK Performance	17.59 PF		
Power	8.2 MW		
System Memory	710 TB total memory		
Interconnect	Gemini High Speed Interconnect	3D Torus	
Storage	Luster Filesystem	32 PB	
Archive	High-Performance Storage System (HPSS)	29 PB	
I/O Nodes	512 Service and I/O nodes		

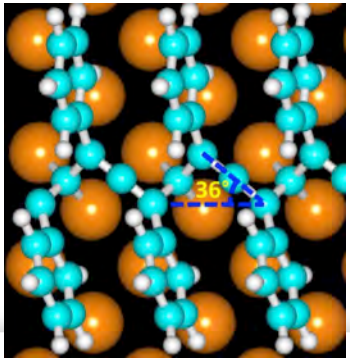
High-impact science across a broad range of disciplines

For example in 2013:

Paleoclimate Science
"Northern Hemisphere forcing of Southern Hemisphere climate during the last deglaciation,"
Feng He (UW Madison), *et al*., *Nature*, February (2013)

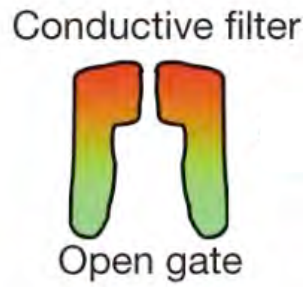


Molecular Biology
"A phenylalanine rotameric switch for signal-state control in bacterial chemoreceptors"
D. Ortega (UTK),
Nature Communications
December (2013)

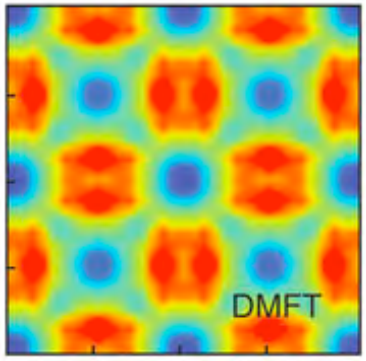


Polymer Science
"Self-Organized and Cu-Coordinated Surface Linear Polymerization"
Qing Li, B. Sumpter (ORNL),
Nature Scientific Reports.
July (2013)

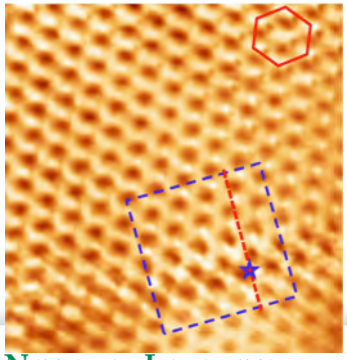
Molecular Biology
MD simulations show selectivity filter of a trans-membrane ion channel is sterically locked open by hidden water
Jared Ostmeyer, et al. (U. Chicago) *Nature*, Sept. (2013)



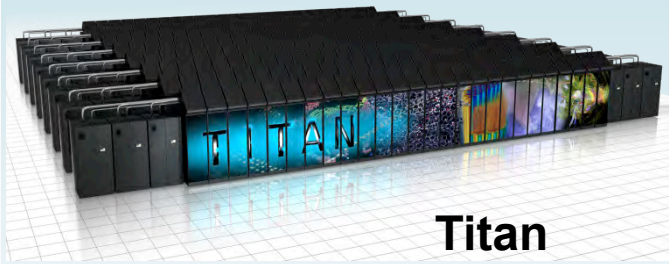
Superconductivity
"Doping dependence of spin excitations and correlations with high-temperature super-conductivity in iron pnictides,"
Meng Wang(IOP CAS Beijing),
Nature Communications.
December (2013)



Complex Oxide Materials
"Atomically resolved spectroscopic studyof Sr2IrO4: Experiment and theory," Qing Li (ORNL), E.G. Eguiluz (UTK)
Nature Scientific Reports.
October (2013)



The Oak Ridge Leadership Computing Facility provides a unique computational user facility for our user community



Titan
Cray XK7

Peak performance	27 PF/s
Memory	710 TB
Disk bandwidth	1 TB/s
Square feet	5,000
Power	8.8 MW

Eos
Cray XC30

Peak performance	248 TF/s
Memory	48 TB
Disk bandwidth	20 GB/s
Square feet	108

Data Storage

- Spider File System
  - 40 PB capacity
  - 1+ TB/s bandwidth
- HPSS Archive
  - 240 PB capacity
  - 6 Tape libraries

Data Analytics & Visualization

- Ewok cluster
- Rhea cluster
- FOCUS cluster
- EVEREST visualization facility
- uRiKA data appliance

Networks

- ESnet – 100 Gbps
- Internet2 – 10 Gbps
- XSEDEnet – 10 Gbps
- Private dark fibre

Who we are

Chief Computational Scientist: Ricky Kendall

Chief Technology Officer: Al Geist

Director of Science: Jack Wells

Director of Operations: Jim Rogers

National Center for Computational Sciences
Oak Ridge Leadership Computing Facility
James Hack, Director

Buddy Bland, OLCF Project Director

Kathlyn Boudwin, Deputy Project Director

Data Systems Architect: Galen Shipman

Industrial Partnerships: Suzy Tichenor

INCITE Programs: Julia White

Computer Science Research
David Bernholdt

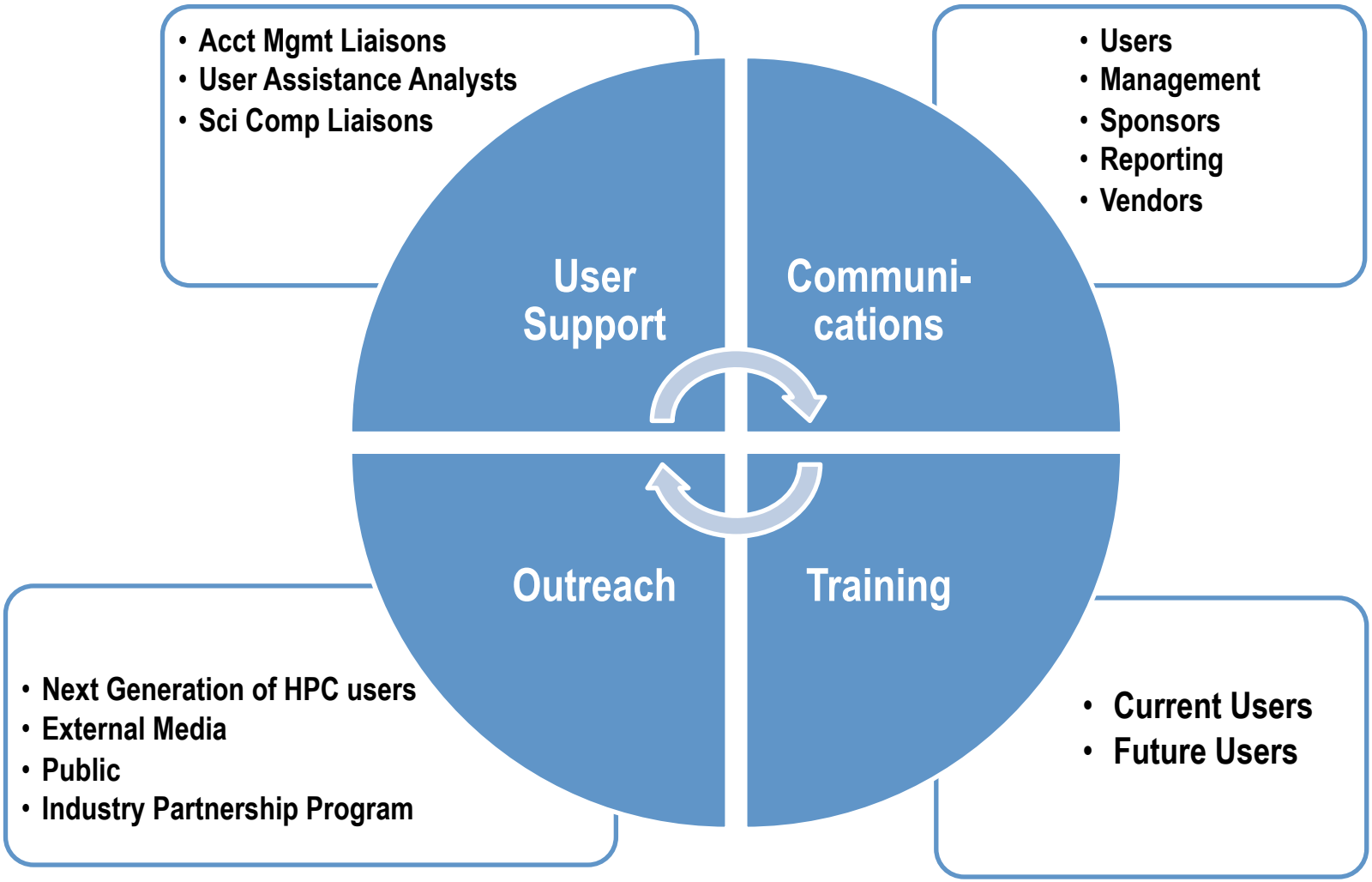
High-Performance Computing Operations
Kevin Thach

Scientific Computing
Tjerk Straatsma

Technology Integration
Sudharshan Vazhkudai

User Assistance and Outreach
Ashley Barker

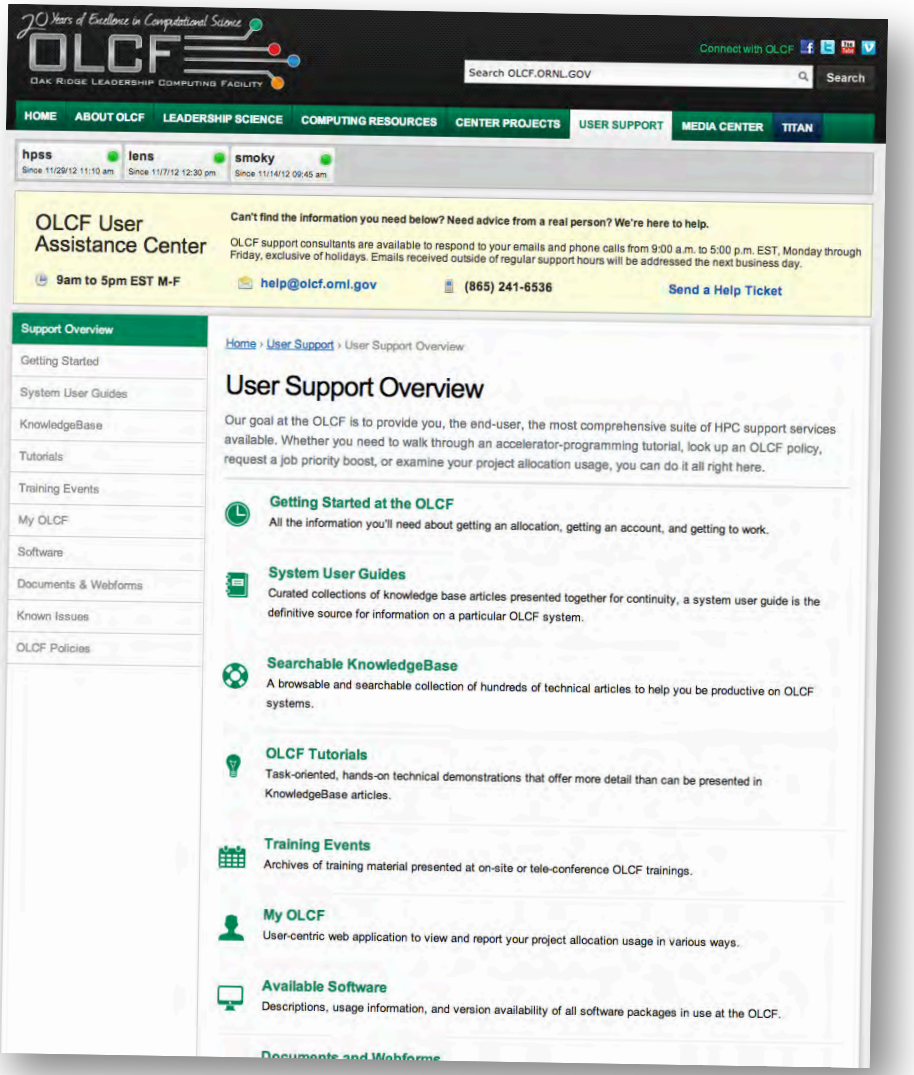
User Assistance and Outreach



OLCF Support Site

- All support information available at:

  http://olcf.ornl.gov/support/
  - Getting started
  - User guides
  - Tutorials
  - Software inventory
  - Knowledge base
  - Known issues
  - Official policies
  - Past & future events



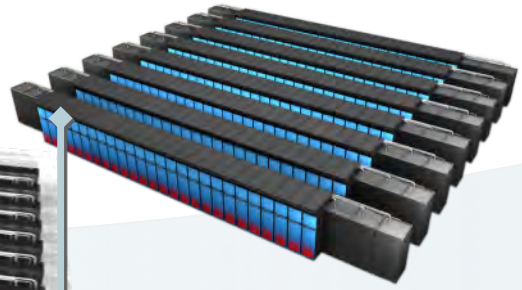
Our Science requires that we advance computational capability 1000x over the next decade.

Mission: Providing world-class computational resources and specialized services for the most computationally intensive global challenges	Vision: Deliver transforming discoveries in climate, materials, biology, energy technologies, etc

Roadmap to Exascale

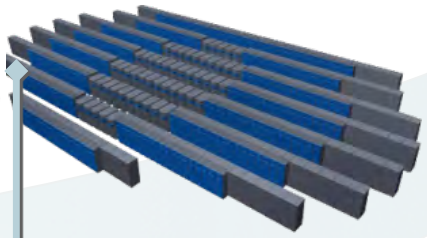
Jaguar 2.3 PF
362 TB DRAM

2010



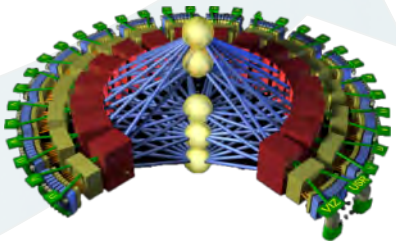
Titan 27 PF
600 TB DRAM
Hybrid GPU/CPU

2012



OLCF-4: 100-250 PF
4000 TB memory
> 20MW
6 day resilience

2017



OLCF-5: 1 EF
20 MW

What are the Challenges?

2022

What is CORAL (Partnership for 2017 System)

- CORAL is a Collaboration of Oak Ridge, Argonne, and Lawrence Livermore Labs to acquire three systems for delivery in 2017.
- DOE's Office of Science (DOE/SC) and National Nuclear Security Administration (NNSA) signed an MOU agreeing to collaborate on HPC research and acquisitions
- Collaboration grouping of DOE labs was done based on common acquisition timings. Collaboration is a win-win for all parties.
  - It reduces the number of RFPs vendors have to respond to
  - It improves the number and quality of proposals
  - It allows pooling of R&D funds
  - It strengthens the alliance between SC/NNSA on road to exascale
  - It encourages sharing technical expertise between Labs

CORAL Joint NNSA & SC Leadership Computing Acquisition Project

Current DOE Leadership Computers

Titan (ORNL)
2012 - 2017

Sequoia (LLNL)
2012 - 2017

Mira (ANL)
2012 - 2017

Objective - Procure 3 leadership computers to be sited at ANL, ORNL and LLNL in CY17

Leadership Computers run the most demanding DOE mission applications and advance HPC technologies to assure continued US/DOE leadership

Approach

Competitive process - one RFP (issued by LLNL) leading to 2 R&D contracts and 3 computer procurement contracts

For risk reduction and to meet a broad set of requirements, 2 architectural paths will be selected

Once Selected, Multi-year Lab-Awardee relationship to co-design computers

Both R&D contracts jointly managed by the 3 Labs

Each lab manages and negotiates its own computer procurement contract, and may exercise options to meet their specific needs

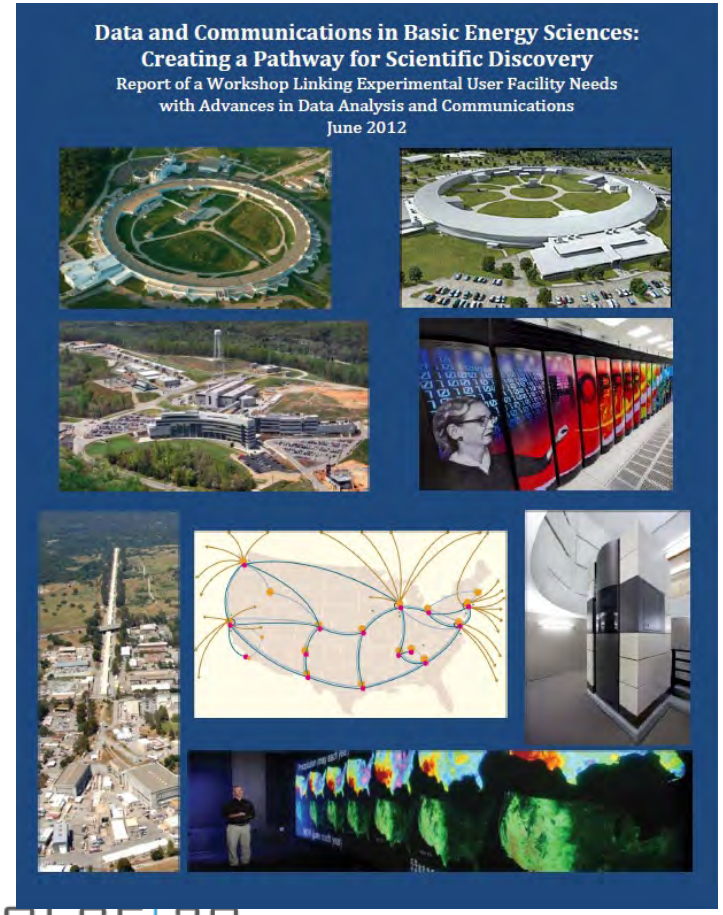
Understanding that long procurement lead-time may impact architectural characteristics and designs of procured computers

CORAL Procurement Model

Two Diverse Architecture Paths

Integrating Compute & Data Capabilities: Creating a Pathway for Scientific Discovery

- Accelerating discovery in neutron sciences
- Enhancing predictive capabilities



- Theory and analysis components should be integrated seamlessly within experimental workflow.
- Move analysis closer to experiment – future possibility of experiment steering.
- Match data management access and capabilities with advancements in detectors and sources.

ORNL has key strengths to address the materials challenge

High Flux Isotope Reactor: Intense steady-stae neutron flux and a high-brightness cold neutron source

Spallation Neutron Source: World's most powerful accelerator-based neutron source

UT-ORNL Joint Institute for Neutron Sciences: User gateway for SNS and HFIR

Titan – Worlds fastest supercomputer

Center for Nanophase Materials Science

ADARA

Enabling real-time feedback from experiment, analysis and computational steering

SNS offers mission-critical capabilities in materials science through neutron diffraction and inelastic neutron scattering at the world's most intense pulsed neutron beam

- ***Challenge***: Realizing the full potential of the SNS requires near real-time feedback to users as the experiment is run and integration of experiment and simulation/modeling
  - Data intensive computing techniques pioneered within CCSD at ORNL can close this gap
- ***Response***: Accelerating Data Acquisition, Reduction and Analysis
  - The ADARA project is a joint CCSD/NScD initiative (PI: Galen Shipman) that has developed a streaming data infrastructure for real-time experiment feedback and instant access to neutron experiment datasets
  - Provides a high-performance data streaming system for SNS forming the basis for future work to integrate experiment and modeling/simulation
- ***Status:*** ADARA is up and running at the SNS
  - Running today on HYSPEC and SEQUOIA, commissioning on VISION, CORELLI, USANS
  - Continued development and deployment of ADARA on subsequent beam lines at SNS

CADES

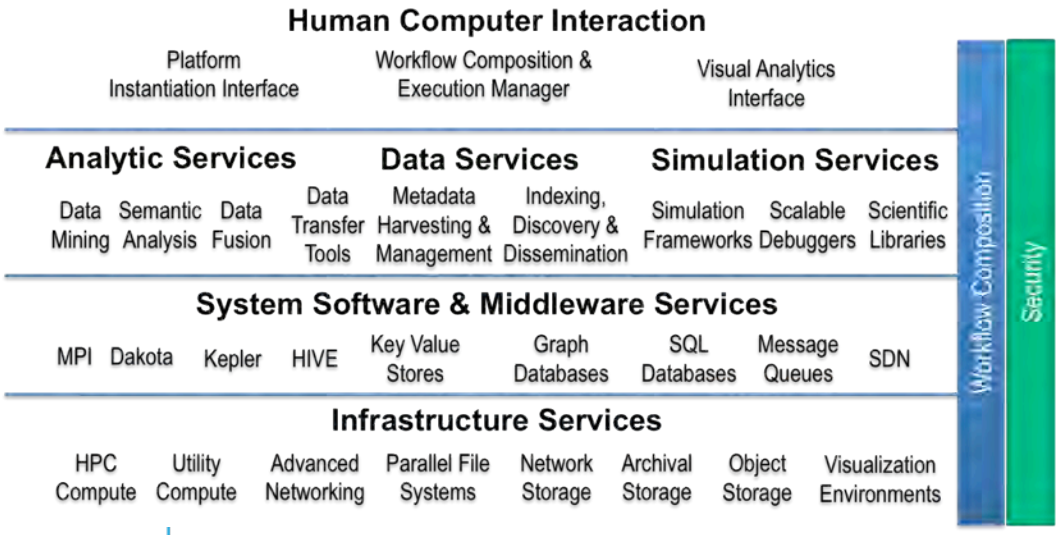
Compute & Data Environment for Science

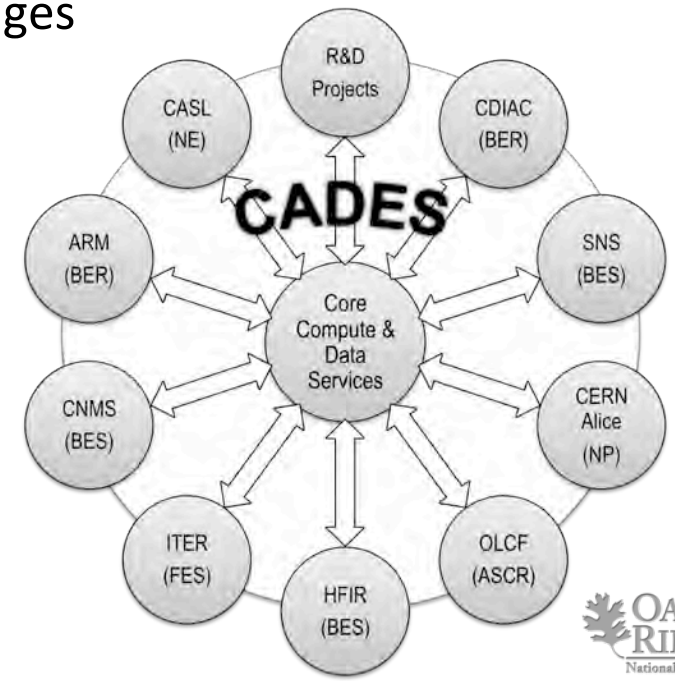
OAK RIDGE National Laboratory

We have established CADES to provide core compute and data services such as those required by ADARA and CAMM projects to major facilities and programs

CADES is a cross-cutting center: it shares both data infrastructure and compute & data science expertise with and among many projects

A rich set of flexibly composable services coupled with experts in data science partnering with domain scientists on their challenges





OAK RIDGE National Laboratory

Conclusions

- Leadership computing facilities for the critically important problems that need the most powerful compute and data infrastructure
- OLCF provides comprehensive support to our user programs in scientific computing, user assistance and outreach, technology integration, and computer science.
- New and exciting opportunities are availed by integrated large experimental facilities with supercomputing facilities.

Acknowledgements

OLCF Users

OLCF Staff

OLCF-3 Vendor Partners: Cray, AMD, NVIDIA, CAPS, Allinea

This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.