Titan: A Heterogeneous Supercomputer for Leadership Science

Buddy Bland
Project Director
Oak Ridge Leadership Computing Facility

Jack Wells
Director of Science
Oak Ridge Leadership Computing Facility

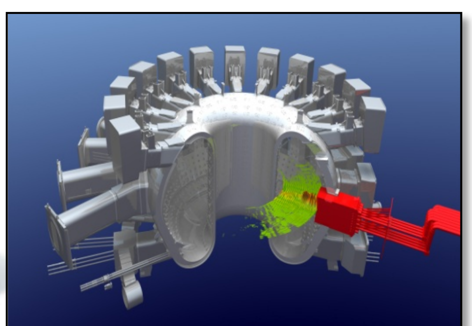
Julia White
INCITE Program Manager

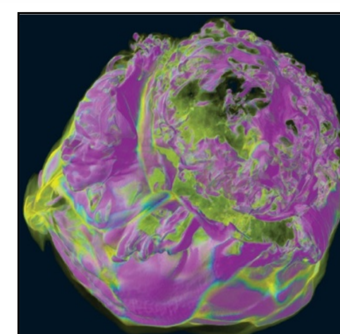
20 Years of Excellence in Computational Science
OLCF
Oak Ridge Leadership Computing Facility
1992–2012

NVIDIA Technology Briefings at ISC 2012
Hamburg, DE
June 19, 2012

Titan System Goals: Deliver breakthrough science for DOE/SC, industry, and the nation

Geosciences	Energy	Fundamental Science
Understanding our earth and the processes that impact it – Sea level rise – Regional climate change – Geologic carbon sequestration – Biofuels – Earthquakes and Tsunamis	**Reducing U.S. reliance on foreign energy & reducing carbon footprint of production** – Carbon free energy production from fusion, fission, solar, wind, and geothermal sources – Improving the efficiency of combustion energy sources	**Understanding the physical processes from the scale of subatomic particles to the universe** – Understanding the makeup of atoms to supernovae – Developing advanced materials for applications such as photovoltaics & electronic components





Accomplishing these missions requires the power of Titan

Titan System Goals: Promote application development for highly scalable architectures through the Center for Accelerated Application Readiness (CAAR)

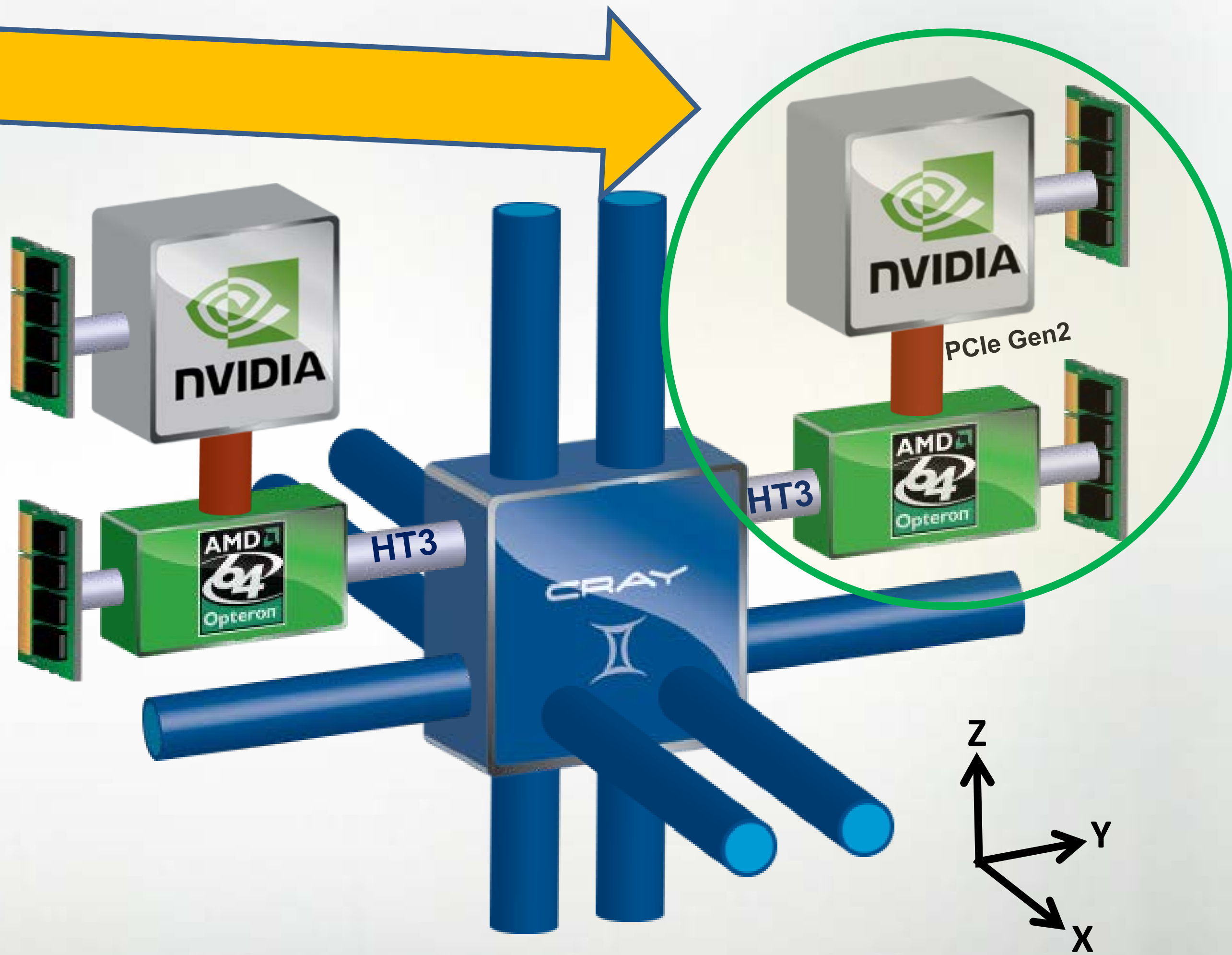
Using six representative apps to explore techniques to effectively use highly scalable architectures

- CAM-SE – Atmospheric model
- Denovo – Nuclear reactor neutron transport
- wl-LSMS - First principles statistical mechanics of magnetic materials
- S3D – Combustion model
- LAMMPS – Molelcular dynamics
- NRDF – Adaptive mesh refinement

- Data locality
- Explicit data management
- Hierarchical parallelism
- Exposing more parallelism through code refactoring and source code directives
- Highly parallel I/O
- Heterogeneous multi-core processor architecture

Cray XK6 Compute Node

XK6 Compute Node Characteristics
AMD Opteron 6200 Interlagos 16 core processor
Tesla X2090 @ 665 GF
Host Memory 16 or 32GB 1600 MHz DDR3
Tesla X090 Memory 6GB GDDR5 capacity
Gemini High Speed Interconnect
Upgradeable to NVIDIA's KEPLER "K20" many-core processor



Slide courtesy of Cray, Inc.

ORNL's "Titan" System

- Upgrade of Jaguar from Cray XT5 to XK6
- Cray Linux Environment operating system
- Gemini interconnect
  - 3-D Torus
  - Globally addressable memory
  - Advanced synchronization features
- AMD Opteron 6274 processors (Interlagos)
- New accelerated node design using NVIDIA multi-core accelerators
  - 2011: 960 NVIDIA x2090 "Fermi" GPUs
  - 2012: 14,592 NVIDIA K20 "Kepler" GPUs
- 20+ PFlops peak system performance
- 600 TB DDR3 mem. + 88 TB GDDR5 mem

Titan Specs	
Compute Nodes	18,688
Login & I/O Nodes	512
Memory per node	32 GB + 6 GB
# of Fermi chips (2012)	960
# of NVIDIA K20 "Kepler" processor (2013)	14,592
Total System Memory	688 TB
Total System Peak Performance	20+ Petaflops
Cross Section Bandwidths	X=14.4 TB/s Y=11.3 TB/s Z=24.0 TB/s

Hybrid Programming Model

- On Jaguar today with 299,008 cores, we are seeing the limits of a single level of MPI scaling for most applications
- To take advantage of the vastly larger parallelism in Titan, users need to use hierarchical parallelism in their codes
  - Distributed memory: MPI, SHMEM, PGAS
  - Node Local: OpenMP, Pthreads, local MPI communicators
  - Within threads: Vector constructs on GPU, libraries, CPU SIMD
- *These are the same types of constructs needed on all multi-PFLOPS computers to scale to the full size of the systems!*

How do you program these nodes?

- Compilers
  - OpenACC is a set of compiler directives that allows the user to express hierarchical parallelism in the source code so that the compiler can generate parallel code for the target platform, be it GPU, MIC, or vector SIMD on CPU
  - Cray compiler supports XK6 nodes and is OpenACC compatible
  - CAPS HMPP compiler supports C, C++ and Fortran compilation for heterogeneous nodes and is adding OpenACC support
  - PGI compiler supports OpenACC and CUDA Fortran
- Tools
  - Allinea DDT debugger scales to full system size and with ORNL support will be able to debug heterogeneous (x86/GPU) apps
  - ORNL has worked with the Vampir team at TUD to add support for profiling codes on heterogeneous nodes
  - CrayPAT and Cray Apprentice support XK6 programming

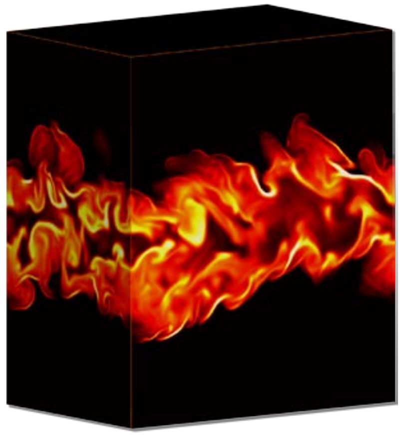
Titan Tool Suite

Compilers	Performance Tools	GPU Libraries	Debuggers	Source Code
Cray PGI CAP-HMPP Pathscale NVIDIA CUDA GNU Intel	CrayPAT Apprentice Vampir VampirTrace TAU HPCToolkit CUDA Profiler	MAGMA CULA Trillinos libSCI	DDT NVIDIA Gdb	HMPP Wizard

Titan: Early Applications & Stretch Goals

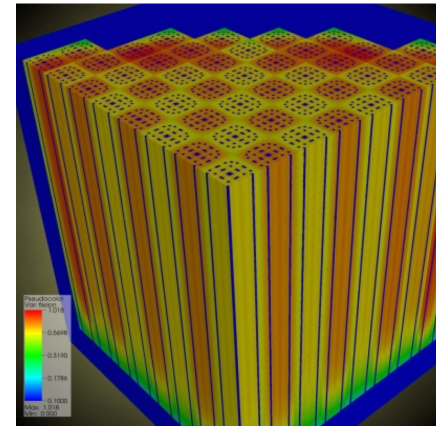
S3D: Turbulent Combustion

Directly solves Navier-Stokes equations. Stretch goals is to move beyond simple fuels to realistic transportation fuels, e.g., iso-octane or biofuels



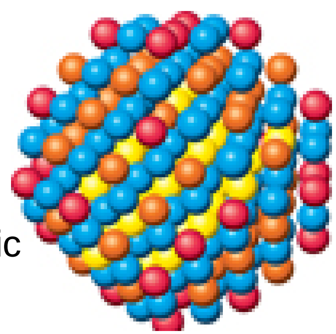
DENOVO: Neutron Transport in Reactor Core

DENOVO is a component of the DOE CASL Hub, necessary to achieve CASL challenge problems



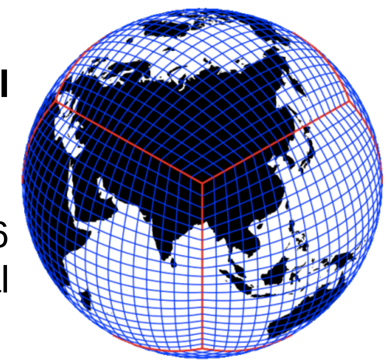
WL-LSMS: Statistical Mechanics of Magnetic Materials

Calculate the free energy for magnet materials. Applications to magnetic recording, magnetic processing of structural materials



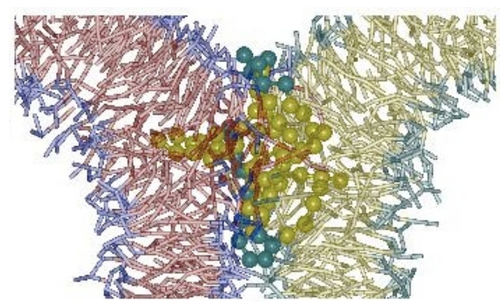
CAM-SE: Community Atmosphere Model – Spectral Elements

CAM simulation using Mozart tropospheric chemistry with 106 constituents at 14 km horizontal grid resolution



LAMMPS: Biological Membrane Fusion

Coarse-grain MD simulation of biological membrane fusion in 5 wall clock days.

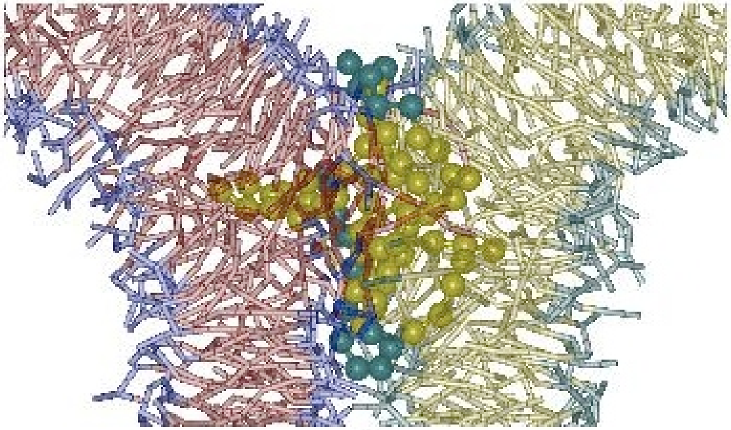


LAMMPS

Large-scale, massively parallel molecular dynamics

Code Description

- Classical N-body problem of atomistic modeling
- Force fields available for chemical, biological, and materials applications
- Long-range electrostatics evaluated using a "particle-particle, particle-mesh" (PPPM) solver.
- 3D FFT in particle-mesh solver limits scaling



Insights into the molecular mechanism of membrane fusion from simulation. Stevens et al., *PRL* 91 (2003)

Porting Strategy

- For PPPM solver, replace 3-D FFT with grid-based algorithms that reduce inter-process communication
- Parallelism through domain decomposition of particle-mesh grid
- Accelerated code builds with OpenCL or CUDA

Early Performance Results on XK6:

- XK6 outperforms XE6 by 3.2x

Science Target (20PF Titan)

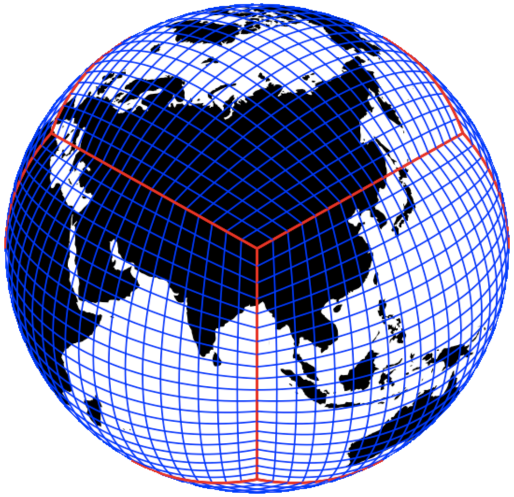
- Simulate biological membrane fusion in coarse-grained MD within 5 wall clock days

CAM-SE

Community Atmosphere Model – Spectral Elements

Code Description

- Employs equal-angle, cubed-sphere grid and terrain-following coordinate system.
- Scaled to 172,800 cores on XT5
- Exactly conserves dry mass without the need for *ad hoc* fixes.
- Original baseline code achieves parallelism through domain decomposition using one MPI task per element



Cubed-sphere grid of CAM spectral element model. Each cube panel is divided into elements.

http://www-personal.umich.edu/~paullric/A_CubedSphere.png

Porting Strategy

- Using realistic “Mozart” chemical tracer network, tracer transport (i.e., advection) dominates the run time.
- Use hybrid MPI/OpenMP parallelism
- Intensive kernels are coded in CUDA Fortran
- Migration in future to OpenACC

Early Performance Results on XK6:

- Refactored code was 1.7x faster on Cray XT5
- XK6 outperforms XE6 by 1.5x

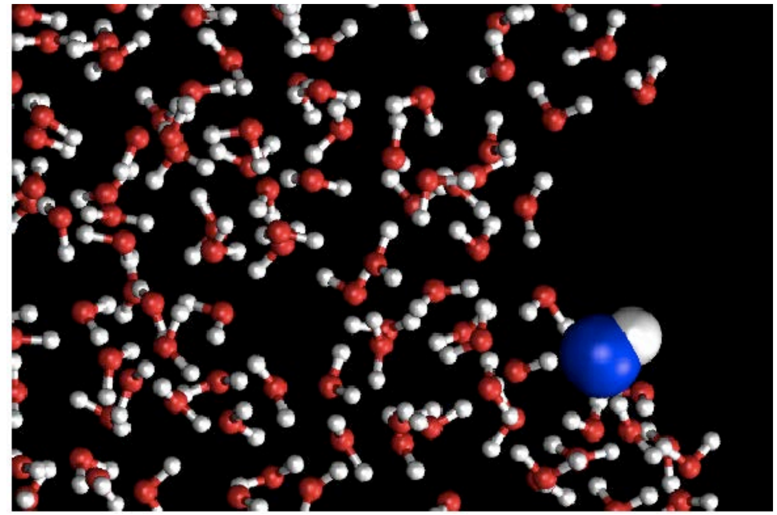
Science Target (20PF Titan)

- CAM simulation using Mozart tropospheric chemistry with 106 constituents at 14 km horizontal grid resolution

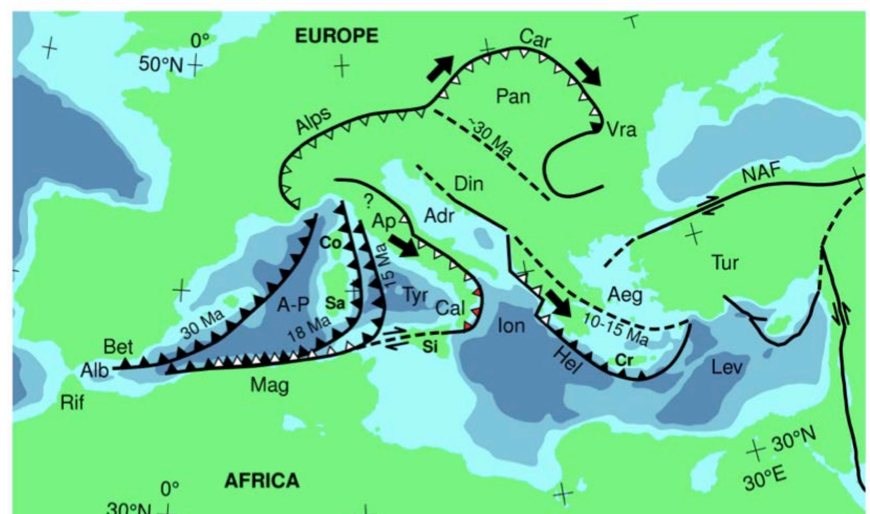
Performance of additional community codes

CP2K: Molecular Simulation

Explore electronic and chemical phenomena in complex systems and materials. Compute the free-energy profiles of individual chemical reactions using ensemble methods to statistically explore the important configurations that give rise to novel phenomena,



Mediterranean-Calabria Paleotectonics



SPECFEM-3D Seismic Imaging

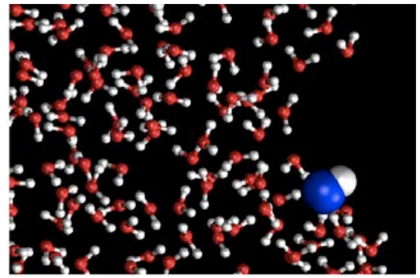
Calculate the free energy for magnet materials. Applications to magnetic recording, magnetic processing of structural materials

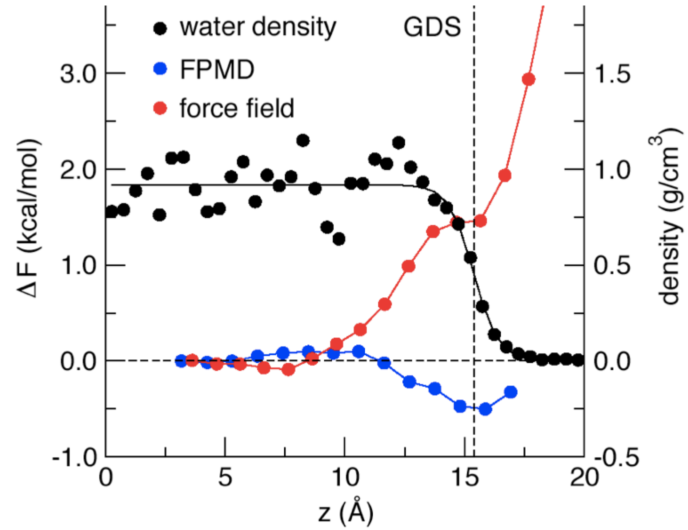
CP2K

"Swiss-army knife for molecular science"

Code Description

- Open source, global user base
  - J. Hutter, J. VandeVondele
- Density-functional theory:
  - Linear-scaling SCF
  - GGA, HFX, & MP2
- Hybrid QM/MM calculations
- Dense & sparse linear algebra
- Regular grid: multigrid and FFT





Is the surface of liquid water is acidic or basic? Mundy et al. show that the hydroxide anion is slightly stabilized at the interface – **Mundy, et al., *Chem. Phys. Letts.* 481, 2 (2009)**

Porting Strategy

(J. VandeVondle, ETH)

- Port sparse linear algebra library
- Matrix multiply kernels are coded in CUDA
- MPI performance is crucial
- Early days, results will improve

Early Performance Results on XK6:

Peter Messmer (NVIDIA)

- XK6 outperforms XE6 by 1.5x

Science Target (20PF Titan)

Chris Mundy (PNNL)

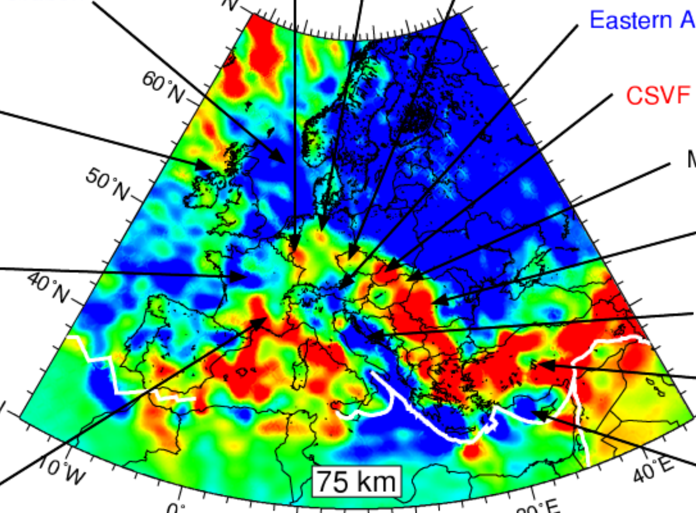
- Ion adsorption at interfaces, for fuel cells, chemical catalysis, biological pumps, atmospheric science

SPECFEM-3D

Seismic imaging

Code Description

- Open source, global user base
- Spectral element wave propagation
- Unstructured grids
- Explicit time propagation
- Full-space waveform inversion
- MPI parallel, excellent scaling
- GB Prize (SC03, SC08 - finalist)



Porting Strategy

(Olaf Schenk - Lugano)

- Force calculation is the heaviest kernel
- Stencil kernels are memory bound, data locality is crucial
- CUDA accelerated

Early Performance Results on XK6:

Peter Messmer (NVIDIA)

- XK6 outperforms XE6 by 2.5x

Science Target (20PF Titan)

Jeroen Tromp (Princeton)

- Adjoint tomography of the entire planet -- requires analysis of 5,000 earthquakes worldwide, using an estimated 739 million core-hours (on Cray XT5)

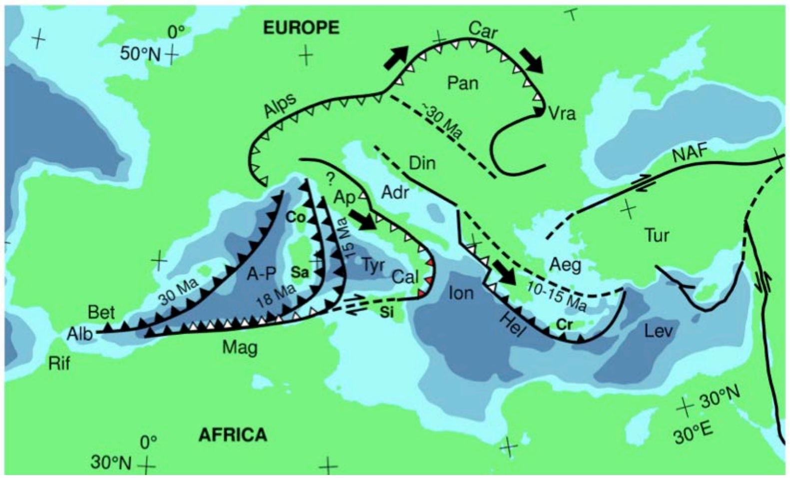
SPECFEM-3D

Seismic imaging

Code Description

- Open source, global user base
- Spectral element wave propagation
- Unstructured grids
- Explicit time propagation
- Full-space waveform inversion
- MPI parallel, excellent scaling
- GB Prize (SC03, SC08 - finalist)

Mediterranean-Calabria Paleotectonics



Analysis of 190 European quakes revealed seismic hotspots in the upper mantle. The imaging detailed the subduction of Africa, volcanism in the Czech Republic, a “hole” under Bulgaria, and Italy’s counterclockwise rotation over the past 6 million years. J. Tromp, ACSS Symposium

Porting Strategy

(Olaf Schenk - Lugano)

- Force calculation is the heaviest kernel
- Stencil kernels are memory bound, data locality is crucial
- CUDA accelerated

Early Performance Results on XK6:

Peter Messmer (NVIDIA)

- XK6 outperforms XE6 by 2.5x

Science Target (20PF Titan)

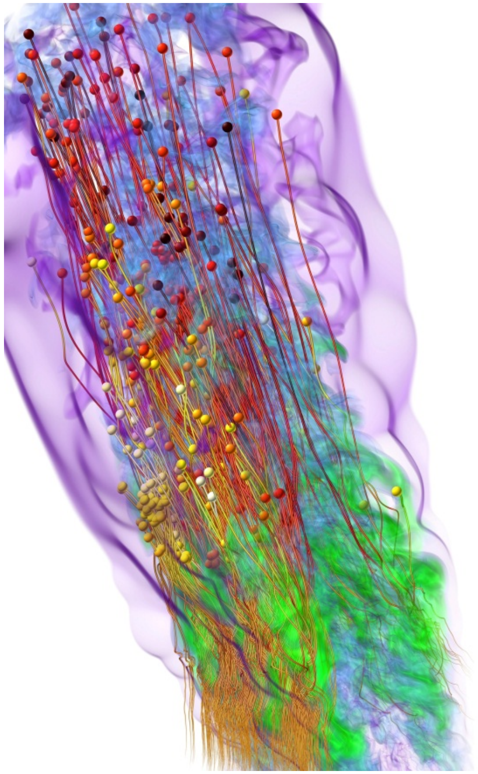
Jeroen Tromp (Princeton)

- Adjoint tomography of the entire planet -- requires analysis of 5,000 earthquakes worldwide, using an estimated 739 million core-hours (on Cray XT5)

Washington Symposium Highlighted Science Enabled by Hybrid Supercomputing

ACSS 2012 gathers novel platform's best, brightest

- ORNL, NCSA, and CSCS gathered experts in science, engineering, and computing from around the world to discuss research advances that are now possible with extreme-scale hybrid supercomputers.
- 100 attendees of the Accelerating Computational Science Symposium 2012 (ACSS 2012), held March 28–30, 2012, in Washington, D.C., explored how hybrid supercomputers speed discoveries.
- Delivering dramatic gains in computational performance and power efficiency compared with CPU-only systems, they enable researchers to accelerate a range of applications.
- The hybrid architecture is the foundation of ORNL's "Titan" supercomputer, which will reach 20 petaflops of performance by the end of this year.



S3D combustion code: lifted ethylene-air jet flame computed from DNS and tracer particle trajectories. C.S. Yoo and J. Chen performed the DNS. H. Yu (Sandia); R. Grout of the NRELperformed volume rendering.

http://www.olcf.ornl.gov/events

How Effective are GPUs on Scalable Applications?

OLCF-3 Early Science Codes -- Performance Measurements on TitanDev

Application	XK6 vs. XE6 Performance Ratio Titan Dev : Monte Rosa
S3D Turbulent combustion	**1.4**
Denovo 3D neutron transport for nuclear reactors	**3.3**
LAMMPS Molecular dynamics	**3.2**
WL-LSMS Statistical mechanics of magnetic materials	**1.6**
CAM-SE Community atmosphere model	**1.5**

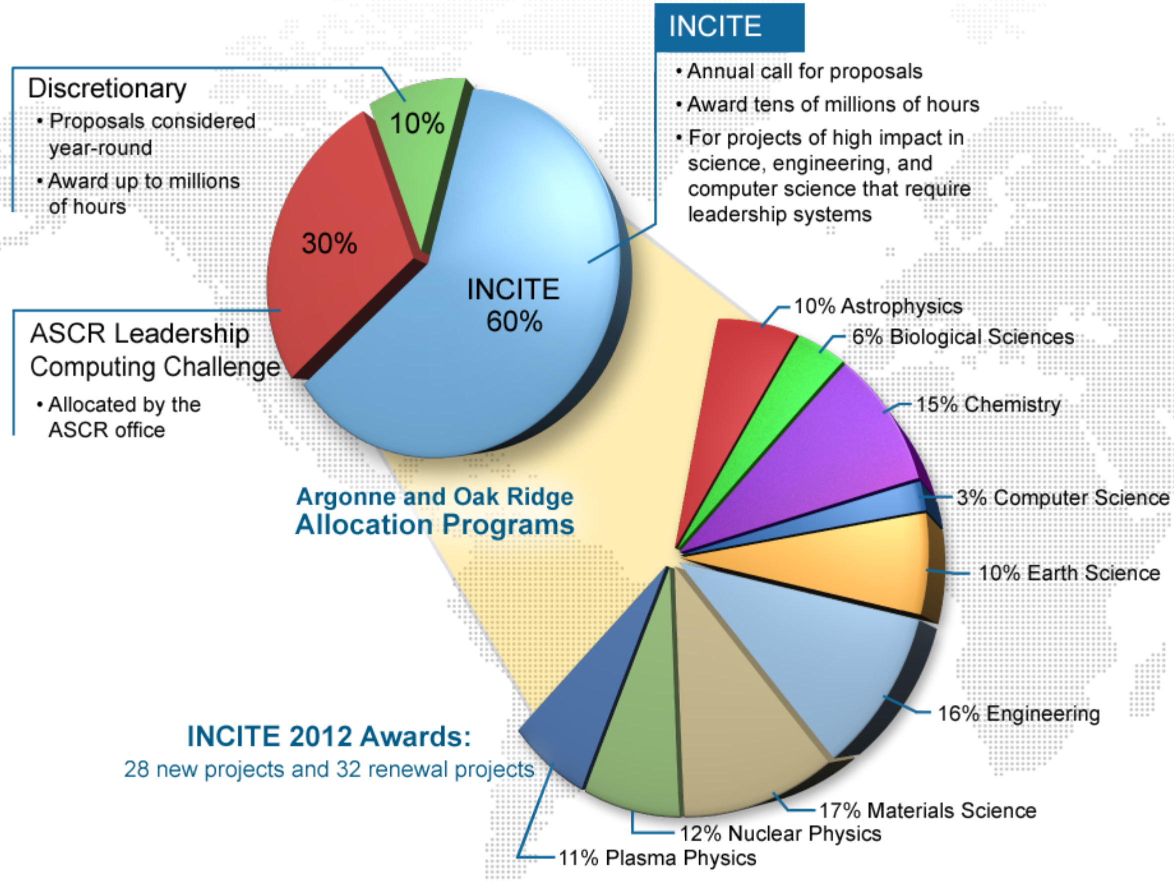
Cray XK6: Fermi GPU plus Interlagos CPU
XE6: AMD Dual Interlagos

Additional Applications from Community Efforts
Current performance measurements on TitanDev

Application	XK6 vs. XE6 Performance Ratio Titan Dev : Monte Rosa
NAMD High-performance molecular dynamics	**1.4**
Chroma High-energy nuclear physics	**6.1**
QMCPACK Electronic structure of materials	**3.0**
SPECFEM-3D Seismology	**2.5**
GTC Plasma physics for fusion-energy	**1.6**
CP2K Chemical physics	**1.5**

Innovative and Novel Computational Impact on Theory and Experiment

INCITE provides awards of time on the Oak Ridge and Argonne Leadership Computing Facility (OLCF and ALCF) systems for researchers to pursue transformational advances in science and technology: **1.7 billion core hours** were awarded in 2012.



Call for Proposals

The INCITE program seeks proposals for high-impact science and technology research challenges that require the power of the leadership-class systems. Allocations will be for calendar year 2013.

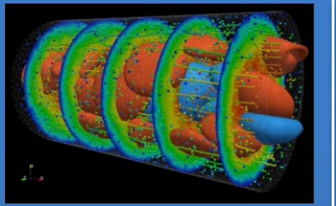
April 11 – June 27, 2012

Contact information
Julia C. White, INCITE Manager
whitejc@DOEleadershipcomputing.org

U.S. DEPARTMENT OF ENERGY INCITE LEADERSHIP COMPUTING

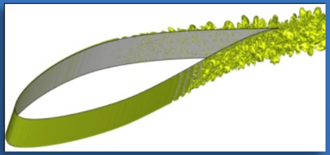
Diversity of INCITE science

Simulating a flow of healthy (red) and diseased (blue) blood cells with a Dissipative Particle Dynamics method.
- George Karniadakis, Brown University

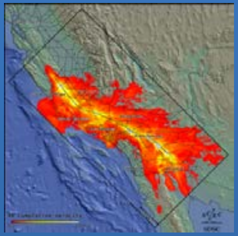


Provide new insights into the dynamics of turbulent combustion processes in internal-combustion engines.
-Jacqueline Chen and Joseph Oefelein, Sandia National Laboratories

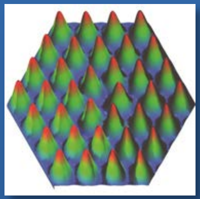
Demonstration of high-fidelity capture of airfoil boundary layer, an example of how this modeling capability can transform product development.
- Umesh Paliath, GE Global Research



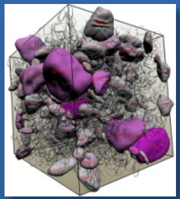
Calculating an improved probabilistic seismic hazard forecast for California.
- Thomas Jordan, University of Southern California



Modeling charge carriers in metals and semiconductors to understand the nature of these ubiquitous electronic devices.
- Richard Needs, University of Cambridge, UK



High-fidelity simulation of complex suspension flow for practical rheometry.
- William George, National Institute of Standards and Technology



Other INCITE research topics

- Glimpse into dark matter
- Supernovae ignition
- Protein structure
- Creation of biofuels
- Replicating enzyme functions
- Global climate
- Regional earthquakes
- Carbon sequestration
- Turbulent flow
- Propulsor systems
- Membrane channels
- Protein folding
- Chemical catalyst design
- Combustion
- Algorithm development
- Nano-devices
- Batteries
- Solar cells
- Reactor design
- Nuclear structure

U.S. DEPARTMENT OF ENERGY INCITE LEADERSHIP COMPUTING

Access to leadership-class resources

INCITE
Eligibility Questions

INCITE is designed for investigators across a wide range of disciplines who are prepared to explore the impact on their research of using tens to hundreds of thousands of processors.

INCITE is open to researchers worldwide. Funding should already be in place for staff, etc, from a recognizable source (state, federal, or private).

- Early access may be requested to prepare for INCITE. See the Director’s Discretionary programs.
  - ALCF: www.alcf.anl.gov
  - OLCF: www.olcf.ornl.gov
- INCITE information
  - Previous awards: www.doeleadershipcomputing.org
  - Call for Proposals: http://hpc.science.doe.gov

Acknowledgements

- OLCF's Center for Accelerated Application Readiness (CAAR)
  - B. Messer (leader), W. Joubert, M. Brown, M. Norman, M. Eisenbach, R. Shankaran,
- Cray Center of Excellence at ORNL
  - J. Levesque, J. Larkin
- NVIDIA Development Technicians
  - C. Lindahl, C. Woolley, P. Micikevicius, J. Luitjens, C. Ponder, S. Tariq, M. Clark, P. Messmer, P. Wang
- Swiss National Supercomputing Centre, CSCS, for their assistance in accessing their Cray XE6 platform, Monte Rosa.
- National Center for Supercomputing Applications, NCSA
- U.S. Department of Energy, Office of Science
- Oak Ridge Leadership Computing Facility at the ORNL