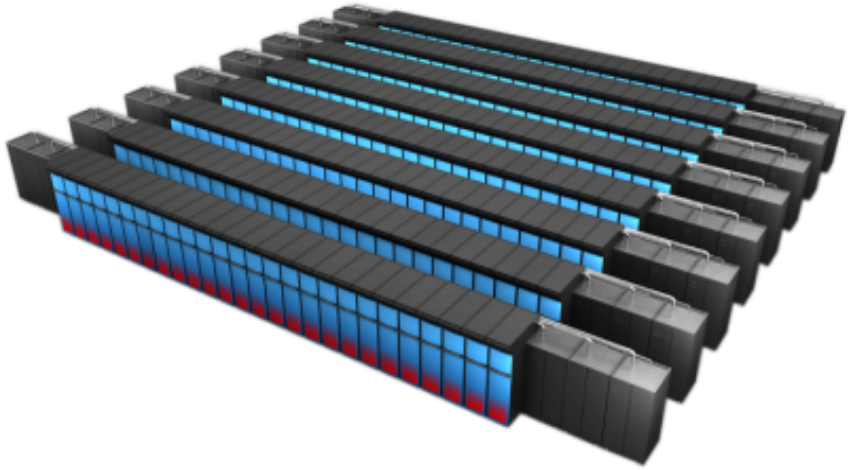
Application Readiness at ORNL



20 Years of Excellence in Computational Science

OLCF

Oak Ridge Leadership Computing Facility

1992–2012

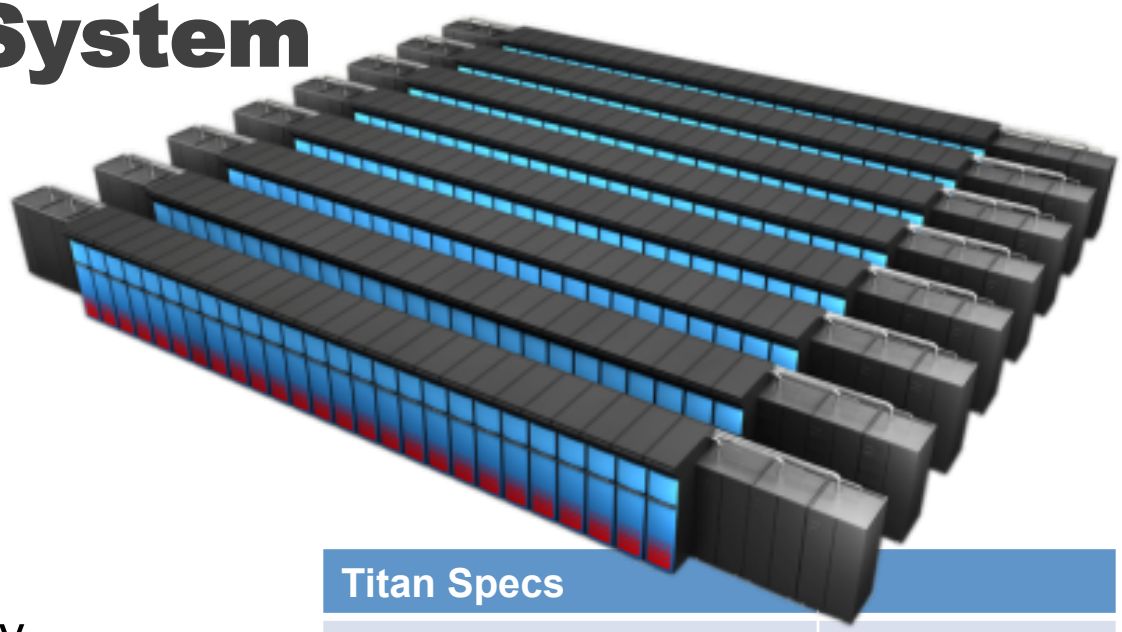
Bronson Messer

OLCF Center for Accelerated Application Readiness (CAAR)

Cray Technical Workshop

October 9, 2012

ORNL's "Titan" System

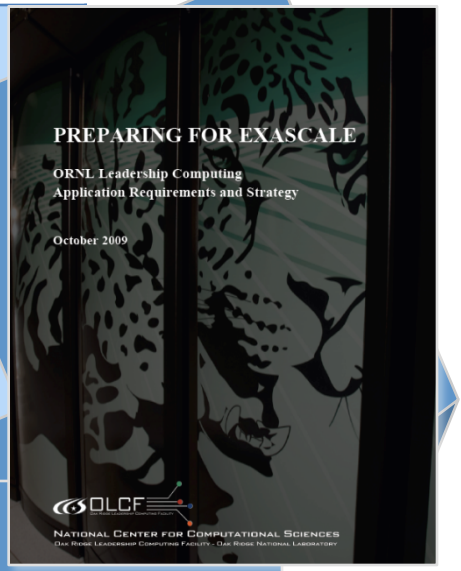


- Upgrade of Jaguar from Cray XT5 to XK6
- Cray Linux Environment operating system
- Gemini interconnect
  - 3-D Torus
  - Globally addressable memory
  - Advanced synchronization features
- AMD Opteron 6274 processors (Interlagos)
- New accelerated node design using NVIDIA multi-core accelerators
  - 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

OLCF-3 Applications Requirements developed by surveying science community

OLCF Application Requirements Document	• Elicited, analyzed, and validated using a new comprehensive requirements questionnaire • Project overview, science motivation and impact, application models, algorithms, parallelization strategy, software, development process, SQA, V&V, usage workflow, performance • Results, analysis, and conclusions documented in 2009 OLCF application requirements document
Science Driver Survey	• Developed in consultation with 50+ leading scientists in many domains • Key questions – What are the science goals and does OLCF-3 enable them? – What might the impact be if the improved science result occurs? – What does it matter if this result is delivered in the 2012 timeframe?



Science Driver Survey

•Science driver

–What science will be pursued on this system and how is it different (in fidelity/quality/ predictability and/or productivity/throughput) from the current system

•Science impact

–What might the impact be if this improved science result occurs? Who cares, and why?

•Science timeliness

–If this result is delivered in the 2010 timeframe, what does it matter as opposed to coming 5 years later (or never at all)? What other programs agencies, stakeholders, and/or facilities are dependent up on the timely delivery of this result, and why?

OLCF-3 Applications Analyzed

Science outcomes were elicited from a broad range of applications

Application area	Application codes	Science target
Astrophysics	Chimera, GenASiS	• Core-collapse supernovae simulation; validation against observations of neutrino signatures, gravitational waves, and photon spectra
	MPA-FT, MAESTRO	• Full-star type Ia supernovae simulations of thermonuclear runaway with realistic subgrid models
Bioenergy	LAMMPS, GROMACS	• Cellulosic ethanol: dynamics of microbial enzyme action on biomass
Biology	LAMMPS	• Systems biology • Genomic structure
Chemistry	CP2K, CPMD	• Interfacial chemistry
	GAMESS	• Atmospheric aerosol chemistry • Fuels from lignocellulosic materials
Combustion	S3D	• Combustion flame front stability and propagation in power and propulsion engines
	RAPTOR	• Internal combustion design in power and propulsion engines: bridge the gap between device- and lab-scale combustion
Energy Storage	MADNESS	• Electrochemical processes at the interfaces; ionic diffusion during charge-discharge cycles

OLCF-3 Applications Analyzed

Science outcomes were elicited from a broad range of applications

Application area	Application codes	Science target
Fusion	GTC	• Energetic particle turbulence and transport in ITER
	GTS	• Electron dynamics and magnetic perturbation (finite-beta) effects in a global code environment for realistic tokamak transport • Improved understanding of confinement physics in tokamak experiments • Address issues such as the formations of plasma critical gradients and transport barriers
	XGC1	• First-principles gyrokinetic particle simulation of multiscale electromagnetic turbulence in whole-volume ITER plasmas with realistic diverted geometry
	AORSA, CQL3D	• Tokamak plasma heating and control
	FSP	• MHD scaling to realistic Reynolds numbers • Global gyrokinetic studies of core turbulence encompassing local & nonlocal phenomena and electromagnetic electron dynamics • Finite ion-orbit-width effects to realistically assess RF wave resonant ion interactions • Multi-scale integrated electromagnetic turbulence in realistic ITER geometry
	GYRO, TGYRO	• Predictive simulations of transport iterated to bring the plasma into steady-state power balance; radial transport balances power input
Geoscience	PFLOTRAN	• Stability and viability of large-scale CO_2 sequestration • Predictive contaminant ground water transport

OLCF-3 Applications Analyzed

Science outcomes were elicited from a broad range of applications

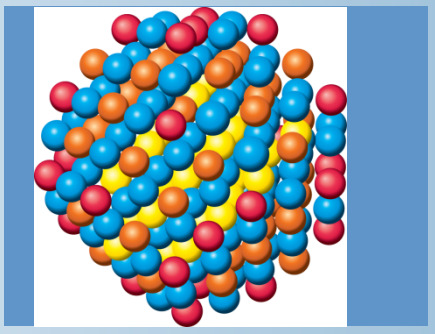
Application area	Application codes	Science target
Nanoscience	OMEN	• Electron-lattice interactions and energy loss in full nanoscale transistors
	LS3DF	• Full device simulation of a nanostructure solar cell
	DCA++	• Magnetic/superconducting phase diagrams including effects of disorder • Effect of impurity configurations on pairing and the high-T superconducting gap • High-T superconducting transition temperature materials dependence in cuprates
	WL-LSMS	• To what extent do thermodynamics and kinetics of magnetic transition and chemical reactions differ between nano and bulk? • What is the role of material disorder, statistics, and fluctuations in nanoscale materials and systems?
Nuclear energy	Denovo	• Predicting, with UQ, the behavior of existing and novel nuclear fuels and reactors in transient and nominal operation
	UNIQ	• Reduce uncertainties and biases in reactor design calculations by replacing existing multi-level homogenization techniques with more direct solution methods
Nuclear Physics	NUCCOR MFDn	• Limits of nuclear stability, static and transport properties of nucleonic matter • Predict half-lives, mass and kinetic energy distribution of fission fragments and fission cross sections
QCD	MILC,Chroma	• Achieving high precision in determining the fundamental parameters of the Standard Model (masses and mixing strengths of quarks)
Turbulence	DNS	• Stratified and unstratified turbulent mixing at simultaneous high Reynolds and Schmidt numbers
	Hybrid	• Nonlinear turbulence phenomena in multi-physics settings

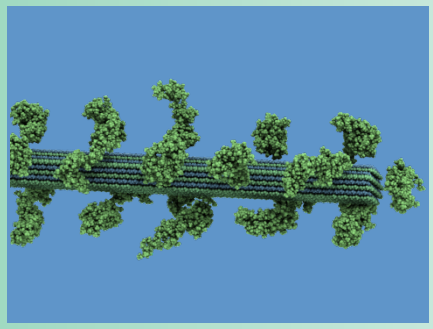
Evaluation Criteria for Selection of Representative Applications

Task	Description
Science	• Science results, impact, timeliness • Alignment with DOE and U.S. science mission (CD-0) • Broad coverage of science domains
Implementation (models, algorithms, software)	• Broad coverage of relevant programming models, environment, languages, implementations • Broad coverage of relevant algorithms and data structures (motifs) • Broad coverage of scientific library requirements
User community (current and anticipated)	• Broad institutional and developer/user involvement • Good representation of current and anticipated INCITE workload
Preparation for steady state (“INCITE ready”) operations	• Mix of low (“straightforward”) and high (“hard”) risk porting and readiness requirements • Availability of OLCF liaison with adequate skills/experience match to application • Availability of key code development personnel to engage in and guide readiness activities

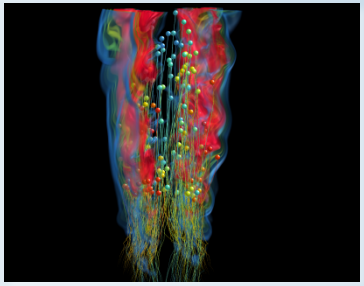
Center for Accelerated Application Readiness (CAAR)

WL-LSMS
Illuminating the role of material disorder, statistics, and fluctuations in nanoscale materials and systems.



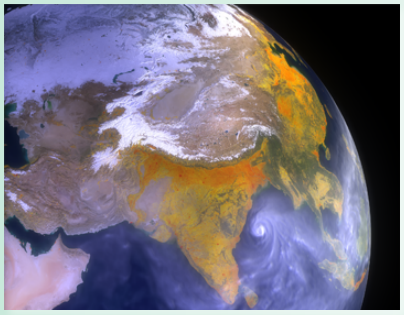


LAMMPS
A molecular description of membrane fusion, one of the most common ways for molecules to enter or exit living cells.

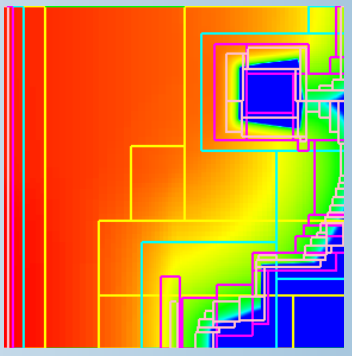


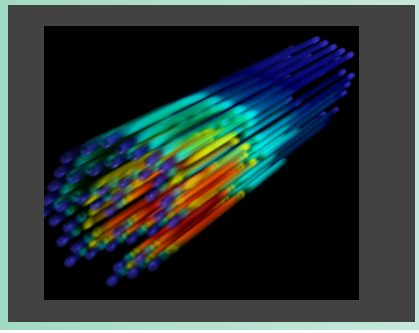
S3D
Understanding turbulent combustion through direct numerical simulation with complex chemistry.
.

CAM-SE
Answering questions about specific climate change adaptation and mitigation scenarios; realistically represent features like precipitation patterns / statistics and tropical storms.



NRDF
Radiation transport – important in astrophysics, laser fusion, combustion, atmospheric dynamics, and medical imaging – computed on AMR grids.





Denovo
Discrete ordinates radiation transport calculations that can be used in a variety of nuclear energy and technology applications.

CAAR Plan

- **Comprehensive team assigned to each app**
  - OLCF application lead
  - Cray engineer
  - NVIDIA developer
  - Other: other application developers, local tool/library developers, computational scientists
- **Single early-science problem targeted for each app**
  - Success on this problem is primary metric for success
- **Particular plan-of-attack different for each app**
  - WL-LSMS – dependent on accelerated ZGEMM
  - CAM-SE– pervasive and widespread custom acceleration required
- **Multiple acceleration methods explored**
  - WL-LSMS – CULA, MAGMA, custom ZGEMM
  - CAM-SE– CUDA, directives
  - Two-fold aim
    - **Maximum acceleration for model problem**
    - **Determination of optimal, reproducible acceleration path for other applications**

Complications

- **All of the chosen apps are under constant development**
  - Groups have, in many cases, already begun to explore GPU acceleration "on their own."
- **Production-level tools, compilers, libraries, etc. are just beginning to become available**
  - Multiple paths are available, with multifarious trade-offs
    - ease-of-use
    - (potential) portability
    - performance

CAAR Algorithmic Coverage

Code	FFT	Dense linear algebra	Sparse linear algebra	Particles	Monte Carlo	Structured grids	Unstructured grids
S3D		X	X	X		X	
CAM	X	X	X	X		X	
LSMS		X					
LAMMPS	X			X			
Denovo		X	X	X	X	X	
NRDF			X				X (AMR)

CAAR Implementation Coverage

App	Science Area	Algorithm(s)	Grid type	Programming Language(s)	Compiler(s) supported	Communication Libraries	Math Libraries
CAM-SE	climate	spectral finite elements, dense & sparse linear algebra, particles	structured	F90	PGI, Lahey, IBM	MPI	Trilinos
LAMMPS	Biology / materials	molecular dynamics, FFT, particles	N/A	C++	GNU, PGI, IBM, Intel	MPI	FFTW
S3D	combustion	Navier-Stokes, finite diff, dense & sparse linear algebra, particles	structured	F77, F90	PGI	MPI	None
Denovo	nuclear energy	wavefront sweep, GMRES	structured	C++, Fortran, Python	GNU, PGI, Cray, Intel	MPI	Trilinos, LAPACK, SuperLU, Metis
WL-LSMS	nanoscience	density functional theory, Monte Carlo	N/A	F77, F90, C, C ++	PGI, GNU	MPI	LAPACK (`ZGEMM`, `ZGTRF`, `ZGTRS`)
NRDF	radiation transport	Non-equilibrium radiation diffusion equation	structured AMR	C++, C, F77	PGI, GNU, Intel	MPI, SAMRAI	BLAS, PETSc, Hypre, SAMRSolvers

Science Target Requirements

Application	First Titan Science Problem	Titan Core-Hours Needed
S3D	3-dimensional DNS of HCCI combustion in a high-pressure stratified turbulent ethanol/air mixture using detailed chemical kinetics (28 chemical species)	• 120M
Denovo	Neutron transport in reactor core, using 128 million spatial grid cells, 16 energy groups, 256 discrete ordinates; steady-state solution in less than 12 wall clock hours	• 14M
LAMMPS	Biological membrane fusion in coarse-grained MD within 5 wall clock days; 850,000,000 coarse-grained particles, roughly 1000× in particle number and spatial extent than current standard.	• 40M
WL-LSMS	Both the magnetization and the free energy for a single 250-atom iron supercell. magnetic materials. (5 million CPU-hours are required on Jaguar today to calculate only the free energy)	• 10M
CAM-SE	CCSM simulation w/ Mozart tropospheric chemistry with 101 constituents at high spatial resolution (1/8 degree); 100-year climate simulation @ 3.8 simulated years per wall clock day.	• 150M

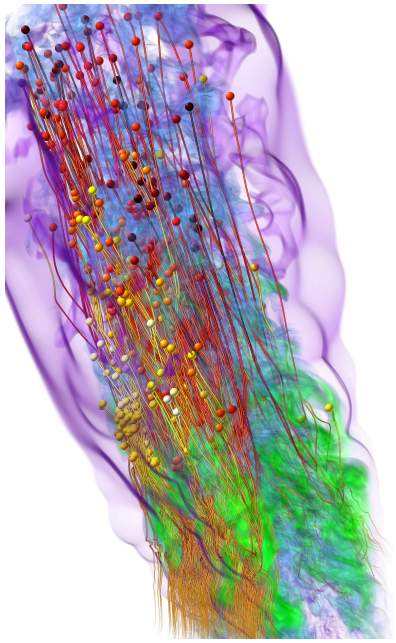
S3D

Direct Numerical Simulation of Turbulent Combustion

OpenACC

Code Description

- Compressible Navier-Stokes equations
- 3D Cartesian grid, 8th-order finite difference
- Explicit 4th-order Runge-Kutta integration
- Fortran, 3D Domain decomposition, non-blocking MPI



DNS provides unique fundamental insight into the chemistry-turbulence interaction

Porting Strategy

- Hybrid MPI/OpenMP/OpenACC application
- All intensive calculations can be on the accelerator
- Redesign message passing to overlap communication and computation

Early Performance Results on XK6

- **Refactored code was 2x faster on Cray XT5**
- OpenACC acceleration with minimal overhead
- XK6 outperforms XE6 by 1.4x

Science Target (20 PF Titan)

Increased chemical complexity for combustion:

- Jaguar: 9-22 species (H_2, syngas, ethylene)
- Titan: 30-100 species (ethanol, n-heptane, iso-octane)

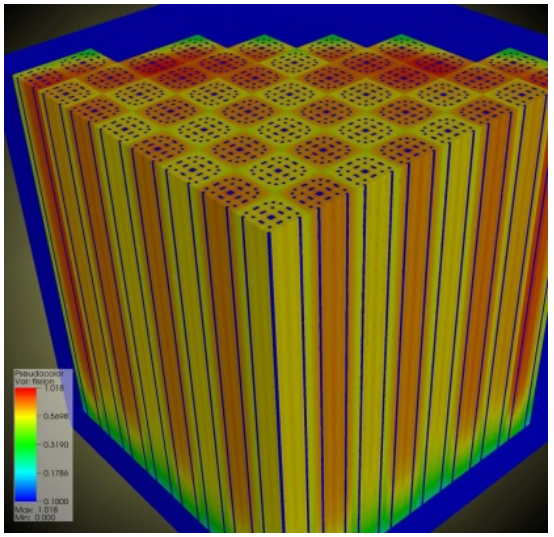
DENOVO

3D Neutron Transport for Nuclear Reactor Design

CUDA

Code Description

- Linear Boltzmann radiation transport
- Discrete ordinates method
- Iterative eigenvalue solution
- Multigrid, preconditioned linear solves
- C++ with F95 kernels



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

Porting Strategy

- SWEEP kernel re-written in C++ & CUDA, runs on CPU or GPU
- Scaling to over 200K cores with opportunities for increased parallelism on GPUs
- Reintegrate SWEEP into DENOVO

Early Performance Results on XK6

- **Refactored code was 2x faster on Cray XT5**
- XK6 performance exceeds XE6 by 3.3x

Science Target (20PF Titan)

- 3-D, Full reactor radiation transport for CASL challenge problems within 12 wall clock hours. This is the CASL stretch goal problem!

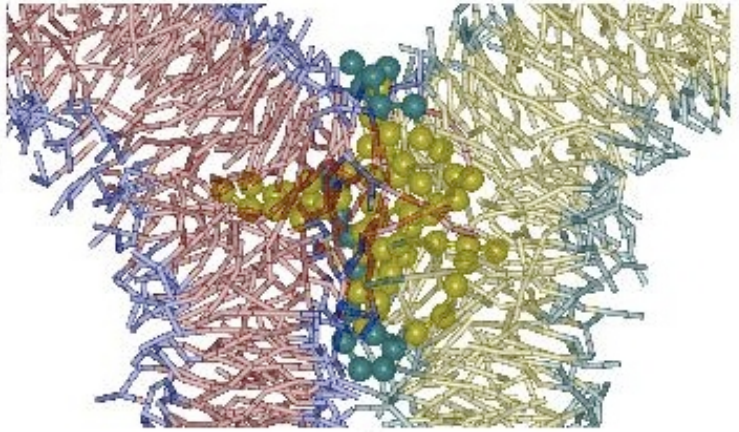
LAMMPS

CUDA

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
- XK6 outperforms XK6 w/o GPU by 6.5x

Science Target (20PF Titan)

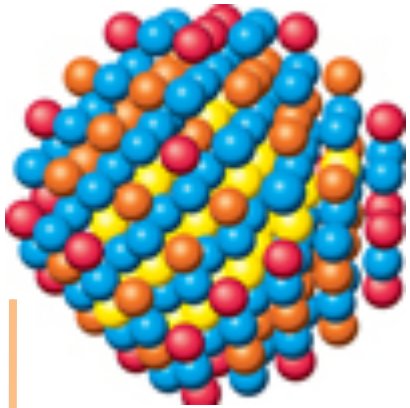
- Simulate biological membrane fusion in coarse-grained MD within 5 wall clock days

Wang-Landau LSMS

First principles, statistical mechanics of magnetic materials

Code Description

- Combines classical statistical mechanics (W-L)for atomic magnetic moment distributions with first-principles calculations (LSMS) of the associated energies.
- Main computational effort is dense linear algebra for complex numbers
- F77 with some F90 and C++ for the statistical mechanics driver.



Compute the magnetic structure and thermodynamics of low-dimensional magnetic structures

Porting Strategy

- Leverage accelerated linear algebra libraries, e.g., cuBLAS + CULA, LibSci_acc
- Parallelization over (1) W-L Monte-Carlo walkers, (2) over atoms through MPI process, (3) OpenMP on CPU sections.
- Restructure communications: moved outside energy loop

Early Performance Results on XK6:

- XK6 outperforms XE6 by 1.6x
- XK6 outperforms XK6 w/o GPU by 3.1x

Science Target (20PF Titan)

- Calculate both the magnetization **and** the free-energy for magnetic materials.

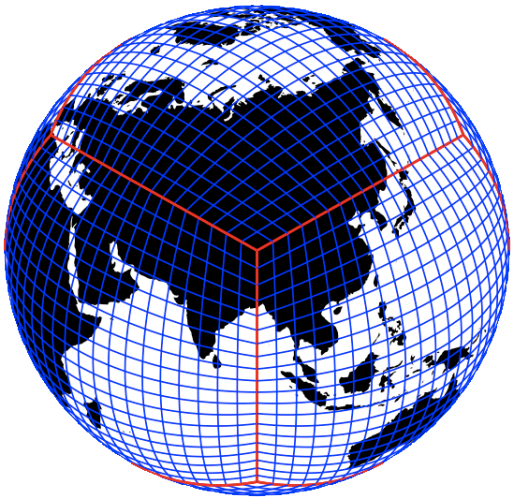
CAM-SE

CUDA Fortran

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
- XK6 outperforms XK6 w/o GPU by 2.6x

Science Target (20PF Titan)

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

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 AMD 16-core Opteron CPU

XE6: 2X AMD 16-core Opteron CPUs (Monte Rosa - 1496 node XE6 @ CSCS)

We expect the results to be even better with the GK110 Kepler processor

- Hyper-Q
  - Fermi allows only a single work queue
  - Kepler allows 32 simultaneous work queues, allowing multiple CPU threads to connect to the GPU, removing false serialization.
  - Ameliorates “not enough work” problem
- GPUDirect
  - Enables GPUs within a single computer, or GPUs in different servers located across a network, to directly exchange data without needing to go to CPU/system memory.
  - Reduces CPU-GPU memory copy penalty for many applications

Both features are expected to significantly increase current application performance and to aid further application ports.

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**

The Accelerating Computational Science Symposium 2012 (ACSS 2012), held March 28–30 in Washington, D.C. Sponsored by: OLCF, NCSA, and CSCS with 100 attendees.

Symposium summary at: http://www.olcf.ornl.gov/media-center/center-reports/

Training: Amplifying the Impact of Early Science

- **Strategy:** Provide conferences, workshops, tutorials, case studies, and lessons learned on tools and techniques for realizing hybrid architecture benefits. *Content will be formulated and field-tested via early-science experiences.*
- **Objective:** Users will be able to expose hierarchical parallelism, use compiler directive-based tools, analyze / optimize / debug codes, and use low-level programming techniques if required

Planned CY 2013 Training Events

West Coast Titan Training Workshop (Jan/Feb 2013)
East Coast Titan Training Workshop (Feb/March 2013)
Object Oriented Programming Workshop (Feb 2013)
2013 OLCF User Meeting (April 2013)
Tutorial Series (Spring/Summer)
- Compilers
- Performance Analysis Tools
- Debugging and Profiling

Fall Titan Training (Fall 2013)

Summary

- CAAR and work by our colleagues at NVIDIA, CSCS, NCSA, and other organizations has produced a set of vanguard GPU-capable applications.
- More than a dozen codes are poised to make immediate use of Titan.
- Strong overlap with INCITE and ALCC awardees, combined with the flexibility afforded by our DD program, will allow us to facilitate first-in-class simulations immediately following Titan acceptance.