GPU-Accelerated Science on Titan

Jack Wells
Director of Science
Oak Ridge Leadership Computing Facility

May 15, 2012
GTC 2012
Oak Ridge Leadership Computing Facility Mission

The OLCF is a DOE Office of Science National User Facility whose mission is to enable breakthrough science by:

• Fielding the most powerful capability computers for scientific research,

• Building the required infrastructure to facilitate user access to these computers,

• Selecting a few time-sensitive problems of national importance that can take advantage of these systems,

• And partnering with these teams to deliver breakthrough science.
FOR IMMEDIATE RELEASE:

SANTA CLARA, Calif. — Sep. 30, 2009 — Oak Ridge National Laboratory (ORNL) announced plans today for a new supercomputer that will use NVIDIA®’s next generation CUDA™ GPU architecture, codenamed “Fermi”. Used to pursue research in areas such as energy and climate change, ORNL’s supercomputer is expected to be 10-times more powerful than today’s fastest supercomputer.

Jeff Nichols, ORNL associate lab director for Computing and Computational Sciences, joined NVIDIA co-founder and CEO Jen-Hsun Huang on stage during his keynote at NVIDIA’s GPU Technology Conference. He told the audience of 1,400 researchers and developers that “Fermi” would enable substantial scientific breakthroughs that would be impossible without the new technology.

“This would be the first co-processing architecture that Oak Ridge has deployed for open science, and we are extremely excited about the opportunities it creates to solve huge scientific challenges,” Nichols said.
Presentations of GPU-accelerated science

Jacqueline H. Chen (Sandia National Laboratories)
"Direct Numerical Simulation of Turbulence-Chemistry Interactions: Fundamental Insights Towards Predictive Models"

Ray Grout (National Renewable Energy Laboratory)
"S3D Direct Numerical Simulation - Preparations for the 10-100PF Era"

William Tang (Princeton Plasma Physics Laboratory (Princeton))
"Fusion Energy Sciences & Computing at the Extreme Scale"

John A. Turner (Oak Ridge National Laboratory)
"Transforming Modeling and Simulation for Nuclear Energy Applications"

Loukas Petridis (Oak Ridge National Laboratory)
"Computer Simulation of Lignocellulosic Biomass"

Jeroen Tromp (Princeton University)
"Toward Global Seismic Imaging based on Spectral-Element and Adjoint Methods"
Outline: Titan project overview

• Goals of Titan project
• Cray XK6 nodes
• Titan system
• Upgrade process
• Programming model
• Tools
• Early Performance Results
### Titan System Goals: Deliver breakthrough science for DOE/SC, industry, and the nation

<table>
<thead>
<tr>
<th>Earth System Science</th>
<th>Energy</th>
<th>Fundamental Science</th>
</tr>
</thead>
<tbody>
<tr>
<td>Understanding our earth and the processes that impact it</td>
<td>Reducing U.S. reliance on foreign energy &amp; reducing carbon footprint of production</td>
<td>Understanding the physical processes from the scale of subatomic particles to the universe</td>
</tr>
<tr>
<td>- Sea level rise</td>
<td>- Carbon free energy production from fusion, fission, solar, wind, and geothermal sources</td>
<td>- Extend our knowledge of the natural world</td>
</tr>
<tr>
<td>- Regional climate change</td>
<td>- Improving the efficiency of combustion energy sources</td>
<td>- Deliver new technologies to advance energy mission</td>
</tr>
<tr>
<td>- Geologic carbon sequestration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Sustainability of biofuels</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Earthquakes and Tsunamis</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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** – Molecular 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 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 “Kepler” GPUs
- 20+ PFlops peak system performance
- 600 TB DDR3 mem. + 88 TB GDDR5 mem

<table>
<thead>
<tr>
<th>Titan Specs</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Compute Nodes</td>
<td>18,688</td>
</tr>
<tr>
<td>Login &amp; I/O Nodes</td>
<td>512</td>
</tr>
<tr>
<td>Memory per node</td>
<td>32 GB + 6 GB</td>
</tr>
<tr>
<td># of Fermi chips (2012)</td>
<td>960</td>
</tr>
<tr>
<td># of NVIDIA “Kepler” (2013)</td>
<td>14,592</td>
</tr>
<tr>
<td>Total System Memory</td>
<td>688 TB</td>
</tr>
<tr>
<td>Total System Peak Performance</td>
<td>20+ Petaflops</td>
</tr>
<tr>
<td>Cross Section Bandwidths</td>
<td>X=14.4 TB/s, Y=11.3 TB/s, Z=24.0 TB/s</td>
</tr>
</tbody>
</table>
Two Phase Upgrade Process

• Phase 1: XT5 to XK6 without GPUs
  – Replace all XT5 nodes with XK6 nodes
  – 16-core processors, 32 GB/node, Gemini
  – 960 nodes (10 cabinets) have NVIDIA Fermi GPUs
  – Users ran on half of system while other half was upgraded

• Phase 2: Add NVIDIA Kepler GPUs
  – Cabinet Mechanical and Electrical upgrades
    • New air plenum bolts on to cabinet
    • Larger fan
    • Additional power supply
  – Rolling upgrade of node boards
    • Pull board, add 4 Kepler GPUs modules, replace board, test, repeat 3,647 times!
    • Keep most of the system available for users during upgrade
  – Acceptance test of system

Image courtesy of Cray Inc.
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: Open standard for compiler directives that allows the expression of 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
  – CrayPAT and Cray Apprentice support XK6 programming
How do you program these nodes?

• 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

<table>
<thead>
<tr>
<th>Compilers</th>
<th>Performance Tools</th>
<th>GPU Libraries</th>
<th>Debuggers</th>
<th>Source Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cray</td>
<td>CrayPAT</td>
<td>MAGMA</td>
<td>DDT</td>
<td>HMPP Wizard</td>
</tr>
<tr>
<td>PGI</td>
<td>Apprentice</td>
<td>CULA</td>
<td>NVIDIA Gdb</td>
<td></td>
</tr>
<tr>
<td>CAP-HMPP</td>
<td>Vampir</td>
<td>Trillinos</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pathscale</td>
<td>VampirTrace</td>
<td>libSCI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NVIDIA CUDA</td>
<td>TAU</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GNU</td>
<td>HPCToolkit</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intel</td>
<td>CUDA Profiler</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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.

DENOV0: Neutron Transport in Reactor Core
DENOV0 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.
Additional Applications from Community Efforts

**GTC: Magnetic confinement of fusion plasmas**
Increasingly faster supercomputers enable new physics insights with GTC, the first fusion code delivering production simulations at petaflops in 2009.

**GROMACS: Biophysics and Biofuels**
Understand the surface structure and function of lignin clumps down to 1 angstrom length scale.

**SPECFEM-3D: Seismological Imaging**
Seismic inverse modeling and imaging to reveal seismic hotspots (red indicates fast seismic shear wave speeds whereas blue indicates slow wave speeds).
How Effective are GPUs on Scalable Applications?
OLCF-3 Early Science Codes -- Performance Measurements on TitanDev

<table>
<thead>
<tr>
<th>Application</th>
<th>XK6 vs. XE6 Performance Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Titan Dev: Monte Rosa</strong></td>
<td></td>
</tr>
<tr>
<td>S3D Turbulent combustion</td>
<td>1.4</td>
</tr>
<tr>
<td>Denovo 3D neutron transport</td>
<td>3.3</td>
</tr>
<tr>
<td>LAMMPS Molecular dynamics</td>
<td>3.2</td>
</tr>
<tr>
<td>WL-LSMS Statistical mechanics of magnetic materials</td>
<td>1.6</td>
</tr>
<tr>
<td>CAM-SE Community atmosphere model</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Cray XK6: Fermi GPU plus Interlagos CPU
XE6: AMD Dual Interlagos
## Additional Applications from Community Efforts

Current performance measurements on TitanDev

<table>
<thead>
<tr>
<th>Application</th>
<th>XK6 vs. XE6 Performance Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAMD High-performance molecular dynamics</td>
<td>Titan Dev : Monte Rosa 1.4</td>
</tr>
<tr>
<td>Chroma High-energy nuclear physics</td>
<td>6.1</td>
</tr>
<tr>
<td>QMCPACK Electronic structure of materials</td>
<td>3.0</td>
</tr>
<tr>
<td>SPECFEM-3D Seismology</td>
<td>2.5</td>
</tr>
<tr>
<td>GTC Plasma physics for fusion-energy</td>
<td>1.6</td>
</tr>
<tr>
<td>CP2K Chemical physics</td>
<td>1.5</td>
</tr>
</tbody>
</table>
Access to Titan via INCITE

INCITE seeks computationally intensive, large-scale research projects with the potential to significantly advance key areas in science and engineering.

1 **Impact criterion**
   - High-impact science and engineering

2 **Computational leadership criterion**
   - Computationally intensive runs that cannot be done anywhere else

3 **Eligibility criterion**
   - INCITE grants allocations regardless of funding source (ex. DOE, NSF, private, etc)
   - Researchers at non-US institutions may apply

Call for 2013 proposals open now through June 27, 2012

The INCITE program seeks proposals for high-impact science and technology research challenges that require the power of the leadership-class systems.

Email: INCITE@DOEleadershipcomputing.org
Web: http://hpc.science.doe.gov

In 2013 over 4 billion core-hours will be available through the INCITE program.
The research and activities described in this presentation were performed using the resources of the National Center for Computational Sciences at Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC0500OR22725.