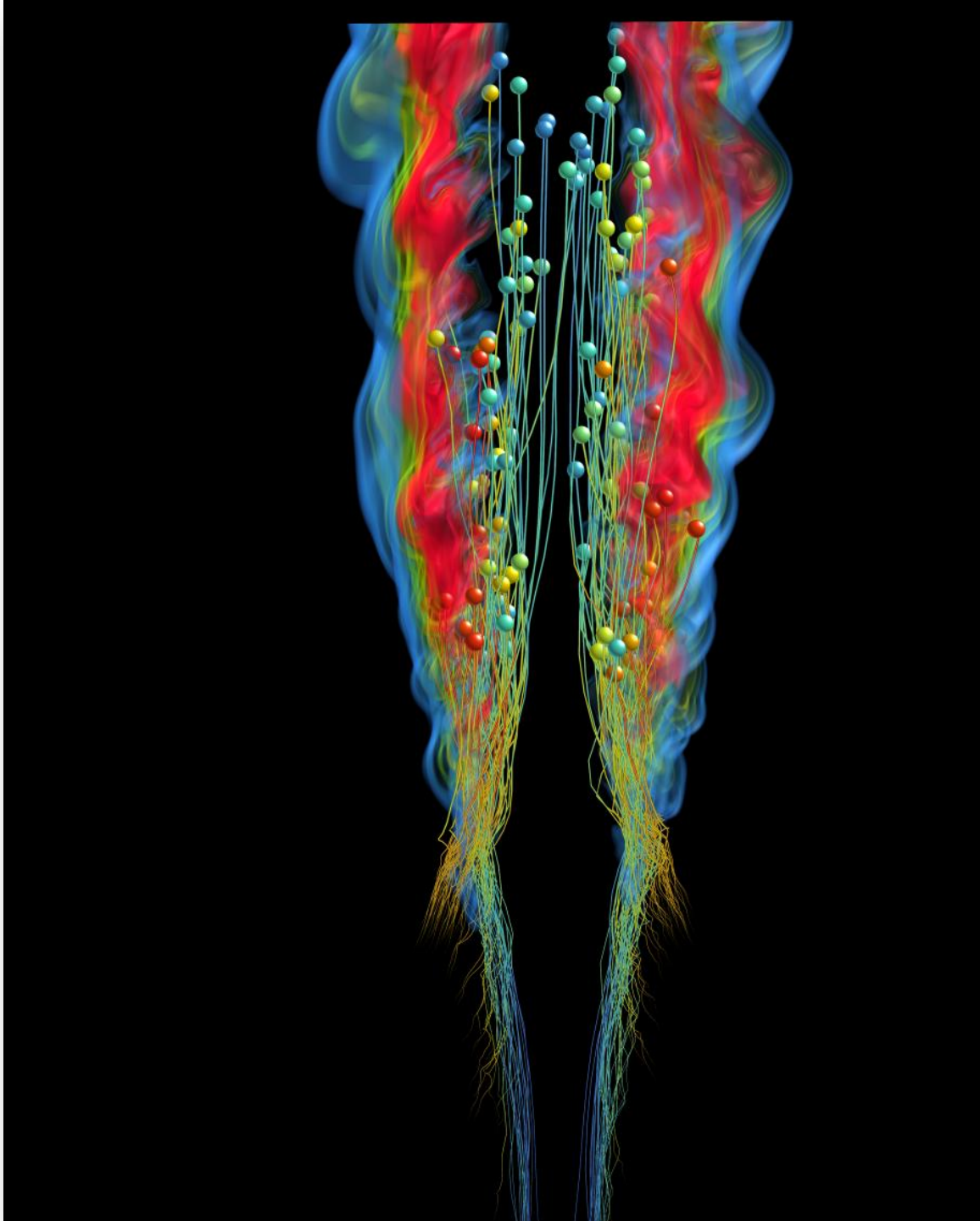
Toward Exascale Computational Science with Graphics Processors

Jeffrey Vetter

With contributions from many…

Presented to

SAAHPC, Knoxville



Oak Ridge National Laboratory

Georgia Tech College of Computing
Computational Science and Engineering

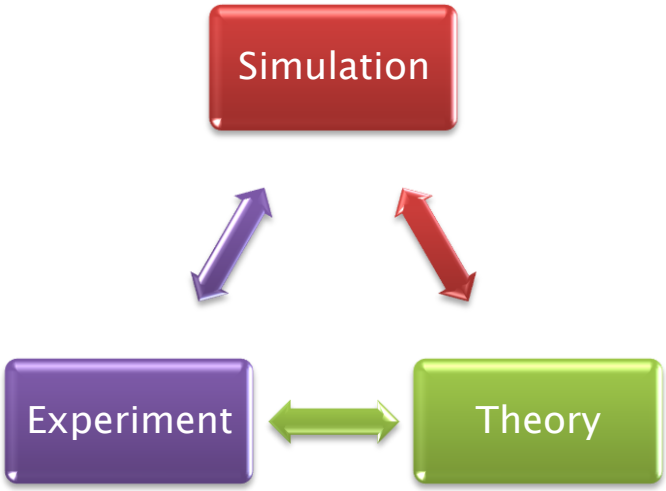
Overview

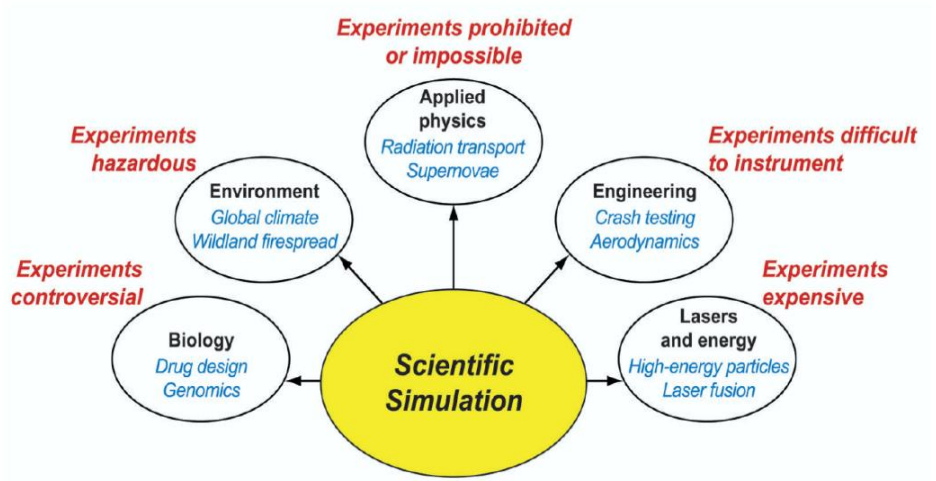
- Predictive scientific simulation is important for scientific discovery
- HPC systems have been highly successful
- The HPC community has several (new) constraints
- Heterogeneous computing with GPUs offers some opportunities and challenges
- Newly awarded NSF partnership will provide heterogeneous supercomputing for open science
- Exascale computing will require fundamental changes to our systems

SCIENTIFIC SIMULATION

Scientific Simulation is an accepted form of science

- **Theory fails for**
  - Complex phenomena
- **Experiment fails for**
  - Ethics
    - Testing on humans, animals
  - Experimentation limits
    - Climate modeling
    - Atomistic scale materials

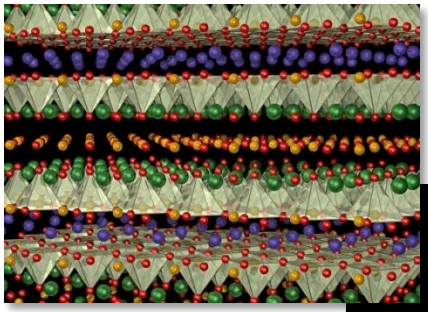


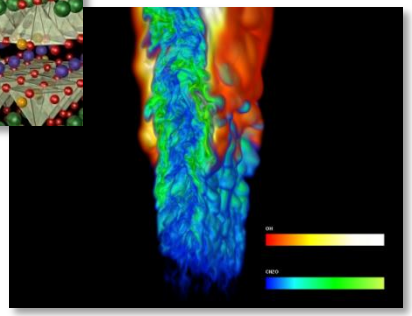


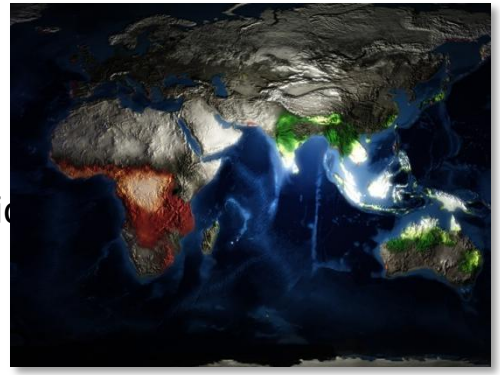
US Department of Energy Office of Science, "A Science-Based Case for Large-Scale Simulation," US Department of Energy Office of Science 2003.

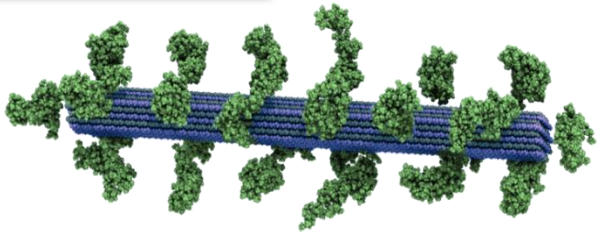
Simulation is Tackling Compelling Problems

- **Energy for environmental sustainability**
  - Climate change: Carbon sequestration, weather event impacts on global climate, decadal climate predictive skill in aerosol forcing, global climate at unprecedented resolution, mantle convection
  - Energy storage: Charge storage and transfer in nano-structured supercapacitors
  - Combustion: Stabilizing diesel jet flames for increased efficiency and decreased emissi
  - Bioenergy: Recalcitrance in cellulosic ethanol
  - Solar: Nonequilibrium semiconductor alloys
  - Energy transmission: Role of inhomogeneities in high-temperature superconducting cuprates
  - Fusion: ITER design, optimization, and operation
  - Nuclear energy: Fully resolved reactor core neutron state



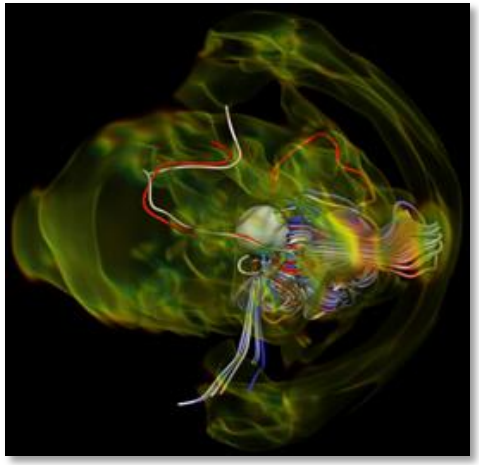


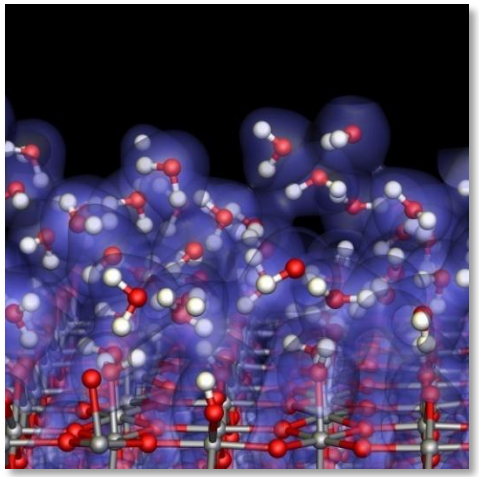




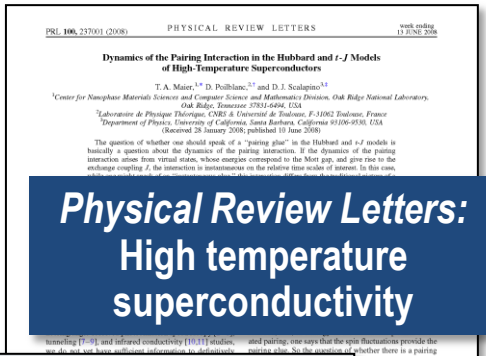
Tackling Compelling Problems (2)

- **Materials and nanoscience**
  - Structure of nanowires, nanoelectronics, nanorods, and strongly correlated materials (magnets)
- **Fundamental science**
  - Astrophysics: Decipher core-collapse supernovae and black hole mergers
  - Chemistry: Water structure in biological and aqueous-phase systems; water bonding on graphene and carbon nanotubes
  - Nuclear physics: Probe the anomalously long lifetime of carbon-14
  - High energy physics: Generate lattices at the physical values of the light quark (u, d, s) masses
  - Turbulence: Dispersion relative to air quality modeling and bioterrorism; physics of shock/turbulence interaction
  - Biology: Effect of an embolization on hemodynamics in the entire neurovasculature

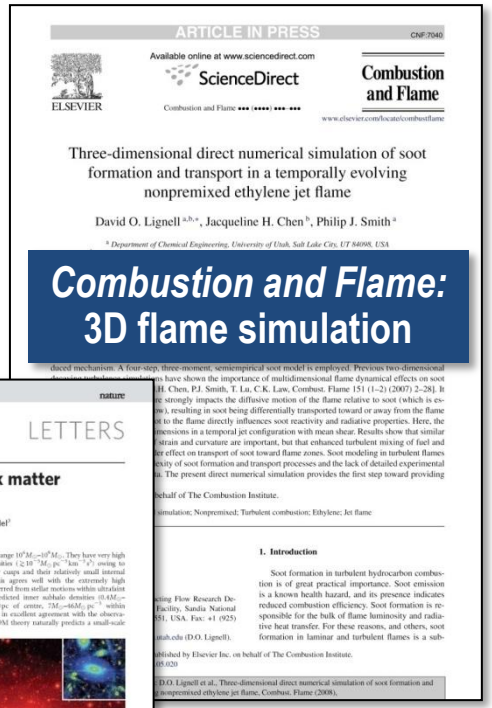




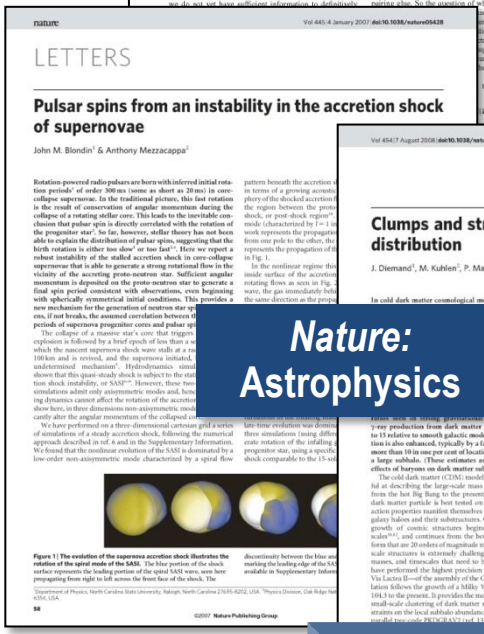
Resulting in Highly Visible Science

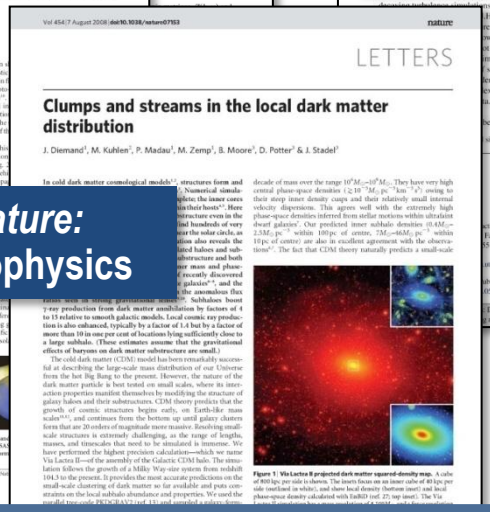


Physical Review Letters:
High temperature superconductivity

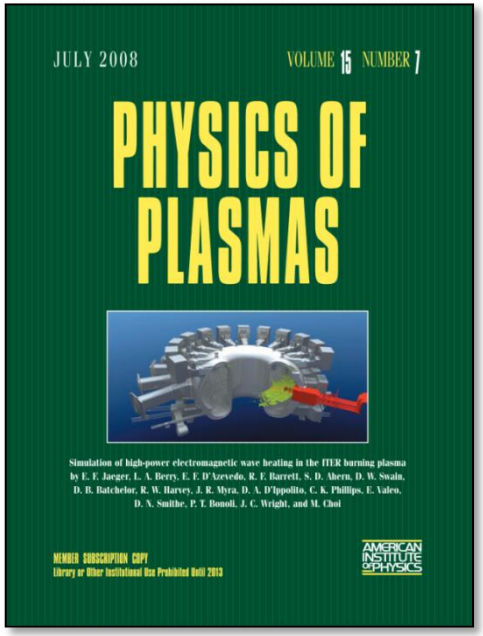


Combustion and Flame:
3D flame simulation





Nature:
Astrophysics

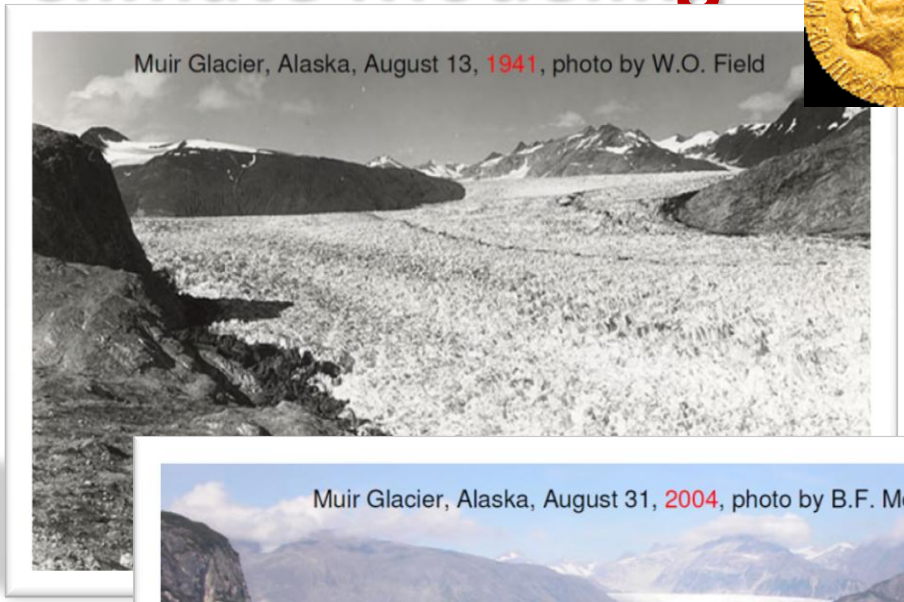


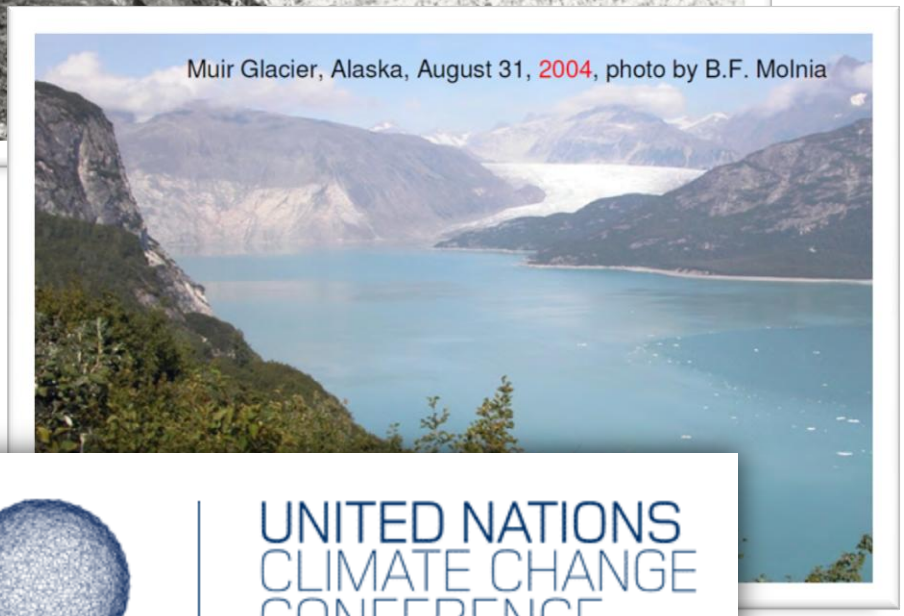
Physics of Plasmas:
ICRF heating in ITER

Transformational science enabled by advanced scientific computing

Motivating Example: Climate Modeling

- **Assess scientific basis, impact, adaptation, vulnerability, mitigation**
  - Observation and simulation play critical roles
- ***Profound impact on worldwide socioeconomic policy, energy, etc***
  - UN Conference - Sep 09
  - Copenhagen Climate conference – Dec 09
- **Intergovernmental Panel On Climate Change**
  - Sponsored by UNEP and WMO
  - Thousands of scientists from around the world





COP15
COPENHAGEN

UNITED NATIONS
CLIMATE CHANGE
CONFERENCE
DEC 7-DEC 18
2009

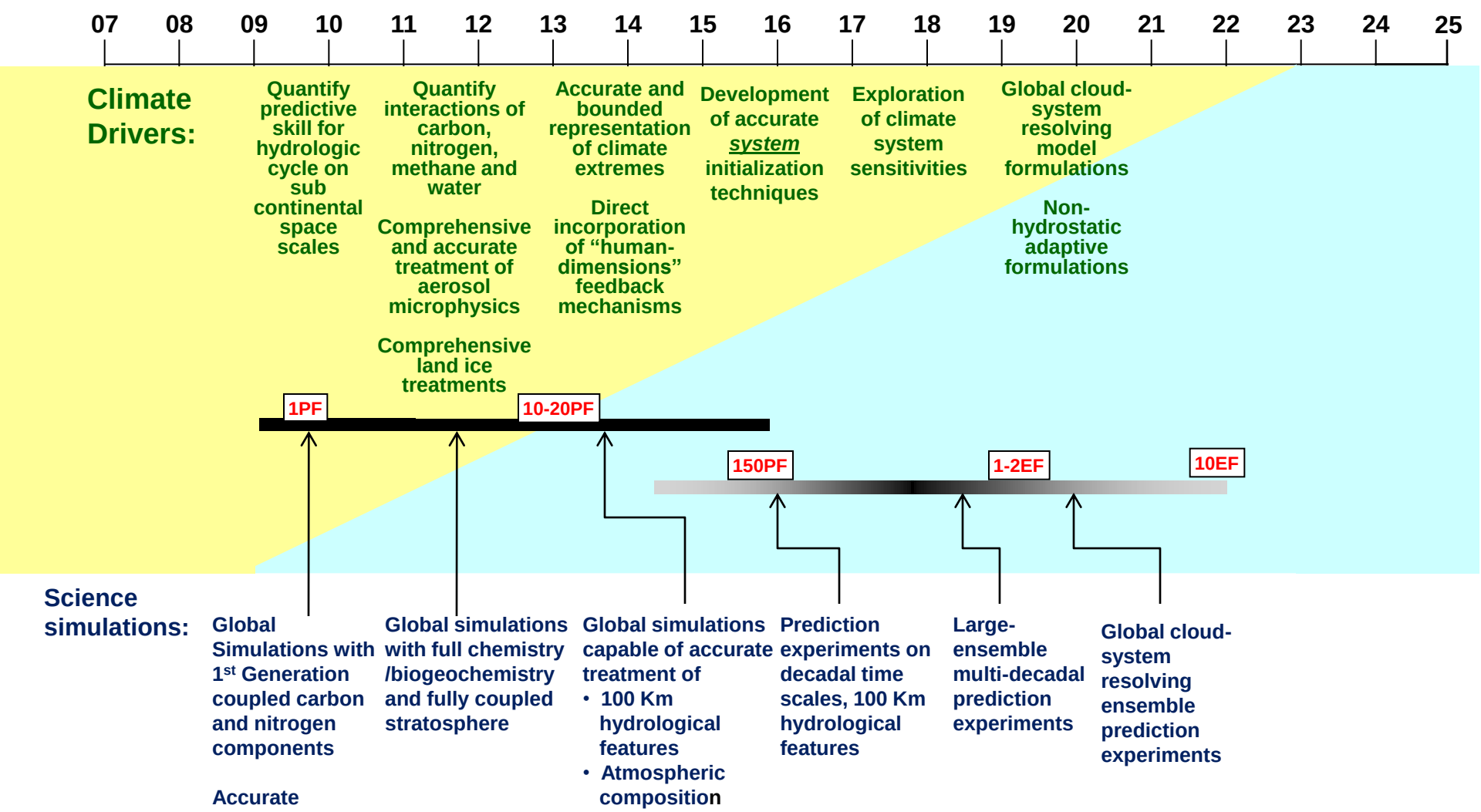
http://www.ipcc.ch

http://en.cop15.dk/

Publicly Released, Feb 2010

U.S. DEPARTMENT OF ENERGY

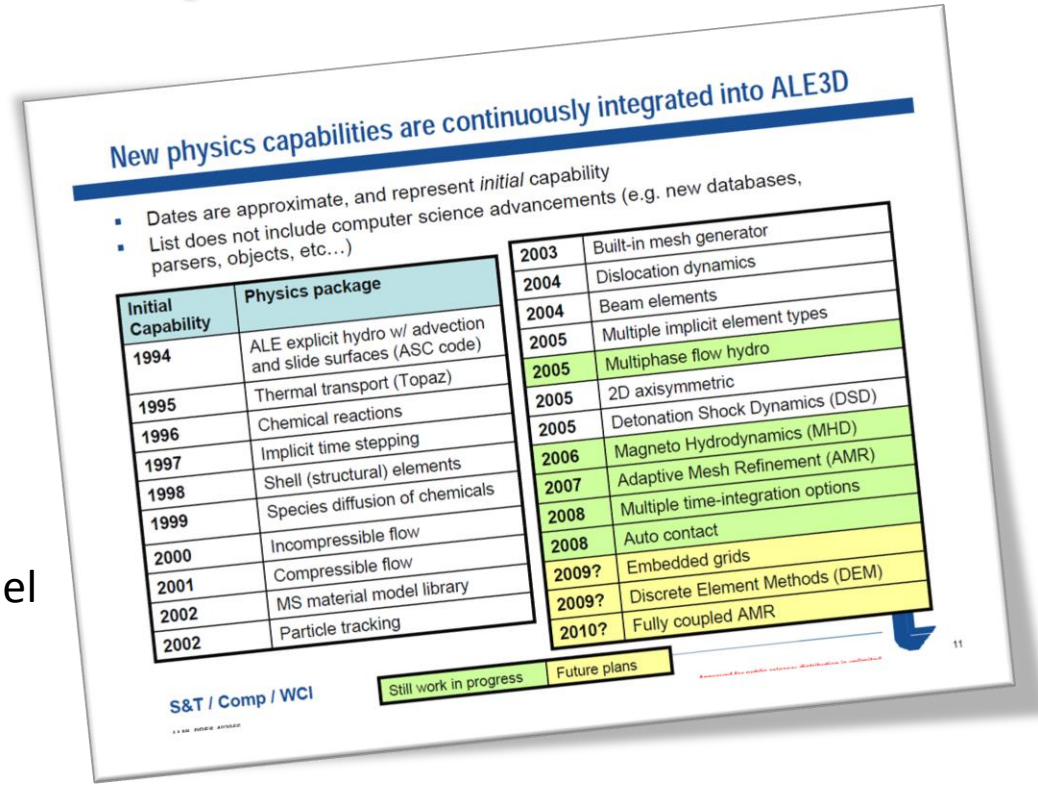
Computational science roadmap for a predictive, integrated climate model



Source: DOE Exascale Initiative
http://extremecomputing.labworks.org/

Application Complexity…

- **Today's applications are quickly becoming more complex than earlier applications**
  - DOE Exascale Initiative [1]
- **Increased fidelity**
- **Answering more complex questions**
  - Scientists [2,3] working on model that combines
    - Climate model
    - Energy economics
    - Population movements, demographics



New physics capabilities are continuously integrated into ALE3D

- Dates are approximate, and represent *initial* capability
- List does not include computer science advancements (e.g. new databases, parsers, objects, etc…)

Initial Capability	Physics package
1994	ALE explicit hydro w/ advection and slide surfaces (ASC code)
1995	Thermal transport (Topaz)
1996	Chemical reactions
1997	Implicit time stepping
1998	Shell (structural) elements
1999	Species diffusion of chemicals
2000	Incompressible flow
2001	Compressible flow
2002	MS material model library
2002	Particle tracking
2003	Built-in mesh generator
2004	Dislocation dynamics
2004	Beam elements
2005	Multiple implicit element types
2005	Multiphase flow hydro
2005	2D axisymmetric
2005	Detonation Shock Dynamics (DSD)
2006	Magneto Hydrodynamics (MHD)
2007	Adaptive Mesh Refinement (AMR)
2008	Multiple time-integration options
2008	Auto contact
2009?	Embedded grids
2009?	Discrete Element Methods (DEM)
2010?	Fully coupled AMR

Still work in progress | Future plans

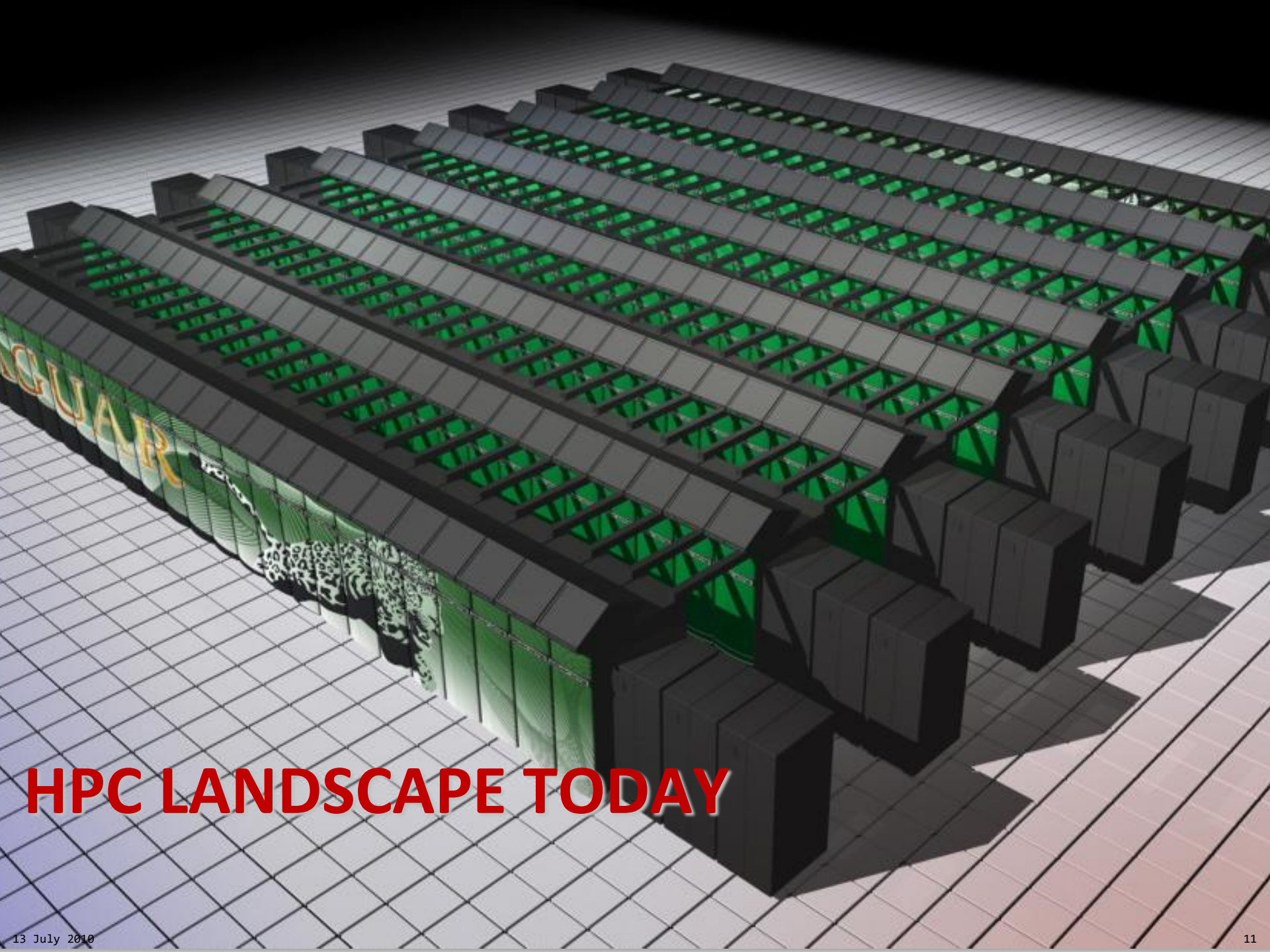
S&T / Comp / WCI

11

Source: LLNL ALE3D Team

Applications design and implementation are already complex;
Writing and optimizing code for each new architecture and programming model is impractical (and only going to happen w/ heroic efforts/funding.)

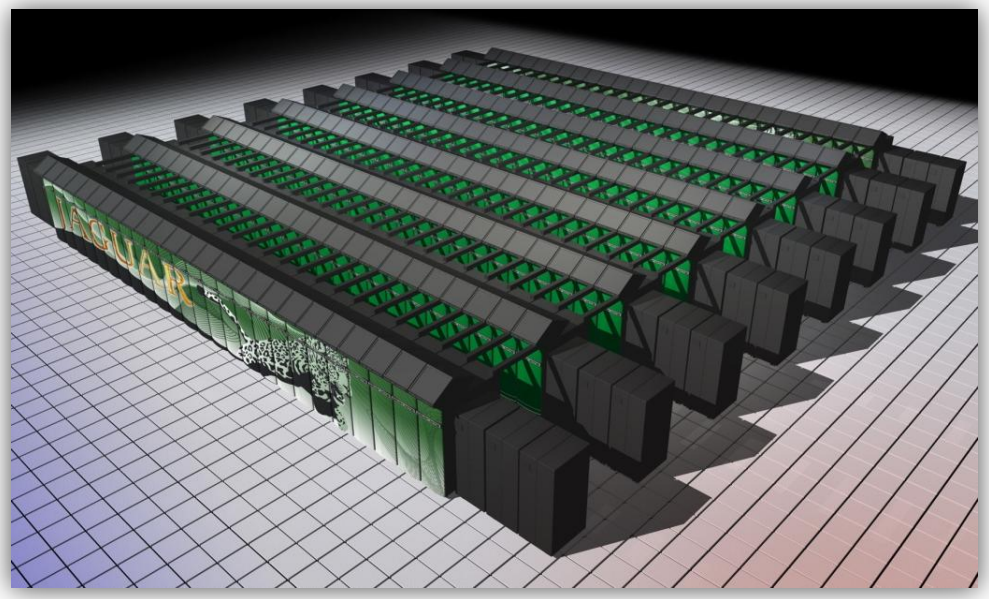
[1] DOE Exascale Initiative, 2010.
[2] S.W. Hadley, D.J. Erickson et al., "Responses of energy use to climate change: A climate modeling study," *Geophysical Research Letters*, 33(17), 2006.
[3] D. Vergano, "Air conditioning fuels warming," in USA Today, 2006.

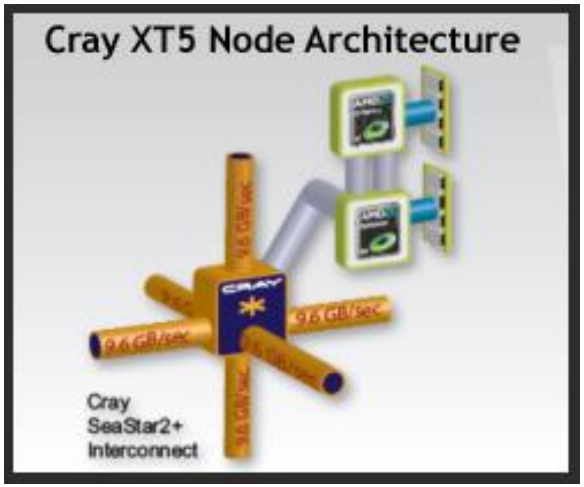


HPC LANDSCAPE TODAY

HPC Systems Today

- **TOP500 #1: Jaguar @ ORNL**
  - Cray XT5
  - 2.3 Petaflops
  - 224k cores, AMD Istanbul
  - 300 TB Main memory
  - Seastar 2.1+
    - 9.6 GBps/link
  - Storage system
    - Lustre
    - 10 PB, 240 GBps
  - ~200 cabinets
  - 4,400 sq ft
  - 7 MWs during HPL run
  - Liquid cooled



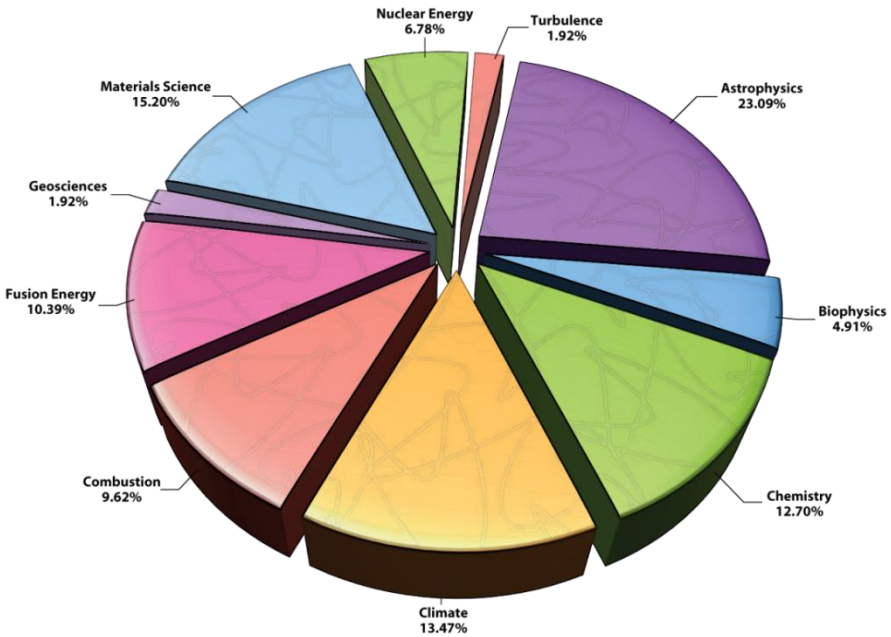


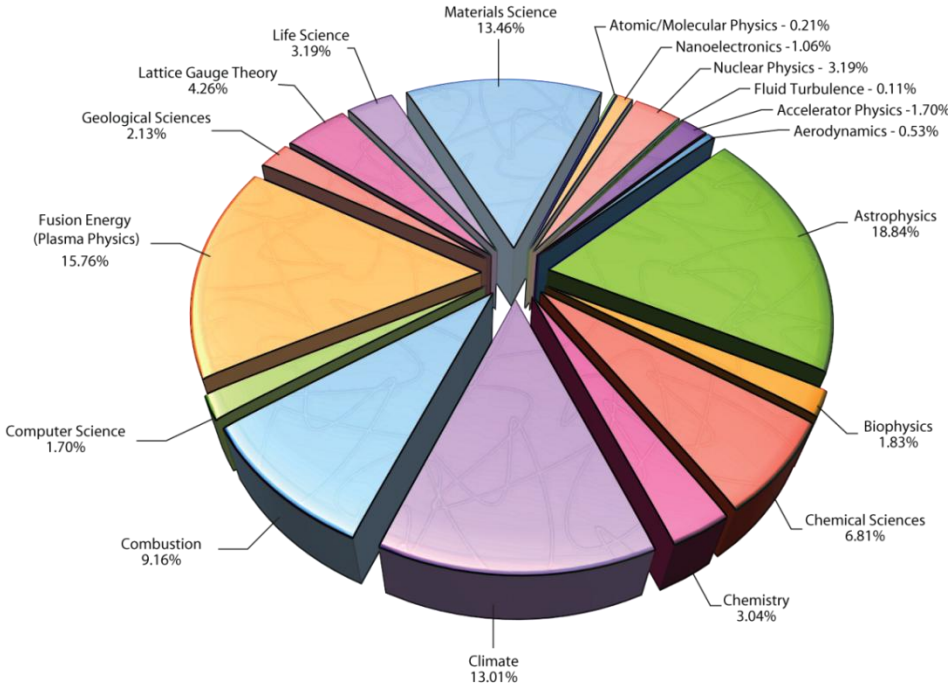
http://nccs.gov

Typical Jaguar XT5 science workload

- **Current: Transition to operations**
- **27 early-access projects, January–July**
  - Total allocation: 540.8M hours
  - Usage since January: 247M hours (159 users)
- **ALCC project: FY09 Joule metric applications**
  - Total allocation: 25M hours
  - Usage since January: 17M hours (17 users)

- **Fall 2009: General availability**
- **38 INCITE projects transitioned to system**
  - Total allocation: 469M hours
  - Usage (XT4) since January: 91M hours (448 users)
- **Discretionary projects: Climate AR5 production**
  - Total allocation: 80M hours
  - 2 projects: NCAR/DOE and NOAA (~50 users)



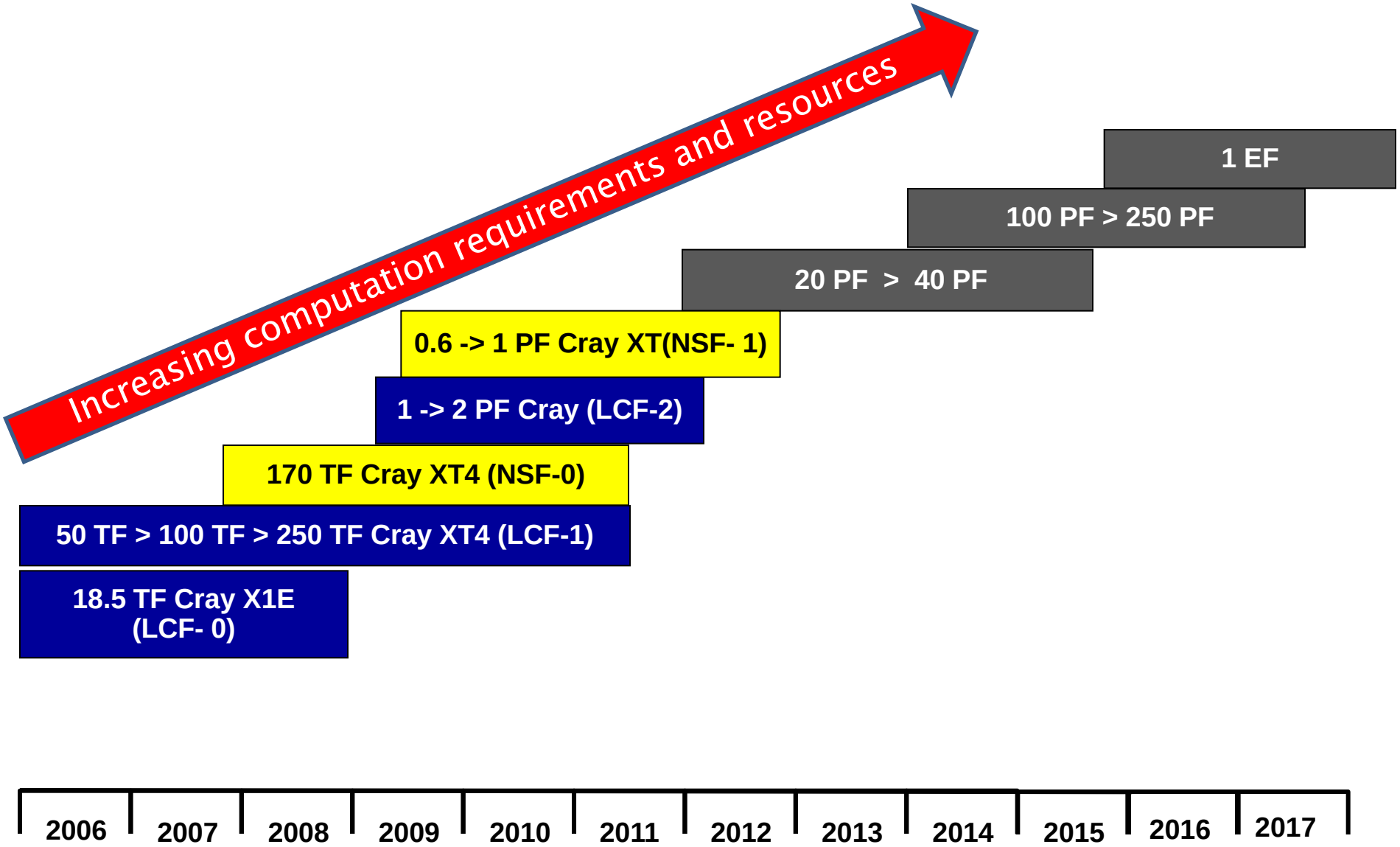


Science applications are scaling on Jaguar

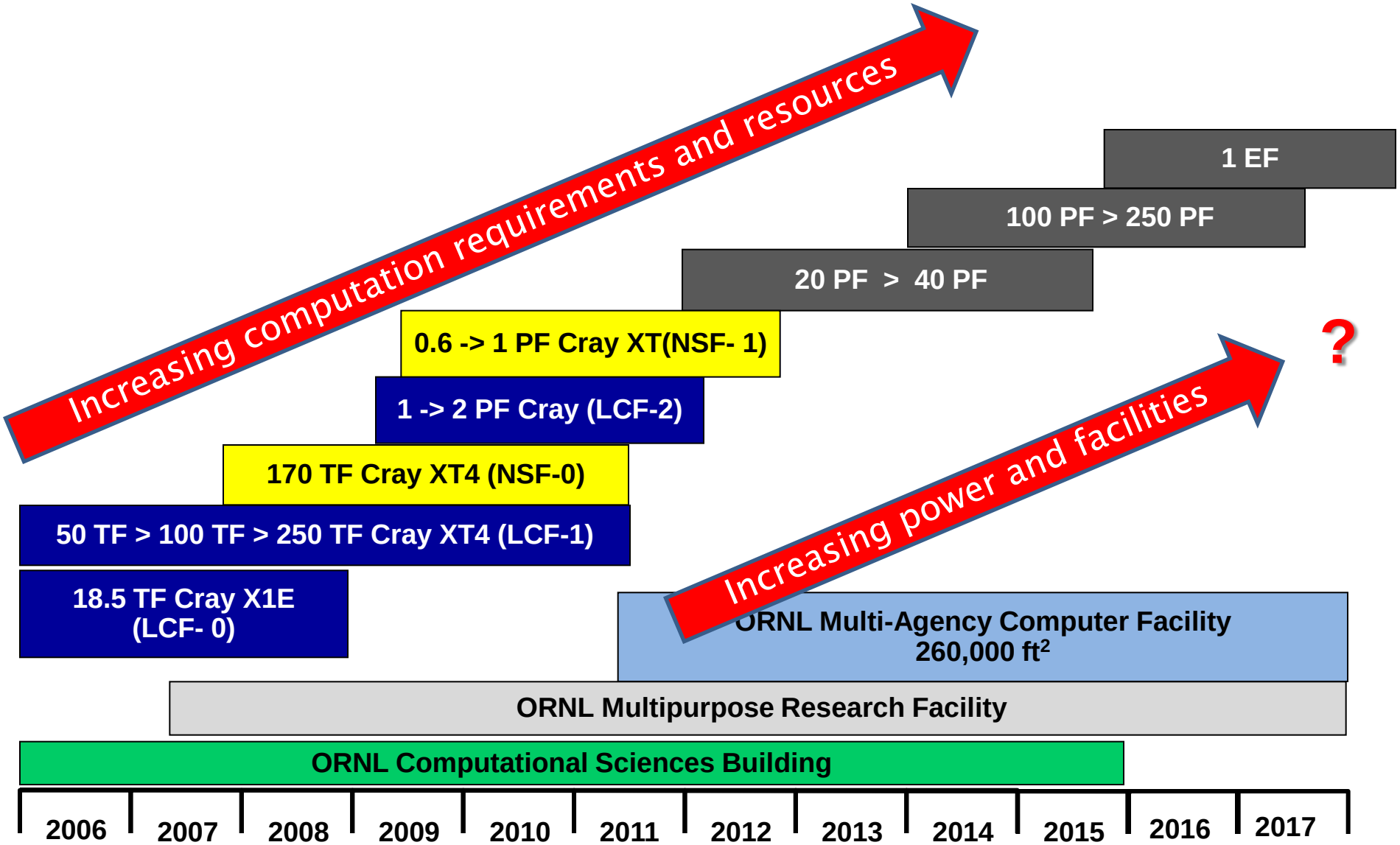
Science area	Code	Contact	Cores	Performance
Materials	**DCA++**	**Schulthess**	**150,144**	**1.3 PF MP**
Materials	LSMS	Eisenbach	149,580	1.05 PF
Seismology	**SPECFEM3D**	**Carrington**	**149,784**	**165 TF**
Weather	WRF	Michalakes	150,000	50 TF
Climate	POP	Jones	18,000	20 simulation years/day
Combustion	S3D	Chen	144,000	83 TF
Fusion	GTC	PPPL	102,000	20 billion particles/second
Materials	**LS3DF**	**Lin-Wang Wang**	**147,456**	**442 TF**
Chemistry	NWChem	Apra	96,000	480 TF
Chemistry	MADNESS	Harrison	140,000	550+ TF


THE ROAD AHEAD

ORNL Roadmap to Exascale



ORNL Roadmap to Exascale



U.S. DEPARTMENT OF ENERGY

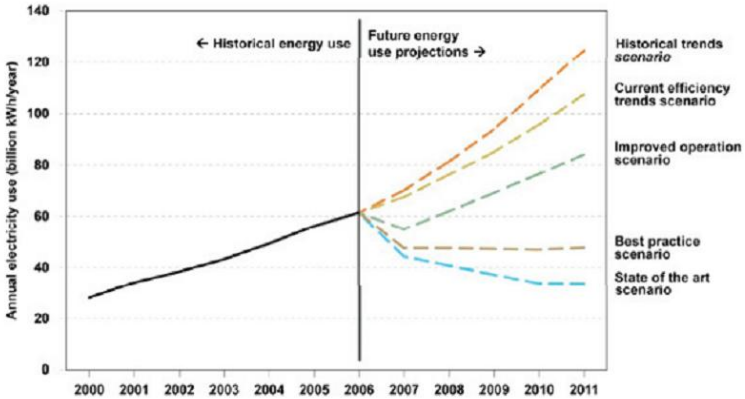
Notional System Architecture Targets and "swim lanes"

System attributes	2010	"2015"		"2018"	
System peak	2 Peta	200 Petaflop/sec		1 Exaflop/sec	
Power	6 MW	15 MW		20 MW	
System memory	0.3 PB	5 PB		32-64 PB	
Node performance	125 GF	0.5 TF	7 TF	1 TF	10 TF
Node memory BW	25 GB/s	0.1 TB/sec	1 TB/sec	0.4 TB/sec	4 TB/sec
Node concurrency	12	O(100)	O(1,000)	O(1,000)	O(10,000)
System size (nodes)	18,700	50,000	5,000	1,000,000	100,000
Total Node Interconnect BW	1.5 GB/s	150 GB/sec	1 TB/sec	250 GB/sec	2 TB/sec
MTTI	day	O(1 day)		O(1 day)	

Tough Decisions for Managing Power and Facilities Demands

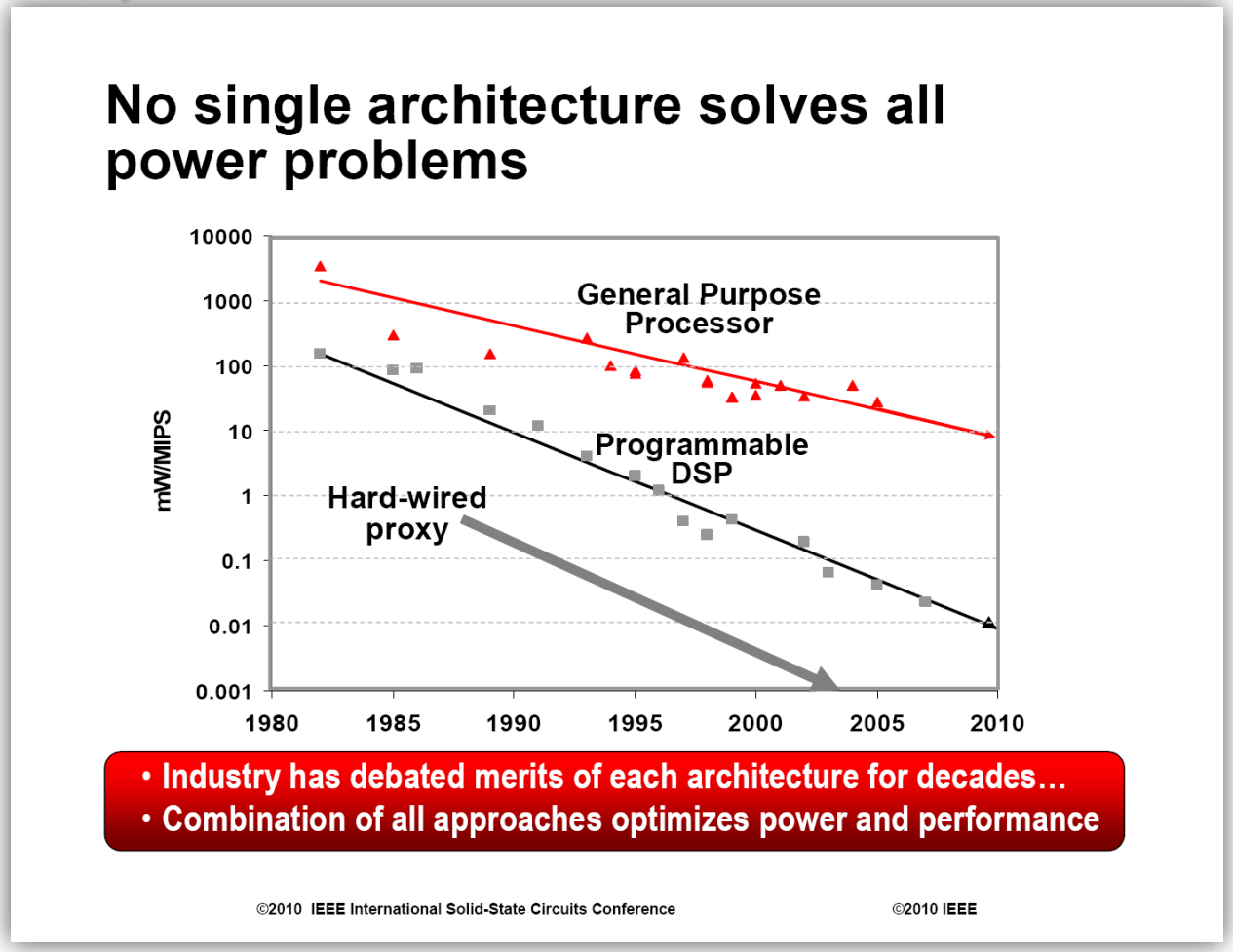
- **Build bigger buildings and plan to pay $$$ for ops**
- **Improve efficiencies**
  - Power distribution
  - Workload scheduling
  - Software
- **Use architectures that 'match' your workload**
- **Design new underlying technologies**
  - Optical networks
  - 3D stacking
  - MRAM, PCM, nanotubes

Comparison of Projected Electricity Use, All Scenarios, 2007 to 2011.



Source: EPA Report to Congress on Server and Data Center Energy Efficiency, 2007.

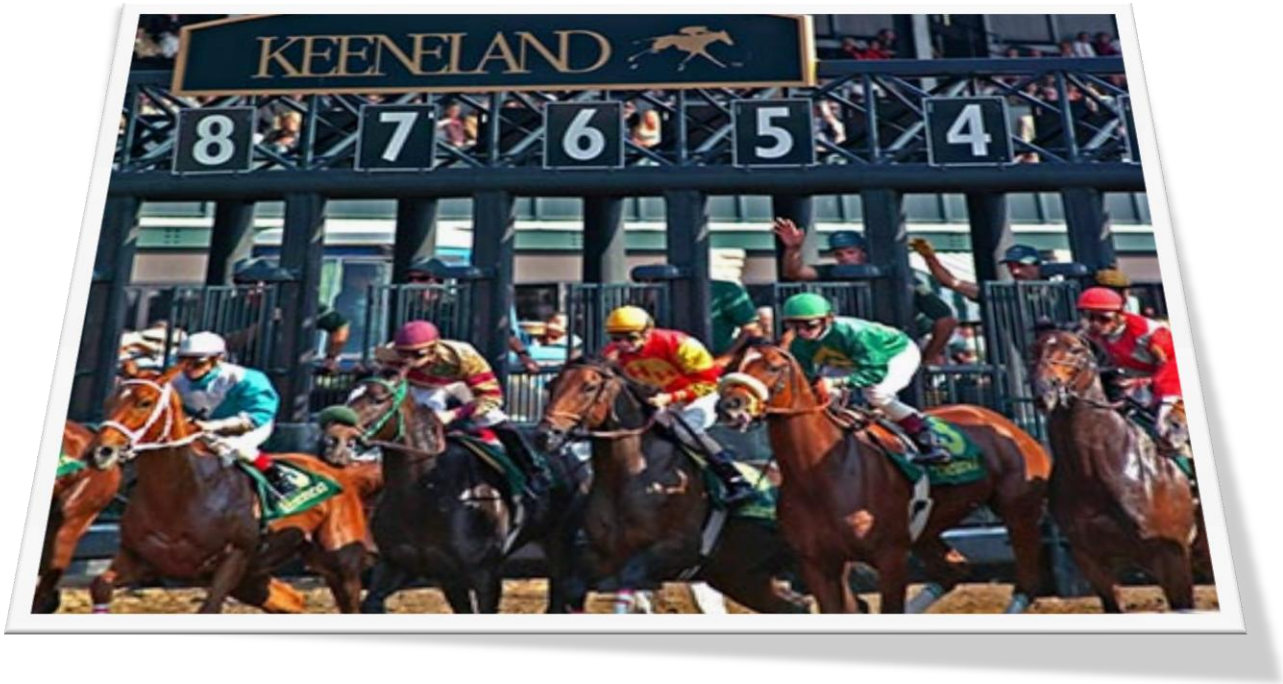
Heterogeneous architectures can offer better performance, power



- **Tukwila – First 2 billion transistor chip**
  - 80486 had ~1.2M transistors, ~50MHz, 1989
  - Specialization can be free

METHOD	BITRATE (MBITS/S)	RELATIVE PERFORMANCE
Intel Nehalem	252914	-
Bitslice	63262	4.0
8-bit LUT	42834	5.9
Naïve	3368	75.1

Source: Delagi, ISSCC 2010



ENABLING HETEROGENEOUS COMPUTING FOR THE OPEN SCIENCE COMMUNITY

NSF Office of Cyber Infrastructure RFP

- NSF 08-573 OCI Track 2D RFP in Fall 2008
  - Data Intensive
  - Experimental Grid testbed
  - Pool of loosely coupled grid-computing resources
  - Experimental HPC System of Innovative Design

An experimental high-performance computing system of innovative design. Proposals are sought for the development and deployment of a system with an architectural design that is outside the mainstream of what is routinely available from computer vendors. Such a project may be for a duration of up to five years and for a total award size of up to $12,000,000. It is not necessary that the system be deployed early in the project; for example, a lengthy development phase might be included. Proposals should explain why such a resource will expand the range of research projects that scientists and engineers can tackle and include some examples of science and engineering questions to which the system will be applied. It is not necessary that the design of the proposed system be useful for all classes of computational science and engineering problems. When finally deployed, the system should be integrated into the TeraGrid. It is is anticipated that the system, once deployed, will be an experimental TeraGrid resource, used by a smaller number of researchers than is typical for a large TeraGrid resource. (Up to 5 years duration. Up to $12,000,000 in total budget to include development and/or acquisition, operations and maintenance, including user support. First-year budget not to exceed $4,000,000.)

Oct 2008 Alternatives Analysis

- STI Cell (?)
- FGPAs
- Cyclops64 (?)
- Cray XMT
- Sun Rock/Niagara (?)
- ClearSpeed (?)
- Tensilica
- Tilera
- Anton
- SGI Molecule (?)
- Intel Larrabee (?)

- Graphics processors

- Others…

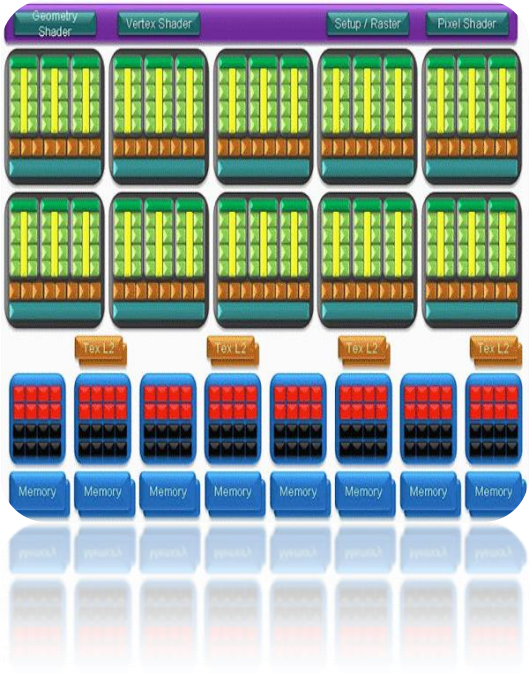
- Performance
  - Must show reasonable performance improvements at scale on real scientific applications of interest

- Programmability
  - Must be easy to re-port and re-optimize applications for each new architecture (generation) without large effort, delays
- Precision - Accuracy
  - Must provide impressive performance accurately
- Reliability
  - Must get high scientific throughput without job failures or inaccurate results

- Power and Facilities Cost
  - Must be reasonably affordable in terms of power and facilities costs

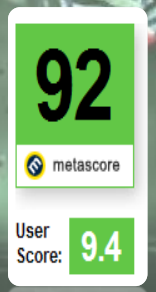
Alternatives analysis concluded GPUs were a competitive solution

- Success with various applications at DOE, NSF, government, industry
  - Signal processing, image processing, etc.
  - DCA++, S3D, NAMD, many others
- Community application experiences also positive
  - Frequent workshops, tutorials, software development, university classes
  - Many apps teams are excited about using GPGPUs



Batman: Arkham Asylum with PhysX!

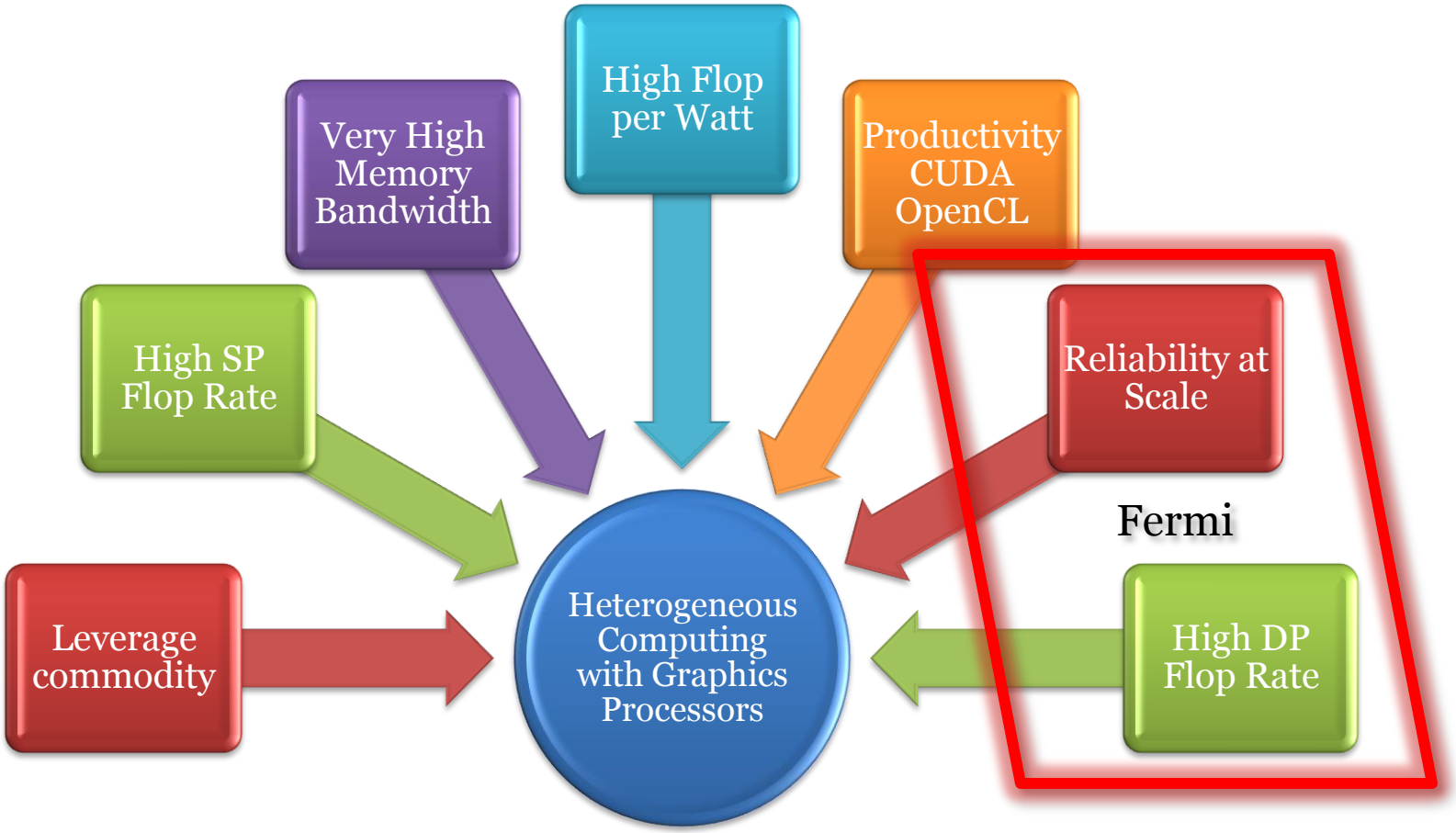
- Smoke reacts with Batman, provides cover
- Walls explode
- Glass shatters
- Tattered curtains react with characters



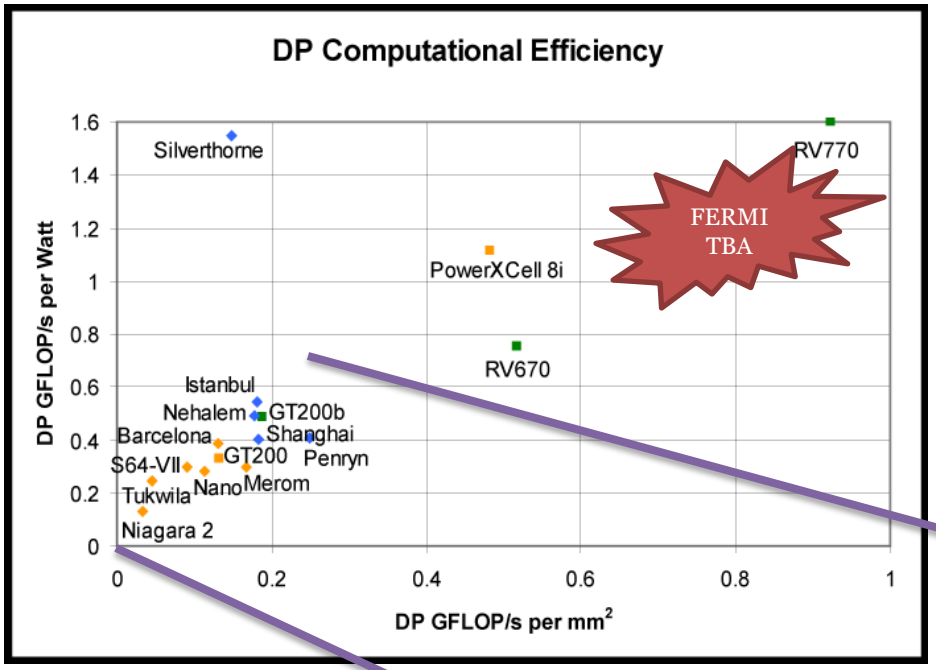
Highest rated PC game since 2007!

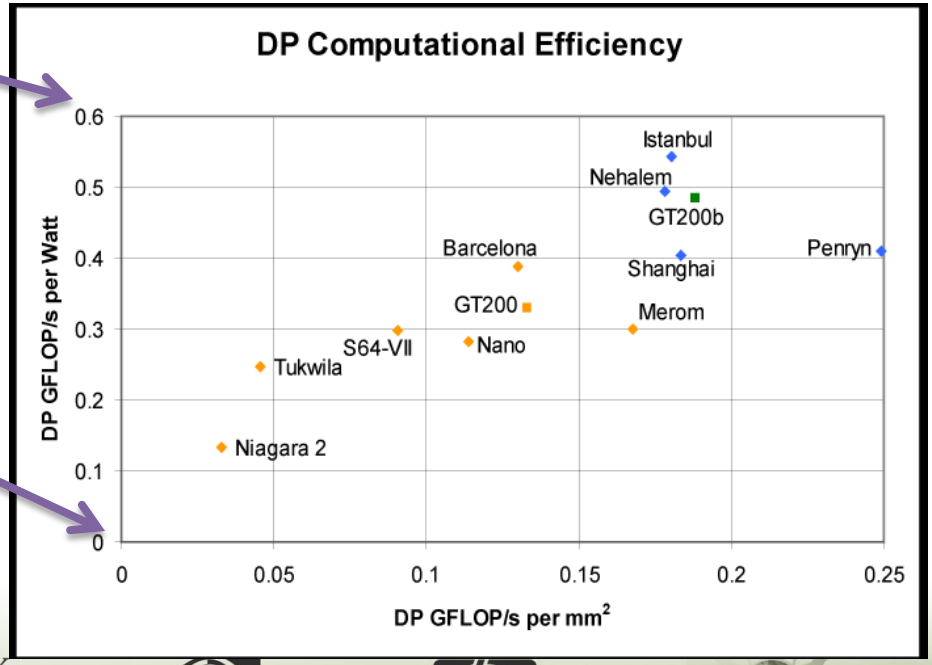
Courtesy of NVIDIA

GPU Rationale – What's different now?



And Power Efficiency!





Source: http://www.realworldtech.com

Keeneland Partners

Georgia Institute of Technology	National Institute of Computational Sciences	Oak Ridge National Laboratory	University of Tennessee, Knoxville	NVIDIA	HP
Project management	Operations and TG/XD Integration	Applications	Scientific Libraries	Tesla	HPC Host System
Acquisition and alternatives assessment	User and Application Support	Facilities	Education, Outreach, Training	Applications optimizations	System integration
System software and development tools	Operational Infrastructure	Education, Outreach, Training		Training	Training
Education, Outreach, Training	Education, Outreach, Training				

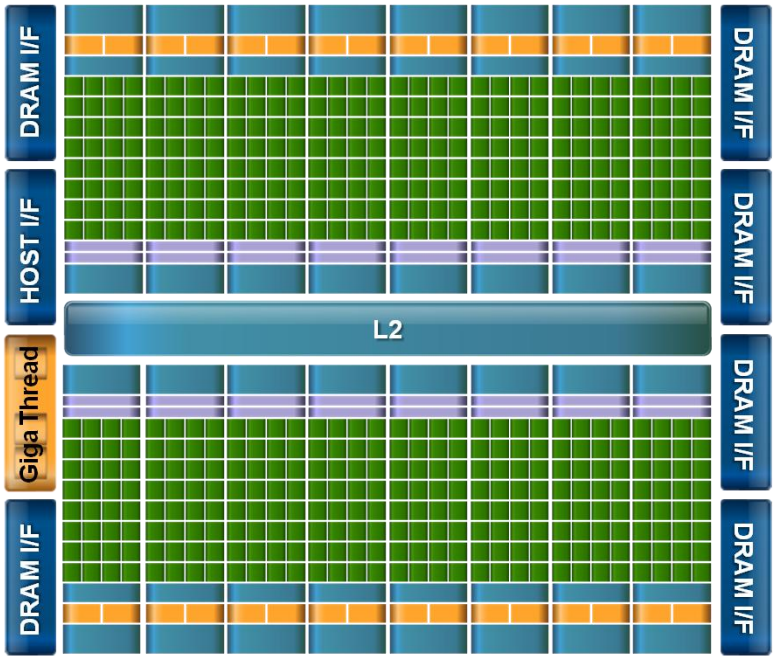
Keeneland Team

- **Jeffrey Vetter, Dick Glassbrook, Jack Dongarra, Richard Fujimoto, Thomas Schulthess, Karsten Schwan**, Sudha Yalamanchili, Kathlyn Boudwin, Jim Ferguson, Doug Hudson, Patricia Kovatch, Bruce Loftis, Jeremy Meredith, Jim Rogers, Philip Roth, Arlene Washington, Phil Andrews, Mark Fahey, Don Reed, Tracy Rafferty, Ursula Henderson, Terry Moore, Kyle, Spafford, and many others
- NVIDIA
- HP

- Keeneland Sponsor: NSF
- Other sponsors: DOE, DOD, DARPA

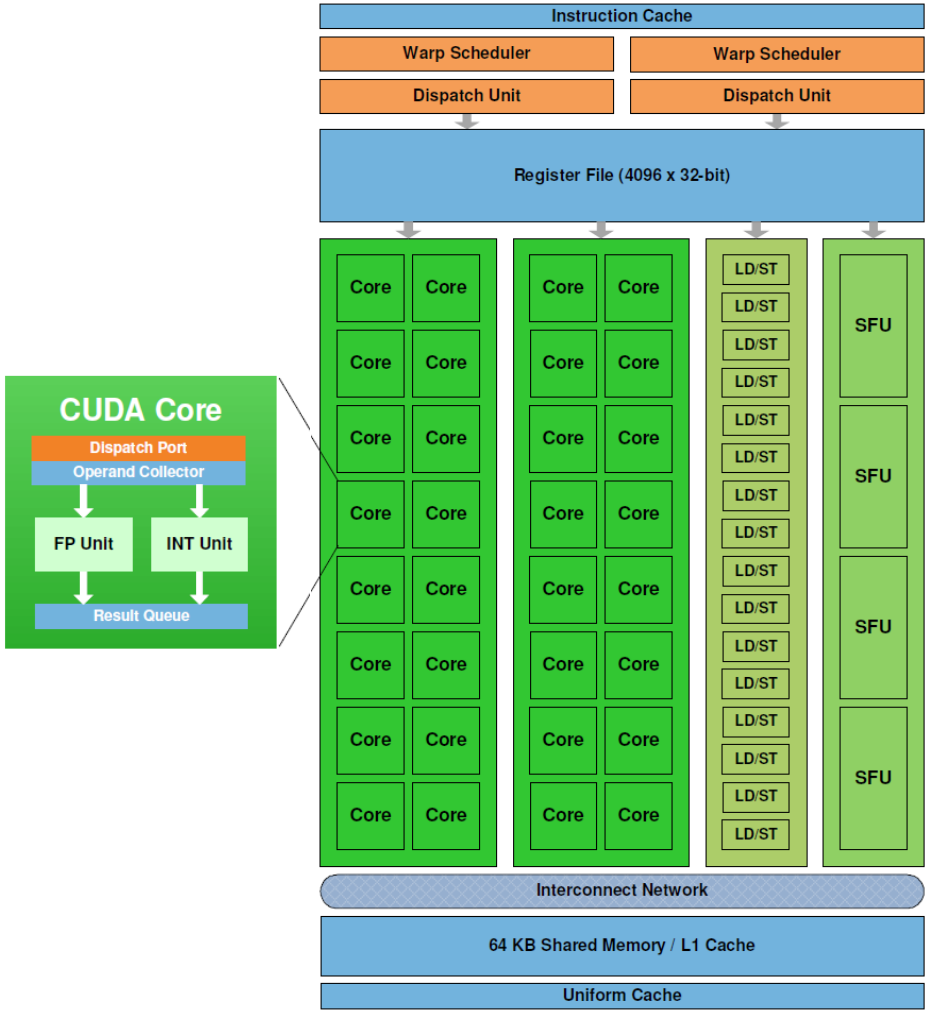
Keeneland Initial Delivery (KID) System

- KID Architecture
  - HP Linux Cluster
  - Fermi 6 GB
  - New integrated CPU/GPU design
  - Details to be announced soon
  - ~200 CPUs + ~300 GPUs
- Traditional Linux software stack augmented with GPU compilers, software tools, libraries
- Delivery and acceptance in Sept 2010



KID will use NVIDIA's Fermi

- 3B transistors
- ECC
- 8x the peak double precision arithmetic performance over NVIDIA's last generation GPU.
- 448 CUDA Cores featuring the new IEEE 754-2008 floating-point standard
- NVIDIA Parallel DataCache
- NVIDIA GigaThread Engine
- Debuggers, language support



Many GPU-enabled systems blooming everywhere ….

China's Defense University builds World Third fastest supercomputer

Posted on 29 October 2009 by admin ShareThis

October.29 (China Military News Reporting by Johnathan Weng) — The Chinese National University of Defense Technology (NUDT) unveiled Thursday China's fastest supercomputer, also the World Third fastest computer, which is able to do more than one quadrillion calculations per second theoretically at its peak speed.

China name this supercomputer
Tianjin. As matter of fact, Tianjin is

China announces that "Tianhe-I"
LINPACK score is 563.1 teraFLOP
(1.0 petaFLOPS) of Jülich Resear

engadget CLASSIC MOBILE HD ALT

DROID X REVIEW | IPHONE 4 REVIEW | NINTENDO 3DS | MICROSOFT AT E3

NEWS HUBS GALLERIES VIDEOS PODCASTS TOPICS REVIEW

FILED UNDER *Desktops*

Tokyo Institute of Technology announces SSD-pac 2.39 petaflop supercomputer

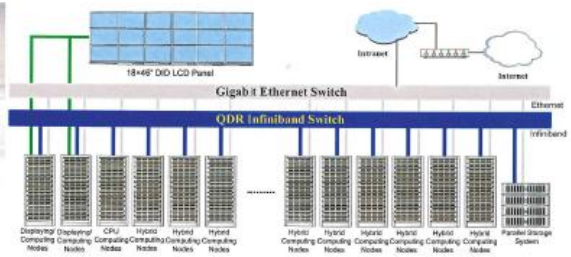
By Joseph L. Flatley posted Jun 23rd 2010 2:06PM

IBM has announced plans to start using SandForce SSDs in its enterprise machines, and now it looks like the Tokyo Institute of Technology is doing one better, working with NEC and HP to produce Tsubame 2.0. This next-gen supercomputer will reportedly operate at 2.39 petaflops (that's a lot of flops!) and uses a new multilevel storage architecture consisting of DRAM as well as SSDs. Not only will this bad boy have thirty times the computing capacity of Tsubame 1.0 (due in part to its some 2,816 Intel Westmere microprocessors and 4,224 NVIDIA Tesla M2050 GPUs), its power draw should be some 1/25th of its predecessor's. If all goes according to plan, it should be in operation this fall, at a cost of ¥3.2 billion (approx $35.5 million).

Development and application of a HPC system for multi-scale discrete simulation—*Mole-8.5*

Xiaowei Wang, Wei Ge, Xianfeng He, Feiguo Chen, Li Guo, Jinghai Li
Institute of Process Engineering, Chinese Academy of Sciences, Beijing, 100190

Mole-8.5 is the first GPGPU supercomputer (Rpeak of about 1100 Tflops) using NVIDIA Tesla C2050 in the world, which includes 372 nodes and is established in April 2010. It is the successor of the first supercomputer with 1.0 Petaflops peak performance in single precision in China, which was a hybrid system including four units integrating NVIDIA and AMD GPUs announced on April 20, 2009. Mole-8.5 was designed and established by Institute of Process Engineering (IPE), Chinese Academy of Sciences, one of the NVIDIA CCOEs. A designing philosophy utilizing the similarity between hardware, software and the problems to be solved is embodied, based on the multi-scale method and discrete simulation approaches developed at IPE. The whole system is connected with Gigabit Ethernet and QDR Infiniband network. *Mole-8.5* has some unique advantages over the HPC system of same performance based on CPU, for example the high performance/price ratio, the area occupied by it is only about 150 M^2. The linpack result of 320 nodes of Mole-8.5 is 2.073e+05 Gflops with a power consumption of about 480 KWatt, therefore the average power efficiency is 431 Mflop/s/Watt, manifesting an energy efficient supercomputer.



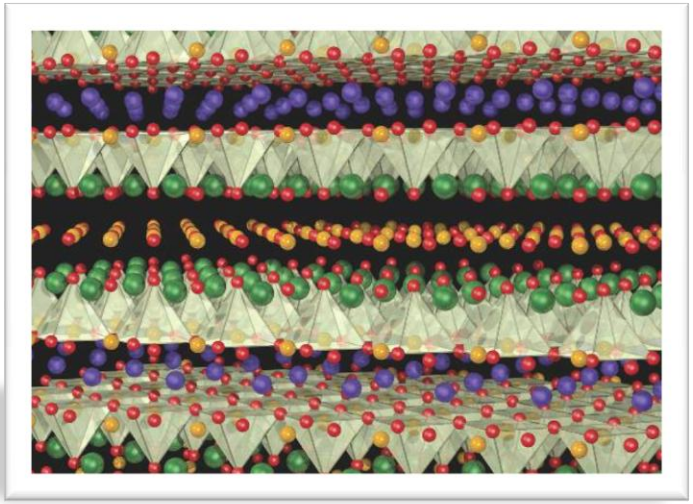
The GPGPU supercomputer *Mole-8.5*

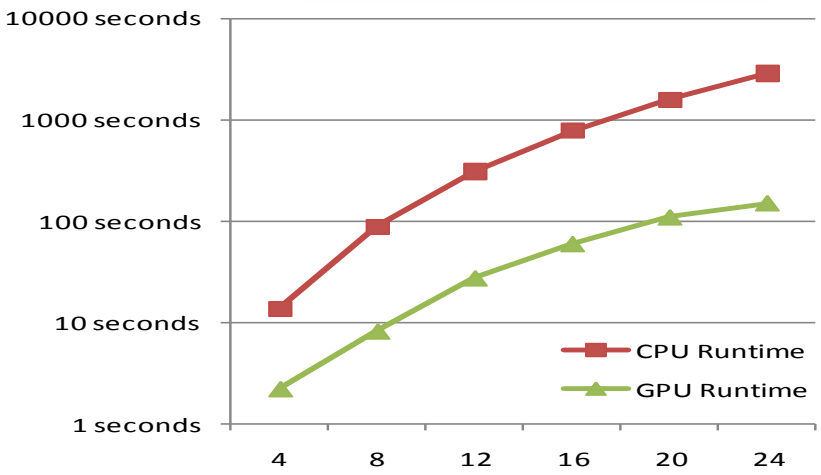
1. Oak Ridge LCF3
2. More announcements expected in coming months …

APPLICATIONS

Computational Materials - Case Study

- Quantum Monte Carlo simulation
  - High-temperature superconductivity and other materials science
  - 2008 Gordon Bell Prize
- GPU acceleration speedup of 19x in main QMC Update routine
  - Single precision for CPU and GPU: target single-precision only cards
  - Required detailed accuracy study and mixed precision port of app
- Full parallel app is 5x faster, start to finish, on a GPU-enabled cluster on Tesla T10



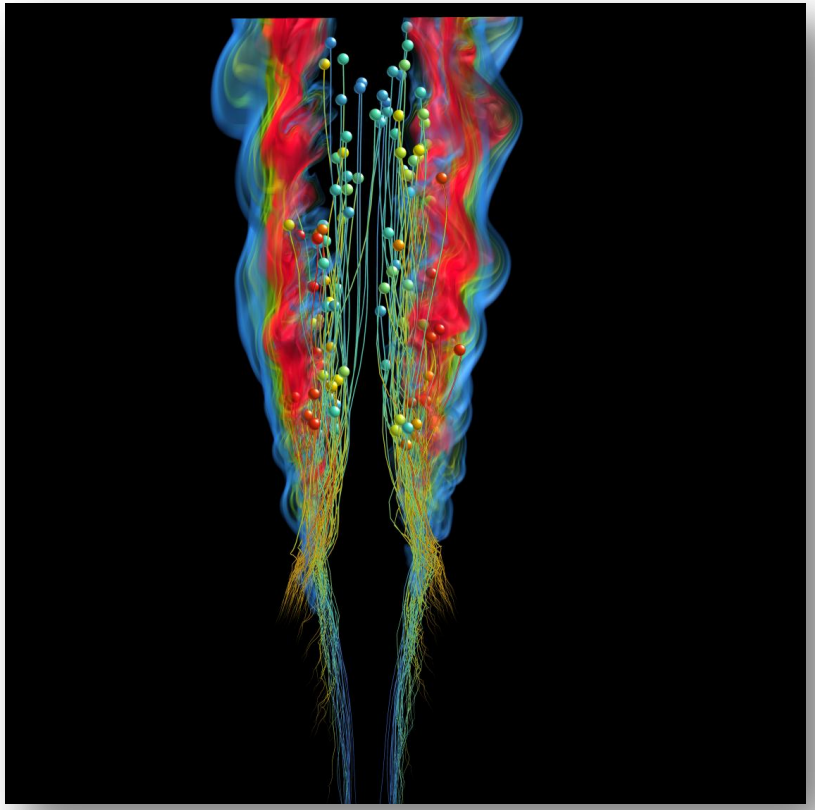


GPU study: J.S. Meredith, G. Alvarez, T.A. Maier, T.C. Schulthess, J.S. Vetter, "Accuracy and Performance of Graphics Processors: A Quantum Monte Carlo Application Case Study", *Parallel Comput., 35(3):151-63, 2009.*

Accuracy study: G. Alvarez, M.S. Summers, D.E. Maxwell, M. Eisenbach, J.S. Meredith, J. M. Larkin, J. Levesque, T. A. Maier, P.R.C. Kent, E.F. D'Azevedo, T.C. Schulthess, "New algorithm to enable 400+ TFlop/s sustained performance in simulations of disorder effects in high-Tc superconductors", SuperComputing, 2008. [*Gordon Bell Prize winner*]

Combustion with S3D – Case Study

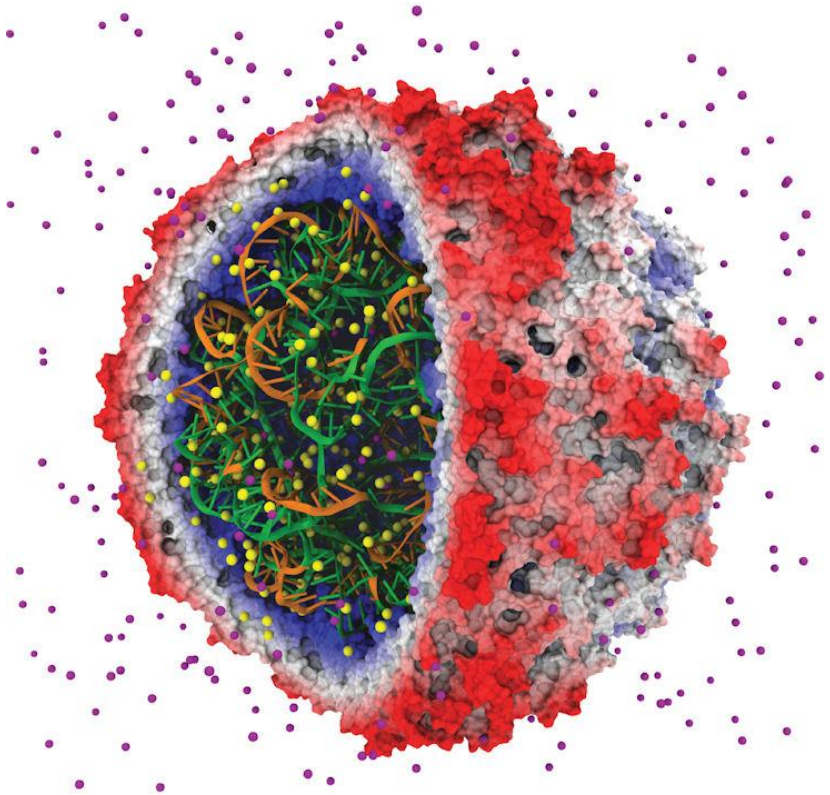
- Application for combustion - S3D
  - Massively parallel direct numerical solver (DNS) for the full compressible Navier-Stokes, total energy, species and mass continuity equations
  - Coupled with detailed chemistry
  - Scales to 150k cores on Jaguar
- Accelerated version of S3D's Getrates kernel in CUDA on Tesla T10
  - 31.4x SP speedup
  - 16.2x DP speedup



K. Spafford, J. Meredith, J. S. Vetter, J. Chen, R. Grout, and R. Sankaran. Accelerating S3D: A GPGPU Case Study. Proceedings of the Seventh International Workshop on Algorithms, Models, and Tools for Parallel Computing on Heterogeneous Platforms (HeteroPar 2009) Delft, The Netherlands.

Biomolecular systems from NAMD Team – Not just us

- NAMD, VMD
  - Study of the structure and function of biological molecules
- Calculation of non-bonded forces on GPUs leads to 9x speedup on Tesla T10
- Framework hides most of the GPU complexity from users



J.C. Phillips and J.E. Stone, "Probing biomolecular machines with graphics processors," *Commun. ACM, 52(10):34-41, 2009.*

Many others … Including Skeptics ?

Debunking the 100X GPU vs. CPU Myth: An Evaluation of Throughput Computing on CPU and GPU

Victor W Lee†, Changkyu Kim†, Jatin Chhugani†, Michael Deisher†, Daehyun Kim†, Anthony D. Nguyen†, Nadathur Satish†, Mikhail Smelyanskiy†, Srinivas Chennupaty⋆, Per Hammarlund⋆, Ronak Singhal⋆ and Pradeep Dubey†

victor.w.lee@intel.com

†Throughput Computing Lab, Intel Corporation

⋆Intel Architecture Group, Intel Corporation

ABSTRACT

Recent advances in computing have led to an explosion in the amount of data being generated. Processing the ever-growing data in a timely manner has made throughput computing an important aspect for emerging applications. Our analysis of a set of important throughput computing kernels shows that there is an ample amount of parallelism in these kernels which makes them suitable for today's multi-core CPUs and GPUs. In the past few years there have been many studies claiming GPUs deliver substantial speedups (between 10X and 1000X) over multi-core CPUs on these kernels. To understand where such large performance difference comes from, we perform a rigorous performance analysis and find that after applying optimizations appropriate for both CPUs and GPUs the performance gap between an Nvidia GTX280 processor and the Intel Core i7 960 processor narrows to only 2.5x on average. In this pa-

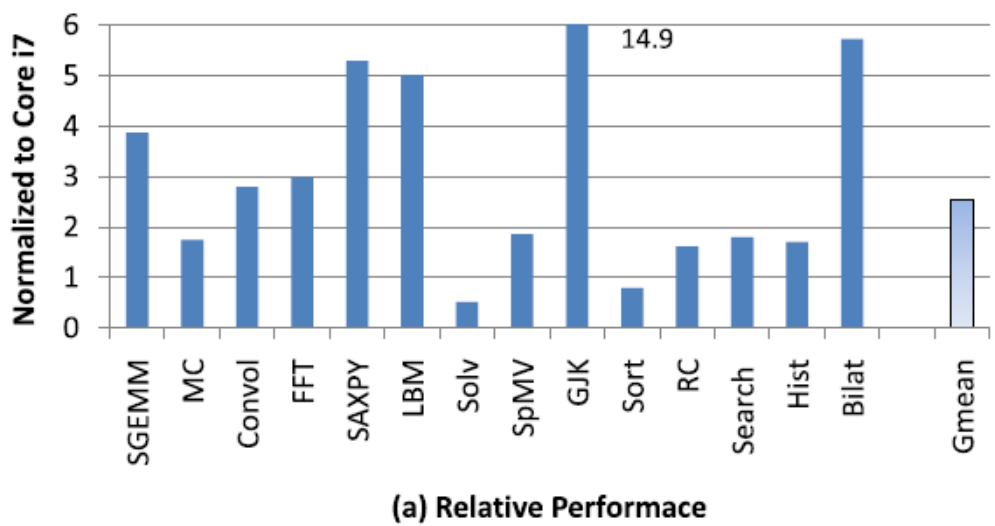


Figure 1: Comparison between Core i7 and GTX280 Performance.

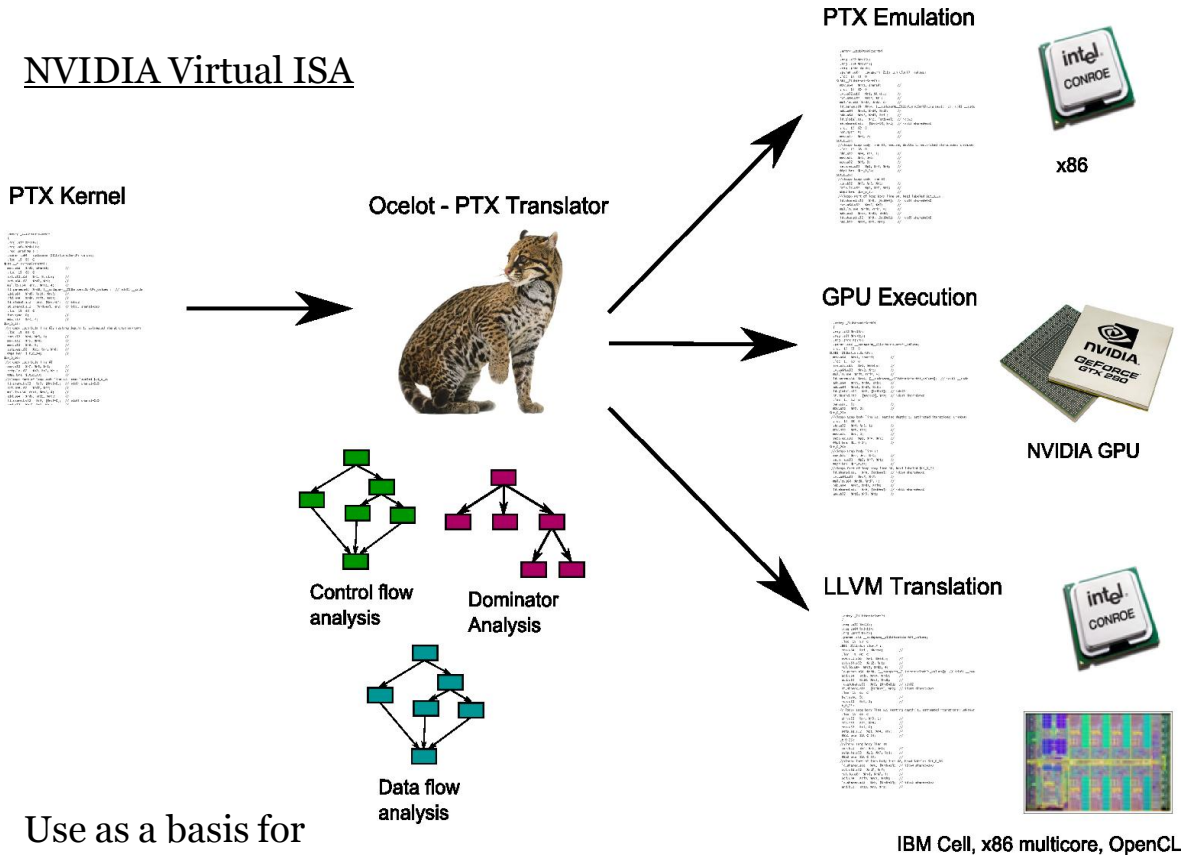
KEENELAND SOFTWARE

Keeneland Software Environment

- Integrated with NSF TeraGrid/XD
  - Including TG and NICS software stack
- Programming environments
  - CUDA
  - OpenCL
  - Compilers
    - PGI
      - Accelerate, CUDA Fortran
    - OpenMP 3.0
  - Scalable debuggers
  - Performance tools
- Additional software activities
  - Benchmarks
  - Performance and correctness tools
  - Scientific libraries
  - Virtualization

Ocelot: Dynamic Execution Infrastructure

NVIDIA Virtual ISA



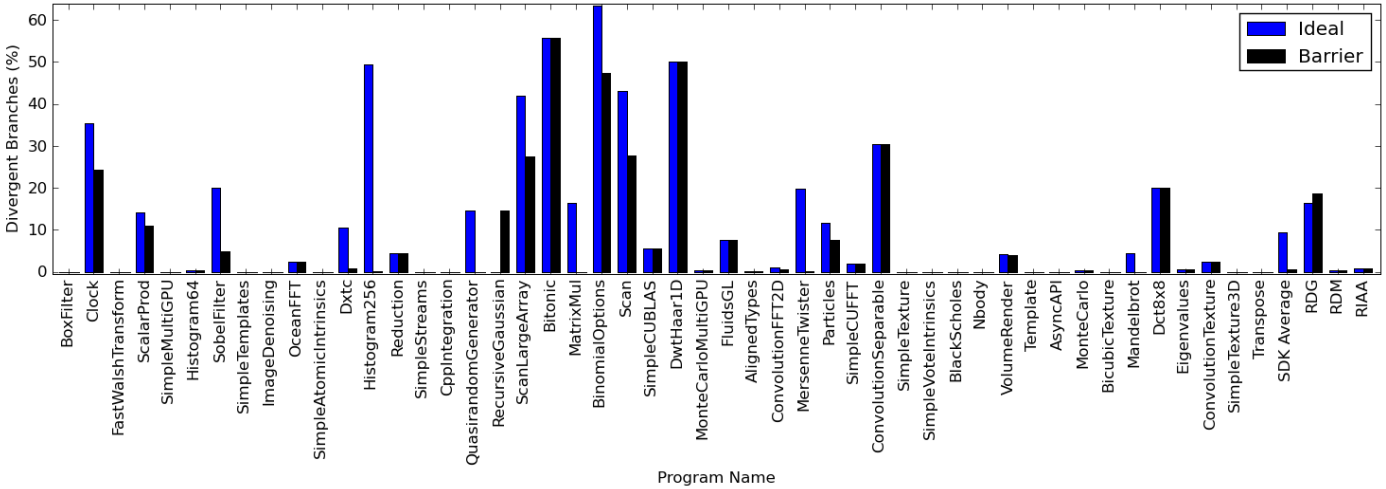
PTX 1.4 compliant Emulation
- Validated on full CUDA SDK
- Open Source version released

http://code.google.com/p/gpuocelot/

Use as a basis for
- Insight → workload characterization
- Performance tuning → detecting memory bank conflicts
- Debugging → illegal memory accesses, out of bounds checks, etc.

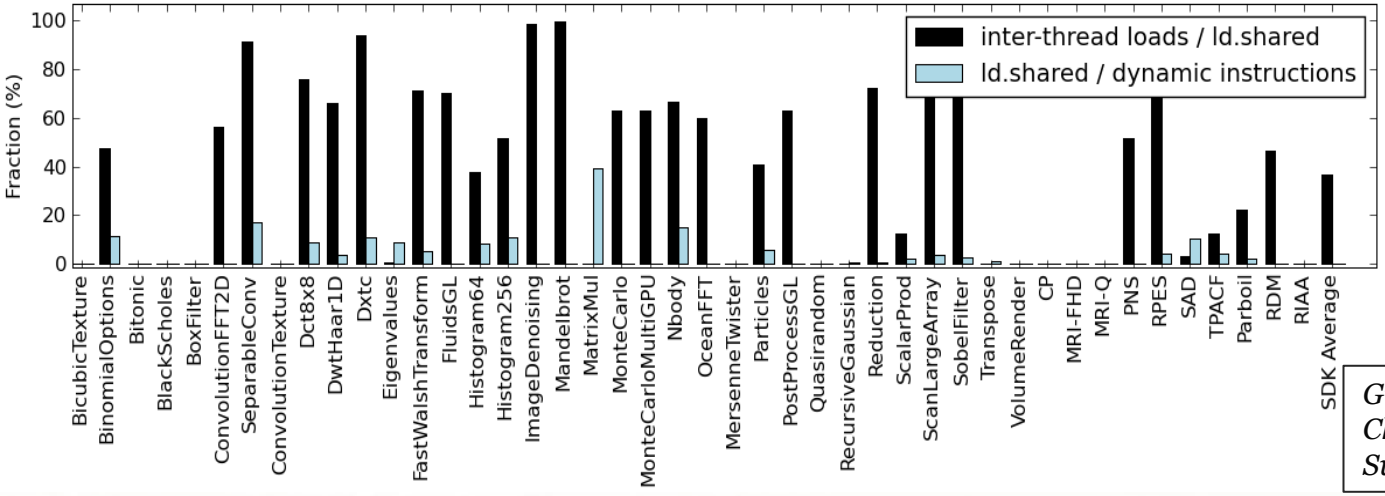
Gregory Diamos, Dhuv Choudhary, Andrew Kerr, Sudhakar Yalamanchili

Workload Analysis: Examples



Branch Divergence

- Study of control Flow behavior
- Motivate synchronization support



Inter-thread Data Flow

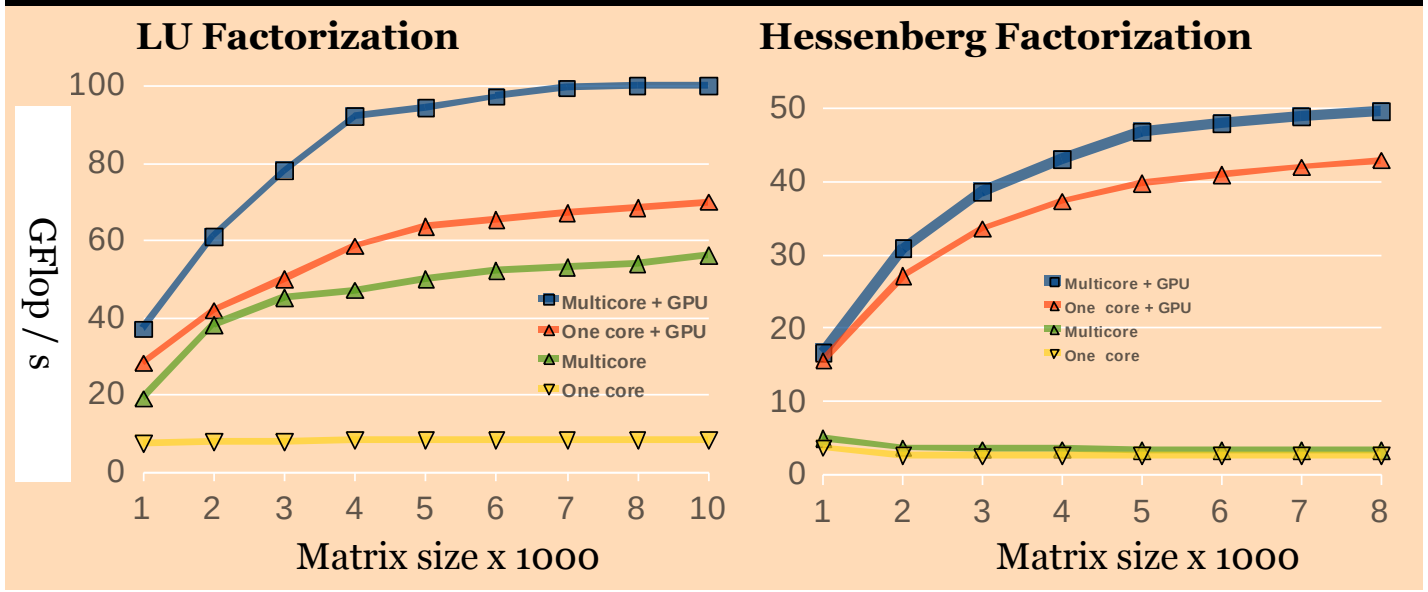
- Study of data sharing patterns
- Motivate architectural support

Gregory Diamos, Dhuv Choudhary, Andrew Kerr, Sudhakar Yalamanchili

Libraries: One and two-sided Multicore+GPU Factorizations

- These will be included in up-coming MAGMA releases
- Two-sided factorizations can not be efficiently accelerated on homogeneous x86-based multicores (above) because of memory-bound operations
  - MAGMA provided hybrid algorithms that overcome those bottlenecks (16x speedup!)

Multicore + GPU Performance in double precision



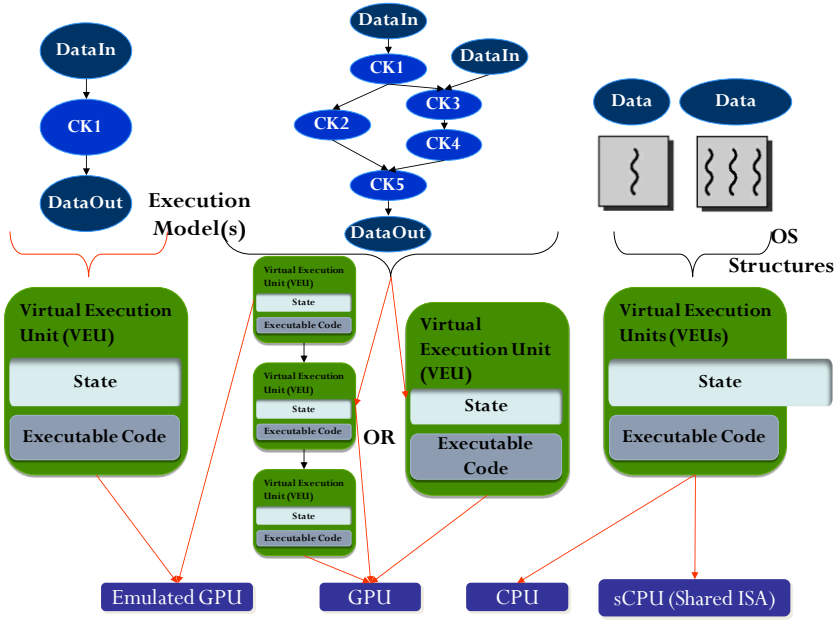
Jack Dongarra, Stan Tomov, and Rajib Nath

GPU : NVIDIA GeForce GTX 280
CPU : Intel Xeon dual socket quad-core @2.33 GHz

GPU BLAS : CUBLAS 2.2, dgemm peak: 75 GFlop/s
CPU BLAS : MKL 10.0 , dgemm peak: 65 GFlop/s

Hybrid Virtual Machine: HyVM

- Uniform runtime model for heterogeneous platforms:
- Hybrid Virtual Machines
  - Uniformity:
    - Virtual Execution Unit (VEU): hypervisor-level, uniform runtime representation for program executables
    - Heterogeneity-aware hypervisors: VMM-level management methods for improved platform utilization (incl. cache and energy) and application performance (SLAs)
    - Dynamic platform emulation: runtime CK compilation or re-writing for diverse accelerator targets (via LLVM)
  - High performance:
    - Commodity and custom VEU 'containers': Virtual Machines (VMs) – processes/threads – commodity cores; Special Execution Environments (e.g., NVIDIA) - Computational Kernels (CKs) - accelerators
    - Runtime and adaptive {CK} optimization for parallelism
    - Standards-compliant CK programming and runtime APIs (OpenCL, CUDA)
    - Compiler-based optimization techniques for {CK}



Ada Gavrilovska, Karsten Schwan, Sudha Yalamanchili, and many PhD students

The Scalable HeterOgeneous Computing (SHOC) Benchmark Suite

- Benchmark suite with a focus on scientific computing workloads, including common kernels like SGEMM, FFT, Stencils
- Parallelized with MPI, with support for multi-GPU and cluster scale comparisons
- Implemented in CUDA and OpenCL for a 1:1 performance comparison
- Includes stability tests

- Level 0
  - **BusSpeedDownload**: measures bandwidth of transferring data across the PCIe bus to a device.
  - **BusSpeedReadback**: measures bandwidth of reading data back from a device.
  - **DeviceMemory**: measures bandwidth of memory accesses to various types of device memory including global, local, and image memories.
  - **KernelCompile**: measures compile time for several OpenCL kernels, which range in complexity
  - **PeakFlops**: measures maximum achievable floating point performance using a combination of auto-generated and hand coded kernels.
  - **QueueDelay**: measures the overhead of using the OpenCL command queue.
- Level 1
  - **FFT**: forward and reverse 1D FFT.
  - **MD**: computation of the Lennard-Jones potential from molecular dynamics, a specific case of the nbody problem.
  - **Reduction**: reduction operation on an array of single precision floating point values.
  - **SGEMM**: single-precision matrix-matrix multiply.
  - **Scan**: scan (also known as parallel prefix sum) on an array of single precision floating point values.
  - **Sort**: sorts an array of key-value pairs using a radix sort algorithm
  - **Stencil2D**: a 9-point stencil operation applied to a 2D data set. In the MPI version, data is distributed across MPI processes organized in a 2D Cartesian topology, with periodic halo exchanges.
  - **Triad**: STREAM Triad operations, implemented in OpenCL.

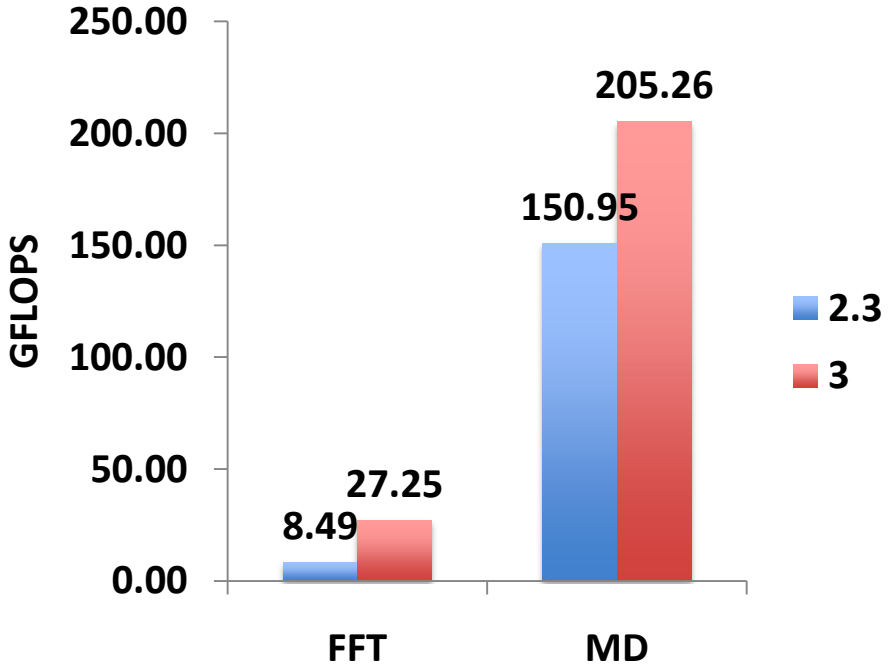
A. Danalis, G. Marin, C. McCurdy, J. Meredith, P.C. Roth, K. Spafford, V. Tipparaju, and J.S. Vetter, "The Scalable HeterOgeneous Computing (SHOC) Benchmark Suite," in Third Workshop on General-Purpose Computation on Graphics Processors (GPGPU 2010)`. Pittsburgh, 2010

Paper also includes energy and CUDA comparisons.

Beta software available at http://ft.ornl.gov/doku/shoc/start

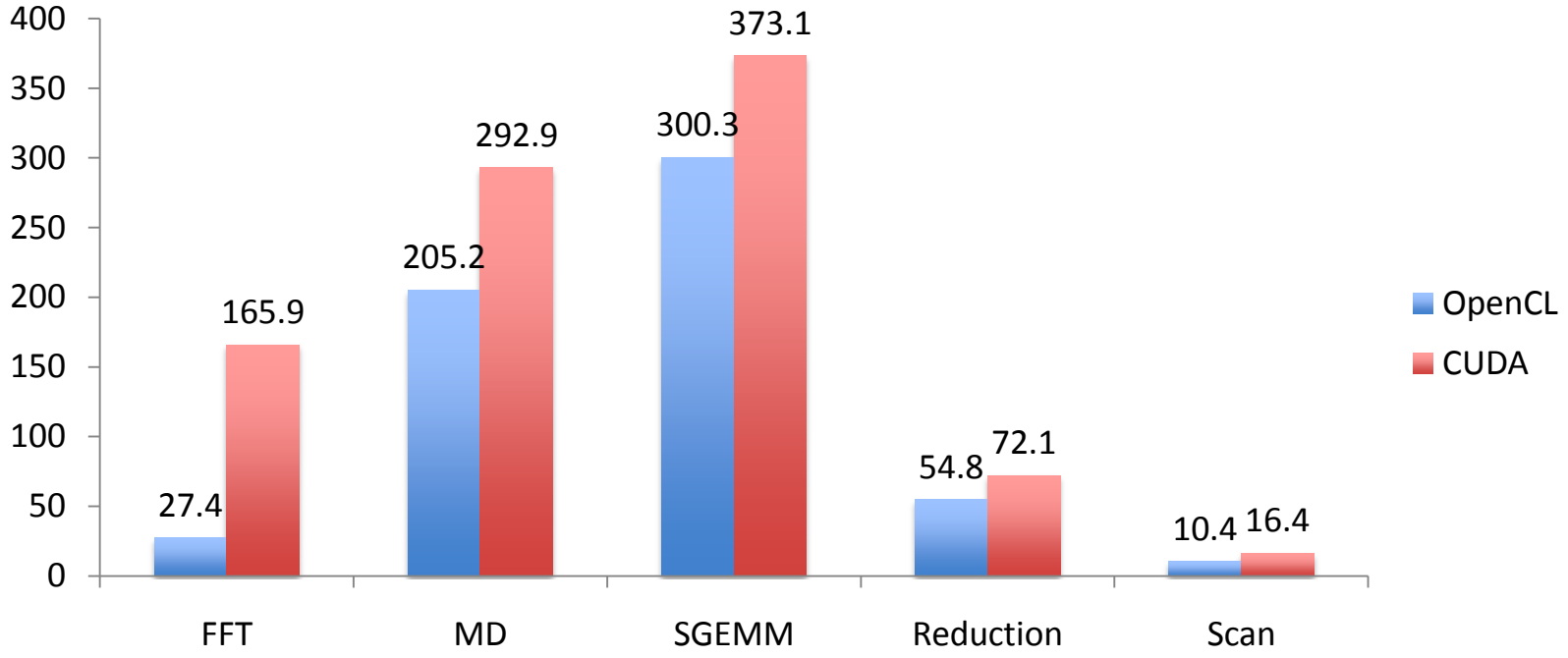
Georgia Tech NICS THE UNIVERSITY of TENNESSEE OAK RIDGE National Laboratory NVIDIA hp NSF

SHOC Example #1: Longitudinal OpenCL Performance



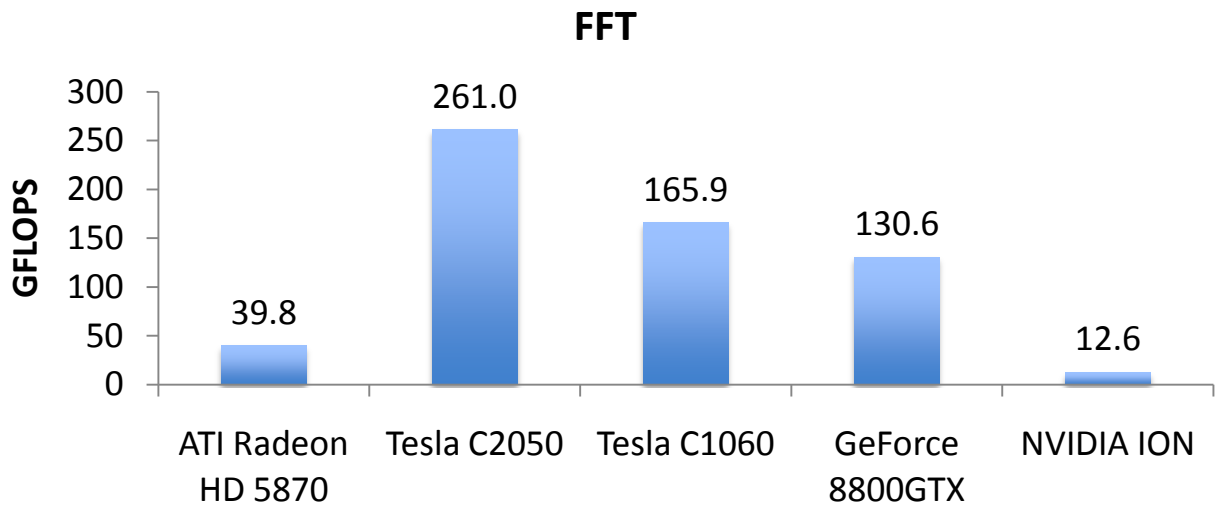
- Single precision, Tesla C1060 GPU
- Comparing NVIDIA OpenCL implementation from 2.3 and 3.0 GPU Computing SDK

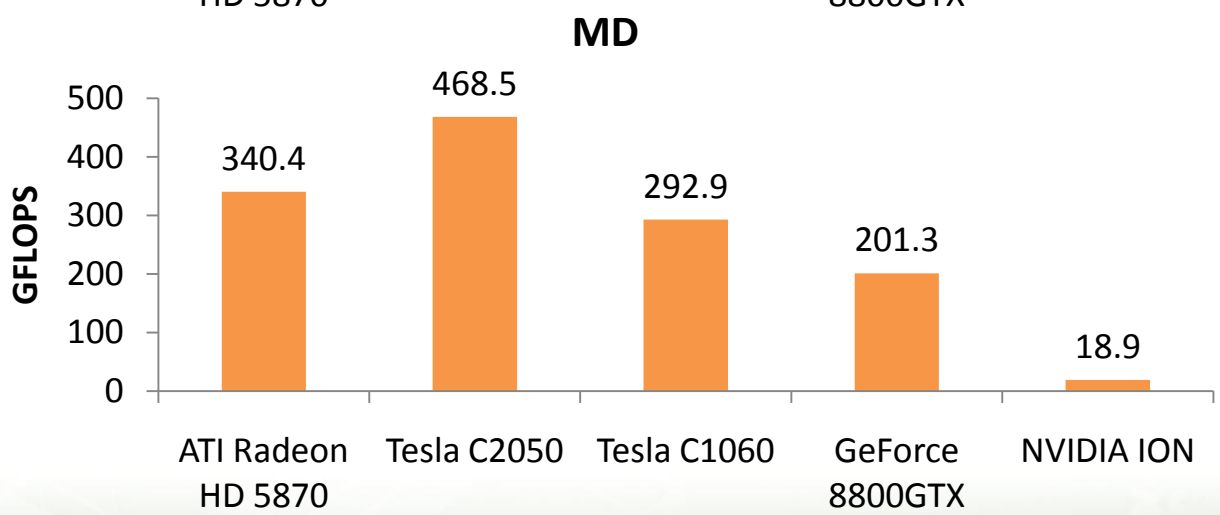
SHOC Example #2: Compare OpenCL and CUDA



- OpenCL improving, but still trailing CUDA
- Tesla C1060, Single Precision, CUDA and OpenCL 3.0
- FFT/MD/SGEMM – GFLOPS, Reduction/Scan – GB/s

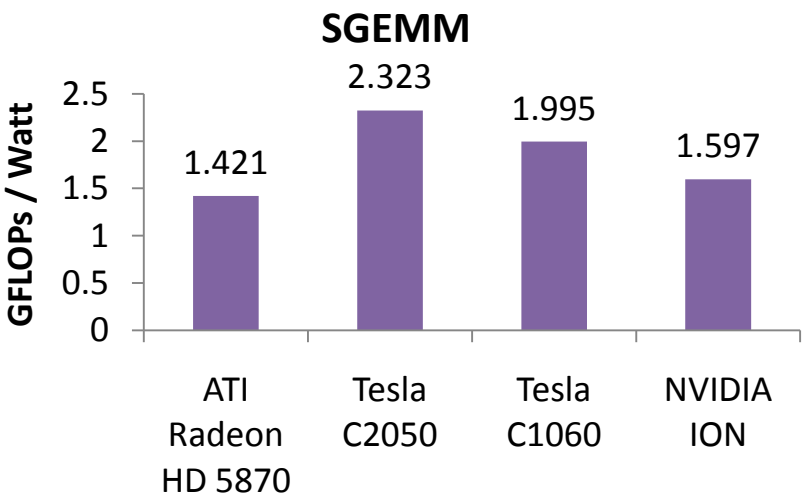
SHOC Example #3: Compare Different GPUs

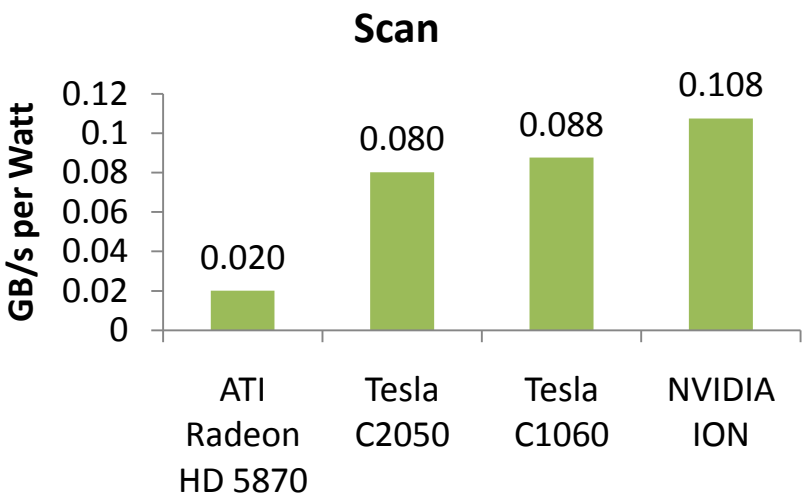


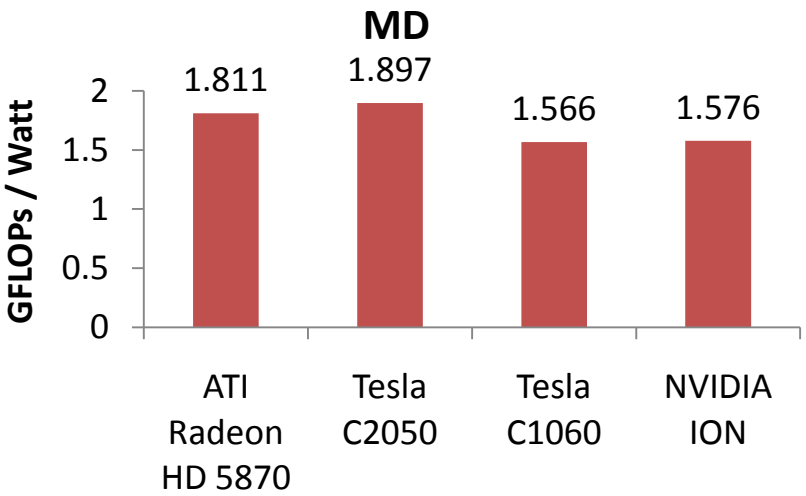


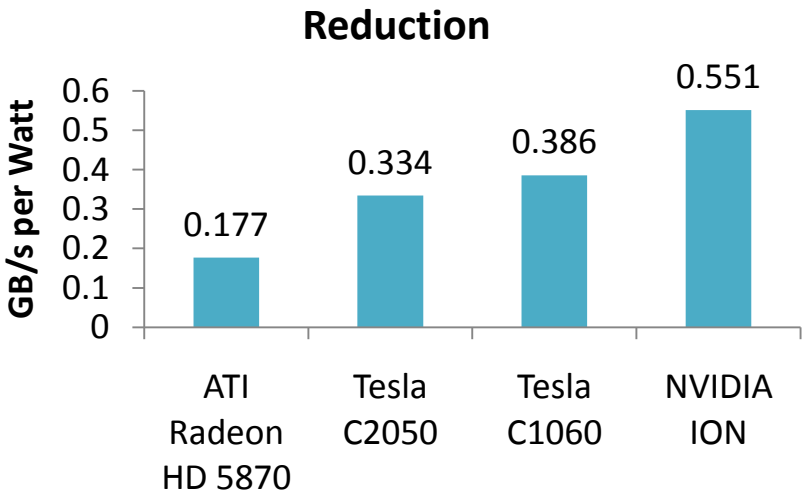
- Single Precision
- ECC On (for Tesla C2050)
- Radeon HD 5870: AMD OpenCL v2.1
- Tesla C2050 CUDA 3.1b
- Others CUDA 3.0

SHOC Example #4 – Energy Efficiency









- Single precision, calculated using vendor's TDP – Ion very efficient for bandwidth bound problems

SHOC 0.9 Beta Release

- Source code available: http://ft.ornl.gov/doku/shoc/downloads
- New Application-Level Benchmark **S3D** – combustion kernel
- Numerous bug fixes
- Coming Soon:
  - Sparse Matrix-Vector Multiply
  - More Double Precision Tests

Recap

- Predictive scientific simulation is important for scientific discovery
  - Advancing science, informing policymakers
- HPC systems have been highly successful
  - Decades of improvement
- The HPC community has several (new) constraints
  - Power, Facilities, Cost
- Heterogeneous computing with GPUs offers some opportunities and challenges
  - High performance; good performance per watt
  - Programmability; limited applicability
- Keeneland - Newly awarded NSF partnership will provide heterogeneous supercomputing for open science

We are hiring Research Scientisits, Postdocs, Students

Develop software for advanced heterogeneous architectures
- Scientific libraries
- Software tools
- Runtime systems

Assess new architectures
Identify and refactor promising applications for Keeneland platforms

http://keeneland.gatech.edu
http://ft.ornl.gov

Thank You!

Thanks to contributors, sponsors

- Many collaborators across apps teams, academia, labs, and industry
- DOE, NSF, ORNL, DOD, DARPA

More information

- http://ft.ornl.gov
- vetter@computer.org
- Publications: http://ft.ornl.gov/pubs

http://keeneland.gatech.edu

http://www.cse.gatech.edu
http://www.cercs.gatech.edu
http://icl.cs.utk.edu
http://www.nics.tennessee.edu/
http://ft.ornl.gov

http://nsf.gov/dir/index.jsp?org=OCI