RYKEN SYMPOSIUM
Feb 16, 2011

Sailing Towards Extreme Levels of Computing

Osni Marques
Lawrence Berkeley National Laboratory
OAMarques@lbl.gov

Outline

- Current HPC Scenario
- DOE's Office of Science (SC) Advanced Scientific Computing Research (ASCR) Program
  - Facilities (@ ANL, LBNL, ORNL)
  - Research
  - Plans
- Closing Remarks

36th top500 (the top 10)

www.top500.org

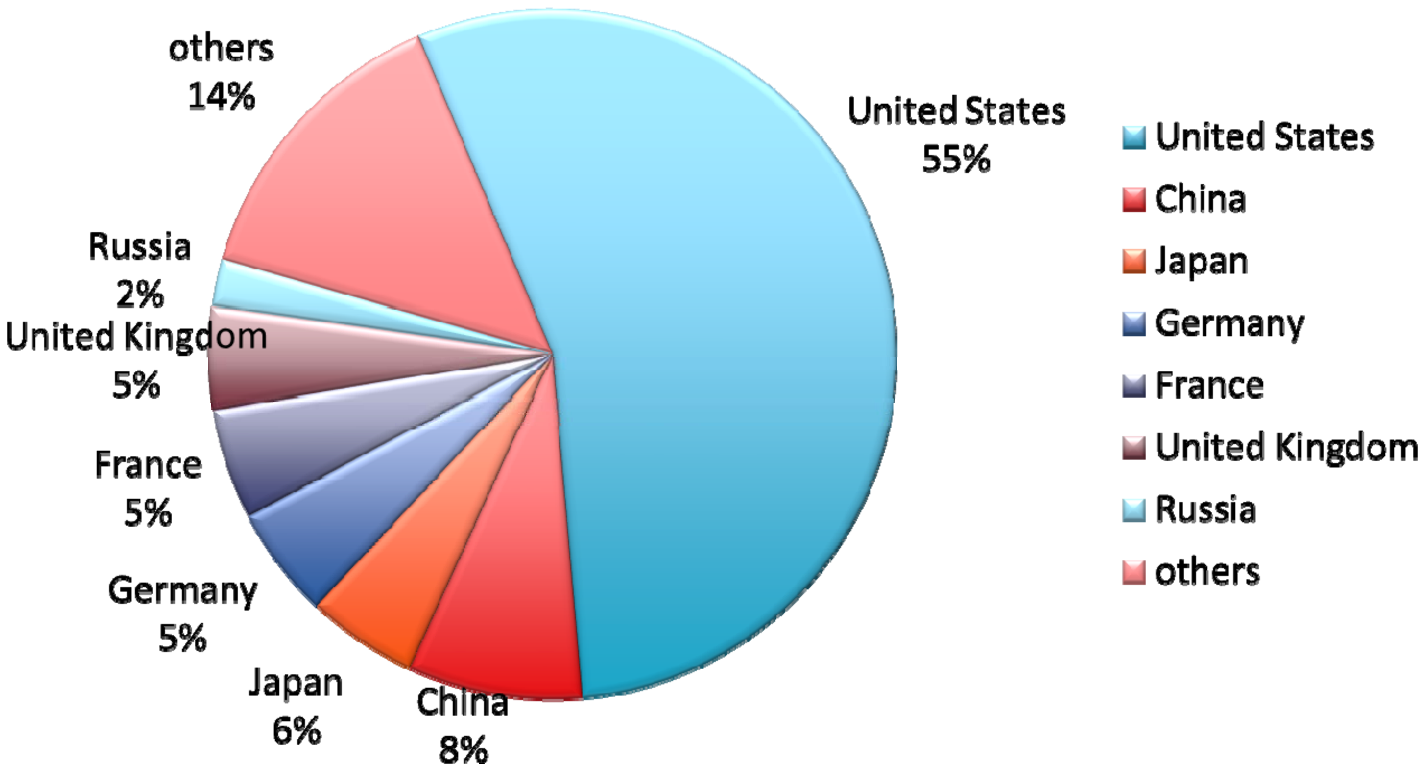
Rank	Site	Manufacturer	Computer	Country	Cores	Rmax [Tflops]	Power [MW]
1	National SuperComputer Center in Tianjin	NUDT	Tianhe-1A NUDT TH MPP, Xeon 6C, NVidia, FT-1000 8C	China	186,368	2,566	4.04
2	Oak Ridge National Laboratory	Cray	Jaguar Cray XT5, HC 2.6 GHz	USA	224,162	1,759	6.95
3	National Supercomputing Centre in Shenzhen	Dawning	Nebulae TC3600 Blade, Intel X5650, NVidia Tesla C2050 GPU	China	120,640	1,271	2.58
4	GSIC, Tokyo Institute of Technology	NEC/HP	TSUBAME-2 HP ProLiant, Xeon 6C, NVidia, Linux/Windows	Japan	73,278	1,192	1.40
5	DOE/SC/ LBNL/NERSC	Cray	Hopper Cray XE6, 6C 2.1 GHz	USA	153,408	1.054	2.91
6	Commissariat a l'Energie Atomique (CEA)	Bull	Tera 100 Bull bullx super-node S6010/S6030	France	138.368	1,050	4.59
7	DOE/NNSA/LANL	IBM	Roadrunner BladeCenter QS22/LS21	USA	122,400	1,042	2.34
8	University of Tennessee	Cray	Kraken Cray XT5 HC 2.36GHz	USA	98,928	831.7	3.09
9	Forschungszentrum Juelich (FZJ)	IBM	Jugene Blue Gene/P Solution	Germany	294,912	825.5	2.26
10	DOE/NNSA/ LANL/SNL	Cray	Cielo Cray XE6, 6C 2.4 GHz	USA	107,152	816.6	2.95

"2011 HPCwire People to Watch"

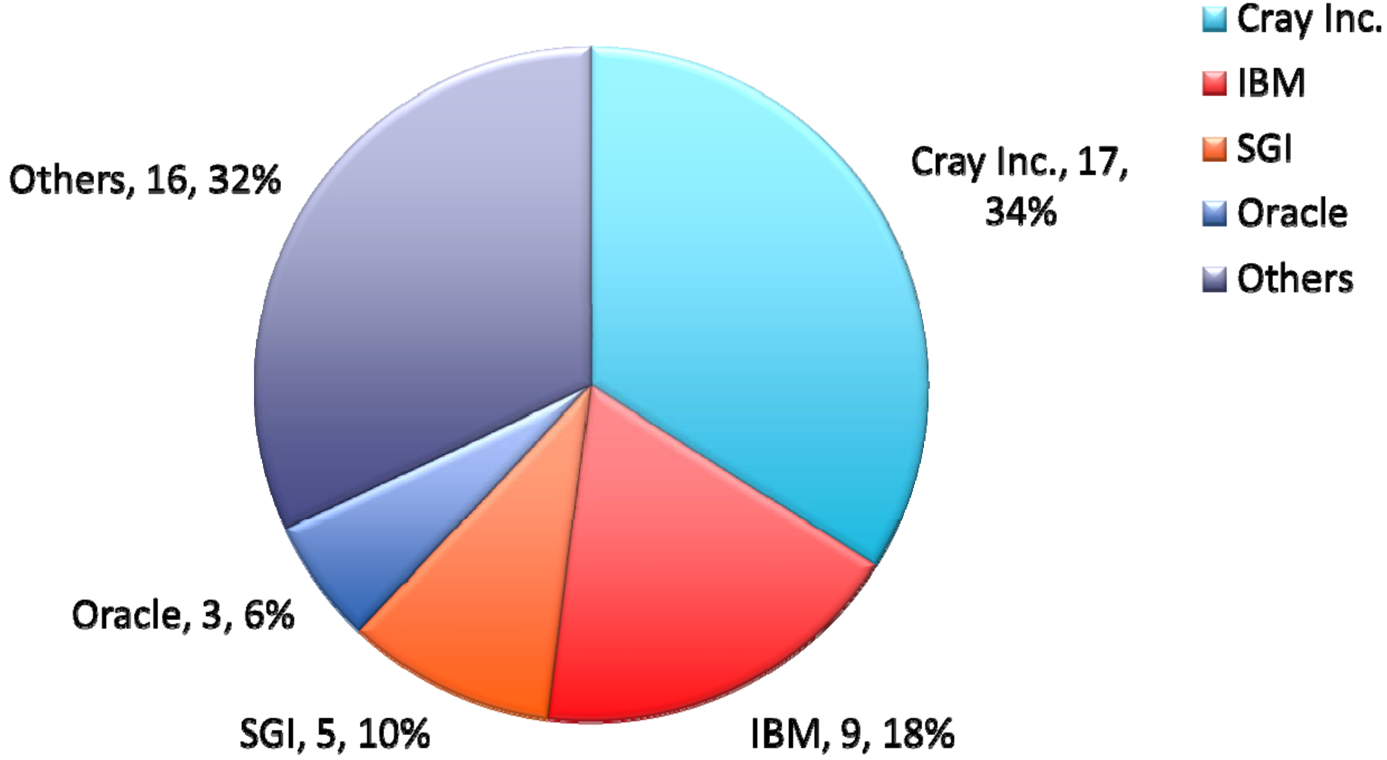
Office of Science
U.S. DEPARTMENT OF ENERGY

BERKELEY LAB

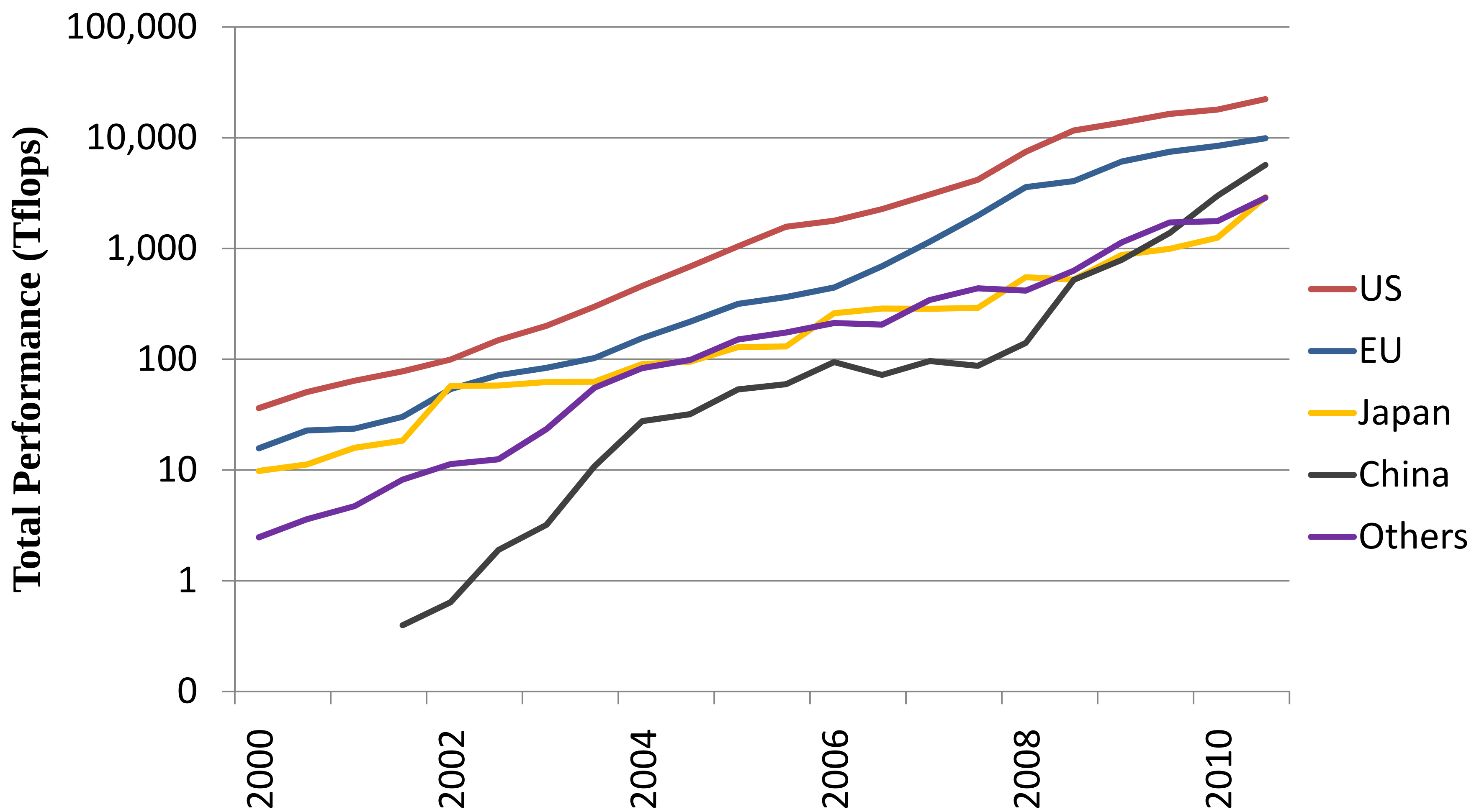
Countries / System Share



Vendors (top 50) / System Share



HPC Resources per Country



Office of Science
U.S. DEPARTMENT OF ENERGY
BERKELEY LAB

HPC in China

THE INTERNATIONAL JOURNAL of HIGH PERFORMANCE COMPUTING APPLICATIONS

The International Journal of High Performance Computing Applications
000(00) 1–46

DOI: 10.1177/1094342010381800
hpc.sagepub.com
SAGE

High-performance Computing in China: Research and Applications

Ninghui Sun[1], David Kahaner[2] and Debbie Chen[3]

Abstract
In this report we review the history of high-performance computing (HPC) system development and applications in China and describe the current status of major government programs, HPC centers and facilities, major research institutions, important HPC application fields, and domestic vendor activities. A comparison between China and developed nations, such as the United States, European countries, and Japan, in terms of system development level, application level, and long-term planning, is also briefly provided. In addition, the authors also include a few predictions about future directions for China's HPC technology and applications.

Keywords
High performance computing, research and applications, China

Table of Contents

[1]Institute of Computer Technology, Chinese Academy of Science, China. Email: snh@ncic.ac.cn
[2]Asian Technology Information Program (ATIP), Japan. Email: Kahaner@atip.or.jp
[3]Asian Technology Information Program (ATIP), China. Email: dchen@atip.org.cn

Corresponding author:
Ninghui Sun, Institute of Computer Technology, Chinese Academy of Science, China
E-mail: snh@ncic.ac.cn

- Scientific computing, oil & gas, meteorology, engineering, internet applications (e.g. gaming), DNA sequencing, stock-index computing, massive data analysis, public opinion analysis
- 2011-2015
  - many petaflops systems
  - at least one 50-100 petaflops system
  - budget 4B RMB (~ 600 USD)
- 2016-2020
  - 1-10 exaflops system
  - budget ?
- HPC training of 1 M people

Source: Xue-bin Chi, Supercomputing Center, CAS (Tokyo, Dec 2010).

Office of Science
U.S. DEPARTMENT OF ENERGY

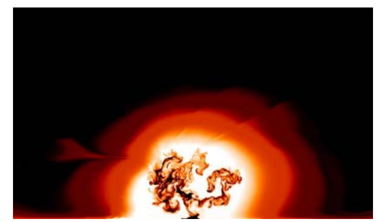
BERKELEY LAB

ASCR Computing Facilities and Research Areas

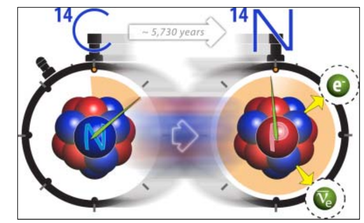
ORNL National Center for Computational Sciences

> *"At the California Institute of Technology, they're developing a way to turn sunlight and water into fuel for our cars. At Oak Ridge National Laboratory, they're using supercomputers to get a lot more power out of our nuclear facilities. With more research and incentives, we can break our dependence on oil with biofuels, and become the first country to have a million electric vehicles on the road by 2015.* (Applause.)" President Obama's State of the Union Address, 01/25/2011.

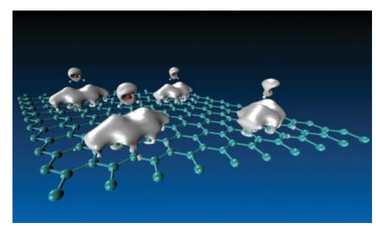
- XT5 partition: 18,688 dual hex-core AMD Opteron (Istanbul proc. running at 2.6GHz, 16GB of DDR2-800 memory) compute nodes; SeaStar 2+ router; 2.3 PF peak
- XT4 partition: 7,832 quad-core AMD Opteron (Budapest proc. running at 2.1 GHz, 8 GB of DDR2-800 memory) compute nodes, SeaStar2 router; 263 TF peak
- Several smaller systems intended for development or specialized purposes
- Kraken (1 PF peak NSF system), Gaea (1.03 PF peak NOAA system)
- OLCF-3 Project ("Titan"): design and cooling "similar" to Jaguar, accelerated node design using GPUs, 20 PF peak (Source: Buddy Bland's talk @ SC10. Also NSF Keeneland @ GA Tech, HP + NVIDIA Tesla)



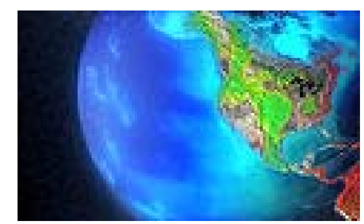
type Ia supernova explosions



carbon-14 studies



carbon–water union



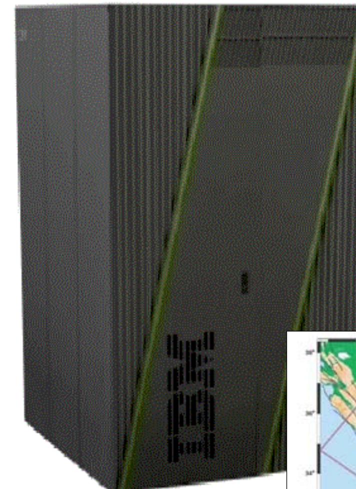
global climate change

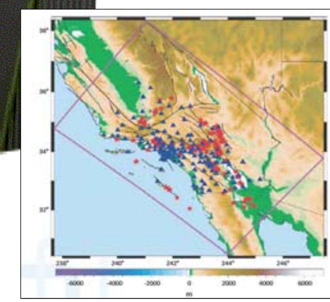
Office of Science
U.S. DEPARTMENT OF ENERGY

BERKELEY LAB

ANL Leadership Computing Facility

- Intrepid (Blue Gene/P): 40,960 quad-core; 557 TF peak
- several smaller systems intended for development or specialized purposes
- Mira (IBM Blue Gene/Q): 750K plus cores; 10 PF peak (Source: HPCwire Feb 08)

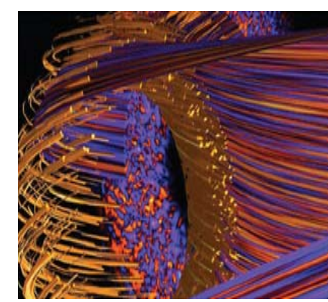




earthquake simulations

www.graph500.org (data intensive applications, "traversed edges / s")

Rank	Machine	Owner	Problem Size	TEPS	Implem.
1	Intrepid (IBM BlueGene/P, 8192 of 40960 nodes / 32k of 163840 cores)	ANL	Scale 36 (Medium)	6.6 GE/s	Optimized
2	Franklin (Cray XT4, 500 of 9544 nodes)	NERSC	Scale 32 (Small)	5.22 GE/s	Optimized
3	cougarxmt (128 node Cray XMT)	PNNL	Scale 29 (Mini)	1.22 GE/s	Optimized
4	graphstorm (128 node Cray XMT)	SNL	Scale 29 (Mini)	1.17 GE/s	Optimized
5	Endeavor (256 node, 512 core Westmere X5670 2.93, IB network)	Intel	Scale 29 (Mini)	533 ME/s	Reference
6	Erdos (64 node Cray XMT)	ORNL	Scale 29 (Mini)	50.5 ME/s	Reference
7	Red Sky (Nehalem X5570 @2.93 GHz, IB Torus, 512 processors)	SNL	Scale 28 (Toy++)	477.5 ME/s	Reference
8	Jaguar (Cray XT5-HE, 512 node subset)	ORNL	Scale 27 (Toy+)	800 ME/s	Reference
9	Endeavor (128 node, 256 core Westmere X5670 2.93, IB network)	Intel	Scale 26 (Toy)	615.8 ME/s	Reference



plasma microturbulence

LBNL National Energy Research Scientific Center

System Name	System Type	CPU		Computational Pool					Node Interconnect	Avg. Power
		Type	Speed	Nodes	SMP Size	Total Cores	Aggregate Memory	Avg. Memory/ core		
Hopper II	Cray XE6	Opteron	2.1 GHz	6,392	24	153,408	216.8 TB	1.33 GB	Gemini	2.9 MW
Hopper †	Cray XT5	Opteron	2.4 GHz	664	8	5,312	10.6 TB	2 GB	SeaStar	
Franklin	Cray XT4	Opteron	2.3 GHz	9,572	4	38,288	78 TB	2 GB	SeaStar	1,600 kW
Carver	IBM iDataPlex	Intel Nehalem	2.67 GHz	400	8	3200	9.6 TB	3 GB	QDR InfiniBand	125 kW
PDSF*	Linux Cluster	AMD/Intel	2+ GHz	~230	2.4	~1000	2.2 TB	2 GB	Ethernet	95 kW
Euclid	Sun Sunfire	Opteron	2.6 GHz	1	48	48	512 GB	10.67 GB	QDR InfiniBand	TBD kW

* hosted by NERSC and dedicated to the High Energy Physics and Nuclear Science communities

- Magellan (NERSC Cloud Testbed) : 560 nodes, 2 quad-core Intel Nehalem 2.67 GHz processors per node, 8 cores per node (4,480 total cores), 24 GB DDR3 1333 MHz memory per node
- Dirac (experimental GPU cluster): 48 nodes with attached NVIDIA Tesla GPUs)

Office of Science
U.S. DEPARTMENT OF ENERGY

BERKELEY LAB

Computing Facilities Facts

NERSC

(flagship facility)

- 1000+ users, 100+ projects
- Allocations
  - 80% ASCR Program Manager control
  - 10% ASCR Leadership Computing Challenge
  - 10% reserves
- All of DOE SC
- Machines procured for application performance

ANL and ORNL

(leadership facilities)

- 100+ users, 10+ projects
- Allocations
  - 60% ANL/ORNL managed INCITE process
  - 30% ASCR Leadership Computing Challenge
  - 10% reserves
- Large scale science (and not only DOE)
- Machines procured for peak performance

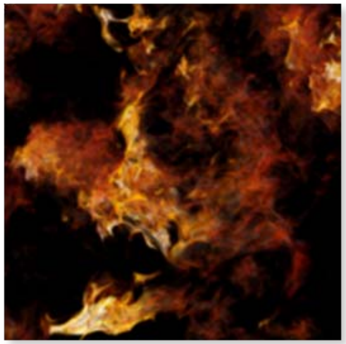
Source: Horst Simon

Office of Science
U.S. DEPARTMENT OF ENERGY

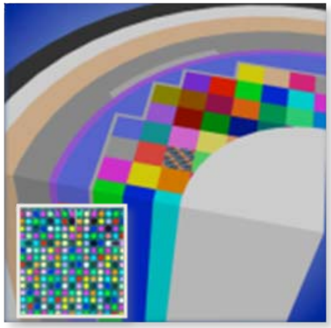
BERKELEY LAB

Leadership Computing: Scientific Progress at the Petascale

All known sustained petascale science applications to date have been run on DOE systems

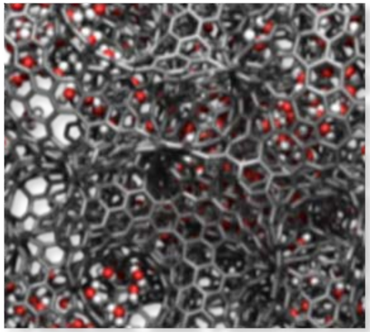


Turbulence: understanding the statistical geometry of turbulent dispersion of pollutants in the environment

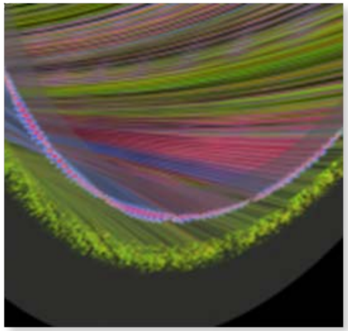


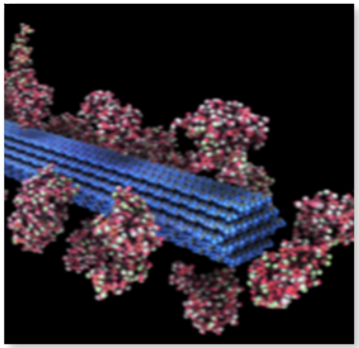
Nuclear Energy: high-fidelity predictive simulation tools for the design of next-generation nuclear reactors to safely increase operating margins

Energy Storage: understanding the storage and flow of energy in next-generation nanostructured carbon tube supercapacitors

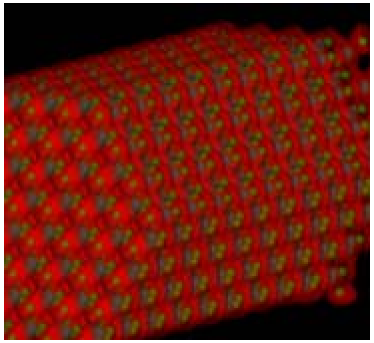


Fusion Energy: substantial progress in the understanding of anomalous electron energy loss in the National Spherical Torus Experiment (NSTX)





Biofuels: a comprehensive simulation model of lignocellulosic biomass to understand the bottleneck to sustainable and economical ethanol production



Nano Science: understanding the atomic and electronic properties of nanostructures in next-generation photovoltaic solar cell materials

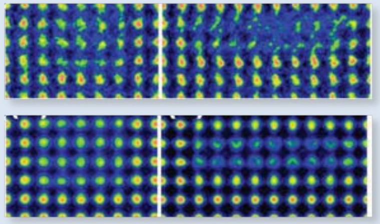
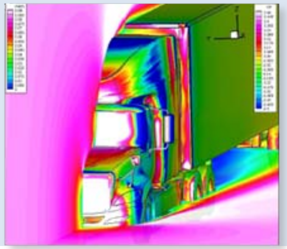
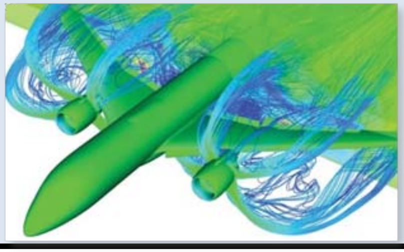
Innovative and Novel Computational Impact on Theory and Experiment

INCITE LEADERSHIP COMPUTING — U.S. DEPARTMENT OF ENERGY

- Initiated in 2004
- Provides SC leadership computing resources to a small number of computationally intensive, high-impact, research projects of large scale
  - Simulating Treatment for Parkinson's Disease (PI: Igor Tsigelny, University of California San Diego)
  - Transporting Hazard Explosives Safely (PI: Martin Berzins, University of Utah)
  - Understanding the Ultimate Battery Chemistry: Rechargeable Lithium/Air (PI: Jack Wells, ORNL)
  - Hydrogen as Alternative Fuel - Simulation (PI: John Bell, LBNL)
  - Simulating Blood Clots in the Brain to Prevent Aneurysms (PI: George Karniadakis, Brown University)
  - Simulating Large Regional Earthquakes (PI: Thomas Jordan, University of Southern California)
  - Modeling Nuclear Reactors for Electrical Power (PI: Thomas Evans, ORNL)
  - Large Eddy Simulation of Two Phase Flow Combustion in Gas Turbines|| (PI: Thierry Poinsot, CERFACS)
  - Detached-Eddy Simulations and Noise Predictions for Tandem Cylinders (PI: Philippe Spalart, The Boeing Company)
- Open to researchers in the USA and abroad, including industry
- DOE SC funding is not a requirement
- 1.7 billion hours awarded in 2010
- Peer-reviewed

Translation of Science to Industrial Solutions

U.S. DEPARTMENT OF ENERGY INCITE LEADERSHIP COMPUTING

High-efficiency thermoelectric materials enabling substantial increases in fuel efficiency	• Atomistic determination of PbTe-$AgSbTe_2$ nanocomposites and growth mechanism explains low thermal conductivity • DFT predictions of Ag atom interstitial position confirmed by high-resolution TEM • GM: Using improved insight to develop new material Actual Simulated  Nanoprecipitates in single crystal $(AgSbTe_2)\text{-}(PbTe)_{18}$
Retrofit parts for improved fuel efficiency and CO_2 emissions for Class 8 long haul trucks	• BMI Corporation: Simulations enable design of retrofit parts, reducing fuel consumption by up to 3,700 gal and CO_2 by up to 41 tons per truck per year • 10–17% improvement in fuel efficiency exceeds regulatory requirement of 5% for trucks operating in California 
Development and correlation of computational tools for transport airplanes	• Boeing: Reduced validation time to transition newer technology (CFD) from research to airplane design and development • Demonstrated and improved correlations between CFD and wind tunnel test data 

Source: T. Zacharia, ORNL

ASCR Research Areas

- **Applied Mathematics:** mathematical descriptions, models, methods and algorithms to enable scientists to accurately describe and understand the behavior of complex systems involving processes that span vastly different time and/or length scales
- **Computer Science:** underlying understanding and software to make effective use of computers at extreme scales; tools to transform extreme scale data from experiments and simulations into scientific insight
- **Integrated Network Environments:** computational and networking capabilities, enabling world-class researchers to work together and to extend the frontiers of science

- **SciDAC:** scientific computing software infrastructure to enable scientific discovery in the physical, biological, and environmental sciences at the petascale; new generation of data management and knowledge discovery tools for large data sets (obtained from scientific user and simulations)

- **CSGF Program:** education and training of science and technology professionals with advanced computer skills

Delivering the Science

Scientific Discovery and the Role of High End Computing

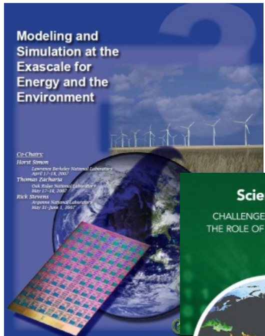
Office of Science
U.S. DEPARTMENT OF ENERGY

BERKELEY LAB

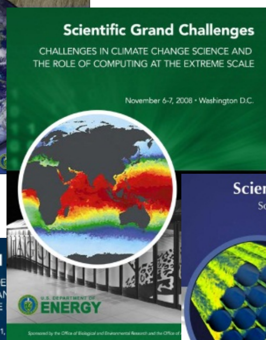
Next Levels of Computing
(Extreme Scales of Computing)

Scientific Grand Challenges: Needs and Potential Outcomes

www.sc.doe.gov/ascr/ProgramDocuments/ProgDocs.html



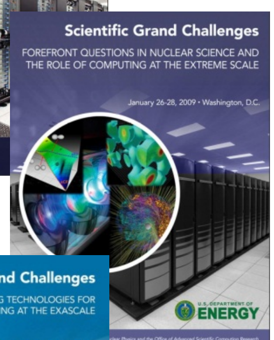
Town Hall Meetings, 2007



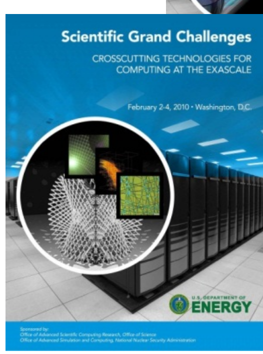
Climate Change Science, 2008

Nuclear Energy Systems, 2009

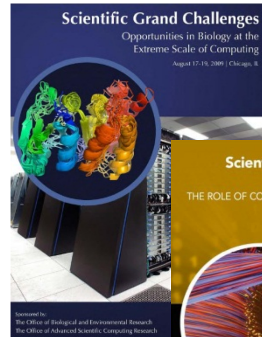
Understanding the Quantum Universe, 2008



Nuclear Science, 2009



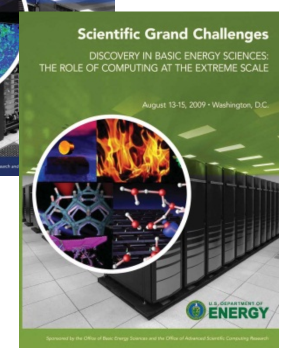
Crosscutting Technologies, 2010



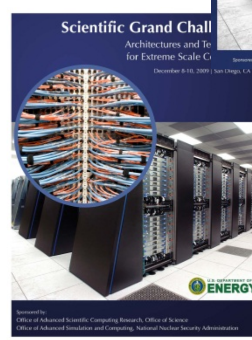
Opportunities in Biology, 2009

Fusion Energy Sciences, 2009

National Security, 2009

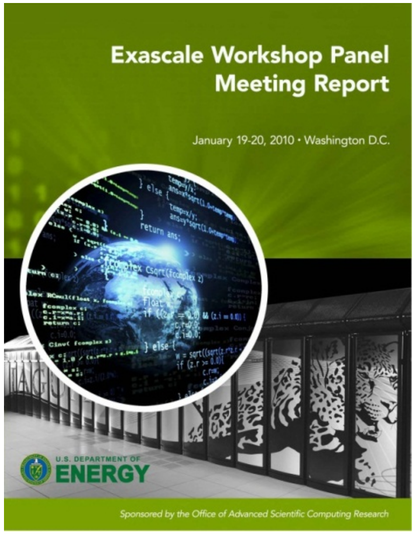


Basic Energy Sciences, 2009



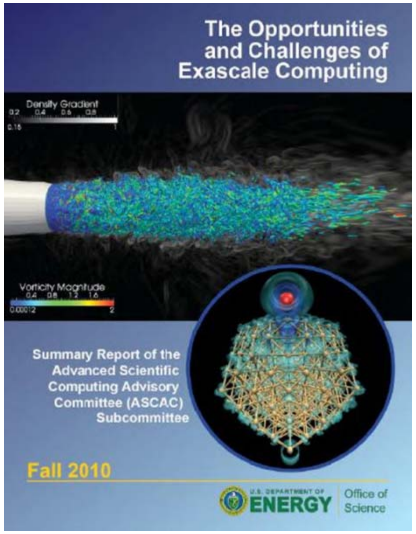
Architectures and Technology for Extreme Scale Computing, 2009

Opportunities and Challenges in General

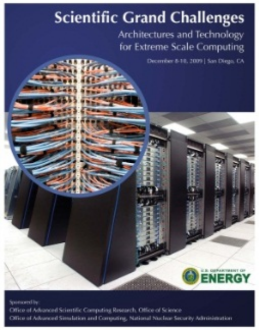


"The key finding of the Panel is that there are compelling needs for exascale computing capability to support the DOE's missions in energy, national security, fundamental sciences, and the environment … Failure to initiate an exascale program could lead to a loss of U.S. competitiveness in several critical technologies." [Trivelpiece Exascale Workshop]

"The mission and science opportunities in going to exascale are compelling … Making the transition to exascale poses numerous unavoidable scientific, algorithmic, mathematical, software, and technological challenges ... The benefits of going to exascale far outweigh the costs … The exascale initiative as described in workshop reports and expert testimony portends an integrated approach to the path forward ... ***Recommendation****: DOE should proceed expeditiously with an exascale initiative so that it continues to lead in using extreme scale computing to meet important national needs."* [Advanced Scientific Computing Advisory Committee]

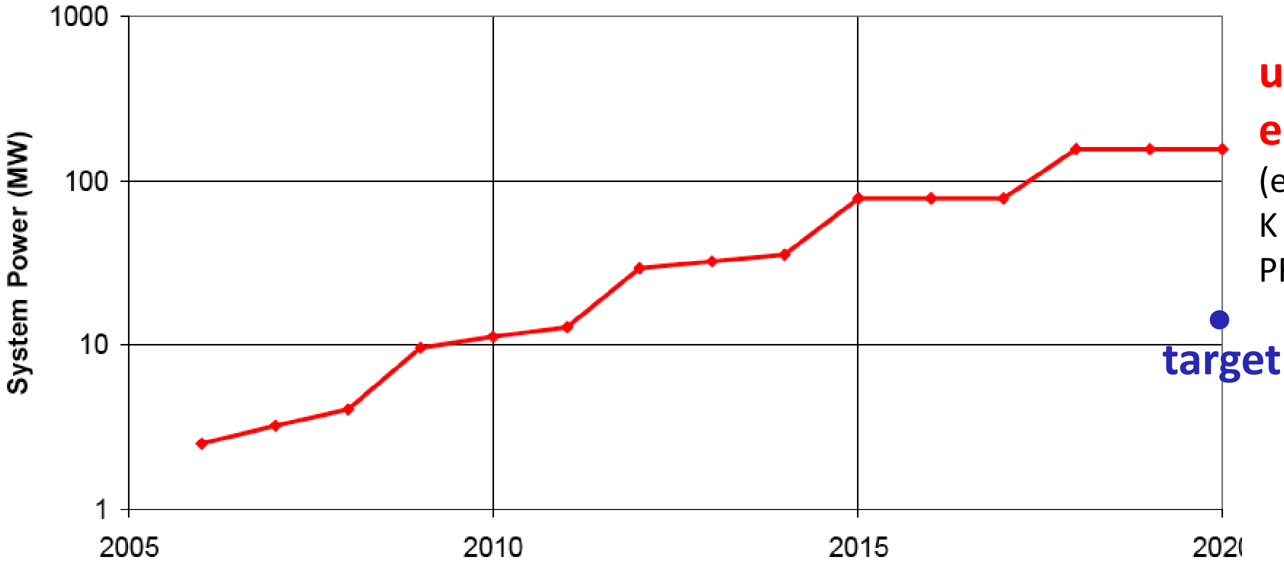


Crosscutting Technologies: *research directions*



- **Algorithm and Model Research**
  - Recast critical algorithms to reflect the impact of anticipated architecture evolution (e.g. memory and communication constraints)
  - Take advantage of architecture evolution to design new algorithms for uncertainty quantification (to establish levels of confidence in computational prediction) and discrete mathematics
  - Mathematical models and formulations that effectively exploit anticipated exascale architectures
  - Extract essential elements of critical science applications as "mini-applications" to be used in the understanding of computational requirements
  - Tools to simulate emerging architectures and performance modeling methods for use in co-design
- **Programming Models to Support Exascale Computing**
  - Programming paradigms to support "billion-way" concurrency
  - Tools and runtime systems for dynamic resource management
  - Programming models that support memory management on exascale architectures
  - Scalable approaches for I/O on exascale architectures
  - Interoperability tools to support the incremental transition of critical legacy science application codes to a exascale
  - Develop programming model support for latency management, fault tolerance and resilience
  - Develop integrated tools to support application performance and correctness
- **Research and Development for System Software at the Exascale**
  - System software tools to support node-level parallelism
  - System support for dynamic resource allocation
  - System software support for memory access (global address space, memory hierarchy, and reconfigurable local memory).
  - Performance and resource measurement and analysis tools for exascale
  - System tools to support fault management and system resilience
  - Capabilities to address the exascale I/O challenge

Power becomes a major concern …



usual scaling → revised estimate: 500 plus MW

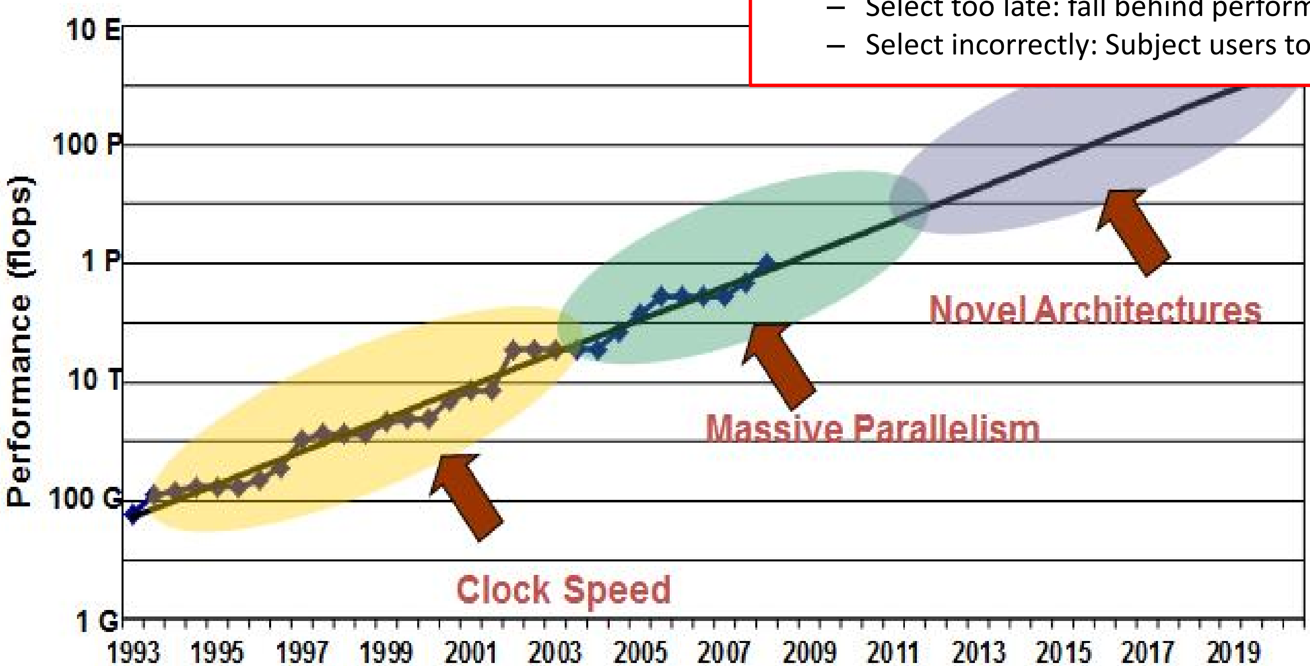
(e.g. Blue Waters @ NCSA: 300 K eight-core IBM Power 7; 10 PF peak, 15 MW)

> $1M per megawatt per year

Source: Peter Kogge, DARPA Exascale Study

Technology Paths

- Leading technology paths (swim lanes)
  - Multicore: maintain complex cores, and replicate (x86 and Power7, Blue Waters, NGSC)
  - Manycore/embedded: use many simpler, low power cores from embedded (BlueGene, Dawning)
  - GPU/Accelerator: use highly specialized processors from gaming/graphics market space (NVidia Fermi, Cell, Intel Knights Corner/Larrabee)
- Risks in “swim lane” selection
  - Select too soon: users cannot follow
  - Select too late: fall behind performance curve
  - Select incorrectly: Subject users to multiple disruptive technology changes

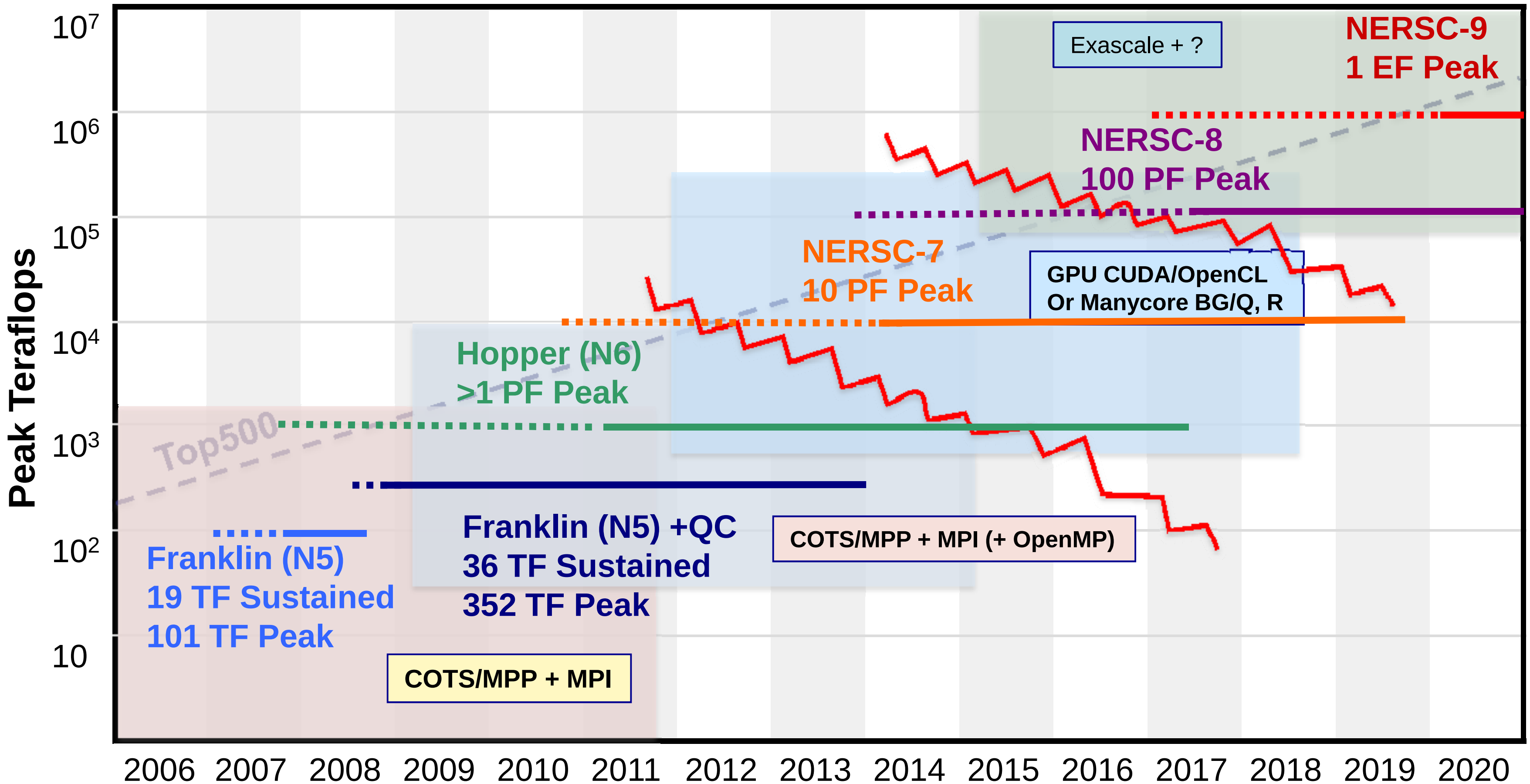


Technology Challenges

- Rapidly changing technology landscape
  - Evolutionary change between nodes (10x more explicit parallelism)
  - Revolutionary change within node (100x more parallelism, with diminished memory capacity and bandwidth)
  - Multiple technology paths (GPUs, manycore, x86/PowerX)
- The technology disruption will be pervasive (and not just at exascale)
  - Assumptions that our current software infrastructure is built upon are no longer valid
  - Applications, algorithms, system software will not work
  - As significant as migration from vector to MPP (early 90's)
- Need a new approach to ensuring continued application performance improvements (at all system scales)

Technology Transition

… and impacts to a facility like NERSC



Source: Horst Simon

Office of Science
U.S. DEPARTMENT OF ENERGY

BERKELEY LAB

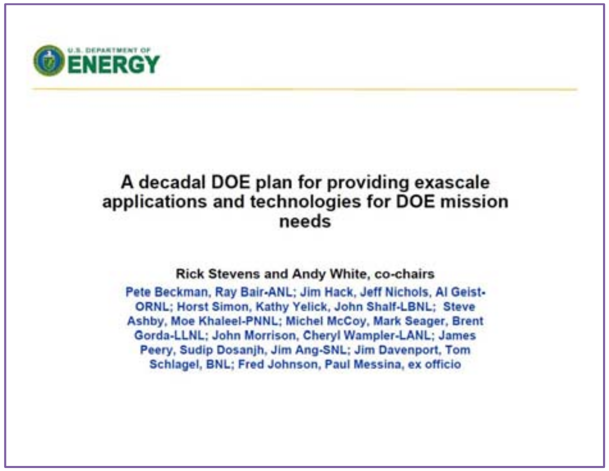
Technology Transition

… and diversity of applications and algorithms at NERSC

Science areas	Dense linear algebra	Sparse linear algebra	Spectral methods	Particle methods	Structured grids	Unstructured grids or AMR
Accelerator Science		●	●	●	●	●
Astrophysics	●	●	●	●	●	●
Chemistry	●	●	●	●		
Climate			●		●	●
Combustion					●	●
Fusion	●	●		●	●	●
Lattice Gauge		●	●	●	●	
Material Science	●		●	●	●	

A 10-year Plan for DOE

www.sc.doe.gov/ascr/ASCAC/Meetings/Presentations3-10.html


U.S. DEPARTMENT OF ENERGY

A decadal DOE plan for providing exascale applications and technologies for DOE mission needs

Rick Stevens and Andy White, co-chairs

Pete Beckman, Ray Bair-ANL; Jim Hack, Jeff Nichols, Al Geist-ORNL; Horst Simon, Kathy Yelick, John Shalf-LBNL; Steve Ashby, Moe Khaleel-PNNL; Michel McCoy, Mark Seager, Brent Gorda-LLNL; John Morrison, Cheryl Wampler-LANL; James Peery, Sudip Dosanjh, Jim Ang-SNL; Jim Davenport, Tom Schlagel, BNL; Fred Johnson, Paul Messina, ex officio

Plan discusses:

- Science and mission applications
- Systems software and programming models
- Hardware technology R&D
- Systems acquisition, deployment and operations

- 2015: co-design and co-development of hardware, software, programming models and applications requires intermediate platforms
- 2018: exascale platform deliveries; robust simulation environment and science and mission applications by 2020

Office of Science
U.S. DEPARTMENT OF ENERGY

BERKELEY LAB

Two Potential Architectures ("swim lanes")

Architectures and Technology for Extreme Scale Computing, 2009

System attributes	2010	2015		2018	
System peak	2 PF	200 PF		1 EF	
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/s	1 TB/s	0.4 TB/s	4 TB/s
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/s	1 TB/s	250 GB/s	2 TB/s
MTTI	day	O(1 day)		O(1 day)	

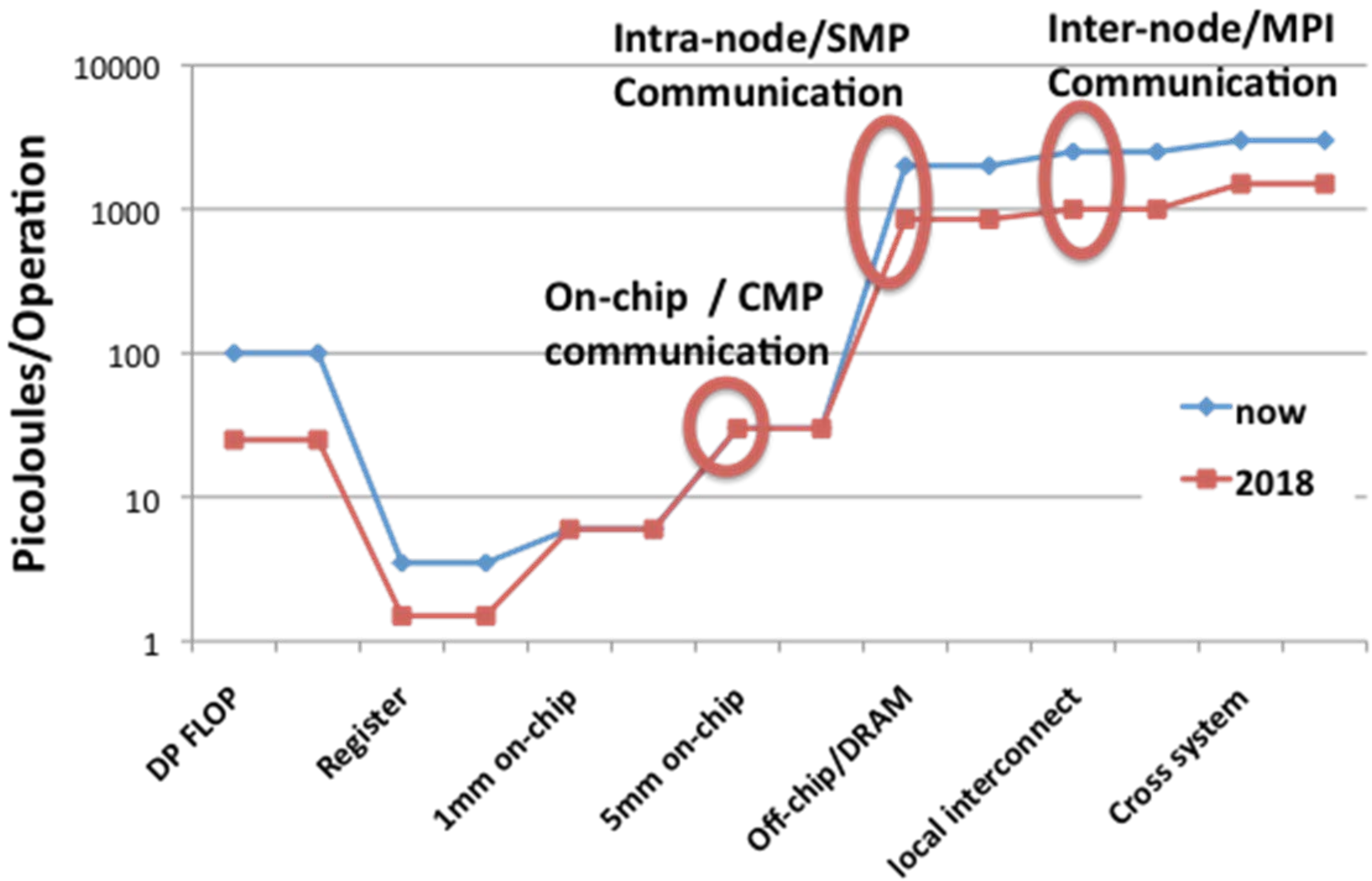
Office of Science
U.S. DEPARTMENT OF ENERGY

BERKELEY LAB

Critical Exascale Technology Investments

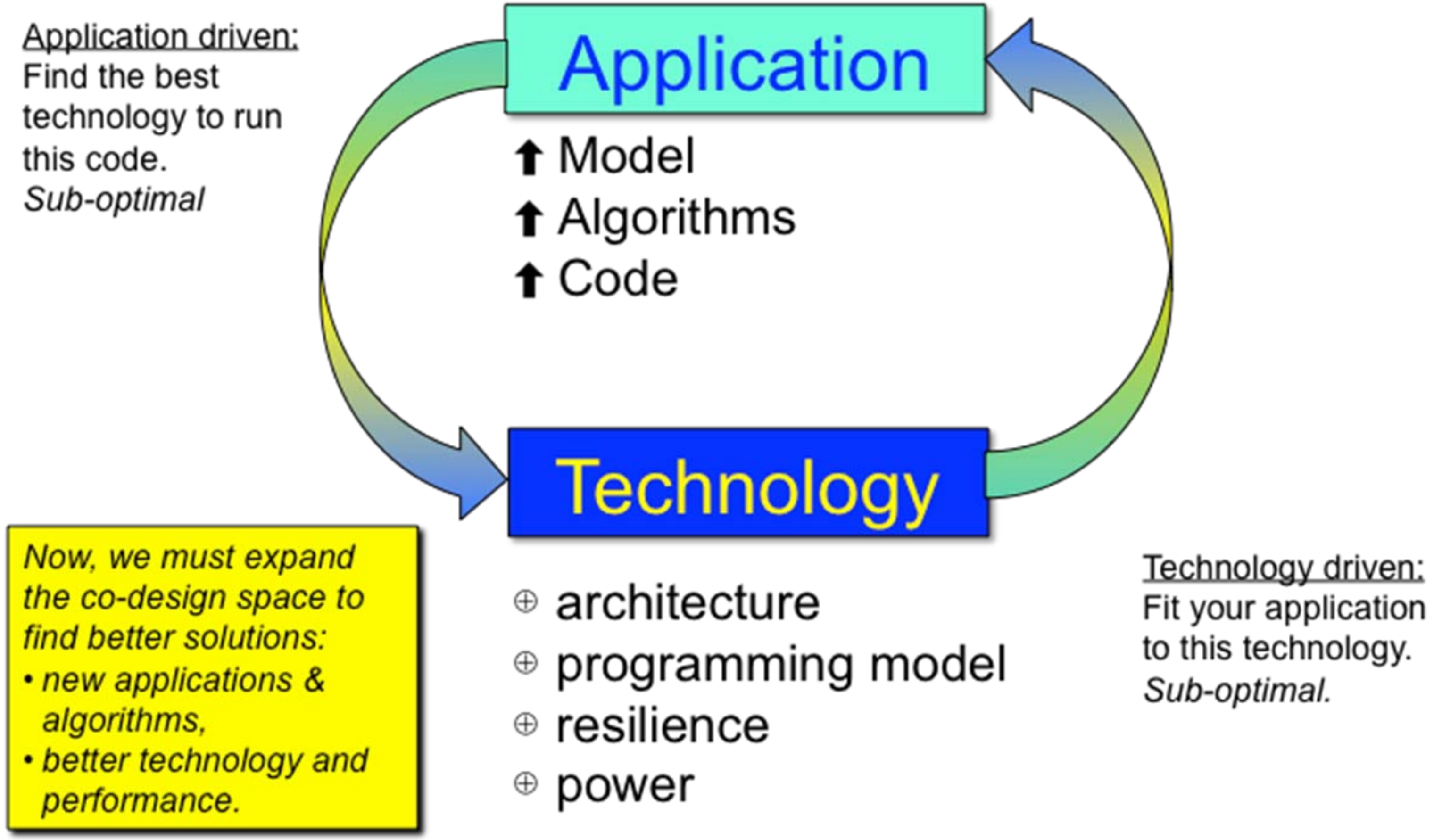
- **System power.** First class constraint on exascale system performance and effectiveness.
- **Memory.** Important component of meeting exascale power and applications goals.
- **Programming model.** Early investment in several efforts to decide in 2013 on exascale programming model, allowing exemplar applications effective access to 2015 system for both mission and science.
- **Investment in exascale processor design.** To achieve an exascale-like system in 2015.
- **Operating system.** Critical for node performance at scale and for efficient support of new programming models and run time systems.
- **Reliability and resiliency.** Critical at this scale and require applications neutral movement of the file system (for check pointing, in particular) closer to the running apps.
- **Co-design strategy.** Requires a set of hierarchical performance models and simulators and commitment from applications, software and architecture communities.

Application locality becomes critical…



Source: Exascale Computing Technology Challenges, Shalf, Dosanjh and Morrison, Proc. of VECPAR'10

Application-driven Hardware/Software Co-design



Source: Exascale Computing Technology Challenges, Shalf, Dosanjh and Morrison, Proc. of VECPAR'10

Office of Science
U.S. DEPARTMENT OF ENERGY

BERKELEY LAB

ASCR's Response to Community Reports

Source: Barbara Helland, IESP Meeting, October 18-19, 2010.

- Proposals processed in Exascale related topics:
  - Applied Mathematics: Uncertainty Quantification (90 proposals requesting ~$45M/year; 6 funded at ~$3M/yr)
  - Computer Science: Advanced Architectures (28 proposals requesting ~ $28M/year, 6 funded at ~$5M/yr)
  - Computer Science: X-Stack (55 proposals requesting ~$40M/year; 11 funded at ~$8.5M/yr)
  - Computational Partnerships: Co-Design (21 proposals requesting ~ $160M/year)
- Exascale Coordination meetings with other Federal Departments and Agencies.
- Partnership with National Nuclear Security Administration (NNSA).

Uncertainty Quantification: *topics of interest*

- Mathematical, statistical and hybrid approaches for quantifying and describing the effects and interactions of uncertainty and errors, potentially from multiple sources and with multiple representations.
- Mathematical and computational frameworks for integrating statistical and deterministic analysis.
- Mathematical theory and/or implementation of algorithms that demonstrably circumvent the "curse of dimensionality" in UQ analysis for complex system simulations.
- Mathematical theory and/or algorithms for reduced-order modeling, inference, and inverse problems.
- Scalable algorithms for numerical solutions of stochastic differential equations.
- Tractable UQ treatment (intrusive or non-intrusive) for high-concurrency architectures.
- Memory-access-efficient algorithms that match current and emerging computer architectures and allow for efficient and tractable sampling-based approaches.

Advanced Architectures: *topics of interest*

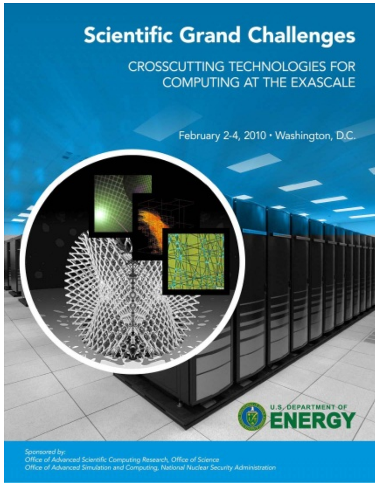
- Approaches for reducing and/or managing power requirements for high performance computing systems, including the memory and storage hierarchy.
- Approaches for reducing and/or managing heat in high performance computing systems.
- Methods for improving system resilience and managing the component failure rate, including approaches for shared information and responsibility among the OS, runtime system, and applications.
- Co-design of systems that support advanced computational science at the extreme scale.
- Scalable I/O systems, which may include alternatives to file systems.
- Approaches to information hiding that reduce the need for users to be aware of system complexity, including heterogeneous cores, the memory hierarchy, etc.

X-stack: *topics of interest*

- System software, including operating systems, runtime systems, adaptable operating and runtime systems, I/O systems, systems management/administration, resource management and means of exposing resources, and external environments.
- Fault management, both by the operating and runtime systems and by applications.
- Development environments, including programming models, frameworks, compilers, and debugging tools
- Application frameworks
- Crosscutting dimensions, including resilience, power management, performance optimization, and programmability
- Design and/or development of high-performance scientific workflow systems that incorporate data management and analysis capabilities

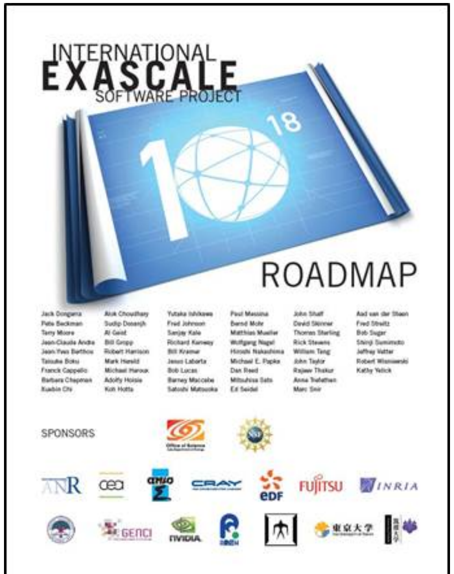
Co-design: *requirements*

- Scientific domain experts, applied mathematicians, computational scientists, computer scientists, hardware architects and software engineers.
- Critical mass of developers organized in a "code-team" who would be able to evaluate and implement multiple alternatives on an aggressive schedule to support the architecture development timeline
- Experience with application codes, i.e. an existing body of work with one or more codes that are under active development, targeting at the exascale design points of approximately 1-10 billion degree concurrency on hundreds of millions cores.
- Experience with scientific kernels and algorithms design, and optimization, uncertainty quantification, verification, and validation techniques.
- Knowledge and practical experience with advanced architectures (i.e. micro-architecture hardware design, circuit design, co-simulation, hardware and software synthesis, formal specification and verification methods, simulation modeling, hardware and system software optimization).



Closing Remarks

The International Exascale Software Project Roadmap



- Initiated by DOE and NSF

- System Software
- Development Environments
- Applications
- Cross-Cutting Dimensions

(Proposed) DOE Exascale Software Center (ESC):

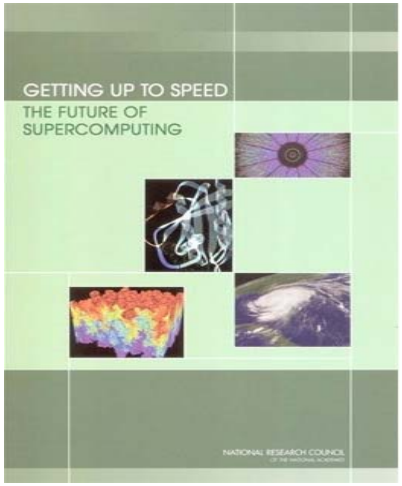
- Ensure successful deployment of coordinated exascale software stack on exascale platforms

- Deliver high quality system software for exascale platforms (~2015, ~2018)
- Identify software gaps, research and development solutions, test and support deployment
- Increase the productivity and capability and reduce the risk of exascale deployments

- Participation in co-design activities?

The Importance of HPC (1/2)

- The 2004 National Research Council Report confirmed the importance of HPC for
  - leadership in scientific research
  - economic competitiveness
  - national security
- Implementation of the report recommendations led to the state of HPC in the US today
- 2010 is very different
  - Processor speeds stagnated
  - Power constraints



2004


The Future of Computing Performance

Game Over or Next Level?

Samuel H. Fuller and Lynette I. Millett, Editors

Committee on Sustaining Growth in Computing Performance

Computer Science and Telecommunications Board

Division on Engineering and Physical Sciences

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, DC
www.nap.edu

2010

The Importance of HPC (2/2)

- China (mentioned earlier)
- Japan
  - US$ 1.5B for 2006-2012
  - 10 PF computing in 2012 (Kobe)
- Europe
  - PRACE (Partnership for Advanced Computing in Europe)
    - World-class HPC systems for world-class science
    - Support Europe in achieving global leadership in public and private research and development
    - Diversity of architectures to meet the needs of European user communities
    - Several prototypes for PF systems in 2009/2010
    - Small number of systems targeting 100 PF in 2014 and 1000 PF by 2020
    - Support and training
  - Exascale Computing Research Centre (*www.exascale-computing.eu)*
  - 1.3 PF system @ Moscow State University (to be delivered by T-platforms early this year)
- G8 Initiative in Exascale: www.nsf.gov/od/oise/g8initiative

Development of a Supercomputing Strategy in Europe, 2010


IDC
Analyze the Future

SPECIAL STUDY

D2 Interim Report: Development of a Supercomputing Strategy in Europe (SMART 2009/0055, Contract Number 2009/S99-142914)

Earl C. Joseph, Ph.D. — Steve Conway
Chris Ingle — Gabriella Cattaneo
Cyril Meunier — Nathaniel Martinez

Author(s)	IDC: Earl C. Joseph, Ph.D., Steve Conway, Chris Ingle, Gabriella Cattaneo, Cyril Meunier and Nathaniel Martinez
Deliverable	D2 Interim Report
Date of delivery	July 7, 2010
Version	2.0
Addressee officers	Bernhard FABIANEK Performance Computing Officer European Commission DG Information Society Unit F03 — office BU25 04/087 25 Avenue Beaulieu, B-1160 Bruxelles Tel: Email:
Contract ref.	Contract Nr 2009/S99-142914

Thank You !

U.S. DEPARTMENT OF ENERGY

Uncertainty comes in a variety of shapes and sizes

	Parametric	Structural	Relational
Theory & models	•Calibrated parameters in models	•Unknown effects omitted from models •Extrapolation	•Multi-scale, multi-physics effects
Algorithms	•Discretization error	•Extrapolation	•Multiple time scales in operator split algorithms
Applications code	•Convergence criteria	•Errors in apps code	•Data mapping among different components
Computation and communication	•Rounding errors	•Silent data corruption	•Race conditions among separate components of system
Operating system & environment	•ECC error rates (chip bit errors)	•System parameters set incorrectly •Chip temperature excursions	•System policy mis-match (e.g. memory management)
Observations & data assimilation	•Statistical variation in experimental data	•Unknown systematic errors in data	•Contextual mismatch of observational and computational data

Exascale Initiative Steering Committee

ENERGY

A decadal DOE plan for providing exascale applications and technologies for DOE mission needs

Rick Stevens and Andy White, co-chairs

Pete Beckman, Ray Bair-ANL; Jim Hack, Jeff Nichols, Al Geist-ORNL; Horst Simon, Kathy Yelick, John Shalf-LBNL; Steve Ashby, Moe Khaleel-PNNL; Michel McCoy, Mark Seager, Brent Gorda-LLNL; John Morrison, Cheryl Wampler-LANL; James Peery, Sudip Dosanjh, Jim Ang-SNL; Jim Davenport, Tom Schlagel, BNL; Fred Johnson, Paul Messina, ex officio

Development of a Supercomputing Strategy in Europe

Some key observations:

- Importance of HPC for scientific leadership, industrial competitiveness
- Analysis of HPC programs around the world (e.g. SciDAC, RIKEN, INCITE)
- Suggestions about the future of HPC
- Actions needed to improve the EU position
- Consequences if the EU does not take additional steps
- What should be avoided (e.g. development of its own hardware)
- EU should pursue HPC in areas where it can excel