

第9回先進スーパーコンピューティング環境研究会（ASE 研究会）実施報告

東京大学情報基盤センター 特任准教授 片桐孝洋

2011年2月17日（木）14時00分～17時30分、東京大学情報基盤センター3階大会議室にて、第9回先進スーパーコンピューティング環境研究会（ASE 研究会）が開催されました。

国内の大学・研究機関および企業からの参加者が19名あり、活発な議論がなされました。招待講演として、米国ローレンス・バークレー国立研究所から Osni Marques 博士をお呼びし、エクサスケールコンピューティングの現状に関する講演を行いました。また情報基盤センターの教員より、最新の研究成果に関する講演を行いました。

Osni Marques 博士の講演は、「Extreme Levels of Computing: Challenges and Opportunities」と題し、現在の高性能計算についてエクサスケールのスパコンへの適用シナリオと、米国エネルギー省（DOE）の科学部門（Office of Science (SC)）における、先進科学計算研究プログラム（Advanced Scientific Computing Research (ASCR) Program）に関する、人員、研究、計画に関するものでした。現在の高性能計算のシナリオは、スパコンランキングで著名な Top500 の第36回ランキングで1位と3位を受賞したスパコンを有する中国における情勢などをご紹介されました。また ASCR プログラムでは、オークリッジ国立研究所、アルゴンヌ国立研究所、ローレンスバークレー国立研究所のスパコンとその運用状況についてご紹介されました。特に、Osni 博士が所属するローレンス・バークレー国立研究所では、スパコンのユーザ数が1000人以上、かつ、100件以上のプロジェクトを管理・運営されています。その他の米国の主要な国立研究所に対し、その数が1桁も大きい規模で研究と運営をされていることが、とても印象的でした。

以上の講演内容について、本原稿の後に当日の発表資料を掲載します。ご興味のある方はご覧ください。

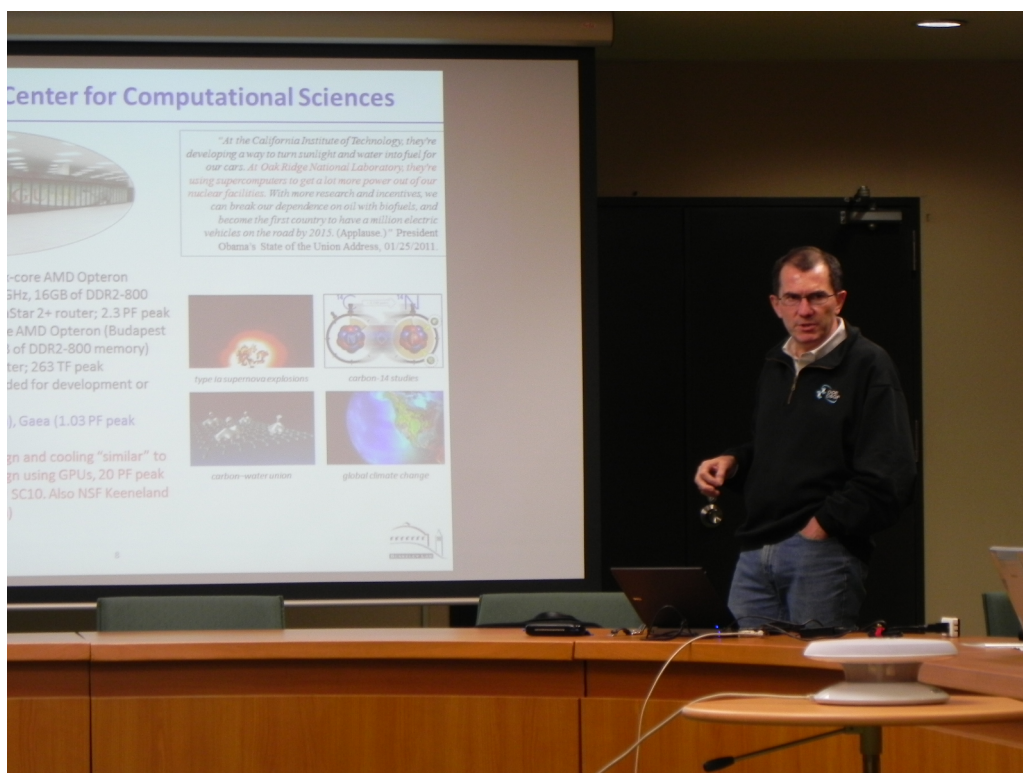


図1 Osni 博士による講演の様子

当日のプログラムを以下に載せます。

●プログラム

○【招待講演】 14:00 - 15:00

Speaker

Dr. Osni Marques (Lawrence Berkeley National Laboratory)

Title

Extreme Levels of Computing: Challenges and Opportunities

Abstract

Exascale computing capabilities are expected to become a reality by the end of this decade. The pursuit of these capabilities is greatly motivated by the needs of a range of applications: exascale offers an unprecedented opportunity for transformative discovery, prediction and design.

Nonetheless, the path to such levels of computing presents formidable technological challenges. Expected changes in hardware will impact the way exascale computers are not only programmed but also built. Meanwhile, it is necessary to have a strategy to make the transition from current to extreme levels of computing less painful to the broad scientific computing community. In this presentation we will give an

overview of efforts, in particular those under the auspices of the US Department of Energy, which aim at extreme levels of computing.

○ 【講演】 15 時 15 分～16 時 15 分

Speaker

Kengo Nakajima (Information Technology Center The University of Tokyo)

Title

Effect of coarse grid solver of parallel multigrid preconditioning for
Krylov iterative method on multi-core cluster with 8,192 cores

Abstract

Strategy of relaxation in coarse grid level of multigrid solver strongly affects the convergence of entire solver.

In this talk, recent experiences on this issue using "T2K Open Supercomputer (Tokyo)" and "Cray-XT4" at Lawrence Berkeley National Laboratory with up to 8,192 cores will be presented.

Both of flat-MPI and Open/MPI hybrid parallel programming models are implemented. Results show that effect of coarse grid solver is significant when number of MPI processes is more than order of 103.

【講演】 16 時 30 分～17 時 30 分

Speaker

Takahiro Katagiri (Information Technology Center The University of Tokyo)

Title

Xabclib2011 : A Sparse Iterative Solver with an Auto-tuning Facility
--- New AT Functions Developed in 2011 ---

Abstract

We have developed a sparse iterative solver with an auto-tuning (AT) facility, named Xabclib. In this presentation, we will present newly developed functions for the AT. The release version including the new AT functions is called Xabclib2011.

The main functions for AT in Xabclib2011 are follows:

In the first AT function, we have developed codes from scratch to establish thread-safety in Xabclib2011. This enables us to establish hybrid MPI programming for the library.

In the second AT function, we have developed new implementations for linear algebra algorithms. The first algorithm is based on the BiCG-Stab algorithm for linear equations solver. This establishes stability of convergence to conventional

preconditioning implementation. The second algorithm is based on the Arnoldi algorithm with explicit restart for eigensolvers. This establishes an efficient restating with real vectors, although its solution vectors are complex in general.

In the third AT function, we have developed a new implementation for sparse matrix vector multiplications (SpMV) with segmented scan (SS) method for scalar processors. This establishes dynamic parameter change for SS on SpMV.

We will show preliminary results of performance evaluation with Xacblib2011 by using one node of the T2K Open supercomputer, which is the AMD quad-core Opteron with 16 cores.

○18 時 00 分～

懇親会（根津周辺）

ASE 研究会の開催情報はメーリングリストで発信をしております。研究会メーリングリストに参加ご希望の方は、ASE 研究会幹事の片桐（katagiri@cc.u-tokyo.ac.jp）までお知らせください。

以上

Extreme Levels of Computing: Challenges and Opportunities

Osni Marques

Lawrence Berkeley National Laboratory
OAMarques@lbl.gov

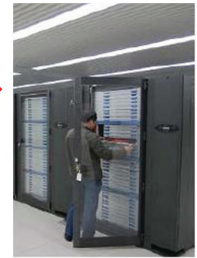
Outline

- 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)

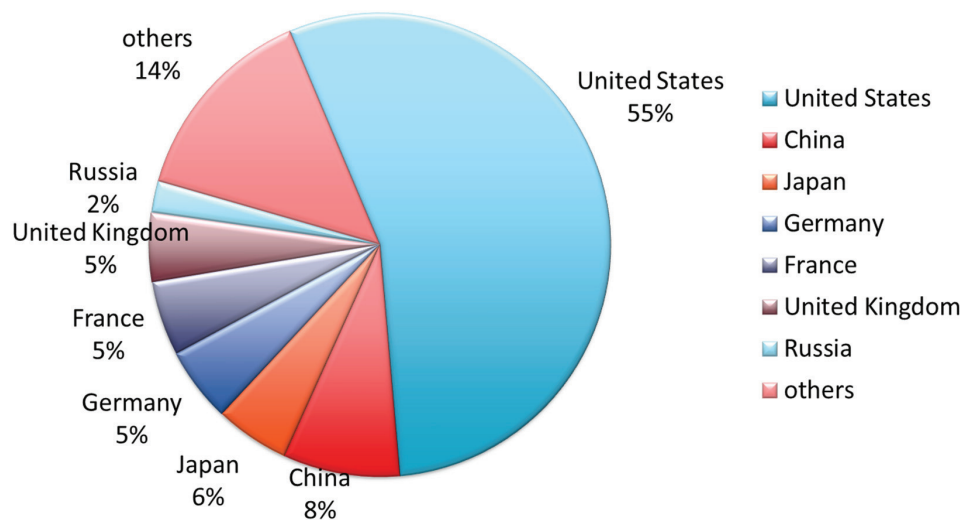
www.top500.org

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

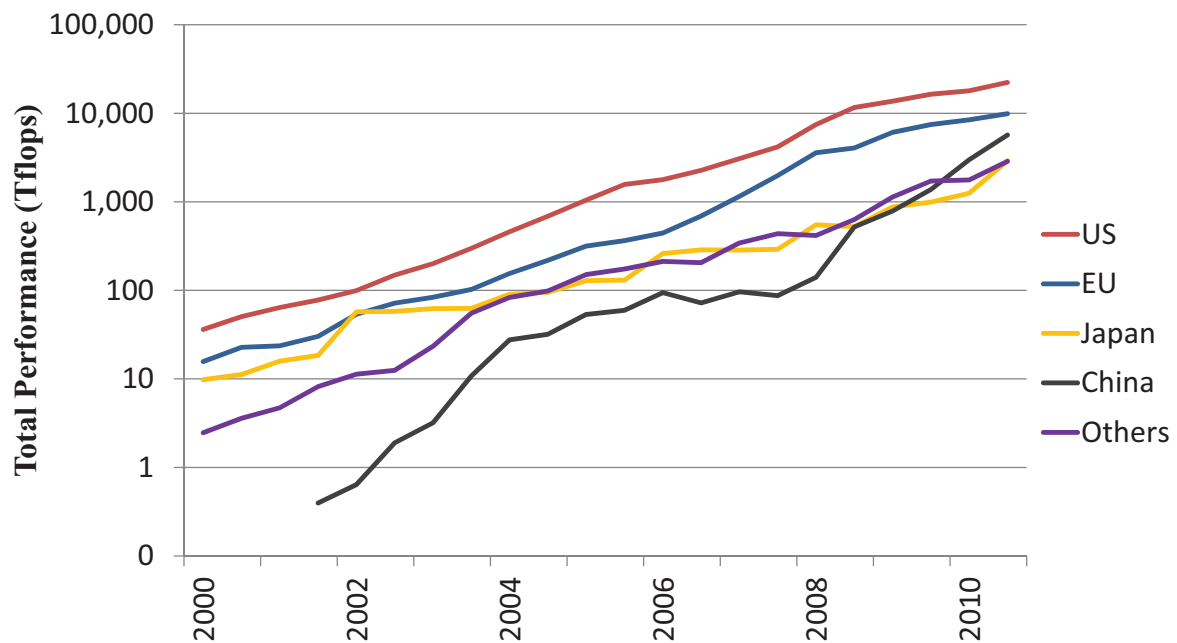


"2011 HPCwire People to Watch"

Countries / System Share



HPC Resources per Country



HPC in China



- 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).

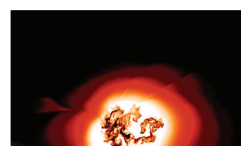
ASCR Computing Facilities and Research Areas

ORNL National Center for Computational Sciences

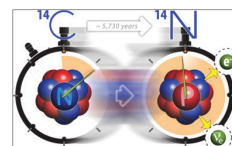


"At the California Institute of Technology, they're developing a way to turn sunlight and water into fuel for our cars. At Oak Ridge National Laboratory, they're using supercomputers to get a lot more power out of our nuclear facilities. With more research and incentives, we can break our dependence 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.

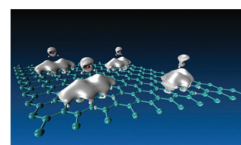
- XT5 partition: 18,688 dual hex-core AMD Opteron (Istanbul proc. running at 2.6GHz, 16GB of DDR2-800 memory) compute nodes; SeaStar 2+ router; 2.3 PF peak
- XT4 partition: 7,832 quad-core AMD Opteron (Budapest proc. running at 2.1 GHz, 8 GB of DDR2-800 memory) compute nodes, SeaStar2 router; 263 TF 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

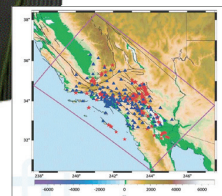
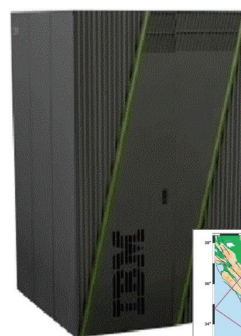


global climate change

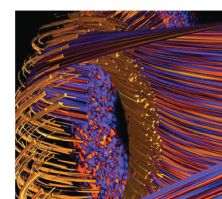
ANL Leadership Computing Facility



- 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



plasma microturbulence

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	Erds (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

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)**

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

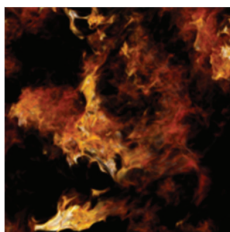
(leadership facilities)

- 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

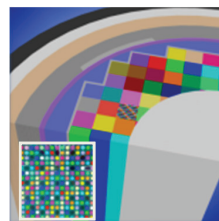
Source: Horst Simon

Leadership Computing: Scientific Progress at the Petascale

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

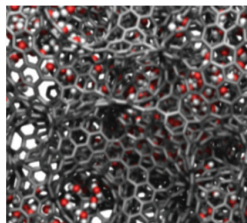


Turbulence: understanding the statistical geometry of turbulent dispersion of pollutants in the environment

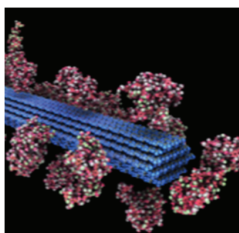
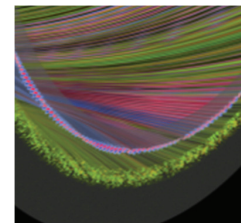


Nuclear Energy: high-fidelity predictive simulation tools for the design of next-generation nuclear reactors to safely increase operating margins

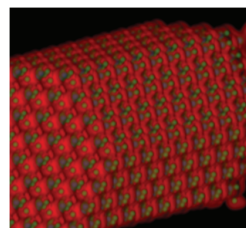
Energy Storage: understanding the storage and flow of energy in next-generation nanostructured carbon tube supercapacitors



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

- 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 Poinot, 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

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

Next Levels of Computing (Extreme Scales of Computing)

Scientific Grand Challenges: Needs and Potential Outcomes

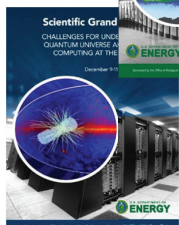
www.sc.doe.gov/ascr/ProgramDocuments/ProgDocs.html



*Town Hall Meetings,
2007*



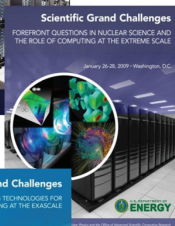
*Climate Change
Science, 2008*



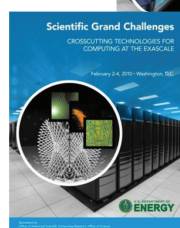
*Understanding the
Quantum Universe,
2008*



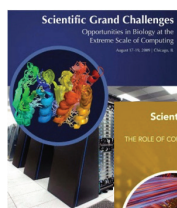
*Nuclear Energy
Systems, 2009*



*Nuclear
Science,
2009*



*Crosscutting
Technologies,
2010*



*Opportunities in Biology,
2009*



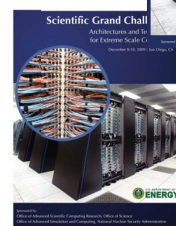
*Fusion Energy Sciences,
2009*



*National Security,
2009*



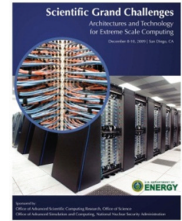
*Basic Energy
Sciences, 2009*



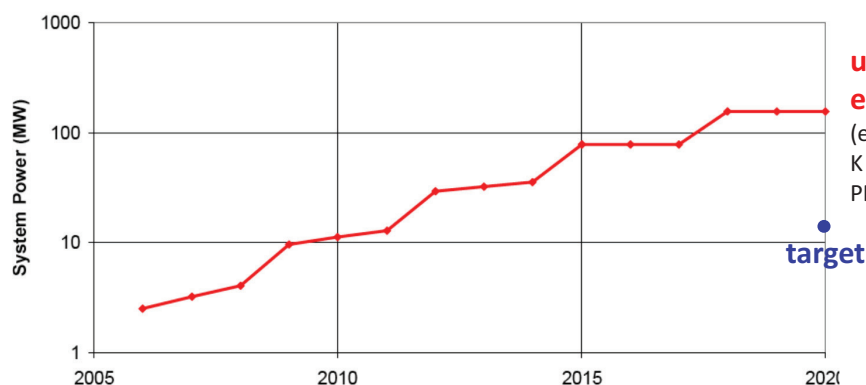
*Architectures and Technology
for Extreme Scale Computing,
2009*

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 ...



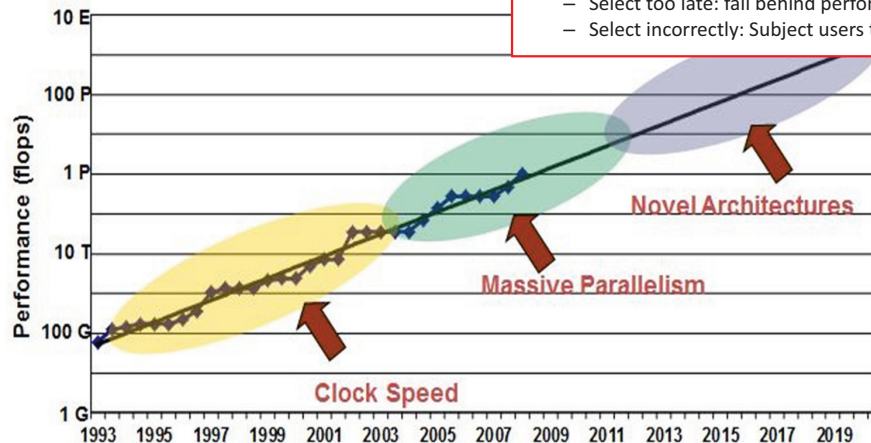
usual scaling → revised estimate: 500 plus MW
(e.g. Blue Waters @ NCSA: 300 K eight-core IBM Power 7; 10 PF peak, 15 MW)

> \$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

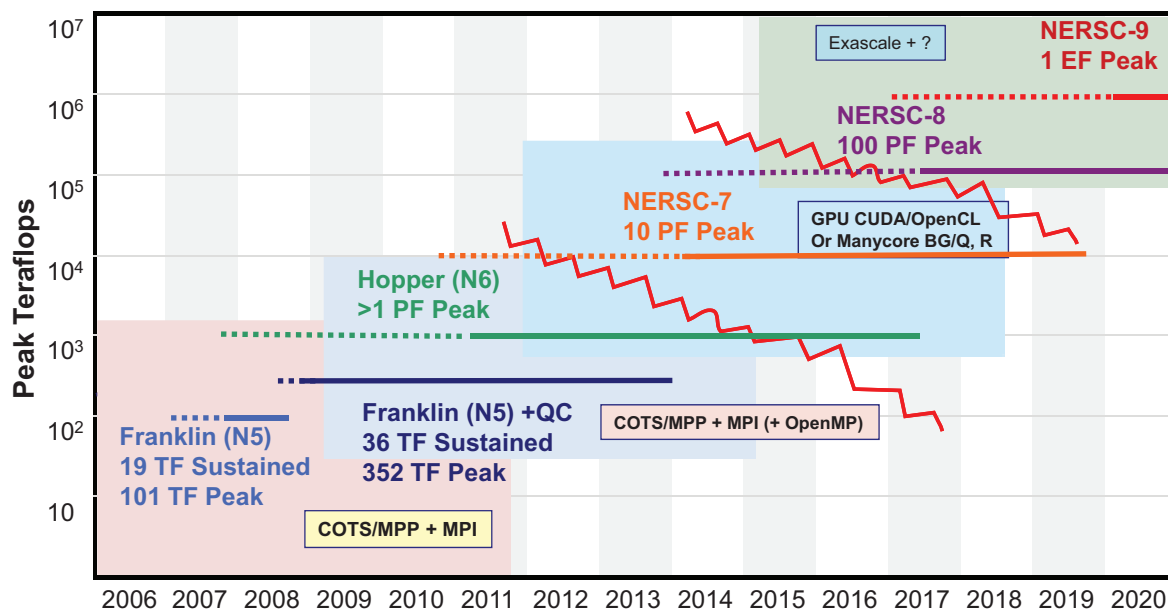


Technology Challenges

- 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

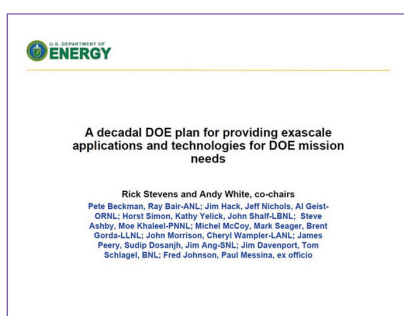
... and impacts to a facility like NERSC



Source: Horst Simon

A 10-year Plan for DOE

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

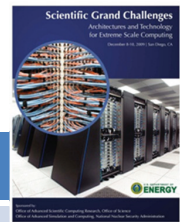


Plan discusses:

- 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

Two Potential Architectures (“swim lanes”)

*Architectures and Technology
for Extreme Scale Computing,
2009*

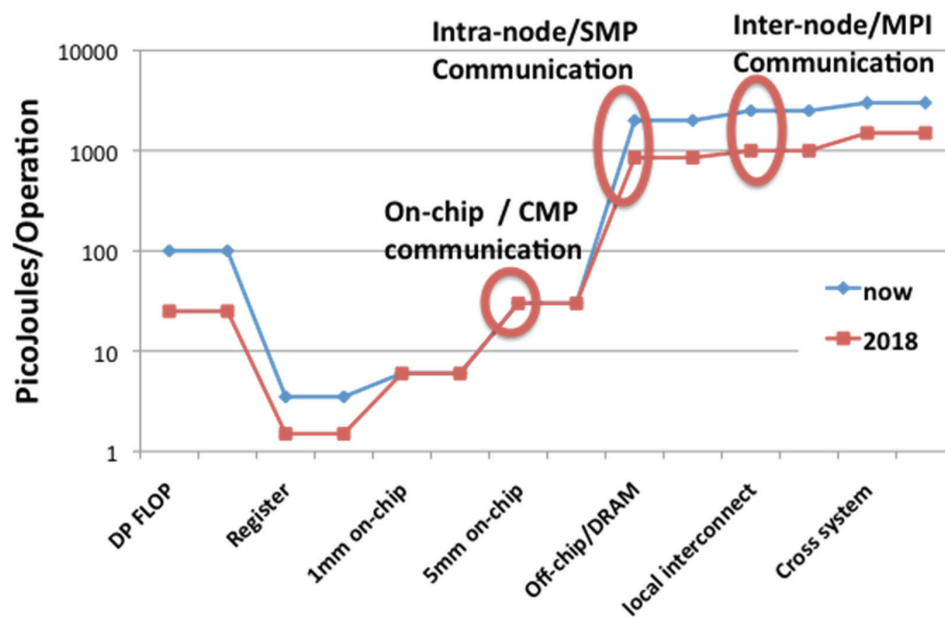


System attributes	2010	2015		2018	
System peak	2 PF	200 PF		1 EF	
Power	6 MW	15 MW		20 MW	
System memory	0.3 PB	5 PB		32-64 PB	
Node performance	125 GF	0.5 TF	7 TF	1 TF	10 TF
Node memory BW	25 GB/s	0.1 TB/s	1 TB/s	0.4 TB/s	4 TB/s
Node concurrency	12	O(100)	O(1,000)	O(1,000)	O(10,000)
System size (nodes)	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)	

Critical Exascale Technology Investments

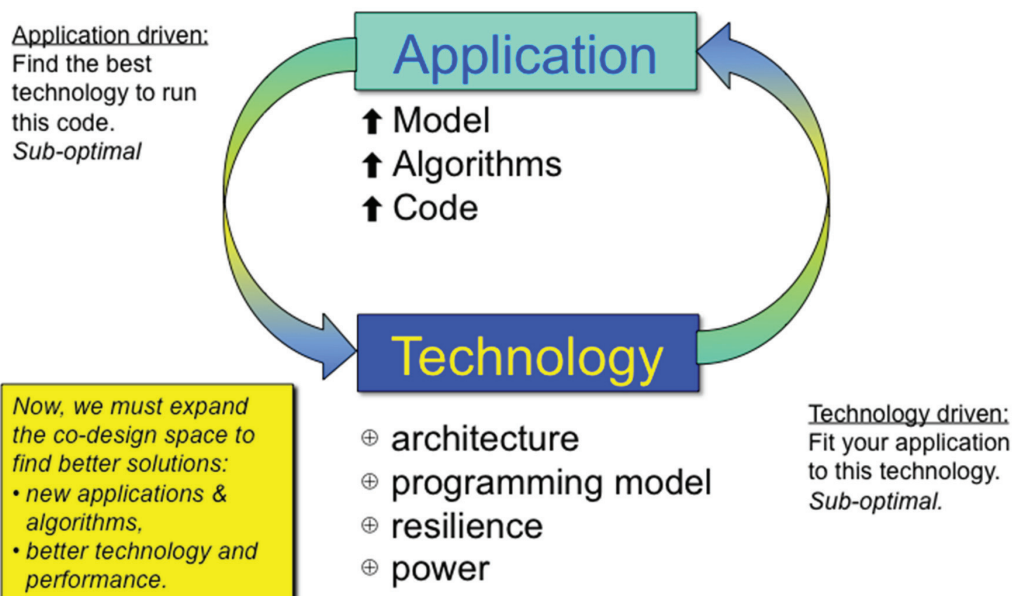
- **System power.** First class constraint on exascale system performance and effectiveness.
- **Memory.** Important component of meeting exascale power and applications goals.
- **Programming model.** Early investment in several efforts to decide in 2013 on exascale programming model, allowing exemplar 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

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*

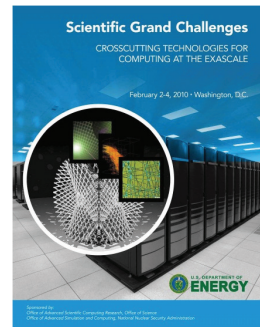
- Approaches for reducing and/or managing power requirements for high performance computing systems, including the memory and storage hierarchy.
- Approaches for reducing and/or managing heat in high performance computing systems.
- Methods for improving system resilience and managing the component failure 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 of 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



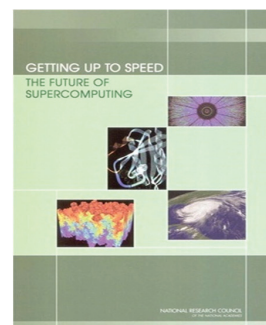
- Initiated by DOE and NSF
- System Software
- Development Environments
- Applications
- Cross-Cutting Dimensions

(Proposed) DOE Exascale Software Center (ESC):

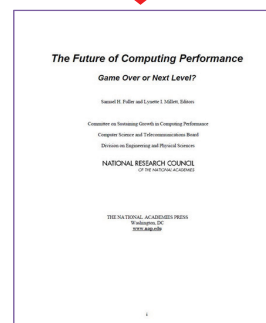
- Ensure successful deployment of coordinated exascale software stack on exascale platforms
- Deliver high quality system software for exascale platforms (~2015, ~2018)
- Identify software gaps, research and development solutions, test and support deployment
- Increase the productivity and capability 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

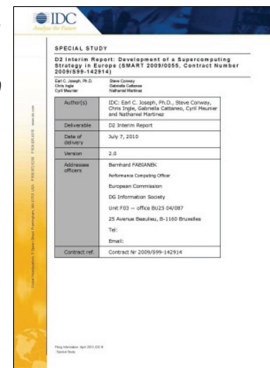


2010

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/oi/se/g8initiative

*Development of a
Supercomputing Strategy in
Europe, 2010*



Thank You !