Future Trends in Advanced Scientific Computing for Open Science

Presented to

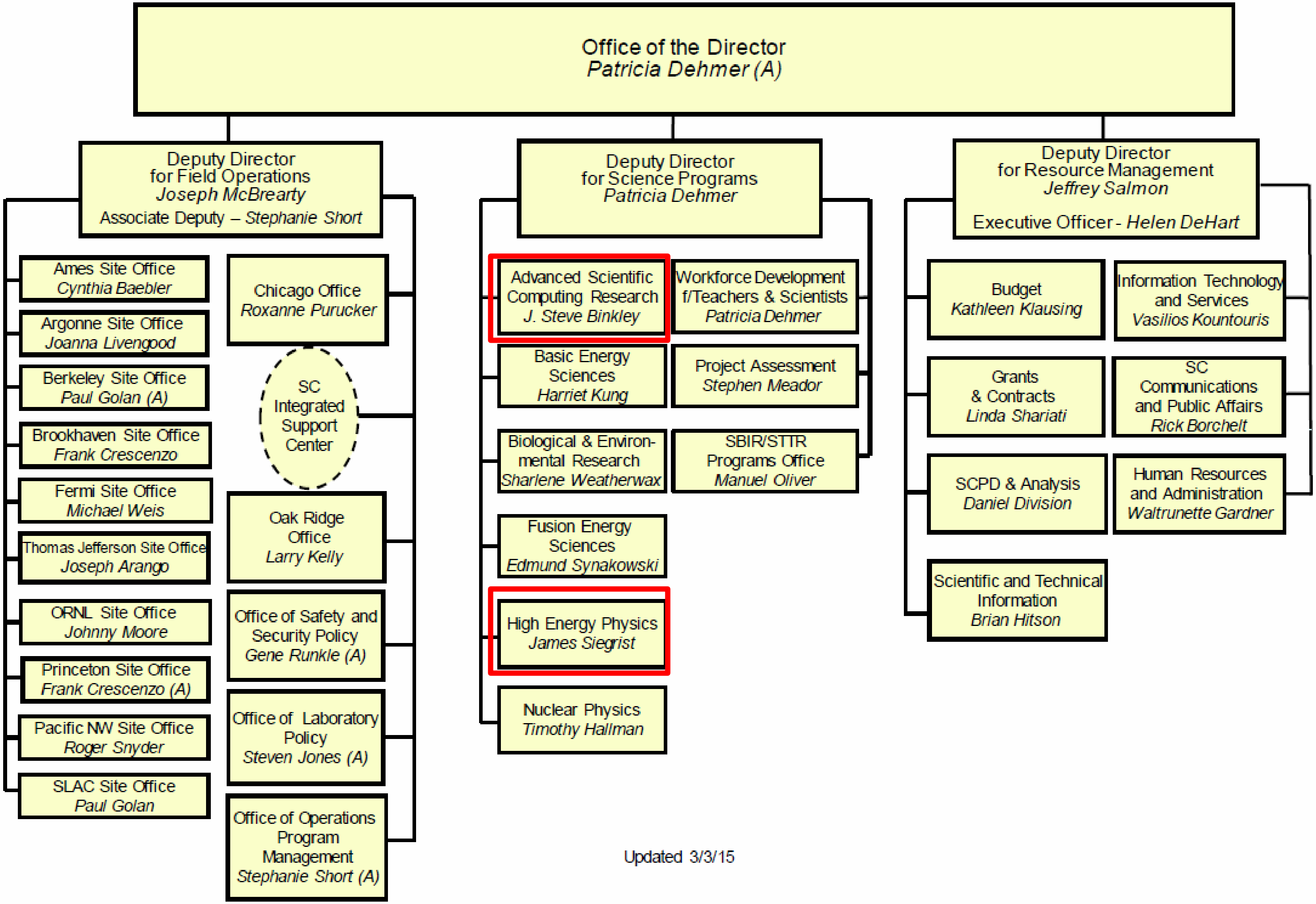
the High-Energy Physics Advisory Panel

by

Steve Binkley
DOE Office of Advanced Scientific Computing Research
Steve.Binkley@science.doe.gov

April 7, 2015

U.S. DEPARTMENT OF ENERGY | Office of Science

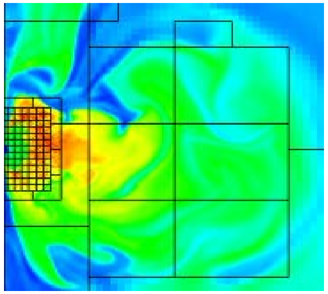


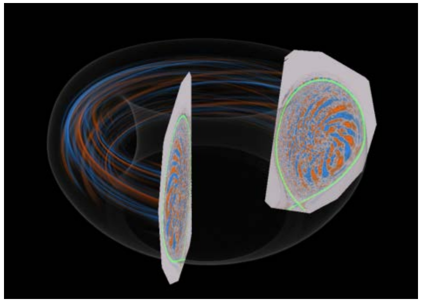
Updated 3/3/15

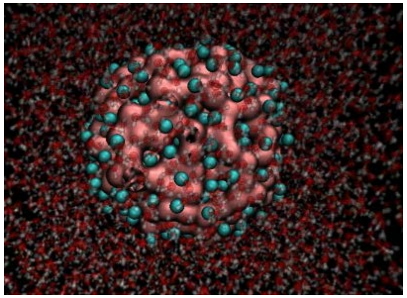
Advanced Scientific Computing Research

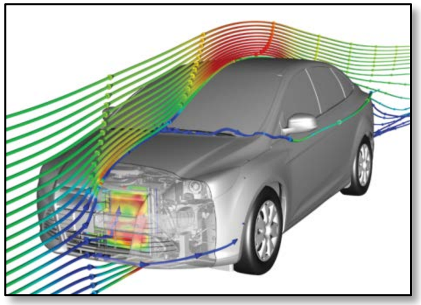
Computational and networking capabilities to extend the frontiers of science and technology

- **Mathematics research** to address challenges of increasing complexity within DOE's mission areas from a mathematical perspective. This requires integrated, iterative processes across multiple mathematical disciplines.
- **Computer science research** to increase the productivity and integrity of HPC systems and simulations, and support data management, analysis, and visualization techniques.
- **SciDAC partnerships** to dramatically accelerate progress in scientific computing that delivers breakthrough scientific results.
- **Exascale computing** research and development of capable exascale hardware architectures and system software, including the deployment of programming environments for energy-efficient, data-intensive applications, and engagement with HPC vendors to deliver systems that address the exascale challenges.
- **Facilities** operate with at least 90% availability while continuing planned upgrades – begin deployment of 10-40 petaflop upgrade at NERSC and continue preparations for 75-200 petaflop upgrades at each LCF.
- Continue a postdoctoral program at the ASCR facilities and provide funding for the Computational Science Graduate Fellowship to address DOE workforce needs.









Leadership and Production Computing Facilities

Titan:
- Peak performance of 27.1 PF
- 18,688 Hybrid Compute Nodes
- 8.9 MW peak power

Mira:
- Peak performance of 10 PF
- 49,152 Compute Nodes
- 4.8 MW peak power

Edison XC30:
- Peak performance 2.4 PF
- 124,608 processing cores
- 2.1 MW peak power

ASCR Computing Upgrades At a Glance

now ← → future

System attributes	NERSC Now	OLCF Now	ALCF Now	*NERSC Upgrade*	*OLCF Upgrade*	*ALCF Upgrade*
Name/Planned Installation	**Edison**	**TITAN**	**MIRA**	***Cori 2016***	***Summit 2017-2018***	***Aurora 2018-2019***
System peak (PF)	2.4	27	10	*>30*	*150*	*>150*
Peak Power (MW)	3	8.2	4.8	*<3.7*	*10*	*~13*
System memory per node	64 GB	38 GB	16 GB	*64-128 GB DDR4 16 GB High Bandwidth*	> 512 GB (High Bandwidth memory and DDR4)	TBA
Node performance (TF)	0.460	1.452	0.204	*>3*	*>40*	*>15 times Mira*
Node processors	Intel Ivy Bridge	AMD Opteron Nvidia Kepler	64-bit PowerPC A2	*Intel Knights Landing many core CPUs Intel Haswell CPU in data partition*	*Multiple IBM Power9 CPUs & mulitple Nvidia Voltas GPUS*	*TBA*
System size (nodes)	5,200 nodes	18,688 nodes	49,152	*9,300 nodes 1,900 nodes in data partition*	*~3,500 nodes*	*~50,000 nodes*
System Interconnect	Aries	Gemini	5D Torus	*Aries*	*Dual Rail EDR-IB*	*TBA*
File System	17.6 PB, 168 GBs, Lustre®	32 PB, 1 TB/s, Lustre®	*GPFS™*	*28 PB, 744 GB/sec , Lustre®*	*120 PB, 1 TB/s, GPFS™*	*TBA*

Delivering Capabilities that Keep the U.S. Competitive
"ASCR inside"

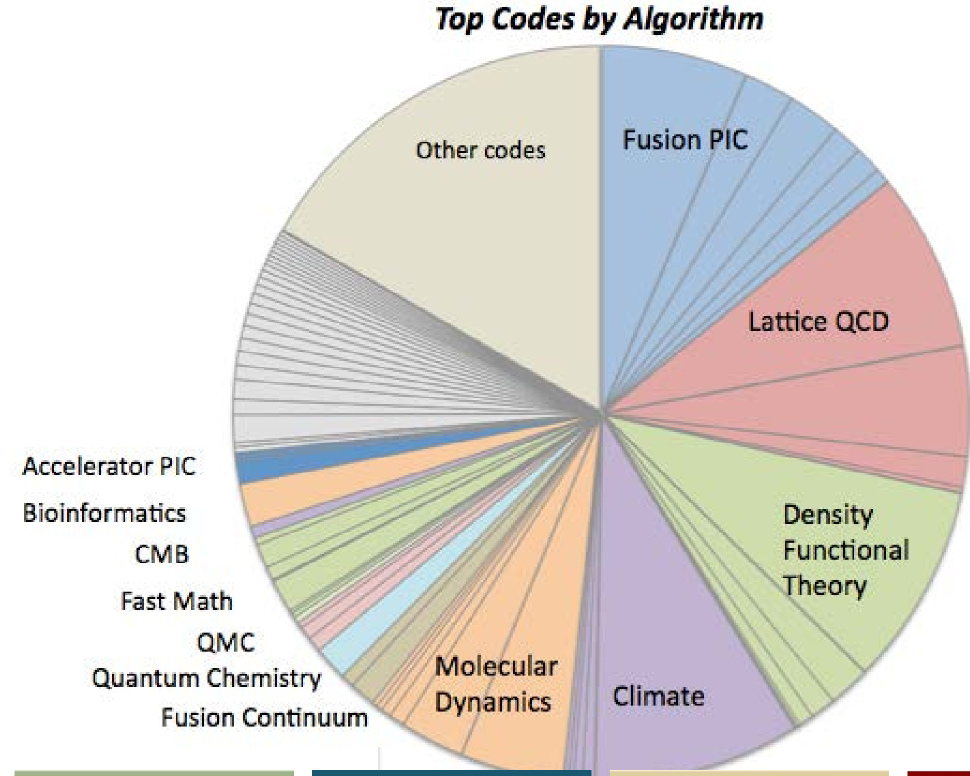
A few ASCR Technologies and the Companies that Use them

- **MPICH** – Message passing library
  "MPICH's impact comes from the fact that since it is open source, portable, efficient, and solid, most computer vendors have chosen it as the foundation of the MPI implementation that they supply to their customers as part of their system software." - Rusty Lusk, MPICH consortia
  ""MPICH is critical to the development of the F135 engine, which will power America's next-generation Joint Strike Fighter," - Robert Barnhardt, VP, Pratt & Whitney
- **Fastbit** – Search algorithm for large-scale datasets
  "FastBit is at least 10 times, in many situations 100 times, faster than current commercial database technologies" – Senior Software Engineer, Yahoo!
- **OSCARS** - On-demand virtual network circuits
  "It used to take three months, 13 network engineers, 250 plus e-mails and 20 international conference calls to set up an inter-continental virtual circuit. With OSCARS and collaborative projects, we can establish this link in 10 minutes." - Chin Guok, ESnet network engineer
- ***perfSONAR*** - network performance monitoring
  "These tools give us better visibility into the network, allowing us to troubleshoot performance issues quickly."
  -- Internet2 Network Performance Workshop participant

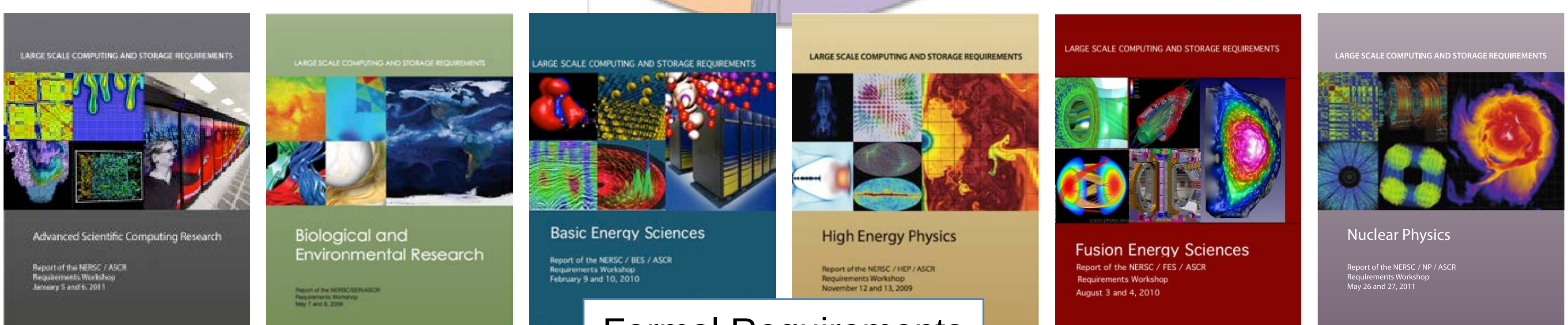
Trends Across the HPC Landscape

- **Increasing importance of computing and simulation within SC and DOE programs (across our missions: science, national security, and quest for cleaner energy)**
  - Continued full subscription of computational resources
  - Increasing importance of effective partnerships (domain + applied math + CS)
  - Difficulties persist in acquiring and retaining highly skilled workforce
- **The drive toward exascale – Scientifically important and challenging questions await exascale**
  - Dennard scaling is driving further increases in concurrency → billion-way concurrency is coming
  - Post-CORAL computer architectures may be significantly different → significant re-coding
  - Data movement will be increasingly costly
  - Power consumption remains an issue ($1M/megawatt-year)
  - A petaflop in a 19-inch rack
- **Nexus of big data and powerful compute is an emerging frontier**
- **The post-Moore's Law epoch is drawing nearer – we need to start preparing**

Computational Capacity is Based on Requirements



Computational Usage at NERSC (2014)



Formal Requirements Gathering Process

Implementation of Exascale Requirements Review

Series of workshops, one per domain program (hybrid of NERSC requirements reviews and Scientific Grand Challenges)

- **Location:** Washington DC area
- **Program Committee:** Representative community leaders from SC domain program office and ASCR facility staff
- **Attendance**: ~50 attendees including DOE program managers, DOE SC community representatives, ASCR supported applied mathematicians and computer scientists and a small number of Postdocs and senior CSGF fellows
- **Agenda**: Plenary session and themed breakout sessions determined by program committee
- **Pre-meeting homework**: Templates will be developed and provided to chairs and attendees of breakout session for discussing and documenting case studies
- **Output**: Summary workshop report written for each workshop.

Proposed Schedule*

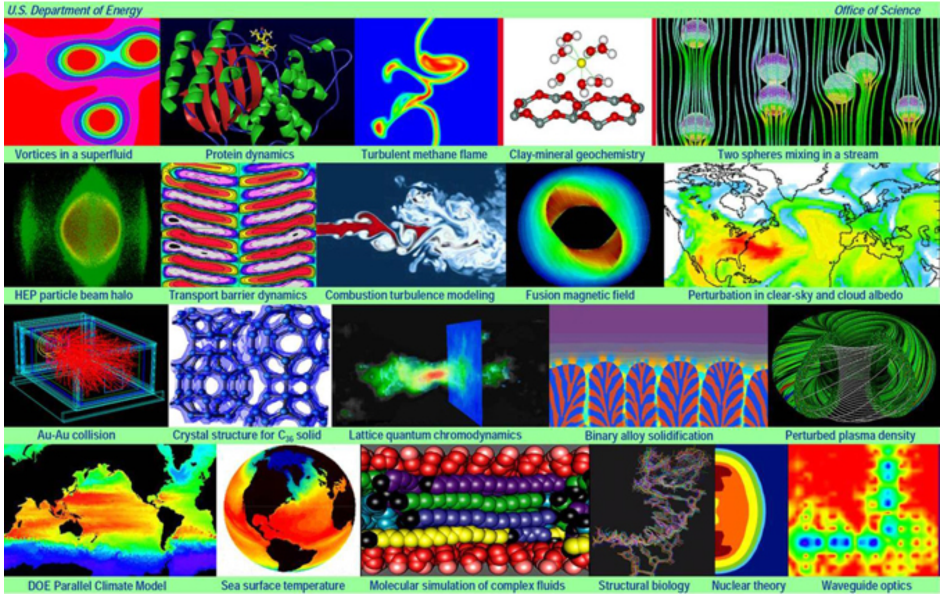
June 2015	HEP
September 2015	BES
December 2015	FES
April/March 2016	BER
June 2016	NP
September 2016	ASCR

*Dates under discussion with domain programs

ASCR Partnerships Across the Office of Science

- Partnerships facilitate transfer of ASCR research results into SC applications
- SciDAC focuses on the high end of high performance computing and addresses two challenges:
  1. to broaden the HPC community to increase the impact of high performance computing on the Department's missions, and
  2. to ensure that further progress at the forefront of HPC technology is rapidly assimilated into DOE applications.
- In its 3rd five-year generation, 18 partnerships span Office of Science programs
- SciDAC is leading U.S. industry into new ways to use computing to improve competitiveness.

SciDAC
Scientific Discovery through Advanced Computing



SciDAC Partnerships 2015

Biological and Environmental Research

- Computationally Efficient Schemes for BioGeochemical Cycles
- Multiscale Methods for Accurate, Efficient & Scale-Aware Models of the Earth System
- Predicting Ice Sheet and Climate Evolution at Extreme Scales

Basic Energy Sciences

- Advanced Modeling of Ions in Solutions, on Surfaces and in Biological Environments
- Developing Advanced Methods for Excited State Chemistry in the NWChem Software Suite
- Optimizing Superconductor Transport Properties through Large-Scale Simulation
- Scalable Computational Tools for Discovery and Design: Excited State Phenomena in Energy Materials
- Simulating the Generation, Evolution and Fate of Electronic Excitations in Molecular and Nanoscale Materials with First Principles Methods
- Discontinuous methods for massively parallel Quantum Molecular Dynamics: Li-ion interface dynamics from first principles

Fusion Energy Sciences

- Partnership for Edge Physics Simulation
- Plasma Surface Interactions

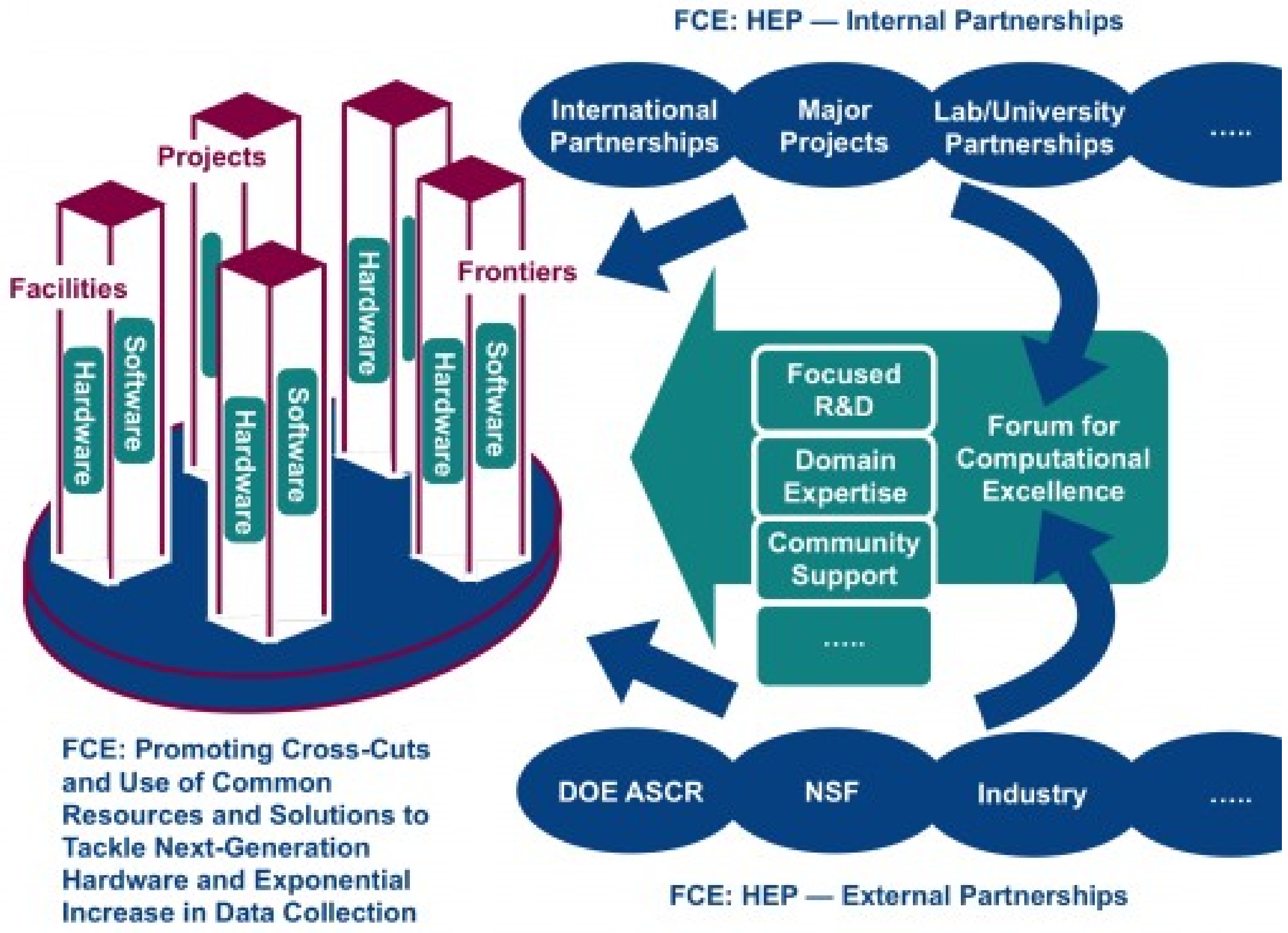
High Energy Physics

- Community Project for Accelerator Science and Simulation
- Computation-Driven Discovery for the Dark Universe
- Searching for Physics Beyond the Standard Model

Nuclear Physics

- A MultiScale Approach to Nuclear Structure and Reactions
- Computing Properties of Hadrons, Nuclei and Nuclear Matter from Quantum Chromodynamics
- Nuclear Computational Low Energy Initiative

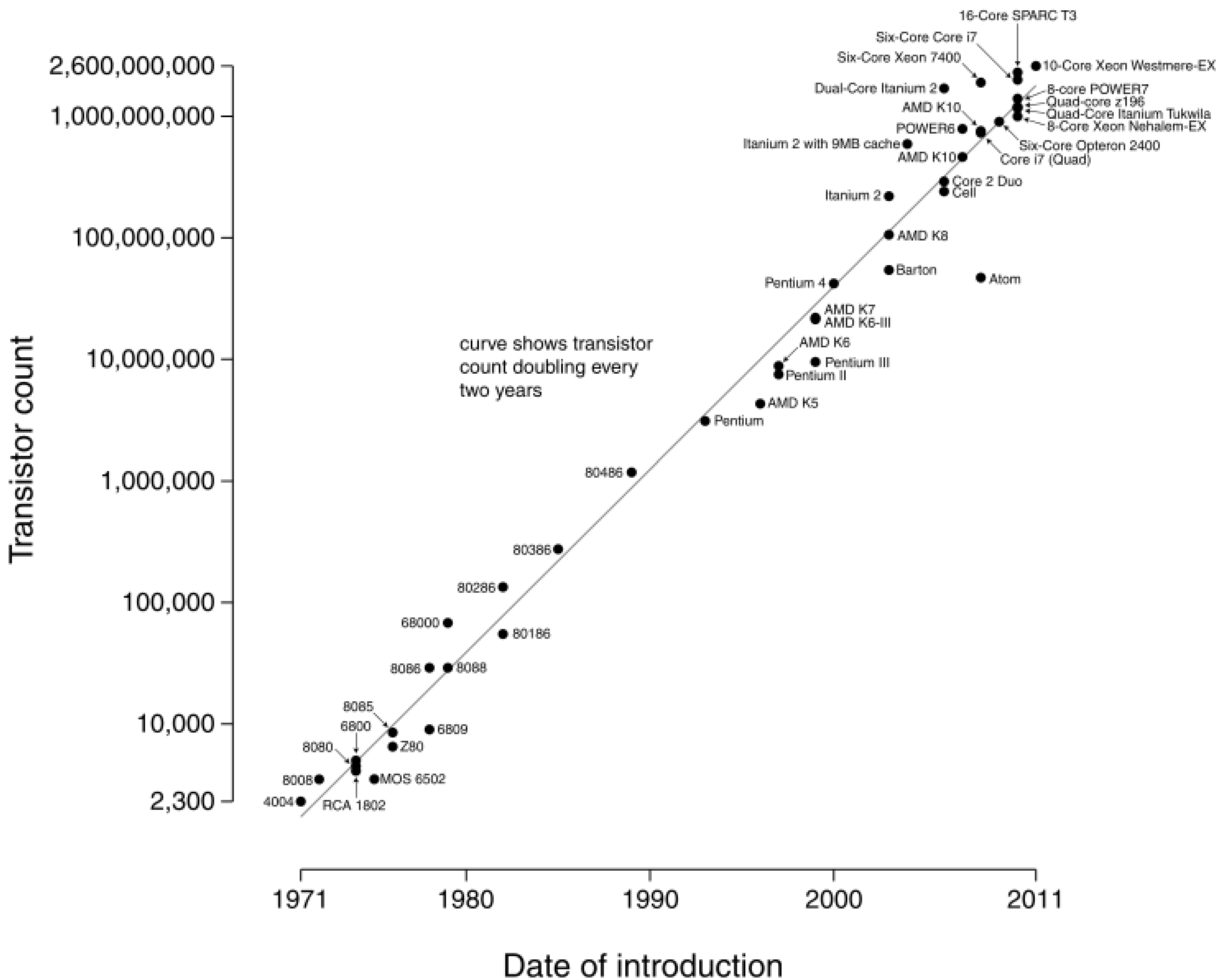
High Energy Physics Forum for Computational Excellence



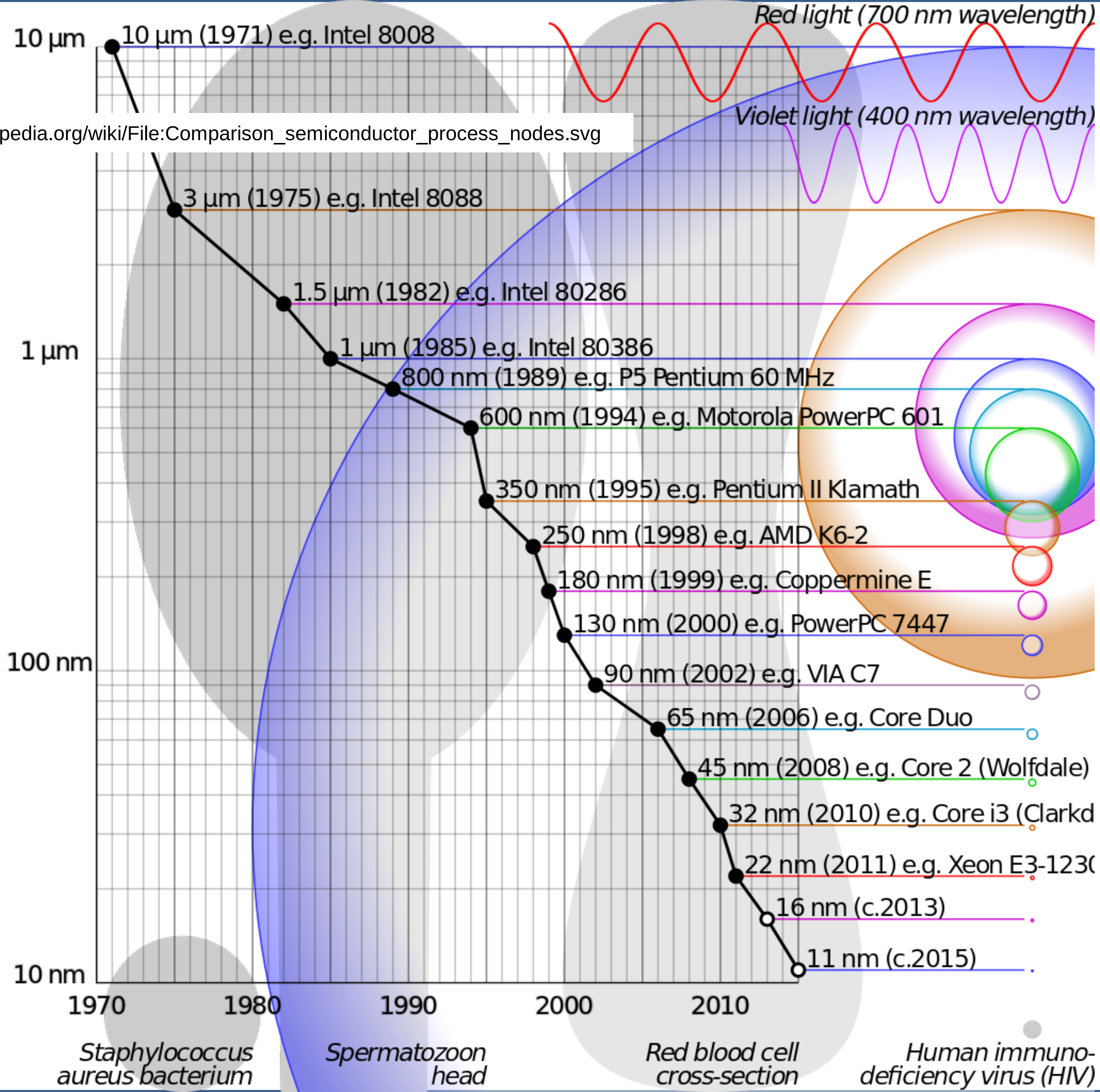
Trends Across the HPC Landscape

- **Increasing importance of computing and simulation within SC and DOE programs (across our missions: science, national security, and quest for cleaner energy)**
  - Continued full subscription of computational resources
  - Increasing importance of effective partnerships (domain + applied math + CS)
  - Difficulties persist in acquiring and retaining highly skilled workforce
- **The drive toward exascale – Scientifically important and challenging questions await exascale**
  - Dennard scaling is driving further increases in concurrency → billion-way concurrency is coming
  - Post-CORAL computer architectures may be significantly different → significant re-coding
  - Data movement will be increasingly costly
  - Power consumption remains an issue ($1M/megawatt-year)
  - A petaflop in a 19-inch rack
- **Nexus of big data and powerful compute is an emerging frontier**
- **The post-Moore's Law epoch is drawing nearer – we need to start preparing**

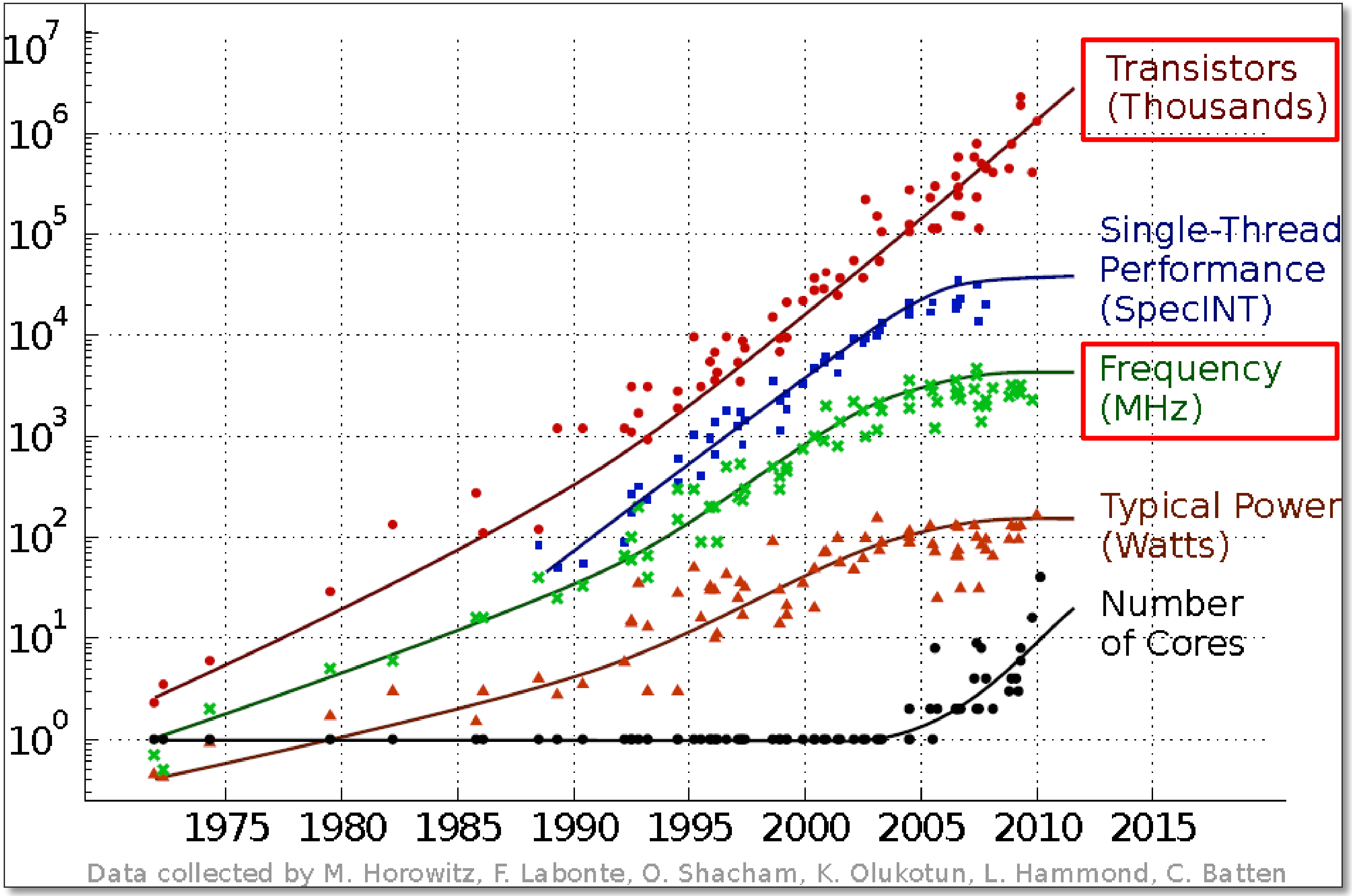
Microprocessor Transistor Counts 1971-2011 & Moore's Law



http://en.wikipedia.org/wiki/File:Comparison_semiconductor_process_nodes.svg



Progress in CMOS CPU Technology



Moore's Law continues

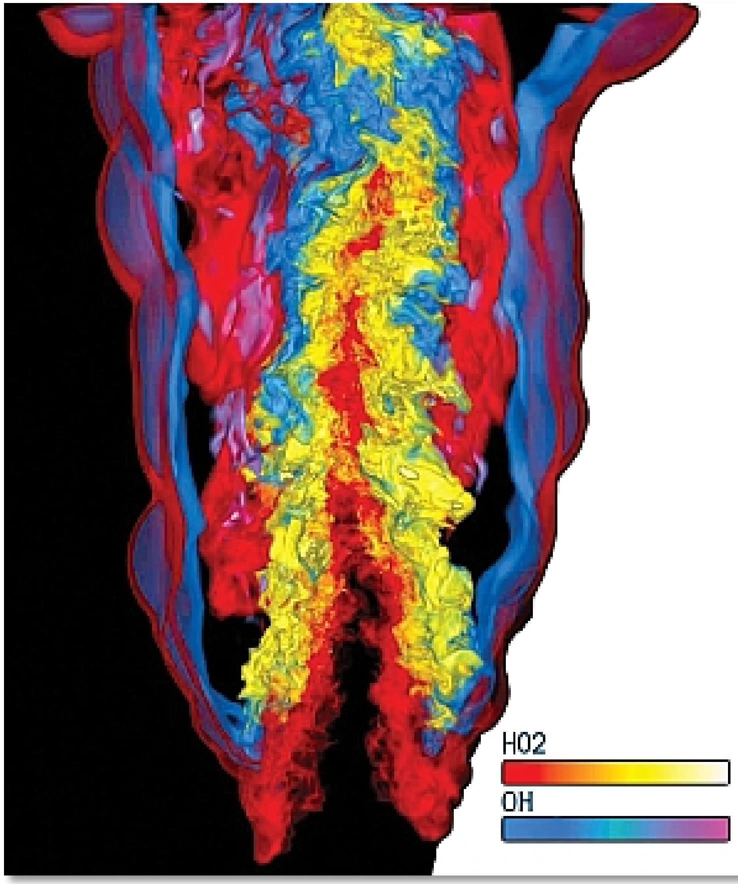
- Transistor count still doubles every 24 months

Dennard scaling stalls

- Voltage
- Clock Speed
- Power
- Performance/clock

Exascale Computing Initiative: Next Generation of Scientific Innovation

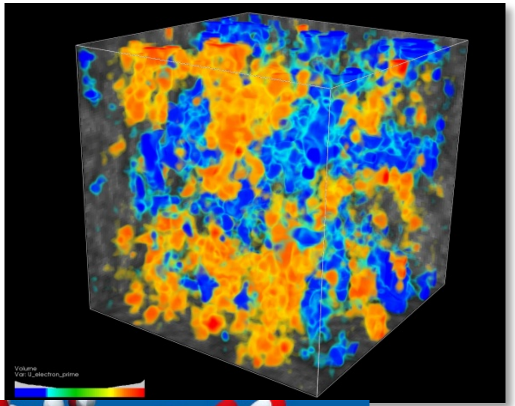
- **In partnership with NNSA**
- **"All-in" approach: hardware, software, applications, large data, underpinning applied math and computer science**
- **Driven by DOE's missions:**
  - National security – Stockpile Stewardship
  - Discovery science – next-generation materials
  - Mission-focused basic science in energy – next-generation climate software
- **DOE goal is to be "first to benefit" from Extreme Scale Computing**
  - Real applications at 100-1,000X capabilities of today's computers with a similar physical size and power footprint
  - Overcoming challenges of power consumption, high concurrency, reliability, usability
- **Extreme Scale Computing, cannot be achieved by a "business-as-usual," evolutionary approach**
  - Exascale systems will be based on marketable technology – Not a "one off" system (but may be "first of a kind")
  - Requires partnerships with U.S. computer vendors for required engineering, research and development for system architectures

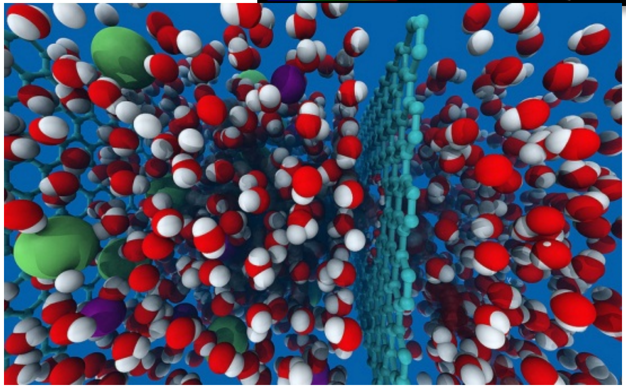


http://science.energy.gov/~/media/ascr/ascac/pdf/meetings/20141121/Exascale_Preliminary_Plan_V11_sb03c.pdf

Exascale Computing
The Vision

- **Exascale computing**
  - Achieve order 10^{18} operations per second and order 10^{18} bytes of storage
  - Address the next generation of scientific, engineering, and large-data workflows
  - Enable extreme scale computing: 1,000X capabilities of today's computers with a similar size and power footprint
  - Barriers: billion-way concurrency, energy consumption, memory
- **Productive system**
  - Usable by a wide variety of scientists and engineers
  - "Easier" to develop software & management of the system
- **Based on marketable technology**
  - Not a "one off" system
  - Scalable, sustainable technology, exploiting economies of scale and trickle-down effect
- **Deployed in early 2020s**



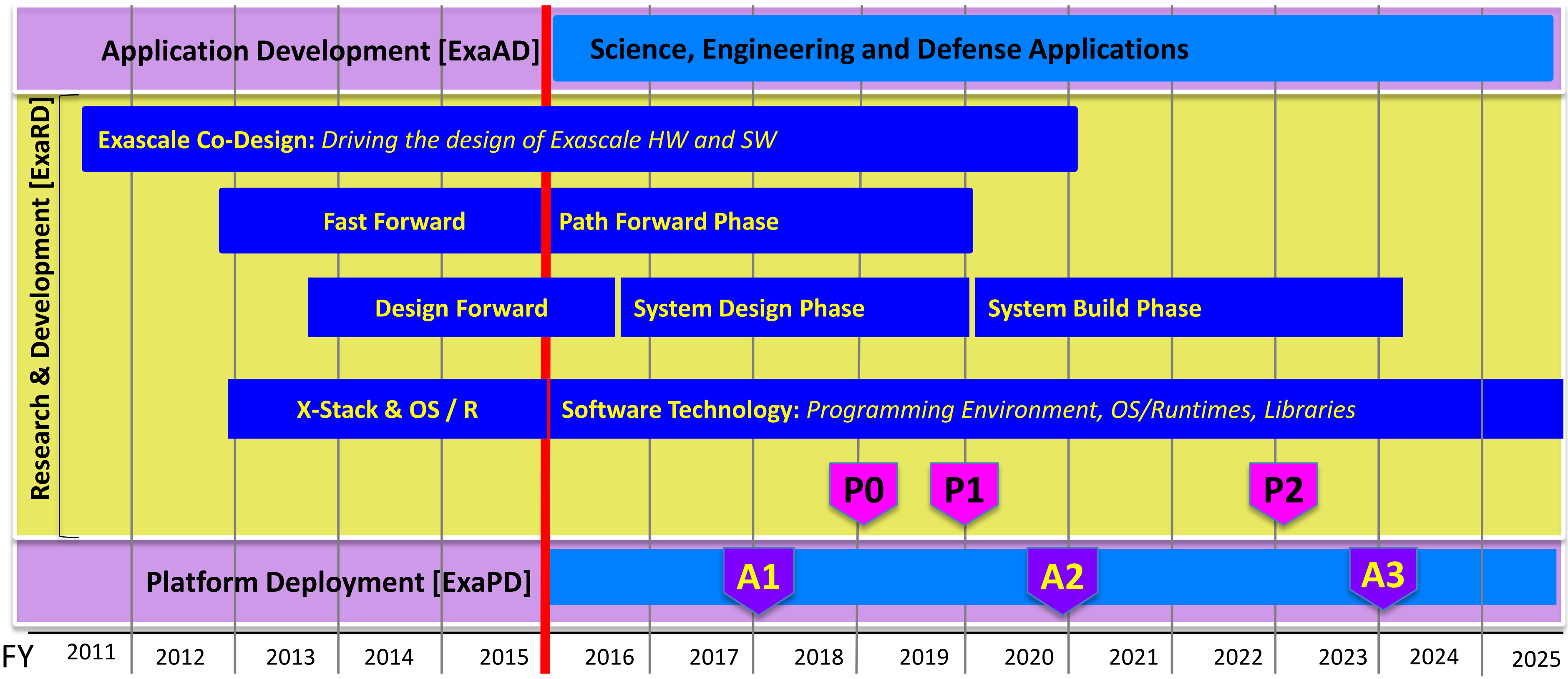


Top Ten Technical Approaches for Exascale
(Advanced Scientific Computing Advisory Committee)

1. **Energy efficiency:** Creating more energy efficient circuit, power, and cooling technologies.
2. **Interconnect technology:** Increasing the performance and energy efficiency of data movement.
3. **Memory technology:** Integrating advanced memory technologies to improve both capacity and bandwidth.
4. **Scalable System Software:** Developing scalable system software that is power and resilience aware.
5. **Programming systems:** Inventing new programming environments that express massive parallelism, data locality, and resilience
6. **Data management:** Creating data management software that can handle the volume, velocity and diversity of data that is anticipated.
7. **Exascale algorithms:** Reformulating science problems and refactoring their solution algorithms for exascale systems.
8. **Algorithms for discovery, design, and decision:** Facilitating mathematical optimization and uncertainty quantification for exascale discovery, design, and decision making.
9. **Resilience and correctness:** Ensuring correct scientific computation in face of faults, reproducibility, and algorithm verification challenges.
10. **Scientific productivity:** Increasing the productivity of computational scientists with new software engineering tools and environments.

http://science.energy.gov/~/media/ascr/ascac/pdf/reports/2013/report.pdf

Exascale Computing Initiative (ECI) Timeline



P0 Node Prototype
P1 Petascale Prototype
P2 Exascale Prototype
A1 CORAL
A2 APEX
A3 Exascale

CORAL – Collaboration ORNL, ANL, LLNL
APEX – Alliance for application Performance at the EXtreme scale

Exascale Key Performance Goals

Parameter	
Performance	Sustained 1 – 10 ExaOPS
Power	20 MW
Cabinets	200 - 300
System Memory	128 PB – 256 PB
Reliability	Consistent with current platforms
Productivity	Better than or consistent with current platforms
Scalable benchmarks	*Target speedup over "current" systems …*
Throughput benchmarks	*Target speedup over "current" systems …*

ExaOPS = 10^{18} Operations / sec

CORAL Node/Rack Layout – ORNL Summit Computer

CORAL rack layout
- 18 nodes
- **779 TF**
- 11 TB RAM
- **55 KW**

CORAL/Summit System
- ~200 racks

CORAL Node/Rack Layout – ORNL Summit Computer

2017
CORAL rack

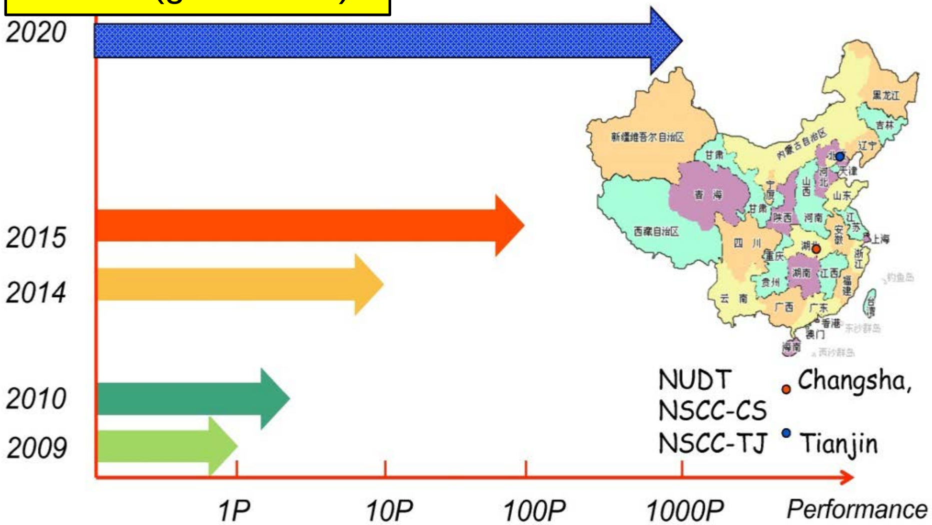
- 18 nodes
- **779 TF**
- 11 TB RAM
- **55 KW**
- **~6 ft^2**

1976
Cray 1

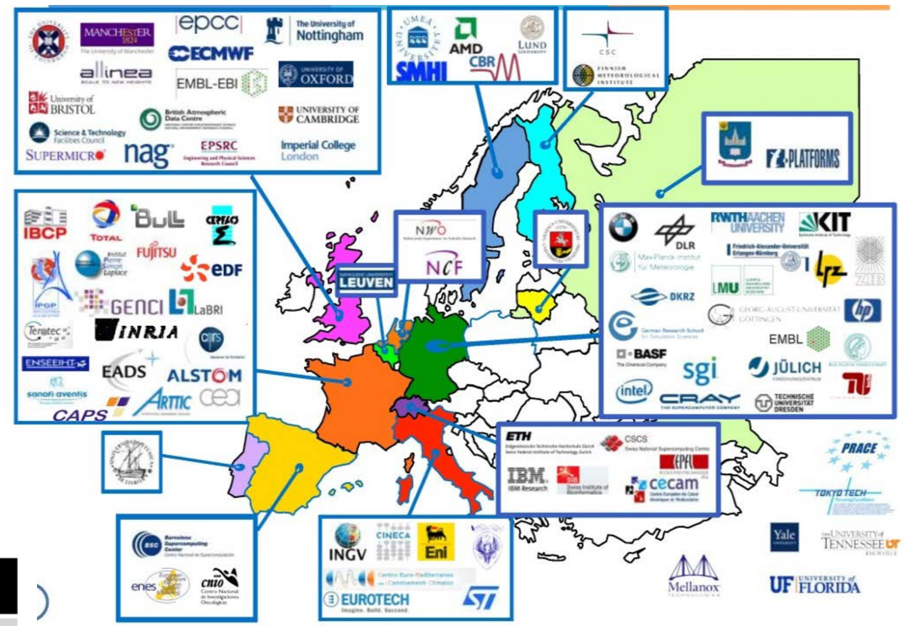
- **150 MF**
- 8 MB RAM
- **110 KW**
- **~150 ft^2**

International Competitors Have Ambitious Plans

China (goal: 2020)



European Union (goal: 2020)



Japan (goal: 2020)

2014 | 2015 | 2016 | 2017 | 2018 | 2019

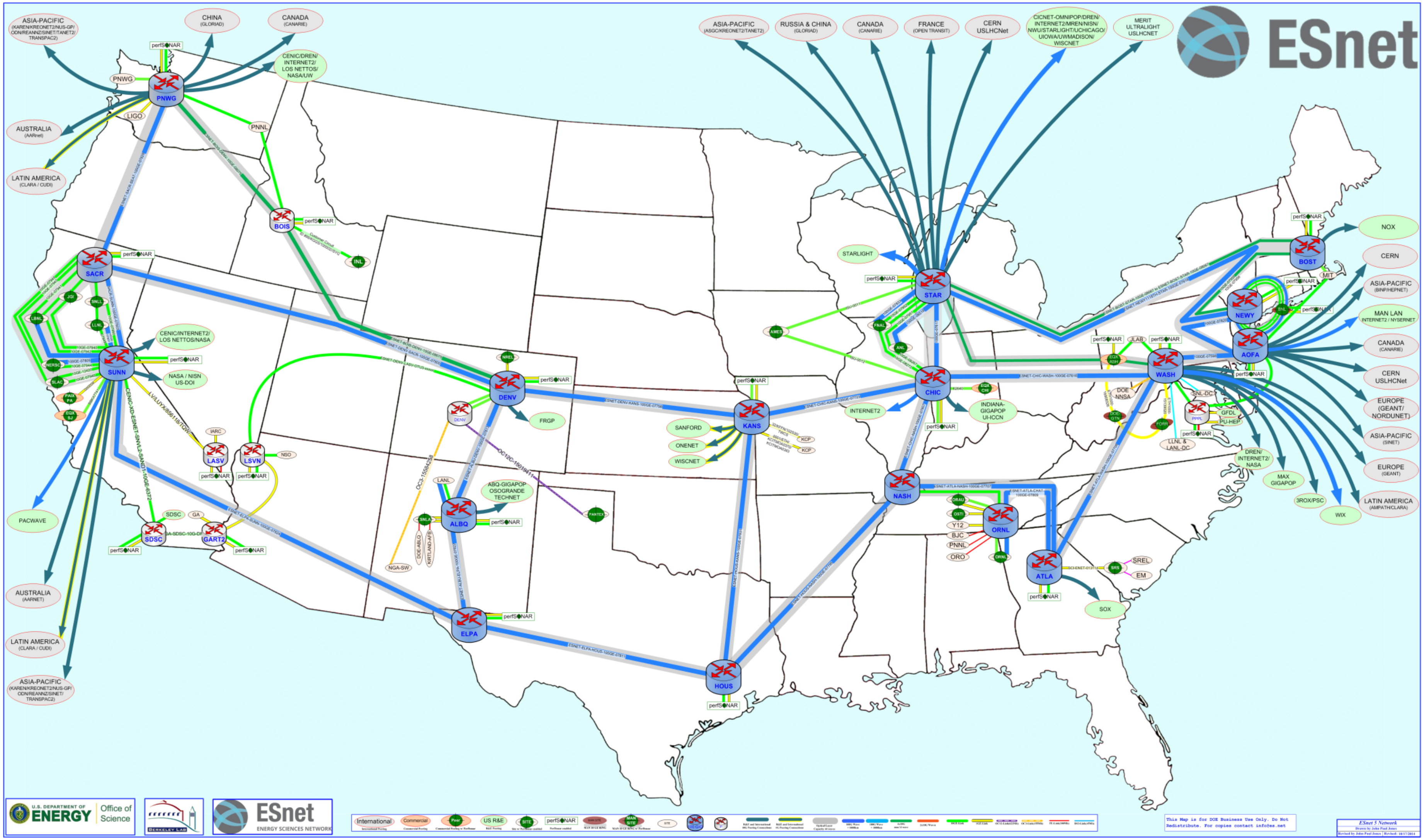
Deployment	Soft Env. → Arch. SIM V1.0 → Soft Env. → Arch. SIM V2.0 → Arch. ES → Post Development → Full System → Post Development
Heterogeneous Architecture	Feasibility Study (FS) → Arch. Selection & Evaluation → Development
Memory	FS → Device development & Evaluation → Development
Interconnect	FS → Component development → Development
Low-Power	FS → Component development → Development
Dependability	FS → Technology devel. with new devices → Development
Co-design	Study of Application → Arch. SIM → Evaluation → Arch. Optimization

Red triangle: To Software Layer
Yellow triangle: From Software Layer

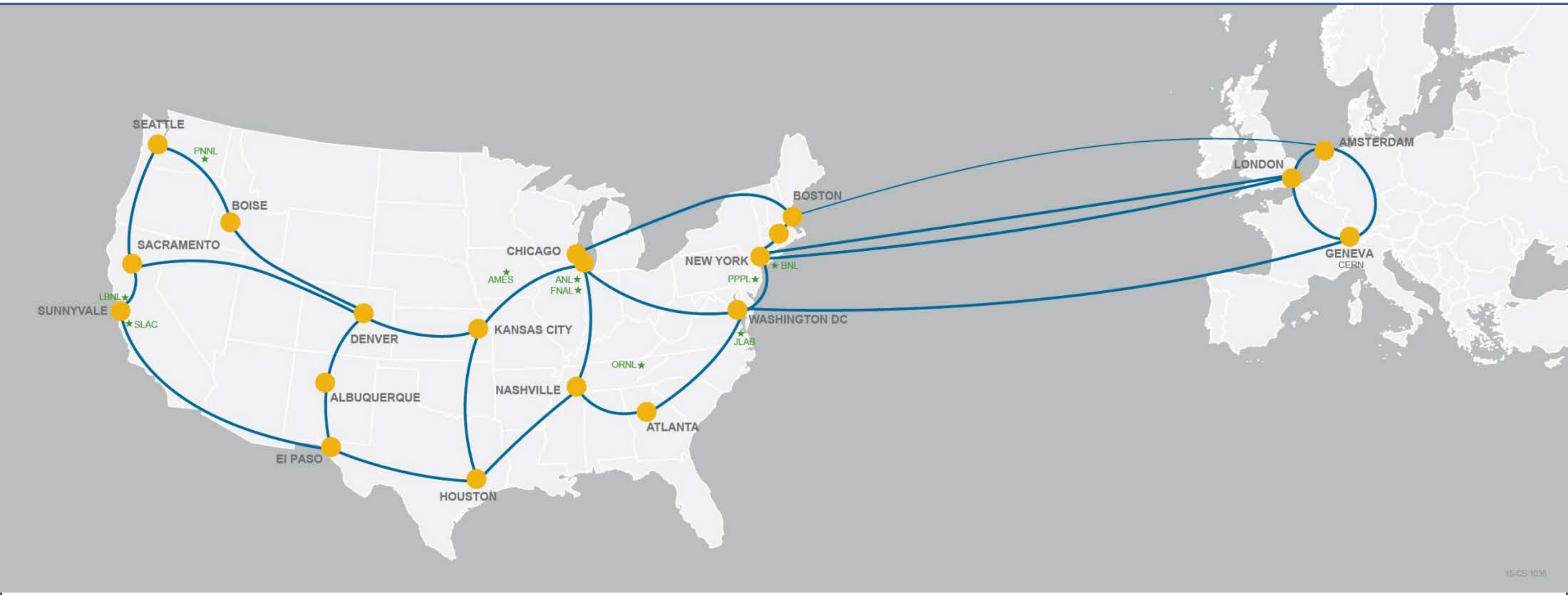
Other countries:
Russia
South Korea
India
Singapore
Australia

EU: J-Y. Berthou, IESP Workshop, Kobe, Japan, April 2012
China: Y. Lu, IESP Workshop, Kobe, Japan, April 2012
Japan: M. Kondo, IESP Workshop, Kobe, Japan, April 2012

A Balanced Infrastructure Requires High-Capacity Networks



ESnet Goes Global: Extension to Europe

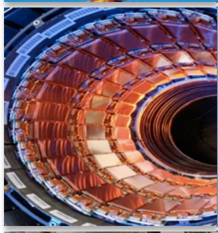


- 25% of all ESnet traffic goes to/from Europe
- 3x100+ Gbps across the Atlantic with redundant paths to serve all DOE missions
- Ready by March 2015 to support LHC Run 2 (was operational in Jan 2015)
- Will support 10x increase in transatlantic traffic from Large Hadron Collider

Extreme-Scale Science Data Explosion

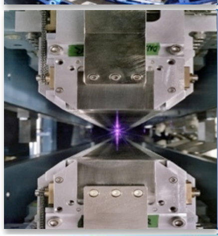
Genomics

Data Volume increases to 10 PB in FY21



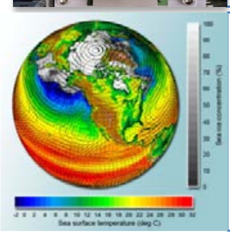
High Energy Physics (Large Hadron Collider)

15 PB of data/year



Light Sources

Approximately 300 TB/day



Climate

Data expected to be hundreds of 100 EB

Driven by exponential technology advances

Data sources

- Scientific Instruments
- Scientific Computing Facilities
- Simulation Results
- Observational data

Big Data and Big Compute

- Analyzing Big Data requires processing (e.g., search, transform, analyze, …)
- Extreme scale computing will enable timely and more complex processing of increasingly large Big Data sets

Data drivers from DOE Scientific User Facilities†

	2013 Current data rate*	2015 Projected need	2018 Projected need
HEP Cosmic Frontier example – Large Synoptic Survey Telescope	~0.2 GB/s	~0.5 GB/s	~1-10 GB/s
HEP Energy Frontier Example – Atlas LHC	1 GB/s*	2 GB/s*	4 GB/s*
HEP Intensity Frontier Example – Belle II	1 GB/s	2 GB/s	20 GB/s
BER Climate	100 GB/s	1000 GB/s	1000 GB /s
BER EMSL – one instrument example - TEM	100 – 1000 images (2Kx2K)/ per day	1000 Images/s = 2GB/s	1,000,000 Images/s = 2 TB/s
BER JGI example - Illumina HiSeq	18 MB/s	72 MB/s	600 MB/s
BES Advanced Photon Source example – 2-BM Beamline	1 GB/s/beamline		10 GB/s
BES Nano Science example – X-Ray Spectroscopy		100 MB x 100 excited atoms x 100 snapshots = 1 TB per point (P,T)	
BES Neutron Facilities	~0.05GB/s	~0.10 GB/s	~0.30GB/s

* Data Rate after 99% reduction in hardware data acquisition system

† http://science.energy.gov/~/media/ascr/pdf/program-documents/docs/ASKD_Report_V1_0.pdf

Post-Moore's Law Computing: What comes after exascale?

- CMOS lithographic feature sizes are approaching fundamental limits
  - Currently at 22 nm (both Intel and Nvidia)
  - 11 nm is projected for ~2015 (both Intel and Nvidia)
    - However, gate lengths may be smaller than 6 nm – corresponding gate dielectric thickness may reach a monolayer or less
  - The Intel roadmap reaches beyond 11 nm (7 nm and 5 nm) but may be unattainable
    - Non-silicon extensions of CMOS, e.g., using III-V materials or nanotubes/nanowires or non-CMOS technoligies, including molecular electronics, spin-based computing, single-electron devices, and graphene have been proposed
    - At scales of ~10 nm, quantum tunneling is expected to become significant
  - Capital costs for tooling are increasing dramatically as feature sizes shrink
- Options:
  - Computing using superconducting technologies
  - Quantum computing
  - Neuromorphic computing
  - Probabilistic computing
  - ???

Considerable R&D required

Recent, Related DOE/NNSA Community Activities

- Workshop on Materials Opportunities for Quantum Computing,* October 7-8, 2014, LANL
- Grand Challenges at the Intersections of QIS, Particle Physics, and Computing, December 11, 2014, DOE/HEP
- NNSA Workshop on Applications of Quantum Computing, February 5-6, 2015, SNL (organized by LANL)
- NSF conference on Mathematical Sciences Challenges in Quantum Information, February 12-13, 2015
- DOE/ASCR Workshop on Quantum Computing in Scientific Applications, Date: February 17-18, 2015

*Community organized

DOE/ASCR Workshop on Quantum Computing in Scientific Applications

Date: February 17-18, 2015
Venue: DoubleTree by Hilton Hotel, Bethesda MD
Contact: Dr. Ceren Susut-Bennett, DOE/SC/ASCR, ceren.susut-bennett@science.doe.gov

The goal of the workshop is to assess the viability of quantum computing technologies to meet the computational requirements in support of DOE's science and energy mission and to identify the potential impact of these technologies.

Research into quantum computing technologies is making rapid progress and it is important for the Office of Advanced Scientific Computing Research (ASCR) to understand the utilization of these new technologies for DOE-relevant applications and their impact on conventional computing systems.

The workshop will explore the following topics: (see next slide)

Workshop Goals

Research into quantum computing technologies is making rapid progress and it is important for ASCR to understand the utilization of these new technologies for DOE-relevant applications and their impact on conventional computing systems. The goal of the workshop is to explore the following topics:

1. Mission relevance: What aspects of DOE's science mission are suitable for quantum computing? What are the early tests that will demonstrate viability, or lack thereof, for the DOE's mission in fundamental and applied sciences?
2. Impact on Computing: How will quantum computing improve the properties of the computation with respect to conventional contemporary computational systems? Such attributes include, but are not limited, to performance, capacity, power, cost, generality and programmability.
3. Challenges: What are the challenges in adopting quantum computing technologies and developing the required infrastructure? What algorithm/application bottlenecks need to be solved before a quantum enabled system can be used for mission critical applications? What can ASCR do to mitigate these challenges?

http://www.csm.ornl.gov/workshops/ascrqcs2015/index.html

Interagency Working Group on Quantum Information Science

Chartered in November 2014 under the National Science and Technology Council, Committee on Science, Subcommittee on Physical Sciences

Purpose and Scope:

The Working Group fosters and coordinates Federal agency efforts to create the scientific basis, infrastructure, future technical workforce, and intellectual property that will be required to address agency missions and secure future U.S. competitiveness in quantum information science. These efforts encompass fundamental and applied research, technology development, and education and training. To accomplish this, the Working Group will develop and coordinate policies, programs, and budgets for quantum information science research and development.

Co-chairs: NIST, NSF, DOE

Quantum Information Science is Broad and Multidisciplinary

- **Spans multiple disciplines**
  - Computer science, applied mathematics, networking, information science
  - High-energy physics (advances in quantum theory)
  - Materials sciences (new materials)
- **Has many, conceivable applications**
  - Quantum simulation
  - Sensors
  - Cryptography
  - Communications, networking
  - Metrology/measurement, accurate timekeeping, …
  - …
- **Quantum theory is evolving**
  - new understanding of entanglement and information theory
- **Practical quantum computers require new technologies**
  - New materials for quantum devices
  - qubit manipulation, error correction
  - System integration
  - Software
  - …

ASCR at a Glance

Office of Advanced Scientific Computing Research

Associate Director – Steve Binkley
Phone: 301-903-7486
E-mail: Steve.Binkley@science.doe.gov

Research

Division Director – William Harrod
Phone: 301-903-5800
E-mail: William.Harrod@science.doe.gov

Facilities

Division Director – Barbara Helland
Phone: 301-903-9958
E-mail: Barbara.Helland@science.doe.gov

Relevant Websites

ASCR: science.energy.gov/ascr/

ASCR Workshops and Conferences:

science.energy.gov/ascr/news-and-resources/workshops-and-conferences/

SciDAC: www.scidac.gov

INCITE: science.energy.gov/ascr/facilities/incite/

Questions?