U.S. DEPARTMENT OF ENERGY

A DOE Laboratory plan for providing exascale applications and technologies for critical DOE mission needs

Argonne National Laboratory
Brookhaven National Laboratory
Oak Ridge National Laboratory
Lawrence Berkeley National Laboratory
Lawrence Livermore National Laboratory
Los Alamos National Laboratory
Pacific Northwest National Laboratory
Sandia National Laboratory

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, BNL;

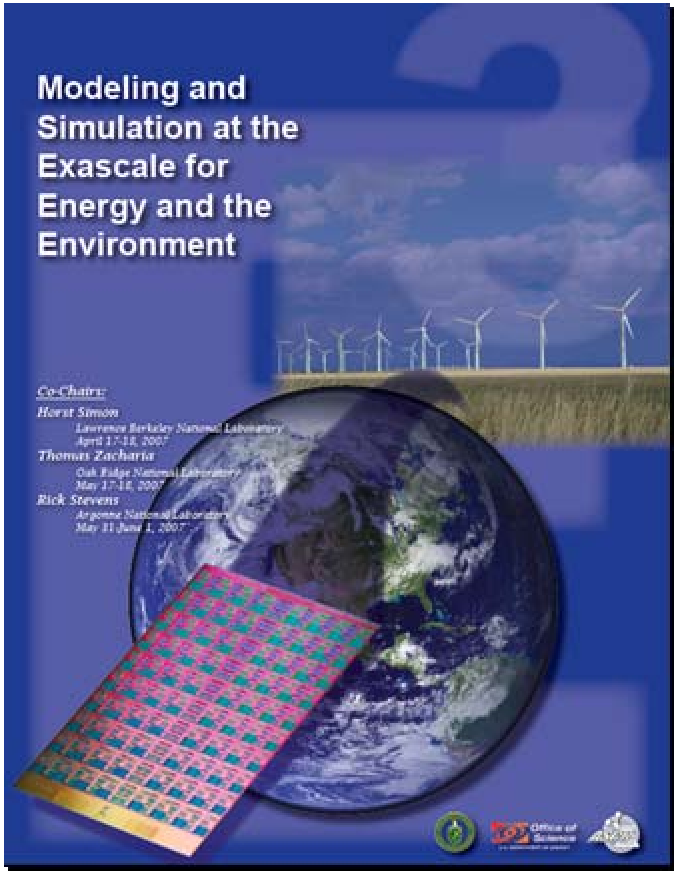
Fred Johnson, Paul Messina, ex officio

Overview of Planning for Exascale

- **Since 2007 US Agencies (DOE, NSF and DARPA) started planning for a *possible* Exascale Initiative**
- **The goal is to develop an exascale simulation and modeling and analysis capability by 2020**
- **Our vision includes significant international partnership**
- **The scope of the initiative concept includes**
  - **Applications Software Development**
  - **Systems Software and Programming Models**
  - **Multiple Hardware Platforms**

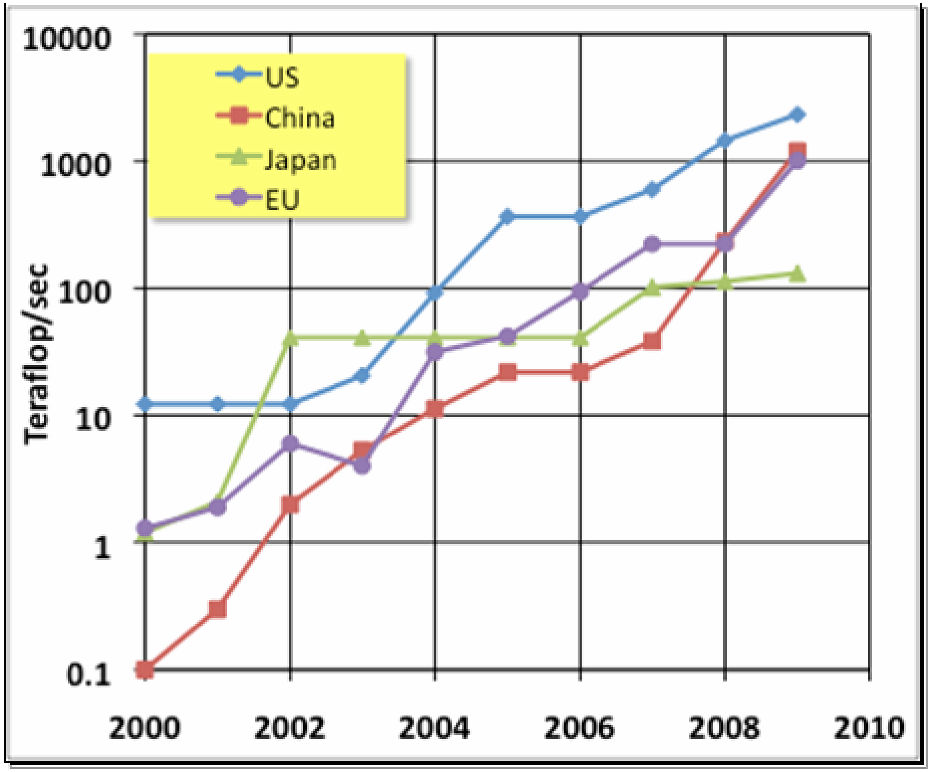
US Exascale Planning Activities

- **Began spring 2007**
- **DOE Lab Town Hall Meetings**
  - **Energy and Environment**
- **DARPA studies**
  - **Hardware and Software**
- **DOE and NSF Launch IESP**
  - **Roadmap for Software**
- **DOE Lab Steering Committee**
  - **Science Case (Applications)**
  - **Technical Roadmap**



Computational science, exascale computing & leadership in science and technology

- **The future will require certification of complex engineered systems and analysis of climate mitigation alternatives with quantified levels of uncertainty**
  - **New fuels and reactors**
  - **Stewardship without nuclear tests**
  - **Carbon sequestration alternatives**
  - **Regional climate impacts**
- **Broader application of exascale computing can provide tremendous advantages for fundamental science and industrial competitiveness**
  - **Renewable energy and energy storage**
  - **Prediction and control of materials in extreme environments**
  - **Understanding dark energy and dark matter**
  - **Clean and efficient combustion in advanced engines**



International Competition in HPC

Chart shows most capable system for each year in TOP500

Preeminence in 21st Century science, technology and engineering requires leadership in computational science and high performance computing == exascale applications & technology.

U.S. DEPARTMENT OF ENERGY

A grand challenge for the 21st century

Development of an Exascale Computing System is a Grand Challenge for the 21st Century

"[Development of] An "exascale" supercomputer capable of a million trillion calculations per second – dramatically increasing our ability to understand the world around us through simulation and slashing the time needed to design complex products such as therapeutics, advanced materials, and highly-efficient autos and aircraft."

Sept 20th 2009
EXECUTIVE OFFICE OF THE PRESIDENT
NATIONAL ECONOMIC COUNCIL OFFICE OF
SCIENCE AND TECHNOLOGY POLICY

The Exascale Draft plan has Four High-Level Components

- **Science and engineering mission applications**
- **Systems software, tools and programming models**
- **Computer hardware and technology development**
- **Systems acquisition, deployment and operations**

The plan is currently under consideration for a national initiative to begin in 2012

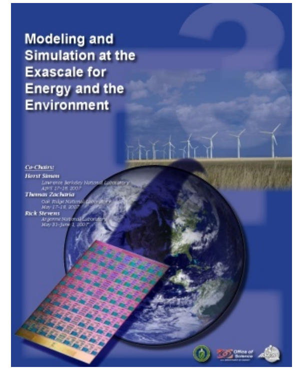
Three early funding opportunities have been release by DOE this spring to support preliminary research

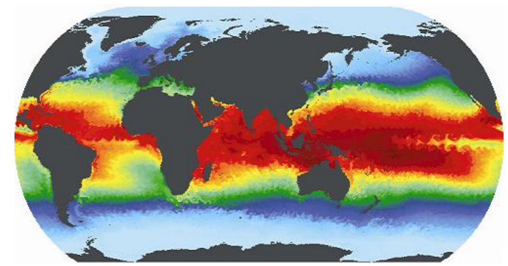
The plan targets exascale platform deliveries in 2018 and a robust simulation environment and science and mission applications by 2020

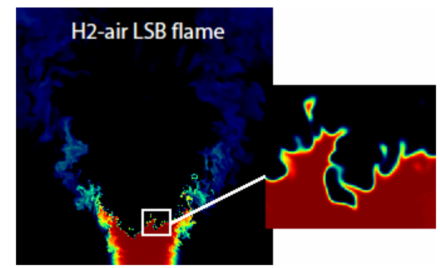
Co-design and co-development of hardware, system software, programming model and applications requires intermediate (~200 PF/s) platforms in 2015

Process Ensures Broad Community Input

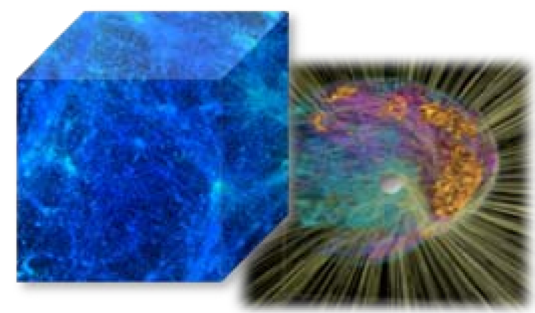
- **Town Hall Meetings April-June 2007**
- **Scientific Grand Challenges Workshops Nov, 2008 – Oct, 2009**
  - **Climate Science (11/08),**
  - **High Energy Physics (12/08),**
  - **Nuclear Physics (1/09),**
  - **Fusion Energy (3/09),**
  - **Nuclear Energy (5/09),**
  - **Biology (8/09),**
  - **Material Science and Chemistry (8/09),**
  - **National Security (10/09)**
- **Exascale Steering Committee**
  - **"Denver" vendor NDA visits 8/2009**
  - **Extreme Architecture and Technology Workshop 12/2009**
  - **Cross-cutting workshop 2/2010**
- **International Exascale Software Project**
  - **Santa Fe, NM 4/2009**
  - **Paris, France 6/2009**
  - **Tsukuba, Japan 10/2009**







MISSION IMPERATIVES



FUNDAMENTAL SCIENCE

Exascale is a Global Challenge

- Formed in 2008
- Goal to engage international computer science community to address common software challenges for Exascale
- Focus on open source systems software that would enable multiple platforms
- Shared risk and investment
- Leverage international talent base

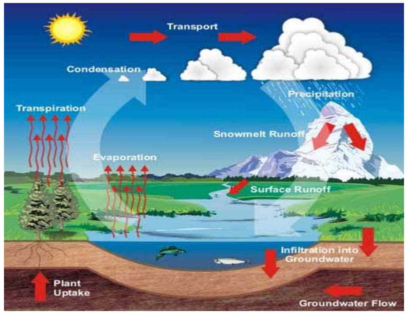
Example Science and Engineering Drivers

- Climate
- Nuclear Energy
- Combustion
- Advanced Materials
- CO_2 Sequestration
- Basic Science

- Common Needs
  - Multiscale
  - Uncertainty Quantification
  - Rare Event Statistics

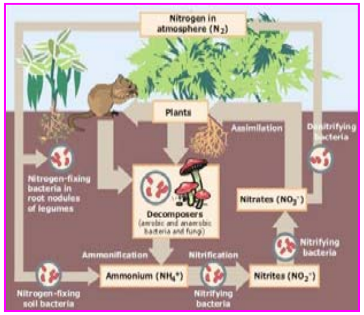
Impact of Exascale on Climate Assessments

- **Ocean and Ice**
  - **Improved modeling of land ice to reduce uncertainties in predictions for sea level rise and changes in ocean circulation.**
- **Hydrological cycle**
  - **Reduce uncertainties in prediction of cloud radiative effects and precipitation due to climate change.**
- **Extreme Weather**
  - **Use multi-scale Earth System Models (ESMs) to understand statistics of cyclones, blizzards, meso-scale storms, heat waves, droughts, and frost.**
- **Carbon, Methane and Nitrogen cycles**
  - **Use ESMs with improved chemical and biological process models to reduce uncertainties in predictions of the evolution of terrestrial ecosystems.**



Hydrological Cycle


Pine Island Glacier

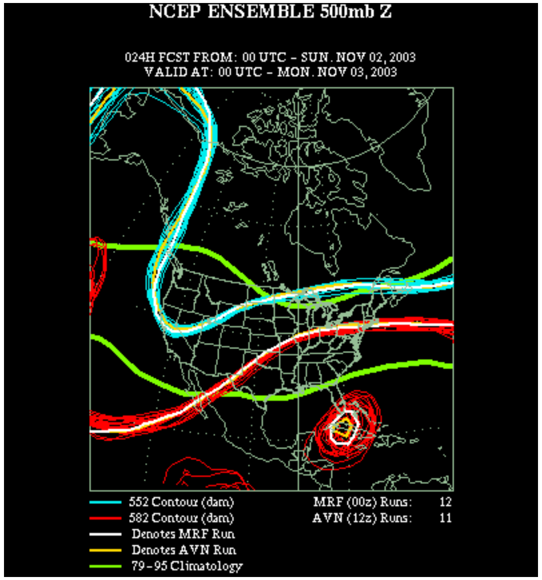


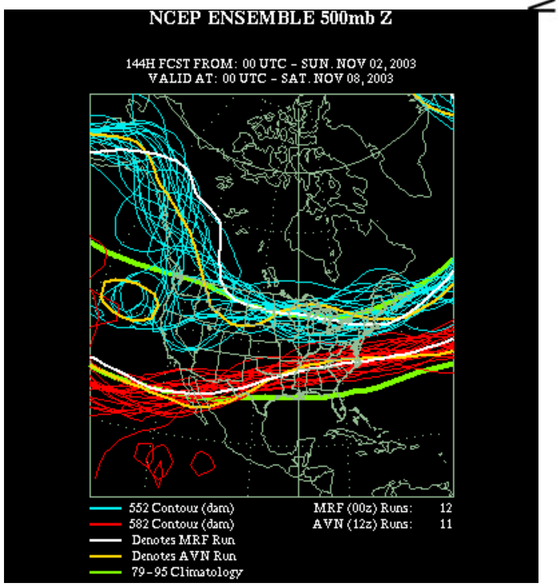
Nitrogen Cycle

Hurricane Floyd

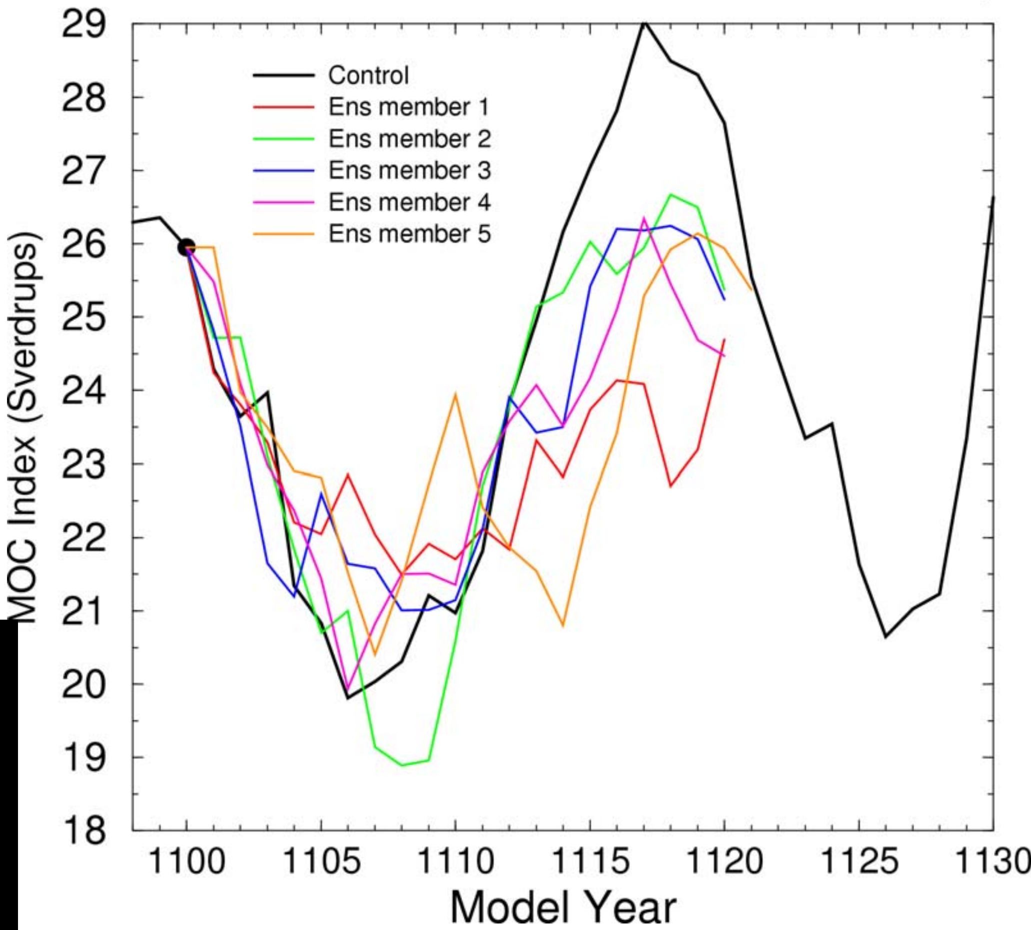
"Given these drivers … it is clear that extreme scale computers and ultra fast networks, data systems and computational infrastructure will be required by 2020."
Challenges in Climate Change Science and the Role of Computing at Extreme Scale, November, 2008

Is the Climate predictable on decadal time scales?





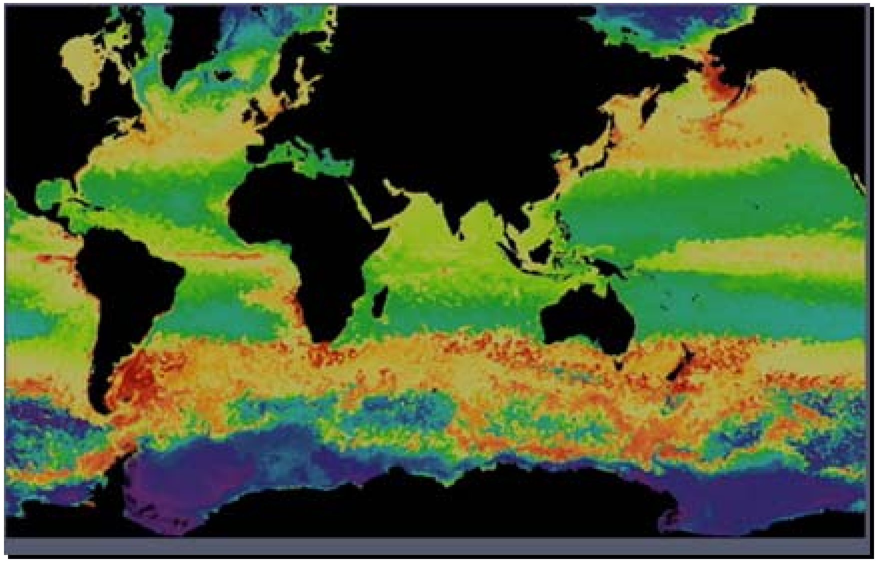
"In view of the inevitable inaccuracy and incompleteness of weather observations, precise very-long range forecasting would seem to be non-existent."
Lorentz 1963



Perturbed ensemble members evolve coherently for two decades, *Hurrel et al, 2009*

Exascale resources are required for predictive climate simulations.

- **Finer resolution**
  - **Provide regional details**
- **Higher realism, more complexity**
  - **Add "new" science**
    - **Biogeochemistry**
    - **Ice-sheets**
  - **Up-grade to "better" science**
    - **Better cloud processes**
    - **Dynamics land surface**
- **Scenario replication, ensembles**
  - **Range of model variability**
- **Time scale of simulation**
  - **Long-term implications**

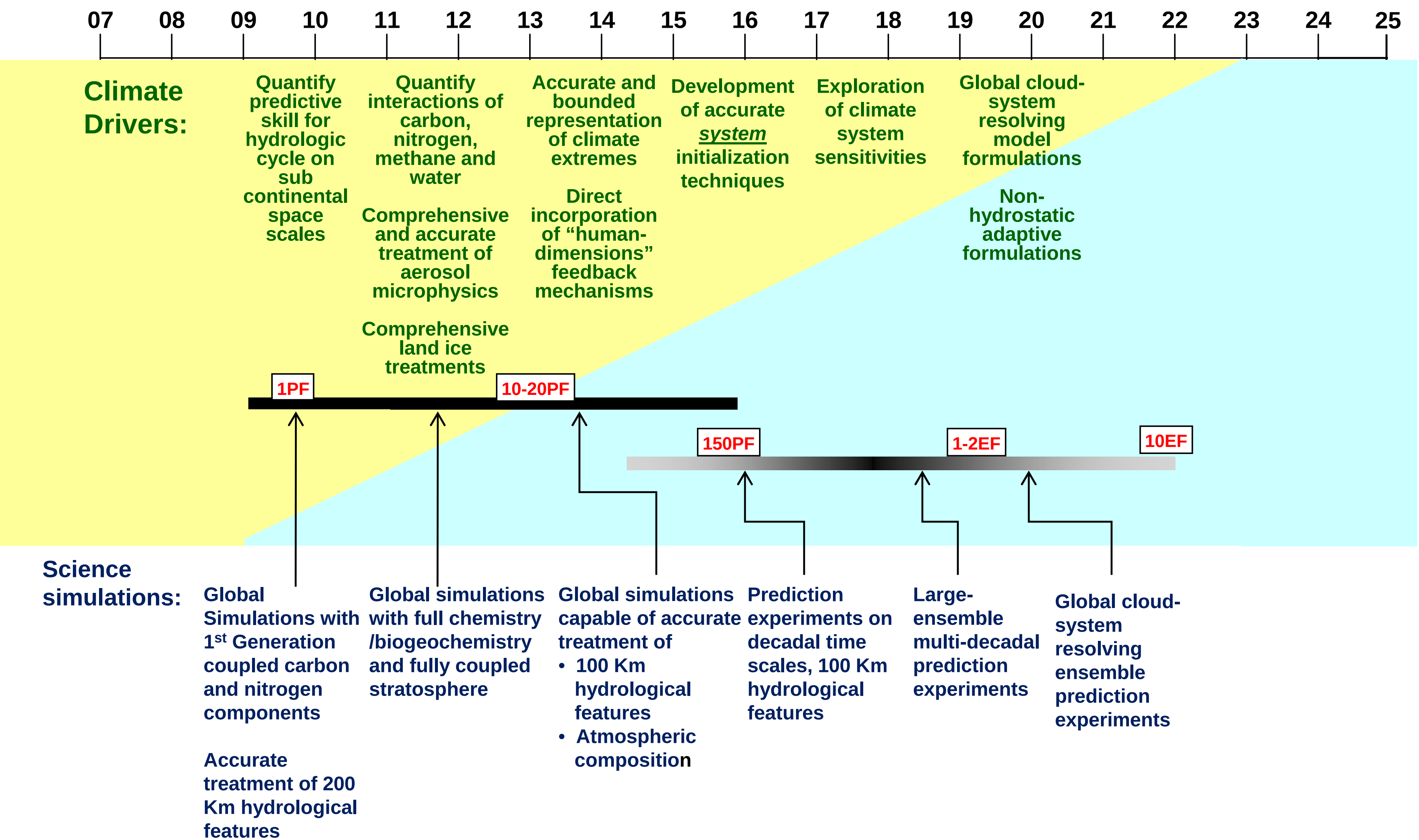


Ocean chlorophyll from an eddy-resolving simulation with ocean ecosystems included

It is essential that computing power be increased substantially (by a factor of 1000), and scientific and technical capacity be increased (by at least a factor of 10) to produce weather and climate information of sufficient skill to facilitate regional adaptations to climate variability and change.
***World Modeling Summit for Climate Prediction*, May, 2008**

Adapted from *Climate Model Development Breakout Background*
Bill Collins and Dave Bader, Co-Chairs

Computational science roadmap for a predictive, integrated climate model

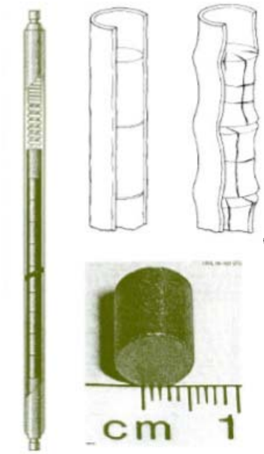


U.S. DEPARTMENT OF ENERGY

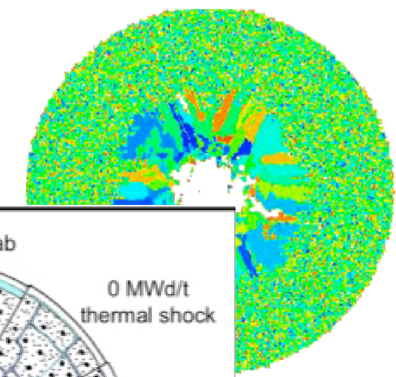
Computational modeling is a critical component of U.S. Nuclear Energy strategy

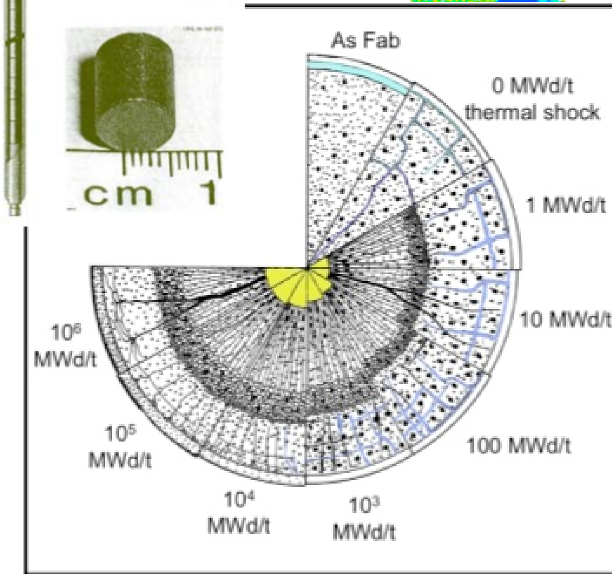
- **Improved GEOMETRIC fidelity**
  - **Sub-10μm resolution**
  - **3D phenomena, microstructure evolution, material failure**
  - **Improved lower-length scale fidelity**
- **Improved NUMERICAL fidelity**
  - **Bridging vastly different time and length scales with multi-physics phenomena**
  - **Bubble/fission fragment interactions (MD)**
  - **Upscale oxide and metal models into pellet simulations**
- **Improved PHYSICS fidelity**
  - **Fission gas bubble formation, transport, and release**
  - **Fuel chemistry and phase stability**
  - **Fuel-cladding mechanical interaction**
  - **Thermal hydraulics, turbulence, and coolant flow in pin assembly → effect on fuel and clad evolution**

Nuclear fuel example



Simulated grain structure in irradiated pellet

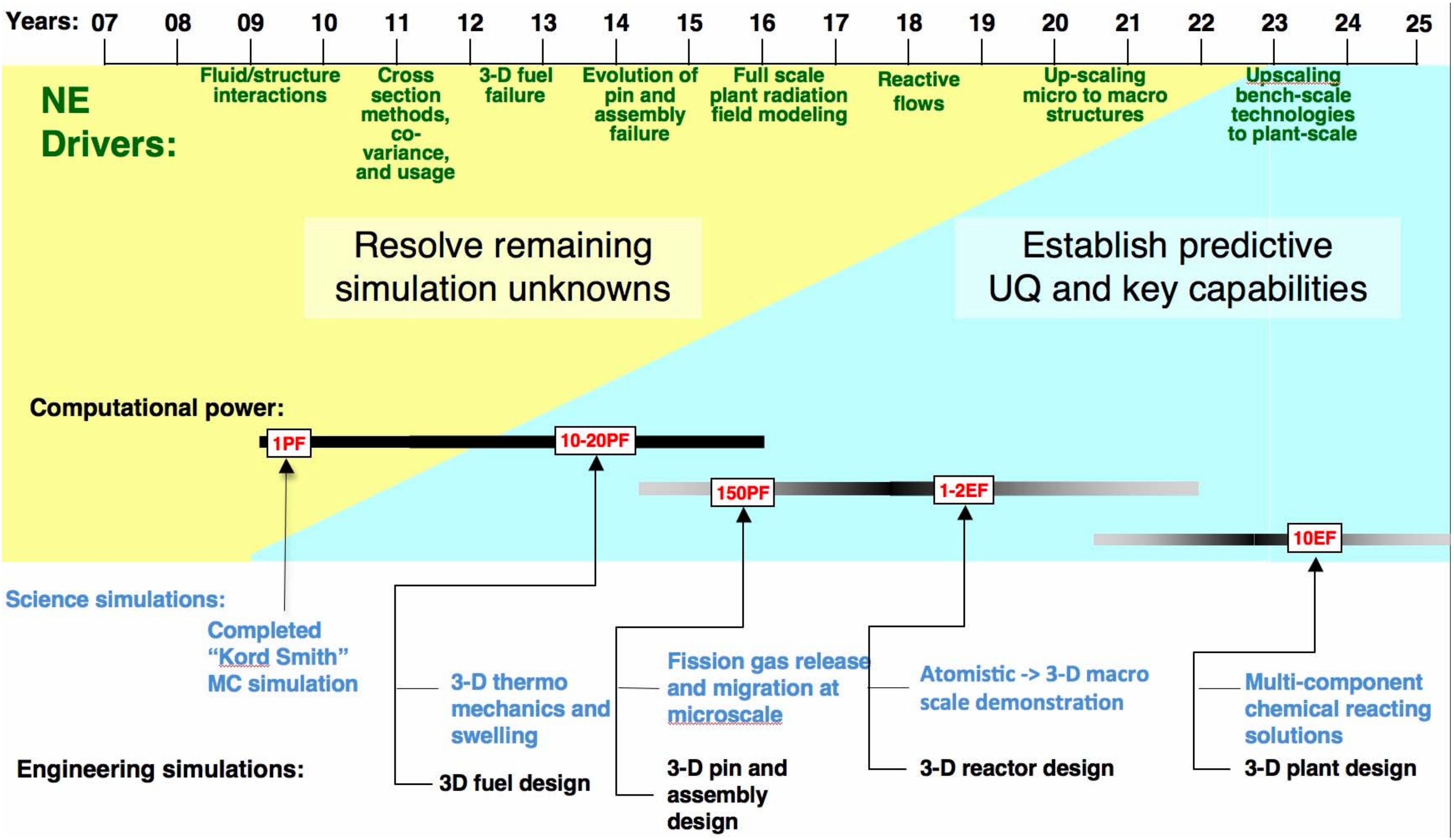




Oxide pellet micostructural evolution as a function of burnup

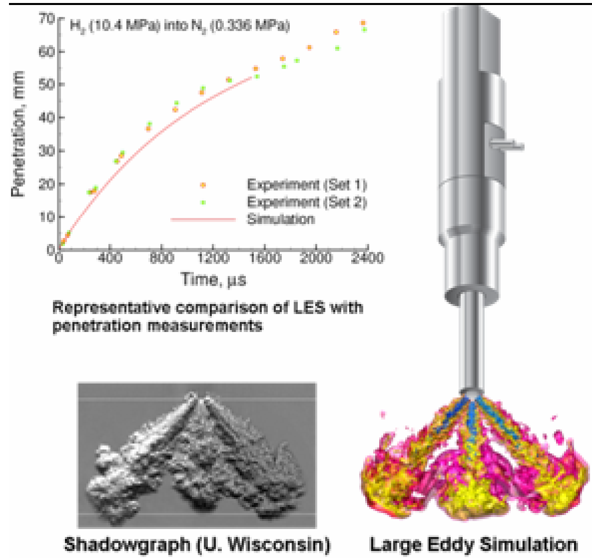
"Nuclear energy science and engineering simulations will drive the need for exaflop-scale, computing power to create robust, predictive simulations that have quantifiable uncertainties."
Science-based nuclear energy systems enabled by advanced modeling and simulation at extreme scale, May 11-12, 2009

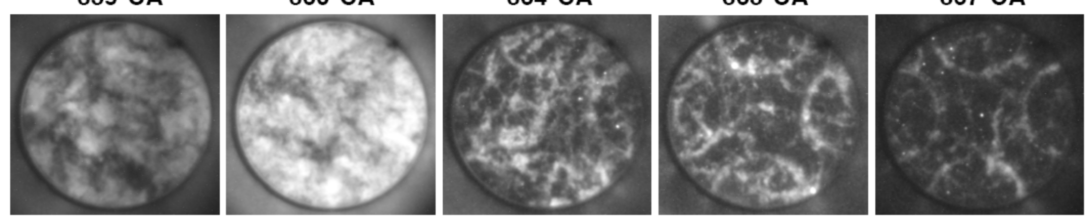
Computational science road map for predictive integrated modeling of nuclear systems



Low temperature combustion for automobile engines can provide high efficiency, low emissions

- **Design challenge:**
  - understand new combustion modes, control ignition timing, rate of pressure rise, etc.
- **Science challenge:**
  - low-temperature ignition kinetics and coupling with turbulence are poorly understood
- **Approach:**
  - perform multi-scale high-fidelity simulations to predict behavior under engine relevant conditions



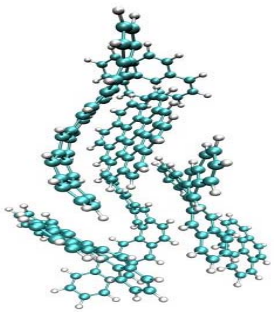


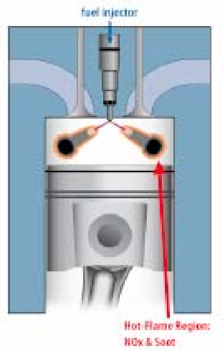
Chemiluminescence images of stratified sequential autoignition in an HCCI engine courtesy John Dec, Combustion Research Facility, Sandia National Laboratories

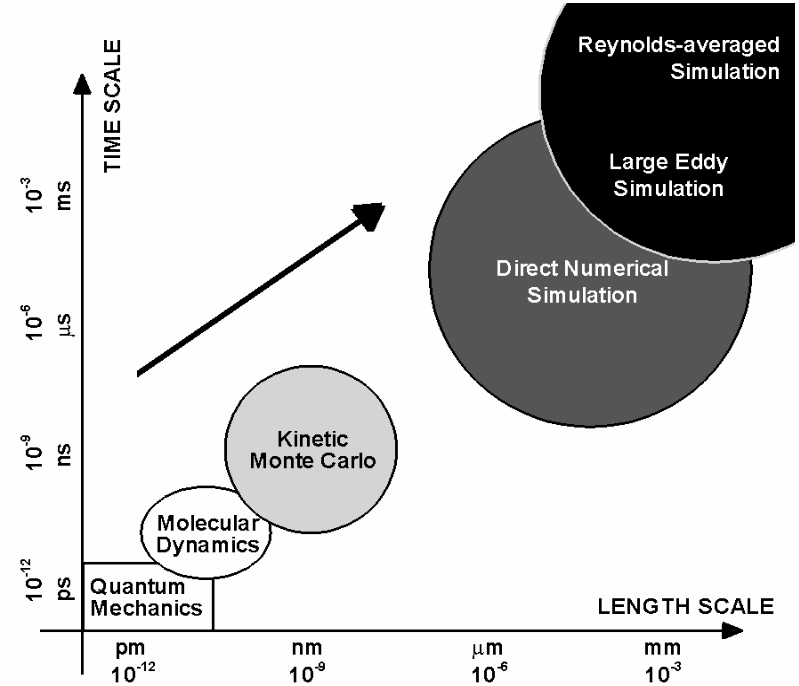
High-fidelity combustion simulations require exascale computing to predict the behavior of alternative fuels in novel fuel-efficient, clean engines, and so facilitate design of optimal combined engine-fuel system

Multi-scale modeling and simulation of combustion at exascale is key for prediction and design

- **Engine design and optimization:**
  - Design and optimization of diesel/HCCI engines, in geometry with surrogate large-molecule fuels representative of actual fuels
- **Direct numerical simulation:**
  - DNS of turbulent jet flames at engine pressures (30-50 atm) with iso-butanol (50-100 species) for diesel or HCCI engine thermochemical conditions
- **Molecular scale:**
  - Full-scale MD with on-the-fly ab-initio force field. Study combined physical and chemical processes affecting nanoparticle and soot formation processes. Size of the problem: >1000 molecules

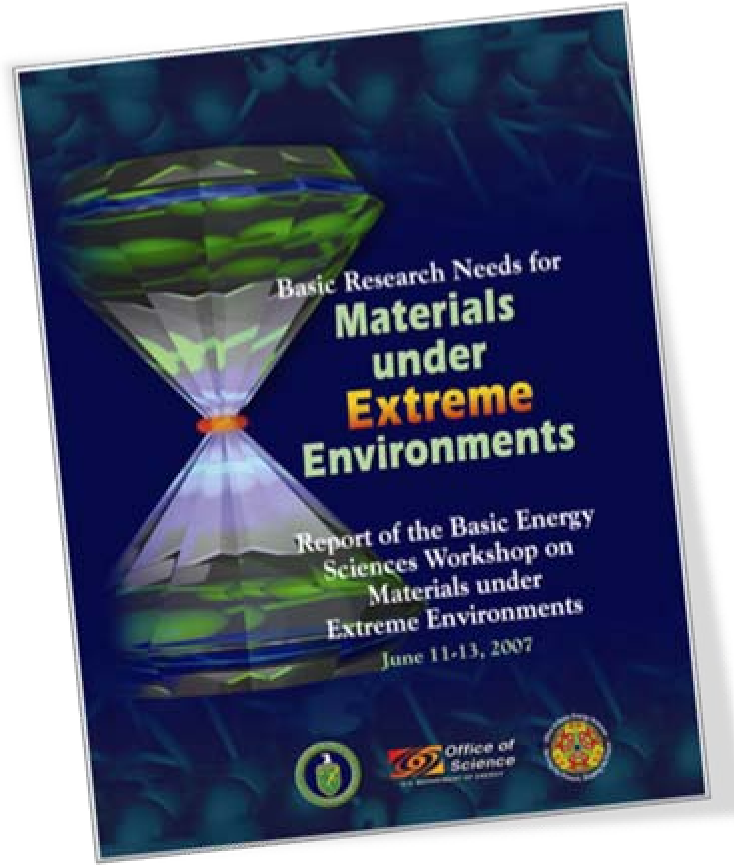






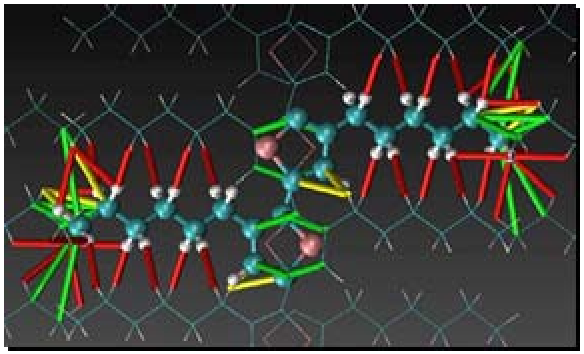
Advanced Nanoscale Materials for Energy

- Virtually all energy systems are limited by the performance of materials under extreme conditions
- Understanding and exploiting the nanoscale is essential to progress in materials
  - Length scale where properties are determined
- Robust simulation and modeling of nanoparticles is within our grasp
- Exascale models offer the potential of predictive design of nanoscale systems
  - Transformational for the development of new materials with desired properties

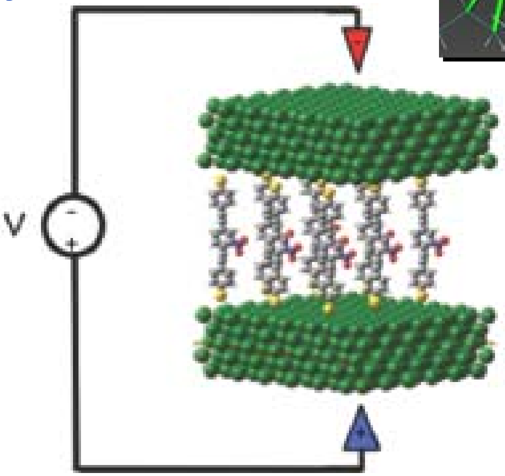


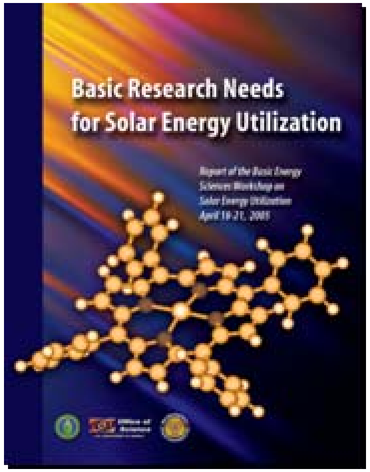
Multi-scale materials design is critical for improving efficiency and cost of photovoltaics.

- **Scientific and engineering challenges**
  - **Prediction of excited states: energies, forces, band-gaps, level alignment**
  - **Understanding and control of carrier generation, recombination and transport**
  - **Understanding and control of defect interactions and migration**
  - **Simulation of interface formation and dynamics**
- **Computational science challenges**
  - **Multi-scale methods**
  - **Increased spatial and time scale simul**
  - **ab initio MD methods**
  - **Electronic structure methods for excite**



From *Photovoltaic Fundamentals Subpane*l, BES workshop 8/2009

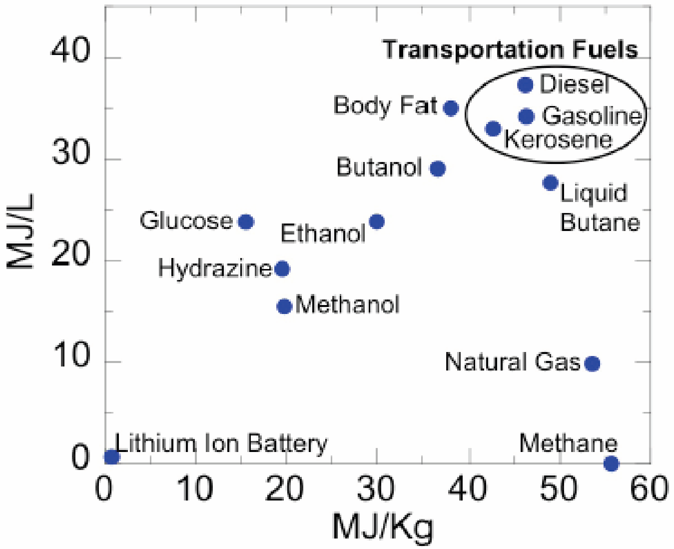


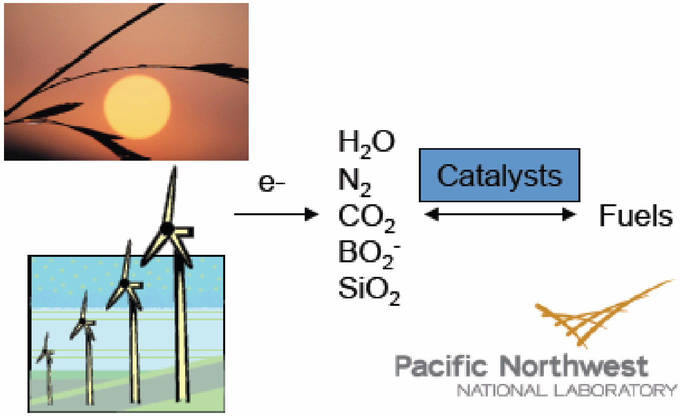


"Extreme-scale computation can create a breakthrough in the accurate prediction of the combined effects of multiple interacting structures and their effects upon electronic excitation and ultimate performance of a PV cell."
Discovery in Basic Energy Sciences: The Role of Computing at the Extreme Scale (Workshop Letter Report)

Full quantum simulations are required for understanding and deploying better catalysts

- **Efficient storage of energy requires chemical bonds**
- **Catalysts are crucial for inter-conversion of electrical and chemical energy**
- **Development of new catalysts requires full quantum simulations**
- **Now: 1000 atoms, 3 nm**
- **Needed: 10,000 atoms, 6 nm**
  - **300,000 basis functions**
  - **> 10^{20} integrals**
- **Achievable with Exascale**
- **Foundation for multi-scale models**





BES/ASCR Extreme scale workshop

Geologic Carbon Sequestration

Single Injection Site
- Moderate-resolution flow + heat transport
- Low-resolution flow + heat transport + geochemistry + geomechanics

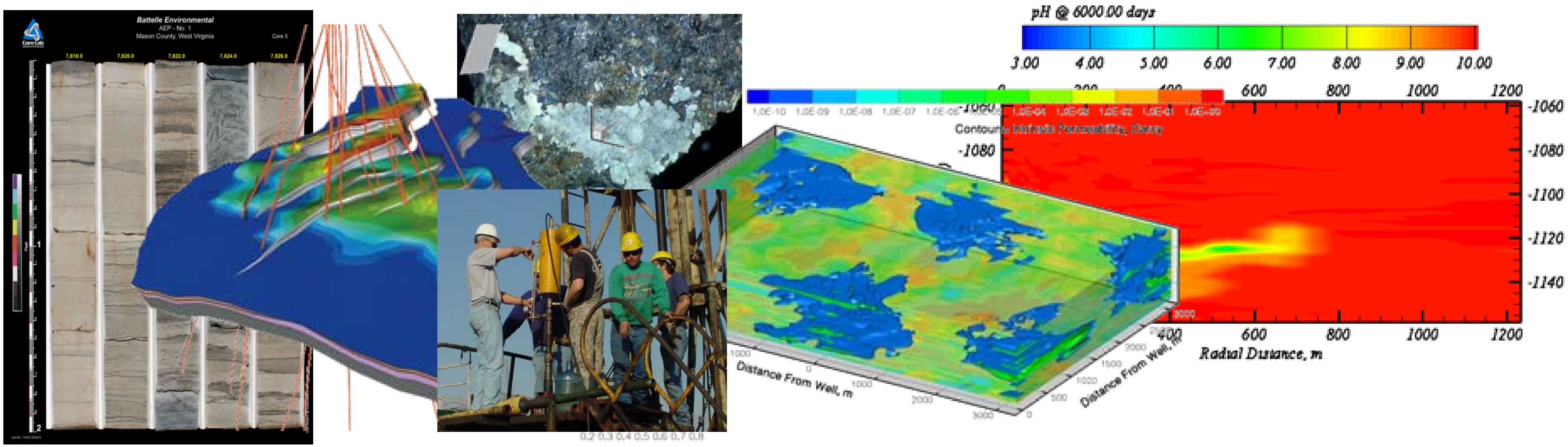
Basin-Scale Injection
- Moderate-resolution flow + heat transport
- Low-resolution flow + heat transport + geochemistry + geomechanics

Basin-Scale Injection
- High-resolution flow + heat transport
- Moderate-resolution flow + heat transport + geochemistry + geomechanics

Multi-Scale Models
- Coupled pore-scale and continuum-scale modeling
- Coupled subsurface and atmospheric modeling

Gigascale — Terascale — Petascale — Exascale

Increasing grid resolution and coupled process complexity

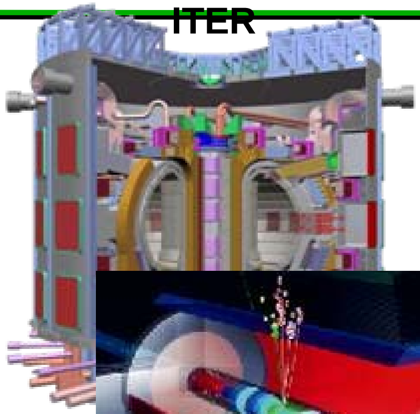


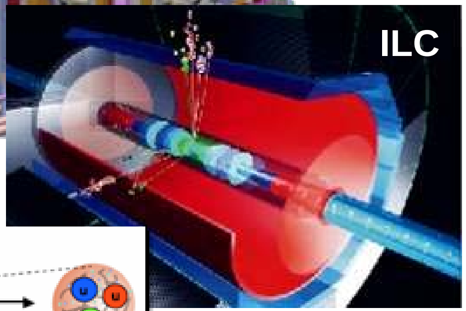
Simulation enables fundamental advances in basic science.

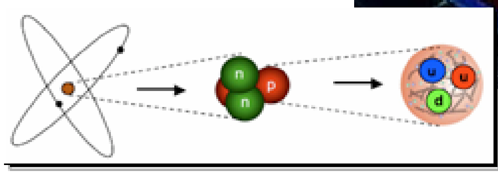
- **Nuclear Physics**
  - **Quark-gluon plasma & nucleon structure**
  - **Fundamentals of fission and fusion reactions**
- **Facility and experimental design**
  - **Effective design of accelerators**
  - **Probes of dark energy and dark matter**
  - **ITER shot planning and device control**
- **Materials / Chemistry**
  - **Predictive multi-scale materials modeling: observation to control**
  - **Effective, commercial, renewable energy technologies, catalysts and batteries**
- **Life Sciences**
  - **Better biofuels**
  - **Sequence *to* structure *to* function**
  - **Genotype to Phenotype**

These breakthrough scientific discoveries and facilities require exascale applications and resources.

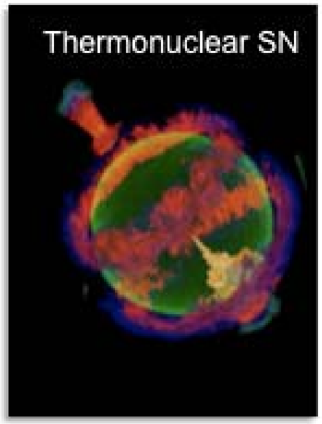
Hubble image of lensing





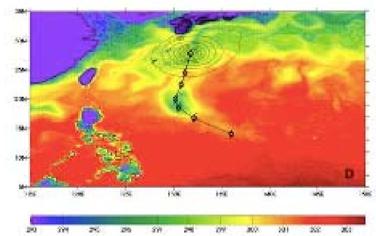


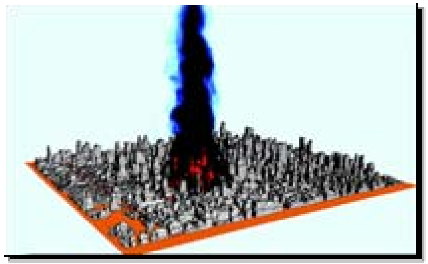
Structure of nucleons

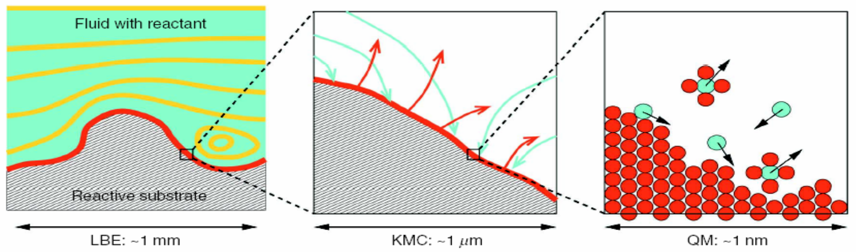


Cross-cutting capabilities are critical to success for DOE mission and science

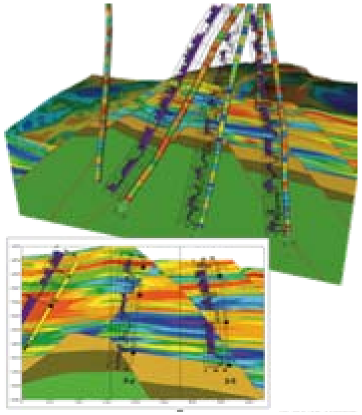
- **Uncertainty quantification**
  - **Predict climate response to energy technology strategies**
  - **Assessment of safety, surety and performance of the aging/evolving stockpile without nuclear testing**
  - **Energy security**
  - **Responding to natural and manmade hazards**
- **Multi-scale, multi-physics modeling**
  - **Multiple physics packages in earth system model: ocean, land surface, atmosphere, ice**
  - **Multiple physics packages in modeling reactor core: neutronics, heat transfer, structures, fluids**
- **Statistics of rare events**
  - **Severe weather and surprises in climate system**
  - **Accident scenarios in nuclear energy**
  - **Nucleation of cracks and damage in materials**


Storm track in ESM


urban fire



Multi-scale modeling in fast sodium reactor


Geologic sequestration

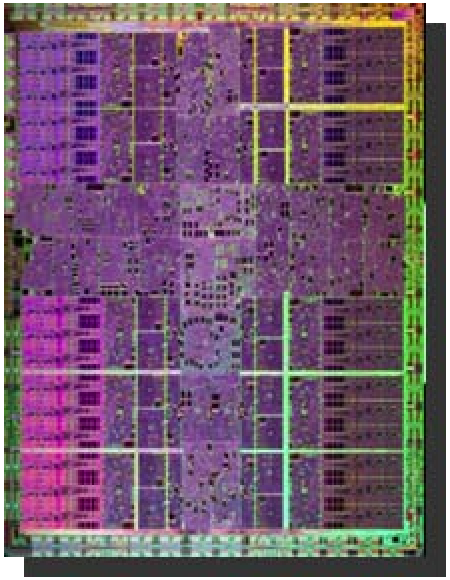
What are critical exascale technology investments?

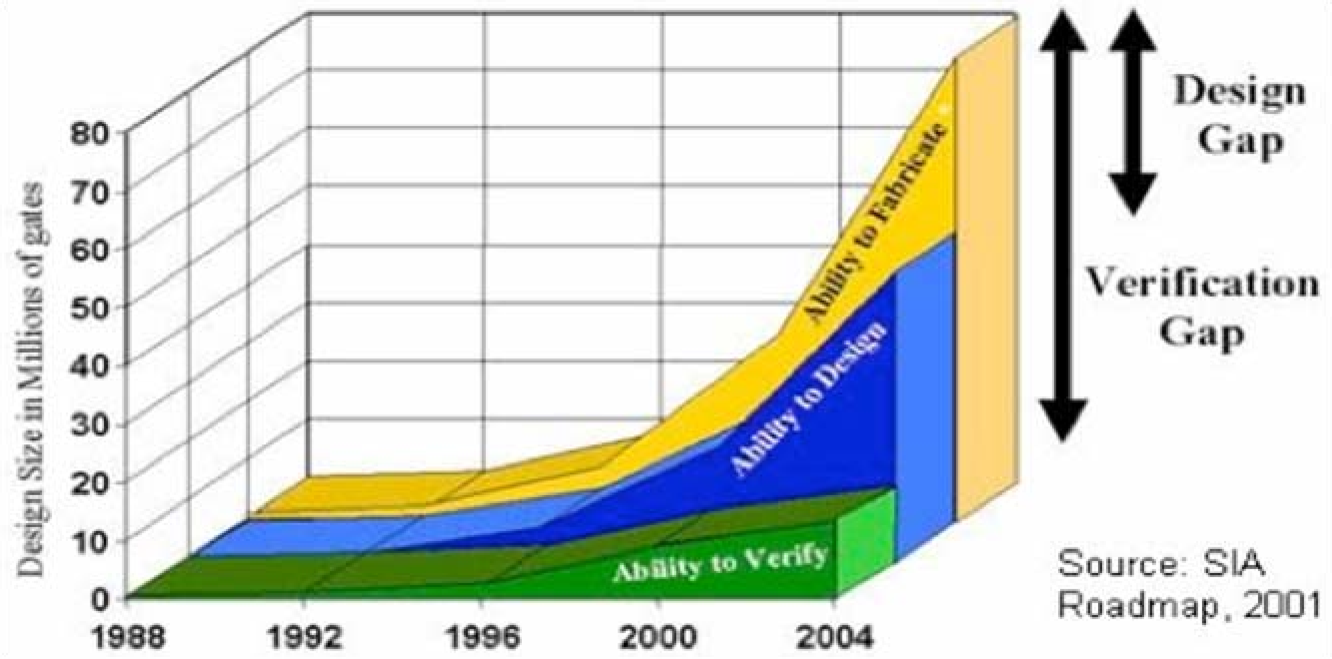
- **System power** is an overall constraint on exascale system performance and effectiveness.
- M**emory** is an important component of meeting exascale power, memory and memory bandwidth goals.
- **Programming model**. Early investment in several efforts (motifs and prototype systems) to decide in 2013 on exascale programming model. Early investment necessary to allow exemplar applications effective access to 2015 system for both science and technology.
- **Exascale processor design** to achieve an exascale-like system in 2015.
- **Operating System strategy for exascale** is critical for node performance at scale and for efficient support of new programming models and run time systems.
- **Reliability and resiliency are critical at this** scale and require applications neutral movement of the file system (for check pointing, in particular) closer to the running apps.
- **HPC co-design strategy and implementation** requires a set of a hierarchical performance models and simulators as well as commitment from apps, software and architecture communities.

Exascale Systems Targets

Systems	2009	2018	Difference Today & 2018
System peak	**2 Pflop/s**	**1 Eflop/s**	**O(1000)**
Power	**6 MW**	**~20 MW (goal)**	
System memory	0.3 PB	32 - 64 PB	O(100)
Node performance	125 GF	1.2 or 15TF	O(10) – O(100)
Node memory BW	25 GB/s	2 - 4TB/s	O(100)
Node concurrency	12	O(1k) or O(10k)	O(100) – O(1000)
Total Node Interconnect BW	3.5 GB/s	200-400GB/s (1:4 or 1:8 from memory BW)	O(100)
System size (nodes)	18,700	O(100,000) or O(1M)	O(10) – O(100)
Total concurrency	225,000	O(billion) + [O(10) to O(100) for latency hiding]	O(10,000)
Storage Capacity	15 PB	500-1000 PB (>10x system memory is min)	O(10) – O(100)
IO Rates	0.2 TB	60 TB/s	O(100)
MTTI	days	O(1 day)	- O(10)

Many-core chip architectures are the future.





The shift toward increasing parallelism is not a triumphant stride forward based on breakthroughs in novel software and architectures for parallelism … instead it is actually a retreat from even greater challenges that thwart efficient silicon implementation of traditional uniprocessor architectures.
Kurt Keutzer

Reducing power is fundamentally about architecture choices & process technology

- **Memory (2x-5x)**
  - New memory interfaces (chip stacking and vias)
  - Replace DRAM with zero power non-volatile memory
- **Processor (10x-20x)**
  - Reducing data movement (functional reorganization, > 20x)
  - Domain/Core power gating and aggressive voltage scaling
- **Interconnect (2x-5x)**
  - More interconnect on package
  - Replace long haul copper with integrated optics
- **Data Center Energy Efficiencies (10%-20%)**
  - Higher operating temperature tolerance
  - Power supply and cooling efficiencies

Reliability and Resilience

- **Barriers**
  - Number of system components increasing faster than overall reliability
  - Silent error rates increasing
  - Reduced job progress due to fault recovery if we use existing checkpoint/restart
- **Technical Focus Areas**
  - Local recovery and migration
  - Development of a standard fault model and better understanding of types/rates of faults
  - Improved hardware and software reliability
    - Greater integration across entire stack
  - Fault resilient algorithms and applications
- **Technical Gap**
  - Maintaining today's MTTI given 10x - 100X increase in sockets will require:

10X improvement in hardware reliability

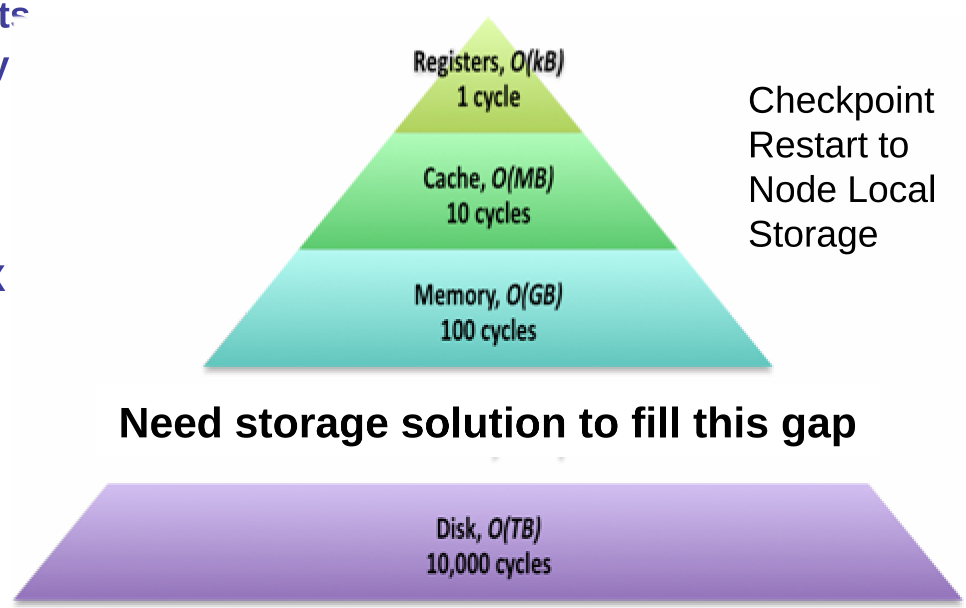
10X in system software reliability, and

10X improvement due to local recovery and migration as well as research in fault resilient applications

- .

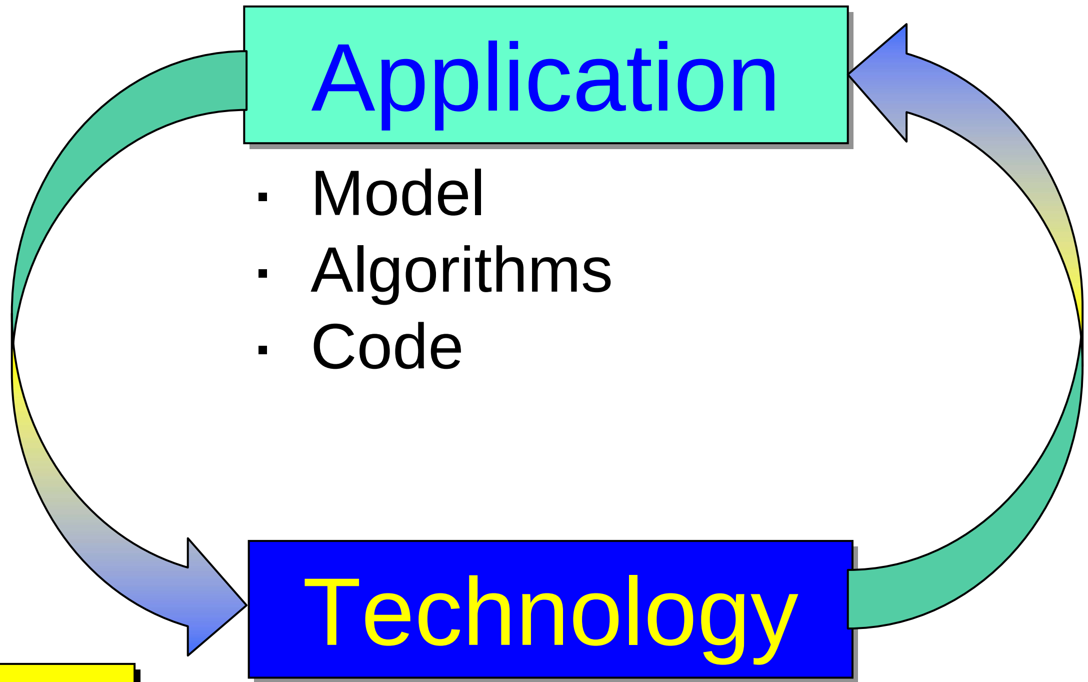
<u>Taxonomy of errors (h/w or s/w)</u>

- **Hard errors**: permanent errors which cause system to hang or crash
- **Soft errors**: transient errors, either correctable or short term failure
- **Silent errors**: undetected errors either permanent or transient. *Concern is that simulation data or calculation have been corrupted and no error reported.*



Co-design is a fundamental tenet of the initiative.

<u>Application driven:</u>
Find the best technology to run this code.
Sub-optimal



- Model
- Algorithms
- Code

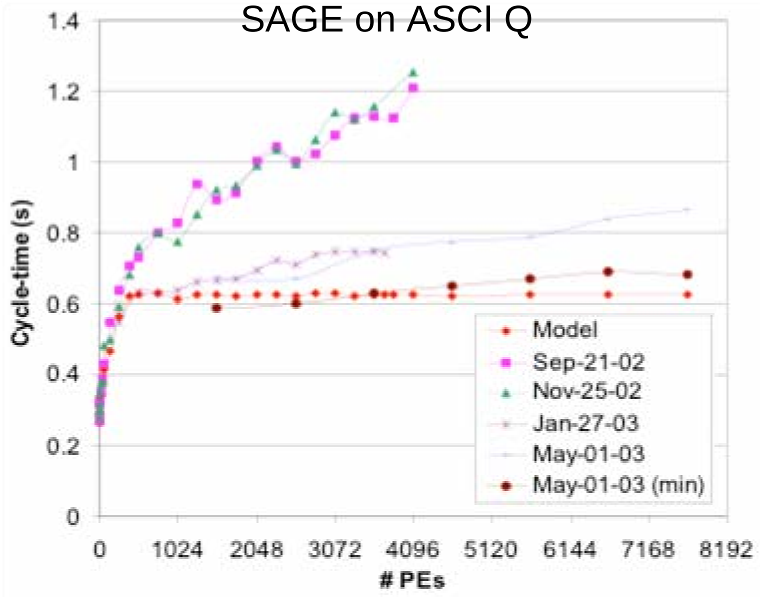
- ⊕ architecture
- ⊕ programming model
- ⊕ resilience
- ⊕ power

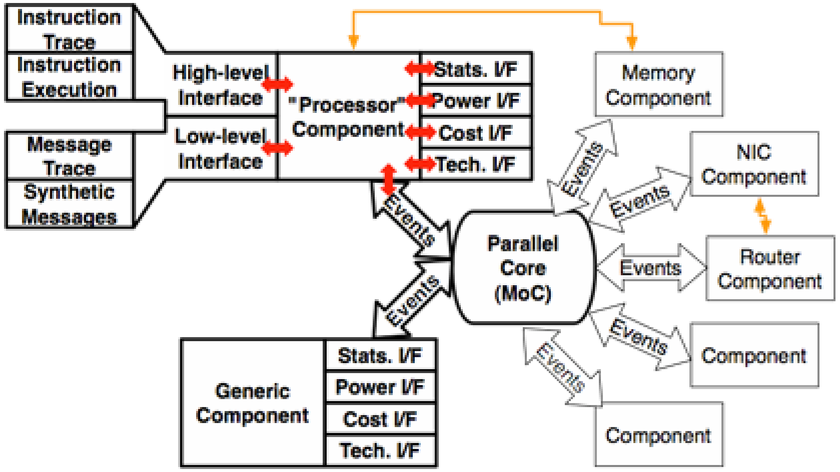
<u>Technology driven:</u>
Fit your application to this technology.
Sub-optimal.

Now, we must expand the co-design space to find better solutions:
•new applications & algorithms,
•better technology and performance.

Hierarchical {system, application} co-simulation a the key for co-design

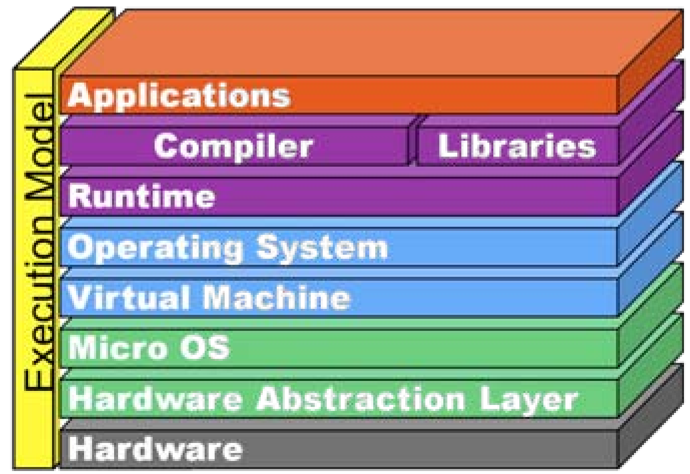
- **Hierarchical co-simulation capability**
  - Discussions between architecture, software and application groups
  - System level simulation based on analytic models
  - Detailed (e.g. cycle accurate) co-simulation of hardware and applications
- **Opportunity to influence future architectures**
  - Cores/node, threads/core, ALUs/thread
  - Logic layer in stacked memory
  - Interconnect performance
  - Memory/core
  - Processor functionality
- **Current community efforts must work together to provide a complete co-design capability**

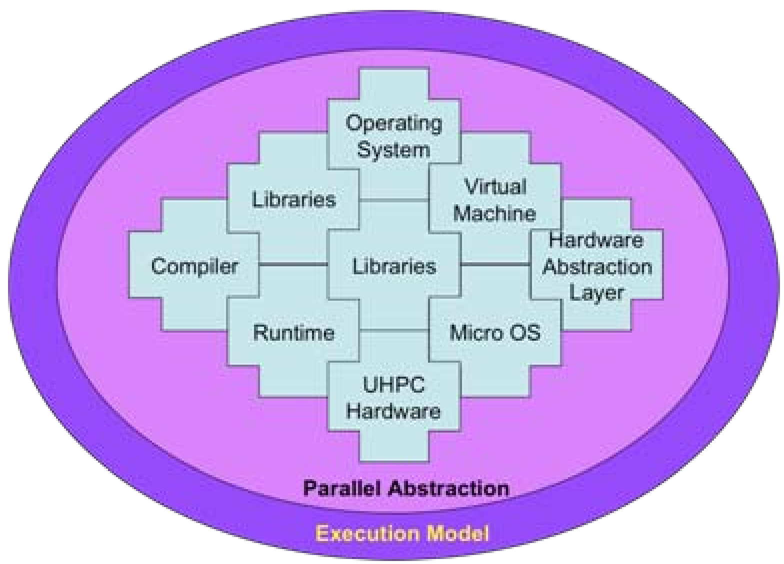




System software as currently implemented is not suitable for exascale system.

- **Barriers**
  - **System management SW not parallel**
  - **Current OS stack designed to manage only O(10) cores on node**
  - **Unprepared for industry shift to NVRAM**
  - **OS management of I/O has hit a wall**
  - **Not prepared for massive concurrency**
- **Technical Focus Areas**
  - **Design HPC OS to partition and manage node resources to support massively concurrency**
  - **I/O system to support on-chip NVRAM**
  - **Co-design messaging system with new hardware to achieve required message rates**
- **Technical gaps**
  - **10X: in affordable I/O rates**
  - **10X: in on-node message injection rates**
  - **100X: in concurrency of on-chip messaging hardware/software**
  - **10X: in OS resource management**

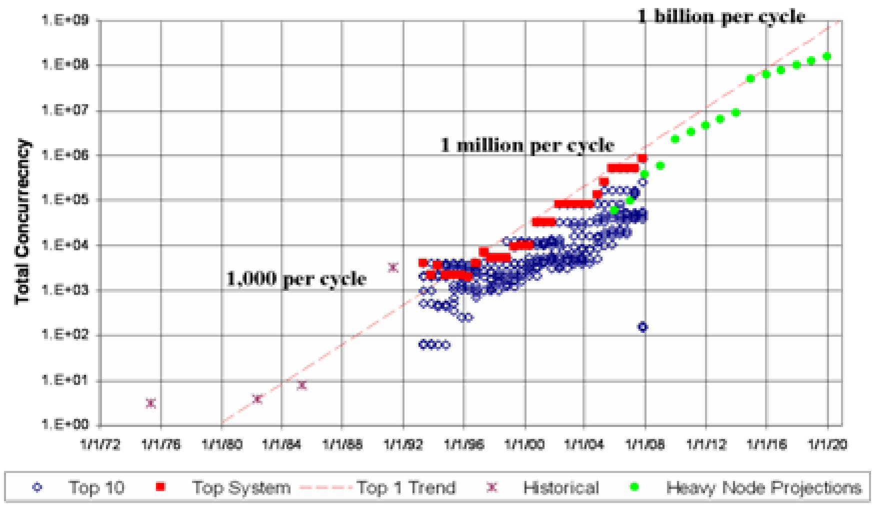




Software challenges in extreme scale systems, *Sarkar, 2010*

Programming models and environments require early investment.

- **Barriers:** **Delivering a large-scale scientific instrument that is productive and fast.**
  - O(1B) way parallelism in Exascale system
  - O(1K) way parallelism in a processor chip
    - Massive lightweight cores for low power
    - Some "full-feature" cores lead to heterogeneity
  - Data movement costs power and time
    - Software-managed memory (local store)
  - Programming for resilience
  - Science goals require complex codes



How much parallelism must be handled by the program?
From Peter Kogge (on behalf of Exascale Working Group), "Architectural *Challenges* at the Exascale Frontier", June 20, 2008

- **Technology Investments**
  - Extend inter-node models for scalability and resilience, e.g., MPI, PGAS (includes HPCS)
  - Develop intra-node models for concurrency, hierarchy, and heterogeneity by adapting current scientific ones (e.g., OpenMP) or leveraging from other domains (e.g., CUDA, OpenCL)
  - Develop common low level runtime for portability and to enable higher level models
- **Technical Gap:**
  - No portable model for variety of on-chip parallelism methods or new memory hierarchies
  - Goal: Hundreds of applications on the Exascale architecture; Tens running at scale

Summary

- **Critical areas of 21st Century Science and Engineering** will be rate limited by improvements in modeling and simulation capabilities
- **Moore's Law is good** for another decade at least, however all improvements in performance will come from increases in concurrency (scale)
- **All types of computing systems** will be impacted by the need to move to scalable many-core processors to improve performance and reduce power
- **Significant investments** will be required to reduce power consumption of future systems and to develop the highly scalable software needed
- **Exascale (10^{18} Operations per second) computer systems** appear to be feasible by the end of the decade if not earlier but significant challenges need be overcome
- **While all computing systems** will need to move to many-core technologies, Exascale systems will have to solve additional problems of extreme scalability, reliability and power
- DOE has historically led the world in the deployment and utilization of HPC for science and engineering, now it is critical that it take on the role of **lead agency for development of HPC technologies**
- The opportunities and challenges are so significant that **international collaboration** and cooperation will be needed to overcome them