WORKSHOP SERIES ON SCIENCE GRAND CHALLENGES ENABLED BY EXTREME SCALE COMPUTING

IESP Workshop

Saclay, France

Paul Messina June 28, 2009

Outline

- Purpose of workshops
- General structure, process, outputs
- List of all “Scientific Challenges” workshops
- Examples of findings from the workshops
- Related and follow-on workshops

Purpose of workshops

- To identify grand challenge scientific problems in [research area] that can exploit computing at extreme scales to bring about dramatic progress toward their resolution.
- The goals of the workshops are to
  - identify grand challenge scientific problems […] that could be aided by computing at the extreme scale over the next decade;
  - identify associated specifics of how and why new high performance computing capability will address issues at the frontiers of […]; and
  - provide a forum for exchange of ideas among application scientists, computer scientists, and applied mathematicians to maximize the use of extreme scale computing for enabling advances and discovery in […].

Science Workshop Series

- Climate, November 6-7, 2008
- HEP, December 9-11, 2008
- Nuclear Physics, January 26-28, 2009
- Fusion Energy Sciences, March 18-20, 2009
- Nuclear Energy, May 11-12, 2009
- BES, August 13-15, 2009
- Biology, August 17-19, 2009
- NNSA, October 6-8, 2009

Process used

- Workshops are organized jointly by US DOE's office of Advanced Scientific Computing Research and other DOE program offices
- Workshop chair(s) work with relevant DOE program offices and colleagues to identify key areas to cover
- Four – six panels defined, panel co-chairs recruited
- White papers for each panel drafted and posted in advance of workshop
- Priority Research Directions (PRDs) identified by each panel
- Panels populated by domain science experts as well as mathematicians and computer scientists, including some international
- Observers from other agencies and math and CS community invited to each workshop, including some international
- Panels report in plenary sessions in middle and at end of workshop

Priority Research Direction (use one slide for each)

Scientific and computational challenges

Brief overview of the underlying scientific and computational challenges

Summary of research direction

What will you do to address the challenges?

Potential scientific impact

What new scientific discoveries will result?

What new methods and techniques will be developed?

Potential impact on SCIENCE DOMAIN

How will this impact key open issues in SCIENCE DOMAIN?

What's the timescale in which that impact may be felt?

This summary is preliminary

- Five workshops have been held
- Only one report is in final form: climate
  - http://extremecomputing.labworks.org/climate/report.stm
  - Several are nearly finished
- All have identified many important problems that would be tractable with exascale system
- Characterization of software and hardware needs is understandably fairly general
  - But future workshops are planned to do a more thorough analysis

Challenges in Climate Change Science and the Role of Computing at the Extreme Scale

November 6-7, 2008, DC

- Chair: Warren Washington
- Panel topics
  - Model Development and Integrated Assessment
  - Algorithms and Computational Environment
  - Data, Visualization and Productivity
  - Decadal Predictability and Prediction

Climate Workshop Goals

- Review and identify the critical scientific challenges.
- Prioritize the challenges in terms of decadal or annual timelines.
- Identify the challenges where computing at the extreme scales is critical for climate change science success within the next two decades.
- Engage international scientific leaders in discussing opportunities to shape the nature of extreme scale scientific computing.
- Provide the high performance computing community with an opportunity to understand the potential future needs of the climate change research community.
- Look for breakthroughs.

Priority Research Directions (PRDs) were established for each of the Breakout sessions

Some PRDs are highlighted as follows:

PRDs for Model Development and Integrated Assessment

- How do the carbon, methane, and nitrogen cycles interact with climate change?
- How will local and regional water, ice, and clouds change with global warming?
- How will the distribution of weather events, particularly extreme events, that determine regional climate change with global warming?
- What are the future sea level and ocean circulation changes?

PRDs for Algorithms and Computational Environment

- Develop numerical algorithms to efficiently use upcoming petascale and exascale architectures
- Form international consortium for parallel input/output, metadata, analysis, and modeling tools for regional and decadal multimodel ensembles
- Develop multicore and deep memory languages to support parallel software infrastructure
- Train scientists in the use of high-performance computers.

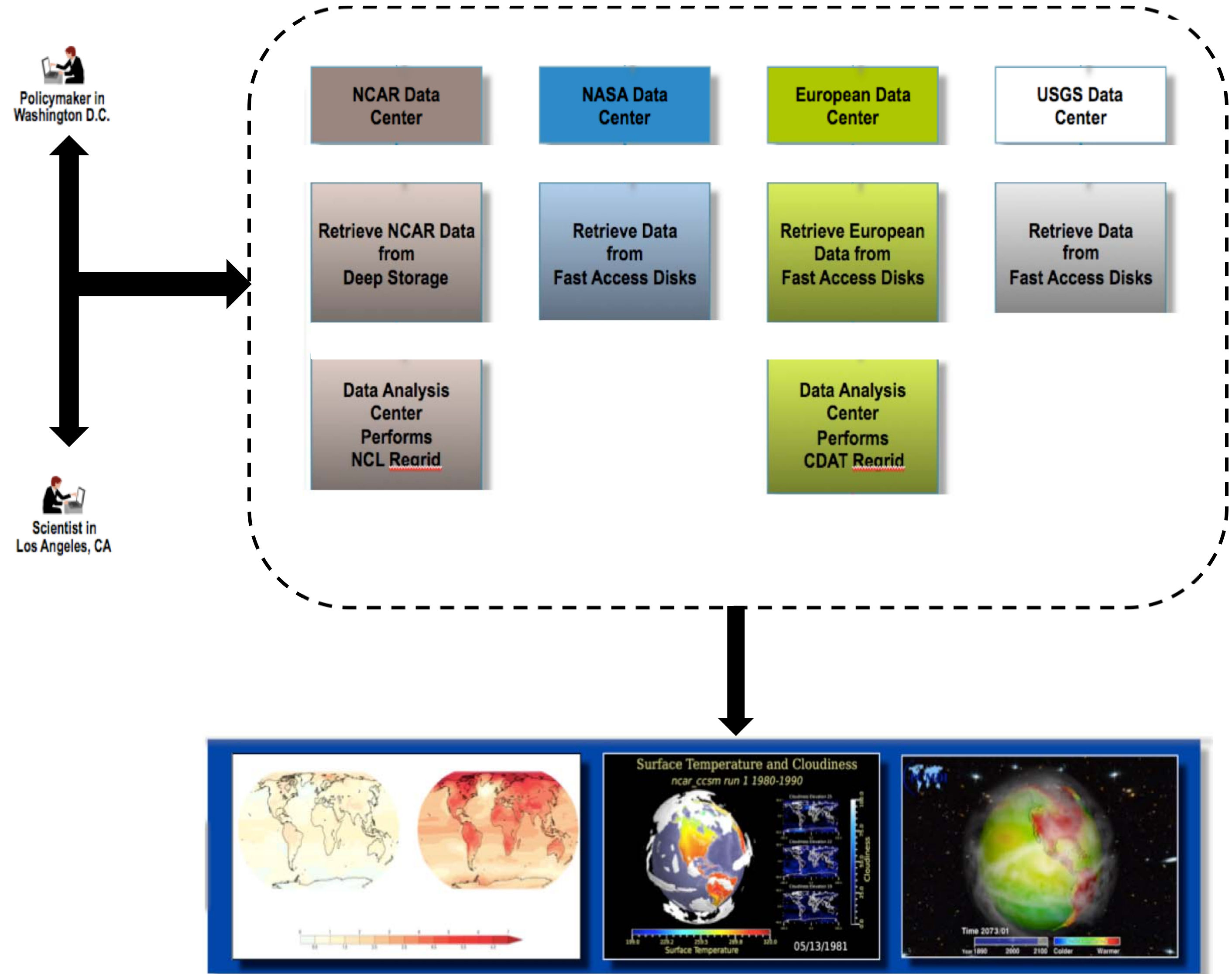
Scientific and Computational Challenges

- Increasing throughput of Earth System Models (ESMs), especially in the 10- to 50-km regime
- Building global cloud-resolving models
- Improving both the fidelity and computational performance of multiphysics and multiscale model coupling approaches
- Improving the speed at which new components can be integrated and evaluated in ESMs
- Developing efficient, scalable data-assimilation approaches for decadal climate predictability work
- Building and running global cloud-resolving climate models at the necessary throughput rates would introduce the largest algorithmic challenges

PRDs for Data Visualization and Computing Productivity

- Develop new, robust techniques for dealing with the input/output, storage, processing, and wide-area transport demands of exascale data
- Integrate diverse and complex data
- Dedicate resources to the development of standards, conventions, and policies, and contribute to related committees

Diverse and complex data are integrated into visualizations to communicate model predictions



Paul Messina June 28, 2009

I/O and Interoperability

- Enhance climate models to exploit parallel input/output and analysis tools.
  - Without an immediate focus on enhancing these tools, it would be difficult to perform global cloud-resolving simulations in the next 2 to 5 years.
- Make new tools accessible to international collaborators.
- Encourage international cooperation between computer hardware and data storage vendors, modeling centers, and research institutes.
  - We recommend that these efforts be catalyzed in the form of an international consortium for parallel input/output, metadata, analysis, and modeling tools.

I/O and Interoperability

- Improve input/output performance
  - to support the dramatic increase in data generated by ensembles of long time simulations required for studying extreme events and the large multimodel ensembles needed for decadal predictability efforts.
- Focus on improving workflow issues.
  - These issues are especially important in the realm of decadal predictability, where many collaborators, including international collaborations, need the ability to analyze petabytes of data distributed at modeling centers worldwide.

I/O and Interoperability

- International collaborations would ensure common metadata and many other requirements to allow analysis of ensembles of multiple models from multiple international efforts.
- Scalable input/output and analysis tools for large distributed data sets are a near-term limiter.
  - Research in this area would enable regional climate change information from ensembles of high-resolution simulations and helping answer outstanding questions of deep-water formation, the role of land-use change on the hydrologic cycle, and decadal prediction.

Scientific Challenges for Understanding the Quantum Universe and the Role of Computing at Extreme Scale (HEP)

- Workshop chair: Roger Blandford, co-chairs: Norman Christ, Young-Kee Kim
- Panel topics
  - Astrophysics data
  - Cosmology and astrophysics simulations
  - Experimental particle physics
  - Accelerator simulation
  - High energy theoretical physics

Selected PRDs identified by HEP workshop

- Cosmology and Astrophysics Simulation
  - Conduct cosmic structure formation probes of dark universe
  - Understand and calibrate supernovae as probes of dark energy
- Accelerator simulation
  - Design lepton collider linac module for cost and risk reduction
  - Predict beam loss and activation in Intensity Frontier accelerators and maximize Energy Frontier accelerators

Cosmic Structure Formation Probes of the Dark Universe

Scientific and computational challenges

Understand cosmic structure to enable the use the universe as a probe of fundamental physics

Perform cosmological hydrodynamical simulations with the dynamic range necessary to interpret future experiments

Summary of research direction

Develop precise predictions of structure formation from the Hubble Volume to the scale of the Solar System

Develop spatially and temporally adaptive codes, algorithms, and workflows for simulations and data on extreme-scale architectures.

Potential scientific impact

Determine the equation of state of dark energy and distinguish between dark energy and modifications of General Relativity

Measure the masses and interactions of dark matter

Measure the sum of the neutrino masses

Probe the fields responsible for primordial fluctuations

Potential impact on High Energy Physics

Revolutionize High Energy Physics by discovering and measuring physics beyond the standard model inaccessible to accelerators.

10 years

The Software Dimension

Consensus view of Astrophysics Simulation and Data Panels

- Identify and support development of low-level modules and libraries, isolating architectural complexity (e.g., MPI, FFT)
- Identify and support development of open-source community application codes, but not to the exclusion of other promising efforts
- Promote development of data models and language for interoperable data analysis (observation <=> simulation)

Forefront Questions in Nuclear Science and the Role of High Performance Computing

- Workshop chair: Glenn Young,
  co-chairs David Dean, Martin Savage
- Panel topics
  - Nuclear structure and nuclear reactions
  - Nuclear astrophysics
  - Nuclear forces and cold QCD
  - Hot and dense QCD
  - Accelerator physics

Selected PRDs from NP workshop

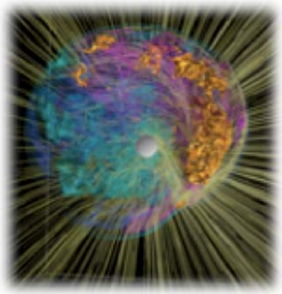
- Physics of extreme neutron-rich nuclei and matter
  - which includes computing properties of nuclei that determine the r-process nucleo-synthesis path in stars and the nucleonic matter in neutron star cores and crusts.
- Microscopic description of nuclear fission
  - which involves a problem of both basic science interest and of great practical importance—computing the paths to fission and its products.
- Nuclei as neutrino physics laboratories
  - which involves computing properties of nuclei in double-beta decay experiments and neutrino-nucleus cross sections for supernova calculations and neutrino decays.
- Reactions that made us
  - which involves computing the triple-alpha process that produces ^{12}C, the nucleus at the core of organic chemistry and thus life, and $^{12}C(\alpha,\gamma)^{16}O$, the element that is key to both water and the reactions that power humankind and our present society.

Nuclear Physics Workshop Summary Observations

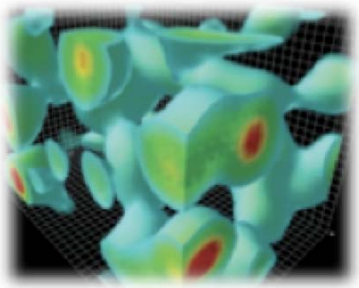
- Exa-scale computing resources are required to achieve the central mission of Nuclear Physics
  - theoretical AND experimental
  - science case is clear
- Nuclear Physics will be considerably unified (transformed) by Exa-scale computational resources
- Organizational infrastructure is required to enable full utilization of Exa-scale resources -- start organizing
- Collaborations within and external to Nuclear Physics are vital
- Prompt access to cutting edge machines is essential (- machine design)

Exa-scale Computational Resources
(slide courtesy Martin Savage)

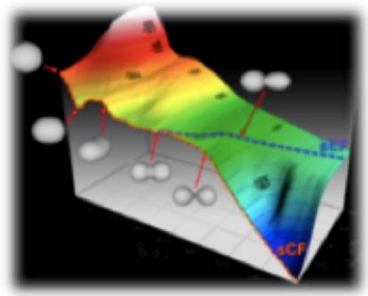
- Meeting structured around present Nuclear Physics areas of effort



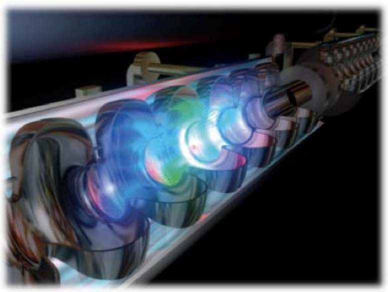
Nuclear Astrophysics



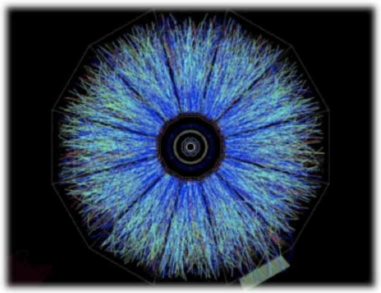
Cold QCD and Nuclear Forces



Nuclear Structure and Reactions



Accelerator Physics

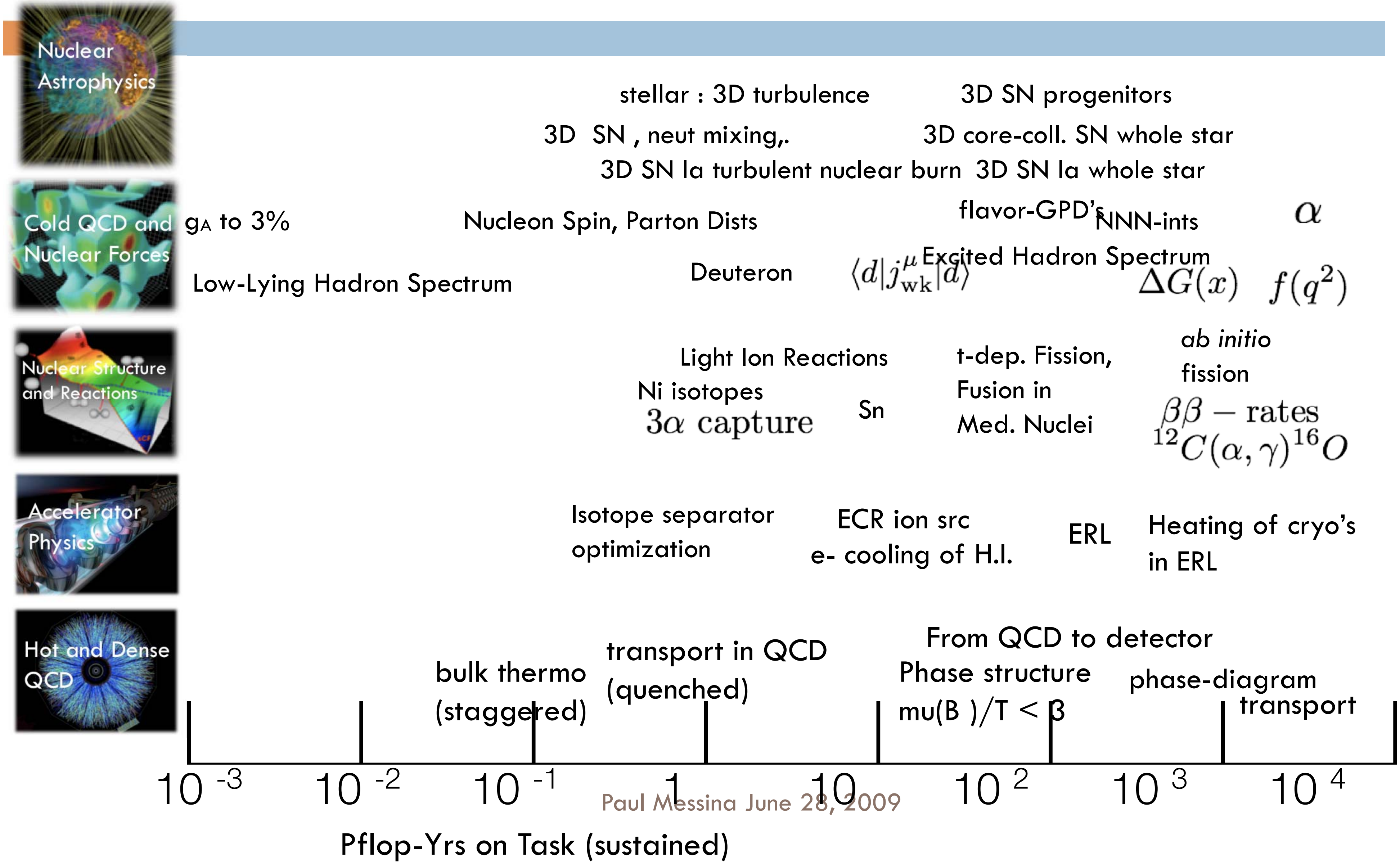


Hot and Dense QCD

- Exa-scale computing is REQUIRED to accomplish the Nuclear Physics mission in each area
- Staging to Exa-flops is crucial :
  - 1 Pflop-yr to 10 Pflop-yrs to 100 Pflop-yrs to 1 Exa-flop-yr (sustained)

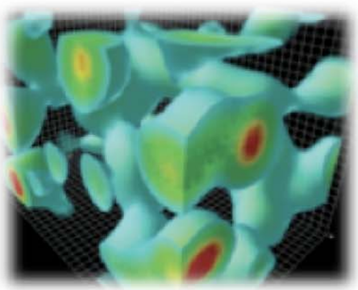
Nuclear Physics Requires Exa-scale Computation

(slide courtesy Martin Savage)

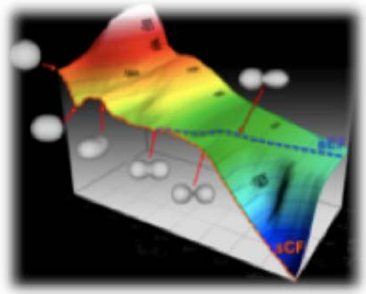


Exa-scale computing will unify Nuclear Physics

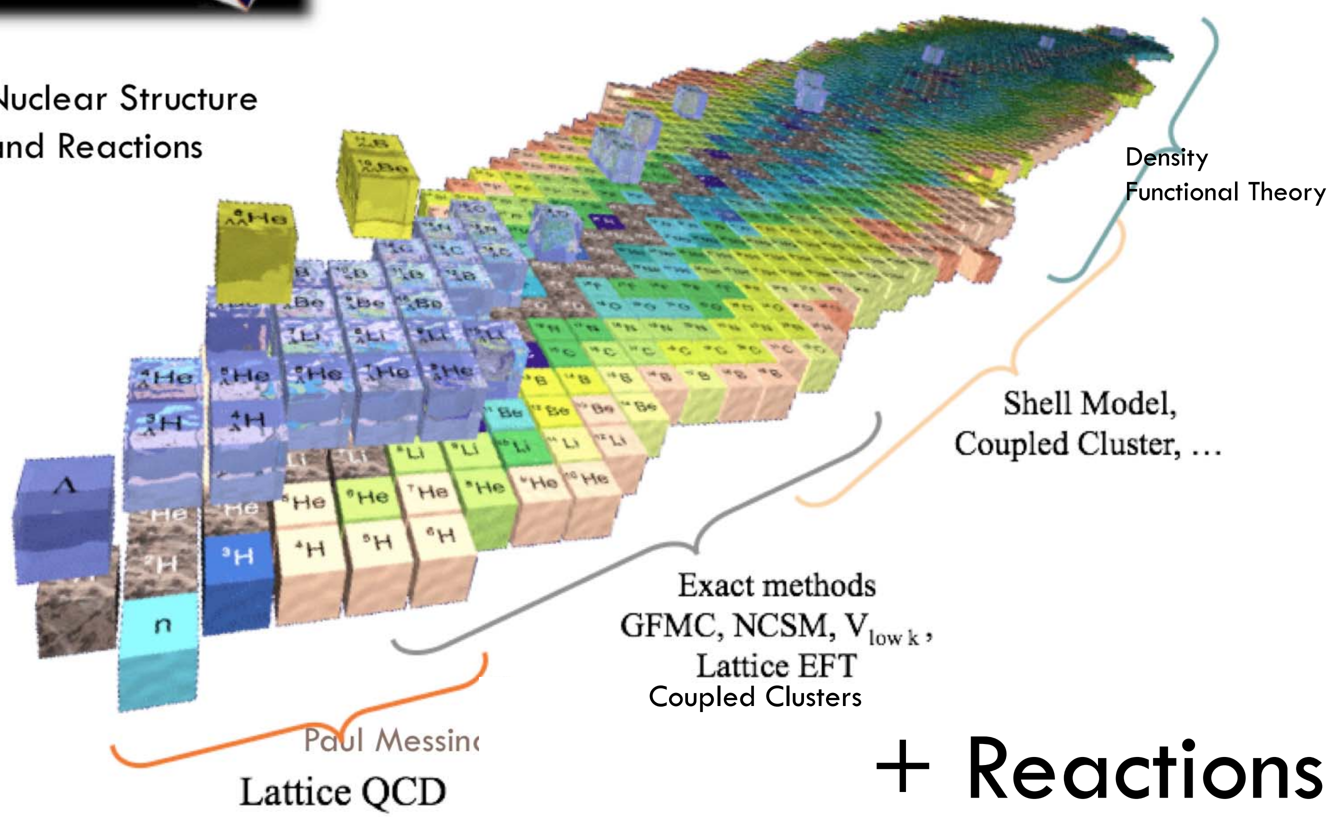
(slide courtesy Martin Savage)



Cold QCD and
Nuclear Forces



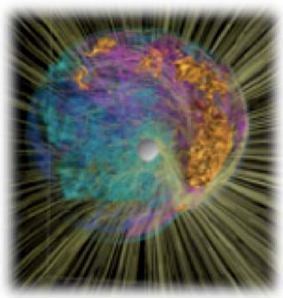
Nuclear Structure
and Reactions



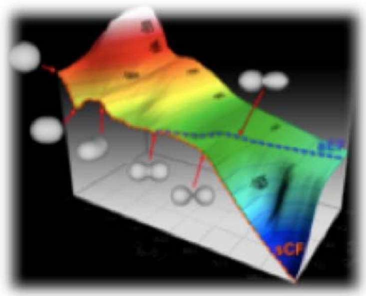
+ Reactions

Different Hardware Requirements

(slide courtesy Martin Savage)

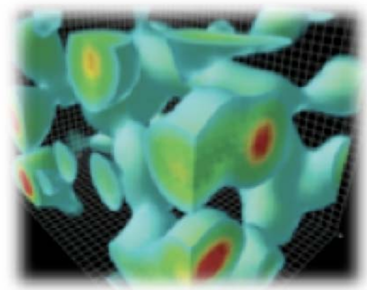


Nuclear Astrophysics

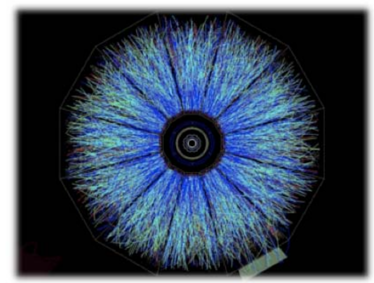


Nuclear Structure and Reactions

- Significant memory/core
- Fast memory access
- somewhat conflicts with present machine design ?
- low efficiency
- in current design : 4 GB/core preferred
- IO



Cold QCD and Nuclear Forces



Hot and Dense QCD

- Modest memory per core
- Large number of cores
- Range of Partitions
- Fast network
- low latency
- IO

Scientific Grand Challenges in Fusion Energy Sciences and the Role of Computing at the Extreme Scale

- Co-Chairs: Bill Tang and David Keyes
- Panel topics – Plasma and fusion energy sciences
  - Burning plasma/ITER
  - Advanced physics integration
  - Plasma-material interaction
  - Laser-plasma interactions
  - Magnetic reconnection physics

Scientific Grand Challenges in Fusion Energy Sciences and the Role of Computing at the Extreme Scale

- Panel topics: Cross-cutting areas
  - Scalable algorithms
  - Data analysis, management, and visualization
  - Mathematical formulations
  - Programming models, frameworks, and tools (incl. languages, optimization, etc.)

PRDs for Data Analysis, Management and Visualization in Fusion Energy Sciences (1)

- *Managing large-scale input/output volume and data movement.*
- *Real-time monitoring of simulations and run-time metadata generation.*
- *Data analysis at extreme scale.*
  - The data analysis challenges in FES applications at the extreme scale stem not only from the large size of the data, but also from data complexity.

PRDs for Data Analysis, Management and Visualization in Fusion Energy Sciences (2)

- *Visualization of very large datasets.*
  - Some codes will model 1 billion cells and 1 trillion particles
  - will need to output 3.5 TB/s of checkpoint data
- *Experiment-simulation data comparison.*
  - Such tools are essential for validation of FES simulations and diagnostics, and for comparing shot data to reduced models for ITER runs.
  - Experimental data are expected to grow to terabyte sizes.

Fusion Panel on Programming Models, Frameworks and Tools: six PRDs (1)

- *Find efficient algorithms and implementations that exploit new multicore, heterogeneous, massively parallel architectures.*
  - languages, libraries, and runtime systems that allow FES programmers to use massive on-chip concurrency in a portable, cross-architecture manner while cooperating with inter-processor parallelism.
- *Find new, productive approaches to writing, integrating, validating, and tuning complex FES application programs.*

Six Priority Research Directions for programming models, frameworks (2)

- *Develop tools for understanding complex application program behavior at scale and for optimizing application performance.*
- *Ensure a migration path from current FES programming approaches to new ones.*
- *Define common framework tools or components that can be reused in multiple FES application domains.*
- *Establish methods and systems that enable pervasive fault resilience.*

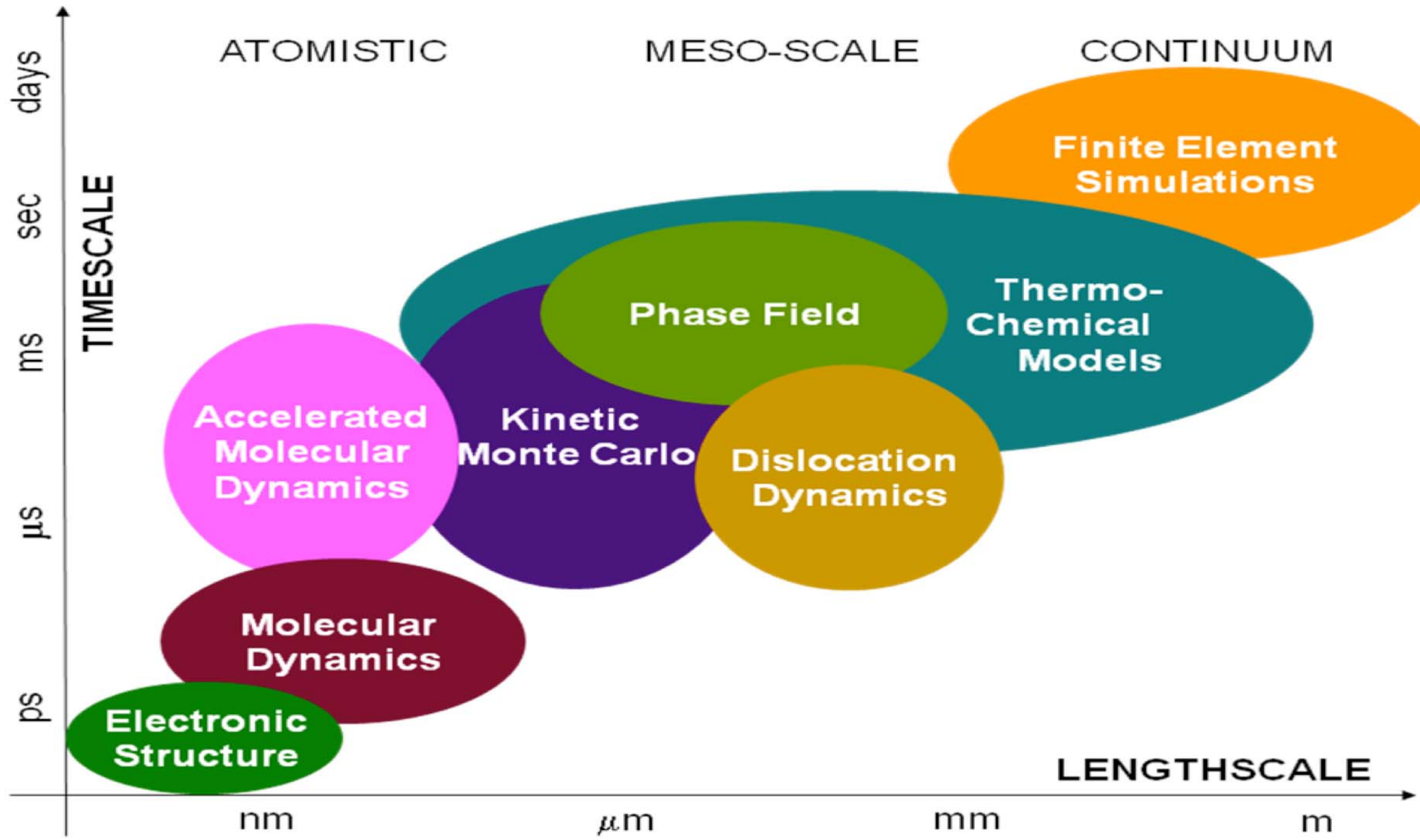
Science Based Nuclear Energy Systems Enabled by Advanced Modeling and Simulation at the Extreme Scale

- Co-Chairs: Ernie Moniz and Bob Rosner
- Panel topics
  - Integrated Performance and Safety Simulations of Nuclear Energy System
  - Advanced Material Behavior Modeling
  - Verification, Validation and Uncertainty Quantification for Nuclear Energy Simulations
  - Nuclear Energy System Integration

Nuclear energy materials modeling

- Applications of high performance (peta-scale and exa-scale) computing carry along both the burden and the opportunity of
  - improved uncertainty evaluations,
  - margins quantifications, and
  - reliable predictions of materials behavior.
- Exa-scale computing will enable
  - simulations of trillions of atoms over seconds or days
  - simulations of complex, coupled physics and chemistry of reactor materials.

Multiscale methods for materials



Paul Messina June 28, 2009

Discovery in Basic Energy Sciences: The Role of Computing at the Extreme Scale

August 12-14, 2009, DC

- Co-Chairs: Giulia Galli and Thom Dunning
- Panel topics
  - Correlation & Time Dependence
  - Photovoltaic Fundamentals
  - Energy Storage
  - Dynamics

Opportunities in Biology at the Extreme Scale of Computing August 17-19, 2009, Chicago

- Co-Chairs: Mark Ellisman and Rick Stevens
- Panel topics
  - Tissues, Organs and Physiology Modeling
  - Pathways, Cells and Organelles
  - Proteins and Protein Complexes
  - Populations, Communities, Ecosystems and Evolutionary Dynamics
  - Genomics and metagenomics etc.
  - Imaging and Computer in the Loop
    - single instruments to sensor loops

Scientific challenges for the NNSA and Office of Science missions October 6-8, 2009, DC

- Co-Chairs: Alan Bishop and Paul Messina
- Panel topics
  - Multiphysics simulation problems
  - Nuclear physics
  - Materials science
  - The science of nonproliferation
  - Chemistry
  - Uncertainty Quantification and Error Analysis

Recurring Topics in the Workshops -- Applications

- Exascale is needed: "the science case is clear"
- Predictive simulations
  - develop experimentally validated, predictive capabilities that can produce accurate and robust simulations
  - Ensemble simulations for V&V, UQ
- Multiphysics simulations
- Data volumes will be huge
  - Observations and simulation results

Recurring Topics in the Workshops – CS & Applied Mathematics

- Multiphysics and multiscale algorithm and software coupling
- Algorithms and software that heal with millions of cores and even more threads
- Data handling
- Interoperability
- Workflow issues
- Fault tolerance

Related Workshops

- We will organize workshops aimed at defining the properties of the computing environments that will enable researchers to tackle the science challenges, guided by the findings of the eight "Science Challenge" workshops
- The workshops that the International Exascale Software Project is organizing will play an important role in identifying the key software environment and tools that are necessary for productive use of present and future leadership computer systems and viable approaches to developing them

For more information

- http://extremecomputing.labworks.org/