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

W. M. TANG
Princeton University, Plasma Physics Laboratory,

SNA+MC2010
Tokyo, Japan

19 October 2010

Korean Prime Minister Kim Hwang-Sik:

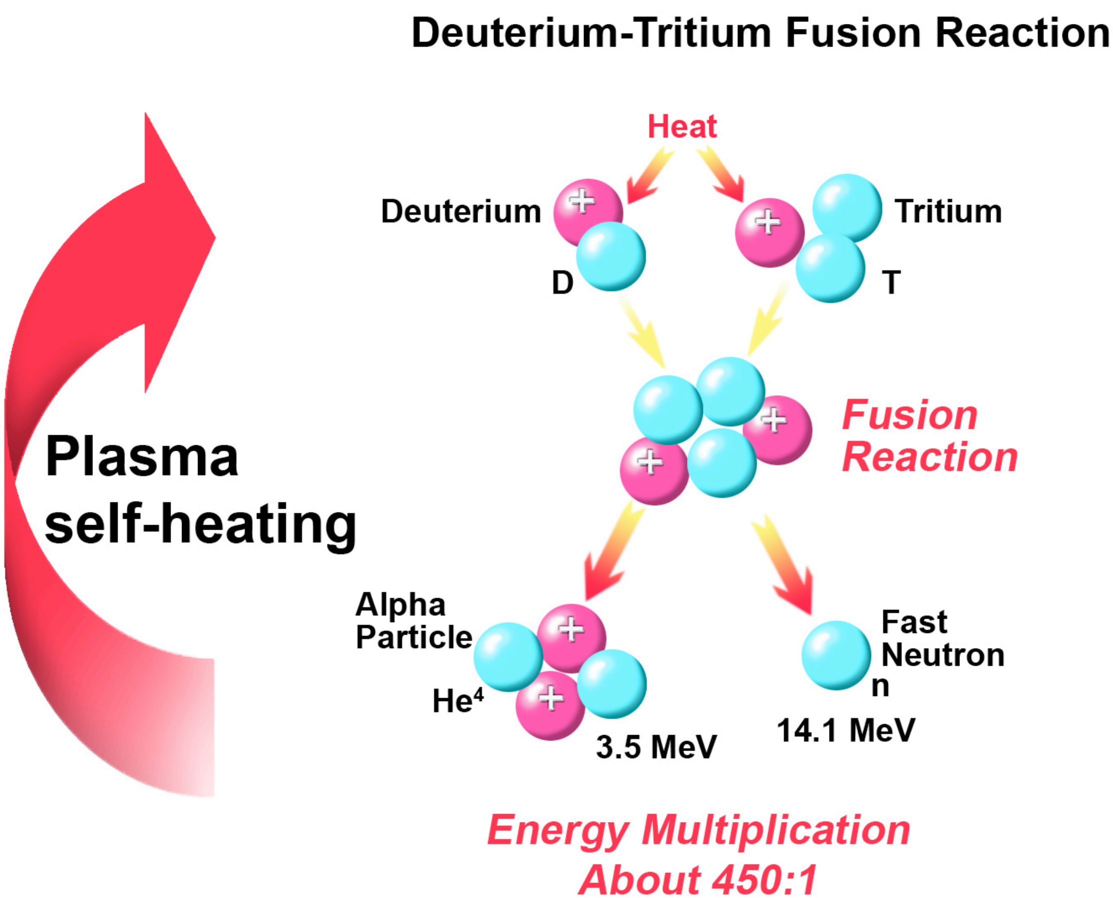
"Convinced that green energy is the key to sustainable growth, Korea has actively taken to pursuing various green R&D development initiatives in areas such as Fusion Energy."

11 October 2010 – Opening Ceremony of the 23rd International Atomic Energy Agency (IAEA) Fusion Energy Conference
Daejon, Korea

"Notwithstanding the looming energy crisis, the high potential of Fusion Energy provides humanity with a silver lining.

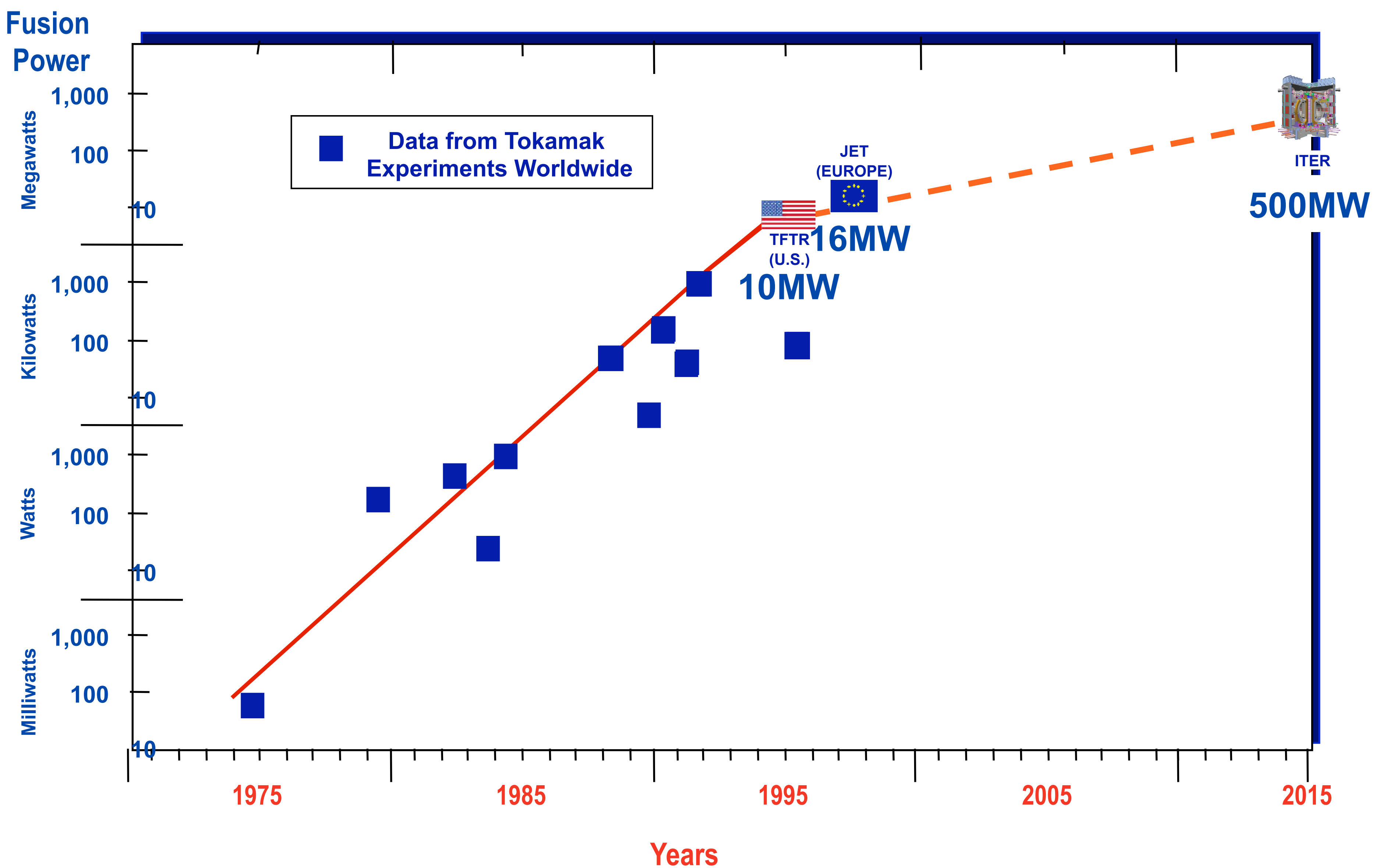
Fusion Energy is a source of green energy which is harmless to the environment and a source of hope for all humanity."

Fusion Energy: *Burning plasmas are self-heated and self-organized systems*



$$D^{+} + T^{+} \rightarrow {}^{4}He^{++}\ (3.5\ MeV) + n^{0}\ (14.1\ MeV)$$

Progress in Magnetic Fusion Research



ITER Goal: *Demonstration of the Scientific and Technological Feasibility of Fusion Power*

- ITER *is a dramatic next-step for Fusion*:

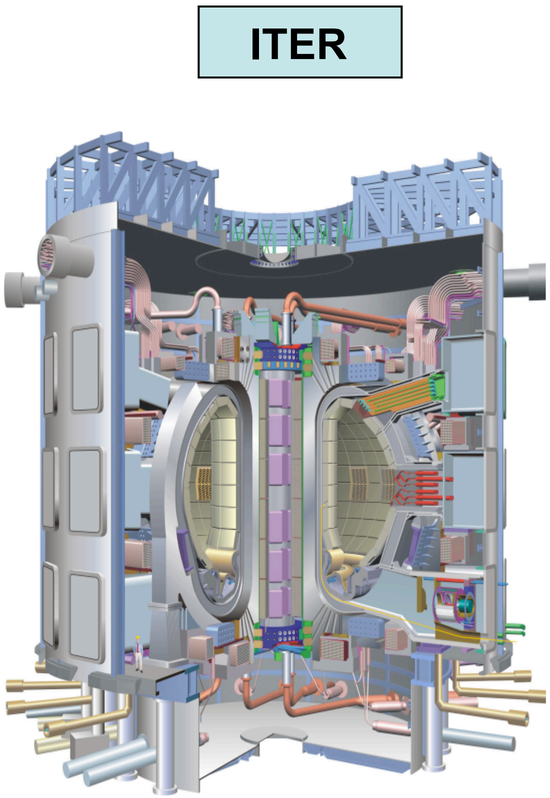
  -- Today: 10 MW(th) for 1 second with gain ~1
  -- ITER: 500 MW(th) for >400 seconds with gain >10

- Many of the technologies used in ITER will be the same as those required in a power plant *but additional R&D will be needed*

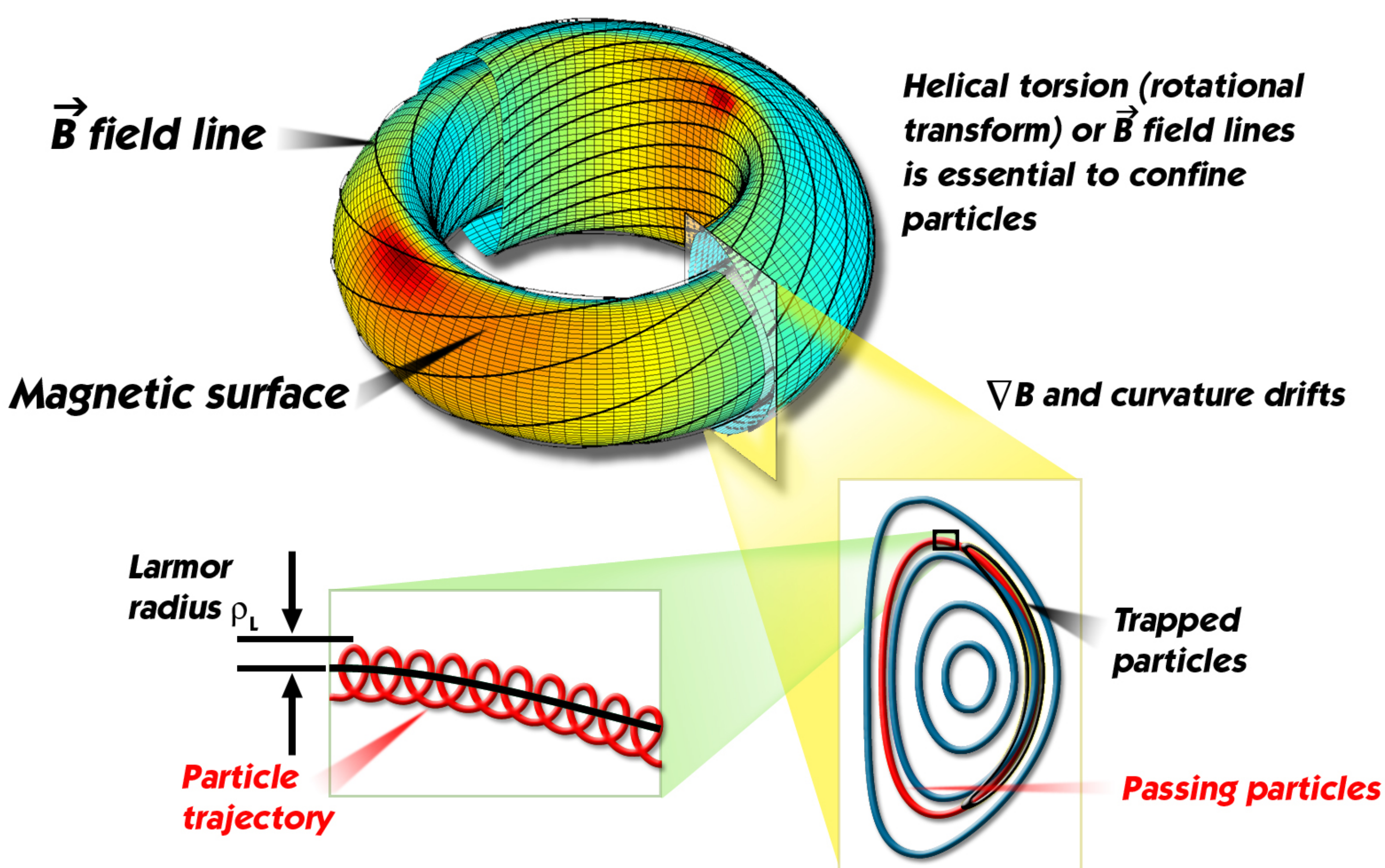
  -- "DEMO": 2500 MW(th) continuous with gain >25, in a device of similar size and field as ITER
  * Higher power density
  * Efficient continuous operation

- Strong R&D programs are required to support ITER and leverage its results.

  -- Experiments, theory, computation, and technology that support, supplement and benefit from ITER



Magnetically confined plasmas in a tokamak are complex and demand integrated analysis

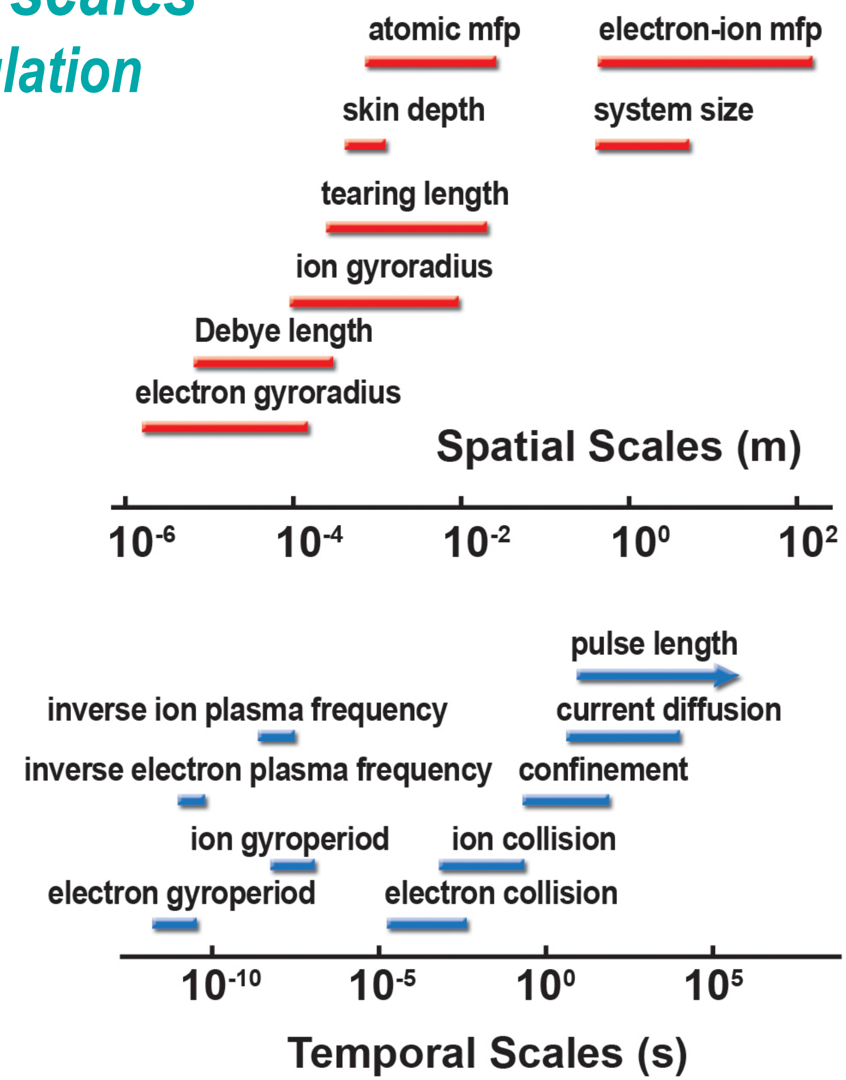


Integrated predictive models must span huge range of spatial & temporal scales
-- major challenges to theory and simulation

- **Overlap in scales often means strong (simplified) ordering is not possible**

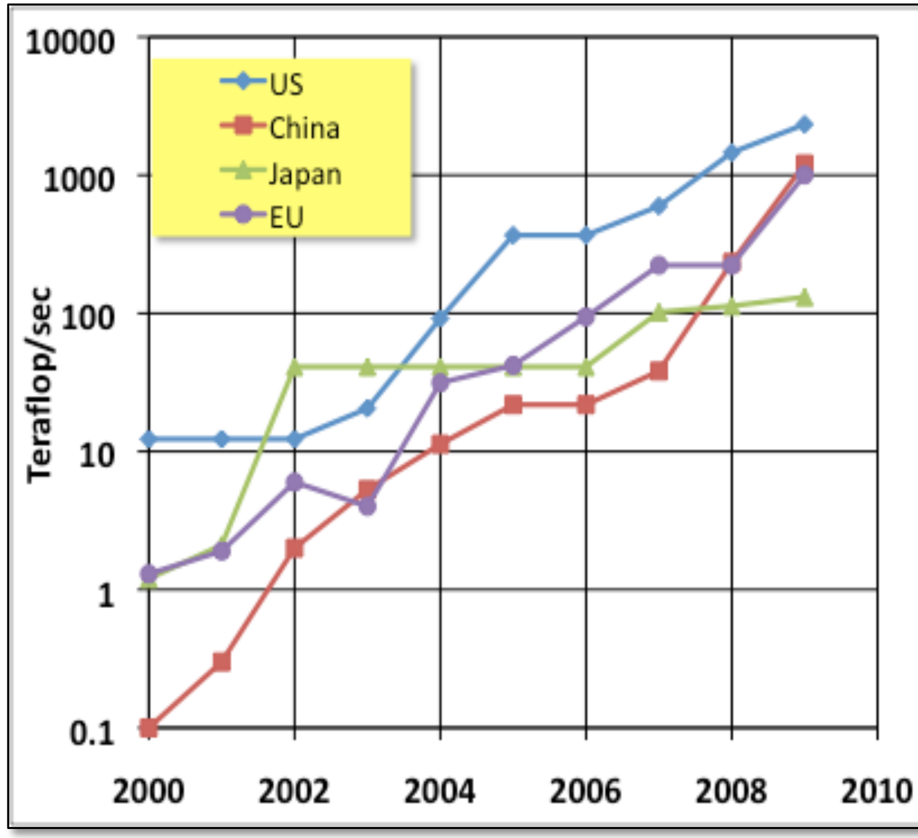
- **Needed to effectively harvest insights from ITER and to plan for DEMO**

- **Effective simulations at the petascale** *(10^{15} floating point operations per second)* **and beyond are required to address grand challenges in plasma science**



Computational Science, Exascale Computing & Leadership in Science & Technology

- **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

Pre-eminence in 21st Century science, technology, & engineering requires leadership in computational science and high performance computing => exascale applications & technology

Advanced Computing can Transform Many Domain Applications Areas (including FES)

- ***Practical Considerations:*** **[achieving "buy-in" from general scientific community]**
  - **Need to distinguish between *"voracious"* (*more of same - just bigger & faster*) *vs.* *"transformational"* (*achievement of major new levels of scientific understanding*)**
  - **Need to improve *significantly* on *experimental validation* together with *verification & uncertainty quantification* to enhance realistic predictive capability**

- ***Associated Extreme Scale Computing Challenges:***
- ***Hardware complexity*: Heterogenous multicore *(e.g., cpu+gpu -- LANL, ORNL, …)*, power management, memory, communications, storage, …**
- ***Software challenges*: Operating systems, I/O and file systems, and coding/algorithmic needs in the face of *increased computer architecture complexity* … *"parallelism doubles every two years" (as a new form of Moore's Law)***

  (MPI + threads; CUDA; rewriting code focused on data movement over arithmetic; …..)

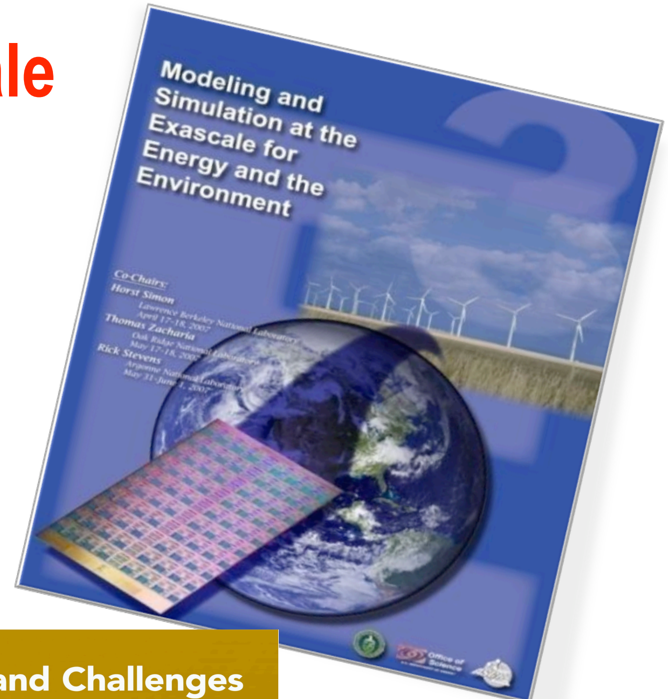
******People:** Major challenge to attract, train, & assimilate the next generation of simulation/modeling-oriented CS, Applied Math and applications-oriented computational scientists and engineers.*

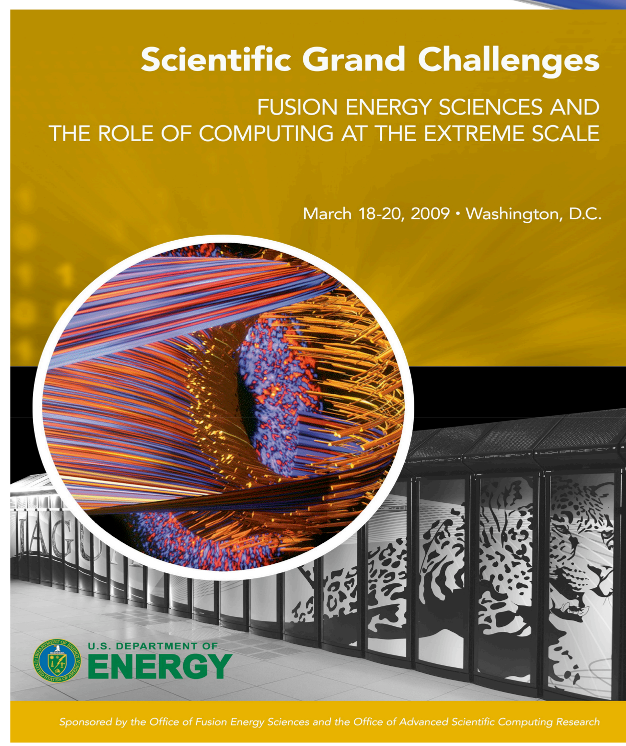
Moving to the Exascale

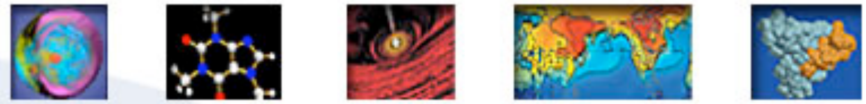
- **DoE (SC/NNSA) held series of workshops in 2009-2010 (including [FES](#)) to assess the *opportunities and challenges of exascale computing for the advancement of science, technology, and Office of Science missions.***
- **ASCR strategy to address the challenges and deliver on such opportunities involves working with:**

  -- domain applications areas such as *FES* to scale applications to each of the new computer systems

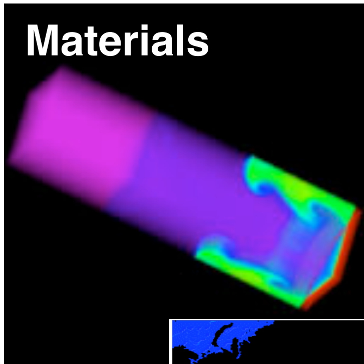
  -- LCF's at ORNL & ANL in providing series of increasingly powerful computer systems

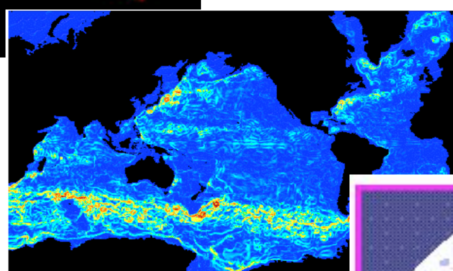




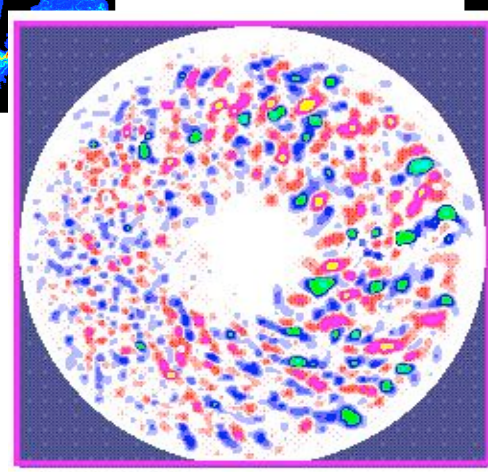


Dramatic Advances
in Simulation Capabilities
NEEDED TO ACCELERATE PROGRESS
in Many Scientific Disciplines

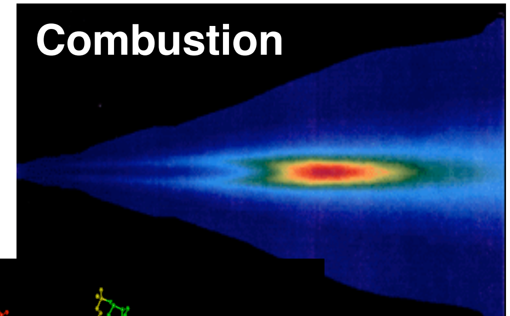


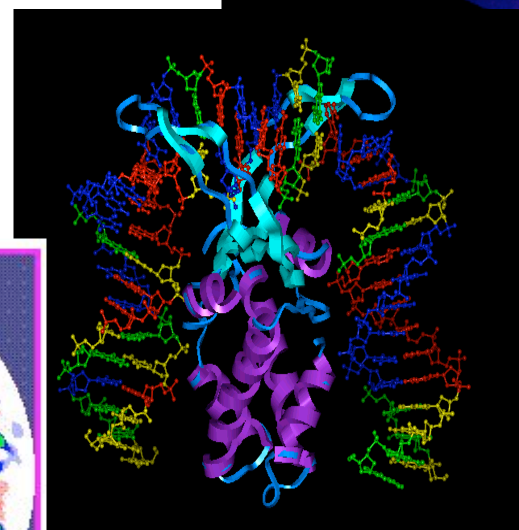


Global Systems



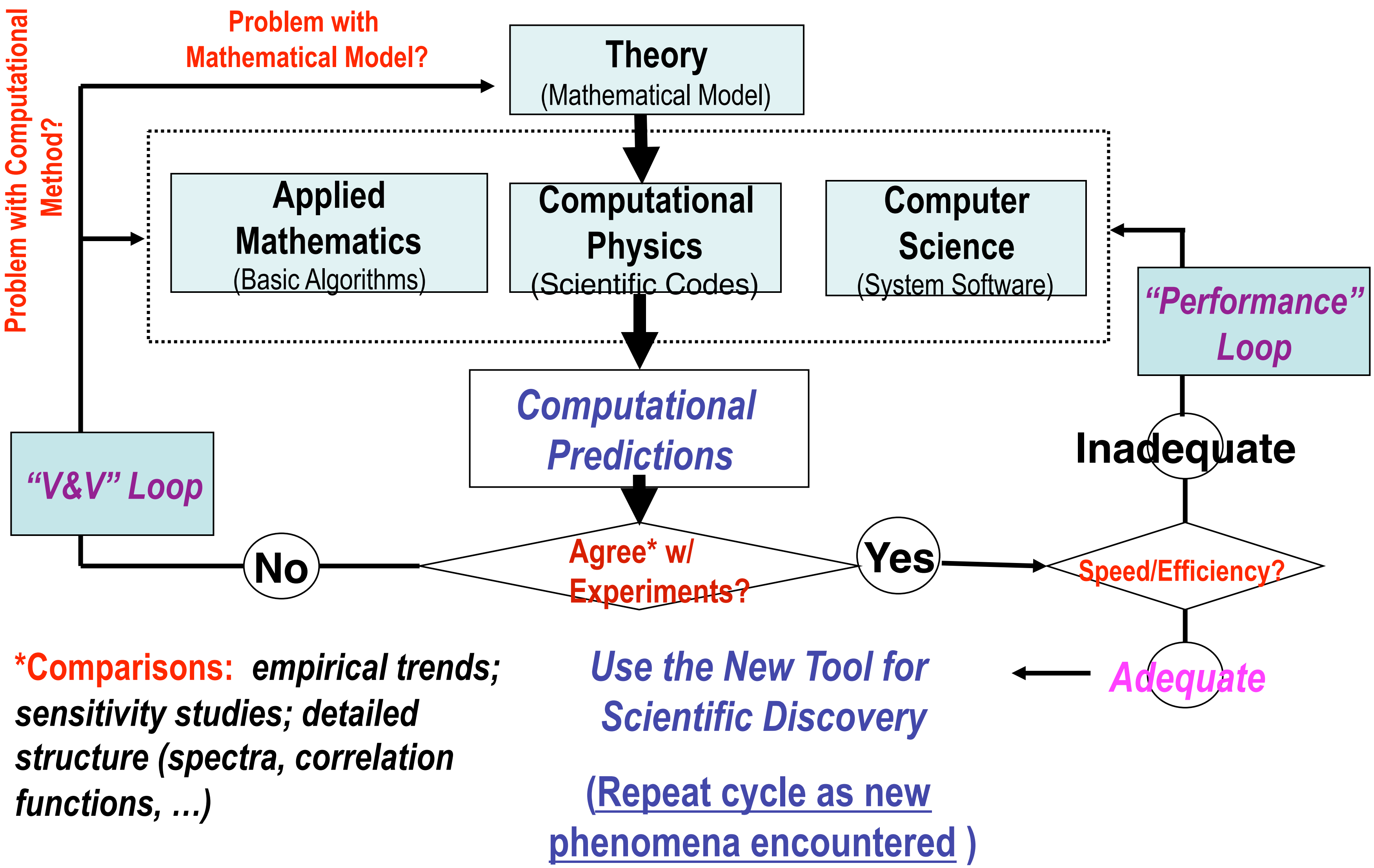
Fusion Energy





Health Effects, Bioremediation

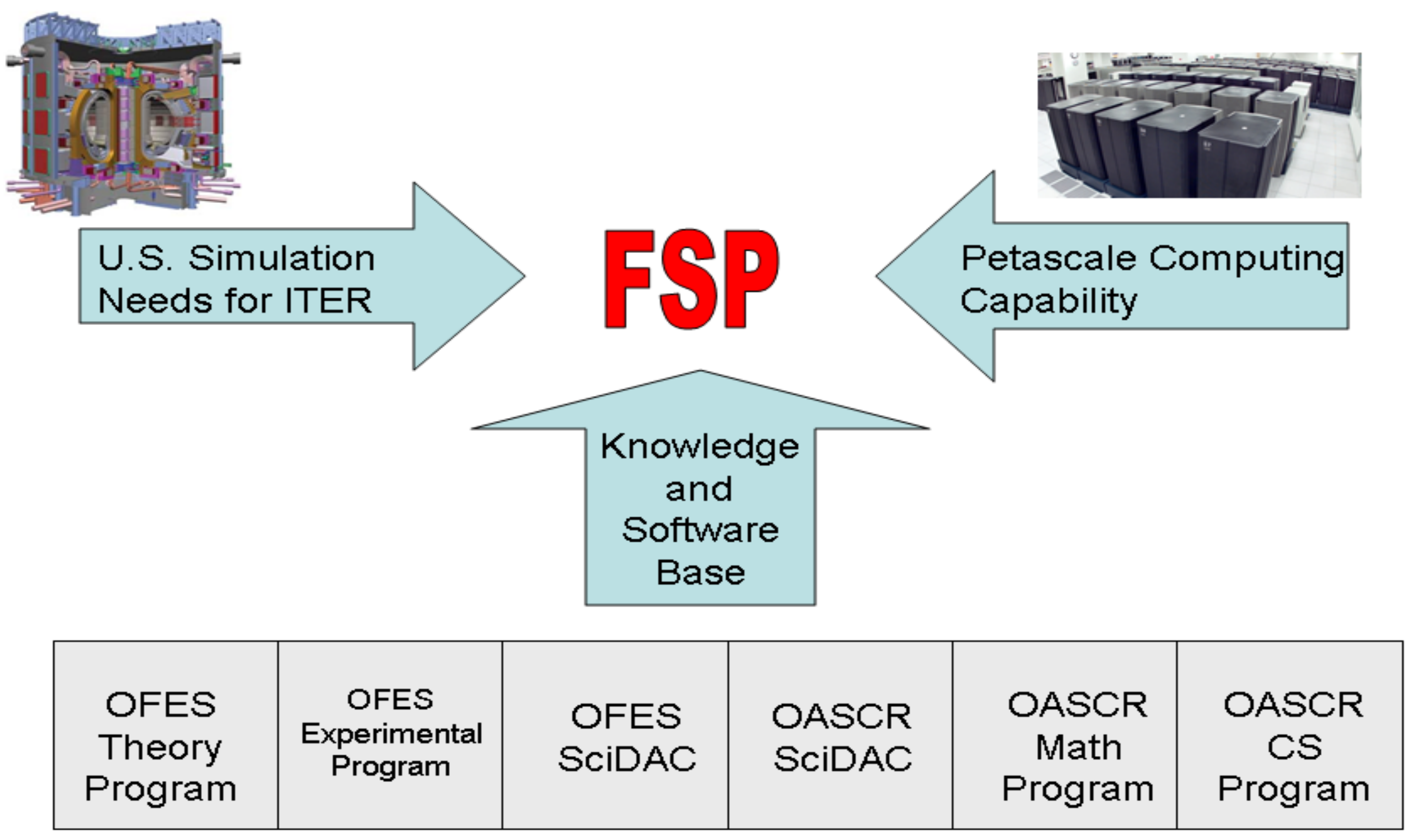
Advanced Scientific Codes --- *"a measure of the state of understanding of natural and engineered systems"* (T. Dunning, 1st SciDAC Director)



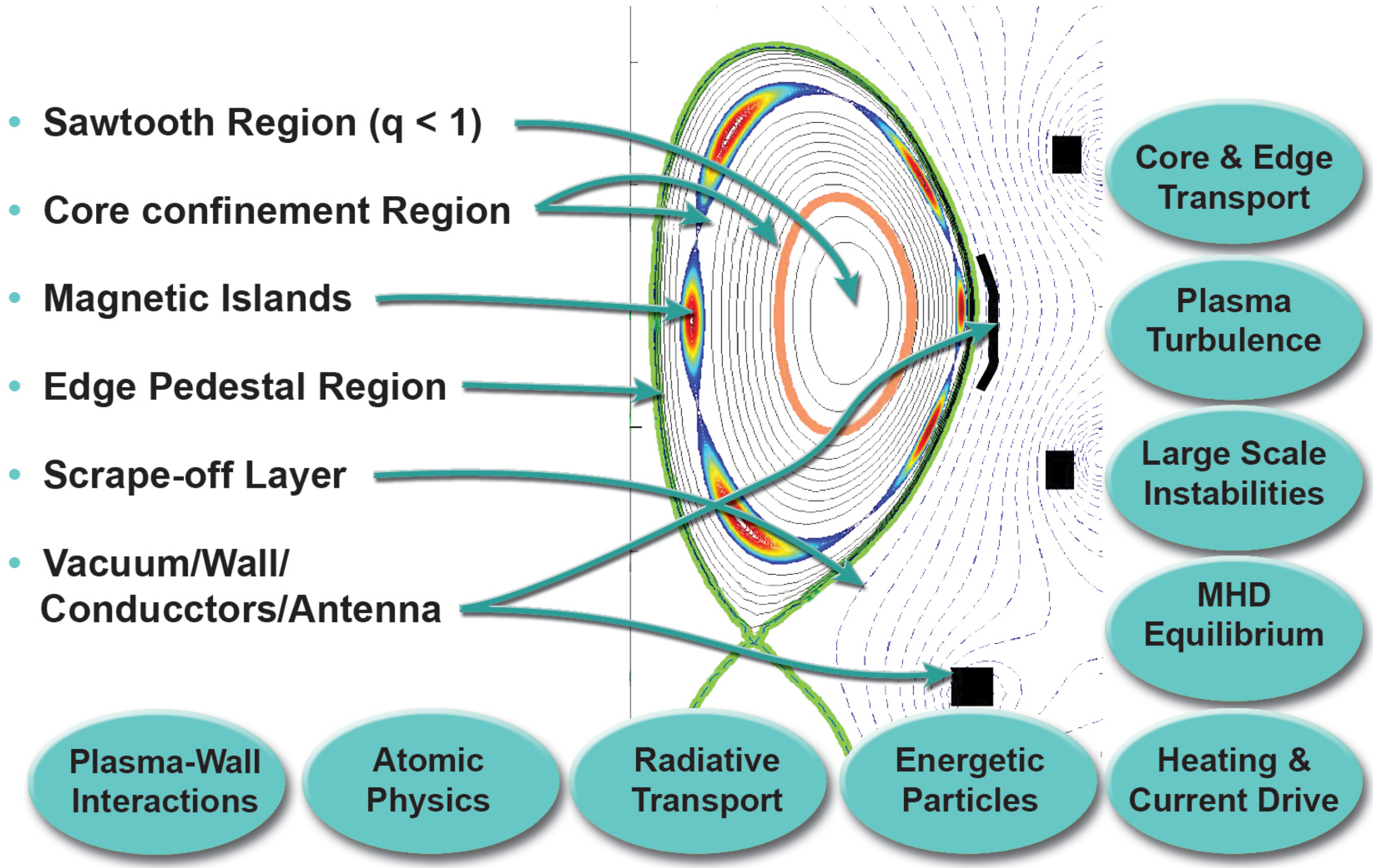
*Comparisons: *empirical trends; sensitivity studies; detailed structure (spectra, correlation functions, …)*

FSP -- A Strategic *Opportunity* to Accelerate Scientific Progress in FES

- Need for reliable predictive simulation capability for *BP/ITER* *(especially in the US)*
- Powerful ("Leadership Class") Computational Facilities moving rapidly toward petascale & beyond
- Interdisciplinary *collaborative experience*, knowledge, & software assembled over the course of nearly a decade under SciDAC plus OFES and OASCR base research programs in the US

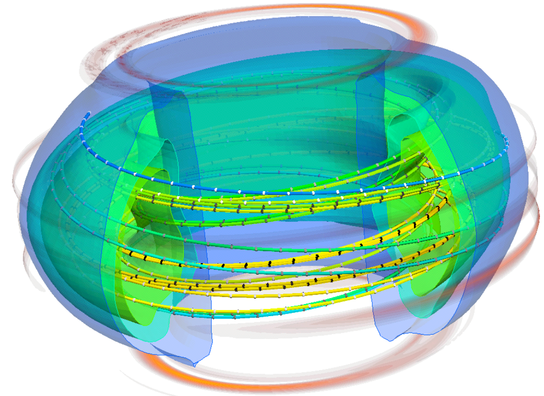


Elements of an FSP Integrated Model

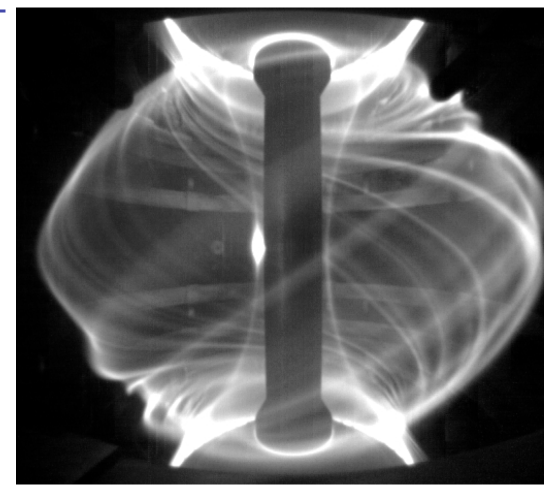


Key Scientific Challenges for Burning Plasmas

- ***Disruptions***: Large-scale macroscopic events leading to rapid termination of plasma discharges
  - Avoid or mitigate because ITER can sustain only a limited number of full-current disruptions
- Pedestals: Formation of steep spatial gradients leading to transient heat loads in plasma periphery (divertor region)
  - Predict onset and growth because pedestal height is observed to control confinement
  - Predict frequency and size of ***edge localized mode (ELM) crashes*** to mitigate erosion of divertor and plasma-facing components
- Tritium migration/retention and impurity transport
- Performance optimization and scenario modeling
- Plasma feedback control
  - *Burning plasma regime is fundamentally new with stronger self-coupling and weaker external control*



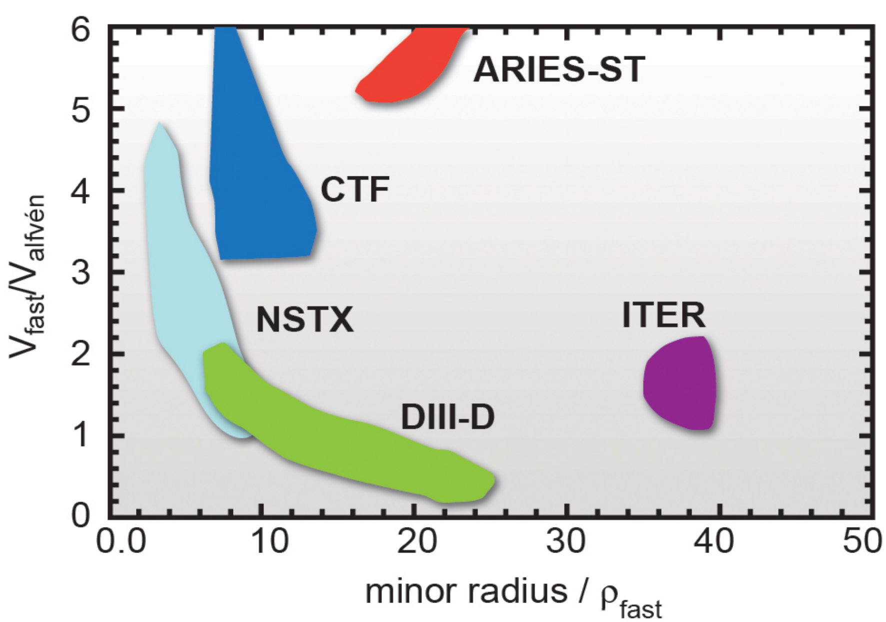
Plasma disruption in DIII-D



ELMs in MAST

Magnetically Confined Burning Plasmas:
Unique opportunities for scientific discoveries

- BP/ITER physics elements raise mission-critical questions
  - Unprecedented size
  - Self-heating
  - Large energetic particle population
  - Multiple instabilities with unknown consequences for fast ion confinement



Predicting fast ion confinement: Critical for sustaining a burning plasma

- What is nonlinear interaction between energetic particles and "sea of Alfvén modes?"
- How is transport affected by presence of multiple instabilities?
- How can predictive numerical tools be properly validated?
- What scale of computational resources will be needed to answer BP/ITER mission-relevant questions?

Verification, Validation, & Uncertainty Quantification Challenges

- Establishing the physics fidelity of modern plasma science simulation tools demands proper Verification & Validation (V&V) and Uncertainty Quantification (UQ) -- ***Reliable codes demand solid theoretical foundations and careful experimental validation***

  - **Verification** assesses degree to which a code (*both in the advanced direct numerical simulation (DNS) and reduced models categories*) correctly implements the chosen physical model
    --- more than "essentially a mathematical problem"
    e.g., accuracy of numerical approximations, mesh/space and temporal discretization, statistical sampling errors, etc.
    --- also requires*:* (1) comparisons with theoretical predictions; and (2) cross-code benchmarking (codes based on different mathematical formulations/algorithms but targeting the same generic physics)

  - **Validation** assesses degree to which a code (within its domain of applicability) "describes the real world"

  - **Uncertainty Quantification** is the quantitative characterization & reduction of uncertainty in applications related to variability of input data/model parameters & uncertainties due to unknown processes or mechanisms (e.g., sensitivity analysis)

Collaboration With Experimental Facilities

- **Strong collaboration with experimental facilities essential for validation of FSP codes** – *regarded as significant value to the facilities for planning and interpretation of experiments.*

- **Draft agreement with major U.S. facilities (DIII-D, C-MOD, NSTX) defining:**
  - *General principles for collaboration and intellectual property (IP) sharing*
  - *Proposed mechanisms for short-term and long-term planning*
  - *Roles & Responsibilities for the FSP and for experimental teams in their collaboration*
  - *Lessons learned from the major experimental facilities useful in planning the FSP research program– e.g. open community research forums*

- **The existing collaboration engagement agreements for and approaches to *research governance* used by the facilities provide a proven model for the FSP.**

- **Elements for Success (from past experience) dictates:**

  -- ongoing partnerships and mutual interactions institutionalized through formal agreements; and

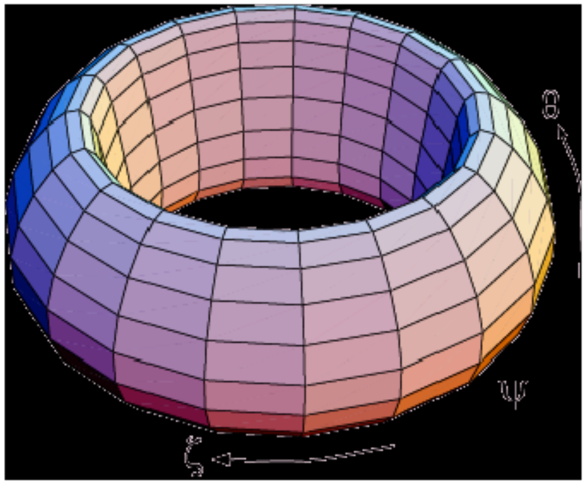
  -- regular participation in planning and reporting activities with cross-membership in planning groups

HPC SIMULATION PROBLEM DESCRIPTION FOR KEY FES TOPIC:

Particle-in-cell (PIC) Simulations of Plasma Turbulence

- **Issue: confinement of high temperature plasmas by magnetic fields in 3D geometry**

Fusion reactor size and cost are determined by balance between loss processes and self-heating rates



- **Pressure gradients drives instabilities producing loss of confinement due to turbulent transport**

- **Plasma turbulence is nonlinear, chaotic, 5-D problem**

- **Particle-in-cell simulation**

→distribution function - integrate along characteristics with particles advanced in parallel

→interaction - self-consistent EM fields

Particle Simulation of the Boltzmann-Maxwell System

- **The Boltzmann equation (Nonlinear PDE in Lagrangian coordinates):**

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \mathbf{v} \cdot \frac{\partial F}{\partial \mathbf{x}} + \left(\mathbf{E} + \frac{1}{c} \mathbf{v} \times \mathbf{B} \right) \cdot \frac{\partial F}{\partial \mathbf{v}} = C(F).$$

- **"Particle Pushing" (Linear ODE's)**

$$\frac{d\mathbf{x}_j}{dt} = \mathbf{v}_j, \qquad \frac{d\mathbf{v}_j}{dt} = \frac{q}{m} \left(\mathbf{E} + \frac{1}{c} \mathbf{v}_j \times \mathbf{B} \right)_{\mathbf{x}_j} .$$

- **Klimontovich-Dupree representation,**

$$F = \sum_{j=1}^{N} \delta(\mathbf{x} - \mathbf{x}_j)\delta(\mathbf{v} - \mathbf{v}_j),$$

- **Poisson's Equation: [Linear PDE in Eulerian coordinates (lab frame)]**

$$\nabla^2 \phi = -4\pi \sum_{\alpha} q_\alpha \sum_{j=1}^{N} \delta(\mathbf{x} - \mathbf{x}_{\alpha j})$$

- **Ampere's Law and Faraday's Law [Linear PDE's in Eulerian coordinates (lab frame)]**

Particle-in-Cell Simulation: Background

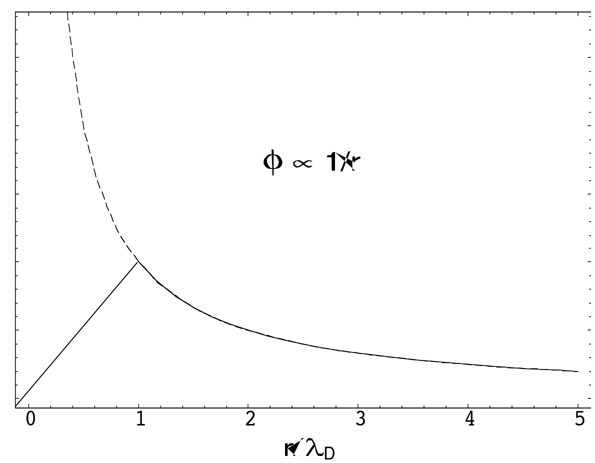
• Early attempts [Buneman (1959); Dawson (1962)]

• Finite-Size Particles and Particle-in-Cell Simulation [Dawson et al. (1968) and Birdsall et al. (1968)]

- Coulomb potential modified by Debye shielding

- Short-range forces within Debye sphere ignored

- *Point particles replaced by finite sized particles -- uniformly charged spheres of Debye-length radius*



• Number of calculations for N particles

- N^2 for direct interactions and N for PIC

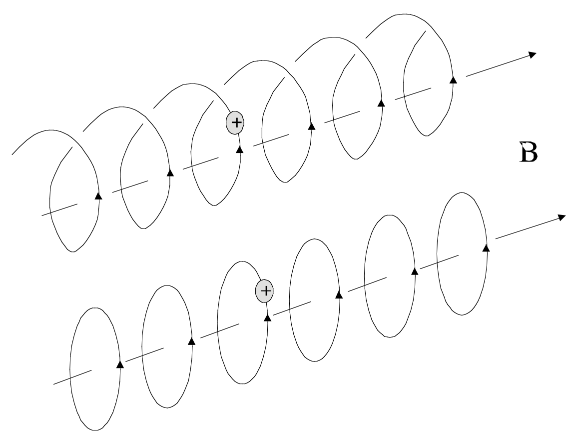
• Collisions are treated as sub-grid phenomena via Monte-Carlo methods [Shanny, Dawson & Greene (1976)]

Gyrokinetic Particle Simulation

- [W. Lee, PF ('83); JCP ('87)]

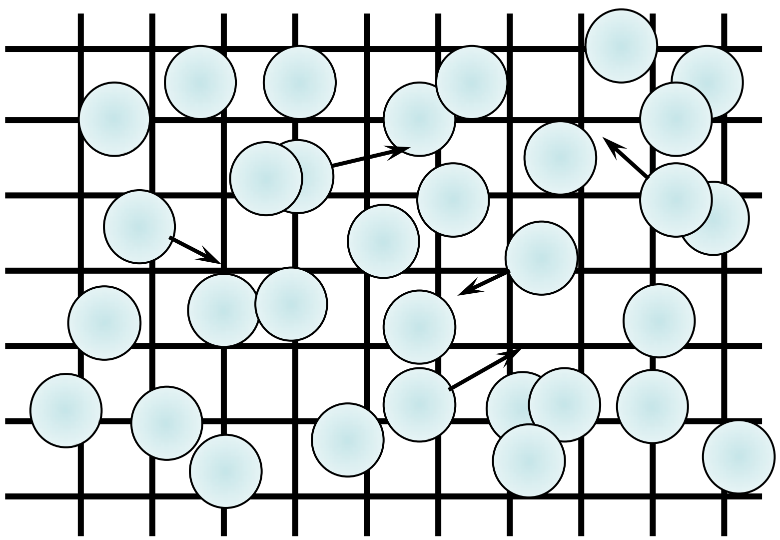
- **Gyrophase-averaged Vlasov-Maxwell equations for low frequency microinstabilities.**

- **Spiral motion of a charged particle is modified as a rotating charged ring subject to guiding center electric and magnetic drift motion as well as parallel acceleration --** speeds up computations by ***3 to 6 orders of magnitude in time steps and 2 to 3 orders in spatial resolution***



Particle-in-cell (PIC) Method

- Particles sample distribution function (markers).
- The particles interact via a grid, on which the potential is calculated from deposited charges.

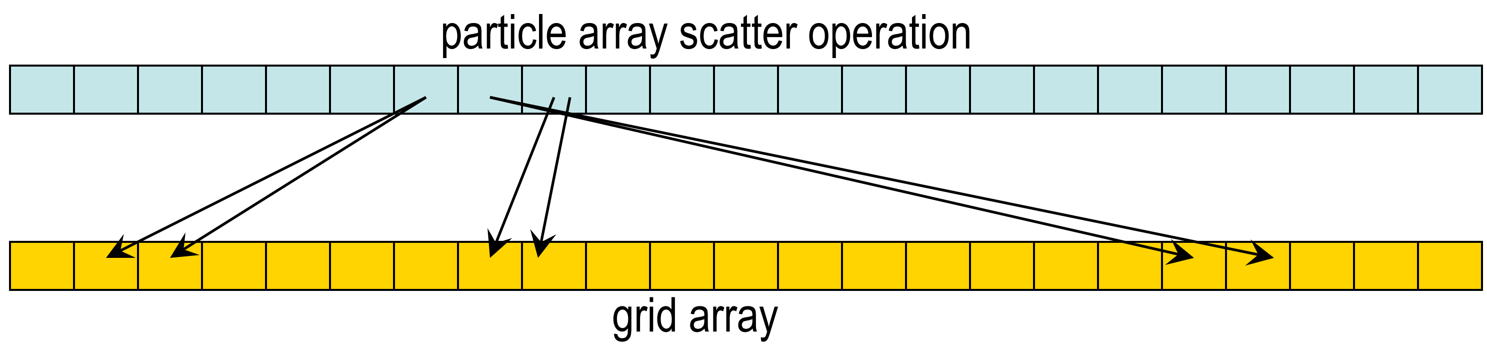


The PIC Steps

- "**SCATTER**", or deposit, charges on the grid (nearest neighbors)
- Solve Poisson equation
- "**GATHER**" forces on each particle from potential
- Move particles (**PUSH**)
- Repeat…

General Nature of PIC Computational Challenge

- ***“Gather-Scatter” operation in PIC codes***
  - **The particles are randomly distributed in the simulation volume (grid).**
  - **Particle charge deposition on the grid leads to indirect addressing in memory (see below).**
    - **need to arrange data to enable “direct-addressing” (at least for some time period)**
    - **also a problem in computer games**
  - **Not cache friendly.**
  - **Need to be tuned differently depending on the architecture.**



General Computational Challenges for FES

- **Fast and Efficient Elliptic (Poisson) Solvers:**
  - Required for both Particle-in-Cell (PIC) kinetic codes and Magneto-hydrodynamics (MHD) fluid codes.
    - PIC applications involve extremely large sparse matrix system (10^8 X 10^8 grid points)
  - Deal with non-Cartesian irregular grid in toroidal geometry.
  - Need efficient pre-conditioner to speed-up the solve (e.g., pre-arranging matrix)
  - Portable parallel solver
- **Optimization of Parallel Algorithms:**
  - Improve scalability and efficient utilization of increasing numbers of processors
  - Properly distribute particles over simulation domain.
  - Improve load balancing

Simulation of Turbulence in Future Ignition-Scale Experiments

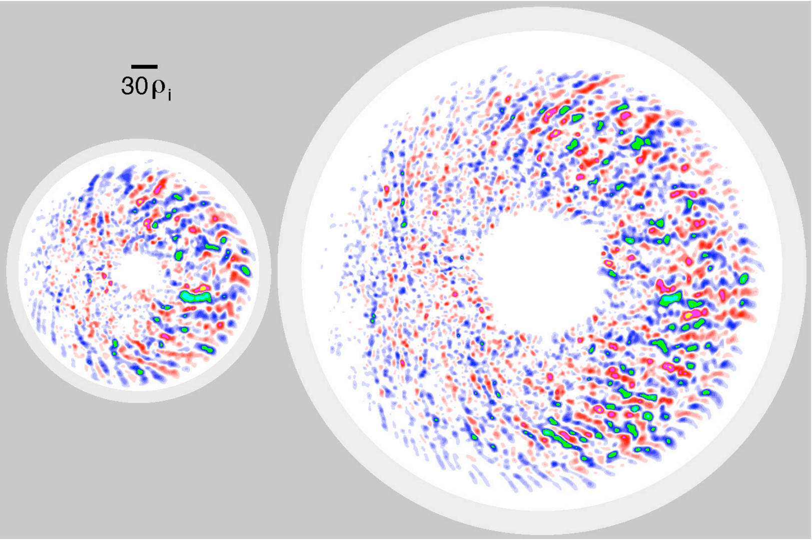
"Scientific Discovery" requiring Leadership-Class Computers

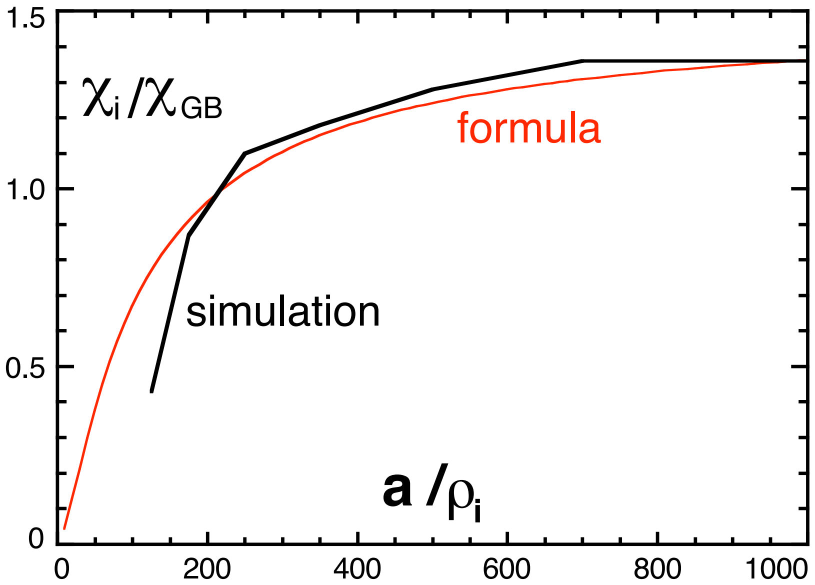
- Microturbulence Simulations for range including:
  - a/ρ_i = 400 (JET, largest present lab experiment) through
  - a/ρ_i = 1000 (ITER, ignition experiment)

- Results (*PRL, 2002*) enabled by use of MPP platforms (e.g., from multi-TF runs @ NERSC)

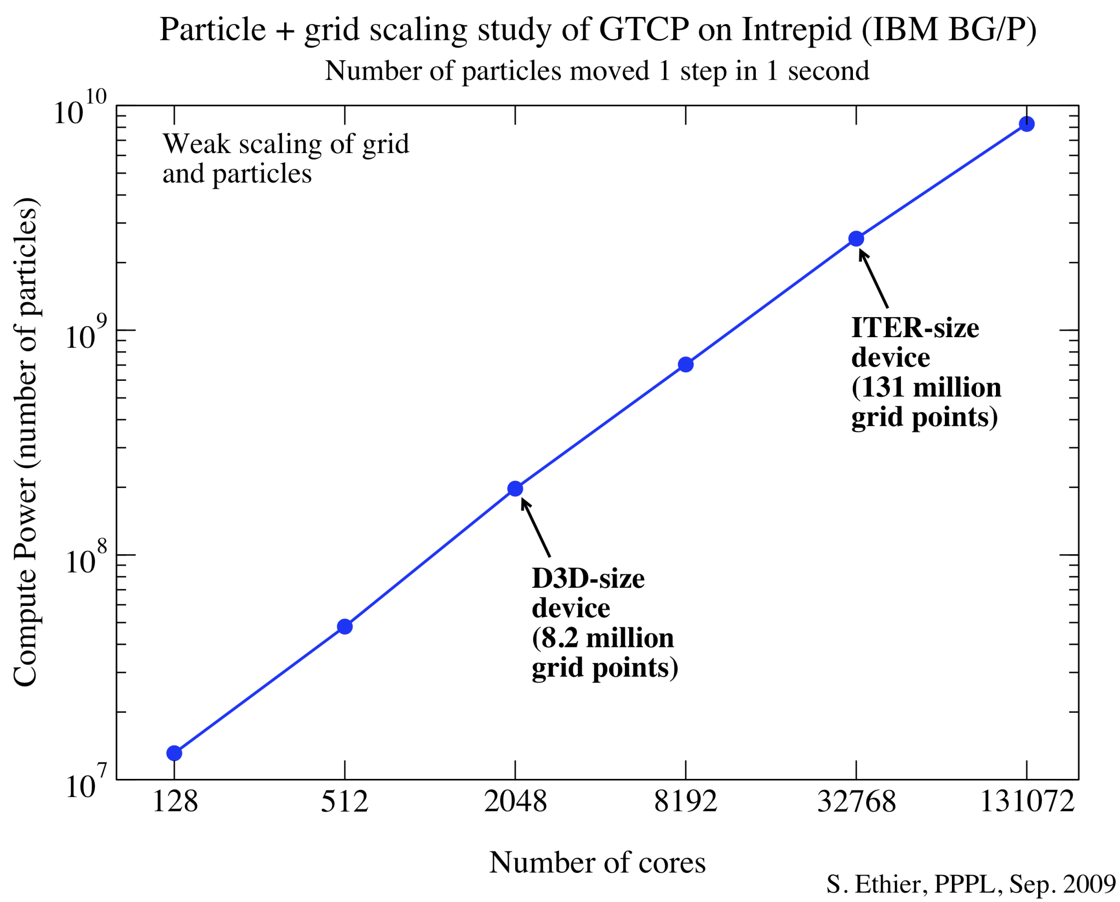
- These PIC simulations: 1 billion particles, 125M spatial grid points; 7000 time steps

- Such larger-scale simulations indicate transition to more favorable scaling of plasma confinement





Scaling GTC-P on IBM BG-P at ALCF



Excellent scalability demonstrated – promising basis for performance on IBM BG-Q

Recent LCF-enabled simulations provide new insights into nature of plasma turbulence

Teraflops-to- petaflops computing power have accelerated progress in understanding heat losses caused by plasma turbulence

Multi-scale simulations accounting for fully global 3D geometric complexity of problem (spanning micro and meso scales) have been carried out on DOE-SC Leadership Computing Facilities

Excellent Scalability of Global PIC Codes *enabled by strong ASCR-FES collaborations* in SciDAC (e.g., XGC1 code in proto-FSP CPES project)

Exascale-level production runs are needed to enable running codes with even higher physics fidelity and more comprehensive & realistic integrated dynamics

e.g. -- *Current petascale-level production runs on ORNL's Jaguar LCF require 24M CPU hours (100,000 cores × 240 hours)*

Mission Importance:

Fusion reactor size and cost are determined by balance between loss processes and self-heating rates

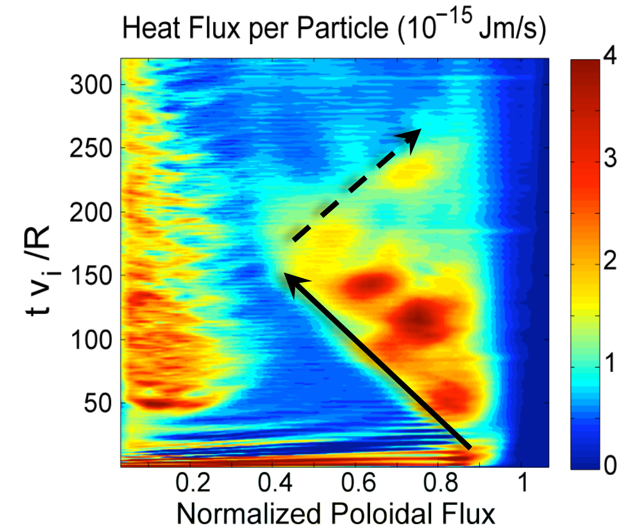
C.S. Chang
New York Univ.

Fusion Science on LCF Platforms

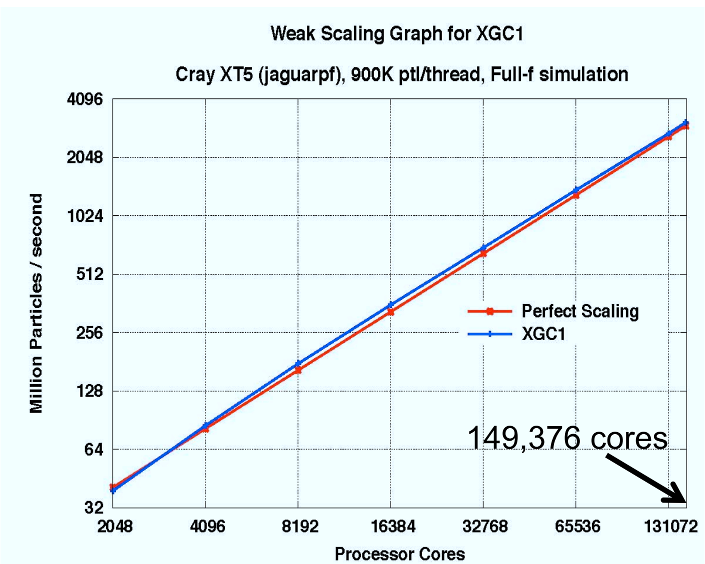
Integrated core-edge simulation of tokamak plasma

Science Objectives and Impact

- **Motivation**: Edge conditioning has been observed to improve the core plasma confinement dramatically in experiments.
- **Goal**: Multiscale nonlocal core-edge simulation of the combined ion-temperature-driven turbulence dynamics and the background ion-temperature profile evolution in realistic DIII-D device geometry
- **Impact**: Practical first-principles predictive simulation capability supporting ITER and DEMO



XGC1 Code Performance



*Science Results***

- First whole volume turbulence simulation in realistic tokamak geometry (2009-10)
- Core turbulence is sum of the incoming intensity from edge and the ambient local fluctuations, self-organizing the temperature gradient in turbulence propagation time scale (similar to experiments).

Image above: turbulent heat flux in time and radius

*** Message: turbulence propagates from edge to core (solid arrow), induces outward heat flux (dashed arrow), and leads to an eventual new self-organized nonlocal state.*

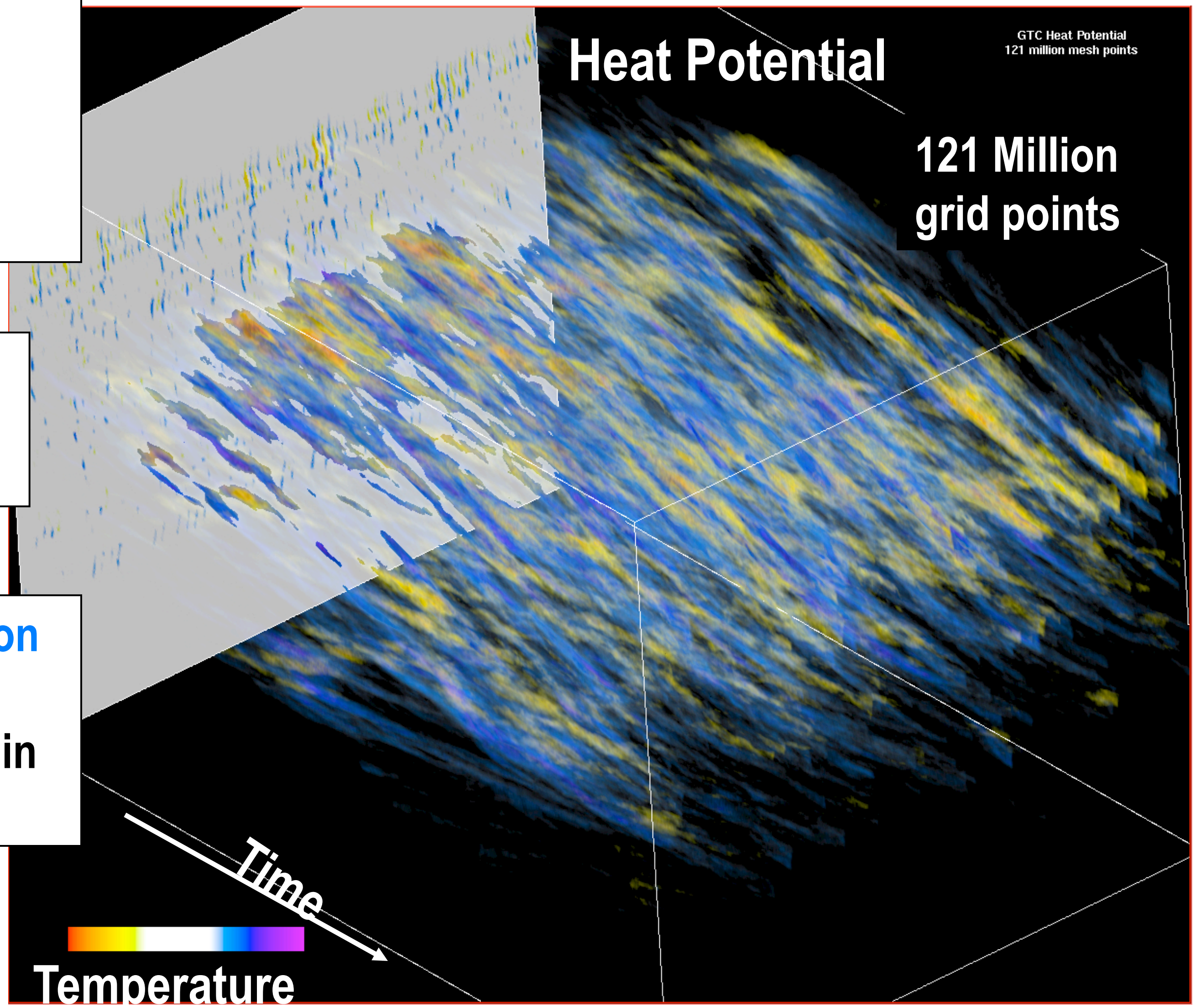
Data Analysis, Management, & Visualization Challenges

Tera- to Peta-bytes of simulation data now generated (Data Management, Data Grid technologies)

Data must be efficiently analyzed to compute derived quantities

New advanced visualization techniques are needed to help identify key features in the data

Particle in Cell Turbulence Simulation



Data Analysis, Management, & Visualization Challenges

- Data-management challenge in some scientific areas already exceeding compute-power challenge in needed resources
- Automated Workflow Environment:
  - Tera- to Peta-bytes of data to be moved automatically from simulations to analysis codes
  - Feature Detection/Tracking to harvest scientific information -- *impossible to understand without new data mining techniques*
- Parallel I/O Development and Support - define portable, efficient standard with interoperability between parallel and non-parallel I/O
  - Massively parallel I/O systems needed since storage capacity growing faster than bandwidth and access times
- Real-time visualization to enable "steering" of long-running simulations
- *Future Experimental Data Challenge*: Current estimates of ITER data size is roughly *40 TB per shot* for long-pulse shots of 400 seconds duration

  -- would demand *100 GB/sec bandwidth*

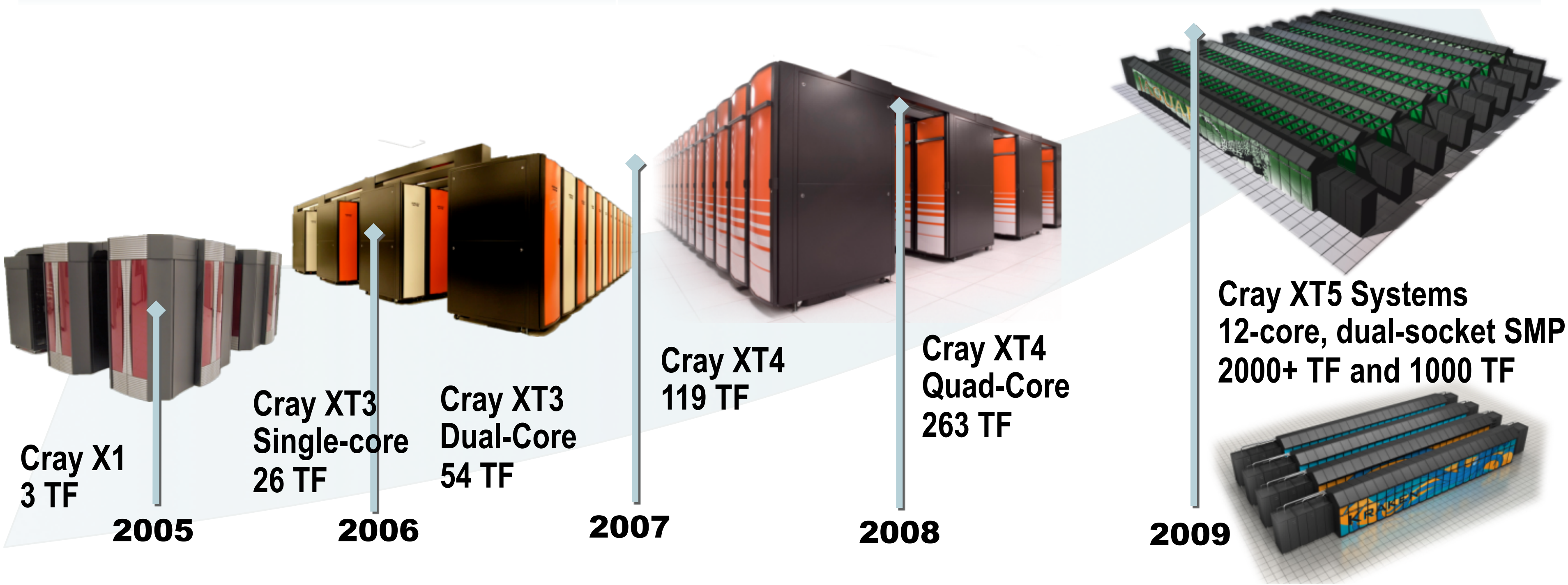
  -- likely need to be able to parallelize at least a significant fraction of this data for streaming

U.S. Leadership Computing Facility (LCF) Resource Capability

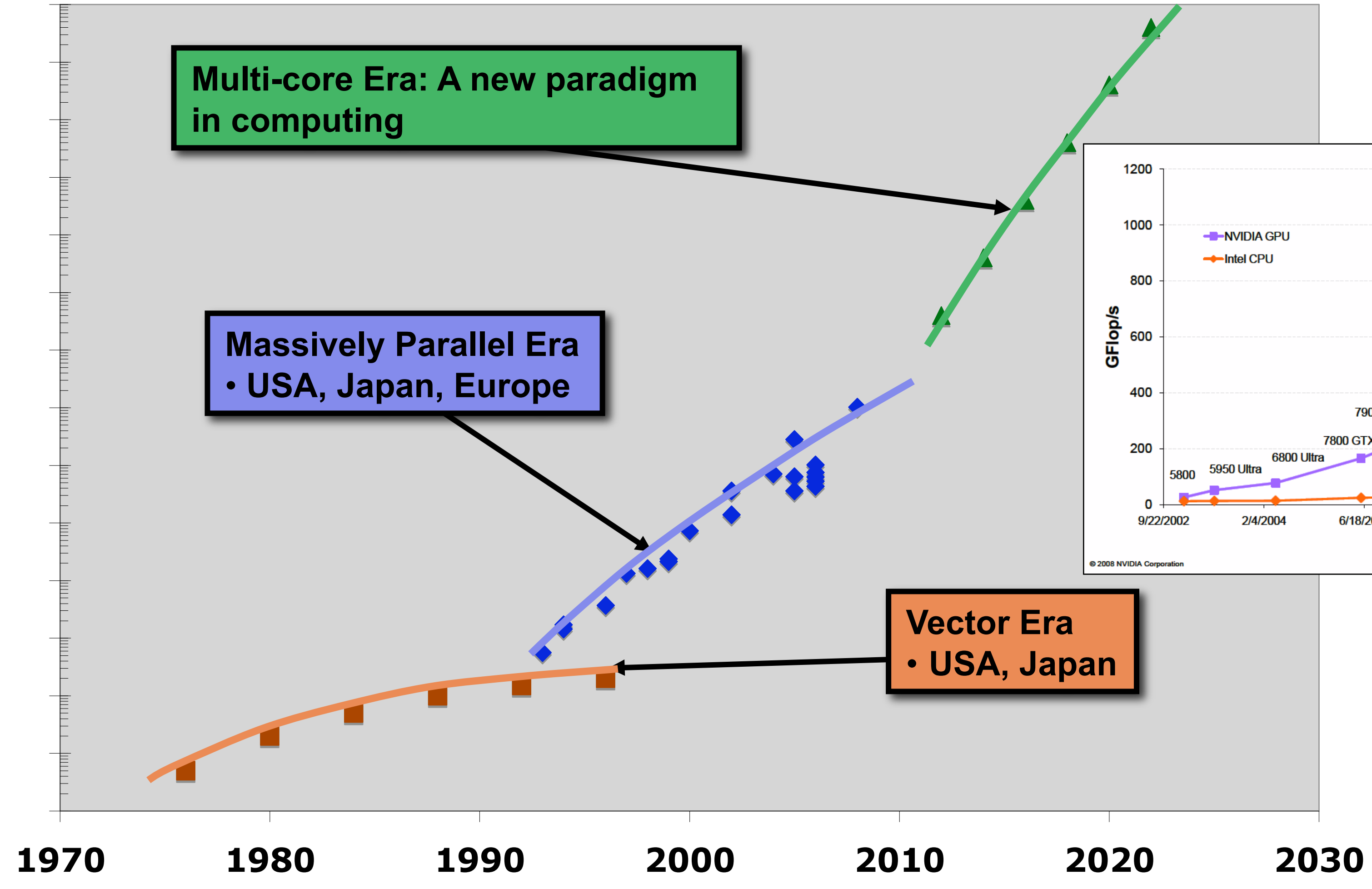
e.g., @ Oak Ridge National Laboratory LCF – "OLCF"

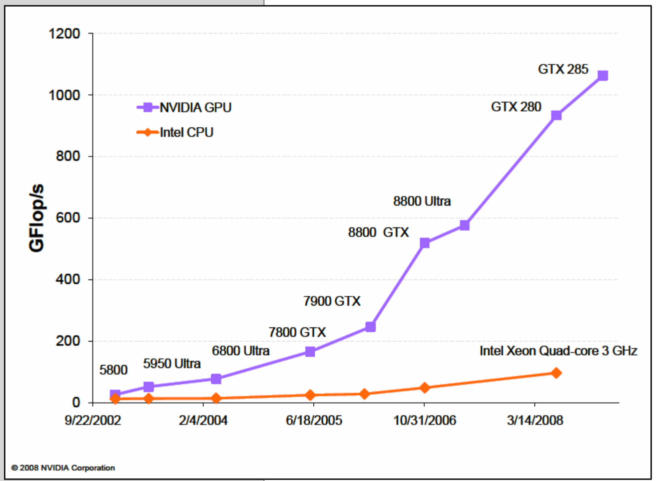
Increased over 750-fold in last 5 years & fusion science applications have been among the largest and most effective consumers

Hardware scaled from single-core through dual-core to quad-core and dual-socket , 12-core SMP nodes	Scaling applications and system software is the biggest challenge



Domain Applications (e.g., FES) Must be Prepared to Exploit *Local Concurrency* to Take Advantage of Coming Hybrid Supercomputing Systems





Future HPC Interests very likely include preparing for Hybrid Architectures

General Purpose GPUs, Floating Point Accelerators, etc.

- Large GPU-based systems springing up everywhere
  - *NSF Track 2D in negotiation with Georgia Tech/ORNL*
  - *Japan: "Tsubame" at Tokyo Institute of Technology*
  - *"Orbit" ORNL 100 TF NVIDIA testbed*
  - *Oil and gas industry deploying large clusters*
- Features for computing on GPUs
  - *Added high-performance 64-bit arithmetic*
  - *Critical for a large system*
  - *Larger memories*
  - *Dual copy engines for simultaneous execution and copy*
  - *Development of CUDA and recently announced work with PGI on Fortran CUDA*
- Large and growing pool of people who know how to program accelerators and who will develop tools
  - *Every laptop has a processor and GPU*
    - *Macintosh, PC, Linux ports of CUDA available*
  - *Most computer science programs now teach GPU programm*

300+ accelerated applications listed on NVIDIA's web site: http://www.nvidia.com/object/cuda_home.html

Current PIC development:
with Exascale Challenges in mind

- Explore different algorithms to improve *data locality* on many-core architecture, including GPUs
  - Sort particles according to position on the grid
  - Various atomic update methods on GPUs: fixed point, mixed precision, full double precision
  - Use texture memory on GPU to store field for particles push
  - Replace expensive replication of local grid data for thread parallelism by atomics and locking mechanisms
- OpenMP tasking to overlap computation and communication in shift routine
- Continue to implement high performance ADIOS I/O calls for data output (also looking at staging…)
- One-sided communication with co-array Fortran (for Cray XE6 Gemini network)

Co-Design Teams: Building Applications for the Exascale

- *US DoE-ASCR is currently promoting this approach*

Code Project Focus

Science Driver
- Model Development
- Physics Component 1
- Physics Component 2
- Physics Component N
- Coupled Multi-physics
- Validation

Code Architecture
- Integration Framework
- Build and Testing Environment
- Programming Model
- Resilience
- Performance/Acceleration
- Scalability
- Dynamic Parallelism
- Task/Data Distribution
- SQA
- Documentation
- Math Libraries
- I/O

Enabling Methods
- Discretization
- Solution Algorithms
- Verification
- Solution Adaptivity
- Sensitivity
- UQ
- Optimization
- Up-scaling Methods
- Time Acceleration

System Architecture
- Apps Micro-kernels
- Performance Models
- Kernel Simulators
- Node Architecture
- Network Architecture
- I/O Architecture
- Runtime Architecture
- Power Efficiency

Science Delivery
- Requirements
- Challenge Problems
- Assessment
- Data Analytics
- Visualization
- Data Management
- Steering & Workflow

FSP expected to include a co-design component & will work with ASCR and FES to share lessons learned and best practices with other co-design science teams within DoE.

U. S. Energy Undersecretary Steven Koonin:
3 November 2009 – American Physical Society Meeting, Atlanta,Georgia

"Validated predictive simulation capability is key to advancing fusion science towards energy"

"Our confidence in validated simulation [close integration of theory, modeling, simulations, and experiments] has to take a major step up

- *moving from description to prediction*
- *use simulation to explore regimes beyond current experimental capabilities*
- *Fusion Simulation Program (FSP) is a start along this path."*

U.S. Energy Secretary Steven Chu:
27 September 2010 – "All Hands Meeting" at the
Princeton Plasma Physics Laboratory, Princeton, NJ

"The world's energy challenge requires a strong continued commitment to plasma and fusion science."
"Progress in fusion has to be grounded in validated predictive understanding: the DoE is clearly interested in your planning and progress for a strong Fusion Simulation Program (FSP)."

Future Challenges and Opportunities

(1) *Energy goal* in FES application domain is to increase availability of clean abundant energy by first moving to a *burning plasma experiment* -- the multi-billion dollar *ITER* facility located in France & involving the collaboration of 7 governments representing over half of world's population
-- ITER targets 500 MW for 400 seconds with gain > 10 to demonstrate *technical feasibility of fusion energy* & *DEMO* (*demonstration power plant*) will target 2500 MW with gain of 25

(2) *HPC goal* is to harness increasing ("Moore's law) power to ensure timely progress on the scientific grand challenges in FES as described in DoE-SC report on *"Scientific Grand Challenges: Fusion Energy Sciences and Computing at the Extreme Scale."*

(3) *Interdisciplinary computational sciences goal* is to leverage advances/"lessons learned" for example from successful U.S. DoE national cross-disciplinary SciDAC Program.
-- impact and heighten the visibility of interdisciplinary research alliances

**** In the FES area, the mission of the FSP (Fusion Simulation Program) is to accelerate progress in delivering reliable predictive capabilities -- benefiting significantly from access to HPC resources – from petascale to exascale & beyond -- together with a vigorous verification, validation, & uncertainty quantification program*