HPC Directions Toward Exascale: An Application Orientation

Kirk Jordan, Ph.D.

Emerging Solutions Exec. & Assoc. Program Director

Computational Science Center

kjordan@us.ibm.com

Outline

- **Background – HPC Challenges & Trends**
- **Some examples that need more HPC**
- **Delivery to the engineer and scientists**
- **Closing Remarks**

Exascale Computing Grand Challenge

1 PetaFlop
72 BG/P Racks

20 PetaFlop
96 BG/Q Racks

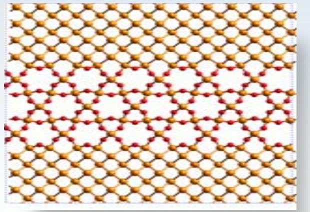
Overall Performance = 1000X
Performance / $ = 1000X

Innovation in technology, architecture, software and algorithms

1 PetaFlop = 1/3 rack
2019

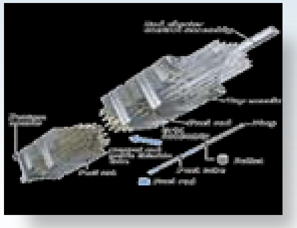
Accelerating Discovery and Innovation in:

Materials Science



Silicon Design

Energy

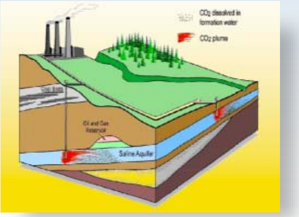


Next Gen Nuclear

Engineering

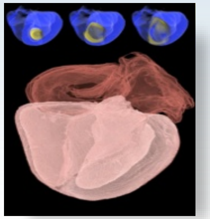
High Efficiency Engines

Climate & Environment

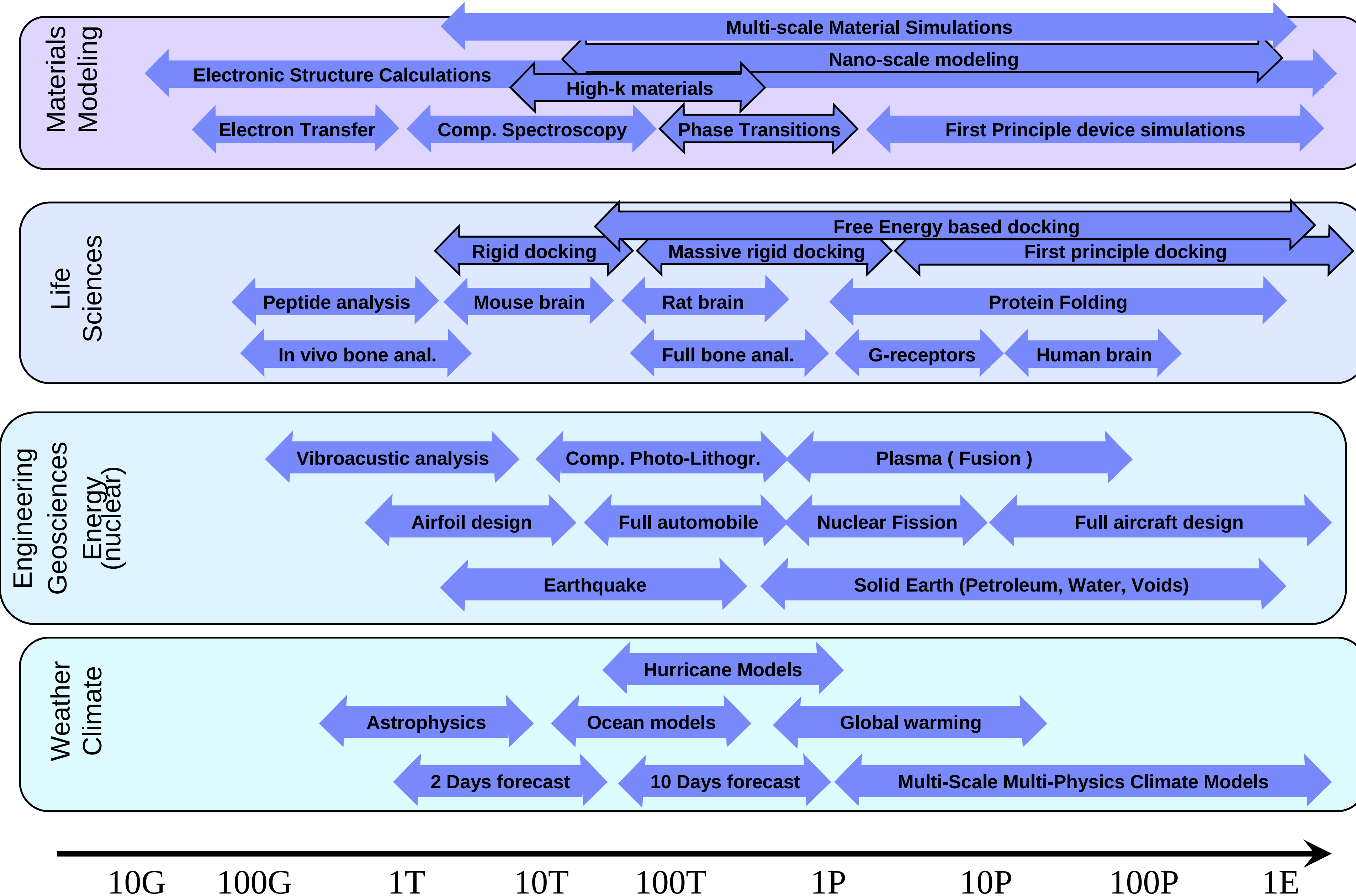


Carbon Sequestration

Life Sciences



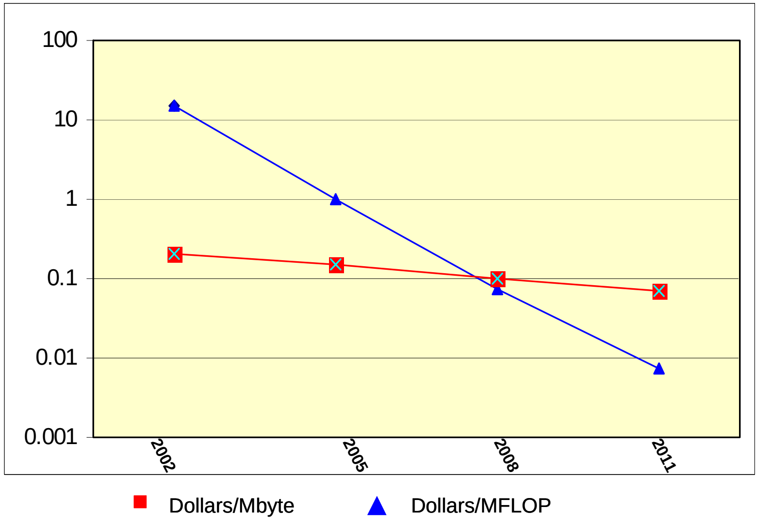
Whole Organ Simulation


Materials Modeling
Multi-scale Material Simulations
Nano-scale modeling
Electronic Structure Calculations
High-k materials
Electron Transfer
Comp. Spectroscopy
Phase Transitions
First Principle device simulations
Life Sciences
Free Energy based docking
Rigid docking
Massive rigid docking
First principle docking
Peptide analysis
Mouse brain
Rat brain
Protein Folding
In vivo bone anal.
Full bone anal.
G-receptors
Human brain
Engineering Geosciences Energy (nuclear)
Vibroacustic analysis
Comp. Photo-Lithogr.
Plasma (Fusion)
Airfoil design
Full automobile
Nuclear Fission
Full aircraft design
Earthquake
Solid Earth (Petroleum, Water, Voids)
Weather Climate
Hurricane Models
Astrophysics
Ocean models
Global warming
2 Days forecast
10 Days forecast
Multi-Scale Multi-Physics Climate Models
10G
100G
1T
10T
100T
1P
10P
100P
1E

Computer Design Challenges

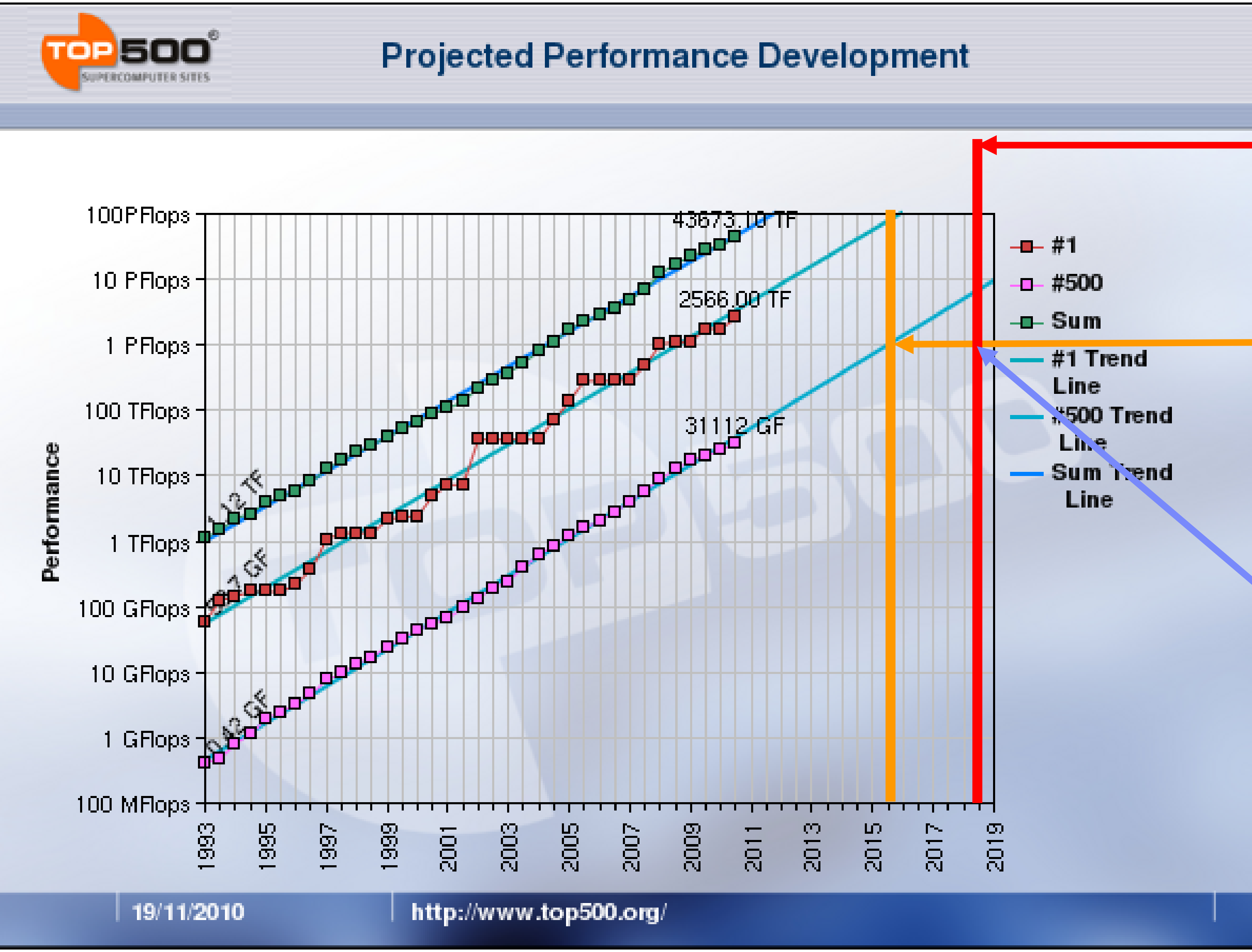
- **Core Frequencies ~**
  - 2-4 GHz, will not change significantly as we go forward
  - 100,000,000 Cores to deliver an Exaflop
- **Power**
  - At today's MegaFlops / Watt: 2 GW needed (~$2B/yr)
  - Power reduction will force simpler chips, longer latencies, more caches, nearest neighbor network
- **Memory and Memory Bandwidth**
  - Much less memory / core (price)
  - Much less bandwidth / core (power / technology)
- **Network Bandwidth**
  - Much less network bandwidth per core (price / core) (Full fat tree ~$1B to $4B)
  - Local network connectivity
- **Reliability**
  - Expect algorithms / applications will have to permit / survive hardware fails.
- **I/O Bandwidth**
  - At 1 Byte / Flop, an EXAFLOP system will have 1 EXABYTE of Memory.
  - No disk system can read / write this amount of data in reasonable time. (BG/P 4TB ~1min but disk array ingest at ~15min)

GFLOPs vs DRAM Price Reductions



- **Exascale Computing**
  - O(100 M) compute engines working together
- **Capability delivered has the potential to be truly revolutionary**
- **However**
  - Systems will be complex
  - Software will be complex
  - Applications will be complex
  - Data Centers will be complex
  - Maintenance / Management will be complex

Trends in Computing Performance



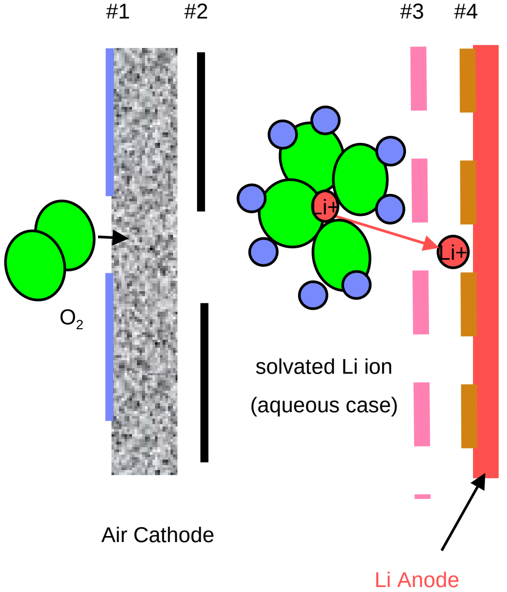
1EFlop ~2nd half of 2018 – est cost $200M

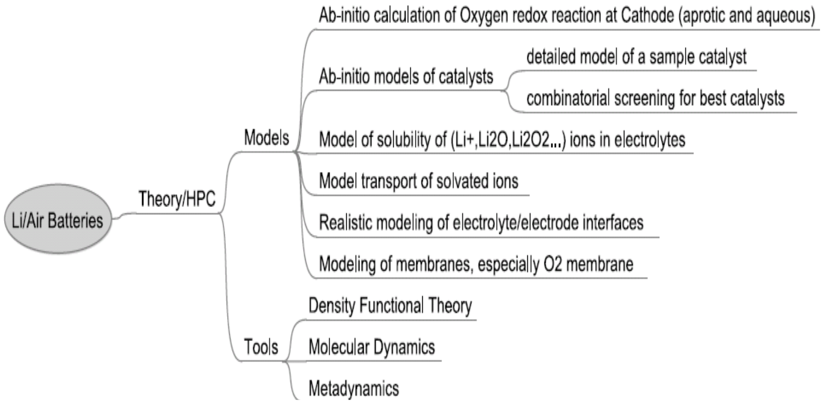
1PFlop ~2nd half of 2015 is bottom of Top500

1PFlop ~2nd half of 2018 – est. cost < $200K

Exascale for Ultimate Batteries Design

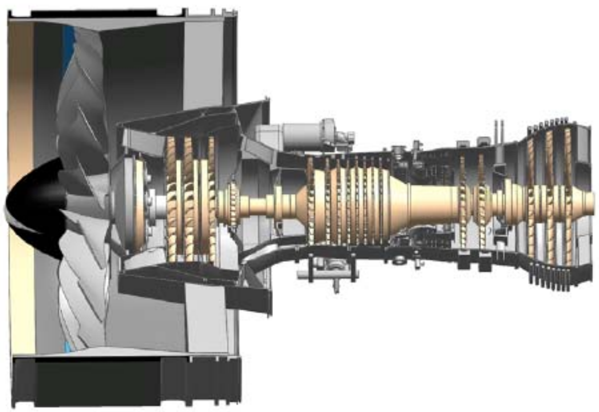
- Li/Air batteries are emerging as a promising technology that could provide a sufficiently high energy density for automotive applications.
- There are several scientific issues that have to be solved hinder the development of rechargeable and high energy density Li/Air batteries. These issues are very difficult and expensive to be solved with experiment only.
- Battery research combines the three most challenging aspects of computational physics
  - ***non-equilibrium***
  - ***Multiphase***
  - ***multiscale* (in space *and* in time)**
- A complete model will require 100's of Petaflops (Exascale) computing.





Pratt & Whitney on Intrepid
INCITE PI : Peter Bradley, Pratt & Whitney

- INCITE 2006-2007 technologies now being applied to next generation low emission engines.
- Important simulations can now be done 3X faster
- A key enabler for the depth of understanding meet emissions goals



Argonne NATIONAL LABORATORY 7

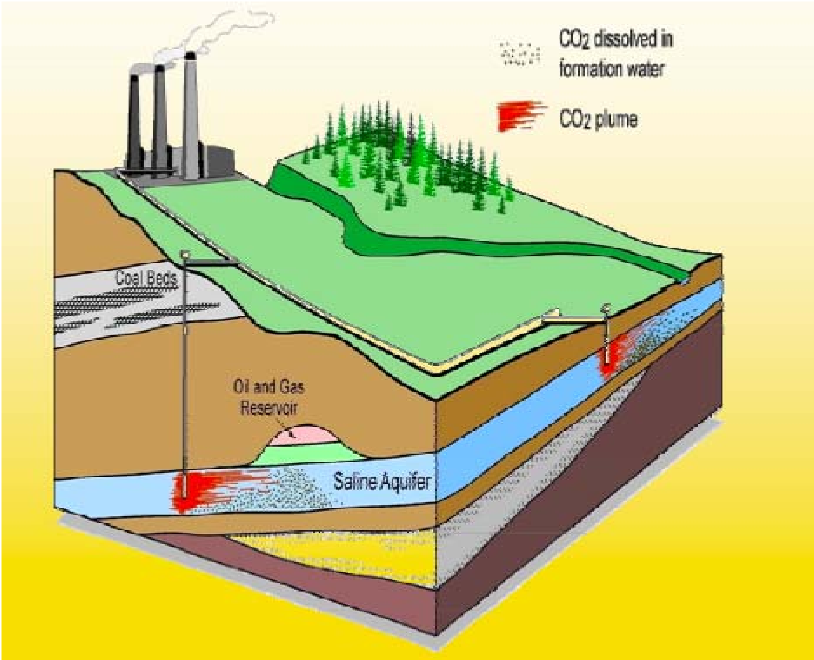
Current focus:

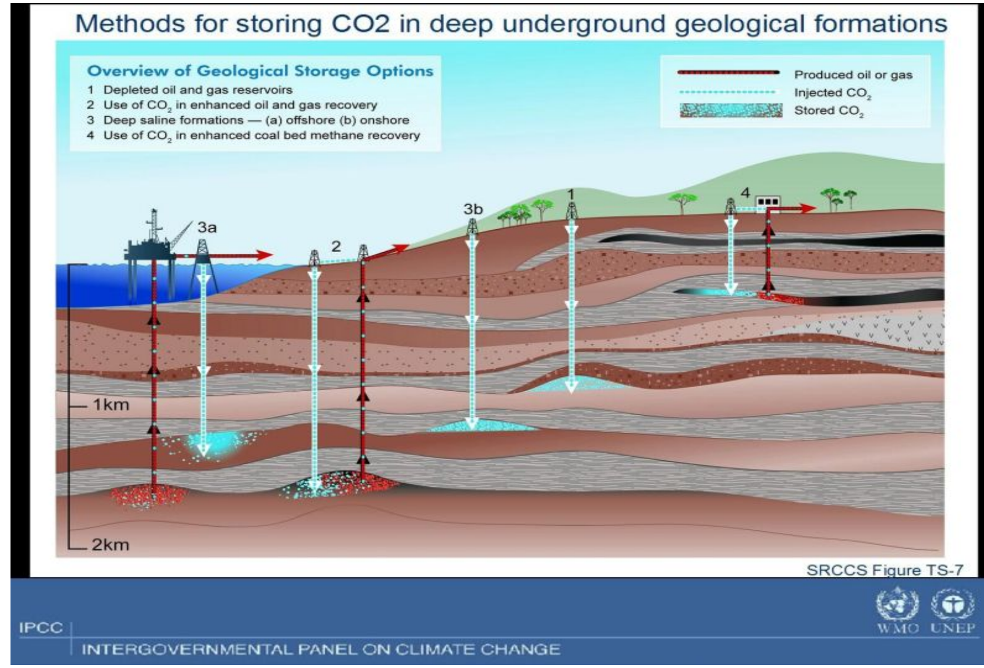
combustion fluid flow modeling only – a petascale problem

Required future enhancements – approach exascale requirements:

Include molecular structure of fuel to assess cleanliness and efficiency of fuel burn
Include engine as part of aircraft model to understand actual performance
Include analysis of effects of different angles of attack and different flight scenarios

CO_2 Sequestration - Storage Options - Model

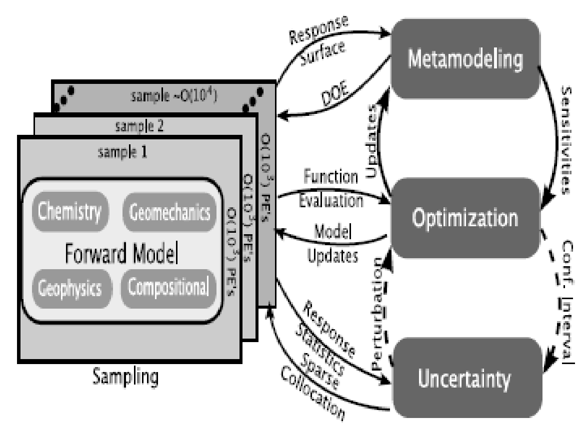




Computational complexity of a 3-D, 3-phase flow and transport in porous media problem on a moderately large subsurface model:

- $N=10^7$ Discretization elements – 3 unknowns/grid element
- N_t=1000 timesteps
- N_{MCS}=1000 Monte Carlo simulation for uncertainty assessment

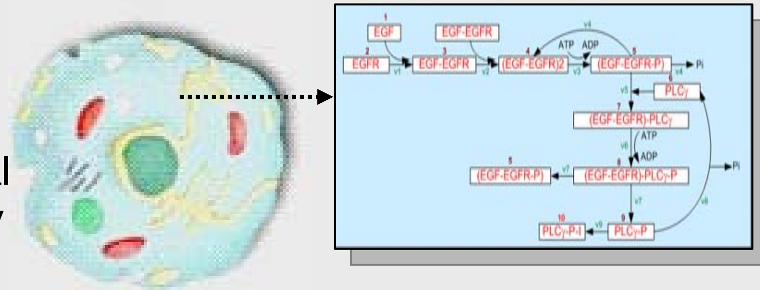
=> FLOPS = $N_t * N_{MCS} (3N)^2 \sim 10^{21}$ (1000 ExaFlops)



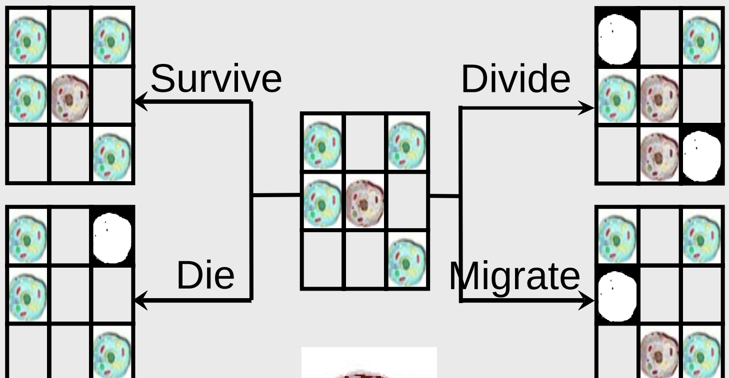
Source: M.F. Wheeler, et. al., NSF PetaApps Proposal 2008

Multiscale Cancer Modeling

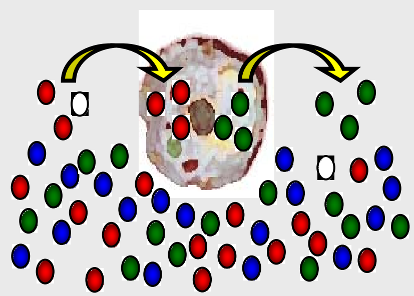
- Model individual tumor cells, each with a unique genome parameterizing a realistic compartmental gene regulatory pathway model

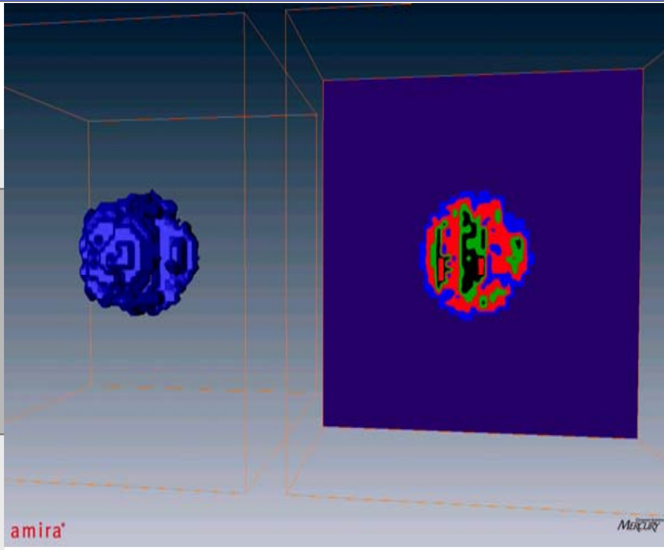


- A continuum model description of the tissue environment in which tumor cells survive, die, divide or migrate via stochastic diffusion and chemotaxis

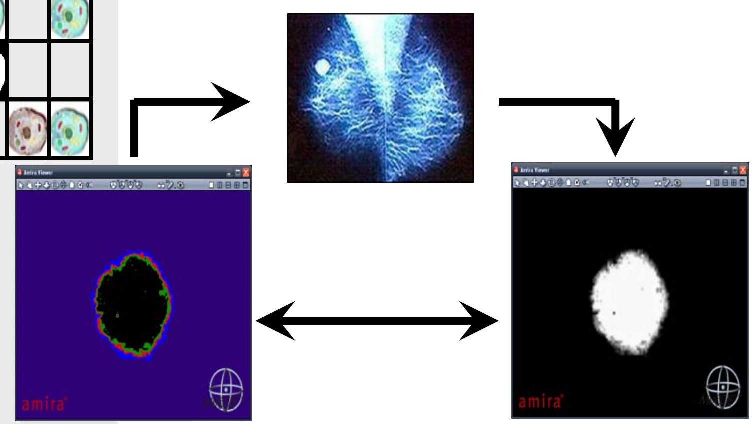


- Molecular interactions between tumor cells (controlled by the gene network) and their tissue/organ environment





Clinical Imaging (e.g. MRI)



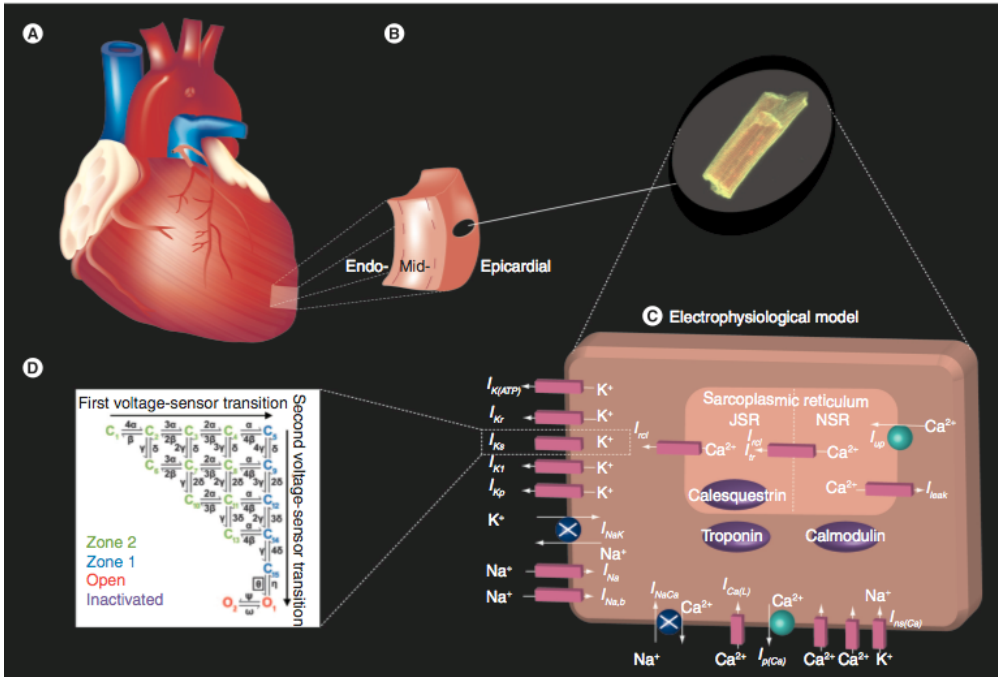
Model Prediction of Tumor Cell Phenotype Distribution

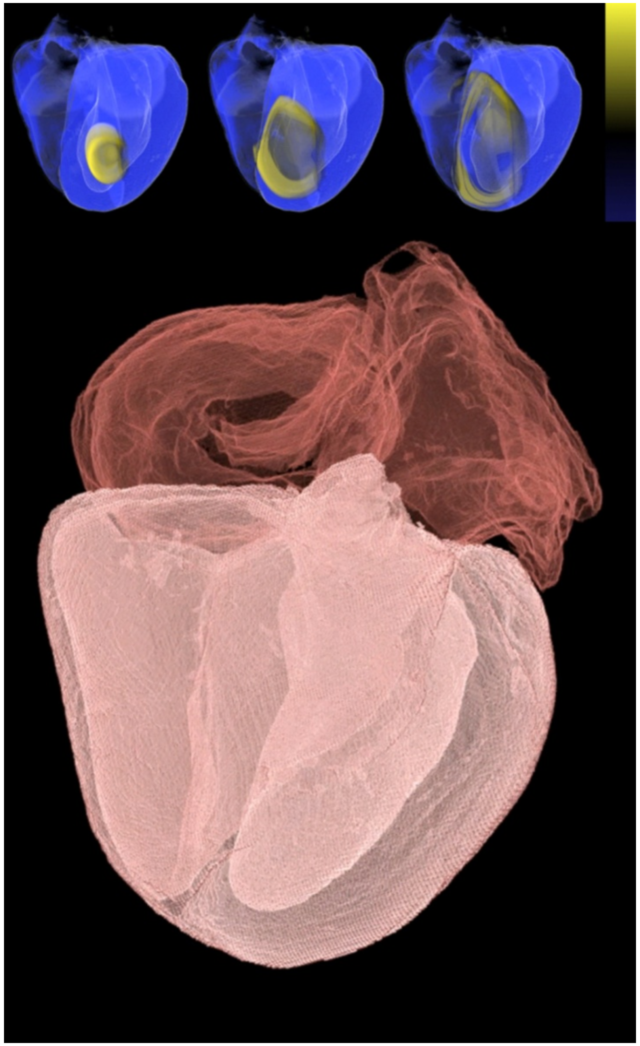
Tumor Cell Density Matched to MRI

John Wagner – IBM, Brian Skjerven – WPI/IBM/U Minn
Thomas Deisboeck, Le (Adam) Zhang – MIT/Harvard/MGH

Computational Medicine: Whole Organ Simulation

- **Predictive Toxicology**
- **Multiscale Model of Organs**
  - From protein function through to cell function through to tissue function through to macroscale organ modeling.
- **Multiple model components and scales require Petascale to Exascale compute capability**
  - Usefulness requires “turnkey” modeling environment where many variations and scenarios can be attempted by the medical or pharmaceutical researcher quickly and accurately
  - Further increases the computational requirements





Fundamental Business Challenges for 100 PF – 1 EF

- **Break out from Research use to large scale Business / Industrial Use.**
  - Traditional research customer base alone cannot justify hardware development costs
    - US Government currently considering $1B investment as its contribution to development costs.
  - Essential to develop useable cost-effective solutions for industrial and commercial domains
    - IBM Smart Planet Solutions, Engineering, Finance, Geophysics, Materials, Energy, Climate …
- **Workflows will be extremely complex, and will require heterogeneous systems solutions**
  - at the processor level
    - heterogeneous cores to service different application functions and algorithm requirements,
  - at the systems level
    - mixed system types – eg. SMP for DB, massive concurrency systems for capability, commodity systems for throughput, pre and post processing,
  - and at the data center level
    - tightly coupled storage, visualization, and web service provision all connected directly to capability systems.
- **Data Scales (Exabytes) will require fundamentally new approaches**
  - Recognition that data can't move from data center
  - Load / Store costs to move data from storage to memory will dominate workflow
    - Need innovative approaches to manage this.
  - Integrated system solution which spans from Desktop to Exascale solution.

Workflow Taxonomy

	Description	Examples	Application Set	Team
Capability				
	Calculations not possible on smaller machines Typically a single application Disparate scales define time to solution	Protein Folding Ab-Initio Materials Modeling 1km grid global air circulation	Single Core Application Pre/Post Processing Steps	Small Core Team Team has expert HPC knowledge Team will have significant code knowledge
Complexity				
	Multiple applications cooperating on single workload Coupling between applications	Combined CFD + Structural Cell to Organ Models Environmental Water Management	Multiple Core Applications Complex Linkages Between Apps Data Prep and Analysis	Multiple core teams Mix of HPC, Science, Domain groups Development activities to establish code linkages
Understanding				
	Multiple executions of complex workflow Optimization, Sensitivity Analysis	Integrated Global Climate Structural, CFD, Combustion for Engine Design Aircraft Airflow + Structural	Robust Individual Codes Significant Test and Verification Frameworks Complex Workflows Significant Database Dependencies	Production Quality Codes Primarily non-HPC customers Commercial Grade Service Delivery

Traditional Laboratory Research

Prototype use only

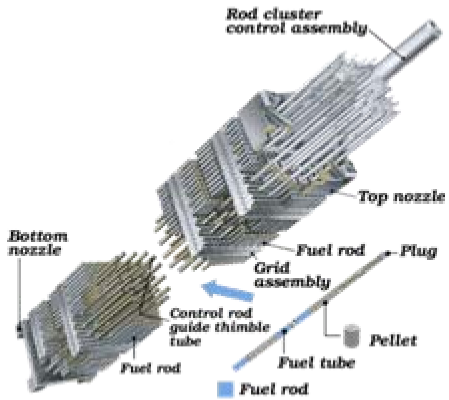
No commercial impact

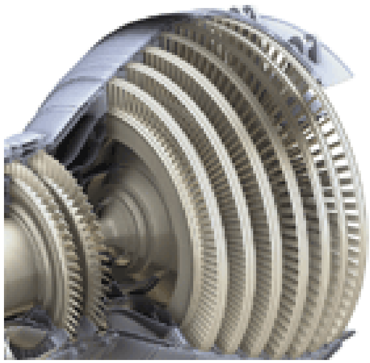
Commercial Opportunity requires sophisticated software management, solutions and services

Examples: Nuclear Energy, Combustion Applications

- **Nuclear:**
  - Next generation reactor design and optimization
  - Develop technologies to improve reliability, safety, increase reactor usable life
  - Develop a sustainable fuel cycle
  - Improve operational management capability
  - Reduce development costs

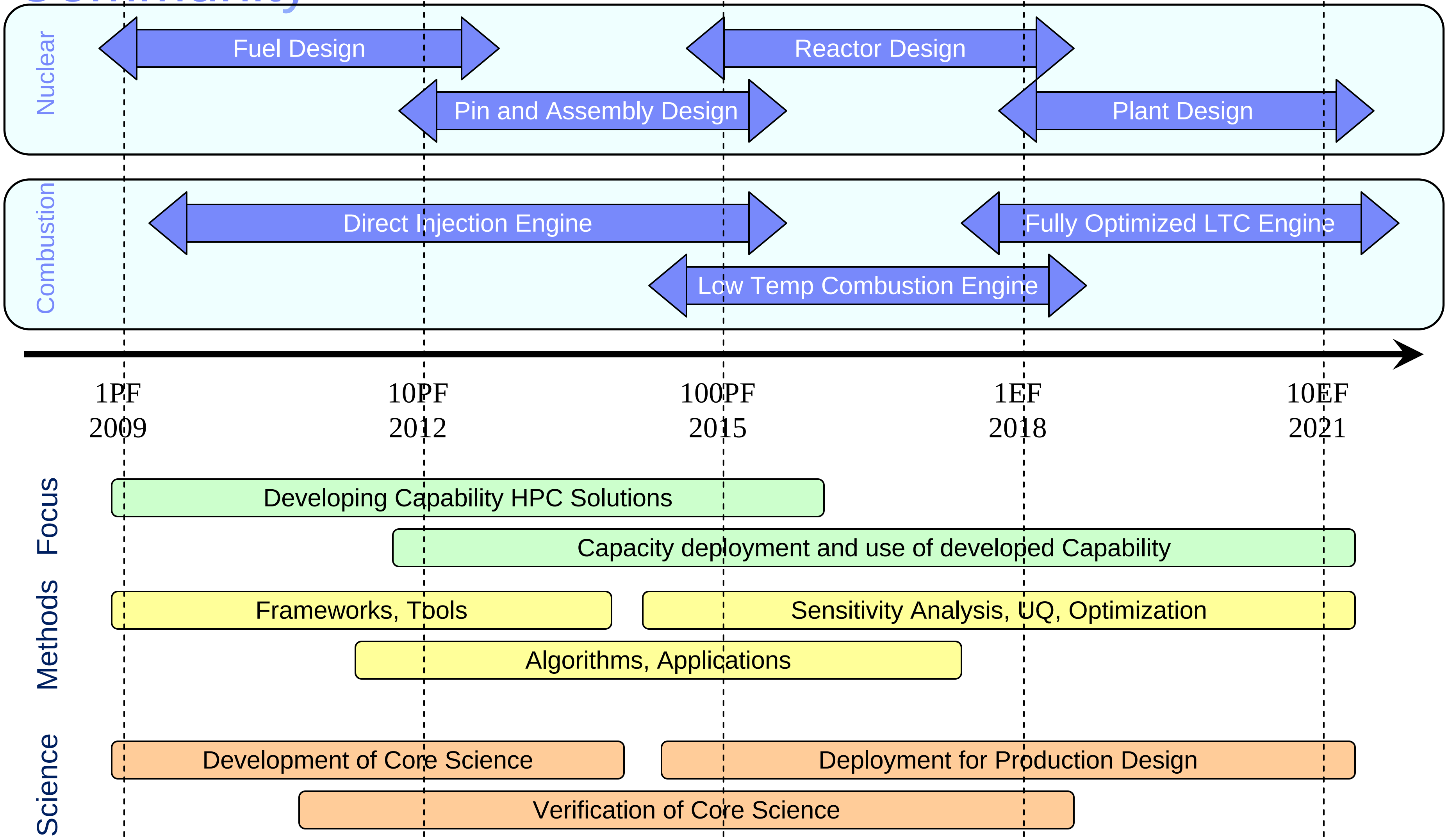
- **Combustion:**
  - Gas Turbines, Gasoline and Diesel Engines
  - Increase efficiency,
  - Reduce emissions,
  - Broaden usable fuels
  - Reduce development costs.





- In both cases, multiscale multiphase physics problem.
- Includes Computational Fluid Dynamics, Thermohydralics, Structural Mechanics.
- Coupling of different physical domains and simulation approaches a significant issue.
- Nuclear codes need to include also neutronics, materials aging under neutron bombardment.
- Combustion codes include fuel injection, combustion analyses.

Roadmap for Research and Development in HPC Community



Early Impacts

- **Multiple systems in the 10-20PF scale arriving 2011, 2012**
  - These systems will have many of the architecture features we expect at Exascale.
  - High concurrency (e.g. BG/Q for 20PF will have 100M computational threads).
  - Evolving programming models & workflow models
- **Significant development on algorithms, programming models, coupling, frameworks, workflows will occur in the 2011-2017 timeframe**
- **Some key scientific deliverables will also be achieved e.g.,**
  - In Biology:
    - full "in-silico" simulation of organs and their interactions.
    - exploratory modeling of different biological systems integrated with biological data for developing novel therapies.
- **Architecture for solutions enabling technologies will be defined**
  - Data requirements, data management solutions
  - Complete service solutions for delivery to non HPC communities
- **Opportunity to establish PathFinder Centers for Innovation, Solutions Development and Solutions Delivery**

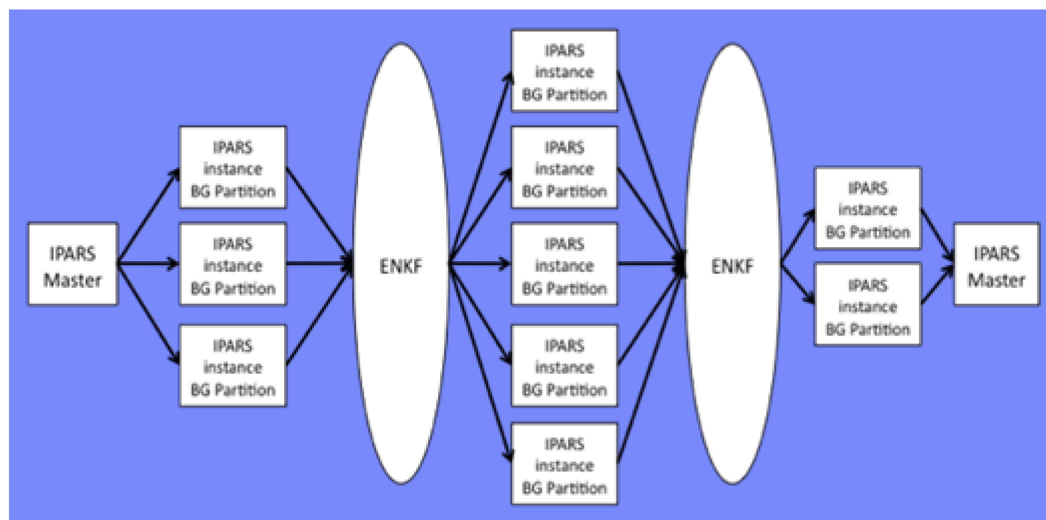
The Vision – Easy Access Delivered Anywhere w/ Elastic Resources

- HPC As-A-Service – the Vision
  - Immediate/Elastic Resources
  - Dynamic Scaling
  - Efficient Utilization
  - Scalable (Multiple BG)
  - Fault Tolerance
  - Delivered to mobile devices – the user's entry point (this is changing from Desk top/Laptop to tablet/smart phone)

- Applicable in many science & engineering domains – e.g.
  - Healthcare & Life Sciences
  - Auto, Aero Engineering
  - Petro, Chem Engineering

- Target at non-HPC users built by HPC Experts
  - Peta & Exascale Systems – Multiple realizations to explore parameter space by engineers, reduce risk etc.

- Deliver to Field Engineers –
  - Mobile devices –
    - Set up runs
    - Deliver results

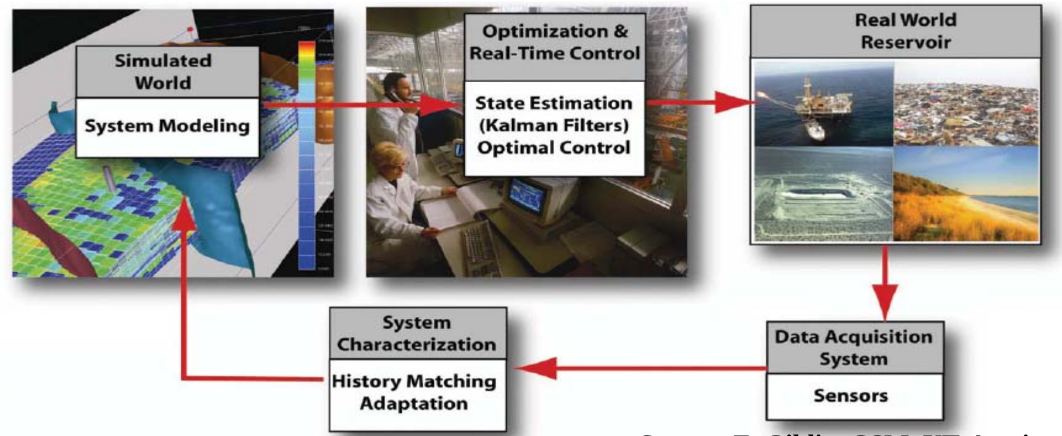


Proof of Concept – Example Problem - - Reservoir Management – Optimal Control while Reducing Risk

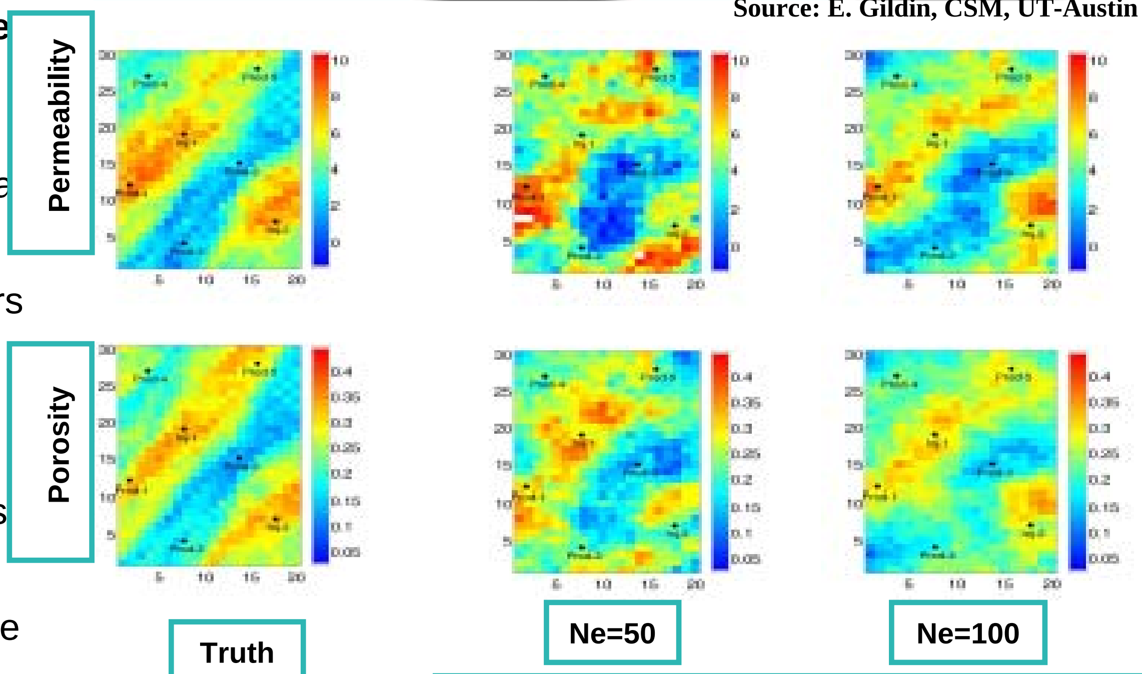
- Smarter Reservoir Management –
  - **Continuous Measurement and Data Analysis for Reservoir characterization and management**
  - **Deliver results where needed – even mobile locations**
  - **Toward Peta² Appliance – Multiple Realizations**

- HPC Methology –
  - **Ensemble Kalman Filter w/ Multiple (parallel) Reservoir Simulations (EnKF) – two steps:**
    - **Forecast step -** set of reservoir simulations to predict data at upda step
    - **Update step -** compute Kalman gain matrix and update parameters
  - **An example Problem**
    - 5 production wells and 2 injection wells
    - Total simulation time is 2700 days with update interval of 30 days
    - Observation data are bottom hole pressure (BHP), oil production rate and water-oil ratio (WOR)

Closed Loop Reservoir Management



Source: E. Gildin, CSM, UT-Austin



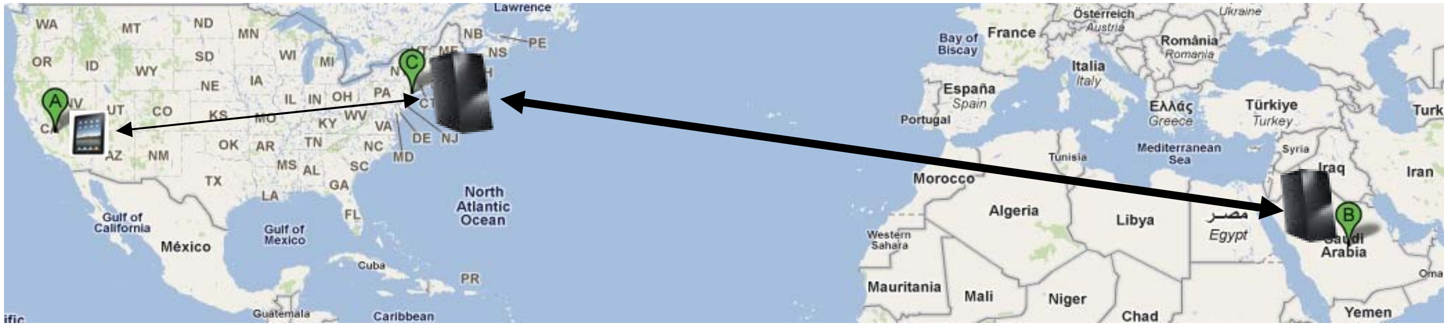
Final estimate of permeability and porosity fields for two different ensemble sizes

IEEE Scale2011 Challenge at CCGrid2011 Conference

- **The Fourth IEEE International Scalable Computing Challenge (SCALE 2011):**

- **The Finalists - - Scale Challenge Presentations (26th May, 2011):**
  - Moustafa AbdelBaky, Hyunjoo Kim, Manish Parashar, Kirk Jordan, Hani Jamjoom, Vipin Sachdeva, James Sexton "*Scalable Ensemble-based Oil-Reservoir Simulations using Blue Gene/P-as-a-Service*".
  - Bhanu Rekepalli and Aaron Vose "*Petascale Genomic Sequence Search*".
  - Ciprian Docan, Fan Zhang, Manish Parashar, and Scott Klasky, "*Coupling Scientific Fusion Simulations at Extreme Scales*".
  - Mathieu Djamai, Bilel Derbel, Nouredine Melab "*A Large-Scale Pure P2P approach for the B&B algorithm*".
  - Suraj Pandey, Letizia Sammut, Andrew Melatos, Rajkumar Buyya "*Scaling Executions of Multiple Workflows from Multiple Users*".

- **1st Place Winner**
  - Moustafa AbdelBaky, Hyunjoo Kim, Manish Parashar, Kirk Jordan, Hani Jamjoom, Vipin Sachdeva, James Sexton "*Scalable Ensemble-based Oil-Reservoir Simulations using Blue Gene/P-as-a-Service*".



Summary

- **Next 10 years:**
  - HPC Capability evolving
    - Fidelity and time to solution relevant for industrial / commercial use
    - Hardware costs continue to fall
  - Focus shifting from Hardware to Solutions
    - Expertise now critical
    - Economic opportunity is development and delivery of robust solutions and services
- **We will have succeeded when**
  - we stop talking about architecture
  - we focus on real impact: Research, Industry, Business
- **Opportunity**
  - Brand New, Green Field Landscape!
  - Focus shifts from single applications to solutions and services
  - Significant opportunities for entry of new players
  - Economic impact is critical

In the end, it's not about the technology; It's what you do with it that counts

Thank you for your attention

Join us as we:

- **Continue to innovate across the whole systems stack to deliver leadership in performance and usability**
- **Help solve problems that are currently intractable or not cost-effective**
- **Accelerate discovery in science, engineering, and business**