What are the Opportunities and Challenges for a new Class of Exascale Applications?

Douglas B. Kothe

Deputy Associate Laboratory Director for Computing and Computational Sciences
Oak Ridge National Laboratory
&
Application Development Director, Exascale Computing Project

Cray Users Group (CUG) Conference 2017
Redmond, WA
May 9, 2017

EXASCALE COMPUTING PROJECT

National Nuclear Security Administration

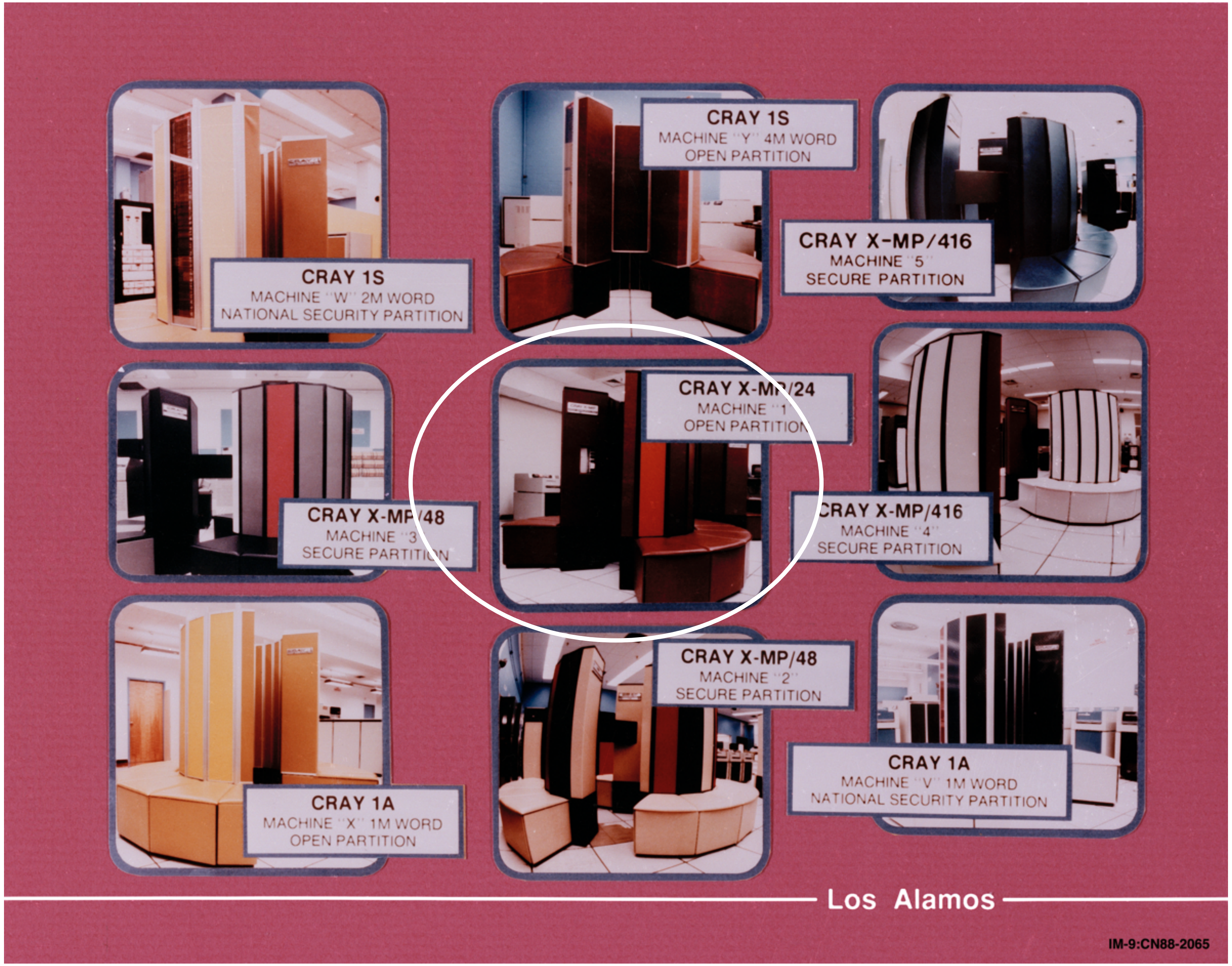
U.S. DEPARTMENT OF ENERGY | Office of Science

www.ExascaleProject.org

Message to Apps Developers: Take the Red Pill!

This is your last chance. After this, there is no turning back. You take the blue pill - the story ends, you wake up in your bed and believe whatever you want to believe. You take the red pill - you stay in Wonderland and I show you how deep the rabbit-hole goes.

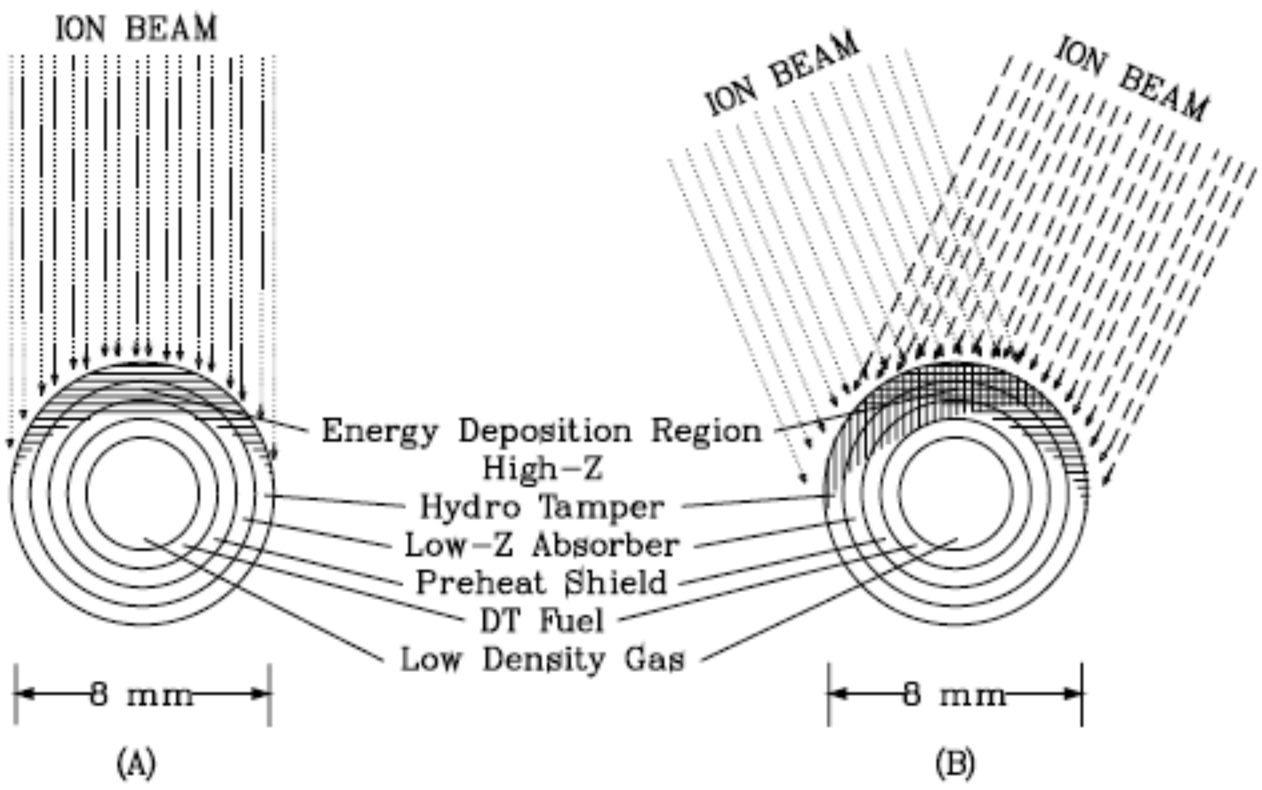
Development Platforms
Cray Systems at Los Alamos in 1980s



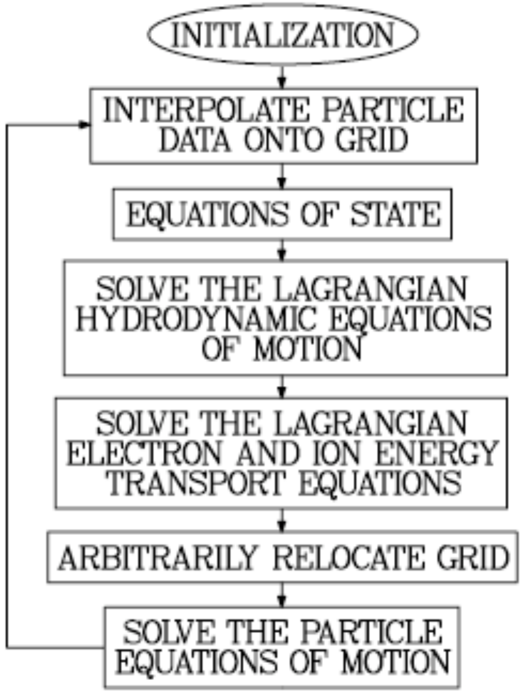
Memories
- DDT debugger
- Cray Assembly Language (CAL)
- Fortran Command Language (FCL)
- Tiny (credit: John Norton, LANL)
- Common File System (CFS)
- Bank points

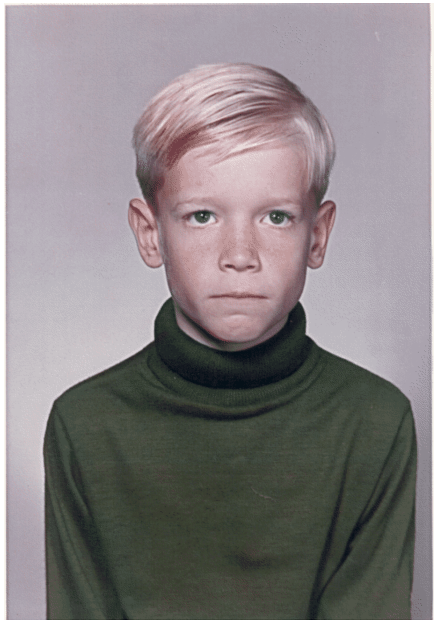
EXASCALE COMPUTING PROJECT

Imploding ICF Target Design (1985-1987)



Fluid-Implicit-Particle (FLIP) particle-in-cell (PIC) method, now a base technology for Disney Animation Studios!





Me (circa 1984):
Prior to engaging in app development

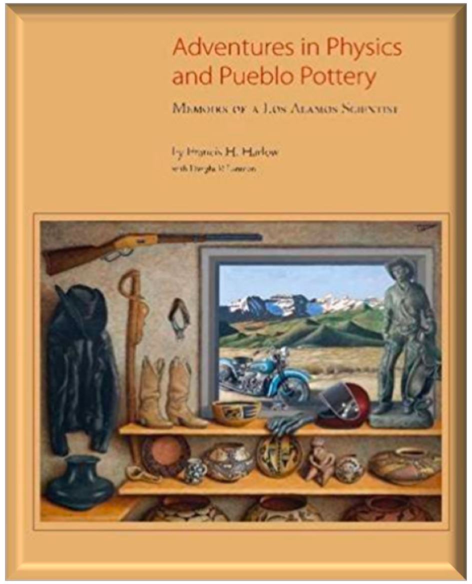
Unique characteristics of numerical method

- 2nd order time/space PIC method (FLIP) with inherent stability and fluid-like collisionality
- Adaptive grid (*not* AMR – it was the pre-AMR days)
- Discrete ray-tracing for ion beam penetration & energy deposition
- Natural ability to track interfaces via particle identity
- Innate sensitivity to unstable hydrodynamics (particle/grid Eulerian/Lagrangian duality)

EXASCALE COMPUTING PROJECT

Particle-in-Cell (PIC) Segway

A life well lived: Francis H. ("Frank") Harlow (1928-2016)



I have lived a lucky life, being in the right place at the time when something new and exciting was happening.
- Francis H. Harlow

- PIC was invented by Frank Harlow (LANL) in mid 1950s for fluid dynamics phenomena
  - Pueblo pottery, paleontology, plot day, motorcycles, leather gloves, fire jumping, Tony Amsden, turbulence, literature reading, hugs, 50th anniversary
  - More: JCP 195 (2004): 414-433
- Motivated by shock physics for weapons phenomena
  - LANL reports: LAMS-2082, LA-2139, LA-2301, LA-2806, LA-3144
  - *Computer Experiments in Fluid Dynamics*, Scientific American (1965)
- Move into plasmas
  - Ned Birdsall, Bruce Langdon, Clare Nielson, Richard Morse (who did what unclear)
- PIC (for fluids) revitalized in 1985 – 1995 (Brackbill - "FLIP")
  - Extended to solids as the "Material Point Method (MPM)" in early 1990s
  - Disney's Frozen, Moana, etc.

Exascale “Killer App”: Simulation “Predictivity”?

- Predictivity = accurate prediction (with quantified uncertainty) of the behavior of complex systems
- Exascale moves us from model systems to real systems
- Greater confidence in decision and policy making
  - Ex: “95-95” - 95% chance the simulation will match 95% of real-world data
- Realizing predictivity requires
  - Additional physics and resolution
  - Ensembles of petascale simulations
  - Uncertainty Quantification (UQ)
  - Application verification and experimental validation
- Strategic areas are poised to cross the predictivity threshold at exascale
  - Can be formally framed with models, e.g. Predictive Capability Maturity Model (PCMM)

Exascale Applications Can Deliver Transformative Technologies

Exascale MARVELs

- **National Security**
  - Materials in Extreme Conditions; Stockpile Stewardship; Electromagnetics in Hostile Environments; Hypersonic Re-Entry
- **Energy**
  - Wind Plants; Small Modular Reactors; Clean Fossil Fuel Combustion; Subsurface Reservoir & Storage; Biofuel Catalysts
- **Transformative Materials**
  - Additive Manufacturing of Metals; Materials for Extreme Environments; Quantum-Level Control of Transition Metal Oxides
- **Economy**
  - Resilient Power Grid; Smart City Dynamics; Earthquake Risk Assessment; High Efficiency & Low Emission IC Engines
- **Earth Systems**
  - Drought Resistant Crops; Metagenomics Analysis of Biogeochemical Cycles; Assessment of Regional Water Cycles
- **Scientific Discovery**
  - Magnetic Fusion Reactor Stability & Control; Origin of Chemical Elements in Universe; Assess Standard Model for Nuclear Force; Plasma Wakefield Accelerator Design; Elucidating Cosmological Structure Formation
- **Healthcare & Biology**
  - Precision Medicine for Cancer; Light-Source Enabled Insights into Protein Dynamics

Exascale Bridges the Gap!

Simulation Gap	Needed to Address the Gap	Impact of Gap Remaining
Simplified or incomplete physics	• Higher fidelity models for all relevant physical phenomena • Compute memory and speed necessary to accommodate their numerical solutions and their non-linear coupling requirements	• Cleaner, more efficient combustion engine designs delayed, or not discovered • Inefficient agriculture & energy production sector planning - insufficient regional water cycle assessments • Fewer (likely more expensive) options for CO_2 sequestration, petroleum extraction, geothermal energy, due to lack of understanding long-term reservoir-scale behavior • Astrophysics discoveries (origin of elements in the universe, gravity waves) remain elusive
Simulation detail insufficient	• Larger simulation domain • Finer partitioning of the simulation domain • Computing speed, memory and I/O to accommodate larger simulation domains with finer partitioning • Workflows integrating an advanced technology and software toolset • Data streaming methods to steer simulation partition real time	• Conservative earthquake retrofits more costly and over-designed • Inability to predict and control material properties at the quantum level precludes advances in high temperature superconductivity • Higher power grid operating margins and lost cost savings potential • Wind plant efficiencies lag theoretical energy extraction potential by 20-30% • Delays in scale-up of new chemical looping reactors for clean fossil fuel combustion • Key cosmology and nuclear physics discoveries – dark matter/energy, standard model of particle physics, inflation of the universe – remain elusive
Can only simulate subset of scenarios of interest	• Robust and fast algorithms for the numerical solution of coupled multi-physics systems that expand limits of applicability • Workflow tools to analyze simulation ensembles	• Continued high rejection rates (80%) of additively manufactured metal alloy parts with tight specifications, increasing waste and cost • Tools for retrofitting & improving urban districts remain empirical • Limited ability to influence ITER design decisions and ultimate operations
Uncertainty insufficiently quantified	• High fidelity in situ data analytic techniques for reliable quantification of simulation uncertainties and sensitivities • Workflow tools to analyze simulation ensembles	• Protracted deployment of small & advanced nuclear reactors • Delays in design of small, low cost, and ultra high intensity plasma wakefield accelerators • Engr & materials design requires more expensive & time-consuming physical experiments
Unable to intersect to design cycle	• Improved computer throughput to shorten simulation turn-around times • Large ensembles of calculations to enable optimization thru rigorous exploration of design space	• Continue deploying materials for extreme environments with a cumbersome make-test cycle rather than by atomistic design • Efficient in silico design of new chemical catalysts not realized
Inadequate analysis and knowledge discovery in big data	• Scalable AI (deep learning) networks of large size and complexity for efficient training on big datasets. • Efficient computational workflows seamlessly integrating simulation, data analytics, and big datasets	• New cancer biology treatment options missed or delayed due to unrealized understanding of precision oncology • Analysis and reduction of data deluge from experimental science facilities requires weeks (instead of near real time) • Manufacture of new products and chemicals delayed or missed because microbiome DNA sequencing unable to keep up with available data • Inability to quantify uncertainties via a nexus of simulation and experimental facility data

Convergence of DAC and M&S* ?

- **Data Analytic Computing (DAC)**
  - Inferring new information from what is already known to enable action on that information
- **Modeling and Simulation (M&S)**
  - Insights into interactions of the parts of a system, and the system as a whole, to advance understanding in science and engineering and inform policy and economic decision-making
- **DAC & M&S have traditionally relied on different hardware and software stacks**
- **M&S community is increasing use of large-scale data analytics, demanding a more dynamic interaction between analysis and simulations**
- **DAC community is demanding increased computational intensity, hence facing barriers to scalability along with new demands for interoperability, robustness, and reliability of results**
  - **Physics-Informed ML (PIML)?**
- **Coherent platform for M&S & DAC benefits both while maximizing returns on R&D investments**
- **Primary challenge: software layers of the HPC environment**
  - A more agile and reusable HPC software portfolio that is equally capable in DAC and M&S
  - Will improve productivity, increase reliability & trustworthiness, establish more sustainable yet agile software
- **Additional primary challenge: application design**
  - Componentization, extensibility, scalability, reusability, interoperability (build on libraries and motif-based toolkits!)
  - Motif-based hardware components?

* ***National Strategic Computing Initiative***: *https://www.whitehouse.gov/sites/whitehouse.gov/files/images/NSCI%20Strategic%20Plan.pdf*

Some M&S + DAC Workflows Evident

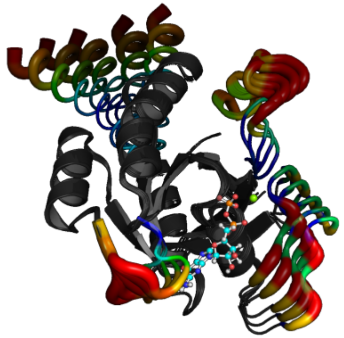
- **In situ (upstream, during, downstream) analytics [online data analysis and reduction]**
  - Cross section processing (*nuclear reactors*)
  - Multi-variate statistical analysis for characterizing turbulence & ignition dynamics; threshold-based topological feature detection (*combustion*)
  - Halo finder for identifying compact structures (*cosmology*)
  - Spatial and temporal feature detection in global ocean and atmospheric models (*climate*)
  - Formation and evolution of nonlinear turbulent plasma structures (*fusion*)
  - Multiphysics coupling within a nonlinear solution framework (*manufacturing, nuclear reactors, fusion, wind, climate, etc.*)
  - Statistically valid spatial and temporal data analyses of billion-atom atomistic simulations (*MD for nuclear materials*)
- **Computational steering**
  - Driving ensembles of MD simulations with unsupervised ML (*precision medicine for cancer*)
  - Robust and error-minimizing ML-driven mesh management (*shock physics of dynamic fluids/solids*)
- **Modeling unresolved physics (”subgrid” models)**
  - Super-parameterization approach for subgrid cloud-resolving model (*climate*)
  - ML-driven subgrid turbulence model parameter setting (*combustion, wind*)

Some M&S + DAC Workflows Evident

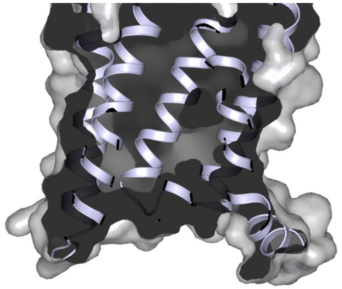
- **Predictive analytics**
  - Scalable supervised and unsupervised ML for tumor-drug response and treatment strategies (*precision medicine for cancer*)
- **Superfacility (HPC + Experimental Facilities)**
  - Integration of simulation and experimental science to improve analysis fidelity and turnaround time and increase knowledge detection/discovery from raw data. Computational steering of experiment via near real-time interpretation of molecular structure revealed by X-ray diffraction (*data analytics of FEL-generated light source facilities*)
  - ML-driven additive manufacturing of metal alloys (*manufacturing*)
- **Graph analytics**
  - Unknown genome reconstruction from collections of short, overlapping DNA segments (*microbiome analysis*)
  - Load balancing with spatial and temporal graph partitioning and clustering techniques (*power grid*)
  - Design of low-rank methods, dynamic load balancing strategies, with combinatorial algorithms, such as graph matching and coloring can help determine the work and data distribution (*chemistry*)

ECP: Heath Care Applications

- **Accelerate and Translate Cancer Research**
  - Build and apply a scalable deep neural network environment - the CANcer Distributed Learning Environment (CANDLE) - to address three top challenges of the National Cancer Institute: understanding the molecular basis of key protein interactions; developing predictive models for drug response, and automating the analysis; and extraction of information from millions of cancer patient records to determine optimal cancer treatment strategies

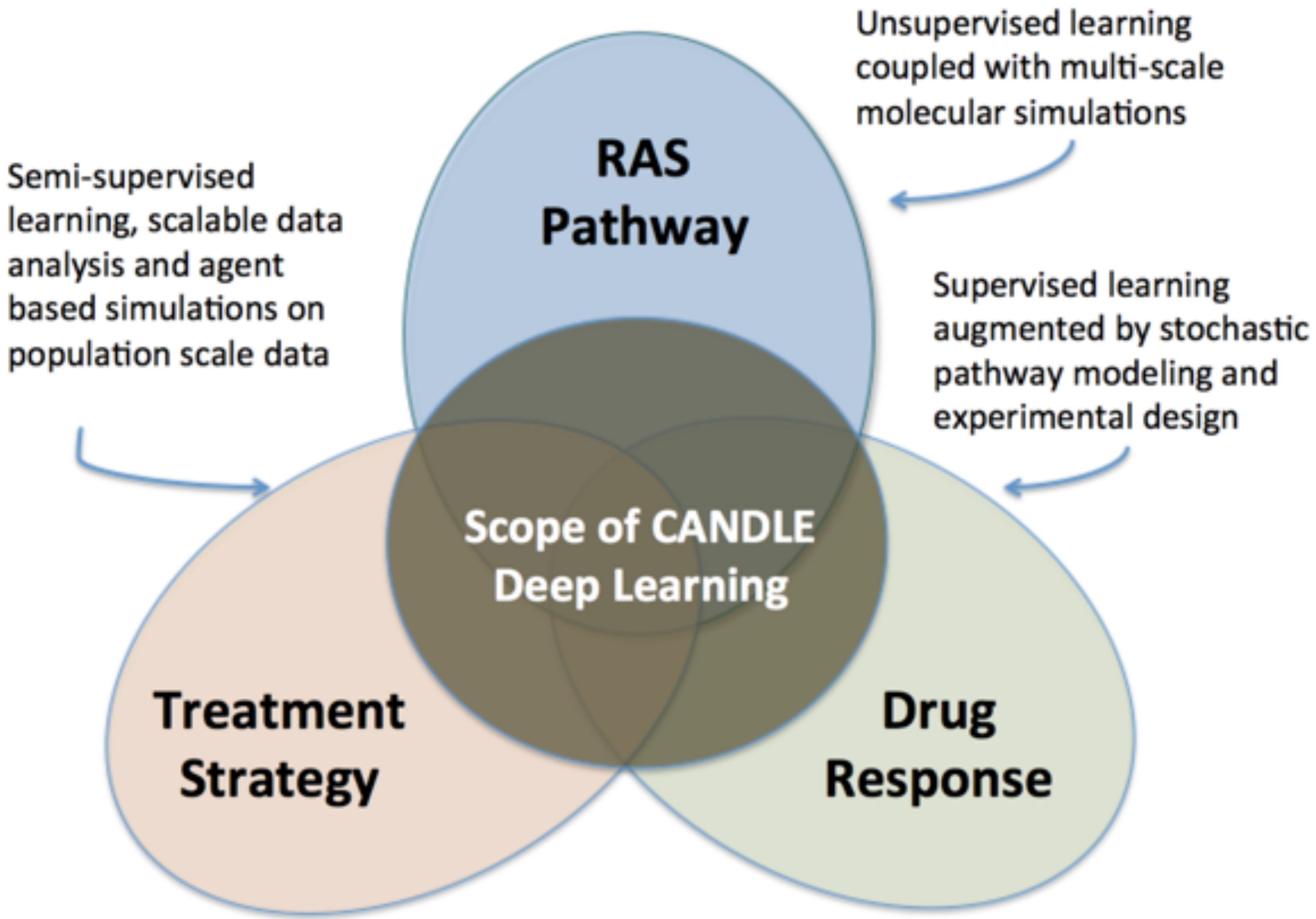


- **Light Source-Enabled Analysis of Protein and Molecular Structure and Design**
  - Linear Coherent Light Source is revealing biological structures in unprecedented atomic detail, helping to elucidate how proteins play key roles in biological function. Advanced solution scattering and coherent imaging simulation techniques are needed to characterize sub-nanometer scale conformational dynamics of heterogeneous ensembles of macromolecules



EXASCALE COMPUTING PROJECT

Pushing the frontiers of cancer research with Exascale
ECP-CANDLE

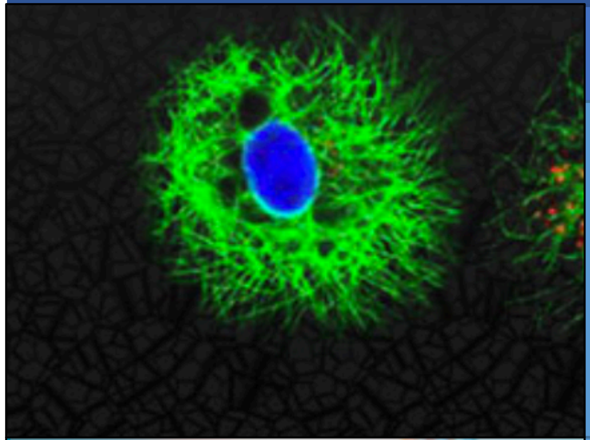
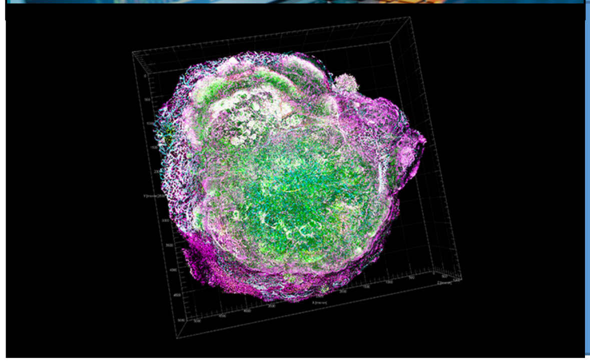


- Focus on building a scalable deep neural network code called the CANcer Distributed Learning Environment (CANDLE)
- CANDLE addresses three top challenges of the National Cancer Institute:
  1. Understanding the molecular basis of key protein interactions
  2. Developing predictive models for drug response, and automating the analysis
  3. Extraction of information from millions of cancer patient records to determine optimal cancer treatment strategies

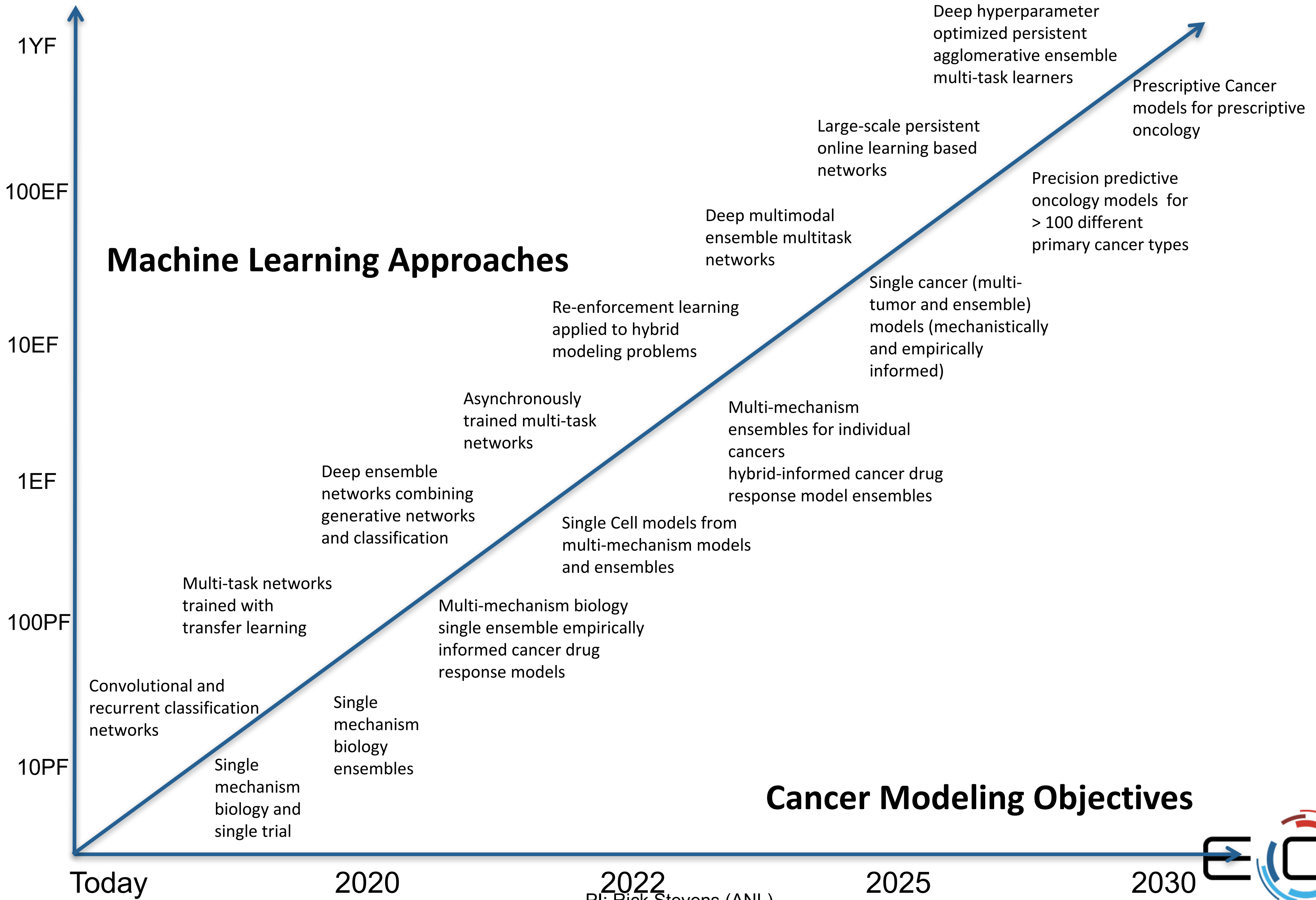
PI: Rick Stevens (ANL)

Pushing the Frontiers of Cancer Research with Exascale

Problems beginning in the 100 Pflop-1Eflop range, and scaling higher as larger ensembles are used for improving statistical outcomes

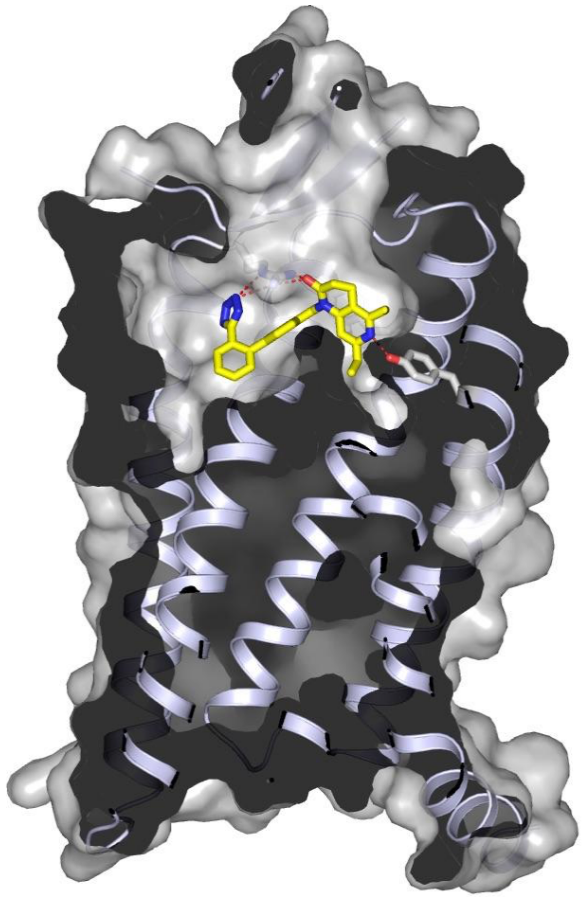
	Challenge Problem	Need for Exascale
	**High-accuracy modeling of cancer drug interactions and reaction dynamics in complex mutational backgrounds:** allow the study and modeling of critical aspects of a drug's mechanism of action across the full spectrum of possible mutational backgrounds	Requires exascale computing to evaluate reaction energies for extremely large systems in complex biological systems orders of magnitude greater than current capabilities
	***In silico* characterization of 100 billion readily synthesizable potential cancer drugs:** provide knowledge base of cost-effective, producible molecules characterized for potential clinical use and for future studies focusing on the system biology and dynamics of cancer	Requires exascale computing to support the generation and *in silico* characterization of more than 100 billion candidate molecules important to clinical and research applications in cancer
	**Computing three-dimensional structures of RNA:** critical first step to understanding their function and activity and ultimately to the development of therapeutic agents targeting RNA molecules	Requires exascale computing to enable data collected from instrument observations to be used as a basis for accurate determination of previously unknown 3D topological structures of RNA

Roadmap for Integration of Deep Learning & Simulation for Predictive Oncology



PI: Rick Stevens (ANL)

Data Analytics at the Exascale for Free Electron Lasers: Revealing Biological Function in Real Time

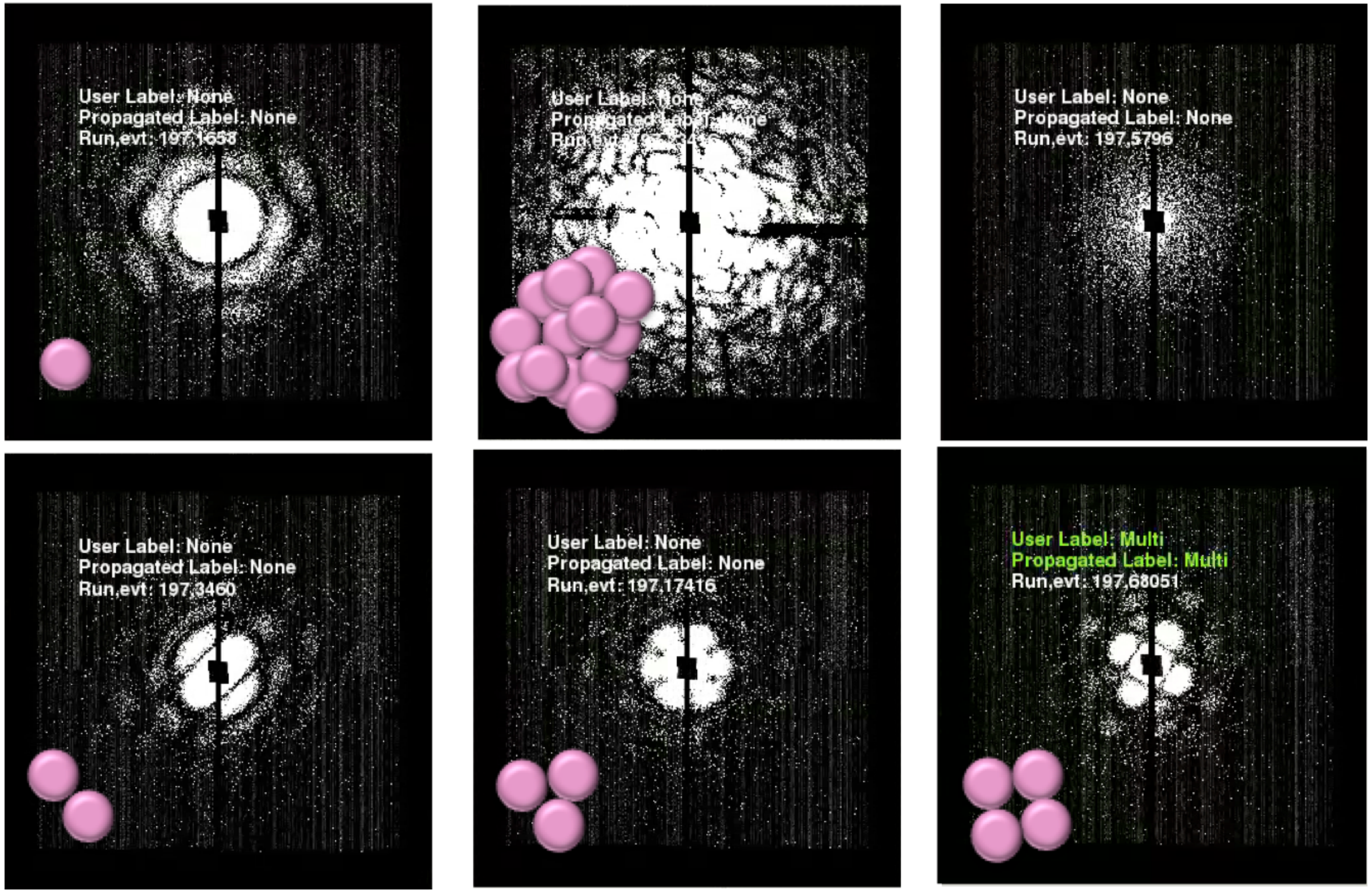


- Free Electron Lasers produce X-ray pulses more than a billion times brighter than the most powerful previously existing sources. The ultrafast X-ray pulses are used much like flashes from a high-speed strobe light, enabling scientists to take stop-action pictures of moving atoms and molecules, shedding light on the fundamental processes of physics and chemistry.
- Modern X-ray crystallography has transformed the field of structural biology by resolving macromolecules at the atomic scale.
- Biological function is profoundly influenced by dynamic changes in protein conformations and by interactions with molecules and other complexes – processes that span a broad range of timescales.
- Advanced solution scattering and coherent imaging techniques can characterize, at the sub-nanometer scale, the conformational dynamics of heterogeneous ensembles of macromolecules – both spontaneous fluctuations of isolated complexes, and conformational changes that may be initiated by the presence of specific molecules, environmental changes, or by other stimuli.

PI: Amedeo Perazzo (SLAC)

Single Particle Imaging (SPI)

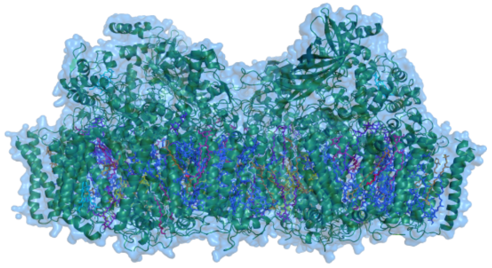
The high repetition rate and ultra-high brightness of the LCLS make it possible to determine the structure of individual molecules, mapping out their natural variation in conformation and flexibility

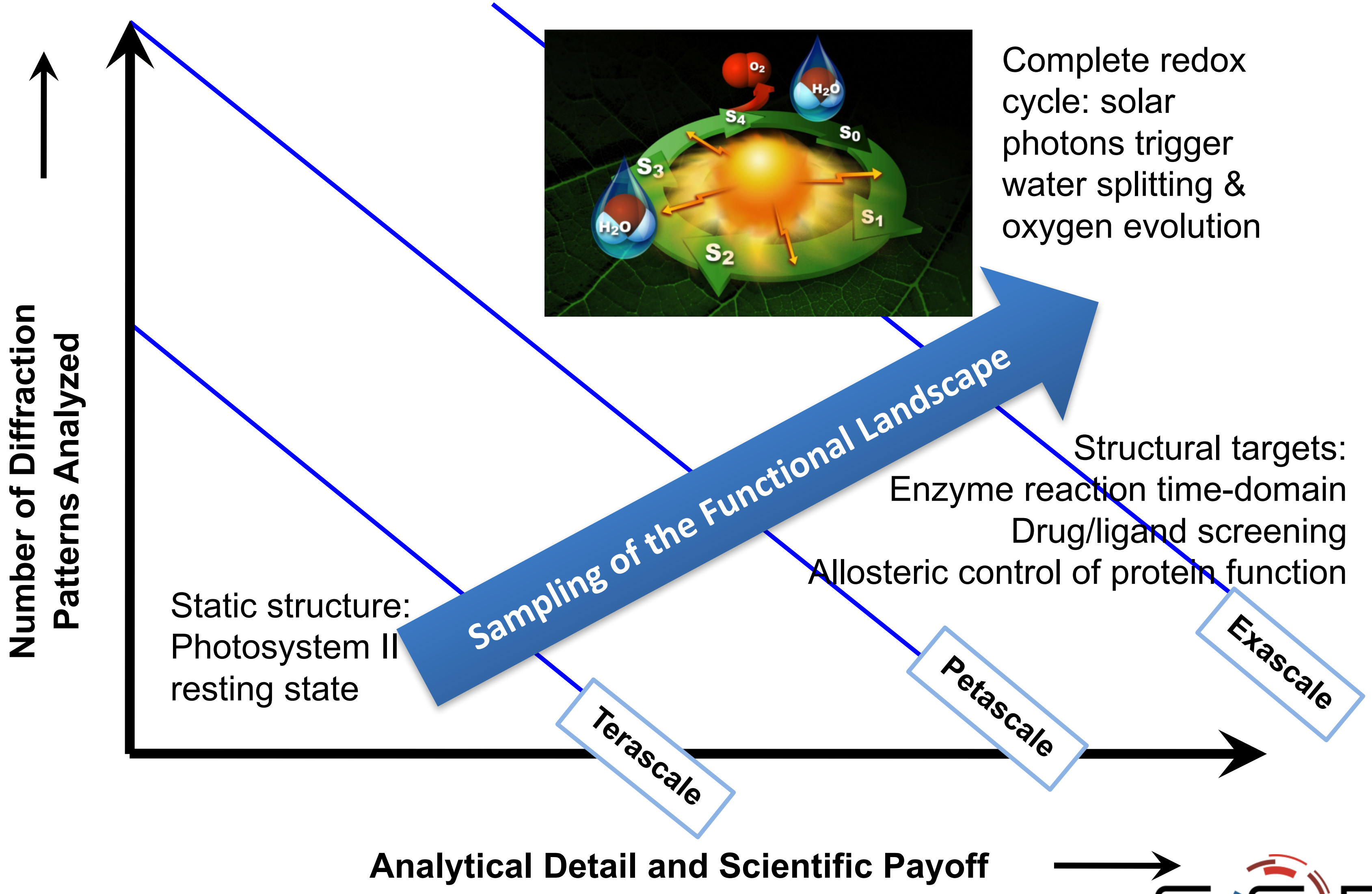


PI: Amedeo Perazzo (SLAC)

EXASCALE COMPUTING PROJECT

X-ray Free-Electron Laser Diffraction Data: From Biological Structure to Functional Understanding



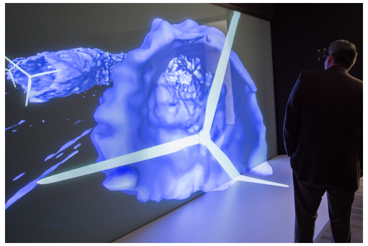


PI: Amedeo Perazzo (SLAC)

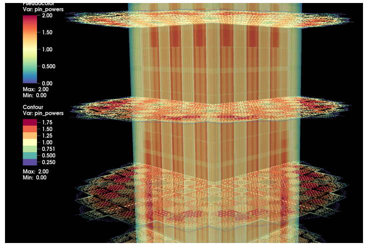
EXASCALE COMPUTING PROJECT

ECP: Energy Security Applications

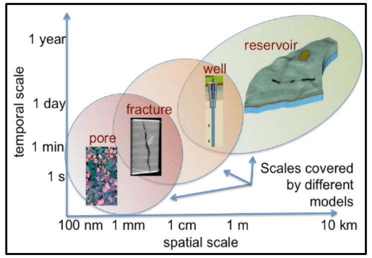
- **Turbine Wind Plant Efficiency**
  – Predict flow physics governing whole wind plant performance: wake formation, complex terrain impacts, turbine-turbine interaction effects. Wide-scale deployment of wind energy on the grid is hampered by significant plant-level energy losses by turbine-turbine interactions in complex terrains



- **Design and Commercialization of Small Modular Reactors (SMRs)**
  – Couple high-fidelity neutronics + fluid dynamics in an integrated toolkit for modeling the operational behavior of SMRs. Model operational behavior of existing Light Water Reactors at full power with full-core multiphase thermal hydraulics and fuel depletion (over the complete reactor lifetime)



- **Subsurface Carbon Capture, Fossil Fuel Extraction, Waste Disposal**
  – Safe and efficient use of the subsurface for geologic CO_2 sequestration, petroleum extraction, geothermal energy and nuclear waste isolation. Predict reservoir-scale behavior as affected by the long-term integrity of hundreds deep wells that penetrate the subsurface for resource utilization



EXASCALE COMPUTING PROJECT

Key to Reducing Wind Energy Cost is Plant-Level Optimization

Gaps and Opportunities

- ✓ Today's computational models cannot predict the complicated flow and structural dynamics of wind farms, which are poorly understood
- ✓ Exascale predictive wind farm simulations will expose new pathways to reduced cost of energy

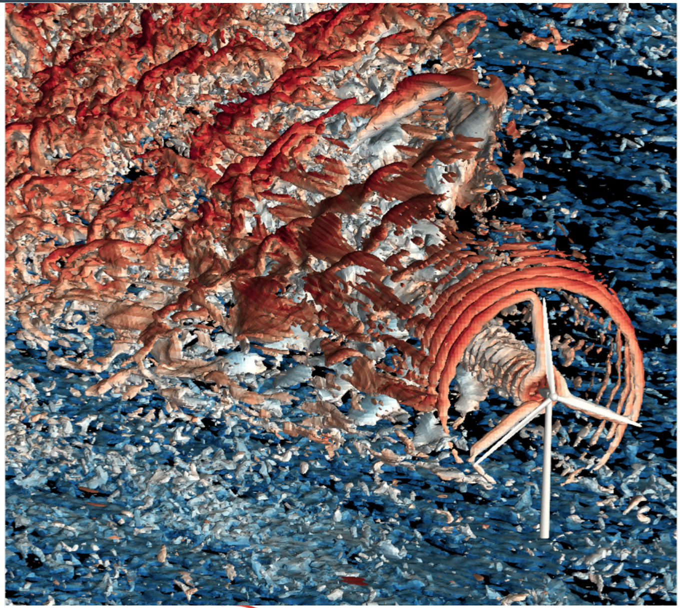
Simulation Challenge Problems

- ✓ Extreme separation of temporal and spatial scales in 10 km x 10 km farm, e.g., from the blade boundary layer to the atmospheric boundary layer
- ✓ Continuous changes in computational-mesh topology due to rotating and deforming blades, rotors, and nacelles
- ✓ Inadequate state of the art in turbulence models

Prospective Outcomes and Impact

- ✓ Minimize plant-level energy losses, which can be 20% or more, especially in complex terrain
- ✓ Reduce turbine failures; much higher in wind farms
- ✓ Quantify and reduce uncertainties in power production and turbine loads, which will reduce project financing costs
- ✓ Enable new technology innovations

Horns Rev wind farm in the eastern North Sea off of Denmark; wake dynamics highlighted by sea smoke



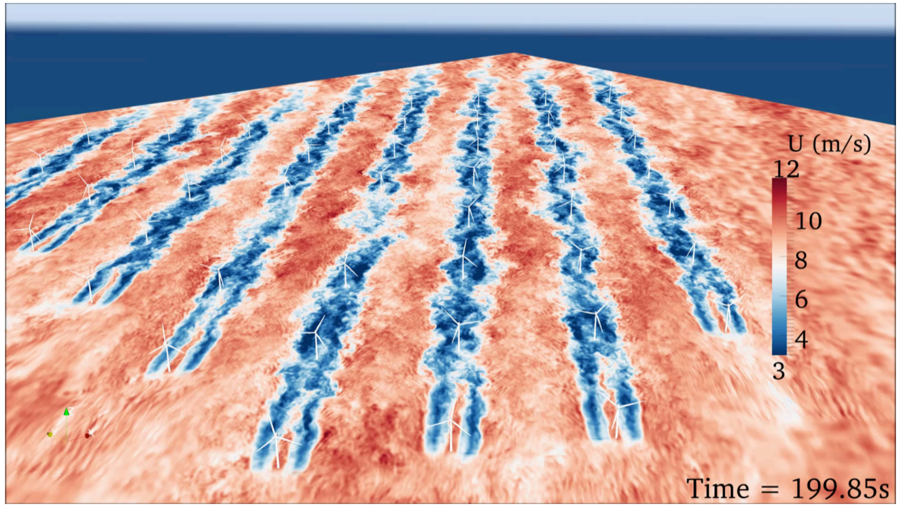
Coarse-grid single-turbine simulation; blades represented as actuator lines

EXASCALE COMPUTING PROJECT

PI: Mike Sprague (NREL)

Today's Wind Farm Simulations Are Not Predictive

- Today's simulations typically model turbines as simplified "line" or "disc" forces in a computational fluid dynamics model
  - Rely on empirical aerodynamic look-up tables
- Other limitations:
  - "Coarse" grid simulations; cannot reach desired extreme-scale resolution
  - Turbulence models need improvement
- Simulations have led to insights into controls system designs, and wake evolution under different atmospheric conditions
- State of the art is inadequate for
  - simulating disruptive technology innovations
  - capturing details of wake formation
  - capturing flow over complex terrain



Coarse-grid wind farm simulation

Wall-resolved LES study of a turbine-blade *section*; 1.6 billion cells

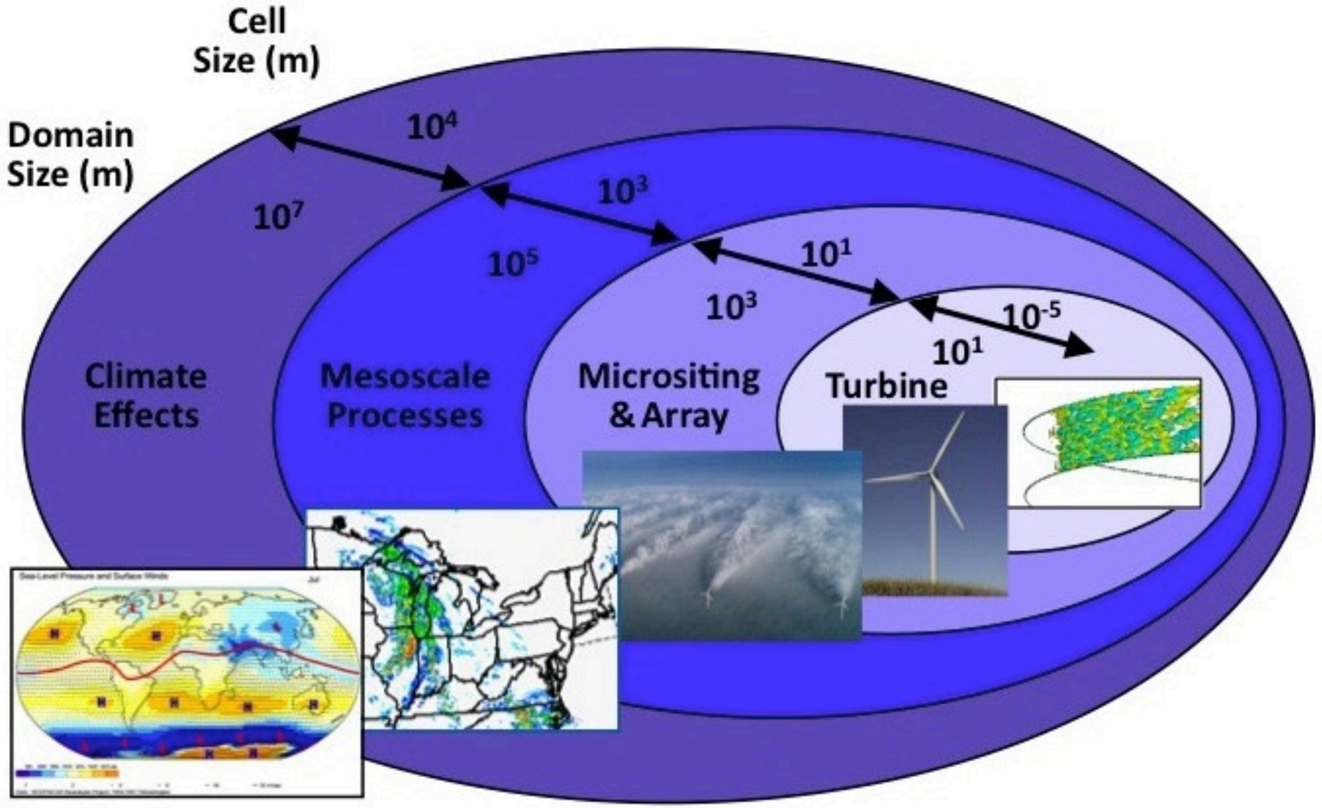
EXASCALE COMPUTING PROJECT

PI: Mike Sprague (NREL)

Exascale is Required for Predictive Wind Farm Simulations

- Spatial scales span from near-blade eddies (mm-scale) to weather-scale eddies (100-km scale)
- Simulation time steps will be about 0.001 sec, but total simulated time will be many hours
- A predictive simulation of a *single* turbine is a petascale-class problem
  - A single-turbine *predictive* simulation has yet to be demonstrated
  - *Single-turbine* simulation would require *O*(10) billions cells

Wind plants require extreme-scale simulations

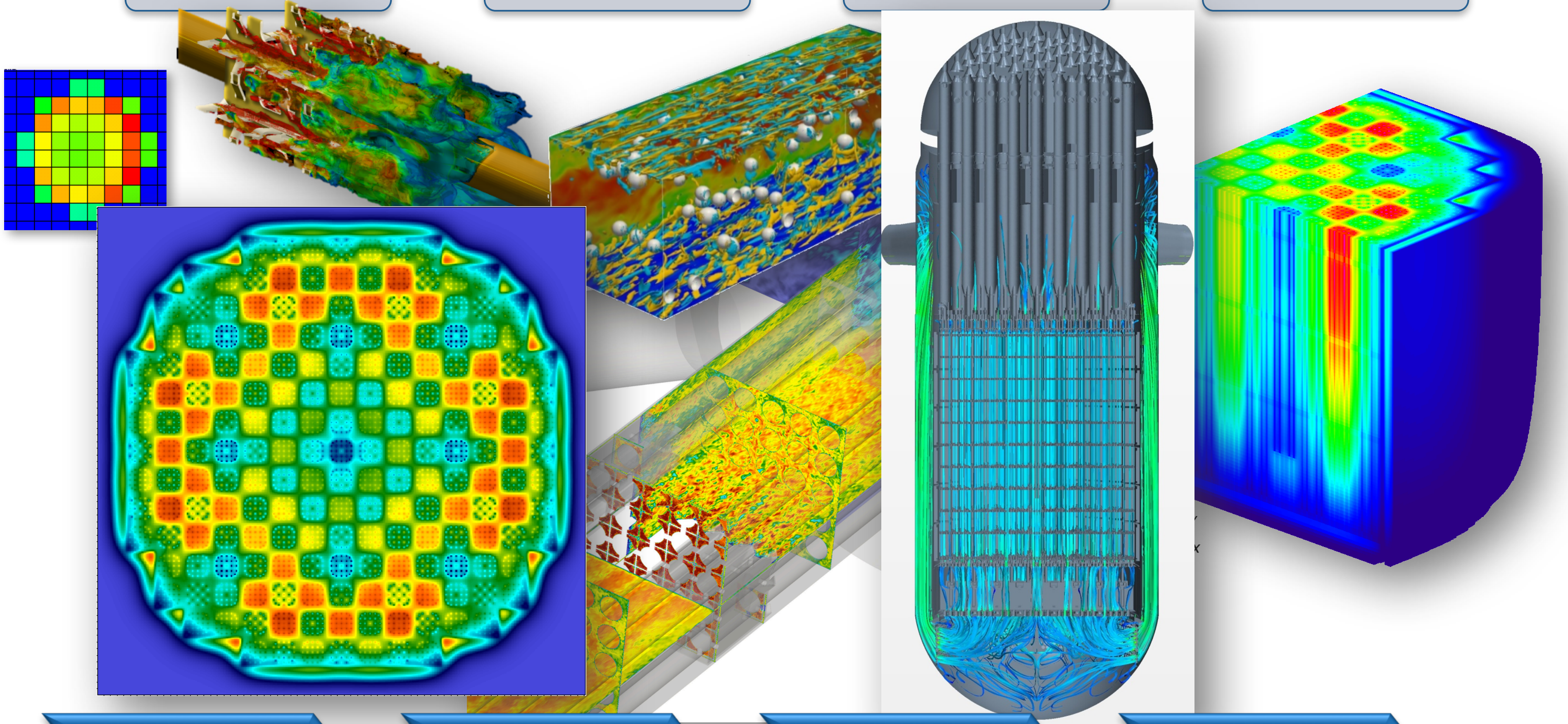


PI: Mike Sprague (NREL)

Exascale Nuclear Reactor Simulation

Accurate Representation of 3D Geometries with Multiple Resolved Scales and Physics

Terascale → Petascale → Exascale → Beyond



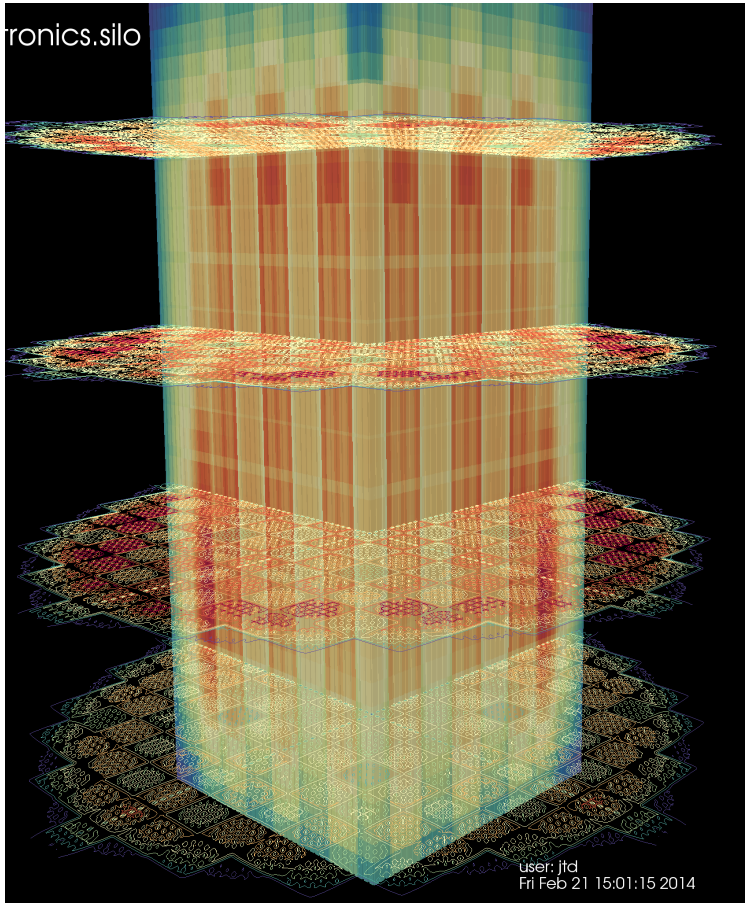
Terascale	Petascale	Exascale	Beyond
2D, homogenization, 1-way coupling	3D, limited 2-way coupling	3D, fully-coupled, micro-informed	3D, multi-scale, UQ/optimization

PI: Thomas Evans (ORNL)

EXASCALE COMPUTING PROJECT

High resolution Monte Carlo neutronics at Petascale

- Monte Carlo methods remove the largest source of error in low-order simulations – *energy discretization*
  - 2012: Shift (MC) neutronics calculations were used to validate **startup** conditions of the new WEC AP1000® reactor
  - 2015 – 2016: VERA – Shift is being used to model WB2 **startup**
- Full core MC calculations require most of Titan
  - 18 688 compute nodes
  - 10^{12} particle histories
  - 3 – 4 hour runtime
- Produce pin-resolved data sets that are used to validate engineering-scale design calculations

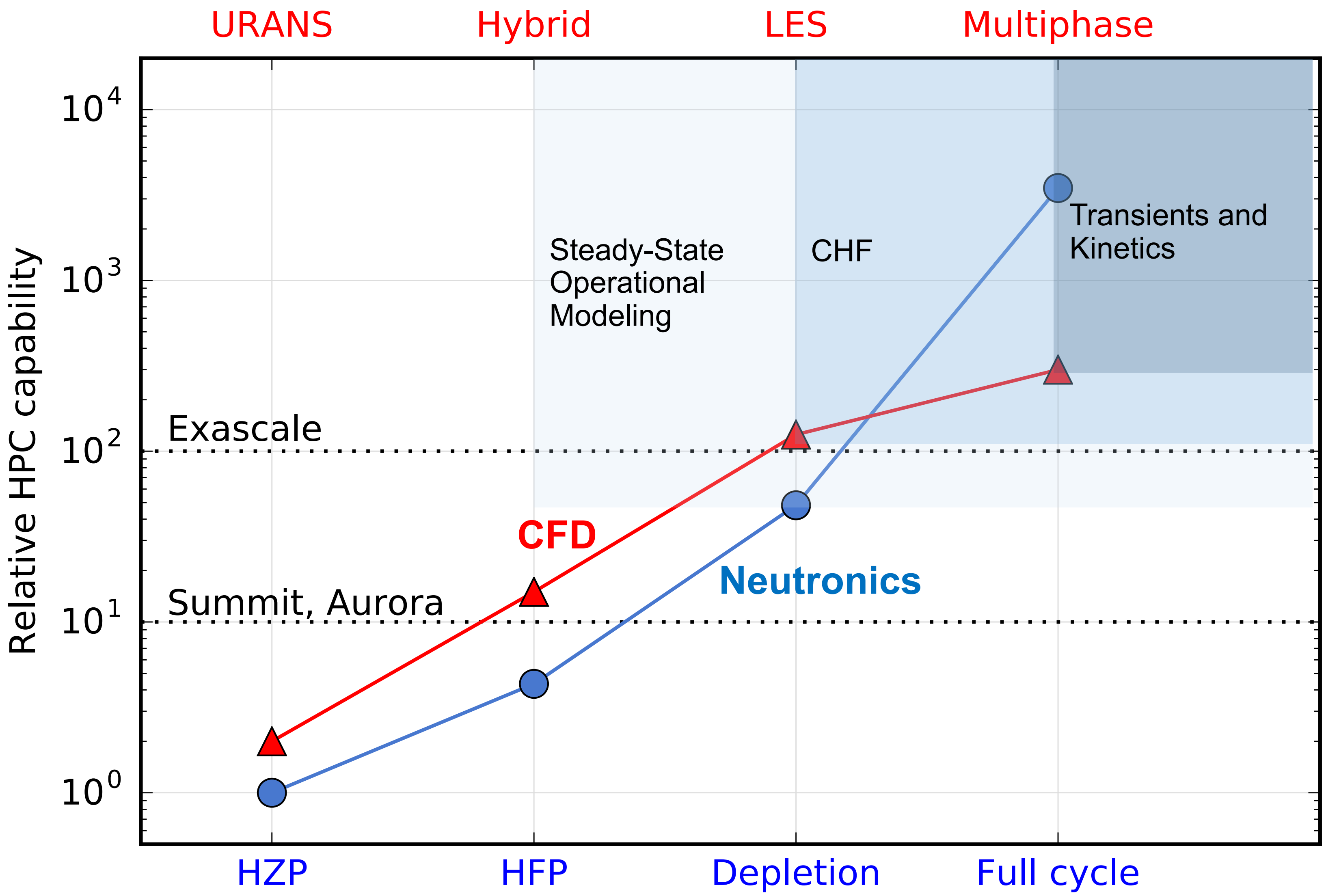


Petascale enables high-resolution for Hot-Zero-Power (HZP) startup simulations

EXASCALE COMPUTING PROJECT

PI: Thomas Evans (ORNL)

Required compute capability for Challenge Problems

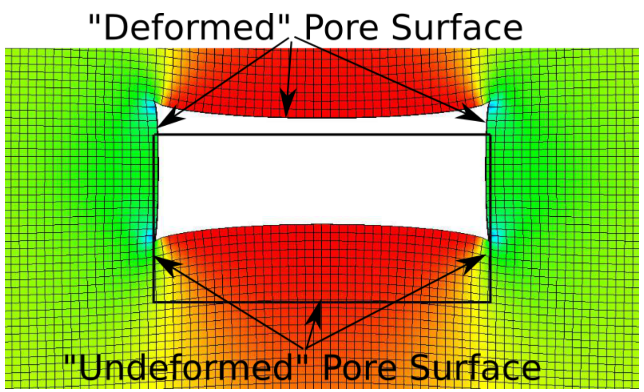


PI: Thomas Evans (ORNL)

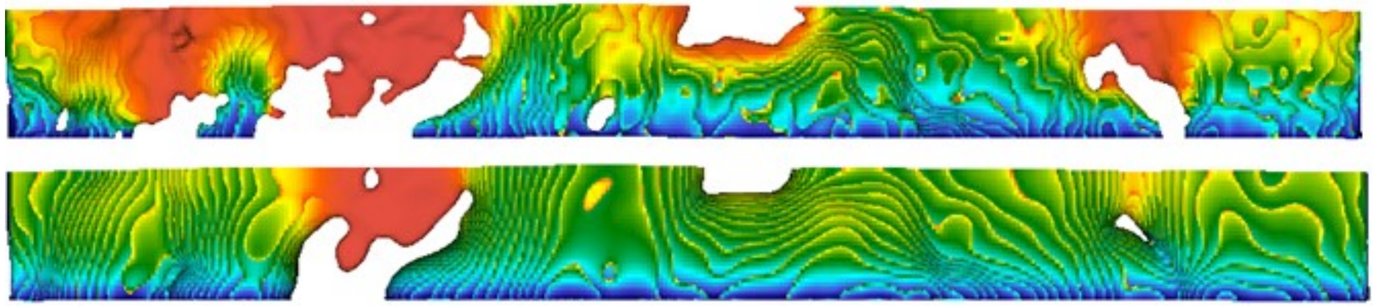
Exascale Can Enable Complex Coupled Physics in Subsurface

Simulation Gap	Needed to Address the Gap	Impact of Gap Remaining
Predictive models for fracture opening/closing and wellbore integrity in the subsurface that account for the coupled physics associated with pore scale flow, multicomponent reactions, and geomechanical processes.	Development of a computational framework based on finite elements for geomechanical deformation coupled at every time step to an embedded boundary/finite volume approach to tracking surfaces for reactive flow.	Inability to consider coupled physical, chemical, and mechanical processes governing fracture and wellbore behavior..

Finite Element-Based Deformation



Embedded Boundary/Finite Volume Simulation of Fracture Evolution due to Reactive Flow

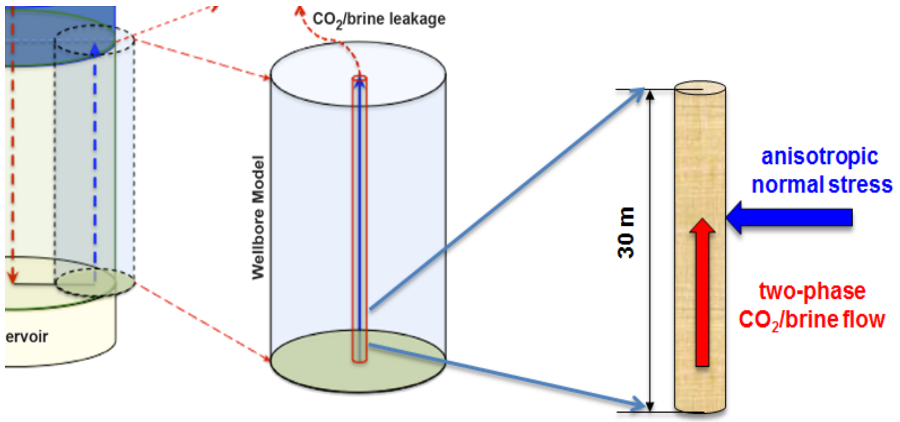


PI: Carl Steefel (LBNL)

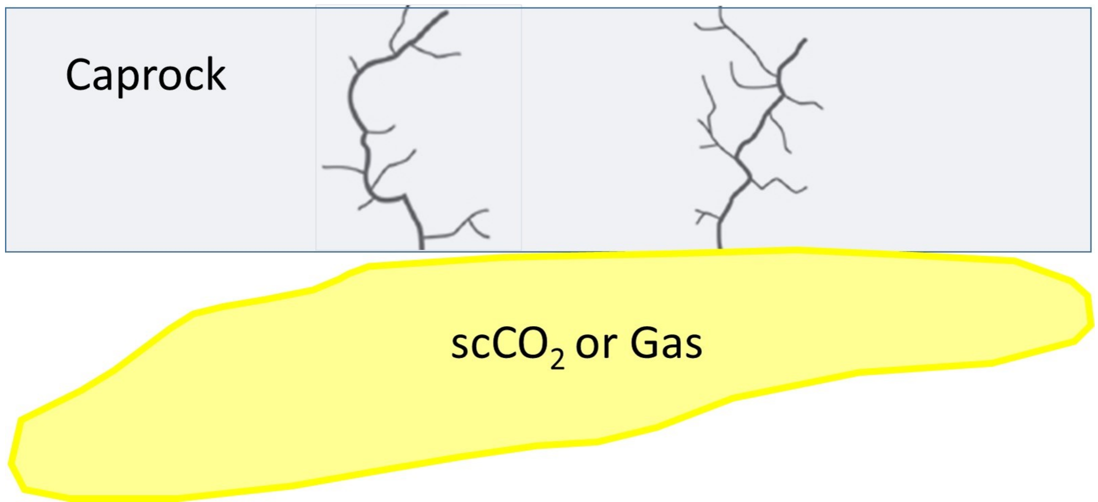
Exascale Can Reconcile Pore and Continuum Scale Multiphase Flow

Simulation Gap	Needed to Address the Gap	Impact of Gap Remaining
Predictive models for how gas, oil, and supercritical CO_2 distributions in the subsurface are impacted by fine-scale physical, mechanical, and geochemical heterogeneity.	Usd DNS to simulate continuum (Darcy) and pore/fracture scale multiphase flow, with seamless integration between distinct mathematical/conceptual model formulations.	Limits to incorporation of geological heterogeneity into simulations of the distribution and migration of multiphase fluids, including gas, oil, and CO_2.

Multiphase Wellbore Problem



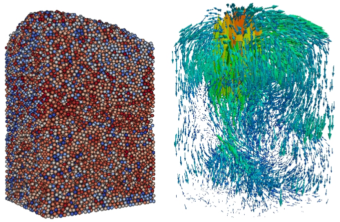
CO_2 Leakage through Caprock



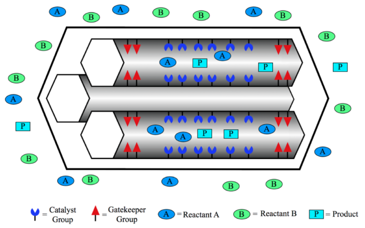
PI: Carl Steefel (LBNL)

ECP: Energy Security Applications

- **Scaleup of Clean Fossil Fuel Combustion**
  - High-fidelity modeling capability to guide scale-up of laboratory designs of multi-phase chemical looping reactors to industrial size; required to impact design decisions and drive large-scale commercial deployment of carbon capture and storage (CCS) and meet DOE CCS goals



- **Biofuel Catalyst Design**
  - Understand and exploit the reaction mechanism and dynamics of heterogeneous catalysis on mesoporous silica nanoparticles - a highly effective and selective heterogeneous catalyst for a wide variety of important reactions that impact development of green energy sources requiring specific conversion of cellulosic based chemicals into fuels or other industrially important products

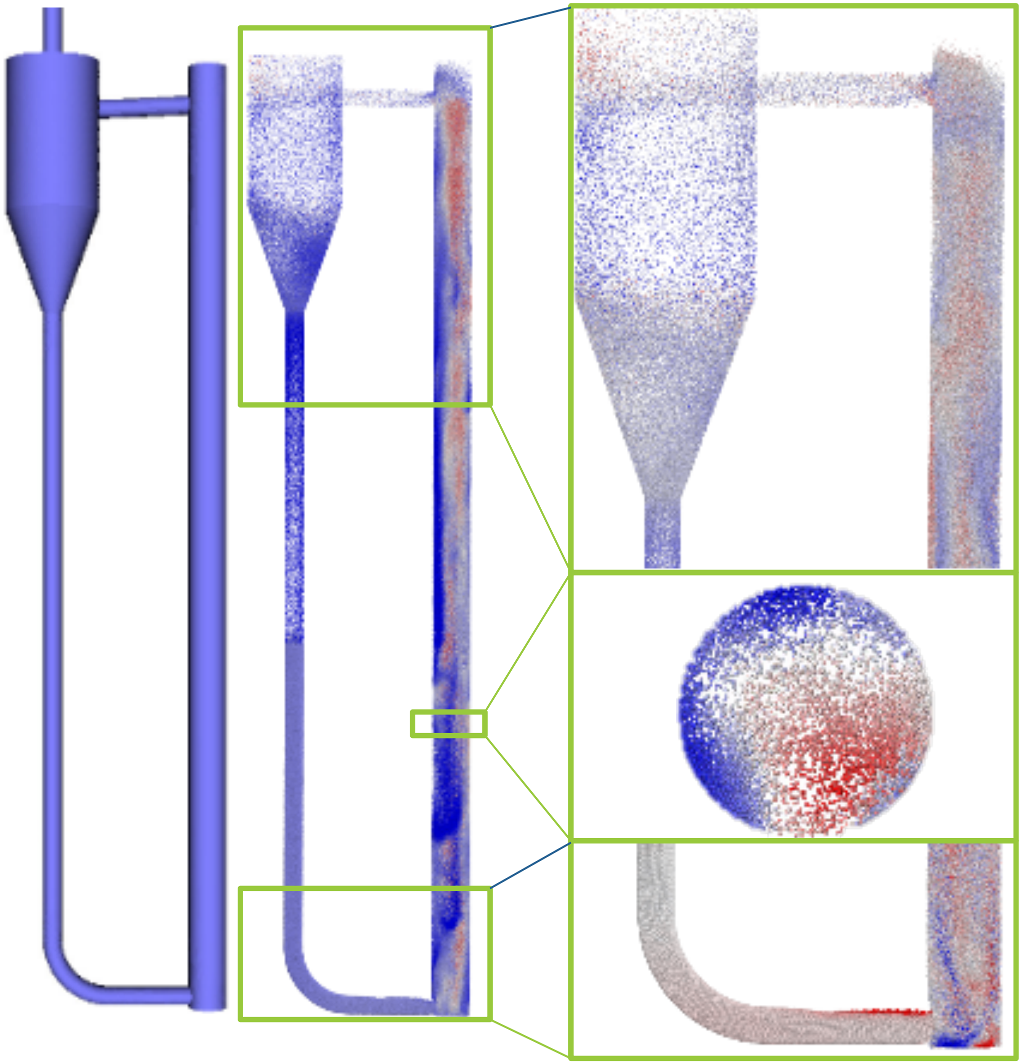


EXASCALE COMPUTING PROJECT

MFIX-Exa: Accelerate DOE-FE Technology Development

Using exascale computations

- Scaling-up of new clean energy technologies from lab- to commercial-scale requires expensive and time-consuming testing at multiple intermediate scales
- The goal of DOE-FE's transformational energy technologies is to reach commercial-scale demo by 2025-2030; that is, very soon
  - Some of the technologies involve gas-solids reactors such as chemical looping reactors
  - High-fidelity gas-solids reactor models that can reduce or replace physical testing will help accelerate the technology development
- MFIX-Exa will develop software for creating gas-solids reactor models that run efficiently on exascale computers



NATIONAL ENERGY TECHNOLOGY LABORATORY BERKELEY LAB MFIX Exa

PI: Madhava Syamlal (NETL)

EXASCALE COMPUTING PROJECT

MFIX-Exa will replace TFM with higher-fidelity CFD-DEM

For modeling pilot-scale reactors

2016
Two-Fluid-Model of Pilot-scale Reactor
Particles treated as a granular fluid

Step change in simulation fidelity:

- Does not require (uncertain) constitutive relations for granular fluid
- Describes slowly shearing or stagnant regions
- Describes the distribution in particle-scale properties: size, density, chemical conversion etc.

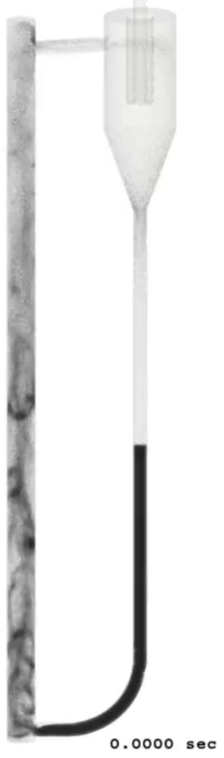
2026
CFD-DEM of Pilot-scale reactors
Particles are individually tracked

EXASCALE COMPUTING PROJECT

PI: Madhava Syamlal (NETL)

New capability demonstrated by solving a challenge problem
Simulation of a 1 MWe chemical looping reactor

2016

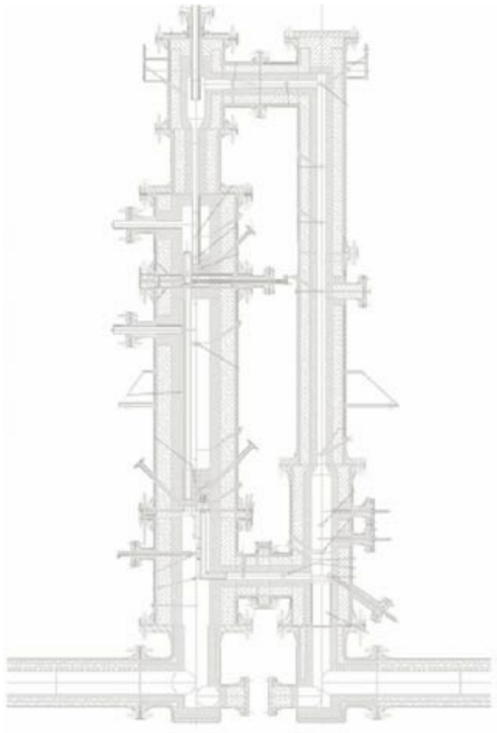


Power: 0 kW
Particle Count: 10^6
Time to Solution: 20 days

2020

Power: 50 kW
Particle Count: $5x10^9$
Time to Solution: 20 days

2026



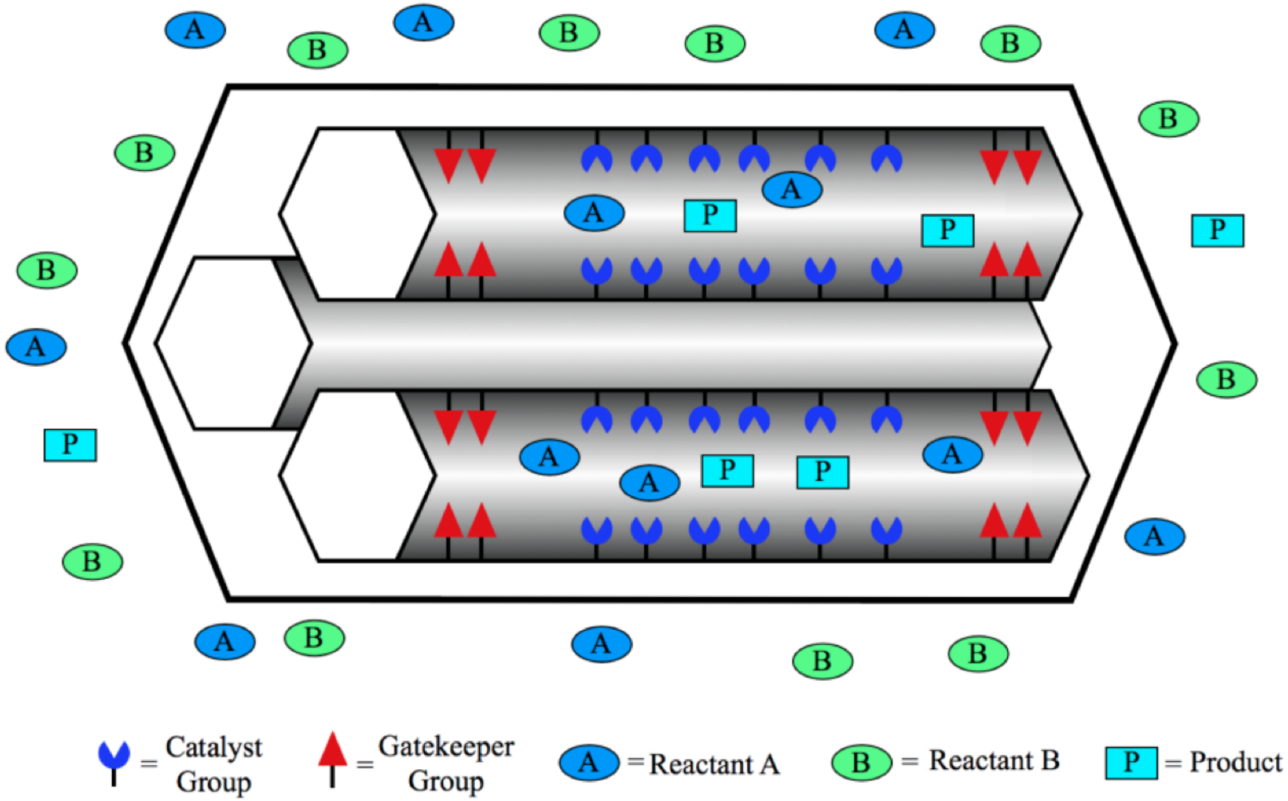
Power: 1 MWe
Particle Count: 10^{11}
Time to Solution: 1day

PI: Madhava Syamlal (NETL)

EXASCALE COMPUTING PROJECT

Enabling GAMESS for Exascale Computing in Chemistry & Materials

Heterogeneous Catalysis on Mesoporous Silica Nanoparticles (MSN)

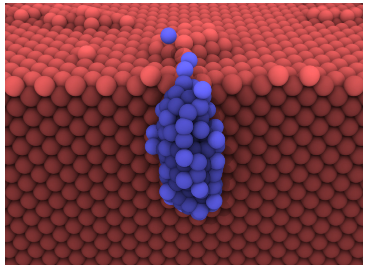


- MSN: highly effective and selective heterogeneous catalysts for a wide variety of important reactions
- MSN selectivity is provided by "gatekeeper" groups (red arrows) that allow only desired reactants A to enter the pore, keeping undesirable species B from entering the pore
- Presence of solvent adds complexity: Accurate electronic structure calculations are needed to deduce the reaction mechanisms, and to design even more effective catalysts
- Narrow pores (3-5 nm) create a diffusion problem that can prevent product molecules from exiting the pore, hence the reaction dynamics must be studied on a sufficiently realistic cross section of the pore
- Adequate representation of the MSN pore requires ~10-100K atoms with a reasonable basis set; reliably modeling an entire system involves >1M basis functions
- Understanding the reaction mechanism *and* dynamics of the system(s) is beyond the scope of current hardware and software – requiring capable exascale

PI: Mark Gordon (Ames)

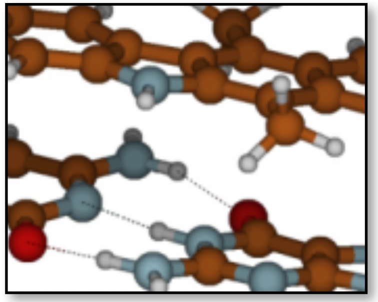
ECP: Transformative Materials Applications

- **Materials for Extreme Environments**
  - Address key fusion and fission energy materials challenges at the atomistic level: extending the burnup of nuclear fuel in fission reactors (dynamics of defects & fission gas clusters in UO_2); and developing plasma facing components (tungsten first wall) to resist the harsh conditions of fusion reactors



- **Additive Manufacturing of Qualifiable Metal Parts**
  - A validated additive manufacturing (AM) simulator enables determination of optimal process parameters for desired material properties and real-time in situ process optimization. Coupled to a modern design optimization tool, allows routine use of AM to build novel and qualifiable metal alloy parts

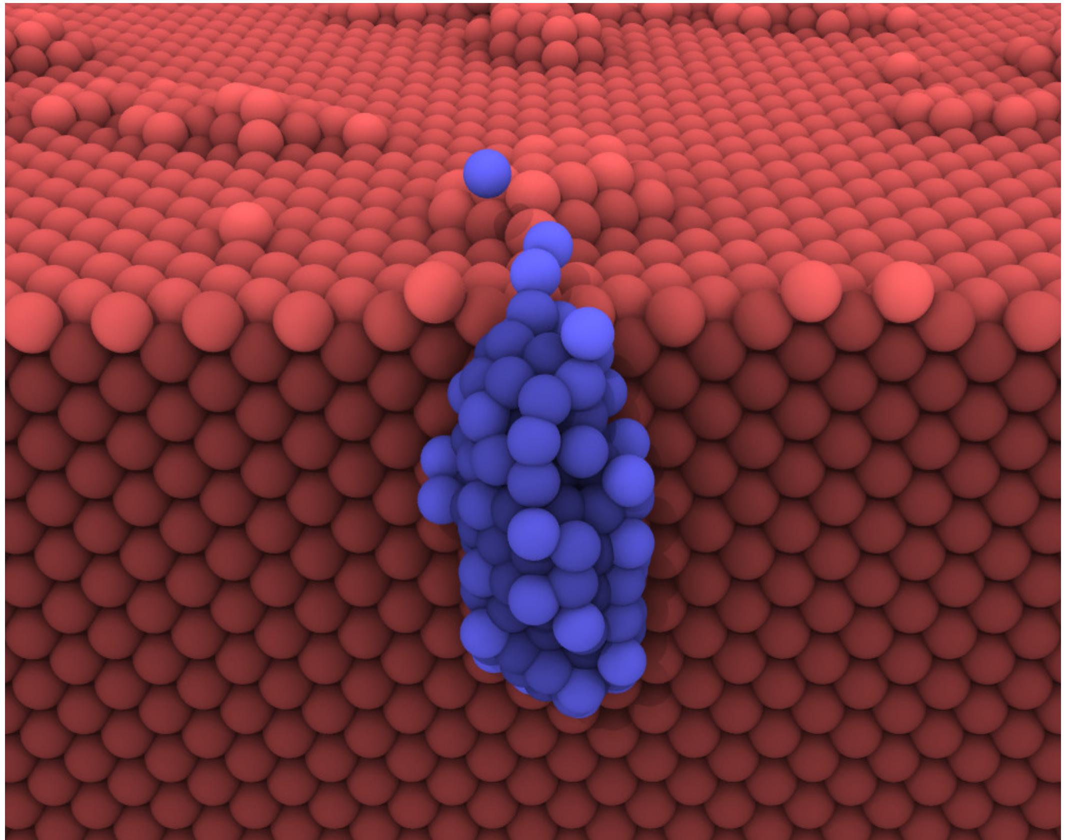
- **Find, Predict, Control Materials and Properties at the Quantum Level**
  - Find, predict and control materials and properties at the quantum level with an unprecedented and systematically improvable accuracy. Simulate transition metal oxide systems, e.g., complex oxide heterostructures that host novel quantum phases, to 10meV statistical accuracy. Major potential materials science impact, e.g., uncovering the mechanisms behind high-temperature superconductivity



EXASCALE COMPUTING PROJECT

Molecular Dynamics at the Exascale: Spanning the Accuracy, Length and Time Scales for Critical Problems in Materials Science (EXAALT)

Combining time-acceleration techniques, spatial decomposition strategies, and high accuracy quantum mechanical and empirical potentials

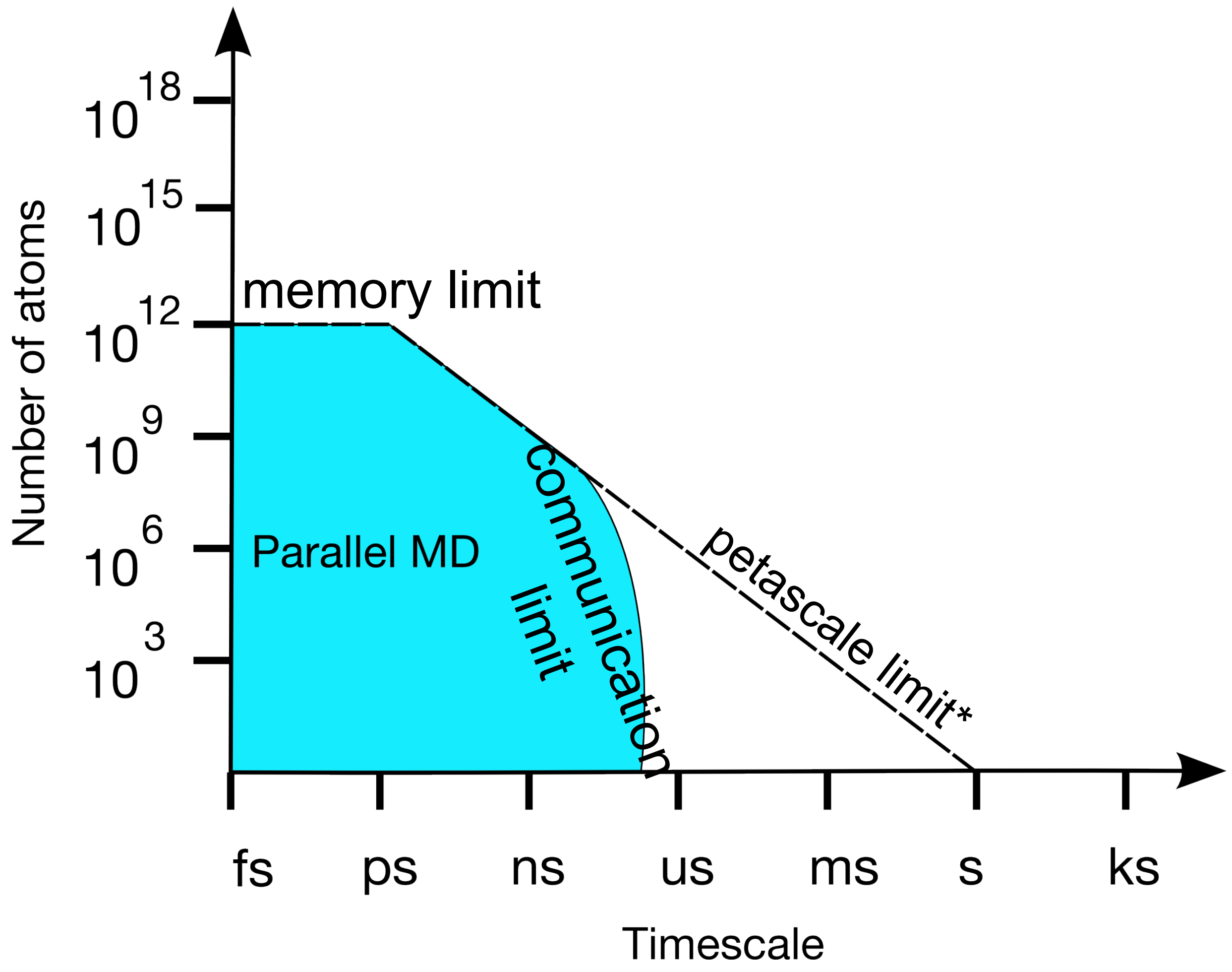


Slowly-growing He bubble in W at bursting

- Tackle materials challenges for energy, especially fission and fusion, by allowing the scientist to target, at the atomistic level, the desired region in accuracy, length, and time space
- Shown here is a simulation aimed at understanding tungsten as a fusion first-wall material, where plasma-implanted helium leads to He bubbles that grow and burst at the surface, ultimately leading to surface "fuzz" by a mechanism not yet understood
- At slower, more realistic growth rates (100 He/μsec), the bubble shows a different behavior, with less surface damage, than the fast-grown bubble simulated with direct molecular dynamics (MD)
- Atomistic simulation allows for complete microscopic understanding of the mechanisms underlying the behavior
- At the slower growth rate, crowdion interstitials emitted from the bubble have time to diffuse over the surface of the bubble, so that they are more likely to release from the surface-facing side of the bubble, giving surface-directed growth.

PI: Arthur Voter (LANL)

Accessible MD time and length scales

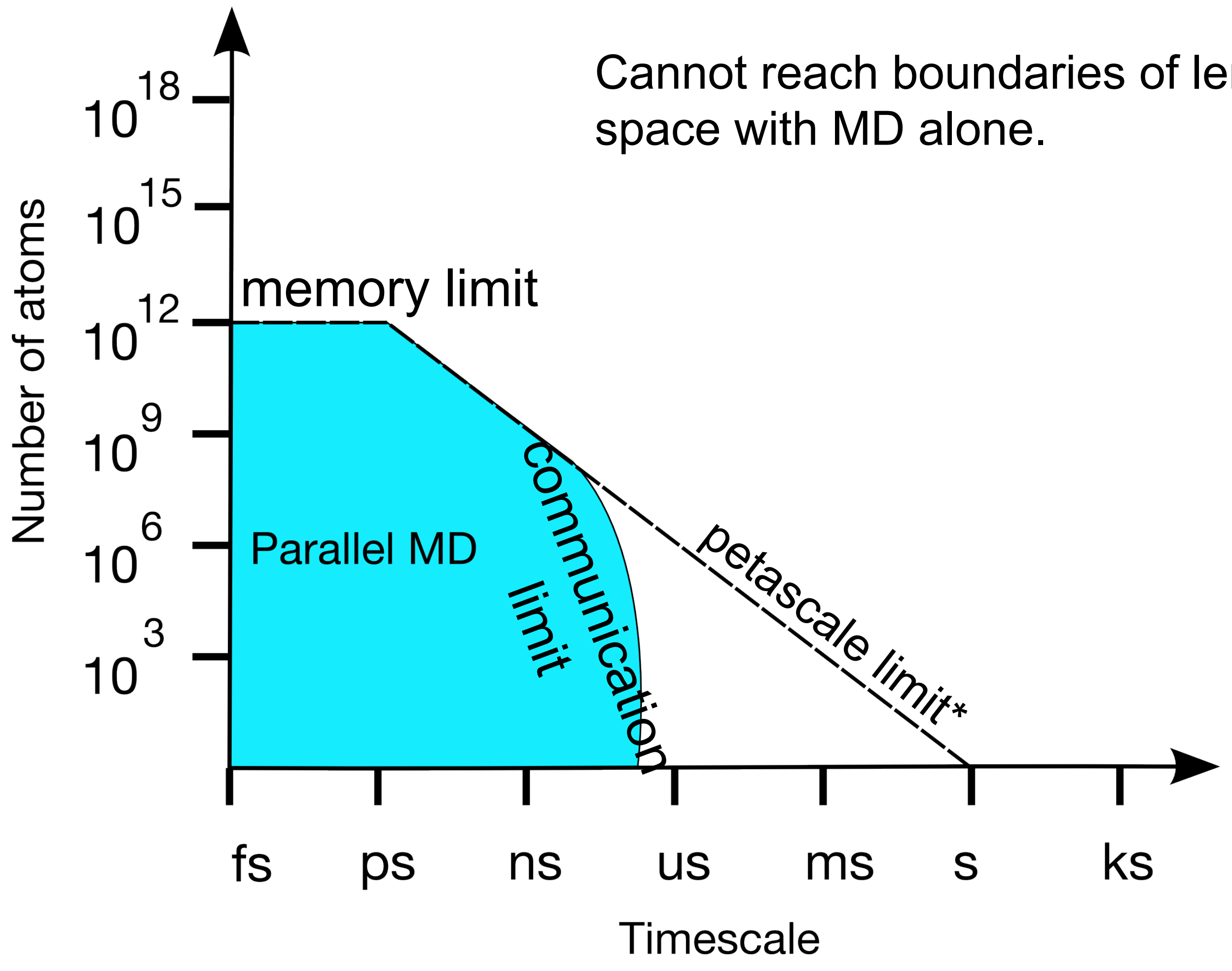


*1 week on Roadrunner – 1 PFLOPS machine, fast EAM potential

PI: Arthur Voter (LANL)

EXASCALE COMPUTING PROJECT

Accessible MD time and length scales



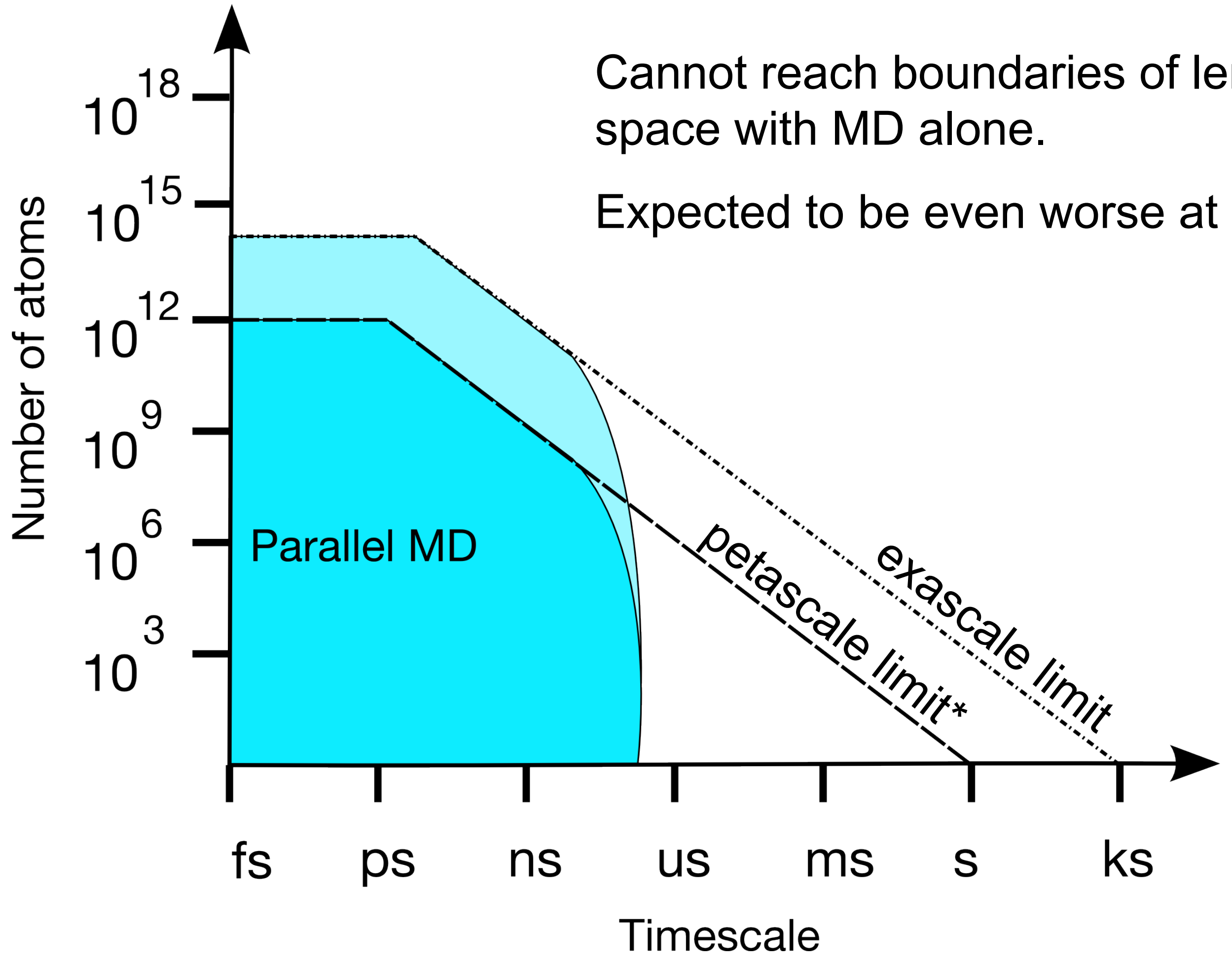
Cannot reach boundaries of length-time space with MD alone.

*1 week on Roadrunner – 1 PFLOPS machine, fast EAM potential

PI: Arthur Voter (LANL)

EXASCALE COMPUTING PROJECT

Accessible MD time and length scales



Cannot reach boundaries of length-time space with MD alone.

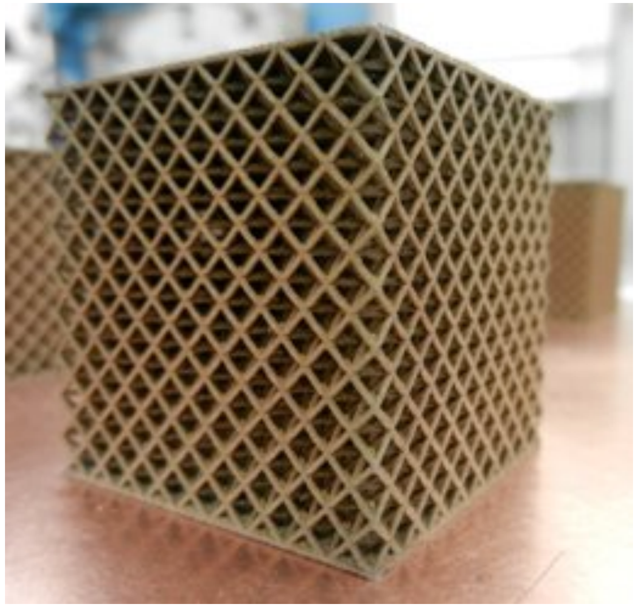
Expected to be even worse at exascale…

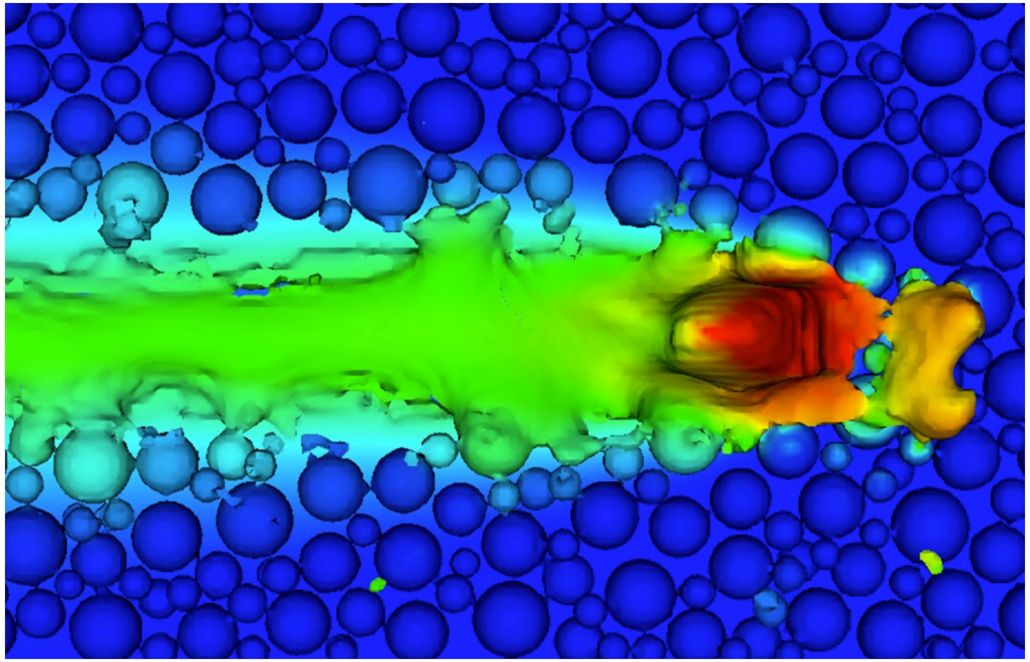
*1 week on Roadrunner – 1 PFLOPS machine, fast EAM potential

PI: Arthur Voter (LANL)

EXASCALE COMPUTING PROJECT

Transforming Additive Manufacturing through Exascale Simulation (ExaAM)



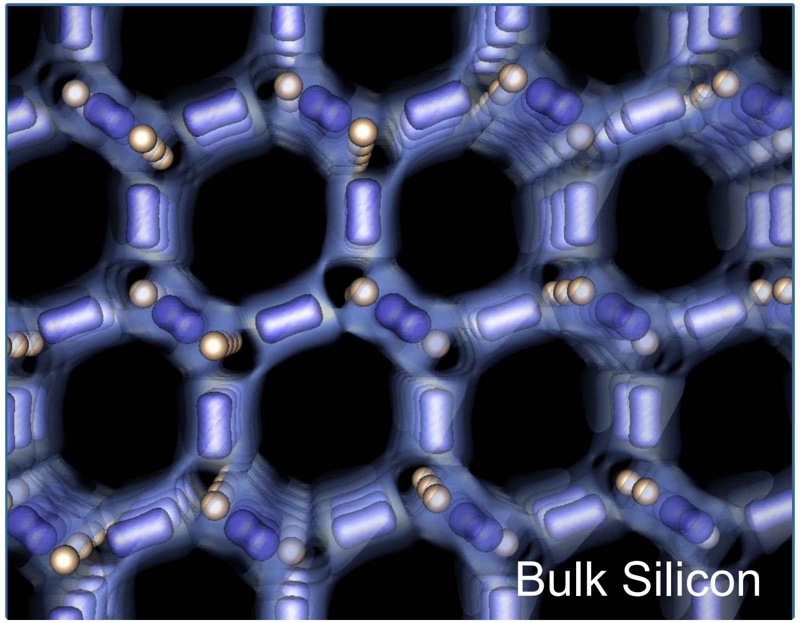


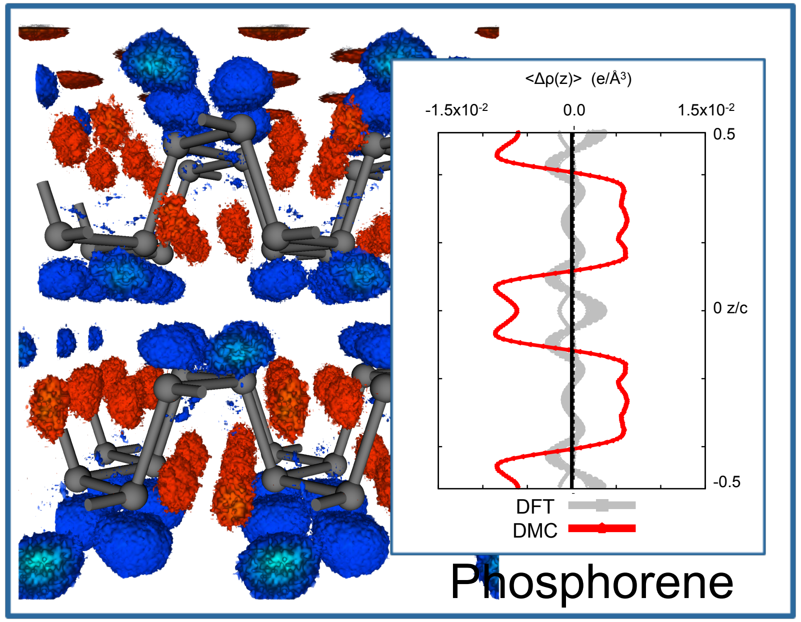
- The Exascale AM (ExaAM) project is building a new multi-physics modeling and simulation toolkit for Additive Manufacturing (AM) to provide an up-front assessment of the manufacturability and performance of additively manufactured parts
- An Integrated Platform for AM Simulation (IPAMS) will be experimentally validated, enabled by in-memory coupling between continuum and mesoscale models to quantify microstructure development and evolution during the AM process
- Microscopic structure determines local material properties such as residual stress and leads to part distortion and failure
- A validated AM simulator enables determination of optimal process parameters for desired material properties, ultimately leading to reduced-order models that can be used for real-time in situ process optimization
- Coupled to a modern design optimization tool, IPAMS will enable the routine use of AM to build novel and qualifiable parts

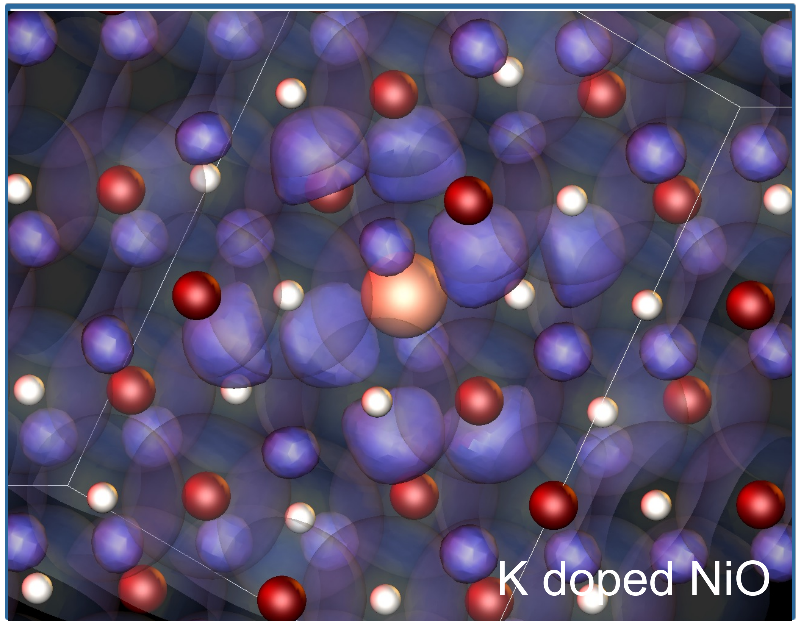
PI: John Turner (ORNL)

ECP EXASCALE COMPUTING PROJECT

Exascale will enable treatment of realistic, complex materials





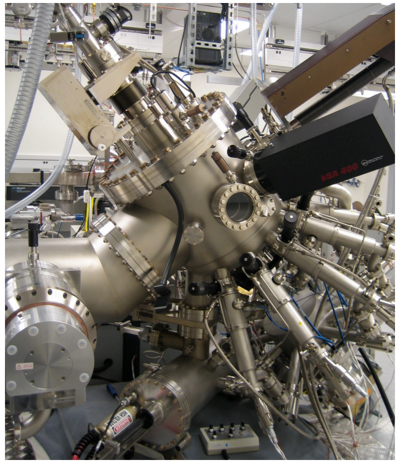


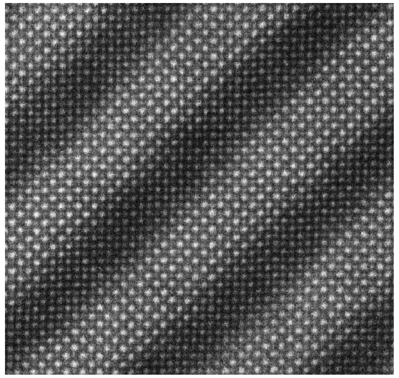
Terascale	Petascale	Exascale
• Simple materials properties • Light elements, e.g. silicon, carbon • Simple molecular systems	• High precision calculations of simple nanostructures, e.g. graphene, phosphorene. • Bulk strongly correlated materials, e.g. binary transition metal oxides • Systematic error reduction in simple molecular systems	• Realistic materials with defects & dopants from across the periodic table • Spin-orbit coupling for heavy elements • Systematic error reduction in simple materials • Sufficient throughput for upscaling

PI: Paul Kent (ORNL)

QMCPACK: Enabling Predictive Simulation of Materials

- To understand & design next generation materials we need reliable, non-empirical, fully atomistic methods
  - e.g. Novel materials for high performance electronics, sensing, and storage that rely on metal to insulator transitions can not be computationally designed today because reliable predictive methods are not yet affordable.
- Quantum Monte Carlo techniques are on track to achieve these goals, but require exascale hardware, software, and algorithms.
- Applications include materials for new electronics and computing technologies, but since all errors will be controlled, the methods can be applied across the materials, physics, and chemistry domains.





Today, complex materials can be grown with near atomic precision, but we can not simulate them sufficiently reliably and therefore design or optimize them.

EXASCALE COMPUTING PROJECT

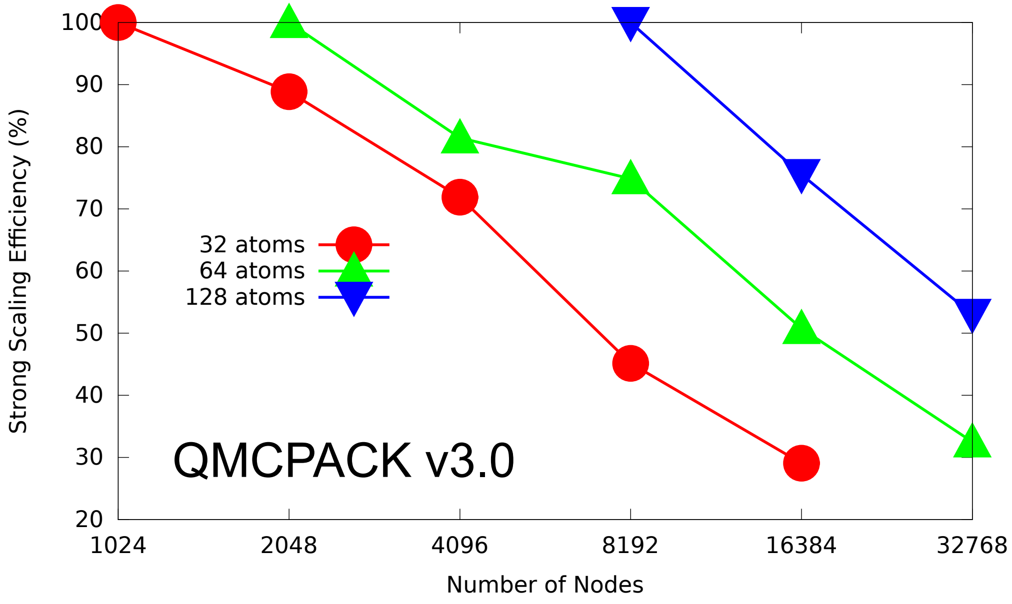
PI: Paul Kent (ORNL)

Quantum Monte Carlo methods

- Calculations on 1000+ atom systems are required to simulate sufficient materials complexity and reduce periodic image effects. The calculations can not run today due to memory bottlenecks.
- These calculations also require significant new on node concurrency to achieve required “time to scientific solution”. Our studies indicate this is realistically achievable.
- Formal scaling is $O(N^{3\text{-}4})$ in electron count.

QMCPACK

http://www.qmcpack.org



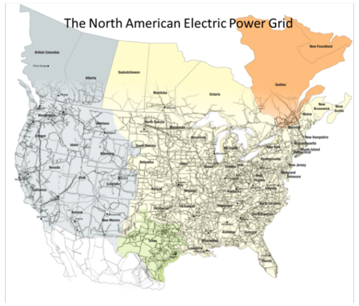
Strong scaling on Mira@ALCF of time to scientific solution for NiO supercells. Additional concurrency is required to increase efficiency and reduce time spent in Monte Carlo equilibration.

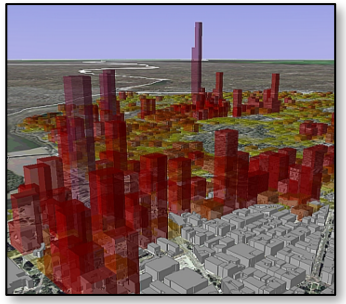
PI: Paul Kent (ORNL)

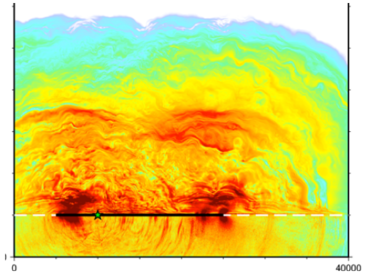
EXASCALE COMPUTING PROJECT

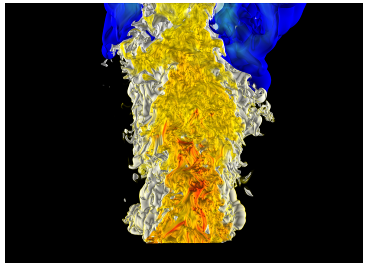
ECP: Economic Security Applications

- **Reliable and Efficient Planning of the Power Grid**
  - Intermittent renewable sources, electric vehicles, and smart loads are changing the behavior of the electric power grid and imposing new stochastics and dynamics for which the grid is not designed nor can accommodate. Move away from current approximations in analyzing the grid uncertainty, dynamics and optimization separately, which results in larger operating margins costing ~$5-15B annually. Deliver achievable margin bounds & potentially save billions of dollars annually
- **Smart City Dynamics**
  - Integrate modules for urban atmosphere and infrastructure heat exchange and air flow; building energy demand at district or city-scale, generation, and use; urban dynamics & activity based decision, behavioral, and socioeconomic models; population mobility and transportation; energy systems; water resources
- **Earthquake Hazard Risk Assessment**
  - Move computational earthquake hazard and risk assessments to a frequency range (~10Hz) relevant to estimating the risk to key engineered systems. Provide first strong coupling and linkage between simulations of earthquake hazards (ground motions) and risk (structural system demands)
- **High-Efficiency, Low-Emission Combustion Engine and Gas Turbine Design**
  - Predictive simulation of the relevant in-cylinder processes in a low-temperature reactivity-controlled compression ignition internal combustion engine that is more thermodynamically favorable then existing engines, with potential for groundbreaking efficiencies yet limiting pollutant formation





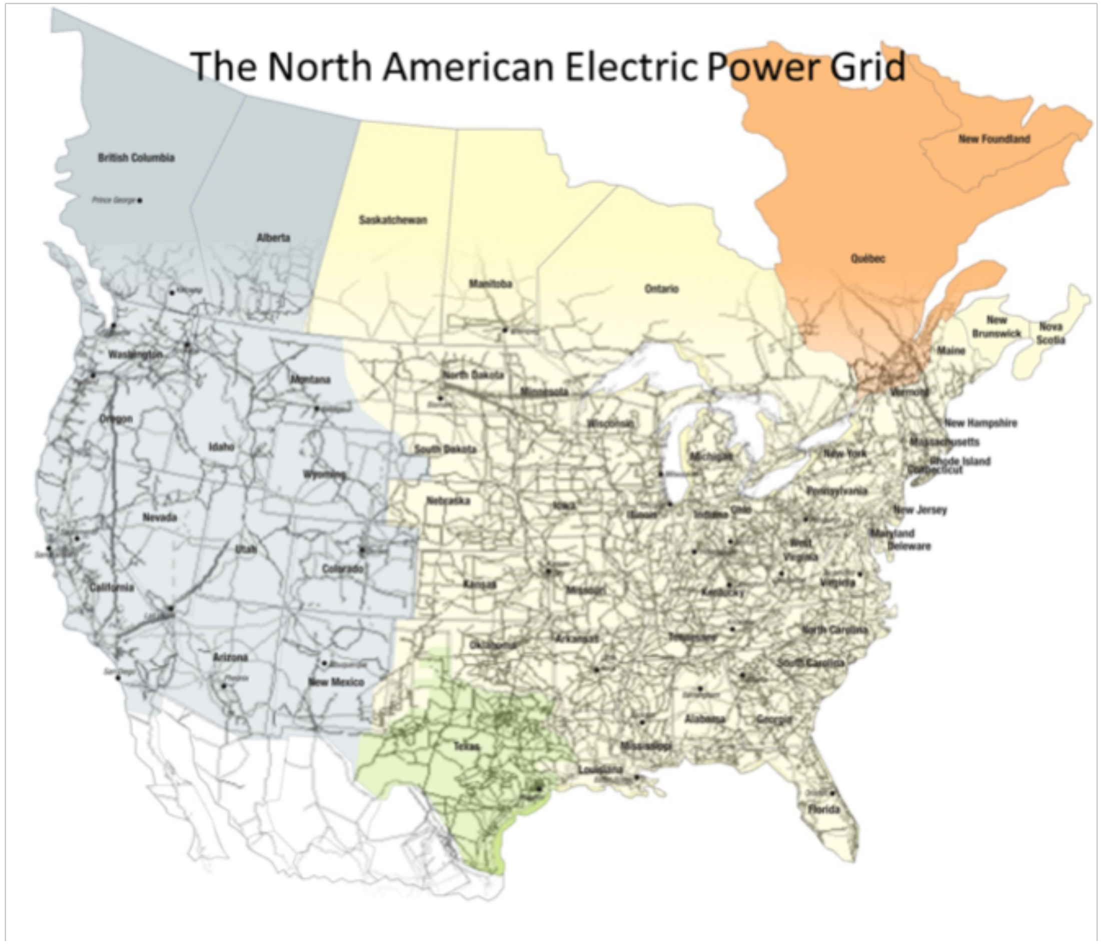




Optimizing Stochastic Grid Dynamics at Exascale

Intermittent renewable sources, electric vehicles, and smart loads will vastly change the behavior of the electric power grid and impose new stochastics and dynamics that the grid is not designed for nor can easily accommodate

- Optimizing such a stochastic and dynamic grid with sufficient reliability and efficiency is a monumental challenge
- Not solving this problem appropriately or accurately could result in either significantly higher energy cost, or decreased reliability inclusive of more blackouts, or both
- Power grid data are clearly showing the trend towards dynamics that cannot be ignored and would invalidate the quasi-steady-state assumption used today for both emergency and normal operation
- The increased uncertainty and dynamics severely strains the analytical workflow that is currently used to obtain the cheapest energy mix at a given level of reliability
- The current practice is to keep the uncertainty, dynamics and optimization analysis separate, and then to make up for the error by allowing for larger operating margins
- The cost of these margins is estimated by various sources to be in $5-15B per year for the entire United States
- The ECP grid dynamics application can result in the best achievable bounds on these errors and thus resulting in potentially billions of dollars a year in savings.

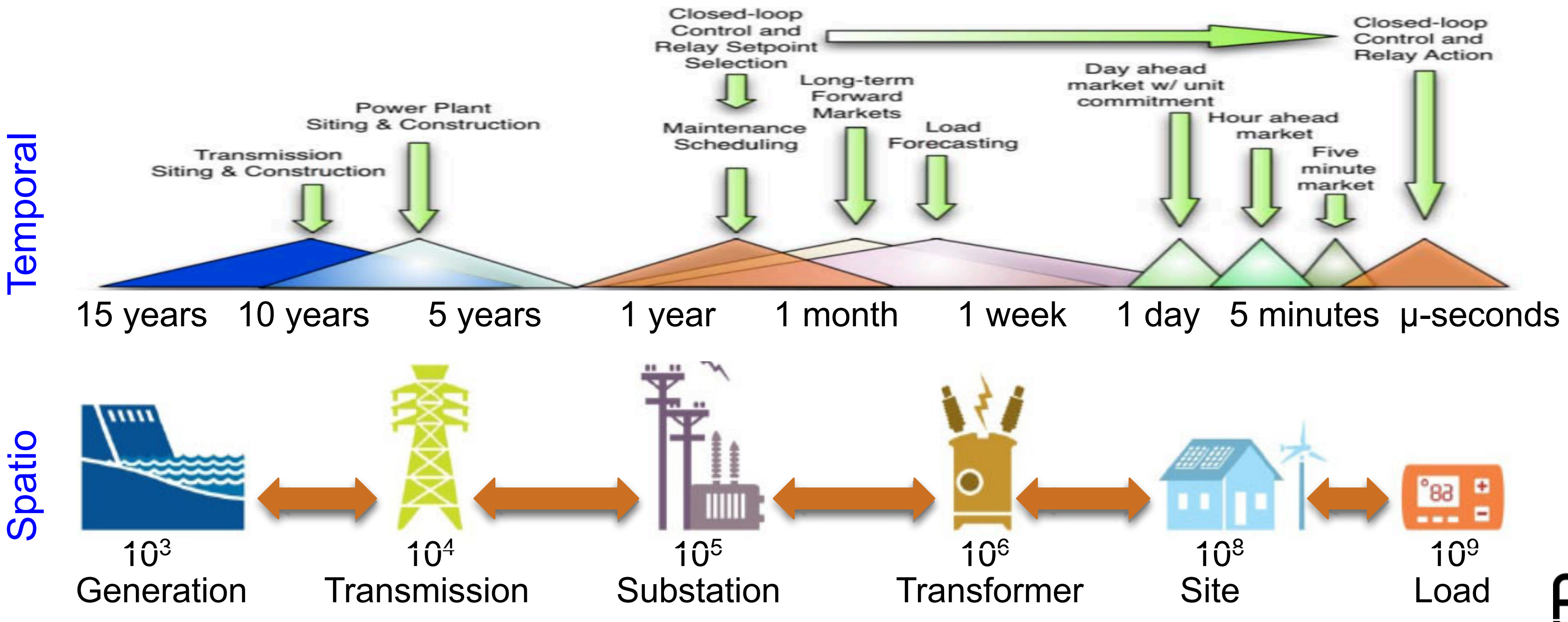


EXASCALE COMPUTING PROJECT

PI: Zhenyu (Henry) Huang (PNNL)

Simulating the Power Grid: Challenges

- Multi-scale spatio-temporal modeling and simulation with stochasticity
  - From micro-second to decades
  - From 10^3 generators nodes to 10^9 end-use devices
- Large-scale data assimilation for state and parameter calibration
  - Petabyte data/year from high-speed sensors and smart meters.
- Modeling of multi-system dynamics and dependency
  - Grid, Communication, Gas-pipeline, Weather/Wind/Solar, Water

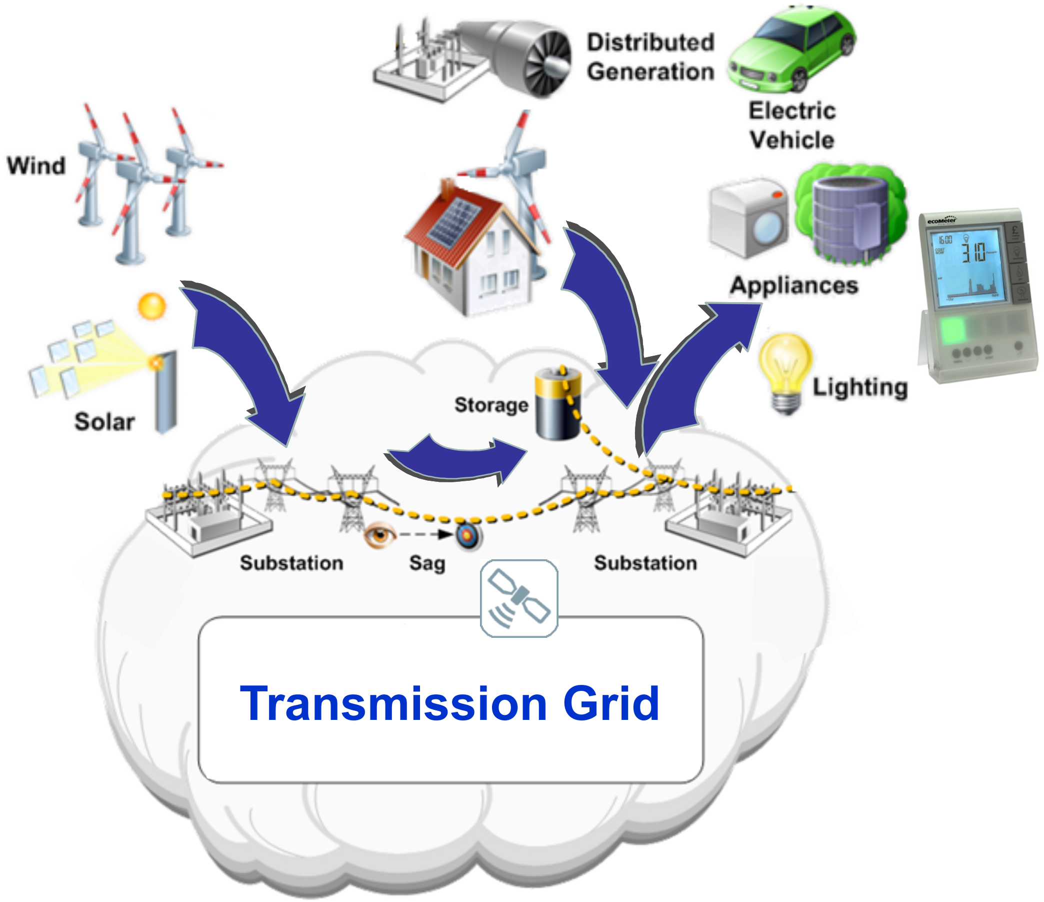


PI: Zhenyu (Henry) Huang (PNNL)

EXASCALE COMPUTING PROJECT

Expected Outcomes – Impact to the Grid

- Understand dynamic behaviors of the large-scale grid with uncertainty in high-impact events such as cascading failures
- Optimize the planning, operation, and control of the future power grid
- Accelerate penetration of clean energy and distributed energy sources (smart loads, electric vehicles, …)
- Improve grid reliability and efficiency in increased complexity, dynamics, and uncertainty

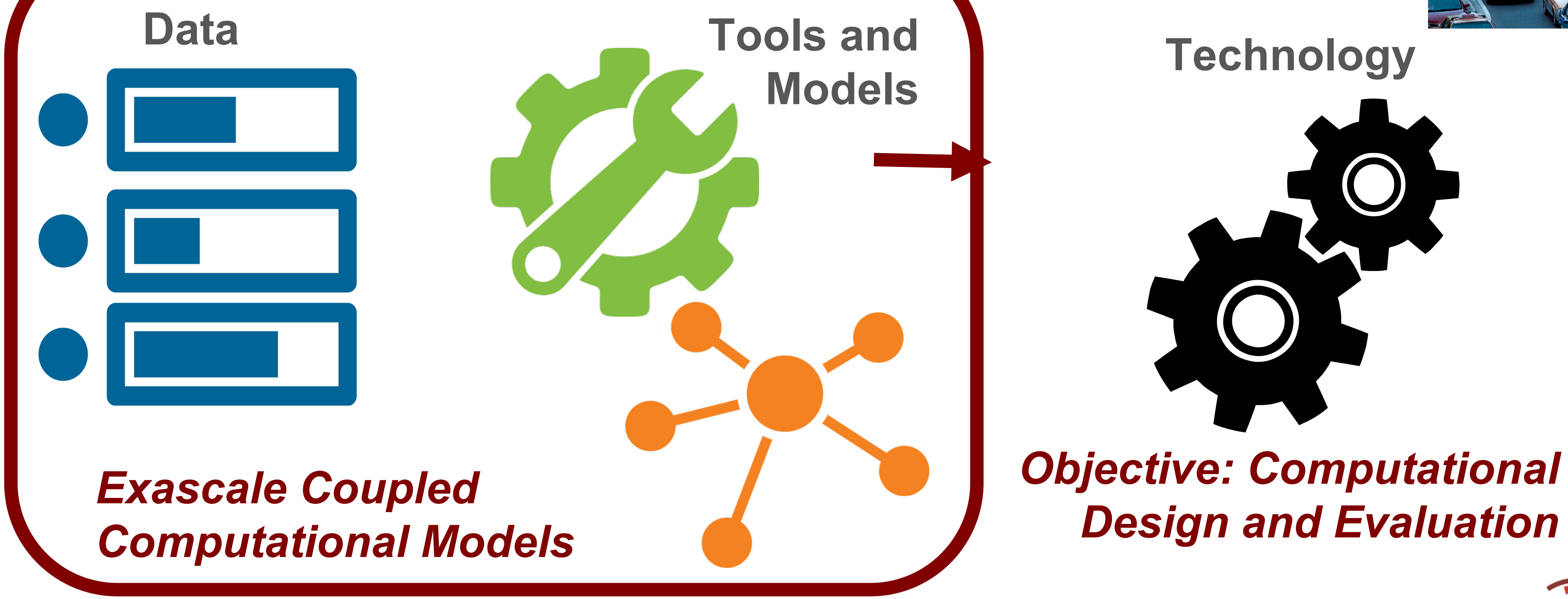


PI: Zhenyu (Henry) Huang (PNNL)

EXASCALE COMPUTING PROJECT

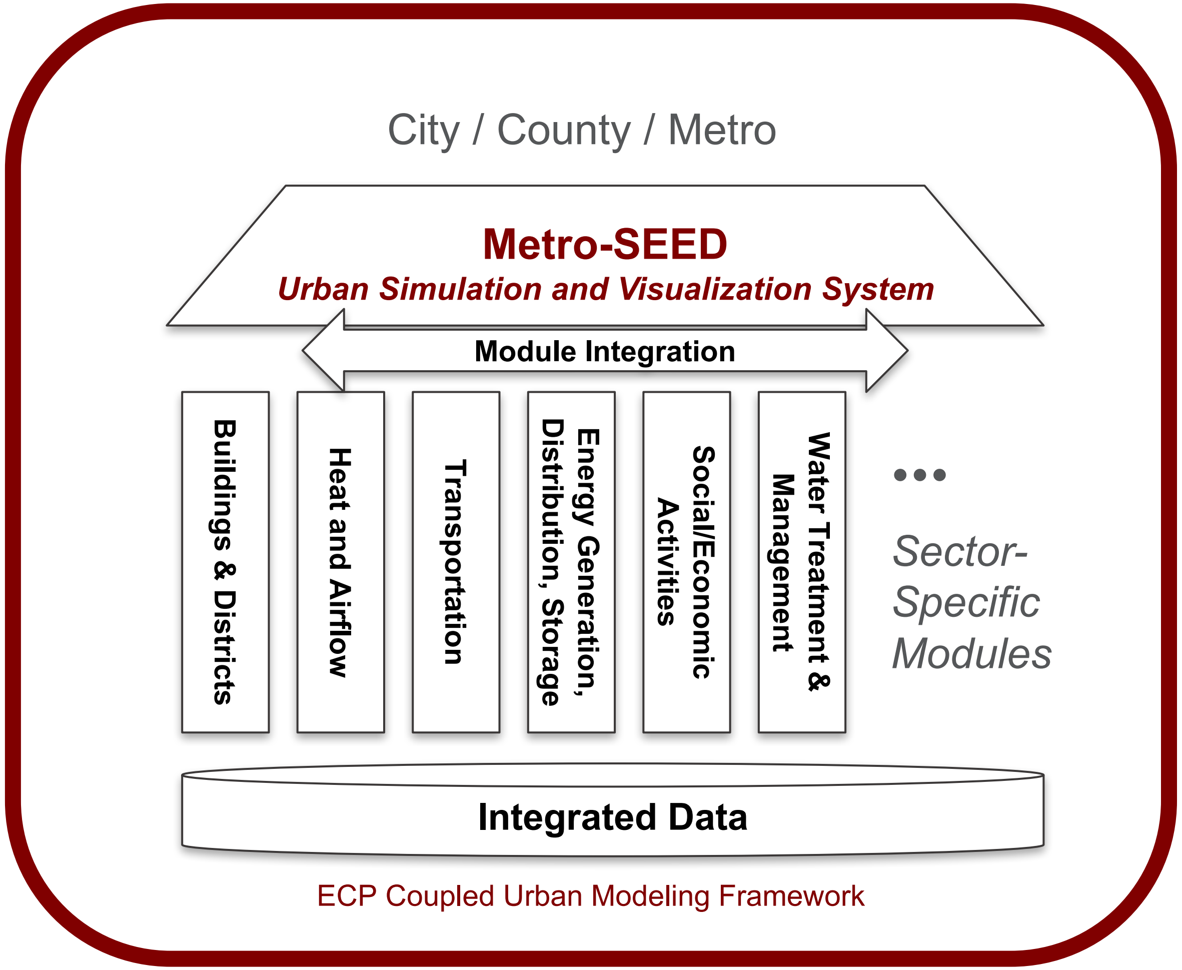
Current Urban Designs Typically Rely on Heuristics or Isolated System Models

- Approach Cities as *integrated, dynamic systems* vs. *places*
- Computational Modeling and integrated data for design, energy planning, and analysis — from buildings to transportation to grid



PI: Charles Catlett (ANL)

Coupled Models will Require New Data Integration and Coupling Mechanisms

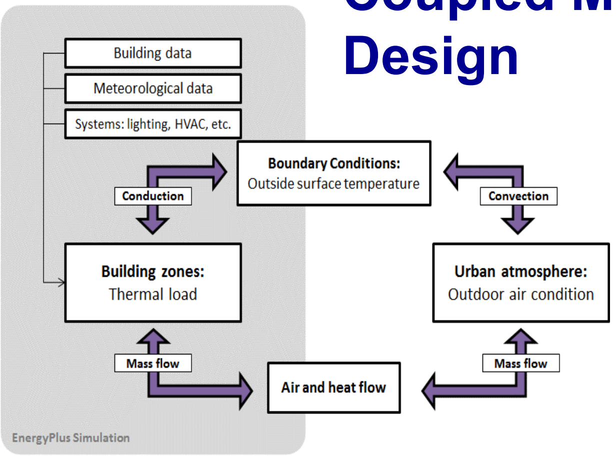


PI: Charles Catlett (ANL)

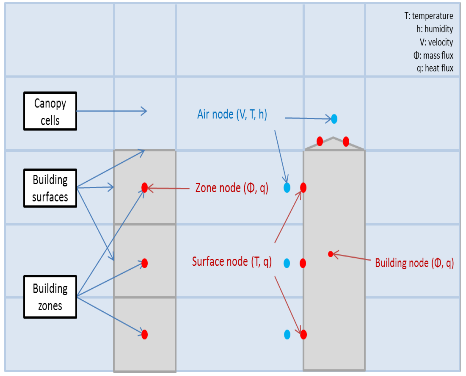
Initial Coupling: Building Energy Demand and Urban Canyon Heat Flow

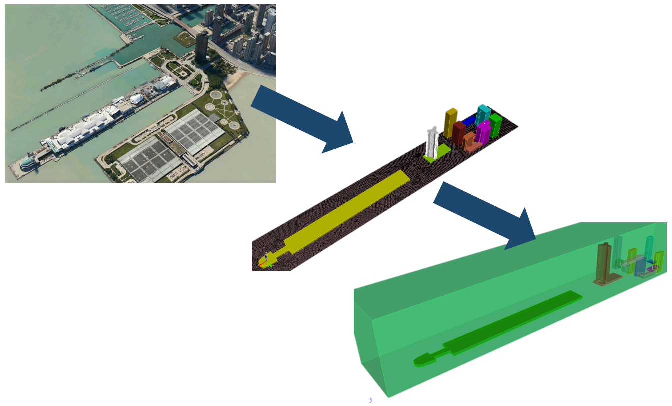
- **Identify and integrate existing and new data sources**
  (environment, air quality, solar loading, vehicle flows, social/economic activity…)
- **Develop capabilities to use urban data sources for model-building.**
- **Quantify and characterize data content, forms, and flow rate/volume for exchanging data between coupled models.**

Coupled Model Design



Implementation





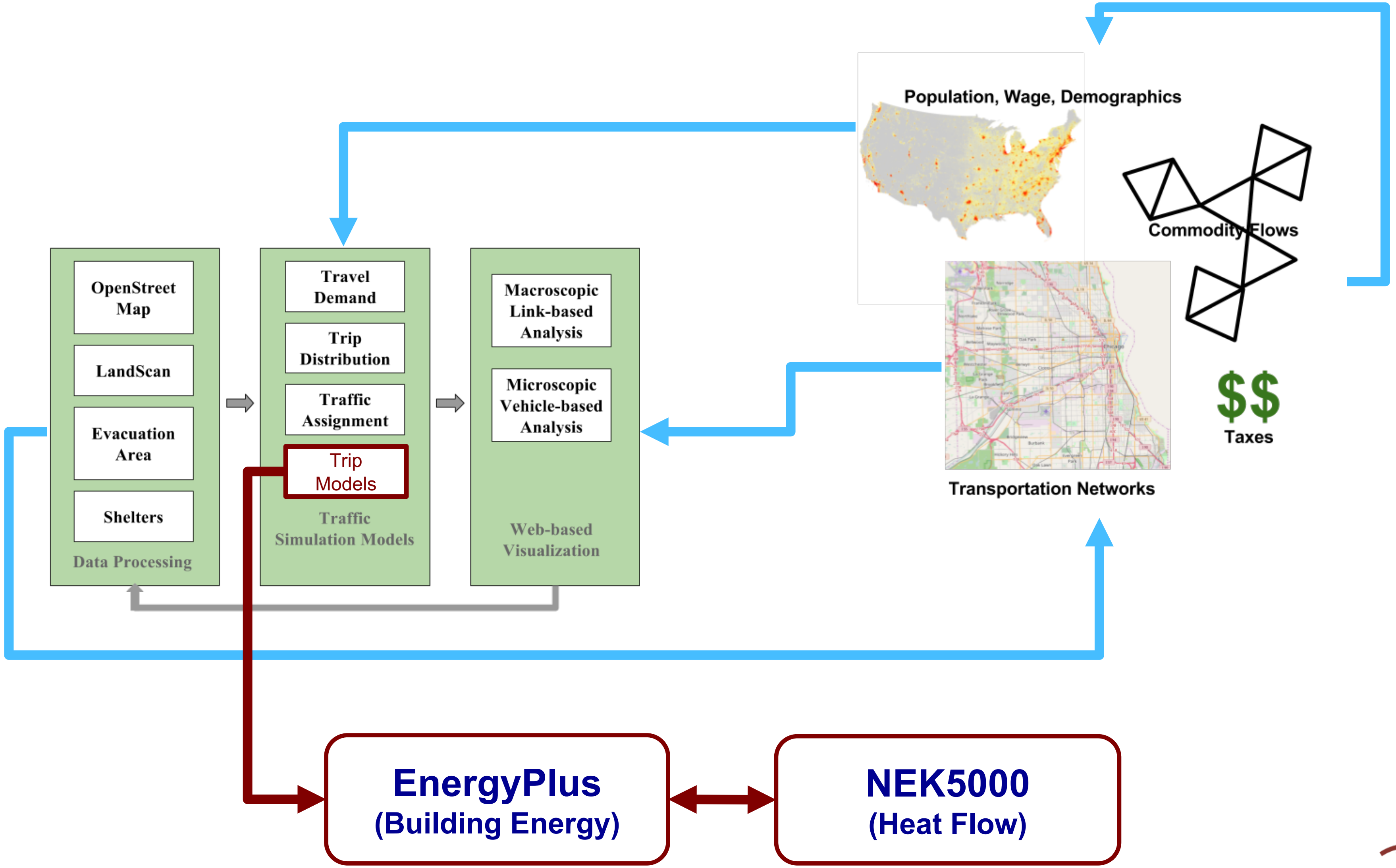
Model Setup from Test Area

(Chicago Loop)

PI: Charles Catlett (ANL)

EXASCALE COMPUTING PROJECT

Next: Transportation and Social-Economic Systems

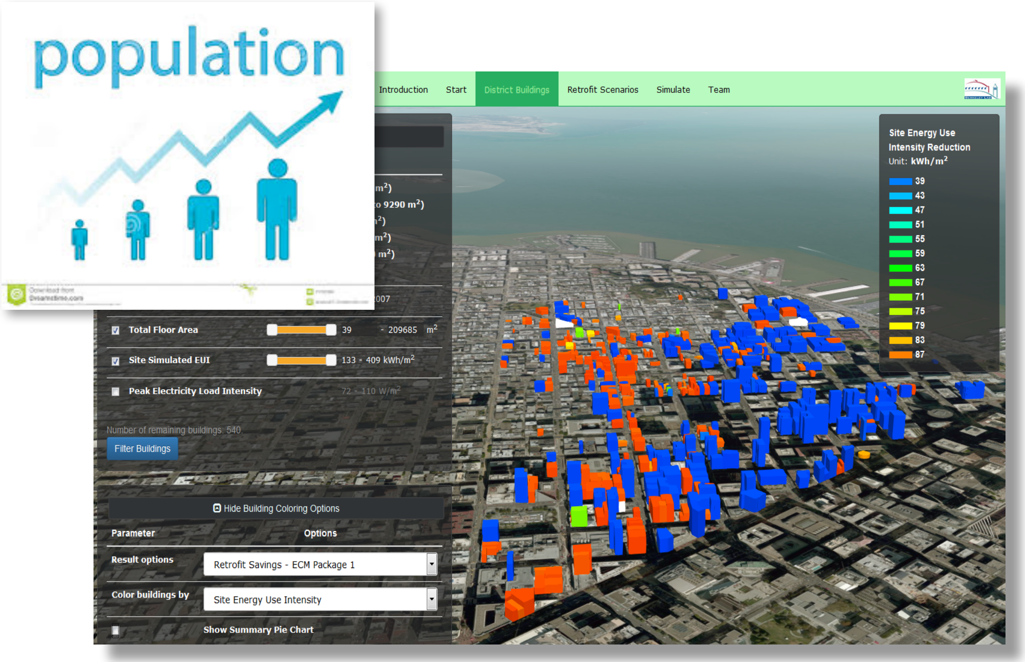


PI: Charles Catlett (ANL)

EXASCALE COMPUTING PROJECT

Integrate New Data Sources Into Coupled Models

- **Urban sensing** to measure air quality, weather, activity
- **National and regional data** combined for custom city analysis
- **Integrate building stock and energy data** to evaluate retrofits



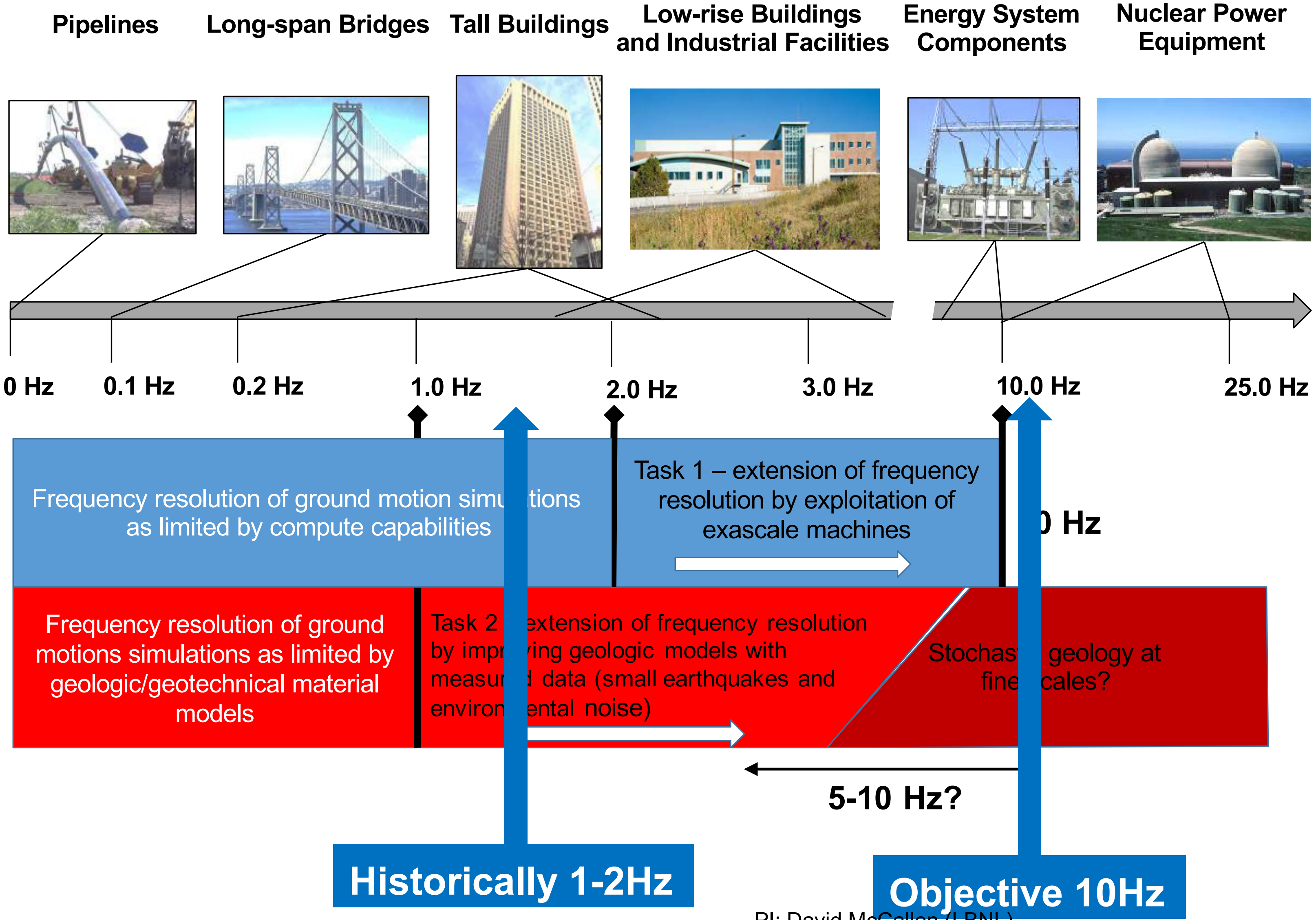
City Building Energy Saver Demo with San Francisco

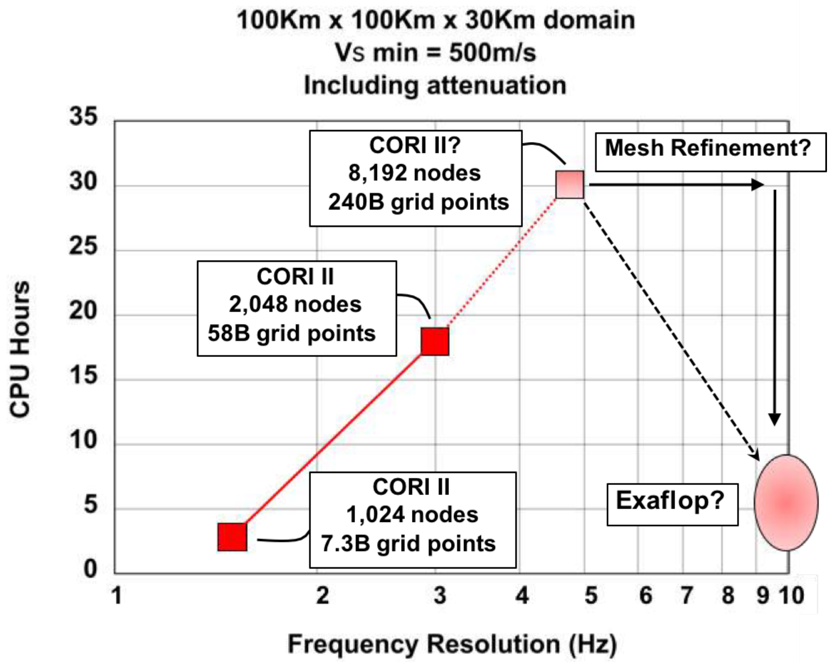
Urban Data Integration

PI: Charles Catlett (ANL)

Exascale Enables Exploitation of Ground Motion Data and Appropriately Defined Stochastic Geology

And allows simulations at the frequencies of interest





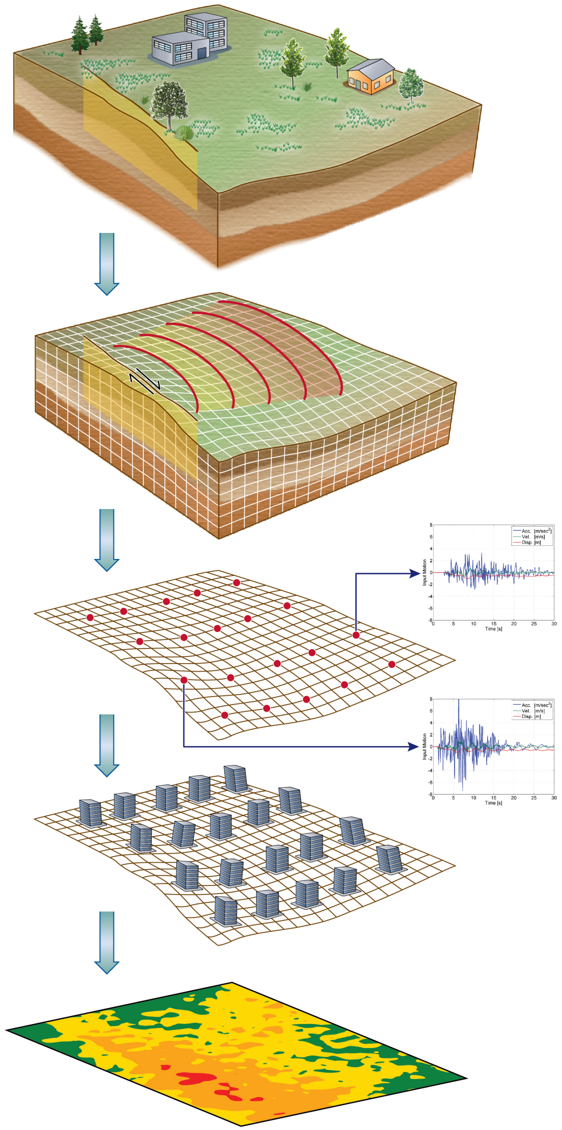
PI: David McCallen (LBNL)

EXASCALE COMPUTING PROJECT

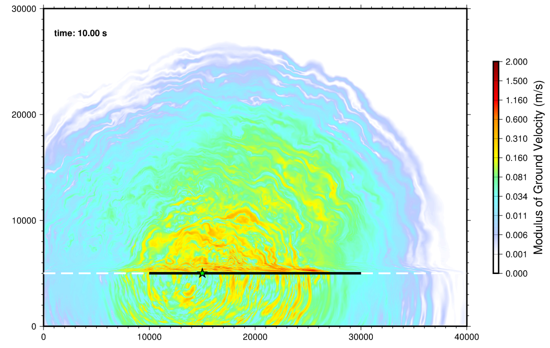
Exascale Enables an Earthquake Framework for Hazard *and* Risk

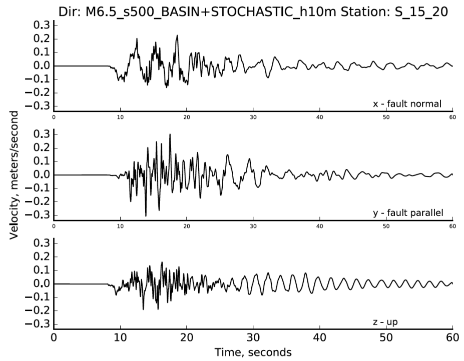
Run much bigger models much faster
- Very large models for resolving 10Hz
- Many realizations to account for uncertainties (e.g. fault rupture)

Representation of fine-scale geology
- Stochastic geology
- Data inversion

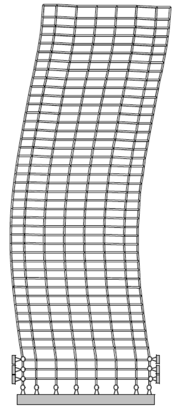


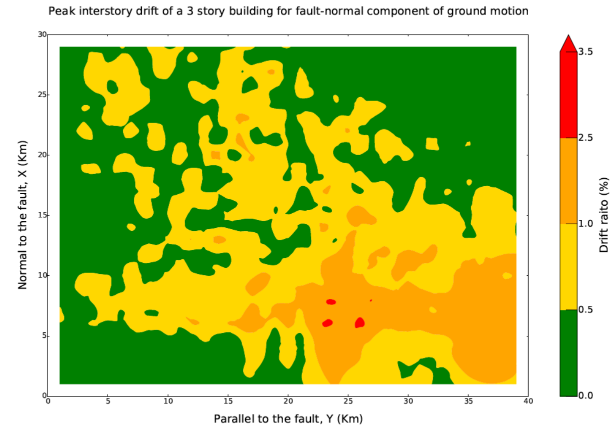
Regional Scale Hazard — Geophysics





Facility Risk — Engineering

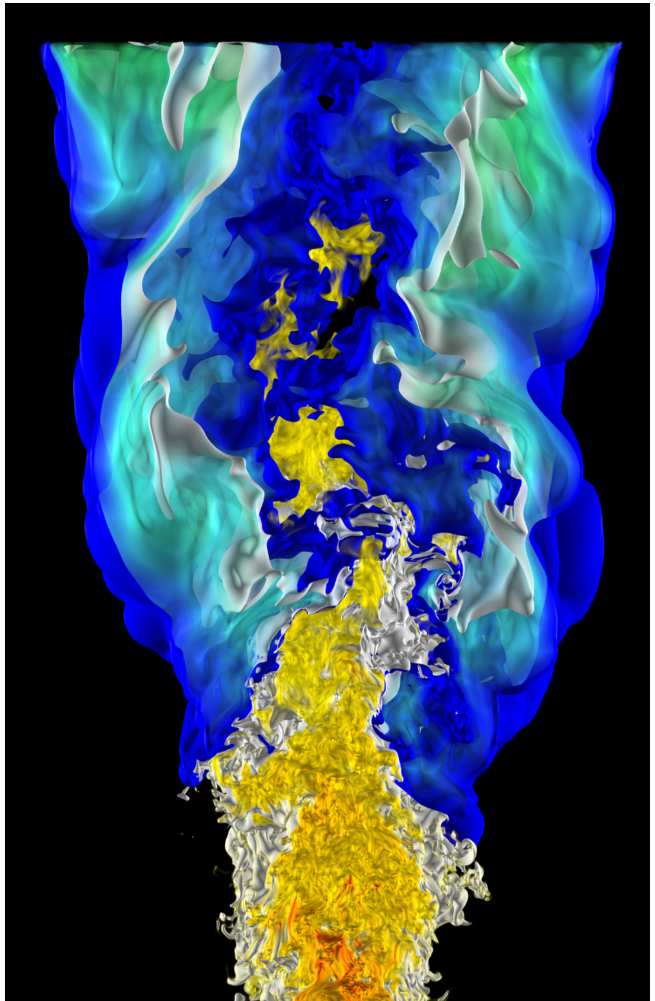




ECP EXASCALE COMPUTING PROJECT

PI: David McCallen (LBNL)

Transforming Combustion Science and Technology with Exascale Simulations



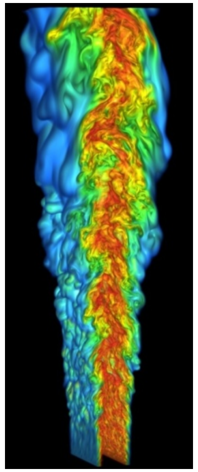
Lifted jet flame showing the separation of low-temperature reactions (yellow to red) and the high-temperature flame (blue to green)

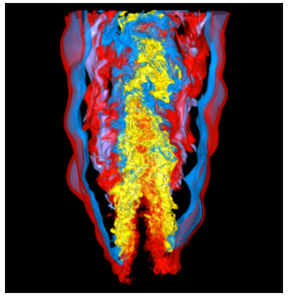
- Direct Numerical Simulation (DNS) of a turbulent lifted jet flame stabilized on pre-ignition species reveals how low-temperature reactions help stabilize the flame against the disruptive effects of high velocity turbulence
- Understand the role of multi-stage ignition in high pressure diesel jet flames
- High-fidelity geometrically faithful simulation of the relevant in-cylinder processes in a low temperature reactivity controlled compression ignition (RCCI) internal combustion engine that is more thermodynamically favorable then existing engines, with potential for groundbreaking efficiencies yet limiting pollutant formation
- Develop DNS and hybrid DNS-LES adaptive mesh refinement solvers for transforming combustion science and technology through capable exascale
- Prediction of the relevant processes - turbulence, mixing, spray vaporization, ignition and flame propagation, and soot/radiation -- in an RCCI internal combustion engine will become feasible

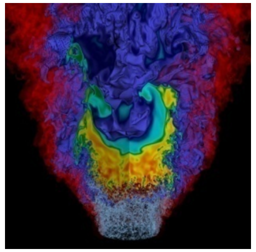
PI: Jacqueline Chen (SNL)

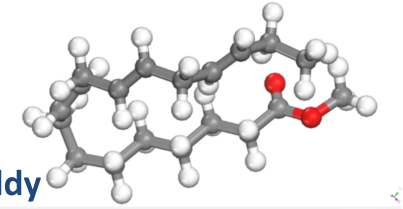
EXASCALE COMPUTING PROJECT

Exascale Accelerates the Design of Fuel Efficient & Clean Burning Automobiles, Trucks, and Gas Turbines for Power Generation









Exaflops Large-Eddy Simulation (LES)

Petaflops Sub-Model Development

Teraflops Direct Numerical Simulation (DNS)

Gigaflops Unsteady Laminar Flame Simulations

Quantum chemical Kinetics Simulations

Foundational fuels (C1-C4)

Hydrogen

Syngas

Methane

Ethylene

Alcohols

Ethanol

Di-methyl Ether

Butanol

Bio-Diesels

Methyl Butanoate

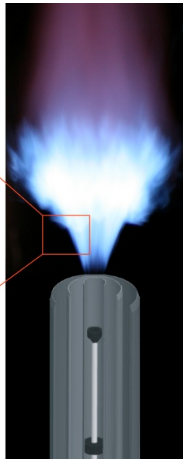
Methyl Decanoate

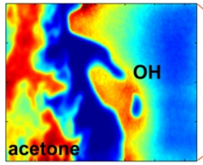
Turbulent Flame Experiments

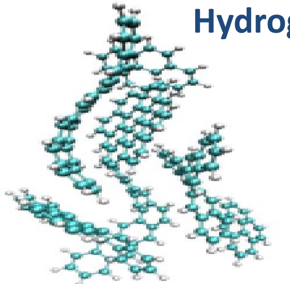
Mechanism Validation Experiments

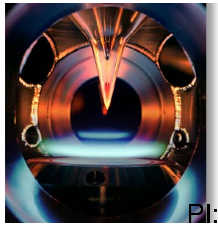
Laminar Flame/Ignition Experiments

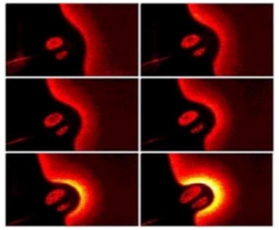
Chemical Kinetic Experiments

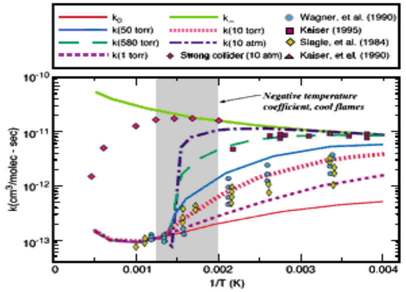








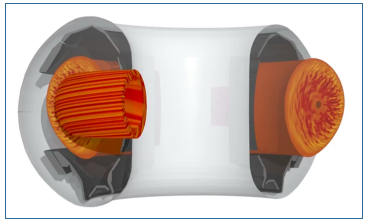




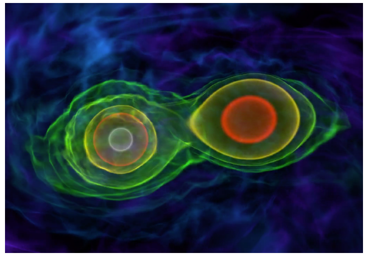
PI: Jacqueline Chen (SNL)

ECP: Scientific Discovery Applications

- **Predict and Control Stable ITER Operational Performance**
  - Simulate and predict the multiple spatio-temporal scale processes in magnetic plasma confinement, predict performance of future fusion reactors such as ITER, and accelerate development of commercial fusion reactors



- **Demystify Origin of Chemical Elements**
  - Stellar explosion simulations help to explain the origin of the elements (those heavier than iron), a longstanding problem in physics. They also help to define the conditions for astrophysical nucleosynthesis that guide key U.S. experimental nuclear data experiments



EXASCALE COMPUTING PROJECT

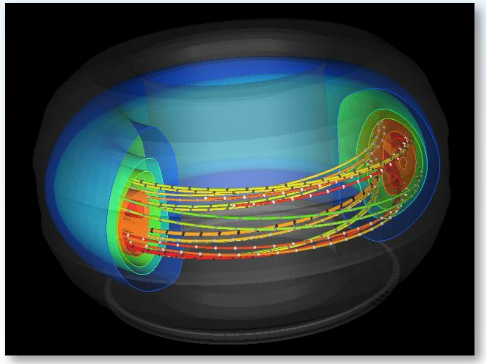
I’m a Big Fusion Fan
and not just in movies . . .

EXASCALE COMPUTING PROJECT

Exascale App: Magnetic Fusion Energy

Gaps and opportunities

- Prepare for and exploit ITER and other planned and future international major experiments (e.g., JET-DT, JT-60SA, Wendelstein 7-X)

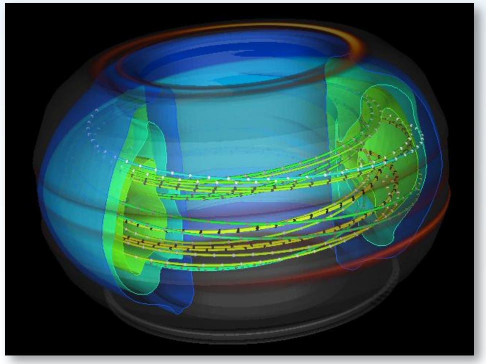


Simulation challenge problems

- Whole-device fusion reactor simulation: Tightly coupled full-f/delta-f models and loosely coupled source/boundary models, including electron/ion kinetics, MHD, and energetic particles in core and edge regions
- Simulate and characterize tokamak disruptions and mitigations, incorporating kinetics, MHD, and fast particles
- Plasma boundary region analysis: Edge kinetic effects, material interaction, radiation and detachment, power and particle exhaust

Prospective outcomes/impact

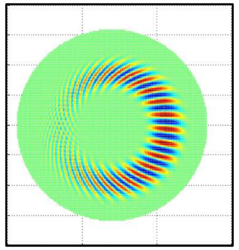
- Prepare for and fully simulate ITER experiments and increase ROI of validation data and understanding
- Prepare for beyond-ITER devices, such as nuclear science facilities and DEMO

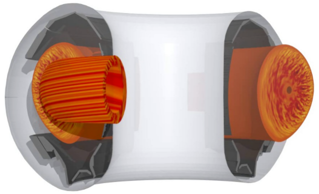


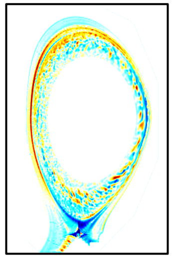
EXASCALE COMPUTING PROJECT

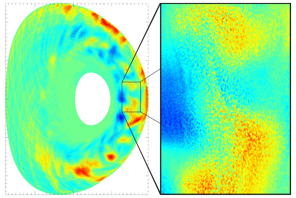
Evolution of Kinetic Modeling Capability with Increasing Compute Power

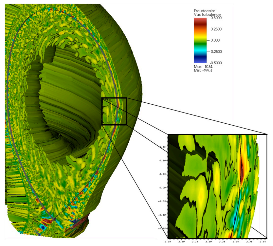
Increasing fidelity in geometry and turbulence physics in evolving mean background

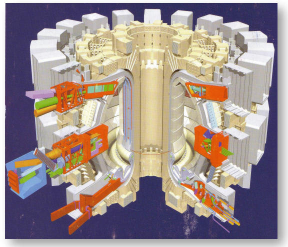


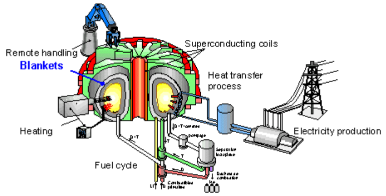












Gigaflops

5-D electrostatic ion physics in simplified circular cylindrical geometry

Teraflops

Core: 5D ion-scale electromagnetic physics in torus

Edge: ion+neutral electrostatic physics in torus

Petaflops

Core: 5D i+e scale electrostatic physics

Edge: i+e+n electroostatic physics

Exaflops

Core-edge coupled 5D electromagnetic study of whole-device ITER, including turbulence, transport, large-scale instability plasma-material interaction, rf heating, and energetic particles

Beyond

6D whole device modeling of all the relevant fusion reactor science with burn control

EXASCALE COMPUTING PROJECT

PI: Amitava Bhattacharjee (PPPL)

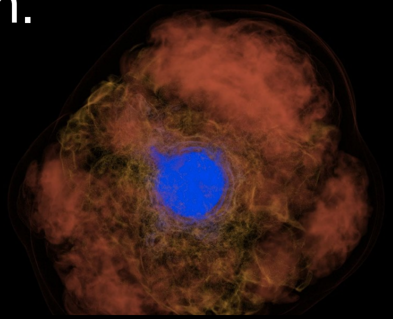
Core Questions in Nuclear Astrophysics

- Where did the elements in the universe come from?
- What is the behavior of matter at extreme densities?
- What are the fundamental properties of neutrinos?
- What are the sources of gravitational waves?

Energetic astrophysical phenomena are our laboratories for studying extreme physics in regimes inaccessible on Earth.

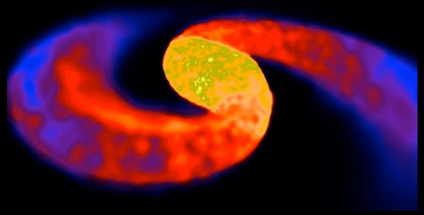
core-collapse supernovae

explosion of massive stars reaching extremes of nuclear burning and neutrino interactions



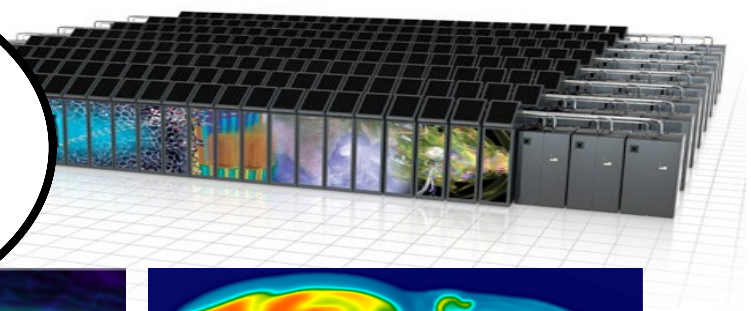
neutron star mergers

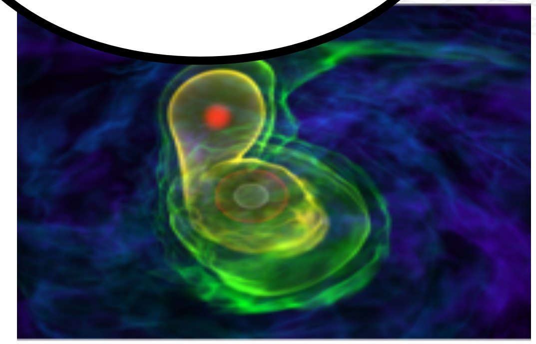
inspiral and coalescence of compact stars reaching extremes of gravity and density

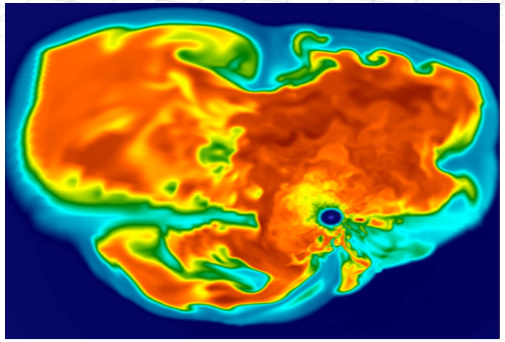


PI: Daniel Kasen (LBNL)

ExaStar nuclear astrophysics simulations







***ExaStar* simulations are essential to:**

Guide future nuclear physics experimental programs
Simulating the astrophysical conditions of r-process nucleosynthesis directly impacts which nuclear data and rates are most important to measure.

Provide reliable templates for gravitational wave and neutrino detectors
Complexity and low signal-to-noise of the data requires templates for matching and physically interpreting signals.

Interpret X-ray, gamma-ray, and optical signals
Modeling the electromagnetic data directly constrains the nucleosynthetic yields, energetics, and explosion geometry.

***ExaStar* capabilities will add value to:**

Experimental nuclear physics data

Satellite observations and gravitational wave detections of astrophysical phenomena

Neutrino experimental data, including solar and nuclear reactor experiments.

Need for multi-physics simulations of reactive fluid flow of broad relevance across science, engineering and industry application

DUNE DEEP UNDERGROUND NEUTRINO EXPERIMENT

LIGO

FRIB

LSST

PI: Daniel Kasen (LBNL)

ExaStar

Quintessential multi-physics simulation

block-structured adaptive mesh refinement

dynamics

Compressible Hydrodynamics
Finite Volume Godunov methods

Magnetohydrodynamics

reactions

Nuclear Reaction Networks
sparse stiff linear systems

Lagrangian tracer Particles
for post-processing

radiation

Two-moment Transport
Semi-implicit Discontinuous Galerkin methods

Boltzmann Transport
Implicit Monte Carlo methods

gravity

Approximate relativity
Conformally flat approx.
Elliptic eq w/ multigrid methods

Full general relativity
e.g., DG methods for Einstein Equations

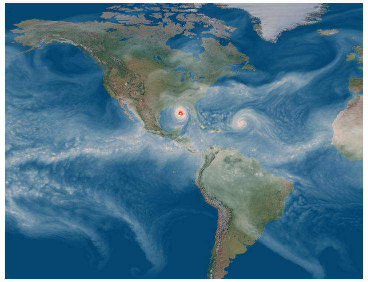
microphysics

Equation of State Module
local, tabulated

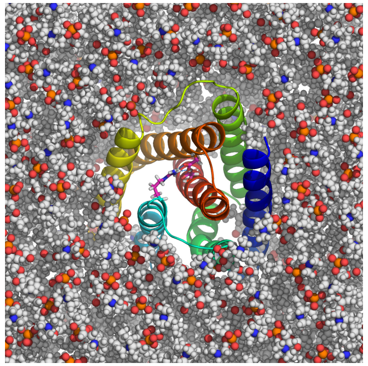
Opacity Module
tabulated, scattering kernels

PI: Daniel Kasen (LBNL)

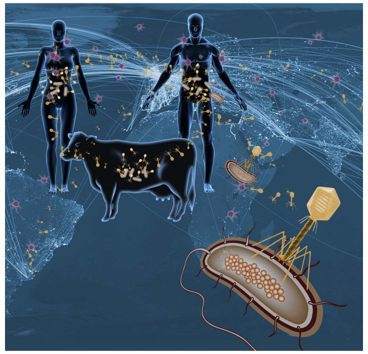
ECP: Earth Systems Applications



- **Accurate Regional Impact Assessments in Earth Systems**
  - Cloud-resolving Earth systems model with throughput necessary for multi-decade, coupled high resolution climate simulations with substantial reduction of major systematic errors in precipitation via a realistic convective storm treatment. Improve ability to assess regional water cycles that directly affect multiple sectors of the U.S. economies (agriculture & energy production)



- **Stress-Resistant Crop Analysis and Catalytic Conversion of Biomass-Derived Alcohols**
  - Deliver molecular and materials modeling capabilities for development of new feedstocks biomass on marginal lands, and new catalysts for the conversion of these feedstocks to usable biofuels and other products



- **Meteganomics for Analysis of Biogeochemical Cycles, Earth Systems, Environmental Remediation**
  - Microorganisms - central players in earth system, environmental remediation, food production, human health - occur naturally as “microbiomes” - communities of thousands of microbial species of varying abundance and diversity, each contributing to the function of the whole. <1% of millions of worldwide microbe species have been isolated, cultivated, and sequenced. Collections of microbial data are growing exponentially, representing untapped info useful for environmental remediation and the manufacture of novel chemicals and medicines.

EXASCALE COMPUTING PROJECT

ACME-MMF Cloud Resolving Climate Model

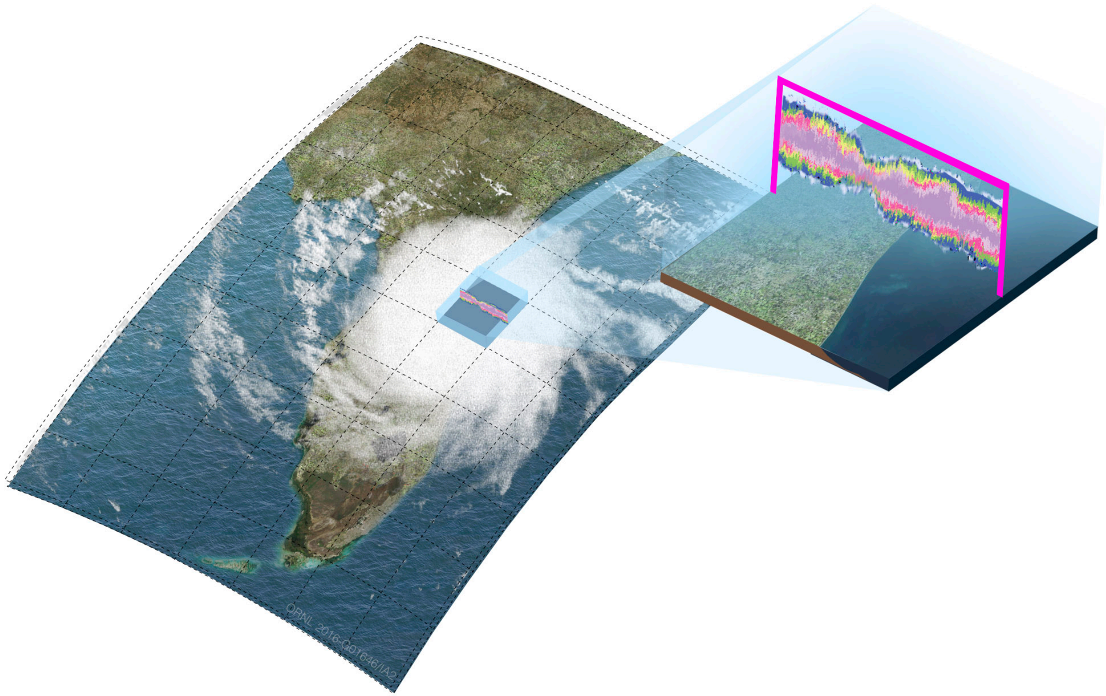
- Develop capability to assess regional impacts of climate change on the water cycle that directly affect the US economy such as agriculture and energy production.
- A cloud resolving climate model is needed to reduce major systematic errors in climate simulations due to structural uncertainty in numerical treatments of convection – such as convective storm systems
- Challenge: Cloud resolving climate model using traditional approaches requires Zettascale resources.
- ACME-MMF: Use a multiscale approach ideal for new architectures to achieve cloud resolving convection on Exascale resources

Convective storm system nearing the Chicago metropolitan area
http://www.spc.noaa.gov/misc/AbtDerechos/derechofacts.htm

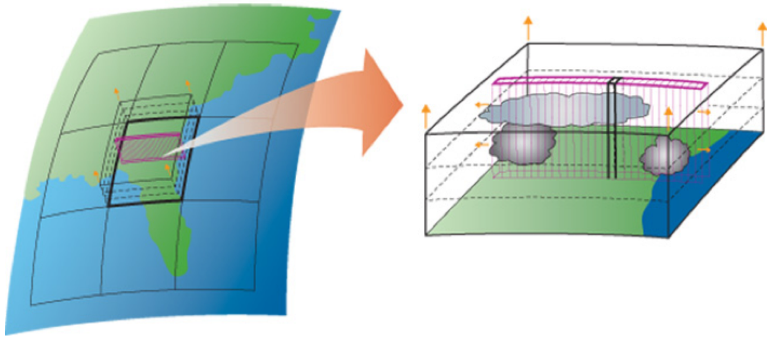
PI: Mark Taylor (SNL)

EXASCALE COMPUTING PROJECT

The ACME Multiscale Modeling Framework (MMF)



- ACME-MMF approach addresses structural uncertainty in cloud processes by replacing traditional parameterizations with cloud resolving “superparameterization” within each grid cell of global climate model
- Super-parameterization dramatically increases arithmetic intensity, making the MMF approach one of the few ways to achieve exascale performance on upcoming architectures.
- Exascale + MMF approach will make it possible for the first time to perform climate simulation campaigns at cloud resolving resolutions.

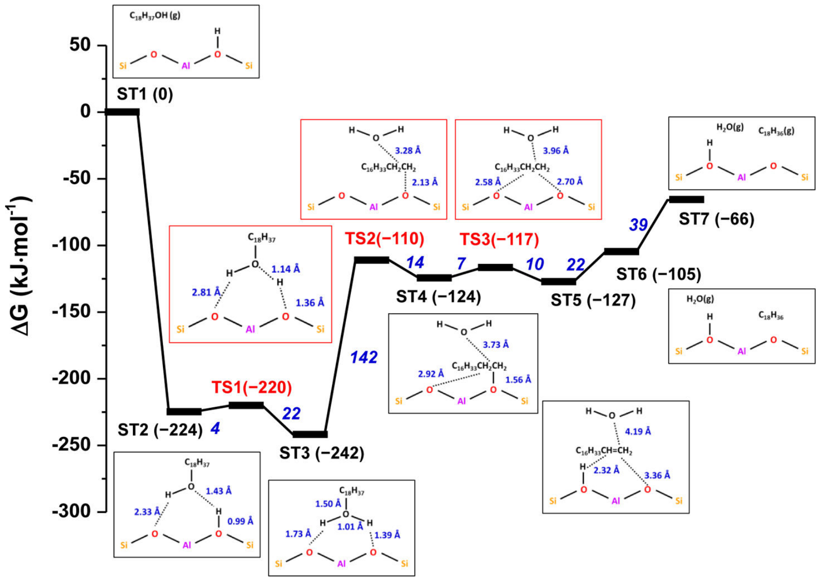


PI: Mark Taylor (SNL)

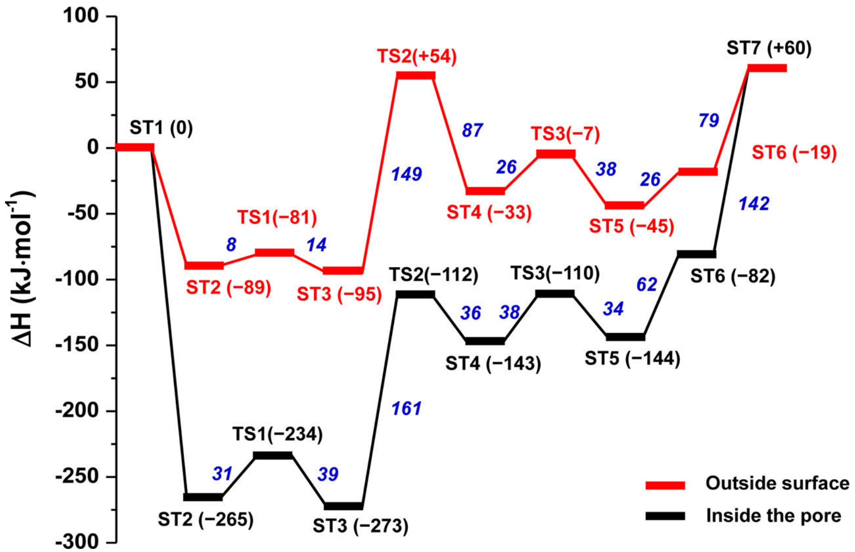
EXASCALE COMPUTING PROJECT

Chemistry: Catalyst Design

Simulation Gap	Needed to Address Gap	Impact of Gap Remaining
Current computational chemistry models used to describe catalysis do not provide the fidelity needed to reliably predict the rates, selectivity, and optimal conditions for the reactions of interest.	Extension of known high fidelity, computationally-intense models to simulation of activation energies, mass transport, and non-equilibrium dynamics in realistic catalytic materials systems.	Simulations are insufficient to efficiently guide the development of new, optimal catalysts; the development processes remain costly, time consuming, and dominated by trial-and-error.



A single reaction pathway



Multiple reaction pathways

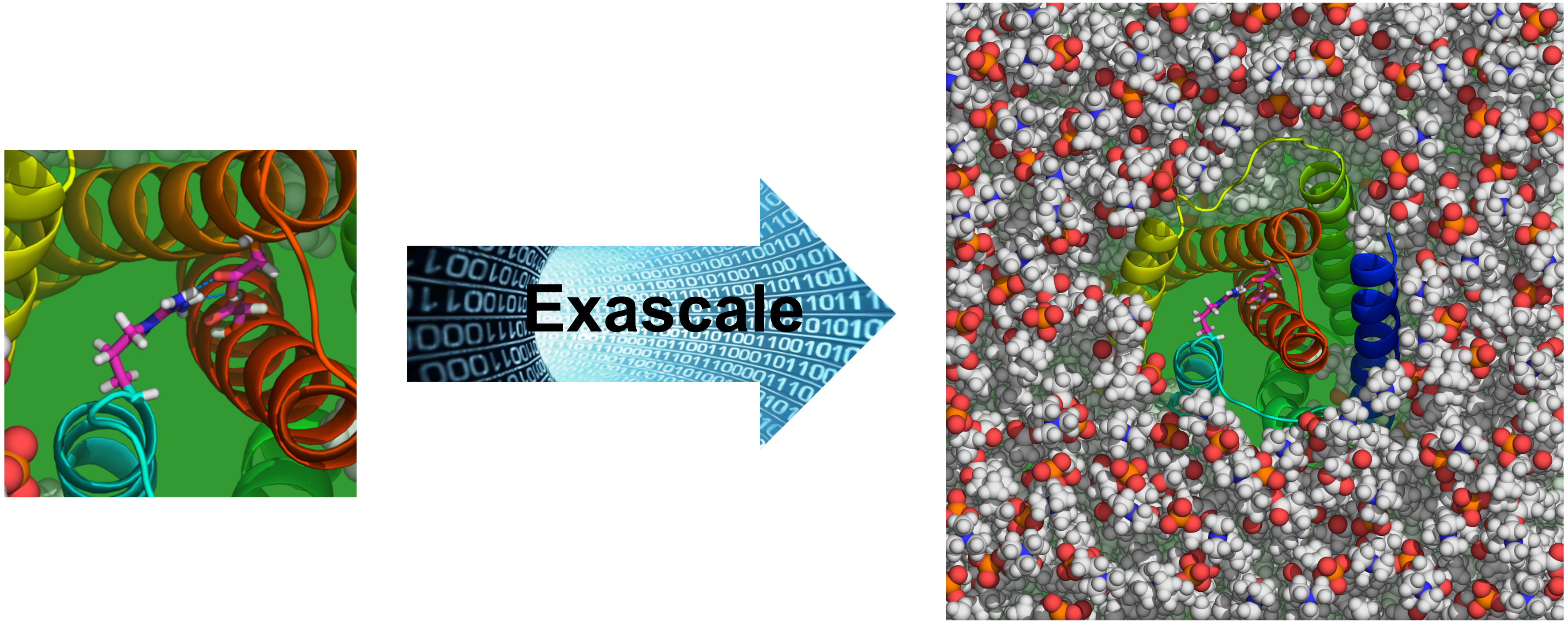
Song, et al., ACS Catalysis, 2016, 6, 878-889, doi:10.1021/acscatal.5b01217

PI: Thom Dunning (PNNL)

EXASCALE COMPUTING PROJECT

Chemistry: Biological Processes in Bioenergy Crops

Simulation Gap	Needed to Address Gap	Impact of Gap Remaining
Bioenergy crops use membrane proteins to sense and respond to unfavorable environmental conditions and stressors. Current simulations cannot represent these biological processes for realistic systems and timescales at the needed fidelity.	Biomolecular simulations of membrane proteins that can accurately describe the these complex systems at the relevant timescales. Such simulations would provide a first-principles basis for understanding how to maximize crop yields on marginal lands.	Lack of this simulation capability will make it difficult, even impossible, for plant systems biologists to improve the yields on or design improved bioenergy crops for marginal lands. This, in turn, will impede the uptake of clean and renewable energy sources.

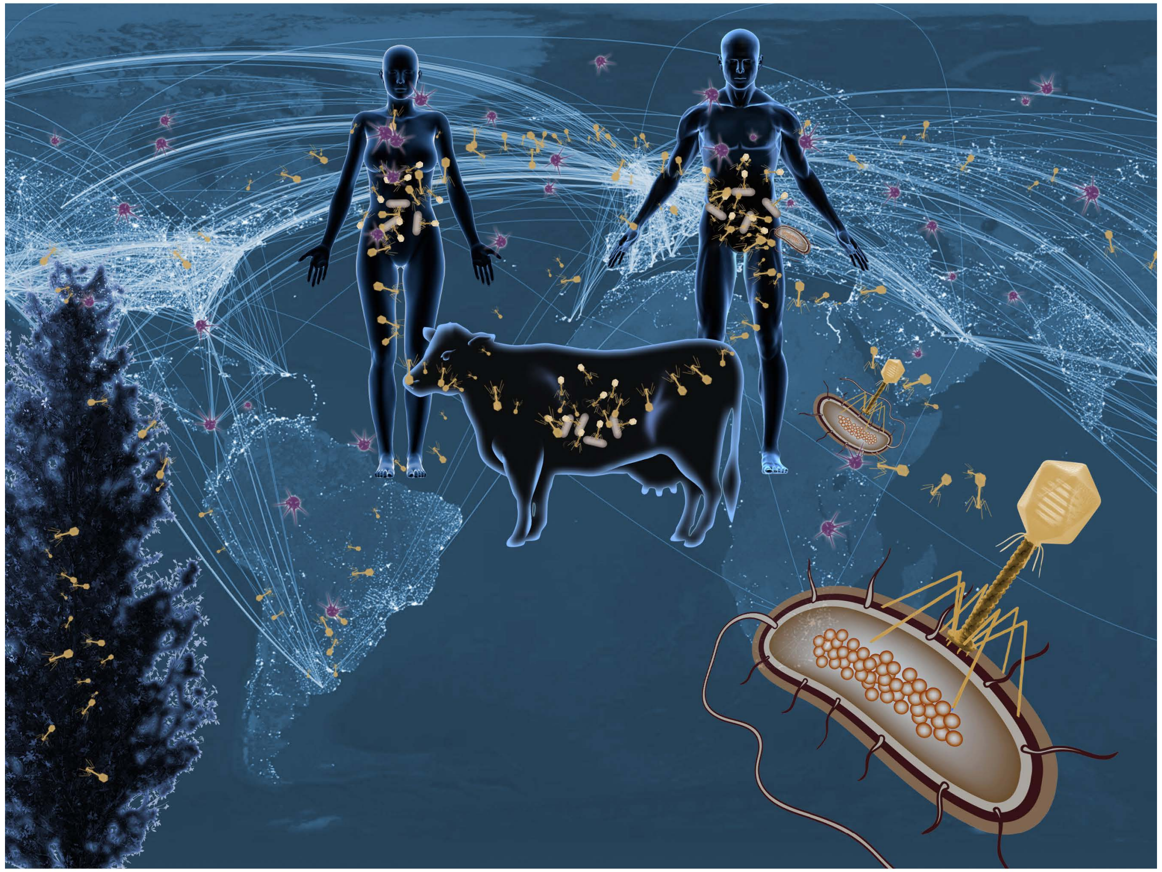


PI: Thom Dunning (PNNL)

EXASCALE COMPUTING PROJECT

Exascale Solutions for Microbiome Analysis

Microbiomes: integral to the environment, agriculture, health and biomanufacturing. Analyzing the DNA of these microorganism communities is a computationally demanding bioinformatic task, requiring exascale computing and advanced algorithms.



- Microorganisms are central players in climate change, environmental remediation, food production, human health
- Occur naturally as “microbiomes” - communities of thousands of microbial species of varying abundance and diversity, each contributing to the function of the whole
- <1% of millions of worldwide microbe species have been isolated and cultivated in the lab, and a small fraction have been sequenced
- Collections of microbial data are growing exponentially, representing untapped info useful for environmental remediation and the manufacture of novel chemicals and medicines.
- “Metagenomics” — the application of high-throughput genome sequencing technologies to DNA extracted from microbiomes — is a powerful method for studying microbiomes
- First assembly step has high computational complexity, like putting together thousands of puzzles from a jumble of their pieces
- After assembly, further data analysis must find families of genes that work together and to compare across metagenomes
- ExaBiome application is developing exascale algorithms and software to address these challenges

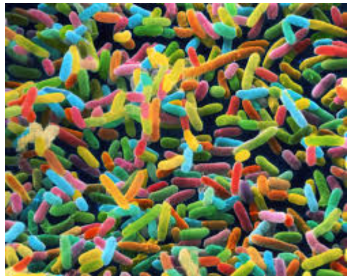
EXASCALE COMPUTING PROJECT

PI: Katherine Yelick (LBNL)

ExaBiome: Exascale Solutions for Microbiome Analysis

- **Microbes:** single cell organisms, such as bacteria and viruses
- **Microbiomes:** communities of 1000s of microbial species, less than 1% individually culturable in a lab (and thus sequenced)
- **Metagenomics:** genome sequencing on these communities (growing exponentially)
- Understanding and controlling microbiomes has many applications

Environment



Health

Bio-Energy

Bio-Manufacturing

PI: Katherine Yelick (LBNL)

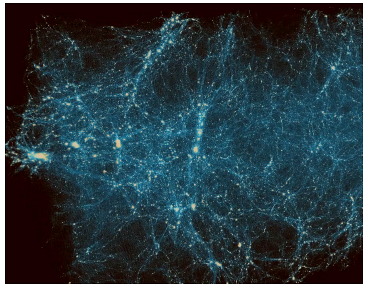
ExaBiome: Exascale Solutions for Microbiome Analysis

2015 Baseline	Challenge	2021 Exascale Enabled
1 thousand meta-genomes assembled with partial data	1 million meta-genomes will be collected by 2020	Assemble 1 million metagenomes based on complete data
50 million proteins clustered to find related families	Trillions of sequences will be in metagenome databases by 2020	Cluster billions of proteins for discovery and to unlock functional behavior
Limited ability to identify low abundance species	Metagenomes have millions of species with variable abundance	Discover new species and functions in large, complex metagenome data sets
Comparative whole-metagenome analysis limited	Comparisons needed for environment and health applications	Use fast assembly and annotation for time-sensitive analyses

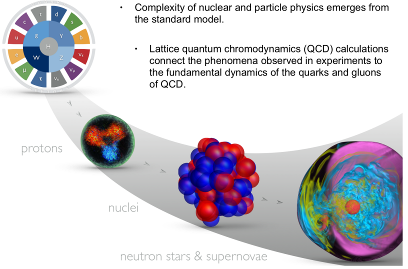
PI: Katherine Yelick (LBNL)

ECP: Scientific Discovery Applications

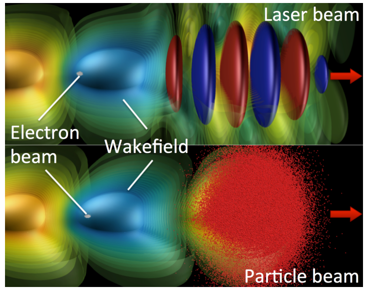
- **Cosmological Probe of the Standard Model of Particle Physics**
  - Elucidate cosmological structure formation by uncovering how smooth and featureless initial conditions evolve under gravity in an expanding universe to eventually form our complex cosmic web. Modern cosmological observations have led to a remarkably successful model for the dynamics of the Universe; 3 key ingredients --- dark energy, dark matter, and inflation --- are signposts to further breakthroughs, as all reach beyond the known boundaries of the particle physics Standard Model



- **Validate Fundamental Laws of Nature**
  - Lattice quantum chromodynamics (QCD) calculations are *the* scientific instrument to connect observed properties of hadrons (particles containing quarks) to fundamental laws of quarks and gluons and critically important to decadal particle and nuclear physics experiments. To elucidate tiny effects of yet-to-be-discovered physics beyond the standard model, particle physics needs QCD simulations accurate to ~0.10% and nuclear physics needs QCD-computed properties and interactions of hadrons and light nuclei on much larger volumes than possible today



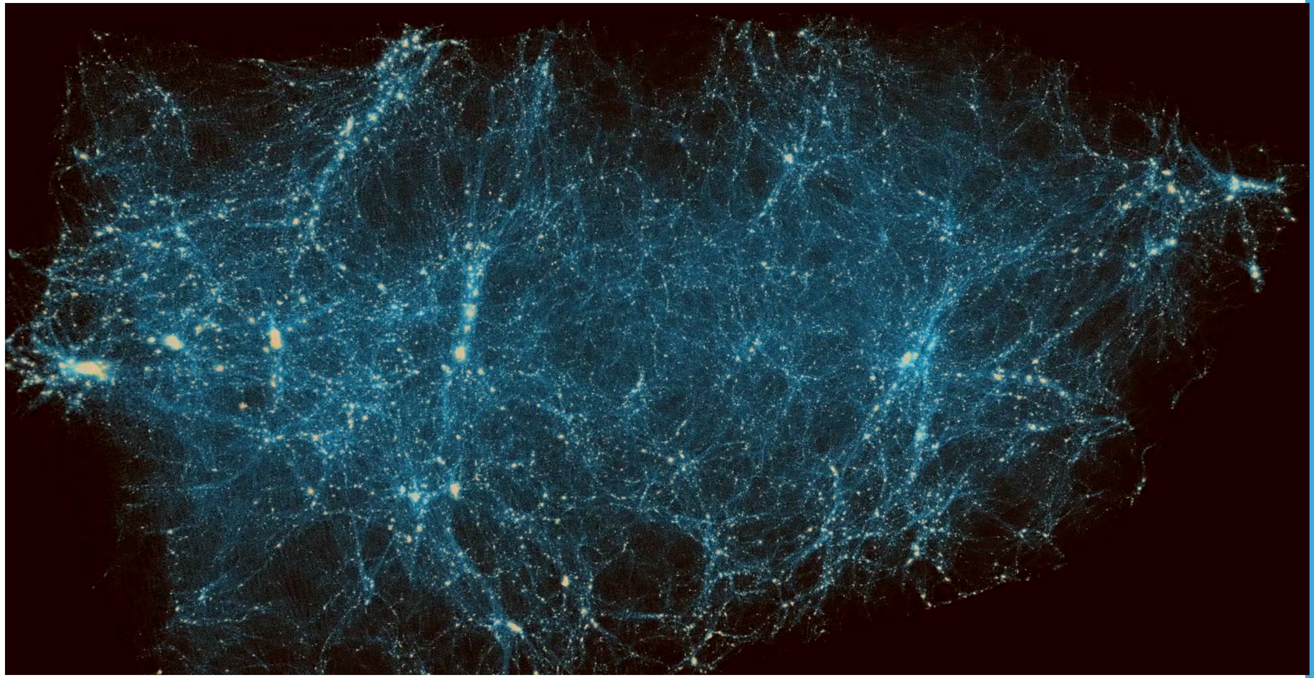
- **Plasma Wakefield Accelerator Design**
  - Particle accelerators are vital in DOE infrastructure for discovery science and university- and private-sector applications, with broad benefits to industry, security, energy, the environment, and medicine. New accelerator designs are needed to reduce size and cost; plasma-based particle accelerators stand apart in their potential for supporting electric fields much larger than current designs



EXASCALE COMPUTING PROJECT

Computing the Sky at Extreme Scales

Elucidating cosmological structure formation by uncovering how smooth and featureless initial conditions evolve under gravity in an expanding universe to eventually form a complex *cosmic web*



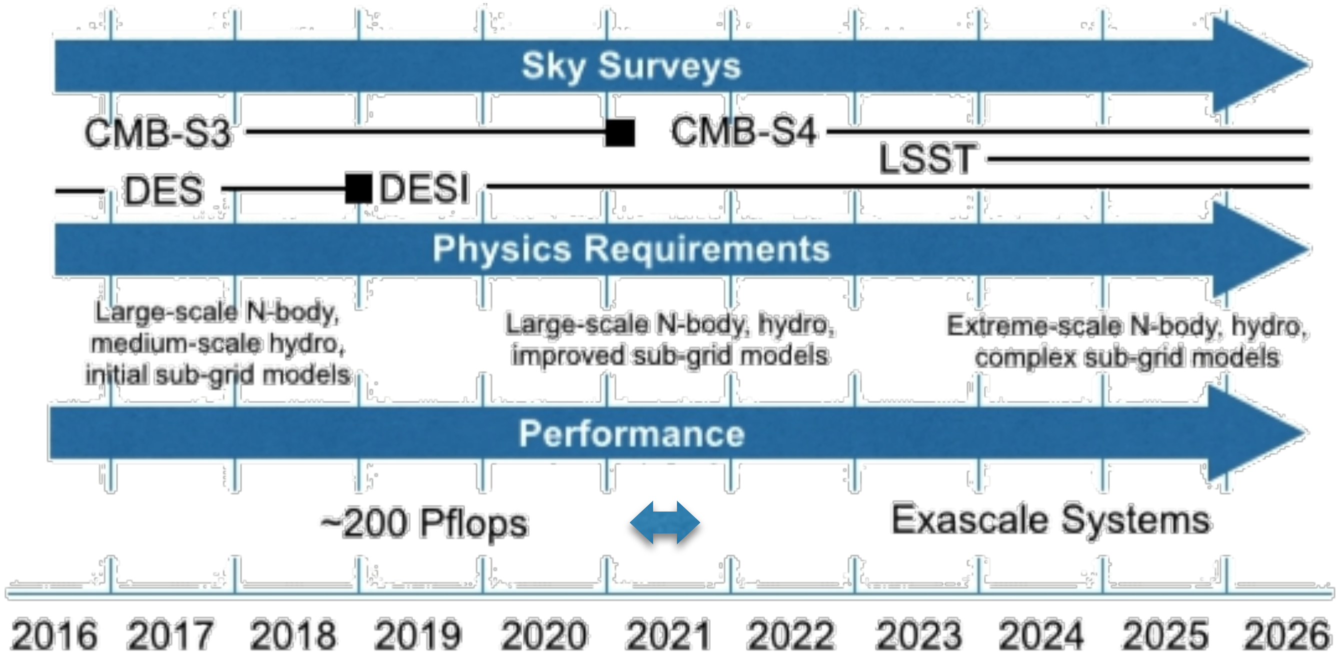
- Modern cosmological observations have led to a remarkably successful model for the dynamics of the Universe; 3 key ingredients --- dark energy, dark matter, and inflation --- are signposts to further breakthroughs, as all reach beyond the known boundaries of the particle physics Standard Model
- A new generation of sky surveys will provide key insights and new measurements such as of neutrino masses
- New discoveries - e.g., primordial gravitational waves and modifications of general relativity - are eagerly awaited
- Capable exascale simulations of cosmic structure formation are essential to shed light on some of the deepest puzzles in all of physical science with a comprehensive program to develop and apply a new extreme-scale cosmology simulation framework for verification of gravitational evolution, gasdynamics, and subgrid models at very high dynamics

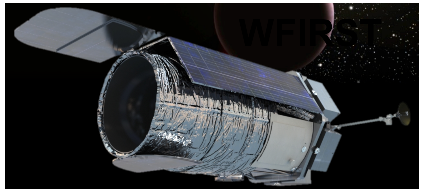
PI: Salman Habib (ANL)

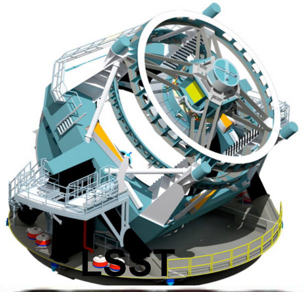
EXASCALE COMPUTING PROJECT

Computing the Sky at Extreme Scales: ExaSky

- **Observations Drive Simulations:** Next-generation cosmological surveys — CMB-S3/4, DES, DESI, eBOSS, Euclid, LSST, WFIRST, **…**
- **Exascale Challenge: Science —** *Dark Energy, Dark Matter, Primordial Gravitational Waves, Neutrino Mass;* **Computation —** factor of **X100** increase in science reach, order of magnitude improvement in modeling accuracy and predictability







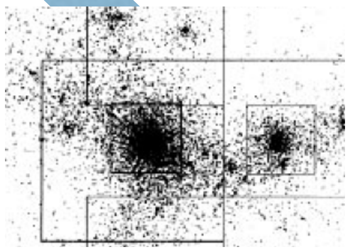
PI: Salman Habib (ANL)

EXASCALE COMPUTING PROJECT

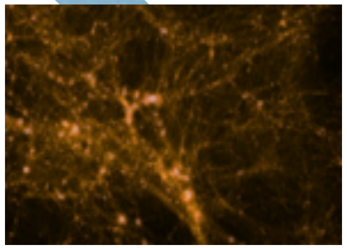
Precision Cosmology: Simulation Frontiers

Fidelity/Complexity

Terascale
- First-generation surveys
- Single-probe simulations



Petascale
- Second-generation surveys
- Multi-probe simulations
- Few precision probes
- Intermediate accuracy parameter estimation



Exascale
- Next-generation surveys
- End-to-end, multi-probe survey-scale simulations
- Multiple cross-calibrated probes
- UQ-enabled cosmic calibration frameworks

Simulation Volume

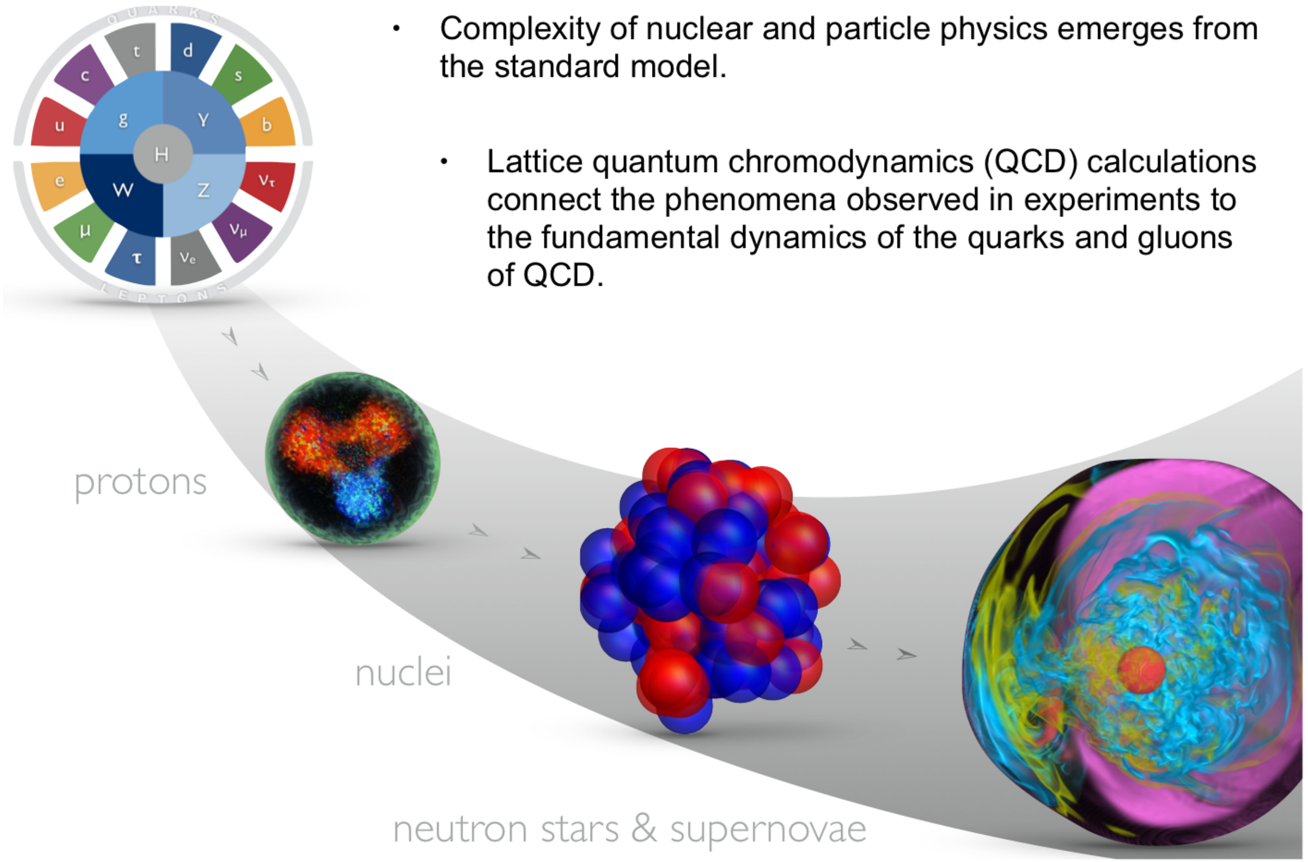
PI: Salman Habib (ANL)

EXASCALE COMPUTING PROJECT

Exascale Lattice Gauge Theory Opportunities and Requirements for Nuclear and High Energy Physics

Lattice quantum chromodynamics (QCD) calculations are the scientific instrument to connect observed properties of hadrons to the fundamental laws of quarks and gluons and critically important to particle and nuclear physics experiments in the decade ahead

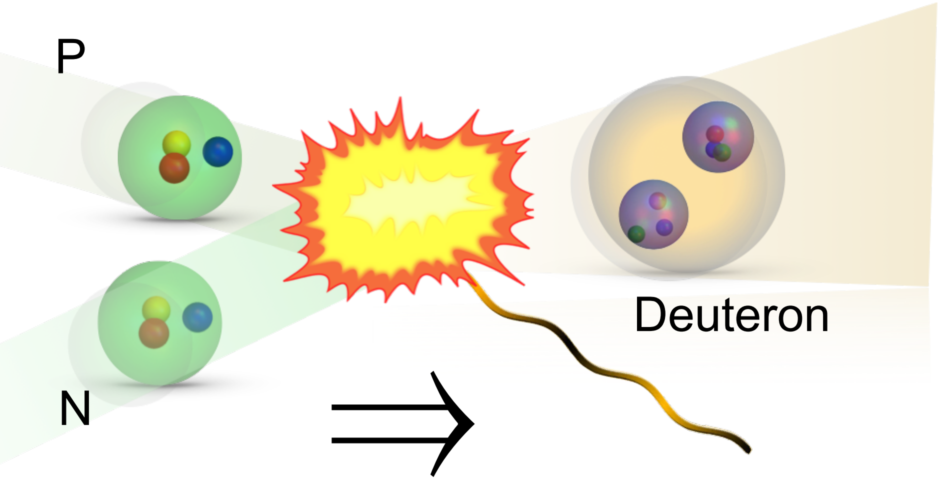
- Complexity of nuclear and particle physics emerges from the standard model.
  - Lattice quantum chromodynamics (QCD) calculations connect the phenomena observed in experiments to the fundamental dynamics of the quarks and gluons of QCD.



- Lattice QCD has made formidable progress in formulating the properties of hadrons (particles containing quarks), but experimental particle and nuclear physics programs require lattice calculations orders of magnitude more demanding still
- Searching for the tiny effects of yet-to-be-discovered physics beyond the standard model, particle physics must have simulations accurate to ~0.10%, an order of magnitude more precise typically realized today
- To accurately compute properties and interactions of hadrons and light nuclei, nuclear physics needs lattice calculations on much larger volumes to investigate multi-hadron states in a reliable controlled way
- Exascale lattice gauge theory will make breakthrough advances possible in particle and nuclear physics

PI: Paul Mackenzie (FNAL)

Simulation Gap	Needed to Address Gap	Impact of Gap Remaining
Nuclei require much larger simulation volumes and have worse signals than mesons. Current computers cannot supply sufficient volumes and light enough quark masses.	A high performance simulation capability for simultaneously reaching the large volumes and light quark masses required for honest nuclear calculations.	An inability to quantitatively understand nuclear physics on the basis of the fundamental laws of physics.



We are on the cusp of being able to calculate the properties of light nuclei such as deuterons, tritons, and alpha particles from first principles.

PI: Paul Mackenzie (FNAL)

EXASCALE COMPUTING PROJECT

Lattice QCD calculations have gone from being essentially crude models to being high-precision tools in the last twenty years. Exascale computing will open up new vistas.

Exascale

Petascale

Terascale

Gigascale

1995 | 2005 | 2015 | 2025

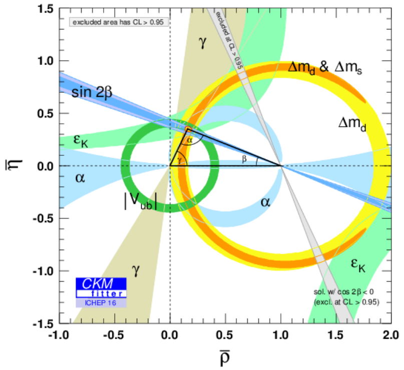
1995: Important effects omitted (quark loops). No realistic uncertainty budgets. Lattice calculations ~ crude models.

2005: Inclusion of light quark loops. Realistic uncertainty budgets possible.

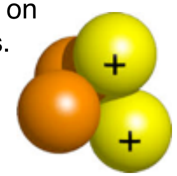
TABLE II: Error budget (in per cent) for $R_{d/s}$, ϕ_s, ϕ_d.

source	$R_{d/s}$	ϕ_s	ϕ_d
statistics	0.5	1.4	1.5
input parameters a and m_c	0.6	2.8	2.9
higher-order $\rho_{A_4^{cq}}$	0	1.3	1.3
heavy-quark discretization	0.5	4.2	4.2
light-quark discretization and χPT fits	5.0	3.9	6.3
static χPT	1.4	0.5	1.5
finite volume	1.4	0.5	1.5
total systematic	5.4	6.5	8.5

2015: Correct light quark masses achieved on medium-size lattices. Few % uncertainties on simple quantities.



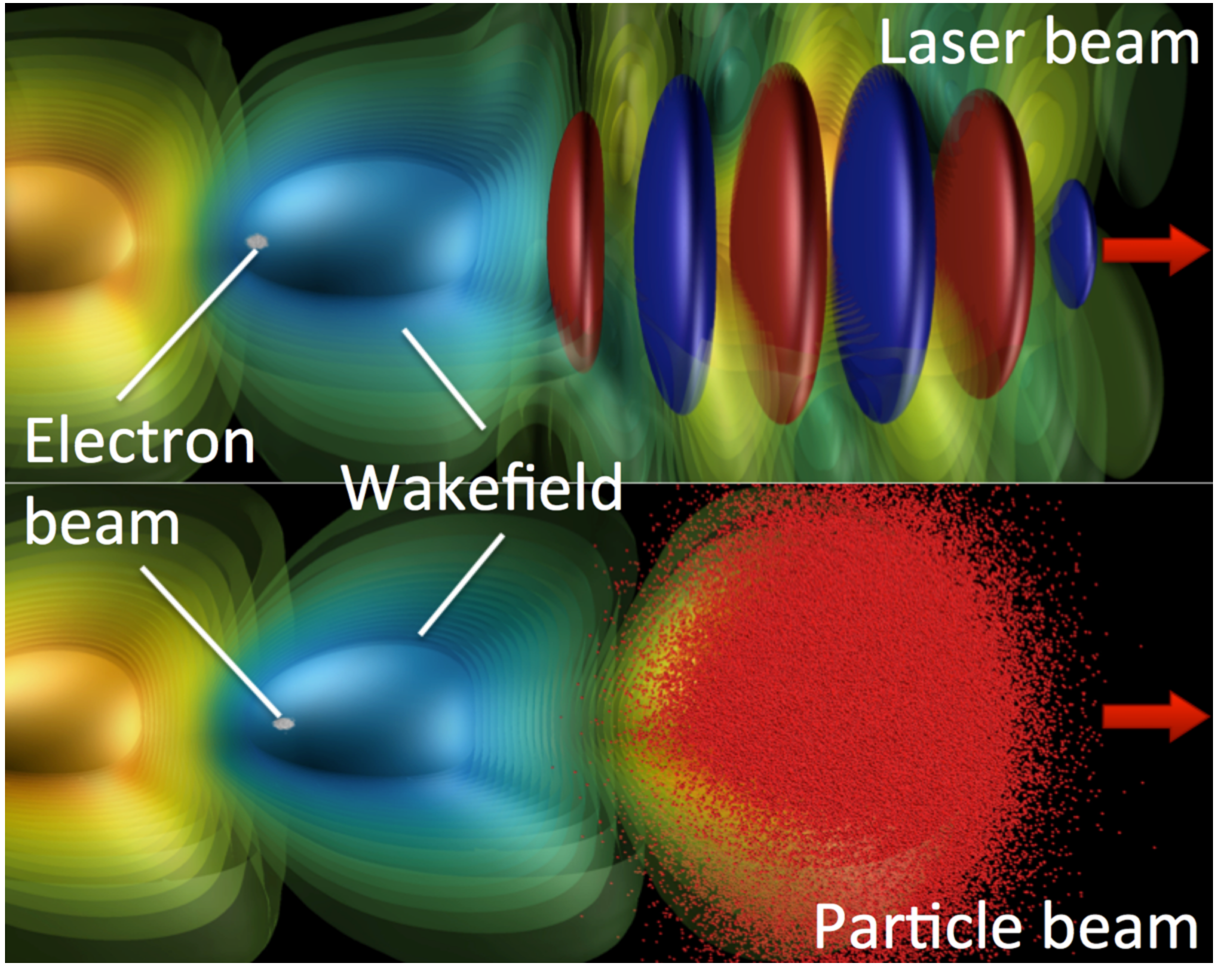
2025: Correct light quark masses large lattices. Properties of light nuclei from first principles. < 1% uncertainties on simple quantities.



PI: Paul Mackenzie (FNAL)

Exascale Modeling of Advanced Particle Accelerators

Toward compact and affordable particle accelerators. A laser beam or a charged particle beam propagating through ionized gas displaces electrons, creates a *wakefield* that supports electric fields orders of magnitude larger than with usual methods, accelerating a charged particle beam to high energy over a very short distance.



- Particle accelerators: a vital part of DOE infrastructure for discovery science and university- and private-sector applications - broad range of benefits to industry, security, energy, the environment, and medicine
- Improved accelerator designs are needed to drive down size and cost; plasma-based particle accelerators stand apart in their potential for these improvements
- Translating this promising technology into a mainstream scientific tool depends critically on exascale-class high-fidelity modeling of the complex processes that develop over a wide range of space and time scales
- Exascale-enabled acceleration design will realize the goal of compact and affordable high-energy physics colliders, with many spinoff plasma accelerator applications likely

PI: Jean-Luc Vay (LBNL)

Particle accelerators are essential tools in modern life that power scientific discovery, cure cancer, secure our borders, and help create a wide range of products

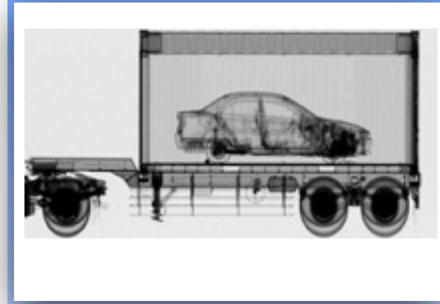
Medicine

- ~9000 worldwide
- 10's of millions of patients treated/yr
- 50 medical isotopes

Industry

- ~20,000 worldwide
- Annual value of all products that use accel. Tech.: $500B

National Security



- Cargo scanning
- Active interrogation
- Stockpile stewardship

Discovery Science

Accel. enabled:
- ~30% Nobel Prizes in Physics since 1939
- 4 of last 14 Nobel Prizes in Chemistry

There are 30,000 Particle Accelerators Making an Impact on Our Lives

Problem: Cost & size is often a limited factor:

- Heidelberg Carbon Therapy Center: ~€119M to build, gantry 670 tons.
- SLAC LCLS-II: ~$1B to build, 5.7 km length.
- CERN LHC: ~$10B to build, 27 km circumference, 150 MW.

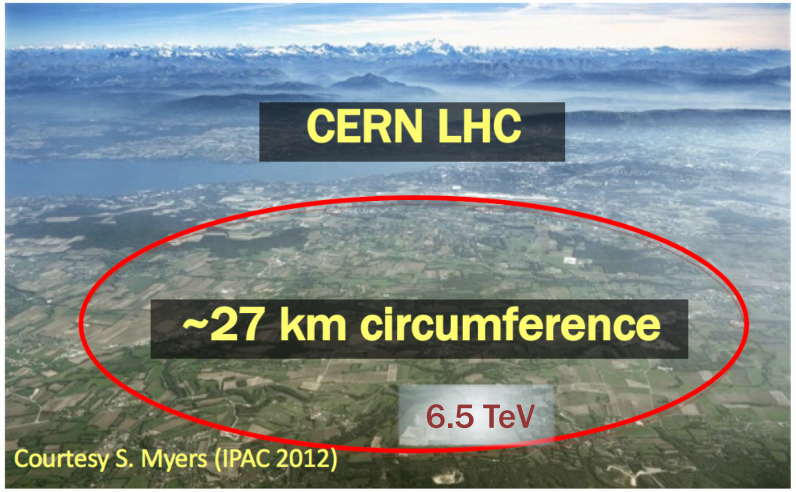
PI: Jean-Luc Vay (LBNL)

Plasma accelerators have the potential to shorten accelerators (and cut cost) dramatically

E. O. Lawrence

First cyclotron ~0.1 MeV

1930



W. Leemans

Plasma accelerator 1 GeV

2009

Let's go from this...

...to this...

9 cm

...and make it ready for application.

Tens of plasma accelerator stages needed for a 1 TeV e^-e^+ collider.

Simulations in 2-D can take days for 1 stage (even at insufficient resolution for collider beam quality).

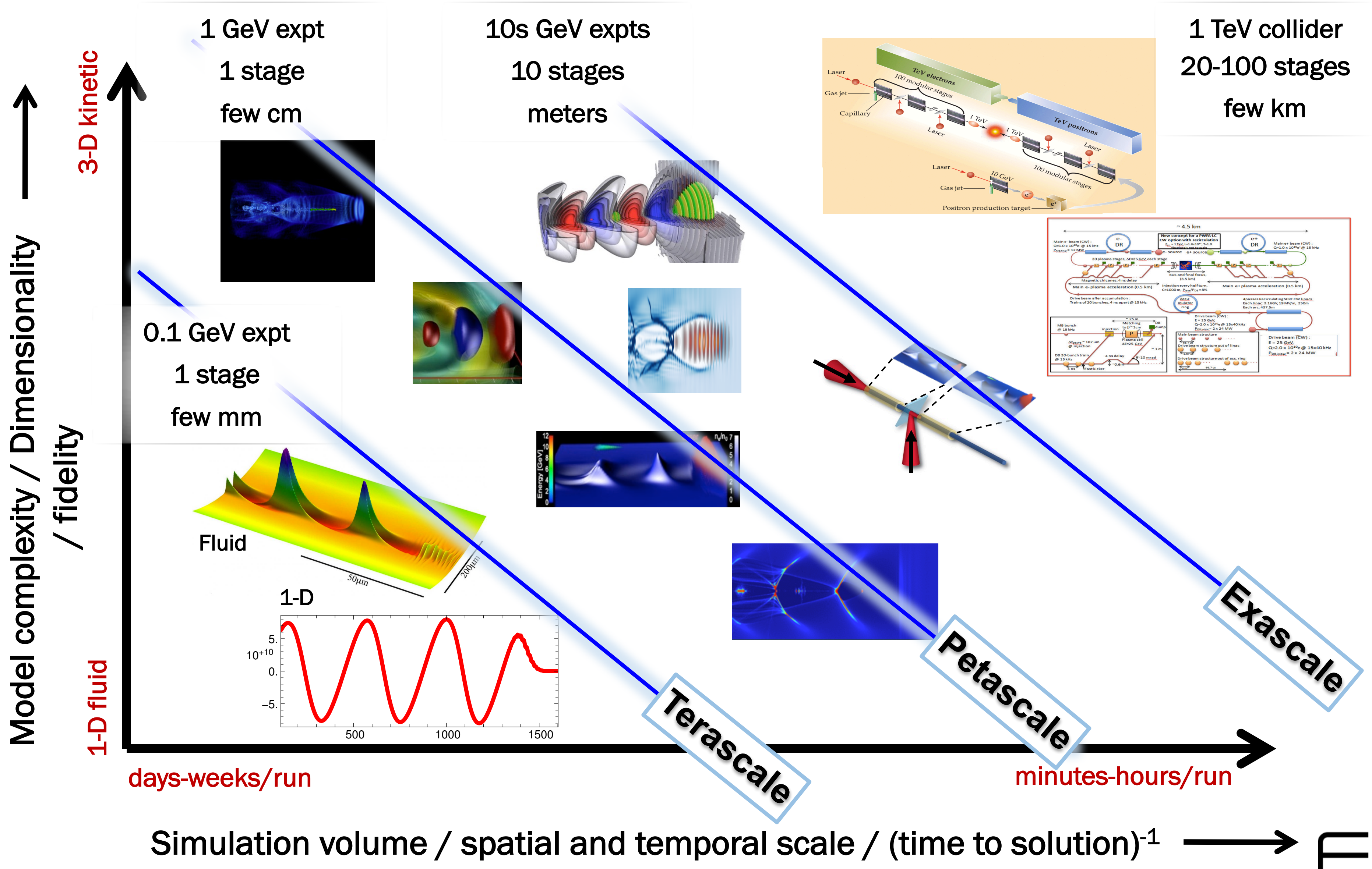
➔ ***Modeling of tens of stages intractable without Exascale computing.***

EXASCALE COMPUTING PROJECT

PI: Jean-Luc Vay (LBNL)

Plasma accelerators

from 2-D cm in days-to-weeks to 3-D m in minutes-to-hours



EXASCALE COMPUTING PROJECT

PI: Jean-Luc Vay (LBNL)

Some Risks and Challenges

- Exploiting on-node memory and compute hierarchies
- Programming models: what to use where and how (e.g., task-based RTS)
- Integrating S/W components that use disparate approaches (e.g., on-node parallelism)
- Developing and integrating co-designed motif-based community components
- Mapping “traditional” HPC applications to current and inbound data hardware
- Infusing data science apps and components into current workflows (e.g., ML for OTF subgrid models)
- Achieving portable performance (without “if-def’ing” 2 different code bases)
- Multi-physics coupling: both algorithms (Picard, JFNK, Anderson Acceleration, HOLO, …) and S/W (e.g., DTK, ADIOS, …); what to use where and how
- Integrating sensitivity analysis, data assimilation, and uncertainty quantification technologies
- Staffing (recruitment & retention)

To exascale
and beyond...

EXASCALE COMPUTING PROJECT

Courtesy Tim Germann, Los Alamos National Laboratory