The U.S. D.O.E. Exascale Computing Project – Goals and Challenges

Paul Messina

Argonne National Laboratory and
Director, Exascale Computing Project

iCAS 2017
Annecy, France
September 13, 2017

EXASCALE COMPUTING PROJECT

National Nuclear Security Administration
U.S. DEPARTMENT OF ENERGY | Office of Science

What is Simulation "Predictivity"?

- Predictivity = accurate prediction (with quantified uncertainty) of the behavior of complex systems
- Exascale moves us from model systems to real systems
  – In some cases
- 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)

First controlled self-sustaining nuclear chain-reaction

December 2, 1942 U. of Chicago, Stagg Field

Enrico Fermi had convinced Arthur Compton that his calculations were reliable enough to rule out a runaway chain reaction or an explosion

The U.S. DOE Exascale Computing Project

- The ECP was established to accelerate delivery of capable exascale computing systems that integrate hardware and software capability to deliver approximately 50 times more performance than today's 20-petaflops machines on mission critical applications
  - DOE is a lead agency for this mission, along with DoD and NSF
    - Deployment agencies: NASA, FBI, NIH, DHS, and NOAA
- Timeline: at least one exascale system will be delivered in 2021 to a DOE Office of Science Leadership Computing Facility (Argonne LCF and/or Oak Ridge LCF)
  - ALCF and OLCF will have diverse architectures
  - A National Nuclear Security Administration (NNSA) facility will field an exascale system in 2022-2023; could be the ALCF or OLCF choice, or a third choice
- ECP's work encompasses
  - Applications
  - System software
  - Hardware technologies and architectures
  - Workforce development to meet scientific and national security mission needs

ECP aims to transform the HPC ecosystem and make major contributions to the nation

Develop applications to tackle a broad spectrum of mission critical problems of unprecedented complexity

Support national security

Partner with vendors to develop computer architectures that support exascale applications

Contribute to the economic competitiveness of the nation

Collaborate with vendors to develop a software stack that is both exascale-capable and usable on industrial and academic scale systems

Train a next-generation workforce of computational scientists, engineers, and computer scientists

I.e., create an exascale ecosystem that will:

- Enable classical simulation and modeling applications to tackle problems that are currently out of reach
- Enable new types applications to utilize exascale systems, including ones that use machine learning, deep learning, and large-scale data analytics
- Support widely used programming models as well as new ones that promise to be more effective on exascale architectures or for applications with new computational patterns, and
- Be suitable for applications that have lower performance requirements currently, thus providing an on ramp to exascale should their future problems require it
-

Key high-level technical challenges that must be tackled to achieve exascale

- *Massive Parallelism* – 100+ times greater than today's largest systems
- *Memory and Storage* – effective use of many levels of hierarchy
  - Memory and storage efficiencies consistent with increased computational rates and data movement requirements
- *Reliability* – system adaptation and recovery from faults in complex system components and designs
- *Energy Consumption* – Energy consumption reduced beyond current industry roadmaps
  - would be prohibitively expensive at this scale
  - hardware and software techniques for minimizing it

What Is a Capable Exascale Computing System?

- Delivers 50× the performance of today's 20 PF systems, supporting applications that deliver high-fidelity solutions in less time and address problems of greater complexity
  - NOTE: no LINPACK or peak FLOPS target
- Operates in a power envelope of 20–30 MW (?)
- Is sufficiently resilient (perceived fault rate: ≤1/week)
- Includes a software stack that supports a broad spectrum of applications and workloads

This ecosystem will be developed using a co-design approach to deliver new software, applications, platforms, and computational science capabilities at heretofore unseen scale

EXASCALE COMPUTING PROJECT

Predictable challenges

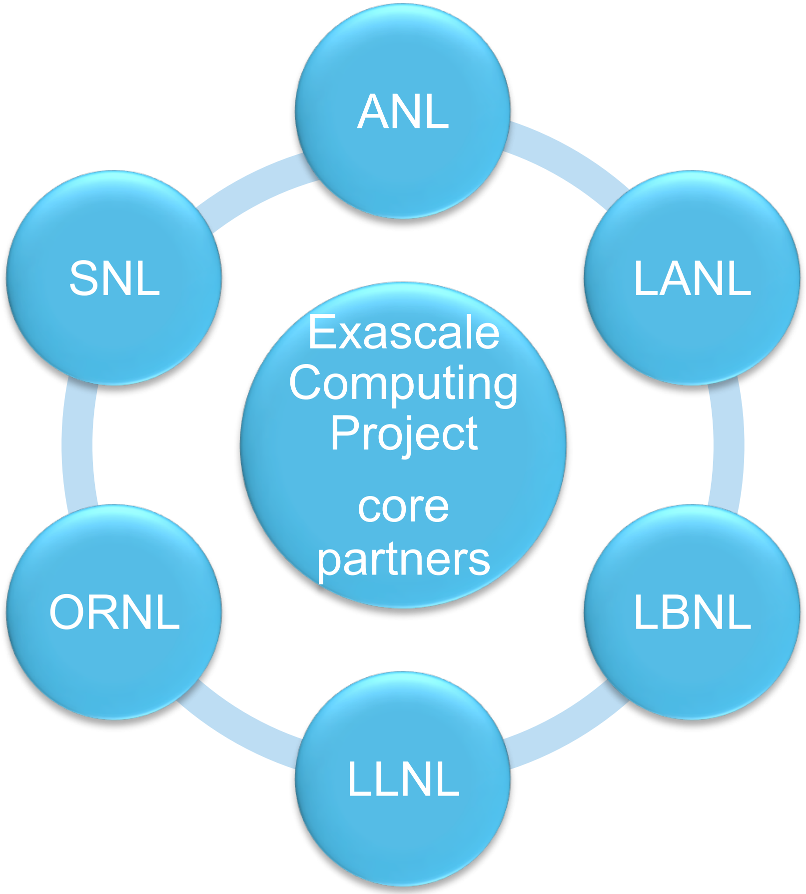
- Node performance: always a major component of overall performance, now even more so
- I/O performance: almost always a bottleneck, now even more so
  - Hardware configuration
  - Software/methods
- Industrial HPC users often rely on commercial software that has not been scaled for use on large systems

Not yet known challenges: Where is that puck going?

- New computational patterns even in traditional simulations
  - UQ, more multi-physics, complex workflows
- New (to HPC) applications
  - Deep Learning, Machine Learning at huge scales
- New hardware architectures
- Squeezing performance (déjà vu)
  - Half precision
  - Single precision
  - Mixed precision

ECP is a Collaboration Among Six US DOE National Laboratories

- The ECP draws from the Nation's 6 premier computing national laboratories
- A Memorandum of Agreement for the ECP was signed by each Laboratory Director defining roles and responsibilities
- Funding comes from two sources: DOE Office of Science and NNSA Advanced Simulation and Computing (ASC) program



But the work is carried out at many institutions

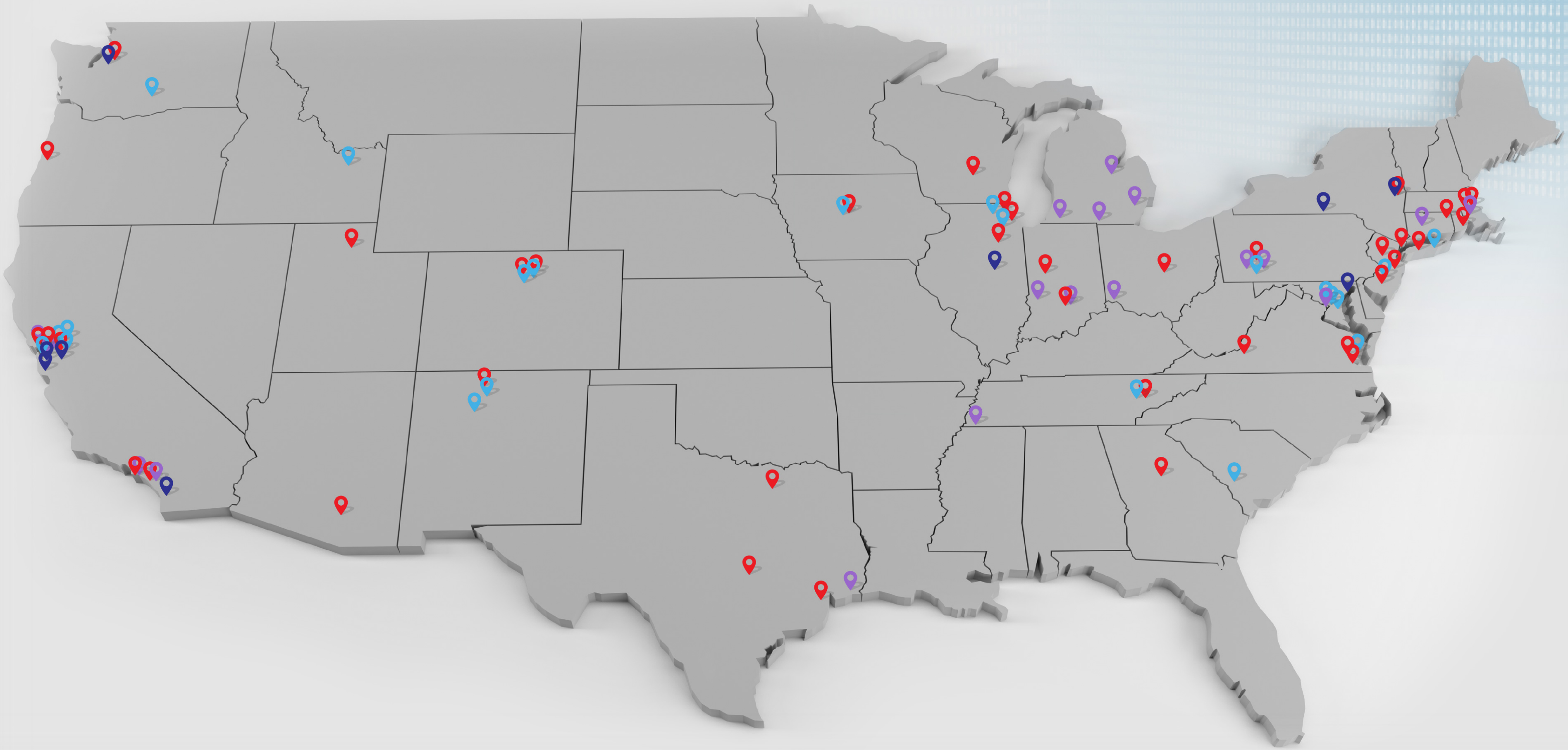
ECP
EXASCALE COMPUTING PROJECT

DOE LABORATORIES & AGENCY PARTNERS
22

PRIVATE SECTOR PARTNERS
10

UNIVERSITY RESEARCH PARTNERS
39

INDUSTRY COUNCIL MEMBERS
18



THE ECP ECOSYSTEM

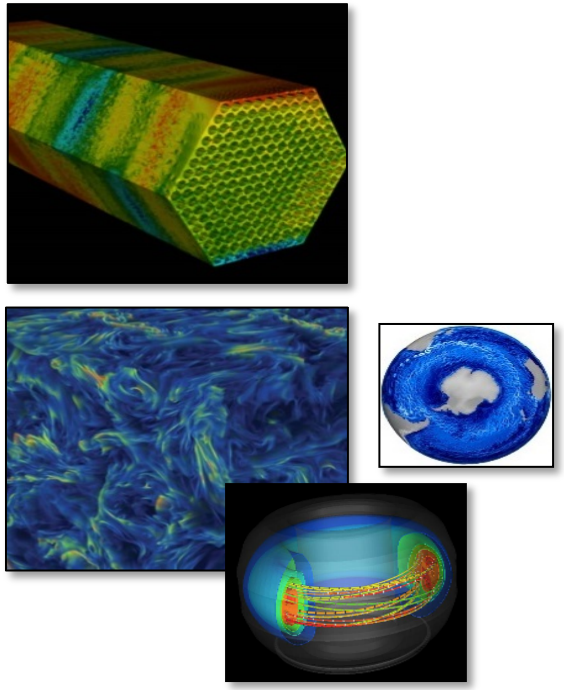
- 800+ Researchers
- 25 Application Development Projects
- 66 Software Development Projects
- 5 Co-Design Centers

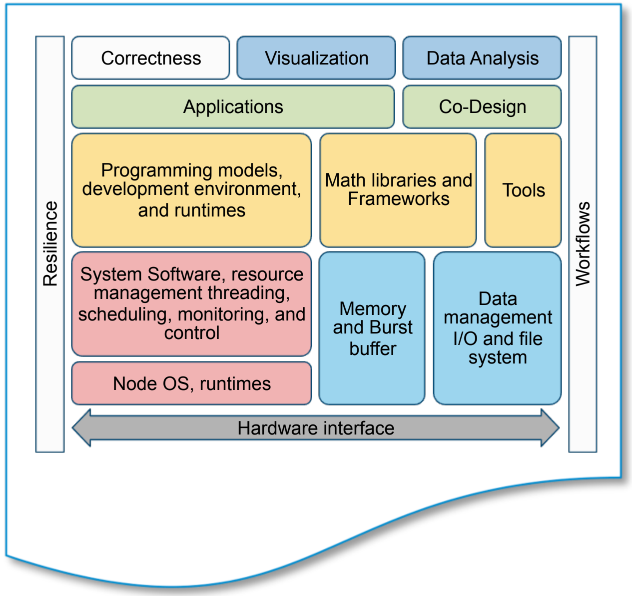
U.S. DEPARTMENT OF ENERGY | Office of Science

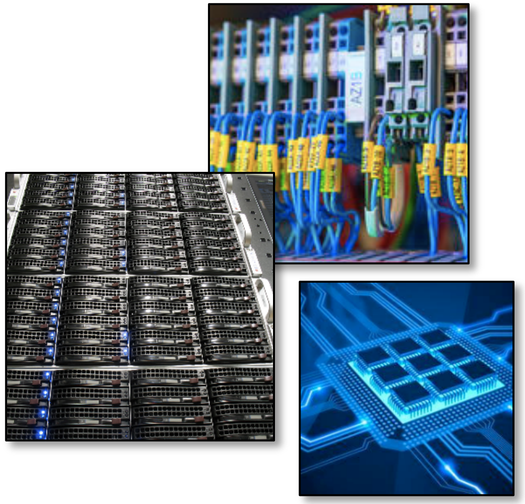
NNSA National Nuclear Security Administration

ECP uses co-design and integration to achieve exascale computing

Application Development	Software Technology	Hardware Technology
Science and mission applications	Scalable software stack	Hardware technology elements







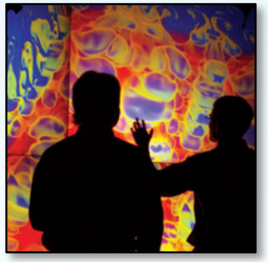
EXASCALE COMPUTING PROJECT

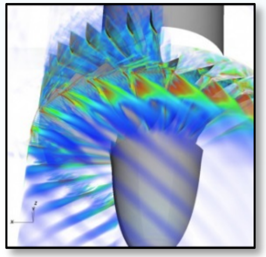
Exascale applications deliver broad coverage of key strategic pillars

National security

Stockpile stewardship

Next-generation electromagnetics simulation of hostile environment and virtual flight testing for hypersonic re-entry vehicles





Energy security

Turbine wind plant efficiency

Design and commercialization of SMRs

Nuclear fission and fusion reactor materials design

Subsurface use for carbon capture, petroleum extraction, waste disposal

High-efficiency, low-emission combustion engine and gas turbine design

Carbon capture and sequestration scaleup

Biofuel catalyst design

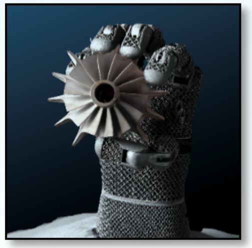
Economic security

Additive manufacturing of qualifiable metal parts

Urban planning

Reliable and efficient planning of the power grid

Seismic hazard risk assessment



Scientific discovery

Cosmological probe of the standard model of particle physics

Validate fundamental laws of nature

Plasma wakefield accelerator design

Light source-enabled analysis of protein and molecular structure and design

Find, predict, and control materials and properties

Predict and control stable ITER operational performance

Demystify origin of chemical elements

Earth system

Accurate regional impact assessments in Earth system models

Stress-resistant crop analysis and catalytic conversion of biomass-derived alcohols

Metagenomics for analysis of biogeochemical cycles, climate change, environmental remediation

Health care

Accelerate and translate cancer research

ECP EXASCALE COMPUTING PROJECT

Application Metrics

1. **Deliver improved and impactful science & engineering (*performance*)**
   – New or improved (ideally step change in) predictability on a problem of national importance (a "challenge problem")
2. **As performance portable as possible and reasonable (*portability*)**
   – No "boutique" one-off applications able to only execute on one (and likely ephemeral) system
3. **Able to make effective use of a capable system (*readiness*)**
   – *Effective* is app specific (weak, strong, ensembles, single-node performance)
4. **Able to integrate latest relevant software technologies (*modern*)**
   – Needed to demonstrate agility, flexibility, modern architecture; overall app portfolio must apply pressure to all key attributes of the system design characteristics
5. **High priority (*strategic*)**
   – Some key stakeholder somewhere really cares about using application to make consequential decisions

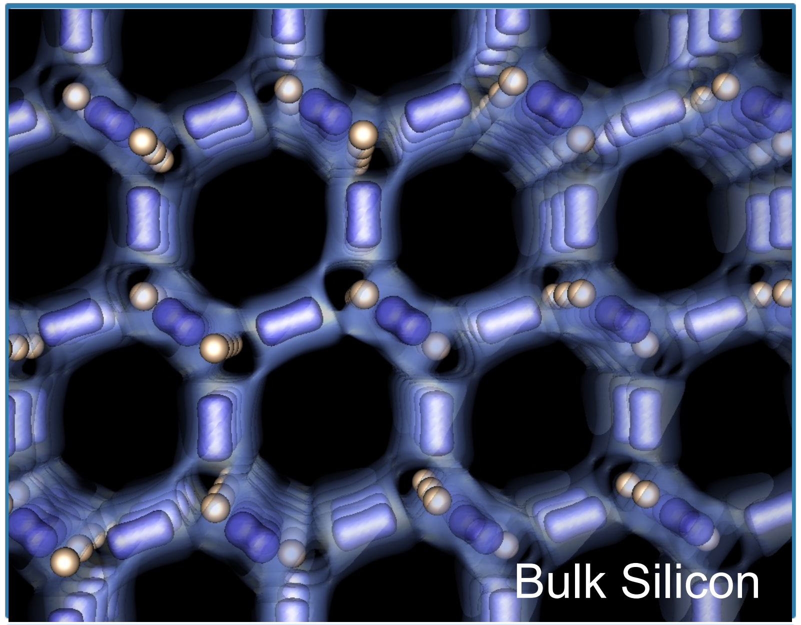
QMCPACK: Enabling predictive simulation of materials

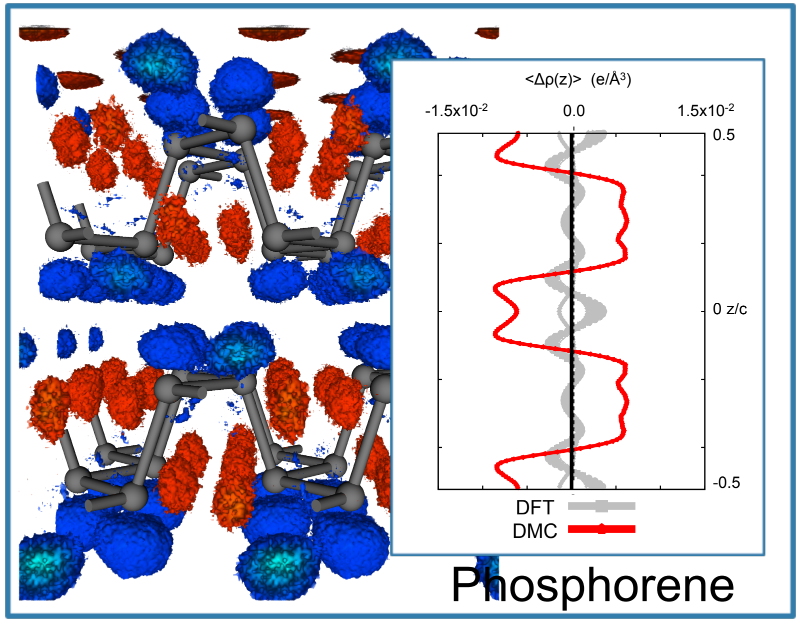
Project Goal

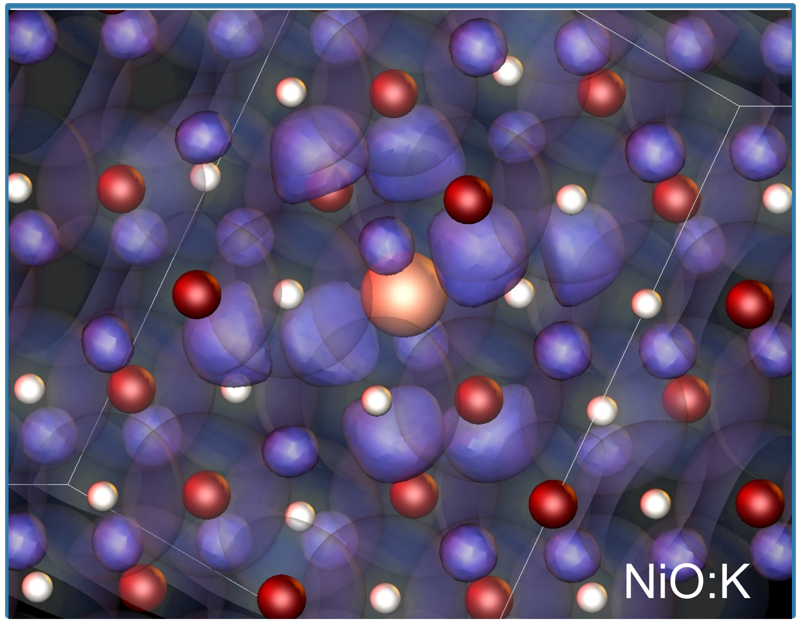
"Deliver a scientifically performant, performance portable, extensible and maintainable open source exascale Quantum Monte Carlo code, **and work within the ECP to help develop the necessary exascale software and hardware ecosystem**."

We will evolve the existing QMCPACK code, not rewrite it from scratch.

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

QMCPACK Challenge problem

- 4 year target problem is an easily specified proxy for a complex transition metal oxide, also an initial CMS material:

"Calculate the cohesive energy of a 1024-atom supercell of NiO to an accuracy of 10 meV per NiO formula unit using a full LCF class machine in a reasonable and scientifically productive amount of wall clock time, ~1 day."

- QMC is ~cubic in the electron count. Here 12288 valence electrons.
- Barely runnable by QMCPACK today due to memory requirements.
- Can readily trade system size, element types, statistical error for general applications.

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

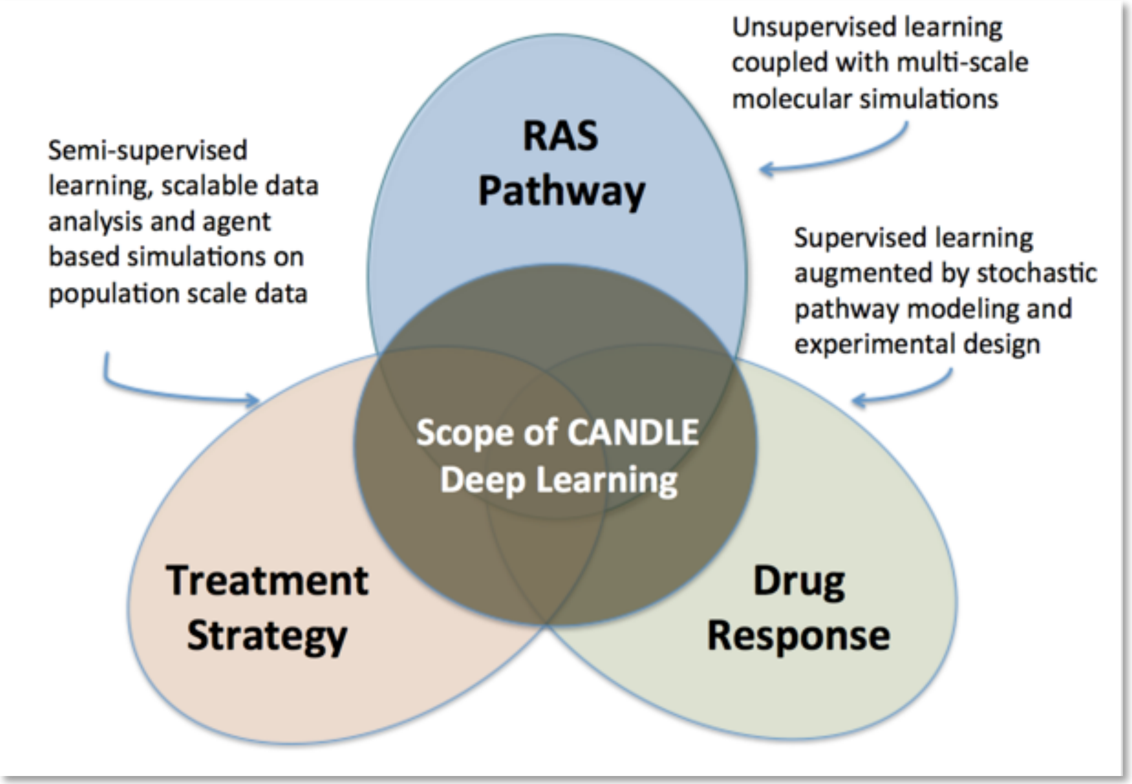
* ACME's new name: Energy Exascale Earth System Model (E3SM)

EXASCALE COMPUTING PROJECT

ACME-MMF Overview by Institution

- SNL: M. Taylor, E. Foster
  - Project leadership
  - Software engineering (testing, code integration), Nonhydrostatic comparisons
- ANL: J. Krishna, D. Wu
  - Parallel I/O performance improvements
- LANL: P. Jones, R. Aulwes, R. Green
  - GPU porting of MPAS components
- LLNL: D. Bader, W. Hannah
  - ACME-MMF development & evaluation, convection
- ORNL: M. Norman, S. Sreepathi
  - GPU porting of CRM
- PNNL: R. Leung, M. Ovchinnikov, C. Jones
  - MAML, ACME-MMF Evaluation
- University subcontracts:
  - CSU: : D. Randall, D. Dazlich, M. Branson
  - UC Irving: M. Pritchard, H. Parishani

ECP-CANDLE : CANcer Distributed Learning Environment



CANDLE Goals

1 Develop an Exscale deep learning environment for Cancer

2 Build on open source deep learning frameworks

3 Optimize for CORAL and Exascale platforms

4 Support all three pilot project needs for deep learning

5 Collaborate with DOE computing centers, HPC vendors, and ECP co-design and software technology projects

PI: Rick Stevens (ANL)

CANDLE's High-Level Development Plan

- Create computational infrastructure to enable large-scale runs of deep learning on leadership computers
  – Training Data Management, Orchestration of Large-Scale Ensembles, Large-scale Model Database, Training Visualization and Monitoring
  – Infrastructure should be deep learning framework agnostic
- Work with each Pilot project to construct deep learning models for key problems
  – Benchmarks, Training Data, Exploratory Models, New DL methods
- Conduct ever increasingly complex series of deep learning studies on leadership machines to refine approach for Exascale
- Work with vendors and framework providers to optimize for Exascale Platforms

PI: Rick Stevens (ANL)

Survey of Application Motifs

Application	Monte Carlo	Particles	Sparse Linear Algebra	Dense Linear Algebra	Spectral Methods	Unstructured Grid	Structured Grid	Comb. Logic	Graph Traversal	Dynamical Program	Backtrack & Branch and Bound	Graphical Models	Finite State Machine
Cosmology													
Subsurface													
Materials (QMC)													
Additive Manufacturing													
Chemistry for Catalysts & Plants													
Climate Science													
Precision Medicine Machine Learning													
QCD for Standard Model Validation													
Accelerator Physics													
Nuclear Binding and Heavy Elements													
MD for Materials Discovery & Design													
Magnetically Confined Fusion													

Survey of Application Motifs

Application	Monte Carlo	Particles	Sparse Linear Algebra	Dense Linear Algebra	Spectral Methods	Unstructured Grid	Structured Grid	Comb. Logic	Graph Traversal	Dynamical Program	Backtrack & Branch and Bound	Graphical Models	Finite State Machine
Combustion S&T		■	■	■			■						
Free Electron Laser Data Analytics	■	■	■	■	■		■			■	■	■	
Microbiome Analysis			■					■	■	■		■	
Catalyst Design	■	■	■	■	■								
Wind Plant Flow Physics			■	■		■							
SMR Core Physics	■	■	■	■	■	■	■						
Next-Gen Engine Design		■	■		■	■							
Urban Systems			■		■	■	■						
Seismic Hazard Assessment			■		■	■	■			■			
Systems Biology	■		■	■				■	■	■	■	■	
Biological Neutron Science		■	■	■	■		■		■		■		
Power Grid Dynamics	■		■	■					■	■	■		

EXASCALE COMPUTING PROJECT

Survey of Application Motifs

Application	Monte Carlo	Particles	Sparse Linear Algebra	Dense Linear Algebra	Spectral Methods	Unstructured Grid	Structured Grid	Comb. Logic	Graph Traversal	Dynamical Program	Backtrack & Branch and Bound	Graphical Models	Finite State Machine
Stellar Explosions	■	■	■	■			■						
Excited State Material Properties				■	■								
Light Sources			■	■	■	■				■			■
Materials for Energy Conversion/Storage		■	■	■	■								
Hypersonic Vehicle Design	■	■	■	■		■							
Multiphase Energy Conversion Devices		■	■			■							

ECP Co-Design Centers

- **A Co-Design Center for Online Data Analysis and Reduction at the Exascale (CODAR)**
  – **Motifs**: Online data analysis and reduction
  – Address growing disparity between simulation speeds and I/O rates rendering it infeasible for HPC and data analytic applications to perform offline analysis. Target common data analysis and reduction methods (e.g., feature and outlier detection, compression) and methods specific to particular data types and domains (e.g., particles, FEM)

- **Block-Structured AMR Co-Design Center (AMReX)**
  – **Motifs**: Structured Mesh, Block-Structured AMR, Particles
  – New block-structured AMR framework (AMReX) for systems of nonlinear PDEs, providing basis for temporal and spatial discretization strategy for DOE applications. Unified infrastructure to effectively utilize exascale and reduce computational cost and memory footprint while preserving local descriptions of physical processes in complex multi-physics algorithms

- **Center for Efficient Exascale Discretizations (CEED)**
  – **Motifs**: Unstructured Mesh, Spectral Methods, Finite Element (FE) Methods
  – Develop FE discretization libraries to enable unstructured PDE-based applications to take full advantage of exascale resources without the need to “reinvent the wheel” of complicated FE machinery on coming exascale hardware

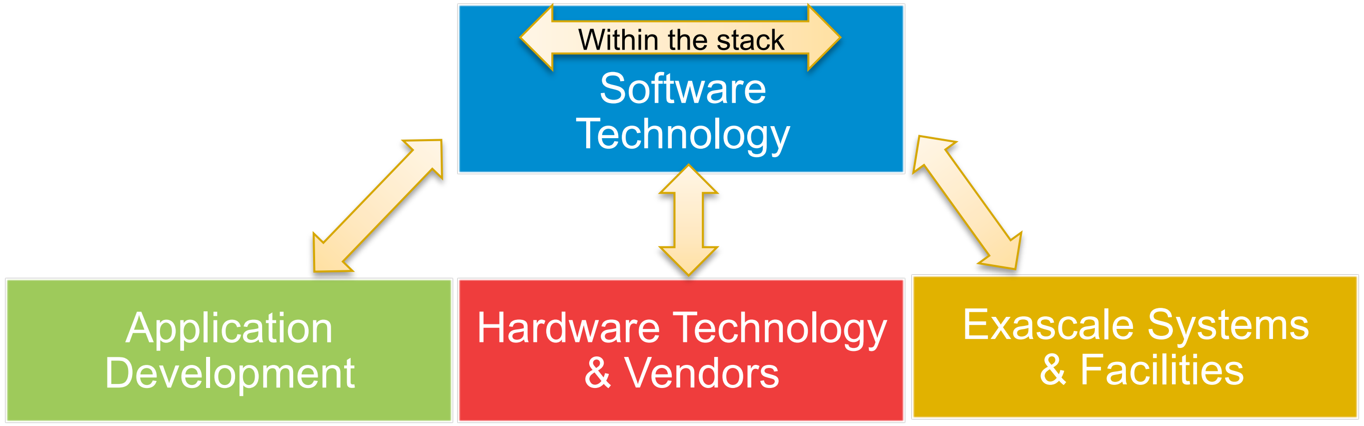
- **Co-Design Center for Particle Applications (CoPA)**
  – **Motif(s)**: Particles (involving particle-particle and particle-mesh interactions)
  – Focus on four sub-motifs: short-range particle-particle (e.g., MD and SPH), long-range particle-particle (e.g., electrostatic and gravitational), particle-in-cell (PIC), and additional sparse matrix and graph operations of linear-scaling quantum MD

- **Combinatorial Methods for Enabling Exascale Applications (ExaGraph)**
  – **Motif(s)**: Graph traversals; graph matching; graph coloring; graph clustering, including clique enumeration, parallel branch-and-bound, graph partitioning
  – Develop methods and techniques for efficient implementation of key combinatorial (graph) algorithms that play a critical enabling role in numerous scientific applications. The irregular memory access nature of these algorithms makes them difficult algorithmic kernels to implement on parallel systems

Additional Application Development activities

- Exascale Proxy Applications Suite
  - Assemble and curate a proxy app suite composed of proxies developed by other ECP projects that represent the most important features (especially performance) of exascale applications.
- Application Assessment
- IDEAS-ECP – Advancing Software Productivity for Exascale Applications
  - Customize and curate methodologies for ECP app productivity & sustainability
  - Create an ECP Application Development Kit of customizable resources for improving scientific software development
- Training
  - Argonne Training Program on Extreme Scale Computing, others

ECP Requires Strong Integration to Achieve Capable Exascale



- To achieve a coherent software stack, we must integrate across all the focus areas
  - Understand and respond to the requirements from the apps but also help them understand challenges they may not yet be aware of
  - Understand and repond to the impact of hardware technologies and platform characteristics
  - Work with the facilities and vendors towards a successful stable deployment of our software technologies
  - Understand and respond to dependencies within the stack, avoiding duplication and scope creep
  - This is a comprehensive team effort — not a set of individual projects!

Software Technology Portfolio

Resilience
Checkpoint-restart, fault injection & modeling

Correctness

Visualization
In situ vis infrastructure and end-user tools

Data Analysis
In-situ data analysis

Applications

Co-Design

Programming Models, Development Environment, and Runtimes
MPI, OpenMP, OpenACC, PGAS, task-based models, RAJA, Kokkos, power steering library

Math Libraries/Frameworks
Dense and sparse linear algebra, nonlinear solvers, discretization, frameworks, etc.

Tools
Performance portability, performance measurement & analysis, debugging, memory, I/O characterization

System Software, Resource Management Threading, Scheduling, Monitoring, and Control
Global resource management and resource montoring, threading library, software stack and installation support

Node OS, low-level runtimes
On-node resource management and runtimes

Memory and Burst buffer
User-level file system for burst buffers, simplified interface for complex memory

Data Management, I/O and File System
I/O libraries, compression, checkpoint-restart, I/O services

Workflows
Federated infrastructure for simulation mgmt.

Hardware interface

The ECP Software Technology Portfolio

Derived from

- Analysis of the software needs of exascale applications
- Inventory of software environments at major DOE HPC facilities (ALCF, OLCF, NERSC, LLNL, LANL, SNL)
  - For current systems and the next acquisition (CORAL, APEX)
- Expected software environment for an exascale system
- Requirements beyond the software environment provided by vendors of HPC systems

Cloud-Resolving Climate Modeling of the Earth's Water Cycle*

Exascale Challenge Problem

- Cloud-resolving (~1 km) Earth system model with throughput necessary for multi-decadal coupled high resolution climate simulations, reducing major systematic errors in precipitation models via explicit treatment of convective storms
- Improve regional impact assessments of climate change on the water cycle, e.g., influencing agriculture and energy production
- Integrate a cloud-resolving GPU-enabled convective parameterization (superparameterization) into the ACME Earth System model using the Multiscale Modeling Framework; refactor key ACME model components for GPU systems
- ACME Earth system model goal: fully weather resolving atmosphere and cloud-resolving superparameterization, eddy resolving ocean and ice components, with throughput (5 SYPD) enabling 10-100 member ensembles of 100 year simulations

Applications & S/W Technologies

Applications

- ACME Earth system model: ACME-Atmosphere; MPAS (Model Prediction Across Scales)-Ocean (ocean); MPAS-Seaice (sea ice); MPAS-Landice (land ice), SAM (System for Atmospheric Modeling)

Software Technologies Cited

- Fortran, C++
- MPI, OpenMP, OpenACC
- Kokkos, Legion
- PIO, Trilinos, PETSc
- ESGF, Globus Online, AKUNA framework

Risks and Challenges

- Inability to obtain sufficient LCF allocations
- Obtaining necessary GPU throughput on the cloud-resolving model
- Cloud-resolving convective parameterization via the multi-scale modeling framework does not provide expected improvements in water cycle simulation quality
- Global atmospheric model cannot obtain necessary throughput
- MPAS ocean/ice components not amenable to GPU acceleration

Development Plan

Y1: Demonstrate ACME-MMF model for “AMIP” (Atmospheric Model Intercomparison Project) configuration. Complete 5 year ACME-MMF simulation with active atmosphere and land components at low resolution and ACME atmosphere diagnostics/metrics

Y2: Demonstrate ACME-MMF model with active atmosphere, land, ocean and ice. Complete 40 year simulation with ACME coupled group water cycle diagnostics/metrics

Y3: Document GPU speedup in performance critical components: Atmosphere, Ocean and Ice. Compare SYPD with and without using the GPU.

Y4: ACME-MMF configuration integrated ACME model; Document highest resolution able to deliver 5 SYPD; Complete 3 member ensemble of 40 year simulations with all active components (atmosphere, ocean, land, ice) with ACME coupled group diagnostics/metrics

*PI: Mark Taylor (SNL)

Example: An Exascale Subsurface Simulator of Coupled Flow, Transport, Reactions and Mechanics*

Exascale Challenge Problem

- 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 of thousands deep wells that penetrate the subsurface for resource utilization
- Resolve pore-scale (0.1-10 μm) physical and geochemical heterogeneities in wellbores and fractures to predict evolution of these features when subjected to geomechanical and geochemical stressors
- Integrate multi-scale (μm to km), multi-physics in a reservoir simulator: non-isothermal multiphase fluid flow and reactive transport, chemical and mechanical effects on formation properties, induced seismicity and reservoir performance
- Century-long simulation of a field of wellbores and their interaction in the reservoir

Applications & S/W Technologies

Applications

- Chombo-Crunch, GEOS

Software Technologies Cited

- C++, Fortran, LLVM/Clang
- MPI, OpenMP, CUDA
- Raja, CHAI
- Chombo AMR, PETSc
- ADIOS, HDF5, Silo, ASCTK
- VisIt

Risks and Challenges

- Porting to exascale results in suboptimal usage across platforms
- No file abstraction API that can meet coupling requirements
- Batch scripting interface incapable of expressing simulation workflow semantics
- **Scalable AMG solver in PETSc**
- Physics coupling stability issues
- Fully overlapping coupling approach results inefficient.

Development Plan

Y1: Evolve GEOS and Chombo-Crunch; Coupling framework v1.0; Large scale (100 m) mechanics test (GEOS); Fine scale (1 cm) reactive transport test (Chombo-Crunch)

Y2: GEOS+Chombo-Crunch coupling for single phase; Coupling framework w/ physics; Multiphase flow for Darcy & pore scale; GEOS large strain deformation conveyed to Chombo-Crunch surfaces; Chombo-Crunch precip/dissolution conveyed to GEOS surfaces

Y3: Full demo of fracture asperity evolution-coupled flow, chemistry, and mechanics

Y4: Full demo of km-scale wellbore problem with reactive flow and geomechanical deformation, from pore scale to resolve the geomechanical and geochemical modifications to the thin interface between cement and subsurface materials in the wellbore and to asperities in fractures and fracture networks

*PI: Carl Steefel (LBNL)

ECP EXASCALE COMPUTING PROJECT

Example: NWChemEx: Tackling Chemical, Materials and Biomolecular Challenges in the Exascale Era*

Exascale Challenge Problem

- Aid & accelerate advanced biofuel development by exploring new feedstock for efficient production of biomass for fuels and new catalysts for efficient conversion of biomass derived intermediates into biofuels and bioproducts
- Molecular understanding of how proton transfer controls protein-assisted transport of ions across biomass cellular membranes; often seen as a stress responses in biomass, would lead to more stress-resistant crops thru genetic modifications
- Molecular-level prediction of the chemical processes driving the specific, selective, low-temperature catalytic conversion (e.g., Zeolites such as H-ZSM-5)) of biomass-derived alcohols into fuels and chemicals in constrained environments

Applications & S/W Technologies

Applications

- NWChemEx (evolved from redesigned NWChem)

Software Technologies Cited

- Fortran, C, C++
- Global arrays, TiledArrays, ParSEC, TASCEL
- VisIt, Swift
- TAO, Libint
- Git, svn, JIRA, Travis CI
- Co-Design: CODAR, CE-PSI, GraphEx

Risks and Challenges

- Unknown performance of parallel tools
- Insufficient performance or scalability or large local memory requirements of critical algorithms
- Unavailable tools for hierarchical memory, I/O, and resource management at exascale
- Unknown exascale architectures
- Unknown types of correlation effect for systems with large number of electrons
- Framework cannot support effective development

Development Plan

Y1: Framework with tensor DSL, RTS, APIs, execution state tracking; Operator-level NK-based CCSD with flexible data distributions & symmetry/sparsity exploitation

Y2: Automated compute of CC energies & 1-/2-body CCSD density matrices; HT & DFT compute of >1K atom systems via multi-threading

Y3: Couple embedding with HF & DFT for multilevel memory hierarchies; QMD using HF & DFT for 10K atoms; Scalable R12/F12 for 500 atoms with CCSD energies and gradients using task-based scheduling

Y4: Optimized data distribution & multithreaded implementations for most time-intensive routines in HF, DFT, and CC.

EXASCALE COMPUTING PROJECT

*PI: Thom Dunning (PNNL)

Software Technologies
Aggregate of technologies cited in all candidate ECP Applications

- **Programming Models and Runtimes**
  - Fortran, C++/C++17, Python, C, Javascript, C#, R, Ruby
  - MPI, OpenMP, OpenACC, CUDA, Global Arrays, TiledArrays, Argobots, HPX, OpenCL, Charm++
  - UPC/UPC++, Co-Array FORTRAN, CHAPEL, Julia, GDDI, DASK-Parallel, PYBIND11
  - PGAS, GASNetEX, Kokkos, Raja, Legion/Regent, OpenShmem, Thrust
  - PARSEC, Panda, Sycl, Perilla, Globus Online, ZeroMQ, ParSEC, TASCEL, Boost
- **Tools** (debuggers, profilers, software development, compilers)
  - LLVM/Clang,HPCToolkit, PAPI, ROSE, Oxbow (performance analysis), JIRA (software development tool), Travis (testing),
  - ASPEN (machine modeling), CMake, git, TAU, Caliper, , GitLab, CDash (testing), Flux, Spack, Docker, Shifter, ESGF, Gerrit
  - GDB, Valgrind, GitHub, Jenkins (testing), DDT (debugger)
- **Mathematical Libraries, Scientific Libraries, Frameworks**
  - BLAS/PBLAS, MOAB, Trilios, PETSc, BoxLib, LAPACK/ScaLAPACK, Hypre, Chombo, SAMRAI, Metis/ParMETIS, SLEPc
  - SuperLU, Repast HPC (agent-based model toolkit), APOSMM (optimization solver), HPGMG (multigrid), FFTW, Dakota, Zero-RK
  - cuDNN, DAAL, P3DFFT, QUDA (QCD on GPUs), QPhiX (QCD on Phi), ArPack (Arnoldi), ADLB, DMEM, MKL, Sundials, Muelu
  - DPLASMA, MAGMA,PEBBL, pbdR, FMM, DASHMM, Chaco (partitioning), libint (gaussian integrals)
  - Smith-Waterman, NumPy, libcchem

Software Technologies
Cited in Candidate ECP Applications

- **Data Management and Workflows**
  - Swift, MPI-IO, HDF, ADIOS, XTC (extended tag container), Decaf, PDACS, GridPro (meshing), Fireworks, NEDB, BlitzDB, CouchDB
  - Bellerophon, Sidre, Silo, ZFP, ASCTK, SCR, Sierra, DHARMA, DTK, PIO, Akuna, GridOPTICS software system (GOSS), DisPy, Luigi
  - CityGML, SIGMA (meshing), OpenStudio, Landscan USA
  - IMG/KBase, SRA, Globus, Python-PANDAS
- **Data Analytics and Visualization**
  - VisIt, VTK, Paraview, netCDF, CESIUM, Pymatgen, MacMolPlt, Yt
  - CombBLAS, Elviz, GAGE, MetaQuast
- **System Software**

Programming Models and Runtime

- **Goal:** A cross-platform, production-ready programming environment that enables and accelerates the development of mission-critical software at both the node and full-system levels

- **Current Portfolio:** Multiple programming models needed by ECP applications
  – MPI (MPICH, Open MPI), OpenMP, OpenACC, PGAS (UPC++, Global Arrays), task--based models (PaRSEC, Legion, DARMA), RAJA, Kokkos

Exascale MPI (MPICH)

- Lead: Pavan Balaji (ANL)
- Co-PI: Marc Snir (UIUC, ANL)
- Improvements to the MPICH implementation
  – Threads, heterogeneous memory, topology-aware communication, fault tolerance, new features that will be added to the MPI Standard, etc.
- Improvements to the MPI Standard for MPI-4
- Transfer technology to vendors to enable vendor-optimized implementations
- Coordinate with Open MPI and OpenMP efforts

Enhancing and Hardening the Legion Programming System for the Exascale Computing Project

- Lead: Galen Shipman (LANL)
- Co-PI: Alex Aiken (Stanford). Partners: ANL, NVIDIA
- High-level task-based programming model
- Runtime system automates the mapping and scheduling of tasks and the movement of data
- Used in S3D (combustion), LANL ATDM next-generation code, and others

Math Libraries and Frameworks Projects (1)

Project	Description
Kokkos Kernels	Performance portable (CPU/GPU) sparse matrix and graph kernels
Trilinos Linear Solvers	Robust scalable preconditioned Krylov solvers for next-gen platforms
Trilinos PDE Components	Meshing, discretization, integration compatible with Trilinos solvers
Trilinos Embedded Analysis	Sensitivity analysis, optimization, uncertainty quantification
xSDK	Compatible, interoperable, turnkey installation & access to libraries
SUNDIALS	Preparing SUNDIALS for next-generation platforms
PETSc/TAO	Preparing PETSc/TAO for next-generation platforms

Math Libraries and Frameworks Projects (2)

Project	Description
STRUMPACK/SuperLU	Next-generation sparse factorizations for multi-node CPU/GPU
ForTrilinos	Native, sustainable Fortran API to Trilinos next-gen capabilities
SLATE	Next-generation dense linear algebra for next-gen platforms
PEEKS	Latency-tolerant, production-quality iterative solvers and APIs
FleCSI	Framework for exploring next-gen component and execution models
MFEM	Advanced, high-order discretizations for next-gen platforms
ALExa	Multi-scale, multi-physics, sparse grid UQ

EXASCALE COMPUTING PROJECT

Performance Portability Tools

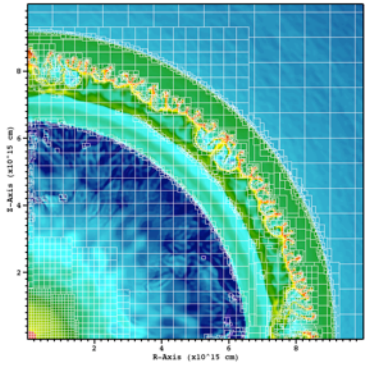
- LLVM, PI: McCormick (LANL)
  - Augment and enhance the HPC toolchain with LLVM based tools and technology
- Autotuning, PI: Hall (Utah)
  - Performance portability across CPUs and GPUs through domain-specific optimization and autotuning using the CHiLL framework and SURF search space navigation
- ROSE, PI: Quinlan (LLNL)
  - Support the automated generation of code for current and future compute architectures
- PROTEAS, PI: Vetter (ORNL)
  - Pathfinding programming solutions based on directive-based methodologies directed at emerging architectural features such as heterogeneous and manycore processors, deep memory hierarchies, and nonvolatile memory systems (NVM)
  - Support for OpenACC

ExaHDF5: Delivering Efficient Parallel I/O on Exascale Computing Systems

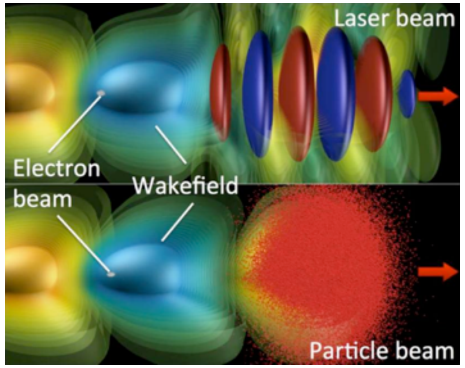
Suren Byna (LBNL, Lead PI), Quincey Koziol (LBNL), Scot Breitenfeld (The HDF Group),
Venkat Vishwanath, and Preeti Malakar (ANL)

- HDF5 is a mature technology being developed and released for 20 years
  - Among the Top-10 SW libraries used at NERSC, ALCF, and OLCF
- Features to be developed
  - Integration of Virtual Object Layer (VOL), data caching and prefetching using storage hierarchy, topology-aware I/O, async I/O, independent metadata updates, full single-writer – multiple-reader (SWMR), querying data and metadata, interoperability with netCDF and ADIOS file formats
- Several ECP applications either currently use or interested in using HDF5

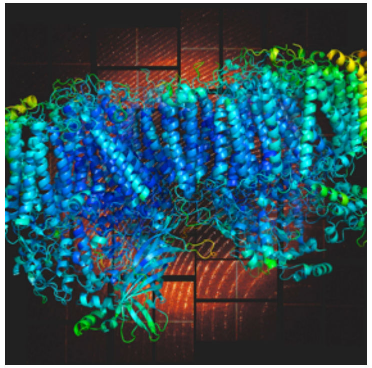
AMReX – 5 apps



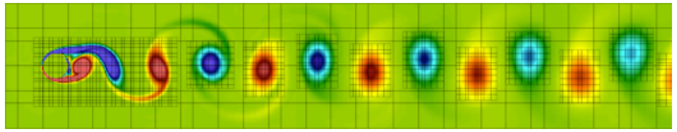
WarpX

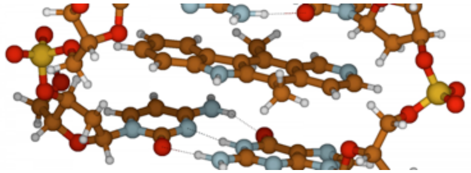


ExaFEL



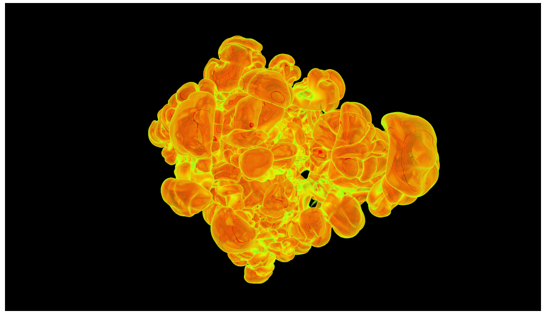
Subsurface simulation





QMCPACK

CLASH



EXASCALE COMPUTING PROJECT

Hardware Technology R&D

- The ECP PathForward program supports DOE-vendor collaborative R&D activities required to develop exascale systems with at least two diverse architectural features; quote from RFP:
  - PathForward seeks solutions that will improve application performance and developer productivity while maximizing energy efficiency and reliability of exascale systems.
- PathForward contracts were awarded recently to these companies:
  - Advanced Micro Devices (AMD)
  - Cray Inc. (CRAY)
  - Hewlett Packard Enterprise (HPE)
  - International Business Machines (IBM)
  - Intel Corp. (Intel)
  - NVIDIA Corp. (NVIDIA)
- Design Space Evaluation
  - Apply laboratory architectural analysis capabilities and Abstract Machine Models to PathForward designs to support ECP co-design interactions

ECP EXASCALE COMPUTING PROJECT

Selected Challenges (1)

- How measure the 50X in performance or complexity?
- Programming models and languages
  - How many are viable to support?
  - Which ones to support in this project for existing applications?
  - What will new applications enabled by exascale need?
- Applications teams often feel they can’t wait for math libraries to be available, have to roll their own – or they do not want to rely on software developed by others
  - The ECP is investing heavily on getting them ready on time AND satisfying application requirements

Selected challenges (2)

- Many moving targets: applications, algorithms, system architectures
- Portability and performance portability
- Co-design and integration for applications, hardware, support software at the scale of the ECP, with the global computational science research community (as opposed to companies)
- Guiding the vendors with our understanding of our applications and support software requirements while workload is changing
  - Difficult to design proxy applications that represent the key features of the full applications
  - Vendors can't deal with too many full or proxy applications

Challenges by area

- Application Development
  - Selecting challenge problems that require exascale and can realistically be developed in time
  - Portability with unknown target architectures
- Software Technology
  - Too many applications want deliverables from some of the ST projects
  - Will the applications use the ST results?
  - Coordinating with the vendors software stack
  - How meet needs of new types of applications or existing applications whose needs evolve due to exascale hardware features
- Hardware Technology
  - Will the vendor R&D projects result in better systems bid for facilities' RFP?

Challenges other than technical

- The "plumbing" aspects are a big challenge
- Co-design with so many sub-projects in so many institutions
  - Different cultures, different terminologies
- Need tight ties with the facilities that will acquire and operate the exascale systems
- Creating an integrated software stack
- Dual funding sources
- Coordinating and collaborating with the vendors' software plans

Communication is the key approach to tackling those challenges

- Co-design is a contact sport
- In this project we create many venues and mechanisms for focused and deep discussions among the participants in all three focus areas
  - Meetings, phone calls
  - Confluence software that enables view into all project activities, discussions, milestones, reports, etc.
- Essential but it comes at a cost: communication takes time and needs to be done continually
- And there are signs of success

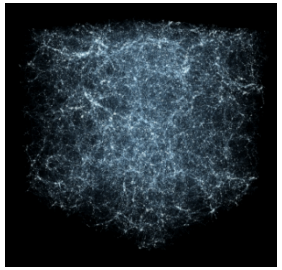
ECP EXASCALE COMPUTING PROJECT

ECP Applications Adopt New Infrastructure for Block-Structured Adaptive Mesh Refinement Developed by AMReX Co-Design Center

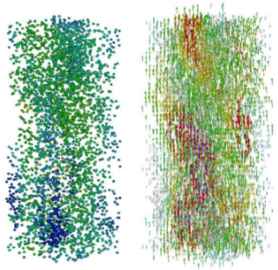
ECP WBS 1.2.5.03: AMReX
PI: John Bell, LBNL
Members: LBNL, ANL, NREL

Scope & Objectives

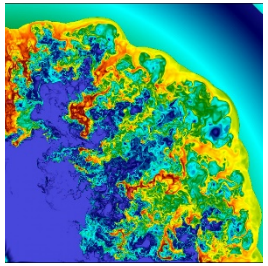
- Develop infrastructure to enable block-structured adaptive mesh refinement on exascale architectures
  - Core mesh, particle & particle-mesh operations on adaptive mesh hierarchy
  - Support for multiple time-stepping approaches
  - Embedded boundary representation of complex geometry
  - Performance portability for different architectures
- Current activities focused on:
  - Establishing support for core AMR functionality
  - Engagement of applications



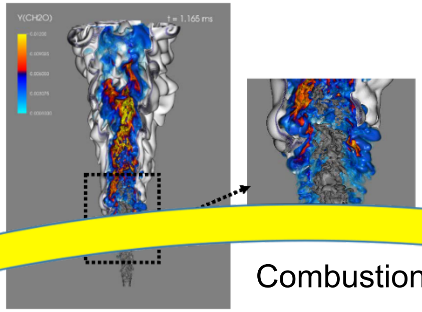
Cosmology



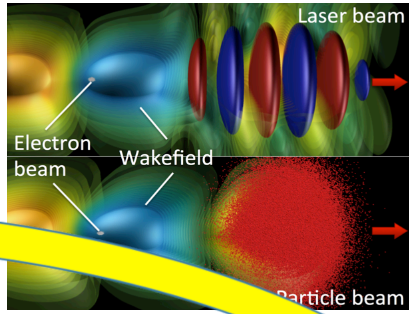
Multiphase flow in chemical reactors



Astrophysics



Combustion



Accelerator design

Impact

- Established a next-generation framework for developing block-structured adaptive mesh refinement algorithms for current and emerging architectures
- Provides a common framework for multiple ECP applications that use AMR
- Provides a common focal point for software technology, hardware technology and vendors to leverage activities over multiple applications
- Broad constituency within Office of Science and NNSA

Project Accomplishment

- New AMReX code framework adopted by multiple ECP applications
  - Accelerator modeling – WarpX
  - Astrophysics -- ExaStar (CASTRO)
  - Combustion – PeleC and PeleLM
  - Cosmology – ExaSky (Nyx)
  - Multiphase flow – MFIX-Exa
- AMReX code framework publicly released

Deliverables: AMReX software available at https://github.com/AMReX-Codes/amrex

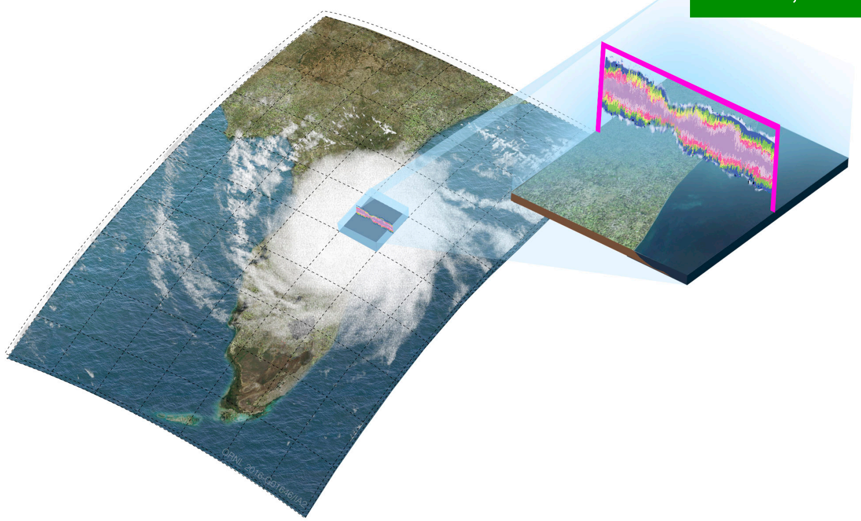
EXASCALE COMPUTING PROJECT

ECP WBS 1.2.1.15/ACME-MMF
Taylor, SNL
Members: ANL, LANL, LLNL, ORNL, PNNL, CSU, UCI

ACME-MMF Cloud Resolving Climate Model

Scope & Objectives

- 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.
- 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.
- ACME-MMF: Use a multiscale approach ideal for new architectures to achieve cloud resolving convection on Exascale resources



MMF framew... scale model (left). Within each grid box in the ... scale model, ACME-MMF will replace ... meterizations with a cloud resolving model

ACME-MMF Impact

- A cloud resolving climate model is needed to reduc... major systematic errors in climate simulations d... to structural uncertainty in numerical treatments o... convection – such as convective storm system...

Project Accomplishments

- Integration of multiscale modeling framework (MMF) into ACME code base
- Integration of the SAM cloud resolving model into the ACME code base
- Development of an ACME model configuration using these new components (ACME-MMF)
- Evaluate computational performance on CPU systems using "smoke tests": short runs that ensure the software can produce a solution that is realistic enough that the simulation can run for several days without crashing.
- This code will serve as the basis of our future work on GPU optimization/porting and development of the MMF approach for modeling the Earth's water cycle

Deliverables: Y1Q2 milestone report at https://confluence.exascalep... /display/ADSE15/ACME-MMF+smoke+test+on+CPU...

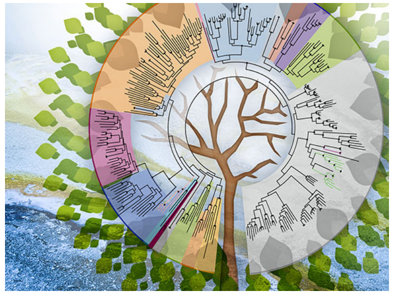
ExaBiome provides first scalable algorithms for high quality metagenome assembly and analysis

ECP WBS 1.2.1.20: ExaBiome
PI: Kathy Yelick, LBNL
Members: LBNL, LANL, JGI

Scope & Objectives

- Develop **genome assembly, protein clustering and comparative analysis** codes for exascale, using the aggregate memory and high speed networks
- **Assemble millions of metagenomes** without filtering data
- **Cluster billions of proteins** for discovery and to unlock functional behavior
- **Compare thousands of metagenomes** for environmental monitoring and analysis
- Discover new species and functions in large, complex metagenome data sets

Understanding ecosystems



New life forms

Impact

- New scalable metagenome assembler (MetaHipmer) reduces runtime by orders of magnitude at petascal
- Demonstrated MetaHipMer assemblies have qual comparable to best assemblers used in productio
- New scalable protein clustering algorithm sped runt e from 15 weeks on single node to 1 hour
- Largest protein clustering from JGI metagenome data

Project Accomplishment and Next Steps

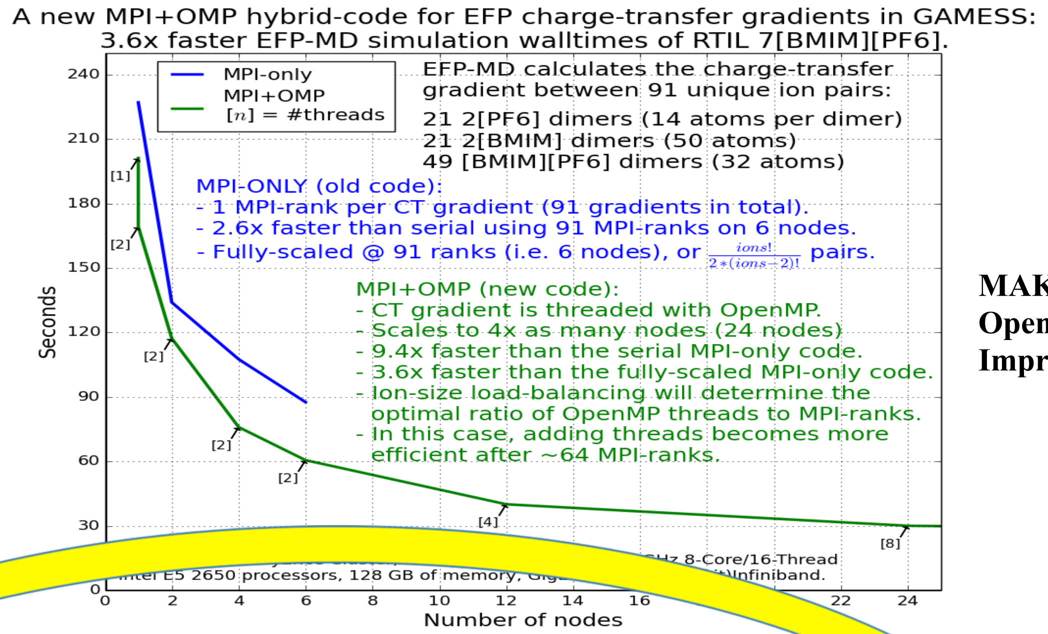
- MetaHipMer code consistently provides the lowest number of mismatches with the highest error-free contiguity compared to state-of-the-art assemblers, including metaSPAdes and MEGAHIT (Milestone 1)
- MetaHipMer scales linearly with cores on KNL (Cori) and to thousands of nodes
- Identified architectural features need for scalable assembly
- Protein clustering (HipMCL) uses communication-avoiding optimizations and novel parallelism for unprecedented performance at petascale (Milestone 2)
- First-of-kind clustering analysis enabled by HipMCL

OpenMP Threads Improve Performance of the Effective Fragment Potential Charge Transfer Gradients in the General Atomic and Molecular Electronic Structure System (GAMESS)

ECP WBS: 1.2.1.16 GAMESS
PI: Mark Gordon, Ames/Iowa State
Members: Ames, Iowa State

Scope & Objectives

- Improve EFMO scaling by implementing OpenMP/threads for MAKEFP code
- Start with most computationally expensive charge transfer (CT) term
- CT term is more expensive than all other terms combined



MAKEFP CT OpenMP Improvements

Impact

- GAMESS users will be able to run EFP and EF... calculations much more efficiently
- The GAMESS-QMCPACK interface will bene... since the main target for that effort is EFMO
- BES Computational & Theoretical Chemistry a... CPIMS programs will benefit from increased GAM... capability

Project Accomplishment

- The CT MAKEFP code has been improved by the OpenMP implementation (s... above figure).
- Next steps: Extend to other MAKEFP terms
- Contributors: Alex Findlater, Sarom Leang

EXASCALE COMPUTING PROJECT

Deliverables: Milestone report Q3; currently hosted at ...

I think we can

- Create an exascale ecosystem that will support many applications and modes of usage
- Foster a culture of co-design, software engineering, and collaboration

- Yes, I am an optimist, but I truly think we can succeed

This research was supported by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of two U.S. Department of Energy organizations (Office of Science and the National Nuclear Security Administration) responsible for the planning and preparation of a capable exascale ecosystem, including software, applications, hardware, advanced system engineering and early testbed platforms, in support of the nation’s exascale computing imperative.

EXASCALE COMPUTING PROJECT

Thank you!

EXASCALE COMPUTING PROJECT