The Exascale Computing Project

Paul Messina, ECP Director

CHEP 2016

October 10, 2016

EXASCALE COMPUTING PROJECT

National Nuclear Security Administration

U.S. DEPARTMENT OF ENERGY | Office of Science

What is the Exascale Computing Project?

- Who in this room has heard of the Exascale Computing Project?

- When we say the Exascale Computing Project – what comes to mind?
  – Hardware / systems / platforms?
  – Software / software stack?
  – Applications?

If you were thinking ‘all the above’ – you were right.

ECP is the vehicle for one of the DOE roles in NSCI

NSCI = National Strategic Computing Initiative, established by Pres. Obama Executive Order

- DOE is a lead agency within NSCI; DOE SC and NNSA will execute a joint effort on advanced simulation through a **capable exascale** computing program emphasizing sustained performance on relevant applications and data analytic computing
- Starting in FY 2016, the Exascale Computing Project (ECP) was initiated as a DOE-SC/NNSA-ASC partnership, using DOE's formal project management processes
- The ECP is a ten-year project led by DOE laboratories and executed in collaboration with academia and industry
- The ECP leadership team has staff from six U.S. DOE labs
  - Staff from most of the 17 DOE national laboratories will take part in the project
- The ECP collaborates with the facilities that operate DOE's most powerful computers

Four key challenges that must be addressed to achieve exascale

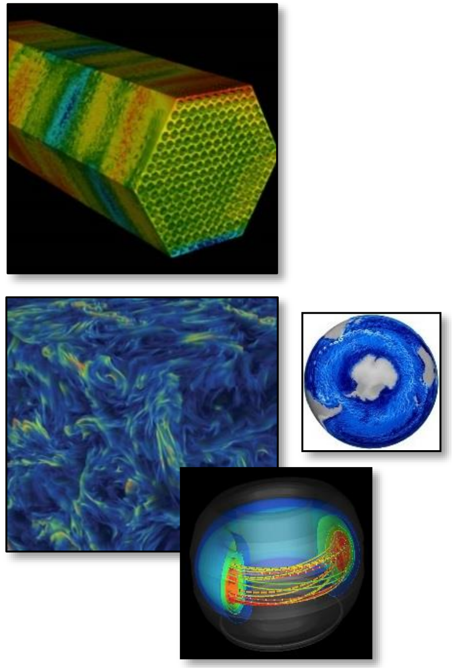
- Parallelism
- Memory and Storage
- Reliability
- Energy Consumption

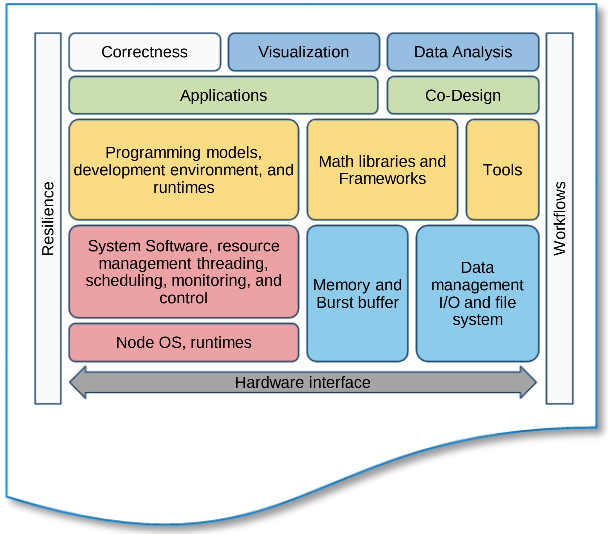
Achieving *capable* exascale computing

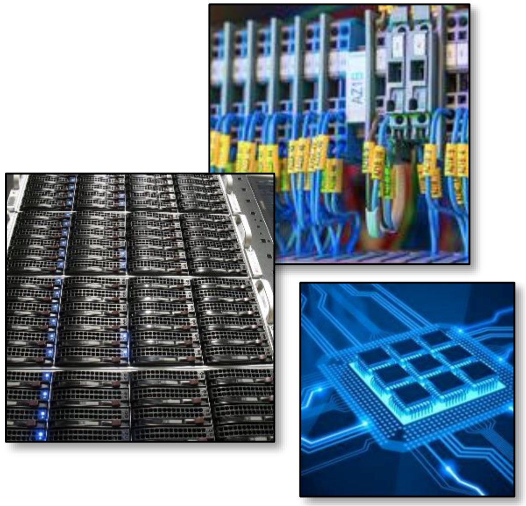
- Support applications solving science problems 50× faster or more complex than today's 20 PF systems
- Operate in a power envelope of 20–30 MW
- Be sufficiently resilient (average fault rate no worse than weekly)
- At least two diverse system architectures
- Possess a software stack that meets the needs of a broad spectrum of applications
- A holistic project approach is needed that uses co-design to develop new platform, software, and computational science capabilities at heretofore unseen scale
  - Essential for tackling much deeper challenges than those that can be solved by hardware scale alone

To achieve capable exascale requires a holistic approach

Application Development	Software Technology	Hardware Technology	Exascale Systems
Science and mission applications	Scalable and productive software stack	Hardware technology elements	Integrated exascale supercomputers







ECP's work encompasses applications, system software, hardware technologies and architectures, and workforce development

Applications Development activities

- Fund applications development teams
  - Each aiming at capability and specific challenge problems
  - Following software engineering practices
  - Tasked to provide software and hardware requirements
  - Execute milestones jointly with software activities
- Establish co-design centers for commonly used methods
    - E.g., Adaptive Mesh Refinement, Particle-in-Cell
- Developer training

ECP Applications Deliver Broad Coverage of Strategic Pillars

Initial (FY16) selections consist of 15 application projects + 7 seed efforts

National Security

- Stockpile Stewardship

Energy Security

- Turbine Wind Plant Efficiency
- Design/Commercialization of SMRs
- Nuclear Fission and Fusion Reactor Materials Design
- Subsurface Use for Carbon Capture, Petro Extraction, Waste Disposal
- High-Efficiency, Low-Emission Combustion Engine and Gas Turbine Design
- Carbon Capture and Sequestration Scaleup (S)
- Biofuel Catalyst Design (S)

Economic Security

- Additive Manufacturing of Qualifiable Metal Parts
- Urban Planning (S)
- Reliable and Efficient Planning of the Power Grid (S)
- Seismic Hazard Risk Assessment (S)

Scientific Discovery

- Cosmological Probe of the Standard Model (SM) of Particle Physics
- Validate Fundamental Laws of Nature (SM)
- 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 (S)

Climate and Environmental Science

- Accurate Regional Impact Assessment of Climate Change
- Stress-Resistant Crop Analysis and Catalytic Conversion of Biomass-Derived Alcohols
- Metegenomics for Analysis of Biogeochemical Cycles, Climate Change, Environ Remediation (S)

Healthcare

- Accelerate and Translate Cancer Research

Exascale Modeling of Advanced Particle Accelerators*

Exascale Challenge Problem

- Design affordable compact 1 TeV electron-positron collider based on plasma acceleration for high-energy physics.
- Develop ultra-compact plasma accelerators with transformative implications in discovery science, medicine, industry and security.
- Enable virtual start-to-end optimization of the design and virtual prototyping of every component before they are built, leading to huge savings in design and construction.
- Develop powerful new accelerator modeling tool (WarpX) designed to run efficiently at scale on exascale supercomputers.

Applications & S/W Technologies

Applications

- Warp, PICSAR

Software Technologies Cited

- Foftran, C, C++, Python
- MPI, OpenMP, GASNet, UPC++
- BoxLib
- HDF5, VisIT, Paraview, YT

Risks and Challenges

- Dynamic load balancing
- Parallel I/O, data analysis & visualization
- Scaling electrostatic solver
- Scaling high-order electromagnetic solver
- Spurious reflections or charges owing to AMR
- Lost signal with AMR
- Numerical Cherenkov instability
- Low temperature plasmas/beams

Development Plan

Y1: Modeling of single plasma-based accelerator stage with WarpX on single grid; verification against previous results.

Y2: Modeling of single plasma-based accelerator stage with WarpX with static mesh refinement. plasma case.

Y3: Optimized FDTD and spectral PIC on to 5-to-10 millions of cores, near-linear weak scaling with AMR, on a uniform plasma case.

Y4:Convergence study in 3-D of ten consecutive multi-GeV stages in linear and bubble regime. Release of software to community.

*PI: Jean-Luc Vay (LBNL)

Computing the Sky at Extreme Scales*

Exascale Challenge Problem

- To evolve and meld aspects of the capabilities of Lagrangian particle-based techniques (gravity + gas) with Eulerian adaptive mesh resolution (AMR) methods to achieve a unified cosmological simulation approach at the exascale
- determine the dark energy equation of state
- search for deviations from general relativity
- determine the neutrino mass sum (to less than 0.1 eV) from galaxy clustering measurements
- characterize the properties of dark matter
- testing the theory of inflation

Applications & S/W Technologies

Applications

- HACC, NYX

Software Technologies Cited

- UPC++, C++17
- MPI, OpenMP, OpenCL, CUDA
- BoxLib, High-Performance Geometric Multigrid (HPGMG), FFTW, PDACS
- Thrust

Risks and Challenges

- Accuracy of subgrid modeling
- Filesystem stability and availability and fast access to storage for post-processing
- Loss of personnel. Team is relatively small; need to avoid single points of failure
- Resilience - machine MTBF

Development Plan

Y1: First major HACC hydro simulation on Theta on full machine;First HACC tests on IBM Power8/NVIDIA Pascal 36 node early-access system; Release HACC & Nyx

Y2: Access to 25% of Summit system as part of CAAR project with HACC, scaling runs and optimization; Nyx scale-up test on Cori/Theta: Clusters of galaxies (deep AMR); Summit CAAR project simulations: HACC hydrodynamic simulations on full machine; Release HACC & Nyx

Y3: HACC and Nyx scaling runs; Scaling of CosmoTools to full scale on Aurora; meeting FOMs; Scaling of CosmoTools to full scale on Summit; Release HACC & Nyx

Y4: Final Major HACC and Nyx code releases

*PI: Salman Habib (ANL)

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

Exascale Challenge Problem

- Develop a software infrastructure that exploits recent compiler advances and improved language support to enable the creation of portable, high-performance QCD code with a shorter software tool-chain
- Focus on two nuclear/HEP applications:
- Compute from first principles the properties and interactions of nucleons and ight nuclei with physical quark masses and achieve the multi-physics goal of incorporating both QCD and electromagnetism
- Search for beyond-the-standard-model physics by increasing the precision of calculations of the properties of quark-anti-quark and three-quark states.

Applications & S/W Technologies

Applications

- MILC, Columbia Physics System, Chroma, QDP++ (all built upon USQCD software infrastructure)

Software Technologies Cited

- MPI, OpenMP, CUDA, Kokkos, OpenACC, C++17, Thrust,
- SyCL, QUDA, QPhiX, LAPACK, ARPACK

Risks and Challenges

- Critical slowing down in gauge evolution
- Correlation functions for large nuclei
- Sub-optimal solver performance
- Multi-level time integration for correlation functions
- Performance in data-parallel GPU offload
- Extending multigrid solver base
- Architectural pathfinding (appropriate parallelization approach)

Development Plan

Y1: Develop adaptive multigrid for domain wall and staggered fermions

Y2: Release new versions of old apps augmented with new implementations of algorithms; Initial release of Workflow framework

Y3: Release of data parallel API with GPU support; Release benchmark suite for non volatile memory storage; Scalable MG based deflation methods with variance reduction applicable to other HP domains

Y4: Validated and documented high performance code implementing the 1-2 most successful algorithms for reducing critical slowing down

*PI: Paul Mackenzie (FNAL)

Data Analytics at the Exascale for Free Electron Lasers*

Exascale Challenge Problem

- LCLS detector data rates up 10^3-fold by 2025; XFEL data analyses times down from weeks to min with real-time interpretation of molecular structure revealed by X-ray diffraction; LCLS-II beam rep rate goes from 120 Hz to 1 MHz by 2020.
- LCLS X-ray beam, @ atomic scale wavelengths & 10^9 brighter than other sources, probes complex, ultra-small structures with ultrafast pulses to freeze atomic motions
- Science drivers to orchestrate compute, network, and storage: Serial Femtosecond Crystallography (SFX) and Single Particle Imaging (SPI)
- SFX: study of biological macromolecules (e.g., protein structure / dynamics) and crystalline nano materials; need rapid image analysis feedback on diffraction data to make experimental decisions
- SPI: discern 3D molecular structure of individual nano particles & molecules; rapid diffraction pattern tuning of sample concentrations needed for sufficient single particle hit rate, adequate data collection

Applications & S/W Technologies

Applications

- Psana Framework, cctbx, Iunus, M-TIP, IOTA

Software Technologies Cited

- Tasking runtime (Legion)
- C++, Python
- MPI, OpenMP, CUDA
- FFT, BLAS/LAPACK
- HDF5, Shifter, XTC

Risks and Challenges

- Schedulability of NERSC & LCLS resources
- LCLS-II data rate > ESnet data rate
- HPC execution overhead
- Scalability of file format(s)
- Keeping up with network infrastructure upgrades
- Experiment calendar uncertainty
- New model for HPC utilization; bursty, short workloads imply lower machine utilization unless resilient jobs can be preempted
- Maturity of tasking runtime/image analysis kernels

Development Plan

Y1: Release cctbx, psana: Benchmark exascale aware M-TIP routines; Prototype psana tasking, image kernels ported to C++ parallel STL; Deploy live streaming HDF5 files from FFB to Cori; SFX experiment on Cori PII

Y2: Release cctbx, psana, M-TIP; port psana- MPI and tasking to Summit, key image kernels ported to image DSL; Decision for psana-MPI vs psana-task; SFX experiment using IOTA on Cori PII & SPI experiment on Cori PII

Y3: Release cctbx, psana, M-TIP; Optimized M-TIP under streaming; cctbx & M-TIP integrated into psana with 50% scaling Cori PII & Sierra; SFX experiment using ray tracing

Y4: End-to-end cctbx on NERSC9 / ESnet6; optimized scheduler for live streaming jobs on Cori; SFX & SPI experimental demos for LCLS users visualizing strructures < 10 min

*PI: Amedeo Perazzo (SLAC)

Application Motifs*

Algorithmic methods that capture a common pattern of computation and communication

1. **Dense Linear Algebra**
   – Dense matrices or vectors (e.g., BLAS Level 1/2/3)
2. **Sparse Linear Algebra**
   – Many zeros, usually stored in compressed matrices to access nonzero values (e.g., Krylov solvers)
3. **Spectral Methods**
   – Frequency domain, combining multiply-add with specific patterns of data permutation with all-to-all for some stages (e.g., 3D FFT)
4. **N-Body Methods (Particles)**
   – Interaction between many discrete points, with variations being particle-particle or hierarchical particle methods (e.g., PIC, SPH, PME)
5. **Structured Grids**
   – Regular grid with points on a grid conceptually updated together with high spatial locality (e.g., FDM-based PDE solvers)
6. **Unstructured Grids**
   – Irregular grid with data locations determined by app and connectivity to neighboring points provided (e.g., FEM-based PDE solvers)
7. **Monte Carlo**
   – Calculations depend upon statistical results of repeated random trials
8. **Combinational Logic**
   – Simple operations on large amounts of data, often exploiting bit-level parallelism (e.g., Cyclic Redundancy Codes or RSA encryption)
9. **Graph Traversal**
   – Traversing objects and examining their characteristics, e.g., for searches, often with indirect table lookups and little computation
10. **Graphical Models**
   – Graphs representing random variables as nodes and dependencies as edges (e.g., Bayesian networks, Hidden Markov Models)
11. **Finite State Machines**
   – Interconnected set of states (e.g., for parsing); often decomposed into multiple simultaneously active state machines that can act in parallel
12. **Dynamic Programming**
   – Computes solutions by solving simpler overlapping subproblems, e.g., for optimization solutions derived from optimal subproblem results
13. **Backtrack and Branch-and-Bound**
   – Solving search and global optimization problems for intractably large spaces where regions of the search space with no interesting solutions are ruled out. Use the divide and conquer principle: subdivide the search space into smaller subregions (“branching”), and bounds are found on solutions contained in each subregion under consideration

The Landscape of Parallel Computing Research: A View from Berkeley, Technical Report No. UCB/EECS-2006-183 (Dec 2006).

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		■	■			■							

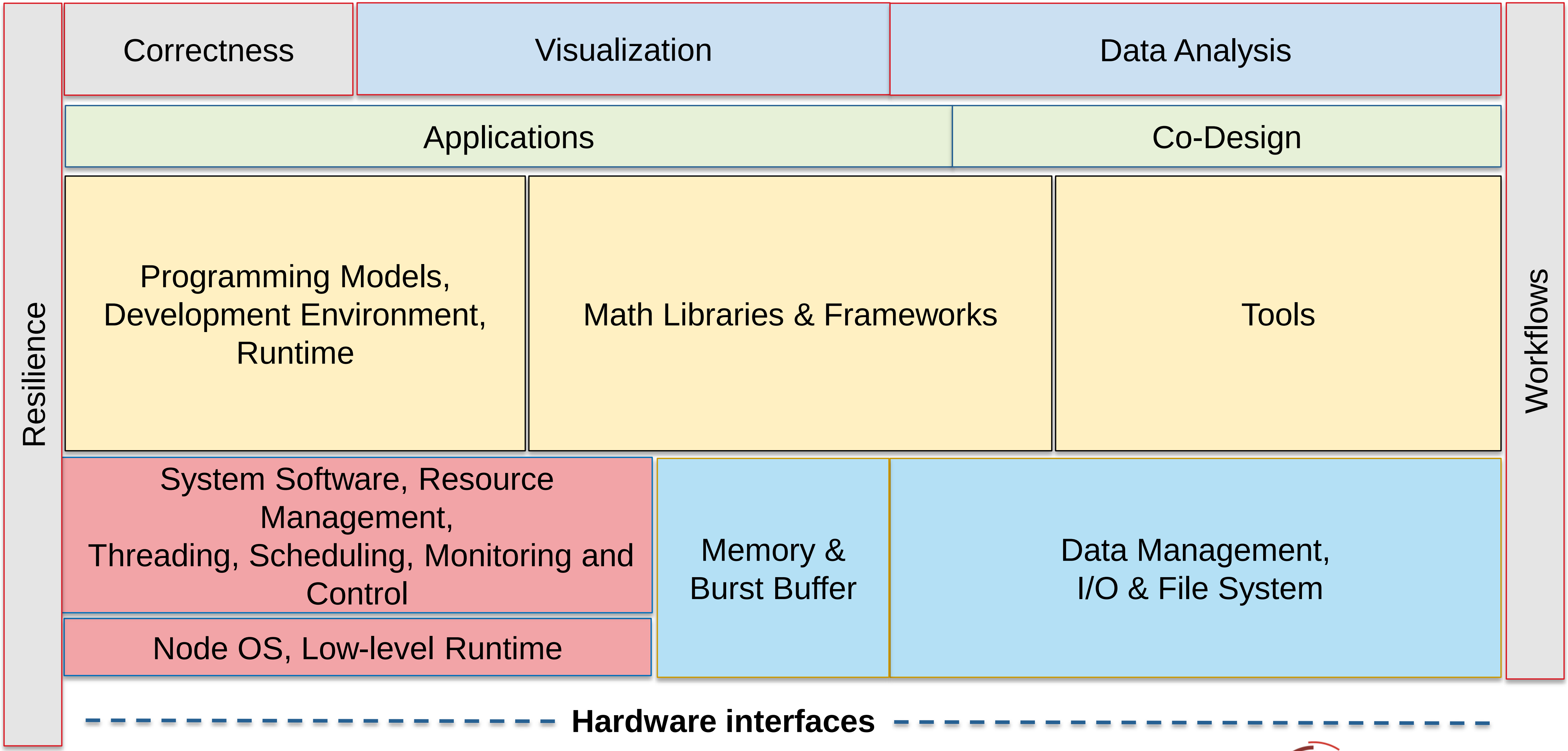
EXASCALE COMPUTING PROJECT

Requirements for Software Technology

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 in 2–3 years
- Expected software environment for an exascale system
- Requirements beyond the software environment provided by vendors of HPC systems

Conceptual ECP Software Stack



Hardware Technology Activities

- PathForward: support 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.
- Design Space Evaluation
  - Apply laboratory architectural analysis capabilities and Abstract Machine Models to PathForward designs to support ECP co-design interactions

EXASCALE COMPUTING PROJECT

Exascale Systems Activities

- Ensure at least two exascale-class systems are accepted no later than 2023 and the systems are diverse, affordable, production-ready, and capable
  - NRE contracts
  - Convey results of ECP R&D to RFP for exascale systems procurement that facilities will issue
- Provide requirements from facilities viewpoint
- Acquire and operate testbeds for application and software development projects and for hardware investigations

ECP phases

- 2016 – 2019
  - Develop applications, conduct R&D&D on software technologies
    - Use current systems, CORAL systems as testbeds
  - Vendor R&D on node and system designs that are better suited for HPC applications
- 2019
  - **ECP insights are used in formulation of RFP for exascale systems**
  - DOE and NNSA laboratories issue RFP for exascale systems, select offers, award build and NRE contracts
- 2019-2023
  - ECP Applications and software technologies are modified with knowledge of systems
  - Software technologies are “productized”
- 2023-2025
  - Exascale systems are in production, applications and software deal with actual system behavior

ECP status

- Solicited and received proposals for
  - applications development,
  - co-design centers, and
  - software technology activities
  - Hardware technology R&D
- 22 application proposals have been selected for funding
- Co-design centers and software technology proposals are being evaluated
- Initial awards will be made this FY
- Responses to PathForward RFP (Hardware Technology R&D by vendors) have been evaluated and proposals selected for funding
  - Contracts expected to be put in place this fall

Exascale Computing Project goals

Develop scientific, engineering, and large-data applications that exploit the emerging, exascale-era computational trends caused by the end of Dennard scaling and Moore’s law	Create software that makes exascale systems usable by a wide variety of scientists and engineers across a range of applications	Enable by 2023 two diverse computing platforms with up to 50× more computational capability than today’s 20 PF systems, within a similar size, cost, and power footprint	Help ensure continued American leadership in architecture, software and applications to support scientific discovery, energy assurance, stockpile stewardship, and nonproliferation programs and policies
Foster application development	Ease of use	Two diverse architectures	US HPC leadership

Thank you!