The DOE Exascale Computing Project:
Overview of Relevant Energy/Power Efforts

Presented to:

8th Annual Workshop for Energy Efficient HPC Working Group

Supercomputing 2017

Colorado Convention Center, Denver, CO

November 12, 2017

James Ang, Ph.D.
Manager, Exascale Computing Program
Center for Computing Research
Sandia National Laboratories

Acknowledgements: Paul Messina, ANL, Doug Kothe, ORNL, Stephen Lee, LANL, Mike Heroux, SNL, Terri Quinn, LLNL and John Shalf, LBNL

EXASCALE COMPUTING PROJECT

National Nuclear Security Administration

U.S. DEPARTMENT OF ENERGY | Office of Science

SAND2017-12300 C

www.ExascaleProject.org

The Exascale Computing Project (ECP)

- A collaborative effort of Two US Department of Energy (DOE) organizations:
  - Office of Science (DOE-SC)
  - National Nuclear Security Administration (NNSA)
- A 7-year project to accelerate the development of *capable* exascale systems
  - Led by DOE laboratories
  - Executed in partnership with academia and industry

A *capable* exascale computing system will leverage a balanced ecosystem (software, hardware, applications)

What is a *capable* exascale computing system?

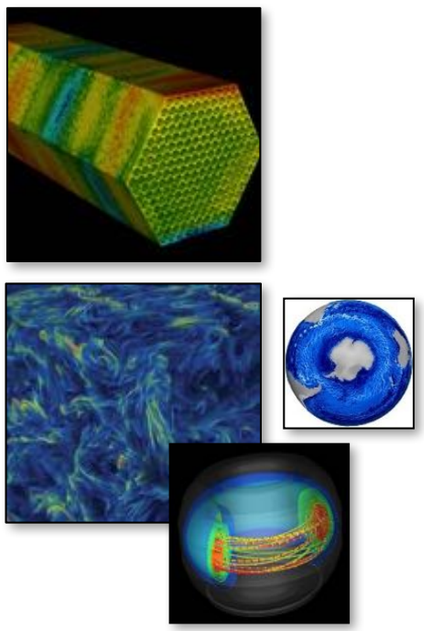
A capable exascale computing system requires an entire computational ecosystem that:

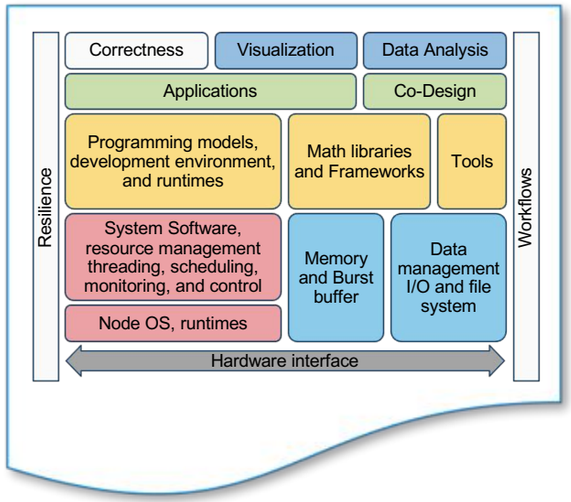
- Delivers 50× the performance of today's 20 PF systems, supporting applications that deliver high-fidelity solutions in less time and address problems of greater complexity
- Operates in a power envelope of 20–30 MW
- Is sufficiently resilient (average fault rate: ≤1/week)
- Includes a software stack 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

ECP has formulated a holistic approach that uses co-design and integration to achieve capable exascale

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

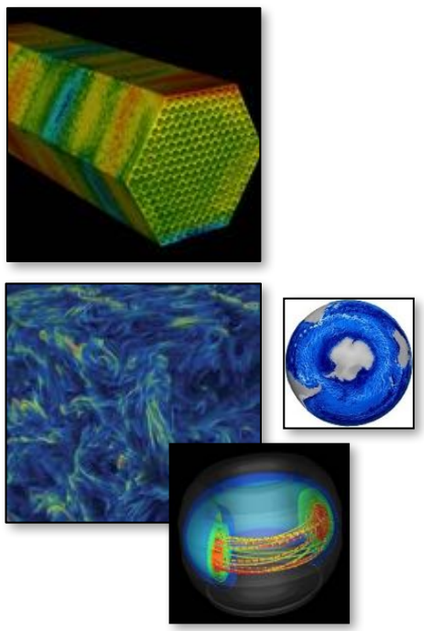


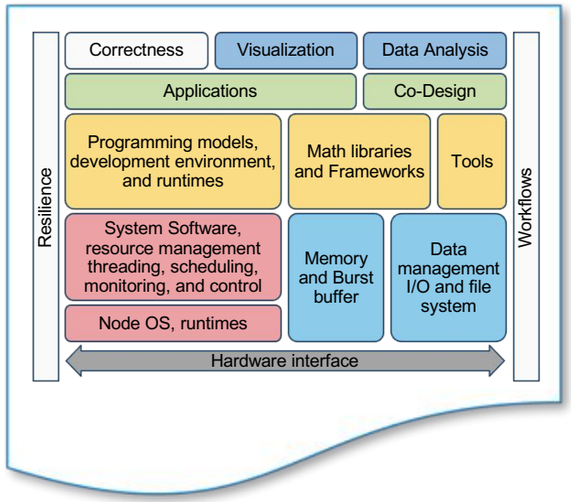


ECP's scope encompasses applications, system software, hardware technologies and architectures, and workforce development

ECP has formulated a holistic approach that uses co-design and integration to achieve capable exascale

Application Development	Software Technology	Hardware & Integration	Facility Preparation & Procurement
Science and mission applications	Scalable and productive software stack	Hardware technology elements	Integrated exascale supercomputers



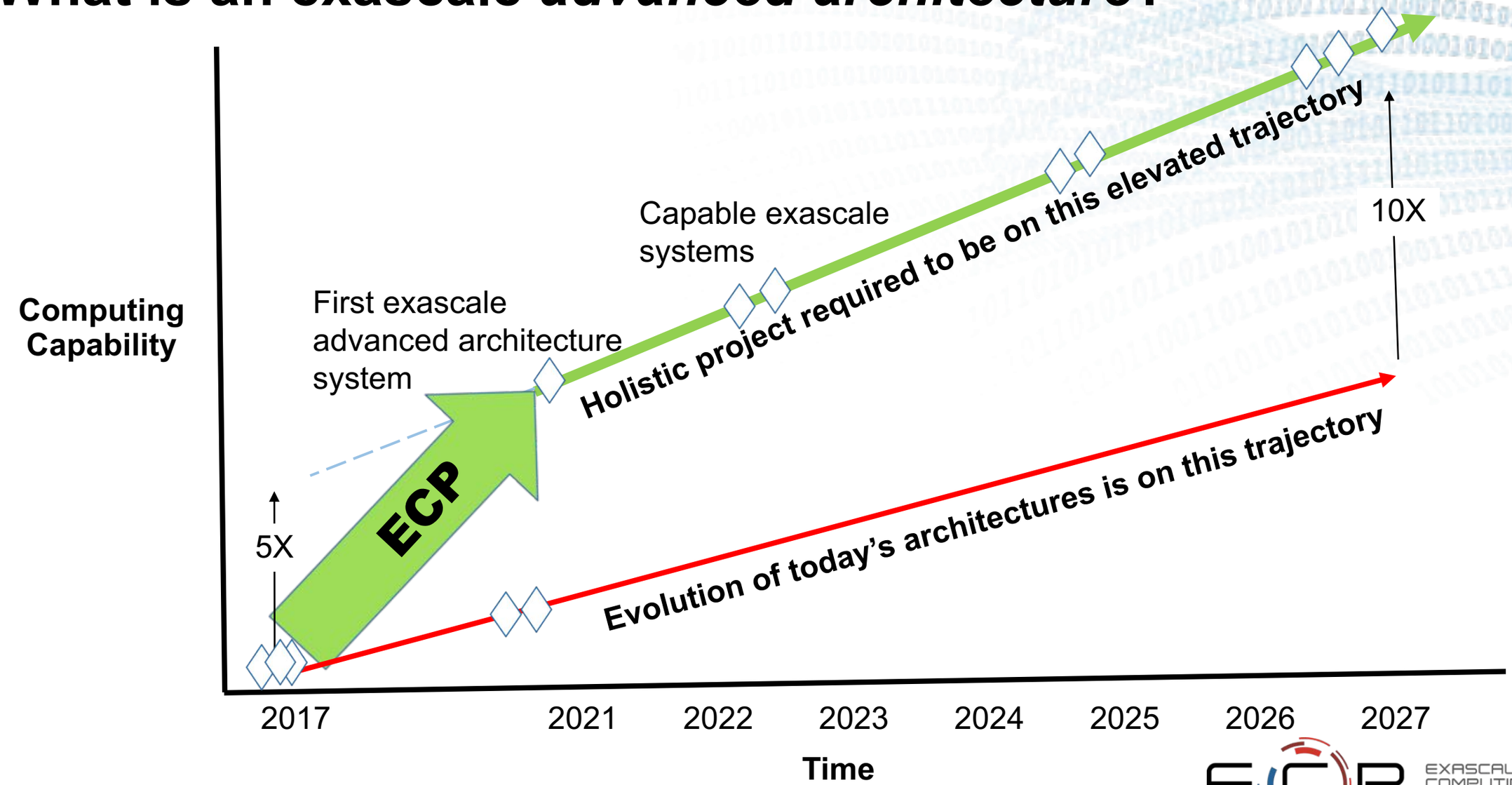


ECP's scope encompasses applications, system software, hardware technologies and architectures, and workforce development

The ECP plan of record

- A 7-year project that follows the *holistic/co-design* approach, which runs through 2023 (including 12 months of schedule contingency)
  - To meet the ECP goals
- Enable initial exascale systems based on an *advanced architectures* and delivered in 2021
- Enable *capable* exascale systems, based on ECP R&D, delivered in 2022 and deployed in 2023 as part of an NNSA and SC facility upgrades
- Acquisition of the exascale systems is outside of the ECP scope, will be carried out by DOE-SC and NNSA-ASC facilities

What is an exascale *advanced architecture*?



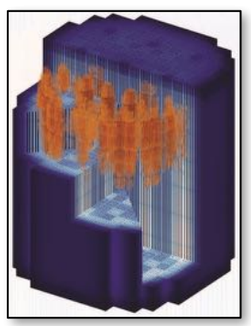
ECP Mission Need Defines the Application Strategy

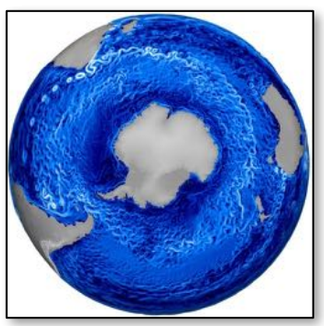
Support DOE science *and* energy missions	Meet national security needs	Key science and technology challenges to be addressed with exascale
• Discover and characterize next-generation **materials** • Systematically understand and improve **chemical processes** • Analyze the extremely large datasets resulting from the next generation of **particle physics** experiments • Extract knowledge from **systems-biology** studies of the microbiome • Advance **applied energy** technologies (e.g., whole-device models of plasma-based fusion systems)	• Stockpile Stewardship Annual Assessment and Significant Finding Investigations • Robust uncertainty quantification (UQ) techniques in support of stockpile lifetime extension programs • Understanding evolving nuclear threats posed by adversaries and in developing policies to mitigate these threats	• Materials discovery and design • Climate science • Nuclear energy • Combustion science • Large-data applications • Fusion energy • National security • Additive manufacturing • Many others!

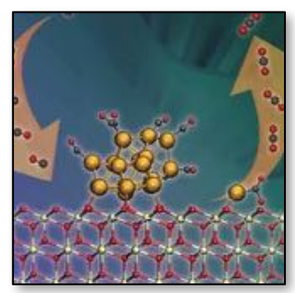
Exascale Applications Will Address National Challenges

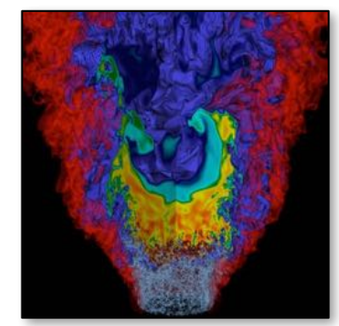
Summary of current DOE Science & Energy application development projects

Nuclear Energy (NE)	Climate (BER)	Chemical Science (BES, BER)	Wind Energy (EERE)	Combustion (BES)
Accelerate design and commercialization of next-generation small modular reactors* Climate Action Plan; SMR licensing support; GAIN	Accurate regional impact assessment of climate change* Climate Action Plan	Biofuel catalysts design; stress-resistant crops Climate Action Plan; MGI	Increase efficiency and reduce cost of turbine wind plants sited in complex terrains* Climate Action Plan	Design high-efficiency, low-emission combustion engines and gas turbines* 2020 greenhouse gas and 2030 carbon emission goals









* Scope includes a discernible data science component

ECP Co-Design Centers (CDCs)

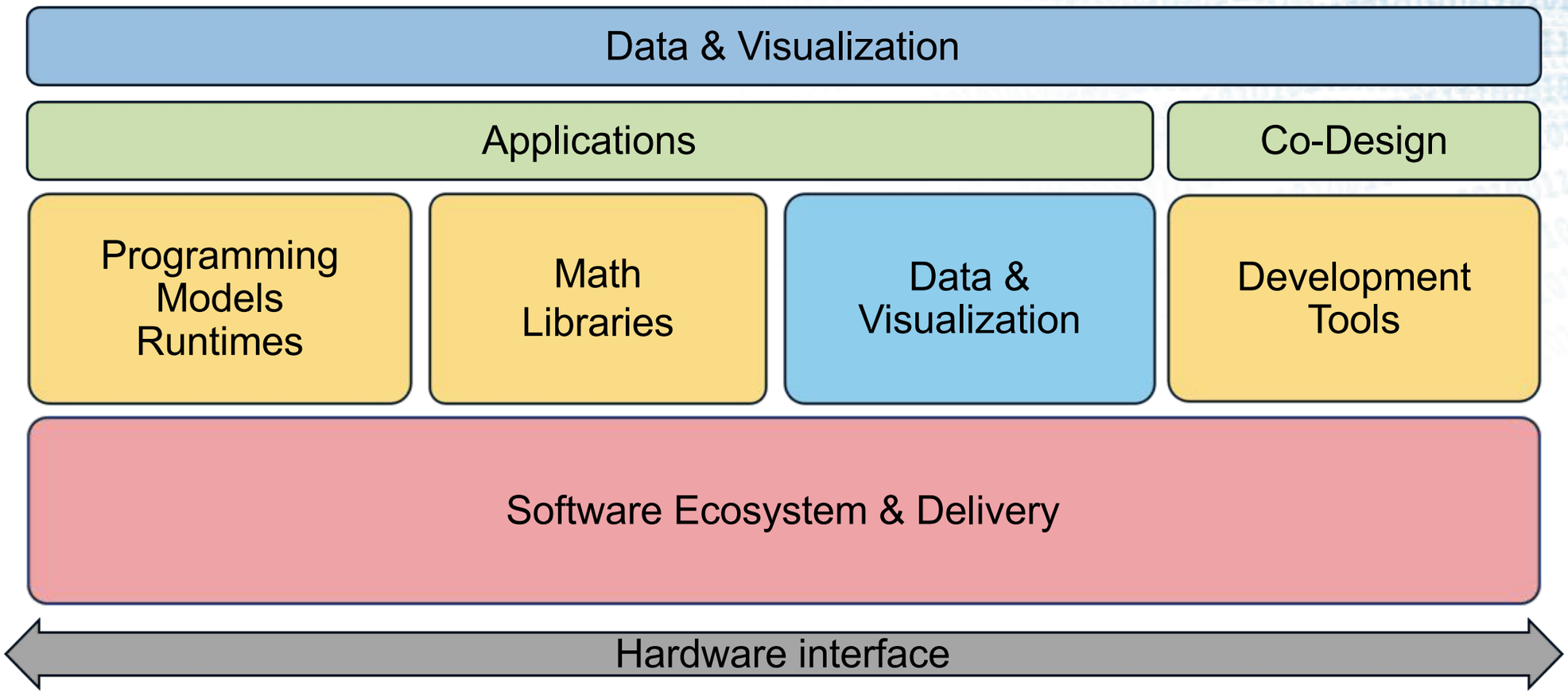
- **CODAR: CDC for Online Data Analysis and Reduction at the Exascale, Ian Foster, ANL**
  - Motifs: Online data analysis and reduction
  - Address growing disparity between simulation speeds and I/O rates rendering it infeasible for HPC and data analytic applications to perform offline analysis. Target 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)
- **CDC for Block-Structured AMR, John Bell, LBNL**
  - 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
- **CDC for Efficient Exascale Discretizations (CEED), Tzanio Kolev, LLNL**
  - 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
- **CDC for Particle-based Methods, Tim Germann, LANL**
  - 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

Software Technology Summary

- ECP will build a comprehensive and coherent software stack that will enable application developers to productively write highly parallel applications that can portably target diverse exascale architectures

- ECP will accomplish this by:
  - Extending current technologies to exascale where possible
  - Performing R&D required to develop new approaches where necessary
  - Coordinating with vendor efforts to develop and deploy high-quality and robust software products

ECP ST Software Stack



ST: Runtime System for Application-Level Power Steering

Scope & Intent	R&D Themes	Delivery Process	Target ECP Users	Support Model
Safe operation and performance optimization of scientific applications in power-limited scenarios.	Maximizing performance under a power bound, adaptive configuration selection, critical path analysis.	Regular releases of software and documentation, open access through GitHub and relevant platforms.	ECP application developers, ECP Proxy apps, other apps such as ParaDis/S3D, ECP projects (Argo, Flux, and Caliper)	GitHub issue tracking, web documentation, tutorials and developer support.

Goal*	Metric
Widespread use of GEOPM across ECP-enabled applications.	Number of ECP benchmarks, scientific applications, system software components, and processor architectures that have been integrated with GEOPM.
Optimize runtime in power constrained environments, with an expected average improvement of 20%*.	Show runtime performance improvement across multiple benchmarks, proxy applications and codes while maintaining power at or under the system-specified bound. Comparison will be made with naïve uniform static power allocation.
Optimize energy consumption with negligible performance overhead, with an expected savings of 20%*.	Show reduced energy across multiple benchmarks, proxy applications and codes while maintaining application execution time. Comparison will be made with full-energy execution.
Optimize runtime in the absence of power and energy constraints via processor configuration, with an expected average improvement of 10%*.	Show runtime performance improvement across multiple benchmarks, proxy applications and codes by accounting for the number of cores used, hyper-threading, and processor variation.
Demonstrate safe execution under either power or energy constraints.	Using multiple benchmarks, proxy applications and applications, sample instantaneous power and measure total energy, and demonstrate system-specified bounds are not exceeded.

Tapasya Patki, LLNL

** Note that the numeric targets specified in the goals are reasonable expectations, actual savings will depend on the specifics of the underlying processor architectures and application implementations*

ST: Exa-PAPI - The Exascale Performance API

Scope & Intent	R&D Themes	Delivery Process	Target ECP Users	Support Model
Consistent interface and methodology for accessing hardware-based and software-based performance events.	Performance and power monitoring and analysis	Regular releases of software and documentation; Anytime open access to production software from BitBucket	Apps, such as NWChemEx and GAMESS; ST libraries, such as SLATE and PEEKS; ST runtimes such as DTE and Kokkos; ST tools, such as TAU and HPCToolkit.	Ongoing developer support; Dedicated mailing lists for developers and users; Comprehensive web-based documentation.

Goal	Metric
Exa-PAPI is used by performance analysis toolkits and ECP applications to access hardware performance counters on ECP hardware.	The number of packages (independently developed performance analysis tools, apps, libraries and runtimes), which use PAPI to access hardware performance counter on ECP hardware.
Exa-PAPI is used by ECP applications, libraries, and runtimes to expose software-defined events.	The number of packages (independently developed apps, libraries and runtimes), which have integrated software-defined events into their code.
Exa-PAPI is used by the ECP community to access power monitoring metrics and manage power usage.	The number of packages (mainly math libraries and runtimes), which use PAPI power management.

Heike Jagode, UT-Knoxville

ST: Argo - Operating System and Resource Management for Exascale

Scope & Intent	R&D Themes	Delivery Process	Target ECP Users	Support Model
Improve and augment existing OS/R components for use with ECP applications	Hierarchical memory, dynamic power management, containers, internode resource management	Periodic release of software and documentation; testing with vendor software stacks and facilities	CANDLE, PETSc, Flux, QMCPACK, etc. Facilities: improved throughput under constrained power, flexibility of isolation and resource management, and new dynamic power control capability.	Mailing lists, wiki pages, website and issue tracking portals for communication with stakeholders. Making code and documentation available through gitlab

Goal	Metric
Provide a high-quality software that helps ECP applicat ions exploit the new capabilities of ECP systems	The number of ECP applications that have been demonstrated to benefit from Argo components
Provide mechanisms and policies to control the available power dynamically on ECP systems	The number of ECP testbeds that deploy Argo components

Pete Beckman, ANL

PathForward Overview

Objective: Fund R&D to design hardware that meets ECP's Targets for application performance, power efficiency, and resilience

Establish *PathForward* (PF) Hardware Architecture R&D contracts that deliver:

- Conceptual component, node and system designs
- Analysis of performance improvement on conceptual exascale system designs
- Technology demonstrators to quantify performance gains over existing roadmaps
- Support for active industry engagement in ECP holistic co-design efforts

DOE labs engage to:

- Participate in evaluation and review of PathForward deliverables
- Lead Design Space Evaluation through Architectural Analysis, and Abstract Machine Models of PathForward designs for ECP's holistic co-design

PathForward Vendor Contracts

- Competitive RFP released in June, 2016
- 6 proposals selected for SOW negotiations of final Work Packages
- Contract Awards Made Feb-May 2017
- DOE announced 6 contract awards on June 15, 2017
  - Firm Fixed Price contracts with milestone deliverables and payments
  - DOE Advance IP Waivers for vendors that provide ≥ 40% cost share
  - Project duration: 3 years
  - Total DOE Investment: $258M
- Six awardees
  - AMD
  - Cray
  - HPE
  - Intel
  - IBM
  - NVIDIA
- Awarded 31 work packages across all projects
  - 259 total milestones/deliverables

PathForward Summary

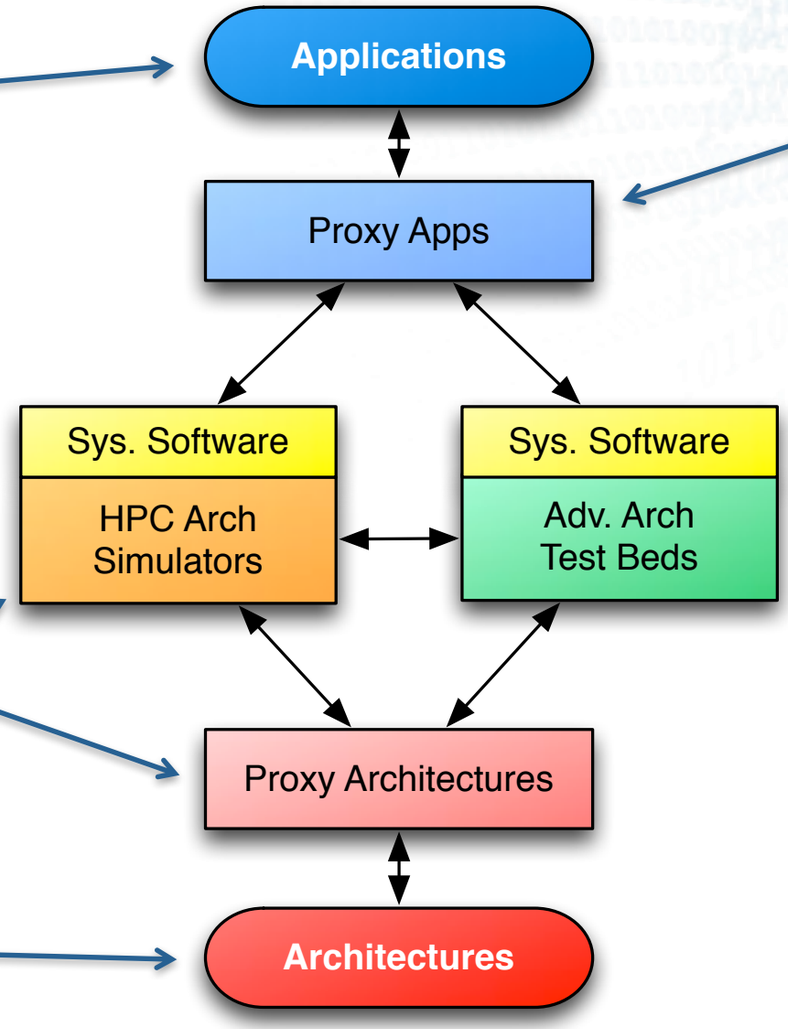
- PathForward reduces the Technical Risk for NRE investments in the 2022/23 capable exascale system(s)
- Establishes a foundation for architectural diversity in the HPC eco-system
- Provides hardware technology expertise and analysis
- Provides an opportunity for inter-agency collaboration
- As of October, 2017, 31 Deliverables Completed
  - 3 application analysis
  - 5 processor design
  - 8 memory
  - 10 interconnect fabric
  - 2 energy and power
  - 3 system design and resilience

Holistic Co-Design

- AD and ST teams should not assume hardware architectures are firmly defined inputs to their project plans
  - AD and ST teams can provide inputs to the PF node and system design projects
- PathForward projects also should not assume that applications and the software stack are fixed inputs to their project plans
  - PF design teams can provide inputs to the ECP AD and ST projects
- *In Holistic Co-Design, each project's output can be another project's input*
  - Individual ECP projects do not operate in a vacuum
  - Assumptions about inputs lead to initial project plans, but the give and take of co-design interactions can lead to changes

Holistic Co-Design: Integration of Focus Area Projects

- ~30 AD projects
- ~50 ST projects
- HI/DSE team
- 6 HI/PathForward projects



- 5 CDC projects + AD projects
- ~50 ST projects + HI/PathForward projects
- Facility Initial Delivery and LCF Systems

EXASCALE COMPUTING PROJECT

National Nuclear Security Administration

U.S. DEPARTMENT OF ENERGY | Office of Science

www.ExascaleProject.org