ECP Overview and Hardware Technology Status

Presented at
SOS-21

James A. Ang, Sandia National Labs
ECP Hardware Technology Director

John Shalf, Lawrence Berkeley National Lab
ECP Hardware Technology Deputy Director

Kongresszentrum Davos, Switzerland
March 21-23, 2017

EXASCALE COMPUTING PROJECT

NNSA National Nuclear Security Administration
U.S. DEPARTMENT OF ENERGY | Office of Science
SAND2017-3134 PE
www.ExascaleProject.org

Exascale Computing Project (ECP) mission and scope

Mission

ECP was established in FY16 with a project office at ORNL to accelerate delivery of a *capable exascale* computing system

Scope

ECP has 4 focus areas:

- Application Development
- Software Technology
- Hardware Technology and Architecture
- Exascale Systems

ECP is committed to workforce development to meet scientific and national security mission needs

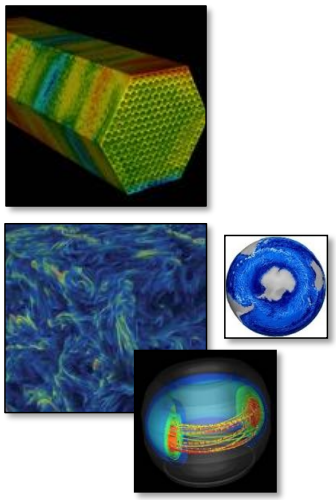
What is a *capable exascale* computing system?

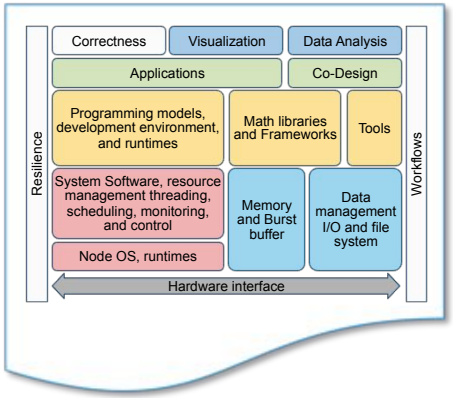
- Delivers 50× the performance of today's 20 PF systems, supporting applications that deliver high-fidelity solutions in less time and address problems of greater complexity
- Operates in a power envelope of 20–30 MW
- Is sufficiently resilient (perceived fault rate: ≤1/ week)
- Includes a software stack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

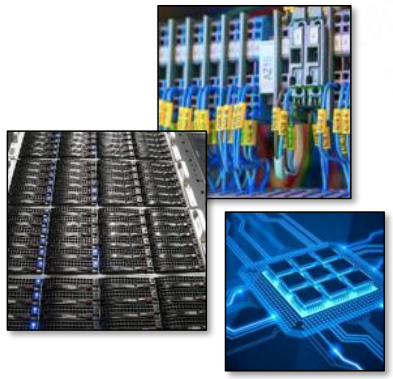
This ecosystem will be developed using a co-design approach to deliver new software, applications, platforms, and computational science capabilities at heretofore unseen scale

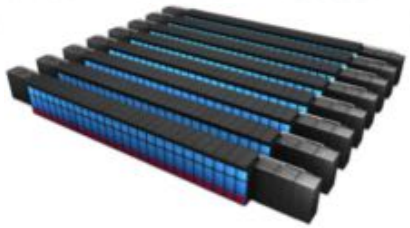
The holistic co-design approach to deliver advanced architecture and capable exascale

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



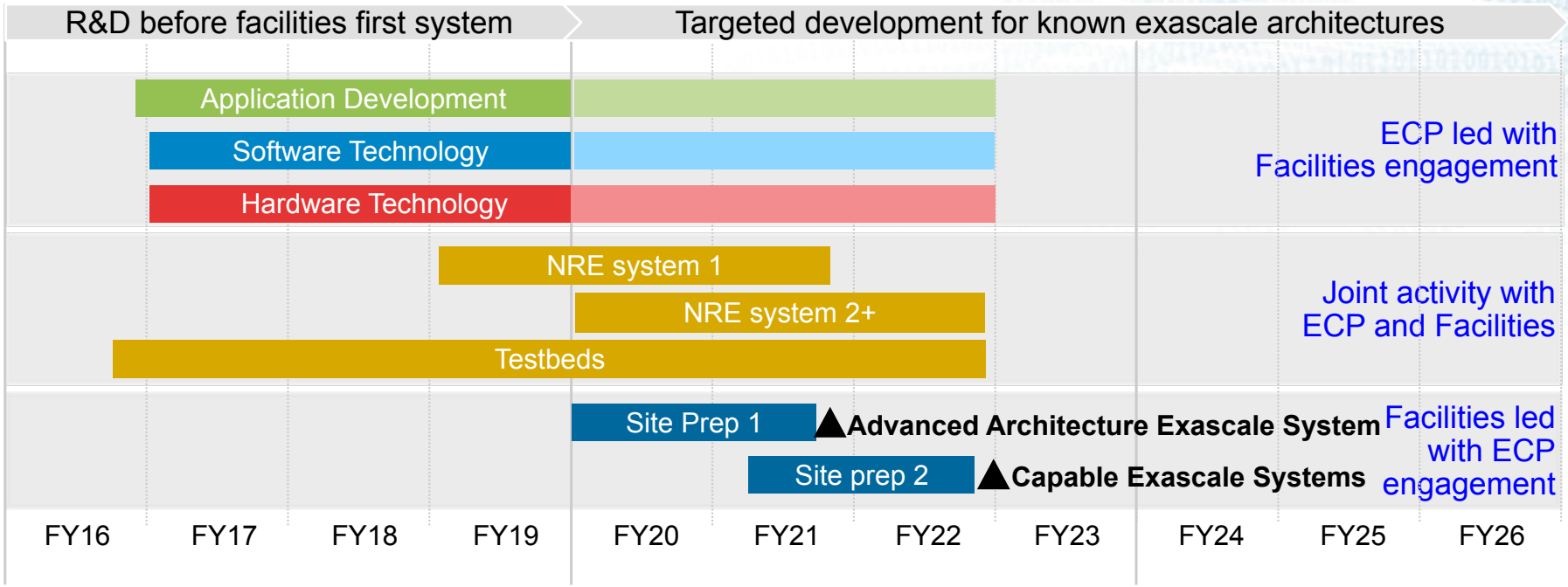




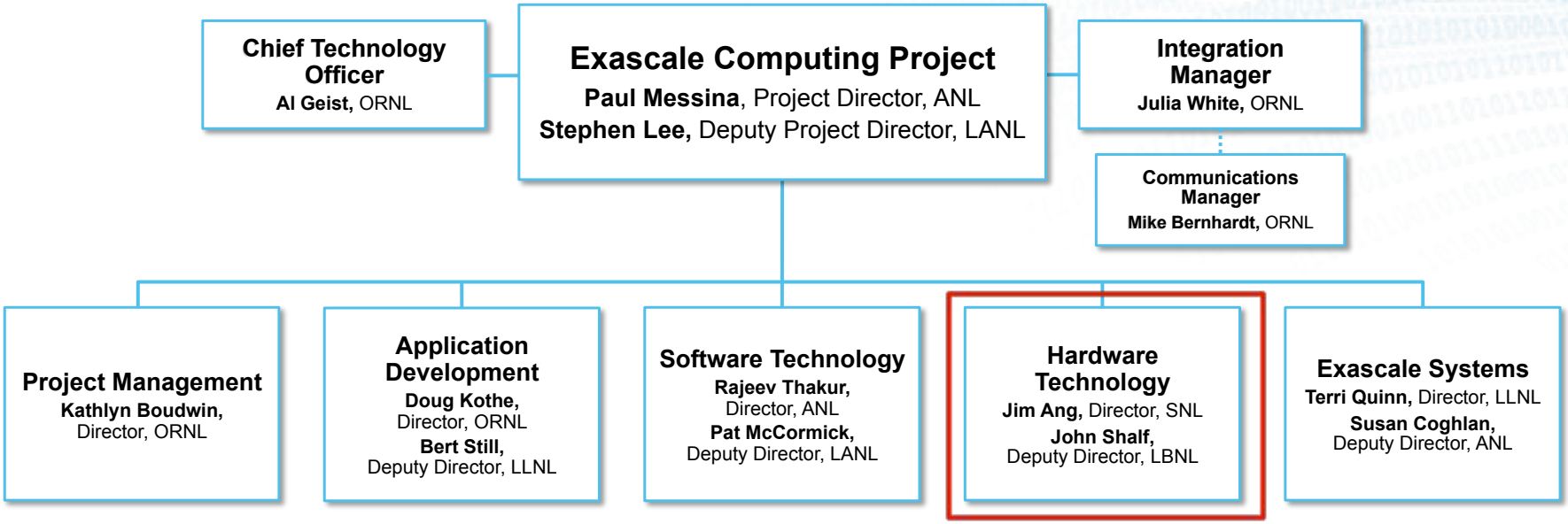


EXASCALE COMPUTING PROJECT

High-level ECP technical project schedule



ECP leadership team



Hardware Technology Focus Area

- Leverage our window of time to support advances in both system and node architectures
- Close gaps in vendor's technology roadmaps or development of new roadmaps to address ECP performance targets *while* affecting and intersecting the Exascale System RFP(s)
- Provide an opportunity for ECP Application Development and Software Technology efforts to *influence* the design of future node and system architecture designs

ECP Hardware Technology mission and scope

Mission

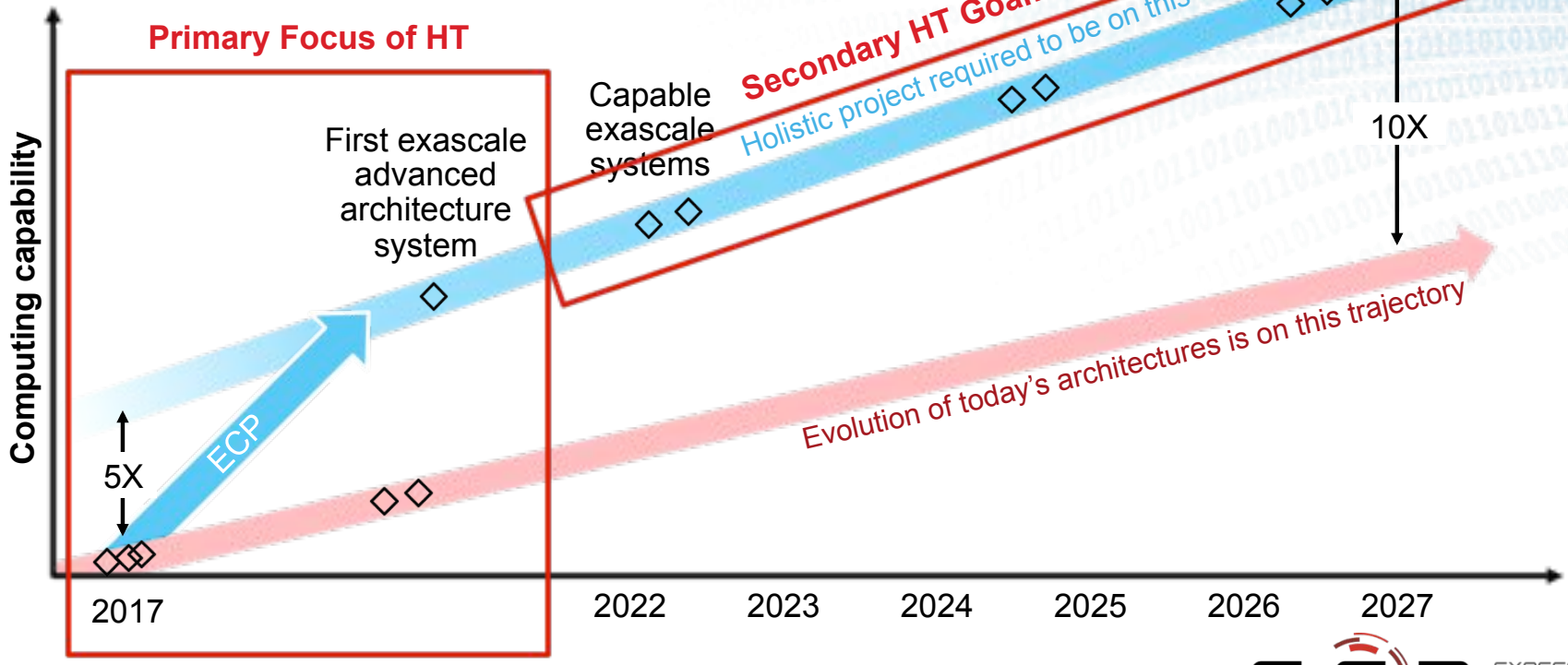
ECP requires Hardware Technology R&D to enhance application and system performance for science, engineering and data-analytics applications on exascale systems

Scope

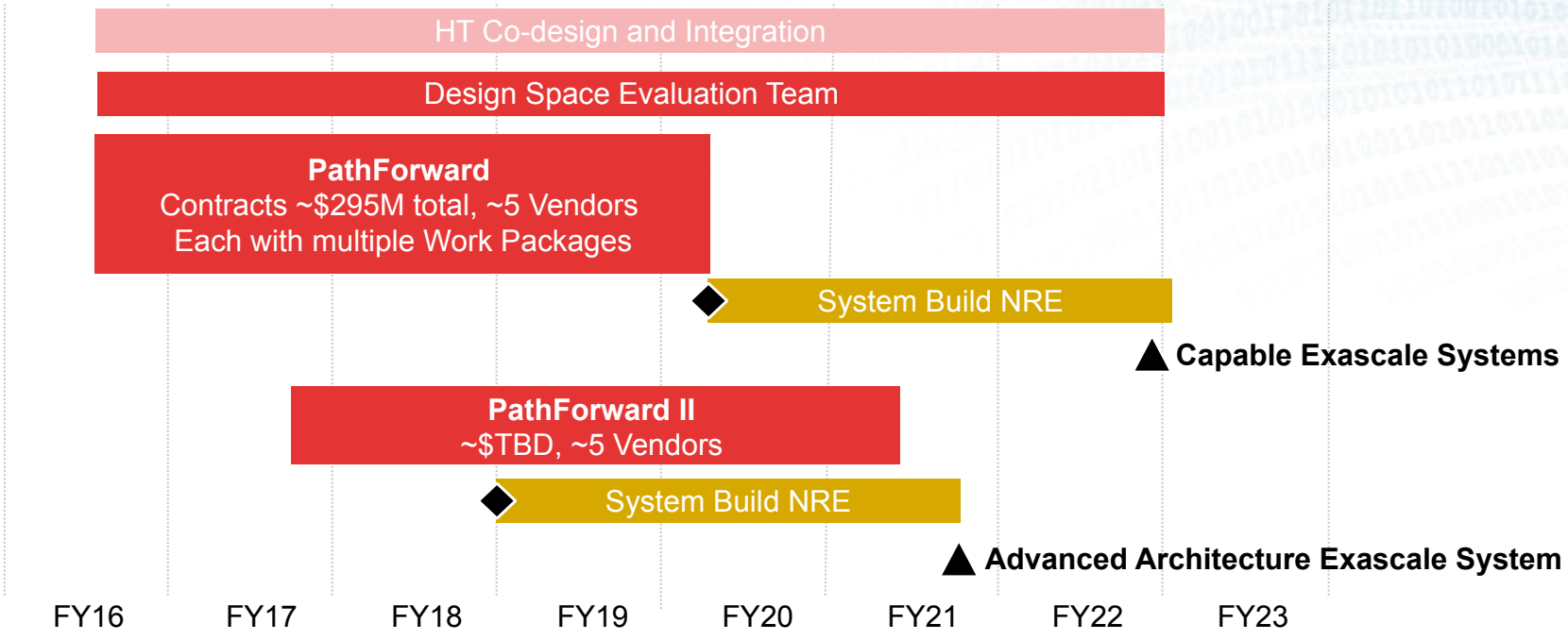
- Support hardware architecture R&D at component, node and system architecture levels
- Prioritize R&D activities that address ECP performance objectives for the initial Exascale System RFPs
- Enable AD, ST, and ES to improve the performance and usability of future HPC hardware platforms

ECP HT provides an opportunity for holistic co-design to develop hardware designs that address ECP technical challenges

Transition to higher trajectory with advanced architecture



HT and ES Timeline:
7-Year Schedule



Hardware Technology Overview

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

Vendor Partnerships for Hardware Architecture R&D contracts that deliver:

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

DOE Design Space Evaluation Team

- Participate in evaluation and review of Vendor HW Architecture R&D deliverables
- Provide Architectural Analysis, and Abstract Machine Models of Vendor designs to support ECP's holistic co-design

PathForward Goals and Objectives

- PathForward supports R&D to quantify the impact of the vendor's advanced architecture concept on ECP technical goals, 50x, 20-30MW, 1 week resilience, application portability
  - Reduce the technical risk for including the subject architecture advances in vendors' Capable Exascale 2022/23 platform bids, and
  - Convince DOE Facility Procurement teams to include technical specifications that call for integration of PathForward designs in the proposed Capable Exascale system architecture
- PathForward leverages Hardware R&D that is in-flight:
  - Prior DOE Architecture R&D: FastForward and DesignForward
  - Other Government Agency Investments
  - Commercial IR&D

PathForward Status

- LLNL is managing the procurement
- Competitive RFP released in June, 2016
- 14 PathForward proposals received in August, 2016
- 6 proposals selected for SOW negotiation of final Work Packages
  - 4 Vendor Contracts have DOE/HQ approval
  - 2 Vendor Contracts are at DOE/HQ or Livermore Field Office for approval
- Firm Fixed Price contracts with milestone deliverables/payments
- DOE Advance IP Waivers for vendors that provide ≥ 40% cost share
- Project duration: 3 years
- Targeting Contract Awards in March 2017

PathForward II: Key Differences from PathForward

- Intersect and enable the initial exascale system based on **advanced architecture** delivered in 2021 and deployed in 2022
- Innovative Hardware Architecture R&D at the component, node, and system level – PathForward II is not focused on accelerating existing product roadmaps, rather it is intended to define new product roadmaps
- Support NSCI Objective #2: *Increasing coherence between the technology base used for modeling and simulation and that used for data analytic computing*
- Our AD and ST projects were not established when we developed the Technical Specifications for the PathForward RFP. *These ECP projects will provide input to the Technical Specifications for the PathForward II RFP*

PathForward II Goals and Objectives

- Improve the competitive offeror responses to the 2018 Advanced Architecture Exascale System RFP for the 2021/22 System
- YES: Starting *later* for a platform that is delivered *earlier* → automatically have a higher risk profile
- PathForward II has a goal to support/encourage high payoff, innovative hardware architecture designs that will likely have higher risk
- Examples of Advanced Architecture Efforts that PathForward II could support
  - Advanced architecture concepts from smaller companies that would not traditionally be considered
  - Advanced architecture concepts that do not start with the existing DOE/NNSA legacy application code base
  - Proposals that support holistic co-design where AD and ST teams identify and propose HT capabilities that support their needs

Design Space Evaluation: Goals and Objectives

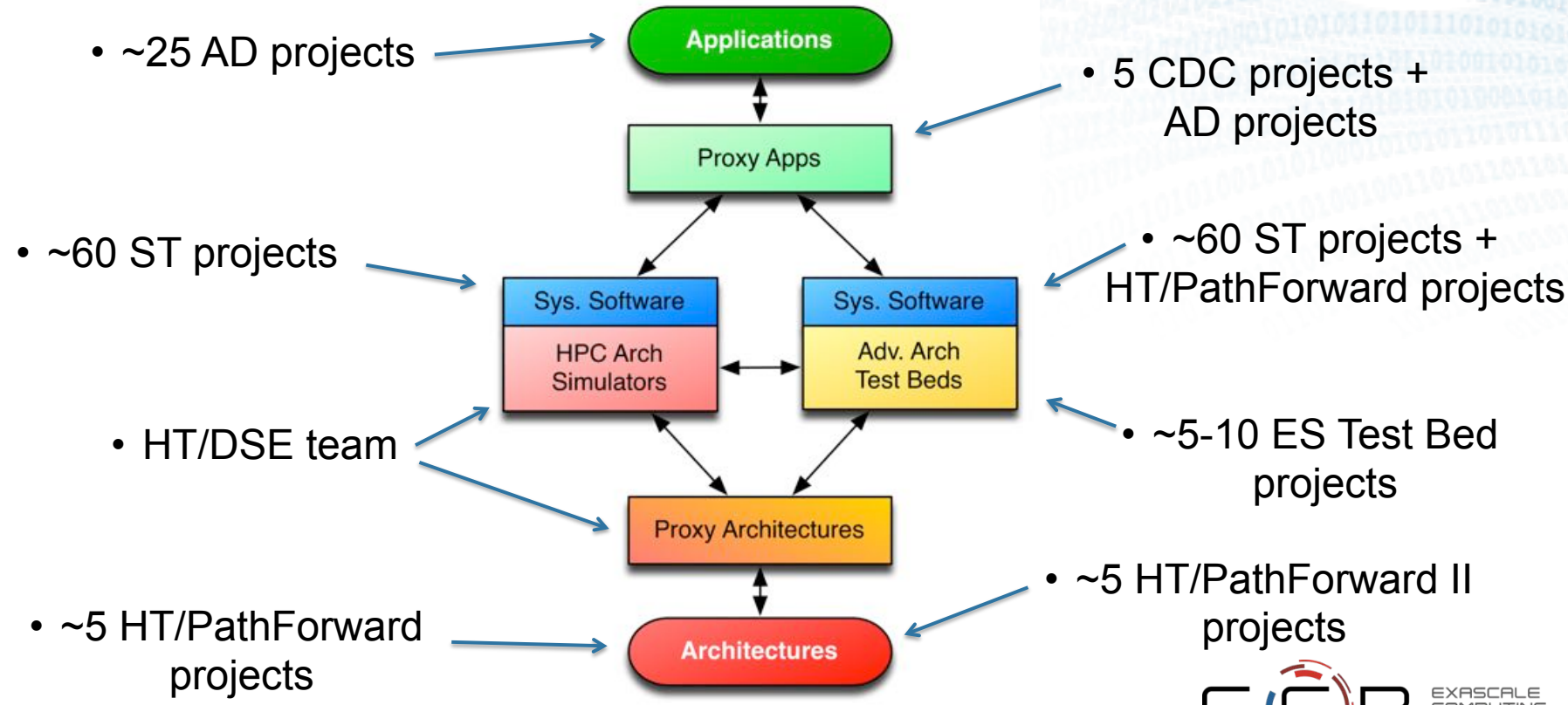
- **Goal:** assemble top experts in hardware architecture analysis and analysis tools at the DOE laboratories and form a single integrated Design Space Evaluation (DSE) team
- **Roles of DSE in ECP & Co-design Interactions**
  - **PathForward I and II (phase 1)**: Stewardship and independent evaluation of PF performers
  - **NRE (phase 2)**: Transition to independent evaluation and assistance for NRE activities
  - **Holistic Co-Design**: Contribute hardware architecture analysis and expertise to ST, AD, and ES activities
  - **Communication**: Distill complex hardware concepts to simpler abstractions to facilitate communication among ST, AD, and ES projects
  - **Tools**: Put existing DOE analysis and architectural simulation tools to work for ECP *(recognizes need to extend existing tools to meet emerging requirements)*

The Importance of Hardware Architecture Models and Simulation for Emerging Hardware

- Changing Hardware is very expensive and takes a long lead time
- Changing Software (rewriting our codes) is also very expensive and takes a long lead time
- Co-design uses quantitative architectural analysis because with cost and power constraints, we need to understand the system design trade-space
  - Easy to ask for more features to be added to the machine
  - It is much harder to evaluate what you are willing to give up
  - particularly when the cost functions are highly non-linear and machines do not yet exist *(need models)*
  - Risk mitigation for expensive decisions

Need an Evidence-based approach to enable rational and thoughtful decision making (for both HW and SW)

Holistic Co-Design: Integration of Focus Area Projects



Hardware Technology Summary

- PathForward *reduces* the Technical Risk for NRE investments in the 2022/23 capable exascale system(s)
- PathForward II *accepts* Technical Risk in return for game-changing impact on the U.S. computing eco-system and potential for early deployment in 2021/22 exascale system
- Establishes a foundation for architectural diversity in the HPC eco-system
- Design Space Evaluation applies DOE hardware technology expertise and analysis to support ECP's holistic co-design opportunities
- Provides an opportunity for inter-agency collaboration under NSCI

Discussion

Alvaneu ~25km SW of Davos

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

www.ExascaleProject.org