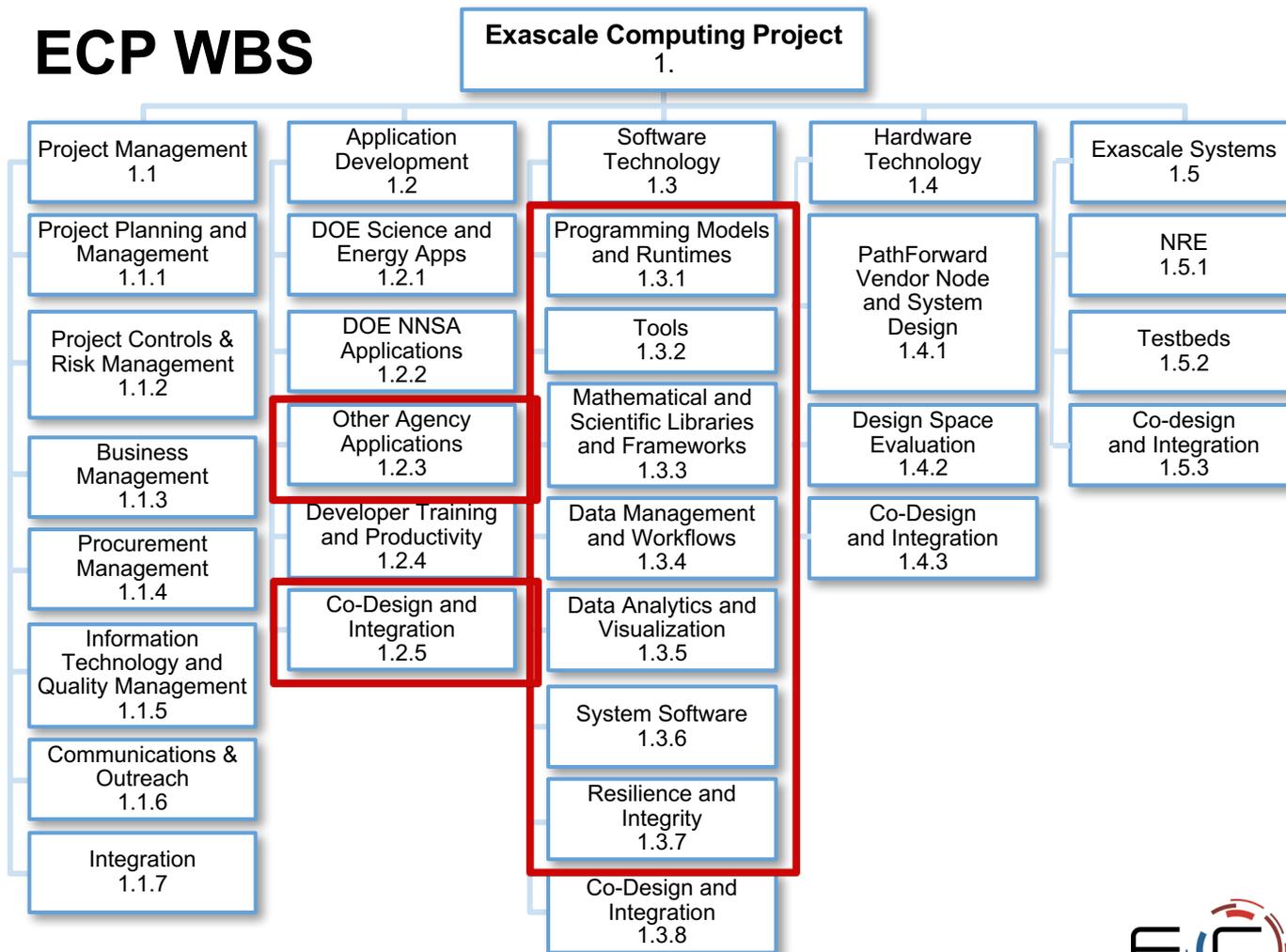


U.S. OVERVIEW/UPDATE: EXASCALE PROJECT HAPPENINGS

ROB ROSS

Mathematics and Computer Science Division
Argonne National Laboratory
ross@mcs.anl.gov

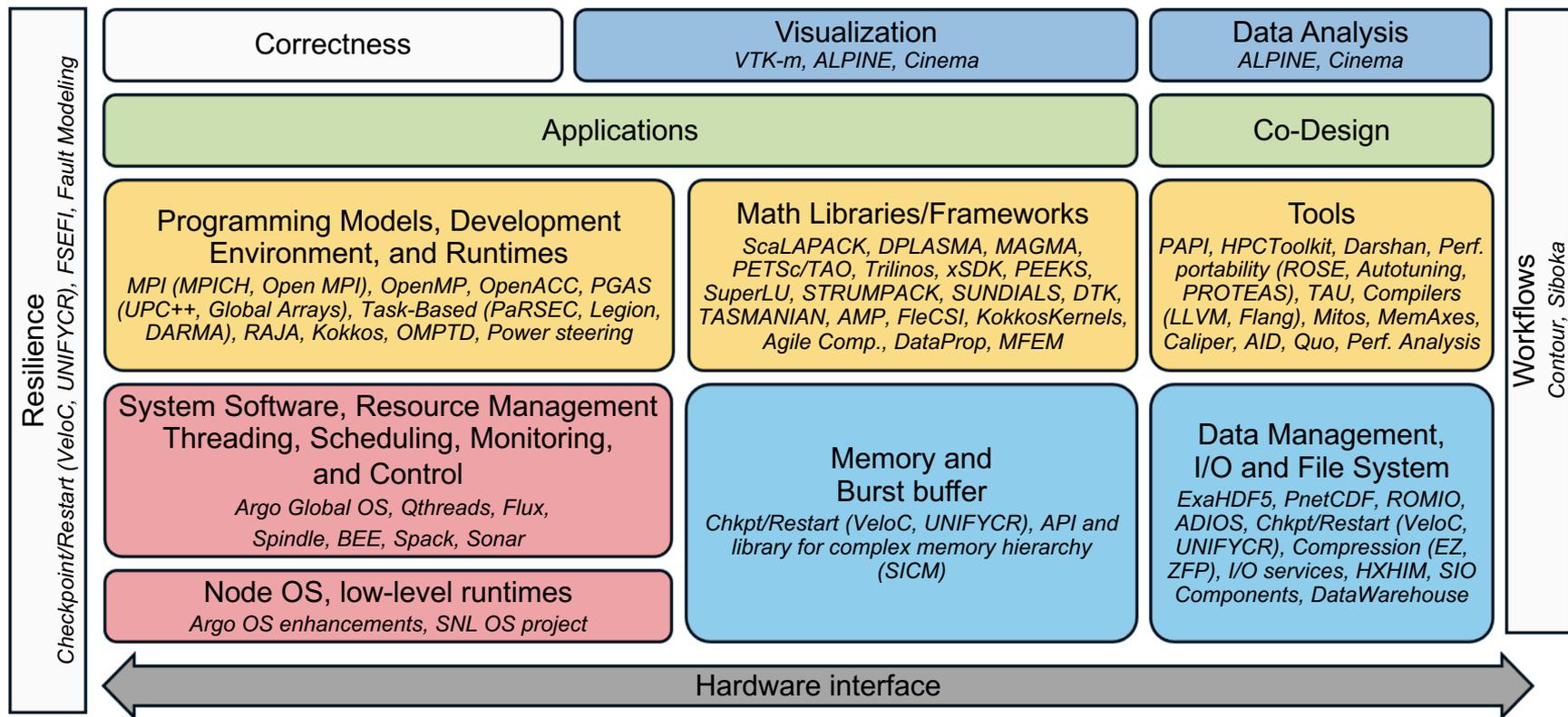
ECP WBS



ECP Software Technology Overview

- 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
- Accomplished by extending current technologies to exascale where possible, performing R&D required to conceive of new approaches where necessary
 - Coordinate with vendor efforts; i.e., develop software other than what is typically done by vendors, develop common interfaces or services
 - Develop and deploy high-quality and robust software products

ST Projects (incl. ATDM) Mapped to Software Stack



The Kinds of Software Technologies ECP Apps Ask For...

Tools (debuggers, profilers, software development, compilers)

- LLVM/Clang, HPCToolkit, PAPI, ROSE, Oxbow, JIRA, Travis
- ASPEN (machine modeling), CMake, git, TAU, Caliper, GitLab, CDash, Flux, Spack, Docker, Shifter, ESGF, Gerrit
- GDB, Valgrind, GitHub, Jenkins (testing), DDT (debugger)

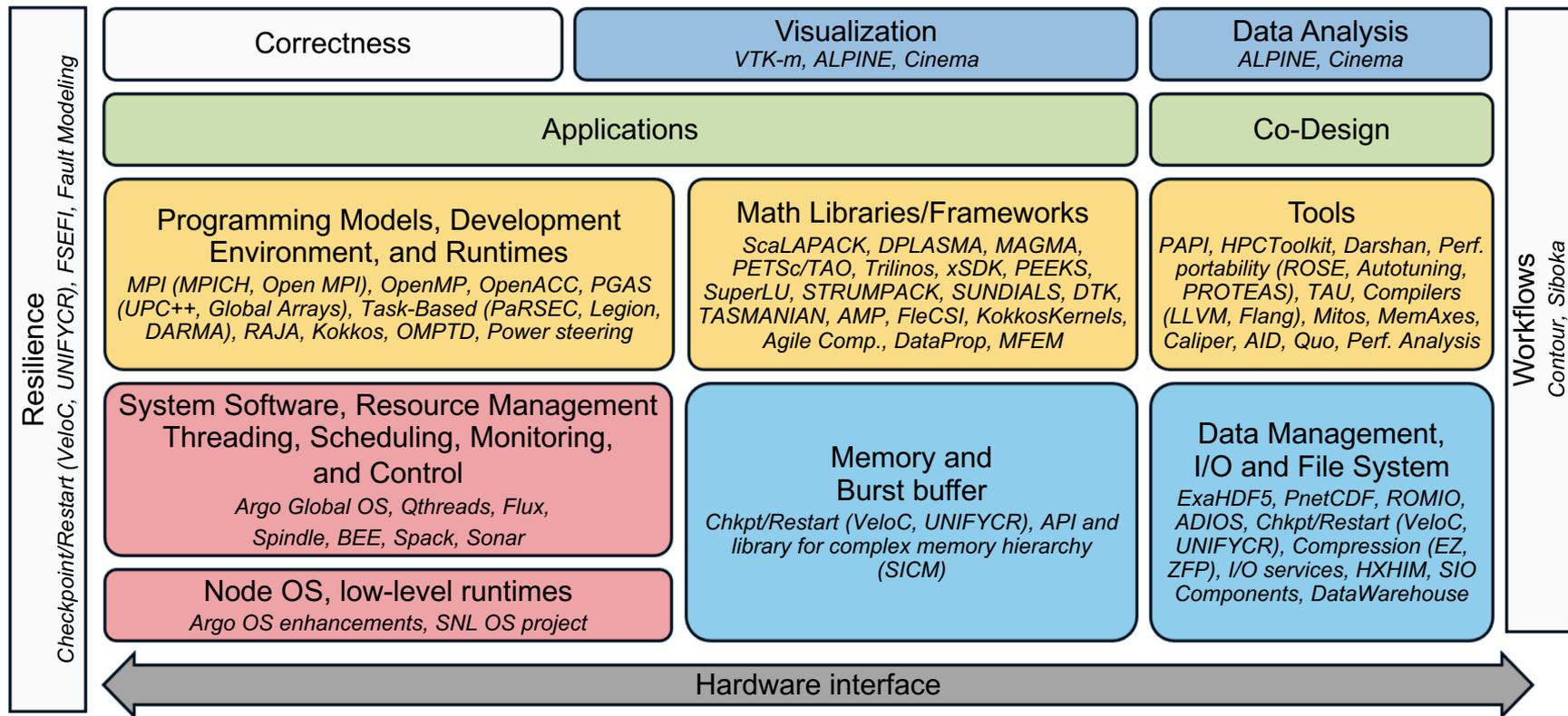
Data Management and Workflows

- 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

ST Projects (incl. ATDM) Mapped to Software Stack



Data Management and Workflow Technology Areas

(Dominant project themes)

- **I/O Interfaces and Data Services**

- The ADIOS framework for Scientific Data on Exascale Systems, Klasky (ORNL)
- ExaHDF5: Delivering Efficient Parallel I/O on Exascale Computing Systems, Byna (LBNL)
- Data Libraries and Services Enabling Exascale Science, Ross (ANL)
- HXHIM Key-value Stores for Applications, Settlemyer (LANL)
- Scalable I/O Components, Ulmer (SNL)

- **Checkpoint/Restart**

- UnifyCR, Mohror (LLNL)
- VeloC: Low Overhead transparent multilevel Checkpoint/restart, Cappello (ANL)

- **Compression**

- ZFP: Compressed Floating-Point Arrays, Lindstrom (LLNL)
- EZ: Exascale lossy compression for scientific data, Cappello (ANL)

- **Workflow**

- Workflow Infrastructure for Increasing User Productivity, Laney (LLNL)
- DataWarehouse: Asynchronous Data Management Layer, Ulmer (SNL)

Data Analysis and Visualization Technology Areas

(Dominant project themes)

- **In Situ Data Reduction**

- ALPINE, Ahrens (LANL)
- Cinema, Ahrens (LANL)

- **Infrastructure Improvement**

- Scalable Visualization, Moreland (SNL)
- VTK-m, Moreland (SNL)

Related Technologies in Other Software Areas

(Dominant project themes)

- **Virtualization and Scheduling**

- BEE: Virtual Environments for Applications, Sewell (LANL)
- Operating System, Pedretti (SNL)
- Argo: OS and Resource Management for Exascale, Beckman (ANL)
- Flux Resource Manager, Ahn (LLNL)

- **Programming and Runtime**

- Distributed Tasking for Exascale, Dongarra (UT Knoxville)
- Enhancing and Hardening Legion for ECP, Shipman (LANL)
- DARMA: Asynchronous Many-Task (AMT) Abstraction Layer, Bennett (SNL)

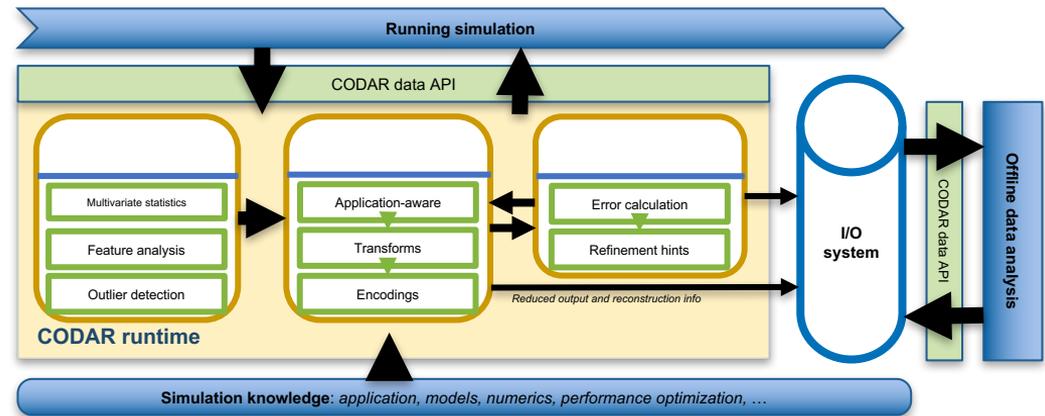
- **Software Package Management**

- Spack Package Manager, Gamblin (LLNL)

CODAR: A Co-Design Center for Online Data Analysis and Reduction at the Exascale

Ian Foster (ANL)

- Address growing disparity between computation and I/O rates
 - Deliver high-performance products for data analysis and reduction
 - Quantify performance tradeoffs for offline vs. online analyses
 - Orchestrate online data analysis and reduction
- Partnership among Argonne, Brookhaven, Oak Ridge, Rutgers, Stony Brook, and Brown University



ExaGraph: Combinatorial Methods for Enabling Exascale Applications Co-Design Center

Mahantesh Halappanavar (PNNL)

- Focus on graph analytics – combinatorial (graph) kernels that play a crucial enabling role in many data analytic computing (DAC) application areas as well as several ECP applications
 - Graph matching
 - Graph coloring
 - Graph clustering: including clique enumeration, parallel branch-and-bound, and graph partitioning
- Partnership among PNNL, Berkeley, Sandia, and Purdue University
- Announced March 2!

THANKS!