SAND2014-19592PE

Exceptional service in the national interest

Sandia National Laboratories

The Sandia Perspective on ATDM

Bill Rider for the ATDM Team

November 6, 2014

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

ATDM @ Sandia in one slide

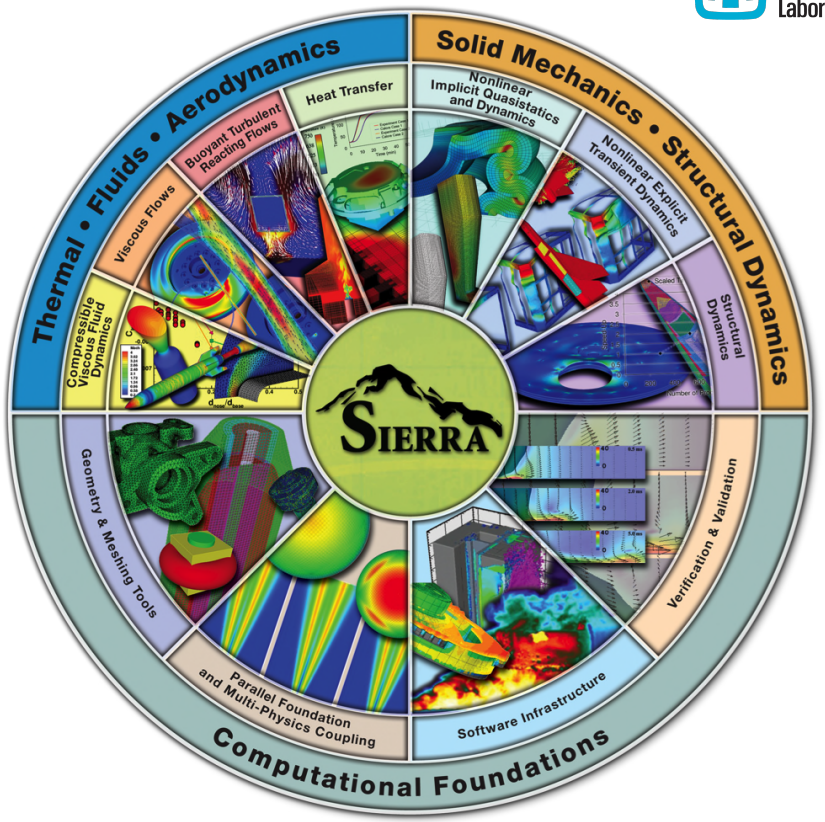
Sandia National Laboratories

Enabled by Next Generation Platforms

- **Software Engineering**
  - Develop single new code-base suitable for future ASC applications
  - Component software design for agility, reuse, and reduced cost
- **Performance**
  - Performance portability through data and loop abstractions
  - Fundamental support for dynamic, asynchronous multi-task
  - Data-management component for analytics, I/O, data-movement
- **Computational Science**
  - Embedded geometry and meshing for full-system models
  - Automatic derivatives, Jacobians, adjoints, etc.
  - Embedded UQ, optimization, and analytics
  - Flexible, robust and dynamic multiscale & multiphysics

ATS + ATDM enables dramatic new design/decision-support for DoE applications

ASC today



- **SIERRA:** thermal, fluids, aero solid mechanics, structural dynamics, materials models, failure models, …
- **RAMSES:** radiation transport, electromagnetics, device modeling, electrical circuits, plasmas and MHD, …
- **Science codes:** hypersonics, molecular dynamics, plasmas, DFT, materials, …
- **Very diverse:** physics, models, scales, geometries, problem specifications, workflows, etc…

ATDM: achieve greater commonality, reuse, agility

Software productivity strategy…

- **Math & Algorithms Research**
- **Agile Components infrastructure**
- **Algorithms Products**
  - Trilinos – general algorithms toolkit
  - Dakota – UQ and optimization
- **Applications & Technologies**
  - Albany – components proving ground
  - Peridigm – fracture mechanics
  - ALEGRA – multiphysics hydrodynamics
  - Drekar – multiphysics MHD
  - LAMMPS – molecular dynamics
  - Sierra– computational mechanics
  - Xyce – electric circuit modeling
  - Charon – semi-conductor drift diffusion
  - Aleph – multiphysics plasma dynamics
  - Sparta – DSMC hypersonics
  - LCM – structural mechanics

Libs

ASCR

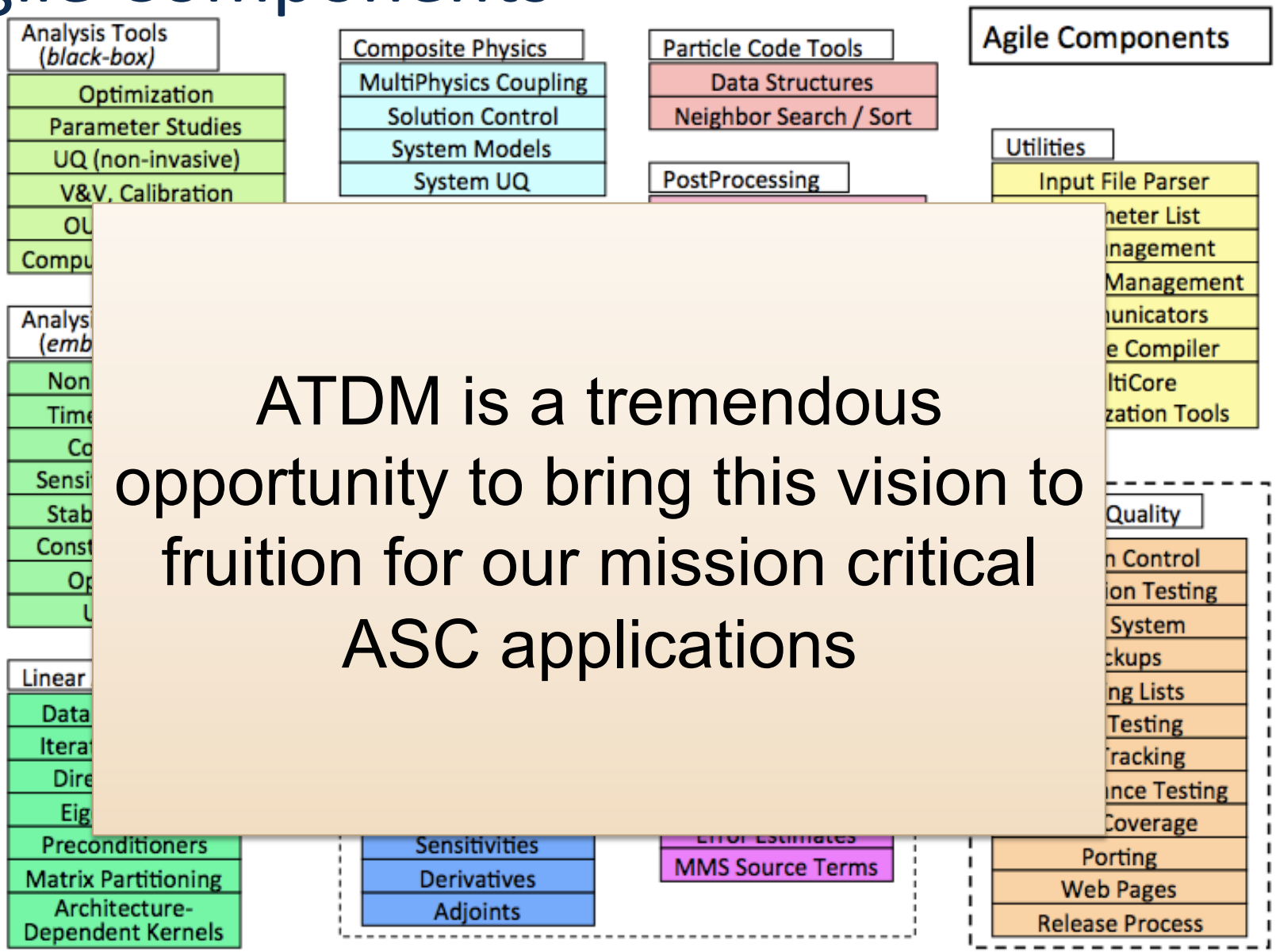
Other

Sierra

Ramses

ASC Science

Agile Components



ATDM is a tremendous opportunity to bring this vision to fruition for our mission critical ASC applications

Properties of ATDM

Sandia National Laboratories

- Uniquely enabled by ATS-3+ computing systems
  - Embrace the dynamic, heterogeneity of future systems
  - Predictive science requires higher fidelity models
  - Multi-physics coupling drives to full-system models
- Shift the analysis paradigm
  - Make geometry and mesh part of the solution
  - Embed uncertainties within the simulation – not as an outer loop
  - Fundamentally change V&V practice through automated calibration, parameter estimation and inversion
- Enable Forward Looking Analysis and Decision Support
  - Design optimization virtual prototypes at component and system level
  - Design for manufacturing, lifecycle, simulation, …
  - Agility and cost reduction by reducing design/test cycles

Algorithms & Solvers (1/2)

Sandia National Laboratories

- **Algorithm Components:** interoperable software components providing unique capabilities exploiting NPG
  - Discretizations: high-order (SE, DG), XFEM, meshless, PIC, …
  - Solvers: AMG for above supporting multi-level parallelism
  - Advanced multiscale/multiphysics coupling (task-based)
  - Hybrid implicit/explicit (IMEX) time integration
- **Analysis Components:** Interoperable software components that enable advanced analysis & workflows on NGP
  - Embedded optimization and UQ
  - Embedded geometry & meshing for robust shape and topological opt
  - Automatic derivative, Jacobian, adjoint, Hessian-vector products, …
  - Embedded analysis: consistent QoIs, data compression, compressed sensing, reduced-order models

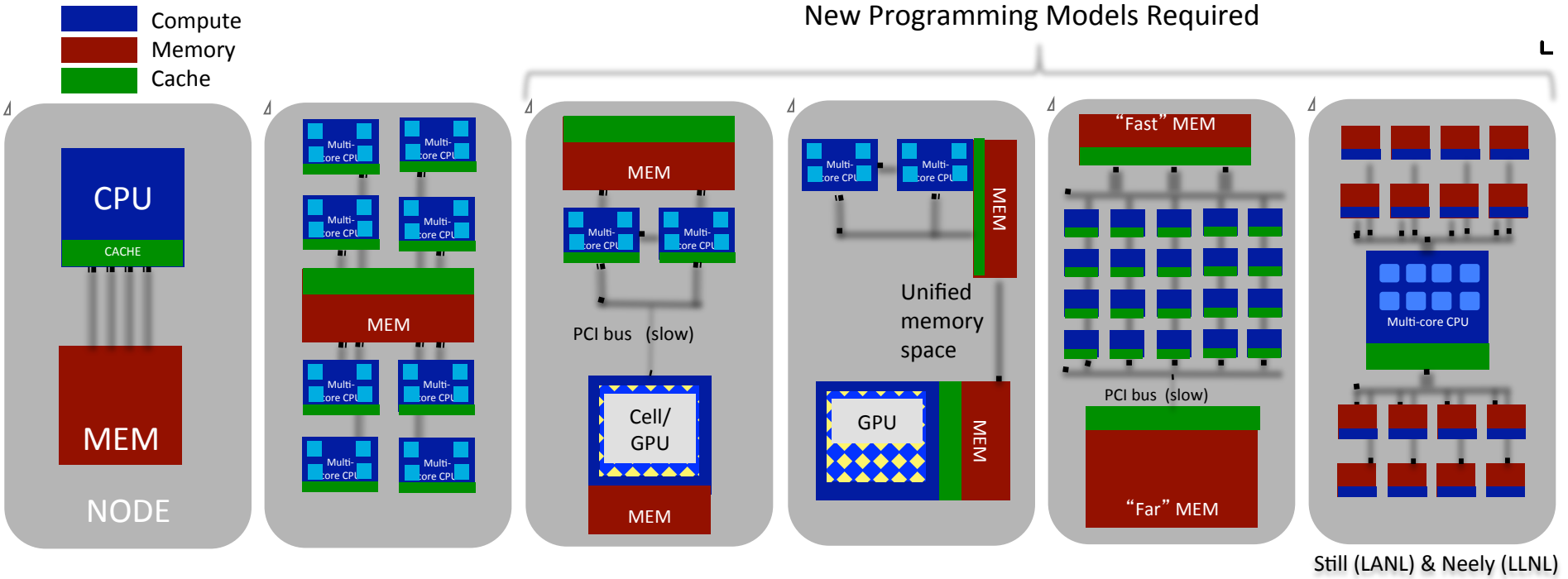
Algorithms & Solvers (2/2)

Sandia National Laboratories

- **Application Components:** interoperable software components/interfaces for modeling physical systems
  - *Mesh component*: meshDB, adaptivity, in-line meshing, geometry refinement, load-balance, …
  - *Particle component*: particleDB, load-balance, collisions, search, …
  - *Solution component*: DOF manager, time-integration, task management and evaluators, interfaces for BCs and ICs
  - *Coupler component*: conservative solution transfer, management of multiscale models and evaluators, global search
  - *Analysis component:* infrastructure for optimization and UQ
- **Physics specific components**
  - Specific element types
  - Specific PIC kernels
  - Contact enforcement
  - Physics evaluators

Test Beds

- Require coverage of expected architectures (see below)
- Most test beds are currently on external networks…
- Must move beyond “few node” systems up to 100’s – 1000’s of nodes to effectively test dynamic task-parallelism



Proxy-Applications

Sandia National Laboratories

Status

- **miniFE & HPCG**: representative of solver dominated implicit codes. Not very representative of the finite element physics and assembly
- **miniAero:** representation of explicit mechanics but does not capture complexities such as contact, special elements and material models
- **miniContact:** represents global search but needs maturation
- **miniPIC:** already started development under ATDM

Needs

- Multiphysics/multiscale coupling & solution transfers
- General finite element (spectral, DG, XFEM) assembly
- Proxy applications with task-based parallelism

Programming Models

Sandia National Laboratories

- Performance portability through data and loop abstractions
  - Accelerate the maturation and production hardening of Kokkos
  - Design higher level abstractions (Field, Particle, Coupler, etc) to ease transition for application programmers
  - Extend Kokkos for processing-element (PE) task and data-parallelism
  - Enable developer training and adoption by current ICs
- Fundamental support for dynamic, asynchronous multi-task
  - Asynchronously schedule tasks to heterogeneous PEs on- or off-node
  - Enable through embedded domain-specific language (i.e. template meta-programming) building, e.g. on Panzer and Uintah technologies
- Data-management for analytics, I/O, data-movement
  - Enhance Nessie/Kelpie for an in-system database that enables fast, flexible storage and data sharing between application tasks

Levels of parallelism (1/2)

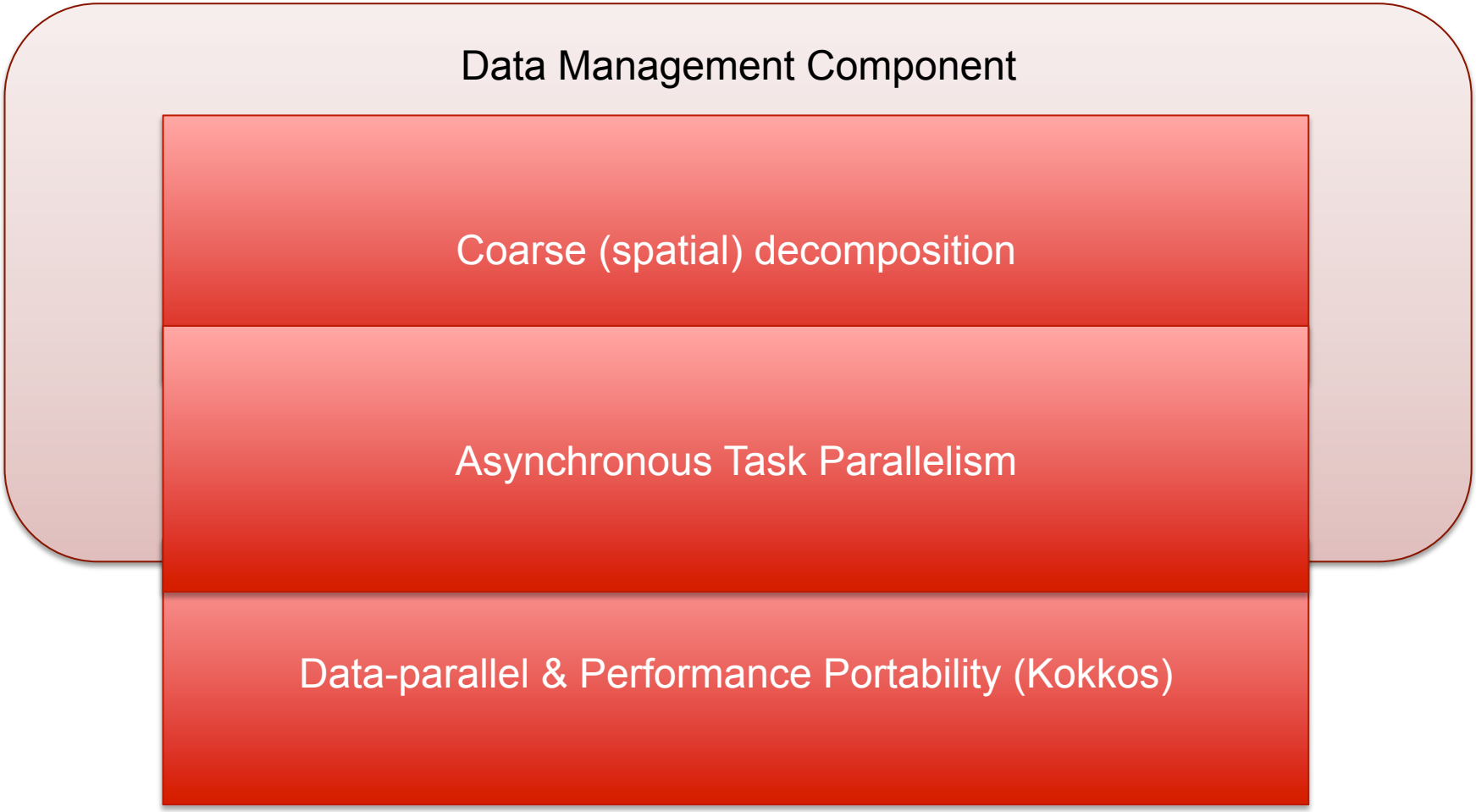
Sandia National Laboratories

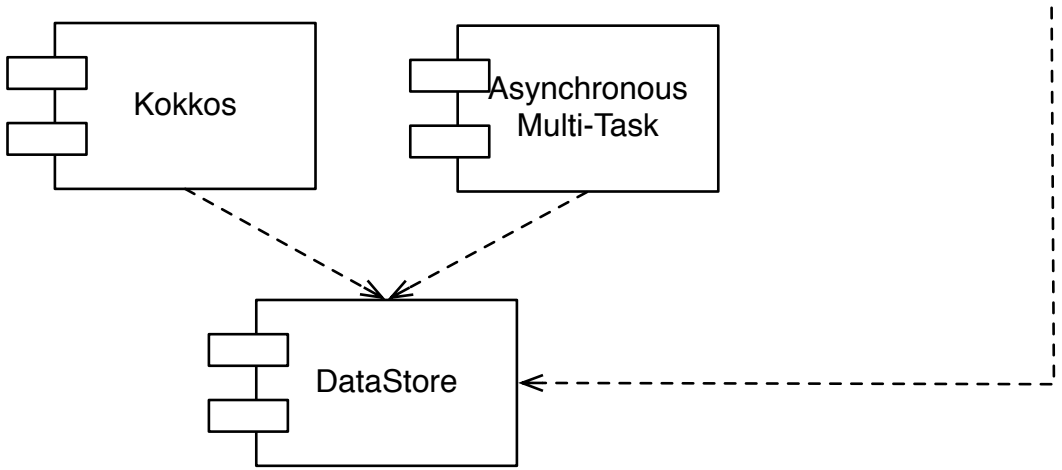
Coarse (spatial) decomposition (MPI)

Asynchronous Task Parallelism

Data-parallel & Performance Portability (Kokkos)

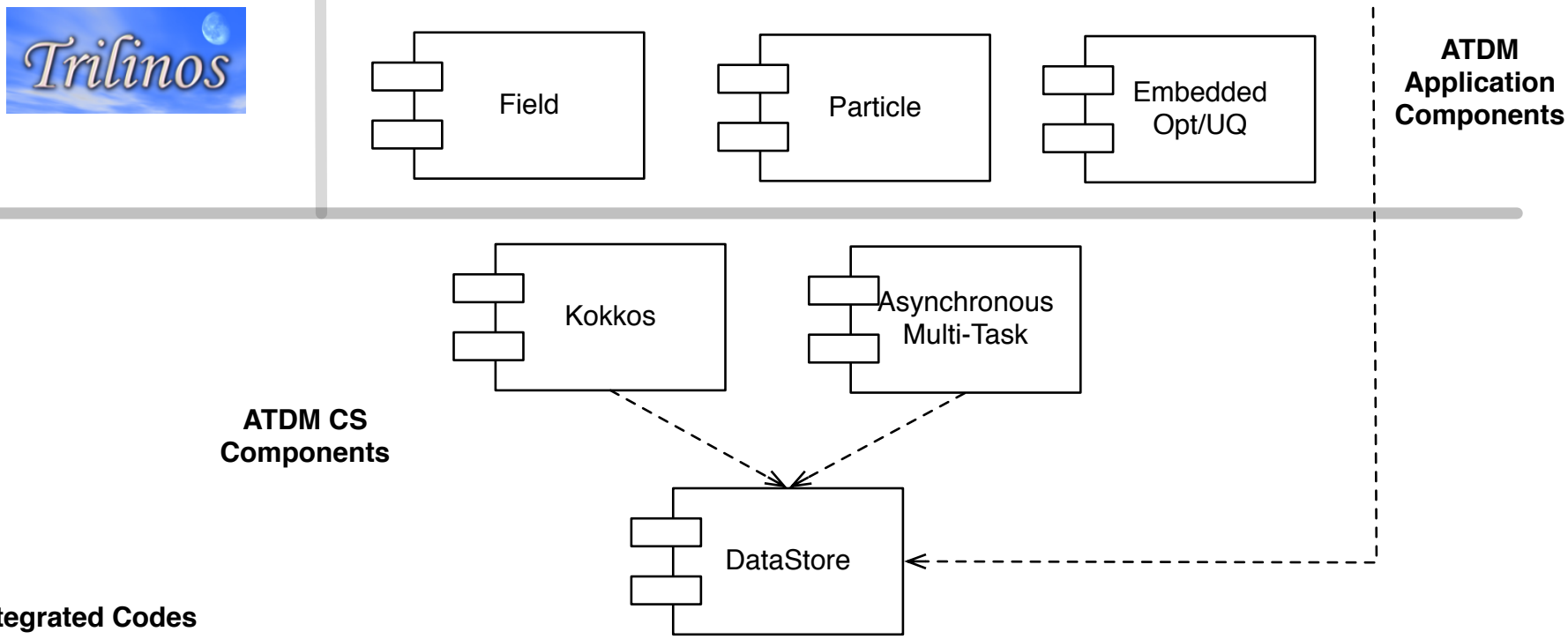
Levels of parallelism (2/2)

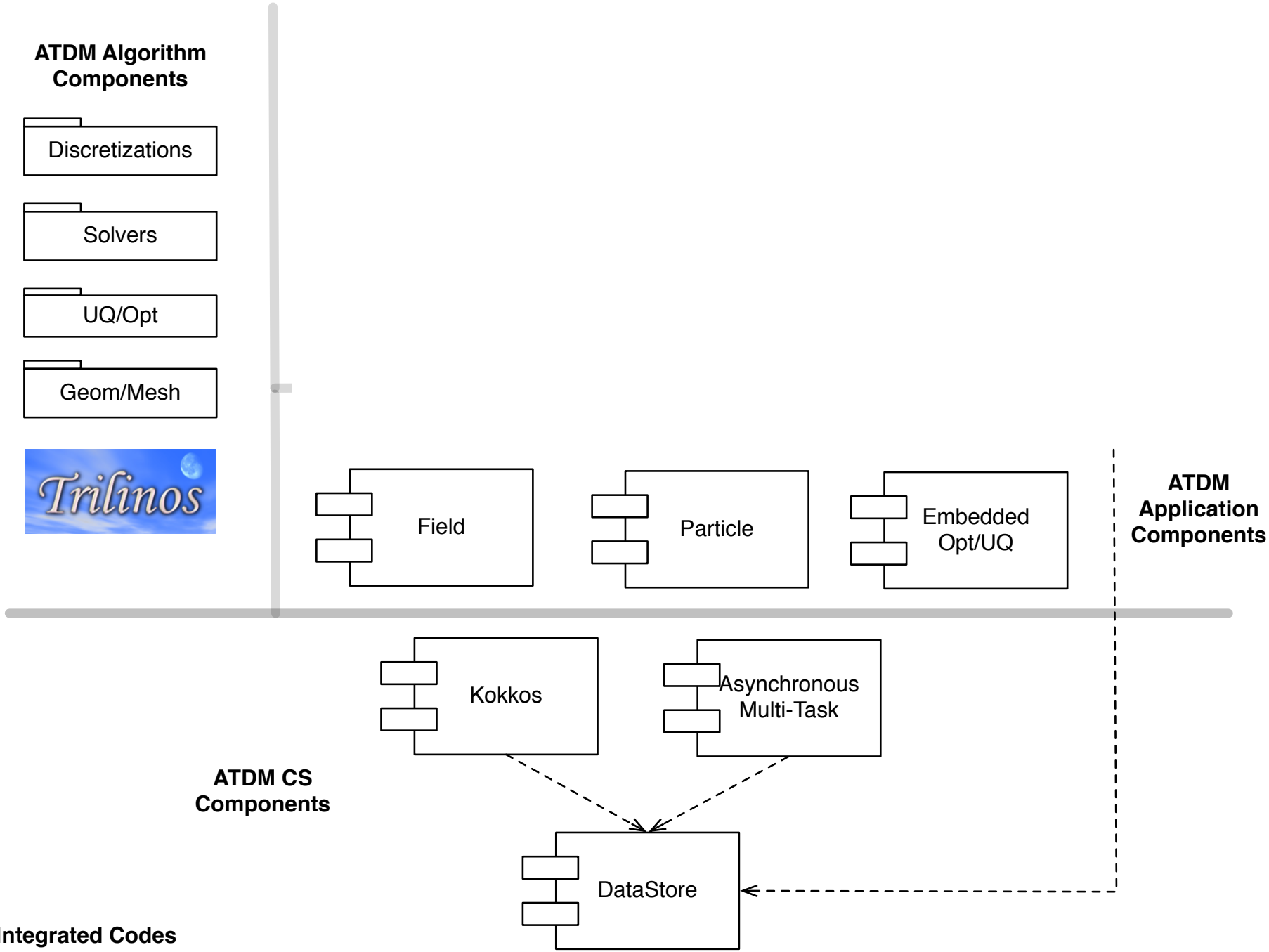


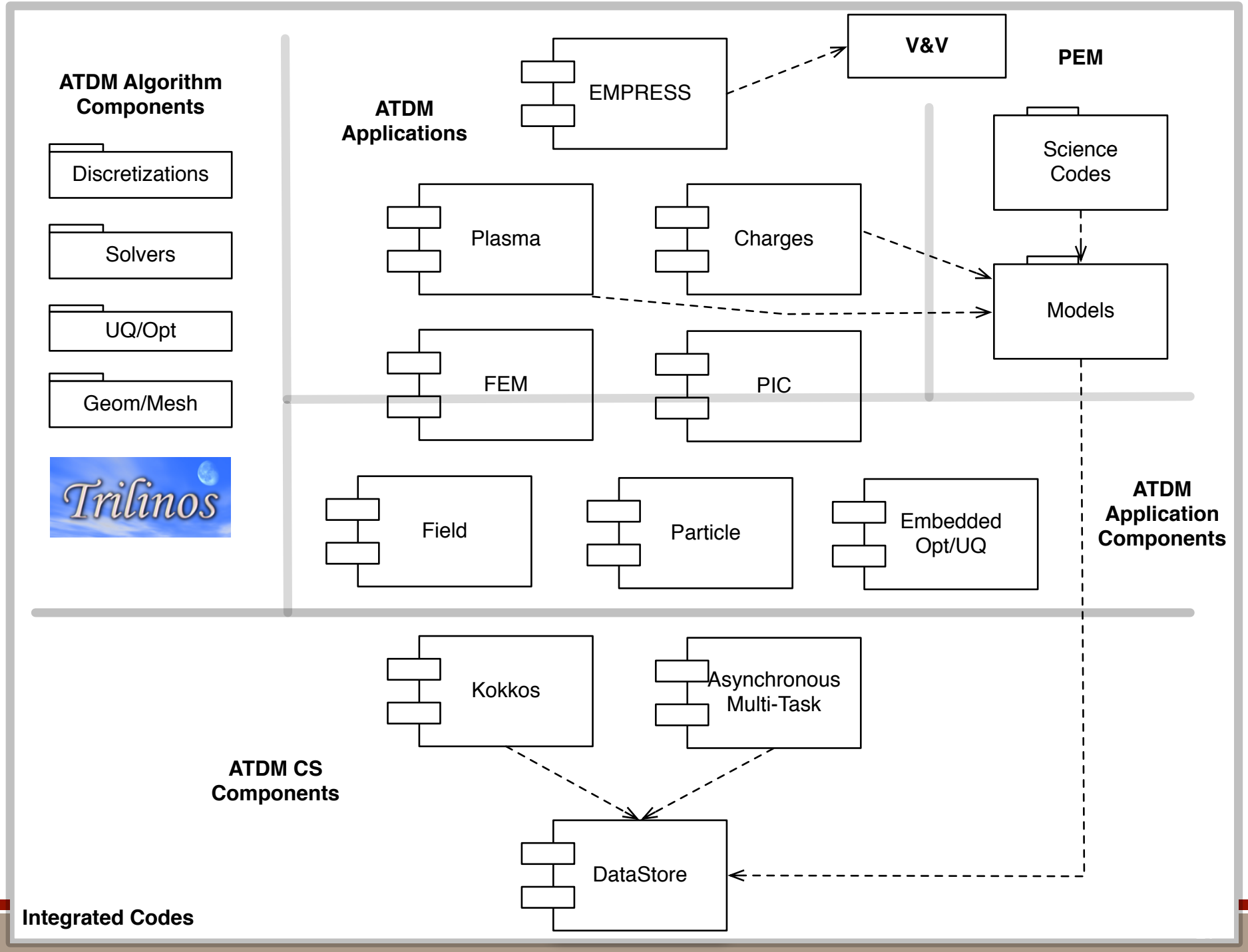


ATDM CS Components

Integrated Codes


Trilinos
Field
Particle
Embedded
Opt/UQ
ATDM
Application
Components
Kokkos
Asynchronous
Multi-Task
ATDM CS
Components
DataStore
Integrated Codes


ATDM Algorithm Components
Discretizations
Solvers
UQ/Opt
Geom/Mesh
Trilinos
Field
Particle
Embedded Opt/UQ
ATDM Application Components
Kokkos
Asynchronous Multi-Task
ATDM CS Components
DataStore
Integrated Codes


ATDM Algorithm Components
Discretizations
Solvers
UQ/Opt
Geom/Mesh
Trilinos
ATDM Applications
EMPRESS
V&V
PEM
Science Codes
Plasma
Charges
Models
FEM
PIC
Field
Particle
Embedded Opt/UQ
ATDM Application Components
Kokkos
Asynchronous Multi-Task
ATDM CS Components
DataStore
Integrated Codes

Code Development Requirements

Sandia National Laboratories

- **Development Workflow**
  - Heavily leverage existing agile components approach
  - Single Git repository for all ATDM code with tight co-dependencies
  - Cmake configuration & build
  - External dependencies managed through superbuild (TriBits) & continuous integration (Jenkins)
  - Automated unit, regression, & performance tests (Ctest)
  - Dashboard to report continuous integration status (Cdash)
  - Code review (Gerrit)
  - Issue tracking and scheduling (Trac)
- **Compilers, debuggers, profilers, etc.**
  - ICC, CLANG, GCC, Eclipse, …
  - Debuggers/profilers a major area of concern…

Specific Technologies

Sandia National Laboratories

Largely driven by FastForward2/DesignForward2 – so we must continue to engage

- **Burst buffers**: Data-management component and embedded analysis will directly leverage
- **Heterogeneous processing elements**: multi-core, many-core, GPU, PIM, FPGA – Kokkos + multi-task
- **Complex memory layout:** data-management component
- **Vectorization:** higher-order for more flops, but…
- Would like hardware support for:
  - gather-scatter,
  - thread synchronization,
  - atomics

In five-years

Sandia National Laboratories

- Core components validated on *production prototypes* that are solving mission relevant use-cases
- A *near production* version of core components
- Evidence to support decisions regarding path forward by demonstrating:
  - performance portability on representative testbeds and ATS-?
  - dynamic algorithms on heterogenous processing elements using asynchronous multi-task
  - extreme scalability
  - enhanced predictivity via tight multiphysics coupling and multiscale
  - Improved workflow via embedded UQ, optimization and analytics
- The foundation and a roadmap for expansion to cover broader mission requirements

Business case

Sandia National Laboratories

- Current workflows require 1000's of simulations for UQ/QMU
  - New code poses UQ problem directly and allows math and CS innovations and ATS computers to achieve speed-up
- Design and qualification that traditionally require many expensive cycles of design/build/test
  - New codes embed optimization, geometry, and meshing to enable rapid virtual prototyping to reduce design cycles (goal is one)
  - Embedded optimization also key to design of qualification experiments
- Predictive failure and modeling of dynamic sub-grid effects are grand challenges
  - By embracing dynamic runtime we simultaneously enable next-generation multiscale while providing agility of the SW to HW changes and resiliency

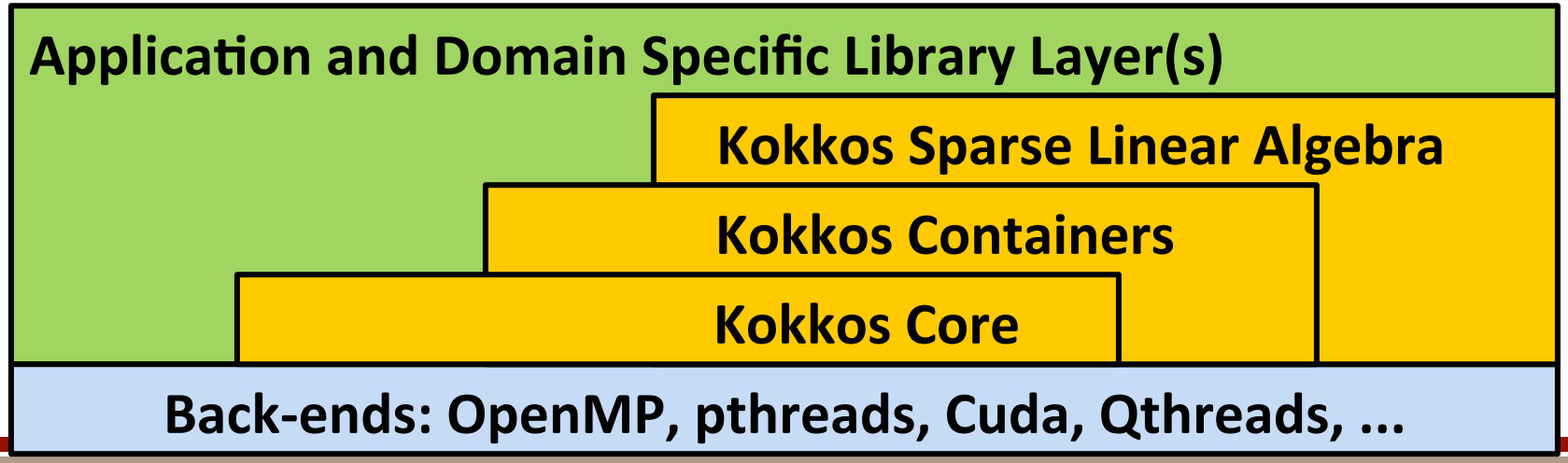
Kokkos for performance-portability

Sandia National Laboratories

Carter Edwards, et al.

- Manycore challenge ("revolution")
  - Terascale workstations, petascale clusters, exascale supercomputers using manycore devices; e.g., Intel Xeon Phi, NVIDIA Kepler
  - Performance requires architecure-specific memory access patterns: caching, coalescing, streaming, vectorizing, ...
  - Goal: refactor applications & domain libraries "1+ε" times
- Intuitive for computational engineers and researchers
  - Performance-portable and extensible for changing architectures
  - Standard C++: avoid language extensions & special compilers

Back



Kokkos: Polymorphic Memory Access

Sandia National Laboratories

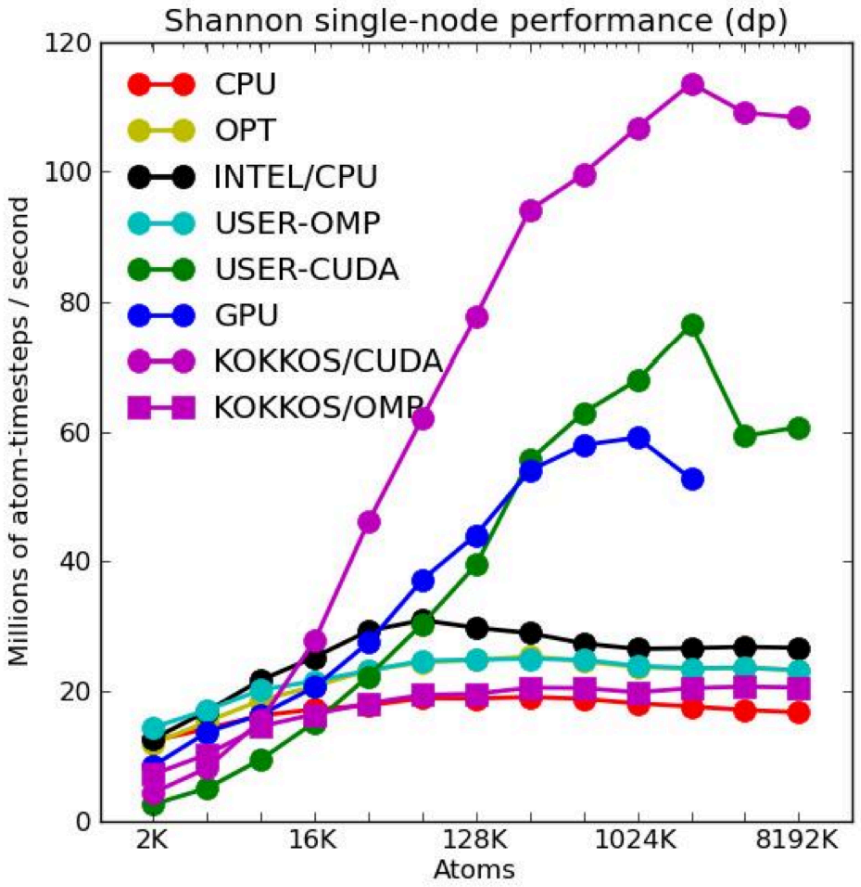
- Polymorphic parallel execution policy + multi-dim array layout
  - Execution policy maps parallel loop body to threads
  - Multidimensional array layout maps array multi-indices to datum
  - Compose to manage memory access pattern
- Compile-time: choose device-appropriate policy & layout
  - Without modifying application's or domain library's source code
  - "Magic" of C++ template meta-programming
- Choose execution space and memory space
  - E.g., CPU, attached accelerator, NUMA regions, coherency domains, ...
- Bonus: transparently utilize special hardware capabilities
  - Random access, read only → GPU texture cache boosts performance
- Agile Components stack will/does use templates + Kokkos…

Back

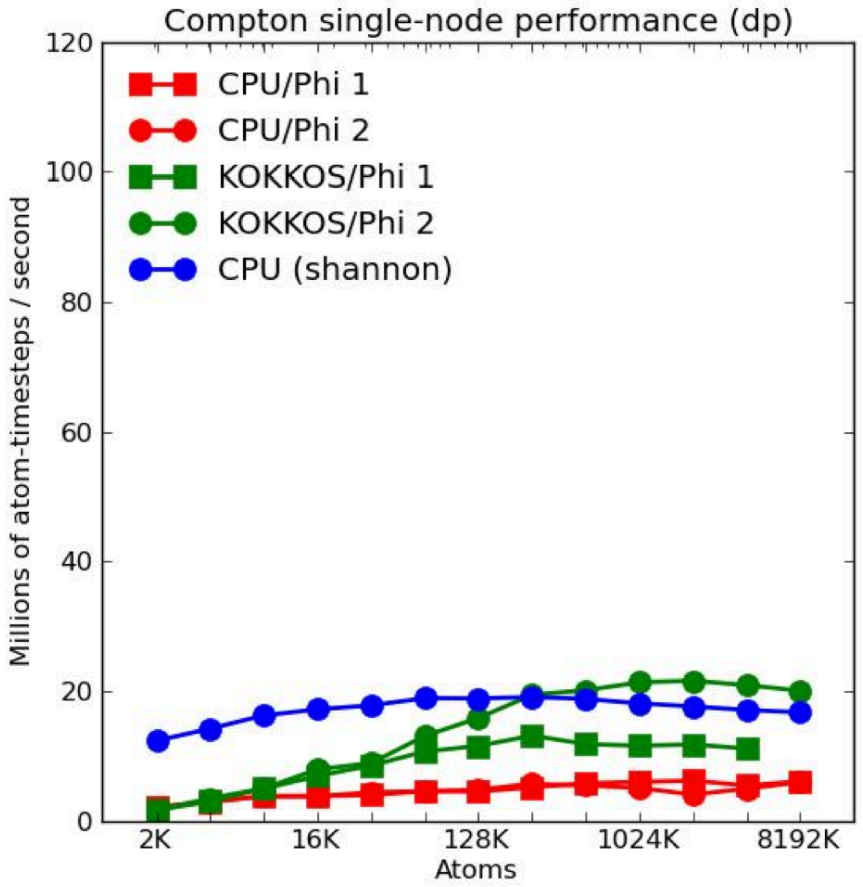
LAMMPS using Kokkos

Sandia National Laboratories

Plimpton & Trott



Good new for Kokkos on GPUs



Kokkos on PHI (expected to be much better on Trinity)

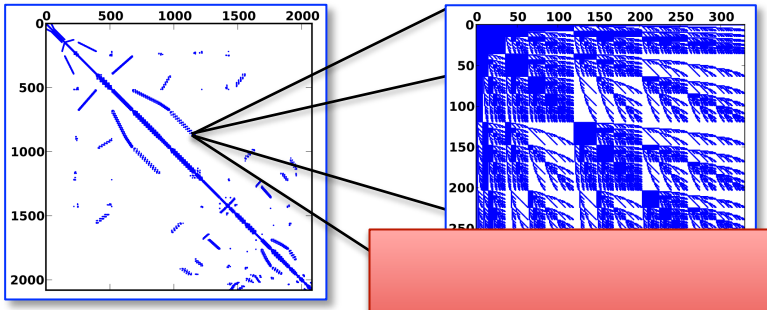
Back

Embedded UQ

Phipps, D'Elia, Edwards, Hu, Rajamanickam

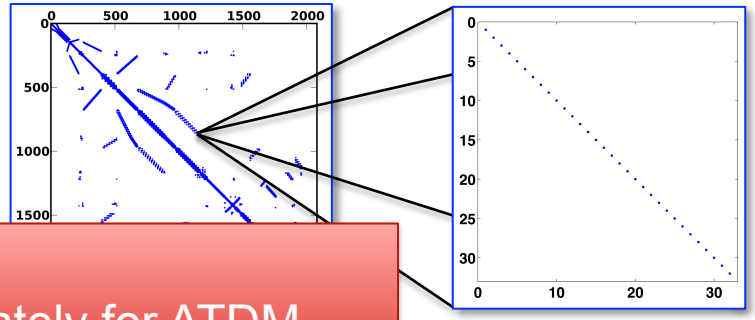
- Propagate UQ info at "scalar" level of PDE discretization/solvers
- Map parallelism in UQ to fine-grained hardware parallelism (SIMD, SIMT, HW threads)
- Effect code transformation through template-based generic programming (TBGP)

Stochastic Galerkin (LDRD)

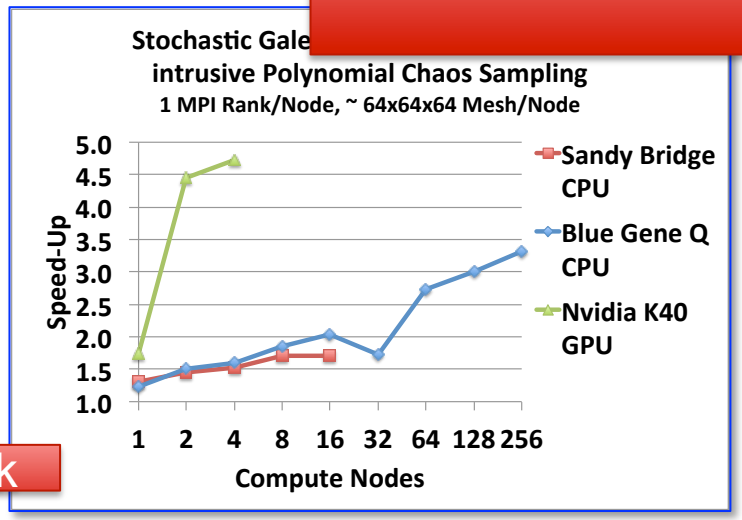


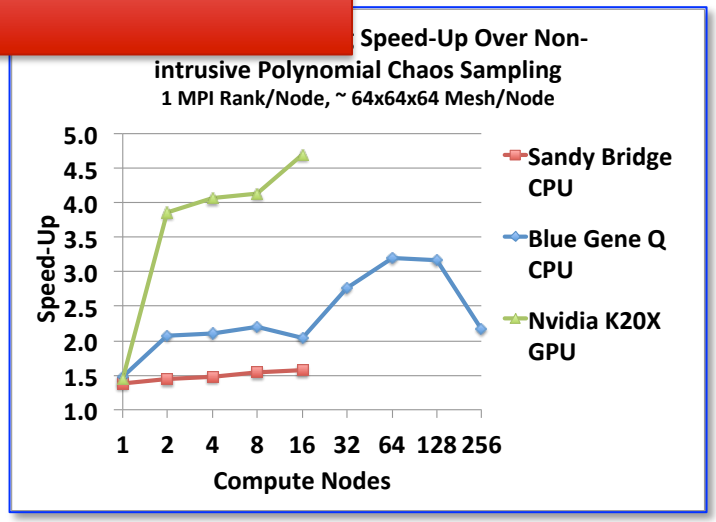
Spatial sparsity

Ensemble (ASCR)



Leveraging immediately for ATDM



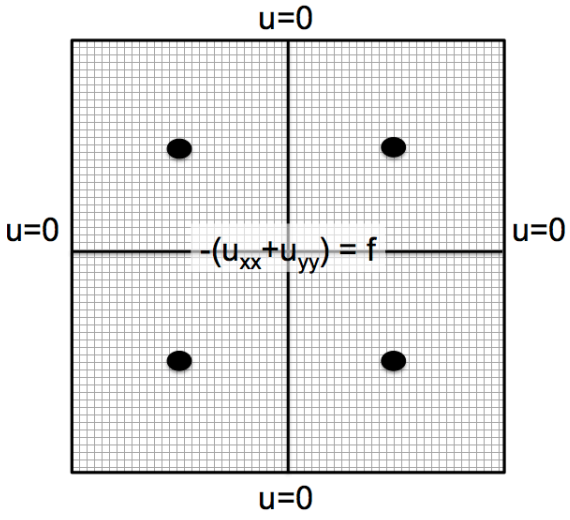


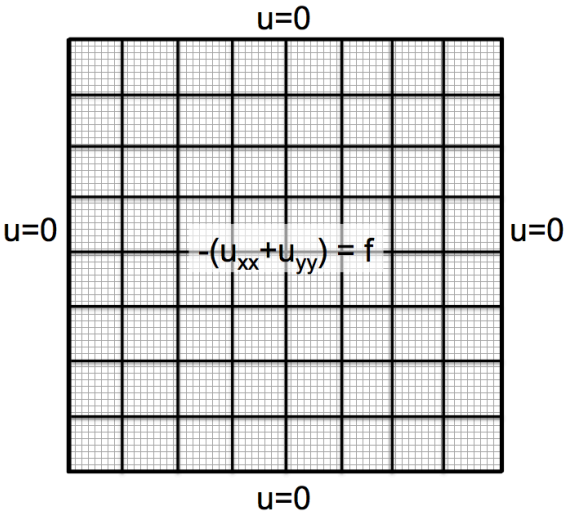
Back

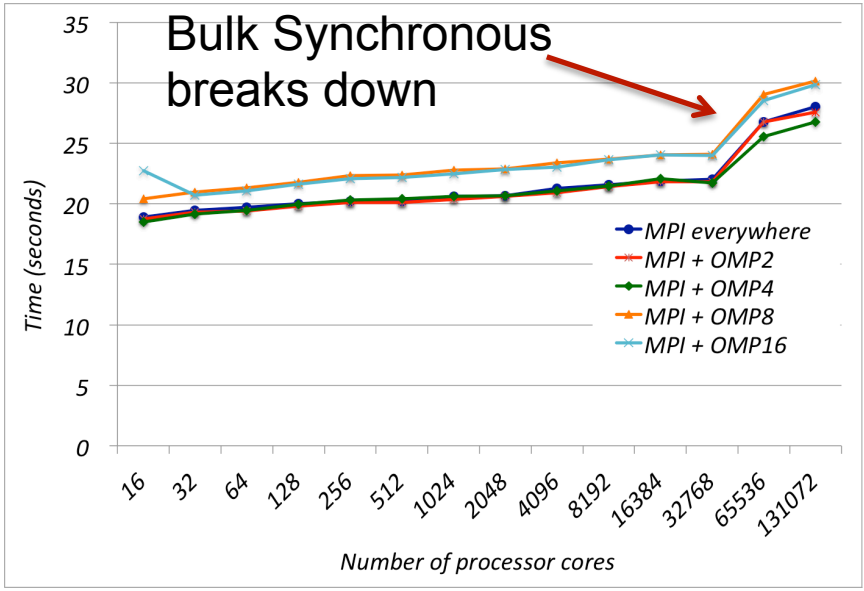
Task Parallel Over Decomposition

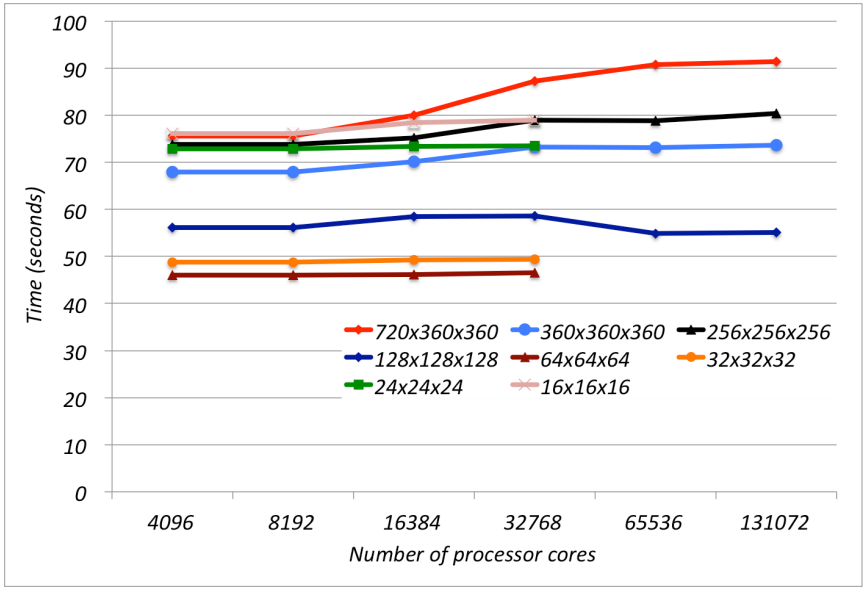
Barrett, Stark, et al.
to appear SC'14

- MPI + Qthreads
- MiniGhost (3d FD)
- Work stealing scheduler
- Over decomposition leads to improved weak scaling
- Overhead and ordering of tasks needs to be reduced…
- Could have used directed acyclic graph…
- ATDM L2 milestone will explore further



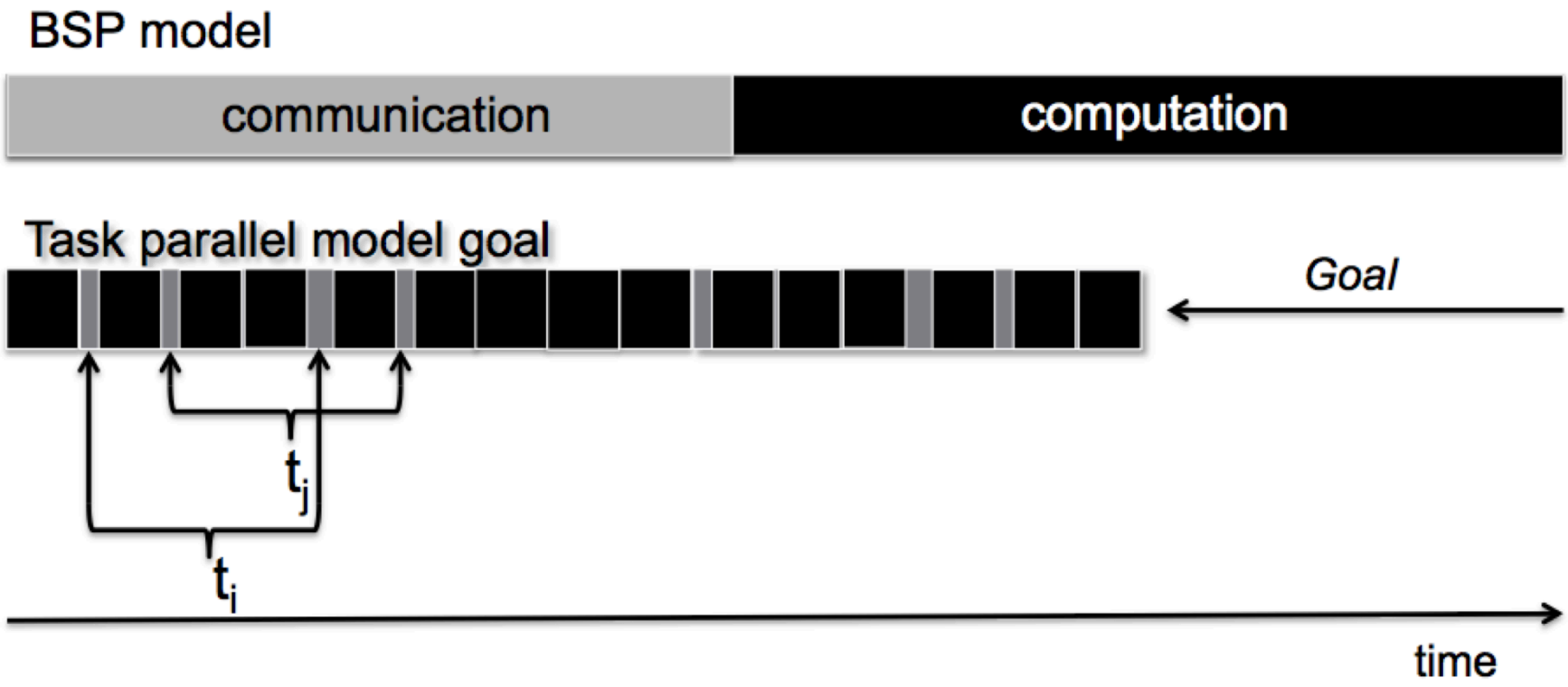






Back

Bulk Synchronous vs Task-based



From Barrett, Stark, et al. to appear SC'14

Back

Data-Management

Sandia National Laboratories

- Capabilities for Integrated Workflows (Nessie, NNTI – CSSE/ATDM)
  - RPC-based framework for developing data services
  - Portable RDMA abstraction over HPC interconnects (Cray XT/XE, IBM BG, IB)
  - Used by ADIOS, Kelpie, Triton (ANL-PFS project)
- Capabilities for data sharing (Kelpie – CSSE/ATDM)
  - In-memory, high-performance key-value store
  - Enables fast storage and sharing complex data structures
  - Use cases: resilience, code coupling, “partial” persistence
- Capabilities for In-situ Analysis and Visualization (CSSE/ATDM)
  - ParaView/Catalyst (SNL/Kitware)
  - Current focus on modularity, low memory footprint, scalability
- OS and Runtime changes to support integrated workflows (CSSE and ASCR)
  - Hobbes and Argo – Both ASCR projects
  - Resource management, data sharing, application composition, prog models.

Back