SAND2007-7081P

Parallel, Heterogeneous, Dynamic unstructured Mesh (phdMesh)

A new package in Trilinos

H. Carter Edwards
Sandia National Laboratories

Trilinos User Group Meeting
Nov 6-8, 2007

Discovery at the Interface of Science and Engineering: Science Matters!

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

R&D Project Spin-off

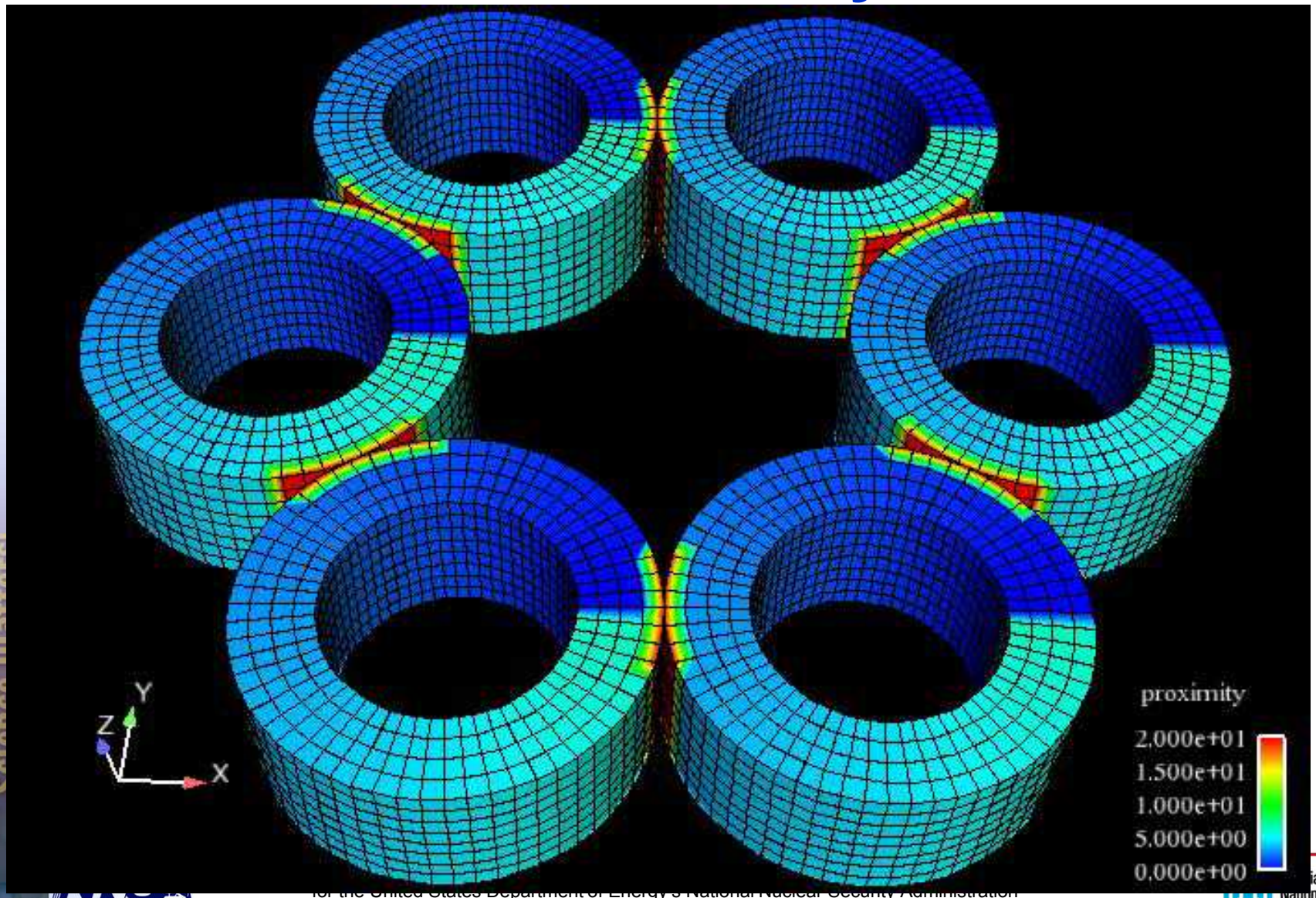
- **Spin-off from Sandia National Laboratory's R&D project for "HPC Application Performance Analysis and Prediction"**
- **Purpose: Enable improved decision making for next generation computer systems and applications.**
- **Approach: Provide targeted compact, highly portable, "mini application" tools that approximate real application performance. Provide guidance to system and application designers. Establish visibility in research community.**
- **phdMesh is a component developed for a "mini-application" intended to approximate performance of parallel geometric proximity search and dynamic load balancing**

Discovery at the Interface of Science and Engineering: Science Matters!

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

Model Problem: Distributed Dynamic Surface-Surface Proximity Detection



for the United States Department of Energy's National Nuclear Security Administration
under contract DE-AC04-94AL85000.

phdMesh exceeded its R&D charter

- **A full capability (instead of approximate) parallel heterogeneous dynamic unstructured mesh library**
  - **Distributed memory HPC database, not a mesh file format**
  - **Parallel geometric proximity search algorithm**
  - **Dynamic load balancing**
  - **Based upon cumulative concepts and lessons learned from many unstructured mesh data models / projects**
  - **Two guiding principles:**
    1. **Keep it simple, i.e. lean & clean**
    2. **Have a well-defined conceptual model**
- **Requirements → *conceptual model* → software design → implementation**

Discovery at the Interface of Science and Engineering: Science Matters!

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Is *Not* a Traditional Mesh Data Structure

- **Minimize scope: just the mesh data**
  - **Discretization entities: node, edge, face, element**
  - **Discretization fields: variables of the problem**
- **NO predefined master elements or element topologies**
  - **Pair with element topology library to fulfill the traditional role**
  - **Should be flexible, pair with any element topology library**
- **NO predefined refinement templates**
- **NOT constrained to a particular mesh file format**

Fundamental Concept: Mesh Database

- **Mesh Database**
  - ***Mesh*: weblike pattern or construction that fills a domain Ω**
  - ***Database*: large collection of data organized for efficient storage, retrieval, and update**

- **Database = Schema + Bulk Data**
  - **Schema is the specification for data to be managed**
  - **Bulk data is stored, retrieved, and updated**

➢ **Mesh Schema + Mesh Bulk Data**

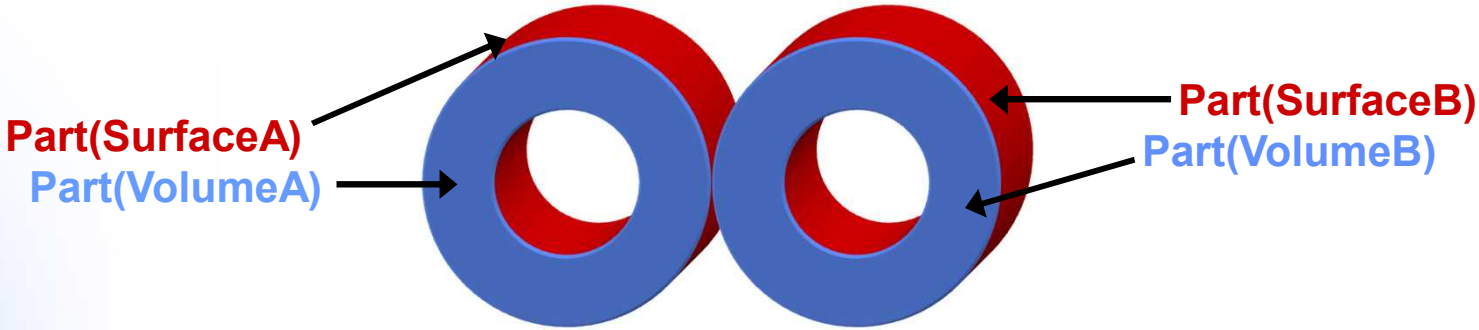
Discovery at the Interface of Science and Engineering: Science Matters!

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

Mesh Schema

- A *specification* for the mesh data to be managed
- *Parts* (subsets) of the problem domain, $\{ \Omega_A \subset \Omega \}$
- Parts' subset / superset relationships, $\Omega_A \subset \Omega_S$



- *Fields* in the discretization, $\{ F_J \}$
  - *Description* of independent, dependent, and auxiliary variables of the problem to be solved over the domain
- Field dimension map, $(F_J , \Omega_A) \rightarrow Dim_{JA}$
  - Variables may have polymorphic dimension
  - e.g., field(n1,n2,n3) such that { n1, n2, n3 } may vary

Mesh Bulk Data

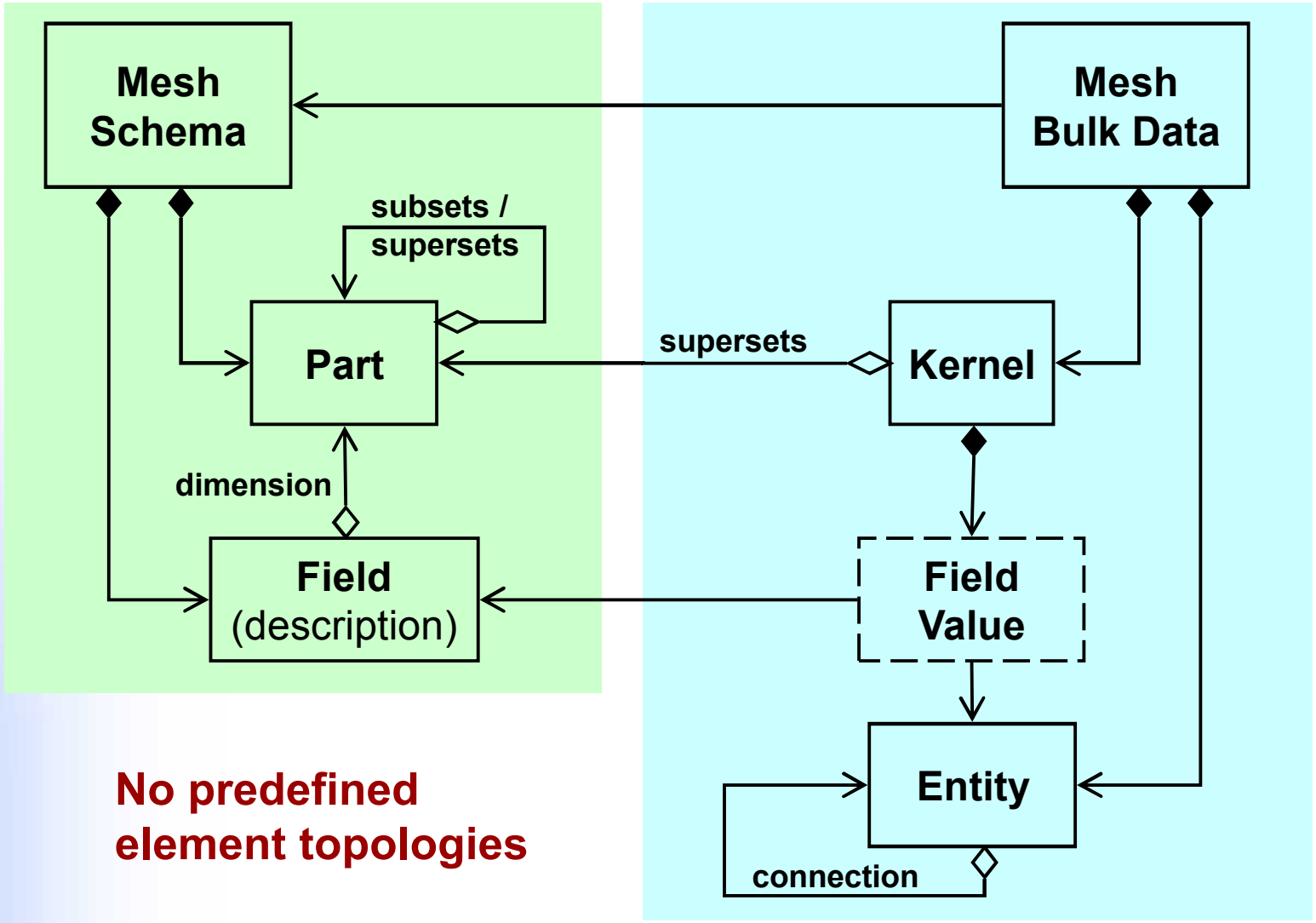
- **Large collection of discretization data conforming to a Mesh Schema**
  - **Many possible discretizations for a single schema**
- ***Entities* of a discretization, e.g. nodes, elements**
- ***Connections* between entities, e.g. element→node**
- ***Field Values* associated with entities**
  - **Per-entity numerical values conforming to field descriptions**
  - **E.g. basis function coefficients, state/material properties**
- ***Kernels* – "chunks" of similar field values**
  - **A contiguous block of memory for arrays of field values**
  - **Needed for performance**

Discovery at the Interface of Science and Engineering: Science Matters!

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

High Level Class Diagram



No predefined element topologies

Discovery at the Interface of Science and Engineering: Science Matters!

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

Parallelization Principles

- **Assumption: Distributed Memory Parallelism**
  - **Not to preclude local task-level (multicore) parallelism**
- **Mesh Schema: replicated on all processors**
- **Mesh Bulk Data: partitioned and distributed among processors**
- ***Minimize impact of global data distribution and local data ordering on algorithms***
  - **Algorithms "see" minimal parallel bookkeeping**
  - **Parallel decomposition and local data ordering *should not effect results*, but may effect performance**
  - **Ideally results are insensitive to global & local parallelism**

Discovery at the Interface of Science and Engineering: Science Matters!

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Summary

- A “lean & clean” (a.k.a. minimalist) approach to the challenge of an in-memory database for parallel, heterogeneous, and dynamic unstructured mesh
- Address local performance: *kernels* of field values
  - Field value arrays in cache friendly “chunks”
  - Anticipating manycore task-based parallelism
- NO predefined / preconceived notion of an element
  - Mesh database doesn’t need element topologies
  - Application needs master elements
  - Pair phdMesh with an element library

Discovery at the Interface of Science and Engineering: Science Matters!

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy’s National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories