LA-UR-17-23218

Using Charm++ to Support Multiscale Multiphysics

On the Trinity Supercomputer

Robert Pavel,
Christoph Junghans,
Susan M. Mniszewski,
Timothy C. Germann

April, 18th 2017

Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

Exascale Co-Design Center for Materials in Extreme Environments (ExMatEx)

- **ExMatEx was one of three* DOE Office of Science application co-design centers (2011-16)**

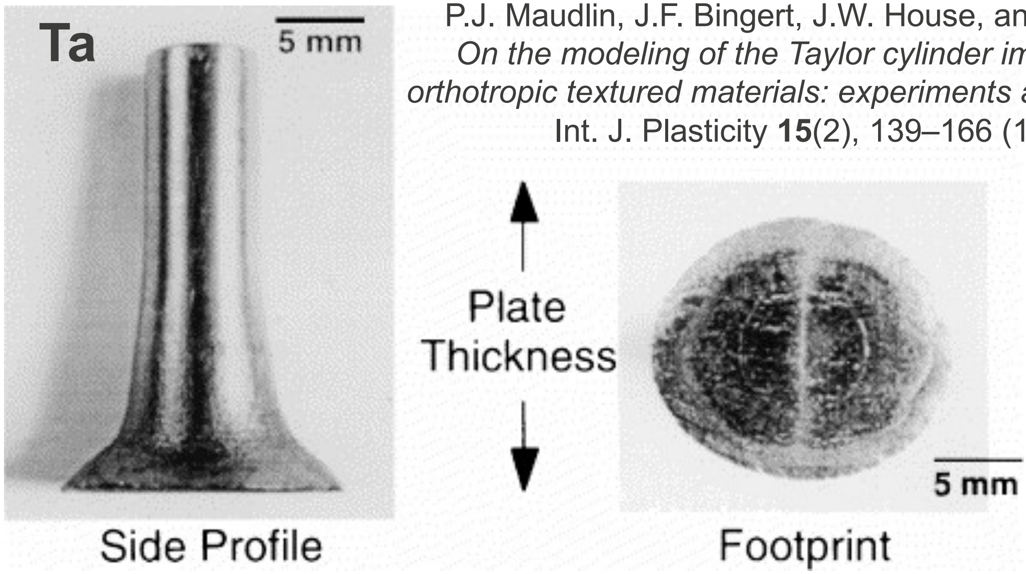
  ***Others are: CESAR (ANL/reactors), ExaCT (SNL-CA/combustion)**

  - Large scale collaborations between national labs, academia, vendors
  - Coordinated with related DOE NNSA co-design efforts
  - Goal: to establish a relationships between algorithms, software stack, and architectures needed to enable exascale-ready science applications

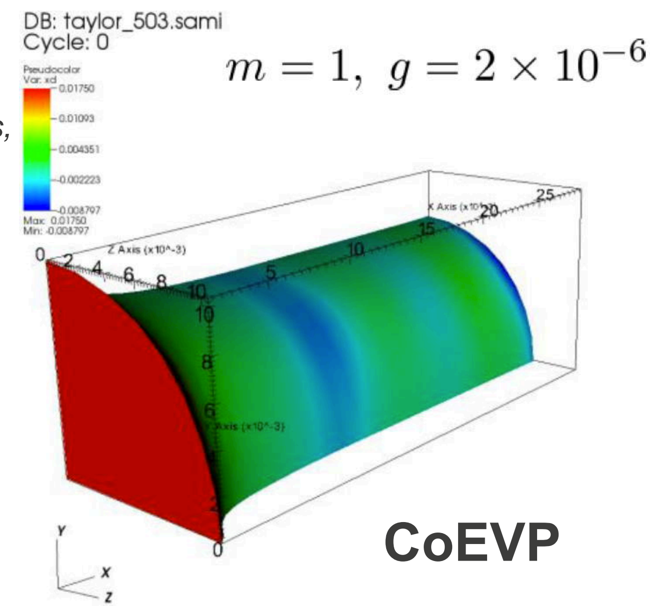
- Two ultimate objectives:
  - Identify the requirements for the exascale ecosystem that are necessary to perform computational materials science simulations (both single- and multi-scale)
  - Demonstrate and deliver a prototype scale-bridging materials science application based upon adaptive physics refinement

Tabasco Test Problem: Modeling a Taylor cylinder impact test

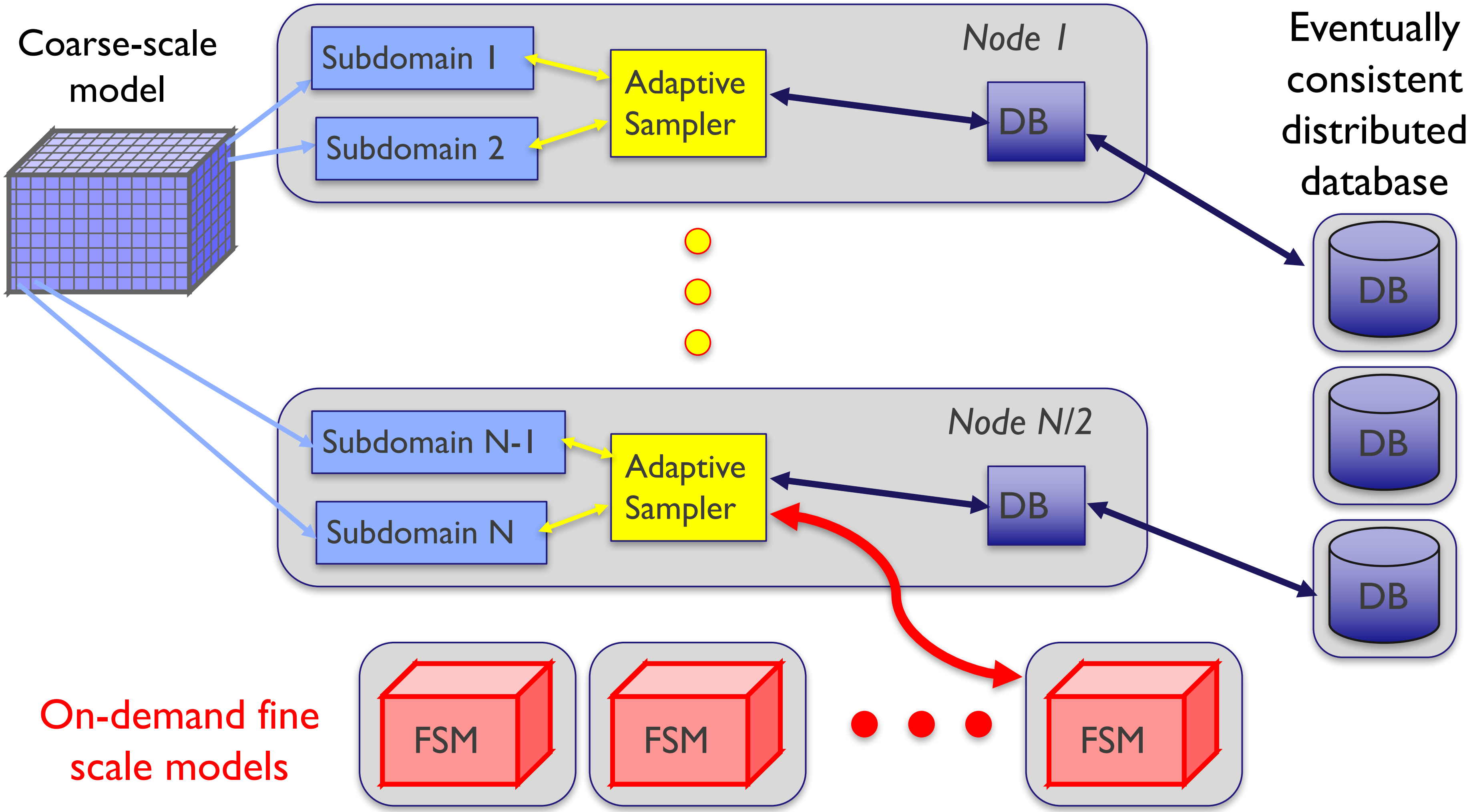
- **The simple Taylor model cannot account for the twinning and anisotropy of a tantalum sample used in a LANL experiments (MST-8), and thus the final shapes do not match.**
- **The physics goal of this demonstration is to show that the more accurate VPSC fine-scale model with an appropriate reduced-dimensionality (~60 degrees of freedom) model of texture can (qualitatively or quantitatively?) reproduce the experimental shape.**



P.J. Maudlin, J.F. Bingert, J.W. House, and S.R. Chen, *On the modeling of the Taylor cylinder impact test for orthotropic textured materials: experiments and simulations,* Int. J. Plasticity **15**(2), 139–166 (1999).



Workflow Overview



Task-Based Scale-bridging Code (TaBaSCo)

- **Demonstrate the feasibility of at-scale heterogeneous computations composed of:**
  - Coarse-scale Lagrangian hydrodynamics
  - Dynamically launched constitutive model calculations
  - Results of fine-scale evaluations stored for reuse in databases
    - Use of Taylor fine scale model evaluation
    - VPSC fine scale model evaluation
  - Adaptive sampling which queries the database, interpolates results, and decides when to spawn fine-scale evaluations
- **Combines an asynchronous task-based runtime environment with persistent database storage**
  - Provides load-balancer and checkpoints for fault tolerant modes

Tabasco Framework

- **Asynchronous Task-based runtimes explored**
  - CHARM++ (built/ran on Trinity)
  - libCircle (built/ran on Darwin, but not Trinity)
  - MPI Task Pool (dual binary version built/ran on Darwin, single binary version ran for small examples on Trinity)
- **Nearest neighbor search**
  - Mtree vs. FLANN (both worked on Trinity)
- **Database storage**
  - In memory HashMap (was limited for long runs)
  - Posix (became our reliable database option for Trinity)
  - Posix/Data Warp (only worked for short runs on Trinity)
  - REDIS (never ran on Trinity)

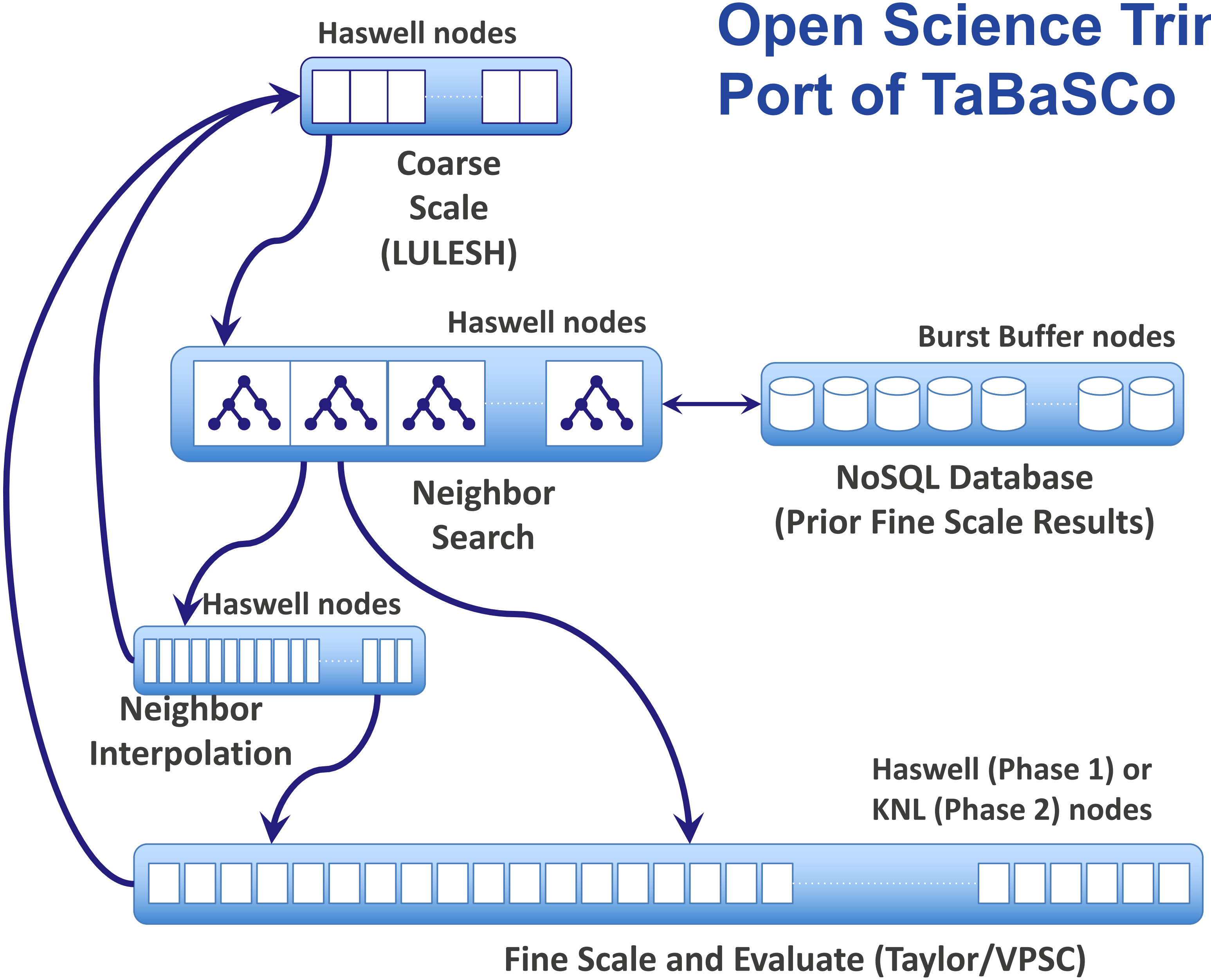
Chare Wrapper mapping

Chare	Dim	Class	Lib	Resolution	Migrate
CoarseScaleModel	1D	Lulesh		MPI Rank	N
FineScaleModel	2D	Constitutive/ ElastoPlasticity	CM	Element	Y
Evaluate	1D	Taylor/VPSC	CM	Element	Y
NearestNeighbor Search	1D	Approx NearestNeighbors	CM/FLANN /Mtree	Request (Service)	N
DBInterface	1D	KrigingDataBase/ SingletonDB	CM/Redis/ libhio/ POSIX	Read/Write (Service)	N

Trinity: Advanced Technology System

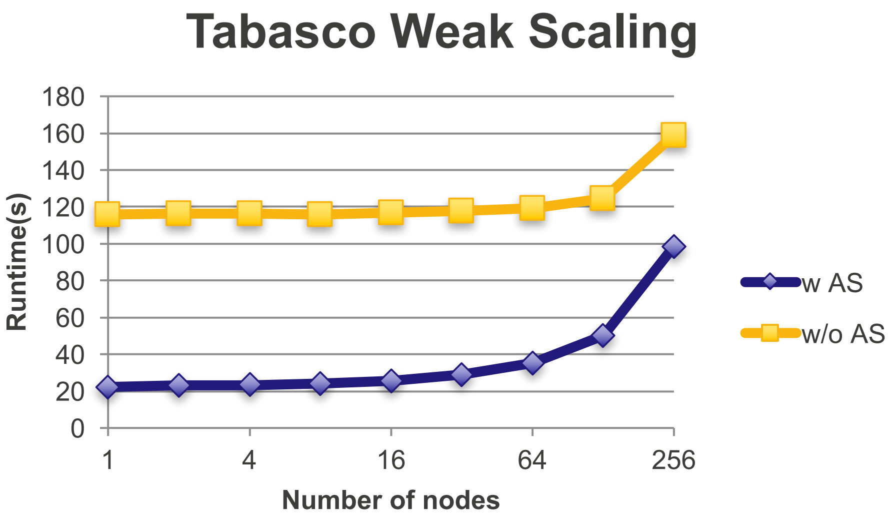
(a mixture of Intel Haswell and Knights Landing (KNL) processors)

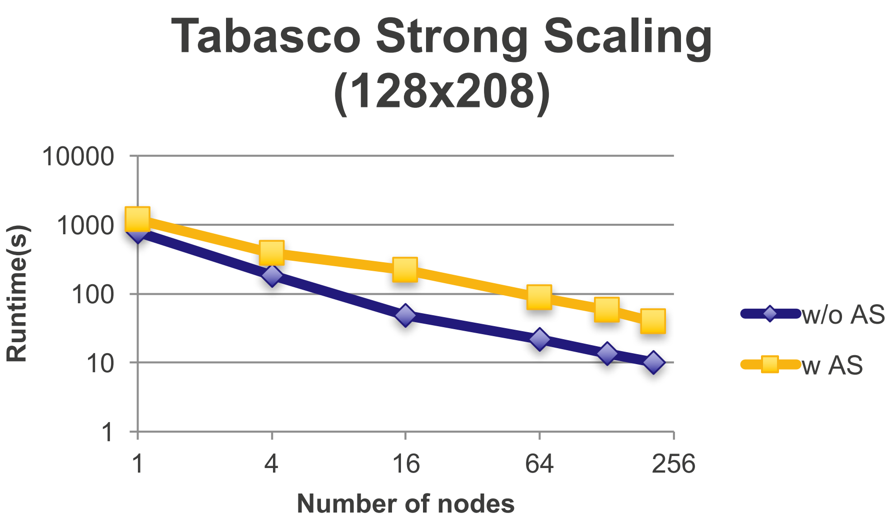
Open Science Trinity Port of TaBaSCo



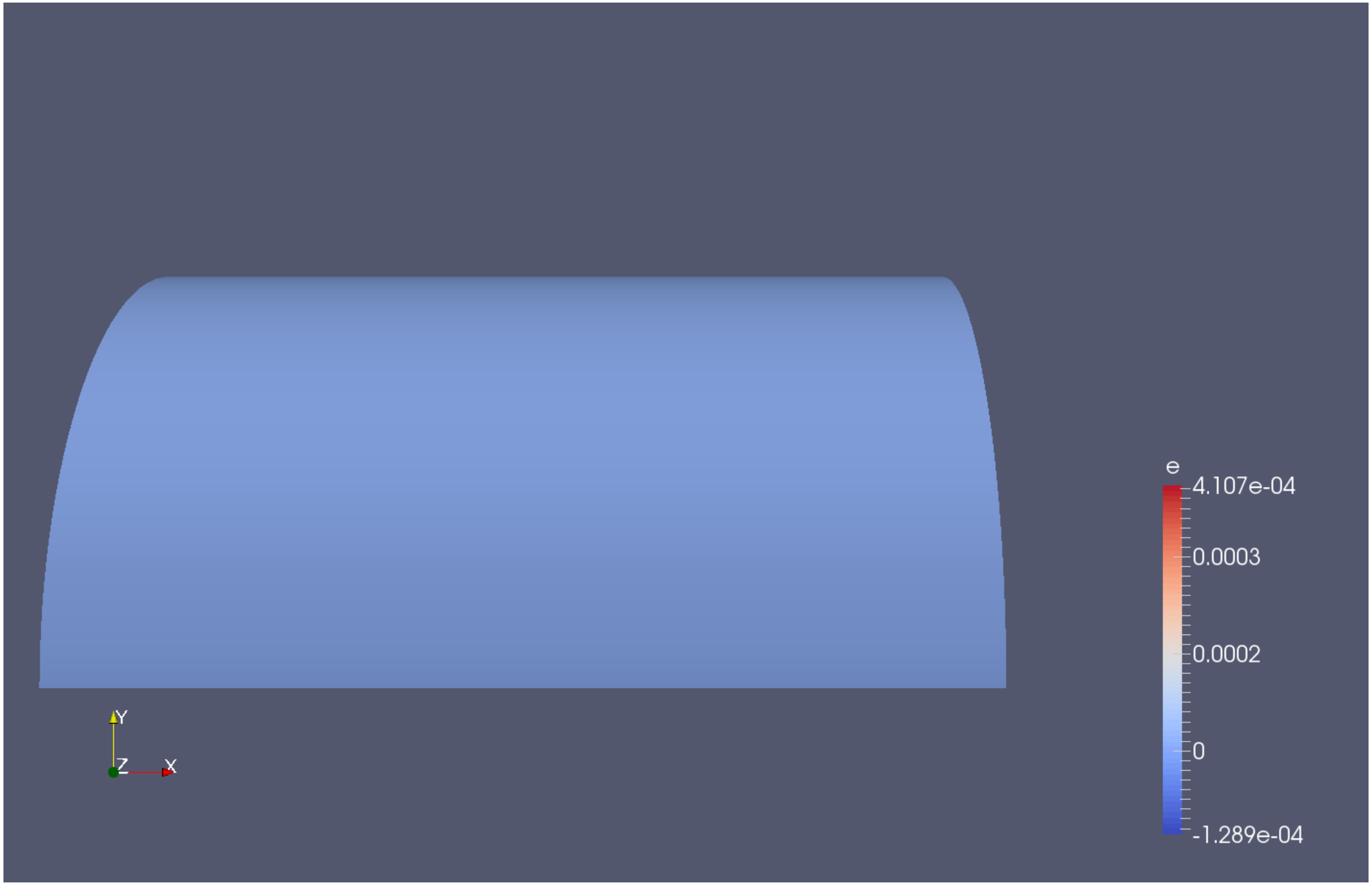
Tabasco Weak and Strong Scaling

- **Weak scaling**
  - Brute force (w/o AS) and Adaptive Sampling (w AS)
  - Edge=64, height=26-13,312, 46,592-23,855,104 elements
  - Good till 128 nodes
  - Communication overhead for >=256
- **Strong scaling**
  - Brute force (w/o AS) and Adaptive Sampling (w AS)
  - Edge=128, height=208, 1,490,944 elements

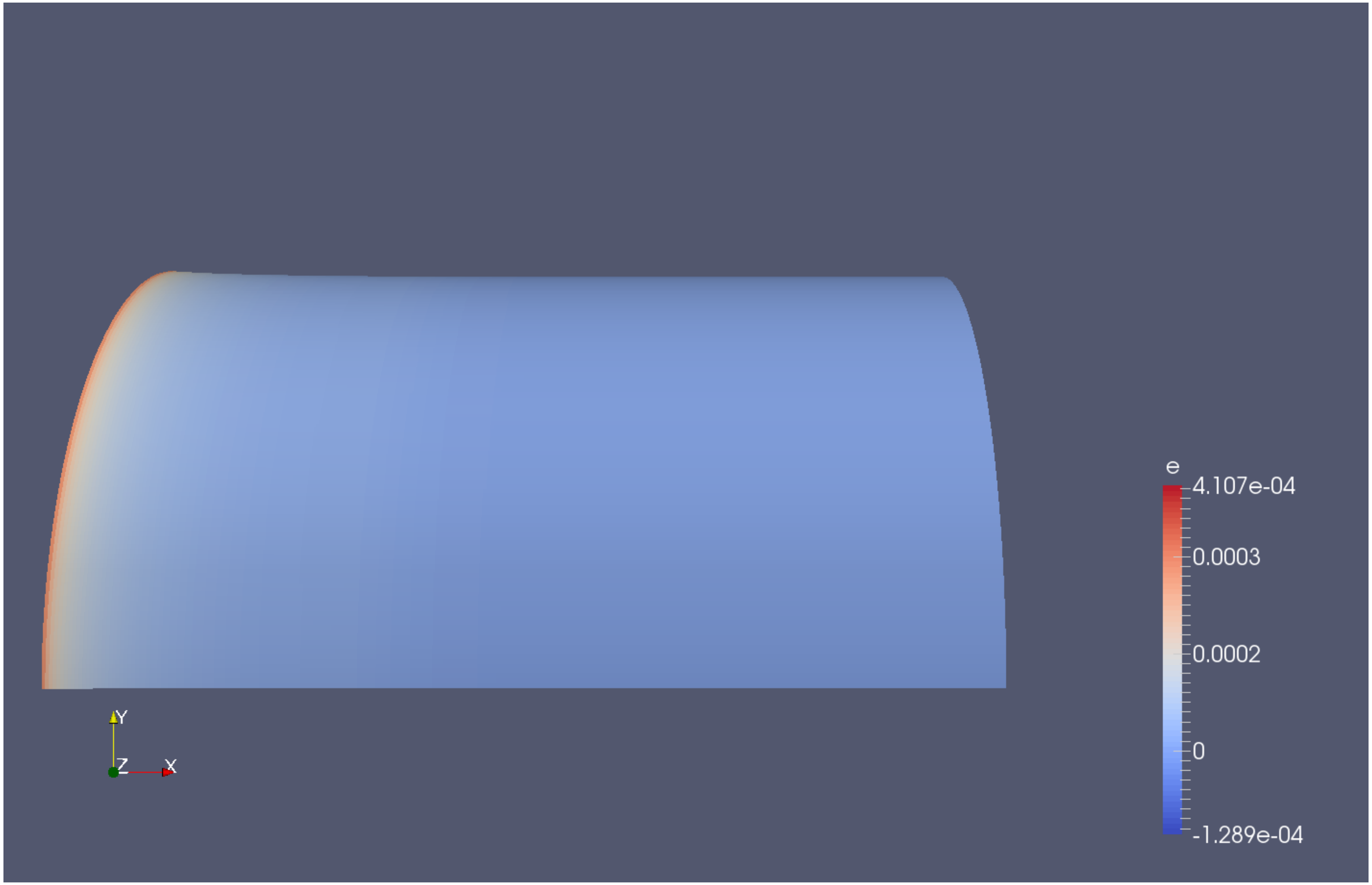




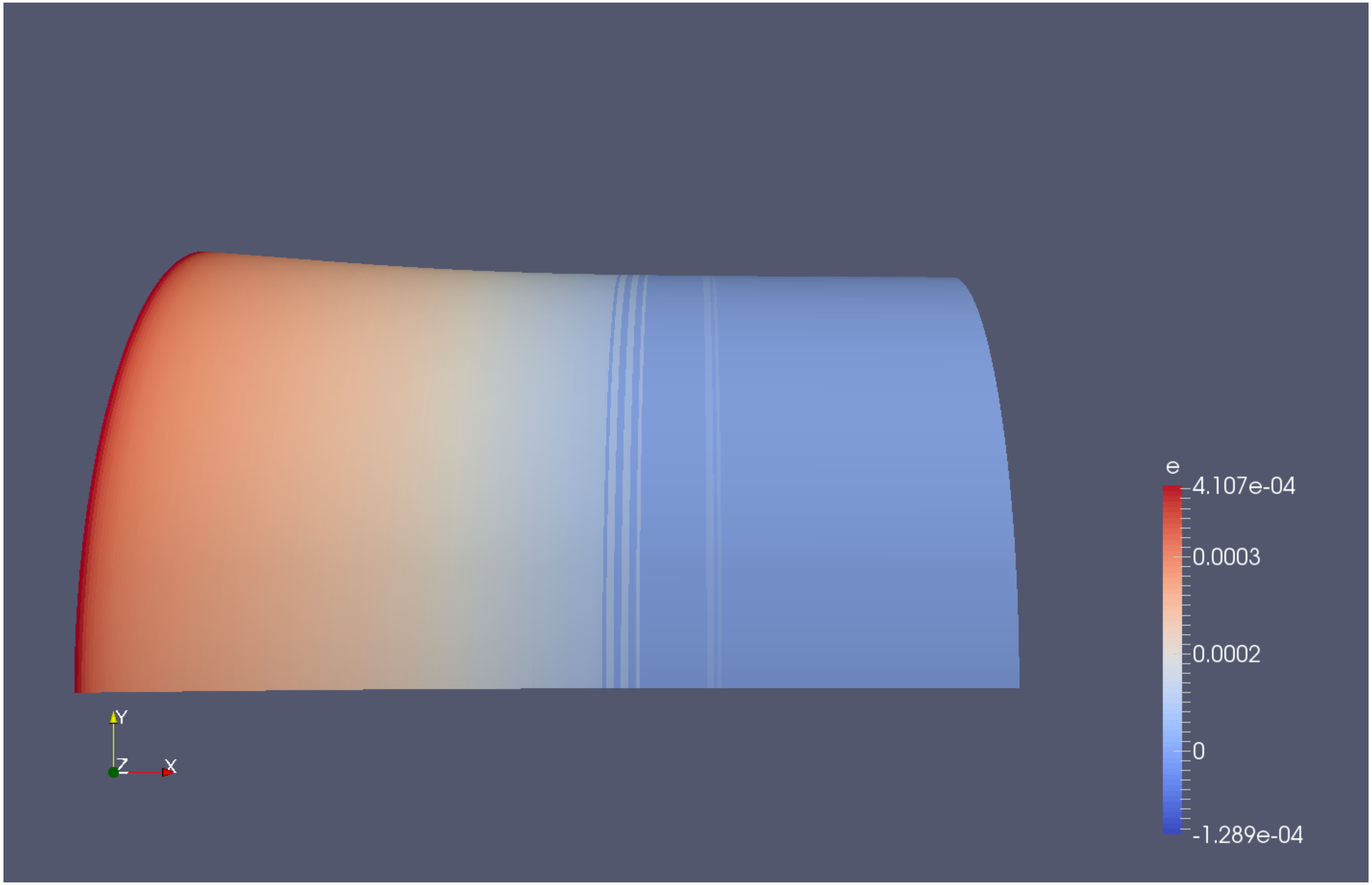
Tabasco Brute Force (w/o Adaptive Sampling) on Trinity (512 nodes) – step 0



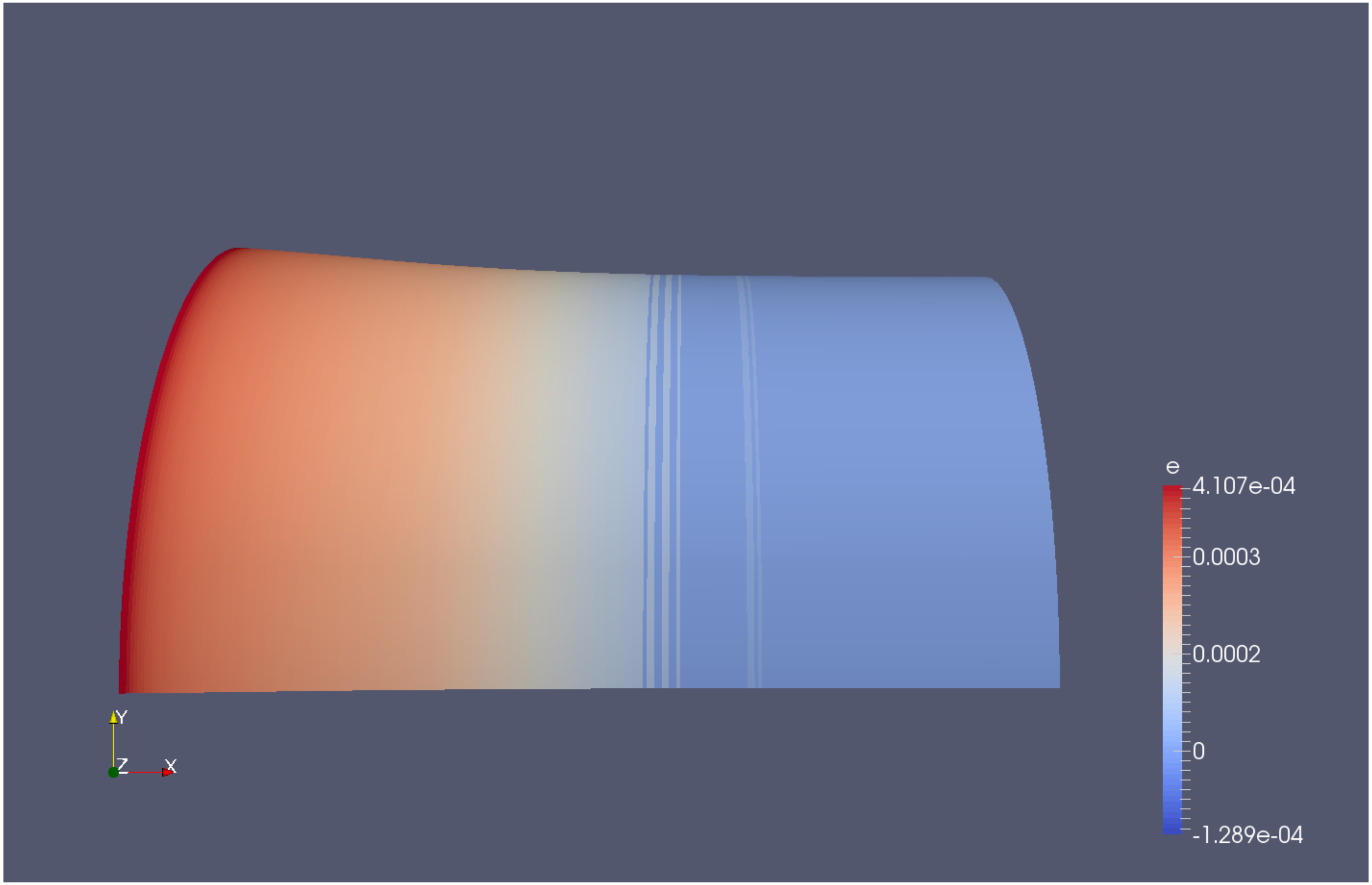
Tabasco Brute Force (w/o Adaptive Sampling) on Trinity (512 nodes) – step 500



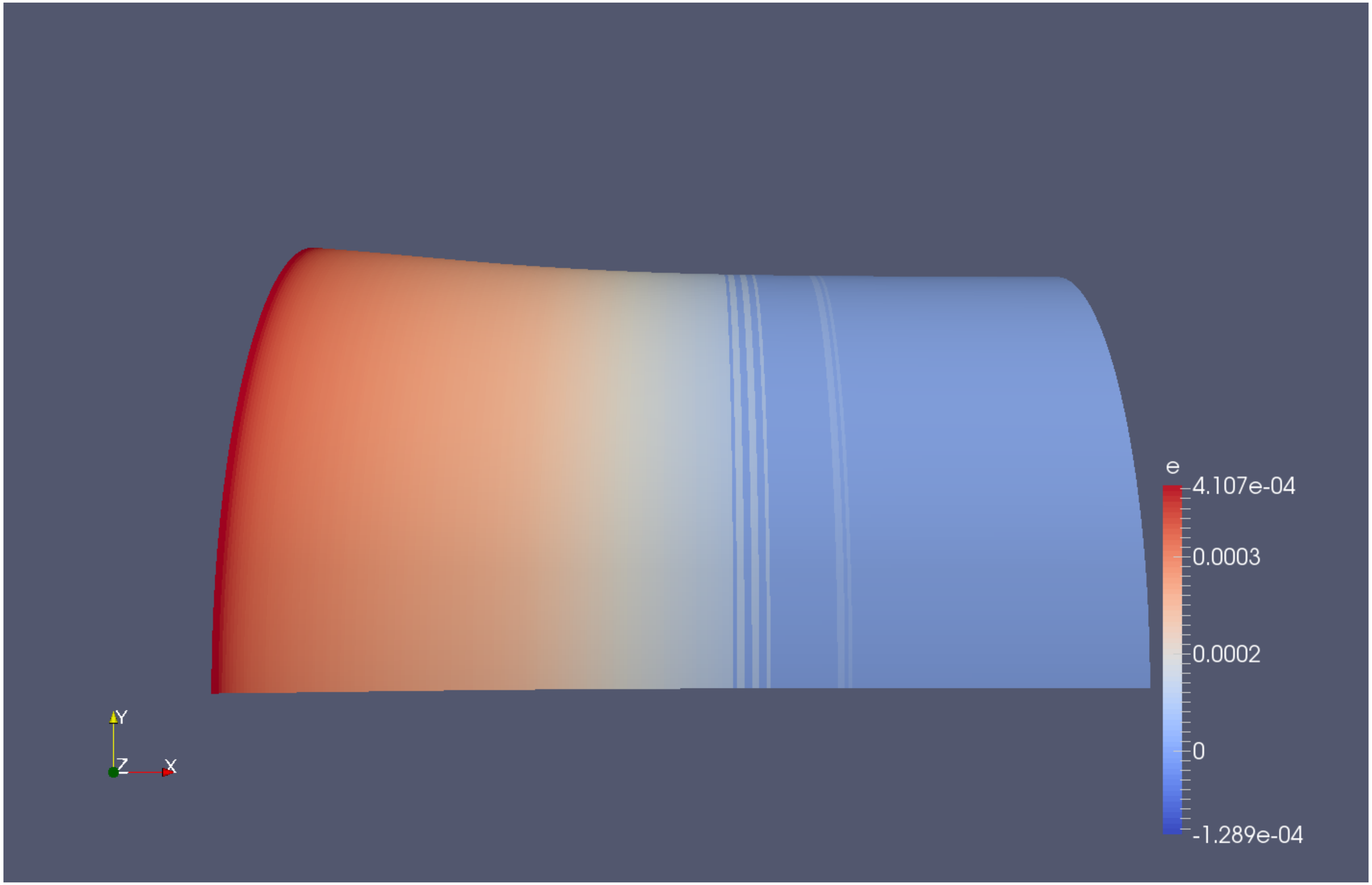
Tabasco Brute Force (w/o Adaptive Sampling) on Trinity (512 nodes) – step 5000



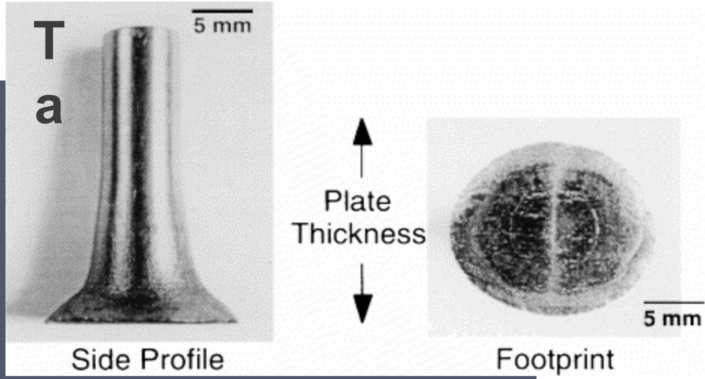
Tabasco Brute Force (w/o Adaptive Sampling) on Trinity (512 nodes) – step 10,000

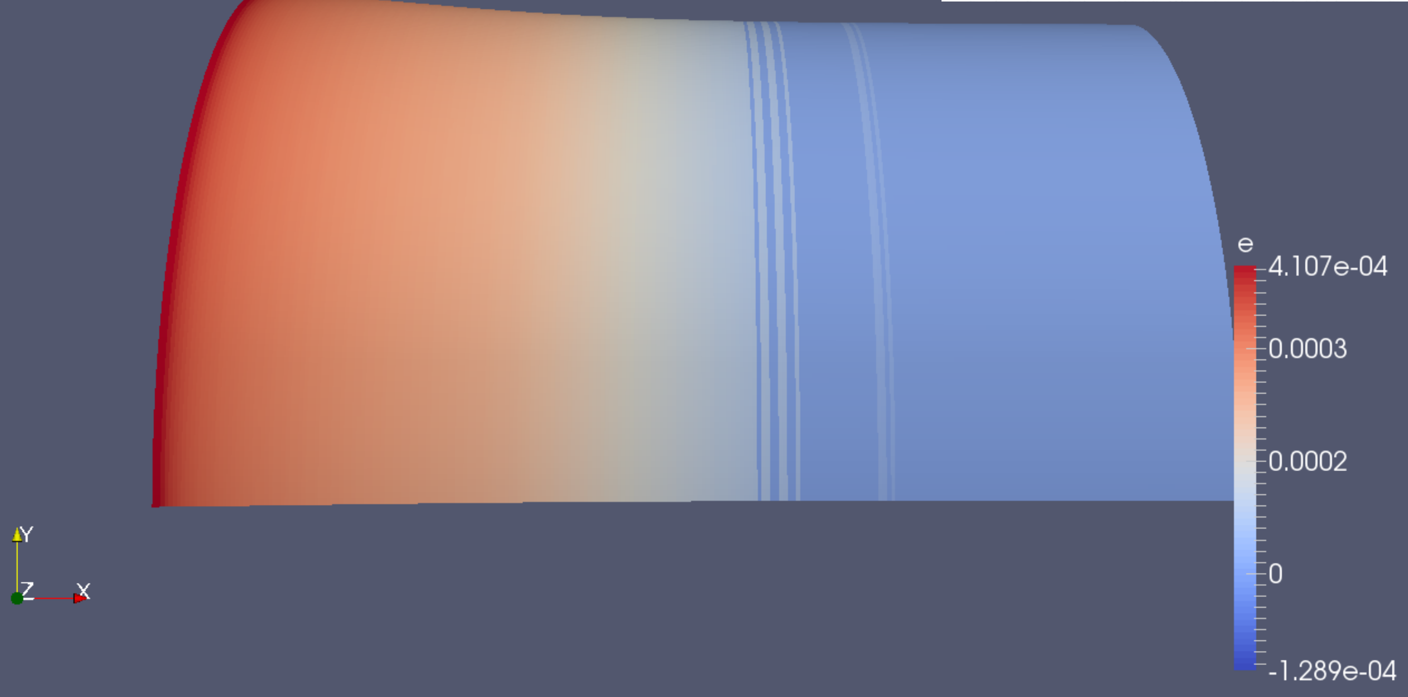


Tabasco Brute Force (w/o Adaptive Samp.) on Trinity (512 nodes) – step 20,000



Tabasco Brute Force (w/o Adaptive Samp.) on Trinity (512 nodes) – step 22,000

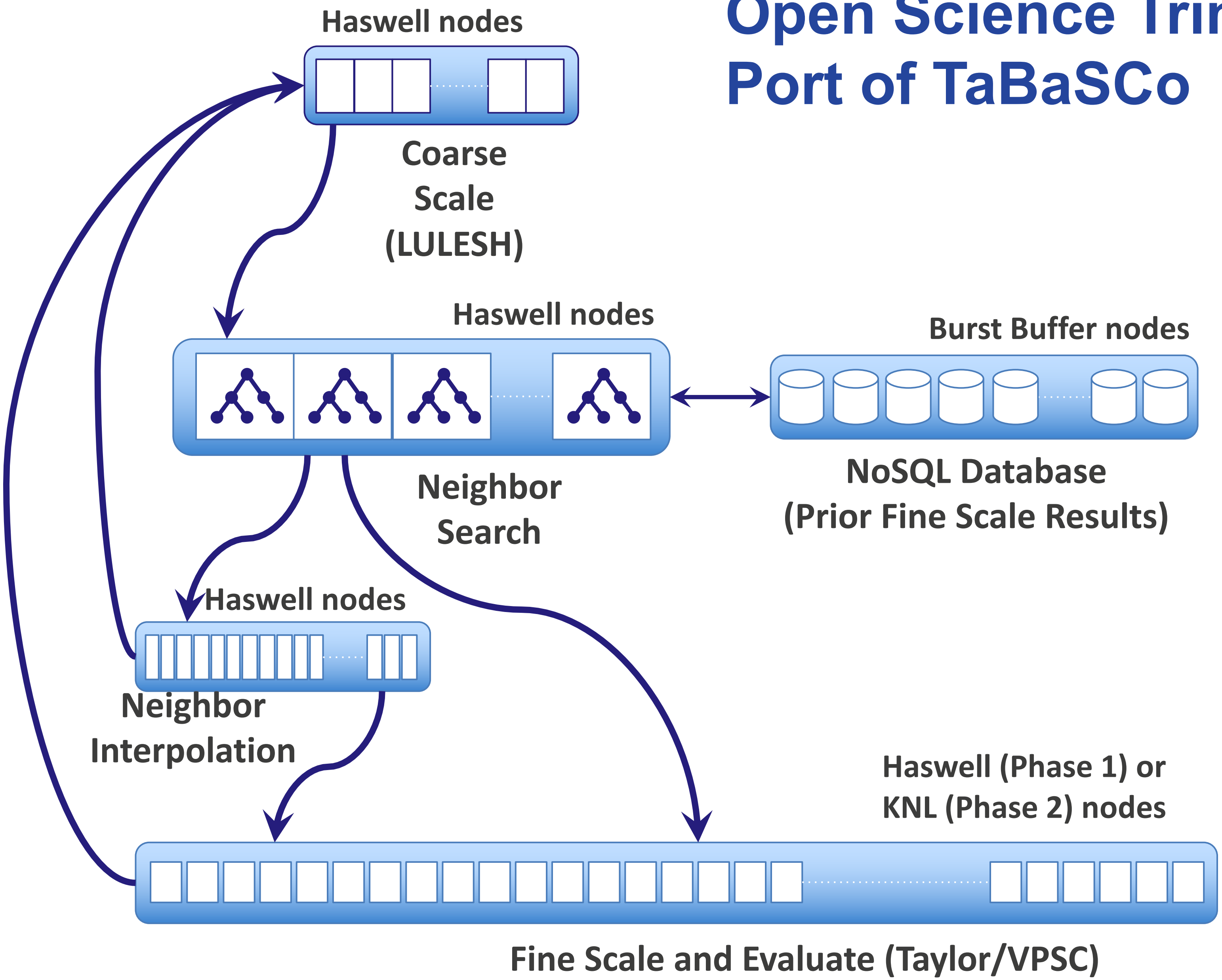




Early Work in Hybrid Runs on Trinity

- **Trinity is a machine that was installed in two stages**
  - 9408 compute nodes with Intel Haswell Processors
    - 32 CPU Cores per node, each with 2x hyperthreads
  - 9500 compute nodes with Intel Knights Landing processors
    - 68 cores per node
- **While similar, nodes have different strengths**
- **Goal of Tabasco is to perform hybrid run in which both stages are utilized**
  - Coarse scale solver and less compute intensive work on Haswell nodes
  - Fine grain solver on KNLs
- **Used Charm++'s Logical Machine Entities to identify KNLs and Haswells**
  - And then used custom mapper to assign chares based on physical node type

Open Science Trinity Port of TaBaSCo



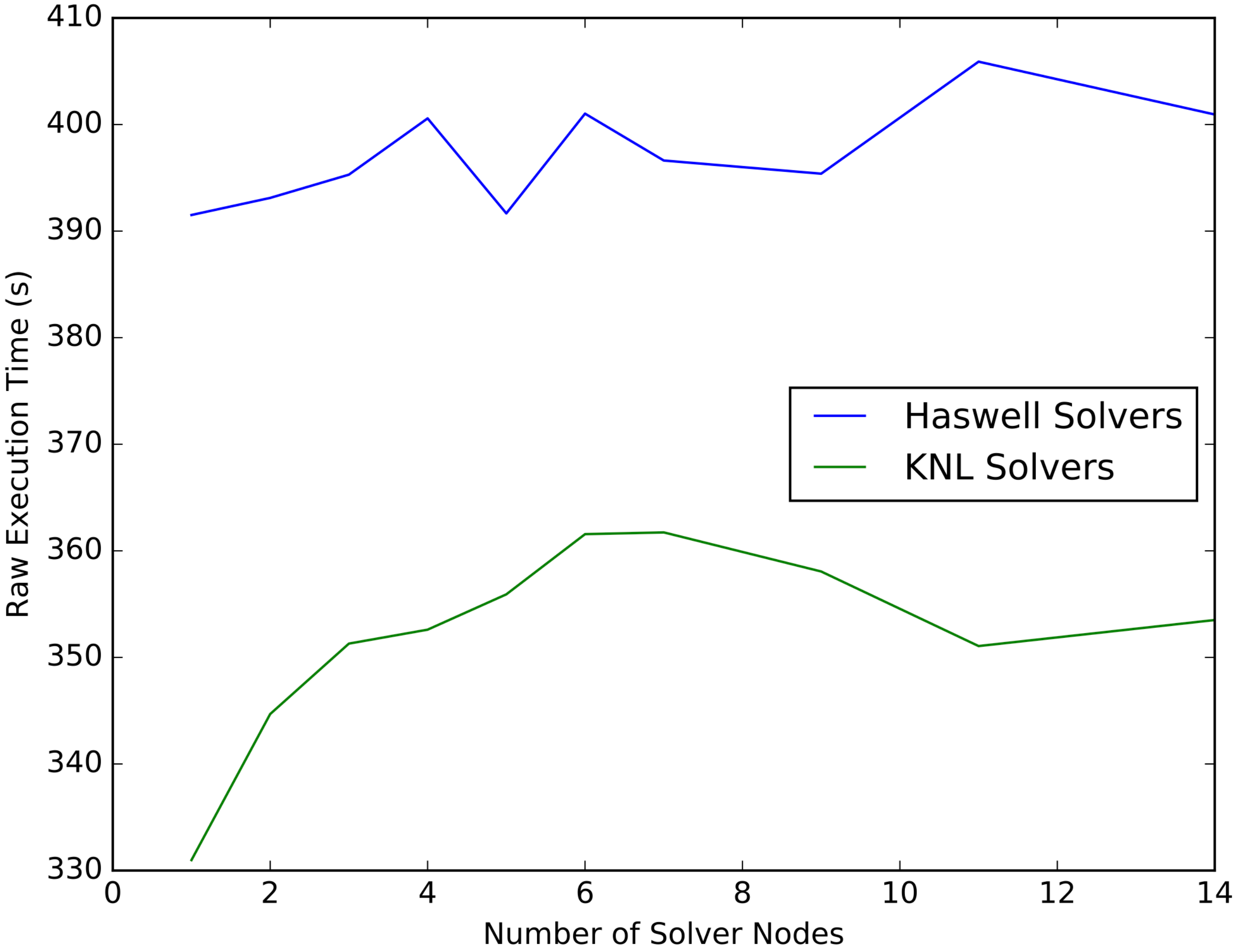
Proof of Concept Hybrid Run: Host Platform

- **Current Proof of Concept implementation running on Trinitite**
  - Run with three types of node
    - Dual Socket Haswell ``Solver'' Node
      - 32 MPI ranks per node
    - KNL ``Solver'' Node
      - 64 MPI ranks per node
    - Dedicated Haswell ``Organizer'' Node
      - 4 MPI ranks per node
  - Run on a subset of Trinitite
    - Goal was to work with the system stack and get initial performance results
    - Larger Runs Planned following unification of Trinity phases

Proof of Concept Hybrid Run: Simulation Configuration

- **Restricted to ``Taylor" Solver**
- **No Adaptive Sampling**
  - Maximize Work, Minimize Communication
- **Re-used a run from Open Science Phase 1**
  - 48 Edge Elements
  - 104 Height Elements
  - 4 Domains (Coarse Scale Chares)
  - 34944 Fine Scale Evaluations
  - 100 Time Steps

Proof of Concept Hybrid Run Results: Raw Execution Time

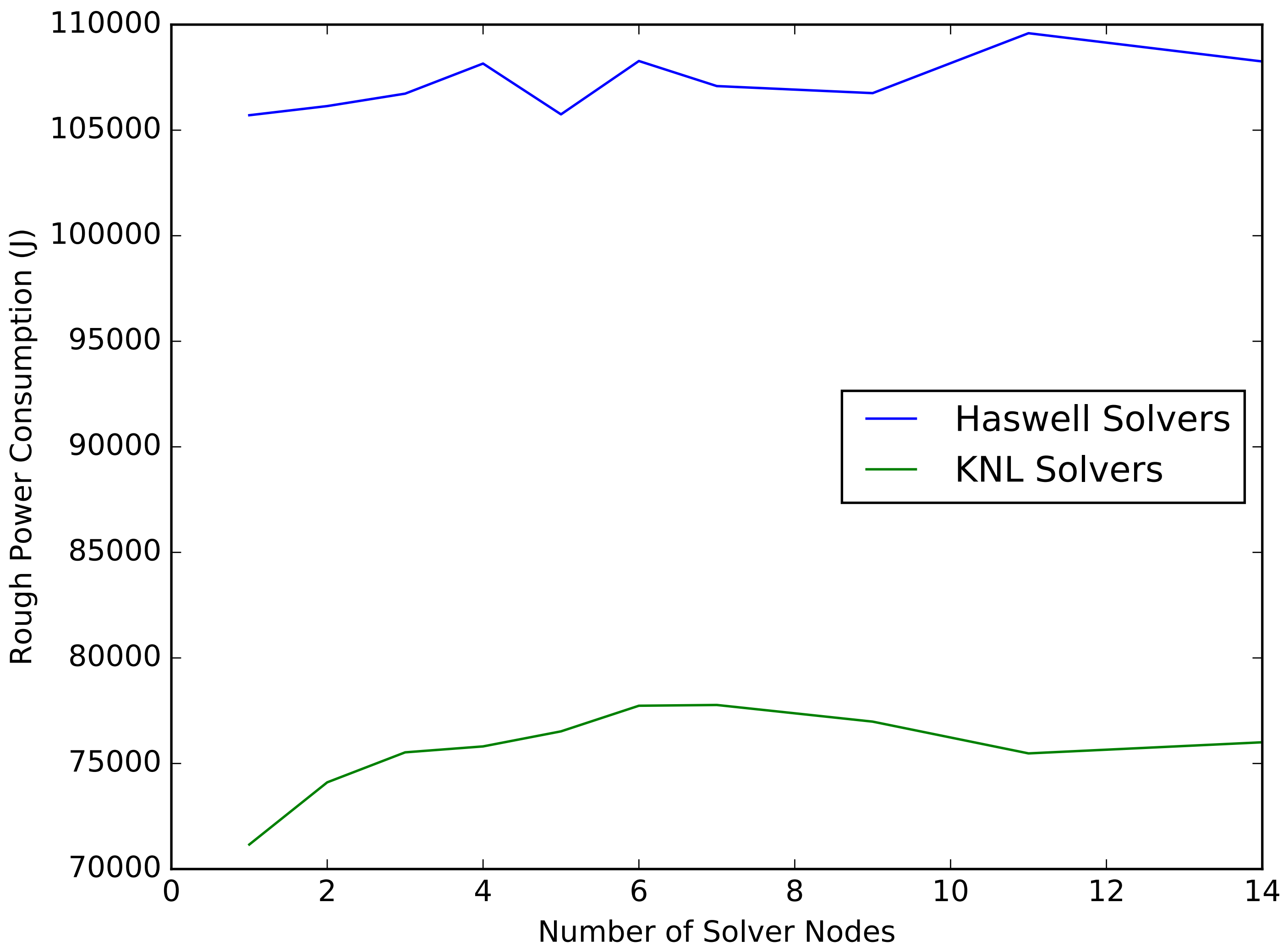


Approximation of Energy Savings

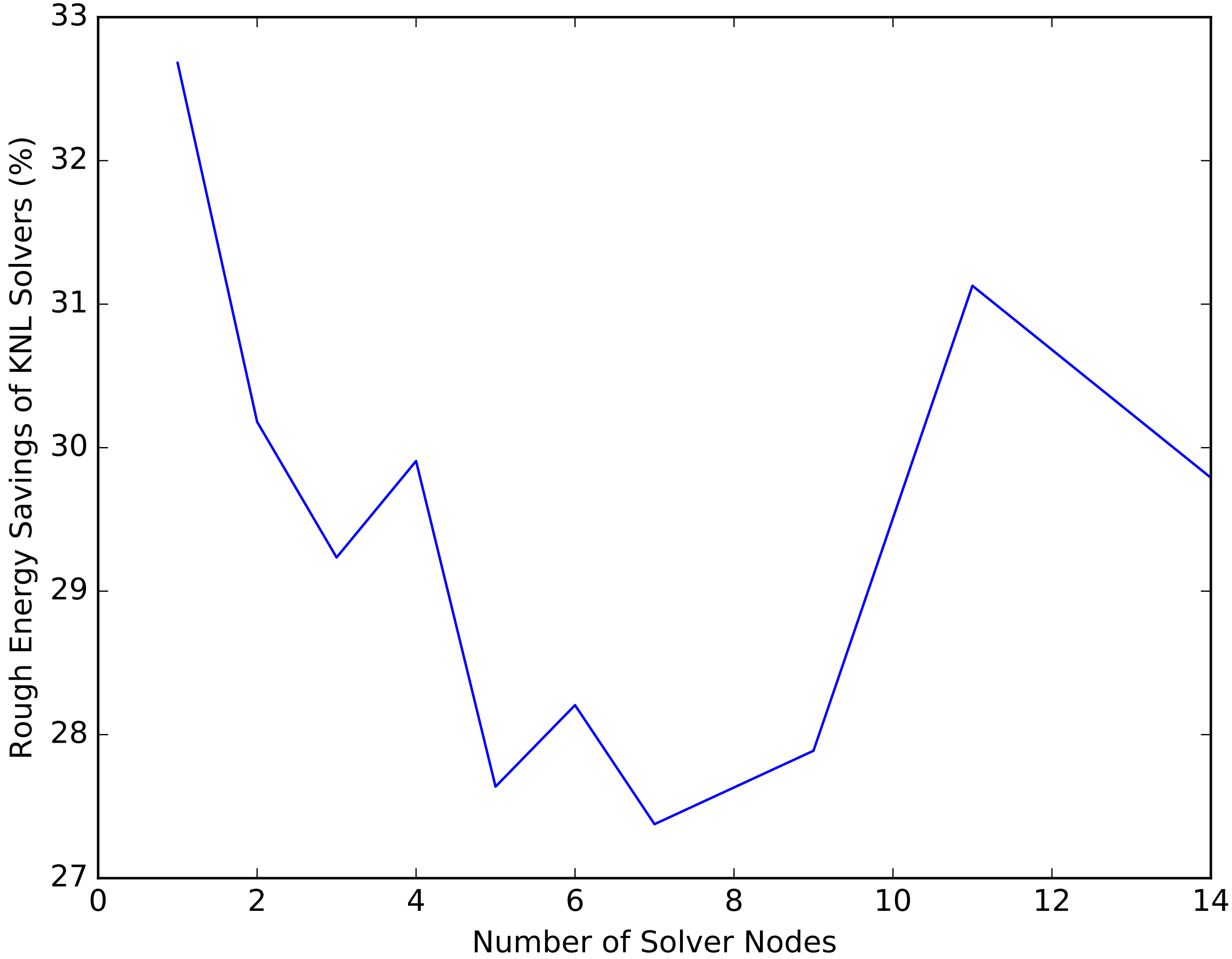
- **Used very rough approximates to estimate energy savings of hybrid run**
  - Execution times of Proof of Concept runs
  - TDP from spec sheets for each processor
    - Don’t do this
  - Assumed all else the same

$$E = \mathrm{TDP} * \mathrm{ExecutionTime}$$

Very Early TDP-Based Energy Results



Energy Savings Through Use of KNL Solvers



Questions?