

Enabling Extreme-Scale Simulations with Next-Generation Trilinos: Low Mach Fluid Case Study

Trilinos is an object-oriented software framework for the solution of large-scale, complex multi-physics engineering and scientific problems. When Trilinos was started in 1999, parallel codes used MPI and enumerated global objects using 32-bit integers. The latter limited problems to less than 2.1 billion unknowns, which now severely restricts simulations' fidelity. MPI-only parallelism also no longer suffices on modern architectures. The next-generation Trilinos can solve arbitrarily large problems, enabling extreme-scale simulations. It also provides a path forward for hybrid (MPI+X) parallelism. A significant recent effort has been on the vertical integration of the new Trilinos with an engineering application code, involving the simulation of low Mach fluid flow. The engineering application has used it to solve a nine billion element large eddy simulation (LES) problem on unstructured meshes with a 27 billion row matrix on 131,072 cores of a Cray XE6.

Low Mach Fluid Flow Case Study

- Integration of next gen Trilinos in an engineering app code: SIERRA thermal fluids low Mach module (Nalu)
- Equations: Variable density, unity Lewis number, acoustically incompressible (pressure projection), subgrid closure models
- Discretization: Control Volume Finite Element (CVFEM), Edge-based Vertex Centered (EBVC)

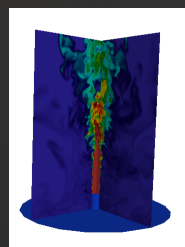


Hydrocarbon JP-8 pool fire: experiment on left, simulation on right

Results

Test case

- Mixture fraction-based turbulent open jet ($Re=6600$)
- Pressure Poisson equation (PPE), momentum, mixture fraction, turbulent kinetic energy (TKE)



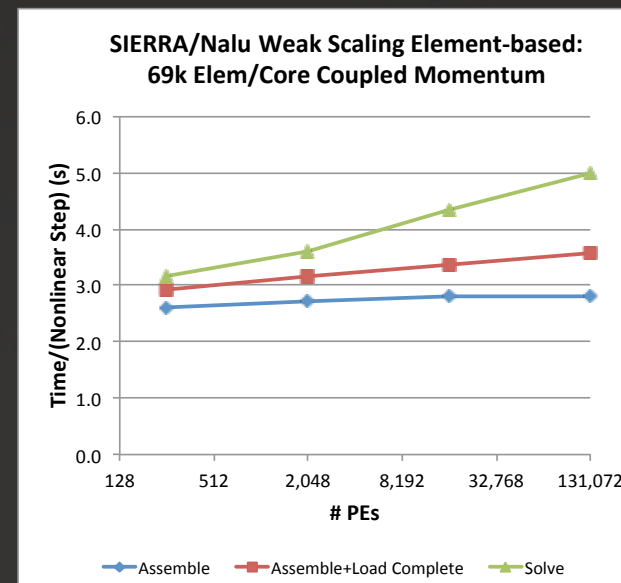
Simulations performed on Cielo Cray XE6: 8944 compute nodes (dual-socket 2.4GHz 8-core AMD Magny-Cours, total 142,000 cores); interconnect: Cray Gemini 3D torus



Results Summary

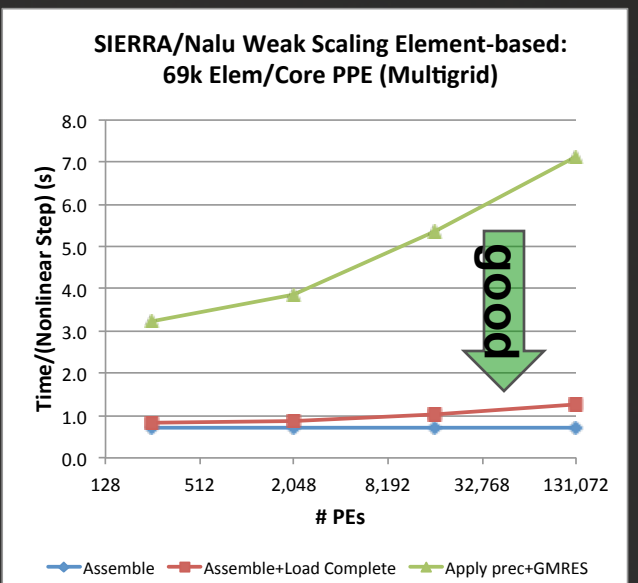
- Enabled 9 billion element, 27 billion row matrix simulation (unstructured mesh)
 - For large eddy simulations on unstructured meshes, rivals any other performed in the field
- Performance of next generation stack competitive with current stack
- Arbitrarily large global objects will enable unprecedented fidelity simulations
 - Enable new science

Weak scaling up to 27 billion row matrix and 131,072 cores



- SGS preconditioned GMRES
- time/(GMRES iteration) (not plotted) almost flat

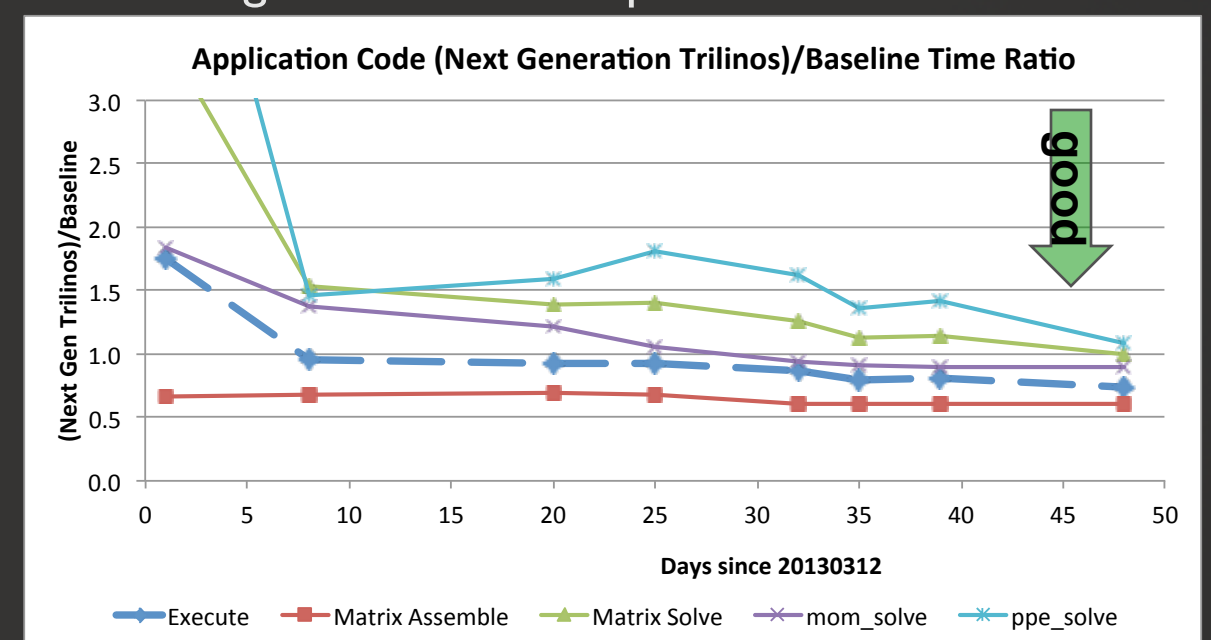
Weak scaling up to 9 billion row matrix with multigrid



- Multigrid preconditioned GMRES

Integration Process and Performance Improvements

Tracking Performance Improvements Over Time



First Cielo comparisons: Mar 13 code (edge-based, 140M elem, 2048 cores)

Lessons Learned

- Trilinos/application code integration considerably more challenging than expected
- Often performance issues cross multiple packages
- Critical for solver library developers to think about performance from the beginning
- Difficulties transitioning from research to production
- Performance analysis tools (e.g. Intel vtune) are critical
- Regular testing critical: correctness, performance, scaling

Future Work

- Optimize new solver stack for other use cases
- Mature additional next generation Trilinos packages
- Future architectures path forward: Trilinos Kokkos library
 - Kokkos: performance portable thread-parallel abstraction layer for HPC scientific/engineering codes
 - Code developers will implement application using Kokkos; Kokkos developers will handle different hardware instead of application developers

FUNDING AGENCY: NNSA/ASC

FUNDING ACKNOWLEDGEMENT: 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.

CONTACT: Paul Lin, Sandia National Labs, ptlin@sandia.gov