Linear Solvers Overview

Ulrike Meier Yang

LLNL

Linear Solvers Section Goals and Objectives

- Linear solvers needed by many DOE applications (icesheet and earth system simulations, plasma physics, tokamaks, accelerator science, CFD, MHD, etc)
- Provide direct and iterative solvers in open source libraries to enable solution of a variety of problems
- Research and development of new algorithms that are better suited for specific applications and/or new computer architectures
- Software development of math libraries that provide new solvers capable of solving more complex problems as well as efficient solvers on a variety of architectures
- Performance evaluation and improvement
- Support of DOE applications

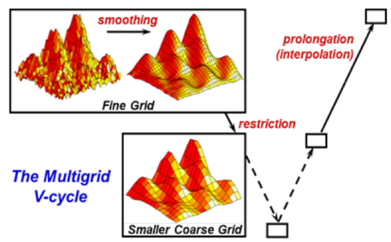
FASTMATH

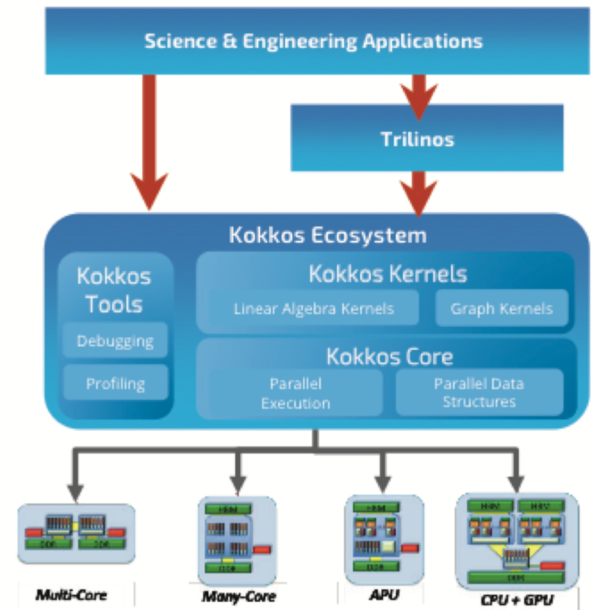
Linear Solvers Team

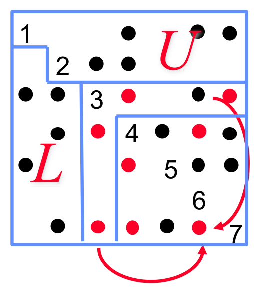
- M. Adams, LBNL
- P. Colella, LBNL
- P. Ghysels, LBNL
- M. Jacquelin, LBNL
- X. Li, LBNL
- Y. Liu, LBNL
- P. McCorquodale, LBNL
- E. Ng, LBNL
- L. C. McInnes, ANL
- B. Smith, ANL
- R. Falgout, LLNL
- R. Li, LLNL
- U. M. Yang, LLNL
- J. Hu, SNL
- S. Rajamanickam, SNL

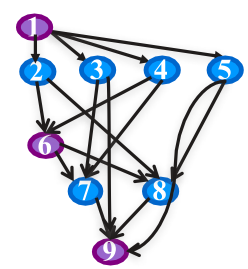
Linear Solvers Software

- KokkosKernels
- *hypre*
- MueLu (Trilinos)
- PETSc
- ShyLU
- STRUMPACK
- SuperLU
- symPACK
- ButterflyPACK (new!)









FASTMATH

ButterflyPACK Overview

- **ButterflyPACK** https://github.com/liuyangzhuan/ButterflyPACK
  - BSD licensed fast direct dense solvers for rank-structured matrices
  - Distributed-memory, OpenMP, Fortran2008 standard
  - Support H-matrix, HODLR formats with LR and butterflies.
  - C++ interface available through STRUMPACK
- **Kernel Functionalities**
  - Fast matrix compression, multiplication, factorization and solution
- **User Input**
  - Option 1: An element evaluation function for Z_{ij}
  - Option 2: A fast matrix-vector multiplication function $\boldsymbol{Zx}$ and $\boldsymbol{Z}^{*}\boldsymbol{x}$
- **Example Drivers**

https://github.com/liuyangzhuan/ButterflyPACK/tree/master/EXAMPLE

  - EMSURF_Driver.f90: 3D IE linear solvers
  - EMSURF_Eigen_Driver.f90: 3D IE eigen solvers
  - . . .

FASTMATH

SciDAC Partnership Involvement

- ProSPect (Probabilistic Sea Level Projections from Ice Sheet and Earth System Models) (BER)
  - AMG software in PETSC improved simulation times
- Energy Exascale Earth System Model (E3SM) (BER): ShyLU
- Exascale Catalytic Chemistry project (BES): KokkosKernels
- COMPASS (Community Project for Accelerator Science and Simulation) (HEP): SuperLU
- Center for High-Fidelity Boundary Plasma Simulationstokamak (HBPS) (FES): PETSc
- SCREAM (Simulation Center for Runaway Electron Avoidance and Mitigation)(FES): multigrid methods, PETSc
- ISEP (Integrated Simulation of Energetic Particles in Burning Plasmas) (FES): hypre
- CTTS: (Center for Tokamak Transient Simulations) (FES)
  - Improvements in SuperLU significantly improved simulation time

FASTMATH

CTTS-RAPIDS-FastMath Partnership

Samuel Williams, Sherry Li, Yang Liu, Nan Ding

Scientific Achievement

Improved sparse triangular solver performance on NERSC's Cori/KNL by 9x at 4096 MPI processes (64 nodes) and 3x at 64 threads on a single node. Achieved 40% speedup of NIMROD simulation through improved triangular solve in SuperLU_dist.

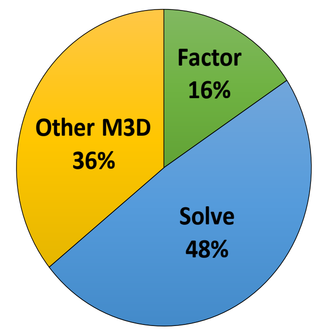
Significance and Impact

SuperLU Preconditioners are essential for the solvers in M3D-C1 and NIMROD. Solver performance is dominated by the sparse triangular (SpTS) preconditioner. When block preconditioners are configured to use a single process per solve, all MPI communication is eliminated, scalability is enhanced, and threading optimizations become paramount. Our performance improvements in SpTS directly improve application performance and scalability.

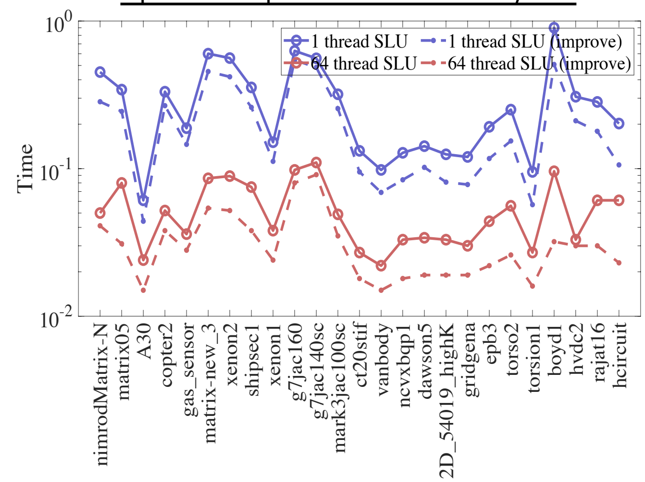
Research Details

- Collaboration between CTTS, RAPIDS, and FastMath.
- Analyzed M3D-C1 and NIMROD performance and thread scalability.
- Developed and deployed optimizations in SuperLU (SpTS) that improve performance and scalability on NERSC's Cori/KNL.

M3D-C1 run time on NERSC's Cori/KNL



Optimized SpTS on NERSC's Cori/KNL



FASTMATH TECH-X PPPL BERKELEY LAB

Accomplishments in new algorithm development

- Development of domain decomposition solvers
  - New solver FROSch (A Fast And Robust Overlapping Schwarz Domain Decomposition Preconditioner) has been developed and included in ShyLU; generalized Dryja Smith Widlund (GDSW) preconditioner that uses the architecture-aware Tpetra stack of Trilinos
  - Integration of the solver with icesheets code Albany/FELIX is in progress.
- Algebraic Multigrid Solver for Nonsymmetric Problems (poster)
  - The new AMG method pAIR (parallel approximate ideal restriction) shows significantly faster convergence and performance for advection dominated problems, as e.g. in transport, and is capable of solving some problems that could not be solved with conventional AMG before
  - Available in hypre

Challenges of and plans for future heterogeneous architectures

- Challenges:
  - New architectures require fine-grained parallelism, reduced communication and favor regular compute patterns
- Develop and take advantage of highly optimized performance portable kernels
- Reduce communication
- Improve scalability through repartitioning and ordering
- Increase structure in solvers

Architecture-Aware Kernels and Linear Solvers

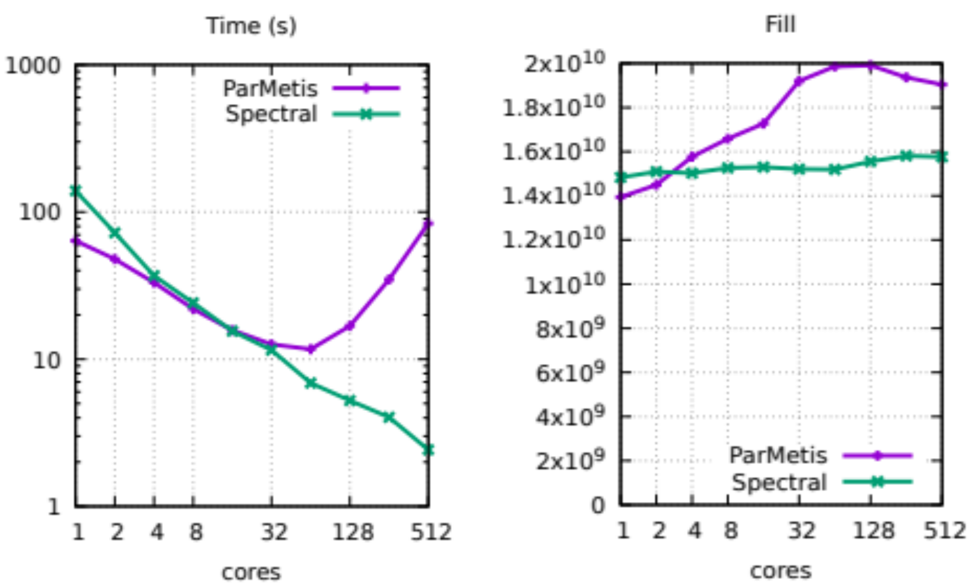
- Kokkos Kernels is a performance-portable on-node library for graph computations, and sparse/dense linear algebra kernels
  - A distance-2 graph coloring has been designed to achieve improved performance for important algebraic multigrid setup kernels
  - A new batched QR factorization has been implemented
- Method of Local Corrections: solves Poisson equation with infinite-domain boundary conditions, on adaptive grids. (poster)
  - High-performance implementation has 1/10th the communication cost of traditional multigrid iteration.
  - Achieves comparable solve times between HPGMG on a uniform grid with 1B grid points, and MLC on 1B grid points that are adaptively distributed, with much higher finest-level grid resolution
- Semi-structured interface and solver (poster)
  - Work has continued on increasing structure in hypre's semi-structured interface and the development of a semi-structured multigrid method that can take advantage of the new interface and is more suitable for highly parallel architectures, including GPUs

FASTMATH

Scalable Krylov and Multi-Level Methods for Spectral Graph Partitioning and Sparse Matrix Ordering

Scientific Achievement

A parallel and scalable spectral nested-dissection code for sparse matrix fill-reducing ordering has been developed, relying on an efficient multi-level scheme and state-of-the-art eigensolvers, such as communication-hiding Lanczos and a stable preconditioned LOBPCG implementation.



Results for the Queen_1417 matrix with 4.1M rows and columns and 333M nonzeros from Tim Davis' sparse matrix collection: (left) Our new spectral nested-dissection code scales much better than the widely used ParMETIS. (right) The fill in a sparse solver (~ memory usage) is similar to that obtained with ParMETIS when running on 1 core, but is less when running at scale, since the quality of ParMETIS degrades with increasing amounts of parallelism. Results obtained on NERSC Cori, Haswell, using 32 MPI processes per node.

Significance and Impact

Currently available parallel graph partitioners and ordering codes do not scale well, or quality degrades at scale, severely hindering performance in HPC codes, such as SuperLU, PETSc, MFEM, ... Spectral partitioning and ordering can leverage available parallel eigensolvers, and achieve good scalability and quality.

Research Details

– Partitioning based on Fiedler eigenvector: computed with a multilevel LOBPCG or Lanczos solver.

– Minimization of the quotient cut / conductance.

– Highly scalable hybrid MPI+OpenMP implementation.

– Demonstrated much improved scalability with quality similar to ParMETIS and PT-Scotch.

– Developing communication-hiding multilevel Lanczos and multilevel preconditioned LOBPCG Fiedler solvers.

FASTMATH

Work was performed at LBNL: P. Ghysels, M. Jacquelin, E. Ng, R. Van Beeumen
Universiteit Antwerpen, Belgium: Siegfried Cools

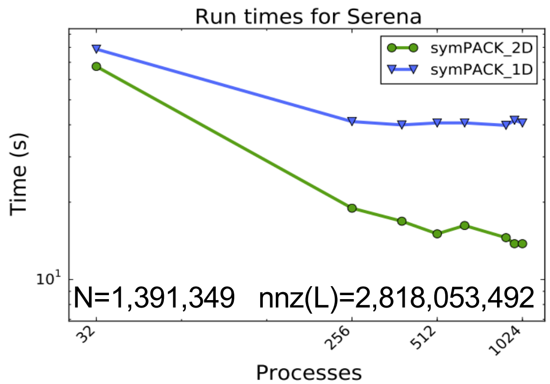
Universiteit Antwerpen

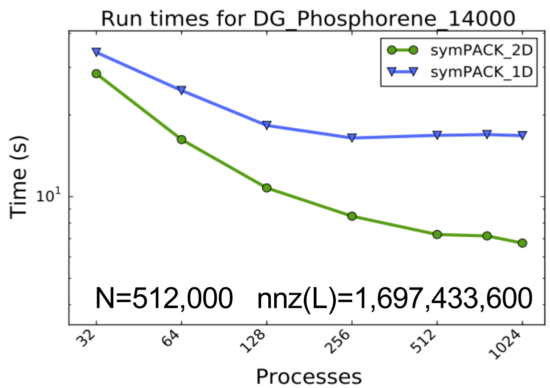
BERKELEY LAB

A new 2D task-based symPACK

Scientific Achievement

A new highly scalable 2D data distribution for the symPACK solver, a direct linear solver for sparse symmetric matrices. New distribution leads to much improved strong scalability and significant speedups over the previous 1D data distribution.





Strong scalability on Cori Haswell:
- Up to **3x** speedup for Serena
- Up to **2.5x** speedup for DG_Phosphorene_14000

Significance and Impact

– **Sparse direct linear solvers are at the heart of many HPC codes. When matrices are symmetric, fewer storage and computations are required.**

– **Factorization is a crucial preprocessing step to PEXSI, a library used in electronic structure computations.**

– **SymPACK 1D supernodal distribution balances flops and memory but lacks strong scalability, the new 2D distribution overcomes the issue.**

Research Details

– New task-based 2D data distribution

– Uses an explicit load balancing, not regular block cyclic mapping to allow more generic data mappings

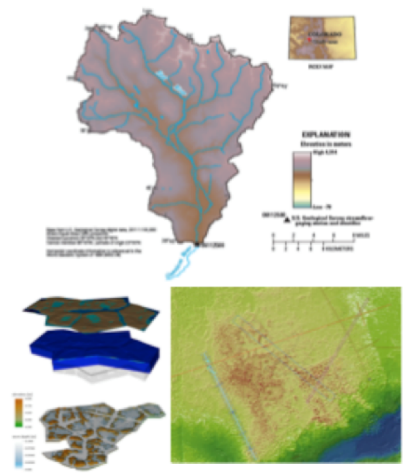
– Balances flops, memory

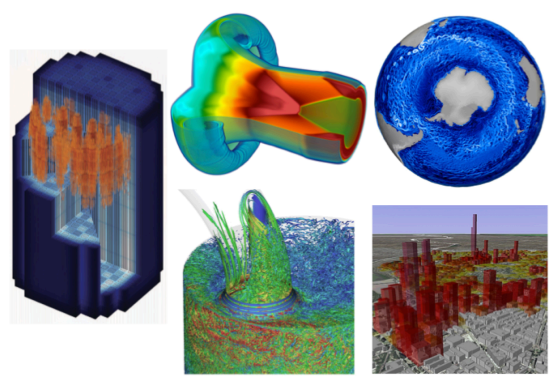
– New distribution leads to higher strong scalability

Work was performed at LBNL: Mathias Jacquelin and Esmond Ng; www.sympack.org www.pexsi.org

Impact on ECP

- FASTMath software is used in numerous ECP applications:
  - ATDM LANL apps
  - ATDM LLNL apps
  - ATDM SNL apps
  - ExaWind (predictive wind plant flow physics modeling)
  - ExaALT
  - ExaAM (additive manufacturing)
  - ExaSGD
  - ExaSky
  - ExaStar
  - Combustion-Pele
  - MFIX-Exa
  - NWChemEx
  - Subsurface
  - WDMApp





FASTMATH

CASC

Center for Applied Scientific Computing

Lawrence Livermore National Laboratory

LLNL-PRES-777297
This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC