Testing the SLATE Linear Algebra Library

Mark Gates

Innovative Computing Laboratory, University of Tennessee, Knoxville

Software Testing for Scientific Computing in HPC BoF
SLATE: Software for Linear Algebra Targeting Exascale

• Distributed, GPU-accelerated, dense linear algebra library
  • BLAS
  • Linear systems
  • Least squares
  • Eigenvalues, SVD, Polar Decomposition

• TestSweeper framework
  • Parses command line parameters
  • Loops over parameters
  • Print output in table
  • Testing & benchmarking, MPI aware
Scope of testing

• Approximately 70 high-level routines to test
  • **BLAS**: gemm, he/symm, he/syrk, he/syr2k, trmm, trsm, trtrm
  • **Auxiliary**: add, scale, scale_row_col, set, ge/he/sy/trnorm
  • **Linear solvers**: {gb, ge, pb, po, he}{sv, trf, trs, tri, condest} + variants
  • **Least squares**: gels, qr, mqr, lq, mlq, cholqr
  • Eig, generalized eig, svd, polar

• Most have backward error formulas; some require reference solution
• 4+ precisions
• 4+ platforms
Explosion of Combinations

- Each has $\approx 10$ parameters to loop over
  - Upper / lower, transpose, dimensions (m, n, k), block size, MPI grid (p, q), OpenMP threads, target (CPU, GPU), ...
- “Quick” check runs 4500 tests x 4 platforms
  - Small matrices
  - CPU, CUDA, ROCm, SYCL
- Need large matrices to stress implementation
- MAGMA’s full test suite can run > 24 hours
Platforms
Platforms

- GitHub CI
  - Test on in-house servers or cloud
  - GNU or Intel compilers, MPI
  - OpenBLAS, Intel MKL
Platforms

• GitHub CI
  • Test on in-house servers or cloud
  • GNU or Intel compilers, MPI
  • OpenBLAS, Intel MKL

• Build Platforms
Platforms

• GitHub CI
  • Test on in-house servers or cloud
  • GNU or Intel compilers, MPI
  • OpenBLAS, Intel MKL

• Build Platforms

• Target: HPC platforms
  • Summit: IBM ESSL, MPI, CUDA
  • Frontier: Cray compilers, LibSci, GPU-aware MPI, ROCm
  • Aurora: Intel compilers, SYCL
More Challenges

• False or unrelated failures
  • Random data yields ill-conditioned matrix ⇒ write better matrix generator
  • Software stack reconfigured ⇒ containers isolate changes
  • Contention with users ⇒ scheduler like slurm or move to cloud
  • Error slighty above tolerance: $2.5e-15 > 2.2e-15 = 10\varepsilon$ FAILED

• Solving these distracts from PR