KAUST BASIC LINEAR ALGEBRA ROUTINES ON GPUs

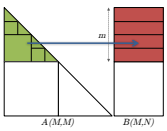
KBLAS

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

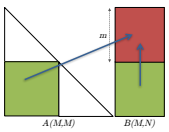
Extreme Computing Research Center

KAUST BLAS (KBLAS) is a high performance CUDA library implementing a subset of BLAS as well as Linear Algebra PACKage (LAPACK) routines on NVIDIA GPUs. Using recursive and batch algorithms, KBLAS maximizes the GPU bandwidth, reuses locally cached data and increases device occupancy. KBLAS represents, therefore, a comprehensive and efficient framework versatile to various workload sizes. Located at the bottom of the usual software stack, KBLAS enables higher-level numerical libraries and scientific applications to extract the expected performance from GPU hardware accelerators.

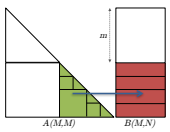
RECURSIVE ALGORITHMS: TRMM and TRSM



Rec. TRMM: $B_1 = \alpha A_1 B_1$
Rec. TRSM: $A_1 X_1 = \alpha B_1$

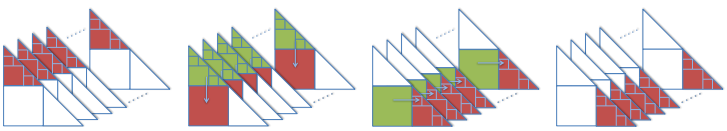


TRMM: GEMM $B_1 = \alpha A_2^T B_2 + B_1$
TRSM: GEMM $B_2 = \alpha B_2 - A_2 B_1$



Rec TRMM: $B_2 = \alpha A_3 B_2$
Rec TRSM: $A_3 X_2 = B_2$.

BATCH ALGORITHMS: Recursive Cholesky POTRF



1. Rec-POTRF
2. Rec-TRSM
3. Rec-SYRK
4. Rec-POTRF

KBLAS 2.0

- Legacy Level-2 BLAS: (↕◇∞○) SYMV, GEMV, HEMV.
- Legacy Level-3 BLAS: (↕◇∞○) TRSM, TRMM, GEMM (∞ only).
- Batch Level-3 BLAS: (↕◈∞○= ✻) TRSM, TRMM, SYRK.
- Batch Triangular: (◈↕∞○= ✻) TRTRI, LAUUM.
- Batch Symmetric: (◈↕∞○= ✻) POTRF, POTRS, POSV, POTRI, POTI.
- Batch General: (◇↕∞○= ✻) GESVJ, GERSVD, GEQRF.

↕ Standard precisions: s/d/c/z.
↕ Real precisions: s/d.
◈ Very small matrix sizes.
◇ Arbitrary sizes.
○ Single-GPU support.
∞ Multi-GPU support.
= Uniform batch sizes.
✻ Non-Strided and Strided variants.

KBLAS HIGHLIGHTS

- KBLAS Level-2 (○): SYMV & HEMV
- KBLAS Level-3 (○): TRMM & TRSM

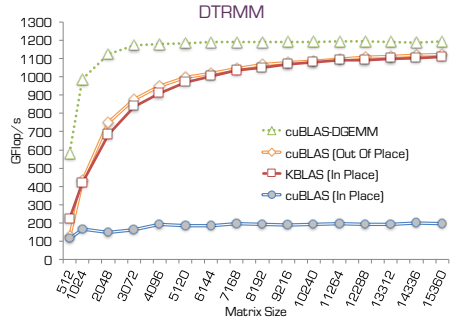
NVIDIA cuBLAS 6.0

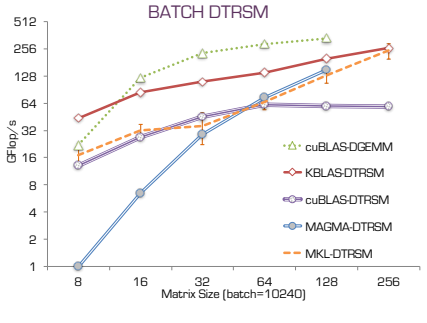
NVIDIA cuBLAS 8.0

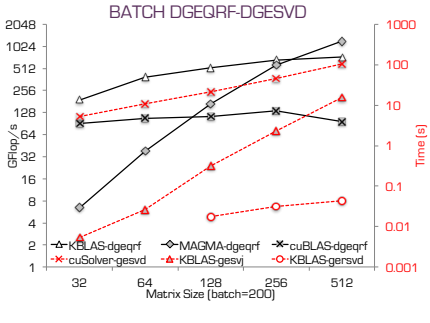
CURRENT RESEARCH

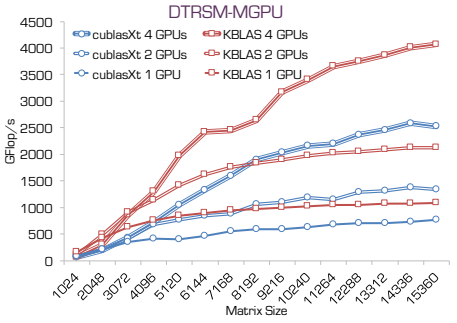
- Half Precision Legacy and Batch BLAS.
- Tile Low-Rank (TLR) BLAS on GPUs.
- Adaptive Cross Approximation (ACA) on GPUs.
- Vectorized Batch BLAS on x86.

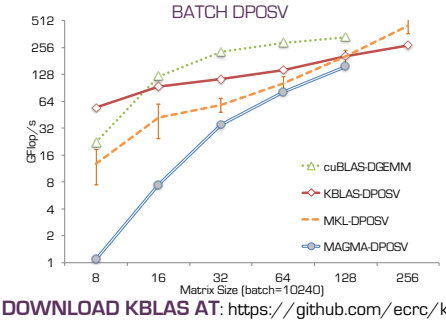
PERFORMANCE RESULTS

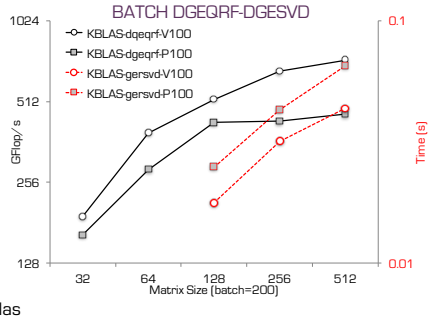












DOWNLOAD KBLAS AT: https://github.com/ecrc/kblas

A collaboration of

With support from

Sponsored by