

LBL Updates

February, 2022

Topics

- Improved GPU-accelerated Factorization
- NERSC Hackathon on SuperLU GPU optimization
- Q&A
- Nan should return to work this month and continue development of the multi-GPU trisolve on Perlmutter(NVIDIA) & Spock(AMD)

Optimizations for GPU-accelerated code

- Achieving GPU speedups for M3DC1 matrices is challenging
- Towards moving everything in numerical factorization to GPU
- Larger Jacobi blocks desirable
- U-solve under-development in the GPU-3d-opt branch

matrix	CPU only (master)	GPU GEMM (master)	GPU GEMM + scatter (master)	GPU GEMM + scatter + panel factor (GPU-3d-opt)	GPU GEMM + scatter + panel factor + diagonal factor (GPU-3d-opt)
s1_mat_0_1269 36	2.639	2.475	2.94	2.14	under development
s1_mat_0_2538 72	8.3	7.9	8.36	5.14	under development

Performance comparison: 5 OMP threads vs 1 V100 on Cori GPU nodes

Q&A

1. Memory on Perlmutter...

https://docs.nersc.gov/systems/perlmutter/system_details/

256GB per 64c EPYC (4GB/core)

4 x 40GB A100 GPUs

can run out of memory

2. Stellar uses 8GB/core