The 38th International Symposium on Lattice Field Theory

LATTICE 21

JULY 26-30 2021, ZOOM/GATHER@MIT

Contribution ID: **598** Type: **Oral presentation**

Strong scaling RHMC on NVIDIA GPUs

Wednesday, 28 July 2021 13:15 (15 minutes)

The ability to strong scale is crucial for Lattice QCD simulations. Therefore Lattice QCD has been constantly craving for higher network and memory bandwidths. While never enough well-balanced systems with favorable GPU-to-network ratios are available, e.g. with the Juelich Booster. However, API overheads and necessary synchronizations between GPU and CPU have become prohibitively expensive, not keeping up with generational improvements of GPUs and networks. This limits the ability to strong scale with MPI communication. A shift towards fine-grained GPU-centric communication provides a way out as it completely removes these bottlenecks by moving the communication to the GPU kernels. Since version 1.1 QUDA implements GPU-centric communication for NVIDIA GPUs using NVSHMEM. We will show low-level Dslash results as well as full RHMC scaling results on modern GPU systems like Selene and the Juelich Booster and discuss further expansions of this approach to even more latency-limited algorithms as Multigrid.

Primary authors: WAGNER, Mathias (NVIDIA); CLARK, Kate (NVIDIA); TU, Jiqun (NVIDIA Corporation); WEINBERG, Evan (NVIDIA Corporation)

Presenter: WAGNER, Mathias (NVIDIA)

Session Classification: Software development and Machines

Track Classification: Software development and Machines