Contribution ID: **463** Type: **Talk**

Rearchitecting QUDA for Multi-RHS Computation

Tuesday, 30 July 2024 15:05 (20 minutes)

It is well-known that computers are no longer getting faster, and only more parallel and hierarchical. Moreover, computations are increasingly bandwidth limited; and with the advent of the end of Moore's Law, often power limited as well. This requires us to rethink how we deploy LQCD computations to maximize science throughput. In this talk we discuss the rearchitecting of QUDA for batch computation to expose more parallelism and locality. Performance results are shown for a variety of workloads, including linear solves, multigrid, and contractions.

Primary author: Dr CLARK, Kate (NVIDIA)

Co-authors: Dr JOÓ, Bálint (NVIDIA); Dr WEINBERG, Evan (NVIDIA); Dr TU, Jiqun (NVIDIA); Dr WAGNER, Mathias (NVIDIA)

Presenter: Dr CLARK, Kate (NVIDIA)

Session Classification: Software development and machines

Track Classification: Software Development and Machines