LATTICE 2024

LIVERPOOL

Contribution ID: **150**   Type: **Talk**

Lattice QCD on the NVIDIA Grace-Hopper architecture

Tuesday, 30 July 2024 13:45 (20 minutes)

The NVIDIA Grace Hopper Superchip architecture, powering upcoming supercomputers like Alps at CSCS, Jupyter in Juelich, Isambard-AI in the UK, and Venado at Los Alamos National Laboratory, offers significant advancements for Lattice QCD applications. Integrating the ARM-based Grace CPU with the Hopper GPU via NVLink-C2C, it provides 7x the bandwidth of PCIe Gen5, enabling coherent memory and efficient data transfer. This architecture's high system memory bandwidth (up to 500 GB/s) and NVLink-C2C bandwidth (900 GB/s) and high GPU memory bandwidth (4TB/s) enhance performance for Lattice QCD. Its balance also helps legacy applications with significant CPU components. This talk will present performance results for QUDA-accelerated workloads like MILC and Chroma, discuss how specific features of Grace Hopper can benefit Lattice QCD and show performance of Lattice QCD applications on the NVIDIA Grace CPU. Furthermore, extracting more science can be achieved by co-scheduling different parts of the workflow on CPU and GPU concurrently.

Primary author: WAGNER, Mathias (NVIDIA)

Presenter: WAGNER, Mathias (NVIDIA)

Session Classification: Software development and machines

Track Classification: Software Development and Machines