Luke Olson Fine-grained parallelism in algebraic multigrid and sparse matrix operations for efficient GPU execution

Department of Computer Science
University of Illinois at Urbana-Champaign
Urbana
IL 61801
lukeo@illinois.edu
Nathan Bell
Steven Dalton

Next generation high-performance computing systems are expected to include high-throughput processors such as GPUs. While algebraic multigrid and the associated sparse matrix operations have been used successfully in a variety of settings, previous techniques do not generally develop the fine-grained data parallelism required for efficient GPU execution. In this talk we detail algebraic multigrid and sparse operations on (multiple) GPUs and discuss directions for further efficiencies. Sparse matrix operations such as the matrix-matrix multiply have long been a bottleneck in the algebraic setup and we present high-throughput optimizations to this approach along with other operations that contribute to the performance in a single and multiple GPU setting.