Accelerated architectures, exascale computing, and high-order finite element methods

Noel Chalmers

The use of accelerators has become ubiquitous in the largest supercomputers in the world, typically providing much higher main memory bandwidth and floating-point operations per second (FLOPS) than their host CPUs. In modern accelerators, specialized sections of hardware, such as matrix cores, further accelerate certain computational workloads leading to increased FLOPS throughput and power efficiency. For finite element methods, which are very often bound by main memory bandwidth, such specialized hardware does not appear to offer much potential performance benefit or is not easily leveraged. On the other hand, high-order methods have operations which naturally map to matrix cores and see significant performance gains. In this talk, I will discuss performance modelling and bottlenecks of high-order finite element operations on accelerators, the details of matrix cores and their architecture, and show how the use of matrix cores can significantly improve the performance of some of these operator kernels.