GPU Considerations of ESMs for the Pre-Exascale HPC Phase

Computational efficiency with increased focus on performance-per-energy-cost has become the overarching driver behind system architectural considerations for Exascale HPC. The development of scientific and engineering application software that achieves full potential from such Exascale systems will require a level of hardware/software co-design that is a relatively new approach in conventional HPC. GPUs are a centerpiece of the first pre-Exascale systems announced in the USA under the U.S. Department of Energy CORAL partnership, and are a key component of co-design collaborations that will strive for accelerated application readiness by system deployment in 2017.

This presentation will first examine the motivation and progress of GPU-based heterogeneous system architectures for the pre-Exascale phase of HPC. The second topic will introduce the requirements for extraction of fine-grain parallelism of application software and current state of GPU-accelerated Earth system models with review of the programming strategies deployed. Select examples will provide relevance to science-scale HPC practice that quantifies the benefits of heterogeneous vs. CPU-only computing. In addition to GPU use with x86 CPUs starting from the mid-2000’s, the POWER and ARM-64 CPU architectures have become available alternatives starting in 2014. The third and final topic will provide roadmaps of GPU hardware and system software, and interoperability with these new host CPU platforms.