Contribution ID: **14** Type: **30 minute oral presentation**

Accelerating Weather Prediction with NVIDIA GPUs

Friday, 28 September 2018 10:00 (30 minutes)

In this talk we will present collaborative work performed as part of the EU-funded ESCAPE project (led by ECMWF) aimed at modernising numerical weather prediction applications for efficient utilisation of modern parallel architectures. The project features the use of "dwarves" designed to represent the main computational properties of the real applications, and this talk will focus on GPU enabling and optimisation for two dwarves. The first captures the key algorithmic patterns from the spherical harmonics techniques used by global modelling applications such as the Integrated Forecasting System (IFS), which combines Legendre and Fourier transforms on a spherical grid. The second represents the "MPDATA" advection scheme and uses an unstructured grid.

For single-GPU benchmarks, we have improved performance, over pre-existing suboptimal GPU implementations, by around a factor of 17 for the spherical harmonics dwarf and a factor of 57 for the MPDATA dwarf. Optimizations include the improvement in memory management and minimization of data movement, algorithmic redevelopment to allow more flexible exposure of parallelism, minimization of launch overhead, data layout and loop structure adaptations to improve memory access patterns and coalescing, and efficient usage of libraries. Performance modelling shows we are approaching maximum available hardware performance. We have performed substantial work to allow efficient multi-GPU scaling, and results will showcase the recently announced DGX-2 server featuring the high-bandwidth NVSwitch interconnect which provides an excellent match for the all-to-all communications required by the algorithms.

We will present our work in a manner to allow our analysis and optimisation techniques to be easily transferable for other researchers working on their own applications.

We will also briefly describe other weather-related projects NVIDIA is involved with including COSMO, MPAS and WRF, as well as research into exploiting artificial intelligence in this domain.

Affiliation

NVIDIA

Primary authors: Dr GRAY, Alan (NVIDIA); Dr MESSMER, Peter (NVIDIA); Mr POSEY, Stan (NVIDIA)

Presenter: Dr GRAY, Alan (NVIDIA)

Track Classification: 18th Workshop on high performance computing in meteorology