5.5 Acceleration of WRF on GPU.

Abdi, Daniel, Sam Elliott, S. M. Iman Gohari, Don Berchoff, Gene Pache and John Manobianco, *TempoQuest*

The Weather Research and Forecasting (WRF) Model is a mesoscale numerical weather prediction system designed for both atmospheric research and operational forecasting applications. TempoQuest (TQI) has achieved an acceleration of WRF computation using Graphical Processing Units (GPUs) that will benefit tens of thousands of users around the world. The TQI goal is to accelerate WRF by 10 times when compared to CPU WRF to reduce operationally run time and increase weather prediction precision and accuracy with the introduction of denser weather data sets.

TQI ported the WRF model to run entirely on the GPU. TQI obtained accelerations of up to 5 times for a typical WRF configuration, and up to 7 times in certain configurations using the NVIDIA Volta GPU. The most time-consuming modules, such as radiation computation kernels, are optimized for best performance; while other less compute intensive modules were ported without much emphasis on optimization. TQI believes the 10 times acceleration is achievable through further optimization. TQI tested the performance against the CPU in a socket-to-socket comparison on standard WRF test cases such as CONUS 2.5 and 12 KM with different dynamics and physics options. In this presentation, TQI will provide performance results on single and multi-node machines.