Geosci. Model Dev. Discuss., 7, C3711–C3713, 2015
www.geosci-model-dev-discuss.net/7/C3711/2015/

Geoscientific Model Development Discussions

Interactive comment on "Development of efficient GPU parallelization of WRF Yonsei University planetary boundary layer scheme" *by* M. Huang et al.

Anonymous Referee #2

Received and published: 6 May 2015

Summary:

This paper presents the GPU porting and optimization process for a sub-process of the WRF model. The authors describe the mathematical background of the existing model sub-process, then delve into the wide variety of optimizations made to achieve performance from the GPU version of the code. The authors compare the results of each successive optimization to the performance of the original CPU code.

General Comments:

This paper does a good job presenting the various optimization techniques used to achieve a clearly excellent speedup result. The optimizations used are described well,

and will likely be useful in other domains beyond the scope of accelerating the YSU PBL scheme.

Questions:

Is there a specific reason that the Yonsei University PBL scheme was chosen for acceleration instead of any other scheme (none of which were named)? Is this the most popular model, or perhaps the most amenable to GPU acceleration? I believe some discussion of other schemes may be warranted, at least as motivation for why the YSU scheme was chosen.

The paper describes a direct mapping from the benchmark dataset's spatial grid size to the implementation's use of thread blocks. Will the accelerated code be easy to apply to other domains or resolutions and other datasets? Has some other test dataset been used to examine speedups for a different test case?

Is this work intended to be incorporated in some future release of WRF? Is there possibly a timeline for when an accelerated (or partially accelerated) version of WRF is available?

How much does the impressive speedup obtained for the PBL scheme improve the performance of the full WRF model? It would be good to see some timing data for the full model with the accelerated PBL scheme incorporated.

Technical Corrections:

Page 8033, line 18: "GPU-accelerated longwave radiation scheme of the rapid radiative transfer model for general circulation models"

Page 8034, line 8: "which is one of the physical models in WRF."

Page 8041, line 12: "built in the WRF model."

Page 8042, line 15: "The driver, in the C language,"

Page 8042, line 20: "into the memory of the CPU."

Page 8042, line 21: "From the viewpoint of CUDA Programming,"

Page 8043, line 10: "contiguous data, and are aligned in memory."

Page 8043, line 24: "Three major reasons for doing this in this way are"

Page 8043, line 26: "CUDA C programs in a short time."

Page 8044, line 3: "accross the entire US. The WRF domain is"

Page 8045, line 6: "one way to do this is to call"

Page 8045, line 20: "to a parallel GPU basis in the next section."

Page 8046, line 6: "one CPU core of an Intel Xeon E5-2603."

Page 8047, line 22: "to the structure of the WRF model,"

Page 8049, line 3: "per thread at 63."

Page 8049, line 15: "make execution more efficient."

Page 8049, line 16: "of the GPU architecture."

Page 8051, line 19: "In the plot, x starts from value"

Page 8062, Table 6a: Column headings need some label relating to what they represent, as opposed to just the description in the caption.

Page 8063, Table 6b: Same as for Table 6a, needs column headings.

Page 8076, Figure 12: Include labels for the x-axis of the graph, instead of just the range mentioned in the paper.

Interactive comment on Geosci. Model Dev. Discuss., 7, 8031, 2014.