

Interactive comment on “Optimizing High-Resolution Community Earth System Model on a Heterogeneous Many-Core Supercomputing Platform (CESM-HR_sw1.0)” by Shaoqing Zhang et al.

Mark Govett (Referee)

mark.w.govett@noaa.gov

Received and published: 25 March 2020

General Comments:

This was a very well organized and written paper. The paper described efforts to port a large legacy, climate code to the Sunway TaihuLight system. The unique architecture of the Sunway processor was described which helps the reader understand some of the changes that were needed for the climate application to run efficiently. The original code was designed and run on Intel-based processors. Performance and baseline

C1

scientific results were made prior to porting the code. The work described detailed efforts to optimize the code, while maintaining sufficient accuracy in the solutions. As the authors admit, determining an acceptable level of scientific accuracy is an ongoing process determined by many factors. With a complex scientific application, thorough testing and evaluation using multiple criteria is needed to build confidence in the solution.

Specific Comments:

The authors describe extensive efforts to optimize the code, which included many common techniques. You spent a lot of time optimizing for the Sunway processor, but did not apply that level of effort to the original code. Some of the changes could have been applied to original code, making the comparison of performance more fairly represented.

The work appears specifically designed to target a single system with a unique processor. Was performance portability considered as a factor in this effort? Could the modified code run on an Intel-based system and how did the results compare to the original code. Given the fine-grain nature of the parallelization, would GPUs or high core count CPU processors be a target for this work? Addressing performance-portability would make the impact of this work much greater than the results you achieved targeting a single system.

The performance impact of the different types of code optimizations you made were not described. This would be a useful way to determine the tradeoff between portability and performance. For example, specific optimizations described in the Stage 1-3 optimizations were closely aligned to the Sunway processor. How much performance benefit were there for each of the stages and was it applied to a large portion of the code?

Regarding porting the code to the TaihuLight system, your approach seemed to be first port the code to two Intel based systems (TAMU and QNLM). It was unclear why

C2

you felt the need to port to both systems. Further, your comparisons were made after only 9 timesteps seemed arbitrary and perhaps not sufficient. Please include some justification. It was also unclear what fields were compared in the UF-CAM-ECT test. A summary of the relevant details from the paper would be useful here.

Are there references for the tools given in Table 1? Most of the tools listed were not referenced in the manuscript. They should be either introduced to the reader in the paper if they add value to the manuscript. For example, you could state how you used them and how it helped identify

Technical Corrections:

Line 103-105: Awkward sentence. Perhaps break into two sentences?

Line 140: Old reference (Govett, 2010) should be replaced with a more comprehensive paper (Govett, et al. 2015) in the Bulletin of the AMS: <https://doi.org/10.1175/BAMS-D-15-00278.1>

Lines 212-218: The authors don't describe what the speed and bandwidth of connections between super-nodes, within a cabinet, and between cabinets. This is essential in understanding the limitations of the Sunway TaihuLight system at scale.

Line 243: change "details Section 3.3" to: "detail in Section 3.3"

Line 245-250: unclear if CPE based parallelism is with MPI or something else??? Line 251: change "details" to "detail"

Lines 260-266: It appears that an MPI-based intelligent programming model would work here. The intelligence would be knowing when comms within a CPU task group or to an MPE are needed.

Lines 268-272: Regarding the power efficiency comparison in Table 2, is this a fair comparison? It seems like there would be size benefits favoring the larger systems especially in terms of infrastructure required no matter the size of the system.

C3

Line 307: *swlu* was italicized but not introduced in section 2.1.2. It appears in capitals as SWLU in Table 1. .. Line 332: TAMU and QNLM are undefined (QNLM is defined on line 602, TASMU on line 603)

Line 341: It seems that such short runs are not sufficient. Did you make a similar test with more than nine time steps. There remains a high potential for variations to show up later in the simulation experiments

Optimizations were made to achieve 1 SYPD on the Sunway system. Did you attempt to incorporate these changes and optimization techniques back into the original model?

Line 372: Can you provide more details or analysis regarding why -O2 fails but -O3 passes How long were the runs made before comparisons were done?

Interactive comment on Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2020-18>, 2020.

C4