Geosci. Model Dev. Discuss., 7, C1146–C1148, 2014
www.geosci-model-dev-discuss.net/7/C1146/2014/

Geoscientific Model Development Discussions
Open Access

Interactive comment on "Implementation and scaling of the fully coupled Terrestrial Systems Modeling Platform (TerrSysMP) in a massively parallel supercomputing environment – a case study on JUQUEEN (IBM Blue Gene/Q)" *by* F. Gasper et al.

Anonymous Referee #2

Received and published: 16 July 2014

This paper reviews the weak scaling of the Terrestrial Systems Modeling Platform (TerrSysMP), a 3-component (COSMO, CLM, ParFlow) coupled system on the IBM BlueGene/Q supercomputer JUQUEEN. It describes the limitations of the previous version of the OASIS coupler, the standalone external OASIS3, and shows how the new version of the coupler, OASIS3-MCT, solves these limitations. However, even in this last case, the paper describes how the initialization of the CLM component eventually prevent the scaling at high number of cores. The paper is nicely written and pro-

vides important information about the scaling of coupled systems on massively parallel computers. However some sections needs clarifications as detailed in the comments below.

Major comments

-Abstract: The abstract is, I think, too general. The specific findings of this study should be summarized in the abstract.

-p.3555, l.10-13: I suppose that the “minimal wait states” described here corresponds to the “optimized balancing” discussed in section 3.2 and Fig. 3B? If so, the link should be made; if not, then I do not understand how the “minimal wait states” was reached.

-p.3557, l.16: The mapfile, assigning MPI ranks to actual CPU cores to optimize the communication pattern on the 5-D torus, is mentioned there but its impact, that may be I think extremely important, is not described in the rest of the paper. More detail on how this mapfile was defined, if so, for the weak scaling experiment should be provided.

-p.3558, l.5-12: Even if one can understand that in Fig. 3a the load is not ideally balanced while in Fig. 3b, the load is improved, it is not clear to me what is the method or principle to follow to improve the load balance. More details should be provided on this important issue.

-p.3559, l.2: I suppose that the fact that only 2304 (CLM, ParFlow) and 1152 (COSMO) number of grid cells were reached is linked to the drastic increase of initialization time of CLM described on p.3561. If so, the link between the two sections of the paper should be made; it not, it should be explained why more grid cells were not reached.

-p.3559, l.8-11: This sentence is not clear at all. I suppose that it means that if the workaround is applied and if a reduced number of processes per node (nnpn) is defined, than this applies to all component models in the coupled system, even the ones not requiring it? This is hard to understand at this point in the text as the workaround is not described yet.

GMDD

7, C1146–C1148, 2014

-p.3573, Fig.5: the legend on the right of the Fig. 5 is very hard to read.

Minor comments

-p.3551, l.19: I do not understand the sentence “remain independent supporting interchangeable executables as a major advantage”; even if the OASIS approach allows to change the different executables of a coupled system, it of course does not guarantee that two executables are interchangeable.

-p.3552, l.22: Writing that “each MPI-process can only access 1GB” is a bit mis-leading as there exists a workaround based on the number of processes per node (nppn) is detailed later on p.3559.

-p.3555, l.26: “T” and “nbn” appearing in formula (1) are not explained.

Interactive comment on Geosci. Model Dev. Discuss., 7, 3545, 2014.