

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 #1

Received and published: 30 June 2014

1 additional information desired for

- i/ section 2.4, page 3555, scaling study setup (2): processor-distribution is based on powers of 2 which is said to imply a non optimal usage of allocated compute resources. Further explanation for this fact would be good.
- ii/ section 3.1, page 3557: compiler options guiding the compiler to make use of C1022

architecture-specific benefits are mentioned that give a speedup of roughly two. Since this is an impressive number and comes for free (no src code modifications needed), the reader should be informed about the options to reproduce or try for his own code.

- iii/ sections 3.2 / 3.4: the weak scaling tests were performed for an idealized 24h simulation (cf. figure 5) but the optimal load balancing seems to be determined for a much shorter simulation (cf. timings within figure 3). Why is a short simulation done a-priori representative for a whole day simulation? What if the setup changes over time (e.g. different parametrizations and/or model internal schemes)? What if different physical aspects appear over the time of 24 simulated hours - does this change the optimal load balancing?
- iv/ section 3.2, page 3558, last paragraph: the improved load balance was found for a "characteristic test case", how is this case related to the fully coupled weak scaling tests done later on? What does the improved load balancing look like? Is it one of the design described in table 1?
- v/ figure 3: the impact of the improved load balancing is not equal for the whole MPMD setup. CLM (`program_off`) and ParFlow (`main`) times decrease by a factor of 4, whereas COSMO (`lmorg->organize_dynamics` and `lmorg->organize_physics`) times increase. Maybe this should be explained further by comparing the used resources for the experiments ran for figure 3 (which might also explain the topology plot in the third column of the cube view).
- vi/ in general: file I/O was disabled as far as possible for the scaling tests. It would be good to have at least a brief comment on the changed scaling behaviour if file I/O is turned on. Does I/O for example hinder scaling by introducing additional synchronisation points for all MPI-tasks? Would it be possible to use an adjusted load balance / task mapping to overcome some issues related to file I/O?

2 clarifications needed

- i/ figure 3: the binaries in the second column of the cube viewer should be matched to the models (`lmorg = COSMO`, `main = ParFlow`, `program_off = CLM`) to see the differences when task distribution is changed

3 typos

- i/ section 2.4, page 3555, line 18: "... powers of two are also used for the ..."
- ii/ table 1(a), page 3567: #processors for scaling step 1 should be "24x16/8x8/8x8" according to the text
- iii/ table 1(b), page 3567: #processors for scaling step 1 should be "16x16/16x8/16x8" according to the text

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