Geosci. Model Dev. Discuss., 8, C2326–C2331, 2015
www.geosci-model-dev-discuss.net/8/C2326/2015/

Geoscientific Model Development Discussions

Interactive comment on "Towards convection-resolving, global atmospheric simulations with the Model for Prediction Across Scales (MPAS): an extreme scaling experiment" *by* D. Heinzeller et al.

Anonymous Referee #2

Received and published: 1 October 2015

General comments

This article is made up of three distinct components;

1. Scaling experiments across four HPC facilities and three model configurations, including two with mesh refinement over Africa.
2. Comparison of results from a regional model to global (regular grid), global with refinement and observations over West Africa.

3. An 'extreme' scaling experiment.

The first component, the scaling and performance analysis, is a valuable addition to the literature, and with some clarifications and discussion of the limitations of this analysis I think should be published. Performance statistics across different systems are vital for planning simulation campaigns, using MPAS or other models, particularly information such as the simulation time per 24h shown in figure 12.

The comparison between MPAS, with and without grid refinement, the WRF WASCAL simulation and observations over the West African monsoon region in section 3 has some serious limitations. The authors put a significant amount of effort into discussing differences between the three models, but this is all based upon a single year of data for each configuration. To demonstrate the ability of MPAS to simulate this region with a reasonable amount of confidence I would want to consider either (a) the climatology over several decades, or (b) the collective performance of a perturbed initial condition/parameter ensemble (or both if computing resources were not a factor). In my experience lowest order bit level perturbations to potential temperature fields are sufficient to cause simulations to take very different trajectories leading to quite a wide range in behaviour after a few days let alone months.

I appreciate that the authors may not be in a position to extend their simulations due to the availability of computational resource, but they must clearly indicate the limitations of the approach taken and data used in section 3 in order not to give the reader more confidence in the model configuration than is really warranted based on the results presented. I have mixed feelings about whether section 3 is appropriate for publication; variable resolution models such as presented here have an important advantage over regional models in that local processes can feed back into large scale behaviour, but the value of a single 11 month simulation as a tool for model validation is extremely limited.

Section 4, on the extreme scaling tests, describes a courageous and successful attempt to run a global 3 km model using a large fraction of a supercomputer and some of the issues found when using hundreds of thousands of MPI tasks. The technical feasibility of these tests will give other groups considering running at similar spacial scales confidence that the underlying computing infrastructure can work, even if it is not currently practical to work on these scales.

Overall I would recommend that this paper is published, but with substantial alterations to section 3 to clearly state the limitations of the analysis performed and to clarify the key results.

Specific comments

Section 2.1: It is very difficult to compare the different systems used here without jumping between sections. I recommend replacing sections 2.1.1 to 2.1.4 with a table comparing the key important parameters of each system with a minimal discussion of the important differences. Additional information that the authors feel necessary to include could be moved to an appendix.

Section 2.2: There are two important details of the model configuration used that I would like to see clarified within the text. First, a clarification of the vertical levels used; from searching through the supplementary information it appears that the top level is at 30 km, this is significantly below most model configurations in preparation for CMIP6 where 80-90 levels covering up to 80-90 km is more more common. Second, could the authors clarify the amount of data being written; it is instructive to consider this in terms of the number of 2D and/or 3D fields rather than the number of grid cells as shown in table G1. This would allow comparisons to other models on other grids more clearly.

Section 2.3: The choice of time step for the regular 120 km grid seems excessively conservative. The authors state that this is to increase the time spent performing the time integration compared to initialisation and I/O, but this makes it more difficult to interpret how such a model would perform if run for multiple decades for a scientific purpose where a 150s time step is unaffordable/unnecessary. One way around this

would be to analyse the performance data you have in terms of a simple model such as Amdahl's law (perhaps with an additional term to allow for initialisation time which may anti-scale with increasing numbers of processes), something which would allow benchmarks for MPAS on other systems to be extrapolated. I strongly recommend that the authors consider their scaling data in the framework of such a model.

Could the authors also clarify whether each performance data point used in the analysis arises from a single test run or whether this is an average over multiple tests? Variability in performance on HPC systems, arising from changing communication/IO loads, can be of the order of a few percent particularly for short runs.

On page 7000 line 27 the authors state that "the communication volume scales linearly with the number of tasks" the figures they refer to indicate a power law relationship, and my back of the envelope calculations for the regular grid suggest an exponent of around 0.8. Please correct this and extend to the other model cases where by eye I can see that this exponent will decrease.

The use of figures here could also be rationalised; figures 1-3 all show simple properties of the grid and its relationship to MPI tasks, and could easily be merged into a single figure (removing one of the globes). Figure 4 is the first of three figures showing a normalised parallel efficiency (6 and 11 follow) that would also be useful to compare and contrast so placing these together would be helpful.

Could I also implore the authors not to use the rainbow colour map, used in many figures in this manuscript, in any publication graphics anywhere, ever. Please read http://www.nature.com/nature/journal/v519/n7543/full/519291d.html , http://www.climate-lab-book.ac.uk/2014/end-of-the-rainbow/ and Light and Bartlein, EOS 85 (40), p385, 2004 for a discussion on how this colour map misleads and prevents a substantial fraction of the readers from interpreting the results. In addition please ensure that line plots are distinguishable by both colour and line style.

Section 2.4: There is a fair amount of discussion of the impact of non-contiguous partitioning, which concludes with the statement on p7004 line 12; "We conclude therefore that the impact of non-contiguous partitions on the run time is negligible for any reasonable number of tasks for a given mesh". To me this indicates that the page of discussion leading up to this statement could be cut along with figures 7 and, 9 which I do not feel adds to the manuscript.

Figure 5 could be merged with figure 10 which would allow comparison of the properties of the two variable grids.

Section 2.5: Figure 12 is perhaps the most useful in this paper, clearly showing the similarity in performance between several systems, where the breakdown in performance happens and the performance issues faced by climate modellers in using Bluegene systems. I find this means of presentation much clearer than the parallel efficiency plots and suggest that the authors include equivalent ones for the other meshes.

Section 2.6: It is pleasing to see the breakdown of scaling data into the different components in figure 13, particularly the integration vs IO balance, but I do not find the inclusion of the parallel efficiency plots here particularly useful. The extension to detailed components, including the pie charts in figure 14, does not add much to the manuscript and could be significantly reduced. In particular I take issue with the use of multi-coloured pie charts in figure 14. This figure is very difficult to read/interpret along side the text and I recommend it is replaced with a table including absolute timings along with percentages should the authors still feel that this information is important to the manuscript.

Section 3: I have stated above that I have concerns over the limitations of the analysis in this section due to the limited amount of data considered.

The allowance of a few months for the spin up of the model should be adequate for the atmosphere, but I would expect the time scales for spin up of soil moisture to be in the 2-5 years range rather than the months suggested. If the soil moisture is out of balance here this should be made more clear.

The orography and land-sea distribution shown in figure 15 shows an odd feature in the 120 km configuration; an apparent retreat of the African coast by 200-300 km. Could the authors confirm whether this is a true property of MPAS or whether this is a graphical artefact that could be addressed in the figure.

The comparison of absolute values in the figures referred to in this section is difficult. The authors should choose an appropriate reference point, be that an observational climatology or the WRF run, and plot differences to make it clearer what and where the differences are rather than relying on the readers colour perception to inform them. In addition please label axes and colour bars with units.

Section 4.3: On page 7019 the issues around model instabilities leading to NaNs is raised. I would expect the occurrence of NaNs in an model run to lead to an immediate failure of the code rather than just a performance degradation. Could the authors note what action is being taken by MPAS when such numerical failures occur.

Could the authors also settle on a single unit for describing the computational size of each model in the scaling tests; MPI tasks, nodes and racks are all used in different places.

Finally, could the authors modify figure 21 such that colour is not the only indicator of the breakdown of run time into different components.

Other technical corrections

Page 7011 line 21: change ‘Golf of Guinea’ to ‘Gulf of Guinea’

Page 7013, line 9: 3000 mm per day seems a very large volume of rainfall, please check the units.

Interactive comment on Geosci. Model Dev. Discuss., 8, 6987, 2015.