Geosci. Model Dev. Discuss., 7, C2628–C2630, 2014
www.geosci-model-dev-discuss.net/7/C2628/2014/

Geoscientific
Model Development
Discussions
Open Access

# *Interactive comment on* "Efficient performance of the Met Office Unified Model v8.2 on Intel Xeon partially used nodes" *by* I. Bermous

**Anonymous Referee #2**

Received and published: 8 December 2014

Paper Summary =============

Year on year, the number of cores offered by modern microchips is steadily increasing. However, the infrastructure delivering data from memory to these processors is improving at a slower rate. This trend is illustrated by the byte/flop metric shown in table 1.

If the bandwidth to memory becomes saturated, the performance of the running program is impeded, as the processors must periodically halt and wait for data to arrive before further computation can take place.

A method to mitigate this which is well established in some High Performance Computing (HPC) communities is to leave some of the cores attached to a socket idle. This

has the effect of improving the memory bandwidth for the remaining cores attached to that socket, which are tasked with the computation.

This paper has two key topics. The first is very useful a record of the practice of leaving some cores idle. The second is a report of some very interesting empirical findings for variants of a particular model run under an idle core regime.

The model in question is Unified Model (UM) developed by the UK Met. Office. The two variants are N512L70, a global model, and UKV, a higher resolution city-scale model.

The performance of these two variants was compared when run on three different clusters. All the cluster were equipped with Intel Xeon processors of various vintages. The Solar cluster (oldest) has 8 cores per node. Ngaimai has 12 and the Raijin cluster (newest) has 16. The memory bandwidth steadily decreases as the core-count increases.

For me the key finding is that–on Raijin, but not on Ngaimai and Solar– both the regional and global models run faster when only partially committed nodes are used. Another finding is that using partially committed nodes can aid scaling for models with a constrained domain decomposition.

General comments ================

I believe that the paper would be clearer and easier to read is the record of the method and the empirical findings were more clearly separated. For example, the nature of the domain decomposition constraint for the UKV model presented in section 4.1 wasn't immediately clear to me.

I believe that the graphs using 'Number of used cores' on the x-axis are not helpful and detract from the core message. A stated aim for the work was to find the most efficient model configuration with regard to computational resources. If cores are left idle, they should still be accounted for in a measure of efficiency and so the graphs using 'Number of reserved cores' on the x-axis are, for me, the right ones to use.

If it were possible, some empirical measures of memory bandwidth (perhaps offered by PAPI calls?) would be very interesting and would bolster the key findings.

I believe that the comments regarding 4D-VAR and other N96 resolution model should be removed from the conclusion. The reason for this is that the conclusion summaries points previously examined in the paper, and these codes were not discussed anywhere else.

Specific comments ================

p7397 l21: Would 'resource contention' be better than 'memory contention'? Since it is the bandwidth to memory rather than the use of particular memory addresses that is in competition.

p7402 l4: Please explain why removing -xHOST ensured reproducibility of results across clusters.

p7402 l24: Please explain why the given environment setting improved the stability of the measured run times.

p7403 l14: Please explain why the given Lustre configuration optimised the I/O performance.

p7407 l13: Please explain why the given Lustre configuration optimised the I/O performance.

Recommendation ==============

I would recommend that the paper not be published in its current form. However, I would strongly recommend the author to resubmit a revised version of the paper.

---

Interactive comment on Geosci. Model Dev. Discuss., 7, 7395, 2014.