Geoscientific
Model Development
Discussions

Interactive comment on "Influence of the compiler on multi-CPU performance of WRFv3" *by* T. Langkamp

D. Morton (Referee)

don.morton@alaska.edu

Received and published: 17 May 2011

Disclaimers -

1) I am cited in this paper 2) To help maintain objectivity and independence, I have not read other reviewers' comments before offering my own.

I think this is a well written paper that offers information of value to the numerous WRF users around the world. The WRF modeling community is often asking questions about what types of hardware and software infrastructures would be optimal for a given WRF deployment. The answer to these questions is typically very complex with many, many degrees of freedom. This paper, in my opinion, offers a valuable contribution to the community in providing an account of experiences gained while examining the

contributions that various compiler and MPI (albeit only OpenMPI) implementations might make to WRF performance. The work is by no means comprehensive, nor should it be. It simply represents a contribution of one research group in its exploration of WRF performance issues. Additionally, given a common code and/or numerical basis for many modeling codes, lessons learned and disseminated in this work apply outside the realm of WRF and may offer valuable ideas for future researchers in yet-to-be-determined application areas.

Although I believe the paper is generally well-written, and would not object to its acceptance as-is, I offer a number of suggestions that I ask the author to at least consider.

Page 548, Line 5: Though "This is currently the most used hardware..." is probably correct, a citation would be nice

Page 549, starting at Line 13: The author should be commended for attempting to reduce WRF performance evaluation to a manageable number of degrees of freedom, and offering compelling justification for doing so.

Section 1.2: I believe this section clearly outlines the goals and justifications of the work.

Page 551, Line 17: Though the claim that ARW is recommended for research and NMM for operational, I think this needs at least a citation and maybe a little qualification.

Page 552, Line 18: The claim that storage and login nodes are "non-crucial components in respect of performance" (should be "...with respect to performance...") might be valid for this particular study, is definitely not true in more general cases. Maybe this should be reworded for clarity.

Page 552, Line 23: The statement is made that "WRF is definitely influenced by the old kernel version." I suggest that this statement "teases" the reader, who may want to understand why (and how) this is the case. I recommend it either be deleted or explained more fully.

Interactive
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Page 553, Line 18: The author makes interesting - apparently verified - statements about the the effects of -O0 and -O3 optimizations on C++ and Fortran code. In our past experiences, -O3 optimizations have sometimes introduced instabilities in the code that seem to get resolved when going back to -O0 optimizations. This has been encountered by us on a number of occasions, and I believe it might warrant mentioning in here that these optimizations just might end up being a source of hard-to-detect problems.

Page 554, Line 20: Somewhere in this area, it might be worth suggesting that large memory support be a "default" setting for netCDF and WRF when compiling. More than once, we have watched people struggle with seemingly mysterious I/O issues that were finally resolved when it was deduced that the output files were larger than 2 GBytes. Although not a factor in this work, it's one of those "hidden" pieces of info that might be prudent to include in searchable articles.

Page 556, Line 13: The author discusses a a mysterious problem in slowdown when using Tornado on 64 cores. Our group has come across two different Sun clusters (and, Tornado appears to be a Sun cluster) in the past year in which this kind of behaviour was tracked down to bad nodes. Isolating the bad node(s) resolved the problem (though users need to be wary that any of the nodes can go bad at any time).

Page 556, Line 25: The author presents an interesting piece of information concerning the use of the numtiles namelist parameter to optimize for cache performance.

Page 557, Line 6: I'm not sure I agree with the author's method/suggestion of basing the timings on a small number of timesteps (45), but the author clearly states methods and reasons, and so the readers can judge for themselves.

Continuing from this point into the discussion of variability, I think this discussion is important, and again, I'm not real happy with this use of "minimal time" performance metric, but the author, as before, gives substantial justification for this approach, and future work by this author or others can easily change this if desired.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Interactive comment on Geosci. Model Dev. Discuss., 4, 547, 2011.

Interactive
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper