Geoscientific

Model Development

Discussions

[GMDD]

Interactive

comment

Interactive comment on "Development and performance of a new version of the OASIS coupler, OASIS3-MCT_3.0" *by* Anthony Craig et al.

Anonymous Referee #2

Received and published: 27 April 2017

General Comments:

This paper provides a good overview of the latest version of OASIS, OASIS3-MCT_3.0, which has a series of new capabilities, such as the usage of offline mapping weights, the ability of coupling intra one executable, and parallel data coupling without separate processes. OASIS3-MCT has already been tested with different configurations on up to 32,000 processes, and has already been used in the community. I recommend to publish this paper after addressing the following specific comments.

Specific comments:

1. The abstract (as well as some other parts in the context, such as P3 L14∼L15, P6 L14∼L15, and P11L31) mentions that "OASIS3-MCT_3.0 is the latest release and includes the ability to couple between components running sequentially on the same

set of tasks". It seems contradictory to P6 L24~L25 that "Each task will be associated with only one executable and one component in any application", which indicates that components cannot share any task. According to the API of "oasis_init_comp", I think the statement in P6 L24~L25 is true.

2. P1 L15~L18, P6 L18~L19, P6 L25~L27 and P12 L1~L2 may indicate that that there can be two different decompositions of the same grid within the same component and these two decompositions can have different subsets of the tasks (processes). To achieve this capability, the API "oasis_def_partition" has been extended with an additional parameter "name". When I read the user manual at the first time, I guested that "name" means the name of the grid. After a careful consideration, I think that "name" should be the keyword of a decomposition but not the name of the corresponding grid, which means that the "name" corresponding to two different decompositions of the same grid within the same component should be different. If that point is true, please clarify it.

3. The ability to define grids has been mentioned several times in the paper. What does it mean when only the API for writing grid data into files are introduced in the user manual. According to Figure 2, is the grid defined implicitly in the definition of decomposition?

4. Compared to OASIS3, OASIS3-MCT_3.0 have a new capability of pre-defined mapping files. After reading the paper as well as the user manual, it is still unclear for me that how to make OASIS3-MCT_3.0 know which mapping file should be used for a specific set of coupling fields (for example, users may want to use bilinear algorithm for state fields and use conservative algorithm for flux fields when coupling fields from an atmosphere model to an ocean model). Is there any restriction when users using the pre-defined mapping file. Concrete examples are welcome for this new capability.

5. P7 L28~L29. It is interesting to know how to make the puts non-blocking. In MCT, the data sending is blocking for example with the MPI_wait, which indicates that such

Interactive comment

MPI_wait should be disabled for the non-blocking puts. It seems that OASIS3-MCT_3.0 does not use another MPI_wait out of MCT. So, one interesting question here is that how OASIS3-MCT_3.0 guarantees the puts constantly non-blocking (for example, we encountered the case that MPI_Isend was blocked when we sent a large message or many small massages) and how OASIS3-MCT_3.0 achieves safe non-blocking puts (for example, how to guarantee that next puts do not flush the data of previous puts in memory buffer).

6. P6 L10~L11 states that "The opt option will however be bit-for-bit reproducible if the same number of processes is used between different runs". Given the same number of processes, bit-for-bit results may fail to be reproduced if the decomposition changed.

7. One suggestion regarding Section 2.4 is that the opt option can use higher-precision of floating-point calculation to achieve faster bit-for-bit identical reduction. For example, using REAL8 when the coupling fields are REAL4 and using REAL16 when coupling fields are REAL8.

8. Some results in Table 4 seem strange to me. Why the time for <10 fields, 10 couplings> is obviously smaller than 10 times of the time of <1 field, 1 coupling>? Why <10 fields, 1 coupling> is not much faster than <10 fields, 10 couplings>? The most significant reason may be the MPI message size of <1 field, 1 coupling> is big because the two components have similar decompositions and the core number is small relative to the big grid size. Given the same core number, more test cases with smaller grid size and different decompositions between the two components are welcome.

9. The year of the first reference should be 2008.

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2017-64, 2017.