

Interactive comment on “A fast input/output library for high resolution climate models” by X. Huang et al.

X. Huang et al.

hxm@tsinghua.edu.cn

Received and published: 9 December 2013

Dear David,

Thanks a lot for your comments. We list our replies to your questions as follows:

(1) “My understanding (correct me if I am wrong) is the model is essentially run with $N+M$ processors where N processors will perform computation only and M processors will perform I/O (I assume it is O , output only since each of those N processor able to read input without any issues).”

Your understanding about N processes for computation and M processes for I/O is correct.

C1550

(2) “Have you compared the timing to run the entire model in this new way versus running the model in the old way with $N+M$ processors?”

In our experiments, we are generally comparing the time for running the model in the new way (N for computation and M for I/O) versus running the model in the old way (N for computation). However, we think that our experimental results already demonstrate that the new way can achieve significantly higher performance than the old way with $N+M$ processes.

As you can see from Fig. 7, the running time for the original CICE with $N=320$ processes is 1233 seconds. Using CFIO ($N=320$, $M=64$), the running time is reduced to 557 seconds, which is very close to the NO-I/O version with $N=320$. Fig. 7 also shows that the running time for the original CICE with $N=640$ and 1280 processes are 966 and 916 seconds respectively. Running the original CICE with $320+64=384$ computation processes shall provide a time between 966 and 1233 seconds, which is much greater than the 557 seconds in the CFIO case. Similar results can also be observed in the POP and the LICOM experiments.

(3) “We have considered this approach many years ago with $M = 1$ and we did not see any benefit with the new approach.”

We agree that $M=1$ would probably not bring significant performance benefits. The only one I/O process in the system will gather the data from all the other N computation processes, and outputs the data. In the scenarios with a large N , the single I/O process will become the bottleneck of both communication and I/O. We believe that a larger M will lead to better load balance and higher performance.

Best, Xiaomeng

Interactive comment on Geosci. Model Dev. Discuss., 6, 4775, 2013.

C1551