Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., referee comment RC1
https://doi.org/10.5194/gmd-2022-33-RC1, 2022

Comment on gmd-2022-33

Anonymous Referee #1

Referee comment on "swNEMO_v4.0: an ocean model based on NEMO4 for the new-generation Sunway supercomputer" by Yuejin Ye et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2022-33-RC1, 2022

This paper develops a unique ultrahigh scalability parallelized NEMO ocean model on the Sunway supercomputer architecture. A new many-core optimization using remote memory access (RMA) blocking and dynamic cache scheduling can effectively accelerate the performance more 90% of ideal bandwidth. The strategical optimization based on mixed precision improves the parallel performance to achieve more than 99% with appropriate 28 million cores. This represents significant progress in the ocean modelling parallelization. The impact will be tremendous. However, there are two major issues to be addressed.

- A very important aspect of improving the parallel performance is to ensure the reproducibility. This study provide a significant speed up combining hardware and software optimization. However, using the mixed precision can change the solution if different cores are used? Can the mixed precision affect the reproducibility and consistency? The authors should address and discuss this issue to ensure the robustness of the proposed model.
- The other issue is related to the commonly used ensemble simulation while different precision is used. Baker et al. (2016) evaluated the consistency and proposed the perturbation requires half precision (see the large variation of SST simulation in Fig. 3 of Baker et al., 2016). The mixed-precision OGCM can causes the Bit-to-bit inconsistency within ocean model. Is that correct? How can this MP approach compare with the reduced convergence accuracy in the solver, which can also speed up the simulation?

Baker, A. H., Hu, Y., Hammerling, D.M., Tseng, Y.H., Xu, H., Huang, X and Bryan, F.O. (2016), “Evaluating statistical consistency in the ocean model component of the Community Earth System Model (pyCECT v2.0),” Geosci. Model Dev., 9, 2391-2406.

Finally, there are some grammatical errors within the text. The text and discussion also require some reorganization for a better presentation. Further improvement in the English and careful proofread by a native speaker are required. This paper is appropriate to be published in GMD after considering the above major issues and the following comments.

- Line 7, Abstract: DMA is not defined. what do you mean by DMA? Do you refer to remote memory access (RMA) or something else (Direct memory access)?
- Line 21, change "the one of most important directions of OGM development" to " one of the most important directions for the OGCM development".
- Line 23, change "horizontal resolution doubled" to "doubled horizontal resolution".
- Line 31, what do you mean by 6.8x? Do you mean by a factor of 6.8? If so, I suggest to change this rather than symbol x. This can be seen elsewhere.
- Line 31, "achieved the performance of 408 Intel Westmere cores on four K20 GPUs". What do you mean by this? What performance is achieved? Equivalent performance of 408 Intel Westmere cores using 4 K20 GPUs? However, how many gpu cores for the K20 GPUs? The cores of Intel processors are not equivalent to the cores of GPU processors, right?
- Line 27-43, Table 1 and the review of performance improvement are impressive. However, are they all for the improvement of ocean models? FUNWAVE seems to be a wave model? What about MUSNUM? I suggest to separate wave model to a different category since the architecture of a wave model is totally different from the ocean dynamical model. Also, what's the difference between POP2 and CESM-HR? It seems they are both 3600x2400 resolution, right? While the performances are similar but the maximum scales quite different (~4 times). I suggest to tabulate the representative ocean model performance development here (exclude other types of models) and discuss the most significant development.
- Line 44-54, the discussion here also mixes the parallelization of atmosphere model, ocean hydrodynamic model and ocean wave models. Particularly, the required global barriers are also different. This can significantly impact the model overall performance. Don't mix the ocean hydrodynamical model with other types of model in the comparison because the solvers are totally different. Also, this paragraph mixes the different limitations of different models to improve their performance without specific focus. I suggest to reorganize this discussion to be more focused and related to the improvement relevant to this study.
- Line 46, change "only improved" to "is only improved".
- Line 56, change "Exa-scale to "Exascale".
- Line 78, is "GYRE-PISCES" abbreviation? If it is not a well-known typical benchmark test name, I suggest to described this briefly here or used a whole name.
- Section 2 describes the architecture of Sunway TaihuLight. The detailed information has been provided extensively. I suggested to remove the technical details but comment and address on the specific features facilitating the performance enhancement used in this paper here.
- Section 2 also describes NEMO model. What's the difference between NEMO and NEMO4

you raised at line 81? I suggest to move NEMO description into section 3 in associated the porting of NEMO.

- Line 120, how "adaptive" works in this four level parallelization? Two levels are using domain decomposition. One level is MPE-CPE asynchronous parallel. Is this performed at compiler level (processor specific) or user specific level? One level is the vector reconstruction. This should be done within the compiler level. Can the author comment which level contributes mostly to help the performance in the current implementation?
- Line 130, a reference is helpful for this MPE-CPE asynchronous parallelization. As described in line 131, IO can be independently separated for sure. However, how boundary data exchange can be parallelized aside from the computation? Normally, the ocean model kernel requires some global communication to solve the pressure equation (normally at least 3, can reduced to 1 in some parallelization). How can the data exchange be performed using MPE-CPE asynchronous parallelization. Some information will be helpful for the readers.
- 1.3, for latitude-depth decomposition, since this depth is not parallel friendly dimension. The parallelization requires level dependence. That means if the depth dimension is changed, the user needs to adjust something for LDA. Is that correct?
- Line 171-line 174, What is alpha_1, beta_2 and beta_1, beta_2 within the equations. The notations are not standard mathematically. Is f a function? or a value represented by the 2nd line? These equations should be labeled numbers. What is x? is x an array? Please rewrite the formula in a more mathematical way?
- 2 discusses the optimization used here. It seems 3.2.1 is used as level 3 described in 3.1. Is that correct? Or combining the 4 level parallelization? Is 3.2.2 used in the MPE-CPE asynchronous parallel or something different? If so, I suggest to reorganize this discussion and make this clear. Section 3.3 discusses the mixed precision optimization, which I believe is different from the four level parallelization. Also, line 108-113, describes three major contributions while the 2nd one is used within the 1st four-level parallelization, right?
- Line 216, the maximum biases reach 0.05%. Are these biases the deviation between DP and HP? However, considering the chaotic behavior with time, can this bias propagate? Can the biases become larger with time? If this is the case, can the model result get bit-to-bit consistency which is a very important feature for ocean model within an earth system model? For the pressure solver within the ocean dynamical kernel, do you still use DP? If you still use DP, the convergence will still take time. Can you compare this optimization with another easier way by reducing the pressure solver criteria to a lower level (change from 10^-13 to 10^-7)? Changing the pressure solver criteria to a lower level can significantly reduce the computational time. Why not just use this simple approach since you already reduce the precision? Do I miss something? Normally for the ocean model, the most intensive computational cost is the pressure solver rather than the tracer equation, right? Why not use this approach while still preserving the precision?
- Line 223, change "periodical" to "periodic". What do you mean by "North Pole folding"? Do you mean "Displaced North Pole"?
- Line 230, change "is equal to" to "equals to".
- Line 229-234, this paragraph is confusing. It describes "three experiments with 2 km, 1 km, and 500m". However, each experiment uses 8 different parallel scales (Table 3), resolution ranging from 9km to 1km. Do you use 2km, 1km and 500m or 9km to 1km? I suggest to clarify these numerical experiments. What's your definition of weak scaling and strong scaling.
- Line 242, what is "CPEs parallel method"? Is this your control experiment? This has nothing to do with the MPE-CPEs parallelization, right? However, does CPEs parallelization still use four-level parallelization? Can you isolate the individual performance enhance resulting from the approaches discussed in section 3?
- Line 248-253, do you include the performance increase due to the mixed-precision approach here or just the DMA and FLOPS for the DP? The timing may be different.
- Line 256, can you describe these five kernels briefly? What's the major differences?

- 8, do you use the real time? Or measure the clock? These are built-in hardware, is it right? Therefore, these values only refer to the access time, right?
- Section 4.2, is the implementation only performed for the tracer equations? Fig. 9 shows only the tracer integration which is only a very small portion of the overall run time. Can the author show the dynamical solver part which requires the most intensive computation (particularly the barotropic solver) instead of this tracer solver?
- Since this is GMD rather than computational journal, can the authors show the final results? It will be useful to examine if the GYRE-PISCES configuration reaches the expected solution as others. A figure with velocity and temperature fields will be enough, particularly what specific features can be found at 500m resolution. The potential impact of mixed precision optimization can also be discussed.
- Line 286, the description is very superficial, any supporting evidence?