Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., referee comment RC1
https://doi.org/10.5194/gmd-2021-319-RC1, 2021

Comment on gmd-2021-319

Anonymous Referee #1

Referee comment on "Parallel implementation of the SHYFEM (System of HydrodYnamic Finite Element Modules) model" by Giorgio Micaletto et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-319-RC1, 2021

The paper presents the parallelization of the software package SHYFEM using distributed memory approach. In particular, optimized libraries were used to solve the free surface equation and to partitioning the computational grid. Parallel approach is described in details and performance analysis is carried out on results.

As claimed by the authors the limiting factors to the overall scalability lie in the sea level computation which involves the PETSc – KSP solver and data decomposition. The efficiency drops below 55% with 144 cores. In particular,

- Matrix decomposition used in PETSc, namely the block row partition, is not the same of SHYFEM. Then, a global communication is required and a loss of efficiency results. My questions: are
  - since the parallelism is being introduced in the original sequential version of SHYFEM, why the data partitioning chosen for SHYFEM is not the same of PETSC?
  - PETSc offers the Distributed Arrays (DMDA) objects that simplify the distribution and the management of the domain data (all the physical quantities on the domain region) in a distributed memory system. Why the authors do not exploit DMDA objects?
  - Otherwise, why do not explore the use of TRILINOS, or HYPRE that are already able to interact with PETSc ?
  - Finally, why the authors do not explore matrix-free solvers?

The authors should clarify these issues.

- The KSP solver used for the free surface equation is known to have synchronization points at each iteration leading to a loss of efficiency for the parallel algorithm. My

question is:

- why do not explore communication avoiding variants of Krylov sparse solvers?

The authors should be aware of this issue

- In addition to these two factors, in my opinion there is another crucial point. Since domain decomposition involves only spacial direction and not the time direction, a global communication is required at each time step of surface equation. My question is:
  - why do not explore parallel -in -time approaches ? Parallelism should be introduced ab initio in any mathematical / numerical model, and this is especially true for time marching models. Otherwise, the efficiency will be ever poor.

The authors should discuss this issue

In conclusion, I think that while the deployment of an application software by means of the use of scientific libraries, such as PETSc, can be considered a good investment, SHYFEM needs to be deeply redesigned to meet scalability requirements.