ACCESS-OM2 scalability on Gadi

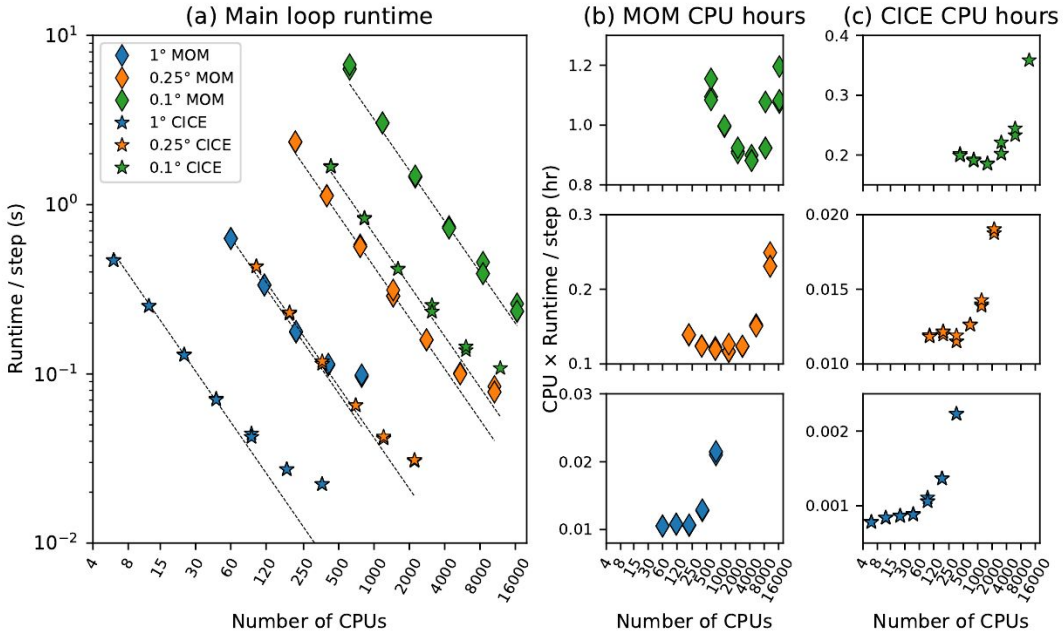
Paul Leopardi
National Computational Infrastructure
20 July 2020

Aims

The aim of the performance testing was to better understand the scalability of ACCESS-OM2 on Gadi to enable:

- Performance improvement by tuning
- More experiment configurations in terms of CPUs vs walltime

Previous report on Raijin performance



Experiments conducted

Testing included:

1. MOM-SIS at 0.25 degrees and 0.1 degrees:
    a. Testing MOM with SIS, unmasked, without writing restart files, to determine the “standalone” scalability of MOM.

2. ACCESS-OM2 at 0.1 degrees:
    a. Testing ACCESS-OM2 with land masking, scaling MOM and CICE proportionally;
    b. Investigating the effects of varying the number of blocks used by CICE;
    c. Investigating the effects of varying the proportion of CICE CPUs to MOM CPUs;
    d. Unsuccessful reproduction of Marshall Ward’s tests as documented in the ACCESS-OM2 report.

Differences to Raijin experiments

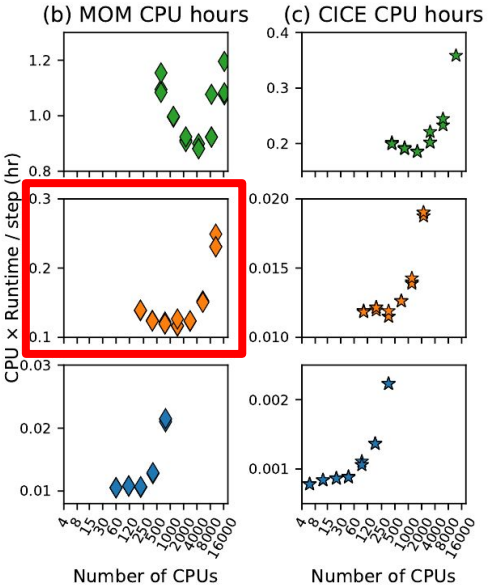
Results presented here are not directly comparable to the Raijin results:

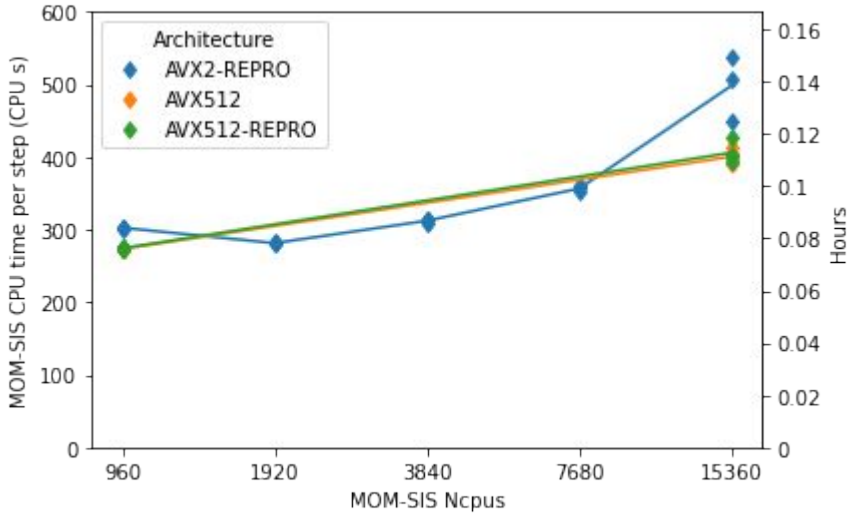
1. MOM-SIS results are based on a coupler and sea ice model different from the ACCESS-OM2 model represented by the Raijin results.
   Number of CPUs used by MOM-SIS is not split into separate MOM and SIS numbers.

2. ACCESS-OM2 results are based on different configurations from those on Raijin:
   a. A continuation run was used, rather than an initial run,
   b. The time step used was 540 seconds rather than 400 seconds,
   c. MOM and CICE were scaled in proportion rather than separately,
   d. Scaling was taken to almost 20,000 CPUs rather than 16,000.

MOM-SIS at 0.25 degrees

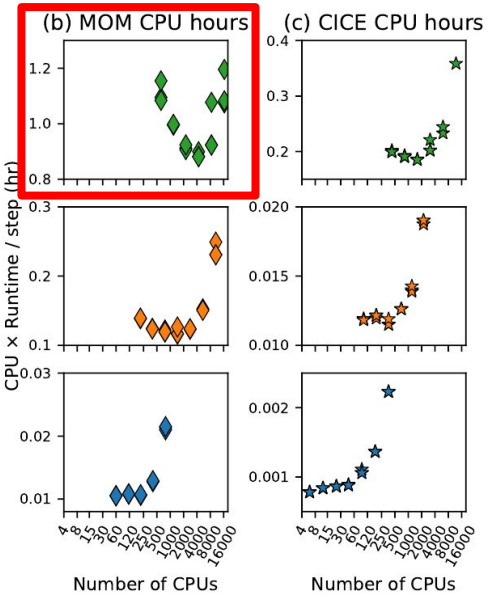
MOM-SIS on Gadi seems between 25% faster than ACCESS-OM2 MOM on Raijin at the low end of CPU scaling, to about twice as fast at the high end.

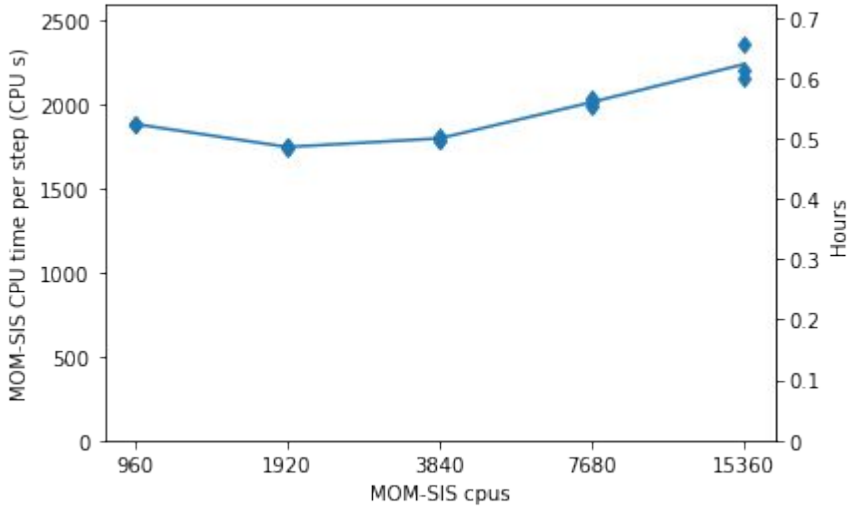




MOM-SIS at 0.1 degrees

MOM-SIS on Gadi seems between 70% faster than ACCESS-OM2 MOM on Raijin, to almost twice as fast at both the low and high ends of CPU scaling.

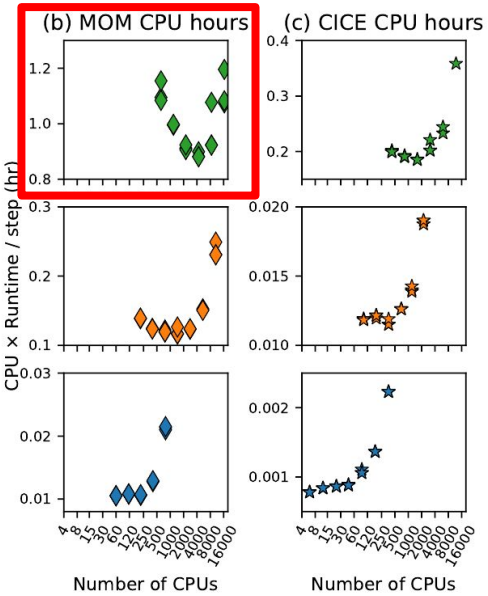


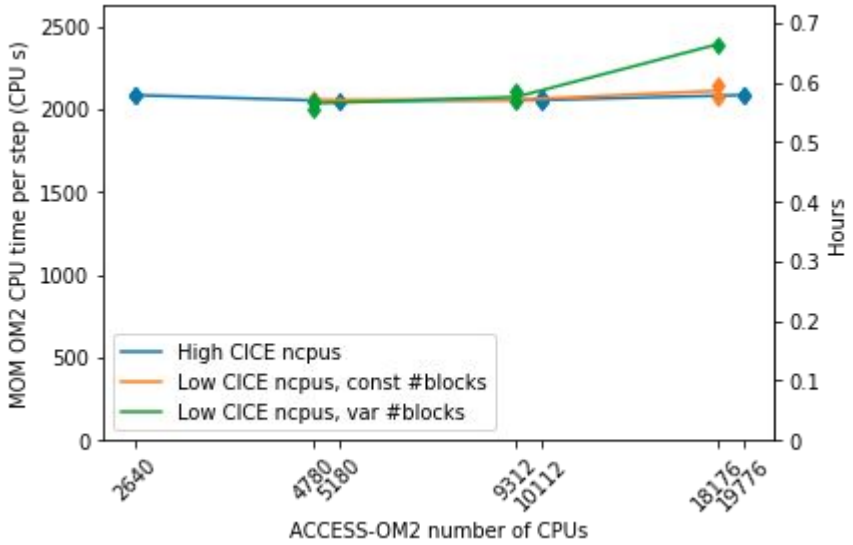


ACCESS-OM2 MOM at 0.1 degrees

At around 5000 ACCESS-OM2 CPUs, MOM on Gadi looks about 50% faster than MOM on Raijin.

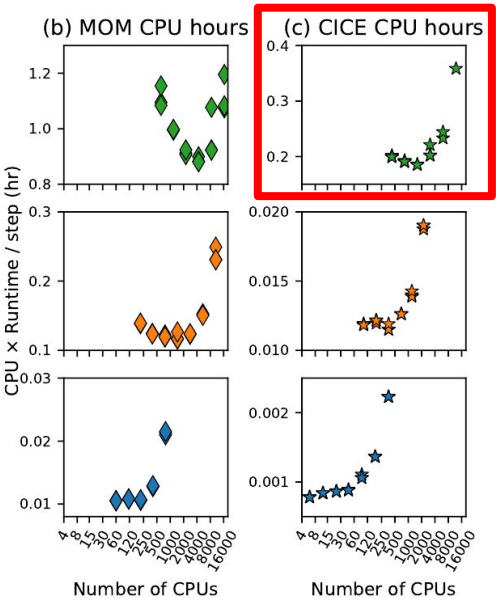
- Almost twice as fast at around 16,000 CPUs, with good scaling to almost 20,000 CPUs.

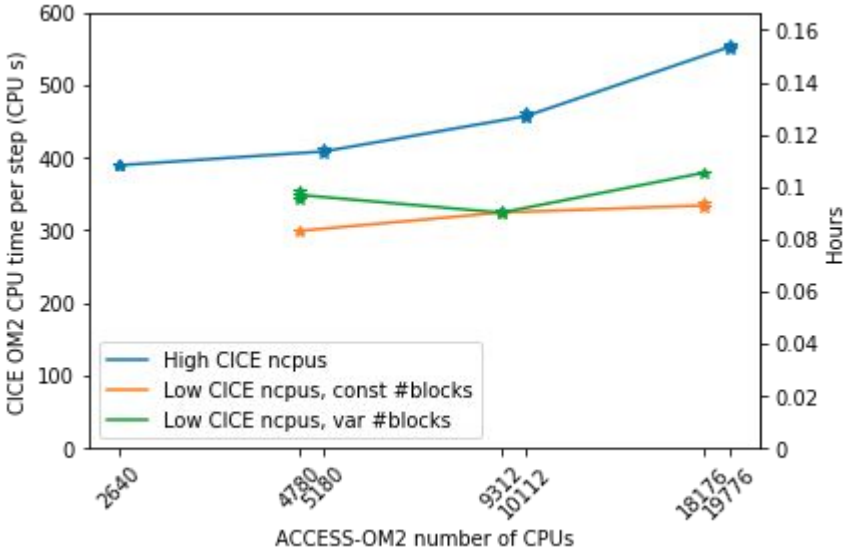




ACCESS-OM2 CICE at 0.1 degrees

At around 2500 ACCESS-OM2 CPUs, CICE on Gadi with a high proportion of CPUs seems between 50% faster than CICE on Raijin. This scales to about 2.5 times as fast at around 16,000 CPUs.



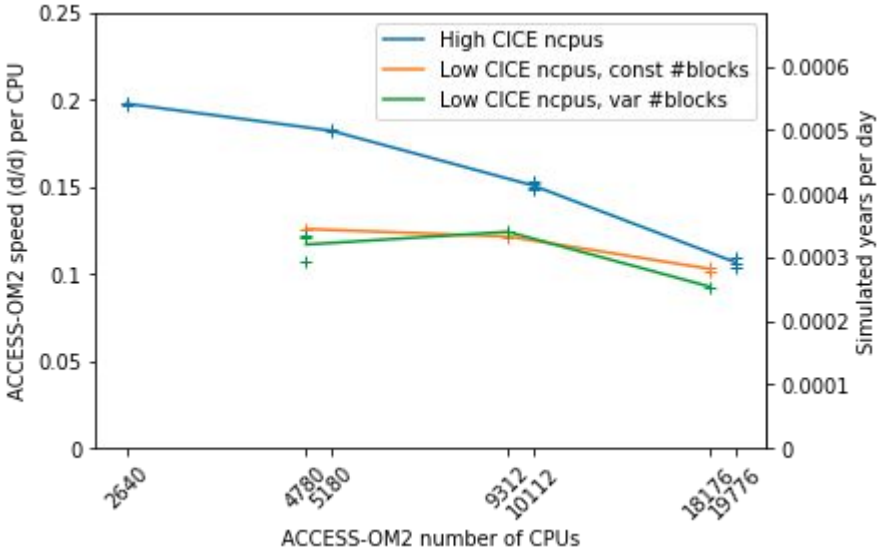


ACCESS-OM2 at 0.1 degrees

1. With the current proportion of CICE CPUs to MOM CPUs (799 to 4358), ACCESS-OM2 does not scale well in terms of walltime.

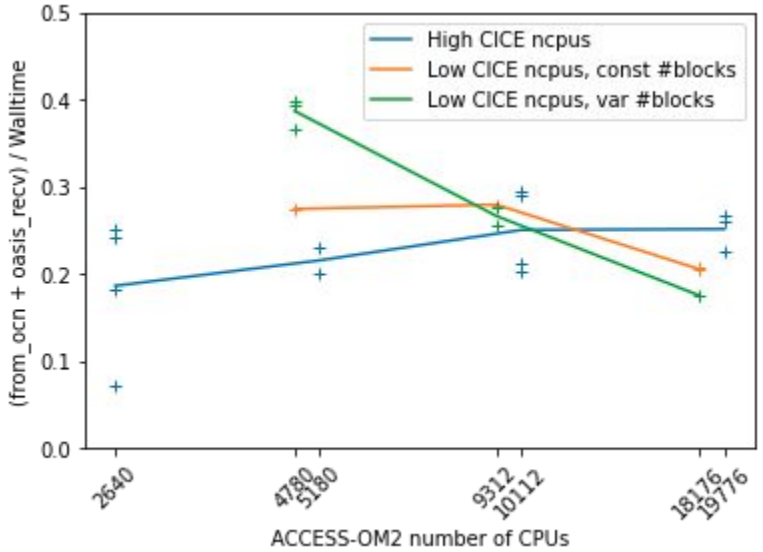
2. With half the proportion of CICE CPUs to MOM CPUs (399 to 4358) and a **constant** number of CICE blocks, ACCESS-OM2 runs more slowly than case 1. but scales better.

3. With half the proportion of CICE CPUs to MOM CPUs and **varying** the number of CICE blocks in proportion with CICE CPUs, ACCESS-OM2 generally runs slightly more slowly than case 2. and scales slightly worse.



ACCESS-OM2 wait times at 0.1 degrees

Relative wait time **`(CICE from_ocn + MOM oasis_recv)/ (OM2 Walltime)`** scales better at the high end with a lower proportion of CICE to MOM CPUs.



ACCESS-OM2 start times at 0.1 degrees

When two outlier runs are excluded, initialization time is seen to increase with increasing numbers of CPUs, with little observable dependence on the proportion of CICE CPUs.