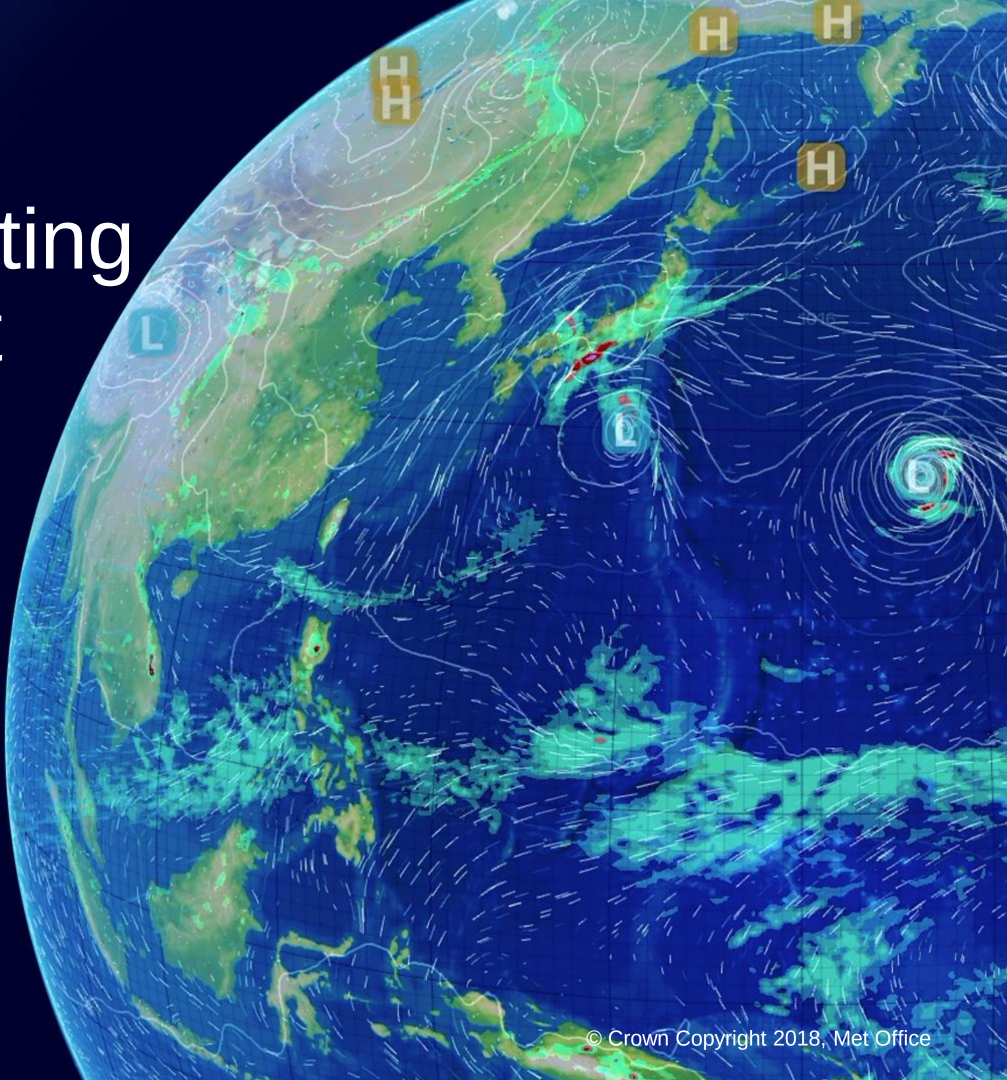
Met Office
Hadley Centre

From performance testing to model improvement

NEMO case study

Mirek Andrejczuk

Met Office



Performance testing

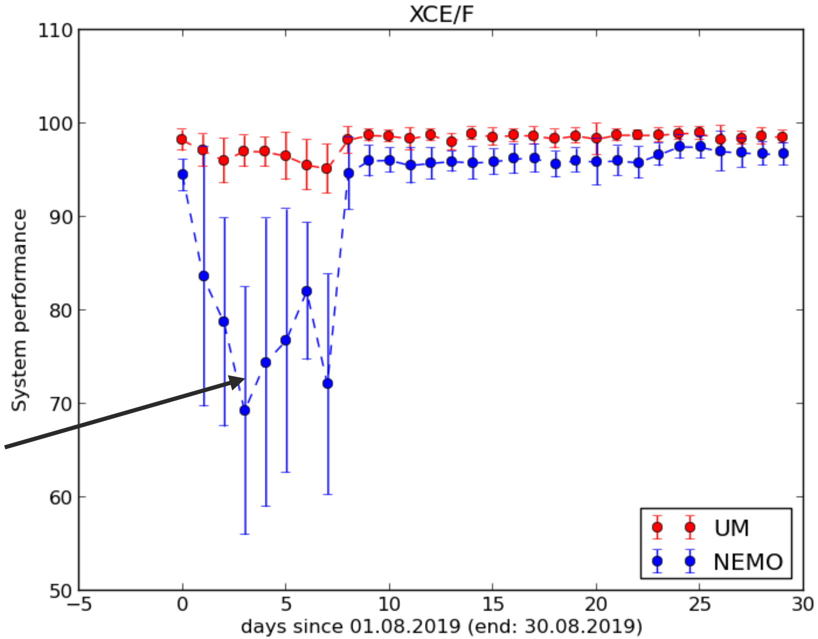
- There is a time limit on each job running on HPC
- How long it takes to run a particular job?
- When it takes longer than usually to run it is it a problem with computer or application itself?
- A way to test changes in settings/compilers/computing environment
- Runtime variability is a big problem in coupled configuration

Met Office test configuration

- Coupled GC3 climate configuration NEMO 3.6 (ORCA25) + UM (N96)
- Runs for 10 years with 10 days restart
- Timing information extracted from timing.output and UM output for each cycle
- Started in 2016 - ~3 years of data
- System performance: By how much run is slower/faster than the fastest ever recorded time for a given cycle.
  - 100% - execution time the same as lowest ever recorded
  - 85% - execution time 15% larger than the lowest ever recorded
  - 110% - execution time 10% lower than the lowest ever recorded

Daily averaged system performance

Representative Results



RAID (disks integrity) check
(1 week/month)

Why there is performance degradation?

- Computing (can’t do much about it)
- I/O (can investigate)
  - Decrease amount of data written (by XIOS, can’t disable restart) - No
  - Improve I/O reading (read namelists/netcdf attributes using single processor) – No
  - Open namelists read only - No
  - Disabling Iceberg output - improvement by ~3%
  - Use Darshan to analyse I/O - Nothing out of ordinary

What was the problem?

Too much text data written by NEMO. Disabling all text output solved this problem!

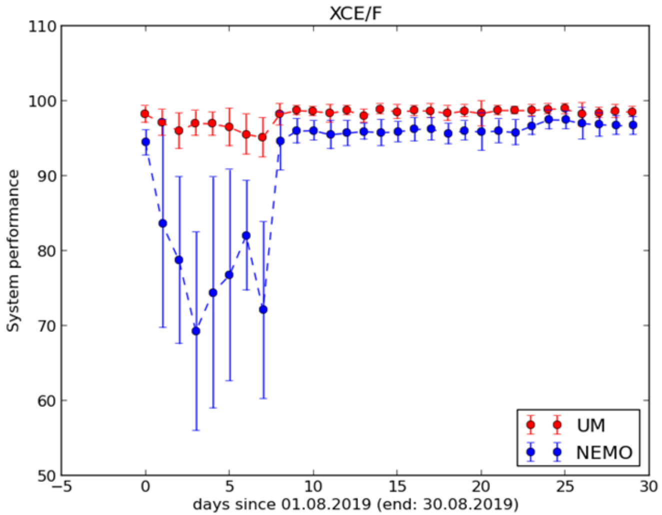
But it’s too radical. We need output produced by NEMO but probably much less than now.

NEMO 4

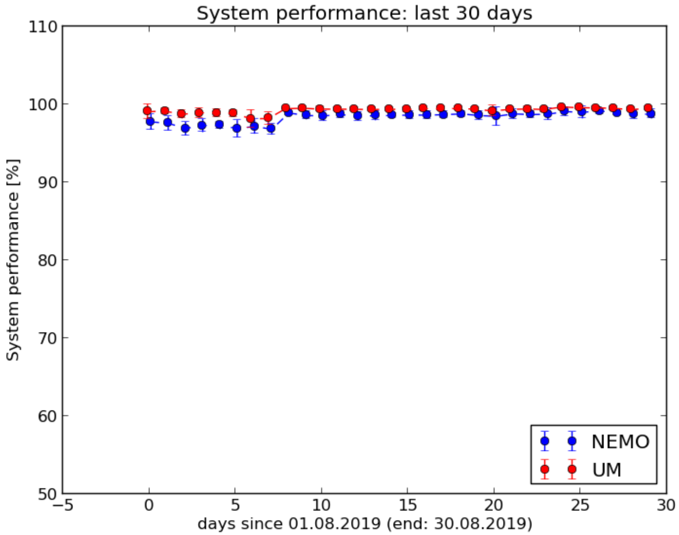
- Some work has been done on text output in NEMO 4:
  - Selective output
  - Iceberg output files
- Our solution is to introduce different levels of output (UKMO/NEMO_4.0_mirror_text_diagnostics)
  - 0 – minimum (ocean.output mostly what settings are used for run, only processor 0 writes iceberg output, run.stat: 384KB)
  - 1 – 0 + basic information related to restart, processors layout, iceberg output as in namelist (65 MB)
  - 2 – 1 + information about communication, forcing read, more restart output (66 MB)
  - 3 – output.namelists, information about text files open, (66 MB)
  - 4 – old nn_print=1
  - An option to flush() after each print or group of prints
  - In addition we set print level for XIOS to -1 for production runs

Does it work (NEMO 3.6)?

Reference run

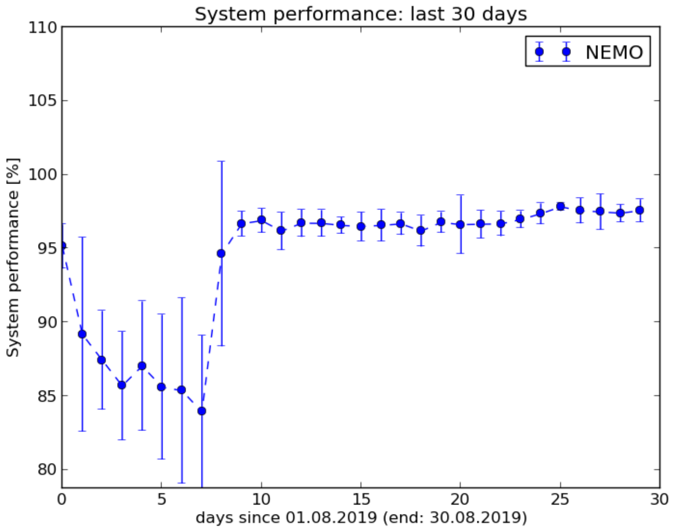


Print level 0 (not the same configuration as on the left!)



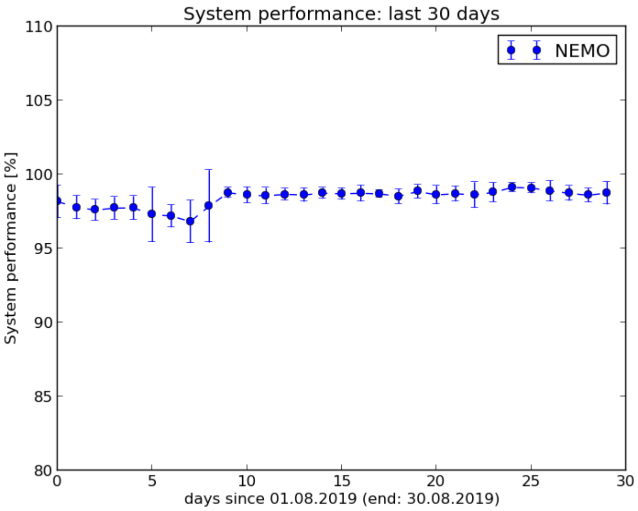
Does it work GO8? (NEMO4@r10881)

Print level 3



Print level 0

Faster by ~5%


System performance: last 30 days
NEMO
System performance [%]
days since 01.08.2019 (end: 30.08.2019)