is-enes
INFRASTRUCTURE FOR THE EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING

NEMO: improving computational performance

I. Epicoco, S. Mocavero, F. Mele – CMCC

M. Acosta, M. Castrillo, - BSC

M. Bell, M. Andrejczuk - MetO

ISENES3 – General Assembly – March, 25-27 2020

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

NEMO: computational performance community

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

NEMO improvements

- Single core performance
  - Tiling
  - Loop fusion
  - Mixed precision
- Communication
  - Neighborhood collective communications
  - Extended halo size
- Macro task parallelization
- Multigrid refinement optimization
- I/O
  - Improving read/write with XIOS
  - Online diagnostics
- Support for different architectures
  - GPU
  - DSL
- Containerization

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Single core performance
Loop fusion

- Efficient exploitation of memory hierarchies and hardware peak performance
- **Loop fusion technique** aims at better exploiting the cache memory by fusing DO loops together

```
DO j=1, n-1
   DO i=1, n
      b (i,j) = in(i,j+1) - in(i,j)
   END DO
END DO

DO j=2, n-1
   DO i=1, n
      out (i,j) = b(i,j) - b(i,j-1)
   END DO
END DO
```

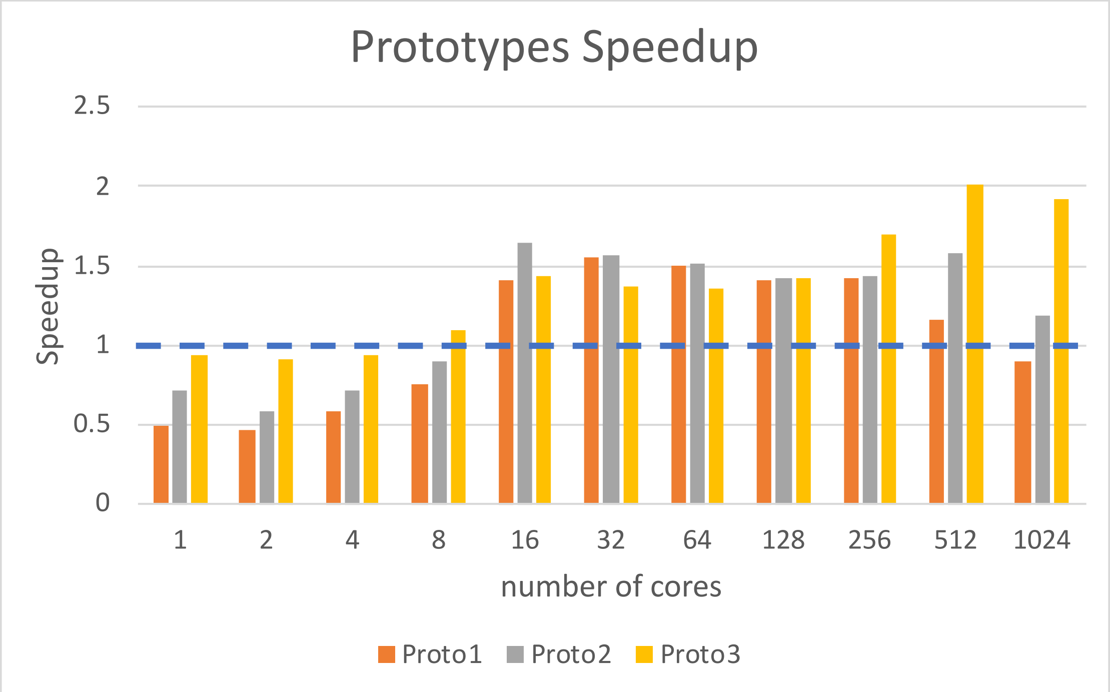
```
DO j=2, n-1; DO i=1, n
      b_0  = in(i,j+1) - in(i,j  ) ! correspond to b(i,j)
      b_m1 = in(i,j  ) - in(i,j-1) ! correspond to b(i,j-1)

      out(i,j) = b_0 - b_m1
END DO; END DO
```

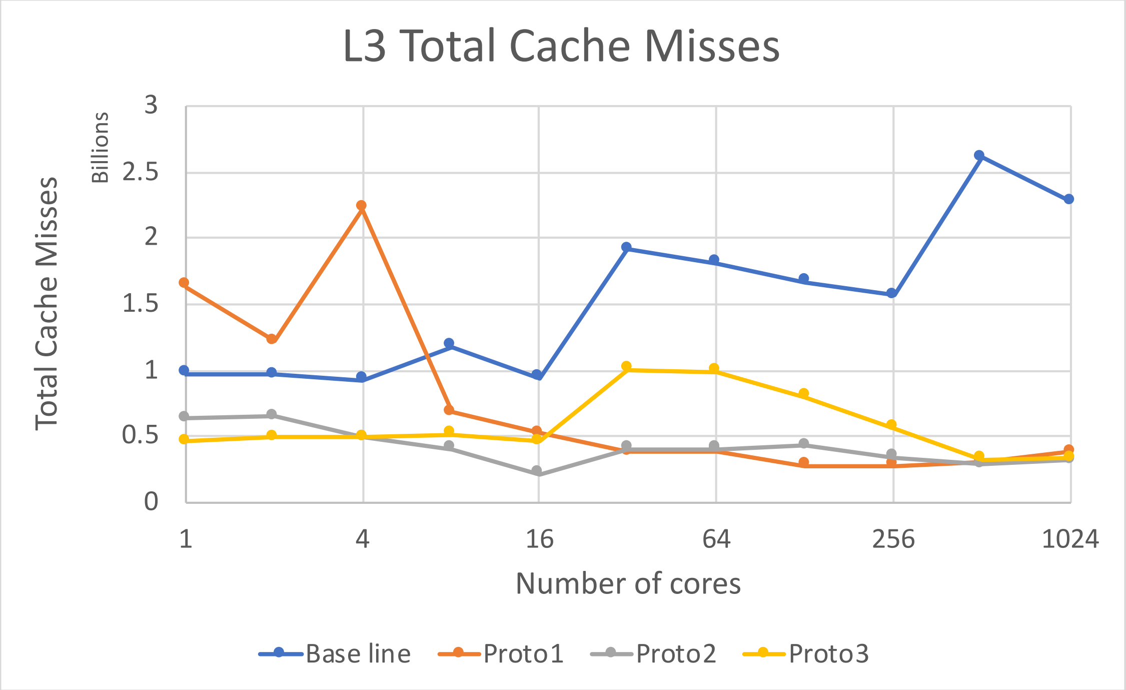
The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

IS-ENES3 1st General Assembly
25-27th March 2020

Single core performance
Loop fusion

- Three different levels of fusion have been implemented
- proto1: has the maximum level of fusion with redundant operations
- proto2: introduces the buffers rotation in the outer loop
- proto3: uses the buffers rotation in the outer and middle loop





The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

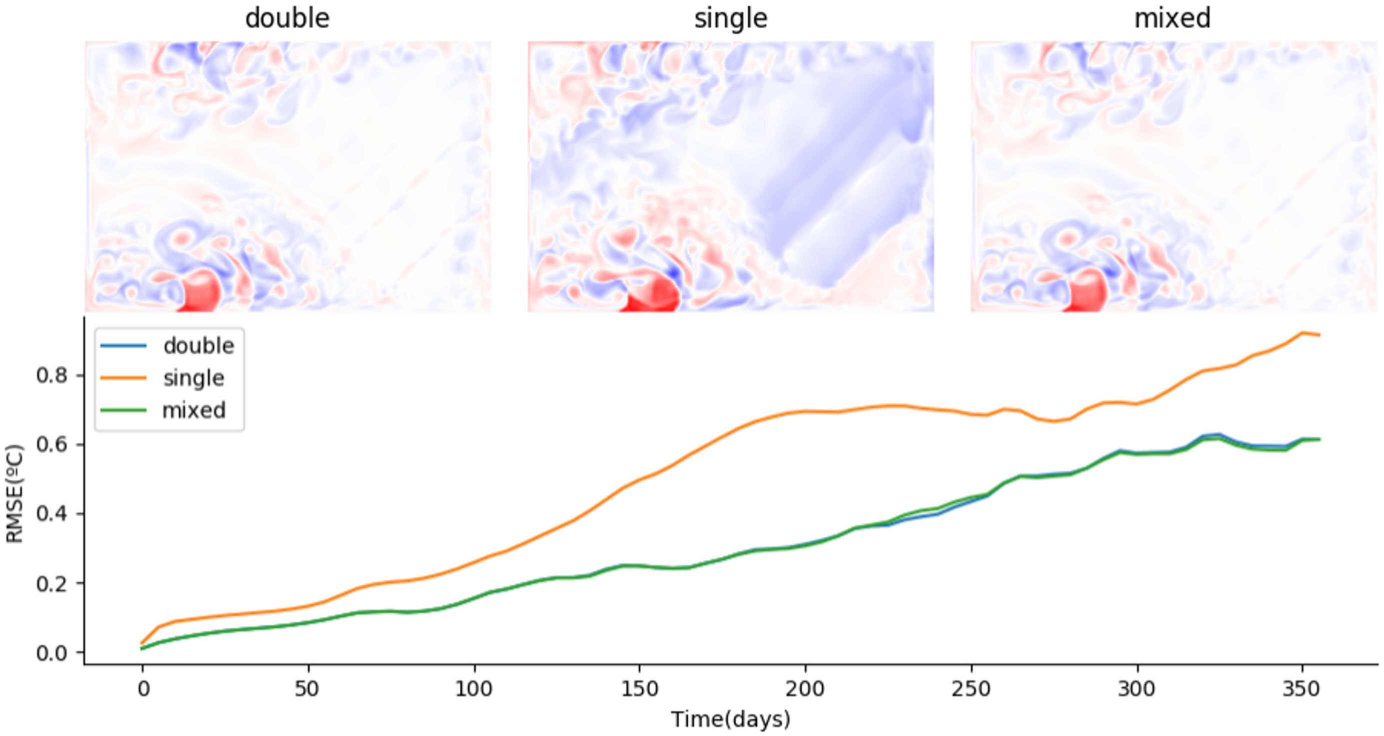
Single core performance
Tiling

- Tiling allows us to divide the calculation into chunks of work that can remain cache-resident for as long as possible.
- The technique leaves the tile size and shape as tunable parameters, which can be tuned appropriately for cache sizes on any platform.
- Preliminary tests established that the CPU time taken by some typical 3D routines within NEMO using current configurations could be reduced by at least a factor of 2 by 3D tiling

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Single core performance Mixed Precision

Impact of precision on sea-surface temperature in NEMO4:
comparison of GYRE1/9º simulations using different precisions



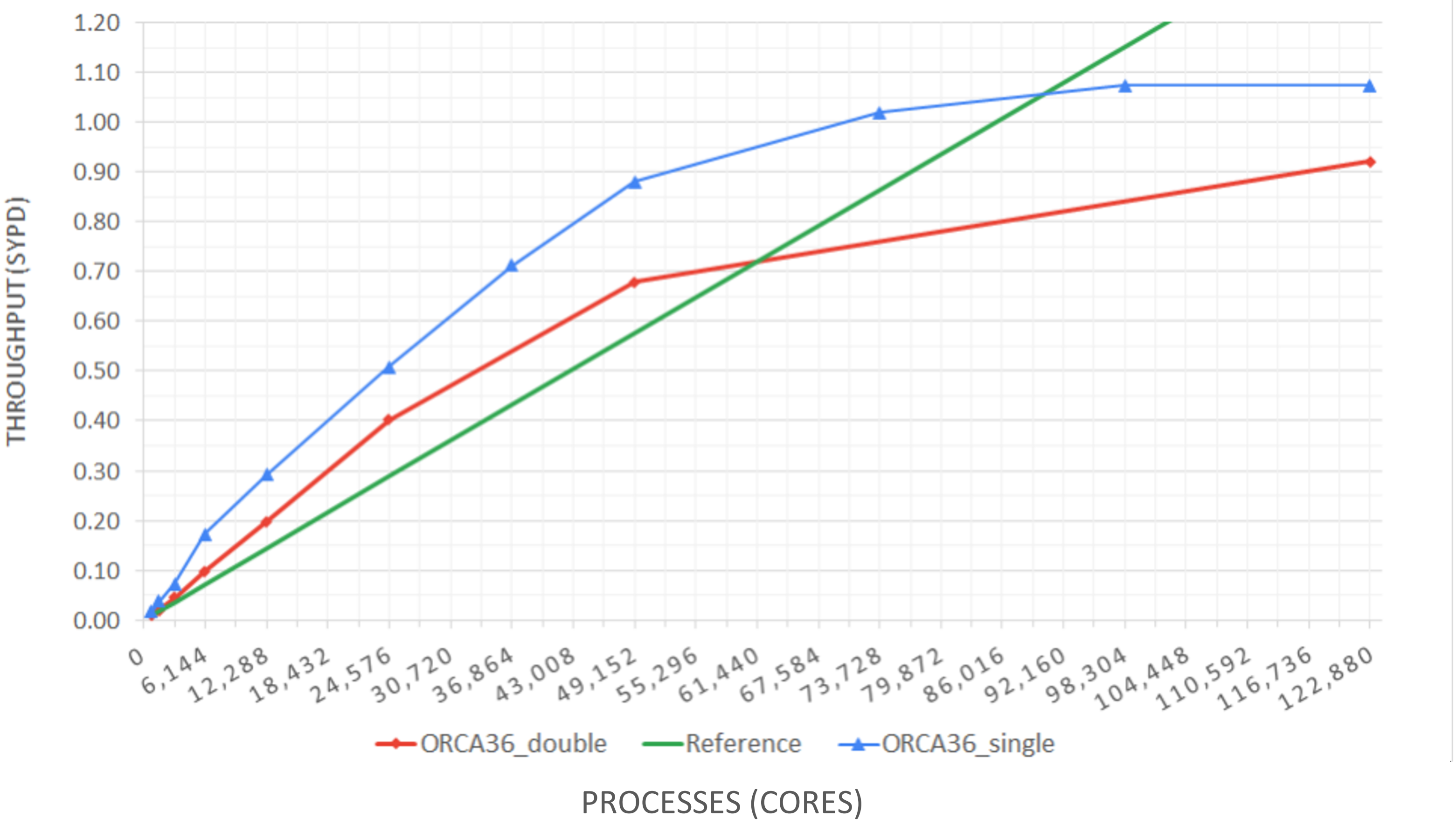
Mixed-precision approaches can provide performance benefits while keeping the accuracy of the results.
In implementation phase this 2020.

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Single core performance
Mixed Precision

ORCA36 scalability – Double vs Single precision – Grand challenge

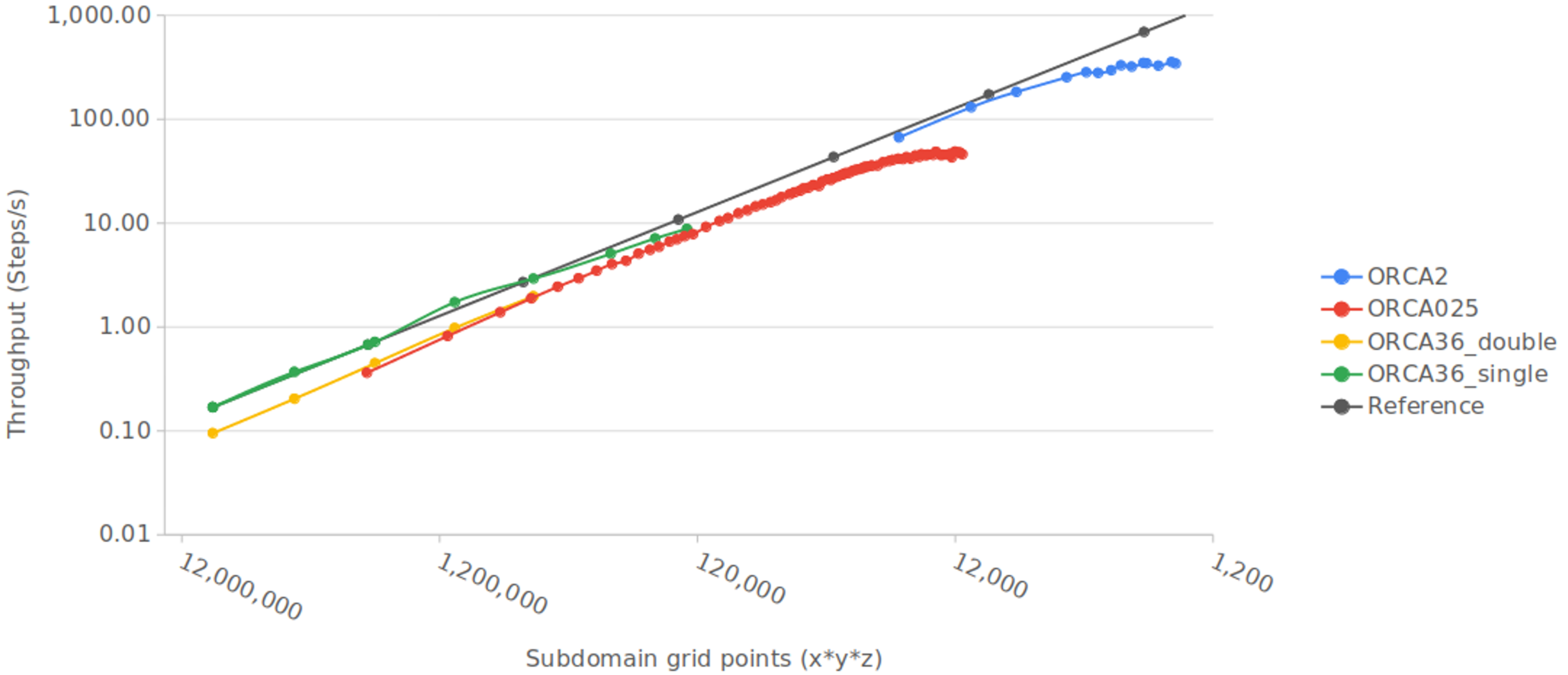


Barcelona Supercomputing Center
Centro Nacional de Supercomputación

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Single core performance Mixed Precision

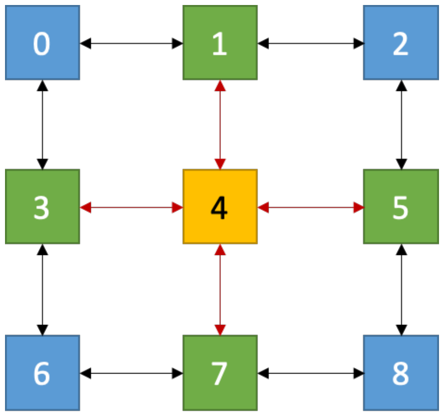
ORCA2, ORCA025 and ORCA36 scalability

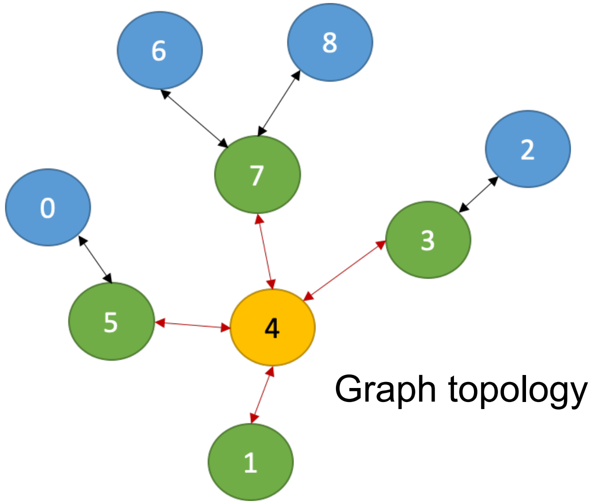


The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

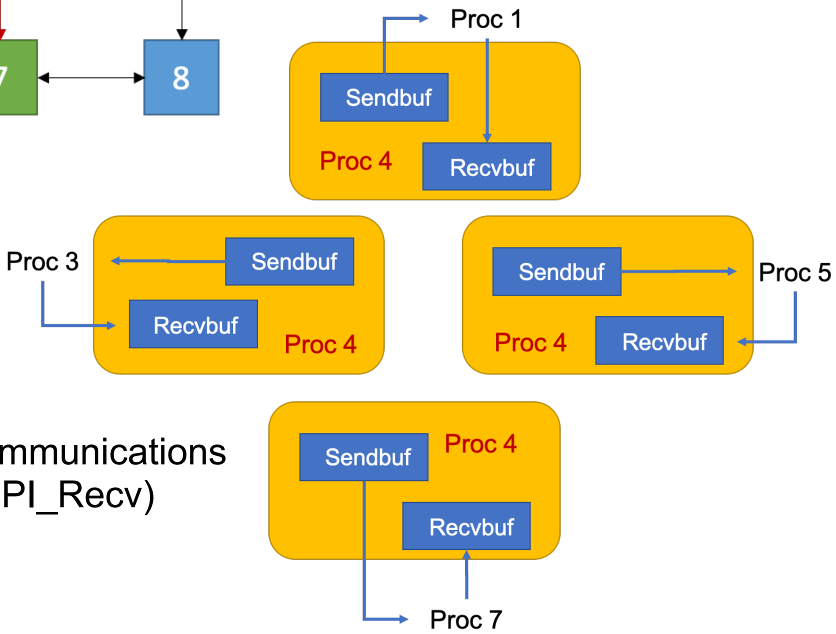
MPI Communication
Neighborhood collectives

Cartesian topology



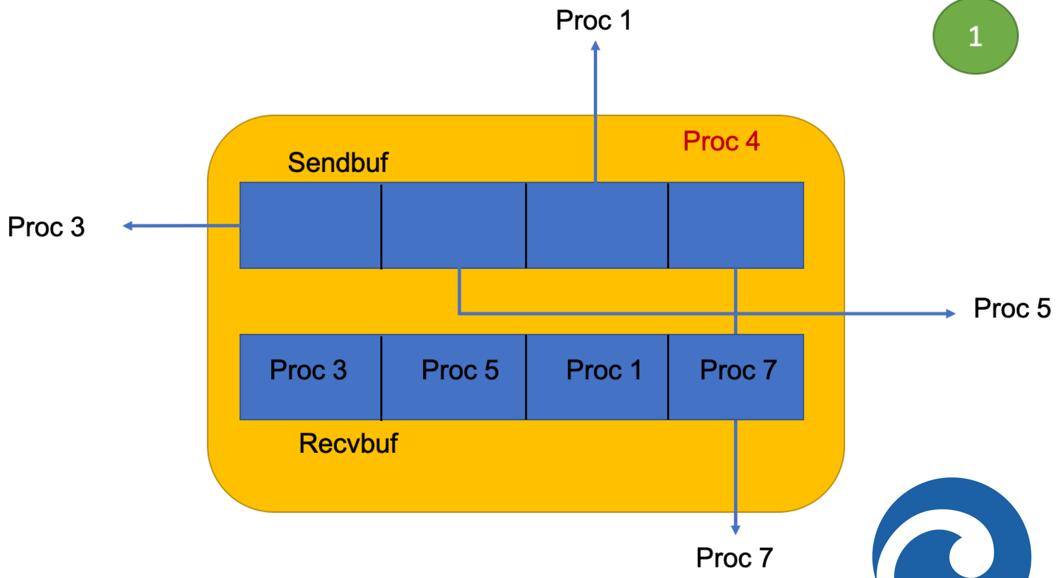


Graph topology



4 Point-2-Point communications
(MPI_Send/MPI_Recv)

1 collective communication
(MPI_Neighbor_alltoall)

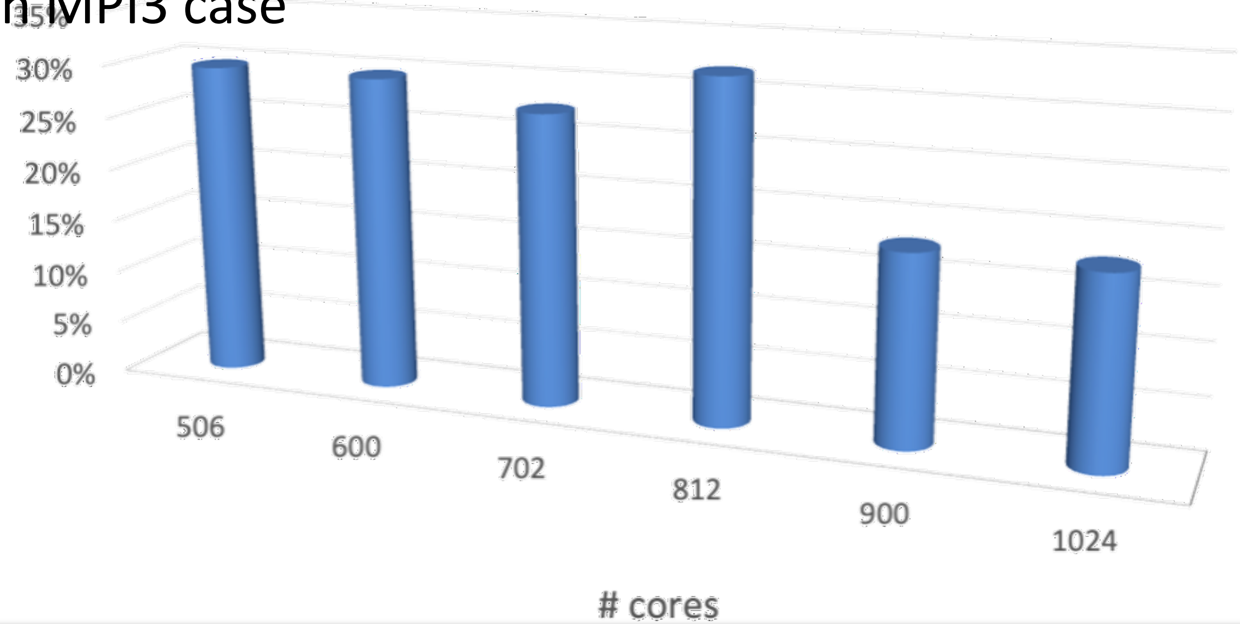


The IS-ENES3 project has received funding from the European Union's Horizon 2020
research and innovation programme under grant agreement No 824084

MPI Communication Neighborhood collectives

- Extension of LBC (Lateral Boundaries Condition) module to support MPI3 Neighborhood Collectives:
  - New Cartesian communicator
  - Ranks reordering to match NEMO processes order
  - Data buffer handling
  - Implementation of multi field exchange in MPI3 case

- Test on the advection scheme
  - GYRE_PISCES configuration (nn_GYRE=200 → ~6000x4000x31 grid resolution)
  - Communication time improved within a range of 18%-32%



The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

MPI Communication Extended halo

- The halo management support has been introduced to provide the developer with a tool for specifying different halo sizes for different NEMO Kernels.
- As an example, in the surface pressure gradient kernel it could be convenient to enlarge the halo region (up to 5 lines or more) to reduce the communication steps in the time sub-stepping loop. In all of the other kernels the halo size could be 2 lines
- A wider halo size reduces the frequency of message exchanges whilst increases the message size at each exchange.

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Macro Task Parallelism

- Parallelize OPA (ocean module) and TOP-PISCES (tracer advection biogeochemistry -BGC- module) into two executables and ensure 3D coupled fields exchange via the community coupler OASIS.
  - Strategy can be decomposed in 6 steps:
  - Set-up of an up-to-date TOP-PISCES stand-alone version
  - Identification of the exchanged coupling fields
  - Duplication of the existing surface (atmosphere) interface (sbccpl)
  - OASIS parameter and input file building
  - Validation of the coupled simulation by comparison with a similarly lasting simulation led with the standard online NEMO-TOP-PISCES configuration
  - Load balancing and performance measurement

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

I/O optimization through XIOS

- Improvement on I/O reading initial conditions and reading regridding weights using XIOS
- the XIOS support has also been adopted for reading and writing of the restart files in the SI3 (sea ice model).

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Online diagnostics – GPU based

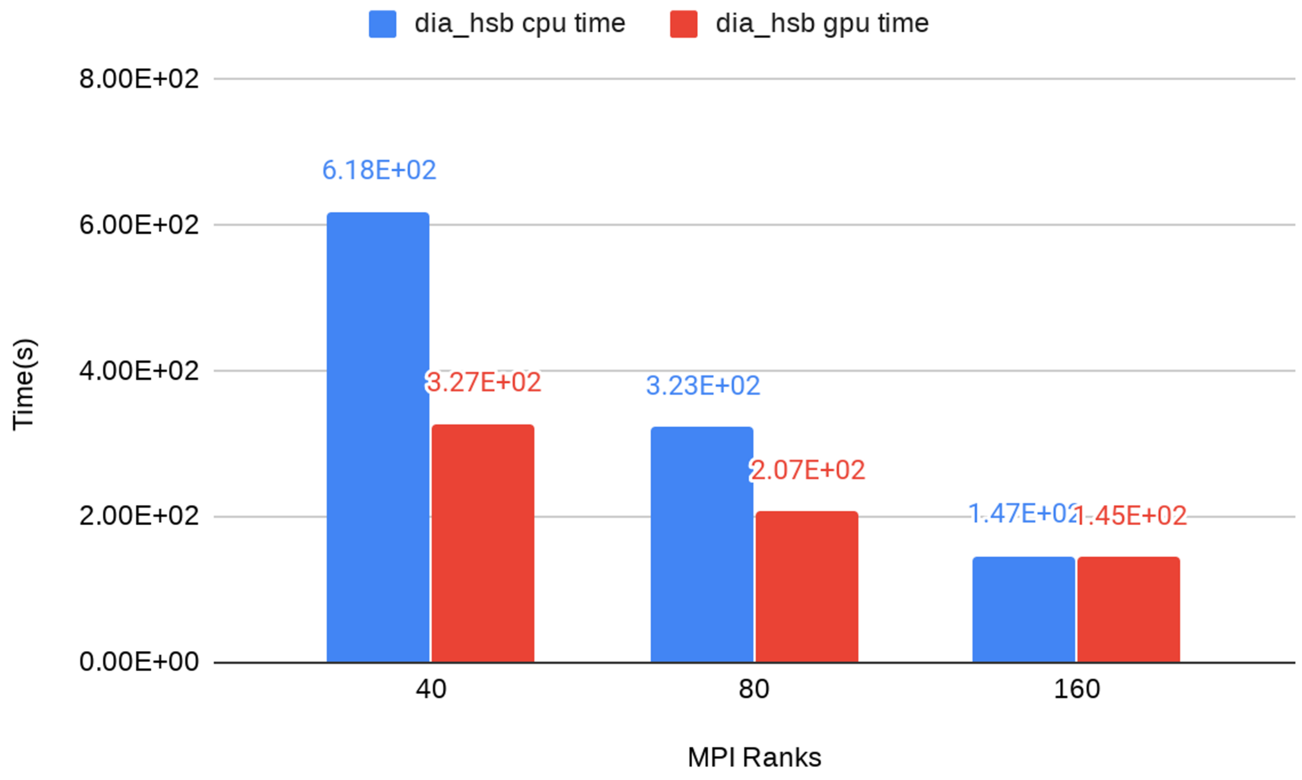
- The rationale of this activity is to improve the NEMO computational performance by offloading the computations for diagnostics on GPU.
- The ocean global heat content, salt content and volume conservation diagnostics (`dia_hsb`) has been chosen as starting point because it is the most expensive.
- The code itself is executed 50x faster than in a single CPU but the data transfer to and from GPU is the main bottleneck.
- Pinned Memory and GPU Directly Attached to the host can be used to mitigate the data transfer penalty

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Online diagnostics – GPU based

`dia_hsb` scalability



Barcelona Supercomputing Center
Centro Nacional de Supercomputación

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Online diagnostics – HPDA based

- The Ophidia High Performance Data Analytics (HPDA) framework has been used for offloading model diagnostics from NEMO to additional parallel cores.
- The Ophidia software architecture decouples the analytics phase from the underlying storage features via simple API, with the aim of supporting different storage devices.
- a diagnostic concerning the Potential Density has been developed in order to compute three different related metrics:
    - the potential 3D density referred to the surface;
    - the mean 3D density over a given time period;
    - the calculation of the 27.8 density isopycnic.

Centro Euro-Mediterraneo sui Cambiamenti Climatici

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

Multigrid capability

- The support for nested multigrid in NEMO is implemented in the AGRIF component
- In realistic configuration, AGRIF is not optimized and lacks an efficient load balancing
- NEMO model has been updated to provide an estimation of the computational cost of each grid; a new load balancing policy can be implemented in AGRIF

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

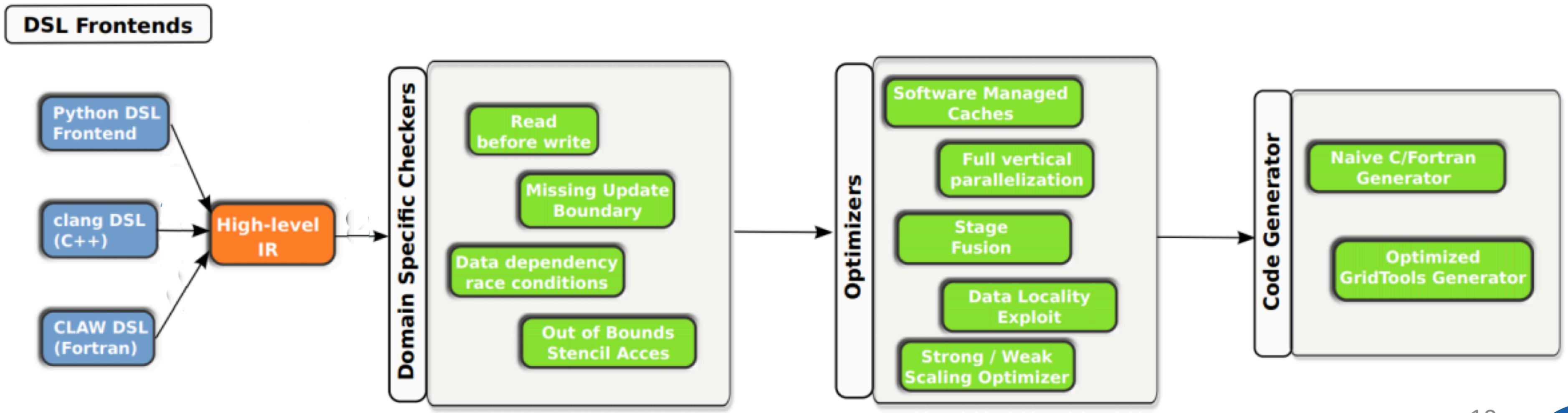
NEMO on GPU

- Investigate NEMO performance on GPUs
- Identify parts of the code preventing efficient execution on GPUs
- Use PSyClone to automatically insert OpenACC directives into the code
- The Met Office GO8: 1 deg. NEMO configuration (362x332x75 points) is used for tests; no MPI
- This configuration (without SI3) is running on 1 GPU 4 times faster than on 1 core; with SI3 the code on GPU is only slightly faster than on 1 core.

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

DSL GTClang for NEMO

- Domain Specific Language GTClang has being enhanced to support NEMO requirements (i.e. regular grid, numerical integration schema, computational kernels)
- Preliminary evaluation of GTClang through porting of specific “dwarf” which represent the advection schema (MUSCL) used in NEMO



The IS-ENES3 project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 824084

cmcc Centro Euro-Mediterraneo sui Cambiamenti Climatici

DSL GTClang for NEMO

```
DO jk = 1, jpkm1
 DO jj = 1, jpjm1
  DO ji = 1, fs_jpim1
   zwx(ji,jj,jk) = umask(ji,jj,jk) * ( ptb(ji+1,jj,jk,jn) - ptb(ji,jj,jk,jn) )
  END DO
 END DO
END DO

DO jk = 1, jpkm1              !-- Slopes
 DO jj = 2, jpj-1
    DO ji = 2, jpi-1
    zslpx(ji,jj,jk) = zwx(ji,jj,jk) + zwx(ji-1,jj,jk)
    END DO
 END DO
END DO

DO jk = 1, jpkm1                 !-- Horizontal advective fluxes
 DO jj = 2, jpj-2
    DO ji = 2, jpi-2
    zu  = pun(ji,jj,jk) / ( e1u(ji,jj) * e2u(ji,jj) * fse3u(ji,jj,jk) )
    zflux(ji,jj,jk) = pun(ji,jj,jk) * ( ptb(ji+1,jj,jk,jn) + zu * zslpx(ji+1,jj,jk) )
    END DO
 END DO
END DO

DO jk = 1, jpkm1                 !-- Tracer advective trend
 DO jj = 3, jpj-2
    DO ji = 3, jpi-2
    zu = 1. / ( e1t(ji,jj) * e2t(ji,jj) * fse3t(ji,jj,jk) )
    pta(ji,jj,jk,jn) = pta(ji,jj,jk,jn) - zu * ( zflux(ji,jj,jk) - zflux(ji-1,jj,jk) )
  END DO
 END DO
END DO
```

```
stencil advection_MUSCL {
  do {
    vertical_region (k_start, k_end - 1) {
      zwx = u_mask * (ptb(i+1) - ptb);
    }
      //-- Slopes of tracer
    vertical_region (k_start, k_end - 1)
      zslpx = zwx + zwx(i-1);
    }
      //-- Horizontal advective fluxes
    vertical_region (k_start, k_end - 1) {
      zu = pun / (e1u * e2u * fse3u);
      zflux = pun * (ptb(i+1) + zu * zslpx(i+1));
    }
      // Tracer advective trend
    vertical_region (k_start, k_end - 1) {
      zu = 1.0 / (e1t * e2t * fse3t)
      pta = pta - zu * (zflux - zflux(i-1));
    }
  }
}
```

Pros

- Easy code maintenance
- Improved code readability
- Seamless support for GPU
- Less error-prone code
- Fast and efficient technical support

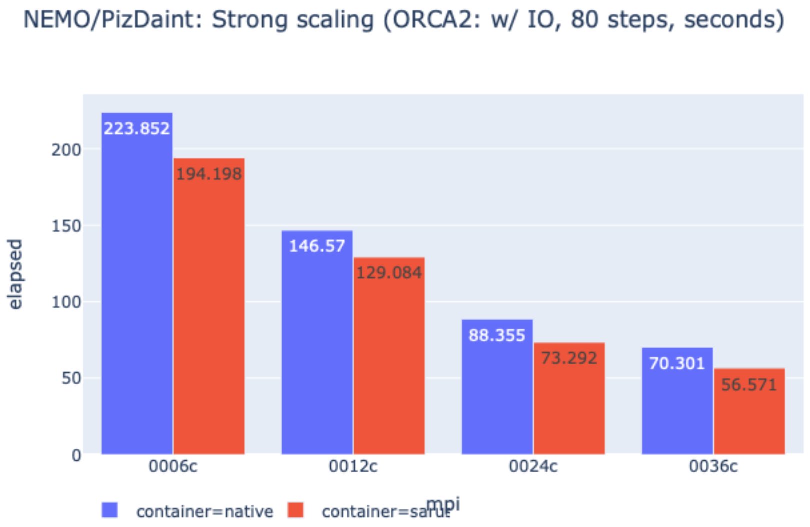
Cons

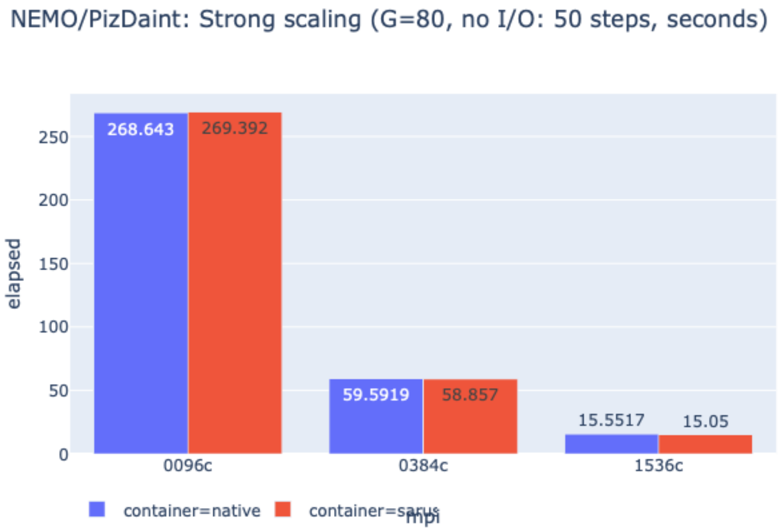
- GTClang environment hard to compile and install
- No documentation
- No MPI support
- Lose of performance w.r.t. the original version

The IS-ENES3 project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824084

NEMO Containerization

- The aim is to build a self-contained installation of the NEMO which includes libraries, compilers and system settings
  - Configuration portability
  - Reproducibility of the same experiments on different machine, on different
- Container manager: SARUS (developed as CSCS)
- Two containers for two configurations were created and tested: GYRE, ORCA2





For more information: https://github.com/eth-cscs/ContainerHackathon

The Consortium

Coordinated by **CNRS-IPSL**, the IS-ENES3 project gathers **22 partners** in **11 countries**

with

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°824084

Our website
https://is.enes.org/

Follow us on Twitter !
@ISENES_RI

Contact us at
is-enes@ipsl.fr

Join the community on ZENODO !