Concurrent Components Execution: the HAMOCC use case

Enrico Degregori (DKRZ)

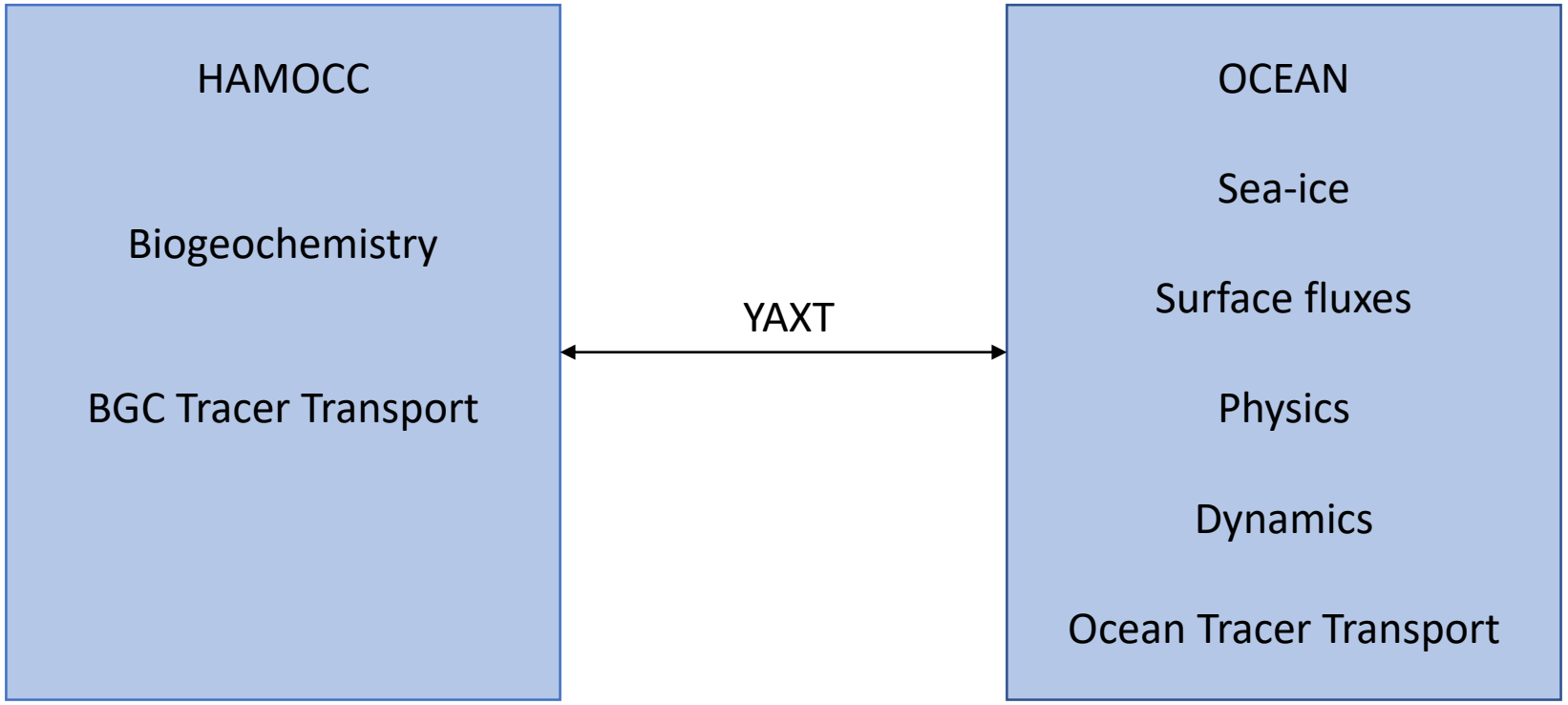
Concept of concurrency

- High Level concurrency: different components running asynchronously
- Concurrent infrastructure components: IO (1 way communication, thus bit-identical results expected)
- Concurrent model components: radiation, biogeochemistry, … (2 way communication, time lag is introduced, thus need scientific validation)

Motivation

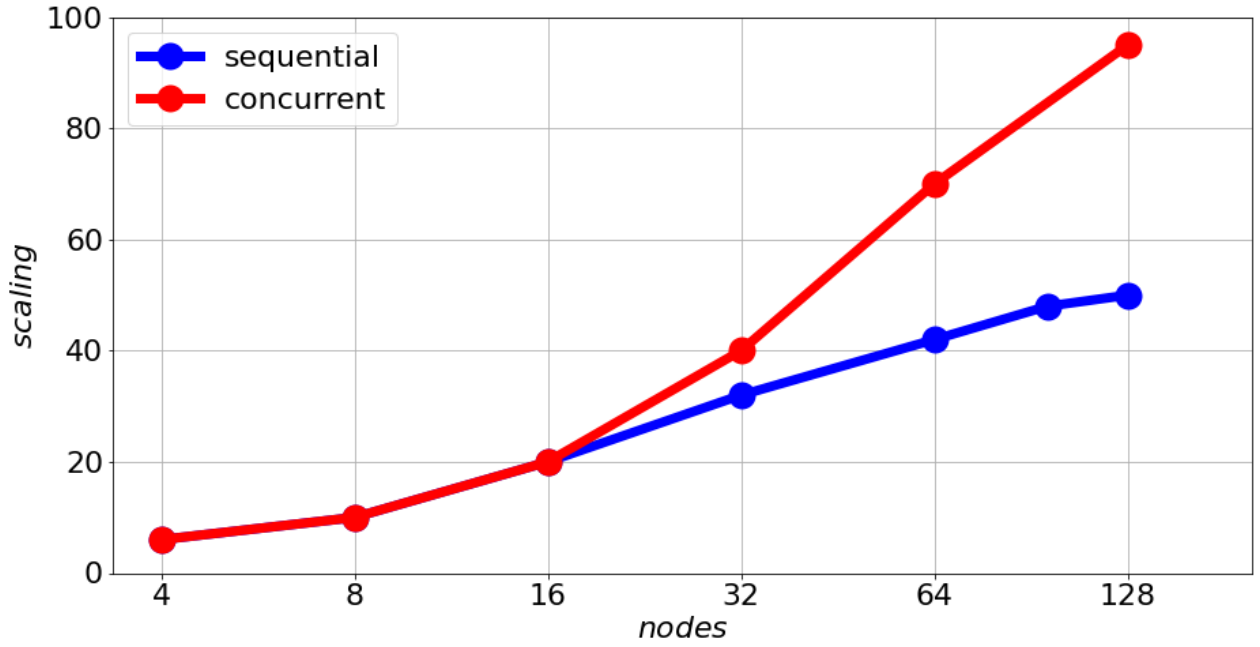
- Scale the model beyond classical domain decomposition
- Use right amount of resources for each component (different scaling)
- Run components on different architectures (heterogeneous systems)

Use Case: HAMOCC (Object Parallelism Approach)



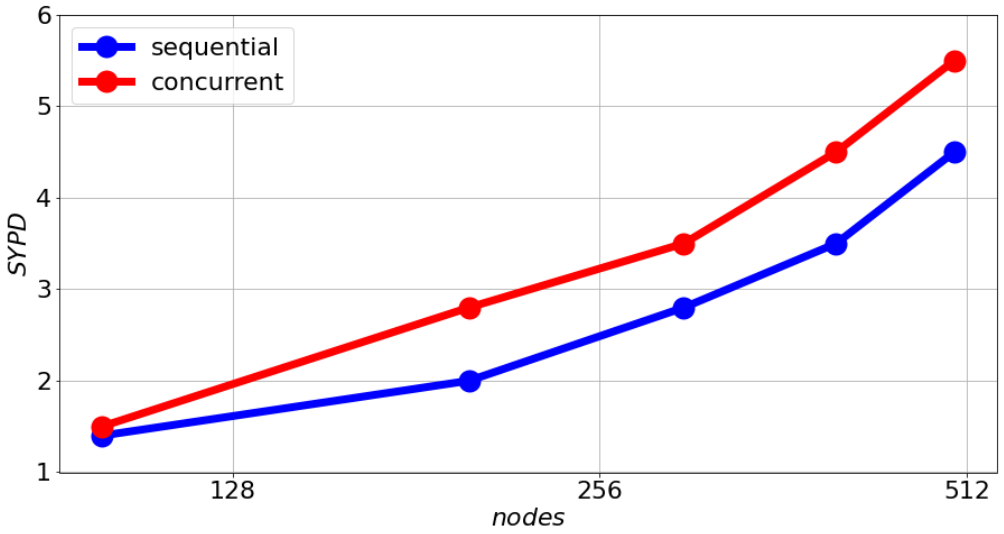
Linardakis et al. 2022

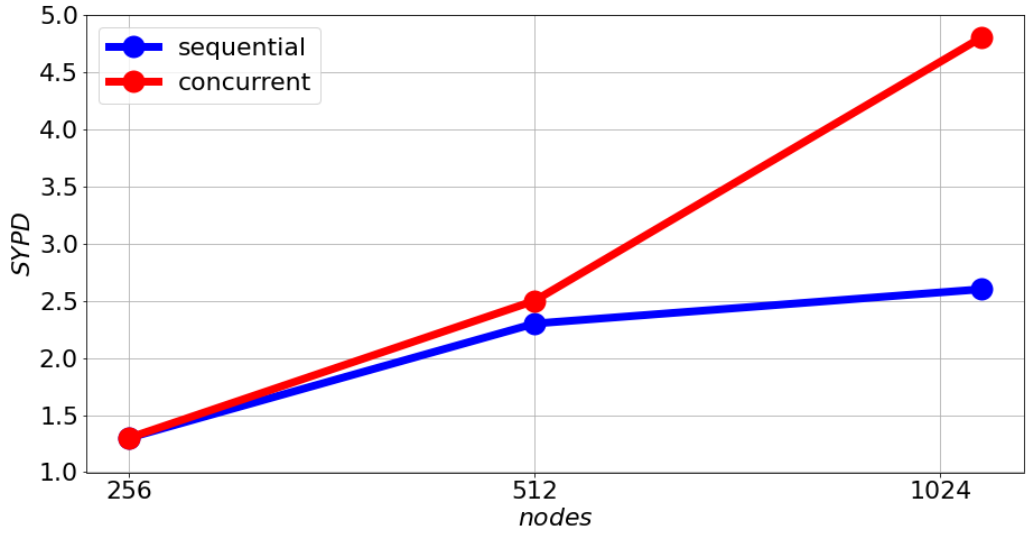
Use Case: HAMOCC (Scaling Model R2B6)



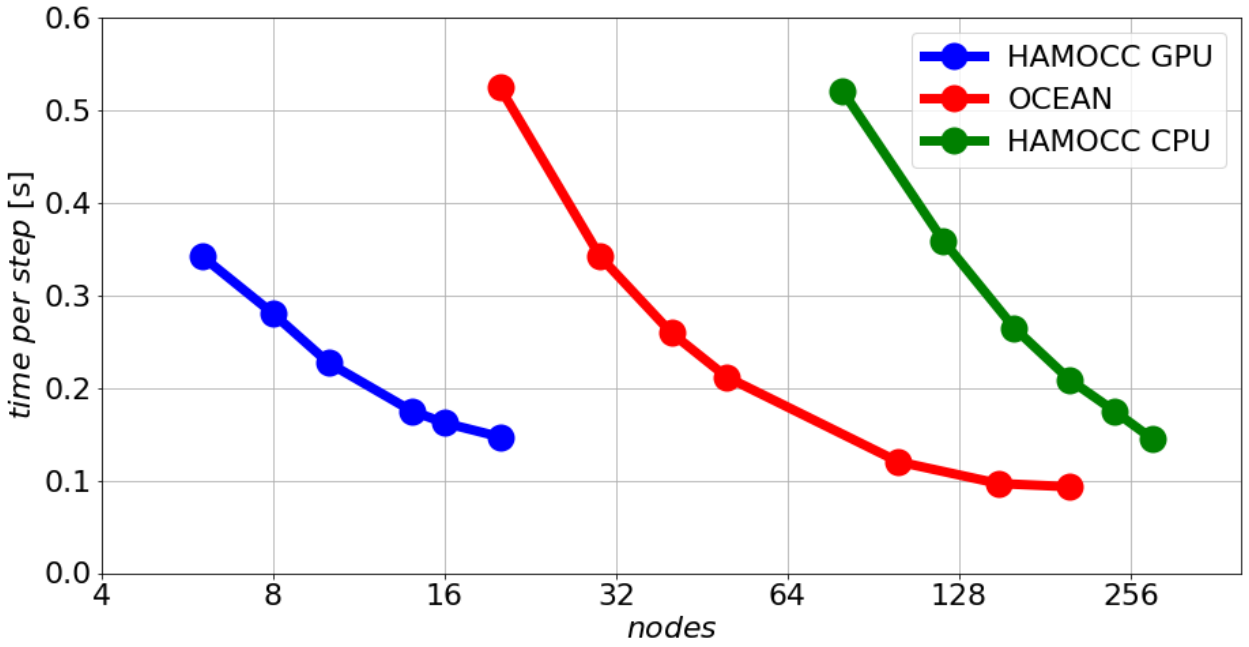
Linardakis et al. 2022

Use Case: HAMOCC (Scaling Model R2B8 / R2B9)





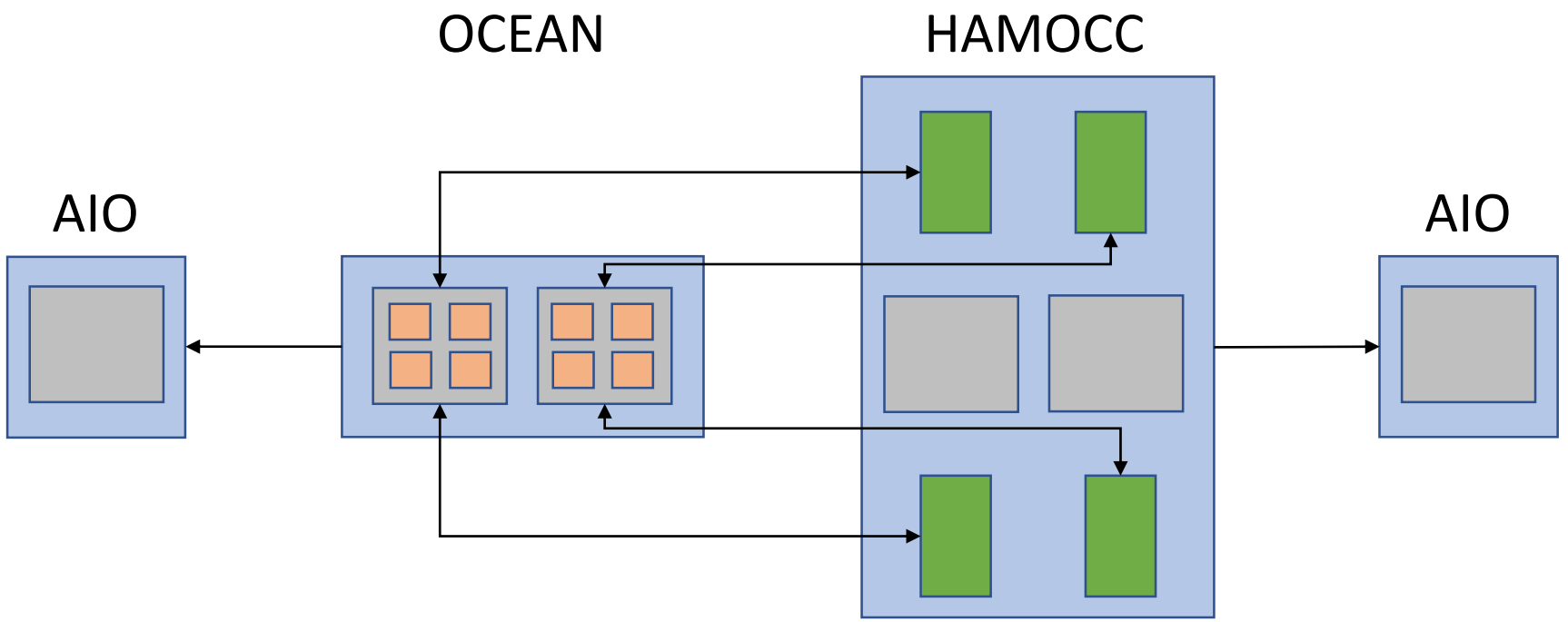
Use Case: HAMOCC (Scaling Components R2B8)



Use Case: HAMOCC (YAXT)

- YAXT simplifies communication between components having **same** grid but **different** domain decompositions
- YAXT is heavily based on **MPI Datatypes** which perform poorly on GPU
- A new **exchanger** is now available which pack/unpack data in a buffer
- It can be set with: XT_CONFIG_DEFAULT_EXCHANGER=irecv_isend_ddt_pack
- If data are on GPU: !$ACC HOST_DATA USE_DEVICE(…)

Use Case: HAMOCC (Heterogeneous Setup)



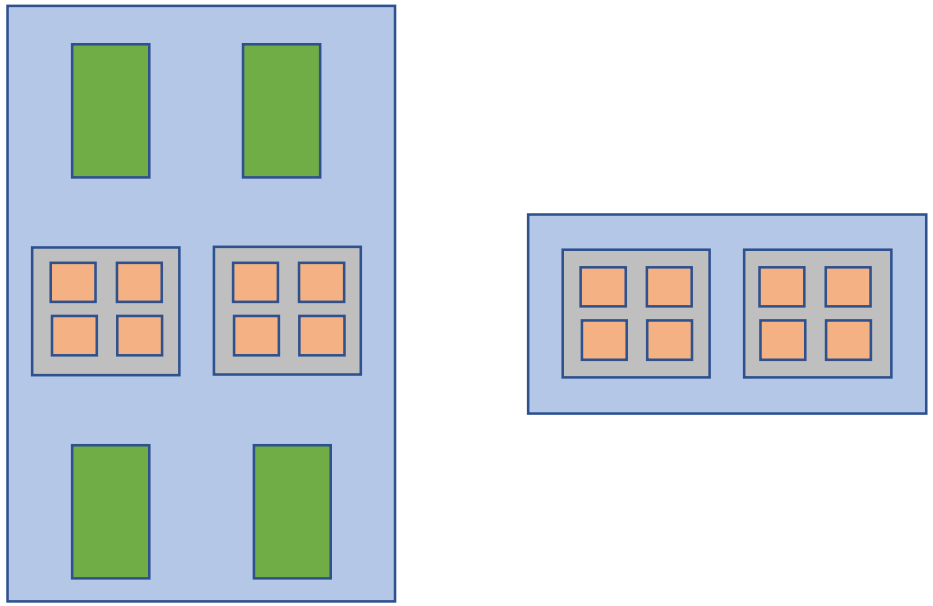
node CPU NUMA GPU

SLURM hetjobs

Use Case: HAMOCC (Heterogeneous Setup)

HAMOCC + OCEAN

OCEAN



- HAMOCC runs on 4 MPI procs per node and 4 GPUs
- OCEAN runs on 28 MPI procs per node and 4 OpenMP threads
- Hetjobs to achieve load balancing

node CPU NUMA GPU

MPMD + SLURM hetjobs

Challenges

- The communication between the two components currently represents the main bottleneck of the heterogeneous setup
- High amount of data to be exchanged because of 3D fields
- High ratio of MPI processes involved in the exchange (50:1)
- Further investigation is needed

Conclusions

- Concurrent approach allows to scale a model beyond domain decomposition
- Component encapsulation is crucial
- It allows to run components on heterogeneous systems
- YAXT hides the complexity of intercomponents communication

Questions?