esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

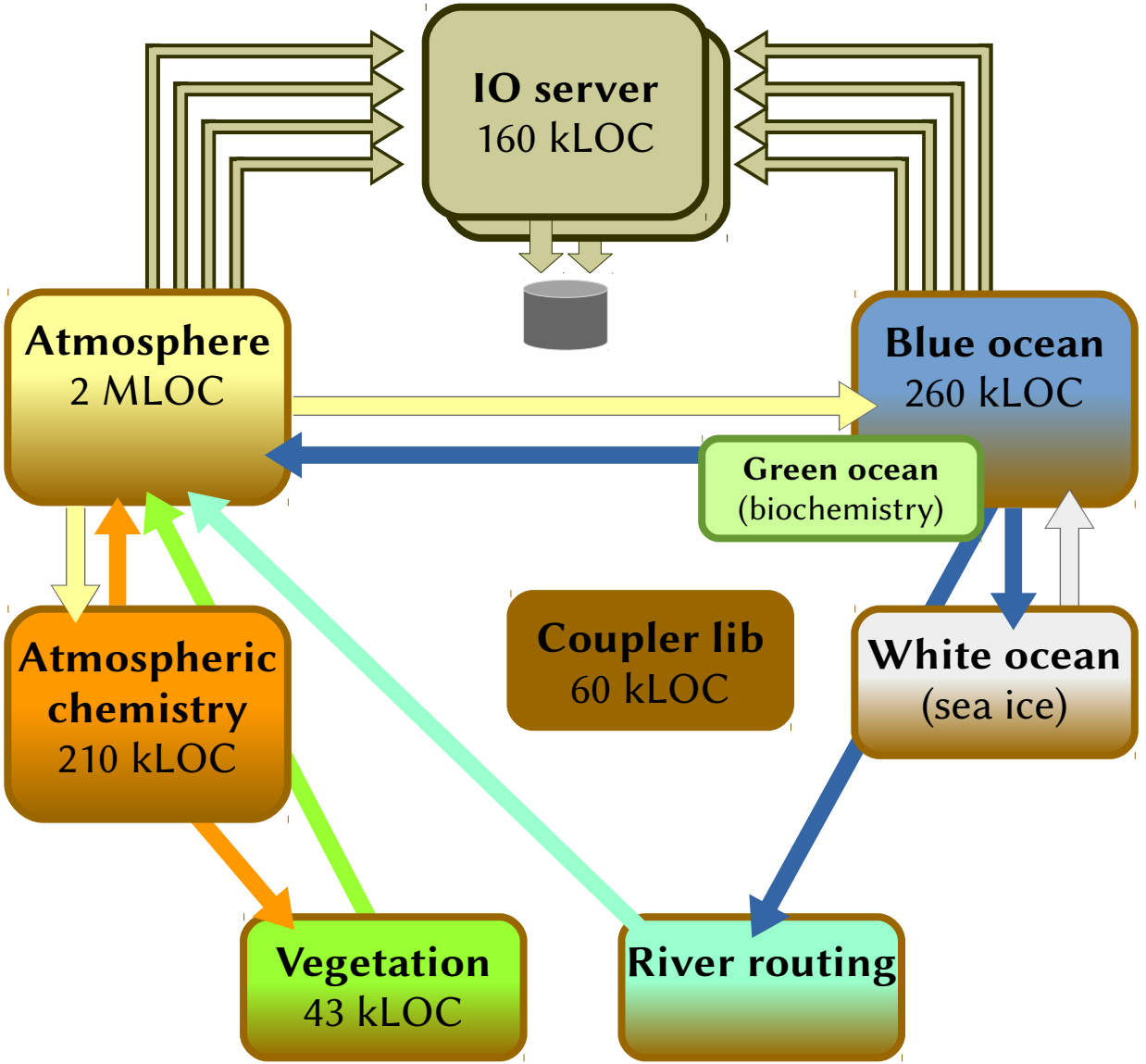

# What is Special about HPC Computing for Weather and Climate?

## A user perspective:
## The EC-Earth Earth System Model

ISC'18 BoF session

Uwe Fladrich, SMHI

# Earth System Models (ESM)

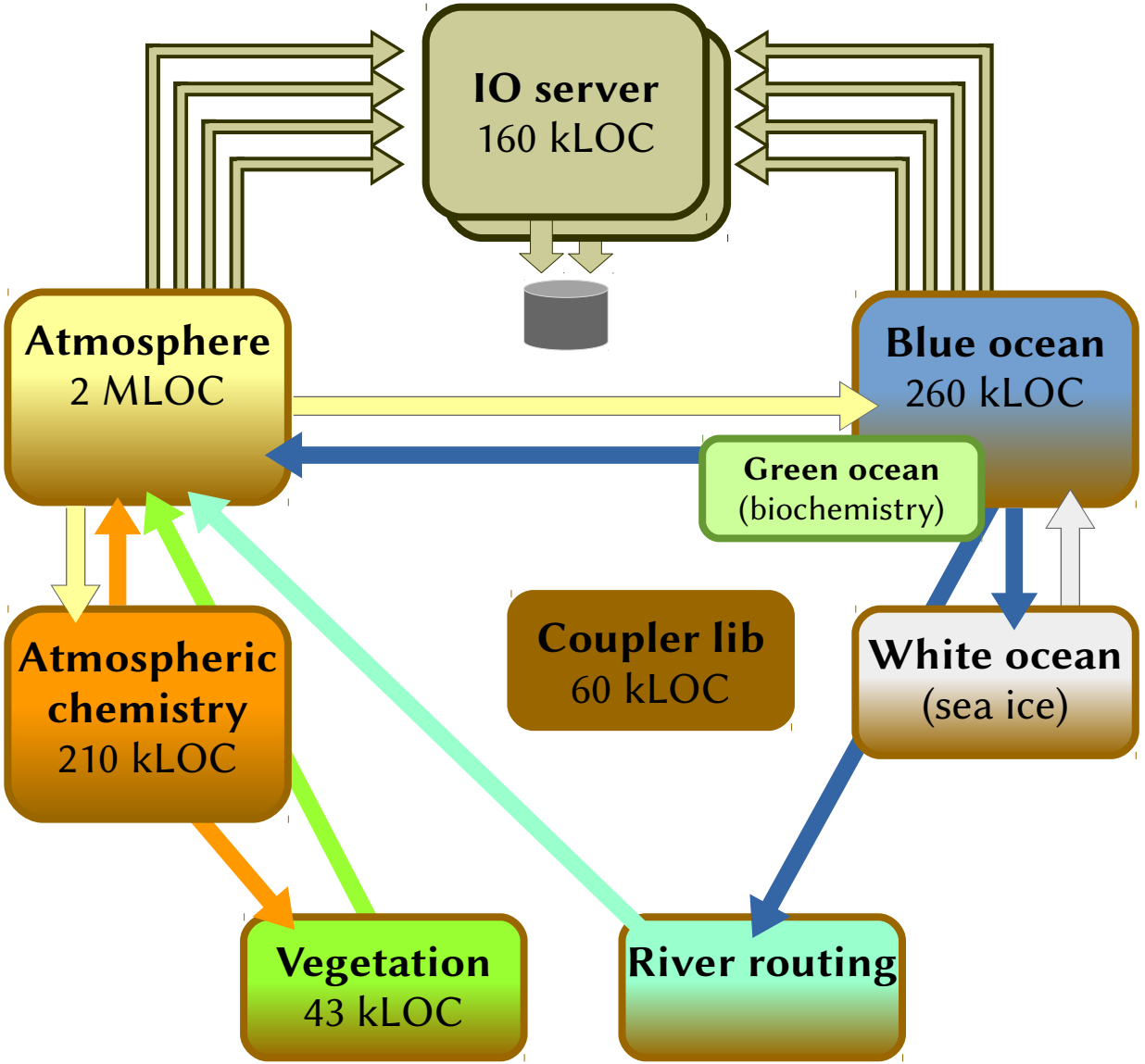



Software/development characteristics:

- Multi-component
- Multi-code
- Multi-language
- Multi-configuration
- Multi-institution

- Legacy code under continuous development

# Earth System Models (ESM)




- Multi-scale, multi-physics
- Sensitivity to initial conditions
- Natural stochastic variability
- Long time integrations
- Weak scalability
- I/O and/or memory boundness

# Need for computing power


ESMs consume computing power along several axes:

- Resolution
  To include processes at finer and finer scale
- Complexity
  To simulate, rather than describe, more and more processes and feedbacks of the climate system
- Ensemble size
  To sample uncertainty across chaotic non-linear dynamics of the underlying complex system

# What's Performance?

- For a given experimental design, what can I afford to run?
- If I add complexity (e.g. biogeochemistry), what will I have to sacrifice in resolution?
- How much computing capacity do I need to participate in a campaign like CMIP6?
  How much data capacity?
- Do the queuing policies on the machine hinder the sustained run of a long-running model?
- During the spinup phase, how long (in wallclock time) before I have an equilibrium state?

# Performance metrics for Climate Computing


Existing performance measures do not provide adequate information about the actual performance during scientific experiments.

Typical questions:

- How long will the experiment take?
- How many nodes can I use efficiently?
- Are there bottlenecks in the work flow?
- How much short/medium/long term storage?
- How many ensemble members to be run in parallel?

# Performance metrics for Climate Computing


- SYPD/ASYPD/JSYPD

  (Actual) Simulated years per day, Joules per simulated year

- Complexity

  Number of prognostic variables per component

- Coupling cost

  Overhead caused by coupling, i.e. cost for coupling algorithm plus load imbalance

- Data output cost

  Cost of performing I/O, i.e. ratio of cost with and without I/O

- Data intensity

  Data produced per core-hour [GB/CH]


esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE
EC EARTH