Building a community hydrological model

Bryan Lawrence

NCAS & University of Reading

Wallingford, 10/09/18

WP1 Description

WP1: Hydro-JULES Community Modelling Framework

- Task 1.1: Design and implement Hydro-JULES modelling framework and interfaces
- Task 1.2: Consult research community and stakeholders on requirements and implementation
- Task 1.3: Build archive of driving data, model configurations, and supporting datasets
- Task 1.4: Provide user training and suport and managed access via JASMIN

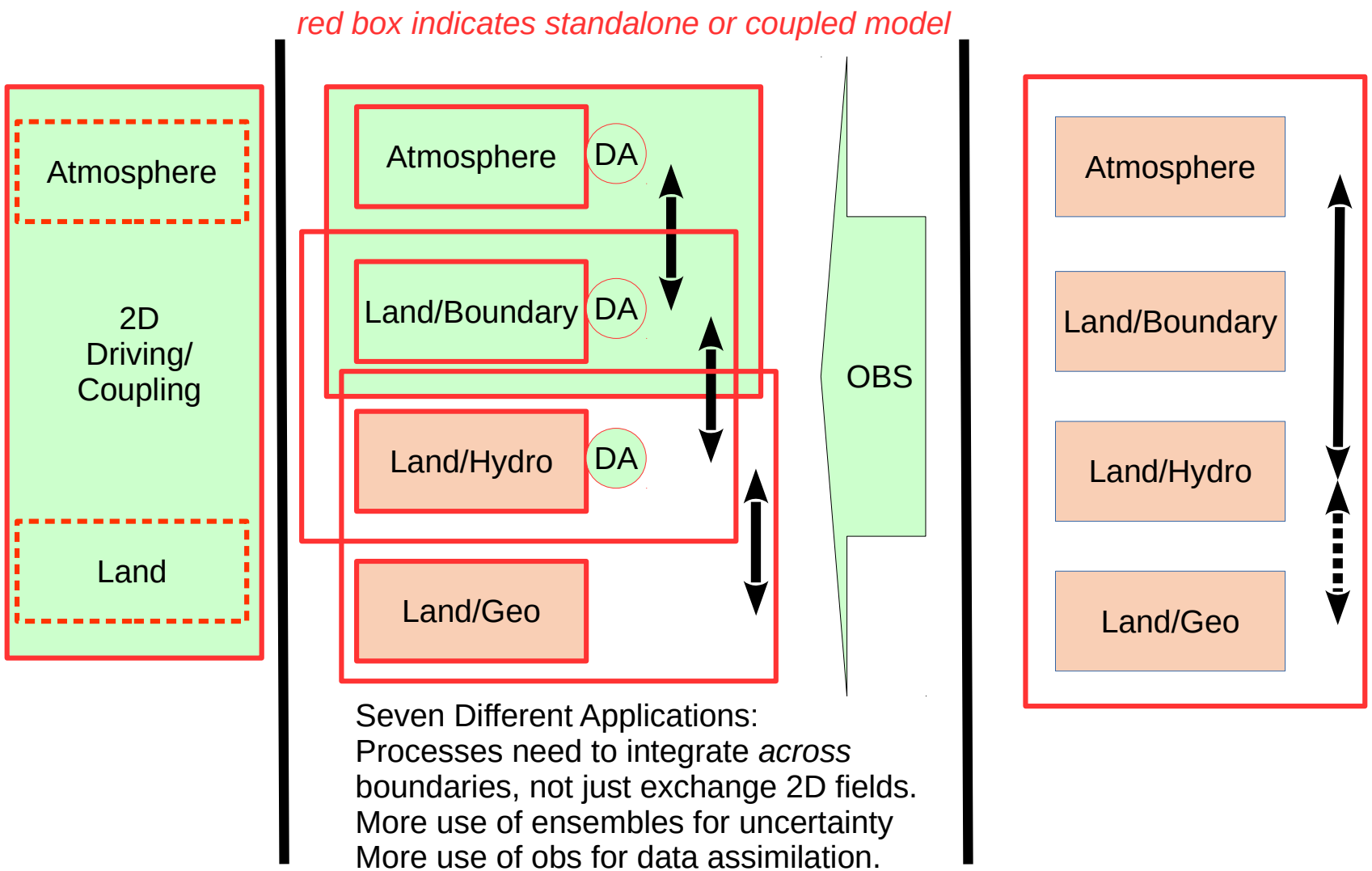
(4 FTE: 2 at CEH, 1.5 at NCAS, 0.5 at BGS)

Framework Objectives (Task 1.1)

Framework Objectives

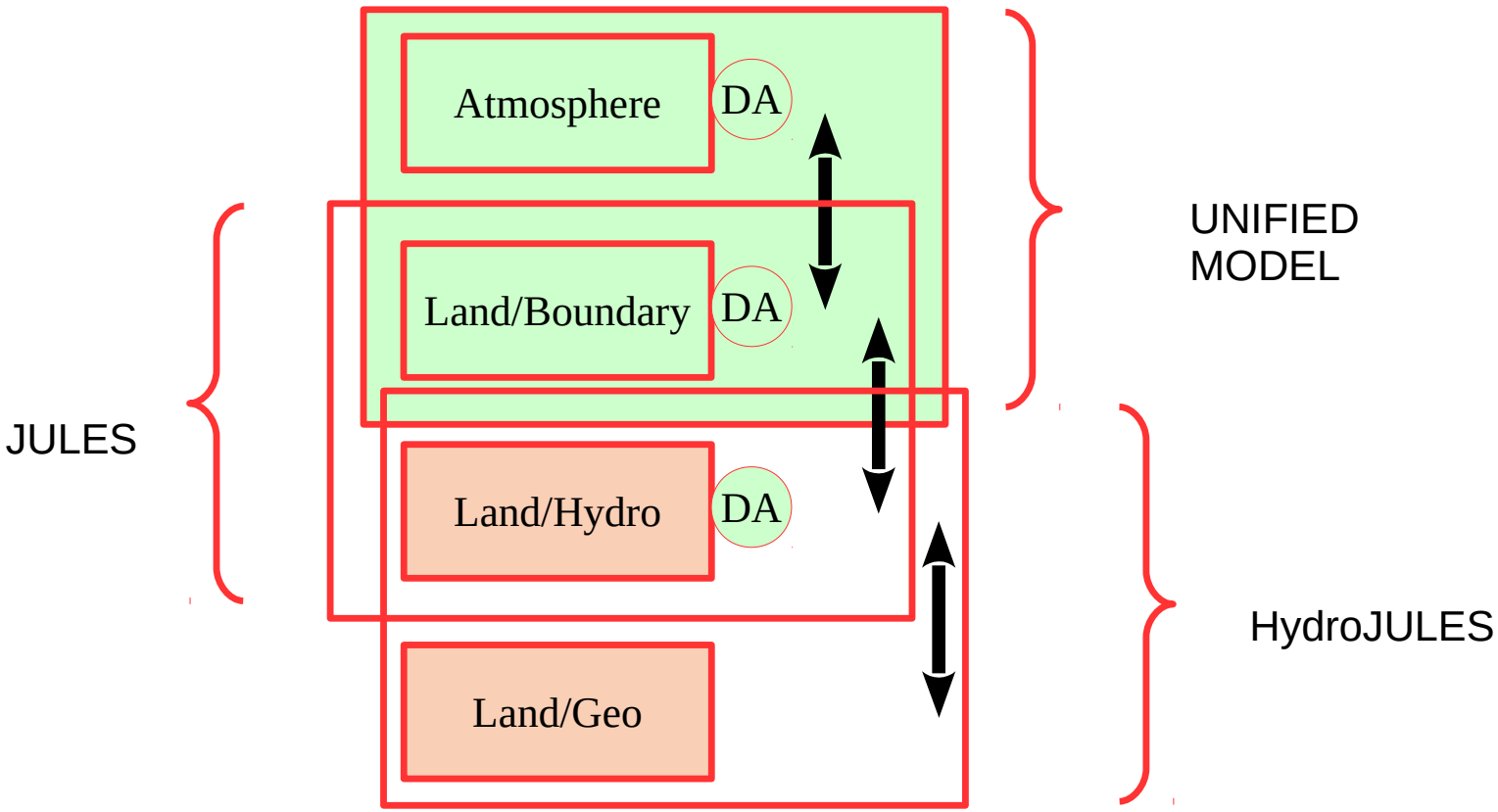
- Provide the "necessary environment": version controlled respository of data and code, and model configurations.
- A specification (or specifications) of the interface necessary for "coupling" between specifically identified model components. Covering
  - Internal component interfaces, and
  - Other models, in particular the UM (via JULES), shelf seas models (which ones?), and models of biogeochemistry and terrestrial nutrient transport (which ones?)
- Functional independence from the science and and support for module interchange (within HydroJules)
- Lifetime of ten to twenty years!

Schematic View of Objectives - I



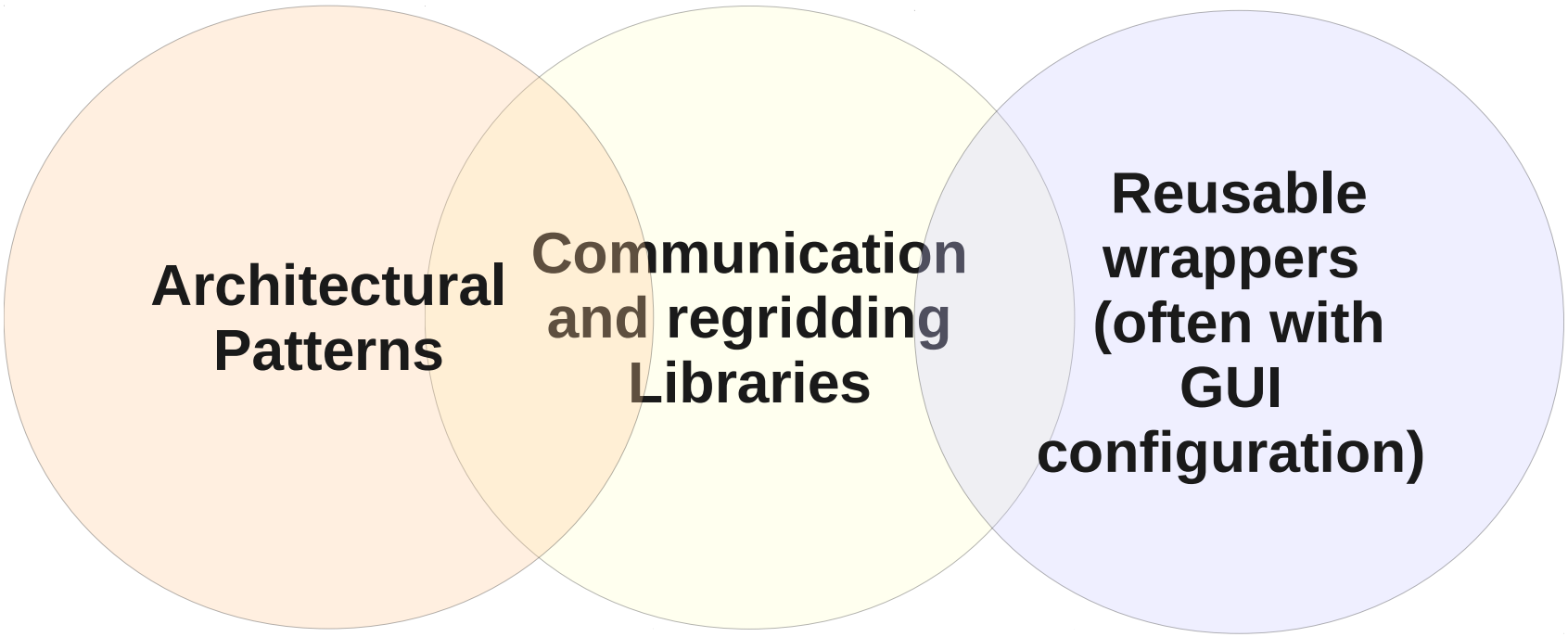
Schematic View of Objectives - II

red box indicates standalone or coupled model

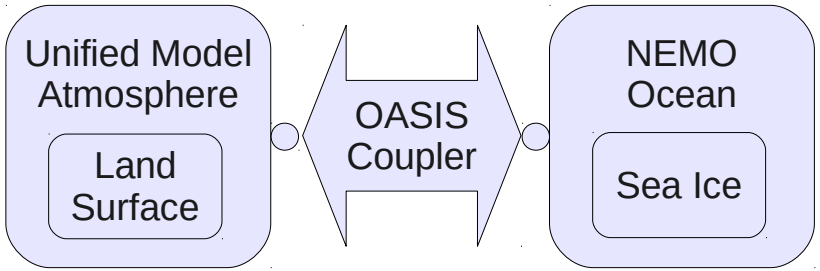


Note expectation that “coupling” may extend through domain of multiple models and doesn’t necessarily simply involve exchanging fields.

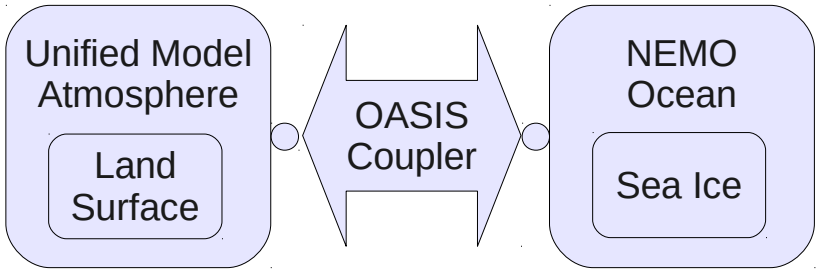
Taxonomy of "Coupling"

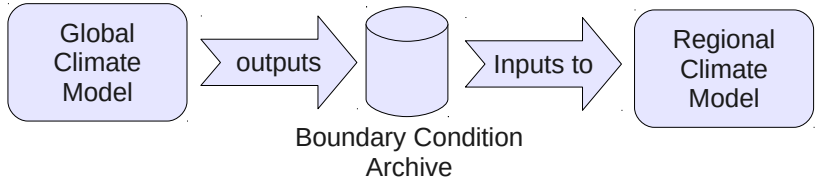


Multiple Modes of "Coupling"

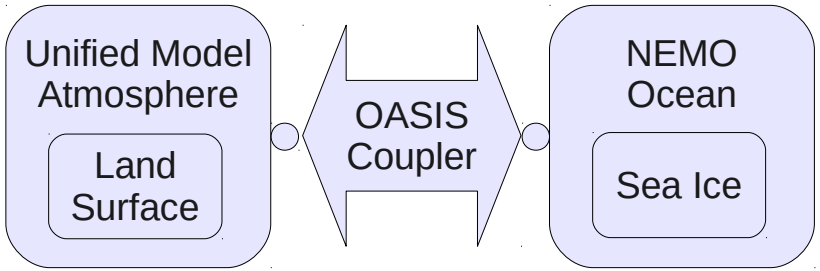


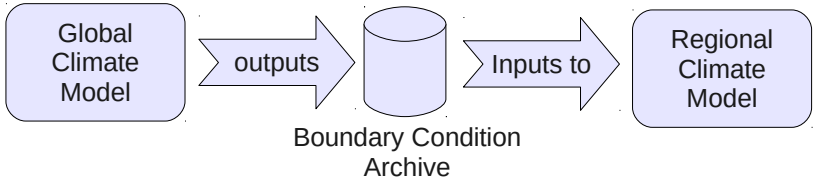
Multiple Modes of "Coupling"

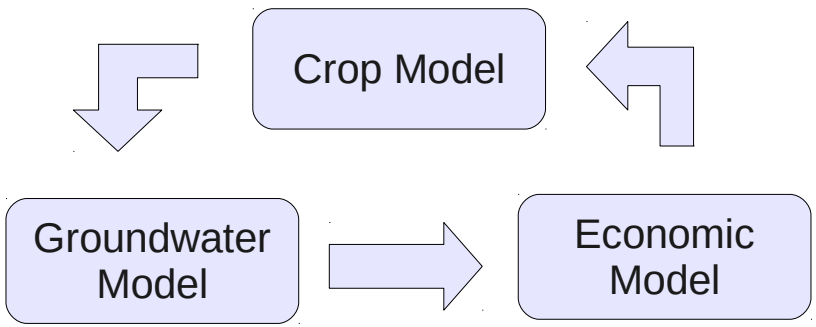




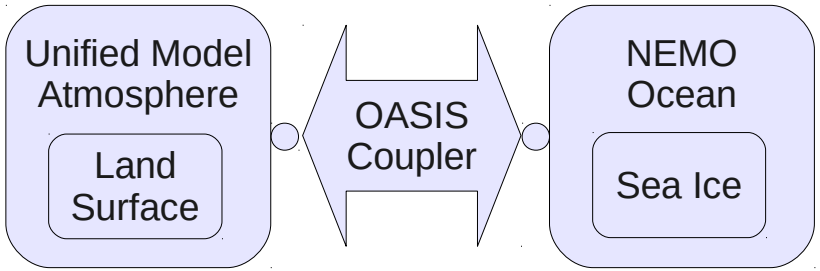
Multiple Modes of "Coupling"

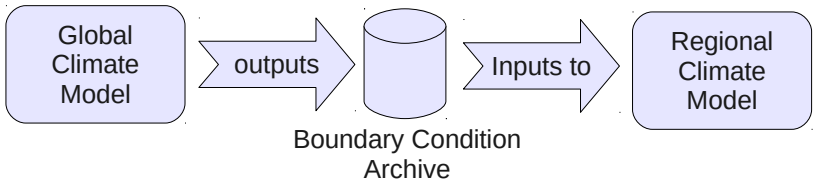


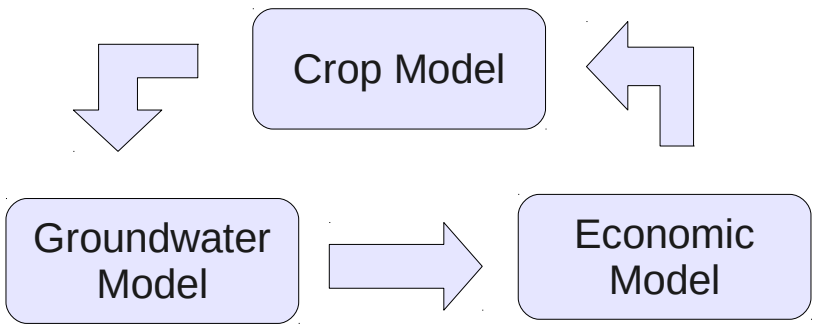


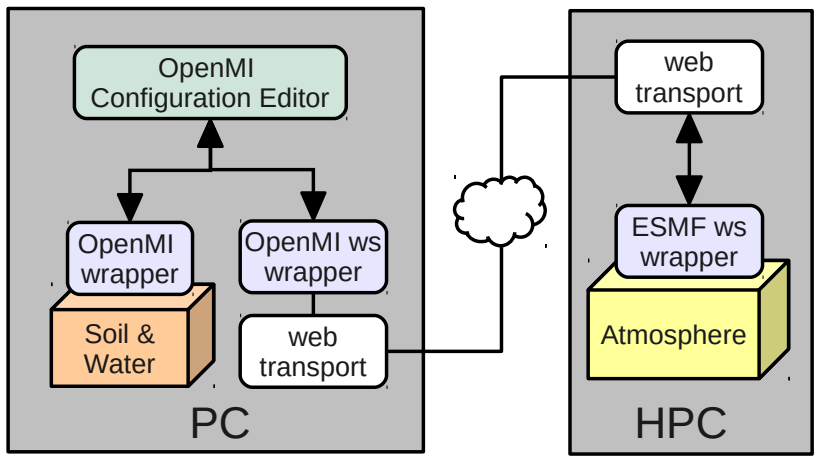


Multiple Modes of "Coupling"

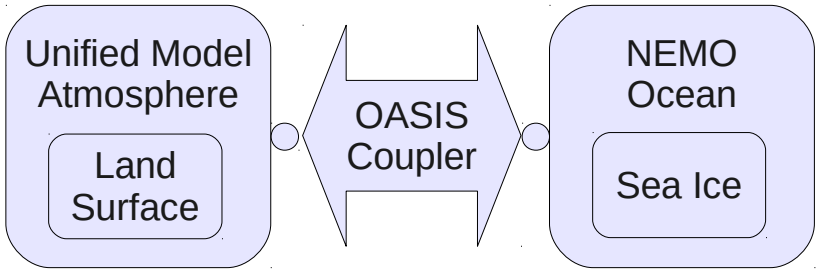


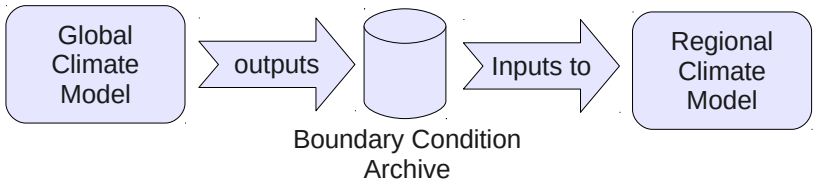


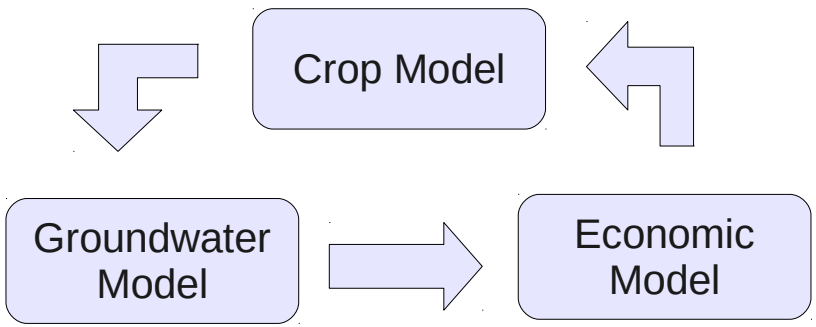


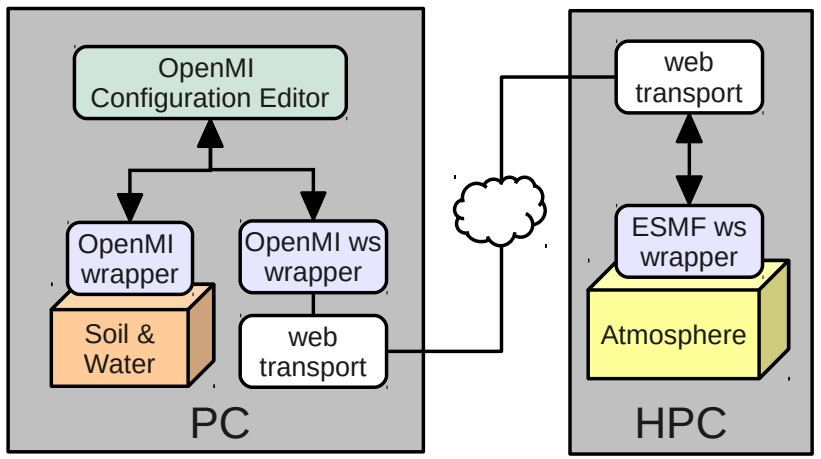


Multiple Modes of "Coupling"



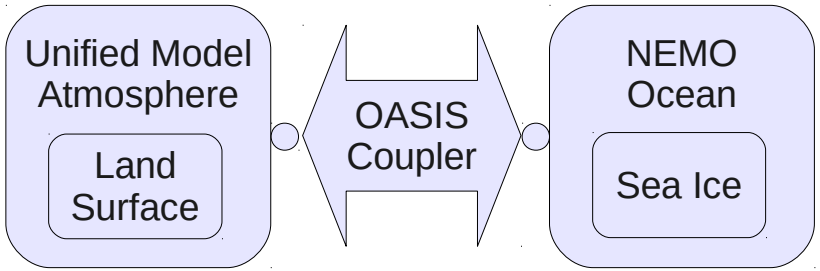


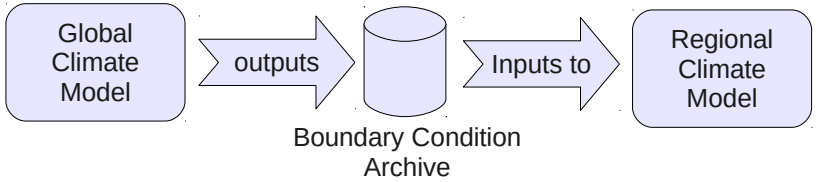


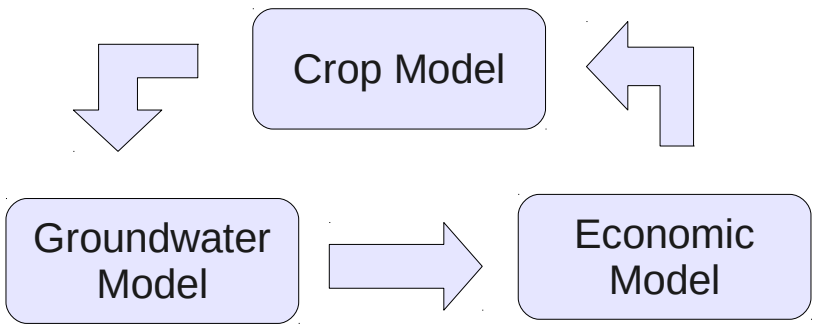


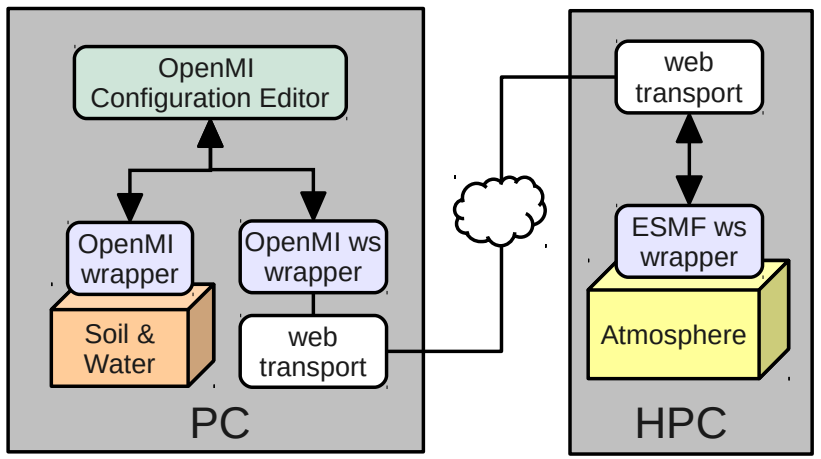
- Not all components should be two-way coupled.

Multiple Modes of "Coupling"









- Not all components should be two-way coupled.
- *Not all coupling is an exchange of fields at a boundary (consider a fast physics solver with input fields from multiple components).*

Requirements

Ensuring interfaces will

- be flexible
- not impact on numerics of any model component
  - (hmm, see last point of previous slide)
- catalyse the development of more comprehensive models of the terrestrial water cycle,
- future-proof Hydro-JULES against forthcoming changes to the UM dynamical core
  - (hmm, radical change coming …)

Some hard asks! Much engagement needed with community to get priorities and requirements, hence:

- *Task 1.2: Consult research community and stakeholders on requirements and implementation*.

Community Based

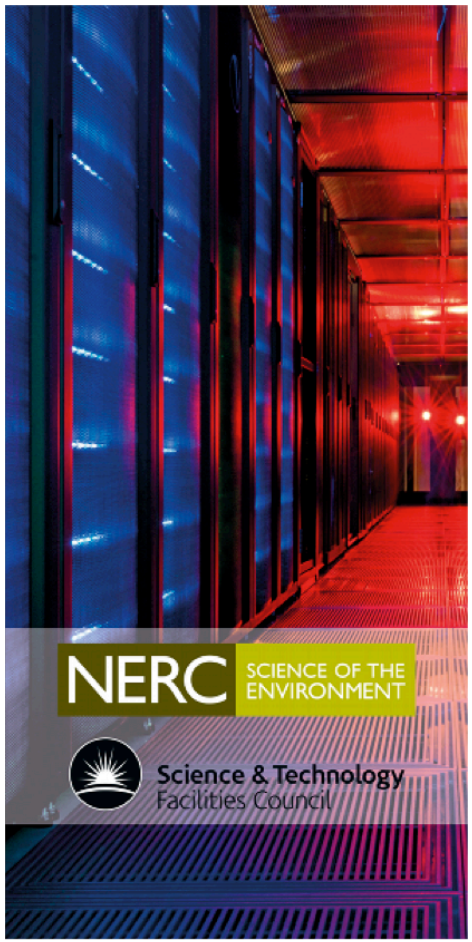
Task 1.3: Build archive of driving data, model configurations and supporting datasets

- **Open** repository of quality controlled driving data, model configurations, and suporting datasets
- Reproducibility! Open-Access Publication! UK participation in international activities!

Task 1.4: Provide user training support and managed access to JASMIN

- Formal Training Courses.
- Software support for installation and use.
- Application support.
- Online training materials
- UI to run HydroJULES model configurations on JASMIN

JASMIN



LOTUS

Optimised High Performance Data Analysis Environment

Community Cloud

Customisable

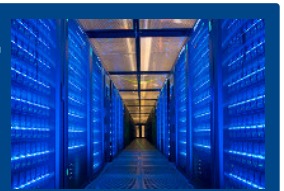
(with high performance route to archive)

CEDA Data Services

Remote access to archive & catalogues. Download etc

CEDA Archives

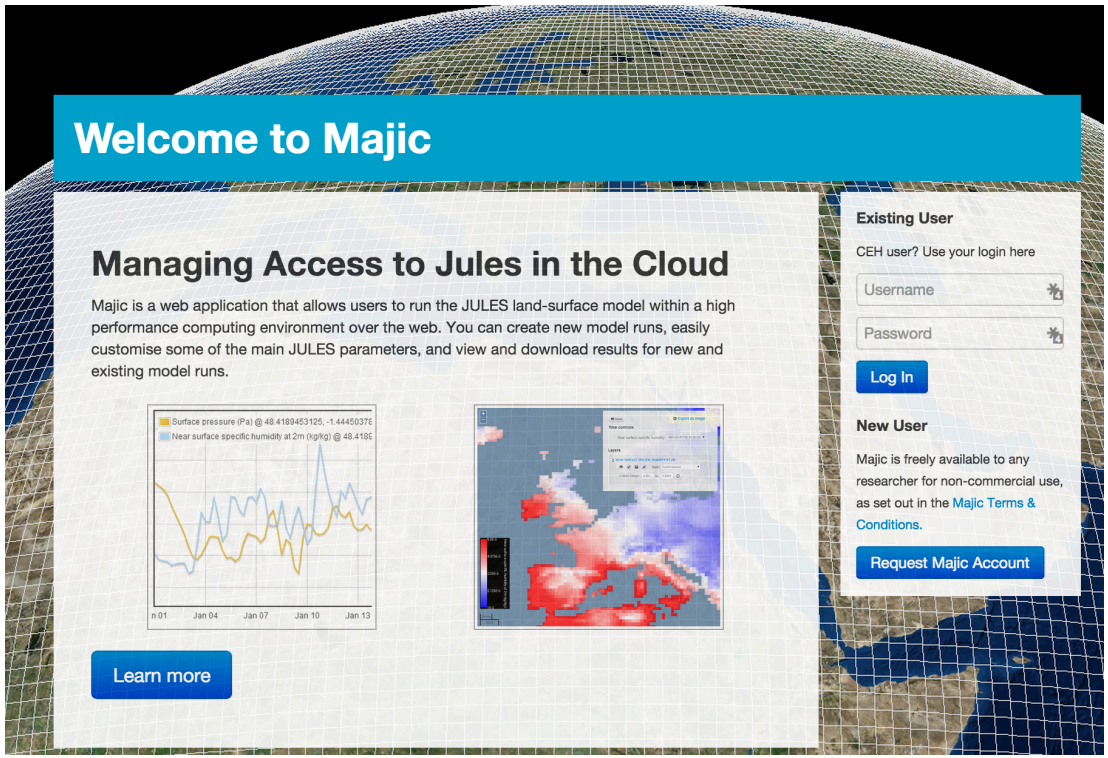
JASMIN – Data Intensive Computer
Storage, Compute and Network Fabric
Batch Compute, Private Cloud, Disk, Tape



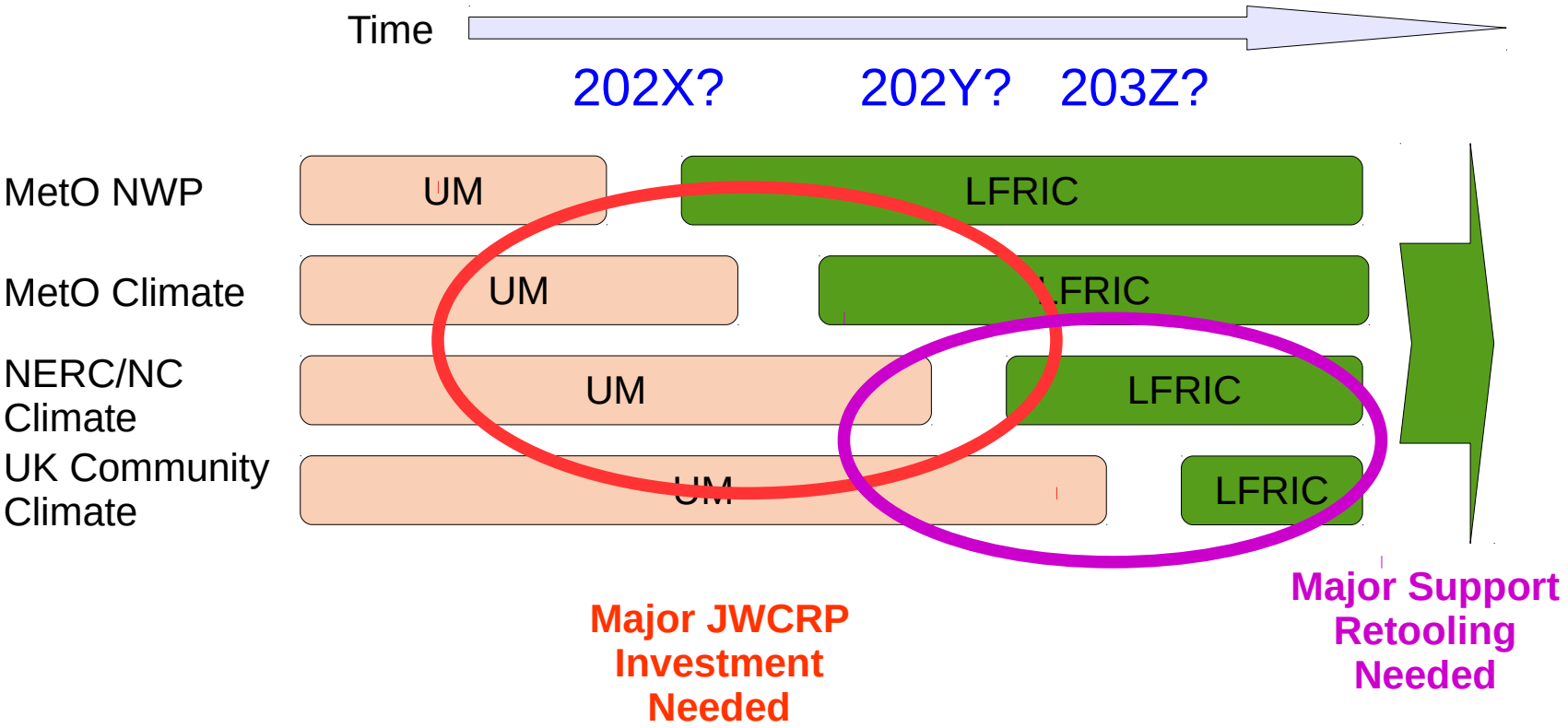
Group Workspace for HydroJULES - dedicated storage for HydroJULES community!

Extending MAJIC

- Advanced users log in to the JASMIN managed cloud, exploit the LOTUS batch cluster, and/or potentially *HydroJULES specific machines* (in the JASMIN external cloud),
- Many users will be able to exploit a (new/updated) web interface to run models and generate data based on:



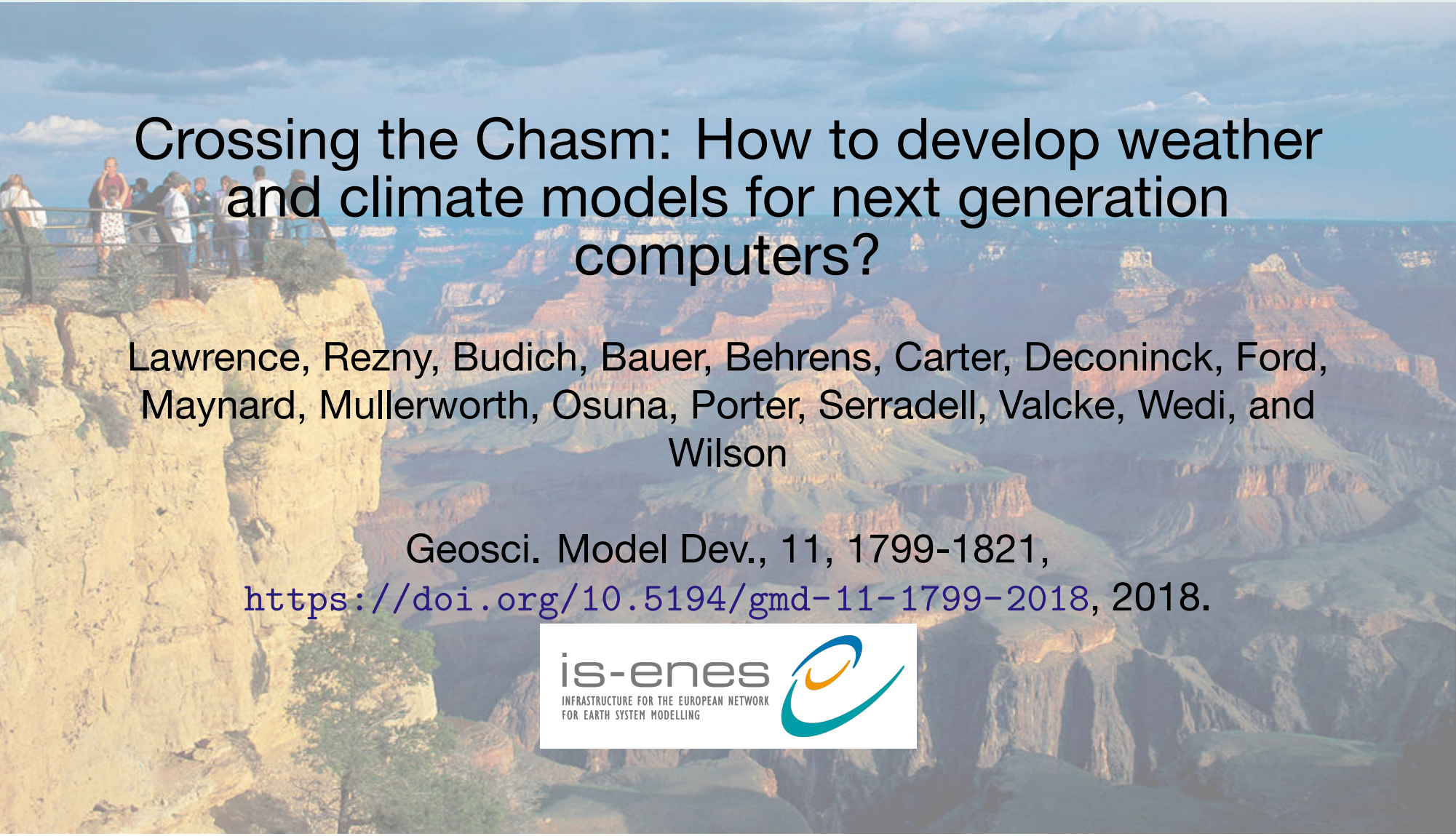
The exascale challenge and LFRIC



The NERC Community has not clocked the scale of this problem!

It is going to be non-trivial to bring a HydroJULES/JULES model interface through this transition!

Crossing the Chasm



Crossing the Chasm: How to develop weather and climate models for next generation computers?

Lawrence, Rezny, Budich, Bauer, Behrens, Carter, Deconinck, Ford, Maynard, Mullerworth, Osuna, Porter, Serradell, Valcke, Wedi, and Wilson

Geosci. Model Dev., 11, 1799-1821, https://doi.org/10.5194/gmd-11-1799-2018, 2018.

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

Software changing slowly & slowing!

Hardware changing rapidly & accelerating!

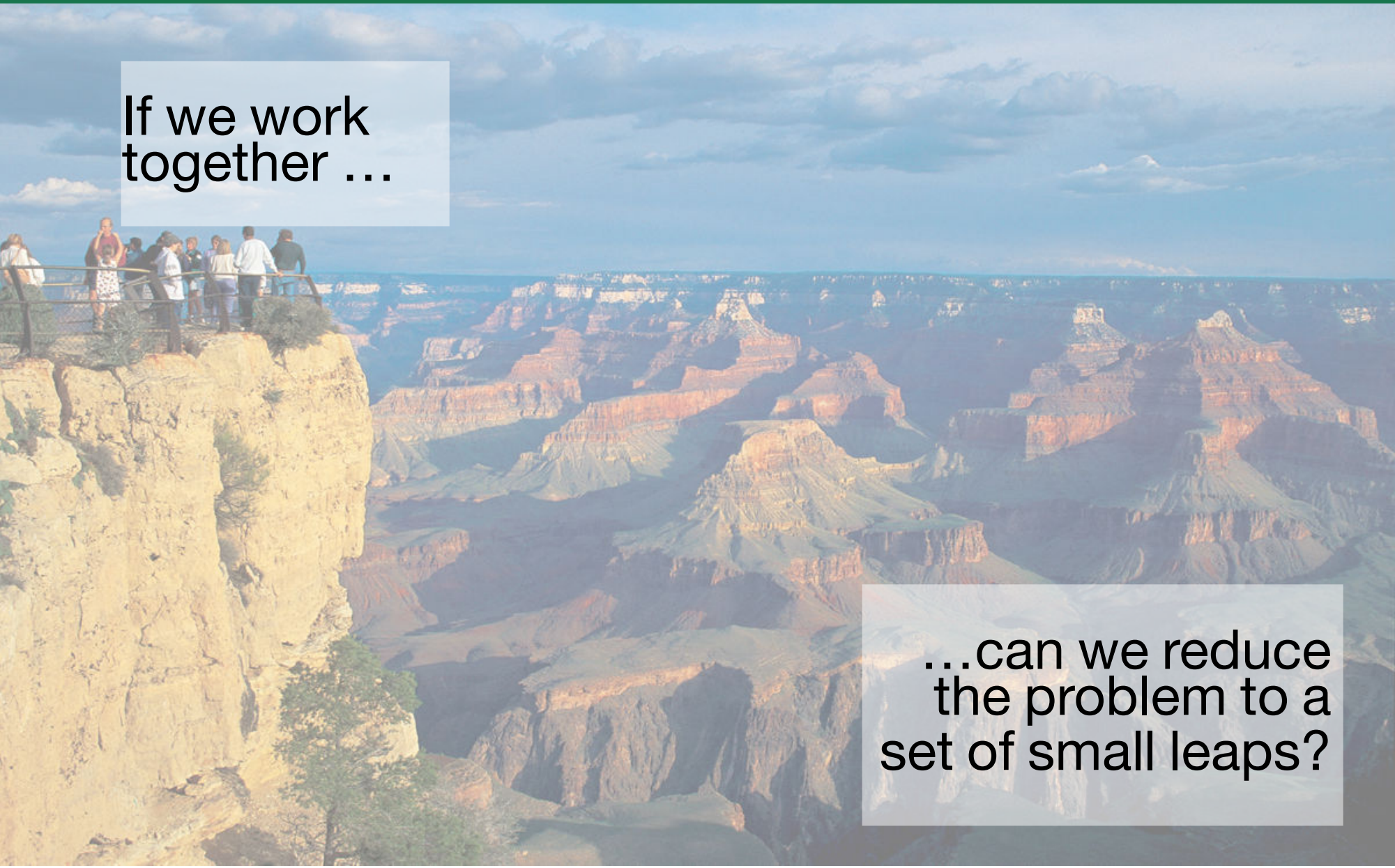
How far is it between our scientific aspiration and our ability to develop and/or rapidly adapt our codes to the available hardware?

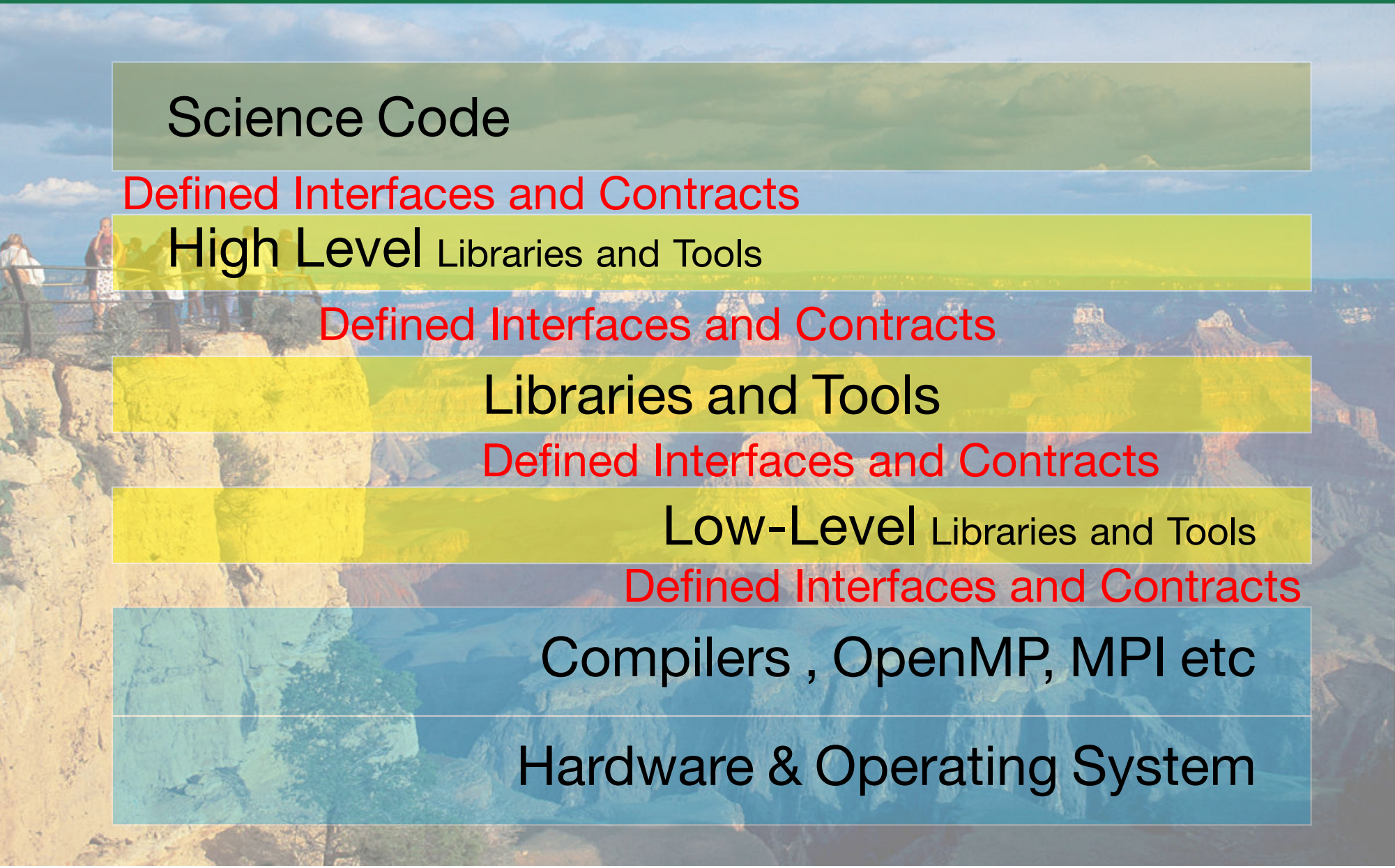
Science Code

How do we bridge the gap?

Compilers , OpenMP, MPI etc

Hardware & Operating System


If we work together …
…can we reduce the problem to a set of small leaps?


Science Code
Defined Interfaces and Contracts
High Level Libraries and Tools
Defined Interfaces and Contracts
Libraries and Tools
Defined Interfaces and Contracts
Low-Level Libraries and Tools
Defined Interfaces and Contracts
Compilers , OpenMP, MPI etc
Hardware & Operating System

What project characteristics are necessary?

They will:

- be open source *and* have an open development process,
- have clear goals, scope, and where appropriate, deliver stable software interfaces,
- have a mechanism to understand and respond to the timescales of collaborators (that is, some sort of governance mechanism which assimilates and responds to requirements),
- potentially be able to accumulate and spend funds to provide user-support, training, and documentation,
- be not initially disruptive of existing solutions, and ideally
- engage both the scientific community and vendors (compare with MPI where vendor implementations are often key to enhanced MPI performance).

HydroJULES objectives in good shape for this …

What institutional characteristics are necessary?

They will most probably:

- Have understood the issue fully at the management level, the science level, and in the infrastructure teams,
- Be able to reward individuals for innovation in, and/or contributions to, external projects,
- Recognise the benefit of external scrutiny and contributions into their own projects,
- Have the courage to stop existing activities and pickup and use/integrate third party libraries and tools, and
- Have the ability to recognise the cost-benefit trade-off between "doing it themselves" and contributing intellectually and financially to third party solutions, and
- Be ready to apply more sophisticated and complex software engineering techniques, and encourage more computational science research.

Summary

Ambitious programme to:

- Design and implement a HydroJULEs modelling framework and set of interfaces,
- Support the community in developing against, and exploiting the resulting models;

…in the presence of:

- a difficulty continuing simultaneously with portability, productivity, and performance arising from:
  - radical changes in the external model software environment, and
  - a widening chasm between scientific aspiration and the hardware environment, requiring new behaviours (which are understood by the project, but maybe not yet by the institutions).

…which wouldn't be worth doing if it wasn't hard!

Deliverables

- D1.1: Community consultation reports gathering user requirements and giving use cases for the Hydro-JULES system.
- D1.2: Design phase report for interface framework in Hydro-JULES code base.
- D1.3: Development of prototype modelling system with version control and links to JULES.
- D1.4: Paper describing interface framework and coupling structure.
- D1.5: Open access repository of shared driving data.
- D1.6: Open access database of quality-controlled supporting datasets and facility to store model configurations (JASMIN).
- D1.7: Test stage review report.
- D1.8: Verification stage report including consultation with stakeholders.
- D1.9: Papers evaluating model performance against global and national datasets; benchmarking studies.