National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

University of Reading

ExCALIData: Exascale I/O & Storage and Workflow

[Bryan Lawrence + many]

https://excalibur.ac.uk/projects/excalidata/

ddn

StackHPC Ltd

UNIVERSITY OF CAMBRIDGE
Research Computing Services

ExCALIBUR

Overview

Exascale Computing ALgorithms & Infrastructures Benefiting UK Research (ExCALIBUR)

ExCALIBUR is a UK research programme that aims to deliver the next generation of high-performance simulation software for the highest-priority fields in UK research. It started in October 2019 and will run through until March 2025, redesigning high priority computer codes and algorithms to meet the demands of both advancing technology and UK research.

ExCALIBUR is built around four pillars – **separation of concerns**, **co-design**, **data science**, **investing in people**

https://excalibur.ac.uk/

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

ExCALIBUR Themes

- High Priority use cases
  - Weather and Climate, Fusion
- Emerging Requirements for High Performance Algorithms
  - Social Sciences, Humanities, Biomedicine…
- Hardware and Enabling Software
  - Testbeds, early access to novel hardware and software
- Cross-cutting Research
- RSE Knowledge Integration
  - Grow interdisciplinary RSE community, fill skills gap, training…

ExCALIBUR Cross-cutting Research

- I/O & storage
- Data workflow
- Coupling
- Domain Specific Languages

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

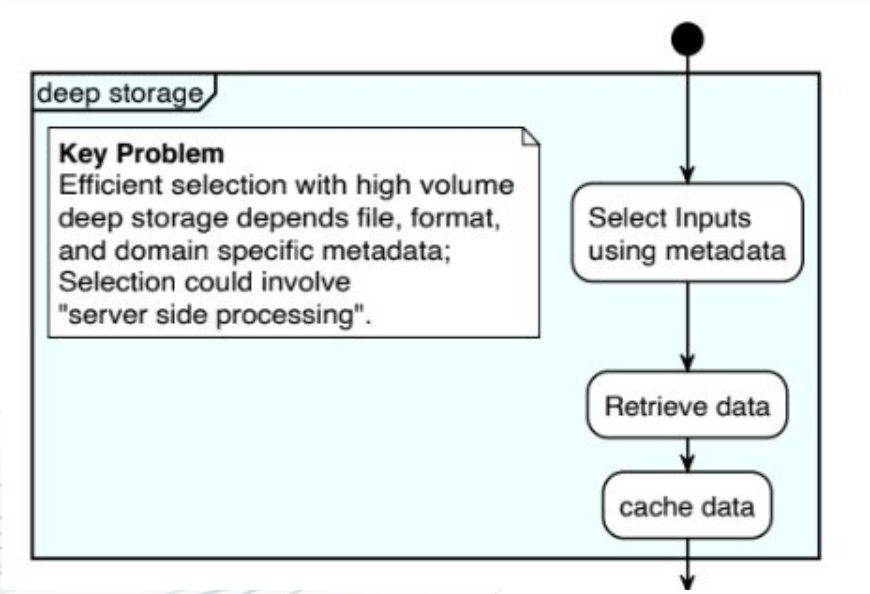
ExCALIBUR Cross-cutting Research

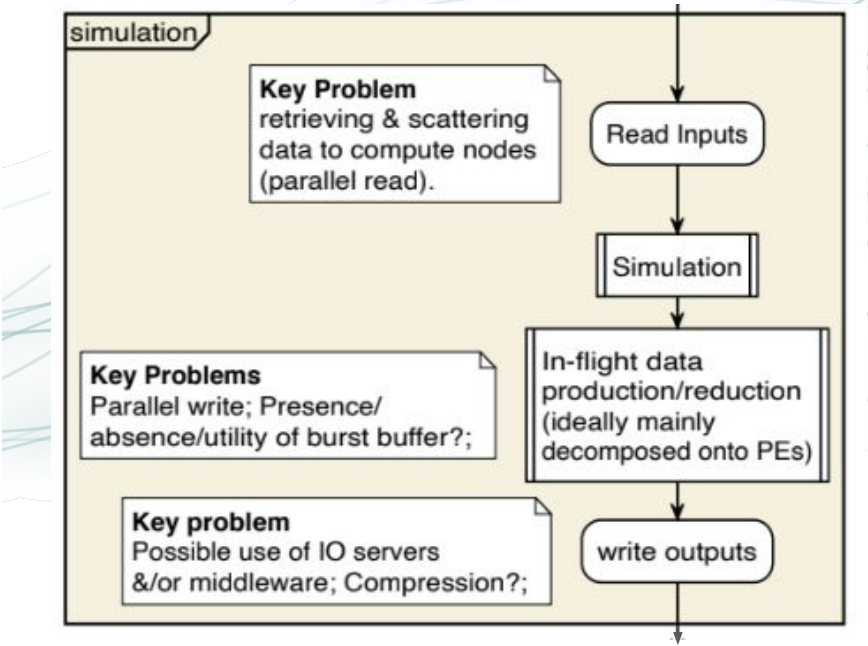
- I/O & storage
  - ExCALIStore
    - Reading, Cambridge
- Data workflow
  - ExCALIWork
    - Reading, DDN, StackHPC
- Coupling
- Domain Specific Languages

ExCALIData

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

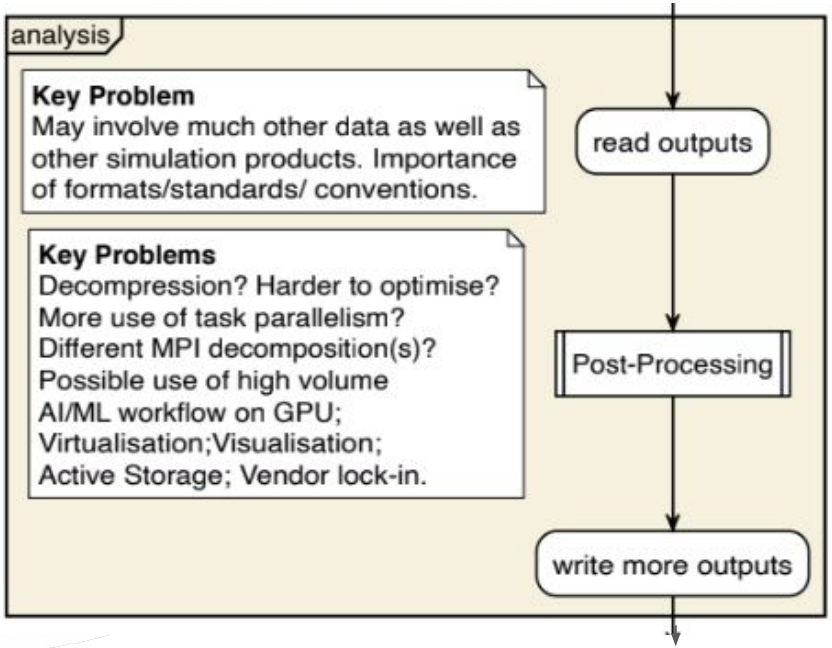
ExCALIData

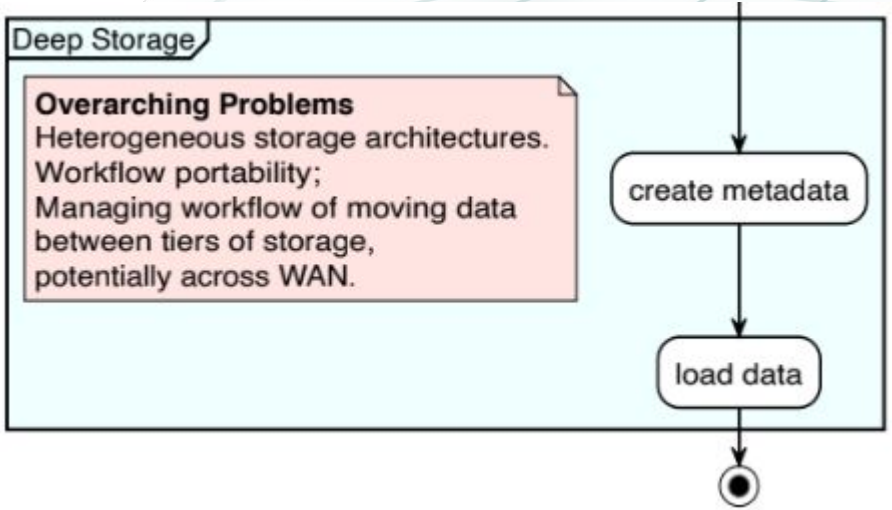




National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

ExCALIData





National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

ExCALIStore

Development and demonstration of novel approaches to optimising aspects of the data flow to improve I/O for large-scale applications.

- Storage interfaces
- Data management across storage tiers and between institutions
- Accelerating IO
  - Network fabric
    - Remote Direct Memory Addressing (RDMA) and/or SmartNICs
  - Advanced burst buffers
  - IO middleware
    - ADIOS
    - ESDM
- Synthetic tests and application to Weather & Climate and Fusion cases

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

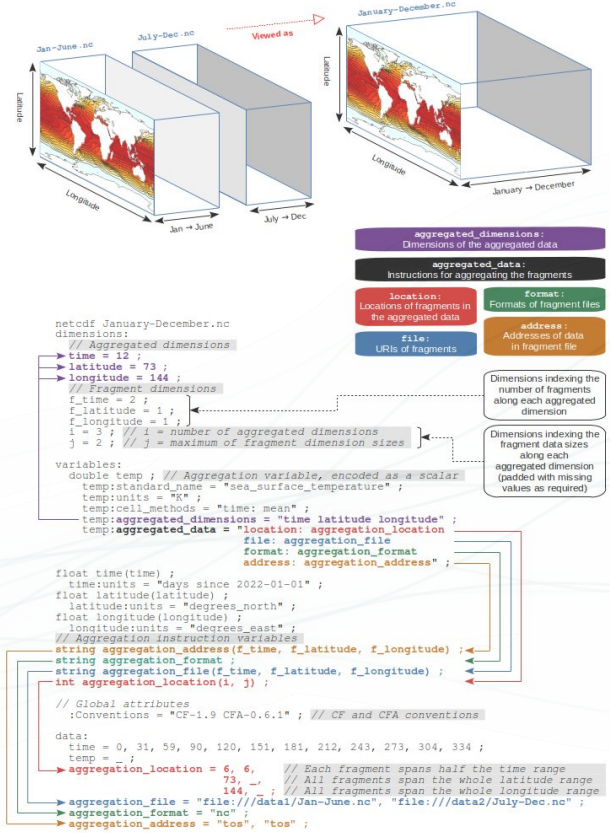
ExCALIStore

extract a particular spatio-temporal variable from a set of files

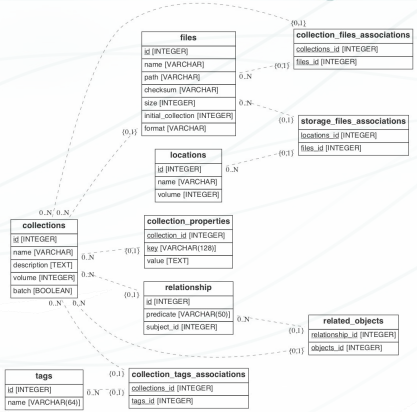
cf aggregation - underlying data remains unaltered, the aggregation presents the user with a single file

keep track of files held in multiple different storage elements

cf store - a method of cataloguing the list of atomic datasets available to user and tools for manipulating the atomic datasets (such as, listing all atomic datasets in a particular storage element, identifying duplicates, and moving subsets between storage elements).



cf aggregation



cf store

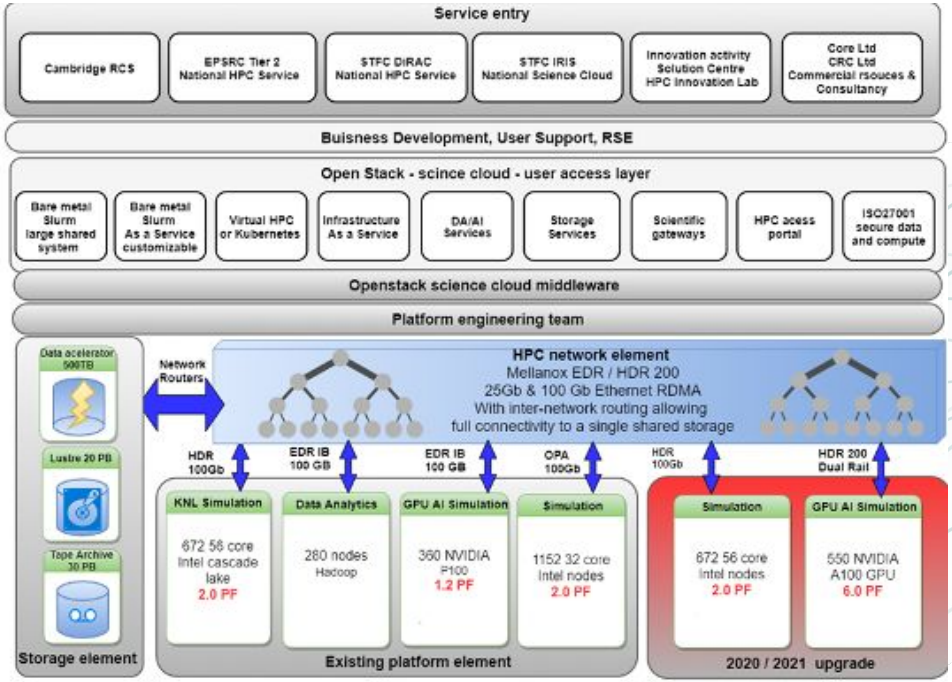
https://ncas-cms.github.io/cf-python
https://ncas-cms.github.io/cf-python/aggregation_rules.html
https://github.com/NCAS-CMS/cfstore

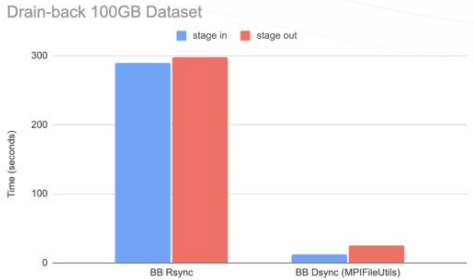
David Hassell, Sadie Bartholomew, George O'Brien

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

ExCALIStore

- CRCS testbed
- Cambridge Data Accelerator (DAC)
  - Burst Buffer
    - Improved NVMe file system life cycle
      - deploying a per job ephemeral directories
    - mpifileutils dsync as opposed to simple rsync for improved stage in and stage out performance
    - SLURM Burst buffer Lua functionality
  - Build a DAOS prototype system, integrate into the burst buffer framework





Paul Calleja, Wojciech Turek, Chris Edsall

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

ExCALIWork

Development and demonstration of novel approaches to taking certain computations to the data to reduce the need for data movement to improve I/O in large-scale applications

- Support for moving computation away from the traditional analysis phase
  - Increase concurrency/Reduce data movement
  - Move data reductions into the storage layer
    - Active Storage
      - User facing
      - Storage compute
  - Move ensemble manipulations into the simulation phase
    - Leverage XIOS capability
      - Current model (UM)
      - LFRic
- Synthetic tests and application to Weather & Climate and Fusion cases

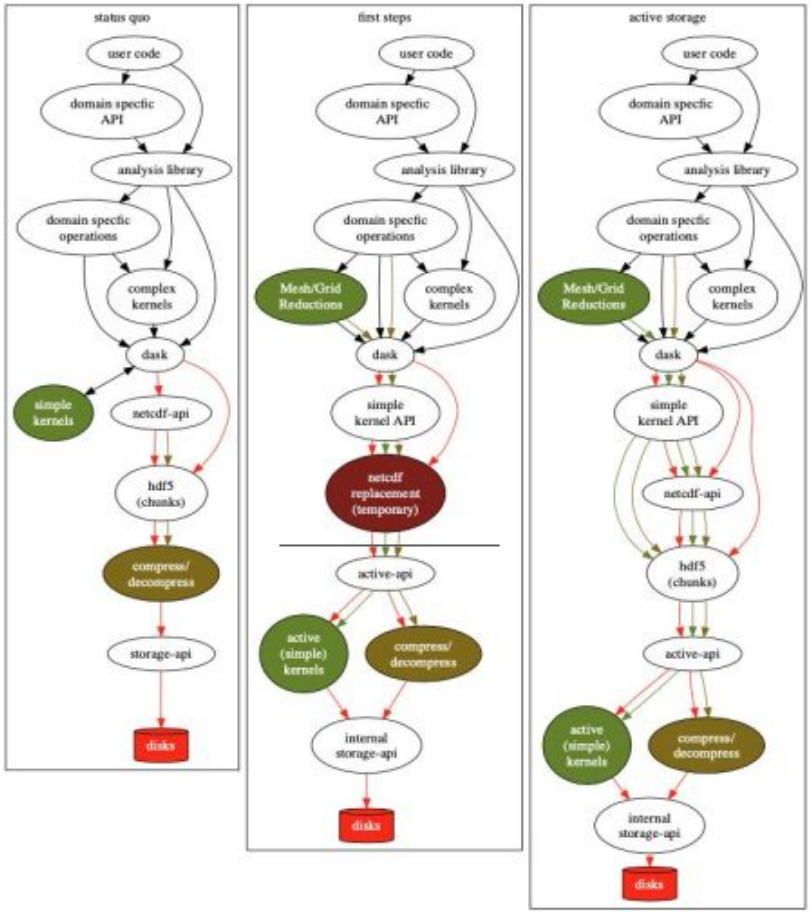
National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

ExCALIWork

Active Storage -

a computer system architecture which utilizes processing power in disk drives to execute application code.

- Interface design
- Simple reduction(s)
- Storage compute
- Client compute (DASK)



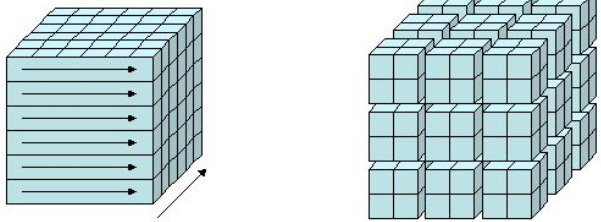
David Hassell, Valeriu Predoi, Stig Telfer, J-T Acquaviva

https://github.com/NCAS-CMS/ActiveStorage

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

Active Storage Servers

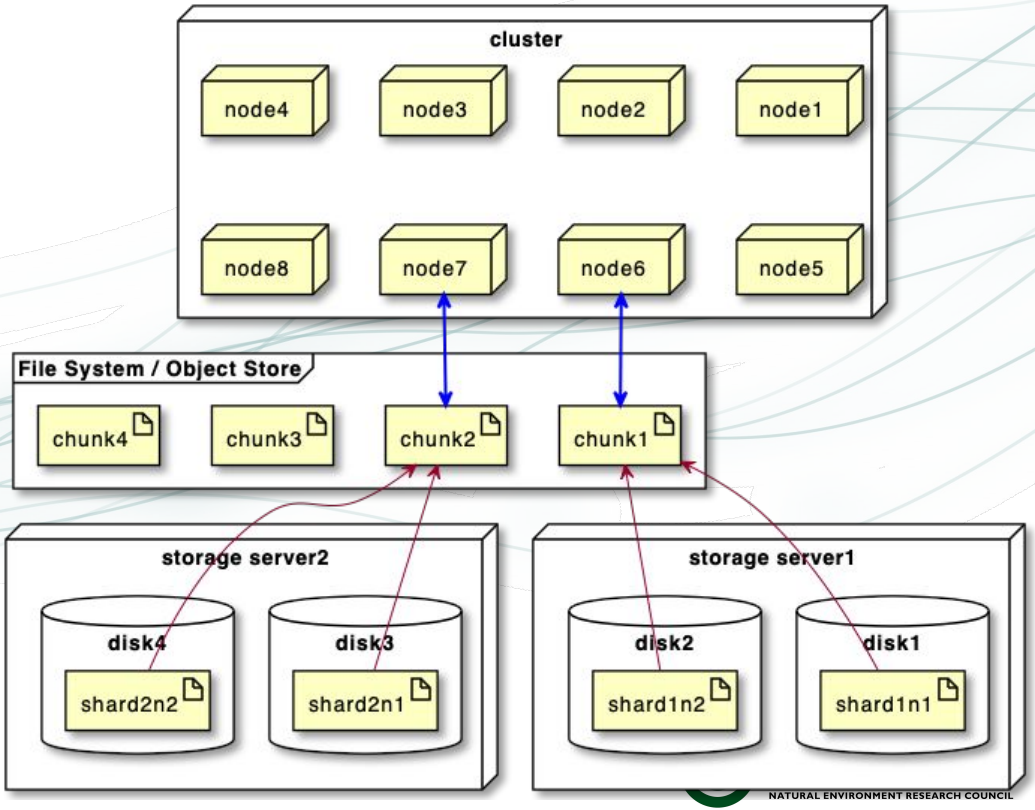
The API talks to “chunks” and needs to work on those



Two implementations:

S3 (with StackHPC)

IME/RED (POSIX) (with DDN)

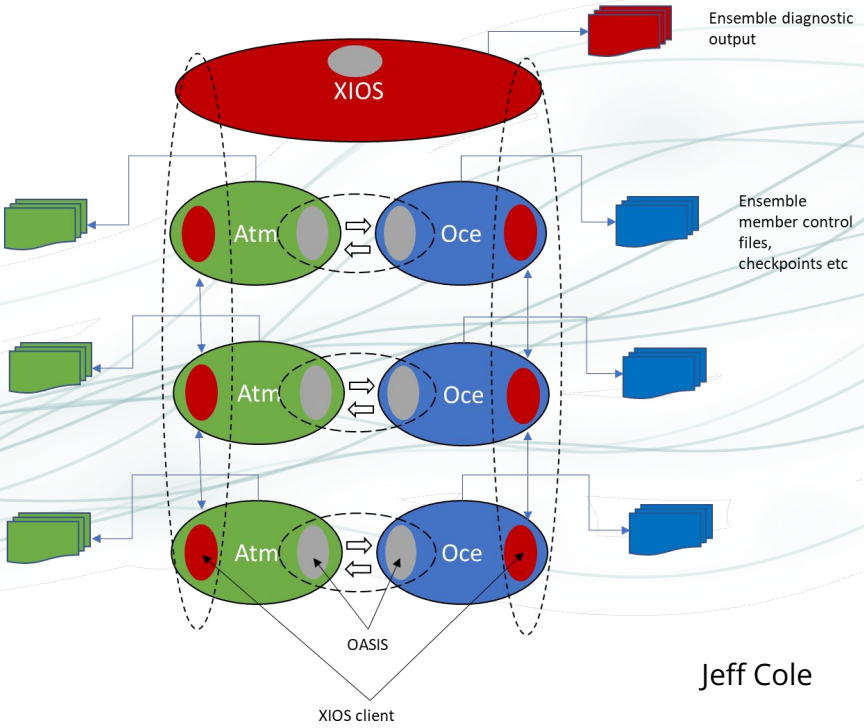


ExCALIWork

Coupled UM Ensemble -

Exploit XIOS capability to deliver *in-flight* ensemble diagnostics

```
<grid id="um-atmos_grid_uv_pfl35225">
    <domain domain_ref="um-atmos_grid_uv" />
    <axis axis_ref="um-atmos_pfl35225" />
    <axis axis_ref="ensemble" />
  </grid>
  <grid id="um-atmos_grid_uv_pfl35225_ensmean">
    <domain domain_ref="um-atmos_grid_uv" />
    <axis axis_ref="um-atmos_pfl35225" />
    <scalar id="um-atmos_grid_uv_pfl35225_ensmean">
      <reduce_axis operation="average" />
    </scalar>
  </grid>
```



Jeff Cole

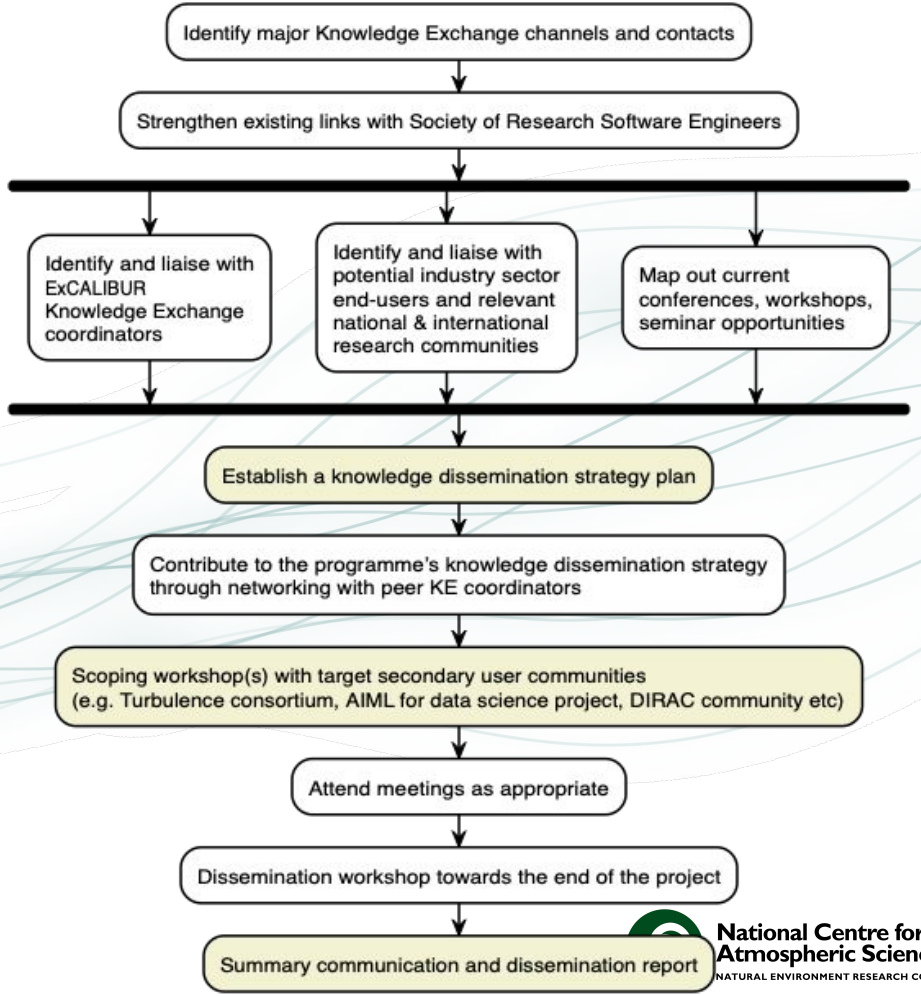
Extend to LFRic - incorporate XIOS ensemble developments

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

ExCALIData

Knowledge Exchange

- Position the next generation of software engineers at the cutting edge of scientific computing
- Ensure integration across the programme activities
- Establish connections with potential beneficiaries in academia, Public Sector Research Establishments and industry



National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

ExCALIData

Watch this space - thanks

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL