From ESiWACE to EXCALIDATA

Bryan Lawrence

NCAS &
University of Reading: Departments of Meteorology and Computer Science

Excalibur Workshop, 2 Sepember 2021

Outline

Motivation

- Modelling Context

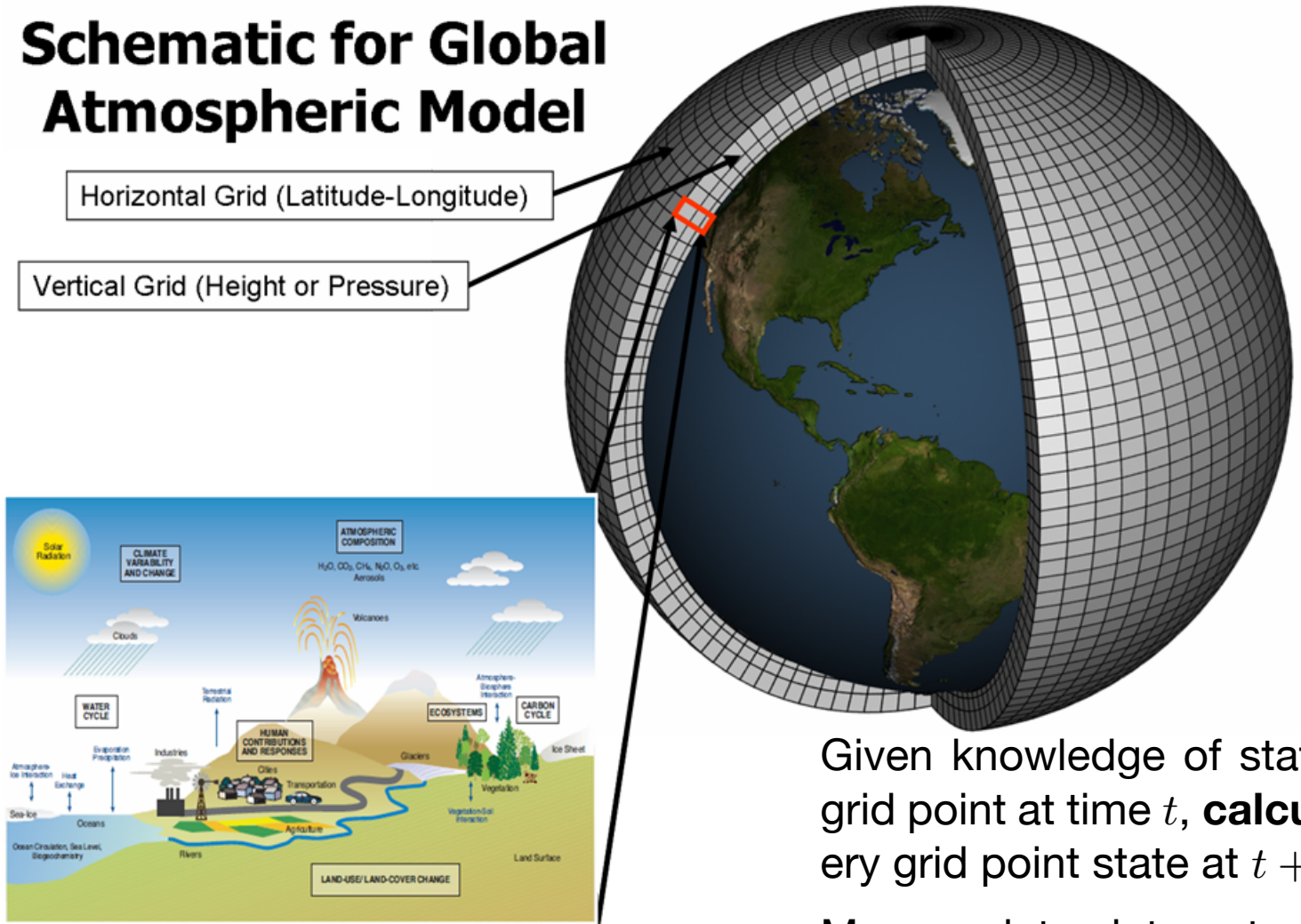
EsiWACE

- Quick history

ExCALIDATA

- Intro to work packages
- Segue to more detailed presentation

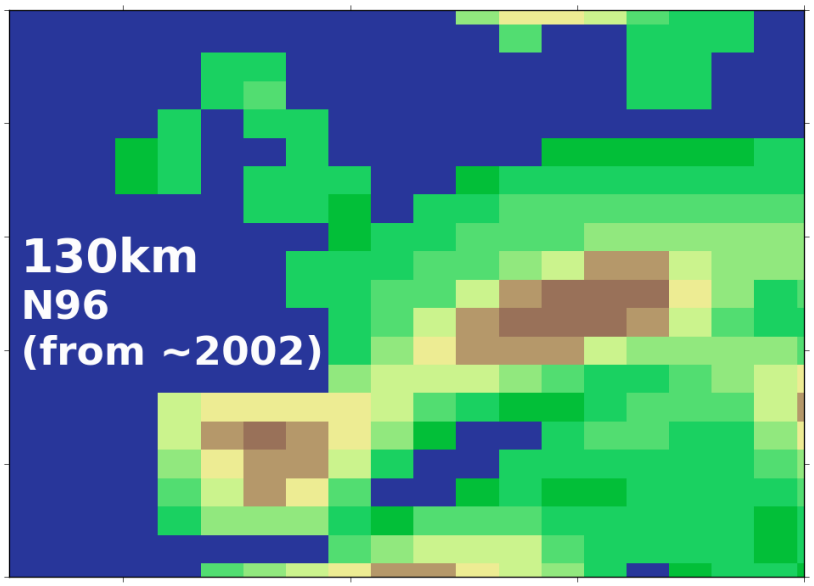
Everything is solved on a grid

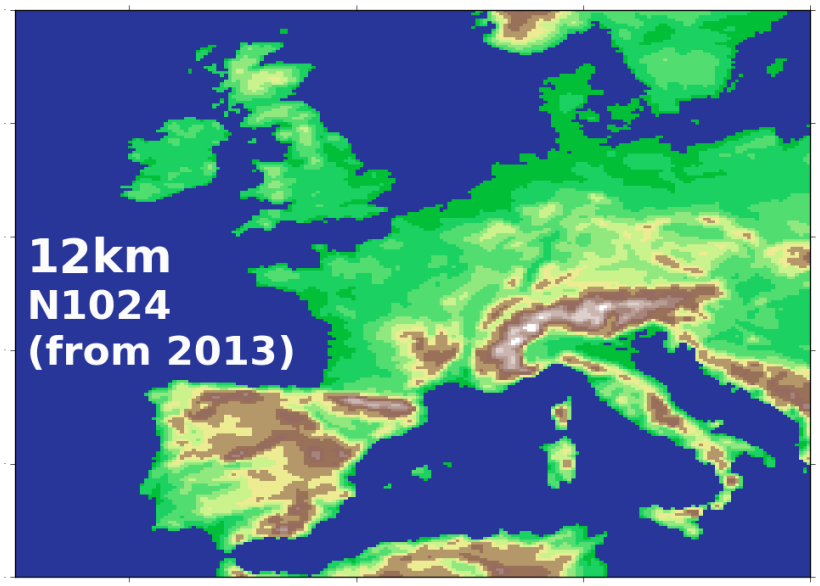


Given knowledge of state at every grid point at time t, **calculate** at every grid point state at $t + \Delta t$.

Many points, integrated for years with timestep of *o(minutes)*!

A modest (?) step ...





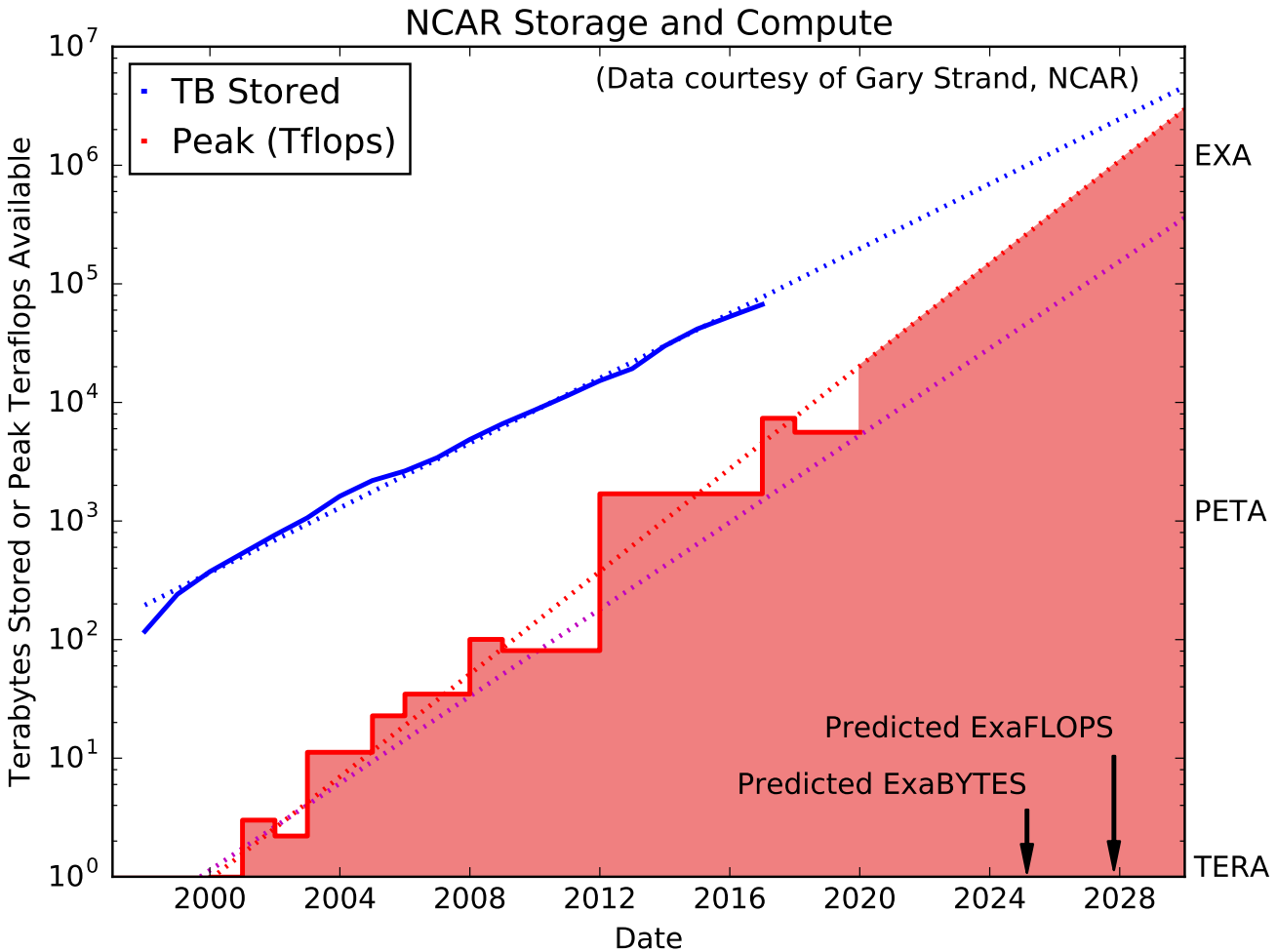
One "field-year" — 26 GB

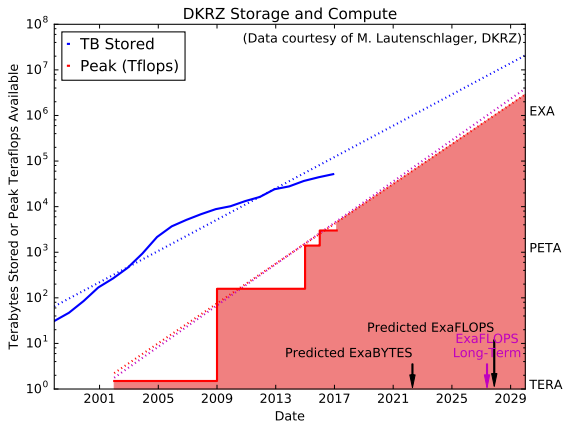
1 field, 1 year, 6 hourly, 80 levels
1 x 1440 x 80 x 148 x 192

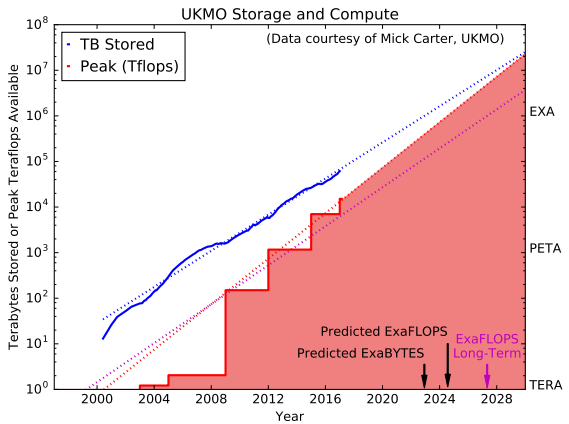
One "field-year" — >6 TB

1 field, 1 year, 6 hourly, 180 levels
1 x 1440 x 180 x 1536 x 2048

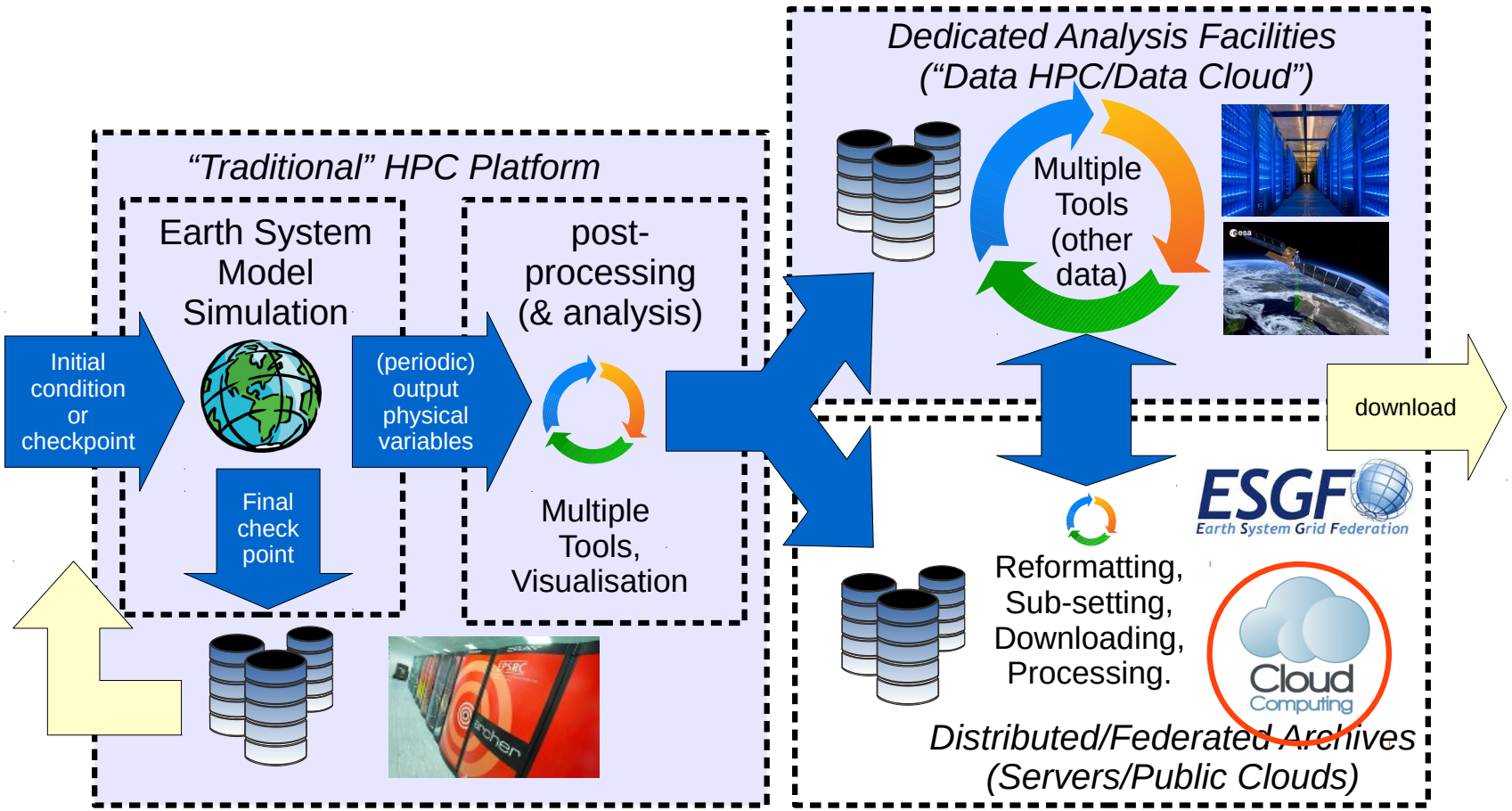
History has given us exponential compute linked to exponential data …



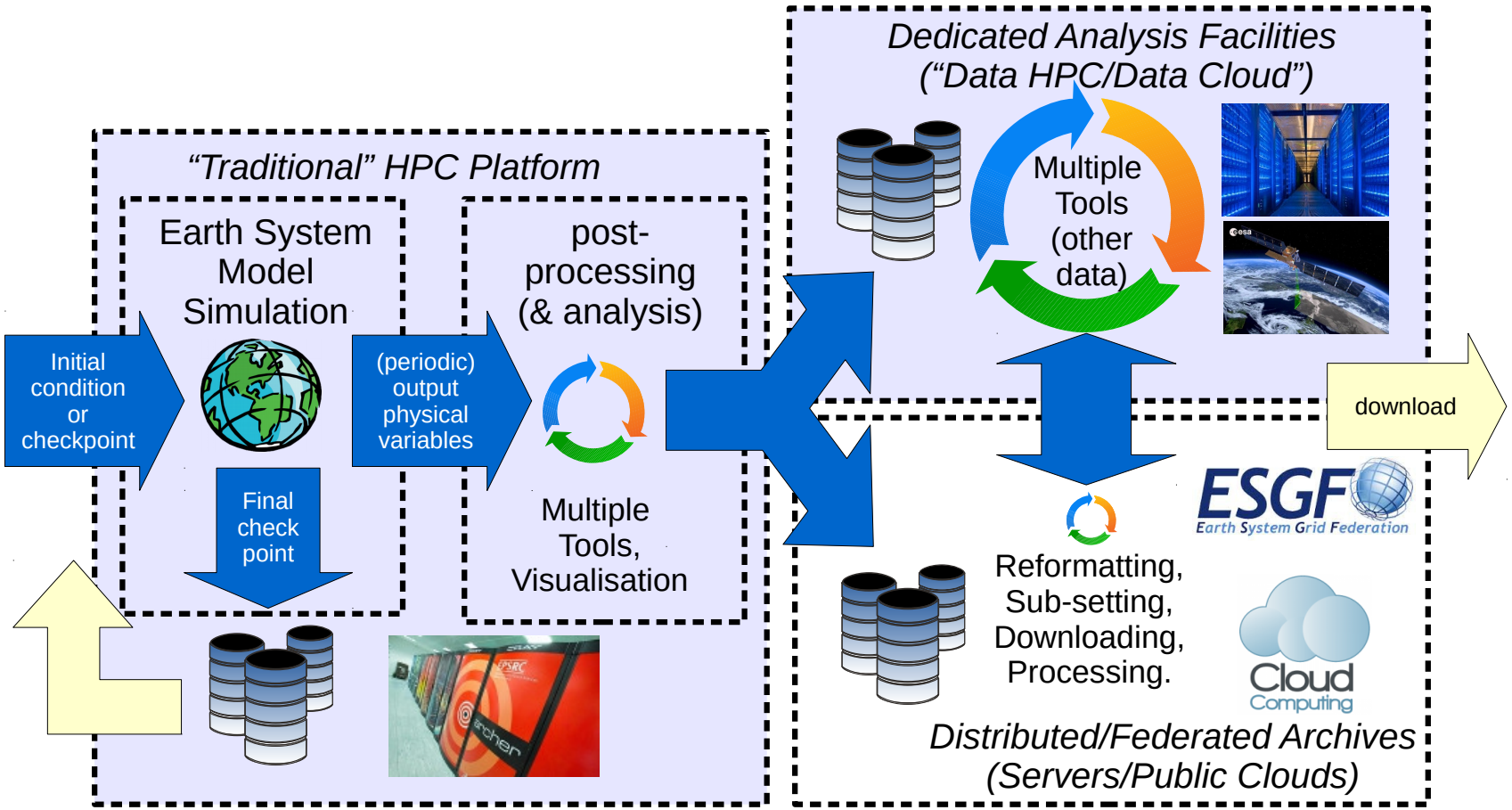




Many different supercomputing environments



Many different supercomputing environments



Multiple Roles, at least:

Model Developer, Model Tinkerer, Runner, Expert Data Analyst, Service Provider, Data Manager, Data User

ESiWACE

The Centre of Excellence in Simulation of Weather and Climate in Europe, enables global storm- and eddy resolving weather and climate simulations on the upcoming (pre-)Exascale supercomputers.

WP4: Data Systems at Scale

Lawrence, Kunkel, et. al.

WP4 Partners:

CNRS-IPSL, CMCC, DDN, DKRZ, METO, Seagate, STFC, UREAD

WP4: Data Systems at Scale

Objectives

to mitigate the effects of the data deluge from high-resolution simulations (project objective-d)

by

1. Supporting data reduction in ensembles by providing tools to carry out ensemble statistics “in-flight” and compress ensemble members on the way to storage, and
2. Providing tools to:
   - 2.1 transparently hide complexity of multiple-storage tiers (middleware between NetCDF and storage) with industrial prototype backends, and
   - 2.2 deliver portable workflow support for manual migration of semantically important content between storage on disk, tape, and object stores.

ensemble tools, storage middleware, storage workflow

WP4 Philosophy and Methodology

Maximum Impact from a Minimum Change Surface

Solutions (tools addressing the data deluge), need to maximise their impact on data handling by

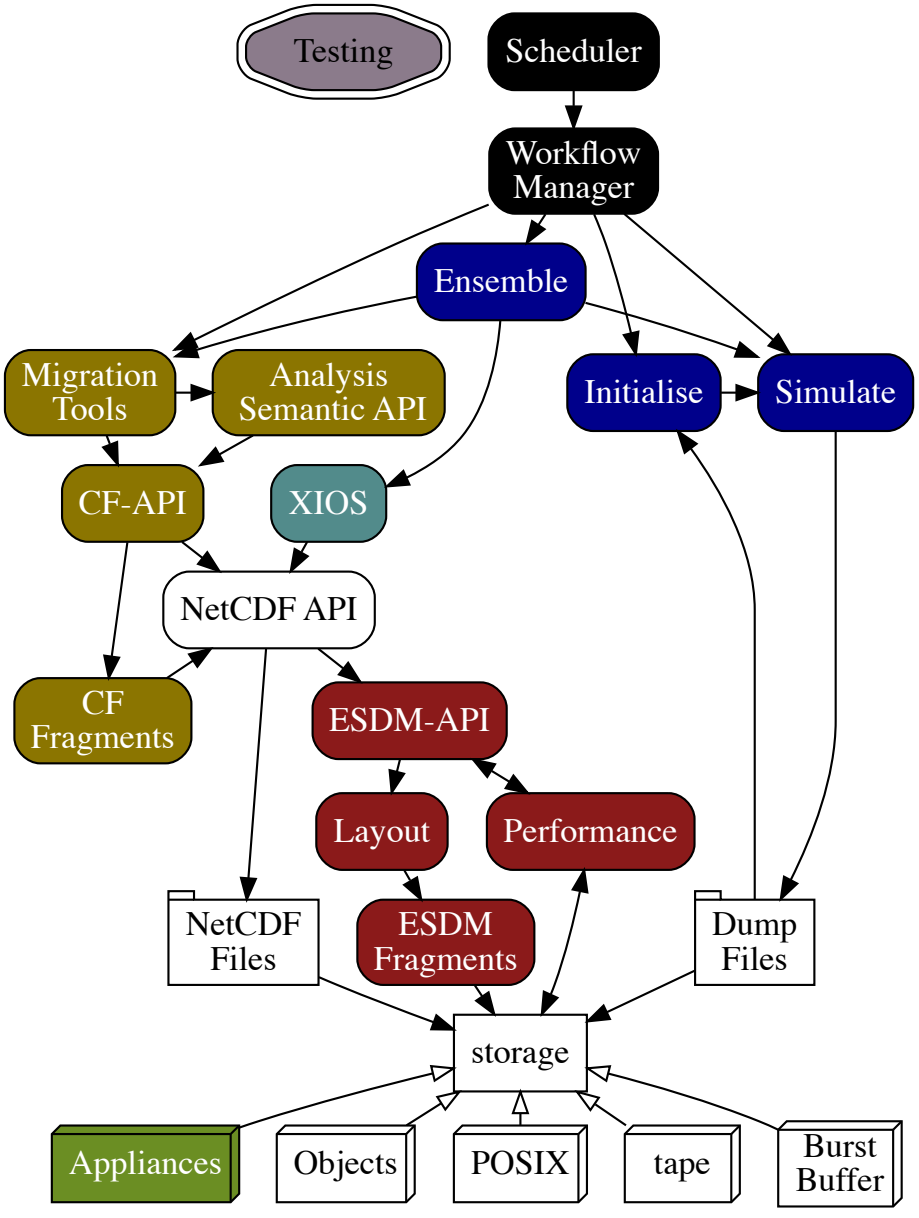
- Minimising the impact of increasing volumes of data, particularly within large-scale ensembles and/or high resolution runs, while
- minimising interference with existing working practice and codes, and
- minimising requirements of the system environment.

Methodology

- Modify existing tools,
- Develop a minimum of new tools and,
- (where possible) Exploit middleware which can be deployed in userspace, but hide complexity from end-users
- (but at the same time, where we do have access to the system) If appropriate, deploy new services and appliances.

Components/Tasks in WP4

- 4.1 Leadership and Design: 12 PM
- 4.2 Ensemble Services (in flight analysis/compression)
- 4.3 Earth System Data Middleware (ESDM) - performance in HPC simulation.
- 4.4 Semantic Storage Tools (SemST) - userspace tools for handing volume.
- 4.5 Workflow Support (enhancements to SLURM/cylc)
- 4.6 Component and End-to-End Testing
- 4.7 Industrial Proof-of-Concept Appliances.



ExCALIData

ExCALIData

EX20-6 ExCALIStore (IO & Storage)

WP-I1: Storage Interfaces

Goal: Users know where their data is, can move it between sites & storage tiers and subset using semantic information exploiting tiered storage locally and across WAN.

WP-I2: Fabric and Solid State Storage

Goals: (1) Deploy and assess RDMA I/O & data analytics acceleration, including emerging SmartNic strategies.
(2) Investigate solid state I/O accel'n strategies; DAOS, Burst Buffers, DAC, Dynamic file systems, GekkoFS, UnifyFS.

WP-I3: I/O middleware

Goal: Test and understand strengths and weaknesses of generic and specific I/O middleware, in particular
(1) ADIOS and (2) ESDM.

EX20-7 ExCALIWork (Workflow)

WP-W1 Active Storage Software

Goal: Develop a full-stack approach to active storage with user tooling which exploits domain knowledge and domain agnostic middleware and new (WP-W3) active storage services.

WP-W2: Active Storage Servers

Goal: Develop, deploy and test systems which exploit the proposed active storage interface to carry out specific operations behind an API suitable for deployment in both (1) cloud and (2) conventional HPC.

WP-W3: I/O Server Prototypes

Goal: Develop an "in-flight" analysis system with interacting MPI communicators and ensemble reductions across model components.

WP4: Community Building & Proj. Man.

(1) Establish a UK community of RSEs and application specialists with interest, aptitude, and experience in managing exascale data flows and exascale data residence.
(2) Coordinate project and formal reporting to Met Office.

Both projects include a WP4 with complementary activities. If both are funded twice as much effort will be deployed as a consolidated work package.

and so to next presentation