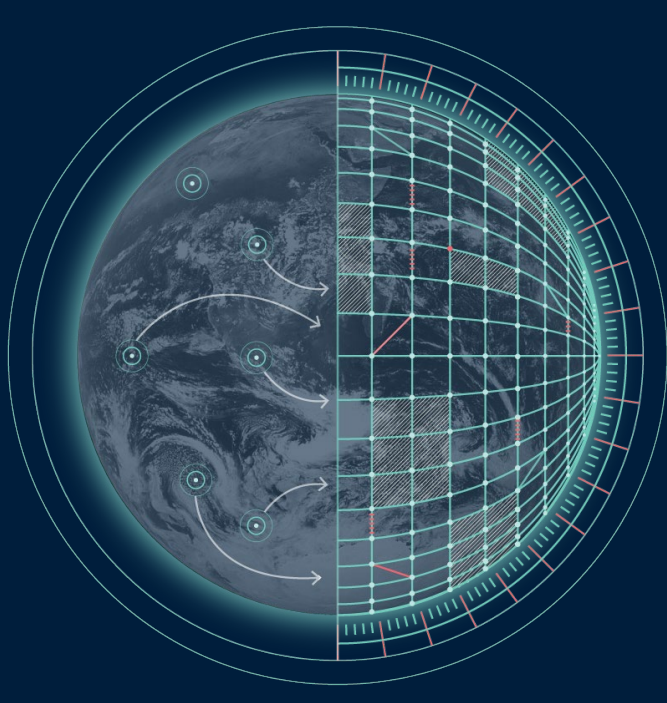


Production Workflow: Continuous Extremes Digital Twin



Emma Kuwertz*, Tryggvi Hjorvar, Bentorey Hernandez Cruz, Johannes Bulin, Paul Burton, Andrew Bennett, Michael Sleight

ECMWF; (*) emma.kuwertz@ecmwf.int

1. The Extremes Digital Twin

The Continuous Weather-Induced & Geophysical Extremes Digital Twin (DT) being delivered by ECMWF as a part of the Destination Earth initiative:

- Provides real-time medium-range weather forecasts at high resolution
- Makes use of EuroHPC computing resources
- Is modelled via ECMWF's Integrated Forecasting System (IFS)
- Is initialised from operational data produced at ECMWF's HPC facility

The DT relies on distributed resources for configuration, initialisation, monitoring, computation and data storage.

ECMWF	EuroHPC	Data Bridge
<ul style="list-style-type: none"> • Model Inputs • EcFlow suite definition and configuration • EcFlow suite deployment 	<ul style="list-style-type: none"> • Compute resource • Short term data storage • Supporting software 	<ul style="list-style-type: none"> • Data Lake archive • EcFlow suite controls and monitoring • Services

To support the implementation, integration and deployment of this DT, dedicated workflows are needed to handle:

- Data preparation, delivery and notifications
- EcFlow Suite configuration, control and deployment

Data produced by the DT is output from the IFS through newly developed MultiIO pipelines to allow encoding, interpolation and statistical processing of data prior to writing to local storage on the host HPC system.

For longer term storage and access by downstream users, a selection of outputs are archived to the Destination Earth Data Lake.

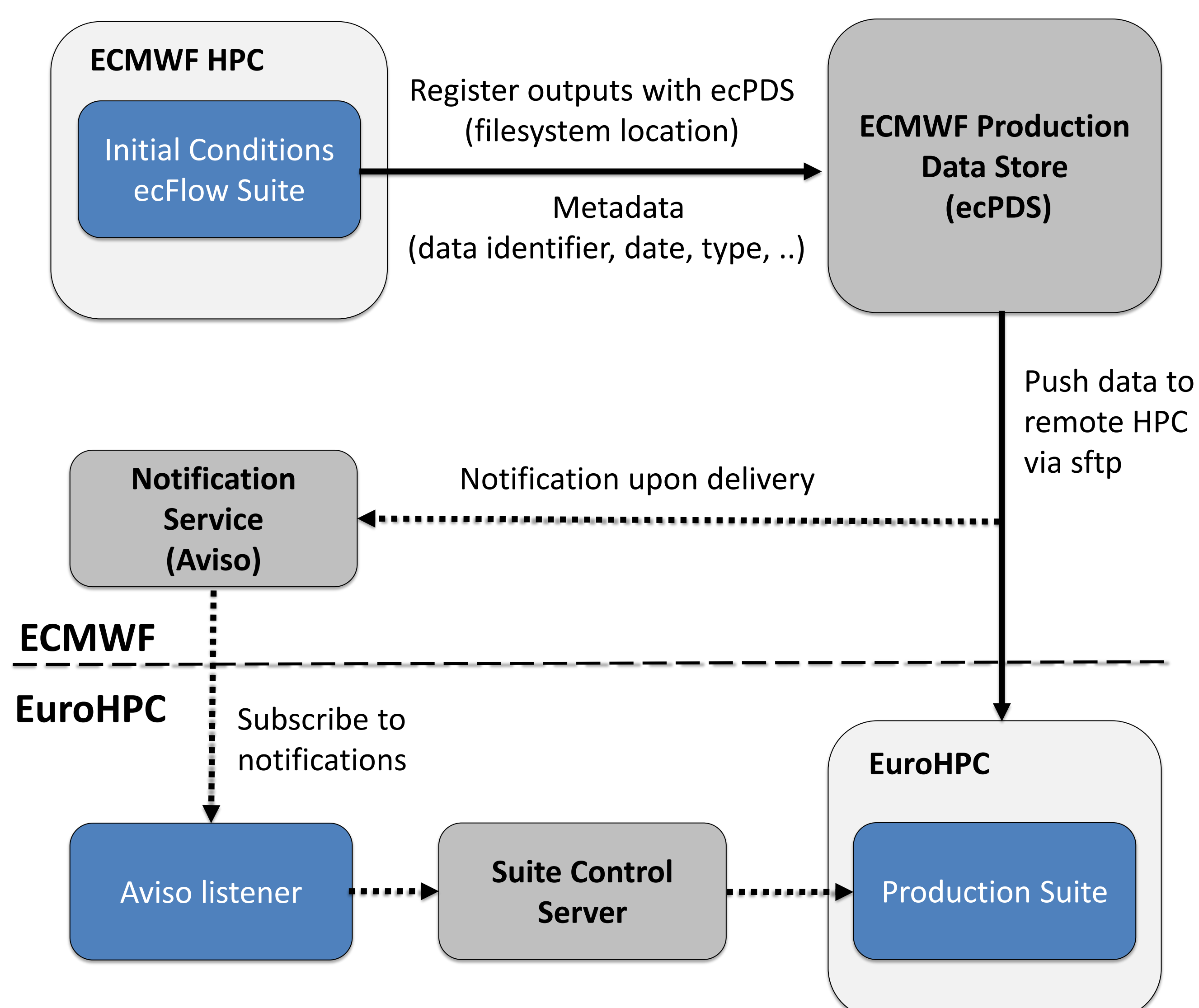
2. Initial Conditions Data

The DT makes use of input data from the ECMWF operational analysis. To make the data available for model ingestion on the EuroHPC host system the following steps are required:

- Scheduled daily extraction of analysis data once available
- Processing of operational data to prepare for high resolution forecast
- Transfer of data to EuroHPC platform
- Notification of data availability

The data preparation process is performed on ECMWF's HPC system. Initial conditions data delivery similar to operational dissemination of data products via the use of:

- ECMWF's Production Data Store (ecPDS)
- Aviso, ECMWF's in-house data availability notification service



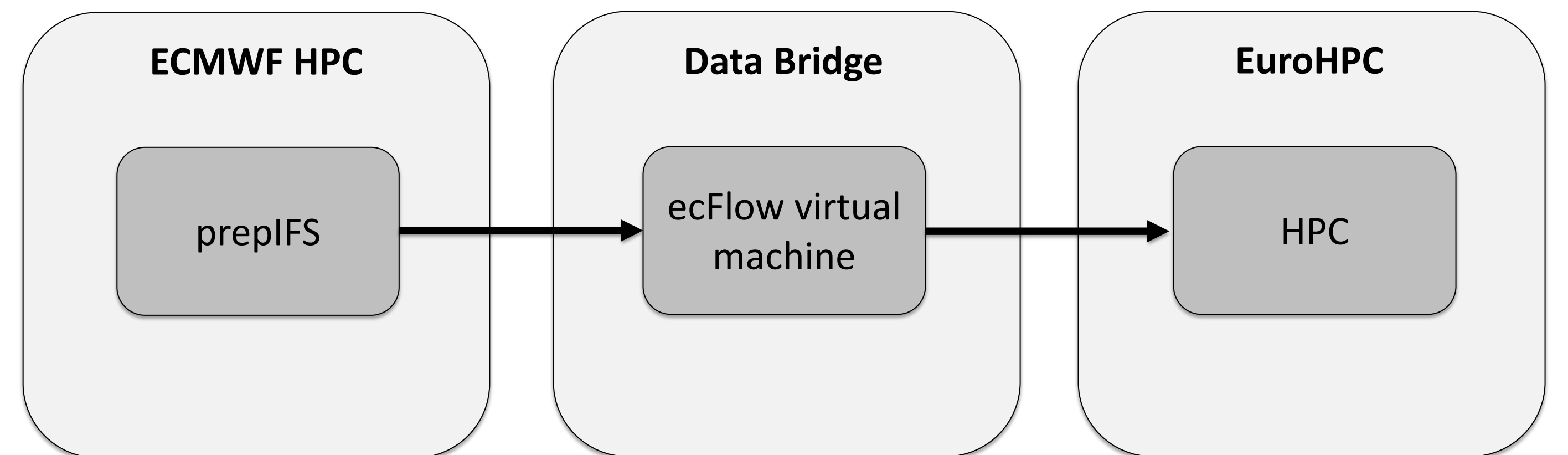
3. Production Suite Configuration and Control

The DT suite is configured via prepIFS, a tool to configure:

- Forecast parameters
- Control server and host HPC
- Computing resource requirements

PrepIFS is routinely used by researchers to configure the IFS and to deploy and submit research suites. Making use of prepIFS to configure the production suite allows one-to-one correspondence of research and production forecast configurations, easing handover between research and production.

The suite is controlled via ecFlow, ECMWF's workflow and job scheduling package.



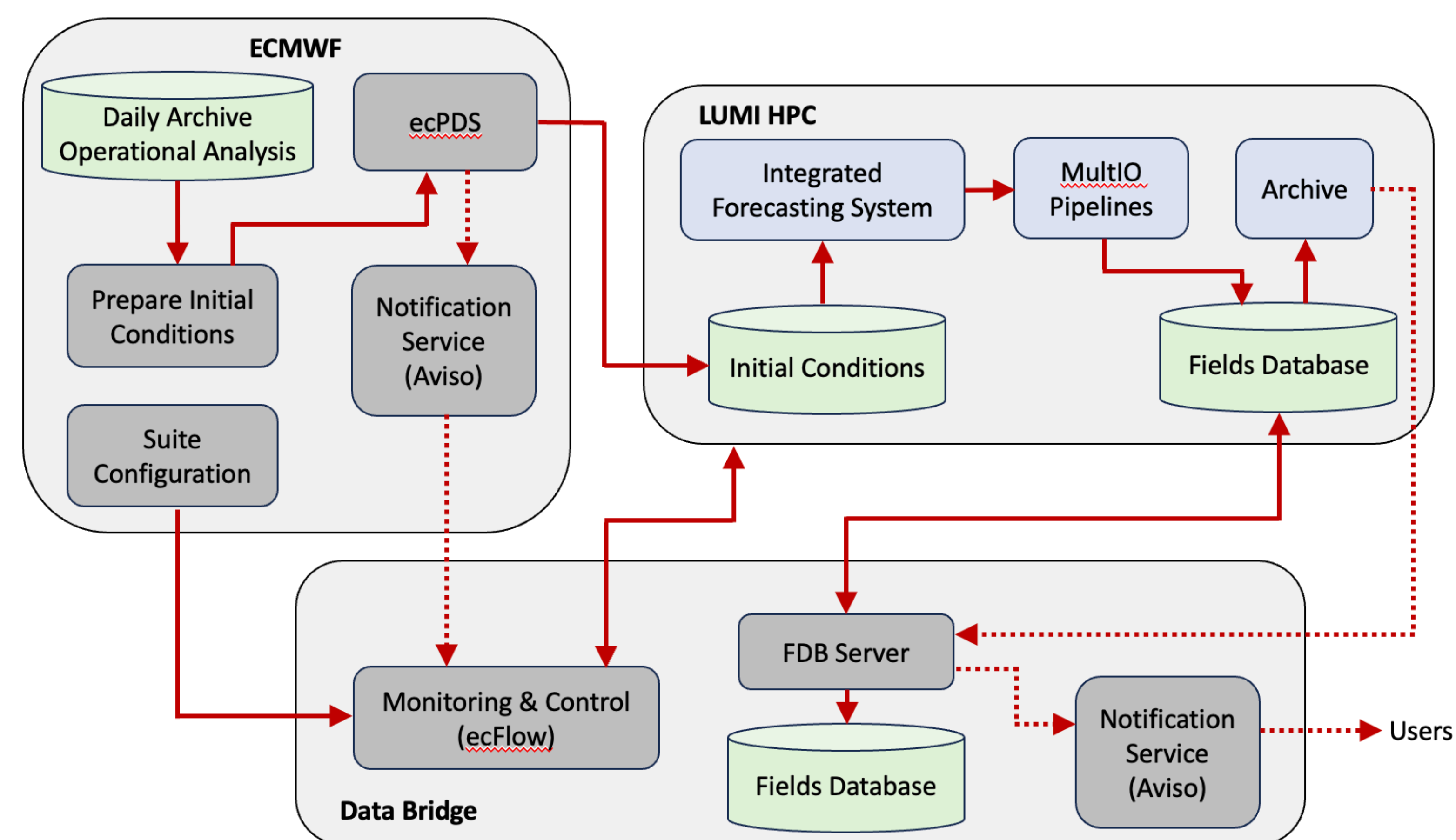
The production suite is configured, deployed and controlled from ECMWF:

1. Suite configuration on ECMWF HPC (prepIFS)
2. Send configuration files to ecFlow production server (hosted on Data Bridge resources)
3. Job submission scripts prepared on ecFlow virtual machine
4. Jobs scheduled and submitted to EuroHPC slurm partitions from ecFlow

A separate development server is used for pre-production testing of ongoing developments.

5. Production Workflow

The data and control workflows established to support the DT come together through a set of complex interactions. A high-level overview of the various workflow components is given below:



6. Workflow and Development Management

The DT is designed to be a production system, and as such access to the control and submission environments is restricted. For different interactions and installations, dedicated service accounts are used, for example:

- A production user account to run production jobs and interact with the control server
- A data transfer account to authenticate for ecPDS data delivery
- A continuous integration account to manage the deployment submission of continuous integration experiments
- A development user account to manage the installation of suite support software and test development workflows

Technical updates can be implemented to the DT after successfully passing through continuous integration pipelines, following successful testing through the development workflow.



Funded by the European Union

Destination Earth

implemented by

