Contribution ID: **177**   Type: **Long Talk**

Copernicus Data Stores as part of the streamlined Cloud-based services strategy at ECMWF

Tuesday, 1 October 2024 16:15 (15 minutes)

ECMWF's IT service provision strategy offers a seamless cloud infrastructure and harmonised services to all ECMWF users, allowing them to effectively use the computing services and data available including Copernicus data and services. To this effect ECMWF procured a "Multi-Purpose Cloud Infrastructure incorporating the European Weather Cloud"(EWC). This infrastructure was recently extended to incorporate the Copernicus Data Stores (CDS/ADS) Services, converging into an ECMWF Common Cloud Infrastructure (CCI). CCI cloud is hosted at ECMWF's data centre in Bologna.

Cloud hosting has provided Data Store (DS) Services since their launch with the capacity for a sustained growth in terms of stored data and computing capabilities which is envisaged to continue at a steady rate in the coming future fostered by its integration within CCI. DSs Services are split into two main layers with different functions and Cloud requirements: Data Repositories and Software Services.

Data Repositories are the foundational base of DSs. They are distributed and diverse in size, format and scope encompassing from a CDS-MARS Archive to other repositories such us a modernised Observations Repository and a series of smaller on-disk datasets. Serving data from cloud disks allow high efficiency and performance able to serve a daily average over 160 TB in the form of more than 500k requests coming from over 3k active users. The Service is also implementing an ARCO (Analysis Ready, Cloud Optimized) Data Lake to improve visualisation and interactivity of C3S and CAMS data on WEkEO and in addition to address the needs of demanding ML/AI solutions and visual-interactive applications.

The Running Services layer is integrated by the software components supporting the range of functionalities of the Data Stores Service. These components are optimized to deploy and run in Cloud environments. They offer to users processing and visualisation capabilities in addition to data access functions. The Climate Data Store (CDS) for C3S, the Atmosphere Data Store (ADS) for CAMS and the recently launched Early Warning Data Store (EWDS) for CEMS are the well-known public facing interfaces of this layer. The operational management of these Services relies heavily on the automatic deployment and configuration of multiple instances facilitated by the elastic resources offered by CCI. As an extension of the Services a Beta version of a Jupiter-Hub is in the pipeline. This will allow users to launch temporal sessions and allocate computing and storage resources to perform computation and visualisation on top of the data using a set of preconfigured expert tools mostly offered by eartkit.

The aim of this presentation is to introduce the scope and plans of the Modernised Data Stores Service and describes how this fits into the streamlined services hosted on the ECMWF Common Cloud Infrastructure (CCI) to foster the future evolution of the Services and the synergies with other platforms such as WEkEO or the Copernicus Data Space Ecosystem.

Topic

Presenter: BIAVATI, Gionata (ECWMF)

Session Classification: Inside Data Spaces: Enabling data sharing paradigms