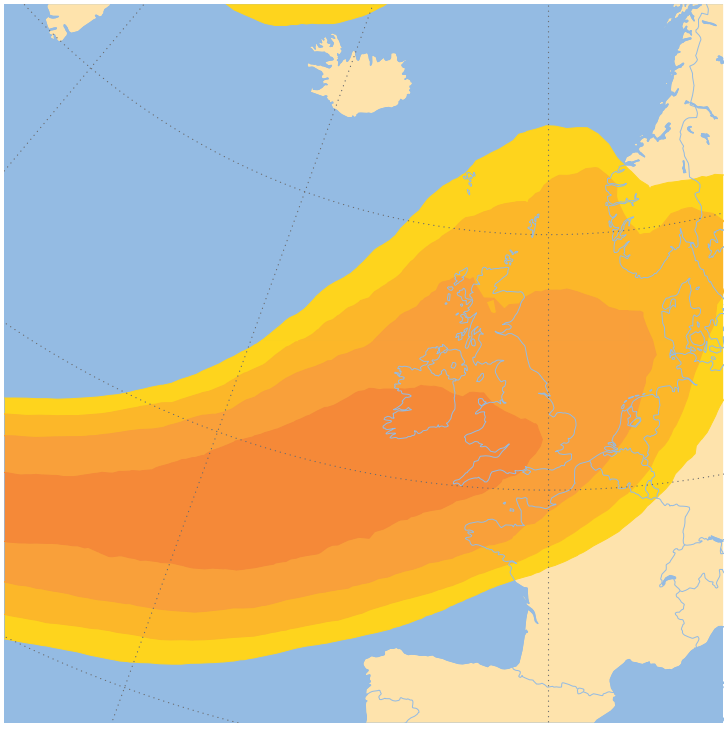
ECMWF Feature article

from Newsletter Number 165 – Autumn 2020

COMPUTING

Progress towards a European Weather Cloud



Cover image: EFI for water vapour flux on 15 February 2020

www.ecmwf.int/en/about/media-centre/media-resources

doi: 10.21957/9pft4uy055

This article appeared in the Computing section of ECMWF Newsletter No. 165 – Autumn 2020, pp. 24–27.

Progress towards a European Weather Cloud

Florian Pappenberger, Martin Palkovič

One of ECMWF's objectives as defined in its Convention is to make the results of the Centre's research and operations available to its Member States "in the most appropriate form". Today, as the volume of data produced at ECMWF continues to grow, cloud technology is becoming the only realistic way to meet this objective. A key advantage of cloud technology is that it can be used to carry out data processing where the data is, without the need to transfer or download large volumes of data. That is why, in December 2018, ECMWF's Council approved a two-year pilot project to create, jointly with EUMETSAT, a federated cloud computing infrastructure focused on meteorological data. Dubbed the European Weather Cloud, it will mainly serve the European Meteorological Infrastructure (EMI) and its users (Box A). As the initial two-year pilot period draws to a close, substantial progress has been made in addressing the project's governance and infrastructure aspects and in trialling use cases. The European Weather Cloud is expected to become operational in 2022 after extending the pilot phase by another year (Figure 1).

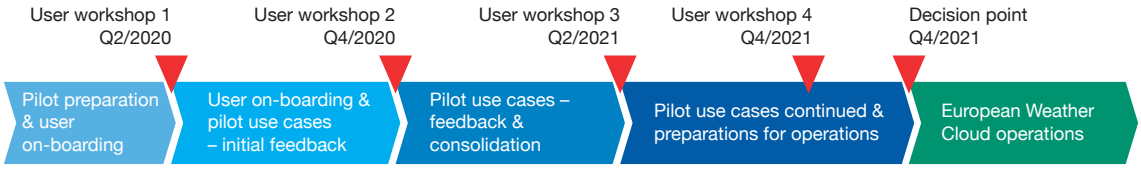


Figure 1 Anticipated timeline from the first quarter of 2020 to the operational phase of the European Weather Cloud.

By creating a joint European Weather Cloud, ECMWF and EUMETSAT will build the foundations for a strong European IT infrastructure enabling direct access to and processing of both organisations' data holdings. This means that observation data, forecast data and meteorological products can be accessed together as if they were collocated. Federation with other relevant cloud infrastructures in our Member States via a dedicated interface will make it possible to further widen the range of accessible data. Such a cloud infrastructure offers significant advantages and has the potential to evolve into a game changer in the way ECMWF, EUMETSAT and Member State data and products are used in the future by their user communities.

Motivation

ECMWF's operational outputs are constantly growing. At the beginning of 2018, the Centre's overall production corresponded to 90 terabytes (TB) of data per day. This is projected to rise to 360 TB by 2022 and to exceed 1 petabyte (PB) by 2026. Today, only about 35 TB of data is disseminated to users every day. The high data volume and lengthy data transfer times are likely to be part of the explanation why only a relatively small proportion of the data produced reaches users. Cloud computing and big data technologies are now sufficiently mature to allow users to make more extensive use of ECMWF products.

Prospective users of the European Weather Cloud **A**

The European Weather Cloud will mainly serve the European Meteorological Infrastructure (EMI) and its users. The EMI comprises several European organisations that are active in the field of meteorology as well as the national meteorological and hydrological services (NMHSs) which they bring together. These organisations are EUMETSAT, ECMWF, EUMETNET (a network of 31 NMHSs) and ECOMET (an economic interest grouping comprising 28 NMHSs).

In addition, in June 2020 ECMWF's Council agreed that the European Weather Cloud can also be used by the NMHSs of the World Meteorological Organization and by research organisations for activities aligned with ECMWF's mission.

doi: 10.21957/9pft4uy055

The basic idea is to bring users to the data instead of transferring the data to users. This is becoming even more important as the Centre moves towards higher-resolution global ensemble forecasts. Decreasing the grid spacing of such forecasts from about 18 km today to a few kilometres, in line with ECMWF's Strategy, will greatly increase the size and complexity of the data volumes produced at the Centre. Cloud technology will enable our Member States and other users to fully exploit the benefits of such data.

In addition, a key idea behind the European Weather Cloud is to bring together data holdings from across the entire EMI in a single federated cloud infrastructure. This will make it possible to access and process multiple datasets, separately or in combination, for the mutual benefit of all entities participating in the European Weather Cloud.

Key concepts

There are three broad layers of services that can be provided on a cloud computing infrastructure:

- Infrastructure as a service (IaaS): this basic cloud computing layer provides computing, network and storage capabilities. Users can commission virtual machines or collections of containers with characteristics of their choice (operating system, number of CPUs, amount of memory, disk volumes…) that they can then start, stop or decommission according to their needs.
- Platform as a service (PaaS): this layer provides access to standard off-the-shelf software and services, such as databases or web servers. Users then build applications that rely on these services but would not necessarily be aware of the systems on which these processes are running.
- Software as a service (SaaS): this layer provides end-user-facing applications. Users interact with them using a web browser and are entirely unaware of the supporting infrastructure.

It is anticipated that all these services will be provided as part of the European Weather Cloud.

A key characteristic of the European Weather Cloud is that it will have a federation capability. This will make it possible for an entity of the EMI (e.g. an ECMWF Member State or a group of Member States) to be connected to the European Weather Cloud via the federation interface (Figure 2). The data holdings of this entity would then become discoverable and accessible for use within the rest of the European Weather Cloud. In turn, the entity's users would be able to access and use all European Weather Cloud data.

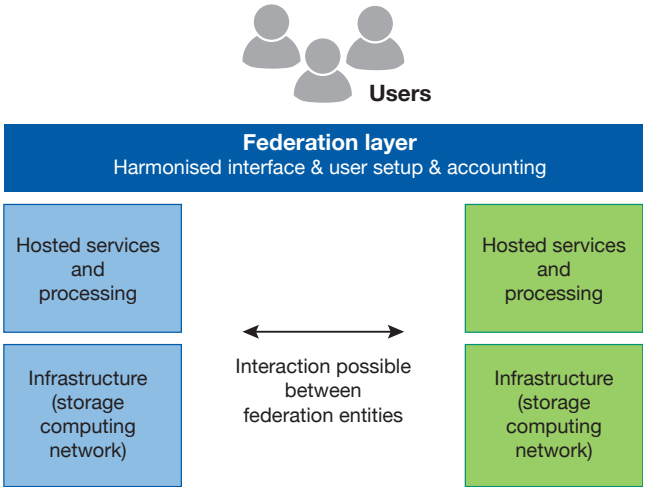


Figure 2 The European Weather Cloud will be a federation of different entities held together by a federation layer. Users can interact with different entities. By interacting with one, they can also access the others. Their European Weather Cloud activities are logged centrally in the federation layer.

Using the federative capability of the European Weather Cloud could be of interest to large national meteorological and hydrological services (NMHS) or to groups of smaller NMHSs with a common thematic or geographic focus. The structure works with two further concepts:

- Portal: The federation layer is delivered through a web portal, which can orchestrate many clouds. Each member of the federation offers a different door to access the federated cloud services through a single European Weather Cloud Portal.
- Tenants: Each project with access to the European Weather Cloud is called a Tenant. Tenants are isolated environments with unique users and workloads, with no access or visibility to other Tenants in the same infrastructure.

doi: 10.21957/9pft4uy055

Governance

The European Weather Cloud will have a three-tier decision-making structure, which is already being applied in the pilot phase. At the top is a common decision layer for governance and management decisions taken by ECMWF and EUMETSAT with their Councils and Committees. There follows a common coordination layer for implementation and operation activities carried out by a Technical Coordination Team. This also provides support for decisions to be taken at the governance level. A local layer is responsible for activities that need to be communicated to other entities of the federation without requiring coordination (Figure 3).

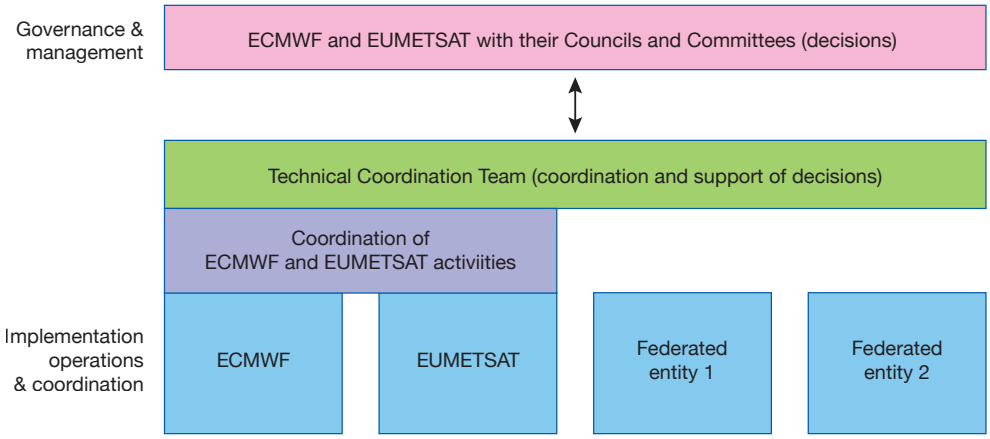


Figure 3 Governance and coordination structure for the European Weather Cloud. While the Councils and Committees are in charge of the governance and direction of the cloud, the cloud entities themselves implement the necessary systems and coordination between the federated entities.

Infrastructure

The design, procurement and deployment of hardware for the initial configuration of the infrastructure have been completed. Below the federation layer, the European Weather Cloud pilot infrastructure at ECMWF comprises two components: a storage platform based on Ceph and a cloud computing platform based on OpenStack, which allow users to create on-demand virtual resources.

Evaluation reports commissioned by ECMWF for its OpenStack and Ceph clusters have found that those systems are performing well and as expected. Recent additions to the infrastructure, like GPUs, make it possible to support resource-demanding artificial intelligence and machine learning workloads. The Ceph cluster has been expanded to further improve its I/O performance.

The ECMWF component of the European Weather Cloud will be part of the Centre's computing capacity available to Member States as outlined in the ECMWF Convention. It will also complement the other computing services available at ECMWF and will be integrated with them. It will be a resource for official Member State duties, predominantly for the purpose of processing ECMWF data and serving products to Member and Co-operating States.

Use cases

For the pilot project, 29 use cases have been defined and are at various stages of development. In addition, an unscheduled use case arose in March this year, when an earthquake damaged the building of the Croatian Meteorological and Hydrological Service (DHMZ). Within days, DHMZ managed to back up its operational production and essential services on ECMWF's High-Performance Computing Facility and the European Weather Cloud. More details on how this was done can be found in a previous ECMWF Newsletter article by Abellan et al. (2020). The pilot project use cases are mostly aimed at users in the developers' own organisations. They include:

- Running web services to explore hosted datasets
- Running an atmospheric dispersion model on ECMWF forecast data
- A platform to support the training of machine learning models on archive datasets
- Research in collaboration with external partners.

doi: 10.21957/9pft4uy055

For example, the German weather service DWD is already feeding maps generated by a server it deployed on the cloud into its public GeoPortal service; a joint EUMETSAT and ECMWF use case will assess bias correction schemes for the assimilation of radiance data based on several satellite data time series; the Royal Netherlands Meteorological Institute (KNMI) will host a climate explorer web application based on KNMI climate explorer data and ECMWF weather and climate reanalyses; the Royal Meteorological Institute of Belgium will prepare ECMWF forecast data for use in a local atmospheric dispersion model; and the EUMETSAT Numerical Weather Prediction Satellite Application Facility (NWP SAF) will develop a training module for a fast radiative transfer model (RTTOV) based on ERA5 reanalysis data.

A joint ECMWF–EUMETSAT user workshop took place on 27 May 2020. The online event was attended by nearly 200 participants at peak times and provided valuable feedback for developing the future cloud service. For more details, see the ECMWF Newsletter article by Siemen et al. (2020). A further virtual ECMWF–EUMETSAT workshop on the European Weather Cloud took place on 10 November 2020.

Pilot federation

The federation capability offered by the European Weather Cloud increases data access for all parties and creates opportunities for processing schemes and applications that were previously not feasible. The underlying principle is to respect the identities of the federated entities and their responsibility for the data. The pilot phase of the European Weather Cloud aims to gather experience in operating such a scheme, initially with a few pilot federations before opening it up to a larger number of NMHSs.

The main aspects of the federation that will be tried out and analysed during the pilot phase relate to identifying technical constraints; validating a framework of reciprocal commitments, rights and responsibilities; validating management and coordination aspects with the federated entities; and demonstrating the benefits of the scheme. The selection of entities participating in the pilot federation is based on their technical readiness and availability for making the required commitments.

Web presence

A European Weather Cloud website has been set up at: *http://www.europeanweather.cloud*. At this stage, the main aim of the site is to enable potential users to find online information on the project's purpose and current status. In the longer term, it is expected that this website will host support pages, the catalogue of available data and services, and system status information. A European Weather Cloud Knowledge Base page has also been created at *https://confluence.ecmwf.int/display/EWCLOUDKB* as a joint effort with EUMETSAT along with a support platform demonstrating the synergies and benefits of working together. This page provides technical details and other information and is intended for pilot users.

Outlook

After nearly two years of the pilot phase of the European Weather Cloud, substantial progress has been made on all aspects of this major project: governance, infrastructure, use cases, federation and web presence. The usefulness of the European Weather Cloud has already been demonstrated by the role it played in backing up Croatia's forecast production following an earthquake earlier this year. Other use cases are also showing great promise by demonstrating the value of combining datasets and processing data in the cloud. The focus for the year ahead will be on further developing and evaluating the use cases. The results will feed into configuration decisions for the operational phase, which is expected to begin in 2022.

Further reading

Abellan, X., K. Horvath, I. Pelajić & A. Stanešić, 2020: Croatian met service backs up its production at ECMWF after earthquake, ECMWF Newsletter No. 164, 5–7.

Siemen, S., X. Abellan, C. Simarro, J. Saalmüller, M. Grant & J. Schulz, 2020: User workshop aids European Weather Cloud development, ECMWF Newsletter No. 164, 16.

doi: 10.21957/9pft4uy055