

EGU2020-10660

<https://doi.org/10.5194/egusphere-egu2020-10660>

EGU General Assembly 2020

© Author(s) 2021. This work is distributed under the Creative Commons Attribution 4.0 License.



## Connecting data streams with On-Demand Services in the Alpine Environmental Data Analysis Centre

**Johannes Munke**<sup>1</sup>, Alexander Götz<sup>1</sup>, Helmut Heller<sup>1</sup>, Stephan Hachinger<sup>1</sup>, Dominik Laux<sup>2</sup>, Oleg Goussev<sup>3</sup>, Jana Handschuh<sup>2,3</sup>, Sabine Wüst<sup>3</sup>, Michael Bittner<sup>2,3</sup>, Roland Mair<sup>4</sup>, Bianca Wittmann<sup>4</sup>, Till Rehm<sup>5</sup>, Inga Beck<sup>5</sup>, and Markus Neumann<sup>5</sup>

<sup>1</sup>Leibniz Supercomputing Centre (LRZ) of the Bavarian Academy of Sciences & Humanities, Garching b. München, Germany

<sup>2</sup>University of Augsburg, Institute of Physics, Augsburg, Germany

<sup>3</sup>German Aerospace Center (DLR), Earth Observation Center, Weßling, Germany

<sup>4</sup>bifa Umweltinstitut GmbH, Augsburg, Germany

<sup>5</sup>Environmental Research Station Schneefernerhaus (UFS), Garmisch, Germany

The AlpEnDAC (Alpine Environmental Data Analysis Center) platform ([www.alpendac.eu](http://www.alpendac.eu)) aims to collect scientific data measured on different high-altitude research stations in the alpine region and beyond. It provides research data management, analysis and simulation services and supports the research activities of the VAO (Virtual Alpine Observatory) community.

With funding from the Bavarian State Ministry of the Environment and Consumer Protection, a new development cycle of the platform was launched in 2019. Novel components for Computing on Demand (CoD), Service on Demand (SoD) and Operating on Demand (OoD) will be integrated into the system. These will help to implement a near-real-time (NRT) decision support for the scientist during measurement processes and a better control of the measurement process.

In this work, the authors present a stream processing architecture to couple the new CoD, SoD and OoD components. Data from measurements (or also simulations) are normally ingested via a representational state transfer application programming interface (REST API) into the AlpEnDAC system. Before such data are stored in the data base, they will be run through a central stream processing engine, based on a message queue (e.g. Apache Kafka) and a series of specialized workers to process the data. A rule engine and analytics tools are connected to this engine and allow the automatic triggering of, e.g., measurements, HPC simulations, or evaluation and notification services in NRT. The services will be usable and configurable, as much as possible, via the AlpEnDAC web portal where also certain measurement device settings can be adjusted. With these developments, we want to make environmental scientists profit from NRT data collection and processing, as it is already an everyday tool, e.g., in the Internet-of-Things sector and in commercial applications.