Contribution ID: **100** Type: **Presentation**

Federated engine for information exchange (Fenix)

Wednesday, 10 October 2018 15:30 (15 minutes)

The neuroscience community has to cope with various data sources each with their specific formats, modalities, spatial and temporal scales (i.e. from multi-electrode array measurements to brain simulations) and with no fixed relationship between them. Thus, the scientific approaches and workflows of this community are typically a moving target, which is much less the case in other disciplines, e.g., high-energy physics. Furthermore, the community is experiencing an increasing demand of computing resources to process data. However, at present, solutions to federate different data sources and couple them with high-end computing capabilities do not exist, or are very limited.

Fenix (https://fenix-ri.eu/) is based on a consortium of five European supercomputing and data centres (BSC, CEA, CINECA, CSCS, and JSC), which agreed to deploy a set of infrastructure services (IaaS) and integrated platform services (iPaaS) to allow the creation of a federated infrastructure and to facilitate access to scalable compute resources, data services, and interactive compute services. The implementation of the Fenix infrastructure is guided by the following considerations:

- It is based on a co-design approach with a set of diverse domain specific use cases which guides both the design of the architecture and its validation.
- Data need to be brought in close proximity to the processing resources at different infrastructure service providers to take advantage of high bandwidth with data repositories and services.
- Federating multiple data resources shall enable easy replication of data at multiple sites to improve resilience, availability as well as access performance of data.
- Services are being implemented in a cloud-like manner that is compatible with the work cultures in scientific computing and data science. Specifically, this entails developing interactive computing capabilities next to extreme-scale computing and data platforms of the participating data centres.
- The level of integration should be kept as low as possible to reduce operational dependencies between the sites (to avoid, e.g., the need for coordinated maintenance and upgrades) and to allow for the local infrastructures to evolve following different technology roadmaps.

Based on the above principles, the Fenix federated infrastructure includes these main components:

- Scalable Compute Services;
- Interactive Compute Services;
- Active Data Repositories based on fast memory and active storage tiers;
- Archival Data Repositories for long term preservation; and
- Information/catalogue services to allow findability and recovery of data.

The major advantages of the Fenix federated architecture are: the use case driven design, the scalability of the services, the easy extensibility which will allow in the future to move to new state of the art solutions or to enable workflows for other scientific communities.

The first steps towards realisation of the Fenix infrastructure will be done within the Interactive Computing E-Infrastructure (ICEI) project, funded by the EC within the Human Brain Project (HBP, https://www.humanbrainproject.eu/). The users of the HBP will be the prime consumers of the resources provided through the infrastructure. Additional resources will be provided to European researchers at large via PRACE (http://www.prace-ri.eu/).

Type of abstract

Presentation

Summary

In this contribution we will present the Fenix federated infrastructure, which aims to facilitate access to scalable compute resources, data services, and interactive compute services for the neuroscience community in particular and for any other European research communities in general.
The presentation will provide an update on the Fenix architecture, on the technological choices made so far, and on the status of its implementation together with the future plans.

Primary author: FIAMENI, Giuseppe (CINECA - Consorzio Interuniversitario)

Co-authors: TESTI, Debora (CINECA); Dr PLEITER, Dirk (Jülich Supercomputing Centre); Mr MUCCI, Roberto (CINECA)

Presenter: FIAMENI, Giuseppe (CINECA - Consorzio Interuniversitario)

Session Classification: Data Management Services

Track Classification: Area 5. Digital Infrastructures for EOSC and/or EDI