FENIX

RESEARCH INFRASTRUCTURE

Overview of the Fenix Infrastructure

The ICEI project has received funding from the European Union's Horizon 2020 research and innovation programme under the grant agreement No 800858.

Contents

Introduction **3**

Fenix Infrastructure Services **4**

- Interactive Computing Services 4
- Scalable Computing Services 4
- Virtual Machine Services 5
- Active Data Repositories 5
- Archival Data Repositories 5

Planned resources **5**

Projects using ICEI resources **6**

Outlook: Example of a use case involving multiple Fenix services **7**

References **9**

1. Introduction

The European supercomputing centers BSC (Spain), CEA (France), CINECA (Italy), ETHZ-CSCS (Switzerland) and JUELICH-JSC (Germany) have joined forces to design and build a federated data and computing infrastructure. The realization of Fenix Research Infrastructure [1], which in the beginning will be used primarily by the Human Brain Project (HBP) [2], started officially in January 2018 with the launch of the "Interactive Computing E-Infrastructure" (ICEI) project. The ICEI project is co-funded by the European Commission through a Specific Grant Agreement (SGA) under the umbrella of the HBP Framework Partnership Agreement (FPA) [3].

The five partners in the Fenix consortium—all of which are Hosting Members of PRACE—are creating a set of e-infrastructure [4] services which serve the HBP and other communities as a basis for the development and operation of community-specific platform tools and services (see Figure 1). To this end, the design and the implementation of the Fenix infrastructure are driven by the needs of the HBP as well as other scientific communities with similar requirements (e.g., materials science) [5].

The key services provided by Fenix will encompass Interactive Computing, Scalable Computing and Virtual Machine (VM) services, as well as Active and Archival Data Repositories (see the next section for short descriptions of these services). A distinctive feature of the Fenix infrastructure is that data storage and scalable computing resources will be in close proximity to each other and tightly integrated. In this manner, Fenix will enable HBP use cases such as GUI-based[1] interaction with large-scale neural network simulations, massive data processing and analysis for the HBP Human Brain Atlas, the validation of simulation results obtained on neuromorphic hardware, and the implementation of the HBP Neurorobotics Platform. The VM services provided as part of the Fenix infrastructure are hosting the HBP Collaboratory, the central gateway to all HBP Platforms, and other HBP platform services.

The ICEI project implements key elements of the Fenix infrastructure through procurements of R&D services and equipment, which are coordinated between the partners. Other significant components of the infrastructure are realized through in-kind contributions from the participating supercomputing centers.

Besides the technical infrastructure, the Fenix consortium has established a governance structure and coordinates the mechanism for the allocation of Fenix resources. While Fenix determines the general principles for the allocation of Fenix resources within stakeholder communities such as the HBP (which must be based on peer review), each community is responsible for the actual distribution of their share within that community. Of the Fenix resources funded through ICEI, a total share of 25% is made available to the HBP, and 15% to European researchers at large (through PRACE). Fenix is open for other communities who want to contribute resources and use the infrastructure.

The first Fenix infrastructure services are already available and used at CSCS (see section 4 for an overview of projects using the services). According to the ICEI project schedule, the deployment of Fenix infrastructure components at the majority of the participating centers and the first demonstration of the key services are expected to start towards the end of 2019. More infrastructure services are planned to be added, with all infrastructure services expected to be operational in early 2021.

[1] GUI: Graphical User Interface

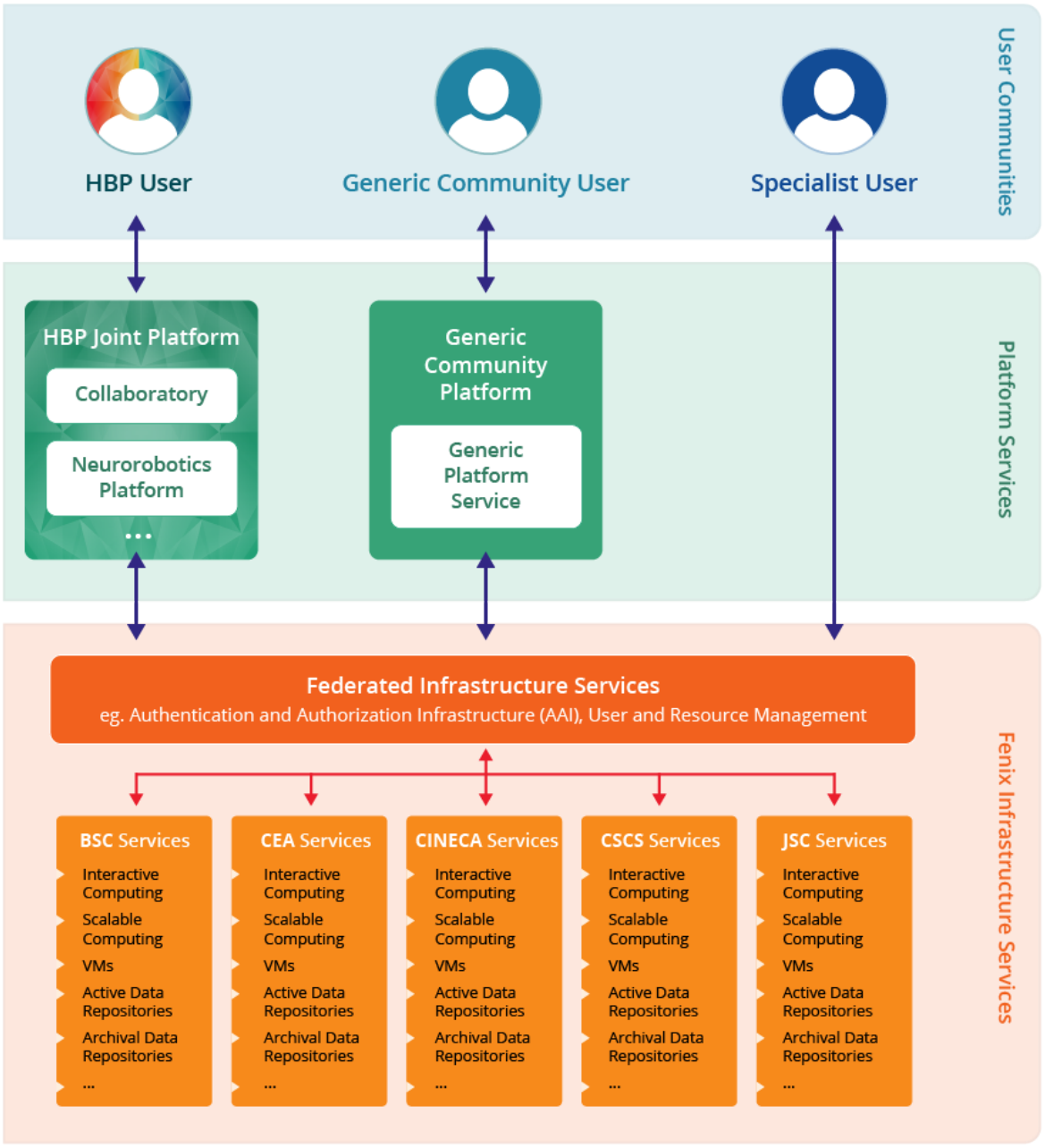

Figure 1: Fenix architecture concept

2. Fenix Infrastructure Services

The Fenix infrastructure provides the following key services:

2.1 Interactive Computing Services

Quick access to single compute servers to analyse and visualise data interactively, or to connect to running simulations, which are using the scalable compute services.

2.2 Scalable Computing Services

Massively parallel HPC systems that are suitable for highly parallel brain simulations or for high-throughput data analysis tasks.

2.3 Virtual Machine Services

Service for deploying virtual machines (VMs) in a stable and controlled environment that is, for example, suitable for deploying platform services like the HBP Collaboratory, image services or neuromorphic computing front-end services.

2.4 Active Data Repositories

Site-local data repositories close to computational and/or visualization resources that are used for storing temporary replicas of data sets. In the near future they will typically be realised using parallel file systems.

2.5 Archival Data Repositories

Federated data storage, optimized for capacity, reliability and availability that is used for long-term storage of large data sets which cannot be easily regenerated. These data stores allow the sharing of data with other researchers inside and outside of HBP.

3. Planned resources

The ICEI partners expect to be able to provide researchers with access to the resources listed below by the end of 2019:

Component	Site (Country)	Total ICEI (100%)	Minimum request
Scalable computing services			
Piz Daint Multicore	CSCS (CH)	250 nodes	1 node
Interactive computing services			
ICCP@JUELICH	JSC (DE)	175 nodes	1 node
Interactive Computing Cluster	CEA (FR)	60 nodes	1 node
Piz Daint Hybrid	CSCS (CH)	400 nodes	1 node
T.B.D.	CINECA (IT)	350 nodes	1 node
T.B.D.	BSC (ES)	6 nodes	1 node
VM services			
ICCP@JUELICH	JSC (DE)	25 nodes	1 VM
Openstack	CEA (FR)	600 VM (20 nodes)	1 VM

compute node			
Pollux Openstack compute node	CSCS (CH)	35 nodes	1 VM
Nord3	BSC (ES)	84 nodes	1 node
Archival data repositories			
Archival	CEA (FR)	7000 TB	0
Archival Data Repository	CSCS (CH)	4000 TB	1 TB
T.B.D.	CINECA (IT)	5000 TB	1 TB
T.B.D.	BSC (ES)	6000 TB	1 TB
Active data repositories			
HPST@JUELICH	JSC (DE)	1 PB	10 TB
Lustre Flash	CEA (FR)	800 TB	1 TB
Data Warp	CSCS (CH)	80 TB	1 TB
T.B.D.	CINECA (IT)	350 TB	1 TB
T.B.D.	BSC (ES)	70 TB	1 TB

4. Projects using ICEI resources

Neuroscientists can apply for resources provided through the ICEI project. As described above, the HBP as a stakeholder community is responsible for deciding on applications and has established an Infrastructure Allocation Committee (IAC) for this purpose.
As of the writing of this document, the following projects have been approved and are carried out using the Fenix infrastructure services available so far (all at CSCS):

1. Virtual Epileptic Patient (Principal Investigator: V. Jirsa): The goal of the Virtual Epileptic Patient is to construct an epilepsy-specific application of The Virtual Brain large-scale modelling approach suitable for inferring an individual patient's pathology from their neuroimaging data, to complement standard clinical practice.
2. Full-scale hippocampus model (PI: M. Migliore): The aim of this project is to study the mechanisms that may contribute to the emergence of higher brain functions at the cellular and behavioural level in the hippocampus.
3. Cerebellum modelling (PI: E. D'Angelo): The goal of this project is to refine the synaptic connectivity, location and strength in order to better match the most critical behaviours seen in experimental traces.

4. Neurorobotics Platform (NRP) development (PI: A. von Arnim): The goal of this project is to enable large-scale NEST simulations running and connected to a virtual robot interacting with a virtual environment.
5. Image segmentation toolkit (ilastik) workflow dev (PI: A. Kreshuk): The goal of this project is to enable the running of ilastik, the image segmentation toolkit developed at EMBL, Heidelberg. The scientific goal of each analysis that uses the service is to be defined by the Neuroscientists that use the ilastik image processing service.
6. NEST network construction and simulation (PI: H. E. Plesser): The goal of this project is to test, benchmark and optimise the NEST simulation engine, especially the network construction, to provide computational neuroscientists with more efficient tools for large-scale cellular-level simulations.
7. SimLab Neuroscience (PI: A. Morrison, B. Orth): This allocation for the SimLab Neuroscience at the Jülich Supercomputing Centre (JSC) is intended for development, testing and demonstration purposes. The goal is to enable the SimLab to support HBP users with their applications for resources as well as the deployment, integration and optimization of software to best exploit the Fenix infrastructure. The allocation is currently used to implement the Learning-to-learn workflow as a demonstrator of the Fenix services available.
8. Model validation Service (PI: A. Davison): The goal of this service is to allow automated validation of computational models against experimental data, with the results being registered on the HBP's validation framework. The HPC resources will provide a platform whereby simulations necessary for validations can be executed remotely, thereby providing a complete web-oriented service.
9. Neuromorphic Computing front-end services (PI: A. Davison): The HBP Neuromorphic Computing (NMC) Platform currently runs a number of servers on a commercial cloud service (Digital Ocean), using Docker. These servers host the NMC job queue, compute quota, benchmarking, and monitoring web services and will be migrated to the ICEI resources as part of this project.

5. Outlook: Example of a use case involving multiple Fenix services

Data management and big data analytics for high-throughput microscopy

The HBP Human Brain Atlas is a multi-scale, multi-model atlas with highly diverse qualitative and quantitative datasets that need to be spatially and semantically registered. These datasets have different file formats, come from different partners and need different kinds of user interface functionality. The data is collected within various HBP Subprojects, first of all in SP2.

This use case is about data management in the context of the HBP Human Brain Atlas: A large amount of imaging data produced on a daily basis by high-throughput microscopes at the Institute of Neuroscience and Medicine in Jülich needs to be processed and stored. To give an example, a single complete human brain at 1-micron resolution requires, depending on the method, a total of 2-6 Petabytes of storage space just for the original data. Since the amount of data is so large, the services used for the

processing and storage should be located close to the data source, i.e., at the Jülich Supercomputing Centre (JSC) in this case, to avoid costly data transfers over long distances. Through web-based services, the final data products are made available to researchers world-wide.

For the processing and storage of the data, the use case has the following requirements with respect to the Fenix infrastructure (see also Figure 2):

- *Scalable compute services* for high-throughput processing of the microscopy data
- *Virtual machines* to run the graphical user interface (GUI) and for the control of the workflow
- *Active data repository* to facilitate fast data access, e.g. for the analysis of specific regions of interest (ROI)
- *Archival data repository* to for the longer-term storage of the original and derived data (about 2 PByte/year)

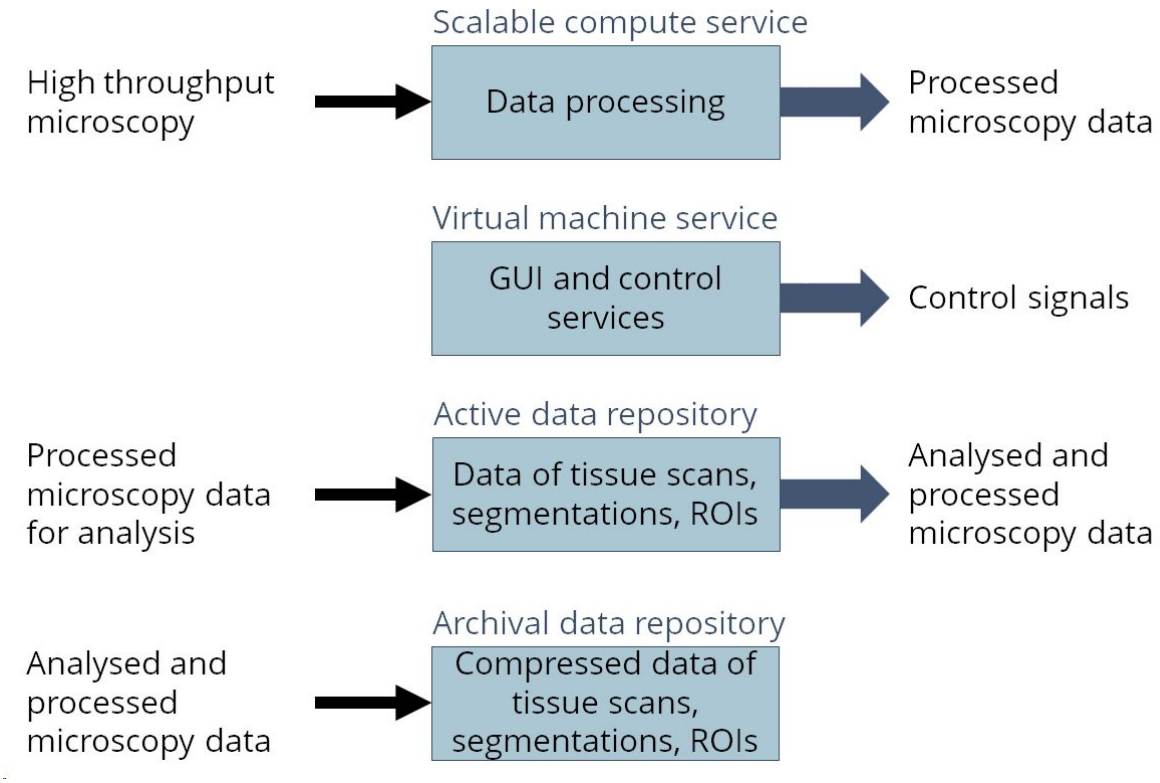

Figure 2: Fenix infrastructure services required for the processing and storage of high-throughput microscopy data for the HBP Human Brain Atlas

The Fenix infrastructure services required for this use case at JSC are expected to become available through the ICEI project in late 2019 / early 2020.

6. References

[1] Fenix Research Infrastructure, https://fenix-ri.eu/

[2] Human Brain Project, https://www.humanbrainproject.eu

[3] Framework Partnership Agreement – Signature (press release), https://www.humanbrainproject.eu/en_GB/fpa-signature

[4] e-Infrastructures, https://ec.europa.eu/digital-single-market/en/e-infrastructures

[5] Human Brain Project: Towards a European infrastructure for brain research, http://inside.hlrs.de/editions/16autumn.html#human-brain-project-towards-a-european-infrastructure-for-brain-research