HPAC Platform Architecture Specification v1

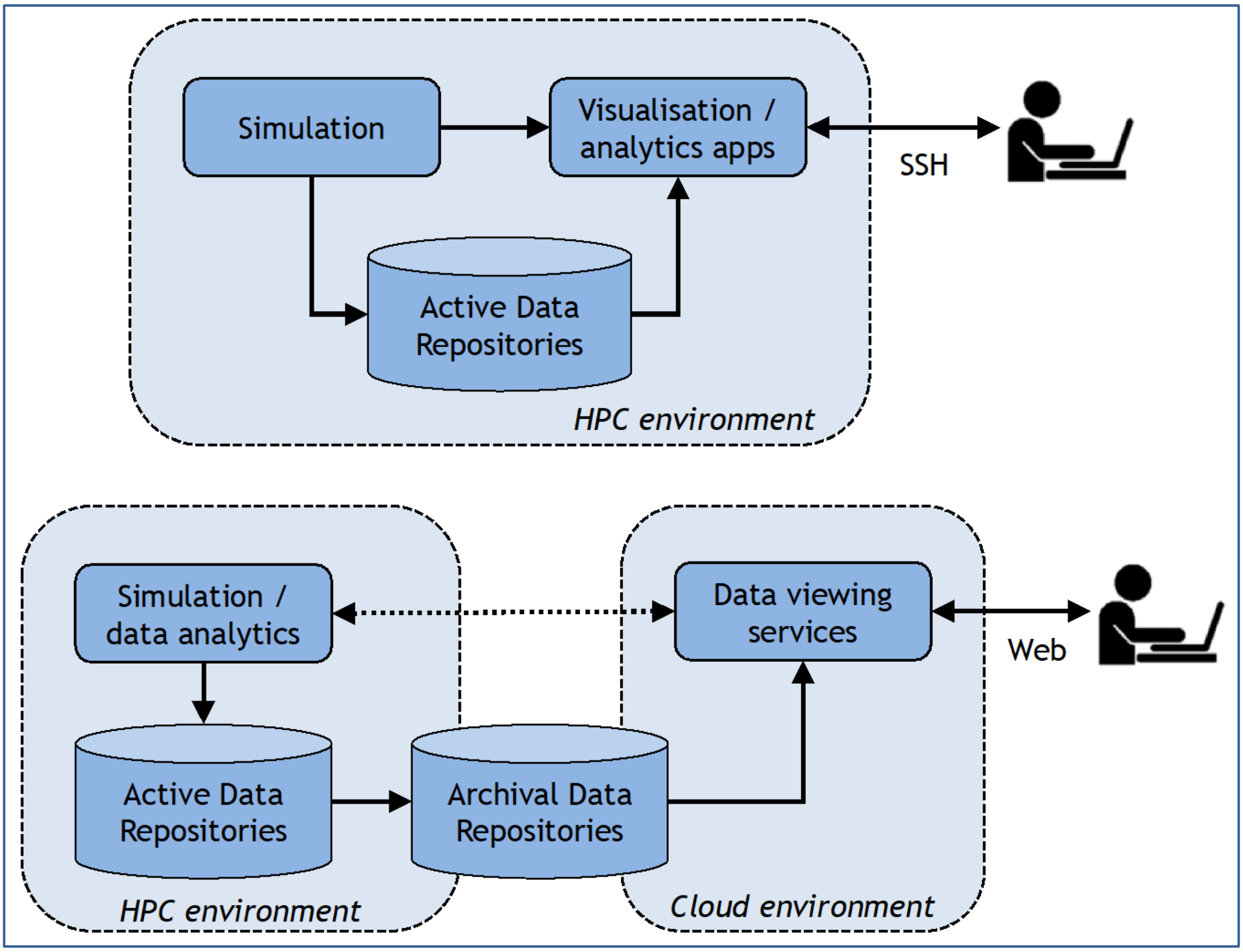

Figure 1: Interactive computing workflows enabled by the HPAC Platform and Fenix

The architecture of the HPAC Platform, as presented in this Deliverable, addresses the challenge of interactivity for workflows using High-Performance Computing and is one of the key contributions of Subproject 7 to the HBP infrastructure. The solid links indicate planned data transport capabilities and connection options, while the dashed links may not be available or restricted for security reasons. When connecting to the HPC environment, a secure protocol like SSH (Secure Shell) needs to be used, while the cloud environment can be accessed using web protocols.

Project Number:	785907	Project Title:	Human Brain Project SGA2

Document Title:	HPAC Platform Architecture Specification v1
Document Filename:	D7.1.1 (D42.1 D81) SGA2 M8 ACCEPTED 200731.docx
Deliverable Number:	SGA2 D7.1.1 (D42.1, D81)
Deliverable Type:	Report
Work Packages:	WP7.1
Dissemination Level:	PU = Public
Planned Delivery Date:	SGA2 M8 / 30 Nov 2018
Actual Delivery Date:	SGA2 M19 / 21 Oct 2019; Accepted: 31 Jul 2020
Author(s):	Sadaf ALAM, ETHZ (P18) Dirk PLEITER, JUELICH (P20)
Compiled by:	Sadaf ALAM, ETHZ (P18)
Contributor(s):	Giuseppe FIAMENI, CINECA (P13) Wouter KLIJN, JUELICH (P20) Anna LÜHRS, JUELICH (P20) Colin MCMURTRIE, ETHZ (P18) Kathrin SAUREN, JUELICH (P20) Bernd SCHULLER, JUELICH (P20)
SciTechCoord Review:	Mehdi SNÈNE, EPFL (P1)
Editorial Review:	Guy WILLIS, EPFL (P1)
Description in GA:	This report will contain a first specification within SGA2 of the HPAC Platform including the Fenix infrastructure.
Abstract:	This document aims at providing the first draft of technical specifications for developing a robust Human Brain Project (HBP) High-Performance Analytics and Computing (HPAC) Platform in SGA2. It brings together the Fenix Infrastructure, a range of operational products and e-infrastructure services federated among five leading supercomputing data centres in Europe, with existing HPAC products and services. Special emphasis is on the level of technical readiness and maturity of services and their composability to deliver an operational, federated e-Infrastructure.
Keywords:	High-Performance Computing, HPC, High-Performance Analytics and Computing Platform, HPAC, Fenix, Interactive Computing e-Infrastructure (ICEI)
Target Users/Readers:	Computational neuroscience community, neuroimaging community, HPC community, HPC industry, consortium members, platform users, platform developers

Table of Contents

Table of Tables

Table of Figures

1. Executive Summary

This document provides the first draft of technical specifications for the ongoing development a robust Human Brain Project (HBP) High Performance Analytics and Computing (HPAC) Platform during SGA2. It brings together the Fenix infrastructure, a range of operational products and e-infrastructure services federated among five leading supercomputing data centres in Europe, with existing HPAC products and services. Special emphasis is placed on the level of technical readiness, maturity of services and their composability to deliver an operational, federated e-Infrastructure.

2. Introduction

One of the key aims of the European Human Brain Project is to build a research infrastructure (RI) to support research and development by the neuroscience communities. The HBP High Performance Analytics and Computing (HPAC) Platform supports that RI by developing and providing supercomputing, storage, visualisation and simulation resources to provision HPC, cloud and large-scale data services.[1] With this RI, scientists can:

- Run large-scale, data intensive, interactive multi-scale brain simulations (up to the scale of a full human brain),
- Manage the large amounts of data used and produced by simulations and in experiments, and
- Manage complex workflows comprising concurrent simulation, data analysis and visualisation workloads.

Hence, the goal of the HPAC Platform is to enable a set of services, which are relevant in the context of the HBP. The architecture of the HPAC Platform, therefore, focuses on enabling complex workflows that exploit high-end computing, networking and storage resources in a distributed and federated e-infrastructure. The aspects of visualisation and simulation software developed and enhanced in the HPAC Platform Subproject (SP7) are not covered in this architecture specification, as these make use of, but are not part of the infrastructure and platform service layer services outlined in this document. Note that the focus on flexible composition of infrastructure services allows for a future migration to support the emerging EBRAINS architecture (HBP SGA3).

The neuroscience community is very diverse, resulting in a very broad range of needs for simulation, data analysis, storage and interactive computing. To address this diversity, a set of e-infrastructure services that will be federated among five leading supercomputing centres in Europe was introduced: the Fenix Infrastructure.[2] The distinguishing characteristic of this e-infrastructure is close proximity and well integration of data repositories and scalable supercomputing systems. An initial version of the Fenix infrastructure is currently being realised through the ICEI project (The European Interactive Computing e-Infrastructure). The ICEI project is funded by the European Commission as a separate Specific Grant Agreement under the HBP Framework Partnership Agreement, and is implemented by the following leading European Supercomputing Centres: BSC (Spain), CEA (France), CINECA (Italy), ETHZ/CSCS (Switzerland) and JUELICH/JSC (Germany). As the initial prime user of this research infrastructure, the HBP is leading the development of community-specific services on top of the Fenix infrastructure services.

An overview and analysis of HPAC and Fenix services is provided in this document in order to highlight the convergence of their infrastructure services. Technical readiness and maturity of products and services are key metrics for these infrastructure services. Additional critical requirements for infrastructure services include security, identity and access management, and accounting and reporting of resources. Research and development (R&D) projects are in place, funded by SGA2 and ICEI, to fill the gaps, where a solution is not of production quality or does not exist yet. Details on the convergence of different efforts, the transition to the newly defined set of Fenix infrastructure

1 https://www.humanbrainproject.eu/en/hbp-platforms/hpac-platform/
2 https://fenix-ri.eu

services, and R&D areas, are highlighted in different sections of this document. Implementation details of the HBP RI Platforms are outside the scope of this document.

There are additional services, namely the platform services that users and developers of community platforms require, in order to exploit infrastructure resources. These platform services require a uniform description and interfaces to infrastructure services, which have been high priorities for the ICEI project. Hence, the HPAC Platform architectural specifications cover the building blocks of infrastructure and platform services, as well as their interfaces and dependencies. As for the infrastructure services, technical readiness and maturity of solutions and products are key metrics. Gaps and plans for transition are identified, where a solution is not production quality or does not exist, to support a federated, HPAC Platform architecture.

3. Architectural Specifications

In order to propose an architecture for the HPAC Platform, two high-level concepts are considered: Infrastructure Services (IS) and Platform Services (PS). We use the term IS to refer to all low-level IT services that could be used to serve the platform requirements of different research communities. IS are typically not neuroscience-specific but rather more inclusive and broader infrastructure elements. These services will be federated by means of Authentication and Authorisation Infrastructure (AAI) services, as well as user and resource management services like FURMS (see Figure 2) to provide users with seamless access. The term PS refers to the set of services that are specifically targeting the neuroscience community and the brain research infrastructure that is being built by the HBP. Some of these services are developed, deployed, managed and operated by the data centres in SP7. However, others may be operated by other entities within the HBP, on top of the infrastructure services provided by these data centres.

Regardless of the classification, each service must exhibit a set of functional attributes and technical implementation options. These attributes include the level of technical readiness and maturity, dependency on other services and composability. A commonly used classification called TRLs (Technology Readiness Levels) is used to specify the maturity of a product or service for production release. As there are different sets of TRL definitions, in this Deliverable, we use the one employed in the EC's H2020 programme[3], which differs from the definition that is being developed within the HBP, as parts of e-infrastructure architecture layers described here extend beyond the HBP. More specifically, the following definitions are used in this document:

- TRL 9 – actual system proven in operational environment
- TRL 8 – system complete and qualified
- TRL 7 – system prototype demonstration in an operational environment
- TRL 6 – technology demonstrated in a relevant environment

Any solution lower than TRL6 is excluded from consideration for the HPAC Platform.

3.1 Infrastructure Services

Table 1 lists the set of infrastructure services that are being deployed in the ICEI project and are operated or planned to be operated by Subproject 7 (SP7) in HBP SGA2 and Work Package 6 (WP6) in SGA3, respectively. The technical readiness levels refer to what we expect to achieve by the end of 2020. An earlier deadline would mean reducing the associated TRL for these services.

[3] https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-trl_en.pdf

Table 1: List of the infrastructure services

	Service Name	Technology Readiness Level	Description
IS-1	SCC (Scalable Computing Services)	TRL9	Massively parallel HPC systems for scalable applications with possibly elastic access. Classic resource management and scheduling systems, such as batch job submission systems, facilitate access to these resources.
IS-2	IAC (Interactive Computing Services)	TRL7	IAC services provide quick access to single servers that provide access to data repositories (i.e. ARD and ACD services) and potentially allow scaling-out using SCC services or connection to running scalable compute jobs. The ICEI project plans to ensure the availability of an integrated product.
IS-3	VM (Virtual Machine Services)	TRL9	Service for deploying virtual machines in a stable and controlled environment that is, for example, suitable for deploying platform services like the Collaboratory, HBP information catalogues, image services, etc.
IS-4	ACD (Active Data Repositories)	TRL9	Site-local data repositories located close to computational and/or visualisation resources at the same site, used for storing temporary slave replicas of large data sets.
IS-5	ARD (Archival Data Repositories)	TRL8	Federated data store optimised for capacity, reliability and availability, used for long-term storage of large data sets that cannot be easily regenerated. These repositories are accessed through an OpenStack SWIFT interface, to ensure coherent access control within a federated infrastructure, through an open protocol. The ICEI project plans to enable this product across all Fenix sites.
IS-6	NETE (External Interconnect)	TRL9	Network connectivity between the different Fenix sites (NETE) to support federation of identity and access management, and Fenix resources.
IS-7	NETI (Internal Interconnect)	TRL9	Site-local, data centre-level interconnect (NETI), to provide classic HPC low-level and high-bandwidth connectivity to compute and storage resources.
IS-8	DM (Data Mover Service)	TRL7	Site-local service for moving data between Archival and Active Data Repositories within the same data centre. The ICEI project plans to ensure the availability of an integrated product.
IS-9	DT (Data Transfer Service)	TRL7	Service for transferring data between different Archival Data Repositories (and possibly external e-Infrastructures). The HBP's WP7.2 aims to ensure the availability of an integrated product.
IS-10	DL (Data Location Service)	TRL7	Central (Fenix) service supporting basic data discovery and access functionalities. The HBP's WP7.2 aims to ensure the availability of an integrated product.
IS-11	AAI (Authentication and Authorization Infrastructure)	TRL8	Set of services needed for authentication and authorisation with all relevant Fenix services. The ICEI project plans to ensure the availability of an integrated product.
IS-12	FURMS (Fenix User and Resource Management Services)	TRL7	Service for managing Fenix Users and Fenix Communities, as well as enabling the allocation and consumption of Fenix Credits. The ICEI project plans to ensure the availability of an integrated product.
IS-13	MONI (Monitoring services for infrastructure services)	TRL9	Set of services for monitoring the state of Fenix IT infrastructure and IT infrastructure services for availability and accountability.
IS-14	SEC (Security services)	TRL9	Compliance with Security for Collaborating Infrastructures Trust Framework such as the Wise Information Security for e-Infrastructures (WISE) trust community.

3.2 Platform Services

The initial concept for HPAC Platform services was tightly coupled with the development of community platforms. The platforms show great diversity, not only in terms of computational needs and the demands of their respective user communities, but also in the level of technical maturity within each area of expertise. The platform services are therefore defined as generic platform interfaces. This will facilitate different rates of adoption by different users and platforms, according to their IT infrastructure needs and demands. Table 2 provides a list of these services, which are largely represented by a set of tools and technologies for contemporary IT workflows, routinely used in public cloud environments and HPC ecosystems. These, however, need to be customized for the neuroscience community. Hence, their technical readiness levels are not at the same level as those of the infrastructure services. Visualisation services are an exception, as, instead of leveraging an existing HPC or cloud technology, there is a continuous need to develop specific tools and solutions for the neuroscience community.

Table 2: List of proposed HPAC Platform services for architectural specification

	Service Name	Technology Readiness Level	Description
PS-1	SWD (Software deployment services)	TRL8	These services facilitate the availability of HBP software products across the Fenix IT infrastructure. Hence, it aggregates services that focus on community-specific support for developing and deploying applications, packages and, where applicable, using container technologies. SWD are taken care of in T7.5.3
PS-2	DEV (Developer services)	TRL8	Platforms rely on standard source code management systems, bug tracking systems and continuous integration systems that need to be instantiated in a way that is specific to the HBP community. DEV are taken care of in T7.5.4
PS-3	DB (Database services)	TRL8	Both structured and unstructured database (DB) and data base management system (DBMS) services are needed by platforms for managing and maintaining access control to data and information. DB are currently not taken care of in SP7.
PS-4	DMGT (Data management)	TRL7	Interfaces to information about data stored in Fenix archival data repositories and their transfer and federation for different platform services. DMGT are taken care of in T7.5.1 (low-level infrastructure) and T7.2.2 (transfer services).
PS-5	MONP (Monitoring services for platforms)	TRL6	Tools and services that allow platform users and the platforms themselves to access information on the availability of IT infrastructure resources and accountability. MONP are taken care of in T7.5.1.
PS-6	VIZ (Visualisation services)	TRL7	Application-specific and HPC visualisation services are available to users and platform developers. These are tuned and optimised for performance and access control when supported by the Fenix IT infrastructure. VIZ are taken care of in T7.2.3.

3.3 Specifications

A high-level view of the relationship between an HBP end user and a developer of an HBP-specific community platform with IT infrastructure, IS, PS and Fenix sites is shown in Figure 2: Overview of Fenix services implemented by the ICEI project and HBP HPAC services and HBP platforms. Fenix IT sites deploy a diverse set of compute, storage and networking infrastructures to accommodate the varied needs of the HBP communities. This IT infrastructure is made available to the HBP platforms and users under a common definition: the Fenix infrastructure services. This is done in order to facilitate ease of access and productivity of users and platform developers. This is particularly critical for complex and distributed workflows that require federated access control and data transfer services. IT infrastructure services provide this layer of federation, access and accounting for diverse and distributed infrastructure, as highlighted in Figure 2. At the platform level, software development, packaging and distribution services facilitate the execution of HPC applications on the Fenix IT ecosystem in a capable and productive manner. HBP users will be able to access Fenix resources via aggregated platform services and infrastructure services.

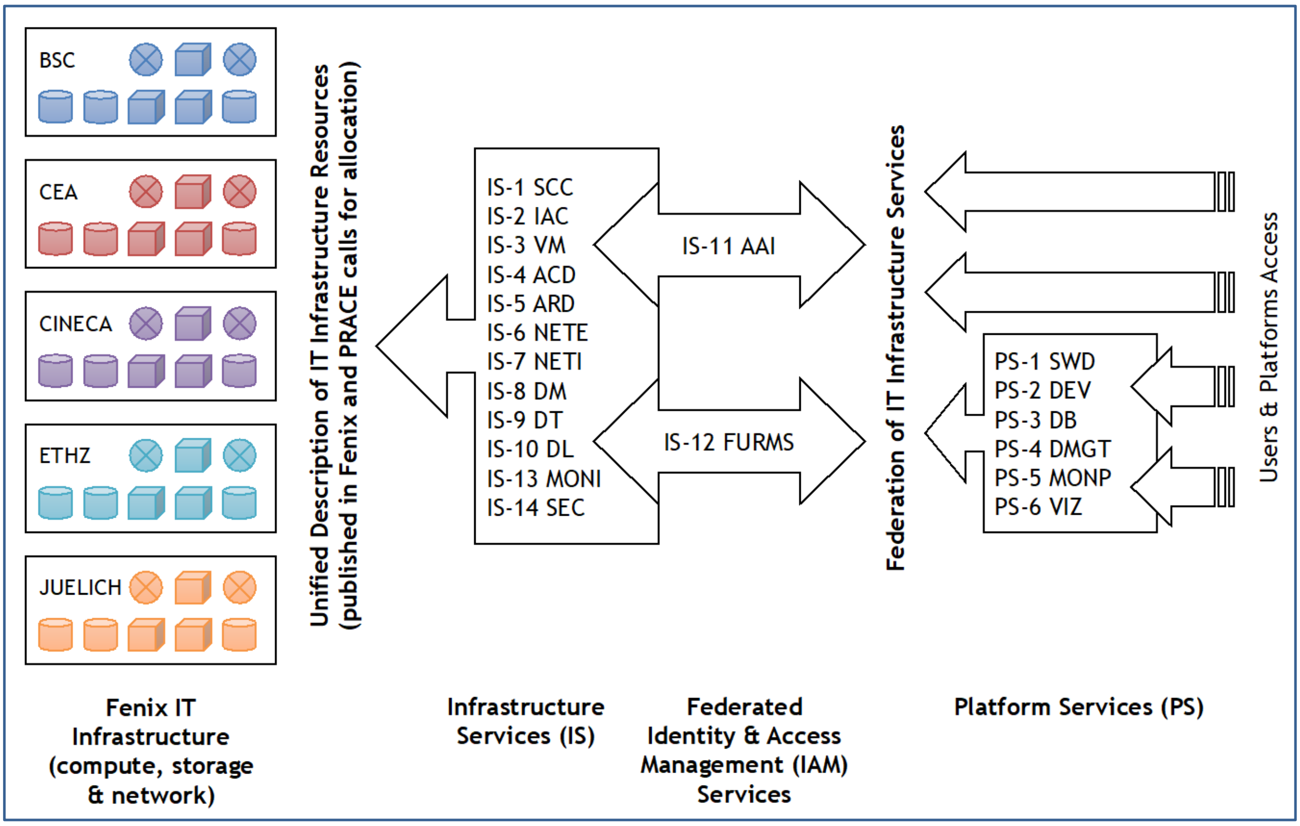

Figure 2: Overview of Fenix services implemented by the ICEI project and HBP HPAC services and HBP platforms

A reference architecture for Fenix for the ICEI implementation project is shown in Figure 3. Each site instantiates an IS by exploiting underlying hardware building blocks and by enabling user interfaces to deliver a service. For instance, SCC at CSCS builds on a hybrid and heterogeneous HPC system (containing both dual-socket, multicore-only and multicore-GPU nodes) with a high-bandwidth, low-latency network fabric (NETI). Users typically access this service through a batch job submission system, such as SLURM. Additionally, there are services in place, such as AAI, for identity management and access control, and the Fenix User and Resource Management Service (FURMS) for accounting and reporting of the usage of resources. The IS services must maintain an open interface across sites and a high level of operational quality by relying on security and monitoring services. Authentication and access control to a diverse set of HPC, cloud-computing and data-management services are management by a unified interface, i.e. AAI and FURMS. This reference architecture provides for different implementation scenarios, thereby allowing Fenix sites to be compliant with the overall architecture, while maintaining their site autonomy.

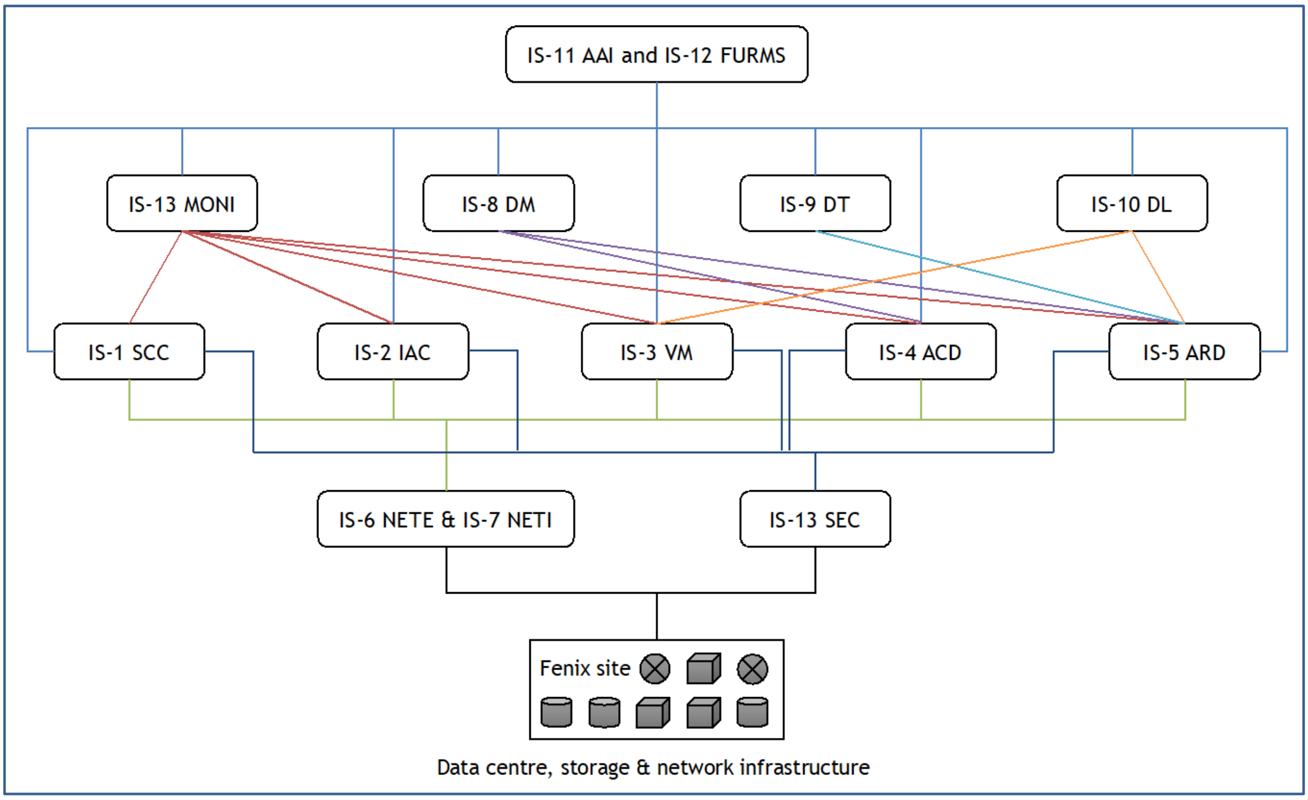

Figure 3: Architecture for the infrastructure services and their interface to the site local infrastructure

In this Deliverable, we document a similar, modular approach, that has been adopted in designing the architecture for the platform services, as shown in Figure 4. Platform services depend on the infrastructure services - more specifically, AAI and FURMS - for identity and access control. A user or a service accessing an IS needs to comply with the access attributes and controls. Please see the initial Fenix AAI specification [ICEI D3.2] and FURMS concepts [ICEI D4.15] for details. We consider container technologies as an example, which can be provisioned and restricted by a Fenix site. For instance, a site can provide an HPC container runtime for Docker containers, like Shifter or Singularity, on SCC resources or a compute service with access to a shared file system like ACD. The same site can provide this service on a VM system, using Docker within a Cloud environment. The choice may be made based on restrictions for access to IS due to the local or national security policies. These may, for example, restrict the access to an HPC system and will therefore only allow access to a container service in a Cloud environment.

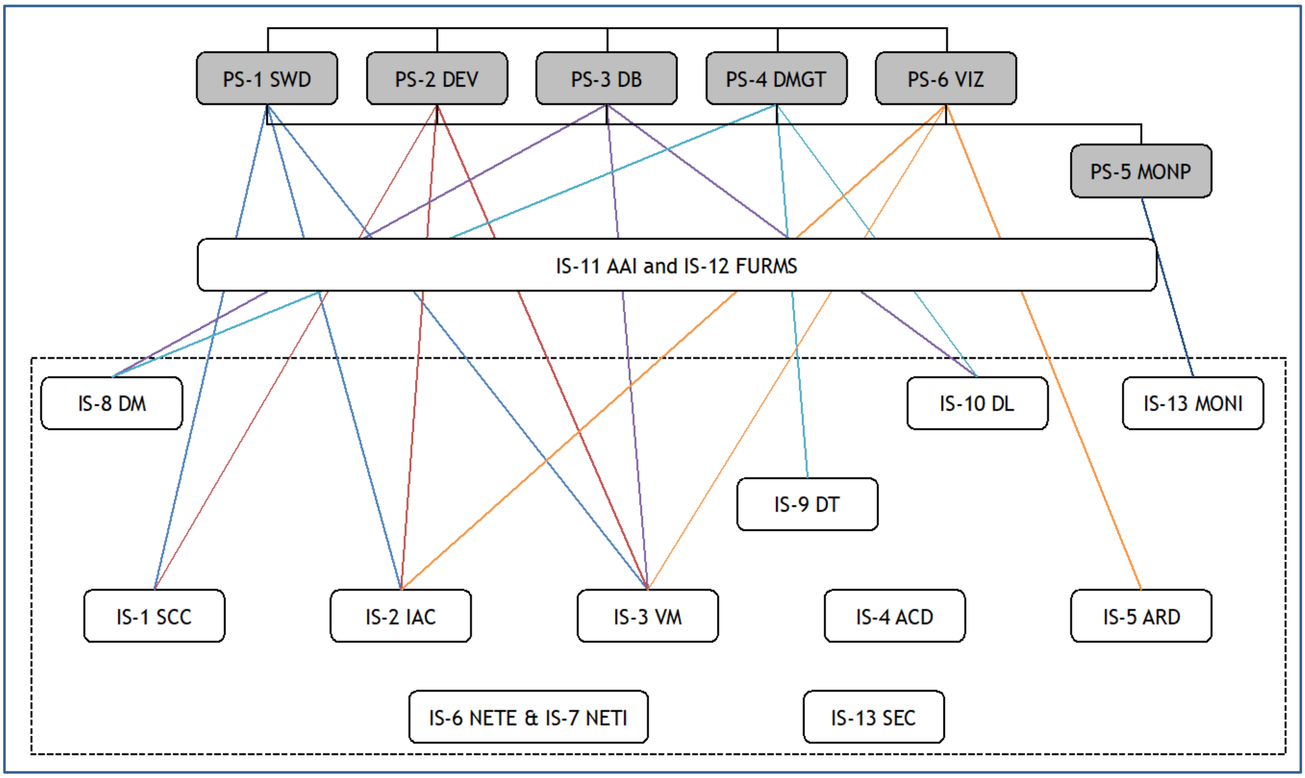

Figure 4: Architecture for HPAC Platform services and interfaces with federated Fenix IT infrastructure services

4. HPAC Platform Services Specifications

Each HPAC Platform service is described in terms of features, implementation status and options, requirements and dependencies to ensure integration into the HPAC architecture. The services are scoped within the HBP by leveraging results from the Co-Design Projects[4], as well as based on the analysis of the use and science cases collected and documented by the ICEI project [ICEI D3.6] and in T7.1.2 in SGA2.

The central concepts of the platform services are development and deployment of software products for the neuroscience communities on the infrastructure. These products have, for example, been developed as part of WP7.3 (Exascale simulator and visualization technology). Therefore, there is a strong emphasis on software build and developer services, as well as community-specific data management interfaces and DB representations. An overview of a software development and deployment pipeline is shown in Figure 5. A similar approach has been adopted in the HPC community in general, as part of the US Exascale computing project called the extreme-scale software stack (e4s) [ECP-E4S].

[4] https://www.humanbrainproject.eu/en/about/project-structure/codesign-projects/

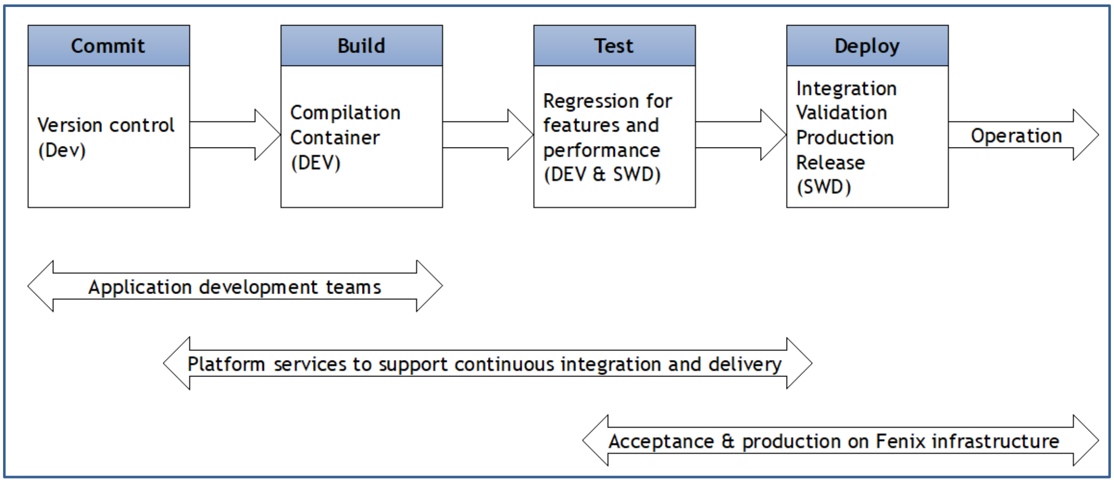

Figure 5: Development and deployment services to enable availability of neuroscience community software stacks that are part of the Work Packages of SP7

4.1 Software Deployment Services (SWD)

Table 3: Software Deployment Services (SWD)

Description	These services facilitate the availability of tested and validated HBP software products across the Fenix IT infrastructure. Platforms rely on the provisioning of virtual machines on the Fenix VM services, for instance, provisioning of platform-specific database services.
Key features	Software stack and runtime to support integration, validation and production readiness of HBP community software. The infrastructure is expected to ensure compliance with Docker image format, OCI runtime and data centre security policy for containerised applications. Flexibility to support custom VM configurations including compute, memory, storage and networking for platform services.
Implementation status and options	Status: ICEI VM services offer options (flavours) to create community DB solutions. Efficient and secure solutions for HPC and cloud software delivery do not exist. Both HPC (Shifter) and cloud container (Docker) technologies are available to HPAC Platform users. Options: Jenkins, Travis Easybuild or similar; Docker or similar container management and runtime systems. Image build service (docker image format compliant); container runtime (Singularity, Shifter, CharlieCloud, and similar); local registry. OpenStack admin controls for creating custom configuration options, integration of telemetry services, integration for Fenix AAI for load balancing and high-availability services, such as DB, etc.
Requirements	Users should be able to both build and deploy their containers. Sites may choose to support different continuous integration (CI) and continuous delivery (CD) models, based on their local security and infrastructure requirements for managing and maintaining CI/CD pipelines. High-availability infrastructure for source and binary distribution of HBP software products. This could be similar to the workflow and tools being employed by the US Exascale computing project.[5] Availability of specific hardware for VM service across the Fenix sites is essential for requesting custom configurations, which will be reflected in FURMS and accounting and reporting.
IS and PS dependencies	SCC, IAC, VM, DEV, AAI, ARD, DB

[5] https://cdn.fs.pathlms.com/G4m3kebCR76rjSVmLLdn (Supporting Continuous Integration at Large-Scale HPC Centers)

4.2 Developer Services (DEV)

Table 4: Developer Services (DEV)

Description	Platforms rely on standard source code management systems, bug tracking systems and continuous integration systems. A standard software packaging and distribution technology that can be supported by HPC and cloud infrastructure services.
Key features	Commonly used version control, build, packaging and testing systems
Implementation status and options	Status: Different tools available at different sites are available to HPAC users; currently not integrated with Fenix AAI and FURMS. Options: EasyBuild, Spack or similar; Jenkins, Travis or similar; GitHub, GitLab or similar
Requirements	Integration of continuous testing part with the batch scheduling system. Dependencies on Fenix computing infrastructure and services
IS and PS dependencies	AAI, FURMS, SWD, SCC, IAC

4.3 Database Services (DB)

Table 5: Database Services (DB)

Description	Both structured and unstructured DB services are needed by platforms for managing and maintaining access control to data and information.
Key features	Support for at least one structured (e.g. SQL) and one unstructured (e.g. noSQL) DB service.
Implementation status and options	Status: Currently a service integrated with Fenix AAI is not available to HPAC users. Options: Widely used options such as mySQL and mongoDB. Dependencies on the VM service model and Fenix AAI may restrict implementation options.
Requirements	Size (capacity and scalability), performance, availability and redundancy. There is dependency on the VM service models of the Fenix infrastructure.
IS and PS dependencies	AAI, FURMS, VM, IAC, ARD, SWD

4.4 Data Management (DMGT)

Table 6: Data Management (DMGT)

Description	Interfaces to information in Fenix archival data repositories and their transfer and federation for different platform services.
Key features	Integration of Fenix AAI architecture, FURMS for user and project management and Swift API.
Implementation status and options	Status: currently not available. Options: Site specific identity and access management, accounting, reporting and monitoring services and integration into storage management system such as keycloak, Swift IAM APIs, and FURMS.
Requirements	FURMS API, Fenix AAI and Swift telemetry API support. Implementation relies on the Fenix data mover and data transfer services.
IS and PS dependencies	DL, DT, DM, AAI, ARD, ACD

4.5 Monitoring Services (MONP)

Table 7: Monitoring Services (MONP)

Description	Platforms users and the platforms themselves may rely on the availability of specific platform services and therefore monitoring services should allow identification of the overall infrastructure status at platform level in a close to real-time manner
Key features	The following features should be enabled at the platform level: • Availability (up, down, reduced functionality, etc.) • Overall health of a service • Metrics (usage, reliability, etc.) • Reporting (periodic accounting and reporting) • Notifications (security, maintenance, etc.) Common features among these environments include programmable interfaces for customisation and integration of 3rd party applications, analysis tools for logs (debugging/troubleshooting) and detailed usage reports and statistics.
Implementation status and options	Status: only discrete HPC and cloud options exist for HPAC Platforms Options: Output can be generated in the form of dashboards, charts and alerts. There are examples in public clouds that could provide a reference for platform- and application-level monitoring. For instance, AWS has CloudWatch, MS has Azure monitor and Google has stackdriver. Other communities such as WLCG (Worldwide LHC Computing Grid) has SAM.
Requirements	Some platform monitoring services will depend on infrastructure monitoring services. However, the aim is to provide these services in a manner that is decoupled from a specific infrastructure. It is expected that generic cloud technologies, such as elasticsearch, logstash, kibana and grafana, can be deployed at the e-Infrastructure or generic platform levels to provide standard APIs or plugins. A centralised service should allow the information to be aggregated and presented to persons in charge of research infrastructure operations, so that issues can be quickly identified.
IS and PS dependencies	AAI, FURMS, MONI, SCC, IAC, VM, ACD, SRD, NETE

4.6 Visualisation Services (VIZ)

Table 8: Visualisation Services (VIZ)

Description	Application-specific and HPC visualisation services are available to users and platform developers. These are tuned and optimised for the Fenix IT infrastructure for performance and access control.
Key features	Visualisation services that are tuned to HPC ecosystems such as ParaView, VisIT or domain specific visualisation tools and collaborative tools with support for Swift backend.
Implementation status and options	Status: currently no coherent service exists across the HPAC Platform, especially for community-specific visualisation solutions; some solutions for *in-situ* visual analysis of simulation data (T7.3.8) and a low-level visualisation back end are under development (T7.3.9), an overview of visualisation software is being prepared (T7.4.3). A service that allows screencast of applications running on Linux has been developed and is currently being tested by end users. Options: CI/CD for visualisation applications, usage of Fenix interactive computing service, archival data storage and Fenix AAI.
Requirements	Interactive access through batch or cloud access models, integration of Fenix AAI and FURMS, support for VM based services for Swift API.
IS and PS dependencies	SCC, IAC, VM, DEV, CON, SWD

5. Building Blocks for Infrastructure Federation Services

FURMS and Fenix AAI are two services that serve as building blocks for a federated e-infrastructure for HPAC IS and PS. Together, these two manage the identification and access of users as well as the allocation and usage of resources for users and projects. Key details of these two critical services are outlined in the following sections.

5.1 Fenix User and Resource Management Service (FURMS)

The Fenix User and Resource Management Service (FURMS) is a one-stop-shop for users and administrators of the infrastructure, dealing with membership management, resource allocations, and accounting. Key definitions and roles are being introduced as part of the technical specifications of ICEI. These include "Fenix Community", "Fenix User", "Fenix Site" and details on "Resources". The resources, which are largely consumable, are published in calls for proposals for ICEI resources.[6]

The high-level objectives of the FURMS service are:

- Integration with the central Fenix Proxy IdP
- Generic attribute provider, including policy document management and SSH key management
- Membership management service
- Resource allocation to communities and projects
- Accounting
- Interaction with sites
- Documentation
- Fenix branding
- Deployment

Core concepts are highlighted in the Fenix architecture document [ICEI D3.1] and the FURMS concept [ICEI D4.15]. The envisaged interaction of FURMS with local sites elements is shown in Figure 6. The ICEI project is addressing two critical features for FURMS: development of workflows and attributes for HPC and cloud resources; and federation across Fenix sites. The resulting solution is designed to be extensible for adding other external IdPs and user and project management workflows.

[6] https://fenix-ri.eu/infrastructure/resources

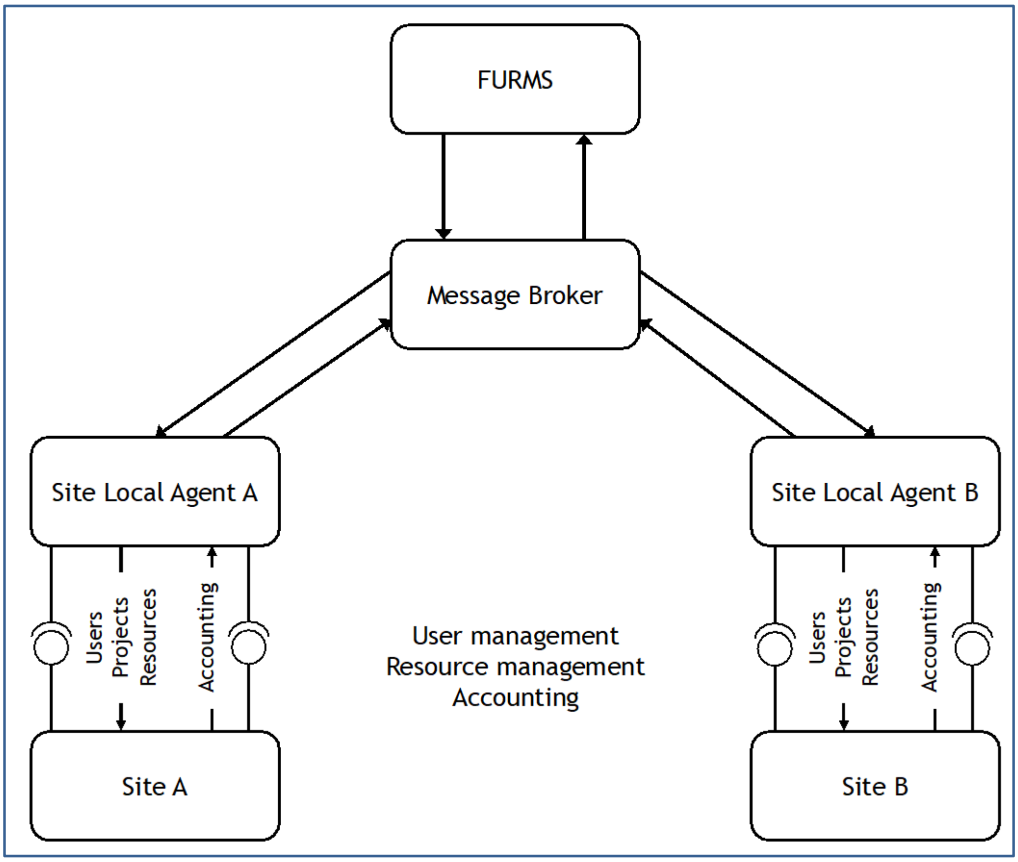

Figure 6: User and project registration and accounting and reporting workflows in FURMS and its interaction with local sites

5.2 Federated AAI Service

In order to provide access to Fenix infrastructure services, HPAC needs to rely on a robust and reliable Authentication and Authorisation Infrastructure (AAI). This assures a trustworthy environment where users can be allowed to access resources securely, as seamlessly as possible, using personal credentials. A user, registered on a “trusted” Identity Provider (IdP)[7] and allowed to consume a certain amount of resources via FURMS, will have seamless access to federated resources using his/her credentials (i.e. username/password, X.509, etc.) without going through a new registration process. For instance, a user registered at CSCS should be able to consume resources provided by JUELICH without registering twice and have the same user experience while accessing services. In other words, the experience that Fenix aims to provide is to allow users to conduct their experiments without having to adjust to different site access mechanisms.

Figure 7 shows the interaction of the different components included in the Fenix AAI architecture. For each user requesting to access the Fenix infrastructure, a new profile will be created on the Fenix Central IdP with a minimum set of attributes, including a permanent, opaque[8] Fenix identifier and a Fenix username. These attributes need to be returned to the SP as part of the authentication process, to permit any SP to identify the user and grant him/her access the service. Further details can be found in the Fenix AAI strategy document and handling of Fenix user agreements.

7 Any IdP part of the federation releasing common attributes for the user
8 https://indieweb.org/opaque

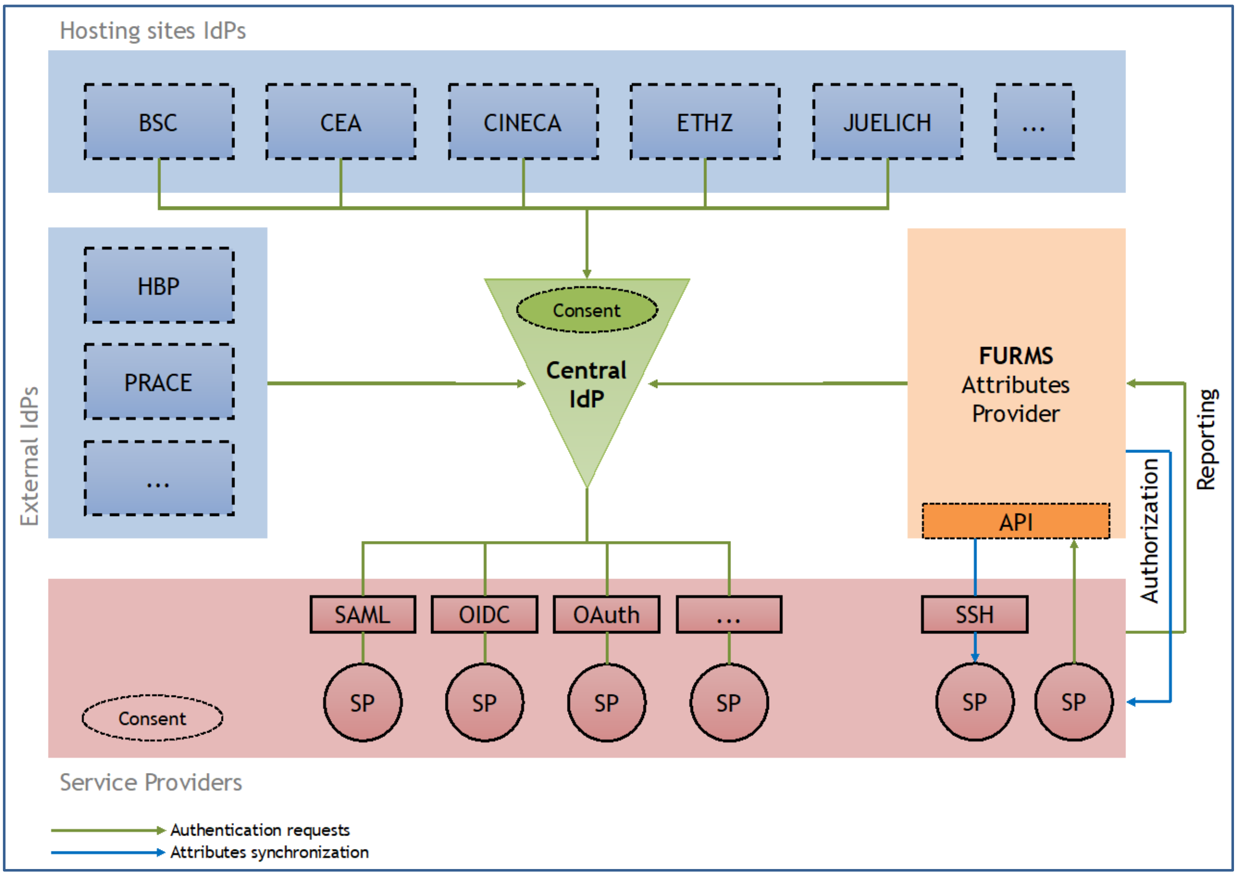

Figure 7: Fenix Authentication and Authorization Architecture for Federation

6. Consolidation and Mapping of HPAC services

Alongside the proposal of the architectural specification of the HPAC Platform, we analysed the HPAC Platform for existing infrastructure and platform services, with a goal of delivering robust operational services for users and developers of HBP community platforms. This analysis resulted in mapping and consolidation of existing services to proposed services which are categorised as follows:

- HPAC infrastructure to infrastructure service mapping (Table 9)
- HPAC Platform to platform services mapping (Table 10)
- HPAC Platform services without mapping (Table 11)

A proposal for transition and plan for migration is discussed in the next section. In some instances, for example, the scalable compute service (SCC), such planning is not required. However, the service must be integrated with Fenix AAI and FURMS for managing complex workflows that require federated identity and access management and accounting and reporting systems (Fenix AAI and FURMS).

Table 9: Consolidated Fenix and HPAC infrastructure services including production services that are necessary for realizing HPAC Platform concepts

HPAC convention	Name	Description	Component IDs	Proposed mapping
IS	HPC systems User managed virtual cluster on HPC systems through SLURM	The following High-Performance Computing (HPC) systems ("supercomputers") are available: JUWELS, JURECA and PCP Pilot: JURON (JSC) Piz Daint (CSCS) MareNostrum (BSC) MARCONI and Pico (CINECA) Irene / Joliot-Curie (CEA) SLURM/PCOCC extension to support container on HPC systems	C330, C331, C332, C333, C2825, C2829	IS-SCC
		This service concept is introduced in Fenix to manage requirements of HBP use cases. Details are available in ICEI D3.6.		IS-2 IAC
IS	VM services	Provision of virtual machines CSCS OpenStack cluster deployment	C565, C968, C1104, C1105, C2874	IS-3 VM
IS	Active data repositories	BCS, CEA, CINECA, CSCS and JSC ACD service	C410, C1108, C1118, C1127, C2838, C2839	IS-4 ACD
IS	Archival data repositories	BCS, CEA, CINECA, CSCS and JSC ARD service. Implementation details are consolidated in ICEI.	C409, C1109, C1119, C1129, C2840, C2841	IS-5 ARD
IS	HPAC network Site network service PRACE network Internet access	The secure high-speed network connects the HPC & data centres of SP7: BSC, CEA, CINECA, CSCS, JUELICH and KIT. These are the network services, which are required for data transfers from and to this site. PRACE network and internet connectivity at BSC, CEA, CINECA, CSCS and JSC.	C334, C1089, C1090, C1091, C1159, C1161, C1162, C1163, C1164, C1165, C1166, C1167, C1168, C1188, C1210, C1215, C1220, C1226, C2833, C2846, C2852, C2853	IS-6 NETE
		This service concept was implicitly defined in HPAC.		IS-7 NETI
		Concept introduced in Fenix to manage requirements of HBP use cases. Details are available in D3.6.		IS-8 DM
IS	Data Transfer Service	Services required for transferring datasets between the sites that are part of the data federation. DT service at BSC, CEA, CINECA, CSCS and JSC.	C416, C1113, C1126, C1131, C2858, C2859	IS-9 DT
	Data Location service	Central Data Location service providing basic functionalities for data discovery and access.	C2826	IS-10 DL
IS	HPAC Authentication & Authorisation Infrastructure Services Federated user database HBP LDAP Unity	HPAC Platform Authentication & Authorisation Infrastructure (AAI) provides single sign-on mechanism. The underlying Lightweight Directory Access Protocol (LDAP) accounts are linked to the user's central HBP account. A federated user database for the federated infrastructure. The LDAPs set up at the HPC and data centres of SP7 are used for the federated user management.	C336, C405, C572, C1085, C1086, C1087, C1171, C1173, C1175, C1176, C1177, C1179, C1187, C2832, C2851	IS-11 AAI

HPAC convention	Name	Description	Component IDs	Proposed mapping
		The Authentication and Authorisation Infrastructure (AAI) at BSC, CEA, CINECA, CSCS and JSC, which is integrated with the HPAC Platform-wide AAI services. https://www.unity-idm.eu/about/		
IS	HPAC Reporting and Accounting Service Site reporting and accounting service Resource usage DB and tools	Mechanisms for reporting and accounting of the resource usage. Accounting of the HPAC Platform access. Reporting and accounting services at BSC, CEA, CINECA, CSCS and JSC. Resource usage DB and tools at BSC, CEA, CINECA, CSCS and JSC.	C400, C571, C1093, C1094, C1096, C1097, C1149, C1150, C1151, C1152, C1153, C1154, C1155, C1156, C2834, C2849, C2854, C2855	IS-12 FURMS
IS	HPAC monitoring service	This services makes basic monitoring information of the systems available at BSC, CEA, CINECA, CSCS and JSC. Nagios at sites. Icinga is a resilient, open source monitoring and metric solution based on a new object-based, rule-driven configuration. Grafana is the leading open-source project for visualising metrics.	C556, C1080, C1081, C1082, C1083, C1133, C1142, C1143, C1144, C1145, C1148, C1193, C2831, C2856	IS-13 MONI
IS	HPAC security services PRACE security services Site security services	Taking care about the security aspects of the federated infrastructure. PRACE network security, intrusion detection and monitoring systems. Firewalls, encryption tools and intrusion detection systems at BSC, CEA, CINECA, CSCS and JSC.	C569, C1189, C1191, C1192, C1194, C1196, C1198, C1199, C1209, C1211, C1212, C1214, C1216, C1217, C1219, C1221, C1222, C1225, C1227, C1228, C2835, C2845, C2847, C2848	IS-14 SEC

Table 10: Consolidated listing of HPAC Platform services together with services in operation

HPAC convention	Name	Description	Component IDs	Proposed mapping
PS	Deployment of software across HPAC Platform Container-based software packaging and deployment technologies Docker Lightweight VM services	This service needs to be developed and introduced in accordance with Section 4.6 specifications. Providing support for container-based software packaging and deployment technologies (e.g. Docker, which can be considered as light-weight VMs) on production HPC systems. CSCS container services (Docker and Shifter). Provision of virtual machines for the users to setup their own services.	C562, C565, C1104, C1105	PS-1 SWD
	Common developer services	Support for common developer services such as bug tracking and source control, as well as continuous integration and development services via tools such as Jenkins for test-driven development.	C563, C1114	PS-2 DEV

HPAC conven-tion	Name	Description	Component IDs	Proposed mapping
PS	DB service	This service needs to be developed and introduced in accordance with section 4.2 specifications.	None	PS-3 DB
PS	Download service Storage management service	These services enable users to download datasets stored in the federated data infrastructure given that they have the necessary access permissions. Storage quota management service.	C415, C576	PS-4 DMGT
PS	HPAC monitoring services	Monitoring the status of the HPAC Platform infrastructure.	C566, C1080, C1081, C1082, C1083, C1133, C1142, C1143, C1144, C1145, C1193, C2831, C2856,	PS-5 MONP
PS	Visualisation as a service Development, extension and deployment of visualisation tools for HPAC Remote high-fidelity visualisation and interaction middleware Backend to establish visualisation and computing sessions for *in-situ* visualisation Integration of FENIX infrastructure into visualisation tools	Community building activities for supporting neuroscientists in the choice of appropriate visualisation tools and methods. Support, extend, develop and deploy neuroscience-specific visualisation software in the HPAC Allows remote, high-fidelity visualisation and data interaction to support knowledge extraction providing both, server and client infrastructure. Backend infrastructure will enable users to establish computing and visualization sessions to run simulations and perform *in situ* visualisation on top of the tools to be developed in T7.3.9. Integration of visualisation tools to be able to visualise distributed data using the FENIX Infrastructure.	C2681, C2682, C2683, C2684, C2708, C2860, C2861	PS-6 VIZ

Table 11: Existing HPAC Platform components with no mapping to new set of infrastructure and platform services

HPAC conven-tion	Name	Description	Component IDs	Comments
PS	UNICORE[9]	UNICORE is a general-purpose federation software suite providing seamless access to compute as well as data resources. Installation at BSC, CEA, CINECA, CSCS and JSC.	C263, C1184, C2857, C1183, C1180, C1181	This service can be replaced with Fenix AAI-compliant REST interfaces. Sites can continue to provide this support, but this will not be mandatory to instantiate HPAC Platform at Fenix sites.
PS	UNICORE Workflow Framework	The workflow framework enables the user to define and run workflows.	C407	
PS	UNICORE clients	Clients and client libraries for UNICORE services	C791	
PS	HPAC UNICORE services	The UNICORE service is installed at various sites. It is accessible via different clients, including a REST API.	C792	
PS	UNICORE/X + XUUDB	BSC, CEA, CINECA, CSCS and JSC installation of UNICORE/X	C1141, C2850, C1136, C1134, C1135	
PS	APIs to access key technologies for interactive data analysis	Report on key technologies supporting fast data access and retrieval for data interactive analysis and visualisation.	C2823	These requirements are merged in the proposal for HPAC Platform services (IAC and VIZ)
IS	SLURM plug-in for the co-allocation of compute and data resources	SLURM plug-in for the co-allocation of compute and data resources.	C2824	This component is still in development within WP7.2. It will not become an explicitly shown service as it will be deployed to manage infrastructure services like SCC, IAC and ACD

7. Recommendations for Transition and Consolidation of Services

The first Fenix e-Infrastructure services are available, allocated and in use by HBP users and customers. Hence, service continuation is one of the critical requirements for preparing plans for transition and consolidation. The following terminology will be used to define the status of a service:

- **Pilot**: Service will be introduced to early access users (by invitation).
- **Deployment**: Service will be deployed to all users (with limited support).
- **Production**: Service fully in production with uptime and user support.
- **Frozen**: Service is in maintenance mode.
- **Deprecated**: Service is deprecated.

[9] https://www.unicore.eu/

Two representative services are considered in this document to demonstrate the steps needed to plan for consolidation. One is AAI, where existing HBP services such as an identity management system and workflow are in place. Another comprises the archival data repositories (ARD), for which the Fenix project agreed on a standard, open format and API (OpenStack Swift). Appendix 9.1 Adapting HBP AAI workflows to Fenix AAI concepts provides details on the transition of HBP identity management to the Fenix identity and access management system. This service is currently in the **Pilot** phase. At the same time, we are comparing and contrasting existing HBP workflows to identify dependencies and integration points with Fenix AAI. The next step will be **deployment** of the services. Both Fenix and HBP identity and access management (AAI) services therefore will continue to operate until **production** readiness has been established. Training and outreach events will be planned to ensure a smooth transition. In addition, the transition phase will involve preparing a schedule when the existing HBP service is **frozen**. Once the full transition is completely validated and tested in operations and at scale, the HBP service will be **deprecated**.

The introduction of the ARD service will have a significant impact on multiple infrastructure and platform services. Any service that would require or enable federation of data within the Fenix AAI and FURMS framework for identity and access management would need to be adapted according to the specification of ARD service. This service is also in the **Pilot** phase (used by limited platforms like the HBP Collaboratory), even though the technology itself is available in production by selected vendors. Infrastructure services such as data mover (DM), data transfer (DT) and data location (DL) services are including requirements to support Fenix AAI-compliant Swift API. Likewise, platform services such as visualisation services (VIZ), which currently support a POSIX compliant file system such as Lustre and GPFS, need to be migrated. This would mean development of the web facing interfaces to a Fenix AAI compliant Swift API. Workflows for applications that use SCC, IAC and ARD as well as VM and ACD would need to integrate DM pipelines and update related applications. The deployment phase of this service is highlighted in Appendices 9.2 and 9.3, with the neuroscience workflow using the archival data repositories.

In short, the HBP SGA2 SP7 and ICEI projects have ongoing integration activities using Co-Design Projects as demonstrators to validate functional and performance requirements for transitioning from pilot to deployment and full production phases.

8. References

[ICEI D3.1] ICEI Deliverable 3.1, "Common Technical Specifications" (confidential)

[ICEI D3.2] ICEI Deliverable 3.2, "Initial Federated AAI Technical Specifications" (confidential)

[ICEI D3.6] ICEI Deliverable 3.6, "Scientific Use Cases Requirements Documentation"
https://fenix-ri.eu/sites/default/files/public/file-uploads/ICEI-D3.6-v3.1_clean.pdf

[ICEI D4.9] ICEI Deliverable 4.9, "Data storage and compute provisioning during M1-M12" (confidential)

[ICEI D4.15] ICEI Deliverable 4.15. "Tender documents for AAI, FURMS and Data Mover services" (confidential)

[ICEI D5.1] ICEI Deliverable 5.1, "Report on User Allocation Specification Mechanism"
https://fenix-ri.eu/sites/default/files/public/file-uploads/ICEI_D5.1_V2.0.pdf

[Fenix-UA] Handling of Fenix Usage Agreements, ICEI internal document, 3.6.2019 (submitted as part of ICEI Deliverable 6.2)

[PRACE-NET] PRACE network and services
http://www.net.prace.fz-juelich.de/

[ECP-SW] ECP Software Technology Initiatives
https://cdn.fs.pathlms.com/G4m3kebCR76rjSVmLLdn

[ECP-E4S] Extreme Scale Software Stack
https://e4s-project.github.io/

9. Appendix

9.1 Adapting HBP AAI workflows to Fenix AAI concepts

The Fenix AAI is not an extension of the HBP AAI. The Fenix AAI pursues the more general objective of federating and supporting both HPC and cloud services across Fenix sites. The resources provided through these services are made available not only to HBP but also to PRACE and national programs. These include federation of compute and data services including HPC and cloud technologies, and extensibility and delegation to external identity providers (IdPs). The Fenix Central IdP is able to support proxy authentication requests coming from external IdPs, such as the HBP OIDC server or the eduGain federation. However, an existing user of HBP OIDC does not automatically become a Fenix user. The Level of Assurance (LoA) is considered “Medium” and therefore additional, site-specific authorization steps may be required to consume Fenix resources. Steps will be taken on a service by service basis, for instance, the Collaboratory, which will be managed by individual platforms. There is a distinction between users registered at hosting sites’ IdPs (high LoA), which already have a UNIX account, and those coming from external IdPs, including HBP, which may not have one. No action is needed if a high LoA entry exists. In the absence of such entry, one of the Fenix sites will initiate the authentication and mapping process. Necessary workflows will be created for HBP OIDC users for accessing Fenix resources. A workflow for adapting HBP AAI to Fenix AAI concepts are shown in Figure 8.

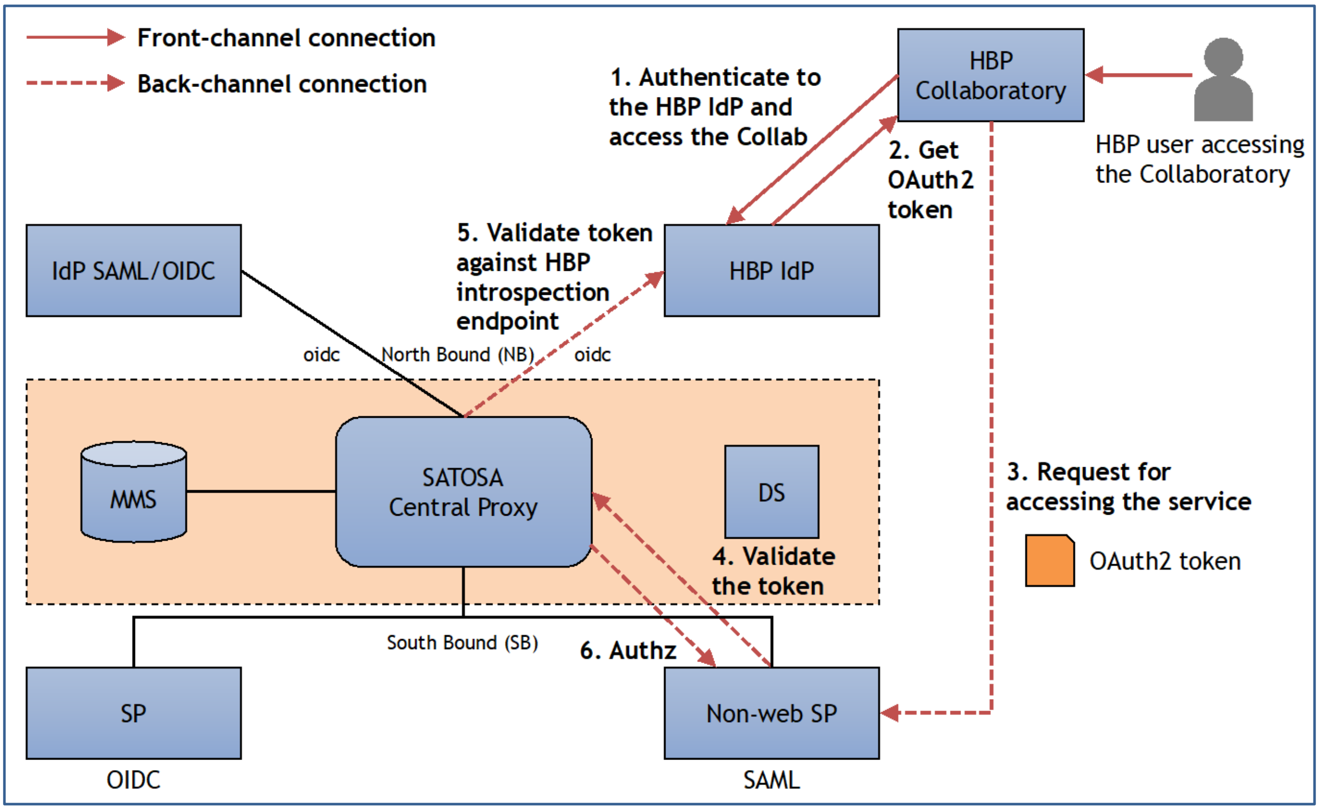

Figure 8: Proposal for OAuth2 token introspection and validation for HBP IdP

9.2 Neuroscience Workflows using Archival Data Repositories

Implementation of HPAC data services on the Fenix infrastructure is guided by the co-designed efforts with use cases. Neuroscience workflow use cases in particular highlight how distributed data services will be orchestrated using the Fenix infrastructure services. Such a use case demonstrator is shown in Figure 9. The workflow shown includes production ready simulations, optimization and analysis of a biologically realistic neuronal model of a single cortical micro-column. The workflow will be paused, saved to long-term storage, retrieved back to active storage and restarted. The dynamics at the start and end of the optimization will be compared with the Elephant toolset. Details of the use case can be found in [ICEI D3.6]. Local and distributed data services i.e. ACD and ARD will be utilized by the applications.

While ACD (typically referred as the scratch file system) requires no modification of the application, ARD does require using an interface called Swift. Swift API support will be provided by all Fenix sites, which is integrated with Fenix AAI. In this specific example, applications that interface with ARD such as the in-situ visualization, will need access to the Swift storage. This work will enable HPAC data transfer services using the Swift API and access protocols.

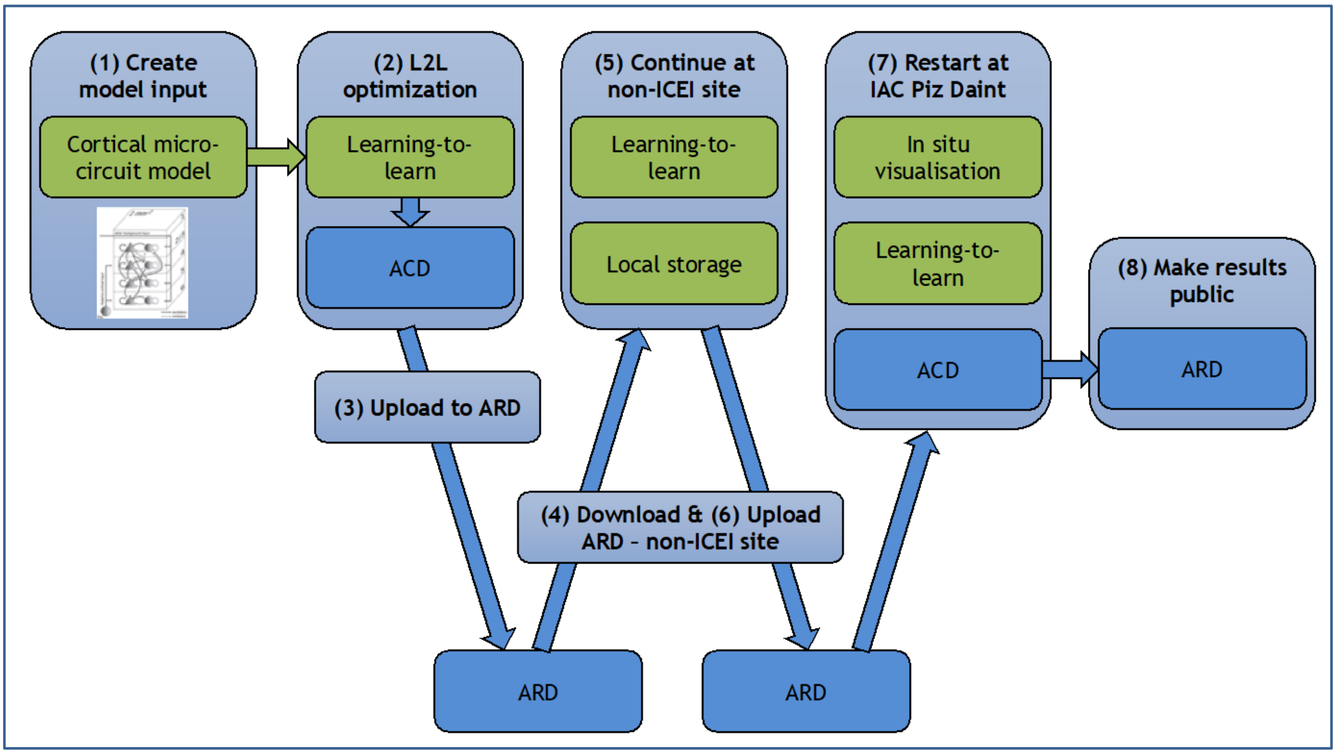

Figure 9: Workflow and infrastructure dependencies for the Learning-to-learn (L2L) workflow

Overview of data flows and components in the demonstrator. Green boxes have been tested and will be delivered as part of the existing roadmap for these components. The blue arrows and boxes are part of ICEI specific functionality and services.

9.3 Visualisation Services Web Interface using Archival Data Repositories

Interactive Computing in the context of ICEI refers to the capability of a system to spawn interactive sessions on HPC systems nodes or on nodes in the vicinity of an HPC system. Note that a classic HPC system employs a batch oriented scheduling system. Such interactive capability would, e.g., enable on-the-fly monitoring and steering of distributed computing workloads by the user. A typical usage scenario for interactive computing regarding the visualization, processing, and reduction of large amounts of data, especially when the processing part cannot be standardized or implemented in a static workflow.

We anticipate the following usage scenarios for interactive access:

1) User connects to ICEI hardware resources via an Interactive Compute Service using secure protocol to, e.g. analyse and visualise data produced by simulations that have been running on Scalable Compute Services, i.e. on a large-scale HPC system. The data may either be directly streamed to the Interactive Compute Node or be staged through an Active Data Repository. This usage scenario stays within a controlled HPC environment. (See Figure 10 for a schematic view on this usage scenario.)

2) User connects through a web protocol to a service running in a cloud environment to analyse and/or visualise data that is made available through an Archival Data Repository. The latter is accessible from both, within an HPC as well as a cloud environment. Optionally, a direct communication channel between, e.g., a simulation running on an HPC system and a cloud service for direct streaming of data for visualisation of steering of the simulation is envisaged. (See Figure 11 for a schematic view on this usage scenario.)

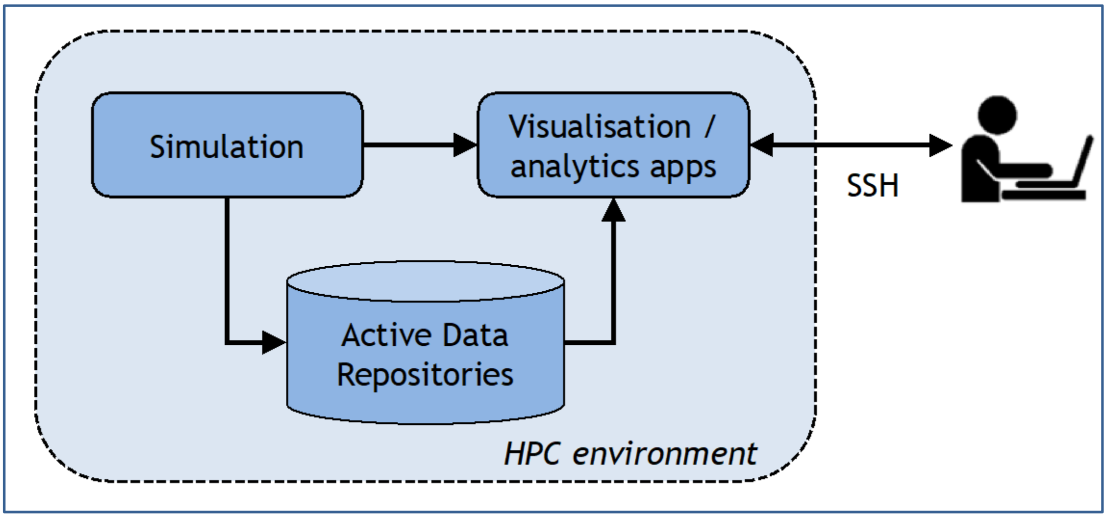

Figure 10: Interaction with the system via an SSH client

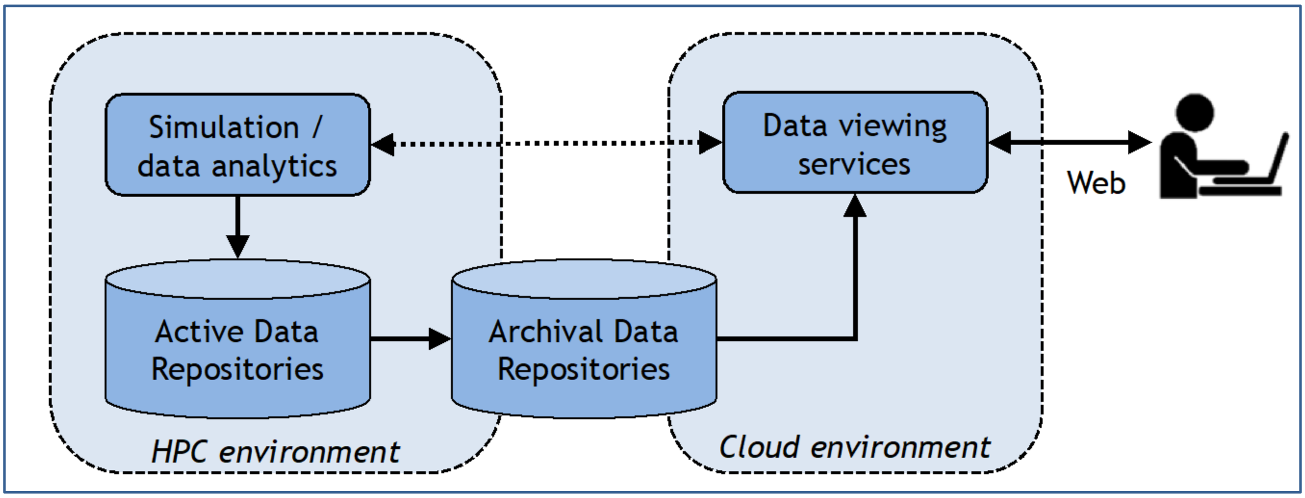

Figure 11: Interaction with the system via a Web application

We envisage the Interactive Compute Services to be provisioned through a set of components, clients and servers, that are part of an HPC cluster, i.e. clusters that provide both Interactive Computer Services as well as Scalable Computer Services, and Cloud services will be running on virtual machines provisioned through an OpenStack instance.

Towards enabling these usage scenarios, the R&D development effort as part of the ICEI project will focus on realisation of supported solutions through R&D services that address several of the following needs:

1) Fast and easy access to Interactive Compute Services through servers that are part of an HPC cluster.
2) Optimise the use of available compute and storage resources including support for efficient suspend-resume mechanisms for HPC jobs running on Scalable Compute Services to facilitate releasing hardware resources for Interactive Compute Services.
3) Easy-to-use tools for efficient remote visualisation through the Interactive Compute Services.