Fenix Strategy Document

(VERSION FOR PUBLIC USE)

Date	Comments	Status
30.10.2021	Aligned with internal version	Final
11.11.2021	Update on the membership evolution	Final

Content

1. Mission statement

Fenix is a collaboration of HPC centres working on the harmonisation and federation of their offerings of e-infrastructure services with the goal of supporting a variety of science and engineering communities. This service portfolio's distinguishing characteristic is that different types of data repositories, scalable supercomputing systems, and private cloud instances are in close proximity and thus well connected and integrated. The different sites that act as e-infrastructure services providers are interconnected via a high-speed network.

Fenix has the ambition of serving in a sustainable manner relevant science and engineering domains that strongly benefit from diverse e-infrastructure services for their collaborative research. For being able to scale to a larger number of such domains, Fenix focuses on a consolidated portfolio of services. To stay aligned with the needs of current and upcoming science and engineering domains, Fenix governance foresees a representation of these domains such that they can drive the evolution of the e-infrastructure services portfolio.

Fenix members leverage national, European and international funding programs to realise the compute, storage and network resources sustaining the e-infrastructure services. The coherence of the approach is ensured through a clear governance model and close technical collaboration at various levels. Furthermore, the Fenix partners share the responsibility for the operation of federation-level services and the integration at each of the sites.

2. Approach

The architectural approach of Fenix is based on the general consideration that a separation between an e-infrastructure service layer and a platform service layer (see Figure 1) is beneficial as it allows to separate concerns[1]. The platform services layer encompasses services that are typically specific for a given research or engineering domain. They are not necessarily useful for other domains or would require significant adaptations. A typical example is web-based portals: while such portals are needed for almost any research infrastructure, their organisation is highly domain-specific. The infrastructure services layer includes a set of services that allow implementing these platform services and are sufficiently generic for being useful for different domains. One example is machines for deploying any of the aforementioned portal services, typically offered in a virtualized environment. The infrastructure services are organised such that they can be provided by multiple, geographically distributed resource providers.

The architectural approach of Fenix has the following benefits:

- It allows providing a consolidated and diverse hardware infrastructure, thus serving a larger number of science and engineering domains. Leveraging scale economy, hardware and operation costs are reduced. The cost savings can be invested in platform services, which are the key to ensure the productivity of the end-users, i.e. the scientists and engineers, given the increasing heterogeneity and diversity of the hardware architectures offered.
- A service portfolio that tightly integrates different types of services including scalable computing services on leading-edge supercomputers will facilitate implementing more complex workflows.[2]
- Harmonising and federation of e-infrastructure services, which include both compute and data services, provides science and engineering domains with opportunities to optimise for data locality, to scale resources if needed and to improve sustainability as services could over time be migrated to other e-infrastructure resource providers.

Fenix's approach to federation is lightweight and realised by a relatively thin layer of federation-level infrastructure services. It includes most notably federated authentication and authorisation services that allow users, e.g., to use e-infrastructure services at different sites using the same identity. Attribute services allow the management of users, e.g. by grouping them in communities and assigning them specific roles, as well as to realise attribute access control models. The federation-level services layer furthermore comprises services for managing resources. They allow them to assign resources to specific projects and track the consumption of these resources centrally even if projects have been granted access to resources at multiple sites.

Fenix endorsed the "European Charter for Access to Research Infrastructures" [ec2016] and will support several of the access policy principles described therein, i.e. excellence-driven[3], market-driven and wide-access models. In its role as e-infrastructure services provider, Fenix aims at supporting different access policies and resource allocation mechanisms while avoiding to be part of the decision-

[1] Such a layered approach is commonly used for creating Cloud infrastructures (see, e.g., [nist2011]), where the terms Infrastructure-as-a-Service (IaaS) and Platform-as-a-Service (PaaS) are widely used.

[2] For instance, workflows considered by the Transcontinuum Initiative [etp4hpc2020].

[3] This is also known as the peer-review mechanism. The PRACE peer-review principles are described here: https://prace-ri.eu/hpc-access/project-access/project-access-the-peer-review-process/. A more generic description, covering different access modes can be found here: https://prace-ri.eu/about/introduction/

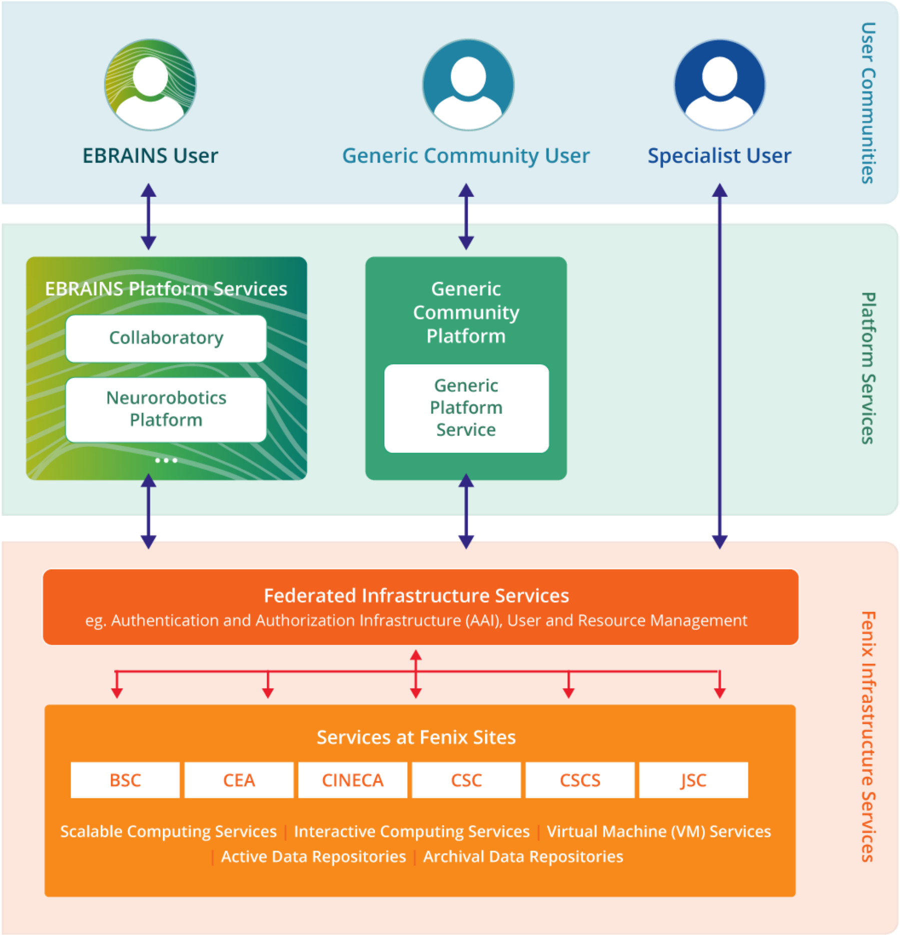
User Communities
EBRAINS User
Generic Community User
Specialist User
Platform Services
EBRAINS Platform Services
Collaboratory
Neurorobotics Platform
...
Generic Community Platform
Generic Platform Service
Fenix Infrastructure Services
Federated Infrastructure Services
eg. Authentication and Authorization Infrastructure (AAI), User and Resource Management
Services at Fenix Sites
BSC
CEA
CINECA
CSC
CSCS
JSC
Scalable Computing Services | Interactive Computing Services | Virtual Machine (VM) Services
| Active Data Repositories | Archival Data Repositories

Figure 1: Schematic overview of Fenix's layered approach for support users from different science and engineering domains.

making itself, in particular in the case of excellence-driven access models. In this case, resource allocation can be delegated to independent organisations like PRACE, to science and engineering domains [4], or to platform service providers.

A first overview of the Fenix e-infrastructure service portfolio including federation-level services have been realised through the ICEI project [alam2019, alam2020]. The portfolio currently includes the following services:

- **Interactive Computing Services**: Quick access to single compute servers to analyse and visualise data interactively, or to connect to running simulations, which are using Scalable Compute Services.
- **Scalable Computing Services**: Massively parallel HPC systems that are suitable for highly parallel brain simulations or for high-throughput data analysis tasks.
- **Virtual Machine Services**: Service for deploying virtual machines in a stable and controlled environment.
- **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.

[4] In the context of the ICEI project, Fenix is providing the Human Brain Project (HBP) with programmatic access and supporting the BRAINS research infrastructure running an excellence-driven access model.

- **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.

3. Science and Engineering Domains

The e-infrastructure services provided by Fenix are designed to accommodate a large diversity of data sources and relevant workflow processing needs. Fenix aims to deliver a flexible set of federated e-infrastructure services as building blocks to enable specific communities or science organizations to develop and operate their domain-specific platform services on top of these services. Specifically, communities that make use of a diverse set of data sources with specific formats, modalities, spatial and temporal scales, coverage, sample sizes, etc. can leverage Fenix resources for operating, developing and scaling their platforms and community services. Data may originate from different experiments or computing systems, where no fixed or predetermined relationship exists between the sources. Furthermore, data analysis capabilities may range from general-purpose computing to distributed, high-throughput computation to large-scale capability simulations.

As an illustrative example, we consider the Human Brain Project (HBP) and the EBRAINS Research Infrastructure (RI), which is being built by the HBP. The current architecture of the Fenix infrastructure has been driven in part by requirements of the neuroscience community as represented by the HBP. It has been developed and implemented within the Interactive Computing E-Infrastructure (ICEI) project, which is co-funded by the European Union under the umbrella of the HBP Framework Partnership Agreement (FPA). The HBP is using Fenix e-infrastructure services created within the ICEI project as the basis for EBRAINS, a digital research infrastructure that gathers an extensive range of data and tools for brain-related research.[5] The EBRAINS RI includes domain-specific platform services for storing, finding and sharing research data, computational models and software, for navigating multi-modal, high-resolution brain atlases, for simulating brain models across multiple scales, for interactive access to brain-inspired technologies such as neuromorphic computing and neurorobotics, and for analysing medical data for diagnosis and research in clinical neuroscience.

3.1 Responsibilities of science and engineering communities

While Fenix provides general-purpose, distributed e-infrastructure services and some common services enabling the federation (e.g., a common authentication and authorisation infrastructure), it is generally the responsibility of the science and engineering communities to build their own domain-specific platform services to utilise the Fenix e-infrastructure resources. Such domain-specific platform services are, e.g., web-based portals through which access to domain-specific databases or simulation services is provided. They may also include additional services to complement or integrate Fenix services according to domain-specific needs (e.g., metadata services or data location services).

Fenix strives to support, within certain limits, the storing and processing also of sensitive data that fall under the European General Data Protection Regulation (GDPR) [eu2016]. Within the ICEI project, a set of regulations and a procedure governing the processing of personal data on the Fenix infrastructure (as required, e.g., by the HBP) have been established.[6] These may evolve in the future to accommodate more data classes. In general, a user uploading personal data to the Fenix infrastructure has the role of the Controller as defined by the GDPR and retains all responsibilities associated with that role (such as protecting the rights of the data subjects, managing consent forms etc.). The Fenix service providers, who will generally be in the Processor role, will support the Fenix users to comply with their obligations as Controllers under the GDPR (e.g., by implementing appropriate technical and organisational measures; see also section 4.3). Details of the respective

[5] https://ebrains.eu/

[6] https://fenix-ri.eu/content/processing-personal-data

roles and responsibilities of the Fenix service providers and the users from science and engineering domains with regard to the processing of sensitive data may be part of a service level agreement.

Science and engineering communities may use Fenix e-infrastructure services to the extent that the costs of their provisioning are covered by the community directly or by third parties, e.g., funding organisations, through fees, co-funding of resources, or by contributing infrastructure resources to Fenix (see also section 6).

3.2 Control by science and engineering communities

Fenix foresees different mechanisms to allow science and engineering domains to impact the evolution of the e-infrastructure services and the underlying resources as well as to exercise control on the operation of these services and resources.

One way in which science and engineering communities may exercise control is to establish service level agreements (SLA) between organisations representing such a community and Fenix e-infrastructure service providers. SLAs can, e.g., define the quality and availability of services as well as the responsibilities of the parties, may include a set of suitable KPIs to monitor the provided services and their federation, and may define security standards for the provisioned services.

Communities may also control resource allocation. Fenix has in the past been in the position to provide science and engineering communities with programmatic access to a share of the available resources. The responsibility for distributing that share within that community has been delegated to the community. Users from the community apply to an entity determined by the community for Fenix resources, which are made available by the Fenix resource providers. The science and engineering communities are responsible for managing the consumption of their resources and for ensuring that the resources allocated to users produce the expected scientific outcomes. Fenix is practising such an access model already for the excellence-driven and wide-access case, but may in the future also allow for a market-driven access model.

Finally, science and engineering domains will impact the evolution of Fenix through the Fenix Council, where communities will have a consulting role in particular with respect to the definition and updating of current and future infrastructure requirements (see section 5). The Fenix Council will be tasked with defining a regularly updated scientific case as well as requirements, which will allow to evaluate whether the realisation and evolution of the Fenix infrastructure meets the needs of relevant science and engineering domains.

4. Joint technical activities

This section describes a selection of technical activities, on which the Fenix partners work jointly and which are considered to be of strategic relevance for further development and enhancement of Fenix.

4.1 Specification of architecture and services portfolio

Development and evolution of federated IT e-Infrastructures has been driven by ever increasing and diverse computing, data processing and networking requirements of scientific workflows across multiple domains. Enabling a convergent use of HPC, cloud, data, and AI have become a common, overlapping theme across scientific communities. The Fenix architecture has been guided by the principle of leveraging inherent capabilities of leading European HPC centres to satisfy IT e-Infrastructure requirements across domains. A lightweight federation layer therefore augments the Fenix e-Infrastructure resources to facilitate uniformity of portfolio of services and identity and access management solutions across all Fenix partners. To this end, users of Fenix can expect a single sign-on, self-service experience across multiple service providers. At the same time, the Fenix architecture ensures site autonomy for integration and life cycle management of technologies by leveraging open standards.

Co-design of e-infrastructure services together with the requirements of the research infrastructure platforms has been and continues being the key method for specifying and evolving the architecture and portfolio of services. For instance, the first project realising the Fenix architecture, ICEI, focuses on designing, building and operating a federated e-Infrastructure for the Human Brain Project (HBP) and other science communities driven by scientific use cases with the following key characteristics:

a) Need for sharing data in a context of diverse and distributed data sources
b) Need for deploying platform services like portals (e.g. HBP Collaboratory) or metadata services (e.g. HBP Knowledge Graph)
c) Need for realising workflows that involve, among others, HPC services

The resulting architecture and technology choices for implementation across Fenix sites are largely driven to support these requirements. One of the distinguishing characteristics of the consortium include unprecedented levels of parallel, high performance computing and data processing capabilities, both locally at sites and within the federation. This in turn enables scientists and platform developers to exploit the full spectrum of HPC and cloud technologies thereby fulfilling the convergent use of HPC, cloud, data, and AI with a uniform, consistent portfolio of services across the Fenix sites. Altogether, the portfolio includes **Scalable Computing** (SCC), **Interactive Computing** (IAC), **Virtual Machine** (VM), **Active Data Repository** (ACD), and **Archival Data Repository** (ARD) services. The storage and data services architecture, in particular, addresses the interoperability between the cloud native and HPC ecosystems to fulfil requirements for inter- and intra-site data sharing, platform deployment and complex workflow management.

Complementing the service portfolio, the unique feature of Fenix has been the development of solutions for identity and access federation and convergence of HPC and cloud technologies. The following services have been introduced to realise the Fenix architecture:

a) Fenix Identity and Access Management (IAM) services
    i) Fenix AAI (authentication and authorisation infrastructure)
    ii) Fenix users and resource management service (FURMS)
b) Data mover service (interfacing HPC and cloud storage systems)

All along, standard and open source technologies have been leveraged. Careful consideration has been given to ensure that technologies enabling federation and HPC-cloud interoperability can be securely integrated into 24/7 operational HPC facilities. The service development process therefore takes into account feedback from the Fenix partners, their operational requirements, and most importantly, engagement of technical staff that are responsible for the day-to-day operation of services. Hence, for the most part, technology readiness level (TRL) 7 or higher are considered for ensuring a consistent, high quality user experience.

4.2 Identity and access management

Identity and access management (IAM) plays an important role in enabling the federation of the Fenix resource providers. Proper deployment of IAM tools and policies allows end-users to access the Fenix offerings securely and in a unified manner, regardless where the services are deployed. The design principles of the Fenix IAM layer include ease of use and consistency for the end-users, standards-compliant, secure and scalable while maintaining site autonomy for site-specific integration. Fenix IAM related processes are designed to be automatable by the member sites. Manual site-specific processes for on-boarding Fenix use cases are discouraged, except in cases where they are required by local regulation. The main scope of the federating layer is to enable the site integration of the Fenix resource providers. Communities building their domain-specific, platform services utilizing the Fenix e-infrastructure are responsible for the platform IAM, which is generally considered out of scope for the Fenix IAM.

The IAM for Fenix consists of two separate parts: identity management and access management. There are established ways for doing federated identity management. Fenix therefore aims to delegate the majority of the identity management to GÉANT whose core competence comprises deploying identity management solutions, and further driving the standardization of policies and technologies for the wider academic and research communities and beyond. The identities could be used more widely, and not limited to Fenix. This allows Fenix to reach the widest audience, in a secure manner, and simplifies the end-users handling of identities. Fenix e-infrastructure service providers retain control of critical aspects such as the Level of Assurance (LoA) of the end-user identities. With collaboration between Fenix, other initiatives, identity providers, and identity federations, Fenix also aims to develop best practices for secure access to a diverse portfolio of resources.

Fenix has significant expertise in access management, therefore the majority of the Fenix IAM efforts have been focused on the development of federated authorization, accounting and reporting services. Fenix is working on the FURMS service, which is responsible for the access management. FURMS will be in the core of the resource integration in Fenix, and is planned be continuously developed to enable the growth of the service portfolio, and support new service offerings. Like identity management, it is expected that some workflows of access management can be delegated to service providers like GÉANT for identity management as this will allow scalability and growth to multiple usage scenarios.

Projects are a core concept in Fenix, and Fenix project data is centrally available to Fenix resource providers. Resources in Fenix are granted to projects. Aggregated views of resource allocation and consumption are supported, for instance, for all projects that belong to a community like HBP. Projects may be granted resources on one or multiple resource providers. Projects are managed by principal investigators who are also responsible for project membership management. User access to individual services by the resource providers is controlled by the project resource allocations. The project information is standardized, and available through APIs. This facilitates site-specific implementations. Resource usage is reported in a standardized way via APIs.

From the access management perspective, granting projects, and handling calls for resources can be organized in multiple ways as long as the resulting flows support an association of projects with resource allocations. This enables Fenix service providers to properly map and control access to the resources. The Fenix IAM layer is a technical layer for enabling federation and integration. Resource discovery, allocation policy, cost recovery and payment of service use are out of scope of the access management layer. Integration of end-to-end pipelines from resource discovery to usage reporting can be supported on a case-by-case basis through the high-level access and policy control alignments.

4.3 Ethics compliance

The Fenix partners closely collaborate to ensure compliance of the Fenix infrastructure with Ethics requirements. This concerns in particular topics such as data protection, dual-use and environmental impact.

Data protection: For communities such as HBP/EBRAINS there is a need to store and process data that represents personal data according to the European General Data Protection Regulation (GDPR). The extent to which Fenix can support the storing and processing of personal data varies between the Fenix sites due to different national, regional and institutional regulations. A set of security measures providing a baseline of security for storing and processing of personal data using the Fenix infrastructure has been documented in the Fenix Security Measures Catalogue [fenix2020]. The first version of this document includes measures that have to be implemented by the user (data encryption in transit and at rest, use of secure data repositories) as well as a Compliance Catalogue for Secure Data Storing to be implemented by the Fenix sites. Fenix has the goal to further improve the security of its services in the future, to enable the storing and processing of personal data beyond what is offered today. Fenix supports users working with personal data in meeting their legal obligations as Controllers under the GDPR, including, e.g., the necessary technical support for data encryption.

Dual-use: Cutting-edge computing and storage systems as provisioned by Fenix can potentially be exploited for non-civilian purposes and therefore fall under the class of dual-use items. There is, e.g., a risk that Fenix resources are abused for purposes other than stated in a user's resource application.

The Fenix approach to mitigating this and other dual-use related risks is based on established standard practices of supercomputing centres, which include in particular the requirement for users to sign one or more usage agreements.

Fenix partners are working together on protecting Fenix resources and services and on improving security as necessary to meet the obligations resulting from risks related to dual-use and data protect requirements (see section 4.4).

Environmental impact: Supercomputers are known to consume large amounts of energy, thus contributing to CO2 emissions and the corresponding impact on the climate. Since energy is also a significant financial cost factor, the supercomputing centres generally have a strong interest in energy-efficiency. To make supercomputing as sustainable as possible, the Fenix partners follow holistic approaches including changes to system architectures, cooling, building design, operations at the centre, application design, and waste heat reuse. Fenix partners support initiatives [7] that aim to track improvement in systems energy efficiency and in systems carbon footprint.

4.4 Security

Fenix aims to provide a harmonised offering of e-infrastructure services based on the major computing centres in Europe. These services need solid and common foundations to provide to users as well as science and engineering communities a unified global view of the distributed services. These foundations are built with all resources brought in by sites, with a common management and a common baseline security policy.

A joint approach of Fenix sites to security is mandatory as security issues on one site may affect the others. The coordinated efforts will also create benefits as they allow leveraging a broader knowledge and skill base and harmonise authentication and authorisation mechanisms from a user perspective, which is important to facilitate single sign-on and use of e-infrastructure services from multiple sites. In this context, the Fenix partners will pay particular attention to the impact of changing security measures on operation of platform services.

The Fenix sites plan to establish a security trust zone. This can be enhanced over time and will make the instantiation of new services much easier. Building this trust zone means that all sites must define and maintain a good common security architecture which needs a huge initial amount of work and further day to day work. First, common definitions and common practices have been defined based on self-assessments using, e.g., the Fenix Security Measure Catalogue [fenix2020].[8] These define the common security level that should be reached by all sites.

It is, furthermore, important that all sites jointly address security related issues while operating the e-infrastructure services. This includes, e.g., suitable communication mechanisms for informing about security incidents.

The security trust zone will be established using the following measures:

- Enforce security constraints for services running on this e-infrastructure with the help of best security guidelines, service security templates, etc. and security audits. Services audits will be required to assess the status of implementation of the agreed security measures. Currently, annual self-assessments are foreseen. In the future, audits by external organisations (e.g., commercial service providers or public security agencies) will be considered.
- Improve the level of security within the security trust zone: The Fenix Security Measures Catalogue will need to evolve and to account for tighter security constraints and changing user needs.

[7] As, for example, in the PPI4HPC project: https://www.ppi4hpc.eu/

[8] This document follows recommendations of the German Federal Office for Information Security [bsi2017] and French National Information Systems Security Agency [anssi2018].

- Improve collaboration on security-related topics between the Fenix sites: exchange on security architectures and concepts (this can give all a global security view of all sites), creating working groups and organizing security meetings twice a year.

Fenix acknowledges that defined security standards and the establishment of a security trust zone are important to establish confidence in using the offered e-infrastructure services. Security requirements will become a topic for service level agreements with organisations (e.g. EBRAINS) representing specific science and engineering domains.

4.5 User and platform service developer support

A key aspect of Fenix is to facilitate the transition of data centres from HPC system providers with a limited set of established end-users to data centres as e-infrastructure services providers with end-users that use these services natively, and science communities that use the e-infrastructure services for deploying domain-specific platform services. The latter includes developers as well as operators of platform services. Note that in the role of e-infrastructure services providers, Fenix is not taking responsibility for supporting users of the platform services.

Fenix partners, therefore, collaborate on facilitating the necessary support of the following categories of users:

- Users that make native use of e-infrastructure services
- Developers of domain-specific platform services
- Operators of domain-specific platform services

Towards that end, created in May 2020, the Fenix User Forum is an online platform which offers user access to specialised Fenix content and enables the interaction between Fenix and users as well as among users themselves. The objective is to create a Fenix user community for exchanging experiences, gathering feedback, facilitating trouble solving, and consequently improving the provision and utilisation of the Fenix services. Regular Fenix User Forum meetings as lone-standing happenings or within key events in the field of HPC also aim at engaging the Fenix user community, addressing their questions, and sharing experiences in using the Fenix infrastructure.

A comprehensive training strategy looks to address user and platform service developer and operator needs, to educate users of science and engineering communities in need of e-infrastructure services to leverage their research projects as well as platform services developers and operators in several aspects of the Fenix access, resources and usage, and to generate a pool of training resources for current and potential users to come back to, as needed.

These efforts include the Fenix webinar series, a set of webinars and hands-on tutorials on how to gain access to and make use of the Fenix resources, the services at the different Fenix sites, and research-specific examples that exploit the Fenix resources, and a portfolio of documentation dedicated to Fenix users addresses technical specifications and usage files. A post-training plan of surveys and Q&As offers the opportunity to follow up with the trainees and incorporate their feedback and suggestions in the development of the Fenix services. It also allows for Fenix experts to look into user issues with using the infrastructure on an individual basis in order to offer troubleshooting support and pass on any important matters to the Fenix consortium.

User support will be a continuous effort of Fenix not only in terms of guiding end-users on how to access and use the e-infrastructure services, but also of training developers and operators on how to build and facilitate their own domain-specific platform services for their users. Accordingly, the webinar series will be expanded to include sessions operated by platform developers for platform developers in order to explain how they have used the Fenix resources for their platforms and offer insights on how other developers can benefit from Fenix to improve their own platforms. In addition, Fenix will conduct a yearly summit dedicated to platform services developers and operators. This meeting will serve for the developers and operators to feed back to the Fenix Technical Board, so that the Fenix providers can improve and update the portfolio of the Fenix services, as needed, and an interaction circle between the two sides is established.

5. Fenix organisation

Fenix as an organisation is structured as follows:

- Fenix Executive Board
- Fenix Technical Board and Fenix Working Groups
- Fenix Council

The Fenix Executive Board comprises one representative per Fenix partner. The board acts as a decision-making body for all major aspects concerning Fenix.

Also, the Fenix Technical Board comprises one representative per Fenix partner. Board meetings are, however, open for other persons affiliated to Fenix partners without these having voting rights. This board manages the daily work of Fenix supported by Fenix Working Groups. The latter are set up by the Technical Board for specific technical topics. Meetings of the Working Groups may be open for external experts.

The Fenix Council brings together representatives of Fenix and science and engineering communities.[9] Representatives of the science and engineering communities are leading researchers, which are expected to have a good overview over their community and are willing to act as representative of their community. The members of the Fenix Council are appointed by the Fenix Executive Board. The Board aims for all communities to be represented, which Fenix is supposed to serve or which happen to be major users of resources provided through Fenix. A number of seats will be reserved for emerging user communities.

The Council has the following tasks:

- Provide advice on the evolution of the Fenix Infrastructure by means of science cases and requirements definitions;
- Provide support for establishing resource allocation mechanisms and service level agreements;
- Review status of the Fenix Infrastructure as well as resource allocation and utilisation.

While the Fenix Council have primarily an advisory role, it can exert a strong influence as Fenix Executive Board must at any point be able to explain, how it supports the science cases and requirements definitions formulated by the Fenix Council.

6. Future evolution

6.1 Cost recovery

The main costs of creating, deploying and operating the Fenix infrastructure can be grouped in the following categories:

- Total cost of ownership (TCO) of equipment
- Data centre and external network connectivity costs
- Development, deployment and operation of site-local as well as federation-level services
- Provisioning of support and training services

To ensure mutual predictability and minimise economic risks, Fenix aims for multi-year agreements that cover the delivery of a defined amount of compute and storage resources as well as the

[9] Note that the Fenix Council does have a different role than the PRACE Council. The latter organizes the PRACE members.

provisioning of services. For example, Fenix sites could receive project funding for providing services to specific user communities as was done for the Interactive Computing e-Infrastructure (ICEI) project, which the European Commission (EC) funded in the framework of the Human Brain Project (HBP). Alternatively, Fenix sites could agree on hosting and operating equipment, which is procured by third-parties as practiced for some of the systems that are deployed within the EuroHPC infrastructure.

6.2 Organisation and business model

For the time being, Fenix plans to continue following a collaborative model where infrastructure service providers commit to the provisioning of a jointly agreed set of e-infrastructure services and work on the federation of these services through a Memorandum of Understanding (MoU). The funding for the resources, the services as well as the support can come from different sources and is typically not part of a common budget.

Other options are being investigated to strengthen the legal basis for the activities, e.g., creation of a common not-for-profit association or company.

6.3 Membership evolution

Fenix is open for new members as it has the ambition to contribute significantly to the EuroHPC pillar on federation and secure service provisioning of supercomputing services and data infrastructures. There is, however, a need to balance the benefits of distributing resources and services, e.g. for optimising data locality, on the one hand and on the other hand the cost benefits achieved by consolidating deployment of hardware at a smaller number of sites. Fenix therefore aims to focus on resource providers that can be considered being large at European scale. New members must commit to supporting a significant fraction of the Fenix services. Furthermore, members must be able to support the ambition of Fenix to serve science communities that use the e-infrastructure services for the development of domain-specific platforms, and therefore must commit to supporting one or more user communities that rely on HPC.

6.4 Service portfolio evolution and service development process

Future development of the service portfolio will be co-designed and driven by requirements of scientific workflows and use cases with a particular attention to emerging user communities. In addition, the fast pace of development of cloud technologies will be reflected in the evolution of services. For instance, as part of the ongoing development of community platforms utilising Fenix resources, the container orchestration service enabled by the Kubernetes technologies has been requested by the platform developers that are currently users of the Fenix Virtual Machine (VM) service. When introduced, such service is expected to be fully compliant with Fenix IAM as well as user and project management guidelines.

In order to continue strengthening and evolving the portfolio of services for multiple scientific domains and workflows, research and development work will continue to enhance the lightweight layer of federation of identity and access management services. Already, Fenix IAM is fully compatible with the AARC2 blueprint architecture [aarc2019] with an important value-added feature for users namely the level of assurance and trust among the Fenix sites. FURMS can provide a similar enhancement for defining a standard schema and lifecycle for an e-Infrastructure allocation project. Such standardisation can go a long way to enable federation goals for a number of allocation and access schemes for research, academic, large-scale facilities (ESFRI) and industrial users and developers of platforms.

Alongside co-designing, evolving portfolio of services and underlying tools, compliance with standards with security, data protection and environmental impacts such as green vision for the 24x7 operational IT infrastructure will remain a priority within Fenix. There has already been ongoing coordination among sites and communities with respect to the policies as well as their evolution and enforcement in a uniform manner across sites. Hence, these community-driven leadership efforts will continue to grow and evolve as an integral part of the evolution of the service portfolio and its development process.

Another dimension for future exploration and development is establishing quality of service for federated access to resources, for instance, data transfer rates between Fenix sites and uptimes of services for migrating and load balancing high availability services. Competence in key areas like GÉANT for the networking capabilities and engagement with domain-specific platform development teams will be critical for augmenting and enhancing the portfolio in a co-design manner.

References

[aarc2019] GÉANT et al. (AARC project), "AARC Blueprint Architecture 2019," November 2019, doi: 10.5281/zenodo.3672785

[alam2019] S. Alam et al., "e-Infrastructure Services for EBRAINS," to appear in the proceedings of Brain-Inspired Computing: Third International Workshop, BrainComp 2019, Cetraro, Italy

[alam2020] S. Alam et al., "Archival Data Repository Services to Enable HPC and Cloud Workflows in a Federated Research e-Infrastructure," 2020 IEEE/ACM International Workshop on Interoperability of Supercomputing and Cloud Technologies (SuperCompCloud), 2020, pp. 39-44, doi: 10.1109/SuperCompCloud51944.2020.00012.

[anssi2018] French National Information Systems Security Agency (ANSSI) "Requirements for cloud computing service providers (SecNumCloud)," June 2018, https://www.ssi.gouv.fr/uploads/2014/12/secnumcloud_referentiel_v3.1_anssi.pdf

[bsi2017] German Federal Office for Information Security, "Cloud Computing Compliance Controls Catalogue (C5)," September 2017, https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/CloudComputing/ComplianceControlsCatalogue-Cloud_Computing-C5.html

[ec2016] EC, "European Charter for Access to Research Infrastructures. Principles and Guidelines for Access and Related Services," 2016, https://op.europa.eu/en/publication-detail/-/publication/78e87306-48bc-11e6-9c64-01aa75ed71a1

[etp4hpc2020] ETP4HPC et al., "TransContinuum Initiative (TCI): our vision", September 2020, https://www.etp4hpc.eu/tci-vision.html

[eu2016] European Parliament and European Council, "Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data," April 2016, https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN

[fenix2020] Fenix, "Fenix Security Measures Catalogue," version 1.0, February 2020 (available on request at https://fenix-ri.eu/contact-us)

[nist2011] NIST, "NIST Cloud Computing Reference Architecture," Special Publication 500-292, September 2011