EOSC Federation: Architecture and Federating Capabilities

16/10/2024

Authors:

Diego Scardaci (EGI.eu), Mark Dietrich (EGI.eu) and Paolo Manghi (OpenAIRE) on behalf of the EOSC Beyond Technical Coordination Board

Abstract

This document describes fundamental concepts to enable the Federation of EOSC Nodes, especially technical aspects, and defines key terms. It presents the EOSC Federation Architecture and introduces the concepts of EOSC Federating Capabilities describing how the Nodes of the Federation should interact to deliver them jointly.

**Funded by
the European Union**

Revision History

Issue	Item	Date	Comments	Author/Reviewer
V 0.1	Draft version	27/09/2024	Full draft for TCB review	Diego Scardaci (EGI.eu) Mark Dietrich (EGI.eu) Paolo Manghi (OpenAIRE)
V 0.2	Final draft version	11/10/2024	Implemented feedback from TCB	Diego Scardaci (EGI.eu) Mark Dietrich (EGI.eu)
V 1.0	Published version	16/10/2024	Implemented final feedback from TCB	Diego Scardaci (EGI.eu) Mark Dietrich (EGI.eu)

Copyright and licence info

Table of contents

Figures

Tables

1. Introduction

The EOSC Beyond project will define the technical architecture of the European Open Science Cloud (**EOSC)** as a **Federation of Nodes** evolving the EOSC Platform architecture defined by EOSC Future and implemented by the EOSC EU Node. A **Federation of Nodes** combines a governance and operational structure with a technical *network of Nodes.* This document will focus on the technical ingredients and specifications required to establish such a network of Nodes and to enable it to deliver services and other resources to European researchers collaboratively. Building on this understanding, the EOSC Beyond project will deliver the next generation of **EOSC Core services** as a reference implementation of the services that enable the operations of the EOSC Nodes.

This paper describes fundamental concepts enabling the Federation of EOSC Nodes and defines key terms. It introduces the concepts of **EOSC Federating Capabilities** and presents a model to describe how the Nodes of the Federation can interact to jointly deliver them. This model sets the grounds for defining the EOSC Federation Architecture and identifies the main elements that are required to make a Federating Capability operational in EOSC.

This document will support and enable community consultation on these topics. It will evolve in the coming months in response to feedback as well as being extended to cover new areas.

2. EOSC Federation

The European Open Science Cloud (EOSC) aspires to enable the sharing, discovery, access, and composition of a portfolio of services and research products that cater to the diverse needs of researchers and scientists regardless of discipline. Through EOSC, providers of services for research can share and contribute to a complementary collection of services ranging from data storage and management to computing resources, streamlining access to essential resources for research and innovation. This portfolio of services, available on a federated basis across the EOSC Federation, allows researchers, policymakers, and organisations to easily discover, access, and utilise the services, data, and infrastructure needed to advance their work, promoting collaboration, efficiency, and impactful scientific outcomes.

The EOSC Federation intends to achieve the vision of EOSC by integrating the highly heterogeneous landscape that exists today, where scientific clusters, research infrastructures, e-infrastructures, and national infrastructures have been created independently, adopting and consolidating different technological stacks, and operating services that are today embedded in the practices of each of the communities served by these infrastructures. Achieving the economies-of-scale, interoperability and convergence promised by EOSC requires an incremental and smart approach, balancing opportunities for harmonisation and consolidation with preservation of today's successful community practices.

2.1. EOSC Nodes - Architecture and Federating Capabilities

The heterogeneous infrastructures described above are the "key components", the **EOSC Nodes**, of the planned **EOSC Federation**. These nodes work together to offer new functionalities to end users, creating a **network of peer-to-peer nodes**[1], see Figure 1 as an example.
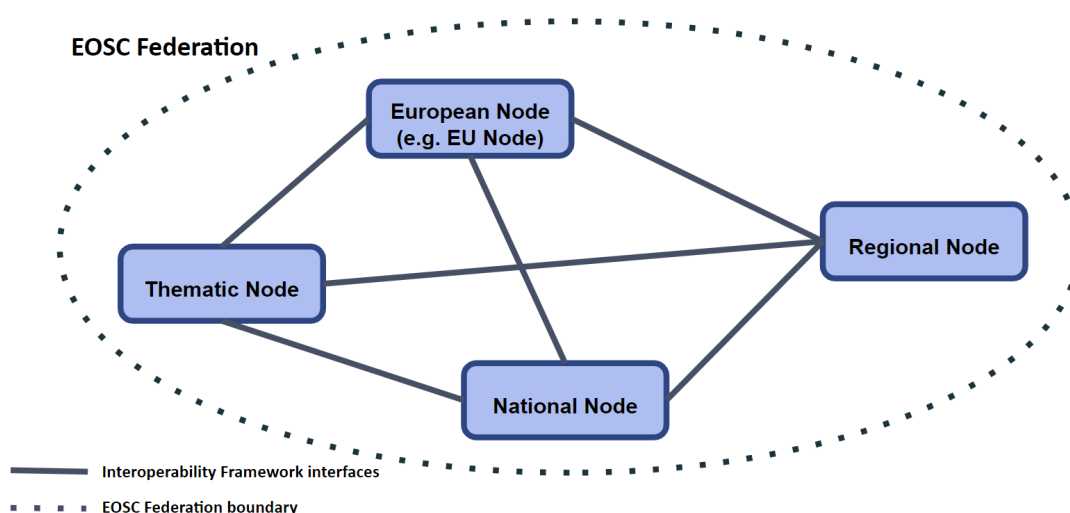

Figure 1. EOSC Network of Nodes

---

[1] https://eosc.eu/news/2024/05/towards-a-fully-fledged-eosc-federation/

> **EOSC Federation: a network of peer-to-peer nodes** that collaboratively deliver added value capabilities to its users[2].

2.1.1. EOSC Nodes

Nodes represent national, regional, European or thematic research collaborations that deliver services, datasets and other resources to its users. The Nodes that choose to enrol as **EOSC Nodes** have their autonomy, governance models, and well-defined offers in terms of services and resources. In order to become an EOSC Node, these Nodes must comply with the EOSC Federation policies, as well as with its legal and technical frameworks.

While this document focuses on the technical aspects, other initiatives will define the governance and operational steps that each of such infrastructures (the Nodes) must take to apply for and be enrolled into the EOSC Federation, thereby formally becoming an EOSC Node.

The Nodes that enrol as EOSC Nodes can be diverse in terms of both organisational structures and technical architectures, but several key features can be identified that are common to them all:

- **Node types and Node Communities:** each Node has been created and is maintained to meet the needs of a defined community of researchers. These communities are typically united by their thematic, regional, national or European focus.
- **Node Users:** users from these communities can be referred to as **Node Users**.
- **Node Objectives:** most Nodes share the vision of EOSC, namely to enable sharing, discovery, access, and composition of a portfolio of services and research products, but, unlike EOSC, the scope of this vision is limited to the needs and interests of the Nodes' specific communities, rather than the needs and interest of the entire research community.
- **Node Core Services.** Within the Node, some services are operated, either centrally or in a collaborative way, to enable the functioning of the Node overall. Examples include: catalogue and similar "discovery" services, repositories storing research data, identity provider (IdP) and attribute provider services, persistent identifier services, and monitoring and accounting services. In a sense, these are all "helper" services, intended to support the functioning of the Node, rather than services or resources that researchers would directly use in their research. These services are referred to as **Node Core services**.
- **Node Resources.** Other services and research products (including datasets) may be offered to Node Users, through inclusion in one or more "discovery" services available in the Node, or through some other "promotion" channels to Node Users. These resources are only discoverable from the Node and/or by Node Users; there may be additional restrictions on their use, but these restrictions are managed within the context of the Node and the community it serves. For example, bioinformatics pipelines might only be offered to users in a specific bioinformatics user community.

---

[2] https://eosc.eu/news/2024/05/towards-a-fully-fledged-eosc-federation/

Resources: Services, catalogues, research products, training resources, interoperability guidelines and other assets. In the context of EOSC, these are understood to be both digital and produced or operated by, or of potential use to, scientists and researchers. Services include data sources, computational resources and data storage resources.

Node Core Services: All the services that enable the operations of a Node such as, for example, a catalogue of services and/or research products, helpdesk, AAI, Monitoring, Accounting, etc. When a Node is enrolled in the EOSC Federation, these Node Core Services are referred to as EOSC Node Core Services and may need to comply with the requirements of the EOSC Federation.

While Nodes are the building blocks of the EOSC Federation, each Node must map its characteristics and the resources it has assembled onto the structures of the EOSC Federation. In particular,

- The governance and operational structure of a Node must be aligned with the structures defined for the EOSC Federation. This alignment is being addressed by other initiatives and not by the EOSC Beyond project.
- The EOSC Federation is establishing a list of technical and operational requirements that Nodes must meet in order to become EOSC Nodes.
- The EOSC Federation also expects EOSC Nodes to make a significant proportion of their Node Resources available to the EOSC Federation as a whole. The Node, and/or the providers of specific Node resources, would need to decide which Node Resources would be made available to the entire EOSC Federation. Those resources made available in this way would become **Node Exchange** resources as part of the overall **EOSC Exchange** and would be discoverable, accessible and usable (potentially) by any EOSC User.
    - For research products, such availability would require a transparent declaration of any access and use policies that might have been "implicit" within the Node. That is, no assumptions can be made about potential users of any EOSC Exchange resource.
    - For services, especially those that involve finite consumable resources such as compute and storage, such availability would require providers to consider the impact of potential use by a larger group of users and possibly to establish mechanisms to manage the extent and rate of this usage.

Node Exchange: services and other resources a given EOSC Node shares with EOSC.

EOSC Exchange: the collection of services and other resources EOSC Nodes share/federate with EOSC (federated resources).

2.1.2. EOSC Node Architecture

The infrastructure on which the EOSC Nodes operate their services and resources has to be compatible with the **EOSC Node Architecture**, a reference architecture where a set of **Core** and **Horizontal services** supports a number of *common and/or enabling capabilities* that allow the Nodes to *share resources across the Federation*. Examples of these capabilities are a catalogue of services and other resources, an authentication and authorisation infrastructure, a helpdesk, data transfer services, etc. The Nodes can implement their services with technologies of choice provided they remain compliant with the interfaces defined for these services in the **EOSC Interoperability Framework** (EOSC IF)[3] to connect with other Nodes.

The EOSC Node Architecture is an evolution of the EOSC Platform[4] (see Figure 2), which has recently been implemented in the EOSC EU Node, taking into account the new node-based EOSC Federation approach.

Figure 2. EOSC Platform Architecture according to EOSC Future

> **EOSC Node Architecture:** A reference architecture that can be implemented by each Node of the federation for the operation of its services and resources and to connect to the Federation enabling resource sharing.

---

[3] EOSC-A Technical Interoperability of Data and Services Task Force: A landscape overview of the EOSC Interoperability Framework - Capabilities and Gaps (https://zenodo.org/records/8399710)
[4] EOSC Future D3.3b Architecture and Interoperability Guidelines for Operational Services of the EOSC Core

> **Horizontal Services:** Services which can be applied in a general way to solve a significantly large class of problems such as a data transfer service or a workflow management system, etc.

> **EOSC Interoperability Framework (EOSC IF):** An extensible framework of interoperability guidelines that acts as the glue to support the interoperability and composability of EOSC Node resources. The EOSC IF is a reference framework to promote standards and best practices while offering the freedom to providers/nodes to develop and operate provider-specific implementations while conforming to the EOSC IF guidelines and standards.

2.1.3. EOSC Federating Capabilities

EOSC Nodes are expected to collaborate to augment their offer to their users by delivering or participating in **EOSC Federating Capabilities**. These are added-value functionalities that allow EOSC users/providers to exploit EOSC services, data and other resources (e.g. a federated catalogue which gives access to all EOSC resources). They are required or offered across all EOSC Nodes and require collaborations between two or more Nodes. Federating capabilities are enabled by **EOSC Federating Services** that can be either Core or Horizontal.

> **EOSC Federating Capabilities:** added-value functionalities that allow EOSC users/providers to exploit EOSC services, data and other resources (e.g. a federated catalogue which gives access to all EOSC resources) jointly delivered by two or more Nodes.

> **EOSC Federating Services:** enablers of EOSC Federating Capabilities.

> **EOSC Node:** An entity complying with the EOSC Federation policies and legal framework, working at local, national, regional, thematic or European level. An EOSC Node offers added value resources to the EOSC Federation and delivers federating capabilities in collaboration with other EOSC Nodes. Each EOSC Node has its autonomy, its own governance model and an offer in terms of resources. It operates its own platform, complying with the technical framework, and the EOSC Federation Platform architecture.

2.2. The EOSC EU Node

To stimulate the establishment of the EOSC Federation, the European Commission (EC) launched a procurement action to deliver the EOSC EU Node[5]. It will be the first recognised Node of the EOSC Federation and will act as a reference for the set-up of other Nodes. It will operate at the European level, offering a multi-disciplinary and multi-national scientific data/service portfolio for all research users and citizen scientists. It will also enable

---

[5] https://open-science-cloud.ec.europa.eu/

federating capabilities such as Identity Management, Federated Resource Catalogue and Application Workflow Management. Common horizontal services for end-users (compute, containers, data transfer, notebooks, file sharing, open research data) complete its offer. Details on the EOSC EU Node Federating Capabilities are provided in Section 3.

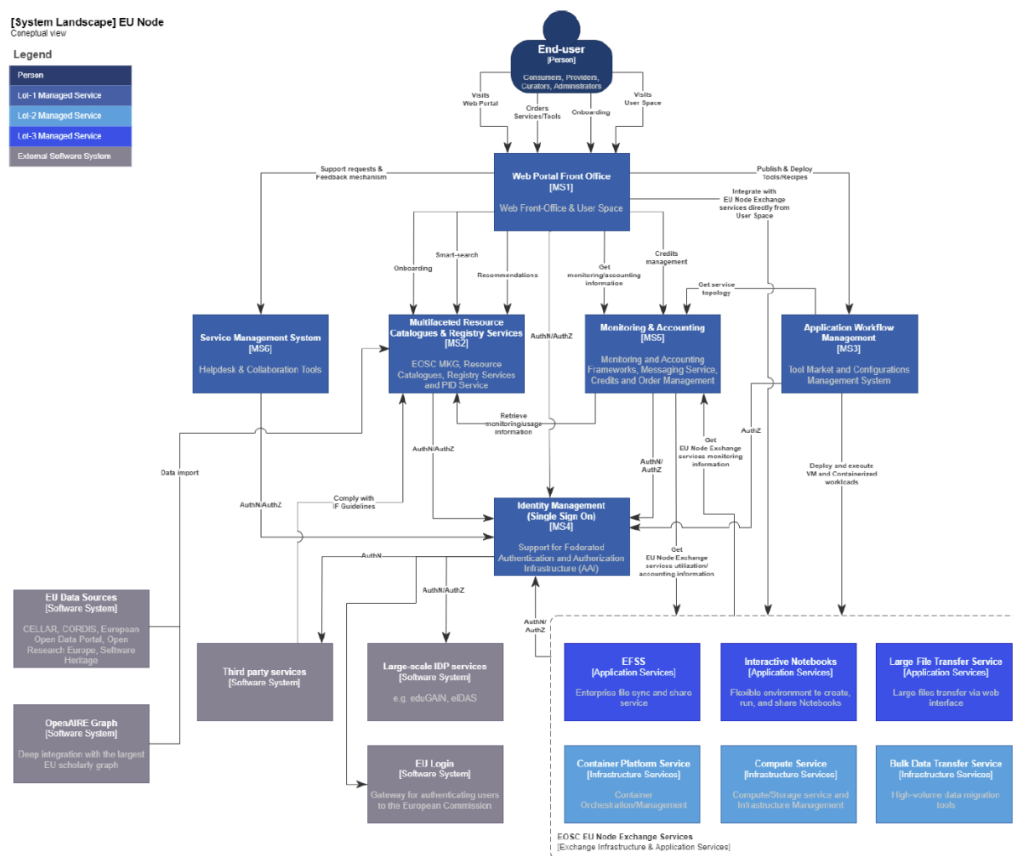Figure 3 depicts the EOSC EU Node Architecture.

Figure 3. The EOSC EU Node architecture.

2.3. EOSC Node Service structure

As briefly mentioned in the section above, EOSC Nodes operate infrastructures that are compatible with the EOSC Node Architecture. In line with this architecture, each Node can be modelled as described in Figure 4. where three main elements are identified:

- **Node Core Services:** to enable the Node operations. Examples of Core Services are a service and/or data catalogue, a monitoring system, an AAI infrastructure, a helpdesk, etc.
- **Node Federating Services:** to enable the nodes to participate in the delivery of Federating Capabilities in collaboration with other nodes. These Node Federating Services may represent versions of the Node Core Services adapted to comply with the Interoperability Framework for respective Federating Capabilities, or may represent new capabilities set up by the Node to participate in the EOSC Federation.

- **Node Resources:** a catalogue of community services and other resources which the end-users of this Node can directly access. This is further split into two subsets:
  - **Node Exchange:** services and other resources a Node shares with the EOSC Federation. They contribute to the overall *EOSC Exchange*.
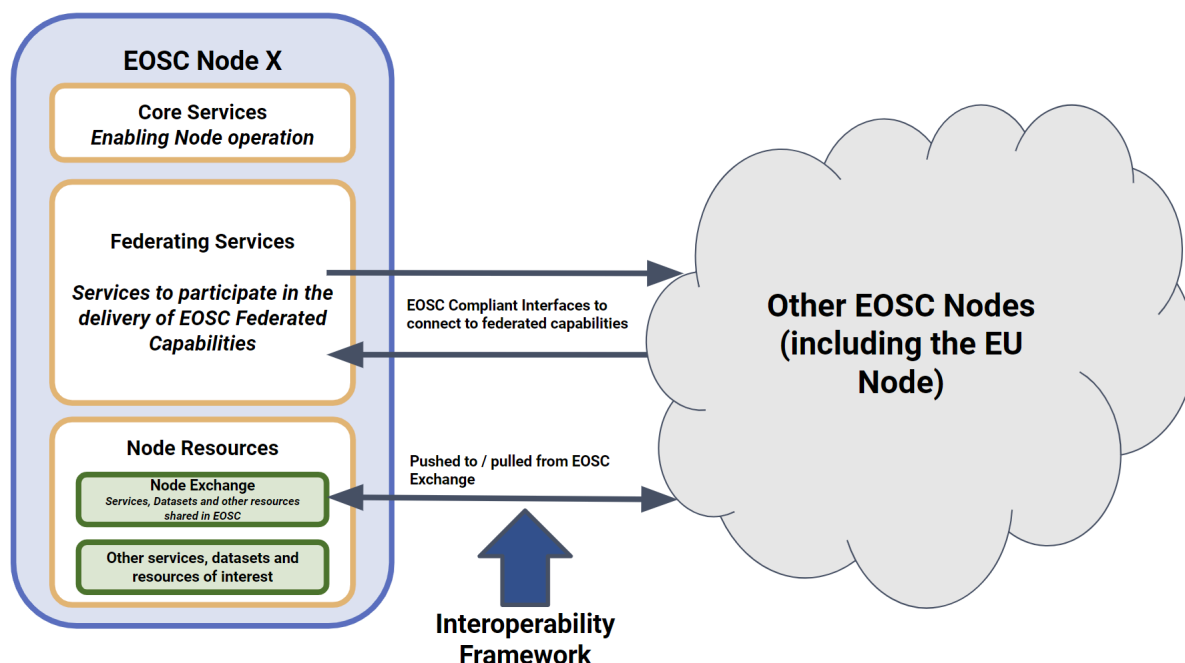 - **Other services and resources:** accessible via a Node but not shared with EOSC.

Figure 4. An EOSC Node's Service structure.

2.4. EOSC Federation Governance

The structure of governance for the EOSC Federation has been proposed by the EOSC Tripartite[6] and is currently under consultation. This discussion addresses "who" will make decisions for the EOSC Federation.

The decisions that will be made by EOSC Federation Governance will include the definition of various policies for the EOSC Federation, including:

- **Rules of Participation (RoP) for Nodes:** What are the requirements for Nodes to be enrolled as "EOSC Nodes"? What conditions (legal, organisational, as well as technical) must they comply with to be enrolled and remain part of the EOSC Federation? These requirements are likely to include specification of Federating Capabilities that each EOSC Node must support (see section 2.1.4 above).
- **RoP for the EOSC Exchange:** What are the requirements or conditions for Node Resources to be made available in the Node Exchange and to become part of the EOSC Exchange?
- **RoP for EOSC Users:** What are the requirements or conditions for Node Users to also become EOSC Users and gain access to the EOSC Exchange?

These and other policies are discussed in the EOSC Federation Handbook[7].

---

[6] https://eosc.eu/tripartite-collaboration/
[7] https://eosc.eu/eosc-federation-handbook/

3. EOSC Federation Architecture - A Network of Collaborative EOSC Nodes

The EOSC Federation is a system of systems (a network of EOSC Nodes) where each system (a Node) is an entity operating an infrastructure delivering services and other resources to its users. End-users access the federation via one of the Nodes and can discover and access the EOSC Resources (from this or other nodes) made available by this node.

The intrinsic added value of the EOSC Federation is defined by what the federation offers to the end users beyond what a single node can do. Organisations behind each node are joining the federation to enrich the number of services and datasets or the amount of additional compute and storage resources their users can access, improving, at the same time, the quality and the reliability of their offers.

To achieve this aim, the EOSC Federation envisages collaborations between Nodes to share services and other resources, adopt common service levels to operate quality services and jointly deliver new **Federating Capabilities** that allow researchers to fully exploit EOSC resources.

This section presents a model to describe how the nodes of the federation should interact to jointly deliver Federating Capabilities. This model sets the grounds to define the **EOSC Federation Architecture** and identifies the main elements required to make a federating capability operational in EOSC.

The list of federating capabilities that the EOSC Federation should offer must still be discussed and agreed upon. An initial set of these capabilities is delivered by the EOSC EU Node and detailed later in this section. High-level functional expectations have been indicated in initial descriptions of the EOSC Federation and its relationship to potential EOSC Nodes. Other capabilities of interest for research communities are being identified through the EOSC Beyond requirements gathering process and the work on the Pilot Nodes[8] and some examples of them are also reported later.

It is expected that the EOSC Governance structure will determine which Federating Capabilities will be required from all EOSC Nodes, and which ones will be recommended and/or optional. Ideally, this decision-making process will balance the high-level objectives of the EOSC Federation with the capabilities of and interest by candidate EOSC Nodes to collectively realise these objectives, and a phased approach may be needed to reflect technical maturity and the realities of implementation.

Once there is agreement at a governance level on requirements for EOSC Federating Capabilities, this must be translated into detailed technical requirements, specifications, and processes for integration and ongoing assessment of compliance. Some of these aspects will be included in the EOSC Interoperability Framework, and the required processes should be supported by the Federation Management Organisation.

---

[8] EOSC Beyond D15.1 Service integration plans of the Data Spaces and EOSC Nodes: https://doi.org/10.5281/zenodo.13152335

3.1. A Model for the EOSC Federating Capabilities

The EOSC Federation offers a certain number of federating capabilities that are enabled by the Nodes themselves. These capabilities can include diverse functionalities that require the participation of multiple Nodes to make them operational for the benefit of the end users, such as a common search engine, a single sign-on (AAI), common interfaces to share, access and analyse data, etc. The minimum set of federating capabilities a Node should support to join the EOSC Federation has to be defined in the EOSC Node Rules of Participation (policy compliance). EOSC can offer other federating capabilities as optional extras (e.g. a federating catalogue for geohazard datasets may be of interest only for a subset of nodes, so connecting to this catalogue would be not required for all nodes).

EOSC Nodes can either:

A. **provide a federating capability** in the EOSC Federation that can be used by other EOSC Nodes (e.g. an EGI Node enabling the EGI Cloud Federation).
B. integrate its own services to **participate in the delivery of a Federating Capability** (e.g. AAI Federation from EU Node) enabled by other Nodes.

A representation of the relationship between an EOSC Node and the federating capabilities is shown in Figure 5. In the picture, the stick represents a federating capability enabled by the EOSC Node X, the circle represents the service the EOSC Node Y uses to connect to this federating capability. In this notation, circles are connected to a stick when nodes connect their services to another node to participate in the delivery of a federating capability represented by this stick.
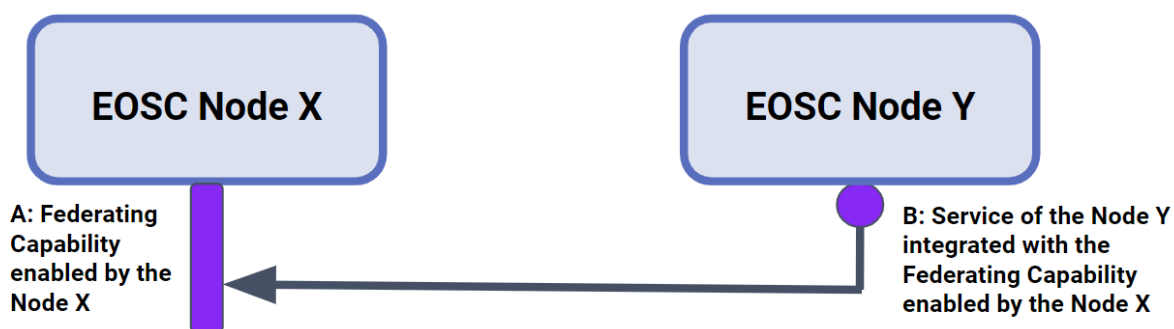
Figure 5. The EOSC Node X enables a Federating Capability (stick) and the EOSC Node Y federates with it connecting its service (circle).

3.1.1. Define and make operational an EOSC Federating Capabilities

Once the objective of a federating capability has been identified (what type of features will be offered), the service components that are required to enable it in the federation need to be specified. These components are the *enablers* of a given federating capability and are named **EOSC Federating Services**. These services can be operated by one or multiple Nodes of the Federation.

The EOSC Federating Services should expose interfaces (e.g. APIs) and adopt metadata schemas (structure of the data that should be provided) in compliance with the EOSC IF and its interoperability guidelines to allow EOSC Nodes to connect to this capability according to the EOSC standards. The ensemble of interfaces and metadata schemas the Federating Services exposes to Nodes for integration are called **EOSC Federating Interfaces**. New interoperability guidelines, specific to the new Federating Capabilities, can be defined and registered in the EOSC IF as needed.

> **EOSC Interoperability Guidelines:** specify the essential technical requirements and configurations that EOSC Nodes should follow in order to connect to a federating capability.

As an example of an EOSC Federating Capability, consider an *EOSC federated search service* able to perform distributed queries across catalogues from different EOSC nodes. To enable this capability the following Federating Services and Interfaces need to be set up:

- *Federating Services:* such capability requires three federating components to operate, namely (1) a registry to manage information and metadata about all the catalogues participating in the federated search (e.g. their API endpoints), (2) a web application offering a UI that allows users to submit queries and (3) a back-end capable of fetching catalogue information from the registry, brokering queries to the federated search engines, and merging the results to feed the UI.
- *Federating Interfaces:* such capability requires the definition of EOSC IF guidelines that identify the interfaces (e.g. APIs) and the metadata schemas that EOSC Nodes are required to implement to connect their catalogues to the EOSC federated search service.

Table 1 summarises the necessary elements to set-up an EOSC Federating Capability.

Elements	Description
Federating Services	Service components that enable a federating capability in EOSC
Federating Interfaces	The ensemble of interfaces and metadata schemas the Federating Services exposes to Nodes for integration

Table 1 - Elements to set up an EOSC Federating Capability.

The Federating Services and the Federating Interfaces are the basic elements for setting up a Federating Capability in EOSC. However, **a Federating Capability becomes fully operational**, bringing benefits to EOSC users, **only when Nodes start to contribute to it with their own data, participating in its delivery**.

To achieve this aim, Nodes have to connect/federate their services to the Federating Services of the capability through its Federating Interfaces and share data in the agreed formats in compliance with the related EOSC IF Interoperability Guidelines.

In the case of our example, the EOSC Federated Search Service, Nodes should onboard (i.e. register) their catalogues to the registry, i.e. complete a profile where API endpoints, metadata format info, etc. are described for the web application to find and reuse. This enables an increased findability and visibility of local resources of participating Nodes.

Indeed, after this process, researchers can find and request resources from all the EOSC Nodes that made them available throughout EOSC via the EOSC federated search engine. This engine can be made accessible by all Nodes of the Federation.

Figure 6 depicts a possible architecture of a common EOSC search engine showing how the EOSC Nodes can connect their catalogues to the federated search service via the interfaces defined in the EOSC IF.
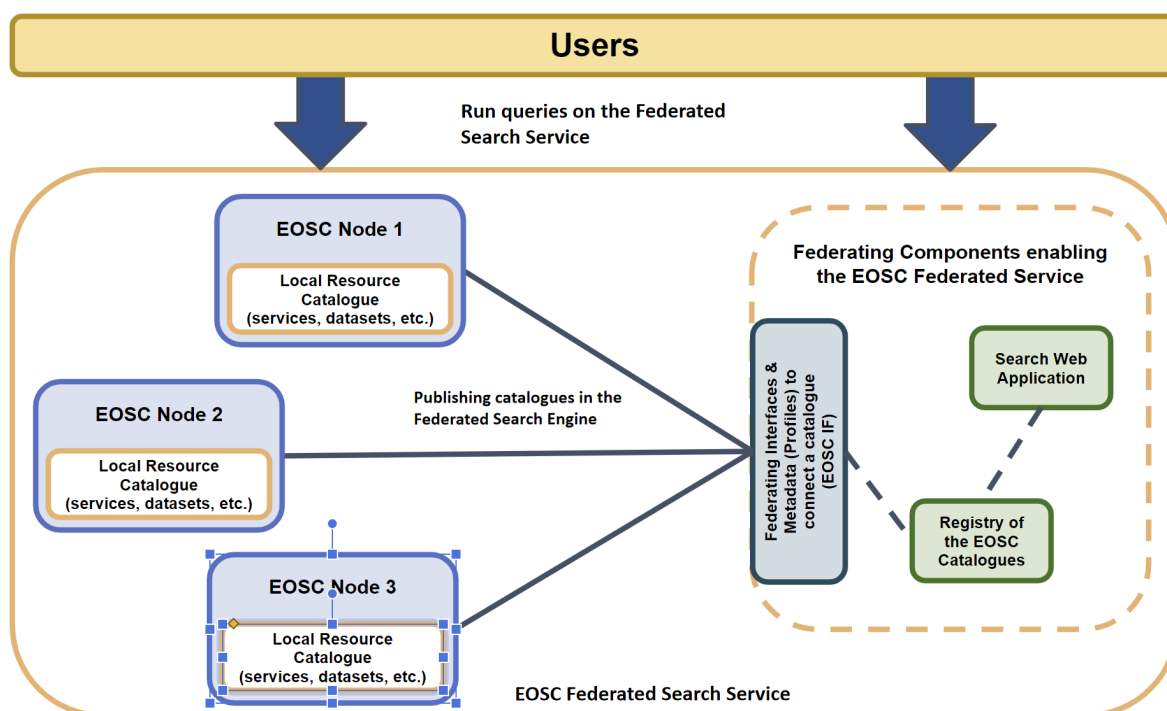
Figure 6. Example of the logical architecture of an EOSC federated search service.

It is worth noting that there are no constraints on who operates the Federating Components and related interfaces and where they are deployed. They can be operated in one or multiple nodes of the federation but require that (1) providers are selected for each of the components and that (2) the behaviour of all components is properly coordinated assuring their quality delivery. Appropriate service levels for each component need to be defined by the EOSC governance.

Figure 7 shows an example of deployment for the EOSC federated search service.

Figure 7. The components enabling the federated search service in the EOSC Federation can be operated by multiple nodes.

3.1.2. Process to set up an EOSC Federating Capability

Table 2 summarises the steps required to fully set up an EOSC Federating Capability. It assumes that there are two different types of federating capabilities:

- *Required Federating Capabilities*: all nodes of the EOSC Federation must federate with these capabilities. These are defined by the EOSC Governance.
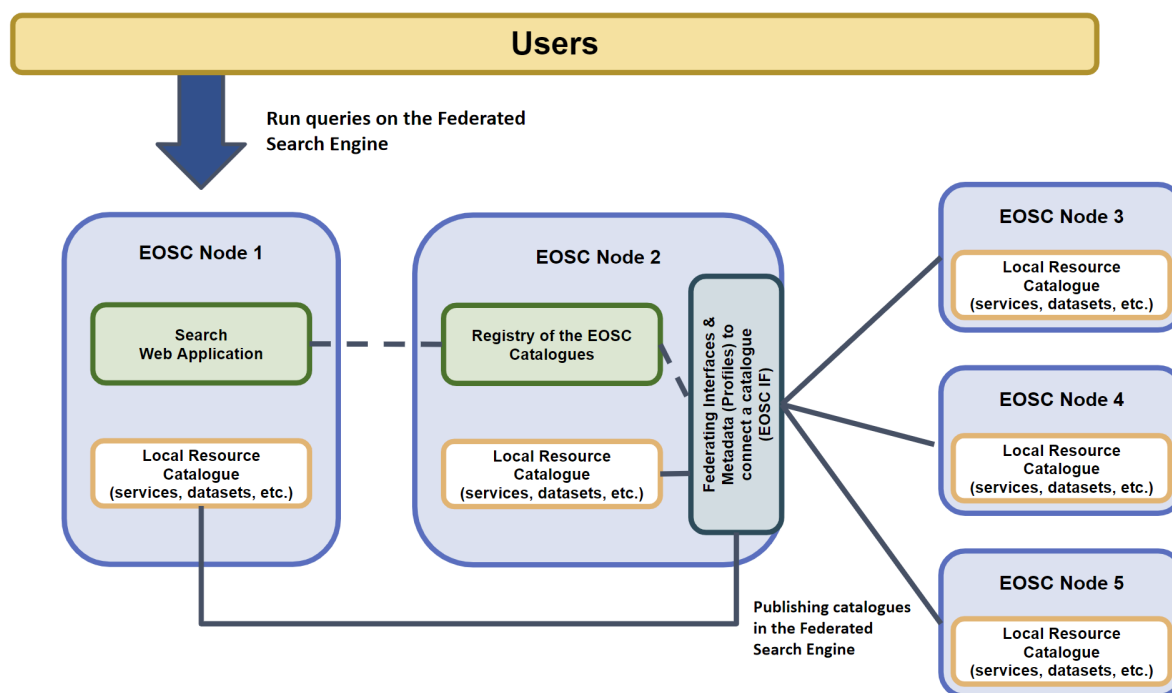- *Optional Federating Capabilities*: a node can decide to federate or not with these capabilities. These can be defined by the EOSC Governance or by the Nodes themselves.

Steps	Description	Actors	Example
Identify an EOSC Federating Capability	A new capability that should be delivered by the EOSC Federation is identified and described. Added value for end-users and nodes is understood.	● EOSC Governance (for required and optional Federating capabilities) ● EOSC Nodes (only for optional Federating Capabilities)	The EOSC Governance decides to set up an *EOSC federated search service* able to perform distributed queries across catalogues from different EOSC nodes
Set up the Federating Services and the	Services components required to enable the federating capability in the federation need to be specified	● EOSC Technology Providers	The EOSC Governance assigns to an EOSC

Steps	Description	Actors	Example
Federating Interfaces	and technologies/tools to implement them need to be identified or developed. Interfaces (e.g. API) towards the EOSC Nodes and metadata schemas (structure of the data that should be provided) are selected from the EOSC IF or new interoperability guidelines specific for the new Federating Capabilities are defined and registered in the EOSC IF.	designed by the EOSC Governance or EOSC Nodes to implement the federating capability	Federation technical body the task to design this new Federating Capability. This technical body designs a registry, a web application offering a UI and a backed tool able to fetch information from EOSC Node catalogues as the Federating Service needed to implement the federating capability. It also defines the interfaces that allow Nodes to connect to them. Procedure and interfaces to connect an EOSC Node catalogue to the Federated Search are described in an EOSC Federating Search IG registered in the EOSC IF.
Identify EOSC Nodes in charge of operating the Federating Services	EOSC Nodes are selected as providers to operate the Federating Service components with appropriate service levels, guaranteeing that the behaviour of all components is properly coordinated, assuring their quality delivery.	● EOSC Governance (for required and optional federating capabilities) ● EOSC Nodes (for optional federating capabilities)	When the design for the three Federating Services of the federating search service is ready and the technology is available, the EOSC Governance selects the providers to operate them. The Governance also agrees appropriate service levels with the providers.
Deployment and operations of the EOSC Federating	Selected Nodes deploy the Federating Services and start to operate them according to the	● EOSC Nodes selected to operate the	Selected providers deploy the registry, the web application

Steps	Description	Actors	Example
Services	agreed service levels. Federating Interfaces are exposed according to the related EOSC IF guidelines.	Federating Services	offering the UI and the backend and start to operate them according to the agreed service levels. Interfaces defined in EOSC IF are exposed to other Nodes.
Federating/ Connecting Nodes to the Federating Capabilities	Nodes connect their services to the Federating Services through the EOSC IF-based Federating Interfaces. Nodes start to populate the Federating Services with their own data. The EOSC Federating Capability is now operational and can be exploited by end users and Nodes.	● All EOSC Nodes for required capabilities ● EOSC Nodes willing to federate for optional capabilities	EOSC Nodes, willing to connect their catalogue to the EOSC Federated Search, access the related IG in the EOSC IF and follow the procedure to start to contribute to the global registry. After they complete the integration, the items in their catalogues will be discoverable via the EOSC Federated Search.

Table 2 - Process to set up a Federating Capability in the EOSC Federation.

3.1.3. Connect to an EOSC Federating Capability

Table 3 details the steps an EOSC Node should follow to participate in the delivery of an EOSC Federating Capability already enabled in the federation by other EOSC Nodes. According to the previous section, this implies that Federating Services for this capability are operational and they expose Federating Interfaces described in EOSC IF guidelines.

Steps	Description	Example
Federating Capability assessment	A Node assesses a Federating Capability to understand if connecting to it can bring value to its end-users. The Node will proceed with the next steps if the value for its community is recognised.	An EOSC Node, operating a catalogue of resources, is interested in participating in the *EOSC Federated Search service* to allow EOSC users to discover items from its catalogue through this federating capability.
Analysis of the	The node access the Interoperability	The EOSC Node gathers

Steps	Description	Example
Federating Interfaces	Guidelines of the Federating Capability and analysis the requirements to connect to it: • Exposed interfaces • Data to be exchanged and their format	information on the API and on the required metadata format to publish information about its catalogue in the registry of the Federated Search service from the related Federated Search Interoperability Guidelines.
Setup the Node services to connect to the Federating Capability	The node can either: • Use an already existing Node service to send data to the Federating Services via the Federating Interfaces. This can require the development of a wrapper to be compliant with the Interoperability Guidelines • Adopt a tool compliant with the interfaces defined in the Interoperability Guidelines • Adopt a service provided by EOSC as a Service to connect to this federating capability Regardless of the chosen option, the node should connect the Node services to the Federating Services following the procedure described in the related Interoperability Guidelines.	The EOSC Node deploys and operates a component that exposes information of its resource catalogue, including API to access it, according to the metadata structure defined in the interoperability guidelines. Instead of developing its own component, EOSC Node may decide to use ready-to-use sw libraries (adapter) that EOSC makes available to facilitate the integration.
Make the Federating Capabilities available to its users	After the Node is connected to a Federating Capability, it can make available features and data enabled by this capability to its users. This may require the development of specific user interfaces or the exposure of the Federated Search service interfaces at node level.	The EOSC Node decides to locally deploy a web application to allow its users to access the EOSC Federated Search service. The node can reuse the web application made available by the Federated Search service or a new one customised for its needs but compliant with the EOSC Federated Search interoperability guidelines.

Table 3 - Process to connect to a Federating Capability in the EOSC Federation.

3.2. Detailing the EOSC Node Service Structure

The previous section clarified the role of EOSC Nodes concerning the Federating Capabilities, highlighting how a Node can be an enabler of a federating capability or can

participate in its delivery by connecting its services to the federating services behind the capability.

Taking into account this, the EOSC Node Service Structure presented in Section 2.3 (Figure 4) can be further defined. The *Federating Services* elements can be split in two sub-elements as described in Section 3.1.

This is graphically represented in Figure 8.
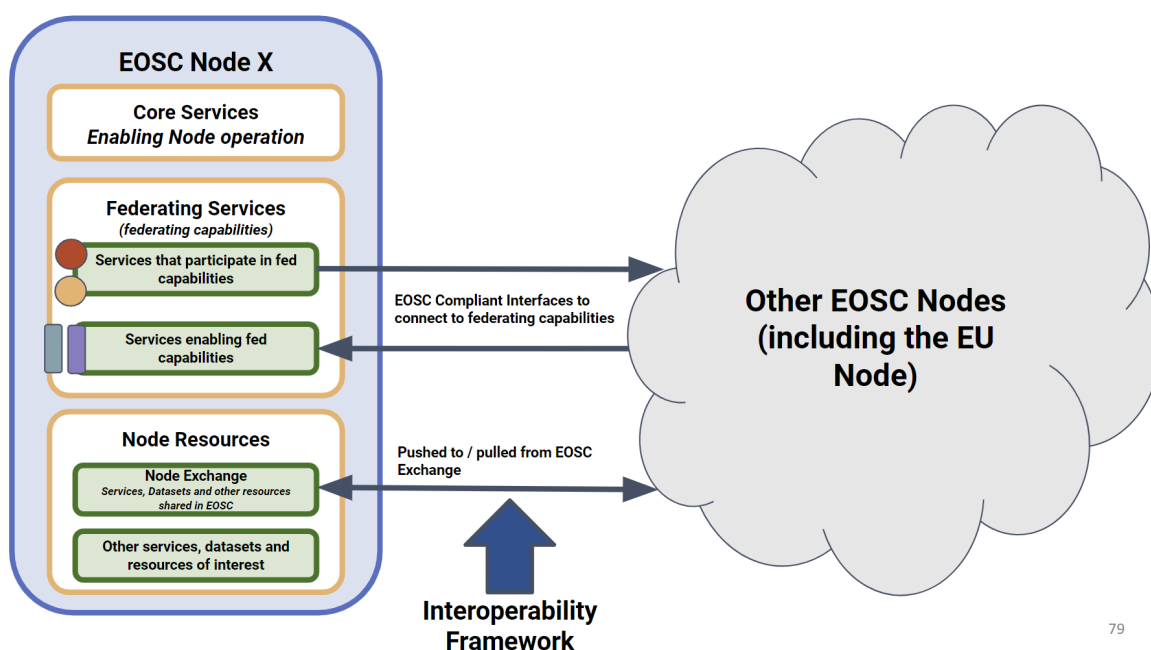
Figure 8. Further detailed EOSC Node service structure

3.3. EOSC EU Node Federating Capabilities

As introduced in Section 2.2, the EOSC EU Node is a node at the European level delivered by the EC via a procurement action that is expected to be the first fully operational node of the EOSC Federation. The EOSC EU Node official launch is planned at the EOSC Symposium 2024[9].

The EOSC EU Node is an enabler for a number of EOSC Federating Capabilities[10]. An overview of these capabilities is provided in Table 4. Because the EOSC EU Node is not yet operational at the time of writing this document, Table 4 contains only the information that is currently public and does not provide details on how other EOSC Nodes can federate with these capabilities. This section will be extended when more information is available.

---

[9] https://eosc.eu/symposium2024/
[10] https://indico.egi.eu/event/6441/contributions/19530/

Federating Capability	Description
Resource Catalogues and Registry services	● Federated EOSC Resource Catalogue (services, datasets and other research products) ● Advanced search engine ● Interfaces to allow EOSC Nodes to publish (push) and retrieve (pull) resources to/from the federated catalogue
Identity Management	Single sign-on over the federated nodes
Application Workflow Management	Deploy applications on compute and storage resources from multiple federated nodes
Service Monitoring	Provide information about the availability and quality of EOSC services (Core and Exchange)
Service Accounting	Provide information about usage of services and resources in EOSC (Core and Exchange)
Order Management	Offer a framework for providers to define offers and a unique interface for end-users to request access to resources
Management System (including Helpdesk)	FitSM-based Service Management System (SMS) that can be federated with SMS from other EOSC Nodes

Table 4 - Federating capabilities enabled by the EOSC EU Node

3.4. Example of Federating Capabilities across the Network of EOSC Nodes

The concept of EOSC Federating Capabilities allows the set up of the EOSC Federation as a real network of Nodes where each Node can be connected to the others to offer new added-value features to the end users. However, it can be envisaged that (1) some Federating Capabilities are of interest for all the nodes of the Federation, requiring or suggesting the participation of all the nodes in their delivery, while (2) others can be more sectorial and relevant only for a small subset of nodes.

The first category can include capabilities such as a global federated catalogue, the AAI or the access to distributed compute and storage resources while the second can foresee capabilities such as a Cultural Heritage or a Biomedical data catalogue.

Figure 9 shows an example of an EOSC Federation made up of five nodes (including the EU Node) highlighting potential connections between the nodes via Federating Capabilities. This is only an example to show how the EOSC Federation could appear when the first set of nodes are enrolled and initial federating capabilities are enabled. The picture uses the notation with sticks (federating capability enabled by a node) and circles (service a node uses to connect to a federating capability) introduced in Section 3.1.

Figure 9. Example of a network of EOSC Nodes. The connections between nodes are only hypothetical.

3.5. An example of a multi-node use case - The Application Workflow Management (AWM) component of the EOSC EU Node

Federating Capabilities bring added value to the end users allowing the implementation of use cases that require the collaboration of multiple nodes.

As an example, consider the case of a community that is using the EOSC EU Node to perform data analysis. The amount of computing resources requested by this community to achieve its task is higher than the capacity the EOSC EU Node can offer (e.g. according to its Acceptable Usage Policy the EU Node can offer a maximum amount of resources), then the community use case cannot be supported by the EOSC EU Node alone. However, the Application Workflow Management (AWM) component of the EOSC EU Node can serve as an integration point for connecting computing infrastructure hosted by other nodes offering federating interfaces for integration. These interfaces will be described in the AWM interoperability guidelines.

Other nodes can decide to federate with the EOSC EU Node AWM Federating Capability to extend the compute and storage capacity and time allocations offered by the EOSC EU Node. This would allow AWM to identify and automatically redirect workloads to the integrated cloud providers, according to user parameters such as required compute capacity and data physical location. It is important to underline that proper agreements between the EOSC EU Node and the connecting Nodes are required to enable the federation of resources. These integrations would allow the EOSC EU Node to scale up its resources and, consequently, be able to serve a wider set of research communities and satisfy more demanding requirements.

Figure 10 describes this use case supposing that the EOSC EU Node is federated with an EGI Node delivering computing services[11].
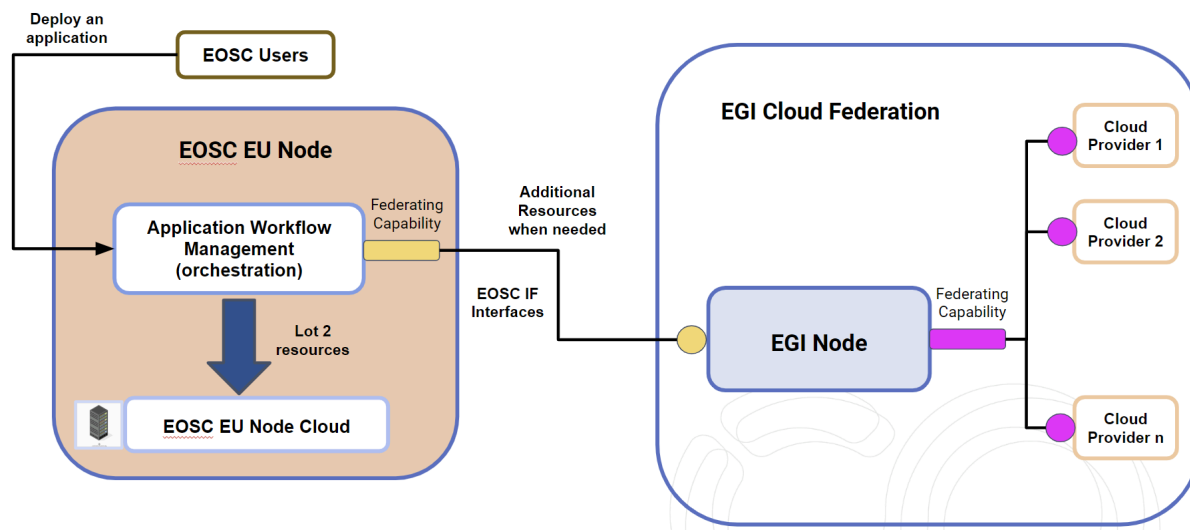

Figure 10. EOSC EU Node AWM scales up with capacity and research data hosted by the EGI Federated Cloud providers and partnering Research Infrastructures[12].

The federation between the EOSC EU Node and other Nodes via AWM can enable other relevant scenarios. Consider again the integration between the EOSC EU Node and the EGI Node, the latter can aggregate user demand from EGI research communities and projects (e.g. communities of practice, long tail of science and Research Infrastructures) and act as a user proxy towards the EOSC EU Node services. For example, this would enable the co-located processing and analysis of research data hosted by the EOSC EU Node with advanced computing, datasets, models, software and user support provided by the EGI Federation. This integration would create an additional delivery channel for the EOSC EU Node that would lead to increased utilisation. AWM recipes, which describe application architectures to be deployed, together with virtualised computing resources, can be made accessible to the EGI user base and served by the EOSC EU Node, as target infrastructure, in conjunction with the EGI infrastructure. Furthermore, EGI can contribute advanced recipes to be registered in the EOSC EU Node AWM and made available to all the EOSC EU Node users.

In both use cases depicted in this section, research communities would be able to access data and infrastructure from two or more EOSC nodes as one integrated offer to implement their scientific solutions.

---

[11] EGI Contribution to EOSC: https://zenodo.org/records/11128540
[12] https://www.egi.eu/egi-infrastructure/

4. Conclusions and next steps

This paper describes the fundamental concepts behind EOSC as a Federation of EOSC Nodes and defines key terms, focussing on the technical aspects. A model to describe how the Nodes of the Federation should interact to jointly deliver **EOSC Federating Capabilities** has been presented together with example use cases of Federating Capabilities.

The work presented in this document will be used as the basis to define an overall EOSC Federation Architecture that takes into account the concept of collaborating EOSC Nodes. This document will evolve in the next months according to the feedback received and the outcome of the EOSC Beyond technical and piloting activities.

One of the most important topics that will be analysed in the coming months will be how to register and make discoverable Nodes of the EOSC Federation. It is evident that there is a need to create a registry that keeps track of the Nodes that are participating in the federation. This registry should record basic information of each EOSC Node, including their offer in terms of services and data, and of active Federating Capabilities. The next version of this document will also include an analysis of this requirement and propose a solution for it.