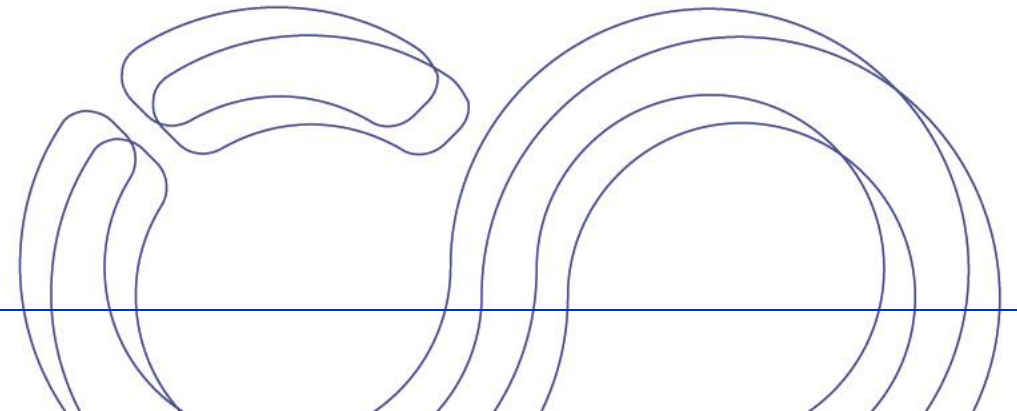


FTS & EOSC Beyond

João Lopes

09/09/2024

FTS&XRootD Workshop 2024



Introduction to EOSC



- **What is EOSC?**

- European Open Science Cloud.
- **A federation of services, tools, data, and other research resources.**
- Consists of resources from hundreds of providers across local, national, regional, and European levels.

- **Mission:**

- Increase **discovery, access, and reuse** of Open Science resources.
- Facilitate collaboration and resource sharing across European research communities.
- Accelerate scientific discoveries by providing access to diverse research resources.
- Enhance European scientific excellence through openness of research data and tools.

Introduction to EOSC Beyond



- **What is EOSC-Beyond [1] ?**
 - **Three year EC project (From 04/2024 to 03/2027):**
 - HORIZON-INFRA-2024-EOSC-01-04 [2].
 - Consortium led by European Grid Initiative Foundation (EGI).
 - +30 Participant institutions (CERN is one of them).
- **Overview of EOSC-Beyond:**
 - Aims to support the growth of EOSC by expanding list of service providers and active users.
 - Focus on developing new technical solutions to enhance Open scientific application environments.
- **Key Goals:**
 - **Advance and enhance the existing EOSC Core** with improved technical solutions and frameworks.
 - **Facilitate the integration and interoperability** of a diverse portfolio of EOSC resources for researchers.

[1] <https://www.eosc-beyond.eu/about>

[2] <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/horizon-infra-2024-eosc-01-04>

EOSC Beyond Overall Objective

Advance Open Science and innovation in research by providing new EOSC core capabilities that allow scientific applications to integrate and compose multiple EOSC Resources from national, regional and thematic EOSC Nodes. [1]

[1] <https://eosc.eu/eu-project/eosc-beyond/>

EOSC Beyond Objectives

Objective 1: Enable a European Network of pilot EOSC Nodes.

- **Enhance the EOSC Federation by offering EOSC Core components as-a-Service, enabling the creation of connected Regional, National, and Thematic Nodes across Europe.**
 - Facilitate the setup of EOSC nodes connected to the EOSC Platform to share data, services or other resources.

Objective 2: Accelerate the development of scientific applications with ready-to-use adapters.

- **EOSC Beyond will provide an "Integration Suite" that will accelerate the development of data-driven tools and services by offering adaptable software libraries for integration of EOSC services across research communities.**
 - Software libraries that are adapters to integrate applications/tools, services, or research data.

[1] <https://github.com/indigo-dc/tosca-templates/blob/v4.0.1/doc/tosca.md>

EOSC Beyond Objectives

Objective 3: Enable Open Science with machine-composability and dynamic deployment of shared Resources in the context of the EOSC Execution Framework.

- **EOSC Beyond will enhance the EOSC "Interoperability Framework" to enable machine-composability, allowing automated, dynamic integration and configuration of services and data using TOSCA templates and DevOps tools.**
 - Deployable services and correspondent guidelines can be discovered in the EOSC.

Objective 4: Increase innovation in EOSC with an environment for the development, testing, integration and deployment of services, data and other research object.

- **EOSC Beyond will create the EOSC "Core Innovation Sandbox", a pre-production environment for testing and verifying EOSC services and resources before their integration into the EOSC Platform.**
 - Validation of services through the Innovation Sandbox.

Objective 5: Enable federated discovery and access of FAIR data across EOSC and European Data Spaces.

EOSC Beyond new components

Integration Suite

- A collection of ready-to-use, pluggable software libraries, or "adapters," that enable easy integration of tools, services, and data across research communities.
- Help standardize interfaces and best practices, allowing different research teams to share and reuse solutions for common challenges in scientific applications.
- Provide adapters for key EOSC Core services making them easily discoverable and usable.

Execution Framework

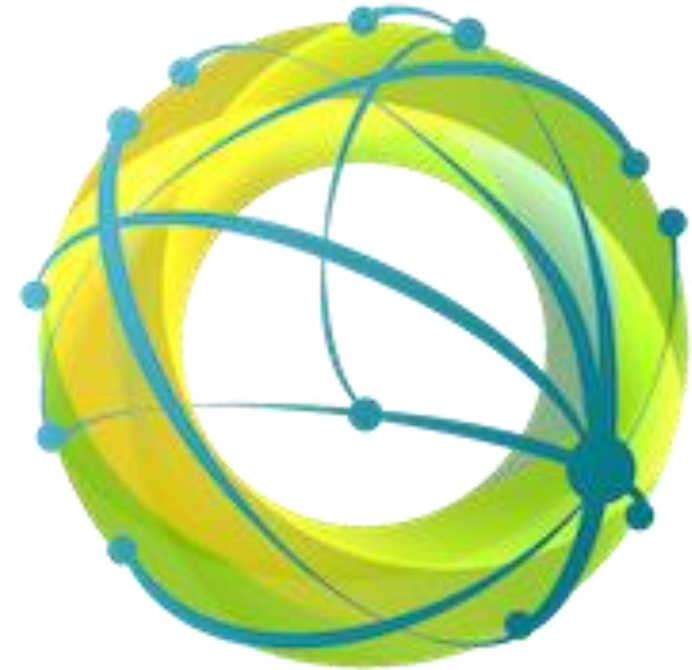
- Enables the automated deployment and orchestration of services across hybrid cloud infrastructures.
- It consists of an enriched registry providing machine-readable configurations and a Deployment Service for on-demand service instantiation.
- Facilitate service discovery and orchestration through structured metadata.

Innovation Sandbox

- A testbed equipped with automation tools for testing and validating services.
- Allow providers to prototype new EOSC Core services and verify their integration with the EOSC Platform before deployment, ensuring resources are ready for production.

FTS Contribution to EOSC Beyond

- Bring to EOSC expertise on how to transfer large amounts of data for the WLCG collaboration.
- Make FTS data transfer capabilities available within EOSC.
- Use the EOSC platform to define pilot new/innovative solutions to transfer data from/to EOSC to/from the European Data Spaces.



FTS Contribution ideas

- Develop libraries to facilitate the use of the FTS within EOSC:
 - The FTS client library exists:
 - Can it be improved to make it easier for new communities to use FTS?
 - Documentation/Guidelines can also be improved.
 - Idea of new "super client" that could submit tasks to FTS and keep track of them could help small users:
 - Running locally as an agent.
 - Remembering which tasks it submitted and keeping track of them "*automagically*".
- Integrate FTS with other non-WLCG IT services:
 - Authentication and Authorization services.
 - Monitoring services:
 - FTS relies on CERN's Monit infra to provide transfer monitoring (ActiveMQ, OpenSearch, Grafana etc.)
 - Storage providers and new transfer protocols:
 - WLCG moved to xrootd&https.
 - In EOSC the SFTP protocol seems to be preferred as it generally available on almost any platform.

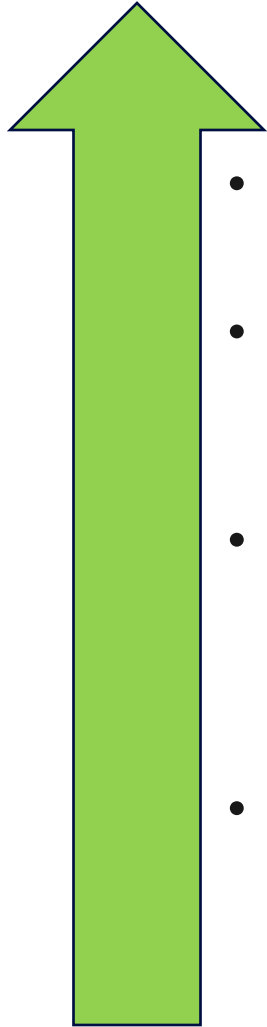
FTS Contribution ideas

- Making deployment of FTS easy:
 - Puppet is used at CERN to deploy FTS.
 - A lot of specific FTS know-how is required.
- Can the deployment of FTS be automatized differently?
 - Integration with INDIGO-DataCloud [1] platform.
 - FTS service TOSCA template [2].
- FTS global deployment:
 - In WLCG generally each collaboration has its own single FTS instance.
 - Different FTS instances don't communicate/collaborate.
 - Prepare for a future where it's possible to "click&deploy" an FTS service on the cloud?

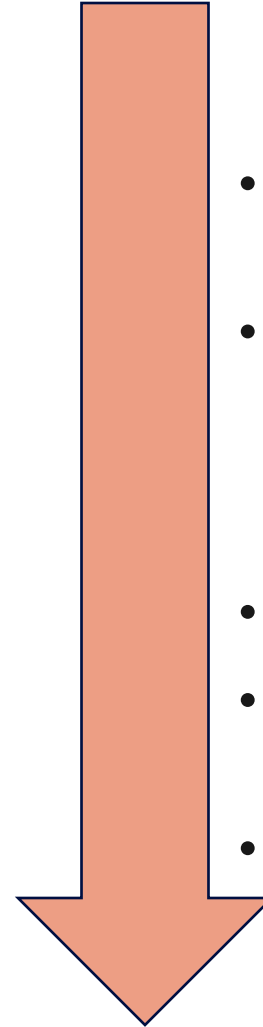
[1] <https://www.indigo-datacloud.eu/>

[2] <https://github.com/indigo-dc/tosca-templates/blob/v4.0.1/doc/tosca.md>

Final Remarks



- Possibility to bring FTS to new communities.
- Ideas/requirements from new communities using the service can be used to improve the FTS service.
- The general FTS community will benefit if the service is made easier to use, the documentation is improved etc.
- Freedom to experiment/prototype ideas.



- Difficult to navigate in a project with +30 partner institutions.
- FTS is not the center of this project:
 - General IT services (AAI, monitoring, helpdesk etc) in addition to scientific specific ones are part of the project.
- Requirements are not always clear.
- Sometimes a solution in search of a problem.
- Too much freedom can lead to a lack of direction.

Conclusion

- EOSC Beyond is a new project that just start in April 2024.
- Is an initiative to drive the evolution of Open Science in Europe by enhancing the EOSC Federation.
- The project aims to improve collaboration, integration, and innovation across scientific disciplines through improved technical infrastructure, interoperability, and automated processes.
- **FTS (File Transfer Service)** will provide data transfer solutions within the EOSC platform, to ensure seamless data sharing across European data spaces.
- FTS's participation in EOSC will not only support the platform but also **enhance the FTS software** itself by improving its **documentation, ease of use**, and potentially **expanding its reach to a wider range of scientific communities**, including the long tail of science.



home.cern