

To the attention of: Simone Campana, Ian Bird, WLCG Coordination
For distribution to: WLCG Management Board
Date: 27 Jan 2020

EGI Call for Collaborations

Project proposal to H2020 INFRAEOSC-07 (A1) Call “*Distributed and cloud computing resources*”

27 Jan 2020

CONFIDENTIAL

EGI invites research communities and industry/SMEs to collaborate with EGI in the context of a project proposal for submission to the H2020 project call: INFRAEOSC-07 subtopic A1¹.

Collaborating organizations are invited to provide use cases to become recipient of free-at-point-of-use dedicated large-scale federated compute and storage capacity and additional adding value services for the secure storage, processing, analysis of sharing of raw data and data products. Organizations will be offered the possibility to host their community-specific data analytics platforms in a federated hybrid compute and storage platform that is interoperable, standard compliant and open to users of the European open Science Cloud (EOSC).

Collaborating organizations will have the choice to directly contribute to the implementation of this large-scale distributed data-centric computing facility by federating in-house compute and storage capacity, data or applications as applicable. These will be federated to EOSC through the EGI federating services and promoted on the EOSC Portal.

| | |
|----------------|--|
| Contact e-mail | projects@mailman.egi.eu |
| Deadline | Friday 28 February |

¹ H2020 INFRAEOSC-07A1 Call submission deadline: 22 April 2020, EC Budget: 8 Million Euro (<https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/infraeosc-07-2020>)

Objectives of the Collaboration

EGI is opening a call for research collaborations and commercial organizations interested in advanced data-driven computing, cloud and compute-intensive computing applications, and big data provisioning, with the purpose of building a community of thematic data and service providers that through the EGI Federation can scale up their in-house IT infrastructure to open the access to data and thematic applications to new user groups of European and international relevance.

The organizations interested in entering this collaboration [the Partners] will be recipient of a free-at-point-of-use dedicated large-scale federated facility collectively provided tens of petabytes of storage and tens of thousands of computing cores for the secure storage, processing, analysis and sharing of raw data and data products.

The Partners will be offered the possibility to host their community-specific data analytics platforms in a federated compute platform that is open to users of the European open Science Cloud (EOSC).

The collaboration is particularly targeted to research communities, commercial operators and e-Infrastructures of pan-European and international interest.

The ultimate goal is to promote open science and innovation through the federation of hybrid HTC, HPC and Cloud infrastructures of national, regional and international relevance. The federation will be enabled by operationally ready exabyte-scale data management and compute orchestration technical solutions, that will be evolved during the project in collaboration with the Partners.

The collaboration will be in the context of a project proposal that the EGI Community is preparing for submission to the H2020 call INFRAEOSC-07 (subtopic A1). The project will address the EC purpose to: *“effectively coordinate at pan-European level the provision through the EOSC Portal of state-of-the-art research enabling services from a wide range of national, regional and institutional public infrastructures in Europe, covering diverse thematic domains, and further non-research resources in order to: 1) scale up the EOSC Portal; and 2) set-up a model for interaction between service providers and the EOSC Portal operators through pan-European e-infrastructure entities, based on transparency and effectiveness of cost compensation.”*

The call addresses the provisioning of “distributed computing and Cloud provisioning services for exploitation in the context of the European Open Science Cloud (EOSC)”. The EGI project proposal leverages 15 years of experience of the EGI Community in providing high performance data processing and analysis solutions for large research collaborations.

The EGI Federation

Through the project, EGI will deliver a dedicated federated HTC/HPC and Cloud compute and storage platform, that will collectively provide tens of petabytes of storage and tens of thousands of computing cores. This hybrid platform will be provided to support a range of different computational workloads from various disciplines and user groups. The facilities and services provided will enable users to process and analyse data in a distributed computing environment. The role of the Partners in the project will be to bring use cases, exploit the infrastructure services offered by the EGI Federation, and manage and operate domain-specific core data resources, disciplinary tools and other open science resources to be contributed to EOSC.

Through the project funding, today's capacity of the EGI Federation will be expanded to procure, operate and support dedicated service capacity to EOSC user communities.

The federated compute environment will offer data hosting capabilities, and secure access across national and organizational boundaries to data and compute facilities and will be complemented by adding value services for compute and data management.

Standard APIs for cloud management will be supported to enable remote compute and storage access, complemented by an integrated Cloud repository of community virtual machine images supporting the registration, sharing and discovery of images, contextualization tools, and federated VM management tools.

The envisaged 'distributed and cloud compute infrastructure' will be realized in a collaborative manner with the Partners, as a fully integrated set of capabilities organized by Tier: different collaborations and deployment models will be possible, as detailed below.

- **The Federating Tier** will provide services and activities for the management and supervision of the services connected to the federation. The Federating Tier will be enabled by a set of mature federation services that have been operated and evolved in the course of 15 years of federated operations in EGI.
 - → Partners willing to integrate their compute platforms are invited to join the project as integrators. The project will sponsor the provisioning of the federating tier and standard procedures and policies to manage services in the federation.
- **The Resource Tier** delivering IaaS/HTC/HPC for computing and data storage. The resource tier is implemented by EGI members and Partners who decide to join the project as integrators extending the existing EGI Federation, which today comprises more than 250 data centres worldwide.
 - → Partners are invited to propose use cases for the exploitation of the resource tier capacity.
- **The PaaS Tier** will deliver data transfer, data caching and workload orchestration services. These services will be managed by technical partners of the EGI Community and provided as managed services to the Partners to be integrated in their workflows.

- → Partners with relevant use cases will receive technical support to operationally integrate these services with their workflow.

Partners entering the collaboration will be key actors in the project and will have the task to deliver a rich portfolio of data, tools and data analytics that provide adding value capabilities to end-users. When possible, these will be promoted through the [EOSC Portal for broader exploitation and support to Open Science and Innovation in EOSC](#).

Benefits for the Partners

Partners entering a collaboration with the project, can play a dual role: of consumer and/or provider.

1. **Consumers.** The project will provide sponsored access to dedicated hybrid HTC, HPC and GPGPU compute and storage resource pools, and Federation and PaaS managed services free-at-point-of-use for integration in user community workflows. The offer of the project includes a number of integrated capabilities as detailed below. All will be accessible in a secure manner through AARC² compliant AAI services.
 - An interoperable **hybrid federated compute platform** offering computation services and capabilities for secure data management in a federated environment, which provides harmonized operating procedures and quality standards.
 - **Co-located storage and compute facilities**, and cloud management capabilities for the storing and processing of large data that cannot otherwise be efficiently analyzed through the download to local IT facilities. This free-at-point-of-use distributed facility, will be complemented by pay-for-use Cloud IaaS services, that will be offered by commercial providers interested in integrating their facilities with the EGI federation.
 - **Data and application hosting** services, complemented by federated data management capabilities to securely stage and process data in the federated facility, plus cloud management capabilities for VM and container-based computing.
 - **Personal workspaces** for secure storage, sharing, processing and analysis.
 - **Orchestration** services for the management complex workflows that require access to distributed data and HTC, HPC and Cloud facilities.
 - **EGI Notebooks**, a centrally managed Jupyter Notebook browser-based tool for interactive analysis of data using the data and computing resources from the Resource Tier.
 - A managed **third-party data transfer** service for the staging of data. The data transfer service will provide robust scheduling for data centre to data centre data transfers in order to move large numbers of files or very large files.

² <https://aarc-project.eu/architecture/>

- An **AAI proxy service** (based on the EGI Check-in solution) for user communities interested in a managed solution for user enrolment and group management, and for service providers willing to easily integrate with external identity providers.
 - **Federation services** to connect HTC, HPC and Cloud providers to EOSC through harmonized operating procedures, policies and service management quality standards. This includes facilities and assets like: harmonized security policies, tools for security management, the central gathering and storage of accounting and monitoring information.
 - **Technical support and training** for co-design and co-development activities.
 - Access to **project meetings and technical workshops**.
2. **Provider.** Research communities who successfully integrate data, tools and other relevant services in the federation with the service tiers of the project, will be supported to join EOSC as supplier and/or integrator of thematic compute services and datasets. Thematic services should be operationally ready, fully supported and showing technology readiness level 8 or higher at the start of the project (Jan 2021).

Requirements

- The Partners will be asked to provide use cases and to collaborate with the project consortium to assess their e-needs. This includes the definition of service targets and of a capacity plan for the duration of the project. IT experts from the Partners will be involved in co-design and technical support activities.
- The Partners will be required to enter a collaboration agreement with the project and will be invited to join the project consortium as applicable (see below for details).
- Thematic services supported by the EGI Community in this project, need to meet Technology Readiness Level 8 or higher at the start of the project. Data will need to meet applicable FAIR policies and align to community data management best practices.

Services sponsored in the project will be managed in compliance with industry standards (ISO 9k and ISO 20k) in accordance with the EGI Service Management System. In particular, the service levels committed by the project will be documented and the delivered performance will be periodically documented for feedback from the Partners.

In order to support international research collaborations beyond Europe, the project welcomes Collaboration Agreements that involve non-EU e-Infrastructure providers and organizations

Funding

All Partners who agree to enter a collaboration agreement with the project will receive a travel budget within the limits of the project funds, to attend project events, hackathons, technical workshops and training events.

Funding to Partners will *not* include person months for community-specific activities and technical developments that concerns the operations and integration of their community services.

Partners that are a member of EGI, or decide to enter the EGI Affiliation Programme, will receive additional budget to support one or more of the following activities: thematic service operations, technical integration, user support and training towards community-specific end-users.

Funding to Partners will be subject to the H2020 financial rules defined by the European Commission.

Participation to the Project

A Partner committing to enter a collaboration agreement with the consortium will be added as third party to the project.

Alternatively, a Partner that is a member of EGI or is committed to become a member, will be invited to join as beneficiary or linked third party, as applicable.

Contact Information

Partners interested in a collaboration or in getting more information, are invited to formally express their interest to the EGI Foundation in writing as specified below.

| | |
|---------------------------|--|
| Contact e-mail | projects@mailman.egi.eu |
| Deadline for applications | Friday 28 February |

Applicants are invited to specify “INFRAEOSC-07 A1 Collaboration” in the subject of the e-mail.