EGI-ACE Open Call No. 1

Checkpoint meeting with Shepherds

Scalable Jupyter backends to complement data holdings in the CS3mesh4EOSC mesh

Zdeněk Šustr / CESNET

Dissemination level: Internal

Disclosing Party:

Recipient Party:

Outline

- *Background about the scientific use case*
- *Ambition, Impact and Challenges*
- *Integration Support*
- *Capacity Requirements*
- *Timeline*

Background about the scientific use case

CS3MESH
- Infrastructure rather than science (Computer science?)
- Initiative to facilitate trust between sync'n'share providers based on the addressee's email address and nothing else.
  - Along federation principles but **not** a federation
  - No authorization attributes obtainable
- Invitation to share
- Identification of a user in their home system, setting up a trust in the sharing system

Ambition, Impact, Challenge(s)

Quoting from the Submission form:

Scientific objective(s): *Starting from scientific data already held in the node of the cs3mesh4eosc consortium, we want to ensure that users can select the correct VREs and scientific workflows to operate on this data, seamlessly. Given that EGI is the eInfra partner for the operation of the VRE instance of many ESFRIs, it stands to reason to try to connect the CS3mesh4EOSC data holdings with a generic test case for EGI's method of running VREs; we think Jupyter Notebooks would make an excellent candidate.*

Computing objective(s): *Connect OCM/Reva fronted EFSS stores to Jupyter backends; ability to select desired VRE toolchain by distributing kernels+environment via cvm-fs; possibility to take jobs via one jupyter frontend but scale them out via EGI's DODAS.*

Integration Support

Unclear!

Quoting again from submission:

federated auth harmonised between CS3mesh4EOSC and EGI identities. As CS3mesh4EOSC uses EduGAIN-type mechanisms, and EGI uses Checkin, SAML-based methods from the AARC2 cookbook seem obvious for a first attempt.

- The principal investigator believes they will be able to work around Check-in

Capacity Requirements

- access to a Jupyter notebooks cluster with at least 200 cores

- a minimum of 5 co-located VMs (4CPU, 8 GB min) that can run interoperability middleware (to be experimented with between EGI and CS3mesh4EOSC)

  This is routine in terms of resources, but the actual usage scenarios are uclear

- Not explicitly states in proposal, but **potentially unrealistic requirements in terms of effort!**
  - To be clarified

Timeline

Current status: formulation of actual requirements
- We were finally able to meet with a representative versed in the AAI setup of CS3MESH
- **Request for description of usage scenarios**, which are unclear

The future remains TBD

Submitter happy to spend the allotted time with designs

Thank you!

Contact: **egi-ace-po@mailman.egi.eu**
Website: **www.egi.eu/projects/egi-ace**

EGI Foundation

@EGI_eInfra