EGI-EUDAT ENES Meeting (2017-09-27)

https://indico.egi.eu/indico/category/165/

Agenda

- Resource needs
- DI4R conference
- Final deliverable
- AARC
- Status update
- Open actions

Present

Michaela Barth (PDC)
Ute Karstens (ICOS)
Christos Kanellopoulos (AARC, Geant)
Hans van Piggelen (SURFsara)
Matthew Viljoen (EGI)
Xavier Pivan
Peter Gille (PDC)

Apologies

Jens Jensen (STFC)

Information

- Resource needs: 500 GB for each of B2SAFE and B2STAGE
  - STFC resources: Can not upload there yet, because of self-signed certificate problems. Recommended to use a more general certificate instead
  - CINECA:
  - When: late October would be good.
- https://www.digitalinfrastructures.eu/ DI4R 30th November - 1st of December Deadline for Abstracts Oct 13th, focus on concrete achieved results, Early Bird deadline will be still in October, Michaela plans to go there, input from EGI?
  - Site note: Michaela will also be there for "Nordic Tier1 as a sustainable model for distributed infrastructures"
- EGI DataHub updated to its new version, should deal with small files in a much better
  - splitting up functionality in two ways:
    - aggregating and searching for data
    - bring your data to VMs for computing (ICOS has requested) <- called EGI Federated Data Manager, as a separate service, but functionality accessible through DataHub

- Final deliverable EUDAT Task 7.2: **due by PM 34 (December 2017)**
- AARC:
  - Accessing services and resources in a scalable and secure way
  - Providing guidelines for infrastructure operators
  - EGI-EUDAT Pilot ongoing: can provide input into AARC
  - Questionnaire (1hour) has been prepared (for gathering requirements) and is almost complete and could be circulated. -> Determine what type of solution could be the best one for this community, not to technical
  - Used as input for common paper, which will act as a starting point for activities later
  - Intended workflow: access portals allowing to submit automated workflow
  - EPOS already part of AARC project
- ENES workflow uses Java API from EGI to create VM, Data at EUDAT (currently: test at ESGF) is accessed through EGI. Also want to download data directly from EUDAT
- ENES status update: Creation of VMs through docker container is fully working. Generates the proxy inside of the container to access EGI, workful until VMs are ready for computing complete.
  - GEF backend on EGI still needs more cleaning and automation
  - working on ESGF API to start calculation on GEF (just started this week)
- No news on PDC testbed: still waiting for hardware
- Hans will give an update on EUDAT documentation on EGI website soon

Actions

- AP Michaela: go back to CINECA with storage numbers
- AP Michaela: structure final deliverable and ask for specific input
- AP Michaela: give AARC access to ICOS user steps documents and walk them through it.
- AP Michaela: give AARC access to EUDAT presentation with figures
- AP Xavier: produce an updated scheme figure on the workflow
- AP Ute, Xavier and EPOS: fill out questionnaire once it is out
- AP all: provide input to suggestested abstract for DI4R conference until next week ICOS meeting, October 4th.

AOB

Next ENES meeting: Wednesday 14:00 CEST, 18th of October
(Next ICOS meeting: Wednesday 10:00 CEST 4th of October)

Suggestion for Abstract to D4IR:

EGI-EUDAT joint access to data and computing resources recapitulation

EGI-EUDAT interoperability collaboration started in 2016 with the goal to harmonise the two e-infrastructures. In order to create seamless access and pairing data and computing resources together into one perceived infrastructure offering both EGI and EUDAT services, user communities have been identified and selected to bring in their requirements on technical interoperability, authentication, authorisation and identity management, policy and operations.

After the definition of a universal use case, this end-user driven approach continued and EGI and EUDAT worked closer together with EPOS, ICOS and later also ENES communities as usecase pilots to test-drive the cross-infrastructure usage of the storage resources managed by EUDAT and the computing resources available through EGI and to finally also validate the results.

During the work the user communities continued to set the priorities right for the aspects most important to them: the user communities put much more emphasizes on automated approaches and quality of end user documentation than foreseen in the beginning. Involving the user communities in this way also meant a sometimes steep learning curve on the technological understanding and effective communicate using the right termes, which presented a major - but necessary - time investment from their side. This time and trust investment had to be properly administered by not misusing them as free beta-users for not production-ready and largely undocumented new features, that were eagerly put forward. Concrete outcomes of the soon finished work include i.e. valuable feedback on data-handling support within the EGI-DataHub and testing to use the EGI Federated Cloud with automatic submission, data transfer tests between the VMs and B2STAGE instances using both OneData and EGI DataHub to access a common storage for several VMs and evaluating the new B2STAGE HTTP API.

This presentation will show the final established design of the workflows for each of the followed user communities, especially adapting to their specific needs and also cross-fertilizing each other.