

8.2 Fusion (CC)

| | |
|-------------------|--------------------------------|
| Short description | Fusion CC |
| Type of community | Competence Centers |
| Community contact | Shaun de Witt Andrew Lahiff |
| Meetings | |
| Supporters | |

Ambition

The CC's ambition is to assess whether the services provided by EOSC are suitable for use cases within the fusion community. This work has been split into two; one storage specific and one compute specific. The reason behind these investigations is in preparation for ITER data handling and analysis, which represents a major technological challenge for the fusion community increasing the volume of output by three orders of magnitude from current experiments.

For storage we wish to investigate replication between sites. Since ITER is an international experiment it is likely there will be at most two European sites which will host a fraction of the data and some portion of that data will need to be readily available for analysis at several centres of excellence. Other sites may wish to access the data, but in this case the analysis is not time critical and so are not considered here. It is envisaged that automated replication of data will be key to this work, but will require the underlying technology to support high speed IO and replication. As this is primarily a technology assessment, we have specifically omitted security implications. However, if a suitable EOSC technology is identified then this will need to be taken into consideration for final usage.

In compute terms, the CC is again being driven by the needs of ITER. It is not anticipated that one single site will be able to meet the needs of ITER data analysis and, indeed, pre-testing. One partner has already demonstrated a service which allows modelling code to be run at any site with available (and suitable) compute resources. This work needs to be extended to support ITER type operations; specifically execution of full workflows. As such we are taking a two fold approach; running an existing 'real life' use case from the MAST tokamak, taking raw output from one of the diagnostic tools and processing it to science products, and also using the ITER Integrated Modelling and Analysis Suite (IMAS) to test prototypical ITER workflows. The idea is that making use of cloud resources will better allow sites to process 'intershot' data at a scalable level and maintain a smaller ecological footprint than would otherwise be necessary.

User stories

- Ambition
- User stories
- Use cases
- Architecture & EOSC-hub technologies considered /assessed
 - The two use cases require two testbeds:
 - EOSC-hub technologies assessed
 - DODAS
 - INDIGO PaaS Orchestrator
 - OneData
 - B2SAFE
 - EGI FedCloud
- Requirements for EOSC-hub
 - Technical Requirements
 - Capacity Requirements
- Validation plan



Instruction

Requirements are based on a user story, which is an informal, natural language description of one or more features of a software system. User stories are often written from the perspective of an end user or user of a system. Depending on the community, user stories may be written by various stakeholders including clients, users, managers or development team members. They facilitate sensemaking and communication, that is, they help software teams organize their understanding of the system and its context. Please do not confuse user story with system requirements. A user story is an informal description of a feature; a requirement is a formal description of need (See section later).

User stories may follow one of several formats or templates. The most common would be:

"As a <role>, I want <capability> so that <receive benefit>"

"In order to <receive benefit> as a <role>, I want <goal/desire>"

"As <persona>, I want <what?> so that <why?>" where a persona is a fictional stakeholder (e. g. user). A persona may include a name, picture; characteristics, behaviours, attitudes, and a goal which the product should help them achieve.

Example:

"As provider of the Climate gateway I want to empower researchers from academia to interact with datasets stored in the Climate Catalogue, and bring their own applications to analyse this data on remote cloud servers offered via EGI."

| No. | User stories |
|-----|---|
| US1 | <p>The Fusion CC wish to demonstrate making use of EOSC computational and storage resources for running containerised modelling applications (primarily HPC and HTC). This requirement derives from the fact that local resources are not scaled for peak demand and we wish to use the infrastructure provides by EOSC (and public cloud providers) as a scalable, non vendor specific resource.</p> <p>At a high level, this is an opportunistic use case where we wish to make use of any spare resources at sites, and thus going through an ordering process would be non optimal since the user would not know local resources are exhausted until they submitted their job. It maybe that some sort of framework agreement would be needed between the community and the sites to allow this opportunistic use beyond the small number of cores already presented through EGI.</p> <p>Different parts of the workflows may involve different computational requirements, from simple single core machines to many core/multi-node. In the first case we would request resources on a single site, but only instantiate the number of machines required for a specific element of the workflow. It is desirable to develop this further to allow the instantiation of machines for different parts of the workflow to meet the requirements of each step. We are also interested in using both using traditional workflows and workflows within the ITER Integrated Modelling and Analysis Suite (IMAS) which is anticipated to become the standard framework for both modelling and analysis work in the future.</p> <p>In most cases the steps of the workflow communicate through files, with each stage producing it's own unique file which acts as an input to the next stage. While running at a single site, it would be possible to request storage at that site of appropriate size. However, in the desirable case of different stages of the workflow running at different sites, either storage system accessible to many sites will be required for these intermediate files, or they would need to be transferred between sites.</p> <p>Final output data (and possibly intermediate data) should be accessible to the end user.</p> |

| | |
|-----|--|
| US2 | <p>As a site data manager I am looking at how we can improve user and computational access to experimental data, in addition to being driven to allow more open access to users beyond the fusion community by either national or international funding bodies (or both!). However, in common with most science disciplines my site places an embargo on experimental data to allow researchers to publish. In addition, some data will not be made public where it has no scientific value (engineering tests for example), or where work is done on behalf of industry. Significant analysis work is performed on the MARCONI/Fusion supercomputer based at CINECA and for data sets which will be accessible it would be beneficial to my users if data could be hosted there. In addition, we want to offload public data to partner sites for hosting and access to the wider science community and general public, so that data used by the fusion community is kept on site and only accessed by fusion users. This, combined with the restricted roadmap for tape technologies is pushing me towards replication as a means of bit preservation in the longer term. the community already has a data access mechanism (UDA) which it uses and must be usable at each site where the community will access data. General access will be via HTTP.</p> <p>Thus as a site manager I would ideally like to put a full copy of my data on a trusted site which will prevent unauthorised access to the data (although not the metadata) during the embargo period but will make it accessible following that. That site should be able to provide me with data download statistics on an annual basis as part of my reporting to senior management and fundholders. Additionally, I would like that data to be copied from the trusted site to CINECA so it can be used optimally for analysis on the MARCONI computer and I would like a third copy to ensure there are four copies on my data availability for high availability (one local and three off site copies).</p> |
| | |

Use cases



Instruction

A use case is a list of actions or event steps typically defining the interactions between a role (known in the Unified Modeling Language as an actor) and a system to achieve a goal.

Include in this section any diagrams that could facilitate the understanding of the use cases and their relationships.

| Step | Description of action | Dependency on 3rd party services (EOSC-hub or other) |
|------|---|---|
| UC1 | On submission of a containerised workflow, sufficient resources are provisioned on a remote site(s) to allow execution of that workflow. The users home credentials (or the services services credentials) must be accepted on the remote site(s) | Orchestrator/Kubernetes EGI Fed Cloud/Other cloud Suitable AAI mechanism (note the fusion community does not have a centralised IdP or AAI system, but each site provides its own authentication based on username/password) PROMINENCE cloud execution service |
| UC2 | During workflow execution, intermediate files should be written to a location which will be accessible to later stages of the workflow. The final results should also be accessible at the users home institute | EGI FedCloud/B2DROP integration based on either user or service credentials |

| | | |
|------------|---|--|
| UC3 | Data hosted off-site shall be accessible to code using the fusion Unified Data Access (UDA) middleware so that the same code can be used to access data regardless of locality. | B2SAFE with three way replication (preferred sites would be CINECA(for the MARCONI link), STFC (geographically close) and PSNC (another fusion site running this service)). Will require suitable AAI mechanism and integration with UDA. |
| UC4 | Any attempt to access data hosted off site should determine whether the data is 'open' or 'embargoed'. In both cases, users should authenticate themselves to allow traceability of who has accessed the data. In cases where the data is embargoed, the hosting site should deny access to unauthorised users. | B2SAFE and embargo periods /OneData |
| UC5 | Data placed at an offsite location should be replicated to CINECA and at least one other partner site within the fusion community | B2SAFE/OneData |

Architecture & EOSC-hub technologies considered /assessed

The two use cases require two testbeds:

1. cloud federation with opportunistic use:
 - a. ... nb of sites
 - b. characteristics of sites
 - c. ...
2. storage sites, linked to CINECA
 - a. ...
 - b. ...

EOSC-hub technologies assessed

DODAS

This has undergone a paper based assessment and found unsuitable for user needs

INDIGO PaaS Orchestrator

In use. Have also tested Kubernetes which is both easier, more feature complete and more widely used. Lack of integration os Orchestrator with other components both in Indigo and EGI has caused issues, but the developers are very responsive

OneData

Under test

B2SAFE

Under discussion

EGI FedCloud

Heavily used. The real limitation is the access model being imposed. We don't want to 'order' CPU time, we just want to make opportunistic use of it. This would be akin to 'pre-emptible' VMs where the process could be killed if needed for other work. the currency order based system would mean I would need to ask for 100 cores at each site but I could not guarantee filling them.

(One issue encountered so far is the availability of licenses for commercial software at sites. Several existing workflows make use of IDL (<https://www.harrisgeospatial.com/Software-Technology/IDL>) and NAG (<https://www.nag.co.uk/content/nag-library>). Currently UKAEA has permission to use the NAG libraries within containers, but no solution has been found for IDL. We are currently working on replacing the IDL calls with suitable python calls.

Requirements for EOSC-hub

Technical Requirements



Instruction

- Requirement number: Use numbers RQ1, RQ2, RQ3, ...
- Requirement title: Use a short but descriptive title. Use the same title in the Jira ticket 'Summary' field
- Link to requirement JIRA ticket: Open a ticket in <this JIRA queue <https://jira.eosc-hub.eu/projects/EOSCWP10/issues/EOSCWP10-4?filter=allopenissues>> (click on 'CREATE' button in the middle-top of JIRA)
- Source use case: Refer back to the use cases above (UC1, 2, ...)

| Requirement number | Requirement title | Link to Requirement JIRA ticket | Source Use Case |
|--------------------|---|--|-----------------|
| RQ1 | Storage requests for WP8.2 Fusion CC |  EOSCW P10-44 - Jira | UC5 |
| RQ2 | Provide a homogenised AAI for single sign on access to all services |  EOSCW P10-62 - Jira | UC4 |
| RQ3 | Support Opportunistic usage of cloud computing |  EOSCW P10-73 - Jira | UC1 |

Capacity Requirements

| EOSC-hub services | Amount of requested resources | Time period | Status |
|-------------------|--------------------------------------|-------------|--|
| B2SAFE | 10TB at 3 sites (CINECA, STFC, PSNC) | 6 months | In progress. CINECA and STFC are now available; PSNC still have issues. Meeting to discuss replication to be held later this week. |

| | | | |
|--------------|-----------|----------------------|---|
| EGI FedCloud | 10 cores | 6 months irregularly | Access to 'open' resources (i.e. without going through the ordering process) makes this service difficult under current EOSC rules. We are able to make use of JSC, but resources available are very limited. |
| EGI FedCloud | 100 cores | 2 months | Not started. This will be needed for heavy duty workflow testing |
| OneData | 10TB | 12 months | CEA and PSNC have installed latest version of OneData which is a significant improvement in terms of installation and maintenance. Initial testing at CEA has shown very poor upload performance. Replication between CEA and PSNC seems to work, but not in the reverse direction. |

Validation plan

Storage: (OneData and B2SAFE)

- Test upload/download speeds to ensure they can meet minimum requirements of 20 MB/s
- Test replication between sites; replication ideally should be asynchronous, but reports will take into account synchronous replication
- Test failover/recovery by deliberately corrupting a copy of the data at one site. Requests to access the data should either fall back to a good replica (preferred) or cause a well defined error (acceptable). Eventually the corrupt replica should be updated automatically from one of the good replicas.
- Test data is accessible from MARCONI/FUSION at CINECA for later work

Compute:

- Ensure a MAST intershot workflow is able to run on EOSC resources within the minimum intershot interval (currently 20 minutes)
 - this should be tested at several sites to take account of different hardware configurations.
- Ensure a prototypical IMAS workflow is able to run successfully on EOSC cloud resources
 - Since there is no real 'intershot' concept in ITER, the data needs to be processed in NRT and returned to the user