INAF participation in OBELICS T3.4.2

INAF tasks in T3.4.2

- Authorisation, Authentication and Accounting:
  - analyse existing requirements and protocols
  - propose and implement, as required, a global infrastructure using agreed-upon standards
- Workflow architectures for the orchestration of compute-intensive data analysis on distributed computing infrastructures
- Liaison and coordination with WP4 (DADI)

INAF T3.4.2 team

- Contributed INAF staff (30 person-months):
  - F. Pasian (OATs/IVOA/Euclid): coordination
  - M. Frailis (OATs/Euclid): workflows
  - C. Knapic (OATs/SKA): A&A
  - M. Molinaro (OATs/IVOA): coordination with DADI
  - G. Taffoni (OATs/IVOA/EGI): A&A, workflows
- 2 additional persons to be hired:
  - 28 person-months on A&A (from April 2016)
  - 21 person-months on workflows (from Sept 2016)

INAF Background

- **IVOA** 2002 F. Pasian, G. Taffoni, M. Molinaro, C. Knapic
- **IA2** 2008 C. Knapic, M. Molinaro
- **SKA** 2013 C. Knapic
- **Erflow** 2010 G. Taffoni
- **StarNET** 2013 G. Taffoni, C. Knapic
- **EGI** A&A 2003 G. Taffoni
- **Euclid** 2007 F. Pasian, M. Frailis

Authentication and Authorization

Scope:

- Authentication: a process by which you verify that someone is who he claim to be.

- Authorization is the process of establishing if the user (who is already authenticated), is permitted to have access to a resource

Users:
Researchers, developers, projects …. But usually customized in house.

The Federated Identity approach

Federated identity managements allow registered users of a certain Institutional domain to access information from other Institutional or trusted domains in a smooth way without having to provide any extra administrative user information:

- Gives a delegated mechanism to manage user identification among different entities and within different subjects;
- Provides a set of attributes to an authenticated users to be used by the final application.
- Advantages: Keep your credential at your institute/company always updated!

Federated Identity is...web oriented. Some technical approaches.

- Designed and developed for services consumed via WEB (e.g. web services, portals, clouds);
- May I access my local computing cluster? Yes but with other technologies (e.g. x509 and ldap and meta-users and ssh (PRACE))

- Technological solutions available are not interoperable:
  - OAuth (Open Authentication)
  - Security Assertion Markup Language
  - OpenID
  - X509

Authorization

Traditionally, identity federations have solved the authorization problems with two opposite approaches:

- Service managed authorizations
- Identity providers managed authorizations

Trend in projects and infrastructures currently is:

- ➔ take care of your own authorization,
- ➔ Identify your own policies,
- ➔ Choose an implementation.

You know your requirements, you develop your AuthZ.

But please do not reinvent the software!

Some technical approaches

- Group Management System (GMS) developed by CADC
  - IVOA compatible;
  - Centralized groups, roles and permissions
- Grouper:
  - Centralized groups, roles, and permissions
  - Delegated control
  - Provision to LDAP/SAML etc.
  - Auditing

https://spaces.internet2.edu/display/Grouper/Grouper+Wiki+Home

- LDAP
- Local relational database
- ....

Astronomical use cases 1/2

Status of Obelics T3.4.2 investigations on some ESFRI projects use cases (others are to be done) :

- **SKA** : Authorization and Authentication is under discussion
  - Federated access to resources / self registration;
  - Grouping service.

- **IVOA** and **EuroVO** : Recomandations for:
  - Single Sign On, Credential Delegation;
  - Authorization under discussion;
  - SSO recommendation “is a profile against existing security standards”. No authentication required. If any: HTTP Basic Authentication, Transport Layer Security (TLS) with passwords, Transport Layer Security (TLS) with client certificates, Cookies, Open Authentication (Oauth), Security Assertion Markup Language (SAML), OpenID

Astronomical use cases 2/2

- **CTA** : Authorization and Authentication is under discussion
  - Different approaches to integrate
    - Ldap;
    - web portals;
    - X.509;
  - UNITY;
  - GROUPER;
- **EUCLID** : Currently A&A foreseen mechanism will be provided by ESA
  - SAML based Authentication ;
  - Custom based Authorization ;
  - Peer to peer mechanism using certificates for computing purposes.

Investigation on EUCLID digital identities management via Federated approach, no actions on Authorization.

Aims and plans of A&A sub task activity

- Investigate general approaches, trends and best practices for A&A;
- Collect ESFRI projects requirements;
- Collect ESFRI projects use cases;
- Analyze ESFRI projects technical solutions, prototypes and activities;
- Contribute to implement the most flexible solution common to the ESFRI projects issues.

Workflows

Scope:
Workflows have emerged as a paradigm for researchers to formalize and structure complex scientific experiments in order to enable and accelerate scientific discoveries

Users:
- Researchers
- Projects (orchestrate high level pipelines and infrastructures)
- Science Gateways

Workflows applications

- Workflows as tools for projects infrastructure:
  - Data acquisition/reduction/analysis;
  - Orchestrate tasks and resources (HPC, HTC, Storage, etc);
  - Macro/Micro pipelines;
- Workflows to support researchers and hide the complexity of computing and storage resources
- Workflows to develop science gateways
- Etc…

Workflows in Astronomy

More than 50 workflows management systems (engines)

- Workflow4Ever project based on TAVERNA
- ESO Recipe flexible execution workbench (Reflex) based on Kepler
- ER-flow project based on gUSE/WSPgrade (able to execute and mix workflows written for different WMS)
- Pegasus used by XSEDE, able to execute tasks on DCIs
- CyberSKA project web based workflow builder that supports image segmentation, image mosaicking, spatial reprojection, and plane extraction from data cubes
- Many others….

Aims and plans of workflows sub task activity

- Identify most common workflows tools;
- Collect requirements and experiences from projects;
- Collect workflows repositories and WMS;
- Produce a working prototype if needed and requested;
- Investigation on Use cases of large projects;
- Requirements for large projects;
- Knowhow on Workflows, Gateways and SSO Protocols and Prototypes;
- Technology evaluation and testing.

Coordination with DADI (WP4)

Data Access, Discovery and Interoperability: builds on topics in common with OBELICS on a different layer in the provider-consumer scenario.
AAA, server-side data processing and access to big data are already established fields of work for DADI (discussed at both the Tech Forum [Sep. '15] and ESFRI Forum [Dec. '15], as well as foreseen at Tech Forum II [Mar. '16]).

OBELICS-DADI coordination/interface aims at

- Avoiding effort duplication
- Assuring interoperability among the solutions/developments arising from the two ASTERICS packages
- Providing a means for WP3 requirements to be taken into account in WP4 activities
- Providing a means for WP4 comments/constraints to feed back WP3 activities