PaN-data ODI

Deliverable D3.3

Production deployment of AAA infrastructure

Grant Agreement Number	RI-283556
Project Title	PaN-data Open Data Infrastructure
Title of Deliverable	Pilot deployment of initial AAI service infrastructure
Deliverable Number	D3.3
Lead Beneficiary	PSI
Deliverable Dissemination Level	Public
Deliverable Nature	Report
Contractual Delivery Date	31 Mar 2013 (Month 12)
Actual Delivery Date	31 Jul 2013

The PaN-data ODI project is partly funded by the European Commission under the 7th Framework Programme, Information Society Technologies, Research Infrastructures.

Production deployment of AAA infrastructure
PaNdata / WP3

Draft

Version 1.0, hjw Jun 28, 2013

Rev Jun 29, 2013

1 Introduction

Since the start of the Umbrella project [1] interest in federated identity services has steadily increased. In the academic sector it is covering practically all research activities involved in activities producing large amounts of data[1].

Specifically, the large research facilities for the photon and neutron user community experience an increased need for transnational and trans-facility services, especially within the IT sector. Users perform their experiments at several facilities by exploiting special experimental potentialities; the data gathered at these facilities have to be provided to the researchers for analysis; researchers need remote access to their experiments while they are performed. All these services need federated identity management as a basic component. Support of the users at these facilities is done via web-based user office (WUO) systems. These services are running since a long time, are very efficient, but they are local. The added value of the Umbrella concept is to offer an additional layer on top of these services, which provides the missing functionalities.

2 Umbrella essentials

In this document Umbrella can be presented only in a nutshell. More detailed descriptions of the Umbrella concept can be found elsewhere [1-3].

The main goal of Umbrella is to provide a *unique and persistent identity* for the users of the large photon and neutron facilities. For that, Umbrella has *one single identity provider* (IP). This IP has just enough capacity to identify a user. All other authentication information and all authorization information reside with the local WUOs (*hybrid data structure*). As a service to the users, Umbrella contains a *synchronization mechanism*, so that new information has to be input only once and is then distributed to the other WUOs, where a user is registered. This loose identity data structure also eases the management of the whole FIM concept. Facilities can join or leave with minimum impact on the other facilities. Also there is no need for a ‘big-bang’-type deployment of Umbrella.

Handling of the user information is done according to a bottom-up, *do-it-yourself (DIY)* scheme. This lowers the administrative load involved. Even more, as the user triggers him/herself the transfer of his own information, administrative and legal complications are thus avoided.

Clearly, Umbrella also provides *single-sign-on* (SSO) functionality, for a sequence of services a user has to register only for the first one and the identity is forwarded for the succeeding ones.

[1] A good overview is provided from the Federated Identity for Research (FIM4R) meetings [4-7]

Concerning technical realization, Umbrella is built on SAML2 using Shibboleth by the Internet2 Middleware Initiative. A detailed technical description is given in the EuroFEL deliverable 2.5 [2].

Initially, the first developments of Umbrella have been performed within WP2 of EuroFEL [2]. Work then has been taken over in work packages WP3 of PaNdata ODI (implementation and deployment) and WP16 of CRISP (further developments).

3 Implementation / deployment

WUOs are very sensitive tools as they are basically involved in the operation of the large research facilities. For example, they are involved in the control of who is when allowed to enter a facility and its equipment; they are involved in the control of who is allowed to access which experimental data. And as Umbrella is strongly linked to the WUOs, these concerns also hold for Umbrella. In addition, very different members at the facilities are involved in these processes. There are the IT experts; there is the user office staff with high knowledge in user-office workflows but not necessarily in IT; and there are the users with a wide range of IT expertise and finally there are the users who need these functionalities for their research activities but who want to use these services as black boxes.

Implementation / deployment at the partner facilities has started in spring 2013 but because of the complexity of the process it is expected to take the full year till all facilities are on the boat.

3.1 Pre-deployments

As a first step, clones of the actual WUO versions have been upgraded at some of the participating facilities (DESY, DIAMOND, ESRF, ILL, PSI) and linked to the Umbrella central service tool. After some initial tests about two dozen ‘friendly users’ from all over Europe have been invited to play with this system and to feedback their experience. The overall experience was very positive, but it turned out that logout handling had to be worked on. This is a problem known also from other security-critical web-based implementations (e.g. banking).

3.2 Deployment phase

3.2.1 Organization and legal issues

The goal of PaNdata / WP3 is the Umbrella is deployed at all participating facilities. This deeply affects the user operation at all these facilities and, therefore, all activities must be pursued in close cooperation. For that, two teams haven been set up, the Umbrella Management Team (UMT) and the Umbrella Technical Team (UTT). The topics for the UMT have been defined in the Holy Umbrella Implementation / Deployment Document which contains the actual status of the project with the following chapters:

- Legal and administrative issues (e.g. MoU)

- Umbrella physical appearance (physical site(s) of the central server, domain service)
- Central user database
- Local user database
- Affiliation database
- Security issues
- User support
- Procedure for handling further developments
- Umbrella production version
- Implementation steps
- Operational issues
- MoU

These topics have been discussed in several telephone conferences and detailed decisions been made. All decisions were unanimous. Technical issues have been forwarded to the UTT.

Concerning management / organization, the only major topic to be realized as soon as possible is the Umbrella legal framework. For the duration of PaNdata ODI (till March 2013) Umbrella operation is covered. At the latest after that time a legal framework has to be set up, which defines resources and legal responsibilities between the partners. The issue is identified and significant progress is expected till the end of the year.

3.2.2 Wave concept

It has already been stated that the introduction of Umbrella affects the basic user operation at the participating facilities. It would be highly demanding to perform a one-shot transition of all these facilities to Umbrella as 30’000+ users are involved. This would pose an unnecessary risk. Umbrella has been designed from the beginning such, that this is not necessary. Instead the transition proceeds in several waves, where only some of the facilities are concerned. Participants of the first wave are ESRF, ILL, and PSI. Then will come DESY, ISIS, and HZB, then the remaining facilities.

3.2.3 In phases

Because of the implications on the user operation it has been decided to perform the deployment in several phases: This allows keeping the system performance under control in case of unforeseen complications.

3.2.3.1 Phase 0, Software tools ready

This phase encompasses the implementation of Umbrella on the operating versions of the user offices as a next step as up to now the test have been performed on clone systems only. In addition, a GeoDNS service has been installed. This service offers the user a single, global URL but the requests are internally dispatched to different local servers. This offers several advantages; it allows load balancing, it enhances significantly failover capabilities, and, in addition, such a topology reflects much better the federated structure of the Umbrella community.

3.2.3.2 Phase I, Facility experts

In this phase Umbrella is accessible to facility IT experts for system checks and initial tests.

3.2.3.3 Phase II, Friendly users

In the next phase, the system is offered to a selected number (about two dozen) of ‘friendly users’ all over Europe. The users are informed about the current status and future plans of Umbrella and are invited to play with the system and to report possible deficiencies and problems. This phase is very important as in this way the system is experienced from the user view but this with limited risk. For most users it is for the first time that they work with a federated identity system and it is very important to incorporate any teething problems into the final Umbrella version.

3.2.3.4 Phase III, all users

Umbrella is open to all users.

3.2.3.5 Present status (July 2013)

The goal of phase I was reached for the 1st-wave facilities and the project is presently in phase II. Transition to phase III is planned for end July beginning of August 2013.

4 Next steps

The final part of the PaNdata ODI project involves several main work branches and, in order to speed up the process, this will be tackled in parallel as much as possible.

4.1 Full Umbrella operation with 1st wave facilities

There will be calls for proposals at two (ESRF: Sept 1and PSI: Sept 15) of the 1st-wave facilities. This will offer a crucial test of the stability of Umbrella. Users will be able to log in to the local user office systems by means of their Umbrella credentials and they will be able to submit proposals to the standard WUO queues.

4.2 Inclusion of higher-wave facilities

According to the PaNdata ODI DoW, Umbrella will be deployed at all partner WUOs. As soon as operation with the 1st wave facilities is stable, deployment will proceed to the next facilities (DESY, ISIS). After that, the remaining facilities will follow.

4.3 Incorporation of the affiliation database

Umbrella is a federated system; it is designed that it offers a central service, but there is only a small central component just sufficient for unique identification, all other user information resides at the local WUOs. With regard to user-friendliness, however, it is highly desirable that the structure of the user information is the same for all partner facilities. This holds especially for the definition of the home affiliation of a user. With synchronized affiliation definition it will be possible to set up system such that in spite of the federated information structure a user can define or modify his affiliation description only once without a need for multiple registrations. Furthermore, there is a high synergy gain at the user offices, as information updates need to be made at one site only and then be available to all partners. This system is presently under development at the ESRF and will be ready for deployment and incorporation in the standard Umbrella in fall 2013.

4.4 Merging Umbrella with ICAT

The first real trans-facility use of Umbrella will be to use it as base for the authentication within ICAT [8]; the open-source metadata-management system developed under the lead of STFC and designed for large research facilities. The metadata-handling part is already very advanced. Concerning data access control ICAT will be merged with Umbrella for providing authentication information. In addition, coupling to the local WUO information access control will gain the necessary fine-graininess (i.e. is this person positively identified and did the person participate in the experiment which produced the specific dataset). This work is planned for fall 2013.

5 Conclusion

Implementation and deployment of the Umbrella system is well on track. A detailed deployment plan, consisting of four phases, has been developed and deployment of the basic tools is proceeding step by step. All components are planned to be in place a the end of the year. The whole process is supervised by two teams (management and technical) consisting of representatives from the partner facilities, which care about the necessary detail planning.

6 References

[1] Functional Description of the EUU / UAA Tools, R. Dimper, D. Porte, O. Schwarzkopf, D. Feichtinger, D. Lauk, H.J. Weyer, 2010, https://www.UmbrellaID.org/euu/documents

[2] The Umbrella system Prototype Web-Based Access Point, EuroFEL Deliverable 2.5, B. Abt, Heinz J Weyer, April 2011, https://www.UmbrellaID.org/euu/documents

[3] Specification of the AAI Infrastructure, PaNdata ODI Deliverable 3.1,Aug 29, 2012, rev. Jul. 10, 2013, <https://www.UmbrellaID.org/euu/documents>

[4] PaNdata / CRISP Harmonization Meeting I, ZRH Airport, Jun 27, 2011 < http://indico.psi.ch/conferenceDisplay.py?confId=1039>

[6] PaNdata / CRISP Harmonization Meeting II, DESY, Dec 8, 2011 < https://indico.desy.de/conferenceDisplay.py?confId=5061>

[5] PaNdata / CRISP Harmonization Meeting III, ZRH Airport, Jun 13, 2012 < http://indico.psi.ch/conferenceDisplay.py?confId=1752>

[7] PaNdata / CRISP Harmonization Meeting IV, EU XFEL, Jan 21/22, 2013 < https://indico.desy.de/conferenceDisplay.py?confId=2159>

[8] ICAT
< http://code.google.com/p/icatproject/>