19-10-2017

Deliverable D3.3
Requirements Gathering from e-Infrastructures

Deliverable D3.3

Contractual Date: 31-08-2017
Actual Date: 19-10-2017
Grant Agreement No.: 731122
Work Package/Activity: 3/NA3
Task Item: Task 3
Nature of Deliverable: R (Report)
Dissemination Level: PU (Public)
Lead Partner: GEANT Limited
Document ID: GN4-2-17-15653C
Authors: A. Grant (GEANT Limited)

Abstract
This document summarises the requirements-gathering activity undertaken by the e-Infrastructure Liaison Task during GN4-2 Period 1. It covers methodology, and findings in the areas of network, AAI, security, training, marketing and communication, and user engagement.

Table of Contents

Table of Figures

Table of Tables

Executive Summary

e-Infrastructure Liaison, Task 3 within Networking Activity 3, Partner, User and Stakeholder Relations (NA3 T3), is a new Task introduced in GN4-2. Previously, e-infrastructures were managed within the project purely as a user-group stakeholder, with requirements relating primarily to their consumption of GÉANT services. The introduction of a dedicated Task reflects a more holistic approach to e-infrastructures, in recognition of their role as collaborative partners as well as service users, with an additional set of requirements relating to, for example, training and information sharing to facilitate joint projects with GÉANT.

The key objectives of Task 3 are to ensure that GÉANT provides excellent service to other e-infrastructure projects, and that GÉANT's work (especially with respect to services and users) is aligned with other related e-infrastructure initiatives.

The five e-infrastructures with which Task 3 is mainly concerned are PRACE, EGI, EUDAT, OpenAIRE and the Research Data Alliance (RDA).

To gather requirements, Task 3 held a series of introductory meetings with each e-infrastructure, involving not only technical representatives but also CEOs, to understand the strategic context of their requirements and their plans for future development. The Task also co-organised the 2016 Digital Infrastructures for Research (DI4R) conference, and signed memoranda of understanding (MoU).

The findings from the requirements gathering fall into the following categories: network, authentication and authorisation infrastructure (AAI), security, training, communications and marketing, and user engagement. The emphasis differs between the e-infrastructures according to their core activities. For example, PRACE had the most significant network requirement, including a new network topology. All the e-infrastructures (except RDA) have AAI requirements, with interoperability pilots planned for EGI and EUDAT; the AARC Engagement Group for InfrastructureS (AEGIS) is an important delivery channel for AAI recommendations and best practices.

The e-infrastructures' primary forum for discussing and addressing security matters is the Wise Information Security for e-Infrastructures (WISE) Community. Version 2 of WISE's Security for Collaboration among Infrastructures (SCI) framework document, which defines best practices, trust and policy standards for collaboration with the aim of managing cross-infrastructure operational security risks, was officially endorsed at TNC17 by GÉANT, PRACE, EGI and EUDAT among other e-infrastructures.

For most of the e-infrastructures, the training requirement is to identify opportunities for cooperation and information sharing. Specific plans include GÉANT's leveraging of PRACE's massive open online

courses (MOOCs) platform and deriving best practice from its successful summer school programme, exchanging training places with EGI, and developing joint training with EUDAT and OpenAIRE.

No specific individual marketing and communications requirements were identified. Rather, there is a general need to explore and develop collaborative marketing and communications opportunities. Past and current collaboration opportunities include hosting joint booths at DI4R 2016 and TNC17, and joint planning of DI4R 2017.

Similarly, no specific individual user-engagement requirements were identified. Rather, there is a general need to explore and develop collaboration opportunities, and to share knowledge and best practice. Past and planned collaboration opportunities include joint organisation of "Design your own e-infrastructure" workshops, the Task Force on Research Engagement Development (TF-RED) and supporting the EOSCpilot project Science Demonstrators.

The e-infrastructures are also working together on the eInfraCentral project, and collaborative agreements are being produced between EOSC-Hub and GN4-2/GN4-3 and between GÉANT and OpenAIRE. GÉANT and PRACE are also planning a joint exascale workshop.

That GÉANT and the e-infrastructures are working together across such a range of areas is testimony to the effectiveness with which Task 3's liaison and requirements-gathering activity has led to closer relationships and synergies that benefit not only the e-infrastructures themselves but ultimately the end users and their research.

1   Introduction

1.1   E-Infrastructure Liaison Task

e-Infrastructure Liaison, Task 3 within Networking Activity 3, Partner, User and Stakeholder Relations (NA3 T3), is a new Task introduced in GN4-2. Previously, e-infrastructures were managed within the project purely as a user-group stakeholder, with requirements relating primarily to their consumption of GÉANT services. The introduction of a dedicated Task reflects a more holistic approach to e-infrastructures, in recognition of their role as collaborative partners as well as service users, with an additional set of requirements relating to, for example, training and information sharing to facilitate joint projects with GÉANT.

Task 3 aims to foster stronger, mutually beneficial relationships through such collaboration. It coordinates the support provided to e-infrastructures by GÉANT, the NRENs and global partners to develop alignment of service portfolios, coordination of engagement and joint initiatives. The Task also seeks to develop and implement GÉANT's role in the emerging integrated e-infrastructure landscape and to coordinate the different strands of integration and consolidation work in this area. This Task works closely with Joint Research Activity 1 Network Infrastructure Evolution, Task 3 Integration with Other e-Infrastructures (JRA1 T3), to reach agreement on a common approach to service design, provision and implementation, and with Service Activity 1, Task 3 Network Implementation (SA1 T3), to whom it passes requirements. It achieves this through:

- Liaison and communication between GÉANT and other e-infrastructure providers to maximise the value of innovation in a cooperative, open and multi-domain ecosystem.
- Supporting the use of GÉANT infrastructure and services by Horizon 2020 projects, to increase uptake of GÉANT and NREN infrastructure and services by other e-infrastructures and projects and their user communities.
- Provision of a dedicated point of contact for e-infrastructures and H2020 projects in order to contribute to the establishment of collaboration agreements, the open catalogue of services and ensure that potential synergies, overlaps and gaps in the overall service offering are identified and addressed appropriately. This work is conducted in close collaboration with NA3 Task 2 Account Management of International User Groups and with the product management team.
- Developing a strategy for an interoperation framework that meets e-infrastructure requirements and service offerings – in alignment with JRA2 Network Services Development, Task 2 Service Provider Architecture, and JRA1 T3, which will deliver the technical aspects of service integration for e-infrastructures.

The key objectives of Task 3 are:

- To ensure that GÉANT provides excellent service to other e-infrastructure projects.
- To ensure that GÉANT's work (especially with respect to services and users) is aligned with other related e-infrastructure initiatives.

1.2 e-Infrastructures

The five e-infrastructures with which Task 3 is mainly concerned are:

- PRACE.
- EGI.
- EUDAT.
- OpenAIRE.
- RDA.

The list represents a progression up the stack, with (to simplify their core activities) PRACE and EGI offering computing resources, EUDAT offering data storage, OpenAIRE offering data services and the Research Data Alliance (RDA) being a promoter of open science (not a specific entity with requirements to be met but a new and valuable collaboration partner nonetheless).

Each of these is briefly described below.

1.2.1 PRACE

The mission of the Partnership for Advanced Computing in Europe (PRACE) is to enable high-impact scientific discovery and engineering research and development across all disciplines to enhance European competitiveness for the benefit of society. PRACE seeks to realise this mission by offering world-class computing and data management resources and services open to all European publicly funded researchers through a peer review process.

PRACE also seeks to strengthen the European users of high-performance computing (HPC) in industry through various initiatives. PRACE has a strong interest in improving the energy efficiency of computing systems and reducing their environmental impact.

For further information, please see [PRACE].

1.2.2 EGI

The European Grid Initiative (EGI) is a federated e-Infrastructure set up to provide advanced computing services for research and innovation. The EGI e-infrastructure is publicly-funded and comprises almost 300 data centres and cloud providers spread across Europe and worldwide. EGI offers a wide range of services for compute, storage, data and support, and provides access to over 650,000 logical CPUs and 500 PB of disk and tape storage.

The federation is coordinated by the EGI Foundation (also known as EGI.eu), created to manage the infrastructure on behalf of the participants of the EGI Council.

EGI's vision is for researchers from all disciplines to have easy, integrated and open access to the advanced scientific computing capabilities, resources and expertise needed to collaborate and to carry out data/compute intensive science and innovation.

Its mission is to create and deliver open solutions for science and research infrastructures by federating digital capabilities, resources and expertise between communities and across national boundaries.

For further information, please see [EGI].

1.2.3   EUDAT

EUDAT's vision is that data is shared and preserved across borders and disciplines. Achieving this vision means enabling data stewardship within and between European research communities through a Collaborative Data Infrastructure (CDI), a common model and service infrastructure for managing data spanning all European research data centres and community data repositories.

European researchers and practitioners from any research discipline can preserve, find, access and process data in a trusted environment, as part of the EUDAT Collaborative Data Infrastructure, a network of collaborating, cooperating centres, combining the richness of numerous generic and community-specific data repositories with the permanence and persistence of some of Europe's largest scientific data centres.

EUDAT offers heterogeneous research data management services and storage resources, supporting multiple research communities as well as individuals, through a geographically distributed, resilient network distributed across 15 European nations. Data is stored alongside some of Europe's most powerful supercomputers.

EUDAT is a service-oriented, community-driven, sustainable and integrated initiative.

One of EUDAT's main ambitions is to bridge the gap between research infrastructures and e-infrastructures through an active engagement strategy, using the communities that are in the consortium as EUDAT beacons and integrating others through innovative partnerships.

For further information, please see [EUDAT].

1.2.4   OpenAIRE

Fifty partners, from all EU countries and beyond, collaborate on the OpenAIRE2020 Project (OpenAIRE), a large-scale initiative that aims to promote open scholarship and substantially improve the discoverability and reusability of research publications and data. The initiative brings together professionals from research libraries, open scholarship organisations, national e-infrastructure and data experts, IT and legal researchers. A network of people, represented by the National Open Access Desks (NOADs), will organise activities to collect H2020 project outputs, and support research data

management. Backing this vast outreach is the OpenAIRE platform, the technical infrastructure that is vital for pulling together and interconnecting the large-scale collections of research outputs across Europe. The project will create workflows and services on top of this valuable repository content, which will enable an interoperable network of repositories (via the adoption of common guidelines), and easy upload into an all-purpose repository (via Zenodo).

OpenAIRE2020 will assist in monitoring H2020 research outputs and will be a key infrastructure for reporting H2020's scientific publications as it will be loosely coupled to the EC's IT backend systems. The EC's Research Data Pilot will be supported through European-wide outreach for best research data management practices and Zenodo, which will provide long-tail data storage. Other activities include: collaboration with national funders to reinforce the infrastructure's research analytic services; an APC Gold OA pilot for FP7 publications with collaboration from LIBER; novel methods of review and scientific publishing with the involvement of hypotheses.org; a study and a pilot on scientific indicators related to open access with CWTS's assistance; legal studies to investigate data privacy issues relevant to the Open Data Pilot; international alignment with related networks elsewhere with the involvement of COAR.

For further information, please see [OpenAIRE].

1.2.5   RDA

The Research Data Alliance (RDA) builds the social and technical bridges that enable open sharing of data.

The RDA vision is researchers and innovators openly sharing data across technologies, disciplines, and countries to address the grand challenges of society.

The current global research data landscape is highly fragmented, by disciplines or by domains, from oceanography, life sciences and health, to agriculture, space and climate. When it comes to cross-disciplinary activities, the notions of "building blocks" of common data infrastructures and building specific "data bridges" are becoming accepted metaphors for approaching the data complexity and enable data sharing. The Research Data Alliance enables data to be shared across barriers through focused Working Groups and Interest Groups, formed of experts from around the world – from academia, industry and government. Participation in RDA is open to anyone who agrees to its guiding principles of openness, consensus, balance, harmonisation, with a community driven and non-profit approach.

The Research Data Alliance was launched as a community-driven organisation in 2013 by the European Commission, the United States National Science Foundation and National Institute of Standards and Technology, and the Australian Government's Department of Innovation with the goal of building the social and technical infrastructure to enable open sharing of data.

With over 6,000 members from 130 countries (September 2017), RDA provides a neutral space where its members can come together through focused global Working and Interest Groups to develop and adopt infrastructure that promotes data-sharing and data-driven research, and accelerate the growth of a cohesive data community that integrates contributors across domain, research, national, geographical and generational boundaries.

For further information, please see [RDA].

1.3   In this Document

This document summarises Task 3's requirements-gathering activity during Period 1 of GN4-2. It covers the following topics:

- Methodology.
- Findings:
  - Network.
  - Authentication and Authorisation Infrastructure (AAI).
  - Security.
  - Training.
  - Marketing and communications.
  - User engagement.
- Conclusion.

2 Methodology

To gather requirements, Task 3 held a series of introductory meetings with each e-infrastructure, involving not only technical representatives but also CEOs, to understand the strategic context of their requirements and their plans for future development. The Task also co-organised the 2016 Digital Infrastructures for Research (DI4R) conference, and signed memoranda of understanding (MoU).

2.1 Meetings

The main meetings with each e-infrastructure at which requirements were gathered included:

2.1.1 PRACE

- 10–13 May 2016 PRACE user conference, including discussions with members of PRACE board of directors.
- 26–27 September 2016 e-IRG meeting in Bratislava – face-to-face side meeting with PRACE's Head of Operations.
- 16 March 2017 at PRACE offices in Brussels – face-to-face all-day meeting between GÉANT senior activity leads and PRACE operational leads.
- 23 March 2017 HPC digital day in Rome.
- 8 September 2017 at GÉANT offices in Cambridge – face-to-face all-day meeting between GÉANT senior activity leads and PRACE operational leads.
- 11 September 2017 – joint GÉANT/PRACE visit to the European Synchrotron Radiation Facility (ESRF) in Grenoble for hands-on user requirements gathering.
- Attendance and plenary presentation at PRACE conference in 2016 and in May 2017. Informal meetings of operational leads and GÉANT activity leads in network and AAI.
- Regular VC meetings with PRACE's Head of Project.

2.1.2 EGI

- Q1: meeting with EGI Technical Director.
- 26–27 September 2016 e-IRG meeting in Bratislava – face-to-face side meeting.
- 28–30 September 2016 DI4R– discussion with EGI CEO.
- November 2016 face-to-face meeting at EGI offices with EGI CEO and strategy lead.

- Q3 – representatives from JRA1 and NA3 T3 met with EGI to discuss how further orchestration and integration work could take place.
- July 2017 – face-to-face meeting at EGI offices with EGI CEO and strategy lead.
- Attendance at EGI conference 12 May 2017 and informal meetings.

2.1.3 EUDAT

- 20–22 September 2016 NORDUnet 2016 conference – meeting with the EUDAT Project Manager.
- Regular VCs with EUDAT Project Director.
- 17 August 2017 – VC with EUDAT Project Director and EUDAT Project Co-ordinator.
- 22 August 2017 – VC with EUDAT Project Director and EUDAT Project Co-ordinator.

2.1.4 OpenAIRE

- 5 May 2016 kick-off meeting.
- 6 May 2016 face-to-face visit.
- 28–30 September 2016 DI4R – face-to-face meeting with OpenAIRE management team.
- November 2016 – face-to-face meeting between the GÉANT management team and the OpenAIRE management and legal teams.
- 6 September 2017 – attendance at Open Science Fair and joint workshop with OpenAIRE and EGI.

2.1.5 RDA

- GÉANT attendance at RDA plenary meetings 5–7 April 2017 in Barcelona and 19–21 September 2017 in Montreal.

2.2 DI4R

Task 3 co-organised the 2016 Digital Infrastructures for Research (DI4R) conference, which took place 28–30 September 2016 in Krakow, Poland [DI4R]. The conference had two key goals:

- To foster communication between the e-infrastructure projects and share knowledge, experience and best practice. Also, to act as a networking forum for the key members of each e-infrastructure so that they could identify opportunities for partnership and collaboration.
- To communicate to external stakeholders the combined activities of the e-infrastructure projects and demonstrate the value of collaboration between them.

The T3 Task Leader gave significant planning and programme support through active participation in the Programme Committee. GÉANT contributed jointly with other e-infrastructures in the following sessions during the conference:

- Plenary panel: GÉANT representative answered questions directly from users.
- Towards interoperable e-infrastructures.
- Federated service management.
- Research user engagement across e-infrastructures
- SME user engagement across e-infrastructures.
- Copernicus data.
- Authentication and Authorisation Infrastructure (AAI).
- Cloud procurement.

The Task also coordinated GÉANT participation in a number of very successful workshops that took place ahead of the DI4R conference including:

- "Design your own e-infrastructure" (with EGI, EUDAT and OpenAIRE), in which five user-group case studies were holistically assessed to see how e-infrastructures could meet their needs.
- WISE (pan e-infrastructure) security full-day workshop.

2.3   MoU

To formalise the relationships, GÉANT has signed or is developing the following memoranda of understanding (MoU):

- PRACE: MoU signed at the end of March 2017.
- EGI: an enhanced and updated MoU, covering additional areas not included in the original 2012 version, was signed at TNC17 (29 May – 2 June 2017).
- EUDAT: a draft MoU has been produced and is under discussion.

With regard to OpenAIRE, both parties have agreed that an MoU is not necessary to pursue the main areas of focus, namely AAI and training, on which they will work collaboratively.

3 Findings

The findings from the requirements gathering fall into the following categories:

- Network.
- Authentication and Authorisation Infrastructure (AAI).
- Security.
- Training.
- Communications and marketing.
- User engagement.

Each of these is discussed below.

3.1 Network

3.1.1 PRACE

Three main requirements were identified for PRACE:

- Network topology.
- Enhanced requirements-gathering template.
- perfSONAR monitoring.

3.1.1.1 Network Topology

PRACE links several high-performance computing sites around Europe. Originally, this was a closed community of connections to a DEISA switch in Frankfurt, hosted by the Jülich Supercomputing Centre. It was a star topology of 10 Gbps optical wavelengths, with a separate 1 Gbps bridge that allowed access to the network for the other members of the PRACE consortium that did not operate supercomputing (SC) centres. This allowed consortium connections to be established in a shorter timescale when compared to an optical network connection, albeit with the disadvantage of the reduced speed.

In order for a new SC site to connect, they would contact their NREN. The NREN would then liaise with the GÉANT business development team in order to establish the optical connection. A similar arrangement existed when non-optical connections needed to be made to the bridge.
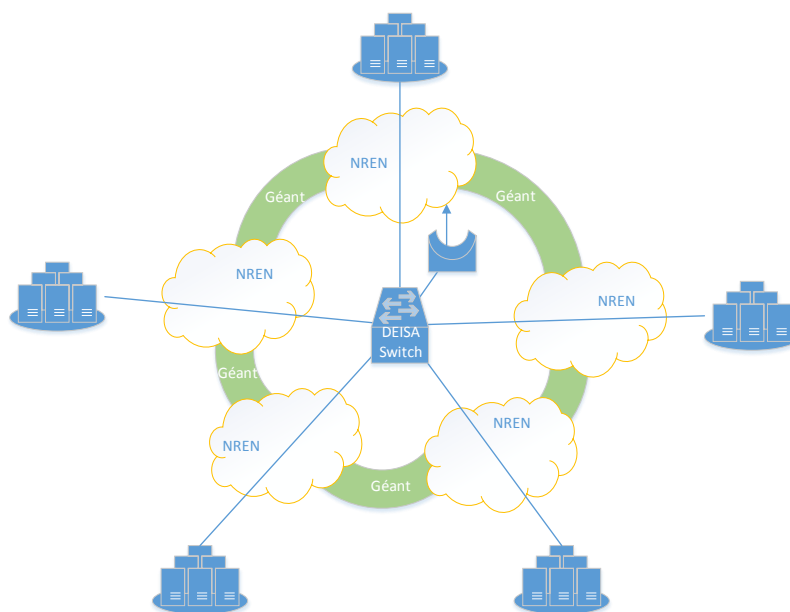
Figure 3.1: High-level overview of PRACE network star topology

The issues which were inherent in this topology were its lack of flexibility, a single point of failure in the central switch and the fact that the switch was soon to come to its end of life.

GÉANT was providing services for the topology, and was also in regular contact as part of the user account management function. It was during these regular interactions that PRACE approached GÉANT about a requirement to update the network topology. Following a requirements-gathering process, GÉANT proposed a mesh network in order to fulfil PRACE's needs. However, due to the costs and complexities involved in moving over to this new topology, the suggested solution was not taken forward.

In 2016, during another regular service call, PRACE again approached GÉANT for support to explore alternative long-term network topology solutions. This time the two primary criteria for a new solution were speed of delivery, as the vendor of the core switch was shortly due to stop supporting it, and reduced cost, to enable the connection of an SC centre for whom cost was currently a prohibiting factor.

During a series of meetings between representatives of PRACE, GÉANT and the lead NREN, DFN, a number of features were identified as requirements in order to help define a suitable future solution. These included:

- Status monitoring of user connectivity.
- Recording of international PRACE traffic volumes.
- Recording of per site PRACE traffic volumes.
- Recording, reporting and analysis of PRACE traffic flow information.
- Use of perfSONAR for additional monitoring and logging purposes.
- NREN endorsement.
- Ease of addition of new users.

Also, upon analysis by GÉANT of the traffic levels across all of the PRACE optical links, it was discovered that bandwidth rarely exceeded 1 Gbps for production traffic. Due to the traffic levels falling within the capabilities of existing NREN connections to the GÉANT backbone, a purely optical solution was not something that had to be adhered to. This meant that a more novel and ultimately cost-effective solution was possible.

Two options were put forward for consideration: L3VPN and MD-VPN. Both of these solutions utilised the existing NREN infrastructure and GÉANT backbone. This would negate the need for new optical circuits to be established between the existing SC centres and would also mean that connections could take advantage of the multiple forms of redundancy across the NREN and GÉANT networks, further strengthening the resiliency of the delivered solution when compared to a purely optical point-to-point circuit. Removing the need for point-to-point circuits would also minimise costs both for existing centres and new centres joining the new topology.

Due to the need for a rapid rollout and strong backing from the NREN community, MD-VPN was the preferred solution.

Following a successful pilot implementation in 2016 Q4 involving the Human Brain project, which linked five supercomputing sites around Europe (located in Spain, France, Germany, Switzerland and Italy), the full solution was ordered in January 2017 and was operational in May 2017, an implementation period of just four months. GÉANT acted as the coordinating function, providing account management, pricing and technical liaison during the planning and deployment of the solution. It is important to note that the establishment of contracts for the MD-VPN service remained between the end sites and their respective NRENs, as this was the most productive way for the end sites to interact with the NREN community. Due to the distributed nature of the PRACE organisation, having GÉANT as the coordinating organisation, rather than each individual NREN, allowed a more productive approach in the delivery of the end service.

Acceptance of the new MD-VPN platform as being fit for purpose was a major milestone, not just for the PRACE community but for the NREN community as a whole, since the technology was developed within an NREN-community project: GÉANT. As the pilot platform for the Human Brain project was in production, PRACE were able to move their existing traffic over to the live MD-VPN system with very limited configuration work required on their part, compared with a brand-new setup. Utilising the production pilot platform saved both time and money for PRACE.
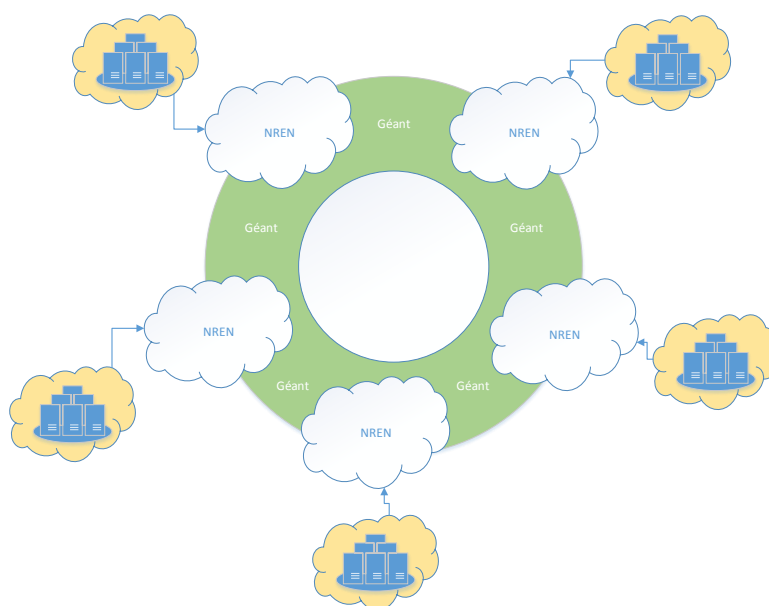
Figure 3.2: High-level overview of PRACE MD-VPN network design

The benefits of the MD-VPN solution include:

- Significant cost savings.
- Shorter timescales and little or no equipment required for implementing new connections.
- Improved resiliency.
- Future-proofing and flexibility.

The first two have the further significant advantage of lowering the barrier for participation in the PRACE network across the world.

For further information, please see the forthcoming *Deliverable 3.4 Case Study: Report on Two Cases of User Account Management* (due M19).

3.1.1.2 *Enhanced Requirements-Gathering Template*

Task 3's work with PRACE and EUDAT has identified the opportunity to offer an extended, improved service to supercomputing (SC) users by providing an enhanced requirements gathering template for HPC projects that helps to identify other services – in particular, data-management services, such as those provided by EUDAT – that might be of relevance to a project in addition to their SC needs.

An enhanced template has been developed as a proof of concept, capturing such information as data flow, transfer and storage. This will be used in Q4 2017 and throughout 2018 to capture more information on the requirements of SC users, which will then be analysed by Task 2.

3.1.1.3 *perfSONAR*

The use of perfSONAR was one of the requirements identified as part of the new network topology solution. Task 3 has facilitated the deployment of the perfSONAR monitoring infrastructure within the MD-VPN service GÉANT provides for PRACE.

3.1.2 EGI

EGI uses GÉANT to connect all its federated data centres in Europe [EGI-Data-Centres]. There is a closer relationship than service provider and client in that many members of GÉANT are also members of EGI.

In the course of the many meetings held with EGI, no specific networking issues or requirements were identified. (The collaboration areas are primarily in the area of AAI.)

3.1.3 EUDAT

EUDAT uses PRACE to reach its storage sites, which are often located at SC centres. While it raised no network requirements of its own, it therefore benefits from any improvements to the PRACE infrastructure, such as the switch to MD-VPN and the resulting cost reductions and increased resiliency.

EUDAT has been involved in the proof-of-concept work to develop an enhanced requirements-gathering template described in Section 3.1.1.2 above.

3.1.4 OpenAIRE

OpenAIRE's technical infrastructure uses the NRENs and GÉANT to interconnect the open access repositories (databases or virtual archives, typically at universities or research institutes) to which it provides a gateway. (A list of European repositories is available at [OpenDOAR-Eu].) OpenAIRE is satisfied with the GÉANT network service it receives and has no additional specific network issues or requirements. (As with EGI, its needs are primarily AAI.)

3.1.5 RDA

RDA has no physical network infrastructure requirements.

3.2 AAI

3.2.1 PRACE

PRACE AAI mainly uses digital certificates to authenticate users to access PRACE resources. The certificates accepted are those issued by the Certification Authorities (CAs) that are accredited by the Interoperable Global Trust Federation (IGTF) [IGTF]. The use of a digital certificate also enables single

sign-on across PRACE systems, which is convenient as PRACE is a distributed infrastructure with different nodes in Europe.

User accounts are kept in a central LDAP database; local accounts on PRACE nodes are provisioned via this database. Each site also runs an LDAP database and has authority on their local branch of the LDAP. When a user requests access to a site, in practice they are added to the LDAP branch for that site. A user needs to submit a request to access PRACE that is then evaluated by PRACE.

Federated access is not provided yet, but PRACE is exploring its adoption at least to access some of their web resources, e.g. wikis.

PRACE requirements to enable federated access include:

- Support for non-web access.
- Authentication process to be documented in English.
- Traceability of all users.
- Users' identifiers should be persistent, i.e. users' identifiers cannot be reassigned.
- Changes in the users' role should be propagated in the related attributes immediately.

GÉANT via the AARC project has engaged with PRACE on two main occasions:

- In 2015, to gather PRACE's requirements in relation to the level of assurance, that is, how much a service provider (in this case PRACE) can trust the credentials issued by a user's home organisation.

    The results of the survey scoped the work on assurance that is now being carried out as part of an open REFEDS working group and to which AARC contributes resources.

- To pilot a way to enable a PRACE user to access EUDAT services seamlessly.

    This pilot started in autumn 2016 and is almost complete. The concept is that a PRACE user, upon giving consent, is tagged in the PRACE LDAP to be EUDAT-enabled. This enables a PRACE-authenticated user to, for instance, store data in EUDAT.

    The user information coming from PRACE LDAP is added into EUDAT B2ACCESS [B2ACCESS] (EUDAT's AAI service), where a PRACE group is created. A script ensures that the information between PRACE and EUDAT is synchronised.

In 2017 AARC created a new engagement forum, the AARC Engagement Group for InfrastructureS (AEGIS) [AEGIS]. AEGIS brings together representatives from research and e-infrastructures, operators of AAI services and the AARC team to bridge communication gaps and make the most of common synergies. PRACE will be engaging more closely with AARC via AEGIS. AEGIS is described in more detail in Section 3.2.6.

At the September 2017 face-to-face meeting with PRACE operational leads, eduTEAMS was identified as a potential solution to PRACE's new AAI requirements and therefore a meeting with PRACE's new AAI lead, Eoin McHugh, is being arranged to take this forward.

3.2.2 EGI

The primary opportunity and requirement is for GÉANT and EGI, together with EUDAT, to work collaboratively to align their AAI systems, to identify and optimise synergies rather than to offer overlapping solutions.

3.2.2.1 *Interoperability Pilot*

GÉANT and EGI are planning an interoperability pilot in which the GÉANT eduTEAMS Membership Management Service (MMS) will be integrated with the EGI Check-in Service as a trusted, external attribute authority. eduTEAMS is a suite of services, built on top of eduGAIN, that supports AAI for R&E collaborations between virtual teams [eduTEAMS]; MMS allows collaborations to record and manage group and role information and deliver that information to their services. EGI Check-in (in Beta phase) is a proxy service that operates as a central hub to connect federated identity providers (IdPs) with EGI service providers [EGI-Check-in]. Check-in allows users to select their preferred IdP so that they can access and use EGI services in a uniform and easy way. The main characteristics of the service are that it:

- Enables multiple federated authentication sources using different technologies.
- Provides increased productivity and security.
- Is federated in eduGAIN as a service provider.
- Is compliant with REFEDS RnS and Sirtfi.
- Providers a user registration portal to allow accounts-linking.
- Combines user attributes originating from various authoritative sources (IdPs and attribute provider services) and delivers them to the connected EGI service providers in a transparent way.

Although Task 3 has already discussed the general principles with EGI representatives, there has been no opportunity to discuss the exact technical details of the pilot. This is expected to happen in Period 2 Q1.

3.2.2.2 *EOSC-Hub Project*

EGI and EUDAT are already working together to develop an AAI solution for the European Open Science Cloud (EOSC)-Hub project, which is due to start in early 2018 [EOSC-Hub]. Led by the EGI Foundation, the project brings together 74 beneficiaries including research infrastructures, national e-infrastructure providers, SMEs and academic institutions. EOSC-Hub builds on existing technology already at technology readiness level (TRL) 8 and addresses the need for interoperability by promoting the adoption of open standards and protocols. EOSC-Hub mobilises e-infrastructures comprising more than 300 data centres worldwide and 18 pan-European infrastructures. The result will be an integrated catalogue of services, software and data from EGI, EUDAT CDI and major research e-infrastructures.

3.2.2.3 *AAI Roadmapping Meeting*

A joint GÉANT and EGI all-day workshop is scheduled to take place on 12 October 2017. The aim of this face-to-face meeting is to ensure alignment of AAI roadmaps. Their requirements will be reviewed following the meeting.

3.2.3 EUDAT

As with EGI, the primary opportunity and requirement is for GÉANT and EUDAT, together with EGI, to work collaboratively to align their AAI systems, to identify and optimise synergies rather than to offer overlapping solutions.

3.2.3.1 *Interoperability Pilot*

GÉANT and EUDAT have also agreed to work together on an interoperability pilot, this time between eduTEAMS and B2ACCESS. B2ACCESS [B2ACCESS] is a secure, versatile easy-to-use authentication and authorisation platform developed by EUDAT. B2ACCESS can be integrated with any service; when integrated, B2ACCESS allows the user to log in by using different methods of authentication, such as a home organisation identity provider via eduGAIN, a Google account or an EUDAT ID.

In the interoperability pilot, EUDAT will be using the eduTEAMS Membership Management Service (MMS) as a way to bring community management capabilities into B2ACCESS. eduTEAMS will be connected to B2ACCESS as a trusted, external attribute authority using OpenID Connect.

In phase 1 of the pilot, a community of users from DKRZ, the German Climate Computing Centre, will be using eduTEAMS to manage user enrolment and their membership of groups and roles within that community. Based on this information, users in this community will be able to access EUDAT services.
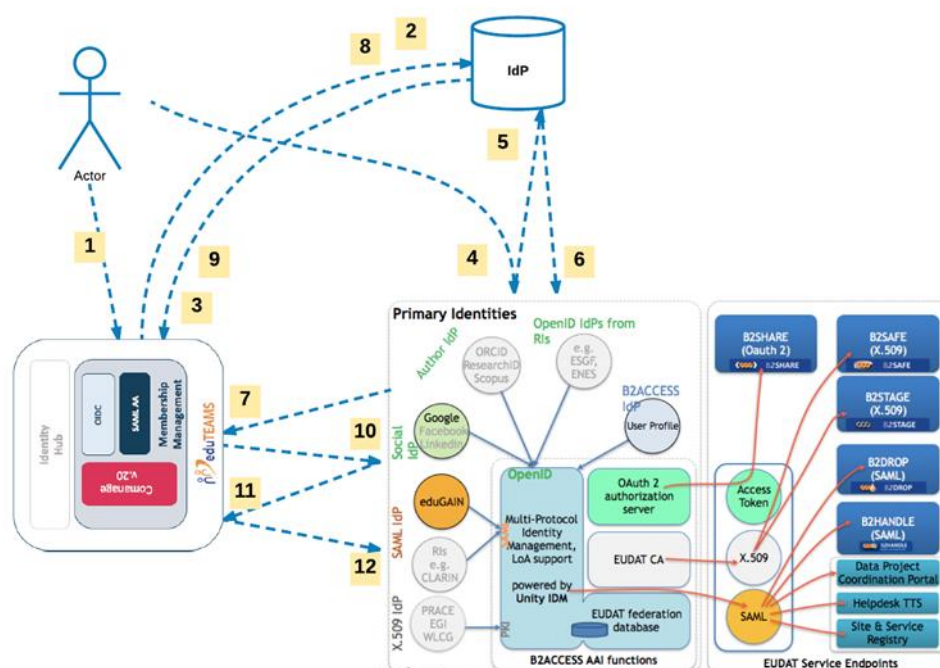
Figure 3.3: eduTEAMS-B2ACCESS pilot phase 1

In phase 2, the pilot will investigate deeper integration between the two platforms, so that communities registered on eduTEAMS can be automatically configured on B2ACCESS and vice versa.

Another point of interest for this pilot is the eduTEAMS Identity Hub service, which allows any service that is already in eduGAIN to authenticate users on external identity providers, such as Google, Facebook, LinkedIn, GitHub, etc. Although B2ACCESS already has support for Google accounts, there are benefits to EUDAT using a GÉANT service with wider external identity provider selection that is generally available via eduGAIN.

3.2.4   OpenAIRE

OpenAIRE has been working with GRNET since last year on an AAI implementation based on the GRNET RCAuth framework, which is also used by the EGI Check-in Service. Recently, Task 3 discussed with representatives from OpenAIRE the possibility of a collaboration in this space. In this regard, OpenAIRE would like GÉANT to validate their current implementation and provide consultancy on its future plans. Furthermore, although OpenAIRE is not an official partner in the AARC2 project, Task 3 has agreed to investigate the possibility of having an AARC2 cross-infrastructure pilot that also involves OpenAIRE. Task 3 and OpenAIRE have agreed to have an initial call on this.

3.2.5   RDA

RDA has no specific AAI requirements although GÉANT, via the AARC project, will produce some how-to documents to support AAI architects and developers in the various research infrastructures that attend RDA meetings and are members.

Further to GÉANT attendance at the September 2017 plenary meeting in Montreal, it became clear that AAI is being discussed in multiple RDA working groups (WGs). Therefore it has been agreed that a cross-WG session will be held at the next RDA in Berlin in March 2018, which will include a presentation by the AARC project on the current state of affairs in AAI.

3.2.6 AEGIS

The AARC2 project, like AARC, has made the strategic choice to entrust the implementation and operation of the AARC results to existing research infrastructure and e-infrastructures. This choice is based on the consideration that existing infrastructures are better positioned to deploy and operate services and to implement any recommendations and best practices to guarantee sustainability and support for the scientific communities. For this reason, whenever possible, AARC2 works with research infrastructures and e-infrastructures, which act as the delivery channel between AARC2 and the user communities they represent.

The AARC Engagement Group for InfrastructureS (AEGIS) [AEGIS], previously called the AARC Competence Centre, establishes bi-directional channels between AARC2 and the infrastructures to advise each other on the developments and production integration aspects of the AARC results. The group is composed of representatives of e-infrastructures and research infrastructures that have committed to implementing the AARC Blueprint Architecture and the accompanying policy frameworks for the benefit of the research communities they are supporting.
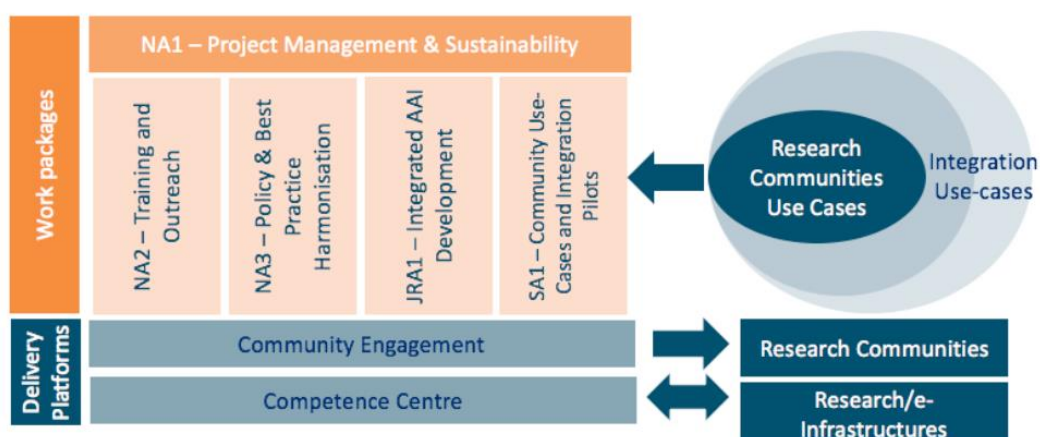
Figure 3.4: Role of AEGIS / Competence Centre in delivering AARC results

AEGIS ensures that:

- The results of AARC2 are known to all research infrastructures and e-infrastructures.

- Infrastructures and AARC2 team can discuss the AARC2 sustainability models, implementation aspects and approaches to use cases that may be received by AARC2.

- All key parties share the same vision and the same information about AARC2 objectives and developments in the trust and identity area, even if they may be in different deployment phases.

- Different infrastructures develop harmonised, interoperable solutions and avoid "reinventing the wheel".

The current list of AEGIS participants includes:

Person(s)	Organisation	Role / Infrastructure
AARC2		
Christos Kanellopoulos	GÉANT	Convener
Licia Florio	GÉANT	AARC2 Project Coordinator
Andrea Biancini	RETI	AARC2 Activity Leader (NA2)
David Groep	Nikhef	AARC2 Activity Leader (NA3)
Nicolas Liampotis	GRNET	AARC2 Activity Leader (JRA1)
Arnout Terpstra	SURFnet	AARC2 Activity Leader (SA1)
Infrastructure Representatives		
Klaas Wierenga, Ann Harding	GÉANT, SWITCH	GÉANT
Johannes Reetz, Willem Elbers	MPCDF, CLARIN	EUDAT
Peter Solagna, Tiziana Ferrari	EGI Foundation	EGI
Mikael Linden, Michal Procházka	CSC, CESNET	ELIXIR
Eoin McHugh	ICHEC	PRACE
Jim Basney, Lee Liming	NCSA, University of Chicago	XSEDE

Table 3.1: Current AEGIS participants

3.3 Security

The e-infrastructures' primary forum for discussing and addressing security matters is the Wise Information Security for e-Infrastructures (WISE) Community [WISE], which provides a trusted global framework where security experts can share information on topics such as risk management, operational security and threat intelligence in the context of e-infrastructures.

A working group within WISE, which includes representatives from GÉANT, PRACE, EGI and EUDAT, has been updating the Security for Collaboration among Infrastructures (SCI) framework document, which defines best practices, trust and policy standards for collaboration with the aim of managing cross-infrastructure operational security risks. Through this work the aim has been to establish a common understanding of the security measures each infrastructure has implemented and to start work on guidelines for interoperation such as the exchange of information during security incident handling. Version 2 of the SCI document, which includes a wider range of stakeholders from research

infrastructures and NRENs, was officially endorsed at TNC17 by GÉANT, PRACE, EGI and EUDAT among other e-infrastructures.

Figure 3.5: SCI endorsement ceremony at TNC17

3.4 Training

3.4.1 PRACE

PRACE delivers some of its training through the massive open online courses (MOOCs) platform FutureLearn [PRACE-MOOC]. It also has an established summer school programme, which is jointly organised with XSEDE, RIKEN and Compute Canada – Calcul Canada [PRACE-SS].

Figure 3.6: PRACE MOOC courses

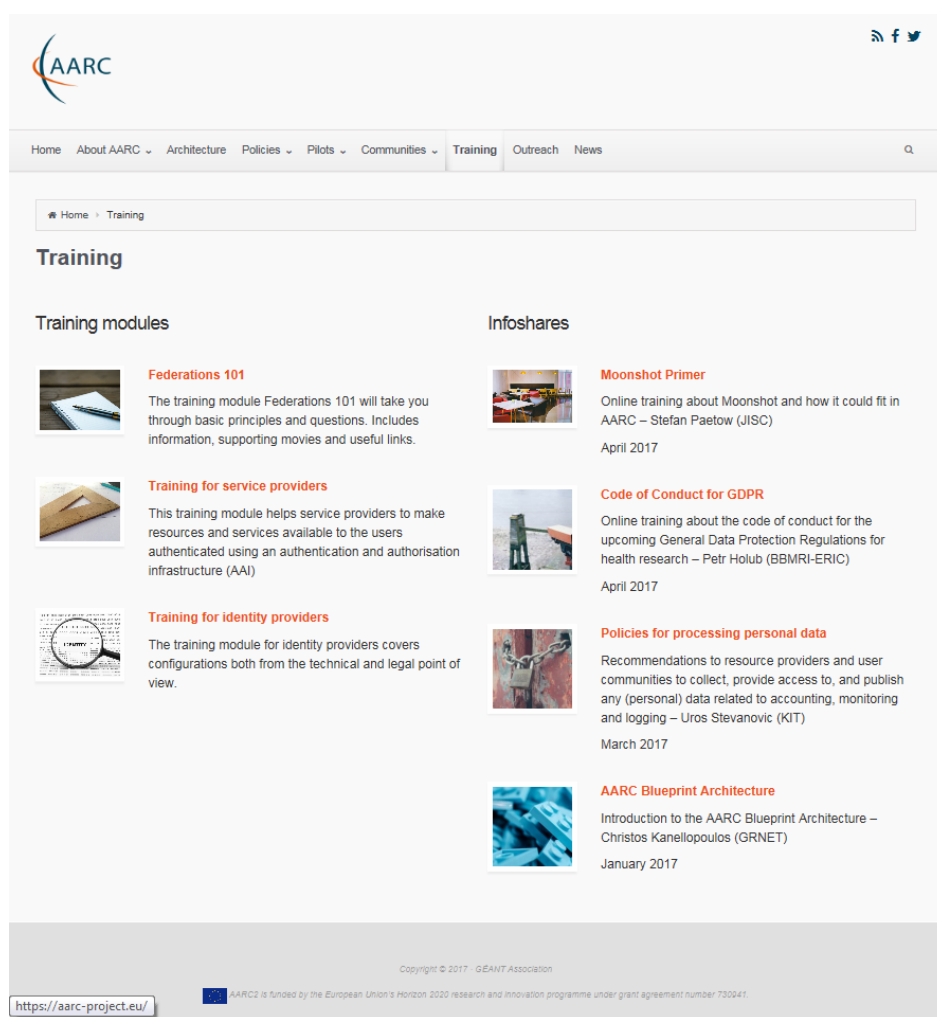GÉANT, meanwhile, has developed a training platform for the AARC2 project.

Figure 3.7: GÉANT AARC2 training platform

The requirement is to investigate opportunities for cooperation, information sharing and best practice, particularly with regard to leveraging PRACE's MOOC platform and deriving best practice from the successful summer school programme. This will take place in Q1 2018.

3.4.2   EGI

As with PRACE, the requirement is to identify opportunities for cooperation and information sharing.

A plan for exchanging training places, and co-designing and co-locating training events has been discussed and a first training exchange has taken place: GÉANT attended EGI's Federated IT Service Management (FITSM) course, and EGI attended GÉANT's Managing People and Performance in Virtual Teams (MPiViT) in 2016 and 2017.

3.4.3 EUDAT

Again, the identified requirement is to investigate opportunities for cooperation, information sharing and best practice. Accordingly, GÉANT and EUDAT have started to develop joint training across their user communities, including NRENs, with EUDAT covering, for example, best practices for data management and storage, particularly in the light of the General Data Protection Regulation (GDPR), and GÉANT delivering, for example, AAI or security training. GÉANT is investigating the possibility of a GDPR session at the EUDAT conference in January 2018.

3.4.4 OpenAIRE

In the interest of addressing the same requirement for cooperation and information sharing, GÉANT and OpenAIRE have agreed to develop and deliver joint training across their communities. GÉANT will promote eduGAIN to the OpenAIRE community via a set of virtual training sessions. OpenAIRE in turn will run a number of training sessions with a focus on data management planning and control, as well as best practice in the area of Open Data, and make them available to the GÉANT community.

Further areas identified for potential knowledge transfer include setting up legal entities, service portfolio management, cost sharing, training and AAI.

3.4.5 RDA

No training requirements were identified for RDA.

3.5 Marketing and Communications

No specific individual marketing and communications requirements were identified. Rather, there is a general need to explore and develop collaborative marketing and communications opportunities, serving to both build and demonstrate a close relationship between the e-infrastructures. Past and planned collaboration opportunities include:

- DI4R 2016 – joint booth.

Figure 3.8: Joint booth at DI4R

- TNC17 – joint booth.

Figure 3.9: Joint booth at TNC17

- Working together on DI4R 2017, which will take place 30 November – 1 December 2017. A joint booth involving all five e-infrastructures is planned.
- Regular meetings of the cross-e-infrastructure marketing group that was set up as an outcome of DI4R 2016. In addition to organising DI4R conferences, the group investigates opportunities for other joint meetings and for promoting e-infrastructures in a holistic way. One such opportunity is the EUDAT conference in January 2018, where the e-infrastructures may have a joint booth.
- Open Science Fair (OSF) – GÉANT attended the OSF and displayed and distributed PRACE marketing materials and took part in a joint e-infrastructure session to support users.

3.6 User Engagement

As with marketing and communications, no specific individual user-engagement requirements were identified. Rather, there is a general need to explore and develop collaboration opportunities, and to share knowledge and best practice. Past and planned collaboration opportunities include:

- DI4R 2016 "Design your own e-infrastructure". In this interactive workshop, jointly organised and delivered by GÉANT, EGI, EUDAT and OpenAIRE, five user-group case studies were holistically assessed to see how e-infrastructures could meet their needs. The workshop analysed the use cases brought by the participating scientific communities, then collaboratively defined and designed suitable e-infrastructure setups and roadmaps to implement them using and customising solutions from the participating e-infrastructures. The use cases analysed were:
  - Data Repository for Design Research.
  - Extreme Light Infrastructures.
  - European Research Initiative in Chronic Lymphocytic Leukaemia (ERIC).
  - Multi-Scale Complex Genomics VRE.
  - Open Digital Research Environment Toolkit for the Advancement of Mathematics (OpenDreamKit).

  Further information is available at [DI4R-DYOeI]. A similar workshop was organised by EGI at their conference in April 2016.

- DI4R 2017. Due to the success of the previous "Design your own e-infrastructure" workshop, this will be repeated again at DI4R this year and all e-infrastructures will take part.
- Task Force on Research Engagement Development (TF-RED). TF-RED is a new Task Force, established in February 2017, that formalises an existing collaboration of NRENs and other research infrastructures, working together to develop methods for engaging researchers and collaborations nationally and internationally, through a set of reports, workshops, and best practices. TF-RED provides an additional forum for the e-infrastructures to communicate and strengthen collaboration. The EGI user engagement lead sits on the steering committee, to ensure a close cooperation of e-infrastructures and a focus on joint collaborative opportunities and events towards the users. Further information is available at [TF-RED].
- EOSCpilot. The e-infrastructures are working together on supporting the EOSCpilot project Science Demonstrators. Science Demonstrators show the relevance and usefulness of EOSC

services and how they enable data reuse, and will drive EOSC development. The Demonstrator set includes environmental & earth sciences, high-energy physics, social sciences, life sciences and physics. Further information is available at [EOSCpilot-SD].

- GÉANT, EGI and OpenAIRE hosted a joint workshop at the Open Science Fair (OSF), 6–8 September 2017 in Athens. The collaborative workshop explored aspects of coordinating EOSC-related activities across large European infrastructures at European and national level, including AAI, technical service provision, training, community engagement and support.

4 Conclusion

NA3 Task 3 is working successfully to enhance relationships with e-infrastructures through identifying and helping to address requirements and identifying and helping to realise opportunities for collaboration and synergy. The liaison work is ongoing, meaning the Task is well placed to identify changing needs.

Current and planned activities that demonstrate the closer relationships built during GN4-2 Period 1 include:

Projects

- eInfraCentral [eInfraCentral]. All the e-infrastructures are working together on this European e-infrastructure services gateway project, whose mission is to ensure that, by 2020, a broader and more varied set of users (including industry) discovers and accesses the existing and developing e-infrastructure capacity. A common approach to defining and monitoring e-infrastructures services will increase their uptake and enhance understanding of where improvements can be made in their delivery.

  eInfraCentral will:

  ○ Structure an open and guided discussion between e-infrastructures to consensually define a common catalogue for their services.

  ○ Develop a single entry point (one-stop shop) – the eInfraCentral portal – for end users to browse the service catalogue, and enhance the monitoring of key performance indicators that focus on availability and quality of services and user satisfaction.

  ○ Draw policy and sustainability lessons for the future development of a European e-infrastructure "market place" as an extension of the common service catalogue and portal so that it includes a much broader range of e-infrastructures and services.

- EOSCpilot [EOSCpilot]. All the e-infrastructures are working together on the European Open Science Cloud for Research Pilot Project, and in particular on supporting the EOSCpilot Science Demonstrators [EOSCpilot-SD].

  The European Open Science Cloud (EOSC) will offer 1.7 million European researchers and 70 million professionals in science and technology a virtual environment with open and seamless services for storage, management, analysis and reuse of research data, across borders and scientific disciplines by federating existing scientific data infrastructures, today scattered across disciplines and Member States.

  The EOSCpilot project has been funded to support the first phase in the development of the EOSC. It will:

- ○ Propose and trial governance frameworks for the EOSC and contribute to the development of European open science policy and best practice.
- ○ Develop a number of demonstrators functioning as high-profile pilots that integrate services and infrastructures to show interoperability and its benefits in a number of scientific domains.
- ○ Engage with a broad range of stakeholders, crossing borders and communities, to build the trust and skills required for adoption of an open approach to scientific research.

Collaborative Agreements

- EOSC-Hub [EOSC-Hub]. An intra-project collaboration agreement is being drafted between EOSC-Hub and GN4-2 (and GN4-3), with the aim of aligning key building blocks of cloud services, including AAI, security, interoperability, training and governance. The planned agreement effective date is 1 January 2018.
- OpenAIRE. A collaborative agreement between GÉANT and OpenAIRE is also due to become effective on 1 January 2018. This will formalise the outcome of the joint discussions regarding GÉANT's validation of OpenAIRE's current AAI implementation and provision of consultancy on future plans, as described in Section 3.2.4.

EDI

- GÉANT/PRACE exascale workshop. As part of the collaborative effort to develop and implement the European data infrastructure (EDI), GÉANT and PRACE are planning a joint workshop to explore aspects of EDI governance, technical architecture requirements and timescales for future joint work in this area, to help ensure that GÉANT's network and services are fit for exascale deployment. This will take place on 13 October 2017.

In addition to ensuring that GÉANT is future-proofed to meet potential exascale requirements, Task 3 is ensuring the e-infrastructures' network and services requirements are met by passing them to JRA1 T3, for incorporation into their infrastructure development plans.

That GÉANT and the e-infrastructures are working together across such a range of areas is testimony to the effectiveness with which Task 3's liaison and requirements-gathering activity has led to closer relationships and synergies that benefit not only the e-infrastructures themselves but ultimately the end users and their research.

References

[AEGIS]	https://aarc-project.eu/communities/aegis/
[B2ACCESS]	https://www.eudat.eu/services/b2access
[DI4R]	https://www.digitalinfrastructures.eu/content/about-di4r-2016
[DI4R-DYOeI]	https://indico.egi.eu/indico/event/3025/
[eduTEAMS]	https://www.geant.org/Innovation/eduteams
[EGI]	http://www.egi.eu/
[EGI-Check-in]	https://www.egi.eu/internal-services/checkin/
[EGI-Data-Centres]	https://www.egi.eu/federation/data-centres/
[eInfraCentral]	http://einfracentral.eu/
[EOSC-Hub]	https://www.egi.eu/news/eosc-hub-project-is-favourably-evaluated/
[EOSCpilot]	https://eoscpilot.eu/
[EOSCpilot-SD]	https://eoscpilot.eu/science-demonstrators
[EUDAT]	https://www.eudat.eu/
[IGTF]	https://www.igtf.net/
[OpenAIRE]	https://www.openaire.eu/
[OpenDOAR-Eu]	http://www.opendoar.org/countrylist.php?cContinent=Europe
[PRACE]	http://www.prace-ri.eu/
[PRACE-MOOC]	http://www.prace-ri.eu/prace-moocs-via-future-learn/
[PRACE-SS]	http://www.prace-ri.eu/training-and-documentation/i-hpc-ss/
[RDA]	https://www.rd-alliance.org/
[TF-RED]	https://wiki.geant.org/display/RED/TF-RED+Home
[WISE]	https://wise-community.org/

Glossary

AAI	Authentication and Authorisation Infrastructure
AARC	Authentication and Authorisation for Research and Collaboration
AEGIS	AARC Engagement Group for InfrastructureS
CA	Certification Authority
CDI	Collaborative Data Infrastructure
CEO	Chief Executive Officer
CPU	Central Processing Unit
DEISA	Distributed European Infrastructure for Supercomputing Applications
DI4R	Digital Infrastructures for Research
EDI	European Data Infrastructure
EGI	European Grid Initiative
EOSC	European Open Science Cloud
ERIC	European Research Initiative in Chronic Lymphocytic Leukaemia
ESRF	European Synchrotron Radiation Facility
FITSM	Federated IT Service Management
GDPR	General Data Protection Regulation
HPC	High-Performance Computing
IdP	Identity Provider
IGTF	Interoperable Global Trust Federation
JRA1	Joint Research Activity 1, Network Infrastructure Evolution
JRA1 T3	JRA1 Task 3, Integration with Other e-Infrastructures
JRA2	Networking Activity 3, Network Services Development
JRA2 t2	JRA2 Task 2, Service Provider Architecture
LDAP	Lightweight Directory Access Protocol
MMS	Membership Management Service
MOOC	Massive Open Online Courses
MoU	Memorandum of Understanding
MPiViT	Managing People and Performance in Virtual Teams
NA3	Networking Activity 3, Partner, User and Stakeholder Relations
NREN	National Research and Education Network
NA3 T2	NA3 Task 2, Account Management of International User Groups
NA3 T3	NA3 Task 3, e-Infrastructure Liaison
NOAD	National Open Access Desk
NREN	National Research and Education Network
OpenDreamKit	Open Digital Research Environment Toolkit for the Advancement of Mathematics
OpenDOAR	Directory of Open Access Repositories
PB	Petabyte
PRACE	Partnership for Advanced Computing in Europe

RDA	Research Data Alliance
REFEDS	Research and Education Federations
RnS	Research and Scholarship
SA1	Service Activity 1, Network Infrastructure and Services Operations
SA1 T3	SA1 Task 3, Network Implementation
SCI	Security for Collaboration among Infrastructures
Sirtfi	Security Incident Response Trust Framework for Federated Identity
SME	Small and Medium-sized Enterprises
TF-RED	Task Force on Research Engagement Development
TNC	The Networking Conference
TRL	Technology Readiness Level
VC	Videoconference
VRE	Virtual Research Environment
WG	Working Group
WISE	Wise Information Security for e-Infrastructures