Available online at www.prace-ri.eu

Partnership for Advanced Computing in Europe

The federated PRACE AAI Pilot

C. Padrin[a]*, S. C. Gorini[b]

[a]*CINECA Consorzio Interuniversitario*
[b]*CSCS Swiss National Supercomputing Centre*

Abstract

The increasing desire by researchers to have easy access to HPC services, and the objective to follow FAIR data access (Findable, Accessible, Interoperable, and Reusable), guided PRACE-6IP WP6 to identify common rules to provide resources in a more user-friendly way. The following white paper describes policies and architecture of an authentication and authorization infrastructure (AAI) agreed by PRACE-6IP partners.
In 2019, PRACE-6IP WP6 was tasked to define a plan for more user-friendly and forward-looking services for the future [1]. PRACE AAI was one of the results of this activity and a first pilot and a demo were released in May 2022 (section "Architecture and Pilot"). The aim was to provide user-friendly access to PRACE services. This document describes the agreed policies to release the service. Additionally, the document describes the policies about how to provide a trustable PRACE IdP (Identity Provider) to PRACE users to facilitate the access to non-PRACE services.
The work has been done in collaboration with Fenix [2], and this document has been partially inspired by work done in the ICEI project [3], and by the AARC Blueprint Architecture [4].

DOI: 10.5281/zenodo.7875077

1 The reasons behind PRACE AAI

PRACE aimed to deliver federated compute and data services to European researchers by allocating heterogeneous computational and data resources from multiple resource providers per project and enabling access from existing community platforms. In order to achieve these goals, the main idea was to define the PRACE IdP to be compliant with Fenix AAI policies and eventually to federate this PRACE IdP in the Fenix AAI. This is because Fenix wants to rely on a robust and reliable Authentication and Authorization Infrastructure (AAI), a trustworthy environment where users can be managed and granted to access resources securely and as seamlessly as possible. Additional advantages were that the Fenix federation involved five PRACE hosting member sites (BSC, CEA, CINECA, CSCS, JSC) and is also able to easily extend the service if other institutions want to join.

The goal of this document is to present the needs behind the realization of such an infrastructure, including requirements and the assumptions made during the designing and the technical implementation of the pilot.

For "seamless access" it is intended for a user, registered on a "trusted" IdP and granted to consume a certain number of resources, to access federated resources using his/her credentials (e.g. username/password, X.509, etc.) without going through a new registration process. In practice, a user registered at PRACE should be able to consume resources provided by a partner local site without registering twice and have the same user experience while accessing services. In other words, the experience that PRACE has aimed at providing, following Fenix, was to allow users to conduct their experiments without necessarily knowing the characteristics of the hosting site.

As the federation of distributed resources through establishing a central AAI has been largely discussed within other research infrastructures in Europe, our strategy is based on the outcomes of the Authentication and Authorisation for Research and Collaboration (AARC) project. The AARC project is a European initiative

providing guidelines, solutions and best practices to address the need for federated access to infrastructures. Thus, according to the design guidelines presented below, and the outcomes of the AARC project, our approach foresees the deployment of one service covering the following high-level functions: (*i*) Users identification and authentication, (*ii*) Federation of multiple IdPs by proxying authentication requests, (*iii*) Validation of user profiles, (*iv*) Policy registry and management of principles of engagement, and (*v*) Managing general PRACE Usage Agreement.

2 Policies of the service

The following policies were discussed by the PRACE partners with the aim to identify common rules which can be applied both for accessing and using a service provided by PRACE to the end-user. We identified guidelines for these common rules, and which identity attributes are needed to manage the authentication and authorization, largely inspired by the analysis performed by Fenix. The PRACE AAI should operate its own proxy, maintaining and following future implementation introduced by Fenix AAI in order to remain aligned.

2.1 Guidelines

The following guidelines are the result of the collaboration of PRACE-6IP WP6 with Fenix/ICEI. These principles have been defined to respect the local site policies of PRACE partners.

1. **Local site autonomy**: each partner providing a PRACE service can deny the authorization to a PRACE user if the local policies are not fully covered by policies of PRACE AAI. The same partner can ask the unauthorised PRACE user to accept missing policies, and this step can be required every time the partner will need it, according to local policies.
2. **Federate local IdPs**: each PRACE partner can choose to federate the local IdP. All federated IdPs will be trusted to authorise the user to a PRACE centralised service (e.g. Peer Review [5], web-portal [6], etc.).
3. **New user subscription**: no new PRACE service is required to manage the user registration to PRACE. The user email is the criteria to identify the user and allow the mapping procedure in case of multiple local accounts linked to the same user (c.f. item 4 below). Every user registered at local sites using the local IdP could become a PRACE user after subscription to the PRACE mailing list [7], accepting the PRACE Privacy Notice [8], the PRACE Data Protection Policy [9], the PRACE Website Legal Notice & Terms of Use [10], and after being provided with PRACE resources. PRACE users may also choose to leave the federation and only retain the local identity.
4. **Multiple accounts**: if a user has multiple accounts at different local sites, this user and all his site accounts must be mapped to a unique PRACE identity to avoid proliferation of identities.
5. **PRACE AAI concept**: the PRACE AAI is conceived as the collaboration of three services: a PRACE Central IdP which is responsible to proxy authentication requests among federated IdPs, an Attributes Provider (PRACE Peer Review [5] process) and a Resources Provider (hosting site) managing users authorization through budget allocation, groups, and roles. All the services put in place by the PRACE AAI (PRACE Central IdP and PRACE Peer Review [5] process), should be security audited by an external company before and during production phase.
6. **Federation**: each federated site needs to expose its user-base through its own IdP, supporting standard identification protocols, e.g. SAML v2 or OIDC. The site is responsible for the operation of its IdP and free to decide how users should authenticate to it, e.g. username/password, X.509, 2FA, RSA, etc. To be part of the federation, each federated IdP should be able to release a common set of user profile attributes in response to the authentication process. These attributes should meet the requirements of the REFEDS Research and Scholarship Entity Category (R&S) [11].
7. **Central IdP**: the PRACE Central IdP needs to be able to proxy authentication requests coming from external IdPs, such as communities' OIDC servers or the eduGain [12] federation. When a site or community does not support any external IdP, the PRACE AAI should be capable of authenticating and validating users against specific data sources (e.g. a local site LDAP).
8. **Services and resources**: since PRACE delivers various classes of services, the PRACE AAI needs to cope with this heterogeneity and thus support various access interfaces, either based on Internet protocols, such as HTTP(S) (Web Portal, SWIFT, etc.) or SSH.

2.2 Policy agreement among sites

A policy agreement will define rules and conditions a user should accept in order to register to a local site (or external IdP) and access the PRACE AAI. Two levels of policies have been identified:

1. **PRACE general policy:** this policy will be agreed on among all PRACE partners and include a common set of rules and criteria the user must respect, (e.g. about providing personal data during registration and so on).
2. **Local site policy:** this is the local policy valid for the specific PRACE partner, which can include some local extension to the subscription process.

2.3 PRACE user profile

The PRACE user profile is composed of a set of common attributes that serve to uniquely identify a user within the federation and it is managed by the PRACE Central IdP. These attributes need to be returned to the SP (Service Provider) as part of the authentication process to permit any SP to identify the user and grant him/her access to the service.

- Name
- Family Name
- Email
- Nationality (critical for the GDPR [13])
- Institution(s) or affiliation(s)
- PRACE identifier
- PRACE username

The "Institution(s) or affiliation(s)" attribute is a multivalued attribute which indicates the organisation (research institution or private company) the user is currently affiliated with and the type of the affiliation. A user can have several affiliations at the same time, and consequently multiple accounts at the various institutes or affiliations.

PRACE services could use such attributes in order to apply specific policies for the organisation as the user authenticates to the service. The syntax of the attribute could follow the eduPersonScopedAffiliation schema.

In general, the user affiliation is directly linked to the user email (e.g. CEA user has @cea.fr email domain), even if a user could also in principle have multiple affiliations.

As the user authenticates, the PRACE Central IdP should also export two more attributes

- Origin IdP the user is authenticating from (e.g. PRACE)
- LoA (Level of Assurance) of the origin IdP:
    - **HIGH** (PRACE site IdP, e.g. CINECA)
    - **MEDIUM** (home organisation IdP, e.g. CERN)
    - **LOW** (unreliable identity verification)

In order to support High-LoA profiles, each core site should provide some evidence that a vetting procedure takes place as soon as the user registers to the site. Moreover, each site should in principle delegate to a Security Officer and/or Project PI the validation of the user profile on the site. The partners will inform each other about the scope of their verification and vetting process.

In the context of AARC, some recommendations have been published about policy and practices of identity and attribute providers. These could help to guide and define more complex PRACE AAI policies in the future.

2.4 Generic Id profile

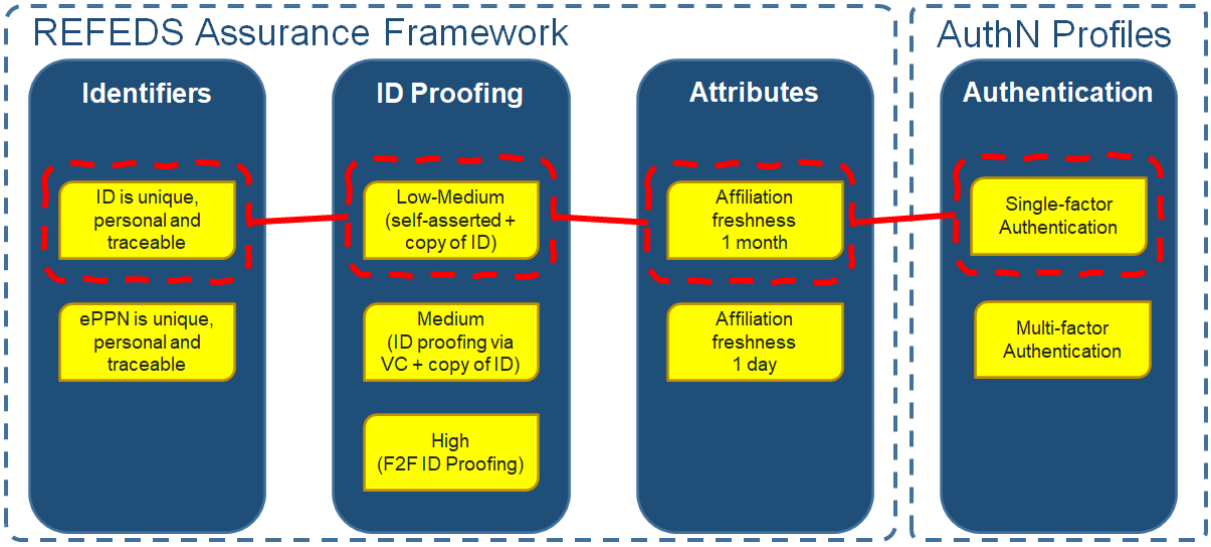
Each local site or community IdP should export a minimum set of attributes:

- Name
- Family Name
- Email
- Institution(s) or affiliation(s)
- Nationality (not mandatory, can be supplied if requested by specific local site policies)

2.5 Level of Assurance

WP6 guaranteed continuous monitoring and maintenance of operational services (like HPC and storage systems, network, repositories, and so on) under common agreed rules among partners. According to the service, it is needed to identify a Level of Assurance (LoA) to assign to users, with the aim to guarantee the autonomy of each partner.

After several meetings focused on this topic, partners' representatives in WP6 agreed to choose a light-medium level from the list of possible applicable levels described by GÉANT in a dedicated remote meeting held in September 2021, where different LoAs from REFEDS framework specification were presented to partners. The REFEDS framework LoA specification is described below, where different LoAs are assigned depending on vetting procedures and identity proofing.



In the REFEDS assurance framework selected, user's identifiers are considered unique, personal, and traceable (mandatory). Furthermore, a user's identity is self-asserted by the user sending a valid identity document (mandatory) and a user's affiliation could be validated periodically, in example each month (optional).

As "pro" we can underline that (*i*) it is not needed to define a new centralised dedicated service to verify the PRACE user's identity, and (*ii*) the user's self-assertion is supported by the user's identity document.

As "cons", we can underline that the LoA has a low-medium value, and it is not acceptable by several partners and external E-infrastructures.

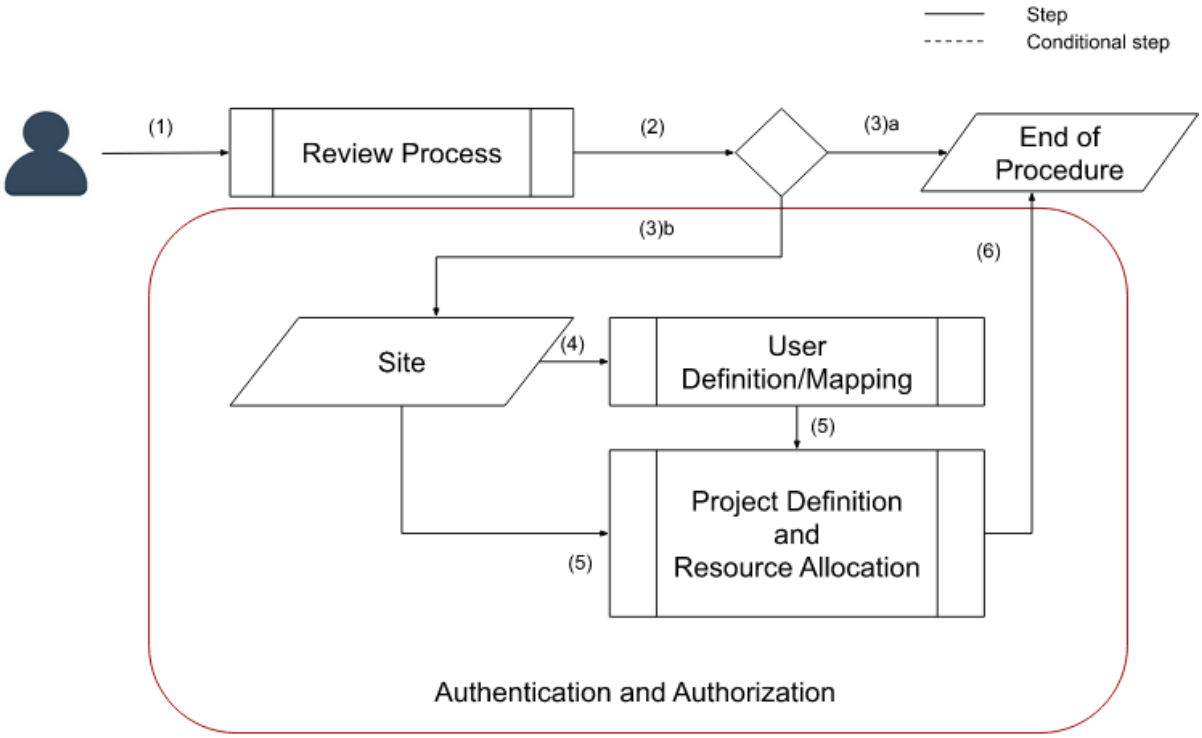
Looking at the cons described, that PRACE partner, who will want to guarantee the access to a PRACE User, can request additional steps to the PRACE User to complete the acceptance of those local policies missed by this level of assurance.

3 Architecture and pilot

The following section describes how a user can obtain PRACE resources with a particular focus on which steps can be managed with an AAI, the general architecture of PRACE AAI, and the pilot demonstrator.

3.1 User resource accounting main workflow

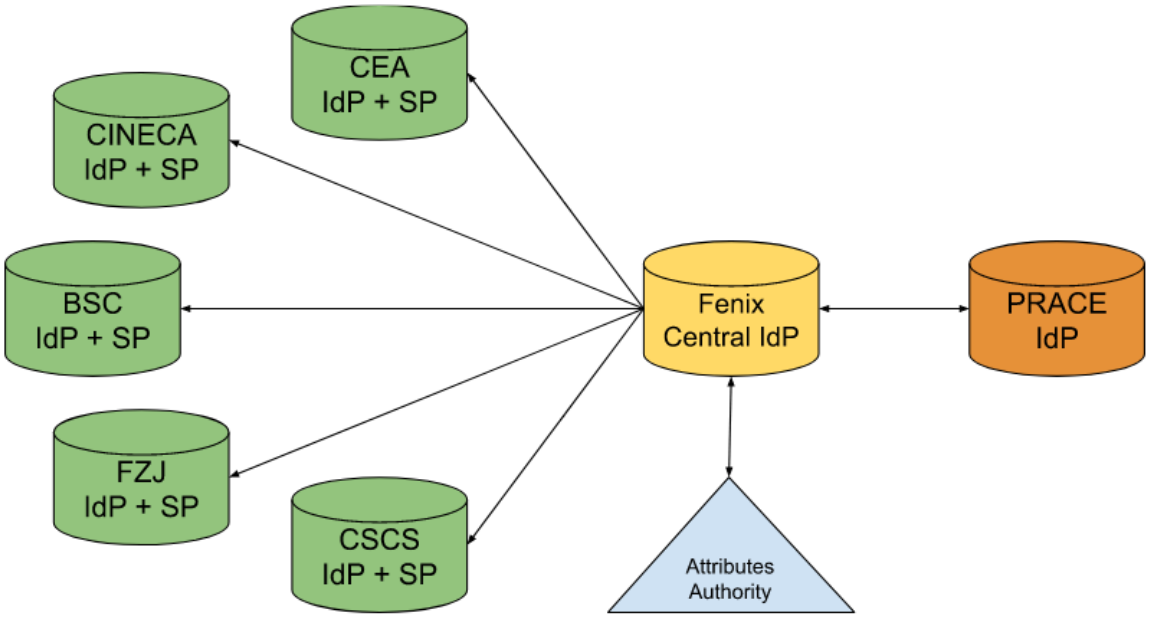
The architecture we have defined is based on satisfying the PRACE general use case described below: in particular, the workflow for the accounting process is presented.



- **Step (1)**: The user submits a request to the PRACE Review Process, asking for PRACE resources and services. *Not Managed by local IdP*
- **Step (2)**: Results of review establish if the request is accepted, partially accepted, or rejected. *Not Managed by local IdP*
- **Step (3)**: *Not Managed by local IdP*
  - ***a.*** If the proposal is rejected, the procedure ends.
  - ***b.*** If the proposal is accepted or partially accepted, the local resource management team receives the request to define a project and allocate resources on local systems hosting PRACE resources and services. The local resource management team checks user identity documents and provide user with access to HPC resources.
- **Step (4)**: The user connects via the Central IdP (e.g. Fenix IdP) with the credential obtained from the parent institution. If the **user** is already **locally defined,** the user can manage the mapping automatically. *Please, note this **does not mean** that the **user** has already **local credentials***. If the **user** is **not** yet **locally defined**, the **site** staff will create a local **user definition that will be automatically mapped.**
- **Step (5)**: The **site** will receive the project definition documents and will allocate resources accordingly. This automation will depend on site by site if they agree or not with the resource allocation mechanism. In the case of Fenix for example the allocation will be managed via FURMS.
- **Step (6)**: End of the procedure

3.2 PRACE AAI general architecture

We have identified the possibility to leverage the synergy with the Fenix infrastructure that is built on top five PRACE sites: BSC, CEA, CINECA, CSCS, and JSC.



Having decided that we will leverage the Fenix infrastructure we have defined a pilot to test if a PRACE user can effectively make use of it.

3.3 Goal of the pilot

The goal of the pilot has been defined as follows:

1. verify that on each site we can use a OAuth2/OIDC token as IdP and that with such an OIDC protocol we will enable any tool that is capable of using it without forcing any site to adopt any specific tool;
2. if most of the sites are aligned on the tool/protocol selected, this will be the pillar to enable the end user to build his workflow without having any additional specific site configuration improving the infrastructure portability (especially for the data transfer or any other similar middleware service based on OIDC).

3.4 Site requirements

Following these two goals in order to run the pilot each site should fulfil these requirements:

- Have an OIDC protocol, or a compatible one
- Setup a proper OIDC token lifetime (making sure site security rules are satisfied) according to the test purpose
- Configure the site as part of the Federation (PRACE AAI Proxy integration)
- Enable the federation mapping procedure
- Allocate resources to run a test
- Setup a UNCIORE service or, alternatively, a data transfer tool supporting OIDC protocol (or a compatible one)

3.5 Workflow

As possible workflow that such an infrastructure could enable is the following:

1. Login using ID credential of site1 thru the federation
2. Connect (automatic user mapping after first login) and submit to a site 2 using the token coming from the federation (point1):
    a. Unity (using UNICORE as submission tool)
    b. ssh session using OAuth2/OIDC token (Please see a possible technical solution description at: https://bscw.zam.kfa-juelich.de/bscw/bscw.cgi/d3730778/OIDC-based-SSH-access.pdf)
    c. Any other submission tool compatible with batch scheduler submission and OAuth2/OIDC token (ex. FirecREST, https://github.com/eth-cscs/firecrest)
3. transfer data [option as per point 2] from site 2 to site 1 using the same token:
    a. Token has to have a life longer than the job submitted if the 2 action are submitted at the same time
    b. If the transfer will be submitted in a separate session it will require another login action as per point 1

3.6 Open points

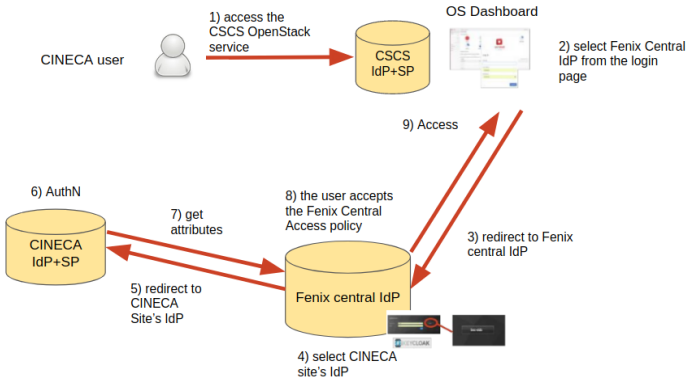
As today unfortunately the existing infrastructure is not able to full-fill the following:

1. MFA (Multi Factor Authentication)
2. Option for Data Transfer:
    a. to be done using FTS3 as it will be available
    b. UNICORE could be an option to transfer data
    c. Any other tool OIDC compatible should be consider usable (ex. FirecREST, https://github.com/eth-cscs/firecrest, Globus Online)

3.7 Pilot

Here below a description of what has been successfully test between CINECA and CSCS:

1. The CINECA user try to access the CSCS OS dashboard and is redirected to the Fenix central IdP where he/she can select his/her own affiliation (CINECA) on a web page
2. The user is consequently redirected to the CINECA IdP web page where authenticates by username and password
3. The user is redirected to the Fenix Central IdP where he/she accepts the Fenix general Usage Agreement, i.e., by ticking a checkbox. In case the user has not been provided with Fenix Credits on FURMS, or in case there are new updates of local site policies to be approved, the authentication flow stops here, and the user is redirected to a predefined web page.
4. Finally, the user is profiled into the Fenix Central IdP and redirected to the CSCS dashboard
5. The OpenStack Keystone SP takes care to map the user on the local user account and to apply local policies



References

[1] Plan for more user friendly and forward-looking services for the future - April 2020 - Final v1.1.docx
[2] https://fenix-ri.eu/about-fenix
[3] https://cordis.europa.eu/project/id/800858
[4] https://aarc-project.eu/architecture/
[5] https://prace-ri.eu/hpc-access/project-access/project-access-the-peer-review-process/
[6] https://hpc-portal.eu/
[7] https://prace-ri.eu/subscribe-to-prace-mailing-lists/
[8] https://prace-ri.eu/legal-corporate/privacy-notice/
[9] https://prace-ri.eu/legal-corporate/data-protection/
[10] https://prace-ri.eu/legal-corporate/legal-notice-terms-of-use/
[11] https://refeds.org/category/research-and-scholarship
[12] https://edugain.org/
[13] https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en

Acronyms

2FA	Two (2) Factor Authentication
AAI	Authentication and Authorization Infrastructure
HTTP(S)	HyperText Transfer Protocol (Secure)
ICEI	Interactive Computing E-Infrastructure for the Human Brain Project
Id	Identity
IdP	Identity Provider
LoA	Level of Assurance
OS	OpenStack
PI	Principal Investigator
SP	Service Provider
SSH	Secure SHell

Glossary

AARC	Authentication and Authorisation for Research and Collaboration initiative
CEA	Commissariat à l'Énergie Atomique et aux énergies alternatives
CERN	Conseil Européen pour la Recherche Nucléaire
CINECA	Consorzio Interuniversitario
eduGain	Interfederation service to connect identity federations around the world
Fenix	Fenix Research Infrastructure
GDPR	General Data Protection Regulation
LDAP	Lightweight Directory Access Protocol
OIDC	OpenID Connect
PRACE	Partnership for Advanced Computing in Europe
REFEDS	Research and Education FEDerations group
RSA	Rivest–Shamir–Adleman public-key cryptosystem
SAML v2	Security Assertion Markup Language version 2.0
x.509	Standard format of public key certificates

Acknowledgements

This work was financially supported by the PRACE project funded in part by the EU's Horizon 2020 Research and Innovation programme (2014-2020) under grant agreement 823767.