Federated Access to Collaborative Data and Compute Infrastructures

Ahmed Shiraz Memon

**Faculty of Industrial Engineering,
Mechanical Engineering and Computer Science
University of Iceland
2021**

Federated Access to Collaborative Data and Compute Infrastructures

Ahmed Shiraz Memon

Dissertation submitted in partial fulfilment of a
Philosophiae Doctor degree in Computer Science

Advisor
Morris Riedel

PhD Committee
Helmut Neukirchen
Matthias Book
Morris Riedel

Opponents
David Wallom
Shukor Bin Abd Razzak

Faculty of Industrial Engineering,
Mechanical Engineering and Computer Science
School of Engineering and Natural Sciences
University of Iceland
Reykjavik, June 2021

Federated Access to Collaborative Data and Compute Infrastructures

Dissertation submitted in partial fulfilment of a *Philosophiae Doctor* degree in Computer Science

Abstract

Distributed data and compute infrastructures aim to provide access to their data or compute services across disciplinary and geographical borders to their users for scientific research. The services are highly collaborative in nature yet independent and shared among multiple scientific communities. Information security and service discovery are two essential functions and precursors for enabling such research collaborations. Given the infrastructure's heterogeneity in data, compute, or other service offerings, the services often require several kinds of authentication protocols. Moreover, the users bring their own organisational identity and relevant attributes to access the infrastructure services. Should the services' authentication protocol differ from that of the user's, the user may not be able to access the target service. Therefore credential translation, attribute harmonisation, scalable trust and authorisation policy management need to be incorporated. In addition to that, enabling service discovery in the federated infrastructures is crucial. Proprietary service registration and query interfaces hinder interoperability across infrastructures. Hence, instead of proprietary and centralised registry approaches, a federated and standard-based registry and discovery model is essential for interoperability across the collaborating infrastructures.

This thesis is motivated by a case study consisting of three multi-national research infrastructures: compute (EGI), data management (EUDAT), and a community infrastructure supporting linguistic research (CLARIN). The thesis contributes EMIR, the European Middleware Initiative (EMI) Registry, a decentralised service registry that supports both hierarchical and peer-to-peer topologies and enables collaboration in large-scale infrastructures. The thesis also contributes the B2ACCESS service which implements a proxy model with credential translation and scalable trust and authorisation policy management. Finally, the thesis contributes an integrative architecture realised as a unified cross-infrastructure (or inter-federation) service access framework, which bridges EMIR and B2ACCESS to enable service discovery and access in federated environments.

Keywords— Federated Identity Management, Service Discovery, Authentication, Open Standards, User Management, Peer-to-peer

Útdráttur

Dreifð gögn og reiknistoðkerfi leitast við að gefa þverfaglegum aðilum, jafnt innlendum sem erlendum, aðgengi að gögnum eða reikniþjónustum til nota við vísindarannsóknir. Reikniþjónusturnar eru samhæfðar en jafnframt sjálfstæðar í eðli sínu, dreifðar um fjöldamörg vísindasamfélög. Upplýsingaöryggi og þjónustuleitir eru tvö nauðsynleg hlutverk og undanfari slíkra vísinda samhæfinga. Vegna þess að stoðkerfin innihalda misleit gögn, reiknigetu, eða annara aðgengilegra þjónustna, þurfa þjónusturnar margar tegundir af auðkenningarleiðum. Ennfremur þurfa notendur að styðjast við auðkenni viðkomandi stofnanna ásamt öðrum viðeigandi eigindum til þess að fá aðgang að stoð-kerfisþjónustunum. Ef auðkenningarleið þjónustunnar er frábrugðin frá leið notandans getur hann mögulega ekki fengið aðgang að tilgreindum þjónustum. Þar af leiðandi er þörf fyrir viðbót með skilríkjahliðrun, eigindasamhæfingu, stigfrjálsu trausti og heimildarstjórnun. Þar fyrir utan er mikilvægt að virkja þjónustuleitir í dreifðum inni-viðum kerfisins. Innskráning í sérþjónustur og fyrirspurnaviðmót hindrar samvirkni stoðkerfanna. Þar af leiðandi er mikilvægt að styðjast við dreift kerfi byggð á staðlaðri skrásetningu og leitarmódeli í stað miðlægrar skrásetningar fyrir samvirkni á milli ólíkra samhæfðra stoðkerfa.

Þessi doktorsritgerð er rökstudd með ferilsathugun sem notast við innviði þriggja fjölþjóðlegra rannsóknastofnanna: reiknistofnun (EGI), gagnaumsýsla (EUDAT), og samfélagsinnviði til styrktar tungumálarannsóknum (CLARIN). Framlag ritgerðinnar er EMIR, skráarsafn (e. registry) fyrir European Middleware Initiative (EMI), ómiðlæg þjónustuskrá sem styður bæði stigveldis grannfræði og deilitækni og býður upp á sam-vinnu í stórtækum innviðum. Að auki er framlag þessarar ritgerðar einnig B2ACCESS þjónustan sem nothæfir vefselsmódel með skilríkjahliðrun og stigfrjálsu trausti og auð-kennisreglustýringum. Að lokum er framlag þessarar ritgerðar einnig samþætt högun innleidd sem þjónstuumgjörð um samtvinnaða margþætta innviði sem brúa EMIR og B2ACCESS til að leyfa þjónustuleit og aðgang í dreifðu sambandsumhverfi.

Lykilorð— Umsýsla dreifðra persónuskilríkja, þjónustuleit, auðkenning, opnir staðlar, notandastýring, deilitækni (e. peer-to-peer).

Table of Contents

List of Figures

List of Tables

List of Publications

Paper I: L. Field, **A.S.Memon**, I. Márton; G. Szigeti, *"The EMI Registry: Discovering Services in a Federated World"*, Journal of Grid Computing 12(1), 29–40 (2014) [DOI: 10.1007/s10723-013-9284-1]

Paper II: **A.S. Memon**, J. Jensen, A. Cernivec, K. Benedyczak, M. Riedel, *"Federated Authentication and Credential Translation in the EUDAT Collaborative Data Infrastructure"*, IEEE/ACM 7th International Conference on Utility and Cloud Computing, IEEE, (2014). [DOI: 10.1109/UCC.2014.118]

Paper III: M. Hardt, A. Hayrapetyan, P. Millar, **A.S. Memon**, *"Combining the X.509 and the SAML Federated Identity Management Systems"*, Second International Conference on Security in Computer Networks and Distributed Systems (SNDS 2014): Recent Trends in Computer Networks and Distributed Systems Security, Trivandrum, India, 13–14 Mar 2014, Communications in Computer and Information Science 420, Springer, Berlin Heidelberg, pp. 404–415 (2014) [DOI: 10.1007/978-3-642-54525-2_36]

Paper IV: **A.S. Memon**, J. Jensen, W. Elbers, M. Riedel, H. Neukirchen, M. Book, *"Implementing an Authorisation Architecture in the EUDAT Services Federation"*, IEEE Conference on Application, Information and Network Security (AINS), Miri, Sarawak, 13–14 Nov 2017, IEEE, pp. 111–117 (2017) [DOI: 10.1109/AINS.2017.8270434]

Paper V: **A.S. Memon,**, J. Jensen, W. Elbers, H. Neukirchen, M. Book, M. Riedel, *"Towards Federated Service Discovery and Identity Management in Collaborative Data and Compute Cloud Infrastructures"*, Journal of Grid Computing 16(4), 663–681 (2018) [DOI: 10.1007/s10723-018-9445-3]

Abbreviations

AAI   Authentication and Authorisation Infrastructure

AARC   Authentication and Authorisation for Research and Collaboration

ABAC   Attribute-Based Access Control

ABFAB   Application Bridging for Federated Access Beyond Web

ACL   Access Control List

API   Application Programming Interface

ARC   Advanced Resource Connector

ATM   Asynchronous Transfer Mode

BBMRI   BioMolecular resources Research Infrastructure

BDII   Berkley Database Information Index

BPA   Blueprint Architecture

BPMN   Business Process Management

CA   Certification Authority

CIA   Confidentiality, Integrity and Availability

CLARIN   Common Language Resources and Technology Infrastructure

CLI   Command Line Interface

CRUD   Create-Read-Update-Delete

CSR   Certificate Signing Request

DARIAH   Digital Research Infrastructure for the Arts and Humanities

DHT   Distributed Hash Table

DLT   Distributed Ledger Technology

DNS   Domain Name System

DoS   Denial of Service

DSR	Domain Service Registry
DUDE	Distributed UDDI Deployment Engine
EAP	Extensible Authentication Protocol
eduGAIN	Global Authentication INfrastructure
EGI	European Grid Infrastructure
EIDA	European Integrated Data Archive
EMI	European Middleware Initiative
EMIR	EMI Service Registry
EOSC	European Open-Science Cloud
EPOS	European Plate Observing System
ERIC	European Research Infrastructure Consortium
EUDAT	European Data Infrastructure
EUGridPMA	European Policy Management Authority
FIM	Federated Identity Management
GDPR	General Data Protection Regulation
GEYSERS	Generalised Architecture for Dynamic Infrastructure Services
GIIS	Grid Index Information Service
GLUE	Grid Laboratory Uniform Environment
GOCDB	Grid Operations Centre Database
GRIS	Grid Resource Information Service
GRRP	Grid Resource Registration Protocol
GSI	Grid Security Infrastructure
GSR	Global Service Registry
HPC	High-Performance Computing
HTC	High-Throughput Computing
IaaS	Infrastructure-as-a-Service
ICOS	Integrated Carbon Observation System
ICT	Information and Communications Technology

ID-FF	Identity Federation Framework
IdM	Identity Management
IdP	Identity Provider
IGTF	Interoperable Global Trust Federation
ISIS	Information System Indexing Service
LAN	Local Area Network
LDAP	Lightweight Directory Access Protocol
LHC	Large Hadron Collider
LoA	Level of Assurance
LSDMA	Large Scale Data Management and Analysis
LTER	Long Term Ecological Research Network
MDS	Metacomputing Directory Service
MFA	Multi-Factor Authentication
MVEL	MVFlex Expression Language
MWSDI	METEOR-S Web Service Discovery Infrastructure
NGI	National Grid Initiative
NIST	National Institute of Standards and Technology
NREN	National Research and Education Network
NSF	National Science Foundation
OGF	Open Grid Forum
OIDC	OpenID Connect
ORFEUS	Observatories & Research Facilities for European Seismology
P2P	Peer-to-Peer
PaaS	Platform-as-a-Service
PAP	Policy Administration Point
PDP	Policy Decision Point
PEP	Policy Enforcement Point
PIP	Policy Information Point

PKI   Public Key Infrastructure

PRACE   Partnership for Advanced Computing in Europe

PR   Policy Repository

PRP   Policy Retrieval Point

RADIUS   Remote Authentication Dial In User Service

RAF   The Research and Education FEDerations group Assurance Framework

RBAC   Role-Based Access Control

REFEDS   Research and Education Identity Federations

RI   Research Infrastructure

SaaS   Software-as-a-Service

SAML   Security Assertion Markup Language

SAML 2.0 Security Assertion Markup Language V2.0

SIRTFI   Security Incident Response Trust Framework for Federated Identity

SOAP   Simple Object Access Protocol

SP   Service Provider

SR   Service Record

SSO   Single Sign-On

STS   Security Token Service

TCS   Terena Certificate Service

TO   thesis objective

TTL   Time-to-Live

UDDI   Universal Description Discovery and Integration

UML   Unified Modeling Language

VO   Virtual Organisation

VOMS   Virtual Organisation Management Service

WAN   Wide Area Network

Web SSO   Web Single Sign-On

WS   Web Services

WSDL   Web Services Description Language

XACML   eXtensible Access Control Markup Language

XSEDE   Extreme Science and Engineering Discovery Environment

Acknowledgements

I would like to express my sincere gratitude to my PhD supervisor Prof. Morris Riedel for giving me the opportunity to pursue my research interests and PhD studies. His insightful discussions and immense knowledge have allowed me to think critically by looking at the bigger picture yet not missing the minor details. The courses from him have added new feathers in my cap. I owe a debt of gratitude to Dr. Jens Jensen (STFC, UK), his rich knowledge and passion for information security and analytical skills have driven my motivation for the subject. It has been enlightening to work with him on the EUDAT, AARC, and EOSC-Hub projects. I also would like to express the deepest appreciation to my co-supervisors Prof. Matthias Book and Prof. Helmut Neukirchen for their invaluable input on the research questions and tremendous support in arranging the resources at the University of Iceland. Besides my supervisors, I am very grateful to Prof. Dr. Lippert, Norbert Attig and Daniel Mallmann from Jülich Supercomputing Centre, Forschungszentrum Jülich GmbH, for providing me adequate financial resources and platform. Without their precious support, it would not have been possible to pursue my PhD project. In addition to that, I would like to thank my project fellows Willem Elbers (CLARIN, Netherlands), Krzysztof Benedyczak (UNITY, Poland), Marcus Hardt (KIT, Germany), and Laurence Field (CERN, Switzerland) for their collaboration and stimulating discussions. I also want to express my gratitude to Ernir Erlingsson for translating the thesis abstract to Icelandic.

I am indebted to my parents, who have always been a source of inspiration and given me encouragement during the PhD studies. I am also very thankful to my wife for her care and unconditional support. The most important support have come to me by my sons Izaan and Subhan, I am very grateful to them for providing me love, affection, and making the whole research experience exciting and fun.

1 Introduction

With evermore research data becoming digitally available, scientific research communities make increasingly use of digital research infrastructures [58], which provide data and compute-related facilities, resources, and services that enable conducting top-level research in their respective fields. Such infrastructures may be "single-sited" or "distributed" (i.e. a network of resources). With an increasing demand for more computing power and data management facilities, the research communities choose to utilise shared resources from external cloud service providers (e.g. Amazon or Microsoft) or scientific e-infrastructures (e.g. the European Open-Science Cloud) [59, 146], this leads to heterogeneous and distributed infrastructures.

1.1 Motivation

This thesis focuses on two important aspects of sharing services and resources of distributed and heterogeneous digital infrastructures: service discovery and security with respect to authentication, authorisation, and identity management.

In cloud, data, or High-Performance Computing (HPC) infrastructures, services are usually deployed on specific middlewares, which enable communication and management of distributed services. Since these services are middleware-specific, they need to be advertised and discovered through middleware-specific registries by client (or service) applications. However, as each middleware registry defines its own protocol to publish, query, or model services, client applications cannot use services provided via another, incompatible middlewares. This restricts the use of infrastructure services offered by other middlewares deployed in a collaborative infrastructure. Consequently, it has become an immediate challenge to discover services in a middleware-agnostic manner, which is essential in composing and executing scientific workflows or even basic HPC jobs with data staging, that involves services from multiple middlewares.

The infrastructure services usually rely on one of the authentication schemes Public Key Infrastructure (PKI), Security Assertion Markup Language (SAML), or Lightweight Directory Access Protocol (LDAP). The users from a scientific community in need of access to an infrastructure service are initially obligated to register separately at each of these services. As a response, users receive new credentials (e.g. username and password or X.509 certificate) specific to each service, which the users have to keep manually in a safe location. This approach has become a barrier as the number of credentials maintained by the user depends on the number of services the infrastructure has to offer.

The end-user's home organisation is usually an authenticating party (also called Identity Provider, IdP), which (among authentication) manages the user's identity and information such as username, password, and other attributes. A relying party (or Service Provider, SP) can then request user authentication from the Identity Provider (IdP) and, based on the attributes, grants users access to the infrastructure service. The infrastructure contains several IdPs – the identity federation – and Service Providers (SPs) – the service federation; however, without any formal agreement on attribute naming between the two parties, the IdPs often release attributes which cannot be understood by the SPs. This is a fragile approach and does not scale well: as the number of IdPs and SPs increases, the number of agreements and mappings of attributes to SP's local representation grows as well. Furthermore, prior to any information exchange between IdPs and SPs, trust has to be created and managed, be it metadata exchange (SAML), client registration (OpenID Connect, OIDC), or adding public-key certificates of each other (PKI). The trust management requires each SP to register all the IdPs and vice-versa in an infrastructure. This trust management can become a bottleneck if there are several IdPs and SPs, because every change in an SP (e.g. endpoint or characteristic update) requires escalation to every single IdP in the infrastructure.

User authorisation is also an essential component of resource sharing and allows users to perform an action on an infrastructure service. Authorisation deals with *what* level of service is allowed to be accessed rather than *who* is accessing it. Bringing both authentication and authorisation together to gain access to the infrastructure services is also called Authentication and Authorisation Infrastructure (AAI).

Depending on the service, the access level or control rights can vary in terms of granularity. The access level is defined in an Access Control List (ACL), which is a list of policies or rules that define mappings of user roles (or other attributes) to individual service functions. In an e-infrastructure, such rules or policies are either kept at the service level or at a centrally managed policy repository. In a distributed and replicated deployment, a service has likely multiple instances deployed across multiple locations (e.g. B2STAGE [52]), and all the service's instances might rely on a central policy repository for managing (Policy Administration Point) as well as enforcing access control (Policy Enforcement and Decision Point). Being the single point of service level authorisation, the central authorisation service can easily become a single point of failure and eventually lead to unavailability.

The main theme of this thesis is to enable sharing of resources by *connecting the infrastructures*, with primary focus on *federated* service discovery, and *federated* service access and security (see Fig. 1.1). As a use case, this thesis uses three European production infrastructures with different sets of capabilities and purpose. It also covers the challenges of bridging the given infrastructures, that is, cross-infrastructure service discovery, security, and connecting the user communities. This thesis addresses the challenges starting from requirements, via the architecture, through to the implementation, and finally concludes with future developments.

Intra-connecting the infrastructure, as well as interconnecting it with other infrastructures, requires binding the services or elements (user or back-end resources) of an infrastructure with other services. The following are the key components that define such bindings:
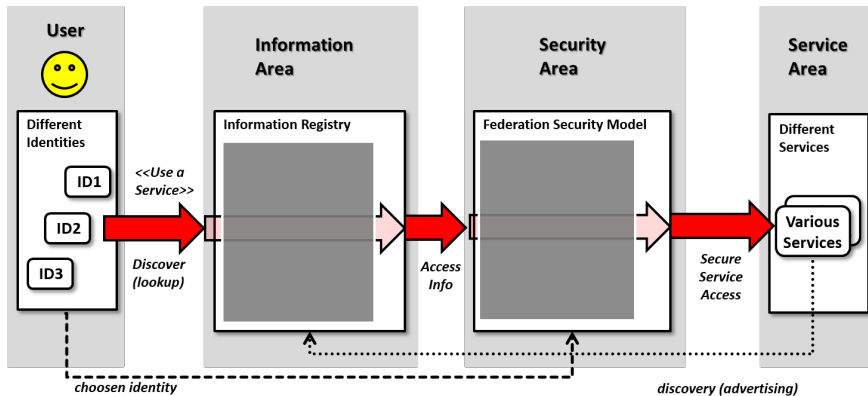
Figure 1.1. Integrated federated authentication and service discovery architecture

- Common fabric security or federation security model (as shown under "Security Area" in Fig. 1.1), i.e., X.509 host certificates from trusted Certification Authorities.

- Service naming: every infrastructure service must have a *name* or an *ID* by which it can be discovered and referenced; a typical type of name is a Web service's endpoint or URI.

- Service discovery/metadata: a set of attributes consisting of service capabilities in order to provide a way to discover which services would be available to the user or client applications.

- Service registry: an information registry (as shown under "Information Area" in Fig. 1.1) that enables registration and query of e-infrastructure services.

- Service information model: refers to an information model that captures (coarse to fine-grained) service details (or attributes) in a flexible manner; additionally, the model must be compliant with open standards to allow interoperable discovery of services in case of cross-infrastructure or low-level cross-middleware service access.

Thus, the central components of e-infrastructures are:

- Common authentication: a single credential to access any service in multiple infrastructures.

- Service discovery mechanisms: there has to be an "entry point" which helps users discover services that are available to them. Typically, this is a Web portal, but could also be hosted on a "user interface" node (to which users log in or connect with remote desktop);

- Service database (may or may not be service discovery): a database storing metadata of the services of the e-infrastructure.

3

- Common authorisation: a unified authorisation mechanism across the infrastructure to enable services to define coarse to fine-grained access control policies. This will allow users to share data and to collaboratively make use of the data and compute services provided.

1.2 Thesis Objectives and Contributions

The main thesis objective (TO) is to enable scientific communities as well as client applications to discover and securely access common digital infrastructure services through a unified set of interfaces and corresponding implementations.

The individual thesis objectives (TOs) of this thesis are as follows:

TO1 Analyse use cases in order to derive requirements for a design of service access (service discovery, authentication and authorisation) in collaborative and research infrastructures.

TO2 Create an architectural design of *secure* and *federated* service access based on a decentralised service registry approach, including aggregation of information and high availability through peering.

TO3 Implement a distributed and decentralised architecture realising the *secure* access to *federated* services using reliable and up-to-date or dynamic information.

TO4 Evaluate the developed architecture and software prototypes using case studies from several scientific communities.

The TOs include gathering and examining the requirements related to discovery and secure federated access to the services in the current European data, compute, and cloud infrastructures. Thereby, a systematic approach is taken while deriving the use cases and then conducting in-depth analysis of the gathered requirements. The requirements identify a unified approach for a robust federated identity management, authentication, authorisation (AAI), and service discovery solution.

The means to address the aforementioned requirements and the use case realisation are based on best practices and experiences from past and existing infrastructures and applications of the identified methodologies therein. Here, robustness and ease of service federation management are of primary interest. The goal is to identify the existing infrastructure management tools for service discovery and security (identity management, authentication and authorisation) applicable to the infrastructures' problems, especially when the domain scientists are required to use the services for their research across multiple infrastructures. If the given infrastructure management tools or implementations are not targeting the use cases, it may be viable to develop new or customise the existing tools to maximise multi-disciplinary research by incorporating multiple infrastructures' services. The novel implementations are validated and evaluated with respect to the scientific community-specific requirements.

The main contribution of this thesis is an AAI framework for federated service discovery, authentication, authorisation, and identity management to enable secure service access across the digital infrastructures. To ensure a broad applicability, the research

described in this thesis uses as input and for validation three different infrastructures as representative examples: one research infrastructure and two e-infrastructures, namely the generic e-infrastructures EUDAT (European Data Infrastructure) [53] and EGI (European Grid Infrastructure) [42] and the digital humanities research infrastructure CLARIN (Common Language Resources and Technology Infrastructure) [30]. The developed framework is open-source and being used in production by the real users of the European research and e-infrastructures.

The Business Process Management (BPMN) diagram [63] in Fig. 1.2 shows how the TOs relate to each other. The diagram begins with TO1 pertinent to the use case analysis — this includes investigating the case studies from three different compute and data research infrastructures, namely EGI, EUDAT, and CLARIN. TO1 also includes detailed requirements and its analysis from the perspective of service discovery and AAI. These requirements pave a way to create a joint model consisting of a system design and architecture (TO2). Based on the architecture, thesis objective TO3 focuses on the implementation of the service registry and AAI that provide a unified framework for federated service discovery, identity management, credential translation, user authentication (supporting multiple authentication protocols), and group management. In order to evaluate (TO4) the developed solution, the services have for testing and feedback been deployed at the service providers and for the end-users.
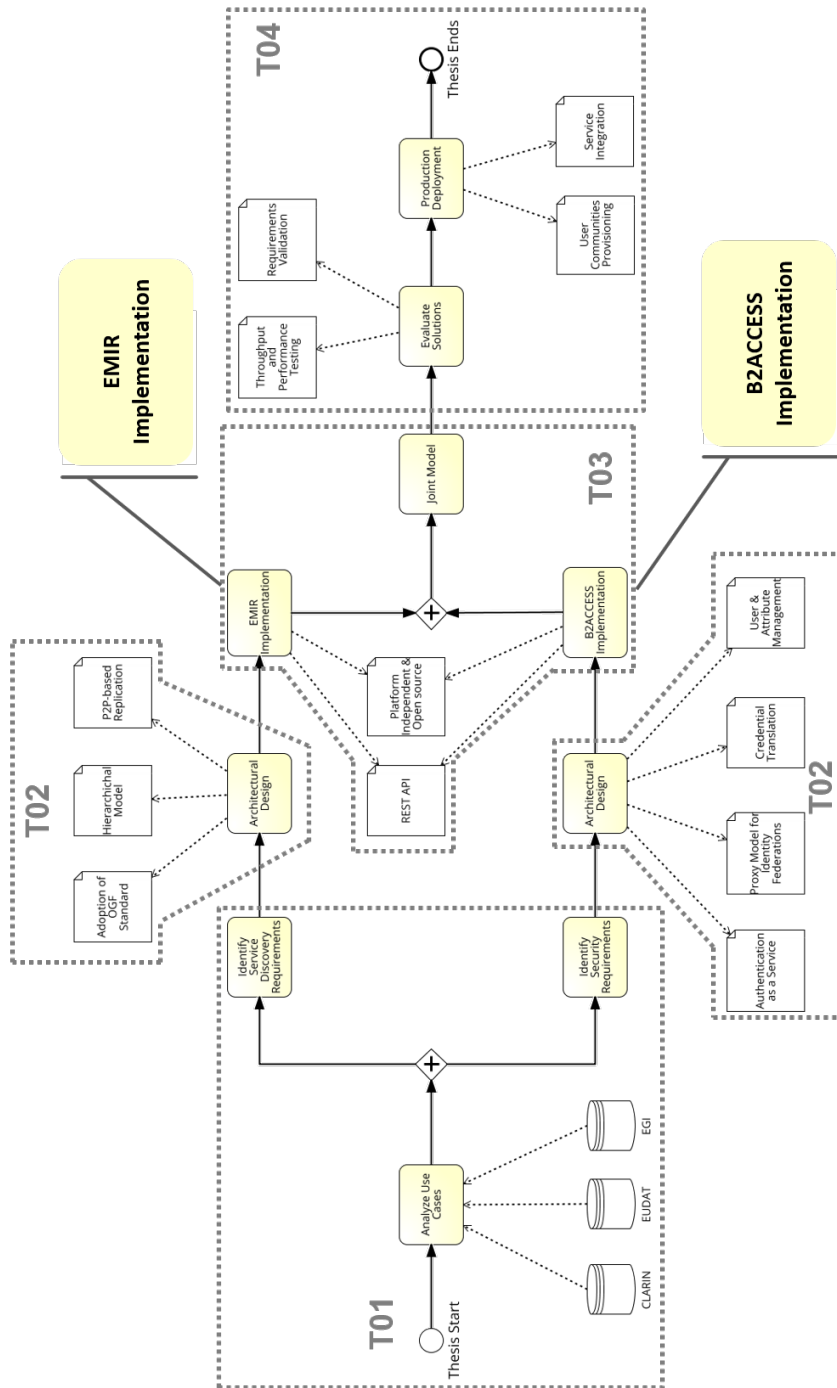
Figure 1.2. BPMN 2.0 diagram depicting the interrelation of thesis objectives

1.3 Outline

This thesis is written in a cumulative style and in addition to the publications, background, related work, and a summary of the main contributions are provided as further chapters. The order the publications has been chosen so that each paper builds on the results of the preceding papers and also according to the order of how an infrastructure end-user uses the services described in the papers (as opposed to the service provider). Therefore, the thesis initially explores the area of information services, wherein the end-user (from a scientific research community) initially discovers the services offered by the infrastructures; followed by the authentication and authorisation processes. Finally, the last publication defines a comprehensive architecture to enable users to discover and securely access the services in federated infrastructures.

1.3.1 Publications

The major contributions and findings have been published in peer-reviewed conference and journal publications:

Paper I: L. Field, **A.S. Memon**, I. Márton; G. Szigeti, *"The EMI Registry: Discovering Services in a Federated World"*, Journal of Grid Computing. 12(1), 29–40 (2014). [DOI: https://doi.org/10.1007/s10723-013-9284-1]

Paper II: **A.S. Memon**, J. Jensen, A. Cernivec, K. Benedyczak, M. Riedel, *"Federated Authentication and Credential Translation in the EUDAT Collaborative Data Infrastructure"*, 2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing (UCC), London, United Kingdom, 8–11 Dec 2014 IEEE 726–731 (2014) [DOI: 10.1109/UCC.2014.118]

Paper III: M. Hardt, A. Hayrapetyan, P. Millar, **A.S. Memon**, *"Combining the X.509 and the SAML Federated Identity Management Systems"*, Second International Conference on Security in Computer Networks and Distributed Systems (SNDS 2014): Recent Trends in Computer Networks and Distributed Systems Security, Trivandrum, India, 13–14 Mar 2014, Communications in Computer and Information Science 420, Springer, Berlin Heidelberg, pp. 404–415 (2014) [DOI: 10.1007/978-3-642-54525-2_36]

Paper IV: **A.S. Memon**, J. Jensen, W. Elbers, M. Riedel, H. Neukirchen, M. Book, *"Implementing an Authorisation Architecture in the EUDAT Services Federation"*, IEEE Conference on Application, Information and Network Security (AINS), Miri, Sarawak, 13 –14 Nov 2017, IEEE, pp. 111–117 (2017) [DOI: 10.1109/AINS.2017.8270434]

Paper V: **A.S. Memon,**, J. Jensen, W. Elbers, H. Neukirchen, M. Book, M. Riedel, *"Towards Federated Service Discovery and Identity Management in Collaborative Data and Compute Cloud Infrastructures"*, Journal of Grid Computing 16(4), 663–681 (2018) [DOI: 10.1007/s10723-018-9445-3]

For a summary of these publications, please refer to Chapter 4. The peer-reviewed papers themselves are provided as an appendix at the end of this thesis.

1.3.2 Thesis Structure

The structure of this thesis is as follows: Subsequent to this introduction, Chapter 2 provides the background on several service discovery approaches, federated identity management, authentication and authorisation models relevant to this thesis (more specific details are provided in the respective publications). Then, Chapter 3 defines a joint framework that addresses the service discovery and AAI requirements in contemporary large research infrastructures, in which services and users are located in different administrative domains. The focus is on two important aspects: first, the service discovery of geographically distributed services, and secondly, the AAI that includes federated authentication, authorisation and identity management. This chapter covers also the implementation of the developed services and provides a use case to evaluate the developed and implemented solution. A summary of the publications underlying this thesis is provided in Chapter 4. Finally, Chapter 5 concludes the thesis with a summary, a description of the impact of the work, and an outlook on future prospects. An appendix provides the full text of the publications that are part of this cumulative thesis.

2 Background

This chapter presents important concepts of service discovery in High-Performance Computing (HPC) and High-Throughput Computing (HTC) infrastructures and the adopted information models that represent the distributed services abstracting the hardware and software resources. Besides service discovery, secure user access to the infrastructure services (and the underlying resources) plays an important role for the scientific user communities in accessing the shared e-infrastructure services or resources. The secure access implies user authentication and authorisation. This chapter also covers identity federation models, authentication protocols, authorisation models, and user management.

2.1 Grid and Cloud Computing

The term *Grid Computing* was conceived in a workshop held at Argonne National Laboratory in September 1997 [36]. The main notion was to access a set of heterogeneous computing resources deployed in different administrative domains as if a single computing resource is accessed (analogous to the electrical Grid concept). Hiding the complexities of resource locations and network connectivities is also referred to as *Metacomputing* [135].

The metacomputing concept was realised 1995 in the I-WAY project [39]. This project aimed to provide an environment to connect multiple supercomputing centers in North America. I-WAY was a collaborative environment at a large scale and it was based on Asynchronous Transfer Mode (ATM) networking technology as the traditional Internet connectivity (due to low bandwidth) was at that time not able to transfer video, audio, or images efficiently enough. Despite being collaborative, I-WAY lacked a uniform software environment to access distributed supercomputing resources. In addition to that, it also lacked real-time status and systems structure information about the distributed resources which was indispensable to make configuration decisions. Given the challenges, Globus Toolkit [67] was first deployed in the I-WAY project. The toolkit offered a Metacomputing Directory Service (MDS) [62] to allow accessing the resource information.

The notion of metacomputing paved the way for Grid computing and then cloud computing. Grid computing [68] based on, e.g., Globus Toolkit [65] and the more recent Globus cloud services [66] enables resource sharing across physical organisational boundaries using virtual organisations in a secure manner. Cloud computing leverages virtualisation technologies in order to offer a set of heterogeneous data or compute resources as a commercial service to its users in a pay-as-you-go manner [8]. Cloud

services are mostly offered at three different levels: Infrastructure-as-a-Service (IaaS), i.e. bare metal or virtualised operating images, Platform-as-a-Service (PaaS), i.e. remote developer-oriented services and APIs for developers to implement end-user applications, and Software-as-a-Service (SaaS), i.e. scalable end-user applications [97]. Grid and cloud middlewares abstract the underlying computing services and resources, though the challenge of service discovery (posed by the middlewares) with respective registries remains. That heterogeneity of registries define new barriers to discover the infrastructure services which this dissertation aims to overcome.

2.2 Research and e-Infrastructures

In 2004, the e-Infrastructure Reflection Group (e-IRG) [146] was founded. The group is a strategic body which aims to facilitate the integration of European e-infrastructures and HPC, HTC, and data management services and sharing of resources. The group's definition of an "e-infrastructure" comprises the fabric (disk, CPU cores, networks) underpinning the "middleware", which is an abstraction layer connecting the infrastructure. The middleware layer also enables the deployment of domain-specific applications (from the research communities) on the e-infrastructure [146].

Grid computing enables the development and use of e-infrastructures with more emphasis on sharing of resources and services in a seamless manner (see Section 2.1). The established e-infrastructures such as TeraGrid [92], European National Grid Initiatives (NGIs), Extreme Science and Engineering Discovery Environment (XSEDE) [153], and the world-wide and global scale Large Hadron Collider (LHC) Computing Grid [125] have been successful in providing resources and support to a wide spectrum of research communities. Digital compute and data infrastructures are either *research infrastructures* or *e-infrastructures*, the main distinction being the target audience and types of services or resources offered.

A research infrastructure (e.g. Common Language Resources and Technology Infrastructure (CLARIN) [30], Digital Research Infrastructure for the Arts and Humanities (DARIAH) [10], or BioMolecular resources Research Infrastructure (BBMRI) [25][1]) is a technology provider offering domain-specific services to the scientists. The European Commission defines a research infrastructure as a scientific community that uses resources, services, and facilities in order to conduct research in their respective fields [58].

An e-infrastructure, however, is independent from any particular scientific community and has a user base with a wide spectrum of requirements. The e-infrastructure thus enables open science by sharing computing and data resources, software, and data across multiple scientific communities, disciplines, and domains. One of the biggest initiatives in Europe is the European cloud Initiative [59], which has paved the way for European Open-Science Cloud (EOSC) [49] to promote data-driven open science.

Despite of being similar in nature, both e-infrastructures and research infrastructures have different scopes, characteristics, challenges, and provide services at different scale. An e-infrastructure is driven by sharing of resources, thus requires accounting at a

---

[1] All having the legal status of being an European Research Infrastructure Consortium (ERIC).

fine-grained level as the offered services are shared among many scientific users from multiple domains [56]. In contrast to a research infrastructure, an e-infrastructure has usually a broader portfolio of services and applications (generic and domain-specific). Moreover, they need to support to handle many users and service provider requests. In addition to the nature of scale, e-infrastructures are also more heterogeneous than research infrastructures in terms of the use of service discovery, authentication and authorisation schemes. Given the various characteristics of services, an e-infrastructure shares its resources among scientists from various scientific domains (e.g. digital humanities, health, and environmental sciences) [57]. Furthermore, they do have means to federate the services and can expand their offerings based on the user's demand. On the contrary, scientific research infrastructures have a limited number of users and services, thus requiring much fewer resources as compared to e-infrastructures. To offer their services, scientific research infrastructures can take advantage of the services of e-infrastructures.

Both scientific and e-infrastructures have a significant relevance for this dissertation as it focuses on the requirements and realisation of use cases from European Data Infrastructure (EUDAT) and European Grid Infrastructure (EGI) as example e-infrastructures and CLARIN as research infrastructure and scientific community.

2.3 GLUE information Model

GLUE 2.0 [40, 130] is an Open Grid Forum (OGF) specification, successor of GLUE Schema 1.x [2], presenting a conceptual and platform independent information model for Grid, compute, and storage entities in human and machine-readable format. The entities and their associations follow a service-oriented approach; hence they are derived from an abstract *Service* entity. The abstraction also enables derivation of other types of infrastructure entities or services, which are not provided in the specification. Moreover the specification allows selecting a subset of entities to model the infrastructure services, thereby not compelling the implementers to use the whole set of entities. One such example of model extensibility is the latest GLUE 2.1 specification [108] which has initially introduced the cloud computing and underlying virtualisation concepts as service extensions and then the extensions have become a part of the standard. The specification can be endorsed by a cloud computing infrastructure to not only represent its service offerings but also to be interoperable with other cloud infrastructures. The associations among all the services and entities are illustrated using Unified Modeling Language (UML) [50] class diagrams. In order to incorporate the GLUE schema into real world applications, the specification provides concrete renderings in XML [131], LDAP [3], and JSON [137] formats in the respective OGF documents.

The GLUE 2.0 specification has played an important role in this dissertation as it provides a means to implement interoperable information systems in the production Grid, data, and cloud infrastructures (see Chapter 3), for example in XSEDE [153, 138], NorduGrid [6], and particularly EGI [24].

2.4 Authentication and Authorisation Infrastructure

An Authentication and Authorisation Infrastructure (AAI) simplifies inter-organisation access by enabling users to access non-Web and Web-based services (data, compute, virtualisation, etc.) deployed on one or more infrastructures [141]. The Web-based services generally use HTTP as a base protocol for communication and can be categorised into browser-based and non-browser-based services. The browser-based services require interactive access by an end-user's Web browser and they enable human-to-machine communication (using HTML). In contrast, the non-browser-based services usually enable machine-to-machine communication through Web services (using REST or SOAP). The non-Web-based services are non-interactive in nature and are generally databases, virtual machines, storage devices, and file transfer services (e.g. FTP, GridFTP [1], UNICORE's UFTP [129], or Git). In certain scenarios, Web-based services can also enable access to its counterpart by generating special security or access tokens (e.g. STS [140], MyProxy [107] or SSH credentials), which can be consumed by a client application to access the non-Web-based services. From the user's perspective, AAIs facilitates the use of a single 'virtualised' digital identity, which is issued from the user's home organisation to access the given Web and non-Web-based services. Furthermore, it saves resources of administering and registering users by reducing the overhead of paper work (by the service or infrastructure operators) and adopts standardised authentication and authorisation mechanisms. As the user information is exchanged between multiple entities, AAI is expected to protect that information while complying with the data privacy standards, such as EU's General Data Protection Regulation (GDPR) [55].

The following sections elaborate on the core concepts and technologies that enable AAI, in particular Federated Identity Management (FIM). AAI also is one of the major components of this dissertation by focusing on the challenges and solutions to provide cross-infrastructure access to the provided services in a secure and reliable manner. AAI facilitates across-infrastructure access in a standardised and secure manner by means of the following key functions: identity management, credential translation, attribute harmonisation, group management, trust management, user (de-)provisioning, and access control.

2.5 Digital Identity

According to ISO/IEC standard 24760-1 [85], identity is an information representing an entity in an Information and Communications Technology (ICT) system. In the context of this thesis, the entity in possession of digital identity can be a real person from a research infrastructure trying to access an e-infrastructure service. Alternatively, the entity can be a service that is performing a scientific task on behalf of a real user or service.

Bertino et al. [21] succinctly describe the building blocks of a digital identity, consisting of three elements: an *identifier* that uniquely identifies the owner, a set of *attributes* corresponding to the identity, and finally an associated *credential* that is

used to establish confidence that the entity corresponds to the claimed person. These elements are essentially used to authenticate and sometimes authorise entities or users to obtain access to the organisation or infrastructure's resources.

Compute or data infrastructures implement *identity management* processes [21] to ensure secure access and mitigate the risks of unauthorised access, which can threaten Confidentiality, Integrity and Availability (CIA) of assets [72]. Unfortunately, insufficient focus on identity management by some organisations can unintentionally cause (sometimes costly) security incidents [69].

This thesis enables management of digital identities within an organisation and also incorporates the following five primary functions [21]:

Creation of an identity may require presenting the applicant's real identity or passport for proofing, or in some cases it can be a lightweight process of self-registering on an organisation's website. In the latter case, the user may not be able to access the privileged services or methods (e.g. uploading data on the cloud or submitting an HPC job).

Use of identity implies authentication of the entity based on the presented identity by the authentication service. Authentication is then further divided into identification and verification processes [132].

Update of identity attributes reflects maintenance, activation, adjustment, archival and restoration activities [85]. Additionally, organisations adopt a procedure to regularly check the identity attributes and keep them up to date.

Revocation of identity refers to the identity deletion or suspension, as soon as the user possessing the identity departs from the organisation. The process should additionally inform all the stakeholders or Service Providers (SPs) of the revocation.

Governance implies the creation and implementation of policies of the identity management lifecycle. For example, in case of creation, the policy defines how the identities will be created and who has the suitable role to perform the privileged operation. The National Institute of Standards and Technology (NIST) suggests implementing consistent and accurate policies to minimise the exploitation of attackers, vulnerabilities, impersonation, and Denial of Service (DoS) attacks [72].

Digital identities are usually linked with the information, such as username and password, that can primarily be used to provide access control. However, in federated environments (such as EUDAT or EGI), diversified services are usually offered or expected to be integrated, and Level of Assurance (LoA) can be seen as a way to provide fine grained access control [104]. The most widely known approach is to assign an assurance level after adequate verification during the credential enrolment process by a registration authority or Identity Provider (IdP). Level of Assurance or a degree of assurance also defines a quality of the identity. The level can be assessed while taking the following factors into account [124]: assurance components and conformance criteria. There are two further essential components: i) uniqueness of issued identity, ii) process or method of identity proofing and credential issuance, renewal and replacement. Moreover, the conformance criteria are applicable to the

Credential Service Provider (CSP) or Identity Provider (IdP): it is pertinent to the adherence of a set of guidelines, such as, if the IdP is operated with organisational-level authority, following generally accepted security practices and whether the contact information (or other metadata) about the IdP is correct. Based on the given components and criteria, The Research and Education FEDerations group Assurance Framework (RAF) [124] defines different profiles (named "Capuccino" and "Espresso"). There are also X.509-specific profiles defined by the Interoperable Global Trust Federation (IGTF) [73] which assigns LoAs to the identities (or generated certificates) based on the type of the issuing authority. On top of the given profiles, the Authentication and Authorisation for Research and Collaboration (AARC) consortium specified a set of guidelines to define LoAs for the linked identities of a user [119]. The guidelines are useful when a user has two or more identities (with different LoAs) issued by different organisations or IdPs – this also includes social media (Google, GitHub, Facebook, ORCID). The assurance level information (generated by single or linked identity) is usually signalled to the SPs in the 'eduPersonAssurance' attribute of the 'eduPerson' object class specification[2] using the Security Assertion Markup Language (SAML).

2.6 Federated Identity Management

One of the important objectives of large-scale research and e-infrastructures is to establish a robust Federated Identity Management (FIM) system that enables resource sharing across corporate boundaries through identity management systems. FIM also facilitates users to authenticate themselves only once, hence avoiding redundant login flows. In FIM-enabled infrastructures, the collaborating partners agree on a set of common Identity Management (IdM) policies and use standards-based or interoperable IdM technologies. They either rely on national identity federations [74] or use their own IdPs. Depending on the requirements, the architecture can be distributed or isolated [136]. In contrast to the isolated or central approach, the distributed approach (see Sect. 2.7) requires implementing identity management processes across security domains [94]. Regardless of the given approaches, a fundamental component and precursor to the usage of the identity management is *trust* [136, 28], which essentially enables the collaborators to rely on each other for identity management functions [21].

2.7 Identity Federation Architectures

An identity federation architecture essentially defines how Service Providers (SPs) and Identity Providers (IdPs) are connected with each other and/or with intermediaries (for example proxies). According to [74], there are three types of identity federation architectures: full mesh federations, hub-and-spoke federations with distributed login, and hub-and-spoke federations with central login. Since each architecture has its merits and demerits, the infrastructures or enterprises choose the suitable architecture that addresses their FIM requirements.

---

[2]https://software.internet2.edu/eduperson/internet2-mace-dir-eduperson-201602.html

Full mesh federation architecture is the most widely adopted architecture, it has no central component, every organisation has an IdP with their users' database and a set of SPs. All the entities in a mesh federation are registered in a SAML [117] (see also Sect. 2.9) metadata file – a building block of trust that describes the entity attributes that will be released and consumed by every registered IdP and SP. The most evident shortcoming of the mesh federations is the management of the metadata file and handling the release of attributes from the IdPs.

A *hub-and-spoke federation with distributed login* is the evolution of mesh federations and precursor to the AARC project's [12] *Blueprint Architecture* (see Sect. 2.8), where a central proxy or hub is deployed, which acts as a SP for IdPs and exposes an IdP interface to the federation SPs. The given architecture has significantly eased the burden of trust management from the federation partners.

Finally, a *hub-and-spoke federation with central login* architecture is also proxy-based, but one of the dissimilarities to federations with distributed login is that it has no IdPs-facing SP interface. Moreover, the establishment of trust requires configuration of the proxy's metadata at each SP and likewise for the proxy, which keeps track of all the underlying SPs' metadata.

2.8 Authentication and Authorisation Research Consortium's Blueprint Architecture (AARC BPA)

The AARC Blueprint Architecture (BPA) [22] is an AAI architecture, collaboratively designed by the research communities, AAI experts, developers, and infrastructure operators within the EU-funded AARC project [12]. BPA consists of essential architectural building blocks and implementation patterns extracted from the existing research and e-infrastructures, such as ELIXIR [44], European Plate Observing System (EPOS) [14], EGI [42]), and also EUDAT which is a part of this thesis (see Sect. 4.2).

BPA allows the current and future infrastructures to design and implement interoperable, interdisciplinary research collaborations through AAI. In addition to the architecture, it defines a set of guidelines and best practices for non-Web-based services, token translation services (TTS), managing authorisation, harmonised expression of group membership and role information, attribute aggregation, and credential delegation.

2.9 Authentication and Authorisation Standards

The Security Assertion Markup Language V2.0 (SAML 2.0) [117] is an open standard from OASIS[3] based on the Liberty Alliance's[4] Identity Federation Framework (ID-FF) that defines syntax and semantics to exchange authentication and authorisation data between security domains. The data is an assertion about an end-user or entity requiring

---

[3] https://www.oasis-open.org
[4] http://www.projectliberty.org

resource access. The assertions are security tokens in XML format, issued by a SAML Identity Provider (IdP) (the authenticating party) and consumed by a Service Provider (SP) (the relying party) – the SAML entities. The specification consists of metadata [26], protocols [117], bindings [83], and profiles to support a wide spectrum of use cases, environments, and technologies. For example, SAML 2.0 standardises Web Single Sign-On (SSO) in its 'Web Browser SSO Profile' [84], which enables the end-users to access multiple Web browser-based services while authenticating themselves once with a set of credentials. The SSO profile on the one hand reduces the overhead of managing (and storing) multiple credentials at the end-users, and on the other hand the SPs can focus on their core functionality (without taking care of authentication). A predominant SAML-based world-wide identity and service federation is eduGAIN [74].

OAuth 2.0 [77] is a delegated authorisation framework that enables third-party applications to acquire temporary access to an HTTP service on behalf of the resource/service owner. The access is granted by an authorisation service and enabled by a security "access token". The token is issued to the third-party application either by orchestrating the approval between the resource owner and the service or by the third-party application on its own. OAuth 2.0 does neither specify how the users are authenticated, nor the authentication flow, nor the content of the access token, nor the validation of access token, nor the representation of user information. In order to standardise these features, OpenID Connect (OIDC) [127] by the OpenID foundation has defined an authentication layer atop the OAuth 2.0 framework that is widely supported by social identity providers (Google, Facebook, Microsoft, ORCID, etc.).

The eXtensible Access Control Markup Language (XACML) [109] is, like SAML, also an open standard from OASIS, focusing mainly on authorisation of resources and HTTP services. XACML specifies a declarative policy language, a request/response scheme, and a reference architecture. It allows implementing Attribute-Based Access Control (ABAC), thereby enabling easier authoring of policies and fine-grained access control. NIST provides a set of guidelines and considerations to implement ABAC with XACML within enterprises and infrastructures [82]. Being an open standard, XACML supports multiple vendors, platforms, and implementation interoperability. In order to support various use cases, XACML-based profiles are specified. For example, the *administration and delegation profile* is used to define the policies to delegate (and limit) administrative rights to local administrators to enforce access control on a subset of resources [112]. Similarly, the *REST profile* [116] defines a RESTful API to communicate with the XACML architectural components (Policy Decision Point (PDP), Policy Administration Point (PAP), and Policy Enforcement Point (PEP)), whereas the *SAML profile* [111] focuses on SAML and XACML integration. In this thesis, the above mentioned authentication and authorisation standards have played an essential role in achieving the interoperability with other infrastructures and identity federations.

2.10 Related Implementations

Service discovery plays a prominent role in the distributed compute and data infrastructures, it is a precursor for enabling user access and scientific workflows that utilise cooperating infrastructure services.

As discussed in Section 2.1, Globus Toolkit [62] introduced a high performance distributed information system Metacomputing Directory Service (MDS) for Grid environments, which consists of the Grid Resource Information Service (GRIS) and the Grid Index Information Service (GIIS), whereby the former enables querying for the resources and the latter responds to queries against an internal cache and forwards them to the other GIISs. The GRIS registers with GIIS using a soft-state registration protocol called Grid Resource Registration Protocol (GRRP) implying that every registration information stored has its own associated Time-to-Live (TTL) property and the registration becomes invalid unless refreshed before the TTL value expires. Since MDS supports hierarchical deployment and is based on a federated model, a hierarchy can be formed where multiple GIIS instances register with a higher level GIIS instance.

Due to the MDS instabilities [93] observed during the European DataGrid project [70] evaluation, a top level cache based on OpenLDAP [29] server (labelled as Berkley Database Information Index (BDII)) was developed. The service had leveraged from the LDAP hierarchical model and therefore significantly increased the performance under high load [93]. BDII as compared to MDS maintains a local file instead of an index that contains the URLs of the lower-level services. In addition to that, BDII does not forward the queries to the indexing service. Instead, it responds directly from an internal cache which gets the updates periodically by querying the indexing service. The local file (consisting of URLs) also gets updated from the Grid Operations Centre Database (GOCDB) [96].

In order to address the shortcomings of MDS, a distributed information system was developed within the NorduGrid project [142]: Information System Indexing Service (ISIS), a part of the Advanced Resource Connector (ARC) middleware [46]. The cache in ISIS was removed and instead the queries are performed in two steps: collect a list of services or resource endpoint URLs from the enhanced GIIS and then, based on the URLs, query the ARC GRIS. The enhanced GIIS contains also the endpoint URLs of other GIIS instances, which makes it a hierarchical structure based on geographical location. The hierarchy consists of a region index, in which all the GIISs within the same region register, a level up is the country index which consolidates all the regions within that country, the country level index then again is registered with a top-level index service. To avoid a single point of failure, multiple top-level indexing services are deployed and connected with each other using a Peer-to-Peer (P2P) network [6]. A client library is provided to conceal the multiple queries from end-users. The disadvantage of the ISIS approach is that the multi-query approach does not perform well at scale due to multiple hops from one GRIS to the other.

A further related implementation is the UNICORE [20] middleware for HPC resources that has a centralised (or global) service registry, which is based on Web services technology.

The Universal Description Discovery and Integration (UDDI) [37] is another standard, which is used to register and discover Simple Object Access Protocol (SOAP)-based Web services. The service information is stored in a platform-independent XML-based registry and uses the SOAP messaging protocol. Despite of being an industry standard it has not been widely adopted, limited to private deployments within organisational boundaries; this is partly due to a lack of replication among registries and lack of autonomous control, and in addition, UDDI does not support discovering of

other UDDI registries. In order to cope with the given shortcomings, Distributed UDDI Deployment Engine (DUDE) was introduced using a Distributed Hash Table (DHT)-based approach that implements querying of multiple UDDI registries by a rendezvous mechanism [15].

Likewise, the METEOR-S Web Service Discovery Infrastructure (MWSDI) is designed for a multi-registry environment, thus enabling publishing and discovery of a group of autonomous registries [134]. MWSDI makes use of a semantic Web approach (using ontologies) by grouping registries into domains and then further grouping domains to federations.

There are service discovery protocols [154], such as Service Location Protocol and Jini, that are mainly designed for Local Area Network (LAN) and are not feasible for Wide Area Network (WAN). Their centralised approach is clearly not suitable for large scale infrastructures. Instead, federated registries that enable usage of two or more autonomous but cooperating registries are essential for service discovery and should be technology and middleware-agnostic.

AAI, including FIM, is an essential part of any infrastructure. Being a challenge for many years [86], service providers are now relying on external identity management solutions – instead of implementing built-in authentication mechanisms. The scientific user communities have their own established AAI, hence they have their own identities (based on different types of credentials).

The e-infrastructures like EUDAT enable collaboration within and across research communities and their infrastructures by offering secure and federated data management services [14]. ELIXIR [44] is a part of the Corbel [35] project and one of the biggest life-science communities in Europe, operating its own AAI [45]. The main goals of ELIXIR are quality control, collection, the archival and long-term sustainability of large amounts of data that is generated from life-science experiments. ELIXIR's AAI is connected with the Global Authentication INfrastructure (eduGAIN) [74] to provide Web identity federation, thus allowing users to authenticate with their institute identity. Moreover, users can also authenticate with their social identities (such as Google, Facebook, or ORCID). The AAI also enables account linking between remote and infrastructure-wide user identities.

XSEDE [153] (the successor of TeraGrid [92]) is another HPC and Grid infrastructure, funded by the National Science Foundation (NSF). It offers digital resources and services (like supercomputers, virtualisation and storage systems, collections of data, software, networks, and expert support) to diverse scientific communities of North America. XSEDE has developed its identity management system (AAI) based on Globus Auth [143]; it allows integration with SAML-based identity federations, account linking, identity brokering (or credential translation), and user attributes (and entitlements) management.

The Terena Certificate Service (TCS) [75] is another credential translation service which converts user identities from IdPs that are members of federations run by National Research and Education Networks (NRENs). Those IdPs authenticate the users and have a right to obtain short-lived user certificates by publishing a particular attribute that asserts sufficient LoA.

The EMI's [48] Security Token Service (STS) [140] was designed for SOAP-based Web services; the STS issues security tokens to the Web services and the generated tokens can be further transferred to other services in the context of WS-Federation [102].

The IETF's Application Bridging for Federated Access Beyond web (ABFAB) [81] standards-based project Moonshot [89], developed by Jisc (UK NREN) [88], aims at providing federated access to non-Web browser-based services. It has been developed using the widely deployed technologies Extensible Authentication Protocol (EAP)/Remote Authentication Dial In User Service (RADIUS) with SAML bindings [80] (also used by eduroam [150]).

Agent-based approaches like InterCloud [133] and the EU-funded VISION cloud project [147] highlight the notion of dynamic federations and distributed trust. Furthermore, Generalised Architecture for Dynamic Infrastructure Services (GEYSERS) [105] uses access tokens in a SAML authentication flow to manage the delegated rights.

It is essential to have an interoperable and distributed authorisation approach to manage and enforce fine-grained access control of the disparate collaborative infrastructure services. Therefore, the focus is given to the XACML-based authorisation services, which is a widely adopted standard in the area of access control. With the release of the XACML v3.0 standard [110], a number of profiles (supporting distributed authorisation) have been published. The most relevant profiles for the collaborative infrastructures are: Administration and delegation profile [112], REST Profile [116], Hierarchical Resource Profile [115], Digital Signature Profile [118], Multiple Decision Profile [113], and Privacy Policy Profile [114]. However, only very few authorisation services provide a reference implementation of the standard and its entire suite of profiles.

Argus [7] was developed within the European Middleware Initiative (EMI) [48] project, it provides a distributed authorisation framework based on the XACML's SAML v2.0, X.509, and Virtual Organisation Management Service (VOMS) specifications. It relies on three standard components PEP, PAP, and PDP. Therefore, based on the security policies stored in its policy repository, the authorisation process takes the authorisation decision to perform a certain action on a particular service.

WSO2 Identity Server [152] is a unified authentication and authorisation service based on the Carbon platform. The service provides integration with LDAP, Active Directory Service (ADS), and JDBC to load user data for authentication. For authorisation, access control is supported through the XACML 2.0 and 3.0 specifications. Similarly, OpenAz [5] from the Apache foundation is based on XACML v3.0, however, it has been discontinued after observing lack of community support.

The AuthzForce CE [60] service is a part of the FIWARE[5] initiative. It provides an API to fetch authorisation decisions based on the stored authorisation policies and authorisation requests from PEPs. The API exposes a REST style HTTP interface, and complies with the XACML v3.0's authorisation policy format and evaluation logic, as well as for the authorisation decision request/response format.

2.11 Discussion

Given the variety of service discovery approaches and authentication and authorisation standards, federated access to the infrastructures' services has posed many challenges to the user and service providers, respectively. Therefore, despite of the availability

---

[5]https://www.fiware.org

Table 2.1. List of related technologies grouped into areas.

Areas	Related Technologies
Service Discovery	GOCDB, UDDI, BDII
Information Model	DMTF CIM
Federated Authentication	Globus Auth, Moonshot, PRACE-AAI, ELIXIR AAI, ORCID
Short-lived credentials	EMI STS, CILogon, TCS
Authorisation Service	Argus, OpenAZ, WSO2-IS

of various technologies and approaches (as described in the previous section and summarised in Table 2.11), a number of gaps remain that are identified in the following subsections.

2.11.1 Service Discovery

Being hierarchical and supporting federations, OpenLDAP-based BDII has a number of limitations. First, the information aggregation at the top-level BDII nodes from the lower level BDII nodes leads to a significant amount of delay. The process to perform such aggregation is scheduled (e.g. hourly or daily), hence some of the services discovered by the client applications might already have been removed from the infrastructure. Secondly, from a technology perspective, LDAP has an archaic design and requires sophisticated approaches to perform trivial operations (such as deployment, maintenance and CRUD operations [95]).

GOCDB is a central database with a graphical Web user interface. It only supports a single level of federation and information management, which requires manual interaction with the service. Therefore, the infrastructure services cannot register themselves and advertise their status programmatically. Furthermore, the 'centralised' GOCDB registry can become overwhelmed by the ever-growing number of services (and the corresponding information) in a large-scale collaborative and research infrastructures.

The UNICORE, AARC, and gLite registries contain all the middleware-specific services endpoint information. Being centralised in nature, they pose a risk of being a single point of failure. Indexing and discovery of services are performed in a middleware-specific fashion, therefore the registries cannot handle discovery requests from clients developed for other technologies. This hinders service usage and adoption in large scale heterogeneous compute and data infrastructures (e.g. EGI, EUDAT, or EOSC-Hub). Both DUDE and MWSDI made an attempt to federate autonomous registries, however, their scalability is limited [13].

2.11.2 Authentication and Authorisation Infrastructure (AAI)

The AAI of the life-science community ELIXIR substantially lacks support for non-Web browser-based federated access such as authenticating with Public Key Infrastructure (PKI)-based X.509 or LDAP credentials to access HPC or HTC services. For any PKI-based access, ELIXIR users have to go through a process of requesting from their organisation's certificate registration authority to issue them an X.509 certificate. Moreover, the users are expected to keep the PKI credentials secure and safe at their end.

The Partnership for Advanced Computing in Europe (PRACE) [120] infrastructure, which is one of the largest HPC infrastructures in Europe, has a PKI-driven AAI. Therefore, only services which can authenticate using the X.509 certificates can be offered. Users with federated identities based on SAML or OIDC requiring access to PRACE HPC resources have to acquire and manage X.509 credentials from their national registration authorities.

The XSEDE-based Globus Auth relies on CILogon [17, 107] is also one of the recent solutions, which (contrary to ELIXIR AAI) provides non-Web browser-based service access to generate X.509 credentials with SAML or OIDC-based federated user identities (through credential translation). With the support of federated authentication, CILogon operates at a national-scale without requiring the large number of certificate registration authorities to perform manual user vetting and identification. Due to the regional policies, the CILogon cannot be deployed within the EU research and e-infrastructures. It requires accreditation from the European Policy Management Authority (EUGridPMA) [54]. CILogon, though, can be deployed internally within an infrastructure for the services or resources that require significantly low LoA.

The short-lived credential services like TCS and STS are not suitable for Web browser-based services as they do not support federated identities. However, they are capable of generating short-lived certificates (for the authenticated users) as they interface with a Certification Authority (CA) (based on the open-source EJBCA [121]).

Shibboleth [27] and SimpleSAMLPhp [139] are SAML-based federated identity management systems mostly used for Web browser-based (interactive) services. Despite supporting standards-based SSO and attribute mapping mechanism, the given identity management systems have a number of limitations when accessing services in a federated environment. This includes the dependence on a single authentication protocol while excluding the services that may rely on other authentication protocols (OIDC, X.509, SSH, LDAP). Also, non-Web browser-based (or non-interactive) service access cannot be achieved. The integration with external attribute providers is also not available with the standard distributions, which is significantly essential to generate adequate LoAs.

While the Moonshot [89] project supporting non-Web browser-based federated access uses the RADIUS protocol, it comes with following shortcomings. First, it requires the infrastructure end-users to install (or deploy) custom client libraries. Second, the identity providers (eduGAIN or infrastructure-based) have to enable the SAML Enhanced Client Proxy profile (ECP), which is uncommon in identity federations (like eduGAIN), and not many identity providers support it. Given the support for only SAML identities, Moonshot does not support end-users with OIDC identities or X.509 certificates, which is one of the core requirements of an e-infrastructure AAI.

ORCID and other publicly available services (e.g. Google, Facebook, GitHub) support OIDC-based IdPs and SPs. Like SAML-based AAI, the systems mentioned above cannot be integrated with services relying on other authentication technologies unless credential translation is implemented. Furthermore, the lightweight identity verification (and vetting) process while issuing the user identities leads to assignment of lower LoA. Therefore, only a very limited number of services (or even none in some cases) or respectively their functions can be accessed with the issued identities. As a matter of fact, ORCID supports SAML-based eduGAIN identity federations (including

Google and Facebook), hence the community or research infrastructure-specific IdPs (and their users) cannot be added as external or upstream authenticating parties. In case of supporting different types of downstream service providers (representing the research and e-infrastructures), ORCID is limited to the OIDC specification, hence the SPs should implement OIDC client-side workflows (e.g., code grant flow) to use ORCID as an authenticating party. In addition to that, ORCID is a commercial service. Thus, despite full integration of the OIDC services, including the release of user claims containing the essential user attributes, paid membership of the service providers is required (which is usually not preferred by the public research organisations).

Several XACML-based systems can (as described in Sect. 2.10) provide authorisation or access control to collaborative or research infrastructures. However, the following gaps can be identified based on the essential requirements collected from the infrastructures: graphical user interface to manage policies by the administrators, delegated access control management (by the service providers), support for hierarchies, audit-ability and above all is the decentralisation of XACML components.

The Argus authorisation service, despite being standards-based, does not offer a graphical user interface to manage policies. Due to the centralised nature of Argus, its core PDP component is susceptible to bottlenecks, and therefore, cannot handle large number of policy decision requests.

The identity server's authorisation component of WSO2 does not offer distributed management and hierarchies of access control policies as required by collaborative infrastructures (e.g. EUDAT or EGI). In addition to that, the WSO2 identity server cannot be decoupled with its set of FIM and authentication components, which is likely daunting for the infrastructures which have an established authentication infrastructure.

The AuthzForce authorisation service from FIWARE does only offer an API, but not any graphical user interface to manage access control policies. It lacks support for administrative delegation of policies, which is one of the essential requirements of services with multiple operators.

3 Federated Service Authentication, Authorisation, and Discovery Framework

This chapter provides a coherent picture of the main contributions that are provided by the individual publications that are collected in this thesis. The chapter begins with the case studies, then covers significant requirements, followed by their realisation into the key functional components, each implementing a group of related functions to support the considered use cases.

3.1 Thesis Case Studies

This section describes the case studies used in the thesis. They haven been chosen to cover on the one hand a compute and on the other hand a data e-infrastructure and in addition a research infrastructure that relies on the e-infrastructures for its compute and data requirements. In addition, a summary of the properties of the case studies that are relevant from a requirements point of view can also be found in Section 3.2, Table 3.1.

3.1.1 CLARIN European Research Infrastructure

CLARIN [30] is one of the largest European research infrastructures with a focus on humanities and social sciences. The infrastructure provides access to language resources in various forms, e.g. text, images, audio, or multimodal. It also offers services [31] and tools to perform several functions such as analyse, aggregate, annotate, or exploit language data. CLARIN builds a federation of data repositories, service and language centres while aiming to support SSO across all partners and stakeholders of the infrastructure. In the context of this thesis, the CLARIN community are the end-users of this research infrastructure. This requires compute and data services that are shared and deployed on multiple e-infrastructures (see *Paper IV and V*).

3.1.2 EUDAT: A European Collaborative Data Infrastructure

EUDAT [53] is a European collaborative data e-infrastructure, which enables and provides data management and federation of research data across Europe. The services being offered to the researchers enable depositing, replicating, and archiving of data and they are geographically distributed (offered by collaborating partners). Along with the services specialised for data, EUDAT also operates services which support the infrastructure itself (e.g., code versioning, ticket monitoring, and resource allocation). In the context of this thesis, EUDAT is a federated data e-infrastructure, and its services are

required to be shared across multiple research communities[6] (e.g., CLARIN, ELIXIR, EPOS, SeaDATA, Integrated Carbon Observation System (ICOS), or Long Term Ecological Research Network (LTER)). Service discovery and AAI-specific requirements have been gathered and services have been implemented to enable the sharing of services (see *Paper II, III, IV, and V*).

3.1.3 EGI: European Grid Infrastructure

The EGI [42] is one of the largest multidisciplinary *federated* cloud e-infrastructures in Europe, and it caters to a vast number of user communities from multiple scientific domains. The infrastructure enables execution of complex computing workflows while collaborating with multiple resource providers and computing centers (which is transparent to the infrastructure users). The EGI's authentication and authorisation infrastructure used to be based on PKI for many years, however it is being upgraded to an advanced authentication service known as *Check-In* [43]. The discovery of the EGI services is based on a centralised approach and is supported by the BDII [18] and GOCDB [96] registries. In the context of the thesis, EGI is an e-infrastructure which supports scientific communities by providing cloud and compute services. Mainly, the service discovery requirements have been extracted from EGI, and a robust and federated service registry has been designed and implemented (see *Paper I and V*).

3.2 Requirements Analysis

This section covers requirements with an emphasis on service discovery and AAI in the research and e-infrastructures. The requirements were acquired through personal interviews, observing user interactions, scientific user applications characteristics [9, 128], and were also guided by the shortcomings in the existing technologies (discussed in Sect. 2.11). This section also discusses how these infrastructures can share users and services such as scientific workflows. The requirement elicitation and analysis presented in this section implements thesis objective *TO1*.

As mentioned in Sect. 1.2, service discovery and AAI functions play an important role in accessing the resources/services. In case of AAI, SSO enables researchers to use their federated identity to access the various remote infrastructure services, while the service providers do not need to maintain separate user accounts and passwords (or credentials). Since scientific research requires international collaboration, it has become indispensable to connect national identity federations, suggesting the federations to use the same set and schema of attributes and LoAs.

eduGAIN [74] provides a framework for interconnecting the national federations. However, the attributes among national federations have not yet been harmonised; this is an on-going work within the Research and Education Identity Federations (REFEDS) group.

Another challenge from the service provider perspective is to support federated identity to access non-Web services. As for the service discovery in the distributed cloud and data infrastructures, the services can support different sets of capabilities,

---

[6]https://eudat.eu/use-cases

Table 3.1. Infrastructures need ways to give access to users and to link services within the infrastructure, e.g. through Command Line Interface (CLI) or Web Services (WS). Some are the infrastructure's own, others are shared or from an external federation.

	Service	**CLARIN**	**EUDAT**	**EGI**
User	Authentication	federated/own	federated/own	X.509/federated/own
	Access methods (Web/CLI/WS)	Web	Web/CLI	CLI/WS
	Authorisation	own	own	VOMS
	Service discovery	Portal/Switchboard	Portal	Wiki
	Workflow	WebLicht	N/A	N/A
Infra	Authentication	IGTF	IGTF	IGTF
	Service discovery	Portal/Switchboard	N/A	BDII
	Service registry	Switchboard	GOCDB	GOCDB

remote interfaces, and end-users; most frequently, the application clients have to support multiple service registries. We shall see how adherence to the standards and distributed architecture helps in scalable service discovery and inter-operation of infrastructures.

Some of the properties of the three infrastructures that have an influence on the AAI and service discovery requirements are shown in see Table 3.1. Despite being distinct in service offerings, some of the requirements of these infrastructures overlap each other (e.g. SSO is a common requirement). Since the requirements to enable inter- and cross-infrastructure federated service access come from heterogeneous infrastructures (involving various user groups and service providers), they are in the following aggregated in an infrastructure-agnostic manner.

The requirements for federated service discovery and AAI that were identified as as part of this thesis (thesis objective *TO1*) are as follows:

R1 *Service discovery:* The collaborative and research infrastructures comprise many distributed heterogeneous services and resources, such as compute, storage, and network resources, and their providers, services, authentication services, etc. Thus, it is essential to know the offered capabilities, types, and other specific characteristics (e.g., data transfer rate or storage capacity) of services. The infrastructure's monitoring systems or service registries should enable users to discover the services based on the service properties. An example of such a registry is the Language Resource Switchboard [155]. Essential in this thesis is the perspective that a service registry plays an important role in enabling the data and compute services including sophisticated scientific workflows (provisioned by workflow engines). Given that, service discovery is an essential requirement of such workflows. The service discovery should also adequately enable service providers to publish or advertise their services and end-user or clients to discover those services seamlessly.

R2 *Common service information model:* As multiple middlewares and heterogeneous services are offered on the compute and data infrastructures, it is crucial to define service features and characteristics using a standardised information model. The model also enables interoperability and makes integration easier across infrastructures.

R3 *Unified service registration and query protocol:* This requirement is related to R2 and inspired by GRRP, whereby the information system should avoid a customised approach and rather adopt a middleware-agnostic unified interface (e.g. request and response messages) to publish and query the service information.

R4 *Service lifecycle management:* It has been observed that stale (or outdated) service information significantly degrades the overall performance of the compute and data services. Therefore, the service registry should provide a mechanism to keep the service information up-to-date. For example, Field et al. [61] provide a quality metric to measure 'freshness' of the indexed information.

R5 *Support for federations using registry hierarchies:* An e-infrastructure (e.g., EGI or EUDAT) collaborates with multiple organisations (called "National Grid Initiatives (NGIs)" within EGI). Each of these organisations offers its own set of hardware and software resources. Instead of managing the organisational resources at the central level, it is required for organisations or partners to manage the service information at their level (which forms a federation). This would make the discovery process not only robust but also traceable.

R6 *Scalability:* The registry or information system should be able to cope with the discovery of innumerable services on a large-scale compute and data infrastructure. The number of services or groups of SPs (or NGIs) can grow dramatically. Therefore, the registry should be capable of distributing the service records in a robust fashion (resilient to failures in case of bottlenecks).

R7 *Single Sign-On (SSO)*: The services within research and e-infrastructures are deployed across multiple organisations and in a distributed fashion. Yet, a single identity should be used to access all the infrastructure services. It also implies that the users should authenticate with either their institute or social identity and credentials to access all the infrastructure services. The underlying AAI technology should enable user attribute management and the services should not maintain user credentials (i.e., passwords). For example, PRACE [120] is currently based on PKI and a user's end-entity X.509 certificate is used for authentication. The services relying on X.509 certificates may require a proxy service (such as MyProxy [107] or IGTF's RCAuth [123] online CA), generating a short-lived or temporary certificate for the users or client applications.

R8 *Multiple authentication protocols:* Most of the services within the research and collaborative e-infrastructures are developed upon existing authentication libraries, and they are using several authentication mechanisms or protocols. Given that, the infrastructure's AAI either supports all the authentication protocols in all the services or endorses a proxy or intermediary [22], which should implement all the protocols and intermediates (or translate the user credentials) between the user and the target services.

R9 *Credential Delegation:* Delegation of user rights is essential for a distributed service-oriented infrastructure [23]. The access to the user's data (which is typically stored in their workspace or directory) has to be processed by the data

analysis services, that are hosted at other centres. However the services require appropriate access rights from the user to access the data in the workspaces. Therefore, the user is authenticated and authorised to the service first and then delegates their identity and permission to the service. In the case of scientific workflows, a sequence of compute or data staging tasks on multiple remote services are executed, which should be carried out on the user's behalf. This, however, requires a delegated access to the services (participating in the workflow). In addition to that, credential translation is also needed as the participating services do not support the same authentication and/or authorisation protocol. Data management services (e.g., transfer of data from one site to another) or HPC compute jobs (involving, for example, data staging-in from a remote storage service) substantially make use of delegated credentials; thereby the users or services should be able to perform tasks on behalf of the service or resource owner.

R10 *Non-Web browser-based federated access:* The infrastructure services are shared across other infrastructures, and some do not offer an interactive user access interface. For example, file transfer services (GridFTP [1], UNICORE FTP [129]), data replication services (iRODS [34]), or other third-party services with a REST API-only interface cannot be accessed in an interactive manner via a Web browser-based interface. Hence, the AAI should be able to support federated authenticated access to such non-Web browser-based infrastructure services.

R11 *Attribute Harmonisation:* Attributes released as part of user authentication by the IdPs most likely cannot be used for authorisation by the target service if the released attributes do not follow the required naming convention. In a large-scale infrastructure, multiple naming conventions (eduPerson schema or OIDC claims [127]) are used by different SPs. The AAI should be able to generate suitable (preferably on the fly) attribute mappings according to the target infrastructure services.

R12 *User Provisioning:* Before accessing any infrastructure service, users are required to be registered or on-boarded. The AAI should provide interactive (i.e. via Web application forms) and non-interactive (through an API) methods. It should also support user de-provisioning, which is far more complicated than provisioning, because de-provisioning cannot be simply achieved by only deleting the user's attributes from any central directory and removing the traces of the user's data from the respective infrastructure services.

R13 *Support for Guest Identities:* Most of the research infrastructures [10, 30] usually maintain their own identity providers (independent from the national identity federations). Social media-based identity providers (Google, GitHub, Facebook) also belong to the same category, but represent 'homeless' users. The AAI should enable identities issued by such IdPs to provide access to the collaborative and research infrastructure services (e.g., data, compute, virtualisation, etc.).

R14 *Support for groups:* One of the core requirement for an e-infrastructure is to support groups, which represent several scientific communities and projects. An

e-infrastructure should provide access to the underlying shared services to these groups (or scientific research communities). The e-infrastructure's AAI should enable community operators to manage their groups. In addition to that, some of the community-specific user attributes from the research infrastructures are required at the e-infrastructure's service, and need to be maintained within the groups to support fine grained authorisation.

R15 *Users with multiple identities:* A user can belong to multiple organisations or scientific communities with multiple affiliations. Should these communities be a part of an e-infrastructure, the AAI should be able to handle such users having multiple identities.

R16 *Different Level of Assurance (LoA):* Often, most of the users perform less sensitive operations, for example reading a data set from a data sharing service (EUDAT's B2SHARE service). For some of the users though, a high LoA is needed to perform privileged operations, for example uploading a dataset or invoking a data archival operation. A low LoA (associated with guest identities) is rather useful for the volunteer scientists (e.g., holding social identities [87]) who are only interested in, say, visualisation of data. Therefore, the AAI should support segregating the service actions into different levels by taking user attributes into account, generating and associating different LoAs according to a digital identity.

R17 *Robust Authorisation:* The infrastructure service should be able to authorise the users after the authentication process. The authorisation service, especially the PDP, should be robust and provide harmonised authorisation policies, which are consistent across the infrastructure. The AAI should essentially support policy decisions in a decentralised fashion.

R18 *Distributed Authorisation Policies Authoring:* Some services in a collaborative infrastructure are distributed in nature, thus having multiple instances of the same service deployed at partner sites (B2SAFE service in EUDAT). Due to organisational and legal boundaries, each site maintains its own set of policies, however the AAI should provide delegated management and synchronisation of policies across different sites.

The above requirements R1–R18 are realised by a unified service discovery and AAI framework for collaborative data and compute (HTC and HPC) infrastructures. *Paper I*, *Paper II*, *Paper III*, and *Paper IV* provide an in-depth description of the realisation of the given requirements.

3.3 Unified Service Discovery and Identity Management

In order to address the requirements collected in Section 3.2, the architecture and design underlying this thesis provides two main parts or building blocks to enable federated access to large scale and heterogeneous infrastructure services: The first building block

focuses on the capabilities definition, advertising, and discovery of the infrastructure services, mainly provided by service operators or providers. The second building block enables a secure access to the infrastructure services with federated identity and provides hence a modern AAI for collaborative and research infrastructures. The service also focuses on management of the authorisation (or access control) policies in a scalable manner. The unified framework combines the aforementioned building blocks in model that enables service in heterogeneous and federated environments. This achieves the thesis objective *TO2*.

The subsections below focus on the core features and architecture and design of service discovery and AAI. While these two building blocks are described separately, an important aspect is that they are in fact integrated with each other to enrich the user experience, service integration, and overall adoption of infrastructure services in the federated infrastructures.

3.3.1 Federated Service Registry

Given the requirements of service discovery in large-scale heterogeneous infrastructures, a distributed and robust service registry was conceived and designed by the thesis author within the EU-funded EMI [48]. The initial registry architecture was driven by the shortcomings posed by several Grid, HPC, data-management and cloud middlewares (UNICORE, ARC, gLite, and dCache) and infrastructures (EUDAT, EGI, and CLARIN):

- publishing and discovery of middleware-specific services,

- centralised architecture,

- restrictive service information model, and

- cross middleware service discovery (for scientific workflows).

To overcome the above list of challenges, the registry adopts a distributed approach with a flexible service information model. The remainder of this section covers the main components and features of the service registry.

The primary goal of the registry is to provide service discovery in large-scale compute and data infrastructures. The discovery process incorporates service publishing and querying in an interoperable and middleware-agnostic manner and supports federations (see Fig 3.1). The services in the registry are grouped as a *domain* (such as an NGI), which can be connected to other domains in a hierarchical fashion, thus forming a *federation* of registries. The services' information or Service Record (SR) in a domain (which is a part of the hierarchy) is indexed in the Domain Service Registry (DSR) node and they are connected with other DSRs nodes in a hierarchical fashion. At the top-level domain, there is always a Global Service Registry (GSR) which aggregates the service information from all the DSRs. In addition to that, a GSR is replicated with other GSRs and that essentially makes the service discovery process robust and resilient to single point of failure. The SPs publish SRs describing the service capabilities using the OGF GLUE 2.0 [130] standard. Following are the core features of the federated service registry.
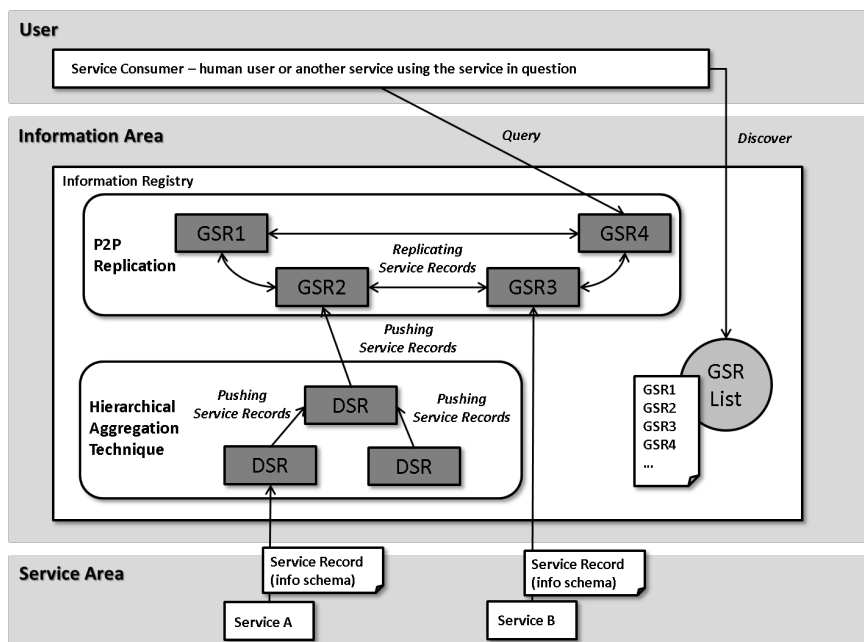
Figure 3.1. Federated service registry: Service discovery in a heterogeneous federated infrastructure

F1 *Standards-based Service Information Model*: It is not uncommon that the large-scale data and compute e-infrastructures (EUDAT, EGI, or CLARIN) offer a wide spectrum of services with the capabilities categorised into HPC, HTC, cloud, Grid, AAI, data management, code versioning, wiki, ticketing system, etc. Some services are more transient than others and dynamic in nature. In order to capture the service characteristics, the registry adopts the standardised Grid Laboratory Uniform Environment (GLUE) [40] information model and adheres to a standard information model for the infrastructure service's metadata and state. Given that, this feature addresses requirement R2. Table 3.2 shows a subset of the mandatory attributes that represent an e-infrastructure service.

F2 *Unified API*: One of the primary goals of the registry is to provide service discovery in a technology-agnostic manner. Therefore, any type of infrastructure service, be it data, compute, or other domain-specific service, can be published using the registry's unified REST-based API to perform query and advertisement in an interoperable manner. By incorporating a unified API, this feature covers requirement R3.

F3 *Hierarchical aggregation*: The architecture enables creating registry hierarchies of DSR nodes with a GSR at the top level. In order to capture dynamic information of the service (hence keeping the information freshness and quality high), the registry implements an on-demand event and push-based aggregation method.

Table 3.2. Core set of service attributes in a Service Record (SR) [98]

Attribute name	Description
Service ID	A globally unique identifier for the service
Name	Human-readable name
Endpoint URL	Location to access the service
Capability	An array of offered capabilities
Service technology	The technology used to implement the service
Service time-to-live (TTL)	The visibility of the service within an infrastructure
Service type	Service type according to namespace-based classification
Service version	Specific service version
Service health	Monitoring information about service state

The given method is inspired by the publish-subscribe messaging pattern [79], the dissimilarity is however that messages containing the service information (registration and modification requests) are published to the parent DSR instead of to a publish-subscribe topic.

A bottom-up aggregation of Service Records (SRs) is depicted in Fig. 3.2, whereby the SPs publish their services' information at the lowest DSR node of their domain or organisation; the DSR pushes the SRs to other *trusted* DSRs in the higher levels of the hierarchy until the SRs gets published onto the top GSR, which makes it *eventually consistent* [148]. The traversal time for the SR to reach the top (partly) depends on the network latency between the registry nodes. The publication of the record in GSR is followed by the replication process (see below feature F5 on the P2P-based replication of service information). By supporting hierarchies with information aggregation (using DSR and GSR), the design supports federations, which addresses requirements R5 and R6. The following two sub-characteristics of the information aggregation are supported by the registry nodes:

- The registries are geographically distributed across different administrative domains (taking into account that failures may always occur). In order to address the intermediary (availability or network) failure of nodes, an in-memory database (dotted database icon in Fig. 3.2) is used that records the (un-synchronised/pushed) modifications.

- In a research or e-infrastructure, the registry can capture a variety of services (independent or deployed at different middlewares) and may contain project- or Virtual Organisation (VO)-specific information (in-addition to what has been mentioned in Table 3.2).

F4 *Service information life-cycle management*: The registry aims at keeping the information up-to-date by making use of TTL signals. However, the service providers advertising their services are expected to refresh the information in predefined time intervals. The services which cannot be updated within the
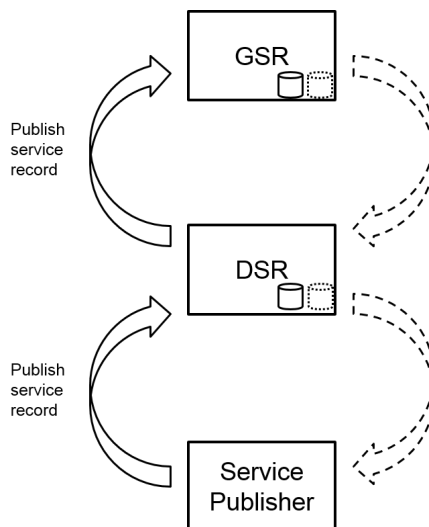
Figure 3.2. Hierarchical aggregation of Service Records (SRs)

given time interval would be removed from the registry. Moreover, life-cycle management of information ensures that the registry have fresh information and prevents the application clients or end-users from discovering stale service endpoints. This feature targets requirement R4.

F5 *P2P-based replication of service information*: A Peer-to-Peer approach to the replication of GSR: the registry supports federations by adopting the Peer-to-Peer *Pastry* algorithm [126] to replicate the SRs in the top level GSR registry nodes of the hierarchy. The approach is inspired by ISIS [106] – the ARC's [6] P2P-based information system. A hybrid approach has been adopted to replicate the information by making the keys redundant across the peer GSR nodes in the network. The given approach combines structured as well as non-structured [64] overlay networks, however depending on the network latency, the information being replicated is available to all the nodes after a certain period of time (eventually consistent [148]). After the replication of information, the services can be queried from any of the nodes in the P2P network. The adopted P2P-based approach results in a robust management of information and a fault-tolerant network of registries. *Sparsity*, i.e. the number of neighbours each peer connects to (see Fig. 3.3), plays a key role in the P2P network performance. Hence choosing an optimal sparsity value significantly impacts the performance and depends on the infrastructure requirements. The P2P approach shows the viability of decentralised information management and discovery of the infrastructure services. Given that, this feature fulfils the core requirement of scalability and failure resiliency R6.

F6 *Authentication and Authorisation*: As mentioned in feature F2, the registry design provides a REST API to allow SPs and end-user applications to publish and query the SRs in a technology-agnostic manner. Additionally, it also provides
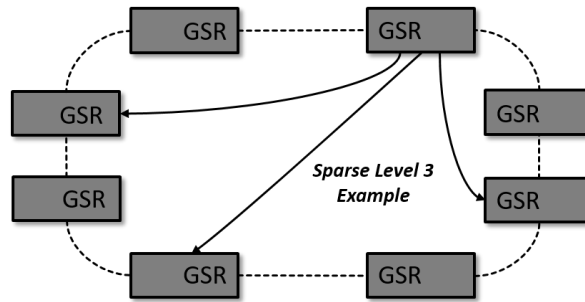
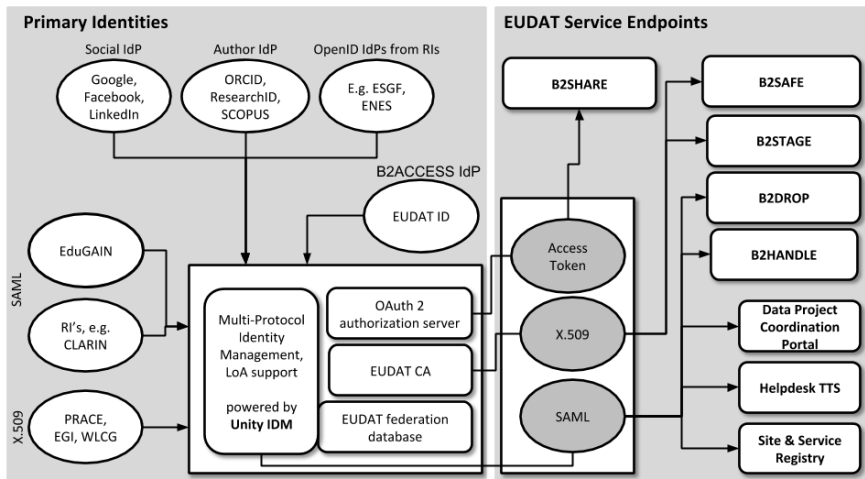*Figure 3.3. A P2P network of registries with replication of service records*

Figure 3.4. Federated user authentication and management components and the target infrastructure services

an API to enable the nodes to connect with other nodes to form a hierarchical or a P2P network. Most of the given functions require an authenticated and authorised access. A higher LoA is also needed to mitigate the risk of maliciously publishing or modifying new or existing services respectively. This feature addresses requirement R1 and is dependent on AAI, which is discussed in the next section.

3.3.2 Authentication Authorisation Infrastructure (AAI) for Collaborative and Research Infrastructures

Based on the collected requirements (see Sect. 3.2), an AAI service has been designed, consisting of an IdM system, an online CA and a distributed authorisation service based on the standardised XACML architecture. Figure 3.4 depicts the federated identity management service: the left-hand side consists of a set of IdPs that manages user identities

(username and password) and attributes. The different types of IdPs (SAML, X.509 certificates, and OpenID Connect) are connected with the AAI (shown in the centre) as external authenticating parties. Domain-specific attributes (e.g., 'eduPersonScoped-Affiliation', 'isMemberOf', or 'dariahRole' [38]) are attached to the user identities once the user is registered. The right-hand side of Fig. 3.4 shows the infrastructure SPs, which are based on multiple authentication mechanisms whereby the AAI being a proxy connects them with the IdPs by generating a suitable credential and passing it on to the target service. The credential is based on SAML, OAuth2, or short-lived X.509 certificates. The target credential is generated at runtime by translating the original credential. The generated credential also contains the enriched and unified attribute set (IdP and infrastructure specific), for example, group membership, community membership, and LoAs. The SPs performs user authorisation based on the attribute set, known as Attribute-Based Access Control (ABAC).

One of the primary goals of the AAI is to provide Authentication as a Service (AaaS) or authentication service to the infrastructures. The main components of the identity management service are authenticators, endpoints, translation profiles and groups (and their attributes). Authenticators, as the name suggests, perform full authentication of the credential (provided by the user being authenticated). The endpoints are associated with the given authenticators, they provide a portfolio of access modules (or endpoint types), each of which can be deployed multiple times (e.g. several SAML IdP for different projects) to deploy multi-federations identity. The user groups and the corresponding attributes can provide a means to segregate users in multiple scientific communities or projects or departments with delegated administrative rights to the group managers (representing those communities). The user attributes can after the authentication process be mapped (regardless of the authentication protocol) according to the infrastructure requirements (e.g. SAML2INT set of attributes to OIDC user claims) which is realised by the input and output translation profiles.

The AAI service developed as part of this thesis offers the following set of essential features to enable federated authenticated access to infrastructure services:

F7 *Multiple authentication providers*: This feature addresses the requirement R8 by supporting several authentication protocols for the authenticating parties (or IdPs) and for the relying parties (or SPs), such as eduGAIN [74] (based on SAML), social IdPs (using OIDC, such as Google, Facebook, or ORCID), community-operated IdPs, or even IGTF X.509 certificates (see Fig. 3.5). In order to support SAML-based IdP and SP entities, the feature allows configuration of metadata (in SAML-specific terms) of the given parties. In case of different authentication protocols between authenticating and relying parties, the AAI service does (as a proxy or intermediary) authenticate the user or principal with the credential (it already has) and generates a suitable credential or security token during the authentication flow.

F8 *Multi-Factor Authentication (MFA)*: Depending on the sensitivity of the resource, the AAI supports users to authenticate using two or more types of credentials. It has been implemented by flexibly combining multiple authenticators, for example, a user can be asked to access a sensitive data repository by typing a code received through Short Message Service (SMS) and then authenticate with their
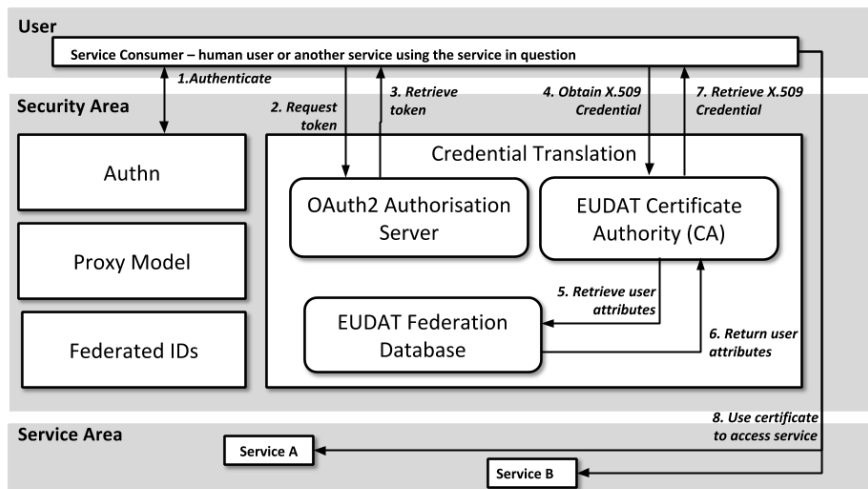
Figure 3.5. Credential translation from user identity to X.509 certificate for non-Web browser-based access

organisation IdP. MFA can also be used as one of the instruments to elevate the identity's LoA (see feature F9) and is therefore needed to fulfil requirement R16.

F9 *Assurance profiles*: Not all the users have the same privileges on all the services, so they could be administrator in one service while being a normal user in the other. The AAI service assigns adequate LoA to the identity, which can be derived from a type of IdP the user is authenticating with and from other credentials (e.g. using MFA). Likewise, the community-specific and organisational or project-specific attributes help in allocating the adequate LoA. The given attributes are processed using the concept of translation profiles (generally known as attribute filters). Since the attribute filters are customisable, standard assurance frameworks such as RAF [124] and IGTF Assurance Profiles [73] can be supported; the service of this particular feature addresses the requirement R16.

F10 *Type of Service providers*: This feature refers to the inverse scenario of feature F7 and aims at requirement R8. The AAI is designed to integrate with the services exposing multiple authentication interfaces (such as SAML, OpenID Connect/OAuth 2.0, X.509, or application-specific credentials).

F11 *Non-Web browser-based federated access*: Non-Web browser-based access is crucial to execute in compute and data environments tasks on the user's behalf. The credentials used in the authentication flow can be short-lived X.509 certificates or JSON Web Token (JWT) [90]. Since short-lived X.509 proxy credentials can be generated either through the online certificate authority or IGTF's accredited RCAuth [123] based on a user's identity, the AAI provides access to non-Web browser-based services (such as GridFTP [1] or iRODS [34]). The given certificate authorities are integrated as OIDC client applications and use access tokens to generate the delegated credentials. OAuth-based access can also be enabled if

35

the target services support the appropriate grant flow (see Figure 3.5) or make use of OAuth refresh tokens. This feature implements requirements R9 and R10.

F12 *User and service (de-)registration*: Users or SPs can register themselves with the AAI using a dedicated Web user interface (with registration forms). This feature of manual user and service registration covers requirement R12.

F13 *User enrolment*: Users can be enrolled with the AAI using: email invitation (initiated by the administrators), based on approval, or filling in a registration form (see the previous feature F12). The AAI also enrols users in an automated fashion by registering them through its REST API. This feature is required when an already established scientific community or research infrastructure joins the e-infrastructure (e.g. EUDAT). This feature implements requirement R12.

F14 *User attribute management*: In order to define attribute release and consume policies, the AAI is designed to offer a rich user interface to create such policies (or translation profiles) in a flexible and declarative manner using the MVFlex Expression Language (MVEL) [101]. Moreover, the infrastructure-specific attribute mappings can also be defined through the given policies. The attribute mapping functionality of the AAI service addresses the requirement R11.

F15 *User account linking*: After the user registration process, while authenticating with one IdP, the AAI allows the user to register with another identity. In this way, multiple identities of the same user can be linked while keeping the corresponding set of attributes. Requirement R15 is covered by this feature.

F16 *Group Management*: The AAI service is designed to support categorising users into multiple groups or group hierarchies representing different scientific communities or projects. In case of large-scale e-infrastructures, this feature is essential when multiple scientific communities are connected with the AAI. This feature enables this kind of delegated group management (including the hierarchies) to manage multiple communities with different policies. The feature implements requirement R14.

F17 *Third-party attribute providers*: Users coming through community portals often possess community-specific attributes, which are managed in specific services called attribute providers. The AAI supports fetching the attributes from external service providers by querying the service provider endpoints using the appropriate query syntax (LDAP or SAML query). The fetched attributes are consequently processed (or harmonised) and merged with the already stored user attribute bundle inside the user database. This feature refers to requirement R14.

F18 *XACML-based access control*: Figure 3.6 depicts an architecture of the authorisation subsystem, which is based on the XACML [110] authorisation standard. The hierarchical architecture allows for harmonised management of the XACML policies in the central service PAP and Policy Repository (PR) (shown at the top of Figure 3.6). Since multiple instances of a single service can be deployed across data centres (such as iRODS), there is a need to run a central PAP and PR combination to manage authorisation policies for the service as a whole, thereby covering all instances running at the individual data centres. The central service
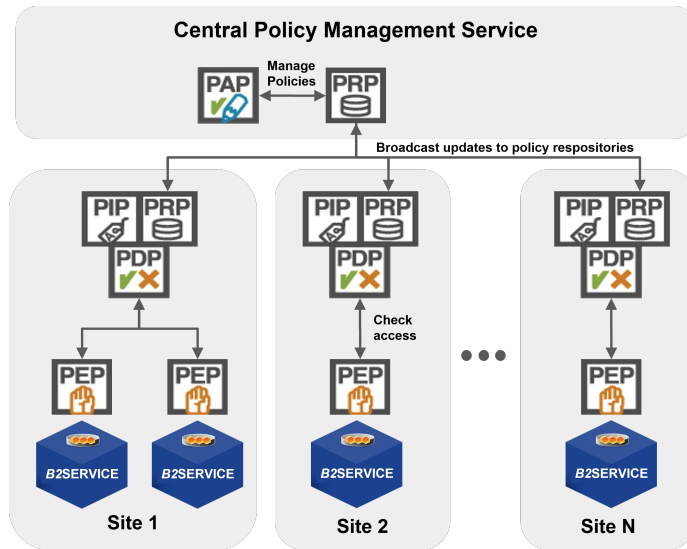
Figure 3.6. An XACML-based distributed authorisation architecture
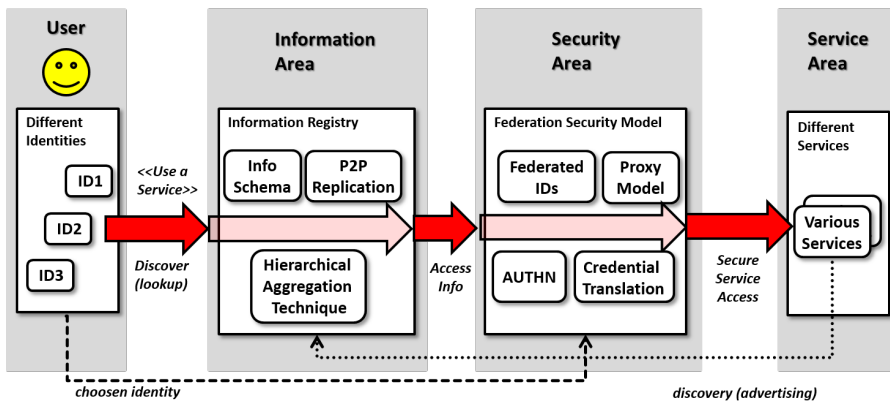
Figure 3.7. Integrative federated authentication and service discovery architecture

PR is replicated amongst data centres. The replication of policies across different sites or organisations addresses requirement R17. Each data centre operates a local PR that caches the policies from the central PR. Besides PR, the data centres are also operating one or more PDPs and PEPs to enforce attribute-based authorisation. The implementation of local caches and policy enforcement within the close proximity of the organisation addresses requirement R18.

3.3.3 Integrative Architecture

Figure 3.7 depicts an integrative architecture, which connects scientific user communities (by scientific research infrastructures) with the e-infrastructures providing data and

compute services (see also Figure 1.1 on Page 3). As mentioned in Section 1.1, the aim of this thesis is the design and implementation of building blocks that enable federated service access to large-scale cloud and data infrastructure resources, which includes robust service discovery and AAI. The combination of the both defines a federated model that can be adopted in e-infrastructures of different scales.

3.4 Implementation

This section covers the implementation of the federated EMI Service Registry (EMIR) (as specified by the features listed in Section 3.3.1) and B2ACCESS, an AAI for federated infrastructures (providing the features listed in Section 3.3.2). This achieves thesis objective *TO3*. (All implementations are open-source and the source code repository locations are provided in Section 3.6.)

3.4.1 EMIR: A federated service registry

The federated registry developed as part of this thesis has been implemented as a service called EMI Service Registry (EMIR). The registry has been implemented as client and server applications. The server application consists of EMIR nodes (DSR or GSR) and implements a REST API using the Jersey (JAX-RS) framework [41]. The RESTful Web services are exposed to the infrastructure service providers and end-user applications. The client-side application of EMIR is implemented as a configurable probe (in Python), which can be used by the service providers to advertise their services.

EMIR uses the GLUE 2.0 information model to capture different kinds of services and depends specifically on its normative XML [131] and JSON [137] formats. EMIR uses the JSON format in the registry and adopts the JSON Schema standard [151] (comparable to an XML Schema Definition), thus allowing annotation and validation of JSON documents.

In order to implement authenticated access to the EMIR Web services, TLS mutual client authentication has been incorporated. Therefore, all the registry nodes and SPs must acquire a valid X.509 certificate issued by a trusted certification authority. The service advertisement requires authorised access, thus every EMIR node maintains an Access Control List (ACL) with in-memory authorisation (implemented using the HERAS-AF XACML library [78]). Unlike conventional SQL, a schema-free or NoSQL (Not-only SQL) approach (using MongoDB [100]) has been incorporated. The database itself offers horizontal scalability to distribute a large number of SRs over multiple database instances.

3.4.2 B2ACCESS: AAI for federated infrastructures

The AAI developed as part of this thesis has been implemented as the following set of services: Unity identity management (IdM) [144], EUDAT online CA (including client libraries), monitoring probe and XACML authorisation service. The suite of services is called B2ACCESS [51][7].

---

[7]The author of this thesis was involved in the development of all of these. However, for UNITY only with respect to the requirements and initial architecture, but for B2ACCESS, online CA, and XACML-based

Unity IdM provides user, group, and attribute management as well as the possibility to integrate with external attribute providers (using both push and pull-based approaches). It can be deployed as a proxy between the identity federations (such as SAML-based eduGAIN [74] or OIDC-based Google and Facebook) and service federations offered by the e-infrastructures (e.g. EUDAT, EGI, or PRACE). To support the OIDC-based specifications, the Unity implementation is based on the Java connect2id library [32] that implements OIDC's authorisation server and client registration services. The base security library to enable communication between the authenticating (identity federations) and relying parties or entities (service federations) is EMI's Common Authentication Library (caNl) [47]. The library also incorporates TLS mutual client authentication for the X.509 certificate-based services. All the functions including SP, IdP, credential and attribute management are offered via an intuitive Web user interface (implemented using the Vaadin framework [145]).

The B2ACCESS online CA is implemented as RESTful Web service. The service generates short-lived X.509 public and private key pairs based on a valid OIDC access token. The service optionally supports embedding SAML assertions (containing user attributes) in relevant certificate's critical extensions. The certificate then can be used as an attribute certificate. The CA service has been integrated with Unity's OIDC authorisation service endpoint, however, any OIDC standard authorisation can be integrated.

The B2ACCESS monitoring probe is a client application implemented in Python to check and report Unity IdM and online CA functions to the target monitoring services. The authorisation service of B2ACCESS's suite is XACML-based and extends the open-source AT&T's XACML library [11]. It offers replication of policy authoring in a distributed fashion and also includes a Web user interface to manage access control policies by the service administrators.

3.5   Use case: Data sharing using federated service discovery and authentication with EMIR and B2ACCESS services

A real use case with a concrete sequence of actions pertaining to service discovery, user authentication, and credential translation within an infrastructure (e.g. EUDAT) is illustrated in Figure 3.8: A scientific community user in the given scenario aims to share results from her experiments with other scientists. The user discovers a B2SHARE data sharing service (that EUDAT has to offer) while sending a query to the EMIR service. Consequently, the information about all the matching and available services would then be sent to the user (or her client application) as a response to the query. After discovering the service, the user authenticates herself with her home IdP through the B2ACCESS service, which is acting as a proxy between B2SHARE and user home organisation's IdP. Subsequently (after successful authentication), the registered user uploads the data set to the B2SHARE service and shares the link with fellow scientific users. After the data sharing process (invoked by an end-user), the EUDAT infrastructure replicates the shared data using the B2SAFE and B2STAGE services

---

architecture also with respect to the implementation.
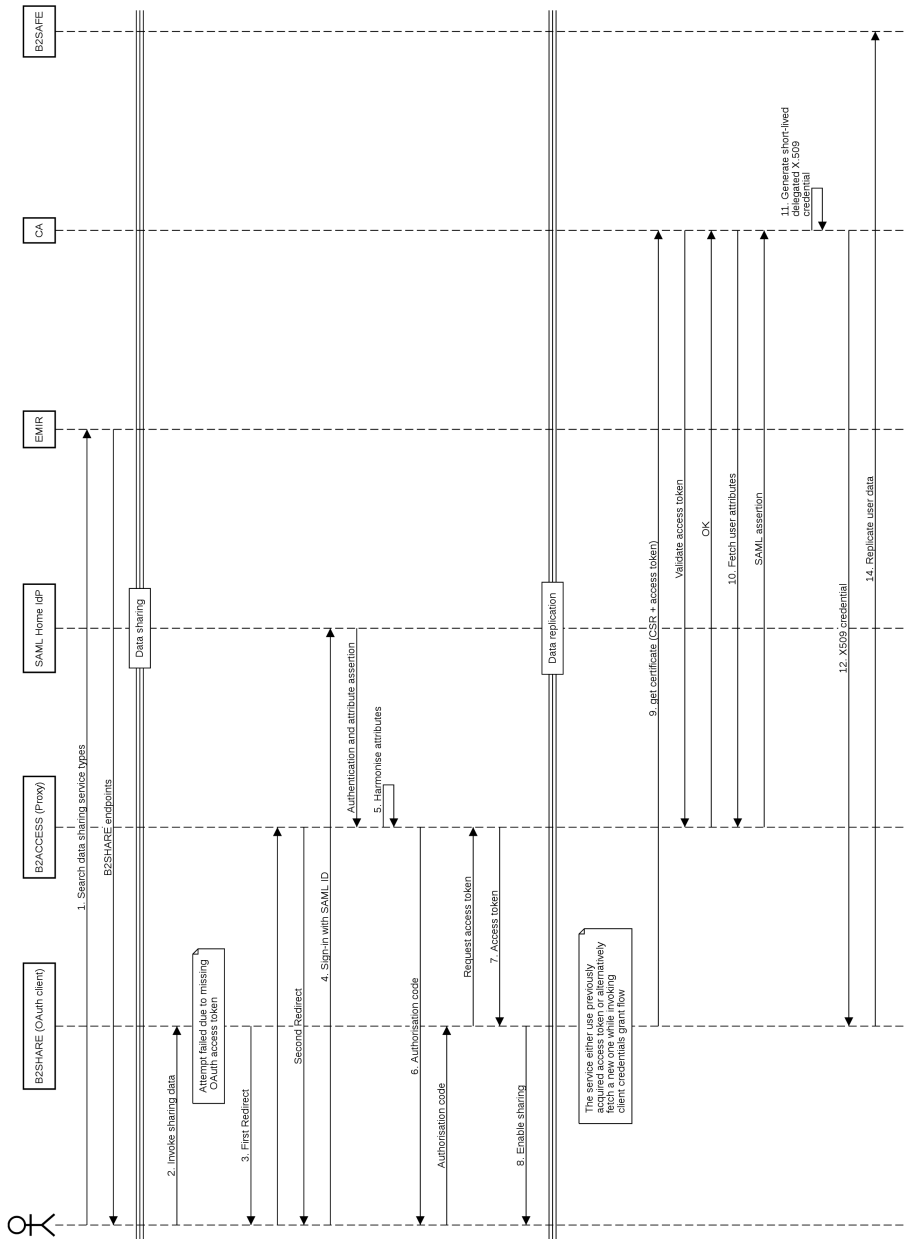
Figure 3.8. Service discovery, federated authentication, attribute harmonisation and credential translation for data sharing and replication in the EUDAT infrastructure

(based on iRODS). The replication is carried out to mitigate the risks of data loss and to enable archival for sustainable data preservation. B2SAFE and B2STAGE require non-Web browser-based authenticated access as the data is replicated with a client application. In order to enable the replication, B2ACCESS provisions a delegated

Figure 3.9. CLARIN data staging use case showing cross-infrastructure federated authentication and service discovery

X.509 credential associated with the user's identity that also contains user attributes, signed by the EUDAT online CA or IGTF accredited RCAuth [123] CA. The attributes embedded therein are used to perform resource-based authorisation. For the delegated access, OAuth token exchange [91] is being investigated for the third-party transfers (e.g., replication of data).
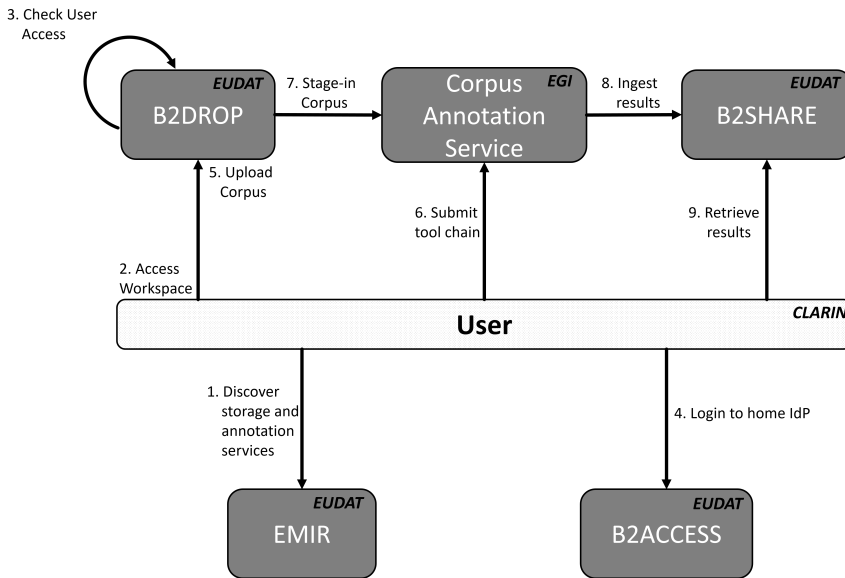
The data sharing, staging and replication mentioned above is focused on a single infrastructure (EUDAT). However, in scientific collaborations like EOSC-Hub [49], users from multiple scientific communities access the services and applications from multiple infrastructures. Such a use case is shown in Figure 3.9 where a CLARIN community user discovers and accesses the data service (storing language resources) from EUDAT, and uses compute resources from EGI, which also hosts a language application (Corpus annotation service) from CLARIN.

The evaluation of the integrative architecture and implementation using this case study achieves the thesis objective *TO4* and is described in *Paper V*. Note that Section 5.2 covers the concrete impact on other research and e-infrastructures which demonstrates as well the applicability of the developed solution.

3.6 Software Repositories

The implementations described in sections 3.4.1 and 3.4.2 are open source software and available via the source code repositories listed in the following subsections.

3.6.1 Federated Service Discovery

1. EMI Service Registry (EMIR): Federated and fault tolerant service registry and client tools. The author of this thesis has mainly contributed to: project structure, REST API, data and application layers.
   `https://github.com/eu-emi/emiregistry`

2. OGF GLUE: A standardised information model to capture compute, storage, and cloud resources. The author of this thesis is one of the editors and authors of the GLUE (2.0 and 2.1) specifications and its (JSON and XML) reference implementations.
   `https://github.com/orgs/OGF-GLUE`

3.6.2 AAI

3. B2ACCESS Unity IdM: The service offers FIM, group management, credential translation, and identity hub. The author of this thesis has contributed to the initial design and implementation of the service and has been responsible for integrating the third-party IdPs and SPs.
   `https://www.assembla.com/spaces/unity-public`

4. B2ACCESS Monitoring Probe: A client application implementing the monitoring plugin of the Argo Monitoring service to probe B2ACCESS authentication and user query functions. The author of this thesis has been involved in all the implementation aspects of the client application.
   `https://github.com/EUDAT-B2ACCESS/b2access-probe`

5. Distributed XACML Authorisation Service: An XACML-based authorisation service to manage and enforce access control policies in an efficient and scalable manner. The author of this thesis has designed, implemented, and deployed the pilot service on the EUDAT infrastructure.
   `https://github.com/EUDAT-B2ACCESS/xacmlDemonstrator`

6. EUDAT Online CA: An online Certification Authority that issues short-lived X.509 credentials. The author of this thesis has forked and revised the CA code (by integrating with Unity IdM) from the EU funded project CONTRAIL [33] on open computing infrastructures for elastic services.
   `https://www.bitbucket.org/eudataai/contrail-ca`

7. B2ACCESS OAuth 2.0 Client Library: A Java-based OAuth 2.0 client library implementing the Code Grant Flow for Web based applications (OAuth 2.0 Clients). The author of this thesis has forked and revised code from the CONTRAIL project by adding a utility to integrate the Java-based OAuth clients.
   `https://www.bitbucket.org/eudataai/contrail-oauth2`

4 Summary of Publications

This chapter summarises the publications that constitute the core of this cumulative thesis. The publications are ordered based on an end-user's perspective, i.e. the order an end-user accesses the infrastructure services: discovery of a service followed by its access.

4.1 Paper I: The EMI registry: discovering services in a federated world

This publication covers all the facets of robust service discovery in the federated infrastructures. Based on the requirements from users, service providers, and infrastructure operators, a robust and unified service registry is designed that enables discovery for federated environments. The publication starts with infrastructure use cases, leads over to building an architecture, and ends with an evaluation (including throughput and performance) of the whole service discovery framework. This publication therefore covers the service discovery component of the thesis objectives (TO1–TO4).

This journal article focuses on the requirements elicitation and usecase analysis pertinent to the service discovery in the contemporary distributed infrastructures. This activity was funded by the EU project EMI. The goal is to provide a unified experience to discover the infrastructure services that are residing in multiple administrative and geographical domains and that are usually deployed on multiple types of middlewares.

Cross-domain and middleware-agnostic service discovery has two major implications. First, services are deployed on a middleware-specific registry if the infrastructure is offering multiple types of services (e.g., HPC, cloud, data management) with proprietary interfaces (e.g., UNICORE's SOAP-based Web service registry or gLite's LDAP-based BDII). Second, the information models that the middlewares are using to capture and expose their deployed services are non-standard, hence not interoperable with other types of service registries. The concept is not even suitable for middleware-specific clients, which plays an important role in the usage of the provided services. The clients can be represented by HPC or HTC Web portals, workflow engines, or other information (and analytical) systems.

To address these challenges in large-scale infrastructures, a service registry with a unified and robust programming interface and information model has been designed and implemented, called EMIR. In order to be fault-tolerant and scalable, the registry implements a semi-structured P2P network to build a network of registry nodes for information replication and hierarchical synchronisation corresponding to the domain hierarchies. This combination of a P2P and hierarchical aggregation approach prevents a single point of failure and enables decentralised management of registries. To express and store the service metadata and its capabilities, EMIR captures the service information in a standardised GLUE 2.0 information model.

Paper contribution: The author of this thesis has mainly contributed to the paper outline, existing approaches, the design of a federated service registry, and its performance analysis, i.e. sections 2, 3, and 4 of the paper.

Supporting publication: **A.S. Memon**, I. Márton; G. Szigeti, L. Field, M. Riedel, *"EMIR: an EMI Service Registry for Federated Grid Infrastructures"*, EGI Community Forum 2012/EMI Second Technical Conference, Munich (Germany), 26 Mar 2012–30 Mar, 2012. Proceedings of Science, Sissa (2012). `http://pos.sissa.it/archive/conferences/162/073/EGICF12-EMITC2073.pdf`

4.2 Paper II: Federated Authentication and Credential Translation in the EUDAT Collaborative Data Infrastructure

A.S. Memon, J. Jensen, A. Cernivec, K. Benedyczak, M. Riedel, *"Federated Authentication and Credential Translation in the EUDAT Collaborative Data Infrastructure"*, 2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing (UCC), London, United Kingdom, 8 Dec 2014–11 Dec 2014, IEEE, pp. 726–731 (2014) [DOI: 10.1109/UCC.2014.118]

This publication puts emphasis on SSO, FIM, and user and trust management challenges in the current research and e-infrastructures. This covers the first three objectives of the thesis (TO1–TO3) with a focus on AAI, namely the requirements analysis, design, and implementation of the B2ACCESS service.

The paper highlights the challenges of Web browser- and non-Web browser-based user authentication, attribute management, and LoA in e-infrastructures, specifically in the EUDAT infrastructure (an EU funded collaborative data infrastructure). Like any other e-infrastructure, EUDAT provides data management services (data sharing, replication, repository) to several scientific user communities ranging from earth scientists (TERENO) to biologists (Instruct) and linguists (CLARIN) to name a few. It also operates internal services to manage own infrastructure (wiki, issue tracker, code versioning service, helpdesk, etc.). Every partner scientific community (i.e. the end-users)

has its own established infrastructure and mechanism to authenticate and manage user identities. For example, CLARIN is based on SAML identity whereas users of the TERENO community possess OIDC-based credentials. In addition to the aforementioned challenges of multiple types of identities, different groups of EUDAT services accept different authentication protocols. Two distinct entities, namely identity and service providers, have incompatible authentication protocols and attribute dictionaries, which leads to a requirement for a proxy or bridge supporting the conversion of credentials to enable access to the EUDAT infrastructure services.

The developed model bridges the identity providers federation and service provider federation to enable authenticated (non-)Web browser-based federated access to the services. The paper highlights a real world use case of data staging or movement from (source) research to (target) e-infrastructure with federated identities, thereby the user credentials are specifically translated from SAML identity to PKI credentials.

Paper contribution: The author of this thesis is the main author of the publication. He has introduced the proxy-based approach and its application in a data staging use-case in the EUDAT e-infrastructure. He has contributed to the outline and majority of the sections of the paper, i.e. sections 1, 2, 3 and 4.

4.3 Paper III: Combining the X.509 and the SAML Federated Identity Management

This publication covers the first two objectives of the thesis (TO1–TO2) with a focus on AAI. The emphasis is on interoperability between authentication systems (deployed on different infrastructures) while maintaining the same level of trust between the authenticating and relying parties. The approach furthermore harmonises user attributes (to support authorisation) across the involved systems.

The primary goal of the published work is to have interoperability between the service and identity providers, where the user accesses an LSDMA service accepting a credential of one type while authenticating the user with a credential of another type. The emphasis is also on translating SAML identifiers to the short-lived X.509 credentials. As a premise, the paper has identified a set of requirements. Based on them, it proposes an integrative architecture for SAML to X.509 credential translation in the LSDMA infrastructure. In addition to the IdP, SP, and online CA components, the architecture also includes VOMS (Virtual Organisation Management Service) to satisfy the requirements of attribute and/or role-based authorisation.

Paper contribution: The author of this thesis has mainly contributed to the overall structure and design of a credential translation approach within the LSDMA project and contributed to the paper sections 1, 2, 3, and 4, i.e. introduction, state of the art, design, architecture, and conclusion.

4.4 Paper IV: Implementing an authorisation architecture in the EUDAT services federation

A.S. Memon, J. Jensen, W. Elbers, M. Riedel, H. Neukirchen, M. Book, *"Implementing an Authorisation Architecture in the EUDAT Services Federation"*, IEEE Conference on Application, Information and Network Security (AINS), Miri, Sarawak, 13–14 Nov 2017, IEEE, pp. 111–117 (2017) [DOI: 10.1109/AINS.2017.8270434]

This publication covers the initial three objectives (TO1–TO3) of the thesis, the focus is on the AAI aspect of the thesis. It includes requirements analysis, architecture and design of an authorisation system targeting multi-disciplinary and multi-stakeholder infrastructures, that supports replication of access control policies in a robust fashion. The paper has derived a new generation of XACML architecture, the prototype of which has been implemented and tested for production deployment, which is a part of B2ACCESS service.

In this paper, a scalable and distributed authorisation system is introduced that is based on the OASIS XACML 3.0 standard [110]. The authorisation service is a part of the B2ACCESS framework and the result of the requirements extracted from the users and operators of B2STAGE [52] and other distributed services within EUDAT. B2STAGE is a data replication service for the EUDAT infrastructure and its core is based on the iRods [34] software stack. The deployment of the B2STAGE service consists of several distributed sites that are deployed across the partners (or service providers). In such a complex and distributed setup, it is a daunting task to maintain data access and site policies at different granularities. Moreover, a change of policy in a particular site requires broadcasting to all other sites. Hence, synchronisation and consistency have become a significant maintenance overhead for the site administrators. The standards-based authorisation service aims to address the aforementioned issues and provides a Web user interface to manage and view the access control policies.

The distributed architecture of the developed XACML-based authorisation service consists of a two layers component hierarchy: central and site level. The central level enables Create-Read-Update-Delete (CRUD) operations on policies using either an API or Web user interface. The main components are the Policy Administration Point (PAP) and Policy Retrieval Point (PRP). The site level on the contrary does not explicitly allow *create, update or delete* and contains PDP, PRP, PEP, and a Policy Information Point (PIP) to enforce service level authorisation based on the user attributes. During the implementation phase, several implementations of the XACML 3.0 standard were evaluated based on the profiles (for example, Administration and Delegation, SAML, REST, Multiple Decision Profiles), support, license, and usability. The authorisation service is based on a reference implementation of the XACML 3.0 standard by AT&T [11], which incorporates most of the required profiles.

Paper contribution: The author of this thesis is the main author of the publication. He has introduced a robust, distributed and standards-based authorisation model for the services that are deployed in a federated environment (and multiple organisational boundaries). He has contributed to the outline, background, architecture, comparative analysis and conclusion sections.

4.5   Paper V: Towards Federated Service Discovery and Identity Management in Collaborative Data and Compute Cloud Infrastructures

This publication presents a joint model (containing AAI and discovery services) by taking into account three multi-national research infrastructures, one that provides data services, one that provides compute services, and one that supports linguistics research. The main objectives are to inter-operate and jointly provide the data and compute infrastructure services to the scientific user communities. This implies building service federations (trust, service status, information systems) and identity federations (identities, authentication, and authorisation). The publication covers both AAI and service discovery aspects of the thesis, thus accomplishing the objectives TO1–TO4 of this thesis.

The two essential functions of any research or e-infrastructure are service access and discovery. It is also necessary to integrate the given functions and implement them as a unified framework. In this journal article, first a number of requirements are summarised to enable service access in a user-centric way, i.e. the user authenticates herself with her home IdP and accesses the service after successful authentication.

The home IdPs do not necessarily support the authentication protocol which the target services are able to accept. One possible way out is to implement translation of the n different user credential types to the m different credential types accepted at the target services. However, with a high number of services supporting many distinct authentication protocols, implementing a direct translation would require $n \times m$ translators. In addition to that, attribute harmonisation is indispensable as the IdPs release a set of attributes which can or cannot be consumed by the end services (for authorisation) depending on the use of non-standard naming conventions. Trust management is another challenge when connecting a service with identity federations (such as eduGAIN [74]).

In order to overcome the complexity and shortcomings both at the identity and the service provider ends, a novel and unified approach of a distributed hub-and-spoke federation model [76] has been developed and is implemented by the B2ACCESS [51] service. It is based on the concept of Identity Management as a Service (IaaS), which enables authenticated and authorised access to the e-infrastructure services. The service is built on the identity management system and distributed authorisation service Unity [144]. B2ACCESS offers the infrastructure operators an administration dashboard to manage the users, services, and attribute (release and consume) policies. The authorisation service provides management and enforcement of access control policies while adhering to the XACML standard and its profiles.

The distributed B2ACCESS architecture divides the XACML components into a two level hierarchy and also makes use of the corresponding XACML profiles [116, 113, 112]. The primary goal is to make the authorisation policy management scalable and user-friendly.

The next part of the article is focused on service discovery, which includes three main elements: service publishing by service providers, information model, and querying for services. HPC or cloud infrastructures are usually based on a single middleware, hence specific service registries are deployed to fulfil the service discovery. However, the service federations usually consists of multiple types of services (HPC, cloud or data), thus requiring multiple service registries. The approaches used to date are based on centralised architectures and thus prone to having single points of failure. In order to support service discovery in such heterogeneous service federations, a common and standards-based service registry EMIR has been developed and implemented as part of the EMI [48] project. The implementation of EMIR is flexible to support registry federations, thus employing hierarchical as well as P2P approaches. That said, the registry persists service information in a robust fashion. The registry hierarchies combine the information from multiple registry nodes called DSR, while the P2P network is used for replicating the information across multiple EMIR nodes. The replication can be done at different hierarchical levels, but mostly at the top level of the EMIR network. The service information stored at the nodes is based on the standardised OGF GLUE 2.0 specification.

The final part of the article describes the derivation and implementation of an integrative model, which enables a unified user experience by integrating service access and discovery together. The joint model is illustrated by a use case scenario of combining three infrastructures: EUDAT, CLARIN, and EGI, offering compute (EGI), data (EUDAT), and a research community infrastructure (CLARIN). The use case covers how a CLARIN user discovers a service, authenticates herself from B2ACCESS, runs

compute job on EGI which then does the staging-in/out of data using the EUDAT B2SHARE service.

Paper contribution: The author of this thesis is the main author of the publication. In the publication, the author has defined a joint approach to enable federated service discovery and secure access to e-infrastructure services for the scientific communities. He has contributed to the majority of the article sections.

4.6   Relation of Publications and Software to Thesis Objectives

As this cumulative thesis is based on publications, the different thesis objectives (TOs) defined in Section 1.2 are covered by different papers. Table 4.1 shows a mapping between publications and TOs.

A result of the research described in this thesis is an integrative architecture addressing service discovery and secure access and corresponding implementations. The implementations EMI Service Registry (EMIR) and B2ACCESS are available as open-source software. The EMIR service registry had been the main discovery service for the EMI compute, data and internal infrastructure services. While the EMI project has finished, EMIR has then been used by the ARC [6] infrastructure. B2ACCESS is being used as a production Authentication and Authorisation Infrastructure (AAI) service in EUDAT[8] (and hence in EOSC-Hub[9]) and various other communities. The developed AAI service is also one of the earliest adopters of the *Hub-and-Spoke Federation with Distributed Login* approach. In addition, the results of the thesis have been integrated into open standards such as GLUE (see Section 2.3). Table 4.2 shows how the created software artefacts are covered by the scientific publications.

Table 4.1. Association matrix of thesis objectives and scientific publications.

Paper / TO	Paper I	Paper II	Paper III	Paper IV	Paper V
TO1	X	X	X	X	
TO2	X	X	X	X	X
TO3	X	X		X	X
TO4	X				X

Table 4.2. Association matrix of software contributions and scientific publications.

Paper / Software	Paper I	Paper II	Paper III	Paper IV	Paper V
Federated Service Registry (Sect. 3.6.1)	X				X
AAI (Sect. 3.6.2)		X	X	X	X

---

[8] https://www.eudat.eu
[9] https://www.eosc-hub.eu

5 Conclusions

This chapter provides a brief summary of the thesis (Section 5.1), describes the impact on production research and e-infrastructures (Section 5.2), and provides an outlook on future work (Section 5.3).

5.1 Summary

This thesis covers the service discovery and AAI models in three European research and e-infrastructures: EUDAT, CLARIN, and EGI. From these infrastructures, concrete requirements have been extracted (thesis objective TO1), which did lead to the design (thesis objective TO2) and implementation (thesis objective TO3) of a robust service registry and AAI services. The developed service registry provides standard interfaces to query and advertise service information that enable discovery of the infrastructure services deployed in a federated (yet secure) environment. The developed services have been evaluated (thesis objective TO4) in and are used in production environments of the above mentioned infrastructures. They serve up to tens of thousands of requests per second. In such an intense environment, low-latency, scalability, secure service (de)provisioning, and accuracy of information are the key factors to surmount the operational challenges of the digital infrastructures.

Authenticated and authorised access to infrastructure services can be easier if all services ideally support the same authentication protocol, attribute schema, and policies. If services require distinct authentication protocols, the users need to keep and secure multiple types of credentials. Furthermore, the attribute providers, as well as Identity Providers (IdPs), use different naming schemes for user attributes which significantly hinders the authorisation (regardless of the granularity). Trust management between IdPs and Service Providers (SPs) is a corner-stone of identity federations, particularly in the widely adopted *Mesh* federation model, in which every IdP can authenticate to every SP. This makes trust management substantially more complex. An infrastructure wishing to register its services with a national federation would have to register all services individually. Finally, the centralised nature of management and enforcement of service access control policies does not scale well and cause bottlenecks.

The research covered in this thesis is contributing solutions that enable federated access to collaborative compute and data infrastructures in a robust manner. In particular, offering scalable and unified federated service discovery and addressing authentication and authorisation management problems. The concepts for solving these problems are:

Federated identity management provides users with a way to use a single credential within and across multiple and heterogeneous infrastructures. This implies

integration of Web and non-Web browser-based services which means that different authentication and authorisation schemes need to be supported. This has been achieved through the use of credential translation, attribute harmonisation, identity linking, and a proxy model within AAI.

Hub-and-spoke (proxy) model unifying identity management, trust management, authentication, authorisation, credential translation, and accounting which makes it possible for users to access and store data across infrastructures and furthermore enables user communities and service providers to build more sophisticated services with lower trust management barriers. The implementation of this research outcome, the B2ACCESS service, is one of the early adopters of the distributed hub-and-spoke federation model. The successful adoption of the model showcases the viability in other research and collaborative infrastructures.

Decentralised authorisation service is a scalable service based on XACML to manage access control policies (while incorporating several of the XACML profiles). The service defines and implements a replication architecture to synchronise and manage the access control policies of the services deployed in different geographical and administrative domains. It also defines a low-latency authorisation model by bringing access control closer to the distributed services deployed on a research or e-infrastructure.

Robust service discovery based on a study of different approaches for service discovery in Grid and cloud infrastructures, that did lead to a unified approach to publish and query services. The design and implementation of the registry uses event-driven replication of service information across registry nodes, which can be deployed as hierarchies (with aggregation) and as Pastry-based Peer-to-Peer network. Both of these approaches, when integrated in a registry, enable publishing and querying of collaborative infrastructure services in a robust fashion; this has been implemented in the EMI Service Registry (EMIR) service.

Integrative and unified architecture used in the two implementations EMIR and B2ACCESS. Both implementations combine federated service discovery and secure authenticated and authorised access to the infrastructure services.

5.2 Impact on Infrastructures and Users

B2ACCESS provides an Authentication and Authorisation Infrastructure (AAI) and is used as the federated identity management, authentication, and user management service in the EUDAT[10] production e-infrastructure. As of December 2020, there were 4246 unique registered EUDAT users, but as EUDAT is connected via B2ACCESS to the eduGAIN identity federation with approximately 27 million students, researchers, and educators, far more users than those having the specific EUDAT account use EUDAT via B2ACCESS. It enables scientific communities to access most of the EUDAT services using their federated identities. In addition, research infrastructures, such as CLARIN

---

[10]https://b2access.eudat.eu

and EPOS and recently the Helmholtz Data Federation (HDF) project, have deployed instances of the B2ACCESS services to manage users and secure infrastructure services.

The Common Language Resources and Technology Infrastructure (CLARIN)[11] is another research infrastructure, focusing on sharing, usage, and sustainability of language data and tools in the area of humanities and social sciences. Its production AAI infrastructure is mainly based on the Security Assertion Markup Language (SAML) and uses the B2ACCESS service. The service provides authentication to currently 2200 CLARIN users.

The European Plate Observing System (EPOS)[12] is one of the largest earth science research infrastructures that enables integrated use of data, data products, and facilities from distributed research infrastructures for solid-earth science in Europe. One of the major requirements of the EPOS community is to have an operational, scalable, federated AAI service that is able to interoperate with community services and is compliant with state-of-the-art technologies. Such a goal was achieved by devising a solution, namely the European Integrated Data Archive (EIDA) Authentication System, which is based on B2ACCESS [122]. The B2ACCESS service was tested in a focused use case by a targeted seismological community, namely the AlpArray seismologic network. Successively, the service was rolled out in production and it is currently operated by the Observatories & Research Facilities for European Seismology (ORFEUS) EIDA[13] project. Even though almost all the relevant data are open and accessible without the need of being authenticated, more than 400 users have already adopted the B2ACCESS service from a base of around 2500 global EPOS users.

B2ACCESS has also been deployed at the Juelich Supercomputing Centre (JSC), where it serves as an identity hub to secure the locally deployed JupyterHub[14] infrastructure and provides an authentication facility to approximately 1278 users (as of March 2020) from the German NREN (DFN-AAI) and local LDAP-based identity providers.

Being a part of EUDAT and being compliant with the AARC [12] Blueprint Architecture (BPA) and the Research and Education Identity Federations (REFEDS) Security Incident Response Trust Framework for Federated Identity (SIRTFI) [16], B2ACCESS has played an important role in achieving interoperability across research and e-infrastructures. In particular, it has enabled inter-federated access across the EOSC-Hub [49] e-infrastructure, which is a "super-federation", currently composed of three large federated e-infrastructures: EGI, INDIGO, EUDAT, and a number of scientific research communities.

EMIR, a unified federated service registry, which provides robust discovery (querying and publishing) of infrastructure services at a large scale. The registry has been adopted within the EMI [48] infrastructure to support its infrastructure operators and integrated within monitoring systems, application clients, and compute / data (UNICORE, ARC, CREAM, dCache) middlewares services. EMIR implements hierarchical aggregation as well as a peering approach to aggregate and replicate the service information across multiple registry nodes. The published service information is based on the standards-based GLUE 2.0 information model to support interoperability with other information services (GOCDB, BDII, and XSEDE information system).

---

[11] https://www.clarin.eu
[12] https://www.epos-ip.org
[13] http://www.orfeus-eu.org
[14] https://unity-jsc.fz-juelich.de

5.3 Future Work

Within the two areas covered by this thesis, service discovery and federated identity, several additional directions remain to be explored through feedback from research communities using B2ACESS and EMIR resulting in new requirements.

Concerning authorisation for federated identity management, it is often necessary to connect multiple attribute providers to the proxy, particularly where each community is a source of authority. Heterogeneity in the attribute sources raises concerns of trustworthiness, quality, and ownership of attributes. Addressing these concerns is crucial to assign adequate Level of Assurances (LoAs) and ensuring authorisation. In order to evaluate the LoA, it is indispensable to design and implement an attribute provenance method [99] to enrich access control policies. Furthermore, a metadata schema has been proposed [71] for the asserted user attributes for enriched authorisation as well as the possibility of data sharing permissions.

Identity tracing is trivial when there are only two parties involved: an IdP and an SP. However, with the emergence of chains of SPs/IdPs, leading to intermediaries between the service and identity federations, it has become more relevant for the service providing infrastructure to know the originating identity provider as well as the intermediate proxies involved in authenticating the user. Currently, proxies convey the information about the immediate authentication provider, which may be, for example, imprecise in calculating the LoA. Blockchain is an established technology for crypto-currencies, such as Bitcoin [103], and it can be applied in identity, trust management, and book-keeping [4]. However, blockchain-enabled identity tracing is still in its infancy and requires further research.

Use cases of scientific workflows sometimes require multiple compute and storage services to co-operate and accomplish the underlying tasks on the user's behalf. Given that, a multiple-hop delegated access to services could be required. The underlying middleware of the existing infrastructures adopts either the explicit trust delegation method [19] or short-lived X.509 certificates (as currently used in B2ACCESS) [149]. The latter approach is a standard and may need the ability to generate certificates on-demand through credential conversion services, for example, RCAuth [123] or CILogon [17]. These services do not have to be IGTF-accredited certification authorities, but if they were, it would help to establish inter-federation trust.

In the future, a standards-based delegation approach may be needed, such as the current draft OAuth 2.0 Token Exchange [91]. The federation proxy implements the Authorisation Server endpoint. Since the draft specification does not consider the deployment of multiple proxies, an ongoing activity in the EOSC-Hub project is attempting to address this.

Registry hierarchies in EMIR are currently static in nature and require each child registry node to pre-configure its parent (registration endpoint) for propagating the service records. This configuration includes setting ACLs which include the parent/child certificate distinguished names. Similarly, the node de-registration from the parent requires manual update in the ACLs. Future work in EMIR will incorporate the dynamic provisioning of registry nodes, thus making EMIR more flexible.

Paper I

The EMI Registry: Discovering Services in a Federated World

L. Field, A.S. Memon, I. Márton and G. Szigeti. 2014.

Journal of Grid Computing 12(1), 29 - 40 (2014) [DOI: 10.1007/s10723-013-9284-1]

Shiraz Memon was the product leader of the EMIR service and therefore substantially contributed in the analysis, design, implementation, and evaluation of the service. He is one of the main authors of this publication and provided the majority of the content of this publication.

The EMI Registry: Discovering Services in a Federated World

**Laurence Field · Shiraz Memon · Iván Márton ·
Gábor Szigeti**

Abstract The Distributed Computing Infrastructure
(DCI) has become an indispensable tool for sci-
entific research. Such infrastructures are composed
of many independent services that are managed by
autonomous service providers. The discovery of ser-
vices is therefore a primary function, which is a
precursor for enabling efficient workflows that utilise
multiple cooperating services. As DCIs, such as the
European Grid Initiative (EGI), are based on a fed-
erated model of cooperating yet autonomous service
providers, a federated approach to service discovery is
required that seamlessly fits into the operational and
management procedures of the infrastructure. Many
existing approaches rely on a centralised service reg-
istry, which is not suited to a federated deployment
and operational model. A federated service registry
is therefore required that is capable of scaling to
handle the number of services and discovery requests
found in a production DCI. In this paper we present
the EMI Registry (EMIR), a decentralised archi-
tecture that supports both hierarchical and peering
topologies, enabling autonomous domains to collab-
orate in a federated infrastructure. An EMIR pilot
service is used in order to evaluate a prototype of
this architecture under real-world conditions with a
geographically-dispersed deployment. The results of
this initial deployment are provided along with a few
performance measurements.

1 Introduction

The Distributed Computing Infrastructure (DCI) has
become an indispensable tool for scientific research
[16]. Such infrastructures are composed of many
independent services [14] that are managed by
autonomous providers. The discovery of services is
therefore a primary function, which is a precursor
for enabling efficient workflows that utilise multiple
cooperating services.

The provision of a service registry can be used to
fulfil such a requirement. Existing service registries,
such as the Advanced Resource Connector (ARC)
Information Index [6] or UNICORE Registry [24],
are examples that have proven themselves in produc-
tion environments. However, these implementations

L. Field (✉)
CERN, Geneva, Switzerland
e-mail: laurence.field@cern.ch

S. Memon
FZJ, Jülich, Germany
e-mail: a.memon@fz-juelich.de

I. Márton · G. Szigeti
NIIF, Budapest, Hungary

I. Márton
e-mail: martoni@niif.hu

G. Szigeti
e-mail: szigeti@niif.hu

@ Springer

follow a centralised approach, whereas DCIs, such as the European Grid Initiative (EGI) [17], are based on a federated model. A service registry needs to mirror such a model in order for it to seamlessly fit into the operational and management procedures of the infrastructure. It must also scale to the number of services in the DCI and the number of discovery requests that are executed.

DCIs are comprised of domains, which are autonomous and can operate in isolation. A good example is a National Grid Initiative (NGI). These in turn are composed of multiple autonomous institutions that provide services. The NGI would like to support collaboration within their national borders and to also participate in multinational initiatives such as EGI. To achieve this goal, the NGI needs to provide the service discovery function for national users and share information on services within its domain with the multinational initiative. The scenario is similar for the organisation that provides the services; they need to support the service discovery function for local users and to share information on services with the NGI. With the advent of cloud computing and the adoption of cloud-based services in Grid infrastructures, interoperability between Grid services and cloud services is a concern. As the discovery of services is a primary function, it must be understood how both Grid and cloud services can be discovered.

This paper provides an overview of related work in the area of federated service discovery and the shortcomings of existing solutions. It presents an architecture for a federated service registry and a prototype based on this architecture, the EMI Registry (EMIR). A world-wide pilot service is used to ensure that the EMIR is robust to the kind of issues that are associated with distributed environments and to understand the operational aspects. The results of some initial performance tests are provided to demonstrate that EMIR is able to handle the scale required for a production Grid infrastructure in terms of services and discovery requests.

The next section provides an overview of existing service registries and their limitations for use within a federated environment. Section 3 presents the proposed architecture and describes the EMIR implementation. The results of the pilot service and performance-testing are given in Section 4 followed by some concluding remarks in Section 5.

2 Overview of Existing Approaches

In the seminal paper [12], *A Directory Service for Configuring High-Performance Distributed Computations*, the need for a high-performance distributed information system for the emerging field of Grid computing was stated. The Metacomputing Directory Service (MDS) was presented, which consisted of two basic elements; Grid Resource Information Services and Grid Index Information Services.

Grid Resource Information Services respond to queries about the resource and Grid Index Information Services evaluate queries against an internal index of registrations and forwards the query or provides a response from its cache as appropriate. A Grid Resource Information Service registers to a Grid Index Information Service using the Grid Resource Registration Protocol which is a soft-state protocol, meaning the state established by a notification may eventually be discarded unless refreshed by subsequent notifications. A Grid Index Information Services can use Grid Resource Registration Protocol to register to another Grid Index Information Service, which creates a hierarchical structure of Index Services.

While MDS provides a complete information service that could be used for service discovery, it was designed for hierarchical deployment. One aspect of a federated model is that there is no concept of *top* as federations have equal status. The Grid Resource Registration Protocol and the information in the Index Services are the core features of MDS that support a service discovery function. However, these are not explicitly exposed and only used internally to support the query functionality of the information service. If the Grid Resource Information Services mapped to a single service, the index could be used for service discovery, however, the internal data model has not been designed with this functionality in mind.

During the evaluation of MDS in the European DataGrid project [15], instabilities with MDS were observed [19]. To work around these issues, a top-level cache based on a standard OpenLDAP [18] server was used. Periodically, information was extracted from the top-level the MDS and added to the OpenLDAP server. This top-level cache showed excellent behaviour under load [19] and was named the Berkeley Database Information Index (BDII) to distinguish it from the MDS. One simplification of the BDII is

the absence of the Grid Resource Registration Protocol, where the index is replaced by a local file that contains the URLs of the lower-level services. Rather than forwarding queries, the BDII responds to them directly using a cache, which updated asynchronously to the queries by periodically querying the lower-level services for all information. It has been noted [9] that this results in inefficient network usage as information that has not changed is transported multiple times. For the production deployment in EGI, the local file is automatically updated by obtaining the URLs from the Grid Operations Centre Database (GOC DB) [20]. The GOC DB is a central authoritative database that contains static service and topology-related information for the purpose of infrastructure operations which is maintained by the appropriate people [20]. As a consequence of the federation requirement, the GOC DB will evolve from a central database to a distributed model that allows regional instances to communicate with one another. The disadvantages with this approach is that it will only support one level of federation and the information management is not automated.

The Advanced Resource Connector (ARC) Grid middleware [7] adopted an alternative strategy to address the instabilities with MDS. The aggregation feature (cache) was disabled and queries were performed in two steps; collect a list of contact URLs from an Enhanced Grid Index Information Service and query the ARC Resource Information Service directly. The Enhanced Grid Index Information Service also contains contact information for other Enhanced Grid Index Information Service instances and the multiple queries are hidden from the end user through the provision of a client library.

The Enhanced Grid Index Information Service instances are organised in a hierarchical structure based on a geographical organisation; services belonging to the same region register under a region index, region indices register to the appropriate country index and country indices register to the top-level index services. In order to avoid a single point of failure, four top-level indices are used and each country index registers to all top-level indices. To improve upon this, an Information System Indexing Service [2] based on a peer-to-peer approach was designed as a proof-of-concept to automatically distribute registrations between the top-level indices. The disadvantage with this approach is that the multi-query approach

does not perform well at scale [3]. The client has to query many information sources and if there are many queries, the information sources are potential bottlenecks in the system.

In UNICORE [24], service discovery is enabled via a global service registry based on Web service technology to which all organisations (hosting the services) are required to publish endpoint information. The organisations themselves host a Common Information Service [21], which provides more detailed information about the service. This centralised approach is the main disadvantage for federated environments. In addition, due to the tight coupling with UNICORE, discovery of non-UNICORE based services is not supported.

Beyond the Grid environment, the Universal Description, Discovery and Integration [5] (UDDI) standard is a mechanism to register and discover Web services. It has three components: White Pages (identifiers), Yellow Pages (categorisations based on standard taxonomies) and Green Pages (technical documents describing the protocol bindings and message formats required to interact with the Web services). These service descriptions are defined in a platform-independent, XML-based registry which is interrogated via SOAP messages. Even though UDDI is the de-facto industry standard for Web services discovery, the imposed requirements of tight-replication among registries and lack of autonomous control, among other things, has severely hindered its widespread deployment and use [4]. Today we find that UDDI has not been widely deployed and in fact, the only known uses of UDDI are as private or semi-private UDDI registries within the enterprise boundary [4]. The latest version (v3) of UDDI does not offer any special features for discovering Web service registries. As a result, it is assumed that Web service clients require prior knowledge of the access points of the registries.

The Distributed UDDI Deployment Engine [4] (DUDE) attempts to overcome this limitation by providing a rendezvous mechanism between multiple UDDI registries based on Distributed Hash Tables. A Distributed Hash Table is a peer-to-peer distributed system that forms a structured overlay allowing more efficient routing than the underlying network [4].

Similarly, the METEOR-S Web Service Discovery Infrastructure [23] attempts to provide support for the discovery and publication over a group of autonomous registries in a multi-registry environment.

<img> Springer

This is achieved using an ontology-based approach where groups of registries are divided into domains and grouped into federations. The semantic metadata of the registries in the infrastructure is stored in an ontology which is used to identify appropriate registries and direct the queries to them. In this context, a *Registry Federation* is defined to be a collection of autonomous but cooperating Web service registries.

Both the DUDE and MWSDI are attempts to federate UDDI registries, however, the coupling only serves to discover relevant registries for a query and as mentioned previously, the multi-query approach does not perform well at scale [3]. In addition, as the approaches have been specifically designed for a Web services environment, it not clear if they are compatible with the sharing polices [13] upon which Grid computing is based.

A number of service discovery protocols exist [25], such as the Service Location Protocol and Jini, that are designed for the discovering networked services in a Local Area Network environment and as they do not scale to a Wide Area Network environment, they have not been considered.

The approaches discussed in this section clearly demonstrate that there is a need for service discovery and services registries play an important role in this respect. It has also been highlighted that a centralised approach is not always feasible and there have been a few attempts to address this. If two or more service registries exist, it must be understood how these can be used together to support the service discovery function. The concept of a *Registry Federation* as a collection of autonomous but cooperating service registries, clearly defines this scenario. There is also little difference between domains of Grid, cloud and Web services with respect to service discovery. The fundamentals of service discovery for online services that are connected to the Internet is the same. The only difference is the information model used, which many contain technology specific information. A technology agnostic approach would therefore enable the discovery of Grid, cloud and Web services. An architecture, along with a concrete implementation, is therefore required to enable service discovery in a federated environment.

3 The EMI Registry

The goal of the EMIR architecture is to provide robust and scalable service discovery in a federated environment. The design is a result of a collaboration between the major European middleware providers (ARC, dCache, gLite and UNICORE) in the European Middleware Initiative (EMI) [1]. It aims to consolidate and evolve the existing middleware stacks and brings together significant experience with many of the approaches outlined in Section 2.

3.1 Concepts

The foundation is that the services to be discovered are online services which are connected to the Internet. The primary use case is Grid services but as generic information model has been adopted, this could be extended to other use cases such as discovering cloud-based services. Services are grouped primarily with other services that are managed by the same organisation that is autonomous and such grouping is call a *Domain*. Domains can be organised in a hierarchical structure, however there is no single *top-level* organisation. Domains can also be organised to form a *federation*, whereby each federation is the authoritative source for services within its federation and shares this information in a peer-like fashion with the other federations.

EMIR provides two main building blocks; the Domain Service Registry (DSR) and Global Service Registry (GSR). It is envisaged that these base components can be used to support different topologies that map to the real operational and deployment models of DCIs. A Service Publisher periodically sends information about the service to the DSR and is a soft-state protocol, similar to the Grid Resource Registration Protocol. The DSR may re-publish this information to a parent DSR/GSR, or in the case of the GSR, replicate this information n-times (where n is the number of GSRs) to the other GSRs. The interactions between the DSR and the other components is shown in Fig. 1.

A peer-to-peer network is formed from GSRs and is configured using a static list as it is assumed that the placement of the GSR will be predetermined when the topology is defined, and that this will not undergo significant changes throughout the lifetime of the infrastructure. An example topology is shown in Fig. 2. The message routing in the GSR peer-to-peer network is
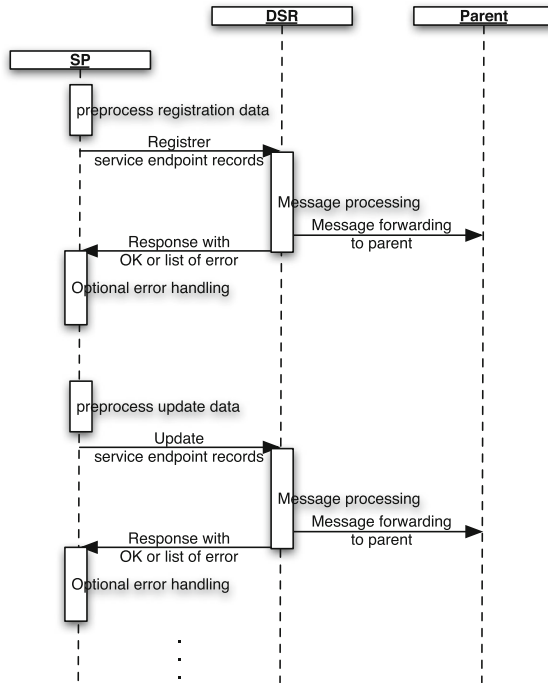
Fig. 1 DSR interactions

based upon the Information System Indexing Service prototype introduced in Section 2. An authentication and authorization mechanism is adopted by all components to ensure that only authorized registration can occur. A policy engine is available to control which registrations can be propagated.

3.2 Service Record

The core aspect of the system is the service registration record. The implementation of the service record adopted by EMIR is based on the GLUE 2.0 information model [8]. The service record is a *profile* of the model that provides the relevant information for this use case. It contains mandatory attributes (described in Table 1) that are required for EMIR to function and optional attributes that can be used to

extend the record. It can be seen from the mandatory attributes in Table 1 that the record can be used to describe any service that can be identified with a URL. The abstract attributes Capability and Technology advertise what the service does and how it does it. To ensure interoperability between federations and infrastructures, standardisation on such a profile is highly desirable. It contains mainly static information that does not change during the lifetime of the service. This is done to avoid the need for updating service records, which would add additional complexity to the system. However, updates to service records is supported, and hence semi-static information such as version information is available. Highly-dynamic information, state information that changes frequently, is not available in the service record.

One aspect of EMIR is that the validation of the service record is only done in the service publisher, the

Fig. 2 Example EMIR
topology
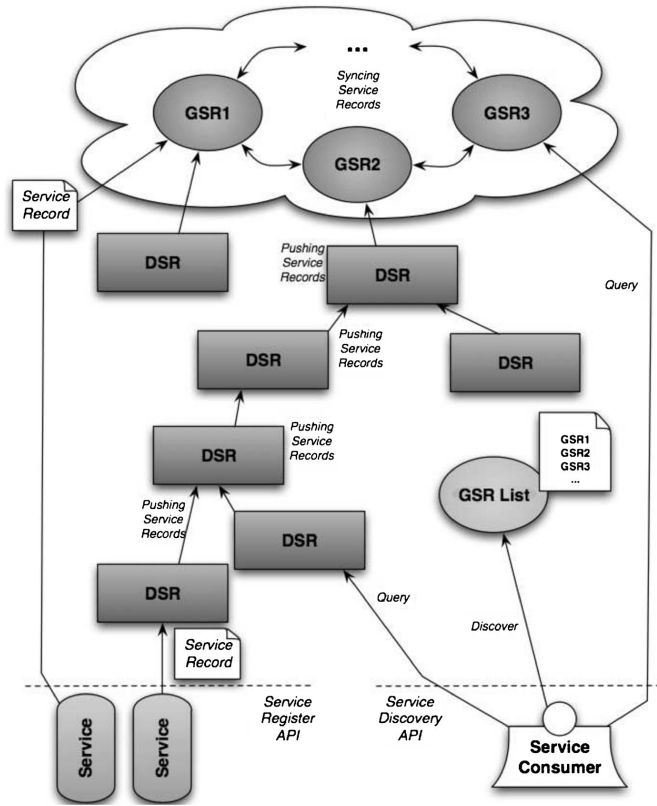
DSR and GSR components are *schema-free* (with the exception of attributes required for EMIR to function). This has been done to simplify any future changes to the service record by avoiding the need to upgrade the DSR or GSR in order to accommodate those changes.

3.3 Implementation

The EMIR is comprised of a number of components, each of which focuses on a specific function. These components are core, database, security, information model, and client side access. The majority have been implemented in Java, however other supporting technologies have been used to implement some capabilities where appropriate.

The core of the EMIR is the service record management functionality that includes the validation of records, validity (time to live) of the records and the synchronization of records (both hierarchical and peer-to-peer). To ensure client independence, a REST architectural style has been provided entailing HTTP URIs which expose CRUD (batch/single) operations.

The management (create, update, delete) of service records must be restricted to the entity that is authoritative for the service record, which in the EMIR architecture is the service itself. Authentication is achieved using X.509 certificates and the trust anchors from the International Grid Trust Federation. Access Control Lists are used to specify which services can

Table 1 The mandatory attributes in the EMIR service record

Attribute	Description
ServiceID	A global unique identifier for the service
ServiceName	Human-readable name
ServiceType	The type of service
ServiceEndpointID	A global unique identifier for the service endpoint
ServiceEndpointURL	Network location of an endpoint
ServiceEndpointCapability	The provided capability
ServiceEndpointTechnology	The technology used to implement the endpoint interface
ServiceEndpointInterfaceName	The name of the primary protocol
ServiceEndpointInterfaceVersion	The version of the primary protocol
ServiceExpireOn	The time after which the record expires

publish the records and from which DSR/GSR records can be accepted. The list contains a Distinguished Name along with associated attributes, which represent roles. As the primary query interface is REST, any HTTP client is able to query the registry to discover services. The service record is considered public information, similar to phone numbers in a telephone directory, and hence there is no restriction on who can perform queries. If the advertisement of a particular service is sensitive, it should not be published and an alternative method to directly communicate the required information on a peer-to-peer basis should be used instead.

Two clients have been developed to support the interaction with EMIR; a Java-based client library for queries and the python-based Service Endpoint Resource Publisher that publishes records on behalf of a service in a periodic manner.

3.4 Advantages

We believe that the EMIR provides a number of improvements over existing approaches. Primarily, the DSRs and GSR support the use of both hierarchical and peering models in the same topology. The hierarchical model supports multiple levels of service registries and the addition of the peering model promotes its adoption in a federated environment. This is the main advantage of EMIR and to the best of our knowledge the first time an architecture combining the two models has been used for a service registry. There are no constraints on the domains themselves as a domain is just a collection of services that are autonomously managed by the same organisation. As

such EMIR offers flexibility with the placement of services registries and how they can be linked. The DSR and GSR respond to queries directly from an internal cache of the service records, which according to [3], should perform better than an approach that forwards the queries. The use of a REST architectural style for interacting with the DSRs and GSRs results in simple publishing and querying interfaces. Custom DSRs that only implement these public interfaces offer the domain flexibility with respect to the provision of service records and can act as gateways [10] facilitating interoperability between infrastructures.

Integral to the EMIR approach is the concept that the services are the authoritative information source for information about themselves and the DSR is the authoritative information source for which services are in their domain. A soft-state registration protocol is used to add the service record and ensure that it is up-to-date. In fact, once the registry topology has been deployed, the handling of service records is fully automated and does not require any manual intervention. This feature of EMIR supports dynamic service provisioning, which is becoming increasingly important with the advent of cloud computing. The service record is based on a standard information model that supports the description of generic services where a service is defined as an endpoint that offers a capability. As such, EMIR can support service discovery in a technology agnostic way which enables the discovery of Grid, cloud and Web services among others, within a single framework. However, initial support for the International Grid Trust Federation trust anchors primes EMIR for immediate use within the Grid environment.

4 Initial Results

In order to evaluate the EMIR architecture, two testing scenarios are used. The first is a pilot service, which evaluates the stability and operational aspects of an EMIR infrastructure under real-world conditions [11], i.e., comparable to the EGI production Grid infrastructure. This considers a geographically-dispersed deployment with real information and system administrators from production Grid sites. The second testing scenario evaluates the performance of the registry under conditions that are comparable and beyond the scale of the EGI.

4.1 Pilot Service

The topology used for the pilot service was a simple hierarchy, with one-level of DSRs that register to a single GSR. An existing tool (ginfo) for querying GLUE 2.0 records, was enhanced to output EMIR-compliant service records in JSON. This tool was used to provide real service records by querying the production site-level BDII for the sites that were participating in the pilot. The roll-out was conducted in two phases. Phase One used plain HTTP (no authentication) and Phase Two switched to HTTPS. Using HTTP reduces the configuration space and hence scope for errors, enabling Phase One to focus mainly on issues with EMIR. In order to verify that a DSR was functioning, service records from the site were extracted hourly from the GSR. This was achieved using ginfo, which had also been enhanced to query the EMIR. The site name was extracted from the service record and a KML file was created to show the sites on Google Maps. This simple visualization made it easy to see if a DSR was publishing correctly. Ten sites contributed to the pilot service; five sites in Europe (CERN, The Helsinki Institute of Physics, Forschungszentrum Jülich, Centre of Supercomputing of Galicia, and National Documentation Centre in Greece) and five sites world-wide (Academia Sinica Grid Computing Centre, The University of Melbourne, Canada's National Laboratory for Particle and Nuclear Physics, The Council for Scientific and Industrial Research (CSIR) in South Africa, and Centro Brasileiro de Pesquisas Fisicas).

The initial feedback from the system administrators was positive. On average it took 30 min for a system administrator to install and configure the DSR.

This deployment experience reinforced a number of software engineering principles for large-scale [22] distributed systems. The probability that at least one component will fail increases with scale and hence with a sufficiently large number of components in a distributed infrastructure it is almost certain that a specific component has failed. Hence, a distributed infrastructure such as EMIR should be made robust to component failures. An example was discovered during the pilot with the handling of malformed service records. As the infrastructure is expected to handle many thousands of records, it is certain that some will be incorrect so the components must be made robust to handle this scenario. Similarly, if a configuration step is required, there is a probability associated that it will be done incorrectly, so configuration should be as simple as possible.

It was also noticed that the scale of deployment can add a significant administrative overhead, which can easily be overlooked. A few mistakes in the documentation resulted in problems with the installation and configuration of the first instance and took approximately three hours combined to investigate and fix. If a staged roll-out approach was not adopted, this problem could have been experienced by all ten administrators resulting in 30 h of lost effort. Scaling up to production-like deployment with 400 instances, 1,200 h would have been lost. It is clear that in a large-scale distributed deployment scenario, this *overhead* needs be minimized and hence the packing of software (including documentation, configuration and diagnostics) is of utmost importance.

After the issues found during the initial deployment of the pilot service were addressed, the infrastructure has been stable (all sites are visible on the map).

4.2 Performance Tests

From a quality of service perspective, there are two important metrics [11]; the query response time for concurrent client requests and the freshness of information returned. A recent study [11] stated that the EGI infrastructure contains over 4000 services and an instance of the EGI information system handles 2 million queries per day (23.1 per second), with an average query response time of between 0.003 s and 0.41 s depending on the query used.

An instance of the EMIR was deployed in a Virtual Machine running Scientific Linux 5, which was
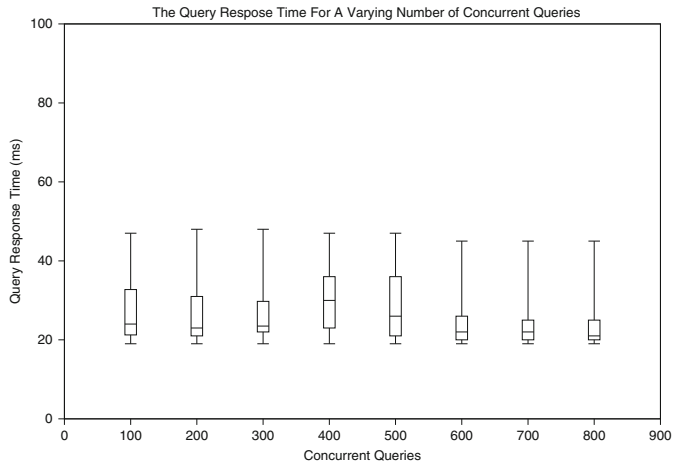
Fig. 3 The query response time for concurrent queries for a registry containing 10 K service records

configured to have 2 cores and 2 GB of RAM. A test client was created that spawned a pre-defined number of threads, each of which queried the EMIR server for service endpoints of a specific service type, resulting in a query response that was approximately 5 % of the total number of services. This client was deployed on a separate Virtual Machine instance with 2 cores and 4 GB of RAM. Both of these machines were located on the same Local Area Network. The registry was populated with a collection of service records containing arbitrary values, each of which was 1.6 KB in size.
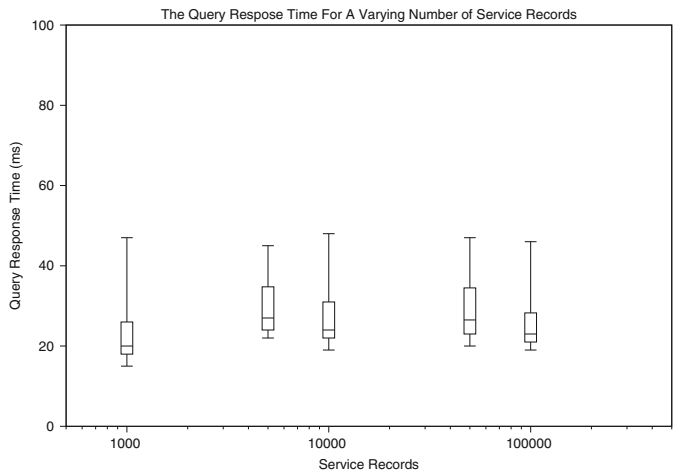
Fig. 4 The query response time for 50 concurrent queries for a varying number of service records

Table 2 The latency for 50 records to be available in the next tier

Tier	Location	Latency (ms)	σ (ms)
1	USA	60.9	3.0
2	Hungary	223.6	4.4
3	USA	207.6	2.9
4	Hungary	248.0	6.0

The query response times for different numbers of concurrent client requests for a registry containing 10,000 service records (more than double the size of the EGI) is shown in Fig. 3. It can be seen that for up to 800 concurrent clients, the query response time was almost constant. The first few initial requests were discarded as they take between 2 and 3 times longer and hence the response time was measured once a steady-state has been reached. There are two underlying factors for this; firstly, an in-memory cache of the query result is created and secondly, indexes are created on receiving the initial query request.

The query response time was measured for a varying number of service records contained in the registry using a query load of 50 concurrent queries (comparable to the EGI). It can be seen in Fig. 4 that for up to 100,000 service records (more than 10 times the size of the EGI), the query response time is almost constant.

The freshness of information is directly related to the latency of registration requests. A distributed multi-tier hierarchy of registries spanning three countries was used to measure the latency of registrations. Two DSR instances (tiers 1 and 3) were deployed in the U.S.A. using separate Virtual Machine instances on the Future Grid infrastructure. Another two DSR instances (tiers 2 and 4) were deployed in Hungary. The machines were synchronised using an NTP server. A client (located in Germany) was used to send service registration requests. The mean latencies for 50 such requests to be available in each tier is shown in Table 2.

Based on these results, the addition of a tier that includes a transatlantic link increases the latency by between 200 ms and 250 ms. According to [11], such latency is sufficient to have a negligible effect on the freshness of service records that contain mainly static information.

5 Conclusion

In this paper we presented the EMIR, a decentralised approach to service discovery that supports both hierarchical and peering topologies enabling autonomous domains to collaborate in a federated infrastructure. Once the registry topology has been deployed, the handling of service records is fully automated and new domains or federations can be added to the infrastructure when required. This feature of EMIR supports dynamic service provisioning which is becoming increasingly important with the advent of cloud computing. In fact, as the service record is based on a standard information model which is supports the description of generic services, EMIR can support service discovery in a technology agnostic way and enable the discovery of Grid, cloud and Web services among others within the same framework. EMIR provides two main building blocks; the DSR and the GSR, which can be used to support different deployment topologies that map to the real operational and deployment models of DCIs. The core aspect of the system is the service registration record, and the EMIR has adopted the GLUE 2.0 information model for attribute names and semantics. This represents a profile of the model and to ensure interoperability between the federations and infrastructures, standardisation of such a record is highly desirable.

A pilot service was used to evaluate the EMIR in real-world conditions with a geographically-dispersed infrastructure. The initial feedback from the system administrators was positive and the deployment experience reinforced a number of software engineering principles. This was complemented with a performance evaluation in a simulated deployment scenario that measured the query response time for concurrent client requests and the freshness of information

returned. The average query response time for 800 concurrent clients requests for a registry containing 10,000 service records was 21 ms with an upper bound of 129 ms. The mean latency for an initial registration is 60.9 ms with $\sigma = 3.0$ ms, with an increase of between 200 ms and 250 ms for each additional tier that includes a transatlantic link and will have a negligible effect on the freshness of service records.

Based on these preliminary results we believe that the EMIR is suitable for providing a solution for service discovery in a federated environment. The work would benefit from further and more detailed performance and scalability tests, especially relating to the replication of service records between the GSRs.

References

1. Aiftimiei, C., Aimar, A., Ceccanti, A., Cecchi, M., Di Meglio, A., Estrella, F., Fuhrmam, P., Giorgio, E., Konya, B., Field, L., Nilsen, J.K., Riedel, M., White, J.: Towards next generations of software for distributed infrastructures: the European middleware initiative. In: Proceedings of the 8th International Conference on E-Science, pp. 1–10. Chicago (2012)

2. Appleton, O., Cameron, D., Cernak, J., Ellert, M., Fragat, T., Gronager, M., Johansson, D., Jonemo, J., Kleist, J., Kocan, M., Konstantinov, A., Konya, B., Marton, I., Mohn, B., Moller, S., Muller, H., Nagy, Z., Nilsen, J.K., Ould Saada, F., Pajchel, K., Qiang, W., Read, A., Rosendahl, P., Roczei, G., Savko, M., Skou Andersen, M., Smirnova, O., Stefan, P., Szalai, F., Taga, A., Toor, S.Z., Waananen, A., Zhou, X.: The next-generation ARC middleware. Ann. Telecommun. **65**(11–12), 771–776 (2010)

3. Baeza-Yates, R., Ribeiro-Neto, B.: Modern Information Retrieval. Addison Wesley, Harlow (1999)

4. Banerjee, S., Basu, S., Garg, S., Garg, S., Lee, S.-J., Mullan, P., Sharma, P.: Scalable Grid service discovery based on UDDI. In: Proceedings of the 3rd International Workshop on Middleware for Grid Computing, pp. 1–6. Grenoble (2005)

5. Curbera, F., Duftler, M., Khalaf, R., Nagy, W., Mukhi, N., Weerawarana, S.: Unraveling the Web services web: an introduction to SOAP, WSDL, and UDDI. IEEE Internet Comput. **6**(2), 86–93 (2002)

6. Ellert, M., Gronager, M., Konstantinov, A., Konya, B., Lindemann, J., Livenson, I., Nielsen, J., Niinimaki, M., Smirnova, O., Waananen, A.: Advanced resource connector middleware for lightweight computational grids. Futur. Gener. Comput. Syst. **23**(2), 219–240 (2007)

7. Ellert, M., Konstantinov, A., Konya, B., Smirnova, O., Waananen, A.: The NorduGrid project: using Globus toolkit for building Grid infrastructure. In: Proceedings of the VIII International Workshop on Advanced Computing and Analysis Techniques in Physics Research, vol. 502, pp. 407–410, Moscow (2002)

8. Field, L., Konya, B., Andreozzi, S.: GLUE Specification v. 2.0. Recomendation GFD.147, Open Grid Forum (2009)

9. Field, L., Harvey, P., Dyce, T.: Designing the next generation Grid information system. In: Journal of Physics: Conference Series, vol. 331, Taipei (2011)

10. Field, L., Laure, E., Schulz, M.W.: Grid deployment experiences: Grid interoperation. J. Grid Computing **7**(3), 287–296 (2009)

11. Field, L., Sakellariou, R.: Benchmarking Grid Information Systems. In: The Proceedings of the 17th International Conference on Parallel Processing, vol. 6852, pp. 479–490, Bordeaux (2011)

12. Fitzgerald, S., Foster, I., Kesselman, C., von Laszewski, G., Smith, W., Tuecke, S.: A directory service for configuring high-performance distributed computations. In: Proceedings of the Sixth IEEE International Symposium on High Performance Distributed Computing, pp. 365–375, Portland (1997)

13. Foster, I.: The anatomy of the Grid: enabling scalable virtual organizations. pp. 6–7. IEEE Comput. Soc. (2001)

14. Foster, I., Kesselman, C., Nick, J.M., Tuecke, S.: Grid services for distributed system integration. Computer **35**(6), 37–46 (2002)

15. Gagliardi, F., Jones, B.: European datagrid project: Experiences of deploying a large scale Testbed for e-Science applications. Perform. Eval. Compl. Syst. Tech. Tools **2459/2002**, 255–264 (2002)

16. Hey, T.: Cyberinfrastructure for e-Science. Science **308**(5723), 817–821 (2005)

17. Kranzlmller, D., Marco Lucas, J., Oster, P.: The European Grid Initiative (EGI). In: Remote Instrumentation and Virtual Laboratories, pp. 61–66. Springer US, Boston (2010)

18. Liebel, O., Ungar, J.M.: OpenLDAP. Galileo Press, Bonn (2006)

19. Loomis, C.: Final evaluation of testbed operation. EU Deliverable DataGrid-06-D6.8-414712-3-0 (2003)

20. Mathieu, G., Richards, A., Gordon, J., Del Cano Novales, C., Colclough, P., Viljoen, M.: GOCDB, a topology repository for a worldwide Grid infrastructure. In: Journal of Physics: Conference Series, vol. 219, Taipei (2010)

21. Memon, A.S., Memon, M.S., Wieder, P., Schuller, B.: CIS: An information service based on the common information model. In: Third IEEE International Conference on e-Science and Grid Computing, pp. 465–473, Bangalore (2007)

22. Ordille, J.J., Miller, B.P.: Database challenges in global information systems. In: Proceedings of the 193 ACM SIGMOD International Conference on Management of Data, pp. 403–407, Washington, DC (1993)

23. Sivashanmugam, K., Verma, K., Sheth, A.: Discovery of Web services in a federated registry environment. In: Proceedings of the IEEE International Conference on Web Services, pp. 270–278, San Diego (2004)

24. Streit, A., Bala, P., Beck-Ratzka, A., Benedyczak, K., Bergmann, S., Breu, R., Daivandy, J.M., Demuth, B., Eifer, A., Giesler, A., Hagemeier, B., Holl, S., Huber, V., Lamla, N., Mallmann, D., Memon, A.S., Memon, M.S., Rambadt, M., Riedel, M., Romberg, M., Schuller, B., Schlauch, T., Schreiber, A., Soddemann, T., Ziegler, W.: UNICORE 6 recent and future advancements. Ann. Telecommun. - Ann. Telecommun. **65**(11–12), 757–762 (2010)

25. Zhu, F., Mutka, M.W., Ni, L.M.: Service discovery in pervasive computing environments. IEEE Pervasive Comput. **4**(4), 81–90 (2005)

Paper II

Federated Authentication and Credential Translation in the EUDAT Collaborative Data Infrastructure

A.S. Memon, J. Jensen, A. Cernivec, K. Benedyczak and M. Riedel. 2014.

IEEE 7th International Conference on Utility and Cloud Computing Proceedings (2014) ISBN 978-1-4799-7881-6 [DOI: 10.1109/UCC.2014.118]

As a part of the EUDAT Authentication and Authorisation Infrastructure (AAI) task force, Shiraz Memon is the main contributor of designing, implementing, and evaluating the B2ACCESS service, which offers federated authentication, identity management, credential translation, and user attribute harmonisation in the EUDAT data management infrastructure. He is the main author of the content of this publication.

Federated Authentication and Credential Translation in the EUDAT Collaborative Data Infrastructure

Ahmed Shiraz Memon*¶, Jens Jensen†, Aleš Černivec‡, Krzysztof Benedyczak§, Morris Riedel*¶

*Juelich Supercomputing Center, Forschungszentrum Juelich GmbH
{a.memon, m.riedel}@fz-juelich.de
†Rutherford Appleton Laboratory,Science and Technology Facilities Council, UK
jens.jensen@stfc.ac.uk
‡XLAB Ljubljana, SI
ales.cernivec@xlab.si
§ICM Warsaw, PL
golbi@icm.edu.pl
¶School of Engineering and Natural Sciences, University of Iceland, Iceland
morris@hi.is

Abstract—One of the challenges in a distributed data infrastructure is how users authenticate to the infrastructure, and how their authorisations are tracked. Each user community comes with its own established practices, all different, and users are put off if they need to use new, difficult tools. From the perspective of the infrastructure project, the level of assurance must be high enough, and it should not be necessary to reimplement an authentication and authorisation infrastructure (AAI).

In the EUDAT project, we chose to implement a mostly-loosely coupled approach based on the outcome of the Contrail and Unicore projects. We have preferred a practical approach, combining the outcome of several projects who have contributed parts of the puzzle. The present paper aims to describe the experiences with the integration of these parts. Eventually, we aim to have a full framework which will enable us to easily integrate new user communities and new services.

Keywords-PKI, EUDAT, federated identity management, OAuth, SAML, OpenID

I. INTRODUCTION

Federated Identity Management enables access to protected web resources via the public internet, or on private intranets by bridging different identity domains: "federation" here means that identity providers (IdPs) and service providers (SP) are separate entities, usually bound together by a common policy. From the user's perspective, they use the same password with every SP; from the SP's perspective, they no longer need to worry about account management such as resetting passwords and keeping details up-to-date [1]. Authentication can in principle range from simple username and password (*e.g* [2].) to smartcards or biometrics (*e.g* [2].)

The EUDAT [3] project supports a number of research communities, currently primarily linguistics and earth, climate, and medical sciences. To build an identity federation, we will need to work with what the communities already use. Thus, it is essential to support multi-technology, different levels of assurance (LoA), and different policies. Beyond this comes (as yet not fully solved, and beyond the scope of this paper), fully harmonised authorisation.

One of the early design decisions was to have X.509 certificates generated internally. While we would not expect the general user to be willing or able to manage X.509 certificates directly, we do need X.509 to drive a number of non-web processes, such as GridFTP transfers, or access to iRODS. Moreover, a delegation mechanism was needed, to enable services to act on behalf of users, and we could then support delegation (or "impersonation" for the purists, as the service essentially acts as the user.) Like many projects before us, we have implemented this via a portal front end to which users authenticate using federated identities, and a key pair is then generated by the portal and signed into a certificate by a federation-level CA. This approach gives a higher level of protection than a shared certificate, and it aids scalability in two ways: first, by disambiguating simultaneous users of the same portal, and second, by enabling the same user to access EUDAT via more than one portal (the user would get the same name regardless of which portal they access.) We also use the X.509 certificate to carry federation-level attributes.

This paper is organised as follows: Section II discusses related work; Section III describes the proposed AAI architecture and Section IV the "data staging" use-case. Section VI concludes our paper.

II. RELATED WORK

As we have mentioned, it is a very common approach to provide a front end portal for the users, and then either have the users share a certificate once they are authenticated, or to generate a certificate for each individual user. The latter approach is obviously slightly more demanding, but also offers greater assurance for the resource.

726

CILogon [4] is an example of the latter, a portal where a certificate is generated on behalf of a user. Indeed, like our goal, CILogon allows authentication with IdPs of different LoA (namely, InCommon silver and bronze, and Google.)

The Terena Certificate Service [5] takes identities from IdPs that are already members of federations–namely, the national federations run by national research and educational networks–and uses them to authenticate users and validate their identities, and their right to obtain a certificate. Only identities of a sufficiently high LoA can be used (and where an agreement exists between Terena and the organisation running the IdP, so IdPs must publish a particular attribute to assert that the LoA is sufficient.

A *Security Token Service* (STS) is a "a Web service that issues security tokens" [6]. Intended for SOAP-based web services, they are designed to pass SAML security tokens [7] in contexts such as WS-Federation [8]. The European Middleware Initiative (EMI) built an STS which interfaces to a CA (using EJBCA) and is therefore capable of generating certificates on behalf of an authenticated user. Note that an STS is really designed for SOAP web services and users would not directly interface with it with a browser except to authenticate to the STS, as per WS-Federation[1].

Another recent example of a token service is the INFN "eTokenServer" [11] where users authenticate to a service which–in this case–issues certificates for use by automated agents (a.k.a. "robots.")

An experience of identity management federation within Storage Clouds is provided in the FP7 European Project VISION Cloud [9] where reference architecture of dynamic federation among storage suppliers is presented. Likewise, agent-based approaches like InterCloud [22] also needed to solve the problem of delegation and distributed trust and security federations. Before that (and before OAuth2), GEYSERS chose to use access tokens (carried in SAML) to manage the delegated rights [20].

Additionally, the authors of [10] provide description of an enhanced "Message Oriented Middleware for Cloud computing" called Security-Enhanced CLEVER, based on the well-known XMPP protocol (originally Jabber).

III. ARCHITECTURE

An early version of the EUDAT AAI was developed by the EU-funded Contrail [13] project – several technologies were examined or trialed, but Contrail came closest to meeting the major goals:

- User Security: The AAI architecture aims to provide transparent registration and account management of the user's federated identity, including allowing the user to present different types of external identities (SAML, X.509, etc...), if they have them;

- Service Security: Establish trust relationship between the services, thus enabling secure inter-services communication, and enable users to trust the infrastructure;
- Traceability: have traceable communication between the infrastructure services – in particular, that users can be traced to their real-life identity if they misuse the system (or their credential is stolen);
- Usability: whenever possible, integrate with the communities' existing identity management system.

These security goals are supported by three main building blocks: Federated Authentication, Public Key Infrastructure (PKI), and Federated Identity Management (FIM):

- Federated Authentication: Referring to *external* federations for AAI, we support authentication with multiple types of IdPs, such as, X.509, OpenID Connect [14] (e.g. Google, Facebook, etc...) and SAML-based IdPs (Shibboleth, SimpleSAMLPhp, etc...);
- Federated Identity Management (FIM): Allowing communities' users to register without having a separate EUDAT or service specific identity, yet make use of the available policies in existing federations;
- Public Key Infrastructure (PKI): Within the EUDAT federation, offer an online Certification Authority (CA) that issues Short-Lived X.509 credentials with embedded authorisation attributes in the form of a SAML assertion, as mentioned above. This is also useful for authenticating with non-browser based applications and authorizing based on the attributes embedded within it (for command line access users will need to download the credential–as command line is mostly used by technically expert users, this is considered acceptable.[2]) Also the infrastructure hosts and services need certificates; these use the same certificates as grids, and are of course not short-lived.

Figure 1 shows the components of the EUDAT AAI. On the left are the core components controlling and managing access to the services. In the middle, EUDAT communities' IdPs and community portals; finally, on the right, the EUDAT infrastructure services (section III-E.)

A. Unity

Unity [15] is a group management and identity provisioning software developed at ICM (icm.edu.pl)and supported by PLGrid. One of the most interesting parts for EUDAT is the support for most of the required authentication protocols, bridging the protocols supported by upstream (or external) identity providers, translate and harmonise the user information. In particular, from EUDAT perspective, Unity:

- bridges OpenID and SAML based identity providers: SAML is used by the linguists, OpenID by the climate

---

[1]The exception being WS-PassiveFederation, also a part of WS-Federation.

[2]Note that we did not use the ECP profile of SAML to manage command line access as our external IdPs are not exclusively SAML-based and most of the ones that are do not support ECP.
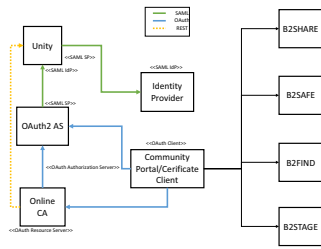
Figure 1. EUDAT federated authentication architecture. B2* are EUDAT services.

modellers[3]

- translates user attributes from external identity providers to EUDAT federation level representation. Like Contrail, it enables the publication of federation-level attributes (community membership, roles, etc.), but provides better features for managing them.
- supports registration of new users, while allowing administrators to design and invoke registration forms automatically. This is crucial and should be as lightweight as possible, as the new EUDAT users do not want to go through a formal registration process in order to become EUDAT member (and may not need it either, in a multi-LoA federation), yet users may have to be asked to provide additional attributes or requests.
- uses internally an API for authentication, so can be extended to other authentication methods via plugins.
- provides other features which can be vital for the EUDAT infrastructure are: high availability, backup and restore, attribute rule processing, and contextualised (for normal or privileged) user interfaces.

B. OAuth 2.0 Authorisation Server (AS)

OAuth [12] is an open standard that enables users ("resource owner" in OAuth) to delegate resource access to third party applications ("Clients"). The protocol itself does not include any authentication protocol or message level security; therefore it is up to the infrastructure to define suitable authentication and cryptographic methods. In EU-DAT, the AS provides SAML Service Provider (SP) end-points to authenticate the users, so must connect to SAML Identity Providers. Like the Contrail federation code before it, Unity's acts here as a SAML Identity Provider to the

---

[3]As it happens, Unity does not support OpenID, so we added another bridge, the SimpleSAMLPhP-based bridge developed by Contrail which consumes OpenID credentials and produces SAML. This is an interim solution as thee climate modellers are expected to switch to OpenID Connect eventually.

authorisation server, regardless of how the user authenticated to Unity. In other words, users authenticate with whichever external credential they have, and an intermediate SAML assertion is used to present a *harmonised* credential to the AS; in turn, the AS issues an *access token* to the portal to enable the portal to obtain an X.509 credential on behalf of the user: the EUDAT Online Certification Authority is implemented as an OAuth "resource" (described in the following section.) This workflow is illustrated in Figures 2 and 3, below. The whole process of authentication and soliciting an access token follows *Authorisation Code* grant flow from OAuth 2.0 specification.

C. Online Certification Authority (CA)

The online CA is a lightweight web service based CA that issues a short-lived X.509 credential to the federation users. As mentioned, it is needed to generate a single federated credential which works also with non-browser clients and non-web services. The CA queries the federation database for user attributes and embeds them in the certificate in the form of SAML assertion (as an extension). Attributes are thus *pushed* to the services with the certificate. EUDAT builds on the Contrail work and uses an OAuth-protected web services interface (in OAuth-speak, the CA is a Resource Server): a portal or other client would be registered beforehand, and the CA checks the access token against the requested name in the certificate (via the OAuth AS) before issuing the X.509 certificate credential, to ensure that a certificate is issued not just to an authorised client (the portal), but also with the federation name of the authorised user (from whom the authority to delegate originates). The X.509 credential has a limited lifetime to eliminate the need for revocation; it is thus necessary for the client to re-request the certificate after (or, more likely, prior to) its expiration.

D. Community Portal / Certificate Client

Each EUDAT portal is both a certificate client (meaning it obtains a certificate) and an OAuth client: it generates the key pair (so the private key is not transmitted across the wire), fetches an access token from the authorisation server upon user's consent, while authenticating the user with their identity provider. The short-lived access token is then sent to the online CA server to acquire a valid short-lived X.509 certificate–a delegated credential.

The community portal is not necessarily the same as the EUDAT portal. User communities tend to have their own portals already; and these can be integrated with a certificate/OAuth client to manage certificates with which they access EUDAT. One of the lessons of EUDAT, however, is that we underestimated the resources required of the communities to perform this integration. So we have developed "hybrid" portals where EUDAT provides parts of the portal functionality and the community provides the other part: in principle these part need come together only in
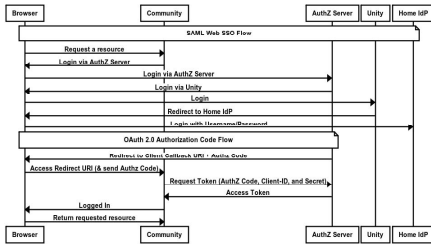
Figure 2.   User authentication endorsing SAML Web Single-Sign-On

Figure 3.   SAML to X509 Conversion

the user's browser. Single sign on ensures that both parts of the portal see the same identity[4]. The EUDAT part provides additional features, such as control of EUDAT services, and the ability to download the credential for command line access. More work is needed to scope the effort required for a full integration with the community portals of not just of the AAI but also the EUDAT services.

E. EUDAT services

There are a number of services within EUDAT, not only the core data services such as B2SAFE, B2Share, B2Find, B2Stage, etc..., but also the services supporting the EUDAT infrastructure operations such as helpdesk, wiki, and registry. Like the community portal, services need integrating with the federated credential management as well. One of the advantages of X.509 is that everything supports it, but full integration still requires work, *e.g.* to implement the required access control. The scalability requirements of EUDAT (in terms of numbers of users) require that we move from identity based account mapping in (for example) B2SAFE to role or community based authorisation, with fine-grained access control. Doing this in a generic way, so it is not specific to EUDAT, and the patches can be contributed back upstream to the developers, is one of the challenges.

F. Service interactions

The authentication perspective of the EUDAT AAI is leveraging the SAML *bridge* concept. Such a bridge consumes external identities, and acts itself as an identity provider to services within the federation. In Contrail, SimpleSAMLPhP [16] was used as such a bridge; in EUDAT this role has been taken over by Unity. Figure 2 depicts the authentication of user, invoked here with a user's request for a web resource, which also requires an "Authorisation Code" grant from an OAuth 2.0 authorisation server (AS). Initially, the user has to go through a login process while following the steps defined by SAML 2 Web Single-Sign-On [17],
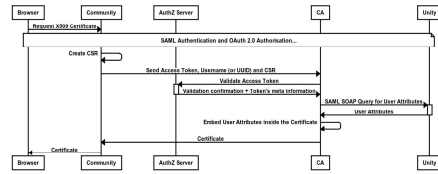
authenticating first via Unity which in turn redirects to the external IdP (unless a previous authentication is cached) After successful authentication, the AS issues the token to the client. This method implements coarse-grained authorisation – indeed, a very useful Contrail extension had users decide beforehand which OAuth clients they trust, as well as inspect the actual delegations made. Thus, if all EUDAT resources were web-based and followed the Web SSO profile, and if coarse grained authentication were sufficient, we would not need the delegated X.509 credentials at all.

Figure 3 depicts the flow of converting the SAML authentication assertion issued from user's home IdP to an X509 certificate, with the interaction of the community (or EUDAT) portal, the OAuth AS, the CA, and, eventually, the EUDAT service. In this sequence, the portal requests a short-lived X.509 credential from the CA server. The requests comprises the valid access token, an X.509 Certificate Signing Request (CSR) generated by the portal, and corresponding user id. When the CA/resource server receives the request, it validates the access token (which, in standard OAuth is opaque); once it is happy, it calls the Unity (database) and to obtain user attributes to embed inside the certificate. Finally, the certificate is returned to the portal.

IV. EXAMPLE USE CASE: "DATA STAGING" WITH FEDERATED IDENTITIES

"Data staging" is EUDAT-speak for moving data, *e.g.* between the compute resource, a data infrastructure (such as EUDAT), and/or the user's personal computer. Within EUDAT itself, data is also replicated, and its metadata is stored to enable discovery.

As an example, we look at moving data from EUDAT to, or from, another e-Infrastructure, PRACE [19]. It is assumed here that the user has accounts in both. The challenge arises because the user does not necessarily use the same IdP to authenticate to both EUDAT and PRACE, and even if they did, the two infrastructures would not necessarily trust each other's credentials[5].

EUDAT normally moves data with GridFTP. In order to move data into, or out of, PRACE, PRACE must publish

---

[4]Except with `eduPersonTargetedId` where the IdP deliberately generates distinct "identities" for the user for each SP (portal).

[5]PRACE also uses X.509 certificates but does not at present trust the EUDAT federation CA.

a GridFTP endpoint. In the general case, we have several options (listed here in order of preference):

- PRACE is asked to also trusts EUDAT's federation-CA (presumably after a review process, and possibly at a certain LoA only), and EUDAT user identities (Distinguished Names, or DNs) are mapped by the external GridFTP server into whichever is the user's correct account on PRACE (this, in turn, may require additional account mapping services, most likely by users requesting that their accounts be linked.)

- PRACE does not trust the EUDAT CA, but issues its own certificates to the users. In this case, the user (or GridFTP client acting on behalf of the user) will need two (possibly delegated) X.509 (or GSI proxy [21]) credentials, one to access the EUDAT endpoint and another to access the PRACE endpoint. Using data channel authentication (DCAU), authentication is established for the session between the endpoints (they will also have to trust each other's host certificates.)

- The file is not moved with user credentials; instead, an automated client with a host or robot certificate moves the file and does the equivalent of `chown`, *i.e.*, locally reassigns ownership to the correct owner (whose identity it will need communicated to it.) In the simpler cases, there would be one such data mover "robot" per community.

- As a final option, one might eschew the use of X.509 certificates for the client, and instead hook GridFTP into the delegation mechanism (or something equivalent). While this should be technically possible, it would not only require some customisation of the GridFTP server, it would likely require the customisation on *both ends*. As EUDAT does not control the software in PRACE, requiring modifications to their software in order to interoperate with EUDAT is not possible.

Note that in all of these cases, both endpoint hosts will have IGTF certificates (www.igtf.net), so the security of the endpoints on either side is established.

In order to keep things simple, we have, for now, chosen the approach of using the community-based data mover. Indeed, the plan for B2SAFE, based currently on iRODS, also called for a data mover credential which is able to *replicate* data across the EUDAT federation, so the "robot-based" approach would be needed anyway. However, in future work we will need to return to the other options, in order to ensure finer grained access control also on the PRACE side.

A. Delegation

In the usual grid contexts, delegation is often based on GSI proxies [21]. In fact, we could have used GSI, as GridFTP supports GSI, as well as iRODS and other services currently used in EUDAT. For the purposes of future

extensions, we chose to follow Contrail and use OAuth to delegate the certificates, since, as we mentioned above, it gives more control over the what is delegated to whom, at the cost of having to run an OAuth AS with high availability – and it would also work with (future) services which do not support GSI.

In the context of the staging use case, using our OAuth-to-X.509 delegation (Fig. 3) rather than a data mover robot would require that the data mover service be registered as a client with the OAuth AS. Once it builds its proxy chain, the authorisation credentials (the embedded SAML extension) will reside in the proxy chain and will still need to be enforced. Firstly, this means that the credential will have to be extracted on the remote side–PRACE in our example–from the location in the chain: EUDAT is working with the Open Grid Forum VOMSPROC working group on documenting the requirements in this step, as well as with the IDEL group on documenting the delegation itself. Secondly, the authorisation will have to be enforced also by PRACE, which would potentially mean calling back to EUDAT to have the SAML assertion checked with a EUDAT federation PDP[6]. But then, as the target is always data, and data access tends to require fine grained access control (at least for some communities), EUDAT and PRACE must have consistent naming not just for the user identity/roles/communities, but also for the data itself: this, too, is a problem we have yet to solve.

V. AAI AS A FRAMEWORK

As mentioned earlier, the AAI is based on using what the communities use already, whenever possible. Promoting the use of standards, and interoperable components implementing these standards, is the first step in providing a framework upon which EUDAT can integrate new services and new communities; this step also helps promote reuse of all or parts in other projects (as EUDAT itself has reused), which in turn could lead to better sustainability models for everybody, as more projects pick up and use the same components.

A true framework would need to be sufficiently powerful to implement what communities require, yet to be sufficiently loosely coupled that components can be replaced or used independently. Beyond the technical interoperation, as we saw in section IV-A, there is a need for interoperation at a semantic level. Not just with a harmonised credential, and with agreed levels of assurance, but also an agreement as to the meaning of attributes (such as roles) and the common enforcement of access control policies. Beyond this, again, we need to look at the harmonisation of policies in existing federations, and plugging the gaps where they do not align. Experiences from the grid world have shown this to be a time consuming task, and much can be achieved with a basic

---

[6]Policy Decision Point, a part of an XACML architecture.

set of levels of assurance Nevertheless, the more extensive work would require a deeper understanding of the elements of policies, again work started in the grid world, and could also be pursued with the emerging interfederation activities such as eduGain and FIM4R.

VI. Conclusion

EUDAT is a data e-Infrastructure, offering users in research communities the ability to upload data, to have it replicated and shared, and further to make it available on other selected infrastructures. In this paper, we have described the role of the federated AAI, how it evolved out of the outcome of the Contrail project and further incorporated the Unity identity manager. From a technology perspective, the work covers a range of diverse technologies– SAML, OpenID, OpenID Connect, and X.509 for authentication, OAuth for delegation, SAML and X.509 as internal credentials. Yet, the technologies are combined in perfectly standard ways, like Lego bricks, one might say; and they are combined in an attempt to pragmatically solve problems: making use of existing identity federations, accessing both web- and non-web services, offering both portal and command line access, supporting delegation. As it is, we have managed to make it work and capable of offering some interesting security features, even if not everything is in place yet: users can control trusted clients and track delegations, both coarse grained (via the OAuth AS) and fine grained (via SAML push attributes) are available, and we have support for delegation of credentials beyond just "cloning" it. From a practical perspective, we noted that the effort of integration with the communities–who mostly already have their portals and established practices–needs effort on the community side; and while X.509 works with practically everything, further work on services would be needed to make use of scalability (specifically, to modify services to not use identity-based accounts.)

Acknowledgment

References

[1] Broeder Daan, Jones Bob, Kelsey David, Kershaw Philip, Lders Stefan, Lyall Andrew, Nyrnen Tommi, Wartel Romain, Weyer Heinz J, *Federated Identity Management for Research Collaborations,* CERN-OPEN-2012-006, https://cdsweb.cern.ch/record/1442597

[2] Shim, S.S.Y.; Geetanjali Bhalla; Vishnu Pendyala; *Federated identity management,* Computer , vol.38, no.12, pp. 120- 122, 2005.

[3] EUDAT, http://www.eudat.eu

[4] CILogon, http://ca.cilogon.org/

[5] Terena Certificate Service, http://www.terena.org/activities/tcs/

[6] A Nadalin and M Goodner and M Gudgin and A Barbir and H Granqvist *(eds)*: *WS-Trust 1.4*, OASIS Standard, February 2009.

[7] A Nadalin and C Kaler and P Hallam-Baker and R Monzillo *(eds.)*: *Web Services Security: SOAP Message Security 1.0*, OASIS Standard 200401 (March 2004).

[8] M Goodner and A Nadalin *(eds.)*: *Web Services Federation Language 1.2*, OASIS Standard 200905 (May 2009).

[9] A Celesti and F Tusa and M Villari and A Puliafito *(eds.)*: *How To Federate Vision Clouds Through Saml/shibboleth Authentication*, Lecture Notes in Computer Science, Springer vol.7592/2012, pp.259-274, 2012

[10] A Celesti and M Fazio and M Villari *(eds.)*: *SE CLEVER: A secure message oriented Middleware for Cloud federation*, IEEE Symposium on Computers and Communications (ISCC), pp.35-40, 7-10 July 2013.

[11] G Larocca and S Monforte and D Scardaci: *Catania Science Gateway eTokenServer*, http://sourceforge.net/p/ctsciencegtwys/wiki/InstalleTokenServ

[12] The OAuth 2.0 Authorization Framework, http://tools.ietf.org/html/rfc6749

[13] Contrail Project, http://www.contrail-project.eu.

[14] N Sakimura and J Bradley and M Jones and B de Medeiros and C Mortimore, *OpenID Connect Core 1.0*, http://openid.net/specs/openid-connect-core-1_0.html, February 25, 2014

[15] Unity, http://www.unity-idm.eu.

[16] https://simplesamlphp.org/

[17] J Hughes and S Cantor and J Hodges and F Hirsch and P Mishra R Philpott and E Maler *(eds.)*: *Profiles for the OASIS Security Assertion Markup Language (SAML) V2.0*, OASIS Standard 200503 (March 2005).

[18] Argus Authorization Service, https://twiki.cern.ch/twiki/bin/view/EGEE/AuthorizationFramework

[19] Prace Research Infrastructure, http://www.prace-ri.eu/

[20] Canh Ngo and Y Demchenko and C de Laat: *Toward a Dynamic Trust Establishment approach for multi-provider Intercloud environment*, Proc. 4th Int'l Conf. on Cloud Computing, Technology and Science (IEEE CloudCom), 2012, pp. 532-538. doi:10.1109/CloudCom.2012.6427548

[21] S Tuecke and V Welch and D Engert and L Pearlman and M Thompson: *Internet X.509 Public Key Infrastructure (PKI) Proxy Certificate Delegation Profile*, RFC 3820, http://www.rfc-editor.org/rfc/rfc3820.txt

[22] Kwang Mong Sim: *Cloud Intelligence: agents within an InterCloud*, doi:10.2417/3201311.005153.

Paper III

Combining the X.509 and the SAML Federated Identity Management Systems

M. Hardt, A. Hayrapetyan, P. Millar and A.S. Memon. 2014.

As a part of the Helmholtz foundation's Large Scale Data Management and Analysis (LSDMA) project, Shiraz Memon was actively involved in the design and development of the Authentication and Authorisation Infrastructure (AAI) in the Helmholtz federation. Shiraz Memon is one of the main authors of this publication and provided the content concerning the proxy model and credential translation in the publication.

Combining the X.509 and the SAML Federated Identity Management Systems

Marcus Hardt[1], Arsen Hayrapetyan[1], Paul Millar[2], and Shiraz Memon[3]

[1] Steinbuch Centre for Computing, Karlsruhe Institute of Technology, Germany
`firstname.lastname@kit.edu`
[2] Deutsches Elektronen Synchrotron, Germany
`p.millar@desy.de`
[3] Jülich Supercomputing Centre, Germany
`a.memon@fz-juelich.de`

Abstract. Every distributed computing infrastructure requires authentication and authorisation infrastructures (AAI) to manage access to resources and content. Several of such so called AAI systems are in use within different groups of users. In the Large Scale Data Management and Analysis project we aim to support and bring together many user communities. We therefore need to harmonise the currently used AAI systems. The approach described is to translate between different authentication systems. We furthermore try to maintain the same trust level wherever possible, and to harmonise authorisation across the involved systems.

1   Introduction

Various scientific communites are developing new techniques and equipment for collecting data at increasing rates. The resulting data-deluge is challenging their ability to manage this data. Also, these communities are geographically distributed. Therefore, existing approaches to data management and access control seize to function well. One of the key questions faced by new approaches is that of how to handle authentication and authorisation in a federated environment. Traditional Authentication and Authorisation Infrastructures (AAI) are based around centrally managed user accounts and groups. This does not scale to the volume and flux of both users and data. Therefore, so called federated identity management approaches are required.

If each of the participating communities independently discovered their own approaches for handling access to its data, that would result in multiplied and therefore wasted efforts. The German Helmholtz Association has therefore funded the Large Scale Data Management and Analysis (LSDMA) portfolio extension, charged to stimulate and drive innovation within the academic sphere and targeted at improving the ways how data are stored, managed and analysed.

The LSDMA project gathers 25 different user communities grouped by their scientific domain. This covers use-cases from electro mobility, battery testing and simulation, climate modelling, human brain image analysis, selective plane

microscopy, synchrotron tomography, high energy physics and more. Common requirements include performant, safe and simple access to data, capabilities to search within the data, archival and analysis of data.

As some of these user communities already have existing AAI systems, the LSDMA project does not attempt to change the ways users are working today. Instead, the goal is to provide a flexible service that can bridge between the existing AAI Infrastructures. Furthermore, many services support only one AAI system. Hence, the ability to bridge between technologies yields at making services that only support one AAI available to users from another AAI system. In addition to the technical challenges of authentication and authorisation, the level of trust needs be maintained wherever possible.

The relevant federated identity management technologies in this field are X.509 and SAML, both token based. Outside the scientific sector, OAuth and OpenID are widely used. While the focus of this paper is on the first two, the latter two can be integrated into the presented approach, because they are also token based. Plain username / password systems are not considered in this paper.

For the sake of clarity, this paper focuses on a hypothetical but practical use-case, which we will use for illustrating the presented solution details. This use case describes a web-portal that provides a visualisation service. The data for visualisation is stored on a third-party storage. The user authenticates to the web-portal with their SAML credentials while the storage service requires X.509 certificates for providing access to the data. Of course, both services could provide alternative authentication methods. However, similar requirements have emerged in the requirement analysis of the LSDMA project. To support this use-case, the authentication token of the user must be translated to one supported by the service. One notable impact of this design will be that distinct scientific communities will be able to access each others services without neither changing their existing tools nor services.

In order to be useful in practice, a solution that combines different AAIs, has to be able to fulfil the following requirements:

1. Hide its complexity from the user
2. Work well on commandline and in a browser
3. Support interactive and non-interactive authentication
4. Support delegation of data access rights to a web portal
5. Support generic delegation of data access rights to any host
6. Translate between the federated identity management tokens
7. Provide support for the mapping to the respective authorisation decisions

The remainder of this paper is organised as follows: In section 2 we give an overview of existing relevant technologies. In section 3 our approach for the integrative architecture is presented. Section 4 gives the conclusion and identifies the future work.

406   M. Hardt et al.

2   State of the Art

A variety of technologies exist that relate to federated authentication and authorisation standards. In this chapter we present those of them which are relevant to our work. The list, however, is not meant to be exhaustive.

2.1   X.509 Based Authentication and Authorisation

The X.509[14] based authentication and authorisation standard is widely used by scientific grid communities. Examples include Worldwide LHC Computing Grid and Open Science Grid. It is used, as well, by numerous web applications, for example, for secure online banking. In this section we are focusing on the grid communities, because they are more relevant to the LSDMA project.

The X.509 based authentication in the grid makes use of several components described briefly below.

The communicating end entities (EEs: users, hosts, services) in the grid, along with one or more Certification Authorities (CAs), are constituents of a Public Key Infrastructure (PKI). The CAs, also known as trusted third parties (TTP), provide EEs with digitally signed X.509 certificates containing their identity information. These certificates are used to authenticate the EEs. The grid CAs are established commonly on a per-country basis. The grid collaborations are geographically distributed over many institutes in different countries. Thus, the structure of a grid collaboration imposes the requirement of trust of the "foreign" CAs by all EEs. This problem is not scalable per se. To solve the pairwise-trust problem, the International Grid Trust Federation (IGTF)[17] has been established. This group maintains a set of minimum requirements for its member CAs (e.g., that a user's identity must be manually vetted using a passport or an equivalent government-issued document before issuing a user certificate). IGTF verifies routinely the CAs' conformance to the requirements. The resource sharing sites in the grid collaboration install the veirified IGTF CA certificate bundle on their resources, thus enabling mutual authentication between EEs on the collaboration scale.

Several CAs in IGTF provide Short-Lived Credential Services (SLCS) for their users. SLCS allows users to get X.509 certificates online without getting through the thorough identity vetting procedure (involving face-to-face meeting) for the standard X.509 certificates. To compensate this "lightweight" identity vetting procedure and improve the security, the SLCS certificates have much shorter maximum validity period than the standard ones (one million seconds versus 13 months). After the old SLCS certificate has expired users can easily request a new one. In Germany, the IGTF-accredited SLCS certificates are provided by DFN SLCS CA.

The X.509 based authorisation in the grid requires the authentication information of an EE and additional data about EE privileges to access shared resources of the collaboration. We describe the relevant components briefly below.

The scientific grid communities contribute their computational and storage resources to Virtual Organisations which are dedicated to specific scientific research. The EEs register with the VO and get assigned specific groups and roles according to the work they perform within their VO. The group and role information is used later by the authorisation services to grant or deny access to the resource. The VO membership, group and role information is managed by the VO via a dedicated central service.

Single sign-on and identity delegation are supported in grid VOs through proxy certificates. A proxy certificate [11] is digitally signed by a user delegating her identity for various grid tasks, for example, accessing files on a remote storage. For improved security each identity delegation is bound to certain public key. Furthermore, the proxy lifetime is limited (the default being 12 hours).

There are many *implementations* for components involved in X.509-based authentication and authorisation in the grid. We list few of them below.

Tools for managing X.509 certificates vary between command line (e.g. OpenSSL [4], gLite UI) and the web interfaces provided by CAs (e.g. German Grid CA web interface).

The SLCS CA front-end is provided by GridShib, a project based on the Shibboleth SP component. Software for CA management and an online interface for web-based certificate requests is usually custom-made by CA. Projects providing such software include OpenCA and EJBCA.

Two standards for grid VO membership management are VOMS[5] provided by gLite software and Unity (formerly known as UVOS [6]) provided by Unicore. Both support grouping of user accounts and assigning specific roles to them.

Proxy certificate management is provided by every grid middleware package. Examples include Globus, gLite, Unicore.

Despite the technical advantages of X.509, many users are uncomfortable with using certificate-based authentication. Reasons include an IGTF-requirement to update certificates annually, the lack of built-in web browser support for proxy certificates, and the fact that grid tools have distinct trust stores from webbrowsers. Since many IGTF CAs use web portals for user interaction, this last item requires users to go through a convoluted export procedure before they may use their certificate for authentication in a grid context.

Regarding the requirements we have identified in section 1, X.509 does support [2, 3, 5]

2.2   SAML Based Authentication and Authorisation

The SAML based authentication and authorisation are used by numerous scientific and non-scientific communities. Examples include the AAI of the German Research Network (Deutsches Forschungsnetz, DFN [2]) and the AAI of the Swiss Research and Education Network (SWITCH).

The Security Assertion Markup Language (SAML) [8] was developed later than X.509. It is an XML-based open standard for specifying authentication and authorisation data. These data are expressed in the form of SAML assertions that

are typically exchanged between Identity Provider (IdP) and Service Provider (SP).

Although there are several profiles for *SAML based authentication*, the most commonly used is the Web Browser Single Sign-On (Web-SSO) profile. In this profile a principal (i.e. a user) wishes to authenticate to a Service Provider (SP) using a web-browser. The SP makes use of web based redirects, so that the user can authenticate himself to the Identity Provider (IdP) of his home institution. After successful authentication, the IdP uses another web redirect, so that the web browser delivers the SAML assertion to the SP. These redirects allow passing information between IdP and SP via the users' browser, therefore no direct connection between IdP and SP is required for authentication.

The Web-SSO profile assumes the use of a web browser as a user agent which makes it not suitable for non-browser applications like desktop applications or server-side code running as a web application. A companion SAML profile known as Enhanced Client or Proxy (ECP) profile is available that removes the limitations of Web-SSO profile designed around limitations of the web browser.

To manage trust relationship between IdPs and SPs, SAML based federations can be formed. The federations set policies which the members of federations adhere to. They also vet their member IdPs and SPs and maintain a list of their members. The SPs can therefore rely on the federation policy to expect the minimum set of information being released by the IdPs to them. On the other hand, the IdPs can rely on federation policy to expect that the released data will be used appropriately. This simplifies the trust relationships, since instead of having multiple bilateral agreements with IdPs, the SPs can only have one agreement with the federation.

In many cases the national research network providers (e.g. DFN in Germany) operate such trust federations. Depending on the level of trust (e.g. the quality level of the user-ID-vetting, information expiry, etc.) different federations may be formed. In case of Germany there are three federations: DFN-advanced, DFN-basic and DFN-test, the first of which has comparable requirements as imposed by the IGTF policies in the X.509 domain.

To enable trustworthy authentication and authorisation, information exchange and to share resources on larger scale, SAML based federations can interfederate. When two federations interfederate, they agree to trust the credentials of each other's member IdPs and SPs. One of the most prominent intiatives in this area is eduGAIN project [9].

The *SAML based authorisation* by the SP is in most cases based on the authentication data provided by the IdP in the SAML assertion. In addition, the users from different IdPs can be grouped according to certain criteria. The group information can be taken into account by the SPs when making authorisation decisions.

SAML does support delegation, but not in the generic manner as X.509 does. It typically involves a lot of overhead, because all involved IdPs and SPs have to authorise delegation. One typical use-case, however, is an exception. This is

the portal delegation, which may be used, in case a web-portal needs to access external information on behalf of the user.

SAML based authentication and authorisation solutions are *implemented* by serveral products. We list some of them below.

The most known SAML implementation is Shibboleth [3], whose developers claim they have the world's most widely deployed federated identity solution. There are also many other Open Source implementations. Examples include OpenSAML, simpleSAML-php, ZXID, Lasso and OpenSSO. SAML based group management is provided by a software called GMT, developed at SWITCH.

Although this paper focuses on SAML Web-SSO, it is worth mentioning that projects exist that aim to add SAML authentication to GSS- and SASL- authentication frameworks. Since many common Internet protocols support either GSS- or SASL- authentication, such approaches (if successful) will bring SAML-based authentication to the majority of non-web applications. There are currently two major approaches: Application Bridging for Federated Access Beyond Web (AB-FAB) [13] implemented by Project Moonshot and ECP-over-GSS [7].

In terms of the identified requirements SAML does support [1, 2, 4]

2.3   Credential Translation

Credential translation refers to a process of generating authentication or authorisation tokens or credentials for a given AAI service based on the authentication with other types of credentials. In this paper we are particularly interested in the token based credential translation from SAML assertions to X.509 certificates or proxies. The credential translation service can be a web service which requires SAML authentication and uses the resulting SAML assertion to request the X.509 certificate on users's behalf. The online CA issuing the X.509 certificate is an SP configured to accept SAML assertions delegated by the user to the credential tranlsation service. The certificates issued by the online CA can be short- or long-lived depending on the trust between the CA and the IdP authenticating the user.

The implementations of credential translation services from SAML to X.509 tokens include gridcertlib [15], a java library developed by SWITCH as well as Security Token Service (STS) developed within the EMI project. The SAML assertion delegation is supported by the GridShib [16] software which makes it suitable as a base for custom-developed credential translation web services.

IGTF-accredited SLCS certificates are issued by several online CAs based on SAML authentication. DFN SLCS CA and SWITCH SLCS CA issue short-lived X.509 certificates while the Terena credential service (TCS) issues long-lived ones.

2.4   Other Federated Authentication and Authorisation Technologies

There is a number of federated authentication and authorisation technologies and standards other than X.509 and SAML based standards. In this section we describe two of the most prominent and widely used ones.

410   M. Hardt et al.

OpenID is an open standard for federated authentication. It offers user-centric authentication mechanism allowing the user to choose the OpenID Provider (i.e. identity provider) for asserting her identity for the relying party (e.g. service provider). Most importantly, the standard does not require the trust relationship to be established between relying party and OpenID Provider in advance. The consumers can work with all possible OpenID providers. This makes it useful for the resource providers who are interested in offering users convenient and fast access to their resources. On the other hand, the relying party should not rely on the OpenID Provider for the trustfullness of the identity information about the user. Thus, the OpenID standard can be the choice in the cases when the trust requirement with respect to identity vetting and end-users' privacy are not the the primary concern of the AAI.

Many content management systems and web-based services provide plugins for OpenID support. It can be enabled also via libraries implemented in many languages.

There are several projects and standards aimed at improving the security level provided by the OpenID protocol. Examples include integration of the OpenID protocol into the SAML IdP (simpleSAMLphp) and integrating OpenID with OAuth (OpenID Connect standard, draft).

OAuth v2.0 is an open standard for authorisation. It allows third-party applications to get access to a web resource with the approval of the resource owner. In the most common OAuth scenario the service provider accepts a third-party Client application to access the data owned by the user based on an access token issued by the Authorisation Server. The access token contains the user's identity information which can be released to the authenticated Client if the user approves it. The user approves the release of her personal data by authenticating to the Authorisation Server. An important implication of this protocol is that the users never share their credentials with their Clients when delegating them the task of accessing their resources.

The implementations of OAuth components are available as libraries for various languages including Java and Python. There are standardisation efforts aimed at integrating OAuth authorisation standard with SAML 2.0 authentication and OpenID authentication (OpenID Connect draft).

2.5   Authorization and Group Management

The ability to define and manage groups is not directly associated with federated identity management systems. However, often the membership in a group is used by a service to make the authorisation decision. Therefore, authentication, group definition and authorisation are closely related.

The two largest computing middlewares globus/gLite and Unicore both provide support for Virtual Organisations or VOs [12]. These VOs are used to form groups of users on side and to allocate hardware resources for VOs on the other side.

VOMS Within WLCG, group membership is managed and asserted through one or more Virtual Organisation Membership Service (VOMS). This service maintains a database of users and their group membership. It also provides a web interface for administrators to add and remove users from groups. When generating the grid-proxy-credential, a user may request an attribute certificate[11] from one or more VOMS server. The supplied attribute certificates are embedded within the proxy certificate so that, when authenticating with a remote service, the remote service is able to extract the attribute certificate. Group membership is then discovered, provided the service trusts the VOMS server that issued the attribute certificate.

Unity SAML also describes how an SP may query an IdP directly, requesting assertions about some particular user. This allows the SP to gather additional assertions about the user from third-party IdPs after the user has delivered an assertion through Web Browser SSO; for example, to query the group membership of this user. Such third-party group-membership services are broadly similar to VOMS servers; however, in contrast to VOMS, SAML group-membership is asserted when the user authenticates with the SP. All the attributes may be employed by the SPs when making an authorisation decision.

The Unicore [10,1] grid middleware comes with the Unicore VO Service, UVOS, now being renamed to Unity. It implements the VOMS concept based on the SAML and XACML standards. For this Unity and Unicore make use of attribute aggregation. This is a SAML technique in which a Service Provider (SP) aggregates a users attributes by querying several attribute services – such as Unity.

Extensions include (among others) support for a hierarchical VO structure and a pluggable support for additional interfaces, so that VOMS style attribute certificates can be generated. Despite the extensibility of Unity, both VO systems are different in structure and not currently able to exchange group definitions among each other.

3 Design and Integrative Architecture

In this section we present the architecture of the AAI for the LSDMA project. The AAI is designed to support different authentication scenarios involving SAML and X.509 credentials. Standard software components of the existing implementations will be used for AAI components (see top of fig. 1). For SAML-based federations these are: Shibboleth Identity Provider (IdP) and Service Provider (SP). For X.509-based PKI these are: Certification Authority (CA) and Virtual Organisation Management Service (VOMS). The translation from SAML to X.509 credentials will be enabled by a component generating an asymmetric key pair and requesting an X.509 certificate from the CA upon user's successful authentication at the IdP. With respect to our initially described hypothetical use-case, three relatively simple authentication scenarios can easily be supported by standard AAI setups as described on Figure 2.
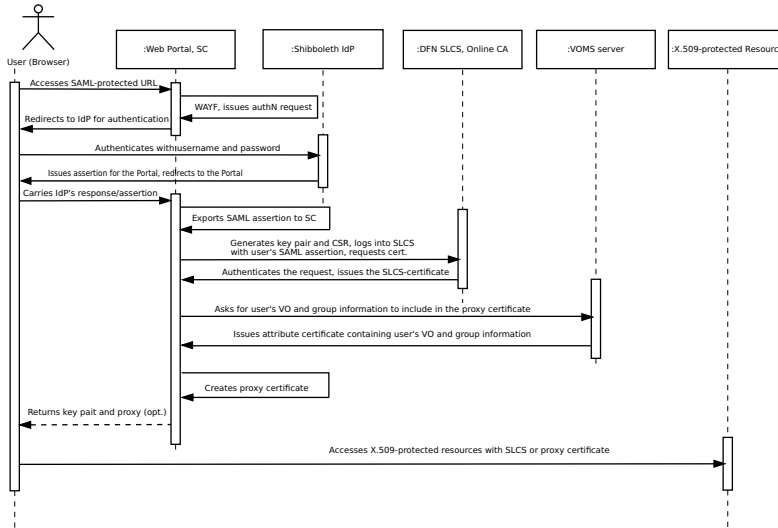
Fig. 1. The components of the LSDMA AAI architecture

In the rest of this section we will focus on the complex case in which credential translation from SAML to X.509 formats is required. We will assume that the access to the visualisation data requires a valid proxy certificate, while the user authenticates with username and password at an IdP and uses the SAML Web-SSO profile. It is mportant to note that the user credentials have to be delegated to the Web portal running the visualisation program that uses the credentials to read and visualise the data. The corresponding sequence diagram is presented in Figure 1. In the following sections we describe the authentication process and components of the federated AAI in more detail.

3.1   SAML-Based Shibboleth Authentication

The user points her browser to the *Web portal* and requests data visualisation. The visualisation service at the web portal is protected by a *Shibboleth SP*. The SP then determines the user's home organisation and redirects her to the corresponding *IdP*. The IdP asks the user to authenticate, e.g. with username and password. Upon successful authentication the user returns to the portal carrying a valid SAML assertion.

3.2   SAML Assertion Delegation and SLCS Certificate Request

The SAML assertion is delegated by the portal to the *SLCS Client* (SC). At this point the SC generates an asymmetric key pair for the user as well as a
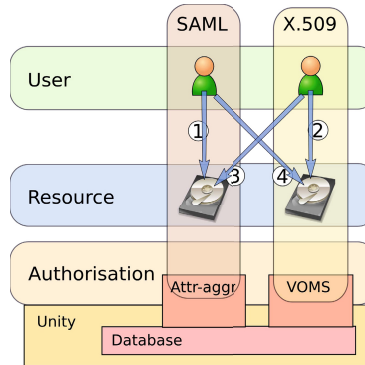
X.509 and the SAML Federated Identity Management Systems   413

Fig. 2. The possible authentication scenarios in LSDMA. The simple ones are (1), where a SAML-authenticated user accesses a SAML-protected resouce and (2) where an X.509-authenticated users access an X.509 protected resouce. Credential translation is required, in cases (3) and (4). The first case (X.509 user accesses SAML resource) can be easily accomplished by allowing X.509 at the IdP. This cannot be centrally provided, because every single IdP has to allow this. Scenario (4), SAML user authenticates to X.509 resource is handled in this paper.

certificate signing request (CSR). The fields of the Distinguished Name (DN) of the user in the CSR, such as user name or organisation, are taken from the SAML assertion provided by IdP. The SC forwards the CSR to the *DFN SLCS* along with the user's SAML assertion for authentication. The DFN SLCS itself acts as Shibboleth SP which supports SAML assertion delegation and therefore can authenticate the delegated request. In particular, it verifies that the request is coming from a trusted delegate, our Web portal, and carries a valid SAML assertion from a trusted IdP. Upon successful authentication of the request DFN SLCS CA will issue the SLCS certificate for the user which is returned to the SC and stored on the portal. This may alreaaddy suffice for the user to access the data and complete his visualisation.

3.3   VOMS Proxy Generation

As part of a more complicated scenario the user and the storage hosting the data for visualisation may be part of a Virtual Organisation (VO). The later controls the access rights on the resources based on the role a user has in the VO. In our case the visualisation application must have a proxy certificate incorporating user VO membership information in order to get access to the data. To fulfill this requirement, the SC will contact the *VOMS server* for the VO to fetch user's VO memebership information and incorporate it into the proxy.

3.4   Putting It All Together

Once the proxy is generated, the visualisation application is able to access the data on the storage on user's behalf and produce visualisation objects and send them to the user's browser.

As it can be seen from the description above, the central component which is acting as a bridge between SAML-based and X.509 based authentication realms, is the SC. For its implementation we are currently considering two possibilities: i) adaptation of the GridCertLib java library, initially developed for SWITCHaai [18], to the existing DFN SLCS; ii) our own implementation of the component based on the example implementation provided by GridShib [16] project.

Regarding the requirements formulated in the introduction, the presented solution supports [1, 2, 3, 4, 5, 6, 7] However, (2) is only supported with IdP that support ECP and (3) supports automated authentication up to 10 days after the initial SAML login was carried out.

4   Conclusions and Future Work

The diverse user communities within the LSDMA project use federated identity management solutions from the two domains of SAML and X.509, both for authentication to services and for authorisation within the service.

In this paper we have presented an integrative architecture that is capable of translating authentication tokens between both domains. We have furthermore shown that group management for authorisation decisions can be supported, too. Currently the presented example of VOMS does however neither take existing attributes of the SAML assertion into account nor would it support group definition for a SAML SP. Future work will therefore include the choice of one group definition platform. This platform needs to contain all group definitions. It will then need to be extended so that interfaces to both, SAML (via attribute-aggregation) and to X.509 (via VOMS).

One option for this is to use Unity for group definition and to extend it so that VOMS proxy certificates can be issued. In this way, SAML and X.509 secured services can base their authorisation decisions on an identical group definition. Furthermore, we foresee to explore the hierarchical group definition concept of Unity to facilitate the creation of subgroups to facilitate data sharing.

Additional technologies, predominantly used outside the scientific sector include OpenID and OAuth. The presented architecture was designed with the goal of being able to include both. The envisaged use cases include authentication to our SC using OpenID (to translate from OpenID to X.509) as well as including OAuth, so that users who have authenticated to our SC, using either SAML, X.509 or OpenID, can subsequently issue authorisations using OAuth. However, work on this is still in a very preliminary state and will be pursued in the future.

Some of the methods we describe are targeted at web-browser based activities. While web browsers are often useful for many workflows, this may not the case when working with specialised clients, such as commandline clients. With the

X.509 and the SAML Federated Identity Management Systems   415

advent of the SAML profile ECP, also commandline access will be available for SAML users.

References

1. Unicore summit (2012), `http://hdl.handle.net/2128/4705` (last visited August 26, 2013)
2. DFN. The German National Research Network Provider, `http://dfn.de` (last visited June 1, 2013)
3. Shibboleth. Project homepage, `http://shibboleth.net`
4. The OpenSSL Team. OpenSSL project homepage, `https://www.openssl.org/` (last visited October 10, 2012)
5. Alfieri, R., Cecchini, R.L., Ciaschini, V., dell'Agnello, L., Frohner, A., Gianoli, A., Lõrentey, K., Spataro, F.: VOMS, an authorization system for virtual organizations. In: Fernández Rivera, F., Bubak, M., Gómez Tato, A., Doallo, R. (eds.) Across Grids 2003. LNCS, vol. 2970, pp. 33–40. Springer, Heidelberg (2004)
6. Benedyczak, K., Biala, P.: Next generation of virtual organizations in unicore. In: Unicore Summit 2012 Proceedings (2012)
7. Cantor, S., Josefsson, S.: SAML Enhanced Client SASL and GSS-API Mechanisms. IETF Draft Document (2013), `https://datatracker.ietf.org/doc/draft-cantor-ietf-kitten-saml-ec/` (last visited November 13, 2013)
8. Cantor, S., Kemp, J., Philpott, R., Maler, E.: Assertions and protocols for the oasis security assertion markup language (SAML) v2.0 (2005)
9. eduGAIN. Project homepage, `http://edugain.org`
10. Erwin, D., Snelling, D.: UNICORE: a grid computing environment. In: Euro-Par 2001 Parallel Processing, pp. 825–834 (2001)
11. Farrell, S., Housley, R.: RFC 3281: An internet attribute certificate profile for authorization. IETF RFC, `http://www.ietf.org/rfc/rfc3281.txt`
12. Foster, I.: The anatomy of the grid: Enabling scalable virtual organizations. In: Sakellariou, R., Keane, J.A., Gurd, J.R., Freeman, L. (eds.) Euro-Par 2001. LNCS, vol. 2150, pp. 1–4. Springer, Heidelberg (2001)
13. Howlett, J., Hartman, S.: Application Bridging for Federated Access Beyond web (ABFAB). IETF Draft, `http://datatracker.ietf.org/wg/abfab/`
14. ITU-T Study Group 17: Security. In: Public-key and attribute certificate frameworks (October 2010), `http://www.itu.int/ITU-T/recommendations/rec.aspx?rec=X.509` (last visisted August 22, 2013)
15. Murri, R., Maffioletti, S., Kunszt, P., Tschopp, V.: Gridcertlib: a single sign-on solution for grid web applications and portals, `http://arxiv.org/abs/1101.4116v3`
16. The GridShib Project. Homepage, `http://gridshib.globus.org` (last visited August 26, 2013)
17. The International Grid Trust Federation, `http://www.igtf.net` (last visited June 12, 2013)
18. The Switch AAI. Homepage, `http://www.switch.ch/aai/` (last visited August 26, 2013)
19. van Wezel, J., Streit, A., Jung, C., Stotzka, R., Halstenberg, S., Rigoll, F., Garcia, A., Heiss, A., Schwarz, K., Gasthuber, M., Giesler, A.: Data life cycle labs, a new concept to support data-intensive science. arXiv e-print 1212.5596 (December 2012)

Paper IV

Implementing an Authorisation Architecture in the EUDAT Services Federation

A.S. Memon, J. Jensen, W. Elbers, M. Riedel, H. Neukirchen and M. Book. 2017.

Shiraz Memon was Authentication and Authorisation Infrastructure (AAI) Engineer in the EUDAT2020 project and substantially contributed in the analysis, design, implementation, and evaluation of the XACML-based B2ACCESS authorisation service. He is the main author of this publication and provided the majority of the content of this publication.

Implementing an Authorisation Architecture in the EUDAT Services Federation

Shiraz Memon*§, Jens Jensen‡, Willem Elbers†, Morris Riedel*§, Helmut Neukirchen§, Matthias Book§

*Jülich Supercomputing Center, Jülich Germany
{a.memon, m.riedel}@fz-juelich.de

†CLARIN ERIC, Utrecht, Netherlands
willem@clarin.eu

‡STFC, Oxford, United Kingdom
jens.jensen@stfc.ac.uk

§University of Iceland, Reykjavik, Iceland
{asm25, morris, helmut, book}@hi.is

Abstract—This paper describes the requirements and architecture of authorisation in a multi-disciplinary, multi-site, multi-stakeholder infrastructure which is using federated identity management. Stakeholders include administrators of sites, infrastructure, services, as well as data owners and community representatives. In order to be able to express and combine policies, we have based the authorisation infrastructure on XACML.

Index Terms—Federated Authorisatiom, XACML, AAI, e-infrastructure, cyberinfrastructure

I. INTRODUCTION

Authorisation is the decision that is taken by a service when a user wishes to perform an action on the service. In its simplest form[1], it has a user accessing a service based on the user's identity or rights granted by an authority (the home organisation in the RFC). In more complicated cases, rights can be *delegated*, for example from the user to a client acting on behalf of the user, or authorisation may be *fine-grained* and depend on the type of action the client wishes to perform, and on which object.

In this paper, we focus on an e-infrastructure/cyberinfrastructure (section II-A) that has already implemented Federated Identity Management (FIM). It is multi-disciplinary, so users may belong to more than one user community, or may be collaborating across communities, so the authorisation system must meet the requirements (section III) of user communities, home organisations, service providers, data owners, as well as the infrastructure operations. In addition, an authorisation service must of course be resilient, have sufficient performance, be sufficiently fine-grained to meet the requirements, and be sufficiently expressive and usable that it is actually used to implement the required protection policies. Further merits include being based on open standards, and having multiple interoperable implementations.

The novelty of the work presented here is precisely this balancing act: we needed a service that interfaces with the existing infrastructure, yet meets the needs of all the stakeholders mentioned earlier. Some of the constraints are technical

(protocol, connecting to extant attribute providers), some are a question of software engineering (service integration, tests) and infrastructure operations (deployment time-scale), some are policy-based (implementing and combining data and service policies), some are architectural (section IV, and some are a question of "inertia" (for example interfacing with existing practices in communities.)

II. BACKGROUND

A. EUDAT

European Data Infrastructure (EUDAT)[2] is an e-/cyberinfrastructure that provides data storage and management for a wide range of research areas[1]. The project grants access to its services based on a FIM service called B2ACCESS[3], which accepts several types of identity providers (IdPs) that are external to the project, and produces a harmonised authentication credential which is consumed by all services in the infrastructure.

EUDAT comprises several data "B2" services (B2SHARE, B2DROP, B2FIND, etc.) for the user communities, plus "internal" services that support the infrastructure itself (Wiki, JIRA, helpdesk, etc.). Nearly all of these services require authentication (through B2ACCESS). B2ACCESS provides user provisioning, attribute management, credential translation, trust management, and entitlements management, as well as the authorisation which is the topic of this paper.

B. Challenges

The main concern of authorisation in EUDAT is to manage the end user's (write) access to shared resources, as well as (read) access to data and meta-data, based on a combination of policies from different stakeholders. In particular, there is a set of infrastructure-specific and community-specific roles that need to be managed, including the rights to grant the roles to others. The community-specific roles are typically managed

[1]To be precise, EUDAT is both a project, EUDAT2020, funded by the European Commission as a part of the Horizon2020 programme, and a sustainable "collaborative data infrastructure" run by a group of organisations, which includes the project partners.

111

92

by people authorised by the communities, that is, they may be maintained on disparate services outside of EUDAT and must be imported. At the service end, EUDAT provides a range of data services which in turn are RFC-developed on top of diverse components developed outside of the EUDAT project.

C1 *Distributed authorisation* The first of the main challenges for EUDAT is thus to implement authorisation in a way that meets the requirements of the user communities, and is technically implementable across the different technologies used in the infrastructure, in particular when a FIM model is used for authentication. It must also be *trustworthy* and *consistent* across services, in order that data owners feel that their data is adequately protected.

C2 *Harmonisation* EUDAT also needs to function as an infrastructure, so authorisation information from the communities may need to be *harmonised* in order to be enforceable consistently across the infrastructure. The second main challenge (C2) is then to make the authorisation implementation *consistent* across all services and across all service providers in the infrastructure, while also making it *scalable* in the number of objects it protects and the frequency of actions, and to make it *resilient* against intermittent network failures.

C3 *Usability* The third main challenge is that the authorisation services should be *usable*. If people do not make use of the features because they are complicated or hard to manage, then the service will not be useful. EUDAT's authentication system takes the simple "portal" approach by default, while providing more complex approaches (such as command line) as options for expert users, and it is likely that the same approach will work for authorisation.

A related, but different, aspect of this challenge is the expressiveness of the policy language: If it is hard for users to express their policy in the language, they will work around the authorisation system, or will go elsewhere.

C4 *standards and interoperation* The fourth and final challenge is to be reasonably future-proof: the authorisation subsystem should be based on mature and open international *standards* and be *interoperable* across multiple implementations (particularly in order to support services implemented in different programming languages.) In particular, this approach should foster harmonisation and best practices between EUDAT and peer infrastructures.

In this paper, our focus is primarily on C1 and C4; we will not have space here to deal with all and, in fact, not all have been fully addressed yet in the project.

C. Context and Related Work

At a high level, the context of the work presented here is the need for "federated authorisation" to support e-/cyber-infrastructures as summarised by [4] and [5]. In particular, the context includes authorisation as implemented by our peer infrastructures [5], and for both these and EUDAT, the use of FIM needs to be followed by an authorisation model which fits with the established authentication methods.

The original concept of Grid computing included Virtual Organisations (VOs) as a means of managing communities, and, as in EUDAT, there was a need to manage policies from multiple stakeholders and resource/quotas in a suitably expressive and flexible way [6]. Authorisation was split into the several parts:

1) The VO negotiates with sites to obtain resources for its members and defines the scope of the work, etc. Roles and their rights are defined.
2) Individual users, once they can authenticate, request membership of the VO.
3) Authorisation attributes are assigned to individual users (by their principal).
4) Users request roles (and, possibly other authorisation attributes).
5) When users access resources, they are typically *mapped* to a local user id based on their role, or if they have no role, a default id assigned to the VO members.

As we shall see (section IV-C), we chose to use eXtensible Access Control Markup Language (XACML) for our implementation. The most important prior work of direct relevance to EUDAT is [7].

Outside of the science world, authorisation in distributed environments have been studied as well

III. REQUIREMENTS

Given the distributed nature of the EUDAT infrastructure, we have identified the following high-level requirements:

R1 : Easy, centralised and delegated management of authorisation policies.
R2 : A resilient, scalable, and highly available authorisation infrastructure.
R3 : Auditable authorisation policies.

Addressing challenges C1-C3, R1 requires there to be an architecturally central service[8] to manage authorisation on behalf of all the stakeholders. R2 is a standard requirement and not specific to our project; and, addresses C2: how to ensure that the implementation of the policies is correct and consistent across the infrastructure.

A. Use case

We use the EUDAT B2SAFE and B2STAGE services as an example use case to illustrate two of the requirements. B2SAFE provides safe (i.e. replicated) storage across several data centres, and because access is *enforced* at the service level, there is a need for harmonised authorisation (R1). B2STAGE is used to move data between EUDAT and other infrastructures, as illustrated in Fig. 1, and introduces the requirements R1 and R2.

A typical workflow is a community user running a simulation and then importing the results into B2SAFE by using the B2STAGE API, making multiple copies in different locations. For this to work, we need to be able to allow write access to specific storage resources in different data centers, based on the user's attributes (such as community membership and role).

112

93

Since multiple data centers need to have access to the authorisation policies in real time, there is a clear need for a scalable and highly available solution (R2). If there is any issue in the communication between EUDAT centers, each B2SAFE should still be able to make authorisation decisions. This demonstrates the need for a distributed authorisation infrastructure (even if, architecturally, it is a single, central, service: There is only one authorisation service, but it must offer multiple service endpoints).
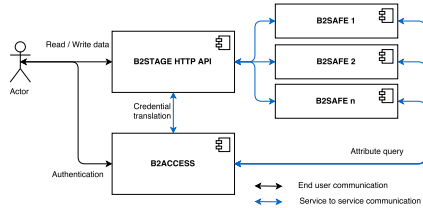

Fig. 1. B2SAFE, B2STAGE use-case

B. RBAC vs ABAC

Attributes are associated with user identities in the EUDAT infrastructure. Therefore, an Attribute-based Access Control (ABAC) approach is preferred over a more traditional Role-based Access Control (RBAC). ABAC provides more flexibility and finer granularity over the more traditional RBAC approach: RBAC requires defining the roles upfront, whereas ABAC requires only the upfront definition of a set of attributes. The ABAC approach matches with the set-up of B2ACCESS, where a fixed set of attributes is already defined and associated with each user. Moreover, as described in [9], an ABAC-based approach does not exclude roles. In ABAC, attributes with role names can be introduced together with rules controlling the modes of access to the protected objects. One of the disadvantages of ABAC, as mentioned in [9], is that auditing (III) is more difficult because the set of attributes/values is dynamic, making it more difficult to enumerate all possibilities. Within B2ACCESS, the set of attributes is fixed, and because of the proxy-like nature of B2ACCESS[8], attribute values are cached at the B2ACCESS service, making it possible to make a snapshot of all identities with access to the EUDAT infrastructure and their associated set of attributes and values. The set of authorisation policies is also centrally available, making it relatively easy to compute the set of permissions of a user at a given moment in time, allowing us to fulfil (RIII). For example, the B2SHARE service requires following (more than merely role) attributes to grant sharing or upload access rights to a user: *community-name* (subject is associated with), *community-role* (the subject has within the community), *email* (to receive/send sharing requests and notifications), *user workspace*, *endpoint URI* (resource information) and *share / upload* data (the invoked action).

IV. ARCHITECTURE

Based on work in existing infrastructures – including EU-DAT – the AARC project identified a common architecture for authentication and authorisation[8].

A. Authentication and user management

The EUDAT authentication service, B2ACCESS, enables users to authenticate, and provides account management. Its features include:

1) support for multiple external authentication protocols (OpenID Connect (OIDC), SAML, X.509, LDAP), and translation of security tokens between different authentication protocols
2) integration with eduGAIN[10], thus supporting identities from hundreds of Universities and Research institutions around the world
3) provisioning of a single user account, and a unique representation of the user identity to the infrastructure
4) user account de-provisioning (i.e. users can request to be "forgotten")
5) support for the proxy Identity Provider (IdP)/Service Provider (SP) concept[8] (acting as an SP to external IdPs and as an IdP to the SPs, i.e. the EUDAT services)
6) "enrichment" of user identities with extra infrastructure-specific attributes (cf. III-B)
7) management of users and attributes in groups, representing user communities (e.g. CLARIN[11], EPOS[12], ENES[13])

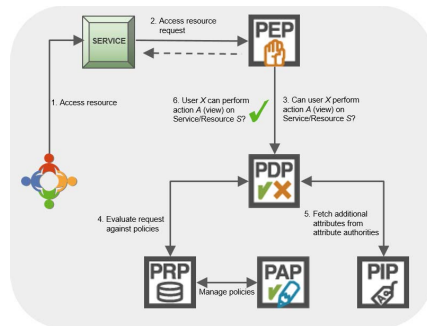### B. Authorisation Model

Fig. 2. XACML architecture

Authorisation in EUDAT is based on OASIS's XACML[14], which specifies an access control policy language instead of the more traditional Access Control Lists (ACLs). XACML implicitly supports (see IV-C) Attribute-Based Access Control (ABAC), and relies on the evaluation of subject, resource, and environment attributes to form an access decision.

113

In addition, XACML provides an abstract architecture, which consists of the following five components (see Figure 2):

1) Policy Administration Point (PAP): The PAP is an administrative service that allows stakeholders to create, manage, debug and store the relevant access control policies. Depending on the reference implementation, the authorisation services can offer a Graphical User Interface (GUI) and/or RESTful Application Programming Interface (API) to administer the policies. The current XACML standard and its profiles do not have a standardised PAP API yet.

2) Policy Decision Point (PDP): Also called Context handler, the Policy Decision Point (PDP) component evaluates access requests (which may contain authorisation attributes) against access control policies and computes a response, i.e. an access decision. Usually the response is *PERMIT* or *DENY*, but it can also decline to take a decision, for instance deferring the decision to another service.

3) Policy Enforcement Point (PEP): Usually, the PEP intercepts the access request, sends a request to PDP and acts upon the response.

4) Policy Information Point (PIP): The PIP is an optional component that is responsible for fetching subject attributes from external attribute providers. The authorisation service can leverage the Policy Repository (PIP) if the Policy Enforcement Point (PEP) does not submit all required attributes with the access decision request.

5) Policy Repository: An infrastructure needs a service which stores the policies, where they can be accessed by the PDP and the Policy Administration Point (PAP).

In addition to the policy language, XACML defines the structure of access requests and responses. For XML implementations, a normative schema XML[2] facilitates standards compliance and interoperation between implementations. The specification has also defined a Java-script Object Notation (JSON) rendering[15], which is only limited to request/response messages.

In addition to this, XACML defines a number of profiles for communication and integration between services, namely:

P1 The *Administration and delegation profile* is used to express administration and delegation policies which enable administrators to delegate – and limit – administrative rights to local administrators to enforce access control on a subset of protected resources[16] (cf. R1.)

P2 The *Security Assertion Markup Language (SAML) profile* enables integration of SAMLv2[17] with XACML. The PDP can consume SAML attribute assertions in order to make authorisation decisions[18].

P3 The *REST profile* partially defines a RESTful API which currently focuses on communication (see 2) between the PEP and PDP [19].

P4 The *Multiple decision profile* allows a requester——typically the PEP—to send several access decision re-

quests in one go, to which the PDP returns one answer with multiple decisions [20].

P5 The *Digital signature profile* [21] defines the authenticity and integrity of XACML schema instances using the W3C XML-Signature Syntax and Processing standard [22].

P6 The *Hierarchical resource profile* provides access control for resources organised as a hierarchy, such as file systems, XML documents, or organisations [23].

P7 The *Hierarchical Role-Based Access Control (RBAC) profile* defines the requirements for core and hierarchical RBAC [24] through XACML policy language.

P8 *Intellectual property control profile:* This profile enables service providers to write and enforce policies for the purpose of providing access control for resources deemed intellectual property [25].

P9 *Privacy policy profile:* This profile lets service providers express privacy policies in XACML, which defines the limits, quality, purpose, and accountability principles of user's personal data [26].

C. The choice of XACML as the policy language

We have briefly discussed XACML in section (IV-C). There are currently two versions v2.0 and v3.0: we mainly focus the latter as it comes with support for all the profiles of the former (P2, P5, P6, P9, and includes a new set of profiles (P1, P3, P4 and P8). As some of the B2 services in EUDAT use OIDC for authentication (see IV-A). These services will need to use the JSON rendering and REST profile to communicate between the service's PEP and the PDP. The profiles in XACML v3 will thus enable us to integrate these services.

However, EUDAT takes into account certain recommendations from earlier v2 based work[7], in order to promote interoperation within the infrastructure, and, eventually, across infrastructures

- subject names are always X.509 distinguished names as in the SAML assertions (section IV-A), irrespective of whether users have a certificate issued to them through the X.509 "gateway"[27]
- attributes are fully qualified, and the PDP matches against the full attribute string
- future extensions will need to look at the obligations, where the PEP specifies which types it is prepared to honour, as they will be important for some user communities[3]. In [7], the issue is versioning; for EUDAT the issue is rather differences between the capabilities of the services, meaning PEPs are likely to handle different types of obligations.

In contrast, notable differences to [7] are

- users may, but need not, have an X.509 certificate,
- they may, but need not, have VOMS assertions assigned to their subject name[7] (by an authority outside of EUDAT);

---

[2]XACML XML Schema: http://docs.oasis-open.org/xacml/3.0/xacml-core-v3-schema-wd-17.xsd

---

[3]Ultimately, it's a question of usability, as the obligations can help communicate additional constraints. However, this use case is beyond the scope of the present paper.

- Users may be members of more than one community ("VO" in [7]) and will need to simultaneously assert membership of both/all, as well as roles in each one.

Policies are defined by the stakeholders. Obligations are defined by them and it is up to the PDP to send it and to the PEP to implement it.

The EUDAT operations team introduced service-specific attributes for each of the services. Requirement R1 enables administrators to give, say, B2SHARE-specific attributes to users. With time, the associated service-specific authorisation policies will become more sophisticated, and will need maintaining by multiple parties. At the same time, each centre will define policies for all its own services. Thus, there is a need to combine policies defined by different stakeholders into policy sets applicable to the request, with appropriate combination algorithms. Although we are not using it yet, we expect the delegated administration profile [16] will make this process easier.

D. Authorisation in EUDAT

Figure 3 depicts the XACML-based hierarchical architecture, which aims to address the requirements of implementing consistent [R1] yet highly available [R2] authorisation in a distributed infrastructure,

From the top of the component hierarchy, EUDAT Authentication and Authorisation Infrastructure (AAI) consists of a central PAP and Policy Repository (PRP), the latter being a database of rules. At this level, rules are defined as *Policy Sets* for each type of services (section IV-C.) The service-specific policies are managed by service administrators through the central PAP service. Delegation of policy administration rights will use the XACML v3.0 Administration and Delegation profile to define the policies for access to the resources. The changes made at the top-level PAP update the policies at the top-level PRP, a database of policies.

Eventually, delegation of administrative rights should encompass all policy stakeholders: data owners, community admins, resource admins, site admins, and the infrastructure admins themselves. The combination of policies needs to resolve based on the target: site admins will have priority for services at their site, community admins are authoritative only for their own data and their own users, etc.

For each EUDAT data center, there should be a full XACML stack with a PEP for each service (or a group of closely co-located services), and a single PDP for the center together with a local, read-only PRP. Although the PIP is displayed in Fig. 3, all the required information (attributes sent by B2ACCESS) is in practice sent via the PEP to the PDP.

Administrator creates or updates policy through the central (read-write) PAP (Fig. 3). The central PRP pushes these policies or policy sets to the site PRP.Each site PRP receives the update and through an *eventually consistent* [28] policy database updates its information. The PDP accesses only the relevant policies from the site PRP in order to evaluate the access decision requests, e.g. for a B2SHARE PEP, it will request only B2SHARE policy sets.
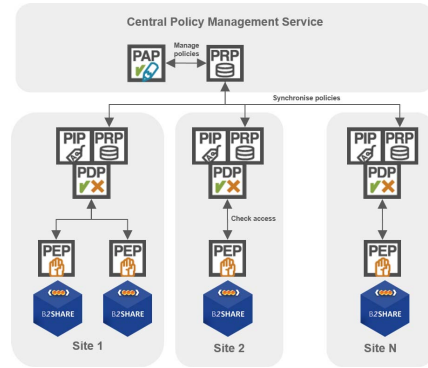

Fig. 3. The EUDAT authorisation architecture

In future work we will introduce a message broker between the site and the central PRP to ensure the reliability of the updates.

E. Access control flow
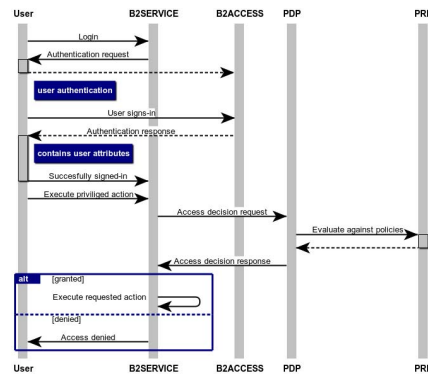
Fig. 4. Service authorisation flow

We now revisit the use case from section III-A, but first from a generic service point of view. Fig. 4 shows the generic authorisation flow for a user accessing a B2-service. The EUDAT community user (e.g. CLARIN, EPOS, ENES) initiates the authorisation flow by trying to execute a privileged action on the B2-service. This, however, requires appropriate rights on the service. Since the user is not authenticated yet, they will be redirected to the B2ACCESS service for authentication (section IV-A).

After successful FIM authentication to B2ACCESS, B2ACCESS returns an authentication token along with authorisation attributes to the B2-service, which will then retry the privileged operation.

During the retry operation, the service PEP sends an authorisation request to the (site-local) PDP service, which contains the *user attributes*, *action* and *resource* information. The PDP service evaluates the request (and attributes within) against the policies stored in the PRP and returns the decision which is then enforced by the PEP.

Returning to the use case (Fig. 1), the user wishes to import result data from their simulation in EUDAT via the B2STAGE service, which in turn ensures replication of data across multiple B2SAFE instances. Both B2STAGE and B2SAFE check authorisation.

Traditionally, users access the B2STAGE service through a HTTP API using command line clients, with a delegated X.509 certificate. The certificate in the current implementation always contains authorisation attributes in SAML format[27] in a custom extension[29]. The B2STAGE can thus extract the authorisation attributes directly after successful authentication.

Assuming the user is authorised by B2STAGE based on the attributes, the service obtains a delegated certificate from the certificate the user client used to authenticate[30], which in turn contains the certificate with authorisation data[4]. Data is copied through B2STAGE to the B2SAFE instances using GridFTP (cite GridFTP), and the B2SAFE services in turn perform their own authorisation check.

As with B2STAGE, the authentication subsystem of B2SAFE extracts the SAML assertion from the relevant certificate(s), and builds a PDP request for the requested action (data ingest) to grant/deny the access.

V. COMPARISON OF IMPLEMENTATIONS

As discussed briefly about well-known XACML implementations in Sect. IV-C, this section provides a high-level analysis of those implementations. It skips those without support for a remote PDP interface. Taking the authorisation requirements of EUDAT into account, the profiles required by the infrastructure services are Core, REST and JSON. XACML version 3.0 seemingly is the adequate standard, as the later two profiles are only available in the specification version. Since it would normally be a daunting task for the operators to define and alter policies in a raw XML format, having a web- or desktop-based graphical user interface, specifically PAP - is a key to integration of the EUDAT services with the authorisation system.

Table I provides a list of open source and commercial implementations, with supported profiles and some offers GUI to manage the access control policies. It can be observed that there is no implementation that can address all of the EUDAT requirements, however, WSO2 Identity Server offers most

[4]OGF VOMS attribute PROCessing Working Group, https://redmine.ogf.org/projects/voms-proc-wg

of the functional features except the replication of XACML policies from the root PAP node to the lower-level PDP servers (see Fig. 3). Therefore, some additional effort is likely required before the different architectural components of the authorisation system can be deployed.

VI. CONCLUSION AND FUTURE WORK

When designing an authorisation service for several types of services in the EUDAT distributed infrastructure, the main challenges were to implement consistent and harmonised authorisation across services and sites, supporting stakeholders from multiple communities through user- (or admin-) friendly interfaces, and based on established standards and interoperable implementations.

The authorisation infrastructure is based on XACML, a declarative policy language, deployed in a hierarchical fashion, with locally cached policies and update propagation. At the time of writing, the current deployment is somewhat limited as it only involves infrastructure administrators as policy managers (no delegation), and authorisation is not yet integrated with all B2 services. The current deployment is simultaneously a feasibility study, a partial implementation, and an indicator of future directions. Apart from the obvious ones, of wider deployment in more production services, future directions also include:

- Implementing the delegated administrative rights, in order to support the multi-stakeholder management of policies;
- Interoperation with other infrastructures, notably EGI, which uses XACML version 2, and PRACE which uses an LDAP based system.
- Implementation of obligations.

REFERENCES

[1] J. Vollbrecht, P. Calhoun, S. Farrell, L. Gommans, G. Gross, B. de Bruijn, C. de Laat, M. Holdrege, and D. Spence, "Aaa authorization framework," Tech. Rep., 2000.

[2] W. Gentzsch, D. Lecarpentier, and P. Wittenburg, "Big data in science and the eudat project," in *2014 Annual SRII Global Conference*, April 2014, pp. 191–194.

[3] "B2access," 2017. [Online]. Available: https://www.eudat.eu/services/b2access

[4] T. A. Study, "TERENA AAA study final report: Advancing technologies and federating communities," 2012. [Online]. Available: https://wiki.geant.org/display/aaastudy/AAA+Study+Home+Page

[5] K. Christos, L. Nicolas, van Dijk Niels, and S. Peter, "Deliverable djra1.1: Analysis of user community and service provider requirements," AARC Project, Project Deliverable AARC-DJRA1.1, 10 2015. [Online]. Available: https://aarc-project.eu/wp-content/uploads/2015/10/AARC-DJRA1.1.pdf

[6] K. Keahey, V. Welch, S. Lang, B. Liu, and S. Meder, "Fine-grain authorization policies in the grid: Design and implementation," in *Proc. 1st Int'l Workshop on Middleware for Grid Computing*, 2003. [Online]. Available: http://toolkit.globus.org/alliance/publications/papers/gauth02.pdf

[7] R. Ananthakrishnan, G. Garzoglio, and O. Koeroo, "An XACML attribute and obligation profile for authorization interoperability in grids," Open Grid Forum, Jan. 2013. [Online]. Available: https://www.ogf.org/documents/GFD.205.pdf

Name	Spec. Version	Supported Profiles	Language	License	UI
AT&T XACML	v2.0, v3.0	Core, Multiple Decision, JSON, REST	Java	Apache 2.0	✓
ndg-xacml	v2.0	Core,SAML 2.0	Python	BSD	✗
ARGUS	v2.0	Core,SAML 2.0	Java & C	Apache 2.0	✗
WSO2 Identity Server	v3.0	Core,Multiple Decision, JSON, REST, Administrative delegation	Java	Apache 2.0	✓
FIWARE AuthzForce CE	v3.0	Core, Hierarchical RBAC, Multiple Decision, JSON, REST, Data Loss Prevention / Network Access Control, Addition Combing Algorithms	Java	GPL	✗
OpenAZ	v3.0	Core, Multiple Decision, JSON, REST	Java	Apache 2.0	✓
Axiomatics Policy Server	v3.0	Core, Multiple Decision Profile, JSON, REST, Hierarchical RBAC, Hierarchical Resource, Privacy Policy, SAML 2, XML Digital signature	Java, .NET	Commercial	✓

TABLE I
ANALYSIS OF IMPLEMENTATIONS

[8] A. Biancini, L. Florio, M. Haase, M. Hardt, M. Jankowski, J. Jensen, C. Kanellopoulos, N. Liampotis, S. Licehammer, S. Memon, N. van Dijk, S. Paetow, M. Prochazka, M. Sallé, P. Solagna, U. Stevanovic, and D. Vaghetti, "AARC: first draft of the blueprint architecture for authentication and authorisation infrastructures," *CoRR*, vol. abs/1611.07832, 2016. [Online]. Available: http://arxiv.org/abs/1611.07832

[9] T. R. W. Ed Coyne, "Abac and rbac: Scalable, flexible, and auditable access management," NIST, Report, 2013. [Online]. Available: http://csrc.nist.gov/groups/SNS/rbac/documents/coyne-weil-13.pdf

[10] "edugain," 2017. [Online]. Available: http://www.edugain.org

[11] "Clarin," 2017. [Online]. Available: https://www.clarin.eu

[12] D. Bailo, K. G. Jeffery, A. Spinuso, and G. Fiameni, "Interoperability oriented architecture: The approach of epos for solid earth e-infrastructures," in *2015 IEEE 11th International Conference on e-Science*, Aug 2015, pp. 529–534.

[13] S. Joussaume and R. Budich, *The Infrastructure Project of the European Network for Earth System Modelling: IS-ENES*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 5–9. [Online]. Available: https://doi.org/10.1007/978-3-642-36597-3_2

[14] B. Parducci, H. Lockhart, and E. Rissanen, "extensible access control markup language (xacml) version 3.0," OASIS, OASIS Standard xacml-3.0-core-spec-en, 1 2013. [Online]. Available: http://docs.oasis-open.org/xacml/3.0/xacml-3.0-core-spec-en.pdf

[15] B. Parducci, H. Lockhart, and D. Brossard, "Json profile of xacml 3.0 version 1.0," OASIS, OASIS Standard xacml-json-http-v1.0, 12 2014. [Online]. Available: http://docs.oasis-open.org/xacml/xacml-json-http/v1.0/xacml-json-http-v1.0.pdf

[16] B. Parducci, H. Lockhart, and E. Rissanen, "Xacml v3.0 administration and delegation profile version 1.0," OASIS, OASIS Standard xacml-json-http-v1.0, 11 2014. [Online]. Available: http://docs.oasis-open.org/xacml/3.0/xacml-3.0-administration-v1-spec-en.pdf

[17] C. P. Cahill, J. Hughes, H. Lockhart, M. Beach, R. M. R. Randall, T. Wisniewski, I. Reid, P. Austel, M. Hondo, M. McIntosh, T. Nadalin, N. Ragouzis, S. Cantor, R. B. Morgan, P. C. Davis, J. Hodges, F. Hirsch, J. Kemp, P. Madsen, S. Anderson, P. Mishra, J. Linn, R. Philpott, J. Moreh, A. Anderson, E. Maler, R. Monzillo, and G. Whitehead, "Assertions and protocols for the oasis security assertion markup language (saml) v2.0," OASIS, OASIS Standard saml-core-2.0-os, 03 2005. [Online]. Available: https://docs.oasis-open.org/security/saml/v2.0/saml-core-2.0-os.pdf

[18] B. Parducci, H. Lockhart, and E. Rissanen, "Xacml saml profile version 2.0," OASIS, OASIS Standard xacml-saml-profile-v2.0, 08 2014. [Online]. Available: http://docs.oasis-open.org/xacml/xacml-saml-profile/v2.0/xacml-saml-profile-v2.0.pdf

[19] B. Parducci, H. Lockhart, and R. Sinnema, "Rest profile of xacml v3.0 version 1.0," OASIS, OASIS Standard xacml-3.0-core-spec-en, 11 2014. [Online]. Available: http://docs.oasis-open.org/xacml/xacml-rest/v1.0/xacml-rest-v1.0.pdf

[20] B. Parducci, H. Lockhart, and E. Rissanen, "Xacml v3.0 multiple decision profile version 1.0," OASIS, OASIS Standard xacml-3.0-multiple-v1-spec-en, 05 2014. [Online]. Available: https://docs.oasis-open.org/xacml/3.0/xacml-3.0-multiple-v1-spec-en.pdf

[21] B. Parducci and H. L. E. Rissanen, "Xacml v3.0 xml digital signature profile version 1.0," OASIS, OASIS Standard xacml-3.0-dsig-v1.0, 05 2014. [Online]. Available: http://docs.oasis-open.org/xacml/3.0/dsig/v1.0/xacml-3.0-dsig-v1.0.pdf

[22] M. Bartel, J. Boyer, B. Fox, B. LaMacchia, and E. Simon, "Xml signature syntax and processing version 2.0," W3C, W3C Standard xmldsig-core2, 07 2015. [Online]. Available: https://www.w3.org/TR/xmldsig-core2/

[23] B. Parducci, H. Lockhart, E. Rissanen, and R. Levinson, "Xacml v3.0 hierarchical resource profile version 1.0," OASIS, OASIS Standard xacml-3.0-hierarchical-v1-spec-en, 05 2014. [Online]. Available: http://docs.oasis-open.org/xacml/3.0/xacml-3.0-hierarchical-v1-spec-en.pdf

[24] B. Parducci, H. Lockhart, and E. Rissanen, "Xacml v3.0 core and hierarchical role based access control (rbac) profile version 1.0," OASIS, OASIS Standard xacml-3.0-rbac-v1-spec-en, 10 2014. [Online]. Available: http://docs.oasis-open.org/xacml/3.0/xacml-3.0-rbac-v1-spec-en.pdf

[25] B. Parducci, H. Lockhart, J. Tolbert, C. Hayes, R. Hill, P. Tyson, A. Han, D. Thorpe, R. Sinnema, E. Rissanen, and D. Brossard, "Xacml intellectual property control (ipc) profile version 1.0," OASIS, OASIS Standard xacml-3.0-ipc-v1-spec-en, 01 2015. [Online]. Available: http://docs.oasis-open.org/xacml/3.0/ipc/xacml-3.0-ipc-v1-spec-en.pdf

[26] B. Parducci, H. Lockhart, and E. Rissanen, "Xacml v3.0 privacy policy profile version 1.0," OASIS, OASIS Standard xacml-3.0-privacy-v1-spec-en, 01 2015. [Online]. Available: http://docs.oasis-open.org/xacml/3.0/xacml-3.0-privacy-v1-spec-en.pdf

[27] A. S. Memon, J. Jensen, A. Cernivec, K. Benedyczak, and M. Riedel, "Federated authentication and credential translation in the eudat collaborative data infrastructure," in *2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing*, Dec 2014, pp. 726–731.

[28] W. Vogels, "Eventually consistent," *Commun. ACM*, vol. 52, no. 1, pp. 40–44, Jan. 2009. [Online]. Available: http://doi.acm.org/10.1145/1435417.1435432

[29] S. Farrell, R. Housley, and S. Turner, "An internet attribute certificate profile for authorization," Internet Requests for Comments, RFC 5755, January 2010. [Online]. Available: https://tools.ietf.org/pdf/rfc5755

[30] S. Tuecke, V. Welch, D. Engert, L. Pearlman, and M. Thompson, "Internet x.509 public key infrastructure (pki) proxy certificate delegation profile," 6 2004, rFC3820.

Paper V

Towards Federated Service Discovery and Identity Management in Collaborative Data and Compute Cloud Infrastructures

A.S. Memon, J. Jensen, W. Elbers, M. Riedel, H. Neukirchen and M. Book. 2018.

Journal of GRID computing 16(4), 663–681 (2018) [DOI: 10.1007/s10723-018-9445-3]

Shiraz Memon was the product leader of the EMIR service and main contributor of the B2ACCESS service in the EMI and EUDAT (now EOSC-hub) infrastructures respectively. He is the main author of this publication and provided the majority of the content of this publication.

CrossMark

Towards Federated Service Discovery and Identity Management in Collaborative Data and Compute Cloud Infrastructures

Shiraz Memon · **Jensen Jens** · **Elbers Willem** ·
Helmut Neukirchen · **Matthias Book** ·
Morris Riedel

Abstract This paper compares three multi-national
research infrastructures, one that provides data ser-
vices, one that provides compute services, and one
that supports linguistics research. The aim is to jointly
provide services to the user communities, and, per-
haps eventually, seamlessly interoperate. To this end,
we look at and compare how the infrastructures build
their service federations (trust, service status, informa-
tion systems), and how they manage users (identities,
authentication, and authorisation).

H. Neukirchen · M. Book
University of Iceland, Reykjavik, Iceland

H. Neukirchen
e-mail: helmut@hi.is

M. Book
e-mail: book@hi.is

S. Memon (✉) · M. Riedel
Jülich Supercomputing Centre, Forschungszentrum Jülich,
Leo-Brandt Straße, 52428 Jülich, Germany
e-mail: a.memon@fz-juelich.de

M. Riedel
e-mail: m.riedel@fz-juelich.de

J. Jens
STFC, Harwell Oxford Campus, Didcot, UK
e-mail: jens.jensen@stfc.ac.uk

E. Willem
CLARIN ERIC, Utrecht, Netherlands
e-mail: willem@clarin.eu

1 Introduction

Distributed compute, data, and more recently, cloud
infrastructures have been successful in providing
resources to a wide variety of research communi-
ties. The e-Infrastructure Reflection Group identified
in 2004 the outline/vision of a distributed infrastruc-
ture comprised of fabric (disk, CPU, networks), and
a "middleware" layer connecting the infrastructure
across sites; user communities would then develop
and deploy their own applications on top of the e-
infrastructure [44]. Also the Foster/Kesselman vision
of grid computing [31], with computing available on
demand through standard interfaces, was hugely influ-
ential in the development and use of e-infrastructures,
leading for example to the middleware that is known
as Globus Toolkit [29] and more recent Globus cloud
services [30].

The established e-infrastructures have been very
successful, having provided resources to researchers
on a national or multinational scale in TeraGrid [36],
European National Grid Initiatives (NGIs), Extreme
Science and Engineering Discovery Environments
(XSEDEs) [52], or, in the case of the world-wide

 Springer

Large Hadron Colliders (LHCs) Computing Grid, a truly global scale [45]. They have provided data and compute resources in support of a vast range of research.

The main contribution of this paper is *connecting the infrastructure*, particularly focusing on security and service discovery (Fig. 1). There is plenty of existing work on e-infrastructure architecture and security, managing users and their communities [2, 8, 13, 18], which we summarise below for the reader's convenience. We are, however, interested in the practical applications, so we have chosen three infrastructures with different purposes and look at the general challenges of bridging them, as well as connecting their user communities. We also look at the specifics of some of the key services involved in this endeavour, going into details of recent developments.

1.1 Connecting the Infrastructure to Itself

The following components are the key components to defining and binding together an infrastructure:

– Common fabric security, i.e., X.509 host certificates from trusted Certification Authorities.
– Service naming: Each relevant service must have a *name* by which it can be discovered and referenced; a typical type of name is a Web services endpoint or URI.
– Service discovery/metadata: a way to discover which services would be available to the user.

– Service registry: a location where each service is registered, typically used to record whether it is a legitimate part of the infrastructure and whether there are scheduled downtimes, etc.
– Service information granularity: The information model representing the service should be sufficiently flexible to capture the service details from a coarse- to fine-grained level. Furthermore, the model must be interoperable as multiple infrastructures are discovering and advertising their services.
– Operations and support: From the user's perspective, there should be a single point of contact for support, and there should be a *team* responsible for operating the service (as opposed to individual admins at each site.)

1.2 Connecting Users to the Infrastructure

Central to the e-infrastructures that are a focus in this paper are:

– Common authentication: This allows each user to access any part of the infrastructure with a single credential (as well as accessing other infrastructures with the same credential);
– Service discovery mechanisms: There has to be an "entry point" which helps users discover services that are available to them. Typically, this is a portal, but could also be hosted on a "user interface" node (to which users log in or connect with remote desktop);
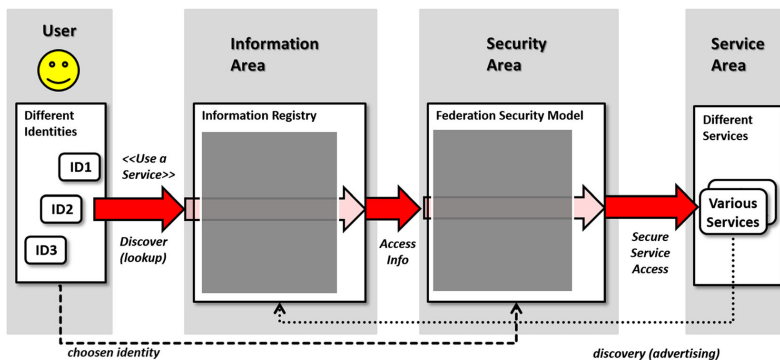

Fig. 1 EUDAT federated authentication and service discovery. The EUDAT architecture without any specific authentication and service discovery architecture. On the right are the EUDAT's B2*services

– Service database (which may or may not be the same as the service discovery): Typically, it is a central database listing the services that are part of the e-infrastructure. By extension, an associated service could be used to monitor service status, announce scheduled downtimes, etc.;

– Common authorisation: This is needed across the infrastructure to provide additional actions to researchers and users, enabling them to share data and to collaboratively make use of the services provided.

Note the difference between the service discovery and registry/database in Sections 1.1 and 1.2: While they might be the same service in some infrastructures, the former is more likely to have an Application Programming Interface (API) to allow programmatic access (cf. R14 below), or technical interfaces for administrators, whereas the latter should be browser-accessible and more user-friendly.

2 Architecture and Concept Backgrounds

Unsurprisingly, the e-infrastructures covered here are architecturally similar; even with independently designed architectures they end up often providing the same types of services. Indeed, one of the achievements of the AARC project was a unified view of the authentication and authorisation parts of the e-infrastructures [32]. Also, common standards and interoperation play an important role, such as the GLUE standard (Section 4.3.1), as they enable service discovery across domains if used correctly [19].

Table 1 shows an overview of how the three different infrastructures provide interfaces for their

users and how they are connected internally. Here, "CLI" is short for "command line interface" (which is generally considered harder to use for novices but saves time for experts); "WS" refers to web services for programmatic access; and X.509 is the standard for certificates [18] provided through IGTF (www.igtf.net). VOMS is the Virtual Organization Membership Service, an attribute authority [2]. Finally, BDII (Berkeley Database Information Index) and GOCDB (Grid Operations Centre DataBase) are information services, used for service discovery and registry, respectively, and are covered in more detail in Section 3.3.

3 Requirements Analysis

In today's research environments, Single Sign-On (SSO) is an important requirement: It enables researchers to use a single account to access remote services, and service providers do not need to maintain separate account data, nor do they need password quality checking, password reset, maintaining user contact details, etc. Importantly, researchers present the same identity and can use the same credential with several different services, so SSO can potentially bridge infrastructures.

Extending SSO, national research networks build identity management federations where Identity Providers (IdPs) are bound by common federation policies, thus ensuring a common level of assurance (LoA) of identities and a common set of attributes being passed to the services. These attributes are used to identify (or at least represent) the user to the service, and/or used for authorisation. Typically, these national

Table 1 Infrastructures need ways to give access to users, and to link services within the infrastructure. Some are the infrastructure's own, others are shared or come from an external federation. Abbreviations are explained in Section 2

Service		CLARIN	EUDAT	EGI
	Authentication	Federated/own	Federated/own	X.509/federated/own
	Access methods (Web/CLI/WS)	Web	Web/CLI	CLI/WS
	Authorisation	Own	Own	VOMS
	Service discovery	Portal/Switchboard	Portal	Wiki
User	Workflow	WebLicht	N/A	N/A
	Authentication	IGTF	IGTF	IGTF
	Service discovery	Portal/Switchboard	N/A	BDII
Infra	Service registry	Switchboard	GOCDB	GOCDB

⁂ Springer

federations use web-based technologies (users use a web browser to access services via portals), such as the SAML Web Single Sign-On (Web SSO) profile, and use (subsets of) the eduPerson schema to publish attributes.

As much research is international, it becomes useful to connect national identity federations, despite their publishing different attributes or having different levels of assurance (LoAs). eduGain [20] is an inter-federation identity management framework, which aims at interconnecting the national federations. However, there is still a need for harmonisation due to the differences between national federations; this is the subject of ongoing work from REFEDS (www.refeds.org) and recent work from the Authentication and Authorisation for Research and Collaboration (AARC) project [1]. As we shall see, one option for infrastructure projects is to implement a *proxy* to harmonise credentials [14], and perhaps, via *credential translation*, provide support for non-web (command line) access. The other main option is to simply implement a project or community-specific independent (non-federated) IdP. Obviously, many of the advantages of SSO are then lost, but as we shall see, the adherence to standards creates opportunities for interoperation between infrastructures.

In the following subsections, we analyze the requirements from three different infrastructures: a research community infrastructure, a data infrastructure, and a compute/cloud infrastructure, the latter two being multi-disciplinary. We look at these as individual infrastructures (cf. Table 1), but also at how they can share users and services such as workflows.

3.1 CLARIN European Research Infrastructure

Common Language Resources and Technology Infrastructure (CLARIN) [15] provides easy and sustainable access for scholars in the humanities and social sciences to digital language data (in written, spoken, or multi-modal form), as well as access to advanced tools to discover, explore, exploit, annotate, analyse or combine the data, regardless of where it is located. CLARIN is building a networked federation of language data repositories, service centres and knowledge centres, with SSO access for all members of the academic community in all participating countries. Tools and data from different centres are interoperable, so that data collections can be combined and tools from different sources can be daisy-chained to perform complex operations.

The CLARIN infrastructure is fully operational in many countries, and a large number of participating centres are offering access services to data, tools and expertise. At the same time, new services are added by countries that joined more recently, and CLARIN's datasets and services are constantly updated and improved. On the services page [16] we show the services accessible at this moment, and explain how and by whom the various services can be accessed.

3.1.1 Requirements

R1 *Single Sign-On (SSO)*: To provide single sign on, users must be able to use a single identity for all CLARIN services, and credentials should only be required for the first authentication. Authorization within the CLARIN infrastructure is not centrally managed, but on a service per service basis. This is a result of the distributed nature of the infrastructure, where each CLARIN centre is responsible for the services it runs.

R2 *Delegation* of user rights is crucial in a distributed service oriented infrastructure such as the CLARIN infrastructure [11]: A user typically stores data in a workspace and wants services, possibly hosted at other centres, to process the data in these workspaces. The user is authenticated and authorized to the service and then wants to delegate his/her identity and permission to the service, so the service can access the workspace on behalf of the user.

R3 *Service discovery*: Given a dataset, what services are available to process this dataset? Given a service, what other services are available to operate on the output of this service? It is necessary to have a discovery service which describes the services' capabilities and provides endpoints for accessible resources and services. An example of such a registry is the Language Resource Switchboard [53]. It is important to point out that such a service registry is not a workflow composition engine itself; instead a workflow composition engine typically queries a service registry during workflow composition.

3.2 EUDAT

European Data Infrastructure (EUDAT) [26] is a European data infrastructure which facilitates management and federation of "big (research) data" across Europe. It operates a number of services to deposit, replicate, and archive data. Services are geographically distributed across different organisations (which are currently the same as the project partners).

3.2.1 Requirements

R4 *Single Sign-On (SSO):* Users should be able to access EUDAT services while authenticating with their "home" credentials issued by their organisation's identity provider. Without SSO, the users would have to register with every service, and each service in EUDAT would have to maintain its own user database. This would not be scalable and might lead to inconsistencies where the same information is stored in multiple databases. Therefore, the Authentication and Authorisation Infrastructure (AAI) technology must be able to support SSO.

R5 *Distributed authorisation:* Once users can *authenticate*, the infrastructure needs to provide an *architecturally central* authorisation service (i.e., there is only one) which is consistently enforced across the *distributed* services. The main goals are: (1) harmonised authorization policy management per service, (2) authorization decisions must be applied even in case any centralised service is unavailable, and (3) based on standards such as eXtensible Access Control Markup Language (XACML) [42].

R6 *Non web-based federated access:* While web-based services are used as "high-level" access points, there is sometimes a need to support command line tools and "delegated" credentials. Typically these drive services based on the data transfer protocol GridFTP [3], the storage service based on iRODS [17], or services offering REST APIs.

R7 *Delegation* of rights to other users or services is important in a data management pipeline where the service or user should be able to perform a task on a behalf of the user/owner of the data or resource.

R8 *Multiple authentication protocols:* None of the EUDAT services were written from scratch; they were all developed around existing software products. However, there was no single authentication mechanism supported by all these products, so EUDAT's choice was to either choose a common mechanism and implement it in all services, or alternatively support multiple authentication mechanisms within the infrastructure. EUDAT, building on previous experiences in its project phase one, chose the latter. Hence, the AAI should act as an intermediary (a *proxy* in [10]) between the user and the services and *translate* the credentials from one form to the other to enable seamless access to the service.

R9 *Different level of assurance (LoA):* Often, most of the users perform less sensitive operations, for example reading a data set from B2SHARE (the EUDAT data sharing service). For some of the users, a high LoA is needed to perform privileged operations, for example uploading a dataset or invoking a data archival operation. A low LoA is rather useful for the volunteer scientists (e.g., holding social identities [34]) who are only interested in, say, visualisation of data. Therefore it is highly desirable for the EUDAT AAI to support segregating the service actions into different levels, hence associating each credential with a different LoA.

R10 *Service discovery:* EUDAT infrastructure is comprised of many distributed heterogeneous services and resources: storage resources, their providers, data services, and authentication services, etc. Thus, it is essential to know the offered capabilities, types, and other specific characteristics (e.g. data transfer rate or storage capacity) of services. The infrastructure should enable users as well as monitoring system to discover the services based on the service properties.

3.3 EGI

The European Grid Infrastructure (EGI) [21] is one of the largest multidisciplinary grid and cloud infrastructures in Europe, hence a wide number of scientific user communities and resource providers are involved. EGI offers a set of distributed services which

enable users to execute complex computing workflows. The authentication and authorisation infrastructure is based on Public Key Infrastructure (PKI), the service discovery is supported by incorporating Berkley Database Information Index (BDII) [9] and Grid Operations Centre Database (GOCDB) [37]. This section focuses on federated authentication [14] and service discovery [22] requirements of the EGI infrastructure.

3.3.1 Requirements

R11 *Single Sign-On (SSO):* The users should be able to use their single institutional identity to access the EGI services. Since EGI is based on PKI, users normally authenticate with their end-entity X.509 certificate. SSO will require a proxy generating a temporary certificate on behalf of the user (via a trusted online Certification Authority).

R12 *Non web-based federated access:* Most of the EGI services are accessed through web portals, but some of them offer command line access.

R13 *Delegation:* Users often submit compute jobs or workflows to the EGI High-Performance Computing (HPC) or cloud resources, and the user job may need to stage-in or stage-out data to a storage resource. Consequently, the compute service may need delegated access to the storage resource on the users' behalf. The delegation of rights is essential in the given use case, and in some cases credential translation is necessary as the services may not necessarily use the same authentication protocol.

R14 *Service discovery:* In addition to providing lists of services for users and administrators, the service registry plays a significant role in composing as well as executing workflows.

3.4 Specific Service Discovery Requirements

R15 *Common service information model:* The federated infrastructure registry should be able to provide a means of publishing information in a standard- and middleware-agnostic manner.

R16 *Unified service registration and query protocol:* EGI uses different middlewares (UNI-CORE, ARC, Globus, HTCondor, etc.), each

potentially with its own native information system. While a provider only needs to talk to their "local" information system (R18), it would be nice if all information systems had a consistent API.

R17 *Service lifecycle management:* It is necessary to have a consistent API to manage the whole lifecycle of the services by the service providers/publishers—registration, discovery, query, downtimes, suspension, deregistration.

R18 *Support for registry hierarchies:* Each domain (NGI) has its own registry since it can act as an infrastructure in its own right. Support for a registry hierarchy provides a unified registry for the infrastructure.

R19 *Replication of service information:* To achieve robustness within the service discovery infrastructure, the technology should support replication of information across distributed entities whereby the failure of one registry node should not hamper the functioning of other registry nodes. Moreover, better performance can also be achieved by routing traffic to less occupied registry nodes. The registry should be able to replicate its state across other registry nodes in an automated fashion.

R20 *Scalability:* The registry should be able to cope with the discovery of large numbers of services in a global scale infrastructure. Since the number of services can also grow dramatically, the underlying database technology should be capable of distributing the service records horizontally and in a cost-effective manner.

3.5 Discussion

Table 2 summarises the AAI and service discovery (SD) requirements from EUDAT, CLARIN, and EGI. It can be observed that most of the requirements are overlapping with each other. This, however, creates a strong motivation for having a common framework for federated service access and discovery. Although they seem similar, it is pertinent to consider certain factors, such as the number of users and services, types of services, cross organisational/domain/country service access, attribute naming, data access policies, user and service provisioning, and attribute mapping.

Table 2 Summary of the requirements analysis

Requirements		CLARIN	EUDAT	EGI
AAI	SSO	✓	✓	✓
	Delegation	✓	✓	✓
	Non-web federated access	✗	✓	✓
	Multiple authentication protocols	✗	✓	✓
	LoAs	✗	✓	✓
	Distributed authorisation	✗	✓	✓
SD	Service discovery	✓	✓	✓
	Unified API	✓	✓	✓
	Replication	✓	✓	✓
	Hierarchies	✗	✓	✓
	Service info. lifecycle management	✓	✓	✓
	Common information model	✓	✓	✓

4 Unified Federated Discovery and Identity Management

4.1 B2ACCESS: The EUDAT AAI Proxy

The B2ACCESS architecture is shown in Fig. 2. On the left-hand side of the diagram, B2ACCESS maps primary user identities, including a (sub)set of associated attributes, from external domains onto the EUDAT domain. The external IdPs can be connected to the B2ACCESS service by using different technologies: Security Assertion Markup Language (SAML), X.509 certificates, and OpenID Connect. For users without access to a suitable IdP, B2ACCESS itself can act as an IdP via a B2ACCESS-specific username and password. On the right-hand side of Fig. 2, the harmonised credential connects to EUDAT services also using different technologies, depending on the target service: SAML, OAuth2, or short-lived X.509 certificates. In all cases, credentials are managed by B2ACCESS and can be delegated to the target service (for credentials that support delegation), and need not be managed by the user at all: Only users who need command line tools need to download and manage credentials (in our case, the X.509 certificate).

In particular, B2ACCESS releases a unified set of attributes (Table 3) to the Service Providers (SPs) in the EUDAT infrastructure. The SPs can define authorization policies to grant certain permissions to a user based on the values of attributes associated with the user's identity. This is known as Attribute-Based Access Control (ABAC), as opposed to the more traditional Role-Based Access Control (RBAC). Examples are group membership, community membership, and LoAs.

B2ACCESS also provides account management, both for the users themselves and administrators. While many of the attribute values are gathered from the external IdP or during the registration process and are fairly stable, group membership can change more often and thus needs a management workflow, as well as delegated permissions (to community/group managers) in B2ACCESS. This is discussed in Section 4.1.3.

The B2ACCESS approach requires a one-time registration step for new users. The first time a user logs in by using B2ACCESS, the user is presented with a registration form. This allows us to require acceptance of license agreements and terms of use and, if needed, to request additional attributes. After completing this registration step, the actual mapping from the external identity onto the EUDAT identity is persisted in the B2ACCESS database.

4.1.1 Example: Accessing B2SHARE and B2SAFE

As an illustration of the process described above, we look at the data sharing service B2SHARE. When authenticating to B2SHARE, the user is directed to B2ACCESS and authenticates via an IdP, say, a SAML IdP. B2SHARE supports OpenID Connect, so B2ACCESS converts the credential into a token which is presented to B2SHARE as an (anonymised) proof of identity. When it needs further attributes, B2SHARE obtains them from B2ACCESS via the "userinfo" API. We shall return to this example in Section 4.4.

When the user logs in, the SAML credential presented by their IdP is also converted into a short-lived X.509 credential.

B2SAFE needs an X.509 certificate. Typically, such a service is accessed through a portal, either one dedicated to the service, or as a feature in the user's community portal. In this case, the *portal* generates the key pair and the certificate request, sends the request to B2ACCESS, and waits for B2ACCESS to return the X.509 certificate. B2ACCESS signs the certificate
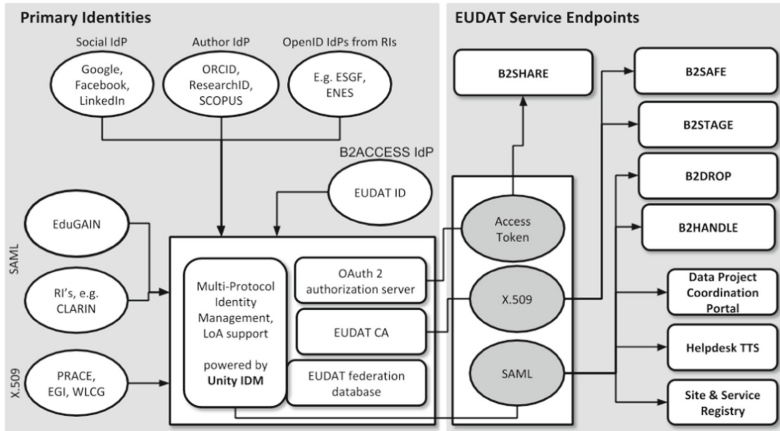
Fig. 2 B2ACCESS: EUDAT AAI federated user authentication and management components and the target B2* services

when the user has authenticated, and embeds relevant attributes into the certificate. For users requiring command line access (e.g. to B2SAFE), B2ACCESS can also generate the key pair and certificate itself, and let the user download both. The user then installs them locally and uses their command line tool. In its current implementation, B2ACCESS supports command line tools for services that use X.509, or for OAuth (via a bearer token).

Generally, converted credentials are only valid for a short period of time (hours instead of days), because they are managed on the user's behalf by services, they are not held by the users themselves (Fig. 3).

### 4.1.2 Attribute Harmonisation

Since B2ACCESS accepts identities from many external IdPs, and different IdPs have different attribute release policies, the incoming set of attributes is very likely heterogeneous. This makes it difficult for SPs to define authorization policies. As mentioned above, B2ACCESS acts as a proxy and tries to harmonize all incoming attribute information. This may imply mapping attribute values from other schemata onto attributes in the EUDAT attribute schema. If any essential attributes are not released by the IdP, B2ACCESS will ask the user to supply these attributes during the initial registration step.

Since users can be asked to supply values for missing attributes, and it is not considered feasible for the B2ACCESS operators to check all these values, we have concluded that a LoA per external IdP is not sufficient, but a LoA per attributes is needed, at least for the more important attributes such as e-mail or organisational affiliation. An attribute provided by a high LoA IdP gets assigned a high LoA while a user-supplied attribute value gets a low(er) level LoA. This is currently under development.

Table 3 EUDAT Attributes

Name	Mandatory	Description
urn:oid:2.5.4.49, distinguishedName	YES	Distinguished name (DN)
unity:persistent	YES	Persistent identifier
urn:oid:2.5.4.3, cn	YES	Common name
urn:oid:1.2.840. 113549.1.9.1, userName	YES	Principal
urn:oid:2.5.4.10, o	YES	Organisational affiliation
email	YES	E-mail address
memberOf	NO	The service will perform the authorisation decision based on these roles.
loa	YES	Level of assurance

Towards Federated Service Discovery and Identity Management in Collaborative Data...

Fig. 3 Credential translation
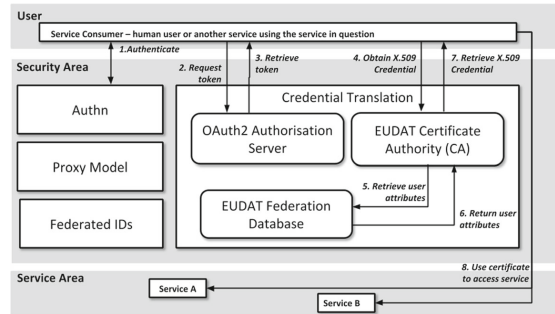
In the current implementation, however, there is a single LoA attribute, namely the LoA associated with the user's (external) IdP (as determined by B2ACCESS operators; we do not ask IdPs to publish their LoA and would not necessarily trust the value if they did.) Typically, X.509 and Academic SAML IdPs are assigned a high level of assurance while the social and direct B2ACCESS IdPs are assigned a lower level of assurance.

4.1.3 Group Management

EUDAT consists of many service providers offering a wide range of services and tools. Some of these tools are publicly accessible, but most apply authorization to at least some of the actions which can be performed in that service or tool. As mentioned earlier, attributes released by B2ACCESS, and group membership especially, are used in these authorization policies. To provide fine-grained control, a hierarchical group structure has been defined providing: (1) a high-level domain directly under the root, defining the infrastructure, community or project, (2) multiple service level domains as children of a high-level domain, one for each service that falls under that specific high-level domain, and (3) the freedom for administrators to define anything below the service level domain to cater for any service-specific needs.

Administrators can be defined on any level to ease the administrative burden of managing the group membership. Typically, the main B2ACCESS administrators have permission in B2ACCESS to manage all groups, including the high-level domains.

4.2 Distributed Authorization within EUDAT

To fulfil the requirements mentioned in R5, a solution based on XACML is under development, based on a proposed architecture shown in Fig. 4. This architecture allows for harmonised management of the XACML policies in the central service Policy Administration Point (PAP). Multiple instances of a single service can be deployed across data centres; thus, there is the need to run a central PAP and Policy Repository (PR) combination to harmonise authorisation policies for the service as a whole, covering all instances running at the individual data centres. The central service PR is replicated to the EUDAT data centres. Each data centre has a local PRs. Changes are only pushed from the central service PR to the local PRs.

Each EUDAT centre is running a Policy Decision Point (PDP) with access to the local PR and each B2-service has a Policy Enforcement Point (PEP) which communicates with the centres PDP. This allows for authorization decisions even if the central service PR is unavailable.

An additional function of the central PAP/PR is to provide a ingest endpoint which can be used to ingest XACML policies from external sources, such as community repositories.

4.3 Federated Service Discovery with the EMI Service Registry (EMIR)

The infrastructures (EUDAT, EGI, CLARIN in our case) offer different types of services: cloud, compute, data, authentication, authorisation, etc. The EMI Service Registry (EMIR) has been designed and implemented in the European Middleware Initiative

Fig. 4 The EUDAT
XACML-based distributed
authorization service
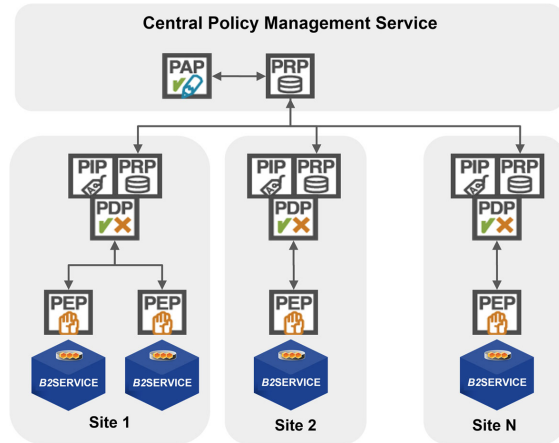
(EMI) [25] project. EMIR aims to provide robust service discovery within large scale infrastructures [27]. The initial implementation was driven by major European grid computing middlewares (UNICORE, Advanced Resource Connector (ARC), gLite, and dCache). However, the scope of the service provisioning and discovery within EMIR is not limited to grid and therefore offers a versatile service discovery utility adequate from small- to large-scale data and cloud infrastructures. The details of EMIR are described in the following subsections (see also Fig. 5).

4.3.1 Concepts

The core notion of EMIR is to enable discovery of services. The set of services can be grouped in a *domain* (such as an NGI), and multiple domains can be organised in a hierarchical structure. The domain is an autonomous entity and can be connected with other domains in a hierarchy to form a *federation*. The top-level domain can replicate its information to other top-level domains in a peer-like fashion. The replication of information at the root of the

Fig. 5 Federated service
registry: Service discovery
in heterogeneous federated
infrastructure
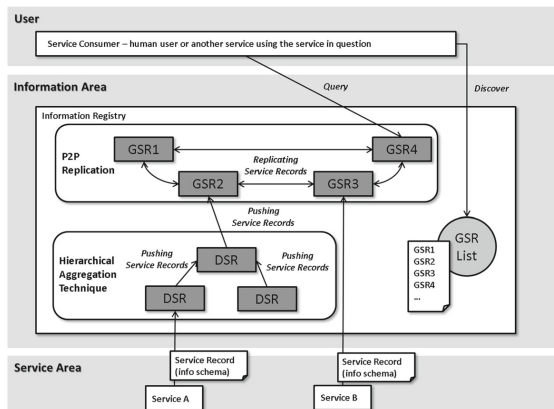
hierarchy makes the federation infrastructure resilient to failures.

EMIR is based on two main components: the Domain Service Registry (DSR) and the Global Service Registry (GSR). The primary difference between the two depends on their position in the hierarchy. The DSR represents any node in the hierarchy, while the GSR always sits on the top (root node). The service to be discovered is published through a Service Record (SR) using the OGF GLUE 2.0 [47] standard.

4.3.2 Service Information Model

According to the requirement R15 of a common service information model, the registry should be capable of representing the infrastructure services of any type. It could be a service having any of the (storage, cloud, network, HPC, etc.) capabilities that may dynamically (dis)appear within an e-infrastructure. In order to address the service discovery use cases from the large spectrum of scientific domains, EMIR adopts the standard GLUE 2.0 information model [47]. Since GLUE is an information model and does not provide a normative realisation, the Open Grid Forum recommendations [48] and [49] were used as a foundation to implement the service registry in the XML and JSON format, respectively. For the latter, the emerging JSON-Spec standard (similar XSD for XML) is used for the implementation. Since the GLUE model can become very extensive, in order to be concise and yet extensible, only the abstract representation (or entities) is taken into account and forms the basis of EMIR's *Service Record (SR)*. Table 4 shows a subset of the mandatory attributes that represent a service. The JSON record in Listing 1 shows a minimal B2SHARE instance.

4.3.3 Hierarchical Aggregation

EMIR allows creating flexible registry hierarchies of DSR nodes with GSR on top. Figure 6 illustrates a simplified hierarchical aggregation model where the service records are published from a leaf node (a service publisher) and traverse the DSRs to the root GSR node. The top level GSR node is *eventually consistent* [51]; however, due to the network latency of service records being published, the freshness of information could be affected. While designing the registry, two major factors must be taken into account:

Table 4 Service record schema containing a set of core service attributes [38]

Attribute name	Description
Service ID	A globally unique identifier for the service
Name	Human-readable name
Endpoint URL	Location to access the service
Capability	An array of offered capabilities
Service technology	The technology used to implement the service
Service time-to-live (TTL)	The visibility of the service within an infrastructure
Service type	Service type according to namespace-based classification
Service version	Specific service version
Service health	Monitoring information about service state

- The registries are geographically distributed across different administrative domains. In order to cope with intermediary (availability or network) failure of nodes, an in-memory database (dotted database icon in Fig. 6) is used that captures the (un-synchronised) modifications.
- A service record may contain a variety of project- or virtual organisation-specific information (apart from what has been mentioned in Table 4). Therefore, unlike conventional SQL, a schema-free or NoSQL approach (using MongoDB [39]) has been implemented. The database also offers horizontal scalability to distribute the large number of service records over multiple database instances.

4.3.4 A Peer-to-Peer Approach to the Replication of GSRs

The notion of replication of GSR top level registry nodes in a hierarchy is based on the Pastry algorithm [46] and inspired by the ISIS [40] algorithm used in the ARC middleware Peer-to-Peer (P2P) information system. Unlike the basic structured P2P concepts of distributing the keys on an overlay network, and non-structured approaches of replicating the information [28], EMIR slightly modifies the algorithm and replicates the keys among the peer GSR nodes in the network and makes the information eventually

S. Memon et al.

Listing 1 Service record in GLUE 2.0 JSON format

```
{
    "Service_Name":"B2DROP",
    "Service_Type":"eu.eudat.b2drop",
    "Service_Capability":["data sharing"],
    "Service_Endpoint_URL":"http://b2drop.eudat.eu",
    "Service_Endpoint_Technology":"technology",
    "Service_Endpoint_InterfaceVersion":["v1.0"],
    "Service_Endpoint_HealthState":"ok",
}
```

consistent [51] after a certain period of time. By replicating the information, all the services can be discovered from any of the available GSRs, which makes the infrastructure resilient to bottlenecks and failures.

The *sparsity*, the number of neighbours each P2P node should replicate to, is another key factor (Fig. 7). Selecting a smaller value would consume less bandwidth at a given time but take longer to reach consistency.

4.3.5 Authentication and Authorisation

The DSR and GSR nodes expose a programmatic interface to the service publishers, as well as to the applications, to publish and query service records. In addition, the nodes must connect with the other nodes
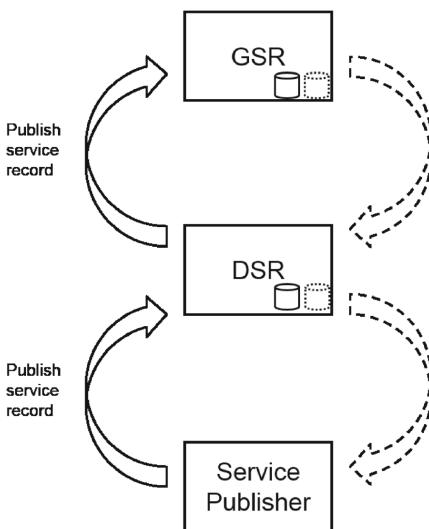
**Fig. 6** Hierarchical (bottom to top) aggregation of information

to form a hierarchy or a P2P network. Publishing a service requires a high LoA credential (X.509 certificate), so attackers can not inject malicious services or modify existing services, so all EMIR nodes, and all entities authorised to publish services, must have X.509 certificates issued by a trusted authority.

4.4 Overall Architecture

Figure 8 is an updated version of Fig. 1, showing the details of the two middle rectangles. To look at this process in more detail, we return to our example from Section 4.1.1. Figure 9 depicts a sequence diagram, with the following steps:

1. A CLARIN user requires a EUDAT data sharing service to deposit her data, and therefore send queries for the "data sharing" service types to EMIR.
2. The user sends a request of depositing their research dataset on B2SHARE.
3. As the access token is missing from the user's request, B2SHARE will redirect (using the HTTP protocol) the user to the B2ACCESS service, the authenticating party, and then further to the user's organisation IdP.
4. The user authenticates themselves, here with a username and password, to the IdP.
5. We assume for this use case that the user is already registered with the B2ACCESS service, so B2ACCESS will not attempt to register them. Instead, B2ACCESS updates its information about the user, if necessary, based on the user attributes in the SAML assertion which has been received from the IdP.
6. The user (or rather the user's browser) receives and then forwards an authorisation code to the B2SHARE service. On the basis of the code, B2SHARE requests an access token from B2ACCESS.

Fig. 7 EMIR P2P network of registries
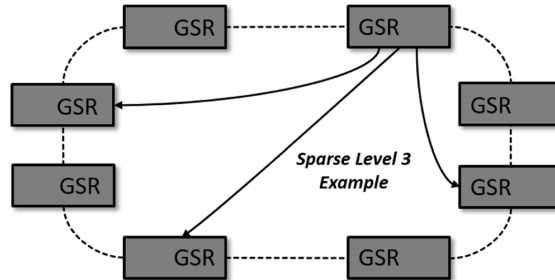
7. B2SHARE receives an access token.
8. B2SHARE validates the access token and eventually grants the user to deposit/publish/share her data. The data stored on the EUDAT resources (B2SHARE) should now be replicated across multiple storage systems.
9. In order to replicate data with B2SAFE, B2SHARE requires a X.509 credential and sends a request and the access token (from previous flow) to the B2ACCESS Certification Authority (CA) server. The CA server validates the access token and the request.
10. The CA requests a full set of attributes (containing the user's role, group, email, etc.) from the B2ACCESS database.
11. A short-lived X.509 credential is generated, containing the user's attributes in its extensions, and returned to the B2SHARE service.

12. The B2SHARE service can now replicate the data to the relevant B2SAFE nodes.

4.4.1 Cross-Infrastructure Federated Service Access

Figure 10 extends the example in Section 4.4. We should point out that the scenario is not possible today; a few components are still missing. Nevertheless, it is instructive, as the missing pieces will help us understand the barriers to interoperation.

Let us assume that a CLARIN user is in possession of a corpus, and wishes to work on a particular data set from it, consisting of video and image data, and the work will result in annotations.

1. The user looks up data exchange services and corpus annotation services on EMIR, and EMIR returns a list of endpoints on multiple infrastructures.
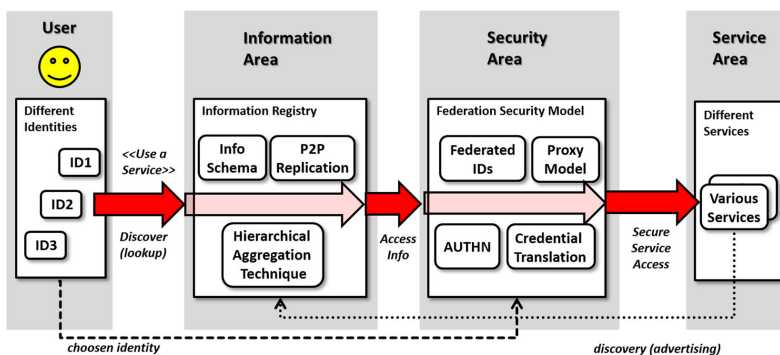

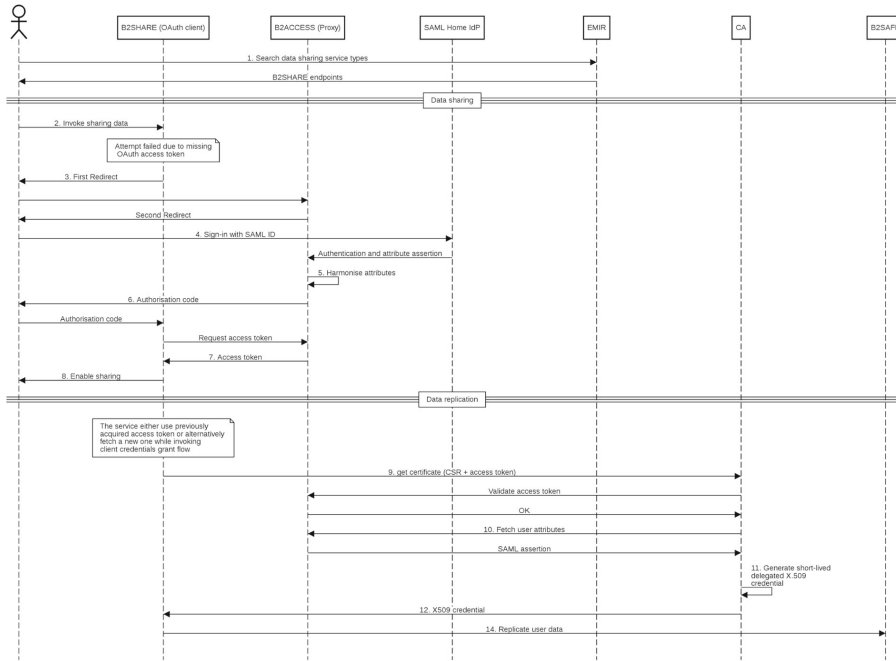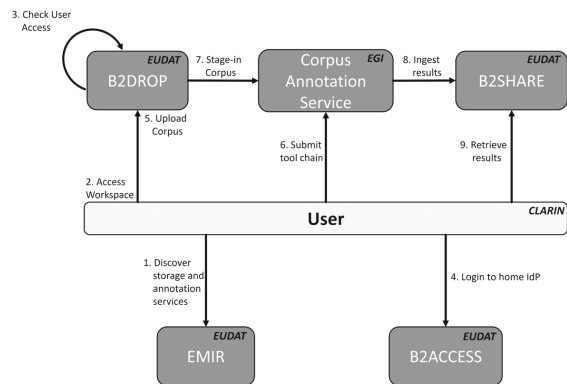**Fig. 8** Integrated federated authentication and service discovery architecture

Fig. 9 Sequence diagram showing service discovery, federated authentication, credential translation, and attribute harmonisation in the EUDAT infrastructure

Fig. 10 CLARIN data staging use case showing cross-infrastructure federated authentication and service discovery

2. The user selects a B2DROP endpoint (an EUDAT data exchange service [4]) and tries to access its workspace to upload the corpus.
3. B2DROP checks whether the user is authenticated and redirects to B2ACCESS service as before.
4. After successful authentication to B2SAFE, the user uploads the corpus and obtains a unique reference to the corpus. Upload is completed through an OAuth token.
5. The user selects Corpus Annotation Service (CAS) from EGI, and submits a compute request which includes the reference to the corpus, as well as a (bearer) token that authorises the service to access the relevant part of the corpus. The user authenticates to EGI using their EUDAT certificate.[1]
6. CAS retrieves (through the provided reference and token) the corpus and processes the relevant parts of it.
7. B2SHARE receives the processed (annotated) data from the CAS service. Here, B2SHARE does not have any prior authorisation from the user, nor does it have the option to ask for one (as CAS is running without the user's direct intervention). Thus, CAS needs to upload data using the delegated certificate. Note that the certificate also contains the e-mail address as metadata, so the service is able to notify the user of the upload, including of course a link to the data.
8. Finally, the user fetches the annotated corpus from the provided link.

As we have mentioned, unlike the scenario in Section 4.4, the scenario above is an aim. It is not possible today, but it serves to highlight the current gaps:

– EUDAT and EGI must both accept the same certificates. Unlike IGTF certificates (www.igtf.net), certificates generated for federation-internal use are not trusted across infrastructures, due to the variation in LoA. Work is in progress to harmonise on RCauth [43].
– Likewise, we have, in this scenario, skipped lightly over the authorisation process. In practice,

EGI would allocate resources to the community, thus requiring community membership attributes to be communicated *with the credential* because an EGI service would not *a priori* be authorised to query user attributes from B2ACCESS. In fact, these attributes are currently communicated with the credential, but will not be after a migration to RCauth. In other words, cross-infrastructure authorisation needs a lot more thought.

– B2ACCESS provides consistent user mapping across OAuth/OpenID Connect credentials and certificates. In the scenario above, a service would sometimes need to use one, sometimes the other: B2SHARE would need to accept a certificate from CAS, but OpenID Connect from the user's browser in step 8. In the current infrastructures, services either use one or the other, but not both.
– The current production instance of B2DROP is, as of this writing, not integrated with B2ACCESS.
– An EMIR service is needed which aggregates services across all three infrastructures. Note that there is no access control on querying service information.
– As with resource allocation, accounting also needs to be consistent.

5 Discussion

This paper presented a federated AAI and service discovery framework. The B2ACCESS service implementing the AAI presented in this paper fulfils requirements R1, R4, R2, R7, R6, R8, and R9 of CLARIN and EUDAT because it manages authentication, user attributes, and credential translation in one service. In addition to that, EMIR addresses requirements R3, R15, R16, R18, R19, and R20 by offering a robust service discovery for EGI infrastructure (or alike). In particular, it combines a hierarchical model that allows subdomains to manage their resources with a peer-to-peer model across the top-level nodes.

In the context of EMIR, the registry nodes are relatively static in nature, so they can rely on PKI and Access Control Lists (ACLs) for authentication and authorisation, respectively. This requires a communication between the administrators to exchange the nodes' information.

---

[1] If B2DROP had used certificates, the EGI service could have used its delegated certificate to access the data.

In terms of B2ACCESS, there are a number of areas (while liaising with EUDAT and AARC) in the future to look into:

– Connecting infrastructures through shared (mutually trusted) authentication. Harmonised communication of the LoA will be useful, which is work in progress through the REFEDS work.
– Supporting multiple LoA and also providing a standard means (e.g. step-up authentication) to augment the assurance levels. Work in progress in AARC should provide guidance on this.
– Integration of a fine-grained and externalised authorisation system based on the XACML standard. However, as we saw in Section 4.4.1, much more research is needed in cross-infrastructure authorisation.
– Unsurprisingly, heterogeneous services which need several different "flavours" of credentials (as in EUDAT) make it harder to build cross-infrastructure (or indeed inter-infrastructure) interoperation.

5.1 Impact on Infrastructures

B2ACCESS already provides production-ready AAI for EUDAT infrastructure,[2] which implies integration as well as enabling federated access (using federated identities) to all the B2 services, with dissimilar authentication protocols (SAML, OIDC, PKI). Given the adoption of B2ACCESS in EUDAT, other scientific communities such as EPOS are also considering to deploy B2ACCESS (independently from EUDAT) in their own research infrastructure. B2ACCESS being EUDAT AAI plays an important role within the AARC consortium as one of its objectives is to achieve interoperability of B2ACCESS across e/cyber/research-infrastructures, such as EGI, PRACE and ELIXIR identity and service federations. This is, however, more than an interoperability exercise as (in particular) EGI and EUDAT will have to collaborate by sharing their services within the future EU-funded EOSC-Hub project, the successor of the EUDAT project. Alongside the interoperation, B2ACCESS has fed its experiences into building the AARC Blueprint Architecture [10]. Being an SP/IdP proxy, B2ACCESS has significantly reduced the

---

[2]https://b2access.eudat.eu

Springer

barrier of trust management between service and identity federations. There are also risks when users' identity is compromised and since EUDAT hosts and manages data from scientific communities, the attacker can delete or rewrite users' datasets with arbitrary data. To cope with such attacks, B2ACCESS adopts the SIRTFI [6] framework to react immediately and mitigate the risks. While the users and services are provisioned into the EUDAT's B2ACCESS service, the registration goes through a formal process for approval by the B2ACCESS administrators, to check whether the identity is compliant with EUDAT policies. As for EMIR, it has been integrated with all the services which are included in the EMI services catalogue, thus it has enabled publishing and querying of the services by the infrastructure operators, monitoring systems and other services (for example workflow). However, EMIR is also being evaluated for service discovery purposes within the EUDAT and EGI infrastructures.

5.2 Impact on Users

With B2ACCESS in EUDAT, end users from various scientific communities (CLARIN, ELIXIR, DARIAH, TERENO, EPOS, ENES, etc.) possessing a single identity have federated access to the EUDAT's B2 services. The underlying credentials of the users can be SAML ID, Social ID (from Google, Facebook or ORCID) and X.509 certificates. The EUDAT services do not rely on any single authentication protocol, thus B2ACCESS enabled the authentication by translation of credentials. As far as service discovery is concerned, EMIR has facilitated the users and software clients by querying of infrastructure services based on the service metadata (service type and capabilities).

6 Related Work

Federated Identity Management (FIM) or AAI in a broader sense has been a challenge for many years [33]; though the social, commercial and research application providers are recently getting more traction towards external rather than built-in identity management solutions. It is also pertinent for a collaborative infrastructure like EUDAT, providing secure and federated data management services to the research

communities [5, 12, 15, 35] in which the earth scientists or linguists would want to collaborate (for example share their data) within or across research communities, given each communities have already their established external or internal identity management system in place, so they bring their own identities.

ELIXIR is one of the largest research infrastructures in Europe, having their own data and identity management infrastructure. The main goals of ELIXIR are to orchestrate the collection, quality control and archiving of large amounts of biological data produced by life science experiments. It also has an aim to improve the long-term sustainability of biological datasets [23]. The ELIXIR AAI [24] provides web identity federation while integrating with Global Authentication INfrastrcuture (eduGAIN) [20] inter-federation service. In addition to that, the AAI allows users to authenticate with their social identities, which are issued from Google, ORCID and Facebook. It supports associating remote user identities with infrastructure-wide identifiers. Unlike B2ACCESS, ELIXIR AAI lacks support for credentials translation. Similarly, ELIXIR's support for multiple authentication protocols is limited, hence it does not provide end user authentication with end-entity X.509 certificates and LDAP based credentials.

XSEDE [52] is the successor to TeraGrid [36], an NSF funded HPC and grid infrastructure. It consists of a collection of advanced digital resources and services (like supercomputers, visualization and storage systems, collections of data, software, networks, and expert support) that support researchers in various scientific domains. XSEDE relies on Globus Auth [50], a framework for identity and access management. Like B2ACCESS, the Globus Auth framework allows integration with SAML-based identity federations, identity linking, identity brokering (or credential translation) and group management. Furthermore, Globus Auth uses MyProxy-based CILogon [7, 41] to enable federated access to non-browser-based resources, which in particular rely on short-lived X.509 credentials. B2ACCESS instead uses its own online CA to generate the short-lived credentials. However, integration of B2ACCESS with RCAuth [43] (a modified version of CILogon service for European infrastructures) is being tested and evaluated, but will have consequences, as mentioned above.

7 Conclusions

In recent years, large-scale infrastructures have substantially evolved where the federated service discovery and access have become increasingly relevant. Users benefit from having a single credential across the whole infrastructure, and benefit further when it is used across multiple infrastructures. With a unified approach to identity management, authentication, authorisation, and accounting, users are able to run workflows and access and store data from one infrastructure to another, thus further enabling user communities and service providers to build more sophisticated services. As with the registry of services for a country, it should be feasible in the near future to extend these into hierarchies of services, similar to the current global grid infrastructures. However, the details matter, and different technologies, varying levels of assurance, different protocols, schemata, conventions and culture can all provide gaps that prevent users from seamlessly interoperating services across infrastructures. However, as we have seen in the present paper, many of the required building blocks are already present, as is the will to interoperate. Also helpful are the harmonisation activities by REFEDS and AARC, and, if needed, the opportunity for standardisation through standards-defining organisations such as DMTF and OGF.

References

1. Authentication and authorisation research consortium. https://aarc-project.eu. Accessed: 19 Nov 2016
2. Alfieri, R., Cecchini, R., Ciaschini, V., dell'Agnello, L., Frohner, Á., Gianoli, A., Lörentey, K., Spataro, F.: Voms, an authorization system for virtual organizations. In: Rivera, F.F., Bubak, M., Gómez-Tato, A., Doallo, R. (eds.) Grid Computing, First European Across Grids Conference, Santiago de Compostela, Spain, February 13–14, 2003, Revised Papers, Lecture Notes in Computer Science, vol. 2970, pp. 33–40. Springer, Berlin (2003). https://doi.org/10.1007/978-3-540-24689-3_5
3. Allcock, W., Bresnahan, J., Kettimuthu, R., Link, M., Dumitrescu, C., Raicu, I., Foster, I.: The Globus striped GridFTP framework and server. In: Proceedings of the 2005 ACM/IEEE Conference on Supercomputing, SC

'05. IEEE Computer Society, Washington, DC (2005). https://doi.org/10.1109/SC.2005.72

4. B2DROP. https://www.eudat.eu/services/b2drop. Accessed: 5 Jan 2017

5. Bailo, D., Jeffery, K.G., Spinuso, A., Fiameni, G.: Interoperability oriented architecture: the approach of epos for solid earth e-infrastructures. In: 2015 IEEE 11th International Conference on e-Science, pp. 529–534 (2015). https://doi.org/10.1109/eScience.2015.22

6. Barton, T., Basney, J., Groep, D., Harris, N., Johansson, L., Kelsey, D., Koranda, S., Wartel, R., West, A., Short, H.: A security incident response trust framework for federated identity (sirtfi). Recommendation Siftfi-1.0. REFEDS. https://refeds.org/wp-content/uploads/2016/01/Sirtfi-1.0.pdf (2015)

7. Basney, J., Fleury, T., Gaynor, J.: Cilogon: a federated x.509 certification authority for cyberinfrastructure logon. Concurr. Comput.: Pract. Exp. **26**(13), 2225–2239 (2014). https://doi.org/10.1002/cpe.3265. CPE-13-0334.R1

8. Baur, T., Breu, R., Kálmán, T., Lindinger, T., Milbert, A., Poghosyan, G., Reiser, H., Romberg, M.: An interoperable grid information system for integrated resource monitoring based on virtual organizations. J. Grid Comput. **7**(3), 319–333 (2009). https://doi.org/10.1007/s10723-009-9134-3

9. Grid information system. http://gridinfo.web.cern.ch. Accessed: 5 Sep 2017

10. Biancini, A., Florio, L., Haase, M., Hardt, M., Jankowski, M., Jensen, J., Kanellopoulos, C., Liampotis, N., Licehammer, S., Memon, S., van Dijk, N., Paetow, S., Prochazka, M., Sallé, R., Solagna, P., Stevanovic, U., Vaghetti, D.: AARC: first draft of the blueprint architecture for authentication and authorisation infrastructures. CoRR arXiv:1611.07832 (2016)

11. Blumtritt, J., Elbers, W., Goosen, T., Hinrichs, M., Qiu, W., Sall, M., Windhouwer, M.: User delegation in the CLARIN infrastructure. In: Selected Papers from the CLARIN 2014 Conference, October 24–25, 2014, Soesterberg, The Netherlands. Linköping University Electronic Press, Linköping (2015). http://www.ep.liu.se/ecp/article.asp?issue=116&volume=&article=002

12. Bogena, H.: Tereno: German network of terrestrial environmental observatories. J. Large-Scale Res. Facil. **2**, A52 (2016). https://doi.org/10.17815/jlsrf-2-98. http://jlsrf.org/index.php/lsf/article/view/98

13. Chadwick, D.W., Siu, K., Lee, C., Fouillat, Y., Germonville, D.: Adding federated identity management to openstack. J. Grid Comput. **12**(1), 3–27 (2014). https://doi.org/10.1007/s10723-013-9283-2

14. Christos, K., Nicolas, L., van Dijk N., Peter, S.: Deliverable djra1.1: analysis of user community and service provider requirements. Project Deliverable AARC-DJRA1.1, AARC Project. https://aarc-project.eu/wp-content/uploads/2015/10/AARC-DJRA1.1.pdf (2015)

15. CLARIN. https://www.clarin.eu. Accessed: 13 July 2017

16. CLARIN services. https://www.clarin.eu/content/services. Accessed: 5 Sep 2017

17. Conway, M., Moore, R., Rajasekar, A., Nief, J.Y.: Demonstration of policy-guided data preservation using iRODS. In: 2011 IEEE International Symposium on Policies for Distributed Systems and Networks (POLICY), pp. 173–174 (2011). https://doi.org/10.1109/POLICY.2011.17

18. Cornwall, L.A., Jensen, J., Kelsey, D.P., Frohner, Á., Kouřil, D., Bonnassieux, F., Nicoud, S., Lőrentey, K., Hahkala, J., Silander, M., Cecchini, R., Ciaschini, V., dell'Agnello, L., Spataro, F., O'Callaghan, D., Mulmo, O., Volpato, G.L., Groep, D., Steenbakkers, M., McNab, A.: Authentication and authorization mechanisms for multidomain grid environments. J. Grid Comput. **2**(4), 301–311 (2004). https://doi.org/10.1007/s10723-004-8182-y

19. Drollette, D.: Standards are the glue 2.0. iSGTW (ScienceNode). https://sciencenode.org/feature/isgtw-feature-standards-are-glue-20.php (2009)

20. eduGAIN. http://www.edugain.org. Accessed: 10 Aug 2017

21. EGI. http://www.egi.eu. Accessed: 5 Sep 2017

22. Federated cloud information discovery. https://wiki.egi.eu/wiki/Federated_Cloud_Information_Discovery. Accessed: 5 Sep 2017

23. ELIXIR. https://www.elixir-europe.org. Accessed: 15 Sept 2017

24. ELIXIR AAI documentation. https://www.elixir-europe.org/services/compute/aai. Accessed: 13 Sep 2017

25. European Middleware Initiative (EMI). http://www.eu-emi.eu. Accessed: 10 June 2016

26. EUDAT collaborative data infrastructure. http://www.eudat.eu. Accessed: 2 Sep 2016

27. Field, L., Memon, A.S., Márton, I., Szigeti, G.: The EMI registry: discovering services in a federated world. J. Grid Comput. **12**(1), 29–40 (2014). https://doi.org/10.1007/s10723-013-9284-1

28. Forestiero, A., Mastroianni, C., Spezzano, G.: Building a peer-to-peer information system in grids via self-organizing agents. J. Grid Comput. **6**(2), 125–140 (2008). https://doi.org/10.1007/s10723-007-9062-z

29. Foster, I.: Globus toolkit version 4: software for service-oriented systems. In: Proceedings of the 2005 IFIP International Conference on Network and Parallel Computing, NPC'05, pp. 2–13. Springer, Berlin (2005). https://doi.org/10.1007/11577188_2

30. Foster, I.: Globus online: accelerating and democratizing science through cloud-based services. IEEE Internet Comput. **15**(3), 70–73 (2011). https://doi.org/10.1109/MIC.2011.64

31. Foster, I, Kesselman, C., Tuecke, S.: The anatomy of the grid: enabling scalable virtual organizations. Int. J. High Perform. Comput. Appl. **15**(3), 200–222 (2001)

32. Hardt, M. (ed.) C.K.: Blueprint architecture. Project deliverable, AARC Project. https://aarc-project.eu/wp-content/uploads/2017/04/AARCBPA-2017.pdf (2017)

33. Jensen, J.: Federated identity management challenges. In: 2012 Seventh International Conference on Availability, Reliability and Security, pp. 230–235 (2012). https://doi.org/10.1109/ARES.2012.68

34. Jensen, J., Stevanovic, U., Kakavas, I., Liampotis, N., Haase, M., Gietz, P., Jankowski, M., Reale, M., Mantovani, M.L., Florio, L.: Design for deploying solutions for "guest identities". Project milestone, AARC Project. https://aarc-project.eu/wp-content/uploads/2016/06/MJRA1.2-Design-for-Deploying-Solutions-for-Guest-Identities.pdf (2016)

35. Joussaume, S., Budich, R.: The Infrastructure Project of the European Network for Earth System Modelling: IS-ENES,

Towards Federated Service Discovery and Identity Management in Collaborative Data...

pp. 5–9. Springer, Berlin (2013). https://doi.org/10.1007/978-3-642-36597-3_2

36. Katz, D.S., Callaghan, S., Harkness, R., Pamidighantam, S., Pierce, M., Plale, B., Song, C., Towns, J.: Science on the teragrid. Special Issue 2010 81–97 (2010)

37. Mathieu, G., Richards, D.A., Gordon, D.J., Novales, C.D.C., Colclough, P., Viljoen, M.: Gocdb, a topology repository for a worldwide grid infrastructure. J. Phys. Conf. Ser. **219**(6), 062021 (2010). http://stacks.iop.org/1742-6596/219/i=6/a=062021

38. Memon, A.S., Riedel, M., Field, L., Szigeti, G., Marton, I.: EMIR: an EMI Service Registry for Federated Grid Infrastructures. In: EGI Community Forum 2012/EMI Second Technical Conference, Munich (Germany), 26 Mar 2012–30 Mar, 2012. Proceedings of Science, Sissa (2012). http://pos.sissa.it/archive/conferences/162/073/EGICF12-EMITC2_073.pdf

39. MongoDB for GIANT Ideas. https://www.mongodb.com. Accessed: 5 Sep 2017

40. NorduGrid: ARC peer-to-peer information system. Documentation and developer's guide NORDUGRID-TECH-21. NorduGrid. http://www.nordugrid.org/documents/infosys_technical.pdf (2013)

41. Novotny, J., Tuecke, S., Welch, V.: An online credential repository for the grid: Myproxy. In: Proceedings 10th IEEE International Symposium on High Performance Distributed Computing, pp. 104–111 (2001). https://doi.org/10.1109/HPDC.2001.945181

42. Parducci, B., Lockhart, H., Rissanen, E.: Extensible access control markup language (XACML) version 3.0. OASIS Standard xacml-3.0-core-spec-en. OASIS. http://docs.oasis-open.org/xacml/3.0/xacml-3.0-core-spec-en.pdf (2013)

43. Research and Collaboration Authentication Certification Authority Service. https://www.rcauth.eu. Accessed: 16 Sep 2017

44. van Rijn, A., Vandenbroucke, R.: Guide to e-infrastructure requirements for european research infrastructures. ISBN 978-90-823661-5-0, E-IRG. http://e-irg.eu/catalogue/eirg-1004 (2017)

45. Robertson, L.: Computing Services for LHC: from Clusters to Grids, pp. 69–89. Springer, Berlin (2012)

46. Rowstron, A.I.T., Druschel, P.: Pastry: scalable, decentralized object location, and routing for large-scale peer-to-peer systems. In: Proceedings of the IFIP/ACM International Conference on Distributed Systems Platforms, Middleware '01, pp. 329–350. Springer, London (2001). http://dl.acm.org/citation.cfm?id=646591.697650

47. Sergio, A., Burke, S., Ehm, F., Field, L., Galang, G., Konya, B., Litmaath, M., Millar, P., Navarro, J.P.: GLUE specification v. 2.0. Recommendation GFD-R-P.147, Open Grid Forum. https://www.ogf.org/documents/GFD.147.pdf (2009)

48. Sergio, A., Burke, S., Field, L., Konya, B., Memon, A.S., Meredith, D., Navarro, J.P., Paganelli, F., Smith, W.: GLUE v, 2.0—reference realisation to XML schema. Recommendation GFD.209, Open Grid Forum. https://www.ogf.org/documents/GFD.209.pdf (2013)

49. Smith, W., Meredith, D., Memon, A.S., Navarro, J.P.: GLUE v, 2.0—reference realisation to JSON schema. Recommendation GFD-RP.219, Open Grid Forum. https://www.ogf.org/documents/GFD.219.pdf (2015)

50. Tuecke, S., Ananthakrishnan, R., Chard, K., Lidman, M., McCollam, B., Rosen, S., Foster, I.: Globus auth: a research identity and access management platform. In: 2016 IEEE 12th International Conference on e-Science (e-Science), pp. 203–212 (2016). https://doi.org/10.1109/eScience.2016.7870901

51. Vogels, W.: Eventually consistent. Commun. ACM **52**(1), 40–44 (2009). https://doi.org/10.1145/1435417.1435432. http://doi.acm.org/10.1145/1435417.1435432

52. XSEDE. https://www.xsede.org. Accessed: 13 Sep 2017

53. Zinn, C., Hinrichs, M., Dima, E., van Uytvanck, D.: CLARIN switchboard specification. CE-2015-0684, CLARIN. https://office.clarin.eu/v/CE-2015-0684-LR_switchboard_spec.pdf (2015)

References

[1] W. Allcock, J. Bresnahan, R. Kettimuthu, M. Link, C. Dumitrescu, I. Raicu, and I. Foster. The Globus Striped GridFTP Framework and Server. In *Proceedings of the 2005 ACM/IEEE Conference on Supercomputing*, SC '05, Washington, DC, USA, 2005. IEEE Computer Society.

[2] S. Andreozzi, S. Burke, F. Donno, L. Field, S. Fisher, J. Jensen, B. Konya, M. Litmaath, M. Mambelli, J. M. Schopf, M. Viljoen, A. Wilson, and R. Zappi. GLUE Schema Specification v1.3. Protocol Specification, Open Grid Forum, June 2007.

[3] S. Andreozzi, S. Burke, F. Ehm, L. Field, G. Galang, D. Horat, B. Konya, M. Litmaath, S. Memon, P. Millar, J. Navarro, and F. Paganelli. GLUE v, 2.0 – Reference realisation to LDAP Schema. Protocol Specification, GFD-R-P.218, Open Grid Forum, December 2015.

[4] A. Anjum, M. Sporny, and A. Sill. Blockchain Standards for Compliance and Trust. *IEEE Cloud Computing*, 4(4):84–90, July 2017.

[5] Apache Software Foundation. OpenAZ. `http://incubator.apache.org/projects/openaz.html`. [Online; Accessed 27 December 2020].

[6] O. Appleton, D. Cameron, J. Cernak, P. Dóbé, M. Ellert, T. Frågåt, M. Grønager, D. Johansson, J. Jönemo, J. Kleist, M. Kočan, A. Konstantinov, B. Kónya, I. Márton, B. Mohn, S. Möller, H. Müller, Z. Nagy, J. K. Nilsen, F. Ould Saada, K. Pajchel, W. Qiang, A. Read, P. Rosendahl, G. Rőczei, M. Savko, M. Skou Andersen, O. Smirnova, P. Stefán, F. Szalai, A. Taga, S. Z. Toor, A. Wäänänen, and X. Zhou. The next-generation ARC middleware. *Annals of telecommunications – Annales des télécommunications*, 65(11):771–776, Dec 2010.

[7] ARGUS. Argus Authorisation Service. `https://argus-documentation.readthedocs.io`. [Online; Accessed 27 December 2020].

[8] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia. A View of Cloud Computing. *Commun. ACM*, 53(4):50–58, April 2010.

[9] C. J. Atherton, T. Barton, J. Basney, D. Broeder, A. Costa, M. van Daalen, S. O. M. Dyke, W. Elbers, C.-F. Enell, E. M. V. Fasanelli, J. Fernandes, L. Florio, P. Gietz, D. L. Groep, M. Junker, C. Kanellopoulos, D. P. Kelsey, P. J. Kershaw, C. Knapic, T. Kollegger, S. Koranda, M. Linden, F. Marinic, L. Matyska, T. H. Nyrönen, S. Paetow, L. Paglione, S. Parlati, M. Prochazka, N. Rees, H. Short, U. Stevanovic, M. Tartakovsky, G. Venekamp, R. Wartel, C. Whalen, J. White, and C. Zwölf. Federated Identity Management for Research Collaborations. 2.0

Final Draft, CERN-OPEN-2012-006, April 2012.

[10] M. Atkinson, R. Baxter, P. Brezany, O. Corcho, M. Galea, M. Parsons, D. Snelling, and J. van Hemert. *Data-Driven Research in the Humanities - the DARIAH Research Infrastructure*, pages 417–429. IEEE, 2013.

[11] AT&T. AT&T XACML GitHub repository. `https://github.com/att/XACML`. [Online; Accessed 6 March 2020].

[12] Authentication and Authorisation for Research and Collaboration (AARC). AARC project web page. `https://aarc-project.eu`. [Online; Accessed 5 March 2020].

[13] R. A. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1999.

[14] D. Bailo, K. G. Jeffery, A. Spinuso, and G. Fiameni. Interoperability Oriented Architecture: The Approach of EPOS for Solid Earth e-Infrastructures. In *2015 IEEE 11th International Conference on e-Science*, pages 529–534. IEEE, 2015.

[15] S. Banerjee, S. Basu, S. Garg, S. Garg, S.-J. Lee, P. Mullan, and P. Sharma. Scalable Grid Service Discovery Based on UDDI. In *Proceedings of the 3rd International Workshop on Middleware for Grid Computing*, MGC '05, pages 1–6, New York, NY, USA, 2005. ACM.

[16] T. Barton, J. Basney, D. Groep, N. Harris, L. Johansson, D. Kelsey, S. Koranda, R. Wartel, A. West, and H. Short. A Security Incident Response Trust Framework for Federated Identity (Sirtfi). Protocol Specification, Sirtfi-1.0, REFEDS, December 2015.

[17] J. Basney, T. Fleury, and J. Gaynor. CILogon: A federated X.509 certification authority for cyberinfrastructure logon. *Concurrency and Computation: Practice and Experience*, 26(13):2225–2239, 2014. CPE-13-0334.R1.

[18] BDII – Grid Information System. `http://gridinfo.web.cern.ch`. [Online; Accessed 5 March 2020].

[19] K. Benedyczak, P. Bała, S. van den Berghe, R. Menday, and B. Schuller. Key aspects of the UNICORE 6 security model. *Future Generation Computer Systems*, 27(2):195 – 201, 2011.

[20] K. Benedyczak, B. Schuller, M. P.-E. Sayed, J. Rybicki, and R. Grunzke. UNICORE 7 — Middleware services for distributed and federated computing. In *2016 International Conference on High Performance Computing Simulation (HPCS)*, pages 613–620, July 2016.

[21] E. Bertino and K. Takahashi. *Identity Management: Concepts, Technologies, and Systems*. Artech House, Inc., Norwood, MA, USA, 2010.

[22] A. Biancini, L. Florio, M. Haase, M. Hardt, M. Jankowski, J. Jensen, C. Kanellopoulos, N. Liampotis, S. Licehammer, S. Memon, N. van Dijk, S. Paetow, M. Prochazka, M. Sallé, P. Solagna, U. Stevanovic, and D. Vaghetti. AARC: First draft of the Blueprint Architecture for Authentication and Authorisation Infrastructures. *CoRR*, abs/1611.07832, 2016.

[23] J. Blumtritt, W. Elbers, T. Goosen, M. Hinrichs, W. Qiu, M. Sallé, and M. Windhouwer. User Delegation in the CLARIN Infrastructure. In *Selected Papers from the CLARIN 2014 Conference, October 24-25, 2014, Soesterberg, The Netherlands*, Linköping, Sweden, August 2015. Linköping University Electronic Press.

[24] S. Burke, M. A. Pradillo, L. Field, and O. Keeble. GLUE 2 deployment: Ensuring quality in the EGI/WLCG information system. *Journal of Physics: Conference Series*, 513(3):032012, June 2014.

[25] A. Cambon-Thomsen, C. Bréchot, G. Dagher, G.-J. van Ommen, H.-E. Wichmann, J.-E. Litton, K. Zatloukal, L. Peltonen, M. Pasterk, M. Yuille, M. Taussig, and U. Landegren. Biobanking for Europe. *Briefings in Bioinformatics*, 9(1):14–24, October 2007.

[26] S. Cantor, J. Moreh, R. Philpott, E. Maler, C. P. Cahill, J. Hughes, H. Lockhart, M. Beach, R. Metz, R. Randall, T. Wisniewski, I. Reid, P. Austel, M. Hondo, M. McIntosh, T. Nadalin, N. Ragouzis, S. Cantor, R. B. Morgan, P. C. Davis, J. Hodges, F. Hirsch, J. Kemp, P. Madsen, S. Anderson, P. Mishra, J. Linn, R. Philpott, J. Moreh, A. Anderson, E. Maler, R. Monzillo, and G. Whitehead. Metadata for the OASIS Security Assertion Markup Language (SAML) V2.0. OASIS Standard, saml-metadata-2.0-os, OASIS, March 2005.

[27] S. Carmody, M. Erdos, K. Hazelton, W. Hoehn, R. B. M. Morgan, T. Scavo, D. Wasley, and S. Cantor. Shibboleth architecture protocols and profiles. Protocol Specification, internet2-mace-shibboleth-arch-protocols-200509, internet2, September 2005.

[28] D. W. Chadwick and G. Inman. Attribute Aggregation in Federated Identity Management. *Computer*, 42(5):33–40, May 2009.

[29] H. Chu. OpenLDAP: Highlights for 2.4. In *21st Large Installation System Administration Conference (LISA 07)*, Dallas, TX, November 2007. USENIX Association.

[30] CLARIN ERIC. `https://www.clarin.eu`. [Online; Accessed 13 July 2017].

[31] CLARIN ERIC. CLARIN services. `https://www.clarin.eu/content/services`. [Online; Accessed 5 February 2019].

[32] connect2id. Nimbus OAuth 2.0 SDK with OpenID Connect extensions. `https://connect2id.com/products/nimbus-oauth-openid-connect-sdk`. [Online; Accessed 4 January 2021].

[33] CONTRAIL: Open Computing Infrastructures for Elastic Services. `https://cordis.europa.eu/project/id/257438`. [Online; Accessed 9 April 2021].

[34] M. Conway, R. Moore, A. Rajasekar, and J. Y. Nief. Demonstration of Policy-Guided Data Preservation Using iRODS. In *2011 IEEE International Symposium on Policies for Distributed Systems and Networks (POLICY)*, pages 173–174, June 2011.

[35] Corbel: Coordinated Research Infrastructures Building Enduring Life-science Services. `http://www.corbel-project.eu`. [Online; Accessed 5 March 2020].

[36] CRPC. Argonne Workshop Explores Construction of a National Computational Grid. `https://www.crpc.rice.edu/newsletters/fal97/news_grid.html`, August 1998. [Online; Accessed 4 March 2020].

[37] F. Curbera, M. Duftler, R. Khalaf, W. Nagy, N. Mukhi, and S. Weerawarana. Unraveling the Web services web: an introduction to SOAP, WSDL, and UDDI. *IEEE Internet Computing*, 6(2):86–93, March 2002.

[38] DARIAH AAI Documentation. `https://wiki.de.dariah.eu/display/publicde/DARIAH+AAI+Documentation#DARIAHAAIDocumentation-AttributesavailableintheDARIAHAAI-Fulllist.` [Online; Accessed 15 March 2021].

[39] T. A. Defanti, I. Foster, M. E. Papka, R. Stevens, and T. Kuhfuss. Overview of the I-Way: Wide-Area Visual Supercomputing. *Int. J. High Perform. Comput. Appl.*, 10(2-3):123–131, June 1996.

[40] D. Drollette. Standards are the GLUE 2.0. iSGTW Feature, ScienceNode, March 2009. `https://sciencenode.org/feature/isgtw-feature-standards-are-glue-20.php`, [Online; Accessed 7 December 2018].

[41] Eclipse Foundation. Eclipse Jersey. `https://eclipse-ee4j.github.io/jersey/`. [Online; Accessed 12 January 2021].

[42] EGI. EGI – EGI Advanced Computing Services for Research. `http://www.egi.eu`. [Online; Accessed 11 March 2020].

[43] EGI. EGI Check-in. `https://www.egi.eu/services/check-in/`. [Online; Accessed 6 March 2021].

[44] ELIXIR. ELIXIR – A distributed infrastructure for life science information. `https://www.elixir-europe.org`. [Online; Accessed 15 February 2020].

[45] ELIXIR. ELIXIR AAI Documentation. `https://www.elixir-europe.org/services/compute/aai`. [Online; Accessed 13 February 2020].

[46] M. Ellert, A. Konstantinov, B. Kónya, O. Smirnova, and A. Wäänänen. The NorduGrid project: using Globus toolkit for building GRID infrastructure. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 502(2):407 – 410, 2003. Proceedings of the VIII International Workshop on Advanced Computing and Analysis Techniques in Physics Research.

[47] EMI. Common Authentication Library (caNl). Reference Implementation, `https://github.com/eu-emi/canl-java`. [Online; Accessed 4 January 2021].

[48] EMI. European Middleware Initiative (EMI). `https://cordis.europa.eu/project/id/261611`. [Online; Accessed 14 April 2021].

[49] EOSC-hub. EOSC Hub. `https://www.eosc-hub.eu`. [Online; Accessed 21 February 2020].

[50] H.-E. Eriksson and M. Penker. *Business Modeling With UML: Business Patterns at Work*. John Wiley & Sons, Inc., New York, NY, USA, 1st edition, 1998.

[51] EUDAT. B2ACCESS – EUDAT. `https://eudat.eu/services/b2access`. [Online; Accessed 11 March 2020].

[52] EUDAT. B2STAGE – EUDAT. `http://www.eudat.eu/b2stage`. [Online; Accessed 11 March 2020].

[53] EUDAT. EUDAT – Research Data Services, Expertise & Technology Solutions. `http://www.eudat.eu`. [Online; Accessed 12 February 2020].

[54] EUGridPMA. Research and Collaboration Authentication Certification Authority Service. `https://www.rcauth.eu`. [Online; Accessed 9 March 2020].

[55] European Commission. Data protection – rules for the protection of personal data inside and outside the eu. `https://ec.europa.eu/info/law/law-topic/data-protection_en`. [Online; Accessed 8 January 2021].

[56] European Commission. eInfrastructures - Shaping Europe's Digital Future. `https://ec.europa.eu/digital-single-market/en/e-infrastructures`. [Online; Accessed 1 March 2021].

[57] European Commission. www.esfri.eu. `https://www.esfri.eu`. [Online; Accessed 1 March 2021].

[58] European Commission. Legal framework for a European Research Infrastructure Consortium – ERIC. Report, European Commission, European Commission, Directorate-General for Research Communication Unit, B-1049 Brussels, April 2010. DOI:10.2777/79873.

[59] European Commission. European Cloud Initiative - Building a competitive data and knowledge economy in Europe. Report, European Commission, European Commission, Directorate-General for Research Communication Unit, B-1049 Brussels, April 2016. `http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=15266`.

[60] FI-WARE. WSO2 On-Premise and in the Cloud. `https://authzforce-ce-fiware.readthedocs.io/en/latest/`. [Online; Accessed 27 December 2020].

[61] L. Field and R. Sakellariou. An Evaluation of Information Consistency in Grid Information Systems. *Journal of Grid Computing*, 15(1):127–137, Mar 2017.

[62] S. Fitzgerald, I. Foster, C. Kesselman, G. von Laszewski, W. Smith, and S. Tuecke. A directory service for configuring high-performance distributed computations. In *Proceedings. The Sixth IEEE International Symposium on High Performance Distributed Computing (Cat. No.97TB100183)*, pages 365–375, Aug 1997.

[63] R. Flowers and C. Edeki. Business Process Modelling Notation. *International Journal of Computer Science and Mobile Computing*, Vol 2:35–40, March 2013.

[64] A. Forestiero, C. Mastroianni, and G. Spezzano. Building a Peer-to-peer Information System in Grids via Self-organizing Agents. *Journal of Grid Computing*, 6(2):125–140, June 2008.

[65] I. Foster. Globus Toolkit Version 4: Software for Service-oriented Systems. In *Proceedings of the 2005 IFIP International Conference on Network and Parallel Computing*, NPC'05, pages 2–13, Berlin, Heidelberg, 2005. Springer-Verlag.

[66] I. Foster. Globus Online: Accelerating and Democratizing Science Through Cloud-Based Services. *IEEE Internet Computing*, 15(3):70–73, May 2011.

[67] I. Foster and C. Kesselman. Globus: A Metacomputing Infrastructure Toolkit. *Int. J. High Perform. Comput. Appl.*, 11(2):115–128, June 1997.

[68] I. Foster, C. Kesselman, and S. Tuecke. The anatomy of the grid: Enabling scalable virtual organizations. *The International Journal of High Performance Computing Applications*, 15(3):200–222, 2001.

[69] L. Fuchs and G. Pernul. inimizing insider misuse through secure Identity Management. *Security and Communication Networks*, 5(8):847–862, 2012.

[70] F. Gagliardi, B. Jones, M. Reale, and S. Burke. European DataGrid Project:

Experiences of Deploying a Large Scale Testbed for E-science Applications. In M. C. Calzarossa and S. Tucci, editors, *Performance Evaluation of Complex Systems: Techniques and Tools*, pages 480–499, Berlin, Heidelberg, 2002. Springer Berlin Heidelberg.

[71] P. A. Grassi, N. B. Lefkowitz, E. M. Nadeau, R. J. Galluzzo, and A. T. Dinh. Attribute Metadata: A Proposed Schema for Evaluating Federated Attributes. Internal Report NISTIR-8112.html, NIST, February 2018.

[72] P. A. Grassi, J. P. Richer, S. K. Squire, J. L. Fenton, E. M. Nadeau, N. B. Lefkovitz, J. M. Danker, Y.-Y. Choong, K. K. Greene, and M. F. Theofanos. Digital Identity Guidelines: Federation and Assertions. Special publication, nist.sp.800-63c, NIST, June 2017.

[73] D. Groep. IGTF Levels of Authentication Assurance. Specification, igtf-authn-assurance-1.1 urn:oid:1.2.840.113612.5.2.6.1, IGTF, January 2016.

[74] GÉANT. eduGAIN – enabling worldwide access. `https://www.edugain.org`. [Online; Accessed 10 March 2020].

[75] GÉANT. Terena Certificate Service (TCS). `https://www.terena.org/activities/tcs`. [Online; Accessed 6 January 2020].

[76] L. Hämmerle and S. Andrėj. Federation Architectures – eduGAIN – GÉANT federated confluence. `https://wiki.geant.org/display/eduGAIN/Federation+Architectures`. [Online; Accessed 9 January 2020].

[77] D. Hardt. The Auth 2.0 Authorization Framework. Internet Requests for Comments, Proposed Standard RFC 6749, IETF, October 2012.

[78] HERAS-AF. HERAS-AF XACML Core. `https://bitbucket.org/herasaf/herasaf-xacml-core/src/master/`. [Online; Accessed 12 January 2021].

[79] G. Hohpe and B. Woolf. *Enterprise Integration Patterns: Designing, Building, and Deploying Messaging Solutions*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2003.

[80] J. Howlett, S. Hartman, and A. Pérez-Méndez. A RADIUS Attribute, Binding, Profiles, Name Identifier Format, and Confirmation Methods for the Security Assertion Markup Language (SAML). Internet Requests for Comment RFC 7833, IETF, May 2016.

[81] J. Howlett, S. Hartman, H. Tschofenig, and J. Schaad. Application Bridging for Federated Access Beyond Web (ABFAB) Architecture. Internet Requests for Comment RFC 7831, IETF, May 2016.

[82] V. C. Hu, D. Ferraiolo, R. Kuhn, A. Schnitzer, K. Sandlin, R. Miller, and K. Scarfone. Guide to Attribute Based Access Control (ABAC) Definition and Considerations. Special Publication NIST.SP.800-162, NIST, January 2014.

[83] J. Hughes, S. Cantor, J. Hodges, F. Hirsch, P. Mishra, R. Philpott, E. Maler, C. P. Cahill, H. Lockhart, M. Beach, R. Metz, R. Randall, T. Wisniewski, I. Reid, P. Austel, M. Hondo, M. McIntosh, T. Nadalin, N. Ragouzis, R. B. Morgan, P. C. Davis, J. Kemp, P. Madsen, S. Anderson, J. Linn, J. Moreh, A. Anderson, R. Monzillo, and G. Whitehead. Bindings for the OASIS Security Assertion Markup Language (SAML) V2.0. OASIS Standard, saml-bindings-2.0-os, OASIS, March 2005.

[84] J. Hughes, S. Cantor, J. Hodges, F. Hirsch, P. Mishra, R. Philpott, E. Maler, C. P.

Cahill, H. Lockhart, M. Beach, R. Metz, R. Randall, T. Wisniewski, I. Reid, P. Austel, M. Hondo, M. McIntosh, T. Nadalin, N. Ragouzis, R. B. Morgan, P. C. Davis, J. Kemp, P. Madsen, S. Anderson, J. Linn, J. Moreh, A. Anderson, R. Monzillo, and G. Whitehead. Profiles for the OASIS Security Assertion Markup Language (SAML) V2.0. OASIS Standard, saml-profiles-2.0-os, OASIS, March 2005.

[85] ISO/IEC. Information technology – Security techniques – A framework for identity management – Part 1: Terminology and concepts. Standard ISO/IEC-24760-1-2011, International Organization for Standardization, Geneva, CH, December 2011.

[86] J. Jensen. Federated Identity Management Challenges. In *2012 Seventh International Conference on Availability, Reliability and Security*, pages 230–235. IEEE, Aug 2012.

[87] J. Jensen, U. Stevanovic, I. Kakavas, N. Liampotis, M. Haase, P. Gietz, M. Jankowski, M. Reale, M.-L. Mantovani, and L. Florio. Design for Deploying Solutions for "Guest Identities". Project Milestone Document MJRA1.2, AARC, June 2016.

[88] Jisc. Jisc. `https://www.jisc.ac.uk`. [Online; Accessed 6 January 2021].

[89] Jisc. Moonshot Project. `https://wiki.moonshot.ja.net`. [Online; Accessed 6 January 2021].

[90] M. Jones, J. Bradley, and N. Sakimura. JSON Web Token (JWT). Internet Requests for Comments RFC 7519, IETF, May 2015.

[91] M. Jones, A. Nadalin, B. Campbell, J. Bradley, and C. Mortimore. OAuth 2.0 Token Exchange. Internet Requests for Comments draft-ietf-oauth-token-exchange-19, IETF, July 2019.

[92] D. S. Katz, S. Callaghan, R. Harkness, S. Pamidighantam, M. Pierce, B. Plale, C. Song, and J. Towns. Science on the TeraGrid. *Computational Methods in Science and Technology*, Special Issue 2010:81–97, January 2010.

[93] C. Loomis. Final Evaluation of Testbed Operation. Project Deliverable D6.8-3.0, DataGrid, December 2003.

[94] E. Maler and D. Reed. The Venn of Identity: Options and Issues in Federated Identity Management. *IEEE Security Privacy*, 6(2):16–23, March 2008.

[95] J. Martin, S. James Martin, and J. Martin. *Managing the Data-base Environment*. (A James Martin book). Prentice-Hall, 1983.

[96] G. Mathieu, D. A. Richards, D. J. Gordon, C. D. C. Novales, P. Colclough, and M. Viljoen. GOCDB, a topology repository for a worldwide grid infrastructure. *Journal of Physics: Conference Series*, 219(6):062021, 2010.

[97] P. Mell and T. Grance. The NIST Definition of Cloud Computing. Special Publication NIST.SP.800-145, NIST, September 2011.

[98] A. S. Memon, M. Riedel, L. Field, G. Szigeti, and I. Marton. EMIR: An EMI Service Registry for Federated Grid Infrastructures. In *EGI Community Forum 2012 / EMI Second Technical Conference, Munich (Germany), 26 Mar 2012 – 30 Mar 2012*, Proceedings of Science, Trieste, 2012. Sissa.

[99] A. Moitra, B. Barnett, A. Crapo, and S. J. Dil. Using Data Provenance to Measure Information Assurance Attributes. In D. L. McGuinness, J. R. Michaelis, and L. Moreau, editors, *Provenance and Annotation of Data and Processes*, pages

111–119, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.

[100] MongoDB for GIANT Ideas. `https://www.mongodb.com`. [Online; Accessed 5 March 2020].

[101] MVEL. `http://mvel.documentnode.com`. [Online; Accessed 27 February 2020].

[102] A. Nadalin and M. Goodner. Web Services Federation Language (WS-Federation) Version 1.2. `URL:http://docs.oasis-open.org/wsfed/federation/v1.2/os/ws-federation-1.2-spec-os.pdf`, 5 2009.

[103] S. Nakamoto. Bitcoin: A Peer-to-Peer Electronic Cash System. *Cryptography Mailing list at https://metzdowd.com*, page 9, 03 2009.

[104] A. Nenadic, N. Zhang, L. Yao, and T. Morrow. Levels of authentication assurance: an investigation. In *Third International Symposium on Information Assurance and Security*, pages 155–160, 2007.

[105] C. Ngo, Y. Demchenko, and C. de Laat. Toward a Dynamic Trust Establishment approach for multi-provider Intercloud environment. In *4th IEEE International Conference on Cloud Computing Technology and Science Proceedings*, pages 532–538, December 2012.

[106] NorduGrid. ARC peer-to-peer information system. Documentation and developer's guide NORDUGRID-TECH-21, NorduGrid, February 2013.

[107] J. Novotny, S. Tuecke, and V. Welch. An online credential repository for the Grid: MyProxy. In *Proceedings 10th IEEE International Symposium on High Performance Distributed Computing*, pages 104–111, 2001.

[108] A. Paolini, S. Burke, E. Fernandez, P. Andreetto, M. Verlato, J. P. Navarro, and S. Memon. GLUE Specification v. 2.1. Open Grid Forum, Protocol Specification, GFD-R-P.239, to appear 2021.

[109] B. Parducci, H. Lockhart, and E. Rissanen. eXtensible Access Control Markup Language (XACML) Version 2.0. Protocol Specification xacml-2.0-core-spec-os, OASIS, February 2005.

[110] B. Parducci, H. Lockhart, and E. Rissanen. eXtensible Access Control Markup Language (XACML) Version 3.0. Protocol Specification xacml-3.0-core-spec-en, OASIS, January 2013.

[111] B. Parducci, H. Lockhart, and E. Rissanen. XACML SAML Profile Version 2.0. Protocol Specification xacml-saml-profile-v2.0, OASIS, August 2014.

[112] B. Parducci, H. Lockhart, and E. Rissanen. XACML v3.0 Administration and Delegation Profile Version 1.0. Protocol Specification xacml-json-http-v1.0, OASIS, November 2014.

[113] B. Parducci, H. Lockhart, and E. Rissanen. XACML v3.0 Multiple Decision Profile Version 1.0. Protocol Specification xacml-3.0-multiple-v1-spec-en, OASIS, May 2014.

[114] B. Parducci, H. Lockhart, and E. Rissanen. XACML v3.0 Privacy Policy Profile Version 1.0. Protocol Specification xacml-3.0-privacy-v1-spec-en, OASIS, January 2015.

[115] B. Parducci, H. Lockhart, E. Rissanen, and R. Levinson. XACML v3.0 Hierarchical Resource Profile Version 1.0. Protocol Specification xacml-3.0-hierarchical-v1-spec-en, OASIS, June 2014.

[116] B. Parducci, H. Lockhart, and R. Sinnema. REST Profile of XACML v3.0

Version 1.0. Protocol Specification xacml-rest-v1.0, OASIS, November 2014.

[117] B. Parducci and H. L. E. Rissanen. Assertions and Protocols for the OASIS Security Assertion Markup Language (SAML) V2.0. OASIS Standard, saml-core-2.0-os, OASIS, March 2005.

[118] B. Parducci and H. L. E. Rissanen. XACML v3.0 XML Digital Signature Profile Version 1.0. Protocol Specification xacml-3.0-dsig-v1.0, OASIS, May 2014.

[119] A. C. Partners and A. members. Guidelines for the evaluation and combination of the assurance of external identities (AARC-G031). Project report, AARC and AppInt, May 2018.

[120] PRACE. PRACE Research Infrastructure. http://www.prace-ri.eu. [Online; Accessed 5 March 2020].

[121] PrimeKey Solutions AB. EJBCA - The Open Source CA. https://www.ejbca.org/. [Online; Accessed 6 March 2021].

[122] J. Quinteros and A. Heinloo. EAS user documentation. Report, Release 0.9b1, EPOS, May 2019.

[123] Research and Collaboration Authentication Certification Authority Service. https://www.rcauth.eu. [Online; Accessed 16 February 2020].

[124] REFEDS. REFEDS Assurance Framework – RAF. https://refeds.org/assurance. [Online; Accessed 27 February 2020].

[125] L. Robertson. *Computing Services for LHC: From Clusters to Grids*, pages 69–89. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[126] A. I. T. Rowstron and P. Druschel. Pastry: Scalable, Decentralized Object Location, and Routing for Large-Scale Peer-to-Peer Systems. In *Proceedings of the IFIP/ACM International Conference on Distributed Systems Platforms*, Middleware '01, pages 329–350, London, UK, 2001. Springer-Verlag.

[127] N. Sakimura, J. Bradley, M. Jones, B. de Medeiros, and C. Mortimore. OpenID Connect Core 1.0. Protocol Specification OpenID Connect Core 1.0, OpenID Foundation, November 2014.

[128] D. Scardaci, L. Florio, D. Huebner, M. Jankowski, J. Jensen, C. Kanellopoulos, D. Kouril, N. Liampotis, M. Linden, S. Memon, M. Salle, and A. Terpstra. Deliverable DJRA1.1: Use-Cases for Interoperable CrossInfrastructure AAI. AARC Project Deliverable, Ref. Ares(2018)4846700 - 21/09/2018 https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5bdced7a1&appId=PPGMS, Sept. 2018.

[129] B. Schuller and T. Pohlmann. UFTP: High-Performance Data Transfer for UNICORE. In *UNICORE Summit 2011, Proceedings, 7-8 July 2011, Torun, Poland, ed.:/ M. Romberg, P. Bala, R. Müller-Pfefferkorn, D.Mallmann, Forschungszentrum Jülich, 2011, IAS Series Vol. 9. - 978-3-89336-750-4. - S. 135 - 142*, 2011. Record converted from VDB: 12.11.2012.

[130] A. Sergio, S. Burke, F. Ehm, L. Field, G. Galang, B. Konya, M. Litmaath, P. Millar, and J. P. Navarro. GLUE Specification v. 2.0. Open Grid Forum, Prtocol Specification, GFD-R-P.147, Mar 2009.

[131] A. Sergio, S. Burke, L. Field, B. Konya, A. S. Memon, D. Meredith, J. P. Navarro, F. Paganelli, and W. Smith. GLUE v, 2.0 – Reference realisation to XML Schema. Open Grid Forum, GFD-R-P.209, Oct 2013.

[132] R. W. Shirey. Internet Security Glossary, Version 2. IETF Internet Requests for Comments, Informational, RFC4949, August 2007.

[133] K. M. Sim. Agent-based approaches for intelligent intercloud resource allocation. *IEEE Transactions on Cloud Computing*, 7(2):442–455, April 2019.

[134] K. Sivashanmugam, K. Verma, and A. Sheth. Discovery of Web services in a federated registry environment. In *Proceedings. IEEE International Conference on Web Services, 2004.*, pages 270–278, July 2004.

[135] L. Smarr and C. E. Catlett. Metacomputing. *Commun. ACM*, 35(6):44–52, June 1992.

[136] D. Smith. The challenge of federated identity management. *Network Security*, 2008(4):7 – 9, 2008.

[137] W. Smith, D. Meredith, A. S. Memon, and J. P. Navarro. GLUE v, 2.0 – Reference realisation to JSON Schema. Protocol Specification, GFD-R-P.219, Open Grid Forum, December 2015.

[138] W. Smith, S. Pamidighantam, and J.-P. Navarro. Publishing and Consuming GLUE V2.0 Resource Information in XSEDE. In *Proceedings of the 2015 XSEDE Conference: Scientific Advancements Enabled by Enhanced Cyberinfrastructure*, XSEDE '15, pages 25:1–25:8, New York, NY, USA, 2015. ACM.

[139] SimpleSAMLphp. `https://simplesamlphp.org`. [Online; Accessed 16 June 2017].

[140] EMI Security Token Service (STS). `https://twiki.cern.ch/twiki/bin/view/EMI/EMISTSDocumentation`. [Online; Accessed 6 March 2020].

[141] SWITCH. Authentication and Authorization Infrastructure (AAI) Preparatory Study. Report, `https://www.switch.ch/aai/docs/AAI_Study_v10.pdf`, June 2002. [Online; Accessed 15 April 2021].

[142] The NorduGrid Collaboration. NorduGrid. `http://www.nordugrid.org`. [Online; Accessed 5 January 2020].

[143] S. Tuecke, R. Ananthakrishnan, K. Chard, M. Lidman, B. McCollam, S. Rosen, and I. Foster. Globus auth: A research identity and access management platform. In *2016 IEEE 12th International Conference on e-Science (e-Science)*, pages 203–212, October 2016.

[144] Unity. Unity IdM - Identity management and authentication (IAM) platform. `https://www.unity-idm.eu`. [Online; Accessed 10 March 2020].

[145] Vaadin. An open platform for building web apps in Java. `https://www.vaadin.com`. [Online; Accessed 10 April 2021].

[146] A. van Rijn and R. Vandenbroucke. Guide to e-Infrastructure requirements for European Research Infrastructures. Specification eirg-1004, E-IRG, March 2017.

[147] M. Villari, F. Tusa, A. Celesti, and A. Puliafito. How to Federate VISION Clouds through SAML/Shibboleth Authentication. In F. De Paoli, E. Pimentel, and G. Zavattaro, editors, *Service-Oriented and Cloud Computing*, pages 259–274, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[148] W. Vogels. Eventually Consistent. *Commun. ACM*, 52(1):40–44, Jan. 2009.

[149] V. Welch, I. Foster, C. Kesselman, O. Mulmo, L. Pearlman, S. Tuecke, J. Gawor, S. Meder, and F. Siebenlist. X.509 Proxy Certificates for Dynamic Delegation. In *3rd annual PKI R&D workshop*, volume 14, 2004.

[150] K. Wierenga, S. Winter, and T. Wolniewicz. The eduroam Architecture for Network Roaming. Request for Comments 7593, IETF, Sept. 2015.

[151] A. Wright, H. Andrews, and B. Hutton. JSON Schema Validation: A Vocabulary for Structural Validation of JSON. Internet-Draft draft-bhutton-json-schema-validation-00, IETF, Dec. 2020. Work in Progress.

[152] WSO2. WSO2 On-Premise and in the Cloud. `https://wso2.com/identity-and-access-management/`. [Online; Accessed 27 December 2020].

[153] XSEDE. Home – XSEDE. `https://www.xsede.org`. [Online; Accessed 13 January 2020].

[154] F. Zhu, M. W. Mutka, and L. M. Ni. Service discovery in pervasive computing environments. *IEEE Pervasive Computing*, 4(4):81–90, October 2005.

[155] C. Zinn, M. Hinrichs, E. Dima, and D. van Uytvanck. CLARIN switchboard specification. Specification CE-2015-0684, CLARIN, 2015.