Authentication and Authorisation for Research and Collaboration

Federations 101

An Introduction to Federated Identity Management

Peter Gietz, Martin Haase

AARC NA2 Task 2 - Outreach and Dissemination

DAASI International

Federate to win! - An AARC Project Workshop at
LIBER 2016 Conference, Helsinki

29 June 2016

DAASI International

Where from did I steal?

- Some of the Slides were originally created by
  - Scott Cantor (Internet 2, lead developer of the Shibboleth Service Provider)
  - the SWITCHaai Team (https://www.switch.ch/aai/)
  - Colleagues from the AARC project, e.g. Mikael Linden on Data Protection Code of Conduct
- Since they were all made in the spirit of cooperation such slides are marked with 

Motivation for Federations
Do these questions sound familiar?

- “What is a federation?”
- “What is SAML?”
- “Institution XYZ just introduced Shibboleth. When will you follow?”
- “What is this Shibboleth all about?”
- “There is a new online research journal, what do we need to do to get access?”
- “I have too many passwords everywhere, why can’t I use my university account?”
- “Students prefer to read their online journals from home. But they must be physically in the library, using one of 3 kiosk PCs. Why is this so?”
- “Are there any other federations technologies I need to know about?”

Motivation for Federations in the library context

- In the past 20 years, publishing houses have all gone online and new online resources have emerged
- The first attempt at securing such resources was by restricting the IP addresses of the institutions that could access the content
  - This would mean that users needed to be physically present in the institution's network or use proxy/VPN solutions
- Whilst this worked in the past, it is not a viable approach in today's world where
  - users move much more frequently
  - remote access is a strong requirement
  - it is fairly simple to circumvent the IP address check
- Federated access and the underlying standard (SAML, Security Assertion Markup Language) has been the answer to these challenges

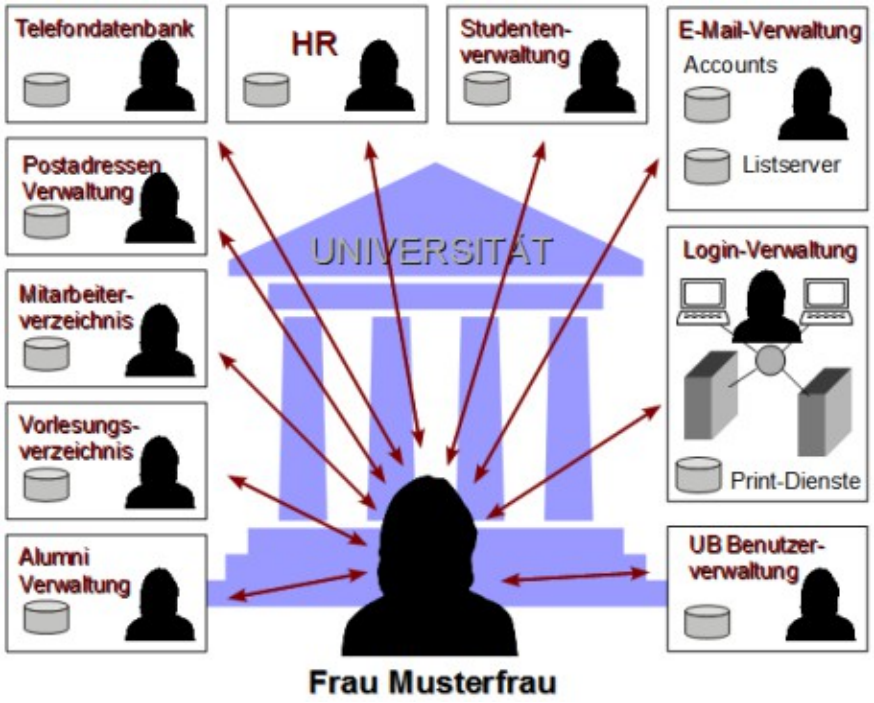
What is Identity Management?

- Identity Management is Software and process for managing identities in such a way that
  - Identity data are only managed at authoritative sources, i.e. HR systems for staff, Student management system for students, a guest database for guests etc.
  - Such data get automatically synchronized to the Identity Management System, whereas the same person coming from more than one authoritative source will be identified as such
  - Automated processes create login accounts with initialisation passwords
  - Automated provisioning processes provide applications with current identity data
  - A user only needs one password for all applications within her organisation (“unified login”)

What is Identity Management good for?

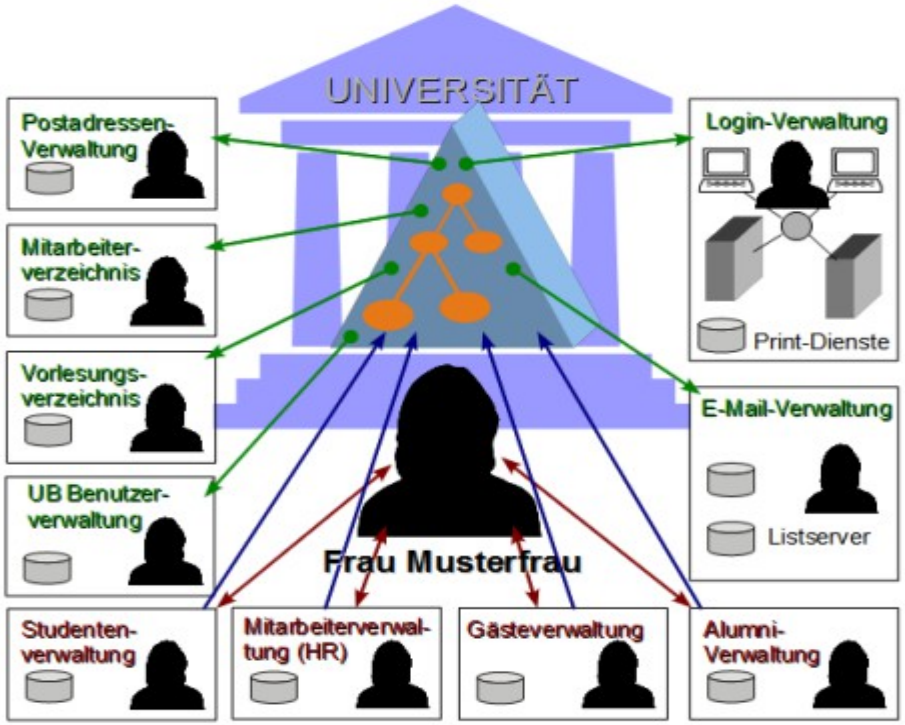
- IdM provide for
  - A structured view on data, data sources and target applications
  - Better data accuracy (the authoritative sources have the most accurate data)
  - More current data (applications always have the current data)
  - Users get their login account and online much faster
  - Users only have to remember one password
  - A stronger password policy can be enforced
  - If a user leaves the organisation her data also get deprovisioned much faster
  - Thus overall more security

What is Identity Management good for?



Without IdM

With IdM



What is Federated Identity Management?

- Federated IdM (FIM) allows for
  - Cooperation of organisations in sharing resources
  - Getting identity management across organisational borders
  - Based on the concept of federation
- Current technologies for FIM also allow for Single Sign-On
  - User authenticates once and is authenticated for all federated applications be they from the own organisation or from another organisation in the federation

What is a Federation?

- A Federation is a group of organisations that trust each other
  - Trust is established by contracts
- A Federation enables user of one organisation access services of another organisation without the need of a second account (e.g.,login name /password) at the service
- A Federation includes different actors in different roles
  - Service Provider – a provider of a service (e.g. a web application)
  - Identity Provider – a provider of identity data that are stored in an Identity or Account Management System
  - Federation Operator – managing the memberships within a federation
- A Federation is implemented by different pieces of software for the different actors / roles that communicate with each other via standard protocols such as SAML

What is an Identity Provider?

- Identity Provider is a **role** within a federation
- Identity Provider (IdP) is a **piece of software** that:
  - Is connected to a user management system, such as an LDAP Server or an MS Active Directory Server
  - Can check an authentication (user sends login name and password, the user management system can verify if the password is correct)
  - Can provide attributes about the user from the user management system
  - Can send messages (statements, assertions) to a Service Provider about authentication success and user attributes using a common protocol (SAML)
- Usually the role Identity Provider operates the IdP software

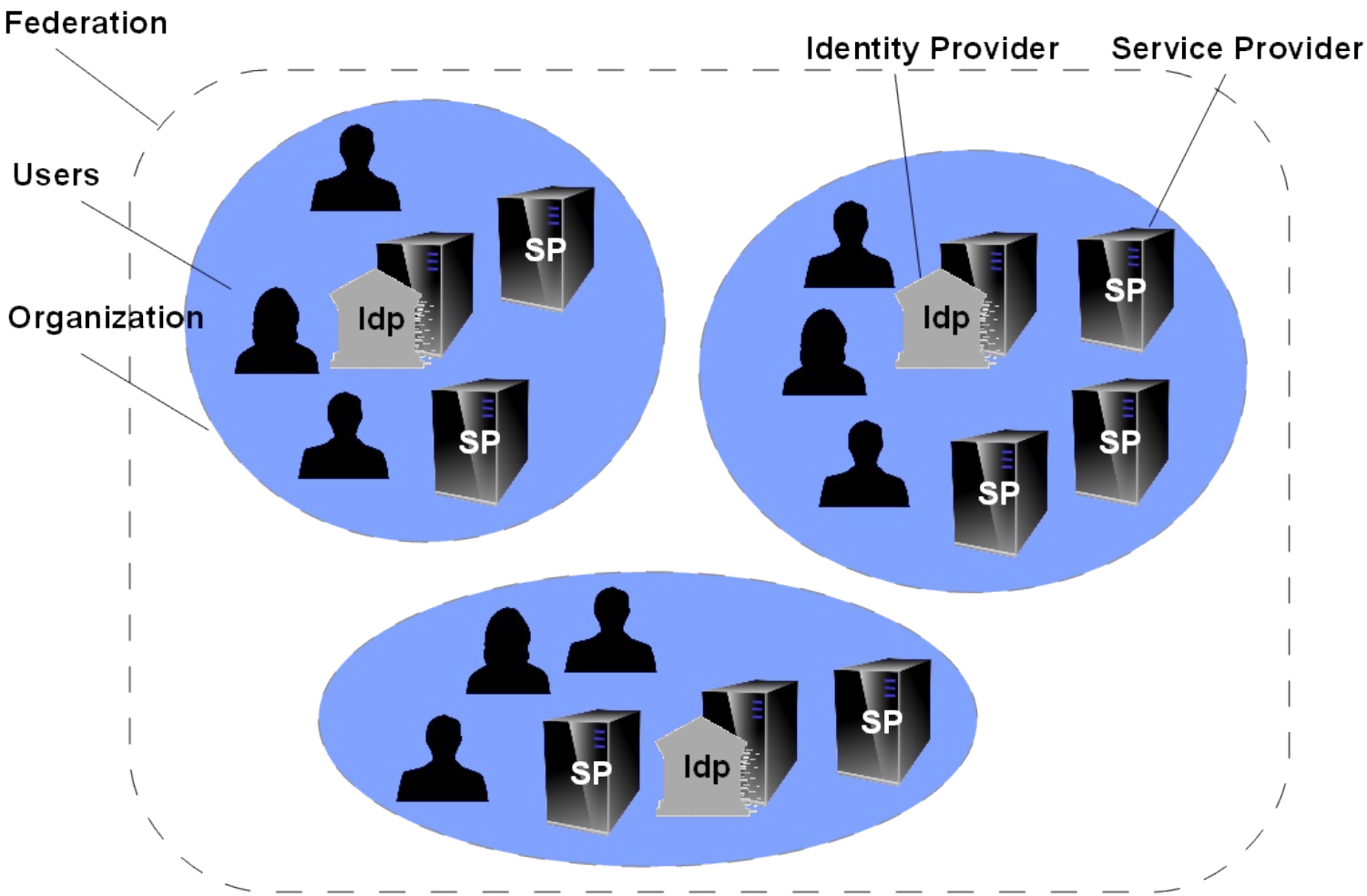
What is a Service Provider?

- Service Provider is a **role** within a federation
- Service Provider (SP) is a **piece of software** that:
  - Is connected to an application software such as a web based application that can be used with a browser
  - Can receive messages (statements, assertions) from an IdP about authentication success and user attributes using a common protocol (SAML)
  - Can transform such messages, so that the application can receive it
- The application then decides about the user's access to the resource (which can be just a protected web page, a menu item, etc.
- Usually the role Service Provider operates the SP software

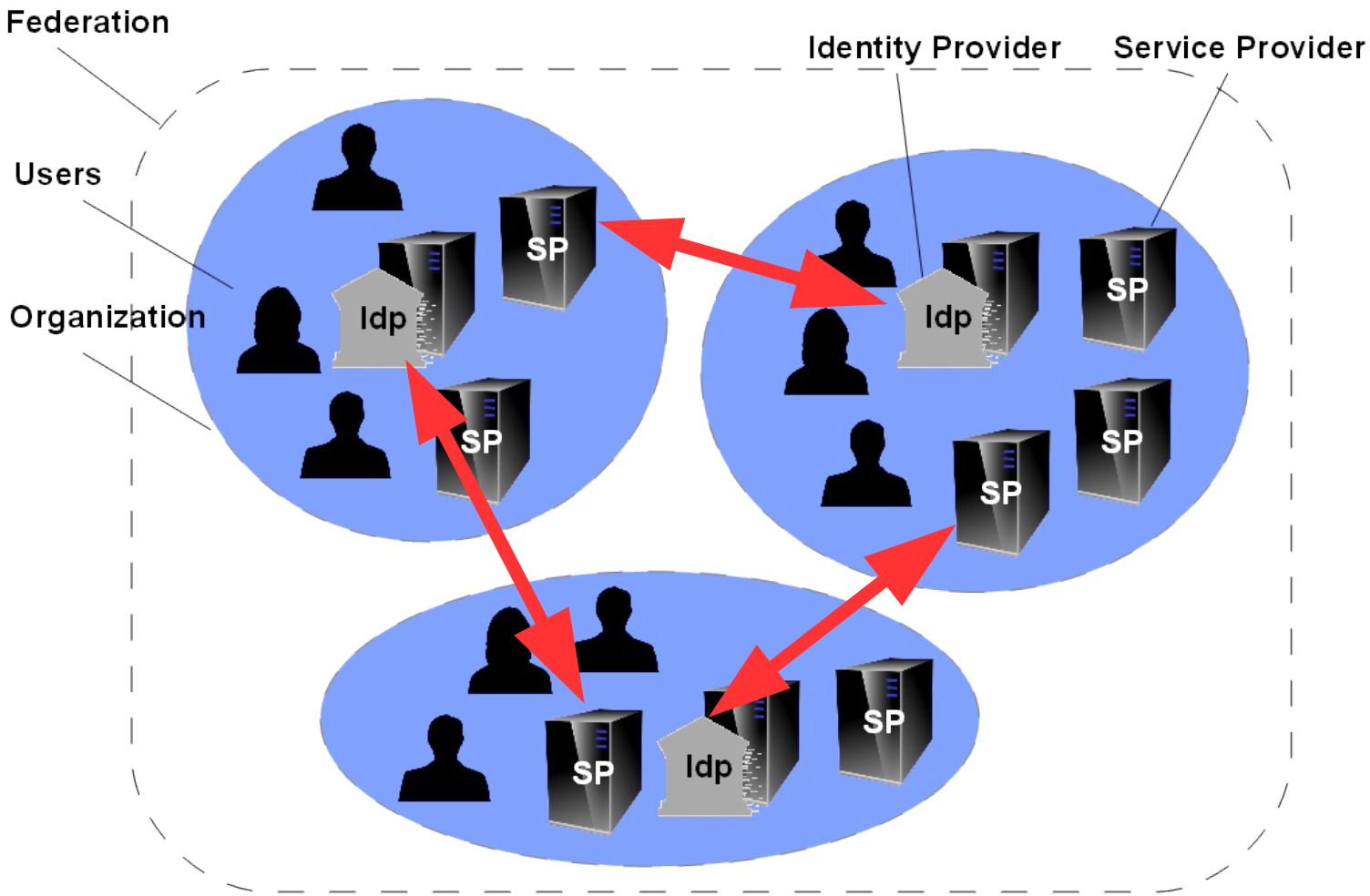
What is a federation provider?

- Federation Provider is a **role** within a federation and is in charge of operating a federation
- The Federation Provider fulfills the following tasks:
  - Defines the policy and contracts of the federation
  - Manage the memberships of the federation, maintaining a list of IdPs and SPs that belong to the federation
    - The list, also called **metadata**, does not contain the role inhabitants but the servers on which the SP or IdP software is running
  - Make this list available to SPs and IdPs, so they can check whether to trust the communication
  - Provide a service with which users can select their home organisation (i.e. their IdP)
- In the higher educational context national federations are often operated by the NRNs (National Research Networks, like SWITCH, or DFN) but also by other state owned organisations (like CSC)

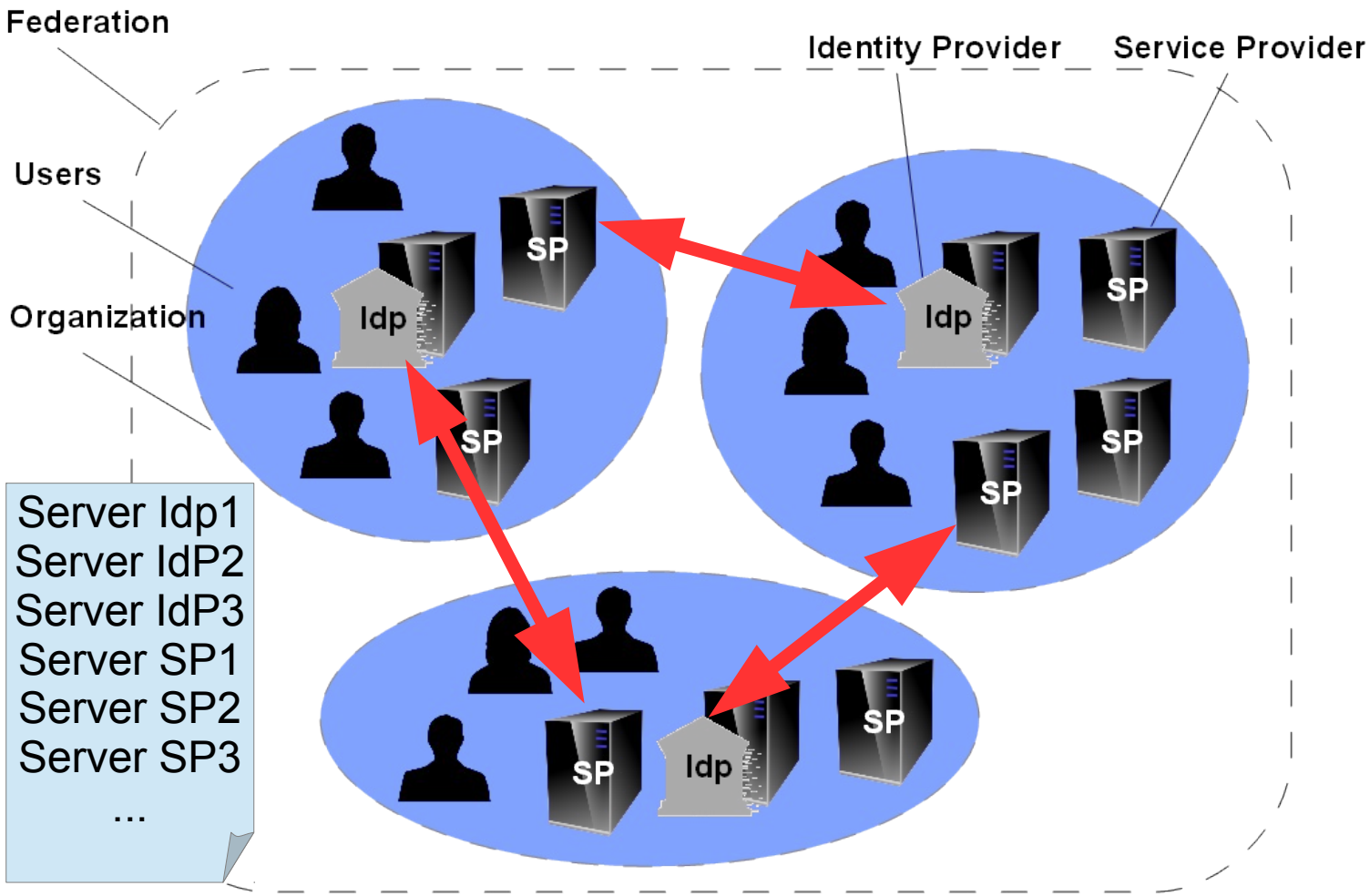
Federation in one Picture



Federation in one Picture



Federation in one Picture



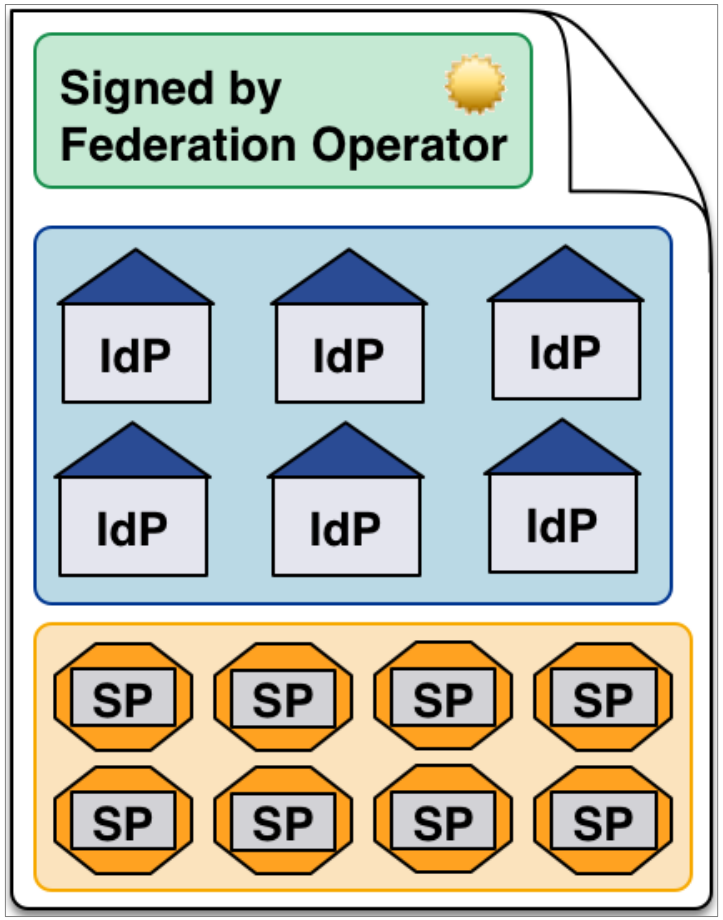
Federation Summary

- A group of organizations running IdPs and SPs that agree on a common set of rules and standards
  - It's a label - to talk about such a collection of organizations
  - An organization may belong to more than one federation at a time

- The grouping can be on a national level or regional level or on a smaller scale (e.g. large campus)

Federation Metadata (The membership list)

- An XML document that describes a federation
- Contains:
  - Unique identifier for each entity known as the entityID
  - Endpoints where each entity can be contacted
  - Certificates used for signing and encrypting data
- May contain
  - Organization and person contact information
  - Information about which attributes an SP wants/needs
- Metadata is usually distributed by a public HTTP URL
  - The metadata should be digitally signed
  - Signature should be verified!
  - Bilateral metadata exchange scales very badly
- Metadata must be kept up to date, so that
  - new entities can interoperate with existing ones
  - old or revoked entities are blocked



Benefits of Federated Identity Management

- Reduces work
  - Authentication-related calls to Penn State University's helpdesk dropped by 85% after they installed Shibboleth
- Provides current data
  - Studies of applications that maintain user data show that the majority of data is out of date
  - Are you "protecting" your app with stale data?
- Insulation from service compromises
  - Data gets pushed to services as needed
  - An attacker can't get everyone's data on a compromised server
- Minimize attack surface area
  - Only the IdP needs to be able to contact user data stores
  - All effort can be focused on securing this single connection instead of one (or more) connection per service.

Benefits of Federated Identity Management

- Users generally find the resulting single sign-on experience to be nicer than logging in numerous times.

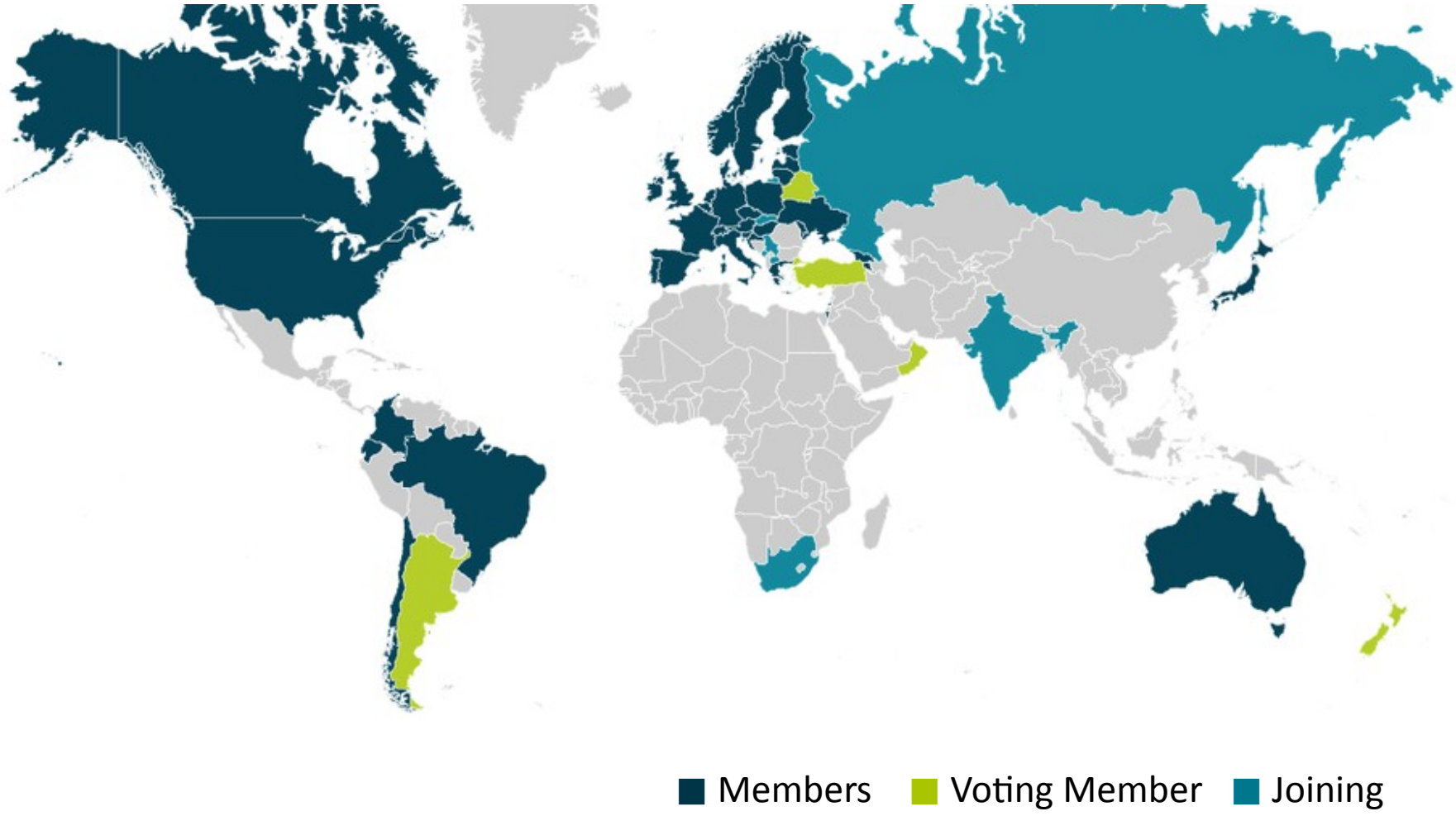
- Usability-focused individuals like that the **authentication process is consistent** regardless of the service accessed (with exception of IdP Discovery).

- A properly maintained federation drastically simplifies the process of integrating new services.

What is an interfederation?

- Interfedarion is the interoperation of several federations
- Interfederation takes place if a user from one federation accesses a service which is registered in another federation
- Interfederation is enabled by interfederation services
- For the higher educational context the interfederation services is called eduGAIN

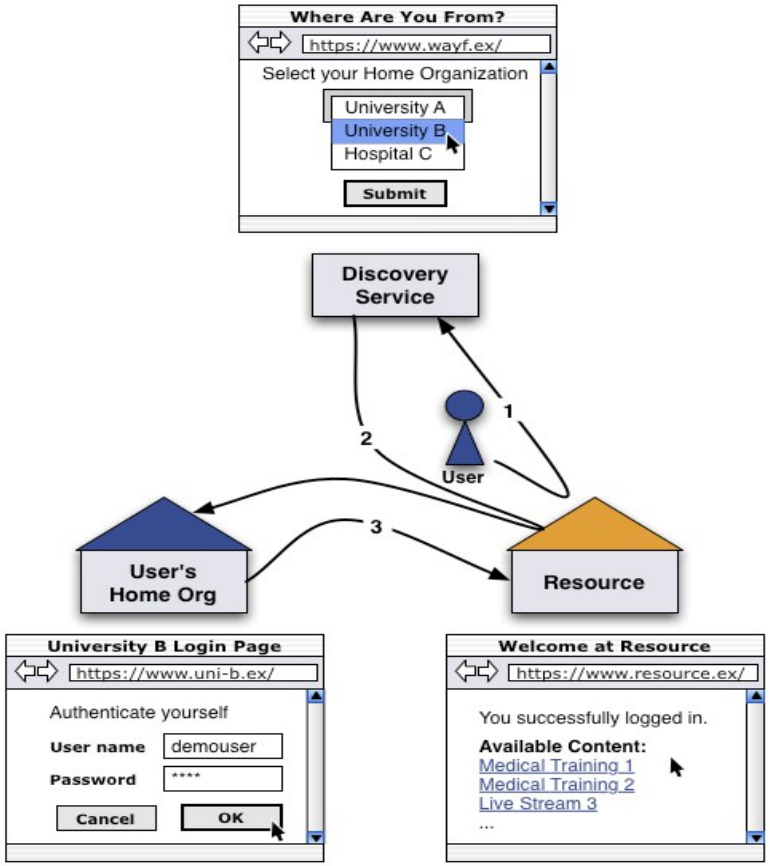
Interfederation eduGAIN



What is SAML?

- OASIS Standard, current version 2.0, March 2005
- Security Assertion Markup Language (SAML) specifies
  - Profiles (e.g. Web Browser SSO, Single Logout, Assertion Query, Attribute Usage)
  - Data exchange formats (esp. Assertions)
  - Protocols and Bindings
  - Metadata
- Components:
  - Identity Provider (IdP), lets users log in using the home organization's user directory
  - Service Provider (SP), protects Web resources and provides for information about the user sent by the IdP
  - IdP Discovery Service (DS, old term: Where-Are-You-From, WAYF)

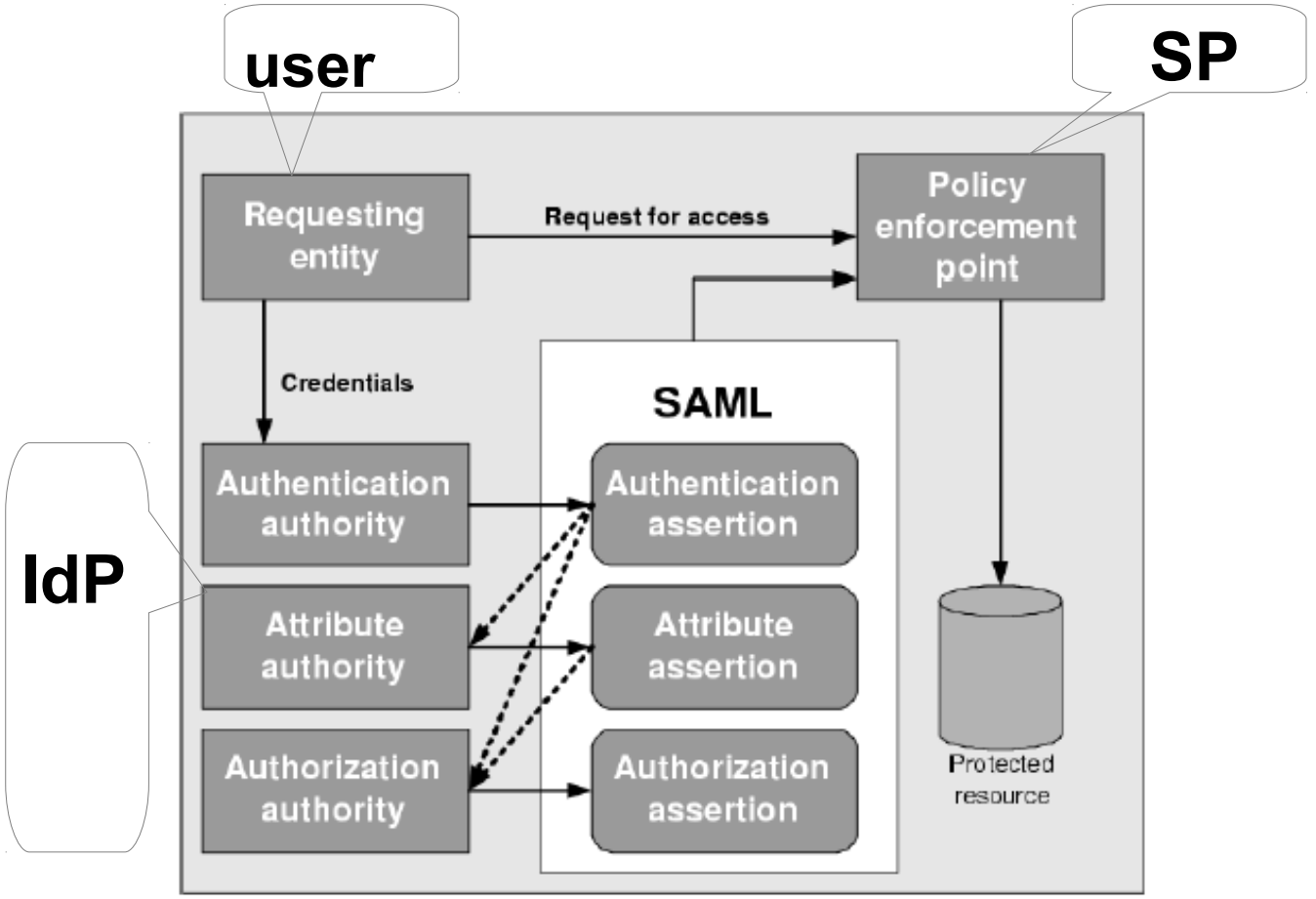
SAML Profile WebSSO



SAML assertions

- „SAML Ticket“ issued to one single SP
- Usually signed and encrypted
- Can contain max. 3 statements:
  - 1 Authentication Statement: AuthN instant + method
  - 1 Attribute Statement: 1..n Attributes, each 1..n values
  - 1 Authorization Decision Statement (very rarely implemented)
- Contains besides the three statements:
  - Issuer EntityID
  - Signature by the IdP
  - Information about the authenticated Subject: Name Identifier (Format + Value), EntityID of IdP and SP
  - Conditions (optional), e.g. AudienceRestriction for the SP

SAML assertions



Nach: RUBENKING, NEIL J.: Securing web services

SAML further core concepts

- Protocols
  - Abstract form: Request / Response
  - Usually one for each of the mentioned profiles
- XML Signature
- XML Encryption
- Bindings: method how a message is transported
  - HTTP Redirect
  - HTTP POST (and HTTP-POST-SimpleSign)
  - SOAP and reverse SOAP (PAOS)
  - „HTTP Artifact“ (plus SOAP)

Base for SP ↔ IdP interaction

- Metadata
  - SPs find information about IdPs, e.g. various endpoint locations for each binding, e.g. the SSO Service
  - IdPs find informationen about SPs, e.g. the ACS
  - embedded public X.509 keys for signature and encryption
  - various extensions (logo urls, contact info, Attr needs)
  - Metadata are public and usually signed
- Common Attribute encoding format, e.g.
  - urn:oid:2.5.4.42 for givenName
  - urn:oid:1.3.6.1.4.1.10126.1.35.3.15 for „TGacceptedTermsOfUse“ (used in TextGrid)
- Synchronized clocks, HTTPS, etc...

What is Shibboleth?

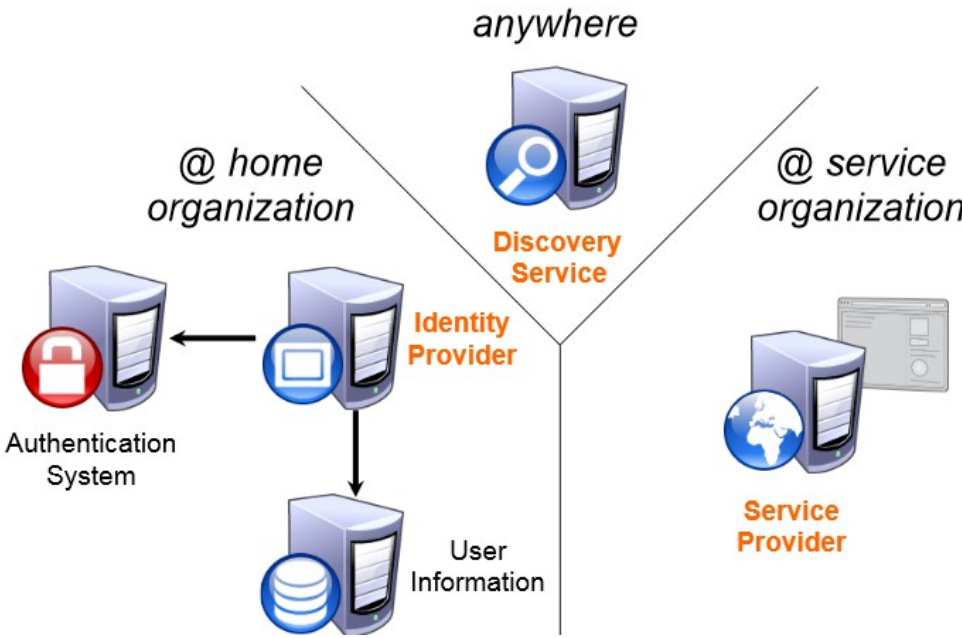
Shibboleth

AARC

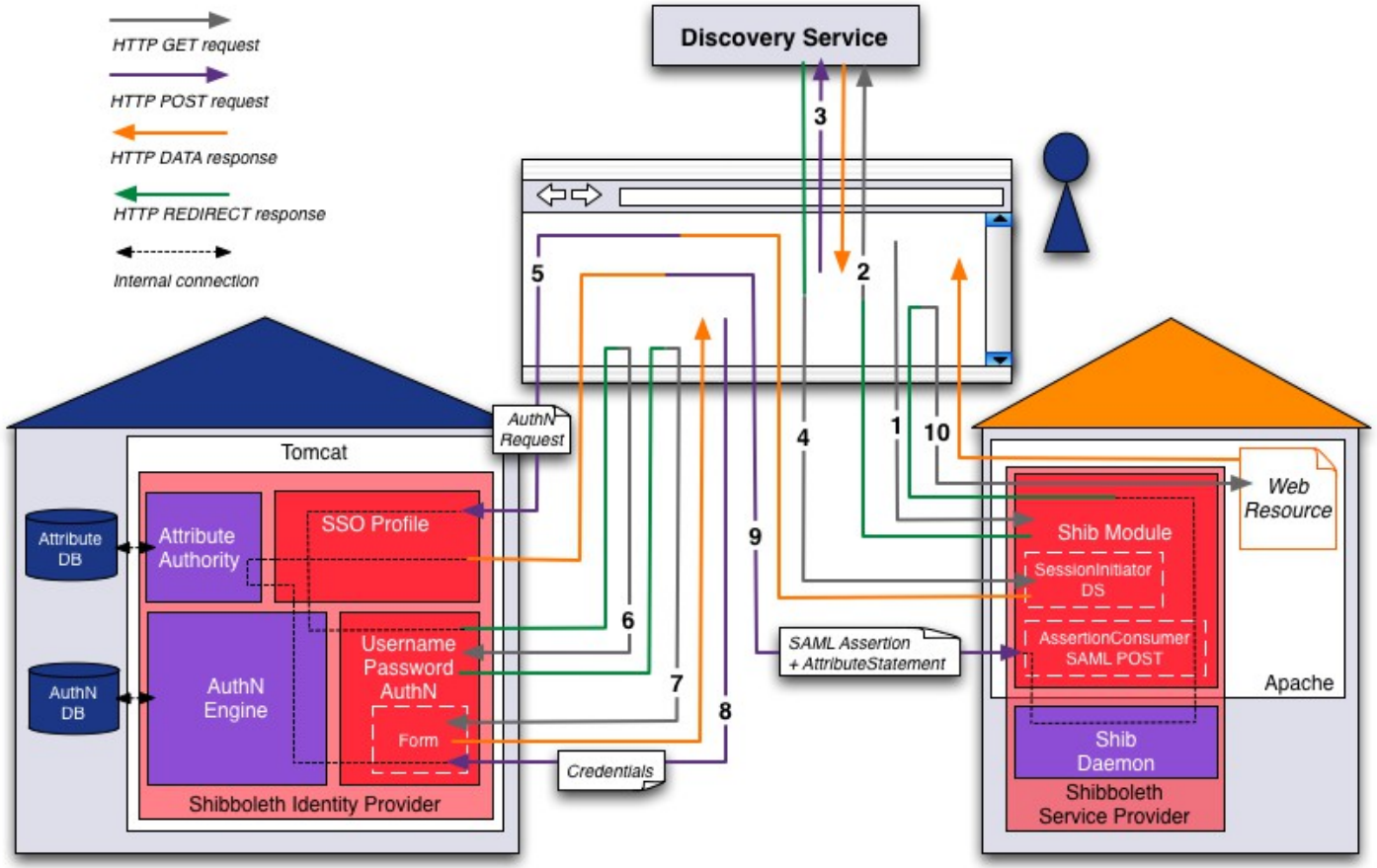
- Open-Source project, originally developed by Internet2
- Now managed by the Shibboleth Consortium
  - The new home for Shibboleth development
  - Collects financial contributions from deployers worldwide
- Implementation of
  - SAML Identity Provider
  - SAML Service Provider
  - SAML Discovery Service (Centralized and Embedded)
- Origin of the word is Hebrew, see the Bible, Judges 12,6
- The Shibboleth software is widely used in the research and education environment

Shibboleth Components

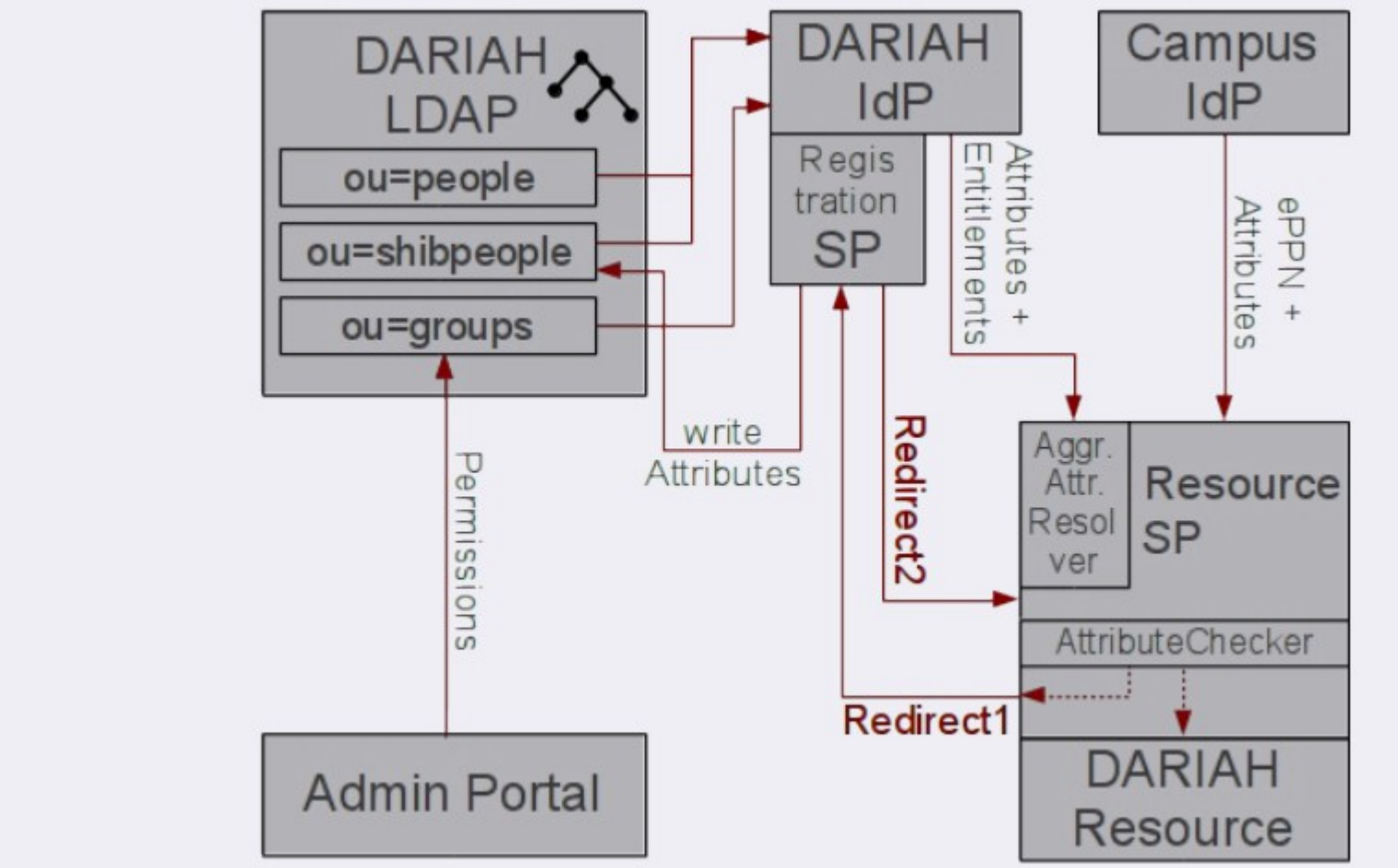
- Most people think of it as the set of software components
  - OpenSAML C++ and Java libraries
  - Shibboleth Identity Provider (IdP)
  - Shibboleth Service Provider (SP)
  - Shibboleth Discovery Service (DS)
  - Shibboleth Metadata Aggregator (MA)
- Together these components make up a federated identity management (FIM) platform.
- The Shibboleth software components are an implementation of the SAML protocols and bindings



Shibboleth WebSSO



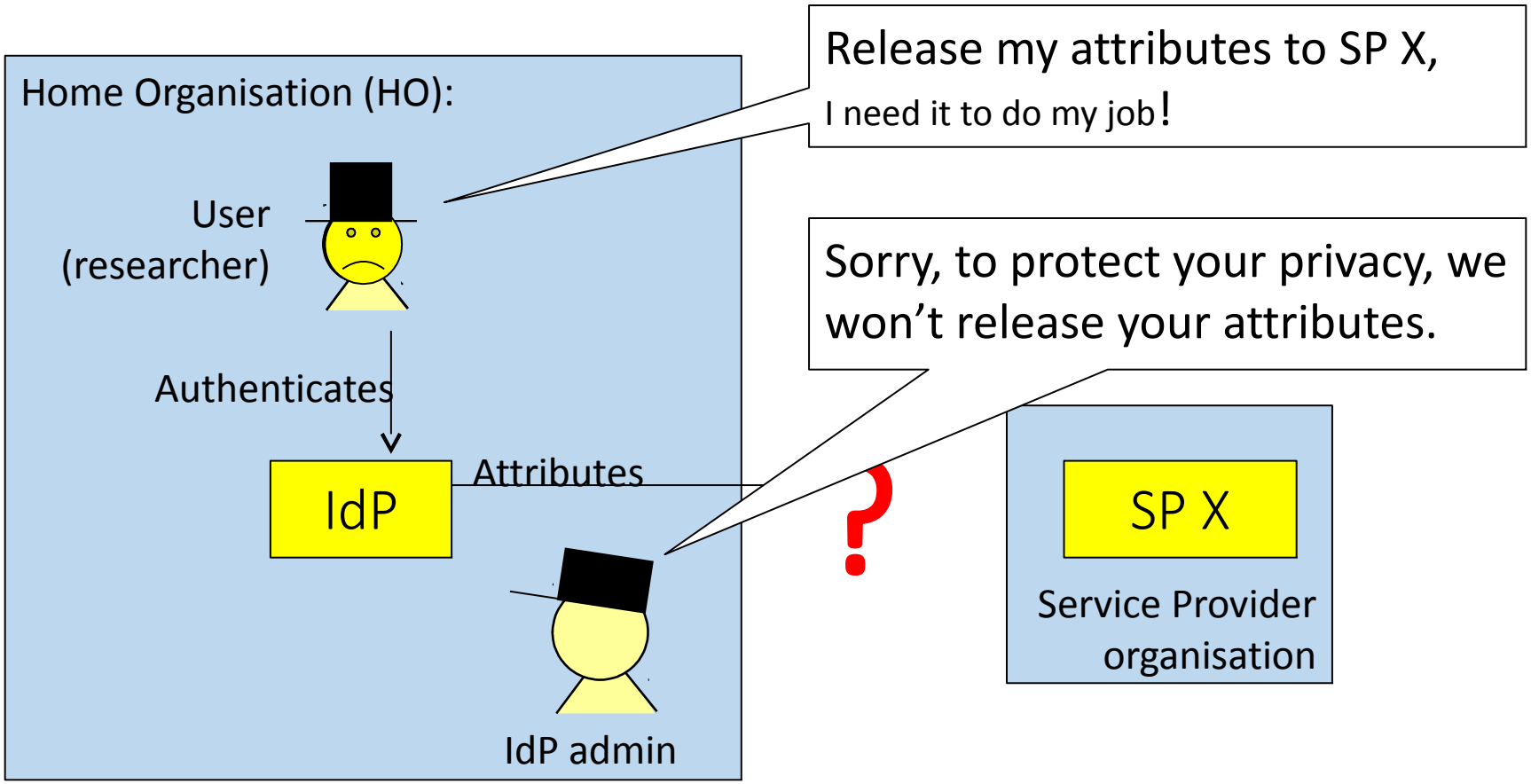
Shibboleth deployment example



Some challenges from research infrastructure point of view

- European-wide federation eduGain has too little outreach
- Not every institution signs federation contracts
- Not every Identity Provider releases personal attributes
- Technologies for non-web-based access only “almost there” (ECP, STS, Moonshot, oAuth2)

The data protection challenge in federated identity



Attributes are personal data. The data protection laws must be followed.
To be on the safe side, many Home Organisations hesitate to release attributes.

GEANT Data Protection Code of Conduct (CoCo)

17 clauses on what the SP can do with the attributes received from an IdP

- What attributes to request?
- For what purposes?
- How to inform the end user?
- How to protect the attributes?
- etc

Based on the EU Data Protection Directive (95/46/EC)

GÉANT Data Protection Code of Conduct

For Service Providers established in European Union, European Economic Area and countries with adequate data protection pursuant to Article 25.6 of the directive 95/46/EC

GN3-12-215

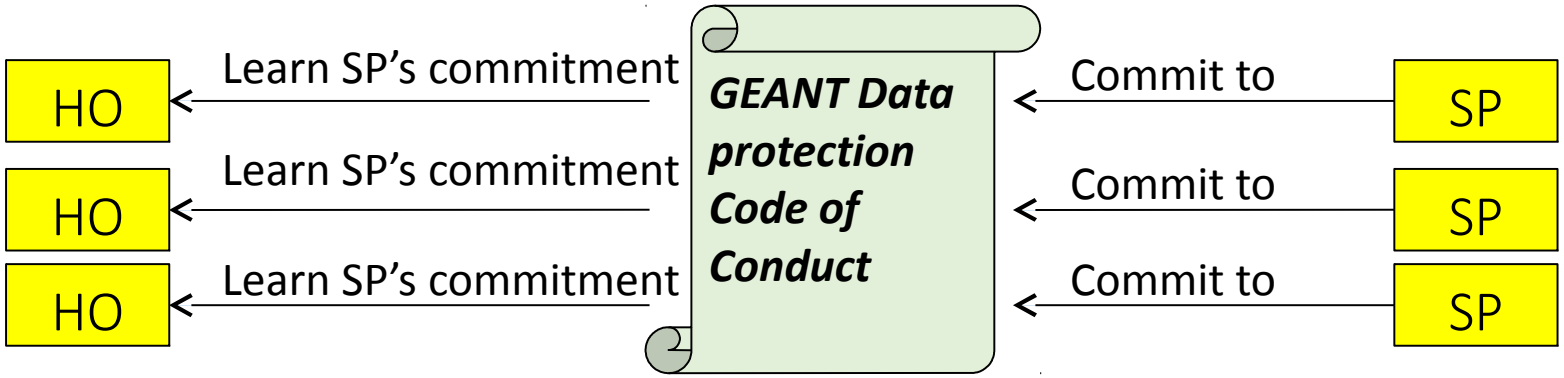
Document URI:
http://www.geant.net/uri/dataprotection-code-of-conduct/v1

Version 1.0, 14 June 2013

https://wiki.refeds.org/display/CODE/

- Normative documents
- Cookbook
- Test tools
- Training material
- Endorsement letter

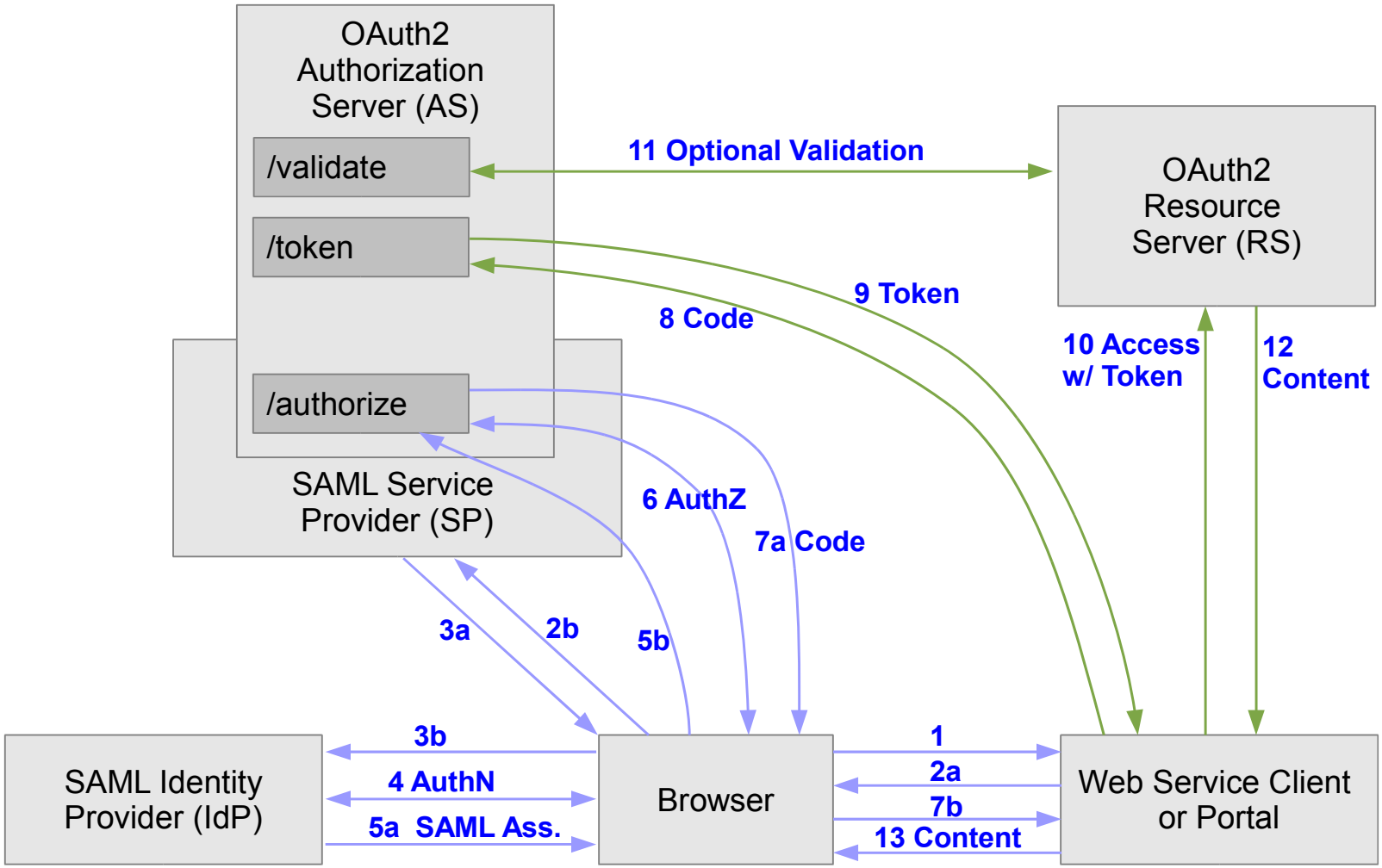
Code of Conduct workflow



1. SP commits to the CoCo
2. Federations (and eduGAIN) relies SPs' commitment to HOs
   - Using SAML2 metadata (Entity Category, etc)
3. HO decides if it feels confident to release attributes to the SP

The practicalities depend on the home federation of the HO/SP.

Non Web challenge: a solution based on OAuth2



Thank you

Any Questions?

info@daasi.de

https://aarc-project.eu

The work leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 653965 (AARC).