

dCache: status update and future directions

Paul Millar

TERENA Storage TF
Uppsala, Sweden



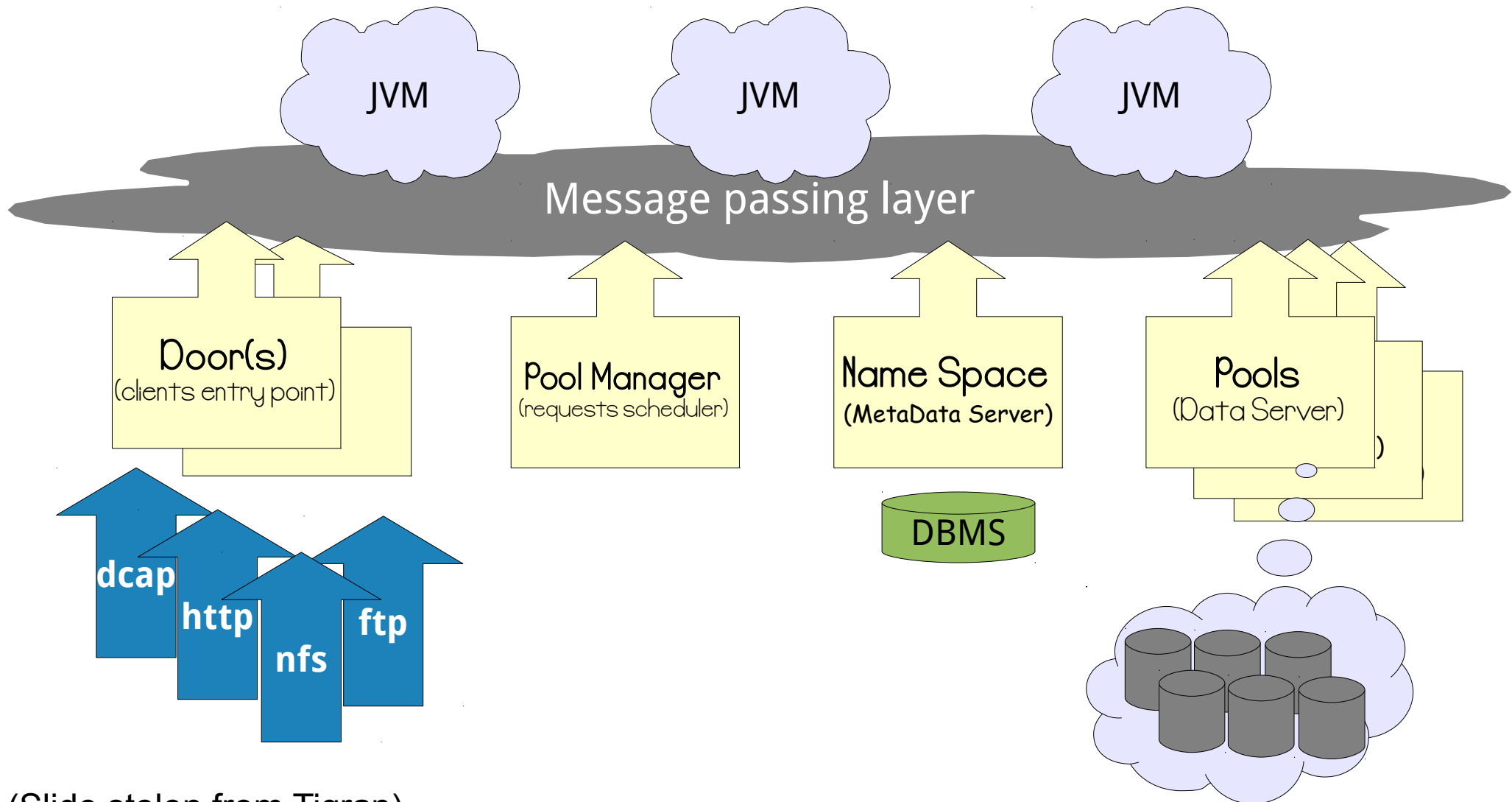
What is dCache?



Introducing dCache

- OpenSource software for aggregating heterogeneous **storage**
 - **Immutable** filesystem with its own namespace **independent** of data location,
 - Integrates with **tertiary storage** (tape)
 - Sophisticated **data-placement**
 - Built-in support for **multiple protocols** (NFS, FTP, HTTP/WebDAV, ...)
 - Consistent and coherent view of the files.
 - **Pluggable** authentication / identity system
 - Supports X.509 client cert, username+password and Kerberos
 - Integrates with site IdM: NIS, LDAP, Active Directory, Kerberos, ...
-

dCache in one slide



(Slide stolen from Tigran)

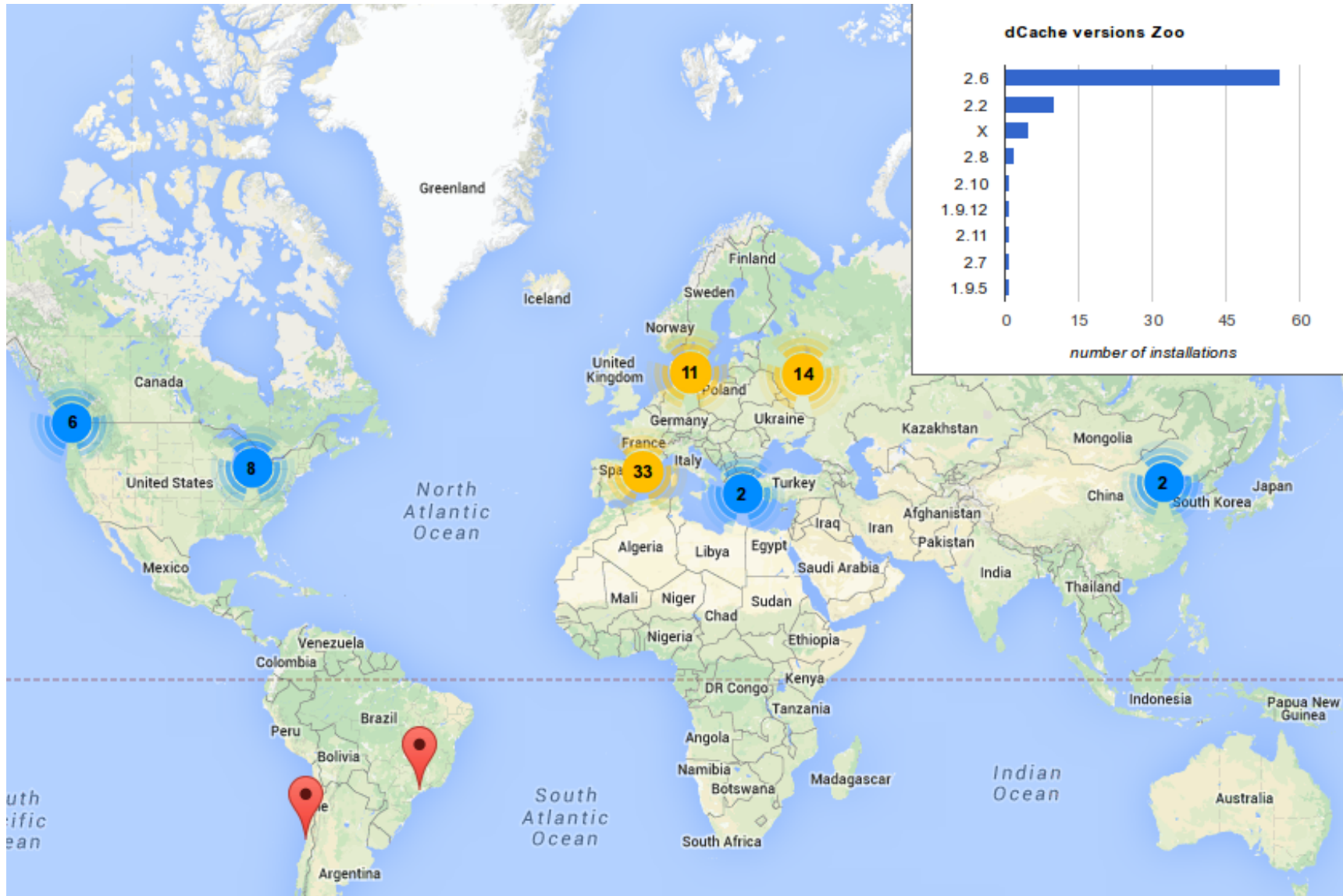
dCache: people and support

- **Core team** (8 people): collaboration between DESY, Fermilab and NEIC,
 - **Students**: HTW Berlin,
 - **External contributors**: people making infrequent contributions
 - **German support group**: volunteer dCache admins who organise and run workshops
 - Support channels:
 - **User forum** where users (i.e., admins) help each other
 - **Direct channels** (support@dcache.org and security@dcache.org)
-

dCache: funding

- **Core partners:** DESY, Fermilab, NEIC
 - **German government:** LSDMA project → PoF
 - **EU projects:**
FP7 projects (EMI) and in three H2020 proposals.
-

WLCG dCache instances (non-WLCG not shown)



Deployments (just some of 'em...)

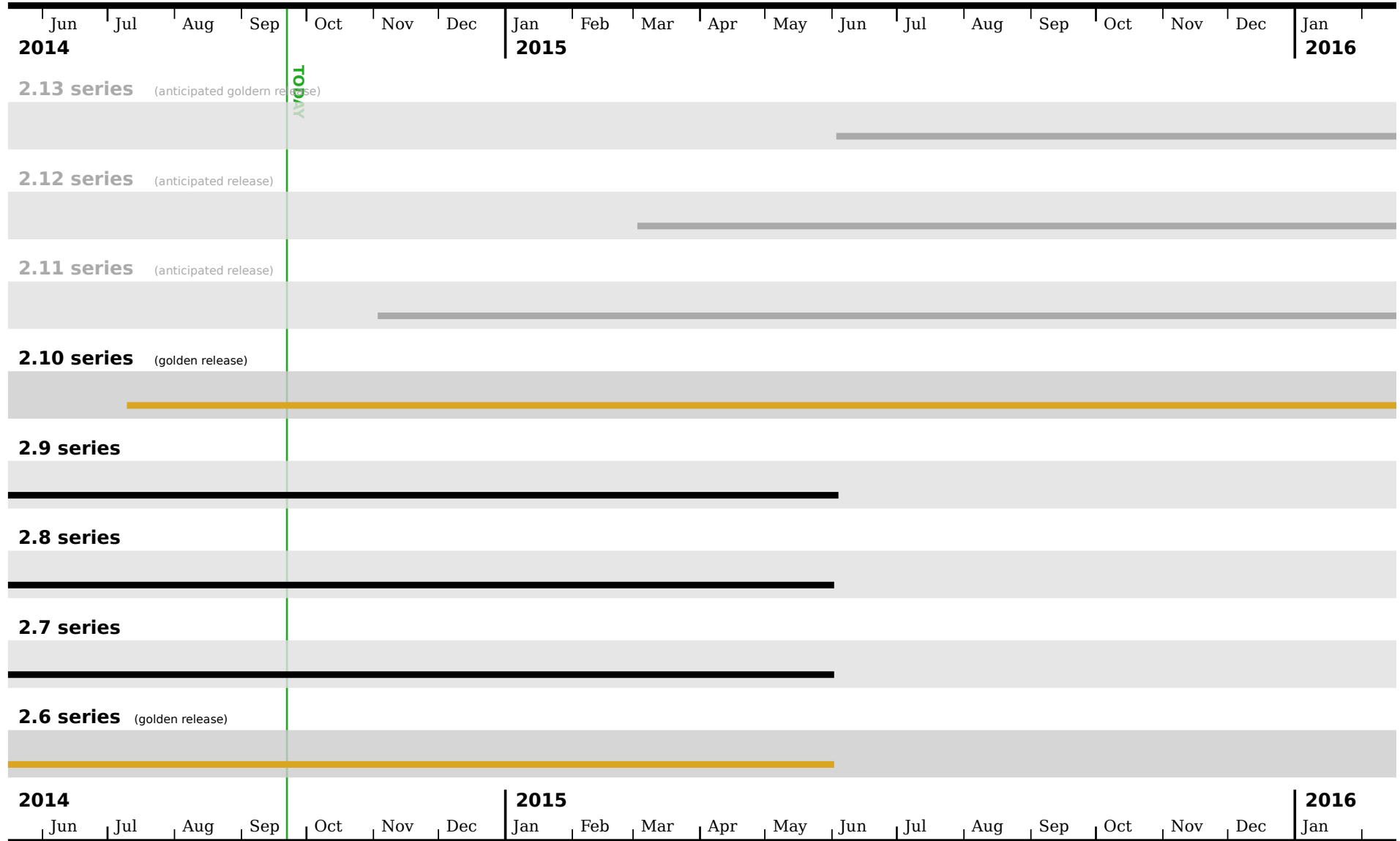
- **WLCG**: 44 sites (world-wide) together provide 100 PB, satisfying ~50% of LHC current requirement.
- **DESY**: HERA, ATLAS, CMS, LHCb, Photon science, ...
- **Fermilab**: CMS, general storage, Intensity Frontier, ...
- **BNL**: ATLAS and RHIC.
- **SNIC**: SweStore.
- **NDGF**: geographically largest single instance, spread over 5 countries.

...

<Your Name Here>

dCache server releases

... along with the series support durations.



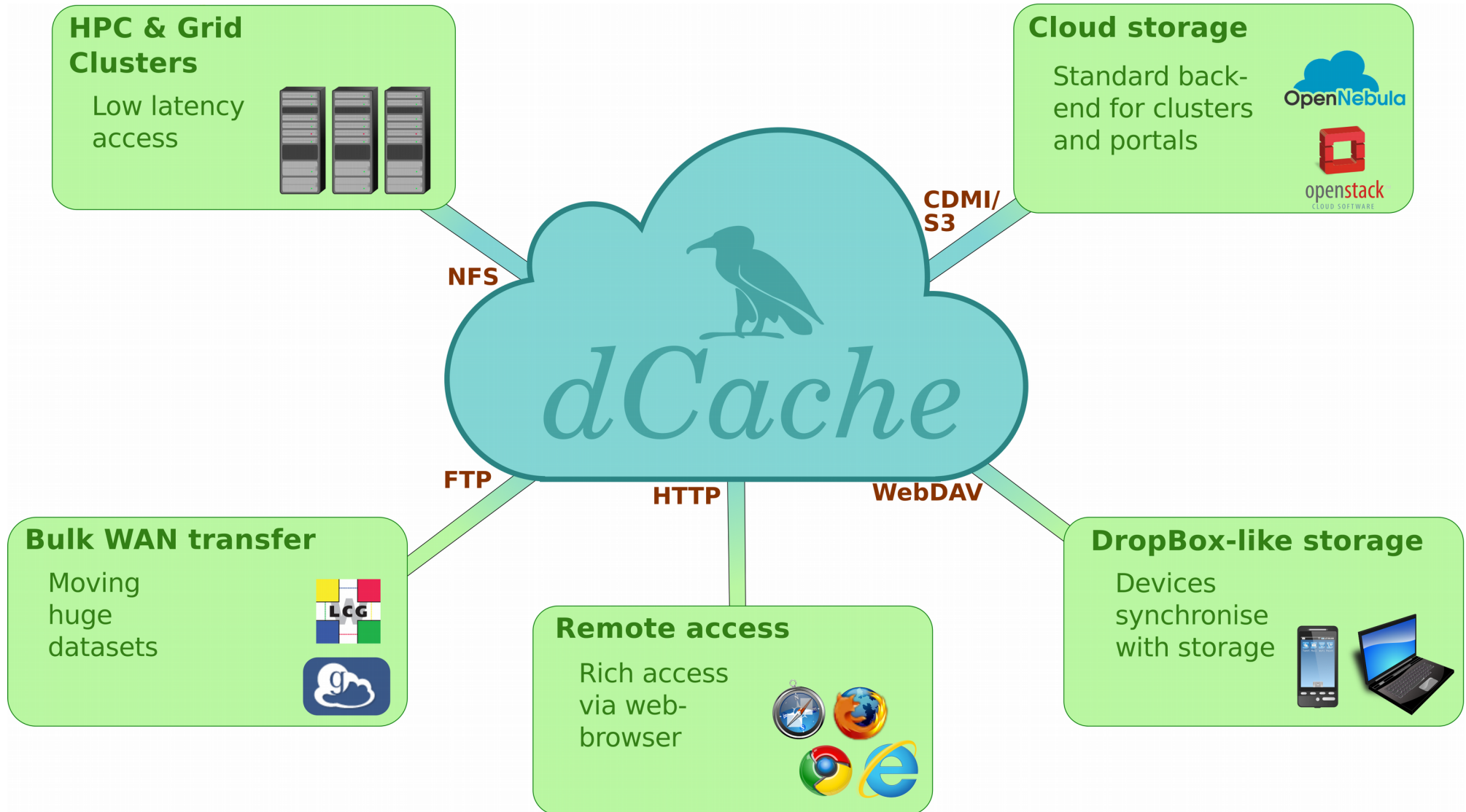
The code-base

- **Open Source** license (AGPL)
 - Code available in **github**
 - four commands (one of which is 'cd') gives you a fully functional, running dCache on your laptop.
 - **All** changes subject to code-review,
 - Large sections of the functionality are **extensible** / pluggable.
 - Spun off some functionality as **independent libraries**:
 - Code used by banks, other storage system vendors, ..
 - We only know who from the bug reports and bugfixes
-

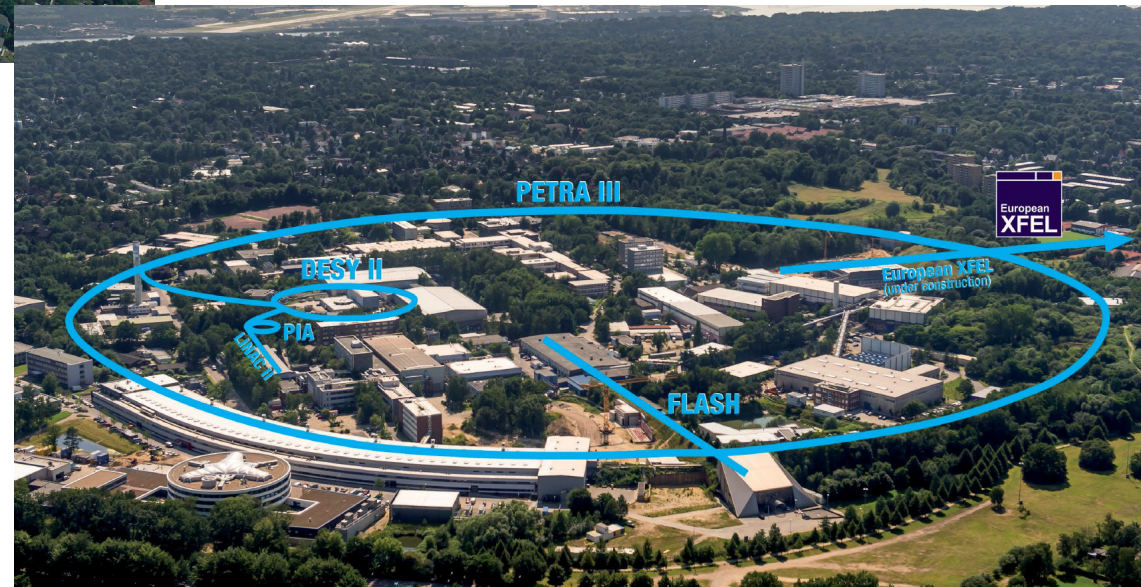
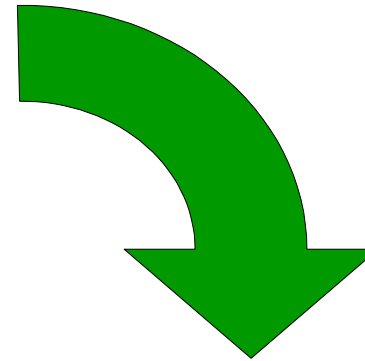
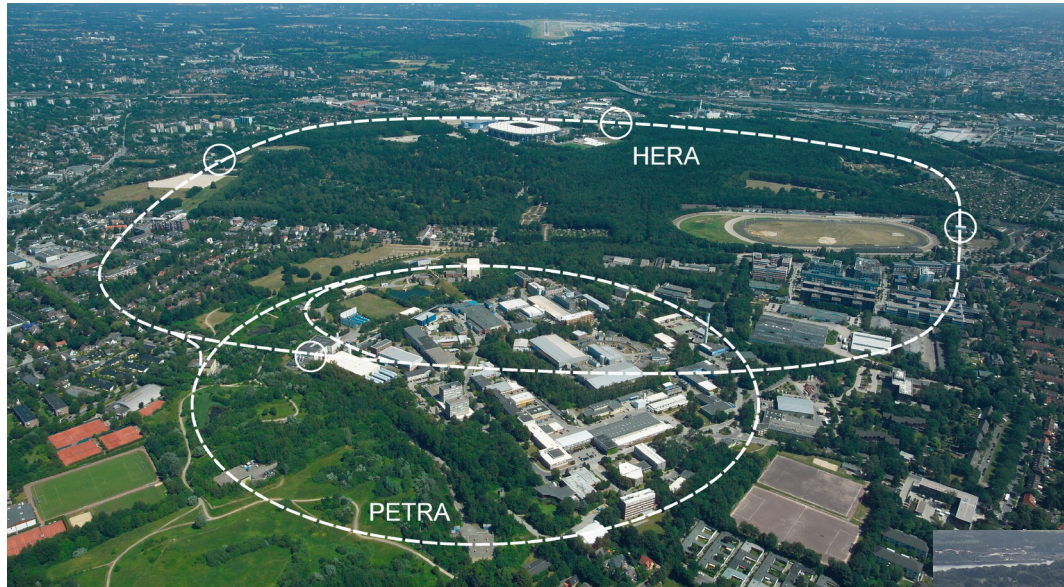
Status updates and Future directions



dCache the scientific cloud



Improving data-injection performance



How to store small files on tape

- Small files are **bad for tapes**
 - Load/seek time vs read time.
 - Random selection → many tape mounts → slow access & broken tapes.
 - Solution: dCache collects files in **a container** (a zip file) before writing to tape
 - Replacing lots of shuttle-buses with one big bus
 - When user **writes** new files:
 - “Small files” are written into dCache,
 - dCache groups files and, based on policies, writes a container back into dCache,
 - Containers are written to tape.
 - When user opens a file for **reading**:
 - Fetch container from tape, if not cached
 - Extract file from container
 - User sees no difference, yet tape is better utilised.
-

HTTP and WebDAV

- Added support for HTTP and WebDAV.
 - Support redirection on read, redirection on write.
 - Metadata operations can be **encrypted**; when redirected, data is transported **unencrypted**.
 - Found problems with (almost) all webdav clients.
 - **Extending WebDAV** to include additional functionality:
 - Added support for triggering 3rd party copy,
 - Added support for recovery in dynamic data federation.
-

HTTP Federation

- Project in collaboration with CERN
 - Multiple HTTP/WebDAV servers provide users an **overlap namespace**
 - Like partial mirrors of some central repository
 - Central server provides an **aggregate view**
 - Assume that if files exists in multiple server, they are identical replicas
 - Client sees all available files
 - When reading, the client is **redirected** to “best” replica.
 - Available as a demo; being evaluated by WLCG experiments
-

Developing dCache sync-n-share

- Provide unlimited storage: ✓
 - Access via web-browser: ✓
 - Synchronisation: ✗ → ✓
 - Sharing: ✗
 - how do we present shared data to the user?
 - how do users share data with others?
-

DESY sync-and-share service

- DESY users needed to **stop using DropBox.**
 - dCache already started working on adding sync-and-share facilities.
 - For DESY, using **dCache and ownCloud** to build a DropBox-like service was the best option.
-

dCache with ownCloud

- Use ownCloud on top of dCache, via NFS
 - Files in dCache **owned by the user** (*not* ownCloud process)
 - Users can write data into dCache
 - Immediately** visible through ownCloud.
 - Users can write data into ownCloud (sync client)
 - Immediately** visible through dCache
 - Limitations:
 - If user shares data with you, you can only read that through ownCloud.
 - If you set ACL in dCache, not reflected in ownCloud
 - Service goes live **today** (for the brave); DESY-wide in two weeks.
-

What is the sync-n-share future?

- Have the client sync directly with dCache.
maybe the ownCloud client
 - Add support for sharing within dCache.
enhanced web interface
 - Drop ownCloud and provide a pure dCache solution.
-

CDMI: managing cloud storage

- **Network protocol** for Cloud storage
 - initially by SNIA, now an ISO standard
 - with many, many features
 - Limited vendor uptake:
 - Catch-22: demand and availability
 - Some **IAAS** systems use CDMI internally,
 - the EGI FedCloud has CDMI as a common requirement
 - **Preliminary support for dCache** from student project,
 - Not available now, but plan to integrate (after code review)
 - What is the demand?
-

gPlazma: flexible identity management

- dCache's **IdM** identity management system:
 - (mostly) authenticates user,
 - figures out their uid, gid(s),
 - rejects banned users,
 - discovers session information: home directory ...
 - **Public API**: anyone can write a plugin.
 - We **supply plugins** for NIS, LDAP, ActiveDirectory, Kerberos, X.509, VOMS, XACML, PAM and some local files (e.g., htaccess).
-

Federated Identity

- Increasing need to “do something”
 - **SAML** seems prevalent system
 - OpenID Connect is also gaining traction.
 - With LSDMA: initial work on **credential translation** (SAML → X.509)
 - Later, add **native SAML** support:
 - Initially with Web-SSO, later maybe Moonshot/AbFab.
-

Globus (Online)

- Globus (Online) provides a **file-movement service**,
- Data connections always authenticate via **X.509**

Globus can use externally-generated credentials

- LSDMA providing a “glue” service:
 - Germany's DFN-AAI run a SLCS (a bit like TCS).
 - The glue service allow Globus users to use the SLCS.
-

Software Defined Storage & QoS

- dCache can already provide **differentiated QoS** (Quality of Service):
 - Different files can have different replication factors, multi-tier (SSD, HDD, tape) usage, utilise different hardware
 - Currently these QoS attributes are most configured by the **dCache admin**.
 - We are investigating SDS to allow:
 - Modification of QoS after data is written,
 - Allow users finer grain control of QoS choices.
-

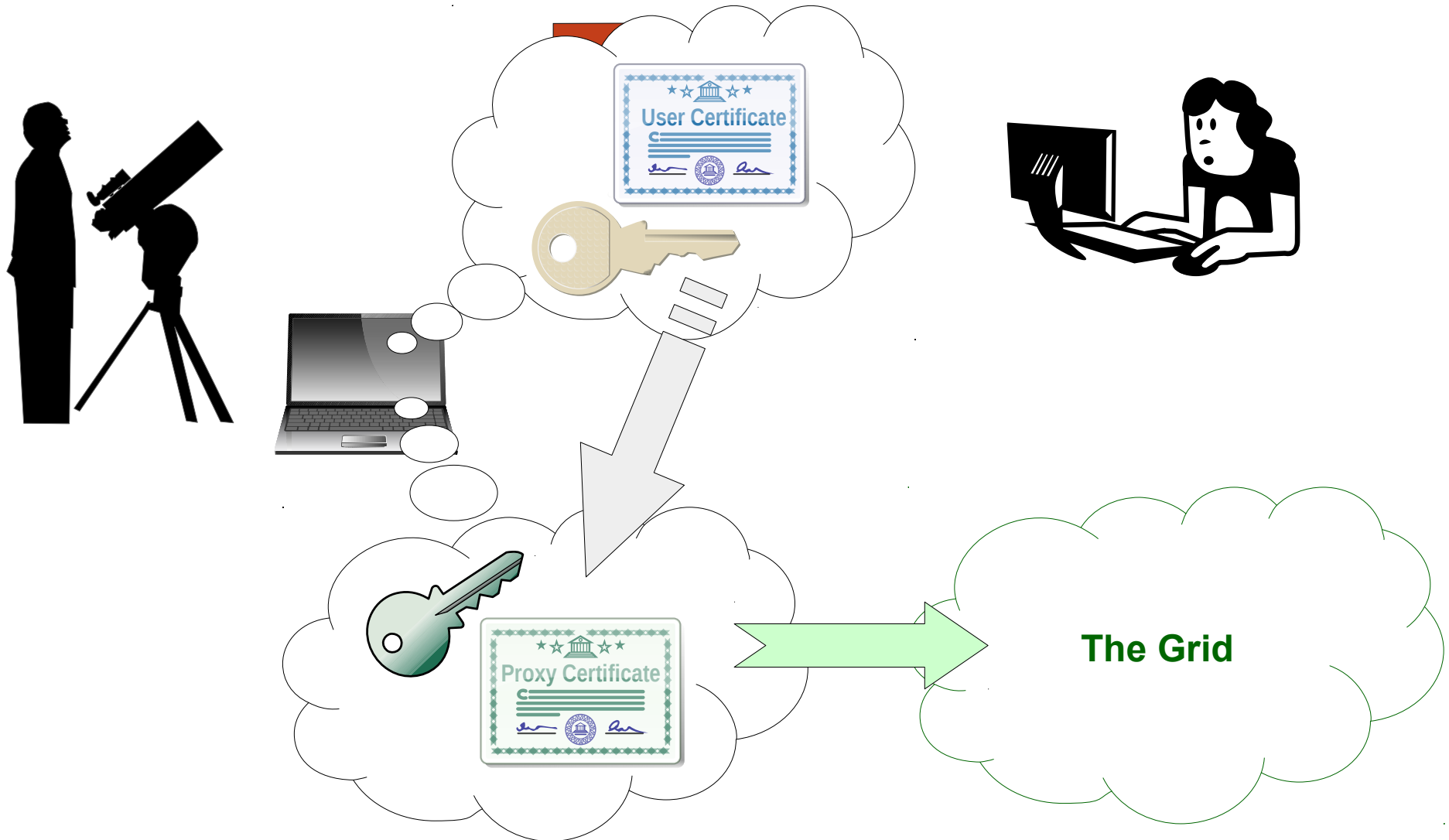
Summary

- We are adding Cloud-like features, both interactive (currently via ownCloud) and through protocols (like CDMI) – rolling out a **production service** at DESY.
 - Investigating how to integrate support for **Federated Identity** into storage software
 - For more than 10 years, dCache provides Big Data storage software that:
 - focuses on users needs,
 - implements state-of-the-art features,
 - pushing user expectations by exposes users to innovation.
-

Backup slides



The grid solution: X.509 (user) certificates



Federated Identity



Check who you are
&
Authorisation decision

Check who you are




Record information




Authorisation decision

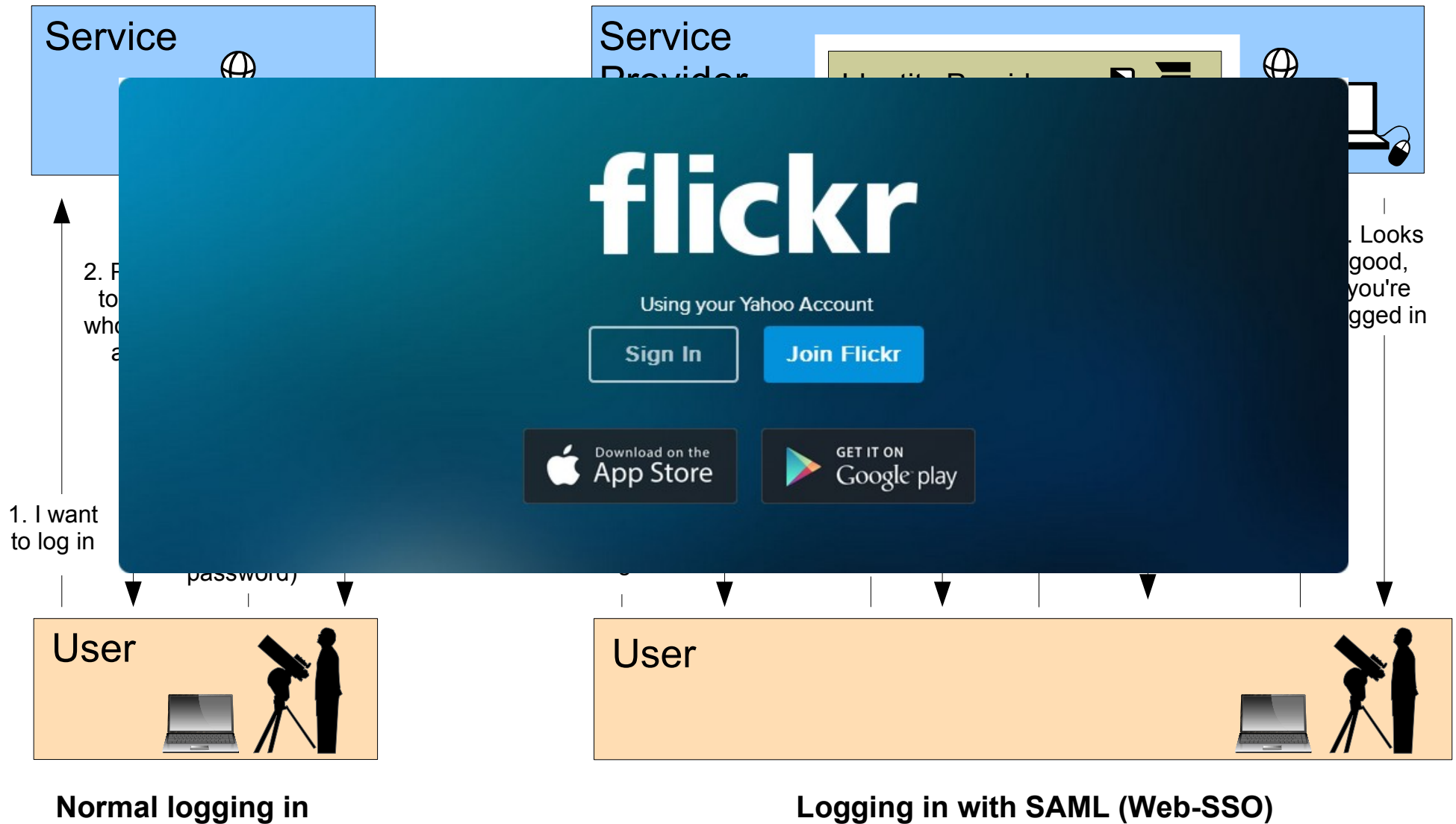


Identity Provider (IdP) 

Assertion

Service Provider (SP) 

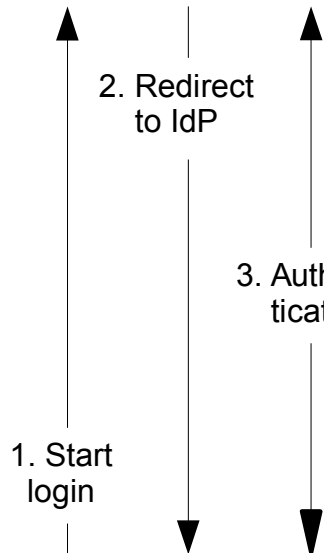
SAML Web Single Sign-On (Web SSO)



“Where”

Service Provider (SP)

IdP (IdP)



User

SAML Web

DFN-AAI - Iceweasel

DFN-AAI

https://wayf.aai.dfn.de/DFN-AAI/wayf/WAYF

DFN
Deutsches Forschungsnetz

DFN-AAI

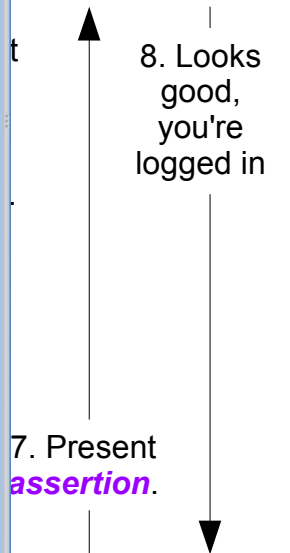
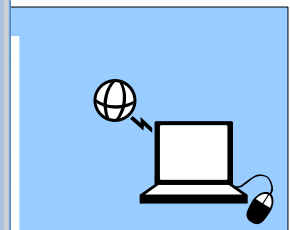
Permanently set your Home Organization

On this page you can set a **default Home Organization** for this web browser. Setting a default Home Organization will hencefort redirect you directly to your Home Organization when you access AAI-Resources. Don't use this feature if you use several AAI accounts.

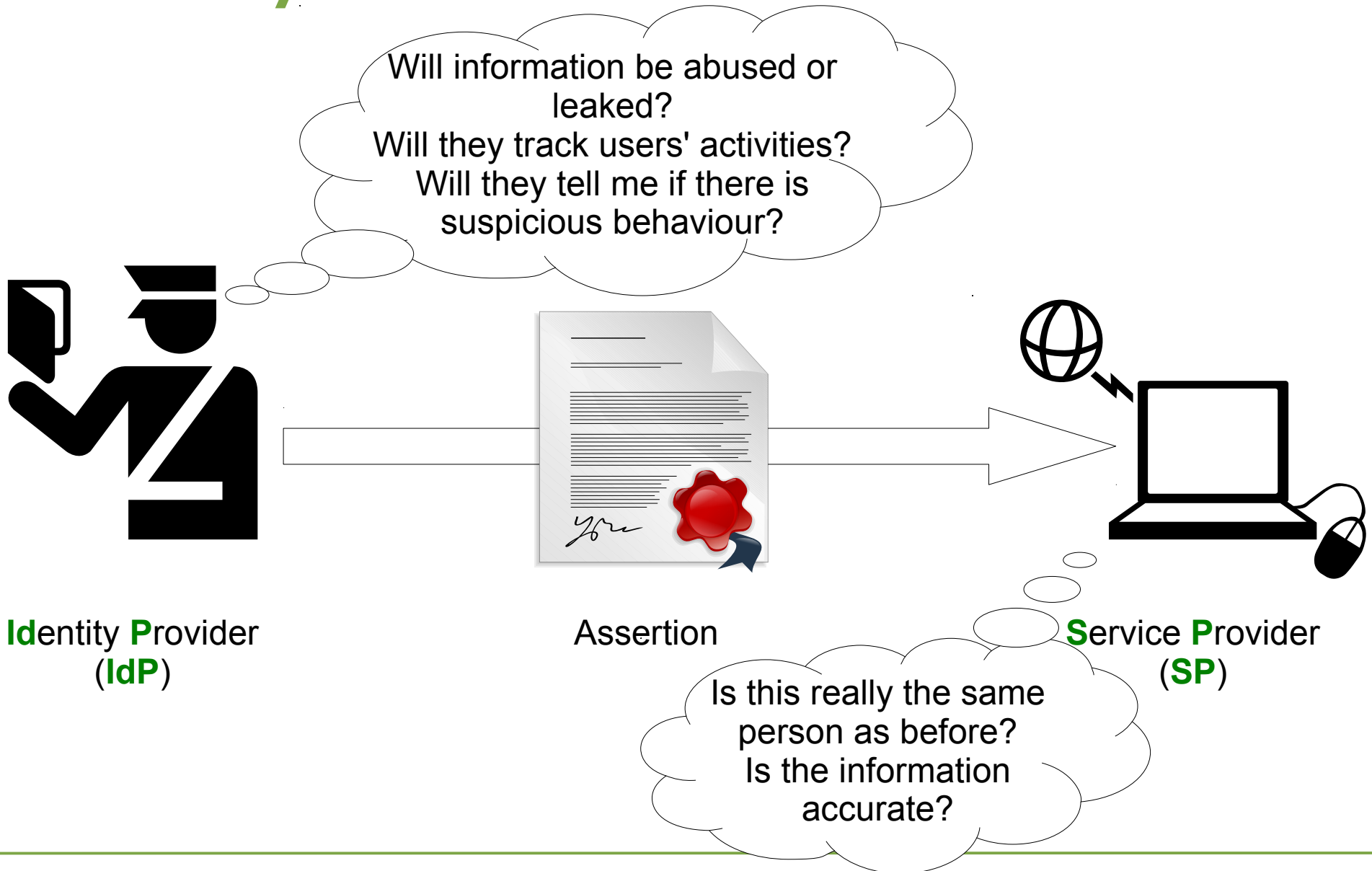
Remember selection permanently and bypass WAYF from now on.

Select the Home Organization you are affiliated with ... Save

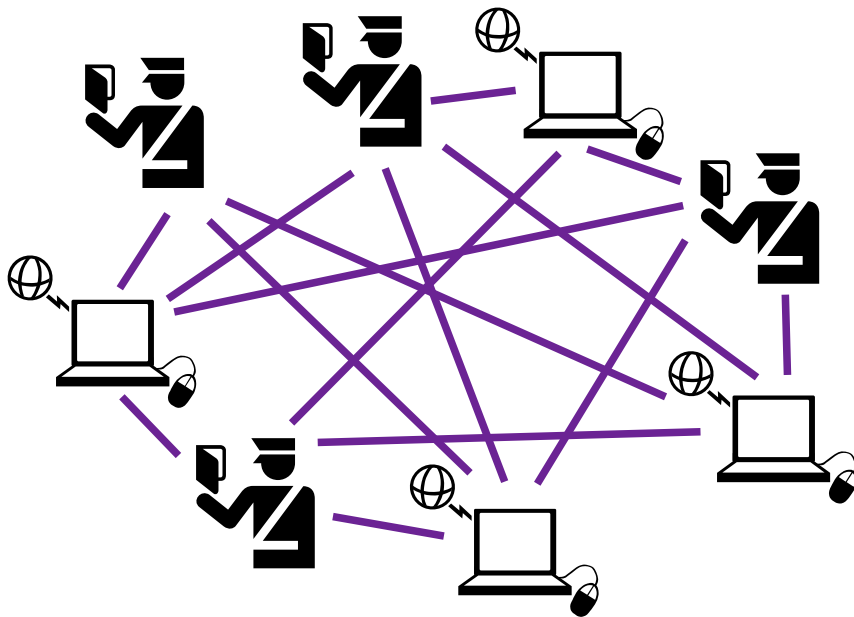
- DFN - Deutsches Zentrum für Luft- und Raumfahrt e.V.
- Europa-Universität Flensburg
- FH Worms
- Fachhochschule Düsseldorf
- Forschungszentrum Jülich GmbH
- Fraunhofer-Gesellschaft
- Freie Universität Berlin
- Friedrich-Schiller-Universität Jena
- GESIS - Leibniz-Institut für Sozialwissenschaften
- Georg-August Universität Göttingen
- Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen
- HS-Harz
- HSU Hamburg
- HTWG Konstanz**
- Helmholtz Zentrum München
- Helmholtz-Zentrum Dresden-Rossendorf e.V.
- Helmholtz-Zentrum für Umweltforschung - UFZ
- HfWU Nürtingen-Geislingen
- Hochschule Aalen - Technik und Wirtschaft
- Hochschule Albstadt-Sigmaringen
- Hochschule Augsburg



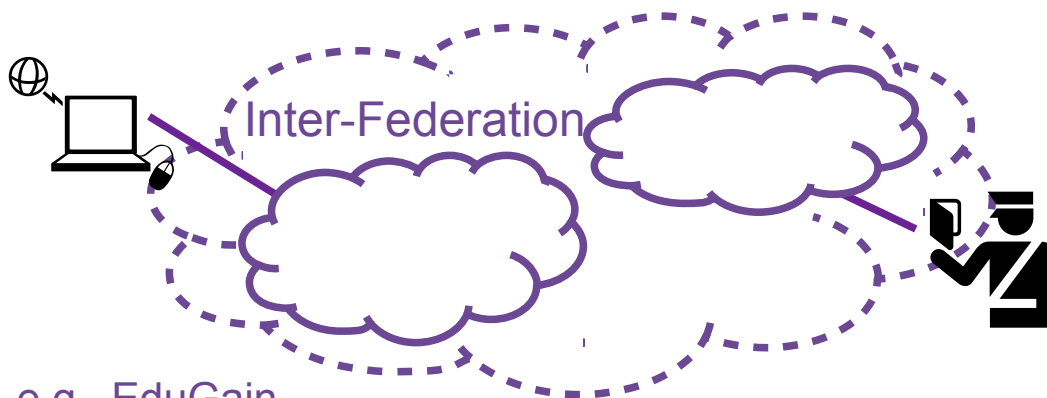
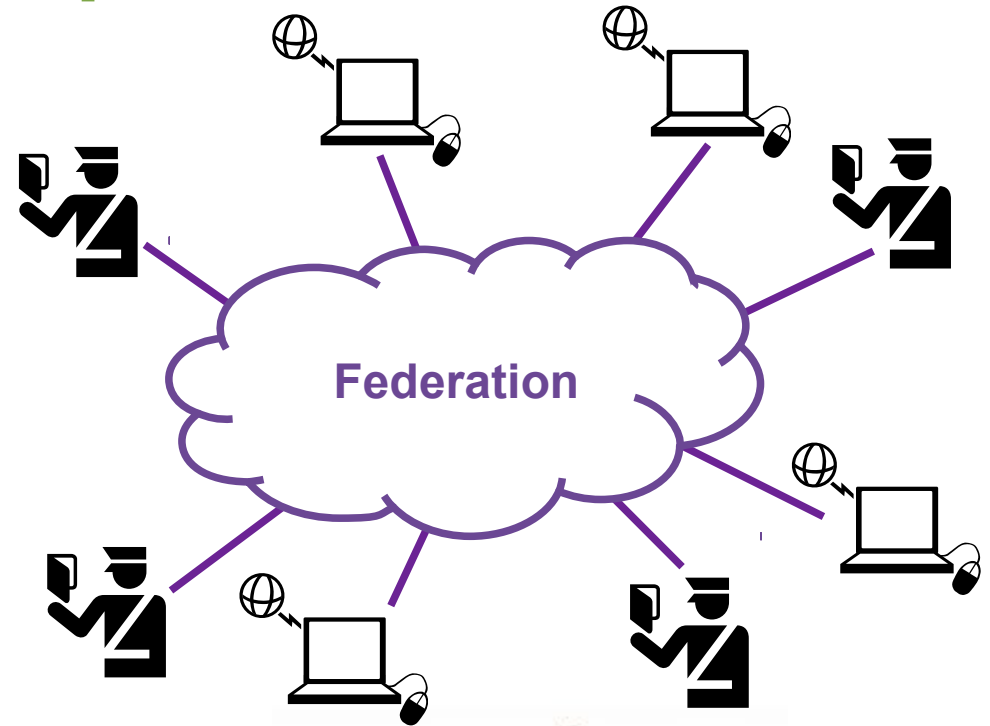
Who do you trust?



How to trust lots of people?



Point-to-point **trust** doesn't scale!

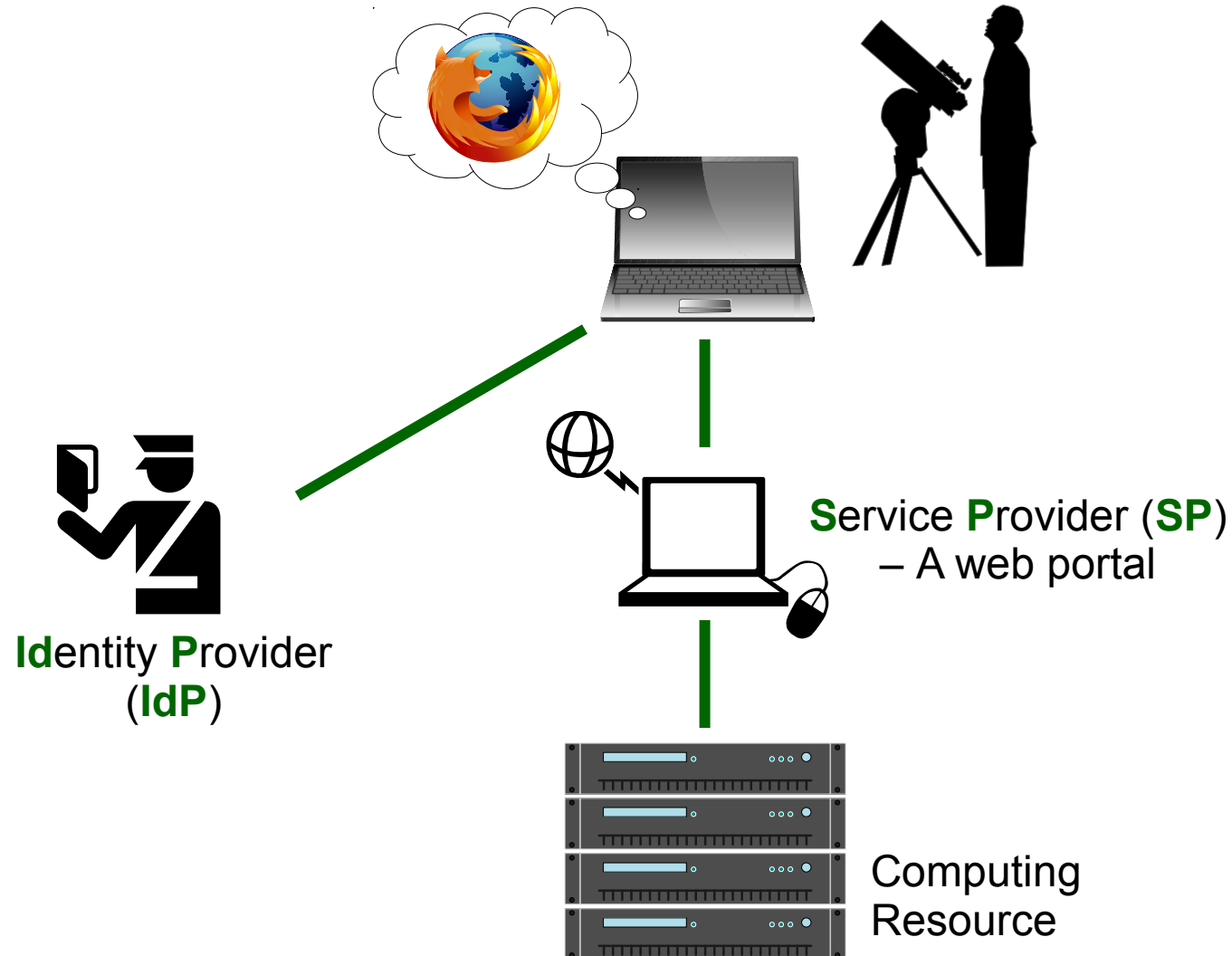


e.g., EduGain



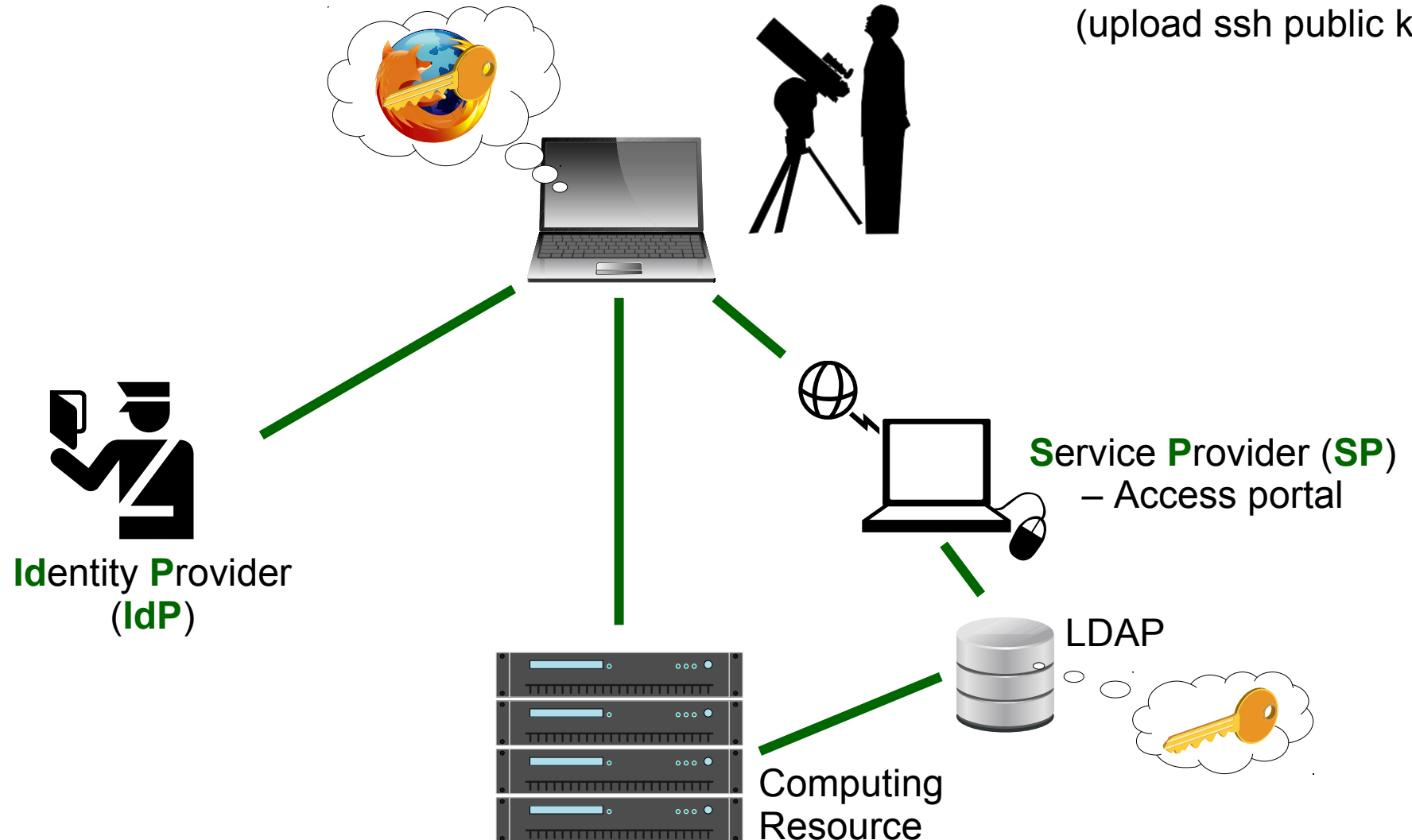
Using (remote) computers

Web portal



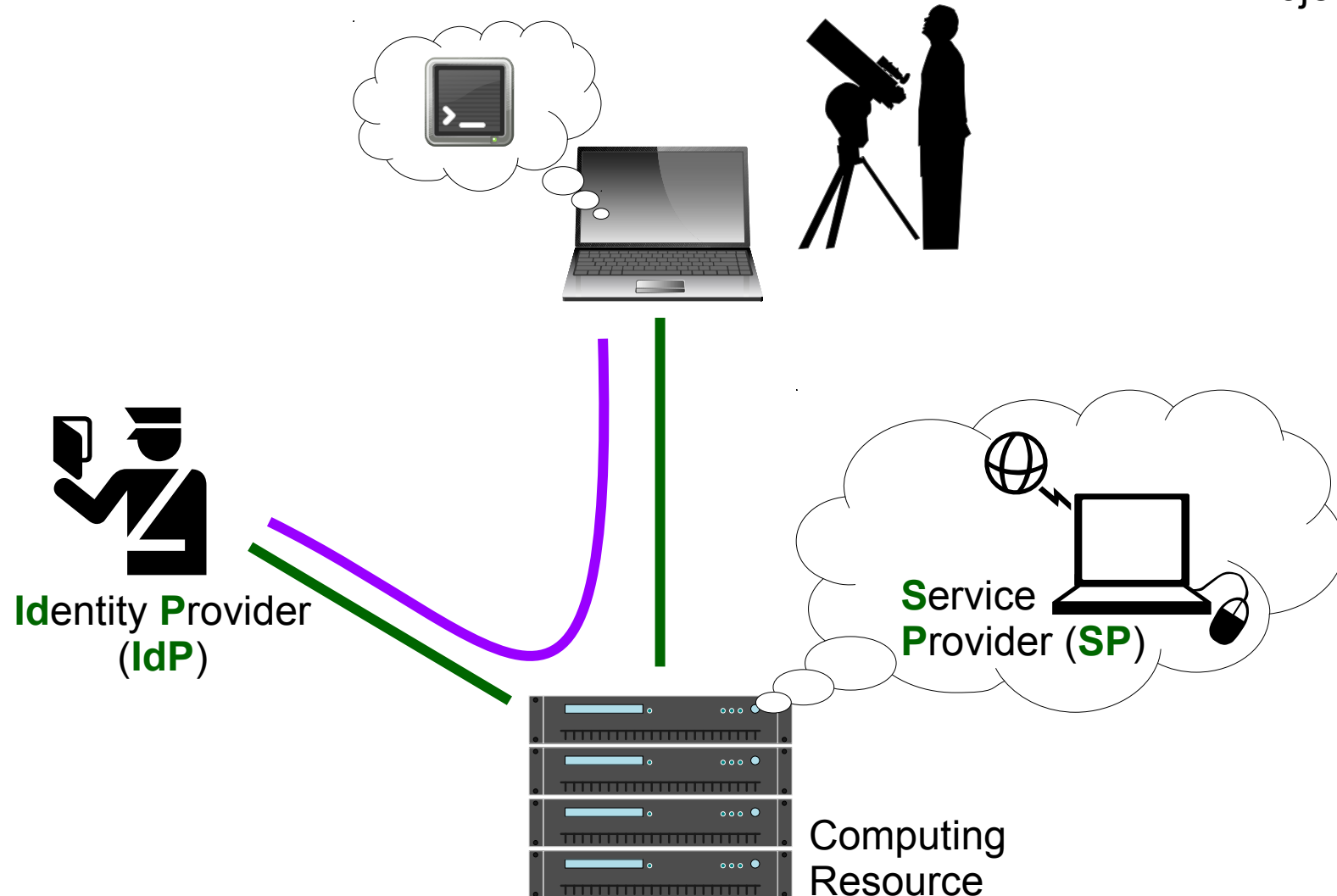
Using (remote) computers

Substitute Credential
(upload ssh public key)



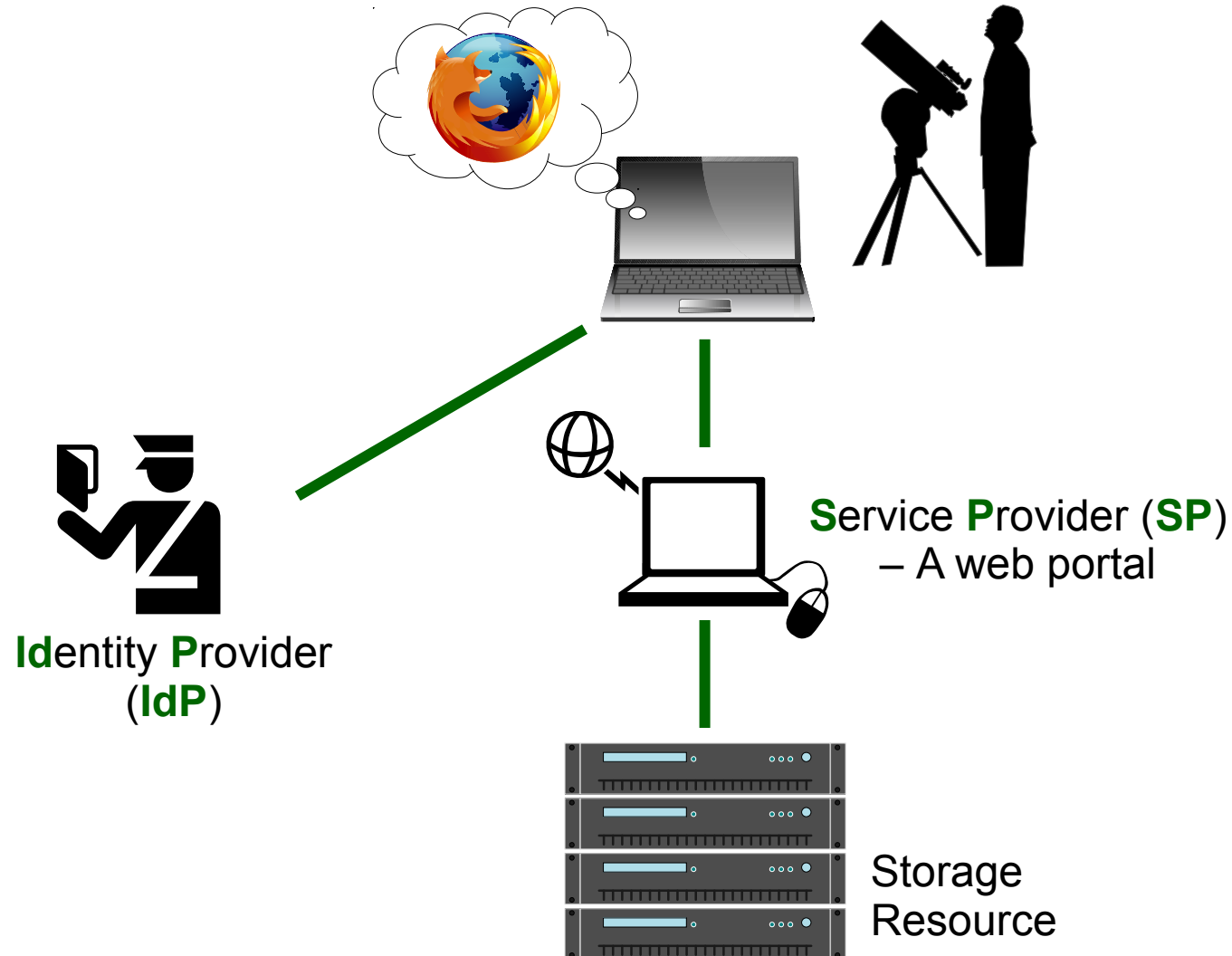
Using (remote) computers

Project Moonshot



Managing (remote) data

Web portal

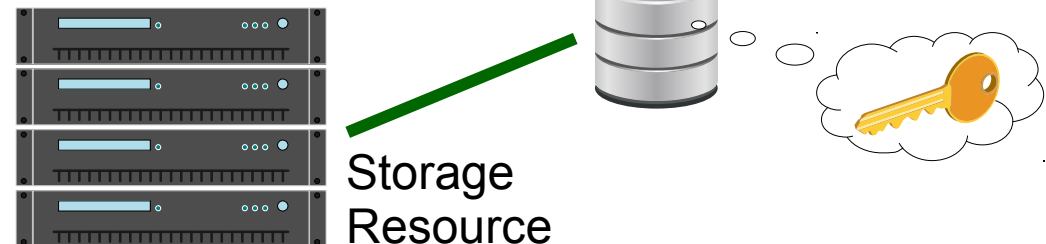
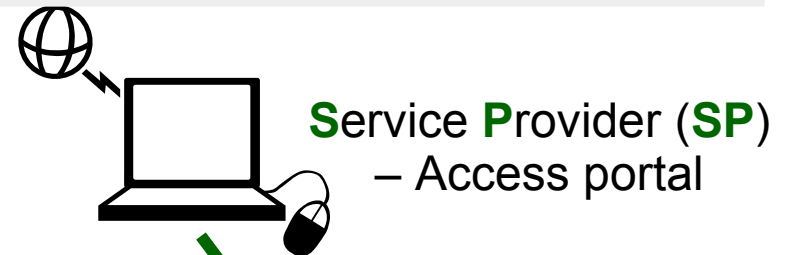


Managing (remote) storage

Fetch Substitute Credential

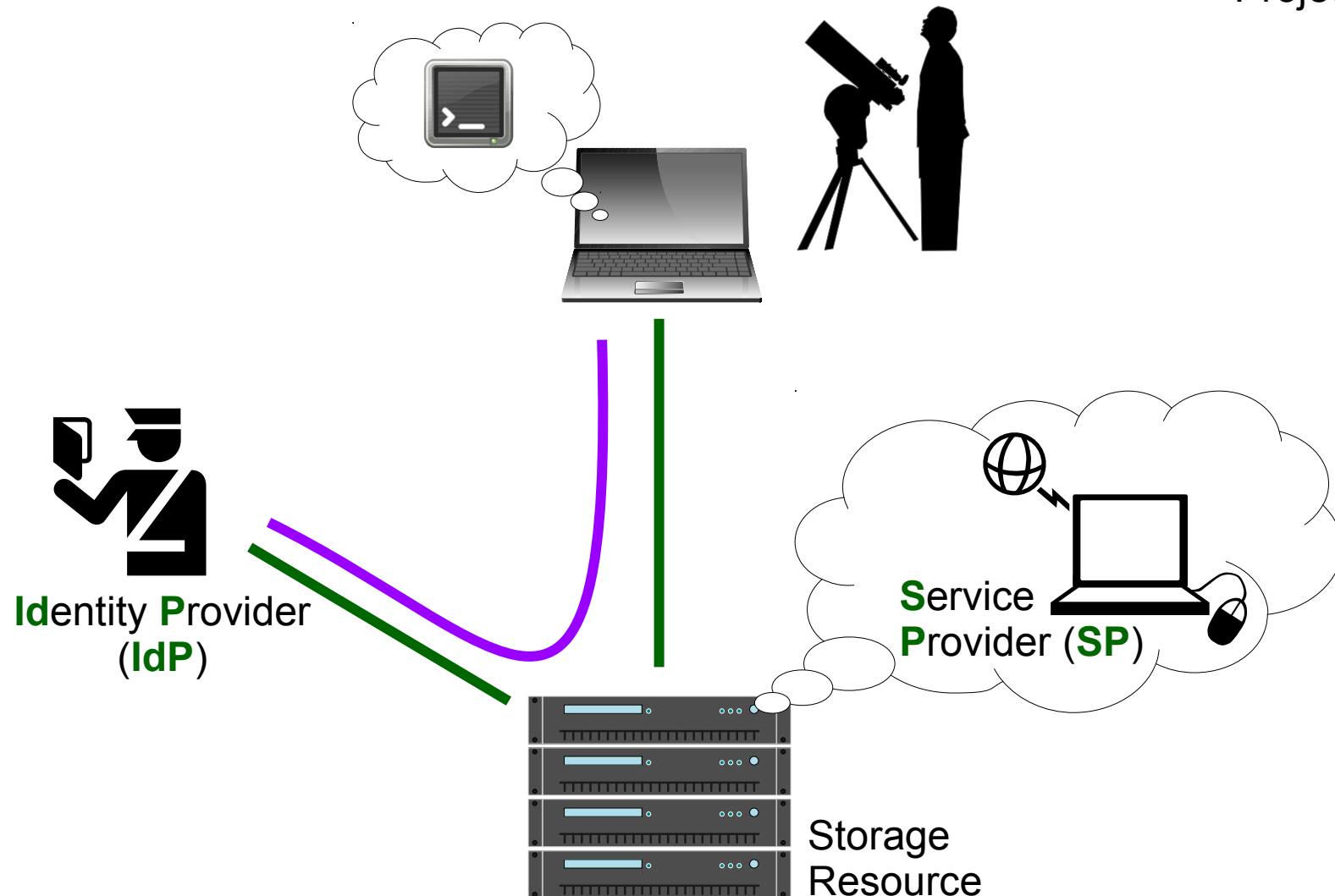


**Token: Amazon AWS/S3 SAML support,
X.509: SLCS, TCS, CI-Login, EMI STS, ...**

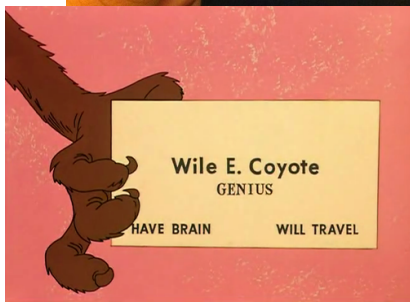


Managing (remote) storage

Project Moonshot



Credential vs Principal



Name: **Wile E. Coyote**

ACME customer ID: **11493**

Passport number: **0008103314**

Bank account number: **001213921**

Banks with: **United ACME Bank**

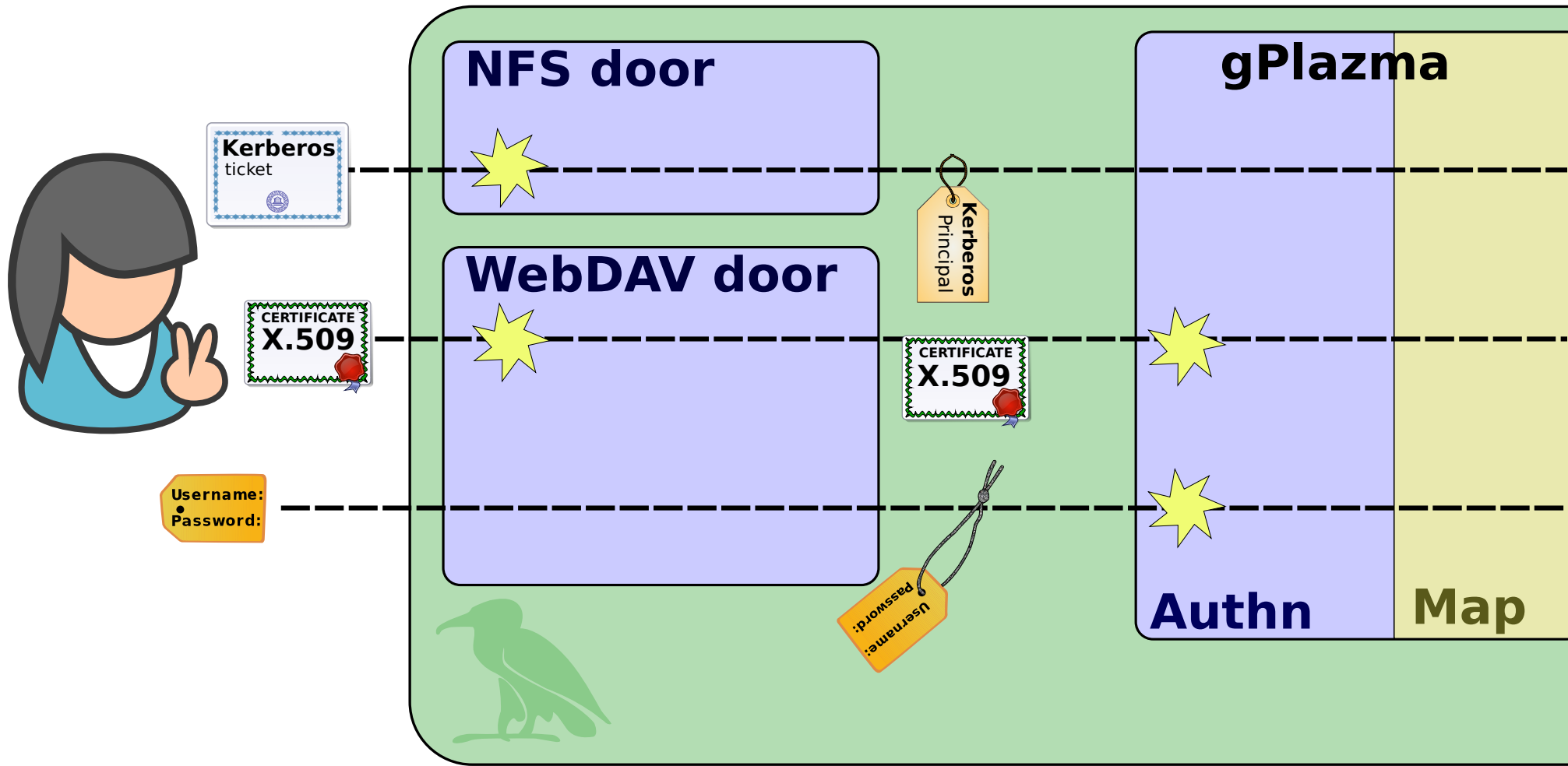
Member-of: **Antagonists Anonymous**



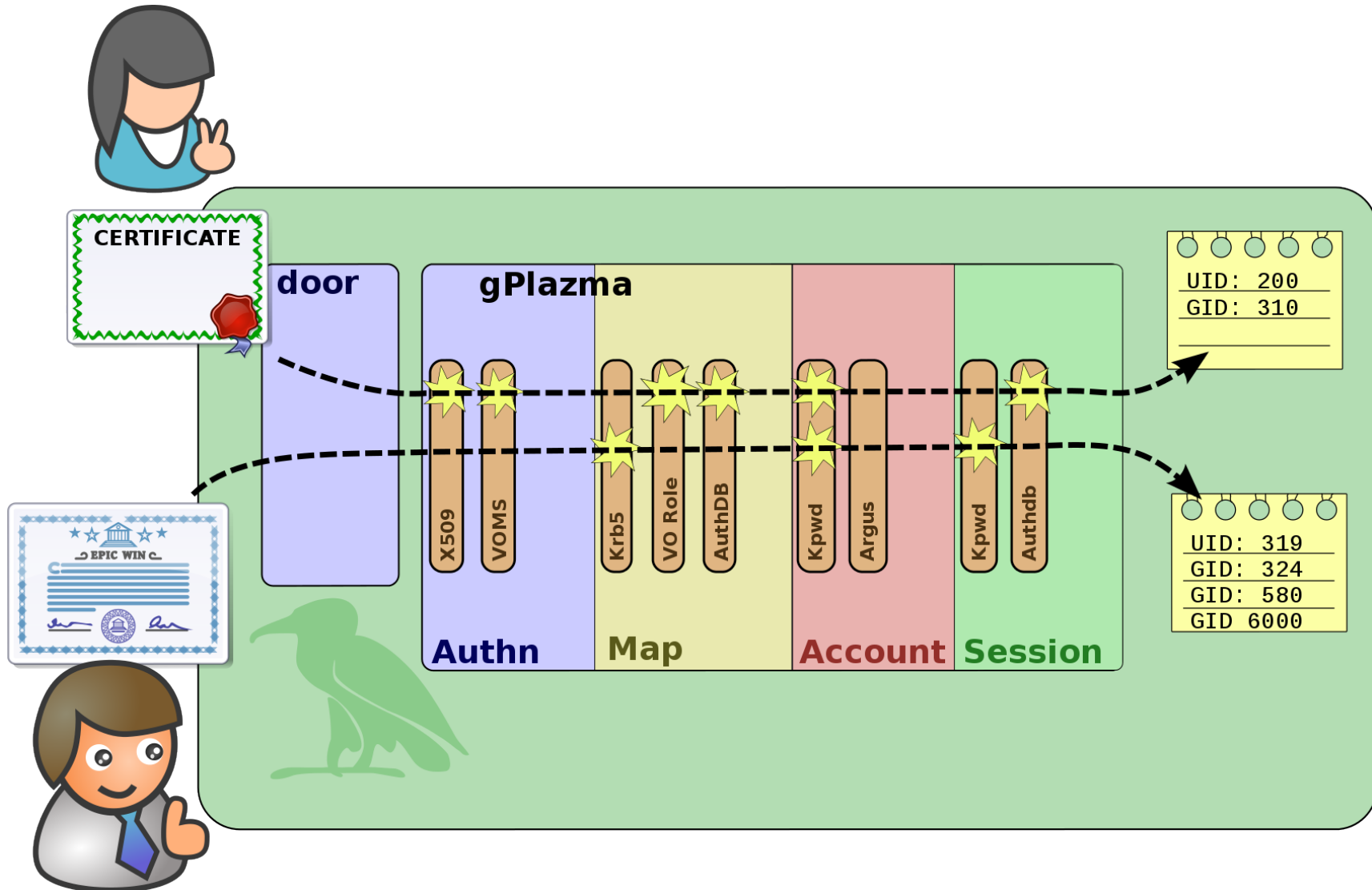
Credentials

Principals

Authentication: door, both or gPlazma



Logging in: four phases, using plugins



Something extra: identity mapping

