dCache integration into HDF

Storage service at DESY for Helmholtz Data Federation (HDF)

Paul Millar
Karlsruhe, 2018-08-30

HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

European Commission

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 653549 and 777367

DESY.

What is dCache

Used by scientists for photon science, astronomy, particle physics, genomic analysis, neutrino experiments, medical, …

Widely deployed storage software, used by institutes throughout the world, often with users throughout the world.

A collaboration between DESY, Fermilab and NEIC

Storage software that allows a site to ...

- move beyond limitations of a single node.
- scale performance with number of nodes.
- choose which hardware to buy.

A software development team, with experience in supporting scientific communities for over 15 years.

Used to provide roughly 50% of the storage needed to find the Higgs particle

Providing cutting-edge features: pluggable authentication, delegated authorisation, event-driven model / “storage events”, …

dCache authn & authz

Protocol layer

- WebDAV door
- Frontend door
- FTP door
- NFS door

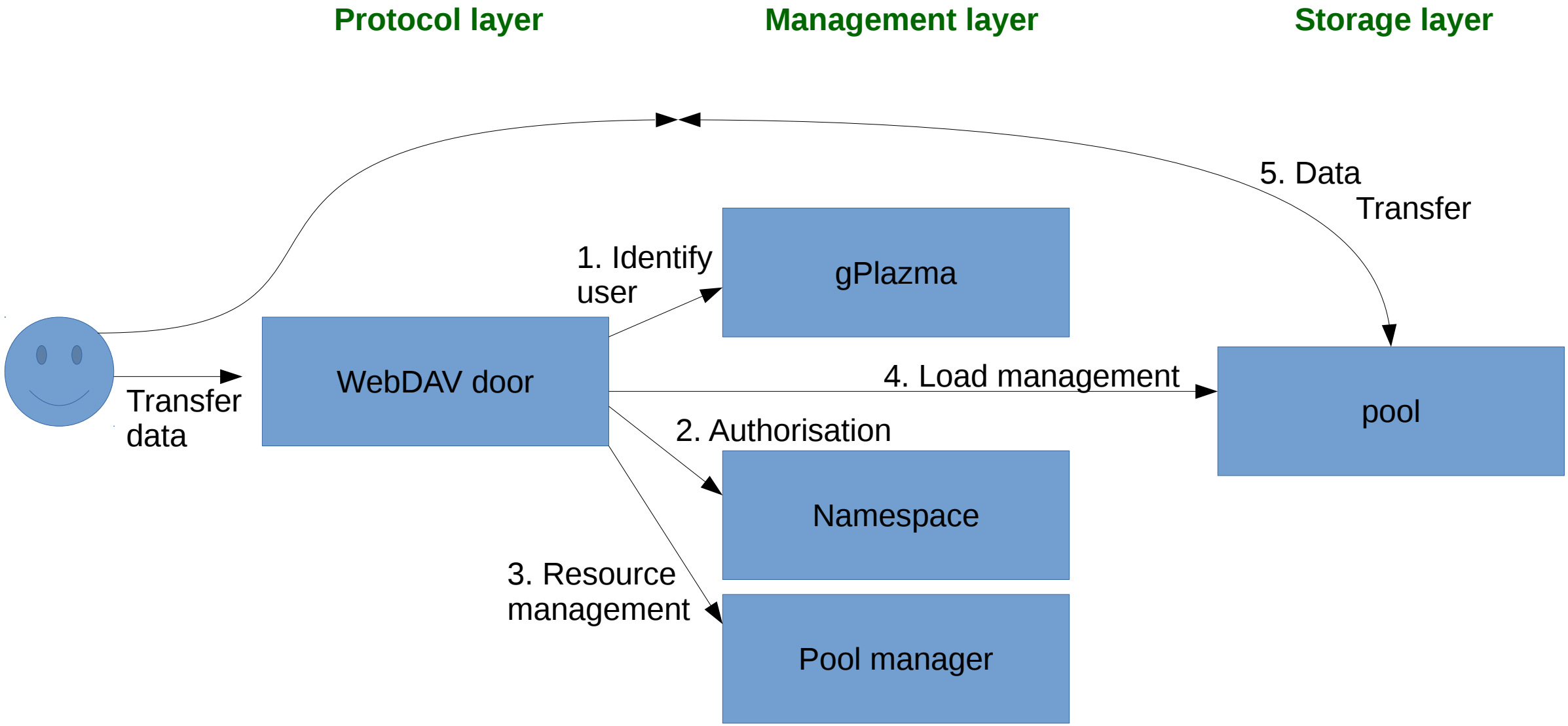
Management layer

- gPlazma
- Namespace
- Pool manager

Storage layer

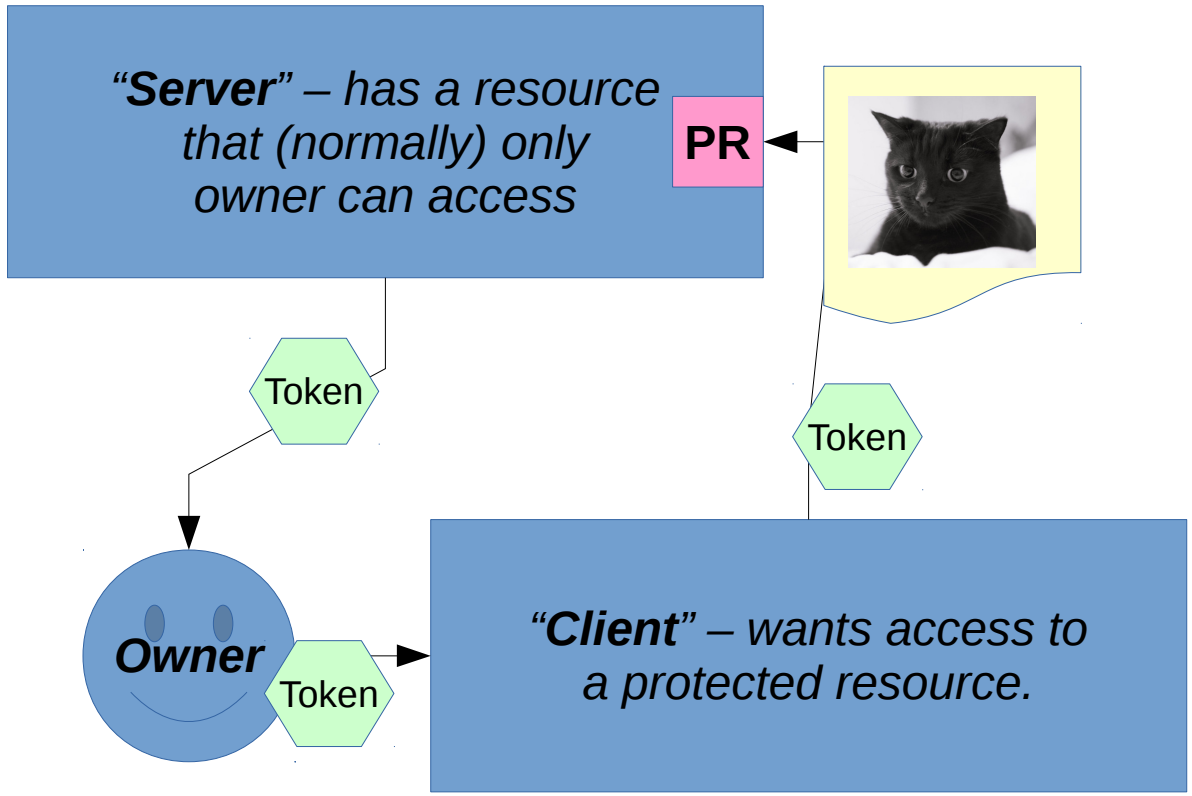
- pool

dCache authn & authz

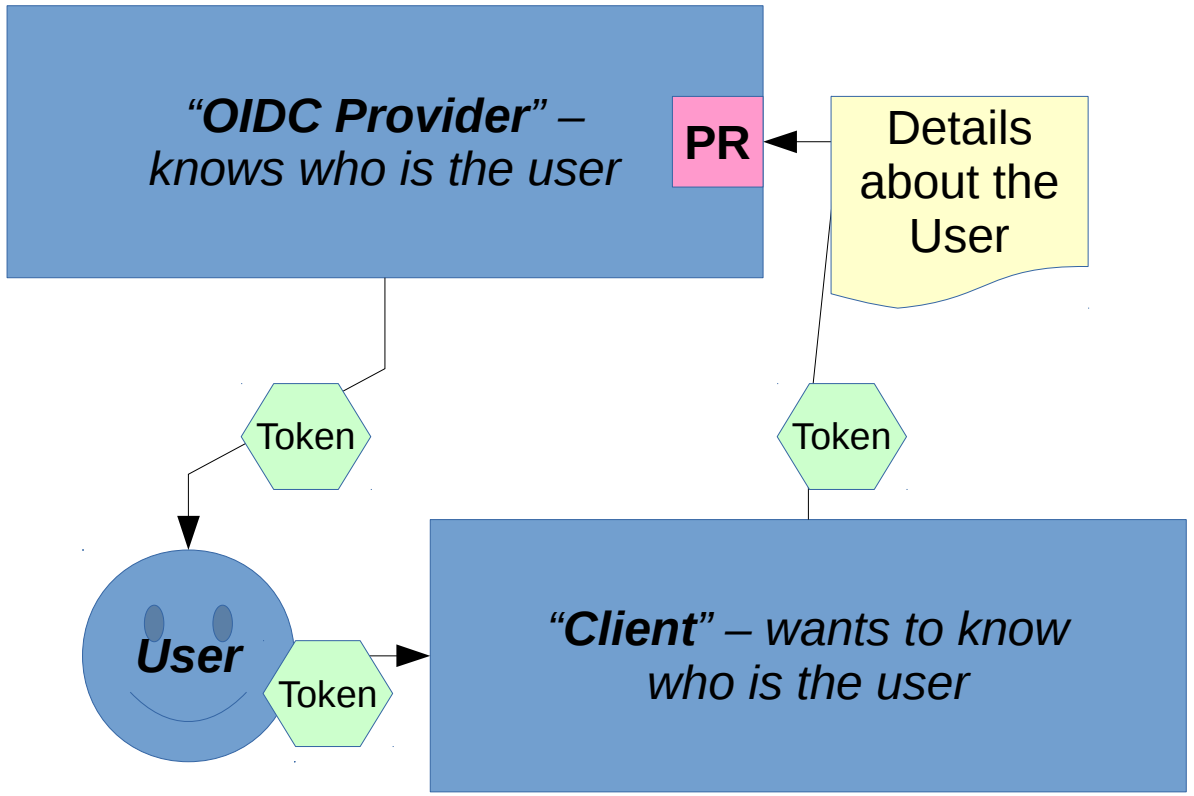


What is OpenID-Connect?

DISCLAIMER: it's more complicated than shown here

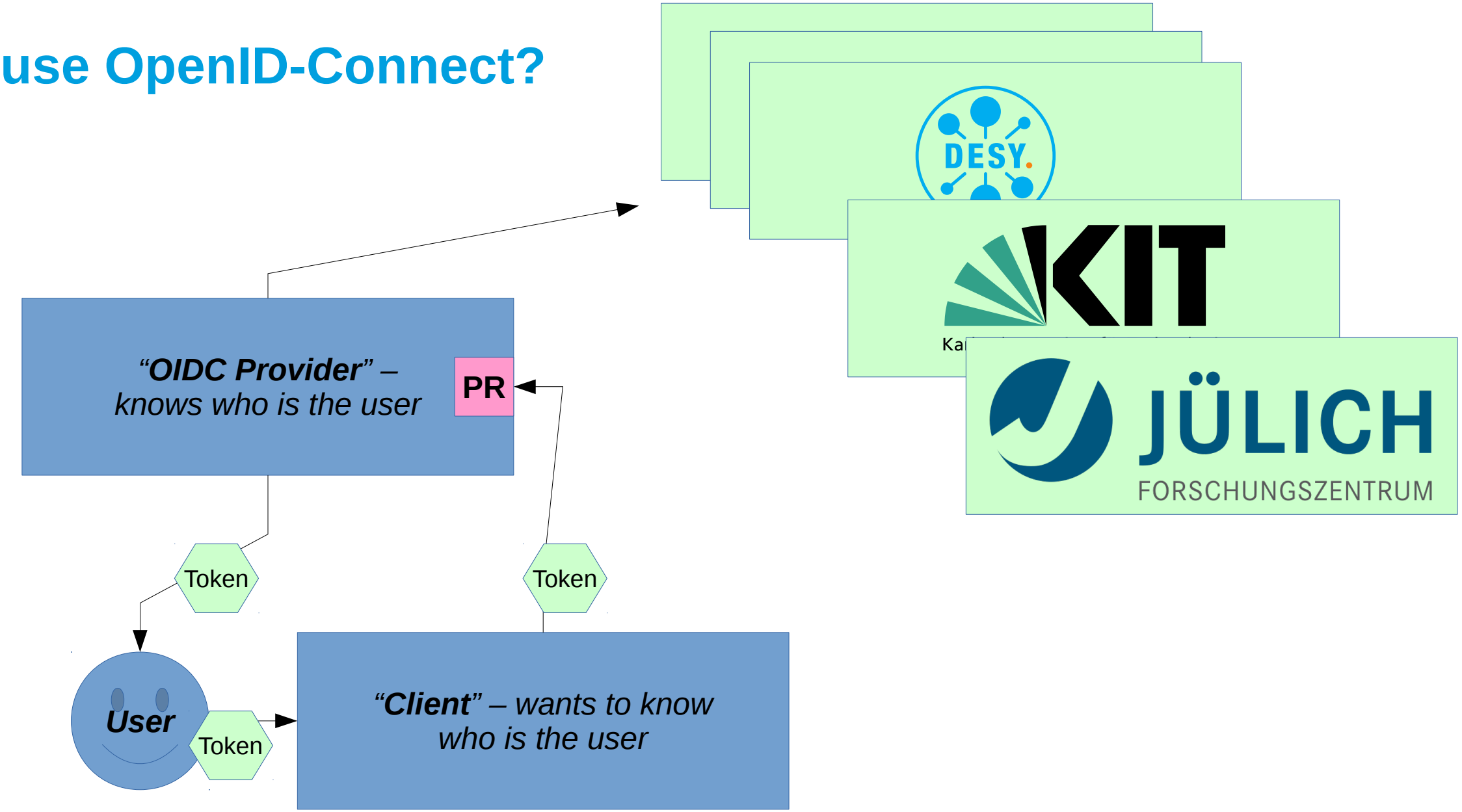


OAuth2

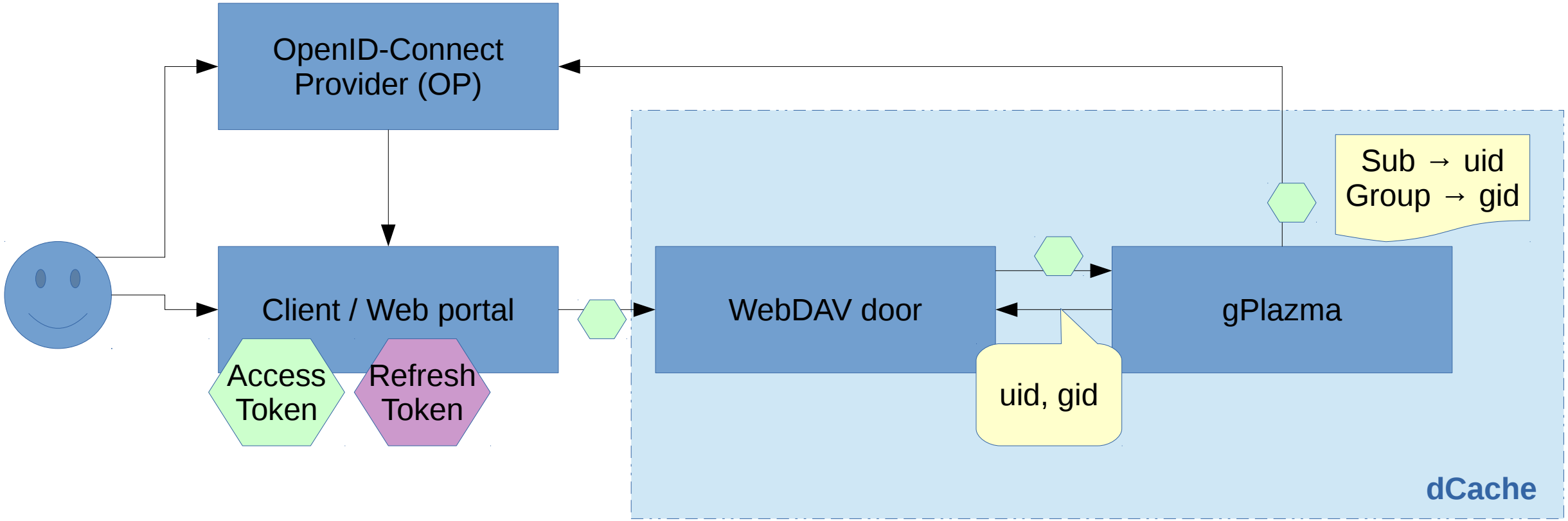


OpenID-Connect

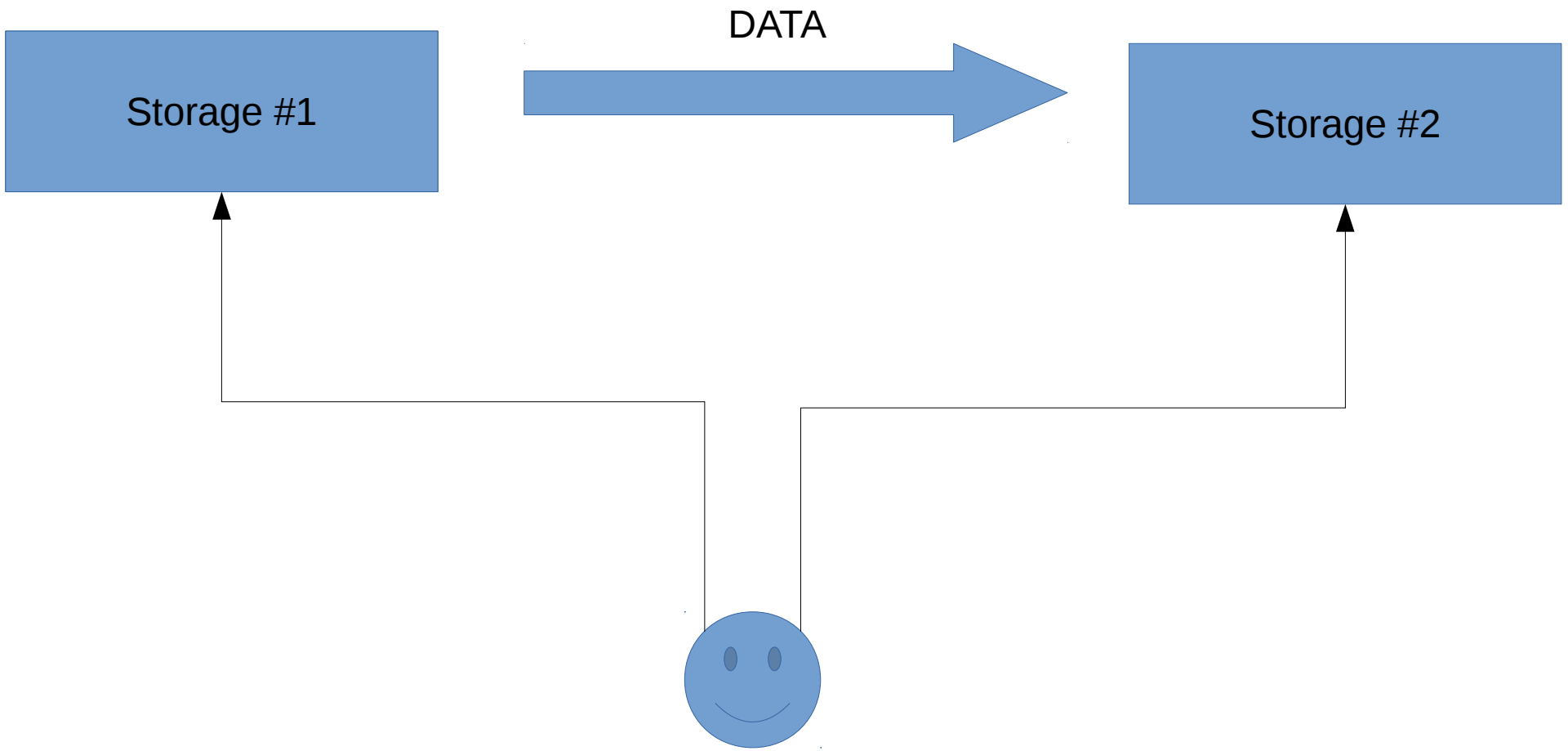
Why use OpenID-Connect?



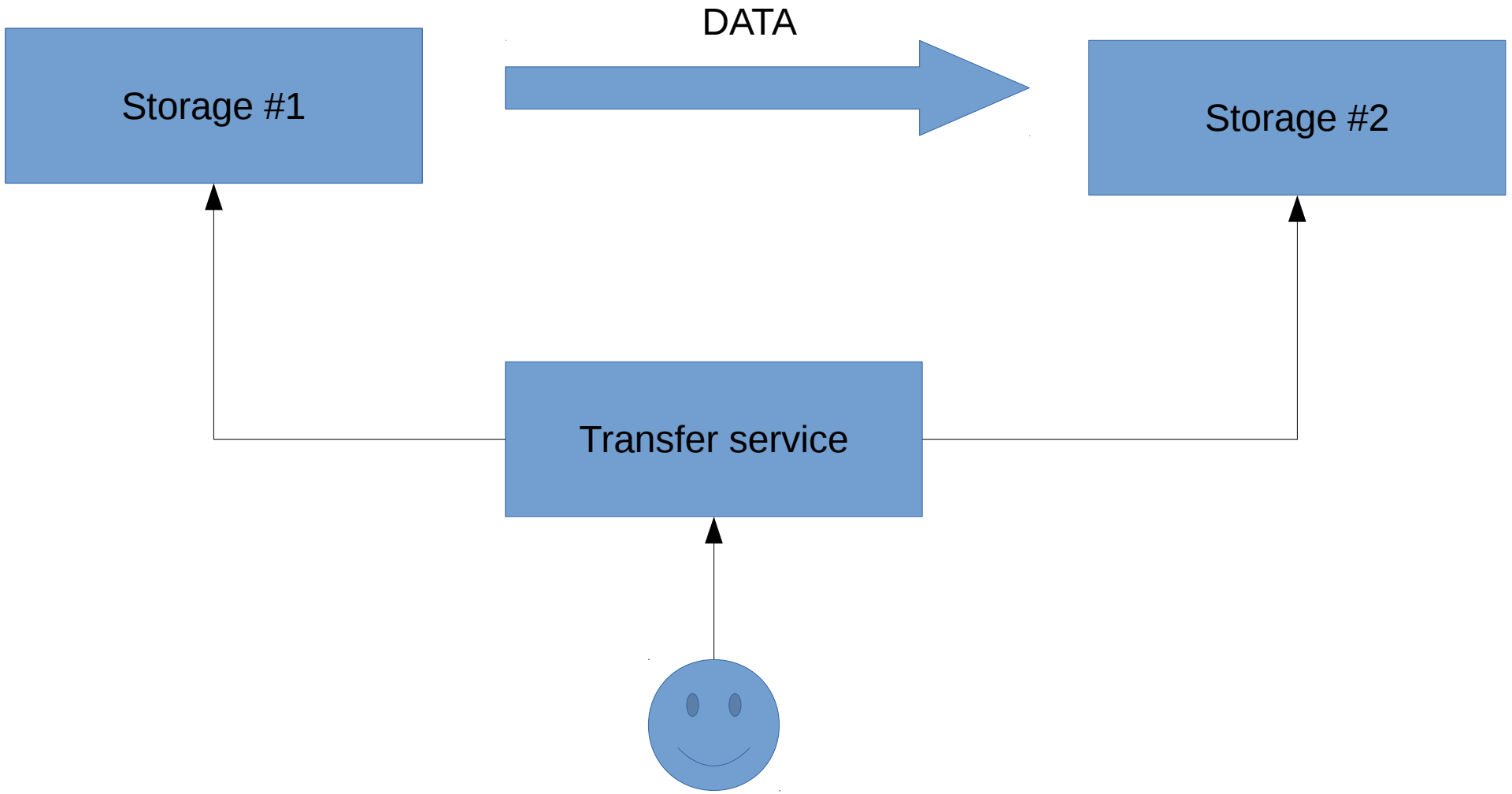
dCache authn & authz: working with OpenID-Connect



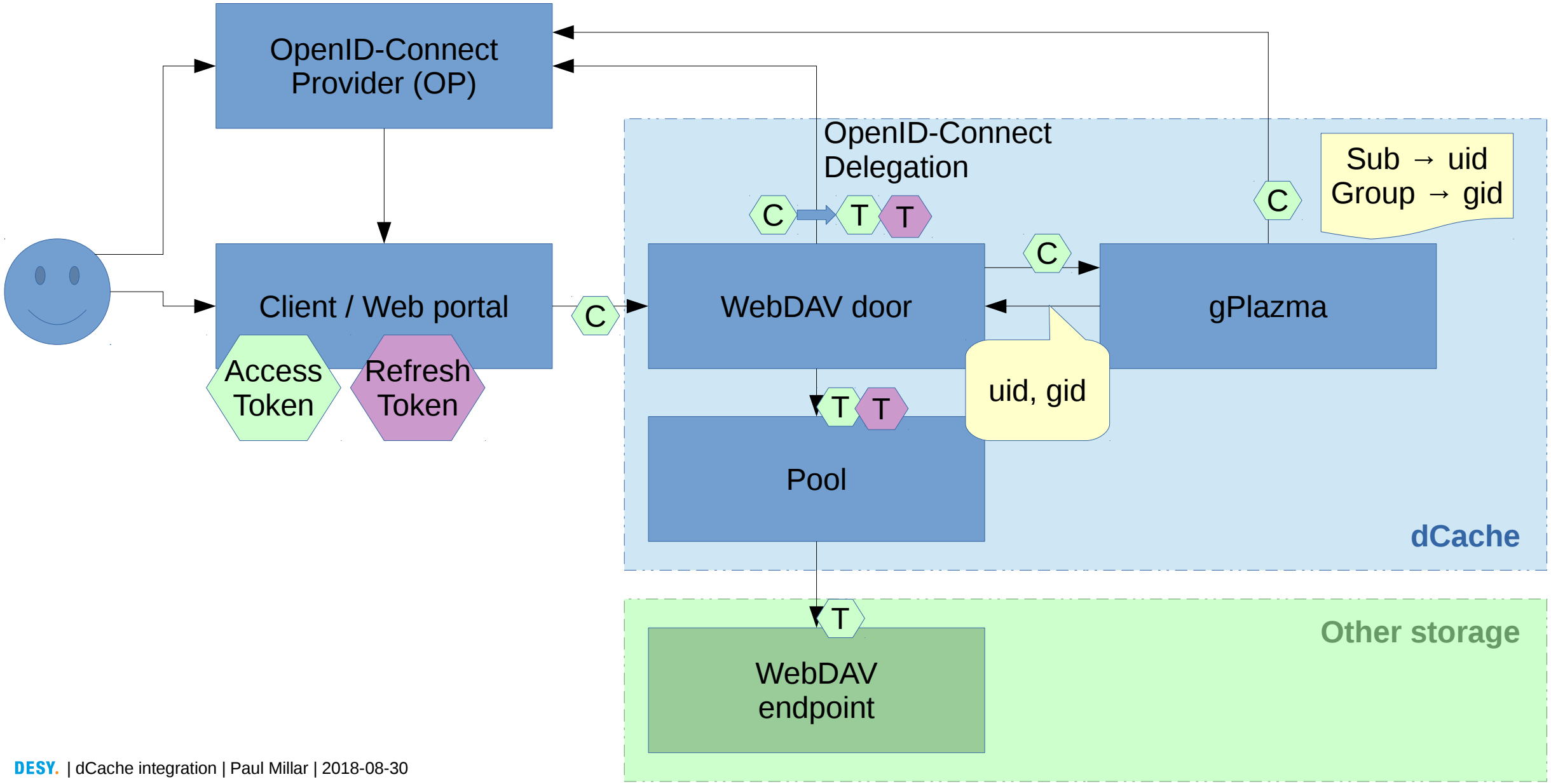
Third-party copy: single file (simplified)



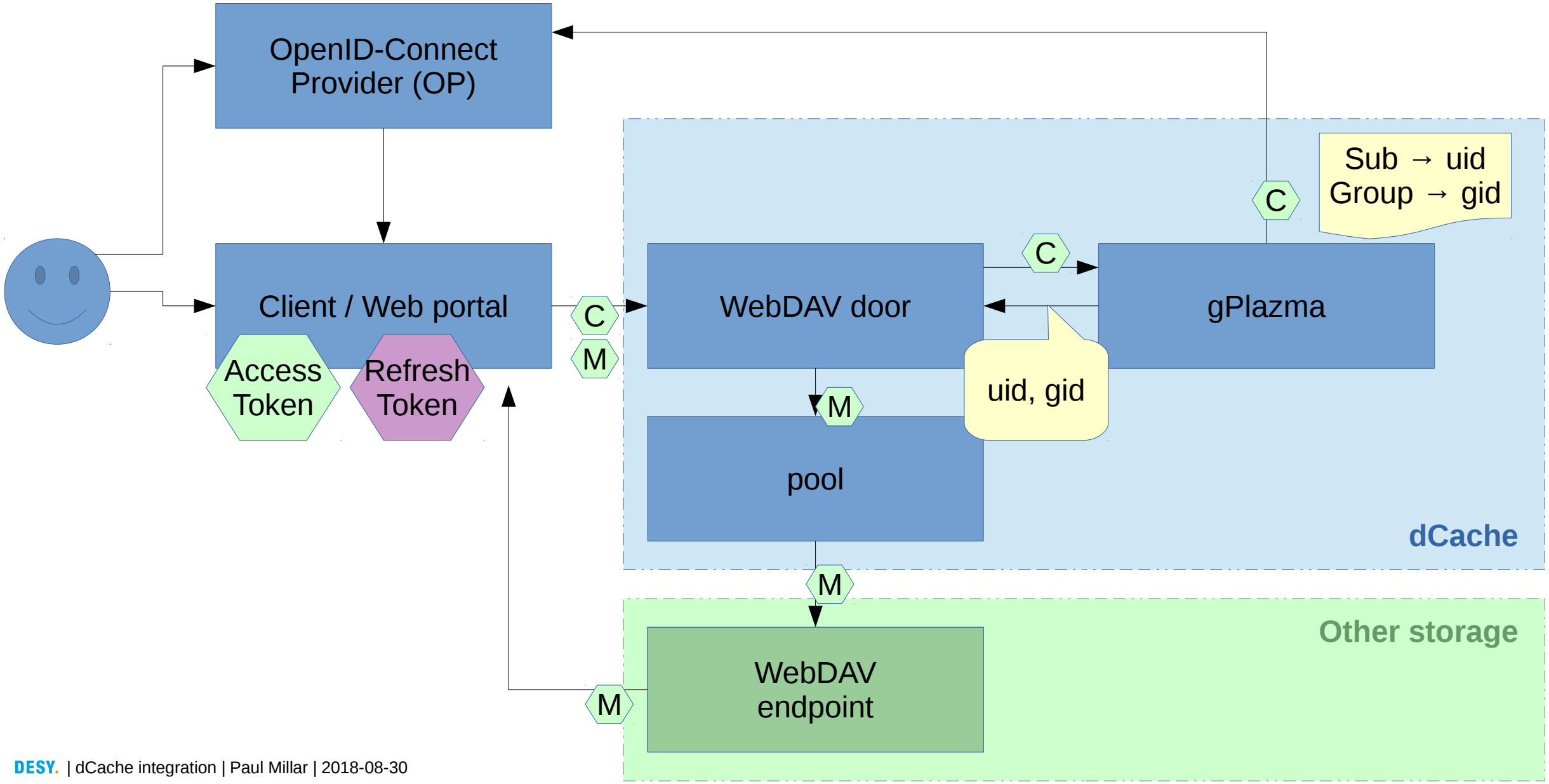
Third-party copy: in practice



dCache authn & authz: third-party copy w/ OIDC delegation



dCache authn & authz: third-party copy with macaroons

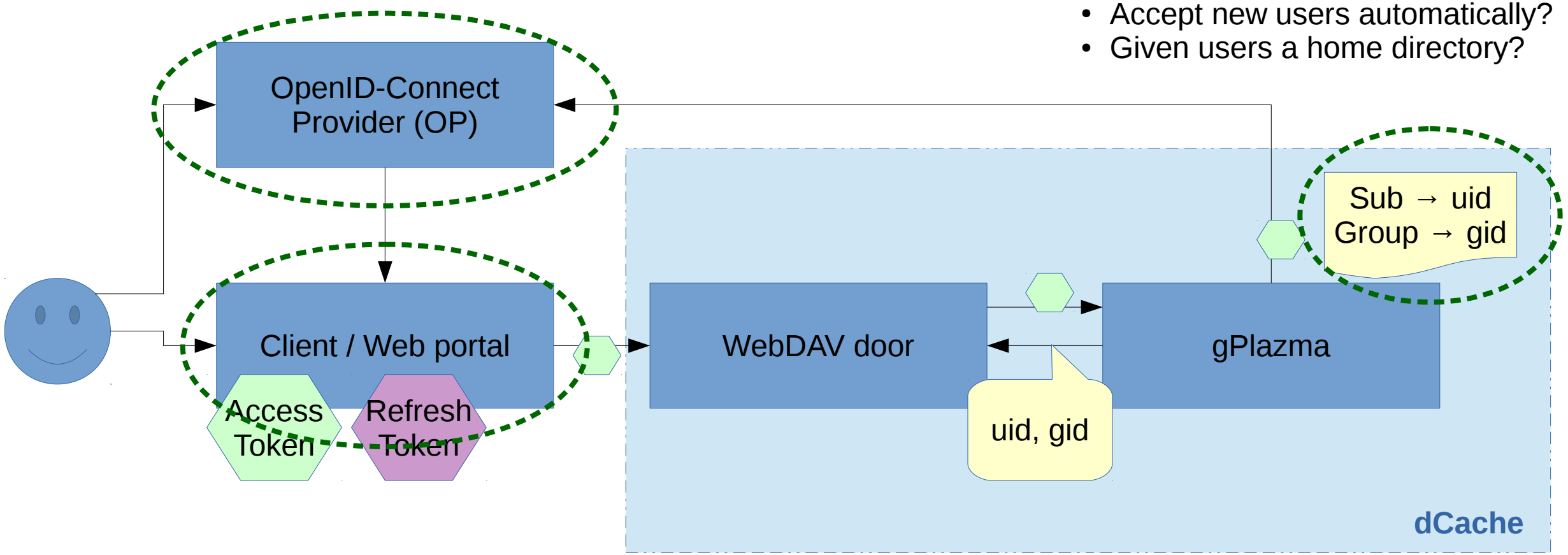


dCache integration points

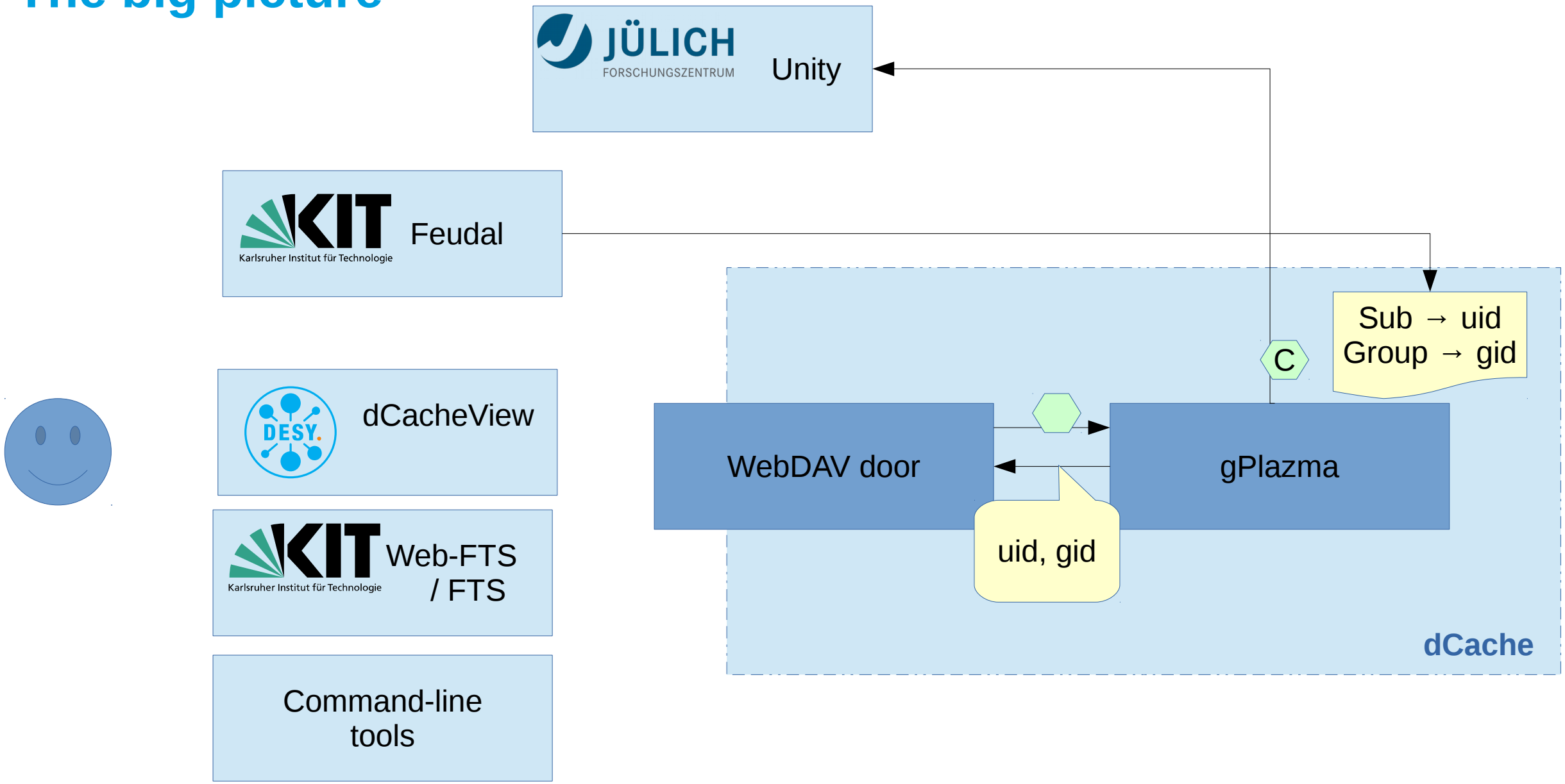
What bits are missing
How these are being added

dCache integration points

Site policy:
- Outside → inside identity
- Accept new users automatically?
- Given users a home directory?



The big picture



Current status

Current status

dCache code

- Support for OIDC in webdav & frontend
- Support for OIDC sub and group mapping
- Third-party transfer support (macaroons and OIDC-delegation)
- OIDC support in dCacheView

Deployment

- Support for Unity in test-bed: prometheus and dcache-xdc
- dCacheView / prometheus support enabled

Initial investigation

- Feudal
- oidc-agent

Future work & time-line

Future work

Development work

- Support dCacheView login longer than the access-token lifetime: either support refresh-token, or add session support.
- Work on command-line clients; work with **KIT** on oidc-agent
- Work with **Jülich** on improvements within Unity; e.g., improve client registration, providing more information, …

Deployment

- Finish enrollment policy at DESY – continue work based on Feudal (collaborating with **KIT**)
- Update a production dCache instance at DESY to support HDF
  - November 2018 – Closed beta testing (selected people)
  - Early 2019 – Open beta testing: users should accept possible unannounced interruption to service.

Possible usage scenarios: leverage dCache advantages

Data for computation work

- Move data to DESY to allow analysis work with low latency access to data.

Sharing data within HDF

- Enable easy sharing data between institutes.
- Support multiple access protocols, with all of dCache features.

Sharing data with people outside of HDF

- Allow others to upload data; e.g., volunteer-computing / BOINC, IoT, …
- Allow non-HDF users access to data without making it public

Make use of other advanced dCache features

- Multiple protocol support
- Event driven work-flows / catalogue synchronisation / …
- Third-party transfer support
- Different Quality-of-Service

Thank you

Contact

DESY. Deutsches
Elektronen-Synchrotron

www.desy.de

Paul Millar
IT
paul.millar@desy.de
+49 40 8998 5378