Sunet Drive

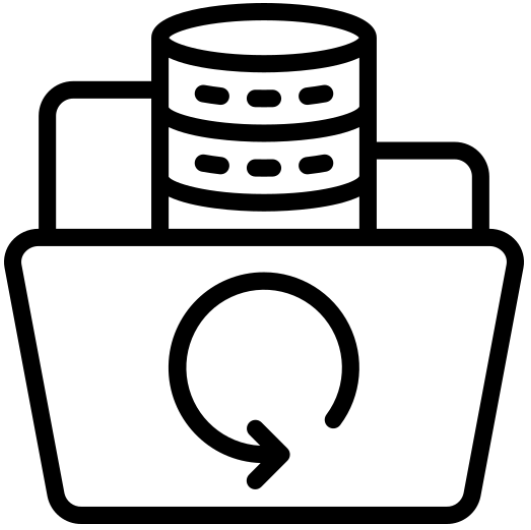
SUNET

Sunet Drive – Under the hood

1) Backups
2)Architecture
3)Login
4)Global Scale
5)Global Site Selector
6)Lookup
7)Login flow
8)MFA in Nextcloud
9)Secure zones
10) Publication with RDS

SUNET

Backups



Sunet Drive now has backups, using duplicity*, built in:

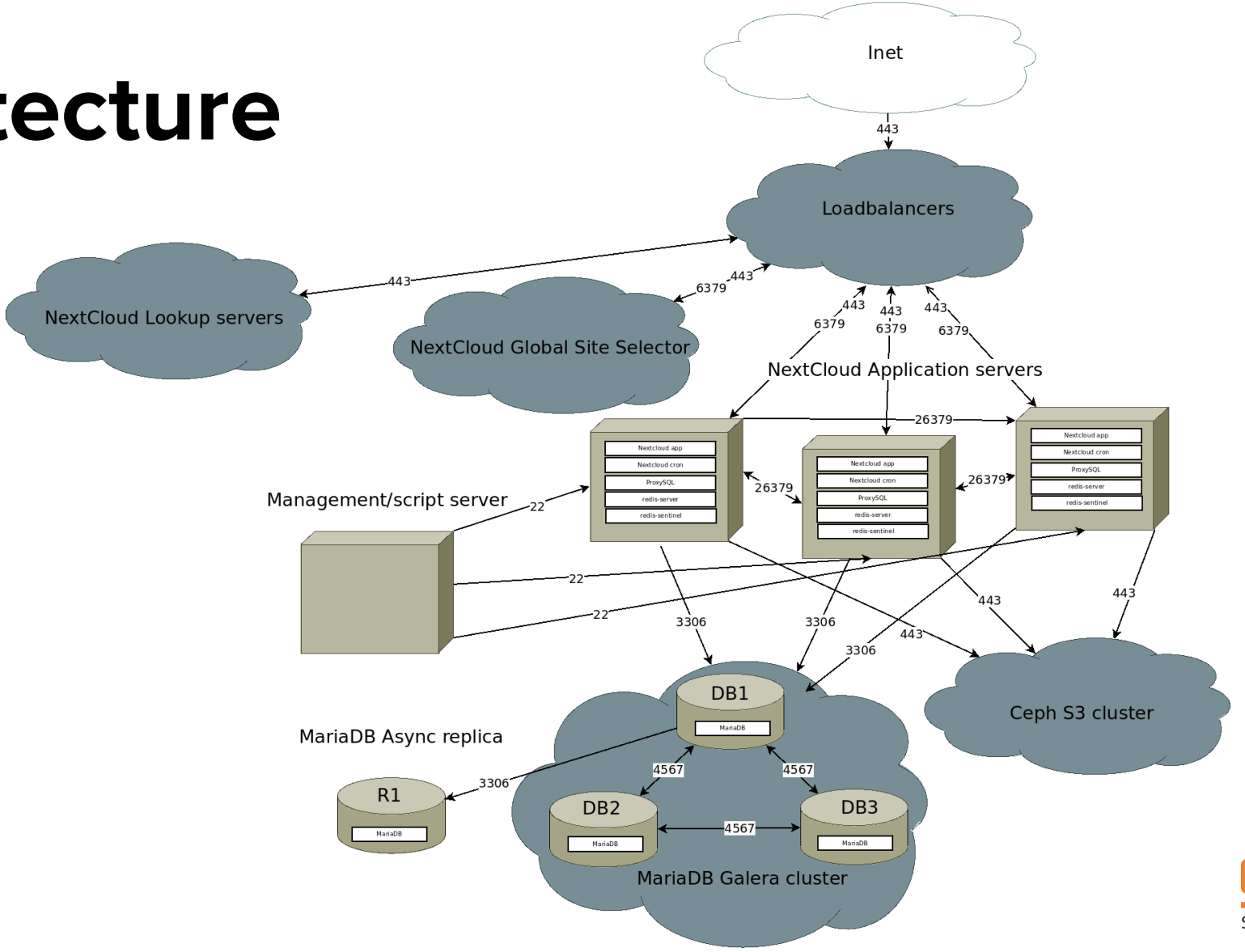
- Used for disaster recovery by us, with one month retention
- Full backups once a month, incremental every night
- Full backup of buckets with more than ~2-3TB will take more than a day
- Retention time can be configured per customer on request
- No added fee**
- Point in time recovery of single files to whole buckets can be done
- However, you will have to implement restore procedures on your end if you want to advertise it to users (i.e. do restores on a regular basis).

* https://duplicity.gitlab.io/

** Except for disk usage. Expect about 105% of bucket size for full backup, and ~5-10% per incremental, depending on how much content changes.

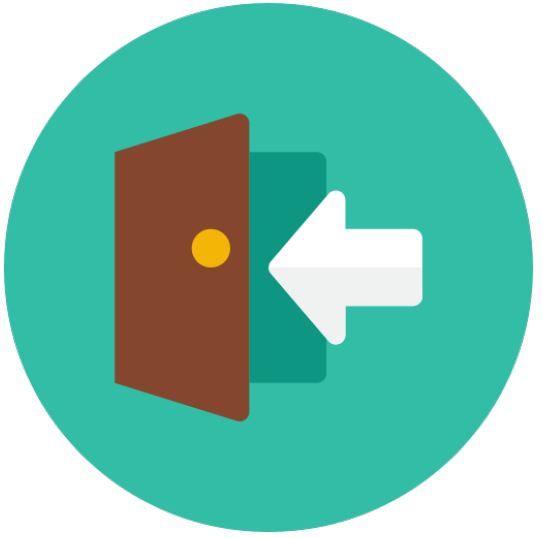
SUNET

Architecture



SUNET

Login



Nextcloud (but not Sunet Drive!) has a lot of different options for logging in users:

- Local database with username and password (default)
- LDAP integration
- SAML
- OpenId Connect
- Social login (Github, Twitter, Facebook, etc)
- And more!

We are using the so called Global Site Selector for user login, which is part of Nextclouds Global Scale solution.

SUNET

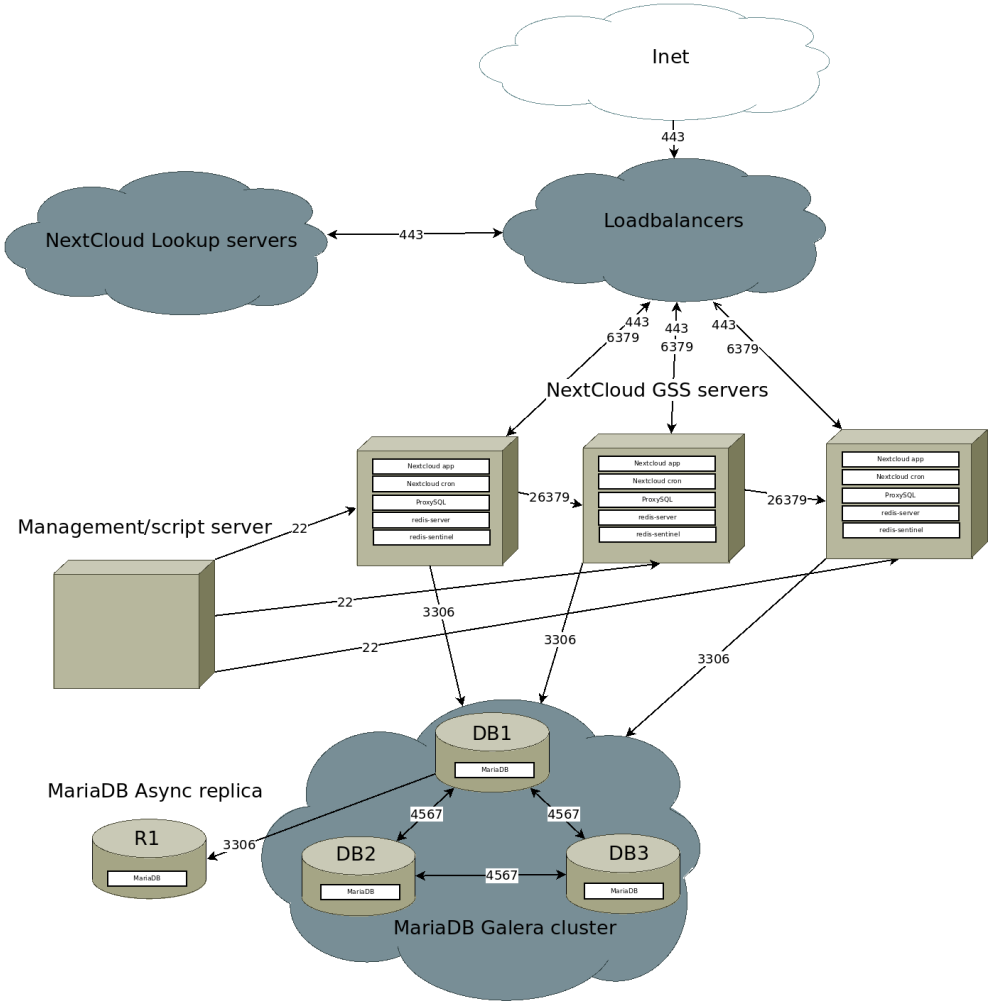
Global Scale

Global Scale is the Nextcloud term for our architecture. Sunet Drive consists of 54 distinct Nextcloud instances, three global site selector servers and two lookup servers GSS is responsible for logging in a user and redirecting the user to the correct instance.

Lookup server is responsible for keeping track of which instance a user belongs to. It is used when searching for users to sh documents with.

SUNET

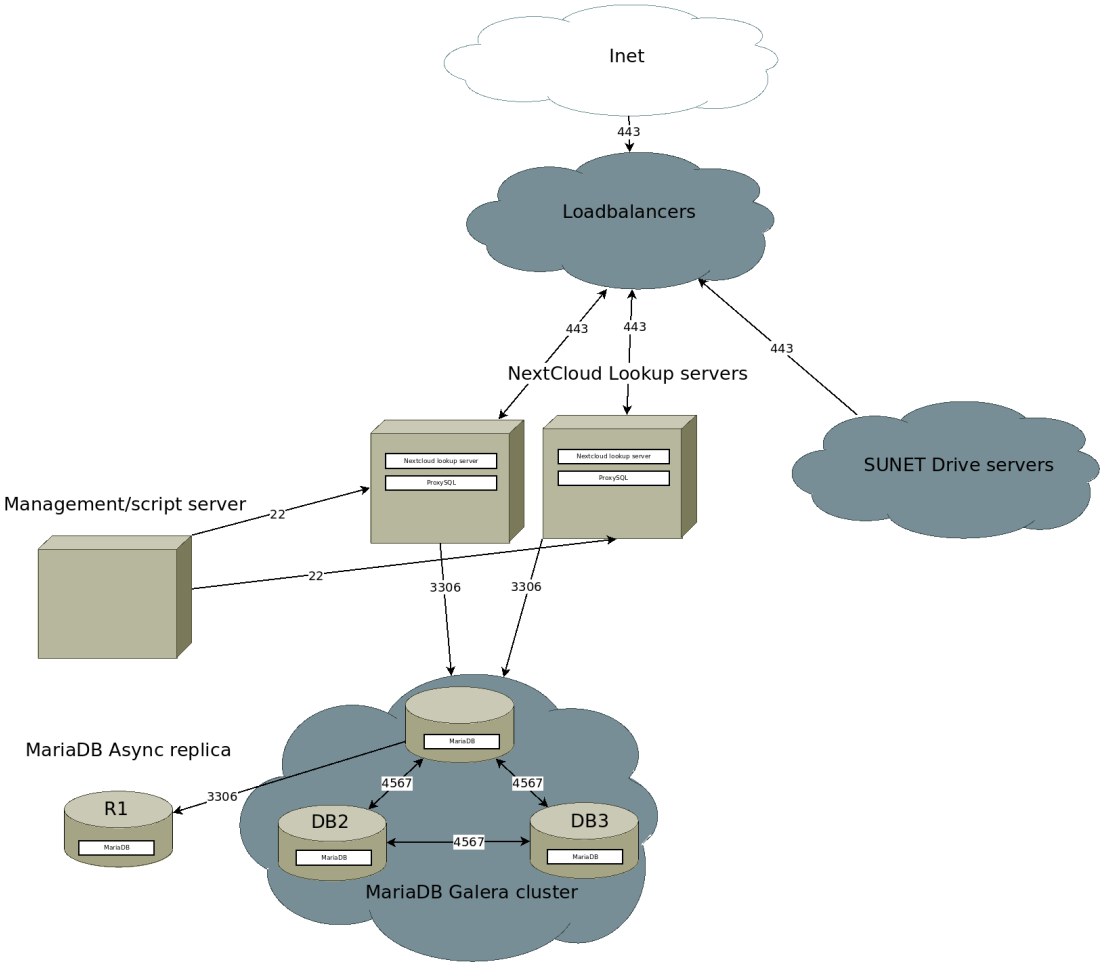
Global Site Selector



GSS servers are very similar to ordinary Nextcloud nodes, the difference is that they are running the GSS app in special mode. They also have a mapping.json file which uses regexes to match a users domain with a Nextcloud node.

SUNET

Lookup



Lookup servers are very simple and has a flat database with key-values which store email, uid and display-name of a user across nodes.

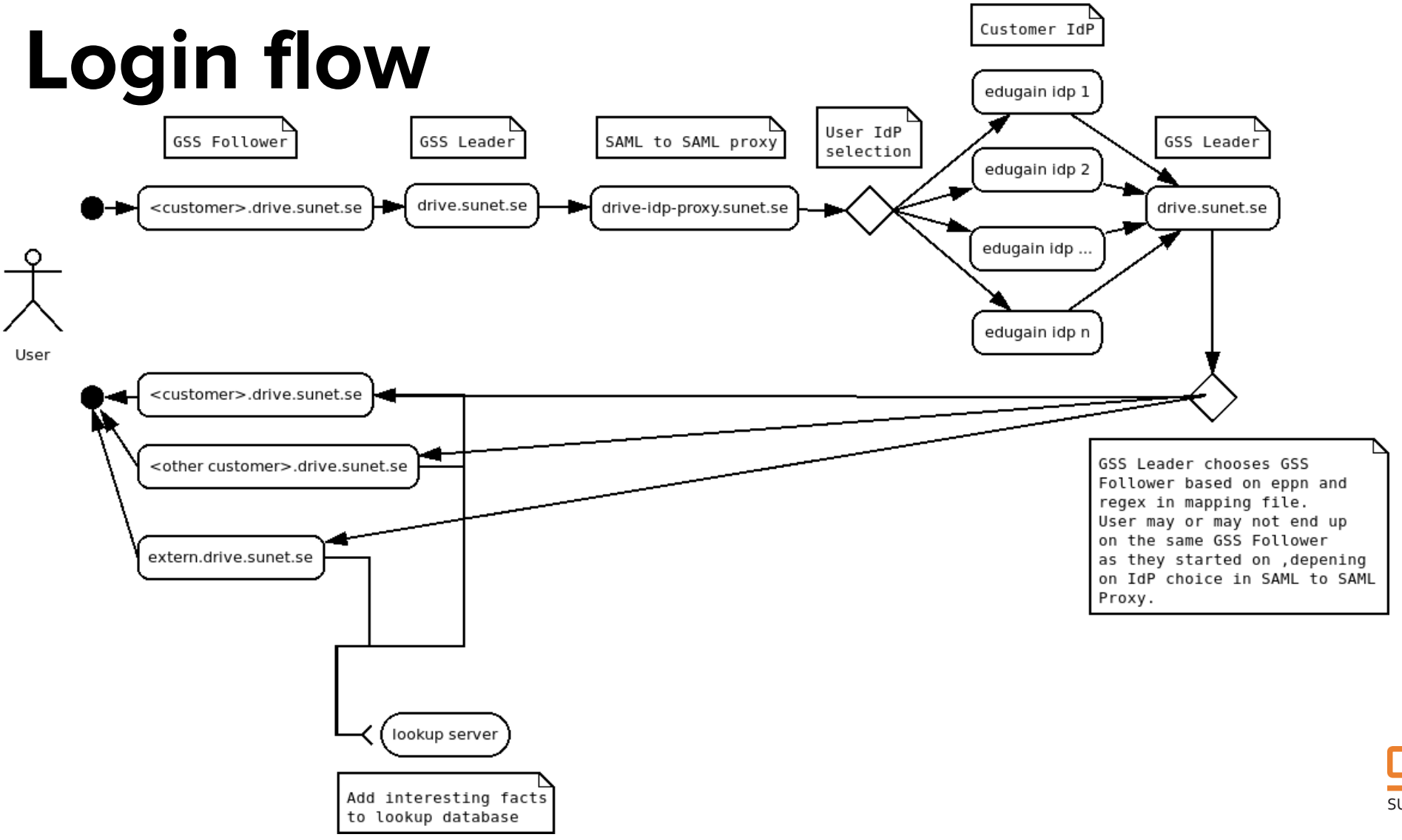
SUNET

Login flow

The login flow with the Global Site Selector is different from all other login flows! The user is, in our case, actually logged in on the Global Site Selector node with SAML, and is then given a valid web token. As a result, the user appears to Nextcloud as an already logged in user. That means that the users node does not trigger any actions for login on it’s own.

SUNET

Login flow



SUNET

MFA in Nextcloud

In Nextcloud MFA is of course a part of the login flow. You can configure Nextcloud with a variety of MFA-providers such as TOTP, Webauthn, single use codes etc.

You can enforce MFA for all users, per groups or per user. The user can also opt in to MFA on their own if enabled.

Enabling MFA does not do anything else, except requiring the user to use a second factor when logging in. It is good for making sure a user is who they say they are, but it does not give the user any extra capabilities.

SUNET

MFA in Nextcloud

If MFA already exists in Nextcloud, why can’t we use it?

Well you can (and some customers do), but it will only work for local logins.

We can not, as of yet, use the Global Site Selector and enforce MFA at the same time.

But you can! You can implement and enforce MFA on the IdP-side, and it will work out of the box with Sunet Drive.

This is not what you, the customer, want to hear though. Because what you want is in fact to have Nextcloud be aware of if the user logged in with MFA or not, and give the user extra capabilities in that case.

SUNET

Secure zones

One primary use case for MFA is to protect certain data, and make sure only appointed users get access.

One primary use case for Sunet Drive is to share data easily between researchers across the globe.

This means that there can be a conflict between these important goals.

Because of this we are developing secure zones for Nextcloud, in partnership with Pondersource whom we are paying to develop secure zones for Nextcloud. The end goal is to have certain areas in Nextcloud marked as “secure” and have them only be accessible via the web interface for MFA-verified users. This is not to stop malicious users who have been given access to such data, but to help users avoid high risk misstakes. After all, if a user can access data, they can of course copy that data, with a pen and paper if necessary.

SUNET

Secure zones

This is a project we are doing in seven steps. First three steps have been delivered, and the fourth is slated for delivery by the end of October. The fourth step of the project is to have a Nextcloud app that can see if the user is logged in via MFA, no matter how the MFA was provided (localy for local users and via SAML attribute passed on to GSS Follower).

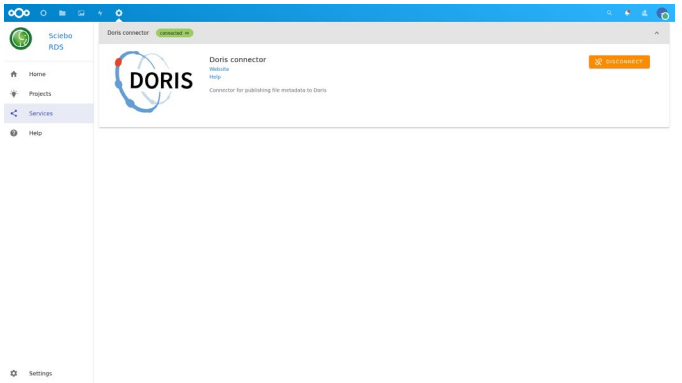
The following steps involve making sure Nextcloud hides certain areas from sync client and users not logged in via MFA.

The final step is an external security audit of the code (which will be made available to the public under the AGPL).

SUNET

Publication with RDS

- Publication of data
  - Under development
  - Will be possible to publish directly from Sunet Drive
  - Zenodo, SND, Harvard Dataverse
- Sciebo-RDS
  - Interoperability layer between EFSS and publication
  - Connectors as microservices to integrate a Data Service
  - https://www.research-data-services.org
- Research Object Crate (RO-Crate)
  - Lightweight approach to packaging research data
  - Based on JSON-LD
  - Focus on workflow-driven analysis
  - Digital repository management
  - FAIRify data

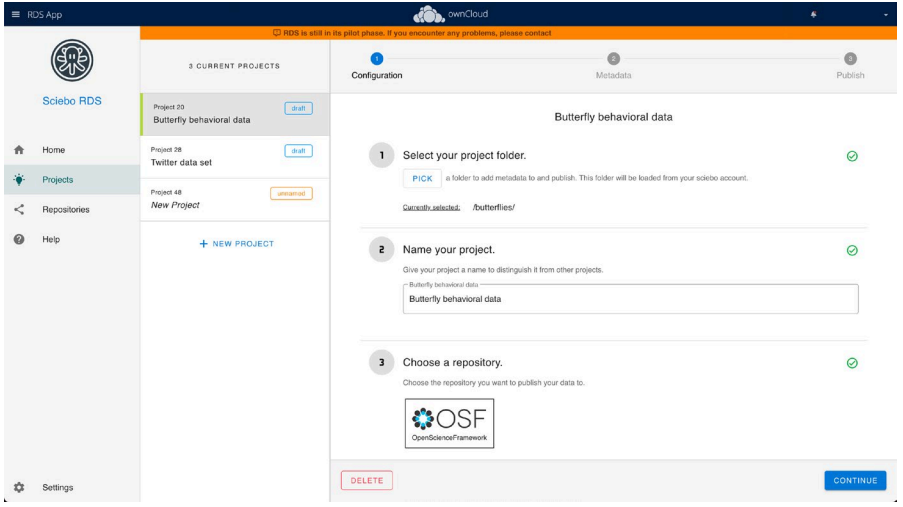


SUNET

Publication with RDS

– Next steps

- Sunet RDS Infrastruture
  - Scale out deployment
  - Map RDS instances with Sunet Drive nodes
- RDS for Nextcloud
  - Improve Nextclud-native RDS-app
  - Deploy RDS to Sunet Drive Test
- Doris SND Connector
  - Define architecture and design in context with Doris and Sunet Drive
  - Specify storage integration architecture
  - Reference users
  - Workflows for sensitive/insensitive data



Report: https://forum.sunet.se/s/sunet-drive/cfiles/browse

SUNET

Sunet Drive

drive@sunet.se

Backup icons created by juicy_fish - Flaticon

SUNET