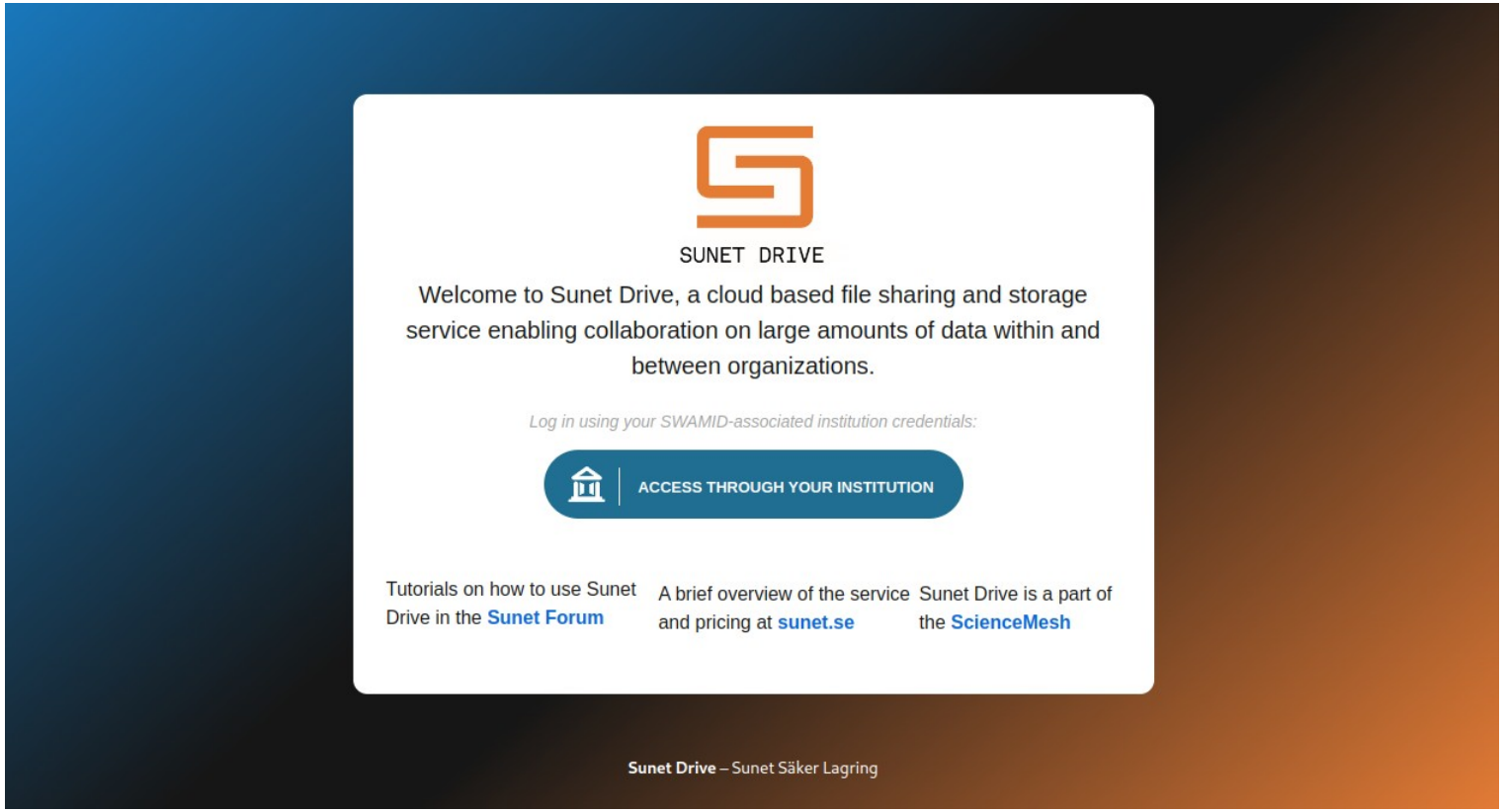
Sunet Drive

An academic EFSS packaged for EOSC



CS3 – 2024 – Community Site Report

Anecdote

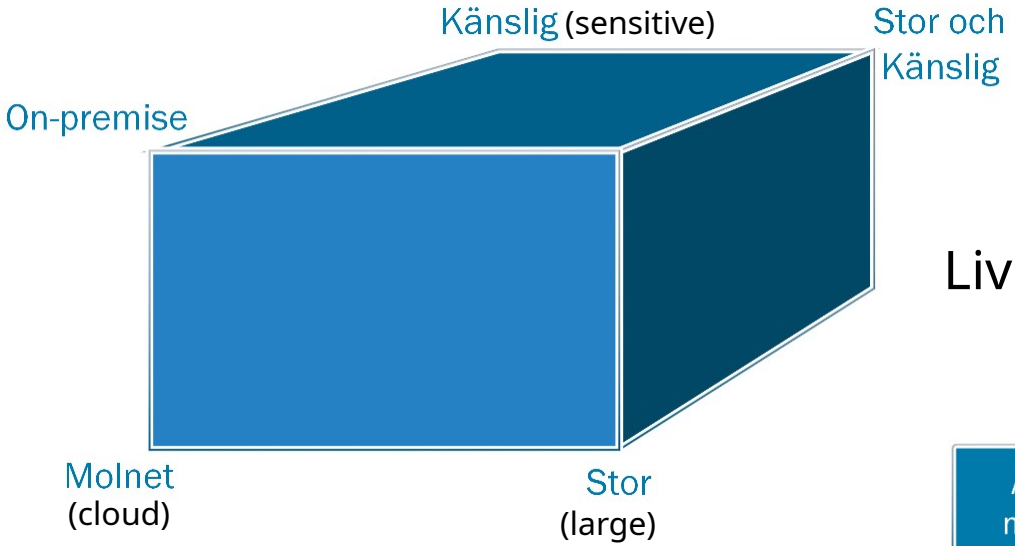
Thanks for the detaljer information! I will be ending my employment this Saturday, so am looking for some solution where the data will not be deleted due to the deactivation of my akka id and where it would be accessible by my PI in the future.

Wednesday, 10:51

SUNET

Research Data – A model

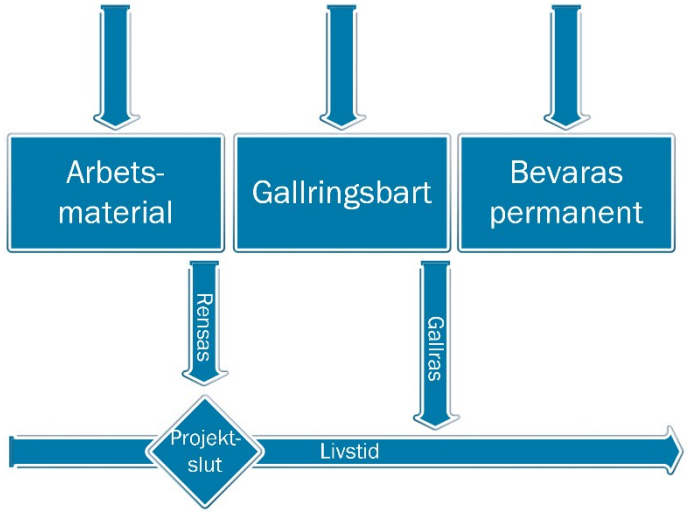
Classification



How do you access data?

Where is my data?

Livecycle



- Working material
- Data retention
- Archival

Access

”Klassisk” Mappar

”Modern” Appar

SUNET

Research Data – A model

Classification

Känslig

Stor och Känslig

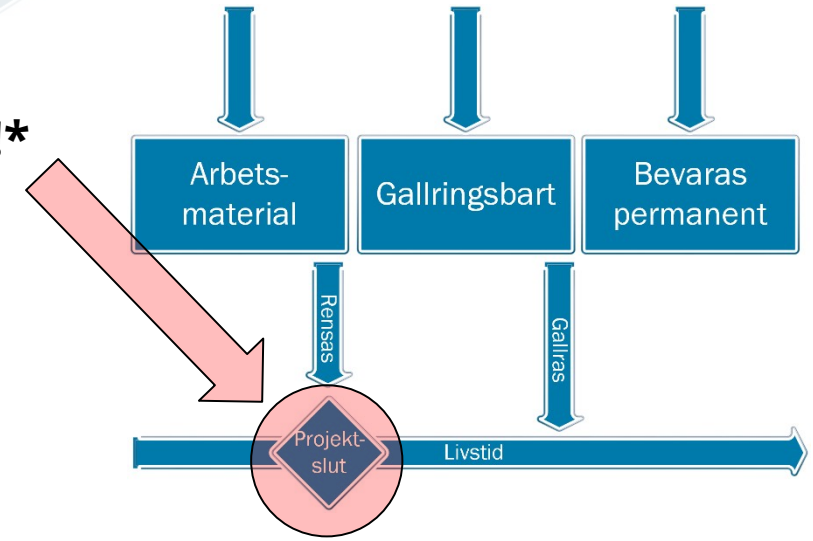
On-premise

How do you access data?

Where is my data?

Livecycle

This is where the money ends!*



Molnet

Stor

Access

”Klassisk” Mappar

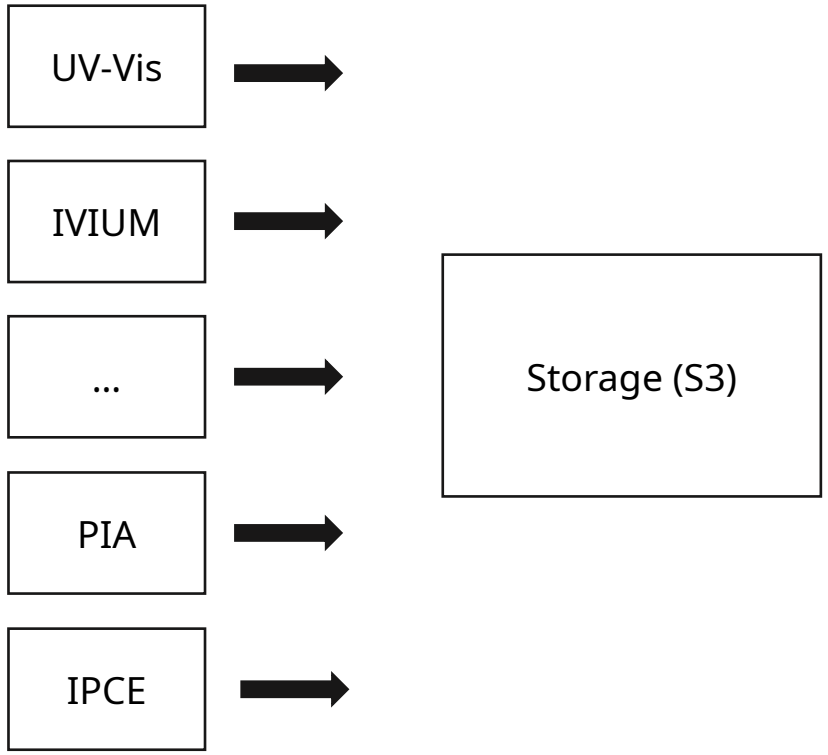
”Modern” Appar

*for many projects

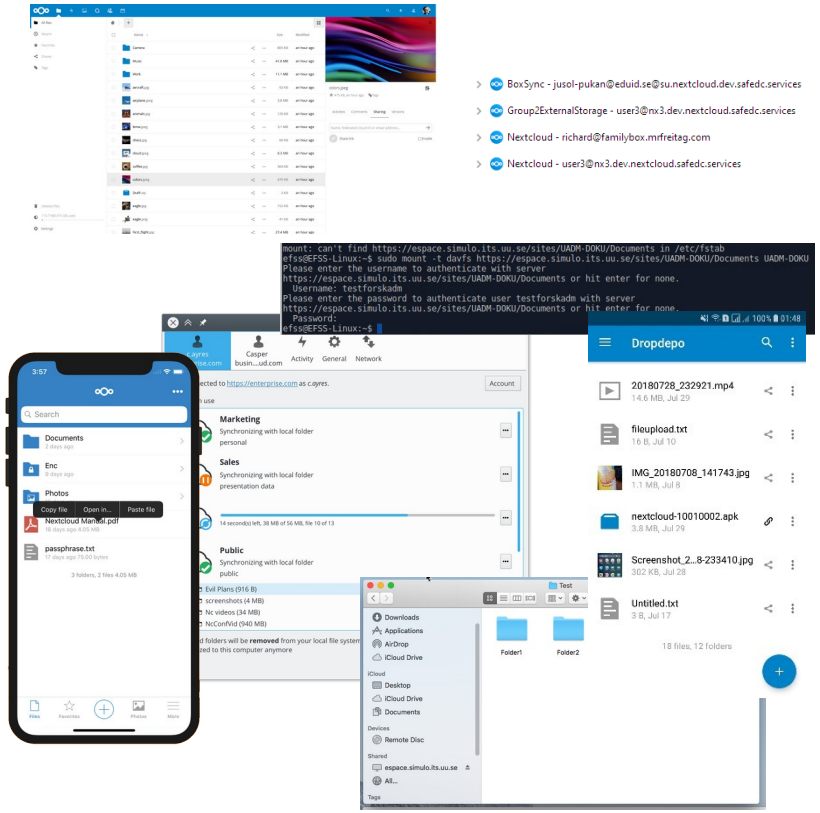
SUNET

How it started

Primary data sources



Access



SUNET

How it is going

Primary data sources

UV-Vis

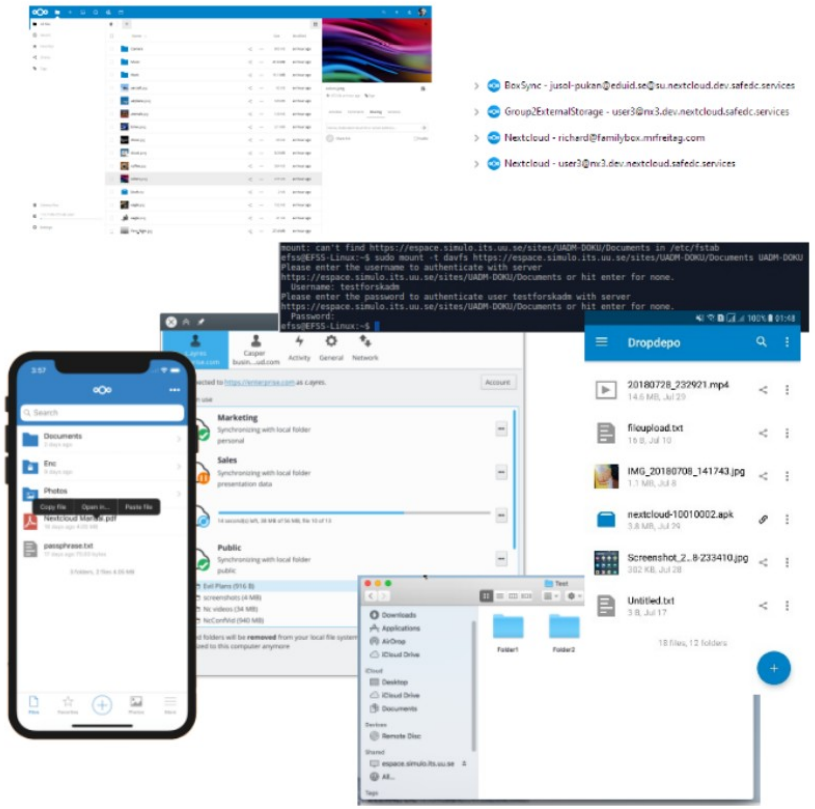
IVIUM

...

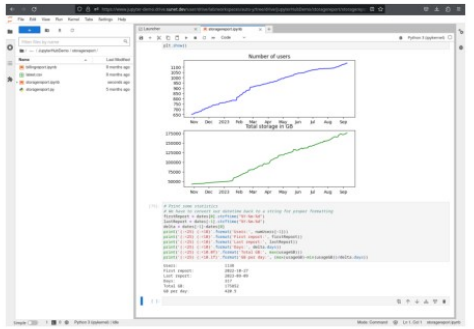
PIA

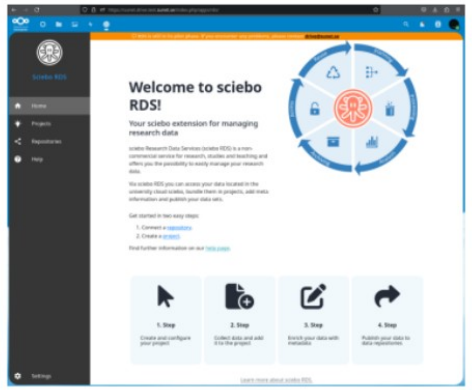
IPCE

Secure Storage



Analyze and Publish





SUNET

Co-management of nodes

The main storage unit is an S3-bucket

Sunet Drive is a cloud-based file storage and collaboration service that utilises the S3 (Simple Storage Service) infrastructure as its main storage unit. S3 is a highly scalable, reliable, and secure object storage service that can store and retrieve any amount of data from anywhere on the web.

Every university has its own Sunet Drive node

Sunet Drive is a service provided by the Swedish University Network (SUNET), which is a collaboration between Swedish universities and colleges. Each university has its own Sunet Drive node, which provides a secure and dedicated file storage environment for its faculty, staff, and students.

Data can be shared between nodes

One of the benefits of the Sunet Drive service is that data can be easily shared between nodes, enabling collaboration and knowledge-sharing between different universities and departments. This is achieved through the use of access control lists (ACLs), which allow node administrators to grant or revoke permissions to specific users or groups.

Separation of access and ownership

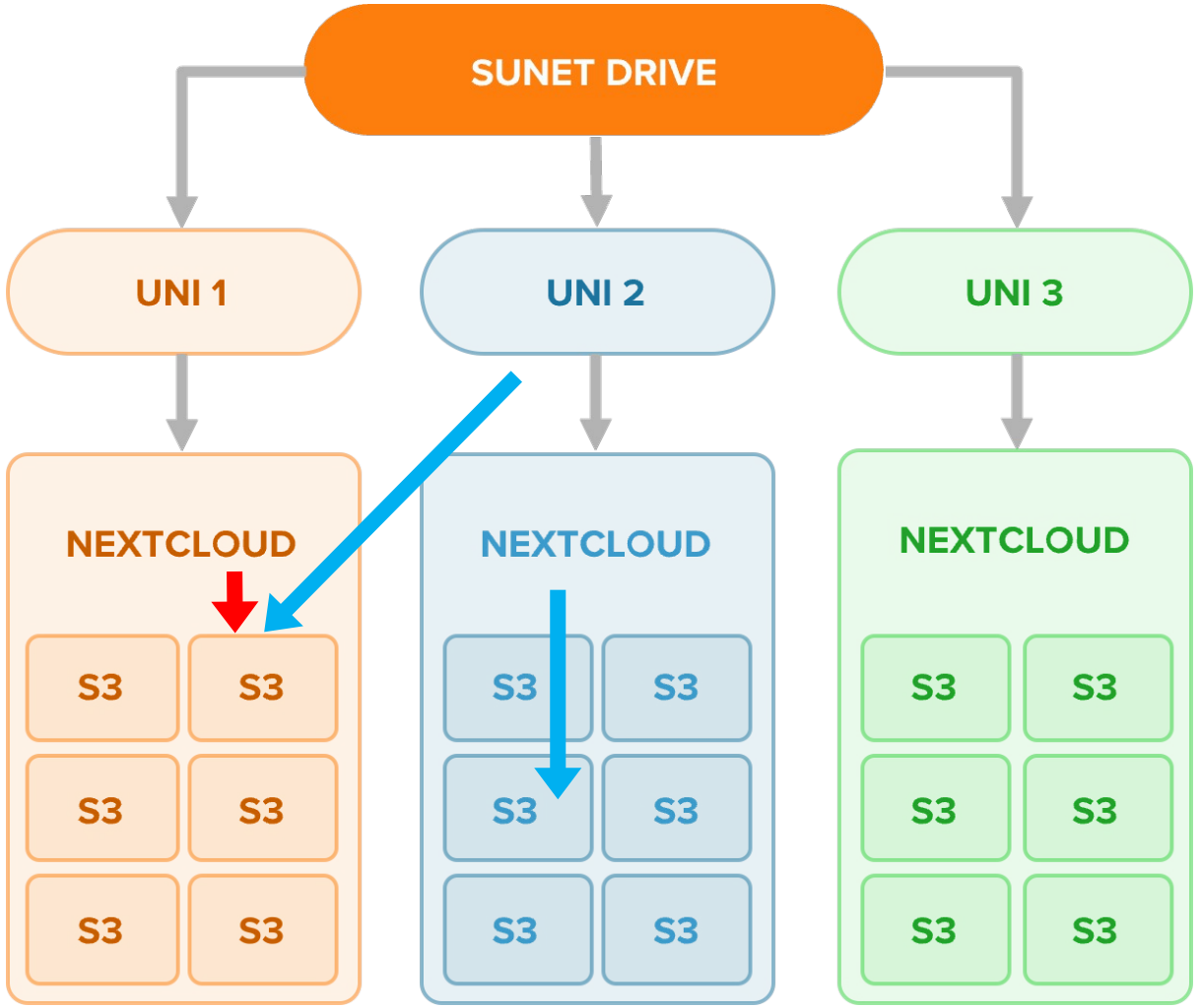
Another important feature of Sunet Drive is the separation of access and ownership. This means that while data can be shared between nodes, the ownership of the data remains with the original uploader. Additionally, access permissions can be finely-grained, allowing administrators to control who can view, edit, or delete files and folders on a per-user or per-group basis. This helps ensure the security and privacy of sensitive data.

Buckets can be lifecycle managed

In S3, buckets (a container for objects stored in S3) can be lifecycle-managed using rules to transition objects to different storage classes automatically, archive them to Glacier for long-term retention, or even delete them when they are no longer needed. Sunet Drive uses this feature to help manage data storage costs and optimise data retention policies.

Co-management of nodes

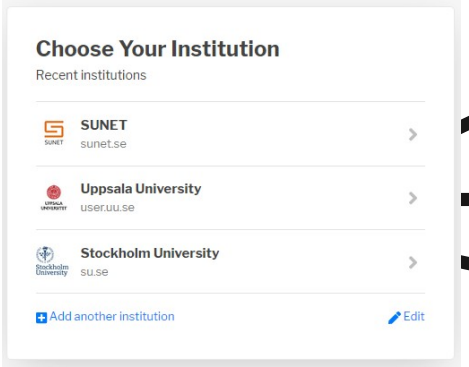
- S3-bucket storage
- Share data between nodes
- Access ownership separation
- Lifecycle managed buckets



SUNET

Identity management: SWAMID/eduGAIN

Swedish Account

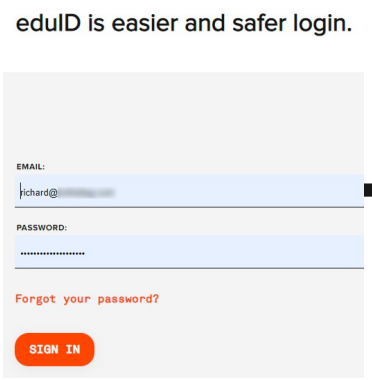


sunet.drive.sunet.se

uu.drive.sunet.se

su.drive.sunet.se

External Account



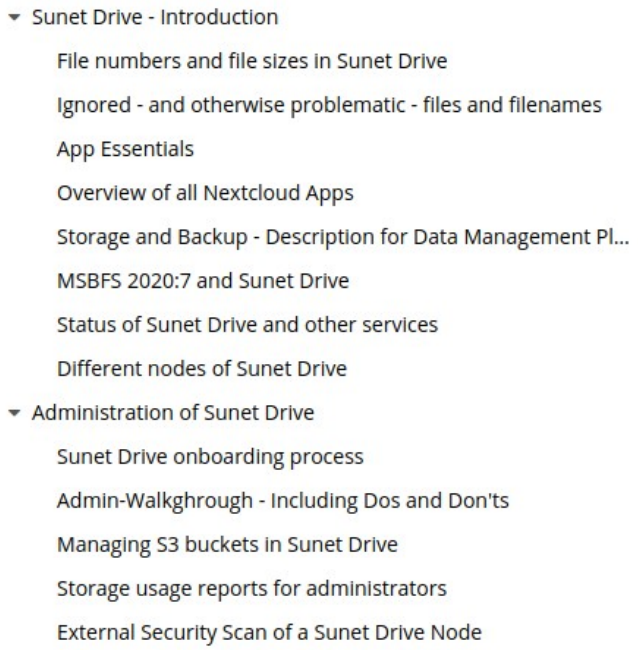
extern.drive.sunet.se

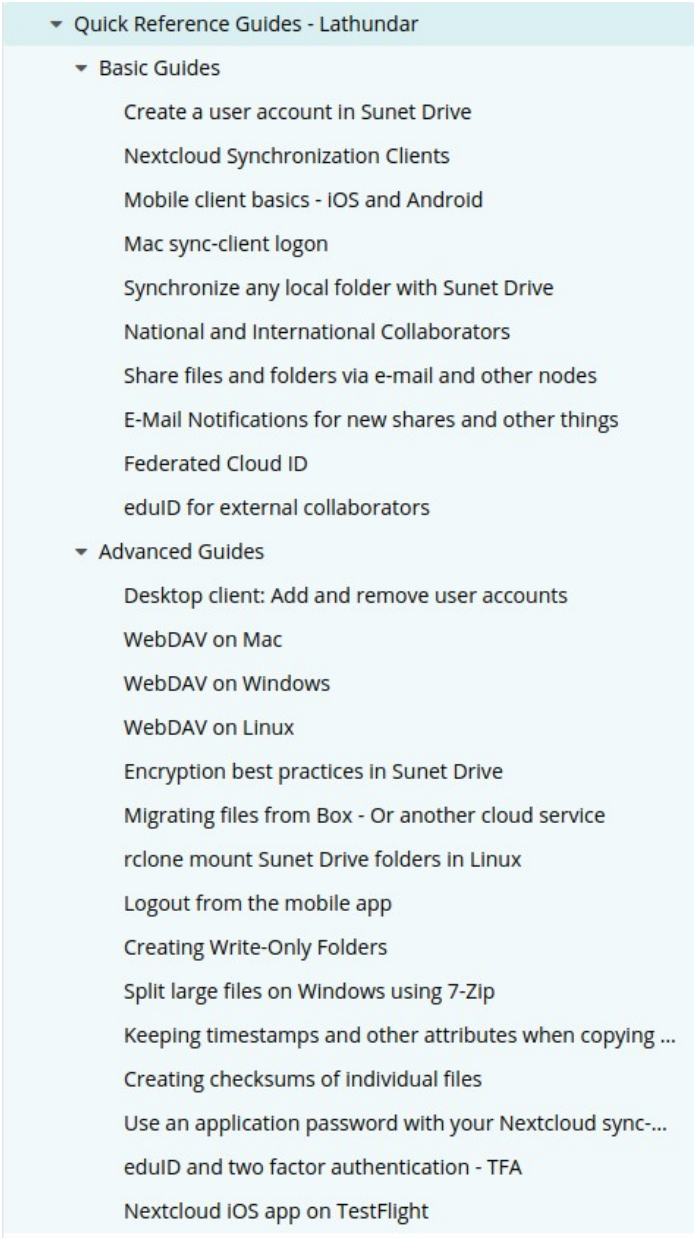
SUNET

Sunet Forum

https://forum.sunet.se/s/sunet-drive/

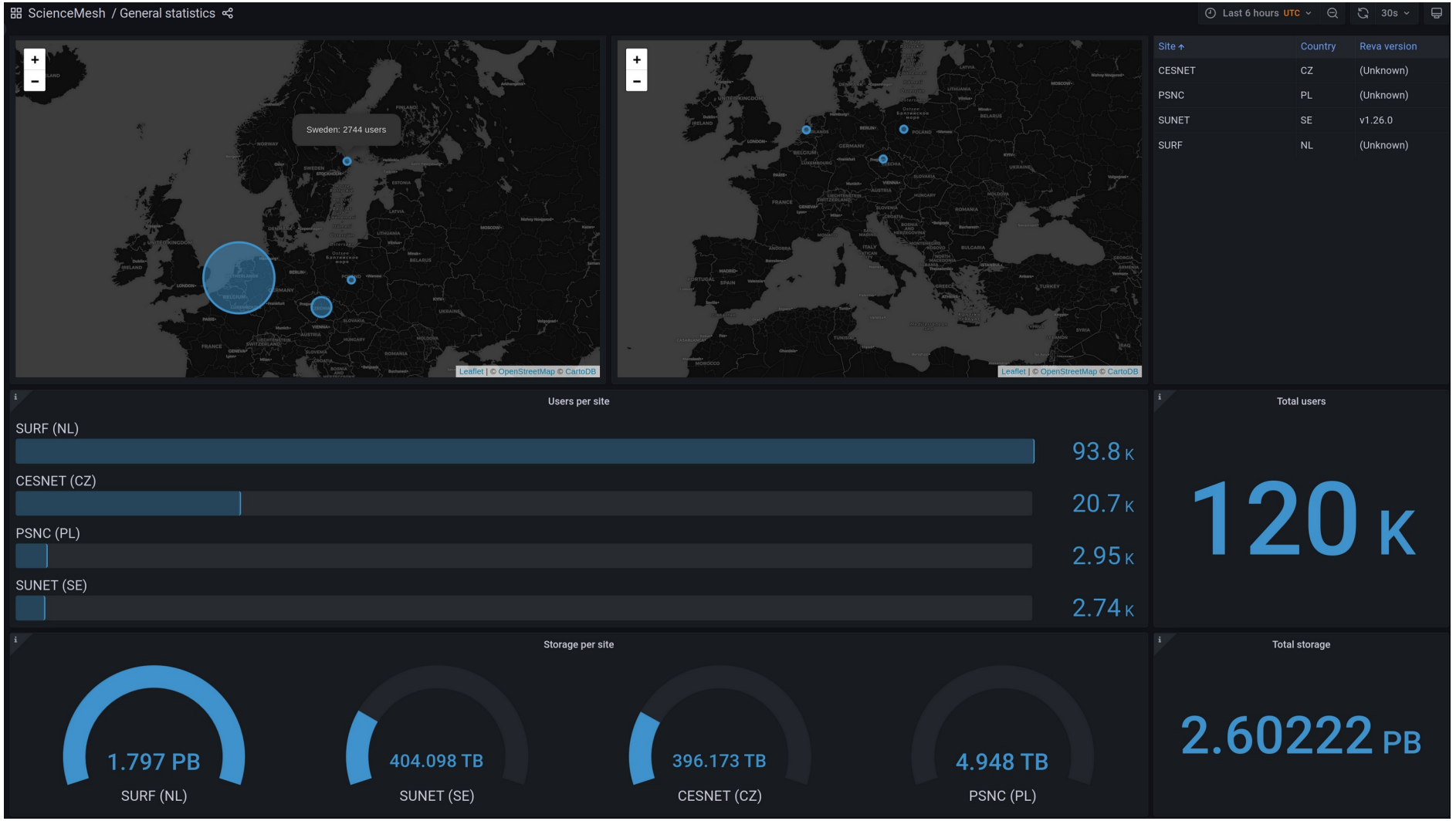
- Central point of documentation for Sunet Drive
- Information Stream
- Files
- Wiki
- Information for users and administrators





SUNET

Sunet Drive in the ScienceMesh



SUNET

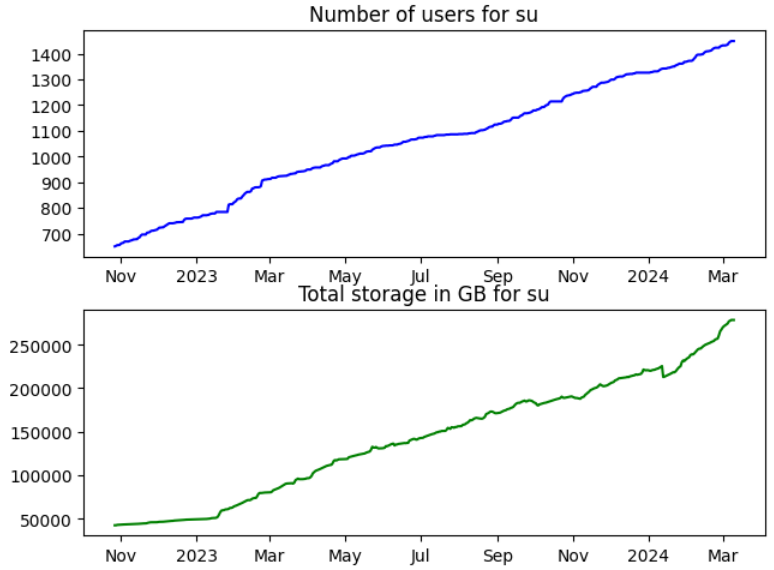
Sunet Drive Operational Statistics

		Jira	Nextcloud	Incidents	Maintenance	Forum	Trainings	Drive-ops	Drive-tests
2023	Jan	26	1	0	3	2	2	340	
	Feb	18	3	1	1	4	2	205	
	Mar	21	2	1	2	5	2	69	
	Apr	17	3	0	1	5	1	40	
	May	18	6	2	2	1	2	154	
	Jun	16	6	1	1	3	2	51	
	Jul	0	3	2	2	0	1	93	21
	Aug	9	0	0	0	2	0	11	10
	Sep	12	5	0	1	6	1	54	12
	Oct	22	2	0	2	2	2	48	27
	Nov	16	3	1	1	3	1	68	28
2024	Dec	11	0	0	1	1	2	21	9
	Jan	7	3	1	1	2	1	37	32
	Feb	14	6	1	2	6	2	140	30
	Mar (so far)	3	6	0	1	1	0	74	18
	Total	**210**	**49**	**10**	**21**	**43**	**21**	**1405**	**187**
	Average	14.0	3.3	0.7	1.4	2.9	1.4	93.7	20.8

SUNET

Sunet drive today

- 2x54 nodes on 300 servers in two data centres
- 100% powered by green energy
- 20 paying customers
- 376TB data stored with >0.5TB added daily
- 2,744 users



Sunet Drive development & projects

- Full RDS-setup in production
- JupyterHub
- Secure Zones/MFA/Step-up Authentication
- Security Audit
- Devops: Third data centre & Kubernetes
- AI Assistant with “Ethical AI”

Core Application
JupyterHub

Core Application
Sciebo Research Data Service

Core Application
ScienceMesh

SUNET

More presentations

Tuesday

Trusted servers and MFA with OCM — *Micke Nordin*

503/1-001 - Council Chamber, CERN — 11:35 - 11:40

SND Doris and Sunet Drive - FAIR and sovereign data publication in a federated world — *Mr Stefan Jakobsson et al.*

503/1-001 - Council Chamber, CERN — 14:15 - 14:30

Wednesday

Continuous Testing at a Global Scale — *Richard Freitag*

503/1-001 - Council Chamber, CERN — 11:45 - 12:00

Utilizing large language models with free and open source software in EFSS while protecting digital sovereignty for Uni

Micke Nordin — 14:15 - 14:30

SUNET

Tack! Frågor?

Anders Nilsson
anders@sunet.se

Richard Freitag
freitag@sunet.se

Micke Nordin
kano@sunet.se

Magnus Andersson
mandersson@sunet.se

Issues and Questions

drive@sunet.se

SUNET