

# SUNET Drive – Swedish Research Storage

A Self-hosted EFSS solution using private cloud infrastructure

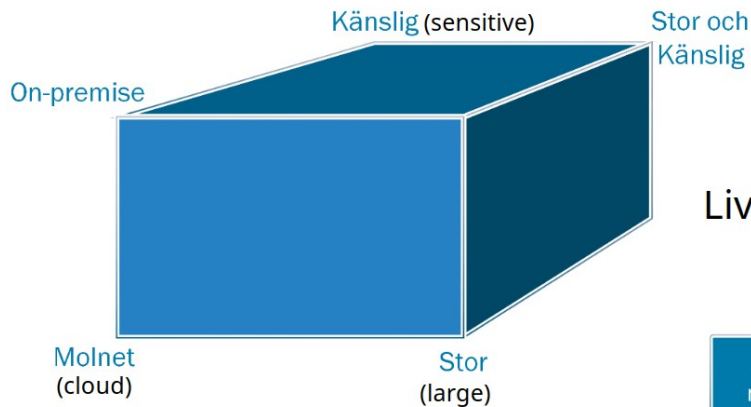
**Richard Freitag**

Room “Paris”

5/6/2023

# Research Data – A model

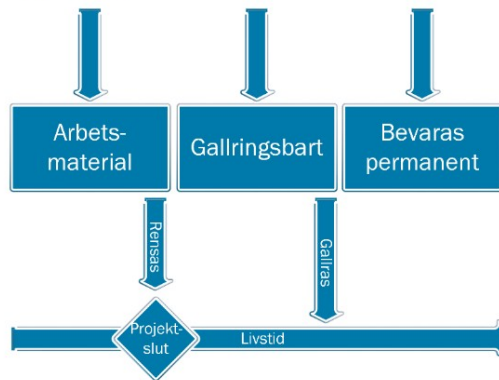
## Classification



How do you access data?

Where is my data?

## Lifecycle



- Working material
- Data retention
- Archival

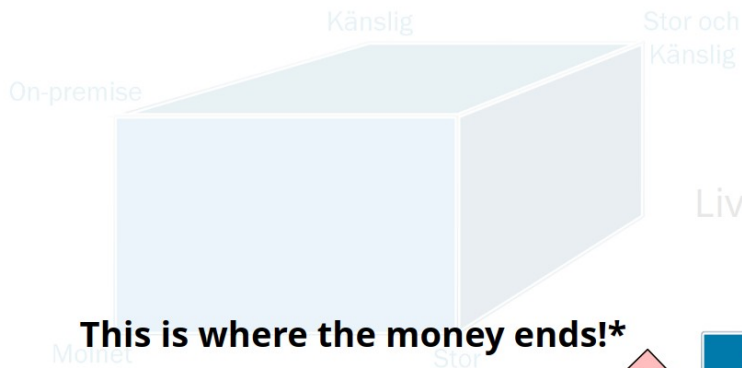
## Access

"Klassisk" Mappar

"Modern" Appar

# Research Data – A model

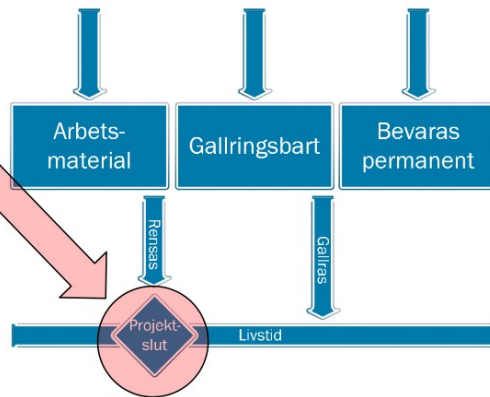
## Classification



How do you access data?

Where is my data?

## Lifecycle



**This is where the money ends!\***

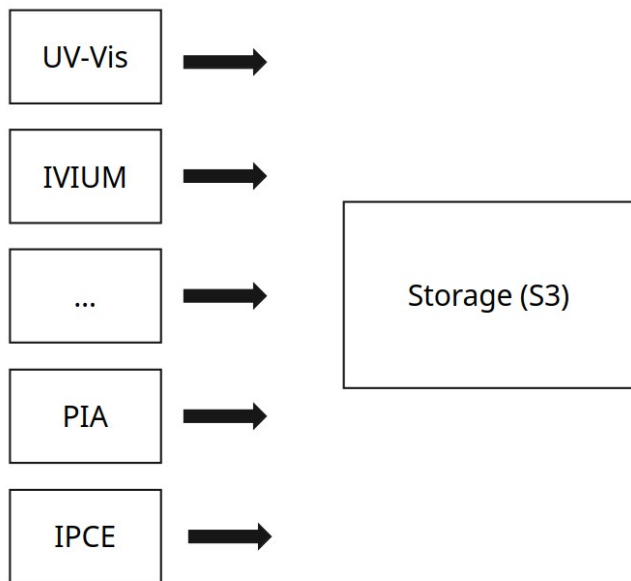
## Access

"Klassisk" Mappar  
\*for many projects

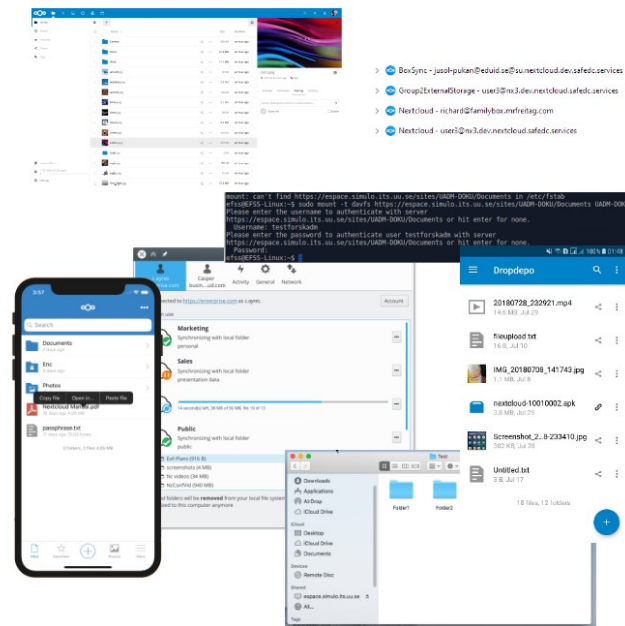
"Modern" Appar

# Data Integration

## Primary data sources

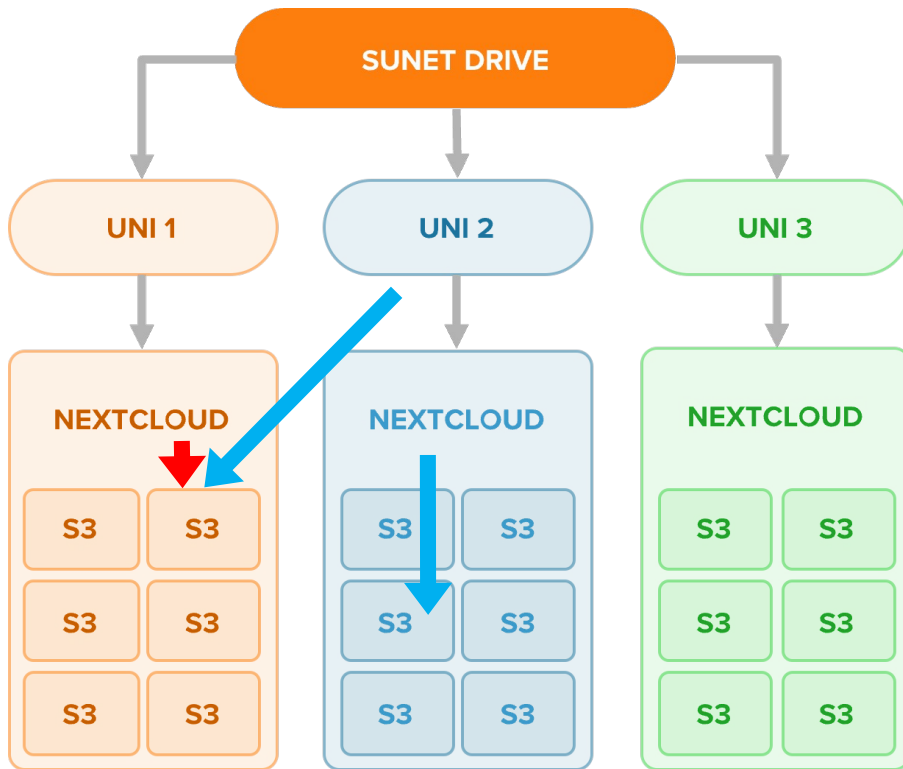


## Access



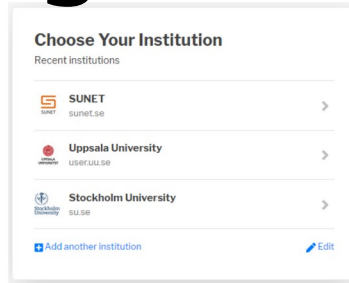
# Sunet Drive

- Federated file sharing
- Co-management of nodes
- Part of ScienceMesh
- S3-bucket storage
- Share data between nodes
- Access ownership separation
- Lifecycle managed buckets
- Private cloud infrastructure
- Nextcloud Enterprise



# Identity Management

Swedish Account



Choose Your Institution

Recent institutions

- SUNET  
sUNET.se
- Uppsala University  
U58F.UU.S8
- Stockholm University  
SU.S8

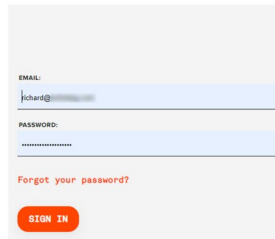
[Add another institution](#) [Edit](#)

[sunet.drive.sunet.se](https://sunet.drive.sunet.se)

[uu.drive.sunet.se](https://uu.drive.sunet.se)

[su.drive.sunet.se](https://su.drive.sunet.se)

External Account



EMAIL:  
richard@

PASSWORD:  
\*\*\*\*\*

[Forgot your password?](#)

[SIGN IN](#)

[extern.drive.sunet.se](https://extern.drive.sunet.se)

edulD is easier and safer login.



# 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.



# SUNET and Hybrid Cloud

## Datacenters/Physical Hardware

- Owned by Sunet and/or universities
- Two datacenters in Stockholm: Sto3/Sto4
- Third datacenter coming: Orion/Kalix
- Hosted by Safespring

## IaaS/StaaS/PaaS

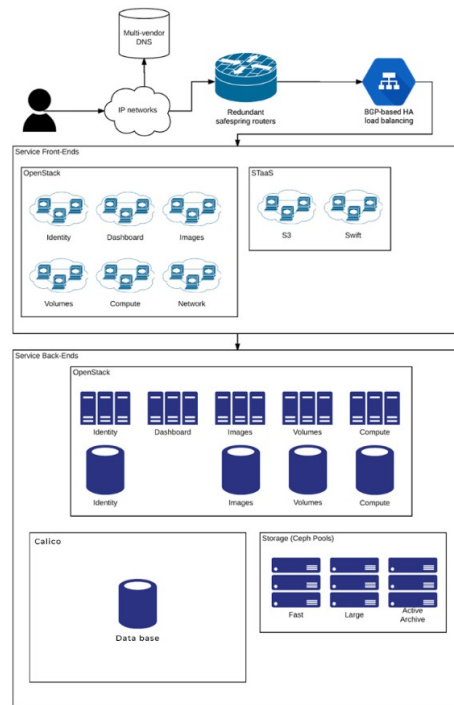
- OpenStack – Virtualization, Storage and more
- Managed by Safespring
- Part of OCRE framework agreement
- Private Cloud Security  
<https://www.safespring.com/documents/sunet/safespring-sunet-private-cloud-security-controls.pdf>

## Universities and Customers

- Identity Provider
- Co-management of nodes
- Data lifecycle management
- Administration and processes

## SUNET – Infrastructure as Code

- Virtual machines
- Puppet-cosmos
- Continuous Delivery - ArgosCD
- Load balancing - HAProxy
- Discovery Service – SeamlessAccess/SaToSa
- MariaDB Cluster/ProxySQL
- Redis Cluster
- Test automation





# Safespring Private Cloud Security

## Personnel control

- Background checks
- Data center staff
- Trusted consultants
- Third party vendors

## Data center security

- Availability
- Fire Protection
- Facility Access Controls
- Physical Intrusion Controls

## Third Party Reviews

## Network Access Controls

- Public Network Access
- Management Network Access

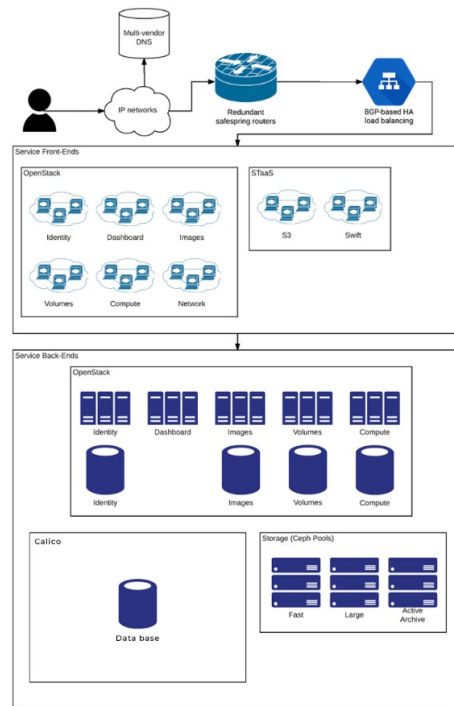
## Systems Management

## Auditing

## Process Description and Configuration Management

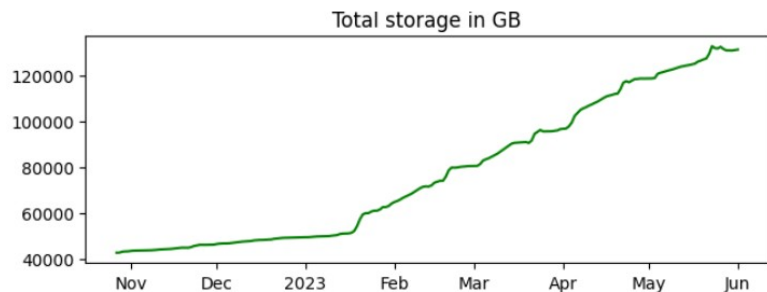
## Infrastructure as a Service

- IaaS high level architecture
- Network exposure management
- IaaS/StaaS – Auditing
- Data Privacy



# Sunet drive today

- 2x54 nodes on 300 servers in two data centres
- 100% powered by green energy
- 17 paying customers
- 200TB data stored with 0.7TB added daily
- ~2,000 users



# Sunet Drive in the future

- JupyterHub integration
- Secure Zones - MFA
- Data migration
- Third data centre (Kalix, Sweden)
- Kubernetes

## FUTURE APPLICATION

Sciebo Research Data Service

## FUTURE APPLICATION

ScienceMesh



**tnc23**

DIGITAL GENERATIONS

TIRANA, ALBANIA | 5-9 JUNE 2023

# Thank you

## Any questions?

[freitag@sUNET.se](mailto:freitag@sUNET.se)



Co-funded by  
the European Union