

# SUNET Drive

An academic EFSS packaged for EOSC



SUNET



nextcloud

News and Development

# Agenda

## Sunet Drive - Now

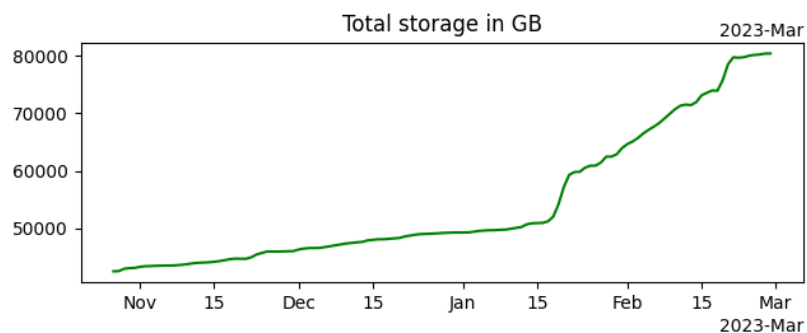
- Sunet Drive Statistics
- Testing
- Sunet Forum and Training
- Sunet Drive Features

## • Planned Functionality

- Secure Zones – Step-up MFA
- Publication of Metadata
- ScienceMesh
- JupyterHub

# Sunet drive today

- 2x54 nodes on 300 servers in two data centres
- 100% powered by green energy
- 15 paying customers
- 120TB data stored with 0.5TB added daily
- Nearly 2,000 users



# Sunet Drive in the future

- Full RDS-setup in production
- Secure Zones
- Data migration
- Third data centre (Kalix, Sweden)
- Kubernetes

FUTURE APPLICATION

Sciebo Research Data Service

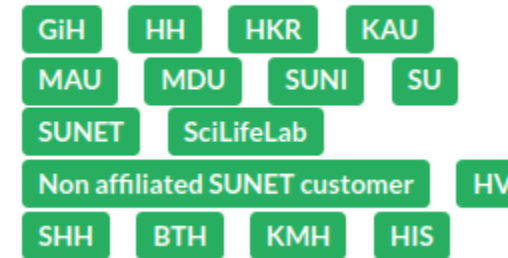
FUTURE APPLICATION

ScienceMesh

# SUNET Drive Operational Status

- Since July 2022: <https://status.sunet.se>
- Since May 2022: Service-account (second-level support) for customers: [drive@sunet.se](mailto:drive@sunet.se)
  - ca. 3-5 tickets per week
- Sunet has Enterprise support from Nextcloud
  - 126 support tickets
  - ca. 1-2 tickets per week
  - Good relationship with Nextcloud

## SUNET Drive



Operational

# Sunet Drive Testing

- Testing on top of monitoring:  
Manual – Automatic – On-demand
- Acceptance Testing
  - General capabilities
  - Status and versions
  - User life-cycle
  - File-handling
- Selenium Testing
  - Node-login, web UI
  - Local and non-MFA SSO accounts
- Collabora testing
  - Document handling and collaboration
- Load testing (for full nodes)
  - 20 virtual machines, 300 users per machine
  - 120-150 full load users (new data)
  - 750-1000 resyncs within 5 minutes
  - 6000 resyncs within 40 minutes

✓	☀️	ScienceMesh Smoketest	23 hr #26	N/A	13 min	▶
✓	☁️	Sunet Drive Acceptance - Prod	48 min #1172	1 hr 53 min #1169	4 min 3 sec	▶
✓	☁️	Sunet Drive Acceptance - Test	51 min #1199	1 hr 57 min #1198	4 min 20 sec	▶
✓	☁️	Sunet Drive Selenium - Prod	44 min #18308	2 hr 3 min #18305	9 min 23 sec	▶
✓	☁️	Sunet Drive Selenium - Test	26 min #17947	1 hr 13 min #17945	11 min	▶
✓	☁️	Sunet Drive Selenium Collabora - Prod	28 min #3407	1 hr 35 min #3405	19 min	▶
✗	☁️	Sunet Drive Selenium Collabora - Test	3 days 5 hr #3182	1 hr 18 min #3243	11 min	▶
✓	☀️	Sunet Drive Selenium RDS - Test	4 hr 46 min #1115	2 days 4 hr #1101	43 sec	▶

# Sunet Forum and Training

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

- Central point of documentation for Sunet Drive
- Information Stream
- Files
- Wiki
- Information for users and administrators
  
- On-demand trainings
  - Admins (ca. 90 minutes)
  - End-users
- Recent training with SU:

<https://sunet.drive.sunet.se/index.php/s/srysAY9oSXoRRSN>

## ☰ Index

- ☰ Admin-Walkthrough - Including Dos and Don'ts
- ☰ Creating Write-Only Folders
- ☰ Different nodes of Sunet Drive
- ☰ E-Mail Notifications for new shares and other things
- ☰ eduID and two factor authentication - TFA
- ☰ eduID for external collaborators
- ☰ Federated Cloud ID
- ☰ File numbers and file sizes in Sunet Drive
- ☰ Ignored - and otherwise problematic - files and filenames
- ☰ Keeping timestamps and other attributes when copying or migrating data
- ☰ Mac sync-client logon
- ☰ Managing S3 buckets in Sunet Drive
- ☰ Migrating files from Box - Or another cloud service
- ☰ MSBFS 2020:7 and Sunet Drive
- ☰ National and International Collaborators
- ☰ Nextcloud iOS app on TestFlight
- ☰ Nextcloud Synchronization Clients
- ☰ rclone mount Sunet Drive folders in Linux
- ☰ Share files and folders via e-mail and other nodes
- ☰ Sunet Drive
- ☰ Sunet Drive onboarding process
- ☰ Synchronizing with multiple accounts and instances
- ☰ Synchronize any local folder with Sunet Drive
- ☰ Use an application password with your Nextcloud sync-client
- ☰ User management in Sunet Drive
- ☰ WebDAV on Linux
- ☰ WebDAV on Windows

# Features and Planned Functionality

- **Production Features**
  - File sync and share
  - Co-management of nodes
  - SSO-Login via SWAMID/EduGAIN
  - Off-site backups using duplicity
  - Document collaboration
- **Under development**
  - Secure Zones – Step-up MFA for certain areas of Sunet Drive
  - Research Data Services (RDS) – Publication of Metadata
  - ScienceMesh integration – European federated EFSS (EOSC)
  - Jupyter Hub – Integration and synchronization
- **Feature ideas/requests**
  - Audit trail/logging
  - EduSign
  - Data migration, ELN, Overleaf
  - Other Apps

# Secure Zones – Background

Enterprise = Among other things: SSO

Nextcloud = Enterprise File Sync and Share

Sunet Drive = Nextcloud + S3 storage

Very early customer feedback on MFA

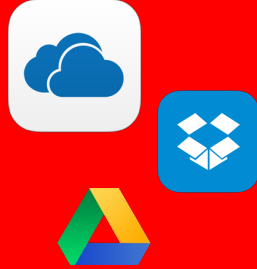








Nextcloud does not supports MFA on SSO accounts

Sunet Drive += MFA





# Secure Zones – Background

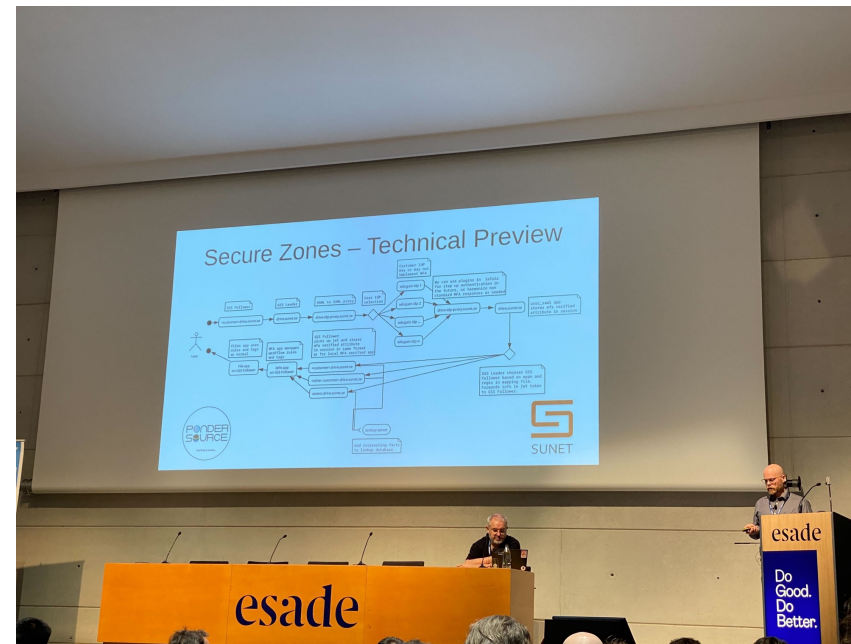
	User friendly	Neutral	Secure
Evil			
Neutral			
Good			



# Secure Zones – Status

Project in 6 steps:

1. Show a boolean flag from user\_saml app in Nextcloud ✓
2. Create a Deny-All zone in Nextcloud ✓
3. Show boolean flag from gss client as well as from local passwordless + MFA ✓
4. In Files app, make MFA-only zone accessible if flag is set ✓
5. GUI for creating and viewing MFA-only zones ✓
6. **Security Audit** 🕒

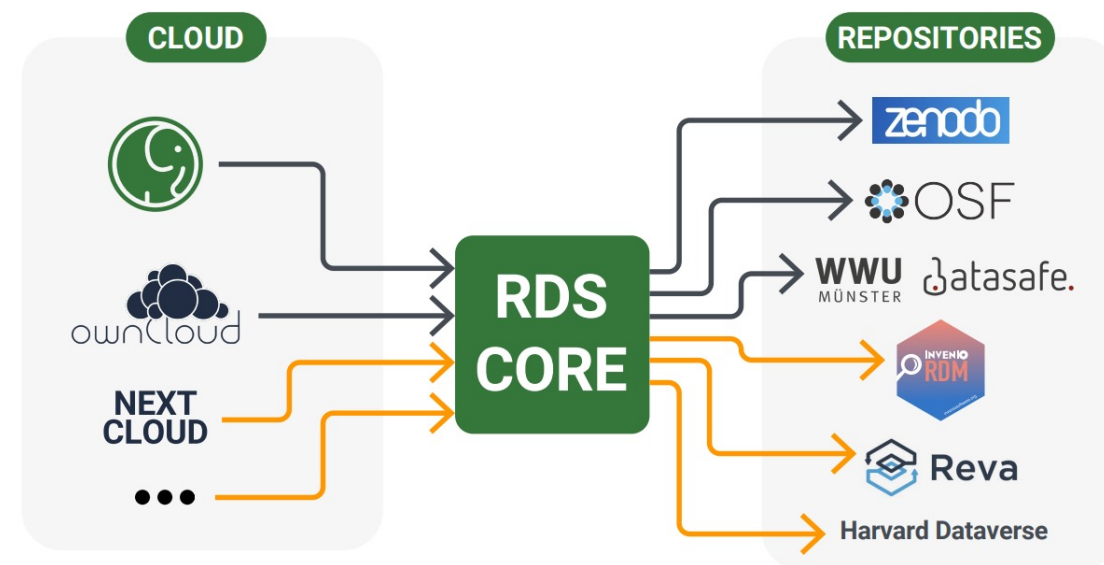


## Sciebo RDS

- Compatibility layer between a host (EFSS) and research data repositories
- Provides metadata annotation via describo, metadata stored as RO-Crate
- Integration via native EFSS plugin system
  - Plugin acts as a wrapper around standalone RDS web app
  - Utilizing basic EFSS functionality like filepickers and oauth for user management
- Each storage provider / repository is implemented as it's own microservice

Easily extendable for your EFSS by writing a native wrapper plugin

Easily extendable for your data repository by writing a microservices that maps to your API



## Porting Sciebo RDS to Nextcloud (Sunet Drive)

### Sunet – RDS Infrastructure

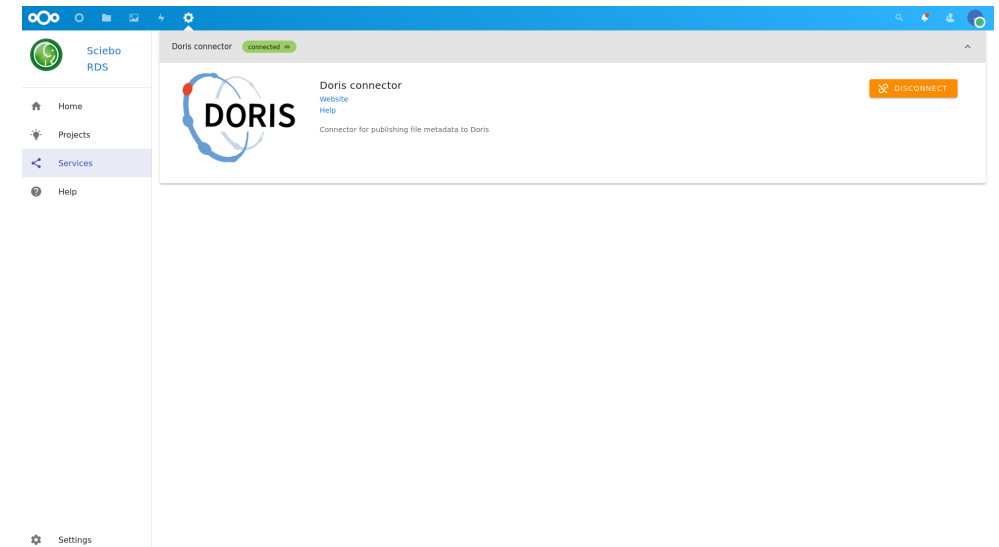
- WWU streamlined infrastructure setup
- Prepare deployment using Sunet infrastructure

### RDS App for Nextcloud

- Initially only minor changes necessary
- Nextcloud updates required further customizations

### SND/Doris Connector

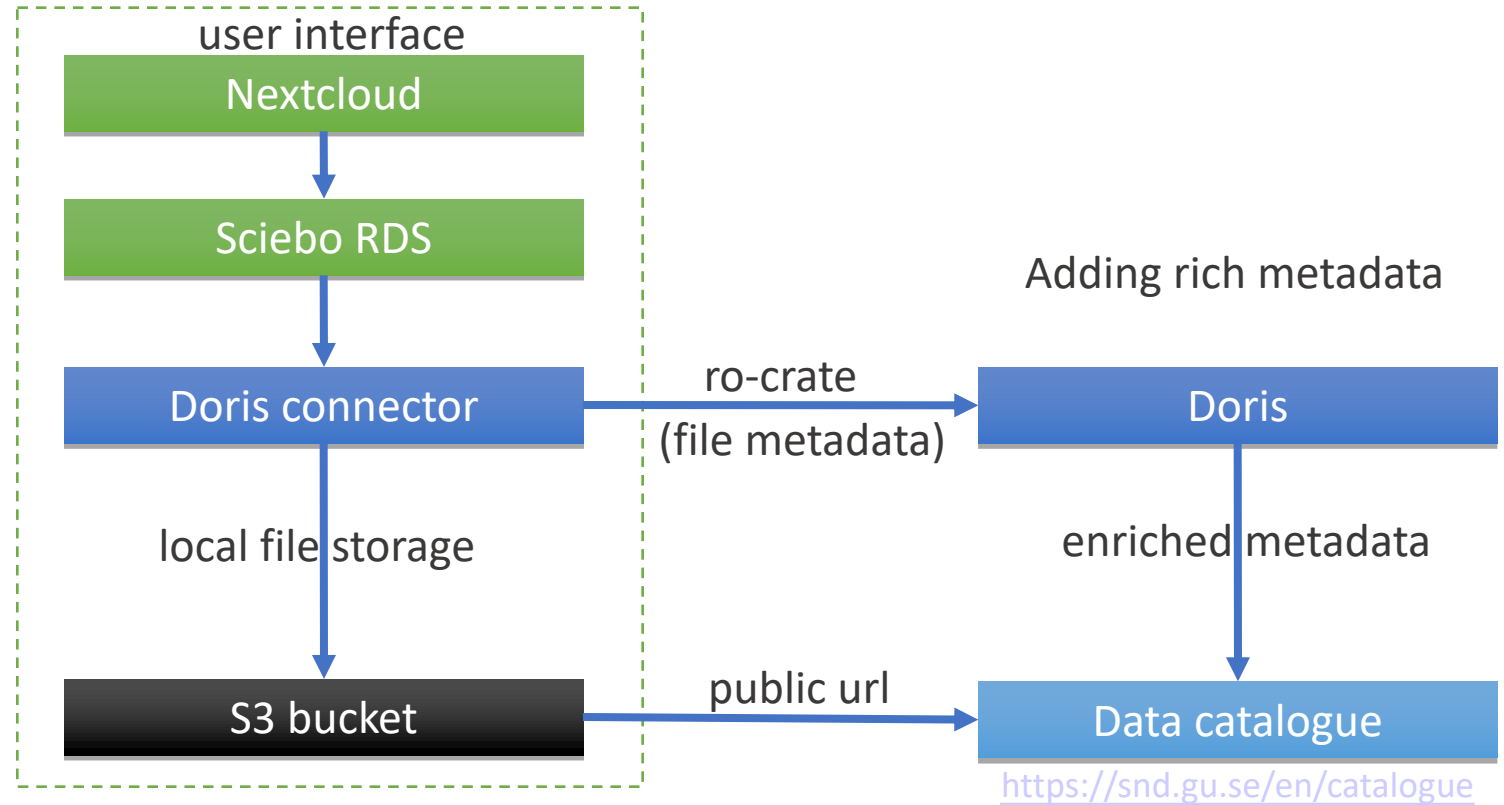
- Define architecture and design within the context of both Doris and Sunet Drive
- Specify integration into storage architecture (S3), ideally avoiding a dependency on Nextcloud
- Deploy app to be used by reference users



Most components still require some development effort, but we are getting there!

## DORIS

- **Status:** (Early Development)
- Metadata management system
- Cross discipline
- Heavy use of controlled vocabularies



More information about sharing data via DORIS  
<https://snd.gu.se/en/describe-and-share-data>

## Development

- Core Components developed at the University of Münster
  - **RDS Core**
  - **Owncloud** plugin
  - Connectors for **Zenodo**, **OpenScienceFramework** and **WWU Datasafe**
- Adaption and extensions by Partners like Sunet, SND and SURF
  - **Nextcloud** port for the Owncloud plugin (Sunet)
  - **Doris** connector (SND)
  - **Figshare** connector (SURF)
  - **Harvard Dataverse** connector (SURF)
- **Everything openly available on Github**

**Sciebo RDS is Open Source and has a strong community**

# RDS – Planning and Rollout

Component	Description	Status	Planning
Sunet RDS App and Infrastructure	Sunet self-hosted RDS infrastructure	Test/Release Candidate	Test: Q2 - 2023 Production: Q2/Q3 - 2023
Sciebo-Connectors	Connectors developed by Sciebo development team	Test/Release Candidate	Test: Q2 - 2023 Production: Q2/Q3 - 2023
SURF-Connector	Figshare, iRods, Harvard Dataverse	Under Development	Dependent on external development
SND-Connector	Connector for Doris	(Early) Development	Dependent on external development

# Science Mesh

- **Status:** Adoption phase aligned with ScienceMesh roadmap
- Federation of EFSS systems with standardized functionality
- Sunet is external participant in EOSC-funded CS3MESH4EOSC

<https://sciencemesh.io/>



**Federated sharing of data** across different sync-and-share services.



**Federated use of applications** such as collaborative document editing, data archiving, and data publishing.



**Fast transfer of large datasets** from one site to another.



**Remote data analysis** through sync-and-share services.

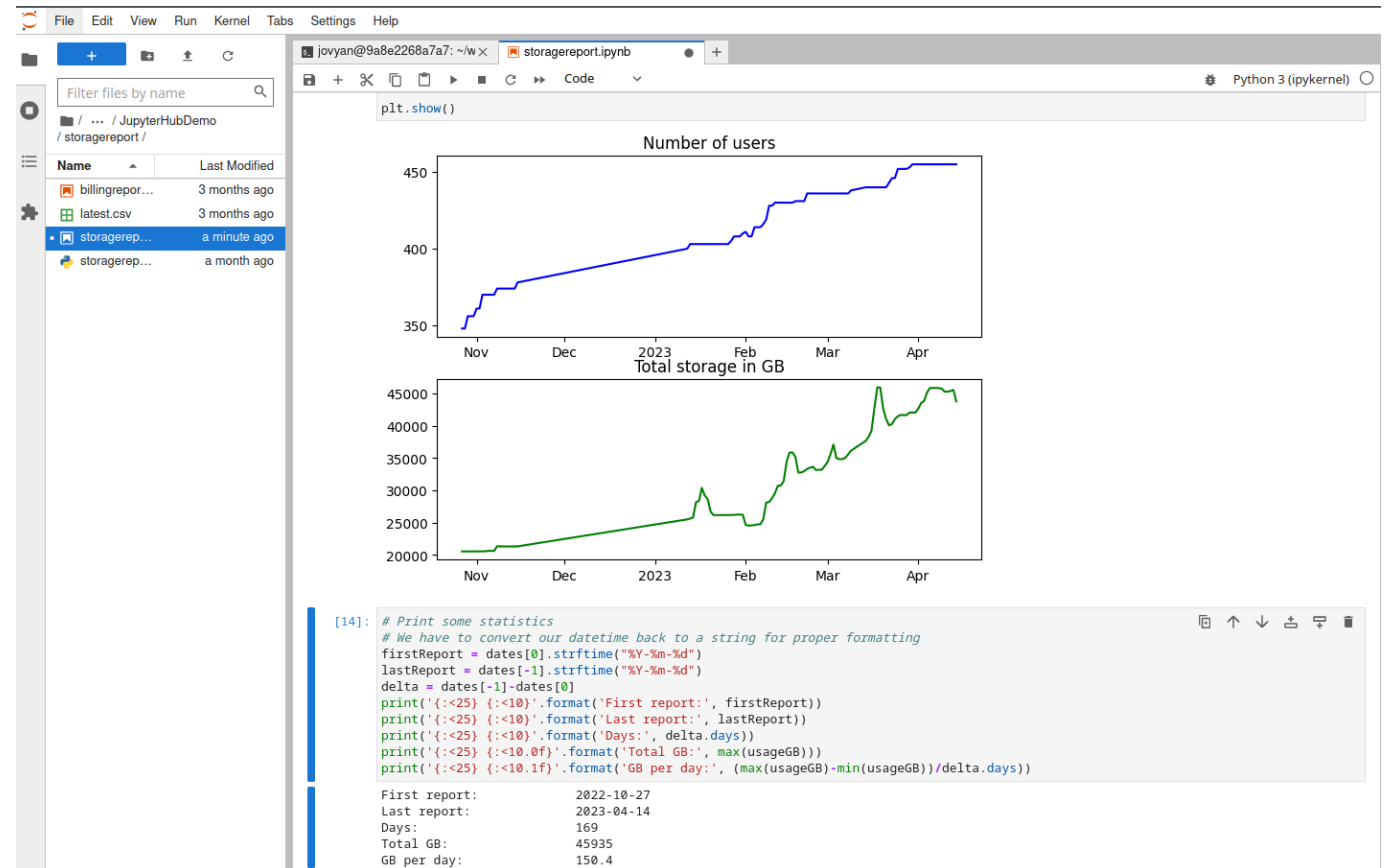


# Jupyter Hub

- Status: **Prototype**

## Features:

- Per-user execution of Jupyter Notebooks
- Synchronization of files between Sunet Drive and Jupyter Hub



# Data and compliance apps

App/Feature	Description	Status
Audit Trail/Logging	Official Nextcloud app for audit trail and logging	Likely
EduSign	Nextcloud app/wrapper for EduSign	Q4/2023
Data Migration	Automatic migration for Box or other services	Under Evaluation Box migrator likely
Electronic Lab Notebooks	Integration of (Open Source) ELN solution	Under Evaluation
Overleaf	Nextcloud app for Overleaf, developed by WWU Münster	Under Evaluation

# General Apps (Under Evaluation)

App/Feature	Description/Comment	Status
Forms	Beneficial for collecting research data	Likely
Polls	Beneficial for collecting research data	Likely
Draw.io	Diagram editor	Likely
Maps	Specific solution needs to be evaluated	Uncertain
Tasks	Global/Federated solution needs to be evaluated	Uncertain
Mind Map	Global/Federated solution needs to be evaluated	Uncertain
Deck	More likely to be part of different product/solution	Unlikely
Calendar	More likely to be part of different product/solution	Unlikely
Talk	More likely to be part of different product/solution	Unlikely
Commercial apps	Integration of other commercial apps	Unlikely

# Feature Survey



<https://forum.sunet.se/content/perma?id=26403>

# Tack! Frågor?

**Richard Freitag**

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

**Anders Nilsson**

[anders@sunet.se](mailto:anders@sunet.se)

**Micke Nordin**

[kano@sunet.se](mailto:kano@sunet.se)