

# LUMI

A white wolf is the central focus, standing in a futuristic, blue-toned digital environment. The background is filled with vertical data streams, particle effects, and a grid-like structure, creating a high-tech, cybernetic atmosphere. The wolf is looking slightly to the right of the viewer.

## Files on LUMI 2: LUMI-O object storage

**Kurt Lust**  
LUMI User Support Team (LUST)  
University of Antwerp

Evolving version, last update Dec. 2024

# Why do I need to know this?

- LUMI-O is the primary option on LUMI to transfer large amounts of data to LUMI
- LUMI-O is the only local option if you want to backup some data
- Some datasets come in a format optimised for object storage rather than a parallel file system
  
- What we will discuss:
  - Properties of object storage
  - Getting started
  - But not a reference manual of the tools that can be used on LUMI-O

# What is LUMI-O?

- Object storage system, based on Ceph
  - Finnish users: similar to Allas, but less functionality at the moment
  - Specific tools to access data, not mounted as a regular file system
- Organisation:
  - Projects: LUMI projects
  - Buckets: “Containers” used to store objects.
    - Flat structure: Buckets cannot contain other buckets
  - Objects: Any type of data, stored in a bucket
    - Atomic access to objects: Put, get, copy, delete, ..., but no partial write
  - Metadata for buckets and objects
    - Bucket: e.g., access rights
    - Custom metadata possible

# What is LUMI-O? (2)

- Objects can be served on the web also
  - This is how recordings of some LUST courses are served
  - But not meant as a data publishing service (e.g., no EUDAT alternative)
- Can be reached easily from outside LUMI
  - So also a mechanism for data exchange
  - Tools of object storage are more performance and more robust than sftp
- Specs:
  - Capacity: 30 PB
  - Quota: 150 TB capacity, 1K buckets and 500K objects per bucket (fixed)
  - Billed at 0.5 TB·hour per TB per hour
  - Persistent for the duration of the project

# Lustre vs LUMI-O (1)

Lustre	LUMI-O object storage
Closely integrated with compute nodes	Separate system
Upgrades with the system	Separate upgrade cycle and on-the-fly
Organisation: Hierarchical directory structure and files	Organisation: Triple flat space of projects, buckets and objects
Files can be read, written, modified, appended, ...	Simple atomic operations on objects: put, get, copy, delete
Optimised for bandwidth to the compute nodes	Optimised for reliability and resilience
Simpler schemes for redundancy	Very complex internal redundancy setup

# Lustre vs LUMI-O (2)

Lustre	LUMI-O object storage
Integrated in the authentication of the supercomputer	Separate key-based authentication mechanism
Seen as any other POSIX file system	Separate range of access tools/APIs, some tools can provide a filesystem view
No external access	External access integrated, includes web
Structure with MDS and ODS	Structure with MDS and ODS, but very different technologies
Parallelism for performance: Access a file in parallel from multiple processes	Parallelism for performance: Different processes access different objects
Fairly expensive to very expensive hardware	Cheaper hardware

# LUSTRE vs LUMI-O (3)

- The optimal way/technology of storing data is very different depending on whether you work from a parallel file system or from object storage. E.g., in earth and climate science:
  - netCDF is a popular data storage format for storing simulation data on a parallel filesystem
    - Not suited for object storage though as it would be a single object
  - Zarr is a format to store similar data on object storage (cloud storage)
    - It is not a single object, but a structured collection of objects
    - Putting it on a parallel filesystem where each object would become a file in a directory tree is a very bad idea!
    - But with the right libraries, you can access Zarr data on object storage directly from your application

# Accessing LUMI-O

- Access is based on authentication keys
  - Generated via a web interface: Separate steps to generate the credentials and to get them on LUMI.
  - Or generated via Open OnDemand: Can put credentials on LUMI for `rc1one` and `s3cmd`.
- Tools:
  - `rc1one`: Easiest tool if you want public and private data
  - `s3cmd`
  - `restic`: More a backup tool
  - `boto3`: Python API from the AWS SDK for programmatic access
    - Needs a more recent Python version than the system Python
  - Additional GUI-based tools exist for clients
  - Open OnDemand web interface is not a substitute!
    - Speed limited by browser protocols

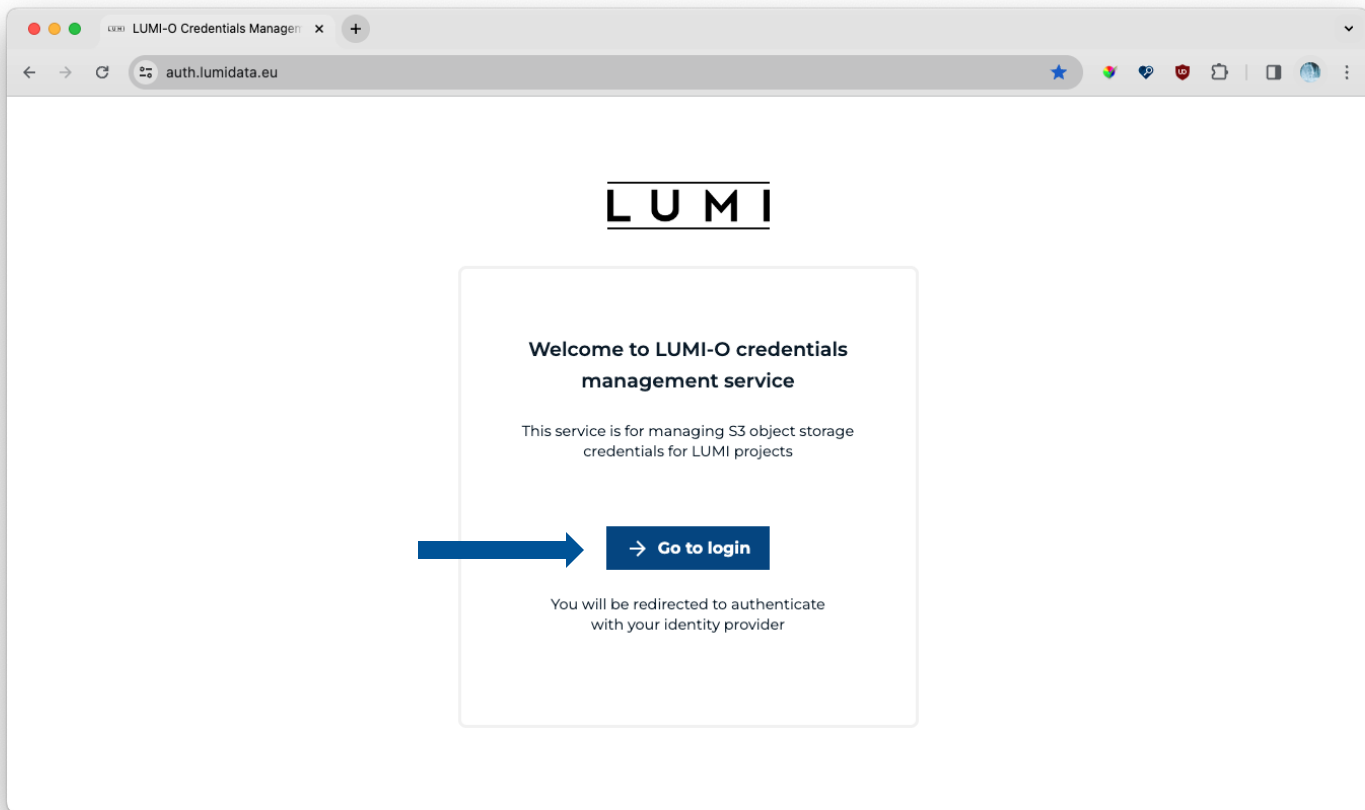
module `lumio`



# The credential management web interface

- Credential management web interface at [auth.lumidata.eu](https://auth.lumidata.eu)
  - Create keys
  - Extend lifetime of a key
  - Create configure scripts for various tools

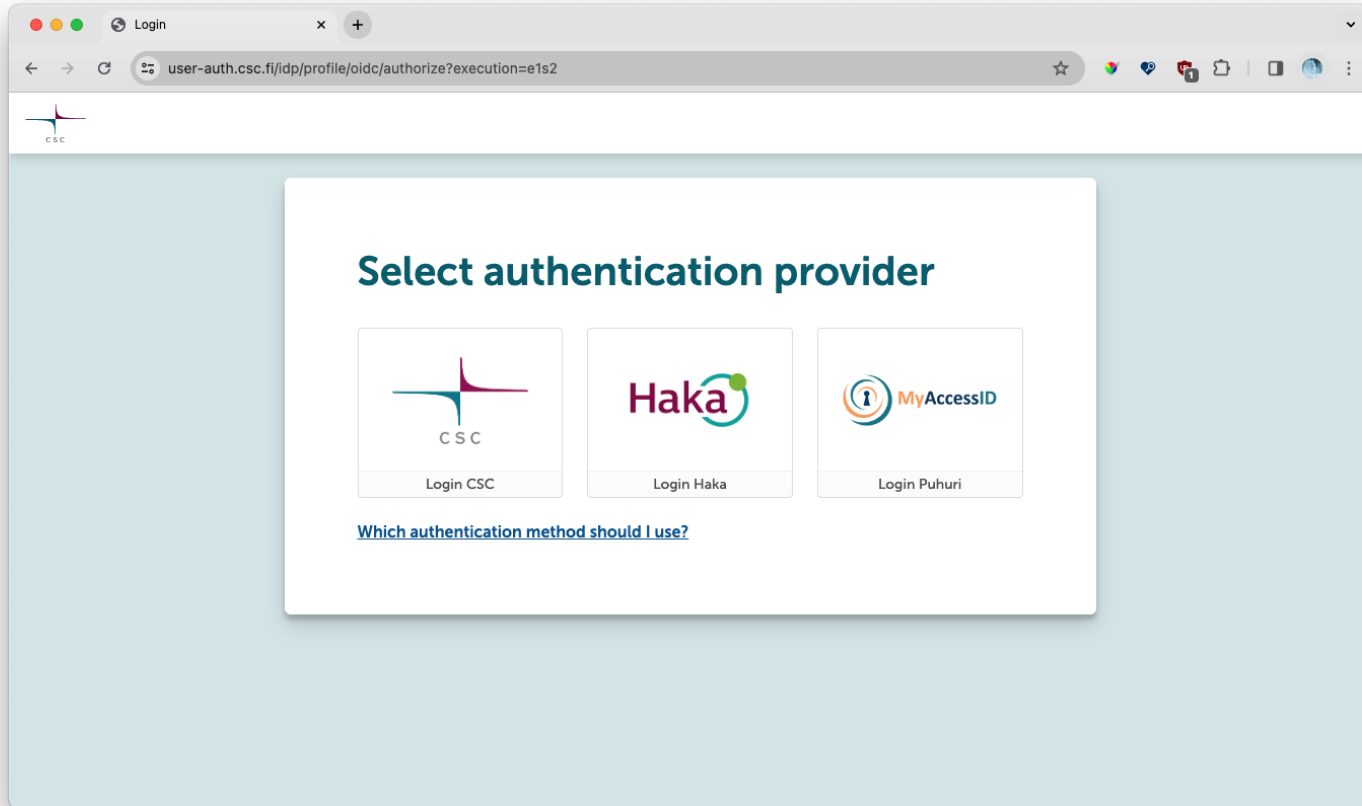
# Credential mangement web interface: Create credentials (1)



# The credential management web interface

- Credential management web interface at [auth.lumidata.eu](https://auth.lumidata.eu)
  - Create keys
  - Extend lifetime of a key
  - Create configure scripts for various tools
- You'll have to select your login method in the same way as for Open OnDemand

# Credential mangement web interface: Create credentials (2)



# The credential management web interface

- Credential management web interface at [auth.lumidata.eu](https://auth.lumidata.eu)
  - Create keys
  - Extend lifetime of a key
  - Create configure scripts for various tools
- You'll have to select your login method in the same way as for Open OnDemand
- After a while you should see a list of projects, select the one for which you want to generate a key
  - The right column will also show active keys for the project, and expired ones

# Credential management web interface: Create credentials (3)

The screenshot shows the LUMI Credentials Manager web interface. The browser address bar displays `auth.lumidata.eu/projects/465001102`. The page header includes the LUMI logo and navigation links for Help & Support, Documentation, and Logout.

### Your projects

Project number	Project description
465000095	VLAAMS SUPERCOMPUTER CENTRUM / VSC-SUPPORT
465000297	LUST Training / Detailed introduction to the LUMI-C environment and architecture (23-24 Nov 22)
465000844	VLAAMS SUPERCOMPUTER CENTRUM / VSC-2023-04-T159-KH-EASYBUILD
465000961	LUST Training / 2024-02-08 LUMI Intro
465001098	LUST Training / 2024-04-23-26 LUMI General
465001102	LUST Training / 2024-05-02-03 Supercomputing with LUMI

The project with ID 465001102 is highlighted in blue, with a blue arrow pointing to it from the left.

### Authentication keys

Project number: 465001102

#### Generate a new authentication key pair

Both fields are required. Key duration may not exceed 168 hours.

Duration (hours)\*:

Key description\*:

**Generate key**

A blue arrow points from the 'Generate key' button to the right.

#### Available keys

There are no keys to show.

The keys you generate will appear in this list.

# Credential mangement web interface: Create credentials (4)

The screenshot shows the LUMI Credentials Manager web interface. The browser address bar displays `auth.lumidata.eu/projects/465001102`. The page header includes the LUMI logo and navigation links for Help & Support, Documentation, and Logout.

### Your projects

Project number	Project description
465000095	VLAAMS SUPERCOMPUTER CENTRUM / VSC-SUPPORT
465000297	LUST Training / Detailed introduction to the LUMI-C environment and architecture (23-24 Nov 22)
465000844	VLAAMS SUPERCOMPUTER CENTRUM / VSC-2023-04-T159-KH-EASYBUILD
465000961	LUST Training / 2024-02-08 LUMI Intro
465001098	LUST Training / 2024-04-23-26 LUMI General
465001102	LUST Training / 2024-05-02-03 Supercomputing with LUMI

The last row of the table is highlighted in light blue, and a blue arrow points from its description to the 'Generate key' button.

### Generate a new authentication key pair

Both fields are required. Key duration may not exceed 168 hours.


**Duration (hours)\***

**Key description\***

**Generate key**

### Available keys

**Access key**

 OGR2KN5PUPW929W9WPP4  
Course demo  
Expires on: Apr 18 2024 11:29:54 GMT+0200

### Expired keys

The are no expired keys to show.

# The credential management web interface

- Credential management web interface at [auth.lumidata.eu](https://auth.lumidata.eu)
  - Create keys
  - Extend lifetime of a key
  - Create configure scripts for various tools
- You'll have to select your login method in the same way as for Open OnDemand
- After a while you should see a list of projects, select the one for which you want to generate a key
  - The right column will also show active keys for the project, and expired ones
- Selecting an active access key changes the right column to one where you get information about the key, can extend the key and can generate templates to configure various tools



# Credential management web interface:

## Check credentials

The screenshot shows the LUMI Credentials Management web interface. The browser address bar displays the URL: `auth.lumidata.eu/projects/465001102/keys/OGR2KN5PUPW929W9WPP4`. The page features a navigation bar with the LUMI logo, 'Help & Support', 'Documentation', and 'Logout' links.

**Your projects**

Project number	Project description
465000095	VLAAMS SUPERCOMPUTER CENTR... SUPPORT
465000297	LUST Training / Detailed introducti... C environment and architecture (2)
465000844	VLAAMS SUPERCOMPUTER CENTR... 04-T159-KH-EASYBUILD
465000961	LUST Training
465001098	LUST Training
465001102	LUST Training / 2024-05-02-03 Sup... with LUMI

**Authentication keys**

### Access key details

<b>Access key</b>	OGR2KN5PUPW929W9WPP4
<b>Secret key</b>	LyYAKH4I5oULHMbwisOskQwdl6xkm03c2G8Jam9
<b>Key description</b>	Course demo
<b>Project number</b>	465001102
<b>Project description</b>	LUST Training / 2024-05-02-03 Supercomputing with LUMI
<b>Creation time</b>	Apr 11 2024 11:29:54 GMT+0200
<b>Expires on</b>	Apr 18 2024 11:29:54 GMT+0200

**Extend key duration**

The expiry time of a key is calculated from the time of its generation and cannot...

**Endpoint URL:** `https://lumidata.eu/`

# Credential management web interface: Extend credential lifetime

The screenshot shows a web browser window with the URL `auth.lumidata.eu/projects/465001102/keys/OGR2KN5PUPW929W9WPP4`. The page features the LUMI logo and navigation links for Help & Support, Documentation, and Logout.

### Your projects

Project number	Project description
465000095	VLAAMS SUPERCOMPUTER CENTI SUPPORT
465000297	LUST Training / Detailed introducti C environment and architecture (2
465000844	VLAAMS SUPERCOMPUTER CENTI 04-T159-KH-EASYBUILD
465000961	LUST Training / 2024-02-08 LUMI Ir
465001098	LUST Training / 2024-04-23-26 LUM
465001102	LUST Training / 2024-05-02-03 Sup with LUMI

A blue arrow points from the 'Project description' column to the 'Extend key duration' section.

### Extend key duration

The expiry time of a key is calculated from the time of its generation and cannot exceed a total of 168 hours.

Extend by (hours)

### Configuration templates

Select configuration format to generate (opens in a new tab)

Select format

- shell
- boto3
- rclone
- s3cmd
- aws

manently disable all connections where this key has been he connection

# Credential management web interface: Tool configuration (1)

The screenshot shows the LUMI Credentials Manager web interface. The browser address bar displays the URL: `auth.lumidata.eu/projects/465001102/keys/OGR2KN5PUPW929W9WPP4`. The page header includes the LUMI logo and navigation links for [Help & Support](#), [Documentation](#), and [Logout](#).

### Your projects

Project number	Project description
465000095	VLAAMS SUPERCOMPUTER CENTI SUPPORT
465000297	LUST Training / Detailed introducti C environment and architecture (2
465000844	VLAAMS SUPERCOMPUTER CENTI 04-T159-KH-EASYBUILD
465000961	LUST Training / 2024-02-08 LUMI Ir
465001098	LUST Training / 2024-04-23-26 LUM
465001102	LUST Training / 2024-05-02-03 Sup with LUMI

Two blue arrows point from the project descriptions to the 'Extend key duration' section. The first arrow points from 'VLAAMS SUPERCOMPUTER CENTI 04-T159-KH-EASYBUILD' to the 'Extend by (hours)' input field. The second arrow points from 'LUST Training / 2024-05-02-03 Sup with LUMI' to the 'Generate' button.

### Extend key duration

The expiry time of a key is calculated from the time of its generation and cannot exceed a total of 168 hours.

Extend by (hours)

### Configuration templates

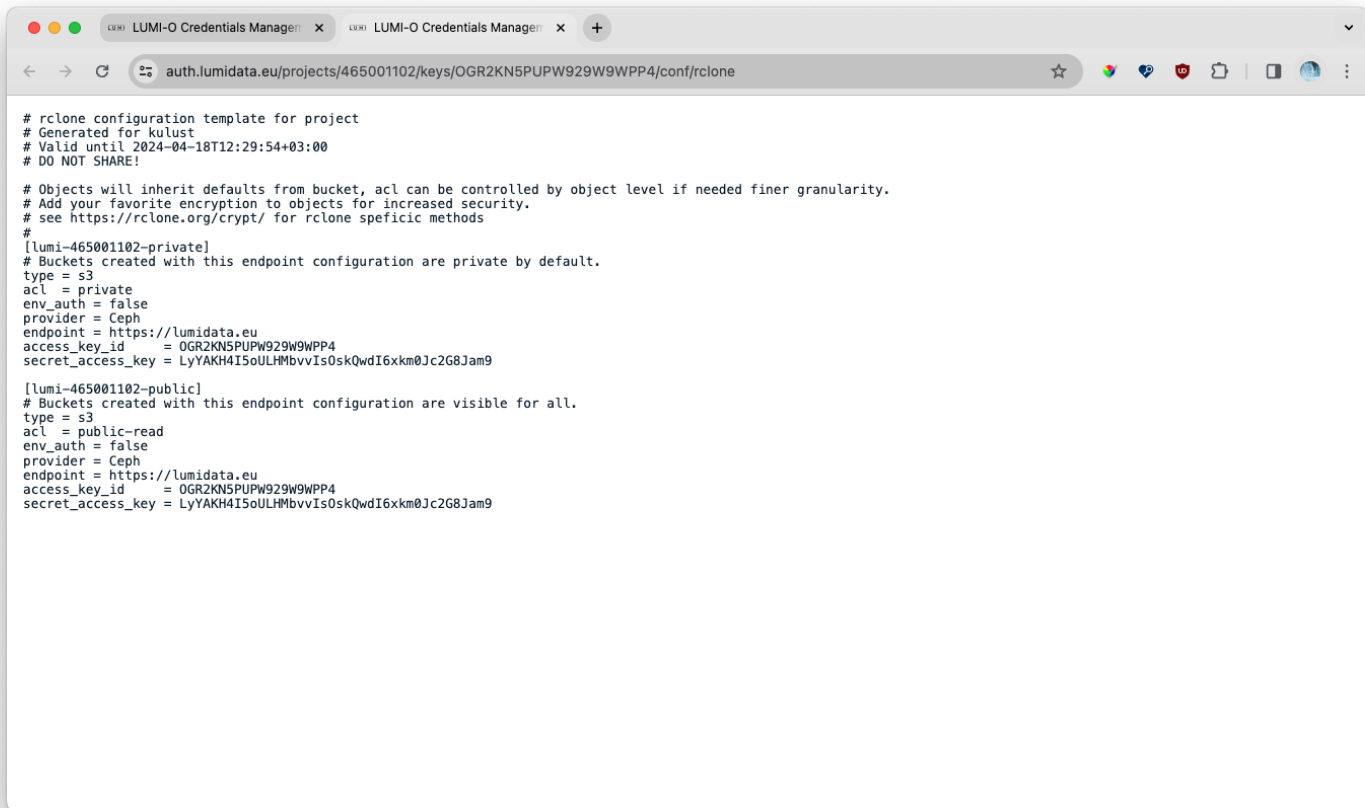
Select configuration format to generate (opens in a new tab)

- shell
- boto3
- rclone
- s3cmd
- aws

manently disable all connections where this key has been he connection

# Credential mangement web interface: Tool configuration (2)



The image shows a web browser window with two tabs. The active tab is titled "LUMI-O Credentials Manager" and displays a configuration template for rclone. The URL in the address bar is "auth.lumidata.eu/projects/465001102/keys/OGR2KN5PUPW929W9WPP4/conf/rclone". The content of the page is a text-based configuration template for rclone, including comments and specific settings for private and public buckets.

```
# rclone configuration template for project
# Generated for kulust
# Valid until 2024-04-18T12:29:54+03:00
# DO NOT SHARE!

# Objects will inherit defaults from bucket, acl can be controlled by object level if needed finer granularity.
# Add your favorite encryption to objects for increased security.
# see https://rclone.org/crypt/ for rclone specific methods
#
[[lumi-465001102-private]
# Buckets created with this endpoint configuration are private by default.
type = s3
acl = private
env_auth = false
provider = Ceph
endpoint = https://lumidata.eu
access_key_id = OGR2KN5PUPW929W9WPP4
secret_access_key = LyYAKH4I5oULHMbvvis0skQwdI6xkm0Jc2G8Jam9
















[[lumi-465001102-public]
# Buckets created with this endpoint configuration are visible for all.
type = s3
acl = public-read
env_auth = false
provider = Ceph
endpoint = https://lumidata.eu
access_key_id = OGR2KN5PUPW929W9WPP4
secret_access_key = LyYAKH4I5oULHMbvvis0skQwdI6xkm0Jc2G8Jam9
```

# Access through Open OnDemand

- Open OnDemand provides
  - An app for simple credential management, including generation of config files for [rclone](#) and [s3cmd](#) on LUMI.
  - The “Home Directory” app can be used to browse, download and upload objects and to create buckets.  
Structure resembles view of a typical rclone-based browser.
- Note that Open OnDemand is not a replacement for proper tools for access to LUMI-O!
  - Uploading and downloading is through web protocols that are not as performant as proper object storage tools.

# Credential management via OOD (1)

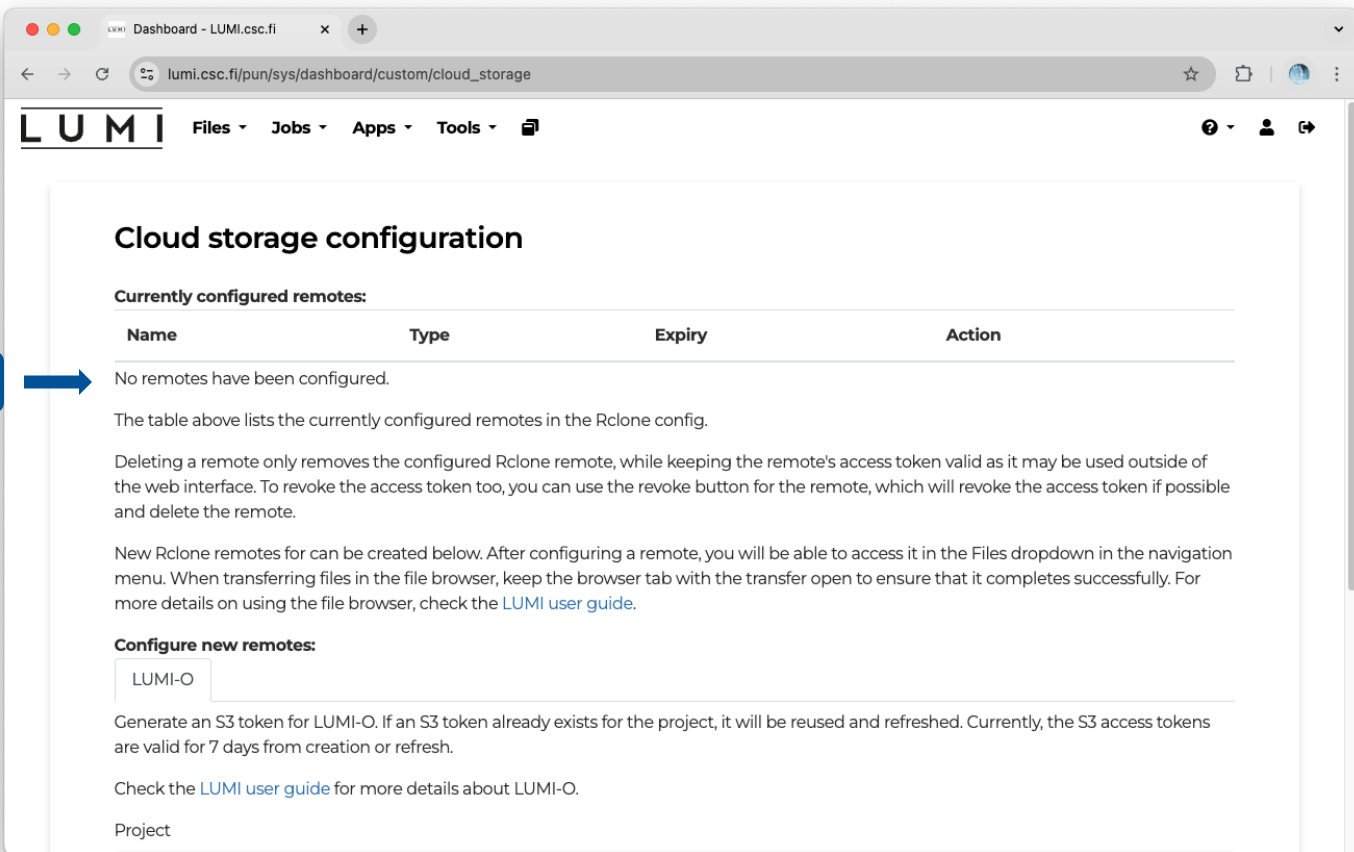
The image shows a web browser window displaying the LUMI dashboard. The browser's address bar shows the URL `lumi.csc.fi/pun/sys/dashboard/`. The dashboard header includes the LUMI logo and navigation menus for Files, Jobs, Apps, and Tools. The main content area is titled "Pinned Apps" and contains a grid of application tiles. A blue arrow points from the "Desktop" tile to the "Cloud storage configuration" tile. Below the pinned apps, a "Notifications" section is visible, indicating that there are no notifications.

Pinned Apps				
 Home Directory	 Compute node shell	 Login node shell	 Desktop	 Cloud storage configuration
 Disk quotas	 Project view	 Active Jobs	 Jupyter	 Jupyter for courses
 Julia-Jupyter	 MATLAB	 MLflow	 TensorBoard	 Visual Studio Code

**Notifications**  
You have no notifications

# Credential management via OOD (2)

## Overview



Dashboard - LUMI.csc.fi

lumi.csc.fi/pun/sys/dashboard/custom/cloud\_storage

**LUMI** Files Jobs Apps Tools

### Cloud storage configuration

**Currently configured remotes:**

Name	Type	Expiry	Action
No remotes have been configured.			

The table above lists the currently configured remotes in the Rclone config.

Deleting a remote only removes the configured Rclone remote, while keeping the remote's access token valid as it may be used outside of the web interface. To revoke the access token too, you can use the revoke button for the remote, which will revoke the access token if possible and delete the remote.

New Rclone remotes can be created below. After configuring a remote, you will be able to access it in the Files dropdown in the navigation menu. When transferring files in the file browser, keep the browser tab with the transfer open to ensure that it completes successfully. For more details on using the file browser, check the [LUMI user guide](#).

**Configure new remotes:**

LUMI-O

Generate an S3 token for LUMI-O. If an S3 token already exists for the project, it will be reused and refreshed. Currently, the S3 access tokens are valid for 7 days from creation or refresh.

Check the [LUMI user guide](#) for more details about LUMI-O.

Project

# Credential management via OOD (3)

The table above lists the currently configured remotes in the Rclone config.

Deleting a remote only removes the configured Rclone remote, while keeping the remote's access token valid as it may be used outside of the web interface. To revoke the access token too, you can use the revoke button for the remote, which will revoke the access token if possible and delete the remote.

New Rclone remotes for can be created below. After configuring a remote, you will be able to access it in the Files dropdown in the navigation menu. When transferring files in the file browser, keep the browser tab with the transfer open to ensure that it completes successfully. For more details on using the file browser, check the [LUMI user guide](#).

**Configure new remotes:**

LUMI-O

Generate an S3 token for LUMI-O. If an S3 token already exists for the project, it will be reused and refreshed. Currently, the S3 access tokens are valid for 7 days from creation or refresh.

Check the [LUMI user guide](#) for more details about LUMI-O.

Project

465000095

Generate s3cmd configuration ?

Configure public remote ?

Submit Add remotes for all projects

Select project  
Configure for s3cmd  
Configure public endpoint also

LUMI web interface: Release 6



# Credential management via OOD (4)

**Cloud storage configuration**

**Success**  
Successfully configured lumi-465000095-private, lumi-465000095-public remotes.

**Currently configured remotes:**

Name	Type	Expiry	Action
lumi-465000095-private	s3	2024-11-29 10:28:42 CET	<button>Delete</button> <button>Revoke</button>
lumi-465000095-public	s3	2024-11-29 10:28:42 CET	<button>Delete</button> <button>Revoke</button>

The table above lists the currently configured remotes in the Rclone config.

Deleting a remote only removes the configured Rclone remote, while keeping the remote's access token valid as it may be used outside of the web interface. To revoke the access token too, you can use the revoke button for the remote, which will revoke the access token if possible and delete the remote.

New Rclone remotes for can be created below. After configuring a remote, you will be able to access it in the Files dropdown in the navigation menu. When transferring files in the file browser, keep the browser tab with the transfer open to ensure that it completes successfully. For more details on using the file browser, check the [LUMI user guide](#).

**Configure new remotes**

Overview

# Credential management via OOD (5)

The screenshot shows a web browser window with two tabs: 'Dashboard - LUMI.csc.fi' and 'LUMI-O Credentials Manager'. The address bar shows the URL 'auth.lumidata.eu/projects/465000095'. The page has a navigation bar with 'Help & Support', 'Documentation', and 'Logout'. The main content is divided into two sections: 'Your projects' and 'Key management'.

## Your projects

Project number	Project description
465000095	VLAAMS SUPERCOMPUTER CENTRUM / VSC-SUPPORT
465000844	VLAAMS SUPERCOMPUTER CENTRUM / VSC-2023-04-TI59-KH-EASYBUILD
465001361	LUST Training / 2024-10-14-18 Hackathon: Optimizing for AMD GPUs 2024
465001362	LUST Training / 2024-10-28-31 Advanced LUMI course Amsterdam

## Key management

Duration (hours)\*:

Key description\*:

[Generate key](#)

### Available keys

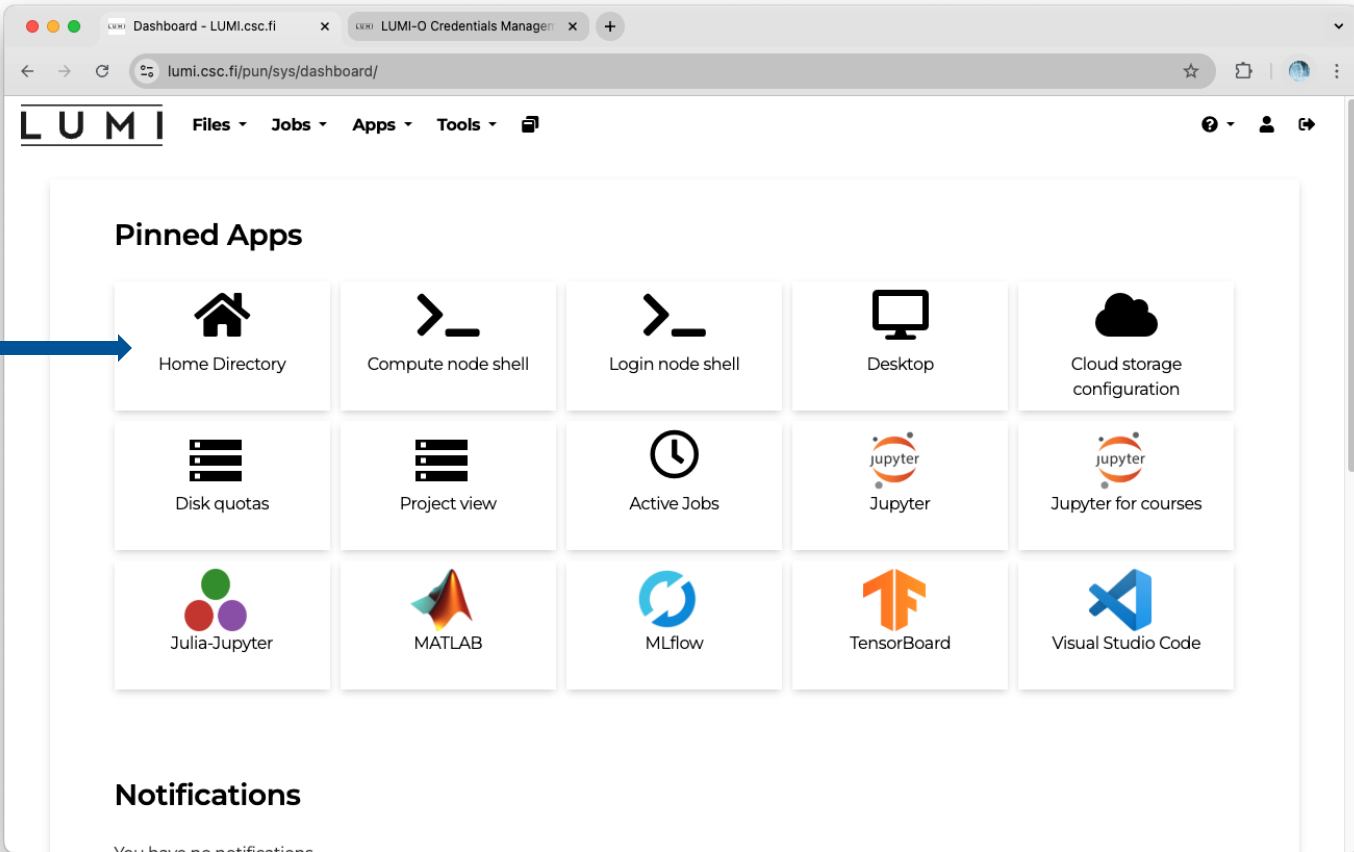
- Access key**  
7CGT5UP049SKVXLDJYI3  
Laptop  
Expires on: Nov 29 2024 10:31:35 GMT+0100
- Access key**  
Q35E9R02NLTPDWMGUVJC  
lumi web interface  
Expires on: Nov 29 2024 10:28:42 GMT+0100

**Expired keys**

The are no expired keys to show.

A blue arrow points to the second 'Access key' entry in the 'Available keys' list.

# Browsing via OOD (1)



The screenshot shows a web browser window with two tabs: "Dashboard - LUMI.csc.fi" and "LUMI-O Credentials Manager". The address bar shows the URL "lumi.csc.fi/pun/sys/dashboard/". The dashboard header includes the LUMI logo and navigation menus for "Files", "Jobs", "Apps", and "Tools".

The main content area is titled "Pinned Apps" and displays a grid of application tiles. A blue arrow points to the "Home Directory" tile. The tiles are:

- Home Directory (house icon)
- Compute node shell (terminal icon)
- Login node shell (terminal icon)
- Desktop (monitor icon)
- Cloud storage configuration (cloud icon)
- Disk quotas (server rack icon)
- Project view (server rack icon)
- Active Jobs (clock icon)
- Jupyter (jupyter logo)
- Jupyter for courses (jupyter logo)
- Julia-Jupyter (Julia logo)
- MATLAB (MATLAB logo)
- MLflow (MLflow logo)
- TensorBoard (TensorBoard logo)
- Visual Studio Code (Visual Studio Code logo)

At the bottom, there is a "Notifications" section with the text "You have no notifications".

# Browsing via OOD (2)

The screenshot shows a web browser window with the URL `lumi.csc.fi/pun/sys/dashboard/files/lumi-465000095-public/`. The interface includes a top navigation bar with buttons for 'Open in Terminal', 'Refresh', 'New File', 'New Directory', 'Upload', 'Download', 'Copy/Move', and 'Delete'. A left sidebar displays a 'Home Directory' tree with various project and scratch directories, and two buckets: 'lumi-465000095-private' and 'lumi-465000095-public'. The main area shows the current directory 'lumi-465000095-public/' with a 'Change directory' button and a 'Copy path' button. Below this, there are checkboxes for 'Show Owner/Mode' and 'Show Dotfiles', and a 'Filter' input field. A table lists the contents of the directory:

<input type="checkbox"/>	Type	Name	Size	Modified at
<input type="checkbox"/>	Folder	hpe	-	2-10-2024 13:50:31
<input type="checkbox"/>	Folder	mytest	-	24-5-2023 13:09:08
<input type="checkbox"/>	Folder	mytest2	-	24-5-2023 13:11:13
<input type="checkbox"/>	Folder	training-materials-web	-	31-10-2023 16:40:48

A blue arrow points from the 'lumi-465000095-public' bucket in the sidebar to the 'training-materials-web' directory in the table. A blue box labeled 'Buckets' has an arrow pointing to the 'training-materials-web' directory.

# Browsing via OOD (3)

The screenshot shows a web browser window with the following elements:

- Address Bar:** Contains the URL `lumi.csc.fi/pun/sys/dashboard/`. Two blue callouts labeled "End point" and "Bucket" point to the domain and path parts of the URL, respectively.
- Navigation Bar:** Includes buttons for "Open in Terminal", "Refresh", "New File", "New Directory", "Upload", "Download", "Copy/Move", and "Delete".
- Sidebar:** Lists a "Home Directory" and several project folders under paths like `/projappl/project_465001362` and `/scratch/project_46500095`. At the bottom, it shows two buckets: `lumi-46500095-private` and `lumi-46500095-public`.
- Main Content Area:** Shows the current directory path `lumi-46500095-public:/ training-materials-web/`. Below this is a table of files and folders.

Type	Name	Size	Modified at
Folder	intro-evolving	-	22-11-2024 10:56:35

A blue callout labeled "Pseudo-folder, not an object or bucket" points to the "intro-evolving" folder in the table.

# Browsing via OOD (4)

The screenshot shows a web browser window with the URL `lumi.csc.fi/pun/sys/dashboard/`. The interface includes a navigation sidebar on the left, a top toolbar with actions like 'Open in Terminal', 'Refresh', 'New File', 'New Directory', 'Upload', 'Download', 'Copy/Move', and 'Delete', and a main content area displaying a file listing.

Annotations in blue boxes and arrows point to the following elements:

- End point**: Points to the URL `lumi.csc.fi/pun/sys/dashboard/`.
- Bucket**: Points to the path `lumi-465000095-public/`.
- intro-evolving/files/**: Points to the current directory path.
- intro-evolving/files/exercises-evolving.tar**: Points to the specific file in the listing.
- Not the object name**: A callout box pointing to the file name `exercises-evolving.tar` in the table.

Type	Name	Size	Modified at
	<code>exercises-evolving.tar</code>	61.44 kB	18-11-2024 18:20:49
	<code>LUMI-BE-Intro-evolving-01-Architecture.pdf</code>	6.44 MB	14-5-2024 10:27:55
	<code>LUMI-BE-Intro-evolving-02-CPE.pdf</code>	1.91 MB	14-5-2024 10:27:44
	<code>LUMI-BE-Intro-evolving-03-Access.pdf</code>	4.92 MB	14-5-2024 12:59:06
	<code>LUMI-BE-Intro-evolving-04-Modules.pdf</code>	3.98 MB	25-10-2024 17:31:53
	<code>LUMI-BE-Intro-evolving-05-SoftwareStacks.pdf</code>	2.59 MB	25-10-2024 17:33:19
	<code>LUMI-BE-Intro-evolving-06-Slurm.pdf</code>	2.08 MB	10-7-2024 16:47:15
	<code>LUMI-BE-Intro-evolving-06-Support.pdf</code>	2.29 MB	14-5-2024 10:26:17

# CLI tool configuration on LUMI: lumio-conf

- On LUMI, you can use `lumio-conf` to configure `rclone`, `s3cmd`, `aws` and `boto3`
  - Need to load the `lumio` module which also provides `rclone`, `s3cmd` and `restic`
  - Will ask for data from the “Access key details” screen
  - Default is to create configurations for `rclone` and `s3cmd` .  
Try `lumio-conf -h` to see other options.
- Generate the configuration snippets via the web interface
  - E.g., for `rclone`: copy manually to `~/.config/rclone/rclone.conf`
  - Can be used to configure tools on your computer also.

# CLI tool configuration on LUMI: lumio-conf, rclone and s3cmd

- The `rclone` configuration contains two endpoints
  - `lumio-46YXXXXXX-private`: Private buckets and objects
  - `lumio-46YXXXXXX-public`: Public buckets and objects

Note that this was different with an earlier version of the tool

- The `s3cmd` config file can only contain data for one project.  
Solution implemented by `lumio-conf`:
  - Store information for project `46YXXXXXX` in `~/ .s3cfg-lumio-46YXXXXXX`
  - Also overwrite `~/ .s3cfg` with the same data
  - So `s3cmd` will by default use the project from `~/ .s3cfg` but if a user has active configurations for multiple projects, the `-c` flag can be used to point `s3cmd` to the right configuration file.



# Bucket and object names

- Bucket name
  - Unique name within a project
  - Must be between 3 and 63 characters long, lowercase letters, numbers, hyphens and dots, must not contain uppercase characters or underscores
  - Must start with a lowercase letter or number
  - When dots are used, each part of the name that is separated by the dots, is also called a label (but there is also a different concept of label in Ceph).
- Object name
  - Unique within a bucket
  - Technically, any UTF-8 string between 1 and 1024 bytes, but client software might add more restrictions
  - Common practice is to implement a folder-like structure using slashes in the name.
  - But creating the folder view is expensive: List all objects in a bucket and sort

# Policies and ACLs

- Access control is managed through bucket policies and bucket and object access control lists (ACLs)
- Policies is a very powerful but also hard to use mechanism
  - Some information in [the "Advanced usage of LUMI-O" section of the docs](#)
  - And there is also [information in the Ceph manual](#)
  - Can be managed through `s3cmd`
- ACLs apply to individual buckets and objects
  - Can only add rights
  - Useful to make a bucket or object public, or give access to another project, but this is done to individual objects (unless applied recursively)
  - This is what rclone uses when uploading to the public or private end points in the configurations created with the various tools for LUMI-O

# Policies and ACLs

## Examples

- Make a bucket and all objects in it public or private

```
s3cmd setacl --recursive --acl-public s3://bucket/
```

```
s3cmd setacl --recursive --acl-private s3://bucket/
```

- Grant or revoke read rights to a bucket

```
s3cmd setacl --acl-grant='read:46YXXXXXX$46YXXXXXX' s3://bucket
```

```
s3cmd setacl --acl-revoke='read:46YXXXXXX$46YXXXXXX' s3://bucket
```

- Note the use of single quotes to make sure that `$46YXXXXXX` is not interpreted as a variable name!
- And similarly to objects
- Check the ACL and other information of a bucket or object

```
s3cmd info s3://2day-20241210
```

```
s3cmd info s3://2day-20241210/img/LUMI-2day-20241210-10-ObjectStorage/Title.png
```

# Sharing data

- Public buckets and objects can be read by anyone in the world, even simply via a web browser.
- It is possible to grant specific projects access to buckets and objects.
  - See the previous slide for setting ACLs
  - Access: e.g., project "46BAAAAAA" can list a bucket from project "46YXXXXXX" (assuming the latter has given sufficient rights to the former) with: 

```
s3cmd ls s3://46YXXXXXX:bucket/
```

```
s3cmd ls --recursive s3://46YXXXXXX:bucket/
```

```
rclone ls lumi-46BAAAAAA-private:"46YXXXXXX:bucket"
```
- Presigned URLs: Next slide
- Invite the other to the project
  - But it is not possible to add a user who came in via puhuri to a 462-project

# Sharing data: Presigned URLs

- Can be generated using `rclone`
- Access to anyone who gets the link without further authentication, for the validity period of the link
- Presigned URLs depend on the authentication key that was used to create them.
  - Link expires when the key expires or is revoked, even if the validity period of the link has not expired.
  - Maximum lifespan of a link is limited to 7 days on LUMI-O.
- Create: with `rclone link`:  
`rclone link lumi-46YXXXXXX-private:bucket/object`  
`rclone link --expire 2d lumi-46YXXXXXX-private:bucket/object`

# Some tips & tricks

- When using the `rclone` command line tool, it is possible to throttle the speed for many commands with the `--bwlimit` command line option
  - May be needed if you upload from home over a very bandwidth-limited connection
- Sharing data from project A with project B does not protect the data from being deleted when project A ends, even if project B is still valid.
  - End of project: 90-days grace period, data read-only
  - End of grace period: Data queued for deletion

**Questions?**

