



RDM ADVANCED

STORING SOLUTIONS



WHAT TO CONSIDER FIRST?

- How much storage space is needed (is the "scale" GiB, TiB or PiB)?
- What type of data is stored and is it sensitive?
- How long the data needs to be stored?
- What software is needed to process the data?
- Who needs access to the data and is sharing it outside UH necessary?
- What is the available budget?
- Consider the whole lifecycle of the data, including long term plans





HOME AND GROUP FOLDERS

	Home Folder ("Z-drive")	Group Folder ("P-drive")
<u>Capacity</u>	100 GiB	10 GiB - 10 TiB
<u>Operating System</u>	Win, Mac, Linux (UH-Linux also has a separate network home directory)	Win, Mac, Linux
<u>Access</u>	Only the UH-account that owns the folder	A group of UH-accounts
<u>Backups</u>	Hourly	Hourly
<u>Physical location</u>	UH premises	UH premises
<u>How to obtain</u>	All users have one by default.	UH employees can order by self-service. The owner of the group can control access by using IAM (Identity and Access Management) –groups.



DEPT-SERVICE ("L-DRIVE")

- Functions as the group folder, but ordered for special purposes
- Bigger investment, usually by joint purchase of research groups or faculty
- Total quota can be shared to smaller units as customer wishes
- If backup is required, also double capacity is needed
- Access control with IAM-groups
- Regarding storage of sensitive data, follow the same principles as with home and group folders

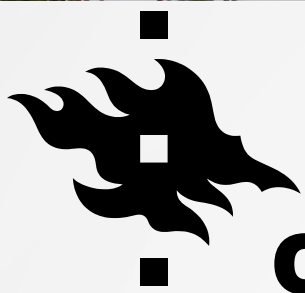




DATA CLOUD

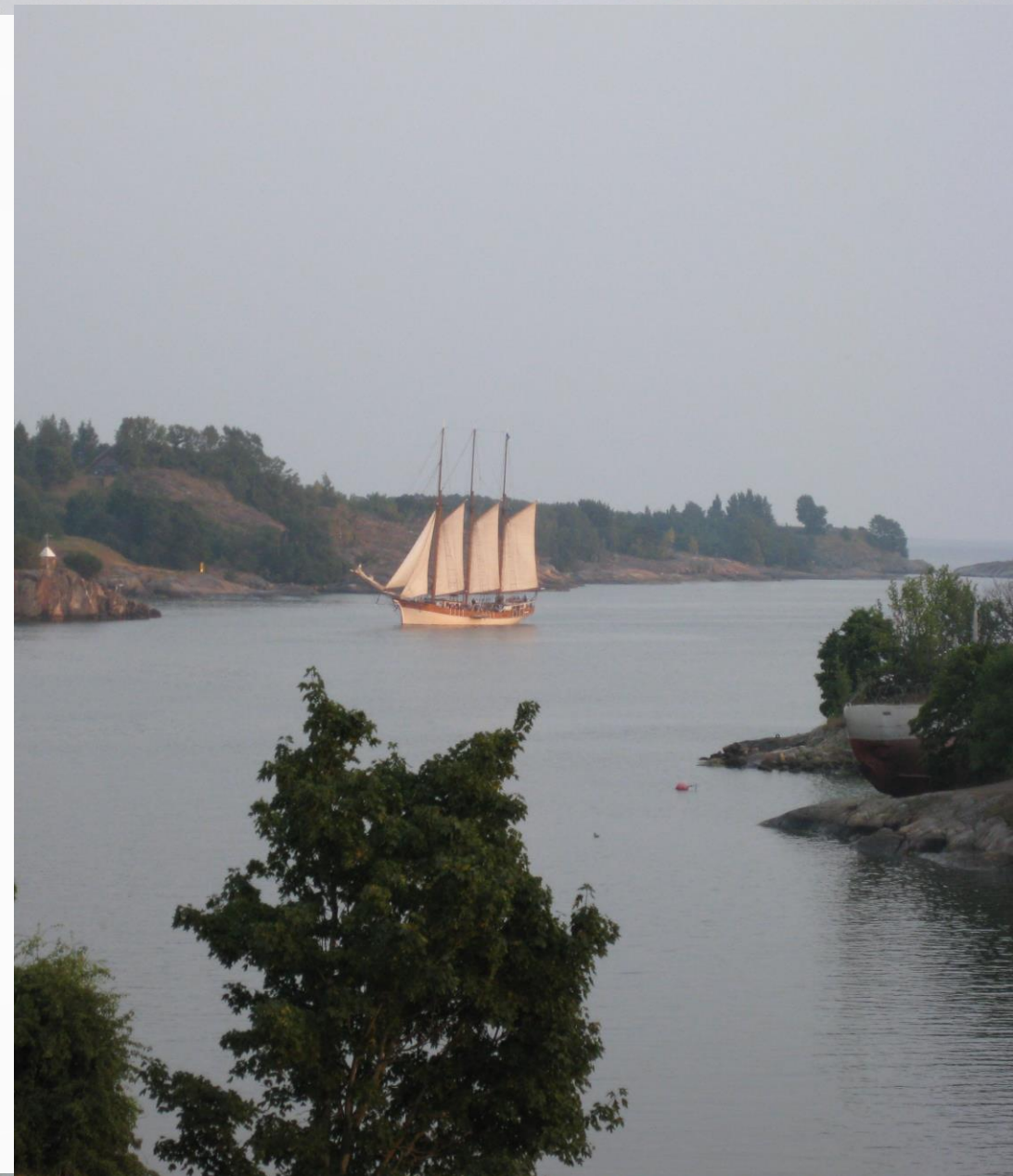


- S3 Object Storage in UH premises
- Can be accessed using Nextcloud client, web browser or command line tools
- Data cannot be modified "online" in Datacloud, but must be downloaded to a computer
- Users can purchase quota in 50 TiB bundles
- Version control, no backup
- Not for sensitive data
- Budget option for large data
- [Instructions](#)



ONEDRIVE AND TEAMS

- Microsoft Office 365 cloud services
- Onedrive storage space is limited to 1 TiB of storage space, in Teams it is limited to 25 TiB
- Onedrive is suitable for storing and sharing non-sensitive material
- Teams is also suitable for collaboration with non-UH members
- Teams can be connected with several applications for project or group management (Trello, Planner, etc.)
- Microsoft cloud storage includes version control but no backups
- Suitability for GDPR regulated data should be done with risk assessment and privacy notice





VDI (VIRTUAL DESKTOP INFRASTRUCTURE)



- Different types of virtual desktops for customer needs (Windows/Linux, data sensitivity, GPU resources)
- Can be used to access UH services from anywhere by connecting to a virtual desktop in a web browser
- Various scientific software are pre-installed in the virtual desktops
- Data must be stored to network storage (home folder, group folder etc.) when using virtual desktops
- Students and non-employees need a license upgrade in order to use virtual Windows desktops outside of university campuses
- It is also possible to acquire custom VDI desktops



■ HIGH PERFORMANCE COMPUTING (HPC)

Turso

- HPC infrastructures of UH
- Combination of Kale, Ukko2, Vorna etc. clusters
- For batch and interactive use
- >7000 CPU cores and 28 GPU's
- >2 PiB storage space (Vakka (in Kumpula) and Kappa (in Viikki))
- Infiniband storage solution (analyze only data which is stored locally in the HPC cluster)
- Parallel computing (not suitable for all needs)

Mahti, Puhti & Lumi

- HPC services from CSC (IT-Center for Science Ltd.)
- Similar functionalities as UH Turso but scales up higher
- Users must apply for capacity





CSC ALLAS

- Quite similar to UH Datacloud
- Capacity up to 200 TiB
- Service is generally available for projects and data should be moved elsewhere after the project
- Data can be transferred to CSC Computing environments
- Several ways to access (cPouta, SD Connect, command line tools, etc)
- Users must have CSC accounts



CSC IDA

- Nextcloud-based service for researchers
- Can be accessed by web browser or command line tools
- Data must be downloaded for editing; no online-edit possible
- Staging area and frozen area: data can be edited until it's frozen
- Etsin-data can be linked to frozen IDA-datasets, but it is also linked to the owner's account and expires with it
- If data is approved for long-term preservation, in the future it may be transferred from IDA to CSC long term storage system
- Users must have CSC accounts

