KU LEUVEN

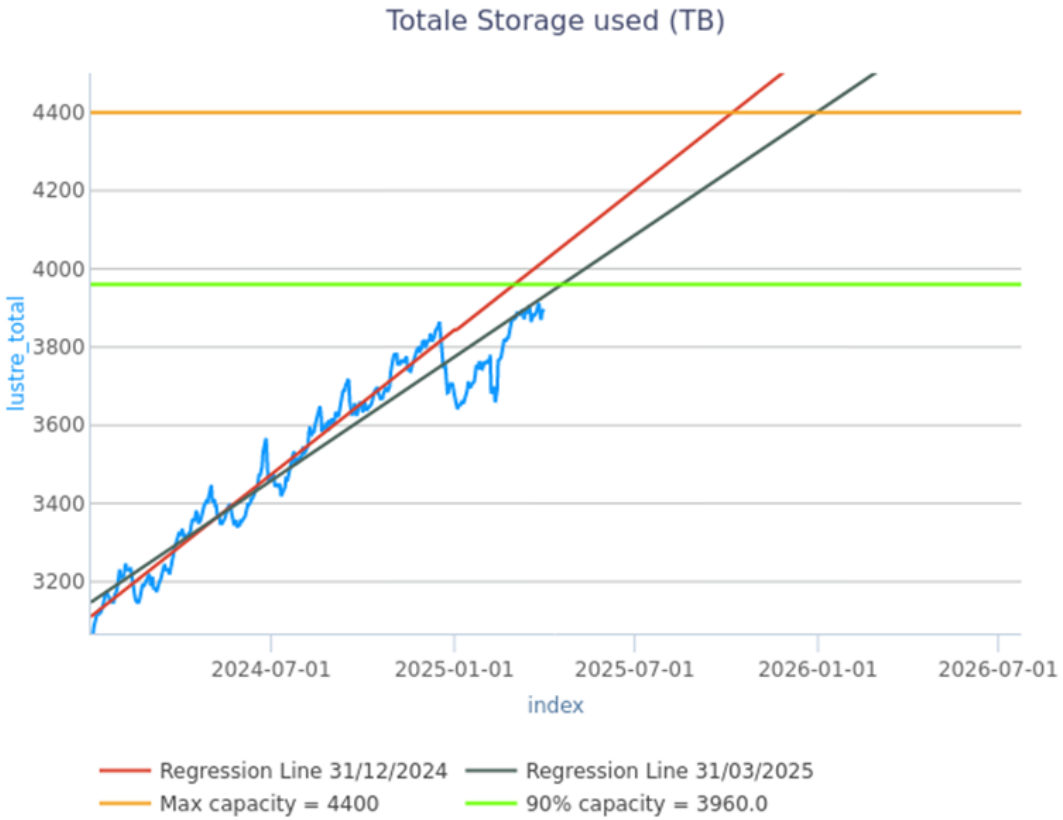
Tier-2 data management

From ingest to outflow

Wouter Lampaert and Jef Scheepers

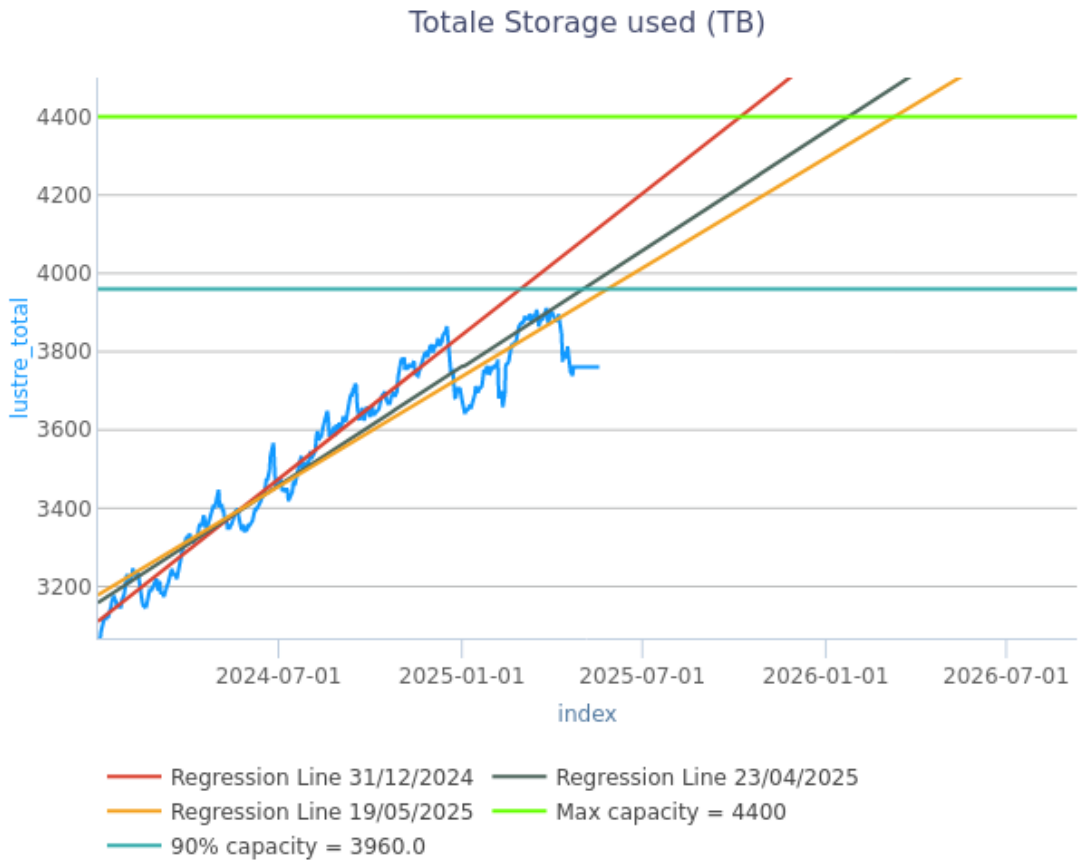
The issue

- HPC integrated storage is expensive and limited



- Moving/deleting unused/old data is essential for
  - System health
  - Useability for all researchers
- Cluster storage capacity will not grow at the same speed as data production

Updated graph



Solutions

Limitations:

- Stricter on staging increases and requests
- Inode quota

Long-term:

- Motivate outflow
  - Large request will require outflow strategy!
- ‘elevated permissions’ for moderators of staging directory
  - Using ACLs
- Extended reporting to staging owners

Overview

1. Choosing the right types of storage
2. Data outflow to external storage: tools
3. Data management plan
4. Structuring your files and directories
5. Keeping track of your data
6. Introduction to Tier-1 Data

Choosing the right types of storage

You'll need two types of storage:

- HPC-integrated storage
- external storage

Choosing the right types of storage: HPC-integrated storage

- $VSC_HOME, $VSC_DATA
- Compute storage
  - Not safe for long-term storage (no double copy, no snapshot)
  - $VSC_SCRATCH
    - For individual projects
    - Recommended for most
    - Automatically cleaned
    - Benefit of a free increase
  - Staging/project storage
    - For sharing data
      - Shared input data, enrichment databases
      - Postprocessing on the same output data
      - Software in some cases
    - Data that cannot be automatically cleaned
      - E.g. input data that needs to be read on an irregular basis
    - Cost based on maximal usage every month

Choosing the right types of storage: external storage

- External storage: for data that is not actively being computed on, could be:
  - ‘in-house’ managed storage
  - K-drive (archive) or L-drive (large volume)
  - ManGO or Tier-1 data
  - Cold storage => KU Leuven pilot project - FriGO
  - Other external solutions ….
- Previously archive storage solution will be phased out
  - Summer 2026

Data outflow to external storage: tools

- Multiple tools available:
  - Globus (and Globus Flows)
  - iCommands/Python iRODS client (for Tier-1 Data/ManGO)
  - Classic command line tools
    - rsync
    - sftp
  - GUI applications
    - FileZilla

Data management plan

- Create a data management plan for your project with your team:
  - Consult the official RDM data management guidelines
  - Create this with your whole team
    - PI for view on past, present and future projects
    - Researchers for day-to-day HPC usage
    - Data support staff: technical knowledge
      - Contact us for HPC related data questions or have a look at: https://www.kuleuven.be/rdm/en for broader support.

Data management plan: HPC-specific suggestions

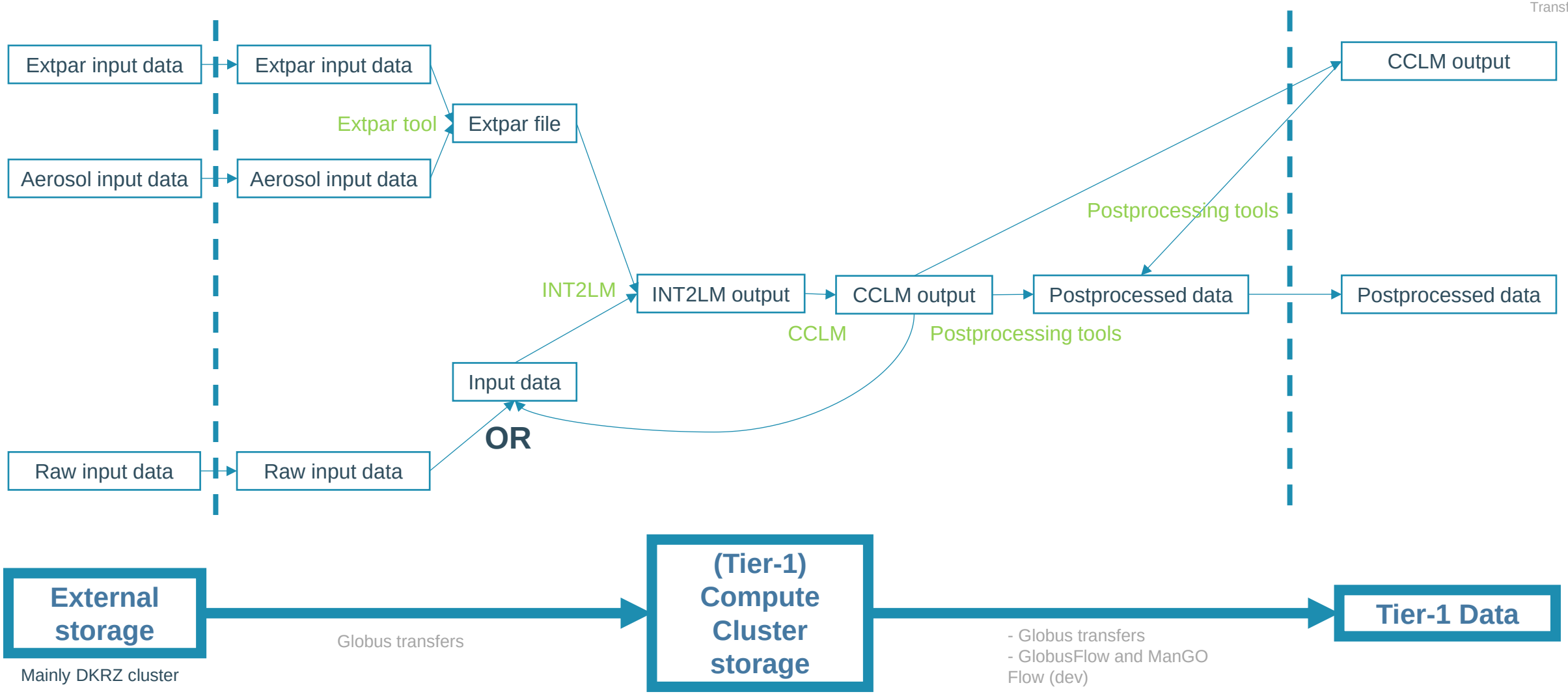
- Map your dataflow (across all cluster/storage platforms)
- Estimate storage capacity
- Identify shared resources
- Identify automation opportunities

Data management plan: HPC-specific suggestions

- Map your dataflow (across all cluster/storage platforms)
- Estimate storage capacity
- Identify shared resources
- Identify automation opportunities

Example data/workflow

Legend:
Data types
Software
Storage locations
Transfer tools



Data management plan: HPC-specific suggestions

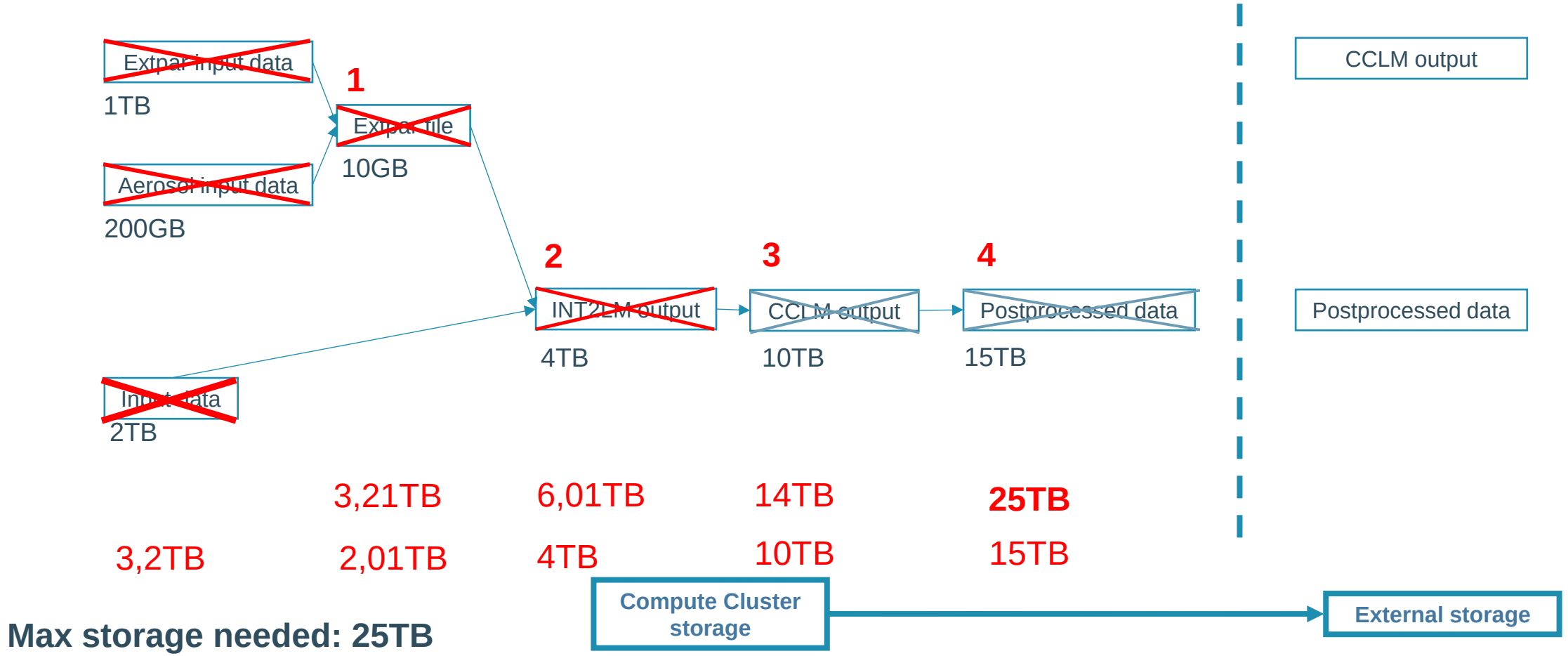
- Map your dataflow (across all cluster/storage platforms)
- Estimate storage capacity
- Identify shared resources
- Identify automation opportunities

Storage estimations: calculate your total usage

Table 5 *Overview of planned simulations and estimated core-hour calculation* (* EUR-11 runs (RCM) are at 12.5 km over the European domain, while CPM, convective permitting simulations are at 2.8 km over the Belgian domain, ** some periods for global warming levels (GWLs) are overlapping, therefore limited time periods)

Computational task: GCM downscaling			number of simulated years (incl spin up)	Core-hour calculation							Storage volume estimate
driving GCM	scenario	domain & GWL*, **		Number of such jobs	Wall clock time (in hours) per job	Number of Tier-1 nodes per job	Number of Tier-1 cores per node per job	Memory factor (memory-per-core (MB) / default-memory-per-core*)	Total core-hours per task*	Estimate of memory usage (GiB) per node per job	OpenMP / MPI / OpenMP + MPI	Tier-1 SCRATCH volume (TiB)
EC-EARTH3-Veg	Historical	EUR-11	23	3	72	8	128	1970	221184	256	MPI	18.975
		CPM	21	7	72	8	128		516096			18.9
	SSP5-8.5	EUR 11	86	10	72	8	128		737280			70.95
		CPM GWL 2	14	5	72	8	128		368640			12.6
		CPM GWL 3	18	6	72	8	128		442368			16.2
	SSP2-4.5	EUR 11	85	10	72	8	128		737280			70.125
		CPM GWL 2	17	6	72	8	128		442368			15.3
		CPM GWL 3	21	7	72	8	128		516096			18.9
MIROC6	Historical	EUR-11	23	3	72	8	128	1970	221184	256	MPI	18.975
		CPM	21	7	72	8	128		516096			18.9
	SSP5-8.5	EUR 11	86	10	72	8	128		737280			70.95
		CPM GWL 2	17	6	72	8	128		442368			15.3
		CPM GWL 3	21	7	72	8	128		516096			18.9
	SSP3-7.0	EUR 11	85	10	72	8	128		737280			70.125
		CPM GWL 2	16	5	72	8	128		368640			14.4
		CPM GWL 3	21	7	72	8	128		516096			18.9
Total Needed									8036352			488.4

Storage estimations: calculate your max need



Max storage needed: 25TB

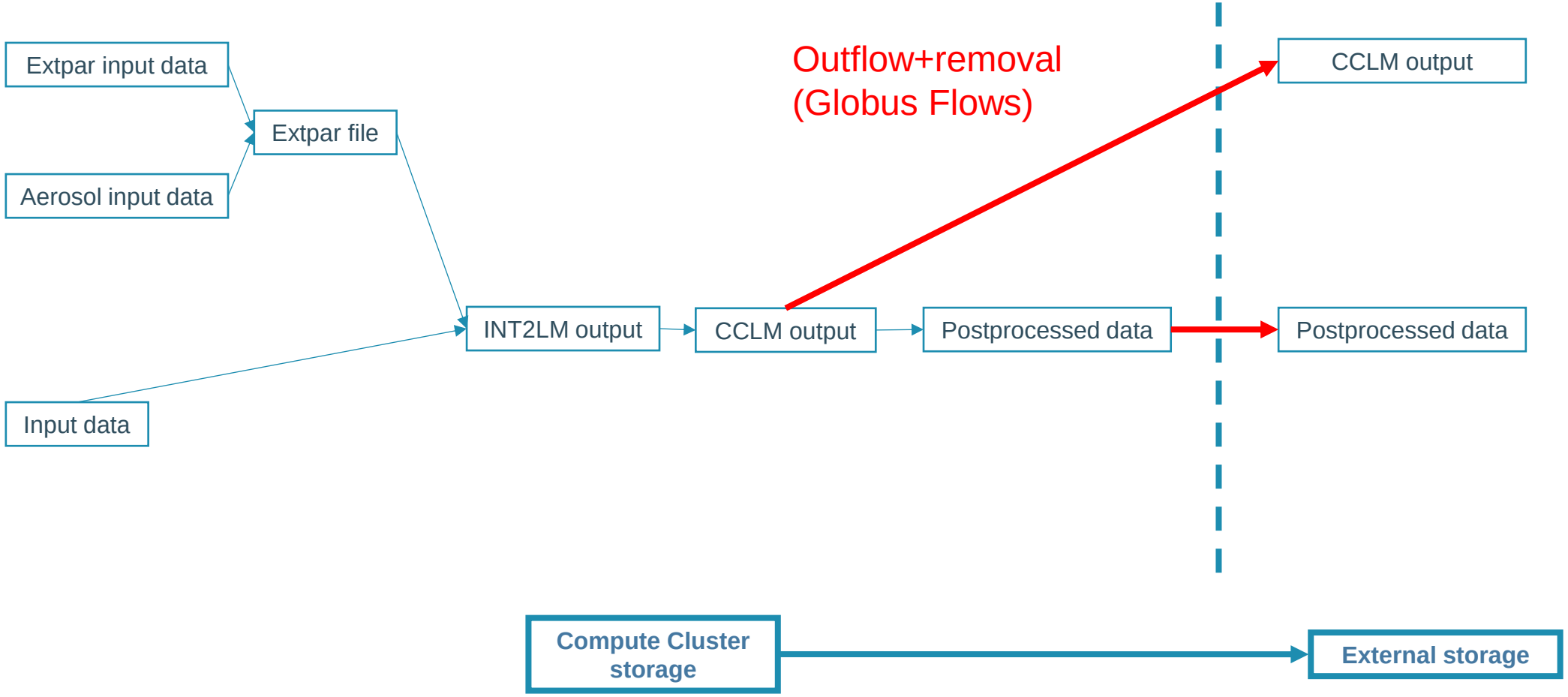
Data management plan: HPC-specific suggestions

- Map your dataflow (across all cluster/storage platforms)
- Estimate storage capacity
- Identify shared resources
  - Enrichment databases
  - Shared (input) data
  - Shared software
    - Consider centrally installed software! (modules)
- Identify automation opportunities

Data management plan: HPC-specific suggestions

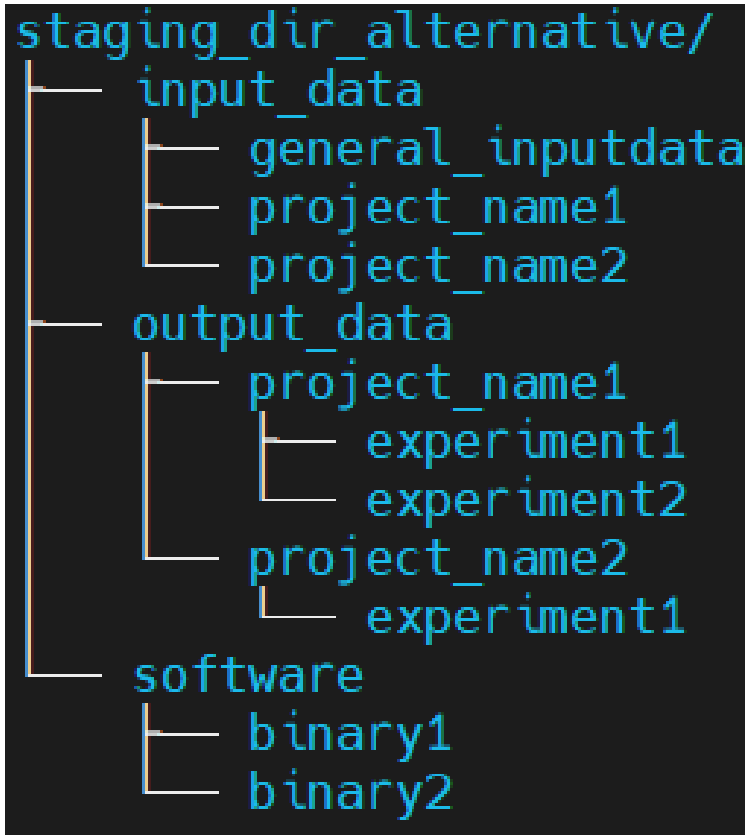
- Map your dataflow (across all cluster/storage platforms)
- Estimate storage capacity
- Identify shared resources
- Identify automation opportunities
  - Even if you cannot immediately automate them
  - Potential automation:
    - Preprocessing
    - Outflow + removal

Automation opportunities



Structuring your files and directories

- For shared storage: create guidelines within your team
- Structure should reflect your structure on your external storage
- Higher level directories contain broader topics
  - e.g. Projects (-> users) -> experiments
  - Consider splitting input, (intermediate) and output data
    - Depends on the level of sharing
  - Separate software from data
- Extra important when using sensitive data
  - Contact us



Data outflow to external storage

- What data needs to copy/move to your external storage
  - Data that is expensive:
    - Computation time
    - Human effort
  - Data that needs to be shared with externals
  - Data that needs to be stored temporarily
    - External storage as a bridge between different machines
    - Data that needs to be processed later, keeping more quota for current tasks
    - Public (input) data with large retrieving times (download, processing steps…)

Data outflow to external storage

- Data that should **NOT** be on your external storage
  - Source code: use version control systems (e.g. Git)
  - Compiled software
    - Store installation recipes and logs instead
  - Public databases (unless retrieving is expensive)
  - Test data, benchmarks…
    - Filter out unnecessary test scripts, test experiments…
      - Point to test status with directory prefixes/suffixes
        - test_
        - tmp_
        - …

Keeping track of your data

- Follow up on your team, for example in regular data meetings (or make it a part of the PI-researcher meetings)
  - Which input data is present on your staging
  - What experiments are finished and can/should be moved out
  - How far from reaching quota?
- Use available tools to monitor storage
  - myquota
  - du
  - duduckdb : new tool for staging directories
- Automate where possible!
  - Irods transfers
  - …
- Use metadata

Keeping track of your data: duduckdb

- Queries database that will be created on a regular basis
- Already available for all staging storages

```
$ duduckdb /data/leuven/public/staging-stats/<your-staging-dir>.parquet --human-readable
directory:              size            inodes
```

```
$ duduckdb /data/leuven/public/staging-stats/<your-staging-dir>.parquet --human --older-than=2023-01-01
direc
=====
:
subdi
subdi
subdi
```

```
$ duduckdb /data/leuven/public/staging-stats/<your-staging-dir>.parquet --per-user --human-readable
directory:              size            inodes
================================================================================
:                       3.0TiB          129.2Ki
--------------------------------------------------------------------------------
    vsc_id1:            2.8TiB          17.3Ki
    vsc_id2:            45.8GiB         111.9Ki
    vsc_id3:            12.0KiB         4.0
    vsc_id4:            16.0KiB         13.0
    vsc_id5:            159.0GiB        14.0
    vsc_id6:            4.0KiB          1.0
...
```

KU LEUVEN

Tier-1 Data

What is Tier-1 Data?

Management of **ACTIVE** research data

Based on open-source software iRODS

More than storage (metadata, automation, sharing)

Accessible via various clients

Free, application based

To help researchers **manage** their data according to best practices

Centralization & Accessibility

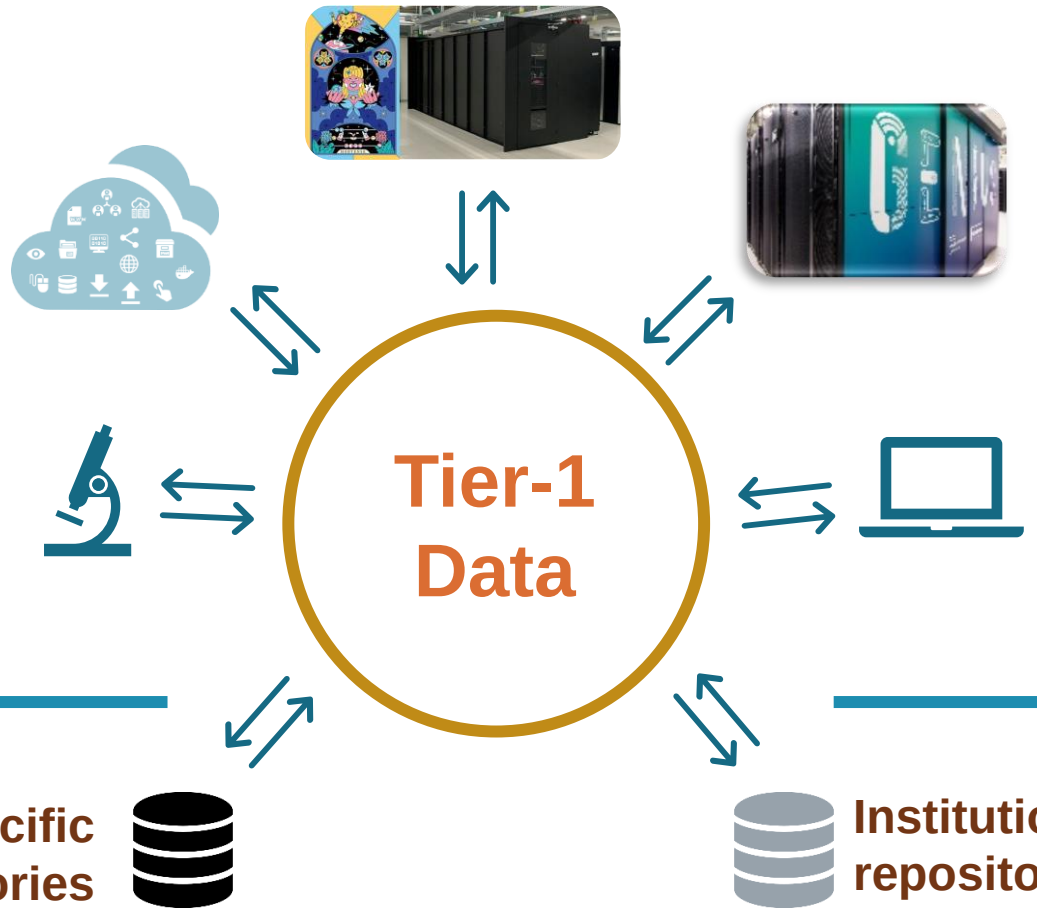
ManGO Portal

iCommands

PRC

Globus

SFTP



Active data

Inactive data

Regional / domain specific repositories

Institutional repositories

Permissions

Group-based

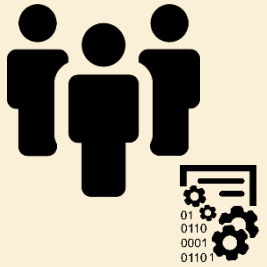
Access Levels:
1. Read
2. Write
3. Own

@ Objects
@ **Collections**

Data creators

Create data
Upload data

Raw Data



Data analysts

Search data
Analyze data
Upload results

Raw Data

Processed Data

Data consumers

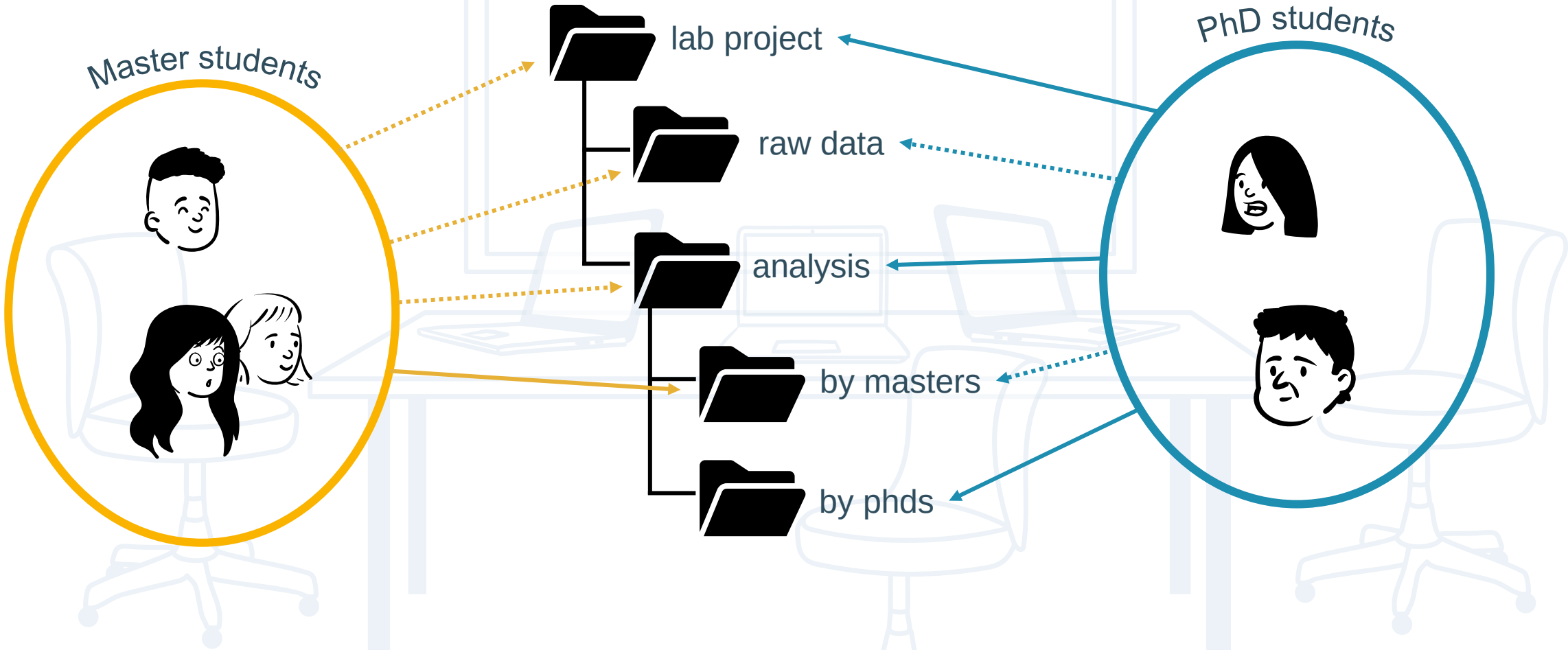
Search data
Read data
Download data

Processed Data

Write Access

Read Access

Permissions



Metadata: for what?

Give context, describe

FWO99-simul

ProjectID: FWO99
Title: Thermodynamics of coffee
PI: ProfessorZonnebloem
Funder: FWO
StartDate: 2021-01-01
EndDate: 2024-12-11
Status: ToArchive

SimulationID: FWO99-sim1
System: Caffeine
JobID: 40000001
System: VSC-Hortense
Software: CP2K
Version: 4.1
Date: 2023-04-20

Data discovery

- List collections/projects with FWO as founder
- List all collections associated to project FWO99
- Find all files that were generated with CP2K v4.1

Automation

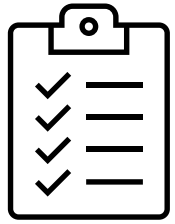
Create a script (e.g in python):

- Look for collections with Metadata ‘ToArchive’
- Make a read-only copy of the collection
- Add metadata: ArticleID=article1

Metadata: how?

- Manual addition

- Via schemas

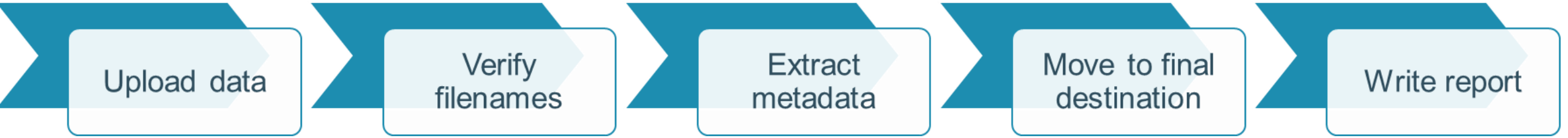


- Automated extraction
  - From file headers
  - From filenames
  - …

Automation

- Client-side automation:
  - Python client, iCommands

- Server-side automation:
  - Monitoring or event-based trigger
  - Implemented by support team



Tier-1 Data clients

iCommands

```
$ icd myPath
$ imkdir test
$ itouch test/myfile
```

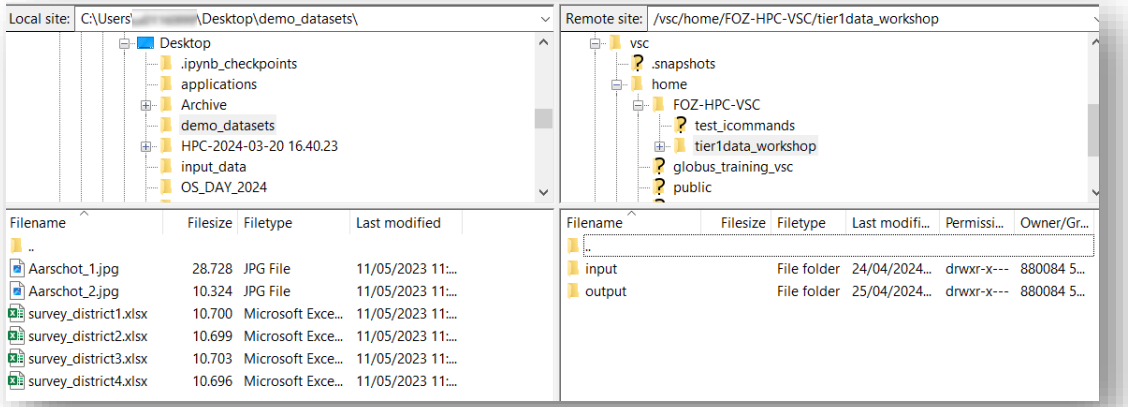
PRC

```python
# this file is used to query which data objects have no attached metadata
import os, os.path
import json
from irods.session import iRODSSession
from irods.column import Criterion
from irods.models import Collection, DataObject, DataObjectMeta

# Get environment file
env_file = os.getenv(
    "IRODS_ENVIRONMENT_FILE",
    os.path.expanduser("~/.irods/irods_environment_python_vsc.json"),
)

# Provide path to iRODs collection
iRODsPath = "/vsc/home/[redacted]/%"

# provide the metadata field attribute that we are looking for
MDname = "mgs.image_attributes.%"

# establish session
session = iRODSSession(irods_env_file=env_file)

# list which of these files do not have metadata attached from the specif
lMDtrue = []
qMDtrue = (
    session.query(Collection.name, DataObject.name, DataObjectMeta.name)
    .filter(Criterion("like", Collection.name, iRODsPath))
    .filter(Criterion("like", DataObject.name, "%.czi"))
    .filter(Criterion("like", DataObjectMeta.name, MDname))
)
for item in qMDtrue:
    lMDtrue.append(f"{item[Collection.name]}/{item[DataObject.name]}")
```

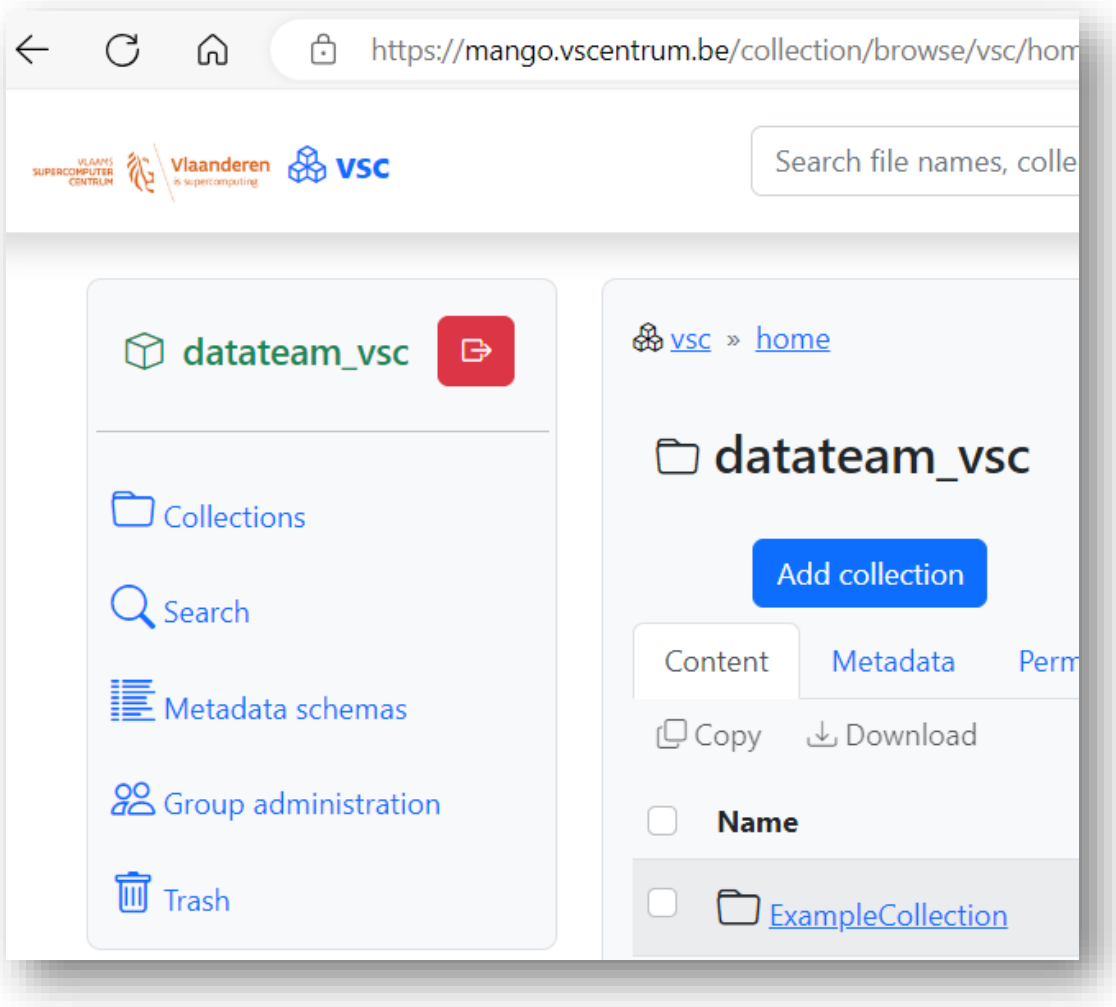
See upload tool **ManGO-Ingest** (https://github.com/kuleuven/mango-ingest)

Tier-1 Data clients

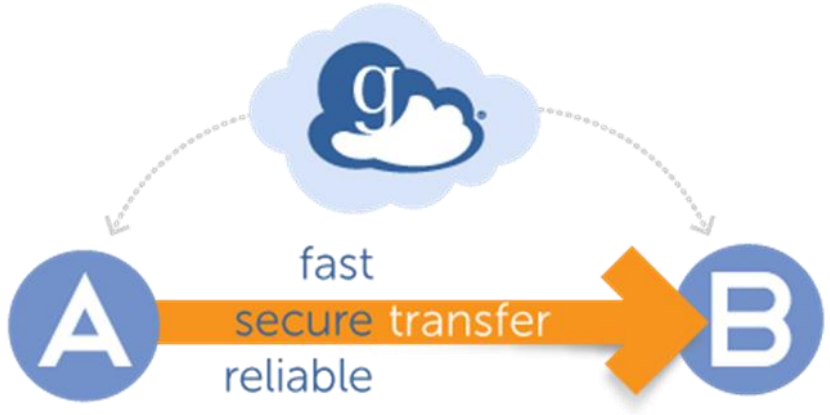
SFTP



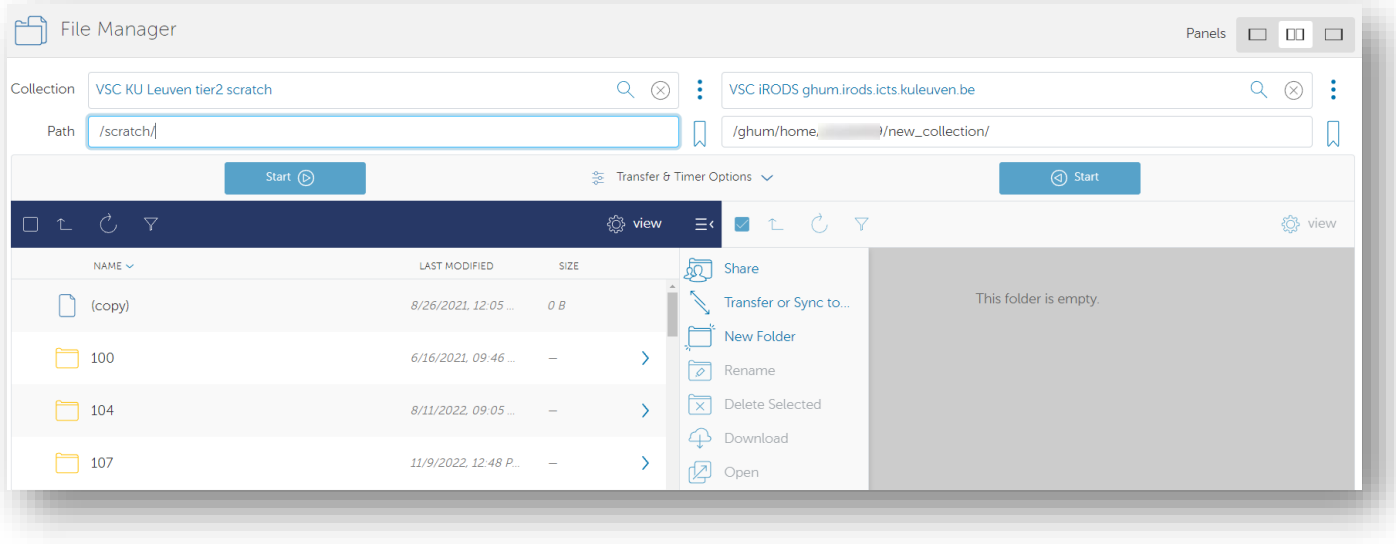
ManGO Portal



Globus



© Globus online



Exit strategies to cold storage

- Choose logical scope for your dataset
- Describe your dataset as a whole
- Export existing file/folder metadata

- Tier-1 Data will allow:
  - To export via multiple clients
  - To send data straight to selected repositories
  - To export metadata in conventional file formats

Practicalities: Get started & Get support

Requesting a Tier-1 Data project

- Project-based access
- Criteria:
  - Flemish Public Research Institution
  - Active data in the context of VSC: Tier-1/Tier-2 compute, Tier-1 Cloud
  - Judged on workflow feasibility and research data management aspects
- Types:
  - Starting grant: 5TB for 8 months
  - Full grant: 5TB-1PB
- Request at any time
- All info on https://www.vscentrum.be/data

User support

- Helpdesk: data@vscentrum.be
- Documentation: https://docs.vscentrum.be/data/tier1_data_service.html