ELIXIR CZ Annual Conference 2024

Onedata for comprehensive management
of distributed scientific data

Lukasz Opiola

10.10.2024

elixir CZECH REPUBLIC

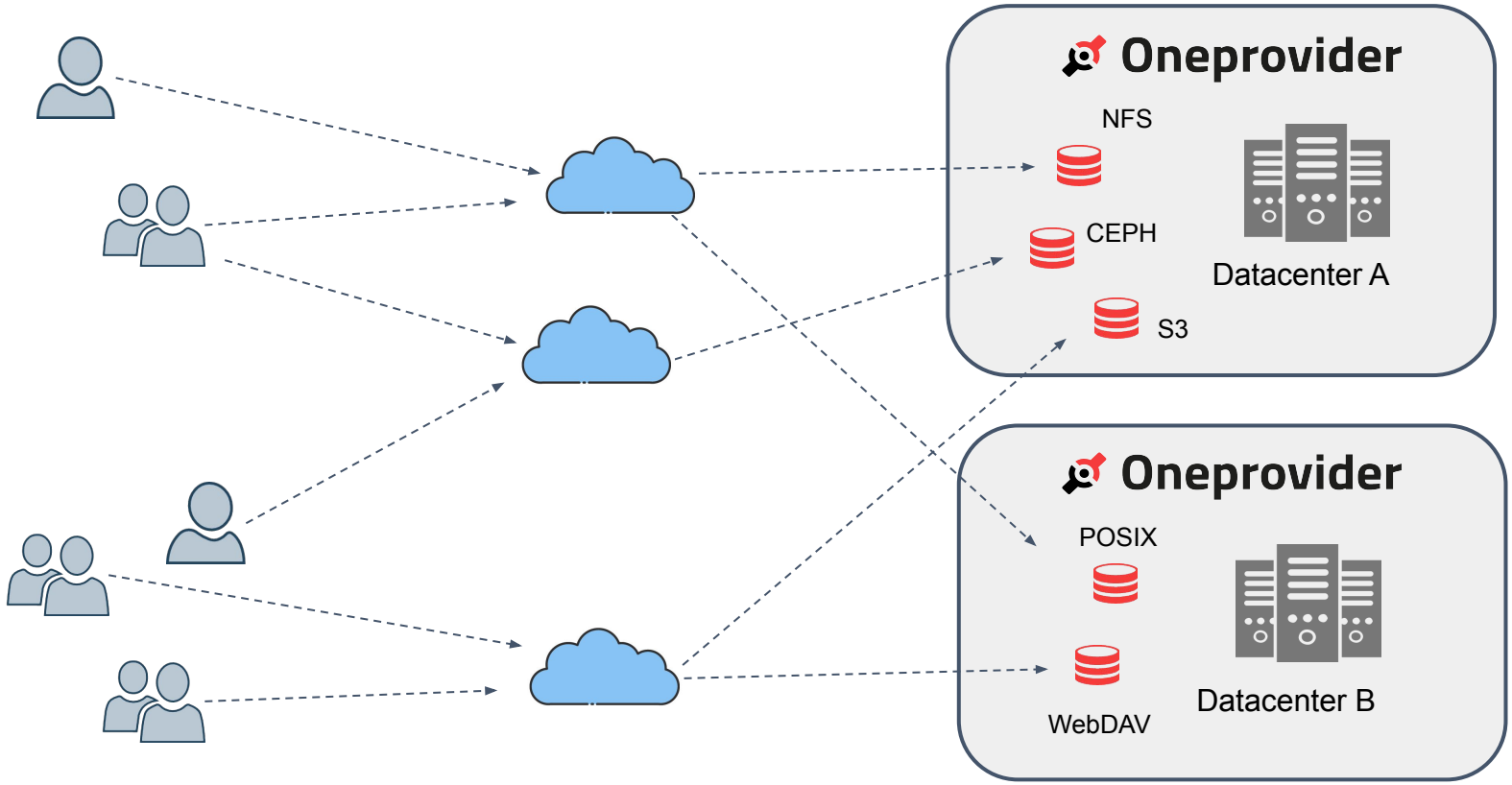
WHO WE ARE

ONEDATA

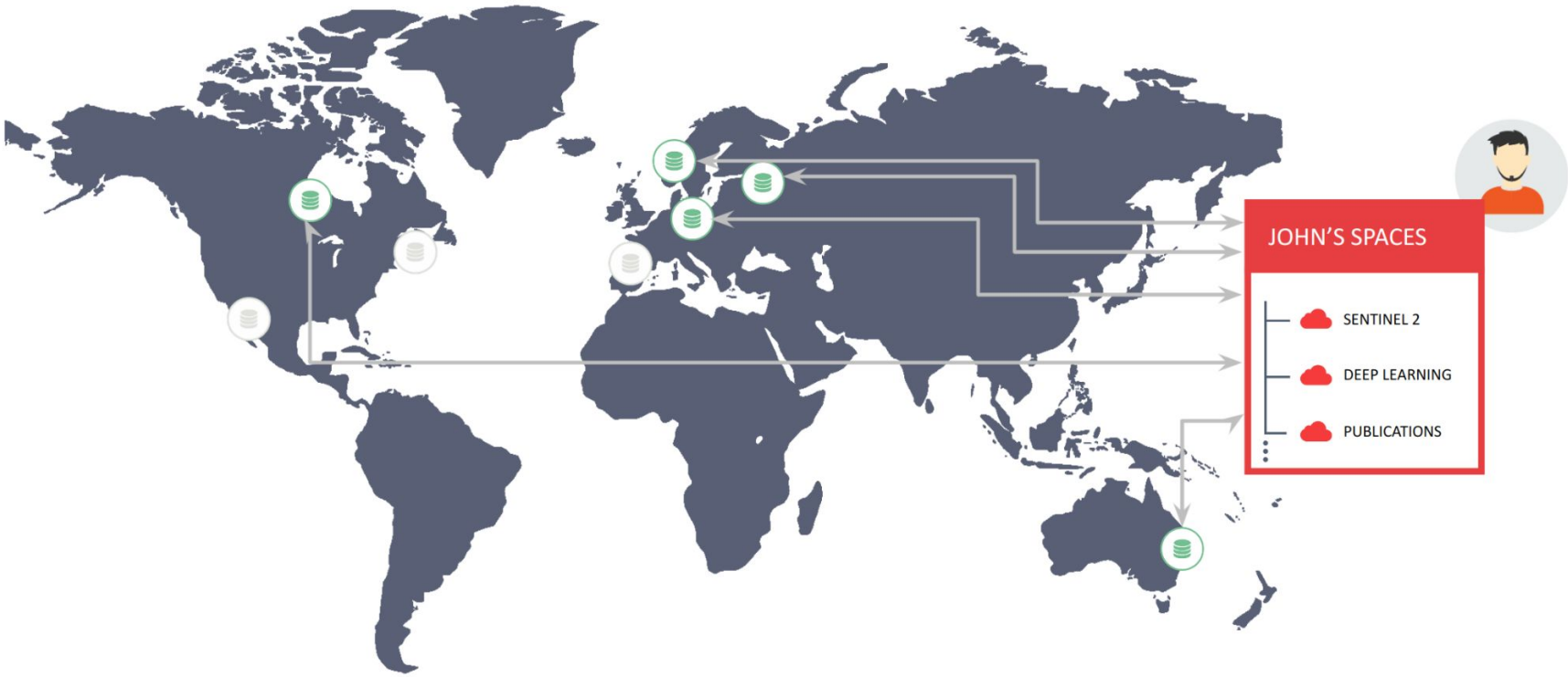
CYFRONET

- Open-source, 11 years of devoted development — github.com/onedata
- Developed @ **AGH University of Krakow** and **Cyfronet** supercomputing center, to:
  - deliver a **data management** platform for large-scale and **distributed** problems,
  - enable global collaborative data sharing **across organizational** domains,
  - streamline data processing in **heterogeneous** data storage setups and **hybrid cloud**.
- Our supporters and partnerships:

UNIFIED ACCESS TO GEOGRAPHICALLY AND ORGANIZATIONALLY DISTRIBUTED DATA



ONEDATA SPACES FOR DISTRIBUTED DATA



elixir CZECH REPUBLIC

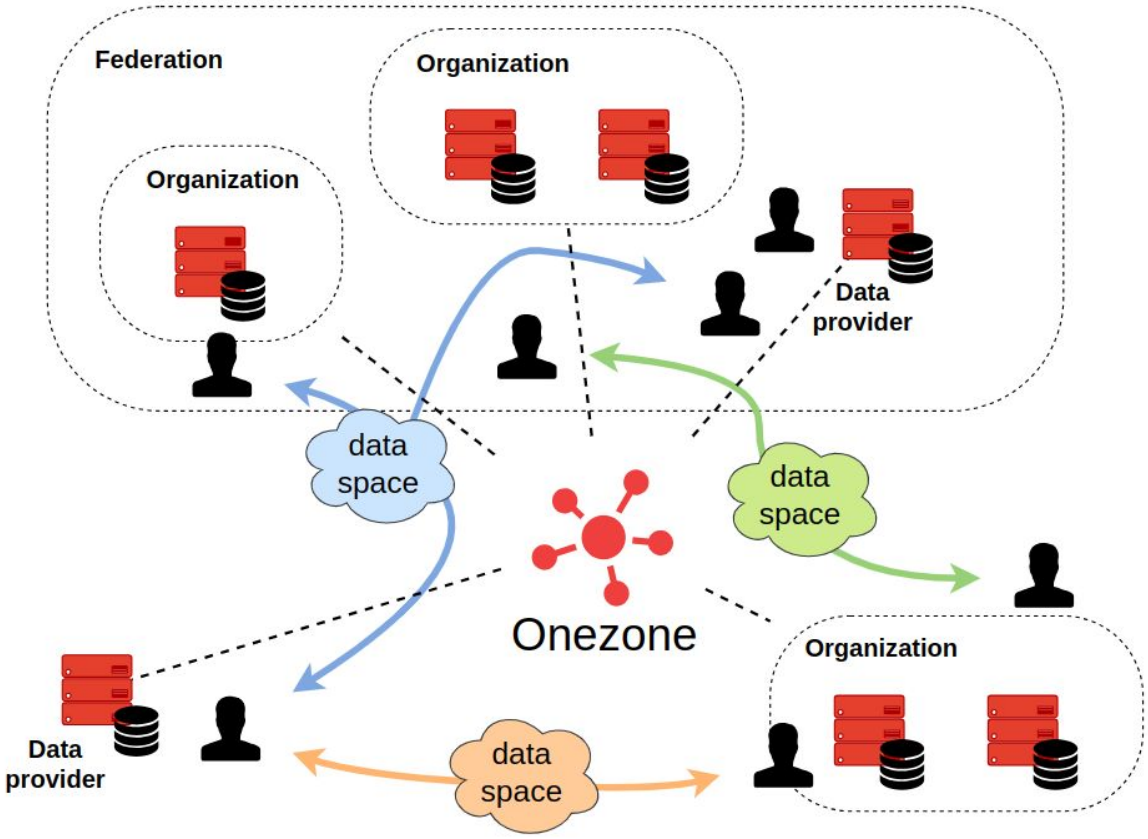
MINISTRY OF EDUCATION, YOUTH AND SPORTS

ONEDATA SPACES FOR DISTRIBUTED DATA

elixir CZECH REPUBLIC

MINISTRY OF EDUCATION, YOUTH AND SPORTS

Onedata architecture



Onezone

- central coordinator
- implements AAI
- entry-point to the system

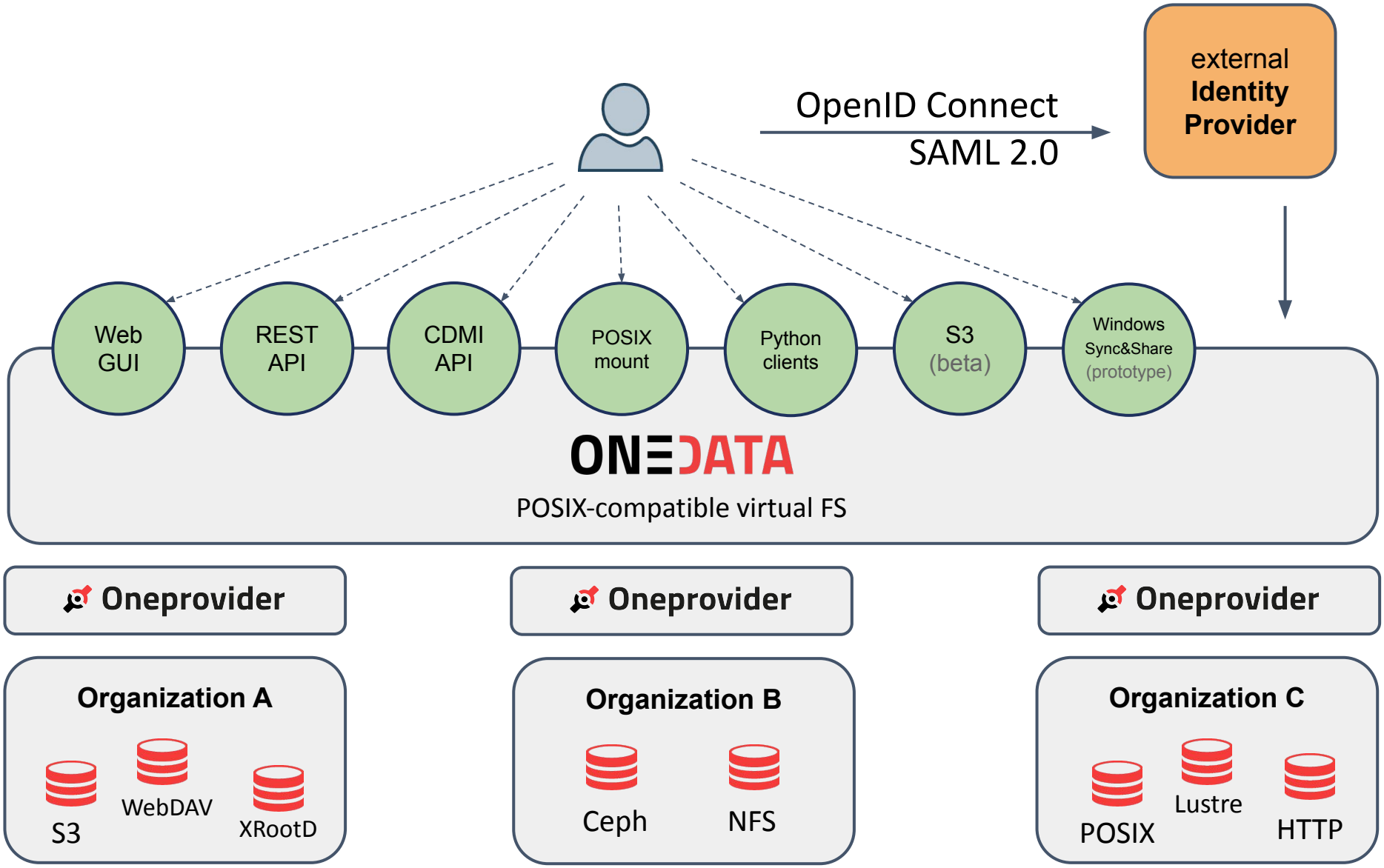
Oneprovider

- installed at a data provider
- realizes data access & mgmt
- virtualizes physical storage

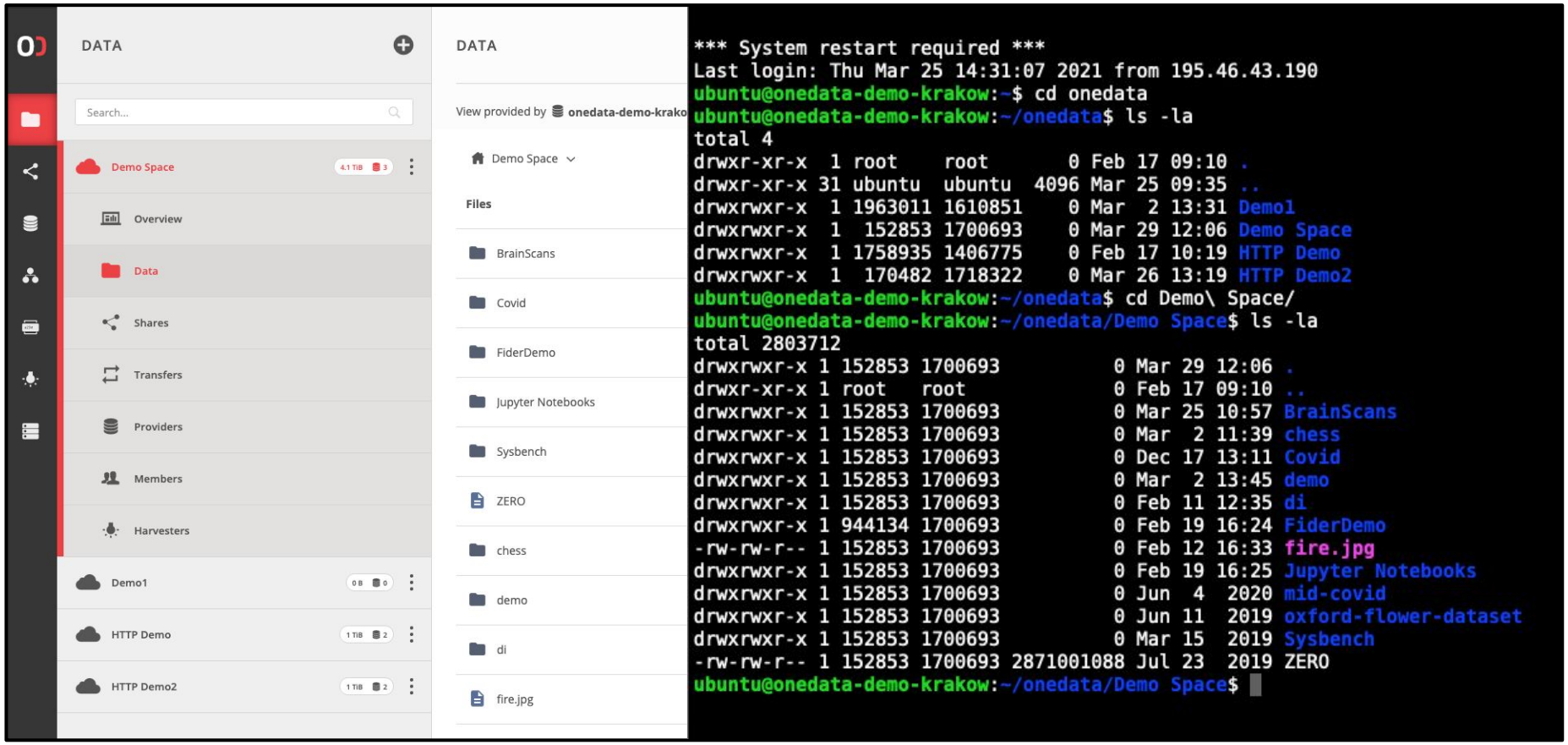
Open-source software stack

Take it and use it, retaining control over your own storage resources.

Virtualized data access



ONECLIENT — FUSE-BASED POSIX MOUNT



Rich REST API defined in swagger (OpenAPI)

```
curl -H "X-Auth-Token: $TOKEN" -H 'Content-type: application/json' \
-X GET "https://$PROVIDER_HOST/api/v3/oneprovider/data/$DIR_ID/children" -d '{
  "attributes": ["name", "size", "xattr.key"],
  "limit": 2
}'

{
    "children": [
        {
            "name": File1.txt,
            "size": 1024,
            "xattr.key": "extended_attribute_value"
        },
        {
            "name": File2.txt,
            "size": 16384,
            "xattr.key": null
        }
    ],
    "isLast": false,
    "nextPageToken": "g2gDZAAKbGlua190b2tlbmgCZAAMY2FjaGVkX3Rva2VuWgADY..."
}
```

Python clients and S3

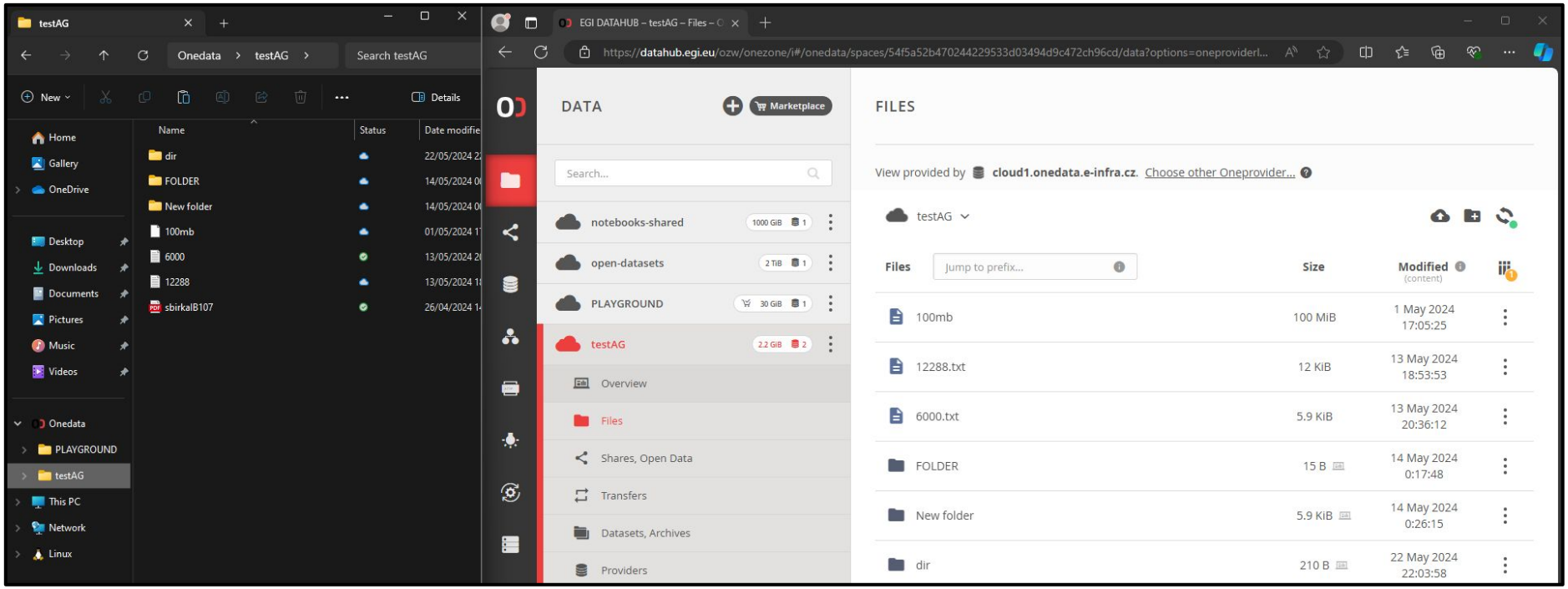
- OnedataFS *(pyfilesystem)*
- OnedataRestFS *(pyfilesystem)*
- OnedataFileRestClient

```
docker run -it --rm -v `pwd`:`pwd` -w `pwd` onedatarestfs bash
root@544de5a58bde:/home/lopiola# python
Python 3.12.6 (main, Sep 27 2024, 06:10:24) [GCC 12.2.0] on linux
Type "help", "copyright", "credits" or "license" for more information.
>>>
>>> from fs.onedatarestfs import OnedataRESTFS
>>> onedata_access_token = 'MDAxY2xvY2F00aW9uIGRhdGFodWIuZWdpLmV1CjAwNmJpZ
GVudGlmaWVyIDIvbm1kL3Vzci0000NWNjZjBjYmRmYTNmNmEyZTAxZTRiNTBiZTRkZTA1Y2NoN
zYwNS9hY3QvOWU4MjY3NjlkZTMzY2ZiNGRhNTg1ZjY00ZmRiZWMzMTBjaDQ1YWQKMDAxOWNpZC
BpbnRlcmZhY2UgPSByZXN00CjAwMThjaWQgc2VydmljZSA9IG9wdy00qCjAwMmZzaWduYXR1cm
UgDBIxbJLiYuB6WMKy3021O6gbPS65bYo00qXZk029epxBw8K'
>>> odfs = OnedataRESTFS('datahub.egi.eu', onedata_access_token)
>>> odfs.listdir('/demo-egi-conf-24')
['bagit-archive.tar.gz', 'colourful-and-horizontal', 'cyfronet-photos', 'c
yfronet.txt', 'lecce.txt', 'porto-photos', 'porto.txt']
>>> odfs.listdir('/demo-egi-conf-24/porto-photos')
['aleksander.jpg', 'andrei.jpg', 'andrews2.jpg', 'artem.jpg', 'baskin.jpg'
, 'bella.jpg', 'deepu.jpg', 'digital.jpg', 'dog.png', 'eberhard2.jpg', 'eb
erhard4.jpg', 'fabian.jpg', 'francesco.jpg', 'hristo.jpg', 'irina3.jpg', '
jakub.jpg', 'lumn.jpg', 'maisa.jpg', 'mali.jpg', 'masood.jpg', 'min.jpg',
'philippe.jpg', 'piccinng.jpg', 'pixabay3.jpg', 'pixabay4.jpg', 'prashant.
jpg', 'sasha.jpg', 'stein.jpg']
>>>
```

Standard S3 protocol (beta)

```
lopiola@rzyleta:~
lopiola@rzyleta ~
➤ aws --endpoint-url https://lecce.datahub.egi.eu:8443 s3 ls s3://demo-
egi-conf-24 --recursive
2024-10-02 14:19:20       6051 bagit-archive.tar.gz
2024-10-02 14:17:33     435507 colourful-and-horizontal/aleksander.jpg
2024-10-02 14:31:08         83 colourful-and-horizontal/baskin.jpg
2024-10-02 14:16:58     689963 colourful-and-horizontal/bella.jpg
2024-10-02 14:17:25     740914 colourful-and-horizontal/eberhard2.jpg
2024-10-02 14:17:29     313295 colourful-and-horizontal/fabian.jpg
2024-10-02 14:31:10         81 colourful-and-horizontal/lumn.jpg
2024-10-02 14:16:34    2528113 colourful-and-horizontal/masood.jpg
2024-10-02 14:17:28     356871 colourful-and-horizontal/min.jpg
2024-10-02 14:17:18     256676 colourful-and-horizontal/pixabay4.jpg
2024-10-02 14:15:59     301629 colourful-and-horizontal/stein.jpg
2024-10-02 14:20:22          0 cyfronet-photos/
2024-10-02 14:20:36    3380936 cyfronet-photos/bagit/alpha.jpeg
2024-10-02 14:20:38    2542904 cyfronet-photos/bagit/beta.jpeg
2024-10-02 14:20:35    1971910 cyfronet-photos/bagit/delta.jpeg
2024-10-02 14:20:35    2260725 cyfronet-photos/bagit/epsilon.jpeg
2024-10-02 14:20:39    1131789 cyfronet-photos/bagit/eta.jpeg
2024-10-02 14:20:35    2574198 cyfronet-photos/bagit/gamma.jpeg
2024-10-02 14:20:43    2639153 cyfronet-photos/bagit/iota.jpeg
2024-10-02 14:20:43    3042851 cyfronet-photos/bagit/kappa.jpeg
2024-10-02 14:20:43     763801 cyfronet-photos/bagit/lambda.jpeg
2024-10-02 14:20:43    1857002 cyfronet-photos/bagit/mu.jpeg
2024-10-02 14:20:51    2599747 cyfronet-photos/bagit/nu.jpeg
2024-10-02 14:20:44     575663 cyfronet-photos/bagit/omicron.jpeg
2024-10-02 14:20:50     855434 cyfronet-photos/bagit/pi.jpeg
2024-10-02 14:20:50     394412 cyfronet-photos/bagit/rho.jpeg
2024-10-02 14:20:51    1411808 cyfronet-photos/bagit/sigma.jpeg
2024-10-02 14:20:53    3764287 cyfronet-photos/bagit/tau.jpeg
2024-10-02 14:20:40    1908966 cyfronet-photos/bagit/theta.jpeg
2024-10-02 14:20:54    2328940 cyfronet-photos/bagit/upsilon.jpeg
```

Windows Sync & Share client (prototype)



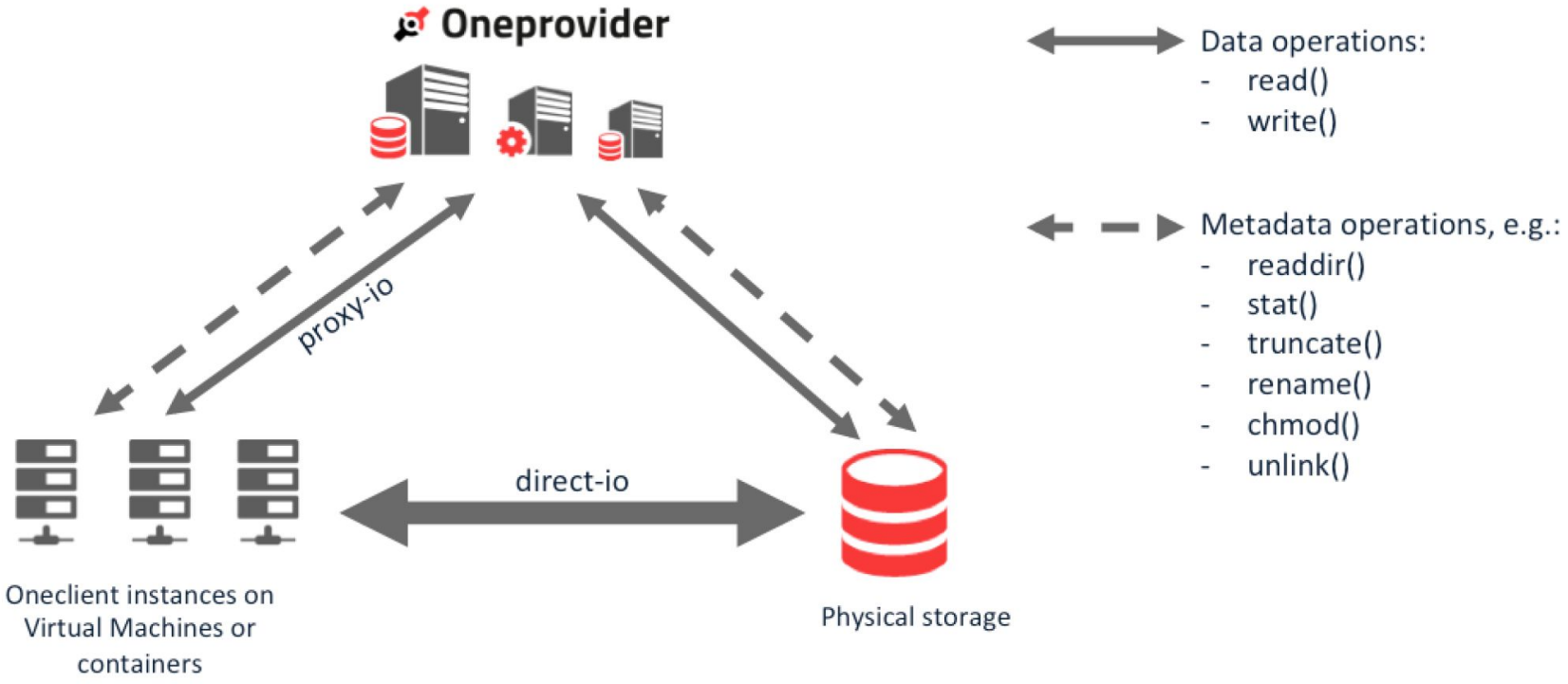
Credits:
Andrej Gáfrik and Tomáš Svoboda
Masaryk University Brno and CESNET

elixir CZECH REPUBLIC

MINISTRY OF EDUCATION, YOUTH AND SPORTS

Direct storage access for HPC

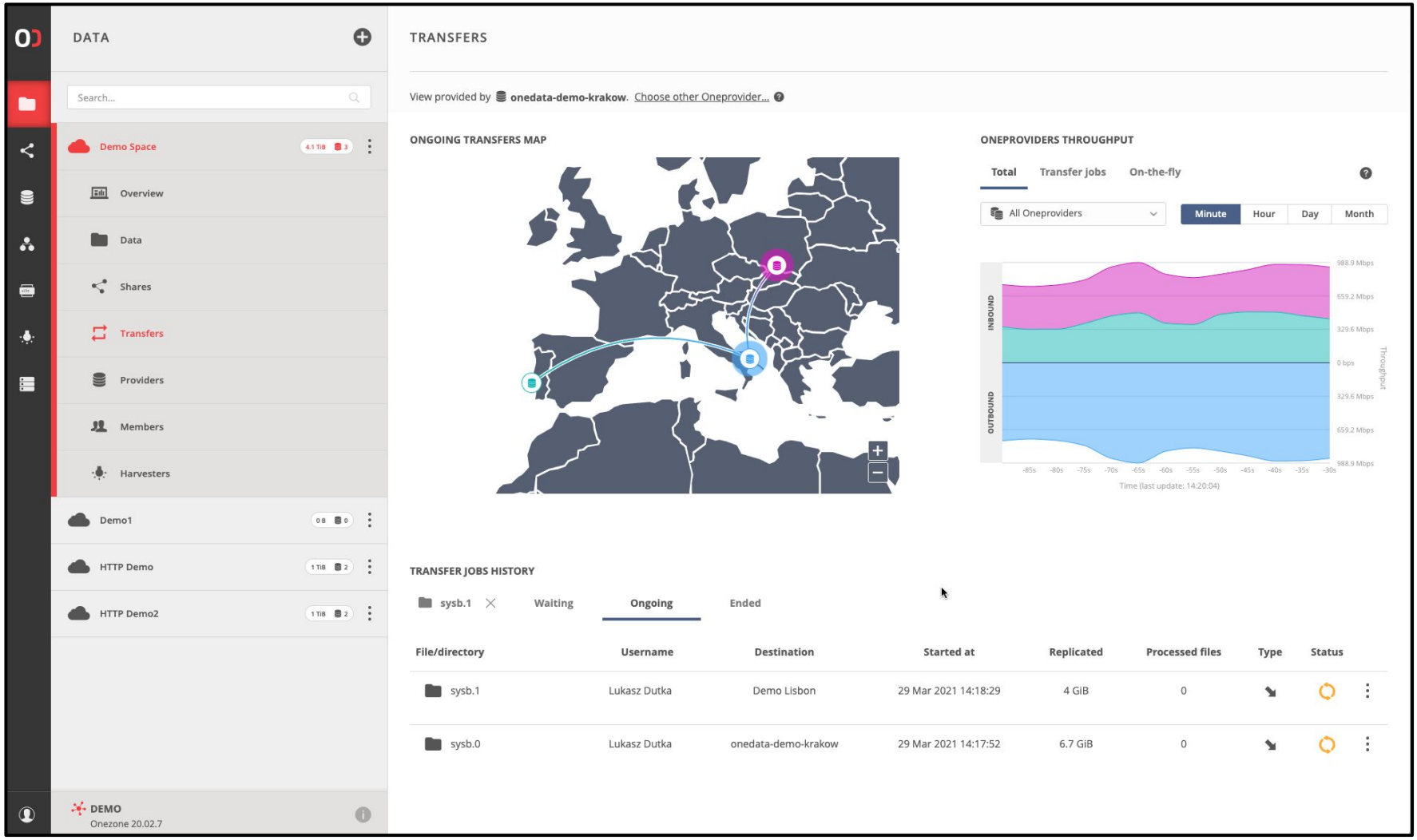
- Oneclient - Linux **POSIX** mount supporting most of the POSIX operations (C++).
- OnedataFS - **Python** `pyfilesystem` plugin reusing the C++ storage drivers.
- All your data accessible in a unified file system mountable on a PC, VM, container, or Cloud/Grid worker node.



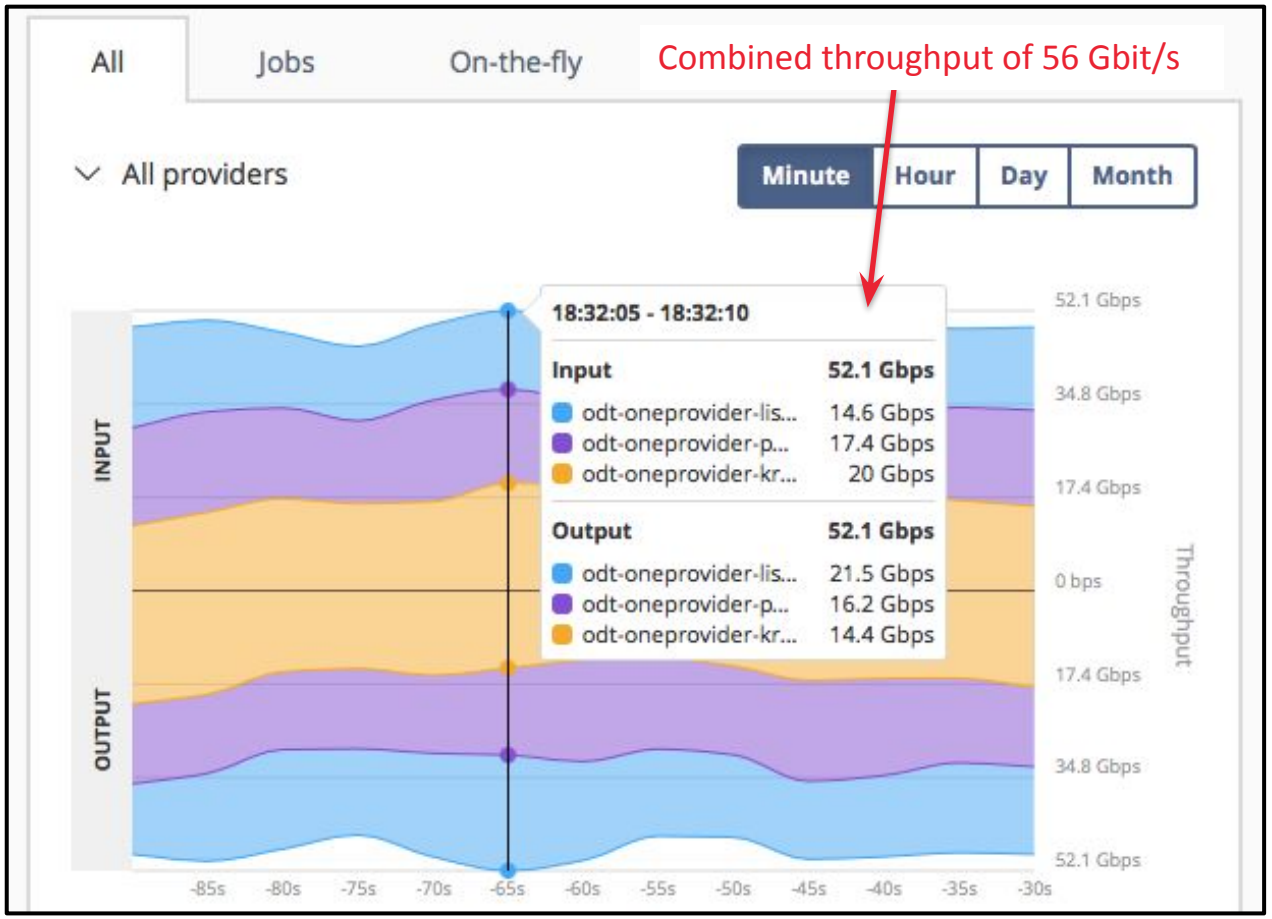
Advanced replica management

- **Chunk-based** replica management
- Missing chunks delivered **on-the-fly**
- Handled **automatically**, but **controllable manually** - prestage, transfer, migrate, define QoS rules (policy-based replica management)

Flexible, high-throughput transfers



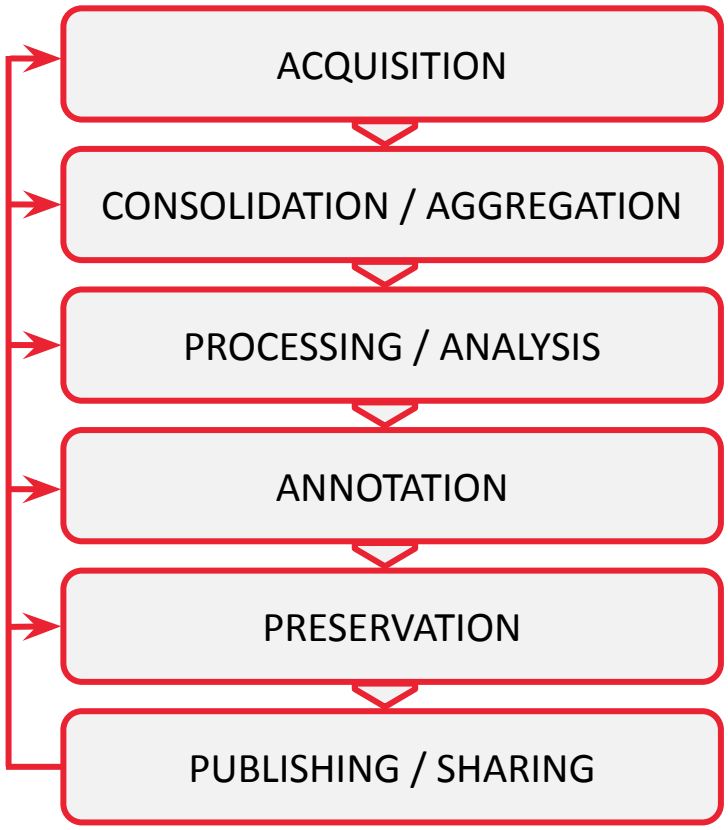
Flexible, high-throughput transfers



Data transfer “mesh”:

- 3 Oneproviders connected by 20+Gbit/s links
- transfer data in all possible directions
- 1 VM per provider
- almost linear scalability thanks to parallel channels

Exemplary data lifecycle in onedata



Ingesting data into the system; manually or automatically using APIs, middleware, data import, legacy data.

Organizing different data sources in a single logical namespace, establishing data structure and replication.

Performant data processing despite the virtualization, also in multiple locations simultaneously (including data writes).

Assigning arbitrary custom metadata (key-value, JSON, JSON-LD, RDF) to files and directories.

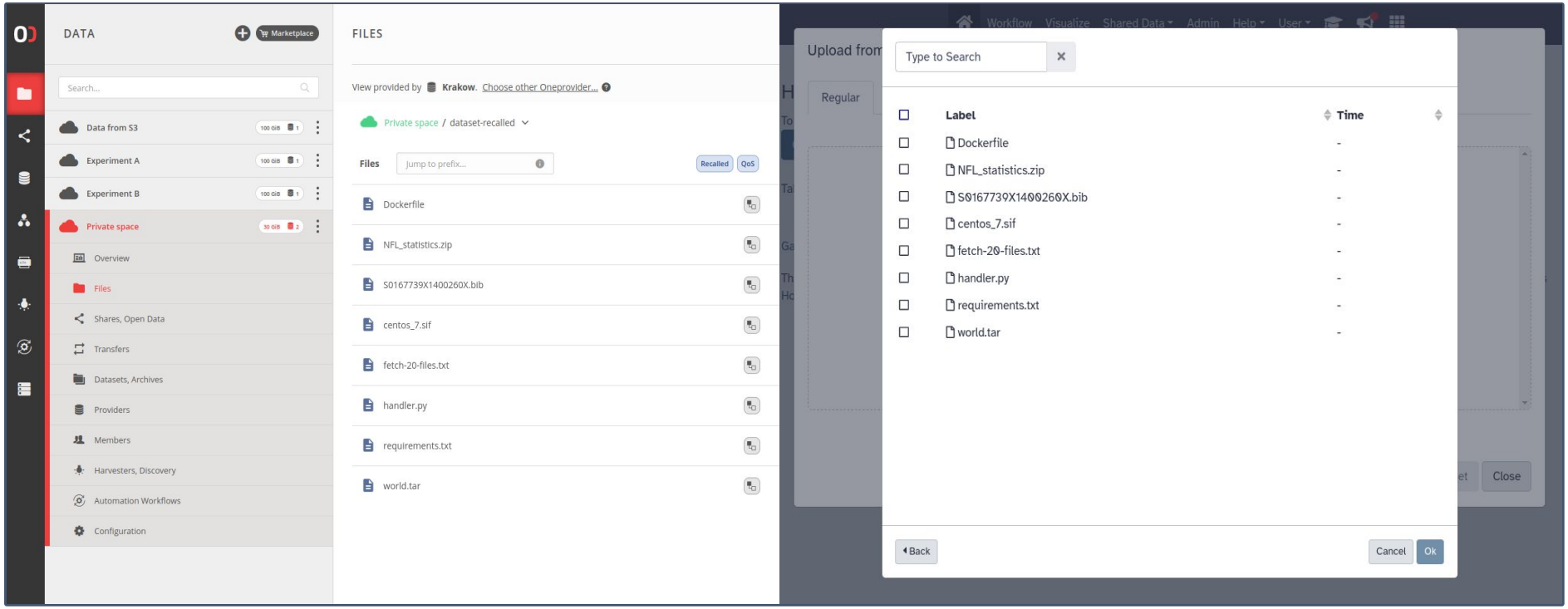
Creation of archives to package the processed and annotated data.

Semi-public link-based sharing or Open Data publishing with integrated PID/DOI minting, pluggable integration with PID services (e.g. DataCite) and advertising via the OAI-PMH protocol.

Onedata instances

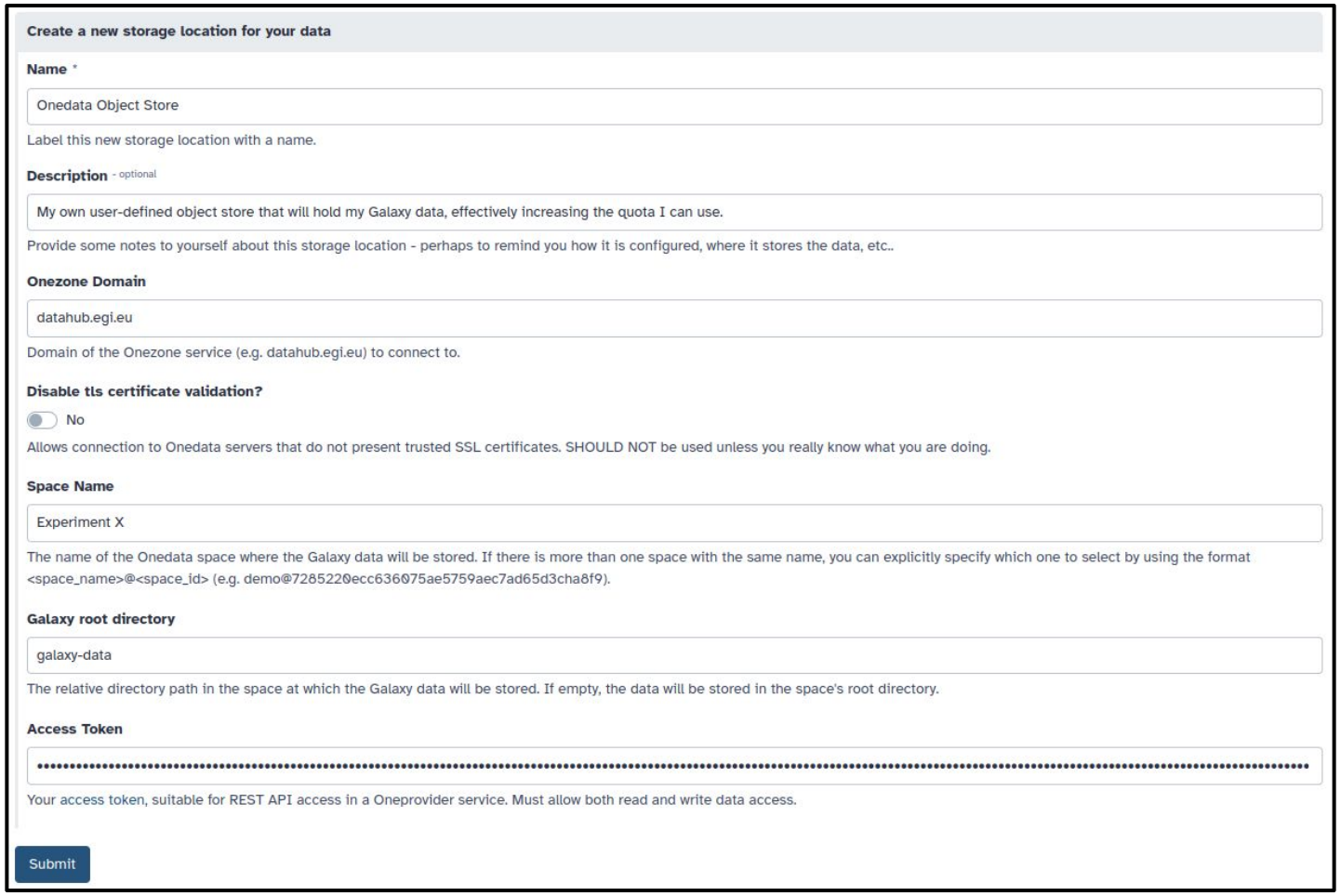
- Between 5 and 10 active Zones in Poland and EU (depending on project lifecycles).
- Several instances not maintained by us.
- DataHub (on the map), long haul project:
  - **18** sites (Oneproviders)
  - **2255** data spaces
  - **~2.16PB** total storage size
  - **800+** users
- Archive for Polish National Museums:
  - **5PB** of data — the current phase
  - **10PB** of data — target scale
  - **~100M** files

Onedata & Galaxy integration



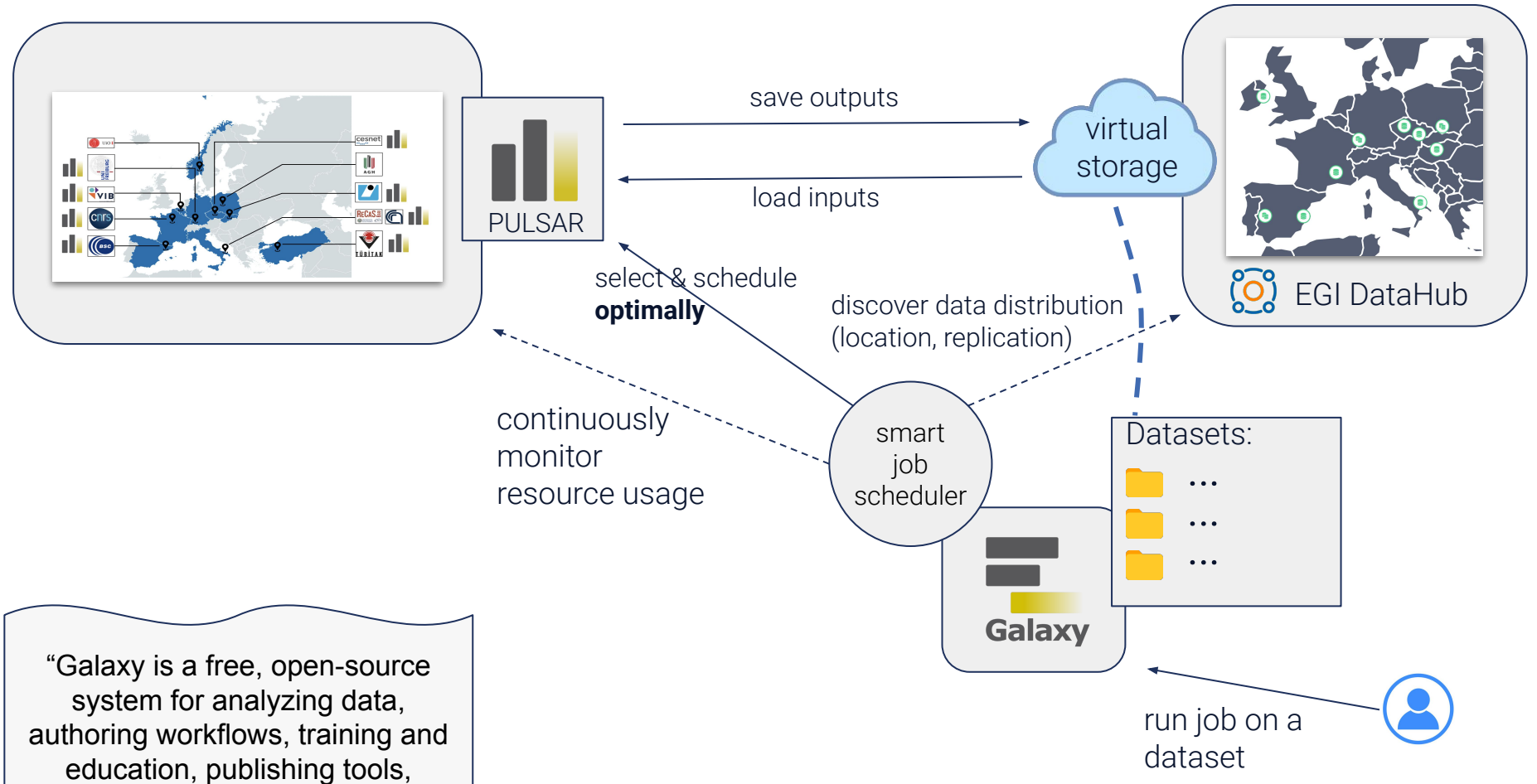
Import/export data from/to your spaces
(using any Onedata ecosystem)

ONEDATA & GALAXY INTEGRATION



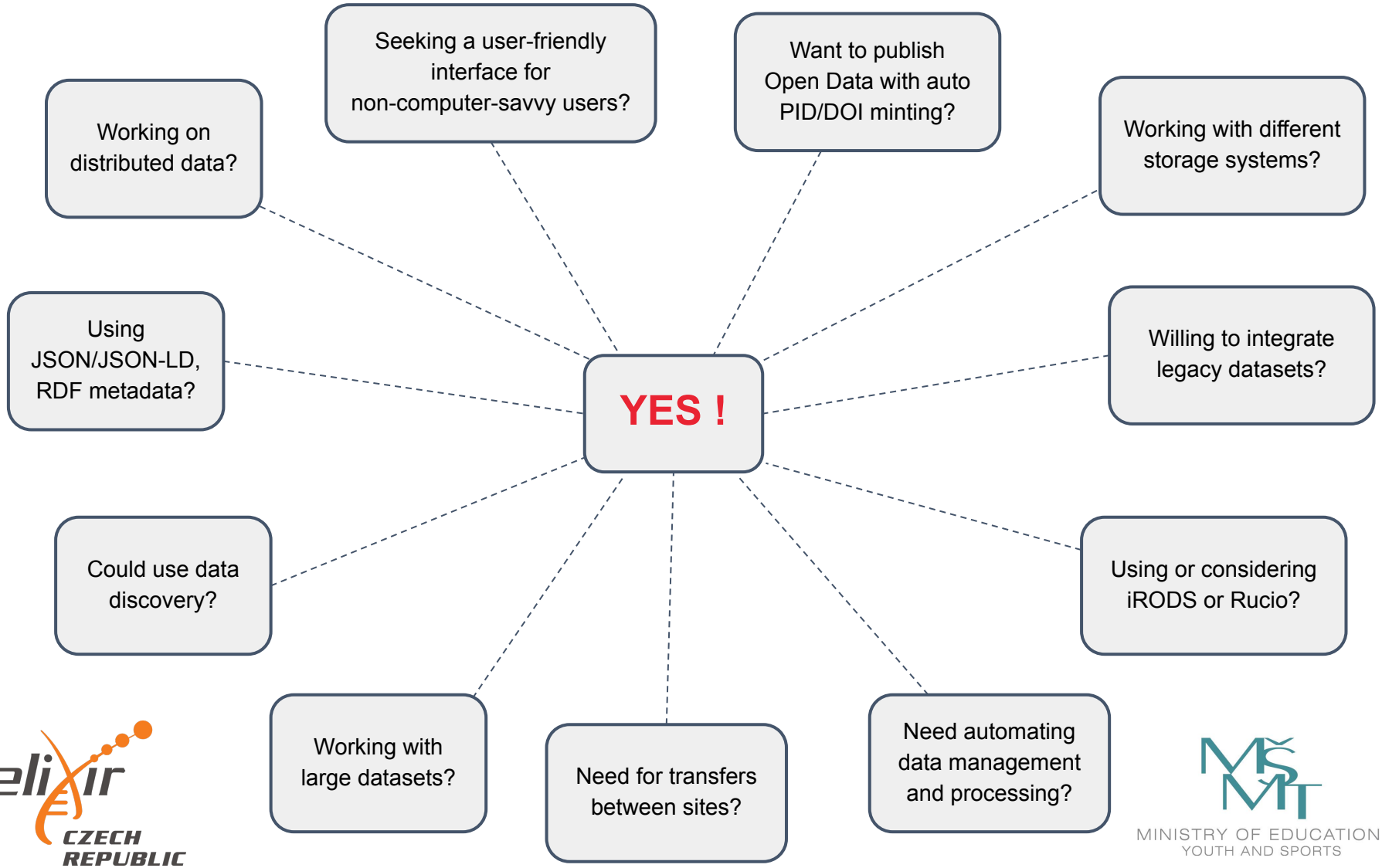
Configure a personal Onedata storage location for Galaxy user data (Object Store)

GALAXY + ONEDATA = WORKFLOWS THAT UNDERSTAND AND EMBRACE DATA DISTRIBUTION



"Galaxy is a free, open-source system for analyzing data, authoring workflows, training and education, publishing tools, managing infrastructure, and more."

Should you look into Onedata?



elixir CZECH REPUBLIC

MINISTRY OF EDUCATION, YOUTH AND SPORTS

THANK YOU

Let's talk about your use-case!

Onedata homepage — **https://onedata.org**

Onedata training — **https://onedata.org/training**

Contact me — **lukasz.opiola@onedata.org**

special offer: ask me for a personalized demo

elixir CZECH REPUBLIC

MINISTRY OF EDUCATION, YOUTH AND SPORTS

ELIXIR CZ Annual Conference 2024

This conference was supported by the Ministry of Education, Youth and Sports, Czech Republic, through Grant LM2023055 for large research infrastructures and access to computing and storage facilities.

elixir CZECH REPUBLIC

MINISTRY OF EDUCATION, YOUTH AND SPORTS