



# Progress Highlights on Integration of data services with computing platforms

EUDAT Conference September 13th – 15th, 2022, Athens, Greece

**Chris Ariyo, CSC IT Center for Science, Finland**



The DICE project has received funding from the European Union's Horizon 2020 project call H2020-INFRAEOSC-2018-2020 under Grant Agreement no. 101017207

- ▶ Integrate data services (B2SAFE, B2SHARE, B2DROP, B2FIND, B2HANDLE) with European computing platforms and infrastructures, like Lumi, FENIX and EuroHPC
- ▶ Enable analysis, data replication, and data publication for HPC and cloud computing environments
  - ▶ B2DROP will be used to maintain “recipes” for analysis
  - ▶ B2SAFE as a tool to register results
  - ▶ B2SHARE as a tool to publish datasets

## ➤ B2DROP

- Ensure that small data (batch queue scripts etc. similar small objects) can be read from B2DROP to HPC environment, and that small data can be written back to B2DROP.
- Access to B2DROP
- NextCloud occ command line tool from computing environments
- WebDAV tools or other similar tools like DaviX



- ▶ B2SAFE
- ▶ Autoingest feature
- ▶ AAI integration
- ▶ AAI integration
- ▶ iRods policy usage
- ▶ Object storage usage
- ▶ Activities for B2SAFE
  - ▶ Evaluate B2SAFE (iRODS) autoingest feature with object storage system and with most common file systems on HPC systems
  - ▶ Connection of B2SAFE to B2ACCESS AAI
  - ▶ Evaluate usage of iRods policy to move data to B2SAFE
  - ▶ Evaluate pros and cons of keeping the data under user space

## ➤ B2SHARE

- Share and publish data objects from HPC and computing environments.
- Share and publish data objects that has been auto-ingested by B2SAFE.
- Share data stored in B2SAFE or other storage services.
- Copy data from B2SAFE to B2SHARE and Share (B2SAFE – B2SHARE bridge).

## ➤ B2DROP integration using share links

- `wget -content-disposition https://b2drop.eudat.eu/s/7mAX6FqfmMz7afF/download`
- `wget`: command-line tool that retrieves content from web servers

## ➤ B2DROP integration using WebDAV

- WebDAV: Web Distributed Authoring and Versioning, HTTP extension that allows a collaborative and writeable usage of the content of a web server
- Authorisation and authentication tokens are generated from B2DROP web interface

## ➤ B2SHARE integration using HTTP REST API and cURL as a tool

- cURL: command-line tool for getting or sending data including files using URL syntax
- Authorisation and authentication tokens are generated from B2SHARE web interface
- Used to share and publish data

## ➤ B2SAFE data upload and download using HTTP-API protocol and cURL as tools for data transfer

- HTTP-API: Provide authorisation header
- cURL: Authentication using username / password

## ➤ Data transfer between Object Storage and HPC using Rclone as tool for data transfer

- Rclone is a command-line tool that allows to manage and transfer content to another storage
- Configuration file needed to simplify process
- Need to load modules for access and to configure connection



## Ongoing work

- Autoingest of data to and from EUDAT services to HPC environment
- B2SAFE integrates data from ObjectStorages
- Better integration of B2SAFE to B2SHARE
- Better connection of B2SAFE to B2ACCESS using OpenID Connect (OIDC)

## Challenges

- Automatic WebDAV mounting of B2DROP requires root access, this limits the user experience when using HPC systems
- Integration of FENIX communities into B2ACCESS require discussions and decisions on the political/policy level



## JOIN OUR COMMUNITY



DICEosc



/company/diceosc



Chris Ariyo

CSC – IT Center for Science, Finland