







# Adopting RUCIO for the Italian datalake

# National experience and perspectives

**A. Troja**, M. Biasotto, D. Ciangottini, M. Delli Veneri, F. Fanzago, A. Italiano, N. Marcelli, L. Morganti, A. Rendina, M. Sgaravatto, D. Spiga, B. Spisso, S. Stalio, M. Verlato

#### 30/09/2024

7° Rucio Community Workshop









# **INFN Computing infrastructure**

The National Institute for Nuclear Physics (INFN) is the Italian research agency dedicated to the study of the fundamental physics.

Distributed heterogeneous computing infrastructure: resources are both integrated in WLCG and exposed through Cloud interfaces.

Resources are about:

- 200k CPU cores;
- 250PB disk storage;
- 200PB tape storage;
- More than 100 international scientific collaborations served.









3



# The TeRABIT project

The Terabit network for Research and Academic Big data in Italy (TeRABIT) envisions the creation of a a distributed, hyper-connected, hybrid HPC-Cloud environment that offers services designed to meet the evolving needs of research and innovation.

The environment will federate and strengthen the three existing research infrastructures GARR-T, PRACE-Italy and HPC-BD-AI (HPC-Big Data-Artificial Intelligence), leveraging their existing of connections to other national and European research infrastructures and data spaces through the GÉANT backbone.

#### Main objectives:

- Enable widespread data transfer, up to Terabits per second, and services on a national scale in Italy, connected to Europe;
- Innovate the central HPC node of PRACE-Italy;
- Innovate the HPC services offered to researchers, beyond the centralized calculation model, adding distributed "HPC-Bubbles"



0



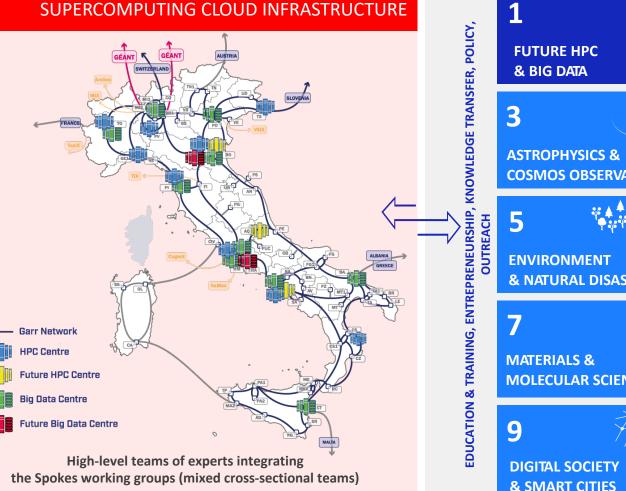


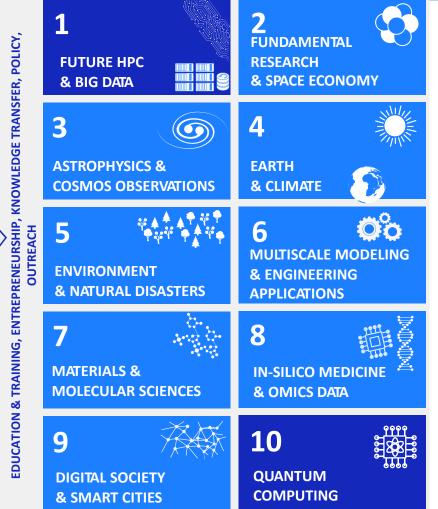


ICSC ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

The Italian Research Center in High Performance Computing, Big Data and Quantum computing conducts R&D at a national and international level.

- 25 universities
- 12 Research institutes
- 14 Strategic private companies











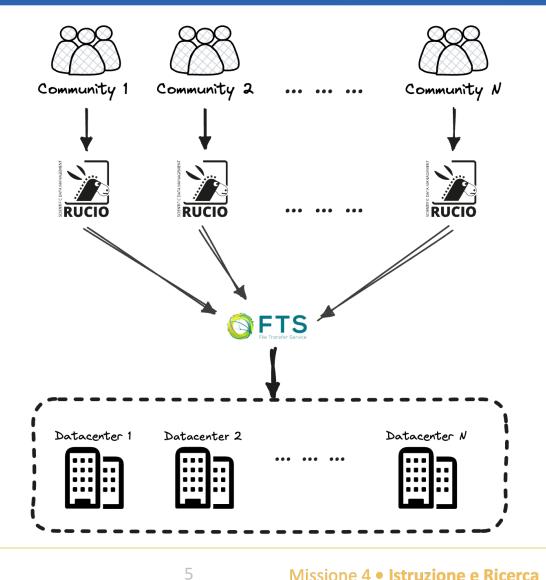


# Data-centric model and the National datalake

Each INFN computing centre hosts a storage system, peculiar to that site, including both Grid and Cloud backend. The need to federate these heterogeneous storages brings to the creation of a National datalake where users are allowed to access and manage data from any federated node. The underlying infrastructure must be transparent to them.

Each community would interact with its Rucio instance, which in turn relies on a single FTS server centrally managed. TPC are performed among the federated datacenter.

Federation will include INFN facilities as well as external ones, just like **CINECA Prace-Italy HPC.** 











# Our experience with Rucio deployment

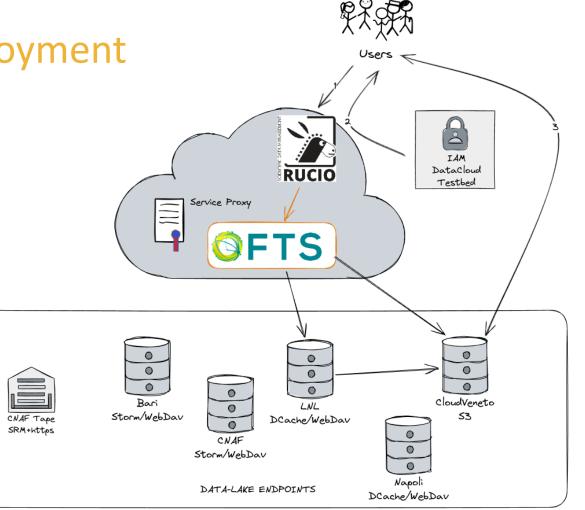
In order to test the federation of geographically distributed, etherogeneous INFN storages, we deployed a **data management testbed**, on which we tested and validated several configurations.

We deployed Rucio on a K8s cluster, hosted in INFN Cloud machines. Same goes for the FTS server. Users interact with Rucio via JWT token, while x.509 VOMS proxies are used by FTS for the TPCs

INDIGO IAM is used to AuthN/Z, and the Rucio database is hosted by a high-availability PostgreSQL infrastructure.

RSEs abstract six INFN storage systems, with different QoS (disk and tape), implementation and protocols.

We use Flux to manage the infrastructure.











## Our experience with scientific communities

We're already supporting scientific communities that decided to adopt Rucio (and also act as beta testers):

- <u>Cygno</u> is an experiment on dark matter detectors. Data produced are stored in the INFN Cloud Backbone, analyzed and then moved to tape storage (and deleted from disks when processed). We completed the implementation of the the Rucio infrastructure early this year;
- <u>DarkSide</u>: research on direct evidence of dark matter from 2013 to 2020. They are developing the Rucio infrastructure by themselves, following the documentation we wrote;
- <u>DAMPE</u>: research on indirect detection of dark matter signature. We're setting up an infrastructure that allows interaction between Tier-1 storage at INFN and data center located in China;
- <u>Euclid Italy</u>: The INFN contribution in the Euclid space mission is remarkable, but it lacks of a way to manage data. The process of developing a Rucio instance started with a kick-off meeting early September.







8



# AuthZ model

We tested several authorization models on the data, in order to fulfill the use cases requirements:

- **Group-based policies**: group A users won't be allowed to access data owned by group B, while it can perform any action on group A data (even deletion);
- Scope-based policies: scopes are defined within the IAM, and passed to Rucio through the token. Scopes give user permission to certain action on certain area in the storages. As an example, we implemented *user isolation* (e.g., each user can read everything but only write in his/her own designated area);
- **RSE-based policies**: users can perform actions only on a specific type RSE and not on others, for example only privileged users can trigger TPC from/to a TAPE RSE. This is useful to avoid misuses of the tape library, which would impact also other user communities.









### **User registration**

Assigning scim:read scope to the RUCIO operator (as required for the integration with IAM) is not compliant with INFN policies unless this operator has been nominated to manage personal data (and this is not always the case)

We therefore **developed a user registration layer**, put in front of the data management infrastructure.

A user that wants to create an account in Rucio, must authenticate through OAuth2 on the WebUI developed specifically for the user registration. Upon accepting the terms, a Rucio account is created, along with:

- User IAM identity attached;
- Quota assigned;
- User-owned scope created.

We think this feature can be **useful for the general Rucio audience**. We're available to talk about the details.









# **S3** integration

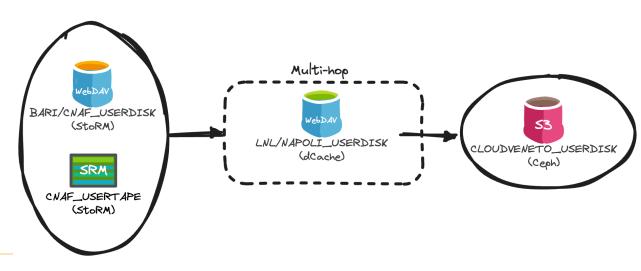
Besides "Grid" storage systems we need to integrate also some S3 endpoints (ceph-rgw and minio)

S3 endpoints must be **manually registered in both RUCIO and FTS** (registering also the S3 credentials that allow read/write access to the relevant buckets)

- As far as we understand this is needed even if the S3 endpoint is configured to authenticate with tokens
- Exposing these credentials to the RUCIO and FTS instances is not very "clean"
- Is this (and more in general the S3 integration model) going to change/evolve?

To implement TPCs between S3 endpoints we rely on multi-hop:

Implemented by deploying an RSE supporting both push and pull mode (dCache webdav in our deployments) The same approach is used also for transfers from STORM based RSEs to S3 RSEs (since STORM doesn't support the push mode)











#### **Issues at deployment**

Thanks to the chat in Mattermost and the help of the developers, we were able to face the issues that arose during development. These include:

- Rucio-server container doesn't include mail templates;
- Reaper daemon doesn't work unless we import "threading" and "concurrent.futures.thread";
- Reaper daemon Helm chart overwrites the certificates in "/etc/grid-security/certificates" with the ones in secret "rucio-cabundle-reaper";
- Hermes daemon doesn't allow the use of smtp port and authentication.

The solution we implemented was to create new containers starting from official ones, with all the necessary patches. (of course, it is a sub-optimal solution).

We may push these code developments upstream if there is some interest (and we don't want to diverge from the official code base).

Could make sense for us to create a PR? Will you review these items?









# **External Metadata catalog**

One of the requirements we got from our early adopters is about the support for **the integration of domain specific metadata catalog**:

- by definition are external to the Rucio system.

We might use the RUCIO DB (i.e. extending it) however also following what already done at SKA we understand that using an external DB properly linked (i.e. via rucio.dids) to the RUCIO DB is a better choice.

Few questions:

- Is this following the right integration model?
- Is already planned to be embedded in Rucio?
- Any forum where discussions related to this topic happen?









#### Documentation

We have produced documentation for both Rucio users and operators on INFN Cloud, especially tackling the gaps in the official documentation (for example rule approval and Hermes configuration.)

This documentation **allows any user to replicate our DM system** anytime is required, without major effort.

<ul> <li>Start from a fresh 'vanilla'</li> </ul>	Pages / / RUCIO deployment in the WP6 testbed 🍙 / Edit Q View inline comments 🏠 Save for later 🛛 Watching < Share …
Start from a INFNCloud da	Av Analytics
Prerequisites     First step towards the data	Rucio server deployment
Kick-off deployment	Deploy rucio-server pod for interaction with Rucio
Rucio server deployment	Deplay Rucio root client
OIDC AuthN/Z with token	Let's move to the core of our application, the Rucio server. In this section, we'll deploy the main Rucio server instance and the client to login as root user within the Rucio framework.
User syncing (oidc-sync)	
Start using Rucio	As stated in the official documentation: "The server layer serves the purpose of authentication & provides a common API for interaction with clients & other external application, as also the Web UI."
VOMS Proxy Authenticatic	We'll need two instances of rucio-server: one to connect the client to the Rucio functions and a second one to authenticate the Rucio user. Let's start from the first
Rucio daemons deployme	one.
RSE Creation	
RSE distance and multihop	Deploy rucio-server pod for interaction with Rucio
Upload a file to the datalal	The first rucio-server instance allows an authenticated user to interact with the Rucio features. We'll deploy it starting form the Helm chart created by the Rucio
Customization of Rucio pc	developers, which repo is located here.
Rucio WebUI	We'll customize this chart by changing some of its value by means of a yaml file (the full list of modifiable value is here). As usual, this yaml file is located at ./apps/rucio-guide/apps, and its existence is notified to FluxCD by adding its name to the list of yaml in the kustomization file, i.e.:
Rule approval	kustomization yaml
Database backup (pg-dun	
Python API	1 aplVersion: kustomize.config.k8s.io/vlbeta1 2 kind: Kustomization
Appendix: AOB	3 resources:
Space tools «	4 - postgres.yaml
INFN Confluence Spaces	
WP6 - Ricerca e Sviluppo, Testbed	v     People     Calendars     Analytics     Create     ···       Upload data to the DataLake
WP6 - Ricerca e Sviluppo, Testbed	
WP6 - Ricerca e Sviluppo, Testbed	
WP6 - Ricerca e Sviluppo, Testbed > WP6 Meetings < Data Management	Upload data to the DataLake Upload a file to a specific RSE
WP6 - Ricerca e Sviluppo, Testbed > WP6 Meetings > Data Management > FTS3	Upload data to the DataLake
WP6 - Ricerca e Sviluppo, Testbed > WP6 Meetings > Data Management > FTS3 > IAM+VOMS	Upload data to the DataLake Upload a file to a specific RSE -bash-4.29 rucio uploadrse CNAF_USERDISKscope user.sgaravat BmmAnalysis.rootname BmmAnalysis-01register-after-ur 2023-03-05 1331124,931 INTO Preparing upload for file BmmAnalysis.root 2023-03-05 1331124,14 INTO Successfully added replica in Rucio catalogue at CNAF_USERDISK
VPG - Ricerca e Sviluppo, Testbed VPG Meetings Data Management > FTS3 > IAM+VOMS • Modello authZ per accesso da	Upload afile to a specific RSE -baah-4.2% rucio uploadrse CNAF_USERDISKscope user.sgaravst BmmAnalysis.rootname BmmAnalysis-01register-after-ur 2023-03-05 13:31:24,931 INFO Preparing upload for file BmmAnalysis.root
WP6 - Ricerca e Sviluppo, Testbec WP6 Meetings Data Management > FTS3 > IAM+VOMS • Modello authZ per accesso da > RUCIO notes	Upload data to the DataLake Upload a file to a specific RSE -baah-4.28 rucio uploadrse CMAP_USERDISKscope user.sgaravat BmmAnalysis.rootname BmmAnalysis-01register-after-ur 2023-03-05 13:31:24,931 INFO Preparing upload for file BmmAnalysis.root 2023-03-05 13:31:24,931 INFO Successfully added replica in Rucio catalogue at CMAP_USERDISK 2023-03-05 13:31:25,748 INFO Trying upload with dava to CMAP_USERDISK 2023-03-05 13:31:25,718 INFO Trying upload vith dava to CMAP_USERDISK 2023-03-05 13:31:25,718 INFO Trying upload vith dava to CMAP_USERDISK 2023-03-05 13:31:25,715 INFO Successfully add of temporary file. dava//xfer.cr.enf.infn.it:84d3/DataCloud-TB/user/g
WP6 - Ricerca e Sviluppo, Testbed WP6 Meetings • Data Management • FTS3 • IAM+VOMS • Modello authZ per accesso da • RUCIO notes • Testbed Data Management	Upload data to the DataLake Upload a file to a specific RSE -bash-4.2% rucio uploadrse CNAF_USERDISKscope user.sgaravat BmmAnalysis.rootname BmmAnalysis-01register-after-ur 2023-03-05 1331125,471 INTO Preparing upload for file BmmAnalysis.root 2023-03-05 1331125,477 INTO Successfully added replication rule at CNAF_USERDISK 2023-03-05 1331125,478 INTO Successfully added replication rule at CNAF_USERDISK 2023-03-05 1331125,488 INTO Trying upload with dava to CNAF_USERDISK
WP6 - Ricerca e Sviluppo, Testbed WP6 Meetings Data Management > FTS3 > IAM-VOMS • Modello authZ per accesso da > RUCIO notes • Testbed Data Management > Data Management Testbed:	Upload data to the DataLake  -bash-4.2% rucio uploadrse CNAF_USERDISKscope user.sgaravit BmmAnalysis.rootname BmmAnalysis-01register-after-ur 2022-03-05 133112,4/331 INFO Preparing upload for file BmmAnalysis.root 2022-03-05 133112,5/471 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,5/471 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,5/471 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,5/180 INFO Trucing upload with dava to CNAF_USERDISK 2022-03-05 133112,6/115 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,6/115 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,6/115 INFO Successfully upload of temporary file.dava://xfer.cr.onaf.infn.its843/DataCloud-TB/user/z 2022-03-05 133112,6/115 INFO Successfully uploaded file BmmAnalysis.root //vufs/cms.cench/rucio/x86_64/rhe17/py3/current/lib/python3.6/site-packages/url1b3/connectionpool.py:1050: InsecureRequee InsecureRequeetFansion
WP6 - Ricerca e Sviluppo, Testbed WP6 Meetings • Data Management • FTS3 • IAM+VOMS • Modello authZ per accesso da • RUCIO notes • Testbed Data Management • Data Management Testbed: • Data Management Testbed: • How to test Grid storage acc	Upload data to the DataLake -baah-4.2% rucio uploadrse CMAF_USERDISKscope user.sgaravat BmmAnalysis.rootname BmmAnalysis-01register-after-ur 2021-03-05 1331124.931 INFO Preparing upload for file BmmAnalysis.root 2021-03-05 1331125,144 INFO Successfully added replication rule at CNAF_USERDISK 2023-03-05 1331125,786 INFO Successfully added replication rule at CNAF_USERDISK 2023-03-05 1331125,186 INFO Successfully added to fempiorary file. davs//xfer.cr.onaf.infn.its843/DataCloud-TB/user/t 2023-03-05 1331126,126 INFO Successfully uploaded file BmmAnalysis.root //vmfs/cmas.cern.ch/rucio/x86_64/rhP/hys/Ourrent/liby/thon3.c/site-package/url11b3/connectionpool.py:1059: InsecureRequeed
WP6 - Ricerca e Sviluppo, Testbed WP6 Meetings • Data Management • FTS3 • IAM+VOMS • Modello auth2 per accesso da • RUCIO notes • Testbed Data Management • Data Management Testbed: • Data Management Testbec • How to test Grid storage acc • Software Management	Upload data to the DataLake  -bash-4.2% rucio uploadrse CNAF_USERDISKscope user.sgaravit BmmAnalysis.rootname BmmAnalysis-01register-after-ur 2022-03-05 133112,4/331 INFO Preparing upload for file BmmAnalysis.root 2022-03-05 133112,5/471 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,5/471 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,5/471 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,5/180 INFO Trucing upload with dava to CNAF_USERDISK 2022-03-05 133112,6/115 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,6/115 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,6/115 INFO Successfully upload of temporary file.dava://xfer.cr.onaf.infn.its843/DataCloud-TB/user/z 2022-03-05 133112,6/115 INFO Successfully uploaded file BmmAnalysis.root //vufs/cms.cench/rucio/x86_64/rhe17/py3/current/lib/python3.6/site-packages/url1b3/connectionpool.py:1050: InsecureRequee InsecureRequeetFansion
WP6 - Ricerca e Sviluppo, Testbed > WP6 Meetings > Data Management > FTS3 > IAM+VOMS • Modello auth2 per accesso da > RUCIO notes > Testbed Data Management > Data Management Testbed: • Data Management Testbed: • How to test Grid storage acc > Software Management > 'datacloud-tb' VO	Upload data to the DataLake  -bash-4.2% rucio uploadrse CNAF_USERDISKscope user.sgaravit BmmAnalysis.rootname BmmAnalysis-01register-after-ur 2022-03-05 133112,4/331 INFO Preparing upload for file BmmAnalysis.root 2022-03-05 133112,5/471 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,5/471 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,5/471 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,5/180 INFO Trucing upload with dava to CNAF_USERDISK 2022-03-05 133112,6/115 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,6/115 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,6/115 INFO Successfully upload of temporary file.dava://xfer.cr.onaf.infn.its843/DataCloud-TB/user/z 2022-03-05 133112,6/115 INFO Successfully uploaded file BmmAnalysis.root //vufs/cms.cench/rucio/x86_64/rhe17/py3/current/lib/python3.6/site-packages/url1b3/connectionpool.py:1050: InsecureRequee InsecureRequeetFansion
WP6 - Ricerca e Sviluppo, Testbed > WP6 Meetings > Data Management > FTS3 > IAM+VOMS - Modello authZ per accesso da > RUCIO notes > Testbed Data Management > Data Management Testbed: - Data Management Testbec - How to test Grid storage acc > Software Management * datacioud-tb* VO > Object storage backbone (altern	Upload data to the DataLake  -bash-4.2% rucio uploadrse CNAF_USERDISKscope user.sgaravit BmmAnalysis.rootname BmmAnalysis-01register-after-ur 2022-03-05 133112,4/331 INFO Preparing upload for file BmmAnalysis.root 2022-03-05 133112,5/471 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,5/471 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,5/471 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,5/180 INFO Trucing upload with dava to CNAF_USERDISK 2022-03-05 133112,6/115 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,6/115 INFO Successfully added replication rule at CNAF_USERDISK 2022-03-05 133112,6/115 INFO Successfully upload of temporary file.dava://xfer.cr.onaf.infn.its843/DataCloud-TB/user/z 2022-03-05 133112,6/115 INFO Successfully uploaded file BmmAnalysis.root //vufs/cms.cench/rucio/x86_64/rhe17/py3/current/lib/python3.6/site-packages/url1b3/connectionpool.py:1050: InsecureRequee InsecureRequeetFansion
WP6 - Ricerca e Sviluppo, Testbed WP6 Meetings • Data Management • JFTS3 • IAM+VOMS • Modello authZ per accesso da • RUCIO notes • Testbed Data Management • Data Management Testbed: • Data Management Testbed: • Idata Management Testbec • How to test Grid storage acc • Software Management • 'datacloud-tb' VO • Object storage backhone (altern • Documentazione howto	Upload data to the DataLake -bash-4.2% rucio uploadrse CNAF_USERDISKscope user.sgaravat BmmAnalysis.rootname BmmAnalysis-01register-after-ur 2023-03-05 1331125,471 INTO Preparing upload for file BmmAnalysis.root 2023-03-05 1331125,471 INTO Successfully added replicat in Rucio catalogue at CNAF_USERDISK 2023-03-05 1331125,471 INTO Successfully added replication rule at CNAF_USERDISK 2023-03-05 1331125,471 INTO Successfully added replication rule at CNAF_USERDISK 2023-03-05 1331125,471 INTO Successfully added replication rule at CNAF_USERDISK 2023-03-05 1331125,471 INTO Successfully upload of themporary file.dvs://sfcr.cr.onaf.infn.it.8443/DataCloud-TB/user/s 2021-03-05 1331125,471 INTO Successfull uploaded file ManAulysis.root 2021-03-05 1331125,471 INTO Successfull uploaded file ManAulysis.root 2021-03-05 1331125,471 Successfull uploaded file ManAulysis.root 2021-03-05 1331125,471 INTO Successfull u
WF6 - Ricerca e Sviluppo, Testbed WF6 Meetings • Data Management > FT53 > IAM+VOMS • Modello auth2 per accesso da > RUCIO notes • Testbed Data Management > Data Management Testbed: • Data Management Testbed: • How to test Grid storage acc > Software Management > 'dateioud-tb' VO > Object storage backbone (altern > Documentazione howto > InterLink based Offloading	Upload data to the DataLake         -bash-4.2% rucio uploadrse CMF_USERDISKscope user.sgaravat BmmAnalysis.rootname BmmAnalysis-01register-after-upload for file BmmAnalysis.root         2023-03-05 1331125,4731       INTO       Preparing upload for file BmmAnalysis.root         2023-03-05 1331125,471       INTO       Successfully added replication rule at CMAP_USERDISK         2023-03-05 1331125,471       INTO       Successfully added replication rule at CMAP_USERDISK         2023-03-05 1331125,471       INTO       Successfully added replication rule at CMAP_USERDISK         2023-03-05 1331125,471       INTO       Successfully added replication rule at CMAP_USERDISK         2023-03-05 1331125,471       INTO       Successfully upload of temporary file. davis://xfer.ct.ct.ads.infn.it:8443/DataCloud-TB/user/s         2023-03-05 1331125,471       INTO       Successfull upload of temporary file. davis://xfer.ct.ct.ads.infn.it:8443/DataCloud-TB/user/s         2023-03-05 1331125,471       INTO       Successful upload of tile BmmAnalysis.root         /cvmfs/cms.cern.ch/rucio/x86_64/rhe17/p3/current/lib/python3.6/site-packages/urllib3/connectionpool.py:1059: InsecureRequest         /msecureRequestMarning,       -ash+4.28         Stind data, find data information       Successful data
WPB - Ricerca e Sviluppo, Testbed WPB Meetings Data Management FTS3 IMM+VOMS Medialio auth2 per accesso da RUCIO notes RUCIO notes Testbed Data Management Data Management Testbed: Data Management Testbec: Data Management Testbec	Upload data to the DataLake
WF6 - Ricerca e Sviluppo, Testbed WF6 Meetings Defat Management TESS IdAN+VOMS Modelio auth2 per accesso da RUCIO notes Testbed Data Management Data Management Testbed: Data Management Testbed: How to test Grid storage acc Software Management Octave Management Octave Management Decumentazione howto InterLink based Offloading Outdated documentation	Upload data to the DataLake         = bash-4.28 rucio uploadrse CNAF_USERDISKscope user.sgaravit BmmAnalysis.rootname BmmAnalysis-01register-after-up         2021-03-05 133112,24,331       INFO         Preparing upload def replication rule at CNAF_USERDISK         2021-03-05 133112,24,71       INFO         2021-03-05 133112,64,115       INFO     <
WF6 - Ricerca e Sviluppo, Testbed WF6 Meetings Data Management FTS3 IGM+VOMS Modello authZ per accesso da UCIC notes Testbed Data Management Testbed: Data Management Testbed: Data Management Testbed: How to test Grid storage acc Software Management Otion to the Soft Soft Soft Soft Soft Soft Soft Soft	Upload data to the DataLake







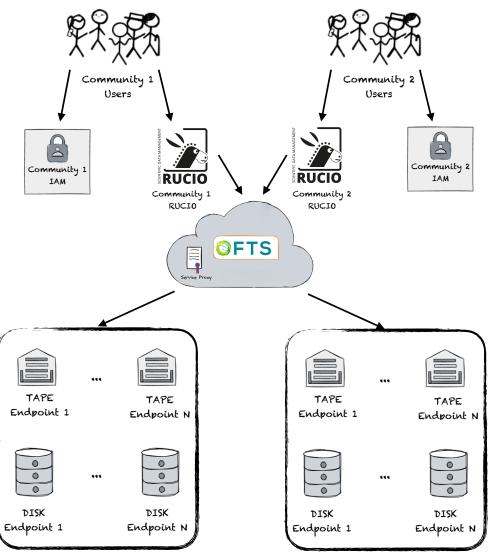


# Rucio as a Service for communities

We aim to add Rucio as a a service in the portfolio of INFN Cloud applications, supporting scientific communities with infrastructure and know-how.

We want to **simplify the deployment** of a community RUCIO istance, e.g. Rucio as a service among the services offered by the INFN Cloud dashboard with the following characteristics:

- Rucio community-specific, managed by them;
- INDIGO IAM community-specific, managing users and policies;
- Centrally managed database in high-availability;
- Centrally managed multi-VO FTS;
- Federation of pledged storages.











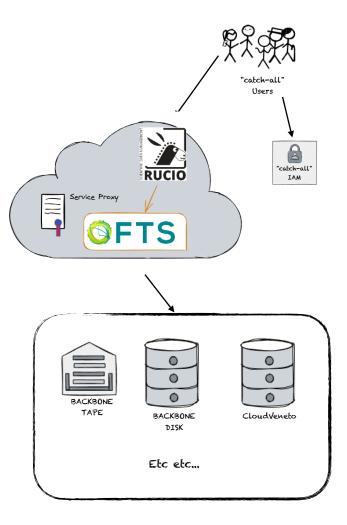
# Centrally managed Rucio for single users

We also plan to have a «central» Rucio instance **to serve also single users**, who can take advantage of a federation of storages created specifically for this purpose.

A single user will be able to use Rucio to manage «personal» data, with a single interface in front of several storage systems  $\rightarrow$  makes it simple.

- Rucio centrally managed;
- INDIGO IAM centrally managed;
- Centrally managed database in high-availability;
- Centrally managed multi-VO FTS;
- Federation of dedicated INFN storages.

#### Are there already similar experiences?











### Summary and final remarks

- Our national computing infrastructure is further **consolidating and improving**;
- We recognize that "Data Management" has a central role and we choosed Rucio as a middleware to abstract the logical layer from the physical one;
- We're moving towards a model where Rucio is within reach of the single user, not only communities;

From the Proof of Concept phase, we want to move to the production phase. Some of the next steps:

- Transfers FTS with token: is it production ready?
- Replace rucio-ui with rucio-webui when the on-going refactoring is done;
- Rucio JupyterLab interface;
- Monitoring: suggestions on tools and technologies that could be used are welcome.

In order to avoid the maintenance of custom Rucio setup, we are **willing to contribute upstream** with possible contributions (i.e. via PR), following the Rucio project's best practices.









