

# iRODS-based REPO

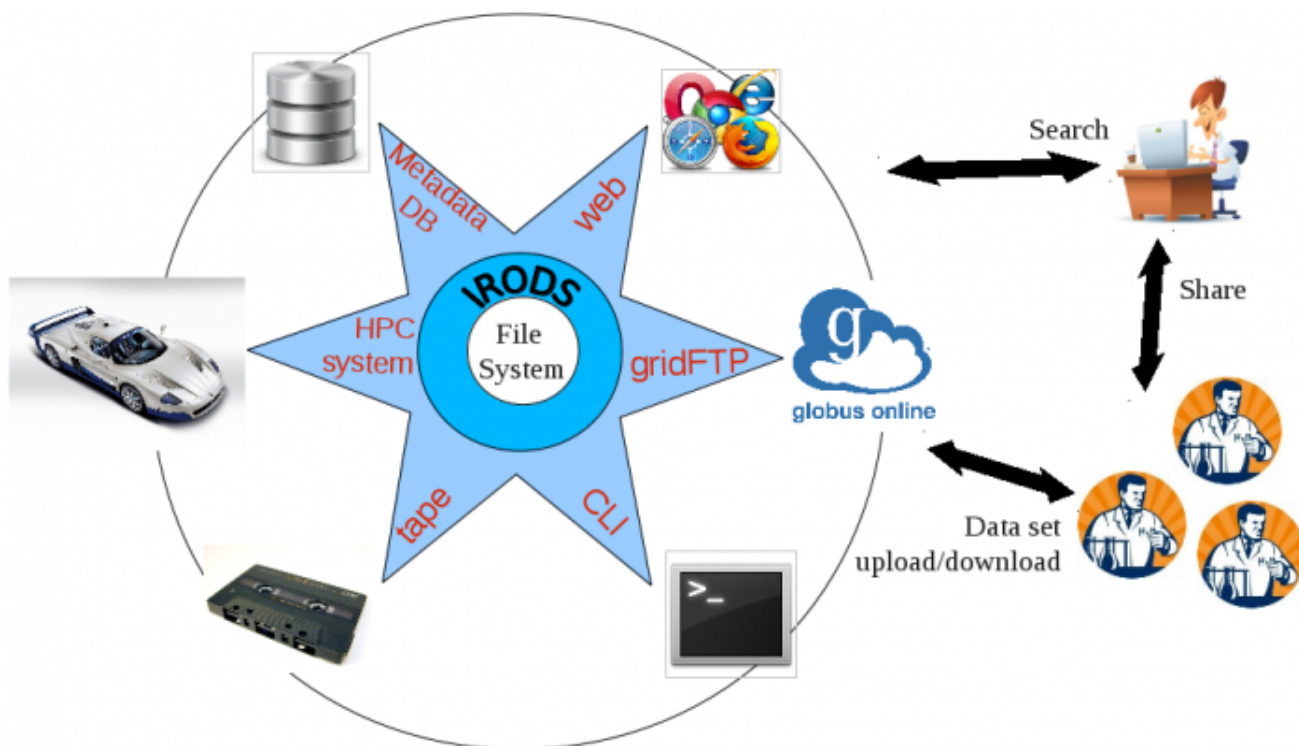
- [Overview](#)
- [How to request a REPO space](#)
- [How to access the REPO space](#)
  - iRODS commands
    - [Configuration](#)
    - [Authentication](#)
  - [GridFTP clients](#)

## Overview

REPO is a Cineca service, implemented through [iRODS](#) (Integrated Rule-Oriented Data System), for the management of long lasting data.

This service aims to store and maintain scientific data sets and it is built in a way that allows a user to safely back-up data and at the same time manage them through a variety of clients, such as web browser, graphical desktop and command line interfaces.

It relies on plain filesystems to store data files and on databases to store the metadata. The service's architecture has been carefully designed to scale to millions of files and petabytes of data, joining robustness and versatility, and to offer to the scientific communities a complete set of features to manage the data life-cycle:



The links for the Data Repository interfaces are listed at the URL <https://www.repo.cineca.it>.

**Upload/Download:** the system supports high performance transfer protocols like GridFTP, or iRODS multi-threads transfer mechanism.

- The GridFTP protocol is supported as described in this [page](#); the GridFTP interface for iRODS is at address: **data.pico.cineca.it:2811**.
- The iRODS commands, official documentation available at <https://docs.irods.org/master/icommands/user/>, but look down to know how to configure them.

**Metadata management:** each object can be associated with specific metadata represented as triplets (name, value, unit), or simply tagged and commented. This operation can be performed at any time, not just before the first upload.

**Preservation:** the long-term accessibility is granted by means of a seamless archiving process, which is able to move the collections of data from the on-line storage space to a tape based off-line space and back, according to general or per-project policies.

**Stage-in/stage-out:** the service is enabled to move data sets, requested as input for computations, towards the HPC machines' local storage space, commonly named "scratch", and backwards as soon as the results are available.

**Sharing:** the capability to share single data objects or whole collections is implemented via a unix-like ownership model, which allows to make them accessible to single users or groups. Moreover, a ticket based approach is used to provide temporary access tokens with limited rights.

**Searching:** the data are indexed and the searches can be based on the objects location or on the associated metadata.

## How to request a REPO space

Archiving on the REPO area is managed through the DRES (Data RESource) space, as discussed in the "[Data Storage Resource](#)" document. You can require a DRES of REPO type by sending an email to [superc@cineca.it](mailto:superc@cineca.it).

## How to access the REPO space

There are three different ways to access data in the REPO:

- iRODS commands
- gridftp clients, such as globus-url-copy or Globus Online
- WebDAV protocol

## iRODS commands

### Configuration

You can use the iCommands from CINECA HPC machines (MARCONI, MARCONI100 and GALILEO100) or from your local linux machine.

1) download iCommands

- On MARCONI the iCommands are available without any module to load
- On MARCONI100 and GALILEO100 the iCommands are available with a module. So on the login node, type:

```
▪ $ module load icommands  
  
$ iinit  
  
$ ils
```

- On your local linux machine you have to install the iCommands downloading it from <http://irods.org/download/> . Packages .deb and .rpm (CentOS7, Ubuntu16 and Ubuntu18) are provided.

If you want to install the iCommands with support for PAM authentication on your linux machine from source code, you have to download it from <https://github.com/irods/irods> .

2) download the file [chain.pem](#) (click to download)

3) create the **.irods/irods\_environment.json** config file in the home directory of the system where you use the icommand (MARCONI, MARCONI100 or GALILEO100, your local linux machine):

```
{
  "irods_host": "data.repo.cineca.it",
  "irods_port": 1247,
  "irods_default_resource": "cinecaRes1",
  "irods_home": "/CINECA01/home/your-group/your-username",
  "irods_cwd": "/CINECA01/home/your-group/your-username",
  "irods_user_name": "your-username",
  "irods_zone_name": "CINECA01",
  "irods_client_server_negotiation": "request_server_negotiation",
  "irods_client_server_policy": "CS_NEG_REFUSE",
  "irods_encryption_key_size": 32,
  "irods_encryption_salt_size": 8,
  "irods_encryption_num_hash_rounds": 16,
  "irods_encryption_algorithm": "AES-256-CBC",
  "irods_default_hash_scheme": "MD5",
  "irods_match_hash_policy": "compatible",
  "irods_server_control_plane_port": 1248,
  "irods_server_control_plane_key": "TEMPORARY_32byte_ctrl_plane_key",
  "irods_server_control_plane_encryption_num_hash_rounds": 16,
  "irods_server_control_plane_encryption_algorithm": "AES-256-CBC",
  "irods_maximum_size_for_single_buffer_in_megabytes": 32,
  "irods_default_number_of_transfer_threads": 4,
  "irods_transfer_buffer_size_for_parallel_transfer_in_megabytes": 4,
  "irods_authentication_scheme": "PAM",
  "irods_ssl_certificate_chain_file": "/path/to/.irods/chain.pem",
  "irods_ssl_ca_certificate_file": "/path/to/.irods/chain.pem",
  "irods_ssl_verify_server": "cert"
}
```

3) type the command **iinit** the first time you use irods . On default, the PAM authentication method is enabled ("irods\_authentication\_scheme parameter" in the json configuration file), so the password of your hpc username will be requested.

Note the after some times (days...), you will need retype the **iinit** command to use the icommands.

4) operate in REPO space, using the icommands. The documentation about the IRODS commands is available [at this link](#).

## Authentication

The installation of iRODS in CINECA supports two authentication mechanisms: username password (PAM) previously seen and [GSI](#) (e.g. X.509 Certificate)

If you want to use GSI authentication instead of PAM authentication, please replace the line "irods\_authentication\_scheme": "PAM" in your ".irods/irods\_environment.json" file with:

```
"irods_authentication_scheme": "GSI",
```

In this case the GSI support should be enabled in the iCommands. The GridFTP server address is: **gftp.repo.cineca.it:2811**

## GridFTP clients

In order to access your REPO space through a GridFTP client as globus-url-copy or Globus Online consult these web pages: [globus-url-copy](#) or [Globus Web App](#).