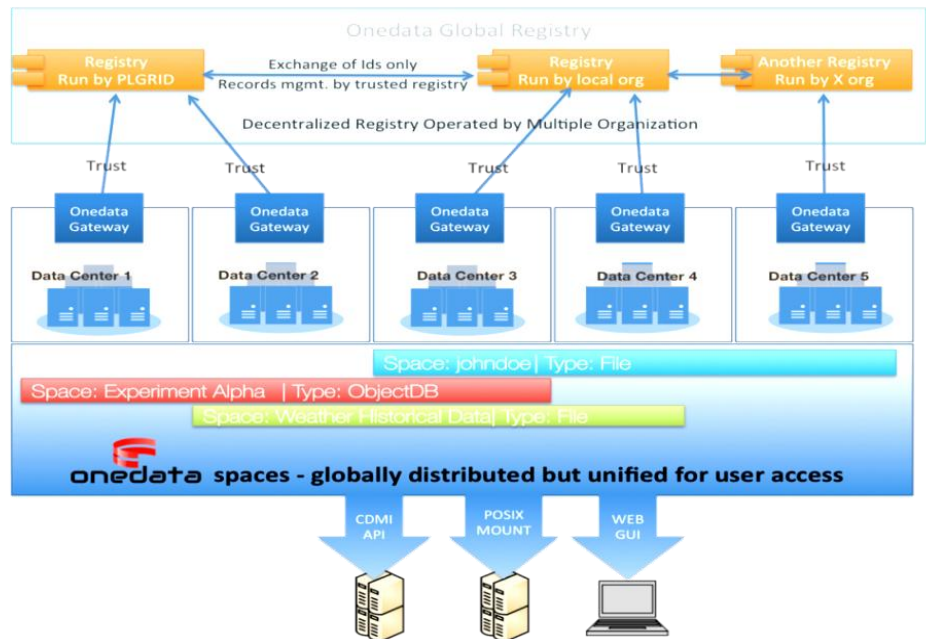


onedata - unified and efficient access to your data



onedata is a system that provides a unified and efficient access to data stored in organizationally distributed environments. A unique feature of onedata is the ability to elastically mix or integrate data from different domains/sources, e.g., from public and private data sources, and from different type of sources, e.g., file systems, databases or cloud storage, into one logical user's defined data space. onedata introduces a very little additional (less than 5%) overhead to the performance of the underlying storage, giving in return a flexible environment for users and system administrators.



Functionalities for end-users

- Provides a uniform and coherent view on all data stored on the storage systems distributed across the infrastructure. Using onedata all users' computing processes see all data stored in all sites of organizations. If the needed data is not stored locally, it is transparently migrated by the system. The user can also mount onedata on his/her PC and access the data as if it was stored on a local hard drive.
- Supports working in groups by creation of an easy-to-use shared workspace for each group. Users are able to share data inside a group by simply moving the data to an appropriate directory.
- Supports publishing data. Until canceled by the owner, anyone is able to download the file using the obtained URL. Serves data efficiently. The system is designed to minimize overheads and provide data from remote storages as fast as possible.

However, in most cases some simplifications of the system for the users increase storage management tasks to be performed by administrators. Hence, onedata is designed to provide functionalities that also facilitate administrators' work and makes the entire storage system more flexible.



Functionalities for the storage administrators

- Gathering and visualizing of the infrastructure state monitoring data – each storage system supervised by onedata is monitored to provide the administrators with an insight into storage utilization.
- Rule-based data management - automatic data management can be performed on the basis of rules specified by administrators, e.g. the rarely used data can be automatically migrated from fast to cost-effective storage. The optimization of use of storage resources is transparent to the users.
- Data protection from unauthorized access - the system is integrated with Grid and Linux security systems to protect the data.

Functionalities for developers

- onedata provides a REST API which allows for direct interfacing from third party applications and services.
- onedata provides standard-based data access using CDMI protocol.
- Applications and services can interact with onedata on behalf of the users with standard X509 proxy certificates, or so there is no need to register services that can cooperate with onedata.
- onedata can show the required data location and estimate the time of transfer to a particular site

3-rd place in the Big Data Challenge of the Copernicus Masters – 2014 for the onedata team



onedata is the result of research projects: PLGrid Plus and PLGrid Core.



PLGrid PLUS is focused on integrated domain specific research communities and helping them use the distributed high performance computing resources, to improve their everyday research activities.

The PLGrid Core project is an initiative of the development of specialized technical Competence Centre in the field of distributed computing infrastructures, with particular emphasis on grid technologies, cloud computing and infrastructure supporting calculations on large data.



onedata is available for all PLGrid users at: <https://veilfs.plgrid.pl>

Developed by:



INNOVATIVE ECONOMY
NATIONAL COHESION STRATEGY

