The LEXIS Platform

Easy access to heterogenous computational workflows execution

Mohamad Hayek (LRZ)
Martin Golasowski (IT4I)

iRODS User Group Meeting 2022

Outline

- The LEXIS platform
- LEXIS Distributed Data Infrastructure
- Authentication with OpenID
- iRODS-Keycloak fine grained access control
- Staging API
- Automated Robot tests
- Bandwidth measurements

THE LEXIS PLATFORM: Timeline

2018 – LEXIS Platform concept

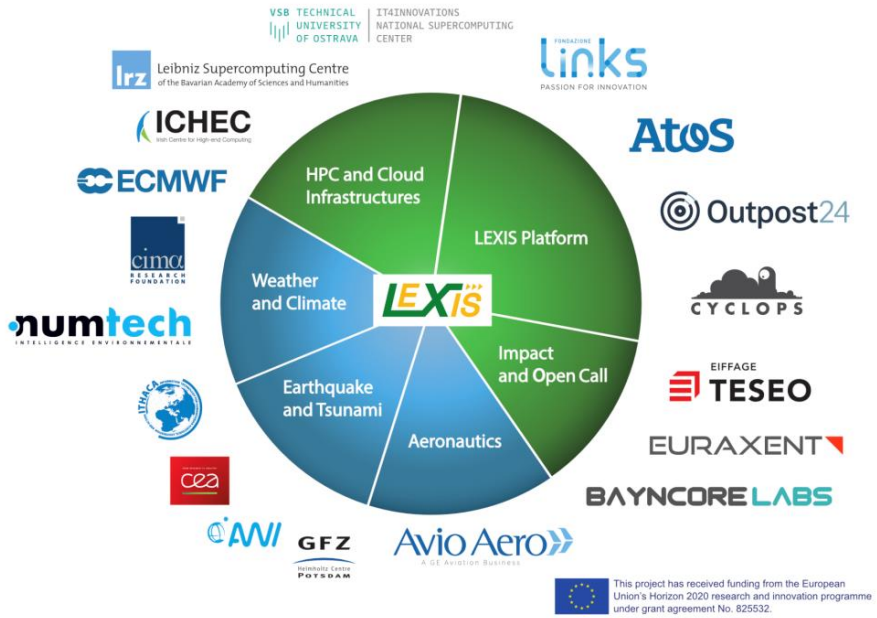
- Technical Concept
- Creation of a Consortium – 17 members, from Germany, Czech Republic, France, Ireland, UK, Italy, Switzerland
- Submission of the LEXIS Project to the EU Commission – Programme H2020
- Approval by the EU Commission - total budget circa 14 millions € - Grant agreement 825532
- Coordinator: IT4Innovations – National Supercomputing Centre of Czech Republic (Dr. Jan Martinovic)

2019 – Start of the project – 3 years timeline

- Project ending December 2021
- Completion validated by the EU Commission June 2022
- Exploitation post-end of project: starting now

2022 – Onward

- Technical tuning & development
- Enrolling new partners (Infrastructures, Data, Services)
- Structuration of legal entity
- Funding



THE LEXIS PLATFORM: an Ecosystem for an Ecosystem for the Digital Continuum

From inception LEXIS has been built by an ecosystem made of representants of:

- Industries, (4)
- Research Organisations (6)
- Super computing centres (4)
- Service companies (4)
- SMEs and Start-ups (5)

The LEXIS Platform will further develop by capitalising on:

- Existing partners,
- Developing the number of infrastructures as members of the federation,
- Strongly reinforcing services by welcoming new Service partners,
- Increase technical ways for integration, interconnections and cooperation with the addition of new components to the federation via development of APIs,
- Digital Sovereignty European framework (GAIA-X or else).

IT4I
Supercomputing Centre

Atos
Industry

LINKS
Research Organisation

TESEO
Industry

CEA
Research Organisation

LRZ
Supercomputing Centre

ECMWF
Supercomputing Centre

ITHACA
Research Organisation

CIMA
Research Organisation

AVIO Aero
Industry

GFZ
Research Organisation

AWI
Research Organisation

Outpost24
SME

Cyclops Labs
SME

BAYNCORE
SME

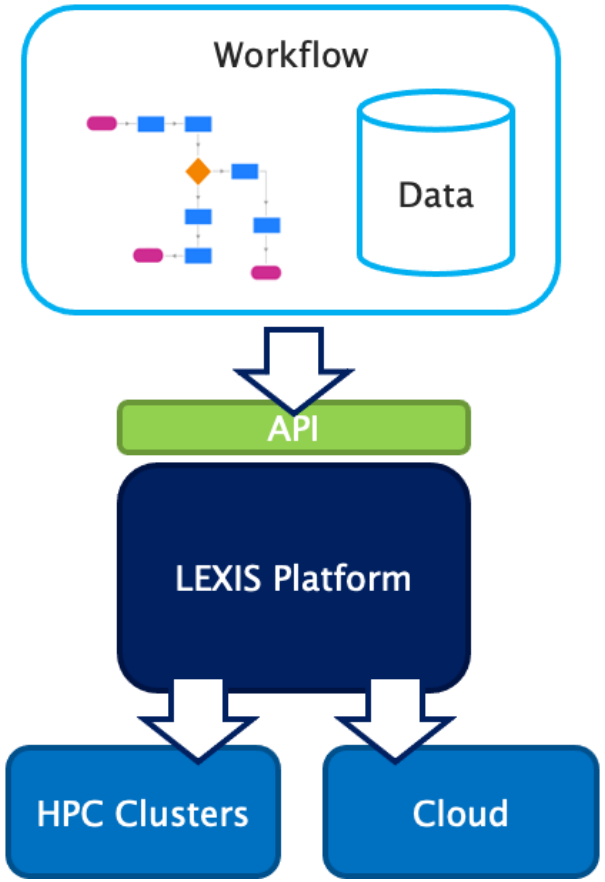
NUMTECH
SME

ICHEC
Supercomputing Centre

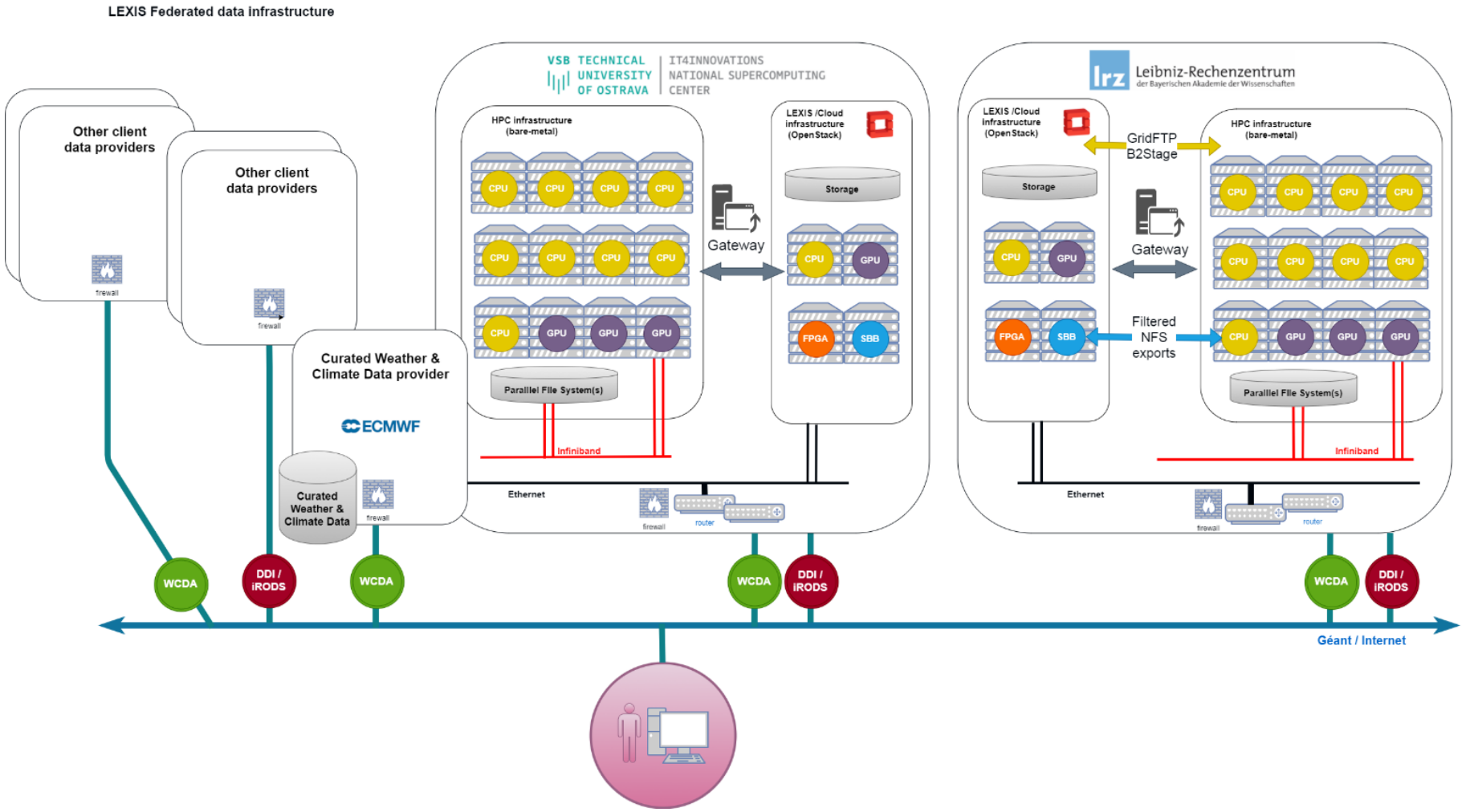
EURAXENT
SME

Key Points of LEXIS Platform

- **Dynamic, complex Cloud- & High-Performance-Computing / Big Data workflows**
  - Orchestration in geographical federation with *YORC, HEAppE*
  - Real-time deadline-aware workflows, etc.
- **Cross-site (meta-)data federation**
  - Distributed data management and data discovery with *EUDAT/iRODS*
  - Data transfers accelerated by Burst Buffer nodes; FPGAs/GPUs for on-line processing
- **Web portal and interfaces for workflow set-up / execution**
  - Unified access to all services via *Keycloak*-based LEXIS AAI
- **Easy HPC/Cloud access for SMEs/Industry – Big Data for everyone**
  - HPC-as-a-Service approach
  - Control over resource usage
  - Fine-grained accounting and billing for multiple HPC centres with *CYCLOPS*



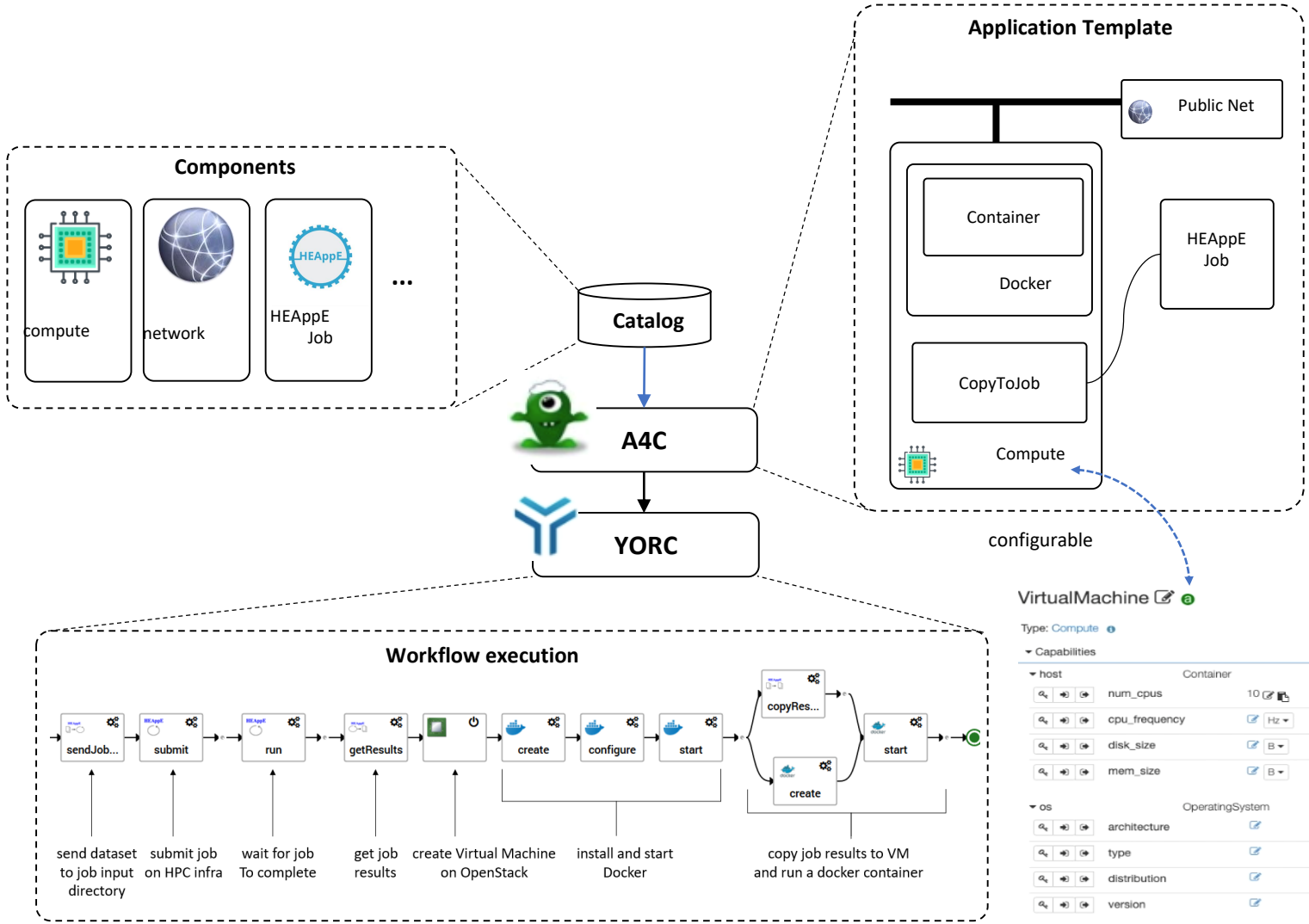
Platform architecture overview



- Federation of European computing centres
- HPC & Cloud service providers, Data providers
- Unified & distributed data management
- Orchestration
- Federated Authentication & Authorization Infrastructure (AAI)
- Masking of technical and operational differences across organizations

Orchestration service & workflow management

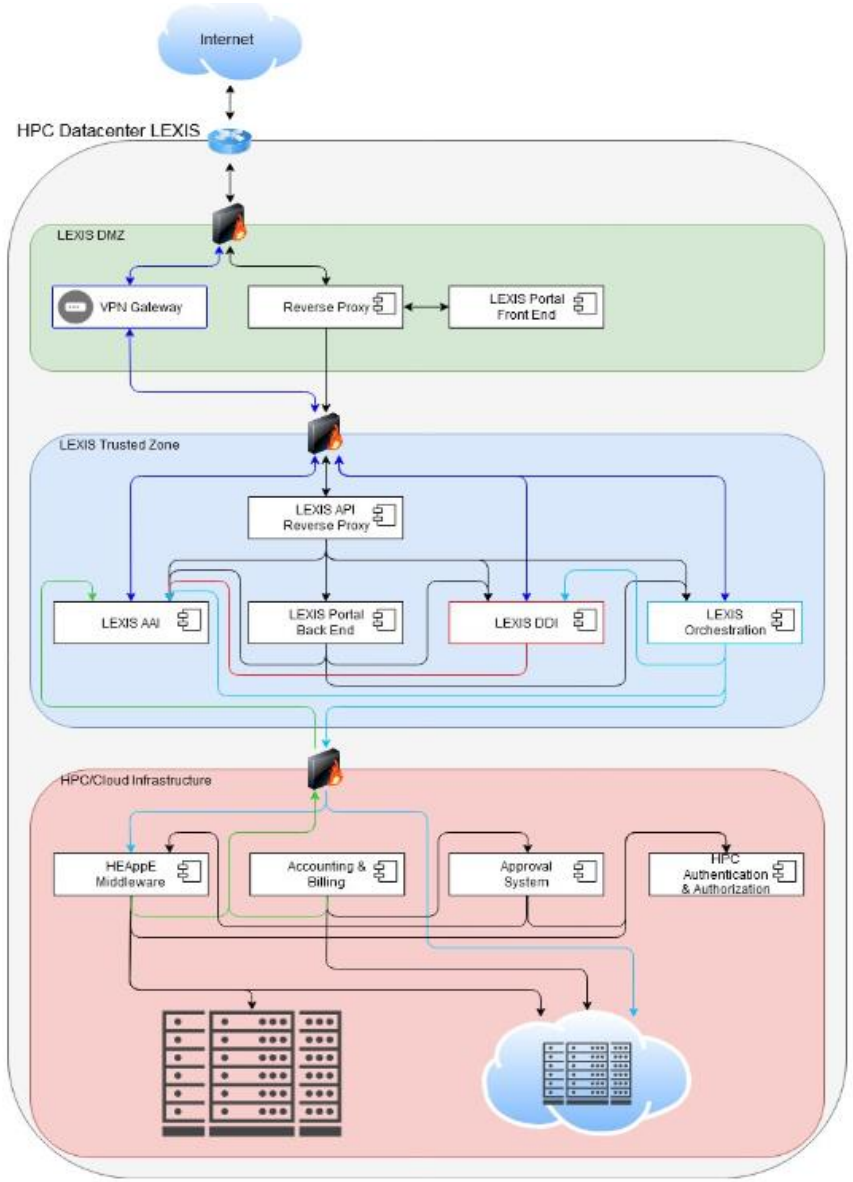
- Execution on geographically distributed HPC and Cloud resources
  - **Cloud**: via OpenStack built-in interface
  - **HPC**: job execution is mediated by HEAppE middleware
- Data management and orchestration policies
  - Leverage the LEXIS DDI service for an effective **data transfer** between systems
  - Placement of **workflow tasks** on the most suitable resource



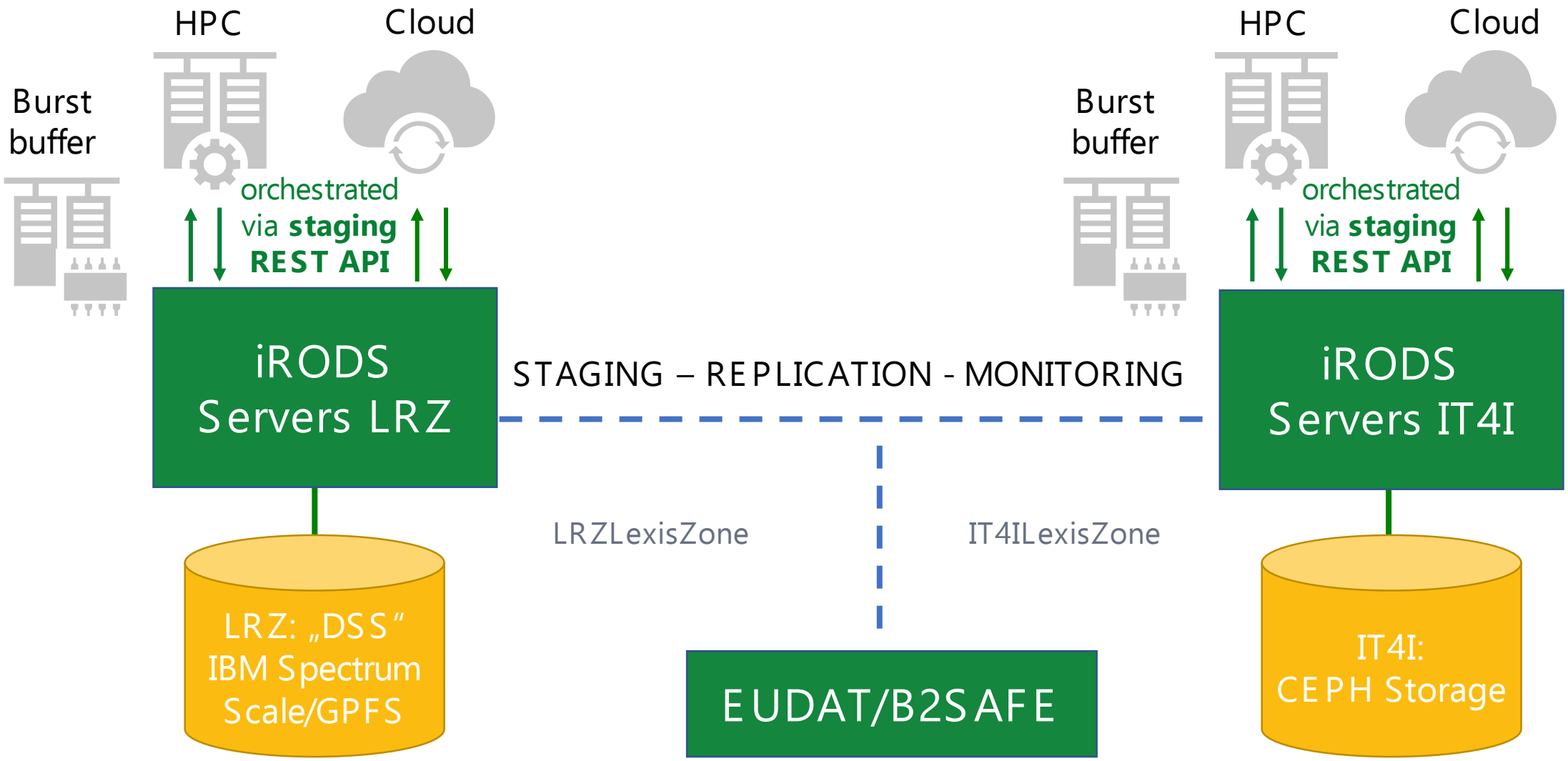
https://github.com/alien4cloud/alien4cloud
https://github.com/ystia
http://heappe.eu

LEXIS Security requirements

- **Custom AAI solution with trusted access to HPC with PI approval**
- Security-by-design
  - **Zero trust, minimal attack surface, separation of concerns**
- Modern frameworks
- HPC infrastructures are protected
  - **Isolated by the HEAppE middleware (developed in IT4I)**
  - **Deployed in both IT4I and LRZ**
- Flexible
  - **Blurs differences between HPC centres**
  - **Provides SSO across the LEXIS federation**



Distributed data infrastructure



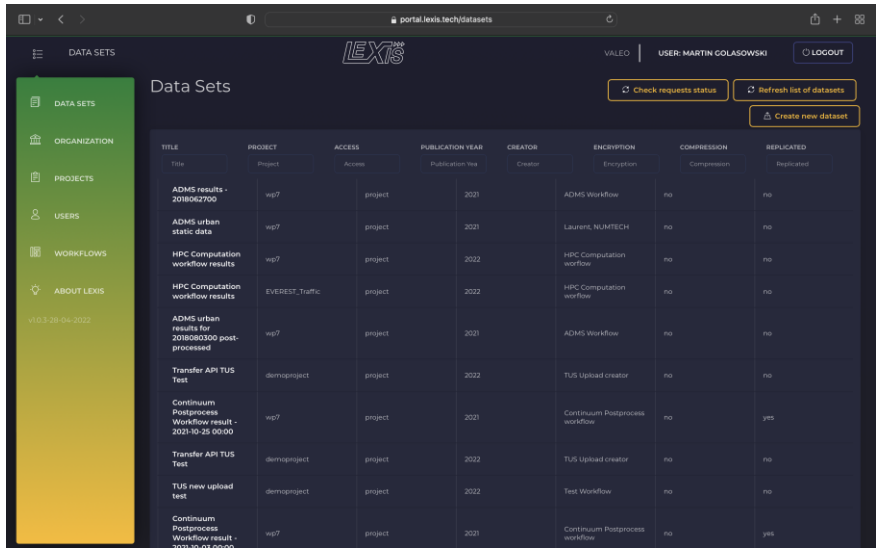
Zone structure organization

One zone per federated center

Hashed project name

The dataset

/ LexisZone/project/proj<shortname hash>/<dataset uuid>




```
$ ils /IT4ILexisZone
/IT4ILexisZone:
        C- /IT4ILexisZone/home
        C- /IT4ILexisZone/project
        C- /IT4ILexisZone/public
        C- /IT4ILexisZone/trash
        C- /IT4ILexisZone/user
```

iRODS user home

Project level datasets

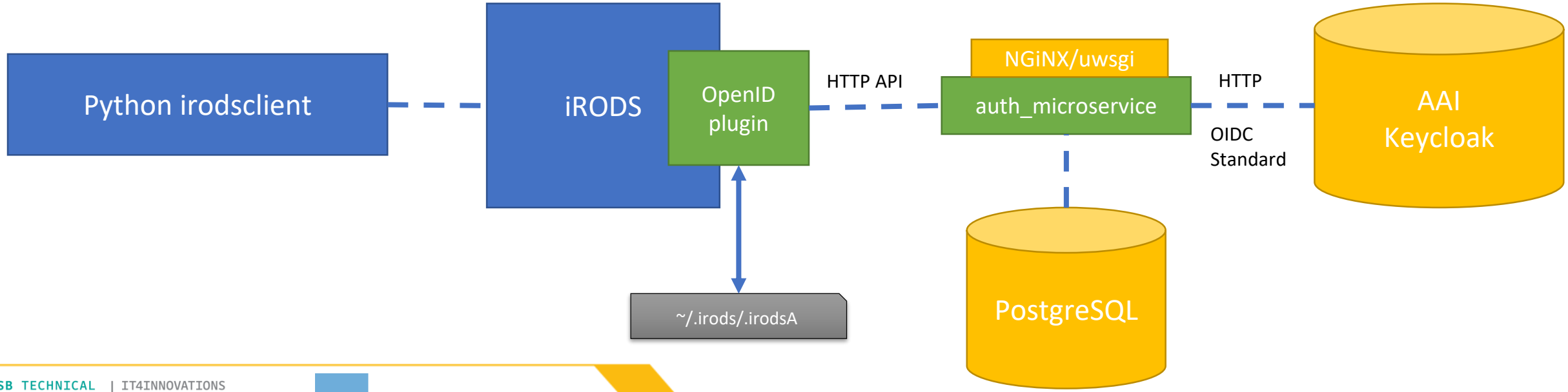
Public datasets

The trash

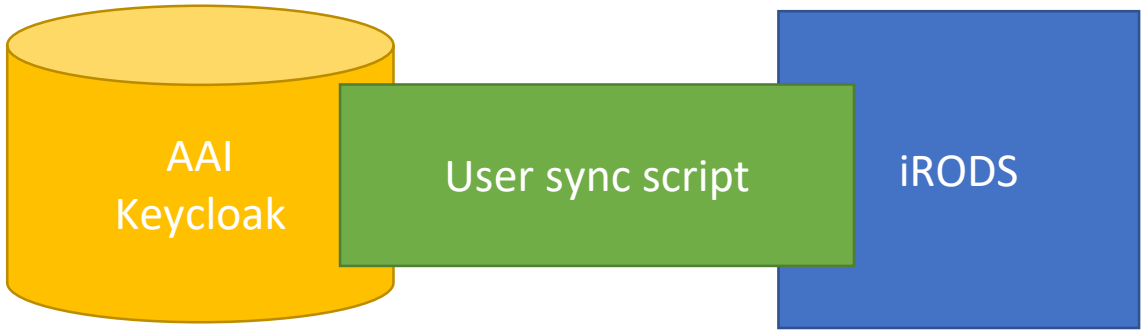
User datasets per project

Using OpenID in iRODS and REST APIs

- Current state
  - auth_microservice broker (https://github.com/lexis-project/auth_microservice)
  - irods_openid_plugin (https://github.com/lexis-project/irods_auth_plugin_openid)
  - patched Python irodsclient (https://github.com/lexis-project/python-irodsclient)
- Obsolete?
- How about iRODS 4.3.0?



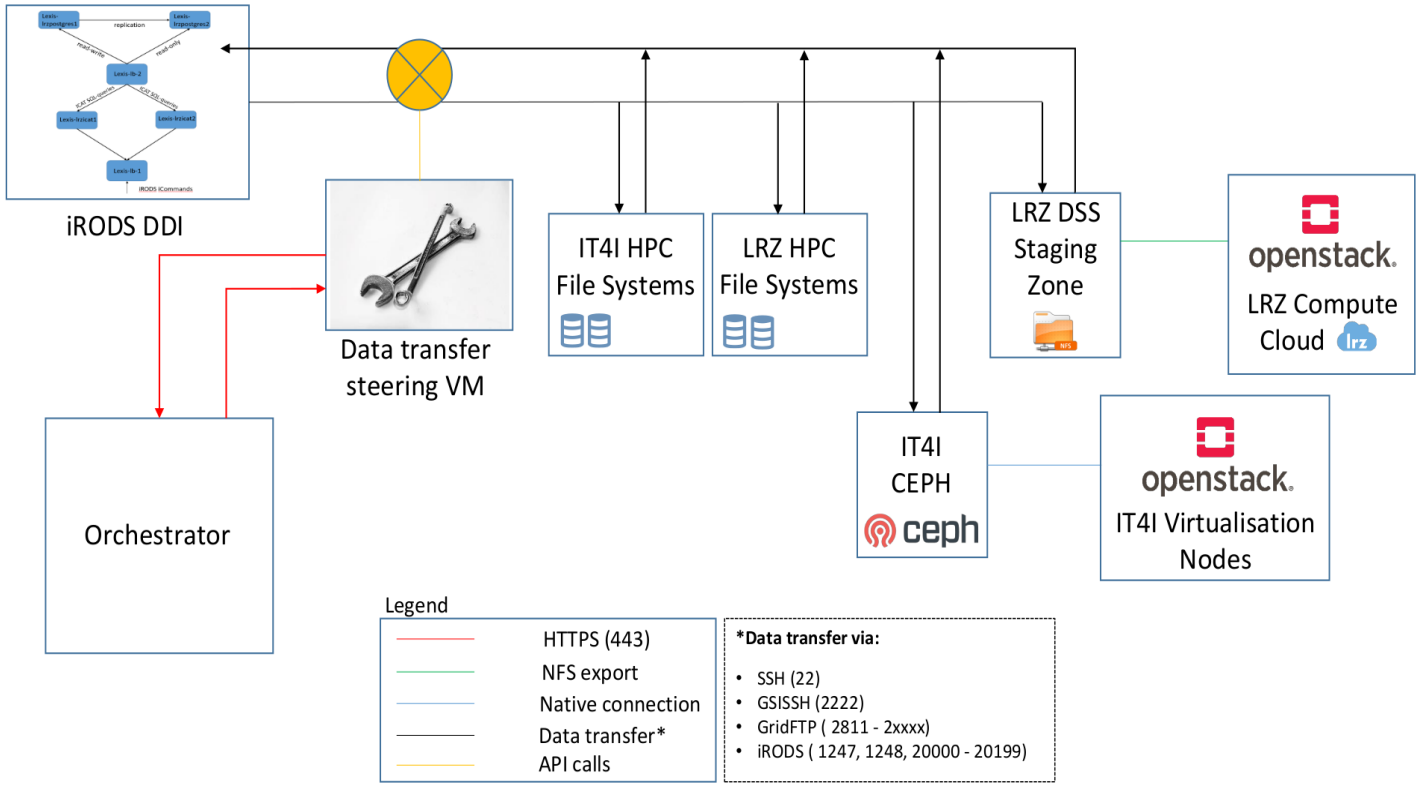
Fine-grained access control to the iRODS



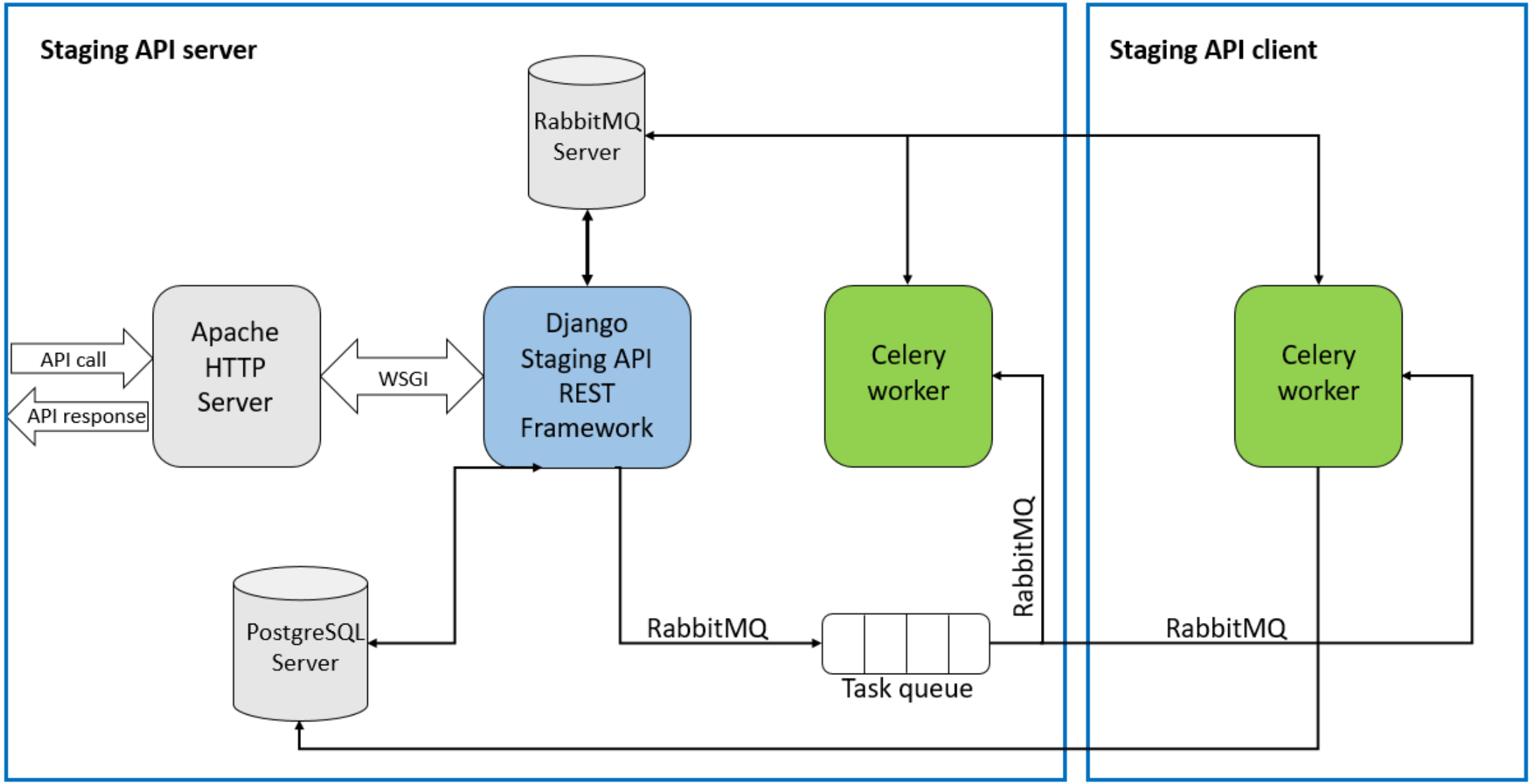
- Users belong to groups according to RBAC matrix
  - DAT_READ / DAT_WRITE – R/W access to datasets
  - DAT_LIST – Listing datasets incl. metadata
  - DAT_PUBLISH - Access to B2SHARE, write permissions to /public collection
- Mapping to iRODS groups
  - Two groups with R/W access and _mgr with publish
- User sync script runs every 3 minutes and performs a sync

LEXIS Staging API

- Django based RESTful API
- Scope: LEXIS orchestrator can move data by simple HTTP request
  - between iRODS,
  - Cloud, and
  - HPC resources at all LEXIS centers.
- Uses LEXIS AAI and the HEAppE middleware to authenticate the requests and the access to the resources
- Deploys a queuing system using Celery and RabbitMQ to allow asynchronous requests.



LEXIS Staging API - Architecture



LEXIS Staging API - Classes

- The source/target combinations are covered by 10 staging classes and served by 7 celery tasks.
- Each center has their separate deployment of the celery worker.
- Based on the source/target combination, the task is pushed to a specific queue.

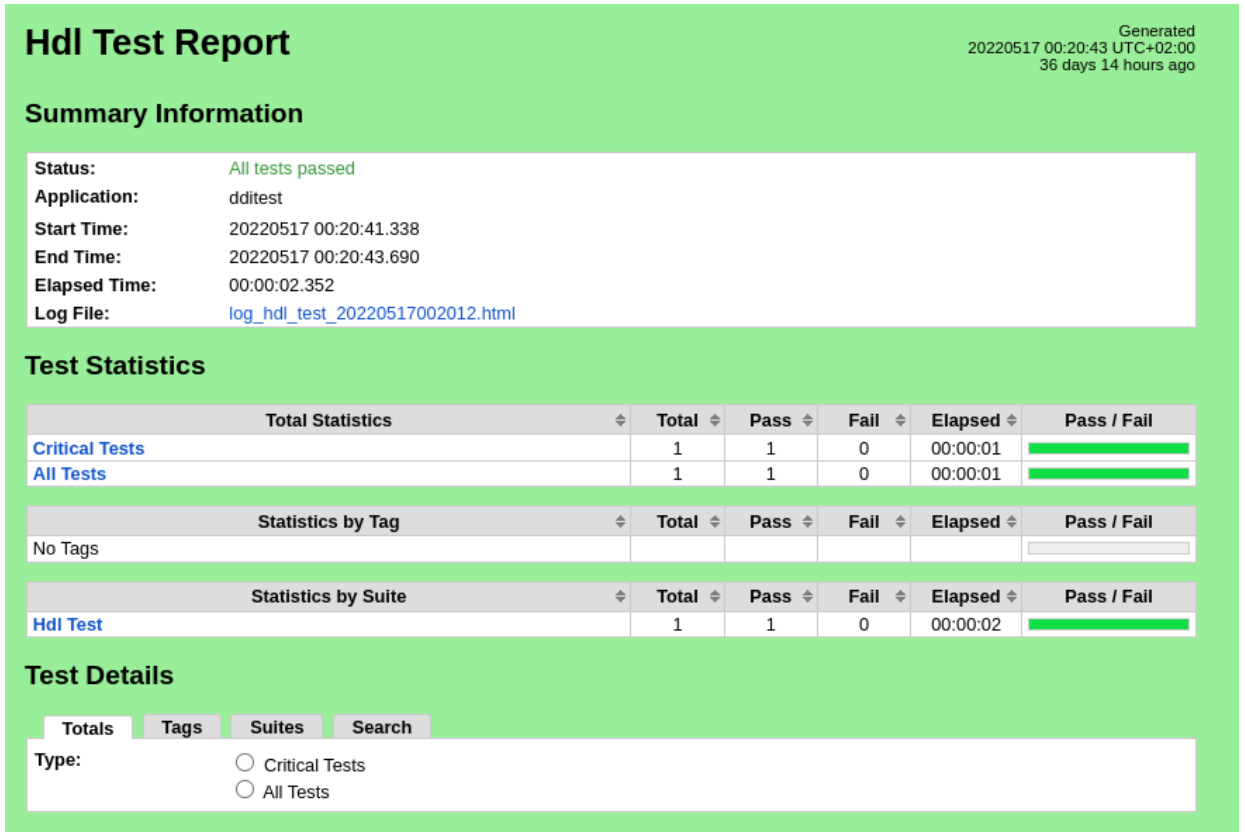
Class	Source	Target	Celery task
1	iRODS	Local cloud/staging area	1
2	Local cloud/staging area	iRODS	2
3	Local cloud/staging area	Local cloud/staging area	3
4	iRODS	local HPC	4
5	local HPC	iRODS	5
6	Local cloud/staging area	local HPC	6
7	local HPC	Local cloud/staging area	7
8	Local cloud/staging area	remote HPC	2 & 4
9	local HPC	Remote cloud/staging area	5 & 1
10	Local cloud/staging area	Remote cloud/staging area	2 & 1

Automated Robot tests

- Uses Python and the Robot framework to run basic checks
- The suite is executed every 10 or 20 minutes
- Tests include:
  - Local iCAT test: obtains session with password and OpenID token and performs a file transfer
  - Federation test: tests file transfer between federated zones with OpenID token
  - Authentication test: obtains and validates OpenID token from LEXIS AAI
  - APIs test: verifies function of DDI APIs
  - Handle test: verifies B2SAFE/HANDLE.NET server function.
- Generates an HTML report with the results
- Sends alert via Slack supported API if a test fails

Automated Robot tests

HTML Reports



Hdl Test Report

Generated
20220517 00:20:43 UTC+02:00
36 days 14 hours ago

Summary Information

Status:	All tests passed
Application:	dditest
Start Time:	20220517 00:20:41.338
End Time:	20220517 00:20:43.690
Elapsed Time:	00:00:02.352
Log File:	log_hdl_test_20220517002012.html

Test Statistics

Total Statistics	Total	Pass	Fail	Elapsed	Pass / Fail
Critical Tests	1	1	0	00:00:01	
All Tests	1	1	0	00:00:01	

Statistics by Tag	Total	Pass	Fail	Elapsed	Pass / Fail
No Tags					

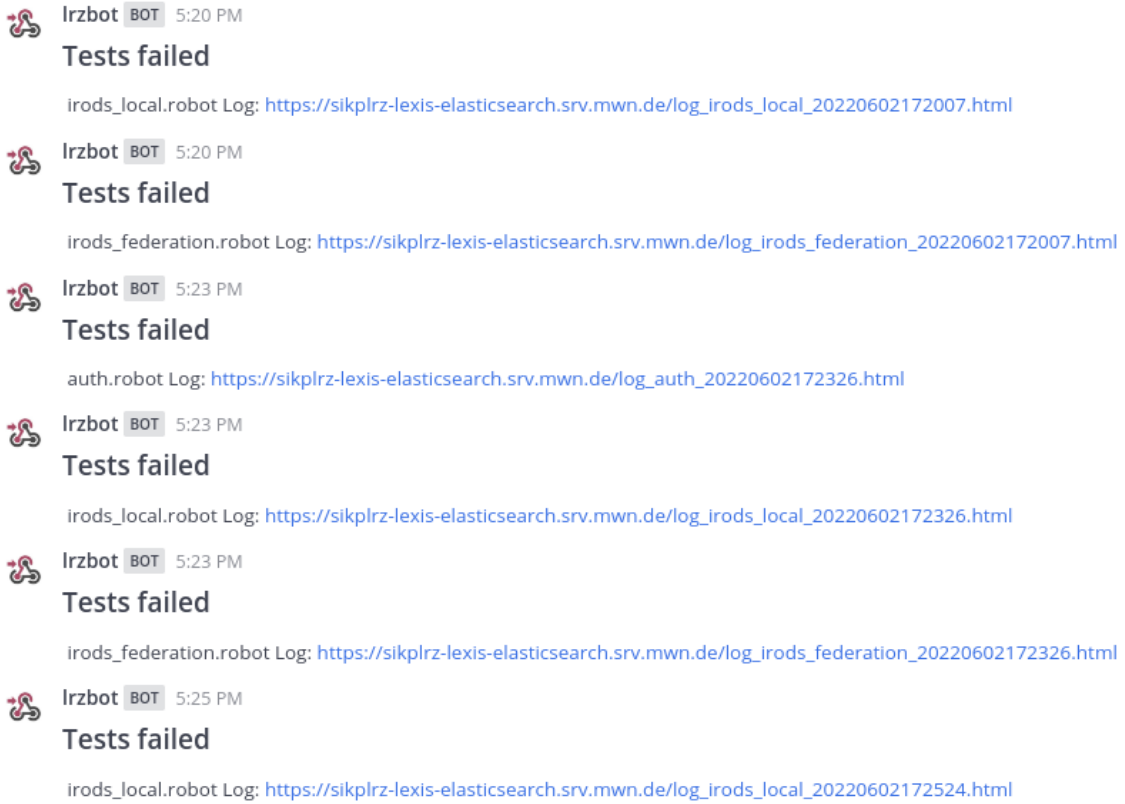
Statistics by Suite	Total	Pass	Fail	Elapsed	Pass / Fail
Hdl Test	1	1	0	00:00:02	

Test Details

Totals | Tags | Suites | Search

Type: ○ Critical Tests ○ All Tests

Mattermost chatbot for alerting



lrzbot BOT 5:20 PM
Tests failed
irods_local.robot Log: https://sikplrz-lexis-elasticsearch.srv.mwn.de/log_irods_local_20220602172007.html

lrzbot BOT 5:20 PM
Tests failed
irods_federation.robot Log: https://sikplrz-lexis-elasticsearch.srv.mwn.de/log_irods_federation_20220602172007.html

lrzbot BOT 5:23 PM
Tests failed
auth.robot Log: https://sikplrz-lexis-elasticsearch.srv.mwn.de/log_auth_20220602172326.html

lrzbot BOT 5:23 PM
Tests failed
irods_local.robot Log: https://sikplrz-lexis-elasticsearch.srv.mwn.de/log_irods_local_20220602172326.html

lrzbot BOT 5:23 PM
Tests failed
irods_federation.robot Log: https://sikplrz-lexis-elasticsearch.srv.mwn.de/log_irods_federation_20220602172326.html

lrzbot BOT 5:25 PM
Tests failed
irods_local.robot Log: https://sikplrz-lexis-elasticsearch.srv.mwn.de/log_irods_local_20220602172524.html

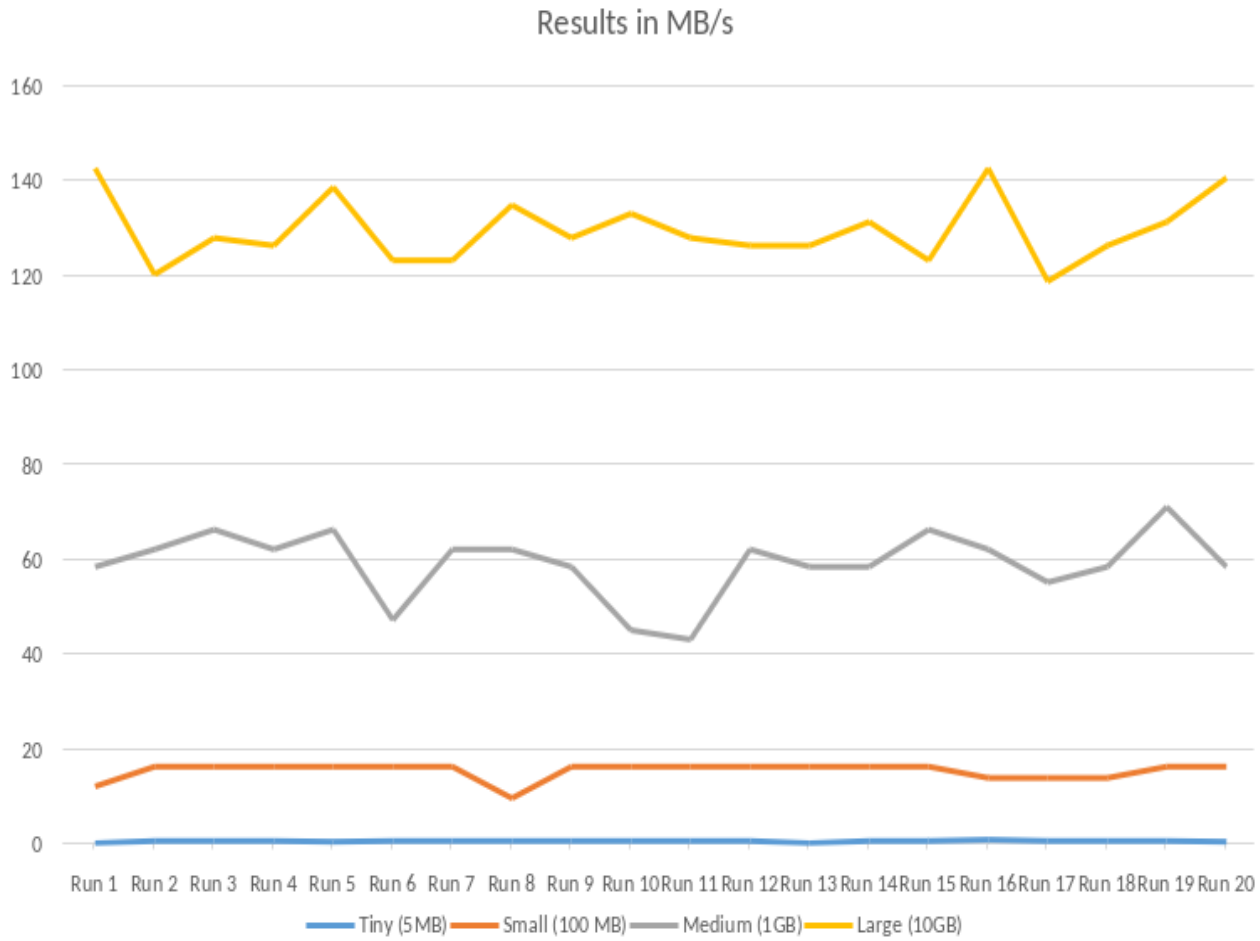
Speed measurement between LRZ and IT4I

- A series of bandwidth tests were executed between LRZ and IT4I to help estimate the data transfer time between the 2 centers.
- The tests were executed 20 times in each direction and with different data sizes
- The tests included:
  - iRODS to iRODS via icp
  - iRODS to iRODS via EUDAT B2SAFE
  - iget from local zone via python client
  - iput to local zone via python client
  - iget from remote zone via python client
  - iput to remote zone via python client

Speed measurement

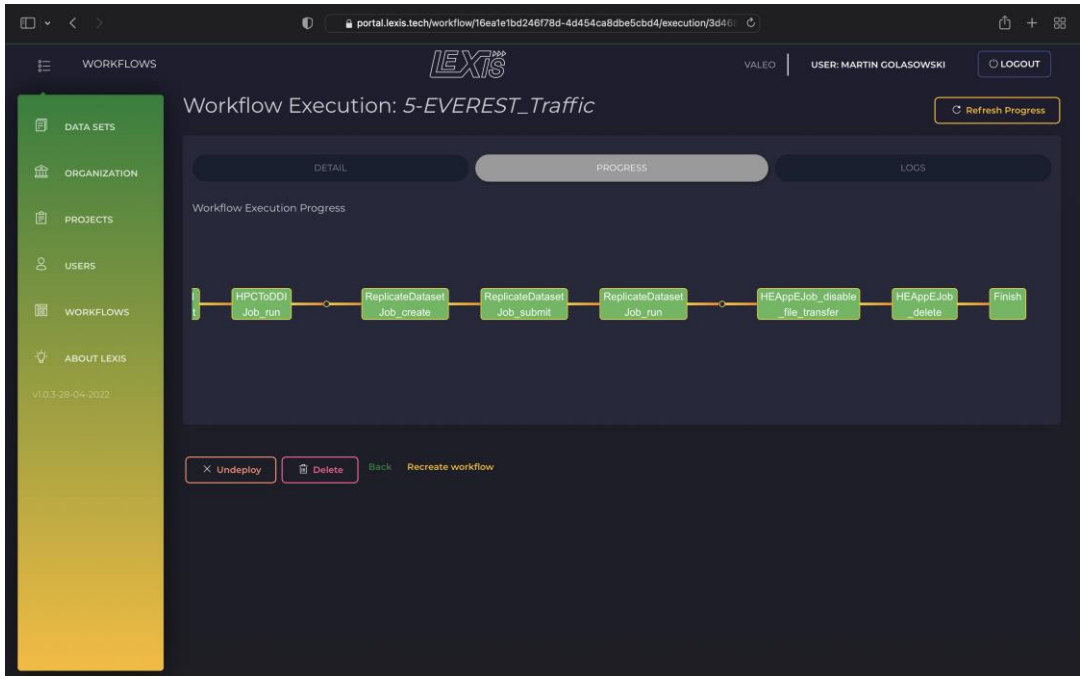
- The tests shows increased performance with larger files.
- Datasets with a big number of files suffers from low transfer rate

- SOLUTION:
  - Compress the data before moving it to iRODS
  - Uses a dedicated machine with 64 VCPUs and NVME disk to perform the compression



The plans and future outlook

- Sustain the platform for commerical use
- Extend Staging APIs for more use-cases (object storage)
- Upgrade to iRODS 4.3.0 and rework OpenID?
- Prepare a DDI node deployment Docker image
  - iRODS iCAT
  - Staging worker
  - HEAppE
  - User sync



Workflow in the LEXIS portal

THANK YOU!

CONTACTS

Mohamad Hayek (LRZ)
hayek@lrz.de

Martin Golasowski (IT4I)
martin.golasowski@vsb.cz

Our GitHub organization
https://github.com/lexis-project

Leibniz-Rechenzentrum
der Bayerischen Akademie der Wissenschaften

VSB TECHNICAL UNIVERSITY OF OSTRAVA | IT4INNOVATIONS NATIONAL SUPERCOMPUTING CENTER