

# LEXIS Platform

**HPC/Cloud Workflows and  
EUDAT-based Data Management  
for Big Data Applications**

**EUDAT Conference '22 | 14/09/2022**

**Stephan Hachinger (LRZ, LEXIS Distributed Data Infrastructure),  
J. Martinovič (IT4I, coordinator), M. Derquennes (EURAXENT)**

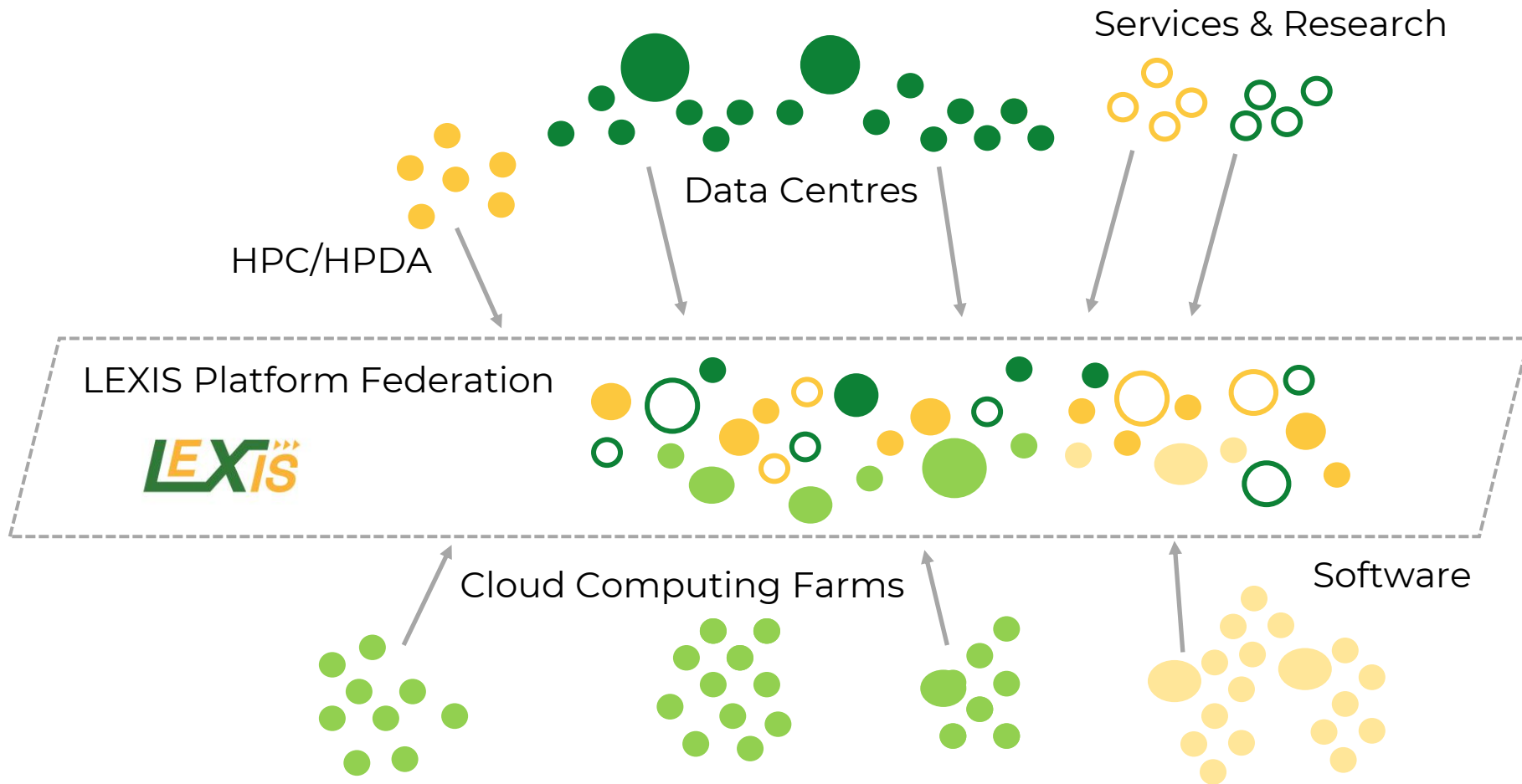
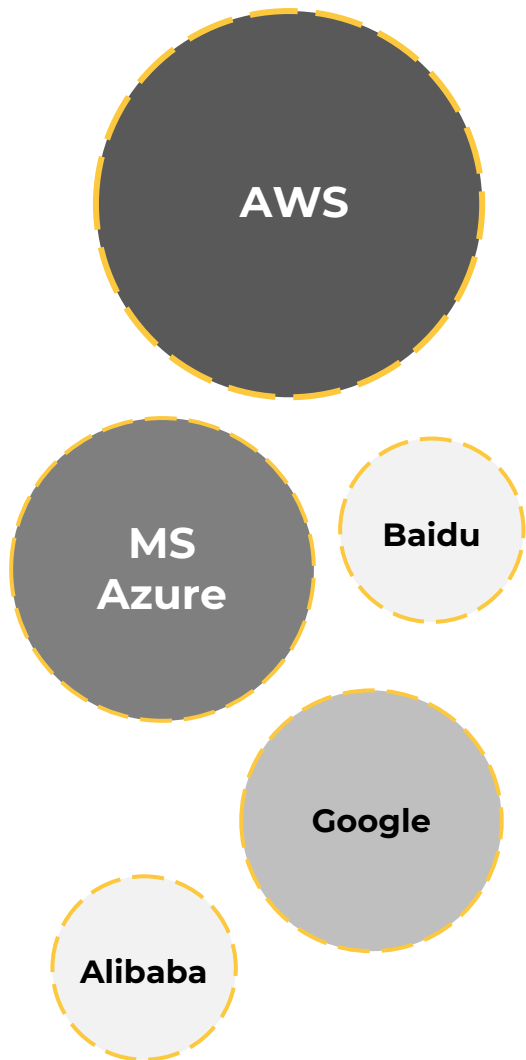


**EURAXENT**

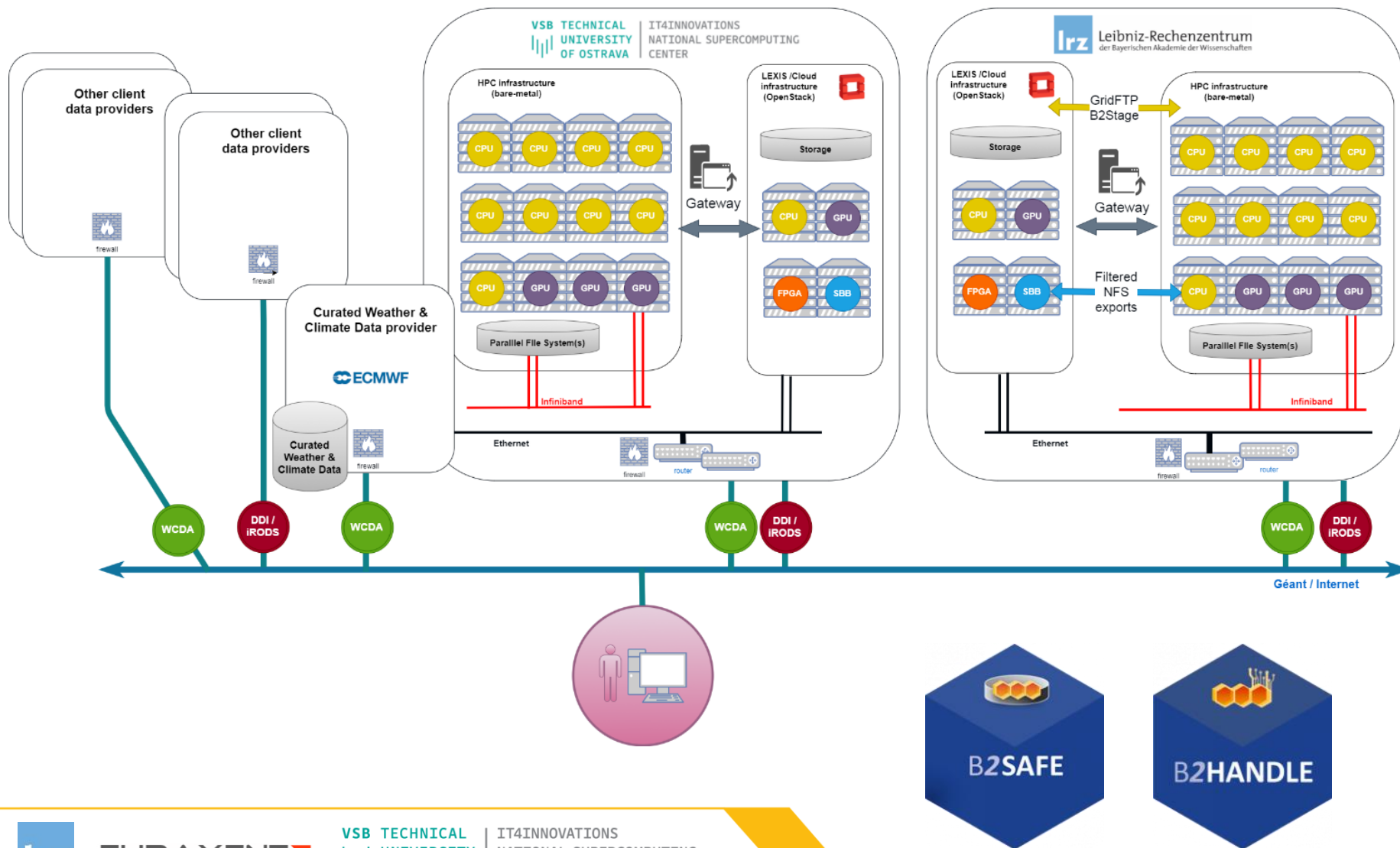
**VSB TECHNICAL  
UNIVERSITY  
OF OSTRAVA**

**IT4INNOVATIONS  
NATIONAL SUPERCOMPUTING  
CENTER**

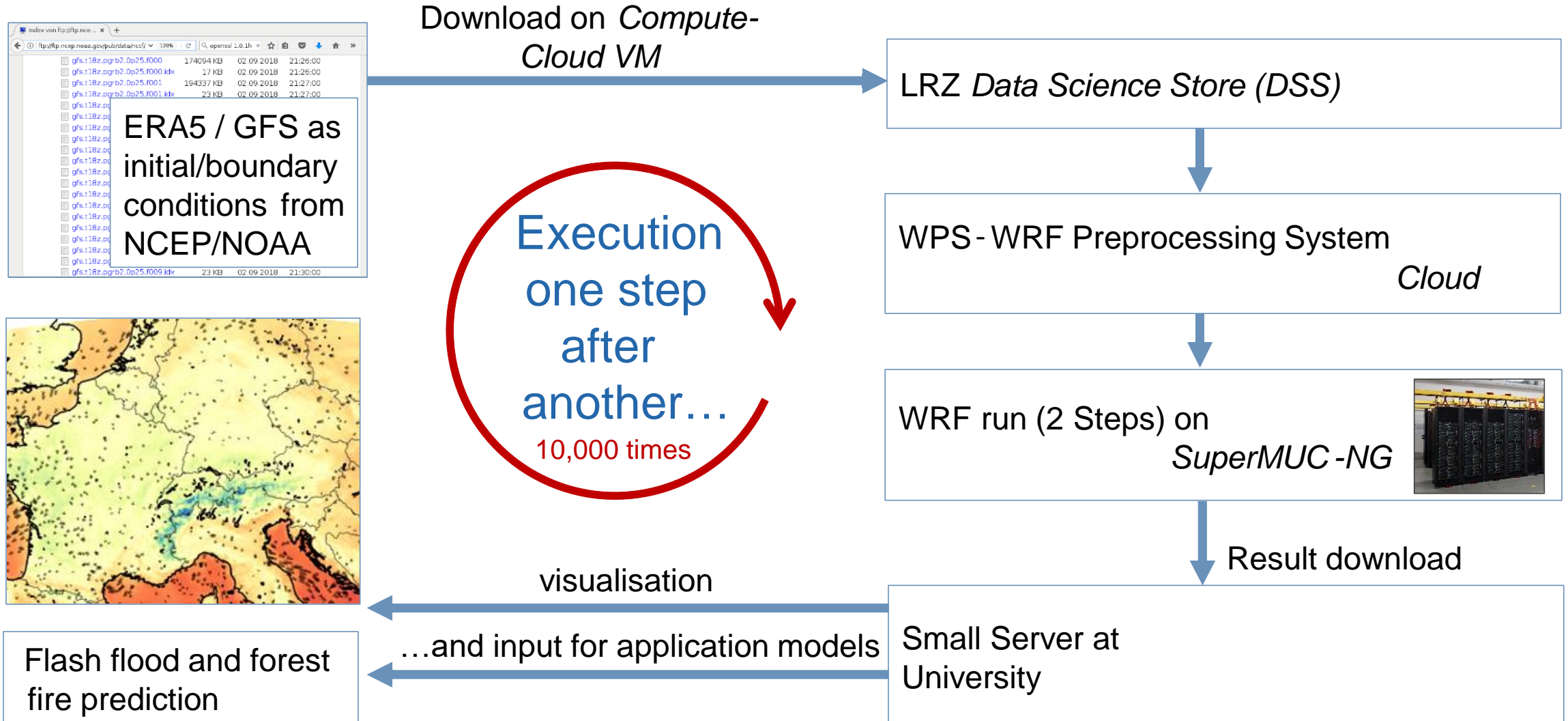


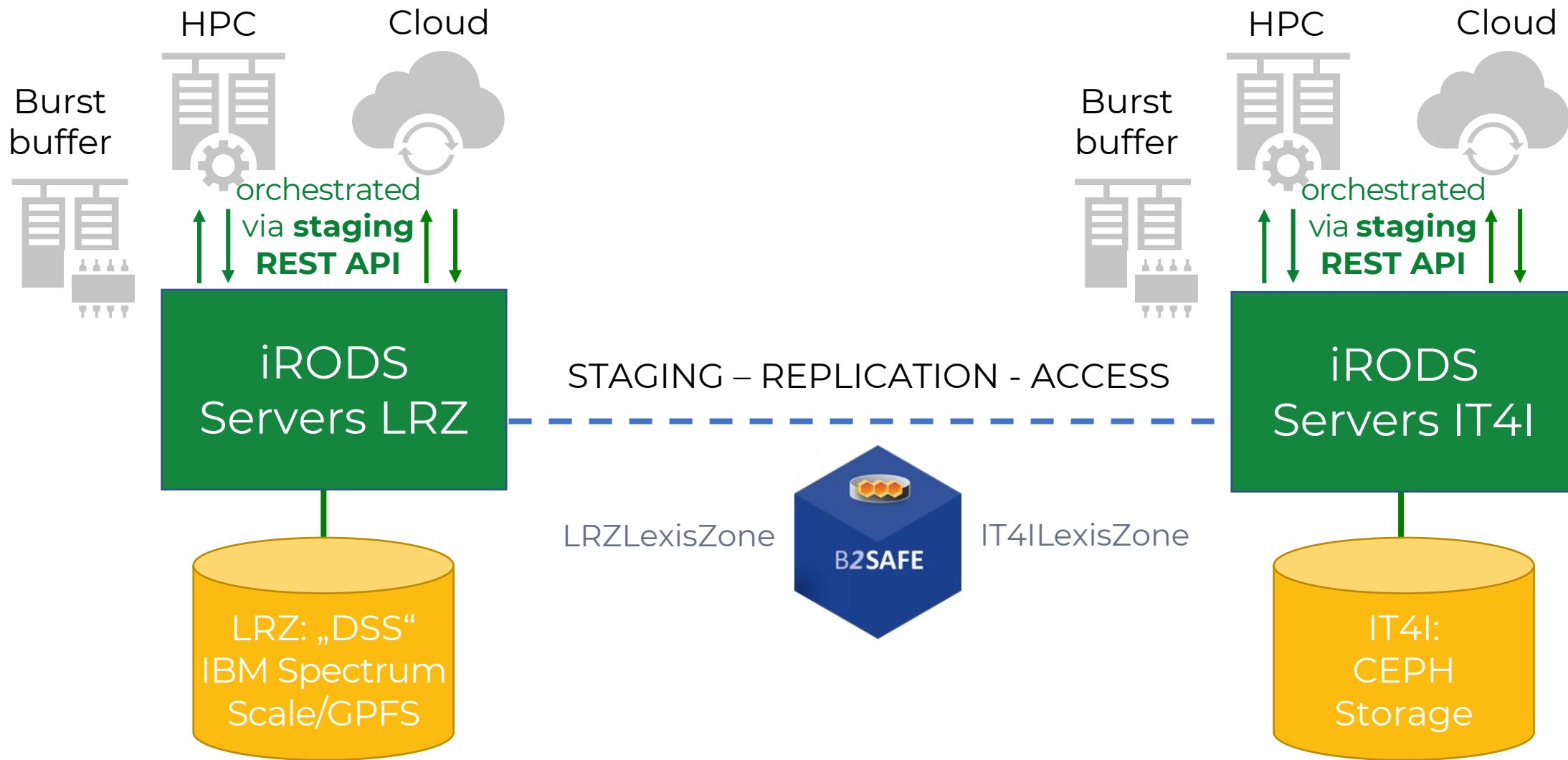


Coherence & Convergence – Solving Private/Public difficulties  
“Large Scale Execution for Industry and Society”



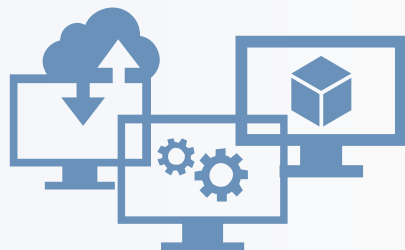
- Federation of European computing centres
- HPC & Cloud service providers, data providers
- Unified & distributed data management
- Orchestration (YORC/TOSCA)
- Federated Authentication & Authorization Infrastructure (AAI)
- User-friendly Data & Compute Portal / UI masks technical issues & organisational differences





# The DDI within the Platform

Portal and Orchestrator



Monitoring System



Data Search API



Data Staging API



Monitoring/Billing API



DDI

(Distributed Data Infrastructure with Metadata Handling / FAIR)



Local Storage Systems



AAI  
(Authentication & Authorization Infrastructure)





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


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

## Upload New Dataset

Upload type: Direct Upload

  
 UPLOAD

  
 METADATA

  
 RESULT

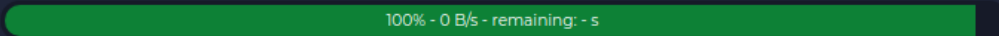
Review of dataset basic information:

<b>Title:</b>	Approval System
<b>Type of upload:</b>	Direct Upload
<b>Encryption:</b>	no
<b>Compression:</b>	no
<b>File to upload:</b>	approval_BLU.png
<b>File type:</b>	FILE
<b>Project:</b>	NewProjectForTest
<b>Access:</b>	user
<b>Zone:</b>	IT4ILexisZone

Review of dataset metadata:

<b>Resource Type:</b>	PNG image
<b>General Type:</b>	Image
<b>Creator:</b>	Foo Barringer
<b>Contributor:</b>	Max M. Painter
<b>Publisher:</b>	LEXIS consortium
<b>Owner:</b>	S. Hachinger
<b>Rights:</b>	MIT
<b>Rights URI:</b>	<a href="https://opensource.org/licenses/MIT">https://opensource.org/licenses/MIT</a>

Uploading...



100% - 0 B/s - remaining: - s

## „Solved“ in partially **LEXIS-platform-specific ways:**

- Connection of DDI to Cloud/HPC for workflows/convergence
- Metadata schema & mandatory annotation vs. „*scientists' tradition*“
- *Trans-national data sharing & PIDs / EU integration* – solved by use of B2SAFE/B2HANDLE
- RDM vs. state-of-the-art IT
  - OpenID-Connect authentication to all LEXIS components (DDI/APIs/Compute)
  - REST APIs, e.g. for automated staging requests
  - Up/download/transfer of large and many-file data sets – compression!
- RDM vs. *open/closed-science policies of stakeholders:* commercial/security requirements, e.g. encryption or access control

**“Access to Big Data, Supercomputing and Cloud Made Easy” – delivered**



## Observed general gaps:

- Problems with deletion/archival policies  
e.g. replica treatment vs. user requirement to delete data
- Missing FAIR standards for large everyday „warm data“ systems
  - Diverging metadata standards – full freedom and uncertainty
  - Deluge of PID assignment policies/implications
  - Missing cross-talk of FAIR data and supercomputing world
- Missing dataset packaging policies and tools

# THANK YOU!

## CONTACTS

**Stephan Hachinger (LRZ)**  
hachinger@lrz.de

**Jan Martinovič (IT4I)**  
jan.martinovic@vsb.cz

**Marc Derquennes (EURAXENT)**  
marc.derquennes@euraxent.com

Our GitHub organization  
[github.com/lexis-project](https://github.com/lexis-project)  
Project and Platform Web Page  
[www.lexisplatform.eu](http://www.lexisplatform.eu)



**EURAXENT** 