HELIPORT

HELmholtz **ScI**entific **P**roject W **OR**kflow PlaT**form**

UNICORE

Overview and recent developments
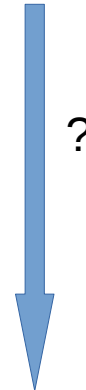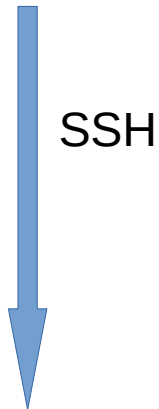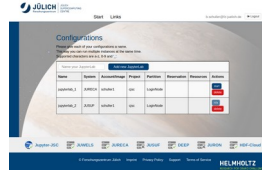
HELIPORT WORKSHOP

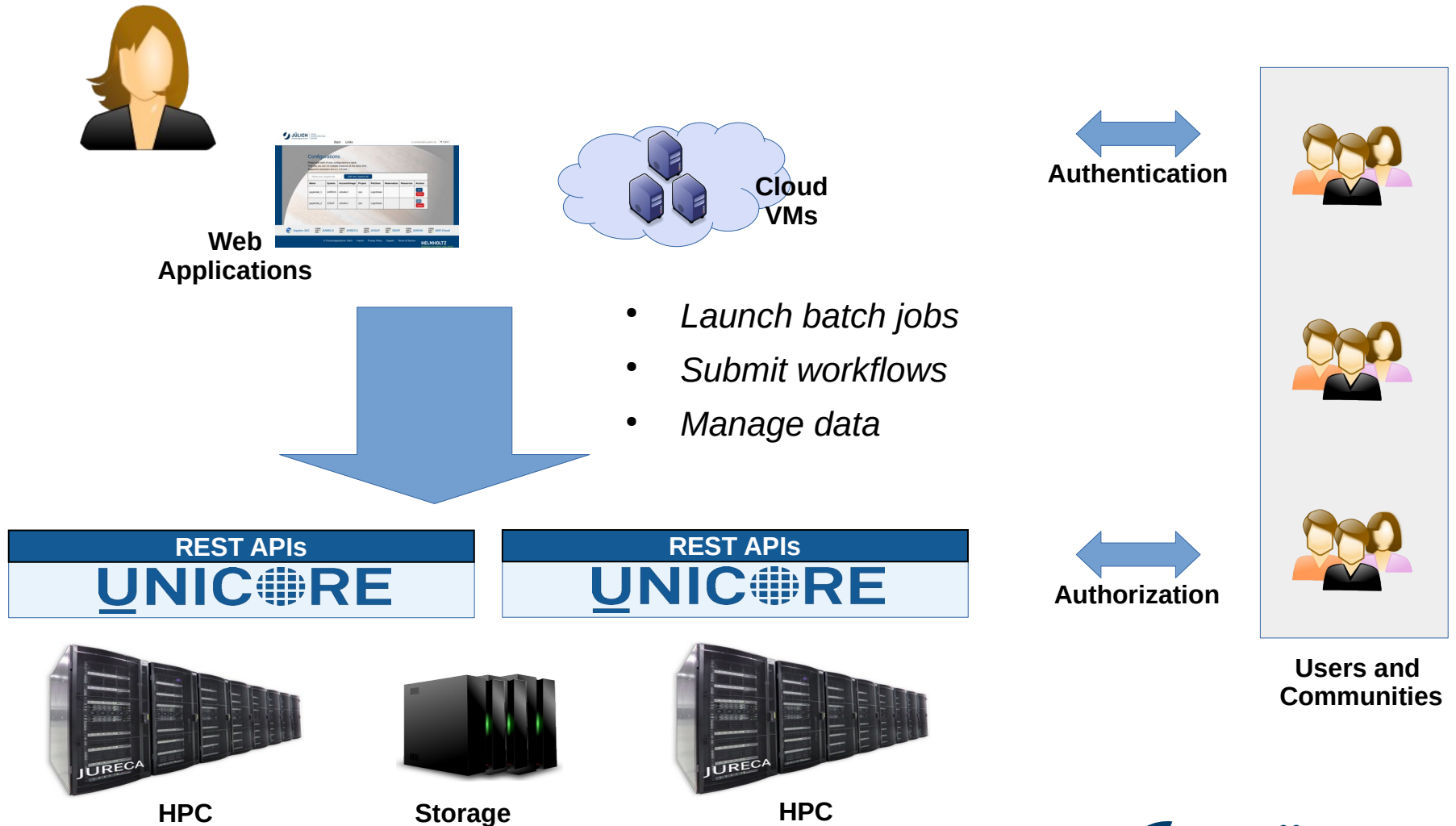HZDR, JUNE 12-14, 2023

BERND SCHULLER

FORSCHUNGSZENTRUM JÜLICH GMBH

JÜLICH
Forschungszentrum

OUTLINE

- Why UNICORE?

- Overview and features

- Security (authentication, permissions, delegation)

- Workflows
    – UNICORE
    – CWL

JÜLICH
Forschungszentrum

ACCESSING HPC SYSTEMS

SSH

?

FEDERATING HPC WITH UNICORE

Web Applications

Cloud VMs

- *Launch batch jobs*
- *Submit workflows*
- *Manage data*

Authentication

Authorization

Users and Communities

REST APIs
UNIC⊕RE

REST APIs
UNIC⊕RE

HPC

Storage

HPC

JÜLICH
Forschungszentrum

UNICORE KEY FEATURES

- Integrates HPC into federated environments

    - Flexible authentication and authorization

    - Site keeps full control over resource usage

- REST APIs and services for

    - Jobs

    - Data access and transfer

    - workflows

- Flexible, non-intrusive, portable, Open Source

JÜLICH
Forschungszentrum

EXAMPLES FOR UNICORE DEPLOYMENTS
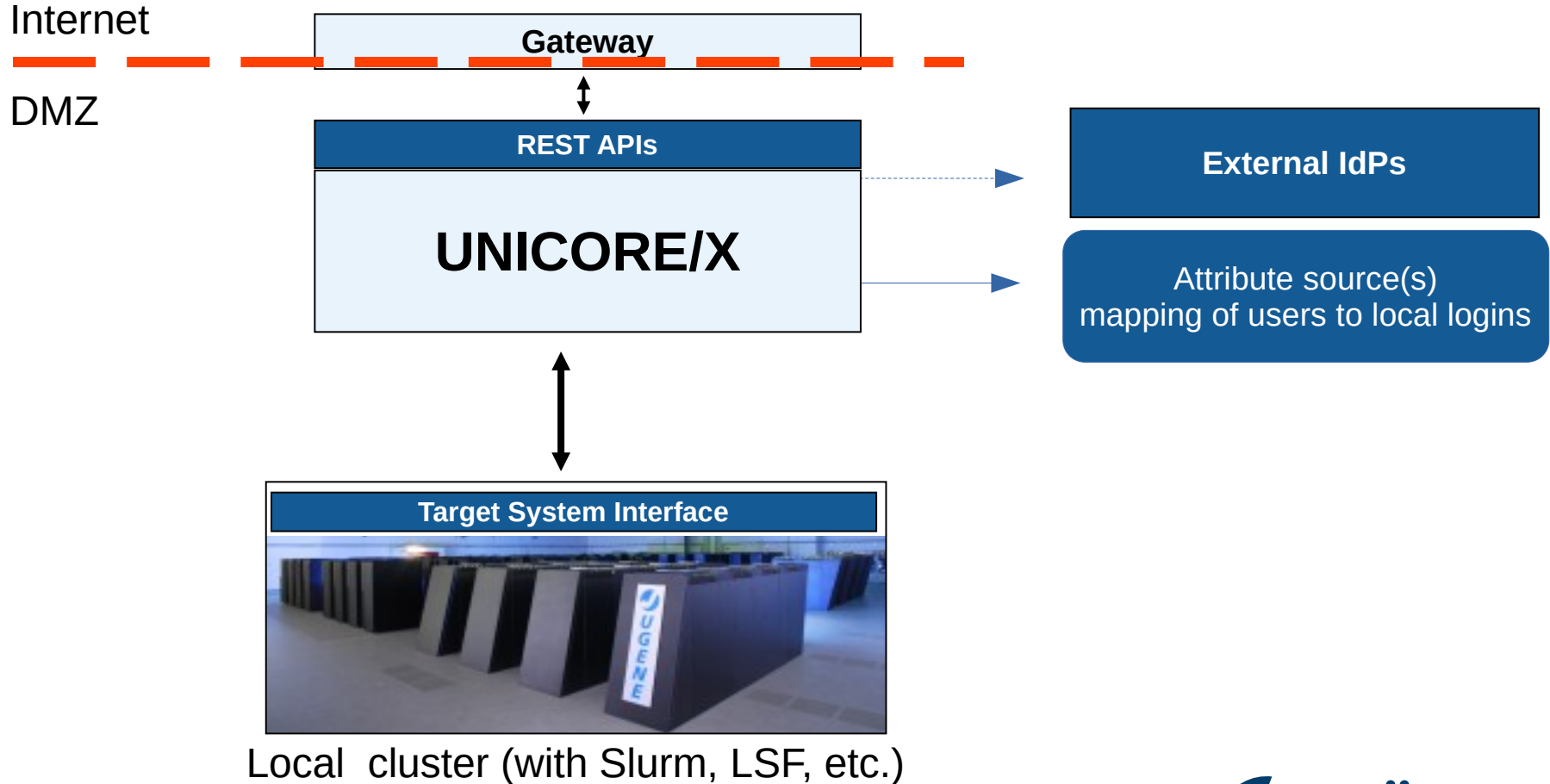
- Jülich

    – All HPC systems
       accessible via UNICORE,
       heavily used by
       Jupyter@JSC
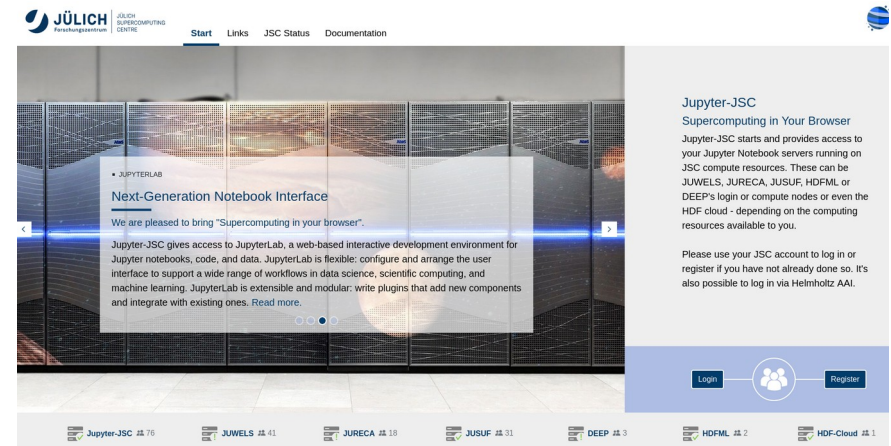
- Gauss Center for Supercomputing (JSC, HLRS, LRZ)

    – UFTP data transfer servers and client nodes

    – JupyterHub service upcoming

- EBRAINS (European research infrastructure for Neuroscience)

    – CSCS, CINECA, BSC, Jülich

UNIC⦿RE COMPONENTS
SINGLE CLUSTER

Internet

DMZ

Gateway

REST APIs

UNICORE/X

External IdPs

Attribute source(s)
mapping of users to local logins

Target System Interface

Local cluster (with Slurm, LSF, etc.)
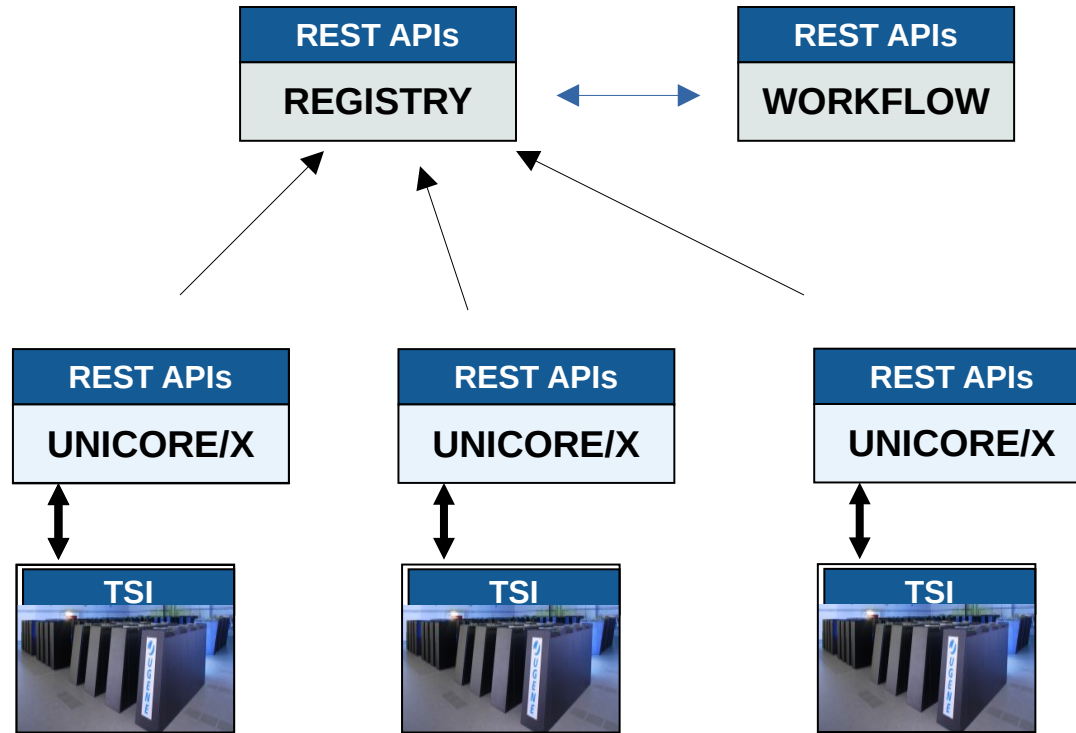
JÜLICH
Forschungszentrum

UNICORE ON A SINGLE CLUSTER

- Jobs

    - Batch or on login node

    - Pre/post processing

    - Data stage-in/out

- File system access, data transfer

- Single site workflows

JÜLICH
Forschungszentrum

UNICORE COMPONENTS
## MULTIPLE CLUSTERS, FEDERATIONS

UNICORE AS A FEDERATION

- Cross-site data transfer

- Workflow service for multi-site workflows

- Federations are defined by a Registry service

    - Strict access control for writing to a Registry

    - Contains service endpoints

    - Contains public keys of servers, important for trust between servers

JÜLICH
Forschungszentrum

SECURITY CONCEPTS

- UNICORE offers a REST API, i.e. JSON documents over HTTPs

- Authentication – identify the current user making a call

- Authorization – assign rights and attributes to the user

- Trusted server-to-server communication

  - SSL with X509 certificates

  - Delegation: securely propagate user identity

JÜLICH
Forschungszentrum

AUTHENTICATION
REST API CALL WITH AUTHENTICATION MATERIAL

User

REST API call

"Authorization: ..."

UNICORE
server
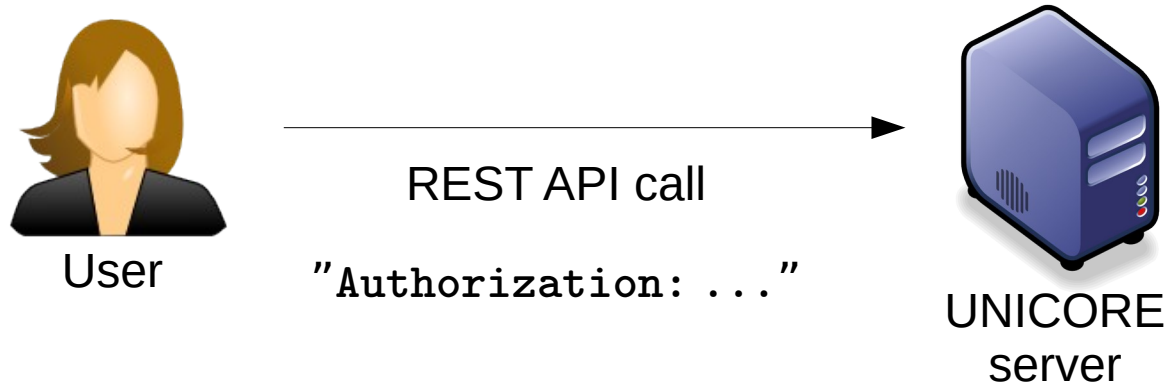
- Configurable list of authentication handlers

    – HTTP Basic (username/password)

    – Bearer token

    – Client X.509 certificate

- Result: user identity   UID=user@heliport.org

JÜLICH
Forschungszentrum

AUTHORIZATION
MAP USER IDENTITY TO LOCAL ATTRIBUTES

UID=user@heliport.org

Role: user,
UID: user1,
Groups: ...

Attribute Sources

UNICORE
server
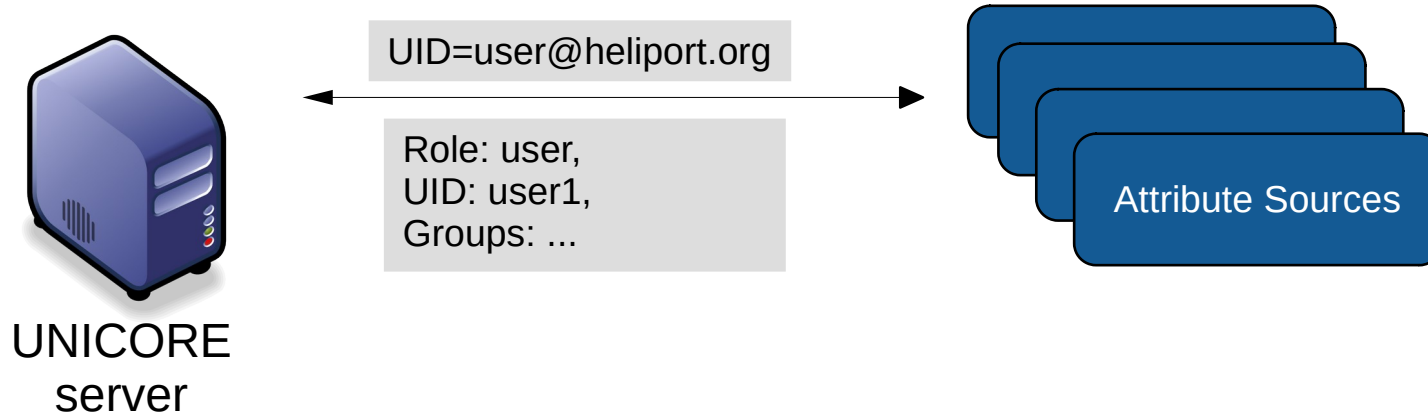
- Configurable list of attribute sources

  - XUUDB, File, LDAP, PAM, ...

- Result: user role, local UNIX login and groups

JÜLICH
Forschungszentrum

DELEGATION
SERVICES WORKING ON BEHALF OF THE USER

- First server authenticates

- Passes on user identity in a JWT token signed with server's private key

- Receiver checks JWT signature, validating with known public key

- Accepts request as if sent by original user (authentication only!!)

User   1. authenticates   UNICORE server A   2. invokes   UNICORE server B

JÜLICH
Forschungszentrum
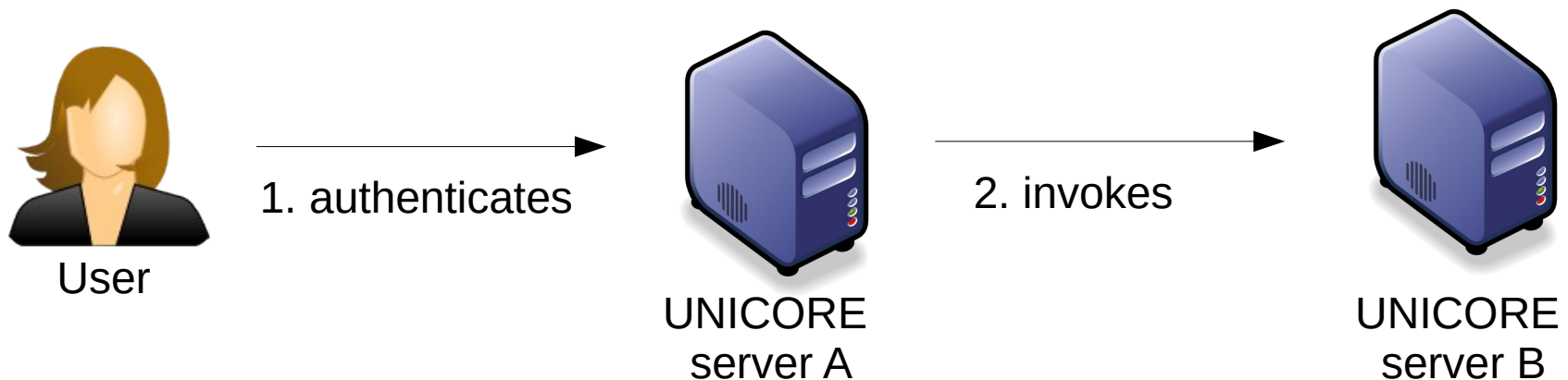
TRUSTED SERVICES
WORK ON BEHALF OF THE USER

- Claims user identity in a JWT token
  signed with trusted service's private key

- Receiver checks JWT signature, **validates** with known public key

- Accepts request as if sent by user

Invokes on behalf of user

```
"sub": ...,
"iss": ...
```

Trusted
service

User

UNICORE
server B

JÜLICH
Forschungszentrum

CLIENTS

- UNICORE Commandline Client (UCC)

    - End-user application, written in Java

    - https://unicore-docs.readthedocs.io

    - Downloadable from Sourceforge

- PyUNICORE

    - Python binding for the REST API

    - https://pyunicore.readthedocs.io

    - Install via

        ```
        pip install pyunicore
        ```

JÜLICH
Forschungszentrum
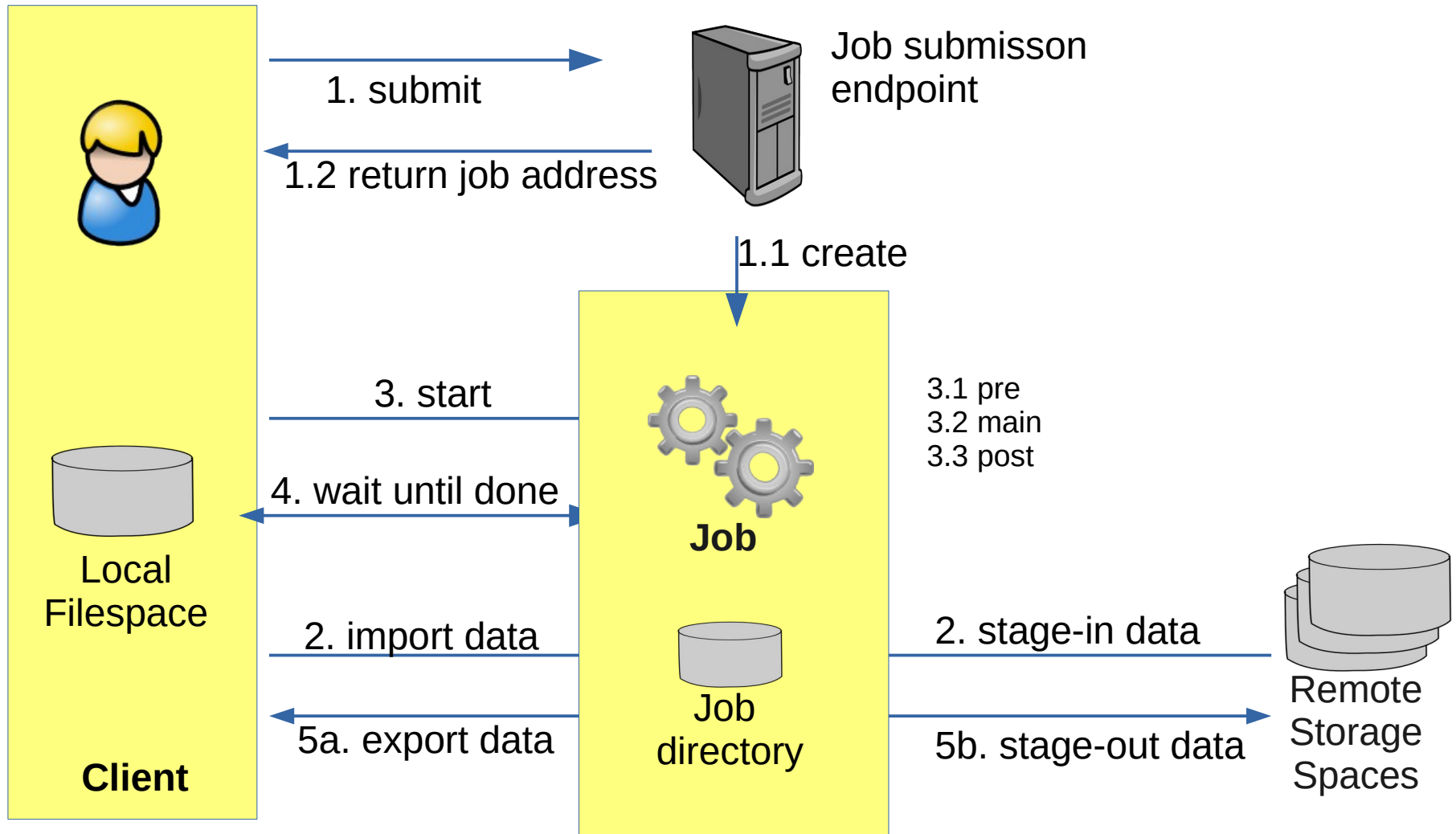
JOB EXECUTION

UNICORE WORKFLOWS

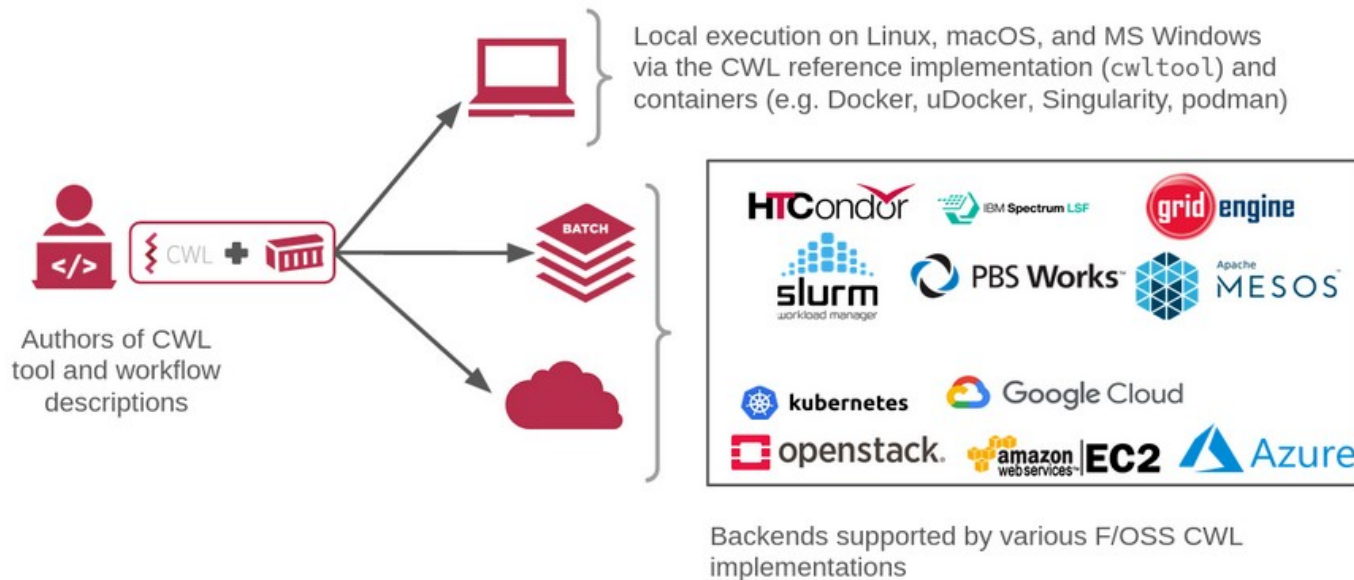- UNICORE Jobs as basic executable units

- Flow control via (conditional) transitions

- For-each loop for parallel execution over input sets

- Variables, modifiable via Groovy scripts and context info

- While/repeat loops

- Hold points can be defined, continue via REST API

- Powerful file management

- JSON syntax

    - … conversion into internal data model

JÜLICH
Forschungszentrum

COMMON WORKFLOW LANGUAGE

- "… for dataflow style batch analysis consisting of command line programs"

- Benefits of standardisation, large community, tooling, etc

- https://commonwl.org

Local execution on Linux, macOS, and MS Windows via the CWL reference implementation (`cwltool`) and containers (e.g. Docker, uDocker, Singularity, podman)

Authors of CWL tool and workflow descriptions

Backends supported by various F/OSS CWL implementations

A ROUGH COMPARISON

	UNICORE	CWL
Simulations / Batch jobs	☑	☑
Flow control	☑	☑
Container support	☑	☑
Loops	☑	☒
Data-parallel execution	☑	☑
Input/output file management	☑	☑
Variables / Scripting	☑	☑
Hold / continue	☑	☒
"Local" execution	☒	☑

JÜLICH
Forschungszentrum

CWL TO UNICORE

HELIPORT HELmholtz ScIentific
Project W ORkflow PlaTform

- Both workflow systems are very similar in scope

- PyUNICORE includes simple converter code
  - Single job / commandline tool only

- Planned / promised:
  - Workflow conversion
  - Executor tool like `cwltool`

JÜLICH
Forschungszentrum

SUMMARY: UNICORE

- Federating HPC – enabling access to HPC
  from portals and web applications like HELIPORT or JupyterHub

- Comprehensive security features: user authentication, attribute mapping, delegation

- RESTful APIs

- Scientific workflows and other integration use cases

- Data access, transfer, sharing

  visit https://unicore-docs.readthedocs.io for more

JÜLICH
Forschungszentrum

THANK YOU!

JÜLICH
Forschungszentrum