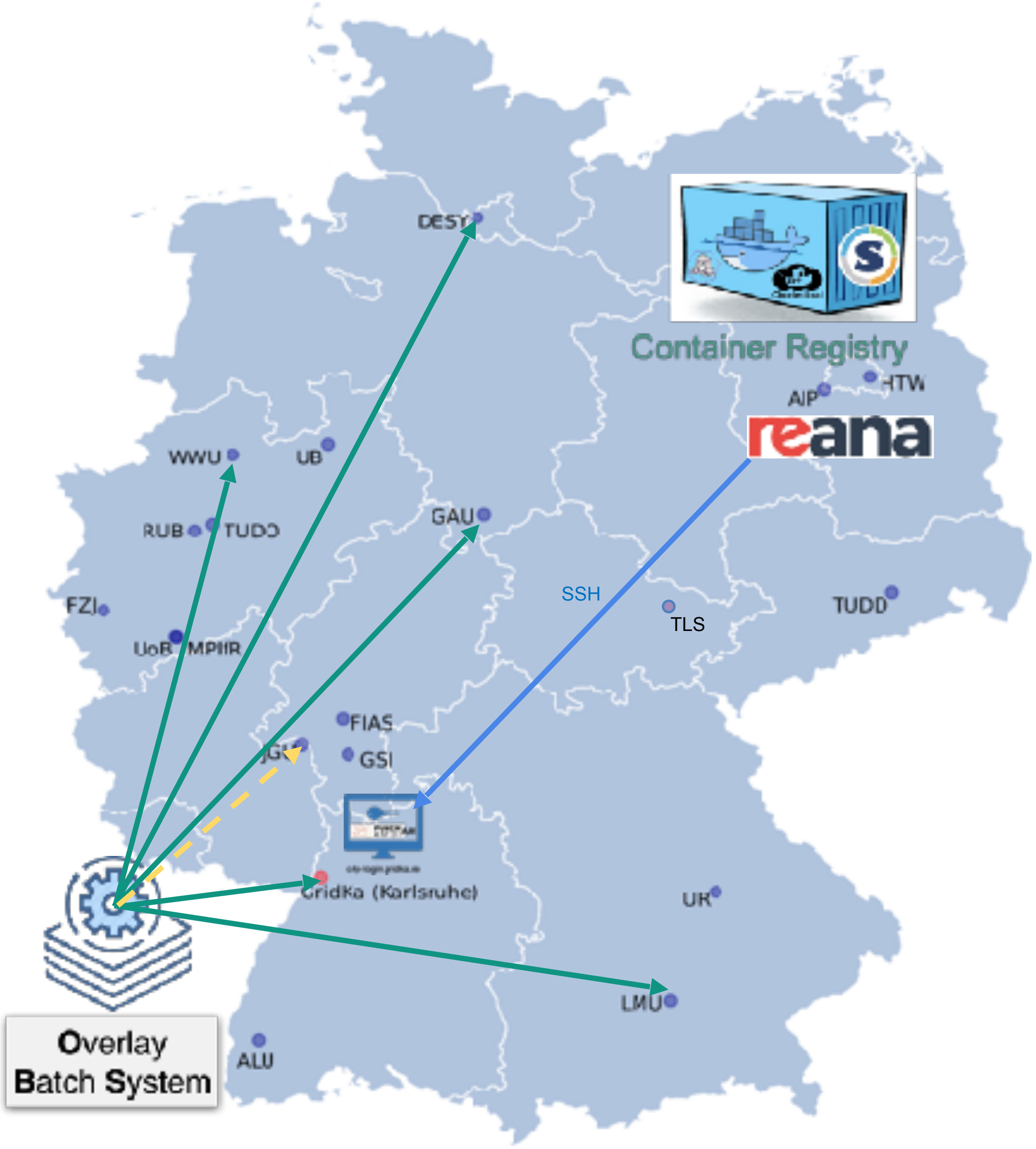
TA2 Status, Highlights and Plans
PUNCH general meeting, 7 July 2024

Particles, Universe, NuClei and Hadrons for the NFDI

Matthias Hoeft on behalf of the TA 2 and others

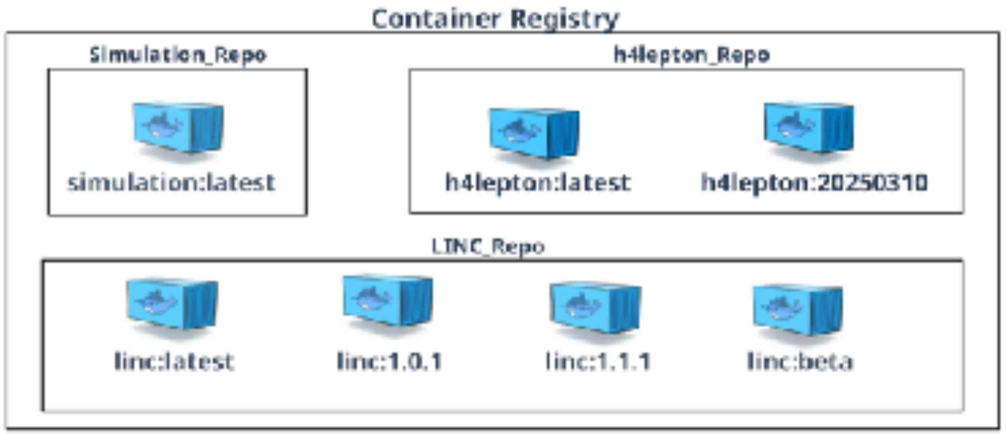
Evolution of Compute4PUNCH

- C4P **fully integrated into REANA** production instance hosted at AIP, discovered already potential improvements
- Dynamic integration of **various heterogeneous compute resource** of providers established
  - ○ First successful HTCondor job flock to Mainz HPC node (no outgoing connectivity)
  - ○ Final integration of CM4 HPC for MeerKLASS use-case is imminent
- Monitoring data of available resources is now continuously transmitted to Influx DB at AIP (REANA Dashboard)
- About to finalise the **Container Registry 2.0** to enhance container provisioning
- Eventually found a viable solution to host the **JupyterHub** developed by the University of Bonn on KIT resources, fulfilling the strict SIRTFI requirements of Helmholtz AAI



CI/CD for Container Registry

- **Docker containers** encapsulate the suitable software environment necessary for running the workflow
- PUNCH4NFDI utilizes this Docker container technology
- **Containers** are maintained in Docker Container Registry at AIP (gitlab-p4n.aip.de) using a CI/CD pipeline

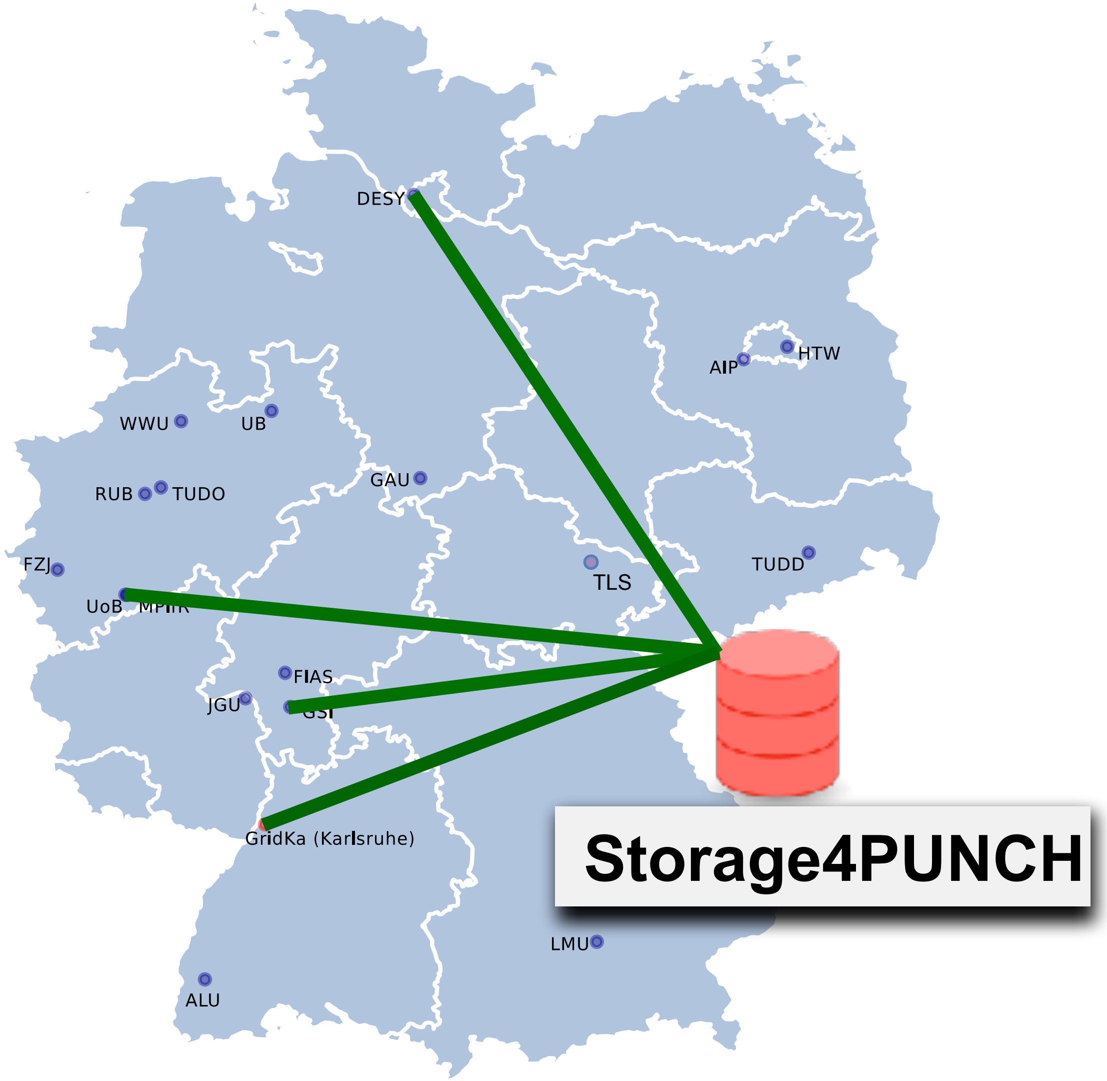


<u>CI/CD Framework v2.0 features:</u>

- A **master CI/CD pipeline** for the continuous integration and deployment of Docker files
  - Template available to include your project in the master CI/CD pipeline
- Each Dockerfile for individual PUNCH project - in a separate repository within the GitLab project group
  - "compute4punch/container-stacks-v2"
- The master CI/CD pipeline only triggers in the repository where a change occurs
- This schema allows for full control over the versioning of Docker containers
- **Developer specified versioning**: use "git tag VERSION_NAME" when pushing your development
  - VERSION_NAME can be any permitted tag, i.e., "v1.0", "v1.2.1", "Release1" or "17072025"
- **Automated deployment in CVMFS**: versioned Docker container is automatically deployed via CVMFS

Evolution of Storage4PUNCH

- Based on two storage technologies
  - dCache (Instances at DESY and KIT)
  - XrootD (Test instances at U Bonn & GSI)
- Token based access using PUNCH AAI
- Supported protocols: WebDAV & XrootD
- Main activity is the testing of AAI features
  - Group based authentication
  - Claim based authentication (using IndigoIAM)
- Pending items
  - Caching solutions
    - Popular in WLCG is Xcache (XrootD based)
    - dCache satellite instance(s)
  - Shared “home” directories
    - Shared/synced directory between Login-Node(s) and C4P instances
  - Prototypes feasible in PUNCH1.0, but perhaps only one of the two.
    Likely same people involved in both



AAI: Managing the Access Control

Local control at (individual) storage

- User typically authenticated centrally via AAI
- On local storage user is mapped to a local account (or group)
- Local (POSIX) file permission grant/deny access

Groups managed at AAI or VO (virtual organisation)

- Users are members of a VO, potentially in dedicated groups for additional privileges
  - VO group memberships are managed centrally
- Access rights depend on mapping of AAI groups to local accounts
  - Possible ambiguities if token contains several groups
  - Recent UNITY feature of reducing capabilities/groups in a token is essential
- Could become challenging, if many **mappings need to maintained at a number of different storage instances**
- Some testing using the LOFAR group is in progress

AAI: Managing the Access Control

Storage claims/capabilities

- Centrally managed
- Central component gives access right to files or directories
  - Tokens like: storage-read:/my-exp/data1
- Storage element trusts the token and grants access accordingly
  - **No need to configure storage systems**
    (all happens in a central service)
- Functionality implemented in Indigo IAM (used by LHC experiments, Belle II, SKA, ILDG)
  - Did some exploration of Indigo IAM “as guests” in the ILDG Indigo instance (hosted at INFN, Italy)
  - Implementation for UNITY is on the roadmap, driven by PUNCH4NFDI (TA6)
    - Requires external policy engine
    - Serious testing is more on PUNCH2.0 timeline, eventually tests in PUNCH1.0 with dummy engine

Metadata and File Catalog for ILDG & Astro

Recent developments

- Moved to fine-grained token-based authentication and access control (Indigo IAM)
- Moved from eXist DB to Postgres for better scaling and sustainability
- Added “Quick Search” (JSON queries) and further functionality for non-ILDG applications
- Add-on service for Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)
- 5 MDC+FC instances deployed in Germany, Japan, and UK
- Prototype for general GUI for markup and search (parametric in XSD)
- CI/CD pipeline for publicly available container with user environment and client tools
- Knowledge Sharing (Hands-on ILDG workshop this week with participants from 9 countries)

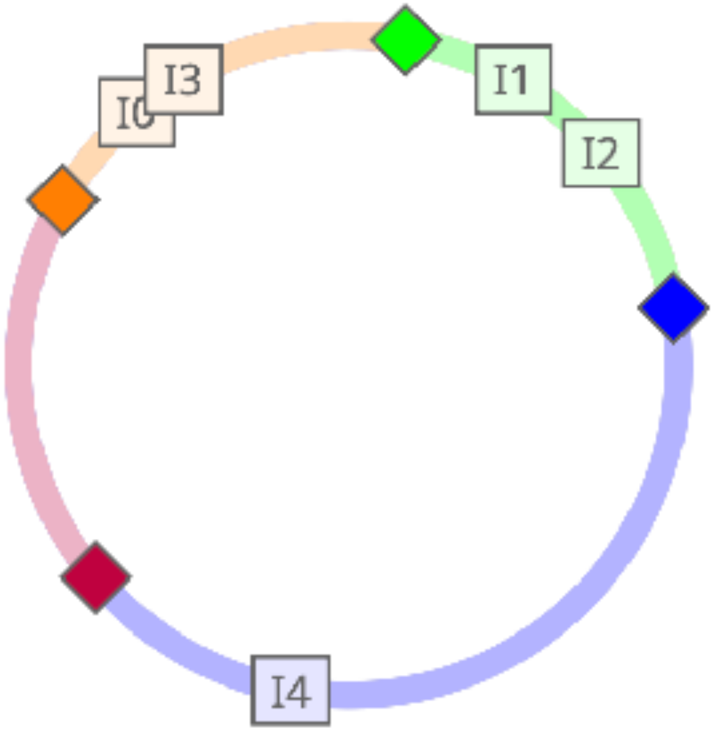
Further Plans

- Still some open issues to work on
- Broader support for other communities through PUNCH 2.0

Distributed Hash-Table for dCache

Main idea

- Metadata is mirrored from central metadata catalogue across all server nodes
  - consistent hashing to determine which node stores requested metadata
  - each of n pool server nodes therefore holds 1/n of metadata
- Metadata lookups are resolved between pool server nodes
  - minimises interaction with central metadata catalogue
  - rebalancing in case of new or leaving servers
- Central manager remains the ground truth
  - updates are sent to central manager which asynchronously forwards them to pool nodes
  - any inconsistency is resolved by querying the manager



PUNCH related activities

- Common project between DESY (Kilian) and U Mainz (Andre Brinkmann)
- First demonstrator feasible still during PUNCH1.0

GoeGrid -
Göttingen Computing Resources for PUNCH4NFDI

1- Göttingen Contribution to Compute4PUNCH

- Göttingen provides **160 CPU cores from the GoeGrid** Cluster to Compute4PUNCH.
- Resources integrated and **operational within the Compute4PUNCH infrastructure since Q2 2024.**
- Contribution cited in the PUNCH4NFDI paper to be published in the proceedings of the **CHEP 2024 conference (EPJ Web of Conferences).**

2- EXPLORE: CERN Open Data Analysis Platform

- **Purpose:** Provides CERN Open Data analysis access to users without CERN or university affiliation using GoeGrid resources.
- **Hosted on** GoeGrid Cluster, **which is operated by** II. Physikalisches Institut, Georg-August-Universität Göttingen.
- **Resource Management:** Dynamic allocation via **COBalD/ TARDIS** with a dedicated **Entry Point** & **HTCondor** Overlay Batch System.
- **System Setup:** 1 Central Manager, 1 Submitter, & 200 CPUs Worker Nodes; monitored with Prometheus & Grafana.
- **User Environment:** Pre-configured, scalable analysis environments via **CVMFS** & **Apptainer**.
- **Scientific Outreach: Abstract accepted for CoRDI 2025**: “EXPLORE: A Scalable Infrastructure for LHC Open Data Analysis and FAIR Data Provisioning” (Track: RDM Infrastructures).

EXPLORE User Access & Compliance
Road-map for Full PUNCH4NFDI Integration

Current Operational State & Optimization of EXPLORE Access

EXPLORE user access currently runs outside the PUNCH AAI infrastructure as an independent service.

- **Optimization**

—> Alpha & beta testing, including with high school students. Improved performance & usability based on feedback.

- **Online Custom Registration System with Minimum Requirements**
  - User-name, Valid email address
  - SSH Key Pair for secure login
  - under/over-18 flag and parental consent for minors
  - Consent to the Terms of Use
  - Register

Compliance & Policy Actions Required for Full PUNCH4NFDI Integration

❖ **Commercial IdPs Removal:** To comply with German law, **all active links to commercial IdPs must be removed** from the PUNCH AAI login page and related services. Their presence **implies tax obligations.** This is a critical compliance issue affecting all PUNCH services.

❖ **Lightweight Registration:** Email-based registration for unaffiliated users technically possible **but needs Executive Board approval**, & **technical changes**.

❖ **Youth Protection: PUNCH AAI** lacks age verification; **an under/over-18 flag and parental consent must be implemented**. Data of under-age users should be automatically **deleted after 3 months, for data privacy** (GDPR)

⍰ ***Follow-up:***

No current process for policy & Governance AAI issues; propose raising this at the Executive Board or creating a dedicated policy track.

MeerKLASS Use-case

- 4 PB of raw data recorded at the MeerKAT array (~2000 hours observation time)

- 10,000 deg² in UHF band, 300 deg² in L band

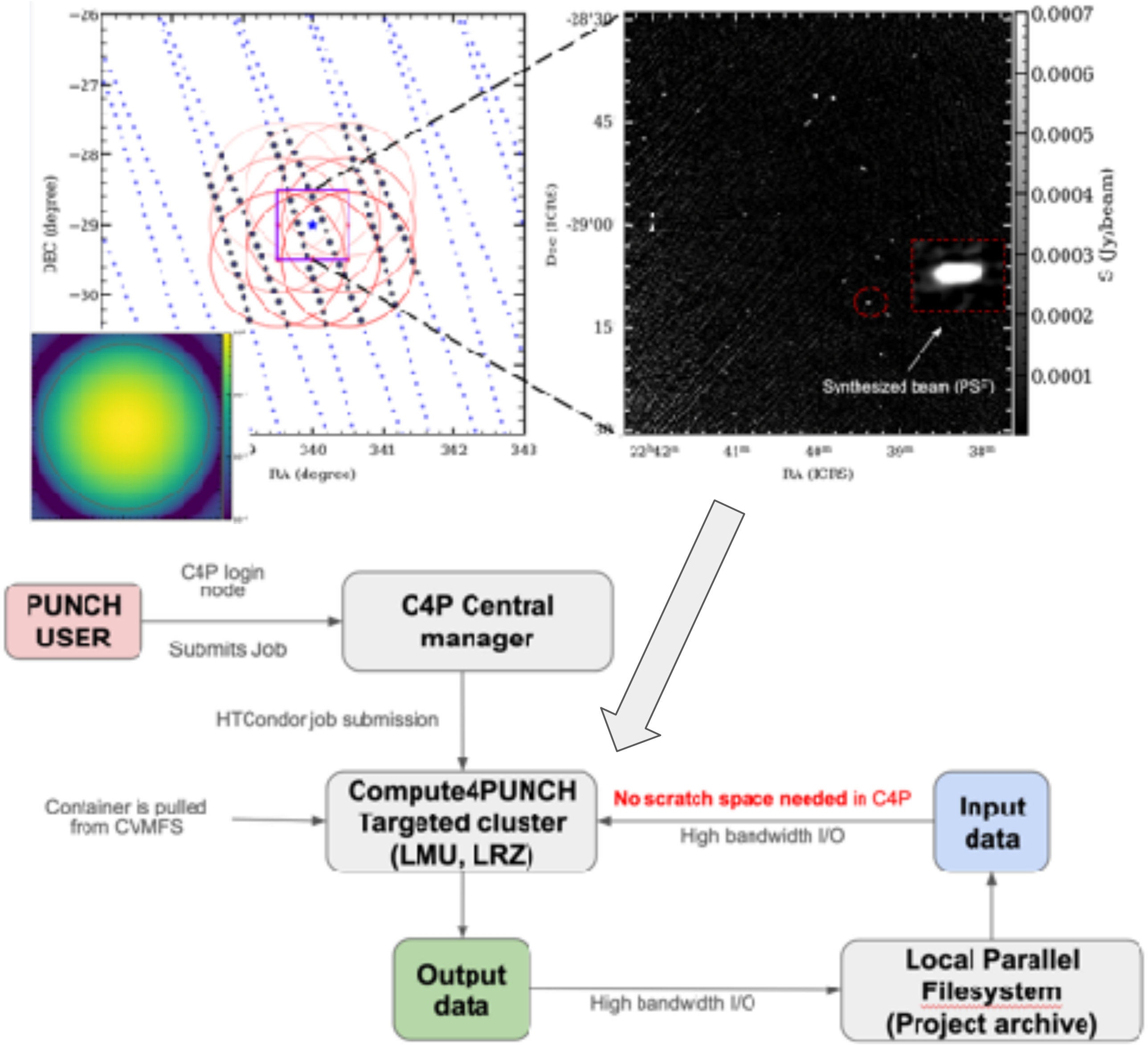
Science goals:

  - Neutral hydrogen intensity mapping for baryon acoustic oscillations
  - Continuum imaging of active galactic nuclei, radio galaxies

- New software enables On The Fly (OTF) interferometric imaging

- Evolving requirements to image 1 deg² sky patch
  - 128 cores, 500 GB RAM, Runtime ~7-8 hours
  - Up to 4 TB intermediate data produced!

- Requires integration of dedicated LRZ HPC cluster hosting data on local parallel FS into C4P
  → Prototype integration of CM4 functional



PUNCH USER

C4P login node

Submits Job

C4P Central manager

HTCondor job submission

Container is pulled from CVMFS

Compute4PUNCH Targeted cluster (LMU, LRZ)

No scratch space needed in C4P

High bandwidth I/O

Input data

Output data

High bandwidth I/O

Local Parallel Filesystem (Project archive)

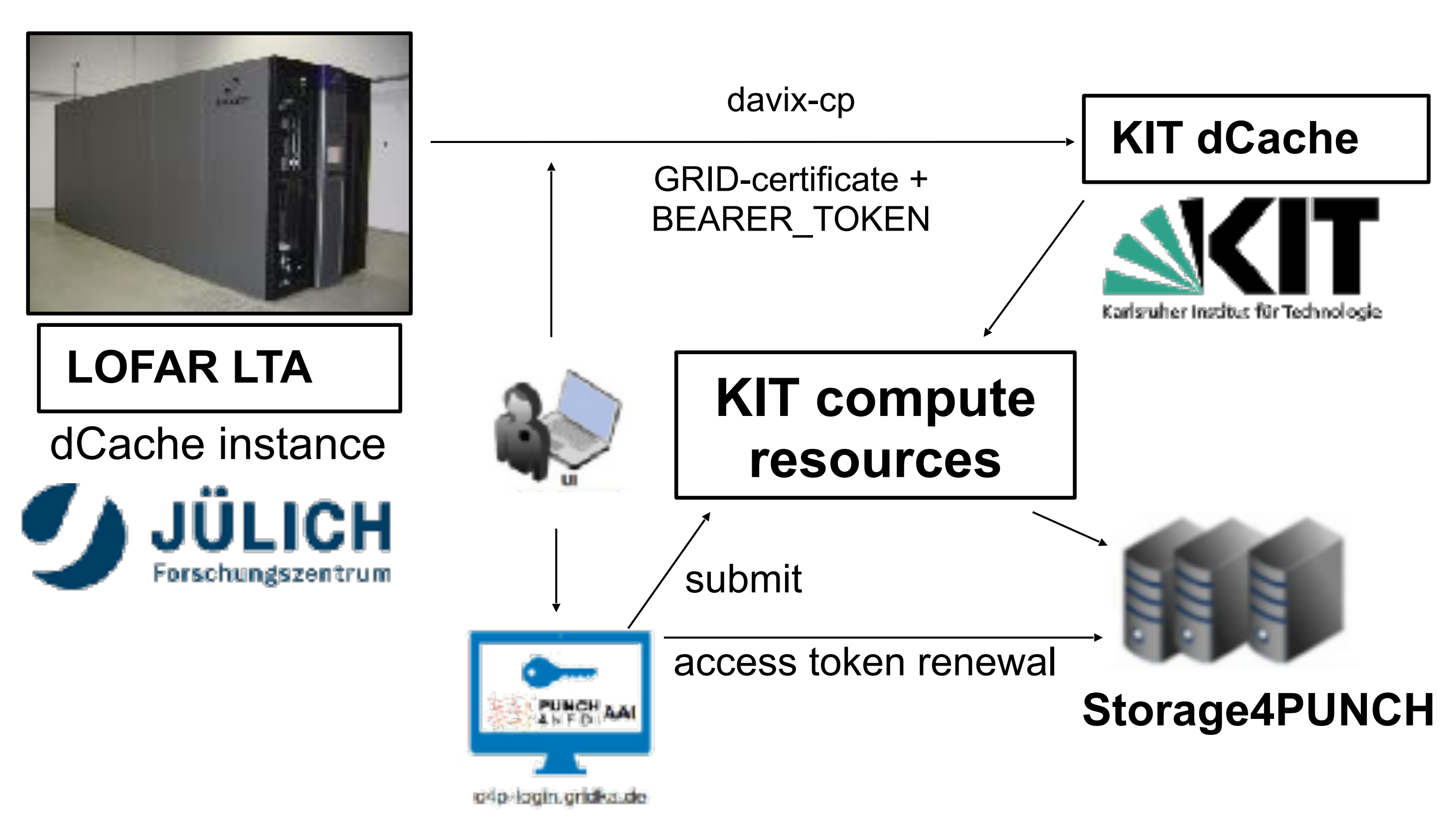
PUNCH 4NFDI

Use case:
Processing the LOFAR Two-Metre Sky Survey (LoTSS)

- Sky survey covers northern hemisphere @ 144 MHz
- **LOFAR** radio interferometric data is large (~ 1 TB/hr)
- Complex interplay of data calibration and sky brightness reconstruction („imaging“)
  → Mixture of high-throughput and high-performance computing

Special requirements:
- Data are stored on tapes (long-term archive)
- Very large data volumes → many copies unfeasible
- Current software design requires powerful (single) nodes and sufficient scratch space (~5 … 20 TB)

Realisation with S4P and C4P @ KIT:
- User only needs to provide URLs of staged archival data
- 3rd party transfer between archive and storage
- Limit job submission to local C4P resources
- Output is written back to local S4P resources



LOFAR LTA
dCache instance
JÜLICH Forschungszentrum

davix-cp
GRID-certificate + BEARER_TOKEN

KIT dCache
KIT Karlsruher Institut für Technologie

UI

KIT compute resources

submit
access token renewal

PUNCH 4NFDI AAI
c4p-login.gridka.de

Storage4PUNCH

TLS
Thüringer Landessternwarte
Tautenburg

Radio Sky Image from LOFAR

LOFAR ERIC

Base4NFDI collaborations

IAM4NFDI

- UNITY development - PUNCH needs coordinated through TA6
- Testing and early adoption in TA2

MultiCloud4NFDI

- Base4NFDI Initialization Phase Proposal, not accepted yet (2x) despite large support base
- Coordinators: Alexander Sczyrba (Bielefeld/BioInformatik), Kilian Schwarz (DESY)
    - Further PUNCH applicants: Harry Enke (AIP), Jörn Künsemöller (Bielefeld/Physik)
- Long-term Goal: Provide infrastructure to easily combine heterogeneous storage and processing resources of different providers and consortia into NFDI-wide execution environment
- Wrote letter together with NFDI Overall Architecture (“Gesamtarchitektur”) to KV for guidance
    - Following recommendation from KV: organise workshop together with OA with resource providers and consortia as target group in autumn. Depending on outcome, will re-submit together with OA in 2026 or wait for better opportunity
- In talks with Base4NFDI liaison officer to get clarity on best approach to continue

Summary on TA2 recent developments

- Compute4PUNCH: new resources available, increased maintainability (login node / container registry 2.0), fully integrated into REANA service
- Storage4PUNCH first exploration of storage scopes being a “guest” of ILDG’s Indigo IAM and “S4P-ILDG”, RUCIO4PUNCH Testbed deployed for Storage4PUNCH storage endpoints
- AAI remains a significant topic of discussion, recent and ongoing developments address some important needs and should arrive still in PUNCH1.0
- Dedicated EXPLORE Service available in Göttingen enabling public access to perform analysis of CERN Open Data
- Recent improvements of the Metadata Catalog for ILDG & Astro
- Supported multiple use-cases utilizing C4P and S4P
    - Processing the LOFAR Two-Metre Sky Survey (LoTSS)
    - MeerKLASS On The Fly (OTF) interferometric imaging
    - First event-level combined ATLAS +CMS CERN Open Data analysis
    - Post-processing of Cosmological Simulations
- Base4NFDI engagement of TA2 in IAM4NFDI and Multicloud4NFDI

Thank you for your attention!