TA2 Report
PUNCH general meeting, 12 April 2022

Particles, Universe, NuClei and Hadrons for the NFDI

PUNCH
4NFDI

TA2 – General Information

- TA coordinators: Matthias Hoeft & Christoph Wissing
- Three Workpackages
  - WP1: Access to Data & Matadata (Dominik Schwarz, Christoph Wissing)
  - WP2: Comput4PUNCH (Manuel Giffels, Matthias Hoeft)
  - WP3: Automation & optimisation of big data workflows (Andreas Henkel, Kilian Schwarz)
    - Has not started yet – as planned in the proposal – but will soonish

- Usual Meeting slot: Monday at 12:00 (noon)
  - Alternating: TA2 general, individual WPs

- Present main activities
  - Prototyping of distributed storage "Storage4PUNCH"
  - Demonstrator Compute4PUNCH

- More detailed information: https://intra.punch4nfdi.de/?md=/docs/TA2/overview.md

Storage Prototyping in PUNCH4NFDI

Common PUNCH-AAI: Based on Helmholtz-AAI

Prototype installation at DESY

- DCache technology
- Test instance shared with other international projects, e.g.
  - ESCAPE
  - PaNOSC (Photon&Neutron Open Science Cloud)

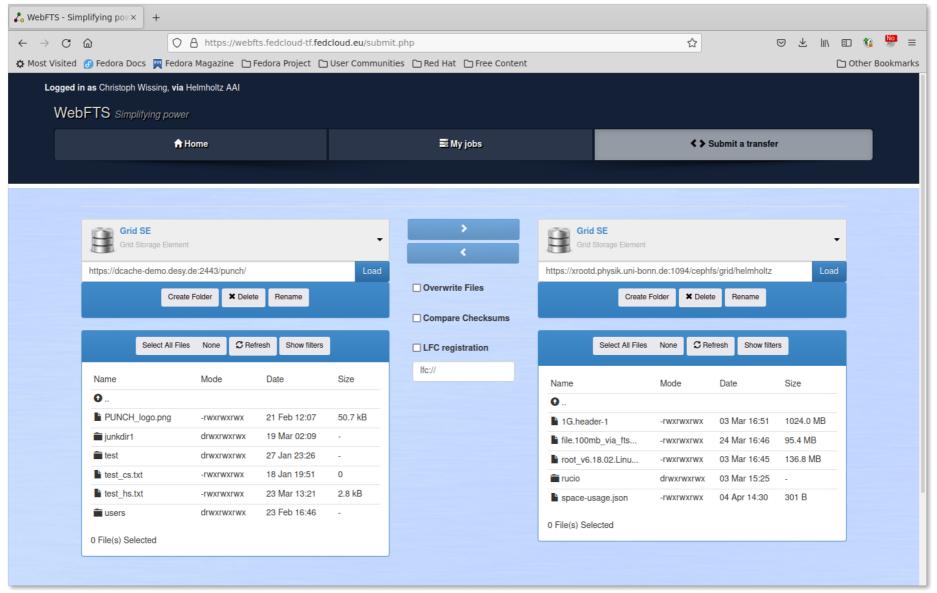## Prototype installation at U Bonn

- XRootD with CEPH backend
- Production instance
  - WLCG Tier-3 for ATLAS
  - Belle-2 storage

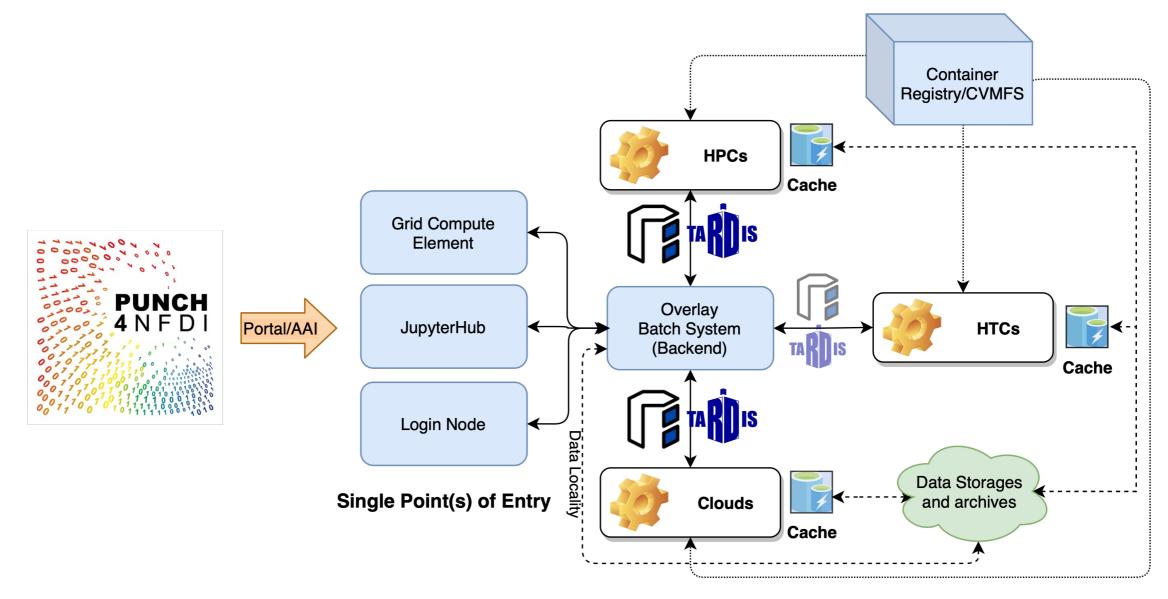## First (still less demanding) workflows from the community are being ported

- HIFIS transfer service already working between the two instances
- Still to be evaluated / implemented: Common namespace and file catalogues
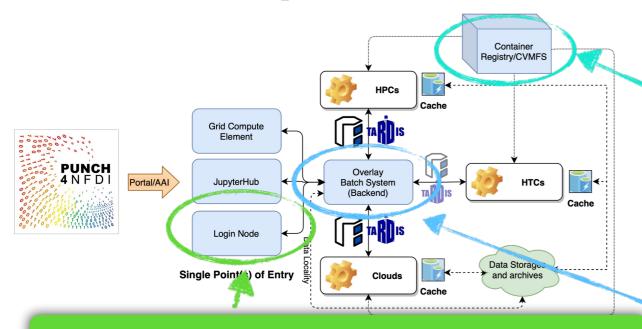
File Transfer Service - WebFTS

Status: Compute4PUNCH Infrastructure

Status: Compute4PUNCH Infrastructure

Container Registry/CVMFS

- Plan to use docker registry (gitlab-p4n.aip.de)
- Dockerfiles managed at GitLab Project (compute4punch/container-stacks)
- CI building and uploading docker containers
- Automated conversion into singularity sandbox format and distribution via CVMFS (unpacked.cern.ch)
- So far: GitHub->Docker Hub -> CVMFS
- Status: Work ongoing

Login Node

- First login node hosted at KIT (c4p-login.gridka.de)
- Managed by RH Satellite & Puppet (easy to re-install)
- Authentication using Helmholtz AAI (OIDC Token, 2FA?)
  - oidc-agent and mccli required for access
  - Local account is created using preferred user name registered in the Helmholtz AAI
- Operates an HTCondor Schedd to submit jobs
- Solution for shared home directories need investigations
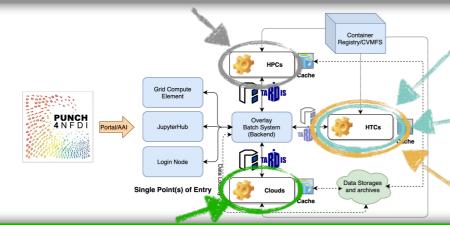- Status: Ready to be tested

Overlay Batch System

- HTCondor Central Manager hosted at KIT (c4p-htcondor.gridka.de)
- Managed by RH Satellite & Puppet (easy to re-install)
- Authentication using IDTOKENS (JWTs)
  - Dedicated token for all nodes of a compute cluster (on request)
  - Dedicated token for login nodes operating HTCondor Schedd to submit jobs (on request)
- Status: Ready to be tested

Status: Compute4PUNCH Infrastructure

MOGON II/Himster II (JGU Mainz)

- Shared access to HPC clusters (~1000 cores for tests)
- Dynamically managed by COBalD/TARDIS
- To be integrated into Compute4PUNCH
- Status: Work ongoing

High Throughput Cluster (TOPAS@KIT)

- 8 NVidia V100, 64 cores, 160 GB RAM (shared)
- More cores or RAM available on request
- Dynamically managed by COBalD/TARDIS
- Fully integrated into Compute4PUNCH
- Status: Ready to be tested!

GridKa Cluster (KIT)

- Up to 2000 cores, 2-3 GB RAM per core (shared)
- Dynamically managed by COBalD/TARDIS
- To be integrated into Compute4PUNCH
- Status: Work ongoing

OpenStack Cluster (WWU Münster)

- Contingent: 96 cores, 256 GB Memory, 1TB disk space
- Dynamically managed by COBalD/TARDIS
- Fully integrated into Compute4PUNCH
- Status: Ready to be tested!

Bonn Analysis Facility BAF (Bonn)

- 10 M core-hours on local compute cluster (3.5kCores)
- Dynamically managed by COBalD/TARDIS
- To be integrated into Compute4PUNCH
- Status: Work ongoing