Technical solutions across communities

FIDIUM: Federated Digital Infrastructures for Research on Universe and Matter

PUNCH: Particles, Universe, NuClei and Hadrons for the NFDI

Oliver Freyermuth

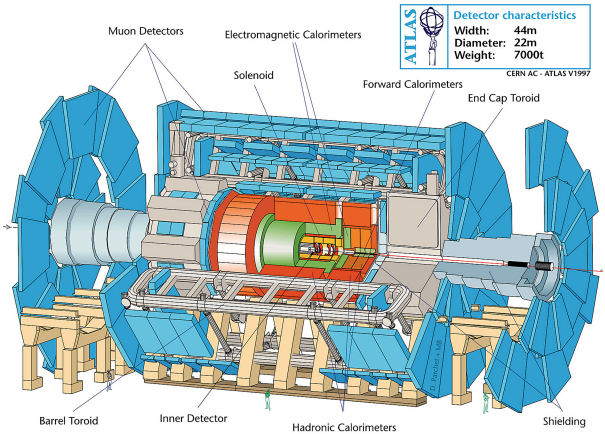
University of Bonn
o.freyermuth@uni-bonn.de

17th January, 2023

FIDIUM

UNIVERSITÄT BONN

ErUM: ‘Erforschung von Universum und Materie’



- Huge experiments, example: ATLAS experiment at CERN, raw data: $60\,^{\mathrm{TB}}/_{\mathrm{s}}$ from 1.7 billion particle collisions
- Hardware trigger selects up to 100 000 events per second
- 40 000 CPU cores are used to decide on the 1000 events to keep per second
- Collected data is the basis for hundreds of different studies

⇒ How is this data analyzed?

Computing model

- Multi-Tier hierarchical model: Tier 0 for Raw Data storage
- Calibration, Simulation, Preprocessing on the Worldwide LHC Computing Grid
- Finally, analysis work done at Universities using local resources (Tier 3)

Example: ATLAS groups at the University of Bonn

- Analysis on local DFG financed HTC resources in 2.5 years:
  20 PhDs, 27 Master theses, 19 Bachelor theses
- Significant compute power with High Throughput and local storage needed with high bandwidth from / to outside world
- Resource requirements increase steadily (e.g. detector upgrades)
  $\Rightarrow$ Action Plan ErUM-Data

FIDIUM Project

- Financed by the BMBF, part of ErUM-Data
- Common project by groups from elementary particle physics, hadron physics, nuclear physics and astroparticlephysics
- One project in a series of projects:
  Q3/2017 → Q3/2021 ErUM-Data Pilot I: IDT-UM
  Q3/2021 → Q3/2024 ErUM-Data Pilot II: **FIDIUM**
  Q3/2024 → Q3/2025 N.N.
  Q3/2025 → Q3/2028 ErUM-Data 'föderierte Digitalinfrastrukturen' I
  Q3/2028 → Q3/2031 ErUM-Data 'föderierte Digitalinfrastrukturen' II

FIDIUM Collaborators

10 institutions

- Rheinisch-Westfälische Technische Hochschule **Aachen**
- Rheinische Friedrich-Wilhelms-Universität **Bonn**
- Goethe Universität **Frankfurt am Main**
- Albert-Ludwigs-Universität **Freiburg**
- Georg-August-Universität **Göttingen**
- Universität **Hamburg**
- **Karlsruhe**r Institut für Technologie
- Johannes Gutenberg-Universität **Mainz**
- Ludwig-Maximilians-Universität **München**
- Bergische Universität **Wuppertal**

4 associated partners

- CERN
- DESY
- GridKa
- GSI Helmholtzzentrum für Schwerionen-forschung

FIDIUM

UNIVERSITÄT BONN

FIDIUM Structure & Goals

Topic Area I **Development of tools to leverage heterogeneous resources**
Coordinators: Oliver Freyermuth, Manuel Giffels

Topic Area II **Data Lakes, Distributed Data, Caching**
Coordinators: André Brinkmann, Kilian Schwarz

Topic Area III **Testing, tuning & optimization in production & analysis environments**
Coordinators: Christian Zeitnitz, Günter Duckeck

Goals

- Enable transparent use of HPCs & Cloud infrastructures
- Flexible use of resources which are only available temporarily (e.g. backfilling)
- Concentration of most production data on large centres with local caches:
  $\Rightarrow$ Concept of data lakes

FIDIUM Activities in Bonn

- Project 'Compute Site in a Box'
  - Usage of COBaID/TARDIS (tools developed at KIT for opportunistic computing)
  - Opening and extending existing Puppet configuration, test deployment with partners
  - Generalize unprivileged container workflows for COBaID/TARDIS resources
  - Leverage publicly accessible container infrastructure
  - Partners Mainz and GSI Darmstadt as 'testbeds'
  - Documentation and workshops for inside and outside particle physics community
- Funding: 1 FTE for three years, equal own contribution
- Close integration with developments in other topic areas (Accounting, Caching, Monitoring)

Transparent integration of Compute resources

Main steps

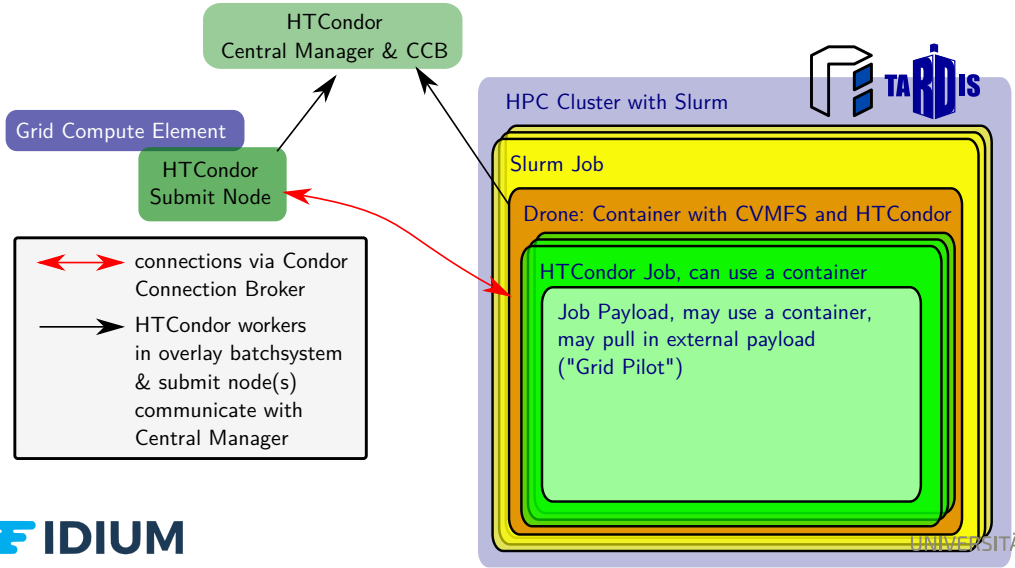
- Jobs submitted locally ⇒ Execution on **Overlay Batch System**
- Software stack as container (shipped via Cern-VM FS)
- Unprivileged containers started as jobs on site: Leverage User Namespaces
- Inside, an HTCondor `startd` (execute node) is started ⇒ 'Drone'
- Drone registers with Central Manager of the Overlay Batch System
- Jobs inside Overlay Batch system can use containers themselves

Scaling of number of 'Drones' based on resource usage

- COBaID — the Opportunistic Balancing Daemon
- TARDIS — The Transparent Adaptive Resource Dynamic Integration System

Successfully used with HPC BONNA

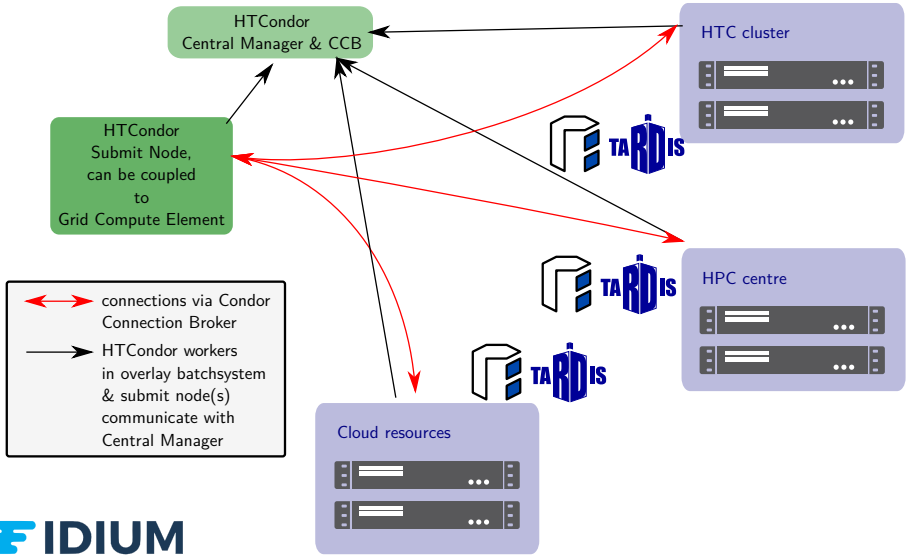
Federating compute resources



IDIUM

UNIVERSITÄT BONN

Federating compute resources



PUNCH: On the way to a Science Data Platform

Particles, **U**niverse, **NuC**lei and **H**adrons for the NFDI

Goal

Federated / FAIR 'Science Data Platform'
Provide infrastructure and interfaces for access to and working with data and compute resources of the communities, breaking community borders

Activities in Bonn with regard to technical infrastructure

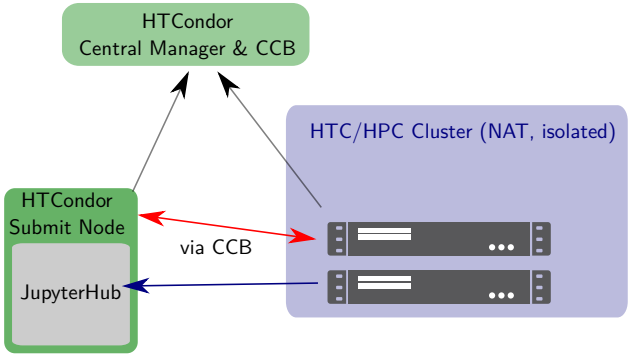
- JupyterHub frontend for federated Compute infrastructure ('Single Point of Entry')
- Including resources in Bonn in the Compute infrastructure
- Federated storage for 'small' experiments

Other activities in Bonn

Metadata structure and definitions, Outreach activities

Interactive use of federated resources

- Leverage reverse connections via the batch system
- Interactive compute jobs in federated infrastructure / opportunistic resources possible
- Use cases: Interactive jobs, JupyterLab environments,...



Oliver Freyermuth, Katrin Kohl, Peter Wienemann

Unleashing JupyterHub:
Exploiting resources without inbound network connectivity using HTCondor

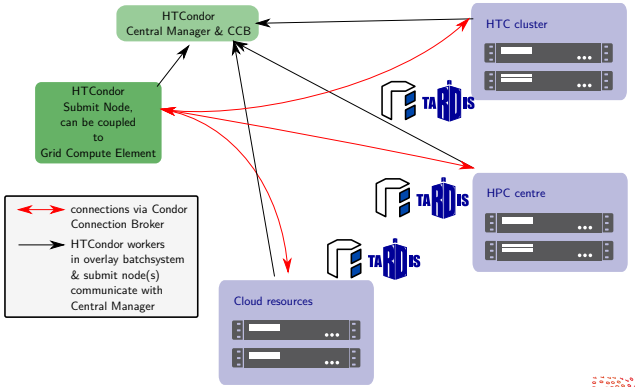
Computing and Software for Big Science 5, 24 (2021)

PUNCH 4NFDI

UNIVERSITÄT BONN

Interactive use of federated resources

- Leverage reverse connections via the batch system
- Interactive compute jobs in federated infrastructure / opportunistic resources possible
- Use cases: Interactive jobs, JupyterLab environments,...



Oliver Freyermuth, Katrin Kohl, Peter Wienemann

Unleashing JupyterHub:
Exploiting resources without inbound network connectivity using HTCondor

Computing and Software for Big Science 5, 24 (2021)

PUNCH 4NFDI

UNIVERSITÄT BONN

Storage for 'small' experiments

- Large experiments use storage systems with X.509 (certificate) authentication to enable distributed data management for their communities
- Ongoing migration to industry standard protocols (WebDAV / HTTPS, XRootD) and industry standard authentication mechanisms (JSON WebTokens / OpenID Connect)
- Necessary key feature: Server to server copies (Third-Party Copy)
- Employment of data catalogue (e.g. Rucio) with a common **global namespace**
- Small experiments (dozens of TB per year) do not have a Distributed Data Management system:
  `scp`, `rsync`, carrying disks, manual management of multiple copies...

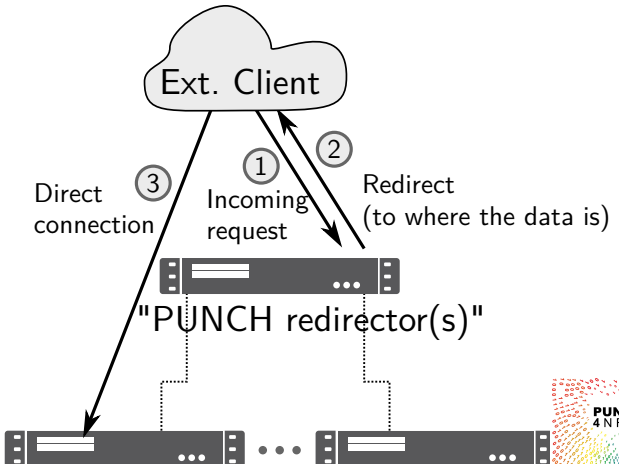
Goals

- Enable them to use lightweight storage software (e.g. Apache, XRootD)
- Allow remote authenticated data access from PUNCH infrastructures, HTC and HPC systems

UNIVERSITÄT BONN

XRootD: Federating data access across sites

- XRootD storage provides HTTPS / WebDAV and XRoottD protocols
- Global namespace: Use **g**lobal **L**ogical **F**ile **N**ames
- Redirection from central redirector to site with physical files ('merge' sites)
- Compute jobs can be routed to sites with data, or close to them
- Concept also works with caches



Conclusions

- Both projects aim to federate compute and storage resources to deal with increasing resource requirements
- Connection of federated compute and storage via federated AAI
- Leverage existing compute resources via an overlay batch system
- PUNCH Goal: break community borders & offer a FAIR Science Data Platform
  - One platform / portal to allow cross-experiment analyses
  - Usage not only for Open Data
  - Allow access to data from all communities via industry standard protocols
  - Common distributed data management infrastructure including caches

IDIUM

UNIVERSITÄT BONN

Thank you

for your attention!