

# Status Report DESY



Status update of developments in the FIDIUM project

Kilian Schwarz

FIDIUM Collaboration Meeting, Aachen, September 30 2024



# Table of contents

## 01 Overview

## 02 Topic Area 1

- Usage of opportunistic resources

## 03 Topic Area 2

- Dynamic data caching
- Data lake technologies
- Data lake prototypes

## 04 Topic Area 3

- Tests and deployment
- Optimisation and adjustments
- support

## 05 FIDIUM prolongation

- TA1
- TA2
- TA3

## 06 summary

# Status report DESY

## introduction

# Status Report DESY

## Introduction

### Application and participation

- DESY is associated partner
- Participation in
  - TA1: WP1 (opportunistic resources)
  - TA2: WP2 (caches), WP3 (data lakes), WP4 (prototypes)
  - TA3: WP1 (deployment), WP2 (optimisation), WP3 (support)

### Included expertise

- Operations of large scale computing centres
- WLCG, Usage of opportunistic resources
- Data lake technologies and dynamic caches
- Monitoring systems

## Föderierte Digitale Infrastrukturen für die Erforschung von Universum und Materie (FIDIUM)

Gemeinsamer Antrag von Gruppen aus den Bereichen Elementarteilchenphysik, Hadronen- und Kernphysik und Astroteilchenphysik

- Rheinisch-Westfälische Technische Hochschule Aachen, Prof. Dr. Alexander Schmid<sup>1</sup>
- Rheinische Friedrich-Wilhelms-Universität Bonn, PD Dr. Philip Bechtle
- Goethe Universität Frankfurt am Main, Prof. Dr. Volker Lindenstruth
- Albert-Ludwigs-Universität Freiburg, Prof. Dr. Markus Schumacher
- Georg-August-Universität Göttingen, Prof. Dr. Arnulf Quadt
- Universität Hamburg, Prof. Dr. Johannes Haller
- Karlsruher Institut für Technologie, Prof. Dr. Günter Quast
- Johannes Gutenberg-Universität Mainz, Prof. Dr. Frank Maas
- Ludwig-Maximilians-Universität München, Prof. Dr. Thomas Kuhr
- Bergische Universität Wuppertal, Prof. Dr. Christian Zeitnitz

Assoziierte Partner sind

- CERN, Dr. Markus Elsing
- DESY, Prof. Dr. Volker Gülzow
- GridKa, Dr. Andreas Petzold
- GSI Helmholtzzentrum für Schwerionenforschung, Dr. Kilian Schwarz<sup>2</sup>

<sup>1</sup> Sprecher des Verbundes

<sup>2</sup> Stellvertretender Sprecher des Verbundes

# Status Report DESY

## Topic Area 1:

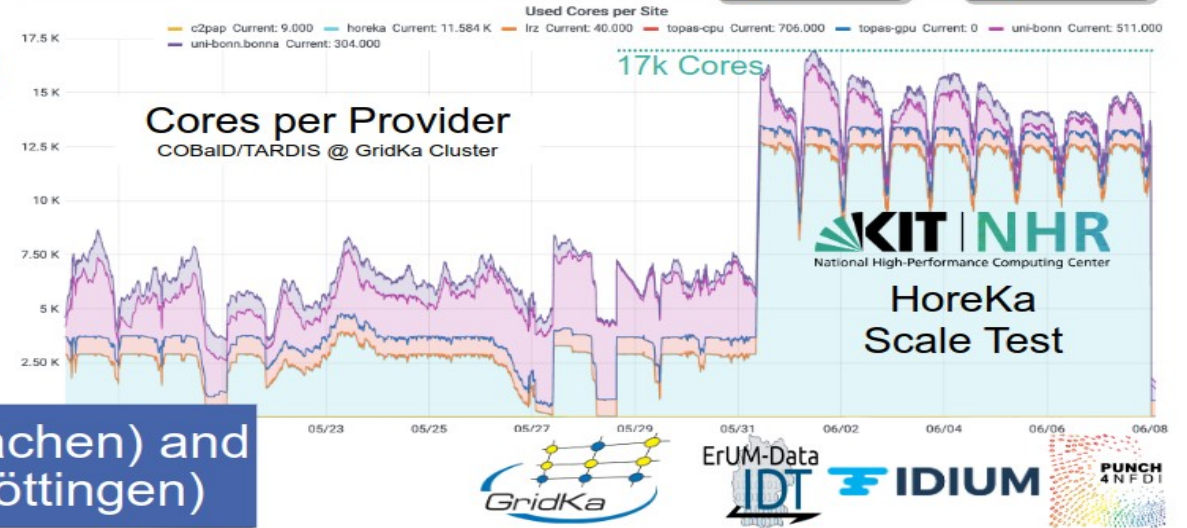
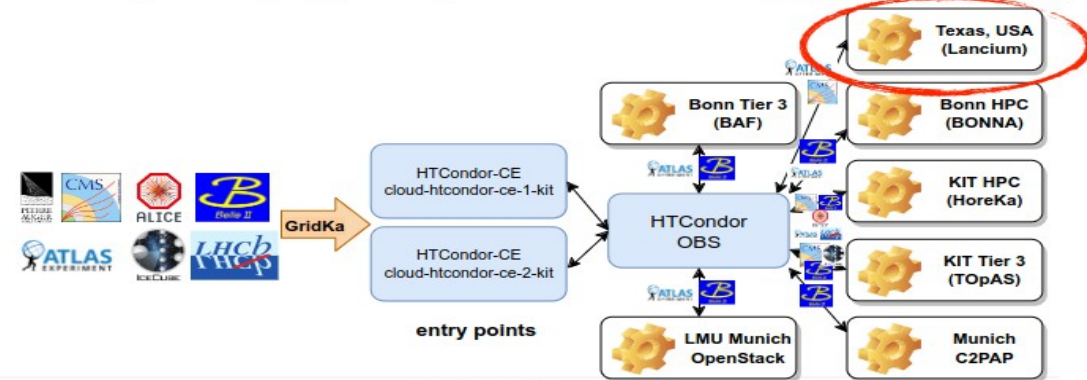
Tools for including  
heterogeneous resources

### Opportunistic Compute @ FIDIUM in a Nutshell

Simplify provisioning and utilization of third-party compute resources for the various communities:

- Dynamic, transparent and on-demand integration via COBaID/TARDIS (in-house development)
- Provide community-overarching unified entry points to a variety of resources (HPCs, Clouds, ...)
- Demonstrated production scale operation during scale test together with HoreKa (KIT HPC cluster)
- Production deployment across HEP institutes & HPC resources coordinated by KIT/GridKa
- Central building block of the Compute4PUNCH infrastructure within PUNCH4NFDI

Similar setup deployed at CLAIX HPC (RWTH Aachen) and on-going deployment at Emmy (University of Göttingen)



# FIDIUM

## Topic Area 1

### WP 1: including opportunistic resources efficiently

- Including DESY NAF resources into COBaID/TARDIS infrastructure
- Status:
  - Works all well in virtual test environment, jobs were running successfully
  - Now transition to production system
    - Local NAF HTCondor scheduler needs to be reinstalled due to upgrade to new OS (RHEL9)
    - No shared cluster file system
      - Need to submit from “cluster remote” submit node via “condor\_submit -spool”
      - This needs to be included into C/T software
  -

# Status Report DESY

## Topic Area 2:

**Data Lakes, Distributed Data  
Caching**



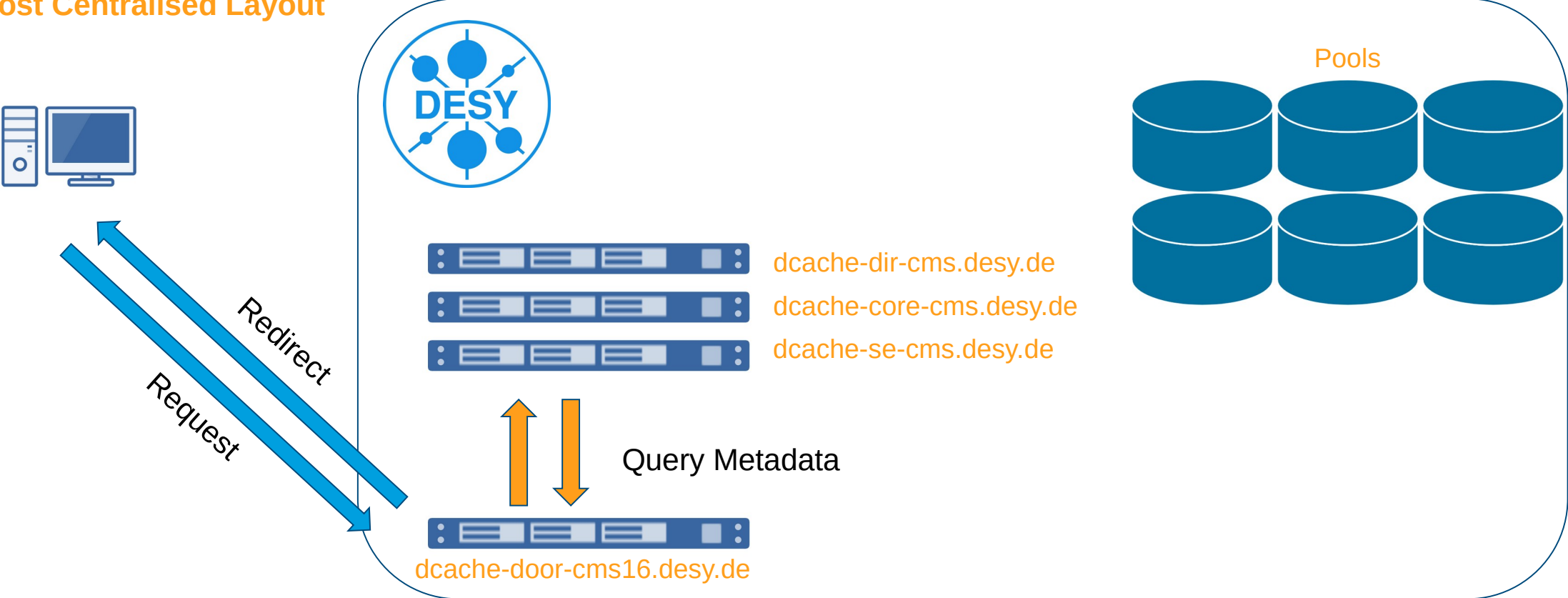
### WP 2 & 3: dynamic data caches & data lake technologies

- DESY IT/SC is lead developer of the federated storage middleware dCache
  - In use at many sites in WLCG as well as in other communities (e.g. photon science, astro particle, accelerator science ...)
  - Tape backend via CTA
- dCache already provides many required technologies for data lakes
- dCache enables lightweight site installations or caches
  - Data access and scaling tests are foreseen with GWDG



# Layout of Federated dCache

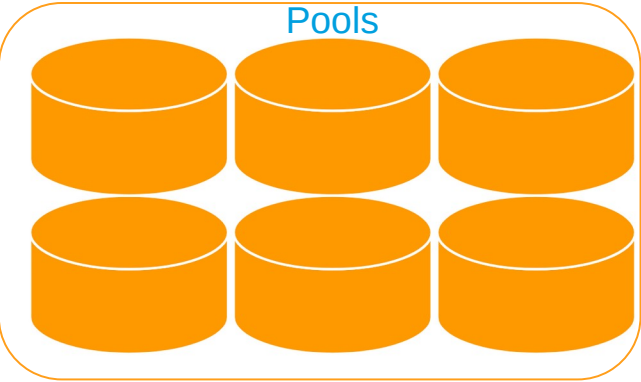
Simplest, most Centralised Layout



- Use dCache: Access to **/punch/<remote-site>/ildg**



Access via protocol of choice

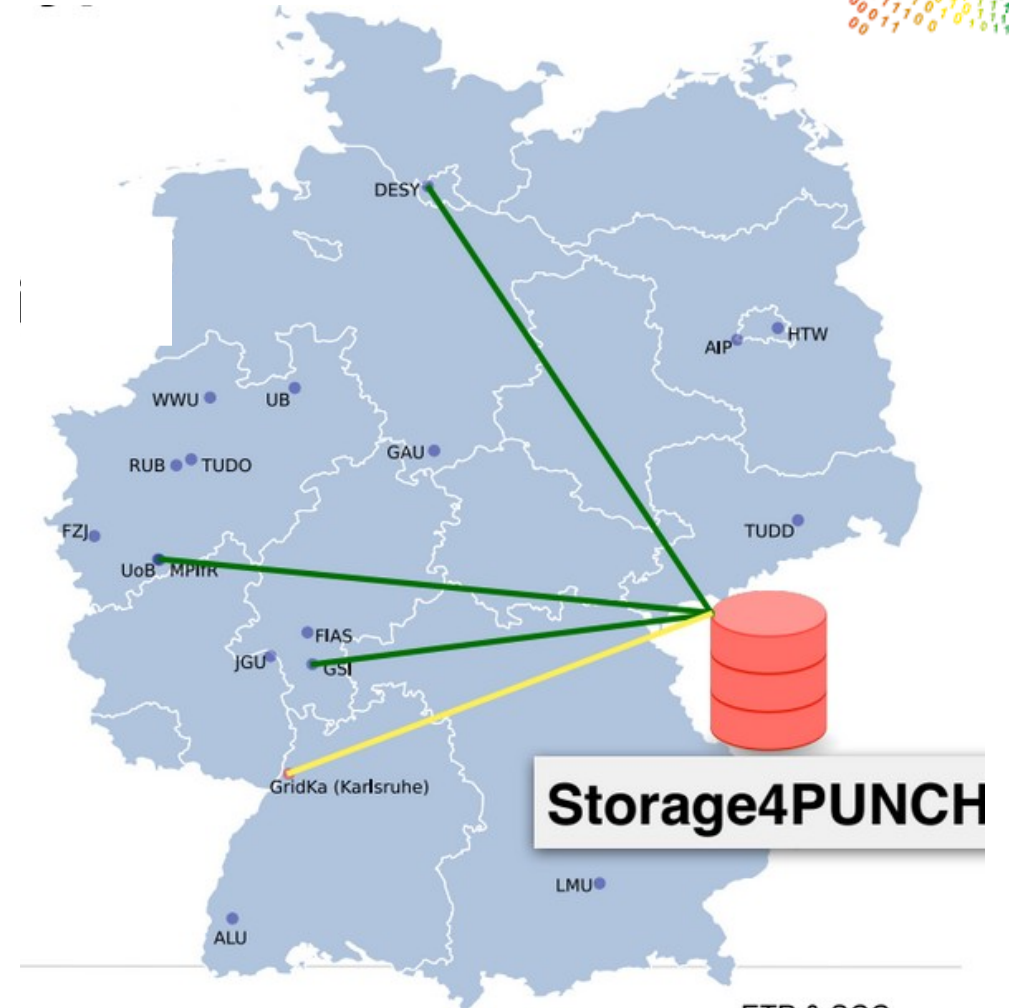


Remote Site

# FIDIUM

## TA2/WP4/prototypes

- Common testbed together with PUNCH4NFDI
- Based on storage middleware XrootD (Bonn, GSI) and dCache (DESY, KIT)
- Token based access



# Status Report DESY

## Topic Area 3:

adjustment, tests and  
optimisation

# FIDIUM

## Topic Area 3

### WP 1: deployment

- DESY contributes to deployment of COBaID/TARDIS and to dCache deployments at various sites

### WP 2: optimisation

- Including DESY NAF resources into COBaID/TARDIS infrastructure
- Optimising dCache installation procedures (also lightweight installations at sites)
- Optimising dCache related storage workflows with respect to functionality and performance

### WP 3: support

- Contributing with support experience at DESY IDAF

# Status Report DESY FIDIUM Prolongation up to Q3/2025

# FIDIUM prolongation

## TA2

### WP2: caching

- Simplified caching systems for opportunistic resources
  - Dynamic data cache via dCache

### WP2: workflows

- Cache-aware data management and workflows
  - Cache awareness in dCache
  - Comparison between dCache and XCache

# FIDIUM prolongation

## TA3

### WP1: tests, optimisation, deployment

- DESY, Hamburg: prototype dCache instance connecting DESY NAF and PhysNet Uni Hamburg
- DESY, Wuppertal: federated dCache instance between DESY and Wuppertal, performance measurement, tests for cache awareness of dCache and XCache

### WP2: adjustments

- DESY, Hamburg: COBaID/TARDIS between DESY NAF and PhysNet Uni Hamburg, performance measurements

### WP3: support

- Contribution to AUDITOR support
- Tutorials, documentation



# FIDIUM

## Summary and outview

# FIDIUM

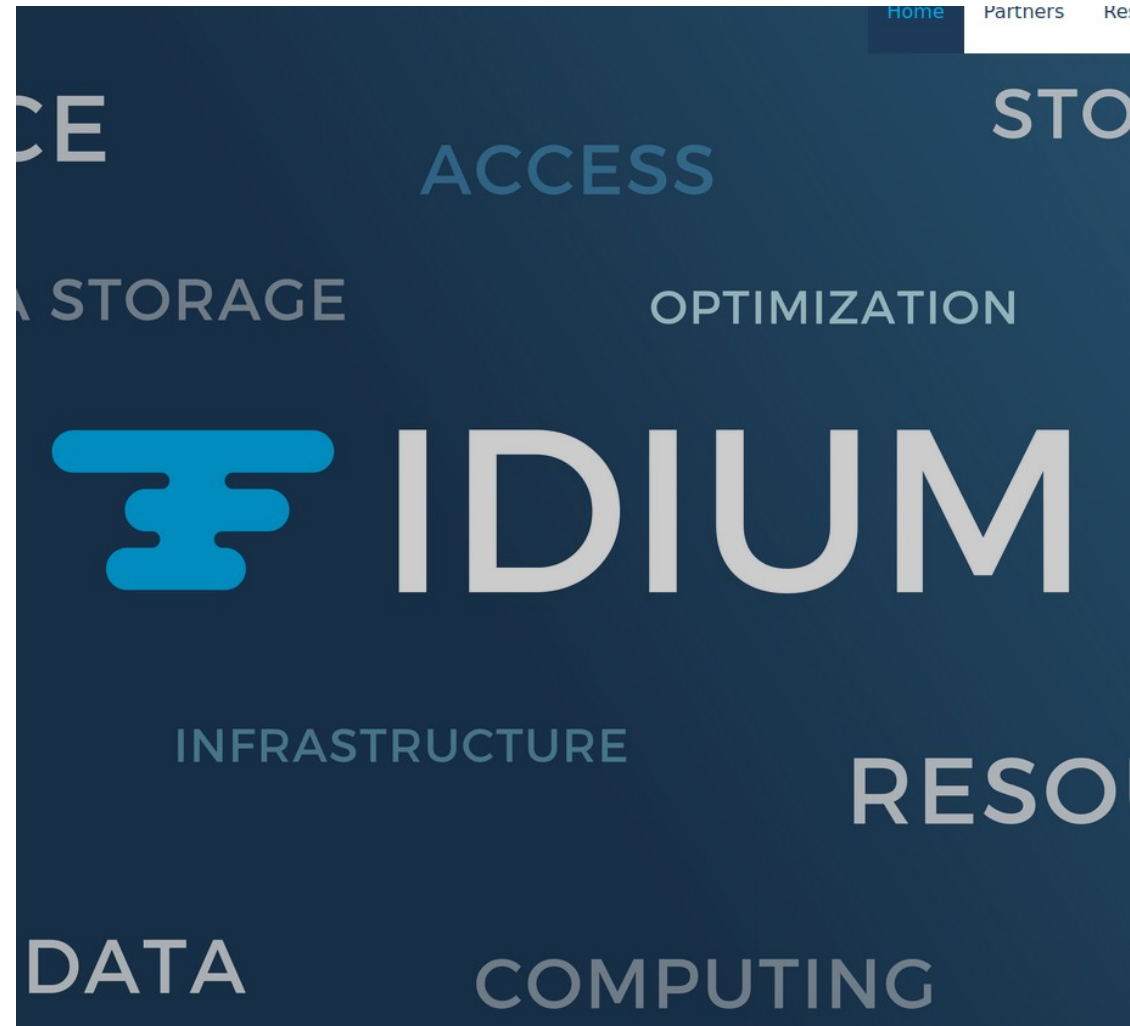
## Summary

### Consortium

- Common application of 10 universities, 3 Helmholtz centres & CERN, 3 communities (KET, KHuK, KAT), submitted autumn 2020
- Funding period Q1 2021 – Q3, 2024, prolonged up to Q3 2025
- See <https://fidium.erumdatahub.de/>

### Mission

- Successfully addressing the mission to develop experiment overarching software so that the experiments are ready to face the challenges of HL-LHC era
- Important software tools for distributed computing, federated storage, monitoring, and accounting are being provided by FIDIUM. DESY contributes to many aspects.



**Thank you**

## Contact

Deutsches Elektronen-  
Synchrotron DESY

[www.desy.de](http://www.desy.de)

Kilian Schwarz  
IT/Scientific Computing  
[Kilian.schwarz@desy.de](mailto:Kilian.schwarz@desy.de)  
040 8998 2596