Status & Plans - Area A & B

ErUM Data IDT Collaboration Meeting 2021

Manuel Giffels & Kilian Schwarz & Christian Zeitnitz, 10.05.2021

TOPICS OF AREA A

Development of Technologies to Enable Utilization of Heterogeneous Computing Resources

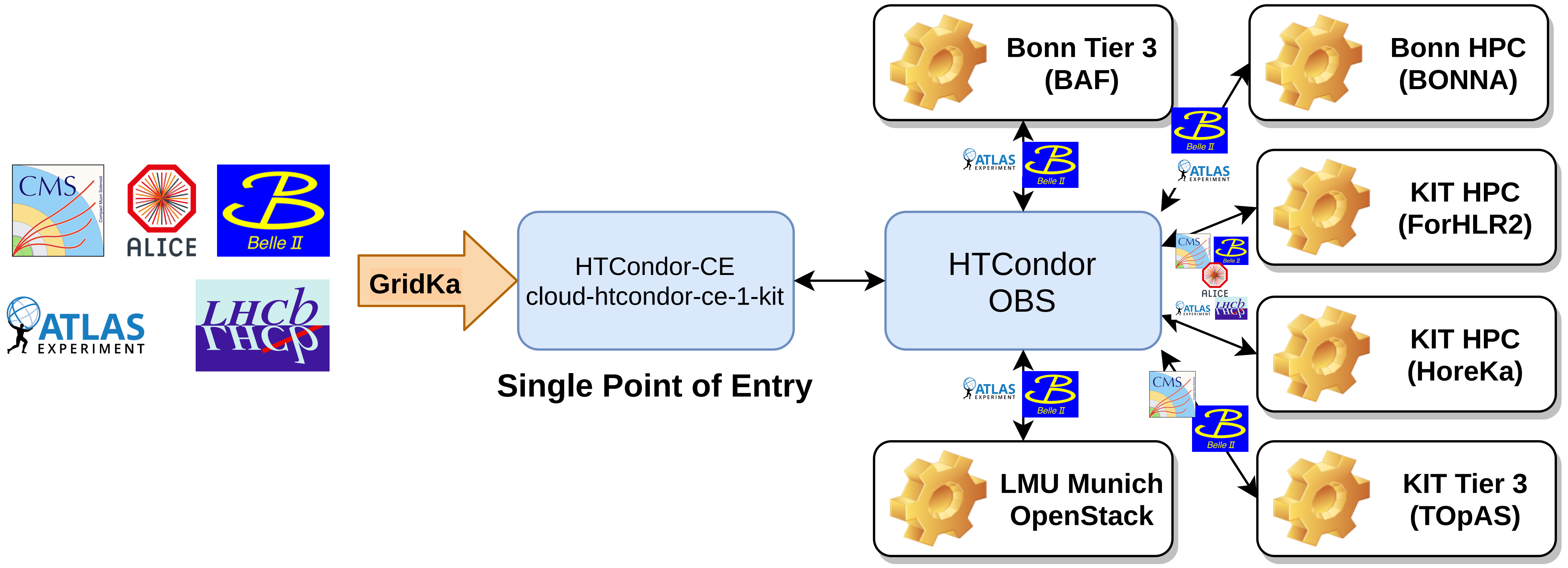
AP 1) Werkzeuge zur Einbindung • Scheduling von Cloud - Jobs • Container - Technologien • Checkpointing • Zugang zu Experiment-Datenbanken	**AP 2) Effiziente Nutzung** • Steigerung der Effizienz von daten-intensiven Anwendungen auf hetero-genen Ressourcen mittels „on the fly“ Datencaches
AP 3) Workflow Steuerung • Identifikation und Steuerung • In - Pilot Job Monitoring • Accounting • Optimierung durch data - mining	

STATUS & PLANS

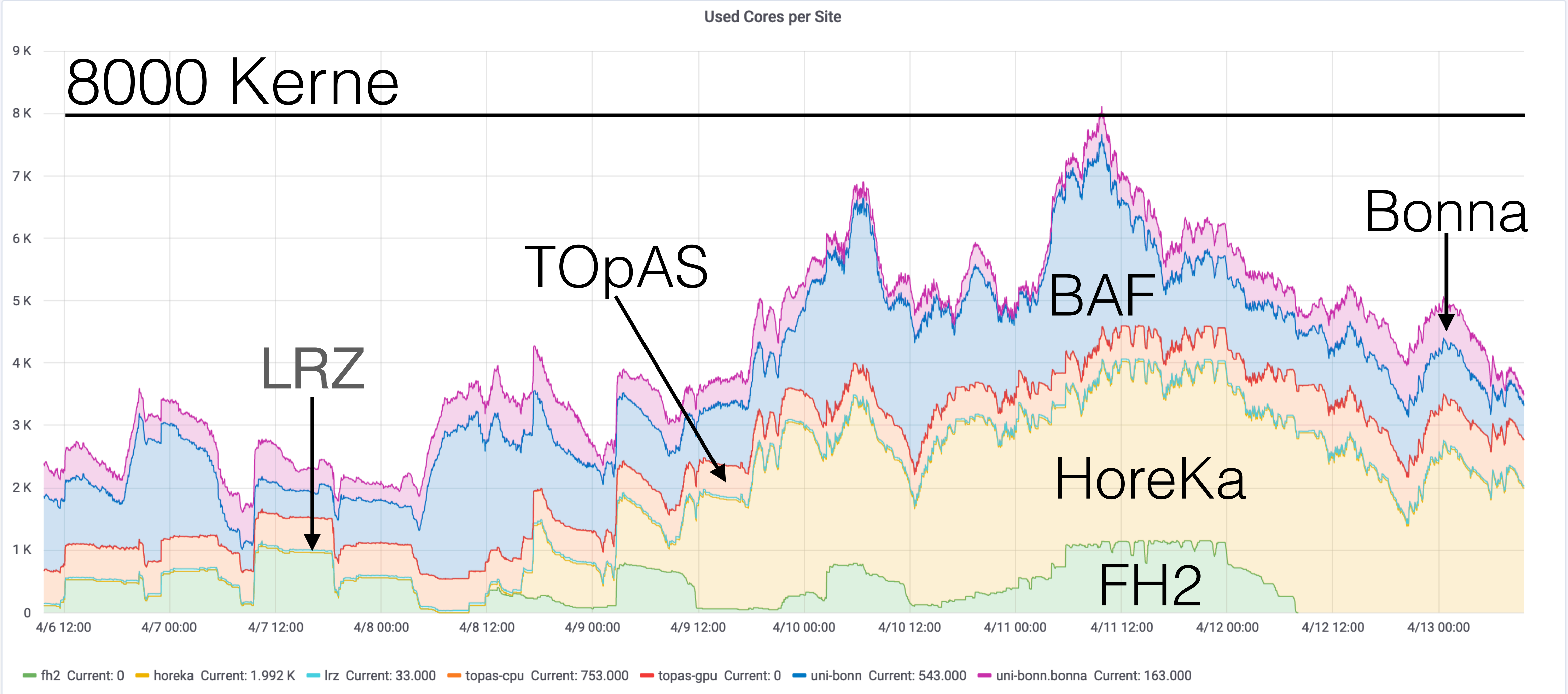
Anything related to compute resources

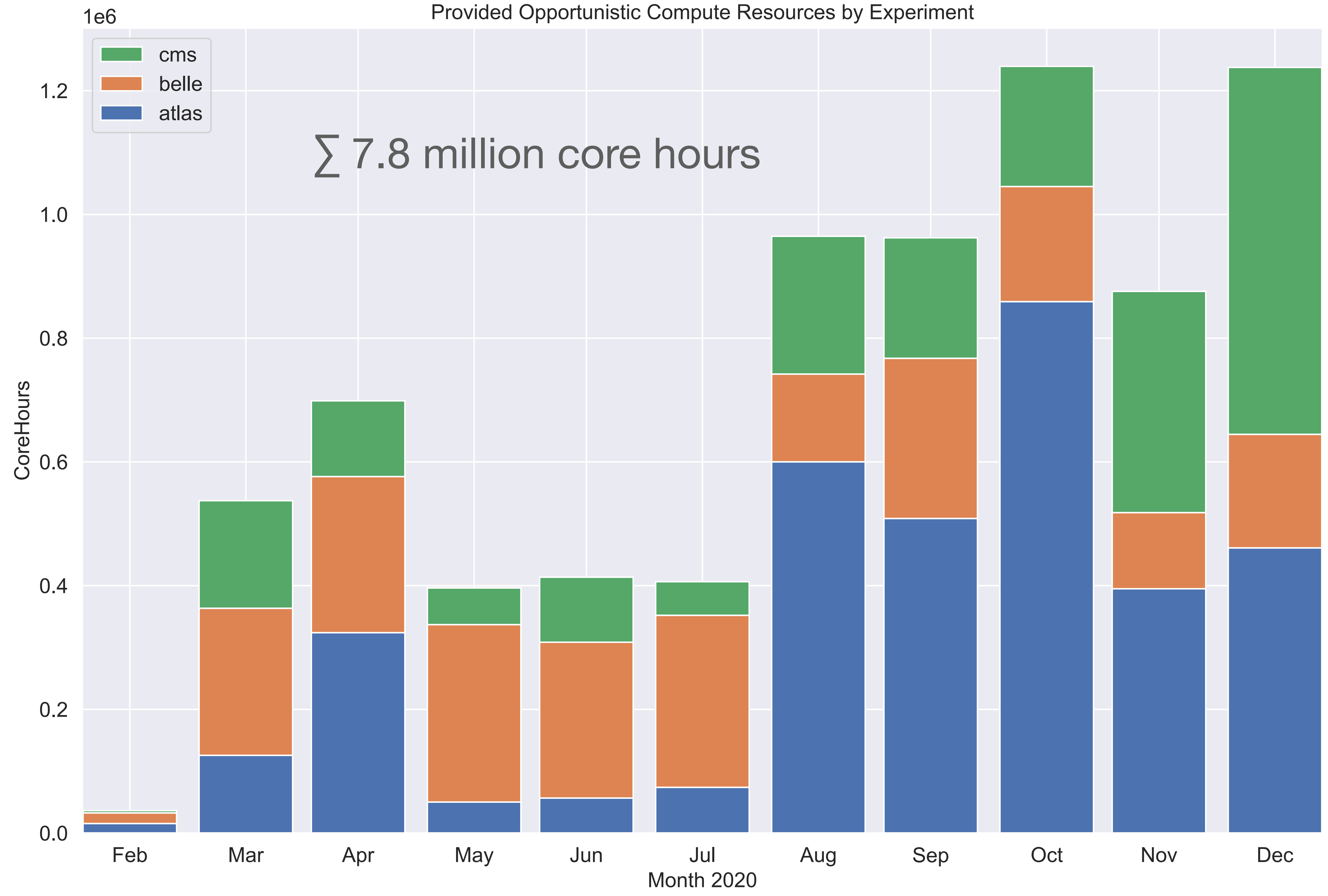
Federated Infrastructure

- Experiment-agnostic single point of entry to plethora of resources (HPCs, Tier-3s and Clouds)
- Dynamic, on demand provisioning and transparent integration of heterogeneous resources with COBalD/TARDIS
- Stable operation since ~ 2 years
- Toped the 8000-cores mark for the first time
- First GPUs integrated (32 NVidia V100, TOpAS)
- Fruitful collaboration (U Bonn, U Freiburg, KIT, LMU, RWTH Aachen, U Wuppertal joint recently)



Next: Integration of the LMU C2PAP cluster

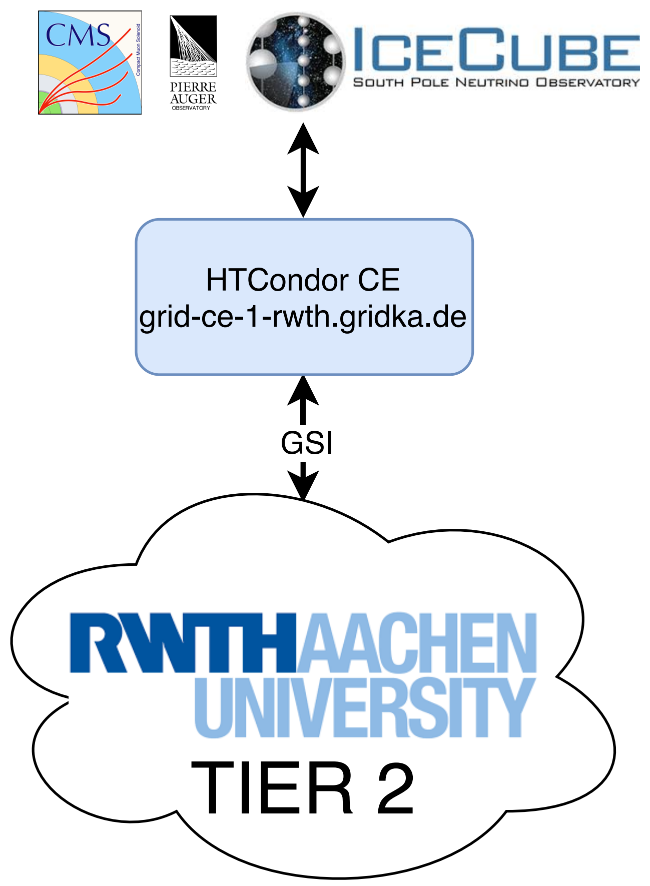




Lightweight Site Operation & Transparent Extension of a Tier-2

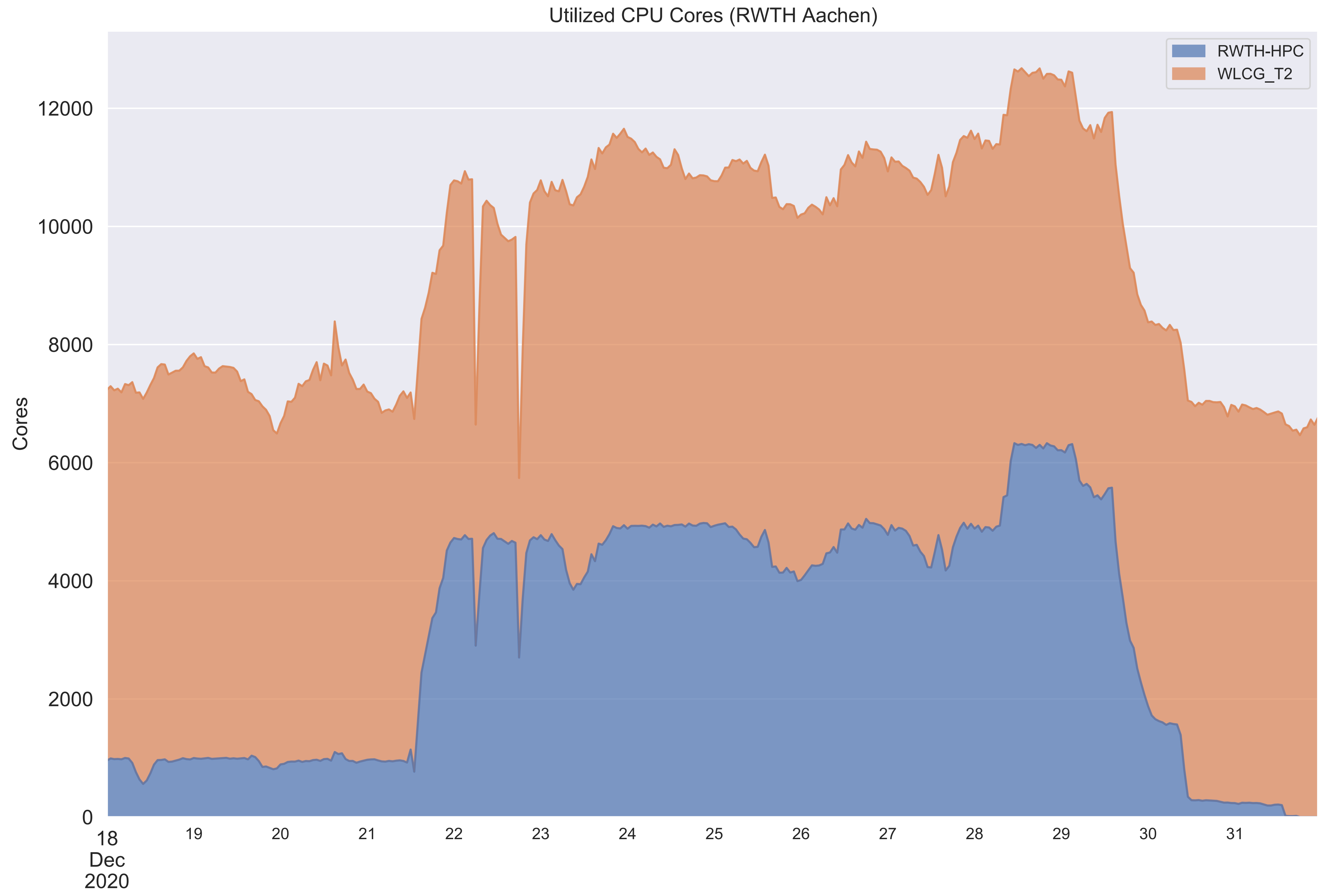
Lightweight Site Operation

- Remote CE for RWTH operated at GridKa (as one of many)
- No need for local CE any more
- Simplifies contribution of compute power for smaller sites



First Transparent Extension of a Tier-2 Site in Germany

- RWTH HPC Cluster CLAIX dynamically integrated in Aachen Tier 2 site using COBalD/TARDIS
- Doubled the CPU capacity of the RWTH Aachen CMS Tier-2 for a week
- Until submission was blocked by HPC team after using ~7 times the monthly quota assigned to the project
- Initial grant over 2 million core hours



OPERATIONAL IMPROVEMENTS

- Bonna HPC cluster:
  - /scratch moved to dedicated 2 TB SSD (used to be <100 GB on HDD)
  - Generic Resource (GRES) for scratch space added (large scratch space usage was an issue in the past), supported by TARDIS (see https://github.com/MatterMiners/tardis/issues/147)
  - Kernel update allowed to use unprivileged namespace FUSE mounts by cvmfsexec
- Added option to ship drone logs to COBalD/TARDIS node (to simplify debugging, see next slides)
- Added switches to ensure drones are pinned to VOs and only one job at a time per drone (to simplify debugging, see next slides)

UNIVERSITÄT BONN

OPERATIONAL IMPROVEMENTS and Bug Hunting

- Observed problematic failure rate of ATLAS jobs in DE-TARDIS Panda queue (up to 50%)
  - Requested memory is strictly enforced on most opportunistic resources (not the case in WLCG [soft limits])
  - Memory measured using proportional set size (PSS), while batch system/kernel uses cgroups/resident set size (RSS)
  - Found bug in ATLAS Pilot that commits parenticide in case PID namespaces are used
- Found race condition in TARDIS (failure rate below 1%)
  - Drone starts and accepts payload
  - Drone removed from batch system before it is visible to TARDIS
- Challenging/time consuming to debug

Annotation by M. Giffels:

Major improvements of the entire infrastructure could be archived with the help of U Bonn

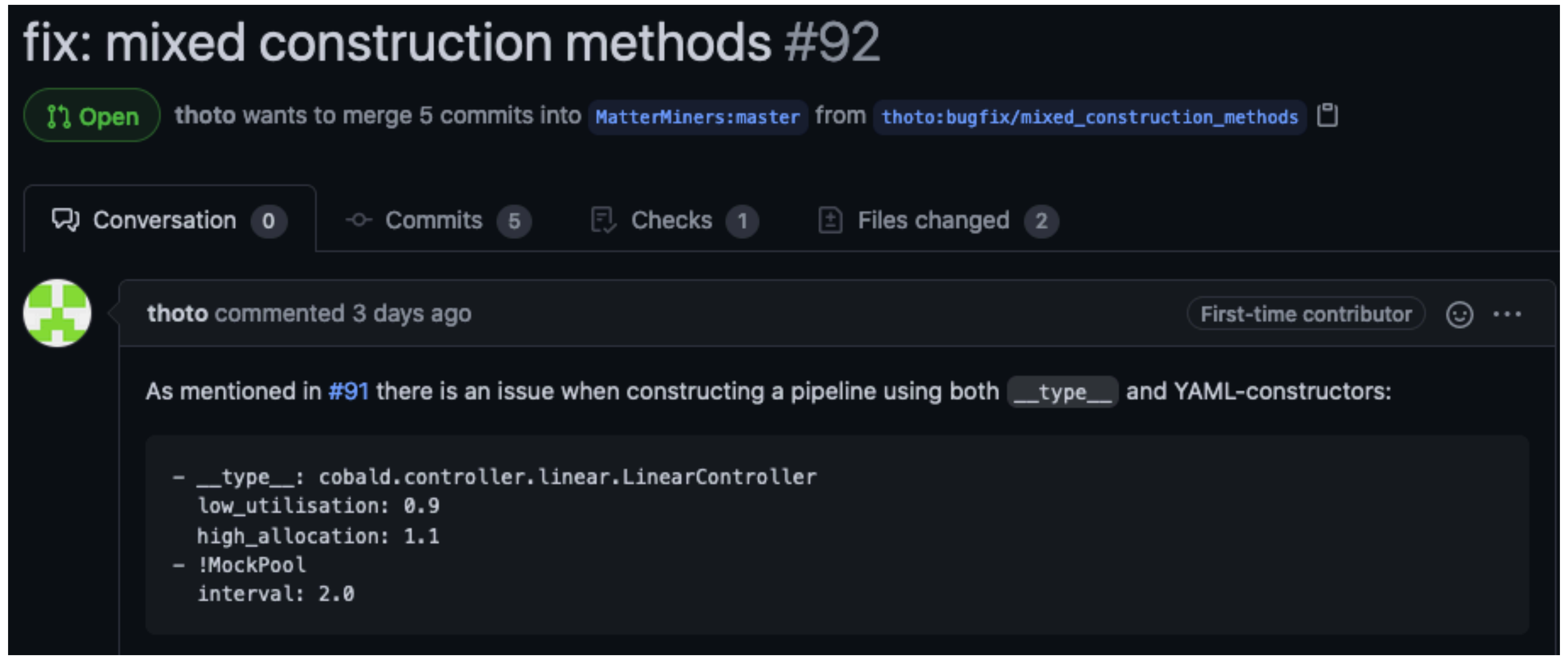
PRESENT STATUS

- After a lot of bug hunting situation looks promising
- Most issues were not caused by the COBalD/TARDIS layer.
- Operation of COBalD/TARDIS itself very smooth and low-maintenance
- Opportunistic usage harmonizes well with local usage

Activities Wuppertal

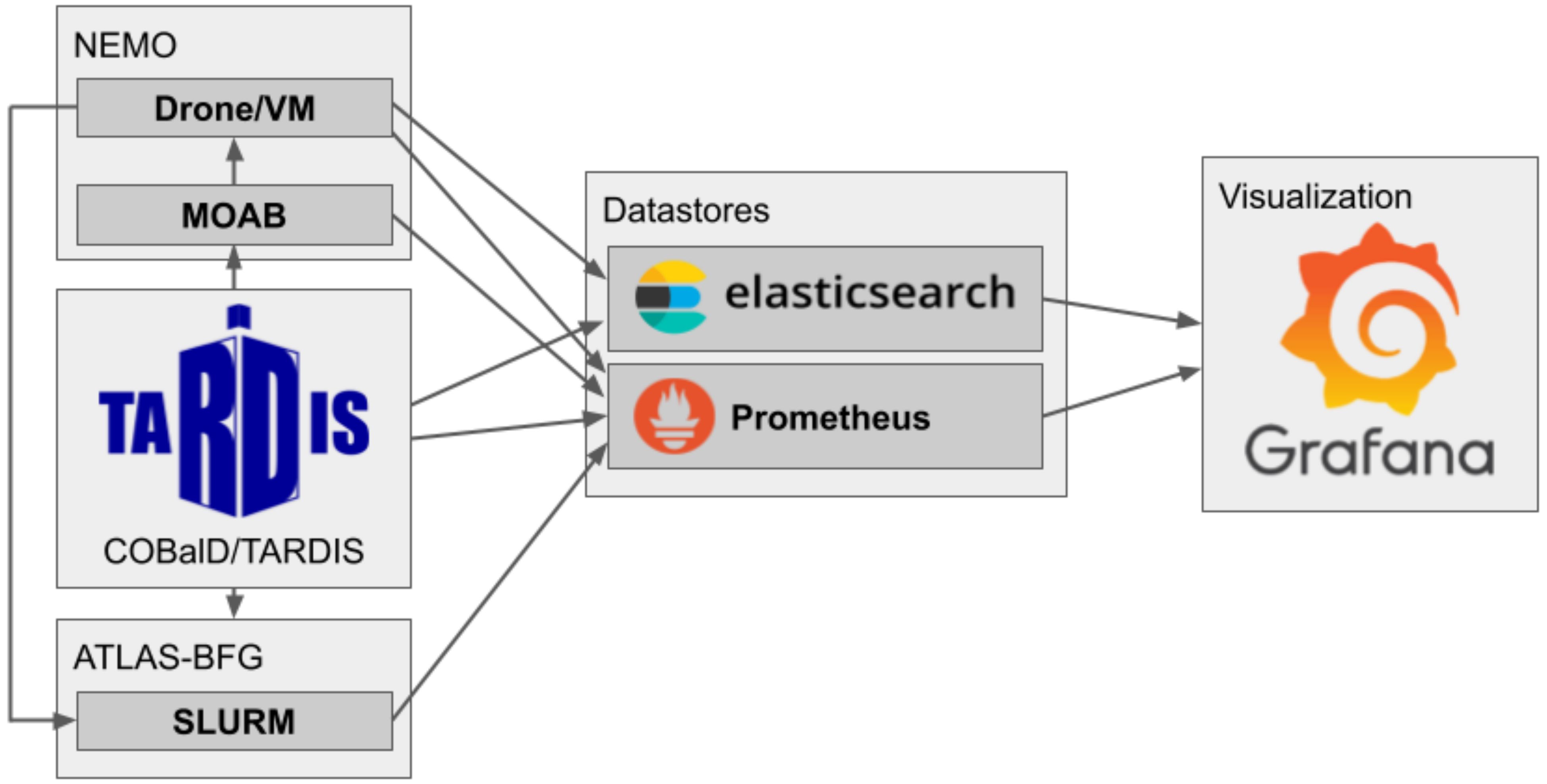
- First steps into using COBalD/TARDIS together with the „Pleiades“ cluster in Wuppertal
- Complicated by very late installation of a new cluster in May/June
- New batch system (SLURM)
- Plan to integrate part of the cluster, at least for a testing period, into the system by the end of the summer

First pull request already arrived:



Monitoring

- Collecting data from various sources in multiple datastores
- Started working on a monitoring puppet module will allow for the deployment of a monitoring setup for sites running COBalD/TARDIS

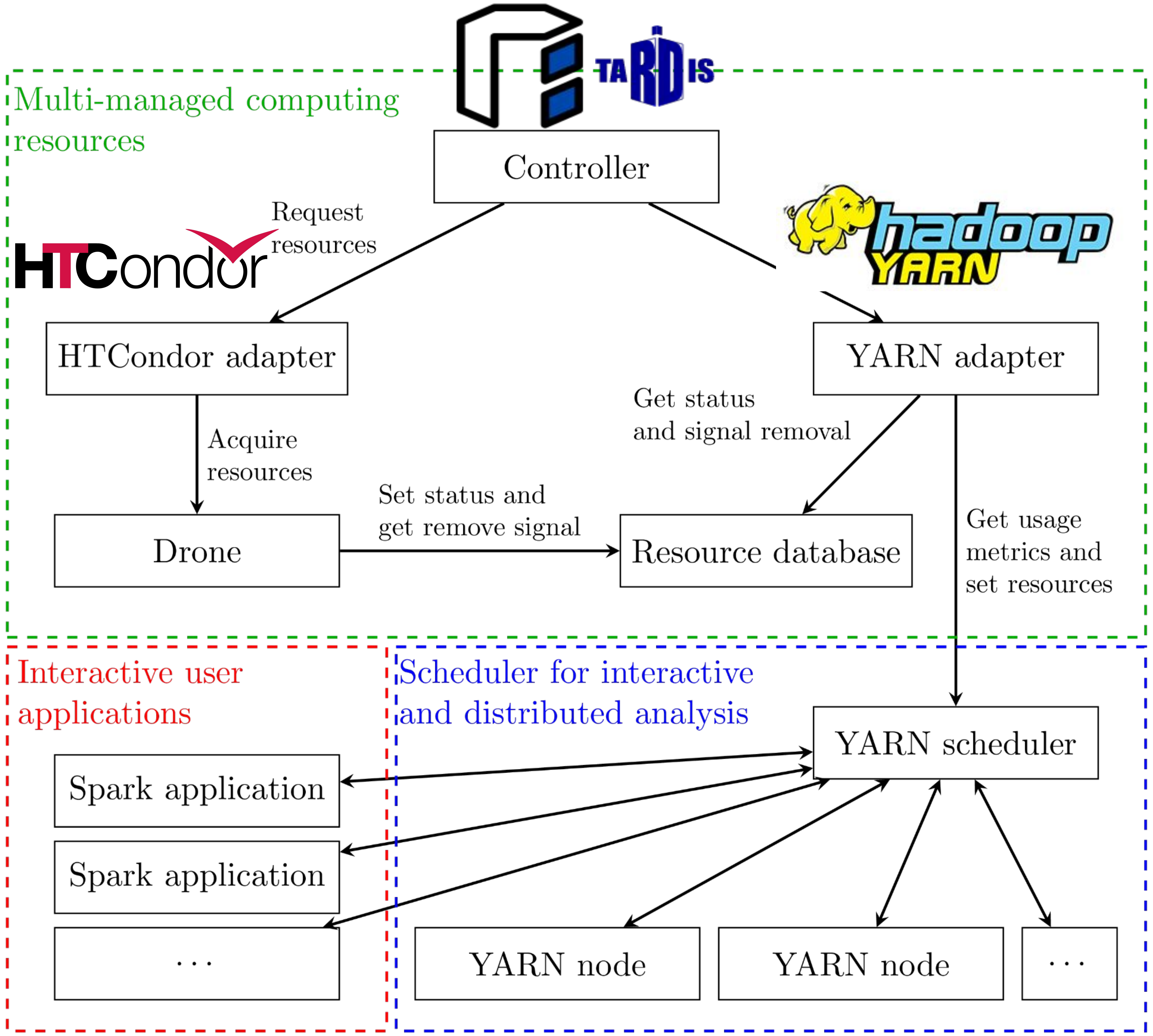


Accounting

- Implemented a prototype of an accounting system for opportunistic resources
- Next step will be to evaluate it on our local setup to transfer fairshares from one batch system to the other
- More details: session Area A+B: 09:40 talk by Stefan Kroboth

Distributed computing on multi-managed cluster resources

- Enable interactive distributed computing **cluster on demand** using available resources
- Coordinate **exchange** of resources between **HTCondor** and Spark/**YARN** thanks to CoBalD/TARDIS
- **Allocate** cores to the Spark cluster when users run their applications, **give** them back to HTCondor during idle time.
- Support **FAIR** scheduling and **multiple user** applications running at the same time.



UNIVERSITAT POLITÈCNICA DE VALÈNCIA
Karlsruher Institut für Technologie

Report of Topic Area A/B Storage/Container

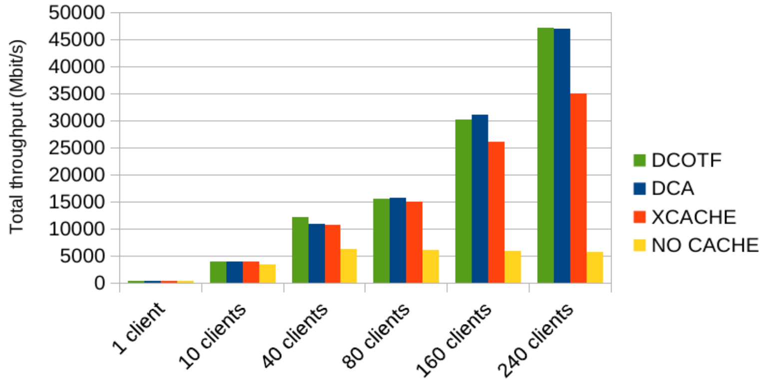
Kilian Schwarz
Manuel Giffels
Christian Zeitnitz

Area A/B -Storage/Container

- With contributions from
  - Frankfurt/GSI
  - Freiburg
  - KIT
  - München
  - Wuppertal

Area A/B – Storage/Container Frankfurt/GSI

- Singularity container for Grid environment running in production at GSI and tested in Frankfurt
  - Container image management in production
- Production ready XrootD based dynamic disk caching system DCOTF developed & provided
- Detailed scaling & performance measurements of DCOTF and XCache
  - DCOTF and Xcache DCA have similar performance
  - For Xcache there are bottlenecks
  - Containerisation does not affect the performance
- Simplified installation methods including documentation for DCOTF and container system
- Support provided for setting up similar Caching systems at KIT and Freiburg



Status report Freiburg

Opportunistic Resources - Disk-Caching-On-The-Fly Testing

- developed automatic provisioning of all components in Freiburg QA-environment (client, data manager, and forward proxy)
  - still some open issues before production-ready: file permissions of cached files, authentication to external sites, prefixing site url
  - **next steps:** put into production (after open issues are solved)
- performed benchmarks of caching setup on ATLAS xAODs
  - caching setup causes no overhead
  - cached files improve event loop time (exception: FR - KIT almost as fast as local)
- created experiment independent benchmark
  - plain ROOT files with pseudo-analysis event loop
  - **next steps:** perform benchmarks on several ATLAS sites (compare with xAOD benchmarks) and distribute plain-ROOT benchmark script and files
- More details: session Area A+B: 10:40 talk by Dirk Sammel

Open issues planned to be addressed in close collaboration with GSI/Frankfurt

Can be used for experiment Overarching comparison of Caching solutions

Area A/B – Storage/Container KIT

- Setup of dynamic disk caching systems at KIT and Freiburg in collaboration with Freiburg and Frankfurt/GSI

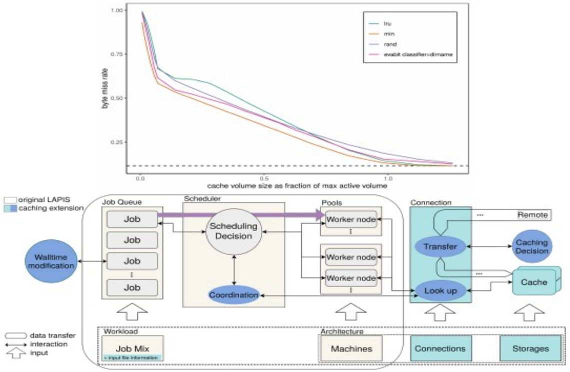
Karlsruhe Institute of Technology

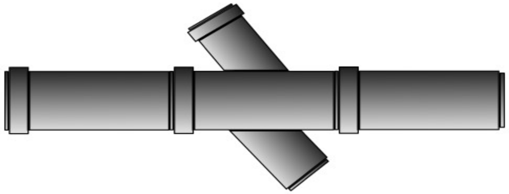
Solving the HEP Computing Challenge

- Optimization of existing workflows via caching poses complex optimization problem (Workflows, Schedulers, Architectures)
- Simulation provides an efficient ansatz

→ LAPIS simulator covers all necessary components, however

- Simplified model with walltime as central metric
- Convoluted basis mixing workload and architecture properties

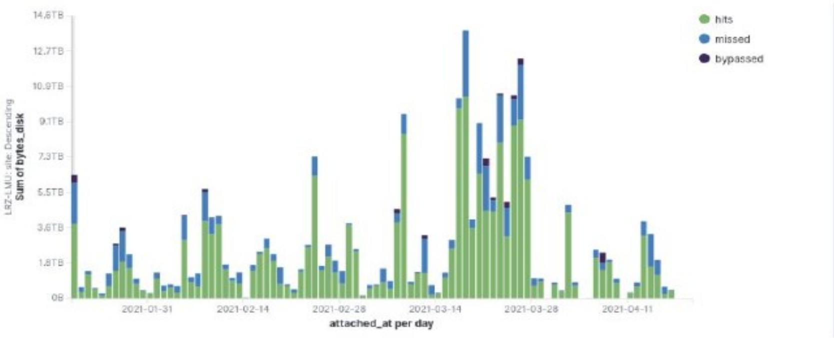




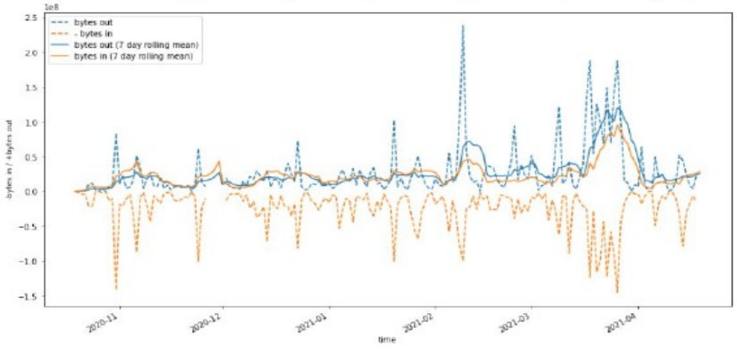
- More elaborate model: INFORMATION FLUX MODEL
  - Avoids convoluted basis
  - Able to represent arbitrary complex architectures and workloads
  - Simulate at any desired precision

Area A/B – Storage/Container München

Hit rate from VP monitoring (XRootD information)



Egress (blue) vs Ingress (orange)



- XCache serves ATLAS analysis job queue with “Virtual Placement system”: XCache acts as a “virtual” Rucio storage element that gets datasets assigned (but not actively transformed, just for job scheduling)
  ➜ helps to increase hit rates
- Sparse data access (read subset of ROOT Branches):
  ➜ Need to utilize XCache blockwise caching well
- Monitoring more difficult: Hit rate from XRootD monitoring does not match data Ingress/Egress
  - conceptual difficulties concerning clear definition of ‘cache hit rate’
- Testing smaller XCache block sizes for more efficient caching

Area A/B – Storage/Container Wuppertal

- Standalone container (no network access needed) integrated into ATLAS production environment
- Monitoring and Job-Reports for containerized user jobs
  - Idea
    - Containerized job (Docker and Sigularity)
    - Start service after the payload has finished
    - Aggregate informations from log files
    - Generate job report with important information about Warnings, Errors and statistics
    - Ship the aggregated data and job report to the user
    - Easy configuration for different job types
  - Allows to pin-point poblems quickly
  - Used software packages
    - Fluentd and Elasticsearch
- Prototype implementation for ATLAS jobs, but the frameworks has NO dependence on ATLAS software!
  - Substantial problems in Singularity to handle services
- Log-Analysis-Framework available at https://github.com/erum-data-idt/LogAnalysisFramework

From IDT-UM to Fidium

Christian Zeitnitz, Manuel Giffels, Kilian Schwarz

Bergische Universität Wuppertal

IDT-UM – Task Area A+B

- Well addressed topics
  - Utilization of different resources via COBalD/TARDIS for productions jobs of the experiments
  - Data caching solutions
  - Containerization of production jobs
  - Monitoring (system and job wise)
- Very limited or no progress (or not funded)
  - Handle complex resource requirements
    - Data locality, MPI, GPUs
  - Accounting and controlling
  - Simplified integration of resources ("Site in a box")
  - Failure tolerance of jobs/Checkpointing

What is needed for the LHC/FAIR Computing in 2025?

- Expect to enter the „Exabyte“ era
  - LHC Run 4
  - FAIR experiments
  - Large scale Astroparticle and Astronomy experiments
- Required resources not absolutely clear
  - Simple scaling for LHC: Factor 60 increase wrt. 2016
  - Technology progress and efficiency increase (hardware and software) hopefully leads to a decrease (~factor 6-10 ?)
- Complexity of workflows will increase
  - Need solutions across experiments
- Complexity of resource requirements will increase
  - Specialized hardware requirements (CPU, MPI, GPU, FPGA …)

Solutions for HL-LHC/FAIR Computing in 2025

- How to handle increase in hardware resources?
  - Concentrate (most) mass storage at larger computing centers (e.g. Helmholtz-centers)
  - Use all available resources – heterogeneous resources (dedicated, Cloud, HPC …)
- Need operational concept and software for different workflows across experiments
  - Often not possible to control the system/network environment
    - Further development of containerization needed
  - Complex resource management and accounting
  - High efficiency required – need fault tolerant solutions
- Adjustment of the software of the experiments needed!
- We have some very good projects and the know-how in IDT-UM
  - COBalD/TARDIS, Data caching, Containerization, Monitoring

Required developments

- Current software is NOT ready for HL-LHC/FAIR!
- Resource Management
  - Critical: never really worked in the past!
  - Further development of COBalD/TARDIS
  - Complex requirements: MPI, GPU, FPGA …
- Simple integration of available resources
- Data and Workflow management
  - Context: data-lakes, data caching
- Accounting
  - Complex and difficult in a heterogeneous environment

Required developments (2)

- Fault tolerance of the whole software infrastructure
  - Checkpointing
- Monitoring
  - Required for Resource management and accounting
  - Need special job monitoring for the users
    - Complexity of the computing environment makes the error analysis much more difficult
- Adaptation of the software of the experiments
- Internationalization of the developed software
  - Get the software into production at the experiments

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 Schmidt[1]
- 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[2]

Submitted to the BMBF end of Oct 2020

Project FIDIUM – required to be ready in 2025

Covers the required developments – three different topic area

- TA I: Development of Tools for the Integration of heterogeneous Resources

1. Erschließung und effiziente Einbindung von opportunistischen Ressourcen
   - Weiterentwicklung und Anpassung des Resource Managers COBalD/TARDIS an zukünftige Gegebenheiten
   - Entwicklung von dynamischer Steuerung des Job-Schedulings (z.B. Berücksichtigung von Datenlokalität, I/O-Raten)
   - Automatisierte Skalierung peripherer Dienste
   - „Compute Site in a Box“: Nutzbarmachung der Ressourcen an Tier-3- und Tier-2-Zentren mit minimalen zusätzlichen administrativen Ressourcen (volle Automatisierung, Skalierbarkeit)
2. Accounting und Controlling von heterogenen Ressourcen
   - Werkzeuge zum Accounting der opportunistisch genutzten Ressourcen
   - Tools für kontinuierliche Überwachung der Nutzungseffizienz

Project FIDIUM (2)

- TA II: Data-Lakes, Distributed Data, Caching

1. Aufbau eines Echtzeit Data-Lake-Monitoring-Systems
   - Erfassung der Auslastung von Data-Lake-Komponenten
   - Erfassung von Datenzugriffsmustern
2. Technologien für Data-Lake-Caching
   - Weiterentwicklung und Konsolidierung von Daten-Cache - Technologien
   - effiziente Einbindung von dynamischen Datencaches in den Data-Lake und an CPU-Ressourcen
   - Einsatz von parallelen ad-hoc Filesystemen als Caches in HPC-Systemen
3. Technologien für Data-Lake-Daten- und Workflow-Management
   - Replikations- und Platzierungsmechanismen
   - bedarfsgetriebene Datenmanagement-Mechanismen
   - effizienter Datenzugriff und Anpassung an Workload-Management-Systeme
4. Data-Lake-Prototypen, Technologien für QoS und effiziente Anbindung
   - Aufbau von Data-Lake-Prototypen
   - effiziente Anbindung von Nutzern, Zentren und Datenquellen
   - Quality of Service

Project FIDIUM (3)

- TA III: Adjustments, Tests and Optimisation of the production and analysis environment

1. Integration, Tests, Optimierung und Deployment der entwickelten Dienste
   - Integration der verschiedenen Komponenten: Workflowmanagement, Caching, Accounting, Ressourcenmanagement und Überwachungsysteme
   - Funktionale Tests auf ausgewählten Zentren
   - Integration in die Produktionsumgebung der beteiligten Experimente
   - Am Ende, nach erfolgreichen Tests, “Deployment” der Gesamtlösung an den verfügbaren WLCG-Tier-Zentren, HPC-Zentren und Cloud Anbietern.
2. Spezifische Anpassung der Dienste an komplexe Workflows und Nutzung spezieller Technologien für die Analyse wissenschaftlicher Daten
   - Optimierung für spezielle Workflows mit hoher IO Last, Speicherbedarf, GPU Nutzung, u.a.
   - Optimierung für schnelle parallele Analyse großer Datenmengen mit modernen vektor-basierten Analysealgorithmen
3. Support
   - Einrichtung eines standortübergreifenden Support-Teams, das die Zentren bei Installation und Betrieb unterstützt

Final remarks

- Very good results from IDT-UM, BUT we are not there yet!
- Very successful collaboration across the experiments and communities
- Need to keep-up the momentum with the existing know-how

Lets hope, that FIDIUM will be funded!