

ERUM DATA IDT COLLABORATION MEETING 2019

STATUS AND PLANS – AREA A

Kilian Schwarz & Manuel Giffels

TOPICS OF AREA A

Development of technologies to enable utilisation of heterogeneous computing resources

<p>AP 1) Werkzeuge zur Einbindung</p> <ul style="list-style-type: none">• Scheduling von Cloud - Jobs• Container - Technologien• Checkpointing• Zugang zu Experiment-Datenbanken	<p>AP 2) Effiziente Nutzung</p> <ul style="list-style-type: none">• Steigerung der Effizienz von datenintensiven Anwendungen auf heterogenen Ressourcen mittels „on the fly“ Datencaches
<p>AP 3) Workflow Steuerung</p> <ul style="list-style-type: none">• Identifikation und Steuerung• In - Pilot Job Monitoring• Accounting• Optimierung durch data - mining	

CONTRIBUTIONS TO AREA A

Aachen (friend):

- ▶ T2_DE_RWTH can be used for dynamic resource management tests [A1]

Bonn (associated partner):

- ▶ Dynamic resource management of T3 resources using (COBaID/TARDIS) [A1]
- ▶ Interests to join: developments and evaluation of coordinated distributed caching solutions [A2]

DESY (associated partner):

- ▶ Smart caching in WLCG data federation using dCache [A2]

Frankfurt/GSI:

- ▶ Singularity Containers to include HPC resources into Grid computing (e.g. ALICE T2@GSI) [A1]
- ▶ Developments of XRootD based coordinated distributed caching solutions [A2]

Freiburg:

- ▶ Dynamic resource management developments (COBaID/TARDIS) [A1]
- ▶ Contribute to the coordinated distributed caching solutions [A2]
- ▶ Development of monitoring, accounting tools and benchmarks [A3]

Karlsruhe:

- ▶ Development of the opportunistic resource manager COBaID/TARDIS [A1]
- ▶ Workflow management in heterogenous environments [A1]
- ▶ Development of a coordinated distributed caching solution using XRootD [A2]

München:

- ▶ Job log files analysis by using ML (anomaly detection) [A3]
- ▶ Development of XRootD based disk caching (XCache) [A2]

Wuppertal:

- ▶ Containerization of user jobs and services (VOMS, DB access, monitoring) focussing on lightweight solutions [A1, A3]

NEWS

FUNDING

Standort	A1	A2	A3	FTE
Bonn	X			-
DESY		X		-
Frankfurt FIAS		X		0,75
Freiburg	X	X	X	1
GSI	X	X		-
KIT	X	X		1
München		X	X	0,5
Wuppertal	X		X	0,5

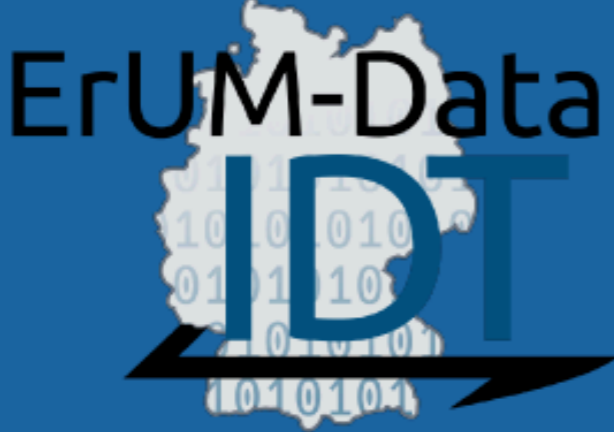
Position will be filled in November!

Two new proposals:

- ▶ Compute Site in a Box (U Bonn)
- ▶ Ad-hoc-Dateisysteme für PANDA an FAIR (U Mainz)

TALKS LATER ON!

ErUM Data Cloud Workshop



27 June 2019
Building 30.23
Europe/Berlin timezone

- Overview
- Timetable
- Contribution List
- My Conference
 - My Contributions
- Registration
- Participant List

1-day workshop on COBaID/TARDIS, our solution to dynamically deploy, integrate and manage opportunistic resources like Clouds, HPC centres, T3s and others. Invited are all parties interested in using COBaID/TARDIS.

The workshop aims at providing an overview of the abilities of COBaID/TARDIS, having technical discussions what is needed to use COBaID/TARDIS on your site and giving assistance to configure your local instances

 **Starts** 27 Jun 2019, 10:30
Ends 27 Jun 2019, 17:00
Europe/Berlin


 [Manuel Giffels](#)

 **Building 30.23**
9-1
Wolfgang-Gaede-Straße 1, 76131 Karlsruhe

 [KIT Map](#)



ERUM DATA IDT CLOUD WORKSHOP


 27 June 2019 Building 30.23 Europe/Berlin time zone	Introduction into COBalDD 9-1, Building 30.23	Max Fischer	10:30 - 11:00
	Dynamic Transparent Integration and Management of Opportunistic Resources with COBalD/TARDIS 9-1, Building 30.23	Manuel Giffels et al.	
	Lunch Restaurant Continent (Indian)		11:30 - 13:00
	COBalD & Tardis in Bonn 9-1, Building 30.23	Oliver Freyermuth	13:00 - 13:20
	Tutorials: Hands-on COBalD/TARDIS 9-1, Building 30.23		14:30 - 16:00

- Overview
- Timetable
- Contributions
- My Conference
- My Contributions
- Registration
- Participant

page interested in

technical configure your

rsruhe



ERUM DATA IDT CLOUD WORKSHOP

Participants:
University Bonn &
Freiburg

Introduction into COBalD Max Fischer 
10:30 - 11:00

Management of Opportunistic Resources with Manuel Giffels et al. 


Restaurant Continent (Indian) 11:30 - 13:00

COBalD & Tardis in Bonn Oliver Freyermuth 
9-1, Building 30.23 13:00 - 13:20

Tutorials: Hands-on COBalD/TARDIS 
9-1, Building 30.23 14:30 - 16:00

27 June 2019
Building 30.23
Europe/Berlin time zone

- Overview
- Timetable
- Contributions
- My Conference
- My Contributions
- Registration
- Participant Information

page interested in
technical configure your
disruptive 

ERUM DATA IDT CLOUD WORKSHOP

Participants:
University Bonn &
Freiburg

Summarized Feedback:
The cloud workshop was
extremely helpful!

Introduction into COBalD Max Fischer

ement of Opportunistic Resourc

27 June 2019
Building 30.23
Europe/Berlin time zone

Restaurant Continent (Indian) 11:30 - 13:00

COBalD & Tardis in Bonn Oliver Freyermuth
9-1, Building 30.23 13:00 - 13:20

Overview
Timetable
Contributions
My Conference
My Contributions
Registration
Participant Information

Tutorials: Hands-on COBalD/TARDIS

9-1, Building 30.23 14:30 - 16:00

Interested in
Technical
Configure your
lsruhe

Participants:
University Bonn &
Freiburg

Summarized Feedback:
The cloud workshop was
extremely helpful!

27 June 2019
Building 30.23
Europe/Berlin time zone

Restaurant Continent (Indian)

11:30 - 13:00

COBalD & Tardis in Bonn

9-1, Building 30.23

Oliver Freyermuth

13:00 - 13:20

Overview

Timetable

Contributions

My Conference

My Contributions

Registration

Participant List

Are there interested parties to have another one soon?

Tutorials: Hands-on COBalD

9-1, Building 30.23

14:30 - 16:00

Are you interested in

Technical
Configure your

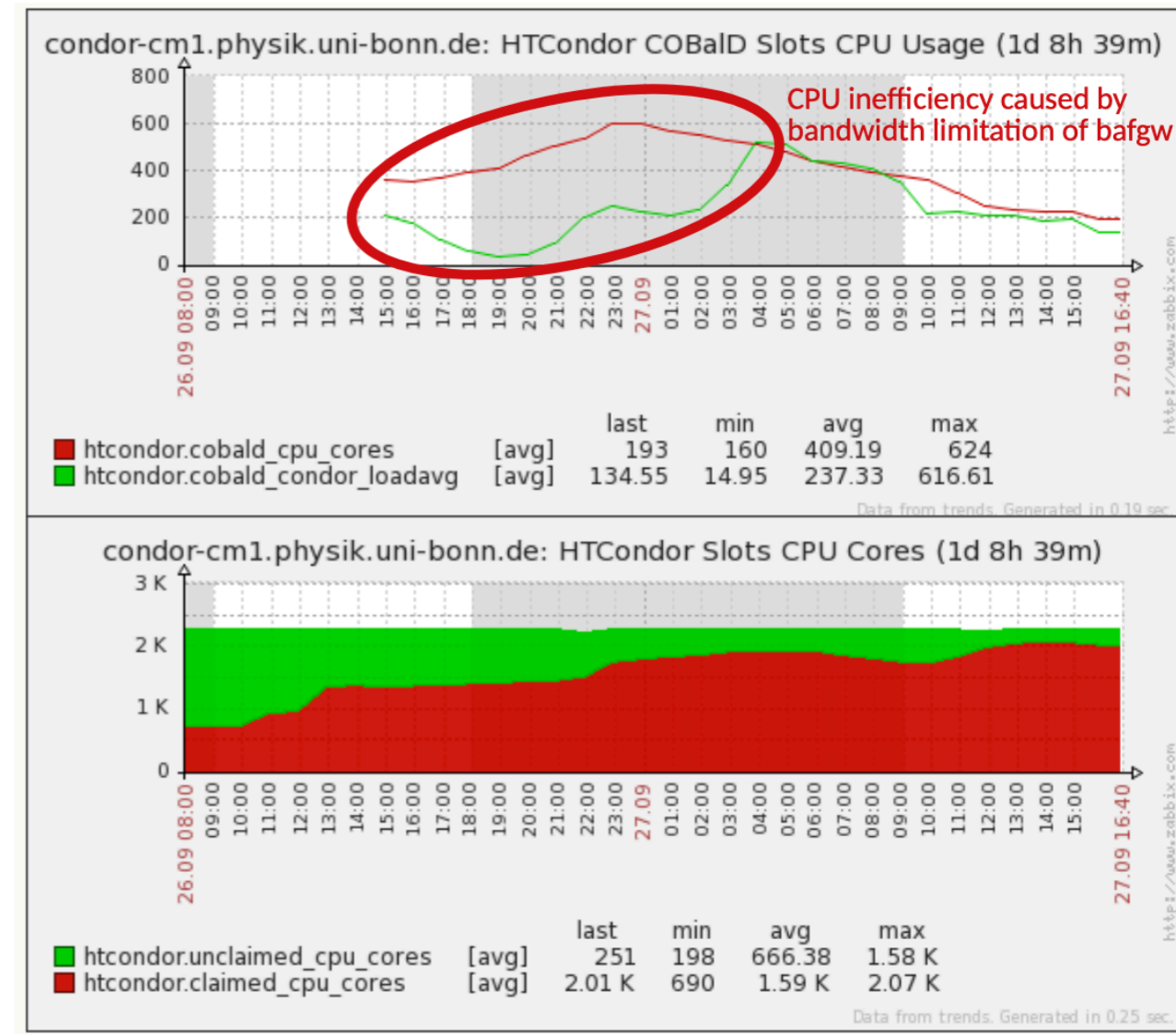
Disruptive



STATUS AND PLANS

PILOT PROJECT @ UNIVERSITY OF BONN

- ▶ COBaID/TARDIS successfully deployed @ U Bonn
- ▶ First ATLAS production jobs are running in Bonn (Tier 3)
- ▶ Jobs are fed by an ARC-CE located at the GridKa Tier 1
- ▶ Completely transparent to the ATLAS experiment



PLANS FROM FRANKFURT FIAS/GSI

A1:

tools for including heterogeneous resources into scientific computing workflows

A prototype for running Grid jobs in Singularity containers in HPC centres has been built.

This prototype shall be intensively tested and improved in context with the ALICE T2 centre at GSI. After that it should be generalised to the use by other experiments and also at other HPC centres.

mile stones:

- M1 (06/19):
 - adapting the prototype to working with the Slurm scheduler at GSI ✓
 - creating a management system for images
- M2 (06/20):
 - demonstrator of the prototype with another batch system (e.g. HTCondor) or at another HPC centre (e.g. GoetheHLR)
- M3 (06/21):
 - generalisation of the developed prototype to the use by other experiments (e.g. CMS)

Position will be filled in November!

PLANS FROM FRANKFURT FIAS/GSI

A2:
efficient usage of heterogeneous resources

A prototype of an XRootD based dynamic data cache for heterogeneous resources is being developed. A job which needs data from an external data server requests these data via an XRootD forward proxy. During this process the data are being cached on a local data cache. In case further jobs would require the same data set XRootD would recognise that the data are already locally available and would redirect these jobs via an XRootD Plug In to the local file system.

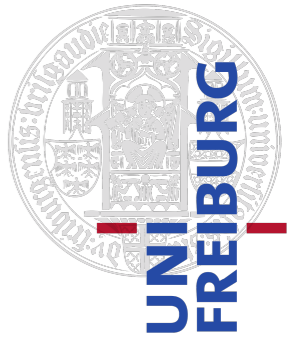
mile stones:

- M1 (12/18):
 - developing an XRootD based „disk cache on the fly“ following the concept described above
- M2 (12/19):
 - first prototype of the „disk cache on the fly“ at GoetheHLR



Position will be filled in November!

STATUS AND PLANNED CONTRIBUTIONS FROM FREIBURG



Plans unchanged. Position only filled 1st of September 2019.

WP1: Tools for Integration of heterogeneous resources in scientific computing

- development of scheduling of cloud-jobs with virtual machines
 - adaption of meta scheduler for “unified queue”
(meta scheduler COBaID/TARDIS developed in G.Quast’ group)
to SLURM/ MOAB frontend/backend combination
 - extension of meta scheduler to other backend-systems and Cloud-APIs
 - extension of meta scheduler to provide different configurations
of VMs depending on job requirements and available resources
next step: implement COBaID/TARDIS in FR (ATLAS-Tier2, NEMO, BWCLOUD)

- development of container solutions
 - inclusion of container solution in workflow e.g. based on “Singularity”
 - develop interface btw. front- and backend systems respecting limited rights in container
 - develop monitoring for/at interface between batch systems



STATUS AND PLANNED CONTRIBUTIONS FROM FREIBURG

WP2: Efficient use of heterogeneous resources

- contribute to development of fast „on the fly“ data caches
 - adapt prototype based on XRootD (developed by groups of G. Quast and K. Schwarz)
for ATLAS use case (next step)
 - implement Dynafed ansatz and prototype and check scalability
 - develop benchmarks and compare different approachesfirst studies by Benjamin → talk in parallel session

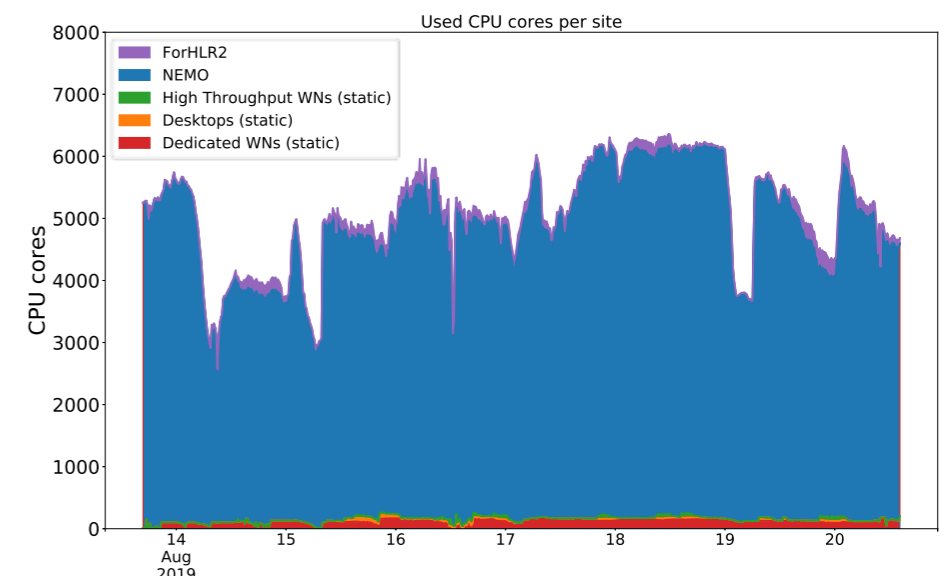
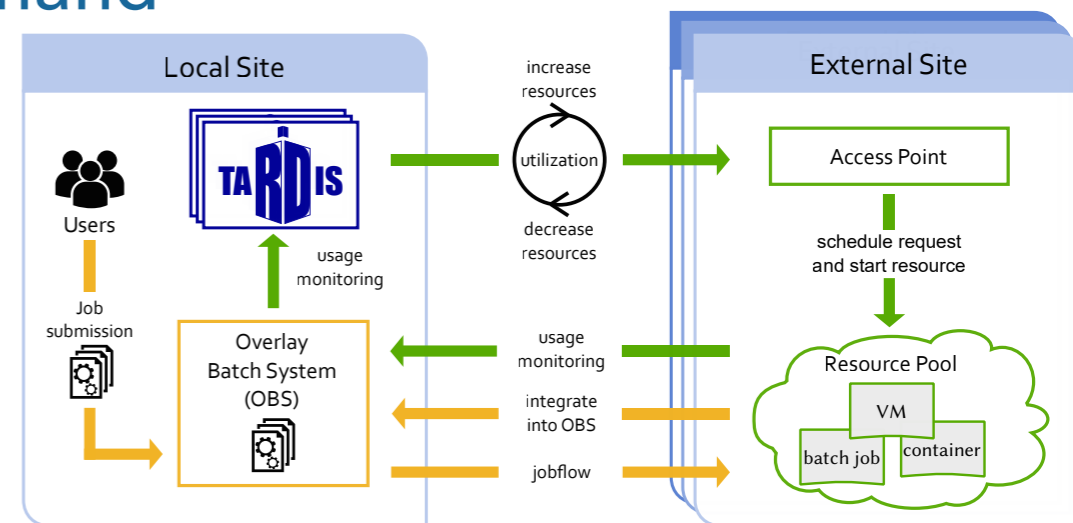
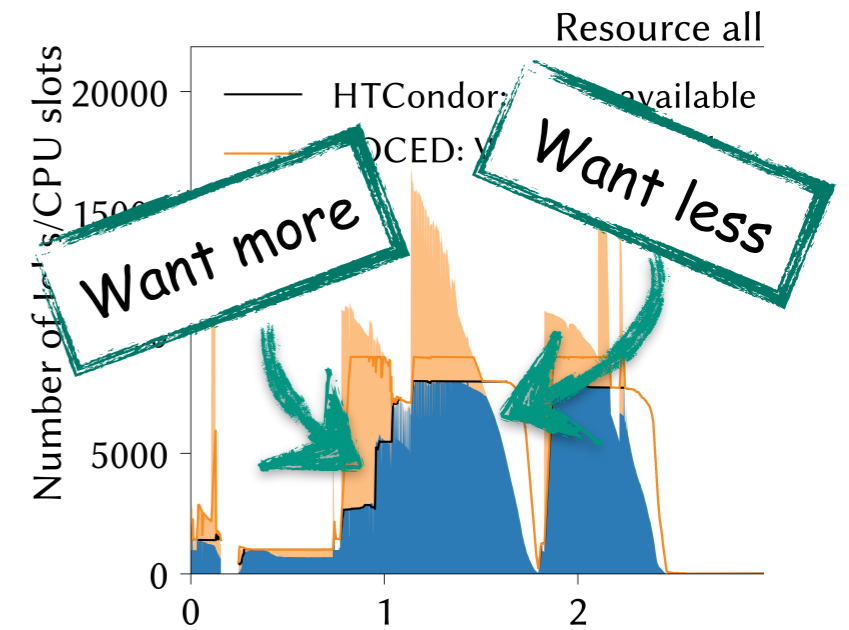
WP3: Identification & steering of workflows on heterogeneous resources

- development of monitoring and accounting tools for different configurations, available number and kind of resources, kind of jobs (1st for ATLAS use case)
 - development of standardized interfaces for automatised tests
(experiment overarching, for monitoring and benchmarking)
 - development and test of benchmarks (for I/O load, cpu load, simulation, user job, ..)
 - evaluate performance, investigate long-term efficiency and reliability
 - analysis and storage of monitoring data via Elastic Search and Kibanafirst studies with various benchmarks on a few heterogeneous resources
by Benoit → talk in parallel session

STATUS AND PLANS FROM KARLSRUHE (A1)

Dynamic on-demand provisioning of resources:

- ▶ COBaID/TARDIS resource manager developed at KIT
- ▶ Enables transparent and dynamic on-demand provisioning of opportunistic resources
- ▶ Enables backfilling of HPC/T3 resources
- ▶ Used at KIT-ETP, GridKa and Uni Bonn. Uni Freiburg is following soon!
- ▶ Ready to add more parties from (B2)
- ▶ Production ready software at scale!



COBaID



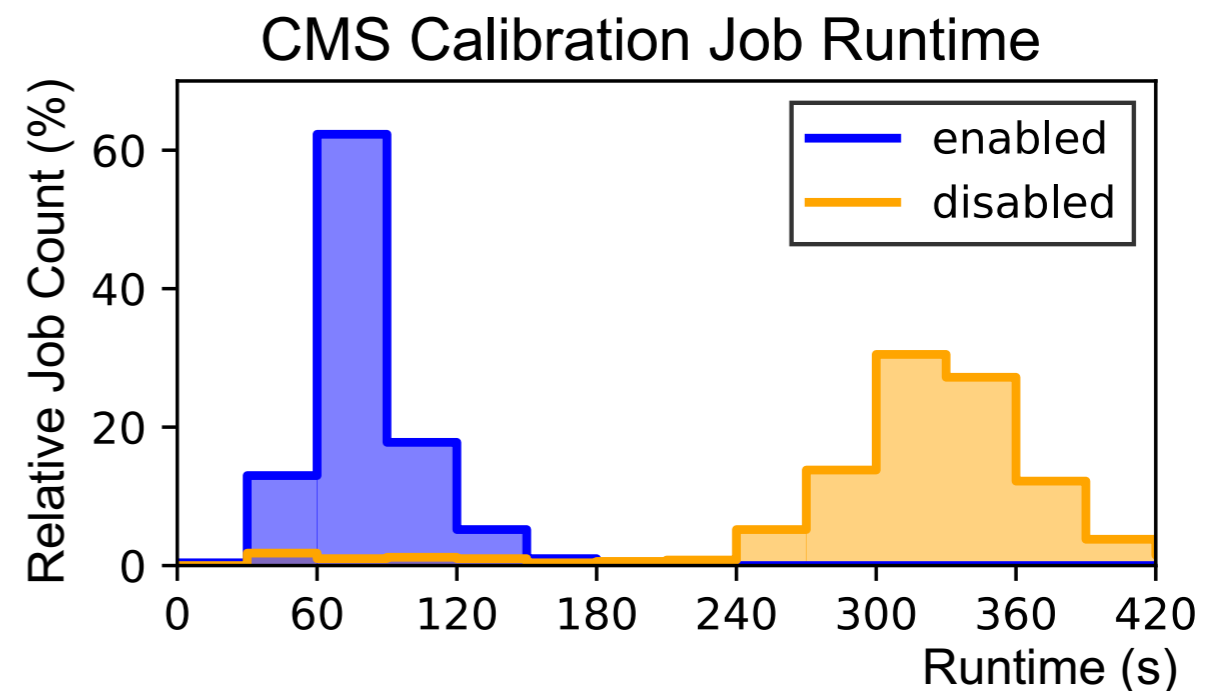
EXPERIMENT INDEPENDENT!

STATUS AND PLANS FROM KARLSRUHE (A2)

Distributed coordinated caching:

- ▶ NaviX coordination service in development at KIT
- ▶ First prototype is available
- ▶ Evaluation ongoing based on data collected so far
- ▶ Working towards a **scalable caching solution** based upon HTCondor and XRootD
- ▶ Stay tuned!

EXPERIMENT INDEPENDENT!

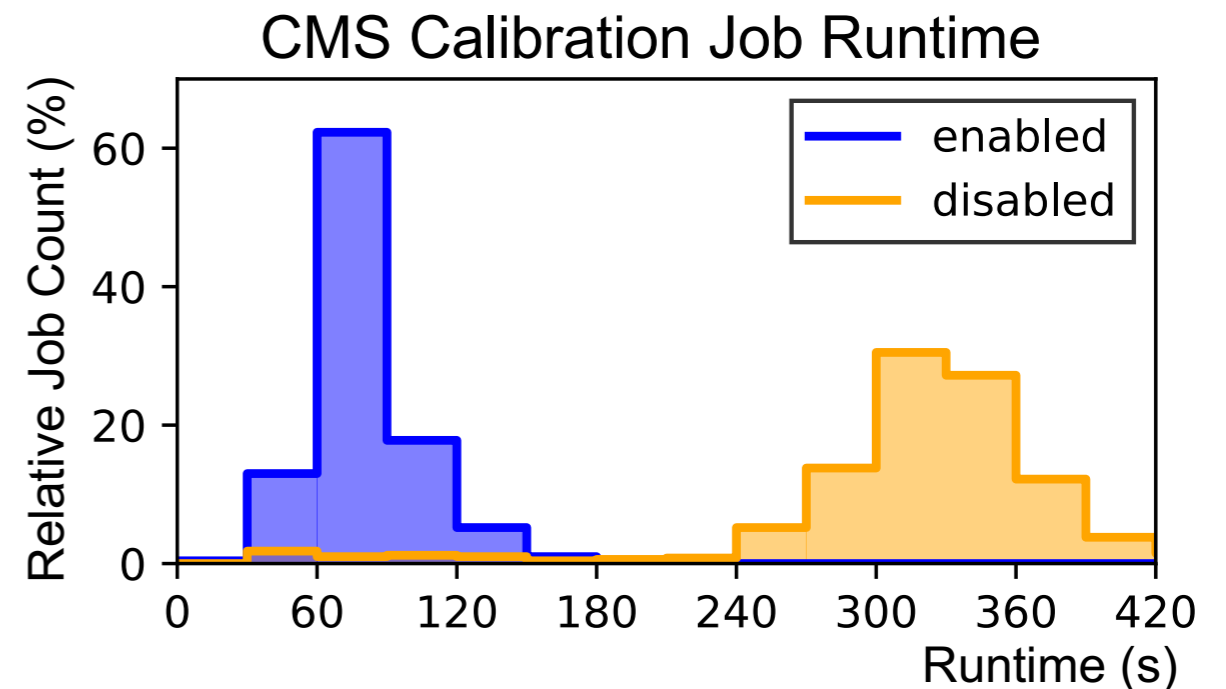
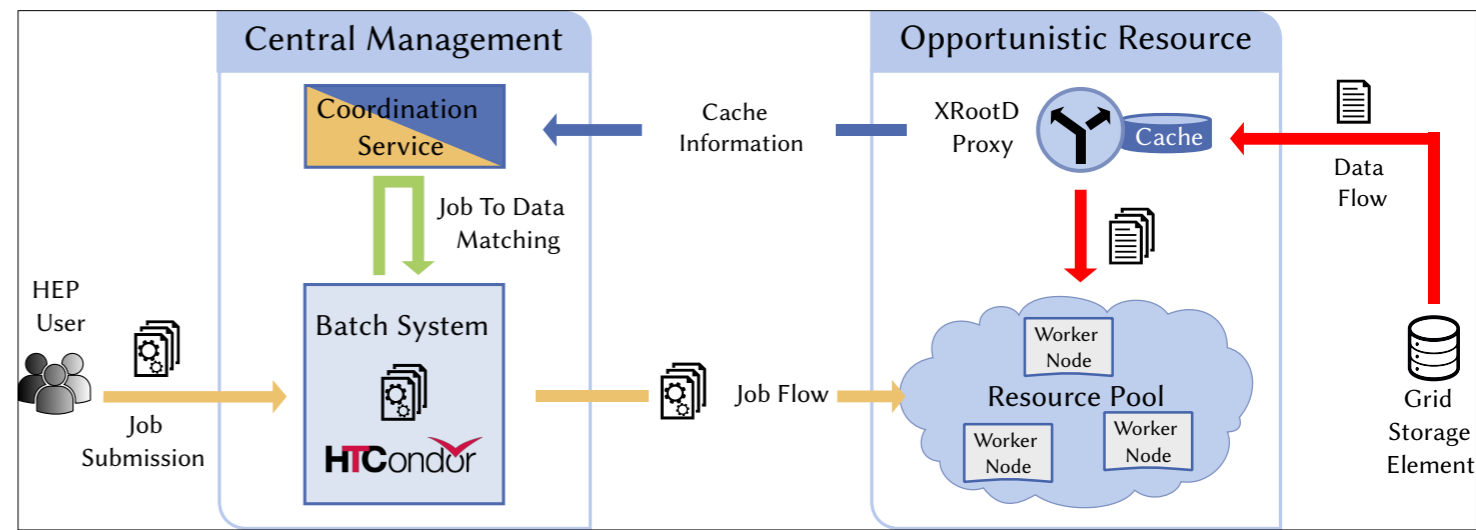


STATUS AND PLANS FROM KARLSRUHE (A2)

Distributed coordinated caching:

- ▶ NaviX coordination service in development at KIT
- ▶ First prototype is available
- ▶ Evaluation ongoing based on data collected so far
- ▶ Working towards a **scalable caching solution** based upon HTCondor and XRootD
- ▶ Stay tuned!

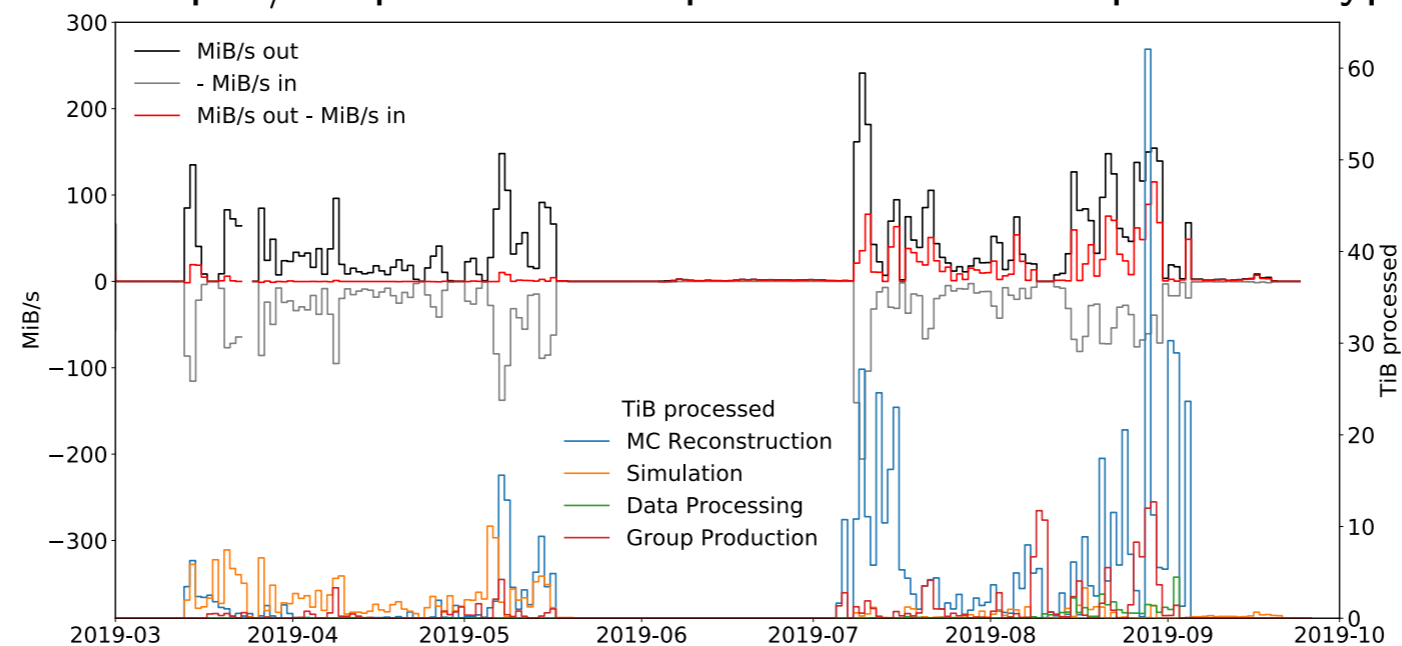
EXPERIMENT INDEPENDENT!



STATUS AND PLANS FROM MUNICH (A2)

Xcache at LMU Munich

Cache input/output rate and processed volume per Job Type:

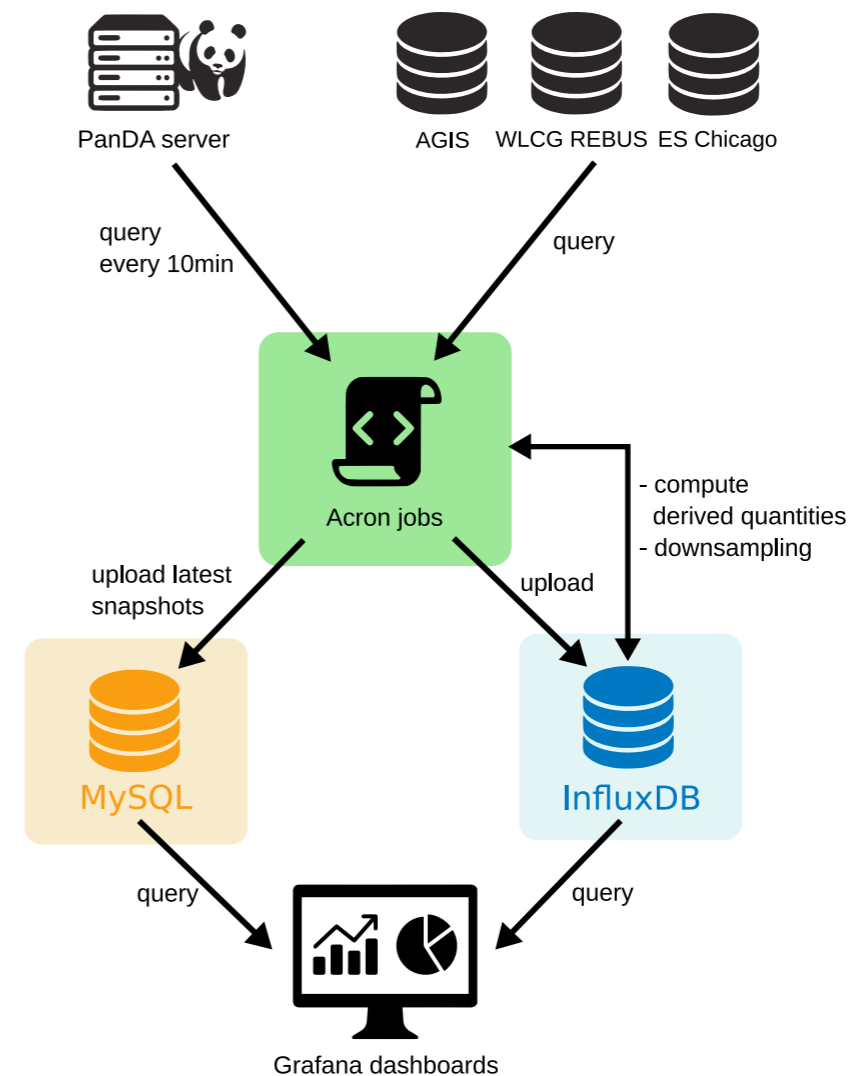


- Xrootd based proxy cache
- After initial studies, tested in ATLAS production environment
- Identified some issues/bugs under heavy load
→ now mostly fixed or mitigated
- Next steps:
 - Test multi-disk mode instead of Raid system
 - Continue investigating use for analysis jobs
 - Potential use case for columnar data analysis (Pandas/Dask)

STATUS AND PLANS FROM MUNICH (A3)

Queue-based Job monitoring for ATLAS

- Problem: Try to find cause for **variations in number of running slots** per queue
- Developed **low-latency and lightweight** system based on existing information
→ PanDa, Grid sites, ATLAS
- **Suspicious sites dashboard** for fast detection and access to relevant monitoring plots (Grafana at CERN)
- Potential to use method for other Experiments
→ checking for Belle II



STATUS AND ACTIVITIES IN WUPPERTAL

- Involved people
 - Marcello Vogel
 - Christian Zeitnitz
- Currently addressed projects
 - Containerization of jobs (simulation, reconstruction, analysis)
 - Database extraction for network-less jobs
 - Determine from metadata the required databases and entries
 - Extract data to SQLITE files
 - Pack job and DB files into container
 - Current status
 - Use case ATLAS
 - Successful for simple case of simulation
 - Reconstruction nearly working, but obstacles because of missing meta data and hardcoded paths and DB access
 - Next step
 - Generalization to other experiments
 - Define interfaces

SUMMARY

WP1 (Tools to integrate opportunistic resources):

- ▶ Have made very good progress!
- ▶ Software ready to add more interested parties from (B2)

WP2 (Efficient Utilization of Resources):

- ▶ Caching prototypes have been set-up using different technologies
- ▶ Evaluations are ongoing
- ▶ Stay tuned!

WP3 (Steering of Workflows/Monitoring):

- ▶ Initial work has been done
- ▶ Need to focus more on experiment independent tools

QUESTIONS?