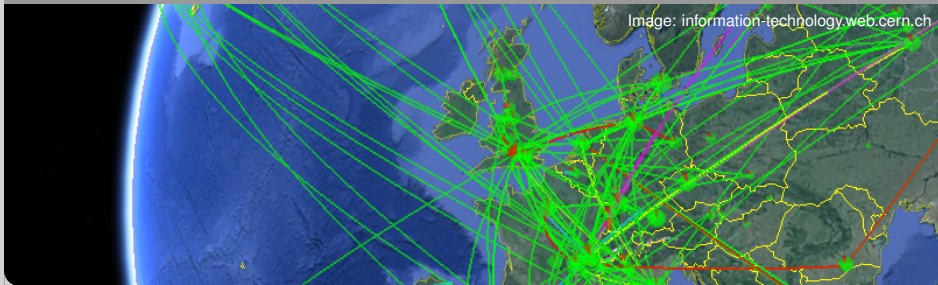


Opportunistic Computing and Lightweight Grid Operations

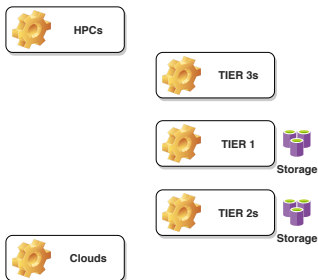
R. Florian von Cube, René Caspart, Tabea Feßenbecker, Max Fischer, Manuel Giffels,
Christoph Heidecker, Maximilian Horzela, Eileen Kühn, Günter Quast, Matthias J. Schnepf
ErUM Data Collaboration Meeting, September 22, 2020

INSTITUTE OF EXPERIMENTAL PARTICLE PHYSICS (ETP), KIT

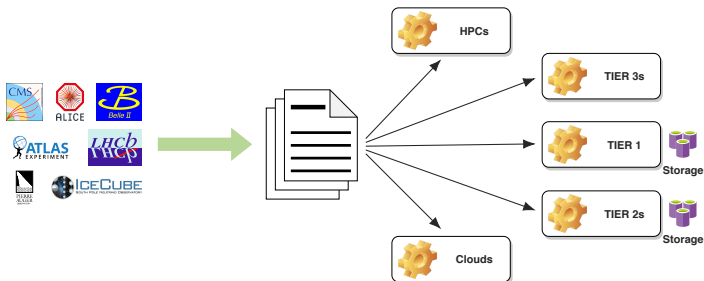




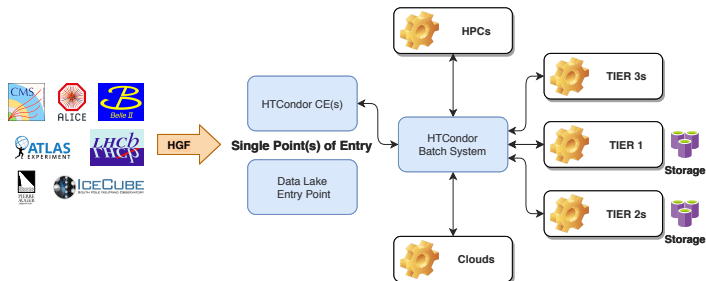
- Multiple experiments



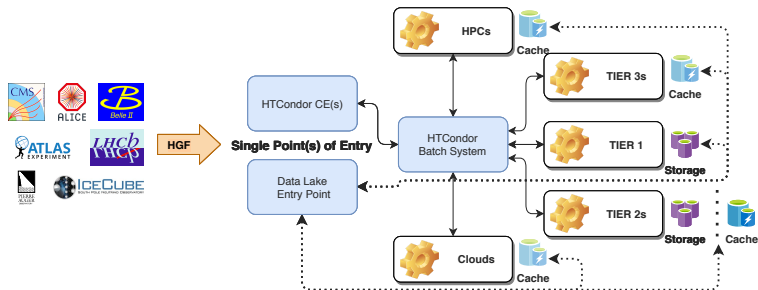
- Multiple experiments on heterogeneous resources



- Multiple experiments on heterogeneous resources



- Multiple experiments on heterogeneous resources
- Opportunistic, transparent **resource integration** and **lightweight site operation**

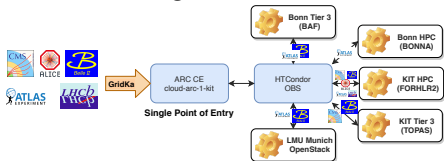


- Multiple experiments on heterogeneous resources
- Opportunistic, transparent **resource integration** and **lightweight site operation**
- **Data locality** for efficient usage

Areas of Development at KIT

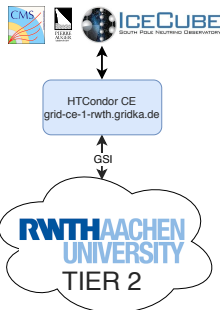
Resource integration

A1



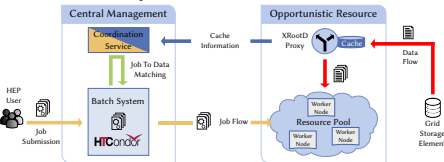
Lightweight site operation

A1/B2



Data locality

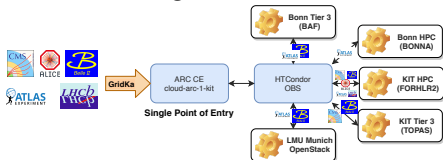
A2



Areas of Development at KIT

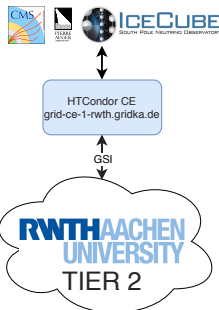
Resource integration

A1



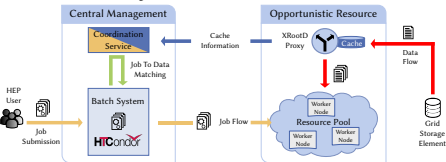
Lightweight site operation

A1/B2



Data locality

A2



Resource Integration

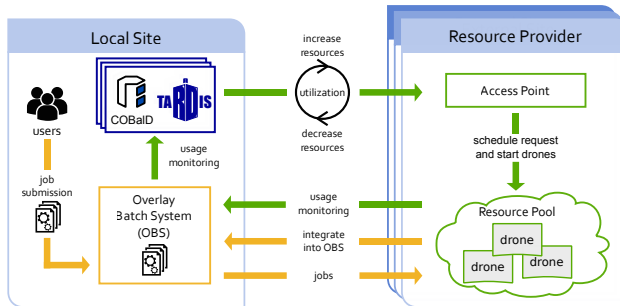
Working with local groups

- knowledge of and access to additional, local resources
 - shares at university clusters
 - collaborations with HPC centers/Cloud providers
- ⇒ Integration of such resources allows to extend local infrastructures
- acquired resources can be shared
 - local group keeps under control *who uses which resources*
 - ⇒ entire community benefits
- overall resource utilization improves
- ⇒ Focus efforts on enabling local collaborations instead of centralized approaches (e.g. experiment-wide negotiations)

Resource Integration

The Concept

- allocate opportunistic resources: Cloud, HPC, university clusters
- provide opportunistic resources as extension to community-specific computing resources
- integrate resources transparently and based on current demand

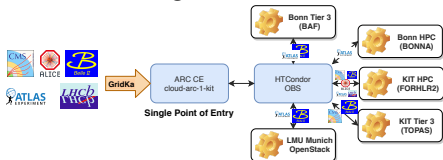


- **COBaID & TARDIS:** developed at KIT

Areas of Development at KIT

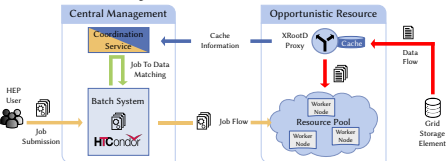
Resource integration

A1



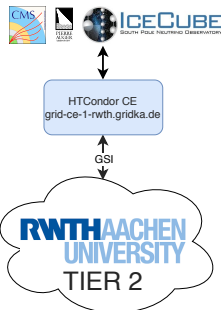
Data locality

A2



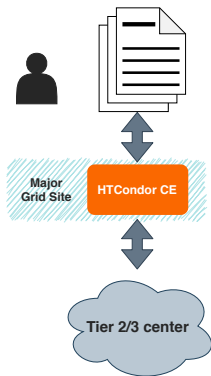
Lightweight site operation

A1/B2

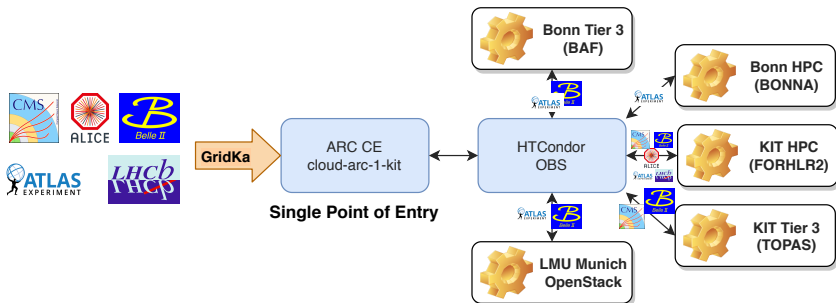


Lightweight Site Operation

- simplifies contribution of computing power for smaller sites
- CEs are operated remotely on major grid site, e.g. GridKa
- reduce technical effort, need for personnel
- since summer 2020 in production for Aachen



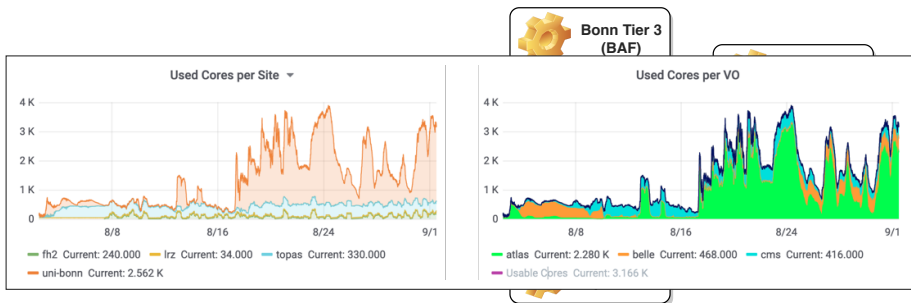
Achievements



- built prototype of federated infrastructure with KIT, U Bonn, LMU, and U Freiburg
- remote CE for RWTH at GridKa in production
- serves multiple VOs / Experiments

→ **wide-range and beneficial collaborations across communities**

Achievements

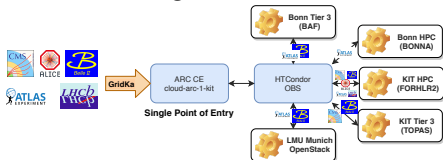


- built prototype of federated infrastructure with KIT, U Bonn, LMU, and U Freiburg
 - remote CE for RWTH at GridKa in production
 - serves multiple VOs / Experiments
- **wide-range and beneficial collaborations across communities**

Areas of Development at KIT

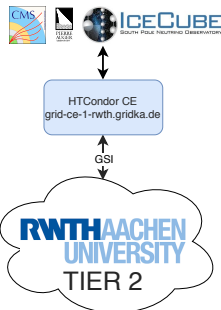
Resource integration

A1



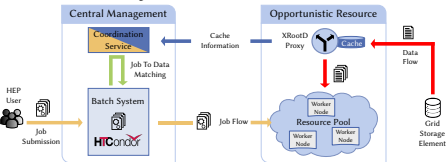
Lightweight site operation

A1/B2



Data locality

A2



Data locality

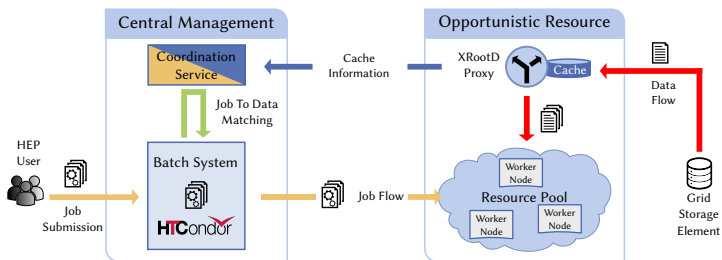
- data transfers in federated systems limited by network throughput

Strategy:

- benefit from existing proximity of CPUs and data storage
- introduce additional caches holding data replicas

⇒ **distributed coordinated caching**

Prototype realization **NaviX** at ETP

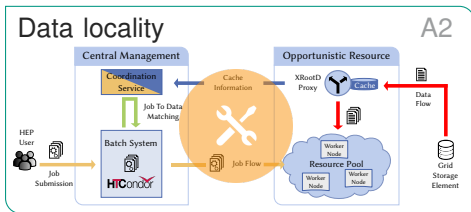
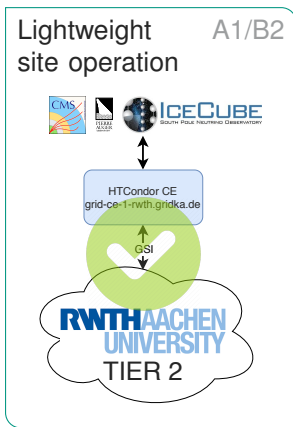
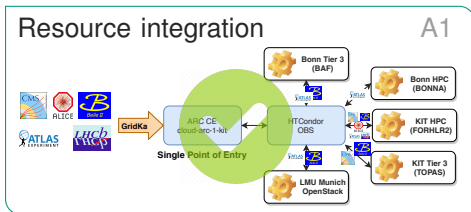


Data locality

Roadmap @ KIT

- goal: increase data locality transparently without user intervention
- ⇒ sophisticated cache operation and workflow management required

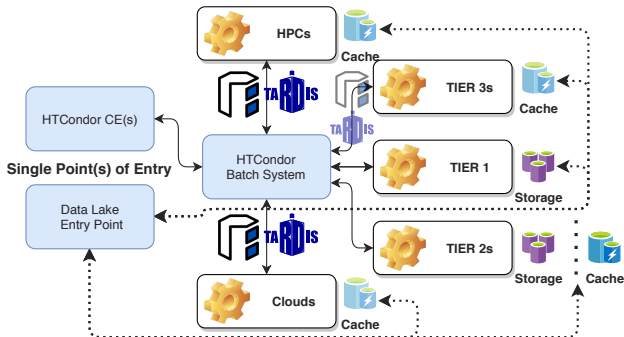
- determine robust caching and coordination strategies
- very complex problem
 - resource infrastructure
 - network bandwidth
 - workload composition
- ⇒ study behavior with extensive **simulations**
- suitable simulator in development at ETP
- ⇒ refine existing prototype to scalable caching solution



- A lot of progress was made in the scope of ErUM Data!
⇒ always looking for new collaborators. Contact us!

Vision

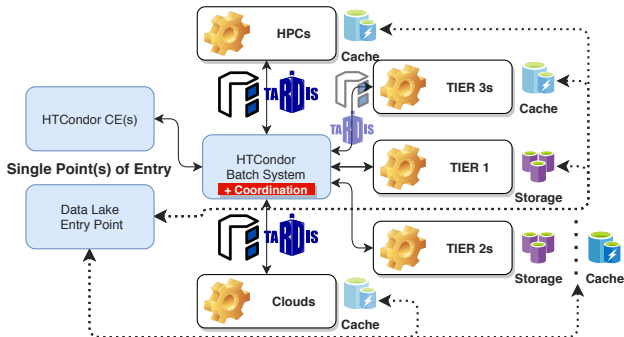
Federated Infrastructures



- extensive and interdisciplinary collaborations
- scalable and intuitive infrastructure

Vision

Workflow Management



- create and use sophisticated workflow management
- efficient usage of existing resources
- be considerate of different resource types