Karlsruhe Institute of Technology

# HPC infrastructure for Particle Physics: The Experience of the Karlsruhe groups

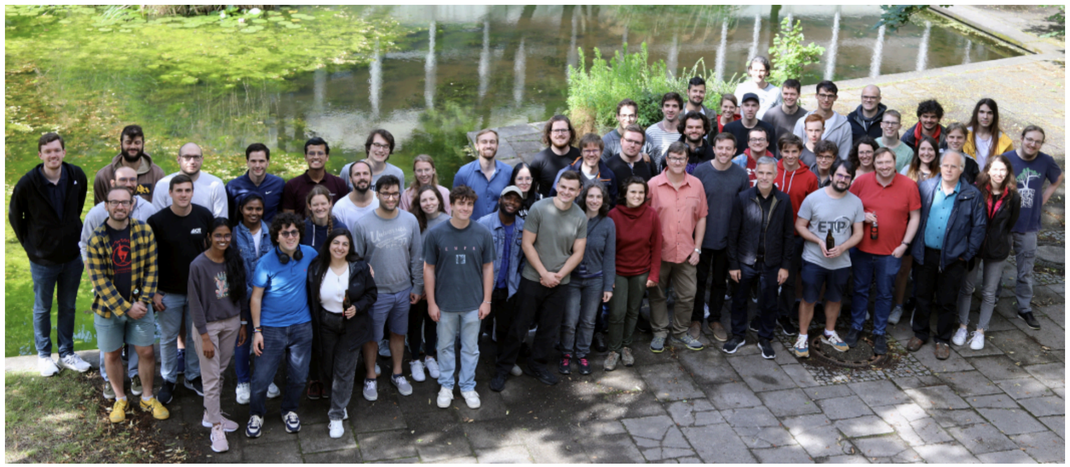
## BWHPC Symposium - Freiburg

**Sebastian Brommer**, Robin Hofsaess, Manuel Giffels, Günter Quast, Matthias Schnepf - 26.09.24

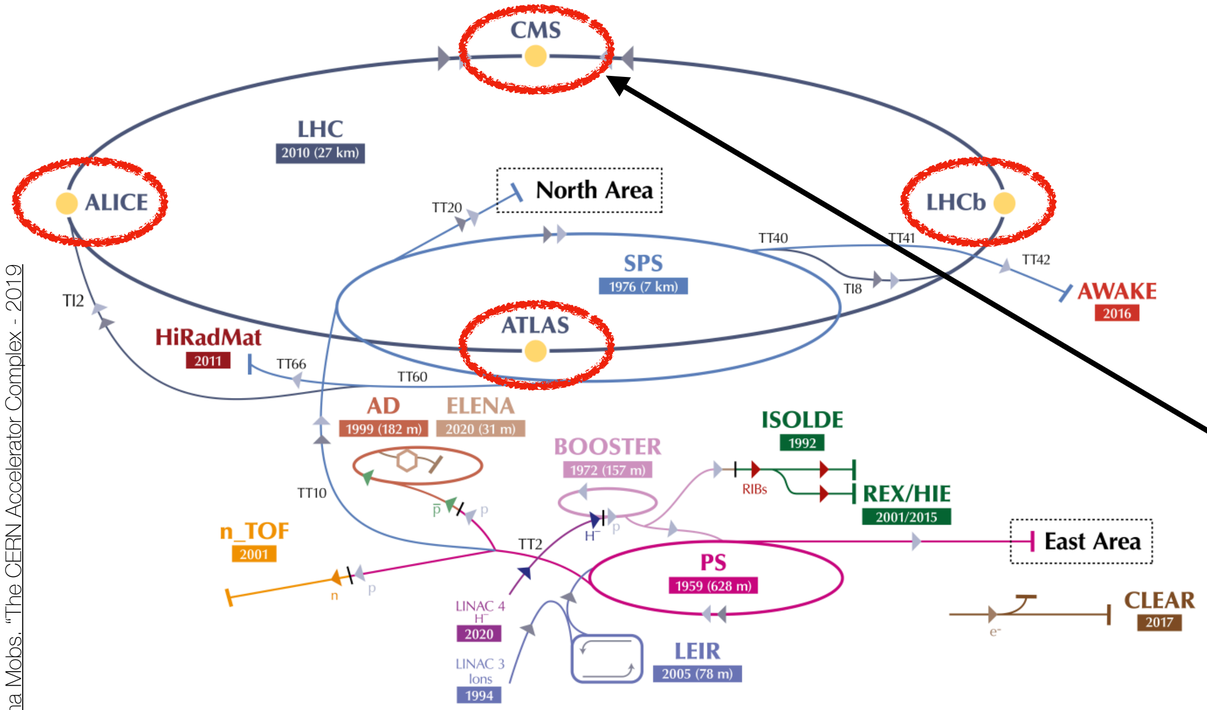
# Institute for Experimental Particle Physics

## At KIT

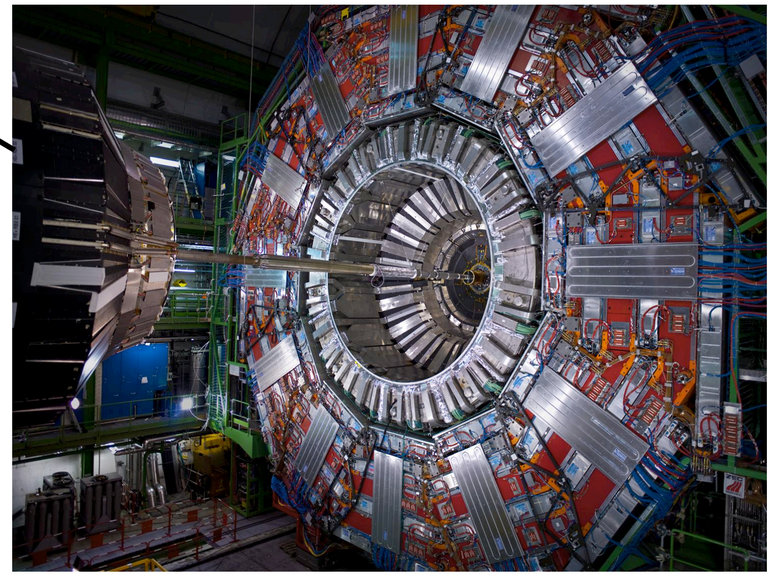
- Particle Physics Institute with ~ 80 members (more than half students)
- Contributions to multiple fields in particle physics including
  - CMS
  - Belle II
  - FCC
  - Computing
  - DM searches

# The Large Hadron Collider at CERN

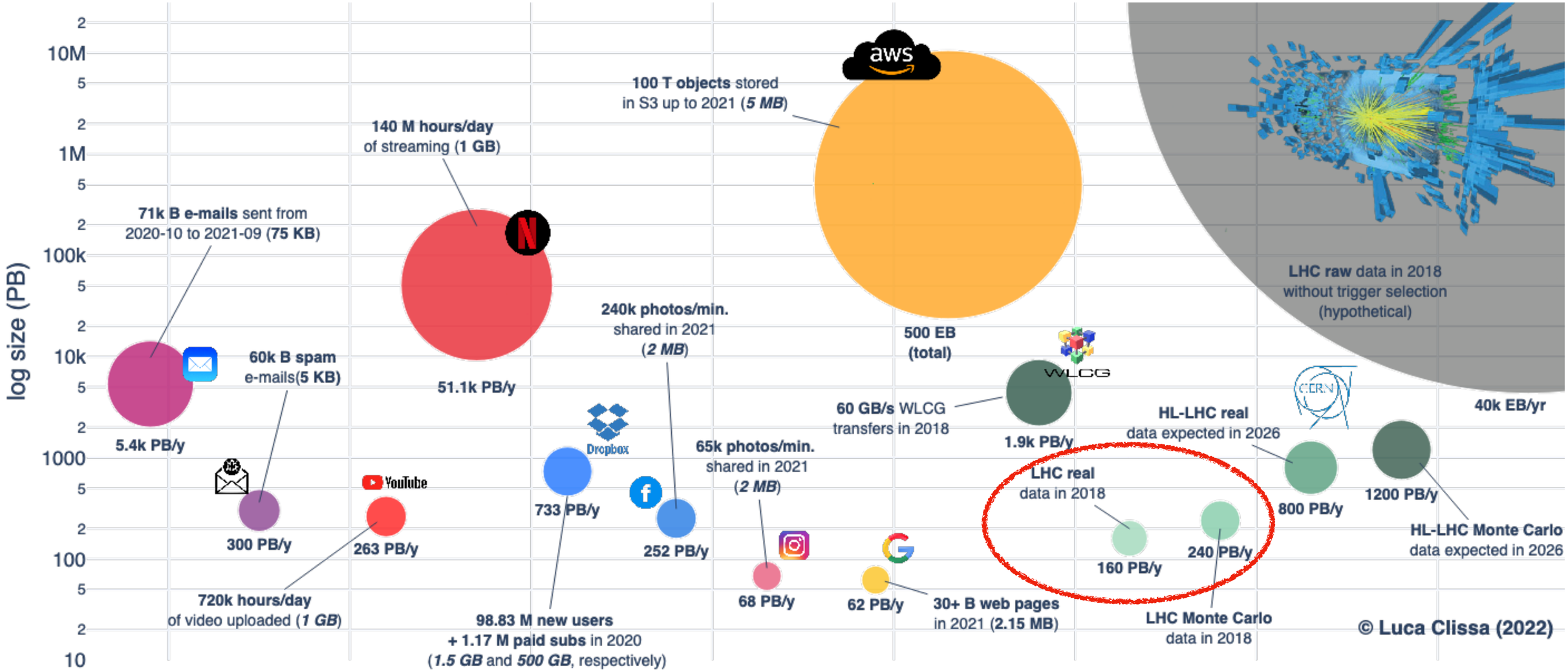



- 27 km proton-proton collider
- collisions every 25 ns with (ideally) ~ 9 months of 24/7 data taking
- 4 large experiments (e.g. for CMS one recorded event is about 1 Mb)



› Source: Esma Mobs, "The CERN Accelerator Complex - 2019

# Computation in Particle Physics




› Source: Clissa - Survey of Big Data sizes in 2021

# Computation in Particle Physics

- The **LHC produces a lot of data** that has to be processed, stored and analysed
- We need **at least the same amount of simulated collisions** to be generated and stored

# Computation in Particle Physics

- The **LHC produces a lot of da** be processed, stored and analysed
- We need **at least the same amoun sions** to be generated and stored

Most done centrally by the collaboration using the WLCG and computing centers like GridKa T1

# Computation in Particle Physics

- The **LHC produces a lot of da** be processed, stored and analysed
- We need **at least the same amoun sions** to be generated and stored

Most done centrally by the collaboration using the WLCG and computing centers like GridKa T1

- Process data and simulation further for **your** specific analysis
  - Filtering of datasets
  - Calculation of high-level variables
  - Derivation and application of corrections
  - ML for classification, fitting, limit calculations, unfolding …
  - […]

# Computation in Particle Physics

- The **LHC produces a lot of da** be processed, stored and analysed
- We need **at least the same amoun sions** to be generated and stored

Most done centrally by the collaboration using the WLCG and computing centers like GridKa T1

- Process data and simulation further for **your** specific analysis
  - Filtering of datasets
  - Calculation of high-level variables
  - Derivation and application of corrections
  - ML for classification, fitting, limit calculatio
  - […]

Possible via collaboration entrypoints but resources are shared between all members

# Computation in Particle Physics

- The **LHC produces a lot of da** be processed, stored and analysed
- We need **at least the same amount** **sions** to be generated and stored

Most done centrally by the collaboration using the WLCG and computing centers like GridKa T1

- Process data and simulation further for **your** specific analysis
  - Filtering of datasets
  - Calculation of high-level variables
  - Derivation and application of corrections
  - ML for classification, fitting, limit calculatio
  - […]

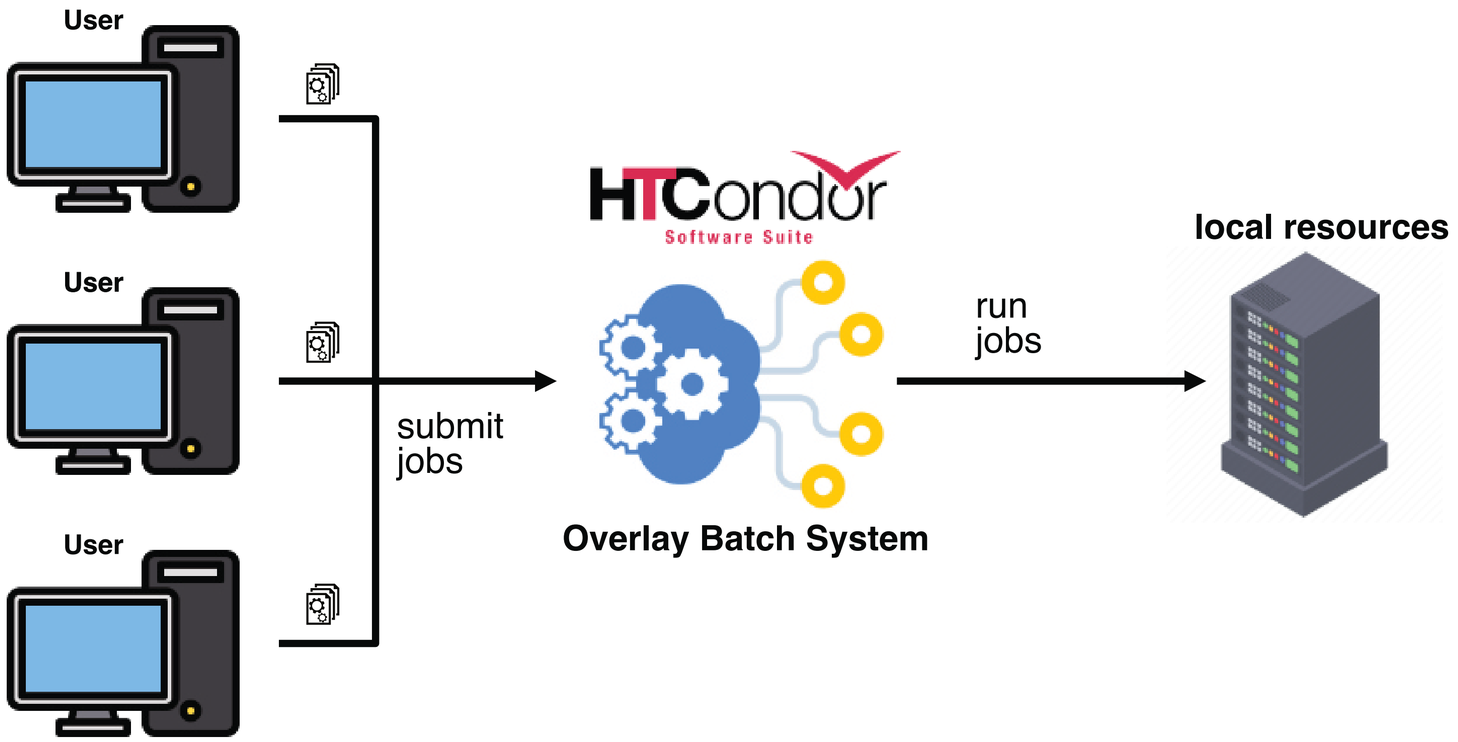
Possible via collaboration entrypoints but resources are shared between all members

**Solution: have dedicated resources for our institute members**

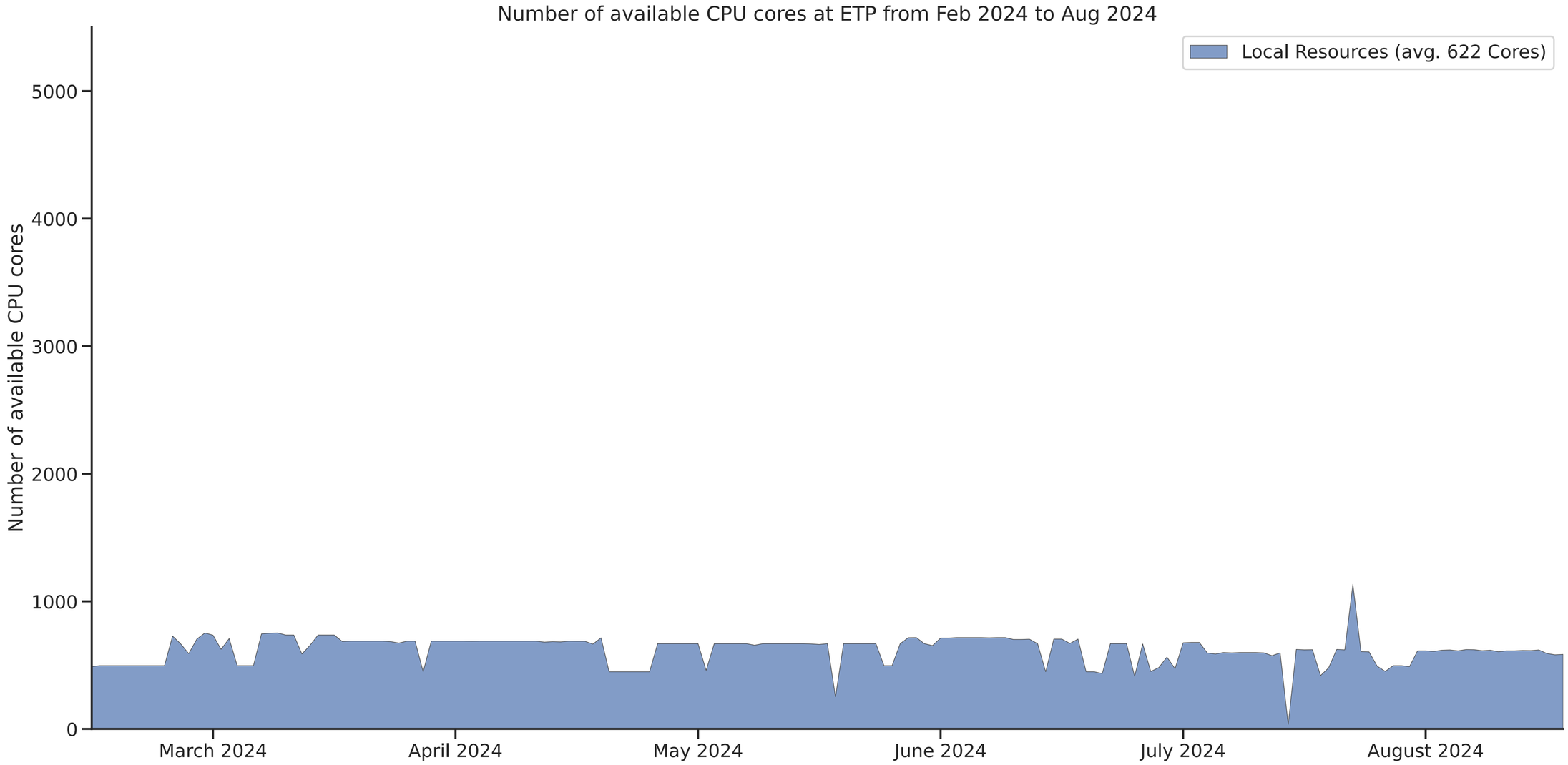
# How to bring computing resources to the analysts ?

# Overlay Batch System

An Overlay Batch System (OBS) based on HTCondor is used as a single entry point to computing resources for all Institute members

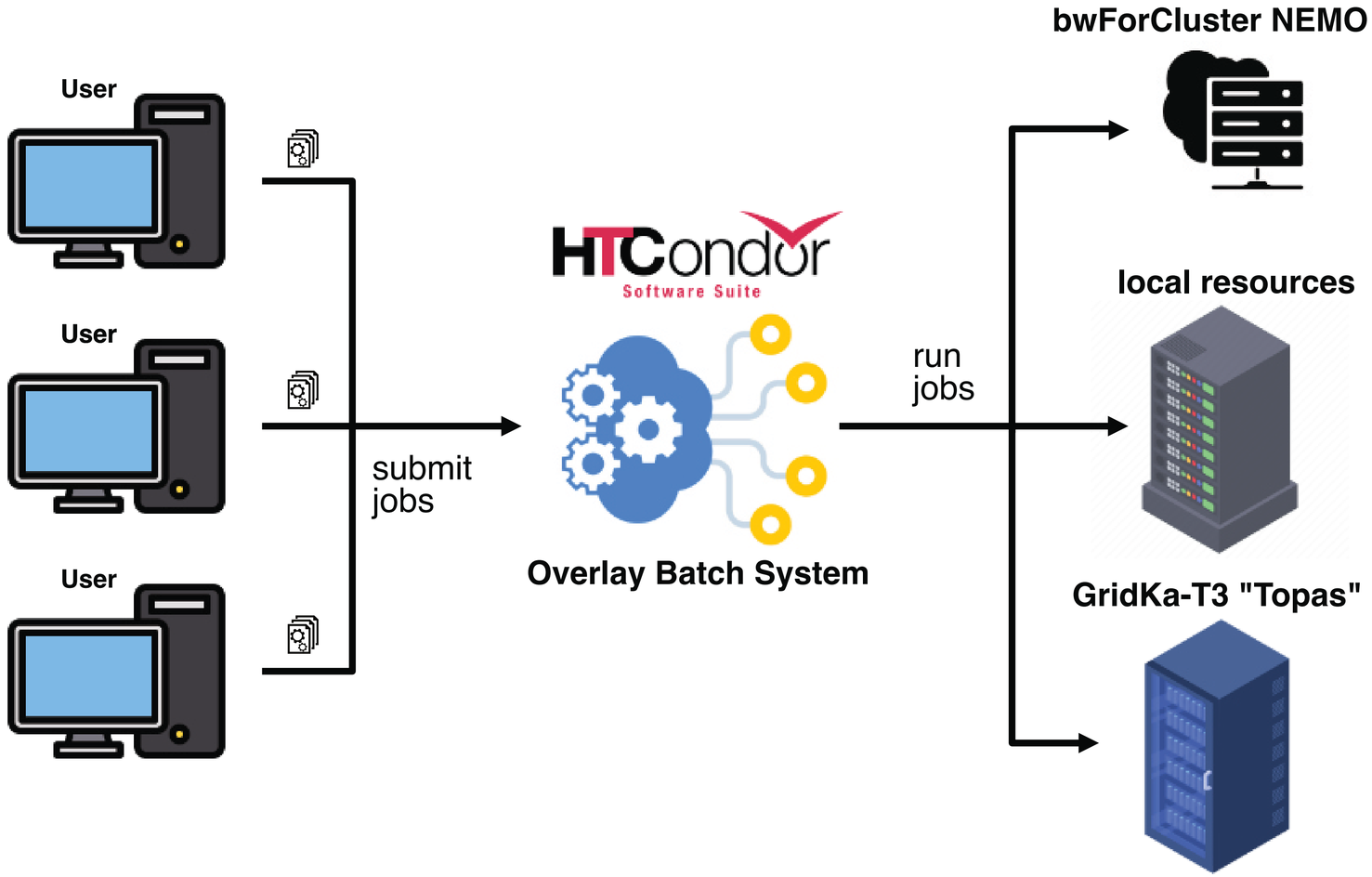


# Overlay Batch System

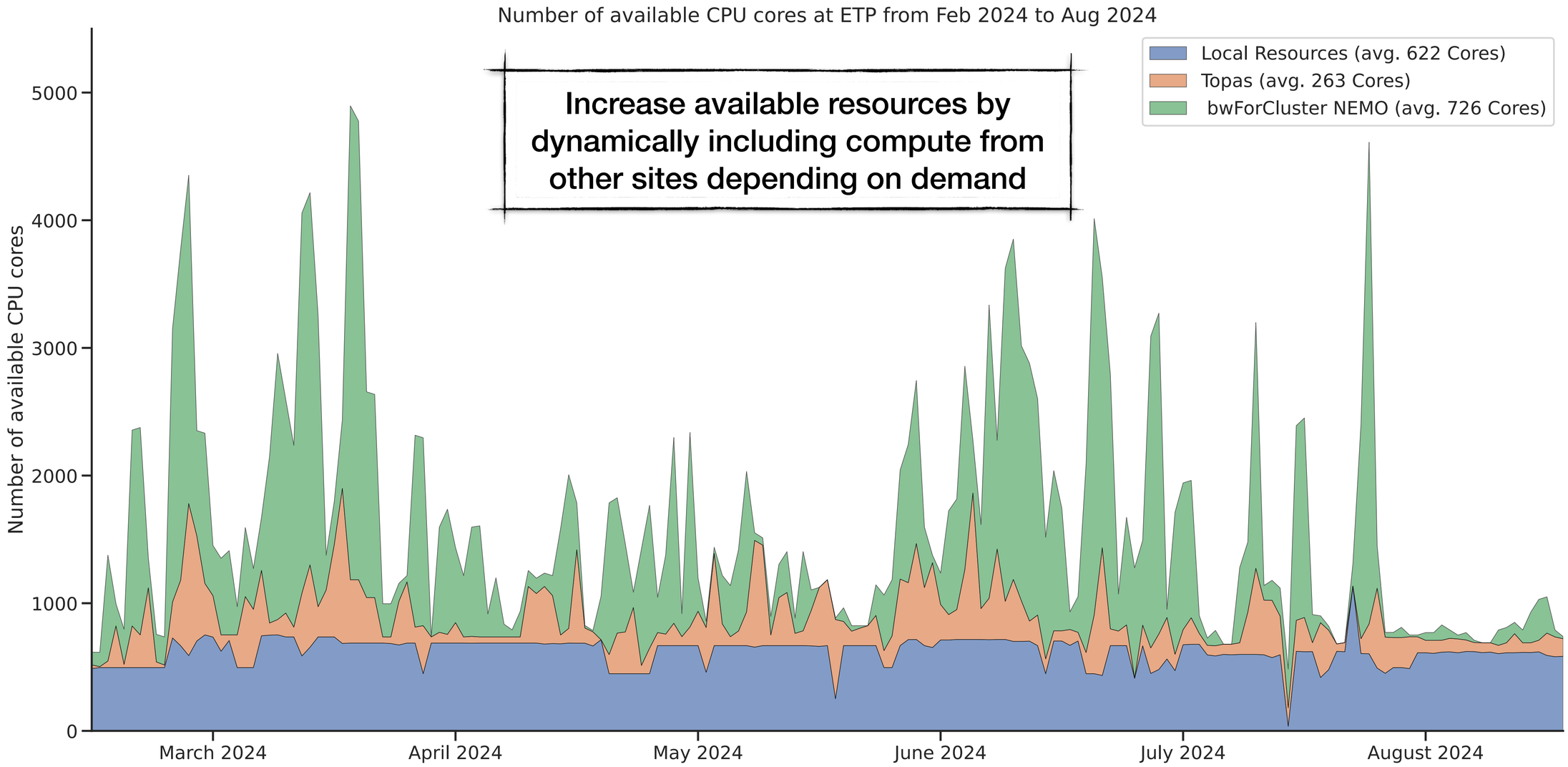


# Overlay Batch System

An Overlay Batch System (OBS) based on HTCondor is used as a single entry point to computing resources for all Institute members

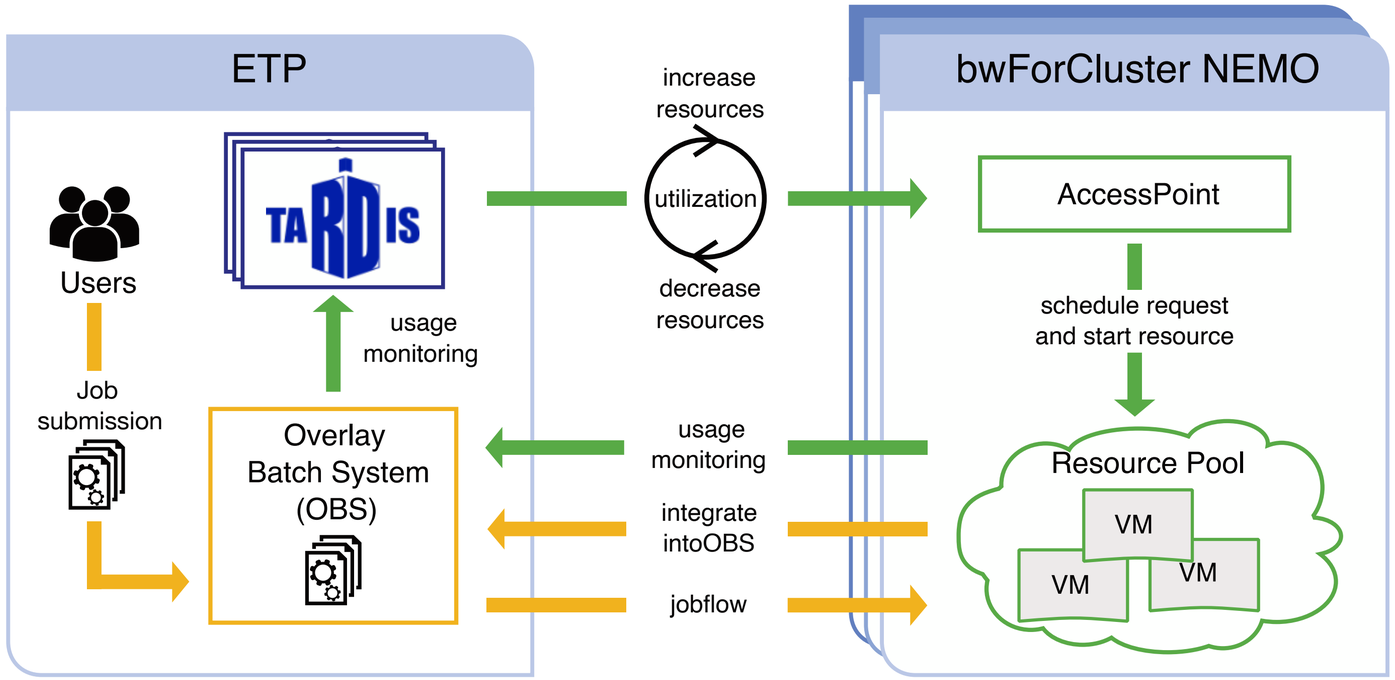


# Overlay Batch System



# Meta-Scheduling with COBaID/TARDIS

- Automatically add additional resources to the institute batch system, if there is demand
- Generic interface allows management of different sites with different schedulers etc

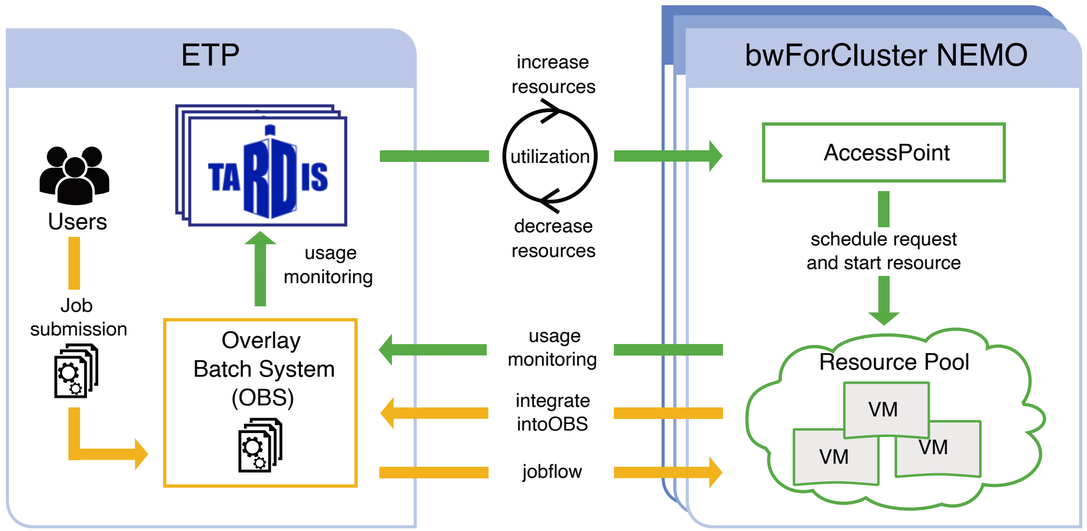



- https://github.com/MatterMiners/tardis
- https://github.com/MatterMiners/cobald

# COBaID/TARDIS on NEMO

1. Book resources via Moab (24/48h jobs)
2. Start a **drone VM** that runs an **HTCondor instance**.
3. Drone HTCondor instance **connects to the Institute batch system** and makes additional resources available
4. Institute jobs run within the VM in **docker containers**, allowing both single-core and multi-core jobs up to the drone limit
5. VMs are killed after latest 48 hours

**No interaction by the user required**



# COBaID/TARDIS on NEMO 2

1. Book resources via **slurm**
2. Start a drone **apptainer container** that runs an HTCondor instance. Container image is provided via cvmfs
3. Drone HTCondor instance connects to the institute batch system and makes additional resources available
4. Jobs run within the drone in **job containers**, allowing both single-core and multi-core jobs up to the drone limit
5. Drones are killed after latest 48 hours

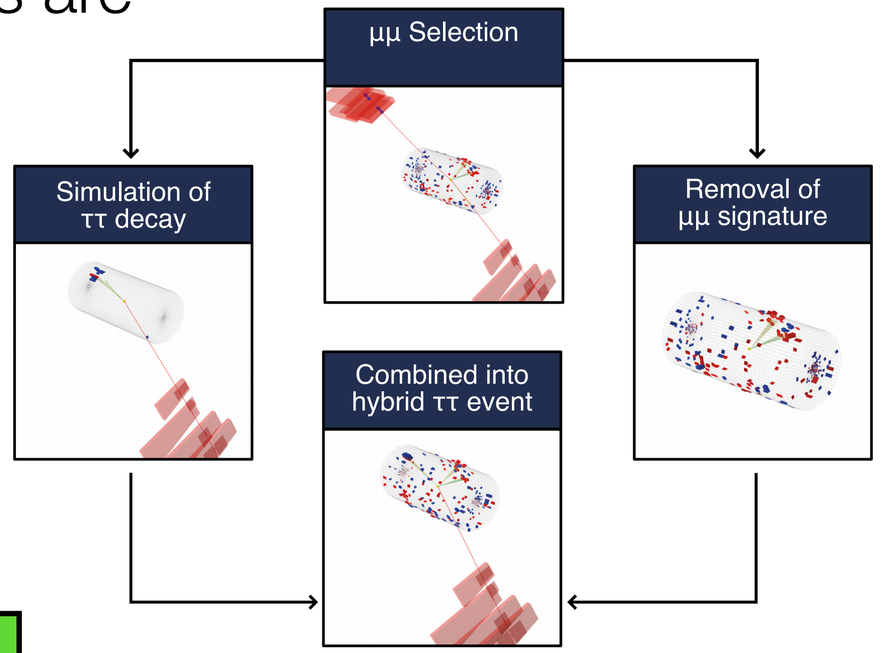
slurm workload manager

**tldr: Switch to slurm and fully containerised setup**

APPTAINER

# Example: Tau Embedding at CMS

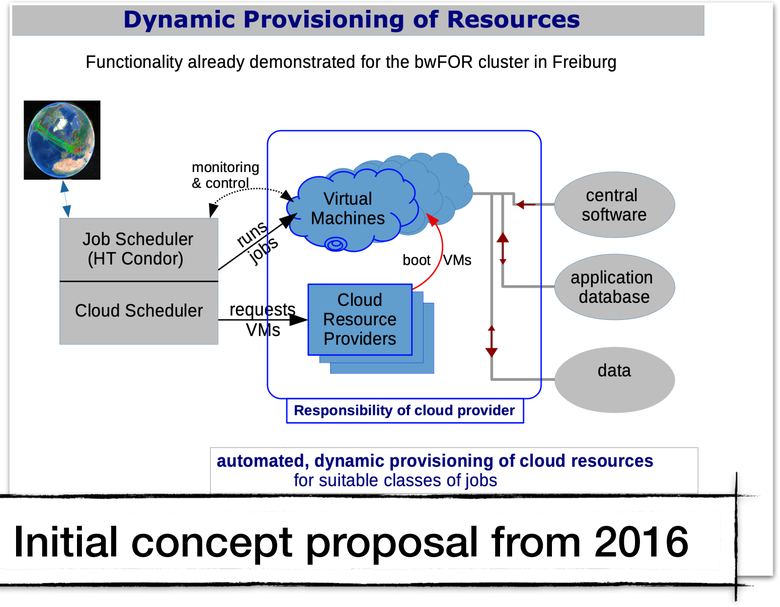
- Event based data driven method of **estimating genuine ττ decays** that **significantly** improves sensitivity of **CMS analyses in the ττ sector**
- Non-Standard procedure developed at KIT, samples are generated by **private productions**
- **Fast network connection** essential since workflow requires the full event information
- Processing takes **O(Mio) CPU hours,** majority provided by NEMO




- Paper: 10.1088/1748-0221/14/06/P06032

**Development and Support of this method only possible with reliable and performant computing resources**

# Conclusion




- Many analysis in HEP require a significant amount of computing resources
- Resources provided by **bwForCluster NEMO** are since many years one of the **most important pillars** of the **computing infrastructure for Karlsruhe particle physics groups**
- Dynamic integration allows us to **use the resources only when available** and **required**
- Modern, containerised setup for NEMO 2 is **ready to be used in production**

# Questions ?