

# Status Report Wuppertal

Raphael Kleinemühl

Bergische Universität Wuppertal

September 30, 2024

# Change in Workplan – Data-Lake Monitoring

- Original proposal assumed the development and setup of data lakes
- Partner at KIT (Astroparticle group) not supported by FIDIUM
- Data lake as distributed storage was never realized as foreseen
- Only local storage systems exist
- Existing tools for the monitoring of local storage systems are sufficient  
→ Shift of focus towards Caching systems  
This is reflected in the proposal for the extension as well

# Plans for the FIDIUM Extension

- Study current status of dCache and XCache as a caching solution
  - Data cache at current University Tier-2 centers, NHR centers etc.
  - Study workflow dependencies
- Development of simplified setup and operation methods
- Setup of dCache instance at federated dCache instance (DESY, HH)
- Setup XCache instance in Wuppertal
- Study performance under real conditions at ATLAS Tier-2 in Wuppertal
- Study in detail interplay with Cobald/Tardis
- Integration of Caching into Data management system RUCIO
  - For XCache a rudimentary solution exists (Virtual Placement)
  - Apply to dCache as well as Xcache (rudimentary solution exists)
  - Avoid unnecessary data transfers
  - Improved job placement

# First XCache instance in Wuppertal

XCache is deployed on Kubernetes via SLATE

“SLATE (Service Layer At The Edge) is a system to enable sites to delegate service deployment and configuration to selected application administrators.” [1]

- XCache Github repository is archived  
<https://github.com/slateci/XCache>
- Docker image is officially deprecated  
<https://slateci.io/XCache/instructions>
- Ilija Vukotic confirms that Kubernetes + SLATE is “the best way” to set up an XCache
- We plan to test XCache using the official way first and consider simplifications to the setup later

# Experience with XCache

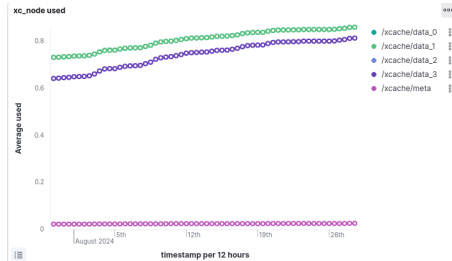
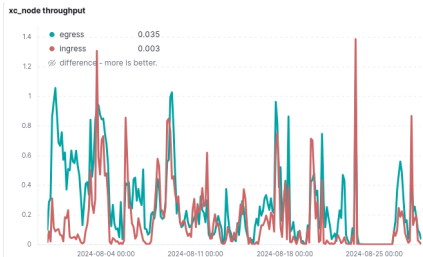
The setup process is non-trivial:

- We built a one-node k8s cluster specifically to host XCache
- Slate requires an additional public IPv4 address for the node running the k8s cluster
- Setup of MetalLB load balancer and associated IP pool necessary
- Linking the cluster to slate.io problematic as SLATE CLI is failing
- Long correspondence with Ilija Vukotic and Lincoln Bryant for configuration of XCache which is (only) done remotely
- Long correspondence for setup of the VP (virtual placement) queue

In conclusion: We need to create a simpler setup, especially for opportunistic use

- We created a Docker Compose Setup to replace Kubernetes and SLATE
- Four Services running
  - XCache itself
  - X509 updater
  - Heartbeat for VP
  - Reporter for Monitoring
- Needs to be supplied with minimal configuration (e.g. paths) and grid certificate

# Experience with XCache



- Typical for a month is 17.44TiB in, 43.39TiB out → Hitrate ~ 60%
- Will the hitrate improve once the “Low Watermark” is reached?
- Not all jobs use the XCache

## Conclusion

- XCache is up and running with a VP queue
- Preparation for extension period
- First simplifications of XCache setup done

## Outlook

- Study interplay with COBaID/TARDIS
- Is further integration with RUCIO possible?
- How does dCache compare?
- Need for better documentation