XRootD monitoring Update

WLCG Monitoring Task Force
Alessandra Forti, Borja Garrido, Derek Weitzel, Fabio Andrijauskas, Julia Andreeva, Katy Ellis, Shawn McKee

Grid Deployment Board

January 10, 2024

Index

- XrootD servers and status
- dCache with XRootD door and status
- Site network monitoring

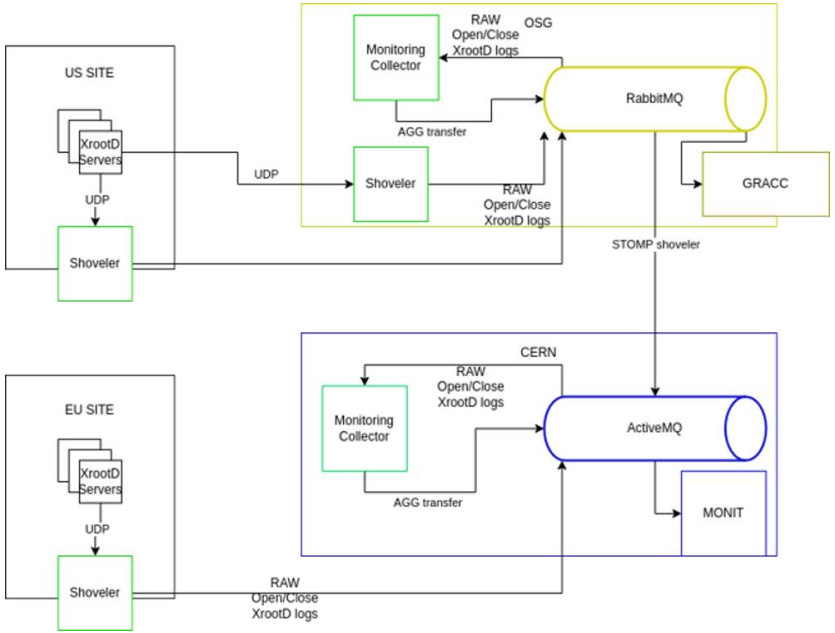
Reliable XRootD Monitoring

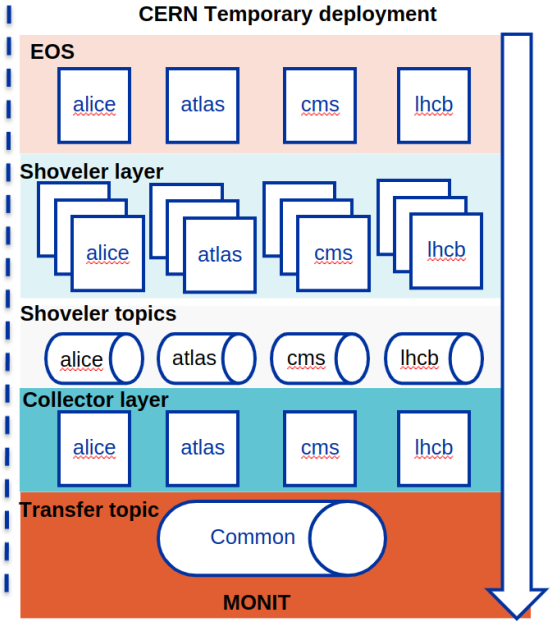
- Current integration (GLED) covers only XRootD servers
  - Lots of the data is lost in transport (unreliable)
  - New flow will aim to fix this issues.
    - Main change is replacing UDP protocol by message queue.
    - First implementation done by OSG colleagues (thanks to Derek Weitzel) has been adopted for the European part of WLCG
- XCache monitoring
  - Covered by new flow as data streams are identical
- Monitoring for dCache with xrootd door use case
  - Not covered under the WLCG scope

Reliable XRootD Monitoring (New flow) I

- Flow has been validated and the data we require for DC is as expected
  - Detected some fields produced by the new flow which might not be reliable but are not important for the exercise
    - (i. e : throughput, since we do approximate it over using the bytes read/write and time of the plot)
- New components are packaged and have been validated
  - Instructions for site managers on shoveler configuration can be found here
  - After first iterations with site managers, there are few request for improvements that need to be worked on
- Grafana dashboard to validate flow configuration (data arriving and being processed)
  - Should not be used for monitoring, that will be provided as part of other dashboard

Reliable XRootD Monitoring (New flow) II





Reliable XRootD Monitoring (New flow) III

- Regarding VO information (clarified after last DC workshop)
  - VO information available
    - If XRootD server uses VOMS authentication
    - If XRootD server uses tokens and is running version 5.6+
  - Otherwise
    - Use specific VO flows

Reliable XRootD Monitoring (New flow): Status

Minimum requirements for DataChallenge 2024:

- Collector and shoveler deployments at CERN
  - All data has been integrated and scalability validated
- EOS servers monitoring using the new flow
- Access to enriched “raw” data (enriched with CRIC topology, not aggregated)
  - To be done by the MONIT team

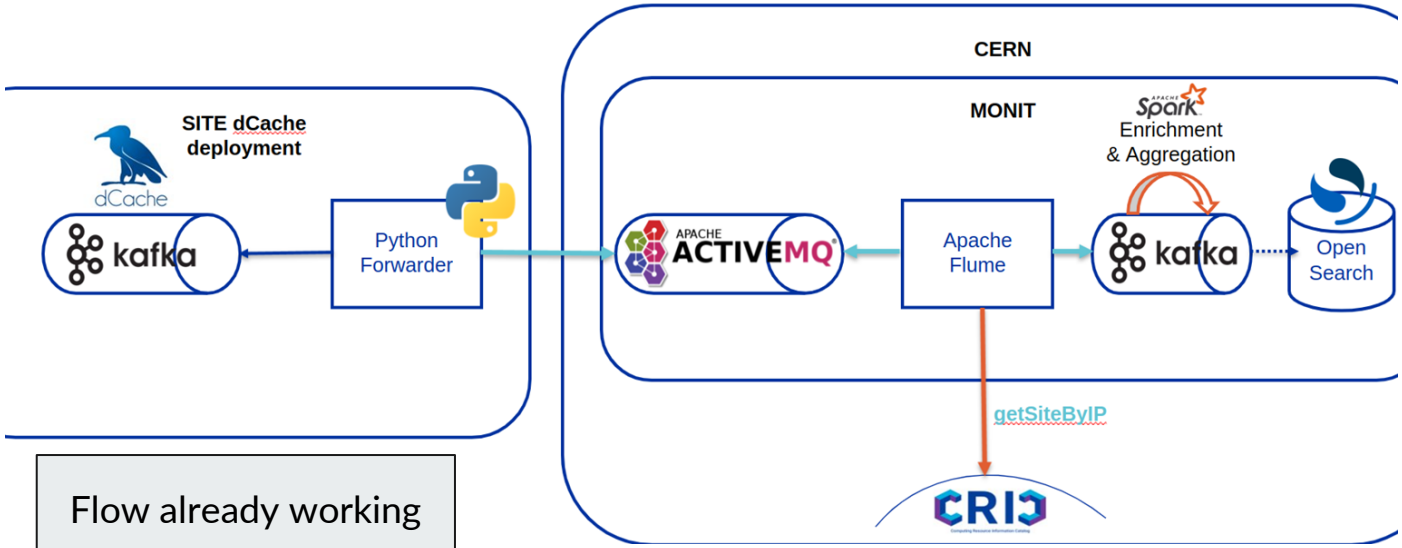
Nice to have:

- Other sites integrated. Part of US sites are already using message queue based implementation for a while. Campaign for European sites used by CMS , ATLAS and LHCb started
- Aggregation of data (in bigger bins with less labels for long term storage), not for DC per-se but good for accounting purposes

Reliable XRootD Monitoring (dCache with xRootd door) I

- Currently not covered by WLCG monitoring
- Billing information for dCache can be enabled by the site. Not done everywhere, not integrated on the global level.n
- Python script to get data from Kafka and push it to CERN MONIT system
  - Developed at DESY (Thanks Sandro Grizzo)
  - Released as part of dCache 9.2
- Resolution of IPs to site and enrichment required
  - To happen on the MONIT site
  - Based on CRIC data and APIs

Reliable XRootD Monitoring (dCache with xRootd door) II



Flow already working

Reliable XRootD Monitoring (dCache with xRootd door): Status

Minimum requirements for DataChallenge 2024:

- Integration script released as part of dCache
- Validation of the workflow with Desy (Thanks to Sandro Grizzo)
- FNAL sending data to CERN activeMQ for monitoring
  - This will cover pileup data access
- Access to enriched “raw” data (enriched with CRIC topology, not aggregated)
  - To be done by the MONIT team

Nice to have:

- Other sites integrated, if interested please get in contact with us (wlcgmon-tf@cern.ch), will require dCache 9.2+
- Aggregation of data, not for DC per-se but good for accounting purposes

Site Network Monitoring

- **GOAL:** To instrument and document site networks to understand usage and identify bottlenecks, filling in missing information identified during DC21.
  - Presentation from April WLCG Ops Coordination covers details
  - See DOMA project description
- Project is hosted on CERN Gitlab: https://gitlab.cern.ch/wlcg-doma/site-network-information
  - Currently have Python snmp monitoring example + new GO version to be added
- Site network monitoring visible (by RCSITE or NetSite) at https://monit-grafana-open.cern.ch/d/MwuxgogIk/wlcg-site-network?orgId=16&from=now-7d&to=now&var-site=All

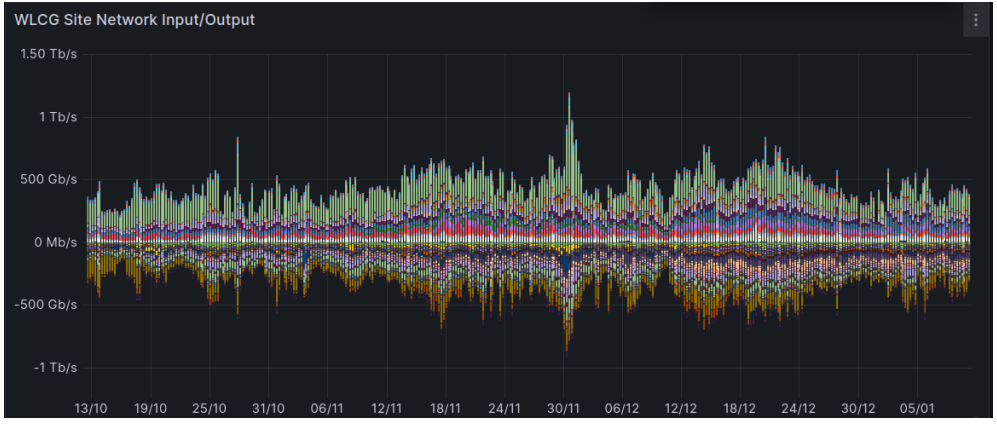
Site Network Monitoring II

Minimum requirements for DataChallenge 2024:

- We expect ALL sites that plan to participate in DC24 to provide the requested network description AND total IN/OUT traffic for the whole site as described in CERN Gitlab.
- We currently have ~70% of the campaign identified sites complete (43/~64).
- May need ticket follow up for non-responsive sites or new tickets for missing DC24 sites.

Nice to have:

- Comprehensive net description
- Site network diagrams



Summary

XRootD servers monitoring

- EOS servers at CERN being monitored
- Enrichment flow in QA, will be available during next week
- Other sites integrated already
    - Few improvements for the shoveler requested (to be worked on during January)

dCache with XRootD door

- Flow validated with DESY
- FNAL integration waiting for dCache upgrade on their side
- Other sites
    - Please contact us so we can help with the integration
    - Remember you need dCache 9.2+