

Data Challenges Monitoring WS

A. Forti

Monitoring workshop

27 April 2021



Data Challenges

- As part of the planning for HL-LHC Data Challenges are foreseen from this year every 2 years until the start of HL-LHC.
- Increasing transfer rates until we reach the required level for HL-LHC
 - 2021, 2023, 2025, 2027
- The overall aim is to prepare the WLCG infrastructure and demonstrate that we can use the bandwidth effectively
 - “The DC should follow and drive the expansion of the network capacity”



DC: infrastructure

- Approach should be step by step
 - Develop the tests methodology
 - **Build on what is already there and plan for evolution**
- All 4 experiments in parallel and all sites
- Use experiments production infrastructure
 - Storage, network, services, tools,....
- In parallel with experiments activities
 - **Both DC and activities use networking**
 - Need to discriminate
- Common Infrastructure for DCs
 - **Services to run the challenges**
 - Resources being requested under DOMA umbrella
 - **Monitoring**

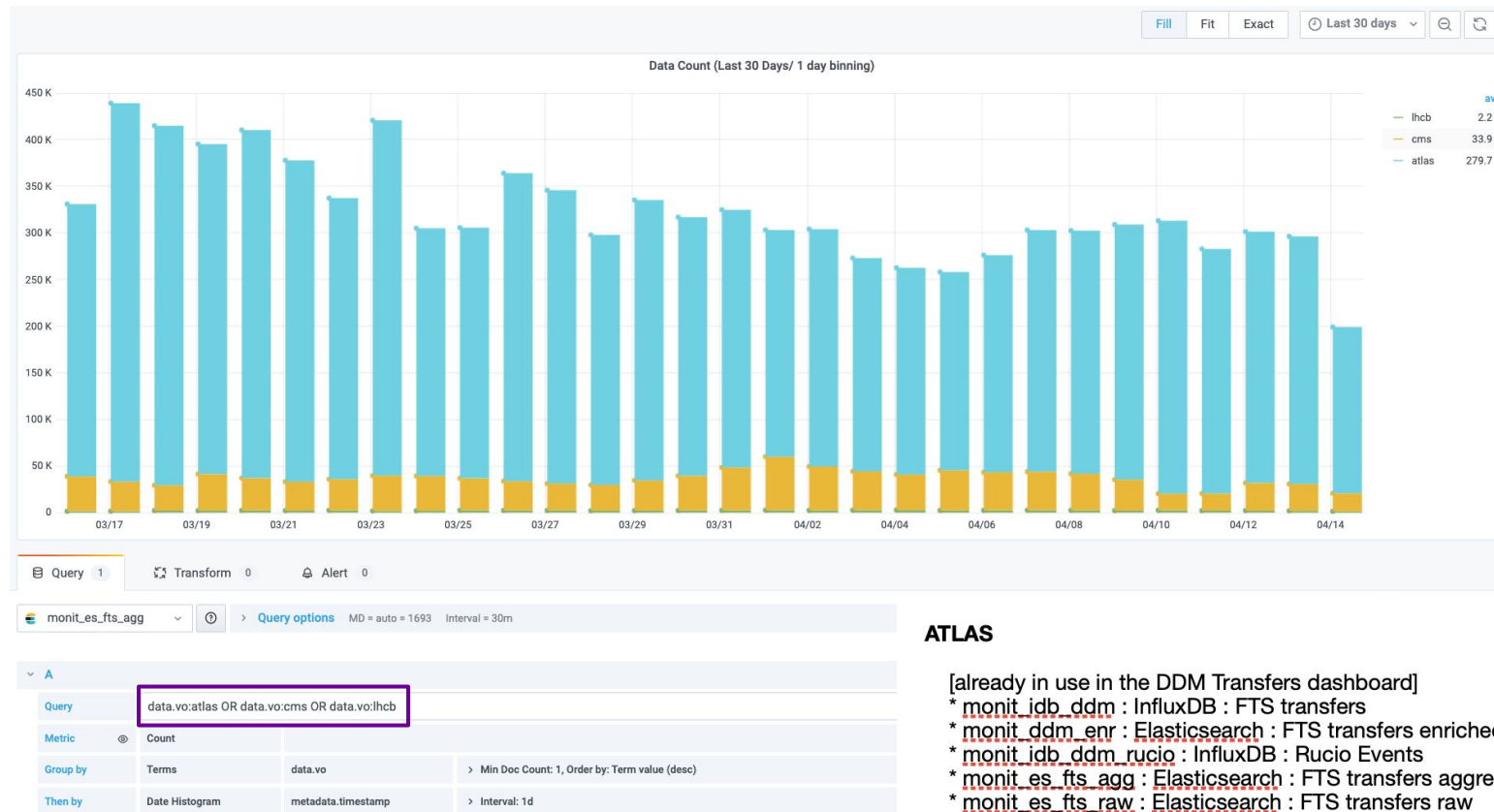


DCs Monitoring

- Most important part to get right
 - What we do now is the baseline for future challenges
 - Baseline for future data challenges
- There is already a large MONIT infrastructure that experiments use to various degrees
 - Experiments have their own dashboards
 - Use of the central monitoring infrastructure is uneven
 - Even when using same tools might not have the same information included
 -
- Getting a more uniform monitoring is the work that needs to be done
 - Even if not all at once
 - Not all the experiments? not all the information? not yet?



DCs Monitoring



- Initial sources overview
 - Hopefully can get more out of today

ATLAS

- [already in use in the DDM Transfers dashboard]
- * monit_idb_ddm : InfluxDB : FTS transfers
- * monit_ddm_enr : Elasticsearch : FTS transfers enriched
- * monit_idb_ddm_rucio : InfluxDB : Rucio Events
- * monit_es_fts_agg : Elasticsearch : FTS transfers aggregated (hourly aggregations)
- * monit_es_fts_raw : Elasticsearch : FTS transfers raw

CMS

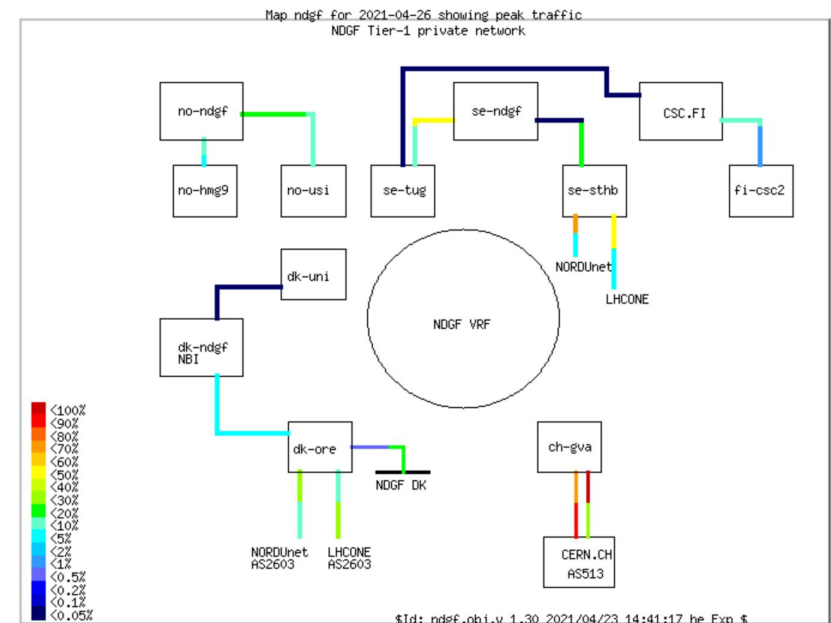
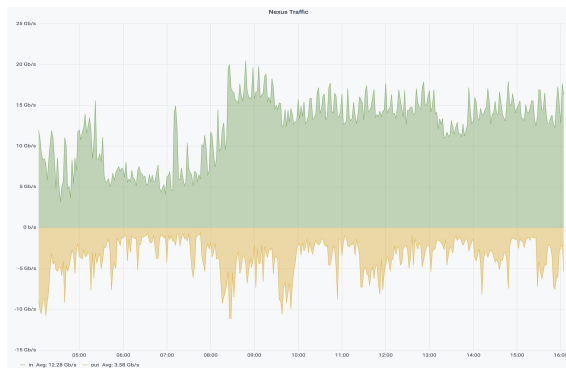
- [already in use in the CMS FTS metrics dashboard]
- * monit_es_fts_agg : Elasticsearch : FTS transfers aggregated (hourly aggregations)
- [already in use in the FTS log clustering dashboard]
- * monit_es_ftsloganalysis : Elasticsearch : FTS transfers logs
- * monit_es_fts_raw : Elasticsearch : FTS transfers raw

LHCB

- * monit_es_fts_agg : Elasticsearch : FTS transfers aggregated (hourly aggregations)
- * monit_es_fts_raw : Elasticsearch : FTS transfers raw

Sites Monitoring

- Sites are the other important component, experiments can only measure what they transfer but do not measure the sites inbound and outbound traffic or the bandwidth
 - Different level of monitoring different level of exposure
 - How can we integrate? Network people and sites participation important!



Common effort

- Data Challenges are by definition a common effort that requires the involvement of experiments, network people, monitoring and sites
- Step by step
- Hopefully today we are going to have a better view of what we need to do for this year and future challenges
- There might be future workshops to track progress



Links

- [HL-LHC network needs and data transfer challenges](#)
- [Data Challenges planning docs](#)
-

