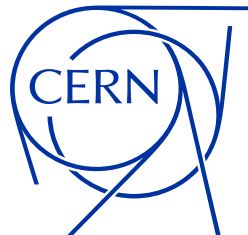


Data Challenge 2024 Retrospective

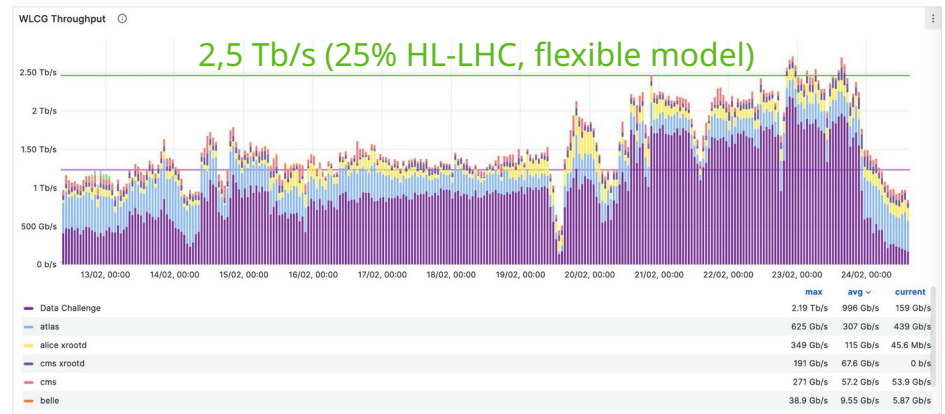
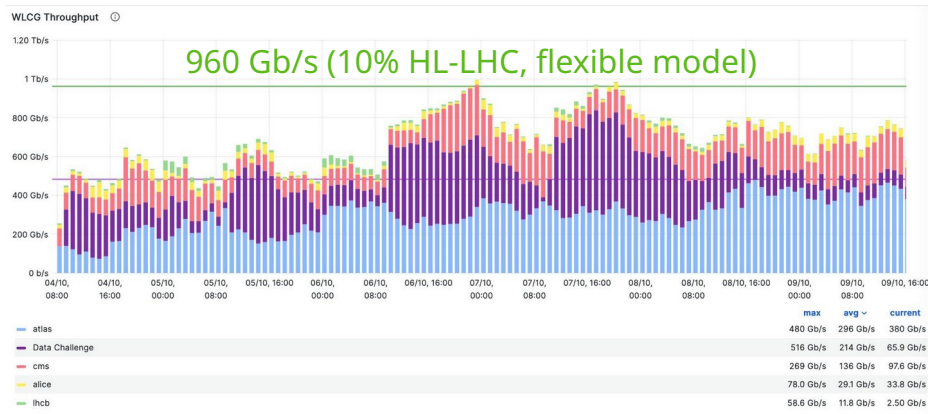
Dimitrios Christidis



Introduction

- The LHC Run-4 will deliver a significant increase in luminosity
- The Data Challenges are major orchestrated tests in anticipation of this increase
 - Main goal is to achieve some target transfer rate
 - Some secondary goals included (e.g. tokens)
 - Active participation from all stakeholders (all LHC and some non-LHC experiments, data-management services, sites, networks)
- Refer to the [final report](#) for more details

Outcomes



Transfer rates during DC21 (left) and DC24 (right).

Summary for Rucio

- The communities using Rucio reported overall success!
 - Rucio proved to be able to scale and meet the demands of the Data Challenge
 - Increasing the number of component instances and/or threads is a manual endeavour, with no obvious recommendations
- Focus on the areas of improvement
 - The injection method is somewhat artificial (e.g. short replication rule lifetimes) and exacerbated some of the observed issues

Transfer submission to FTS

- Following an FTS outage, ATLAS reported poor submission performance
 - The top of the queue was dominated by transfer requests of expired replication rules
 - Contention between Rucio daemons
- Tracked in issue [#6505](#)
 - ATLAS applied a primitive patch directly on production

Handling of expired replication rules

- ATLAS reported inability to delete expired replication rules on large datasets in their early stages of replication
 - In one case, more than 24 hours
 - Contention between Rucio daemons
- No significant effect on overall performance
- Tracked in issue [#6511](#)

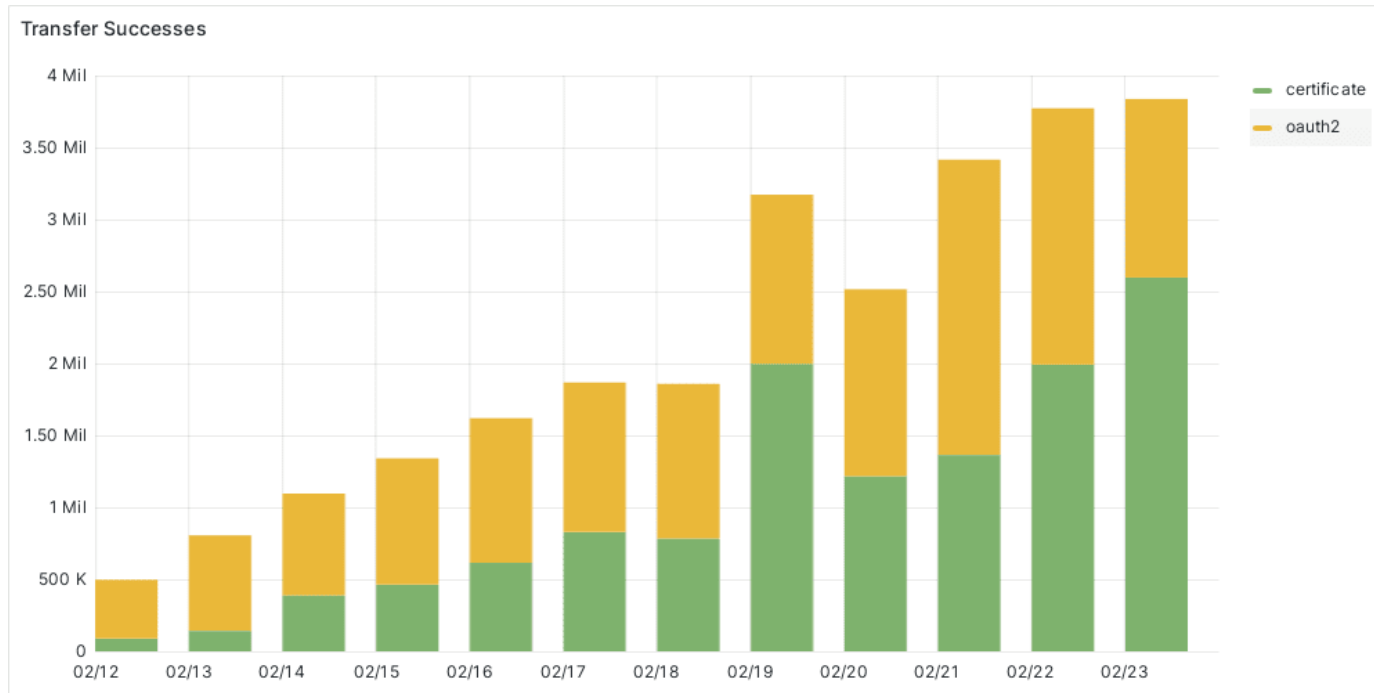
Deletion overlap on slower sites

- CMS reported poor deletion performance at some sites
 - Underlying issue is the rate of deletion at the sites themselves
 - But a design in Rucio does not handle this optimally, leading to multiple threads working on the same files
 - This hinders performance even further
- Affected dataset reuse and storage occupancy
- Tracked in issue [#6512](#)

Tokens (1)

- Both the Rucio and FTS advertised their token implementations a 'technology preview'
- Despite this, and the fact that it was a secondary goal, it was without a doubt a success!
- However, significant concerns were raised, mainly in terms of robustness and security
- Refer to [tomorrow's talk](#) for more details

Tokens (2)



Use of tokens over certificates (ATLAS & CMS, 'Data Challenge' activity). [Source]

Injection method

- Currently, an [external collection](#) of scripts, SQL queries, and configuration files, under WLCG DOMA
- Expressed interest in evolving and integrating it in Rucio
 - As a generic stress-testing component
- Expecting some dedicated effort by the end of the year

Questions?