

The 2021 WLCG Data Challenges

Thursday, March 24, 2022 3:30 PM (20 minutes)

The gradual approach of the High-Luminosity LHC (HL-LHC) era poses interest in the status and in the expected behaviour of the WLCG infrastructure. In particular, efficient networks and tape storage are expected pillars of success in the upcoming LHC phase. Considering the current computing models and data volume of the LHC experiments, the requirements are mainly driven by custodial storage of data on tape, export of RAW data from CERN to the Tier-1 centres, as well as data reprocessing. For this reason, an activity under the WLCG Data Organisation, Management, and Access (DOMA) forum has been started to assess the readiness of the WLCG infrastructure for HL-LHC.

A hierarchical model referred to as “Minimal”, which considers a T0-T1-T2 traffic flow only, and a realistic model referred to as “Flexible” are taken into account during the assessment exercise. They represent the range within which future planning should occur. A capability to fill around 50% of the full bandwidth for the Minimal scenario with production-like storage-to-storage traffic should be demonstrated by the HL-LHC start. Consequently, increasingly larger challenges are scheduled throughout the years anticipating HL-LHC start in 2027, aiming at upsizing targets and goals for both scenarios.

The 2021 challenge was the first of these series of challenges, pivotal to set a baseline, and represented the cornerstone of the LHC Run-3 preparation. Moreover, it provided the playground for other activities, such as the commissioning of HTTPS as a Third Party Copy (TPC) protocol instead of gsiftp at pledged sites. During the challenge, which was split into a Network Challenge and a Tape Challenge, production traffic has been backfilled by additional dedicated, centrally-managed, Data Challenge traffic to reach the 2021 target. The usage of the production infrastructure of the experiments has the benefit of testing the presence of bottlenecks across production services, focusing on the currently deployed infrastructure. Despite production services of the WLCG sites and experiments being used, innovative and original work has been produced with respect to a centralised infrastructure handling the injection of data to transfer and a common monitoring solution to provide the WLCG community with a unified picture of the four LHC experiments. Finally, this challenge boosted ongoing activities and created new task forces on specific topics that have been identified to be crucial for future challenges in order to successfully reach the HL-LHC target.

Primary authors: FORTI, Alessandra (ATLAS); DI MARIA, Riccardo (CERN); MAIER, Benedikt (CERN); GARRIDO, Borja (CERN); HAEN, Christophe (CERN); CHRISTIDIS, Dimitrios (ATLAS); GOMEZ-CORTEZ, Felipe (CMS); GARZON, Fernando (CERN); PASPALAKI, Garyfalia (CMS); ELLIS, Katy (CMS); BETEV, Latchezar (CERN); MASCETTI, Luca (CERN); RIOS, Maria Arsuaga (CERN); LASSNIG, Mario (CERN); PATRASCOIU, Mihai (CERN); DONA, Rizart (CERN); CAMPANA, Simone (CERN); ZHAO, Xin

Presenter: DI MARIA, Riccardo (CERN)

Session Classification: Data Management & Big Data

Track Classification: Track 6: Data Management & Big Data