

# Caching for Analysis Workloads of the ATLAS LHC experiment

Since the start of the Large Hadron Collider (LHC) in 2008, the Worldwide LHC Computing Grid (WLCG) has been serving the needs of the largest LHC experiments' detectors - ATLAS, CMS, LHCb, and ALICE. The volume of the data coming from these experiments every year exceeds 90 PB per year, and the rate of the raw data reaches 100 GB/s, which requires the best approaches in computing and storage system architectures.

WLCG has a hierarchical three-tier structure. The tasks of data archiving and serving, computing, and running the infrastructure services are distributed amongst the grid sites depending on the Tier they belong to.

Optimizing WLCG is crucial for the success of the whole LHC program, especially in the light of the upgrade of the accelerator - the High Luminosity LHC (HL-LHC). This will make it possible to obtain higher energies and a much greater number of collisions, which may lead physicists to new scientific discoveries. The increase in data volume and complexity outpaces the expected hardware gains, and staying within the budget requires significant improvements in the efficiency.

ATLAS is the largest of the LHC detectors and is built and operated by a collaborative effort of physicists, engineers, and technicians from around the world. The typical data processing workflow of this experiment could be roughly divided into reconstruction and analysis activities [1]. The reconstruction campaigns are aimed at preparing the raw data coming from the LHC detectors for the physics users. They are run in a well-organized, scheduled manner, while the analysis tasks are relatively sporadic, as they are independently triggered by different users according to their needs. The analysis data are stored in a form of events, which are further united into files and datasets. This hierarchical data organization implies specific data access patterns.

Caches have proved to be an effective technique to hide the latency of the file accesses and reduce the network traffic. From the detailed log information coming from the CERN Data Center (Tier 0), we extract traces of user accesses and evaluate the potential benefit of adding a caching layer under the different cache eviction policies. In particular, we compared the effects of the widely used and state-of-the-art techniques (LRU, 2-LRU) and compare them to the optimal policies (we use the PFOO-L approximation [2]). To evaluate the cache performance, we use both the Object Miss Ratio and the Bytes Miss Ratio. For the latter, we suggest a modification of the PFOO-L algorithm to estimate the optimal for the Bytes Miss Ratio. Additionally, we analyse the dependencies between the individual file accesses within the same dataset and evaluate the performance of dataset-specific policies, which are suitable for any workloads reading files in clusters (datasets).

We also consider the remote computations model, where we explore the possibility for some Tier 2 sites to move to the storageless mode and be solely used as a processing unit. In this case, the temporary storage to couple the CPU capacities will be mostly used as a cache for the processing data. In this model, we take into account the constraint on the connectivity between the storage and the processing units. This in turn may lead to the phenomenon called 'delayed hits', which affects the cache performance. In this work, we evaluate models with both file-based and dataset-based caching algorithms in order to estimate the influence on the cache performance and evaluate the connectivity bandwidth needed for the transition to the storageless sites.

[1] ATLAS, Collaboration, et al. "ATLAS Computing: technical design report." LHCC Reports; ATLAS Technical Design Reports (2005).

[2] Berger, Daniel S., Nathan Beckmann, and Mor Harchol-Balter. "Practical bounds on optimal caching with variable object sizes." Proceedings of the ACM on Measurement and Analysis of Computing Systems 2.2 (2018): 1-38.

**Primary authors:** CHUCHUK, Olga (CERN, University of Cote d'Azur); Dr SCHULZ, Markus (CERN); Dr NEGLIA, Giovanni (Inria)

**Presenter:** CHUCHUK, Olga (CERN, University of Cote d'Azur)

**Session Classification:** Network, Security, Infrastructure & Operations

**Track Classification:** Track 7: Network, Security, Infrastructure & Operations