Contribution ID: **578** Contribution code: **contribution ID 578**　　　　　　　　Type: **Oral**

Exploring data merging methods for a distributed processing system

Wednesday 1 December 2021 19:00 (20 minutes)

The ALICE experiment at the CERN LHC (Large Hadron Collider) is undertaking a major upgrade during the LHC Long Shutdown 2 in 2019-2021, which includes a new computing system called O2 (Online-Offline). The raw data input from the ALICE detectors will increase a hundredfold, up to 3.5 TB/s. By reconstructing the data online, it will be possible to compress the data stream down to 100 GB/s before storing it permanently.

The O2 software is a message-passing system. It will run on approximately 500 computing nodes performing reconstruction, compression, calibration and quality control of the received data stream. As a direct consequence of having a distributed computing system, locally generated data might be incomplete and could require merging to obtain complete results.

This paper presents the O2 Mergers, the software designed to match and combine partial data into complete objects synchronously to data taking. Based on a detailed study and results of extensive benchmarks, a qualitative and quantitative comparison of different merging strategies considered to reach the final design and implementation of the software is discussed.

Significance

The presentation will summarize how we merge incomplete objects in our large computing system and explain our thought-process when choosing the final solution.
According to our research, other data acquisition and quality monitoring systems at CERN LHC perform these tasks as well, but the solutions were never presented in detail.
We would like to share our experiences in this topic to help any current and future developers of similar software.

References

The software was shortly described in the context of the ALICE data Quality Control system presentation: https://cds.cern.ch/record/2756296/files/10.1051_epjconf_202024501027.pdf

Speaker time zone

Compatible with Europe

Primary authors: VON HALLER, Barthelemy (CERN); KONOPKA, Piotr (CERN)

Presenter: KONOPKA, Piotr (CERN)

Session Classification: Track 1: Computing Technology for Physics Research