

Data distribution and load balancing for the ALICE Online-Offline (O2) system

Thursday 12 July 2018 12:00 (15 minutes)

ALICE (A Large Ion Collider Experiment), one of the large LHC experiments, is undergoing a major upgrade during the next long shutdown. Increase in data rates planned for LHC Run3 (3TiB/s for Pb-Pb collisions) with triggerless continuous readout operation requires a paradigm shift in computing and networking infrastructure.

The new ALICE O2 (online-offline) computing facility consists of two types of nodes: First Level Processors (FLP), containing the read-out PCI cards, and Event Processing Nodes (EPN), responsible for the online reconstruction. Each FLP node buffers detector data, of a predefined time interval, called SubTimeFrame (STF). The central task of the data distribution is to aggregate corresponding STFs from all FLP nodes into an object called Time Frame (TF). The FLP-EPN network must support the high aggregate data rate and sustain a large number of concurrent transfers. An application-level scheduling of data transfers as well as the selection of the receiving EPNs will be necessary to keep the high quality of service.

We give an overview of the TF building process including FLP node synchronization, traffic shaping and balancing for even utilization of processing and network components of the O2 facility.

Primary authors: NESKOVIC, Gvozden (Johann-Wolfgang-Goethe Univ. (DE)); KIRSCH, Stefan (Johann-Wolfgang-Goethe Univ. (DE))

Presenter: NESKOVIC, Gvozden (Johann-Wolfgang-Goethe Univ. (DE))

Session Classification: T1 - Online computing

Track Classification: Track 1 - Online computing