Sample Project: Data encoding and compression for the upgrade of the ATLAS experiment

Code	PH3802
Programme	FCT
Department	PH
Responsible	22012 - Dr. Nicolas Ellis
Created by	42992 - Dr. Giovanna Lehmann Miotto
Updated by	96245 - Mr. Vasco Miguel Chibante Barroso
Date Created	21-MAY-15
Date updated	19-JUN-15

Title

Data encoding and compression for the upgrade of the ATLAS experiment

Description

The ATLAS experiment so far produced several PB of raw data, organised into 30 millions data files. These data and files use simp binary encoding schemes based on fixed headers and allow data compression via the well-known DEFLATE algorithm.

Future upgrade plans will drive a fifty-fold increase in the data storage rates. In addition, the ATLAS experiment is looking forward strengthen the data preservation and open access policies. As a consequence, there is the need to review, modernise and future proof the raw data storing format. The successful candidate will investigate different aspects like:

-- candidate technologies for the replacement of the current encoding scheme. These should allow simpler backward and forwa compatibility, possibly providing self-describing features, while remaining space efficient and fast;

-- alternative, and possibly custom, data compression schemes. The capability of reducing the data volume has a direct impact o the requirements and cost of the on-line data-acquisition system as well as on the long-term data storage and distribution

-- review and redesign the C++ API of the raw data handling libraries incorporating the latest features of the standard

Skills

Disciplines

Applied Physics, Information Technologies