Contribution ID: **249** Type: **Poster presentations**

Using binary file format description languages for verifying raw data in astroparticle physics experiments

Monday, 10 September 2018 16:00 (1 hour)

The exponential growth of the amount of astroparticle data is expected to happen in near future. This trend gives rise to a number of emerging issues of big data management. One of the important issues is how to describe and verify raw binary data to support their availability and reuse in future. The present work demonstrates a possible solution for this issue in application to data of TAIGA observatory, which consists of a number of facilities featuring five different formats of raw data. The long-term preservation of raw binary data as originally generated is essential for rerunning analyses and reproducing research results in future. In this case, the raw data should be well documented and accompanied by the parsing and verifying tools. There are some declarative languages for describing binary file format description (e.g. DFDL, FLEXT, KAITAI STRUCT). The present work shows the progress of the application of KAITAI STRUCT to specify, parse and verify raw data of the TAIGA binary file formats. The format specifications implemented using this framework allow us to generate program code for parsing and verifying the raw binary data in the target languages (C++, Java, Python, etc.). The libraries were tested on real data, have shown good performance and indicated the parts with corrupted data. This study can be interested in other experiments which raw binary data formats remain weakly documented. This work was financially supported by the Russian Scientific Foundation (grant 18-41-06003).

Primary author: Dr KRYUKOV, Alexander (SINP MSU)

Co-authors: KHMELNOV, Alexey (Matrosov Institute for System Dynamics and Control Theory SB RAS); SHIGAROV, Alexey (ISDCT SB RAS); MIKHAILOV, Andrey (Matrosov Institute for System Dynamics and Control Theory SB RAS); KOSTUNIN, Dmitriy (Karlsruhe Institute of Technology (KIT)); KOROSTELEVA, Elena (Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University); BYCHKOV, Igor (Matrosov Institute for System Dynamics and Control Theory SB RAS); FEDOROV, Oleg (Applied Physics Institute Irkutsk State University)

Presenter: KOROSTELEVA, Elena (Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University)

Session Classification: Poster Session