Contribution ID: **6**   Type: **Submitted talk**

Recent improvements in the HDF5/Blosc2 plugin systems

Tuesday 19 September 2023 11:00 (30 minutes)

Recently, the hdf5plugin (https://www.silx.org/doc/hdf5plugin) has gained support of the Blosc2 library. This allows for HDF5/h5py to use many of the technologies that Blosc2 already supports.

In our talk, we will be describing recent work that we have conducted in enhancing Blosc2, namely:

1) A new dynamic plugin system, that can be easily installed via Python wheels.

2) A new dynamic plugin for the HTJ2K (https://github.com/osamu620/OpenHTJ2K) codec. This codec has better performance and image quality scores than e.g. JPEG (https://htj2k.com/htj2k-versus-ye-olde-jpeg/).

3) Support for Blosc2 NDim inside HDF5/PyTables. Blosc2 NDim leverages a double partition (chunks and blocks) for storing data, allowing for a better utilization of L1/L2/L3 cache hierarchy in CPU caches. This makes for increased performance when reading general slices in multi-dimensional datasets. This implementation can be leveraged to do a port for h5py, and will provide hints on doing this.

4) We will briefly introduce Btune (btune.blosc.org), a tool for automatically selecting the best combination of codecs and filters based on a user-specified tradeoff between compression ratio and speed.

Most of these enhancements should be available for the HDF5/h5py via the hdf5plugin, with minimal modifications.

Website

www.blosc.org

Primary author: Mr ALTED, Francesc (Blosc project)

Presenter: Mr ALTED, Francesc (Blosc project)

Session Classification: Day 1