

Down to the bytes: can we simplify alternative access to HDF5?

Tuesday 27 May 2025 09:25 (25 minutes)

HDF5 is an enormously powerful and flexible file format. There are many different ways to use it, and it's difficult to provide one API that works efficiently for all the possible use cases. However, the complexity of the on-disk file format is a high barrier to alternative implementations, so with a few heroic exceptions, most code reading & writing HDF5 does so through the canonical C implementation, including its bindings in various other languages.

In this open-ended talk, I aim to give an overview of a range of different ways to access HDF5 files, bypassing part or all of libhdf5. This will include practical techniques to speed up access in specific circumstances, projects building additional indexes around HDF5 files, and alternative implementations of the file format. I'll conclude with some ideas on how we might lower the barrier and make it easier to implement optimised HDF5 file access for specific scenarios.

May we record your session?

Yes

Primary author: KLUYVER, Thomas (Eur.XFEL (European XFEL))

Presenter: KLUYVER, Thomas (Eur.XFEL (European XFEL))