Contribution ID: **20**   Type: **20-minute presentation + 5-minute Q&A**

Using a HDF5 File as a Zarr v3 Shard

Monday 26 May 2025 14:10 (25 minutes)

Version 3 of the Zarr specification includes a sharding codec that allows for chunks to contain small inner chunks. The format of the resulting binary file format of shards is reminiscent of a HDF5 file. Both HDF5 files and Zarr v3 shards may contain compressed chunks. Furthermore, the Zarr v3 shard specification is similar to the Fixed Array Data Block structure within a HDF5 file. Additionally, the sharding concept is similar to a subset of the HDF5 virutal dataset feature. I will discuss how a standard HDF5 file could be used a Zarr v3 shard and how a set of such files could be used as a Zarr array or assembled into a HDF5 virtual dataset. Finally, I will discuss potential cooperation between Zarr and HDF5 and alternatives to my approach.

As a bonus topic, time permitting, I could discuss how a single file could simultaneously act as a valid TIFF file, HDF5 file, and Zarr v3 shard, taking advantage of cloud optimization of each of these formats.

May we record your session?

Yes

Primary author: KITTISOPIKUL, Mark (Howard Hughes Medical Institute)

Presenter: KITTISOPIKUL, Mark (Howard Hughes Medical Institute)