

## NCZarr Filters -- Proof of Concept

- NCZarr has **Proof-Of-Concept** support for filters/compressors
- **DRAFT: not official**; Require further considerations
- Currently based on the HDF5 model
  - Because C
  - Filters are dynamically loaded from specific directory
  - Filter IDs are integers assigned by HDF5 filter authority
  - Filter parameters consist of a sequence of unsigned integers
- **Tentatively** stored in .zarray in this form
  - `{"id": "320001", "parameters": ["9", "17", "..."]}`

## Standardization Issues

- HDF group is the standard for the numeric ids for its filters
  - Acts as authority to assign numeric ids
- NumCodecs is the official(?) standard for Zarr V2 compressors
  - <https://numcodecs.readthedocs.io/en/stable/index.html>
- As authority for codecs it assigns
  - codec id names
  - codec parameter keys
- HDF5 filter collection overlaps NumCodecs
  - Each has codecs not in the other

## NCZarr Filter Requirements

- Make use of existing HDF5 filter repository unchanged
  - Do not want to have to duplicate
- Support NumCodec JSON definitions where possible
  - versus id+parameter form
- I have to address HDF5 filters not defined by NumCodecs
  - *Someone* must define a JSON representation for those filters
  - Under what authority?
    - HDF and NumCodecs not likely interested
    - Perhaps some kind of appendix to Zarr V2 specification?

## NCZarr Possible solution

- “Compressor” and “Filters” keys in .zarray will
  1. Use NumCodec JSON where available
  2. Use id+parameters JSON if necessary
    - Not compatible with Zarr-Python
- Extend HDF5 dynamic loaded codecs to support both of these APIs over time:
  1. Existing HDF5 filter interface
  2. A new NumCodec-style JSON interface (does not yet exist)