NeXus and PHDF

Mark Könnecke

Paul Scherrer Institute
Switzerland

October 19, 2011

Why PHDF?

- HDF-5 is threadsafe
- But serializes access to data internally
- Cannot be different: a disk can only write a block at a time
- PHDF allows parallell writing to multiple datasets
- Stream data from multiple detectors in parallel to disk

PHDF Prerequisites

- C or Fortran interface
- MPI
- MPI-IO
- Parallell filesystem (GPFS, PVFS, Lustre)

PHDF Features

- Files HDF-5 compatible
- Parallell writing and reading of datasets
- Performance like underlying MPI-IO
- Independent I/O: separate processes write to different datasets
- Collective I/O: separate processes write to same dataset

PHDF Features

- Files HDF-5 compatible
- Parallell writing and reading of datasets
- Performance like underlying MPI-IO
- Independent I/O: separate processes write to different datasets
- Collective I/O: separate processes write to same dataset
- No parallell write for compressed datasets!, reading: YES
    - According to the hdfgroup this is not going to change soon
    - Compression requires chunks, need to be preallocated in HDF-5 data structure

PHDF Features

- Files HDF-5 compatible
- Parallell writing and reading of datasets
- Performance like underlying MPI-IO
- Independent I/O: separate processes write to different datasets
- Collective I/O: separate processes write to same dataset
- No parallell write for compressed datasets!, reading: YES
  - According to the hdfgroup this is not going to change soon
  - Compression requires chunks, need to be preallocated in HDF-5 data structure
- BUT: HDF-5 files can be compressed after writing with h5repack

PHDF Programming

- File/dataset connection/group/attribute management happens collectively
- Only datasets can be read wnd written in parallel
- Need to pass in MPI parameters on opening
- Need to tell dataset if independent or collectively accessed
- Data I/O modes:
  - Contiguous hyperslab
  - Chunked
  - Interleaving data

PHDF & NeXus

- Loss of compression is worrysome
- Do we want this?
- PHDF NeXus driver is a modification of existing HDF-5 driver
- Not much work,development can happen on standard multi core machine
- How to pass in additional MPI parameters?
  - Can be done via special attributes
  - Or additional NAPI calls