

HDF Group report to LLNL
October 2011
Quincey Koziol

Summary:

During the month of October, 2011 the HDF Group worked on the following tasks:

- Metadata Aggregation (94.7 hours)
- Support starting core VFD from file image in memory (80.2 hours)
- Port and test HDF5 on LLNL machines (17.0 hours)
- Page Buffering (11.0 hours)
- Misc. Admin Tasks (7.0 hours)
- Parallel performance benchmark tool (5.7 hours)
- Support “single chunk” indexing method for chunked datasets (4.9 hours)
- Collaborations w/LANL developers about stackable VFD ideas (0.0 hours)
- Investigate and correct issues reported by Klocwork tool (0.0 hours)

The **total number of hours** worked is **220.5** hours.

New tasks:

During this time period the following tasks were begun:

- ***Page Buffering***
 - Create infrastructure to perform page-level caching of HDF5 I/O operations

Completed tasks:

During this time period the following tasks or sub-tasks were completed:

- ***Support starting core VFD from file image in memory***
 - Added Java & Fortran wrapper design to RFC
 - Finished “high-level” API routine coding
- ***Port and test HDF5 on HPC machines***
 - Set up daily tests for both Silo and HDF5 on uDawn system @ LLNL
 - Set up daily test of Silo on HDF Group system

Deferred tasks:

During this time period the following tasks or sub-tasks were deferred:

- *none*

Tasks in progress:

During this period of time The HDF Group worked on the following tasks:

- ***Metadata Aggregation, John Mainzer, Quincey Koziol*** (94.7 hours)
 - Lots of design discussions
 - Investigation of library's free space allocation code
 - Writing and revising RFC for feature
- ***Support starting core VFD from file image in memory, John Mainzer, Christian Chilan, Peter Cao, Scot Breitenfeld*** (80.2 hours)
 - Added Fortran and Java API wrapper design to RFC
 - Minor revisions on RFC
 - Finished "high-level" API routine coding
 - Testing high-level API routines for feature
 - Code review and revisions
- ***Port and test HDF5 on HPC machines, Albert Cheng*** (17.0 hours)
 - Set up daily tests for both HDF5 and Silo on uDawn @ LLNL
 - Set up daily test of Silo on HDF Group Linux system
 - Test HDF5-1.8.8 pre-release on INL system and aztec @ LLNL
- ***Page Buffering, Jacob Gruber*** (11.0 hours)
 - Investigate current behavior of HDF5 VFD code
 - Prototyping new page buffering VFD
- ***Miscellaneous Admin Tasks, Ruth Ayd, Quincey Koziol, Albert Cheng, John Mainzer*** (7.0 hours)
 - Set up user accounts
 - Planning and reporting activities.
 - User discussions, status telecons & e-mail.
 - Make snapshots, etc.
- ***Parallel performance benchmark tool, Ruth Ayd, Quincey Koziol*** (5.7 hours)
 - Reviewed h5perf tool's features, to include in parallel performance tool comparisons, etc
 - Design discussions for new parallel performance tool
- ***Support "single chunk" indexing method for chunked datasets, Vailin Choi*** (4.9 hours)
 - Revisions after code review

- Send back for 2nd code review
- ***Collaborations w/LANL developers about stackable VFD ideas, - (0.0 hours)***
 - *Nothing to report during this time period*
- ***Investigate and correct issues reported by Klocwork tool, - (0.0 hours)***
 - *Nothing to report during this time period*

Current Projects for People:

- Quincey Koziol:
 - Design & architecture guidance
 - Project management
- Albert Cheng:
 - Port and test HDF5 on HPC machines
- Vailin Choi:
 - “Single chunk” chunked dataset indexing method
- John Mainzer:
 - Metadata aggregation and Page buffering design
 - “stackable” VFD design, implementation and collaborations w/LANL developers
 - Investigate issues reported by Klocwork tool
 - Design VFDs to enable poor man’s parallel I/O
- Ruth Aydt:
 - Parallel performance benchmarking tool
- Christian Chilan
 - Enable starting “core” VFD from existing buffer: implement “load image” feature, from RFC
- Jacob Gruber
 - Prototype page buffering implementation

Ongoing tasks for next reporting period:

- ***Enable starting “core” VFD from file image, John Mainzer, Christian Chilan***
 - Implement feature.
 - Code reviews.
- ***Single Chunk Index Method for Chunked Datasets, Vailin Choi***
 - Second review.
- ***Parallel performance benchmark tool, Ruth Aydt***
 - Gather requirements, use cases and goals of project
 - Write RFC describing new tool
 - Implement tool.

- ***Metadata Aggregation, John Mainzer***
 - Gather requirements, use cases and goals of project
 - Write RFC describing new feature
 - Implement feature.

- ***Page Buffering, John Mainzer***
 - Gather requirements, use cases and goals of project
 - Write RFC describing new feature
 - Implement feature.

- ***Port and test HDF5 on LLNL machines, Albert Cheng***
 - Stand up daily testing on LLNL machines.
 - Investigate and add tests for “poor man’s parallel” I/O to HDF5 regression test suite.

- ***Investigate and correct issues reported by Klocwork tool, John Mainzer***
 - Investigate issues reported by Klocwork and correct them.

Deferred/Future tasks:

- ***Scope effort for implementing “stackable” VFDs***
 - Discuss feature and write RFC for allowing VFDs to be “stacked” on top of each other.

- ***Design VFDs to enable poor man’s parallel I/O***
 - Discuss feature and write RFC for VFDs that can improve “Poor Man’s Parallel” I/O on HPC systems.