2019-01-07 Meeting notes

Date

07 Jan 2019

Meeting Info

1. Please join my meeting.
   https://global.gotomeeting.com/join/840709413

2. Use your microphone and speakers (VoIP) - a headset is recommended. Or, call in using your telephone.
   Dial +1 (224) 501-3412
   Access Code: 840-709-413
   Audio PIN: Shown after joining the meeting
   Meeting ID: 840-709-413

Attendees

- John Mainzer
- Elena Pourmal
- Mark Miller (LLNL)
- Suren Byna (LBNL)
- Robert B. Ross (ANL)

Goals

- Go over agenda items

Discussion items

<table>
<thead>
<tr>
<th>Time</th>
<th>Item</th>
<th>Who</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 min</td>
<td>Updates to Charter</td>
<td>Elena</td>
<td>HDF5 Technical Advisory Board Charter</td>
</tr>
<tr>
<td>10 min</td>
<td>Prabhat's feedback on strengths and weaknesses</td>
<td>All</td>
<td>HDF5 Software Strengths and Weaknesses</td>
</tr>
</tbody>
</table>
From Charter: *At the beginning of the calendar year the Board will identify goals and objectives to be achieved during the year. These goals will be available to the HDF5 community. By the end of the year the Board will publish a report on its activities and achievements.*

**Emerged motives from the previous discussions of strengths and weaknesses:**

- A set of large-scale use cases to understand performance issues and develop necessary improvements and tactics to bring HDF5 up to par with these other libraries (ADIOS).
  - [http://conferences.computer.org/scw/2018/pdfs/PDSW-DISC502018-7eAhkUBB2dWhq2lmKsykIN/5DJPALrRZcBy](http://conferences.computer.org/scw/2018/pdfs/PDSW-DISC502018-7eAhkUBB2dWhq2lmKsykIN/5DJPALrRZcBy)
  - Part of SciDac is an effort to create benchmarks

  > Elena will look at the paper and coordinate with the existing projects; report at the next TAB meeting.

- How well benchmarks are accepted by large community? Maybe the first thing is to evaluate and if they are appropriate.
- If the goal is how to use correctly then this is not the right approach.
- Benchmarks are baseline for performance on different systems, understanding the performance and having a roadmap.
- Not use word “benchmark”.

- **Discussion of different motives/examples users can chose from to avoid misuse of HDF5 library.**
  - Approach this one first and then come back to the first one.
  - Have several examples - particle example, etc., intentions of I/O operations are, and codes to demonstrate how to do it.

Identify examples to cover for this year, which examples make sense to cover. Shared file I/O, multiple files I/O - two different approaches to performance of HDF5 on top of DAOS. There were several synthetic codes - legion, climate code. We need a balance of the set of use cases looks like.

Another dimension: do we want the use cases to focus on callers of HDF5 directly, or something like CGNS and netCDF-4. Do they use HDF5 correctly or there are weaknesses in HDF5?

- Functionality for both reads and writes; search - data and metadata, query interfaces.
- Capability should be performant.
- Examples should be done at the workflow level. Emphasize schema and constraints.

> Come up with the use case example and a description of schema, storage

> Put on Confluence.