2018-06-07 Meeting notes

Date

07 Jun 2018

Recording

Attendees

THG
- Elena Pourmal
- John Mainzer
- David Pearah
- Tom Brady
- Ann Johnson
- Mark Miller (LLNL)
- Mr. Prabhat (NERSC)
- Suren Byna (LBNL)
- Robert B. Ross (ANL)
- Venkat Vishwanath (ANL)
- Ward Fisher (Unidata)
- Quincey Koziol (NERSC)
- Lee Ward (Sandia)

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>Welcome</td>
<td>David Pearah</td>
<td>See presentation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Elena Pourmal</td>
<td></td>
</tr>
<tr>
<td>10 min</td>
<td>State of HDF5</td>
<td>Elena Pourmal</td>
<td>2018-06-07-HDF5 status.pptx</td>
</tr>
<tr>
<td>10 min</td>
<td>Members introductions</td>
<td>Why are you interested in being on TAB? See notes below.</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------</td>
<td>---------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2018-06-07 16.03 HDF5 Technical Advisory Committee Meeting.mp4 (Introductions start around minute 20)</td>
<td></td>
</tr>
<tr>
<td>30 min</td>
<td>Discussion of governance</td>
<td>All</td>
<td>We didn't have time and postponed it for the next meeting.</td>
</tr>
</tbody>
</table>
| 5 min  | Closing remarks (next steps) | Elena Pourmal and All | • Discuss governance  
• Identify discussion topics |

### Action items

- Venkat Vishwanath will recommend someone from Genomics community.

### Notes (summary):

**Lee:** works in storage; hopes to see a sustainable library that addresses Labs needs, specifically in the HPC area; wants to see sustainable maintenance effort and have dedicated THG resources; contribute complimentary improvements.

**Mark:** supports IO libraries SILO, new library Conduit, and multiple applications based on HDF5; would like to help HDF5 code base get to the point where it has broader community adoption and development, large number of developers who can contribute and evolve the code in that direction; could contribute wisdom and people, summer projects, small projects that full time people at the Labs do as part of their orientation.

**Prabhat:** Works with data stack and libraries that use HDF5; hopes to coordinate other DOE and DOD labs as well as create stability of HDF5 in terms of personal, remove dependence on Quincey, wants to make sure that there is a mitigation strategy in place. Contributions: LBNL will continue contribute to code base (there will be a new hire for ExaHDF5 project), summer, interns will be able to help too.

**Quincey:** Involved in HDF5 leadership, is an architect of new pieces and components; at NERSC started new development group; interested passing knowledge via working on new code, creating documentation, and training new people.

**Rob:** Started to play with HDF5 around 2000, got a performance paper in 2001; in general, interested and advocating for applications to use high-level libraries; HDF5 is a great example, and Rob wants to see it succeed. Contributions: Rob's team built 3-4 packages that had seen some success, and there are some things that can be applicable to HDF5 project too. Focus on building software infrastructure that is accessible; ANL team thought a lot about how to make it a little bit easier for people to have their head around a package and move faster.

**Suren:** Interested in I/O research; gets good performance from HDF5, leads projects ASCR and ECP project that develop a lot of features for HDF5 and productize them (VOL). New EOD project focuses on prototyping new features for HDF5. Quincey works closely with Suren is training and brings a lot of expertise, especially in developing VOL plugins that allow to extend HDF5 features and talk to different storage.

**Venkat** – leads data science group at ANL, long time HDF user, used HDF4; would like to have design documentation for HDF5, hopes that will finally happen. Part of this will have more community involvement, use of HDF5 by applications on their system. Paul Kaufmann already contributes. ANL sees a lot of HDF5 uses in non-HPC applications, for example, in genomics. It would be good to invite non-HPC members to TAB; want us to consider this suggestion since it would be good to have a variety of perspectives to consider.

**Ward:** Represents UNIDATA, netCDF4 lead, represents netCDF community, netCDF4 developers would like to stay abreast of HDF5 development. UNIDATA was very successful in building contributing community, and this is where Ward can contribute.

**Mark:** favors of idea having Advisory Board that represents diversity of needs. Is a good idea.

**Rob** – good representation of HPC; supports to have members from other communities.

**Mark** – to what extend the Technical Advisory Board affect CE vs ESE of HD5, is advisory Board advising long term planning in both? Elena explained that all new features will be in CE and than in ESE, ESE is driven by CE. Frequency of the CE releases is a concern, technical advisory Board can discuss the issue.

**John** – gather questions from the Board, how technical advisory Board wants to get go forward.

Some members represent organizations that have active contracts; We agreed not to dilute TAB discussions in deep technical detail, we have other venues to do that, but HPC representatives can bring wisdom in respect with planning HPC relevant features, how they are fostered and not negatively impacted by other activities that are of interest to other HPC users.

**John:** major things such as adding support for sparse datasets, some people Prabhat will be interested.

Suren: TAB can help to prioritize which features can be useful, and bring in develop in a timely manner.

Elena need to talk about it Development process; how current projects affect the library and alignment of HPC features that are coming.

Mark likes an idea of are major activities are going on. Even the sequence of things – will help.

Rob suggested to have 2-3 questions to steer specific conversations, how we prioritize on feature integration, how we encourage students to work with the code, some folks will be more interested and we will have more focused conversation in a timely manner.