

Reponses / comments on 2018 SCAP Panel Recommendations

Updated 01/17/2021 – 16:15 CST

Dear SCAP Panel

These responses have been updated January 2021. Due to the length of time between SCAP meetings, now almost 3 years, we are appending the updates to the original responses to provide additional context, also due to the large fraction of new members to the advisory panel. New material added will be indicated with a solid vertical bar at the left of the text:

| New material from January 2021 is indicated as such.

On behalf of the IceCube Collaboration and the ICNO management, I again thank you for the time you put into providing critical feedback and suggestions for improvement. I am providing responses to your recommendations below.

- 2018-1** IceCube high level management is strongly urged to review the organizational structure of the M&O software and computing domain. In particular, management is invited to name a manager dedicated to maintain and execute an overall, global vision of this area (a “Global Computing Coordinator” reporting to the IceCube Neutrino Observatory Director of Operations.)
- 2018-2** IceCube high level management is strongly urged to review the functions of the ICC committee pertaining to the coordination between computing and analysis. In particular, management is invited to define a small, dedicated coordination group, co-chaired by the Global Computing Coordinator and the Analysis Coordinator, which would be responsible for preparing issues and recommend actions and priorities, to be brought for review and approval by the ICC.
- 2018-3** IceCube high level management is strongly urged to empower the Global Computing Coordinator with the ability to dialog with responsables of collaborating institutions in order to harvest additional resources for the software and computing domain. This should be done through the definition of specific work items to be accomplished within well-defined periods of time, in a spirit similar – but not limited to – the way the Software area has been recently refocused.

To these first three points we announce the hiring of Dr. Benedikt Riedel who began working again in IceCube in December 2018. Prior to that Benedikt spent many years working in OSG at UChicago. His vast experience with massive distributed computing and his familiarity with IceCube (he graduated with

his PhD from UW-Madison on IceCube) make him an ideal leader for IceCube Computing. He is replacing the vacant position of Gonzalo Merino who left to become Director of the Port d'Informacio Cientifica in Barcelona. The first discussion of the expanded scope of responsibility of the position with the appointment of Benedikt took place at the last IceCube Collaboration meeting and will be put before the ICB during the upcoming meeting.

We have reassessed these recommendations following significant evolution of the IceCube computing needs and with the successful integration of the working group technical leads who have brought to the ICC the particular needs of IceCube's diverse scientific program. Because of the broad scientific scope of the working groups that range from cosmic ray science to low-energy neutrinos from supernovae to identification of transient and continuously emitting neutrino point sources, we feel the LIGO-like top-down model proposed in 2018 where the computing coordinator and analysis coordinator interact at a high level would not be as effective as a bottom-up approach. The WG tech leads serve the role of the dedicated group referenced in the 2nd recommendation. We feel that the ICC has been effectively using these inputs in managing the IceCube common resources. These improvements in the coordination structures are complemented by recent and upcoming roll-outs of processes and tools to aid the coordinators: a formalized simulation request process; web-based tools for resource pledging, matching, and tracking; processes for internal data releases.

2018-4 The IceProd2 team is urged to complete the implementation of multi-user features, which should enable the inclusion of collaboration-wide activities. A wider use of an IceCube AAI framework should be considered, preferably in the direction of a single-sign-on for as many IceCube resources and services as possible.

We are evaluating several frameworks that allow SSO (single sign-on) and authentication and authorization. The current frontrunner is [COmanage](#). We are also working on a new authorization scheme using JSON Web Tokens and SciTokens.

IceProd2 multi-user has been completed and is currently being used by some WGs. Expansion of the user-base to provide the benefits to all users is planned but requires effort to improve usability.

2018-5 Architectural and Technical frameworks should be defined to consolidate data management and metadata related activities. Evaluation of software products to support the architectural framework should be performed as focused, time limited activities.

Some progress has been made indexing metadata for the experimental and simulation data (File Catalog). Progress of insertion of data into LTA will be presented.

See also response to 7.

2018-6 The focused action scheme of the Software area should be brought into steady-state operation.

Several remedial actions to address these points are:

- Biweekly calls that happen even if there are no agenda items
- Themed code sprints: "Simulation Production", "Reconstruction Performance", Machine Learning Frameworks"

- Releases have a due date and are released on that date
- Monthly reviews of outstanding pull requests on calls
- Release cycle as a deadline
- Coding camps during Summer

Biweekly Calls - Happen fairly regularly (15 calls in 2020 - most in last 5 years). Cancellations due to lack of agenda items, holidays, and collaboration meetings. Also included public call for agenda items one week prior and collected through the week. Automatic slack reminders sent to the #software channel.

Code Sprints - Theme has been ticket resolution over the last year. Historically low number of open tickets. Future plans to track and report critical ticket lifetimes in quarterly reports.

Releases - Committed to seasonal releases. 4 seasonal releases in 2020 with subsequent incremental releases. (2019? I think I remember the first full 'combo' release taking longer than expected, but we had two incremental simulation releases that year.)

PR Reviews - Migrating to GitHub is still in-progress. We reached a significant milestone on the development model and policy. Working on a timely PR model, which will be a challenge for us - basing this on M&O L3s and WG Tech Leads.

Bootcamps - Alternating Strike Team leadership between Olivas, Diaz Velez, Meagher, Schultz, and La Dieu. Each leader will be in charge of the seasonal bootcamp topic. Each leader signs up for one season. C++ refresher bootcamp in the works for Winter 2021. All collaboration members are welcome to join.

2018-7 Science reproducibility and public data releases should be considered different aspects of a more global Data Management and Preservation framework. An end-to-end architecture, from DAQ to public data releases, should be arrived at, possibly in incremental steps which are coherently orchestrated. This should be coordinated with metadata-related activities.

A re: DOMA: this is a large, multi-year project because we need to create software and most likely deploy hardware. It includes several areas: I3Live, JADE, Long Term Archival, Data Center Infrastructure, SNDAQ, etc. We need to review the current DOMA strategy from pole to public data release. Data transfer from pole to Madison is on solid footing. The biggest issues are:

- Various different data sources that seemingly don't talk to each other, e.g. Is there physics in the I3Live data? How users interact with I3Live data beyond monitoring shifts? etc.
- How do people access data?
- How can we move away from POSIX access?
- What to do about the file catalog
- How can we design data to be more accessible and better organized?
- How and when should we publish data?

RE: Reproducibility: Internally discussions have started for a grant proposal focused on reproducing IceCube results with new software and/or knowledge – Codename: Continuous Science. Question is how to staff and fund. We are considering [this CSSI solicitation](#). The main goal is to make sure large (for start) changes to the codebase trigger a series of tests, such as redoing the HESE analysis, that facilitate high level checks, similar to sanity checkers.

The ICC has introduced a set of general guidelines to establish the IceCube internal data repository, where data analyses are properly documented to insure reproducibility of the published results. The guidelines introduced the concept of documenting and storing the final event sample used in specific analyses (including simulation data), in revisioning control the software used in the event selection and/or analysis, and in properly documenting the analysis stages and the software usage to reach the results (e.g., a set of final physical quantities and graphs). The next step will be to provide adequate tools to make documentation simpler, enable re-analysis, and make new analyses on legacy data seamless.

2018-8 IceCube high level management should take note that fully establishing and maintaining these policies requires non-negligible human resources which are currently not identified.

As WIPAC embarks on future facility enhancements additional resources do become available. We are exploring the possibility to hire an additional developer and are considering experience with DOMA as a requirement.

The beginning of the IceCube Upgrade project has *increased* the overall support requirements for ICNO without providing backing funding. The proposal to operate the IceCube Neutrino Observatory from 2021-2026 proposal seeks support for 5 additional FTE (actually 4.05 averaged over the 5 years) to provide support with the additional context of the Upgrade. These position are:

Workflow programmer : develop middleware to allow IceCube software to exploit more platforms

Simulation programmer: maintain existing physics and detector simulation codes (i.e. PMT digitization) and implement/support new Upgrade features. This also helps address simulation software efficiency - a noted bottleneck.

Visualization programmer: something as central as the IceCube event viewer has been without identified maintainer. We ask support for maintaining this tool as well as other visualization tools (i.e. web visualization)

Reconstruction support: support / implementation of new reconstruction tools, especially ML based. Again maintain current with eye on software efficiency.

GPU programmer: dedicated support for the photon tracking is needed in particular the specialized nature of skills here demand a dedicated professional to maintain.

2018-9 Efforts should continue in a highly focused manner in order to maintain workflows which can run efficiently on systems where IceCube can request resources. This requires work on the workflows themselves, but also on monitoring and job scheduling and on the handling of intermediate results.

IceProd2 is now stable and working well in production. We are adding new features including IceProd2 “campaign mode” to take advantage of supercomputer sites. In campaign mode a single dataset or large portions of a dataset would be produced at a single site.

I think we can claim victory on the “campaign mode” right (viz Cloudburst)? Human resources remain short, however, with the progress made to expand the usability of IceProd2 we now focus available effort on: (1) improving efficiency in utilization of GPU and memory through resource prediction; (2) addressing remaining scheduler issues.

2018-10 IceCube is encouraged to request time from research computing providers, such as OSG, XSEDE, supercomputers, etc. with a target of achieving within 18 months a computing power about 10 times higher than currently available. This should be done through WIPAC and through IceCube collaborating institutions around the globe, and the applications should receive maximum support from their principal investigators. The M&O team should be the catalyst that puts these resources to the best use for IceCube. This requires an immediate corresponding effort to scale the computing and data analysis infrastructure to be able to efficiently and robustly handle such an order of magnitude increase in resources.

In the US we have applied for and been awarded cloud credits through ECAS (Exploring Clouds for Acceleration of Science) and a Mid-Scale Research Infrastructure-1 proposal with the cryoEM group at UCSD for a GPU cluster hosted at SDSC. Allocations on existing and upcoming XSEDE resources, e.g. Frontera, are being pursued. In the the EU we are discussing with EU PIs the possibility to apply for EU computing programs such PRACE.

In November 2019, IceCube and SDSC completed the [Cloudburst Experiment](#) harnessing over 50,000 GPUs simultaneously to produce IceCube simulated data. This represents a successful scaling of IceCube simulation production by a factor of over 50x.

Kael HANSON
Director of Operations, IceCube Neutrino Observatory