

Skimmer Discussion October 2008

Attendees: David Chamont, Richard Dubois, Tom Glanzman, Heather Kelly, Leon Rochester, Tracy Usher

Current Status

The current production release is v7r1. David does encourage users to try this latest version, it makes it that much easier to address problems.

David has migrated the skimmer to a new Subversion repository. Bug tracking will now take place through the TRAC server. David will send mail explaining how Tom, Tony, and Heather can submit tickets to the TRAC server. This is also the location of the current user guide. See:

<http://lir.in2p3.fr/trac/tskim>

David will consider putting the [FERMI extension to the user guide](#) in our Confluence area, rather than in the TRAC web page, but wonder if it is such a good idea not to keep all documentation together.

Please note that the v7r1p1 tag is a non-Fermi patch to address generalizing the skimmer.

Reported Problems and Priorities

Francesco has reported trouble with skimming many (~10k) small files. David asks how typical this use case is and what priority he should assign this issue. Tom responds that this tends to be typical for MC jobs, and may become more typical as the number of available data runs increases. Tom also notes, that we may be able to work around this for now, by merging these small files before submitting them for skimming.

Leon has reported trouble concerning environment variables that appear to remain set even after a skim job. David explained that the few env variables that the skimmer sets on its own are created in a subshell which should not affect the users environment. Leon notes that this issue does not occur in his batch jobs...rather it is in interactive calls to the skimmer that notes his environment remains updated. David notes that it depends on how one is invoking the skimmer... the user environment could be impacted if the eventual script which prepare the variable is sourced instead of executed in a subshell. It is probably a general unix shell issue, rather than something specific to the skimmer. Leon says he will investigate further.

Leon's second issue concerns skimming periodic triggers for overlays, which are rather large (2- 3 GB) files. Leon attempted to run with `SK_REBUILD_INDEX = false`, to avoid rebuilding indices, in the hopes that was causing the problem. This had no effect, the job still fails due to excessive memory use. Leon does not see how to work around this problem, and feels that a fix needs to be found soon.

We looked briefly at the current list of JIRA issues. We do need to talk to Tony to determine priorities. David plans to migrate these JIRAs to the new TRAC server. He asks that all future problems be submitted as tickets to the TRAC server. He will send email detailing how that works.

Tom also mentioned our desire to upgrade to ROOT v5.20.00. This will likely happen in the next few weeks. David stated that his first priority will be to test the new ROOT v5.20.00. After our above discussion, is sounded as though Leon's bug report resulting in large memory consumption might take priority over Francesco's trouble with many small files. [editor's note: Tony later mentioned that he is receiving many complaints concerning Francesco's reported bug - so perhaps we need to revisit the priorities as David migrates the JIRA issues to TRAC]

Where To From Here

Tom suggested we should talk to Tony about using the v7r1 version of the skimmer in the web front end. David reports that the latest skimmer is backwards compatible with the v6 series, so that the old environment variables will still work - however we are encouraged to migrate to the new `TK_*` environment variables as soon as possible. Tom asked about the skimmer wrapper, and David did say we should continue to use the usual skimmer executable, which now sets the defaults as typically expected by Fermi users.

Tom broached the issue of how much time David could hope to spend on the skimmer in the future. David stated that he is still available part of his time. He can meet with us from time to time, though he is not actively following the GLAST news and issues. David also encouraged us to use the new TRAC server for bug reports. Tom and Richard reiterated our appreciation for David's work on the skimmer and hope he can continue to contribute in the future.

Concerning SVN write access, Heather wondered if some subset of GLAST developers could be allowed to make branches and tags to address emergency issues concerning the skimmer. David would like to avoid branches, if possible. He also wants to make sure we know how to use subversion, as it is a little different from cvs. He offered to communicate to Heather the necessary commands if we feel that is necessary. Alternatively, we do have the skimmer source and can modify that as necessary and communicate the changes to David for inclusion in the repository. Heather stated that is fine, though there is a concern of allowing our copy of the skimmer to diverge from the one that David supports.

David reiterated the potential issue concerning the event ids being Long64. The ROOT indices do not support keys > 31 bits. The skimmer relies heavily upon the native ROOT index class. We do need to keep an eye out for the day when our event ids are truly too large. There has been no promise from the ROOT development team that their `TVirtualIndex` will be updated to handle 64 bit keys.

Action Items

1. Leon will try the skimmer v7r1 to see if that fixes the memory issues he has been experiencing with skimming periodic trigger runs.
2. Tom, Leon and Heather will talk to Tony concerning the version of the skimmer used in the web front end, as well as the outstanding list of JIRAs to determine priorities.
3. David will look over and migrate the JIRA issues to the TRAC server.
4. David will send mail to Tom, Tony and Heather detailing how to submit tickets to the TRAC server

<http://llr.in2p3.fr/trac/tskim>

5. David and Heather will discuss and practice accessing the SVN server so that emergency branches and tags can be made if absolutely necessary.
6. David will investigate whether the skimmer works with ROOT v5.20.00, so that we may proceed with our planned upgrade to ROOT.