Database Meeting 2023-01-04

Date

04 Jan 2023

Attendees

- Igor Gaponenko Colin Slater Fritz Mueller Andy Salnikov Fabrice Jammes

Notes from the previous meeting

- Database Meeting 2022-12-14

Discussion items

Discussed	Topic	Notes
✓	Project news	Fritz Mueller: • All green work is classified (within FPD at least) as "administrative work" and is cleared to continue. • (for Fritz Mueller , Igor Gaponenko , John Gates and Kian-Tat Lim) Office move to an upstairs room in B048 is scheduled on January 26th (Thursday). Please plan appropriately for the site visit(s) to pack up your offices.
✓	Investigating worker lockups (Note: John Gates wasn't present during the discussion)	Context: • continue discussing the topic (see Database Meeting 2022-12-14) • improvements to the worker-side Qserv monitoring have been made to help with the investigation: ○ ⚠ DM-36965 - Jira project doesn't exist or you don't have permission to view it. • the ongoing investigation of the lockup using the Qserv instance at SLAC (`slac6`) : ○ ⚠ DM-37334 - Jira project doesn't exist or you don't have permission to view it. • summary of findings (so far): ○ the problem seems to be triggered (for some queries) `czar` by connection timeouts to the MySQL server when merging results. This has a ripple effect on Qserv's ability to process the affected query (as well as subsequent queries). More details can be found in a conversation at: https://lsstc.slack.com/archives/G2JPZ3GC8/p1671556554242149 Further steps: • Monitoring the status of the MySQL connections at `czar` using `SHOW PROCESSLIST` to see if there were any outstanding queries/locks before the event? • (Perhaps) reconfiguring MySQL to allow longer timeouts? • TBC...
✓	Status of `qserv-operator` and `qserv-ingest`	Context: • migrating the ingest application to the latest version of the ingest API: The workflow developer's guide for the Qserv Ingest system: The version history of the Ingest API • the CI randomly fails as reported in: https://github.com/lsst-dm/qserv-ingest/actions Next steps: • Fabrice Jammes and Igor Gaponenko will investigate the issue by running the same ingest workflow against `qserv-dev` (of IDF) as it's used in the CI

✓	Progress on APDB	Andy Salnikov: • made progress in implementing a mechanism for transporting data between APDB and PPDB (only the APDB side to produce results to be moved to another side) • the next step will be to have the full implementation • (problems, concerns) with very little input/feedback from the pipeline folks on this subject ○ Fritz Mueller is also concerned about this. An effort to establish better coordination in this development will have to be made.
✓	Performance of the partitioned InnoDB tables in MySQL version 8.	Context: • Investigating an option of replacing MariaDB with MySQL and switching from MyISAM to InnoDB. • Findings have been reported in Experiments with partitioned InnoDB tables in the MySQL server version 8 Fritz Mueller: • we need to test the performance of the SELECT queries against InnoDB to see how it would compare against MyISAM. A concern here is the performance of the full-scan queries. • we may consider the partial migration of Qserv by switching the `czar` database from MariaDB to MySQL while leaving MariaDB at worker databases. Igor Gaponenko: • there is still one important use case where we could benefit from relying on the MySQL-partitioned InnoDB. A problem is that the current implementation of MariaDB doesn't support parallel ingest into partitions of the MyISAM tables. The ingest streams get serialized during ingest. This limits the overall performance of the ingest at a level of fewer than **100 MB/s** even on the high-performance filesystem (based NVMe). This wouldn't be a problem if we had many chank tables to ingest into (in the extreme case we might have up to 150k such tables). However, there might be a case when the input data would be arriving from the LSST Pipeline in the spatially-constrained batches. That would reduce the total number of chunks in each such batch to some small number, thus preventing the Ingest system to reach the desired (high) level of parallelism.

Action items

□