MariaDB + Qserv Meeting 2016-03-01

Date

01 Mar 2016

Attendees

Andy Salnikov, Brian Van Klaveren, Fritz Mueller, John Gates, Jacek Becla, Kian-Tat Lim, Serge Monkewitz, Vaikunth Thukral, Fabrice Jammes, Monty Widenius

Goals

- Technical Q&A with MariaDD team
- Get implementation recommendations
- Get insight into scheduling of relevant upcoming features

Discussion items

MaxScale

- Feeback provided Qserv->MariaDB re. MaxScale: not much documentation, lots of implementation work to do compared to mysqlproxy+LUA for our use case.
- mysqlproxy drawbacks are single-threadedness, no decent query parser available, lack of support
- Monty agrees that if scaling perf. is not an issue (for Qserv use case it is not) then using MaxScale may be more work than we need to do.
- MariaDB is currently looking to replace the parser in MaxScale with a more modern / lighter-weight parser. Timeframe to choose direction here is ~2mo. Monty interested in hearing about parsers we are considering, and what our technical needs from a parser might be.

GIS

- GIS point indexing on 2D plane and sphere are implemented for AriaDB; will be implemented for InnoDB. Performance is beating postgres in preliminary tests.
- Near-neighbor query for 2D plane is implemented; for 2D sphere postponed (but we could get it back on the plate for $).
- Indexable 3D Cartesian is not currently implemented (but we could get on plate for $).

MyISAM Decompression

- Currently done redundantly each time a row is accessed.
- Monty asserts decompression overhead will be very small compared to block I/O.
- But Qserv full-table scans will only read once, access many?
- Monty asserts yes more CPU, but overall memory savings will help.
- We'll need a refresh on benchmarks to be sure; last set was done 3 or 4 years ago.
- Feature is used mostly by users archiving tables at this point; has been stable code for many years.

Secondary Index

- Monty recommends trying a single table, but beware error recovery times. If single table runs into problems, would next try breaking into ~16 tables (along the lines of the experimental design we have already been investigating).
- Monty recommends we look into Fusion io cards from SanDisk for hosting this table (like NVM express, but includes file system and compression, ~16x spinning disk perf. and ~20% cost increase).

Result Table Aggregation

- MariaDB currently working with third-party storage subsystem vendors on spec-ing their own clustered shard + aggregate system.
- ~1mo. until partners are announced and architecture could be discussed with us.

Auth*n

- Monty says no off-the-shelf oauth/openid authentication plugin is available, but should be easy to write our own (1 or 2 FTE-days) and would be happy to provide support for this effort.

Connection Timeouts

- Most/all clients have a configurable client-side timeout (called connect or query timeout) and we should investigate our settings/defaults here.
- Server has configurable timeout as well for reaping idle connections.
- There is no server-side timeout for running queries.

Monitoring

- Enterprise aggreement includes monitoring tools from MONyog and SeveralNines (third-party, closed-source).
- MariaDB working on their own open-source monitoring solution, but probably 1-2yrs away.

Vertical Partition Joins

- Monty recommends sizing up the join buffer
- Currently MariaDB makes no assumptions that data in multi-table joins is sorted by primary key.
- Possible to work to add a merge-join solution, but would need a detailed use-case to consider.

Misc.

- Jacek to have follow-up meeting with Rasmus Johansson to figure out contract mechanics so MariaDB team can be paid for needed consulting /development work.