

Database Meeting 2022-02-02

Date

02 Feb 2022

Attendees

- Igor Gaponenko Fritz Mueller John Gates Unknown User (npease) Andy Hanushevsky Andy Salnikov Fabrice Jammes Colin Slater

Goals

- See a list of topics below

Discussion items

Time	Item	Who	Notes
	Project news	Fritz Mueller	<ul style="list-style-type: none">no major news so far
	Progress on topics discussed at the previous meeting Database Meeting 2022-01-26		<p>Qserv readiness for DP0.2:</p> <ul style="list-style-type: none">Fritz Mueller: the operator is ready to be tested in IDF using int. We need to do the MariaDB migration as well.Igor Gaponenko will try loading a slice of the catalog into the small cluster at NCSA. <p>Fritz Mueller's proposal for the "Grand Unification" of the relevant Git packages (qserv, qserv-operator, etc.). Are we going for this?</p> <ul style="list-style-type: none">Andy Salnikov: it works for the LHC/ATLAS projectFritz Mueller: make qserv_testdata (*300 MB*) a submodule of qserv? This has non-trivial implications. Or, better, include qserv_testdata directly into qserv?Igor Gaponenko: what if we replace the existing data with the generated data produced before running the integration test:<ul style="list-style-type: none">Fritz Mueller: this has <i>pros</i> (extend the coverage of the integration tests) and <i>cons</i>, such as losing reproducibility.Unknown User (npease): same concernsColin Slater: (on Zoom) "The data generator is deterministic"Fabrice Jammes: submodules are complicated.Igor Gaponenko: we should merge qserv_web (Qserv Web Dashboard) into the source tree of qserv. Fritz Mueller - are you going to do this?Fritz Mueller Adopted! <p>Fritz Mueller's proposal to move the documentation tools into the Qserv build container</p> <ul style="list-style-type: none">Fritz Mueller: Keep using Sphinx. Need help with making it run in the Qserv build container.Igor Gaponenko: we need a coordinated effort to design the documentation<ul style="list-style-type: none">Fritz Mueller: will create a "straw man" proposal to be discussed at the next meetingIgor Gaponenko: the next meeting will be dedicated to discussing the documentation

	Mike Reed's idea of dealing with the very wide Object (alike) tables	Igor Gaponenko	<p>The context:</p> <ul style="list-style-type: none"> see a discussion at the Slack channel https://lsstc.slack.com/archives/C2B70NXFD Mike is proposing to split the wide "Object" table (a few thousand columns) into a very narrow director table (objectid, ra, and dec) and split the remaining columns between the linked (band-specific) child tables <p>Questions:</p> <ul style="list-style-type: none"> what performance implications would this have? <p>Fritz Mueller: this idea has been around for a while and was known as "Object-lite". That table could be stored in a conventional database. Some folks are still interested in it (because Qserv is expensive and complex to set up and maintain).</p> <p>Colin Slater: no actual effort has been made yet. Though, the idea is known.</p> <p>Igor Gaponenko: Qserv may treat queries with table JOINs differently, which may affect the performance. We know that we had this issue in the past after fixes made in the Linux kernel to address a problem of https://meltownattack.com/. These fixes resulted in a dramatic drop of the Qserv performance ("KPM30") for the JOIN queries.</p> <ul style="list-style-type: none"> Fritz Mueller: KPM50 is meant to help here with testing the effects. <ul style="list-style-type: none"> Igor Gaponenko should resume working on KPM50 later this year while we still have KPM50 at NCSA. John Gates: the problem was seen in the query mixes (queries with JOIN and shared scan queries). Igor Gaponenko: still need to test it at a lower-level (directly in MariaDB) to see where the performance "cliff" is happening Colin Slater optimizing the Object table won't gain much as most of the slowness comes from reading (disk I/O) the large tables (Source and ForceSource) Fritz Mueller: there is no definitive answer here. A lot depends on a specific use case. Action items: 1) encourage experiments made by Qserv users, 2) test this by ourselves, 3) feed suggestions/recommendations to users.
	What shall we do about existing Qserv catalogs at NCSA?	Igor Gaponenko	<p>The context:</p> <ul style="list-style-type: none"> We're expected to vacate NCSA by the end of FY22 GPFS is full at NCSA A number of the large-scale (and still very useful) catalogs (SDSS_Stripe82, wise, gaia dr2, kpm50) exists at NCSA in two forms: <ul style="list-style-type: none"> Within Qserv Raw and partitioned (ready to be ingested) files at GPFS (~300 TB*) (except kpm50) <p>Possible solutions:</p> <ul style="list-style-type: none"> delete files from GPFS and use Qserv as a "source of truth" should we still need those catalogs past FY22. The Replication/Ingest system already implements the catalog export services allowing data extraction from Qserv. retain the files at GPFS and delete Qservs <p>Discussion:</p> <ul style="list-style-type: none"> Fritz Mueller: we shall put those input files into the "cold" storage (to tape). Can't use NCSA as we will lose it at the end of the FY22. Will talk to Richard and SLAC admins to see what can be done here to put the files to SLAC's HPSS? Andy Hanushevsky proposed using Rucio. Rucio is already working at SLAC and IN2P3. This will require setting up a test version of Rucio at NCSA. <ul style="list-style-type: none"> Fritz Mueller will talk to Michelle today to discuss setting up Rucio endpoint at NCSA.
	What's next for the transition to endpoint?	Unknown User (npease)	<p>Context:</p> <ul style="list-style-type: none"> Transition to launching qserv nodes with the endpoint script <p>Question:</p> <ul style="list-style-type: none"> What is known to be missing from endpoint to do this work? Should we set up some group-work time to start using endpoint in Operator and/or Igor-mode? <p>Igor Gaponenko is ready to begin (gradually) migrating Qserv setups at NCSA to use endpoints.</p> <p>Igor Gaponenko: asked if it's possible to pass a file with template variables instead of specifying multiple --target parameters.</p> <p>Unknown User (npease): the JINJA template processor is configured to catch unresolved templates in the config files.</p> <p>Fabrice Jammes: still working on migrating qserv-operator to use the endpoints. 90% ready. The work will resume working on this next week with Unknown User (npease)'s help.</p>

Action items

