Database Meeting 2022-09-21

Date

21 Sep 2022

Attendees

- Igor Gaponenko Andy Hanushevsky Fritz Mueller Fabrice Jammes Joanne Bogart John Gates Colin Slater Andy Salnikov

Notes from the previous meeting:

- Database Meeting 2022-09-07

Discussion items

Discussed	Item	Who	Notes
✅	Project news	Fritz Mueller, Colin Slater	Fritz Mueller: - JSR started yesterday and it's ongoing. Most of the DM is busy with that. - they're focused on the operations and commissioning activities - DM breakout section will be today Colin Slater: - There is a video on ComCam being installed on LSST video channel
✅	Qserv at SLAC	Igor Gaponenko	Latest news: - The 6-node Qserv is up and running - It's wired into the monitoring system, that's really useful. Here is the direct link to the Graphana monitoring page for the cluster: https://grafana.slac.stanford.edu/d/m_Xura8Zz/batch-cluster?orgId=1&var-hosts=sdfk8sc&from=now-1h&to=now&refresh=30s - Loaded with 4 catalogs: `DP01`, `DP02`, `Gaia DR2`, `Wise` - Single replica configuration as we don't have enough storage for 2 replicas - Used **9.5 TB** (out of **12 TB** available) at each node - Wrote the cluster usage and data management guide at: Managing Qserv instances at SLAC that is linked from Qserv Deployments - Made a short report at the joined facilities (USDF, FrDF, UKDF) meeting on Monday: Joint Data Facility Meeting - 2022-09-19 - The preliminary ETA for the new hardware (15 nodes) is the end of November (heard at the same meeting) Issues: - the network configuration of these 6 nodes prevents apps from connecting to the Internet - had to use a proxy - the real performance of the proxy while ingesting **40 TB** of `DP02` was limited by **200 MB/s** - in the latest change of networking configuration, direct access to the Internet was enabled. ⚠ Though, all external traffic now goes via a software router, limiting the available bandwidth to something like 10 Gbps or less. According to Yee, it was decided by Amedeo P and SLAC Cybersecurity Platoon. Yee thinks that in the future all ingested products should be put onto the SLAC/USDF-local data sources (shared POSIX filesystems, object stores). - the Replication system's algorithm for populating the *director* indexes has poor performance (has restricted implementation preventing it from utilizing available resources). This will be addressed in: ○ ⚠ DM-36241 - Jira project doesn't exist or you don't have permission to view it. What's next: - TAP service - performance tests (to compare with IDF) Fritz Mueller - S3 storage needs to be installed locally for storing TAP service results or enable using the one from the LSST Google Cloud as a workaround

✅	Status of `DP02`	team	**Colin Slater** has it been enabled in the TAP service of `qserv-prod`? Are the truth tables of the `qserv-prod` behaving the way it's expected (like we saw this at `qserv-int`)? **Igor Gaponenko** on the status of the latest version of the remaining tables at IN2P3: - they got the "green light" to proceed with updating/ingesting these tables - according to my last comms with Gabriele Mainetti, this will happen shortly **Colin Slater**: - joins from TAP haven't been tested yet (blocked by the ongoing DM reviews) - latest changes have already been pushed to IDF
✅	Progress on `qserv-operator` and `qserv-ingest`	**Fabrice Jammes**	**Igor Gaponenko** on the known issues: - the short hostnames for the Replication/Ingest workers that were introduced recently do not seem to work for `qserv-ingest` due to Kubernetes namespaces. - issues with the LSST Logger configuration reported in this development: <blockquote>⚠ **DM-36004** - Jira project doesn't exist or you don't have permission to view it.</blockquote>

Action items

- [ ]