Database Meeting 2022-06-01

Date

01 Jun 2022

Attendees

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

Notes from the previous meeting

- Database Meeting 2022-05-25

Discussion items

Discussed	Item	Who	Notes
✅	Project news	Fritz Mueller	- new news from DMLT - PCW2022: everyone is encouraged to get registered - DMLT Face-to-Face: next week, an update on the current status of APDB/PPDB is expected. Fritz Mueller will need help to prepare a report. - Epic S22 is dead, change of cycle this morning. It's been replaced by F22.
✅	DP02		Igor Gaponenko: on the status of processing and ingesting the remaining tables into Qserv - The table `Source` (over 1.6 million contributions) has been successfully ingested into the "small" cluster. - The table `ForcedSource` has (almost) ingested into the "small" cluster. In total, **12** out of **16021** contributions failed to be ingested (HTTP protocol error 18 when pulling contributions from GCS). These contributions will be retried. The table has nearly **30 TB**. The contribution files are rather large. Some have many **GB**s. - Other tables will be ingested later. - Ran into a number of limitations of the R-I system. Most have been addressed in: ⚠ DM-34979 - Jira project doesn't exist or you don't have permission to view it. ⚠ DM-28626 - Jira project doesn't exist or you don't have permission to view it. Colin Slater: on priorities of the ingests into IDF: - first, we need the tables `Visit`, `CcdVisit` and all 3 `Dia*` tables. - the large tables `Source` and `ForcedSource` could be ingested later On priorities: - get the PR of DM-28626 reviewed by the end of the day tomorrow - get this new code deployed in IDF (`qserv-int`) - start ingesting the high-priority tables mentioned above Fritz Mueller: on the status of the IVOA tables (schema, data, ...): - integration test in IDF was conducted last Friday - ObsCore schema has been published in the TAP service in IDF - we have a new (bigger) version of the `ObsCore` table (data) to replace what we have in the catalog `ivoa` in IDF. There is a new schema too. It's on GitHub. Fritz Mueller will send a link to Igor Gaponenko. The input data will be available later. - improved SciSql UDF is ready. It will be deployed in the new version of the MariaDB container - Igor Gaponenko: it would be nice to come up with a better naming convention

| ✅ | Large scale tests of Qserv and the Replication/INgest system | team | Qserv at USDF:

 • 15 workers, 32 TB of the usable NVMe-based storage per worker.
   ○ Fritz Mueller: potentially, the cluster could be increased in size next year. This still needs to be discussed (justified).
 • availability time - sometimes this Fall
 • R&D for the DAX team and the warm-up of the USDF team in caring of Qserv
 • maybe also used for hosting the `ComCam` catalog data (`DC1`).

 Large-scale catalogs for testing in USDF:

 • KPM50-like catalog
 • DP02
 • maybe `DC1` (`ComCam`)

 IDF could be also dramatically increased in size should this decision be justified |

Action items

- TBC