

Level 3, User Tables, and External Data Sets

Related ticket(s): *none yet*

In addition to production databases, DAC users will be allocated database (and file storage) space for their own use. This data may include products derived from the production LSST databases (e.g., results from previously executed queries), or may be brought in from outside (from other sources, e.g., other surveys). At least part of the user database space needs to be distributed on the same nodes as the Level 2 production database to facilitate joins with the large tables.

The database architecture needs to support managing such user tables, including creating, modifying, listing, and deleting them. In some cases user data will require special protection, e.g., users may chose to not share their data with others.

Architecture:

- each user will have her/his own database, users will not be allowed to write/modify/delete any production data
- when running a query, user will be able to specify how the results should be handled (return to the client, store centrally, store as distributed tables)
- qserv metadata keeps track of every user table, including including information such as:
 - ◆ is this table partitioned/distributed?
 - ◆ is it private or can it be publicly readable
- non trivial queries (eg queries that join all many user partitions with production Level 2 catalogs) will rely on shared scans
- user code will run "near the database" (not as UDFs executed in db process)