

# Transaction Handling in the Database

## LSST Database

There are two major cases that we foresee in DC2 and beyond:

- Bulk appends from the image processing and detection pipelines
- Updates in the association pipeline

For bulk appends, the safest way to go is to persist the data into CSV files. These CSV files can be passed to the Catalog Ingest Service atomically if the creating stage has succeeded. They would then be ingested using `LOAD DATA INFILE`. An extra appended column, an integer load identifier, can be used to enable dropping all of the loaded data at once if necessary.

Updates will go to the in-memory copies of table partitions. These will be written out as entire partitions in the post-process phase. The existing partitions could be moved to other names so that they can be put back in place easily if a failure occurs.

Control of ingest and updates and thus control of rollback if a failure occurs will need to be handled at a higher level than the low-level Persistence objects. Instead, it should be part of the pipeline control or Catalog Ingest Service.

## Open questions

- Are there any additional cases than those above? In particular, are individual table rows going to be updated and yet need to be rolled back as a group?