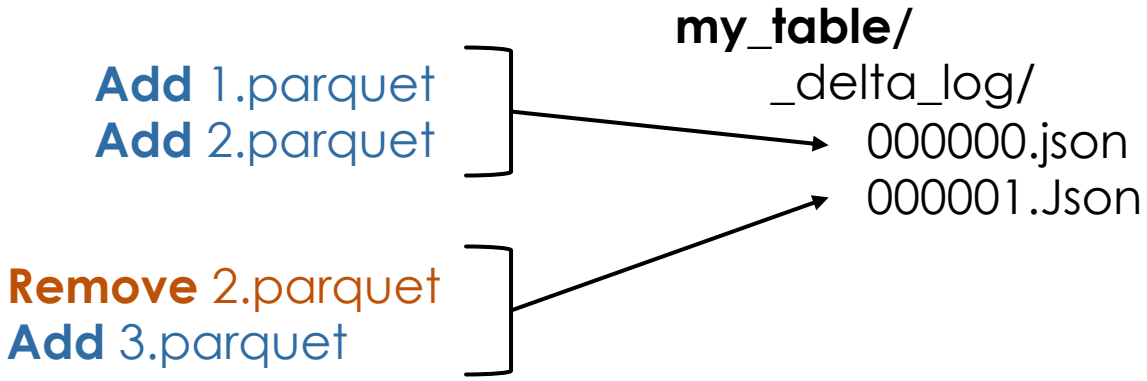
Delta Lake Transaction Log

Learning Objectives

- Transaction Log Checkpoints
- Delta Lake File Statistics

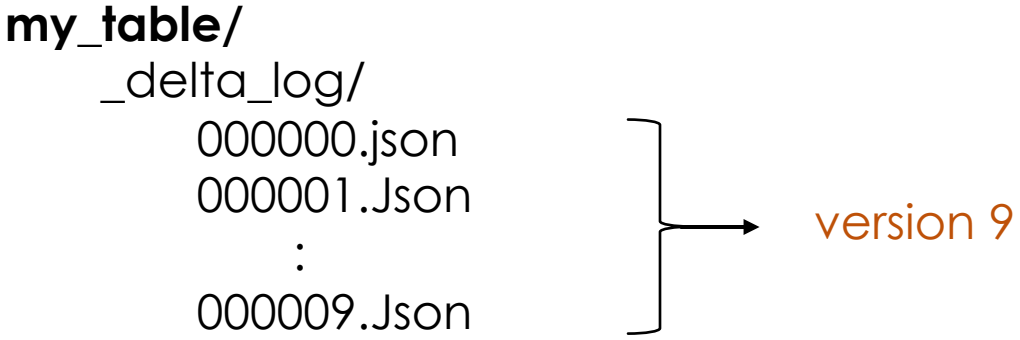
Transaction Log

- Each commit to the table is written out as a JSON file



Transaction Log

- Spark needs to process many tiny, inefficient JSON files in order to the resolve the current table state



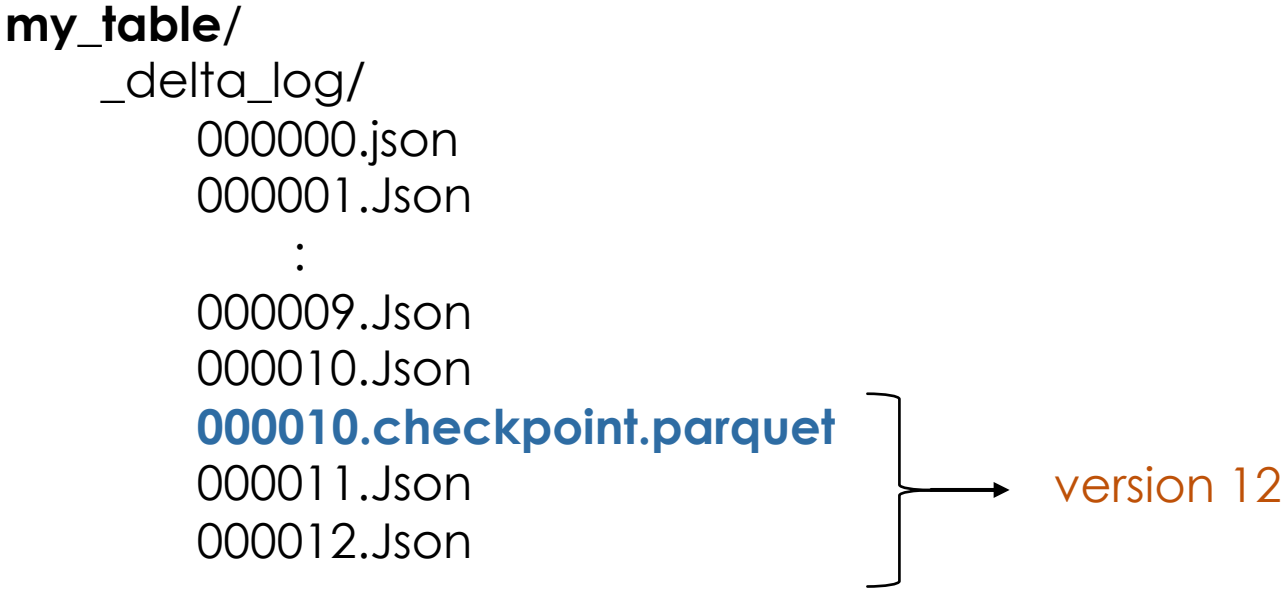
Transaction Log Checkpoints

- Databricks automatically creates Parquet checkpoint files every 10 commits to accelerate the resolution of the current table state.

my_table/
- _delta_log/
  - 000000.json
  - 000001.Json
  - :
  - 000009.Json
  - 000010.Json
  - **000010.checkpoint.parquet** → version 10

Transaction Log Checkpoints

- Then, Spark only has to perform incremental processing of newly added JSON files

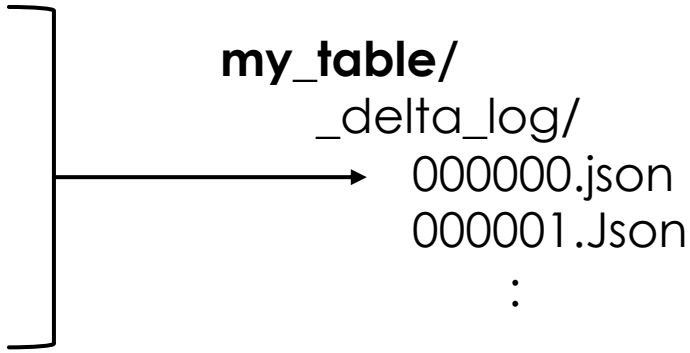


Delta Lake File Statistics

- Delta Lake captures statistics in the transaction log for each added data file

Add 1.parquet
Stats {numRecords, minValues, maxValues, … }

Add 2.parquet
Stats {numRecords, minValues, maxValues, … }



Delta Lake File Statistics

- These statistics indicate per file:
  1. Total number of records
  - Statistics on the **first 32 columns** of the table
    2. Minimum value in each column
    3. Maximum value in each column
    4. Null value counts for each of the columns

- Statistics will always be leveraged for **file skipping**.

Delta Lake File Statistics

- Nested fields count when determining the first 32 columns
  - Example: 4 struct fields with 8 nested fields will total to the 32 columns.
- Statistics are uninformative for string fields with very high cardinality
  - Example: free text fields
  - Time consuming! Move them outside the first 32 columns

Log Retention Period

- Running the VACUUM does not delete Delta log files
- Log files are **automatically** cleaned up by Databricks
  - Each time a checkpoint is written, Databricks automatically cleans up log entries older than the log retention interval (default: **30 days**)
- By default, you can time travel to a Delta table **only** up to 30 days old
  - `delta.logRetentionDuration` controls how long the history for a table is kept