See also Section 17.5.1.14, “Replication and System Functions”.

17.5.1.34 Replication and Transaction Inconsistencies

Inconsistencies in the sequence of transactions that have been executed from the relay log can occur depending on your replication configuration. This section explains how to avoid inconsistencies and solve any problems they cause.

The following types of inconsistencies can exist:

- *Half-applied transactions*. A transaction which updates non-transactional tables has applied some but not all of its changes.

- *Gaps*. A gap in the externalized transaction set appears when, given an ordered sequence of transactions, a transaction that is later in the sequence is applied before some other transaction that is prior in the sequence. Gaps can only appear when using a multithreaded replica. To avoid gaps occurring, set `slave_preserve_commit_order=1`. Up to and including MySQL 8.0.18, this setting requires that binary logging (`log_bin`) and replica update logging (`log_slave_updates`) are also enabled, which are the default settings from MySQL 8.0. From MySQL 8.0.19, binary logging and replica update logging are not required on the replica to set `slave_preserve_commit_order=1`, and can be disabled if wanted. In all releases, setting `slave_preserve_commit_order=1` requires that `slave_parallel_type` is set to `LOGICAL_CLOCK`, which is *not* the default setting. Note that in some specific situations, as listed in the description for `slave_preserve_commit_order`, setting `slave_preserve_commit_order=1` cannot preserve commit order on the replica, so in these cases gaps might still appear in the sequence of transactions that have been executed from the replica's relay log.

- *Source binary log position lag*. Even in the absence of gaps, it is possible that transactions after `Exec_master_log_pos` have been applied. That is, all transactions up to point `N` have been applied, and no transactions after `N` have been applied, but `Exec_master_log_pos` has a value smaller than `N`. In this situation, `Exec_master_log_pos` is a “low-water mark” of the transactions applied, and lags behind the position of the most recently applied transaction. This can only happen on multithreaded replicas. Enabling `slave_preserve_commit_order` does not prevent source binary log position lag.

The following scenarios are relevant to the existence of half-applied transactions, gaps, and source binary log position lag:

1. While replication threads are running, there may be gaps and half-applied transactions.

2. `mysqld` shuts down. Both clean and unclean shutdown abort ongoing transactions and may leave gaps and half-applied transactions.

3. `KILL` of replication threads (the SQL thread when using a single-threaded replica, the coordinator thread when using a multithreaded replica). This aborts ongoing transactions and may leave gaps and half-applied transactions.

4. Error in applier threads. This may leave gaps. If the error is in a mixed transaction, that transaction is half-applied. When using a multithreaded replica, workers which have not received an error complete their queues, so it may take time to stop all threads.

5. `STOP REPLICA | SLAVE` when using a multithreaded replica. After issuing `STOP REPLICA | SLAVE`, the replica waits for any gaps to be filled and then updates `Exec_master_log_pos`. This ensures it never leaves gaps or source binary log position lag, unless any of the cases above applies, in other words, before `STOP REPLICA | SLAVE` completes, either an error happens, or another thread issues `KILL`, or the server restarts. In these cases, `STOP REPLICA | SLAVE` returns successfully.