Auto Clean Up for Hive ORC

Since, openlookeng is now supporting compaction/vacuum operation on Hive transactional tables which merges all base and delta files together (according to the compaction type). So that upcoming read queries can be redirected to use newly created, compacted directories which improves performance. This makes old, smaller delta directories eligible for cleanup. But cleaner operation has to wait for any read operation started before vacuum operation since it might still be using old delta directories.

Therefore, a background process is required which will be responsible for removing any redundant directory from table path.

Directory structure of table `table_txn_001` after 3 insert operations:

```
drwxrwxrwx   - root supergroup          0 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000001_0000001_0000
-rw-r--r--   3 root supergroup          1 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000001_0000001_0000/_orc_acid_version
-rw-r--r--   3 root supergroup        901 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000001_0000001_0000/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000002_0000002_0000
-rw-r--r--   3 root supergroup          1 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000002_0000002_0000/_orc_acid_version
-rw-r--r--   3 root supergroup        909 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000002_0000002_0000/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000003_0000003_0000
-rw-r--r--   3 root supergroup          1 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000003_0000003_0000/_orc_acid_version
-rw-r--r--   3 root supergroup        905 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000003_0000003_0000/bucket_00000
```

➢ After minor compaction on table `table_txn_001`

Command:

VACUUM TABLE table_txn_001;

```
drwxrwxrwx   - root supergroup          0 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000001_0000001_0000
-rw-r--r--   3 root supergroup          1 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000001_0000001_0000/_orc_acid_version
-rw-r--r--   3 root supergroup        901 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000001_0000001_0000/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 20:55 /user/hive/warehouse/table_txn_001/delta_0000001_0000003
-rw-r--r--   3 root supergroup          1 2020-07-09 20:55 /user/hive/warehouse/table_txn_001/delta_0000001_0000003/_orc_acid_version
-rw-r--r--   3 root supergroup        916 2020-07-09 20:55 /user/hive/warehouse/table_txn_001/delta_0000001_0000003/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000002_0000002_0000
-rw-r--r--   3 root supergroup          1 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000002_0000002_0000/_orc_acid_version
-rw-r--r--   3 root supergroup        909 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000002_0000002_0000/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000003_0000003_0000
-rw-r--r--   3 root supergroup          1 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000003_0000003_0000/_orc_acid_version
-rw-r--r--   3 root supergroup        905 2020-07-09 20:50 /user/hive/warehouse/table_txn_001/delta_0000003_0000003_0000/bucket_00000
```

Directory created by Vacuum:

delta_0000001_0000003

Directories ready for clean up:

delta_0000001_0000001_0000

delta_0000002_0000002_0000

delta_0000003_0000003_0000

➢ After major compaction on table `table_txn_001`

Command:

VACUUM TABLE table_txn_001 FULL;

```
drwxrwxrwx   - root supergroup          0 2020-07-09 20:57 /user/hive/warehouse/table_txn_001/base_0000003
-rw-r--r--   3 root supergroup          1 2020-07-09 20:57 /user/hive/warehouse/table_txn_001/base_0000003/_orc_acid_version
-rw-r--r--   3 root supergroup        916 2020-07-09 20:57 /user/hive/warehouse/table_txn_001/base_0000003/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 20:56 /user/hive/warehouse/table_txn_001/delta_0000001_0000001_0000
-rw-r--r--   3 root supergroup          1 2020-07-09 20:56 /user/hive/warehouse/table_txn_001/delta_0000001_0000001_0000/_orc_acid_version
-rw-r--r--   3 root supergroup        903 2020-07-09 20:56 /user/hive/warehouse/table_txn_001/delta_0000001_0000001_0000/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 20:56 /user/hive/warehouse/table_txn_001/delta_0000002_0000002_0000
-rw-r--r--   3 root supergroup          1 2020-07-09 20:56 /user/hive/warehouse/table_txn_001/delta_0000002_0000002_0000/_orc_acid_version
-rw-r--r--   3 root supergroup        910 2020-07-09 20:56 /user/hive/warehouse/table_txn_001/delta_0000002_0000002_0000/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 20:56 /user/hive/warehouse/table_txn_001/delta_0000003_0000003_0000
-rw-r--r--   3 root supergroup          1 2020-07-09 20:56 /user/hive/warehouse/table_txn_001/delta_0000003_0000003_0000/_orc_acid_version
-rw-r--r--   3 root supergroup        905 2020-07-09 20:56 /user/hive/warehouse/table_txn_001/delta_0000003_0000003_0000/bucket_00000
```

Directory created by Vacuum:

    base_0000003

Directories ready for clean up:

    delta_0000001_0000001_0000

    delta_0000002_0000002_0000

    delta_0000003_0000003_0000

➢ After major compaction followed by minor compaction on table `table_txn_001`

```
drwxrwxrwx   - root supergroup          0 2020-07-09 21:21 /user/hive/warehouse/table_txn_001/base_0000003
-rw-r--r--   3 root supergroup          1 2020-07-09 21:21 /user/hive/warehouse/table_txn_001/base_0000003/_orc_acid_version
-rw-r--r--   3 root supergroup        917 2020-07-09 21:21 /user/hive/warehouse/table_txn_001/base_0000003/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 21:19 /user/hive/warehouse/table_txn_001/delta_0000001_0000001_0000
-rw-r--r--   3 root supergroup          1 2020-07-09 21:19 /user/hive/warehouse/table_txn_001/delta_0000001_0000001_0000/_orc_acid_version
-rw-r--r--   3 root supergroup        913 2020-07-09 21:19 /user/hive/warehouse/table_txn_001/delta_0000001_0000001_0000/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 21:21 /user/hive/warehouse/table_txn_001/delta_0000001_0000003
-rw-r--r--   3 root supergroup          1 2020-07-09 21:21 /user/hive/warehouse/table_txn_001/delta_0000001_0000003/_orc_acid_version
-rw-r--r--   3 root supergroup        918 2020-07-09 21:21 /user/hive/warehouse/table_txn_001/delta_0000001_0000003/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 21:19 /user/hive/warehouse/table_txn_001/delta_0000002_0000002_0000
-rw-r--r--   3 root supergroup          1 2020-07-09 21:19 /user/hive/warehouse/table_txn_001/delta_0000002_0000002_0000/_orc_acid_version
-rw-r--r--   3 root supergroup        903 2020-07-09 21:19 /user/hive/warehouse/table_txn_001/delta_0000002_0000002_0000/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 21:20 /user/hive/warehouse/table_txn_001/delta_0000003_0000003_0000
-rw-r--r--   3 root supergroup          1 2020-07-09 21:20 /user/hive/warehouse/table_txn_001/delta_0000003_0000003_0000/_orc_acid_version
-rw-r--r--   3 root supergroup        910 2020-07-09 21:20 /user/hive/warehouse/table_txn_001/delta_0000003_0000003_0000/bucket_00000
```

Directory created by last Vacuum:

    base_0000003

Directories ready for clean up:

    delta_0000001_0000003

    delta_0000001_0000001_0000

    delta_0000002_0000002_0000

    delta_0000003_0000003_0000

➢ After minor compaction on table with delete delta directories

```
drwxrwxrwx   - root supergroup          0 2020-07-09 21:23 /user/hive/warehouse/table_txn_001/delete_delta_0000001_0000004
-rw-r--r--   3 root supergroup          1 2020-07-09 21:23 /user/hive/warehouse/table_txn_001/delete_delta_0000001_0000004/_orc_acid_version
-rw-r--r--   3 root supergroup        786 2020-07-09 21:23 /user/hive/warehouse/table_txn_001/delete_delta_0000001_0000004/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 21:22 /user/hive/warehouse/table_txn_001/delete_delta_0000004_0000004_0000
-rw-r--r--   3 root supergroup          1 2020-07-09 21:22 /user/hive/warehouse/table_txn_001/delete_delta_0000004_0000004_0000/_orc_acid_version
-rw-r--r--   3 root supergroup        786 2020-07-09 21:22 /user/hive/warehouse/table_txn_001/delete_delta_0000004_0000004_0000/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 21:22 /user/hive/warehouse/table_txn_001/delta_0000001_0000001_0000
-rw-r--r--   3 root supergroup          1 2020-07-09 21:22 /user/hive/warehouse/table_txn_001/delta_0000001_0000001_0000/_orc_acid_version
-rw-r--r--   3 root supergroup        907 2020-07-09 21:22 /user/hive/warehouse/table_txn_001/delta_0000001_0000001_0000/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 21:23 /user/hive/warehouse/table_txn_001/delta_0000001_0000004
-rw-r--r--   3 root supergroup          1 2020-07-09 21:23 /user/hive/warehouse/table_txn_001/delta_0000001_0000004/_orc_acid_version
-rw-r--r--   3 root supergroup        938 2020-07-09 21:23 /user/hive/warehouse/table_txn_001/delta_0000001_0000004/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 21:22 /user/hive/warehouse/table_txn_001/delta_0000002_0000002_0000
-rw-r--r--   3 root supergroup          1 2020-07-09 21:22 /user/hive/warehouse/table_txn_001/delta_0000002_0000002_0000/_orc_acid_version
-rw-r--r--   3 root supergroup        911 2020-07-09 21:22 /user/hive/warehouse/table_txn_001/delta_0000002_0000002_0000/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 21:22 /user/hive/warehouse/table_txn_001/delta_0000003_0000003_0000
-rw-r--r--   3 root supergroup          1 2020-07-09 21:22 /user/hive/warehouse/table_txn_001/delta_0000003_0000003_0000/_orc_acid_version
-rw-r--r--   3 root supergroup        908 2020-07-09 21:22 /user/hive/warehouse/table_txn_001/delta_0000003_0000003_0000/bucket_00000
drwxrwxrwx   - root supergroup          0 2020-07-09 21:22 /user/hive/warehouse/table_txn_001/delta_0000004_0000004_0000
-rw-r--r--   3 root supergroup          1 2020-07-09 21:22 /user/hive/warehouse/table_txn_001/delta_0000004_0000004_0000/_orc_acid_version
-rw-r--r--   3 root supergroup        925 2020-07-09 21:22 /user/hive/warehouse/table_txn_001/delta_0000004_0000004_0000/bucket_00000
```

Directory created by last Vacuum:

delta_0000001_0000004

delete_delta_0000001_0000004

Directories ready for clean up:

delta_0000001_0000001_0000

delta_0000002_0000002_0000

delta_0000003_0000003_0000

delta_0000004_0000004_0000

delete_delta_0000004_0000004_0000

Approach:

To perform cleanup after vacuum, CleanerTask will be submitted to ScheduledExecutorService from HiveMetadata#commit(). Till this stage, new directory is created and old/compacted directories can be cleaned. But before cleanup we have to wait for finish of any readers using old/compacted directories.

CleanerTask#readyToClean() is used to insure that each read lock taken on that table before compaction must finish. This function makes call to ThriftHiveMetastoreClient#showLocks to get a list of read locks on table. This function returns true if all locks taken before compaction, are released. Otherwise returns false.

On receiving false from readyToClean(), CleanerTask will stop current thread execution and wait for ExecutorService to schedule CleanerTask again, that would be after configurable value (hive.compaction-cleanup-interval).

True from readyToClean() signals CleanerTask to start remove files. List of directories which have to be removed will be received from AcidUtils.Directory.getObsolete().

➢ After each vacuum commit, a CleanerTask will be submitted to scheduled executor service for cleaning obsolete files.
➢ Configuration added:
  o *hive.compaction-cleanup-interval* : (Milis) Interval after which cleanup task will be resubmitted if there is any unreleased read lock on table. Default value is 2000.
  o *hive-compaction-cleanup-threads* : Number of threads to run in compaction cleanup service. Default value is 2.

Flow diagram: