

Supporting weekly, monthly, quarterly and yearly MOB compaction partition policy in HBase

Huaxiang Sun 10/24/16

1. Current Implementation

The current mob compaction partition is daily based, i.e, mob files for one day is one partition. All mob files belonging to a partition will be compacted into one file. For one specific region, there could be as many as 365 mob files after major mob compaction for one year. Mob files for all the regions in one table are created in one specific mob region. In theory, there will be lots of mob files under that directory. Given that HDFS limits the number of files under one directory, this is not going to scale. For example, table t1 has two regions, r1, r2, it has one column family f1 with mob enabled. For 2016/01/01, there will be two mob files under directory of mobdir/data/default/t1/78e317a6e78a0fceb27b9fa0cb9dcf5b/f1 (the mob region directory).

D279186428a75016b17e4df5ea43d080 20160101f9d9713ab2fb4a8b825485f6a8acfd5
D41d8cd98f00b204e9800998ecf8427e 20160101fc94af623c2345f1b241887721e32a48

D279186428a75016b17e4df5ea43d080 is the hash value of region r1.
D41d8cd98f00b204e9800998ecf8427e is the hash value of region r2.

The number of mob files for one specific family in one year will be 365 * number of regions.

2. Proposal

To reduce number of files after mob compaction, we propose the weekly, monthly, quarterly and yearly mob compact partition policy. All these policies are based on the calendar date.

For an example, today is 10/20/2016, the calendar week starts from 10/17/2016 to 10/23/2016(inclusive), i.e., from Monday to Sunday. For monthly policy, it will be from 10/1/2016 to 10/31/2016. With these new policies, after major mob compaction, there will be one file per region for one week/one month/one quarter/one year, i.e, there will be 52 files for weekly policy, 12 files for monthly policy, 3 files for quarterly policy and 1 file for yearly policy.

3. How file names are created with new policies

The file name format for new policies is same as the current one, i.e, (hash of region) + date + (uuid). When compacting mob files with policy other than daily, the date in newly created file name will be the last day of the policy configured. For an example, for one region, there are three mob files for 10/19/2016, 10/20/2016, 10/22/2016 and the compact policy is weekly,

```
D279186428a75016b17e4df5ea43d080 20161020f9d9713ab2fb4a8b825485f6a8acfd5  
D279186428a75016b17e4df5ea43d080 20161019adf9713ab2fb4a8b825485f6a8acfd5  
D279186428a75016b17e4df5ea43d080 20161022adf9713ab2fb4a8b825485f6a8acafd1
```

After major mob compact, there will be one file with date 10/23/2016 (the last day of that week).

```
D279186428a75016b17e4df5ea43d080 20161023adf9713ab2fb4a8b8245644f6a8acafd1
```

The reason is that the above 3 days belong to the same week ending on 10/23/2016. After compaction, 10/23/2016 is date for the new mob file.

4. How compaction works

When mob compaction happens, it tries to partition files based on the currently policy and it compacts files with the same partition into one file, similar as the current daily partition policy.

5. when compaction partition policy changes

Since there is no format change, the new compacted file is treated as a daily file. Changing partition policy does not need special handling.

6. How is mob expired file policy affected?

Today Mob file expiration policy is daily based. With the new policies, cells will be compacted into a daily file with date as the last day of the policy. Per example in [3], the final file's date is 10/23/2016. Let's say cells created on 10/20/2016 is supposed to be cleaned up on 10/22/2016, in this case, it will be cleaned up on 10/25/2016 by the expiredMobFileCleaner.

7. Snapshot changes

Snapshot will not be impacted and will still work as it is.

8. Config and verification

There will be a new attribute added for mob column family, MOB_COMPACT_PARTITION_POLICY. Its value for now is “daily”, “weekly”, “monthly”, “quarterly” and “yearly”. User can set this attribute from the shell.

```
>create 't1', {NAME => 'f1', IS_MOB => true, MOB_THRESHOLD => 10,  
MOB_COMPACT_PARTITION_POLICY => 'weekly'}
```

```
>alter 't1', {NAME => 'f1', MOB_COMPACT_PARTITION_POLICY => 'monthly'}
```

User can check the number of files in mob region to see if the policy works or not. They can also check the region server logs to confirm if the new mob partition policy works or not.

9. Tests

The following unittest cases will be added.

1. Major compaction works as expected for weekly, monthly, quarterly and yearly mob compaction policies.
2. Minor compaction with merge size works as expected for newly added compaction policies.
3. Compaction with Delete files works as expected for newly added weekly, monthly compaction policies.
4. When policy changes from one policy to another, after compaction, the result is correct.