

git_comments:

1. * * The number of invisible segment info which will be preserved in tablestatus file, * if it exceeds this value, they will be removed and write to tablestatus.history file.
2. * * default value is 200, it means that it will preserve 200 invisible segment info * in tablestatus file. * The size of one segment info is about 500 bytes, so the size of tablestatus file * will remain at 100KB.
3. if execute command 'clean files' or the number of invisible segment info exceeds the value of 'carbon.invisible.segments.preserve.count', it need to append the invisible segment list to 'tablestatus.history' file.
4. * * Separate visible and invisible segments into two array.
5. * * This method reads the load history metadata file * * @param metadataFolderPath * @return
6. * * Return an array containing all invisible segment entries in appendList and historyList.
7. can not remove segment 0, there are some info will be used later for example: updateStatusFileName
8. * * Get the number of invisible segment info from segment info list.
9. * * Get the number of invisible segment info which will be preserved in tablestatus file.
10. * * Return table status history file path based on `tablePath`

git_commits:

1. **summary:** [CARBONDATA-2258] Separate visible and invisible segments info into two files to reduce the size of tablestatus file.
message: [CARBONDATA-2258] Separate visible and invisible segments info into two files to reduce the size of tablestatus file. The size of the tablestatus file is getting larger, there are many places will scan this file and it will impact the performance of reading this file. According to the discussion on thread, it can append the invisible segment list to the file called 'tablestatus.history' when execute command 'CLEAN FILES FOR TABLE' (in method 'SegmentStatusManager.deleteLoadsAndUpdateMetadata') every time, separate visible and invisible segments into two files(tablestatus file and tablestatus.history file). This closes #2091
label: code-design

github_issues:

github_issues_comments:

github_pulls:

1. **title:** [CARBONDATA-2258] Separate visible and invisible segments info into two files to reduce the size of tablestatus file.
body: The size of the tablestatus file is getting larger, there are many places will scan this file and it will impact the performance of reading this file. According to the discussion on [thread](<http://apache-carbondata-dev-mailing-list-archive.1130556.n5.nabble.com/The-size-of-the-tablestatus-file-is-getting-larger-does-it-impact-the-performance-of-reading-this-fi-td41941.html>), it can *append* the invisible segment list to the file called 'tablestatus.history' when execute command 'CLEAN FILES FOR TABLE' (in method 'SegmentStatusManager.deleteLoadsAndUpdateMetadata') every time, separate visible and invisible segments into two files(tablestatus file and tablestatus.history file). Be sure to do all of the following checklist to help us incorporate your contribution quickly and easily: - [] Any interfaces changed? No - [] Any backward compatibility impacted? No - [] Document update required? - [x] Testing done Please provide details on - Whether new unit test cases have been added or why no new tests are required? - How it is tested? Please attach test report. - Is it a performance related change? Please attach the performance test report. - Any additional information to help reviewers in testing this change. - [] For large changes, please consider breaking it into sub-tasks under an umbrella JIRA.

github_pulls_comments:

1. Build Failed with Spark 2.2.1, Please check CI
<http://88.99.58.216:8080/job/ApacheCarbonPRBuilder/3283/>
2. Build Failed with Spark 2.1.0, Please check CI
<http://136.243.101.176:8080/job/ApacheCarbonPRBuilder1/4511/>

3. SDV Build Fail , Please check CI <http://144.76.159.231:8080/job/ApacheSDVTests/4017/>
4. retest this please
5. Build Failed with Spark 2.2.1, Please check CI
<http://88.99.58.216:8080/job/ApacheCarbonPRBuilder/3288/>
6. Build Failed with Spark 2.1.0, Please check CI
<http://136.243.101.176:8080/job/ApacheCarbonPRBuilder1/4516/>
7. Build Failed with Spark 2.2.1, Please check CI
<http://88.99.58.216:8080/job/ApacheCarbonPRBuilder/3303/>
8. Build Failed with Spark 2.1.0, Please check CI
<http://136.243.101.176:8080/job/ApacheCarbonPRBuilder1/4531/>
9. retest sdv please.
10. SDV Build Success , Please check CI <http://144.76.159.231:8080/job/ApacheSDVTests/4031/>
11. Build Failed with Spark 2.2.1, Please check CI
<http://88.99.58.216:8080/job/ApacheCarbonPRBuilder/3308/>
12. Build Failed with Spark 2.1.0, Please check CI
<http://136.243.101.176:8080/job/ApacheCarbonPRBuilder1/4536/>
13. Build Success with Spark 2.1.0, Please check CI
<http://136.243.101.176:8080/job/ApacheCarbonPRBuilder1/4542/>
14. SDV Build Success , Please check CI <http://144.76.159.231:8080/job/ApacheSDVTests/4035/>
15. retest this please
16. retest sdv please
17. Build Failed with Spark 2.2.1, Please check CI
<http://88.99.58.216:8080/job/ApacheCarbonPRBuilder/3318/>
18. Build Success with Spark 2.1.0, Please check CI
<http://136.243.101.176:8080/job/ApacheCarbonPRBuilder1/4545/>
19. SDV Build Success , Please check CI <http://144.76.159.231:8080/job/ApacheSDVTests/4040/>
20. SDV Build Success , Please check CI <http://144.76.159.231:8080/job/ApacheSDVTests/4043/>
21. SDV Build Success , Please check CI <http://144.76.159.231:8080/job/ApacheSDVTests/4044/>
22. retest this please
23. Build Success with Spark 2.1.0, Please check CI
<http://136.243.101.176:8080/job/ApacheCarbonPRBuilder1/4549/>
24. Build Success with Spark 2.2.1, Please check CI
<http://88.99.58.216:8080/job/ApacheCarbonPRBuilder/3322/>
25. @jackylk @ravipesala please help to review, thanks
26. Build Success with Spark 2.2.1, Please check CI
<http://88.99.58.216:8080/job/ApacheCarbonPRBuilder/3339/>
27. Build Success with Spark 2.1.0, Please check CI
<http://136.243.101.176:8080/job/ApacheCarbonPRBuilder1/4565/>
28. SDV Build Success , Please check CI <http://144.76.159.231:8080/job/ApacheSDVTests/4061/>
29. LGTM

github_pulls_reviews:

1. The default is 200 right. Please update comment
2. change to `Return an array containing all invisible segment entries in appendList and historyList`
3. Done
4. Done

jira_issues:

1. **summary:** Separate visible and invisible segments info into two files to reduce the size of tablestatus file.
description: The size of the tablestatus file is getting larger, there are many places will scan this file and it will impact the performance of reading this file. According to the discussion on [thread|<http://apache-carbondata-dev-mailing-list-archive.1130556.n5.nabble.com/The-size-of-the-tablestatus-file-is-getting-larger-does-it-impact-the-performance-of-reading-this-fi-td41941.html>], it can *append* the invisible segment list to the file called 'tablestatus.history' when execute command 'CLEAN FILES FOR TABLE' (in method 'SegmentStatusManager.deleteLoadsAndUpdateMetadata') every time, separate visible and invisible segments into two files(tablestatus file and tablestatus.history file). If later it needs to support

listing all segments(include visible and invisible) list when execute 'SHOW SEGMENTS FOR TABLE', it just need to read from two files.

jira_issues_comments: