Item 167
git_comments:

1. **comment:** TODO: make it mutable and update it after each writing back to HDFS
   label: requirement
2. * * Check if we can start the write (back to HDFS) now. If there is no hole for * writing, and there is no other threads writing (i.e., asyncStatus is * false), start the writing and set asyncStatus to true. * * @return True if the new write is sequencial and we can start writing * (including the case that there is already a thread writing).
3. * Dump data into a file
4. update nextOffset
5. offset > nextOffset. check if we need to dump data
6. * * The current offset of the file in HDFS. All the content before this offset * has been written back to HDFS.
7. * * Creates and adds a WriteCtx into the pendingWrites map. This is a * synchronized method to handle concurrent writes. * * @return A non-null {@link WriteCtx} instance if the incoming write * request's offset >= nextOffset. Otherwise null.
8. The writes is added to pendingWrites. Check and start writing back if necessary
9. shouldn't happen since we do sync for overlapped concurrent writers
10. Fail non-append call
11. In test, noticed some Linux client sends a batch (e.g., 1MB) of reordered writes and won't send more writes until it gets responses of the previous batch. So here send response immediately for unstable non-sequential write
12. * Process an overwrite write request
13. * * Get (and remove) the next WriteCtx from {@link #pendingWrites} if possible. * * @return Null if {@link #pendingWrites} is null, or the next WriteCtx's * offset is larger than nextOffSet.
14. * Invoked by AsynDataService to write back to HDFS
15. after writing, remove the WriteCtx from cache
16. Handle repeated write requests (same xid or not). If already replied, send reply again. If not replied, drop the repeated request.
17. not a repeated write request
18. check if alive again
19. We use the ReverseComparatorOnMin as the comparator of the map. In this way, we first dump the data with larger offset. In the meanwhile, we retrieve the last element to write back to HDFS.
20. wake up the dumper thread to dump the data
21. update the memory size
22. The write is not protected by lock. asyncState is used to make sure there is one thread doing write back at any time
23. stop the dump thread
24. make sure we reset asyncStatus to false
25. check if there is a WriteCtx with the same range in pendingWrites
26. **comment:** TODO: let executor handle perfect overwrite
    label: code-design
27. read back from dumped file
28. **comment:** it is possible that while we dump the data, the data is also being written back to HDFS. After dump, if the writing back has not finished yet, we change its flag to DUMPED and set the data to null. Otherwise this WriteCtx instance should have been removed from the buffer.
    label: requirement
29. * * Data belonging to the same {@link OpenFileCtx} may be dumped to a file. * After being dumped to the file, the corresponding {@link WriteCtx} records * the dump file and the offset.
30. * * Writing the data into a local file. After the writing, if * {@link #dataState} is still ALLOW_DUMP, set {@link #data} to null and set * {@link #dataState} to DUMPED.
31. 10 minutes 10 seconds

git_commits:

1. **summary:** HDFS-4971. Move IO operations out of locking in OpenFileCtx. Contributed by Jing Zhao and Brandon Li.

message: HDFS-4971. Move IO operations out of locking in OpenFileCtx. Contributed by Jing Zhao and Brandon Li. git-svn-id: https://svn.apache.org/repos/asf/hadoop/common/trunk@1525681 13f79535-47bb-0310-9956-ffa450edef68

github_issues:

github_issues_comments:

github_pulls:

github_pulls_comments:

github_pulls_reviews:

jira_issues:

1. **summary:** Move IO operations out of locking in OpenFileCtx
   description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.
2. **summary:** Move IO operations out of locking in OpenFileCtx
   description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.
3. **summary:** Move IO operations out of locking in OpenFileCtx
   description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.
4. **summary:** Move IO operations out of locking in OpenFileCtx
   description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.
5. **summary:** Move IO operations out of locking in OpenFileCtx
   description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.
6. **summary:** Move IO operations out of locking in OpenFileCtx
   description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.
7. **summary:** Move IO operations out of locking in OpenFileCtx
   description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.
8. **summary:** Move IO operations out of locking in OpenFileCtx
   description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.
9. **summary:** Move IO operations out of locking in OpenFileCtx
   description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.
10. **summary:** Move IO operations out of locking in OpenFileCtx
    description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.
11. **summary:** Move IO operations out of locking in OpenFileCtx
    description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to

optimize OpenFileCtx and move the IO operations out of the locking.

12. **summary:** Move IO operations out of locking in OpenFileCtx
    description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.
13. **summary:** Move IO operations out of locking in OpenFileCtx
    description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.
14. **summary:** Move IO operations out of locking in OpenFileCtx
    description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.
15. **summary:** Move IO operations out of locking in OpenFileCtx
    description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.
16. **summary:** Move IO operations out of locking in OpenFileCtx
    description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.
17. **summary:** Move IO operations out of locking in OpenFileCtx
    description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.
18. **summary:** Move IO operations out of locking in OpenFileCtx
    description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.
    label: code-design
19. **summary:** Move IO operations out of locking in OpenFileCtx
    description: Currently some IO operations (such as writing data to HDFS and dumping to local disk) in OpenFileCtx may hold a lock which can block processing incoming writing requests. This jira aims to optimize OpenFileCtx and move the IO operations out of the locking.

jira_issues_comments:

1. Initial patch just for review.
2. Rebase the patch.
3. Update the patch: 1) simply the locking mechanism, 2) handle the concurrent overlap writer scenario, 3) handle a race on nextOffset, 4) refactor OpenFileCtx#receivedNewWrite and OpenFileCtx#executeWriteBack.
4. {color:red}-1 overall{color}. Here are the results of testing the latest attachment http://issues.apache.org/jira/secure/attachment/12603698/HDFS-4971.002.patch against trunk revision . {color:green}+1 @author{color}. The patch does not contain any @author tags. {color:green}+1 tests included{color}. The patch appears to include 1 new or modified test files. {color:green}+1 javac{color}. The applied patch does not increase the total number of javac compiler warnings. {color:green}+1 javadoc{color}. The javadoc tool did not generate any warning messages. {color:green}+1 eclipse:eclipse{color}. The patch built with eclipse:eclipse. {color:red}-1 findbugs{color}. The patch appears to introduce 1 new Findbugs (version 1.3.9) warnings. {color:green}+1 release audit{color}. The applied patch does not increase the total number of release audit warnings. {color:green}+1 core tests{color}. The patch passed unit tests in hadoop-hdfs-project/hadoop-hdfs-nfs. {color:green}+1 contrib tests{color}. The patch passed contrib unit tests. Test results: https://builds.apache.org/job/PreCommit-HDFS-Build/4984//testReport/ Findbugs warnings: https://builds.apache.org/job/PreCommit-HDFS-Build/4984//artifact/trunk/patchprocess/newPatchFindbugsWarningshadoop-hdfs-nfs.html Console output: https://builds.apache.org/job/PreCommit-HDFS-Build/4984//console This message is automatically generated.
5. {color:green}+1 overall{color}. Here are the results of testing the latest attachment http://issues.apache.org/jira/secure/attachment/12603711/HDFS-4971.003.patch against trunk revision .

{color:green}+1 @author{color}. The patch does not contain any @author tags. {color:green}+1 tests included{color}. The patch appears to include 1 new or modified test files. {color:green}+1 javac{color}. The applied patch does not increase the total number of javac compiler warnings. {color:green}+1 javadoc{color}. The javadoc tool did not generate any warning messages. {color:green}+1 eclipse:eclipse{color}. The patch built with eclipse:eclipse. {color:green}+1 findbugs{color}. The patch does not introduce any new Findbugs (version 1.3.9) warnings. {color:green}+1 release audit{color}. The applied patch does not increase the total number of release audit warnings. {color:green}+1 core tests{color}. The patch passed unit tests in hadoop-hdfs-project/hadoop-hdfs-nfs. {color:green}+1 contrib tests{color}. The patch passed contrib unit tests. Test results: https://builds.apache.org/job/PreCommit-HDFS-Build/4989//testReport/ Console output: https://builds.apache.org/job/PreCommit-HDFS-Build/4989//console This message is automatically generated.

6. Rebase the patch.
7. {color:green}+1 overall{color}. Here are the results of testing the latest attachment http://issues.apache.org/jira/secure/attachment/12603874/HDFS-4971.004.patch against trunk revision . {color:green}+1 @author{color}. The patch does not contain any @author tags. {color:green}+1 tests included{color}. The patch appears to include 1 new or modified test files. {color:green}+1 javac{color}. The applied patch does not increase the total number of javac compiler warnings. {color:green}+1 javadoc{color}. The javadoc tool did not generate any warning messages. {color:green}+1 eclipse:eclipse{color}. The patch built with eclipse:eclipse. {color:green}+1 findbugs{color}. The patch does not introduce any new Findbugs (version 1.3.9) warnings. {color:green}+1 release audit{color}. The applied patch does not increase the total number of release audit warnings. {color:green}+1 core tests{color}. The patch passed unit tests in hadoop-hdfs-project/hadoop-hdfs-nfs. {color:green}+1 contrib tests{color}. The patch passed contrib unit tests. Test results: https://builds.apache.org/job/PreCommit-HDFS-Build/4997//testReport/ Console output: https://builds.apache.org/job/PreCommit-HDFS-Build/4997//console This message is automatically generated.
8. Update the patch after offline discussion with [~brandonli]. The main change includes: # Add processing logic for overlapped write # Handling closed FSDataOutputStream # Start dumper thread only when necessary # Initialize nextOffset
9. {color:red}-1 overall{color}. Here are the results of testing the latest attachment http://issues.apache.org/jira/secure/attachment/12603970/HDFS-4971.005.patch against trunk revision . {color:green}+1 @author{color}. The patch does not contain any @author tags. {color:green}+1 tests included{color}. The patch appears to include 1 new or modified test files. {color:green}+1 javac{color}. The applied patch does not increase the total number of javac compiler warnings. {color:green}+1 javadoc{color}. The javadoc tool did not generate any warning messages. {color:green}+1 eclipse:eclipse{color}. The patch built with eclipse:eclipse. {color:red}-1 findbugs{color}. The patch appears to introduce 1 new Findbugs (version 1.3.9) warnings. {color:green}+1 release audit{color}. The applied patch does not increase the total number of release audit warnings. {color:green}+1 core tests{color}. The patch passed unit tests in hadoop-hdfs-project/hadoop-hdfs-nfs. {color:green}+1 contrib tests{color}. The patch passed contrib unit tests. Test results: https://builds.apache.org/job/PreCommit-HDFS-Build/5000//testReport/ Findbugs warnings: https://builds.apache.org/job/PreCommit-HDFS-Build/5000//artifact/trunk/patchprocess/newPatchFindbugsWarningshadoop-hdfs-nfs.html Console output: https://builds.apache.org/job/PreCommit-HDFS-Build/5000//console This message is automatically generated.
10. Fix the findbug.
11. {color:green}+1 overall{color}. Here are the results of testing the latest attachment http://issues.apache.org/jira/secure/attachment/12603983/HDFS-4971.006.patch against trunk revision . {color:green}+1 @author{color}. The patch does not contain any @author tags. {color:green}+1 tests included{color}. The patch appears to include 1 new or modified test files. {color:green}+1 javac{color}. The applied patch does not increase the total number of javac compiler warnings. {color:green}+1 javadoc{color}. The javadoc tool did not generate any warning messages. {color:green}+1 eclipse:eclipse{color}. The patch built with eclipse:eclipse. {color:green}+1 findbugs{color}. The patch does not introduce any new Findbugs (version 1.3.9) warnings. {color:green}+1 release audit{color}. The applied patch does not increase the total number of release audit warnings. {color:green}+1 core tests{color}. The patch passed unit tests in hadoop-hdfs-project/hadoop-hdfs-nfs. {color:green}+1 contrib tests{color}. The patch passed contrib unit tests. Test results: https://builds.apache.org/job/PreCommit-HDFS-Build/5002//testReport/ Console output: https://builds.apache.org/job/PreCommit-HDFS-Build/5002//console This message is automatically generated.
12. Thanks for the patch. Some quick comments: 1. in the cleanup(), it should check if dumpOut==null before trying to delete the dump file. 2. in the dumper thread, it should check if dump is still enable

before each loop: while (activeState && enabledDump) {
13. Update the patch. Thanks Brandon for offline review, testing, and bug fix!
14. +1. Thanks you, Jing!!
15. {color:green}+1 overall{color}. Here are the results of testing the latest attachment http://issues.apache.org/jira/secure/attachment/12604472/HDFS-4971.007.patch against trunk revision . {color:green}+1 @author{color}. The patch does not contain any @author tags. {color:green}+1 tests included{color}. The patch appears to include 1 new or modified test files. {color:green}+1 javac{color}. The applied patch does not increase the total number of javac compiler warnings. {color:green}+1 javadoc{color}. The javadoc tool did not generate any warning messages. {color:green}+1 eclipse:eclipse{color}. The patch built with eclipse:eclipse. {color:green}+1 findbugs{color}. The patch does not introduce any new Findbugs (version 1.3.9) warnings. {color:green}+1 release audit{color}. The applied patch does not increase the total number of release audit warnings. {color:green}+1 core tests{color}. The patch passed unit tests in hadoop-hdfs-project/hadoop-hdfs-nfs. {color:green}+1 contrib tests{color}. The patch passed contrib unit tests. Test results: https://builds.apache.org/job/PreCommit-HDFS-Build/5019//testReport/ Console output: https://builds.apache.org/job/PreCommit-HDFS-Build/5019//console This message is automatically generated.
16. Thanks again Brandon! I will commit the patch soon.
17. SUCCESS: Integrated in Hadoop-trunk-Commit #4458 (See [https://builds.apache.org/job/Hadoop-trunk-Commit/4458/]) HDFS-4971. Move IO operations out of locking in OpenFileCtx. Contributed by Jing Zhao and Brandon Li. (jing9: http://svn.apache.org/viewcvs.cgi/?root=Apache-SVN&view=rev&rev=1525681) * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/AsyncDataService.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/OffsetRange.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/OpenFileCtx.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/WriteCtx.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/WriteManager.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/test/java/org/apache/hadoop/hdfs/nfs/nfs3/TestOffsetRange.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt
18. I've committed this to trunk, branch-2 and branch-2.1-beta.
19. SUCCESS: Integrated in Hadoop-Yarn-trunk #342 (See [https://builds.apache.org/job/Hadoop-Yarn-trunk/342/]) HDFS-4971. Move IO operations out of locking in OpenFileCtx. Contributed by Jing Zhao and Brandon Li. (jing9: http://svn.apache.org/viewcvs.cgi/?root=Apache-SVN&view=rev&rev=1525681) * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/AsyncDataService.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/OffsetRange.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/OpenFileCtx.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/WriteCtx.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/WriteManager.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/test/java/org/apache/hadoop/hdfs/nfs/nfs3/TestOffsetRange.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt
20. FAILURE: Integrated in Hadoop-Mapreduce-trunk #1558 (See [https://builds.apache.org/job/Hadoop-Mapreduce-trunk/1558/]) HDFS-4971. Move IO operations out of locking in OpenFileCtx. Contributed by Jing Zhao and Brandon Li. (jing9: http://svn.apache.org/viewcvs.cgi/?root=Apache-SVN&view=rev&rev=1525681) * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/AsyncDataService.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/OffsetRange.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/OpenFileCtx.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/WriteCtx.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/WriteManager.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-

nfs/src/test/java/org/apache/hadoop/hdfs/nfs/nfs3/TestOffsetRange.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt

21. SUCCESS: Integrated in Hadoop-Hdfs-trunk #1532 (See [https://builds.apache.org/job/Hadoop-Hdfs-trunk/1532/]) HDFS-4971. Move IO operations out of locking in OpenFileCtx. Contributed by Jing Zhao and Brandon Li. (jing9: http://svn.apache.org/viewcvs.cgi/?root=Apache-SVN&view=rev&rev=1525681) * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/AsyncDataService.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/OffsetRange.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/OpenFileCtx.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/WriteCtx.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/main/java/org/apache/hadoop/hdfs/nfs/nfs3/WriteManager.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs-nfs/src/test/java/org/apache/hadoop/hdfs/nfs/nfs3/TestOffsetRange.java * /hadoop/common/trunk/hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt