



Hadoop HDFS / HDFS-9434

Recommission a datanode with 500k blocks may pause NN for 30 seconds

Details

Type:	Improvement	Status:	CLOSED
Priority:	Major	Resolution:	Fixed
Affects Version/s:	None	Fix Version/s:	2.8.0, 2.7.2, 2.6.3, 3.0.0-alpha1
Component/s:	namenode		
Labels:	None		
Hadoop Flags:	Reviewed		

Description

In BlockManager, processOverReplicatedBlocksOnReCommission is called within the namespace lock. There is a (not very useful) log message printed in processOverReplicatedBlock. When there is a large number of blocks stored in a storage, printing the log message for each block can pause NN to process any other operations. We did see that it could pause NN for 30 seconds for a storage with 500k blocks.

I suggest to change the log message to trace level as a quick fix.

Attachments

h9434_20151116_branch-2.6.patch	1 kB	24/Nov/15 22:22
h9434_20151116.patch	1 kB	17/Nov/15 00:24

Issue Links

relates to

[HDFS-6425](#) Large postponedMisreplicatedBlocks has impact on blockReport latency **CLOSED**

Activity

- [Tsz-wo Sze](#) added a comment - 17/Nov/15 00:24
- h9434_20151116.patch: **changes the log message to trace and makes it shorter.**
- [Xiaoyu Yao](#) added a comment - 17/Nov/15 00:49
- Thanks [szetszwo](#) for reporting the issue and posting the fix. The patch LGTM.
+1 pending Jenkins.
- [Mingliang Liu](#) added a comment - 17/Nov/15 01:20
- Thanks for working on this . The patch looks good to me. +1 (non-binding) pending on Jenkins.
- [Hadoop QA](#) added a comment - 17/Nov/15 04:37
- 1 overall**

Vote	Subsystem	Runtime	Comment
0	reexec	1m 26s	docker + precommit patch detected.
+1	@author	0m 0s	The patch does not contain any @author tags.


-1	test4tests	0m 0s	The patch doesn't appear to include any new or modified tests. Please justify why no new tests are needed for this patch. Also please list what manual steps were performed to verify this patch.
+1	mvninstall	3m 24s	trunk passed
+1	compile	0m 44s	trunk passed with JDK v1.8.0_66
+1	compile	0m 38s	trunk passed with JDK v1.7.0_79
+1	checkstyle	0m 18s	trunk passed
+1	mvnsite	0m 50s	trunk passed
+1	mvneclipse	0m 16s	trunk passed
+1	findbugs	2m 11s	trunk passed
+1	javadoc	1m 22s	trunk passed with JDK v1.8.0_66
+1	javadoc	2m 8s	trunk passed with JDK v1.7.0_79
+1	mvninstall	0m 46s	the patch passed
+1	compile	0m 41s	the patch passed with JDK v1.8.0_66
+1	javac	0m 41s	the patch passed
+1	compile	0m 37s	the patch passed with JDK v1.7.0_79
+1	javac	0m 37s	the patch passed
-1	checkstyle	0m 18s	Patch generated 1 new checkstyle issues in hadoop-hdfs-project/hadoop-hdfs (total was 161, now 161).
+1	mvnsite	0m 49s	the patch passed
+1	mvneclipse	0m 15s	the patch passed
+1	whitespace	0m 0s	Patch has no whitespace issues.
+1	findbugs	2m 25s	the patch passed
+1	javadoc	1m 18s	the patch passed with JDK v1.8.0_66
+1	javadoc	2m 5s	the patch passed with JDK v1.7.0_79
-1	unit	76m 37s	hadoop-hdfs in the patch failed with JDK v1.8.0_66.
-1	unit	73m 38s	hadoop-hdfs in the patch failed with JDK v1.7.0_79.
-1	asflicense	0m 24s	Patch generated 56 ASF License warnings.
		176m 17s	

Reason	Tests
JDK v1.8.0_66 Failed junit tests	hadoop.hdfs.server.datanode.TestBlockScanner
	hadoop.hdfs.server.namenode.ha.TestEditLogTailer
	hadoop.hdfs.TestFileAppend2
	hadoop.hdfs.TestRollingUpgrade
	hadoop.hdfs.server.datanode.TestDataNodeHotSwapVolumes
	hadoop.hdfs.server.namenode.ha.TestSeveralNameNodes
	hadoop.hdfs.qjournal.client.TestQuorumJournalManager
	hadoop.hdfs.server.datanode.fsdataset.impl.TestLazyPersistReplicaRecovery
	hadoop.hdfs.TestDFSStripedOutputStreamWithFailure190
JDK v1.7.0_79 Failed junit tests	hadoop.hdfs.server.namenode.TestNameNodeMetadataConsistency
	hadoop.hdfs.TestDFSStripedOutputStreamWithFailure170
	hadoop.hdfs.server.namenode.ha.TestRequestHedgingProxyProvider

Subsystem	Report/Notes
-----------	--------------

Docker	Image:yetus/hadoop:date2015-11-17
JIRA Patch URL	https://issues.apache.org/jira/secure/attachment/12772619/h9434_20151116.patch
JIRA Issue	HDFS-9434
Optional Tests	asflicense compile javac javadoc mvninstall mvnsite unit findbugs checkstyle
uname	Linux 75227a3aad68 3.13.0-36-lowlatency #63-Ubuntu SMP PREEMPT Wed Sep 3 21:56:12 UTC 2014 x86_64 x86_64 x86_64 GNU/Linux
Build tool	maven
Personality	/home/jenkins/jenkins-slave/workspace/PreCommit-HDFS-Build/patchprocess/apache-yetus-4659377/precommit/personality/hadoop.sh
git revision	trunk / 6502d59
findbugs	v3.0.0
checkstyle	https://builds.apache.org/job/PreCommit-HDFS-Build/13534/artifact/patchprocess/diff-checkstyle-hadoop-hdfs-project_hadoop-hdfs.txt
unit	https://builds.apache.org/job/PreCommit-HDFS-Build/13534/artifact/patchprocess/patch-unit-hadoop-hdfs-project_hadoop-hdfs-jdk1.8.0_66.txt
unit	https://builds.apache.org/job/PreCommit-HDFS-Build/13534/artifact/patchprocess/patch-unit-hadoop-hdfs-project_hadoop-hdfs-jdk1.7.0_79.txt
unit test logs	https://builds.apache.org/job/PreCommit-HDFS-Build/13534/artifact/patchprocess/patch-unit-hadoop-hdfs-project_hadoop-hdfs-jdk1.8.0_66.txt https://builds.apache.org/job/PreCommit-HDFS-Build/13534/artifact/patchprocess/patch-unit-hadoop-hdfs-project_hadoop-hdfs-jdk1.7.0_79.txt
JDK v1.7.0_79 Test Results	https://builds.apache.org/job/PreCommit-HDFS-Build/13534/testReport/
asflicense	https://builds.apache.org/job/PreCommit-HDFS-Build/13534/artifact/patchprocess/patch-asflicense-problems.txt
modules	C: hadoop-hdfs-project/hadoop-hdfs U: hadoop-hdfs-project/hadoop-hdfs
Max memory used	226MB
Powered by	Apache Yetus http://yetus.apache.org
Console output	https://builds.apache.org/job/PreCommit-HDFS-Build/13534/console

This message was automatically generated.

▼  [Tsz-wo Sze](#) added a comment - 19/Nov/15 23:01

Thanks Xiaoyu and Mingliang for reviewing the patch.

I have committed this.

▼  [Hudson](#) added a comment - 19/Nov/15 23:15

FAILURE: Integrated in Hadoop-trunk-Commit #8831 (See <https://builds.apache.org/job/Hadoop-trunk-Commit/8831/>) [HDFS-9434](#). Recommission a datanode with 500k blocks may pause NN for 30 (szetszwo: rev 8e03e855b6a0cf650f43aac47b1ec642caf493f5)

- hadoop-hdfs-project/hadoop-hdfs/src/main/java/org/apache/hadoop/hdfs/server/blockmanagement/BlockManager.java
- hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt

▼  [Hudson](#) added a comment - 20/Nov/15 00:14

FAILURE: Integrated in Hadoop-Mapreduce-trunk-Java8 #687 (See <https://builds.apache.org/job/Hadoop-Mapreduce-trunk-Java8/687/>) [HDFS-9434](#). Recommission a datanode with 500k blocks may pause NN for 30 (szetszwo: rev 8e03e855b6a0cf650f43aac47b1ec642caf493f5)

- hadoop-hdfs-project/hadoop-hdfs/src/main/java/org/apache/hadoop/hdfs/server/blockmanagement/BlockManager.java
- hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt

▼  Hudson added a comment - 20/Nov/15 00:31

FAILURE: Integrated in Hadoop-Yarn-trunk #1425 (See <https://builds.apache.org/job/Hadoop-Yarn-trunk/1425/>)
[HDFS-9434](#). Recommission a datanode with 500k blocks may pause NN for 30 (szetszwo: rev 8e03e855b6a0cf650f43aac47b1ec642caf493f5)

- hadoop-hdfs-project/hadoop-hdfs/src/main/java/org/apache/hadoop/hdfs/server/blockmanagement/BlockManager.java
- hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt

▼  Hudson added a comment - 20/Nov/15 00:35

SUCCESS: Integrated in Hadoop-Yarn-trunk-Java8 #699 (See <https://builds.apache.org/job/Hadoop-Yarn-trunk-Java8/699/>)
[HDFS-9434](#). Recommission a datanode with 500k blocks may pause NN for 30 (szetszwo: rev 8e03e855b6a0cf650f43aac47b1ec642caf493f5)

- hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt
- hadoop-hdfs-project/hadoop-hdfs/src/main/java/org/apache/hadoop/hdfs/server/blockmanagement/BlockManager.java

▼  Hudson added a comment - 20/Nov/15 00:53

FAILURE: Integrated in Hadoop-Mapreduce-trunk #2629 (See <https://builds.apache.org/job/Hadoop-Mapreduce-trunk/2629/>)
[HDFS-9434](#). Recommission a datanode with 500k blocks may pause NN for 30 (szetszwo: rev 8e03e855b6a0cf650f43aac47b1ec642caf493f5)

- hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt
- hadoop-hdfs-project/hadoop-hdfs/src/main/java/org/apache/hadoop/hdfs/server/blockmanagement/BlockManager.java

▼  Hudson added a comment - 20/Nov/15 02:22


FAILURE: Integrated in Hadoop-Hdfs-trunk-Java8 #622 (See <https://builds.apache.org/job/Hadoop-Hdfs-trunk-Java8/622/>)
[HDFS-9434](#). Recommission a datanode with 500k blocks may pause NN for 30 (szetszwo: rev 8e03e855b6a0cf650f43aac47b1ec642caf493f5)

- hadoop-hdfs-project/hadoop-hdfs/src/main/java/org/apache/hadoop/hdfs/server/blockmanagement/BlockManager.java
- hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt


▼  Hudson added a comment - 20/Nov/15 04:43

FAILURE: Integrated in Hadoop-Hdfs-trunk #2560 (See <https://builds.apache.org/job/Hadoop-Hdfs-trunk/2560/>)
[HDFS-9434](#). Recommission a datanode with 500k blocks may pause NN for 30 (szetszwo: rev 8e03e855b6a0cf650f43aac47b1ec642caf493f5)

- hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt
- hadoop-hdfs-project/hadoop-hdfs/src/main/java/org/apache/hadoop/hdfs/server/blockmanagement/BlockManager.java

▼  Sangjin Lee added a comment - 24/Nov/15 20:12


szetszwo, can this be backported to branch-2.6? It would be as good a change there too.

▼  Tsz-wo Sze added a comment - 24/Nov/15 20:41

Done. Thanks for the suggestion.

▼  Jason Darrell Lowe added a comment - 24/Nov/15 21:14

This broke the 2.6 build. The patch is assuming SLF4J but that hasn't happened on the 2.6 branch.


▼  Sangjin Lee added a comment - 24/Nov/15 21:18

szetszwo, could you post a patch that fixes the build? We should revert back to the commons logging style logging (as well as isTraceEnabled() call around it).

▼  Hudson added a comment - 24/Nov/15 21:27

FAILURE: Integrated in Hadoop-trunk-Commit #8880 (See <https://builds.apache.org/job/Hadoop-trunk-Commit/8880/>)
Move [HDFS-9434](#) to 2.6.3 in CHANGES.txt. (szetszwo: rev 56493cda04e30ab737fc6cecc8c43a87d5b006b7)

- hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt

▼  Hudson added a comment - 24/Nov/15 21:28


FAILURE: Integrated in Hadoop-Mapreduce-trunk #2658 (See <https://builds.apache.org/job/Hadoop-Mapreduce-trunk/2658/>)
Move [HDFS-9434](#) to 2.6.3 in CHANGES.txt. (szetszwo: rev 56493cda04e30ab737fc6cecc8c43a87d5b006b7)

- hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt

▼  Hudson added a comment - 24/Nov/15 21:57


FAILURE: Integrated in Hadoop-Yarn-trunk-Java8 #727 (See <https://builds.apache.org/job/Hadoop-Yarn-trunk-Java8/727/>)
Move [HDFS-9434](#) to 2.6.3 in CHANGES.txt. (szetszwo: rev 56493cda04e30ab737fc6cecc8c43a87d5b006b7)

- hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt

▼  Hudson added a comment - 24/Nov/15 22:07


FAILURE: Integrated in Hadoop-Mapreduce-trunk-Java8 #717 (See <https://builds.apache.org/job/Hadoop-Mapreduce-trunk-Java8/717/>)
Move [HDFS-9434](#) to 2.6.3 in CHANGES.txt. (szetszwo: rev 56493cda04e30ab737fc6cecc8c43a87d5b006b7)

- hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt


▼  Tsz-wo Sze added a comment - 24/Nov/15 22:09

> This broke the 2.6 build. ...


Reverting it now.

▼  Tsz-wo Sze added a comment - 24/Nov/15 22:17


Reopen for backporting to branch-2.6

▼  Tsz-wo Sze added a comment - 24/Nov/15 22:22

h9434_20151116_branch-2.6.patch: for 2.6.

▼  Sangjin Lee added a comment - 24/Nov/15 22:27

+1. LGTM.

▼  Tsz-wo Sze added a comment - 24/Nov/15 22:37

Sangjin, thanks for the review.

I have committed the branch-2.6 patch.

▼  Hudson added a comment - 25/Nov/15 00:15

FAILURE: Integrated in Hadoop-Yarn-trunk #1450 (See <https://builds.apache.org/job/Hadoop-Yarn-trunk/1450/>)
Move [HDFS-9434](#) to 2.6.3 in CHANGES.txt. (szetszwo: rev 56493cda04e30ab737fc6cecc8c43a87d5b006b7)

- hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt

▼  Hudson added a comment - 25/Nov/15 00:26

FAILURE: Integrated in Hadoop-Hdfs-trunk-Java8 #639 (See <https://builds.apache.org/job/Hadoop-Hdfs-trunk-Java8/639/>)
Move [HDFS-9434](#) to 2.6.3 in CHANGES.txt. (szetszwo: rev 56493cda04e30ab737fc6cecc8c43a87d5b006b7)

- hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt


▼  Hudson added a comment - 25/Nov/15 00:52

FAILURE: Integrated in Hadoop-Hdfs-trunk #2576 (See <https://builds.apache.org/job/Hadoop-Hdfs-trunk/2576/>)
Move [HDFS-9434](#) to 2.6.3 in CHANGES.txt. (szetszwo: rev 56493cda04e30ab737fc6cecc8c43a87d5b006b7)


- hadoop-hdfs-project/hadoop-hdfs/CHANGES.txt

▼ People

Assignee:

 Tsz-wo Sze

Reporter:

 Tsz-wo Sze

Votes:

 0 Vote for this issue

Watchers:

 16 Start watching this issue

▼ **Dates**

Created:

17/Nov/15 00:13

Updated:

06/Jan/17 07:38

Resolved:

24/Nov/15 22:37