# Issue 4819:

**Problem statement :** Namenode crashes every 4hrs .

**Observation :** Namenode dies whenever application team executes there jobs . The jobs are batch jobs and scheduled approximately every 4hrs .

Observed issue in the namenode logs :

There has been 2 occurance reported by Shashang during the morning call . Both the occurance has been captured below and scenario has been explained below .

## Namenode Log :

First Occurance :

at java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1145)

at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:615)

at java.lang.Thread.run(Thread.java:744)

2017-10-12 00:17:24,208 WARN security.Groups (Groups.java:fetchGroupList(361)) - Potential performance problem: getGroups(user=hsgctcrp) took 145673 milliseconds.

2017-10-12 00:17:24,209 FATAL namenode.FSEditLog (JournalSet.java:mapJournalsAndReportErrors(398)) - Error: flush failed for required journal (JournalAndStream(mgr=QJM to [172.20.17

8.13:8485, 172.20.178.14:8485, 172.20.178.100:8485], stream=QuorumOutputStream starting at txid 2403281340))

org.apache.hadoop.hdfs.qjournal.client.QuorumException: Got too many exceptions to achieve quorum size 2/3. 3 exceptions thrown:

172.20.178.100:8485: IPC's epoch 190 is less than the last promised epoch 191

at org.apache.hadoop.hdfs.qjournal.server.Journal.checkRequest(Journal.java:428)

at org.apache.hadoop.hdfs.qjournal.server.Journal.checkWriteRequest(Journal.java:456)

at org.apache.hadoop.hdfs.qjournal.server.Journal.journal(Journal.java:351)

at org.apache.hadoop.hdfs.qjournal.server.JournalNodeRpcServer.journal(JournalNodeRpcServer.java:152)

at org.apache.hadoop.hdfs.qjournal.protocolPB.QJournalProtocolServerSideTranslatorPB.journal(QJournalProtocolServerSideTranslatorPB.java:158)

at org.apache.hadoop.hdfs.qjournal.protocol.QJournalProtocolProtos$QJournalProtocolService$2.callBlockingMethod(QJournalProtocolProtos.java:25421)

at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:640)

at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:982)

at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2313)

at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2309)

at java.security.AccessController.doPrivileged(Native Method)

at javax.security.auth.Subject.doAs(Subject.java:415)

at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1724)

at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2307)

172.20.178.13:8485: IPC's epoch 190 is less than the last promised epoch 191

at org.apache.hadoop.hdfs.qjournal.server.Journal.checkRequest(Journal.java:428)

at org.apache.hadoop.hdfs.qjournal.server.Journal.checkWriteRequest(Journal.java:456)

at org.apache.hadoop.hdfs.qjournal.server.Journal.journal(Journal.java:351)

at org.apache.hadoop.hdfs.qjournal.server.JournalNodeRpcServer.journal(JournalNodeRpcServer.java:152)

:

Next occurance :

2017-10-12 08:18:15,102 INFO authorize.ServiceAuthorizationManager (ServiceAuthorizationManager.java:authorize(137)) - Authorization successful for uxqvadm (auth:TOKEN) via hive/am

1plccmrhdn04.r1-core.r1.aig.net@R1-CORE.R1.AIG.NET (auth:TOKEN) for protocol=interface org.apache.hadoop.hdfs.protocol.ClientProtocol

2017-10-12 08:18:16,028 WARN security.Groups (Groups.java:fetchGroupList(361)) - Potential performance problem: getGroups(user=solr) took 80422 milliseconds.

2017-10-12 08:18:29,917 INFO provider.BaseAuditHandler (BaseAuditHandler.java:logStatus(312)) - Audit Status Log: name=hdfs.async.batch, finalDestination=hdfs.async.batch.hdfs, int

erval=01:02.217 minutes, events=110904, succcessCount=986, totalEvents=202987349, totalSuccessCount=801164

2017-10-12 08:20:15,551 WARN security.Groups (Groups.java:fetchGroupList(361)) - Potential performance problem: getGroups(user=uxqvadm) took 120448 milliseconds.

2017-10-12 08:20:15,560 WARN security.Groups (Groups.java:fetchGroupList(361)) - Potential performance problem: getGroups(user=hsgctcrp) took 156920 milliseconds.

2017-10-12 08:20:15,560 INFO authorize.ServiceAuthorizationManager (ServiceAuthorizationManager.java:authorize(137)) - Authorization successful for uxqvadm (auth:TOKEN) via hive/am

1plccmrhdn04.r1-core.r1.aig.net@R1-CORE.R1.AIG.NET (auth:TOKEN) for protocol=interface org.apache.hadoop.hdfs.protocol.ClientProtocol

2017-10-12 08:20:15,561 INFO namenode.FSEditLog (FSEditLog.java:printStatistics(716)) - Number of transactions: 4914 Total time for transactions(ms): 166 Number of transactions bat

ched in Syncs: 21 Number of syncs: 3286 SyncTimes(ms): 3218 775

2017-10-12 08:20:15,563 WARN ipc.Server (Server.java:processResponse(1245)) - IPC Server handler 250 on 8020, call org.apache.hadoop.hdfs.protocol.ClientProtocol.getFileInfo from 1

72.20.178.115:41324 Call#16 Retry#0: output error

2017-10-12 08:20:15,563 WARN ipc.Server (Server.java:processResponse(1245)) - IPC Server handler 540 on 8020, call org.apache.hadoop.hdfs.protocol.ClientProtocol.getFileInfo from 1

72.20.177.109:46132 Call#14 Retry#0: output error

2017-10-12 08:20:15,563 INFO ipc.Server (Server.java:run(2368)) - IPC Server handler 250 on 8020 caught an exception

java.nio.channels.ClosedChannelException

at sun.nio.ch.SocketChannelImpl.ensureWriteOpen(SocketChannelImpl.java:265)

at sun.nio.ch.SocketChannelImpl.write(SocketChannelImpl.java:474)

at org.apache.hadoop.ipc.Server.channelWrite(Server.java:2869)

at org.apache.hadoop.ipc.Server.access$2200(Server.java:136)

at org.apache.hadoop.ipc.Server$Responder.processResponse(Server.java:1195)

at org.apache.hadoop.ipc.Server$Responder.doRespond(Server.java:1260)

at org.apache.hadoop.ipc.Server$Connection.sendResponse(Server.java:2228)

at org.apache.hadoop.ipc.Server$Connection.access$600(Server.java:1340)

at org.apache.hadoop.ipc.Server$Call.sendResponse(Server.java:739)

at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2357)

2017-10-12 08:20:15,564 WARN ipc.Server (Server.java:processResponse(1245)) - IPC Server handler 85 on 8020, call org.apache.hadoop.ha.HAServiceProtocol.getServiceStatus from 172.2

0.178.11:42909 Call#592234 Retry#0: output error

2017-10-12 08:20:15,564 WARN ipc.Server (Server.java:processResponse(1245)) - Socket Reader #1 for port 8020, call null from 172.20.178.12:40696 Call#-33 Retry#-1: output error

2017-10-12 08:20:15,563 INFO ipc.Server (Server.java:run(2368)) - IPC Server handler 540 on 8020 caught an exception

java.nio.channels.ClosedChannelException

at sun.nio.ch.SocketChannelImpl.ensureWriteOpen(SocketChannelImpl.java:265)

at sun.nio.ch.SocketChannelImpl.write(SocketChannelImpl.java:474)

at org.apache.hadoop.ipc.Server.channelWrite(Server.java:2869)

at org.apache.hadoop.ipc.Server.access$2200(Server.java:136)

at org.apache.hadoop.ipc.Server$Responder.processResponse(Server.java:1195)

at org.apache.hadoop.ipc.Server$Responder.doRespond(Server.java:1260)

:

Journal logs & ZKFC show the errors around similar timelines .

## Description of Cause :

In a distributed architecture, resolving user group membership generate unexpectedly high load on LDAP servers .

In case of the Hadoop cluster the user group membership resolution happens on the namenode and slow response from ldap/AD( directory service ) server may result in timeouts .

Usually in such case below warning message is thrown :

security.Groups (Groups.java:getGroups(181)) - Potential performance problem: getGroups……

The impact of such slow resolution may result in to exceed the NameNode's timeout for JournalNode calls.

The NameNode must be able to log edits to a quorum of JournalNodes (i.e. 2 out of 3 JournalNodes). If the calls time out to 2 or more JournalNodes, then it's a fatal condition. The NameNode must be able to log transactions successfully, and if it fails, then it aborts intentionally.

This condition would trigger an unwanted HA failover. The problem might reoccur after failover, resulting in flapping. If this happens, then the JournalNode logs will show the "performance problem" mentioned above, and the NameNode logs will show a message about "Timed out waiting for a quorum of nodes to respond" before a FATAL shutdown error.

Please note as in the above highlighted logs in red , the time to resolve the groups has been 2 minutes plus which results in the above mentioned scenario and the namenode dies.

Please refer to the below reference link for details of the situation . Section “Load Patterns” describes the complete scenario in detail

<https://community.hortonworks.com/articles/38591/hadoop-and-ldap-usage-load-patterns-and-tuning.html>

Also , there are suggestions for tuning . Please discuss which might apply best to your current scenario .

## References :

* <https://community.hortonworks.com/questions/41255/how-to-debug-the-issue-ipcs-epoch-x-is-less-than-t.html>
* <https://community.hortonworks.com/articles/38591/hadoop-and-ldap-usage-load-patterns-and-tuning.html>