[ZooKeeper-1496](https://www.google.com/url?q=https://issues.apache.org/jira/browse/ZOOKEEPER-1496&sa=D&source=editors&ust=1751358799115297&usg=AOvVaw3YGGbWvHZgtoHsVKKJSLEu)

Ephemeral node not getting cleared even after client has exited

**(1)  Log information**

**(1.1) Roles in this case**

zkClient        zkServer-91 -> Leader            zkServer-55 -> Follower                  zkServer-102 -> Follower

**(1.2) Symptoms**

Environment:

“Hadoop setup. We were using Namenode HA with bookkeeper as shared storage and auto failover enabled. NN102 was active and NN55 was standby. FailoverController at 102 got shut down due to ZK connection error. The lock-*ActiveStandbyElectorLock* created (ephemeral node) by this failovercontroller is not cleared from ZK”.

Background: Hadoop HA uses Zookeeper to coordinate and share states. Name node HA uses an active/standby mode so that there is only one active NN at any given time. To compete for the active NN among multiple hosts, an ActiveStandbyElector is designed to integrate with zookeeper to achieve this goal. ActiveStandbyElector is more like a zookeeper handler working between zookeeper and the controller such as ZKFailoverControllerc(ZKFC).

The main idea is to create an ephemeral node*"ActiveStandbyElectorLock"* and whoever creates this node successfully becomes active. Since this node is ephemeral, the current active one could lose this node if the session expires or connection is lost so that another node has an opportunity to create the same node and become active.

Based on the environment and background above, the symptom should be: in Hadoop cluster, the active NN 102 does not work, and the standby NN 55 cannot become active (due to the remaining ephemeral node), so the HDFS is not working.

The zookeeper logs show

2015-02-24 18:01:05,616 INFO org.apache.zookeeper.server.PrepRequestProcessor: Got user-level KeeperException when processing sessionid:0x34bb91dede303c9 type:create cxid:0x1 zxid:0xe00002ce3 txntype:-1 reqpath:n/a Error Path:/hadoop-ha/mynameservice/ActiveStandbyElectorLock Error:KeeperErrorCode = NodeExists for /hadoop-ha/mynameservice/ActiveStandbyElectorLock

2012-06-26 13:10:12,566 [myid:3] - DEBUG [ProcessThread(sid:3 cport:-1)::CommitProcessor@171] - Processing request:: sessionid:0x1382791d4e50004 type:createSession cxid:0x0 zxid:0x200000070 txntype:-10 reqpath:n/a

2012-06-26 13:10:13,001 [myid:3] - DEBUG [ProcessThread(sid:3 cport:-1)::Leader@716] - Proposing:: sessionid:0x1382791d4e50004 type:createSession cxid:0x0 zxid:0x200000070 txntype:-10 reqpath:n/a

2012-06-26 13:10:13,202 [myid:3] - DEBUG [LearnerHandler-/xx.xx.xx.102:13846:CommitProcessor@161] - Committing request:: sessionid:0x1382791d4e50004 type:createSession cxid:0x0 zxid:0x200000070 txntype:-10 reqpath:n/a

2012-06-26 13:10:18,653 [myid:3] - INFO [SessionTracker:ZooKeeperServer@325] - Expiring session 0x1382791d4e50004, timeout of 5000ms exceeded

2012-06-26 13:10:18,803 [myid:3] - INFO [ProcessThread(sid:3 cport:-1)::PrepRequestProcessor@627] - Got user-level KeeperException when processing sessionid:0x1382791d4e50004 type:create cxid:0x1 zxid:0x200000072 txntype:-1 reqpath:/hadoop-ha/hacluster/ActiveStandbyElectorLock Error Path:null Error:KeeperErrorCode = Session expired

2012-06-26 13:10:18,834 [myid:3] - INFO [ProcessThread(sid:3 cport:-1)::PrepRequestProcessor@476] - Processed session termination for sessionid: 0x1382791d4e50004

2012-06-26 13:10:18,834 [myid:3] - DEBUG [ProcessThread(sid:3 cport:-1)::CommitProcessor@171] - Processing request:: sessionid:0x1382791d4e50004 type:closeSession cxid:0x0 zxid:0x200000074 txntype:-11 reqpath:n/a

2012-06-26 13:10:19,892 [myid:3] - DEBUG [ProcessThread(sid:3 cport:-1)::Leader@716] - Proposing:: sessionid:0x1382791d4e50004 type:closeSession cxid:0x0 zxid:0x200000074 txntype:-11 reqpath:n/a

2012-06-26 13:10:19,919 [myid:3] - DEBUG [LearnerHandler-/xx.xx.xx.102:13846:CommitProcessor@161] - Committing request:: sessionid:0x1382791d4e50004 type:closeSession cxid:0x0 zxid:0x200000074 txntype:-11 reqpath:n/a

2012-06-26 13:10:20,608 [myid:3] - DEBUG [CommitProcessor:3:FinalRequestProcessor@88] - Processing request:: sessionid:0x1382791d4e50004 type:closeSession cxid:0x0 zxid:0x200000074 txntype:-11 reqpath:n/a

HOSt-xx-xx-xx-91:/home/Jun25\_LR/install/zookeeper/logs # grep -E "/hadoop-ha/hacluster/ActiveStandbyElectorLock|0x1382791d4e50004" \*|grep 0x200000075

2012-06-26 13:10:19,893 [myid:3] - DEBUG [ProcessThread(sid:3 cport:-1)::CommitProcessor@171] - Processing request:: sessionid:0x1382791d4e50004 type:create cxid:0x2 zxid:0x200000075 txntype:1 reqpath:n/a

2012-06-26 13:10:19,920 [myid:3] - DEBUG [ProcessThread(sid:3 cport:-1)::Leader@716] - Proposing:: sessionid:0x1382791d4e50004 type:create cxid:0x2 zxid:0x200000075 txntype:1 reqpath:n/a

2012-06-26 13:10:20,278 [myid:3] - DEBUG [LearnerHandler-/xx.xx.xx.102:13846:CommitProcessor@161] - Committing request:: sessionid:0x1382791d4e50004 type:create cxid:0x2 zxid:0x200000075 txntype:1 reqpath:n/a

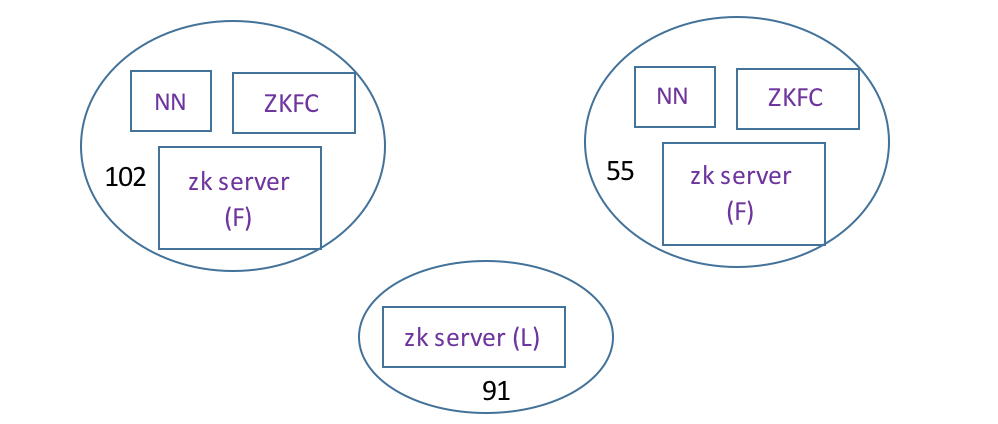
2012-06-26 13:10:20,608 [myid:3] - DEBUG [CommitProcessor:3:FinalRequestProcessor@88] - Processing request:: sessionid:0x1382791d4e50004 type:closeSession cxid:0x0 zxid:0x200000074 txntype:-11 reqpath:n/a

2012-06-26 13:10:20,752 [myid:3] - DEBUG [CommitProcessor:3:FinalRequestProcessor@88] - Processing request:: sessionid:0x1382791d4e50004 type:create cxid:0x2 zxid:0x200000075 txntype:1 reqpath:n/a

**Environment details:**

ZK-91 -> Leader                 ZK-55 -> Follower              ZK-102 -> Follower

ZKClient session timeout -> 5seconds



When the user connects to the zk server, and checks ActiveStandbyElectorLock, the result is:

[zk: xx.xx.xx.55:2182(CONNECTED) 0] get /hadoop-ha/hacluster/ActiveStandbyElectorLock

haclusternn2HOSt-xx-xx-xx-102 ï¿½ï¿½

cZxid = 0x200000075

ctime = Tue Jun 26 13:10:19 IST 2012

mZxid = 0x200000075

mtime = Tue Jun 26 13:10:19 IST 2012

pZxid = 0x200000075

cversion = 0

dataVersion = 0

aclVersion = 0

ephemeralOwner = 0x1382791d4e50004

dataLength = 42

numChildren = 0

[zk: xx.xx.xx.55:2182(CONNECTED) 1]

The result shows that the ephemeral node ActiveStandbyElectorLock exists, and it belongs to session 0x1382791d4e50004. The transaction id for creating this ephemeral node is cZxid = 0x200000075