[HBase-9593](https://www.google.com/url?q=https://issues.apache.org/jira/browse/HBASE-9593&sa=D&source=editors&ust=1751363081497045&usg=AOvVaw0QffCBrfTONY1_QNAI1AQn)

Region server left in online servers list forever if it went down after registering to master and before creating ephemeral node

**(1)  Log information**

**(1.1) Roles in this case**

HRegionServer (slave)     HMaster(master)               ZooKeeper sends notification to HMaster

HRegionServer (client)     ZooKeeper (server)

**(1.2) Symptoms**

In some of our tests we found that RegionServer always showing online in master UI but its actually dead.

If RegionServer went down in the middle of following steps then the region server will always show in master online servers list.

1) register to master

2) create ephemeral znode (on ZK)

The logs are as follows:

1) RegionServer HOST-10-18-40-153 is registered to master. (HMaster log)

2013-09-19 19:47:41,123 DEBUG org.apache.hadoop.hbase.master.ServerManager: STARTUP: Server HOST-10-18-40-153,61020,1379600260255 came back up, removed it from the dead servers list

2013-09-19 19:47:41,123 INFO org.apache.hadoop.hbase.master.ServerManager: Registering server=HOST-10-18-40-153,61020,1379600260255

(RegionServer log)

2013-09-19 19:47:41,119 INFO org.apache.hadoop.hbase.regionserver.HRegionServer: Connected to master at HOST-10-18-40-153/10.18.40.153:61000

2013-09-19 19:47:41,119 INFO org.apache.hadoop.hbase.regionserver.HRegionServer: Telling master at HOST-10-18-40-153,61000,1379600055284 that we are up with port=61020, startcode=1379600260255

2) Terminated before creating ephemeral node.   (RegionServer log)

Thu Sep 19 19:47:41 IST 2013 Terminating regionserver

3) Since the HMaster doesn’t know the RS’s down (the RS doesn’t create ephemeral znode in ZooKeeper, so no notification will be sent from ZK), the RS can be selected for assignment by HMaster and it will fail. (HMaster log)

When HMaster sends an OPEN RPC to the specified RS(HOST-10-18-40-153,61020) to open the specified region(ROOT) by calling function ServerManager.sendRegionOpen(), it fails to set up the connection, and ConnectException happens for the first attempt of assignment (retry=0).

2013-09-19 19:47:54,049 WARN org.apache.hadoop.hbase.master.AssignmentManager: Failed assignment of -ROOT-,,0.70236052 to HOST-10-18-40-153,61020,1379600260255, trying to assign elsewhere instead; retry=0

java.net.ConnectException: Connection refused

         at sun.nio.ch.SocketChannelImpl.checkConnect(Native Method)

         at sun.nio.ch.SocketChannelImpl.finishConnect(SocketChannelImpl.java:567)

         at org.apache.hadoop.net.SocketIOWithTimeout.connect(SocketIOWithTimeout.java:206)

         at org.apache.hadoop.net.NetUtils.connect(NetUtils.java:529)

         at org.apache.hadoop.net.NetUtils.connect(NetUtils.java:493)

         at org.apache.hadoop.hbase.ipc.HBaseClient$Connection.setupConnection(HBaseClient.java:390)

         at org.apache.hadoop.hbase.ipc.HBaseClient$Connection.setupIOstreams(HBaseClient.java:436)

         at org.apache.hadoop.hbase.ipc.HBaseClient.getConnection(HBaseClient.java:1127)

         at org.apache.hadoop.hbase.ipc.HBaseClient.call(HBaseClient.java:974)

         at org.apache.hadoop.hbase.ipc.WritableRpcEngine$Invoker.invoke(WritableRpcEngine.java:86)

         at $Proxy15.openRegion(Unknown Source)

         at org.apache.hadoop.hbase.master.ServerManager.sendRegionOpen(ServerManager.java:533)

         at org.apache.hadoop.hbase.master.AssignmentManager.assign(AssignmentManager.java:1734)

         at org.apache.hadoop.hbase.master.AssignmentManager.assign(AssignmentManager.java:1431)

         at org.apache.hadoop.hbase.master.AssignmentManager.assign(AssignmentManager.java:1406)

         at org.apache.hadoop.hbase.master.AssignmentManager.assign(AssignmentManager.java:1401)

         at org.apache.hadoop.hbase.master.AssignmentManager.assignRoot(AssignmentManager.java:2374)

         at org.apache.hadoop.hbase.master.handler.MetaServerShutdownHandler.verifyAndAssignRoot(MetaServerShutdownHandler.java:136)

         at org.apache.hadoop.hbase.master.handler.MetaServerShutdownHandler.verifyAndAssignRootWithRetries(MetaServerShutdownHandler.java:160)

         at org.apache.hadoop.hbase.master.handler.MetaServerShutdownHandler.process(MetaServerShutdownHandler.java:82)

         at org.apache.hadoop.hbase.executor.EventHandler.run(EventHandler.java:175)

         at java.util.concurrent.ThreadPoolExecutor$Worker.runTask(ThreadPoolExecutor.java:886)

         at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:908)

         at java.lang.Thread.run(Thread.java:662)

2013-09-19 19:47:54,050 DEBUG org.apache.hadoop.hbase.master.AssignmentManager: Found an existing plan for -ROOT-,,0.70236052 destination server is HOST-10-18-40-153,61020,1379600260255

2013-09-19 19:47:54,050 DEBUG org.apache.hadoop.hbase.master.AssignmentManager: No previous transition plan was found (or we are ignoring an existing plan) for -ROOT-,,0.70236052 so generated a random one; hri=-ROOT-,,0.70236052, src=, dest=HOST-10-18-40-153,61020,1379600260255; 1 (online=1, available=1) available servers

2013-09-19 19:47:54,050 DEBUG org.apache.hadoop.hbase.zookeeper.ZKAssign: master:61000-0x14135a277ff017d Creating (or updating) unassigned node for 70236052 with OFFLINE state

2013-09-19 19:47:54,070 DEBUG org.apache.hadoop.hbase.master.AssignmentManager: Handling transition=M\_ZK\_REGION\_OFFLINE, server=HOST-10-18-40-153,61000,1379600055284, region=70236052/-ROOT-

2013-09-19 19:47:54,071 DEBUG org.apache.hadoop.hbase.master.AssignmentManager: Found an existing plan for -ROOT-,,0.70236052 destination server is HOST-10-18-40-153,61020,1379600260255

2013-09-19 19:47:54,071 DEBUG org.apache.hadoop.hbase.master.AssignmentManager: Using pre-existing plan for region -ROOT-,,0.70236052; plan=hri=-ROOT-,,0.70236052, src=, dest=HOST-10-18-40-153,61020,1379600260255

2013-09-19 19:47:54,071 DEBUG org.apache.hadoop.hbase.master.AssignmentManager: Assigning region -ROOT-,,0.70236052 to HOST-10-18-40-153,61020,1379600260255

2013-09-19 19:47:54,072 WARN org.apache.hadoop.hbase.master.AssignmentManager: Failed assignment of -ROOT-,,0.70236052 to HOST-10-18-40-153,61020,1379600260255, trying to assign elsewhere instead; retry=1

org.apache.hadoop.hbase.ipc.HBaseClient$FailedServerException: This server is in the failed servers list: HOST-10-18-40-153/10.18.40.153:61020

         at org.apache.hadoop.hbase.ipc.HBaseClient$Connection.setupIOstreams(HBaseClient.java:425)

         at org.apache.hadoop.hbase.ipc.HBaseClient.getConnection(HBaseClient.java:1127)

         at org.apache.hadoop.hbase.ipc.HBaseClient.call(HBaseClient.java:974)

         at org.apache.hadoop.hbase.ipc.WritableRpcEngine$Invoker.invoke(WritableRpcEngine.java:86)

         at $Proxy15.openRegion(Unknown Source)

         at org.apache.hadoop.hbase.master.ServerManager.sendRegionOpen(ServerManager.java:533)

         at org.apache.hadoop.hbase.master.AssignmentManager.assign(AssignmentManager.java:1734)

         at org.apache.hadoop.hbase.master.AssignmentManager.assign(AssignmentManager.java:1431)

         at org.apache.hadoop.hbase.master.AssignmentManager.assign(AssignmentManager.java:1406)

         at org.apache.hadoop.hbase.master.AssignmentManager.assign(AssignmentManager.java:1401)

         at org.apache.hadoop.hbase.master.AssignmentManager.assignRoot(AssignmentManager.java:2374)

         at org.apache.hadoop.hbase.master.handler.MetaServerShutdownHandler.verifyAndAssignRoot(MetaServerShutdownHandler.java:136)

         at org.apache.hadoop.hbase.master.handler.MetaServerShutdownHandler.verifyAndAssignRootWithRetries(MetaServerShutdownHandler.java:160)

         at org.apache.hadoop.hbase.master.handler.MetaServerShutdownHandler.process(MetaServerShutdownHandler.java:82)

         at org.apache.hadoop.hbase.executor.EventHandler.run(EventHandler.java:175)

         at java.util.concurrent.ThreadPoolExecutor$Worker.runTask(ThreadPoolExecutor.java:886)

         at java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:908)

         at java.lang.Thread.run(Thread.java:662)

2013-09-19 19:47:54,072 DEBUG org.apache.hadoop.hbase.master.AssignmentManager: Found an existing plan for -ROOT-,,0.70236052 destination server is HOST-10-18-40-153,61020,1379600260255

…                       (keep assigning the ROOT region to the current RS)

Then HMaster attempts to assign the ROOT region to the same RS for the second time(retry=1), and fails with “FailedServerException: This server is in the failed servers list”.

HMaster will keep assigning the ROOT region to the current RS until the attempt times equals the maximumAssignmentAttempts(10) in assign(), and it will also fail in the following attempts with FailedServerException.

When finish the all assignment attempts for this terminated server, HMaster will select the other available servers to assign the ROOT region.