[HBase-7799](https://www.google.com/url?q=https://issues.apache.org/jira/browse/HBASE-7799&sa=D&source=editors&ust=1751363081525082&usg=AOvVaw3sfrubib-AOYDcG58uX4yx)

reassigning region stuck in open still may not work correctly due to leftover ZK node

**(1)  Log information**

**(1.1) Roles in this case**

HMaster        keeps sendRegionOpen to       RS

RS     fails to open the region due to the wrong znode state in    ZooKeeper

**(1.2) Symptoms**

When the region 871d1c3bdf98a2c93b527cb6cc61327d is being opened on RegionServer 10.11.2.92,64485,1360362800564, the RS dies, and the current state for the region is “RS\_ZK\_REGION\_OPENING”.

HMaster#ServerShutdownHandler (SSH) is responsible for dealing with this dead RS.

**①**HMaster tries to force this region offline and reassign it:

2013-02-08 14:35:06,500 INFO  [MASTER\_SERVER\_OPERATIONS-10.11.2.92,64483,1360362800340-1] master.RegionStates(347): Found opening region {IntegrationTestRebalanceAndKillServersTargeted,7333332c,1360362805563.871d1c3bdf98a2c93b527cb6cc61327d. state=OPENING, ts=1360362901596, server=10.11.2.92,64485,1360362800564} to be reassigned by SSH for 10.11.2.92,64485,1360362800564

2013-02-08 14:35:06,500 INFO  [MASTER\_SERVER\_OPERATIONS-10.11.2.92,64483,1360362800340-1] master.RegionStates(242): Region {NAME => 'IntegrationTestRebalanceAndKillServersTargeted,7333332c,1360362805563.871d1c3bdf98a2c93b527cb6cc61327d.', STARTKEY => '7333332c', ENDKEY => '7ffffff8', ENCODED => 871d1c3bdf98a2c93b527cb6cc61327d,} transitioned from {IntegrationTestRebalanceAndKillServersTargeted,7333332c,1360362805563.871d1c3bdf98a2c93b527cb6cc61327d. state=OPENING, ts=1360362901596, server=10.11.2.92,64485,1360362800564} to {IntegrationTestRebalanceAndKillServersTargeted,7333332c,1360362805563.871d1c3bdf98a2c93b527cb6cc61327d. state=CLOSED, ts=1360362906500, server=null}

2013-02-08 14:35:06,505 DEBUG [10.11.2.92,64483,1360362800340-GeneralBulkAssigner-1] master.AssignmentManager(1530): Forcing OFFLINE; was={IntegrationTestRebalanceAndKillServersTargeted,7333332c,1360362805563.871d1c3bdf98a2c93b527cb6cc61327d. state=CLOSED, ts=1360362906500, server=null}

...

**②**When reassigning this region, HMaster fails to asynchronously create the znode with OFFLINE state, because the node already exists, which indicates that the original znode is not deleted.

2013-02-08 14:35:06,506 DEBUG [10.11.2.92,64483,1360362800340-GeneralBulkAssigner-1] zookeeper.ZKAssign(176): master:64483-0x13cbbf1025d0000 Async create of unassigned node for 871d1c3bdf98a2c93b527cb6cc61327d with OFFLINE state

2013-02-08 14:35:06,509 WARN  [main-EventThread] master.OfflineCallback(59): Node for /hbase/region-in-transition/871d1c3bdf98a2c93b527cb6cc61327d already exists

2013-02-08 14:35:06,509 DEBUG [main-EventThread] master.OfflineCallback(69): rs={IntegrationTestRebalanceAndKillServersTargeted,7333332c,1360362805563.871d1c3bdf98a2c93b527cb6cc61327d. state=OFFLINE, ts=1360362906506, server=null}, server=10.11.2.92,64488,1360362800651

2013-02-08 14:35:06,512 DEBUG [main-EventThread] master.OfflineCallback$ExistCallback(106): rs={IntegrationTestRebalanceAndKillServersTargeted,7333332c,1360362805563.871d1c3bdf98a2c93b527cb6cc61327d. state=OFFLINE, ts=1360362906506, server=null}, server=10.11.2.92,64488,1360362800651

...

**③**When RS receives the open region request, it fails to transit the znode from M\_ZK\_REGION\_OFFLINE to RS\_ZK\_REGION\_OPENING since the znode state is RS\_ZK\_REGION\_OPENING which is set by the previous RS. Then transition from M\_ZK\_REGION\_OFFLINE to RS\_ZK\_REGION\_FAILED\_OPEN also fails due to the same reason.

So it goes into infinite cycle of failing to assign:

2013-02-08 14:35:06,517 INFO  [PRI IPC Server handler 7 on 64488] regionserver.HRegionServer(3435): Received request to open region: IntegrationTestRebalanceAndKillServersTargeted,7333332c,1360362805563.871d1c3bdf98a2c93b527cb6cc61327d. on 10.11.2.92,64488,1360362800651

...

2013-02-08 14:35:06,521 WARN  [RS\_OPEN\_REGION-10.11.2.92,64488,1360362800651-0] zookeeper.ZKAssign(762): regionserver:64488-0x13cbbf1025d0004 Attempt to transition the unassigned node for 871d1c3bdf98a2c93b527cb6cc61327d from M\_ZK\_REGION\_OFFLINE to RS\_ZK\_REGION\_OPENING failed, the node existed but was in the state RS\_ZK\_REGION\_OPENING set by the server 10.11.2.92,64488,1360362800651           [wrong server name, see HBASE-7798: should be the original RS, rather than new RS]

...

2013-02-08 14:35:06,528 WARN  [RS\_OPEN\_REGION-10.11.2.92,64488,1360362800651-0] zookeeper.ZKAssign(762): regionserver:64488-0x13cbbf1025d0004 Attempt to transition the unassigned node for 871d1c3bdf98a2c93b527cb6cc61327d from M\_ZK\_REGION\_OFFLINE to RS\_ZK\_REGION\_FAILED\_OPEN failed, the node existed but was in the state RS\_ZK\_REGION\_OPENING set by the server 10.11.2.92,64488,1360362800651

...

...

2013-02-08 14:36:09,644 INFO  [PRI IPC Server handler 3 on 65343] regionserver.HRegionServer(3435): Received request to open region: IntegrationTestRebalanceAndKillServersTargeted,7333332c,1360362805563.871d1c3bdf98a2c93b527cb6cc61327d. on 10.11.2.92,65343,1360362907858

...

2013-02-08 14:36:09,658 WARN  [RS\_OPEN\_REGION-10.11.2.92,65343,1360362907858-0] zookeeper.ZKAssign(762): regionserver:65343-0x13cbbf1025d000f Attempt to transition the unassigned node for 871d1c3bdf98a2c93b527cb6cc61327d from M\_ZK\_REGION\_OFFLINE to RS\_ZK\_REGION\_OPENING failed, the node existed but was in the state RS\_ZK\_REGION\_OPENING set by the server 10.11.2.92,65343,1360362907858

...

2013-02-08 14:36:09,662 WARN  [RS\_OPEN\_REGION-10.11.2.92,65343,1360362907858-0] zookeeper.ZKAssign(762): regionserver:65343-0x13cbbf1025d000f Attempt to transition the unassigned node for 871d1c3bdf98a2c93b527cb6cc61327d from M\_ZK\_REGION\_OFFLINE to RS\_ZK\_REGION\_FAILED\_OPEN failed, the node existed but was in the state RS\_ZK\_REGION\_OPENING set by the server 10.11.2.92,65343,1360362907858

...

...

2013-02-08 14:38:17,696 INFO  [PRI IPC Server handler 6 on 49317] regionserver.HRegionServer(3435): Received request to open region: IntegrationTestRebalanceAndKillServersTargeted,7333332c,1360362805563.871d1c3bdf98a2c93b527cb6cc61327d. on 10.11.2.92,49317,1360362971256

...

2013-02-08 14:38:17,713 WARN  [RS\_OPEN\_REGION-10.11.2.92,49317,1360362971256-2] zookeeper.ZKAssign(762): regionserver:49317-0x13cbbf1025d0011 Attempt to transition the unassigned node for 871d1c3bdf98a2c93b527cb6cc61327d from M\_ZK\_REGION\_OFFLINE to RS\_ZK\_REGION\_OPENING failed, the node existed but was in the state RS\_ZK\_REGION\_OPENING set by the server 10.11.2.92,49317,1360362971256

...

2013-02-08 14:38:17,720 WARN  [RS\_OPEN\_REGION-10.11.2.92,49317,1360362971256-2] zookeeper.ZKAssign(762): regionserver:49317-0x13cbbf1025d0011 Attempt to transition the unassigned node for 871d1c3bdf98a2c93b527cb6cc61327d from M\_ZK\_REGION\_OFFLINE to RS\_ZK\_REGION\_FAILED\_OPEN failed, the node existed but was in the state RS\_ZK\_REGION\_OPENING set by the server 10.11.2.92,49317,1360362971256

...

2013-02-08 14:39:22,086 DEBUG [main-EventThread] master.OfflineCallback(69): rs={IntegrationTestRebalanceAndKillServersTargeted,0ccccccc,1360362805563.f6daa3cccefad73f8141d200ce58645b. state=OFFLINE, ts=1360363161900, server=null}, server=10.11.2.92,49476,1360363036574

2013-02-08 14:39:22,090 TRACE [HBaseWriterThread\_8] client.HConnectionManager$HConnectionImplementation$Process(2171): Retrying due to errors: Error from [10.11.2.92:64486] for [ed777f62b485b3975ac2d4ca644d619e-24475]java.util.concurrent.ExecutionException: org.apache.hadoop.hbase.NotServingRegionException: org.apache.hadoop.hbase.NotServingRegionException: Region is not online: 763919a2e8ebd05f3735805ba2541ed6

at java.util.concurrent.FutureTask$Sync.innerGet(FutureTask.java:222)

at java.util.concurrent.FutureTask.get(FutureTask.java:83)

at org.apache.hadoop.hbase.client.HConnectionManager$HConnectionImplementation$Process.processBatchCallback(HConnectionManager.java:2113)

at org.apache.hadoop.hbase.client.HConnectionManager$HConnectionImplementation$Process.access$900(HConnectionManager.java:1950)

at org.apache.hadoop.hbase.client.HConnectionManager$HConnectionImplementation.processBatchCallback(HConnectionManager.java:1939)

at org.apache.hadoop.hbase.client.HConnectionManager$HConnectionImplementation.processBatch(HConnectionManager.java:1918)

at org.apache.hadoop.hbase.client.HTable.flushCommits(HTable.java:1026)

at org.apache.hadoop.hbase.client.HTable.doPut(HTable.java:695)

at org.apache.hadoop.hbase.client.HTable.put(HTable.java:670)

at org.apache.hadoop.hbase.util.MultiThreadedWriter$HBaseWriterThread.insert(MultiThreadedWriter.java:175)

at org.apache.hadoop.hbase.util.MultiThreadedWriter$HBaseWriterThread.run(MultiThreadedWriter.java:145)

Caused by: org.apache.hadoop.hbase.NotServingRegionException: org.apache.hadoop.hbase.NotServingRegionException: Region is not online: 763919a2e8ebd05f3735805ba2541ed6

at sun.reflect.GeneratedConstructorAccessor39.newInstance(Unknown Source)

at sun.reflect.DelegatingConstructorAccessorImpl.newInstance(DelegatingConstructorAccessorImpl.java:27)

at java.lang.reflect.Constructor.newInstance(Constructor.java:513)

at org.apache.hadoop.ipc.RemoteException.instantiateException(RemoteException.java:95)

at org.apache.hadoop.ipc.RemoteException.unwrapRemoteException(RemoteException.java:79)

at org.apache.hadoop.hbase.ipc.ProtobufRpcClientEngine$Invoker.invoke(ProtobufRpcClientEngine.java:148)

at $Proxy20.multi(Unknown Source)

at org.apache.hadoop.hbase.protobuf.ProtobufUtil.multi(ProtobufUtil.java:1092)

at org.apache.hadoop.hbase.client.HConnectionManager$HConnectionImplementation$3$1.call(HConnectionManager.java:1783)

at org.apache.hadoop.hbase.client.HConnectionManager$HConnectionImplementation$3$1.call(HConnectionManager.java:1781)

at org.apache.hadoop.hbase.client.ServerCallable.withoutRetries(ServerCallable.java:219)

at org.apache.hadoop.hbase.client.HConnectionManager$HConnectionImplementation$3.call(HConnectionManager.java:1792)

at org.apache.hadoop.hbase.client.HConnectionManager$HConnectionImplementation$3.call(HConnectionManager.java:1778)

at org.apache.hadoop.hbase.client.HConnectionManager$HConnectionImplementation$Process$1.call(HConnectionManager.java:2290)

at org.apache.hadoop.hbase.client.HConnectionManager$HConnectionImplementation$Process$1.call(HConnectionManager.java:2280)

at java.util.concurrent.FutureTask$Sync.innerRun(FutureTask.java:303)

at java.util.concurrent.FutureTask.run(FutureTask.java:138)

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:680)

Caused by: org.apache.hadoop.ipc.RemoteException: org.apache.hadoop.hbase.NotServingRegionException: Region is not online: 763919a2e8ebd05f3735805ba2541ed6

at org.apache.hadoop.hbase.regionserver.HRegionServer.getRegionByEncodedName(HRegionServer.java:2538)

at org.apache.hadoop.hbase.regionserver.HRegionServer.getRegion(HRegionServer.java:3745)

at org.apache.hadoop.hbase.regionserver.HRegionServer.multi(HRegionServer.java:3211)

at sun.reflect.GeneratedMethodAccessor37.invoke(Unknown Source)

at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:25)

at java.lang.reflect.Method.invoke(Method.java:597)

at org.apache.hadoop.hbase.ipc.ProtobufRpcServerEngine$Server.call(ProtobufRpcServerEngine.java:197)

at org.apache.hadoop.hbase.ipc.HBaseServer$Handler.run(HBaseServer.java:1786)

at org.apache.hadoop.hbase.ipc.HBaseClient.call(HBaseClient.java:1304)

at org.apache.hadoop.hbase.ipc.ProtobufRpcClientEngine$Invoker.invoke(ProtobufRpcClientEngine.java:139)

... 14 more

2013-02-08 14:39:22,091 DEBUG [HBaseWriterThread\_8] client.MetaScanner(252): Current INFO from scan results = {NAME => 'IntegrationTestRebalanceAndKillServersTargeted,e6666658,1360362805564.763919a2e8ebd05f3735805ba2541ed6.', STARTKEY => 'e6666658', ENDKEY => 'f3333324', ENCODED => 763919a2e8ebd05f3735805ba2541ed6,}

2013-02-08 14:39:22,091 DEBUG [HBaseWriterThread\_8] client.MetaScanner(199): Scanning .META. starting at row=IntegrationTestRebalanceAndKillServersTargeted,e6666658,00000000000000 for max=10 rows using hconnection 0x610f7612

2013-02-08 14:39:22,092 DEBUG [HBaseWriterThread\_8] client.MetaScanner(252): Current INFO from scan results = {NAME => 'IntegrationTestRebalanceAndKillServersTargeted,e6666658,1360362805564.763919a2e8ebd05f3735805ba2541ed6.', STARTKEY => 'e6666658', ENDKEY => 'f3333324', ENCODED => 763919a2e8ebd05f3735805ba2541ed6,}

2013-02-08 14:39:22,092 DEBUG [HBaseWriterThread\_8] client.HConnectionManager$HConnectionImplementation(1387): Cached location for IntegrationTestRebalanceAndKillServersTargeted,e6666658,1360362805564.763919a2e8ebd05f3735805ba2541ed6. is 10.11.2.92:64486

2013-02-08 14:39:22,093 DEBUG [HBaseWriterThread\_8] client.MetaScanner(252): Current INFO from scan results = {NAME => 'IntegrationTestRebalanceAndKillServersTargeted,f3333324,1360362805564.d362fd7914d7e74e3cff35e8bcaa8c9f.', STARTKEY => 'f3333324', ENDKEY => '', ENCODED => d362fd7914d7e74e3cff35e8bcaa8c9f,}

2013-02-08 14:39:22,093 TRACE [HBaseWriterThread\_8] client.HConnectionManager$HConnectionImplementation$Process(2036): Will retry requests to [10.11.2.92:64486] after delay of [32041] for rows [ed777f62b485b3975ac2d4ca644d619e-24475;]

2013-02-08 14:39:22,107 WARN [main-EventThread] master.OfflineCallback(59): Node for /hbase/region-in-transition/871d1c3bdf98a2c93b527cb6cc61327d already exists

2013-02-08 14:39:22,107 DEBUG [main-EventThread] master.OfflineCallback(69): rs={IntegrationTestRebalanceAndKillServersTargeted,7333332c,1360362805563.871d1c3bdf98a2c93b527cb6cc61327d. state=OFFLINE, ts=1360363161906, server=null}, server=10.11.2.92,49780,1360363101797

2013-02-08 14:39:22,107 WARN [main-EventThread] master.OfflineCallback(59): Node for /hbase/region-in-transition/763919a2e8ebd05f3735805ba2541ed6 already exists

2013-02-08 14:39:22,107 DEBUG [main-EventThread] master.OfflineCallback(69): rs={IntegrationTestRebalanceAndKillServersTargeted,e6666658,1360362805564.763919a2e8ebd05f3735805ba2541ed6. state=OFFLINE, ts=1360363161906, server=null}, server=10.11.2.92,64486,1360362800604

2013-02-08 14:39:22,117 INFO [10.11.2.92,64483,1360362800340-GeneralBulkAssigner-1] master.AssignmentManager(1331): 10.11.2.92,49780,1360363101797 unassigned znodes=5 of total=5

2013-02-08 14:39:22,117 INFO [10.11.2.92,64483,1360362800340-GeneralBulkAssigner-1] master.RegionStates(242): Region {NAME => 'IntegrationTestRebalanceAndKillServersTargeted,7333332c,1360362805563.871d1c3bdf98a2c93b527cb6cc61327d.', STARTKEY => '7333332c', ENDKEY => '7ffffff8', ENCODED => 871d1c3bdf98a2c93b527cb6cc61327d,} transitioned from {IntegrationTestRebalanceAndKillServersTargeted,7333332c,1360362805563.871d1c3bdf98a2c93b527cb6cc61327d. state=OFFLINE, ts=1360363161906, server=null} to {IntegrationTestRebalanceAndKillServersTargeted,7333332c,1360362805563.871d1c3bdf98a2c93b527cb6cc61327d. state=PENDING\_OPEN, ts=1360363162117, server=10.11.2.92,49780,1360363101797}

2013-02-08 14:39:22,117 DEBUG [main-EventThread] master.OfflineCallback$ExistCallback(106): rs={IntegrationTestRebalanceAndKillServersTargeted,f3333324,1360362805564.d362fd7914d7e74e3cff35e8bcaa8c9f. state=OFFLINE, ts=1360363161913, server=null}, server=10.11.2.92,64486,1360362800604

2013-02-08 14:39:22,117 INFO [10.11.2.92,64483,1360362800340-GeneralBulkAssigner-1] master.RegionStates(242): Region {NAME => 'IntegrationTestRebalanceAndKillServersTargeted,99999990,1360362805563.569ec0c144f9033bb42e5e4936fdda4f.', STARTKEY => '99999990', ENDKEY => 'a666665c', ENCODED => 569ec0c144f9033bb42e5e4936fdda4f,} transitioned from {IntegrationTestRebalanceAndKillServersTargeted,99999990,1360362805563.569ec0c144f9033bb42e5e4936fdda4f. state=OFFLINE, ts=1360363161911, server=null} to {IntegrationTestRebalanceAndKillServersTargeted,99999990,1360362805563.569ec0c144f9033bb42e5e4936fdda4f. state=PENDING\_OPEN, ts=1360363162117, server=10.11.2.92,49780,1360363101797}

2013-02-08 14:39:22,118 INFO  [PRI IPC Server handler 6 on 49780] regionserver.HRegionServer(3435): Received request to open region: IntegrationTestRebalanceAndKillServersTargeted,7333332c,1360362805563.871d1c3bdf98a2c93b527cb6cc61327d. on 10.11.2.92,49780,1360363101797

...

2013-02-08 14:39:22,121 WARN  [RS\_OPEN\_REGION-10.11.2.92,49780,1360363101797-1] zookeeper.ZKAssign(762): regionserver:49780-0x13cbbf1025d0015 Attempt to transition the unassigned node for 871d1c3bdf98a2c93b527cb6cc61327d from M\_ZK\_REGION\_OFFLINE to RS\_ZK\_REGION\_OPENING failed, the node existed but was in the state RS\_ZK\_REGION\_OPENING set by the server 10.11.2.92,49780,1360363101797

...

2013-02-08 14:39:22,135 WARN  [RS\_OPEN\_REGION-10.11.2.92,49780,1360363101797-1] zookeeper.ZKAssign(762): regionserver:49780-0x13cbbf1025d0015 Attempt to transition the unassigned node for 871d1c3bdf98a2c93b527cb6cc61327d from M\_ZK\_REGION\_OFFLINE to RS\_ZK\_REGION\_FAILED\_OPEN failed, the node existed but was in the state RS\_ZK\_REGION\_OPENING set by the server 10.11.2.92,49780,1360363101797

...

The bug symptom is not error or exception, but the repeated logs printed in RS, which keeps receiving the request to open the region, but fails (fail to transit the znode state for this region).