HBASE-24292

ndimiduk

2020-04-30

Contents

Chapter 1

Root issue HBASE-24292

1.1 Summary

A "stuck" master should not idle as active without taking action

1.2 Description

The master schedules a SCP for the region server hosting meta. However, due to a misconfiguration, the cluster cannot make progress. After fixing the configuration issue and restarting, the cluster still cannot make progress. After the configured period (15 minuets), the master enters a "holding pattern" where it retains Active master status, but isn't taking any action.

This "brown-out" state is toxic. It should either keep trying to make progress, or it should abort. Staying up and not doing anything is the wrong thing to do.

1.3 Attachments

No attachments

1.4 Comments

1. **ndimiduk:** Here's an example of the "holding pattern" message.

```
2020-04-30 17:46:31,649 WARN org.apache.hadoop.hbase.master.HMaster:
hbase:meta,,1.1588230740 is NOT online; state={1588230740 state=OPEN,
ts=1588268531929, server=host-a.example.com,16020,1588033841562}; ServerCrashProced
Master startup cannot progress, in holding-pattern until region onlined.
```

2. **shahrs87:**

```
1
2  private boolean isRegionOnline(RegionInfo ri) throws InterruptedException {
3
4     RetryCounter rc = null;
5
6     while (!isStopped()) {
7
8       RegionState rs = this.assignmentManager.getRegionStates().getRegionState(ri);
9
10      if (rs.isOpened()) {
11
12        if (this.getServerManager().isServerOnline(rs.getServerName())) {
13
14          return true;
15
16        }
17
18      }
19
20      // Region is not OPEN.
21
22      Optional<Procedure<MasterProcedureEnv>> optProc = this.procedureExecutor.
            getProcedures().
23
24          stream().filter(p -> p instanceof ServerCrashProcedure).findAny();
25
26      // TODO: Add a page to refguide on how to do repair. Have this log message
            point to it.
27
28      // Page will talk about loss of edits, how to schedule at least the meta WAL
            recovery, and
29
30      // then how to assign including how to break region lock if one held.
31
32      LOG.warn("{} is NOT online; state={}; ServerCrashProcedures={}. Master
            startup cannot " +
33
34          "progress, in holding-pattern until region onlined.",
35
36          ri.getRegionNameAsString(), rs, optProc.isPresent());
37
38      // Check once-a-minute.
39
40      if (rc == null) {
41
42        rc = new RetryCounterFactory(1000).create();
43
44      }
45
46      Threads.sleep(rc.getBackoffTimeAndIncrementAttempts());
47
48    }
49
50    return false;
51
52  }
```

If I understand the code correctly, the code sleeps until hbase:meta region comes online and *doesn't* give up. Only one thing I see problematic is the sleeps time never max out. It will grow exponentially without limit. Maybe we should cap it at 1 or 2 minutes.

3. **anoop.hbase:** Ya seems the sleep time can really grow high like 10+ mins even. (Just was trying to test on a 2.0.x cluster)

   Also the while loop is with !isStopped() condition.. No one is stopping this.

   I believe in 1.x there was a timeout like 24 mins or so by default after which the active HM will get stopped. I have seen it in a cluster where it was an issue with getting the NS table online.. After the META table wait, we can see below there is a wait for the META region also. Did not check 1.x code to see what was different there.