

Infrastructure - Task #3116

Metacat CN startup and initialization time needs to be reduced an order of magnitude

2012-08-03 01:14 - Chris Jones

Status:	Closed	Start date:	2012-07-02
Priority:	High	Due date:	
Assignee:	Ben Leinfelder	% Done:	100%
Category:	Metacat	Estimated time:	0.00 hour
Target version:	Sprint-2012.27-Block.4.2	Story Points:	
Milestone:	None		
Product Version:	*		

Description

The nature of a CN environment is such that the definitive list of all known objects can't be determined by a single CN. Even well-behaved CNs may have missing pids because of network partitioning problems (split brain), and so CNs merging into the cluster need to add their known list of objects to the master list (hzIdentifiers), and each CN needs to check locally to ensure it has a persisted copy of the system metadata for that pid.

The startup problem is two-fold. Population of the hzIdentifiers set seems to be slow due to Hazelcast internally migrating ownership of items in the set. This needs more investigation, but see [#3045](#) for recent work on this. The second issue is that each CN broadcasts its full set of pids via put() calls to the hzSystemMetadata map. All CNs listen for these add/update events, and attempt to locally save a copy of the system metadata. This introduces a large delay because of the iteration over every pid. At 1 pid/sec, the stage environment may take over a day to fully synchronize (currently 143K objects).

To alleviate this, Ben and I have discussed a means for each CN to calculate a short list of locally missing pids by comparing to the hzIdentifiers et once it is fully initialized. Once the CN has its short list, it will publish its list to the (newly created) hzMissingIdentifiers set. All CNs will be listening for itemAdded events on this set, which will trigger all CNs to look for a local copy of the system metadata for that pid. If it has it, it will first attempt to gain a shared locked (via tryLock()) based on the pid (like "missing-{pid}"). If it gains the lock, it will call hzSystemMetadata.put(pid), which will cause all CNs to save the system metadata locally if they don't have it. The lock reduces the number of put(s) to the system metadata map since most of the time multiple CNs will indeed have a copy, and only one needs to do the put().

In theory, by each CN creating a short list of missing pids, and only 'requesting' a copy of the those pids via the hzMissingIdentifiers set, we will reduce the number of hzSystemMetadata.put() calls at least an order of magnitude. There will still be redundancies, but this is a start.

Subtasks:

Task # 3045: Reduce Hazelcast.init() startup time for populating hzIdentifiers

Closed

Task # 3117: Reduce Hazelcast.init() startup time for populating hzSystemMetadata

Closed

History

#1 - 2012-08-07 15:24 - Ben Leinfelder

- Status changed from New to In Progress

Will be deploying from the -stable channel on sandbox soon.

#2 - 2012-08-14 23:53 - Ben Leinfelder

- Status changed from In Progress to Closed

This is in production on the CNs now