CN REST - Task #8777

Story # 8756 (New): Ensure replica auditor is effective

Configure CN to audit objects greater than 1GB

2019-03-12 16:47 - Chris Jones

Status:	New	**Start date:**	2019-03-12
Priority:	Normal	**Due date:**	
Assignee:	Chris Jones	**% Done:**	0%
Category:	d1_replication_auditor	**Estimated time:**	0.00 hour
Target version:			
Story Points:			

Description

The replication auditor currently limits auditing of objects at 1GB. There are currently 4 objects greater than 1TB in size, and 3,588 objects greater than 1GB in size, both being very small counts compared to the 2,769,111 objects less than 1GB in size in the network. Nonetheless, they should still be audited if feasible. The limiting factor is likely HTTP timeout limits during the call to MN.getChecksum(). For reference, I'm seeing the following general times for calculating MD5 and SHA-1 checksums:

```
Size   MD5         SHA-1
----   -------     -------
1GB    00m02.5s    00m02.6s
10GB   00m25.9s    00m30.0s
100GB  03m28.0s    04m01.8s
1TB    50m14.2s    67m38.6s
```

10GB and 100GB objects seem pretty feasible if we set the HTTP client timeout to > 5 minutes, whereas the few > 1TB files may be challenging just due to the timeouts. The other factor is that the AbstractReplicationAuditor sets a default timeout to 60 seconds, and if the task future doesn't return in that time frame, the future gets cancelled. So the HTTP timeout and this timeout need to be increased and coordinated in order to handle larger object auditing.

History

#1 - 2019-03-13 20:09 - Chris Jones

- Description updated

#2 - 2019-03-13 20:50 - Dave Vieglais

Need to be smarter about verifying content. It is prohibitive for the CN to go around checking millions of objects, and won't scale. Perhaps MNs should be responsible for ensuring their copy is accurate according to the checksum reported by the authoritative MN or the CN?