Arvados - Bug #16842

[keepstore] S3 driver should handle sub-second precision in timestamps

09/16/2020 05:44 PM - Tom Clegg

Status:	Resolved		Start date:	09/16/2020
Priority:	Normal		Due date:	
Assigned To:	Tom Clegg		% Done:	100%
Category:	Keep		Estimated time:	0.00 hour
Target version:	2020-09-23 Sprint			

Description

Before deleting a block, in order to detect GC races, keepstore's trash worker compares the timestamp of the stored block to the timestamp provided by keep-balance in the trash list, which in turn comes from keepstore's "index" worker.

The stored block's timestamp comes from the Last-Modified header from an S3 "head object" response, which is RFC1123-formatted and therefore has no better than 1-second precision.

The index worker's timestamp comes from the LastModified XML tag in the S3 "list objects" response, which is RFC3339-formatted and potentially has a non-zero nanosecond part.

AWS and Google implementations of S3 always return zero nanoseconds in the "list objects" response, but Ceph/nautilus returns timestamps with milliseconds, which as a result will never match the RFC1123-formatted timestamp used by the trash worker.

To resolve this, the S3 index handler should truncate the nanosecond part when parsing timestamps.

Subtasks:

Task # 16844: Review 16842-s3-timestamp-precision	**Resolved**

Associated revisions

Revision 2a3c7865 - 09/17/2020 05:31 PM - Tom Clegg

Merge branch '16842-s3-timestamp-precision'

fixes #16842

Arvados-DCO-1.1-Signed-off-by: Tom Clegg <tom@tomclegg.ca>

History

#1 - 09/16/2020 05:48 PM - Tom Clegg

- Description updated

- Subject changed from [keepstore] S3 driver trash worker should accept 1-second-precision timestamps to [keepstore] S3 driver should handle sub-second precision in timestamps

#2 - 09/16/2020 06:08 PM - Tom Clegg

Aside: the rationale for truncating the index timestamps (rather than truncating before comparing in the trash worker) is to avoid a situation where keep-balance sees multiple copies with differing timestamps (say, 1234567.123 and 1234567.456) and concludes it's safe to trash one of them, but then keepstore only checks the 1234567 part, and trashes both copies.

16842-s3-timestamp-precision @ f317fc0d8e77ce950b6a650149600b0c8f6c38f3 --
https://ci.arvados.org/view/Developer/job/developer-run-tests/2100/

#3 - 09/16/2020 11:59 PM - Ward Vandewege

Developer tests very unhappy/stuck, I cleared them all out and restarted them at https://ci.arvados.org/view/Developer/job/developer-run-tests/2102/, they passed finally.

#4 - 09/17/2020 01:54 PM - Tom Clegg

- File keepstore added

16842-s3-timestamp-precision @ 7e503a238d67b79a9f87e50086ca217d4873726e --
https://ci.arvados.org/view/Developer/job/developer-run-tests/2103/

#5 - 09/17/2020 05:28 PM - Ward Vandewege

Tom Clegg wrote:

> 16842-s3-timestamp-precision @ 7e503a238d67b79a9f87e50086ca217d4873726e --
> https://ci.arvados.org/view/Developer/job/developer-run-tests/2103/

LGTM

#6 - 09/17/2020 05:35 PM - Anonymous

- Status changed from In Progress to Resolved

Applied in changeset arvados|2a3c78659dfdb264cd65ef6ababbd973062b4eae.

#7 - 09/17/2020 06:10 PM - Ward Vandewege

- File keepstore_2.0.4.dev20200917174905-1_amd64.deb added

Attaching a keepstore Ubuntu 18.04 package built from the 2.0-dev branch, with f317fc0d8e77ce950b6a650149600b0c8f6c38f3 cherry-picked on top.

Functionally, that means version 2.0.4 plus the fix from this ticket.

#8 - 09/21/2020 02:47 PM - Tom Clegg

- Related to Story #16360: Keep-web supports S3 compatible interface added

#9 - 09/21/2020 02:47 PM - Tom Clegg

- Related to deleted (Story #16360: Keep-web supports S3 compatible interface)

Files

keepstore	23.4 MB	09/17/2020	Tom Clegg
keepstore_2.0.4.dev20200917174905-1_amd64.deb	8.26 MB	09/17/2020	Ward Vandewege