

Arvados - Bug #13513

[keep-balance] hang on ComputeChangeSets

05/22/2018 03:07 PM - Ward Vandewege

Status:	Resolved	Start date:	05/29/2018
Priority:	Normal	Due date:	
Assigned To:	Tom Clegg	% Done:	100%
Category:		Estimated time:	0.00 hour
Target version:	2018-06-06 Sprint		

Description

After the merge of 9918-index-timeouts, I'm observing that keep-balance hangs (?) on ComputeChangeSets:

```
May 22 14:06:37 dhhck.arvadosapi.com keep-balance[11166]: 2018/05/22 14:06:37
dhhck-bi614-pkwwh8mhe0ggmu6 (keep2.dhhck.arvadosapi.com:25107, s3): done
May 22 14:08:40 dhhck.arvadosapi.com keep-balance[11166]: 2018/05/22 14:08:40
zzzzz-ivpuk-v2udip63fnkdyxf (s3:///dhhck-keep-0) on dhhck-bi614-oynapdlh4hzydcf
(keep0.dhhck.arvadosapi.com:25107, s3): add 1043919 replicas to map
May 22 14:08:40 dhhck.arvadosapi.com keep-balance[11166]: 2018/05/22 14:08:40
zzzzz-ivpuk-v2udip63fnkdyxf (s3:///dhhck-keep-0) on dhhck-bi614-oynapdlh4hzydcf
(keep0.dhhck.arvadosapi.com:25107, s3): done
May 22 14:08:40 dhhck.arvadosapi.com keep-balance[11166]: 2018/05/22 14:08:40
dhhck-bi614-oynapdlh4hzydcf (keep0.dhhck.arvadosapi.com:25107, s3): done
May 22 14:08:40 dhhck.arvadosapi.com keep-balance[11166]: 2018/05/22 14:08:40 GetCurrentState: too
k 10m6.992266703s
May 22 14:08:40 dhhck.arvadosapi.com keep-balance[11166]: 2018/05/22 14:08:40 ComputeChangeSets: s
tart
```

I stopped it after ~42 minutes.

```
May 22 14:50:02 dhhck.arvadosapi.com systemd[1]: Stopping Arvados Keep Balance...
May 22 14:50:02 dhhck.arvadosapi.com systemd[1]: Stopped Arvados Keep Balance.
```

Command line:

```
/usr/bin/keep-balance -commit-trash
```

I also tried with -commit-pull enabled, and the behavior was unchanged.

Config file:

```
# cat /etc/arvados/keep-balance/keep-balance.yml
#####
# THIS FILE IS MANAGED BY PUPPET -- CHANGES WILL BE OVERWRITTEN #
#####
Client:
  APIHost: dhhck.arvadosapi.com:443
  AuthToken: STRIPPED
  Insecure: false
KeepServiceTypes:
  - s3
RunPeriod: 14400s
CollectionBatchSize: 100000
CollectionBuffers: 1000
```

Bisecting:

0.1.20180322172032.41e612b59-1	(with extra patch to increase timeout to 20 minutes)	OK
1.1.4.20180403215323-1	(with extra patch to increase timeout to 20 minutes)	OK
1.1.4.20180420195921-1	(with extra patch to increase timeout to 20 minutes)	OK

	minutes)	
1.1.4.20180426154228-1	(with extra patch to increase timeout to 20 minutes)	OK
1.1.4.20180426193406-1	(with extra patch to increase timeout to 20 minutes)	HANGS
1.1.4.20180510200716-1	(with extra patch to increase timeout to 20 minutes)	HANGS
1.1.4.20180518195015-1		HANGS

So, it looks like the problem was introduced between version 1.1.4.20180426154228-1 ([fcfbddf572db32008fcdc7d0750a13b8d6f3b1c](#)) and version 1.1.4.20180426193406-1 ([932e3d6e9a899cc662ea3934b79057d39cd88fed](#)).

Subtasks:

Task # 13529: Review 13513-balance-deadlock

Closed

Associated revisions

Revision 05fdfcec - 05/29/2018 08:05 PM - Tom Clegg

Merge branch '13513-balance-deadlock'

fixes #13513

Arvados-DCO-1.1-Signed-off-by: Tom Clegg <tclegg@veritasgenetics.com>

History

#1 - 05/22/2018 03:30 PM - Ward Vandewege

- Description updated

#2 - 05/22/2018 04:07 PM - Ward Vandewege

- Description updated

#3 - 05/22/2018 04:40 PM - Ward Vandewege

- Description updated

#4 - 05/22/2018 05:12 PM - Ward Vandewege

- Description updated

#5 - 05/23/2018 02:49 PM - Tom Morris

- Target version set to 2018-06-06 Sprint

#6 - 05/23/2018 02:50 PM - Ward Vandewege

- Project changed from Arvados Private to Arvados

#7 - 05/23/2018 02:52 PM - Ward Vandewege

- Subject changed from keep-balance hangs on ComputeChangeSets to [keep-balance] hang on ComputeChangeSets

#8 - 05/23/2018 03:14 PM - Ward Vandewege

- Assigned To set to Brad Chapman

#9 - 05/23/2018 03:14 PM - Ward Vandewege

- Assigned To changed from Brad Chapman to Tom Clegg

#10 - 05/26/2018 01:35 AM - Tom Clegg

Indeed, 1.1.4.20180426193406-1 was [932e3d6e9a899cc662ea3934b79057d39cd88fed](#), which merged [#12708](#), including this:

```
close(todo)
go func() {
    wg.Wait()
```

```
    close(results)
  }()
  bal.collectStatistics(results)
```

close(todo) waits for workers to read everything from todo, but every worker waits forever for bal.collectStatistics() to receive its first result.

Should use the pattern "sending goroutine closes channel": run Apply() in a goroutine which closes todo when finished, and run the nWorkers loop in a goroutine which closes results after wg.Wait(). Then collectStatistics() can be called synchronously and nobody outside the nWorkers loop needs to know about the wg.

#11 - 05/29/2018 04:22 PM - Tom Clegg

13513-balance-deadlock @ [f1cf835d576de364faa7e94f5ba8caca5ad7be43](#)

#12 - 05/29/2018 08:01 PM - Ward Vandewege

Tom Clegg wrote:

13513-balance-deadlock @ [f1cf835d576de364faa7e94f5ba8caca5ad7be43](#)

Test run of this code:

```
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 keep-balance dev started
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 starting up: will scan every 4h0
m0s and on SIGUSR1
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 Run: start
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 skipping
dhhck-bi614-blb6dtx0sx07d7k with service type "proxy"
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 clearing existing trash lists, i
n case the new rendezvous order differs from previous run
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 dhhck-bi614-pkwwh8mhe0qgm6
(keep2.dhhck.arvadosapi.com:25107, s3): send trash list: start
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 dhhck-bi614-jt6xs3d6d531oaa
(keep1.dhhck.arvadosapi.com:25107, s3): send trash list: start
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 dhhck-bi614-u6ez2qmdvdde1xc
(keep3.dhhck.arvadosapi.com:25107, s3): send trash list: start
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 dhhck-bi614-oynapdlh4hzydcf
(keep0.dhhck.arvadosapi.com:25107, s3): send trash list: start
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 dhhck-bi614-u6ez2qmdvdde1xc
(keep3.dhhck.arvadosapi.com:25107, s3): send trash list: took 830.003µs
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 dhhck-bi614-oynapdlh4hzydcf
(keep0.dhhck.arvadosapi.com:25107, s3): send trash list: took 909.802µs
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 dhhck-bi614-jt6xs3d6d531oaa
(keep1.dhhck.arvadosapi.com:25107, s3): send trash list: took 1.686482ms
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 dhhck-bi614-pkwwh8mhe0qgm6
(keep2.dhhck.arvadosapi.com:25107, s3): send trash list: took 1.735439ms
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 GetCurrentState: start
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 dhhck-bi614-u6ez2qmdvdde1xc
(keep3.dhhck.arvadosapi.com:25107, s3): retrieve indexes
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 dhhck-bi614-pkwwh8mhe0qgm6
(keep2.dhhck.arvadosapi.com:25107, s3): retrieve indexes
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 zzzzz-ivpuk-aj576elt62use9i
(s3:///dhhck-keep-2) on dhhck-bi614-u6ez2qmdvdde1xc (keep3.dhhck.arvadosapi.com:25107, s3): retrieve index
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 zzzzz-ivpuk-mkmes11v77mlcso
(s3:///dhhck-keep-3) on dhhck-bi614-pkwwh8mhe0qgm6 (keep2.dhhck.arvadosapi.com:25107, s3): retrieve index
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 dhhck-bi614-jt6xs3d6d531oaa
(keep1.dhhck.arvadosapi.com:25107, s3): retrieve indexes
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 zzzzz-ivpuk-aexopqahib74kpp
(s3:///dhhck-keep-1) on dhhck-bi614-jt6xs3d6d531oaa (keep1.dhhck.arvadosapi.com:25107, s3): retrieve index
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 dhhck-bi614-oynapdlh4hzydcf
(keep0.dhhck.arvadosapi.com:25107, s3): retrieve indexes
May 29 19:34:16 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:16 zzzzz-ivpuk-v2udip63fnkdyxf
(s3:///dhhck-keep-0) on dhhck-bi614-oynapdlh4hzydcf (keep0.dhhck.arvadosapi.com:25107, s3): retrieve index
May 29 19:34:17 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:17 collections: 0/137062
May 29 19:34:19 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:19 collections: 1000/137062
May 29 19:34:20 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:20 collections: 2000/137062
May 29 19:34:21 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:21 collections: 3000/137062
May 29 19:34:22 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:22 collections: 4000/137062
May 29 19:34:23 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:23 collections: 5000/137062
May 29 19:34:24 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:24 collections: 6000/137062
May 29 19:34:25 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:25 collections: 7000/137062
May 29 19:34:25 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:25 collections: 8000/137062
May 29 19:34:26 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:26 collections: 9000/137062
May 29 19:34:27 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:34:27 collections: 10000/137062
```


May 29 19:47:46 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:46 [zzzzz-ivpuk-v2udip63fnkdxyf](#)
(s3:///dhhck-keep-0) on [dhhck-bi614-oynapdlh4hzydcf](#)
(keep0.dhhck.arvadosapi.com:25107, s3): add 1043916 replicas to map
May 29 19:47:47 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:47 [zzzzz-ivpuk-v2udip63fnkdxyf](#)
(s3:///dhhck-keep-0) on [dhhck-bi614-oynapdlh4hzydcf](#) (keep0.dhhck.arvadosapi.com:25107, s3): done
May 29 19:47:47 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:47 [dhhck-bi614-oynapdlh4hzydcf](#)
(keep0.dhhck.arvadosapi.com:25107, s3): done
May 29 19:47:47 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:47 GetCurrentState: took 13m30.4014
16854s
May 29 19:47:47 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:47 ComputeChangeSets: start
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 ComputeChangeSets: took 10.95570
2441s
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 ===
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 0 replicas (0 blocks, 0 bytes) l
ost (0=have<want)
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 0 replicas (0 blocks, 0 bytes) u
nderreplicated (0<have<want)
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 3357390 replicas (3357390 blocks
, 199210561559859 bytes) just right (have=want)
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 154 replicas (150 blocks, 101771
5394 bytes) overreplicated (have>want>0)
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 0 replicas (0 blocks, 0 bytes) u
nreferenced (have>want=0, new)
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 0 replicas (0 blocks, 0 bytes) g
arbage (have>want=0, old)
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 ===
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 storage class "default": 3357540
replicas (3357540 blocks, 199211579243382 bytes) desired
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 storage class "default": 0 repli
cas (0 blocks, 0 bytes) short
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 storage class "default": 154 rep
licas (150 blocks, 1017715394 bytes) surplus
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 storage class "default": 0 repli
cas (0 blocks, 0 bytes) unachievable
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 ===
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 3357540 replicas (3357540 blocks
, 199211579243382 bytes) total commitment (excluding unreferenced)
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 3357694 replicas (3357540 blocks
, 199212596958776 bytes) total usage
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 ===
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 [dhhck-bi614-jt6xs3d6d531oaa](#)
(keep1.dhhck.arvadosapi.com:25107, s3): ChangeSet{Pulls:326784, Trashes:0}
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 [dhhck-bi614-pkwwh8mhe0qgmu6](#)
(keep2.dhhck.arvadosapi.com:25107, s3): ChangeSet{Pulls:1427, Trashes:0}
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 [dhhck-bi614-u6ez2qmdvddelxc](#)
(keep3.dhhck.arvadosapi.com:25107, s3): ChangeSet{Pulls:1480, Trashes:0}
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 [dhhck-bi614-oynapdlh4hzydcf](#)
(keep0.dhhck.arvadosapi.com:25107, s3): ChangeSet{Pulls:343, Trashes:0}
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 ===
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 Replication level distribution (
counting N replicas on a single server as N):
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 0: 0
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 1: 3357390 #####

May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 2: 146 #####
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 3: 4 #####
May 29 19:47:58 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:57 ===
May 29 19:47:59 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:59 [dhhck-bi614-pkwwh8mhe0qgmu6](#)
(keep2.dhhck.arvadosapi.com:25107, s3): send trash list: start
May 29 19:47:59 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:59 [dhhck-bi614-oynapdlh4hzydcf](#)
(keep0.dhhck.arvadosapi.com:25107, s3): send trash list: start
May 29 19:47:59 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:59 [dhhck-bi614-jt6xs3d6d531oaa](#)
(keep1.dhhck.arvadosapi.com:25107, s3): send trash list: start
May 29 19:47:59 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:59 [dhhck-bi614-u6ez2qmdvddelxc](#)
(keep3.dhhck.arvadosapi.com:25107, s3): send trash list: start
May 29 19:47:59 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:59 [dhhck-bi614-oynapdlh4hzydcf](#)
(keep0.dhhck.arvadosapi.com:25107, s3): send trash list: took 780.98µs
May 29 19:47:59 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:59 [dhhck-bi614-pkwwh8mhe0qgmu6](#)
(keep2.dhhck.arvadosapi.com:25107, s3): send trash list: took 1.555394ms
May 29 19:47:59 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:59 [dhhck-bi614-jt6xs3d6d531oaa](#)
(keep1.dhhck.arvadosapi.com:25107, s3): send trash list: took 1.537235ms
May 29 19:47:59 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:59 [dhhck-bi614-u6ez2qmdvddelxc](#)
(keep3.dhhck.arvadosapi.com:25107, s3): send trash list: took 1.595105ms
May 29 19:47:59 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:59 Run: took 13m43.311678264s

May 29 19:47:59 dhhck.arvadosapi.com keep-balance[21398]: 2018/05/29 19:47:59 run succeeded

So from a functional perspective, this LGTM.

#13 - 05/29/2018 09:11 PM - Tom Clegg

- *Status changed from New to Resolved*
- *% Done changed from 0 to 100*

Applied in changeset [arvados|05fdfcec08e280a03c532a1a98b3aeebc716f6eb](#).

#14 - 07/23/2018 06:52 PM - Tom Morris

- *Release set to 13*