Arvados - Feature #11184

[Keep] Support multiple storage classes

02/28/2017 11:23 PM - Tom Morris

Status:	Resolved		**Start date:**	
Priority:	Normal		**Due date:**	
Assigned To:	Tom Morris		**% Done:**	0%
Category:			**Estimated time:**	0.00 hour
Target version:				

Description

As an Arvados system administrator, I want to take advantage of the cool/cold storage classes offered by cloud vendors.

This involves designating desired storage class(es) in some way (collection, all collections in a project, etc) as well as a way to migrate between storage classes.

Keep storage classes

Overview

- Each keep volume offers one or more storage classes (the default is just the "default" class).
- Each collection has one or more desired storage classes (the default is just the "default" class).
- When writing, clients *may* specify one or more required storage classes; if not, the required class is "default". A keepstore server will only write the data on a volume that offers *all* of the required classes.
- Keep-balance moves data to volumes that have the desired attributes and updates collection records to reflect the storage classes currently satisfied by all blocks in the collection (much like replication level, these are not necessarily equal to the desired classes).

If overlapping collections (i.e., with common data blocks) request different storage classes, keep-balance will maintain multiple copies of the common blocks if necessary to satisfy all collections' requirements.

Simplifying restrictions in initial implementation (#12708):

- No client side support.
- No keepstore support for writing data to a given storage class.
- API server configuration specifies the set of classes that can be requested (in addition to "default", which is always available).
- Keepstore configuration specifies the set of classes offered by each volume.

Related issues:

Related to Arvados - Story #7931: [keep-balance] Count block replication by v...	**Resolved**	**12/02/2015**	
Related to Arvados - Feature #11644: [keepstore] mount-oriented APIs	**Resolved**	**05/09/2017**	
Related to Arvados - Feature #13382: [keepstore] Write new blocks to appropri...	**Resolved**	**04/02/2021**	
Related to Arvados - Feature #17392: Support writing blocks to correct storag...	**Resolved**	**04/12/2021**	
Related to Arvados Epics - Story #16107: Storage classes	**New**	**03/01/2021**	**09/30/2021**
Blocked by Arvados - Feature #11645: [keepstore] Add "StorageClasses" field t...	**Resolved**	**05/09/2017**	
Blocked by Arvados - Story #12707: [API] Add columns for desired/actual stora...	**Resolved**	**02/22/2018**	
Blocked by Arvados - Story #12708: [keep-balance] Move blocks to satisfy stor...	**Resolved**	**04/19/2018**	
Blocked by Arvados - Feature #13431: [keepproxy] [GoSDK] Propagate desired st...	**Resolved**	**06/05/2018**	
Blocked by Arvados - Story #13429: [API] [arvados-cwl-runner] Save workflow o...	**Closed**	**06/04/2018**	
Blocked by Arvados - Story #13430: [arv-put] [Python] Allow caller to specify...	**Resolved**	**05/22/2018**	

History

#1 - 03/28/2017 06:53 PM - Tom Morris

- Assigned To set to Tom Clegg

#2 - 04/25/2017 06:45 PM - Tom Clegg

- Description updated

#3 - 06/13/2017 06:52 PM - Tom Morris

- Assigned To changed from Tom Clegg to Tom Morris

#4 - 06/13/2017 09:14 PM - Tom Morris

- Description updated

#5 - 09/19/2017 06:18 PM - Tom Clegg

suggest

1. specify desired functionality (user can push X buttons, result is Y)
2. enumerate features/improvements needed to achieve desired functionality, and define any new APIs needed
3. estimate points (or maybe split into separate stories if needed)

#6 - 10/05/2017 06:09 PM - Peter Amstutz

https://dev.arvados.org/projects/arvados/wiki/Keep_storage_groups

#7 - 10/05/2017 06:52 PM - Tom Morris

- Description updated

#8 - 10/05/2017 06:52 PM - Tom Morris

- Story points set to 3.0

#9 - 11/29/2017 06:53 PM - Peter Amstutz

- Target version changed from Arvados Future Sprints to To Be Groomed

#10 - 11/29/2017 06:53 PM - Peter Amstutz

- Assigned To changed from Tom Morris to Tom Clegg

#11 - 11/29/2017 08:16 PM - Tom Clegg

- Blocked by Feature #11645: [keepstore] Add "StorageClasses" field to volume config added

#12 - 11/29/2017 08:30 PM - Tom Clegg

- Subject changed from [Keep] Support multiple storage tiers to [Keep] Support multiple storage classes

- Description updated

Pricing examples (moved from description)

Microsoft Azure (pricing at 50-500 TB level)

- LRS-COOL $0.01/GB/mo, $0.01/10Kops + $0.01GB
- LRS-HOT $0.0177/GB/mo, $0.05/10Kops
- GRS-COOL $0.02/GB/mo, $0.20/10Kops + $0.01/GB
- GRS-HOT $0.0354/GB/mo, $0.10/10Kops
- RAGRS-COOL $0.025/GB/mo, $0.20/10Kops + $0.01/GB
- RAGRS-HOST $0.0442/GB/mo, $0.10/10Kops

Amazon S3 (pricing at 50-500TB level)

- Standard - $0.022/GB, $0.004/10Kops (get)
- Infrequent Access - $0.0125/GB, $0.01/10Kops (get)
- Glacier - $0.004/GB + variable retrieval charge depending on speed

Google

- Multi-Regional Storage $0.026/GB/mo
- Regional Storage $0.02/GB/mo
- Nearline Storage $0.01/GB/mo, $0.01/GB retrieval charge
- Coldline Storage $0.007, $0.05/GB retrieval charge
- Optional bucket versioning

#13 - 11/29/2017 08:45 PM - Tom Clegg

- Blocked by Story #12707: [API] Add columns for desired/actual storage classes for each collection added

#14 - 11/29/2017 08:46 PM - Tom Clegg

- Blocked by Story #12708: [keep-balance] Move blocks to satisfy storage_classes_desired added

#15 - 11/29/2017 08:58 PM - Tom Clegg

- Description updated

#16 - 12/05/2017 07:10 PM - Tom Clegg

- Related to Story #7929: [SDKs] PySDK KeepClient considers volume IDs when replicating added

#17 - 12/05/2017 07:10 PM - Tom Clegg

- Related to Story #7930: [SDKs] GoSDK KeepClient considers volume IDs when replicating added

#18 - 12/05/2017 07:10 PM - Tom Clegg

- Related to Story #7931: [keep-balance] Count block replication by volume IDs added

#19 - 12/05/2017 07:10 PM - Tom Clegg

- Related to Story #7932: [Keep] keepproxy aggregates and reports volume IDs from Keepstores added

#20 - 02/28/2018 08:42 PM - Tom Morris

- Target version changed from To Be Groomed to 2018-03-14 Sprint

#21 - 03/14/2018 07:03 PM - Tom Clegg

- Target version changed from 2018-03-14 Sprint to 2018-03-28 Sprint

#22 - 03/14/2018 07:17 PM - Tom Morris

- Target version changed from 2018-03-28 Sprint to Arvados Future Sprints

#23 - 03/21/2018 07:44 PM - Tom Clegg

- Related to Feature #11644: [keepstore] mount-oriented APIs added

#24 - 04/11/2018 03:32 PM - Tom Morris

- Related to deleted (Story #7932: [Keep] keepproxy aggregates and reports volume IDs from Keepstores)

#25 - 04/24/2018 01:28 PM - Tom Clegg

- Related to Feature #13382: [keepstore] Write new blocks to appropriate storage class added

#26 - 05/02/2018 02:24 PM - Tom Morris

- Related to Story #13429: [API] [arvados-cwl-runner] Save workflow outputs to desired storage classes added

#27 - 05/02/2018 02:25 PM - Tom Morris

- Related to Story #13430: [arv-put] [Python] Allow caller to specify storage classes when writing data to Keep added

#28 - 05/02/2018 03:22 PM - Tom Clegg

- Blocked by Feature #13431: [keepproxy] [GoSDK] Propagate desired storage classes in PUT request headers added

#29 - 05/02/2018 04:18 PM - Tom Clegg

- Related to deleted (Story #7929: [SDKs] PySDK KeepClient considers volume IDs when replicating)

#30 - 05/02/2018 04:18 PM - Tom Clegg

- Related to deleted (Story #7930: [SDKs] GoSDK KeepClient considers volume IDs when replicating)

#31 - 05/02/2018 04:19 PM - Tom Clegg

- Related to deleted (Story #13429: [API] [arvados-cwl-runner] Save workflow outputs to desired storage classes)

#32 - 05/02/2018 04:19 PM - Tom Clegg

- Blocked by Story #13429: [API] [arvados-cwl-runner] Save workflow outputs to desired storage classes added

#33 - 05/02/2018 04:19 PM - Tom Clegg

- Related to deleted (Story #13430: [arv-put] [Python] Allow caller to specify storage classes when writing data to Keep)

#34 - 05/02/2018 04:19 PM - Tom Clegg

- Blocked by Story #13430: [arv-put] [Python] Allow caller to specify storage classes when writing data to Keep added

#35 - 05/02/2018 04:21 PM - Tom Clegg

- Description updated

#36 - 06/06/2018 03:06 PM - Tom Morris

- Target version changed from Arvados Future Sprints to 2018-06-20 Sprint

#37 - 06/06/2018 03:19 PM - Tom Morris

- Assigned To changed from Tom Clegg to Tom Morris

- Story points deleted (3.0)

#38 - 06/06/2018 03:24 PM - Tom Morris

- Status changed from New to In Progress

#39 - 06/20/2018 03:37 PM - Tom Morris

- Target version changed from 2018-06-20 Sprint to Arvados Future Sprints

#40 - 04/12/2021 03:29 PM - Tom Clegg

- Related to Feature #17392: Support writing blocks to correct storage classes in Go SDK added

#41 - 04/12/2021 03:29 PM - Tom Clegg

- Related to Story #16107: Storage classes added

#42 - 07/06/2021 09:13 PM - Peter Amstutz

- Status changed from In Progress to Resolved

#43 - 07/07/2021 05:12 PM - Ward Vandewege

- Target version changed from Arvados Future Sprints to 0

#44 - 07/07/2021 05:14 PM - Ward Vandewege

- Target version deleted (0)