Arvados - Feature #6611

[Data Manager] Act on pull lists for underreplicated blocks

07/14/2015 02:37 PM - Tom Clegg

Status:	Closed		**Start date:**	07/14/2015
Priority:	Normal		**Due date:**	
Assigned To:			**% Done:**	0%
Category:	Keep		**Estimated time:**	0.00 hour
Target version:				

Description

By design, Data Manager fixed underreplicated blocks by sending "pull lists" to the keepstore servers.

Currently, Data Manager computes lists of underreplicated blocks, but doesn't send the pull lists.

Before this can be deployed:

- data manager must send its pull lists to keep servers
- data manager must have *at least* one comprehensive integration test covering a miniature version of normal operation:
    - bring up api, keepstore, and keepproxy services
    - store some collections
    - delete some blocks (using keepstore DELETE method with a suitable token) so the desired replication level is no longer met for some collections
    - run data manager
    - wait for all keepstores to finish working their trash and pull lists (by getting each keepstore's /status.json until the queues are empty and workers are idle)
    - make sure all collections' blocks are back up to their desired replication level
    - run data manager again
    - wait for all keepstores to finish working their trash and pull lists
    - make sure no more (excess) replicas have been made

Subtasks:

Task # 6612: Report trash and pull queue & worker status in keepstore /status.json	**New**
Task # 6631: Review 6611-send-pull-list	**In Progress**

Related issues:

Related to Arvados - Story #9162: [Block Manager] Trashes unreferenced blocks...	**Resolved**	**05/16/2016**

History

#1 - 07/14/2015 02:38 PM - Tom Clegg

- Description updated

#2 - 07/14/2015 02:48 PM - Tom Clegg

- Description updated

#3 - 07/14/2015 02:49 PM - Tom Clegg

- Description updated

#4 - 09/25/2015 01:16 PM - Peter Amstutz

Note: there is already a preliminary 6611-send-pull-list branch that I wrote when I wrote at the same time as the "send trash list" code.

#5 - 04/27/2016 07:30 PM - Peter Amstutz

- Target version changed from Arvados Future Sprints to 2016-05-11 sprint

#6 - 04/27/2016 07:33 PM - Tom Clegg

Fixing underreplicated blocks is a special case of rebalancing: "rebalance only the blocks that have less than desired replication."

If

- block B has replication A ("actual"), and

- block B should have replication D ("desired"), and
- A < D, and
- keep service K is one of the first D services in the probe order for block A, and
- block B is not in service K's index

...then

- pick some service K2 that *does* have block B in its index
- add K2/B to the pull list for K.

#7 - 04/27/2016 07:40 PM - Radhika Chippada

- Assigned To set to Radhika Chippada

#8 - 04/27/2016 07:48 PM - Radhika Chippada

- Story points set to 2.0

#9 - 04/27/2016 08:21 PM - Brett Smith

- Target version changed from 2016-05-11 sprint to Arvados Future Sprints

#10 - 06/09/2016 10:07 PM - Radhika Chippada

- Assigned To deleted (Radhika Chippada)

#11 - 10/12/2016 08:27 PM - Tom Clegg

- Status changed from New to Closed

#12 - 01/18/2017 07:10 AM - Tom Morris

- Target version deleted (Arvados Future Sprints)