

rd - Bug #16855

rd mirror: after promote, the mirror image often be up+error

07/29/2016 07:28 AM - de lan

Status:	Resolved	Start date:	07/29/2016
Priority:	Normal	Due date:	
Assignee:	Jason Dillaman	% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:		Spent time:	0.00 hour
Source:	other	Reviewed:	
Tags:		Affected Versions:	
Backport:	jewel	ceph-qa-suite:	
Regression:	No	Pull request ID:	
Severity:	3 - minor		

Description

HI!

when i do some demote and promote operation.the mirror image often become up+error.

my operating record:

```
ubuntu@plana140:~$ rbd --cluster cluster2 -p mirror create --size 128 --image-feature layering,exclusive-lock,journaling test123
2016-07-29 15:15:48.033155 7f89304a8700 0 -- 192.168.34.140:0/348281884 >> 192.168.34.140:6800/5665 pipe(0xb3ac60 sd=8 :51524 s=2 pgs=241 cs=1 l=1 c=0xb14b50).injecting socket failure
ubuntu@plana140:~$ rbd --cluster cluster1 -p mirror mirror image status test123
rbd: error opening image test123: (2) No such file or directory
ubuntu@plana140:~$ rbd --cluster cluster2 -p mirror mirror image status test123
test123:
  global_id:  be56d70d-3bd7-4c4e-b299-98b7af31463c
  state:      down+unknown
  description: status not found
  last_update: 1970-01-01 08:00:00
ubuntu@plana140:~$ rbd --cluster cluster2 -p mirror mirror image status test123
2016-07-29 15:16:17.456285 7f3be4362700 0 -- 192.168.34.140:0/3001668689 >> 192.168.34.140:6804/5666 pipe(0x24af440 sd=8 :0 s=1 pgs=0 cs=0 l=1 c=0x24b06f0).fault
test123:
  global_id:  be56d70d-3bd7-4c4e-b299-98b7af31463c
  state:      up+stopped
  description: remote image is non-primary or local image is primary
  last_update: 2016-07-29 15:16:13
ubuntu@plana140:~$ rbd --cluster cluster1 -p mirror mirror image status test123
test123:
  global_id:  be56d70d-3bd7-4c4e-b299-98b7af31463c
  state:      up+replaying
  description: replaying, master_position=[object_number=3, tag_tid=1, entry_tid=3], mirror_position=[], entries_behind_master=3
  last_update: 2016-07-29 15:16:13
ubuntu@plana140:~$
ubuntu@plana140:~$ *rbd --cluster=cluster2 mirror image demote mirror/test123*
2016-07-29 15:16:45.271475 7f4d857fa700 0 -- 192.168.34.140:0/1404170656 >> 192.168.34.140:6804/5666 pipe(0x864f90 sd=8 :53652 s=2 pgs=209 cs=1 l=1 c=0x866240).injecting socket failure
2016-07-29 15:16:45.271644 7f4d867fc700 0 -- 192.168.34.140:0/1404170656 submit_message osd_op(client.4235.0:19 2.daeeba7f journal.1085580bd78f [call journal.get_minimum_set,call journal.get_active_set] snapc 0=[] ack+read+known_if_redirected e24) v7 remote, 192.168.34.140:6804/5666, failed lorry con, dropping message 0x7f4d6c00da20
Image demoted to non-primary
ubuntu@plana140:~$ rbd --cluster cluster1 -p mirror mirror image status test123
2016-07-29 15:16:54.758774 7efff07f8700 0 -- 192.168.34.140:0/113640562 >> 192.168.34.135:6804/5633 pipe(0x22310d0 sd=9 :48824 s=2 pgs=270 cs=1 l=1 c=0x222e770).injecting socket failure
```

```

test123:
  global_id:   be56d70d-3bd7-4c4e-b299-98b7af31463c
  state:       up+stopped
  description: stopped
  last_update: 2016-07-29 15:16:45
ubuntu@plana140:~$ rbd --cluster cluster2 -p mirror mirror image status test123
test123:
  global_id:   be56d70d-3bd7-4c4e-b299-98b7af31463c
  state:       up+stopped
  description: remote image is non-primary or local image is primary
  last_update: 2016-07-29 15:16:42
ubuntu@plana140:~$ *rbd --cluster=cluster2 mirror image promote mirror/test123*
2016-07-29 15:17:16.756244 7f85547f8700 0 -- 192.168.34.140:0/3364772875 >> 192.168.34.140:6804/5
666 pipe(0xd56200 sd=8 :53734 s=2 pgs=214 cs=1 l=1 c=0xd574b0).injecting socket failure
Image promoted to primary
ubuntu@plana140:~$ rbd --cluster cluster2 -p mirror mirror image status test123
test123:
  global_id:   be56d70d-3bd7-4c4e-b299-98b7af31463c
  state:       up+stopped
  description: remote image is non-primary or local image is primary
  last_update: 2016-07-29 15:17:12
ubuntu@plana140:~$ rbd --cluster cluster2 -p mirror mirror image status test123
test123:
  global_id:   be56d70d-3bd7-4c4e-b299-98b7af31463c
  state:       up+stopped
  description: remote image is non-primary or local image is primary
  last_update: 2016-07-29 15:17:12
ubuntu@plana140:~$ rbd --cluster cluster1 -p mirror mirror image status test123
test123:
  global_id:   be56d70d-3bd7-4c4e-b299-98b7af31463c
  state:       up+stopped
  description: remote image is non-primary or local image is primary
  last_update: 2016-07-29 15:17:10
ubuntu@plana140:~$ rbd --cluster cluster1 -p mirror mirror image status test123
test123:
  global_id:   be56d70d-3bd7-4c4e-b299-98b7af31463c
  state:       up+stopped
  description: remote image is non-primary or local image is primary
  last_update: 2016-07-29 15:17:10
ubuntu@plana140:~$ rbd --cluster cluster1 -p mirror mirror image status test123
test123:
  global_id:   be56d70d-3bd7-4c4e-b299-98b7af31463c
  state:       up+stopped
  description: remote image is non-primary or local image is primary
  last_update: 2016-07-29 15:17:10
ubuntu@plana140:~$ rbd --cluster cluster1 -p mirror mirror image status test123
test123:
  global_id:   be56d70d-3bd7-4c4e-b299-98b7af31463c
  state:       up+stopped
  description: remote image is non-primary or local image is primary
  last_update: 2016-07-29 15:17:10
ubuntu@plana140:~$ rbd --cluster cluster1 -p mirror mirror image status test123
test123:
  global_id:   be56d70d-3bd7-4c4e-b299-98b7af31463c
  state:       up+stopped
  description: remote image is non-primary or local image is primary
  last_update: 2016-07-29 15:17:10
ubuntu@plana140:~$ rbd --cluster cluster1 -p mirror mirror image status test123
test123:
  global_id:   be56d70d-3bd7-4c4e-b299-98b7af31463c
  state:       up+syncing
  description: bootstrapping, OPEN_REMOTE_IMAGE
  last_update: 2016-07-29 15:17:39
ubuntu@plana140:~$ rbd --cluster cluster1 -p mirror mirror image status test123
test123:
  global_id:   be56d70d-3bd7-4c4e-b299-98b7af31463c
  * state:     up+error*

```

```
description: error bootstrapping replay
last_update: 2016-07-29 15:17:40
```

```
ubuntu@plana140:~$ ceph -v
ceph version v11.0.0-1071-g066661c5 (066661c536cfac48ea1f3d11b8d46bc5a6a0d7c44)
ubuntu@plana140:~$
```

Related issues:

Copied to rbd - Backport #17065: jewel: rbd mirror: after promote, the mirror...

Resolved

History

#1 - 07/29/2016 12:34 PM - Jason Dillaman

- Status changed from New to Need More Info
- Assignee deleted (Jason Dillaman)
- Priority changed from High to Normal

@de lan: can you please provide a debug log from the rbd-mirror daemon executing against cluster 'cluster1'? You need to add "debug rbd_mirror = 20" and "debug rbd = 20" to the cluster1's ceph config. I am going to guess that the image "test123" is split-brained between your two clusters and you need to request a resync via "rbd --cluster cluster1 --pool mirror mirror image resync test123", but the logs should provide more detail as to why it cannot sync the two images.

#2 - 08/02/2016 06:30 AM - de lan

- File cluster1-client.mirror.3769.log added
- File cluster1-client.mirror.3772.log added

@Jason Dillaman

Hi.

I have reproduced it, and taken the log.

It shows that the image is split-brained, and it didn't recover when I do "rbd --cluster cluster1 --pool mirror mirror image resync test"

I have doubt that the probability of error is very high when I do the demote and promote operation. If I do something wrong?

thanks!

#3 - 08/08/2016 01:09 PM - Jason Dillaman

@de lan: it doesn't look like debugging was enabled for the first log. Can you please provide the exact steps you perform and how your system is configured?

#4 - 08/08/2016 01:10 PM - Jason Dillaman

- Description updated

#5 - 08/08/2016 01:25 PM - Jason Dillaman

- Status changed from Need More Info to In Progress
- Assignee set to Jason Dillaman
- Backport set to jewel

Actually, I was able to repeat the issue. Thanks.

#6 - 08/12/2016 01:38 AM - Jason Dillaman

- Status changed from In Progress to Need Review

PR: <https://github.com/ceph/ceph/pull/10703>

#7 - 08/17/2016 03:17 PM - Mykola Golub

- Status changed from Need Review to Pending Backport

#8 - 08/17/2016 07:46 PM - Loic Dachary

- Copied to Backport #17065: jewel: rbd mirror: after promote, the mirror image often be up+error added

#9 - 08/19/2016 05:33 AM - Loic Dachary

- Subject changed from rbd mirror: after promote ,zhe mirror image often be up+error to rbd mirror: after promote, the mirror image often be up+error

#10 - 10/17/2016 03:14 PM - Loic Dachary

- Status changed from Pending Backport to Resolved

Files

cluster1-client.mirror.3769.log	2.11 KB	08/02/2016	de lan
cluster1-client.mirror.3772.log	226 KB	08/02/2016	de lan