

Ceph - Backport #15171

hammer: osd: corruption when min_read_recency_for_promote > 1

03/17/2016 12:22 PM - Sage Weil

Status:	Resolved	Spent time:	0.00 hour
Priority:	Urgent		
Assignee:	Sage Weil		
Target version:	v0.94.7		
Release:	hammer		
Description			
https://github.com/ceph/ceph/pull/8187			
<p>just got done with a test against a build of 0.94.6 minus the two commits that were backported in PR 7207. everything worked as it should with the cache-mode set to writeback and the min_read_recency_for_promote set to 2. assuming it works properly on master, there must be a commit that we're missing on the backport to support this properly.</p>			
<p>sage, i'm adding you to the recipients on this so hopefully you see it. the tl;dr version is that the backport of the cache recency fix to hammer doesn't work right and potentially corrupts data when the min_read_recency_for_promote is set to greater than 1.</p>			
<p>mike</p>			
<p>On Wed, Mar 16, 2016 at 4:41 PM, Mike Lovell <mike.lovell@endurance.com> wrote:</p>			
<pre>> robert and i have done some further investigation the past couple days on > this. we have a test environment with a hard drive tier and an ssd tier as > a cache. several vms were created with volumes from the ceph cluster. i did > a test in each guest where i un-tarred the linux kernel source multiple > times and then did a md5sum check against all of the files in the resulting > source tree. i started off with the monitors and osds running 0.94.5 and > never saw any problems. > > a single node was then upgraded to 0.94.6 which has osds in both the ssd > and hard drive tier. i then proceeded to run the same test and, while the > untar and md5sum operations were running, i changed the ssd tier cache-mode > from forward to writeback. almost immediately the vms started reporting io > errors and odd data corruption. the remainder of the cluster was updated to > 0.94.6, including the monitors, and the same thing happened. > > things were cleaned up and reset and then a test was run > where min_read_recency_for_promote for the ssd cache pool was set to 1. we > previously had it set to 6. there was never an error with the recency > setting set to 1. i then tested with it set to 2 and it immediately caused > failures. we are currently thinking that it is related to the backport of > the fix for the recency promotion and are in progress of making a .6 build > without that backport to see if we can cause corruption. is anyone using a > version from after the original recency fix (PR 6702) with a cache tier in > writeback mode? anyone have a similar problem?</pre>			
<p>> mike</p>			
<p>> On Mon, Mar 14, 2016 at 8:51 PM, Mike Lovell <mike.lovell@endurance.com> wrote:</p>			
<pre>>> something weird happened on one of the ceph clusters that i administer >> tonight which resulted in virtual machines using rbd volumes seeing >> corruption in multiple forms.</pre>			

```
>>
>> when everything was fine earlier in the day, the cluster was a number of
>> storage nodes spread across 3 different roots in the crush map. the first
>> bunch of storage nodes have both hard drives and ssds in them with the hard
>> drives in one root and the ssds in another. there is a pool for each and
>> the pool for the ssds is a cache tier for the hard drives. the last set of
>> storage nodes were in a separate root with their own pool that is being
>> used for burn in testing.
>>
>> these nodes had run for a while with test traffic and we decided to move
>> them to the main root and pools. the main cluster is running 0.94.5 and the
>> new nodes got 0.94.6 due to them getting configured after that was
>> released. i removed the test pool and did a ceph osd crush move to move the
>> first node into the main cluster, the hard drives into the root for that
>> tier of storage and the ssds into the root and pool for the cache tier.
>> each set was done about 45 minutes apart and they ran for a couple hours
>> while performing backfill without any issue other than high load on the
>> cluster.
>>
>> we normally run the ssd tier in the forward cache-mode due to the ssds we
>> have not being able to keep up with the io of writeback. this results in io
>> on the hard drives slowing going up and performance of the cluster starting
>> to suffer. about once a week, i change the cache-mode between writeback and
>> forward for short periods of time to promote actively used data to the
>> cache tier. this moves io load from the hard drive tier to the ssd tier and
>> has been done multiple times without issue. i normally don't do this while
>> there are backfills or recoveries happening on the cluster but decided to
>> go ahead while backfill was happening due to the high load.
>>
>> i tried this procedure to change the ssd cache-tier between writeback and
>> forward cache-mode and things seemed okay from the ceph cluster. about 10
>> minutes after the first attempt a changing the mode, vms using the ceph
>> cluster for their storage started seeing corruption in multiple forms. the
>> mode was flipped back and forth multiple times in that time frame and its
>> unknown if the corruption was noticed with the first change or subsequent
>> changes. the vms were having issues of filesystems having errors and
>> getting remounted RO and mysql databases seeing corruption (both myisam and
>> innodb). some of this was recoverable but on some filesystems there was
>> corruption that lead to things like lots of data ending up in the
>> lost+found and some of the databases were un-recoverable (backups are
>> helping there).
>>
>> i'm not sure what would have happened to cause this corruption. the
>> libvirt logs for the qemu processes for the vms did not provide any output
>> of problems from the ceph client code. it doesn't look like any of the qemu
>> processes had crashed. also, it has now been several hours since this
>> happened with no additional corruption noticed by the vms. it doesn't
>> appear that we had any corruption happen before i attempted the flipping of
>> the ssd tier cache-mode.
>>
>> the only think i can think of that is different between this time doing
>> this procedure vs previous attempts was that there was the one storage node
>> running 0.94.6 where the remainder were running 0.94.5. is is possible that
>> something changed between these two releases that would have caused
>> problems with data consistency related to the cache tier? or otherwise? any
>> other thoughts or suggestions?
>>
>> thanks in advance for any help you can provide.
>>
>> mike
>>
>
>
```

Related issues:

Related to Ceph - Bug #12814: flipping the overlay from forward to seems to r...

Can't reproduc08/27/2015

Related to Ceph - Bug #12469: RadosModel.h: read returned error code -2 (hammer)

Can't reproduc07/26/2015

Associated revisions

Revision 7eae05e1 - 03/17/2016 05:28 PM - Sage Weil

osd/ReplicatedPG: do not proxy read **and** process op locally

If we proxy the read, we cannot return false, even if we fail to also trigger a promotion.

Fixes: #15171

Signed-off-by: Sage Weil <sage@redhat.com>

History

#1 - 03/17/2016 01:00 PM - Irek Fasikhov

Sage Weil wrote:

[...]

I confirm the problem. When `min_read_recency_for_promote > 1` data failure.

#2 - 03/17/2016 01:09 PM - Sage Weil

<https://github.com/ceph/ceph-qa-suite/pull/888>

#3 - 03/17/2016 01:14 PM - Sage Weil

<https://github.com/ceph/ceph-qa-suite/pull/889>

#4 - 03/17/2016 01:15 PM - Sage Weil

- Related to Bug #12814: flipping the overlay from forward to seems to reorder writes added

#5 - 03/17/2016 05:33 PM - Sage Weil

- Subject changed from `osd: corruption when min_read_recency_for_promote > 1 (hammer, master?)` to `osd: corruption when min_read_recency_for_promote > 1 (hammer)`

- Status changed from *Verified* to *Testing*

<https://github.com/ceph/ceph/pull/8187>

#6 - 03/17/2016 07:52 PM - Robert LeBlanc

- Release set to *hammer*

- Affected Versions *v0.94.6* added

<https://github.com/ceph/ceph/pull/8187> appears to resolve the issue for us, we can't replicate the corruption with either `ceph_test_rados` or the VM tests.

#7 - 03/18/2016 01:45 AM - David Zafman

- Related to Bug #12469: RadosModel.h: read returned error code -2 (hammer) added

#8 - 03/18/2016 03:12 PM - Loic Dachary

- Tracker changed from Bug to Backport

#9 - 03/18/2016 03:14 PM - Loic Dachary

- Description updated

#10 - 03/18/2016 03:15 PM - Loic Dachary

- Subject changed from *osd: corruption when min_read_recency_for_promote > 1 (hammer)* to *osd: corruption when min_read_recency_for_promote > 1*

#11 - 03/18/2016 03:15 PM - Loic Dachary

- Subject changed from *osd: corruption when min_read_recency_for_promote > 1* to *hammer: osd: corruption when min_read_recency_for_promote > 1*

#12 - 03/18/2016 03:19 PM - Loic Dachary

@sage do you have a run of <https://github.com/ceph/ceph-qa-suite/pull/889> somewhere already ?

#13 - 04/30/2016 04:46 PM - Nathan Cutler

- Status changed from *Testing* to *Resolved*

#14 - 10/10/2016 01:55 PM - Loic Dachary

- Target version set to v0.94.7