

rbd - Bug #14373

Cannot reliably create snapshot after freezing QEMU IO

01/14/2016 03:28 PM - Jason Dillaman

Status: Resolved	% Done: 0%
Priority: High	Spent time: 0.00 hour
Assignee:	
Category:	
Target version:	
Source: Community (user)	Affected Versions:
Tags:	ceph-qa-suite:
Backport: hammer, infernalis	Pull request ID:
Regression: No	Crash signature (v1):
Severity: 3 - minor	Crash signature (v2):
Reviewed:	
Description	
<pre>root@slpeah001:[~]:# rbd snap create volumes/volume-26c89a0a-be4d-45d4-85a6-e0dc134941fd --snap test 2016-01-13 12:04:39.107166 7fb70e4c2880 -1 librbd::ImageWatcher: 0x427a710 no lock owners detected</pre>	
http://www.spinics.net/lists/ceph-users/msg24639.html	
Related issues:	
Copied to rbd - Backport #14542: infernalis: Cannot reliably create snapshot ...	Resolved
Copied to rbd - Backport #14543: hammer: Cannot reliably create snapshot afte...	Resolved

History

#1 - 01/14/2016 05:00 PM - Angapov Vasily

Actually, for now snapshots are sometimes hanging even without calling FSfreeze in libvirt, so I believe the problem lies not in quiescing probably. The bad thing is that most of the time snapshots go well, but only sometimes they fail, resulting in total disk loss by VM, meaning that is unable to write data to RBD image anymore until VM hard reboot.

Attached is one-second log from one of failed snapshot guests with "debug rbd = 20".

#2 - 01/14/2016 06:16 PM - Angapov Vasily

For some reason file attachment is not working.

Here is the log:

<https://www.dropbox.com/s/dzgux81enrukd90/qemu-guest-73037.log?dl=0>

#3 - 01/14/2016 07:12 PM - Jason Dillaman

Here is what I see in the provided log:

- volume-26c89a0a-be4d-45d4-85a6-e0dc134941fd is opened @ 04:49:13.858469
- first write to image @ 04:49:27.332004
 - fails to acquire exclusive lock -- owned by client.3425137@192.168.254.17:0/2112235
 - did you force kill this VM process shortly before rebooting it?

- blacklists old exclusive lock owner @ 04:49:40.856433
- blocked IO resumes

This log unfortunately doesn't include any requests to create a snapshot and is what I would expect to see if you force-killed a qemu-kvm process while it owned the exclusive lock (i.e. it waits 30 seconds for the watch to time-out before blacklisting the old client and stealing the lock).

#4 - 01/15/2016 03:42 PM - Angapov Vasily

Jason, qemu-kvm process was alive all the time, but when snapshot taking became stuck - VM completely lost its disk. And actually the snapshot taking process was hanging forever, so there is not 30 seconds timeout. Can I provide any more logs for you?

#5 - 01/15/2016 03:58 PM - Angapov Vasily

You are telling that blocked IO will resume after 30 seconds - how can I decrease that timeout? And also I was trying to break the exclusive lock right before the snapshot but libvirt immediately switches VM disk to read-only mode somehow (disk inside VM immediately becomes read-only), so is this expected also? My libvirt version is 1.2.17, qemu-kvm is 1.5.3.

#6 - 01/15/2016 06:21 PM - Jason Dillaman

The logs only cover 04:49:13.858469 (when qemu-kvm first opened the device) to 04:49:59.784578. In these logs, there is no mention of a snapshot being created or the exclusive lock being lost by qemu-kvm. Instead, there is a 22 second delay to acquire the exclusive lock because a previous process did not cleanly shutdown. After the watch timeout expires, the lock is correctly acquired and AIO operations resume through the end of the provided log. You can reduce the watch timeout on the OSDs via the 'osd client watch timeout' setting.

I am not sure what you are referring to when you say you tried to break the exclusive lock. What were you doing?

If at all possible, I would like to see logs from both qemu-kvm and 'rbd snap create' -- with the qemu-kvm log encompassing the timespan before and after the snapshot create attempt.

#7 - 01/20/2016 07:50 AM - Angapov Vasily

Here is it. Hope that will help.

```
2016-01-20 10:33:47.954001 7fadb64ac700 20 librbd::AioRequest: send_pre 0x56262fd06670 rbd_data.35f26e2ae8944a.0000000000000221
3928064~4096
2016-01-20 10:33:47.954003 7fadb64ac700 20 librbd::AioRequest: send_write 0x56262fd06670 rbd_data.35f26e2ae8944a.0000000000000221
3928064~4096
2016-01-20 10:33:47.954037 7fadb64ac700 20 librbd::AioCompletion: AioCompletion::finish_adding_requests 0x5626368d6180 pending 2
2016-01-20 10:33:47.954042 7fadb64ac700 20 librbd::AioCompletion: AioCompletion::complete_request() 0x5626368d6180
complete_cb=0x56262c633440 pending 2
2016-01-20 10:33:47.956558 7fadb84b0700 20 librbd::AioRequest: write 0x56262fd06670 rbd_data.35f26e2ae8944a.0000000000000221
3928064~4096 should_complete: r = 0
2016-01-20 10:33:47.956566 7fadb84b0700 20 librbd::AioRequest: WRITE_FLAT
2016-01-20 10:33:47.956567 7fadb84b0700 20 librbd::AioRequest: complete 0x56262fd06670
2016-01-20 10:33:47.956569 7fadb84b0700 20 librbdwriteback: C_OrderedWrite completing 0x562635bd2740
2016-01-20 10:33:47.956570 7fadb84b0700 20 librbdwriteback: complete_writes() oid rbd_data.35f26e2ae8944a.0000000000000221
2016-01-20 10:33:47.956572 7fadb84b0700 20 librbdwriteback: complete_writes() completing 0x562635bd2740
2016-01-20 10:33:47.956577 7fadb84b0700 20 librbd::AioCompletion: AioCompletion::complete_request() 0x5626368d6180
complete_cb=0x56262c633440 pending 1
2016-01-20 10:33:47.956579 7fadb84b0700 20 librbd::AioCompletion: AioCompletion::finalize() 0x5626368d6180 rval 0 read_buf 0 read_bl 0
2016-01-20 10:33:47.956581 7fadb84b0700 20 librbd::AsyncOperation: 0x5626368d62b0 finish_op
2016-01-20 10:33:47.956589 7fadb84b0700 20 librbdwriteback: C_OrderedWrite finished 0x562635bd2740
```

2016-01-20 10:33:48.284879 7fad64ac700 20 librbd: aio_write 0x56262fd76000 off = 1048576 len = 4096 buf = 0x562636838000
2016-01-20 10:33:48.284902 7fad64ac700 20 librbd: ictx_check 0x56262fd76000
2016-01-20 10:33:48.284906 7fad64ac700 20 librbd::AsyncOperation: 0x5626368d6730 start_op
2016-01-20 10:33:48.284921 7fad64ac700 20 librbd: oid rbd_data.35f26e2ae8944a.0000000000000000 1048576~4096 from [0,4096]
2016-01-20 10:33:48.285000 7fad64ac700 20 librbd::AioCompletion: AioCompletion::finish_adding_requests 0x5626368d6600 pending 1
2016-01-20 10:33:48.285009 7fad64ac700 20 librbd: aio_write 0x56262fd76000 off = 1056768 len = 4096 buf = 0x5626368de000
2016-01-20 10:33:48.285014 7fad64ac700 20 librbd: ictx_check 0x56262fd76000
2016-01-20 10:33:48.285016 7fad64ac700 20 librbd::AsyncOperation: 0x5626368d6bb0 start_op
2016-01-20 10:33:48.285012 7fad64ae700 20 librbd::AioCompletion: AioCompletion::complete_request() 0x5626368d6600
complete_cb=0x56262c633440 pending 1
2016-01-20 10:33:48.285022 7fad64ac700 20 librbd: oid rbd_data.35f26e2ae8944a.0000000000000000 1056768~4096 from [0,4096]
2016-01-20 10:33:48.285025 7fad64ae700 20 librbd::AioCompletion: AioCompletion::finalize() 0x5626368d6600 rval 0 read_buf 0 read_bl 0
2016-01-20 10:33:48.285029 7fad64ae700 20 librbd::AsyncOperation: 0x5626368d6730 finish_op
2016-01-20 10:33:48.285071 7fad64ac700 20 librbd::AioCompletion: AioCompletion::finish_adding_requests 0x5626368d6a80 pending 1
2016-01-20 10:33:48.285076 7fad64ae700 20 librbd::AioCompletion: AioCompletion::complete_request() 0x5626368d6a80
complete_cb=0x56262c633440 pending 1
2016-01-20 10:33:48.285077 7fad64ac700 20 librbd: aio_write 0x56262fd76000 off = 15033438208 len = 4096 buf = 0x562635e57000
2016-01-20 10:33:48.285082 7fad64ac700 20 librbd: ictx_check 0x56262fd76000
2016-01-20 10:33:48.285082 7fad64ae700 20 librbd::AioCompletion: AioCompletion::finalize() 0x5626368d6a80 rval 0 read_buf 0 read_bl 0
2016-01-20 10:33:48.285084 7fad64ac700 20 librbd::AsyncOperation: 0x56262fe4f8b0 start_op
2016-01-20 10:33:48.285084 7fad64ae700 20 librbd::AsyncOperation: 0x5626368d6bb0 finish_op
2016-01-20 10:33:48.285092 7fad64ac700 20 librbd: oid rbd_data.35f26e2ae8944a.00000000000000e0 1052672~4096 from [0,4096]
2016-01-20 10:33:48.285117 7fad64ac700 20 librbd::AioCompletion: AioCompletion::finish_adding_requests 0x56262fe4f780 pending 1
2016-01-20 10:33:48.285121 7fad64ae700 20 librbd::AioCompletion: AioCompletion::complete_request() 0x56262fe4f780
complete_cb=0x56262c633440 pending 1
2016-01-20 10:33:48.285124 7fad64ae700 20 librbd::AioCompletion: AioCompletion::finalize() 0x56262fe4f780 rval 0 read_buf 0 read_bl 0
2016-01-20 10:33:48.285121 7fad64ac700 20 librbd: aio_write 0x56262fd76000 off = 15033499648 len = 4096 buf = 0x562632354000
2016-01-20 10:33:48.285126 7fad64ac700 20 librbd: ictx_check 0x56262fd76000
2016-01-20 10:33:48.285128 7fad64ac700 20 librbd::AsyncOperation: 0x56262fe4efb0 start_op
2016-01-20 10:33:48.285126 7fad64ae700 20 librbd::AsyncOperation: 0x56262fe4f8b0 finish_op
2016-01-20 10:33:48.285133 7fad64ac700 20 librbd: oid rbd_data.35f26e2ae8944a.00000000000000e0 1114112~4096 from [0,4096]
2016-01-20 10:33:48.285149 7fad64ac700 20 librbd::AioCompletion: AioCompletion::finish_adding_requests 0x56262fe4ee80 pending 1
2016-01-20 10:33:48.285153 7fad64ac700 20 librbd: aio_write 0x56262fd76000 off = 15033643008 len = 4096 buf = 0x562632918000
2016-01-20 10:33:48.285156 7fad64ac700 20 librbd: ictx_check 0x56262fd76000
2016-01-20 10:33:48.285154 7fad64ae700 20 librbd::AioCompletion: AioCompletion::complete_request() 0x56262fe4ee80
complete_cb=0x56262c633440 pending 1
2016-01-20 10:33:48.285158 7fad64ac700 20 librbd::AsyncOperation: 0x562631ec8230 start_op
2016-01-20 10:33:48.285157 7fad64ae700 20 librbd::AioCompletion: AioCompletion::finalize() 0x56262fe4ee80 rval 0 read_buf 0 read_bl 0
2016-01-20 10:33:48.285161 7fad64ae700 20 librbd::AsyncOperation: 0x56262fe4efb0 finish_op
2016-01-20 10:33:48.285163 7fad64ac700 20 librbd: oid rbd_data.35f26e2ae8944a.00000000000000e0 1257472~4096 from [0,4096]
2016-01-20 10:33:48.285181 7fad64ac700 20 librbd::AioCompletion: AioCompletion::finish_adding_requests 0x562631ec8100 pending 1
2016-01-20 10:33:48.285185 7fad64ac700 20 librbd: aio_write 0x56262fd76000 off = 15035969536 len = 4096 buf = 0x562633369000
2016-01-20 10:33:48.285187 7fad64ac700 20 librbd: ictx_check 0x56262fd76000
2016-01-20 10:33:48.285185 7fad64ae700 20 librbd::AioCompletion: AioCompletion::complete_request() 0x562631ec8100
complete_cb=0x56262c633440 pending 1
2016-01-20 10:33:48.285190 7fad64ac700 20 librbd::AsyncOperation: 0x5626368d8530 start_op
2016-01-20 10:33:48.285190 7fad64ae700 20 librbd::AioCompletion: AioCompletion::finalize() 0x562631ec8100 rval 0 read_buf 0 read_bl 0
2016-01-20 10:33:48.285193 7fad64ae700 20 librbd::AsyncOperation: 0x562631ec8230 finish_op
2016-01-20 10:33:48.285195 7fad64ac700 20 librbd: oid rbd_data.35f26e2ae8944a.00000000000000e0 3584000~4096 from [0,4096]
2016-01-20 10:33:48.285211 7fad64ac700 20 librbd::AioCompletion: AioCompletion::finish_adding_requests 0x5626368d8400 pending 1
2016-01-20 10:33:48.285214 7fad64ac700 20 librbd: aio_write 0x56262fd76000 off = 15070990336 len = 4096 buf = 0x562631cad000
2016-01-20 10:33:48.285217 7fad64ac700 20 librbd: ictx_check 0x56262fd76000
2016-01-20 10:33:48.285215 7fad64ae700 20 librbd::AioCompletion: AioCompletion::complete_request() 0x5626368d8400
complete_cb=0x56262c633440 pending 1
2016-01-20 10:33:48.285219 7fad64ac700 20 librbd::AsyncOperation: 0x562631ec7330 start_op
2016-01-20 10:33:48.285219 7fad64ae700 20 librbd::AioCompletion: AioCompletion::finalize() 0x5626368d8400 rval 0 read_buf 0 read_bl 0
2016-01-20 10:33:48.285222 7fad64ae700 20 librbd::AsyncOperation: 0x5626368d8530 finish_op
2016-01-20 10:33:48.285223 7fad64ac700 20 librbd: oid rbd_data.35f26e2ae8944a.00000000000000e0 856064~4096 from [0,4096]
2016-01-20 10:33:48.285241 7fad64ac700 20 librbd::AioCompletion: AioCompletion::finish_adding_requests 0x562631ec7200 pending 1
2016-01-20 10:33:48.285246 7fad64ae700 20 librbd::AioCompletion: AioCompletion::complete_request() 0x562631ec7200
complete_cb=0x56262c633440 pending 1
2016-01-20 10:33:48.285250 7fad64ae700 20 librbd::AioCompletion: AioCompletion::finalize() 0x562631ec7200 rval 0 read_buf 0 read_bl 0
2016-01-20 10:33:48.285252 7fad64ae700 20 librbd::AsyncOperation: 0x562631ec7330 finish_op
2016-01-20 10:33:48.315363 7fad64b0700 10 librbd::ImageWatcher: 0x56262fcc0b40 remote snap_create request: test
2016-01-20 10:33:48.315378 7fad64b0700 20 librbd: snap_create_helper 0x56262fd76000 test
2016-01-20 10:33:48.315380 7fad64b0700 20 librbd: ictx_check 0x56262fd76000
2016-01-20 10:33:48.315445 7fad64b0700 20 librbdwriteback: write will wait for result 0x5626368a4ee0
2016-01-20 10:33:48.315463 7fad64b0700 10 librbd::ImageCtx: prune_parent_extents image overlap 0, object overlap 0 from image extents []
2016-01-20 10:33:48.315479 7fad64b0700 20 librbd::AioRequest: send 0x56262fd06670 rbd_data.35f26e2ae8944a.0000000000000000
1048576~4096
2016-01-20 10:33:48.315484 7fad64b0700 20 librbd::AioRequest: send_pre 0x56262fd06670 rbd_data.35f26e2ae8944a.0000000000000000
1048576~4096
2016-01-20 10:33:48.315489 7fad64b0700 20 librbd::AioRequest: send_write 0x56262fd06670 rbd_data.35f26e2ae8944a.0000000000000000
1048576~4096
2016-01-20 10:33:48.315567 7fad64b0700 20 librbdwriteback: write will wait for result 0x562630e1e020

2016-01-20 10:33:48.315576 7fad84b0700 10 librbd::ImageCtx: prune_parent_extents image overlap 0, object overlap 0 from image extents []
2016-01-20 10:33:48.315581 7fad84b0700 20 librbd::AioRequest: send 0x56262fd06d00 rbd_data.35f26e2ae8944a.000000000000e001052672~4096
2016-01-20 10:33:48.315584 7fad84b0700 20 librbd::AioRequest: send_pre 0x56262fd06d00 rbd_data.35f26e2ae8944a.000000000000e001052672~4096
2016-01-20 10:33:48.315587 7fad84b0700 20 librbd::AioRequest: send_write 0x56262fd06d00 rbd_data.35f26e2ae8944a.000000000000e001052672~4096
2016-01-20 10:33:48.315639 7fad84b0700 20 librbdwriteback: write will wait for result 0x562630e1eb40
2016-01-20 10:33:48.315643 7fad84b0700 10 librbd::ImageCtx: prune_parent_extents image overlap 0, object overlap 0 from image extents []
2016-01-20 10:33:48.315654 7fad84b0700 20 librbd::AioRequest: send 0x562630675ef0 rbd_data.35f26e2ae8944a.000000000000e001257472~4096
2016-01-20 10:33:48.315656 7fad84b0700 20 librbd::AioRequest: send_pre 0x562630675ef0 rbd_data.35f26e2ae8944a.000000000000e001257472~4096
2016-01-20 10:33:48.315660 7fad84b0700 20 librbd::AioRequest: send_write 0x562630675ef0 rbd_data.35f26e2ae8944a.000000000000e001257472~4096
2016-01-20 10:33:48.315695 7fad84b0700 20 librbdwriteback: write will wait for result 0x562636ef9640
2016-01-20 10:33:48.315701 7fad84b0700 10 librbd::ImageCtx: prune_parent_extents image overlap 0, object overlap 0 from image extents []
2016-01-20 10:33:48.315705 7fad84b0700 20 librbd::AioRequest: send 0x56262fd051d0 rbd_data.35f26e2ae8944a.0000000000000001056768~4096
2016-01-20 10:33:48.315708 7fad84b0700 20 librbd::AioRequest: send_pre 0x56262fd051d0 rbd_data.35f26e2ae8944a.0000000000000001056768~4096
2016-01-20 10:33:48.315711 7fad84b0700 20 librbd::AioRequest: send_write 0x56262fd051d0 rbd_data.35f26e2ae8944a.0000000000000001056768~4096
2016-01-20 10:33:48.315747 7fad84b0700 20 librbdwriteback: write will wait for result 0x562630e20820
2016-01-20 10:33:48.315750 7fad84b0700 10 librbd::ImageCtx: prune_parent_extents image overlap 0, object overlap 0 from image extents []
2016-01-20 10:33:48.315754 7fad84b0700 20 librbd::AioRequest: send 0x56262fd07930 rbd_data.35f26e2ae8944a.000000000000e003584000~4096
2016-01-20 10:33:48.315757 7fad84b0700 20 librbd::AioRequest: send_pre 0x56262fd07930 rbd_data.35f26e2ae8944a.000000000000e003584000~4096
2016-01-20 10:33:48.315760 7fad84b0700 20 librbd::AioRequest: send_write 0x56262fd07930 rbd_data.35f26e2ae8944a.000000000000e003584000~4096
2016-01-20 10:33:48.315790 7fad84b0700 20 librbdwriteback: write will wait for result 0x5626305c3900
2016-01-20 10:33:48.315793 7fad84b0700 10 librbd::ImageCtx: prune_parent_extents image overlap 0, object overlap 0 from image extents []
2016-01-20 10:33:48.315796 7fad84b0700 20 librbd::AioRequest: send 0x562630674780 rbd_data.35f26e2ae8944a.000000000000e09856064~4096
2016-01-20 10:33:48.315800 7fad84b0700 20 librbd::AioRequest: send_pre 0x562630674780 rbd_data.35f26e2ae8944a.000000000000e09856064~4096
2016-01-20 10:33:48.315802 7fad84b0700 20 librbd::AioRequest: send_write 0x562630674780 rbd_data.35f26e2ae8944a.000000000000e09856064~4096
2016-01-20 10:33:48.315860 7fad84b0700 20 librbdwriteback: write will wait for result 0x562630fc1ce0
2016-01-20 10:33:48.315872 7fad84b0700 10 librbd::ImageCtx: prune_parent_extents image overlap 0, object overlap 0 from image extents []
2016-01-20 10:33:48.315876 7fad84b0700 20 librbd::AioRequest: send 0x56262fd05a40 rbd_data.35f26e2ae8944a.000000000000e001141112~4096
2016-01-20 10:33:48.315879 7fad84b0700 20 librbd::AioRequest: send_pre 0x56262fd05a40 rbd_data.35f26e2ae8944a.000000000000e001141112~4096
2016-01-20 10:33:48.315882 7fad84b0700 20 librbd::AioRequest: send_write 0x56262fd05a40 rbd_data.35f26e2ae8944a.000000000000e001141112~4096
2016-01-20 10:33:48.315914 7fad84b0700 20 librbd::ImageCtx: waiting for cache to be flushed
2016-01-20 10:33:52.360668 7fad84b0700 20 librbd: aio_write 0x56262fd76000 off = 15238873088 len = 16384 buf = 0x562630ca6000
2016-01-20 10:33:52.360682 7fad84b0700 20 librbd: ictx_check 0x56262fd76000
FROM HERE NOTHING MORE HAPPENS IN QEMU LOG!
At the same time VM disk switches to RO and instance effectively hangs...

```
root@slpeah001:[~/angapov/OpenStack/scripts]:# rbd snap create vms/e77120b5-dbd0-4245-ae82-c2bad069fba6_disk --snap test  
2016-01-20 10:33:53.315344 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:33:58.317617 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:34:03.319805 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:34:08.322164 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:34:13.324486 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:34:18.327584 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:34:23.330858 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:34:28.333413 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:34:33.338182 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:34:38.341258 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:34:43.343679 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:34:48.346791 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:34:53.349327 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:34:58.353363 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:35:03.356472 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:35:08.359554 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:35:13.362610 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:35:18.365671 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:35:23.368732 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:35:28.371793 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected  
2016-01-20 10:35:33.374854 7f8e077a3880 -1 librbd::ImageWatcher: 0x396ba00 no lock owners detected
```


#8 - 01/20/2016 08:58 AM - Angapov Vasily

I also should say it is a very floating issue, meaning it may suddenly appear on one VM and after some time suddenly disappear. It seems like it is not a single RBD image problem, but somewhere more globally.

BTW, is there any difference between shell "rbd snap create" and python create_snap() RBD API call?

#9 - 01/20/2016 09:02 AM - Angapov Vasily

One more observation: sometimes there are such records in qemu logs when instance starts hanging:

```
2016-01-20 10:44:20.561026 7f1deba92700 10 librbd::ImageCtx: prune_parent_extents image overlap 0, object overlap 0 from image extents []
2016-01-20 10:44:20.561032 7f1deba92700 20 librbd::AioRequest: send 0x56278963b2c0 rbd_data.35f26e2ae8944a.0000000000000601
188416~4096
2016-01-20 10:44:20.561033 7f1deba92700 20 librbd::AioRequest: send_pre 0x56278963b2c0 rbd_data.35f26e2ae8944a.0000000000000601
188416~4096
2016-01-20 10:44:20.561035 7f1deba92700 20 librbd::AioRequest: send_write 0x56278963b2c0 rbd_data.35f26e2ae8944a.0000000000000601
188416~4096
2016-01-20 10:44:20.561065 7f1deba92700 20 librbd::AioCompletion: AioCompletion::finish_adding_requests 0x5627890f3000 pending 2
2016-01-20 10:44:20.561069 7f1deba92700 20 librbd::AioCompletion: AioCompletion::complete_request() 0x5627890f3000
complete_cb=0x562786e02440 pending 2
2016-01-20 10:44:20.563543 7f1deda96700 20 librbd::AioRequest: write 0x56278963c670 rbd_data.35f26e2ae8944a.0000000000000401
2912256~4096 should_complete: r = 0
2016-01-20 10:44:20.563600 7f1deda96700 20 librbd::AioRequest: WRITE_FLAT
2016-01-20 10:44:20.563604 7f1deda96700 20 librbd::AioRequest: complete 0x56278963c670
2016-01-20 10:44:20.563606 7f1deda96700 20 librbdwriteback: C_OrderedWrite completing 0x56278a52e500
2016-01-20 10:44:20.563608 7f1deda96700 20 librbdwriteback: complete_writes() oid rbd_data.35f26e2ae8944a.0000000000000401
2016-01-20 10:44:20.563611 7f1deda96700 20 librbdwriteback: complete_writes() completing 0x56278a52e500
2016-01-20 10:44:20.563621 7f1deda96700 20 librbdwriteback: C_OrderedWrite finished 0x56278a52e500
2016-01-20 10:44:20.563625 7f1deda96700 20 librbd::AioRequest: write 0x56278963d660 rbd_data.35f26e2ae8944a.0000000000000600
2416640~4096 should_complete: r = 0
2016-01-20 10:44:20.563627 7f1deda96700 20 librbd::AioRequest: WRITE_FLAT
2016-01-20 10:44:20.563628 7f1deda96700 20 librbd::AioRequest: complete 0x56278963d660
2016-01-20 10:44:20.563629 7f1deda96700 20 librbdwriteback: C_OrderedWrite completing 0x56278a52d200
2016-01-20 10:44:20.563630 7f1deda96700 20 librbdwriteback: complete_writes() oid rbd_data.35f26e2ae8944a.0000000000000600
2016-01-20 10:44:20.563632 7f1deda96700 20 librbdwriteback: complete_writes() completing 0x56278a52d200
2016-01-20 10:44:20.563635 7f1deda96700 20 librbdwriteback: C_OrderedWrite finished 0x56278a52d200
2016-01-20 10:44:20.564275 7f1deda96700 20 librbd::AioRequest: write 0x56278963b2c0 rbd_data.35f26e2ae8944a.0000000000000601
188416~4096 should_complete: r = 0
2016-01-20 10:44:20.564284 7f1deda96700 20 librbd::AioRequest: WRITE_FLAT
2016-01-20 10:44:20.564286 7f1deda96700 20 librbd::AioRequest: complete 0x56278963b2c0
2016-01-20 10:44:20.564287 7f1deda96700 20 librbdwriteback: C_OrderedWrite completing 0x56278a52d320
2016-01-20 10:44:20.564289 7f1deda96700 20 librbdwriteback: complete_writes() oid rbd_data.35f26e2ae8944a.0000000000000601
2016-01-20 10:44:20.564291 7f1deda96700 20 librbdwriteback: complete_writes() completing 0x56278a52d320
2016-01-20 10:44:20.564295 7f1deda96700 20 librbd::AioCompletion: AioCompletion::complete_request() 0x5627890f3000
complete_cb=0x562786e02440 pending 1
2016-01-20 10:44:20.564297 7f1deda96700 20 librbd::AioCompletion: AioCompletion::finalize() 0x5627890f3000 rval 0 read_buf 0 read_bl 0
2016-01-20 10:44:20.564299 7f1deda96700 20 librbd::AsyncOperation: 0x5627890f3130 finish_op
2016-01-20 10:44:20.564307 7f1deda96700 20 librbdwriteback: C_OrderedWrite finished 0x56278a52d320
2016-01-20 10:44:20.644808 7f1deba92700 20 librbd: aio_flush 0x56278901a000 completion 0x5627890f0d80
2016-01-20 10:44:20.644822 7f1deba92700 20 librbd: ictx_check 0x56278901a000
2016-01-20 10:44:20.644825 7f1deba92700 20 librbd::AsyncOperation: 0x5627890f0eb0 start_op
2016-01-20 10:44:20.644829 7f1deba92700 20 librbd::AioCompletion: AioCompletion::complete_request() 0x5627890f0d80
complete_cb=0x562786e02440 pending 2
2016-01-20 10:44:20.644833 7f1deba92700 20 librbd::AioCompletion: AioCompletion::finish_adding_requests 0x5627890f0d80 pending 1
2016-01-20 10:44:20.644836 7f1deba92700 20 librbd::AioCompletion: AioCompletion::complete_request() 0x5627890f0d80
complete_cb=0x562786e02440 pending 1
2016-01-20 10:44:20.644837 7f1deba92700 20 librbd::AioCompletion: AioCompletion::finalize() 0x5627890f0d80 rval 0 read_buf 0 read_bl 0
2016-01-20 10:44:20.644839 7f1deba92700 20 librbd::AsyncOperation: 0x5627890f0eb0 finish_op
2016-01-20 10:44:20.645044 7f1deba92700 20 librbd: aio_flush 0x56278901a000 completion 0x5627890f2400
2016-01-20 10:44:20.645051 7f1deba92700 20 librbd: ictx_check 0x56278901a000
2016-01-20 10:44:20.645053 7f1deba92700 20 librbd::AsyncOperation: 0x5627890f2530 start_op
2016-01-20 10:44:20.645054 7f1deba92700 20 librbd::AioCompletion: AioCompletion::complete_request() 0x5627890f2400
complete_cb=0x562786e02440 pending 2
2016-01-20 10:44:20.645057 7f1deba92700 20 librbd::AioCompletion: AioCompletion::finish_adding_requests 0x5627890f2400 pending 1
2016-01-20 10:44:20.645059 7f1deba92700 20 librbd::AioCompletion: AioCompletion::complete_request() 0x5627890f2400
complete_cb=0x562786e02440 pending 1
2016-01-20 10:44:20.645060 7f1deba92700 20 librbd::AioCompletion: AioCompletion::finalize() 0x5627890f2400 rval 0 read_buf 0 read_bl 0
2016-01-20 10:44:20.645061 7f1deba92700 20 librbd::AsyncOperation: 0x5627890f2530 finish_op
2016-01-20 10:44:20.645090 7f1deba92700 20 librbd: aio_flush 0x56278901a000 completion 0x5627890f0f00
2016-01-20 10:44:20.645094 7f1deba92700 20 librbd: ictx_check 0x56278901a000
2016-01-20 10:44:20.645096 7f1deba92700 20 librbd::AsyncOperation: 0x5627890f1030 start_op
```

2016-01-20 10:44:20.645097 7f1deba92700 20 librbd::AioCompletion: AioCompletion::complete_request() 0x5627890f0f00
complete_cb=0x562786e02440 pending 2
2016-01-20 10:44:20.645099 7f1deba92700 20 librbd::AioCompletion: AioCompletion::finish_adding_requests 0x5627890f0f00 pending 1
2016-01-20 10:44:20.645101 7f1deba92700 20 librbd::AioCompletion: AioCompletion::complete_request() 0x5627890f0f00
complete_cb=0x562786e02440 pending 1
2016-01-20 10:44:20.645102 7f1deba92700 20 librbd::AioCompletion: AioCompletion::finalize() 0x5627890f0f00 rval 0 read_buf 0 read_bl 0
2016-01-20 10:44:20.645103 7f1deba92700 20 librbd::AsyncOperation: 0x5627890f1030 finish_op
2016-01-20 10:44:20.645121 7f1e0647fac0 20 librbd: close_image 0x56278901a000
2016-01-20 10:44:20.645128 7f1e0647fac0 10 librbd::ImageCtx: canceling async requests: count=0
2016-01-20 10:44:20.649064 7f1e0647fac0 10 librbd::ImageWatcher: 0x562788f64b40 releasing exclusive lock
2016-01-20 10:44:20.684422 7f1e0647fac0 10 librbd::ObjectMap: 0x56278901a000 unlocking object map
2016-01-20 10:44:20.686348 7f1e0647fac0 10 librbd::ImageWatcher: 0x562788f64b40 current lock owner: [0,0]
2016-01-20 10:44:20.686366 7f1e0647fac0 10 librbd::ImageWatcher: 0x562788f64b40 unregistering image watcher

#10 - 01/27/2016 01:25 AM - Tyler Wilson

Having similar issue on <http://tracker.ceph.com/issues/14522>

#11 - 01/28/2016 02:31 AM - Jason Dillaman

- Status changed from *Need More Info* to *In Progress*
- Priority changed from *Normal* to *High*

Sorry for the delay, I somehow missed the email notification for this ticket. The logs are perfect and I can see the issue now:

```
2016-01-20 10:33:48.315914 7fad84b0700 20 librbd::ImageCtx: waiting for cache to be flushed
```

librbd never gets the notification that it has safely written the cache writeback to disk. I'll start working on a fix for this ASAP.

#12 - 01/28/2016 07:42 PM - Jason Dillaman

- Status changed from *In Progress* to *Pending Backport*
- Backport set to *hammer, infernalis*

#13 - 01/28/2016 07:43 PM - Jason Dillaman

- Copied to Backport #14542: *infernalis: Cannot reliably create snapshot after freezing QEMU IO added*

#14 - 01/28/2016 07:43 PM - Jason Dillaman

- Copied to Backport #14543: *hammer: Cannot reliably create snapshot after freezing QEMU IO added*

#15 - 02/12/2016 01:37 PM - Nathan Cutler

- Status changed from *Pending Backport* to *Resolved*