

bluestore - Bug #23840

Bluestore OSD hit assert((log_reader->buf.pos & ~super.block_mask()) == 0)

04/24/2018 04:36 PM - Rafal Wadolowski

Status: Resolved	% Done: 0%
Priority: High	
Assignee:	
Category:	
Target version:	
Source:	Reviewed:
Tags:	Affected Versions: v12.2.4
Backport: luminous	ceph-qa-suite:
Regression: No	Pull request ID:
Severity: 1 - critical	Crash signature:
Description	
We're using ceph 12.2.4, where db and wal are on separate nvme.	
After restart on some OSDs we see the following error:	
<pre>-3> 2018-04-24 16:18:33.940657 7f95ee81de00 10 bluefs _replay 0xffffe000: txn(seq 1550851 len 0xa6a crc 0x8cdd70c6) -2> 2018-04-24 16:18:33.940663 7f95ee81de00 10 bluefs _read h 0x55a7876b6f00 0xfffff000~1000 from file(ino 1 size 0xfffff000 mtime 0.000000 bdev 0 extents -1> 2018-04-24 16:18:33.940884 7f95ee81de00 10 bluefs _replay 0xfffff000: txn(seq 1550852 len 0xa6a crc 0x5e1c1b4f) 0> 2018-04-24 16:18:33.943750 7f95ee81de00 -1 /build/ceph-12.2.4 /src/os/bluestore/BlueFS.cc: In function 'int BlueFS::_replay(bool)' thread 7f95ee81de00 time 2018-04-24 16:18:33.940909 /build/ceph-12.2.4/src/os/bluestore/BlueFS.cc: 551 : FAILED assert((log_reader->buf.pos & ~super.block_mask()) == 0)</pre>	
Right now we have some pgs inactive.	
Full log in attachment.	
Related issues:	
Copied to bluestore - Backport #23881: luminous: Bluestore OSD hit assert((lo...	Resolved

History

#1 - 04/24/2018 05:12 PM - Rafal Wadolowski

What is interesting, it looks like the max read of file is limited to 0xfffffff.

```
-11> 2018-04-24 16:18:33.939739 7f95ee81de00 10 bluefs _replay 0xfffffa000: txn(seq 1550847 len 0xa6a crc 0x95281bbd)
  -10> 2018-04-24 16:18:33.939745 7f95ee81de00 10 bluefs _read h 0x55a7876b6f00 0xfffffb000~1000
    from file(ino 1 size 0xfffffb000 mtime 0.000000 bdev 0 extents [1:0x15300000+100000,0:0x3d800000+c00000,
  -9> 2018-04-24 16:18:33.939970 7f95ee81de00 10 bluefs _replay 0xfffffb000: txn(seq 1550848 len 0xa6a crc 0x86579e18)
    -8> 2018-04-24 16:18:33.939976 7f95ee81de00 10 bluefs _read h 0x55a7876b6f00 0xfffffc000~1000
    from file(ino 1 size 0xfffffc000 mtime 0.000000 bdev 0 extents [1:0x15300000+100000,0:0x3d800000+c00000,
  -7> 2018-04-24 16:18:33.940196 7f95ee81de00 10 bluefs _replay 0xfffffc000: txn(seq 1550849 len 0xa6a crc 0x61eaa8e2)
    -6> 2018-04-24 16:18:33.940203 7f95ee81de00 10 bluefs _read h 0x55a7876b6f00 0xfffffd000~1000
    from file(ino 1 size 0xfffffd000 mtime 0.000000 bdev 0 extents [1:0x15300000+100000,0:0x3d800000+c00000,
  -5> 2018-04-24 16:18:33.940427 7f95ee81de00 10 bluefs _replay 0xfffffd000: txn(seq 1550850 len 0xa6a crc
```

```
0xbcc8bda3)
-4> 2018-04-24 16:18:33.940433 7f95ee81de00 10 bluefs _read h 0x55a7876b6f00 0xfffffe000~1000
from file(ino 1 size 0xfffffe000 mtime 0.000000 bdev 0 extents [1:0x15300000+100000,0:0x3d800000+c00000,
-3> 2018-04-24 16:18:33.940657 7f95ee81de00 10 bluefs _replay 0xfffffe000: txn(seq 1550851 len 0xa6a crc
0x8cdd70c6)
-2> 2018-04-24 16:18:33.940663 7f95ee81de00 10 bluefs _read h 0x55a7876b6f00 0xfffff000~1000
from file(ino 1 size 0xfffff000 mtime 0.000000 bdev 0 extents [1:0x15300000+100000,0:0x3d800000+c00000,
-1> 2018-04-24 16:18:33.940884 7f95ee81de00 10 bluefs _replay 0xfffff000: txn(seq 1550852 len 0xa6a crc
0x5e1c1b4f)
0> 2018-04-24 16:18:33.943750 7f95ee81de00 -1 /build/ceph-12.2.4/src/os/bluestore/BlueFS.cc: In function
'int BlueFS::_replay(bool)' thread 7f95ee81de00 time 2018-04-24 16:18:33.940909
/build/ceph-12.2.4/src/os/bluestore/BlueFS.cc: 551
: FAILED assert((log_reader->buf.pos & ~super.block_mask()) == 0)
```

In each line of bluefs _read, we see iterating address of log file, and it is breaking at 0xffff000~1000.

#2 - 04/24/2018 05:17 PM - Rafal Wadolowski

- File shutdown.log added

Moment of planned shutdown

#3 - 04/24/2018 10:11 PM - Sage Weil

- Status changed from New to 7

<https://github.com/ceph/ceph/pull/21629>

#4 - 04/24/2018 10:11 PM - Sage Weil

- Project changed from Ceph to bluestore

- Category deleted (OSD)

- Priority changed from Normal to High

- Backport set to luminous

#5 - 04/25/2018 10:09 AM - Rafal Wadolowski

This change is working, I think it could be merge with master.

#6 - 04/26/2018 04:40 AM - Kefu Chai

- Status changed from 7 to Pending Backport

#7 - 04/26/2018 06:28 AM - Enming Zhang

I have met the same issue in Luminous.

#8 - 04/26/2018 08:48 AM - Nathan Cutler

- Copied to Backport #23881: luminous: Bluestore OSD hit assert((log_reader->buf.pos & ~super.block_mask()) == 0) added

#9 - 05/05/2018 01:48 PM - Nathan Cutler

- Status changed from Pending Backport to Resolved

Files

osd.log	19.1 KB	04/24/2018	Rafal Wadolowski
shutdown.log	8.15 KB	04/24/2018	Rafal Wadolowski