

bluestore - Bug #46270

mimic:osd can not start

06/30/2020 06:40 AM - ☐ ☐ ☐

Status: Can't reproduce	% Done: 0%
Priority: Low	
Assignee:	
Category:	
Target version:	
Source:	Affected Versions: v13.2.5
Tags:	ceph-qa-suite:
Backport:	Pull request ID:
Regression: No	Crash signature (v1):
Severity: 3 - minor	Crash signature (v2):
Reviewed:	

Description

My env:

```
[root@mon1 test]# ceph -v
ceph version 13.2.5 (cbff874f9007f1869bfd3821b7e33b2a6ffd4988) mimic (stable)
[root@mon1 test]# cat /etc/redhat-release
CentOS Linux release 7.6.1810 (Core)
```

and my cluster has stored a lot of data:

```
pools: 8 pools, 34048 pgs
objects: 5.78 G objects, 783 TiB
usage: 1.5 PiB used, 10 PiB / 12 PiB avail
pgs: 0.162% pgs not active
7238/70303191729 objects degraded (0.000%)
1714/70303191729 objects misplaced (0.000%)
33539 active+clean
216 active+undersized
175 active+undersized+degraded
48 active+clean+scrubbing+deep+repair
39 undersized+peered
16 undersized+degraded+peered
11 active+clean+remapped
2 active+undersized+remapped
1 active+clean+remapped+scrubbing+deep
1 active+clean+scrubbing+deep
```

this cluster focuses on object storage and i put index log meta control in ssd based osds then i suddenly found two of them are broken

```
-2> 2020-06-30 14:23:39.698 7f09adeadd80 -1 bluefs _allocate failed to allocate 0x100000 on bdev 2
, dne
-1> 2020-06-30 14:23:39.698 7f09adeadd80 -1 bluefs _flush_range allocated: 0x3600000 offset: 0
x35eb2c1 length: 0xe9031
0> 2020-06-30 14:23:39.703 7f09adeadd80 -1 /home/jenkins-build/build/workspace/ceph-build/ARCH/x86_64/AVAILABLE_ARCH/x86_64/AVAILABLE_DIST/centos7/DIST/centos7/MACHINE_SIZE/huge/release/13.2.5/rpm/el7/BUILD/ceph-13.2.5/src/os/bluestore/BlueFS.cc: In function 'int BlueFS::_flush_range(BlueFS::FileWriter*, uint64_t, uint64_t)' thread 7f09adeadd80 time 2020-06-30 14:23:39.698917
/home/jenkins-build/build/workspace/ceph-build/ARCH/x86_64/AVAILABLE_ARCH/x86_64/AVAILABLE_DIST/centos7/DIST/centos7/MACHINE_SIZE/huge/release/13.2.5/rpm/el7/BUILD/ceph-13.2.5/src/os/bluestore/BlueFS.cc: 1687: FAILED assert(0 == "bluefs enospc")
```

```
ceph version 13.2.5 (cbff874f9007f1869bfd3821b7e33b2a6ffd4988) mimic (stable)
```

```
1: (ceph::__ceph_assert_fail(char const*, char const*, int, char const*)+0xff) [0x7f09a52b7fbf]
2: (()+0x26d187) [0x7f09a52b8187]
3: (BlueFS::_flush_range(BlueFS::FileWriter*, unsigned long, unsigned long)+0x1ac6) [0x562e69bd7496]
4: (BlueRocksWritableFile::Flush()+0x3d) [0x562e69bed15d]
5: (rocksdb::WritableFileWriter::Flush()+0x196) [0x562e69dce676]
6: (rocksdb::WritableFileWriter::Append(rocksdb::Slice const&)+0x15b) [0x562e69dcebbb]
7: (rocksdb::BlockBasedTableBuilder::WriteRawBlock(rocksdb::Slice const&, rocksdb::CompressionType, rocksdb::BlockHandle*, bool)+0x1f6) [0x562e69e48f46]
8: (rocksdb::BlockBasedTableBuilder::WriteBlock(rocksdb::Slice const&, rocksdb::BlockHandle*, bool)+0xdb) [0x562e69e493eb]
9: (rocksdb::BlockBasedTableBuilder::Finish()+0x920) [0x562e69e4cbc0]
10: (rocksdb::BuildTable(std::string const&, rocksdb::Env*, rocksdb::ImmutableCFOptions const&, rocksdb::MutableCFOptions const&, rocksdb::EnvOptions const&, rocksdb::TableCache*, rocksdb::InternalIterator*, std::unique_ptr<rocksdb::InternalIterator, std::default_delete<rocksdb::InternalIterator>&gt;, rocksdb::FileMetaData*, rocksdb::InternalKeyComparator const&, std::vector<std::unique_ptr<rocksdb::IntTblPropCollectorFactory, std::default_delete<rocksdb::IntTblPropCollectorFactory>&gt;&gt;, std::allocator<std::unique_ptr<rocksdb::IntTblPropCollectorFactory, std::default_delete<rocksdb::IntTblPropCollectorFactory>&gt;&gt;, std::string const&, std::vector<unsigned long, std::allocator<unsigned long>&gt;&gt;, unsigned long, rocksdb::SnapshotChecker*, rocksdb::CompressionType, rocksdb::CompressionOptions const&, bool, rocksdb::InternalStats*, rocksdb::TableFileCreationReason, rocksdb::EventLogger*, int, rocksdb::Env::IOPriority, rocksdb::TableProperties*, int, unsigned long, unsigned long, rocksdb::Env::WriteLifeTimeHint)+0x1a2d) [0x562e69df4b5d]
11: (rocksdb::DBImpl::WriteLevel0TableForRecovery(int, rocksdb::ColumnFamilyData*, rocksdb::MemTable*, rocksdb::VersionEdit*)+0xbe6) [0x562e69c87ad6]
12: (rocksdb::DBImpl::RecoverLogFiles(std::vector<unsigned long, std::allocator<unsigned long>&gt;& const&, unsigned long*, bool)+0x10f1) [0x562e69c893c1]
13: (rocksdb::DBImpl::Recover(std::vector<rocksdb::ColumnFamilyDescriptor, std::allocator<rocksdb::ColumnFamilyDescriptor>&gt;& const&, bool, bool, bool)+0xa59) [0x562e69c8aa69]
14: (rocksdb::DBImpl::Open(rocksdb::DBOptions const&, std::string const&, std::vector<rocksdb::ColumnFamilyDescriptor, std::allocator<rocksdb::ColumnFamilyDescriptor>&gt;& const&, std::vector<rocksdb::ColumnFamilyHandle*, std::allocator<rocksdb::ColumnFamilyHandle>&gt;& const&, >, rocksdb::DB*, bool)+0x689) [0x562e69c8b819]
15: (rocksdb::DB::Open(rocksdb::DBOptions const&, std::string const&, std::vector<rocksdb::ColumnFamilyDescriptor, std::allocator<rocksdb::ColumnFamilyDescriptor>&gt;& const&, std::vector<rocksdb::ColumnFamilyHandle*, std::allocator<rocksdb::ColumnFamilyHandle>&gt;& const&, >, rocksdb::DB*)+0x22) [0x562e69c8d042]
16: (RocksDBStore::do_open(std::ostream&, bool, std::vector<KeyValueDB::ColumnFamily, std::allocator<KeyValueDB::ColumnFamily>&gt;& const&)+0x164e) [0x562e69b6eb3e]
17: (BlueStore::_open_db(bool, bool)+0xd6a) [0x562e69afa8fa]
18: (BlueStore::_mount(bool, bool)+0x4d1) [0x562e69b2b6b1]
19: (OSD::init()+0x28f) [0x562e696d508f]
20: (main()+0x23a3) [0x562e695b3363]
21: (__libc_start_main()+0xf5) [0x7f09a0e3b3d5]
22: (()+0x384ab0) [0x562e6968bab0]
NOTE: a copy of the executable, or `objdump -rdS &lt;executable>` is needed to interpret this.
```

History

#1 - 06/30/2020 04:48 PM - Neha Ojha

- Project changed from RADOS to bluestore
- Priority changed from Normal to Low

This just looks like bluefs is running out of space. Mimic is EOL, I'd recommend you to upgrade and report back if you see similar issues.

#2 - 06/08/2021 12:07 PM - Igor Fedotov

- Status changed from New to Can't reproduce