

## bluestore - Bug #36625

### \_aio\_log\_start inflight overlap of 0x10000~1000 with [65536~4096]

10/30/2018 06:55 AM - Honggang Yang

<b>Status:</b>	Resolved	<b>Start date:</b>	10/30/2018
<b>Priority:</b>	High	<b>Due date:</b>	
<b>Assignee:</b>		<b>% Done:</b>	0%
<b>Category:</b>		<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>		<b>Reviewed:</b>	
<b>Source:</b>	Community (user)	<b>Affected Versions:</b>	v14.0.0
<b>Tags:</b>		<b>ceph-qa-suite:</b>	
<b>Backport:</b>	mimic,luminous	<b>Pull request ID:</b>	
<b>Regression:</b>	No		
<b>Severity:</b>	1 - critical		

#### Description

### discription

```
-11> 2018-10-30 04:26:56.898 7f5e0e8cf680 20 bluestore(bluestore.test_temp_dir) _fsck referenc
ed 0x1 for Blob(0x55c1b71c0cb0 blob([0x10000~10000] csum+has_unused crc32c/0x1000 unused=0xffff) u
se_tracker(0x10000 0x1e) SharedBlob(0x55c1b71c0e00 sbid 0x0))
-10> 2018-10-30 04:26:56.898 7f5e0e8cf680 20 bluestore(bluestore.test_temp_dir) _do_read 0x0~28
size 0x28 (40)
-9> 2018-10-30 04:26:56.898 7f5e0e8cf680 20 bluestore(bluestore.test_temp_dir) _do_read defaul
ting to buffered read
-8> 2018-10-30 04:26:56.898 7f5e0e8cf680 20 bluestore(bluestore.test_temp_dir) _do_read blob
Blob(0x55c1b71c0cb0 blob([0x10000~10000] csum+has_unused crc32c/0x1000 unused=0xffff) use_tracker(
0x10000 0x1e) SharedBlob(0x55c1b71c0e00 sbid 0x0)) need 0x0~5 cache has 0x[]
-7> 2018-10-30 04:26:56.898 7f5e0e8cf680 20 bluestore(bluestore.test_temp_dir) _do_read blob
Blob(0x55c1b71c0cb0 blob([0x10000~10000] csum+has_unused crc32c/0x1000 unused=0xffff) use_tracker(
0x10000 0x1e) SharedBlob(0x55c1b71c0e00 sbid 0x0)) need 0xa~5 cache has 0x[]
-6> 2018-10-30 04:26:56.898 7f5e0e8cf680 20 bluestore(bluestore.test_temp_dir) _do_read blob
Blob(0x55c1b71c0cb0 blob([0x10000~10000] csum+has_unused crc32c/0x1000 unused=0xffff) use_tracker(
0x10000 0x1e) SharedBlob(0x55c1b71c0e00 sbid 0x0)) need 0x14~14 cache has 0x[]
-5> 2018-10-30 04:26:56.898 7f5e0e8cf680 20 bluestore(bluestore.test_temp_dir) _do_read blob
Blob(0x55c1b71c0cb0 blob([0x10000~10000] csum+has_unused crc32c/0x1000 unused=0xffff) use_tracker(
0x10000 0x1e) SharedBlob(0x55c1b71c0e00 sbid 0x0)) need 0x0:0~5,0xa:a~5,0x14:14~14
-4> 2018-10-30 04:26:56.898 7f5e0e8cf680 20 bluestore(bluestore.test_temp_dir) _do_read reg
ion 0x0: 0x0~5 reading 0x0~1000
-3> 2018-10-30 04:26:56.898 7f5e0e8cf680 20 bluestore(bluestore.test_temp_dir) _do_read reg
ion 0xa: 0xa~5 reading 0x0~1000
-2> 2018-10-30 04:26:56.898 7f5e0e8cf680 -1 bdev(0x55c1b7f50000 bluestore.test_temp_dir/block)
_aio_log_start inflight overlap of 0x10000~1000 with [65536~4096]
-1> 2018-10-30 04:26:56.914 7f5e0e8cf680 -1 /home/yhg/work/ceph/src/os/bluestore/KernelDevice.
cc: In function 'void KernelDevice::_aio_log_start(IOContext*, uint64_t, uint64_t)' thread 7f5e0e8
cf680 time 2018-10-30 04:26:56.900490
/home/yhg/work/ceph/src/os/bluestore/KernelDevice.cc: 576: abort()

ceph version 14.0.0-4631-g039e29b (039e29b5ddc8a50fd9a623b267bfc6e326d9de9c) nautilus (dev)
1: (ceph::__ceph_abort(char const*, int, char const*, std::string const&)+0xfe) [0x7f5e039462b2]
2: (KernelDevice::_aio_log_start(IOContext*, unsigned long, unsigned long)+0x417) [0x55c1b4dd6a57
]
3: (KernelDevice::aio_read(unsigned long, unsigned long, ceph::buffer::list*, IOContext*)+0x26a)
[0x55c1b4dda2d8]
4: ((()+0xaac146) [0x55c1b4c04146]
5: ((()+0xaeaf17) [0x55c1b4c42f17]
6: (BlueStore::_do_read(BlueStore::Collection*, boost::intrusive_ptr<BlueStore::Onode>, unsigned
long, unsigned long, ceph::buffer::list&, unsigned int, unsigned long)+0x1bf3) [0x55c1b4c05ead]
7: (BlueStore::_fsck(bool, bool)+0x488f) [0x55c1b4bf93a9]
```

```

8: (BlueStore::fsck(bool)+0x28) [0x55c1b4c5623a]
9: (BlueStore::umount()+0x472) [0x55c1b4bf2f6c]
10: (StoreTestFixture::TearDown()+0x71) [0x55c1b4aabf0f]
11: (void testing::internal::HandleSehExceptionsInMethodIfSupported<testing::Test, void>(testing:
:Test*, void (testing::Test::* )(), char const*)+0x65) [0x55c1b4e1bf67]
12: (void testing::internal::HandleExceptionsInMethodIfSupported<testing::Test, void>(testing::Te
st*, void (testing::Test::* )(), char const*)+0x5a) [0x55c1b4e16b6f]
13: (testing::Test::Run()+0x11b) [0x55c1b4dfc571]
14: (testing::TestInfo::Run()+0x104) [0x55c1b4dfcdfe]
15: (testing::TestCase::Run()+0x107) [0x55c1b4dfd497]
16: (testing::internal::UnitTestImpl::RunAllTests()+0x2d2) [0x55c1b4e03f7c]
17: (bool testing::internal::HandleSehExceptionsInMethodIfSupported<testing::internal::UnitTestIm
pl, bool>(testing::internal::UnitTestImpl*, bool (testing::internal::UnitTestImpl::* )(), char cons
t*)+0x65) [0x55c1b4e1d077]
18: (bool testing::internal::HandleExceptionsInMethodIfSupported<testing::internal::UnitTestImpl,
bool>(testing::internal::UnitTestImpl*, bool (testing::internal::UnitTestImpl::* )(), char const*)
+0x5a) [0x55c1b4e178c3]
19: (testing::UnitTest::Run()+0xc9) [0x55c1b4e02b73]
20: (RUN_ALL_TESTS()+0x11) [0x55c1b4a4798a]
21: (main()+0xcc8) [0x55c1b4a45b3f]
22: (__libc_start_main()+0xf5) [0x7f5e00364c05]
23: (()+0x864a25) [0x55c1b49bca25]

```

## howto reproduce

enable the following three options.

```

# git diff test
diff --git a/src/test/objectstore/store_test.cc b/src/test/objectstore/store_test.cc
index 8366469..88cd04c 100644
--- a/src/test/objectstore/store_test.cc
+++ b/src/test/objectstore/store_test.cc
@@ -7359,13 +7359,14 @@ int main(int argc, char **argv) {
    g_ceph_context->_conf.set_val_or_die("filestore_op_thread_suicide_timeout", "10000");
    //g_ceph_context->_conf.set_val_or_die("filestore_fiemap", "true");
    g_ceph_context->_conf.set_val_or_die("bluestore_fsck_on_mkfs", "false");
-   g_ceph_context->_conf.set_val_or_die("bluestore_fsck_on_mount", "false");
-   g_ceph_context->_conf.set_val_or_die("bluestore_fsck_on_umount", "false");
+   g_ceph_context->_conf.set_val_or_die("bluestore_fsck_on_mount", "true");
+   g_ceph_context->_conf.set_val_or_die("bluestore_fsck_on_umount", "true");
    g_ceph_context->_conf.set_val_or_die("bluestore_debug_misc", "true");
    g_ceph_context->_conf.set_val_or_die("bluestore_debug_small_allocations", "4");
    g_ceph_context->_conf.set_val_or_die("bluestore_debug_freelist", "true");
    g_ceph_context->_conf.set_val_or_die("bluestore_clone_cow", "true");
    g_ceph_context->_conf.set_val_or_die("bluestore_max_alloc_size", "196608");
+   g_ceph_context->_conf.set_val_or_die("bdev_debug_inflight_ios", "true");

    // set small cache sizes so we see trimming during Synthetic tests
    g_ceph_context->_conf.set_val_or_die("bluestore_cache_size_hdd", "400000");

```

run test:

```

ceph_test_objectstore --gtest_filter=ObjectStore/StoreTest.BufferCacheReadTest/2 --debug_bluestore
=22 --debug_osd=10 bluestore

```

**Related issues:**

Copied to bluestore - Backport #36754: mimic: \_aio\_log\_start inflight overlap...  
Copied to bluestore - Backport #36755: luminous: \_aio\_log\_start inflight over...

**Resolved**  
**Rejected**

## History

---

### #1 - 10/30/2018 07:00 AM - Honggang Yang

<https://github.com/ceph/ceph/pull/24820>

### #2 - 10/30/2018 02:41 PM - Sage Weil

- Status changed from New to In Progress
- Priority changed from Normal to High
- Backport set to mimic,luminous

### #3 - 11/08/2018 10:29 PM - Sage Weil

- Status changed from In Progress to Pending Backport

### #4 - 11/10/2018 08:54 AM - Nathan Cutler

- Copied to Backport #36754: mimic: \_aio\_log\_start inflight overlap of 0x10000~1000 with [65536~4096] added

### #5 - 11/10/2018 08:54 AM - Nathan Cutler

- Copied to Backport #36755: luminous: \_aio\_log\_start inflight overlap of 0x10000~1000 with [65536~4096] added

### #6 - 01/14/2019 12:41 PM - Nathan Cutler

- Status changed from Pending Backport to Resolved