# RADOS - Bug #37507

## osd_memory_target: failed assert when options mismatch

12/03/2018 04:02 PM - Dan van der Ster

| | | | |
|---|---|---|---|
| **Status:** | Resolved | **% Done:** | 0% |
| **Priority:** | Normal | **Spent time:** | 0.00 hour |
| **Assignee:** | Mark Nelson | | |
| **Category:** | | | |
| **Target version:** | | | |
| **Source:** | Community (user) | **Affected Versions:** | v12.2.10 |
| **Tags:** | | **ceph-qa-suite:** | |
| **Backport:** | luminous,mimic | **Component(RADOS):** | OSD |
| **Regression:** | No | **Pull request ID:** | |
| **Severity:** | 3 - minor | **Crash signature (v1):** | |
| **Reviewed:** | | **Crash signature (v2):** | |

**Description**

We tried setting osd_memory_target to 1GB and this results in the following assertion early after startup:

```
     0> 2018-12-03 11:55:31.049163 7f46c91cc700 -1 /home/jenkins-build/build/workspace/ceph-build/
ARCH/x86_64/AVAILABLE_ARCH/x86_64/AVAILABLE_DIST/centos7/DIST/centos7/MACHINE_SIZE/huge/release/12
.2.10/rpm/el7/BUILD/ceph-12.2.10/src/os/bluestore/BlueStore.cc: In function 'void BlueStore::Mempo
olThread::_balance_cache(const std::list<PriorityCache::PriCache*>&)' thread 7f46c91cc700 time 201
8-12-03 11:55:31.047247
/home/jenkins-build/build/workspace/ceph-build/ARCH/x86_64/AVAILABLE_ARCH/x86_64/AVAILABLE_DIST/ce
ntos7/DIST/centos7/MACHINE_SIZE/huge/release/12.2.10/rpm/el7/BUILD/ceph-12.2.10/src/os/bluestore/B
lueStore.cc: 3488: FAILED assert(mem_avail >= 0)

 ceph version 12.2.10 (177915764b752804194937482a39e95e0ca3de94) luminous (stable)
 1: (ceph::__ceph_assert_fail(char const*, char const*, int, char const*)+0x110) [0x5597ce4fb7b0]
 2: (()+0x8fc754) [0x5597ce357754]
 3: (BlueStore::MempoolThread::entry()+0x332) [0x5597ce35bc72]
 4: (()+0x7e25) [0x7f46d3a81e25]
 5: (clone()+0x6d) [0x7f46d2b72bad]
 NOTE: a copy of the executable, or `objdump -rdS <executable>` is needed to interpret this.
```

Perhaps we should more gracefully enforce a min practical value? Also the docs don't seem to mention any min usable value?
http://docs.ceph.com/docs/master/rados/configuration/bluestore-config-ref/?highlight=osd%20memory%20target#automatic-cache-sizing

**Related issues:**

| | |
|---|---|
| Copied to RADOS - Backport #37697: luminous: osd_memory_target: failed assert... | **Resolved** |
| Copied to RADOS - Backport #37698: mimic: osd_memory_target: failed assert wh... | **Resolved** |

### History

#### #1 - 12/08/2018 01:44 PM - Konstantin Shalygin

```
Dec 08 19:32:51 ceph-osd0 ceph-osd[171040]: starting osd.0 at - osd_data /var/lib/ceph/osd/ceph-0 /var/lib/cep
h/osd/ceph-0/journal
Dec 08 19:32:58 ceph-osd0 ceph-osd[171040]: 2018-12-08 19:32:58.409947 7f1391a7ed80 -1 osd.0 53990 log_to_moni
tors {default=true}
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: /home/jenkins-build/build/workspace/ceph-build/ARCH/x86_64/AVAILAB
LE_ARCH/x86_64/AVAILABLE_DIST/centos7/DIST/centos7/MACHINE_SIZE/huge/release/12.2.10/rpm/el7/BUILD/ceph-12.2.1
0/src/os/bluestore/BlueStore.cc: In function 'void BlueStore::MempoolThread::_balance_cache(const std::list<Pr
iorityCache::PriCache*>&)' thread 7f1382673700 time 2018-12-08 19:33:00.001144
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: /home/jenkins-build/build/workspace/ceph-build/ARCH/x86_64/AVAILAB
```

```
LE_ARCH/x86_64/AVAILABLE_DIST/centos7/DIST/centos7/MACHINE_SIZE/huge/release/12.2.10/rpm/el7/BUILD/ceph-12.2.1
0/src/os/bluestore/BlueStore.cc: 3488: FAILED assert(mem_avail >= 0)
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: ceph version 12.2.10 (177915764b752804194937482a39e95e0ca3de94) lu
minous (stable)
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 1: (ceph::__ceph_assert_fail(char const*, char const*, int, char c
onst*)+0x110) [0x561dd4c2d7b0]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 2: (()+0x8fc754) [0x561dd4a89754]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 3: (BlueStore::MempoolThread::entry()+0x332) [0x561dd4a8dc72]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 4: (()+0x7e25) [0x7f138ef2ce25]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 5: (clone()+0x6d) [0x7f138e01dbad]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: NOTE: a copy of the executable, or `objdump -rdS <executable>` is
needed to interpret this.
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 2018-12-08 19:33:00.003444 7f1382673700 -1 /home/jenkins-build/bui
ld/workspace/ceph-build/ARCH/x86_64/AVAILABLE_ARCH/x86_64/AVAILABLE_DIST/centos7/DIST/centos7/MACHINE_SIZE/hug
e/release/12.2.10/rpm/el7/BUILD/ceph-12.2.10/src/os/bluestore/BlueStore.cc: In function 'void BlueStore::Mempo
olThread::_balance_cache(const std::list<PriorityCache::PriCache*>&)' thread 7f1382673700 time 2018-12-08 19:3
3:00.001144
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: /home/jenkins-build/build/workspace/ceph-build/ARCH/x86_64/AVAILAB
LE_ARCH/x86_64/AVAILABLE_DIST/centos7/DIST/centos7/MACHINE_SIZE/huge/release/12.2.10/rpm/el7/BUILD/ceph-12.2.1
0/src/os/bluestore/BlueStore.cc: 3488: FAILED assert(mem_avail >= 0)
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: ceph version 12.2.10 (177915764b752804194937482a39e95e0ca3de94) lu
minous (stable)
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 1: (ceph::__ceph_assert_fail(char const*, char const*, int, char c
onst*)+0x110) [0x561dd4c2d7b0]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 2: (()+0x8fc754) [0x561dd4a89754]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 3: (BlueStore::MempoolThread::entry()+0x332) [0x561dd4a8dc72]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 4: (()+0x7e25) [0x7f138ef2ce25]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 5: (clone()+0x6d) [0x7f138e01dbad]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: NOTE: a copy of the executable, or `objdump -rdS <executable>` is
needed to interpret this.
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: -2836> 2018-12-08 19:32:58.409947 7f1391a7ed80 -1 osd.0 53990 log_
to_monitors {default=true}
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 0> 2018-12-08 19:33:00.003444 7f1382673700 -1 /home/jenkins-build/
build/workspace/ceph-build/ARCH/x86_64/AVAILABLE_ARCH/x86_64/AVAILABLE_DIST/centos7/DIST/centos7/MACHINE_SIZE/
huge/release/12.2.10/rpm/el7/BUILD/ceph-12.2.10/src/os/bluestore/BlueStore.cc: In function 'void BlueStore::Me
mpoolThread::_balance_cache(const std::list<PriorityCache::PriCache*>&)' thread 7f1382673700 time 2018-12-08 1
9:33:00.001144
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: /home/jenkins-build/build/workspace/ceph-build/ARCH/x86_64/AVAILAB
LE_ARCH/x86_64/AVAILABLE_DIST/centos7/DIST/centos7/MACHINE_SIZE/huge/release/12.2.10/rpm/el7/BUILD/ceph-12.2.1
0/src/os/bluestore/BlueStore.cc: 3488: FAILED assert(mem_avail >= 0)
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: ceph version 12.2.10 (177915764b752804194937482a39e95e0ca3de94) lu
minous (stable)
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 1: (ceph::__ceph_assert_fail(char const*, char const*, int, char c
onst*)+0x110) [0x561dd4c2d7b0]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 2: (()+0x8fc754) [0x561dd4a89754]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 3: (BlueStore::MempoolThread::entry()+0x332) [0x561dd4a8dc72]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 4: (()+0x7e25) [0x7f138ef2ce25]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 5: (clone()+0x6d) [0x7f138e01dbad]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: NOTE: a copy of the executable, or `objdump -rdS <executable>` is
needed to interpret this.
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: *** Caught signal (Aborted) **
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: in thread 7f1382673700 thread_name:bstore_mempool
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: ceph version 12.2.10 (177915764b752804194937482a39e95e0ca3de94) lu
minous (stable)
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 1: (()+0xa618e1) [0x561dd4bee8e1]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 2: (()+0xf6d0) [0x7f138ef346d0]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 3: (gsignal()+0x37) [0x7f138df55277]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 4: (abort()+0x148) [0x7f138df56968]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 5: (ceph::__ceph_assert_fail(char const*, char const*, int, char c
onst*)+0x284) [0x561dd4c2d924]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 6: (()+0x8fc754) [0x561dd4a89754]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 7: (BlueStore::MempoolThread::entry()+0x332) [0x561dd4a8dc72]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 8: (()+0x7e25) [0x7f138ef2ce25]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 9: (clone()+0x6d) [0x7f138e01dbad]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 2018-12-08 19:33:00.018742 7f1382673700 -1 *** Caught signal (Abor
ted) **
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: in thread 7f1382673700 thread_name:bstore_mempool
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: ceph version 12.2.10 (177915764b752804194937482a39e95e0ca3de94) lu
minous (stable)
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 1: (()+0xa618e1) [0x561dd4bee8e1]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 2: (()+0xf6d0) [0x7f138ef346d0]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 3: (gsignal()+0x37) [0x7f138df55277]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 4: (abort()+0x148) [0x7f138df56968]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 5: (ceph::__ceph_assert_fail(char const*, char const*, int, char c
onst*)+0x284) [0x561dd4c2d924]
```

```
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 6: (()+0x8fc754) [0x561dd4a89754]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 7: (BlueStore::MempoolThread::entry()+0x332) [0x561dd4a8dc72]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 8: (()+0x7e25) [0x7f138ef2ce25]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 9: (clone()+0x6d) [0x7f138e01dbad]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: NOTE: a copy of the executable, or `objdump -rdS <executable>` is
needed to interpret this.
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 0> 2018-12-08 19:33:00.018742 7f1382673700 -1 *** Caught signal (A
borted) **
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: in thread 7f1382673700 thread_name:bstore_mempool
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: ceph version 12.2.10 (177915764b752804194937482a39e95e0ca3de94) lu
minous (stable)
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 1: (()+0xa618e1) [0x561dd4bee8e1]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 2: (()+0xf6d0) [0x7f138ef346d0]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 3: (gsignal()+0x37) [0x7f138df55277]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 4: (abort()+0x148) [0x7f138df56968]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 5: (ceph::__ceph_assert_fail(char const*, char const*, int, char c
onst*)+0x284) [0x561dd4c2d924]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 6: (()+0x8fc754) [0x561dd4a89754]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 7: (BlueStore::MempoolThread::entry()+0x332) [0x561dd4a8dc72]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 8: (()+0x7e25) [0x7f138ef2ce25]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: 9: (clone()+0x6d) [0x7f138e01dbad]
Dec 08 19:33:00 ceph-osd0 ceph-osd[171040]: NOTE: a copy of the executable, or `objdump -rdS <executable>` is
needed to interpret this.
```

Same. It would be nice if such changes were laid out at least approximate expectations of the *new* default settings. My osd hosts used 50-53% RAM when Bluestore cache size is default (1GB). Now I had to put 2GB for osd_memory_target  in order to start and observe what the memory consumption will be.

#### #2 - 12/10/2018 10:37 PM - Greg Farnum

*- Assignee set to Mark Nelson*

Thoughts, Mark?

#### #3 - 12/10/2018 11:18 PM - Mark Nelson

Hi Folks,

I'm guessing this is related to https://github.com/ceph/ceph/pull/25421  Basically a stupid uint64_t bug on my part where the target size wraps to be huge when the osd_memory_target is set too low.  My goal was not to restrict the minimum size since ideally the autotuner will just set the bluestore cache size to osd_memory_cache_min and leave it at that.

Can you verify that you don't see the crash if you set the osd_memory_target to something a little larger? IE make sure this expression is true:

(((1.0 - osd_memory_expected_fragmentation) * osd_memory_target) - osd_memory_base) > osd_memory_cache_min

ie for default settings:

osd_memory_target = (134217728 + 805306368) / 0.85 = 1105322466

If that fixes it then it's most likely the same issue as 25421.

Edit: I verified that setting osd_memory_target to 1105322466 avoids the assert on one of our test boxes. Alternatively if you tweak the osd_memory_base and/or the osd_memory_expected_fragmentation you can avoid the assert. We'll issue a fix to master tomorrow.

Mark

#### #4 - 12/11/2018 03:08 PM - Dan van der Ster

Hi Mark,

You got it: 1105322466 boots, and 1105322465 crashes with the above trace.

Cheers, Dan

#### #5 - 12/14/2018 11:36 PM - Neha Ojha

*- Backport set to luminous,mimic*

#### #6 - 12/14/2018 11:49 PM - Greg Farnum

*- Subject changed from osd_memory_target: enforce or at least document a min usable value to osd_memory_target: failed assert when options mismatch*

#### #7 - 12/17/2018 02:44 PM - Yuri Weinstein

merged https://github.com/ceph/ceph/pull/25421

#### #8 - 12/18/2018 12:39 AM - xie xingguo

*- Status changed from New to Pending Backport*

#### #9 - 12/18/2018 11:10 AM - Nathan Cutler

*- Copied to Backport #37697: luminous: osd_memory_target: failed assert when options mismatch added*

#### #10 - 12/18/2018 11:10 AM - Nathan Cutler

*- Copied to Backport #37698: mimic: osd_memory_target: failed assert when options mismatch added*

#### #11 - 01/28/2019 02:55 PM - Nathan Cutler

*- Status changed from Pending Backport to 15*

**#12 - 01/30/2019 01:00 PM - Nathan Cutler**

*- Status changed from 15 to Resolved*