

Ceph - Bug #11484

OPT_INT option interprets "3221225472" as -1073741824, and crashes in Throttle::Throttle()

04/27/2015 11:48 PM - elder one

Status: Resolved	% Done: 60%
Priority: High	Spent time: 0.00 hour
Assignee: Kefu Chai	
Category:	
Target version:	
Source: Community (user)	Affected Versions:
Tags:	ceph-qa-suite:
Backport: hammer	Pull request ID:
Regression: No	Crash signature (v1):
Severity: 3 - minor	Crash signature (v2):
Reviewed:	
Description	
Upgraded one of my cluster nodes to Hammer 0.94.1. After upgrade OSD-s on that node fail to start with error on log: Throttle.cc: 38: FAILED assert(m >= 0)	
See osd log for details.	
Ubuntu 14.04, 3.18.12 kernel	
Related issues:	
Related to Ceph - Feature #11525: improve OPTION macro to add a check functio...	New 05/04/2015
Copied to Ceph - Backport #11738: OPT_INT option interprets "3221225472" as -...	Resolved 04/27/2015

Associated revisions

Revision d62f80dc - 04/29/2015 12:36 PM - Kefu Chai

common/config: detect overflow of int values

- #include "strtol.h" in strtol.cc, to ensure the function defintions are consistent.
- add a test accordingly
- fix the testcase of converting 1024E.
- do not accept integers overflow after adding SI suffix
- do not accept integers underflow (i.e. negative values)

Fixes: #11484

Signed-off-by: Kefu Chai <kchai@redhat.com>

Revision 9b947fa3 - 06/07/2015 09:02 AM - Kefu Chai

common/config: detect overflow of int values

- #include "strtol.h" in strtol.cc, to ensure the function defintions are consistent.
- add a test accordingly
- fix the testcase of converting 1024E.

- do not accept integers overflow after adding SI suffix
- do not accept integers underflow (i.e. negative values)

Fixes: #11484

Signed-off-by: Kefu Chai <kchai@redhat.com>

(cherry picked from commit d62f80dc7b25d312ff05b65b7be854aae15b66a8)

History

#1 - 04/28/2015 02:48 PM - Kefu Chai

```
common/Throttle.cc: 38: FAILED assert(m >= 0)
```

```
ceph version 0.94.1 (e4bfad3a3c51054df7e537a724c8d0bf9be972ff)
1: (ceph::__ceph_assert_fail(char const*, char const*, int, char const*)+0x8b) [0xbc271b]
2: (Throttle::Throttle(CephContext*, std::string, long, bool)+0x290) [0xba7670]
3: (FileJournal::FileJournal(uuid_d, Finisher*, Cond*, char const*, bool, bool, bool)+0x6ad) [0x9312dd]
4: (FileStore::open_journal()+0xd9) [0x90a3c9]
5: (FileStore::mount()+0x2824) [0x90f9f4]
6: (OSD::init()+0x259) [0x6c47b9]
7: (main()+0x2860) [0x651fc0]
8: (__libc_start_main()+0xf5) [0x7f5d977c4ec5]
9: /usr/bin/ceph-osd() [0x66aff7]
NOTE: a copy of the executable, or `objdump -rds <executable>` is needed to interpret this.
```

looks like the "journal_queue_max_ops" and/or "journal_queue_max_bytes" is negative.

elder one, what are the values of these settings in your ceph cluster? or could you post the output of following commands:

```
ceph-conf --lookup journal_queue_max_bytes
ceph-conf --lookup journal_queue_max_ops
```

#2 - 04/28/2015 04:37 PM - elder one

Here it is:

```
ceph-conf -s osd --lookup journal_queue_max_bytes
3G
ceph-conf -s osd --lookup journal_queue_max_ops
5000
```

Is it a "G" suffix causing troubles?

#3 - 04/28/2015 04:46 PM - elder one

Yeah, seems to be so:

```
ceph --admin-daemon /var/run/ceph/ceph-osd.23.asok config show | grep journal_queue_max_bytes
"journal_queue_max_bytes": "-1073741824"
```

#4 - 04/28/2015 05:27 PM - elder one

That's strange - changed journal_queue_max_bytes from "3G" to 3221225472 in ceph.conf and restarted all osd's on host but admin socket gives the same negative output:

```
ceph-conf -s osd --lookup journal_queue_max_bytes
3221225472
ceph --admin-daemon /var/run/ceph/ceph-osd.23.asok config show | grep journal_queue_max_bytes
"journal_queue_max_bytes": "-1073741824",
```

#5 - 04/29/2015 06:08 AM - Kefu Chai

- Status changed from New to 12
- Assignee set to Kefu Chai
- Source changed from other to Community (user)

elder one, yes, this is a bug, because the 3221225472 (i.e. 0xc0000000) is out of range of `numeric_limits<int>::max()` (i.e. 0x7fffffff on LP64 arch⁰), and `journal_queue_max_bytes` is a config whose type is `int`. so your config value of 3221225472 cannot be represented by `int`.

so we need to:

- fix the config parsing for `int`
- probably change the type of `journal_queue_max_bytes` from `int` to unsigned long long (if we have a good reason to do so)

is it a "G" suffix causing troubles?

not exactly, you can still use the "G" suffix. but 3G exceeds the limit of `int`. so **2G-1** should be fine.

mean while, you could use "2147483647" instead of 3G for this config as a work around. just out of curiosity, is there any specific reason you need a 3G `journal_queue_max_bytes` for your system?

[0] most modern unix platforms on amd64 are LP64. see <http://www.unix.org/whitepapers/64bit.html>. and yes, i believe your linux is not an exception.

#6 - 04/29/2015 06:12 AM - Kefu Chai

- Subject changed from *OSD-s die after upgrade from Giant to Hammer to OPT_INT option interprets "3221225472" as -1073741824, and crashes in Throttle::Throttle()*

#7 - 04/29/2015 08:38 AM - Kefu Chai

- Status changed from *12* to *In Progress*

- % Done changed from *0* to *40*

#8 - 04/29/2015 12:35 PM - Kefu Chai

- Status changed from *In Progress* to *Fix Under Review*

- % Done changed from *40* to *60*

<https://github.com/ceph/ceph/pull/4494>

#9 - 04/29/2015 01:37 PM - elder one

You are right, apparently I don't need 3G journal queue. Value left info config after benchmarking/tuning cluster. That 2G limit I figured out yesterday and now my cluster runs happily on Hammer:)

On Giant, OSD admin socket gives also negative output, but osd-s start fine.

I propose if any value in config file is out of range, Ceph could use default value for that and warn about that on startup.

#10 - 04/29/2015 02:54 PM - Kefu Chai

elder one wrote:

I propose if any value in config file is out of range, Ceph could use default value for that and warn about that on startup.

thank you, elder one, =)

Yeah, i had the same feeling when looking at this issue. will step back and see what we can do to make the error handling of bad config value more user friendly. maybe we should shutdown gracefully at seeing a bad config value when starting up, but reject it at runtime. will file another ticket when i have a better idea.

#11 - 05/04/2015 03:38 AM - Kefu Chai

elder one, i created [#11525](#) to track the feature request for improving the invalid option handling.

#12 - 05/12/2015 07:09 AM - Kefu Chai

- Status changed from *Fix Under Review* to *Resolved*

- Backport set to *hammer*

- Regression set to *No*

#13 - 05/12/2015 07:09 AM - Kefu Chai

- Status changed from Resolved to Pending Backport

#14 - 06/07/2015 09:05 AM - Abhishek Lekshmanan

- Description updated

#15 - 06/07/2015 11:22 AM - Loïc Dachary

@abhishek did you intend to edit the summary to include a link to the master pull request <https://github.com/ceph/ceph/pull/4494> ? You added <https://github.com/ceph/ceph/pull/4889> which is the hammer backport matching the issue <http://tracker.ceph.com/issues/11738>.

#16 - 06/07/2015 04:31 PM - Abhishek Lekshmanan

@loic sorry it was unintended.. Nathan informed me of the mistake I had done; basically if we use the script and feed it some other tracker id; it goes and updates the description to that instead..

#17 - 06/07/2015 04:31 PM - Abhishek Lekshmanan

- Description updated

#18 - 09/02/2015 10:42 AM - Loïc Dachary

- Status changed from Pending Backport to Resolved

Files

ceph-osd.23.log	595 KB	04/27/2015	elder one
-----------------	--------	------------	-----------