Ceph - Bug #5205

mon: FAILED assert(ret == 0) on config's set_val_or_die() from pick_addresses()

05/30/2013 09:24 AM - Joao Eduardo Luis

Status:	Resolved	**% Done:**	0%
Priority:	Urgent	**Spent time:**	0.00 hour
Assignee:			
Category:	Monitor		
Target version:			
Source:	Community (user)	**Affected Versions:**	
Tags:		**ceph-qa-suite:**	
Backport:		**Pull request ID:**	
Regression:	No	**Crash signature (v1):**	
Severity:	3 - minor	**Crash signature (v2):**	
Reviewed:			

Description

This is the crash's log (from saaby @ #ceph):

```
root@ceph1-cph1c16-mon1:/var/lib/ceph/mon# ceph-mon -i ceph1-cph1c16-mon1 --debug-osd 20 -d
2013-05-29 13:39:23.760439 7f58c4b07780  0 ceph version 0.61.2-26-g1071736 (1071736d3b6611b6c5edeb
9b225f32b4e9afdc6d), process ceph-mon, pid 22748
2013-05-29 13:39:23.933491 7f58c4b07780  0 mon.ceph1-cph1c16-mon1 does not exist in monmap, will a
ttempt to join an existing cluster
common/config.cc: In function 'void md_config_t::set_val_or_die(const char*, const char*)' thread
7f58c4b07780 time 2013-05-29 13:39:23.933914
common/config.cc: 621: FAILED assert(ret == 0)
 ceph version 0.61.2-26-g1071736 (1071736d3b6611b6c5edeb9b225f32b4e9afdc6d)
 1: ceph-mon() [0x668046]
 2: ceph-mon() [0x69e889]
 3: (pick_addresses(CephContext*)+0x8d) [0x69e9ed]
 4: (main()+0x1a6b) [0x4a146b]
 5: (__libc_start_main()+0xfd) [0x7f58c2d86ead]
 6: ceph-mon() [0x4a3609]
 NOTE: a copy of the executable, or `objdump -rdS <executable>` is needed to interpret this.
```

Related issues:

Related to Ceph - Bug #5195: "ceph-deploy mon create" fails when adding addit...		**Resolved**	**05/29/2013**

Associated revisions

Revision eb86eebe - 06/24/2013 11:15 PM - Sage Weil

common/pick_addresses: behave even after internal_safe_to_start_threads

ceph-mon recently started using Preforker to working around forking issues.
As a result, internal_safe_to_start_threads got set sooner and calls to
pick_addresses() which try to set string config values now fail because
there are no config observers for them.

Work around this by observing the change while we adjust the value. We
assume pick_addresses() callers are smart enough to realize that their
result will be reflected by cct->_conf and not magically handled elsewhere.

Fixes: #5195, #5205
Backport: cuttlefish
Signed-off-by: Sage Weil <sage@inktank.com>
Reviewed-by: Dan Mick <dan.mick@inktank.com>

Revision 4d57c12f - 06/24/2013 11:16 PM - Sage Weil

common/pick_addresses: behave even after internal_safe_to_start_threads

ceph-mon recently started using Preforker to working around forking issues.
As a result, internal_safe_to_start_threads got set sooner and calls to
pick_addresses() which try to set string config values now fail because
there are no config observers for them.

Work around this by observing the change while we adjust the value. We
assume pick_addresses() callers are smart enough to realize that their
result will be reflected by cct->_conf and not magically handled elsewhere.

Fixes: #5195, #5205
Backport: cuttlefish
Signed-off-by: Sage Weil <sage@inktank.com>
Reviewed-by: Dan Mick <dan.mick@inktank.com>
(cherry picked from commit eb86eebe1ba42f04b46f7c3e3419b83eb6fe7f9a)

Revision 7ed6de9d - 08/09/2013 09:48 PM - Joao Eduardo Luis

common: pick_addresses: fix bug with observer class that triggered #5205

The Observer class we defined to observe conf changes and thus avoid
triggering #5205 (as fixed by eb86eebe1ba42f04b46f7c3e3419b83eb6fe7f9a),
was returning always the same const static array, which would lead us to
always populate the observer's list with an observer for 'public_addr'.

This would of course become a problem when trying to obtain the observer
for 'cluster_add' during md_config_t::set_val() -- thus triggering the
same assert as initially reported on #5205.

Backport: cuttlefish
Fixes: #5205

Signed-off-by: Joao Eduardo Luis <jecluis@gmail.com>
Reviewed-by: Sage Weil <sage@inktank.com>

Revision 2a34df68 - 08/09/2013 10:30 PM - Joao Eduardo Luis

common: pick_addresses: fix bug with observer class that triggered #5205

The Observer class we defined to observe conf changes and thus avoid
triggering #5205 (as fixed by eb86eebe1ba42f04b46f7c3e3419b83eb6fe7f9a),

was returning always the same const static array, which would lead us to always populate the observer's list with an observer for 'public_addr'.

This would of course become a problem when trying to obtain the observer for 'cluster_add' during md_config_t::set_val() -- thus triggering the same assert as initially reported on #5205.

Backport: cuttlefish
Fixes: #5205

Signed-off-by: Joao Eduardo Luis <jecluis@gmail.com>
Reviewed-by: Sage Weil <sage@inktank.com>
(cherry picked from commit 7ed6de9dd7aed59f3c5dd93e012cf080bcc36d8a)

History

#1 - 06/04/2013 10:11 PM - Adam Compton

I've also encountered this problem, running 0.61.2 on CentOS 6.4 (uname 2.6.32-220.el6.x86_64 #1 SMP Tue Dec 6 19:48:22 GMT 2011 x86_64 x86_64 x86_64 GNU/Linux).

I believe I've narrowed it down to the following circumstances:

1. Have an existing cluster to which you're trying to join a new mon.
2. Specify the "public network" configuration option in the [globals] section of ceph.conf
3. Follow the instructions in http://ceph.com/docs/next/rados/operations/add-or-rm-mons/ to set up a new mon (install the software, get the key and monmap, run --mkfs, edit ceph.conf, run "ceph mon add")

In this condition, when I get to step 8 (actually running ceph-mon), it immediately aborts with the same crash log as saaby provided above. This is repeatable in my cluster, although I haven't tried destroying and recreating the other mons to start from scratch. I dug around with gdb and it looks like the problem is in fill_in_one_address (IP addresses obscured):

#8 0x0000000000666c9e in fill_in_one_address (cct=0x1250000, ifa=<value optimized out>, networks="10.x.x.x/24",
conf_var=0x6cd350 "public_addr") at common/pick_address.cc:78
78   cct->_conf->set_val_or_die(conf_var, buf);
(gdb) p conf_var
$3 = 0x6cd350 "public_addr"
(gdb) p buf
$4 = "10.x.x.x\000\377\177\000\000X\316\377\377\005\000\000\000\000\000\210\066\001\000\000\000\000\000\017\000\000\000sK\000\000\351a\256Q\377"

I'm pretty sure buf isn't supposed to have all that gunk at the end. I did not dig further to figure out where it's coming from. As a workaround, removing the "public network" option in [globals] will let it get joined to the cluster, after which you can put the "public network" option back; ceph-mon won't break on subsequent startups.

#2 - 06/05/2013 04:14 AM - Joao Eduardo Luis

Thanks Adam, this provides great insight on what's going on.

#3 - 06/18/2013 01:04 PM - Sage Weil

- Status changed from New to 12

- Priority changed from Normal to High

#4 - 06/24/2013 11:14 AM - Sage Weil

- Priority changed from High to Urgent

#5 - 06/24/2013 12:53 PM - Sage Weil

- Status changed from 12 to Fix Under Review

#6 - 06/24/2013 04:24 PM - Sage Weil

- Status changed from Fix Under Review to Resolved