

## CephFS - Bug #43039

### client: shutdown race fails with status 141

11/27/2019 12:24 AM - Patrick Donnelly

<b>Status:</b>	Resolved	<b>% Done:</b>	0%
<b>Priority:</b>	High		
<b>Assignee:</b>	Xiubo Li		
<b>Category:</b>			
<b>Target version:</b>	v16.0.0		
<b>Source:</b>	Q/A	<b>ceph-qa-suite:</b>	
<b>Tags:</b>		<b>Component(FS):</b>	qa-suite
<b>Backport:</b>	octopus,nautilus	<b>Labels (FS):</b>	
<b>Regression:</b>	No	<b>Pull request ID:</b>	36515
<b>Severity:</b>	3 - minor	<b>Crash signature (v1):</b>	
<b>Reviewed:</b>		<b>Crash signature (v2):</b>	
<b>Affected Versions:</b>			
<b>Description</b>			
<pre>2019-11-26T11:33:09.502 INFO:tasks.workunit.client.0.smithi067.stdout:[ RUN      ] LibCephFS.ShutdownRace 2019-11-26T11:33:11.910 DEBUG:teuthology.orchestra.run:got remote process result: 141 2019-11-26T11:33:11.910 INFO:tasks.workunit:Stopping ['libcephfs/test.sh'] on client.0... 2019-11-26T11:33:11.911 INFO:teuthology.orchestra.run.smithi067:&gt; sudo rm -rf -- /home/ubuntu/ceph test/workunits.list.client.0 /home/ubuntu/ceph/test/clone.client.0 2019-11-26T11:33:12.126 ERROR:teuthology.run_tasks:Saw exception from tasks. Traceback (most recent call last):   File "/home/teuthworker/src/git.ceph.com_git_teuthology_master/teuthology/run_tasks.py", line 86 , in run_tasks     manager = run_one_task(taskname, ctx=ctx, config=config)   File "/home/teuthworker/src/git.ceph.com_git_teuthology_master/teuthology/run_tasks.py", line 65 , in run_one_task     return task(**kwargs)   File "/home/teuthworker/src/git.ceph.com_ceph-c_wip-pdonnell-testing-20191126.005014/qa/tasks/workunit.py", line 123, in task     timeout=timeout, cleanup=cleanup)   File "/home/teuthworker/src/git.ceph.com_git_teuthology_master/teuthology/parallel.py", line 87, in __exit__     for result in self:   File "/home/teuthworker/src/git.ceph.com_git_teuthology_master/teuthology/parallel.py", line 101 , in __next__     resurrect_traceback(result)   File "/home/teuthworker/src/git.ceph.com_git_teuthology_master/teuthology/parallel.py", line 37, in resurrect_traceback     reraise(*exc_info)   File "/home/teuthworker/src/git.ceph.com_git_teuthology_master/teuthology/parallel.py", line 24, in capture_traceback     return func(*args, **kwargs)   File "/home/teuthworker/src/git.ceph.com_ceph-c_wip-pdonnell-testing-20191126.005014/qa/tasks/workunit.py", line 409, in _run_tests     label="workunit test {workunit}".format(workunit=workunit)   File "/home/teuthworker/src/git.ceph.com_git_teuthology_master/teuthology/orchestra/remote.py", line 198, in run     r = self._runner(client=self.ssh, name=self.shortname, **kwargs)   File "/home/teuthworker/src/git.ceph.com_git_teuthology_master/teuthology/orchestra/run.py", lin e 433, in run     r.wait()   File "/home/teuthworker/src/git.ceph.com_git_teuthology_master/teuthology/orchestra/run.py", lin e 158, in wait</pre>			

```
self._raise_for_status()
File "/home/teuthworker/src/git.ceph.com_git_teuthology_master/teuthology/orchestra/run.py", line 180, in _raise_for_status
node=self.hostname, label=self.label
CommandFailedError: Command failed (workunit test libcephfs/test.sh) on smithi067 with status 141:
'mkdir -p -- /home/ubuntu/ceph/test/mnt.0/client.0/tmp && cd -- /home/ubuntu/ceph/test/mnt.0/client.0/tmp && CEPH_CLI_TEST_DUP_COMMAND=1 CEPH_REF=bd192f5b5e5ae468f12d6612b5c3fa83e8e7be38 TESTDIR="/home/ubuntu/ceph/test" CEPH_ARGS="--cluster ceph" CEPH_ID="0" PATH=$PATH:/usr/sbin CEPH_BASE=/home/ubuntu/ceph/test/clone.client.0 CEPH_ROOT=/home/ubuntu/ceph/test/clone.client.0 adjust-ulimits ceph-coverage /home/ubuntu/ceph/test/archive/coverage timeout 3h /home/ubuntu/ceph/test/clone.client.0/qa/workunits/libcephfs/test.sh'
```

From: /ceph/teuthology-archive/pdonnell-2019-11-26\_04:58:35-fs-wip-pdonnell-testing-20191126.005014-distro-basic-smithi/4543849/teuthology.log

```
2019-11-26T11:33:11.882+0000 7fb6d9fdb700 2 client.6481 unmounted.
2019-11-26T11:33:11.882+0000 7fb6d9fdb700 1 client.6481 shutdown
2019-11-26T11:33:11.883+0000 7fb6d9fdb700 1 -- 172.21.15.67:0/3399245369 >> [v2:172.21.15.67:3300/0,v1:172.21.15.67:6789/0] conn(0x7fb7bd749d40 msgr2=0x7fb7bd6aebc0 secure :-1 s=STATE_CONNECTION_ESTABLISHED l=1).mark_down
2019-11-26T11:33:11.883+0000 7fb6d9fdb700 1 --2- 172.21.15.67:0/3399245369 >> [v2:172.21.15.67:3300/0,v1:172.21.15.67:6789/0] conn(0x7fb7bd749d40 0x7fb7bd6aebc0 secure :-1 s=READY pgs=759 cs=0 l=1 rx=0x7fb75cb9a450 tx=0x7fb75caba6f0).stop
2019-11-26T11:33:11.883+0000 7fb6d9fdb700 1 -- 172.21.15.67:0/3399245369 shutdown_connections
2019-11-26T11:33:11.883+0000 7fb6d9fdb700 1 --2- 172.21.15.67:0/3399245369 >> [v2:172.21.15.22:6834/3196009262,v1:172.21.15.22:6835/3196009262] conn(0x7fb7bcad2ba0 0x7fb7bd75bab0 unknown :-1 s=CLOSED pgs=315 cs=0 l=0 rx=0 tx=0).stop
2019-11-26T11:33:11.883+0000 7fb6d9fdb700 1 --2- 172.21.15.67:0/3399245369 >> [v2:172.21.15.67:3300/0,v1:172.21.15.67:6789/0] conn(0x7fb7bd749d40 0x7fb7bd6aebc0 unknown :-1 s=CLOSED pgs=759 cs=0 l=1 rx=0 tx=0).stop
2019-11-26T11:33:11.883+0000 7fb6d9fdb700 1 --2- 172.21.15.67:0/3399245369 >> [v2:172.21.15.22:3301/0,v1:172.21.15.22:6790/0] conn(0x7fb7bcb72f80 0x7fb7bc01efd0 unknown :-1 s=CLOSED pgs=0 cs=0 l=0 rx=0 tx=0).stop
2019-11-26T11:33:11.883+0000 7fb6d9fdb700 1 -- 172.21.15.67:0/3399245369 shutdown_connections
2019-11-26T11:33:11.883+0000 7fb6d9fdb700 1 -- 172.21.15.67:0/3399245369 wait complete.
2019-11-26T11:33:11.883+0000 7fb6d9fdb700 1 -- 172.21.15.67:0/3399245369 >> 172.21.15.67:0/3399245369 conn(0x7fb7bd6bbc00 msgr2=0x7fb7bc01e5e0 unknown :-1 s=STATE_NONE l=0).mark_down
2019-11-26T11:33:11.883+0000 7fb6d9fdb700 20 client.6481 trim_cache size 0 max 16384
```

From: /ceph/teuthology-archive/pdonnell-2019-11-26\_04:58:35-fs-wip-pdonnell-testing-20191126.005014-distro-basic-smithi/4543849/remote/smithi067/log/ceph-client.admin.40393.log.gz

Related issues:	
Related to CephFS - Bug #50495: libcephfs: shutdown race fails with status 141	<b>New</b>
Duplicated by CephFS - Bug #46853: ceph_test_libcephfs: LibCephFS.TestUtime g...	<b>Duplicate</b>
Copied to CephFS - Backport #47020: nautilus: client: shutdown race fails wit...	<b>Resolved</b>
Copied to CephFS - Backport #47021: octopus: client: shutdown race fails with...	<b>Resolved</b>

## History

### #1 - 11/27/2019 10:33 AM - Jeff Layton

I think that's probably indicative of a SIGPIPE error, which probably means some task was writing to a pipe that did not have a reader.

Last week, I was getting some failures just after the ShutdownRace test ran. In that test we increase RLIMIT\_NOFILE to the max, and then reduce it once the test is done. What I was seeing was that as soon as the old RLIMIT\_NOFILE was restored, the next test after ShutdownRace would fail with EMFILE when it attempted to do an open().

I think there may have been some changes to the low-level messaging code that cause it to keep fd's open for longer than it used to. close() calls seem to have been deferred to a different task, and now when we go to reset RLIMIT\_NOFILE after the ShutdownRace test it sometimes causes

subsequent open calls to fail because the process is still over the limit.

**#2 - 12/09/2019 02:42 PM - Jeff Layton**

- Assignee changed from Jeff Layton to Patrick Donnelly

(Handing back to Patrick for now)

Is this problem still occurring in teuthology?

**#3 - 12/17/2019 05:18 PM - Patrick Donnelly**

- Status changed from New to Need More Info

- Assignee deleted (Patrick Donnelly)

Jeff Layton wrote:

(Handing back to Patrick for now)

Is this problem still occurring in teuthology?

Havne't seen it again. Let's see if it shows up again...

**#4 - 01/17/2020 10:53 PM - Patrick Donnelly**

- Target version deleted (v15.0.0)

**#5 - 02/17/2020 11:50 PM - Patrick Donnelly**

- Status changed from Need More Info to New

- Assignee set to Jeff Layton

- Target version set to v16.0.0

- Backport changed from nautilus to octopus,nautilus

/ceph/teuthology-archive/pdonnell-2020-02-15\_16:51:06-fs-wip-pdonnell-testing-20200215.033325-distro-basic-smithi/4767574/teuthology.log

**#6 - 02/18/2020 12:32 PM - Jeff Layton**

I took a look at the logs but there is nothing conclusive there. Again, I suspect that this is a problem down in the low-level messenger code -- we're probably hitting RLIMIT\_NOFILE somehow.

**#7 - 04/13/2020 01:26 PM - Jeff Layton**

- Status changed from New to Can't reproduce

I haven't heard of this cropping up anymore, and it seemed like more of a problem in low-level message handling code rather than anything in cephfs. Closing.

#### #8 - 07/01/2020 08:15 PM - Patrick Donnelly

- Status changed from *Can't reproduce* to *New*
- Assignee deleted (*Jeff Layton*)

/ceph/teuthology-archive/pdonnell-2020-07-01\_06:37:23-fs-wip-pdonnell-testing-20200701.033411-distro-basic-smithi/5193080/teuthology.log

#### #9 - 07/13/2020 05:17 PM - Patrick Donnelly

/ceph/teuthology-archive/pdonnell-2020-07-11\_02:43:08-fs-wip-pdonnell-testing-20200711.001802-distro-basic-smithi/5214184/teuthology.log

#### #10 - 08/07/2020 01:51 AM - Patrick Donnelly

- Duplicated by Bug #46853: *ceph\_test\_libcephfs: LibCephFS.TestUtime gets core dumped randomly added*

#### #11 - 08/07/2020 05:33 AM - Xiubo Li

Hi Jeff, Patrick

From my test locally, this issue is very similar to <https://tracker.ceph.com/issues/45829>, which dues to the open files reach up to the limit.

Locally I can reproduce the core dump very easily without any change in `src/tests/libcephfs/test.c` and I have add some log info about the open file limit:

```
[ RUN      ] LibCephFS.ShutdownRace
rold.rlim_cur:1024 rold.rlim_max:262144
```

The limitation is 1024 is too small, this is the reason why I can reproduce it so easy.

After setting it to 262144 at the end of `ShutdownRace()` test case, I couldn't reproduce it any more.

In the teuthology case, I think we should increase the limit to a larger number, just like in <https://github.com/ceph/teuthology/pull/1508>.

This could also explain why the core dump info are different many times.

#### #12 - 08/07/2020 06:39 AM - Xiubo Li

The root cause should be the test process reached the open files limit, and the fd returned is -1, or something, when accessing the `FileEvent` vector by using it, it trigger the following abort trace:

```
/usr/include/c++/8/bits/stl_vector.h:932
: std::vector<_Tp, _Alloc>::reference std::vector<_Tp, _Alloc>::operator[](std::vector<_Tp, _Alloc>::size_type
) [with _Tp = EventCenter::FileEvent; _Alloc = std::allocator<EventCenter::FileEvent>; std::vector<_Tp, _Alloc
>::reference = EventCenter::FileEvent&; std::vector<_Tp, _Alloc>::size_type = long unsigned int]: Assertion '_
_builtin_expect(__n < this->size(), true)' failed.
Aborted (core dumped)
```

Since we have enabled `\_GLIBCXX\_ASSERTIONS`, then we will always see the above Abort logs.

**#13 - 08/07/2020 08:46 AM - Xiubo Li**

- Pull request ID set to 36515

**#14 - 08/07/2020 08:46 AM - Xiubo Li**

- Status changed from New to Fix Under Review

- Assignee set to Xiubo Li

**#15 - 08/18/2020 03:40 AM - Patrick Donnelly**

- Status changed from Fix Under Review to Pending Backport

**#16 - 08/18/2020 04:38 PM - Nathan Cutler**

- Copied to Backport #47020: nautilus: client: shutdown race fails with status 141 added

**#17 - 08/18/2020 04:38 PM - Nathan Cutler**

- Copied to Backport #47021: octopus: client: shutdown race fails with status 141 added

**#18 - 04/26/2021 01:43 PM - Patrick Donnelly**

- Related to Bug #50495: libcephfs: shutdown race fails with status 141 added

**#19 - 06/15/2021 10:50 AM - Loïc Dachary**

- Status changed from Pending Backport to Resolved

While running with `--resolve-parent`, the script "backport-create-issue" noticed that all backports of this issue are in status "Resolved" or "Rejected".