

Linux kernel client - Bug #24976

can not umount and kernel crash

07/18/2018 07:56 AM - ☐ ☐ ☐

Status:	New	% Done:	0%
Priority:	Normal	Spent time:	0.00 hour
Assignee:	Zheng Yan		
Category:	fs/ceph		
Target version:			
Source:		Reviewed:	
Tags:		Affected Versions:	v12.2.5
Backport:		ceph-qa-suite:	
Regression:	No	Crash signature:	
Severity:	1 - critical		

Description

My environment:

```
[cephfsd@gz-ceph-52-205 ceph-deploy]$ cat /etc/redhat-release
CentOS Linux release 7.2.1511 (Core)
[cephfsd@gz-ceph-52-205 ceph-deploy]$ sudo ceph -v
ceph version 12.2.5 (cad919881333ac92274171586c827e01f554a70a) luminous (stable)
```

and my client is :

```
[root@gz-open-dev-c221 ~]# cat /etc/redhat-release
CentOS Linux release 7.2.1511 (Core)
```

cephfs client is not a ceph-node

after read-write file on cephfs,i try to umount mount point, but umount operation blocked and give me those:

```
Jul 18 15:13:59 gz-open-dev-c221 kernel: INFO: task umount:28175 blocked for more than 120 seconds.
Jul 18 15:13:59 gz-open-dev-c221 kernel: "echo 0 > /proc/sys/kernel/hung_task_timeout_secs" disables this message.
Jul 18 15:13:59 gz-open-dev-c221 kernel: umount          D ffff8808cf58bc60  0 28175  6658 0x00000080
Jul 18 15:13:59 gz-open-dev-c221 kernel: ffff8808cf58bb00 00000000000000086 ffff8801244f6780 ffff8808cf58bfd8
Jul 18 15:13:59 gz-open-dev-c221 kernel: ffff8808cf58bfd8 ffff8808cf58bfd8 ffff8801244f6780 ffff8808601147c0
Jul 18 15:13:59 gz-open-dev-c221 kernel: 00000000000000000 7fffffff81168d00 ffff8808cf58bc60
Jul 18 15:13:59 gz-open-dev-c221 kernel: Call Trace:
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff81168d00>] ? wait_on_page_read+0x60/0x60
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff81163b809>] schedule+0x29/0x70
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff8116394f9>] schedule_timeout+0x209/0x2d0
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff810c1a26>] ? dequeue_entity+0x106/0x520
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff810c3e38>] ? enqueue_task_fair+0x208/0x6c0
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff8101c859>] ? read_tsc+0x9/0x10
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff81168d00>] ? wait_on_page_read+0x60/0x60
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff81163ae3e>] io_schedule_timeout+0xae/0x130
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff81163aed8>] io_schedule+0x18/0x20
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff81168d0e>] sleep_on_page+0xe/0x20
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff811639680>] __wait_on_bit+0x60/0x90
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff81168a96>] wait_on_page_bit+0x86/0xb0
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff810a6b60>] ? wake_atomic_t_function+0x40/0x40
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff81168bd1>] filemap_fdatawait_range+0x111/0x1b0
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff81168c97>] filemap_fdatawait+0x27/0x30
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff811208fcd>] sync_inodes_sb+0x15d/0x1e0
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff8112100eb>] sync_filesystem+0x5b/0xa0
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff811e10e0>] generic_shutdown_super+0x30/0xe0
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff811e14e2>] kill_anon_super+0x12/0x20
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff8107f9130>] ceph_kill_sb+0x30/0x60 [ceph]
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff811e1899>] deactivate_locked_super+0x49/0x60
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff811e1e96>] deactivate_super+0x46/0x60
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff811fedf5>] mntput_no_expire+0xc5/0x120
Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffff811fff2f>] SyS_umount+0x9f/0x3c0
```

Jul 18 15:13:59 gz-open-dev-c221 kernel: [<ffffffff81646889>] system_call_fastpath+0x16/0x1b

after this tragedy,i restarted my ceph cluster and i can umount the point but can never mount cephfs again,also get those messages:

```
Jul 18 15:23:59 gz-open-dev-c221 kernel: INFO: task umount:28175 blocked for more than 120 seconds.
Jul 18 15:23:59 gz-open-dev-c221 kernel: "echo 0 > /proc/sys/kernel/hung_task_timeout_secs" disables this message.
Jul 18 15:23:59 gz-open-dev-c221 kernel: umount      D ffff8808cf58bc60   0 28175   6658 0x00000084
Jul 18 15:23:59 gz-open-dev-c221 kernel: ffff8808cf58bb00 00000000000000086 ffff8801244f6780 ffff8808cf58bfd8
Jul 18 15:23:59 gz-open-dev-c221 kernel: ffff8808cf58bfd8 ffff8808cf58bfd8 ffff8801244f6780 ffff8808601147c0
Jul 18 15:23:59 gz-open-dev-c221 kernel: 0000000000000000 7fffffffffffff ffffffff81168d00 ffff8808cf58bc60
Jul 18 15:23:59 gz-open-dev-c221 kernel: Call Trace:
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff81168d00>] ? wait_on_page_read+0x60/0x60
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff8163b809>] schedule+0x29/0x70
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff816394f9>] schedule_timeout+0x209/0x2d0
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff810c1a26>] ? dequeue_entity+0x106/0x520
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff810c3e38>] ? enqueue_task_fair+0x208/0x6c0
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff8101c859>] ? read_tsc+0x9/0x10
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff81168d00>] ? wait_on_page_read+0x60/0x60
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff8163ae3e>] io_schedule_timeout+0xae/0x130
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff8163aed8>] io_schedule+0x18/0x20
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff81168d0e>] sleep_on_page+0xe/0x20
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff81639680>] __wait_on_bit+0x60/0x90
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff81168a96>] wait_on_page_bit+0x86/0xb0
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff810a6b60>] ? wake_atomic_t_function+0x40/0x40
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff81168bd1>] filemap_fdatawait_range+0x111/0x1b0
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff81168c97>] filemap_fdatawait+0x27/0x30
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff81208fcd>] sync_inodes_sb+0x15d/0x1e0
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff812100eb>] sync_filesystem+0x5b/0xa0
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff811e10e0>] generic_shutdown_super+0x30/0xe0
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff811e14e2>] kill_anon_super+0x12/0x20
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffffffa07f9130>] ceph_kill_sb+0x30/0x60 [ceph]
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff811e1899>] deactivate_locked_super+0x49/0x60
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff811e1e96>] deactivate_super+0x46/0x60
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff811fedf5>] mntput_no_expire+0xc5/0x120
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff811fff2f>] SyS_umount+0x9f/0x3c0
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff81646889>] system_call_fastpath+0x16/0x1b
Jul 18 15:23:59 gz-open-dev-c221 kernel: INFO: task mount.ceph:32509 blocked for more than 120 seconds.
Jul 18 15:23:59 gz-open-dev-c221 kernel: "echo 0 > /proc/sys/kernel/hung_task_timeout_secs" disables this message.
Jul 18 15:23:59 gz-open-dev-c221 kernel: mount.ceph  D ffff8808c9e8e078   0 32509   1 0x00000080
Jul 18 15:23:59 gz-open-dev-c221 kernel: ffff88011b8fbc78 00000000000000086 ffff880371c05c00 ffff88011b8bfd8
Jul 18 15:23:59 gz-open-dev-c221 kernel: ffff88011b8bfd8 ffff88011b8bfd8 ffff880371c05c00 ffff880371c05c00
Jul 18 15:23:59 gz-open-dev-c221 kernel: ffff8808c9e8e068 ffff8808c9e8e070 ffffffff00000000 ffff8808c9e8e078
Jul 18 15:23:59 gz-open-dev-c221 kernel: Call Trace:
Jul 18 15:23:59 gz-open-dev-c221 kernel: INFO: task mount.ceph:32509 blocked for more than 120 seconds.
Jul 18 15:23:59 gz-open-dev-c221 kernel: "echo 0 > /proc/sys/kernel/hung_task_timeout_secs" disables this message.
Jul 18 15:23:59 gz-open-dev-c221 kernel: mount.ceph  D ffff8808c9e8e078   0 32509   1 0x00000080
Jul 18 15:23:59 gz-open-dev-c221 kernel: ffff88011b8fbc78 00000000000000086 ffff880371c05c00 ffff88011b8bfd8
Jul 18 15:23:59 gz-open-dev-c221 kernel: ffff88011b8bfd8 ffff88011b8bfd8 ffff880371c05c00 ffff880371c05c00
Jul 18 15:23:59 gz-open-dev-c221 kernel: ffff8808c9e8e068 ffff8808c9e8e070 ffffffff00000000 ffff8808c9e8e078
Jul 18 15:23:59 gz-open-dev-c221 kernel: Call Trace:
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff8163b809>] schedule+0x29/0x70
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff8163cfc5>] rwsem_down_write_failed+0x115/0x220
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffffffa07f9aa0>] ? ceph_inode_init_once+0x20/0x20 [ceph]
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff81301fe3>] call_rwsem_down_write_failed+0x13/0x20
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff8116c512>] ? mempool_create_node+0x72/0x140
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffffffa081c856>] ? ceph_mdsc_init+0x66/0x2f0 [ceph]
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff8163aa5d>] ? down_write+0x2d/0x30
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff811e1abe>] grab_super+0x2e/0xa0
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff811e2158>] sget+0x2a8/0x3d0
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffffffa07f9160>] ? ceph_kill_sb+0x60/0x60 [ceph]
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffffffa07fa103>] ceph_mount+0x303/0x850 [ceph]
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff811e2d39>] mount_fs+0x39/0x1b0
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff811fe5ff>] vfs_kern_mount+0x5f/0xf0
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff81200b4e>] do_mount+0x24e/0xa40
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff812013d6>] SyS_mount+0x96/0xf0
Jul 18 15:23:59 gz-open-dev-c221 kernel: [<ffffffff81646889>] system_call_fastpath+0x16/0x1b
```

History

#1 - 07/18/2018 08:02 AM - ☐ ☐ ☐

oh!by the way,i have turned off some feature in crushmap , like this:

```
1. begin crush map
   tunable choose_local_tries 0
   tunable choose_local_fallback_tries 0
   tunable choose_total_tries 50
   tunable chooseleaf_descend_once 1
   tunable chooseleaf_vary_r 0 ----->this one
   tunable chooseleaf_stable 0 ----->and this one
   tunable straw_calc_version 1
   tunable allowed_bucket_algs 54
```

#2 - 07/23/2018 04:48 PM - Patrick Donnelly

- *Project changed from Ceph to Linux kernel client*
- *Subject changed from cephfs:can not umount and kernel crush to can not umount and kernel crush*
- *Category deleted (librados)*
- *Assignee set to Zheng Yan*

#3 - 03/15/2019 06:38 PM - Ilya Dryomov

- *Category set to fs/ceph*