

CephFS - Bug #40746

client: removing dir reports "not empty" issue due to client side filled wrong dir offset

07/12/2019 08:30 AM - Peng Xie

Status:	Resolved	% Done:	0%
Priority:	Normal		
Assignee:	Peng Xie		
Category:			
Target version:	v15.0.0		
Source:	Community (dev)	ceph-qa-suite:	
Tags:		Component(FS):	Client
Backport:	nautilus,mimic,luminous	Labels (FS):	
Regression:	No	Pull request ID:	29005
Severity:	3 - minor	Crash signature (v1):	
Reviewed:		Crash signature (v2):	
Affected Versions:			

Description

recently, during use nfs-ganesha+cephfs, we found some "directory not empty error" when removing existing directory.

after deeper investigating the interaction between nfs-ganesha and cephfs, we root cause the problem due to the readdir missing some of the dir entries prior to the rmdir operation.

the problem is: during the first time of "ls" directory "a", ceph client readdir_cache is empty and fetches the entries from mds and fill the dir entry with the correct cookie for ganasha mdcache. however, the second time "ls" operation of directory "a" for the rmdir operation, missing the last entry in ganasha mdcache and filled from ceph client side readdir_cache, in which _readdir_cache_cb wrongly calculate the cookie, finally causing the misbehaving in ganasha mdcache.

for example, in our error case, directory "a" contains 2 ordinary files: "f6b6" and "f78c" here is the problematic logs we got in the ceph client, the detailed explanation was in the error log's comment:

ceph client log :

```
.....
2019-07-10 14:34:03.500338 7ff66a2d0700 10 client.13224161 readdir_r_cb 100001d7cac.head(faked_ino=0 ref=4 ll_ref=3
cap_refs={} open={} mode=40777 size=0/0 mtime=2019-07-10 12:24:29.172806 mds_cap_wantedFx
caps=pAsLsXsFsx(0=pAsLsXsFsx/0/0) COMPLETE parents=0x4b76920 0x49d1900) offset ff8f9a040000003 at_end=0
hash_order=1
2019-07-10 14:34:03.500349 7ff66a2d0700 10 client.13224161 offset ff8f9a040000003 snapid head (complete && ordered) 1 issued
pAsLsXsFsx
2019-07-10 14:34:03.500353 7ff66a2d0700 10 client.13224161 _readdir_cache_cb 0x2bb71e0 on 100001d7cac last_name f6b6
offset ff8f9a040000003 <<<---- the previous filled up dir entries "f6b6"
2019-07-10 14:34:03.500356 7ff66a2d0700 10 client.13224161 fill_stat on 100001d74ab snap/devhead mode 0100666 mtime
2019-07-10 12:52:13.901023 ctime 2019-07-10 12:29:48.863330
2019-07-10 14:34:03.500361 7ff66a2d0700 10 client.13224161 fill_dirent 'f78c' >100001d74ab type 8 w/ next_off
1000000000000000-
<<<<-- the next_off will be filled up 'f78c' dir entry's off and reply to nfs-ganesha as its mdcache
dirent's cookie which now be set wrongly to dir_result_t::END (1000000000000000) in the following code logic:
```

```
Client::_readdir_cache_cb(...) {
.....
uint64_t next_off = dn->offset + 1;
++pd;
if (pd == dir->readdir_cache.end())
next_off = dir_result_t::END; <<<<---- dir entry is "f78c" whose filled off should not be END
```

```
Inode *in = NULL;
    fill_dirent(&de, dn->name.c_str(), stx.stx_mode, stx.stx_ino, next_off)
    .....
}
```

At last, in nfs-ganesha, the expected the 'f78c's lookup cookie was "ffe9bfe10000003" mismatching with the "1000000000000000", so return the nfs client without 'f78c' and leaving directory "a" as not empty

Related issues:

Copied to CephFS - Backport #41855: nautilus: client: removing dir reports "n...	Resolved
Copied to CephFS - Backport #41856: mimic: client: removing dir reports "not ...	Resolved
Copied to CephFS - Backport #41857: luminous: client: removing dir reports "n...	Resolved

History

#1 - 07/12/2019 05:27 PM - Patrick Donnelly

- Project changed from Ceph to CephFS
- Subject changed from *nfs-ganesha removing dir reports "not empty" issue due to cephfs client side filled wrong dir offset* to *client: removing dir reports "not empty" issue due to client side filled wrong dir offset*
- Status changed from *New* to *Fix Under Review*
- Assignee set to *Peng Xie*
- Target version set to *v15.0.0*
- Start date deleted (*07/12/2019*)
- Source set to *Community (dev)*
- Backport set to *nautilus,mimic,luminous*
- Severity changed from *2 - major* to *3 - minor*
- Pull request ID set to *29005*
- Component(FS) Client added

#2 - 07/13/2019 03:12 AM - Zheng Yan

I don't see any problem. last paramter of fill_dirent() should be offset for next readdir. With your change, offset of current dentry is passed to fill_dirent

#3 - 09/13/2019 03:34 PM - Patrick Donnelly

- Status changed from *Fix Under Review* to *Pending Backport*

#4 - 09/16/2019 07:17 AM - Nathan Cutler

- Copied to Backport #41855: *nautilus: client: removing dir reports "not empty" issue due to client side filled wrong dir offset* added

#5 - 09/16/2019 07:17 AM - Nathan Cutler

- Copied to Backport #41856: *mimic: client: removing dir reports "not empty" issue due to client side filled wrong dir offset* added

#6 - 09/16/2019 07:17 AM - Nathan Cutler

- Copied to Backport #41857: *luminous: client: removing dir reports "not empty" issue due to client side filled wrong dir offset* added

#7 - 07/10/2020 05:57 PM - Nathan Cutler

- 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".