

## CephFS - Bug #47918

### cephfs client and nfs-ganesha have inconsistent reference count after release cache

10/21/2020 03:21 AM - wei liu

<b>Status:</b>	Resolved	<b>% Done:</b>	0%
<b>Priority:</b>	Normal		
<b>Assignee:</b>	wei liu		
<b>Category:</b>			
<b>Target version:</b>	v16.0.0		
<b>Source:</b>	Community (dev)	<b>Affected Versions:</b>	v14.2.10
<b>Tags:</b>		<b>ceph-qa-suite:</b>	
<b>Backport:</b>	octopus,nautilus	<b>Component(FS):</b>	Client
<b>Regression:</b>	No	<b>Labels (FS):</b>	
<b>Severity:</b>	3 - minor	<b>Pull request ID:</b>	37735
<b>Reviewed:</b>		<b>Crash signature:</b>	

#### Description

After nfs-ganesha has released the cache, cephfs client still holds many inodes in pin state.  
The number of caches is as follows[]

nfs-ganesha:

```
dbus-send --print-reply --system --dest=org.ganesha.nfsd /org/ganesha/nfsd/ExportMgr
org.ganesha.nfsd.exportstats.ShowCacheNode
method return time=1603248091.070865 sender=:1.12680 -> destination=:1.12986 serial=1158 reply_serial=2
boolean true
string "OK"
```

...

```
struct {
string " FSAL opened FD count : "
uint64 0
string " System limit on FDs : "
uint32 1048576
string " FD usage : "
string " Below Low Water Mark "
string " LRU entries in use : "
uint64 3
string " Chunks in use : "
uint64 0
}
```

cephfs Client:

```
{
"metadata": {
...
"pid": "30165",
"root": "/test10",
"timeout": "300",
"uuid": "ganesha-host-192-168-8-106-0002"
},
"dentry_count": 866,
"dentry_pinned_count": 866,
"id": 724497169,
"inst": {
"name": {
"type": "client",
"num": 724497169
},
"addr": {
"type": "v1",
```

```

"addr": "100.100.8.126:0",
"nonce": 875083291
},
"addr": {
"type": "v1",
"addr": "100.100.8.126:0",
"nonce": 875083291
},
"inst_str": "client.724497169 100.100.8.126:0/875083291",
"addr_str": "100.100.8.126:0/875083291",
"inode_count": 1181,
"mds_epoch": 162008,
"osd_epoch": 410332,
"osd_epoch_barrier": 0,
"blacklisted": false
}

```

nfs-ganesha holds 3 entries, but Client still holds 866 dentries in pin state. Even though client\_cache\_size=1 is set, it cannot be released.

ceph version: 14.2.10

#### Related issues:

Copied to CephFS - Backport #47988: nautilus: cephfs client and nfs-ganesha h...

Resolved

Copied to CephFS - Backport #47989: octopus: cephfs client and nfs-ganesha ha...

Resolved

#### History

##### #1 - 10/21/2020 07:25 AM - wei liu

Tracing to the client log the ll\_ref is increased 2 times each time lookup parent directory(..).

```

@2020-10-21 16:45:27.690 7fc9fb95f700 20 client.699904984 _ll_get 0x7fca37c67400 0x10003f85239 -> 169
2020-10-21 16:45:27.690 7fc9fb95f700 10 client.699904984 fill_statx on 0x10003f85239 snap/devhead mode 040755
mtime 2020-10-14 08:49:42.560573 ctime 2020-10-14 08:49:42.560573
2020-10-21 16:45:27.690 7fc9fb95f700 3 client.699904984 ll_getattrx 0x10003f85239.head = 0
2020-10-21 16:45:27.690 7fc9fb95f700 3 client.699904984 ll_lookupx 0x10003f85239.head ..
2020-10-21 16:45:27.690 7fc9fb95f700 20 client.699904984 may_lookup 0x10003f85239.head(faked_ino=0 ref=2 ll_re
f=169 cap_refs={} open={} mode=40755 size=0/0 nlink=1 btime=2020-10-13 09:52:02.759385 mtime=2020-10-14 08:49:
42.560573 ctime=2020-10-14 08:49:42.560573 caps=pAsLsXsFs(0=pAsLsXsFs) 0x7fca37c67400); UserPerm(uid: 0, gid:
0)
2020-10-21 16:45:27.690 7fc9fb95f700 10 client.699904984 _getattr mask AsXs issued=1
2020-10-21 16:45:27.690 7fc9fb95f700 3 client.699904984 may_lookup 0x7fca37c67400 = 0
2020-10-21 16:45:27.690 7fc9fb95f700 10 client.699904984 choose_target_mds resend_mds specified as mds.0
2020-10-21 16:45:27.690 7fc9fb95f700 20 client.699904984 mds is 0
2020-10-21 16:45:27.690 7fc9fb95f700 10 client.699904984 send_request rebuilding request 362280 for mds.0
2020-10-21 16:45:27.690 7fc9fb95f700 20 client.699904984 encode_cap_releases enter (req: 0x7fca2e01c100, mds:
0)
2020-10-21 16:45:27.690 7fc9fb95f700 20 client.699904984 send_request set sent_stamp to 2020-10-21 16:45:27.69
2342
2020-10-21 16:45:27.690 7fc9fb95f700 10 client.699904984 send_request client_request(unknown.0:362280 lookupa
parent #0x10003f85239 2020-10-21 16:45:27.692331 caller_uid=0, caller_gid=0{0,}) v4 to mds.0
2020-10-21 16:45:27.690 7fc9fb95f700 20 client.699904984 awaiting reply|forward|kick on 0x7fc9fb95bb50
2020-10-21 16:45:27.691 7fca35bfb700 20 client.699904984 handle_client_reply got a reply. Safe:1 tid 362280
2020-10-21 16:45:27.691 7fca35bfb700 10 client.699904984 insert_trace from 2020-10-21 16:45:27.692342 mds.0 is
_target=1 is_dentry=0
2020-10-21 16:45:27.691 7fca35bfb700 10 client.699904984 features 0xffffffffffffffff
2020-10-21 16:45:27.691 7fca35bfb700 10 client.699904984 update_snap_trace len 48
2020-10-21 16:45:27.691 7fca35bfb700 20 client.699904984 get_snap_realm 0x1 0x7fca37c281e0 40 -> 41
2020-10-21 16:45:27.691 7fca35bfb700 10 client.699904984 update_snap_trace snaprealm(0x1 nref=41 c=0 seq=1 par
ent=0x0 my_snaps=[] cached_snapc=1=[]) seq 1 <= 1 and same parent, SKIPPING
2020-10-21 16:45:27.691 7fca35bfb700 10 client.699904984 hrm is_target=1 is_dentry=0
2020-10-21 16:45:27.691 7fca35bfb700 12 client.699904984 add_update_inode had 0x10003bf25b7.head(faked_ino=0 r
ef=2 ll_ref=38 cap_refs={} open={} mode=40777 size=0/0 nlink=1 btime=2020-09-14 17:08:42.822517 mtime=2020-10-
13 11:01:31.986878 ctime=2020-10-13 11:01:31.986878 caps=pAsLsXsFs(0=pAsLsXsFs) 0x7fca37c64400) caps pAsLsXsFs
2020-10-21 16:45:27.691 7fca35bfb700 20 client.699904984 dir hash is 2
2020-10-21 16:45:27.691 7fca35bfb700 10 client.699904984 add_update_cap issued pAsLsXsFs -> pAsLsXsFs from mds
.0 on 0x10003bf25b7.head(faked_ino=0 ref=2 ll_ref=38 cap_refs={} open={} mode=40777 size=0/0 nlink=1 btime=202
0-09-14 17:08:42.822517 mtime=2020-10-13 11:01:31.986878 ctime=2020-10-13 11:01:31.986878 caps=pAsLsXsFs(0=pAs
LsXsFs) 0x7fca37c64400)
2020-10-21 16:45:27.692 7fca35bfb700 15 inode.get on 0x7fca37c64400 0x10003bf25b7.head now 3

```

2020-10-21 16:45:27.692 7fca35bfb700 20 client.699904984 put\_snap\_realm 0x1 0x7fca37c281e0 41 -> 40  
2020-10-21 16:45:27.692 7fca35bfb700 15 inode.get on 0x7fca37c64400 0x10003bf25b7.head now 4  
2020-10-21 16:45:27.692 7fca35bfb700 20 client.699904984 handle\_client\_reply signalling caller 0x7fc9fb95bb50  
2020-10-21 16:45:27.692 7fca35bfb700 20 client.699904984 handle\_client\_reply awaiting kickback on tid 362280 0  
x7fca35bf8730  
2020-10-21 16:45:27.692 7fc9fb95f700 20 client.699904984 sendrecv kickback on tid 362280 0x7fca35bf8730  
2020-10-21 16:45:27.692 7fc9fb95f700 15 inode.get on 0x7fca37c64400 0x10003bf25b7.head now 5  
2020-10-21 16:45:27.692 7fc9fb95f700 20 client.699904984 make\_request target is 0x10003bf25b7.head(faked\_ino=0  
ref=5 ll\_ref=38 cap\_refs={} open={} mode=40777 size=0/0 nlink=1 btime=2020-09-14 17:08:42.822517 mtime=2020-1  
0-13 11:01:31.986878 ctime=2020-10-13 11:01:31.986878 caps=pAsLsXsFs(0=pAsLsXsFs) 0x7fca37c64400)  
2020-10-21 16:45:27.692 7fc9fb95f700 20 client.699904984 lat 0.001131  
2020-10-21 16:45:27.692 7fc9fb95f700 20 client.699904984 ll\_get 0x7fca37c64400 0x10003bf25b7 -> 39  
2020-10-21 16:45:27.692 7fc9fb95f700 15 inode.get on 0x7fca37c64400 0x10003bf25b7.head now 6  
2020-10-21 16:45:27.692 7fc9fb95f700 8 client.699904984 ll\_lookup found target 0x10003bf25b7  
2020-10-21 16:45:27.692 7fc9fb95f700 10 client.699904984 put\_inode on 0x10003bf25b7.head(faked\_ino=0 ref=6 ll\_  
ref=39 cap\_refs={} open={} mode=40777 size=0/0 nlink=1 btime=2020-09-14 17:08:42.822517 mtime=2020-10-13 11:01  
:31.986878 ctime=2020-10-13 11:01:31.986878 caps=pAsLsXsFs(0=pAsLsXsFs) 0x7fca37c64400)  
2020-10-21 16:45:27.692 7fc9fb95f700 15 inode.put on 0x7fca37c64400 0x10003bf25b7.head now 5  
2020-10-21 16:45:27.692 7fc9fb95f700 10 client.699904984 ll\_lookup 0x10003f85239.head(faked\_ino=0 ref=2 ll\_ref=1  
69 cap\_refs={} open={} mode=40755 size=0/0 nlink=1 btime=2020-10-13 09:52:02.759385 mtime=2020-10-14 08:49:42.  
560573 ctime=2020-10-14 08:49:42.560573 caps=pAsLsXsFs(0=pAsLsXsFs) 0x7fca37c67400) .. = 0x10003bf25b7.head(fa  
ked\_ino=0 ref=5 ll\_ref=39 cap\_refs={} open={} mode=40777 size=0/0 nlink=1 btime=2020-09-14 17:08:42.822517 mti  
me=2020-10-13 11:01:31.986878 ctime=2020-10-13 11:01:31.986878 caps=pAsLsXsFs(0=pAsLsXsFs) 0x7fca37c64400)  
2020-10-21 16:45:27.692 7fc9fb95f700 10 client.699904984 fill\_statx on 0x10003bf25b7 snap/devhead mode 040777  
mtime 2020-10-13 11:01:31.986878 ctime 2020-10-13 11:01:31.986878  
2020-10-21 16:45:27.692 7fc9fb95f700 20 client.699904984 ll\_get 0x7fca37c64400 0x10003bf25b7 -> 40  
2020-10-21 16:45:27.692 7fc9fb95f700 3 client.699904984 ll\_lookupx 0x10003f85239.head .. -> 0 (10003bf25b7)  
2020-10-21 16:45:27.692 7fc9fb95f700 10 client.699904984 put\_inode on 0x10003bf25b7.head(faked\_ino=0 ref=5 ll\_  
ref=40 cap\_refs={} open={} mode=40777 size=0/0 nlink=1 btime=2020-09-14 17:08:42.822517 mtime=2020-10-13 11:01  
:31.986878 ctime=2020-10-13 11:01:31.986878 caps=pAsLsXsFs(0=pAsLsXsFs) 0x7fca37c64400)  
2020-10-21 16:45:27.692 7fc9fb95f700 15 inode.put on 0x7fca37c64400 0x10003bf25b7.head now 4  
2020-10-21 16:45:27.692 7fca35bfb700 10 client.699904984 put\_inode on 0x10003bf25b7.head(faked\_ino=0 ref=4 ll\_  
ref=40 cap\_refs={} open={} mode=40777 size=0/0 nlink=1 btime=2020-09-14 17:08:42.822517 mtime=2020-10-13 11:01  
:31.986878 ctime=2020-10-13 11:01:31.986878 caps=pAsLsXsFs(0=pAsLsXsFs) 0x7fca37c64400)  
2020-10-21 16:45:27.692 7fca35bfb700 15 inode.put on 0x7fca37c64400 0x10003bf25b7.head now 3  
2020-10-21 16:45:27.692 7fca35bfb700 10 client.699904984 put\_inode on 0x10003bf25b7.head(faked\_ino=0 ref=3 ll\_  
ref=40 cap\_refs={} open={} mode=40777 size=0/0 nlink=1 btime=2020-09-14 17:08:42.822517 mtime=2020-10-13 11:01  
:31.986878 ctime=2020-10-13 11:01:31.986878 caps=pAsLsXsFs(0=pAsLsXsFs) 0x7fca37c64400)  
2020-10-21 16:45:27.692 7fca35bfb700 15 inode.put on 0x7fca37c64400 0x10003bf25b7.head now 2  
2020-10-21 16:45:27.692 7fc9fb95f700 8 client.699904984 ll\_forget 0x10003bf25b7 1  
2020-10-21 16:45:27.692 7fc9fb95f700 20 client.699904984 ll\_put 0x7fca37c64400 0x10003bf25b7 1 -> 39  
2020-10-21 16:45:27.693 7fc9a49fd700 10 client.699904984 fill\_statx on 0x10003bf25b7 snap/devhead mode 040777  
mtime 2020-10-13 11:01:31.986878 ctime 2020-10-13 11:01:31.986878  
2020-10-21 16:45:27.693 7fc9a49fd700 3 client.699904984 ll\_getattrx 0x10003bf25b7.head = 0  
2020-10-21 16:45:27.693 7fc9a49fd700 10 client.699904984 fill\_statx on 0x10003bf25b7 snap/devhead mode 040777  
mtime 2020-10-13 11:01:31.986878 ctime 2020-10-13 11:01:31.986878  
2020-10-21 16:45:27.693 7fc9a49fd700 3 client.699904984 ll\_getattrx 0x10003bf25b7.head = 0  
2020-10-21 16:45:27.898 7fca373fe700 20 client.699904984 trim\_cache size 16 max 0  
2020-10-21 16:45:28.898 7fca373fe700 20 client.699904984 trim\_cache size 16 max 0  
2020-10-21 16:45:29.653 7fca403ff700 8 client.699904984 ll\_forget 0x10003f85239 1  
2020-10-21 16:45:29.653 7fca403ff700 20 client.699904984 ll\_put 0x7fca37c67400 0x10003f85239 1 -> 168@

**#2 - 10/21/2020 02:46 PM - Patrick Donnelly**

- Status changed from New to Fix Under Review
- Assignee set to wei liu
- Source set to Community (dev)
- Backport set to octopus,nautilus
- Pull request ID set to 37735

**#3 - 10/25/2020 11:28 PM - Patrick Donnelly**

- Status changed from Fix Under Review to Pending Backport

**#4 - 10/26/2020 10:33 AM - Nathan Cutler**

- Copied to Backport #47988: nautilus: cephfs client and nfs-ganesha have inconsistent reference count after release cache added

**#5 - 10/26/2020 10:33 AM - Nathan Cutler**

- Copied to Backport #47989: octopus: cephfs client and nfs-ganesha have inconsistent reference count after release cache added

**#6 - 12/02/2020 11:06 AM - 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".