

fs - Bug #40960

client: failed to drop dn and release caps causing mds stary stacking.

07/25/2019 02:32 PM - Xiaoxi Chen

Status:	Pending Backport	Start date:	
Priority:	Normal	Due date:	
Assignee:	Xiaoxi Chen	% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:	v15.0.0	Affected Versions:	v14.2.2
Source:	Community (dev)	ceph-qa-suite:	
Tags:		Component(FS):	Client
Backport:	nautilus,mimic,luminous	Labels (FS):	
Regression:	No	Pull request ID:	29321
Severity:	3 - minor	Crash signature:	
Reviewed:			

Description

when client get notification from MDS that a file has been deleted(via getting CEPH_CAP_LINK_SHARED cap for inode with nlink = 0, see handle_cap_grant), if the client hasnt touch the inode in the past, the ll_ref will be zero.

In previous code, we only call Client::unlink when ll_ref > 0, which is wrong and will leave the dn in cache, keeping the caps and resulting the inode stays in stray till the dn cache is dropped by kernel.

Under certain workload(write intensive and rotate intensive), this issue can cause stray stacking to several MILLIONS and causing huge space "leaking".

```
2019-07-24 02:09:03.527 7eff12ffd700 5 client.231690 handle_cap_grant on in 0x1000000279f mds.0 s
eq 3 caps now pAsLsXsFscr was pAsXsFscr (stale)
```

```
2019-07-24 02:09:03.527 7eff12ffd700 10 client.231690 update_inode_file_time 0x1000000279f.head(fa
ked_ino=0 ref=1 ll_ref=0 cap_refs={} open={} mode=100644 size=0/0 nlink=0 btime=0.000000 mtime=201
9-07-24 02:04:01.319475 ctime=2019-07-24 02:04:01.326440 caps=pAsXsFscr(0=pAsXsFscr) objectset[0x1
000000279f ts 0/0 objects 0 dirty_or_tx 0] parents=0x1000000279d.head["b"] 0x7eff18002940) pAsXsFs
cr ctime 2019-07-24 02:09:02.074122 mtime 2019-07-24 02:04:01.319475
```

```
2019-07-24 02:09:03.527 7eff12ffd700 10 client.231690 grant, new caps are Ls
```

Reproduce:

I have two client(14.2.2 fuse) mounting same ceph-fs,
/mnt/xiaoxi has 3 files, a, b and c.

Client A:
ls /mnt/xiaoxi
Client B:
ls /mnt/xiaoxi

Client A:
rm /mnt/xiaoxi/b
Client B(right after the rm):
ls /mnt/xiaoxi

After that, the b will stay in stray forever as client B holding pAsLsXsFscr, Client A does release all its caps

Related issues:

Copied to fs - Backport #41000: luminous: client: failed to drop dn and relea...	In Progress
Copied to fs - Backport #41001: mimic: client: failed to drop dn and release ...	In Progress
Copied to fs - Backport #41002: nautilus:client: failed to drop dn and releas...	Resolved

History**#1 - 07/25/2019 02:33 PM - Xiaoxi Chen**

- File Screen Shot 2019-07-25 at 10.30.22 PM.png added

#2 - 07/25/2019 05:51 PM - Patrick Donnelly

- Status changed from New to Need Review
- Target version set to v15.0.0
- Start date deleted (07/25/2019)
- Backport changed from nautilus,mimic to nautilus,mimic,luminous
- Pull request ID set to 29321

#3 - 07/29/2019 09:46 PM - Patrick Donnelly

- Subject changed from client failed to drop dn and release caps causing mds stary stacking. to client: failed to drop dn and release caps causing mds stary stacking.
- Status changed from Need Review to Pending Backport
- Component(FS) deleted (ceph-fuse)

#4 - 07/30/2019 01:38 AM - Xiaoxi Chen

- Copied to Backport #41000: luminous: client: failed to drop dn and release caps causing mds stary stacking. added

#5 - 07/30/2019 01:41 AM - Xiaoxi Chen

- Copied to Backport #41001: mimic: client: failed to drop dn and release caps causing mds stary stacking. added

#6 - 07/30/2019 01:42 AM - Xiaoxi Chen

- Copied to Backport #41002: nautilus:client: failed to drop dn and release caps causing mds stary stacking. added

#7 - 07/31/2019 09:08 AM - Xiaoxi Chen

some more background of this issue is under <https://tracker.ceph.com/issues/38679#note-9>

Files

Screen Shot 2019-07-25 at 10.30.22 PM.png	71.8 KB	07/25/2019	Xiaoxi Chen
-------------------------------------------	---------	------------	-------------