## CephFS - Bug #38679

**mds: behind on trimming and "[dentry] was purgeable but no longer is!"**

03/11/2019 05:25 PM - Patrick Donnelly

<table>
<thead>
<tr>
<th>Status:</th>
<th>Resolved</th>
<th>% Done:</th>
<th>0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priority:</td>
<td>High</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assignee:</td>
<td>Zheng Yan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Target version:</td>
<td>v15.0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source:</td>
<td>Q/A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Backport:</td>
<td>nautilus,mimic,luminous</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regression:</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severity:</td>
<td>3 - minor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reviewed:</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Tags:**
- Component(FS): MDS
- Labels (FS): multimds

**Affected Versions:**

**Description**

```
```

06/04/2021
2019-03-07 22:25:01.153 INFO:tasks.ceph.mds.b.smithi167.stderr:2019-03-07 22:25:01.153 7efd6c00570 0 -1 mds.0.cache.strays Dentry [dentry #0x100/stray/1000000a533 [2,head] auth (dversion lock) v=4 696 ino=0x1000000a533 state=1879048193 new | request=0 lock=0 inodrop=1 dirty=1 authpin=0 0x555e5d74f180] was purgeable but no longer is!

2019-03-07 22:25:01.153 INFO:tasks.ceph.mds.b.smithi167.stderr:2019-03-07 22:25:01.153 7efd6c00570 0 -1 mds.0.cache.strays Dentry [dentry #0x100/stray/1000000a533 [2,head] auth (dversion lock) v=4 696 ino=0x1000000a533 state=1879048193 new | request=0 lock=0 inodrop=1 dirty=1 authpin=0 0x555e5d74f180] was purgeable but no longer is!

2019-03-07T22:25:01.153 INFO:tasks.ceph.mds.b.smithi167.stderr:2019-03-07 22:25:01.153 7efd6c00570 0 -1 mds.0.cache.strays Dentry [dentry #0x100/stray/10000004524 [2,head] auth (dversion lock) v=4 696 ino=0x100000004524 state=1879048193 new | request=0 lock=0 inodrop=1 dirty=1 authpin=0 0x555e5d74f180] was purgeable but no longer is!

2019-03-07T22:25:01.153 INFO:tasks.ceph.mds.b.smithi167.stderr:2019-03-07 22:25:01.153 7efd6c00570 0 -1 mds.0.cache.strays Dentry [dentry #0x100/stray/10000004524 [2,head] auth (dversion lock) v=4 696 ino=0x100000004524 state=1879048193 new | request=0 lock=0 inodrop=1 dirty=1 authpin=0 0x555e5d74f180] was purgeable but no longer is!

2019-03-07 22:25:15.787 INFO:tasks.ceph.mds.b.smithi167.stdout:2019-03-07 22:25:15.787 7efd6c00570 0 -1 mds.0.cache.strays Dentry [dentry #0x100/stray/10000000a547 [2,head] auth (dversion lock) v=4 724 ino=0x10000000a547 state=1879048193 new | request=0 lock=0 inodrop=1 dirty=1 authpin=0 0x555e5d76f700] was purgeable but no longer is!

2019-03-07 22:25:15.787 INFO:tasks.ceph.mds.b.smithi167.stdout:2019-03-07 22:25:15.787 7efd6c00570 0 -1 mds.0.cache.strays Dentry [dentry #0x100/stray/10000000a547 [2,head] auth (dversion lock) v=4 724 ino=0x10000000a547 state=1879048193 new | request=0 lock=0 inodrop=1 dirty=1 authpin=0 0x555e5d76f700] was purgeable but no longer is!

2019-03-07 22:25:15.787 INFO:tasks.ceph.mds.b.smithi167.stdout:2019-03-07 22:25:15.787 7efd6c00570 0 -1 mds.0.cache.strays Dentry [dentry #0x100/stray/10000000a547 [2,head] auth (dversion lock) v=4 724 ino=0x10000000a547 state=1879048193 new | request=0 lock=0 inodrop=1 dirty=1 authpin=0 0x555e5d76f700] was purgeable but no longer is!

2019-03-07 22:25:15.787 INFO:tasks.ceph.mds.b.smithi167.stdout:2019-03-07 22:25:15.787 7efd6c00570 0 -1 mds.0.cache.strays Dentry [dentry #0x100/stray/10000000a547 [2,head] auth (dversion lock) v=4 724 ino=0x10000000a547 state=1879048193 new | request=0 lock=0 inodrop=1 dirty=1 authpin=0 0x555e5d76f700] was purgeable but no longer is!

2019-03-07T22:25:15.787 INFO:tasks.ceph.mds.b.smithi167.stdout:2019-03-07 22:25:15.787 7efd6c00570 0 -1 mds.0.cache.strays Dentry [dentry #0x100/stray/10000000a547 [2,head] auth (dversion lock) v=4 724 ino=0x10000000a547 state=1879048193 new | request=0 lock=0 inodrop=1 dirty=1 authpin=0 0x555e5d76f700] was purgeable but no longer is!

2019-03-07T22:25:15.787 INFO:tasks.ceph.mds.b.smithi167.stdout:2019-03-07 22:25:15.787 7efd6c00570 0 -1 mds.0.cache.strays Dentry [dentry #0x100/stray/10000000a547 [2,head] auth (dversion lock) v=4 724 ino=0x10000000a547 state=1879048193 new | request=0 lock=0 inodrop=1 dirty=1 authpin=0 0x555e5d76f700] was purgeable but no longer is!
was purgeable but no longer is!

2019-03-07T22:25:20.741 INFO:tasks.ceph.mds.b.smithi167.stderr:2019-03-07 22:25:20.748 7efd6c005700 -1 mds.0.cache.strays Dentry [dentry #0x100/stray5/1000000999f [2,head] auth (dversion lock) v=6 519 ino=0x1000000999f state=1879048192 | request=0 lock=0 inodepin=1 dirty=1 authpin=0 0x555e57da7500] was purgeable but no longer is!


From:
2/teuthology.log

Related issues:
Copied to CephFS - Backport #39221: luminous: mds: behind on trimming and "[d...
Resolved
Copied to CephFS - Backport #39222: nautilus: mds: behind on trimming and "[d...
Resolved
Copied to CephFS - Backport #39223: mimic: mds: behind on trimming and "[dent...
Resolved

History
#1 - 03/19/2019 03:42 AM - Zheng Yan

For the "was purgeable but no longer is" errors.

2019-03-07 22:25:01.126 7efd6c005700 10 mds.0.cache.ino(0x10000009bc2) auth_pin by 0x555e5ddfa700 on [inode 0x10000009bc2 [2,head] ~mds0/stray6/10000009bc2 auth v4642 ap=1 dirtyparent s=0 n1=0 n(v0 rc2019-03-07 22:24:26.391179 1=1+0) (iversion lock) | ptrwaiter=0 request=0 lock=0 caps=0 dirtyparent=1 dirty=1 authpin=1 0x555e5ddfa700] now 1

2019-03-07 22:25:01.126 7efd6c005700 20 mds.0.cache.ino(0x10000009bc2) store_backtrace: no dirtypool or no old pools

2019-03-07 22:25:01.126 7efd6c005700 1 -- [v2:172.21.15.167:6834/679747418,v1:172.21.15.167:6836/679747418] -- [v2:172.21.15.167:6816/6682,v1:172.21.15.170:6817/6682] -- osd_op(unknown.0.5:41412 3.5 3:b5d1c589::10000009bc2,00000000:head [create,setxattr parent (99),setxattr layout (30)] snapc 0=[] ondisk+write+known_if_redire芀eted+full_force e20) v8 -- 0x555e5f4bb80 con 0x555e52d7a000

2019-03-07 22:25:01.153 7efd6c005700 10 mds.0.cache.strays eval_stray [dentry #0x100/stray6/10000009bc2 [2,head] auth (dversion lock) v=4642 ino=0x10000009bc2 state=1879048209 | request=0 lock=0 inodepin=1 dirty=1 authpin=0 0x555e5c1c4fc0]

2019-03-07 22:25:01.153 7efd6c005700 10 mds.0.cache.strays inode is [inode 0x10000009bc2 [2,head] ~mds0/stray6/10000009bc2 auth v4642 ap=1 dirtyparent s=0 n1=0 n(v0 rc2019-03-07 22:24:26.391179 1=1+0) (iversion lock) | ptrwaiter=0 request=0 lock=0 caps=0 dirtyparent=1 dirty=1 authpin=1 0x555e5ddfa700]

2019-03-07 22:25:01.153 7efd6c005700 10 mds.0.cache.strays Dentry [dentry #0x100/stray6/10000009bc2 [2,head] auth (dversion lock) v=4642 ino=0x10000009bc2 state=1879048192 | request=0 lock=0 inodepin=1 dirty=1 authpin=0 0x555e5c1c4fc0] was purgeable but no longer is!

I think we can just remove the code that prints error
The "behind on trimming" was caused by a frozen dirfrag. The dirfrag stayed in frozen state for a long time is because Migrator::handle_export_prep() took long time.

diff --git a/src/mds/Migrator.cc b/src/mds/Migrator.cc
index 180fb90b2d..59f721921c 100644
--- a/src/mds/Migrator.cc
+++ b/src/mds/Migrator.cc
@@ -2534,6 +2534,7 @@ void Migrator::handle_export_prep(const MExportDirPrep::const_ref &m, bool did_a
    if (!bound) {
        dout(7) << " opening bounding dirfrag " << leaf << " on " << *in << dendl;
+            // this may take long if there are lots of bounding dirfrags
        cache->open_remote_dirfrag(in, leaf, cf.build());
        return;
    }

It seems not a valid fix as this is not the only path to have the error.

We hit this in another way, not due to frozen and migration but due to client has caps.

2019-07-10 03:27:51.458 7fc882aac700 10 mds.9.cache.strays  inode is [inode 0xa0001908553 [2,head] -mds9/stray 2/a0001908553 auth v12728687 s=7890291 n1=0 n(v0 rc2019-07-10 02:39:11.984908 b7890291 l=1+0) (iversion lock) caps={217702=pAsLsXsFscr/-@6} | ptrwaiter=0 request=0 lock=0 caps=1 dirtyparent=0 dirty=0 waiter=0 authpin=0 0 x55a7342f9100]...
2019-07-10 03:27:51.458 7fc882aac700 20 mds.9.cache.strays  caps | leases
2019-07-10 03:27:51.458 7fc882aac700 -1 mds.9.cache.strays Dentry [dentry #0x109/stray2/a0001908553 [2,head] a uth (iversion lock) v=12728687 ino=0xa0001908553 state=1342177280 | request=0 lock=0 inodepin=1 dirty=0 authpin=0 0x55a86e584f00] was purgeable but no longer is!

06/04/2021
root@cal-rno-mds-new-27:~# ceph daemon --cluster rno_ceph_cal mds.`hostname -s` dump inode 0xa0001908553
{
    "path": "/mds9/stray2/a0001908553",
    "ino": 1099514256291,
    "rdev": 0,
    "ctime": "2019-07-10 02:39:11.984908",
    "btime": "2019-07-09 06:33:06.046917",
    "mode": 33188,
    "uid": 0,
    "gid": 0,
    "nlink": 0,
    "dir_layout": {
        "dir_hash": 0
    },
    "layout": {
        "stripe_unit": 4194304,
        "stripe_count": 1,
        "object_size": 4194304,
        "pool_id": 2,
        "pool_ns": ""
    },
    "old_pools": [],
    "size": 7890291,
    "truncate_seq": 1,
    "truncate_size": 18446744073709551615,
    "truncate_from": 0,
    "truncate_pending": 0,
    "mtime": "2019-07-09 06:34:49.419935",
    "atime": "2019-07-09 06:33:06.046917",
    "time_warp_seq": 0,
    "change_attr": 1,
    "export_pin": -1,
    "client_ranges": [],
    "dirstat": {
        "version": 0,
        "mtime": "0.000000",
        "num_files": 0,
        "num_subdirs": 0
    },
    "rstat": {
        "version": 0,
        "rbytes": 7890291,
        "rfiles": 1,
        "rsubdirs": 0,
        "rsnaps": 0,
        "rctime": "2019-07-10 02:39:11.984908"
    },
    "accounted_rstat": {
        "version": 0,
        "rbytes": 7890291,
        "rfiles": 1,
        "rsubdirs": 0,
        "rsnaps": 0,
        "rctime": "2019-07-10 02:39:11.984908"
    },
    "version": 12728687,
    "file_data_version": 0,
    "xattr_version": 1,
    "backtrace_version": 12728687,
    "stray_prior_path": "ORIGINAL_PATH_HIDDEN",
    "symlink": "",
    "old_inodes": [],
    "dirfragtree": {
        "splits": []
    },
    "is_auth": true,
    "auth_state": {
        "replicas": []
    },
    "replica_state": {
               }
"authority": [
  9,
  -2
],
"replica_nonce": 0,
"auth_pins": 0,
"is_frozen": false,
"is_freezing": false,
"pins": {
  "ptrwaiter": 0,
  "request": 0,
  "lock": 0,
  "caps": 1,
  "dirtyparent": 0,
  "dirty": 0,
  "waiter": 0,
  "authpin": 0
},
"nref": 1,
"versionlock": {
  "gather_set": [],
  "is_leased": false,
  "num_rdlocks": 0,
  "num_wrlocks": 0,
  "num_xlocks": 0,
  "xlock_by": []
},
"authlock": {},
"linklock": {},
"dirfragtreeLock": {},
"filelock": {},
"xattrlock": {},
"nestlock": {
  "gather_set": [],
  "is_leased": false,
  "num_rdlocks": 0,
  "num_wrlocks": 0,
  "num_xlocks": 0,
  "xlock_by": []
},
"flocklock": {},
"policylock": {},
"states": [
  "auth",
  "notifyref",
  "orphan"
],
"client_caps": {
  "client_id": 217702,
  "pending": "pAsLxXsFscr",
  "issued": "pAsLxXsFscr",
  "wanted": "-",
  "last_sent": 6
},
"loner": -1,
"want_loner": -1,
"mds_caps_wanted": []
#11 - 07/11/2019 03:44 AM - Zheng Yan

not big issue even inode becomes to have caps

#12 - 07/11/2019 03:55 AM - Xiaoxi Chen

I think we need to try reclaim the caps in this case. I am seeing num_stray accumulated in my env due to the inodes have caps.

The clients that holding the caps are the cleanup clients --- i.e the one that trigger the unlink. What it does and only does is
find /import/ceph/callogs/ -maxdepth 3 -type f -name ".mul.gz" -cmin +1200
so the fd should closed and caps can be released however it doesn't.

    echo 3 > /proc/sys/vm/drop_caches on client side solve the issue, also shrink the #inodes in cache(fuse) from 1.7M down to 93k.

#13 - 07/11/2019 01:35 PM - Zheng Yan

is cephfs exported to nfs

#14 - 07/11/2019 02:53 PM - Xiaoxi Chen

Zheng Yan wrote:

    is cephfs exported to nfs

No, it is ceph-fuse (13.2.5).

It seems like customer has ~10 nodes mounted same share, all do something like
"find /import/ceph/callogs/ -maxdepth 3 -type f -name "*.mul.gz" -cmin +1200 -exec rm -f {} +" but it is a hand written java application...

We did a change this afternoon that 1) reduce concurrent from 10 to 2; 2) change from fuse to kernel and is seeing positive result.

Also, reducing mds_max_caps_per_client helps. So it seems like the ll_forget in ceph-fuse is not get called from kernel to release caps for the deleted file?

#15 - 10/18/2019 11:40 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".