

CephFS - Bug #46360

mgr/volumes: fs subvolume clones stuck in progress when libcephfs hits certain errors

07/05/2020 11:24 AM - Ramana Raja

Status:	Resolved	% Done:	0%
Priority:	Normal		
Assignee:	Ramana Raja		
Category:	Correctness/Safety		
Target version:	v16.0.0		
Source:	Community (dev)	Affected Versions:	
Tags:		ceph-qa-suite:	
Backport:	octopus,nautilus	Component(FS):	cephfs.pyx, mgr/volumes
Regression:	No	Labels (FS):	
Severity:	3 - minor	Pull request ID:	35934
Reviewed:		Crash signature:	

Description

During `fs subvolume clone`, libcephfs hit the "Disk quota exceeded error" that caused the subvolume clone to be stuck in progress instead of entering failed state. I could see the following traceback in the mgr log,

```
File "/home/r raja/git/ceph/src/pybind/mgr/volumes/fs/fs_util.py", line 117, in copy_file
    written += fs.write(dst_fd, data[written:], -1)
File "cephfs.pyx", line 1463, in cephfs.LibCephFS.write
cephfs.Error: error in write: Disk quota exceeded [Errno 122]
```

Traceback (most recent call last):

```
File "/home/r raja/git/ceph/src/pybind/mgr/volumes/fs/async_job.py", line 44, in run
    self.async_job.execute_job(vol_job[0], vol_job[1], should_cancel=lambda: thread_id.should_cancel())
File "/home/r raja/git/ceph/src/pybind/mgr/volumes/fs/async_cloner.py", line 309, in execute_job
    clone(self.vc, volname, job[0].decode('utf-8'), job[1].decode('utf-8'), self.state_table, should_cancel)
File "/home/r raja/git/ceph/src/pybind/mgr/volumes/fs/async_cloner.py", line 222, in clone
    start_clone_sm(volume_client, volname, index, groupname, subvolname, state_table, should_cancel)
File "/home/r raja/git/ceph/src/pybind/mgr/volumes/fs/async_cloner.py", line 202, in start_clone_sm
    (next_state, finished) = handler(volume_client, volname, index, groupname, subvolname, should_cancel)
File "/home/r raja/git/ceph/src/pybind/mgr/volumes/fs/async_cloner.py", line 159, in handle_clone_in_progress
    do_clone(volume_client, volname, groupname, subvolname, should_cancel)
File "/home/r raja/git/ceph/src/pybind/mgr/volumes/fs/async_cloner.py", line 155, in do_clone
    bulk_copy(fs_handle, src_path, dst_path, should_cancel)
File "/home/r raja/git/ceph/src/pybind/mgr/volumes/fs/async_cloner.py", line 144, in bulk_copy
    cptree(source_path, dst_path)
File "/home/r raja/git/ceph/src/pybind/mgr/volumes/fs/async_cloner.py", line 129, in cptree
    copy_file(fs_handle, d_full_src, d_full_dst, mo, cancel_check=should_cancel)
File "/home/r raja/git/ceph/src/pybind/mgr/volumes/fs/fs_util.py", line 120, in copy_file
    raise VolumeException(-e.args[0], e.args[1])
TypeError: bad operand type for unary -: 'str'
```

Digging further found that if a libcephfs return code is not converted into a python exception by cephfs.pyx, then cephfs.pyx raises an exception with a different argument than it normally does. See in cephfs.pyx,

```
cdef make_ex(ret, msg):
```

```

"""
Translate a librados return code into an exception.
"""
ret = abs(ret)
if ret in errno_to_exception:
    return errno_to_exception[ret](ret, msg)
else:
    return Error(msg + ': {} [Errno {:d}]'.format(os.strerror(ret), ret))

```

So it sometimes raises `cephfs.Error(ret, msg)` and sometimes `cephfs.Error(msg)`. The `mgr/volumes` only handles `cephfs.Error(ret, msg)` correctly.

Related issues:

Duplicated by CephFS - Bug #47798: pybind/mgr/volumes: TypeError: bad operand...	Duplicate
Copied to CephFS - Backport #46463: octopus: mgr/volumes: fs subvolume clones...	Resolved
Copied to CephFS - Backport #46464: nautilus: mgr/volumes: fs subvolume clone...	Resolved

History

#1 - 07/05/2020 11:43 AM - Ramana Raja

- Subject changed from *mgr/volumes: fs subvolume clones stuck in progress when certain errors are raised by libcephfs to mgr/volumes: fs subvolume clones stuck in progress when libcephfs hits certain errors*
- Description updated

#2 - 07/05/2020 11:55 AM - Ramana Raja

- Pull request ID set to 35934

#3 - 07/06/2020 04:18 PM - Patrick Donnelly

- Status changed from *New* to *Fix Under Review*
- Target version set to *v16.0.0*
- Backport set to *octopus,nautilus*

#4 - 07/09/2020 12:04 AM - Patrick Donnelly

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

#5 - 07/10/2020 05:51 PM - Nathan Cutler

- Copied to Backport #46463: *octopus: mgr/volumes: fs subvolume clones stuck in progress when libcephfs hits certain errors added*

#6 - 07/10/2020 05:51 PM - Nathan Cutler

- Copied to Backport #46464: *nautilus: mgr/volumes: fs subvolume clones stuck in progress when libcephfs hits certain errors added*

#7 - 10/13/2020 01:51 PM - Patrick Donnelly

- Duplicated by Bug #47798: *pybind/mgr/volumes: TypeError: bad operand type for unary -: 'str' for errno ETIMEDOUT added*

#8 - 10/23/2020 07:33 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".