NFSv4 Courteous Server

Implementation Status, Late 2021

Dai Ngo – October 7, 2021

What Is NFSv4 Courteous Server?

- A server which does not immediately expunge the client's state on lease expiration
- Expired client's states are destroyed only when there are conflicts with other clients
- Conflicts include lock, delegation and share reservation conflicts

Courtesy Server Motivation

- To help deal with transient network partition or temporary lost of network connectivity
  - Client does not recover locks after the network partition heals
  - Client cannot reclaim open states after the network partition heals - NFS4ERR_NO_GRACE
  - Might take awhile for client to re-establish all previous states

Courtesy Client

- An expired client that still has states
- No waiter for any locks owned this client, *fl_blocked _requests* is empty
- Maximum idle time is 24-hr
- Entire client lease is destroyed, not just the conflicted state, when there are conflicts with other clients' requests

Sources of conflict

- Local thread
- Another NFSv4 client
- A NFSv3 client with NLM

Type of Conflicts

- Delegation

  `_break_lease/lm_break/nfs4_break_deleg_cb`

- Lock & Test Lock

  `posix_lock_inode/lm_expire_lock/nfsd4_fl_expire_lock`

  `posix_test_lock/lm_expire_lock/nfsd4_fl_expire_lock`

- Share reservation: access conflict & deny conflict

  `nfs4_get_vfs_file`

- LOCKT conflict behavior

Admin

- To display courtesy client status

```
# cat /proc/fs/nfsd/clients/XX/info
clientid: 0xb08d86ef6153a96c
address: "10.80.62.94:743"
status: confirmed
courtesy client: no
seconds from last renew: 10
name: "Linux NFSv4.1 nfsvme14.us.oracle.com"
minor version: 1
Implementation domain: "kernel.org"
Implementation name: "Linux 5.15.0-rc1_ori+ #1 SMP Thu Feb 18 18:41:59 GMT 2021 x86_64"
Implementation time: [0, 0]
callback state: UP
callback address: 10.80.62.94:0
```

- To manually destroy courtesy client

```
# echo “expire” > /proc/fs/nfsd/clients/XX/ctl
```

Future Works

- Resolve conflict per conflicted state instead of destroy entire client lease.

Acknowledgment

- Thanks to Bruce Fields, Chuck Lever and Bill Baker for guidance with this project