Leases on Clustered Filesystem

Felix Blyakher
SGI
setlease overview

• setlease when successful allows local cashing
• lease is broken on conflicting operation
• lease could be F_RDLCK or F_WRLCK
• used by nfsv4 and samba
  – samba uses fcntl(F_SETLEASE)
  – nfsv4 delegations call internally vfs_setlease
setlease overview

- setlease when successful allows local caching
- lease is broken on conflicting operation
- lease could be F_RDLCK or F_WRLCK
- used by nfsv4 and samba
  - samba uses fcntl(F_SETLEASE)
  - nfsv4 delegations call internally vfs_setlease
- needs to be made cluster wide aware
State of affair yesterday

VFS
generic_setlease
  validate lease
  set FLEASE into inode i_flock

Filesystem
  unaware
State of affair today

VFS
  ->setlease

--------------------------------

Filesystem
  return -EINVAL

Correct behavior, but missing on leases optimizations
Cluster leases

VFS

->setlease

-------------------------------------

Filesystem

fs_setlease

validate lease cluster wide

set FLEASE into inode i_flock
Cluster leases prototype implementation

MDC
->setlease
fs_setlease
  generic_setlease
rpc to server --> generic_setlease
if (failure) delete lease

MDS
Cluster leases prototype implementation (cont.)

MDC open
  ->open

fs_open -------> fs_break_lease --> break_lease

break_lease
Issues

• cluster filesystem is interested in fl_break the same way as the original caller of setlease, both needs to be notified on lease break
Issues

• cluster filesystem is interested in fl_break the same way as the original caller of setlease, both needs to be notified on lease break

• Similar conflict exists in the file_lock_operations between nfs and the clustered filesystem, both needs to be notified on lock insert and removal
Issues

- cluster filesystem is interested in fl_break the same way as the original caller of setlease, both needs to be notified on lease break
- Similar conflict exists in the file_lock_operations between nfs and the clustered filesystem, both needs to be notified on lock insert and removal
- Some issues with leases on local filesystem, i.e. leases broken for hardlinked files
Conclusion

- The prototype showed it workable for the leases set with fcntl
- Some adjustments in handling lock_manager_operations may be needed to make it working with nfs leases