

CMD Orphans Handling HLD

Mike Pershin

February 8, 2008

1 Introduction

The orphans problem is occur due to concurrent unlink/rename operations on opened file. While file is open it shouldn't be removed. To provide correct functionality we don't remove file but save it under special directory named PENDING. After closing the file or server failure all orphans should be removed.

2 Requirements

Review the orphan handling scheme and adapt it to the CMD environment.

3 Functional specification

4 Use cases

5 Logic specification

6 State management

6.1 State invariants

6.2 Scalability & performance

6.3 Recovery changes

6.4 Disk format changes

6.5 Wire format changes

6.6 Protocol changes

6.7 API changes

6.8 RPCs order changes

7 Alternatives

The alternative idea is based on bottom FS functionality and on fact that we don't need to duplicate it actually:

1. do small patch to the `ldiskfs` which will deny `orphans_cleanup()` at mounting
2. add `fsfilt` ops that call `orphan_handling()`
3. after failure the `mds` will do recovery and all open replays will do `iget()` on orphans again, because they were not deleted due to 1)
4. call `fsfilt` ops mentioned 2) and all orphans will be processed in natural way - referenced will remain and others - deleted.

This way is very cheap but rely on FS functionality so `lustre` will depend on it.

8 Focus for inspections