Migration, Replication, and *Referrals*

Some Issues with RFC3530

Dave Noveck

60th IETF: August 3, 2004

Fs_locations Facilities

- Migration
  - Fs moves, client get MOVED error
  - Fs_locations tells him where it went
- Replication
  - Fs_locations tells client where replicas are
  - When server unresponsive, client looks there
- Referrals
  - What are referrals? Spec doesn't mention them.

What are Referrals?

- They're migrations when client is a bit late
- If client tries to access fs after it moves,
  - Could say "Never heard of it. You lose. Them's the breaks."
    - Client says "What do I do now?"
  - Or you could tell him using fs_locations
    - Client does a subset of migration
    - No state, fh's to worry about

But Spec Doesn't Mention Them

- But, it does support them
  - Some confusion, lack of clarity. Is not explicit.
  - Many descriptions assume fs has been there
  - But if you follow the spec *carefully,* it works
- Big issues:
  - Look at FH at beginning of op (for MOVED)
  - GETFH can return MOVED
  - How to do READDIR

READDIR Issues

- Dir contains mount points of absent fs's
- Returns MOVED when getting attributes
  - unless RDATTR_ERROR requested
  - Then RDATTR_ERROR gets MOVED
- Attr's to return
  - Fs_locations OK, fsid OK
  - Fileid not OK, bur mounted_on_fileid is OK

Evanescent Filehandles

- They’re the QM version of v4 filehandles
  - Yes, this is strange
- If you do GETFH at the root of absent fs
  - Get a moved error. Never see the fh
- You can do GETATTR(fs_locations).
- Fs root fh is … not persistent, not volatile
  - Until you do the migration and look
  - Then it chooses and you know which it was!

Pure Referrals

- Referrals are migrations after-the-fact
  - How long after?
  - Could be a very long time
- Pure referrals are fs was never really there
  - Notionally, fs moved during Jurassic
  - Doesn’t matter to client
- Allows a multi-server namespace

Referrals and Global Namespace

- Referrals do not provide global namespace
  - Does not provide any way for servers to co-operate
    - Namespace definition
    - Namespace discovery
- Situation like migration
  - No server-to-server migration protocol
  - Anybody interested in working on one?

Paths to Global Namespace

- Define a new server-to-protocol
  - Hasn't been much interest
- Use existing protocol together with a set of conventions
  - Could use v4
    - Servers could act as clients of master server which has the namespace description
  - Could use LDAP schema

What's in my Draft

- How to do referrals
  - Let me know of problems you see
- Places where spec is
  - Confusing, self-contradictory, generally obscure
  - Suggestion for fixing
- Includes referrals and other migration issues

Issues for the NFSv4.1 Spec

- What to do about a case in which,
  - V4.0 protocol is sound (no op changes)
  - But the description needs work
- New description is definitive for v4.1
- V4.0 is more troublesome.
  - You want greater clarity
  - But v4.1 spec cannot change v4.0

How about this?

- New description definitive for v4.x+1
- Descriptions for V4.x and v4.x+1 should be compatible, but
  - When there is a conflict, v4.x description is definitive for v4.x
  - Where the v4.x description is unclear or ambiguous, clarification may be provided by the v4.x+1 description.