Multi-Server Namespace
Chapter 10 in draft-03

Dave Noveck

Overview of Talk

- What's in it
  - Stuff from RFC3530
  - Explicit discussion of referrals
  - Stuff from *draft-noveck-nfsv4-migrep-00*
- Current status
- Going forward
- Questions and discussion

From RFC3530

- Migration and Replication
- Same as before except,
  - Eliminated contradictions
  - Clarified stuff
  - Cleanup and corrections from I-D:
    - draft-noveck-nfsv4-migrep-00
  - Tighter rules for GETATTR and READDIR
  - Unified through concept of an *absent* fs, which in turn allows explicit discussion of …

Referrals

- Explicitly mentioned and discussed
- Uses approach of *draft-ietf-nfsv4-referrals-00*
- Replaces fs movement discussions by fs absence as the basic concept

From draft-noveck-nfsv4-migrep-00

- Cleanup and corrections as mentioned
- Options when effecting fs transitions
- New attributes:
  - fs_absent: Simple boolean "is it here?"
  - fs_locations_info: fs_locations on steroids
    - Reasons to bulk up discussed below
  - fs_status: migration/replication related info on the current fs replica

fs_locations_info

- Basic theme: **More**
  - To deal with today's problems and the future's
  - More replicas
  - More different kinds of replicas
  - More knowledge helpful to client
  - More ability for server control
  - More support for continuous access under a variety of situations
  - More kinds of server deployments, e.g. clustering

Selection Priorities

- fs_locations had a list, no explicit priority
- fs_locations_info adds replica priorities
  - Also for referral selection
  - Levels of backup (server down, site disaster)
  - Separate prio's for writable and read-only.
  - Server can use to direct load-balancing

Types of Replicas

- What's it mean for A to be a replica of B?
  - They are exactly the same (incl. metadata, eg. fileids)
  - They have the same data
  - They have almost the same data
    - One's a point-in-time-copy of another
    - They're both point-in-time copies of a third replica
- What does RFC3530 say about this?
  - Nothing?
  - What's the right answer?
  - There isn't one

Replica Types, continued

- Different types of replicas are useful
  - Depends on requirements
  - Cannot legislate a single answer
  - Let client know the nature of replicas
    - Client can select appropriately
    - Client can adapt (and be more efficient).

Extreme Replica Types

- Two paths to the same thing:
  - Can be a replica from client point of view
- Different versions of "two paths"
  - Two paths to the same server
  - Two path to different servers with same clustered fs.
- Needs to connect with trunking piece of sessions, currently an open issue.

fs_status

- Descriptive information on this fs
  - With many replicas there will be issues
  - Helps in tracking them down
- Type of replica
  - Writable, fixed, periodic-update, versioned
- Supports to ensure time doesn't go bkwrd.
  - Requires client support

Changes from
draft-noveck-nfsv4-migrep-00

- Re-organized fs_locations_info for expandability
- Re-organized continuity information
- Added requested stuff
  - RDMA capability bit
  - how-current field to fs_status
- Deleted stuff due to perceived lack of interest
  - fh-replacement stuff
  - VLCACHE bit
  - Support for transparently splitting an fs

Status

- Ready for a thorough review
  - Converted from proposal to spec chapter
- Pending issues/items
  - Support for non-uniform namespaces
  - Updates for mandatory sessions
  - Integrate with session trunking support
    - Depends on what we decide to do there

Going Forward on Chapter 10

- Hope you've enjoyed the powerpoint
- There's no movie
- There's no TV show
- There's no video game
- Would those rights would belong to IETF?
- Anyway, read the chapter
- Send comments to the working group