



Federated FS Update

Robert Thurlow

Sun Microsystems

Robert.Thurlow@sun.com

A Brief History

- First, there was replication and failover
 - > ... with lists of locations a client could follow
- Then came migration and a MOVED error
 - > “It's not here anymore, see the list”
- Then we extended MOVED to referrals
 - > “It's not here (and never was, actually)”
- Then we baked it into NFSv4.1
 - > And sometimes even tested it
- Now, we need to manage server referrals
 - > That's FedFS!

What is FedFS about again?

- We want to build uniform namespaces
 - > /nfs/sun.com/home/thurlow same across company (or beyond)
 - > With any type of hardware or OS, of course!
 - > Making use of existing server contents
 - > With no changes to NFSv4
- We want to take advantage of V4 stuff
 - > Replica lists for right geo/load balancing
 - > Migrate files to new storage
- We want to get this **DEPLOYED!**
 - > Too cool to let languish

What are we defining?

- We spec a way for a client to find the top
 - > Bonjour, DNS!
 - > (This is new – don't have a cow, man)
- Clients see and follow referrals as they go
 - > ... some clients do this now
- We spec how referrals are managed
 - > Admin tool changes LDAP and server state
 - > Servers consult an LDAP server

“But doesn't pNFS get me this?”

- pNFS does give you some nice location independence
 - > ... but not a real namespace
- And would you put data servers in different time zones? Really?
- pNFS is great for a large amount of storage in a distance you would walk.
- A referrals-based global namespace can span the globe!

Terms we use

- Fileset: a directory tree that can be managed
- FSN (fileset name): a unique fileset identifier
 - > UUID plus an NSDB name (see below)
- FSL (fileset location): network location of a fileset instance
 - > Server name plus relative path
- Junction: an object on a server that stores an FSN
- NSDB (namespace database): a service that tracks the mapping between FSNs and FSLs; implemented with LDAP

Three main protocols

- Admin to NSDB
 - > Creates logical namespace of FSLs and FSNs
- Admin to file server
 - > Installs and maintains junctions on server
- Fileserver to NSDB
 - > Server dereferences junction

Referral Example (1)

- Client mounts server B: /
- User does "cd sales"



NFS server B



NFS Client



NFS server A



NSDB Server

Referral Example (2)

- Client mounts server B: /
- User does "cd sales"

PUTROOTFH
LOOKUP sales
GETFH



NFS server B



NFS Client



NFS server A



NSDB Server

Referral Example (3)

junction



NFS server B

- Client mounts server B:/"
- User does "cd sales"
- NFS server B determines sales is a junction



NFS Client



NFS server A



NSDB Server

Referral Example (4)

- Client mounts server B: /
- User does "cd sales"
- NFS server B determines sales is a junction
- NFS server B queries NSDB for FSL



NFS server B



NFS server A



NSDB Server

LDAP query



NFS Client

Referral Example (5)

- Client mounts server B: /
- User does "cd sales"
- NFS server B determines sales is a junction
- NFS server B queries NSDB for FSLs



NFS server B



NFS server A



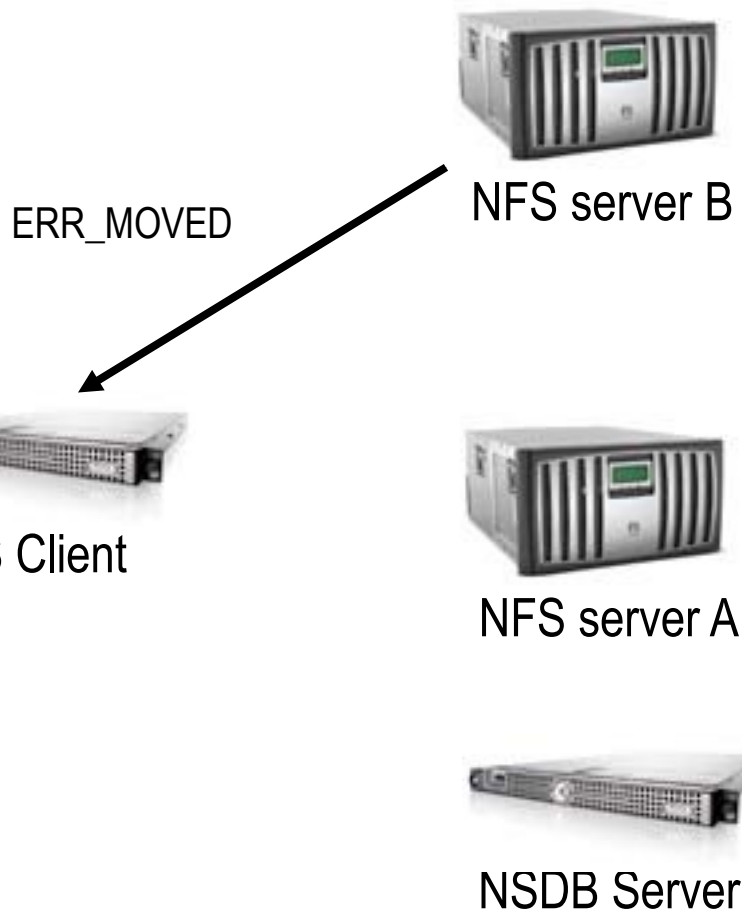
NSDB Server

LDAP reply



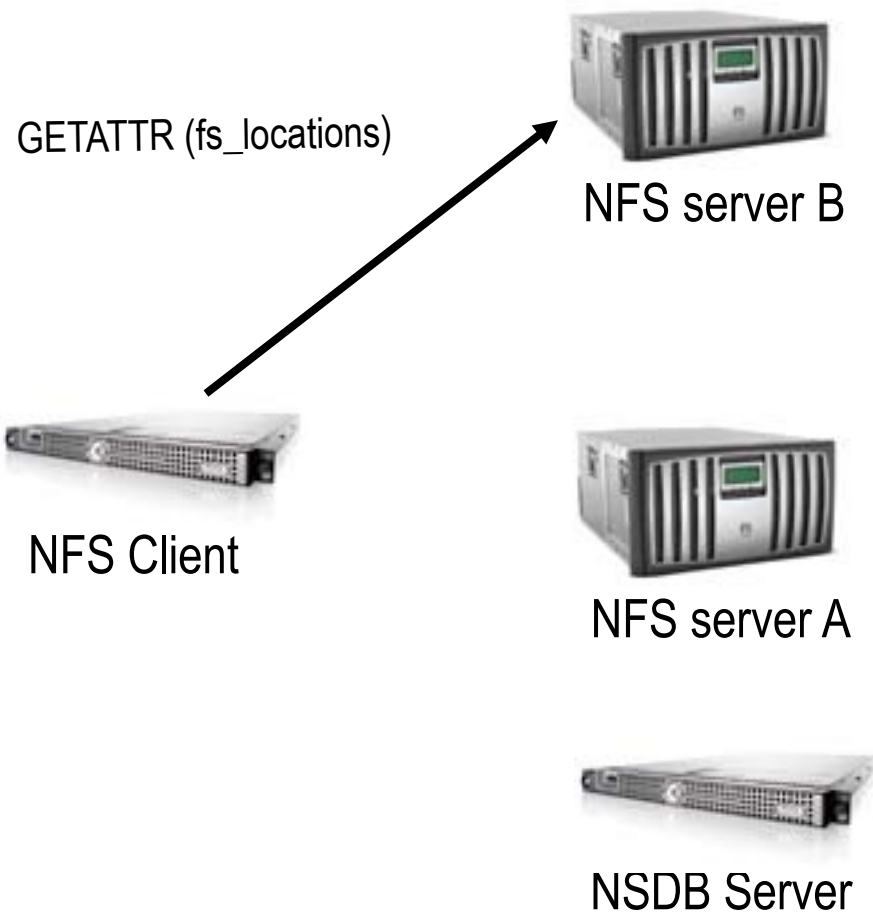
NFS Client

Referral Example (6)



- Client mounts server B:/'
- User does "cd sales"
- NFS server B determines sales is a junction
- NFS server B queries NSDB for FSLs
- NFS server B returns ERR_MOVED

Referral Example (7)



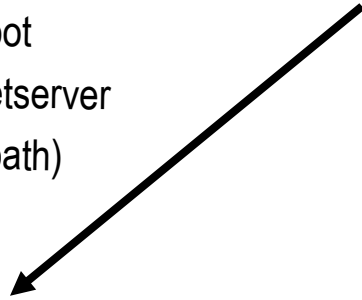
- ❑ Client mounts server B: /
- ❑ User does "cd sales"
- ❑ NFS server B determines sales is a junction
- ❑ NFS server B queries NSDB for FSLs
- ❑ NFS server B returns ERR_MOVED
- ❑ Client gets fs_locations

Referral Example (8)

fs_locations
 ■ fs_root
 ■ targetserver
 ■ rootpath)



NFS server B



NFS Client



NFS server A



NSDB Server

- ❑ Client mounts server B: /
- ❑ User does “cd sales”
- ❑ NFS server B determines sales is a junction
- ❑ NFS server B queries NSDB for FSLs
- ❑ NFS server B returns ERR_MOVED
- ❑ Client gets fs_locations

Referral Example (9)



NFS server B



PUTROOTFH



NFS server A



NSDB Server

- ❑ Client mounts server B: /
- ❑ User does "cd sales"
- ❑ NFS server A determines sales is a junction
- ❑ NFS server B queries NSDB for FSLs
- ❑ NFS server B returns ERR_MOVED
- ❑ Client gets fs_locations
- ❑ Client mounts NFS server A

Current Decision Points

- Will LDAP work for us?
- Should FedFS support CIFS?

Futures

- An admin would like point-and-click to:
 - > Browse the junctions
 - > Replicate a fileset
 - > Migrate a fileset
 - > Set QOS on a fileset
- Do we work on replication & migration some day?

Current status

- Charter addition in the hands of IESG
- Four NFSv4 working group drafts
 - > See <http://www.ietf.org/html.charters/nfsv4-charter.html>
 - > Requirements for Federated File Systems
 - > Using DNS SRV to Specify a Global File Name Space with NFS version 4
 - > Administration Protocol for Federated Filesystems
 - > NSDB Protocol for Federated Filesystems

The Admonition

- This is an official IETF work item now, so:
 - > Get involved with spec development
 - > Get those prototypes ready!
 - > Start working on product plans
- What can we test?
 - > This week?
 - > At the next Bake-a-thon?
- Please help me brainstorm Billboard data for next year



Federated FS Update

Robert Thurlow

Robert.Thurlow@sun.com