Ideas for an NFS namespace

Robert Thurlow,

Staff Engineer

Solaris Data Technology

Sun microsystems

We make the net work.

Customers see a gap

- They want a global namespace
  - Like AFS, DCE/DFS
    - one big enterprise (or better) namespace
    - delegation of authority to match data
- They don't think NFS has one
  - Automounter need not apply :-(
    - perception that it is Sun-only (or -mostly)
    - reality that it is not standardized
    - Reality that we have poor backing store story
  - But there's really nothing else out there

What do we actually have?

- Sun Automounter since 1988
  - User-level and kernel implementations
  - Works – Sun thrives on it
  - Permits creation of a large namespace
  - … but all subject to backing store :-(
- Surprisingly well deployed
  - HP-UX, Tru-64, AIX, IRIX, Linux, MacOS X
  - Also some PC presence: Hummingbird, WRQ, Solstice NFS Client

What do customers see?

- Doesn't work when they need it!
  - Map syntax not standardized
  - NIS and NIS+ not standardized
  - NIS is quite insecure
  - NIS+ not secure, and EOL'd by Sun
  - LDAP use with maps not standardized
  - LDAP maps not widely supported
  - Implementation inconsistencies hurt
  - Many don't know anything beyond /net
- “ Why doesn't somebody fix this”?

Requirements

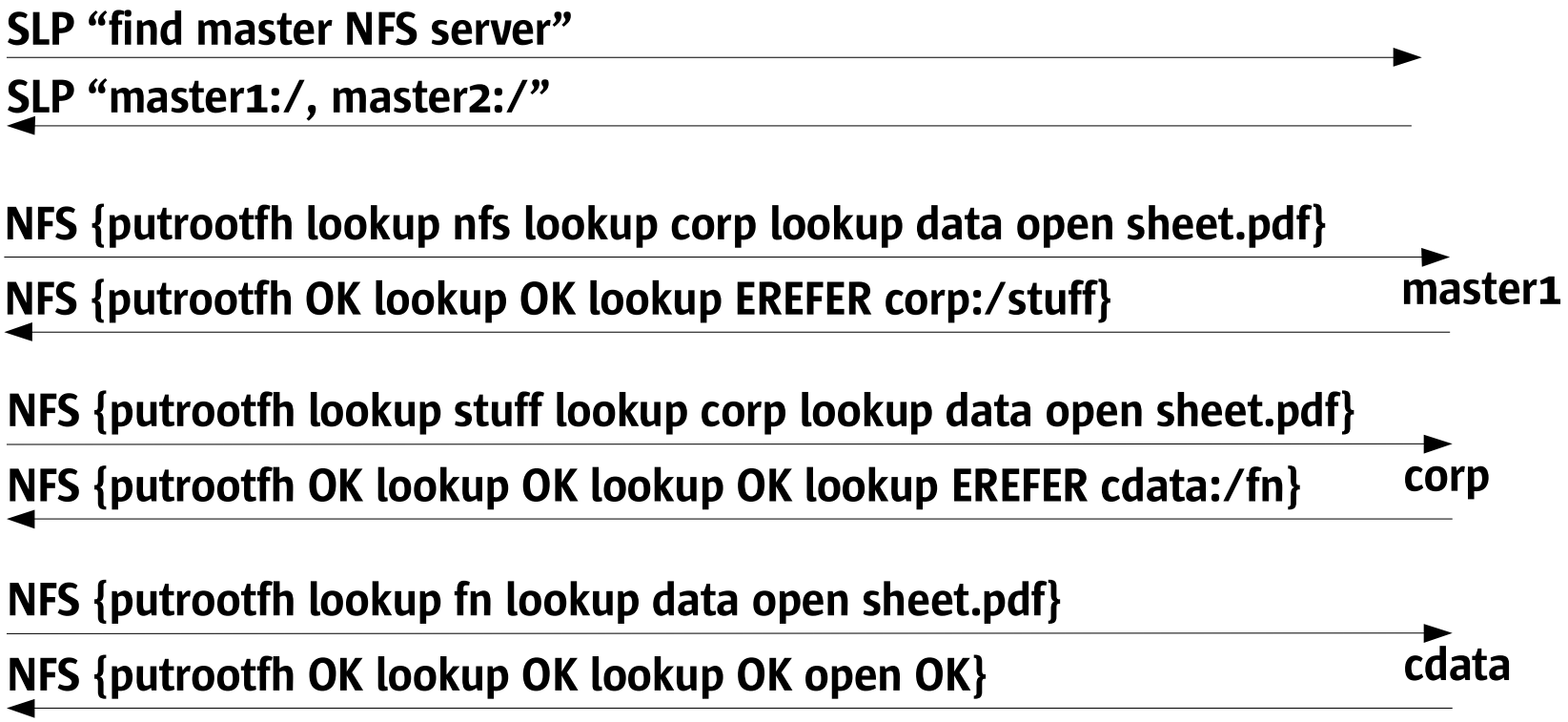
- Permit me to build enterprise-wide namespaces
- Permit me to delegate management of parts of namespace to owners of data
- Make this manageable from anywhere
- Don't make me add new naming service
- Permit some backwards compatibility

What do we need to do?

- Simplify client interaction!
  - LOOKUP extension (+EREFERRAL?) could return all info needed within NFS core protocol (including replica location, ro/rw, etc.)
  - Could use SLP to find your local “master browser” holding global root
  - Can then find referrals to the rest of an enterprise namespace
  - Could define an Internet-wide root with some cooperation

Client Example

fd = open("/nfs/corp/data/sheet.pdf", ...);



What do we need to do (2)?

- Start work on server-side
  - LDAP use seems likely
    - Document Sun LDAP map use as informational RFC
    - Discard specifics and start anew
    - Work towards a standards-track RFC
  - But other ideas are possible
    - Embed info in special “referral” nodes in filesystem
    - Use replication support to propagate changes!
    - Have to be able to enumerate referral nodes easily

What do we need to do (3)?

- Build a namespace manager
  - Able to import whole namespace from several sources
  - Pluggable backends to push changes to multiple name services for backwards compatibility
  - Add a way to “publish” new filesystem into a particular point in the namespace

End result

- Customers win
  - They can build and maintain global namespace
  - Can delegate authority properly
- Clients can get simpler in future
  - They can approach zero-config
- Can fix our name service mess
- … All for one amazing low price :-)

Robert Thurlow

robert.thurlow@sun.com

Sun microsystems

We make the net work.