NFS Namespace Futures: What's In A Name?

Rob Thurlow, Sun Microsystems

Overview

- What's a global namespace?
- Some use cases
- What do you need to make them work?
- Finding the top of a namespace
- NFSv4 referrals
- NFSv4.1 additions
- Global namespace management
- Further out: FAN

What's a Global Namespace?

- What if all people in your company:
  - Knew the same name for the same file
  - Could access the same file from anywhere
  - Could know their files are replicated without having a jumble of different names for them
  - Could use a file despite a server move
- That's what a global namespace does for you!
- Examples: AFS and DCE/DFS did this well

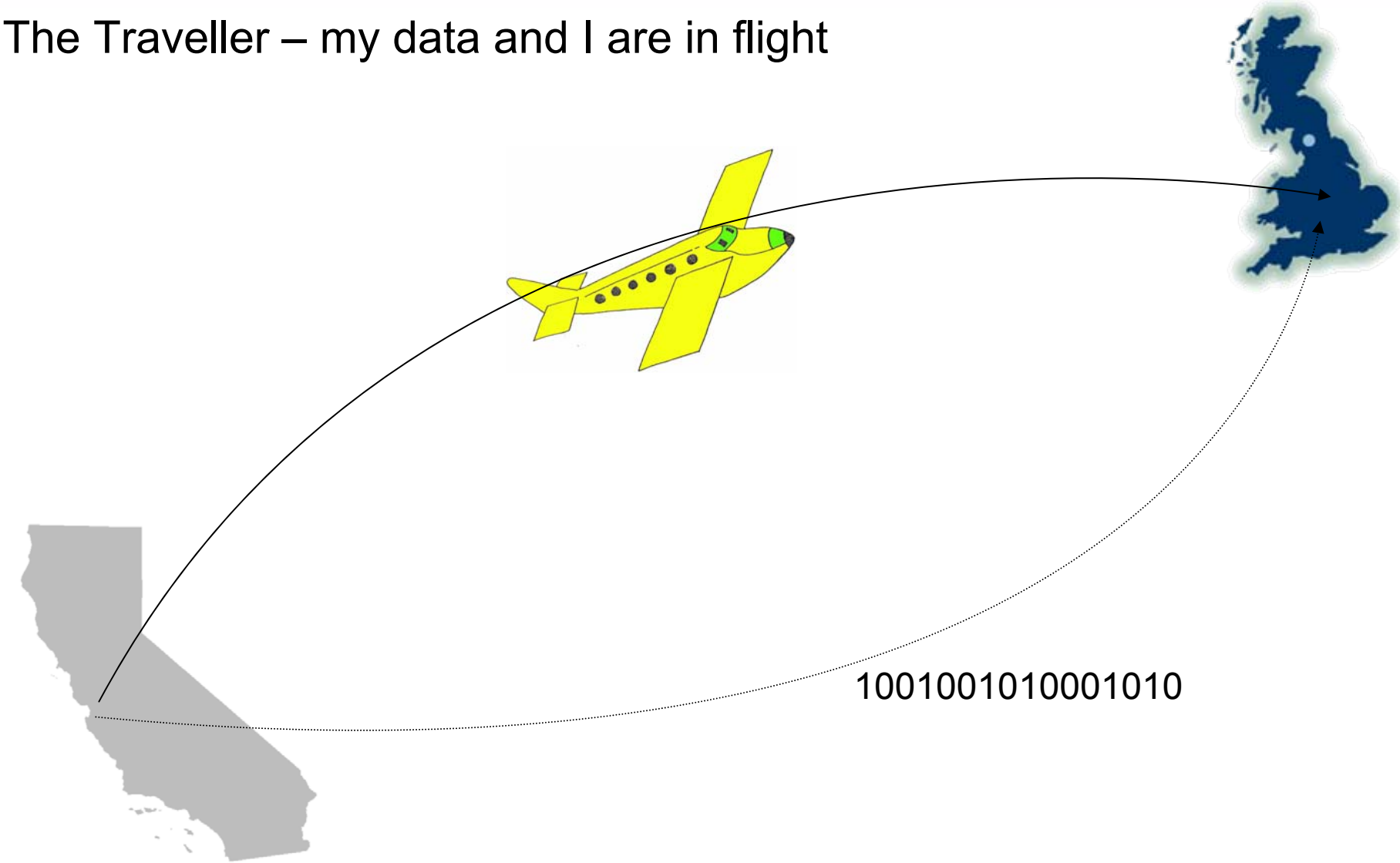
What's not a Global Namespace?

- pNFS
  - pNFS is about how you would create a BIG honking server, with great LAN scalability
  - Nice location independence inside the server
  - Limited ultimately by metadata server
- **Automounters**
  - They were a start, but maps don't scale
- “Global” is the key word

Some Use Cases

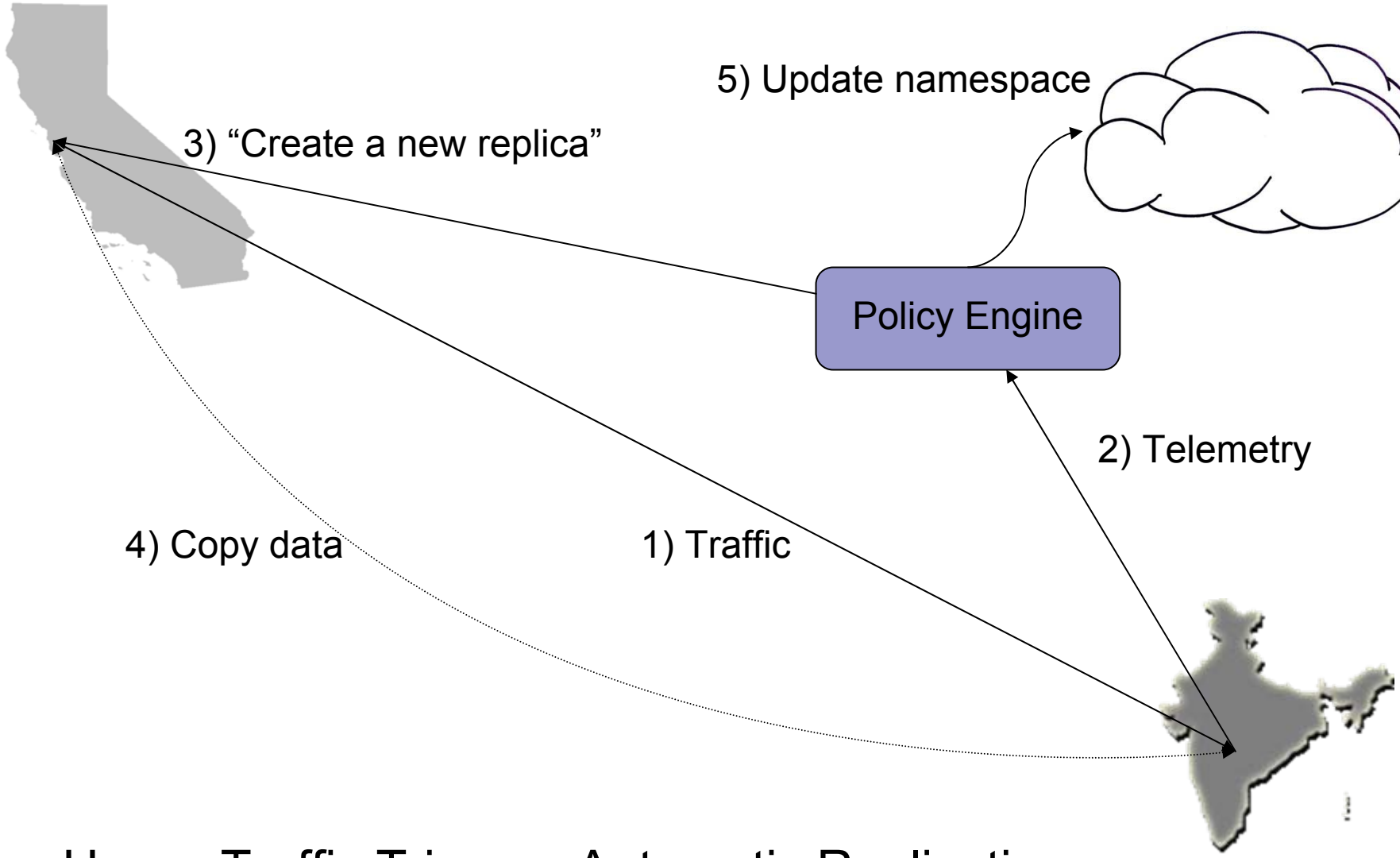
- The Traveller
  - Suppose I was working from the UK next week
  - In the past, I'd access my home dir across the Atlantic
  - With a global namespace, I could get a new read-write replica created in the UK
  - As I return, the master copy could revert

The Traveller – my data and I are in flight



Some Use Cases

- Automatic Replication
  - I hear of a rig in India which I could use for crunching some large datasets I have
  - Telemetry alerts a policy engine of the gap between provider and consumer
  - Through the global namespace, I get a new replica created in India
  - Computations produce results I can access at the same path via the namespace

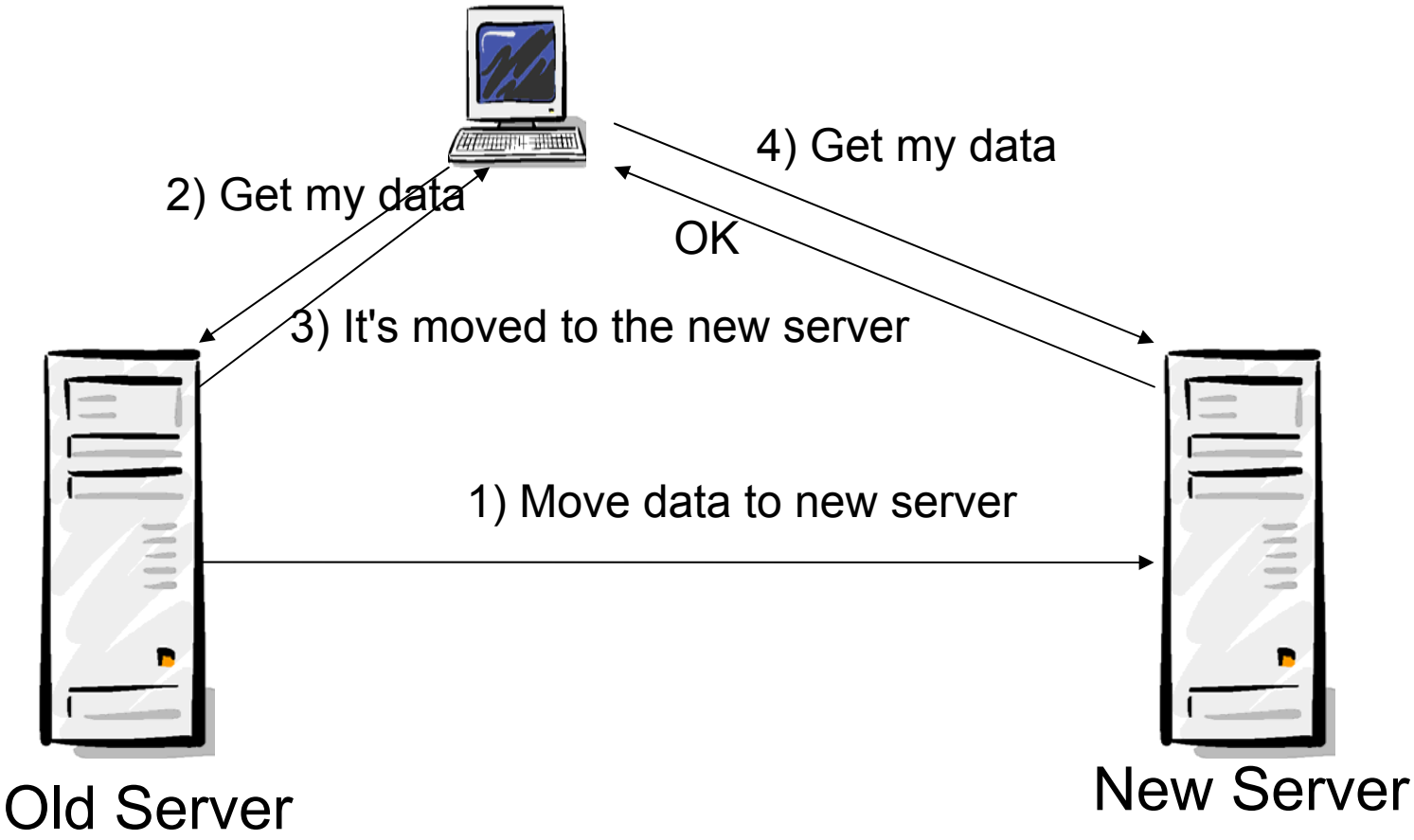


Heavy Traffic Triggers Automatic Replication

Some Use Cases

- Server Retirement
  - A server is being retired
  - In the past, this could be disruptive (log out or even reboot clients)
  - Now, the admin uses namespace tools to create and publish new locations
  - Then the server can point to its replacement and clients can just follow along

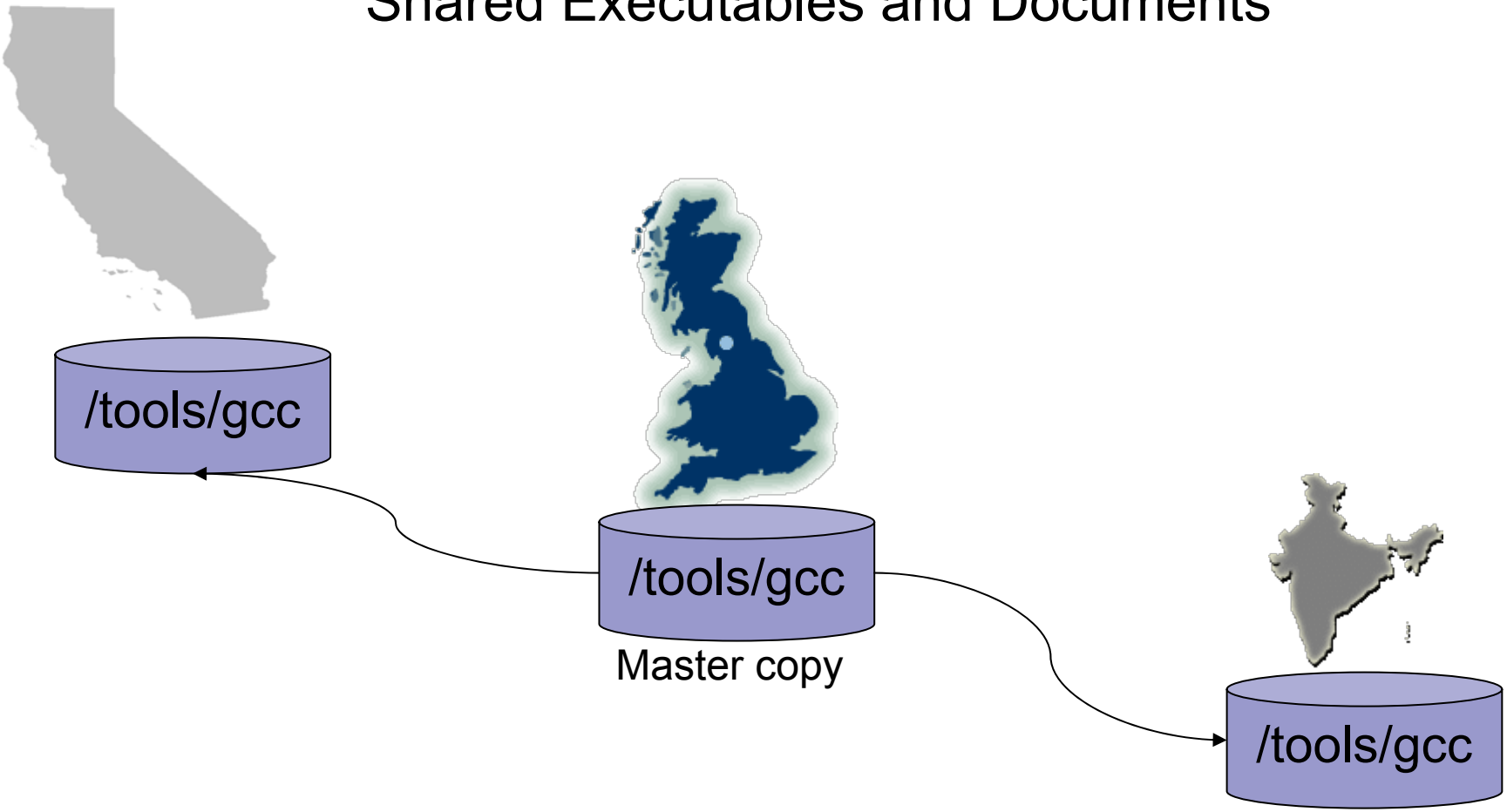
Server Retirement



Some Use Cases

- Shared executables and documents
  - All over your enterprise, people need access to some common files, be it Firefox or the current benefits document
  - Now, massive replication can put those in every branch with redundancy if the branch server is down
  - Overloaded servers can redirect clients away from them, knowing they can adapt

Shared Executables and Documents



What do you need?

- For a Global Namespace, you really only need:
  - A way to find the top level(s)
  - A way to get a “referral” from one file system point to a new server
  - A server that can issue referrals
  - A client that can follow referrals
- NFSv4 is growing up fast in these areas!

Finding the top

- How might we find the top of a namespace?
  - Very likely: Ask DNS
    - A DNS SRV record has been proposed:
    - http://tools.ietf.org/wg/nfsv4/draft-everhart-nfsv4-namespace-via-dns-srv-01.txt
  - Could extend this idea as well
    - Multicast DNS as in MacOS's “Bonjour”
- Once you know the top, a client can do deeper lookups
  - Getting referrals as it goes

NFSv4 Referrals

- Outgrowth of file system migration handling
- ObGeek: GETFH fails with NFS4ERR_MOVED
- FSID check reveals a new file system
  - so we know it's not a migration event
- Client asks for fs_locations
- Client selects and mounts a file system
  - Likely from a server it hasn't used before

NFSv4.1 Additions

- NFSv4 fs_locations work, but ...
  - Not much information to go on
- So we add more detail about the locations
  - fs_locations -> fs_locations_info
  - Which locations are writable?
  - Are any in fact sharing storage?
  - How current are they?
  - Rank and order to control “herding”

NFSv4.1 Additions

- Server can also know if the client supports referrals and migration
  - EXCHANGE_ID has capability bits
  - It's easier to decide to load balance if you know clients can handle it

Namespace management

- Current standards talk about client and server conversations
- But this misses some key things:
  - How does a server get referral info to hand out?
  - How does an admin make changes?
  - How does a server publish new locations?
- Work starting to try to standardize this:
  - http://www.ietf.org/internet-drafts/draft-ellard-federated-fs-01.txt
- Gathering requirements almost done, wire protocol next

Namespace management, cont'd

- New terms:
  - Fileset: a collection of files to be managed
  - NSDB: Namespace Database Service
  - FSN: Fileset Name (globally unique)
  - FSL: Fileset Location
- Servers have “junctions” pointing to FSNs
- Initially get a new FSN from an NSDB
- Then tell NSDB of FSLs for this FSN
- Then add more FSLs as they're created

Further Out - FAN

- SNIA File Area Network task force
  - Reference model includes a global namespace
  - Namespace decoupled from other operations
  - Federated nature to smooth out CIFS vs. NFS
  - Some talks with Open Grid Forum
  - FAN Task Force's ideas will go to SNIA and IETF

Conclusion

- Global namespaces are coming!
  - And they should make your life easier