



NetApp™

Go further, faster™

Federated Namespace BOF: Applications and Protocols

James Lentini
Advanced Technology Group
NetApp, Inc.

jlentini@netapp.com





OVERVIEW

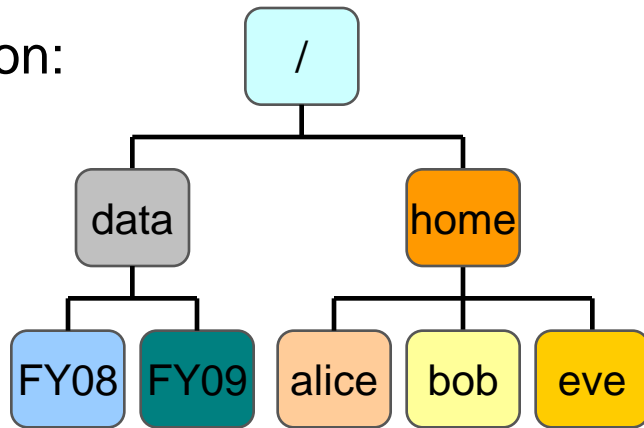


What is FedFS?

- FedFS is a set of open protocols that permit the construction of a scalable, federated file system namespace accessible to unmodified NFSv4[.1] clients.
- Key points:
 - Unmodified clients
 - Open: cross-platform, multi-vendor
 - Federated: participants retain control of their systems
 - Scalable: supports large namespaces with many clients and servers in different geographies

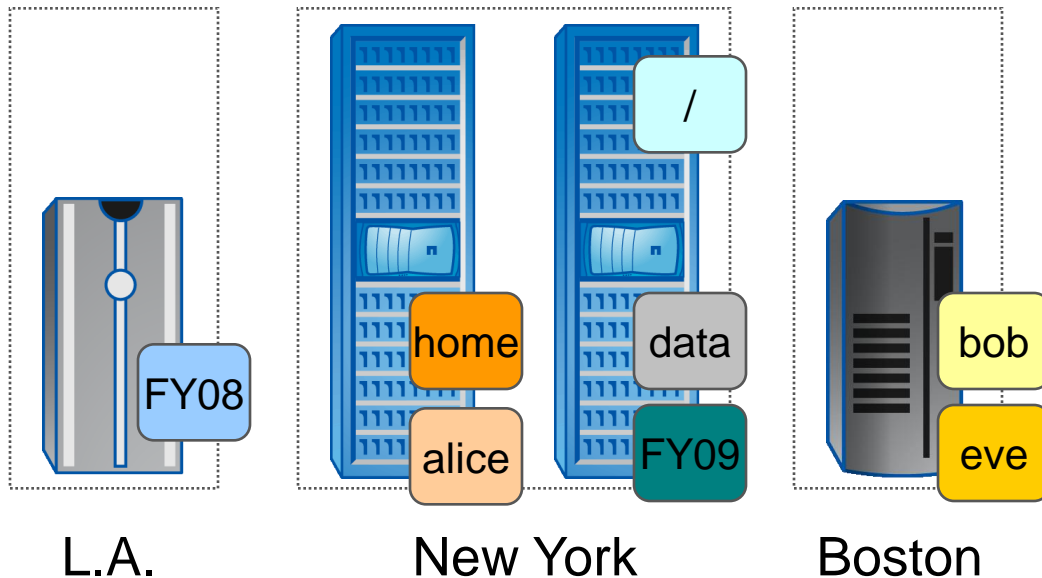
A Federated Namespace Example

The illusion:



The user and application software see a simple, hierarchical namespace.

The reality:



Behind the scenes, simple management operations allow data mobility for high performance, high reliability, and high availability.



What are the benefits?

- Simplified management: eliminates complicated software such as the automounter.
- Separates logical and physical data location: allows data movement for cost/performance tiering, worker mobility, and application mobility.
- Enhances:
 - Data Replication: for load balancing or high availability.
 - Data Migration: for moving data closer to compute or decommissioning systems.

TECHNICAL DETAILS

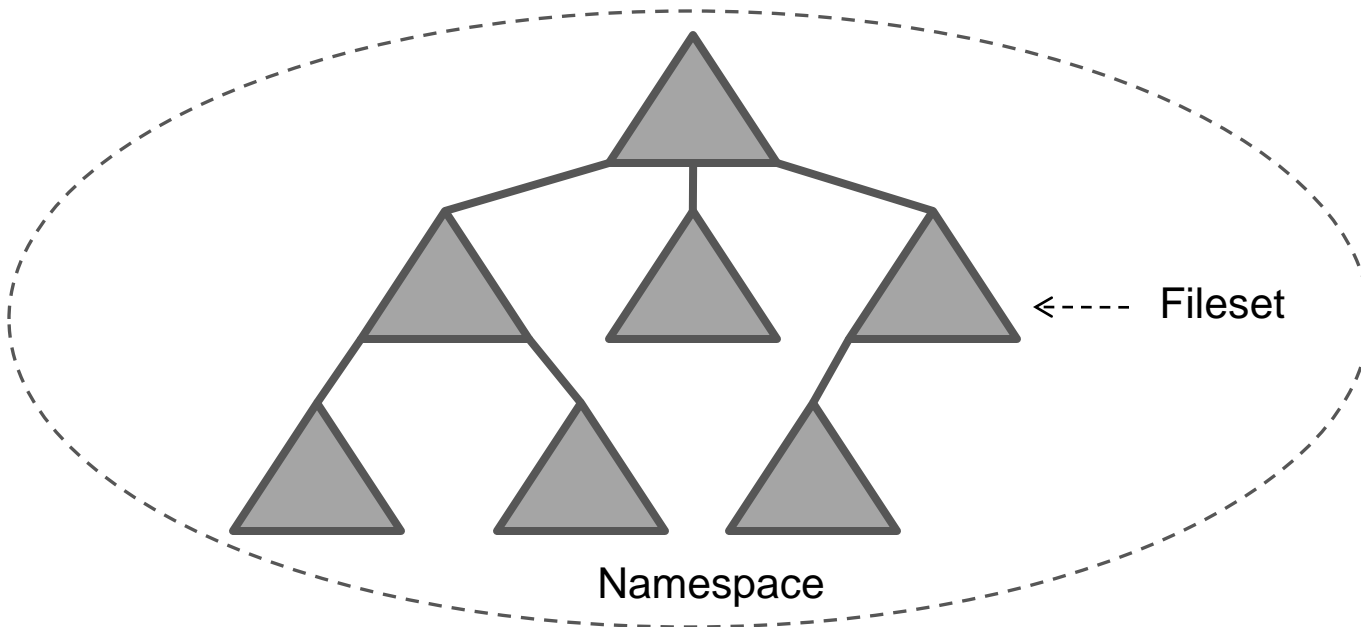


Referrals

- NFSv4 referrals permit a server to redirect a client to another location, possibly on another server.
- The server refers a client to another location by returning an `ERR_MOVED` error. The client can discover the object's new location via information in the
 - `fs_locations` attribute (v4)
 - `fs_locations_info` attribute (v4.1)
- FedFS allows administrators to construct a federated file system namespace using the referral mechanism.

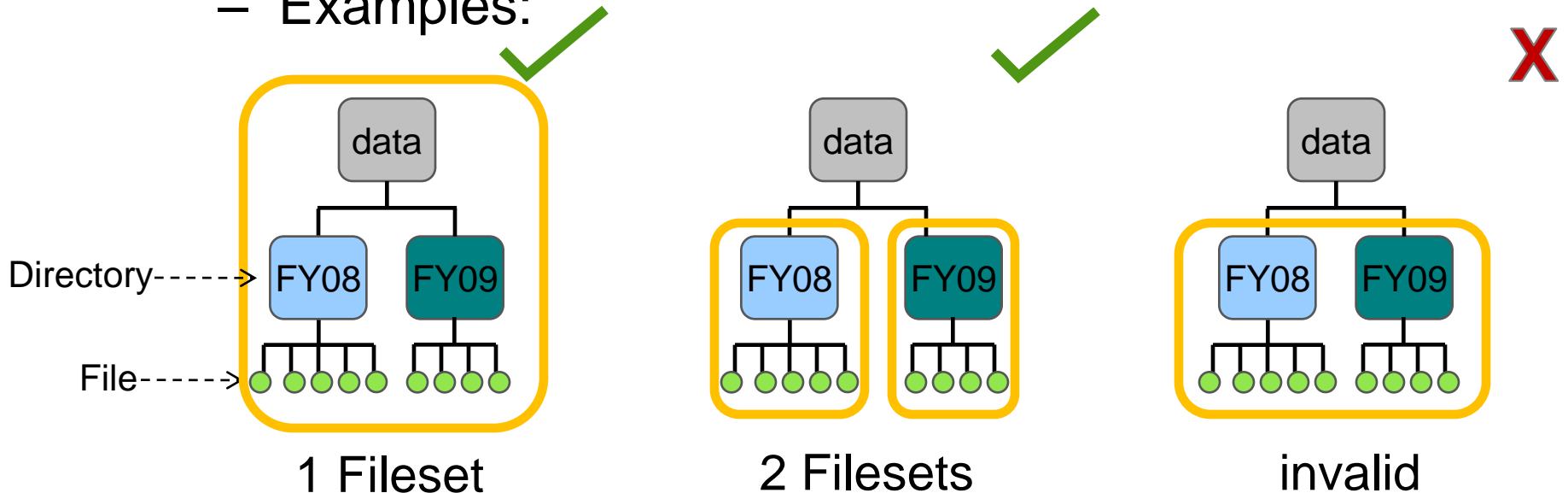
Federated File System Namespace

- Appears to the user as a single directory tree.
- Comprised of multiple filesets.
- Referrals allow the client to navigate the namespace.
 - No special client configuration is required.



Filesets

- A **fileset** is a set of files and the directory tree that contains them.
 - All members of a fileset are descendants of the same directory tree.
 - Abstract concept; no on disk changes (unless the fileset contains a junction).
 - Examples:





Fileset Names and Locations

- A **Fileset Name (FSN)** uniquely identifies a fileset. Includes:
 - RFC4122 Universal Unique Identifier (UUID)
 - DNS name of namespace database
- A **Fileset Location (FSL)** describes an instantiation of a FSN. Includes:
 - Fileserver location (DNS name or IP)
 - Export path
- An FSN maps to one or more FSLs.



Junction

- A file system object used to link a directory name in the current fileset to the root of the target fileset
- A leaf object of a fileset
- Stitches together the federated namespace

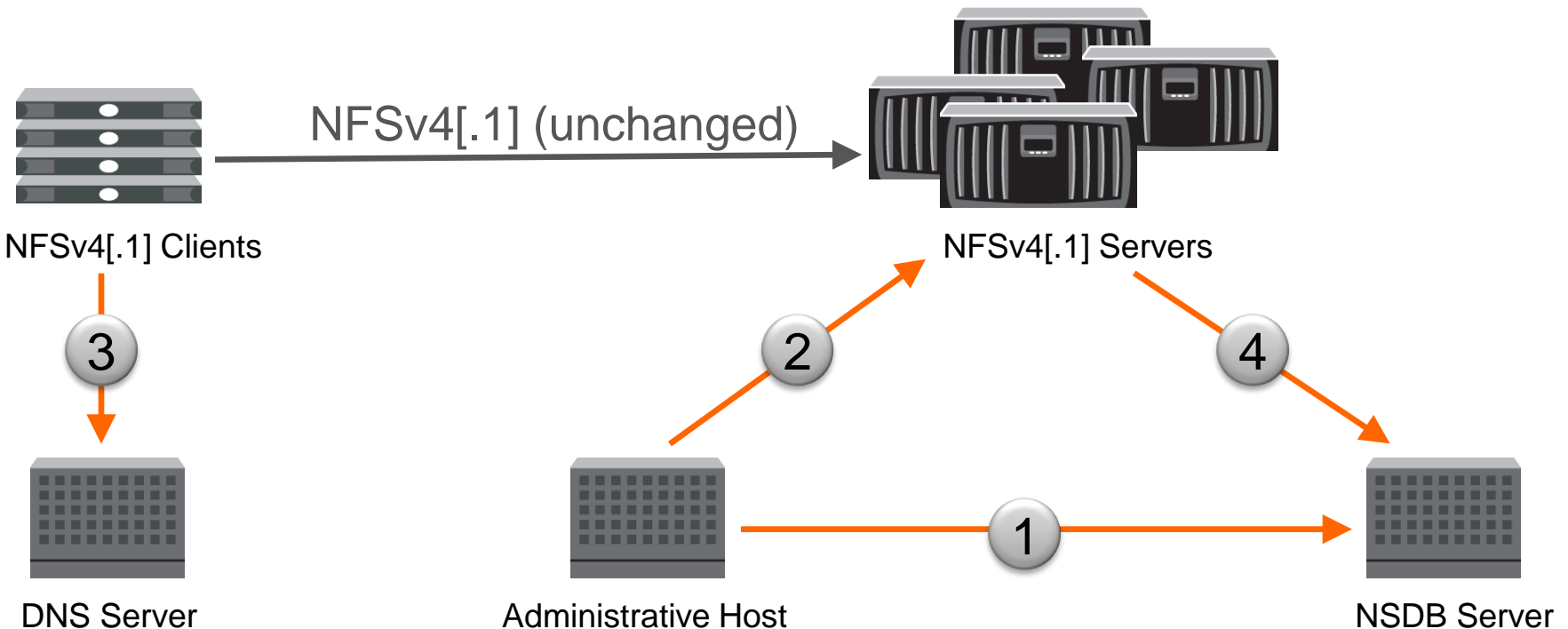
FedFS Protocols (in orange)

Namespace Management

- 1 NSDB Management (LDAP)
- 2 Junction Management (ONC RPC)

Namespace Navigation

- 3 Namespace Root discovery (DNS)
- 4 Junction Resolution (LDAP)



Namespace Root Discovery

- Allows an NFS client to locate the root of a domain's namespace using the DNS.
 - A domain's root file servers are published in DNS using SRV Resource Records (RRs).
 - SRV RRs priority and weight fields allow multiple root file servers to be specified with associated preferences.
 - Client can locate the root of any published domain, not just its own domain.
- Recommended convention is for domains to appear in local file system under special /nfs4 directory.



Namespace Root Discovery Example

- Special /nfs4 directory in NFSv4 client's local filesystem:

```
> ls /nfs4
example.net/  jumble.edu/  foo.com/
```
- Application accesses /nfs4/example.net

```
> cd /nfs4/example.net
```
- A DNS SRV query is performed automatically, SRV RR lists example.net's root servers.
- NFSv4 client automatically mounts most preferred example.net NFSv4 server at /nfs4/example.net.



Namespace Root Discovery Details

- Read-only and read-write locations can be defined.
- Namespace root NFS servers can also export non-root data. Reserved names for root filesystems:
 - `/.domain-root-example.net` for read-only
 - `/.domain-root-write-example.net` for read-write
- A fileserver could export a special directory and perform DNS lookups for legacy NFS clients.



Namespace Database (NSDB)

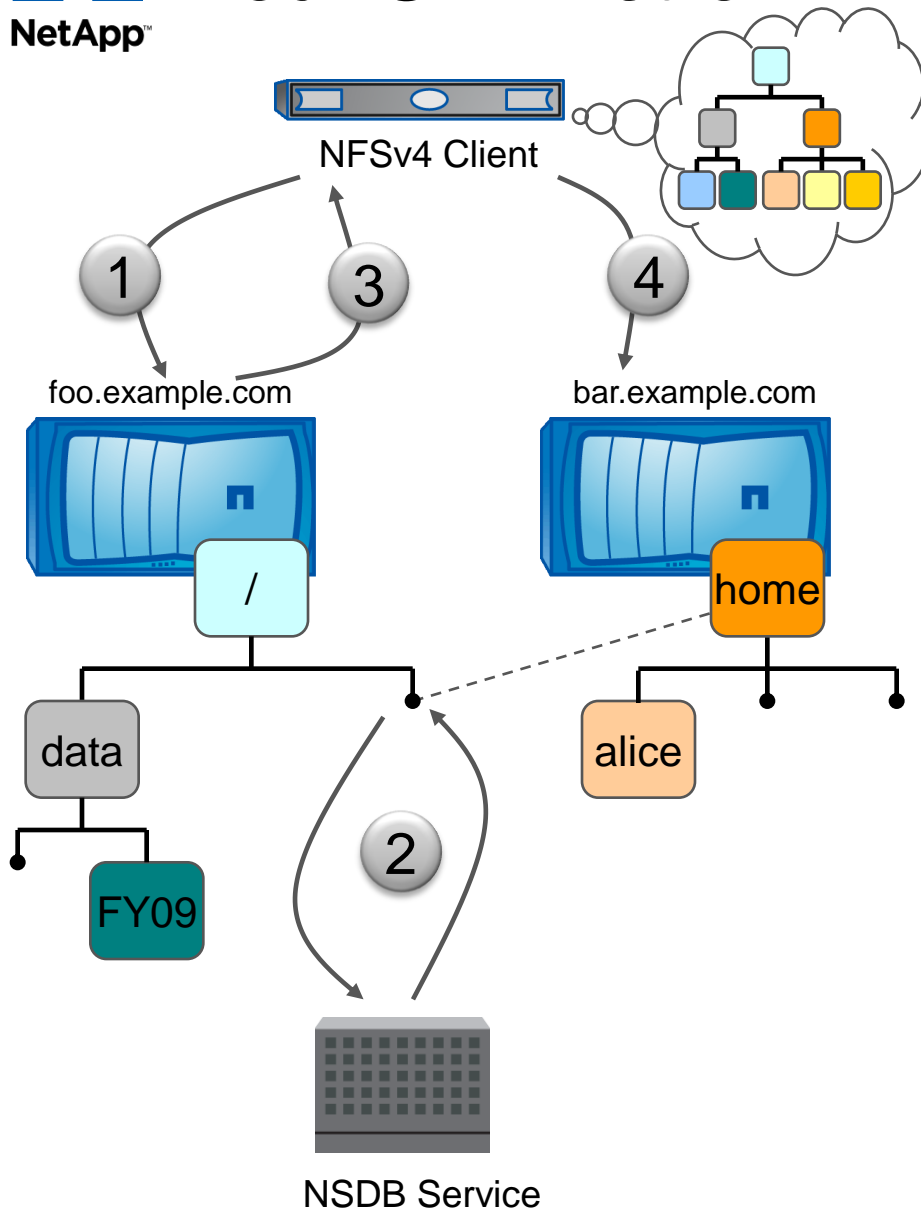
- The **NSDB** maps FSNs to FSLs.
 - An LDAPv3 service.
 - LDAP schema with FedFS attributes and objects
- Administrator creates FSN and FSL records.
- Fileserver looks up FSL records for a given FSN.



Administration Protocol

- An ONC RPC protocol is defined for administering file servers.
- Allows administrator to:
 - Create a junction
 - Delete a junction
 - Query a junction

FedFS in Action



The user requests /home/alice:

- 1 The client attempts to access /home/alice on server foo.
- 2 Server foo discovers that home is a namespace junction and determines its location using the FedFS NSDB service.
- 3 Server foo returns an NFSv4 referral to the client directing it to server bar.
- 4 The client accesses home/alice on server bar.

Access to the federated namespace may be protected using the full flexibility and capabilities of NFSv4 ACLs.



STATUS



Standardization Status

- Protocols are being standardized in the IETF's NFSv4 working group.
 - Four documents total.
- Documents are being reviewed by the NFS and Internet community for publication as RFCs.
 - Requirements published as RFC 5716.
 - Moving towards WG Last Call on remaining documents.
- Participation from the multiple NFS vendors: EMC, IBM, NetApp, Sun, and others.

References

- Requirements for Federated File Systems
<http://www.ietf.org/rfc/rfc5716.txt>
- Using DNS SRV to Specify a Global File Name Space with NFS version 4
<http://tools.ietf.org/html/draft-ietf-nfsv4-federated-fs-dns-srv-namespace>
- NSDB Protocol for Federated Filesystems
<http://tools.ietf.org/html/draft-ietf-nfsv4-federated-fs-protocol>
- Admin Protocol for Federated Filesystems
<http://tools.ietf.org/html/draft-ietf-nfsv4-federated-fs-admin>



Questions?