Federated FS

IETF'74 NFSv4 WG Meeting
March 25, 2009

James Lentini   Renu Tewari

jlentini@netapp.com   tewarir@us.ibm.com

Overview

- FedFS is a set of open, portable protocols that permit the construction of a cross-platform, 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

Federated Namespace Benefits

- Simplified Management
  - Separates logical and physical location
  - Clients only need to mount the root (or some other part) of the namespace. No more automounter?
- Allows future enhancements to be transparent to the client (these enhancements are outside the scope of the FedFS effort):
  - Replication: may be used for load balancing or high availability
  - Migration: may be used for moving data closer to compute or decommissioning systems

Influences

- IBM Almaden Research Center's Glamour Project

http://www.almaden.ibm.com/StorageSystems/projects/glamour/

- University of Michigan CITI's NFSv4 Namespace Project

http://www.citi.umich.edu/techreports/reports/citi-tr-06-1.pdf

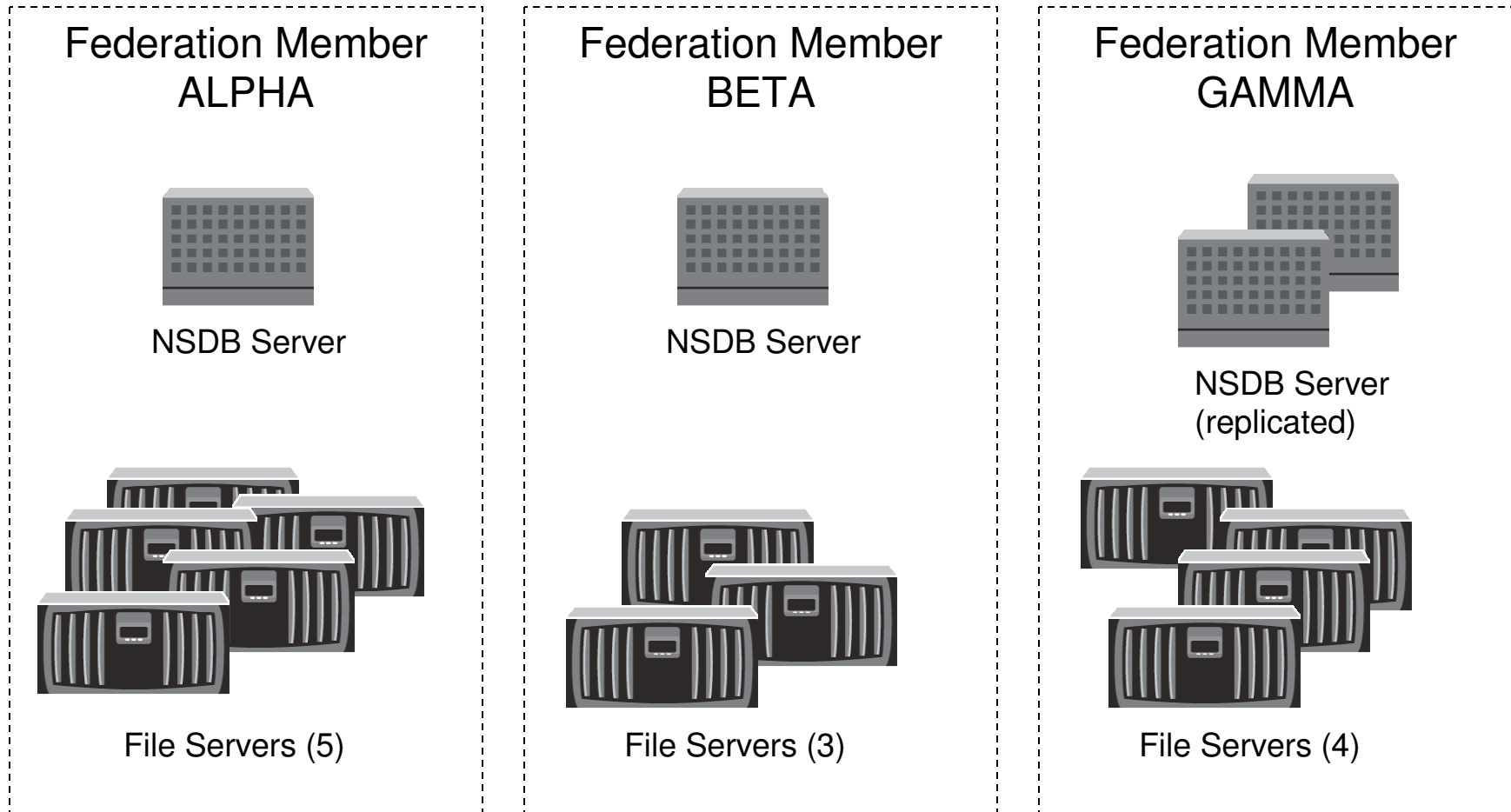- DCE/DFS, Andrew File System (AFS), …

Terms

- **Fileset**: a directory tree (as small as a single directory; multiple FSIDs allowed)
- **FSN** (fileset name): a fileset identifier that is independent of the representation of the fileset
  - ☐ Each FSN contains an FsnUuid (a UUID) and an NSDB location
- **FSL** (fileset location): network location of a fileset instance
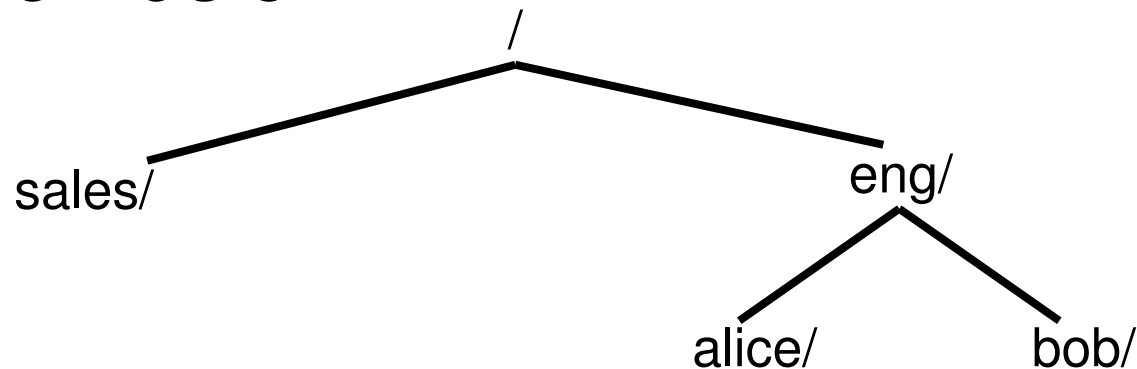
Terms (continued)

- **Junction**: an object that provides a way for one fileset to reference another

- **NSDB** (namespace database): a service that tracks the mapping between FSNs and FSLs; implemented with LDAP

An Example Federation

Federation Member ALPHA	Federation Member BETA	Federation Member GAMMA
NSDB Server	NSDB Server	NSDB Server (replicated)
File Servers (5)	File Servers (3)	File Servers (4)

An example namespace

- The illusion:

```
                    /
           /                 \
      sales/                 eng/
                          /        \
                      alice/        bob/
```
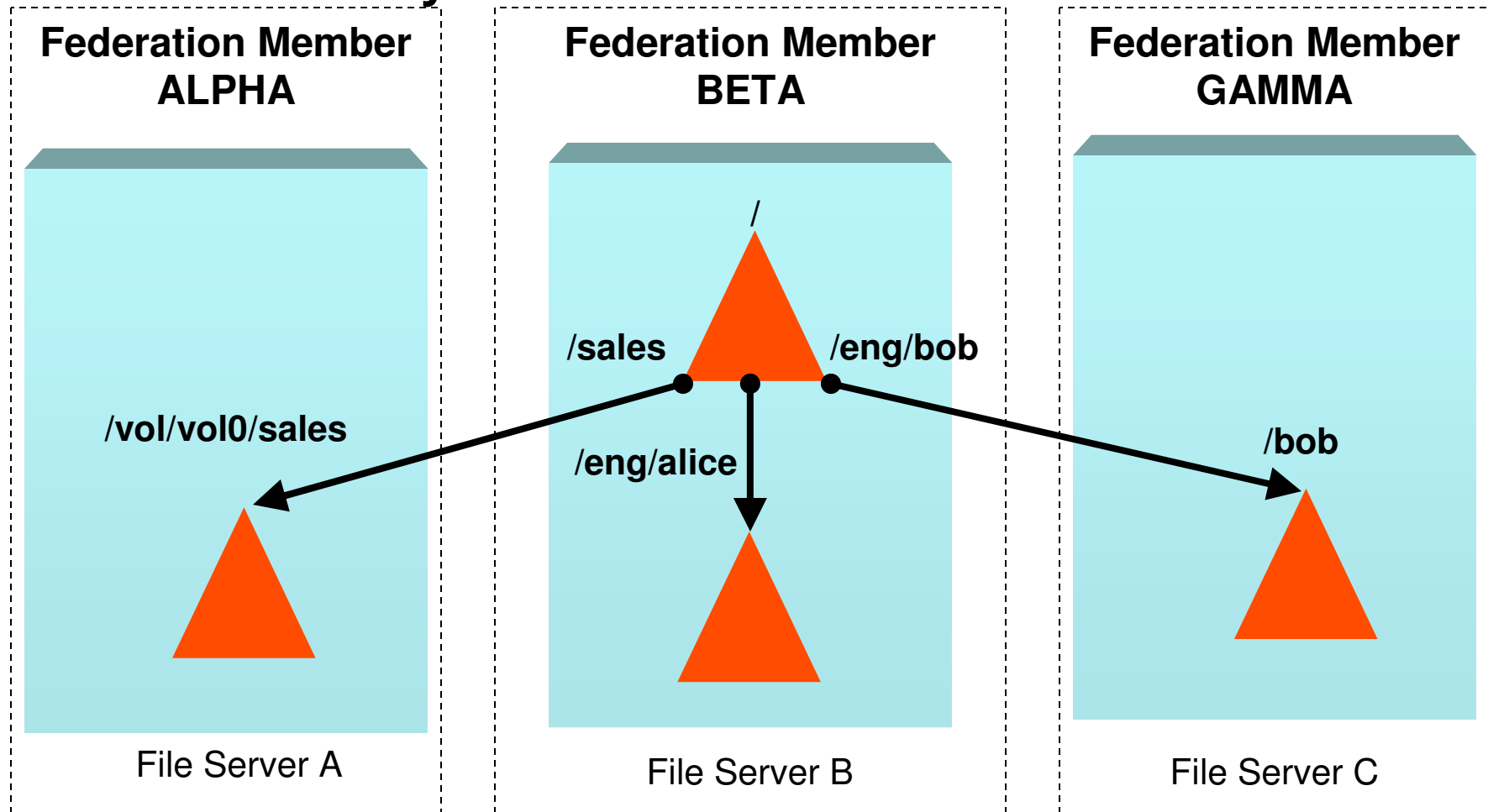
- A simple hierarchical namespace is what we want the client (and user) to see.

- Behind the scenes, things may be somewhat more complicated…

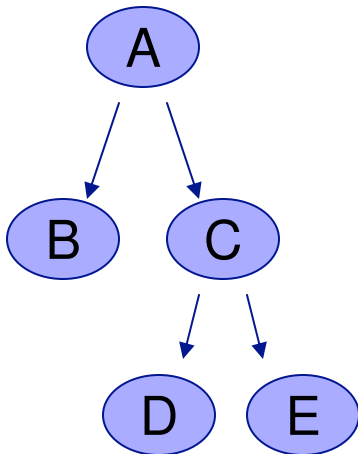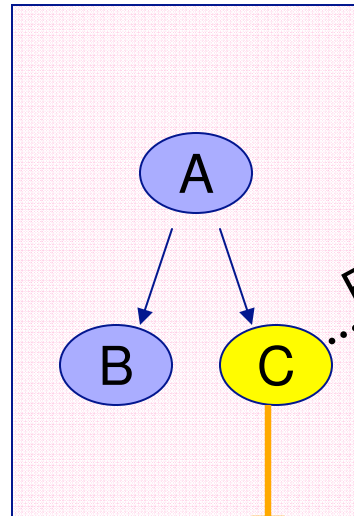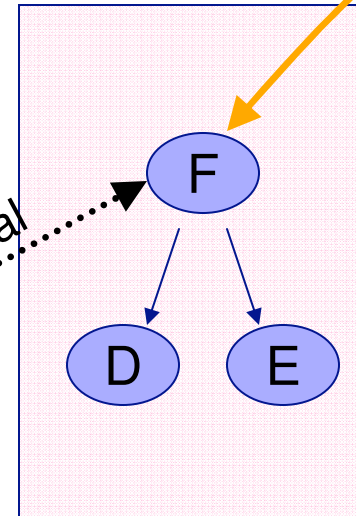# An example namespace

- The reality:

Federation Member ALPHA — File Server A
Federation Member BETA — File Server B
Federation Member GAMMA — File Server C

/
/sales
/eng/bob
/vol/vol0/sales
/eng/alice
/bob

© The IETF Trust (2009).

FSN to FSL Mapping

Namespace

Server X

Server Y

A

B   C

D   E

A

B   C

F

D   E

Referral

FSN = <NSDB=Y', FsnUuid=xyz>

NSDB Y'

...

...

xyz ➔ Y:/F

...

...
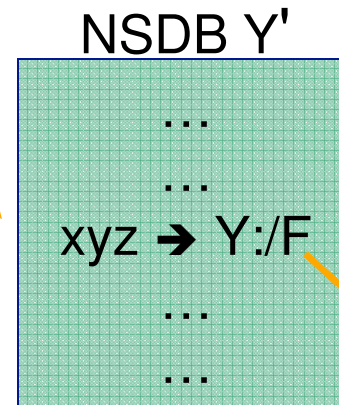
- ❑ Node C contains the FSN of the fileset
- ❑ The NSDB in the FSN knows the current FSLs for the fileset
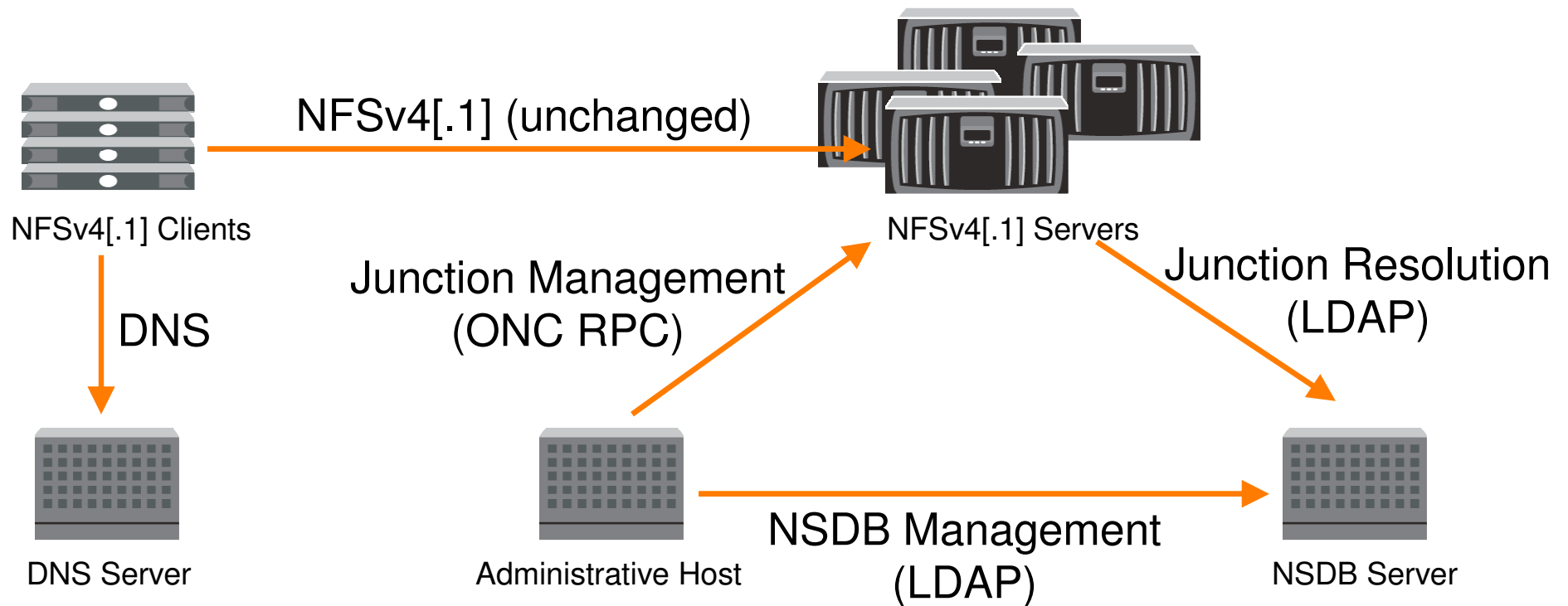- ❑ FSLs can change without updating FSNs

Goals

- No modifications to NFSv4[.1] client protocol
  - Client configuration (e.g. automounter) changes allowed
- An open, cross-platform, multi-vendor standard
- Support for decentralized (federated) administrative control of storage systems
- Highly scalable to large namespaces supporting many clients and servers in different geographies

Protocols

NFSv4[.1] (unchanged)

NFSv4[.1] Clients

NFSv4[.1] Servers

Junction Management
(ONC RPC)

Junction Resolution
(LDAP)

DNS

DNS Server

Administrative Host

NSDB Management
(LDAP)

NSDB Server

Drafts

Four drafts published as working group documents:

- Requirements for Federated File Systems
- NSDB Protocol for Federated Filesystems
- Admin Protocol for Federated Filesystems
- Using DNS SRV to Specify a Global File Name Space with NFS version 4

Requirements

Draft name: draft-ietf-nfsv4-federated-fs-reqts-01

Intended status: Informational

WG Last Call Estimate: May, 2009

- Lists assumptions and requirements
- Assumptions:
  - No changes to existing client protocols, may add new client protocols (A1)
  - Locations can be correctly interpreted throughout the federation (e.g. a FQDN will map to the same IP address and be routed to the same (logical) server throughout the federation). (A8)
  - Additional assumptions are listed in the draft.

Requirements (continued)

- Binding an FSN to FSL does not require interaction with the fileserver (R2).

- Junctions are not stored in NSDB; fileservers MUST be able report if a specific path ends in a junction and MAY be able to report all junctions in a fileset (R8).

- A common root MUST be provided, possibly using a root fileset (R15).

Namespace Root Discovery
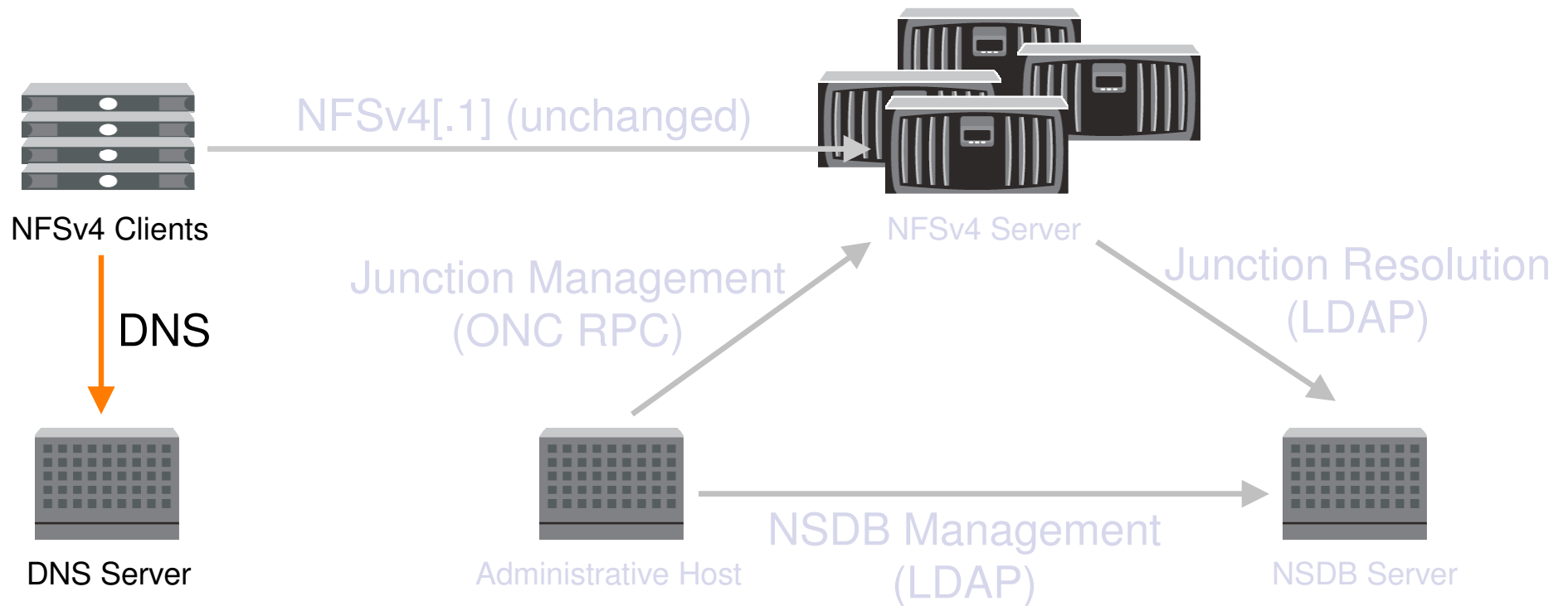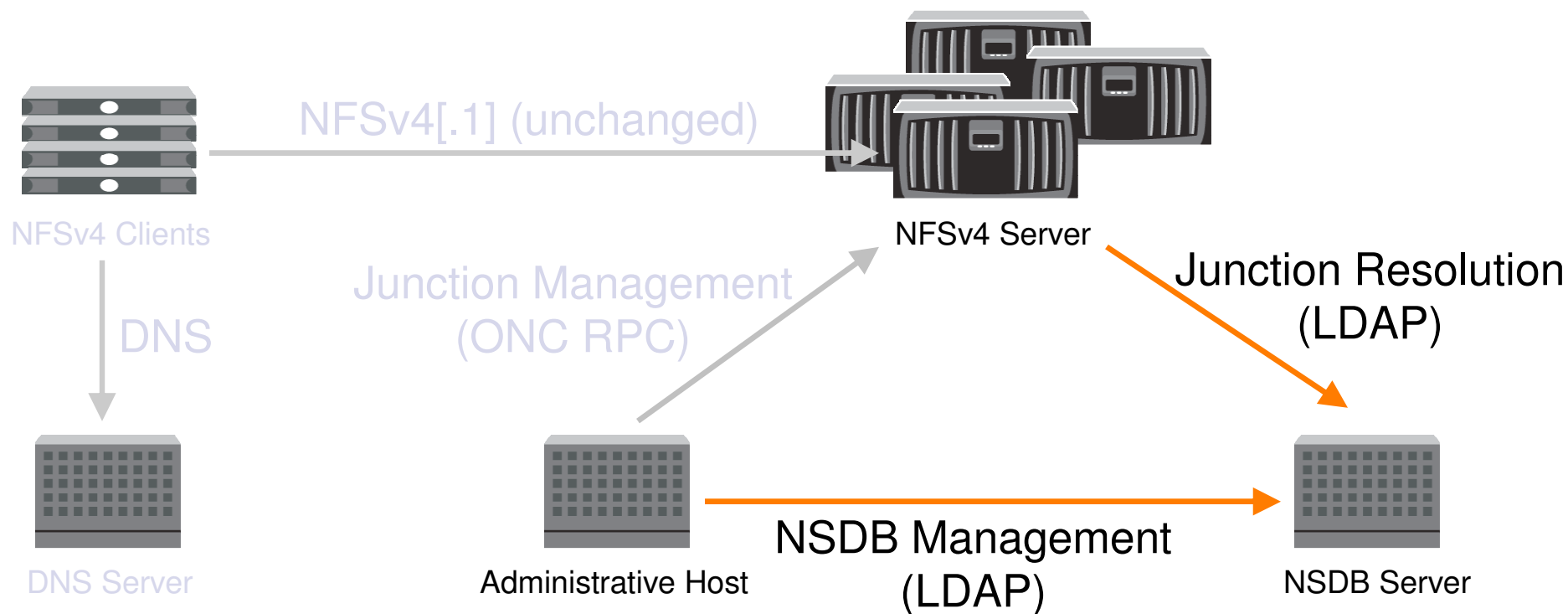
Namespace Root Discovery

Draft Name: draft-ietf-nfsv4-federated-fs-dns-srv-namespace-00

Intended Status: Standards Track

WG Last Call Estimate: October, 2009

- Describes how to publish a NFSv4 domain's namespace root to NFS clients.

- Feedback from the DNS working group (a few years ago) endorsed the approach of using DNS SRV resource records.

- The capability exists to authenticate the role of an NFS server as the namespace root.

- RECOMMENDS the client convention of mounting the root at "/nfs4/".

NSDB Protocol

NFSv4[.1] (unchanged)

NFSv4 Clients

NFSv4 Server

Junction Resolution
(LDAP)

DNS

Junction Management
(ONC RPC)

DNS Server

Administrative Host

NSDB Management
(LDAP)

NSDB Server

NSDB Protocol

Draft name: draft-ietf-nfsv4-federated-fs-protocol-01

Intended Status: Standards Track

WG Last Call Estimate: October, 2009

- Describes the LDAP junction resolution protocol (NFS server to NSDB) and LDAP NSDB management protocol (administrative host to NSDB)

NSDB Management

- Admin host sends LDAP commands to the NSDB
  - ☐ Create FSN
  - ☐ Create FSL for an FSN
  - ☐ Delete, Update FSLs
  - ☐ Junction mapping between parent and target FSNs
- Only one NSDB is involved in any operation.

NSDB LDAP Schema

- LDAP stores mapping between FSN and FSL
  - ☐ Defines an FSN object, FSL object
- Use of LDAP (replication limitations) and schema (flexibility)
- Need support in fileserver implementations (deployment).
- Other options are available (e.g. ONC RPC protocol for FSN to FSL mapping that "wraps" LDAP).

Junction Resolution

- The fileserver resolves the FSLs for a given FSN (using LDAP to query the NSDB)

- The fileserver needs to support junction resolution based on an FSN

- The fileserver needs to convert the FSLs returned by the NSDB to the v4[.1] fs_locations [info] attributes
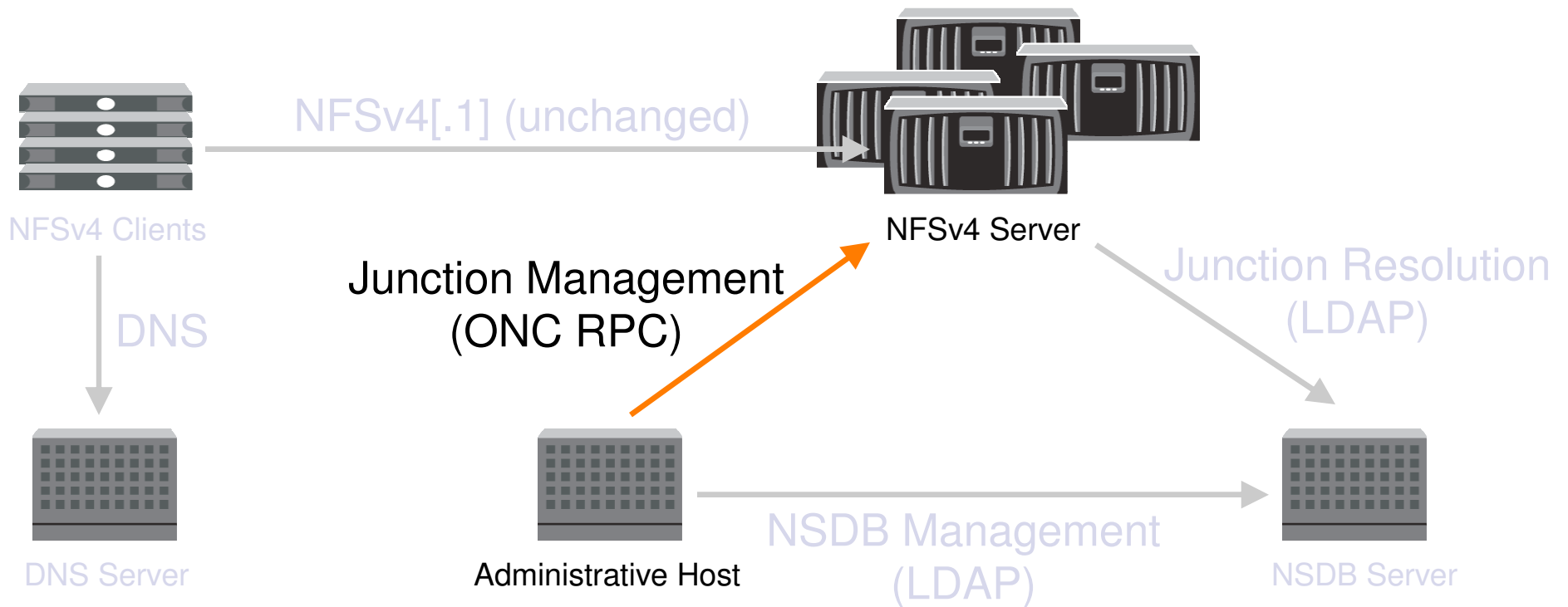
Root Fileset

- A root fileset is exported by one or more fileservers in the federation as the top level of a namespace (multiple namespaces are supported).

- Multiple fileservers can export the root and clients can find the root fileserver using an external protocol.

- Root Fileset layout stored in one NSDB and replicated
  - build a FedFS-specific mechanism (using LDAP) or require a general replication mechanism (possibly provided outside the FedFS protocols)
  - Fileservers need to export the layout to clients and update the layout to reflect the latest stored in the NSDB.

Security Considerations

- FedFS does not introduce new security mechanisms.

- RPCSEC_GSS is RECOMMEDED for NFS client to NFS server communication

- TLS is RECOMMENDED for LDAP access to NSDB

- NSDB must be as secure as the NFS server

Admin Protocol

NFSv4[.1] (unchanged)

NFSv4 Clients

NFSv4 Server

DNS

Junction Management
(ONC RPC)

Junction Resolution
(LDAP)

DNS Server

Administrative Host

NSDB Management
(LDAP)

NSDB Server

Admin Protocol

Draft name: draft-ietf-nfsv4-federated-fs-admin-01
Intended Status: Standards Track
WG Last Call Estimate: October, 2009

- Describes an ONC RPC protocol for managing junctions on a fileserver. Procedures to
    - Create a junction
    - Delete a junction
    - Query the existence of a junction at a specific path
- Decisions to review:
    - Security considerations: RECOMMENDs the use of RPCSEC_GSS
    - No concept of administrator (user) names or login within ONC RPC Admin protocol; server may use RPCSEC_GSS identity to limit capabilities

Status

- Prototype of NSDB protocols demonstrated at the summer WG meeting in Dublin (Summer 2008)

- Four drafts published as NFSv4 WG documents (Fall 2008).

- Federated namespace added to the NFSv4 WG charter (Spring 2009)

Federation-wide Identity

- Within a single NFSv4 user domain, user and group mapping (e.g. in Access Control Lists) is consistent.

- When multiple NFSv4 user domains are part of a federation, how can the users (groups) be mapped consistently? One possibility is schema extensions to the LDAP database format (RFC 2307). See presentation from Andy Adamson.

Client Support for Referrals

NFSv4 clients supporting referrals are available on many platforms. For example:

- **AIX**: referrals and replication (including failover) supported since 5.3 (released August, 2004)

- **HPUX**: referrals supported in HP-UX 11iv3 with ONCplus B.11.31.03 (released May, 2008)

- **Linux**: referrals supported since 2.6.18 (released September, 2006)
  - ☐ Migration/replication support under development

Meetings

Open meetings are held each week to resolve issues and review proposals.

- Thursdays, 1:30 – 2:30 PM Eastern

    (10:30 - 11:30 AM Pacific)

- Conference Number: 1-888-765-3653
- Conference ID: 2354843

Acknowledgements

Many people have contributed!
- Andy Adamson (NetApp)
- Dan Ellard (BBN Technologies)
- Craig Everhart (NetApp)
- Paul Lemahieu (EMC)
- James Lentini (NetApp)
- Manoj Naik (IBM)
- Renu Tewari (IBM)
- Robert Thurlow (Sun)
- Mario Würzl (EMC)