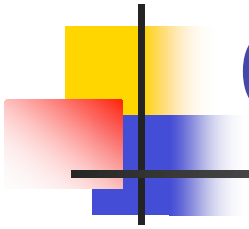
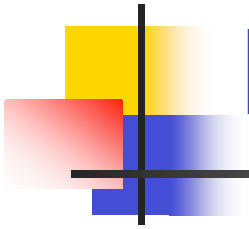
Federated-fs protocol overview

R. Tewari, M. Naik
IBM Almaden
D. Ellard, C. Everhart
Network Appliance



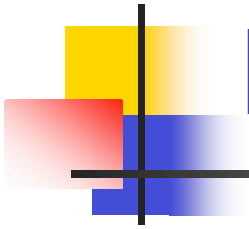
Outline

- Background and motivation
- Terms and definitions
- The protocol
- Next steps



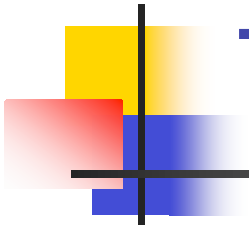
Background

- Idea: build a cross-platform federated file system with a shared common namespace.
  - Independent fileservers are federation members
  - Federation members are peers.
  - Could have multiple administrative boundaries
- Use Cases:
  - Single enterprise with multi vendor fileservers and fileserver collections
  - Multiple enterprises sharing independently managed namespace(s)



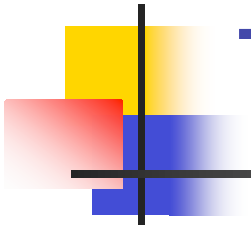
Protocol Goals

- Specify a set of interfaces by which fileservers and collections of fileservers with different administrators can form a fileserver federation
  - that provides a namespace composed of the filesystems physically hosted on and exported by the fileservers of the federation.
- It should be possible, using a system that implements the interfaces, to share a *common* namespace across all the fileservers in the federation.
- It should also be possible for different fileservers in the federation to project *different* namespaces and enable clients to traverse them.



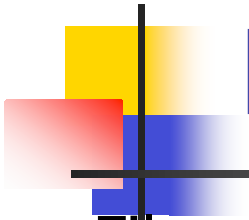
Terms and Definitions

- Fileset: the basic organizational unit for data
  - Container of data abstraction
  - To the user, it's a filesystem or a directory tree
  - Behind the scenes, can be implemented as a replicated, mobile storage container
- FSN: "Fileset Name"
  - Symbolic name of a fileset
  - UUID + Name of the repository responsible for the fileset
  - FSN is globally unique



Terms…

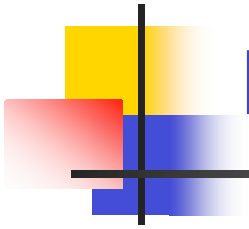
- FSL: “Fileset Location”
  - Location of an implementation of a fileset
  - A fileset may be implemented by several FSLs
  - FSLs can come and go, change, etc.
- NSDB: “namespace database”
  - Repository of fileset and fileset relationship information
  - Keeps track of the mapping from FSNs to FSLs
  - Every FSN has a single authoritative NSDB
- Junction: binds a path within a fileset to a target FSN
  - Requires fileserver support
  - a junction can be viewed as a reference from a directory in one fileset to the root of the target fileset



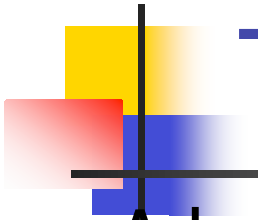
How it all fits together

- Filesets are the building blocks.
- Junctions define the relations between the filesets
- Namespace built by "mounting" filesets
- When an NFSv4 client accesses a junction (which looks to the client like a directory):
  - The fileserver finds the FSN of the junction's target fileset
  - The fileserver finds (from the FSN) the NSDB location responsible for that FSN
  - The fileserver queries the NSDB for the current set of FSLs
  - The fileserver uses the NFSv4 referral mechanism
    - informs the client that the directory has "moved"
    - Returns the FSL information in the fs_locations attribute
  - The client is redirected to the root of the fileset that is the target of the junction

Protocol basic functional requirements

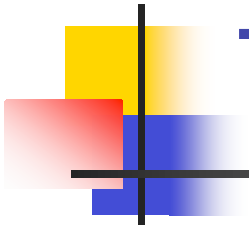


- An interface to add/delete/modify FSN → FSL mappings in the NSDB
- An interface to query the NSDB for FSN → FSL mappings
- An interface to create a junction at a fileserver
  - that a junction to some other fileset exists within a directory of a fileset location hosted by that server



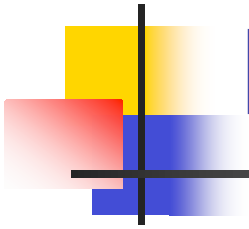
Three players

- Administrator
  - Initiates creation/updates/deletes of filesets, junctions, fileset locations
- NSDB
  - Maintains the fileset state
  - Responds to inserts/updates/queries
  - An NSDB is an LDAP server + persistent database
- NFSv4 server
  - In response to client requests, queries the NSDB for FSLs
  - In response to admin requests, updates/queries server state (junctions)



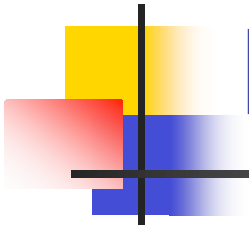
Three protocols

- Admin---NSDB
  - Update the NSDB state
  - Implemented as LDAP using fed-fs schema
- Fileserver---NSDB
  - Query the NSDB
  - Implemented as LDAP queries using fed-fs schema
- Admin---Fileserver
  - Update/query the NFSv4 server state (junctions)
  - Implemented via SUNRPC
  - Why not LDAP?
    - We don't want every NFS server to be an LDAP server as well (it can be an LDAP client). This seems like a simpler approach.



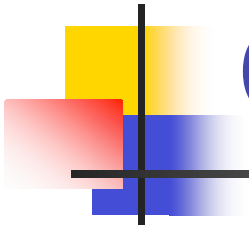
NSDB State

- Each NSDB uses the LDAP fed-fs schema
- Define objects
  - Fileset
  - Fileset location
  - Junctions
- Need fileserver to parse fed-fs schema
  - Query based on FSN not pathname
  - Convert to fs_locations
  - Not identical to `rnfs` experimental schema.



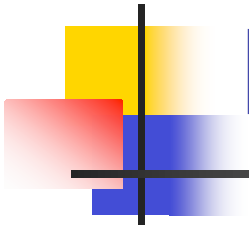
Discussion questions

- Choice of protocol for fileserver—admin interaction
- LDAP schema rnfs variant
- Common namespace
  - finding root fileservers not defined by protocol
- FSL consistency not defined
- Caching of namespace information across NSDBs
  - Default local NSDB not defined
- Security Issues
- Access/authorization ID mapping issues



Open source plans

- IBM Research Glamour base code on Linux to be made available using common public license



More Information at

- list federated-fs@sdsc.edu
- IETF Draft
  - draft-tewari-federated-fs-protocol-01.txt
- Emails:
  - Daniel dot Ellard at netapp dot com
  - tewarir at us dot ibm dot com
- BOF at FAST?