Reparse Points

Afshin Salek
afshin@sun.com

Agenda

- Context/background
- Scope
  - Umbrella case (2009/399)
  - Reparse Point case (2009/387)
- Reparse point proposal

File/Storage Management Challenges

- Namespace Problem
  - Explosive growth of data generation
  - Proliferation of file servers and NAS appliances
  - It's a *management nightmare*
- Administration Complexity
  - Adding file servers, rebalancing storage, setting up failover, replication, migration, consolidation, data distribution
- Issues:
  - Complex, tedious and time-consuming for admins
  - Disruptive to users
  - Expensive for companies

Global Namespace

- Namespace
  - A logical view or abstraction between clients (data consumers) and file systems (data storage)
- Namespace Administration
  - Logical presentation rather than physical location
  - add, change, move, and reconfigure physical file storage without affecting the logical view

Global Namespace Implementation

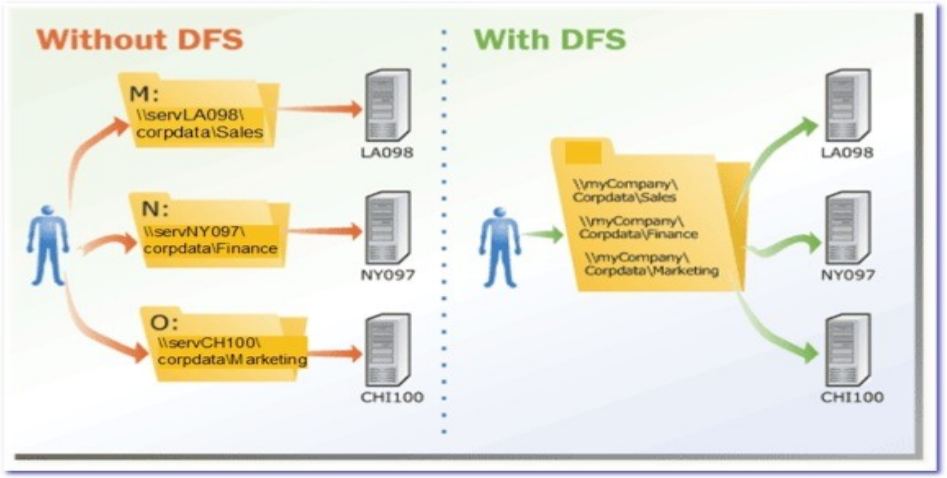
- NFS and SMB offer a mechanism to implement a global namespace
  - NFSv4.x *referrals*
  - SMB *DFS (Distributed File System)*

- Solaris can support native referrals for both NFS and SMB
  - Unique heterogeneous support for referrals

Umbrella Case (2009/399)

- Reparse Points (2009/387)
- NFSv4 Referrals
- NFS Support for FedFS
- Standalone MS-DFS
- Domain based MS-DFS

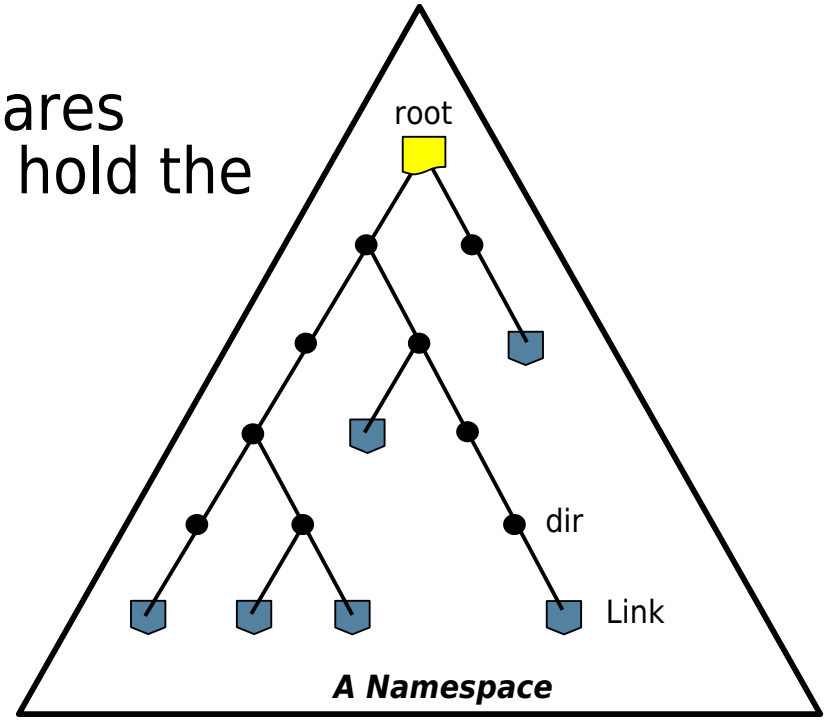
Namespace Illustration (MS-DFS)



- Namespace
  - a group of exports or shared folders located on different servers
  - presented as a virtual tree of directories on a local file system

Namespace Representation

- A namespace is managed centrally on the server that hosts it
- A root folder which is exported/shared by the host server so the namespace is visible over network
- Regular folders
- Links that represent exports/shares [typically on other systems] that hold the actual data behind the virtual namespace



Link Representation (Reparse Points)

- A generic mechanism for location redirection
  - DFS/NFS referrals, HSM, mounted file systems
- Need a file system object to represent the reparse point and store the target location information (reparse data)
- The file system is not cognizant of reparse point format or content
  - Reparse point consumers know how to use the reparse data

Reparse Points Object

- Options Considered
  - introduce a new object type
  - a mount point
  - use existing types i.e. file, directory or symbolic link
  - use existing types tagged as a reparse point
- Object Criteria
  - specific functionality: location redirection
  - reparse points have some data associated with them
  - minimum effect on existing utilities and applications, especially backup utilities and software

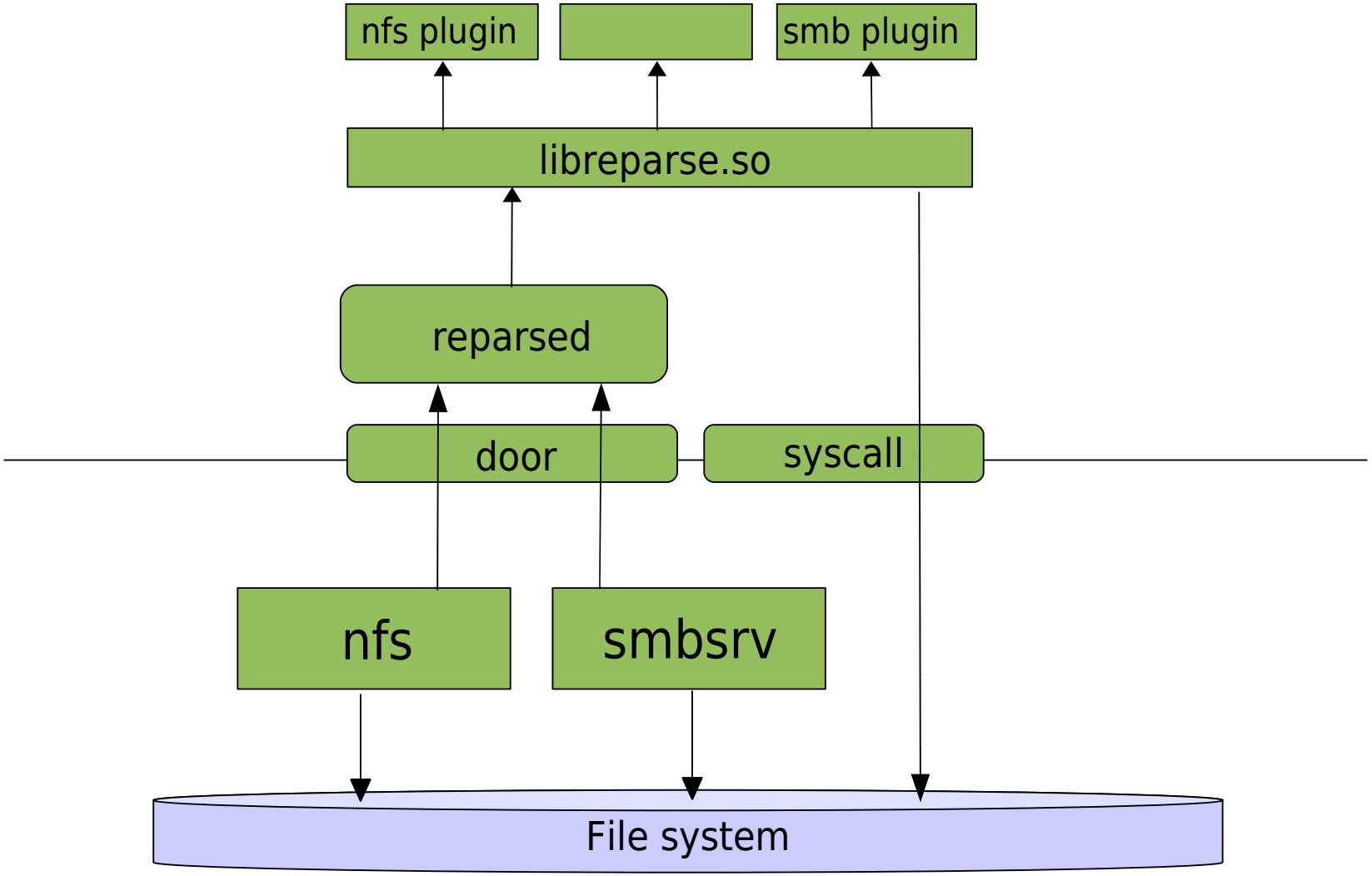
Reparse Points Proposal

- Reparse point
  - symlink with an extensible attribute tag (XAT_REPARSE)
- Reparse data (symlink content)
  - @{REPARSE@{service-type1:data} [@{service-type2:data}]...}
- Basic API
  - Userspace: symlink(2), readlink(2)
  - Kernel: VOP_SYMLINK, VOP_READLINK
  - VOP_SYMLINK: Set XAT_REPARSE if symlink is a reparse point

Reparse Points Proposal cont.

- VFS feature registration to announce reparse point support
- Enhance VOP_LOOKUP to return attributes
- Userpace daemon (*reparsed*) and libraries with plugin architecture
  - unified API for managing reparse points and reparse data for existing and future consumers

Reparse Daemon and Library



Reparse Points

Afshin Salek
afshin@sun.com